MŰEGYETEM 1782

TYPOTEX

WETTL FERENC

LINEÁRIS ALGEBRA

azoknak, akik érteni is szeretnék

2011

Tartalomjegyzék

Pályázati támogatás

Gondozó

Copyright

Kulcsszavak: Lineáris algebra, vektorok, lineáris egyenletrendszerek, mátrixok, lineáris leképezések.

Rövid ismertetés: A könyv a szerző mérnökhallgatók számára tartott előadásainak tapasztalataira építve a lineáris algebra több témáját újszerű módon tárgyalja. A fogalmakhoz és tételekhez a szokásos helyett igyekszik motiválhatóbb, természetesebb utakat találni, és ezzel érthetőbbé tenni az Olvasó számára. Különösen azokra a témákra koncentrál, melyek ismerete a modern mérnöki, természettudományos és közgazdasági alkalmazások megértéséhez szükséges.

A könyv jelen változata az első oktatási eredmények nyomán folyamatosan változik.

Támogatás: Készült a TÁMOP – 4.1.2. – 08/2/A/KMR-2009-0028 számú pályázat „Természettudományos (matematika és fizika) képzés a műszaki és informatikai felsőoktatásban” című projekt keretében.

Készült: a BME TTK Matematika Intézet gondozásában
Szakmai felelős vezető: dr. Ferenczi Miklós
Projektmenedzser: dr. Ádám Katalin

A projekt webcíme: http://tankonyvtar.ttk.bme.hu

Címlap grafikai terve: Csépány Gergely László, Tóth Norbert

Tartalomjegyzék

Listák

Tételek, állítások, következmények

Definíciók

Kidolgozott példák

Bevezetés

Két motiváló emlékem Néhány éve, az akkor legkiválóbb mérnökhallgatómat megkérdeztem, hogy mi a véleménye a szemeszter anyagáról. Néhány óvatos, tartózkodó mondat után egy váratlan, és akkor számomra teljesen érthetetlen mondattal lepett meg „A lineáris algebrát nem lehet érteni.” Hiába próbálkoztam azzal, hogy minden dolgozatát maximális pontszámmal írta meg, még a nehéz, gondolkodtató, bizonyítást kérő kérdésekre is tudott válaszolni. Semmi magyarázattal nem tudta feloldani ezt az ellentmondást, csak makacsul megismételte állítását, és a végén még annyit tett hozzá, „az analízist lehet érteni, az szép”. Mire gondolhatott? Hamar én is úgy gondoltam, igaza lehet. Például a függvény határértékének vagy folytonosságának Cauchy-féle definíciója igen nehéz sok diák számára, pedig már középiskolában is tanulták. Nehéz, de valami fogalma mégis minden hallgatónak van róla. Akár tudja, akár nem a definíciót, akár jók, akár zavarosak az elképzelései, többnyire tudja miről van szó. De nincs ez így például a determinánssal. Aki megtanulta azt, hogy „vedd az első sor elemeit, és mindegyiket szorozd meg a hozzá tartozó előjeles aldeterminánssal” (egy rekurzív módon definiált fogalom!), az ezzel elélhet, megoldhat feladatokat, de úgy érzi, nem ért ebből semmit. És mit gondol, ha azt látja, hogy ezt az érthetetlen fogalmat használva egy rejtélyesnek tűnő szabállyal (Cramer-szabály) meg tudja oldani azt az egyenletrendszert, amit már az általános iskolában is meg tudott? Csak akkor értette is, hogy mit miért csinál!

A másik történet 30 éves. Fiatal oktatóként kérdőre vontam az egyik mérnöki kar dékánját, hogy az oktatási reformjában miért csökkentette a matematikaórák számát! Azt válaszolta, hogy „mert szükségünk van időre, hogy megtaníthassuk a diákokat gondolkodni”. Közbevetésemre, hogy a matematika épp ezt teszi, röviden csak annyit mondott, hogy a „matematika csak kaptafákat ad nekik, gondolkodni mi tanítjuk őket”.

Kinek készül ez a könyv és miért Ez a könyv főiskolai és egyetemi BSc és MSc szintű lineáris algebra kurzusaihoz és részben az azt megelőző félévek vektorgeometriát is tartalmazó kurzusaihoz készült. Szemléletében eltér a Magyarországon megjelent hasonló témájú tankönyvektől. Megírását és a sok tekintetben újfajta megközelítést az alábbi tények indukálták:

- A felsőoktatás reformjának hatásaként a hallgatók mind hozott tudásukat, mind matematikai képességeiket tekintve heterogénebbek, mint korábban.
- A felsőfokú oktatás változó szemlélete nagyobb hangsúlyt fektet az alkalmazásokra, mind a tananyag összeállításában, mind azoknak a képességeknek a kifejlesztésében, amelyek a megszerzett tudás alkalmazásához szükségesek.
- A matematika felsőfokú oktatásával foglalkozó nemzetközi kutatások eredményei, a számítógép használatának elterjedése új oktatási szemlélet kialakítását kívánják.

A könyvben követett elvek

Didaktikai célszerűség A könyv megírásakor fő célunk az volt, hogy a lineáris algebra absztrakt fogalmait a lehető legegyszerűbben, legérthetőbben vezessük be. Sosem a legáltalánosabb megfogalmazás, a legaxiomatikusabb felépítés, a matematikailag legtömörebb tárgyalásmód megtalálása volt a cél, hanem a didaktikai célszerűség, a tananyag minél hatékonyabb tanulhatóságának elérése.

Moduláris szerkesztés A könyv anyagát igyekeztünk modulárisan, apró egységekre bontva megszerkeszteni, ezzel nem csak az áttekinthetőségét növelni, de a többcélú, különböző szintű, különböző időtartamú kurzusokhoz való alkalmasságát is megkönnyíteni.

Alapfogalmak korai bevezetése Tapasztalataink szerint nem elég hatékonyak a lineáris algebra alapfogalmainak megértetésében azok a kurzusok, melyek a kurzus elejét az egyszerű mátrix- és determinánsszámítási, egyenletrendszer-megoldási, sajátérték-számítási technikákkal töltik, majd a kurzus végén a hallgatók nyakába öntik a lineáris függetlenség, test feletti vektortér, altér, bázis, lineáris leképezés… fogalmakat. De nem jobb a hatásfoka a fordított felépítésű kurzusoknak sem, melyek az általános fogalmakkal és eredményekkel kezdik, melyekből a végén „könnyedén” kipottyan pl. az egyenletrendszerek elmélete.

Az a határozott véleményünk, hogy (az absztrakt gondolkodásban kiválóak szűk csoportját leszámítva) a hallgatók gyorsabban és mélyebb ismeretekhez jutnak, ha az absztakt fogalmakkal konkrét esetekben már korábban megismerkednek, és az absztrakt fogalomalkotás valóban absztrakció, és nem kinyilatkoztatás útján történik. A lineáris algebra legtöbb fontos, e könyvben tárgyalt fogalmával az első fejezetekben találkozik az olvasó, az általános fogalomalkotás csak ezt követi.

Fokozatosság A könyv egészen elemi – az első fejezetekben középiskolás szintig visszanyúló – tárgyalásmóddal kezdődik, melyet egyre összetettebb, nehezebb anyagrészek, és fokozatosan egyre tömörebb tárgyalásmód követ.

Többirányú megközelítés A lineáris algebrai ismeretek, hasonlóan egyéb ismeretekhez, több különböző módon is feldolgozhatók. Bár e könyv semmiképp sem sorolható a formális definíció-tétel-bizonyítás ciklusokra épülő tankönyvek közé, gerincét a *klasszikus* megközelítés adja, mely a definíciók és tételek, valamint a köztük lévő összefüggések precíz megfogalmazására, az algoritmikus ismeretek mintapéldákon való bemutatására, és a tudásnak feladatok megoldásán való elmélyítésére épül. Emellett *több újkeletű technikát* is segítségül hívunk. Ezek egy részének felsőfokú matematika tankönyvben való alkalmazása hazánkban nem gyakori.

Fogalmi és procedurális gondolkodás Elsőként az *absztrakt összefüggések* megértését segítő elemi, konkretizáló, szemléltető példák használatát említjük. Ezek első sorban a *procedurális* gondolkodásban erősebb, a valami fajta *kézzelfoghatóságot* igénylő hallgatóknak készültek. Az absztrakt, fogalmi gondolkodásban otthonos olvasó számára ezek többnyire egyszerű trivialitások, az előbb jelzett hallgatók számára viszont a megértés első lépését jelenthetik.

Vizuális, geometriai megközelítés A második technika a – mérnökhallgatók közt érthetően erős, de korunk kultúrájára egyébként is jellemző – *vizuális* gondolkodásra, és ennek matematikai megfelelőjére, a geometriai intuícióra épít. Szerencsére erre egy lineáris algebra könyv különösen alkalmas a téma számtalan geometriai kapcsolata okán. Könyvünk a geometriai tartalom megismertetése mellett a vizualizáció egyéb lehetőségeit is igénybe veszi (összefüggések absztrakt ábrázolása, dinamikus geometriai programok a könyvet kísérő weboldalon,...).

Algoritmikus megközelítés Részben a számítógépes kultúra elterjedtsége, részben az alkalmazásokban való fontossága miatt könyvünk fontos szerepet szán egyes témák *algoritmikus megközelítésének*.

Alkalmazások A harmadik technika az *alkalmazások* bemutatásához kapcsolódik, ami nem csak matematikán kívüli alkalmazást jelent. A könyvben szereplő alkalmazások nem csak a megtanult anyag felhasználási lehetőségeit tekinti át, ami a lineáris algebra tanulásának motiváló tényezője is lehet, de sok helyütt a megértést – a matematikai fogalmak megértését – segíti, sőt a matematikai fogalomalkotásban, és az absztrakciós készség mélyítésében is szerepet játszik.

Számítógép használata A negyedik eszköz a *számítógép* bevonása az oktatásba. Az életszerűbb problémákkal való foglalkozáshoz, valóságos alkalmazások megértéséhez ma már nélkülözhetetlen a számítógépes eszközök használata. Ezek ráadásul oktatási segédeszközként is használhatók (pl. szemléltetés, vizualizáció), és több numerikus példa vizsgálatát is lehetővé teszik. A diákok számára kínált szoftverek kiválasztásában fontos szempont volt a szabad elérhetőség és az ingyenesség. Bár a számítógép hasznos segédeszköz, a könyv számítógépet nem használó kurzusokhoz is teljes értékű.

Kitekintések Egy ismeret elsajátítását nagyban segíti, ha több szálon kapcsolódik már korábban megszerzett ismeretekhez. A matematika sokak számára idegen terület, mely elvontsága miatt nehezen kapcsolódik bármi máshoz. A könyv szövegét a margón apró *kitekintő* megjegyzések kísérik, melyek a közvetlen alkalmazásokon túli egyéb kapcsolatokat igyekeznek létrehozni. Ilyenek például a történeti megjegyzések, életrajzok, a lineáris algebra fogalmaira vonatkozó etimológiai magyarázatok, lineáris algebrai számítógépes programokhoz kapcsolódó ismeretek, programkódok, de ide tartoznak a további tanulmányokat motiváló, a matematika más területeire kitekintő megjegyzések is. E kitekintéseket néhol internetes linkek erősítik.

Feladatok Didaktikai célból a könyv sok kidolgozott mintapéldát tartalmaz. A feladatokat a könyv többcélú felhasználása érdekében nehézség és tartalom szerint osztályoztuk. A feladat sorszámát a felső indexbe tett – az 'A' betűre emlékeztető – háromszög jelzi, ha *alkalmazási* feladatról van szó (pl. 2.11$^{\blacktriangle}$), és – a számítógép monitorára emlékeztető – négyzet jelzi a *számítógéppel megoldható* feladatokat (pl. 2.12$^{\blacksquare}$). A *fontosnak ítélt* feladatokat egy díszpont (pl. 2.13$^{\bullet}$), a *nehéz*, több időt és némi matematikai képességet igénylő feladatokat csillag jelzi (pl. 2.15*). Végül az *elemi rutinfeladatokat*, egészen egyszerű – néha a képletbehelyettesítés szintjén lévő – alapfeladatokat, amelyek megoldása minden hallgatótól elvárható, egy bíztatásnak szánt karakter jelzi (pl. 2.19$^{\heartsuit}$). Reményeink szerint ezek a matematika iránt kevesebb fogékonyságot mutató hallgatókat is sikerélményhez juttathatják.

Angol szótár Mára a legtöbb szakterületen való előrelépés feltétele az angol szakkifejezések ismerete. A további tanulmányokhoz számtalan forrás érhető el angol nyelven, ezért fontosnak tartottuk, hogy e könyvben használt fontosabb szakszavakat angolul is megadjuk.

A könyv felépítése

A könyv részei A könyv *első részét a lineáris algebra két fő forrásának* tanulmányozására szántuk. E két forrás jól jellemezhető egy-egy alapfogalommal: a vektorral és a lineáris egyenletrendszerrel. Egyikük geometriai, másikuk algebrai jellegű. E fogalmakat az Olvasó korábbi tanulmányaiból már ismeri. E részben ezekből kiindulva, de a lineáris vektortér absztrakt fogalmának ismerete és a mátrixműveletek bevezetése nélkül közel jutunk a lineáris algebra mélyebb fogalmaihoz.

A könyv *második része a „Mátrixok algebrája és geometriája"* címet viseli. Megszívlelve a „Linear Algebra Curriculum Study Group" ajánlásait [1], e rész az első lineáris algebra kurzus középpontjába helyezi a mátrix fogalmát, de a legtöbb könyvvel ellentétben a mátrixok algebrája mellé helyezi a mátrixok hatásának geometriai vizsgálatát is. Ez néhány későbbi fogalom szemléletesebbé tételében is segít, de fontos több modern alkalmazás miatt is (pl. komputer grafika). E részben vezetjük be a determináns fogalmát is, mivel annak egyértelműen geometriai motivációt adunk.

[1]

A könyv *harmadik részének kulcsfogalma a sajátérték*, amit a cím a *mátrix sajátságai* szójátékkal jelez. E részben nem csak a mátrixok diagonalizálása, vagy Jordan-féle normálalakja szerepel, de ide vettük a szinguláris értéket is, melynek fontossága az alkalmazásokban rohamosan növekszik.

A szokásostól eltérő tartalmi megoldások Kiemelünk néhány témát, melynek tárgyalásában eltérünk a bevezető lineáris algebra könyvek többségétől.

1. A vektorok geometriai-fizikai bevezetését fontosnak tartottuk szemben az egyszerűbb, de a kevésbé motivált koordinátás bevezetéssel.
2. Az egyenes és sík egyenletei/egyenletrendszerei osztályozásában a szokásosak helyett (paraméteres, normál) az implicit és explicit elnevezéseket használjuk, ami sokkal szorosabbá teszi e geometriai alakzatok és a lineáris egyenletrendszerek és azok megoldásai közti kapcsolatot. Nevezetesen természetessé válik az egyenletrendszer–implicit alak, egyenletrendszer megoldása–explicit alak párosítás.
3. Az $\mathbb{R}^n$ alterének fogalmát sokkal előbb bevezetjük, mint a vektortér fogalmát. Fontosnak tartjuk annak megmutatását egészen elemi szinten, hogy egy homogén lineáris egyenletrendszer megoldásai alteret alkotnak, és hogy az inhomogén egyenletrendszer megoldásait ennek eltolása adja.
4. Egészen elemi szinten olyan fogalmakat is tárgyalunk, mint az alterek merőlegessége és direkt összege, hogy megértsük az egyenletrendszer megoldásainak szerkezetét.
5. Az első rész végén eljutunk a lineáris algebra alaptételének kimon-

dásáig (a mátrix sortere és nulltere merőleges kiegészítő alterei egymásnak).

6. Az alterek szemléltetésére egy teljesen új módszert, a levéldiagrammokat használjuk.
7. A mátrixok szorzását motivált módon vezetjük be, úgy, mint ami a valósok közti szorzás számtáblázatokra való természetes általánosításából adódik.
8. Csak a magyar nyelvű tankönyvirodalomban újszerű, hogy miután az egyenletrendszerek megoldása az elemi sorműveletekre épül, a mátrixműveletek tárgyalásában fontos szerep jut az elemi mátrixoknak, és az elemi sorműveletek bizonyos sorozatát magában őrző LU-felbontásnak.
9. A determinánsok tárgyalásában is fontosnak tekintettük, hogy e fogalomnak ne valami érthetetlen, égből pottyant definícióját adjuk. A parallelepipedon előjeles térfogatán keresztül való szemléletes bevezetés e cél elérésére kiváló, ráadásul szerencsés módon a felsőbb matematika modern definíciójához vezet.
10. A determinánsok tárgyalásában új a fejezet két alfejeztre osztása. Az első a determinánst, mint sorvektorainak függvényét tárgyalja. Itt szerepel a determináns definíciója, és kiszámításának a gyakorlatban is használt elemi technikája. A másik alfejezet a determinánst, mint elemeinek függvényét vizsgálja. Ez a kifejtési tételt és az ún. elemi szorzatok összegeként való előállítást az általunk ismert könyveknél egyszerűbb módon teszi érthetővé és emészthetővé.
11. A mátrixleképezések geometriája tartalmas fejezet, melyben a forgatás és vetítés transzformációiból messzire jutunk (legkisebb négyzetek módszere, Gram–Schmidt-eljárás, ortogonális mátrixok). E fejezet igen sok része opcionális, egy első kurzusból kihagyható.
12. Ebben a fejezetben tárgyaljuk a pszeudoinverz fogalmát, amelynek egészen elemi, egyszerű és szemléletes definícióját adjuk, mellyel másutt nem találkoztunk.
13. A sajátérték-sajátvektor fogalmának tárgyalása nem tér el a hagyományostól, de mindjárt az első pillanattól nagy hangsúlyt helyezünk a sajátaltér fogalmára is, melynek megértése nélkül nem lehet e témában sokra jutni.
14. Bár egy első lineáris algebra kurzusba nem fér el, de kiemelten fontos helyet kap a szinguláris érték és az SVD is. E fogalmakat is egészen elemi és természetes módon, két olyan ortonormált bázis meghatározásával vezetjük be, melyek egyikének képe a másik elemeinek skalárszorosaiból áll. Ez a sajátérték fogalmának természetes általánosítása.

Szoftverek

Lineáris algebra kurzusokhoz többnyire kétféle szoftver valemelyikét használják: MATLAB-típusú vagy komputer algebra rendszert. Egy kurzus alatt elegendő egyetlen szoftver használata.

Mátrix alapú nyelvek A lineáris algebra a programnyelvek felől legtermészetesebb módon valamely mátrix alapú numerikus matematikai szoftveren keresztül közelíthető meg. A MATLAB-nak és a hozzá hasonló nyelveknek e területen meghatározó szerepük van, ezért a továbbiakban *mátrix alapú nyelveken* csak ezeket értjük. E nyelvek közül négyet emelünk ki. A mintaadó és egyúttal a legelterjedtebb közöttük a MATLAB, mely egy másik, O-Matrix nevű programmal az üzleti szoftverek közé tartozik. A több, főként francia kutatóintézet és egyetem (pl. École Polytechnique, École Centrale Paris, INRIA) valamint cég (pl. a nagy francia autógyárak) konzorciuma által támogatott SciLab és a GNU szoftverek közé tartozó Octave nyílt forráskódú ingyenes szoftverek. E szoftverek mindegyike igen megbízható, nagy tudású, mindegyik komoly referenciákat szerzett valódi műszaki, pénzügyi és tudományos számítások elvégzésével, ezért nyugodt szívvel ajánlható oktatási célokra is. Körültekintő mérlegelés után az Octave mellett döntöttünk, annak ingyenessége és a MATLAB-bal való nagyobb kompatibilitása miatt, így a könyvünkben szereplő mátrix alapú nyelven írt kódok ebben készültek.

Komputer algebra rendszerek A komputer algebra rendszerek (Computer Algebra Systems, rövidítve CAS) oktatásban való használhatósága ma már nem kérdés. Legismertebb ilyen rendszerek a Maple és a Mathematica. Mindkét rendszer igen nagy tudású, képességeik messze felülmúlják azt, amire egy lineáris algebra kurzusnak szüksége lehet. Mivel e szoftverek beszerzése nem olcsó, itt is érdemes az ingyen elérhető lehetőségeket keresni. Egy friss fejlesztés a Sage nevű program. Ennek egyik előnye, hogy saját programnyelv helyett egy széles körben elterjedt és könnyen tanulható nyelvre, a Pythontra épül. További jellemzői: felhasználói felületének egy web-es kereső, melyen keresztül számtalan egyéb computer algebra program is elérhető. Mindez gyors fejlődést és nagy lehetőségeket kínál. A fent felsorolt szoftverek bármelyike ajánlható lineáris algebra kurzushoz. Könyvünk CAS-kódjai a Sage-rendszert használják. A támogatás oka a rendszer ingyenessége és nagy tudása mellett az, hogy mivel webes keresőkben futhat, ezért nem csak saját gépről, hanem az Interneten keresztül valamely szerverről, így akár netbookon, vagy okostelefonon is használható, és ezzel igen rugalmas hozzáférést biztosít.

Jelölések

Képlet	**oldal**	**megjegyzés**
$\mathrm{proj}_{\mathbf{b}}\,\mathbf{a}$	42	**a** vektor **b**-re eső vetülete
$\mathbf{a}\cdot\mathbf{b}$	39	**a** és **b** skaláris szorzata
$\mathbf{a}\times\mathbf{b}$	45	**a** és **b** vektori szorzata
$(\mathbf{a},\mathbf{b})_\angle$	31	**a** és **b** szöge
$(\mathbf{a},\mathbf{b})_\sphericalangle$	44	**a** és **b** irányított szöge
$:=$		definiáló egyenlőség
i, i		imaginárius egység, és az i változó
e, e		az e szám, és az e változó
$\mathbb{C}$, $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{Z}$		komplex, valós, racionális, illetve egész számok
$\mathbb{Z}_m$	446	modulo m vett maradékosztályok
$\mathbb{F}_p=\mathbb{Z}_p$	447	a modulo p (p prím) vett maradékosztályok, a prímelemű test
$\lvert\mathbf{a}\rvert$	31	az **a** vektor abszolút értéke
$\lVert\mathbf{a}\rVert$	31	az **a** vektor normája
a_{ij}, $a_{i,j}$	170	az **A** mátrix i-edik sorának, j-edik oszlopának eleme
$\mathbf{a}_{i*}$	170	az **A** mátrix i-edik sorvektora
$\mathbf{a}_{*j}$, $\mathbf{a}_j$	170	az **A** mátrix j-edik oszlopvektora
$(\mathbf{v})_{\mathcal{B}}$, $[\mathbf{v}]_{\mathcal{B}}$	139	a **v** vektor $\mathcal{B}$ bázisra vonatkozó koordinátás alakja
$[L]_{\mathcal{B}}$		az L lineáris leképezés $\mathcal{B}$ bázisra vonatkozó mátrixa
A, $\mathbf{A}$		az A lineáris leképezés és annak **A** mátrixa a standard bázisban

A jelölések kiválasztásánál azt az elvet követtük, hogy a fontosabb jelölések esetén a nemzetközi angol nyelvű matematikai szakirodalomban elterjedt jelölések valamelyikét követtük. Ez a lebegőpontos számok írására is vonatkozik, tehát nem a magyar irodai szabványt követjük, így nem *tizedesvesszőt*, hanem *tizedespontot* használunk.

I. rész

A lineáris algebra forrásai

A lineáris algebra két fő forrásának egyike a geometria, másika az algebra vidékéről ered. Mindkét forrás jól jellemezhető egy-egy elemi fogalommal: az egyik a vektor, a másik a lineáris egyenletrendszer. E könyv első része e két fogalmat vizsgálja egészen elemi, középiskolai szintről indulva. A lineáris algebra mélyebb fogalmai már itt fölbukkannak, de csak nagyon egyszerű és a legkevésbé absztrakt formájukban. Az első rész végére látni fogjuk, hogy e két forrás már ezen a bevezető szinten szétválaszthatatlanul egyetlen folyammá válik.

1 Vektorok

Általánosan elterjedt nézet szerint a természeti jelenségek leírásakor sok összefüggést számszerű adatokkal, ún. *skalárokkal* vagy *skalármennyiségekkel* fejezünk ki, míg mások leírásához a számadat mellett egy irány megadása is szükséges, és ez utóbbiakat nevezzük *vektoroknak*. A valóság ennél sokkal színesebb: a téridő 4-dimenziós vektoraitól, a bitvektorokon, a gazdasági számítások többszázezer-dimenziós, vagy az internetkeresők által kezelt sokmillió-dimenziós vektorain át a matematika különböző területein gyümölcsöző absztrakt vektorfogalomig széles a skála.

Vektorok a 2- és 3-dimenziós térben

E szakaszban a vektor szemléletes, geometriai fogalmával ismerkedünk. A vektorok összeadásán és skalárral való szorzásán keresztül két kulcsfogalomig – a lineáris kombináció és a lineáris függetlenség fogalmáig – jutunk.

Irányított szakasz, kötött és szabad vektor Tekintsünk egy sárkányrepülőt repülés közben. Számtalan skalár- és vektormennyiség írja le állapotát. A földtől való távolság, a légnyomás, a légellenállási együttható vagy az emelkedés szöge skalármennyiségek, míg vektormennyiségek a sebesség- és gyorsulásvektor, a szárnyra ható felhajtóerő, a gravitációs erő, a szél ereje vagy az elmozdulást leíró vektor.

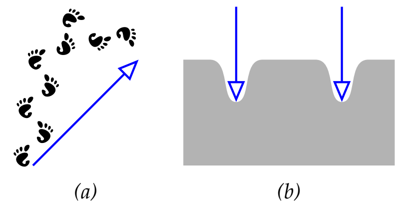
Skalár, skaláris: a *lépcső*, *létra* jelentésű latin scalae (scālae) szóból ered. E szó származéka a skála szó is, mely jól őrzi az eredeti jelentést. A skalár vagy skaláris szót a matematikában szám vagy számszerű értelemben használjuk, például olyankor, amikor egy mennyiségről azt akarjuk hangsúlyozni, hogy irány nélküli, azaz nem vektor jellegű.

A vektor fogalma kapcsolatban van az irányított szakasz fogalmával. Irányított szakaszon olyan szakaszt értünk, melynek végpontjain megadunk egy sorrendet, azaz kijelöljük, hogy melyik a *kezdő*- és melyik a *végpontja*. Más szóhasználatban az irányított szakaszt szokás *kötött vektornak* is nevezni. Az A kezdőpontú és B végpontú irányított szakaszt $\overrightarrow{AB}$ jelöli.

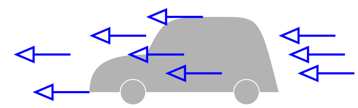
Több jelenség leírására a kötött vektor alkalmas. Természetes példa az elmozdulásvektor, mely megadja, hogy egy tárgy a tér mely pontjából melyik pontjába jutott. Másik példa kötött vektorra a rugalmas testen alakváltozást okozó erőt leíró vektor (1.1. ábra).

Alkalmazásokban gyakran előfordul, hogy egy jelenség különböző irányított szakaszokkal is ugyanúgy leírható. Például ha egy tárgy mozgását egy olyan irányított szakasszal jellemezzük, melynek hossza az időegység alatt megtett út hosszával egyenlő, iránya pedig a mozgás irányát jelzi, akkor mindegy hogy a tér melyik pontjából indítjuk e szakaszt, a mozgást ugyanúgy leírja (1.2. ábra). Ekkor tehát nem a két pont, hanem azok viszonya a kérdés, vagyis például hogy az egyik pont a másiktól milyen *távolságra*, és milyen *irányban* van. Az, hogy a két pont pontosan hol van, nem lényeges. Ekkor bármely két irányított szakasz, mely párhuzamosan egymásba tolható, ugyanazt a viszonyt fejezi ki. Az így kapott fogalmat a fizikában *szabad vektornak* nevezik. Ez a fogalom a lineáris algebra vektor-fogalmának egyik forrása. A *vektor* fogalma az irányított szakaszéból származtatható, annak a feltételnek a hozzáadásával, hogy két irányított szakasz pontosan akkor reprezentálja ugyanazt a vektort, ha párhuzamosan egymásba tolhatók (ld. 1.3 ábra).

Vektorok jelölésére félkövér kisbetűket használunk, pl. $\mathbf{x}$, $\mathbf{u}$, $\mathbf{v}$, stb. A műszaki és fizikai szakirodalomban a félkövér nagy betű is előfordul, pl. az $\mathbf{F}$ erő, a $\mathbf{B}$ indukció is vektormennyiségek.



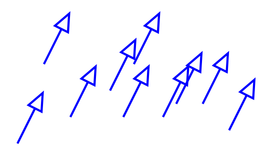
1.1. ábra: Kötött vektorok: *(a)* elmozdulásvektor (lábnyomokkal), *(b)* rugalmas testen alakváltozást okozó erő vektora



1.2. ábra: Példa szabad vektorra

Vektor: a *hordozó, vivő, utazó* jelentésű latin *vector* szóból származik. A tudomány más területein hordozó anyag, az élettanban vírushordozó értelemben használják.

Vektor magadása egy irányított szakasszal Egy vektor megadható egy irányított szakasszal, azaz két pont és a köztük lévő sorrend kijelölésével. Valójában ennyi adat felesleges, hisz egy irányított szakasz önmagával párhuzamosan eltolva ugyanazt a vektort adja meg, ezért például kiköthető, hogy a kezdőpont a sík (tér) egy előre kijelölt rögzített pontja legyen. Ezt a közös kezdőpontot nevezzük *origónak*. Egy origóból induló irányított szakaszt egyértelműen definiálja a végpontja, így a vektorok megadásához elég egyetlen pont, a végpont megadása. Ezzel a sík vagy tér pontjai és vektorai közt kölcsönösen egyértelmű megfeleltetést létesíthetünk (1.4. ábra). Az origóból egy P pontba húzott irányított $\overrightarrow{OP}$ szakaszt a ponthoz tartozó *helyvektornak* is szokás nevezni. Világos, hogy minden vektor reprezentánsai közt pontosan egy helyvektor van.

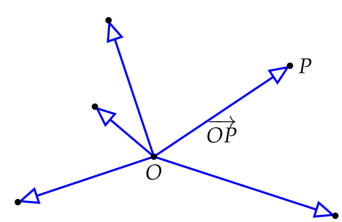


1.3. ábra: Ugyanazt a vektort reprezentáló irányított szakaszok

Vektorok jelölése: Műszaki, fizikai szövegek szedésének tipográfiai szabályait az ISO 31-11 szabvány írja le. Eszerint a vektorok félkövér betűkkel szedendők. Kézírásban aláhúzással, vagy fölé írt nyíllal szokás jelezni a vektort (pl. $\underline{x}$, $\underline{u}$, $\vec{v}\dots$), de körültekintő jelölésrendszer és jegyzetelés esetén elhagyhatók a jelzések. Felsőbb matematikai művek nem használják e szabványt, mondván, kiderül a szövegből, hogy vektort jelölnek-e a betűk (x, u, $v\dots$).

A későbbiekben gyakran fogunk egy ponthalmazt úgy jellemezni, hogy az origóból a ponthalmaz pontjaiba mutató vektorokat jellemezzük. Amikor vektorok végpontjairól beszélünk, mindig a vektorokat megadó, az origóból indított irányított szakaszok végpontjaira gondolunk.

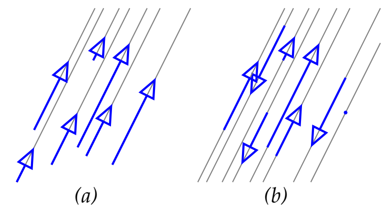
Az olyan vektort, melynek kezdő és végpontja egybeesik, *zérusvektornak* vagy *nullvektornak* nevezzük. A zérusvektort általában félkövér zérussal, azaz $\mathbf{0}$-val jelöljük. A pontok és vektorok közti megfeleltetésben a zérusvektornak az origó felel meg.



1.4. ábra: A sík pontjai és vektorai közti kölcsönösen egyértelmű megfeleltetés: egy P pontnak az $\overrightarrow{OP}$ vektor felel meg, az origónak a nullvektor.

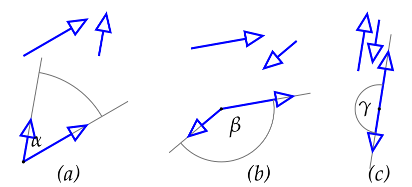
Vektor megadása hossz és irány segítségével Ha tudunk távolságot mérni, és irányt meghatározni, akkor a vektor megadható hosszával és irányával is. Vektor *hosszát*, azaz két végpontjának távolságát a vektor *abszolút értékének* is nevezzük. Az $\mathbf{a}$ vektor abszolút értékét $|\mathbf{a}|$ jelöli. Az abszolút érték másik neve *euklideszi norma*, ugyanis speciális esete egy később részletezendő fogalomnak, a normának. Az $\mathbf{a}$ vektor (euklideszi) normájának jelölése az abszolút értékre emlékeztet: $\|\mathbf{a}\|$.

Az irány fogalmát az 1.20. feladatban definiáljuk. Itt megelégszünk annyival, hogy két nemzérus vektort *azonos irányú*nak vagy *egyirányú*nak nevezünk, ha a kezdőpontjukból induló, és a végpontjukon áthaladó félegyenesek párhuzamos eltolással fedésbe hozhatók (1.5 *(a)* ábra). Két nemzérus vektort *kollineáris*nak vagy *párhuzamos*nak nevezünk, ha az őket tartalmazó egyenesek párhuzamosak. Két vektort, amely párhuzamos, de nem egyirányú, *ellenkező irányú*nak nevezünk (1.5 *(b)* ábra). A zérusvektor irányát tetszőlegesnek tekintjük, így az bármely vektorral egyirányú. Belátható, hogy a vektort egyértelműen meghatározza hossza és iránya.



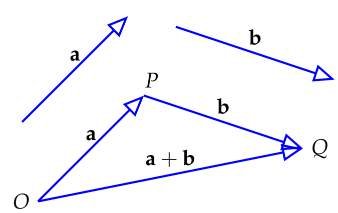
1.5. ábra: *(a)* egyirányú vektorok, *(b)* kollineáris (párhuzamos) vektorok, vannak köztük egyirányúak és ellenkező irányúak

Vektor irányának meghatározásakor gyakran hívjuk segítségül a szög fogalmát. Két vektor szögén azt a szöget értjük, melyet a sík vagy tér egy tetszőleges pontjából kiinduló és az adott vektorokkal egyirányú félegyenesek zárnak be (1.6. ábra). Az $\mathbf{a}$ és $\mathbf{b}$ vektorok szögét $(\mathbf{a}, \mathbf{b})_\angle$ jelöli. Két vektor szöge tehát mindig $0°$ és $180°$ – radiánban mérve 0 és π – közé esik, beleértve a határokat is. Egyirányú vektorok szöge 0, ellenkező irányúaké π.



1.6. ábra: Két vektor szöge ($0 \leq \alpha, \beta, \gamma \leq \pi$). Az ábra felső felén a két adott vektor, alatta szögük meghatározásának módja szerepel.

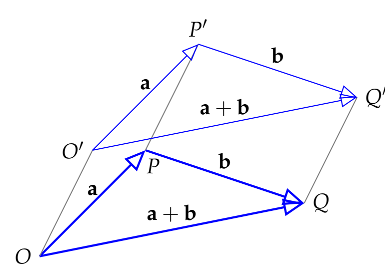
Vektorműveletek a 2- és 3-dimenziós térben A vektorműveletek – az összeadás és a számmal való szorzás – definíciója természetes módon adódik, ha a vektorok legtipikusabb alkalmazásaira gondolunk. Pl. magától értetődő, hogy két elmozdulás összegén az elmozgatások egymás után való elvégzését, egy eltolás kétszeresén egy azonos irányú, de kétszer olyan hosszú eltolást értsünk.



1.7. ábra: Az $\mathbf{a}$ és $\mathbf{b}$ vektor összege

1.1. DEFINÍCIÓ (KÉT VEKTOR ÖSSZEGE – HÁROMSZÖGMÓDSZER). *Legyen adva két vektor,* $\mathbf{a}$ *és* $\mathbf{b}$. *Vegyünk föl egy tetszőleges* O *pontot. Indítsunk belőle egy* $\mathbf{a}$*-val egyenlő* $\overrightarrow{OP}$ *vektort, ennek végpontjából pedig egy* $\mathbf{b}$*-vel egyenlő* $\overrightarrow{PQ}$ *vektort. Az* $\overrightarrow{OQ}$ *vektort az* $\mathbf{a}$ *és* $\mathbf{b}$ *vektorok összegének nevezzük és* $\mathbf{a}+\mathbf{b}$*-vel jelöljük (ld. 1.7. ábra).*

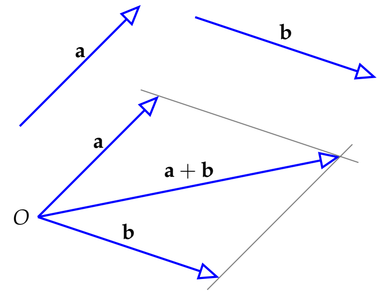
Könnyen belátható, hogy az eredmény független az O pont megválasztásától, tehát vektorok összeadásának művelete definiálható e módszerrel (a bizonyítás leolvasható az 1.8. ábráról).



1.8. ábra: Az összeg független az O pont megválasztásától, ugyanis $\overrightarrow{OQ} = \overrightarrow{O'Q'}$.

Egy másik módszert is ismertetünk két nem kollineáris vektor összegének megszerkesztésére:

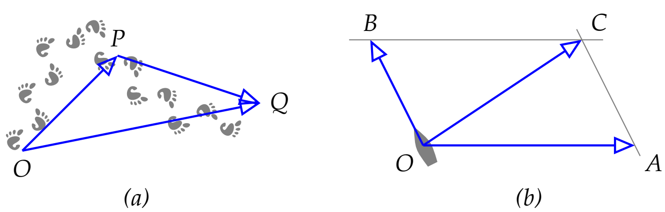
1.2. állítás (Parallelogramma-módszer). *A közös kezdőpontból indított* $\mathbf{a}$ *és* $\mathbf{b}$ *vektorok összege megkapható abból a parallelogrammából, melynek két szomszédos oldala* $\mathbf{a}$ *és* $\mathbf{b}$*, ekkor az összeg a közös kezdőpontból indított, és a parallelogramma szemközti csúcsába futó vektor.*



1.9. ábra: Parallelogramma-módszer

► Ha $\mathbf{a}$ és $\mathbf{b}$ *nem kollineárisak*, akkor összegük pl. megkapható úgy, hogy $\mathbf{a}$ végpontján át egy $\mathbf{b}$ egyenesével, $\mathbf{b}$ végpontján át egy $\mathbf{a}$ egyenesével párhuzamos egyenest húzunk. A közös kezdőpontból e két egyenes metszéspontjába futó vektor lesz az összeg (ld. 1.9. ábra).

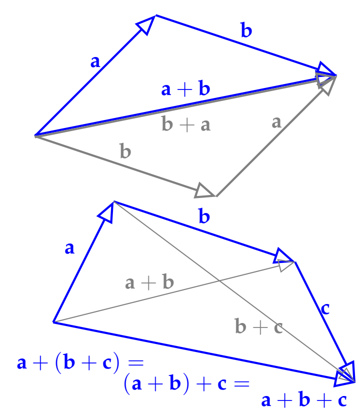
Az alkalmazásokban hol a háromszög-, hol a parallelogramma-módszer tűnik kézenfekvőbbnek (ld. 1.10).



1.10. ábra: Az *(a)* ábrán a lábnyomok O-ból P-be, majd onnan Q-ba vezetnek. Az $\overrightarrow{OP}$ és a $\overrightarrow{PQ}$ elmozdulásvektorok összege $\overrightarrow{OQ}$ (háromszögmódszer). A *(b)* ábrán a csónak az $\overrightarrow{OB}$ irányba evez, de a folyó $\overrightarrow{OA}$ irányba folyik. A két sebesség eredője, azaz összege $\overrightarrow{OC}$ (parallelogramma módszer).

Ha $\mathbf{a}$ és $\mathbf{b}$ két térbeli vektor, akkor a háromszögmódszerben és a parallelogramma-módszerben is az $\mathbf{a}$, $\mathbf{b}$ és $\mathbf{a}+\mathbf{b}$ vektorokat reprezentáló irányított szakaszok egy síkba esnek. Általában azt mondjuk, hogy néhány térbeli vektor egy síkba esik, más szóval *komplanáris*, ha van olyan sík, hogy mindegyik vektort reprezentáló irányított szakasz párhuzamosan betolható e síkba. Eszerint tehát az $\mathbf{a}$, $\mathbf{b}$ és $\mathbf{a}+\mathbf{b}$ vektorok mindig komplanárisak.

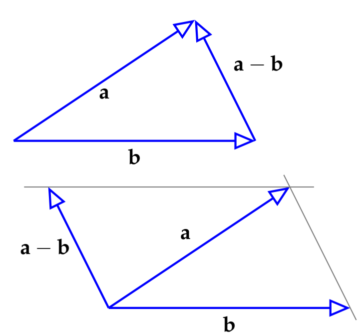
A vektorösszeadás két fontos tulajdonsága, kommutativitása ($\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$) és asszociativitása ($\mathbf{a}+(\mathbf{b}+\mathbf{c})=(\mathbf{a}+\mathbf{b})+\mathbf{c}$) könnyen leolvasható az 1.11. ábráról. Az asszociativitás következtében több tag összeadásánál elhagyható a zárójel, például az ábrabeli három vektor összegére $\mathbf{a}+\mathbf{b}+\mathbf{c}$ írható.



1.11. ábra: A vektorösszeadás kommutativitása és asszociativitása.

Az $\mathbf{a}$ és $\mathbf{b}$ vektorokat közös kezdőpontból indítva – a háromszögmódszerrel – azonnal látható, hogy csak egyetlen olyan $\mathbf{x}$ vektor létezik, melyre $\mathbf{a}=\mathbf{b}+\mathbf{x}$ (ld. 1.12 (a) ábra). Ennek felhasználásával definiálható vektorok különbsége.

1.3. definíció (Vektorok különbsége). *Adva van az* $\mathbf{a}$ *és* $\mathbf{b}$ *vektor. Azt az egyértelműen létező* $\mathbf{x}$ *vektort, melyre* $\mathbf{a}=\mathbf{b}+\mathbf{x}$*, az* $\mathbf{a}$ *és* $\mathbf{b}$ *különbségének nevezzük és* $\mathbf{a}-\mathbf{b}$*-vel jelöljük.*

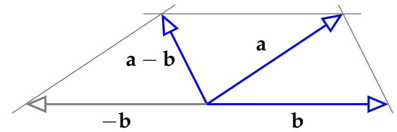


1.12. ábra: A különbségvektor meghatározása háromszög- és parallelogramma-módszerrel.

Könnyen fejben tartható a különbségvektor megszerkesztése akár a háromszög-, akár a parallelogrammamódszerrel (ld. 1.12. ábra), ha arra gondolunk, hogy $\mathbf{a}-\mathbf{b}$ az a vektor, melyet $\mathbf{b}$-hez adva $\mathbf{a}$-t kapunk, azaz

$$\mathbf{a}=\mathbf{b}+(\mathbf{a}-\mathbf{b}).$$

Az 1.13. ábráról az is leolvasható, hogy ha a $\mathbf{b}$ vektorral egyenlő hosszúságú, de ellenkező irányú vektort $-\mathbf{b}$ jelöli, akkor fönnáll az $\mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b})$ összefüggés, és így az is igaz, hogy $\mathbf{b} + (-\mathbf{b}) = \mathbf{0}$.



1.13. ábra: Az $\mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b})$ szemléltetése.

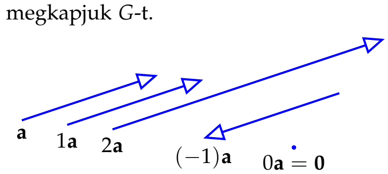
Érdekes megjegyezni, hogy ha P és Q két tetszőleges pont, akkor az $\overrightarrow{OQ} - \overrightarrow{OP}$ vektort akkor is ismerjük, ha az O pontot nem, hisz az a $\overrightarrow{PQ}$ vektor. Sok hasonló jelenség vezetett a *torzor* fogalmához, melyet egy rövid széljegyzetben ismertetünk.

Torzor: a modern matematika fogalma. Néhány példa, mielőtt definiálnánk: (1) Az energiát a newtoni fizikában nem tudjuk mérni, csak az energiakülönbséget. Ha viszont megállapodunk abban, hogy egy adott rendszernek melyik állapota tartozik a 0 energiaszinthez, beszélhetünk a rendszer energiájáról is. (2) A pontba mutató vektor fogalmának nincs értelme, amíg nincs kijelölve az origó, viszont két pontba mutató vektor különbségét az origótól függetlenül is meg tudjuk határozni. (3) Egy f függvény I intervallumon vett határozatlan integrálja $F + C$ alakú, ahol C konstans. Nincs értelme megkérdezni, hogy f egy konkrét primitív függvényében mennyi a C értéke, de két primitív függvény különbsége mindig egy konstans. (4) Egy hasonló jelenség a zenéből: bármely két hang közti távolság meghatározható, de azt nem mondhatjuk egy hangra, hogy az a „fá”, amíg nem rögzítjük, melyik a „dó”.

A torzort egy *kommutatív csoport* nevű algebrai struktúrával definiálhatjuk, mely egy kommutatív, asszociatív, nullelemes, invertálható művelettel ellátott és e műveletre zárt halmaz. Kommutatív csoport például a valósok az összeadásra nézve, a vektorok az összeadásra nézve, vagy $\mathbb{Z}_{12}$ az összeadásra nézve (ld. az 1.9.. példát és az 1.10.. definíciót). Legyen G egy kommutatív csoport, és X egy nem üres halmaz, melyen definiálva van bármely két elem különbsége, ami G-beli. Ekkor X-et G-*torzornak* nevezzük, ha bármely $x_0, x_1, x_2 \in X$ elem esetén, ha $x_1 - x_0 = g_1$ és $x_2 - x_0 = g_2$, akkor $x_1 - x_2 = g_1 - g_2$. Másként fogalmazva, X őrzi G struktúráját a zéruselem nélkül úgy, hogy bármely elemét zéruselemnek választva azonnal megkapjuk G-t.

1.4. definíció (Vektor szorzása skalárral). *Legyen k valós szám. Az $\mathbf{a}$ vektor k-szorosán azt a vektort értjük, melynek hossza az $\mathbf{a}$ hosszának $|k|$-szorosa, iránya*

- *tetszőleges, ha $k = 0$ vagy $\mathbf{a} = \mathbf{0}$,*
- *megegyezik $\mathbf{a}$ irányával, ha $k > 0$, és*
- *ellentétes, ha $k < 0$ (ld. 1.14. ábra).*



1.14. ábra: Vektor skalárszorosai

A skalárral való szorzás definíciójából azonnal látszik, hogy minden $\mathbf{a}$ vektorra $1\mathbf{a} = \mathbf{a}$, $0\mathbf{a} = \mathbf{0}$ és $(-1)\mathbf{a} = -\mathbf{a}$.

E paragrafus végén összefoglaljuk a vektorműveletek legfontosabb tulajdonságait, melyek segítségével később általánosítani fogjuk a vektor fogalmát. Az eddig nem bizonyított tulajdonságok igazolását az olvasóra hagyjuk.

1.5. tétel (A vektorműveletek tulajdonságai). *Ha $\mathbf{a}$, $\mathbf{b}$ és $\mathbf{c}$ a 2- vagy 3-dimenziós tér tetszőleges vektorai, $\mathbf{0}$ a zérusvektor és r, s két tetszőleges valós szám, akkor fönnállnak az alábbi azonosságok:*

a)	$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$	*e)*	$r(s\mathbf{a}) = (rs)\mathbf{a}$
b)	$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$	*f)*	$r(\mathbf{a} + \mathbf{b}) = r\mathbf{a} + r\mathbf{b}$
c)	$\mathbf{a} + \mathbf{0} = \mathbf{a}$	*g)*	$(r + s)\mathbf{a} = r\mathbf{a} + s\mathbf{a}$
d)	$\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$	*h)*	$1\mathbf{a} = \mathbf{a}$ *és* $0\mathbf{a} = \mathbf{0}$

A lineáris kombináció definíciója Ha vektorokra a skalárral való szorzás és az összeadás műveletét alkalmazzuk, akkor e vektorok egy lineáris kombinációját kapjuk. Pontosabban:

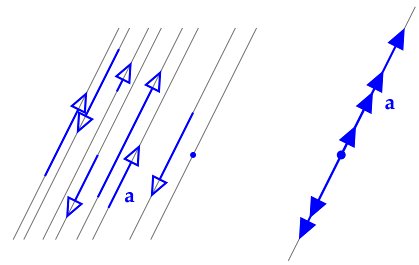
1.6. definíció (Lineáris kombináció). *Az $\mathbf{a}_1$, $\mathbf{a}_2$,… $\mathbf{a}_k$ vektorok* lineáris kombináció*ján egy*

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \ldots + c_k\mathbf{a}_k$$

alakú vektort értünk, ahol $c_1, c_2, \ldots, c_k$ valós számok. Azt mondjuk, hogy a $\mathbf{v}$ vektor előáll *az $\mathbf{a}_1$, $\mathbf{a}_2$,…, $\mathbf{a}_k$ vektorok lineáris kombinációjaként, ha vannak olyan $c_1, c_2, \ldots, c_k$ valós számok, hogy $\mathbf{v} = c_1\mathbf{a}_1 + \ldots + c_k\mathbf{a}_k$.*

Ha egy vektort egy skalárral beszorzunk, az előző definíció szerint egy lineáris kombinációját kapjuk, mely vele párhuzamos, azaz kollineáris. Így egy nemzérus vektor összes lineáris kombinációja csupa vele párhuzamos vektor (ld. 1.15. ábrát). Ennél több is igaz:

1.7. TÉTEL (VEKTORRAL PÁRHUZAMOS VEKTOROK). *Ha* $\mathbf{a}$ *nem zérusvektor, akkor bármely vele párhuzamos* $\mathbf{v}$ *vektor az* $\mathbf{a}$ *skalárszorosa, azaz van olyan* c *valós szám, hogy* $\mathbf{v} = c\mathbf{a}$*, más szóval* $\mathbf{v}$ *előáll az* $\mathbf{a}$ *valamely lineáris kombinációjaként. Ez az előállítás egyértelmű.*



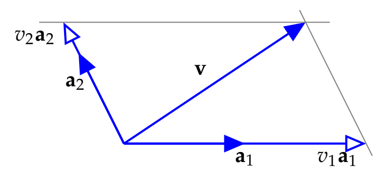
1.15. ábra: Egy nemzérus $\mathbf{a}$ vektor, és néhány lineáris kombinációja kétféle reprezentációban.

BIZONYÍTÁS. Ha a két vektor egyirányú, az előállításban szereplő c konstans egyszerűen a $\mathbf{v}$ és $\mathbf{a}$ vektorok abszolút értékének hányadosa, ha ellenkező irányúak, e hányados (-1)-szerese. □

E tétel következménye, hogy ha $\mathbf{a}$ nem zérusvektor, akkor az $\mathbf{a}$ összes lineáris kombinációjának halmaza és az $\mathbf{a}$-val párhuzamos vektorok halmaza megegyezik. Másként fogalmazva: egy nemzérus vektor összes lineáris kombinációjának végpontja egy *origón átmenő egyenest ad*.

A háromszögmódszerből jól látszik, hogy tetszőleges két vektor bármely lineáris kombinációja velük komplanáris vektor lesz. Az állítás megfordítása is igaz:

1.8. TÉTEL (KÉT VEKTORRAL EGY SÍKBA ESŐ VEKTOROK). *Ha* $\mathbf{a}_1$ *és* $\mathbf{a}_2$ *nem párhuzamos vektorok, akkor bármely velük egy síkba eső* $\mathbf{v}$ *vektor előáll az* $\mathbf{a}_1$ *és* $\mathbf{a}_2$ *valamely lineáris kombinációjaként, azaz van olyan* v_1 *és* v_2 *konstans, hogy* $\mathbf{v} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2$*. Ez az előállítás egyértelmű.*



1.16. ábra: A $\mathbf{v}$ egyértelműen előáll $\mathbf{v} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2$ alakban, ha $\mathbf{a}_1$ és $\mathbf{a}_2$ nem párhuzamos.

BIZONYÍTÁS. A bizonyításnak a felbontás létezését biztosító része könnyen leolvasható az 1.16. ábráról. A $\mathbf{v}$ végpontjából húzzunk az $\mathbf{a}_1$ és az $\mathbf{a}_2$ vektorokkal párhuzamos egyeneseket. Az így létrejött nem elfajuló parallelogramma két oldala az előző tétel szerint $\mathbf{a}_1$, illetve $\mathbf{a}_2$ konstansszorosa, melyek összege a parallelogramma szabály szerint épp $\mathbf{v}$. Előállítottuk tehát $\mathbf{v}$-t $\mathbf{a}_1$ és $\mathbf{a}_2$ lineáris kombinációjaként. Meg kell még mutatnunk, hogy ez az előállítás egyértelmű. Legyen

$$\mathbf{v} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2 = w_1\mathbf{a}_1 + w_2\mathbf{a}_2.$$

a $\mathbf{v}$ vektor két előállítása. Ekkor átrendezés után $(v_1 - w_1)\mathbf{a}_1 = (w_2 - v_2)\mathbf{a}_2$. Mivel $\mathbf{a}_1$ és $\mathbf{a}_2$ nem kollineárisak, konstansszorosaik csak akkor egyezhetnek meg, ha mindkettő a zérusvektor. Ugyanakkor $\mathbf{a}_1 \neq \mathbf{0}$ és $\mathbf{a}_2 \neq \mathbf{0}$, ezért az előző egyenlőség csak akkor áll fönn, ha $(v_1 - w_1) = (w_2 - v_2) = 0$, azaz ha $v_1 = w_1$ és $v_2 = w_2$. Tehát a felbontás egyértelmű. □

Látható tehát, hogy két nem párhuzamos vektor összes lineáris kombinációjának halmaza megegyezik a két vektorral komplanáris vektorok halmazával, egyszerűbben fogalmazva két nem párhuzamos vektor összes lineáris kombinációjának végpontja *egy origón átmenő síkot ad*.

Abban nincs semmi meglepő, hogy a tér három nem egy síkba eső vektorának bármely lineáris kombinációja térbeli vektor, az állítás megfordítása viszont igen fontos:

1.9. TÉTEL (TÉRBELI VEKTOROK). *Ha* $\mathbf{a}_1$, $\mathbf{a}_2$ *és* $\mathbf{a}_3$ *nem egy síkba eső vektorok, akkor a tér bármely* $\mathbf{v}$ *vektora előáll az* $\mathbf{a}_1$, $\mathbf{a}_2$ *és* $\mathbf{a}_3$ *valamely lineáris kombinációjaként, azaz van olyan* v_1, v_2 *és* v_3 *konstans, hogy*

$$\mathbf{v} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3. \tag{1.1}$$

Ez az előállítás egyértelmű.

BIZONYÍTÁS. A bizonyítás alapötlete az, hogy a $\mathbf{v}$ vektor V végpontján át párhuzamos egyenest húzunk az $\mathbf{a}_3$ vektorral, mely az $\mathbf{a}_1$ és $\mathbf{a}_2$ vektorok síkját egy C pontban metszi. Az $\overrightarrow{OC}$ vektor az előző tétel szerint egyértelműen előáll $\mathbf{a}_1$ és $\mathbf{a}_2$ lineáris kombinációjaként (1.17. ábra), azaz $\overrightarrow{OC} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2$. Másrészt $\mathbf{v} = \overrightarrow{OV} = \overrightarrow{OC} + \overrightarrow{CV}$, ahol $\overrightarrow{CV} \parallel \mathbf{a}_3$, így $\overrightarrow{CV} = v_3\mathbf{a}_3$ valamely v_3 valósra. Tehát $\mathbf{v} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3$.

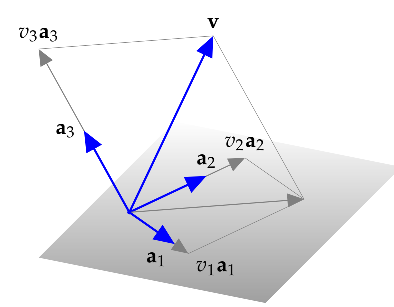
Be kell még látnunk az előállítás egyértelműségét! Tegyük fel, hogy

$$\mathbf{v} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3 = w_1\mathbf{a}_1 + w_2\mathbf{a}_2 + w_3\mathbf{a}_3$$

a $\mathbf{v}$ két felbontása. Ekkor $(v_1 - w_1)\mathbf{a}_1 + (v_2 - w_2)\mathbf{a}_2 + (v_3 - w_3)\mathbf{a}_3 = \mathbf{0}$. Így ha $v_1 \neq w_1$, akkor $\mathbf{a}_1$ kifejezhető $\mathbf{a}_2$ és $\mathbf{a}_3$ lineáris kombinációjaként:

$$\mathbf{a}_1 = -\frac{v_2 - w_2}{v_1 - w_1}\mathbf{a}_2 - \frac{v_3 - w_3}{v_1 - w_1}\mathbf{a}_3.$$

Ez ellentmond annak, hogy $\mathbf{a}_1$, $\mathbf{a}_2$ és $\mathbf{a}_3$ nem esnek egy síkba. Így tehát $v_1 = w_1$. Hasonlóan kapjuk, hogy $v_2 = w_2$ és $v_3 = w_3$, azaz az (1.1) előállítás egyértelmű. □



1.17. ábra: A térbeli $\mathbf{v}$ vektor előállítása három nem egy síkba eső vektor lineáris kombinációjaként.

Lineáris függetlenség Az előző két tételből világos, hogy a tér három vektora vagy egy síkba esik, ekkor valamelyikük a másik kettő lineáris kombinációja, vagy nem esik egy síkba, és akkor egyikük sem áll elő a másik kettő lineáris kombinációjaként. Ekkor viszont a tér minden vektora előáll az ő lineáris kombinációjukként. Látjuk, alapvető, hogy egy vektor kifejezhető-e más vektorok lineáris kombinációjaként.

1.10. DEFINÍCIÓ (VEKTOROK FÜGGETLENSÉGE). *Azt mondjuk, hogy egy* $\mathbf{v}$ *vektor* lineárisan független *az* $\mathbf{a}_1$, $\mathbf{a}_2$,... $\mathbf{a}_n$ *(*$n \geq 1$*) vektoroktól, ha* $\mathbf{v}$ *nem fejezhető ki e vektorok lineáris kombinációjaként. Azt mondjuk, hogy az* $\mathbf{a}_1$, $\mathbf{a}_2$,... $\mathbf{a}_n$ *(*$n \geq 2$*) vektorok* lineárisan függetlenek *ha e vektorok egyike sem fejezhető ki a többi lineáris kombinációjaként. Ha legalább egyikük kifejezhető a többi lineáris kombinációjaként, azaz legalább egyikük* lineárisan függ *a többitől, akkor e vektorokat* lineárisan összefüggőknek *nevezzük. Az egyetlen vektorból álló vektorrendszert lineárisan függetlennek tekintjük, ha a vektor nem a zérusvektor.*

Például egy térbeli vektor, mely nem esik egy adott síkba, független a síkba eső vektorok bármely rendszerétől (**??**. ábra). Egy kocka egy csúcsból kiinduló élvektorai lineárisan függetlenek (**??**. ábra). Általában: bármely két nem kollineáris vektor lineárisan független, hasonlóképp, a tér bármely három nem komplanáris, azaz nem egy síkba eső vektora lineárisan független.

Az 1.8. tétel tehát a következőképp fogalmazható át:

1.11. TÉTEL (SÍKBELI VEKTOR FELBONTÁSA). *Ha $\mathbf{a}_1$ és $\mathbf{a}_2$ egy sík két lineárisan független vektora, akkor a sík minden $\mathbf{v}$ vektora egyértelműen előáll e vektorok lineáris kombinációjaként, azaz egyértelműen léteznek olyan v_1 és v_2 valós számok, hogy*

$$\mathbf{v} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2.$$

Hasonlóképp az 1.9. tétel így fogalmazható át:

1.12. TÉTEL (TÉRBELI VEKTOR FELBONTÁSA). *Ha $\mathbf{a}_1$, $\mathbf{a}_2$ és $\mathbf{a}_3$ három lineárisan független térbeli vektor, akkor a tér minden $\mathbf{v}$ vektora egyértelműen előáll e vektorok lineáris kombinációjaként, azaz egyértelműen léteznek olyan v_1, v_2 és v_3 valós számok, hogy*

$$\mathbf{v} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3.$$

A koordinátákról szóló szakaszban e két tétel lesz alapja a koordinátarendszer bevezetésének.

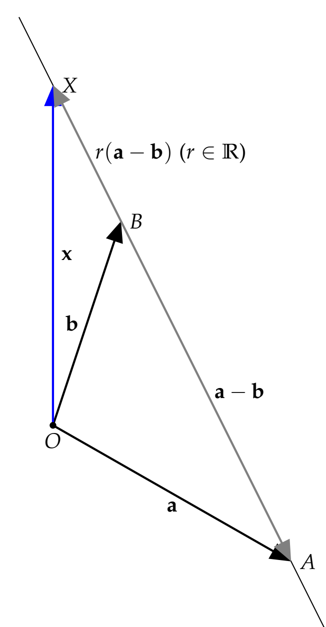
Speciális lineáris kombinációk A sík és a tér bizonyos konfigurációi jól jellemezhetők lineáris kombinációkkal, ha a kombinációs együtthatókra bizonyos feltételeket kötünk ki.

1.13. ÁLLÍTÁS (KÉT PONTON ÁTMENŐ EGYENES JELLEMZÉSE). *Legyen O, A és B a sík vagy a tér három pontja. Az $r\overrightarrow{OA} + s\overrightarrow{OB}$ alakú lineáris kombináció végpontja pontosan akkor mutat az A és B ponton átmenő egyenes egy pontjába, ha $r + s = 1$.*

BIZONYÍTÁS. Legyen $\mathbf{a} = \overrightarrow{OA}$, $\mathbf{b} = \overrightarrow{OB}$, és $\mathbf{x}$ mutasson az AB egyenes valamely X pontjára, azaz legyen $\mathbf{x} = \overrightarrow{OB} + r\overrightarrow{BA}$ valamilyen r valós számra, tehát

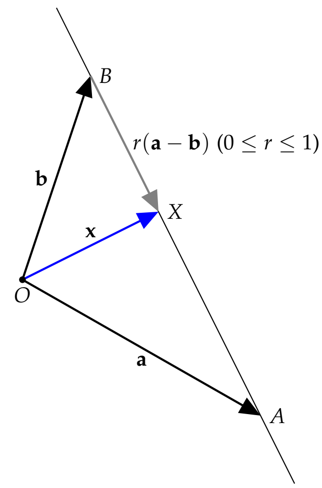
$$\mathbf{x} = \mathbf{b} + r(\mathbf{a} - \mathbf{b}), \text{ azaz } \mathbf{x} = r\mathbf{a} + (1 - r)\mathbf{b}.$$

A fenti gondolatmenet lépésein visszafelé haladva látható, hogy minden valós r számra az $r\mathbf{a} + (1 - r)\mathbf{b}$ vektor végpontja az AB egyenesen van. Fogalmazhatunk úgy is, hogy az $\mathbf{a}$ és $\mathbf{b}$ vektorok végpontján átmenő egyenes összes pontját pontosan azok az $r\mathbf{a} + s\mathbf{b}$ alakú lineáris kombinációk adják, amelyeknél $r + s = 1$. □



1.18. ábra: Az X pont pontosan akkor van az AB egyenesen, ha valamely r és s valósokra $\overrightarrow{OX} = r\overrightarrow{OA} + s\overrightarrow{OB}$, és $r + s = 1$.

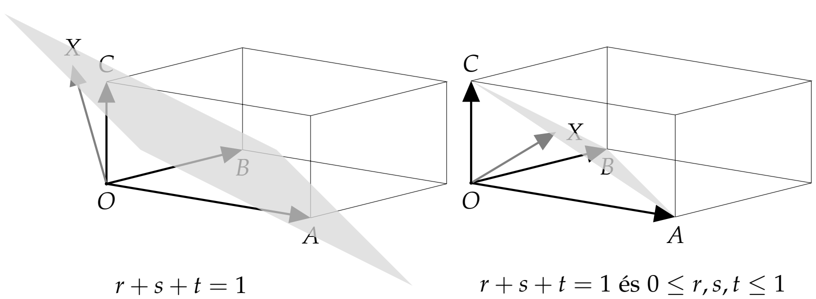
1.14. ÁLLÍTÁS (INTERVALLUM PONTJAINAK JELLEMZÉSE). *Legyen O, A és B a sík vagy a tér három pontja. Az $r\overrightarrow{OA} + s\overrightarrow{OB}$ vektor pontosan akkor mutat az A és B pontot összekötő szakasz valamely pontjába, ha $r + s = 1$ és $0 \le r, s \le 1$.*



1.19. ábra: Az X pont pontosan akkor van az $\overline{AB}$ intervallumban, ha valamely 0 és 1 közé eső r és s valósokra $\overrightarrow{OX} = r\overrightarrow{OA} + s\overrightarrow{OB}$, és $r + s = 1$.

BIZONYÍTÁS. Megismételjük az előző feladat megoldását azzal a különbséggel, hogy itt a $\overrightarrow{BX} = r\overrightarrow{BA}$ összefüggés csak 0 és 1 közé eső r értékekre igaz. Tehát $\mathbf{x} = r\mathbf{a} + (1-r)\mathbf{b}$, ahol $0 \le r \le 1$. Másként fogalmazva az $\mathbf{a}$ és $\mathbf{b}$ vektorok végpontjait összekötő szakasz összes pontját pontosan azok a $r\mathbf{a} + s\mathbf{b}$ alakú lineáris kombinációk adják, amelyekben $r + s = 1$ és $0 \le r, s \le 1$. □

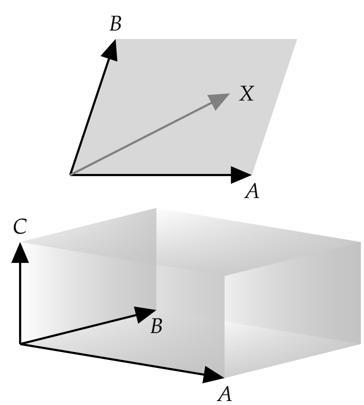
Hasonló összefüggés igaz három vektor esetén is, azaz megmutatható, hogy a nem kollineáris $\mathbf{a}$, $\mathbf{b}$ és $\mathbf{c}$ vektorok végpontjaira fektetett sík pontjaiba pontosan azok a vektorok mutatnak, melyeket $r\mathbf{a} + s\mathbf{b} + t\mathbf{c}$ alakba írva $r + s + t = 1$. Ha még azt is kikötjük e három számról, hogy legyen $0 \le r, s, t \le 1$, akkor az $r\mathbf{a} + s\mathbf{b} + t\mathbf{c}$ alakú vektorok a három vektor végpontja által meghatározott háromszög pontjaiba mutatnak (ld. az 1.20. ábrát és a **??** feladatot).



1.20. ábra: Az X pont pontosan akkor esik az A, B és C pontokon átmenő síkba, ha $\overrightarrow{OX} = r\overrightarrow{OA} + s\overrightarrow{OB} + t\overrightarrow{OC}$ és $r + s + t = 1$. Az X az ABC háromszögbe pedig pontosan akkor esik, ha ezen kívül még $0 \le r, s, t \le 1$ is fönnáll.

Szemléletesen világos, például a mellékelt 1.21. ábráról leolvasható, de nem bizonyítjuk, hogy két tetszőleges nem kollineáris vektor összes olyan lineáris kombinációja, amelyben az együtthatók 0 és 1 közé esnek, egy parallelogrammát ad. Pontosabban fogalmazva egy $r\mathbf{a} + s\mathbf{b}$ alakú vektor végpontja pontosan akkor tartozik az $\mathbf{a}$ és $\mathbf{b}$ által meghatározott (kifeszített) *parallelogrammához*, ha $0 \le r, s \le 1$.

Hasonló mondható három, nem egy síkba eső vektorról: egy $r\mathbf{a} + s\mathbf{b} + t\mathbf{c}$ alakú vektor végpontja pontosan akkor tartozik az $\mathbf{a}$, $\mathbf{b}$ és $\mathbf{c}$ által kifeszített *parallelepipedonhoz*, ha $0 \le r, s, t \le 1$ (1.21. ábra).



1.21. ábra: A *parallelogramma* és a *parallelepipedon* olyan lineáris kombinációkkal állítható elő, ahol az együtthatók nem negatívak, és összegük legföljebb 1.

Feladatok

Vektorműveletek a 2- és 3-dimenziós térben

1.1. Adva van a síkban két tetszőleges vektor, $\mathbf{a}$ és $\mathbf{b}$. Szerkesszük meg a következő vektorokat: *a)* $\mathbf{c} = 2\mathbf{a} + \mathbf{b}$, *b)* $\mathbf{d} = 2\mathbf{a} - \mathbf{b}$, *c)* $\mathbf{e} = \frac{2}{3}\mathbf{a} + \frac{1}{3}\mathbf{b}$, *d)* $\mathbf{f} = \frac{2}{5}\mathbf{a} + \frac{3}{5}\mathbf{b}$.

1.2. Legyen $\mathbf{u} = \mathbf{a} + \mathbf{b}$, $\mathbf{v} = \mathbf{a} - \mathbf{b}$. Fejezzük ki az $\mathbf{a}$ és $\mathbf{b}$ vektor segítségével a következő vektorokat: *a)* $2\mathbf{u} + 2\mathbf{v}$, *b)* $3\mathbf{u} - 3\mathbf{v}$, *c)* $3\mathbf{u} - \mathbf{v}$, *d)* $2\mathbf{u} - \frac{1}{2}\mathbf{v}$.

1.3. Tekintsük az $ABCD$ négyzetet. Határozzuk meg a következő összegeket! *a)* $\overrightarrow{AB} + \overrightarrow{CD}$, *b)* $\overrightarrow{AB} + \overrightarrow{BC} + \overrightarrow{CD}$, *c)* $\overrightarrow{AB} - \overrightarrow{AC}$, *d)* $\overrightarrow{AB} - \overrightarrow{CB}$, *e)* $\overrightarrow{AC} + \overrightarrow{DB}$, *f)* $\overrightarrow{AC} - \overrightarrow{DB}$, *g)* $2\overrightarrow{AB} + \overrightarrow{BD}$

1.4. Tekintsük az $ABCD$ négyzetet. Jelölje a BC oldal felezőpontját E, a CD oldal felezőpontját F, a négyzet középpontját O. Fejezzük ki az egymásra merőleges $\mathbf{a} = \overrightarrow{AB}$ és $\mathbf{d} = \overrightarrow{AD}$ vektorok segítségével az $\overrightarrow{AE}$, $\overrightarrow{AF}$, $\overrightarrow{AO}$, $\overrightarrow{EF}$, $\overrightarrow{OF}$ vektorokat!

1.5. Tekintsük az $ABCD$ tetraédert! Határozzuk meg az
a) $\overrightarrow{AB} + \overrightarrow{BC} + \overrightarrow{CD} + \overrightarrow{DA}$,
b) $\overrightarrow{AB} - \overrightarrow{CB} + \overrightarrow{CD} - \overrightarrow{AD}$,
c) $\overrightarrow{AD} - \overrightarrow{AC} - \overrightarrow{BD}$
vektorokat.

1.6. Tekintsük a szabályos $ABCDEF$ hatszöget, melynek geometriai középpontját jelölje O. Fejezzük ki az $\mathbf{a} = \overrightarrow{OA}$ és $\mathbf{b} = \overrightarrow{OB}$ vektorok segítségével az *a)* $\overrightarrow{OC}$, *b)* $\overrightarrow{OE}$, *c)* $\overrightarrow{OF}$, *d)* $\overrightarrow{AC}$, *e)* $\overrightarrow{BD}$, *f)* $\overrightarrow{BF}$, *g)* $\overrightarrow{AB} + \overrightarrow{CD} + \overrightarrow{EF}$ vektorokat!

1.7. Adva van n tetszőleges, nem feltétlenül különböző P_1, $P_2, \dots, P_n$ pont a térben. Mivel egyenlő a

$$\overrightarrow{P_1P_2} + \overrightarrow{P_2P_3} + \overrightarrow{P_3P_4} + \dots + \overrightarrow{P_{n-1}P_n}$$

és a

$$\overrightarrow{P_1P_2} + \overrightarrow{P_2P_3} + \overrightarrow{P_3P_4} + \dots + \overrightarrow{P_{n-1}P_n} + \overrightarrow{P_nP_1}$$

összeg?

1.8. Mutassuk meg, hogy az $\mathbf{a}$, $\mathbf{b}$ és $\mathbf{c}$ vektorok pontosan akkor lehetnek (egy esetleg szakasszá vagy ponttá elfajuló) háromszög oldalvektorai, ha az

$$\mathbf{a} + \mathbf{b} + \mathbf{c}, \quad \mathbf{a} + \mathbf{b} - \mathbf{c}, \quad \mathbf{a} - \mathbf{b} + \mathbf{c}, \quad \mathbf{a} - \mathbf{b} - \mathbf{c}$$

vektorok legalább egyike zérus. Másként fogalmazva: ha a három vektor összege $\mathbf{0}$, vagy valamelyik vektor egyenlő a másik kettő összegével.

1.9. Legyen $\mathbf{a}$ és $\mathbf{b}$ két tetszőleges vektor. Mutassuk meg, hogy van olyan (esetleg elfajuló) háromszög, melynek oldalvektorai $2\mathbf{a} - \mathbf{b}$, $\mathbf{a} + 2\mathbf{b}$ és $3\mathbf{a} + \mathbf{b}$.

Lineáris kombináció, lineáris függetlenség

Speciális lineáris kombinációk

1.10.• SZAKASZT $m : n$ ARÁNYBAN OSZTÓ PONT Ha az $\overline{AB}$ szakaszt a P pont úgy bontja ketté, hogy $|\overline{AP}| : |\overline{PB}| = m : n$, akkor bármely O pontra igaz, hogy

$$\overrightarrow{OP} = \frac{n}{m+n}\overrightarrow{OA} + \frac{m}{m+n}\overrightarrow{OB}.$$

Speciálisan, az $\overline{AB}$ szakasz felezőpontjába az

$$\frac{\overrightarrow{OA} + \overrightarrow{OB}}{2}$$

vektor mutat.

Alkalmazás: geometriai bizonyítások vektorokkal

Alkalmazás: szív.....

Távolság, szög, orientáció

A címben jelzett három alapfogalomhoz három – vektorok közti – szorzásművelet visz közelebb. Mindegyik művelet igen szokatlan tulajdonságokkal rendelkezik: az egyik eredményül nem vektort, hanem skalárt ad, a másik nem felcserélhető, és kétváltozós (bináris) műveletként csak a 3-dimenziós térben definiálható, a harmadik pedig nem két- hanem háromváltozós művelet.

Skaláris szorzás A fizikában az erő által végzett munka az út hosszának és az erő elmozdulás irányába eső merőleges vetülete hosszának szorzata. Vagyis két vektorjellegű mennyiségből egy skalármennyiséget kapunk eredményül. Ha $\mathbf{F}$ jelöli az erővektort, $\mathbf{s}$ az elmozdulásvektort, $\mathbf{F}_s$ az erőnek az elmozdulás irányába eső merőleges vetületi vektorát és γ az $\mathbf{F}$ és $\mathbf{s}$ vektorok hajlásszögét, akkor a munka értéke $|\mathbf{F}_s||\mathbf{s}| = |\mathbf{F}||\mathbf{s}|\cos\gamma$. Ez vezet a következő definícióhoz:

1.15. DEFINÍCIÓ (KÉT VEKTOR SKALÁRIS SZORZATA). *Két vektor* skaláris szorzatán *a vektorok abszolút értékének és az általuk bezárt szög koszinuszának szorzatát értjük. Az* $\mathbf{a}$ *és* $\mathbf{b}$ *vektorok skaláris szorzatát* $\mathbf{a}\cdot\mathbf{b}$ *jelöli, tehát*

$$\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos(\mathbf{a},\mathbf{b})_\angle,$$

ahol a két vektor által bezárt szög $(\mathbf{a},\mathbf{b})_\angle$.

Ha $\mathbf{a}$ és $\mathbf{b}$ valamelyike zérusvektor, akkor a két vektor szöge, s így annak koszinusza sem határozható meg egyértelműen, a skaláris szorzat viszont ekkor is egyértelmű, éspedig 0, hisz a zérusvektor abszolút értéke 0, és 0 bármivel vett szorzata 0.

Szokás $\mathbf{a}$ és $\mathbf{b}$ skaláris szorzatát $\mathbf{ab}$-vel is jelölni, de ezt egy később bevezetendő művelettel (a mátrixszorzással) való összekeverés elkerülése érdekében e könyvben nem fogjuk használni.

1.16. PÉLDA (SKALÁRIS SZORZAT KISZÁMÍTÁSA A DEFINÍCIÓ ALAPJÁN). *Mennyi a skaláris szorzata egy* 1 *és egy* 2 *egység hosszú, és egymással* $60°$*-os szöget bezáró két vektornak?*

MEGOLDÁS. A szorzat $1\cdot 2\cdot\cos 60° = 1\cdot 2\cdot\frac{1}{2} = 1$. □

1.17. TÉTEL (MIKOR 0 A SKALÁRIS SZORZAT?). *Két vektor skaláris szorzata pontosan akkor* 0, *ha a két vektor merőleges egymásra.*

BIZONYÍTÁS. ($\Longleftarrow$) Ha $\mathbf{a}\perp\mathbf{b}$, akkor $(\mathbf{a},\mathbf{b})_\angle = \pi/2$, azaz $\cos(\mathbf{a},\mathbf{b})_\angle = 0$, tehát $\mathbf{a}\cdot\mathbf{b} = 0$.

($\Longrightarrow$) Ha $\mathbf{a}\cdot\mathbf{b} = 0$, azaz $|\mathbf{a}||\mathbf{b}|\cos(\mathbf{a},\mathbf{b})_\angle = 0$, akkor $|\mathbf{a}| = 0$, $|\mathbf{b}| = 0$ vagy $\cos(\mathbf{a},\mathbf{b})_\angle = 0$. Ha valamelyik vektor zérusvektor, akkor iránya bármely vektoréra merőlegesnek tekinthető. Ha viszont $\mathbf{a}\neq\mathbf{0}$ és $\mathbf{b}\neq$

$\mathbf{0}$, akkor $\cos(\mathbf{a},\mathbf{b})_\angle = 0$, a cos függvénynek a $[0,\pi]$ intervallumban csak $\pi/2$-ben van zérushelye, tehát a két vektor merőleges egymásra. □

1.18. TÉTEL (A SKALÁRIS SZORZÁS MŰVELETI TULAJDONSÁGAI). *Ha* $\mathbf{a}$, $\mathbf{b}$ *és* $\mathbf{c}$ *tetszőleges térbeli (síkbeli) vektorok és r tetszőleges valós szám, akkor igazak az alábbi összefüggések:*

a) $\mathbf{a}\cdot\mathbf{b} = \mathbf{b}\cdot\mathbf{a}$ *(kommutativitás)*
b) $(\mathbf{a}+\mathbf{b})\cdot\mathbf{c} = \mathbf{a}\cdot\mathbf{c}+\mathbf{b}\cdot\mathbf{c}$ *(disztributivitás)*
c) $r(\mathbf{a}\cdot\mathbf{b}) = (r\mathbf{a})\cdot\mathbf{b} = \mathbf{a}\cdot(r\mathbf{b})$
d) $\mathbf{a}\cdot\mathbf{a} > 0$, *ha* $\mathbf{a}\neq\mathbf{0}$, *és* $\mathbf{a}\cdot\mathbf{a} = 0$, *ha* $\mathbf{a}=\mathbf{0}$.

A bizonyítást az olvasóra hagyjuk.

Mivel két vektor skaláris szorzata skalár, ezért az asszociativitás (csoportosíthatóság) kérdése föl sem vethető, hisz az $(\mathbf{a}\cdot\mathbf{b})\mathbf{c}$ szorzatban két különböző szorzásművelet szerepel. Mindezzel együtt $(\mathbf{a}\cdot\mathbf{b})\mathbf{c}\neq\mathbf{a}(\mathbf{b}\cdot\mathbf{c})$ (ld. az 1.12. feladatot).

Hosszúság és szög Egy vektor hossza, és ezzel két pont távolsága, valamint két vektor hajlásszöge kifejezhető a skaláris szorzat segítségével.

Egy tetszőleges $\mathbf{a}$ vektorra $\mathbf{a}\cdot\mathbf{a} = |\mathbf{a}||\mathbf{a}|\cos 0 = |\mathbf{a}||\mathbf{a}|$, tehát

$$|\mathbf{a}|^2 = \mathbf{a}\cdot\mathbf{a}, \text{ azaz } |\mathbf{a}| = \sqrt{\mathbf{a}\cdot\mathbf{a}}.$$

E képlet szerint tehát egy *vektor hossza* megegyezik az önmagával vett skaláris szorzat gyökével. Ebből az is adódik, hogy *két pont távolsága* megegyezik az őket összekötő vektor önmagával vett skaláris szorzatának négyzetgyökével.

Két pontot összekötő vektor egyenlő az oda mutató helyvektorok különbségével, így ha a két pontba mutató helyvektor $\mathbf{a}$ és $\mathbf{b}$, akkor a pontok távolsága – és ezt fogjuk a vektorok távolságának is tekinteni

$$d(\mathbf{a},\mathbf{b}) = |\mathbf{a}-\mathbf{b}|.$$

Két vektor skaláris szorzatának és a vektorok hosszának ismeretében a *szögük* meghatározható:

$$\cos(\mathbf{a},\mathbf{b})_\angle = \frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}||\mathbf{b}|} \qquad (1.2)$$

Pithagorász-tétel A távolságot vagy hosszúságot skaláris szorzattal is ki tudjuk fejezni, így segítségével a rá vonatkozó összefüggések is vizsgálhatók.

1.19. TÉTEL (PITHAGORÁSZ-TÉTEL). *Az* $\mathbf{a}$ *és* $\mathbf{b}$ *vektorokra pontosan akkor teljesül az*

$$|\mathbf{a}+\mathbf{b}|^2 = |\mathbf{a}|^2+|\mathbf{b}|^2$$

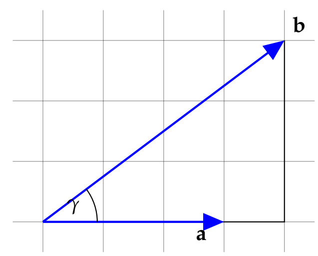
összefüggés, ha $\mathbf{a}$ *és* $\mathbf{b}$ *merőlegesek egymásra.*

BIZONYÍTÁS.

$$
\begin{aligned}
|\mathbf{a}+\mathbf{b}|^2 &= (\mathbf{a}+\mathbf{b})\cdot(\mathbf{a}+\mathbf{b}) \\
&= \mathbf{a}\cdot\mathbf{a}+\mathbf{a}\cdot\mathbf{b}+\mathbf{b}\cdot\mathbf{a}+\mathbf{b}\cdot\mathbf{b} && \text{disztributivitás} \\
&= \mathbf{a}\cdot\mathbf{a}+2(\mathbf{a}\cdot\mathbf{b})+\mathbf{b}\cdot\mathbf{b} && \text{kommutativitás} \\
&\stackrel{?}{=} \mathbf{a}\cdot\mathbf{a}+\mathbf{b}\cdot\mathbf{b} && ? \\
&= |\mathbf{a}|^2+|\mathbf{b}|^2,
\end{aligned}
$$

Világos, hogy a ?-lel megjelölt egyenlőség pontosan akkor teljesül, ha $\mathbf{a}\cdot\mathbf{b}=0$, azaz ha **a** és **b** merőlegesek egymásra. □

1.20. PÉLDA (SKALÁRIS SZORZAT KISZÁMÍTÁSA). *Számítsuk ki az ábrán látható két vektor skaláris szorzatát (a szomszédos rácsvonalak távolsága 1 egység).*



1.22. ábra: Két vektor skaláris szorzata

MEGOLDÁS. Az **a** vektor hossza 3, a **b** vektor hossza a Pithagorász-tétellel kiszámolható: $\sqrt{4^2+3^2}=5$, a vektorok által közbezárt szög koszinusza $\cos\gamma=\frac{4}{5}$, így a skaláris szorzat $\mathbf{a}\cdot\mathbf{b}=3\cdot5\cdot\frac{4}{5}=12$. □

Két fontos egyenlőtlenség Mivel a koszinusz függvény értéke abszolút értékben sosem nagyobb 1-nél, ezért a skaláris szorzat definíciójából azonnal látszik, hogy

$$|\mathbf{a}\cdot\mathbf{b}|=|\mathbf{a}||\mathbf{b}||\cos(\mathbf{a},\mathbf{b})_\angle|\le|\mathbf{a}||\mathbf{b}|.$$

Ezzel bizonyítottuk a következő tételt:

1.21. TÉTEL (CAUCHY–BUNYAKOVSZKIJ–SCHWARZ-EGYENLŐTLENSÉG). *Két vektor skaláris szorzatának abszolút értéke sosem nagyobb abszolút értékeik szorzatánál, azaz*

$$|\mathbf{a}\cdot\mathbf{b}|\le|\mathbf{a}||\mathbf{b}|.$$

A Cauchy–Bunyakovszkij–Schwarz-egyenlőtlenség segítségével bizonyítjuk a geometriából jól ismert háromszög-egyenlőtlenséget. Ennek az az értelme, hogy e bizonyítás változtatás nélkül működni fog sokkal általánosabb körülmények között is.

1.22. TÉTEL (HÁROMSZÖG-EGYENLŐTLENSÉG). *Bármely két **a** és **b** vektorra*

$$|\mathbf{a}+\mathbf{b}|\le|\mathbf{a}|+|\mathbf{b}|.$$

BIZONYÍTÁS. Mivel az egyenlőtlenség mindkét oldalán nemnegatív szám áll, ezért vele ekvivalens egyenlőtlenséghez jutunk, ha mindkét oldalt négyzetre emeljük.

$$\begin{aligned}
|\mathbf{a}+\mathbf{b}|^2 &= (\mathbf{a}+\mathbf{b})\cdot(\mathbf{a}+\mathbf{b}) \\
&= |\mathbf{a}|^2 + 2(\mathbf{a}\cdot\mathbf{b}) + |\mathbf{b}|^2 \qquad \text{ld. az 1.19. tétel bizonyítását} \\
&\le |\mathbf{a}|^2 + 2|\mathbf{a}\cdot\mathbf{b}| + |\mathbf{b}|^2 \\
&\le |\mathbf{a}|^2 + 2|\mathbf{a}||\mathbf{b}| + |\mathbf{b}|^2 \\
&= (|\mathbf{a}| + |\mathbf{b}|)^2 .
\end{aligned}$$

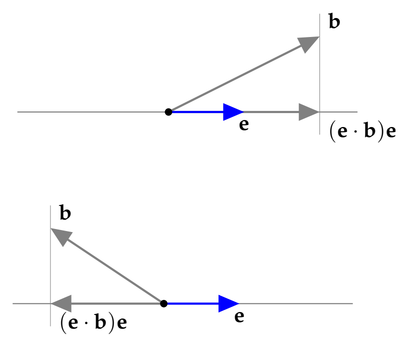
És ezt akartuk bizonyítani. □

Egységvektorral való szorzás és a merőleges vetítés Minden olyan vektort, melynek abszolút értéke 1, *egységvektornak* nevezünk.

Ha $\mathbf{a}$ egy tetszőleges nemzérus vektor, akkor $\mathbf{a}/|\mathbf{a}|$ egységvektor, ugyanis abszolút értéke 1:

$$\left|\frac{\mathbf{a}}{|\mathbf{a}|}\right| = \frac{1}{|\mathbf{a}|}|\mathbf{a}| = 1.$$

1.23. tétel (Egységvektorral való szorzás geometriai jelentése). *Ha $\mathbf{e}$ egységvektor, akkor a $\hat{\mathbf{b}} = (\mathbf{e}\cdot\mathbf{b})\mathbf{e}$ vektor a $\mathbf{b}$ vektornak az $\mathbf{e}$ egyenesére való merőleges vetülete. Az $\mathbf{e}\cdot\mathbf{b}$ szorzat e vetület előjeles hossza, mely pozitív, ha $\hat{\mathbf{b}}$ és $\mathbf{e}$ egyirányúak, és negatív, ha ellenkező irányúak.*



1.23. ábra: Az $\mathbf{b}$ vektor és az $\mathbf{e}$ egységvektor egyenesére eső vetülete. A felső ábrán $\mathbf{e}\cdot\mathbf{b} > 0$, az alsón $\mathbf{e}\cdot\mathbf{b} < 0$.

Bizonyítás. Ha $\mathbf{e}$ egységvektor, azaz abszolút értéke 1, akkor $\mathbf{e}\cdot\mathbf{b} = |\mathbf{b}|\cos(\mathbf{e},\mathbf{b})_\angle$, ez pedig a koszinusz függvény definíciója szerint $\mathbf{b}$ merőleges vetületének előjeles hosszát jelenti. E szám $\mathbf{e}$-szerese pedig egy $\mathbf{e}$ irányú, és ilyen hosszú vektort ad, mely épp $\mathbf{b}$ vetületi vektora. □

Jelölje a továbbiakban a $\mathbf{b}$ vektornak az $\mathbf{a}$ egyenesére eső merőleges vetületi vektorát $\operatorname{proj}_\mathbf{a}\mathbf{b}$. Eszerint ha $\mathbf{e}$ egységvektor, akkor

$$\operatorname{proj}_\mathbf{e}\mathbf{b} = (\mathbf{e}\cdot\mathbf{b})\mathbf{e}.$$

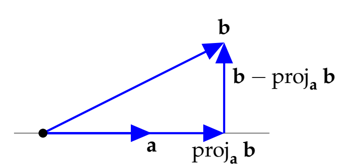
Alapvető feladat egy vektornak egy másikkal párhuzamos és rá merőleges vektorok összegére való felbontása, amit másként *merőleges összetevőkre bontásnak* nevezünk.

1.24. tétel (Vektor felbontása merőleges összetevőkre). *Ha $\mathbf{a}$ és $\mathbf{b}$ a sík vagy a tér két vektora, és $\mathbf{a} \neq \mathbf{0}$, akkor $\mathbf{b}$-nek az $\mathbf{a}$ egyenesére eső merőleges vetülete*

$$\operatorname{proj}_\mathbf{a}\mathbf{b} = \frac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}.$$

A $\mathbf{b}$-nek az $\mathbf{a}$ egyenesére merőleges összetevője

$$\mathbf{b} - \operatorname{proj}_\mathbf{a}\mathbf{b} = \mathbf{b} - \frac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}.$$



1.24. ábra: Az $\mathbf{b}$ vektor felbontása az $\mathbf{a}$ vektorral párhuzamos és rá merőleges vektorok összegére.

BIZONYÍTÁS. Az első képlet az egységvektorral szorzás geometriai jelentéséről szóló 1.23. tételből következik. Legyen $\mathbf{e} = \frac{\mathbf{a}}{|\mathbf{a}|}$ az $\mathbf{a}$-irányú egységvektor. Ekkor

$$\operatorname{proj}_{\mathbf{e}} \mathbf{b} = (\mathbf{e} \cdot \mathbf{b})\mathbf{e} = \left(\frac{\mathbf{a}}{|\mathbf{a}|} \cdot \mathbf{b} \right) \frac{\mathbf{a}}{|\mathbf{a}|} = \frac{1}{|\mathbf{a}|^2} (\mathbf{a} \cdot \mathbf{b})\mathbf{a} = \frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{a} \cdot \mathbf{a}} \mathbf{a}.$$

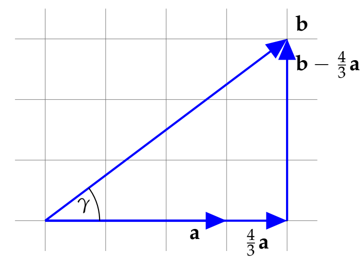
(Az utolsó egyenlőségnél kihasználtuk, hogy $|\mathbf{a}|^2 = \mathbf{a} \cdot \mathbf{a}$.) Mivel $\mathbf{e}$ és $\mathbf{a}$ párhuzamosak, ezért $\operatorname{proj}_{\mathbf{a}} \mathbf{b} = \operatorname{proj}_{\mathbf{e}} \mathbf{b}$, ami bizonyítja első állításunkat. Az állítás második fele abból következik, hogy a két összetevő összege $\mathbf{b}$. □

1.25. PÉLDA (MERŐLEGES ÖSSZETEVŐKRE BONTÁS). *Az 1.22 ábrabeli* $\mathbf{b}$ *vektort bontsuk fel az* $\mathbf{a}$*-val párhuzamos és rá merőleges összetevőkre.*

MEGOLDÁS. Bár a megoldás az 1.25 ábráról is azonnal látszik, kövessük végig a számítást: az 1.20. példa szerint $\mathbf{a} \cdot \mathbf{b} = 12$, $|\mathbf{a}| = 3$, ezért

$$\operatorname{proj}_{\mathbf{a}} \mathbf{b} = \frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{a} \cdot \mathbf{a}} \mathbf{a} = \frac{12}{9} \mathbf{a} = \frac{4}{3} \mathbf{a},$$

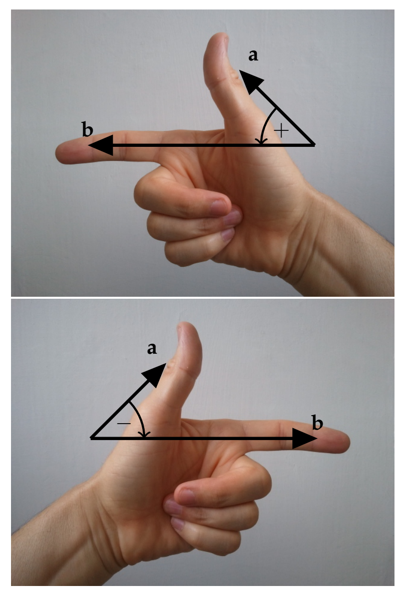
míg az $\mathbf{a}$-ra merőleges összetevő $\mathbf{b} - \frac{4}{3}\mathbf{a}$. □



1.25. ábra: Vektor felbontása merőleges összetevőkre

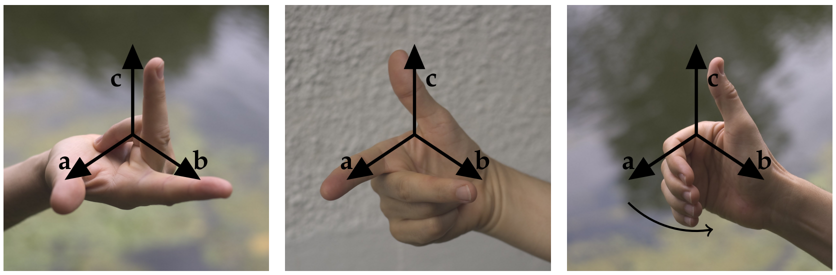
Merőlegesség és orientáció Legyenek $\mathbf{a}$ és $\mathbf{b}$ egymásra merőleges nemzérus vektorok a síkban. Ekkor $\mathbf{a}$ és $-\mathbf{b}$ is merőlegesek, azaz $(\mathbf{a}, \mathbf{b})_\angle = (\mathbf{a}, -\mathbf{b})_\angle = \pi/2$. Csak az $\mathbf{a}$ ismeretében meg tudjuk-e különböztetni a $\mathbf{b}$ és $-\mathbf{b}$ vektorokat? Hasonló kérdés a térben is fölmerül: ha $\mathbf{c}$ merőleges a nem kollineáris $\mathbf{a}$ és $\mathbf{b}$ vektorok mindegyikére, akkor $-\mathbf{c}$ is. Vajon $\mathbf{c}$ és $-\mathbf{c}$ megkülönböztethető-e egymástól csak $\mathbf{a}$-hoz és $\mathbf{b}$-hez való viszonyukat tekintve?

A válaszhoz az *irányítás*, más szóval *orientáció* fogalma vezet. E fogalmat precízen később definiáljuk (ld. ???. szakasz), az alapgondolat viszont egyszerű. A síkban a két független vektorból álló párokat két osztályba sorolhatjuk aszerint, hogy a tenyérrel fölfelé fordított jobb vagy bal kezünk első két ujjával szemléltethetőek (1.26. ábra) (hüvelyk az első, mutató a második vektor).

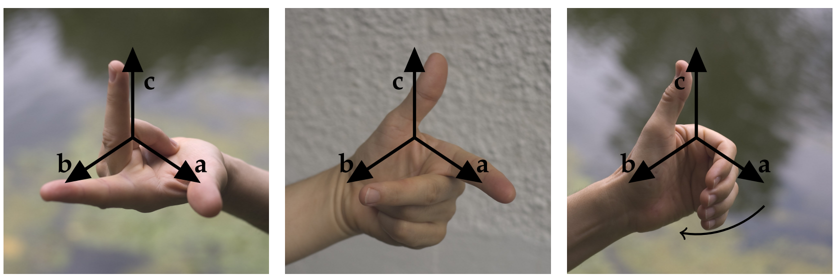


1.26. ábra: Két vektor egymáshoz való viszonya jobbrendszert (felső ábra) vagy balrendszert (alsó ábra) alkot. A közbe zárt irányított szög az előbbi esetben pozitív, utóbbiban negatív.

Hasonlóképp: a térben a független vektorokból álló hármasokat két osztályba sorolhatjuk aszerint, hogy jobb vagy bal kezünk első három ujjával szemléltethetőek (1.27. és 1.28. ábra). Kézenfekvőnek tűnik az első vektornak a hüvelyk, a másodiknak a mutató, a harmadiknak a középső ujjunkat megfeleltetni, de azokban a kultúrákban, ahol a mutató és középső ujjal mutatják a kettőt, ott a mutató-középső-hüvelyk a sorrend. Aszerint, hogy egy vektorpár a síkban, illetve egy vektorhármas a térben melyik osztályba esik, azt mondjuk, hogy *jobbrendszert*, illetve *balrendszert* alkot. Az 1.27. ábra harmadik képén látható mód (az ökölbe szoruló jobb kéz mozgása) azt is megmutatja, hogy milyen egy egyenes körül való pozitív forgás iránya. A síkban ezt azzal is ki tudjuk fejezni, hogy két független vektor szögét előjellel látjuk el, nevezetesen pozitívval, ha jobbrendszert, és negatívval, ha balrendszert alkotnak. Az így kapott szöget a két vektor *irányított szögének* nevezzük. Az $\mathbf{a}$ és $\mathbf{b}$ irányított szögét $(\mathbf{a},\mathbf{b})_\sphericalangle$ jelöli. Tehát míg $(\mathbf{a},\mathbf{b})_\angle = (\mathbf{b},\mathbf{a})_\angle$, addig $(\mathbf{a},\mathbf{b})_\sphericalangle = -(\mathbf{b},\mathbf{a})_\sphericalangle$, és ha $(\mathbf{a},\mathbf{b})_\sphericalangle = \pi/2$, akkor $(\mathbf{a},-\mathbf{b})_\sphericalangle = -\pi/2$. Ez tehát a válasz a paragrafus elején feltett kérdésre.



1.27. ábra: Az **a**, **b** és **c** vektorok ebben a sorrendben jobbrendszert alkotnak, ha irányuk a jobb kezünkkel mutatható a mellékelt három ábra bármelyike szerint: (1) hüvelyk–mutató–középső ujj, (2) mutató–középső–hüvelykujj, (3) a hüvelyk mutatja a **c** vektort, ökölbe szoruló kezünk ujjai pedig az **a** felől a **b** felé haladnak.



1.28. ábra: Az **a**, **b** és **c** vektorok ebben a sorrendben balrendszert alkotnak, ha irányuk a bal kezünkkel mutatható a következők bármelyike szerint: (1) hüvelyk–mutató–középső ujj, (2) mutató–középső–hüvelykujj, (3) a hüvelyk mutatja a **c** vektort, ökölbe szoruló kezünk ujjai pedig az **a** felől a **b** felé haladnak.

Vektori szorzás A fizikában több olyan jelenség is van, melyben két térbeli vektorhoz keresünk egy mindkettőre merőleges harmadikat. Legismertebb példa a *forgatónyomaték*.

Hasson egy $\mathbf{F}$ erő egy test $\mathbf{P}$ pontjában, és legyen a test rögzítve az O pontjában. A P ponton átmenő, $\mathbf{F}$ irányú egyenesnek az O-tól való távolságát az erő karjának nevezzük. Az $\mathbf{F}$ hatására a test O körül elfordul. Ennek jellemzésére tudnunk kell a forgás tengelyét, a forgás „nagyságát”, és azt, hogy a tengely körüli két forgásirány közül melyikről van szó. Erre alkalmas lehet egy vektor – ezt nevezzük *forgatónyomatéknak* –, melynek iránya a forgástengellyel párhuzamos, hossza a forgás nagyságát írja le, és a forgástengellyel párhuzamos két vektorirány a két forgásirányhoz tartozik. Hogyan definiálható a forgatónyomaték-vektor, ha tudjuk, hogy abszolút értéke az erőkar hosszának és az erő abszolút értékének szorzata?

Az erő karja $|\overrightarrow{OP}|\sin(\overrightarrow{OP},\mathbf{F})_\angle$, így az $\mathbf{M}$ forgatónyomaték abszolút értéke:

$$|\mathbf{M}| = |\mathbf{F}||\overrightarrow{OP}|\sin(\overrightarrow{OP},\mathbf{F})_\angle.$$

A forgás tengelye nyilván merőleges **F**-re és $\overrightarrow{OP}$-re is, csak abban kell megegyezni, hogy az $\overrightarrow{OP}$, **F** és **M** vektorok jobb- vagy balrendszert alkossanak. A fizikusok a jobbrendszert választották.

A forgatónyomaték, és több hasonló fizikai fogalom a következő definícióhoz vezet:

1.26. DEFINÍCIÓ (VEKTORI SZORZAT). *A 3-dimenziós tér két vektorának* vektori szorzatán *azt a vektort értjük, melynek*

a) abszolút értéke *a két vektor abszolút értékének és a közbezárt szög szinuszának szorzata,*

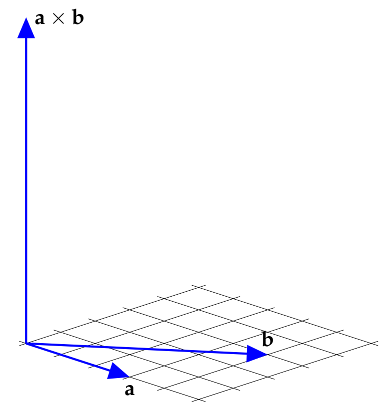
b) iránya *merőleges mindkét vektor irányára és – ha a két tényező és a szorzat egyike sem a nullvektor, akkor – az első tényező, a második tényező és a szorzat ebben a sorrendben jobbrendszert alkot.*

► E definíció bármely két 3-dimenziós vektor vektori szorzatát egyértelműen definiálja, ugyanis minden olyan esetben, amikor nem tudnánk eldönteni, hogy a vektorok jobbrendszert alkotnak-e, a szorzat a nullvektor (gondoljuk meg!).

► Az **a** és **b** vektorok vektori szorzatát $\mathbf{a} \times \mathbf{b}$ jelöli, amit „a kereszt b”-nek olvasunk. Képletekkel megfogalmazva: $\mathbf{a} \times \mathbf{b}$ egy vektor, melyre

$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}||\mathbf{b}| \sin(\mathbf{a}, \mathbf{b})_\angle,$$

$\mathbf{a} \times \mathbf{b} \perp \mathbf{a}$, $\mathbf{a} \times \mathbf{b} \perp \mathbf{b}$, továbbá **a**, **b** és $\mathbf{a} \times \mathbf{b}$ ebben a sorrendben jobbrendszert alkot.



1.29. ábra: Az **a**, **b** és az $\mathbf{a} \times \mathbf{b}$ vektorok

1.27. PÉLDA (VEKTORI SZORZAT MEGHATÁROZÁSA). *Tegyük fel, hogy a tér két vektora 3 illetve 5 hosszú, az általuk bezárt szög koszinusza $\frac{4}{5}$, mint az 1.20. példában. Mit tudunk a vektori szorzatról?*

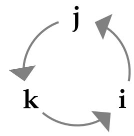
MEGOLDÁS. Ha $\cos\gamma = \frac{4}{5}$, akkor $\sin\gamma = \sqrt{1 - (\frac{4}{5})^2} = \frac{3}{5}$, így a vektori szorzat hossza $|\mathbf{a}||\mathbf{b}| \sin(\mathbf{a}, \mathbf{b})_\angle = 3 \cdot 5 \cdot \frac{3}{5} = 9$, iránya merőleges mindkét vektorra és **a**, **b**, $\mathbf{a} \times \mathbf{b}$ ebben a sorrendben jobbrendszert alkot (ld. 1.29. ábra). □

1.28. PÉLDA (**i**, **j**, **k** VEKTORI SZORZATA). *Legyen* **i**, **j**, **k** *három, páronként egymásra merőleges, ebben a sorrendben jobbrendszert alkotó egységvektor. Készítsünk művelettáblát vektori szorzataikról!*

MEGOLDÁS. Mivel $(\mathbf{i}, \mathbf{i})_\angle = 0$, ezért $|\mathbf{i} \times \mathbf{i}| = 0$, így $\mathbf{i} \times \mathbf{i} = \mathbf{0}$. Hasonlóan $\mathbf{j} \times \mathbf{j} = \mathbf{0}$ és $\mathbf{k} \times \mathbf{k} = \mathbf{0}$.

Mivel $|\mathbf{i}| = |\mathbf{j}| = 1$ és $(\mathbf{i}, \mathbf{j})_\angle = 90°$, ezért $|\mathbf{i} \times \mathbf{j}| = 1$, azaz $\mathbf{i} \times \mathbf{j}$ is egységvektor. Ráadásul $\mathbf{i} \times \mathbf{j}$ merőleges **i**-re és **j**-re, és **i**, **j** valamint $\mathbf{i} \times \mathbf{j}$ jobbrendszert alkotnak épp úgy, mint **i**, **j** és **k**. Ebből következik, hogy $\mathbf{i} \times \mathbf{j} = \mathbf{k}$. Hasonlóképp $\mathbf{j} \times \mathbf{k} = \mathbf{i}$ és $\mathbf{k} \times \mathbf{i} = \mathbf{j}$. Ha **i**, **j**, **k** jobbrendszert alkot, akkor **j**, **i** és **k** balrendszert, így $\mathbf{j} \times \mathbf{i} = -\mathbf{k}$. Mindezeket összefoglalva a következő művelettáblát kapjuk.

$\times$	$\mathbf{i}$	$\mathbf{j}$	$\mathbf{k}$
$\mathbf{i}$	$\mathbf{0}$	$\mathbf{k}$	$-\mathbf{j}$
$\mathbf{j}$	$-\mathbf{k}$	$\mathbf{0}$	$\mathbf{i}$
$\mathbf{k}$	$\mathbf{j}$	$-\mathbf{i}$	$\mathbf{0}$



E három vektor közti szorzatok könnyen megjegyezhetőek, ha egy szabályos háromszög csúcsaira írjuk őket pozitív körüljárás szerint, mint azt a táblázat melletti ábra mutatja. Ekkor két különböző vektor szorzata a harmadik, ha a két vektor pozitív körüljárás szerint követi egymást. Ha negatív körüljárás szerint követik egymást, a szorzat a harmadik vektor -1-szerese. □

1.29. TÉTEL (MIKOR $\mathbf{0}$ A VEKTORI SZORZAT?). *Két térbeli vektor vektori szorzata pontosan akkor zérusvektor, ha a két vektor párhuzamos.*

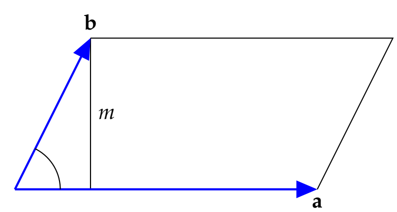
BIZONYÍTÁS. Ha $\mathbf{a}$ vagy $\mathbf{b}$ valamelyike zérusvektor, akkor egyrészt a két vektor tekinthető párhuzamosnak, másrészt $\mathbf{a} \times \mathbf{b} = \mathbf{0}$, az állítás tehát igaz, ezért a továbbiakban feltesszük, hogy a két tényező egyike sem zérusvektor.

($\Longleftarrow$) Ha $\mathbf{a}$ és $\mathbf{b}$ párhuzamosak (más szóval kollineárisak), akkor $(\mathbf{a}, \mathbf{b})_\angle = 0$ vagy π, tehát $\sin(\mathbf{a}, \mathbf{b})_\angle = 0$, így $|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}||\mathbf{b}|0 = 0$, azaz $\mathbf{a} \times \mathbf{b} = \mathbf{0}$.

($\Longrightarrow$) Ha $\mathbf{a} \times \mathbf{b} = \mathbf{0}$, azaz $|\mathbf{a}||\mathbf{b}| \sin(\mathbf{a}, \mathbf{b})_\angle = 0$, akkor $|\mathbf{a}| = 0$, $|\mathbf{b}| = 0$ vagy $\sin(\mathbf{a}, \mathbf{b})_\angle = 0$. Az $|\mathbf{a}| = 0$ vagy $|\mathbf{b}| = 0$ eseteket már elintéztük a bizonyítás elején, így marad a $\sin(\mathbf{a}, \mathbf{b})_\angle = 0$ eset. A szinusz függvénynek a $[0, \pi]$ intervallumban a 0 és a π helyen van zérushelye, tehát a két vektor vagy egyirányú, vagy ellenkező irányú, vagyis párhuzamos. □

1.30. TÉTEL (VEKTORI SZORZAT ABSZOLÚT ÉRTÉKÉNEK GEOMETRIAI JELENTÉSE). *Két vektor vektori szorzatának abszolút értéke a két vektor által kifeszített parallelogramma területének mérőszámával egyenlő.*

BIZONYÍTÁS. Az $\mathbf{a}$ és $\mathbf{b}$ vektorok által kifeszített parallelogramma oldalainak hossza $|\mathbf{a}|$ és $|\mathbf{b}|$, az $\mathbf{a}$ oldalhoz tartozó magassága pedig $m = |\mathbf{b}| \sin(\mathbf{a}, \mathbf{b})_\angle$. A parallelogramma területe $|\mathbf{a}|m = |\mathbf{a}||\mathbf{b}| \sin(\mathbf{a}, \mathbf{b})_\angle = |\mathbf{a} \times \mathbf{b}|$ (1.30. ábra). □



1.30. ábra: $|\mathbf{a} \times \mathbf{b}|$ megegyezik a parallelogramma területével

1.31. TÉTEL (VEKTORI SZORZÁS MŰVELETI TULAJDONSÁGAI). *Tetszőleges* $\mathbf{a}$, $\mathbf{b}$ *és* $\mathbf{c}$ *vektorokra, valamint tetszőleges* r *valós számra igazak az alábbi összefüggések:*

a) $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$ *(alternáló tulajdonság)*

b) $(\mathbf{a} + \mathbf{b}) \times \mathbf{c} = \mathbf{a} \times \mathbf{c} + \mathbf{b} \times \mathbf{c}$
$\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$ *(disztributivitás)*

c) $r(\mathbf{a} \times \mathbf{b}) = (r\mathbf{a}) \times \mathbf{b} = \mathbf{a} \times (r\mathbf{b})$

d) $|\mathbf{a} \times \mathbf{b}| = \sqrt{|\mathbf{a}|^2|\mathbf{b}|^2 - |\mathbf{a} \cdot \mathbf{b}|^2}$

A tétel bizonyítását feladatnak tűzzük ki.

► E tétel *a)* pontja szerint a vektori szorzás *nem kommutatív*!
► Egy egyszerű példa mutatja, hogy a vektori szorzás nem is asszociatív. Az 1.28. példa eredményét használva könnyen látható, hogy

$$(\mathbf{i} \times \mathbf{j}) \times \mathbf{j} \neq \mathbf{i} \times (\mathbf{j} \times \mathbf{j}),$$

ugyanis $(\mathbf{i} \times \mathbf{j}) \times \mathbf{j} = \mathbf{k} \times \mathbf{j} = -\mathbf{i}$, másrészt $\mathbf{i} \times (\mathbf{j} \times \mathbf{j}) = \mathbf{i} \times \mathbf{0} = \mathbf{0}$.

Parallelepipedon térfogata, és előjeles térfogata Az 1.30. tételben megmutattuk, hogy a vektori szorzat abszolút értéke a két vektor által kifeszített parallelogramma területét adja. Hogy számítható ki a parallelepipedon térfogata?

1.32. PÉLDA (PARALLELEPIPEDON TÉRFOGATA). *Határozzuk meg az* $\mathbf{a}$, $\mathbf{b}$ *és* $\mathbf{c}$ *vektorok által kifeszített parallelepipedon térfogatát!*

MEGOLDÁS. Az $\mathbf{a}$ és $\mathbf{b}$ által kifeszített parallelogramma területe $|\mathbf{a} \times \mathbf{b}|$, és mivel $\mathbf{a} \times \mathbf{b}$ merőleges a parallelogramma síkjára, ezért a parallelepipedon magassága $\mathbf{c}$-nek az $\mathbf{a} \times \mathbf{b}$ egyenesére eső merőleges vetületi hosszával egyenlő. Ez az $\mathbf{a} \times \mathbf{b}$ irányú egységvektorral való skaláris szorzással számolható. Az egységvektor

$$\mathbf{e} = \frac{\mathbf{a} \times \mathbf{b}}{|\mathbf{a} \times \mathbf{b}|},$$

a magasság $|\mathbf{e} \cdot \mathbf{c}|$, és így a térfogat (azaz az alapterületszer magasság) értéke

$$|\mathbf{a} \times \mathbf{b}| \left| \frac{\mathbf{a} \times \mathbf{b}}{|\mathbf{a} \times \mathbf{b}|} \cdot \mathbf{c} \right| = |(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}| \,.$$

Tehát a parallelepipedon térfogata $|(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}|$. □

A térfogat tehát az $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$ skalár abszolút értéke. Kérdés: vajon mit jelent e skalár előjele? Ez pontosan akkor negatív, ha a $\mathbf{c}$ vektor $\mathbf{a} \times \mathbf{b}$ egyenesére eső merőleges vetülete és $\mathbf{a} \times \mathbf{b}$ ellenkező irányú. Vagyis ha a $\mathbf{c}$ vektor az $\mathbf{a}$ és $\mathbf{b}$ síkjának másik oldalán van, mint az $\mathbf{a} \times \mathbf{b}$ vektor, azaz ha $\mathbf{a}$, $\mathbf{b}$ és $\mathbf{c}$ balrendszert alkot!

Összefoglalva: Az $\mathbf{a}$, $\mathbf{b}$ és $\mathbf{c}$ vektorokból képzett $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$ skalár abszolút értéke a három vektor által kifeszített parallelepipedon térfogatával egyenlő, míg előjele a három vektor orientációját adja, nevezetesen pontosan akkor pozitív, ha a három vektor jobbrendszert alkot. Az is következik a fentiekből, hogy a három vektor pontosan akkor esik egy síkba, azaz pontosan akkor lineárisan összefüggők, ha $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = 0$.

Vegyes szorzat Az előző paragrafusban megmutattuk az $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$ kifejezés fontosságát. Ez vezet a következő definícióhoz.

1.33. DEFINÍCIÓ (VEGYES SZORZAT). *A 3-dimenziós tér három tetszőleges* $\mathbf{a}$, $\mathbf{b}$ *és* $\mathbf{c}$ *vektorából képzett*

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$$

kifejezést a három vektor vegyes szorzatának *nevezzük.*

- ▶ A vegyes szorzat eredménye skalár.
- ▶ Az $\mathbf{a}$, $\mathbf{b}$ és $\mathbf{c}$ vektorok vegyes szorzatának szokás jelölése $\mathbf{abc}$, de mi a későbbi fejezetekben nem fogjuk használni.
- ▶ Mivel a skaláris szorzás kommutatív, ezért $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$.
- ▶ A parallelepipedon térfogatára ugyanazt az értéket kell kapnunk, bármelyik oldallapot is választjuk alapnak, így e három vektorból a vektorok különböző sorrendjeivel képzett vegyes szorzatok csak előjelében térhetnek el egymástól. Mivel az előjel az orientáció függvénye, ezért – figyelembe véve az előző megjegyzést is – kapjuk, hogy

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = (\mathbf{b} \times \mathbf{c}) \cdot \mathbf{a} = (\mathbf{c} \times \mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}).$$

Az ellenkező előjelű szorzatok:

$$(\mathbf{c} \times \mathbf{b}) \cdot \mathbf{a} = (\mathbf{b} \times \mathbf{a}) \cdot \mathbf{c} = (\mathbf{a} \times \mathbf{c}) \cdot \mathbf{b} = \mathbf{c} \cdot (\mathbf{b} \times \mathbf{a}) = \mathbf{b} \cdot (\mathbf{a} \times \mathbf{c}) = \mathbf{a} \cdot (\mathbf{c} \times \mathbf{b}).$$

Ezeket a vegyes szorzatra használt jelöléssel fölírva:

$$\mathbf{abc} = \mathbf{bca} = \mathbf{cab} = -\mathbf{acb} = -\mathbf{cba} = -\mathbf{bac}.$$

1.34. PÉLDA (VEGYES SZORZAT). *Határozzuk meg egy egységélű kocka egy csúcsból induló három lapátló-vektorának vegyes szorzatát!*

MEGOLDÁS. Jelölje a kocka egyik csúcsából induló három élvektorát $\mathbf{i}$, $\mathbf{j}$ és $\mathbf{k}$. E három vektor ebben a sorrendben alkosson jobbrendszert. Ekkor az előző megjegyzés szerint $\mathbf{ijk} = \mathbf{jki} = \mathbf{kij} = 1$, $\mathbf{kji} = \mathbf{jik} = \mathbf{ikj} = -1$. Mivel a vegyes szorzat egy parallelepipedon térfogatát, vagy annak ellentettjét adja, ezért ha egy szorzatban egy vektor többször is szerepel, akkor annak értéke 0. Például $\mathbf{iji} = (\mathbf{i} \times \mathbf{j}) \cdot \mathbf{i} = \mathbf{k} \cdot \mathbf{i} = 0$. A három lapátló-vektor: $\mathbf{i} + \mathbf{j}$, $\mathbf{j} + \mathbf{k}$, $\mathbf{k} + \mathbf{i}$. Ezek vegyes szorzata

$$\begin{aligned}((\mathbf{i} + \mathbf{j}) \times (\mathbf{j} + \mathbf{k})) \cdot (\mathbf{k} + \mathbf{i}) &= \mathbf{ijk} + \mathbf{iji} + \mathbf{ikk} + \mathbf{iki} + \mathbf{jjk} + \mathbf{jji} + \mathbf{jkk} + \mathbf{jki} \\ &= 1 + 0 + 0 + 0 + 0 + 0 + 0 + 1 \\ &= 2,\end{aligned}$$

tehát a három lapátló-vektor vegyes szorzata 2. □

Feladatok

Skaláris szorzás

1.11. Mutassuk meg, hogy $|\mathbf{a}+\mathbf{b}+\mathbf{c}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2 + |\mathbf{c}|^2$ pontosan akkor áll fenn, ha **a**, **b** és **c** három egymásra páronként merőleges vektor.

1.12. Igazoljuk, hogy

$$(\mathbf{a}\cdot\mathbf{b})\mathbf{c} \neq \mathbf{a}(\mathbf{b}\cdot\mathbf{c}).$$

Merőlegesség és orientáció: vektori szorzás Vektori szorzat

1.13. Tekintsünk egy egységélű kockát, melynek egyik csúcsát jelölje P. Számítsuk ki a P-ből induló

a) valamelyik két lapátló-vektor,

b) egyik lapátló- és a testátló-vektor

skaláris szorzatát, valamint a P-ből induló

c) valamelyik élvektor és egy vele egy lapon lévő lapátló-vektor,

d) valamelyik élvektor és a vele nem egy lapon lévő lapátló-vektor

vektori szorzatát.

1.14. Mutassuk meg, hogy ha **u** merőleges a **v** és **w** vektorokra, akkor merőleges minden lineáris kombinációjukra is.

1.15. Három lineárisan független vektornak hány sorrendje van? E sorrendek közül hány alkot jobb- és hány balrendszert?

1.16. [Szögfelező] Legyenek **a** és **b** nemzérus vektorok. Mutassuk meg, hogy a $|\mathbf{b}|\mathbf{a} + |\mathbf{a}|\mathbf{b}$ vektor felezi **a** és **b** szögét!

1.17. [Háromszög szögfelezője] Az előző feladat eredményét felhasználva mutassuk meg, hogy a háromszög egyik szögének szögfelezője a szemközti oldalt a két szomszédos oldal hosszának arányában osztja fel.

1.18. [Mit cserél föl a tükör?] Hogy lehet az, hogy a tükör fölcseréli a jobbat a ballal, de nem cseréli föl a föntet a lenttel?

Projekt: ekvivalencia reláció

Egy X halmazon értelmezett *reláción* (pontosabban *bináris reláción*) az X elempárjainak egy R halmazát értjük. Ha egy (a, b) pár benne van ebben a halmazban, azt mondjuk, hogy a az R relációban van b-vel, és úgy jelöljük, hogy $a\ R\ b$. Például, ha X az összes valaha élt ember halmaza, akkor az összes olyan (a, b) emberpár halmaza, ahol a anyja b-nek, egy reláció. Ez a köznyelvből is ismert anyagyermek reláció. Egy matematikai példa: ha X a valósok halmaza, és R azokból az (a, b) párokból áll, melyekre a kisebb vagy egyenlő, mint b, akkor R egy reláció, melyet a valósok rendezési relációjának nevezünk. E reláció szokásos jele $\leq$, így ha $(a, b) \in R$, akkor az $a \leq b$ jelölést használjuk.

Az X halmazon értelmezett *ekvivalencia reláció* egy olyan reláció, mely az X halmaz elemeinek egy osztályozását írja le. Ez azt jelenti, hogy ha az X halmazt páronként kizáró részhalmazok uniójára bontjuk, azaz megadjuk az X elemeinek egy *osztályozását*, vagy más szóval X egy *particionálását*, akkor létezik egy reláció, melyben az $a, b \in X$ elemek pontosan akkor vannak relációban, ha egy osztályba (partícióba) tartoznak. Az így kapott relációnak van három olyan tulajdonsága, melynek csak az ilyen módon megkapható relációk felelnek meg.

1.35. TÉTEL (EKVIVALENCIA RELÁCIÓ). ....

Matematikán kívüli hasonlattal élve: ha az irányított szakasz a hal, a szabad vektor a halraj, melyet mindegyik hal ugyanúgy reprezentál (**??**. ábra).

1.19. [Szabad vektor fogalma] Blabla

1.20. [Vektor iránya] Blabla

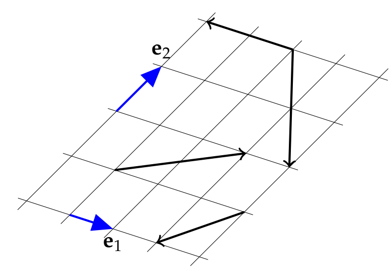
Vektorok koordinátás alakban

A koordináták bevezetésével egyrészt új algebrai eszközökhöz jutunk a vektorok és a különféle geometriai alakzatok vizsgálatában, másrészt lehetővé válik a vektor fogalmának kiterjesztése. Így jutunk a sokdimenziós terek fogalmához, ami nélkülözhetetlen a közgazdaságtanban, az internetes keresők matematikájában, vagy véges struktúrák fölötti változatában a kódelméletben és a kriptográfiában.

René Descartes (Renatus Cartesianus) (1596–1650) francia filozófus és matematikus, a modern filozófia atyja, az analitikus geometria egyik megalkotója. Filozófiáját a puszta hitre alapozott állításokkal szemben a racionális érvelések útján kívánta fölépíteni (lásd *descartesi kételkedés* és „gondolkodom, tehát vagyok"). Orvostudományt és jogot tanult, végül hadmérnöki képesítést szerzett. Több háborúban is részt vett. 1619-ben egy Magyarországot is érintő hosszú útján egy Ulm melletti parasztházban három álmot látott, melyek megfejtése „egy csodálatos tudományhoz" vezette, ami filozófiája alapjává vált.

Descartes-féle koordinátarendszer Descartes 1637-ben *La Géométrie* című művében egy szép ötlettel összekapcsolta a geometriát az algebrával. Alapgondolata az volt, hogy a geometria alapelemei (pl. pontok) és a valós számok/számpárok/számhármasok közt kölcsönösen egyértelmű megfeleltetés hozható létre, így bizonyos geometriai alakzatok algebrai egyenletekkel leírhatóvá és vizsgálhatóvá válnak.

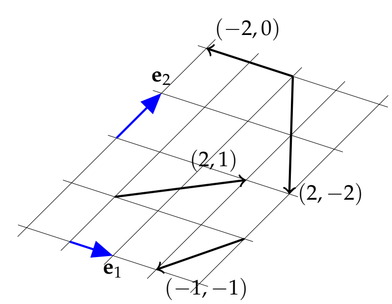
Az 1.11. tétel szerint a sík bármely $\mathbf{v}$ vektora felírható két adott lineárisan független $\mathbf{e}_1$ és $\mathbf{e}_2$ vektor lineáris kombinációjaként, és e felírás egyértelmű. Ha e lineáris kombináció $\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2$ alakú, akkor a $\mathbf{v}$ vektorhoz a (v_1, v_2) számpárt rendeljük, amit a $\mathbf{v}$ vektor *koordinátás alakjának* nevezünk, a v_1 és v_2 skalárokat pedig a $\mathbf{v}$ *koordinátáinak* nevezzük. Azt mondjuk, hogy az $\{\mathbf{e}_1, \mathbf{e}_2\}$ vektorpár e koordinátázási rendszer – egyszerűbben *koordinátarendszer* – *bázisa*, az $\mathbf{e}_1$ és $\mathbf{e}_2$ vektorok a *bázisvektorok* vagy *alapvektorok*. Tetszőleges vektor koordinátáinak meghatározásához elég a bázisvektorokat ismerni.



1.32. ábra: Mik a vektorok koordinátái?

1.36. PÉLDA (VEKTOROK KOORDINÁTÁI). *Határozzuk meg az 1.32. ábrán megadott vektoroknak – az $\mathbf{e}_1$ és $\mathbf{e}_2$ vektorokra, mint bázisra vonatkozó – koordinátáit!*

MEGOLDÁS. A megoldás könnyen leolvasható az 1.33. ábráról. Áttekinthetőbb, ha az összes vektort egyetlen pontból indítjuk. Ezt mutatja az 1.34. ábra. □



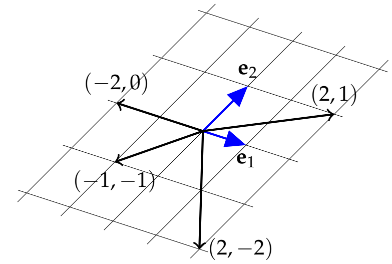
1.33. ábra: A megoldás

A koordinátarendszer a 3-dimenziós térben is hasonló módon építhető fel. Az 1.12. tétel szerint a tér bármely $\mathbf{v}$ vektora felírható három adott lineárisan független $\mathbf{e}_1$, $\mathbf{e}_2$ és $\mathbf{e}_3$ vektor lineáris kombinációjaként, és e felírás egyértelmű. Ha e lineáris kombináció $\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3$ alakú, akkor a $\mathbf{v}$ vektorhoz a (v_1, v_2, v_3) számhármast rendeljük, amit a $\mathbf{v}$ vektor *koordinátás alakjának* nevezünk, a v_1, v_2 és v_3 skalárokat pedig a $\mathbf{v}$ *koordinátáinak* nevezzük, a *bázis* pedig az $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ vektorhármas.

Sőt, a koordinátázás az 1-dimenziós térben is megvalósítható: ha $\mathbf{e}$ egy nemnulla vektor (tehát lineárisan független vektorrendszer!), akkor bármely vele párhuzamos $\mathbf{v}$ vektor egyértelműen felírható $\mathbf{v} = v\mathbf{e}$

alakban. E v skalár lesz a $\mathbf{v}$ koordinátás alakja (a zárójel használata itt szükségtelen). Így a $\mathbf{v} \leftrightarrow v$ hozzárendelés kölcsönösen egyértelmű megfeleltetést létesít a vektorok és a skalárok közt.

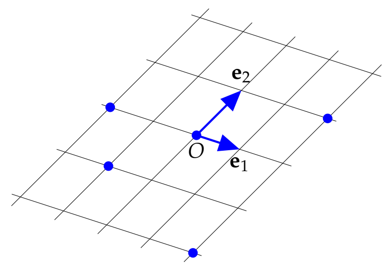
Ha kijelölünk egy pontot az egyenesen/síkban/térben – ez lesz az origó –, akkor az egyenes/sík/tér pontjai és a helyvektorok végpontjai közti kölcsönösen egyértelmű megfeleltetéssel egyúttal a pontok is koordinátát kapnak.



1.34. ábra: A megoldás helyvektorokkal ábrázolva.

1.37. PÉLDA (PONTOK KOORDINÁTÁI). *Határozzuk meg az 1.35. ábrán kijelölt pontok – megadott bázisvektorokra és origóra vonatkozó – koordinátáit!*

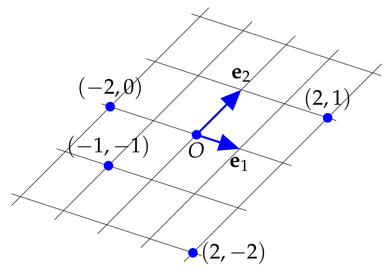
MEGOLDÁS. E feladat megoldása lényegében azonos az előzőével, minthogy a kijelölt pontokba mutató helyvektorok megegyeznek az ott megadott vektorokkal (ld. 1.36. ábra). □



1.35. ábra: Mik a pontok koordinátái?

A helyvektorok és a pontok közti kölcsönösen egyértelmű megfeleltetést a jelölésben is kifejezzük azzal, hogy nem teszünk különbséget a vektor és a pont koordinátás alakja közt, azaz a $\mathbf{v} = (a, b)$ vektorhoz adott origó mellett rendelt pontot is (a, b) jelöli. Ha a pontnak nevet is adunk, pl. e pont a P pont, akkor a $P(a, b)$ jelölés a nevet és a koordinátákat is megadja. Ekkor a helyvektort $\overrightarrow{OP}$ is jelölheti. Tehát $\mathbf{v} = \overrightarrow{OP}$. Szokás a vektorok koordinátás alakját úgynevezett *oszlopvektor* alakba is írni, mi e könyvben ekkor kerek helyett szögletes zárójelet használunk, például:

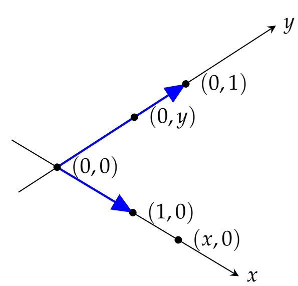
$$\overrightarrow{OP} = \begin{bmatrix} a \\ b \end{bmatrix}.$$



1.36. ábra: Pontok és koordinátáik

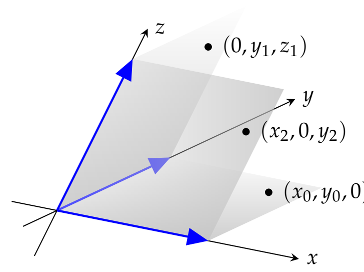
E jelölés előnyeivel hamarosan találkozunk.

Látható, hogy ha egy pont az első tengelyen van, és azon az egyenesen x az 1-dimenziós koordinátája, akkor síkbeli koordinátás alakja $(x, 0)$ lesz. Hasonlóképp a második tengely minden pontjának $(0, y)$ a koordinátás alakja. Az origóé $(0, 0)$ (ld. 1.37. ábra). Az alapvektorok koordinátás alakja $\mathbf{e}_1 = (1, 0)$ és $\mathbf{e}_2 = (0, 1)$.



1.37. ábra: Pontok a koordinátarendszer tengelyein.

Hasonlóképp a 3-dimenziós esetben a koordinátatengelyekre eső pontok 3-dimenziós koordinátás alakja $(x, 0, 0)$, $(0, y, 0)$, illetve $(0, 0, z)$ attól függően, hogy melyik tengelyről van szó. Az origón átmenő és 2 tengelyt tartalmazó síkokat *koordinátasíkoknak* nevezzük. Ezekből is három van. Könnyen látható, hogy a koordinátasíkok pontjainak alakja $(x, y, 0)$, $(x, 0, z)$, illetve $(0, y, z)$. Az origóé $(0, 0, 0)$, míg az alapvektoroké $\mathbf{e}_1 = (1, 0, 0)$, $\mathbf{e}_2 = (0, 1, 0)$, $\mathbf{e}_3 = (0, 0, 1)$ (ld. 1.38. ábra).



1.38. ábra: Pontok a koordinátasíkokon

Műveletek koordinátás alakban megadott vektorokkal Legyen adva a térben egy koordinátarendszer, és abban két tetszőleges $\mathbf{u} = (u_1, u_2, u_3)$ és $\mathbf{v} = (v_1, v_2, v_3)$ vektor. E paragrafusban megkeressük a vektorműveletek koordinátás alakját. A kérdés tehát az, hogy hogyan kapható meg $\mathbf{u} + \mathbf{v}$, $\mathbf{u} - \mathbf{v}$, $c\mathbf{u}$, $\mathbf{u} \cdot \mathbf{v}$, $\mathbf{u} \times \mathbf{v}$ koordinátás alakja.

Mindegyik műveletnél felhasználjuk a vektoroknak a bázisvektorok lineáris kombinációjaként való előállítását. Az adott két vektor összege

$$\begin{aligned}\mathbf{u}+\mathbf{v} &= (u_1,u_2,u_3)+(v_1,v_2,v_3)\\ &= (u_1\mathbf{e}_1+u_2\mathbf{e}_2+u_3\mathbf{e}_3)+(v_1\mathbf{e}_1+v_2\mathbf{e}_2+v_3\mathbf{e}_3)\\ &= (u_1+v_1)\mathbf{e}_1+(u_2+v_2)\mathbf{e}_2+(u_3+v_3)\mathbf{e}_3\\ &= (u_1+v_1,u_2+v_2,u_3+v_3).\end{aligned}$$

Hasonló képlet adódik a különbségre

$$\mathbf{u}-\mathbf{v}=(u_1,u_2,u_3)-(v_1,v_2,v_3)=(u_1-v_1,u_2-v_2,u_3-v_3).$$

A skalárral való szorzás is a koordinátánként való végrahajtás lehetőségét mutatja:

$$\begin{aligned}c\mathbf{u} &= c(u_1,u_2,u_3)=c(u_1\mathbf{e}_1+u_2\mathbf{e}_2+u_3\mathbf{e}_3)\\ &= cu_1\mathbf{e}_1+cu_2\mathbf{e}_2+cu_3\mathbf{e}_3\\ &= (cu_1,cu_2,cu_3).\end{aligned}$$

Összefoglalva tehát a következő állítást kapjuk:

1.38. ÁLLÍTÁS (VEKTORMŰVELETEK KOORDINÁTÁS ALAKJA). *Adva van a térben egy koordinátarendszer, és abban két tetszőleges* $\mathbf{u}=(u_1,u_2,u_3)$ *és* $\mathbf{v}=(v_1,v_2,v_3)$ *vektor, valamint egy tetszőleges* $c\in\mathbb{R}$ *valós szám. Ekkor a vektorok összegének, különbségének és skalárszorosának koordinátás alakja rendre*

$$\begin{aligned}\mathbf{u}+\mathbf{v} &= (u_1,u_2,u_3)+(v_1,v_2,v_3)=(u_1+v_1,u_2+v_2,u_3+v_3),\\ \mathbf{u}-\mathbf{v} &= (u_1,u_2,u_3)-(v_1,v_2,v_3)=(u_1-v_1,u_2-v_2,u_3-v_3),\\ c\mathbf{u} &= c(u_1,u_2,u_3)=(cu_1,cu_2,cu_3).\end{aligned}$$

Az oszlopvektor jelölést használva

$$\mathbf{u}\pm\mathbf{v}=\begin{bmatrix}u_1\\u_2\\u_3\end{bmatrix}\pm\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}=\begin{bmatrix}u_1\pm v_1\\u_2\pm v_2\\u_3\pm v_3\end{bmatrix},\quad c\mathbf{u}=c\begin{bmatrix}u_1\\u_2\\u_3\end{bmatrix}=\begin{bmatrix}cu_1\\cu_2\\cu_3\end{bmatrix}.$$

A síkbeli vektorokra hasonló állítások igazak, csak két koordinátával. Érdekesebb a helyzet a skaláris szorzással. Kezdjük két síkbeli vektorral.

1.39. PÉLDA (SKALÁRIS SZORZÁS KOORDINÁTARENDSZERBEN). *Tekintsünk egy olyan síkbeli koordinátarendszert, ahol az első alapvektor hossza* 1*, a másodiké* 2*, és a kettőjük közti szög* $\pi/3$*. Számítsuk ki az* $\mathbf{a}=(1,1)$ *és a* $\mathbf{b}=(-5/2,1)$ *vektorok skaláris szorzatát.*

MEGOLDÁS. Az alapvektorok hosszát és szögét ismerve ki tudjuk számítani az alapvektorok skaláris szorzatait:

$$\mathbf{e}_1 \cdot \mathbf{e}_1 = 1, \quad \mathbf{e}_2 \cdot \mathbf{e}_2 = 2^2 = 4, \quad \mathbf{e}_1 \cdot \mathbf{e}_2 = 1 \cdot 2 \cdot \cos \frac{\pi}{3} = 1.$$

Ezt fölhasználva kapjuk, hogy

$$\begin{aligned}
\mathbf{a} \cdot \mathbf{b} &= (\mathbf{e}_1 + \mathbf{e}_2) \cdot (-\frac{5}{2}\mathbf{e}_1 + \mathbf{e}_2) \\
&= -\frac{5}{2}\mathbf{e}_1 \cdot \mathbf{e}_1 + (1 - \frac{5}{2})\mathbf{e}_1 \cdot \mathbf{e}_2 + \mathbf{e}_2 \cdot \mathbf{e}_2 \\
&= -\frac{5}{2} - \frac{3}{2} + 4 \\
&= 0,
\end{aligned}$$

tehát a két vektor merőleges egymásra.

Érdekességként meghatározzuk e koordinátarendszerben a skaláris szorzás általános képletét:

$$\begin{aligned}
\mathbf{u} \cdot \mathbf{v} &= (u_1\mathbf{e}_1 + u_2\mathbf{e}_2) \cdot (v_1\mathbf{e}_1 + v_2\mathbf{e}_2) \\
&= u_1v_1\mathbf{e}_1 \cdot \mathbf{e}_1 + (u_1v_2 + u_2v_1)\mathbf{e}_1 \cdot \mathbf{e}_2 + u_2v_2\mathbf{e}_2 \cdot \mathbf{e}_2 \\
&= u_1v_1 + u_1v_2 + u_2v_1 + 4u_2v_2.
\end{aligned}$$

Ebből látható, hogy a skaláris szorzat koordinátás alakja függ a koordinátarendszertől is. □

A derékszögű koordinátarendszer A természeti törvények különös fontosságot adnak az egymásra merőleges irányoknak, ezért például igen gyakran érdemes olyan koordinátarendszert választani, amelyben az alapvektorok merőlegesek, más szóval *ortogonálisak* egymásra. A bázisvektorok szöge mellett azok hosszát is érdemes standardizálni, nevezetesen egységnyi hosszúnak választani, így mindegyik koordináta egyúttal távolságot is jelent. Az egységvektorokból álló ortogonális bázist *ortonormáltnak* nevezzük.

Az egységes tárgyalás érdekében a bázisvektorok körüljárását is előírhatjuk: általánosan elterjedt szokás a jobbrendszert választani. Az így konstruált bázis vektorait síkban gyakran $\mathbf{i}$, $\mathbf{j}$, térben $\mathbf{i}$, $\mathbf{j}$ és $\mathbf{k}$ jelöli.

A két és háromdimenziós térben a skaláris szorzat egyszerű alakot ölt, ha a koordinátarendszer alapvektorai ortonormáltak.

1.40. ÁLLÍTÁS (SKALÁRIS SZORZAT ORTONORMÁLT KOORDINÁTARENDSZERBEN). *A síkbeli* $\mathbf{u} = (u_1, u_2)$ *és* $\mathbf{v} = (v_1, v_2)$*, illetve a térbeli* $\mathbf{u} = (u_1, u_2, u_3)$ *és* $\mathbf{v} = (v_1, v_2, v_3)$ *vektorok skaláris szorzata ortonormált koordinátarendszerben*

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2, \text{ illetve } \mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + u_3v_3.$$

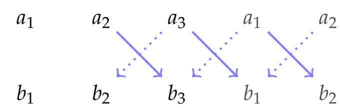
BIZONYÍTÁS. A síkbeli esetben kihasználjuk, hogy $\mathbf{i} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{j} = 1$ és $\mathbf{i} \cdot \mathbf{j} = 0$:

$$\begin{aligned}
\mathbf{u} \cdot \mathbf{v} &= (u_1\mathbf{i} + u_2\mathbf{j}) \cdot (v_1\mathbf{i} + v_2\mathbf{j}) \\
&= u_1v_1\mathbf{i} \cdot \mathbf{i} + (u_1v_2 + u_2v_1)\mathbf{i} \cdot \mathbf{j} + u_2v_2\mathbf{j} \cdot \mathbf{j} \\
&= u_1v_1 + u_2v_2
\end{aligned}$$

A térbeli eset hasonlóan bizonyítható. □

1.41. ÁLLÍTÁS (VEKTORI SZORZAT ORTONORMÁLT KOORDINÁTARENDSZERBEN). *A térbeli $\mathbf{a} = (a_1, a_2, a_3)$ és $\mathbf{b} = (b_1, b_2, b_3)$ vektorok vektori szorzata derékszögű koordinátarendszerben*

$$\mathbf{a} \times \mathbf{b} = (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1).$$



1.39. ábra: A vektori szorzat kiszámítása a két vektor koordinátáiból: írjuk a két vektort egymás alá, majd az első két koordinátát másoljuk a vektorok végére, végül az X alakba rakott nyílpároknál a ↘ nyíl végein lévő számok szorzatából vonjuk ki a ↙ szerinti szorzatot. Az eredmény: $(a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1)$.

BIZONYÍTÁS. Az alapvektorok egymással való vektori szorzatait már kiszámoltuk az 1.28. példában. Kihasználva, hogy $\mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = \mathbf{0}$, $\mathbf{i} \times \mathbf{j} = \mathbf{k}$, $\mathbf{j} \times \mathbf{i} = -\mathbf{k}, \ldots$, kapjuk hogy

$$\begin{aligned}
\mathbf{a} \times \mathbf{b} &= (a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \times (b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}) \\
&= a_2b_3\mathbf{j} \times \mathbf{k} + a_3b_2\mathbf{k} \times \mathbf{j} + a_3b_1\mathbf{k} \times \mathbf{i} + a_1b_3\mathbf{i} \times \mathbf{k} + a_1b_2\mathbf{i} \times \mathbf{j} + a_2b_1\mathbf{j} \times \mathbf{i} \\
&= a_2b_3\mathbf{i} - a_3b_2\mathbf{i} + a_3b_1\mathbf{j} - a_1b_3\mathbf{j} + a_1b_2\mathbf{k} - a_2b_1\mathbf{k} \\
&= (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1).
\end{aligned}$$

□

1.42. PÉLDA (PARALLELOGRAMMA TERÜLETE). *Mutassuk meg, hogy az (a, b) és a (c, d) vektorok által kifeszített parallelogramma területe*

$$|ad - bc|.$$

Mi a jelentése az $ad - bc$ előjelének?

MEGOLDÁS. Két 3-dimenziós vektor által kifeszített parallelogramma területe a vektori szorzatuk abszolút értéke. Ágyazzuk be a megadott két vektort a tér egyik koordinátasíkjába, tekintsük például az $(a, b, 0)$ és a $(c, d, 0)$ vektorokat. Vektori szorzatuk

$$(a, b, 0) \times (c, d, 0) = (0, 0, ad - bc),$$

ennek abszolút értéke $|ad - bc|$.

Mivel az $(a, b, 0)$, $(c, d, 0)$ és $(0, 0, ad - bc)$ vektorok jobbrendszert alkotnak, ezért $ad - bc$ pontosan akkor pozitív, ha a síkban az (a, b) és a (c, d) vektorok jobbrendszert alkotnak, és $ad - bc$ pontosan akkor negatív, ha az (a, b) és a (c, d) vektorok balrendszert alkotnak (gondoljuk meg!). □

Az $\mathbb{R}^n$ halmaz Láttuk, hogy 2-dimenziós, illetve 3-dimenziós vektorjellegű mennyiségek leírhatók egy rendezett számpárral, illetve számhármassal. Izgalmas a fordított helyzet, amikor legalább 4, de akár több millió szorosan összefüggő adatból képzett, rendezett szám-n-essel dolgozunk. Vajon értelmes dolog-e e szám-n-eseket egy n-dimenziós tér vektorainak, vagy pontjainak tekinteni? És van-e értelme a 2- és 3-dimenziós térben használt fogalmak általánosításának n dimenzióra? A válasz mindegyik kérdésre határozott igen, amit a fizika 4-dimenziós tér-idő fogalma, számtalan gazdasági, vagy internettel kapcsolatos probléma megoldása fényesen bizonyít.

Egy tetszőleges H halmaz elemeiből képzett rendezett elem-n-esek halmazát H^n-nel jelöljük.

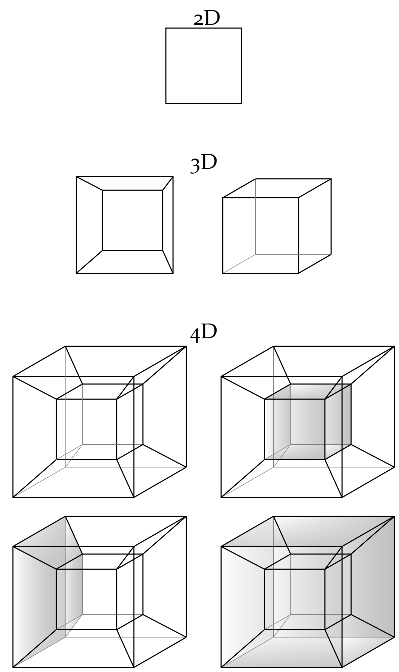
Például ha $H = \{0, 1\}$, akkor H^3 a következő halmaz:

$$\{(0,0,0),(0,0,1),(0,1,0),(1,0,0),(0,1,1),(1,0,1),(1,1,0),(1,1,1)\},$$

hisz H^3 a H elemeiből képzett rendezett elemhármasok halmaza. Eszerint a sík pontjait és vektorait $\mathbb{R}^2$, a térét $\mathbb{R}^3$ elemeivel koordinátáztuk.

A fenti jelölésnek megfelelően $\mathbb{R}^n$ *a valós számokból képzett rendezett szám-n-esek halmazát jelöli.* Így pl. mondhatjuk azt, hogy az $(1, 3, 0, 9, 5)$ számötös eleme az $\mathbb{R}^5$ halmaznak. Később $\mathbb{R}^n$ elemein vektorműveleteket fogunk bevezetni, és $\mathbb{R}^n$-ről, mint vektortérről fogunk beszélni. Hasonlóképp fogjuk $\mathbb{R}^n$-t geometriai vagy ponttérnek tekinteni, ha elemeire, mint pontokra fogunk gondolni, és köztük geometriai műveleteket fogunk végezni. E sokféleség sosem fog gondot okozni: $\mathbb{R}^n$ szerepét mindig az fogja meghatározni, hogy mit teszünk elemeivel, vagyis a szám-n-esekkel.

Az $\mathbb{R}^n$ megismerésében az *analógia* fonalán haladunk, a 2- és 3-dimenziós tér fogalmait fogjuk átvinni, általánosítani n dimenzióra. Ez az analógia fog segíteni abban, hogy magabiztosan érezzük magunkat n dimenzióban is, akkor is, ha ott „nem látunk olyan jól”, mint 3 dimenzióban. Bevezető példaként az analógiára – képernyővédők kedvenc témáját – egy 4-dimenziós kocka 2-dimenziós vetületét mutatjuk az 1.40. ábrán.



1.40. ábra: 4-dimenziós kocka ábrázolása 2-dimenzióban.

$\mathbb{R}^n$ vektorainak összeadása és skalárral szorzása A 2- és 3-dimenziós vektorok műveleteinek koordinátás alakja az összeadás, kivonás, skalárral szorzás és a skaláris szorzás esetén analóg módon kiterjeszthető az n-dimenziós térre. Kezdjük a lineáris kombinációban szereplő műveletek vizsgálatával.

1.43. DEFINÍCIÓ (VEKTORMŰVELETEK $\mathbb{R}^n$-BEN). *Legyen* $\mathbf{u} = (u_1, u_2, \dots, u_n)$ *és* $\mathbf{v} = (v_1, v_2, \dots, v_n)$ *az* $\mathbb{R}^n$ *tér két tetszőleges vektora,* $c \in \mathbb{R}$ *egy tetszőleges valós. E két vektor összegét, különbségét,* *c-szeresén és skaláris szorzatát a következő képletek definiálják:*

$$\begin{aligned}
\mathbf{u}+\mathbf{v} &= (u_1+v_1, u_2+v_2, \ldots, u_n+v_n)\\
\mathbf{u}-\mathbf{v} &= (u_1-v_1, u_2-v_2, \ldots, u_n-v_n)\\
c\mathbf{u} &= (cu_1, cu_2, \ldots, cu_n)\\
\mathbf{u}\cdot\mathbf{v} &= u_1v_1+u_2v_2+\cdots+u_nv_n.
\end{aligned}$$

Összefoglaljuk e műveletek legfontosabb tulajdonságait. Ezekre többször is hivatkozni fogunk a későbbiekben. Először a lineáris kombinációban szereplő műveleteit vizsgáljuk:

1.44. TÉTEL (AZ ÖSSZEADÁS ÉS SKALÁRRAL SZORZÁS TULAJDONSÁGAI). *Legyen* $\mathbf{u}$, $\mathbf{v}$ *és* $\mathbf{w}$ *az* $\mathbb{R}^n$ *három tetszőleges vektora, és legyen* c, d *két tetszőleges valós, jelölje* $\mathbf{0}$ *a* $(0,0,\ldots,0)$ *vektort és* $-\mathbf{u}$ *a* $(-u_1,-u_2,\ldots,-u_n)$ *vektort. Ekkor*

a)	$\mathbf{u}+\mathbf{v}\in\mathbb{R}^n$	*a + művelet nem vezet ki* $\mathbb{R}^n$*-ből*
b)	$\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$	*a művelet fölcserélhető (kommutatív)*
c)	$\mathbf{u}+(\mathbf{v}+\mathbf{w})=(\mathbf{v}+\mathbf{u})+\mathbf{w}$	*csoportosítható (asszociatív)*
d)	$\mathbf{u}+\mathbf{0}=\mathbf{u}$	*zérusvektor*
e)	$\mathbf{u}+(-\mathbf{u})=\mathbf{0}$	*ellentett vektor*
f)	$c\mathbf{u}\in\mathbb{R}^n$	*e szorzás nem vezet ki* $\mathbb{R}^n$*-ből*
g)	$c(d\mathbf{u})=(cd)\mathbf{u}$	*a két szorzás kompatibilis*
h)	$c(\mathbf{u}+\mathbf{v})=c\mathbf{u}+c\mathbf{v}$	*disztributív*
i)	$(c+d)\mathbf{u}=c\mathbf{u}+d\mathbf{u}$	*disztributív*
j)	$1\mathbf{u}=\mathbf{u}$	*szorzás* 1*-gyel*

E tíz tulajdonság később kitüntetett szerepet fog játszani, ezért elkülönítjük a tovább felsorolandóktól:

1) $0\mathbf{u}=\mathbf{0}$
2) $(-1)\mathbf{u}=-\mathbf{u}$
3) $\mathbf{u}-\mathbf{v}=\mathbf{u}+(-\mathbf{v})$
4) a skalár hátra is írható, azaz $c\mathbf{u}=\mathbf{u}c$ *és így* $\mathbf{u}/c=\frac{1}{c}\mathbf{u}$

► E tulajdonságok mindegyike könnyen visszavezethető a valós számok algebrai tulajdonságaira, ezért ezek ellenőrzését (bizonyítását) az Olvasóra hagyjuk. Mintaként megmutatjuk a *b)* bizonyítását:

$$\begin{aligned}
\mathbf{u}+\mathbf{v} &= (u_1+v_1, u_2+v_2, \ldots, u_n+v_n)\\
&= (v_1+u_1, v_2+u_2, \ldots, v_n+u_n)\\
&= \mathbf{v}+\mathbf{u}.
\end{aligned}$$

► Az *a)–e)* tulajdonságok az összeadás, az *f)–j)* tulajdonságok a skalárral szorzás tulajdonságait írják le. Mindkét csoport első tulajdonsága csak annyit mond, hogy a művelet eredménye is ugyanabba a vektorhalmazba, azaz $\mathbb{R}^n$-be esik, ahová a műveletben szereplő vektor való.

1.45. PÉLDA. *Mutassuk meg, hogy az $\mathbb{R}^n$-beli $\mathbf{e}_1 = (1, 0, \ldots, 0)$, $\mathbf{e}_2 = (0, 1, \ldots, 0)$,…, $\mathbf{e}_n = (0, 0, \ldots, 1)$ vektorok lineárisan függetlenek, és hogy $\mathbb{R}^n$ minden vektora egyértelműen előáll ezek lineáris kombinációjaként!*

MEGOLDÁS. Az $\mathbf{e}_1$ biztosan nem áll elő a többi vektor lineáris kombinációjaként, hisz a többi vektor első koordinátája 0, így azok bármely lineáris kombinációjában is 0 az első koordináta, $\mathbf{e}_1$-ben pedig 1. Hasonlóan igazolható, hogy egyik $\mathbf{e}_i$ sem áll elő a többi vektor lineáris kombinációjaként ($i = 2, 3, \ldots, n$). A megadott vektorok tehát lineárisan függetlenek.

Mivel az i-edik koordináta egyedül csak az $\mathbf{e}_i$ vektorban 1, a többiben 0, ezért ha egy tetszőleges $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ vektor előáll az $\mathbf{e}_1$, $\mathbf{e}_2$,…, $\mathbf{e}_n$ vektorok lineáris kombinációjaként, akkor abban $\mathbf{e}_i$ együtthatója csak v_i lehet. Másrészt az is világos, hogy

$$(v_1, v_2, \ldots, v_n) = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \cdots + v_n\mathbf{e}_n.$$

Ezzel igazoltuk, hogy $\mathbb{R}^n$ minden vektora egyértelműen áll elő az $\mathbf{e}_1$, $\mathbf{e}_2$,…, $\mathbf{e}_n$ vektorok lineáris kombinációjaként. □

Lineáris kombináció, lineáris függetlenség, lineáris összefüggőség Hiába definiáltuk vektorok lineáris függetlenségének fogalmát tetszőleges számú vektorból álló vektorhalmazra, láttuk, hogy a 3-dimenziós térben legföljebb csak 3 vektor lehet lineárisan független. Viszont az $\mathbb{R}^n$ térben találtunk n lineárisan független vektort is. Térjünk vissza e fogalom vizsgálatára.

Az 1.10. definíció szerint egy legalább kételemű vektorrendszer lineárisan független, ha mindegyik vektor független a többitől, azaz egyik sem fejezhető ki a többi lineáris kombinációjaként, egyetlen vektor pedig lineárisan független, ha nem a zérusvektor. Ez nehézkes feltétel, hisz mindegyik vektorra külön ellenőrizni kell, ezért egy könnyebben ellenőrizhető, de ekvivalens feltételt keresünk. A háromdimenziós térben láttuk, hogy ha három vektor független, akkor a tér bármely vektora *egyértelműen* előáll lineáris kombinációjukként. Ez igaz a nullvektorra is. A nullvektor egyféleképp biztosan előáll: a három vektor nullákkal vett lineáris kombinációjaként. Ezt nevezzük a nullvektor triviális előállításának. És a fentiek szerint más előállítása nincs is, ha a három vektor lineárisan független. Ez az alapja a következő tételnek:

1.46. TÉTEL (LINEÁRIS FÜGGETLENSÉG). *Tetszőleges $\mathbb{R}^n$-beli $\mathcal{V} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ vektorrendszerre az alábbi két állítás ekvivalens:*

1. $\mathcal{V}$ lineárisan független, *azaz $k > 1$ esetén egyik vektora sem fejezhető* *ki a többi lineáris kombinációjaként, $k = 1$ esetén pedig a vektor nem a zérusvektor.*

2. *A zérusvektor csak egyféleképp – a triviális módon – áll elő $\mathcal{V}$ lineáris kombinációjaként. Másként fogalmazva, a c_1, c_2,. . . ,c_k skalárokkal vett lineáris kombináció csak akkor lehet a nullvektor, azaz*

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k = \mathbf{0}$$

csak akkor állhat fenn, ha

$$c_1 = c_2 = \ldots = c_k = 0.$$

Bizonyítás. Először tegyük fel, hogy a vektorrendszer csak egyetlen $\mathbf{v}$ vektorból áll. Ekkor a tétel azt állítja, hogy e vektor pontosan akkor lineáris független, azaz pontosan akkor nem a nullvektor, ha a $c\mathbf{v} = \mathbf{0}$ csak $c = 0$ esetén állhat fenn. Ez nyilvánvaló, hisz ha $\mathbf{v} \neq \mathbf{0}$ és $c \neq 0$, akkor $c\mathbf{v} = \mathbf{0}$ sem állhat fenn. A továbbiakban tegyük fel, hogy a vektorrendszer legalább két vektorból áll.

($\Longleftarrow$) Megmutatjuk, hogy ha $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k = \mathbf{0}$ csak $c_1 = c_2 = \ldots = c_k = 0$ esetén állhat fenn, akkor semelyik $\mathbf{v}_i$ vektor sem fejezhető ki a többi lineáris kombinációjaként ($i = 1, 2, \ldots, k$). Tegyük fel indirekt módon, hogy valamelyik vektor – például a $\mathbf{v}_1$ – kifejezhető a többi lineáris kombinációjaként, azaz

$$\mathbf{v}_1 = d_2\mathbf{v}_2 + \ldots + d_k\mathbf{v}_k,$$

vagyis átrendezés után

$$(-1)\mathbf{v}_1 + d_2\mathbf{v}_2 + \ldots + d_k\mathbf{v}_k = \mathbf{0}.$$

Mivel $\mathbf{v}_1$ együtthatója nem 0, ezért feltevésünkkel ellentmondásban előállítottuk a nullvektort olyan lineáris kombinációként, melyben nem minden együttható 0. Ez az ellentmondás bizonyítja az állítást.

($\Longrightarrow$) Megmutatjuk, hogy ha a vektorrendszer egyik vektora sem áll elő a többi lineáris kombinációjaként, akkor a egyedül csak a csupa zérus együtthatójú lineáris kombinációja lehet zérusvektor. Ismét indirekt módon bizonyítunk: tegyük fel, hogy van olyan – nem csupa 0 együtthatójú – lineáris kombináció, mely a nullvektorral egyenlő, azaz

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k = \mathbf{0},$$

de valamelyik együttható – például a c_1 – nem 0. Ekkor $\mathbf{v}_1$ kifejezhető a többi vektor lineáris kombinációjaként:

$$\mathbf{v}_1 = -\frac{c_2}{c_1}\mathbf{v}_2 - \ldots - \frac{c_k}{c_1}\mathbf{v}_k,$$

ami bizonyítja az állítást. □

Egy vektorrendszert *lineárisan összefüggőnek* nevezünk, ha nem független, azaz egyelemű vektorrendszer esetén ha az a vektor a zérusvektor, többelemű vektorrendszer esetén pedig ha van olyan vektora, mely kifejezhető a többi lineáris kombinációjaként. Az előző tétel szerint ez azzal ekvivalens, hogy a vektorrendszernek van olyan zérusvektort adó lineáris kombinációja, melyben nem mindegyik együttható zérus. A lineáris összefüggőség definíciója kicsit élesíthető:

1.47. TÉTEL (LINEÁRIS ÖSSZEFÜGGŐSÉG). *Egy nullvektortól különböző elemekből álló, legalább kételemű $\mathbb{R}^n$-beli $V = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ vektorrendszer pontosan akkor lineárisan összefüggő, ha valamely $2 \le t$ esetén $\mathbf{v}_t$ a $\mathbf{v}_1$, $\mathbf{v}_2, \dots$, $\mathbf{v}_{t-1}$ vektorok lineáris kombinációja.*

Másként fogalmazva, ha egy nullvektort nem tartalmazó vektorrendszerben találunk olyan vektort, mely a többi lineáris kombinációja, akkor olyat is találunk, mely sorrendben csak az őket megelőző vektor(ok) lineáris kombinációja.

BIZONYÍTÁS. Legyen t az a legkisebb egész, melyre a $\mathbf{v}_1$, $\mathbf{v}_2, \dots$, $\mathbf{v}_t$ vektorok már összefüggők. Mivel $\mathbf{v}_1 \ne \mathbf{0}$, ezért az első vektor nem lehet összefüggő, ezért $t \ge 2$. E vektorok összefüggősége miatt vannak olyan c_i konstansok, melyekkel

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_t\mathbf{v}_t = \mathbf{0}.$$

Biztos, hogy $c_t \ne 0$, különben már a $\mathbf{v}_1$, $\mathbf{v}_2, \dots$, $\mathbf{v}_{t-1}$ vektorok is lineáris összefüggők lennének, és ez ellentmond t definíciójának. Így

$$\mathbf{v}_t = \frac{-c_1}{c_t}\mathbf{v}_1 + \frac{-c_2}{c_t}\mathbf{v}_2 + \dots + \frac{-c_{t-1}}{c_t}\mathbf{v}_{t-1},$$

ami bizonyítja, hogy összefüggő vektorrendszerben létezik ilyen vektor. Az állítás másik fele definíció szerint igaz, hisz ha létezik ilyen $\mathbf{v}_t$ vektor, akkor ez valóban lineáris kombinációja az összes többi vektornak. □

Skaláris szorzás $\mathbb{R}^n$-ben A skaláris szorzás koordinátákkal megadott definíciója (ld. 1.43.) segítségével először csak azokat a legfontosabb tulajdonságait említjük, melyek később a fogalom további általánosítására is használhatók lesznek.

1.48. TÉTEL (A SKALÁRIS SZORZÁS TULAJDONSÁGAI). *Legyen* $\mathbf{u}$, $\mathbf{v}$ *és* $\mathbf{w}$ *az* $\mathbb{R}^n$ *három tetszőleges vektora, és legyen* c *egy teszőleges valós. Ekkor*

a)	$\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$	*a művelet fölcserélhető (kommutatív)*
b)	$\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$	*disztributív*
c)	$(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v})$	*a két szorzás kompatibilis*
d)	$\mathbf{u} \cdot \mathbf{u} \ge 0$ *és* $\mathbf{u} \cdot \mathbf{u} = 0$ *pontosan akkor teljesül, ha* $\mathbf{u} = \mathbf{0}$.	

BIZONYÍTÁS. A bizonyítás itt is igen egyszerű, ezért csak az *a)* pontét mutatjuk meg, a többit az Olvasóra hagyjuk:

$$\begin{aligned}\mathbf{u}\cdot\mathbf{v} &= u_1v_1 + u_2v_2 + \cdots + u_nv_n \\ &= v_1u_1 + v_2u_2 + \cdots + v_nu_n \\ &= \mathbf{v}\cdot\mathbf{u}.\end{aligned}$$ □

Távolság és szög $\mathbb{R}^n$-ben Két 2- vagy 3-dimenziós vektor (végpontja) távolságának, és szögének a skaláris szorzatukkal való kapcsolatát használjuk e fogalmaknak a magasabb dimenziós terekben való definíciójához.

1.49. DEFINÍCIÓ (ABSZOLÚT ÉRTÉK, SZÖG, MERŐLEGESSÉG, TÁVOLSÁG). *Legyen* $\mathbf{u}$ *és* $\mathbf{v}$ *az* $\mathbb{R}^n$ *tér két tetszőleges vektora.*

a) *Az* $\mathbf{u}$ *vektor hosszán önmagával vett skaláris szorzatának gyökét értjük, azaz*

$$|\mathbf{u}| = \sqrt{\mathbf{u}\cdot\mathbf{u}}. \tag{1.3}$$

b) *Az* $\mathbf{u}$ *és* $\mathbf{v}$ *vektorok* (hajlás)szögének *koszinusza az alábbi törttel definiálható:*

$$\cos(\mathbf{u},\mathbf{v})_\angle := \frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}||\mathbf{v}|} \tag{1.4}$$

c) *Azt mondjuk, hogy az* $\mathbf{u}$ *és* $\mathbf{v}$ *vektorok* merőlegesek *egymásra, ha*

$$\mathbf{u}\cdot\mathbf{v} = 0. \tag{1.5}$$

d) *A két vektor végpontjának távolságán, amit egyszerűen a két* vektor távolságának *nevezünk, a*

$$d(\mathbf{u},\mathbf{v}) = |\mathbf{u}-\mathbf{v}| \tag{1.6}$$

értéket értjük.

► A fenti definíciókat érdemes megtekinteni koordinátás alakjukba átírva. Eszerint például

$$\begin{aligned}|\mathbf{u}| &= \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}, \\ \cos(\mathbf{u},\mathbf{v})_\angle &= \frac{u_1v_1 + u_2v_2 + \cdots + u_nv_n}{\sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}\sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}}\end{aligned}$$

► A fenti definíciók közül a vektorok hajlásszögének definíciója még hiányos. Egy szög koszinusza csak a $[-1,1]$ intervallumba eső szám lehet, ezért e definíció csak akkor értelmes, ha az (1.4) képlet számlálójára és nevezőjére igaz, hogy $|\mathbf{u}\cdot\mathbf{v}| \leq |\mathbf{u}||\mathbf{v}|$, azaz ha n-dimenziós vektorokra is fönnáll a CBS-egyenlőtlenség. Ezt hamarosan igazolni fogjuk!

1.50. PÉLDA (VEKTOROK SZÖGE ÉS TÁVOLSÁGA). *Az* $\mathbf{u} = (2,3,4,14)$ *vektornak mennyi az abszolút értéke, mennyi a* $\mathbf{v} = (4,6,-10,10)$ *vektortól való távolsága, és mennyi a* $\mathbf{w} = (0,3,6,-2)$ *vektorral bezárt szögének koszinusza?*

MEGOLDÁS. A válaszhoz az (1.3), az (1.6) és az (1.4) képleteket használjuk:

$$\begin{aligned}
|\mathbf{u}| &= \sqrt{2^2+3^2+4^2+14^2} = \sqrt{225} = 15,\\
d(\mathbf{u},\mathbf{v}) &= \sqrt{(2-4)^2+(3-6)^2+(4-(-10))^2+(14-10)^2}\\
&= \sqrt{2^2+3^2+14^2+4^2} = 15\\
\cos(\mathbf{u},\mathbf{w})_\angle &= \frac{2\cdot 0+3\cdot 3+4\cdot 6+14\cdot(-2)}{\sqrt{2^2+3^2+4^2+14^2}\sqrt{0^2+3^2+6^2+(-2)^2}} = \frac{1}{21}.
\end{aligned}$$
□

1.51. TÉTEL (CAUCHY–BUNYAKOVSZKIJ–SCHWARZ-EGYENLŐTLENSÉG). *Tetszőleges* $\mathbf{u},\mathbf{v} \in \mathbb{R}^n$ *vektorokra*

$$|\mathbf{u}\cdot\mathbf{v}| \le |\mathbf{u}||\mathbf{v}|. \tag{1.7}$$

Egyenlőség pontosan akkor áll, ha $\mathbf{u}$ *és* $\mathbf{v}$ *lineárisan összefüggők, azaz egyik vektor a másik skalárszorosa.*

BIZONYÍTÁS. Tegyük fel először, hogy $\mathbf{v} = \mathbf{0}$. Ekkor a tétel állításának mindkét része nyilván igaz, hisz egyenlőség áll fenn, és a két vektor lineárisan összefüggő. Ha $\mathbf{v} \ne \mathbf{0}$, akkor legyen $\mathbf{e} = \mathbf{v}/|\mathbf{v}|$ a $\mathbf{v}$ irányú egységvektor. Az $\mathbf{u}$ vektor $\mathbf{e}$ egyenesére merőleges összetevőjének hossza, illetve annak négyzete nyilván nem negatív, azaz

$$\begin{aligned}
0 &\le |\mathbf{u}-(\mathbf{u}\cdot\mathbf{e})\mathbf{e}|^2\\
&= |\mathbf{u}|^2 + |\mathbf{u}\cdot\mathbf{e}|^2 - 2\,|\mathbf{u}\cdot\mathbf{e}|^2\\
&= |\mathbf{u}|^2 + |\mathbf{u}\cdot\mathbf{e}|^2\\
&= |\mathbf{u}|^2 - \frac{|\mathbf{u}\cdot\mathbf{v}|^2}{|\mathbf{v}|^2},
\end{aligned}$$

Innen átrendezéssel azonnal megkapjuk a bizonyítandó állítást. Másrészt az is világos, hogy a merőleges összetevő pontosan akkor egyenlő a nullvektorral, ha $\mathbf{u}$ megegyezik a $\mathbf{v}$ egyenesére eső merőleges vetületével, azaz ha $\mathbf{u}$ a $\mathbf{v}$ skalárszorosa, azaz ha a két vektor lineárisan összefüggő. □

1.52. TÉTEL (HÁROMSZÖG-EGYENLŐTLENSÉG $\mathbb{R}^n$-BEN). *Tetszőleges* $\mathbf{u},\mathbf{v} \in \mathbb{R}^n$ *vektorokra*

$$|\mathbf{u}+\mathbf{v}| \le |\mathbf{u}| + |\mathbf{v}|.$$

A bizonyítás megegyezik a 3-dimenziós változatra, azaz a az 1.22. tételre adott bizonyítással.

A vektor abszolút értékét a skaláris szorzat segítségével definiáltuk, de fordítva, a skaláris szorzat is kifejezhető a vektor abszolút értékével.

1.53. TÉTEL (SKALÁRIS SZORZAT ÉS ABSZOLÚT ÉRTÉK $\mathbb{R}^n$-BEN). *Tetszőleges $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ vektorokra*

$$\mathbf{u} \cdot \mathbf{v} = \frac{1}{4}\left(|\mathbf{u}+\mathbf{v}|^2 - |\mathbf{u}-\mathbf{v}|^2\right) \tag{1.8}$$

$$\mathbf{u} \cdot \mathbf{v} = \frac{1}{2}\left(|\mathbf{u}+\mathbf{v}|^2 - |\mathbf{u}|^2 - |\mathbf{v}|^2\right) \tag{1.9}$$

BIZONYÍTÁS. A bizonyítás az abszolút érték (1.3)-beli definícióját használjuk:

$$\begin{aligned}
\frac{1}{4}\left(|\mathbf{u}+\mathbf{v}|^2 - |\mathbf{u}-\mathbf{v}|^2\right) &= \frac{1}{4}\left((\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) - (\mathbf{u}-\mathbf{v})\cdot(\mathbf{u}-\mathbf{v})\right)\\
&= \frac{1}{4}\left(\mathbf{u}\cdot\mathbf{u} + \mathbf{u}\cdot\mathbf{v} + \mathbf{v}\cdot\mathbf{u} + \mathbf{v}\cdot\mathbf{v} - \mathbf{u}\cdot\mathbf{u} + \mathbf{u}\cdot\mathbf{v} + \mathbf{v}\cdot\mathbf{u} - \mathbf{v}\cdot\mathbf{v}\right)\\
&= \frac{1}{4}(4\mathbf{u}\cdot\mathbf{v}) = \mathbf{u}\cdot\mathbf{v}.
\end{aligned}$$

A másik formula hasonlóan bizonyítható. □

Korrelációs együttható A következőkben egy fontos alkalmazást mutatunk, melyben az n-dimenziós térben szerzett friss szemléletünk segít a fogalom megértésében.

Adva van két adatsor: $x_1, x_2, \dots x_n$ és $y_1, y_2, \dots y_n$. Átlagukat jelölje $\bar{x}$, illetve $\bar{y}$, azaz

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}, \quad \bar{y} = \frac{y_1 + y_2 + \ldots + y_n}{n}.$$

Az r-rel jelölt ún. *korrelációs együttható* azt méri, hogy a két adatsor közti lineáris függvénykapcsolat milyen erős. Az erre használt képlet a következő:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}.$$

Vajon hogyan méri r a lineáris függvénykapcsolat erősségét? Az eddig tanultak szemléletes választ adnak. A fogalom mélyebb megértése a valószínűségszámítás ismeretét is megkívánja, ezzel mi itt nem foglalkozunk.

Tegyük fel, hogy a két adatsor közt fönnáll az $y_i = cx_i + d$ lineáris függvénykapcsolat minden $i = 1, 2, \dots n$ indexre valamely c, d konstansokkal. Ha mindkét adatsorból levonjuk az átlagukat (a két adatsort normáljuk), akkor az így kapott

$$a_i = x_i - \bar{x}, \quad b_i = y_i - \bar{y} \quad (i = 1, 2, \ldots, n)$$

adatsorokra igaz a $b_i = ca_i$ összefüggés. Ugyanis

$$\bar{y} = \frac{1}{n}\left(\sum_{i=1}^{n} cx_i + d\right) = c\bar{x} + d,$$

amiből $b_i = y_i - \bar{y} = cx_i + d - c\bar{x} - d = c(x_i - \bar{x}) = ca_i$.

Tehát, az adatsorokat n-dimenziós vektorokba foglalva az $\mathbf{y} = c\mathbf{x} + d$ lineáris függvénykapcsolat pontosan akkor áll fönn, ha $\mathbf{b} = c\mathbf{a}$, azaz ha a normált vektorok kollineárisak. A korrelációs együttható nem más, mint e két utóbbi vektor szögének koszinusza, ugyanis

$$\begin{aligned}\cos(\mathbf{a}, \mathbf{b})_\angle = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|} &= \frac{\sum_{i=1}^{n} a_i b_i}{\sqrt{\sum_{i=1}^{n} a_i^2}\sqrt{\sum_{i=1}^{n} b_i^2}} \\ &= \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} = r.\end{aligned}$$

Valóban, a két vektor szögének koszinusza pontosan akkor 1, ha a vektorok szöge 0, és akkor -1, ha a szög π. A korreláció tehát -1 és 1 közt változik, és abszolút értéke annál kisebb, minél nagyobb az **a** és **b** vektorok hajlásszöge, azaz minél kevésbé kollineárisak, azaz minél kevésbé erős a két számsorozat közti lineáris függvénykapcsolat.

Ha $r = 0$, akkor **a** és **b** merőlegesek, ekkor lineáris függvénykapcsolat nincs az eredeti két mennyiség közt (más kapcsolat még lehet, tehát nem feltétlenül független a két adatsor egymástól).

Bitvektorok A modern számítógépek memóriájában vagy háttértárolóin az adatok tárolásának legkisebb egysége a bit. Egy bittel két állapot tárolható, melyeket a 0 és 1 számokkal jelölünk, de amelyek több mindent is reprezentálhatnak: hamis/igaz, nem/igen, ki/be,.... A biteket a hardver lehetőségei és a feladat igényei szerint csoportokba, sorozatokba, vektorokba gyűjtik, melyekkel különféle műveletek végezhetők. Ezek attól is függnek, hogy a bitvektorok milyen adatokat kódolnak. E műveletek közül minket azok fognak érdekelni, melyek algebrailag a korábban megismert vektorműveletekre hasonlítanak. Később látni fogjuk, hogy itt többről van szó, mint véletlen hasonlóságról.

Bit: az angol ***b**inary dig**it*** kifejezésből képzett szó, ami magyarul bináris, azaz kettes számrendszerbeli számot jelent. A szoftver (software) szót is megalkotó John W. Tukey ötlete.

A bitvektorok jelölése a számítástechnikában leginkább a biteket jelölő számjegyek egyszerű egymás mellé írásával történik, pl. 01110101 a $(0, 1, 1, 1, 0, 1, 0, 1)$ vektort jelöli. Mi mindkét jelölést használni fogjuk.

Kódvektorok, kódok Az információ tárolásához, továbbításához szükség van kódolásra, gyakran egymás után több különböző kódolási eljárás alkalmazására is. Például ha valaki egy gondolatát leírja egy papírra, majd onnan begépeli egy számítógépbe és egy adott formátumban elmenti, egy tömörítő programmal összetömöríti, egy titkosító szoftverrel titkosítja, végül egy biztonságos internetes csatornán továbbküldi valakinek, hat kódolási lépést hajthatott végre.

A kódoláshoz mi a továbbiakban mindig egy rögzített, véges kódábécét használunk, amelynek betűi általában a 0-tól $n-1$-ig terjedő egészek lesznek. Az ábécé „betűiből”, azaz elemeiből képzett vektorokat *kódvektor*oknak vagy *kódszavak*nak nevezzük. A bitvektorok is kódvektorok, ahol a kódábécé a kételemű $\{0,1\}$ halmaz.

A kódvektorok koordinátáinak számát, vagyis a kódvektor dimenzióját a kód *hosszának* nevezzük. Ez természetesen nem analóg fogalom a vektor abszolút értékével.

A személyi szám egy olyan kódvektorra példa, melynek hossza 11, az ábécé pedig a 10-elemű $\{0,1,\ldots,9\}$ halmaz. Nem minden 11-hosszú decimális kódvektor érvényes személyi szám, vagyis a személyi számok a 11-hosszú kódvektorok halmazának egy részhalmazát alkotják. Az ilyen részhalmazokat nevezzük kódnak.

1.54. definíció (Kód). *A* kód *azonos ábécéből képzett azonos hosszúságú kódszavak egy halmaza.* Kódolás *az a folyamat, amikor az üzenethez kódszavakat rendelünk,* dekódolás *az ellenkező irányú folyamat.*

Decimális számok egyik szokásos kódolása a BCD (Binary Coded Decimal), mely a számok kódolására a kettes számrendszerbe való átírás helyett a számjegyenként való kódolást választja. Több változata is van, a legegyszerűbbikben egy-egy számjegynek 4-4 bit felel meg, így a 16 lehetséges 4-hosszú kódszó helyett csak tízet használ: a 0, 1,…, 9 jegyek kódja rendre 0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001.

1.55. példa (BCD-kód). *Hasonlítsuk össze* 561 *bináris alakját és BCD-kódját!*

Megoldás. Az 561 BCD-kódja három kódvektorból áll: 0101 0110 0001. A kettes számrendszerbeli alak 1000110001, amit úgy kapunk meg, hogy 561-et osztjuk 2-vel, a maradék a bináris alak utolsó jegye, a hányadost ismét elosztjuk 2-vel, annak maradéka az utolsó előtti jegy…

Vektorműveletek $\mathbb{Z}_m^n$-ben A $\mathbb{Z}_m$-re vonatkozó ismereteket a függelékben részletezzük. A korábbiakban bevezetett jelölések szerint $\mathbb{Z}_m^n$ a $\mathbb{Z}_m$-beli n-hosszú vektorokból áll. E vektorok összeadása, skalárral való szorzása és skaláris szorzása a $\mathbb{Z}_m$-beli műveletekkel az $\mathbb{R}^n$-beli vektorműveletekhez hasonlóan végezhető el. Ennek következtében a lineáris kombináció, lineáris függetlenség itt is ugyanúgy definiálható és használható.

1.56. példa (Lineáris kombináció $\mathbb{Z}_m^n$-ben). *Számítsuk ki a $\mathbb{Z}_2^5$-beli*

$$\mathbf{a}=(1,0,0,1,1,0),\ \mathbf{b}=(0,1,0,1,0,1) \textit{ és } \mathbf{c}=(0,0,1,0,1,1)$$

vektorok összes lineáris kombinációját $\mathbb{Z}_2$-beli együtthatókkal, valamint a $\mathbb{Z}_3^3$-beli

$$\mathbf{u} = (1,1,0) \text{ és } \mathbf{v} = (0,1,1)$$

vektorok összes lineáris kombinációját $\mathbb{Z}_3$-beli együtthatókkal.

MEGOLDÁS. A lehetséges $x\mathbf{a} + y\mathbf{b} + z\mathbf{c}$ alakú lineáris kombinációk száma 8, hisz $x, y, z \in \mathbb{Z}_2$, vagyis mindegyik együtthatónak 0 vagy 1 az értéke, és ez $2 \cdot 2 \cdot 2 = 8$ eshetőség. Az $x = y = z = 0$ eset a zérusvektort adja. Ha x, y és z közül csak egyikük értéke 1, a többi 0, akkor a három adott vektort kapjuk vissza. Azok az esetek maradnak, amikor legalább két vektort kell összeadni. Például $1\mathbf{a} + 1\mathbf{b} + 0\mathbf{c} = (1,0,0,1,1,0) + (0,1,0,1,0,1) = (1,1,0,0,1,1)$. Az összes lineáris kombináció az 1.1. (a) táblázatban látható.

Az $x\mathbf{u} + y\mathbf{v}$ alakú lineáris kombinációk száma 9, mivel $x, y \in \mathbb{Z}_3$, azaz lehetséges értékük 0, 1 vagy 2, ami $3 \cdot 3 = 9$ lehetőséget ad. Példaként egy lineáris kombináció, a többi az 1.1. (b) táblázatban látható: $2\mathbf{u} + 1\mathbf{v} = 2(1,1,0) + (0,1,1) = (2,2,0) + (0,1,1) = (2,0,1)$. □

$x\ y\ z$	$x\mathbf{a} + y\mathbf{b} + z\mathbf{c}$
0 0 0	$(0,0,0,0,0,0)$
1 0 0	$(1,0,0,1,1,0)$
0 1 0	$(0,1,0,1,0,1)$
0 0 1	$(0,0,1,0,1,1)$
1 1 0	$(1,1,0,0,1,1)$
1 0 1	$(1,0,1,1,0,1)$
0 1 1	$(0,1,1,1,1,0)$
1 1 1	$(1,1,1,0,0,0)$

(a)

$x\ y$	$x\mathbf{u} + y\mathbf{v}$
0 0	$(0,0,0)$
1 0	$(1,1,0)$
2 0	$(2,2,0)$
0 1	$(0,1,1)$
1 1	$(1,2,1)$
2 1	$(2,0,1)$
0 2	$(0,2,2)$
1 2	$(1,0,2)$
2 2	$(2,1,2)$

(b)

1.1. táblázat: Vektorok lineáris kombinációi *(a)* $\mathbb{Z}_2$ és *(b)* $\mathbb{Z}_3$ fölött.

E paragrafus további részében a vektorműveletekre két alkalmazást mutatunk.

1.57. PÉLDA (ONE TIME PAD – A TÖKÉLETES TITKOSÍTÁS). *Tegyük fel, hogy az üzenet küldése előtt a küldő és a fogadó megegyezett egy titkos kulcsban, mely egy olyan hosszú véletlen bitvektor, mint amilyen az üzenet legföljebb lehet. Legyen a kulcs $\mathbf{k} \in \mathbb{Z}_2^m$. Legyen a titkosítandó üzenet $\mathbf{u} \in \mathbb{Z}_2^m$. A titkosítás egyszerűen annyi, hogy küldő kiszámolja az $\mathbf{u} + \mathbf{k}$ vektort, és azt küldi a fogadónak, aki a titkosított üzenethez maga is hozzáadja a kulcsot, és mivel bármely $\mathbf{x} \in \mathbb{Z}_2^m$ vektorra $\mathbf{x} + \mathbf{x} = \mathbf{0}$, ezért $(\mathbf{u} + \mathbf{k}) + \mathbf{k} = \mathbf{u} + (\mathbf{k} + \mathbf{k}) = \mathbf{u}$, vagyis a fogadó így valóban megfejti az üzenetet. Példaként egy üzenet, egy kulcs és a kettő összege a tömör bitvektor-jelöléssel:*

$$\begin{aligned} \mathbf{u} &= 01010101000011111111111 \\ \mathbf{k} &= 00101100010110100101101 0, \text{ ekkor} \\ \mathbf{u} + \mathbf{k} &= 01111001010101011010010 1 \end{aligned}$$

E titkosítás hátránya, hogy a $\mathbf{k}$ kulcs csak egyszer használható fel, mert két különböző $\mathbf{u}_1 + \mathbf{k}$ és $\mathbf{u}_2 + \mathbf{k}$ üzenetet elcsípve és összeadva az $(\mathbf{u}_1 + \mathbf{k}) + (\mathbf{u}_2 + \mathbf{k}) = \mathbf{u}_1 + \mathbf{u}_2$ vektorban már nem szerepel $\mathbf{k}$, és ez statisztikai módszerekkel már megfejthető.

A kódelmélet egyik célja, hogy redundáns információ hozzáadásával elérje az elküldött üzenet megérkezését zajos, veszteséges csatornán keresztül is. Ennek két gyakran alkalmazott típusa a hibajelző és a hibajavító kód: az előbbi az átvitel során bekövetkezett bizonyos hibákat jelez a fogadó számára, míg az utóbbi bizonyos hibák kijavítását is lehetővé teszi. Az 1.56. példában előállított lineáris kombinációk hibajelző kódok, egyikük hibajavító is. Az 1.21. feladat arra kérdez, hogy milyen hibát jeleznek, illetve javítanak.

Az elektronikus számítógépek adatkezelésének egyik első ötlete az adattárolás vagy továbbítás biztonságosabbá tételére a paritásbit. Ha egy n-hosszú $\mathbf{b}$ bitvektorhoz még egy bitet csatolunk, melynek értéke 1, ha $\mathbf{b}$-ben páratlan sok bit egyenlő 1-gyel, egyébként 0, akkor olyan $(n+1)$-hosszú vektort kapunk, melyben páros sok 1-esnek kell lenni. Ha egy bit elromlik, páratlan sok 1-es lesz, tehát ez az $(n+1)$-edik bit hibajelző. Ezt nevezzük *paritásbitnek*.

1.58. példa (Paritásbit). *Írjuk fel a paritásbitet skaláris szorzatként, és döntsük el, mely vektorokra igaz, hogy minden bitjük a vektor maradék részének paritásbitje.*

Megoldás. A paritásbit $\mathbb{Z}_2$ fölött $\mathbf{1}\cdot\mathbf{b}$ alakba írható, ahol $\mathbf{1}$ a $\mathbf{b}$-vel azonos hosszúságú és csupa 1-esből álló vektor.

Minden olyan vektor, amelyben páros sok 1-es van, eleget tesz a feladatbeli feltételnek, a többi nem. □

A paritásbit általánosítása az ún. *ellenőrző összeg*, melyre számtalan példát találunk a mindennapi életben.

A magyar személyi szám a személyre jellemző 10 jegyből, és az azt követő e ellenőrző összegből áll. Az e kiszámítási képlete

$$\mathbb{Z}_{11}\text{-ben számolva:}\quad e = (1,2,3,4,5,6,7,8,9,10)\cdot\mathbf{u},$$

ahol $\mathbf{u}$ a személyi szám első 10 jegye. A személyi szám 8-10-edik jegye úgy van kiválasztva, hogy $e \neq 10$, így az ellenőrző összeg egyjegyű. 2007 előtt egy hasonló képlettel számolták a könyvek ún. ISBN-10 kódjának ellenőrző jegyét, ami ha 10 volt, X-et írtak (római 10-es).

A termékek EAN-kódja (European Article Number) egy 13-jegyű, a termék azonosítására szolgáló kód, melyhez egy vonalkód is tartozik. A 13-dik jegy az ellenőrző összeg. Ha az EAN kódvektort $\mathbf{v}$ jelöli, akkor fönn kell állni

$$\mathbb{Z}_{10}\text{-ben számolva az } (1,3,1,3,1,3,1,3,1,3,1,3,1)\cdot\mathbf{v} = 0$$

összefüggésnek. E képlet az ellenőrző összeg meghatározására is használható. 2007 óta a könyvek ISBN-száma (ISBN-13) megegyezik EAN-kódjával (1.41. ábra).

ISBN 978-963-545-398-6

9 789635 453986 >

1.41. ábra: Egy könyv ISBN kódja, ami egyúttal az EAN kódja is, a hozzá tartozó vonalkóddal

1.59. példa (Ellenőrző összeg). *Csak az ellenőrző összeget nézve érvényes személyi szám-e a 26012310018, és érvényes EAN-kód-e a 9998887776665?*

Megoldás. Ez a személyi szám érvényes, mert
$(1,2,3,4,5,6,7,8,9,10)\cdot(2,6,0,1,2,3,1,0,0,1) =$

$1 \cdot 2 + 2 \cdot 6 + 3 \cdot 0 + 4 \cdot 1 + 5 \cdot 2 + 6 \cdot 3 + 7 \cdot 1 + 8 \cdot 0 + 9 \cdot 0 + 10 \cdot 1 \bmod 11 = 63 \bmod 11 = 8$.

Ez az EAN-kód nem érvényes, mert $(1,3,1,3,1,3,1,3,1,3,1,3,1) \cdot (9,9,9,8,8,8,7,7,7,6,6,6,5) = (9 + 3 \cdot 9 + 9 + 3 \cdot 8 + 8 + 3 \cdot 8 + 7 + 3 \cdot 7 + 7 + 3 \cdot 6 + 6 + 3 \cdot 6 + 5) \bmod 10 = 183 \bmod 10 = 3 \neq 0$. $\square$

Feladatok

1.21. Az 1.1. (a) táblázatban egy 8 bináris kódszóból, a (b) táblázatban egy 9 kódszóból álló ternér kód kódszavai vannak felsorolva.

1.22. Mely vektorokra igaz, hogy minden bitjük a vektor maradék részének paritásbitje.

2

Lineáris egyenletrendszerek és megoldásuk

A lineáris egyenletrendszerek geometriai megközelítése után megismerjük a megoldás technikáit, végül a megoldások halmazának szerkezetét!

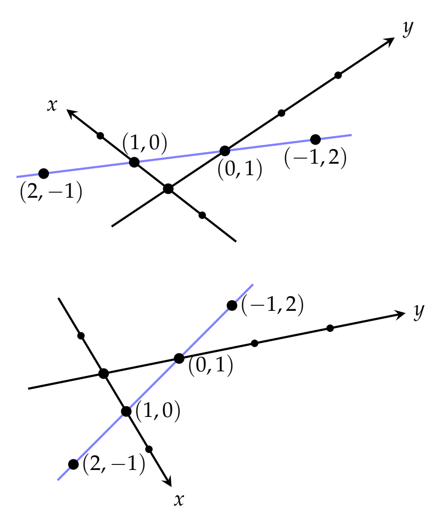
Egyenes és sík egyenletei

A sík és tér pontjainak és vektorainak koordinátáit használva lehetővé válik geometriai alakzatok algebrai vizsgálata, vagy algebrai problémák jobb megértése geometriai szemléltetéssel.

2.1. PÉLDA (AZ $x + y = 1$ EGYENLET). *Egy tetszőleges síkbeli koordinátarendszerben ábrázoljunk néhány pontot, melynek koordinátái kielégítik az $x + y = 1$ egyenletet. Fogalmazzunk meg sejtést az egyenlet összes megoldásának megfelelő pontok halmazáról!*

Alakzatok és egyenletek

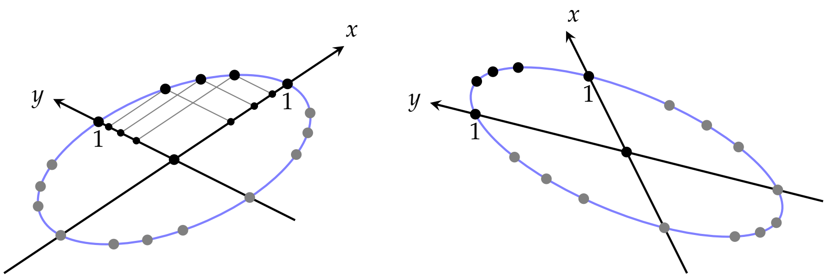
MEGOLDÁS. Az alábbi ábrán két különböző koordinátarendszert ábrázolunk, és azokban a fenti egyenletet kielégítő pontok közül néhányat. Ennek alapján azt sejthetjük, hogy az $x + y = 1$ egyenletet kielégítő pontok egy egyenesen vannak. Ezt az egyenest is berajzoltuk. A sejtést hamarosan bizonyítjuk. □



2.1. ábra: Az $x + y = 1$ egyenletet kielégítő néhány pont két különböző koordinátarendszerben.

2.2. PÉLDA (AZ $x^2 + y^2 = 1$ EGYENLET). *Egy tetszőleges síkbeli koordinátarendszerben ábrázoljunk néhány pontot, melynek koordinátái kielégítik az $x^2 + y^2 = 1$ egyenletet. Fogalmazzunk meg sejtést az egyenlet összes megoldásának megfelelő pontok halmazáról!*

MEGOLDÁS. Az alábbi ábrán néhány koordinátarendszert ábrázoltunk, az $x^2 + y^2 = 1$ egyenletet kielégítő néhány ponttal. A ??? fejezetben visszatérünk e feladatra, és meg fogjuk mutatni, hogy az egyenletet kielégítő pontok egy ellipszisen vannak. □



2.2. ábra: Az $x^2 + y^2 = 1$ egyenletet kielégítő (x, y) pontok halmaza két koordinátarendszerben.

A latin eredetű *implicit* szó jelentése *nem kifejtett*, *rejtett*, ami az összeköt, összefügg, összekever, körülcsavar jelentésű *implico* (implicō) szó származéka. E szó a matematikában az implicit alak, implicit függvény, stb. kifejezésekben arra utal, hogy valamely fontosnak tekintett mennyiség, változó, stb. nincs kifejezve a képletből. Ugyanennek a szónak a származéka a magába foglal, maga után von jelentésű *implikál* szó is, mely a matematikai logika „ha..., akkor...” szerkezetű műveletével, az *implikációval* is kapcsolatban van.

2.3. definíció (Alakzat (implicit) egyenletrendszere). *Egy geometriai alakzat egy adott koordinátarendszerre vonatkozó* (implicit) egyenletrendszerén *olyan egyenletrendszert értünk, melynek egyszerre minden egyenletét kielégítik a térnek az alakzathoz tartozó pontjai, de más pontok nem. Ha az egyenletrendszer egy egyenletből áll, az alakzat* egyenletéről *beszélünk. Az egyenletet* vektoregyenletnek *nevezzük, ha nem a pontok koordinátáira, hanem a pontokba mutató vektorokra írjuk fel. Egy alakzat m egyenletből álló egyenletrendszerének, illetve m vektoregyenletből álló egyenletrendszerének általános alakja*

$$\begin{cases} F_1(x_1, x_2, \ldots, x_n) = 0 \\ F_2(x_1, x_2, \ldots, x_n) = 0 \\ \qquad \vdots \\ F_m(x_1, x_2, \ldots, x_n) = 0 \end{cases} \quad \textit{illetve} \quad \begin{cases} F_1(\mathbf{r}) = 0 \\ F_2(\mathbf{r}) = 0 \\ \quad \vdots \\ F_m(\mathbf{r}) = 0 \end{cases}$$

ahol $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ a tér egy pontja, és $\mathbf{r}$ az oda mutató vektor.

Középiskolai tanulmányainkban több példát láttunk alakzat egyenletére, például tudjuk, hogy a síkban a koordinátatengelyek szögét felező egyenes egyenlete $y = x$, azaz $x - y = 0$. Ortonormált bázist választva az origó középű egységsugarú kör egyenlete $x^2 + y^2 = 1$. Az előző két egyenlet mindegyikéből *kifejezhető* a két koordináta egy paraméter bevezetésével. Az $y = x$ egyenlet ekvivalens az

$$\begin{aligned} x &= t \\ y &= t \end{aligned}$$

egyenletrendszerrel, míg az $x^2 + y^2 = 1$ egyenlet ekvivalens az

$$\begin{aligned} x &= \cos t \\ y &= \sin t \end{aligned}$$

egyenlettel. Mindkettő átírható vektoralakba is. Használjuk a oszlopvektoros jelölést:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} t \\ t \end{bmatrix}, \quad t \in \mathbb{R}, \quad \text{illetve} \quad \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos t \\ \sin t \end{bmatrix}, \quad t \in [0, 2\pi) \subseteq \mathbb{R}.$$

E két példa vezet a következő általános fogalomhoz.

A latin eredetű *explicit* szó jelentése *kifejtett, világosan kimondott*, ami a kibont, szétterít, kiszabadít, átvitt értelemben tisztáz, kifejt, megfejt jelentésű *explico* (explicō) szó származéka. E szó a matematikában az explicit alak, explicit függvény, stb. kifejezésekben arra utal, hogy valamely fontosnak tekintett mennyiség, változó, stb. ki van fejezve a többi segítségével.

2.4. definíció (Alakzat (explicit) egyenletrendszere). *Egy geometriai alakzat egy adott koordinátarendszerre vonatkozó* (explicit) egyenletrendszerén *olyan egyenletrendszert értünk, melyben az egyenletek bal oldalán a pontok koordinátáit megadó változók, jobb oldalán adott paraméterek függvényei szerepelnek. Általános alakja*

$$\begin{aligned} x_1 &= f_1(t_1, t_2, \dots, t_k) \\ x_2 &= f_2(t_1, t_2, \dots, t_k) \\ &\vdots \\ x_n &= f_n(t_1, t_2, \dots, t_k) \end{aligned}$$

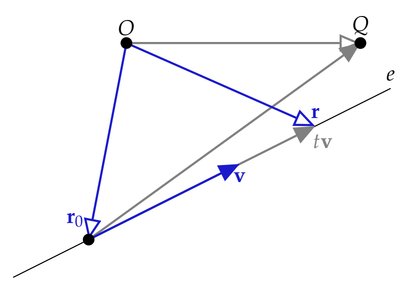
ahol $t_1 \in I_1$, $t_2 \in I_2, \dots, t_n \in I_n$, és $I_1, \dots, I_n \subseteq \mathbb{R}$. Az ilyen egyenletrendszer egyetlen vektoregyenletté fogható össze:

$$\mathbf{r} = \mathbf{f}(t_1, t_2, \dots, t_k),$$

ahol $\mathbf{f}$ *egy $\mathbb{R}^k \to \mathbb{R}^n$ függvény. Az explicit egyenletrendszereket szokás* paraméteres egyenletrendszernek *is nevezni.*

A következő paragrafusokban egyenes és sík különböző egyenleteit, egyenletrendszereit fogjuk áttekinteni példákat adva a fenti két általános definícióra.

Síkbeli egyenes egyenletei Tekintsük a sík egy tetszőleges e egyenesét, és jelöljük ki a síkban az O origót. Legyen a nemzérus $\mathbf{v}$ egy tetszőleges, az egyenessel párhuzamos vektor. Az ilyen vektorokat az egyenes *irányvektorának* nevezzük. Mutasson $\mathbf{r}_0$ az egyenes egy tetszőleges, kijelölt pontjába. Világos, hogy az e egyenes bármely pontjába mutató $\mathbf{r}$ vektor előáll $\mathbf{r}_0 + t\mathbf{v}$ alakban, ahol t valós szám. Másrészt ha Q a sík egy tetszőleges, nem az e egyenesre eső pontja, akkor az $\overrightarrow{OQ} - \mathbf{r}_0$ vektor nem párhuzamos $\mathbf{v}$-vel, tehát nem is konstansszorosa, azaz $\overrightarrow{OQ} - \mathbf{r}_0 \neq t\mathbf{v}$ semmilyen t-re sem, így $\overrightarrow{OQ}$ nem áll elő $\mathbf{r}_0 + t\mathbf{v}$ alakban. Tehát az e tetszőleges pontjába mutató $\mathbf{r}$ vektor felírható $\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}$ alakban, és ez csak e pontjaira igaz.



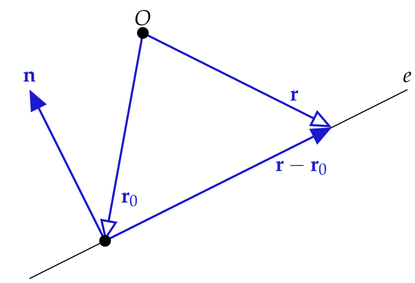
2.3. ábra: Egyenes explicit vektoregyenlete: $\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}$.

2.5. állítás (Síkbeli egyenes explicit vektoregyenlete). *A sík minden egyenesének van*

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{v} \tag{2.1}$$

alakú vektoregyenlete, és minden ilyen alakú egyenlet egy egyenes egyenlete, ahol $\mathbf{v}$ *az egyenes egy irányvektora, és* $\mathbf{r}_0$ *az egyenes egy tetszőleges, de rögzített pontjába mutató vektor.*

A síkbeli egyenesre merőleges vektorokat az egyenes *normálvektorainak* nevezzük. Legyen $\mathbf{n}$ egy tetszőleges, a $\mathbf{v}$-re merőleges vektor, azaz legyen $\mathbf{n}$ az e egy normálvektora. Azt, hogy $\mathbf{r} - \mathbf{r}_0$ az e tetszőleges pontjába mutató $\mathbf{r}$ vektorra párhuzamos $\mathbf{v}$-vel, úgy is kifejezhetjük, hogy $\mathbf{r} - \mathbf{r}_0$ merőleges $\mathbf{n}$-re. A merőlegesség pedig kifejezhető a skaláris szorzattal. Így az egyenes egy implicit vektoregyenletéhez jutunk: $\mathbf{r}$ pontosan akkor mutat az e egy pontjába, ha $\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0$. Ez az egyenlet átrendezés után $\mathbf{n} \cdot \mathbf{r} = \mathbf{n} \cdot \mathbf{r}_0$ alakra, majd a $C = \mathbf{n} \cdot \mathbf{r}_0$ jelölés bevezetésével $\mathbf{n} \cdot \mathbf{r} = C$ alakra hozható.



2.4. ábra: Síkbeli egyenes implicit vektoregyenlete: $\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0$.

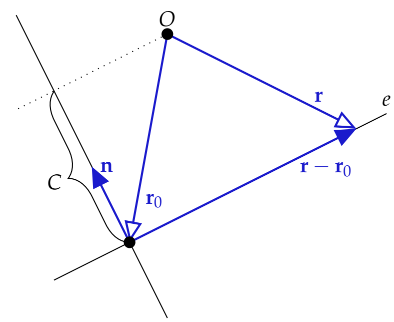
2.6. ÁLLÍTÁS (SÍKBELI EGYENES IMPLICIT VEKTOREGYENLETE). *A sík minden egyenesének van*

$$\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0, \tag{2.2}$$

és vele ekvivalens

$$\mathbf{n} \cdot \mathbf{r} = C \tag{2.3}$$

alakú vektoregyenlete, és minden ilyen alakú egyenlet egy egyenes egyenlete, ahol $\mathbf{n}$ *az egyenes egy normálvektora,* $\mathbf{r}_0$ *az egyenes egy tetszőleges, de rögzített pontjába mutató vektor és* C *konstans.*



2.5. ábra: Síkbeli egyenes (implicit) vektoregyenlete: $\mathbf{n} \cdot \mathbf{r} = C$. Ha $\mathbf{n}$ *egységvektor*, akkor az $\mathbf{n} \cdot \mathbf{r} = C$ geometriai jelentése az, hogy az egyenes bármely pontjába mutató vektornak az $\mathbf{n}$ egyenesére eső merőleges vetülete C. Ez az ábra is ezt az esetet szemlélteti.

A (2.2) alakú egyenlet könnyen átírható (2.3) alakúvá a $C = \mathbf{n} \cdot \mathbf{r}_0$ jelöléssel. Az átalakítás fordított irányban is egyszerű, hisz ha $\mathbf{n} \cdot \mathbf{r} = C$, akkor találunk olyan $\mathbf{r}_0$ vektort, melyre $\mathbf{n} \cdot \mathbf{r}_0 = C$. Ez azért igaz, mert ha tetszőleges $\mathbf{n}$-re nem merőleges $\mathbf{v}$ vektorra $\mathbf{n} \cdot \mathbf{v} = D$, akkor $\mathbf{n} \cdot (\frac{C}{D}\mathbf{v}) = C$, így az $\mathbf{r}_0 = \frac{C}{D}\mathbf{v}$ megfelel.

Az $\mathbf{r} = (x, y)$, $\mathbf{r}_0 = (x_0, y_0)$ és $\mathbf{v} = (a, b)$ jelöléseket használva az explicit vektoregyenlet azonnal egyenletrendszerré alakítható.

2.7. ÁLLÍTÁS (SÍKBELI EGYENES EXPLICIT EGYENLETRENDSZERE). *A sík minden egyenesének van*

$$\begin{aligned} x &= x_0 + at \\ y &= y_0 + bt \end{aligned} \tag{2.4}$$

alakú egyenletrendszere, ahol (a, b) *az egyenes egy irányvektora, és* (x_0, y_0) *az egyenes egy tetszőleges rögzített pontja.*

A következőkben megmutatjuk, hogy az explicit egyenletrendszerből a t paraméter kiküszöbölhető, és így egy implicit egyenletet kapunk.

2.8. ÁLLÍTÁS (SÍKBELI EGYENES (IMPLICIT) EGYENLETE). *A sík minden* *egyenesének van*

$$Ax + By = C \tag{2.5}$$

alakú egyenlete, és minden ilyen alakú egyenlet egy egyenes egyenlete, ahol A és B közül nem mindkettő nulla, és $(-B, A)$ az egyenes egy irányvektora.

A bizonyítás előtt érdemes megjegyezni, hogy az egyenes fenti implicit egyenlete az egyenes $\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0$ alakú vektoregyenletéből azonnal megkapható, de ezt egyelőre csak *ortonormált* koordinátarendszerben tudjuk könnyen igazolni. Legyen $(A, B) = (b, -a)$. Ez az egyenes egy normálvektora, hisz merőleges az (a, b) irányvektorra. Továbbá $\mathbf{r} - \mathbf{r}_0 = (x - x_0, y - y_0)$, ezért a vektoregyenlet

$$(A, B) \cdot (x - x_0, y - y_0) = 0$$

alakú lesz, ami a skaláris szorzást elvégezve az $A(x - x_0) + B(y - y_0) = 0$ formulát adja. Ha a koordinátarendszer nem ortonormált, az (A, B) vektor nem szükségképpen normálvektor, és a skaláris szorzás képlete is más, de azért a (2.5) egyenletről mondott állítás igaz. A ??. fejezetben tanultak alapján az egyenes egyenletét a vektoregyenletből majd nem ortonormált koordinátarendszer esetén is le tudjuk vezetni, most viszont egy olyan bizonyítást adunk, mely az explicit egyenletrendszerre épül.

BIZONYÍTÁS. Ha a vagy b valamelyike 0, akkor a két egyenlet egyike felesleges, például ha $a = 0$, akkor az egyenletrendszer alakja

$$\begin{aligned} x &= x_0 \\ y &= y_0 + bt \end{aligned}$$

ami ekvivalens az $x = x_0$ egyenlettel, hisz az $y = y_0 + bt$ semmi mást nem mond, mint hogy y egy valós szám. Mivel $(a, b) \neq (0, 0)$, ezért csak az az eset marad, amikor a és b egyike sem 0. Ekkor mindkét egyenletből kifejezhető t, és a két értéket egyenlővé téve kapjuk, hogy

$$\frac{x - x_0}{a} = \frac{y - y_0}{b},$$

azaz

$$bx - ay = bx_0 - ay_0, \text{ vagy } b(x - x_0) - a(y - y_0) = 0.$$

Legyen a továbbiakban $A = b$ és $B = -a$. Ekkor a fenti egyenlet $Ax + By = Ax_0 + By_0$ lesz. Az egyenlet jobb oldalán lévő konstanst C-vel jelölve az egyenes egyenlete $Ax + By = C$ alakot ölt. Másrészt könnyen látható, hogy minden ilyen alakú egyenlet egy egyenes egyenlete, mert ekvivalens egy egyenes paraméteres egyenletrendszerével. Nevezetesen az $Ax + By = C$ egyenlet visszaírható $Ax + By = Ax_0 + By_0$ alakba, hisz az $Ax_0 + By_0 = C$ egyenletben $A \neq 0$ esetén egy tetszőleges y_0-t választva, egyértelműen kifejezhető x_0. (A $B \neq 0$ eset analóg.) Ennek alapján felírható a (2.4) egyenletrendszer. □

2.9. példa (Síkbeli egyenes egyenletei). *Írjuk fel annak az egyenesnek összes egyenletét vagy egyenletrendszerét, mely átmegy a $(2,3)$ és az $(1,1)$ koordinátájú pontokon.*

Megoldás. Ha egy egyenes átmegy e két ponton, akkor irányvektora a két pontba mutató vektorok különbsége, azaz $\mathbf{v} = (2,3) - (1,1) = (1,2)$. Legyen $\mathbf{r}_0 = (1,1)$, de az $\mathbf{r}_0 = (2,3)$ választás is megfelelő.

Az irányvektor segítségével kapjuk, hogy

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Explicit (paraméteres) egyenletrendszer alakban:

$$\begin{aligned} x &= 1 + t \\ y &= 1 + 2t. \end{aligned}$$

Az irányvektorból $(A, B) = (2, -1)$, innen az egyenes egyenlete $2x - y = 2 \cdot 1 - 1 \cdot 1$, azaz

$$2x - y = 1.$$

Ortonormált koordináta-rendszerben a

$$(2, -1) \cdot (x - 1, y - 1) = 0$$

egyenletet kapjuk vektoregyenletként, mely kiszámolva az előző egyenletet adja. □

Síkbeli pont egyenletei Tekintsük a síkbeli (x_0, y_0) pontot. Ennek explicit egyenletrendszere, illetve vektoregyenlete:

$$\begin{aligned} x &= x_0 \\ y &= y_0 \end{aligned} \quad \text{illetve} \quad \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}.$$

Ez annyira nyilvánvaló, semmitmondó, hogy a gyakorlatban nem is szoktunk pont egyenleteiről beszélni, e könyvbe is csak didaktikai okokból került, ugyanis a matematikai fogalmak megértésében gyakran nagy segítségünkre van a szélső, extremális esetek megértése, vizsgálata.

Mivel itt az alakzat csak egyetlen pontból áll, nincs szükség paraméterre, így ez az alak egyúttal implicitnek is tekinthető. Ekkor úgy tekintünk ugyanerre az egyenletrendszerre, mint két egyenes egyenletére, melyek normálvektorai $(1,0)$ illetve $(0,1)$, és amelyek metszéspontja a keresett pont:

$$\begin{aligned} x &= x_0 \\ y &= y_0 \end{aligned} \quad \text{vagy minden együtthatót kiírva} \quad \begin{aligned} x + 0y &= x_0 \\ 0x + y &= y_0 \end{aligned}$$

Ez adja az ötletet, egy pont implicit egyenletrendszerének tekinthetnénk két egyenletet, melyek egymást az adott pontban metsző egy-egy egyenes egyenletei. Tehát mondhatjuk, hogy a pont implicit egyenletrendszerének általános alakja:

$$\begin{aligned} A_1x + B_1y &= C_1 \\ A_2x + B_2y &= C_2 \end{aligned}$$

Az azonban itt nem igaz, hogy minden ilyen alakú egyenletrendszer egy pont egyenletrendszere, mert két egyenes metszheti egymást egyetlen pontban, de lehet, hogy nincs közös pontjuk, és lehet végtelen sok közös pontjuk is. Épp ennek a kérdésnek a részletes vizsgálata lesz a 2. fejezet témája.

A 3-dimenziós tér síkjainak egyenletei Tudjuk, hogy két lineárisan független $\mathbf{u}$ és $\mathbf{v}$ vektor bármely lineáris kombinációja a két vektor által meghatározott síkban van, továbbá hogy e sík bármely vektora előáll a megadott két vektor lineáris kombinációjaként (ld. 1.8.. és 1.11. tételek). Ebből azonnal adódik, hogy a sík egy rögzített pontjába mutató $\mathbf{r}_0$ vektor segítségével a sík bármelyik pontjába mutató $\mathbf{r}$ vektor felírható $\mathbf{r} = \mathbf{r}_0 + s\mathbf{u} + t\mathbf{v}$ alakban.

2.10. Állítás (Sík explicit vektoregyenlete). *Bármely síknak van*

$$\mathbf{r} = \mathbf{r}_0 + s\mathbf{u} + t\mathbf{v} \tag{2.6}$$

alakú vektoregyenlete, és minden ilyen alakú egyenlet egy sík egyenlete, ahol $\mathbf{u}$ *és* $\mathbf{v}$ *a sík két lineárisan független vektora, és* $\mathbf{r}_0$ *a sík egy tetszőleges, de rögzített pontjába mutató vektor.*

Hasonlóan a síkbeli egyeneshez, a térbeli sík egyenletéből is kiküszöbölhető a paraméter a merőlegesség felhasználásával. Az 1.14. feladat állítása szerint, ha egy vektor merőleges két tetszőleges vektor mindegyikére, akkor merőleges azok lineáris kombinációjára is. Mivel az $\mathbf{n} = \mathbf{u} \times \mathbf{v}$ merőleges $\mathbf{u}$-ra és $\mathbf{v}$-re is, ezért merőleges azok minden lineáris kombinációjára is, azaz az $\mathbf{r} - \mathbf{r}_0 = s\mathbf{u} + t\mathbf{v}$ vektorra is. Ez az észrevétel az alapja az alábbi tételnek.

2.11. Állítás (Sík implicit vektoregyenlete). *A háromdimenziós térben minden síknak van*

$$\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0, \tag{2.7}$$

és a vele ekvivalens

$$\mathbf{n} \cdot \mathbf{r} = C \tag{2.8}$$

alakú vektoregyenlete, és minden ilyen alakú egyenlet egy sík egyenlete, ahol $\mathbf{n}$ *a sík egy normálvektora,* $\mathbf{r}_0$ *a sík egy tetszőleges, de rögzített pontjába mutató vektor és C konstans.*

Az állítás igazolása analóg a síkbeli egyenesnél leírtakkal.

Az $\mathbf{r} = (x, y, z)$, $\mathbf{r}_0 = (x_0, y_0, z_0)$ és $\mathbf{u} = (a_1, b_1, c_1)$ $\mathbf{v} = (a_2, b_2, c_2)$ jelöléseket használva az explicit vektoregyenlet azonnal egyenletrendszerré alakítható.

2.12. ÁLLÍTÁS (SÍK EXPLICIT EGYENLETRENDSZERE). *A háromdimenziós tér minden síkjának van*

$$\begin{aligned} x &= x_0 + a_1 s + a_2 t \\ y &= y_0 + b_1 s + b_2 t \\ z &= z_0 + c_1 s + c_2 t \end{aligned} \tag{2.9}$$

alakú egyenletrendszere, ahol (a_1, b_1, c_1) *és* (a_2, b_2, c_2) *a sík két lineárisan független vektora, és* (x_0, y_0, z_0) *a sík egy tetszőleges rögzített pontja.*

Az explicit egyenletrendszerből kiküszöbölhető a két paraméter, ha például két egyenletből kifejezzük a paramétereket, és behelyettesítjük a harmadik egyenletbe. Így egy implicit egyenletet kapunk. A számításokat nem részletezzük, az eredmény

$$(b_1c_2 - b_2c_1)(x - x_0) + (c_1a_2 - c_2a_1)(y - y_0) + (a_1b_2 - b_1a_2)(z - z_0) = 0.$$

Az $(A, B, C) = (b_1c_2 - b_2c_1, c_1a_2 - c_2a_1, a_1b_2 - a_2b_1)$ jelöléssel a sík egyenlete $A(x - x_0) + B(y - y_0) + C(z - z_0) = 0$ alakra hozható, vagy ami vele ekvivalens, $Ax + By + Cz = D$ alakra.

2.13. ÁLLÍTÁS (SÍK IMPLICIT EGYENLETE). *A háromdimenziós térben minden síknak van*

$$Ax + By + Cz = D \tag{2.10}$$

alakú egyenlete, és minden ilyen alakú egyenlet egy sík egyenlete, ha A, B *és* C *legalább egyike nem nulla, és* $D = Ax_0 + By_0 + Cz_0$, *ahol* (x_0, y_0, z_0) *a sík valamely pontja.*

A sík fenti egyenlete a sík $\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0$ alakú vektoregyenletéből is megkapható, de ezt egyelőre csak *ortonormált* koordinátarendszerben tudjuk könnyen igazolni. Mivel

$$(A, B, C) = (b_1c_2 - b_2c_1, c_1a_2 - c_2a_1, a_1b_2 - a_2b_1), \tag{2.11}$$

ami épp az $\mathbf{u} \times \mathbf{v}$ vektorral egyenlő, ezért (A, B, C) merőleges a sík minden vektorára, vagyis a sík egy normálvektora. Az $\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0$ egyenletet koordinátás alakba átírva kapjuk, hogy

$$(A, B, C) \cdot (x - x_0, y - y_0, z - z_0) = 0.$$

2.14. PÉLDA (SÍK EGYENLETEI). *Írjuk fel annak a síknak az egyenleteit, mely átmegy a* $(0, -1, 2)$, *a* $(-1, 0, 7)$ *és a* $(2, 1, 4)$ *pontokon.*

Megoldás. A három pontba mutató vektorok különbségei a síkkal párhuzamos vektorok, így azokkal felírható a sík mindegyik egyenlete. Két vektor a lehetséges háromból:

$$\begin{aligned}\mathbf{u} &= (2,1,4) - (0,-1,2) = (2,2,2), \text{ és}\\ \mathbf{v} &= (-1,0,7) - (0,-1,2) = (-1,1,5).\end{aligned}$$

Ezek alapján például az $\mathbf{r}_0 = (0,-1,2)$ választás mellett a sík explicit vektoregyenlete

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix} + s \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} + t \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix},$$

explicit egyenletrendszere

$$\begin{aligned} x &= \phantom{-1+{}} 2s - t \\ y &= -1 + 2s + t \\ z &= 2 + 2s + 5t \end{aligned}$$

Mivel a (2.11) képlet szerint $(A,B,C) = (8,-12,4)$, ezért a sík implicit egyenlete $8(x-0) - 12(y-(-1)) + 4(z-2) = 0$, azaz 4-gyel való osztás és átrendezés után

$$2x - 3y + z = 5.$$

Így ortonormált koordinátarendszerben a

$$(2,-3,1)\cdot(x,y,z) = 5, \text{ vagy } (2,-3,1)\cdot(x,y+1,z-2) = 0$$

a sík implicit vektoregyenlet alakja. □

Térbeli egyenes egyenletei Mindaz, amit a síkbeli egyenes explicit vektoregyenletéről mondtunk a 71. oldalon, lényegében változtatás nélkül megismételhető. Jelöljük ki a térben az origót, és tekintsük azt az e egyenest, melynek irányvektora $\mathbf{v}$, és amely átmegy azon a ponton, melybe az $\mathbf{r}_0$ vektor mutat. Világos, hogy az e egyenes bármely pontjába mutató $\mathbf{r}$ vektor előáll $\mathbf{r}_0 + t\mathbf{v}$ alakban, ahol t valós szám, és az e-re nem illeszkedő pontokra ez nem áll. Így igaz a következő állítás:

2.15. állítás (Térbeli egyenes explicit vektoregyenlete). *A háromdimenziós tér minden egyenesének van*

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{v} \tag{2.12}$$

alakú vektoregyenlete, és minden ilyen alakú egyenlet egy egyenes egyenlete, ahol $\mathbf{v}$ az egyenes egy irányvektora, és $\mathbf{r}_0$ egy tetszőleges, de rögzített pontjába mutató vektor.

Itt nem tudjuk a paramétert egyetlen vektoregyenletben kiküszöbölni, de az explicit egyenletrendszerré való átírás megy, ha felveszünk egy koordinátarendszert, melyben $\mathbf{r} = (x, y, z)$, $\mathbf{r}_0 = (x_0, y_0, z_0)$ és $\mathbf{v} = (a, b, c)$:

2.16. ÁLLÍTÁS (TÉRBELI EGYENES EXPLICIT EGYENLETRENDSZERE). *A tér minden egyenesének van*

$$\begin{aligned} x &= x_0 + at \\ y &= y_0 + bt \\ z &= z_0 + ct \end{aligned} \tag{2.13}$$

alakú egyenletrendszere, ahol (a, b, c) *az egyenes egy irányvektora, és* (x_0, y_0, z_0) *az egyenes egy tetszőleges rögzített pontja.*

A fenti explicit (paraméteres) egyenletrendszerből a paraméter kiküszöbölhető. Ha az a, b és c számok valamelyike 0, akkor a neki megfelelő fenti egyenletben már nem szerepel t, akkor nincs is mit tennünk. Ha legalább két egyenletben szerepel t, akkor mindegyikből kifejezve t-t, majd egyenlővé téve őket paraméter nélküli egyenleteket kapunk. Például ha a, b és c egyike sem 0, akkor

$$t = \frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}.$$

A t-t elhagyva valójában három egyenletet kaptunk:

$$\frac{x - x_0}{a} = \frac{y - y_0}{b}, \qquad \frac{x - x_0}{a} = \frac{z - z_0}{c}, \qquad \frac{y - y_0}{b} = \frac{z - z_0}{c}.$$

Annak az egyenletnek nincs értelme, amelyikben a nevező 0, de a nevezőkkel való bővítés után kapott

$$b(x - x_0) = a(y - y_0), \quad c(x - x_0) = a(z - z_0), \quad c(y - y_0) = b(z - z_0)$$

egyenletek mindegyike korrekt akkor is, ha 0 valamelyik együttható. E három egyenlet három sík egyenlete, melyek metszésvonala az adott egyenes. Kivétel az az eset, amikor az egyik egyenlet $0 = 0$ alakú, ilyenkor a másik két egyenlet egy-egy sík egyenlete. Egy egyenes azonban megadható két sík metszésvonalaként, így adódik a következő tétel, melynek bizonyítását feladatként tűzzük ki:

A három síkból azonban már kettő is meghatározza az egyenest, így két egyenletből álló egyenletrendszerrel is megadható az egyenes. Bizonyítható az alábbi állítás:

2.17. ÁLLÍTÁS (TÉRBELI EGYENES IMPLICIT EGYENLETRENDSZERE). *A tér minden egyenesének van két egyenletből álló egyenletrendszere. Ha az egyenes egy irányvektora* (a, b, c), *akkor a két egyenlet az alábbi három közül* *bármelyik kettő, amelyik nem* $0 = 0$ *alakú:*

$$\begin{aligned} b(x-x_0) &= a(y-y_0) \\ c(x-x_0) &= a(z-z_0) \\ c(y-y_0) &= b(z-z_0) \end{aligned} \tag{2.14}$$

► A (2.14) egyenletrendszer a következő alakba is átírható:

$$\begin{array}{rrrlrr} bx & -ay & & = bx_0 & -ay_0 & \\ cx & & -az & = cx_0 & & -az_0 \\ & cy & -bz & = & cy_0 & -bz_0 \end{array} \tag{2.15}$$

Erről könnyen leolvasható, hogy ha pl. $a \neq 0$, akkor a második egyenlet b-szereséből kivonva az első egyenlet c-szeresét, a harmadik egyenlet a-szorosát kapjuk. Hamarosan látni fogjuk, hogy eszerint a harmadik egyenlet elhagyható, anélkül, hogy az egyenletrendszert kielégítő pontok halmaza megváltozna.

2.18. PÉLDA (TÉRBELI EGYENES EGYENLETRENDSZEREI). *Írjuk fel annak az egyenesnek az explicit és implicit egyenletrendszerét, mely átmegy az* $A(1,3,4)$ *és az a)* $B(3,3,1)$ *b)* $B(5,5,-2)$ *ponton.*

MEGOLDÁS. *a)* A két pontot összekötő vektor $= (2,0,-3)$. Innen az egyenes explicit egyenletrendszere

$$\begin{aligned} x &= 1+2t \\ y &= 3 \\ z &= 4-3t \end{aligned}$$

második egyenlete $y = 3$ egy xz-síkkal párhuzamos sík egyenlete. A másik két egyenlet mindegyikéből kifejezve t-t, majd egyenlővé téve őket, egy másik sík egyenletét kapjuk. Az egyenes ennek a két síknak a metszésvonala. Az első egyenletből $t = \frac{1}{2}(x-1)$, a harmadikból $t = -\frac{1}{3}(z-4)$ hogy $3x + 2z = 11$. Így az előző egyeneshez a következő implicit (paraméter nélküli) egyenletrendszer tartozik, mely két sík egyenletéből áll:

$$\begin{array}{rrrl} 3x & & +2z & = 11 \\ & y & & = 3 \end{array}$$

b) A két pontot összekötő vektor itt $= (4,2,-6)$. Innen az egyenes explicit egyenletrendszere

$$\begin{aligned} x &= 1+4t \\ y &= 3+2t \\ z &= 4-6t \end{aligned}$$

Mindegyik egyenletből kifejezve t-t, és ezeket egyenlővé téve kapjuk, hogy
$$\frac{x-1}{4} = \frac{y-3}{2} = \frac{z-4}{-6}.$$
Ez a következő három sík egyenletét adja:
$$\frac{x-1}{4} = \frac{y-3}{2}, \quad \frac{z-4}{-6} = \frac{x-1}{4}, \quad \frac{y-3}{2} = \frac{z-4}{-6}.$$
Átrendezve
$$\begin{aligned} x - 2y \phantom{{}+2z} &= -5 \\ 3x \phantom{{}-2y} + 2z &= 11 \\ 3y + z &= 13 \end{aligned}$$
E három sík közül bármely kettő meghatározza az adott egyenest, így e három egyenlet közül bármely kettő az egyenes (implicit) egyenletrendszere. □

Térbeli pont egyenletei Csak a teljesség és az analógiák megértése céljából vizsgáljuk meg a tér egy pontjának lehetséges egyenleteit. A térbeli (x_0, y_0, z_0) pont explicit egyenletrendszere, illetve vektoregyenlete:
$$\begin{aligned} x &= x_0 \\ y &= y_0 \\ z &= z_0 \end{aligned} \quad \text{illetve} \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}.$$
Az explicit egyenletrendszert implicit alaknak is tekinthetjük, ekkor három – a koordinátasíkokkal párhuzamos – sík egyenletét látjuk, melyek egyetlen közös pontban metszik egymást.
$$\begin{aligned} x \phantom{{}+y+z} &= x_0 \\ y \phantom{{}+z} &= y_0 \\ z &= z_0 \end{aligned} \quad \text{vagy minden együtthatót kiírva} \quad \begin{aligned} x + 0y + 0z &= x_0 \\ 0x + y + 0z &= y_0 \\ 0x + 0y + z &= z_0 \end{aligned}$$
A síkbeli esethez hasonlóan egy pont implicit egyenletrendszerének tekinthetnénk három egyenletet, melyek egymást az adott pontban metsző egy-egy sík egyenletei. Tehát mondhatjuk, hogy a pont implicit egyenletrendszerének általános alakja:
$$\begin{aligned} A_1x + B_1y + C_1z &= D_1 \\ A_2x + B_2y + C_2z &= D_2 \\ A_3x + B_3y + C_3z &= D_3 \end{aligned}$$
Itt is óvatosnak kell lennünk, mert nem minden ilyen alakú egyenletrendszer egy pont egyenletrendszere. Például három sík metszheti egymást egy egyenesben, de párhuzamos síkok esetén az is előfordulhat, hogy nincs közös pontjuk. E kérdés vizsgálatára visszatérünk a 2. fejezetben.

Egyenletek $\mathbb{R}^n$-ben Az egyenes és a sík explicit vektoregyenlete $\mathbb{R}^n$-ben is ugyanolyan alakú, mint $\mathbb{R}^3$-ben, azaz az egyenes explicit vektoregyenlete $\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}$, a síké $\mathbf{r} = \mathbf{r}_0 + s\mathbf{u} + t\mathbf{v}$ alakú.

2.19. PÉLDA (EGYENES ÉS SÍK EXPLICIT VEKTOREGYENLETE). *Írjuk fel az $A(1,1,1,1)$, $B(2,3,2,4)$ pontokon átmenő egyenes, és az A, B és $C(3,2,1,0)$ pontokon átmenő sík explicit vektoregyenletét!*

MEGOLDÁS. Az $\overrightarrow{AB} = (1,2,1,3)$ és az $\overrightarrow{AC} = (2,1,0,-1)$ vektorok segítségével azonnal fölírható az egyenes és a sík egyenlete is:

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 2 \\ 1 \\ 3 \end{bmatrix}, \text{ illetve } \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + s \begin{bmatrix} 1 \\ 2 \\ 1 \\ 3 \end{bmatrix} + t \begin{bmatrix} 2 \\ 1 \\ 0 \\ -1 \end{bmatrix}.$$ □

A síkbeli egyenes és a térbeli sík vektoregyenlete $\mathbf{n} \cdot \mathbf{r} = c$ alakú. E két esetben ez az egyenlet az n-dimenziós tér egy $n-1$-dimenziós alakzatának egyenlete ($n = 2,3$). A későbbiekben látni fogjuk, hogy ez általában is igaz, de e pillanatban még a dimenzió fogalmát sem definiáltuk, ezért egyelőre csak nevet adunk ennek az alakzatnak. Az $\mathbb{R}^n$ térben az $\mathbf{n} \cdot \mathbf{r} = c$ egyenletet kielégítő $\mathbf{r}$ vektorok végpontjainak halmazát *hipersíknak* nevezzük. Koordinátás alakban

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = c,$$

ahol $\mathbf{n} = (a_1, a_2, \ldots, a_n)$, $\mathbf{r} = (x_1, x_2, \ldots, x_n)$.

2.20. PÉLDA (HIPERSÍK EGYENLETE). *Mutassuk meg, hogy az*

$$x + 2y + 3z + 6w = 12$$

egyenletű hipersík bármely két pontját összekötő vektor merőleges az $(1,2,3,6)$ *vektorra, azaz vele való skaláris szorzata* 0.

MEGOLDÁS. Ha az (x_1, y_1, z_1, w_1) és az (x_2, y_2, z_2, w_2) pontok a megadott hipersíkon vannak, akkor

$$\begin{aligned} x_1 + 2y_1 + 3z_1 + 6w_1 &= 12 \\ x_2 + 2y_2 + 3z_2 + 6w_2 &= 12 \end{aligned}$$

Ha a két egyenletet kivonjuk egymásból, akkor az

$$(x_1 - x_2) + 2(y_1 - y_2) + 3(z_1 - z_2) + 6(w_1 - w_2) = 0$$

egyenlethez jutunk, amely skalárszorzat alakban

$$(1,2,3,6) \cdot (x_1 - x_2, y_1 - y_2, z_1 - z_2, w_1 - w_2) = 0,$$

ami épp azt jelenti, hogy a két pontot összekötő vektor merőleges az $(1,2,3,6)$ vektorra. □

A következő táblázat összefoglalja geometriai alakzatoknak a további szempontjából legfontosabb egyenleteit. Az $\mathbb{R}^n$-beli egyenletek közül többet még nem ismerünk, ezeket három kérdőjel jelzi, viszont arra bíztatjuk az Olvasót, hogy az analógia fonalán haladva fogalmazza meg sejtéseit.

		Explicit vektoregyenlet	Implicit egyenlet(rendszer)
Síkban	egyenes	$\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}$	$Ax + By = C$
	pont	$\mathbf{r} = \mathbf{r}_0$	$A_1x + B_1y = C_1$ $A_2x + B_2y = C_2$
Térben	sík	$\mathbf{r} = \mathbf{r}_0 + s\mathbf{u} + t\mathbf{v}$	$Ax + By + Cz = D$
	egyenes	$\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}$	$A_1x + B_1y + C_1z = D_1$ $A_2x + B_2y + C_2z = D_2$
	pont	$\mathbf{r} = \mathbf{r}_0$	$A_1x + B_1y + C_1z = D_1$ $A_2x + B_2y + C_2z = D_2$ $A_3x + B_3y + C_3z = D_3$
$\mathbb{R}^n$-ben	hipersík	???	$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$
	sík	$\mathbf{r} = \mathbf{r}_0 + s\mathbf{u} + t\mathbf{v}$	???
	egyenes	$\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}$	???
	pont	$\mathbf{r} = \mathbf{r}_0$	???

Feladatok

A lineáris egyenletrendszer és két modellje

E szakasz témája a lineáris egyenletrendszerek fogalma és a lineáris egyenletrendszer megoldásának két geometriai interpretációja: hipersíkok metszetének meghatározása és egy vektor lineáris kombinációként való előállítása. A számítások kényelmes könyvelésére bevezetjük a mátrix fogalmát.

Lineáris egyenlet és egyenletrendszer Az előző rész végén láttuk, hogy a síkbeli egyenes egyenletének általános alakja $Ax + By = C$, ahol A, B és C konstansok. Ennek általánosításaként jutunk a lineáris egyenlet fogalmához.[1]

[1] *Lineáris:* a *vonalas* jelentésű latin lineāris szóból ered, mely a *lenfonal*, *horgászzsinór*, átvitt értelemben *vonal*, *határvonal* jelentésű linea (līnea) szó származéka. A matematikában *egyenessel kapcsolatba hozható*, illetve *elsőfokú* értelemben szokás használni.

2.21. DEFINÍCIÓ (LINEÁRIS EGYENLET). *Az*

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \tag{2.16}$$

alakra hozható egyenletet az x_1, $x_2 \dots x_n$ ismeretlenekben lineáris egyenletnek *nevezzük, ahol a_1, a_2,… és a_n, valamint b konstansok. Az a_1, a_2,… és a_n konstansokat az egyenlet* együtthatóinak *b-t az egyenlet* konstans tagjának *nevezzük.*

2.22. PÉLDA (LINEÁRIS EGYENLET). *Az alábbi egyenletek lineárisak:*

$$x - 2y = 1, \quad \frac{1}{2}x_1 - \sqrt{2}x_2 + (5 - \pi)x_3 = 0, \quad a\cos 0.87 - 0.15c = 0.23.$$

A következő egyenletek nem lineárisak az x, y és z ismeretlenekben:

$$xz - y = 0, \quad x + 2y = 3^z, \quad x\sin z + y\cos z + y = z^2,$$

viszont mindegyikük lineáris az x és y ismeretlenekben, hisz ekkor z paraméter, melynek bármely értéke mellett lineárisak az egyenletek.

2.23. PÉLDA (LINEÁRIS EGYENLET AZONOS ÁTALAKÍTÁSA). *Az*

$$x = y, \quad x = 3 - y + 2z$$

egyenletek az x, y és z ismeretlenekben lineárisak, mert azonos átalakítással a definícióbeli alakra hozhatók:

$$x - y + 0z = 0, \quad x + y - 2z = 3.$$

Másrészt az

$$\frac{x}{z} + \frac{y}{z} + 2 = 0$$

nem lineáris, mert a z-vel való beszorzás nem azonos átalakítás, tehát a lineáris $x + y + 2z = 0$ egyenlettel nem ekvivalens.

Lineáris egyenletek egy halmazát *lineáris egyenletrendszernek* nevezzük. Az egyenletrendszer ismeretlenei mindazok az ismeretlenek, amelyek legalább egy egyenletben szerepelnek. Ha egy ismeretlen egy egyenletben nem szerepel, akkor úgy tekintjük, hogy 0 az együtthatója. A jobb áttekinthetőség kedvéért az egyenletrendszereket úgy írjuk fel, hogy az ismeretlenek mindegyik egyenletben ugyanabban a sorrendben szerepeljenek. Egy egyenletrendszer egy egyenletből is állhat.

2.24. PÉLDA (LINEÁRIS EGYENLETRENDSZEREK). *Lineáris egyenletrendszerek például a következők:*

$$\begin{aligned} 3x - y &= 2 \\ -x + 2y &= 6 \\ x + y &= 6 \end{aligned} \qquad \begin{aligned} x_1 &= 3 \\ x_2 &= 1 \\ x_3 &= 4 \end{aligned} \qquad 2x - 3y + z - w = 6. \tag{2.17}$$

Elképzelhető, hogy egy egyenletrendszer átalakítása közben olyan egyenletet kapunk, melyben minden együttható 0, azaz amely $0 = b$ alakú. Az is lehet, hogy egy egyenletrendszerben egyes együtthatók paraméterek. Ilyenkor tudnunk kell, mely változók az ismeretlenek, melyek a paraméterek. Így a következő egyenletrendszerek is lineárisak az x, y ismeretlenekben:

$$\begin{aligned} ax + \phantom{\frac{1}{a}}y &= 2a \\ x - \frac{1}{a}y &= 0 \end{aligned} \qquad \begin{aligned} 3x - y &= 2 \\ -x + 2y &= 6 \\ 0 &= 0 \end{aligned} \qquad \begin{aligned} x + y &= 1 \\ 0 &= 2. \end{aligned} \tag{2.18}$$

Paraméterek használatával felírható az összes olyan egyenletrendszer, mely adott számú egyenletből áll és adott számú ismeretlent tartalmaz. Például az összes 3-ismeretlenes, 2 egyenletből álló egyenletrendszer a következő alakú, illetve ilyenné alakítható:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \end{aligned}$$

2.25. DEFINÍCIÓ (LINEÁRIS EGYENLETRENDSZER). Lineáris egyenletrendszer*en ugyanazokban a változókban lineáris egyenletek egy véges halmazát értjük. Általános alakja m egyenlet és n ismeretlen esetén*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \vdots \qquad \vdots \qquad\quad \vdots \quad &\ \ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned} \tag{2.19}$$

ahol x_1, x_2,… x_n az ismeretlenek, a_{ij} az i-edik egyenletben az x_j ismeretlen együtthatóját jelöli, és b_i az i-edik egyenlet konstans tagja. Ha mindegyik *egyenlet konstans tagja* 0*, a lineáris egyenletrendszer* homogén*, ha csak egy is különbözik* 0*-tól* inhomogén*.*

2.26. DEFINÍCIÓ (LINEÁRIS EGYENLETRENDSZER MEGOLDÁSA). *Azt mondjuk, hogy a rendezett* $(u_1, u_2, \ldots, u_n)$ *szám-n-es* megoldása *a (2.19) egyenletrendszernek, ha megoldása minden egyenletnek, azaz minden egyenletet kielégít az* $x_1 = u_1$, $x_2 = u_2, \ldots$ $x_n = u_n$ *helyettesítéssel. Ha e szám-n-est vektornak tekintjük,* megoldásvektorról *beszélünk. Az összes megoldás halmazát az egyenletrendszer* megoldáshalmaz*ának nevezzük. Egy egyenletrendszert* megoldhatónak *vagy* konzisztensnek *nevezünk, ha van megoldása, azaz ha megoldáshalmaza nem üres. Ellenkező esetben az egyenletrendszer* nem megoldható *vagy* inkonzisztens*.*

2.27. PÉLDA (EGYENLETRENDSZER EGY MEGOLDÁSA). *Keressük meg a (2.17) és a (2.18) egyenletrendszereinek egy-egy megoldását!*

MEGOLDÁS. $(x, y) = (2, 4)$, $(x_1, x_2, x_3) = (3, 1, 4)$, $(x, y, z, w) = (2, 0, 2, 0)$, $(x, y) = (1, a)$, $(x, y) = (2, 4)$. A (2.17)-beli harmadik egyenletrendszernek végtelen sok megoldása van, például egy másik megoldás az $(x, y, z, w) = (3, 0, 0, 0)$. A (2.18) utolsó egyenletrendszerének nincs megoldása, hisz nincs olyan x és y érték, melyre fönnállna a $0x + 0y = 2$ egyenlőség. □

Általában, a

$$0x_1 + 0x_2 + \cdots + 0x_n = 0$$

egyenletnek minden szám-n-es megoldása, míg a

$$0x_1 + 0x_2 + \cdots + 0x_n = b, \quad (b \neq 0)$$

egyenletnek egyetlen megoldása sincs.

Ekvivalens lineáris egyenletrendszerek Tekintsük az alábbi három egyenletrendszert:

$$\begin{aligned} x + y &= 3 \\ x + 2y &= 4 \end{aligned} \qquad \begin{aligned} x + y &= 3 \\ y &= 1 \end{aligned} \qquad \begin{aligned} x &= 2 \\ y &= 1 \end{aligned} \tag{2.20}$$

Mindháromnak $(x, y) = (2, 1)$ az egyetlen megoldása.

2.28. DEFINÍCIÓ (EKVIVALENS EGYENLETRENDSZEREK). *Azonos ismeretlenekkel felírt két egyenletrendszert* ekvivalensnek *nevezünk, ha megoldásaik halmaza azonos.*

2.29. TÉTEL (EKVIVALENS ÁTALAKÍTÁSOK). *Az alábbi transzformációk minden egyenletrendszert ekvivalens egyenletrendszerbe visznek át:*

1. *két egyenlet felcserélése;*
2. *egy egyenlet nem nulla számmal való szorzása;*

Ha egy egyenletrendszer több egyenletből áll, mint ahány ismeretlene van, *túlhatározottnak* nevezzük, míg ha kevesebb egyenletből áll, *alulhatározottnak*. E fogalmak időnként félrevezető megfogalmazásokhoz és téves következtetésekre vezetnek, ha az az elképzelés alakul ki, hogy a túlhatározottság azt jelenti: az egyenletek (a feltételek) már „túl sokan” vannak ahhoz, hogy akár csak egy szám-n-es is kielégítse. Később látni fogjuk, hogy ezzel ellentétben nem a „túl sok” egyenlet, hanem az egymásnak ellentmondó egyenletek okozzák az inkonzisztenciát. Hasonlóképp az alulhatározottság nem jelenti azt, hogy szükségképpen több megoldás is van. Alulhatározott egyenletrendszer is lehet inkonzisztens. Egyedül annyi mondható: alulhatározott egyenletrendszernek nem lehet csak egyetlen megoldása.

3. egy egyenlet konstansszorosának egy másikhoz adása.
Ezen kívül
4. egy $0 = 0$ alakú egyenlet elhagyása
is ekvivalens átalakítás, de ez eggyel csökkenti az egyenletek számát.

Bizonyítás. Az első kettő és a negyedik átalakítás nyilvánvalóan nem változtatja meg a megoldások halmazát (a negyedikkel kapcsolatban lásd a 2.8. feladatot). Nézzük a harmadik átalakítást: legyen c egy tetszőleges konstans. Egy megoldást az egyenletrendszerbe helyettesítve, majd az i-edik c-szeresét hozzáadva egy másik egyenlethez, például a j-edikhez, könnyen látható, hogy a *régi* egyenletrendszer minden megoldása az *újnak* is megoldása. Ezután az új egyenletrendszer i-edik egyenletének $-c$-szeresét hozzáadjuk a j-edikhez. Így visszakapjuk a régi egyenletrendszert, tehát az előző gondolatmenet szerint az *új* egyenletrendszer minden megoldása a *réginek* is megoldása. Vagyis a két megoldáshalmaz megegyezik. Tehát ez az átalakítás is ekvivalens. □

Mátrixok Az egyenletrendszer megoldásában az ekvivalens átalakítások során a műveleteket csak az egyenletrendszer együtthatóival és konstans tagjaival végezzük, az ismeretlenek másolgatása felesleges, ezért az együtthatókat és a konstans tagokat egy táblázatba gyűjtjük – megőrizve az egyenletrendszerbeli egymáshoz való helyzetüket –, és az egyenletrendszer megoldásainak lépéseit csak ezen hajtjuk végre. Az ilyen számtáblázatokat *mátrixoknak* nevezzük, ezekkel később külön fejezetben foglalkozunk. A mátrixba írt számokat a *mátrix elemeinek* nevezzük.

Mátrix: a latin mater (māter) (*anya*, *szülőanya*, *forrás*) szó származéka a matrix (mātrīx), melynek jelentése az európai nyelvekben a következő változásokon ment át: *anyaállat*, *vemhes állat*, *anyaméh*, *bezárt hely, ahonnan valami kifejlődik*, *bezárt*, *körülzárt dolgok sokasága*, *tömbje*. Jelentése az élettanban méh, a geológiában finomszemcsés kő, melybe fossziliák, kristályok, drágakövek vannak zárva, az anatómiában a körmöt, fogat kialakító szövet.

A mátrix méretének jellemzéséhez mindig előbb a sorok, majd az oszlopok számát adjuk meg, tehát egy $m \times n$-es mátrixnak m sora és n oszlopa van. Egy ilyen mátrix általános alakja:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

A mátrix *főátlójába* azok az elemek tartoznak, amelyek ugyanannyiadik sorban vannak, mint ahányadik oszlopban, azaz a például a fenti mátrixban a főátló elemei $a_{11}, a_{22}, \dots$.

A vektorokat is szokás *mátrix jelöléssel*, *mátrix alakban*, azaz egy 1-soros vagy 1-oszlopos mátrixszal leírni – ahogy azt az első fejezetben mi is tettük. Az $n \times 1$-es mátrixot *oszlopvektornak* vagy *oszlopmátrixnak*, az $1 \times n$-es mátrixot *sorvektornak sormátrixnak* is szokás nevezni. Annak a kérdésnek az eldöntése, hogy egy n-dimenziós vektort sor- vagy oszlopvektorral reprezentáljunk, döntés (szokás, ízlés) kérdése. Manapság jobban el van terjedve a vektorok oszlopvektoros jelölése, ezért e könyvben alapértelmezésként mi is ezt a jelölést fogjuk használni, de egyes témáknál a másik használatát is bemutatjuk. Így tehát az $(1, 2)$ vektornak megfelelő sorvektor és oszlopvektor alakja

$$\begin{bmatrix} 1 & 2 \end{bmatrix}, \text{ illetve } \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

amelyek közül, ha mást nem mondunk, az utóbbit fogjuk a vektor mátrixos jelöléseként használni.

VEKTOROK MAGYAR IRODAI és általános iskolában használt jelölése – a tizedes vessző használata miatt – pontosvesszőt tesz a *vektor koordinátái* közé *elválasztójelként*. Magyar nyelvű felsőbb matematika szövegekben ez nem szokás, mi is elkerüljük, és tizedespontot, vektor koordinátái közt vesszőt használunk. Vegyük észre, hogy vektorok sorvektorral (sormátrixszal) való megadásnál írásjelet nem használunk, csak szóközzel választjuk el a koordinátákat!

Egyenletrendszer mátrixa és bővített mátrixa Az egyenletrendszer *együtthatómátrixa* az egyenletek együtthatóit, míg *bővített* mátrixa, vagy egyszerűen csak *mátrixa* az egyenletek együtthatóit és konstans tagjait tartalmazza. Az áttekinthetőség érdekében a bővített mátrixban egy függőleges vonallal választhatjuk el az együtthatókat a konstans tagoktól. A 2.25. definícióbeli általános alak együttható- és bővített mátrixa:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \qquad \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right]$$

A gyakorlatban nagy méretű egyenletrendszereket, s így nagy méretű mátrixokat is kezelni kell. Ezekben az elemek nagy része általában 0. Az ilyen mátrixokat *ritka mátrixoknak* nevezzük. Szokás az ilyen együtthatómátrixú egyenletrendszereket is ritkának nevezni. A nem ritka mátrixokat *sűrűnek* nevezzük. Előbb a kis méretű sűrű mátrixokra hatékony módszerekkel ismerkedünk meg.

2.30. PÉLDA (MÁTRIX HASZNÁLATA A MEGOLDÁSHOZ). *Oldjuk meg a következő egyenletrendszert!*

$$\begin{aligned} 2x + 3y + 2z &= 7 \\ x + y + z &= 3 \\ 2x + 2y + 3z &= 6 \end{aligned}$$

MEGOLDÁS. Két lehetséges megoldást mutatunk. A (2.20) egyenletrendszereinél látott háromszög, illetve átlós alak elérése a cél. Először írjuk fel az egyenletrendszer bővített mátrixát!

$$\begin{aligned} 2x + 3y + 2z &= 7 \\ x + y + z &= 3 \\ 2x + 2y + 3z &= 6 \end{aligned} \qquad \left[\begin{array}{ccc|c} 2 & 3 & 2 & 7 \\ 1 & 1 & 1 & 3 \\ 2 & 2 & 3 & 6 \end{array}\right]$$

Kicseréljük az első két egyenletet:

$$\begin{aligned} x + y + z &= 3 \\ 2x + 3y + 2z &= 7 \\ 2x + 2y + 3z &= 6 \end{aligned}$$

Kicseréljük az első két sort:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 2 & 3 & 2 & 7 \\ 2 & 2 & 3 & 6 \end{array}\right]$$

Az első egyenlet 2-szeresét kivonjuk a második, majd a harmadik egyenletből (azaz -2-szeresét hozzáadjuk a második majd a harmadik egyenlethez).

$$\begin{aligned} x + y + z &= 3 \\ y &= 1 \\ z &= 0 \end{aligned}$$

Az első sor kétszeresét kivonjuk a második és harmadik sorból (azaz az első sor -2-szeresét hozzáadjuk a második majd a harmadik sorhoz).

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

Az egyenletrendszerről azonnal leolvasható y és z értéke. Ezeket az első egyenletbe helyettesítve megkapjuk x értékét is, nevezetesen $x + y + z = 3$, azaz $y = 1$ és $z = 0$ behelyettesítése után: $x + 1 + 0 = 3$, vagyis $x = 2$. Másik megoldási módszerhez jutunk, ha a visszahelyettesítés helyett folytathatjuk az ekvivalens átalakítások sorozatát:

Kivonjuk a második, majd a harmadik egyenletet az elsőből:

$$\begin{aligned} x &= 2 \\ y &= 1 \\ z &= 0 \end{aligned}$$

Kivonjuk a második, majd a harmadik sort az elsőből:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

Így olyan alakra hoztuk az egyenletrendszert, illetve a bővített mátrixot, amiből azonnal leolvasható a megoldás: $(x, y, z) = (2, 1, 0)$. □

Sormodell: hipersíkok metszete A lineáris egyenletrendszerek szemléltetésére két geometriai modellt mutatunk, melyek segíteni fognak az általánosabb fogalmak megértésében, szemléltetésében.

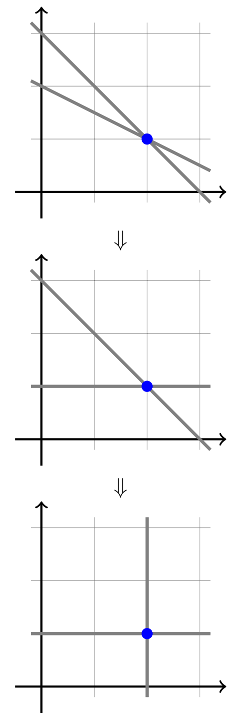
A sormodell lépései jól nyomon követők a SagePlayer sormodell című demonstrációján. Ott saját bővített mátrixokkal is lehet kísérletezni.

Tekintsük a kétváltozós lineáris $ax + by = c$ egyenletet, ahol a, b és c valós konstansok. Ha a és b legalább egyike nem 0, akkor az egyenletet kielégítő pontok halmaza egy egyenes, vagyis a megoldáshalmaz egy egyenest alkot. (Ha $a = b = c = 0$, akkor az egyenlet alakja $0x + 0y = 0$, azaz $0 = 0$, ami minden (x, y) számpárra fennáll, vagyis az egyenletet kielégítő (x, y) pontok halmaza az egész sík. Ha pedig $a = b = 0$, de $c \neq 0$, akkor az egyenletnek nincs megoldása.)

2.31. PÉLDA (SORMODELL KÉT KÉTISMERETLENES EGYENLETTEL). *Oldjuk meg az*

$$\begin{aligned} x + y &= 3 \\ x + 2y &= 4 \end{aligned}$$

egyenletrendszert ekvivalens átalakításokkal, és ábrázoljuk minden lépésben.



2.6. ábra: Egyenletrendszer megoldásának szemléltetése

MEGOLDÁS. Két lépésben megoldhatjuk az egyenletrendszert, ha az első egyenletet kivonjuk a másodikból, majd az így kapott egyenletrendszerben a második egyenletet kivonjuk az elsőből, azaz:

$$\begin{aligned} x + y &= 3 \\ x + 2y &= 4 \end{aligned} \quad \Rightarrow \quad \begin{aligned} x + y &= 3 \\ y &= 1 \end{aligned} \quad \Rightarrow \quad \begin{aligned} x &= 2 \\ y &= 1 \end{aligned}$$

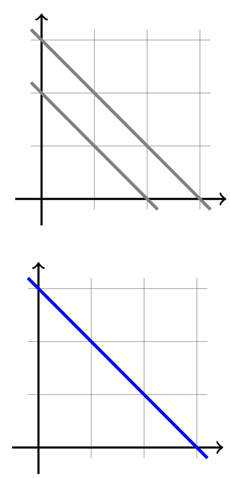
Az egyenletrendszer két egyenlete egy-egy egyenes egyenlete a síkban. Az, hogy az egyenletrendszer megoldható, pontosan azt jelenti, hogy a két egyenesnek van közös pontja, példánkban a $(2, 1)$ pont. A 2.6 ábra a megoldás lépéseit szemlélteti az egyenletek grafikonjával. □

2.32. PÉLDA (HA 0 LESZ A BAL OLDAL). *Vizsgáljuk meg az alábbi egyenletrendszert a sormodellben!*

$$\begin{aligned} x + y &= 3 \\ x + y &= 2 \end{aligned}$$

MEGOLDÁS. Látható, hogy ez két párhuzamos egyenes egyenlete, melyeknek nincs közös pontjuk, így az egyenletrendszer nem oldható meg. Ha az első egyenletet kivonjuk a másodikból, az ellentmondó $0 = -1$ egyenletet kapjuk, vagyis így is arra jutottunk, hogy az egyenletrendszer nem oldható meg. Az

$$\begin{aligned} x + y &= 3 \\ x + y &= 3 \end{aligned}, \quad \text{vagy az} \quad \begin{aligned} x + y &= 3 \\ 2x + 2y &= 6 \end{aligned}$$

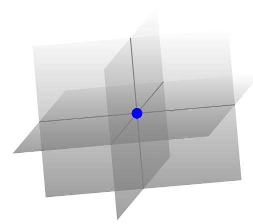


2.7. ábra: A megoldás szemléltetése, ha a két egyenlet egyikének bal oldala nullává tehető

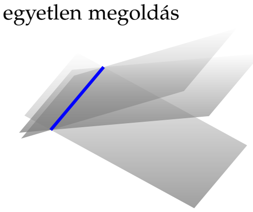
egyenletrendszerekben az első egyenlettel a második $0 = 0$ alakra hozható, aminek minden számpár megoldása, így elhagyható. Így csak az $x + y = 3$ egyenlet marad. Ennek összes megoldása paraméteres alakba írva például $(x, y) = (3 - t, t)$. □

2.33. PÉLDA (SORMODELL HÁROM HÁROMISMERETLENES EGYENLETTEL). *Vizsgáljuk meg, hogy három háromismeretlenes egyenletből álló egyenletrendszer megoldásainak halmaza milyen geometriai alakzatot adhat!*

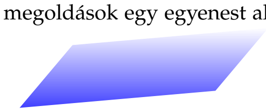
Megoldás. Ha a három egyenlettel meghatározott három sík általános helyzetű, akkor az egyenletrendszernek egyetlen megoldása van (ld. 2.8. (a) ábra). Például a 2.30. példában szereplő egyenletrendszernek egyetlen megoldása van: $(x, y, z) = (2, 1, 0)$.



(a) Három általános helyzetű sík: egyetlen megoldás



(b) Egy egyenesen átmenő, de nem csupa azonos sík: végtelen sok megoldás, a megoldások egy egyenest alkotnak



(c) Azonos síkok: végtelen sok megoldás, a megoldások egy síkot alkotnak

2.8. ábra: Megoldható egyenletrendszerek ábrázolása (a megoldáshalmazt kék szín jelzi)

Tekintsük a

$$\begin{aligned} 2x + y + 2z &= 5 \\ x + y + z &= 3 \\ 3x + 2y + 3z &= 8 \end{aligned} \tag{2.21}$$

egyenletrendszert. Ennek egy megoldása $(x, y, z) = (2, 1, 0)$, ugyanakkor a három sík normálvektorai egy síkba esnek, ugyanis

$$(2, 1, 2) + (1, 1, 1) = (3, 2, 3).$$

Mivel mindhárom normálvektorra merőleges például a

$$(2, 1, 2) \times (1, 1, 1) = (-1, 0, 1)$$

vektor, ezért e vektor párhuzamos mindhárom síkkal. A $(2, 1, 0)$ ponton átmenő, és $(-1, 0, 1)$ irányvektorú egyenes tehát benne van mindhárom síkban (ilyen esetet ábrázol a 2.8. (b) ábra). Az összes megoldást tehát megadja ennek az egyenesnek a paraméteres vektoregyenlete:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Hamarosan ugyanezt a megoldást ekvivalens átalakításokkal is meg fogjuk tudni határozni.

Hasonló esetet ábrázol a 2.9. *(b)* ábra, ahol mindhárom sík párhuzamos egy egyenessel, de a síkok egymással nem párhuzamosak, viszont a 2.8. *(b)* esettel ellentétben az egyenletrendszernek nincs megoldása. Ilyen például a (2.21) kis változtatásával kapott

$$\begin{aligned} 2x + y + 2z &= 5 \\ x + y + z &= 3 \\ 3x + 2y + 3z &= 9 \end{aligned} \tag{2.22}$$

egyenletrendszer. Vegyük észre, hogy míg a (2.21) egyenletrendszerben a harmadik egyenletből kivonva az első kettőt az elhagyható $0 = 0$ egyenletet kapjuk, addig a (2.22) egyenletrendszerben az ellentmondó $0 = 1$ egyenletre jutunk. Így ennek az egyenletrendszernek nincs megoldása.

Végül tekintsük az

$$\begin{aligned} x + y + z &= 3 \\ 2x + 2y + 2z &= 6 \\ 3x + 3y + 3z &= 9 \end{aligned} \tag{2.23}$$

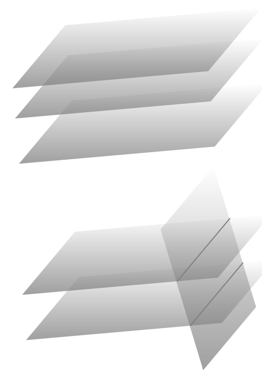
egyenletrendszert! Látható, hogy a második és a harmadik egyenlet az első konstansszorosa, azaz ugyanannak a síknak az egyenletei, az egyenletrendszer tehát ekvivalens az egyetlen

$$x + y + z = 3$$

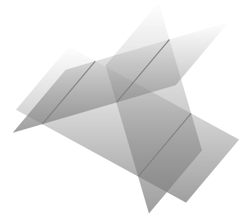
egyenletből álló egyenletrendszerrel. Az y-nak és z-nek tetszőleges értéket választunk, például legyen $y = s$, $z = t$, akkor $x = 3 - y - z$, azaz $x = 3 - s - t$. Így az összes megoldás: $(x, y, z) = (3 - s - t, s, t)$. Ezt oszlopvektorokkal fölírva kapjuk, hogy a megoldás

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Ez a 2.8. *(c)* ábra szerint eset. A 2.9. *(a)*-beli eseteknek megfelelő egyenletrendszerek felírását az olvasóra hagyjuk. □



(a) A síkok közül legalább kettő párhuzamos, de nem azonos



(b) Egy egyenessel párhuzamos, de közös egyenest nem tartalmazó három sík

2.9. ábra: Nem megoldható egyenletrendszerek

Az előző példákban a 2- és 3-dimenziós térben szemléltettük a 2-, illetve 3-ismeretlenes egyenletrendszer megoldásait. E geometriai modell lényege a következőképp foglalható össze az n-ismeretlenes esetben:

2.34. ÁLLÍTÁS (SORMODELL). *Ha egy n-ismeretlenes egyenlet bal oldalán nem minden együttható 0, akkor az egyenletet kielégítő pontok (azaz az egyenlet megoldásai) egy hipersíkot alkotnak $\mathbb{R}^n$-ben. Ha egy n-ismeretlenes egyenletrendszer m ilyen egyenletből áll, akkor az egyenletrendszer megoldása a nekik megfelelő m hipersík közös része $\mathbb{R}^n$-ben.*

Az m egyenlet a skaláris szorzás segítségével tömörebb alakban is fölírható. Az $m \times n$-es $\mathbf{A}$ együtthatómátrixú lineáris egyenletrendszer i-edik egyenletének alakja

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i.$$

Ha $\mathbf{a}_{i*}$ jelöli az $\mathbf{A}$ mátrix i-edik sorát, ez az egyenlet vektorok skaláris szorzataként is felírható:

$$\mathbf{a}_{i*} \cdot \mathbf{x} = b_i. \tag{2.24}$$

Ez különösen akkor lesz érdekes, ha homogén lineáris egyenletrendszereket fogunk vizsgálni, mert ott mindegyik egyenlet $\mathbf{a}_{i*} \cdot \mathbf{x} = 0$ alakot ölt, ami azt jelenti, hogy olyan $\mathbf{x}$ vektort keresünk, mely merőleges az $\mathbf{a}_{i*}$ vektorok mindegyikére.

Oszlopmodell: vektor előállítása lineáris kombinációként E modellben az egyenletrendszerre úgy tekintünk, mint egy olyan vektoregyenletre,

Az oszlopmodell lépései jól nyomon követetők a SagePlayer oszlopmodell című demonstrációján. Ott saját bővített mátrixokkal is lehet kísérletezni.

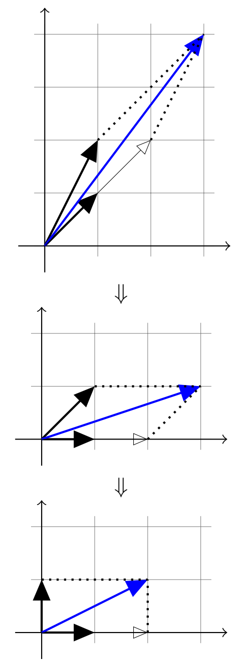
amelyben egy vektort kell előállítani adott vektorok lineáris kombinációjaként. Például a 2.31. példabeli

$$\begin{aligned} x + y &= 3 \\ x + 2y &= 4 \end{aligned}$$

egyenletrendszer ekvivalens az

$$\begin{bmatrix}1\\1\end{bmatrix} x + \begin{bmatrix}1\\2\end{bmatrix} y = \begin{bmatrix}3\\4\end{bmatrix}$$

vektoregyenlettel. Itt az a feladat, hogy megkeressük az $(1,1)$ és $(1,2)$ vektoroknak azt a lineáris kombinációját, amely egyenlő a $(3,4)$ vektorral.



2.10. ábra: A megoldás lépései a vektormodellben.

2.35. példa (A megoldás lépései az oszlopmodellben). *Kövessük végig a fenti egyenletrendszer 2.31. példában adott megoldásának lépéseit e modellben is.*

Megoldás. Az ekvivalens átalakítások lépései:

$$\begin{aligned} x + y &= 3 \\ x + 2y &= 4 \end{aligned} \quad\Rightarrow\quad \begin{aligned} x + y &= 3 \\ y &= 1 \end{aligned} \quad\Rightarrow\quad \begin{aligned} x &= 2 \\ y &= 1 \end{aligned}$$

Vektoros alakban:

$$\begin{bmatrix}1\\1\end{bmatrix} x + \begin{bmatrix}1\\2\end{bmatrix} y = \begin{bmatrix}3\\4\end{bmatrix} \Rightarrow \begin{bmatrix}1\\0\end{bmatrix} x + \begin{bmatrix}1\\1\end{bmatrix} y = \begin{bmatrix}3\\1\end{bmatrix} \Rightarrow \begin{bmatrix}1\\0\end{bmatrix} x + \begin{bmatrix}0\\1\end{bmatrix} y = \begin{bmatrix}2\\1\end{bmatrix}.$$

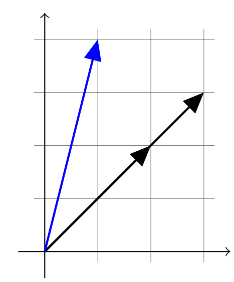
E lépéseket szemléltetjük a 2.10 ábrán. □

Általánosan kimondható a következő:

2.36. állítás (Oszlopmodell). *A 2.25. definícióban megadott (2.19) egyenletrendszer a következő vektoregyenlettel ekvivalens:*

$$\begin{bmatrix}a_{11}\\a_{21}\\\vdots\\a_{m1}\end{bmatrix} x_1 + \begin{bmatrix}a_{12}\\a_{22}\\\vdots\\a_{m2}\end{bmatrix} x_2 + \ldots + \begin{bmatrix}a_{1n}\\a_{2n}\\\vdots\\a_{mn}\end{bmatrix} x_n = \begin{bmatrix}b_1\\b_2\\\vdots\\b_m\end{bmatrix}$$

Az egyenletrendszer megoldása ekvivalens egy vektoregyenlet megoldásával, ahol az egyenletrendszer konstans tagjaiból álló vektort kell az együtthatómátrix oszlopvektorainak lineáris kombinációjaként előállítani.



2.11. ábra: Tekintsük a nyilvánvalóan ellentmondó $2x + 3y = 1$, $2x + 3y = 4$ egyenletrendszert, és annak oszlopmodell alakját: $(2,2)x + (3,3)y = (1,4)$. Mivel a $(2,2)$ és a $(3,3)$ vektorok lineáris kombinációjaként csak (c,c) alakú vektor állítható elő, az $(1,4)$ vektor nem, ezért a fenti egyenletrendszernek nincs megoldása.

E modellben egy egyenletrendszer pontosan akkor oldható meg, ha az együtthatómátrix oszlopvektorainak összes lineáris kombinációjából álló halmazban a konstans tagokból álló vektor is szerepel (ld. 2.10. feladat).

Feladatok

2.1.• Melyek lineáris egyenletek az x, y és z változókban az alábbiak közül?

a) $3x-(\ln 2)y+e^3z=0.4$ *b)* $a^2x-b^2y=0$

c) $xy-yz-zx=0$ *d)* $(\sin 1)x+y-\pi z=0$

e) $\dfrac{x}{a}+\dfrac{y}{b}+\dfrac{z}{c}=1$ *f)* $\dfrac{1}{x}+\dfrac{1}{y}+\dfrac{1}{z}=1$

Igazoljuk, hogy az alábbi egyenletrendszerek ekvivalensek!

2.2. $\left.\begin{aligned} x+3y&=5\\ y&=1\end{aligned}\right\}$ $\left.\begin{aligned} x+y&=3\\ x\quad&=2\end{aligned}\right\}$

2.3. $\left.\begin{aligned} 2x+3y&=2\\ 0x+0y&=3\end{aligned}\right\}$ $\left.\begin{aligned} x+y&=2\\ x+y&=7\end{aligned}\right\}$

Oldjuk meg (fejben számolva) az alábbi lineáris egyenletrendszereket az $a=1$, $b=2$, $c=3$ paraméterválasztás esetén!

2.4. $\left.\begin{aligned} (2a-b)x+(3a-c)y&=0\\ (3b-2c)x+(b-2a)y&=0\end{aligned}\right\}$

2.5. $\left.\begin{aligned} (b-a)x+(3a-c)y&=1\\ (3b-2c)x+(b-2a)y&=0\end{aligned}\right\}$

2.6. $\left.\begin{aligned} (b-a)x+(3a-c)y&=1\\ (3b-2c)x+(b-2a)y&=1\end{aligned}\right\}$

2.7. $\left.\begin{aligned} (b-a)x+(3a-c)y&=1\\ (3b-2c)x+\ (c-b)y&=2\end{aligned}\right\}$

2.8. EGYENLETRENDSZEREK KÖZÖS MEGOLDÁSA Tekintsük az azonos ismeretleneket tartalmazó $\mathcal{E}_1$ és $\mathcal{E}_2$ egyenletrendszereket. Legyen ezek megoldáshalmaza $\mathcal{M}_1$, illetve $\mathcal{M}_2$. Mutassuk meg, hogy ha $\mathcal{E}$ az $\mathcal{E}_1$ és $\mathcal{E}_2$ egyenletrendszerek egyesítése, azaz $\mathcal{E}=\mathcal{E}_1\cup\mathcal{E}_2$, és $\mathcal{M}$ az $\mathcal{E}$ megoldáshalmaza, akkor $\mathcal{M}$ az $\mathcal{M}_1$ és $\mathcal{M}_2$ közös része, azaz $\mathcal{M}=\mathcal{M}_1\cap\mathcal{M}_2$.

Vizsgáljuk meg ezt állítást az alábbi esetekben:

a) $\mathcal{E}_1=\{\,x+y=2\}$, $\mathcal{E}_2=\{\,x-y=0\,\}$;
b) $\mathcal{E}_1=\{\,x+y=2,\ x-y=0\}$, $\mathcal{E}_2=\{\,x-y=0\,\}$;
c) $\mathcal{E}_1=\{\,x+y=2,\ x-y=0\}$, $\mathcal{E}_2=\{\,x-y=1\,\}$;
d) $\mathcal{E}_1=\{\,x+y=2,\ x-y=0\}$, $\mathcal{E}_2=\{\,0x+0y=0\,\}$;
e) $\mathcal{E}_1$ tetszőleges egyenletrendszer, $\mathcal{E}_2=\{\,0=0\,\}$.

2.9.• SOR ÉS OSZLOPMODELL Rajzoljuk fel a következő két egyenletrendszerhez tartozó sormodell és oszlopmodell szerinti ábrát!

a) $\begin{aligned} 2x+3y&=7\\ 3x-2y&=4\end{aligned}$ *b)* $\begin{aligned} 2x+4y&=3\\ 3x+6y&=4\end{aligned}$

2.10. SOR- ÉS AZ OSZLOPMODELL 3D-BEN Vizsgáljuk meg az alábbi két – azonos együtthatómátrixú – egyenletrendszer megoldhatóságát a sor- és az oszlopmodellben:

$$\begin{aligned} x+\ y+2z&=3\\ x+2y+4z&=3\\ 3x+4y+8z&=9\end{aligned}\qquad\qquad \begin{aligned} x+\ y+2z&=3\\ x+2y+4z&=3\\ 3x+4y+8z&=1\end{aligned}$$

2.11. SOR ÉS OSZLOPMODELL $m\neq n$ ESETÉN Vizsgáljuk meg az alábbi három egyenletrendszer megoldhatóságát a sor- és az oszlopmodellben:

$$a)\ \begin{aligned} x+\ y&=3\\ x+\ y&=4\\ x+3y&=5\end{aligned}\qquad b)\ \begin{aligned} x+\ y&=3\\ x+2y&=4\\ x+3y&=5\end{aligned}\qquad c)\ \begin{aligned} x+\ y&=3\\ x+2y&=3\\ x+3y&=5\end{aligned}$$

2.12.• IGAZ – HAMIS Mely állítások igazak, melyek hamisak az alábbiak közül?

a) Ha egy n-ismeretlenes egyenletrendszer olyan hipersíkok egyenleteiből áll, melyek közt van két párhuzamos, akkor az egyenletrendszer nem oldható meg.

b) Ha egy n-ismeretlenes egyenletrendszer nem oldható meg, akkor az egyenletek olyan hipersíkok egyenletei, melyek közt van két párhuzamos, de nem azonos hipersík.

c) Ha egy n-ismeretlenes egyenletrendszer csak két egyenletből áll, akkor az oszlopmodell szerint pontosan akkor oldható meg tetszőleges jobb oldal esetén, ha a vektoregyenlet bal oldalán szereplő vektorok közt van kettő lineárisan független.

2.13.• Egészítsük ki az alábbi állításokat úgy, hogy igazak legyenek!

a) Egy két egyenletből álló háromismeretlenes egyenletrendszer sormodellje szerinti ábra a(z) ..-dimenziós térben .. darab ......ból/ből áll, melyek ha ...., akkor az egyenletrendszernek nincs megoldása, egyébként megoldásainak száma .... Oszlopmodellje a(z) ..-dimenziós térben .. darab ....ból/ből áll.

b) Egy három egyenletből álló kétismeretlenes egyenletrendszer sormodellje szerinti ábra a(z) ..-dimenziós térben .. darab ....ból/ből áll, míg az oszlopmodellje a ..-dimenziós térben .. darab ......ból/ből.

c) Egy négy egyenletből álló ötismeretlenes egyenletrendszer sormodellje szerinti ábra a(z) ..-dimenziós térben .. darab ....ból/ből áll. Oszlopmodellje a(z) ..-dimenziós térben .. darab ....-ból/ből áll.

Megoldás kiküszöböléssel

Elemi sorműveletek A lineáris egyenletrendszerek egyik megoldási módszerének lényege, hogy ekvivalens átalakításokkal olyan alakra hozzuk az egyenletrendszert, melyből visszahelyettesítések után, vagy azok nélkül azonnal leolvasható az eredmény. Az átalakításokat a bővített mátrixon hajtjuk végre úgy, hogy a nekik megfelelő átalakítások az egyenletrendszeren ekvivalens átalakítások legyenek. A 2.29. tételben felsorolt első három ekvivalens átalakítás nem változtatja meg az egyenletrendszer egyenleteinek számát sem. Az egyenletrendszer bővített mátrixán az ezeknek megfelelő átalakításokat elemi sorműveleteknek nevezzük.[2]

[2] Lineáris egyenletrendszerek felírása és megoldása már időszámításunk előtt 300 körül babiloni iratokban szerepelt. Az első századra teszik a kínai Jiǔzhāng Suànshù című mű megjelenését, mely az előző ezer évben összegyűlt matematikai tudást foglalja össze (címének magyar fordítása „A matematikai művészet kilenc fejezete” vagy „Kilenc fejezet a matematikai eljárásokról” lehet). E műben már a kiküszöbölés (azaz a Gauss-elimináció) néven ismert technikát alkalmazzák lineáris egyenletrendszer megoldására. A két fenti műben szereplő egyenletrendszerek, és további történeti részletek olvashatók a The MacTutor History of Mathematics archive című weboldalon.

2.37. DEFINÍCIÓ (ELEMI SORMŰVELETEK). *Egy mátrix sorain végzett alábbi műveleteket* elemi sorműveleteknek *nevezzük:*

1. Sorcsere: *két sor cseréje.*
2. Beszorzás: *egy sor beszorzása egy nemnulla számmal.*
3. Hozzáadás: *egy sorhoz egy másik sor konstansszorosának hozzáadása.*

Természetesen egy sort el is oszthatunk egy nemnulla c számmal, hisz az az $1/c$-vel való beszorzással egyenértékű. Hasonlóképp levonhatjuk egy sorból egy másik sor c-szeresét, hisz az a $-c$-szeresének hozzáadásával ekvivalens.

Az elemi átalakításokra a következő jelöléseket fogjuk használni:

1. $S_i \leftrightarrow S_j$: az i-edik és a j-edik sorok cseréje.
2. cS_i: az i-edik sor beszorzása c-vel.
3. $S_i + cS_j$: a j-edik sor c-szeresének az i-edik sorhoz adása.

Az elemi sorműveletek mintájára elemi oszlopműveletek is definiálhatók, de azokat ritkán használjuk. Jelölésükre értelemszerűen az $O_i \leftrightarrow O_j$, cO_i, $O_i + cO_j$ formulákat használjuk.

Lépcsős alak Az eddig megoldott egyenletrendszerekben igyekeztünk átlós, vagy legalább átló alatt kinullázott alakra hozni az egyenletrendszert, mint azt például a 2.30. példában tettük. Ez nem mindig sikerül, mert néha nem kívánt elemek is kinullázódnak, de a következőkben definiált lépcsős alakhoz mindig el tudunk jutni.

2.38. DEFINÍCIÓ (LÉPCSŐS ALAK). *Egy mátrix* lépcsős, *vagy* sorlépcsős alakú, *ha kielégíti a következő két feltételt:*

1. *a csupa* 0*-ból álló sorok (ha egyáltalán vannak) a mátrix utolsó sorai;*
2. *bármely két egymás után következő nem-o sorban az alsó sor elején több* 0 *van, mint a fölötte lévő sor elején.*

A nemnulla sorok első zérustól különböző elemét főelemnek, vezérelemnek *vagy* pivotelemnek *hívjuk. Egy főelem oszlopának* főoszlop *vagy* bázisoszlop *a neve.*

2.39. PÉLDA (LÉPCSŐS ALAK). *A következő mátrixok lépcsős alakúak:*

$$\begin{bmatrix} 3 & 2 \\ 0 & 4 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & -2 & 3 & -4 \\ 0 & 0 & -5 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Gauss-módszer A *Gauss-módszer* vagy más néven *Gauss-kiküszöbölés* vagy *Gauss-elimináció* a lineáris egyenletrendszerek megoldásának egy módszere. Lényege, hogy a lineáris egyenletrendszer bővített mátrixát elemi sorműveletekkel lépcsős alakra hozzuk, és abból visszahelyettesítéssel meghatározzuk a megoldás általános alakját. A Gauss-módszer könnyen algoritmizálható, ha sorban haladunk az oszlopokon. A módszerre láttunk már példát, ilyen volt a 2.30. példa első megoldása. Most lássunk két további példát.

2.40. PÉLDA (GAUSS-MÓDSZER, EGY MEGOLDÁS). *Oldjuk meg az alábbi egyenletrendszert Gauss-módszerrel:*

$$\begin{aligned} x + y + 2z &= 0 \\ 2x + 2y + 3z &= 2 \\ x + 3y + 3z &= 4 \\ x + 2y + z &= 5 \end{aligned}$$

MEGOLDÁS. Írjuk fel az egyenletrendszer bővített mátrixát, és oszloponként haladva küszöböljük ki – nullázzuk ki – a főelemek alatti elemeket!

$$\left[\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 2 & 2 & 3 & 2 \\ 1 & 3 & 3 & 4 \\ 1 & 2 & 1 & 5 \end{array}\right] \xrightarrow[S_4-S_1]{\substack{S_2-2S_1 \\ S_3-S_1}} \left[\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 0 & -1 & 2 \\ 0 & 2 & 1 & 4 \\ 0 & 1 & -1 & 5 \end{array}\right] \xrightarrow{S_2 \leftrightarrow S_3} \left[\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 2 & 1 & 4 \\ 0 & 0 & -1 & 2 \\ 0 & 1 & -1 & 5 \end{array}\right] \xrightarrow{S_4-\frac{1}{2}S_2}$$

$$\left[\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 2 & 1 & 4 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & -\frac{3}{2} & 3 \end{array}\right] \xrightarrow{S_4-\frac{3}{2}S_3} \left[\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 2 & 1 & 4 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right] \Longrightarrow \begin{aligned} x + y + 2z &= 0 \\ 2y + z &= 4 \\ -z &= 2 \end{aligned}$$

A harmadik egyenletből $z = -2$, ezt a másodikba helyettesítve $y = 3$, ezeket az elsőbe helyettesítve kapjuk, hogy $x = 1$, azaz az egyetlen megoldás $(x, y, z) = (1, 3, -2)$. □

Mit csinálunk akkor, ha a lépcsős alak szerint kevesebb a főelemek, mint az oszlopok száma? Egyelőre bevezetünk két elnevezést, melyek jelentése hamarosan világos lesz: az egyenletrendszer azon változó-it, melyek főelemek oszlopaihoz tartoznak, *kötött változóknak*, míg az összes többi változót *szabad változónak* nevezzük.

2.41. példa (Gauss-módszer, végtelen sok megoldás). *Oldjuk meg az alábbi egyenletrendszert Gauss-módszerrel:*

$$\begin{aligned} x_1 + 2x_2 + x_3 + 2x_4 + x_5 &= 1 \\ x_1 + 2x_2 + 3x_3 + 3x_4 + x_5 &= 0 \\ 3x_1 + 6x_2 + 7x_3 + 8x_4 + 3x_5 &= 1 \end{aligned}$$

Megoldás. Írjuk fel az egyenletrendszer bővített mátrixát, és oszloponként haladva küszöböljük ki a főelemek alatti elemeket!

$$\left[\begin{array}{ccccc|c} 1 & 2 & 1 & 2 & 1 & 1 \\ 1 & 2 & 3 & 3 & 1 & 0 \\ 3 & 6 & 7 & 8 & 3 & 1 \end{array}\right] \overset{\substack{S_2-S_1 \\ S_3-3S_1}}{\Longrightarrow} \left[\begin{array}{ccccc|c} 1 & 2 & 1 & 2 & 1 & 1 \\ 0 & 0 & 2 & 1 & 0 & -1 \\ 0 & 0 & 4 & 2 & 0 & -2 \end{array}\right] \overset{S_3-2S_2}{\Longrightarrow}$$

$$\left[\begin{array}{ccccc|c} 1 & 2 & 1 & 2 & 1 & 1 \\ 0 & 0 & 2 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \Longrightarrow \begin{aligned} x_1 + 2x_2 + x_3 + 2x_4 + x_5 &= 1 \\ 2x_3 + x_4 &= -1 \end{aligned}$$

Az egyenletrendszer kötött változói a lépcsős alak főoszlopaihoz tartozó változók, azaz x_1 és x_3. A szabad változók: x_2, x_4, x_5. A szabad változóknak tetszőleges értékeket adhatunk, a kötöttek értéke kifejezhető velük. Legyen például a szabad változók értéke $x_2 = s$, $x_4 = t$, $x_5 = u$. Ezek behelyettesítése után a fenti egyenletek közül először a másodikból kifejezzük x_3-at, majd azt behelyettesítjük az elsőbe, ahonnan kifejezzük az x_1-et, azaz a fenti egyenletekből kifejezzük a kötött változókat:

$$\begin{aligned} x_1 &= \frac{3}{2} - 2s - \frac{3}{2}t - u \\ x_3 &= -\frac{1}{2} - \frac{1}{2}t \end{aligned}$$

. Innen az egyenletrendszer megoldása

$$(x_1, x_2, x_3, x_4, x_5) = (\frac{3}{2} - 2s - \frac{3}{2}t - u, s, -\frac{1}{2} - \frac{1}{2}t, t, u),$$

vagy mátrixjelöléssel

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} \frac{3}{2} - 2s - \frac{3}{2}t - u \\ s \\ -\frac{1}{2} - \frac{1}{2}t \\ t \\ u \end{bmatrix} = \begin{bmatrix} \frac{3}{2} \\ 0 \\ -\frac{1}{2} \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -\frac{3}{2} \\ 0 \\ -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + u \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Később különösen az utóbbi – vektorok lineáris kombinációjára bontással való – felírás lesz hasznos. □

Világos, hogy ha a szabad változóknak tetszőleges értéket adhatunk, melyből a kötött változók egyértelműen kifejezhetők, akkor a fenti példában mutatott módszerrel az egyenletrendszer összes megoldását leírtuk. Az ilyen módon megadott megoldást az egyenletrendszer *általános megoldásának*, a konkrét paraméterértékekhez tartozó megoldásokat *partikuláris megoldásnak* nevezzük. Például az előző példabeli egyenletrendszer egy partikuláris megoldása az $s = 0$, $t = 1$, $u = 2$ értékekhez tartozó

$$(x_1, x_2, x_3, x_4, x_5) = (-2, 0, -1, 1, 2).$$

Lényeges e megoldási módban, hogy a bővített mátrixot lépcsős alakra tudtuk hozni. Ez vajon mindig sikerül?

2.42. TÉTEL (LÉPCSŐS ALAKRA HOZÁS). *Bármely mátrix elemi sorműveletekkel* lépcsős *alakra hozható.*

BIZONYÍTÁS. Tekintsünk egy tetszőleges $m \times n$-es mátrixot. A következő eljárás egyes lépéseiben e mátrixnak le fogjuk takarni egy-egy sorát vagy oszlopát. Az egyszerűség kedvéért a letakarás után keletkezett mátrix sorainak és oszlopainak számát ismét m és n fogja jelölni, a_{ij} pedig a letakarások után maradt mátrix i-edik sorának j-edik elemét.

1. Ha az első oszlopban csak 0 elemek állnak, takarjuk le ezt az oszlopot, és tekintsük a maradék mátrixot. Ha ennek első oszlopában ismét csak 0 elemek vannak, azt is takarjuk le, és ezt addig folytassuk, míg egy olyan oszlopot nem találunk, amelyben van nem 0 elem. Ha ilyen oszlopot nem találunk, az eljárásnak vége, a mátrix lépcsős alakú.
2. Ha az első oszlop első sorában álló elem 0, akkor cseréljük ki e sort egy olyannal, melynek első eleme nem 0. Így olyan mátrixot kaptunk, amelyben $a_{11} \neq 0$.
3. Vegyük az i-edik sort $i = 2$-től $i = m$-ig, és ha a sor első eleme $a_{i1} \neq 0$, akkor az első sor $-a_{i1}/a_{11}$-szeresét adjuk hozzá. Mivel $a_{i1} - \frac{a_{i1}}{a_{11}} a_{11} = 0$, ezért e lépés után az a_{11} alatti elemek 0-vá válnak.
4. A fenti átalakítás után takarjuk le az első sort és az első oszlopot. Ha ekkor nem marad a mátrixban több sor, vége az eljárásnak, a korábban letakart sorokat feltárva megkaptuk a lépcsős alakot. Egyébként ugorjunk vissza az 1. lépéshez, és folytassuk az eljárást.

Világos, hogy ez az eljárás véges sok lépésben véget ér, melynek eredményeként eljutunk az eredeti mátrix egy lépcsős alakjához. □

Egy *inhomogén lineáris egyenletrendszerhez tartozó homogén lineáris egyenletrendszeren* azt a homogén egyenletrendszert értjük, melyet az inhomogénből a konstans tagok 0-ra változtatásával kapunk.

2.43. PÉLDA (HOMOGÉN LINEÁRIS EGYENLETRENDSZER MEGOLDÁSA). *Oldjuk meg a 2.41. példabeli egyenletrendszerhez tartozó homogén lineáris*

$$\begin{aligned} x_1 + 2x_2 + x_3 + 2x_4 + x_5 &= 0 \\ x_1 + 2x_2 + 3x_3 + 3x_4 + x_5 &= 0 \\ 3x_1 + 6x_2 + 7x_3 + 8x_4 + 3x_5 &= 0 \end{aligned}$$

egyenletrendszer.

MEGOLDÁS. Mivel homogén lineáris egyenletrendszerről van szó, a megoldáshoz szükségtelen a bővített mátrixot használni, hisz annak utolsó oszlopa csak nullákból áll, így az elemi sorműveletek alatt nem változik. Az együtthatómátrix lépcsős alakja ugyanazokkal a sorműveletekkel megkapható, mint a 2.41. példa megoldásában, azaz

$$\begin{bmatrix} 1 & 2 & 1 & 2 & 1 \\ 1 & 2 & 3 & 3 & 1 \\ 3 & 6 & 7 & 8 & 3 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & 2 & 1 & 2 & 1 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \Longrightarrow$$

$$\begin{aligned} x_1 + 2x_2 + x_3 + 2x_4 + x_5 &= 0 \\ 2x_3 + x_4 &= 0 \end{aligned}$$

Innen a megoldás is ugyanúgy kapható meg, sőt, ugyanaz a lineáris kombináció szerepel benne, csak a konstans tagok nem szerepelnek:

$$(x_1, x_2, x_3, x_4, x_5) = (-2s - \frac{3}{2}t - u, s, -\frac{1}{2}t, t, u),$$

vagy mátrixjelöléssel

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2s - \frac{3}{2}t - u \\ s \\ -\frac{1}{2}t \\ t \\ u \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -\frac{3}{2} \\ 0 \\ -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + u \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

A homogén és inhomogén egyenletrendszerek e példából sejthető kapcsolatára még visszatérünk a 3.18. tételben. □

Végül egy alkalmazás:

2.44. PÉLDA (SÍKOK METSZÉSVONALÁNAK MEGHATÁROZÁSA). *Határozzuk meg az alábbi két sík metszésvonalának explicit (paraméteres) alakját!*

$$\begin{aligned} x + y + z &= 1 \\ 3x + 4y &= 2 \end{aligned}$$

MEGOLDÁS. A fenti egyenletekkel megadott két sík metszésvonalának meghatározásához, pontosabban a metszésvonal explicit, paraméteres egyenletrendszerének felírásához egyszerűen meg kell oldani a két egyenletből álló egyenletrendszert:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 3 & 4 & 0 & 2 \end{array}\right] \overset{S_2-3S_1}{\Longrightarrow} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -3 & -1 \end{array}\right] \Longrightarrow \begin{aligned} x+y+\ z &= \ \ 1 \\ y-3z &= -1 \end{aligned}$$

Ebből $z = t$ paraméterválasztással $y = -1 + 3t$ és $x = 2 - 4t$, azaz

$$(x, y, z) = (-4t + 2, 3t - 1, t) = (2, -1, 0) + t(-4, 3, 1),$$

vagy mátrixjelöléssel

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -4 \\ 3 \\ 1 \end{bmatrix}. \qquad \square$$

Redukált lépcsős alak A 2.30. példa második megoldási módszerében átlós alakra hoztuk az együtthatómátrixot, azaz nem elégedtünk meg azzal, hogy a főelemek alatt kinulláztunk minden együtthatót, hanem a főelemeket 1-re változtattuk a sor beszorzásával, és a főelemek fölött is kinulláztunk minden elemet, más szóval elimináltunk.

2.45. DEFINÍCIÓ (REDUKÁLT LÉPCSŐS ALAK). *Egy mátrix* redukált lépcsős, *vagy* redukált sorlépcsős alakú, *ha kielégíti a következő feltételeket:*

1. *lépcsős alakú;*
2. *minden főelem egyenlő* 1*-gyel;*
3. *a főelemek oszlopaiban a főelemeken kívül minden elem* 0*;*

A főelemet itt vezéregyesnek *vagy* vezető egyesnek *is szokás nevezni.*

2.46. PÉLDA (REDUKÁLT LÉPCSŐS ALAK). *A következő mátrixok redukált lépcsős alakúak:*

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & -2 & 0 & -4 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Minden valós, vagy racionális elemű mátrix redukált lépcsős alakra hozható, azonban az egészegyütthatós mátrixok általában nem, ha az egészeken belül akarunk maradni. De az egészegyütthatós mátrixok is redukált lépcsős alakra hozhatók a racionálisok számkörében. Az kiküszöbölés lépéseinek követésére használható a SagePlayer redukált lépcsős alak című szemléltető munkalapja.

2.47. PÉLDA (REDUKÁLT LÉPCSŐS ALAKRA HOZÁS). *Hozzuk redukált lépcsős alakra az*

$$\begin{bmatrix} 1 & 3 & 0 \\ 1 & 1 & 2 \\ 2 & 2 & 4 \end{bmatrix}$$

mátrixot!

MEGOLDÁS. Egy lehetséges megoldás:

$$\begin{bmatrix} 1 & 3 & 0 \\ 1 & 1 & 2 \\ 2 & 2 & 4 \end{bmatrix} \overset{\substack{S_2-S_1 \\ S_3-2S_1}}{\Longrightarrow} \begin{bmatrix} 1 & 3 & 0 \\ 0 & -2 & 2 \\ 0 & -4 & 4 \end{bmatrix} \overset{-\frac{1}{2}S_2}{\Longrightarrow}$$
$$\begin{bmatrix} 1 & 3 & 0 \\ 0 & 1 & -1 \\ 0 & -4 & 4 \end{bmatrix} \overset{S_3+4S_2}{\Longrightarrow} \begin{bmatrix} 1 & 3 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \overset{S_1-3S_2}{\Longrightarrow} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Egy másik lehetséges megoldás:

$$\begin{bmatrix} 1 & 3 & 0 \\ 1 & 1 & 2 \\ 2 & 2 & 4 \end{bmatrix} \overset{S_1 \leftrightarrow S_2}{\Longrightarrow} \begin{bmatrix} 1 & 1 & 2 \\ 1 & 3 & 0 \\ 2 & 2 & 4 \end{bmatrix} \overset{\substack{S_2-S_1 \\ S_3-2S_1}}{\Longrightarrow} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & -2 \\ 0 & 0 & 0 \end{bmatrix} \overset{-\frac{1}{2}S_2}{\Longrightarrow}$$
$$\begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \overset{S_1-3S_2}{\Longrightarrow} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Bár különböző úton, de mindkétszer azonos eredményre jutottunk. Valóban, hamarosan be fogjuk látni, hogy a redukált lépcsős alak egyértelmű, míg e példából is látjuk, hogy a lépcsős alak nem: a megadott mátrixnak a megoldás során négy különböző lépcsős alakját is előállítottuk. □

Gauss – Jordan-módszer A *Gauss – Jordan-módszer*, más néven *Gauss – Jordan-kiküszöbölés* vagy *Gauss – Jordan-elimináció* a lineáris egyenletrendszerek egy megoldási módszere. Lényege, hogy a lineáris egyenletrendszer bővített mátrixát elemi sorműveletekkel *redukált* lépcsős alakra hozzuk. Ebből az alakból azonnal leolvasható a megoldás. Adjunk új megoldást a Gauss-módszernél bemutatott egyenletrendszerekre.

2.48. PÉLDA (GAUSS – JORDAN-MÓDSZER, EGY MEGOLDÁS). *Oldjuk meg a 2.40. példában felírt egyenletrendszert Gauss – Jordan-módszerrel!*

MEGOLDÁS. Felírjuk az egyenletrendszer bővített mátrixát, és a 2.40. példában látott módon eljutunk a lépcsős alakhoz, majd folytatjuk:

$$\left[\begin{array}{ccc|c}1&1&2&0\\2&2&3&2\\1&3&3&4\\1&2&1&5\end{array}\right] \implies \ldots \implies \left[\begin{array}{ccc|c}1&1&2&0\\0&2&1&4\\0&0&-1&2\\0&0&0&0\end{array}\right] \overset{\substack{\frac{1}{2}S_2\\-S_3}}{\implies} \left[\begin{array}{ccc|c}1&1&2&0\\0&1&\frac{1}{2}&2\\0&0&1&-2\\0&0&0&0\end{array}\right] \overset{S_1-S_2}{\implies}$$

$$\left[\begin{array}{ccc|c}1&0&\frac{3}{2}&-2\\0&1&\frac{1}{2}&2\\0&0&1&-2\\0&0&0&0\end{array}\right] \overset{\substack{S_2-\frac{1}{2}S_3\\S_1-\frac{3}{2}S_3}}{\implies} \left[\begin{array}{ccc|c}1&0&0&1\\0&1&0&3\\0&0&1&-2\\0&0&0&0\end{array}\right] \implies \begin{aligned}x\quad &= 1\\ y\quad &= 3\\ z &= -2\end{aligned}$$

Tehát az egyenletrendszer egyetlen megoldása $(x, y, z) = (1, 3, -2)$. □

2.49. PÉLDA (GAUSS–JORDAN-MÓDSZER, VÉGTELEN SOK MEGOLDÁS). *Oldjuk meg a 2.41. példabeli*

$$\begin{aligned}x_1 + 2x_2 + x_3 + 2x_4 + x_5 &= 1\\ x_1 + 2x_2 + 3x_3 + 3x_4 + x_5 &= 0\\ 3x_1 + 6x_2 + 7x_3 + 8x_4 + 3x_5 &= 1\end{aligned}$$

egyenletrendszert Gauss–Jordan-módszerrel!

MEGOLDÁS. A 2.41. példában eljutottunk egy lépcsős alakig. Az eljárást folytatjuk, míg a redukált lépcsős alakra nem jutunk.

$$\left[\begin{array}{ccccc|c}1&2&1&2&1&1\\1&2&3&3&1&0\\3&6&7&8&3&1\end{array}\right] \implies \ldots \implies \left[\begin{array}{ccccc|c}1&2&1&2&1&1\\0&0&2&1&0&-1\\0&0&0&0&0&0\end{array}\right] \overset{\substack{1/2S_2\\S_1-S_2}}{\implies}$$

$$\left[\begin{array}{ccccc|c}1&2&0&3/2&1&3/2\\0&0&1&1/2&0&-1/2\\0&0&0&0&0&0\end{array}\right] \implies \begin{aligned}x_1 + 2x_2 \quad + \frac{3}{2}x_4 + x_5 &= \frac{3}{2}\\ x_3 + \frac{1}{2}x_4 \quad &= -\frac{1}{2}\end{aligned}$$

Végezzük el az $x_2 = s$, $x_4 = t$, $x_5 = u$ helyettesítést, az x_1 és x_3 változók azonnal kifejezhetők. Így a megoldás:

$$(x_1, x_2, x_3, x_4, x_5) = (\frac{3}{2} - 2s - \frac{3}{2}t - u, s, -\frac{1}{2} - \frac{1}{2}t, t, u),$$

vagy mátrixjelöléssel

$$\begin{bmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{bmatrix} = \begin{bmatrix}\frac{3}{2} - 2s - \frac{3}{2}t - u\\ s\\ -\frac{1}{2} - \frac{1}{2}t\\ t\\ u\end{bmatrix} = \begin{bmatrix}\frac{3}{2}\\0\\-\frac{1}{2}\\0\\0\end{bmatrix} + s\begin{bmatrix}-2\\1\\0\\0\\0\end{bmatrix} + t\begin{bmatrix}-\frac{3}{2}\\0\\-\frac{1}{2}\\1\\0\end{bmatrix} + u\begin{bmatrix}-1\\0\\0\\0\\1\end{bmatrix}.$$

Természetesen ugyanazt a megoldást kaptuk, mint a 2.41. példában. □

A redukált lépcsős alak egyértelműsége Fontos következményei vannak a következő tételnek:

2.50. TÉTEL (A REDUKÁLT LÉPCSŐS ALAK EGYÉRTELMŰ). *Minden mátrix redukált lépcsős alakra hozható, amely egyértelmű.*

BIZONYÍTÁS. Tegyük fel indirekt módon, hogy van egy olyan mátrix, mely elemi sorműveletekkel két különböző redukált lépcsős alakra hozható. Jelölje ezeket $\mathbf{R}$ és $\mathbf{S}$. Mivel mindketten ugyanazzal a mátrixszal ekvivalensek, elemi sorműveletekkel egymásba alakíthatóak, vagyis egymással is ekvivalensek. Válasszuk ki oszlopaik közül azt a balról első oszlopot, melyben különböznek, valamint az összes előttük álló vezéroszlopot. Az így kapott mátrixokat jelölje $\hat{\mathbf{R}}$ és $\hat{\mathbf{S}}$. Tehát $\hat{\mathbf{R}} \neq \hat{\mathbf{S}}$, mert különböznek az utolsó oszlopukban. Például, ha

$$\mathbf{R} = \begin{bmatrix} 1 & 2 & 0 & 4 & 5 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{és} \quad \mathbf{S} = \begin{bmatrix} 1 & 2 & 0 & 4 & 5 \\ 0 & 0 & 1 & 9 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

akkor

$$\hat{\mathbf{R}} = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{és} \quad \hat{\mathbf{S}} = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 9 \\ 0 & 0 & 0 \end{bmatrix}.$$

Ez az oszlop, melyben különböznek, nem lehet az első oszlop, mert ha az a zérusvektor az egyik mátrixban, akkor a sorekvivalencia miatt a másikban is az lenne, egyébként pedig ez az oszlop mindenképp az első helyen 1-est, alatta 0-kat tartalmaz.

Tekintsük az így kapott $\hat{\mathbf{R}}$, $\hat{\mathbf{S}}$ mátrixokat egy-egy egyenletrendszer bővített együtthatómátrixának. Ezek általános alakja tehát a következő:

$$\hat{\mathbf{R}} = \left[\begin{array}{cccc|c} 1 & 0 & \dots & 0 & r_1 \\ 0 & 1 & \dots & 0 & r_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & r_k \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \end{array}\right] \quad \text{vagy} \quad \hat{\mathbf{R}} = \left[\begin{array}{cccc|c} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \end{array}\right]$$

és

$$\hat{\mathbf{S}} = \left[\begin{array}{cccc|c} 1 & 0 & \dots & 0 & s_1 \\ 0 & 1 & \dots & 0 & s_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & s_k \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \end{array}\right] \quad \text{vagy} \quad \hat{\mathbf{S}} = \left[\begin{array}{cccc|c} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \end{array}\right]$$

Mivel oszlopok kihagyása nem változtat a sorekvivalencián – hisz elemi sorműveletekben műveletet csak egy oszlopon belül végzünk –, ezért az $\hat{\mathbf{R}}$ és $\hat{\mathbf{S}}$ mátrixok ekvivalensek, azaz a hozzájuk tartozó két egyenletrendszernek ugyanaz a megoldása. Ez csak úgy lehet, ha vagy minden $i = 1, \dots, k$ indexre $r_i = s_i$, vagy egyik egyenletrendszer sem oldható meg, azaz mindkét esetben azt kaptuk, hogy $\hat{\mathbf{R}} = \hat{\mathbf{S}}$, ami ellentmondás. Ez az ellentmondás bizonyítja, hogy a kiinduló $\mathbf{R} \neq \mathbf{S}$ feltevés helytelen volt, tehát $\mathbf{R} = \mathbf{S}$. (E bizonyítás Holzmann[3] Interneten publikált cikkén alapul). □

[3] Wolf Holzmann. Uniqueness of reduced row echelon form. http://www.cs.uleth.ca/~holzmann/notes/reduceduniq.pdf, 2002

Mivel a redukált lépcsős alak egyértelmű, definiálhatunk egy függvényt, mely minden mátrixhoz annak ezt az alakját rendeli. Az $\operatorname{rref}(\mathbf{A})$ jelölést mi arra a függvényre fogjuk alkalmazni, mely egy $m \times n$-es r-rangú mátrixhoz a redukált lépcsős alakjának a zérussorok elhagyásával kapott alakját rendeli, mely egy $r \times n$-es mátrix. Például

$$\operatorname{rref}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Szimultán egyenletrendszerek Gyakori feladat az alkalmazásokban, hogy sok olyan egyenletrendszert kell megoldani, amelyek csak a konstans tagokban térnek el egymástól. A kiküszöböléses módszerekkel ezek egyszerre is megoldhatók alig több erőforrás felhasználásával, mint ami egyetlen egyenletrendszer megoldásához szükséges.

2.51. DEFINÍCIÓ (SZIMULTÁN EGYENLETRENDSZEREK). *Több egyenletrendszer halmazát* szimultán egyenletrendszernek *nevezünk, ha együtthatómátrixaik azonosak.*

2.52. PÉLDA (SZIMULTÁN EGYENLETRENDSZER MEGOLDÁSA). *Oldjuk meg az alábbi egyenletrendszereket!*

$$\begin{aligned} x + y + z &= 3 \\ 2x + 3y + 2z &= 7 \\ 2x + 2y + 3z &= 6 \end{aligned} \qquad \begin{aligned} u + v + w &= 3 \\ 2u + 3v + 2w &= 7 \\ 2u + 2v + 3w &= 7 \end{aligned} \qquad \begin{aligned} r + s + t &= 0 \\ 2r + 3s + 2t &= 0 \\ 2r + 2s + 3t &= 1 \end{aligned}$$

MEGOLDÁS. Mivel e három egyenletrendszer együtthatómátrixa azonos, a bal oldal átalakítását elég egyszer elvégezni, a jobb oldalak átalakítását pedig vele együtt. Ehhez a szimultán egyenletrendszerre a következő bővített mátrixot érdemes képezni:

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 3 & 3 & 0 \\ 2 & 3 & 2 & 7 & 7 & 0 \\ 2 & 2 & 3 & 6 & 7 & 1 \end{array}\right]$$

A megoldáshoz használjuk a Gauss–Jordan-módszert:

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 3 & 3 & 0 \\ 2 & 3 & 2 & 7 & 7 & 0 \\ 2 & 2 & 3 & 6 & 7 & 1 \end{array}\right] \overset{\substack{S_2-2S_1 \\ S_3-2S_1}}{\Longrightarrow} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 3 & 3 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{array}\right] \overset{\substack{S_1-S_2 \\ S_1-S_3}}{\Longrightarrow}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & 1 & -1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{array}\right],$$

és ebből leolvasható mindhárom egyenletrendszer megoldása:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} r \\ s \\ t \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

□

Ha tudjuk, hogy több egyenletrendszerből álló szimultán egyenletrendszerről van szó, mindegyik egyenletrendszerben használhatjuk ugyanazokat a változókat.

2.53. PÉLDA (SZIMULTÁN EGYENLETRENDSZER BŐVÍTETT MÁTRIXA). *Oldjuk meg azt a szimultán egyenletrendszert, melynek bővített mátrixa*

$$\left[\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 5 & 3 & 0 & 1 \end{array}\right]$$

MEGOLDÁS. A Gauss – Jordan-módszer lépései:

$$\left[\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 5 & 3 & 0 & 1 \end{array}\right] \xRightarrow{S_2-\frac{5}{2}S_1} \left[\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 0 & \frac{1}{2} & -\frac{5}{2} & 1 \end{array}\right] \xRightarrow{2S_2} \left[\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 0 & 1 & -5 & 2 \end{array}\right] \xRightarrow{S_1-S_2}$$

$$\left[\begin{array}{cc|cc} 2 & 0 & 6 & -2 \\ 0 & 1 & -5 & 2 \end{array}\right] \xRightarrow{\frac{1}{2}S_1} \left[\begin{array}{cc|cc} 1 & 0 & 3 & -1 \\ 0 & 1 & -5 & 2 \end{array}\right].$$ □

*Kiküszöbölés $\mathbb{Z}_p$-ben** Ha p prím, akkor a modulo p maradékosztályok közti műveletek rendelkeznek minden olyan tulajdonsággal, melyet a kiküszöbölés során a valós számok körében használtunk. Ennek következtében a Gauss- és Gauss – Jordan-módszerek minden további nélkül használhatók $\mathbb{Z}_p$ fölötti egyenletrendszerekre is. (Lásd még a 447. oldalon az algebrai testről írtakat.)

2.54. PÉLDA (EGYENLETRENDSZER $\mathbb{Z}_2$ FÖLÖTT). *4-bites kódszavakat küldünk, bitjeit jelölje a, b, c és d. Hibajavító kódot készítünk úgy, hogy minden kódszó végére három paritásbitet teszünk, nevezetesen a $b+c+d$, $a+c+d$ és $a+b+d$ bitet Az összeadás itt természetesen $\mathbb{Z}_2$ fölött értendő. Például a* 0110 *kódszó helyett a* 0110011 *kódszót küldjük. Egy üzenetben az egyik ilyen 7-bites kódszó első 4 bitjét a vevő szerkezet bizonytalanul érzékeli, amit kapunk, az a* $(?,?,?,?,1,0,1)$ *kódvektor. Mi lehetett az eredeti üzenet, ha az utolsó 3 bit biztosan jó?*

MEGOLDÁS. Az a, b, c és d bitek ismeretlenek, csak annyit tudunk, hogy

$$\begin{aligned} b+c+d&=1 \\ a+\quad c+d&=0 \\ a+b+\quad d&=1 \end{aligned}$$

Oldjuk meg ezt az egyenletrendszert Gauss – Jordan kiküszöböléssel $\mathbb{Z}_2$ fölött. Ne felejtsük, hogy $\mathbb{Z}_2$-ben $1+1=0$, így $1=-1$, azaz a kivonás nem különbözik az összeadástól.

$$\left[\begin{array}{cccc|c} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \end{array}\right] \xRightarrow{S_1\leftrightarrow S_2} \left[\begin{array}{cccc|c} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \end{array}\right] \xRightarrow{S_3+S_1} \left[\begin{array}{cccc|c} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{array}\right] \xRightarrow{S_3+S_2}$$

$$\left[\begin{array}{cccc|c} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right] \xRightarrow[S_1+S_3]{S_2+S_3} \left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right]$$

Visszaírva egyenletrendszerré:

$$\begin{aligned} a\quad +c\quad &=0 \\ b+c\quad &=1 \\ d&=0 \end{aligned}$$

Az utolsó egyenletből $d = 0$, a másodikból $b = 1 - c$, azaz $b = 1 + c$ és az elsőből $a = -c$, azaz $a = c$. Eszerint c szabad változó, legyen értéke s, a többi kötött. A megoldás általános alakban $(a, b, c, d) = (s, 1 + s, s, 0)$, azaz $(a, b, c, d) = (0, 1, 0, 0) + s(1, 1, 1, 0)$. Az $s = 0$ és az $s = 1$ értékekhez tartozó megoldások tehát: $(0, 1, 0, 0)$ és $(1, 0, 1, 0)$.

Ha az egyenletrendszert vektoregyenletnek tekintjük, akkor az első megoldás azt mutatja, hogy az együtthatómátrix második oszlopa megegyezik a jobb oldallal (és valóban), a második megoldás pedig azt, hogy az első és a harmadik oszlop összege a jobb oldalt adja.

Megjegyezzük e kódról, melyet $[7, 4, 3]_2$ bináris *Hamming-kódnak* neveznek, hogy a kód 16 szóból áll, bármely szavának bármely 4 bitje egyértelműen meghatározza a maradék hármat. Így ha legföljebb 3 bit megváltozik egy szóban, akkor az kimutatható, és ha csak egy bit változik meg, az kijavítható. □

2.55. PÉLDA (EGYENLETRENDSZER $\mathbb{Z}_5$ FÖLÖTT). *Oldjuk meg az alábbi két egyenletrendszert $\mathbb{Z}_5$ fölött.*

$$\begin{aligned} 2x + 3y &= 1 \qquad & 2x + 3y &= 1 \\ 3x + 2y &= 4 \qquad & 3x + 4y &= 3 \end{aligned}$$

MEGOLDÁS. A számolás megkönnyítésére vagy készítsünk osztási táblát, vagy használjuk a 447. oldalon található A.8. ábra szorzástábláját.

$$\left[\begin{array}{cc|c} 2 & 3 & 1 \\ 3 & 2 & 4 \end{array}\right] \overset{3S_1}{\Longrightarrow} \left[\begin{array}{cc|c} 1 & 4 & 3 \\ 3 & 2 & 4 \end{array}\right] \overset{S_2-3S_1}{\Longrightarrow} \left[\begin{array}{cc|c} 1 & 4 & 3 \\ 0 & 0 & 0 \end{array}\right],$$

azaz az egyenletrendszernek több megoldása van. Itt ez nem azt jelenti, hogy végtelen sok, hanem azt, hogy legalább egy paraméter végigfut $\mathbb{Z}_5$ összes elemén. Szabad változó az y, legyen $y = s$, így $x = 3 - 4s = 3 + s$, tehát $(x, y) = (3 + s, s)$, azaz a vektorok mátrixjelölésével:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} + s \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad s \in \mathbb{Z}_5$$

Mivel $\mathbb{Z}_5$-nek öt eleme van, ezért s-nek is ennyi értéke lehet, azaz az első egyenletrendszer összes megoldása $(3, 0)$, $(4, 1)$, $(0, 2)$, $(1, 3)$, $(2, 4)$. A másik egyenletrendszer megoldása:

$$\left[\begin{array}{cc|c} 2 & 3 & 1 \\ 3 & 4 & 3 \end{array}\right] \overset{3S_1}{\Longrightarrow} \left[\begin{array}{cc|c} 1 & 4 & 3 \\ 3 & 4 & 3 \end{array}\right] \overset{S_2-3S_1}{\Longrightarrow} \left[\begin{array}{cc|c} 1 & 4 & 3 \\ 0 & 2 & 4 \end{array}\right],$$

amiből $2y = 4$, azaz $y = 2$, $x + 4 \cdot 2 = 3$, azaz $x = 0$. Tehát a megoldás $(x, y) = (0, 2)$. □

Feladatok

2.14.• LÉPCSŐS ALAK: IGAZ – HAMIS Melyek igazak, melyek hamisak az alábbi állítások közül?

a) Egy mátrix minden lépcsős alakjában ugyanannyi nemzérus sor van.

b) Egy mátrix minden lépcsős alakjában ugyanannyi főoszlop (bázisoszlop) van.

c) Minden valós mátrixnak van lépcsős alakja, ami egyértelmű.

d) Különböző mátrixoknak különböző a redukált lépcsős alakjuk.

e) Ha egy mátrix elemi sorműveletekkel egy másikba vihető, akkor redukált lépcsős alakjuk megegyezik.

2.15.• EGYENLETRENDSZEREK: IGAZ – HAMIS Melyek igazak, melyek hamisak az alábbi állítások közül?

a) Elemi sorműveletek közben az egyenletrendszer megoldáshalmaza nem változik.

b) Egy lineáris egyenletrendszer nem konzisztens (nincs megoldása), ha több egyenletből áll, mint ahány ismeretlenes (más szóval, ha túlhatározott).

c) Ha egy valósegyütthatós lineáris egyenletrendszernek van két különböző megoldása, akkor végtelen sok is van.

d) Egy homogén lineáris egyenletrendszer mindig megoldható.

Határozzuk meg valamely lépcsős alakját, majd a redukált lépcsős alakját az alábbi mátrixoknak!

2.16. $\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 3 & 2 & 3 & 4 \\ 1 & 2 & 1 & 2 & 3 \end{bmatrix}$

2.17. $\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 3 & 2 & 3 & 4 \\ 3 & 2 & 1 & 2 & 3 \end{bmatrix}$

Ekvivalensek-e az alábbi egyenletrendszerek?

2.18. $\begin{cases} 3x+2y-2z=1 \\ 2x+3y-3z=-1 \\ 4x+2y =8 \end{cases}$ $\qquad \begin{cases} 2x+2y-2z=0 \\ 3x+3y-2z=3 \\ 5x-3y+2z=5 \end{cases}$

2.19. $\begin{cases} 2x+3y+5z=0 \\ 3x+2y+2z=3 \\ 5x -4z=9 \end{cases}$ $\qquad \begin{cases} x-y-3z=3 \\ 5x+5y+7z=3 \end{cases}$

2.20. Csak egész számokkal számolva megoldható-e az az egyenletrendszer, melynek bővített mátrixa a következő:

a) $\begin{bmatrix} 3 & 4 & 1 & 1 \\ 7 & 8 & 3 & 7 \\ 11 & 7 & -2 & 2 \end{bmatrix}$ $\qquad$ *b)* $\begin{bmatrix} 3 & 4 & 1 & 1 \\ 7 & 8 & 3 & 7 \\ 11 & 7 & 2 & 2 \end{bmatrix}$

2.21. SORMŰVELETEK REVERZIBILITÁSA Mutassuk meg, hogy minden elemi sorművelettel átalakított mátrixhoz van egy olyan sorművelet, mely azt visszaalakítja.

Megoldás a gyakorlatban

Bár e szakasz tartalma elsősorban nem a lineáris algebra, hanem a numerikus analízis témakörébe tartozik, ismerete elengedhetetlen annak, aki a gyakorlatban lineáris algebrai eszközöket alkalmaz. Először a Gauss- és Gauss–Jordan-kiküszöbölés műveletigényét, majd numerikus megbízhatóságának kérdését vizsgáljuk. Ezután az iterációs módszerek lényegét vázoljuk, melyek alkalmazásakor az együtthatómátrix nem változik, így a számítási hibák sem halmozódnak. Ráadásul e módszerek a ritka mátrixokat sem „rontják el", mint a Gauss-módszer, mely sok zérust ír fölül.

A kiküszöbölés műveletigénye Ahhoz, hogy a lineáris egyenletrendszerek különböző megoldási módszereit össze tudjuk hasonlítani, azt is tudnunk kell, mennyi a műveletigényük. A flop mértékegységről részletesen a függelékben írunk a 440. oldalon.

2.56. tétel (A kiküszöbölés műveletigénye). *A Gauss- és a Gauss–Jordan-módszer műveletigénye egy n-ismeretlenes, n egyenletből álló egyenletrendszer esetén egyaránt*

$$\frac{n^3}{3}+\frac{n^2}{2}-\frac{5n}{6}\ \text{összeadás/kivonás},\quad \frac{n^3}{3}+n^2-\frac{n}{3}\ \text{szorzás/osztás}.$$

azaz összesen

$$\frac{2}{3}n^3+\frac{3}{2}n^2-\frac{7}{6}n\,\text{flop},$$

azaz jó közelítéssel $2n^3/3$ flop.

Bizonyítás. Először felelevenítünk két elemi összefüggést, amire a bizonyításban szükség van:

$$1+2+\ldots+n=\frac{n(n+1)}{2},\quad 1^2+2^2+\ldots+n^2=\frac{n(n+1)(2n+1)}{6}.$$

A továbbiakban feltételezzük, hogy a kiküszöbölés során a főátlóba kerülő elemek egyike sem 0. A Gauss-módszernél a főátló alatti elemek eliminálásához $\frac{1}{3}n^3-\frac{1}{3}n$ összeadás és $\frac{1}{3}n^3+\frac{1}{2}n^2-\frac{5}{6}n$ szorzás szükséges. A visszahelyettesítés $\frac{1}{2}n^2-\frac{1}{2}n$ összeadásból és $\frac{1}{2}n^2+\frac{1}{2}n$ szorzásból áll. Ha a Gauss–Jordan-módszernél a főátló alatti elemek kiküszöbölése mellett a főátló elemeit is 1-re változtatjuk, az $\frac{1}{3}n^3-\frac{1}{3}n$ összeadás mellett $\frac{1}{3}n^3+\frac{1}{2}n^2+\frac{1}{6}n$ szorzás szükséges. A főátló feletti elemek eliminálásához $\frac{1}{2}n^2-\frac{1}{2}n$ összeadás és ugyanennyi szorzás kell. A számítások részletezését az olvasóra bízzuk. □

Numerikusan instabil egyenletrendszerek A gyakorlati feladatokban gyakran mérési eredményekkel, így nem pontos adatokkal dolgozunk.

2.57. PÉLDA (INSTABIL EGYENLETRENDSZER). *Oldjuk meg a következő egyenletrendszert!*

$$\begin{aligned} 6.73x - 8.97y &= 5.61 \\ 4.79x - 6.39y &= 3.99 \end{aligned}$$

Mutassuk meg, hogy az együtthatók 0.01*-dal való megváltoztatása a megoldások nagy megváltozását okozhatja, sőt az is elérhető, hogy az egyenletrendszernek ne legyen, vagy épp végtelen sok megoldása legyen!*

MEGOLDÁS. Az egyenletrendszer megoldása: $x = 1.5$, $y = 0.5$. Az első egyenletben az x együtthatóját újra mérjük, és másodszorra egy századdal kevesebbnek, 6.72-nek adódik. Oldjuk meg ezt az egyenletrendszert is! Az eredmény meglepő módon nagyon messze van az előzőtől: $x \approx -2.26$, $y \approx -2.32$. Újabb mérés az y együtthatóját -8.96-nak mutatja. E két együttható egy századdal való megváltozása után a megoldás messze van mindkét előző eredménytől: $x \approx 4.35$, $y \approx 2.64$. Ha végül az első egyenlet konstans tagját is megváltoztatjuk egy századdal 5.62-re, akkor $x \approx 7.21$, $y \approx 4.78$ lesz az eredmény, ha pedig 5.60-ra, akkor – csemegeként – ismét a kerek $x = 1.5$, $y = 0.5$ értékeket kapjuk.

A fenti egyenletrendszeren tovább változtatva az együtthatókat az is elérhető, hogy végtelen sok megoldása legyen:

$$\begin{aligned} 6.72x - 8.96y &= 5.60 \\ 4.80x - 6.40y &= 4.00 \end{aligned}$$

ugyanis itt a két egyenlet egymás konstansszorosa. Ha pedig a második egyenlet konstans tagját visszaírjuk 3.99-re, egy ellentmondó egyenletrendszert kapunk. □

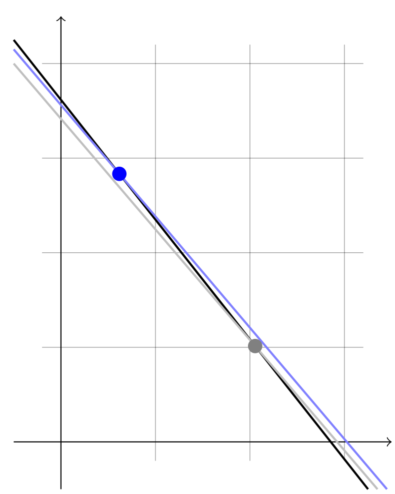
Ilyen megbízhatatlan eredményekkel a gyakorlatban semmit nem lehet kezdeni!

Az olyan egyenletrendszert, melyben az együtthatók vagy a konstans tagok kis változása a megoldásban nagy változást okoz, *numerikusan instabilnak* vagy *rosszul kondicionáltnak* nevezzük. Egyébként *numerikusan stabil*, illetve *jól kondicionált* egyenletrendszerről beszélünk.

Világos, hogy a fentiek nem precíz matematikai fogalmak. Később precízen definiálva számmal fogjuk mérni a kondicionáltság fokát, de azt, hogy egy adott egyenletrendszer megoldásai elfogadhatóak-e vagy nem, csak a feladat döntheti el.

A numerikus instabilitás okát szemlélteti a 2.12. ábra. Kétváltozós egyenletrendszerek esetén, ha a két egyenes grafikonja „közel" van egymáshoz, azaz majdnem egybe esnek, akkor kis változások az egyeneseken messze vihetik a metszéspontot, de párhuzamossá is tehetik a két egyenest.

Ha a gyakorlatban numerikusan instabil egyenletrendszerrel találkozunk, vizsgáljuk meg, hogy az egyenleteink közti „majdnem" lineá-



2.12. ábra: Instabil egyenletrendszer, melyben az egyenletek együtthatóinak kis megváltoztatása a megoldás nagy megváltozását okozza.

ris összefüggőség mögött nem valódi lineáris összefüggőség van-e kis mérési hibával.

Részleges főelemkiválasztás A következő példákban csak tízes számrendszerbeli aritmetikát használunk. A számításokat úgy kell elvégezni, hogy az adott pontosságnak megfelelően minden részeredményt p értékes jegyre kerekítünk.

2.58. PÉLDA (GAUSS-MÓDSZER LEBEGŐPONTOS SZÁMOKKAL). *Oldjuk meg az alábbi – numerikusan stabil – egyenletrendszert pontosan, majd 3 értékes jegy pontossággal számolva.*

$$\begin{aligned} 10^{-4}x + y &= 2 \\ x - y &= 0 \end{aligned}$$

MEGOLDÁS. Pontosan számolva

$$\left[\begin{array}{cc|c} 10^{-4} & 1 & 2 \\ 1 & -1 & 0 \end{array}\right] \overset{S_2 - 10^4 S_1}{\Longrightarrow} \left[\begin{array}{cc|c} 10^{-4} & 1 & 2 \\ 0 & -1-10^4 & -2\cdot 10^4 \end{array}\right]$$

amiből az eredmény $x = y = \frac{2\cdot 10^4}{1+10^4}$. Igazolható, hogy az egyenletrendszer numerikusan stabil, ami azt jelenti, hogy például 10^{-4} helyébe 0-t helyettesítve, vagyis kicsit változtatva egy együtthatót, a kapott

$$\left[\begin{array}{cc|c} 0 & 1 & 2 \\ 1 & -1 & 0 \end{array}\right]$$

egyenletrendszer megoldása csak kicsit különbözik az előzőtől: $x = y = 2$. Végezzük most el a Gauss-kiküszöbölést 3 értékes jeggyel számolva:

$$\left[\begin{array}{cc|c} 10^{-4} & 1 & 2 \\ 1 & -1 & 0 \end{array}\right] \overset{S_2 - 10^4 S_1}{\Longrightarrow} \left[\begin{array}{cc|c} 10^{-4} & 1 & 2 \\ 0 & -1-10^4 & -2\cdot 10^4 \end{array}\right] \approx \left[\begin{array}{cc|c} 10^{-4} & 1 & 2 \\ 0 & -10^4 & -2\cdot 10^4 \end{array}\right],$$

ahol a közelítésnél a $\mathrm{fl}(-1-10^4) = -10^4$ összefüggést használtuk. Az így kapott egyenletrendszernek viszont $x = 0$, $y = 2$ a megoldása, ami nagyon messze van az eredeti egyenletrendszer megoldásától! Most végezzünk egy apró változtatást: először cseréljük fel a két egyenletet!

$$\left[\begin{array}{cc|c} 1 & -1 & 0 \\ 10^{-4} & 1 & 2 \end{array}\right] \overset{S_2 - 10^{-4} S_1}{\Longrightarrow} \left[\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 1+10^{-4} & 2 \end{array}\right] \approx \left[\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 1 & 2 \end{array}\right],$$

amelynek megoldása $x = y = 2$, ami nagyon közel van a pontos megoldáshoz! □

Mi az oka a két megoldás közti különbségnek, és fel tudnánk-e használni minél jobb megoldás megtalálásában?

Mindkét megoldásban az első egyenlet konstansszorosát hozzáadtuk a második egyenlethez, de az első esetben a kisebb, a másodikban az első oszlop nagyobb elemét választottuk főelemnek. Amikor a kisebbet választottuk, akkor az első sort egy kis számmal osztottuk, vagyis repciprokával – egy nagy számmal – szoroztuk, és ezt adtuk a második sorhoz. A nagy számmal való beszorzás következtében a második egyenlet együtthatóit „elnyomták” e nagy számok, nagyon megváltoztatva az egyenletet, aminek következtében a megoldások is nagyon megváltoztak! A $\mathrm{fl}(-1-10^4) = -10^4$ kerekítés hatása, vagyis a -1 „eltüntetése”, ekvivalens azzal, mintha az eredeti egyenletrendszer helyett a következőt kéne megoldani:

$$\left[\begin{array}{cc|c} 10^{-4} & 1 & 2 \\ 1 & 0 & 0 \end{array}\right].$$

És ennek valóban $x = 0$, $y = 2$ a megoldása! Amikor viszont az első oszlop nagyobbik elemét választottuk főelemnek, a sort egy kis számmal kellett beszorozni, és ezt hozzáadni a másik sorhoz, vagyis kerekítéskor az eredeti egyenlet együtthatói megmaradtak, az egyenletrendszer kevésbé torzult. Ennek alapján megfogalmazható egy széles körben elterjedt szabály: a Gauss-féle kiküszöbölési eljárás során, lebegőpontos adatokkal dolgozva minden oszlopban a szóbajöhető elemek közül – sorcserék segítségével – mindig a legnagyobb abszolút értékűt válasszuk főelemnek! E módszert *részleges főelemkiválasztásnak*, illetve *részleges pivotálásnak* nevezik.

Bizonyos – a gyakorlatban ritkán előforduló – esetekben jobb eredmény kapható a *teljes főelemkiválasztás* módszerével. Ekkor főelemnek az összes még hátralévő elem abszolút értékben legnagyobbikát választjuk. Itt oszlopcserékre is szükség van, és műveletigényesebb is ez az eljárás, ezért ritkán alkalmazzák.

2.59. PÉLDA (RÉSZLEGES FŐELEMKIVÁLASZTÁS). *Részleges főelemkiválasztással hozzuk lépcsős alakra az alábbi mátrixot!*

$$\begin{bmatrix} 1.8 & 3.0 & 3.0 & 3.7 & 7.5 \\ 3.6 & 3.2 & 3.6 & 6.2 & 7.8 \\ 7.2 & 3.6 & 4.8 & 2.4 & 1.2 \\ 2.4 & 5.4 & 5.2 & 2.6 & 5.2 \end{bmatrix}$$

MEGOLDÁS. Az első oszlop legnagyobb eleme a harmadik sorban van, így az első és a harmadik sor cseréjével kezdünk:

$$\xRightarrow{S_1 \leftrightarrow S_3} \begin{bmatrix} \mathbf{7.2} & 3.6 & 4.8 & 2.4 & 1.2 \\ 3.6 & 3.2 & 3.6 & 6.2 & 7.8 \\ 1.8 & 3.0 & 3.0 & 3.7 & 7.5 \\ 2.4 & 5.4 & 5.2 & 2.6 & 5.2 \end{bmatrix} \xRightarrow{\substack{S_2 - S_1/2 \\ S_3 - S_1/4 \\ S_4 - S_1/3}} \begin{bmatrix} \mathbf{7.2} & 3.6 & 4.8 & 2.4 & 1.2 \\ 0.0 & 1.4 & 1.2 & 5.0 & 7.2 \\ 0.0 & 2.1 & 1.8 & 3.1 & 7.2 \\ 0.0 & \mathbf{4.2} & 3.6 & 1.8 & 4.8 \end{bmatrix}$$

$$\xRightarrow{S_1 \leftrightarrow S_3} \begin{bmatrix} \mathbf{7.2} & 3.6 & 4.8 & 2.4 & 1.2 \\ 0.0 & \mathbf{4.2} & 3.6 & 1.8 & 4.8 \\ 0.0 & 2.1 & 1.8 & 3.1 & 7.2 \\ 0.0 & 1.4 & 1.2 & 5.0 & 7.2 \end{bmatrix} \xRightarrow{\substack{S_3 - S_2/2 \\ S_4 - S_2/3}} \begin{bmatrix} \mathbf{7.2} & 3.6 & 4.8 & 2.4 & 1.2 \\ 0.0 & \mathbf{4.2} & 3.6 & 1.8 & 4.8 \\ 0.0 & 0.0 & 0.0 & 2.2 & 4.8 \\ 0.0 & 0.0 & 0.0 & \mathbf{4.4} & 5.6 \end{bmatrix}$$

$$\xRightarrow{S_3 \leftrightarrow S_4} \begin{bmatrix} \mathbf{7.2} & 3.6 & 4.8 & 2.4 & 1.2 \\ 0.0 & \mathbf{4.2} & 3.6 & 1.8 & 4.8 \\ 0.0 & 0.0 & 0.0 & \mathbf{4.4} & 5.6 \\ 0.0 & 0.0 & 0.0 & 2.2 & 4.8 \end{bmatrix} \xRightarrow{S_4 - S_3/2} \begin{bmatrix} \mathbf{7.2} & 3.6 & 4.8 & 2.4 & 1.2 \\ 0.0 & \mathbf{4.2} & 3.6 & 1.8 & 4.8 \\ 0.0 & 0.0 & 0.0 & \mathbf{4.4} & 5.6 \\ 0.0 & 0.0 & 0.0 & 0.0 & \mathbf{2.0} \end{bmatrix} \quad \square$$

Skálázás A részleges főelemkiválasztásban mindig az oszlop legnagyobb elemét választottuk. Nem lehet egy elemet nagyobbá tenni, és ezzel az egész módszert elrontani úgy, hogy egy sorát egyszerűen beszorozzuk?

2.60. példa (Sor szorzása). *A 2.58. példában szorozzuk meg az első egyenletet 10^5-nel, azaz a kisebb elemből csináljunk nagyot, és ezt az egyenletrendszert is oldjuk meg részleges főelemkiválasztással.*

$$\begin{aligned} 10x + 10^5 y &= 2 \cdot 10^5 \\ x - \quad y &= \quad 0 \end{aligned}$$

Megoldás. Egy egyenlet beszorzása egy nemzérus számmal ekvivalens átalakítás, így ennek az egyenletrendszernek is $x = y = \frac{2 \cdot 10^4}{1+10^4}$ a pontos megoldása. Ha 3 értékes jegyre számolunk, és alkalmazzuk a részleges főelemkiválasztás módszerét, akkor ismét rossz eredményt kapunk:

$$\left[\begin{array}{cc|c} 10 & 10^5 & 2 \cdot 10^5 \\ 1 & -1 & 0 \end{array}\right] \xRightarrow{S_2 - 10S_1} \left[\begin{array}{cc|c} 10 & 10^5 & 2 \cdot 10^5 \\ 0 & -1 - 10^4 & -2 \cdot 10^4 \end{array}\right] \approx \left[\begin{array}{cc|c} 10^{-4} & 1 & 2 \\ 0 & -10^4 & -2 \cdot 10^4 \end{array}\right],$$

amiből $x = 0$ és $y = 2$. $\square$

Hasonlóképp zavart okozhat az együtthatómátrix egy oszlopának beszorzása is, ami az egyenletrendszeren például úgy valósítható meg, ha egyik változó mértékegységét megváltoztatjuk. (Ha például a korábban kilométerben meghatározott ismeretlent milliméterben keressük, együtthatóját minden egyenletben 10^6-nal kell osztani.)

Az együtthatók ilyen „egyenetlenségeiből” származó számítási hibák csökkentésére a *skálázás* nevű gyakorlati módszer ajánlható. Ez a következő két skálázási szabály követéséből áll, mely a tapasztalatok szerint a gyakorlati feladatok nagy részében nagyon jó eredményt ad a részleges főelemkiválasztással együtt alkalmazva:

1. *Oszlopok skálázása:* Válasszunk a feladatban szereplő mennyiségeknek természetes mértékegységet, ezzel általában elkerülhetők az együtthatók közti tetemes nagyságrendi különbségek. Ezen kívül nincs szükség az oszlopok elemeinek beszorzására.
2. *Sorok skálázása:* Az egyenletrendszer $[\mathbf{A}|\mathbf{b}]$ bővített mátrixának minden sorát osszuk el az $\mathbf{A}$ együtthatómátrix adott sorbeli legnagyobb abszolút értékű elemével. Így $\mathbf{A}$ minden sorának 1 a legnagyobb eleme.

Nem ismeretes olyan módszer, mely a lebegőpontos ábrázolás korlátai mellett hatékonyan megtalálná a lehető legpontosabb eredményt. Az elmélet és a tapasztalatok alapján sűrű, nem túlzottan nagy méretű egyenletrendszerekre a skálázott főelem kiválasztásos Gauss-módszer ajánlható. A ritka egyenletrendszerekre a következőkben ismertetendő iteratív módszerek általában jobb eredményt adnak.

Iteratív módszerek A továbbiakban is csak olyan egyenletrendszerekkel foglalkozunk, melyek n-ismeretlenesek és n egyenletből állnak, tehát melyek együtthatómátrixa négyzetes.

Az iteratív módszerek lényege, hogy olyan

$$\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_k, \ldots$$

vektorsorozatot generálunk, mely az adott egyenletrendszer megoldásvektorához konvergál. Első pillanatra meglepőnek tűnhet egy végtelen sorozat generálásával keresni a megoldást, de mivel számításaink eleve csak véges pontosságúak, gyakran igen kevés lépésben elérhetjük a megkívánt pontosságot. Ráadásul a kerekítési hibák még növelhetik is a konvergencia sebességét.

A kiindulási pont – a matematika több más területén is gyümölcsöző módszer – a fixpontkeresés. Ennek lényegét egy egyváltozós függvény példáján mutatjuk be. Legyen f egy minden valós helyen értelmezett függvény, mely bármely két a és b pontot két olyan pontba visz, melyek távolsága a és b távolságának legföljebb a fele. Képletben:

$$|f(b) - f(a)| \leq \frac{1}{2}|b - a|, \text{ azaz } \frac{|f(b) - f(a)|}{|b - a|} \leq \frac{1}{2}.$$

Ez azt jelenti, hogy f összes különbségi hányadosa legfeljebb $1/2$. A sokkal általánosabban megfogalmazható Banach-féle fixpont tétel szerint ekkor egyetlen olyan $\bar{x}$ pont létezik, hogy $\bar{x} = f(\bar{x})$, és ez megkapható úgy, hogy egy tetszőleges x_0 pontból kiindulva képezzük az

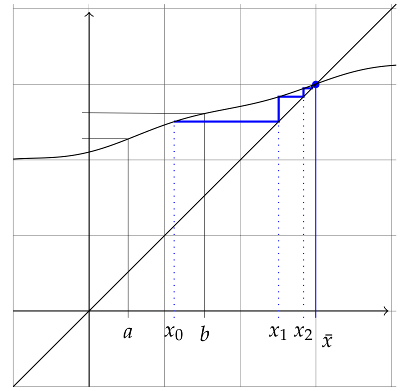
$$x_0,\ x_1 = f(x_0),\ x_2 = f(x_1), \ldots,\ x_{k+1} = f(x_k), \ldots$$

sorozatot, és vesszük a határértékét. Ekkor

$$\bar{x} = \lim_{k\to\infty} x_k.$$

A 2.13. ábra szemlélteti a fenti állítást. Az $1/2$-es szorzó kicserélhető tetszőleges 0 és 1 közé eső konstansra.

A Banach fixponttétel könnyen szemléltethető hétköznapi módon: képzeljük el, hogy egy nagyobb gumilapot néhányan körbeállva egy kerek asztal tetején széthúznak az asztal széléig, majd (most jön a leképezés!) visszaengedik eredeti állapotába. Ekkor igaz az, hogy az asztalon pontosan egy olyan pont van, mely fölött a gumilap helyben marad. E pont megkapható, ha kiválasztunk az asztalon egy tetszőleges P_0 pontot, és megnézzük, hogy a kinyújtott gumilap e fölötti pontja összehúzódáskor hová ugrik, legyen ez a P_1 pont az asztalon. A kinyújtott gumilap P_1 fölötti pontja összehúzódáskor az P_2 pont fölé ugrik, stb. Az így kapott pontsorozat a fixponthoz konvergál.



2.13. ábra: Egy függvény, mely bármely a és b pontot két olyan pontba visz, melyek távolsága a és b távolságának legföljebb a fele, így a függvény minden különbségi hányadosa abszolút értékben legföljebb $1/2$. E függvénynek pontosan egy fixpontja van, mely megkapható egy tetszőleges x_0 pontból induló $x_k = f(x_{k-1})$ sorozat határértékeként.

Jacobi-iteráció Az előző paragrafusban leírtakat követve megpróbáljuk az egyenletrendszert átrendezni úgy, hogy az $\mathbf{x} = f(\mathbf{x})$ alakú legyen, ahol $\mathbf{x}$ jelöli az ismeretlenek vektorát.

2.61. PÉLDA (JACOBI-ITERÁCIÓ). *Oldjuk meg a*

$$\begin{aligned} 4x - y &= 2 \\ 2x - 5y &= -8 \end{aligned}$$

egyenletrendszert, majd hozzuk $\mathbf{x} = f(\mathbf{x})$ *alakra, és egy tetszőleges* $\mathbf{x}_0$ *vektorból indulva végezzünk az* $\mathbf{x}_{k+1} = f(\mathbf{x}_k)$ *formulával iterációt. Számoljunk 3 tizedes pontossággal. Hová tart az így kapott vektorsorozat?*

MEGOLDÁS. Az egyenletrendszert kiküszöböléssel megoldva kapjuk, hogy $\mathbf{x} = (1, 2)$ az egyetlen megoldás.

Hozzuk az egyenletrendszert $\mathbf{x} = f(\mathbf{x})$, azaz $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] = f(\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right])$ alakra. Erre több lehetőség is adódik. Ezek közül talán az a legkézenfekvőbb, hogy az első egyenletből az x-et, a másodikból y-t kifejezzük:

$$\begin{aligned} x &= \frac{y+2}{4} \\ y &= \frac{2x+8}{5} \end{aligned}$$

Válasszunk egy $\mathbf{x}_0$ vektort tetszőlegesen, legyen pl. $\mathbf{x}_0 = (0, 0)$, azaz $x = y = 0$. A fenti képletekbe helyettesítve kapjuk, hogy $\mathbf{x}_1 = (\frac{0+2}{4}, \frac{0+8}{5}) = (0.5, 1.6)$. A további értékeket egy táblázatban adjuk meg:

	$\mathbf{x}_0$	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_3$	$\mathbf{x}_4$	$\mathbf{x}_5$	$\mathbf{x}_6$	$\mathbf{x}_7$	$\mathbf{x}_8$
x	0	0.5	0.9	0.95	0.99	0.995	0.999	1.000	1.000
y	0	1.6	1.8	1.96	1.98	1.996	1.998	2.000	2.000

E példa esetén tehát a végtelen sorozat konvergensnek mutatkozott, de a kerekítési hiba folytán véges sok lépés után megtalálta a konvergenciapontot. □

Az általános eset hasonlóan írható le. Tegyük fel, hogy az

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\
\vdots \qquad \vdots \qquad\qquad \vdots \quad &\quad \vdots \\
a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= b_n
\end{aligned}$$

egyenletrendszer egyértelműen megoldható és főátlójának minden eleme különbözik 0-tól. A *Jacobi-iteráció* menete tehát a következő. A k-adik egyenletből fejezzük ki az x_k változót:

$$\begin{aligned}
x_1 &= \frac{1}{a_{11}}(b_1 \qquad\qquad - a_{12}x_2 - \ldots - a_{1,n-1}x_{n-1} - a_{1n}x_n) \\
x_2 &= \frac{1}{a_{22}}(b_2 - a_{21}x_1 \qquad\qquad - \ldots - a_{2,n-1}x_{n-1} - a_{2n}x_n) \\
&\;\;\vdots \\
x_n &= \frac{1}{a_{nn}}(b_n - a_{n1}x_1 - a_{n2}x_2 - \ldots - a_{n,n-1}x_{n-1} \qquad\qquad)
\end{aligned} \tag{2.25}$$

Válasszunk az ismeretlenek $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ vektorának egy $\mathbf{x}_0$ kezdőértéket, pl. legyen $\mathbf{x}_0 = (0, 0, \ldots, 0)$. A (2.25) egyenletrendszer jobb oldalába helyettesítsük be $\mathbf{x}_0$ koordinátáinak értékét, a bal oldal adja $\mathbf{x}_1$ koordinátáit. Ezt a lépést ismételjük meg, generálva az $\mathbf{x}_2$, $\mathbf{x}_3$,... vektorokat addig, míg el nem érjük a megfelelő pontosságot.

Gauss – Seidel-iteráció A Jacobi-iteráció gyorsaságán könnyen javíthatunk, ha a (2.25) egy egyenletének jobb oldalába való behelyettesítés után a bal oldalon kapott változó értékét azonnal fölhasználjuk, nem várunk vele a ciklus végéig. Ezt a módosított algoritmust nevezzük *Gauss – Seidel-iterációnak*.

2.62. PÉLDA (GAUSS – SEIDEL-ITERÁCIÓ). *Oldjuk meg a*

$$\begin{aligned}
4x - y &= 2 \\
2x - 5y &= -8
\end{aligned}$$

egyenletrendszert Gauss – Seidel-iterációval.

MEGOLDÁS. Itt is, mint a Jacobi-iterációnál az

$$\begin{aligned}
x &= \frac{y+2}{4} \\
y &= \frac{2x+8}{5}
\end{aligned}$$

egyenleteket használjuk, de míg a Jacobi-iterációnál $\mathbf{x}_0 = (0, 0)$ kezdőérték után az $x = \frac{0+2}{4} = \frac{1}{2}$, $y = \frac{0+8}{5} = \frac{8}{5}$ értékek következtek, a Gauss–Seidel-iterációnál a második egyenletben 0 helyett már az első egyenletben kiszámolt $x = \frac{1}{2}$ értéket helyettesítjük, azaz $y = \frac{2\frac{1}{2}+8}{5} = \frac{9}{5}$. A további értékeket egy táblázatban adjuk meg de úgy, hogy jelezzük a kiszámítás sorrendjét:

	$\mathbf{x}_0$	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_3$	$\mathbf{x}_4$
x	0	0.5	0.95	0.995	1.000
y	0	1.8	1.98	1.998	2.000

Hasonlítsuk össze az eredményt a Jacobi iterációnál készített táblázattal. □

Mindkét iteráció a 2-ismeretlenes esetben jól szemléltethető. A **??**. és a **??**. ábra

Az iterációk konvergenciája A fenti példákból nem látszik, hogy vajon a Jacobi- és a Gauss–Seidel-iterációk mindig konvergálnak-e.

2.63. példa (Divergens iteráció). *Oldjuk meg Jacobi- és Gauss–Seidel-iterációval a következő egyenletrendszert:*

$$
\begin{aligned}
x - y &= 2 \\
2x - y &= 5
\end{aligned}
$$

Megoldás. Alakítsuk át az egyenletrendszert:

$$
\begin{aligned}
x &= y + 2 \\
y &= 2x - 5
\end{aligned}
$$

Először próbálkozzunk Jacobi-iterációval:

	$\mathbf{x}_0$	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_3$	$\mathbf{x}_4$	$\mathbf{x}_5$	$\mathbf{x}_6$	$\mathbf{x}_7$	$\mathbf{x}_8$
x	0	2	-3	1	-9	-1	-21	-5	-45
y	0	-5	-1	-11	-3	-23	-7	-47	-15

Úgy tűnik, nem konvergens a vektorsorozat, mint ahogy nem tűnik annak a Gauss–Seidel-iterációnál sem:

	$\mathbf{x}_0$	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_3$	$\mathbf{x}_4$	
x	0	2	1	-1	-5	-13
y	0	-1	-3	-7	-15	-31

A divergencia leolvasható az iterációkat szemléltető ábrákról is! □

2.64. DEFINÍCIÓ (SORONKÉNT DOMINÁNS FŐÁTLÓJÚ MÁTRIX). *Azt mondjuk, hogy az $n \times n$-es $\mathbf{A}$ mátrix* soronként (szigorúan) domináns főátlóval rendelkezik, *vagy* soronként (szigorúan) domináns főátlójú, *ha a főátló minden eleme abszolút értékben nagyobb a sorában lévő többi elem abszolút értékeinek összegénél, azaz képletben*

$$\begin{aligned}
|a_{11}| &> & & |a_{12}| + \dots + |a_{1,n-1}| + & & |a_{1n}| \\
|a_{22}| &> |a_{21}| & & + \dots + |a_{2,n-1}| + & & |a_{2n}| \\
&\vdots & & \ddots & & \vdots \\
|a_{n-1,n-1}| &> |a_{n-1,1}| + |a_{n-1,2}| + \dots & & & & + |a_{n-1,n}| \\
|a_{nn}| &> |a_{n1}| + & & |a_{n2}| + \dots + |a_{n,n-1}| & &
\end{aligned}$$

Hasonlóan definiálható az oszloponként domináns főátlójú mátrix is.

Világos, hogy az alábbi mátrixok soronként domináns főátlójúak:

$$\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} -10 & 1 & 2 \\ 1 & 10 & -3 \\ -1 & -2 & 10 \end{bmatrix}, \quad \begin{bmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -5 \end{bmatrix}, \quad \begin{bmatrix} 1 & .25 & .25 & .25 \\ .25 & 1 & .25 & .25 \\ .25 & .25 & 1 & .25 \\ .25 & .25 & .25 & 1 \end{bmatrix}.$$

Az alábbi mátrixok nem soronként domináns főátlójúak, de sorcserékkel azzá tehetők:

$$\begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}, \quad \begin{bmatrix} -1 & -2 & 10 \\ -10 & 1 & 2 \\ 1 & 10 & -3 \end{bmatrix}, \quad \begin{bmatrix} 0 & -3 & 0 \\ 0 & 0 & -5 \\ 2 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} .25 & .25 & .25 & 1 \\ 1 & .25 & .25 & .25 \\ .25 & .25 & 1 & .25 \\ .25 & 1 & .25 & .25 \end{bmatrix}.$$

Az alábbi egyenletrendszer együtthatómátrixa soronként domináns főátlójú:

$$\begin{aligned}
4x - y &= 11 \\
2x - 5y &= -17
\end{aligned}$$

2.65. TÉTEL (ELÉGSÉGES FELTÉTEL AZ ITERÁCIÓK KONVERGENCIÁJÁRA). *Ha az n egyenletből álló n-ismeretlenes egyenletrendszer együtthatómátrixa soronként domináns főátlójú, akkor bármely indulóvektor esetén a Jacobi- és a Gauss–Seidel-iteráció is konvergens.*

E tételt nem bizonyítjuk. Megjegyezzük, hogy a tételbeli feltétel nem szükséges, csak elégséges, azaz olyan egyenletrendszeren is konvergens lehet valamelyik iteráció, melynek nem domináns főátlójú az együtthatómátrixa. Hasonló tétel igaz oszloponként domináns főátlójú együtthatómátrixok esetén is.

A domináns főátlójú mátrixokon a Gauss–Seidel-iteráció sosem lassabb, mint a Jacobi-iteráció, sőt, gyakran érezhetően gyorsabb. Az viszont előfordulhat, hogy a Gauss–Seidel-iteráció divergens, míg a Jacobi-iteráció konvergens (ld. 2.23. feladat).

A gyakorlatban ezeknek az iterációknak különböző, hatékonyabb javításait használják. E témában az olvasó figyelmébe ajánljuk a „numerikus módszerek” témában írt könyveket, web-oldalakat.

Feladatok

Jacobi-iteráció

2.22. Oldjuk meg a

$$\begin{aligned} 4x - y &= 8 \\ 2x - 5y &= -5 \end{aligned}$$

egyenletrendszert Jacobi-iterációval! Számoljunk 3, majd 4 értékes jegyre!

Vegyes feladatok

2.23. JACOBI-ITERÁCIÓ KONVERGÁL, GAUSS – SEIDEL-ITERÁCIÓ NEM Írjunk programot annak az állításnak az ellenőrzésére, hogy a

$$\begin{aligned} x \phantom{{}+5/6y} &+ z = 0 \\ -x + 5/6y &\phantom{{}+z} = 0 \\ x + 2y &- 3z = 1 \end{aligned}$$

egyenletrendszeren a Jacobi-iteráció konvergál, a Gauss – Seidel-iteráció nem.

2.24. GAUSS-ELIMINÁCIÓ DOMINÁNS FŐÁTLÓJÚ MÁTRIXON Bizonyítsuk be, hogy ha az $\mathbf{A}$ mátrix főátlója soronként domináns, akkor végrehajtható rajta a főelemkiválasztásos Gauss-elimináció sorcsere nélkül!

2.25. AZ ITERÁCIÓK SZEMLÉLTETÉSE A A városból elindul egy A jelű vonat a B város felé, vele egyidőben a B városból egy B jelű A felé. A B vonat indulásával egyidőben a B vonat orráról elindul egy légy is A felé, de amint találkozik az A vonattal megfordul, és addig repül, míg a B vonattal nem találkozik, amikor ismét megfordul, stb. Mindhármuk sebessége konstans, de a légy sebessége nagyobb mindkét vonaténál.

1. Egy táblázatban megadjuk mindkét vonat távolságát az indulási helyüktől km-ben mérve azokban a pillanatokban, amikor a légy épp a B vonattal találkozik.

	(x_0, y_0)	(x_1, y_1)	(x_2, y_2)	(x_3, y_3)
x: távolság A-tól	0	40	48	49.6
y: távolság B-től	0	80	96	99.2

   Számítsuk ki a táblázat egy-két további oszlopát!

2. Milyen messze van A város B-től, ha a légy sebessége 200 km/h?

3. Most egy másik táblázatban megadjuk annak a vonatnak a távolságát az indulási helyétől, amelyik épp találkozik a léggyel:

	y_0	x_1	y_1	x_2	y_2	x_3	y_3
x: távolság A-tól		30		46		49.2	
y: távolság B-től	0		80		96		99.2

   Számítsuk ki a táblázat egy-két további oszlopát!

4. Milyen messze van A város B-től, ha a légy sebessége 200 km/h?

5. Mi köze van e feladatnak a Jacobi- és a Gauss – Seidel-iterációhoz?

3
Megoldhatóság és a megoldások tere

E fejezetet az egyenletrendszer megoldásainak jellemzésére szánjuk. Ennek során megismerkedünk az altér fogalmával és a velük végezhető legegyszerűbb műveletekkel. Végül megmutatjuk, hogy minden konzisztens lineáris egyenletrendszer origóhoz legközelebbi megoldása az egyetlen, mely a sortérbe esik, és bármely két megoldásának különbsége merőleges rá.

Homogén és inhomogén egyenletrendszerek megoldásai

Az előzőekben a megoldások megtalálásának módszereit tanulmányoztuk. E szakaszban a megoldhatóság kérdését és a megoldások halmazának legfontosabb tulajdonságait vizsgáljuk. A vizsgálatokban a lineáris egyenletrendszerek mindkét geometriai interpretációja fontos szerepet kap.

Kötött változók száma, mátrix rangja A redukált lépcsős alak egyértelműségének egy nyilvánvaló, de fontos folyománya az alábbi eredmény:

3.1. KÖVETKEZMÉNY (FŐELEMEK OSZLOPAI). *Egy valós mátrix bármely lépcsős alakjában a főelemek ugyanazokban az oszlopokban vannak, tehát ezek száma is független a lépcsős alaktól.*

A bizonyítás azonnal adódik abból, hogy bármely lépcsős alak főelemeiből kapjuk a redukált lépcsős alak vezéregyeseit, így bármely lépcsős alak főelemei ugyanott vannak, ahol a vezéregyesek, a redukált lépcsős alak pedig egyértelmű.

Ebből az is következik, hogy bármely valós mátrix esetén

$$\boxed{\begin{array}{c}\text{bármely lépcsős}\\ \text{alak főelemeinek}\\ \text{száma}\end{array}} = \boxed{\begin{array}{c}\text{bármely lépcsős}\\ \text{alak nemzérus}\\ \text{sorainak száma}\end{array}} = \boxed{\begin{array}{c}\text{a redukált lépcsős}\\ \text{alak vezéregye-}\\ \text{seinek száma.}\end{array}}$$

Ez a következő definícióhoz vezet.

3.2. DEFINÍCIÓ (MÁTRIX RANGJA). *Egy mátrix valamely lépcsős alakjában a nemnulla sorok számát a mátrix* rangjának *nevezzük. Az* $\mathbf{A}$ *mátrix rangját* $\mathrm{r}(\mathbf{A})$ *jelöli.*

3.3. PÉLDA (MÁTRIX RANGJÁNAK KISZÁMÍTÁSA). *Számítsuk ki az alábbi mátrixok rangját!*

$$\begin{bmatrix} 2 & 3 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 3 & 2 & 1 \\ 0 & 0 & 4 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}.$$

MEGOLDÁS. Az első és második mátrix lépcsős alakú, rangjuk azonnal leolvasható: 1 és 2. A harmadik és negyedik mátrix elemi sorműveletekkel

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -2 & 0 & -2 \\ 0 & 0 & -2 & 2 \\ 0 & 0 & 0 & -4 \end{bmatrix}, \text{ illetve } \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

alakra hozható, tehát a rang 4, illetve 2. □

3.4. ÁLLÍTÁS (KÖTÖTT ÉS SZABAD VÁLTOZÓK SZÁMA). *Ha egy n-ismeretlenes egyenletrendszer megoldható, és együtthatómátrixának rangja r, akkor a Gauss- vagy a Gauss–Jordan-módszerrel kapott megoldásában a kötött változók száma r, a szabad változók száma $n-r$.*

Fontos megjegyezni, hogy egyelőre csak annyit tudunk bizonyítani, hogy ha az együtthatómátrix és a bővített mátrix rangja r, akkor az egyenletrendszernek *van olyan megoldása*, amelyben a kötött változók száma r, a szabad változóké $n-r$, és egy ilyen megoldás megkapható a Gauss- vagy a Gauss–Jordan-módszerrel. Ez abból következik, hogy a rang megegyezik a főelemek számával, az pedig a kötött változók számával. Mi történik azonban, ha más módszerrel keresünk megoldást, vagy ha például felcseréljük az ismeretlenek sorrendjét? Arról még nem tudunk semmit, hogy a mátrix rangja vajon megváltozik-e, ha fölcseréljük az oszlopait, a változók sorrendjének cseréje pedig oszlopcserét okoz az együtthatómátrixon. A következő fejezetben be fogjuk azt is látni, hogy a kötött és szabad változók száma független a változók sorrendjétől, sőt a megoldás módszerétől is.

3.5. PÉLDA (KÖTÖTT ÉS SZABAD VÁLTOZÓK SZÁMA). *Egy lineáris egyen-*

letrendszer bővített együtthatómátrixának egy lépcsős alakja a következő:

$$\left[\begin{array}{ccccccc|c} 0.145 & 0.016 & 0.755 & 0.649 & 0.625 & 0.426 & 0.859 & 0.040 \\ 0.000 & 0.901 & 0.015 & 0.397 & 0.076 & 0.690 & 0.007 & 0.755 \\ 0.000 & 0.000 & 0.000 & 0.969 & 0.085 & 0.889 & 0.304 & 0.296 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.368 & 0.027 & 0.497 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \end{array}\right].$$

Látható, hogy a mátrix rangja és így kötött változóinak száma 4, az egyenletrendszer ismeretleneinek száma 7, a szabad változók száma 3.

Egyenletrendszer megoldhatóságának feltétele Tudjuk, hogy egy lineáris egyenletrendszer pontosan akkor *nem* oldható meg, ha a bővített mátrix lépcsős alakjának van olyan sora, melyben csak a legutolsó elem nem nulla. Ez ugyanis egyenletté visszaírva $0 = c$ alakú, ahol $c \neq 0$, és ennek az egyenletnek nincs megoldása. Ez viszont azt jelenti, hogy ilyenkor a bővített mátrix rangja nagyobb az együtthatómátrix rangjánál. E megállapítás azonnali következménye a következő tétel.

3.6. TÉTEL (A MEGOLDHATÓSÁG MÁTRIXRANGOS FELTÉTELE). *Legyen egy n-ismeretlenes lineáris egyenletrendszer együtthatómátrixa $\mathbf{A}$, a konstans tagokból álló vektora $\mathbf{b}$.*

1. *Ez az egyenletrendszer* pontosan akkor oldható meg, *ha együtthatómátrixának és bővített mátrixának rangja megegyezik, azaz*

$$\mathrm{r}(\mathbf{A}) = \mathrm{r}(\mathbf{A}|\mathbf{b}).$$

2. *Ez az egyenletrendszer pontosan akkor oldható meg* egyértelműen, *ha együtthatómátrixának és bővített mátrixának rangja megegyezik az ismeretlenek számával, azaz*

$$\mathrm{r}(\mathbf{A}) = \mathrm{r}(\mathbf{A}|\mathbf{b}) = n.$$

BIZONYÍTÁS. Az egyenletrendszer pontosan akkor oldható meg, ha bővített mátrixának lépcsős alakjában nincs olyan sor, melynek csak az utolsó eleme nem 0. Ez épp azt jelenti, hogy $\mathrm{r}(\mathbf{A}) = \mathrm{r}(\mathbf{A}|\mathbf{b})$.

A második állítás abból következik, hogy az egyenletrendszer akkor oldható meg egyértelműen, ha megoldható, és nincs szabad változója, vagyis az együtthatómátrix rangja megegyezik az ismeretlenek számával. □

Az előzőekből az is adódik, hogy egy lineáris egyenletrendszernek pontosan akkor van egynél több megoldása, ha

$$\mathrm{r}(\mathbf{A}) = \mathrm{r}(\mathbf{A}|\mathbf{b}) < n.$$

(Miért nem lehet $\mathrm{r}(\mathbf{A}) > n$?)

Egy valós együtthatós egyenletrendszernek csak úgy lehet egynél több megoldása, ha van szabad változója. Viszont, annak minden értékéhez egy-egy másik megoldás tartozik, vagyis ekkor az egyenletrendszernek végtelen sok megoldása van. Így, ha $\mathbf{A}$ valós mátrix, akkor a megoldások száma, a két rang és az ismeretlenek száma közt a következő a kapcsolat:

Feltétel	Megoldások száma
$\mathrm{r}(\mathbf{A}) < \mathrm{r}(\mathbf{A}\|\mathbf{b})$	0
$\mathrm{r}(\mathbf{A}) = \mathrm{r}(\mathbf{A}\|\mathbf{b}) = n$	1
$\mathrm{r}(\mathbf{A}) = \mathrm{r}(\mathbf{A}\|\mathbf{b}) < n$	∞

Ha az egyenletrendszer homogén lineáris, azaz mindegyik konstans tag 0, akkor az elemi sorműveletek közben a bővített mátrix utolsó oszlopában minden elem 0 marad, így ebben az oszlopban biztosan nem lesz főelem. Eszerint a homogén lineáris egyenletrendszerek mindig megoldhatók, hisz ekkor a $\mathrm{r}(\mathbf{A}) = \mathrm{r}(\mathbf{A}|\mathbf{b})$ összefüggés mindig fönnáll. A megoldhatóság persze e feltétel ellenőrzése nélkül is látszik, hisz az $x_1 = x_2 = \cdots = x_n = 0$ mindig megoldás! Mivel $\mathrm{r}(\mathbf{A})$ megegyezik a redukált lépcsős alak főelemeinek számával, ezért $\mathrm{r}(\mathbf{A}) \leq m$ és $\mathrm{r}(\mathbf{A}) \leq n$ is fönnáll, ahol m az egyenletek, n az ismeretlenek száma. Így viszont $m < n$ esetén $\mathrm{r}(\mathbf{A}) = n$ nem állhat fönn, tehát a homogén lineáris egyenletrendszernek van az $\mathbf{x} = \mathbf{0}$ vektoron kívül is megoldása. Ezzel bizonyítottuk az alábbi tételt:

3.7. TÉTEL (HOMOGÉN LINEÁRIS EGYENLETRENDSZER MEGOLDHATÓSÁGA). *Az* $\mathbf{A}$ *együtthatómátrixú homogén lineáris egyenletrendszer mindig megoldható, mert a nullvektor – az ún.* triviális megoldás *– mindig megoldás. Pontosan akkor van* nemtriviális, *vagyis a* $\mathbf{0}$*-vektortól különböző megoldása is, ha*

$$\mathrm{r}(\mathbf{A}) < n,$$

ahol n az ismeretlenek – azaz $\mathbf{A}$ *oszlopainak – számát jelöli. Speciálisan, az m egyenletből álló homogén lineáris egyenletrendszernek $m < n$ esetén mindig van nemtriviális megoldása.*

Valós együtthatós homogén lineáris egyenletrendszerekre az előző táblázat a következő alakot ölti:

Feltétel	Megoldások száma
$\mathrm{r}(\mathbf{A}) = n$	1
$\mathrm{r}(\mathbf{A}) < n$	∞

3.8. PÉLDA (EGYENLETRENDSZER MEGOLDÁSAINAK SZÁMA). *Az a paraméter mely értékei mellett van az alábbi egyenletrendszernek* 0, 1, *illetve* ∞

sok megoldása?

$$\begin{aligned} x_1 + x_2 + ax_3 &= 1 \\ x_1 + ax_2 + x_3 &= a \\ ax_1 + x_2 + x_3 &= a^2 \end{aligned}$$

MEGOLDÁS. Hozzuk a bővített mátrixot lépcsős alakra:

$$\left[\begin{array}{ccc|c} 1 & 1 & a & 1 \\ 1 & a & 1 & a \\ a & 1 & 1 & a^2 \end{array}\right] \overset{S_2-S_1}{\underset{}{\overset{S_3-aS_1}{\Longrightarrow}}} \left[\begin{array}{ccc|c} 1 & 1 & a & 1 \\ 0 & a-1 & 1-a & a-1 \\ 0 & 1-a & 1-a^2 & a^2-a \end{array}\right] \overset{S_3+S_2}{\Longrightarrow}$$

$$\left[\begin{array}{ccc|c} 1 & 1 & a & 1 \\ 0 & a-1 & 1-a & a-1 \\ 0 & 0 & -(a-1)(a+2) & (a+1)(a-1) \end{array}\right]$$

Látható, hogy $a = 1$ esetén az utolsó két sorban minden elem 0, tehát az együtthatómátrix és a bővített mátrix rangja is 1, így az egyenletrendszer az $x_1 + x_2 + x_3 = 1$ egyenlettel ekvivalens. Ennek megoldása: $(x_1, x_2, x_3) = (1 - s - t, s, t)$, azaz oszlopvektor alakba írva:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Ha $a = -2$, akkor az együtthatómátrix rangja 2, a bővített mátrix rangja 3, tehát az egyenletrendszer nem oldható meg (az utolsó sor egyenletté visszaírva $0 = 3$ alakú). Minden egyéb esetben, azaz ha $a \neq 1$ és $a \neq -2$, akkor a két rang 3, ami megegyezik az ismeretlenek számával, tehát egyetlen megoldás van. Ez ki is fejezhető:

$$x_1 = \frac{(a+1)^2}{a+2}, \quad x_2 = \frac{1}{a+2}, \quad x_3 = -\frac{a+1}{a+2}. \qquad \square$$

Homogén lineáris egyenletrendszer megoldásai Tekintsünk egy tetszőleges homogén lineáris egyenletrendszert. Mint a 3.7. tételben láttuk, ez biztosan megoldható, és a megoldások halmazában a nullvektor benne van. Mit mondhatunk a megoldások halmazáról, ha több megoldása is van a homogén egyenletrendszernek?

3.9. ÁLLÍTÁS (MEGOLDÁSOK LINEÁRIS KOMBINÁCIÓJA). *Egy homogén lineáris egyenletrendszer megoldásainak bármely lineáris kombinációja is megoldás.*

BIZONYÍTÁS. Elég az állítást két megoldásra bizonyítani. Jelölje $\mathbf{a}_1$, $\mathbf{a}_2,\dots,\mathbf{a}_n$ az egyenletrendszer együtthatómátrixának oszlopvektorait.

Legyen $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ és $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ két tetszőleges megoldás, azaz

$$\begin{aligned}\mathbf{a}_1x_1 + \mathbf{a}_2x_2 + \ldots + \mathbf{a}_nx_n &= \mathbf{0}\\ \mathbf{a}_1y_1 + \mathbf{a}_2y_2 + \ldots + \mathbf{a}_ny_n &= \mathbf{0},\end{aligned}$$

és c, d legyen két tetszőleges skalár. Megmutatjuk, hogy ekkor $c\mathbf{x} + d\mathbf{y}$ is megoldás, ugyanis

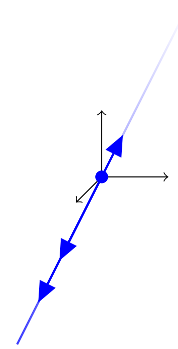
$$\begin{aligned}&\mathbf{a}_1(cx_1 + dy_1) + \mathbf{a}_2(cx_2 + dy_2) + \ldots + \mathbf{a}_n(cx_n + dy_n)\\ =&(c\mathbf{a}_1x_1 + d\mathbf{a}_1y_1) + (c\mathbf{a}_2x_2 + d\mathbf{a}_2y_2) + \ldots + (c\mathbf{a}_nx_n + d\mathbf{a}_ny_n)\\ =&c(\mathbf{a}_1x_1 + \mathbf{a}_2x_2 + \ldots + \mathbf{a}_nx_n) + d(\mathbf{a}_1y_1 + \mathbf{a}_2y_2 + \ldots + \mathbf{a}_ny_n)\\ =&\mathbf{0} + \mathbf{0} = \mathbf{0}.\end{aligned}$$

azaz $c\mathbf{x} + d\mathbf{y}$ is megoldás. Ez bizonyítja állításunkat.

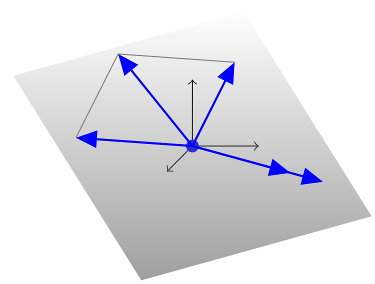
E bizonyítás az oszlopmodellre épült, de hasonlóan egyszerű bizonyítás adható a sormmodellben is (ld. 3.6. feladat). □

Altér Az, hogy vektorok egy H halmaza olyan, hogy a H-beli vektorokból képzett lineáris kombinációk is mind H-ban vannak, azzal ekvivalens, hogy bármely H-beli vektor skalárszorosa és bármely két H-beli vektor összege is H-beli (ld. a 3.14. feladatot). Másként fogalmazva: a vektorok skalárral való szorzása és a vektorok összeadása nem vezet ki H-ból.

Az $\mathbb{R}^n$ tér ilyen típusú részhalmazai igen fontosak. Például a síkban egy origón átmenő egyenes vektorai (az egyenes pontjaiba mutató helyvektorok) ilyenek. Hasonlóképp, a térben bármely origón átmenő sík vagy egyenes vektorai ilyenek (ld. a 3.1. ábra). Ez a következő definícióhoz vezet.



a)

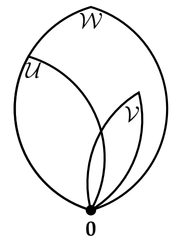


b)

3.1. ábra: *a)* Egy origón átmenő egyenes bármely vektorának konstansszorosa és bármely két vektorának összege az egyenesbe esik, *b)* egy origón átmenő sík bármely vektorának konstansszorosa és bármely két vektorának összege a síkba esik.

3.10. DEFINÍCIÓ (ALTÉR). *Az $\mathbb{R}^n$ tér vektorainak olyan részhalmazát, mely zárt a vektorok skalárral való szorzásának és a vektorok összeadásának műveletére, az $\mathbb{R}^n$* alterének *nevezzük. Képlettel kifejezve: a $H \subseteq \mathbb{R}^n$ vektorhalmaz $\mathbb{R}^n$* altere, *ha*
1. *$\mathbf{u}, \mathbf{v} \in H$ esetén $\mathbf{u} + \mathbf{v} \in H$,*
2. *$\mathbf{u} \in H$, és $c \in \mathbb{R}$ esetén $c\mathbf{u} \in H$.*

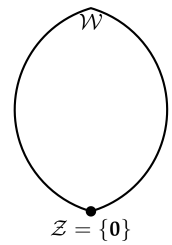
Az $\mathbb{R}^2$ vagy az $\mathbb{R}^3$ altereit könnyen tudjuk szemléltetni, ahogy azt a 3.1. ábrán is tettük. De hogyan szemléltethetők $\mathbb{R}^n$ alterei? Természetesen itt csak egy leegyszerűsítő, az elemibb összefüggések szemléltetésére használható ábrázolás jöhet szóba. A Venn-diagramok példáját követjük: az ábrázolás alakjának nem „hasonlítania” kell az ábrázolt dologra, csak tudnunk kell, hogy annak melyik tulajdonságát hogyan szemlélteti. Vektorterekre a *levéldiagramot* fogjuk használni. Ez egy egyszerű levélformával szemlélteti a teret, melynek a levél száránál lévő töve jelzi a nullvektort. Az altereket olyan kisebb levelek szemléltetik, melyek töve – azaz a nullvektor – közös (ld. 3.2. ábra). A levelek felső csúcsába a tér nevét vagy dimenzióját írhatjuk.



3.2. ábra: Egy $\mathcal{W}$ vektortér és annak $\mathcal{U}$ és $\mathcal{V}$ alterei, valamint a mindannyiukban közös nullvektor

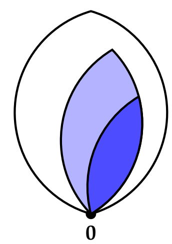
Felsoroljuk az alterek néhány egyszerűen belátható tulajdonságát:

- ▶ Minden altérnek eleme a nullvektor, hisz ha egy $\mathbf{x}$ vektor eleme az altérnek, akkor $-\mathbf{x}$ is és e két vektor összege is, azaz $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$. (Bizonyíthatunk úgy is, hogy bármely altérbeli vektorral együtt annak 0-szorosa is, vagyis a $\mathbf{0}$-vektor is eleme az altérnek.)
- ▶ Minden vektortér maga is altér, hisz bármely két vektorának összes lineáris kombinációját is tartalmazza.
- ▶ A nullvektor önmagában alteret alkot, ez a *zérustér*, amit $\mathcal{Z}$ jelöl. A nulltér kifejezést másra használjuk, ne keverjük össze. A zérusteret és a teljes teret szokás triviális altereknek nevezni (ld. 3.3. ábra).



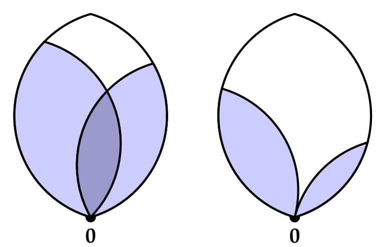
3.3. ábra: Egy $\mathcal{W}$ vektortér két triviális alterével: az egyik maga $\mathcal{W}$, a másik a nullvektort tartalmazó $\mathcal{Z}$ tér

- ▶ Altér altere is altér (ld. 3.4. ábra).



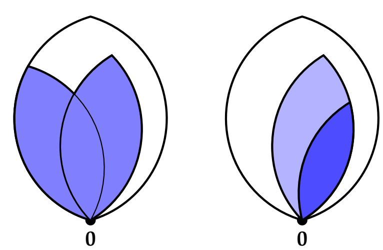
3.4. ábra: Altér altere is altér

- ▶ Két altér metszete is altér. Ha $\mathcal{U}$ és $\mathcal{V}$ egy vektortér két altere, és $\mathcal{W}$ a közös részük, akkor $\mathcal{W}$ nem üres, hisz a nullvektor benne van. Másrészt bármely két $\mathbf{x}, \mathbf{y} \in \mathcal{W}$ vektor összes lineáris kombinációja benne van $\mathcal{U}$-ban és $\mathcal{V}$-ben is, így metszetükben is. Alterek metszetére is a $\cap$ jelet használjuk, tehát az előbbi alterekre $\mathcal{U} \cap \mathcal{V} = \mathcal{W}$ (ld. 3.5. ábra).



3.5. ábra: Alterek metszete is altér, de az megeshet, hogy ez a metszet csak az egyetlen nullvektorból álló zérustér.

- ▶ Az előző gondolat tetszőleges számú altérre is megismételhető, tehát egy vektortér tetszőleges számú alterének közös része is altér. Az alterek száma végtelen is lehet.
- ▶ Két altér egyesítése csak akkor altér, ha egyik altere a másiknak. Például a térben egy origón átmenő egyenes vektorait és egy origón átmenő sík vektorait egyesítve csak akkor kapunk alteret, ha az egyenes a síkba esik.



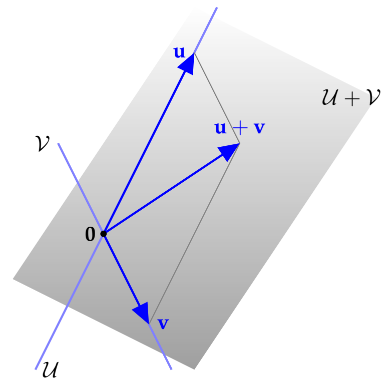
3.6. ábra: Két altér egyesítése csak akkor altér, ha egyik a másik altere

Ha $\mathcal{U}$ és $\mathcal{V}$ ugyanannak a vektortérnek az alterei, akkor az egyesítésük által generált alteret $\mathcal{U} + \mathcal{V}$-vel jelöljük, és a két altér összegének nevezzük.

3.11. ÁLLÍTÁS (ALTEREK ÖSSZEGE). *Ha $\mathcal{U}$ és $\mathcal{V}$ a $\mathcal{W}$ altér két altere, akkor az egyesítésük által generált $\mathcal{U} + \mathcal{V}$ altér pontosan azokból a vektorokból áll, melyek egy $\mathcal{U}$- és egy $\mathcal{V}$-beli vektor összegeként előállnak.*

BIZONYÍTÁS. Ha $\mathbf{x}$ egy $\mathcal{U} + \mathcal{V}$-beli vektor, akkor előáll néhány $\mathcal{U}$- és $\mathcal{V}$-beli vektor lineáris kombinációjaként. De a lineáris kombináció $\mathcal{U}$-beli vektorokat tartalmazó része egy $\mathcal{U}$-beli $\mathbf{u}$ vektort ad, míg a többi egy $\mathcal{V}$-beli $\mathbf{v}$ vektort, így $\mathbf{x} = \mathbf{u} + \mathbf{v}$. Fordítva világos, minden $\mathbf{u} + \mathbf{v}$ alakú vektor $\mathcal{U}$- és $\mathcal{V}$-beli vektorok lineáris kombinációja, tehát benne van $\mathcal{U} + \mathcal{V}$-ben. □

Szemléltetésül: ha például $\mathcal{W} = \mathbb{R}^3$, és $\mathcal{U}$ és $\mathcal{V}$ egy-egy egymástól különböző 1-dimenziós altere, akkor az egyesítésük által generált 2-dimenziós altérbe pontosan azok a vektorok tartoznak, melyek egy $\mathcal{U}$-beli $\mathbf{u}$ és egy $\mathcal{V}$-beli $\mathbf{v}$ vektor összegei (ld. 3.7 ábra).



3.7. ábra: $\mathcal{U}+\mathcal{V}$ bármely vektora előáll $\mathbf{u}+\mathbf{v}$ alakban

3.12. PÉLDA (ALTÉR). *Alteret alkotnak-e az alábbi vektorhalmazok $\mathbb{R}^3$-ben?*
1. $\{(x, y, z) \mid x = y,\ z = xy\}$,
2. $\{(s + 2t, s - 1, 2s + t) \mid s, t \in \mathbb{R}\}$.

MEGOLDÁS. Az első halmaz nem altér, mert nem elégíti ki a definícióbeli feltételeket. Például az $(1, 1, 1)$ vektor benne van e halmazban, azonban kétszerese nem, mert nem elégíti ki a $z = xy$ egyenlőséget!

A második feladatban szereplő halmazban nincs benne a nullvektor, ugyanis az $s + 2t = 0$, $s - 1 = 0$, $2s + t = 0$ egyenletrendszernek nincs megoldása, így ez a halmaz sem alkot alteret! □

Hamarosan látni fogjuk, hogy $\mathbb{R}^2$ alterei az alábbiak:
1. a zérusvektorból álló egyelemű halmaz,
2. egy origón átmenő egyenes összes vektora,
3. a sík összes vektora.

$\mathbb{R}^3$ alterei:
1. a zérusvektorból álló egyelemű halmaz,
2. egy origón átmenő egyenes összes vektora,
3. egy origón átmenő sík összes vektora,
4. a tér összes vektora.

Mivel egy egyetlen n-ismeretlenes egyenletből álló egyenletrendszer megoldáshalmaza egy $\mathbb{R}^n$-beli hipersík, ezért az origón átmenő hipersíkok is alterek. Lássunk további példákat! A 3.9. állítás az altér fogalmát használva a következő alakot ölti:

3.13. ÁLLÍTÁS (MEGOLDÁSOK ALTERE). *Egy n-ismeretlenes* homogén *lineáris egyenletrendszer megoldáshalmaza alteret alkot $\mathbb{R}^n$-ben.*

3.14. DEFINÍCIÓ (NULLTÉR). *Az $\mathbf{A}$ együtthatómátrixú homogén lineáris egyenletrendszer megoldásainak alterét az $\mathbf{A}$ mátrix* nullterének *nevezzük és $\mathcal{N}(\mathbf{A})$-val jelöljük.*

Kérdés: általában hogyan adhatjuk meg $\mathbb{R}^n$ egy tetszőleges alterét?

Kifeszített altér A homogén lineáris egyenletrendszer összes megoldását néhány vektor lineáris kombinációjaként állítottuk elő. A megoldások alterét tehát „generálja” vagy geometrikusabb szóhasználattal „kifeszíti” néhány megoldásvektor. Erre utal a következő definíció. Azt, hogy az altér szót a definícióban használhatjuk, a rákövetkező állítás bizonyítja.

3.15. DEFINÍCIÓ (KIFESZÍTETT ALTÉR). *A $\mathbf{v}_1$, $\mathbf{v}_2, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ vektorok összes*

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k$$

alakú lineáris kombinációját a $\mathbf{v}_1$, $\mathbf{v}_2, \ldots, \mathbf{v}_k$ *vektorok által* kifeszített altérnek *nevezzük, és* $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$*-val jelöljük. Képletben:*

$$\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k) = \{\, c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k : c_1, c_2 \ldots, c_k \in \mathbb{R} \,\}.$$

Valóban, az elnevezés jogos:

3.16. ÁLLÍTÁS (A KIFESZÍTETT ALTÉR ALTÉR). *A* $\mathbf{v}_1$, $\mathbf{v}_2, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ *vektorok által kifeszített* $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ *vektorhalmaz* $\mathbb{R}^n$ *egy altere.*

BIZONYÍTÁS. Be kell látni, hogy $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ bármely vektorának skalárszorosa és bármely két vektorának összege is ide tartozik. Legyen

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k, \quad \text{és} \quad \mathbf{v} = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \ldots + d_k\mathbf{v}_k$$

a $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ két tetszőleges vektora, és legyen $x \in \mathbb{R}$ tetszőleges valós. Ekkor

$$x\mathbf{u} = (xc_1)\mathbf{v}_1 + (xc_2)\mathbf{v}_2 + \ldots + (xc_k)\mathbf{v}_k \in \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k),$$

és

$$\mathbf{u} + \mathbf{v} = (c_1 + d_1)\mathbf{v}_1 + \ldots + (c_k + d_k)\mathbf{v}_k \in \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k). \qquad \square$$

3.17. PÉLDA (NULLTÉR). *Határozzuk meg az*

$$\begin{bmatrix} 1 & 2 & 1 & 2 & 1 \\ 1 & 2 & 3 & 3 & 1 \\ 3 & 6 & 7 & 8 & 3 \end{bmatrix}$$

mátrix nullterét. Keressünk véges sok olyan vektort, melyek kifeszítik e teret!

MEGOLDÁS. Az adott mátrix a 2.43. példabeli homogén lineáris egyenletrendszer együtthatómátrixa. A homogén lineáris egyenletrendszer megoldáshalmaza épp a nulltér vektoraiból álló halmaz. A nulltér vektorai tehát

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2s - \frac{3}{2}t - u \\ s \\ -\frac{1}{2}t \\ t \\ u \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -\frac{3}{2} \\ 0 \\ -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + u \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

alakba írhatók, amiből leolvasható, hogy e teret a megoldásban megadott három vektor feszíti ki. $\square$

Az inhomogén lineáris egyenletrendszer megoldásai Az inhomogén lineáris egyenletrendszer megoldásai nem alkotnak alteret. Legegyszerűbben ez onnan látszik, hogy a zérusvektor minden altérnek eleme, viszont egyetlen inhomogén egyenletrendszernek sem megoldása! Ugyanakkor az inhomogén lineáris egyenletrendszer és a hozzá tartozó homogén egyenletrendszer megoldásai közt egy igen fontos kapcsolat van.

3.18. TÉTEL (HOMOGÉN ÉS INHOMOGÉN EGYENLETRENDSZER MEGOLDÁSAI). *Az inhomogén lineáris $[\mathbf{A}|\mathbf{b}]$ mátrixú egyenletrendszer általános megoldása megkapható úgy, hogy egy partikuláris megoldásához hozzáadjuk a hozzá tartozó homogén $[\mathbf{A}|\mathbf{0}]$ mátrixú egyenletrendszer általános megoldását.*

inhomogén általános megoldása	=	inhomogén egy partikuláris megoldása	+	homogén általános megoldása

BIZONYÍTÁS. Jelölje az egyenletrendszer együtthatómátrixát $\mathbf{A}$, annak sorvektorait $\mathbf{a}_{1*}$, $\mathbf{a}_{2*}$,…,$\mathbf{a}_{m*}$, és jelölje $\mathbf{b} = (b_1, b_2, \dots, b_m)$, a konstans tagok vektorát. Legyen $\mathbf{x}$ az inhomogén egyenletrendszer egy partikuláris megoldása, és jelölje H a homogén, I az inhomogén egyenletrendszer általános megoldását. Megmutatjuk, hogy $\mathbf{x} + H = I$, ahol a bal oldali összeadást elemenként értjük.

$\mathbf{x} + H \subseteq I$: Meg kell mutatnunk, hogy ha $\mathbf{x}$-hez adjuk a H egy tetszőleges $\mathbf{y}$ elemét, az inhomogén egyenletrendszer egy megoldását kapjuk. Valóban, $\mathbf{x}$, illetve $\mathbf{y}$ eleget tesz az

$$\begin{aligned} \mathbf{a}_{i*} \cdot \mathbf{x} &= b_i, \\ \mathbf{a}_{i*} \cdot \mathbf{y} &= 0, \qquad (i = 1, 2, \dots, m) \end{aligned}$$

egyenleteknek. Ebből

$$\mathbf{a}_{i*} \cdot (\mathbf{x} + \mathbf{y}) = \mathbf{a}_{i*} \cdot \mathbf{x} + \mathbf{a}_{i*} \cdot \mathbf{y} = b_i + 0 = b_i.$$

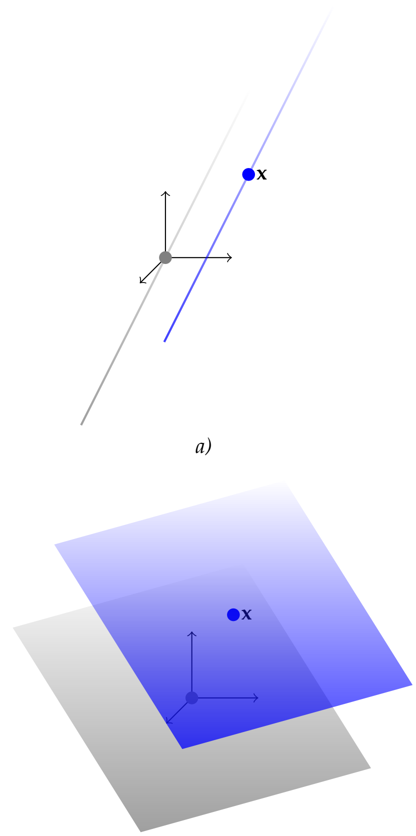
tehát $\mathbf{x} + \mathbf{y}$ megoldása az inhomogén egyenletrendszernek, azaz $\mathbf{x} + \mathbf{y} \in I$.

$\mathbf{x} + H \supseteq I$: Meg kell mutatnunk, hogy ha $\mathbf{z}$ az inhomogén egy tetszőleges megoldása, azaz $\mathbf{z} \in I$, akkor található olyan $\mathbf{y} \in H$, hogy $\mathbf{z} = \mathbf{x} + \mathbf{y}$. Valóban, az $\mathbf{y} = \mathbf{z} - \mathbf{x}$ megteszi, mert

$$\mathbf{a}_{i*} \cdot (\mathbf{z} - \mathbf{x}) = \mathbf{a}_{i*} \cdot \mathbf{z} - \mathbf{a}_{i*} \cdot \mathbf{x} = b_i - b_i = 0.$$

fennáll minden $i = 1, 2, \dots, m$ indexre, azaz $\mathbf{z} - \mathbf{x} \in H$. Ezzel kész a bizonyítás. □

E tétel azt jelenti, hogy ugyan az inhomogén lineáris egyenletrendszer megoldásainak halmaza *nem altér*, de egy *altér eltoltja*. Az ilyen

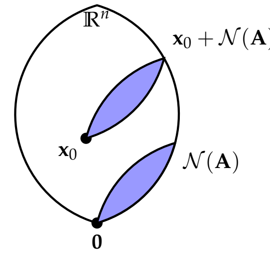


3.8. ábra: *a)* Egy háromismeretlenes inhomogén lineáris egyenletrendszer megoldáshalmaza, ha az általános megoldás egyparaméteres; *b)* Egy háromismeretlenes inhomogén lineáris egyenletrendszer megoldáshalmaza, ha az általános megoldás kétparaméteres.

halmazokat geometriai nyelven *affin altereknek* is szokás nevezni. Eltolt altereket mutat a 3.8. ábra.

E tételt szemléltetik a 2.41. és a 2.43. példák is.

Ha általában szeretnénk ábrázolni az inhomogén egyenletrendszer megoldását a levéldiagramon, azt könnyen megtehetjük egy eltolt altér szemléltetésével, ahogy azt a 3.9. ábra mutatja.

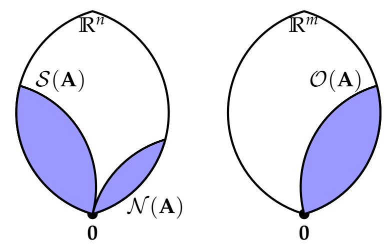


3.9. ábra: Az $\mathbf{A}$ együtthatómátrixú homogén egyenletrendszer megoldása a nulltér, azaz $\mathcal{N}(\mathbf{A})$, az inhomogéné e tér egy $\mathbf{x}_0 + \mathcal{N}(\mathbf{A})$ eltoltja, ahol $\mathbf{x}_0$ az inhomogén egyenletrendszer egy megoldása.

Az előző tétel szerint az inhomogén egyenletrendszer összes megoldása a homogén összes megoldásának – azaz $\mathcal{N}(\mathbf{A})$-nak – az inhomogén valamelyik megoldásával való eltoltja. Fontos látnunk, hogy mindegy melyik megoldást választjuk az inhomogén megoldásai közül, bár az eltolás mértéke változik, az eredmény ugyanaz lesz. Ez jól leolvasható a 3.8 ábráról: ha az origón átmenő egyenes (sík) origónál lévő pontját nem $\mathbf{x}$-be, hanem az eltolt egyenes (sík) egy másik pontjába toljuk, ugyanahhoz az eltolt altérhez jutunk.

Az inhomogén $[\mathbf{A}|\mathbf{b}]$ egyenletrendszer megoldásai nem alkotnak alteret, de azok a $\mathbf{b}$ vektorok, melyre az egyenletrendszer megoldható, igen. Ezek ugyanis az oszlopmodell szerint épp azok a vektorok, melyek az együtthatómátrix oszlopvektorainak lineáris kombinációjaként előállnak. Az ilyen vektorok pedig alteret alkotnak, mégpedig az $\mathbf{A}$ oszlopvektorai által kifeszített alteret. Ezt nevezzük oszloptérnek.

3.19. DEFINÍCIÓ (SORTÉR, OSZLOPTÉR). *Egy mátrix oszlopvektorai által kifeszített alteret* oszloptérnek*, a sorvektorai által kifeszített alteret* sortérnek *nevezzük.*

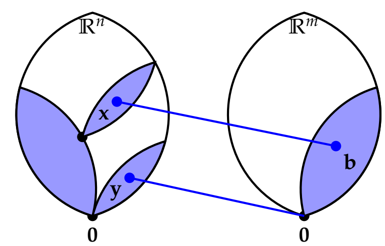


3.10. ábra: Az $\mathbf{A}$ mátrix sortere ($\mathcal{S}(\mathbf{A})$), oszloptere ($\mathcal{O}(\mathbf{A})$) és nulltere ($\mathcal{N}(\mathbf{A})$).

Az $m \times n$-es valós $\mathbf{A}$ mátrix sortere $\mathbb{R}^n$ altere, oszloptere $\mathbb{R}^m$ altere. Az $\mathbf{A}$ sorterét $\mathcal{S}(\mathbf{A})$-val, oszlopterét $\mathcal{O}(\mathbf{A})$-val jelöljük (ld. 3.11 ábra).

Az oszlopmodellt használva az is látható, hogy egy lineáris kombináció együtthatóinak meghatározása egy egyenletrendszer megoldását jelenti.

3.20. KÖVETKEZMÉNY (INHOMOGÉN EGYENLETRENDSZER MEGOLDHATÓSÁGA). *Az* $[\mathbf{A}|\mathbf{b}]$ *mátrixú egyenletrendszer pontosan akkor oldható meg, ha* $\mathbf{b}$ *előáll az* $\mathbf{A}$ *oszlopainak lineáris kombinációjaként, azaz* $\mathbf{b}$ *benne van az* $\mathbf{A}$ *oszlopterében. A lineáris kombináció együtthatói megegyeznek a megoldásvektor koordinátáival.*



3.11. ábra: A nulltér, a sortér, és az oszloptér, valamint a homogén $\mathbf{Ay} = \mathbf{0}$ és az inhomogén $\mathbf{Ax} = \mathbf{b}$ egyenletrendszer egy-egy megoldása a levéldiagramban.

3.21. PÉLDA (KIFESZÍTETT ALTÉR VEKTORAI). *Határozzuk meg, hogy a* $\mathbf{v}_1 = (1, 0, 1, 2)$*,* $\mathbf{v}_2 = (-1, 2, -2, 1)$ *és* $\mathbf{v}_3 = (1, 1, 1, 1)$ *vektorok által kifeszített altérnek eleme-e az* $\mathbf{u} = (-1, 2, -3, 6)$ *vektor! Ha igen, adjunk meg egy ezt bizonyító lineáris kombinációt is! Mutassuk meg, hogy a* $\mathbf{w} = (-1, 2, -3, 4)$ *vektor nem eleme az altérnek!*

MEGOLDÁS. Olyan x_1, x_2, x_3 valósokat keresünk, melyekkel $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{u}$ fennáll. Ez egy négy egyenletből álló egyenletrendszerrel ekvivalens, melynek bővített mátrixa a $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ és $\mathbf{u}$ vektorokból áll. Az egyenletrendszer bővített mátrixát lépcsős alakra hozva kapjuk, hogy

$$\left[\begin{array}{rrr|r} 1 & -1 & 1 & -1 \\ 0 & 2 & 1 & 2 \\ 1 & -2 & 1 & -3 \\ 2 & 1 & 1 & 6 \end{array}\right] \Rightarrow \left[\begin{array}{rrr|r} 1 & -1 & 1 & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

amiből $(x_1, x_2, x_3) = (3, 2, -2)$.

Ha $\mathbf{u}$ helyett $\mathbf{w}$-vel számolunk, olyan egyenletrendszert kapunk, melynek nincs megoldása, tehát $\mathbf{w}$ valóban nincs benne a megadott altérben. Másként fogalmazva, a $\mathbf{v}_1$, $\mathbf{v}_2$ és $\mathbf{v}_3$ vektorokból, mint oszlopvektorokból képzett mátrix oszlopterében az $\mathbf{u}$ vektor benne van, de a $\mathbf{w}$ vektor nincs benne. □

Lineáris függetlenség és összefüggőség A lineáris egyenletrendszerek megoldása és vektorok lineáris függetlenségével vagy összefüggőségével kapcsolatos kérdések szoros kapcsolatban vannak egymással.

Az előző 3.21. példa tanulsága úgy is összefoglalható, hogy egy $\mathbf{w}$ vektor pontosan akkor független az $\mathbf{A}$ mátrix oszlopvektoraitól, vagyis az $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ vektorrendszertől, ha az $[\mathbf{A}|\mathbf{w}]$ egyenletrendszer nem oldható meg.

Egy $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ vektorrendszer lineáris függetlenségének eldöntéséhez meg kell oldani az

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{0}$$

homogén lineáris egyenletrendszert. Ha van nemtriviális megoldása, akkor a vektorrendszer lineárisan összefüggő, egyébként lineárisan független. Ez igazolja az alábbi ekvivalenciákat:

3.22. következmény (Lineáris függetlenség eldöntése). *Tekintsük az* $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_k \end{bmatrix}$ *mátrixot! Az alábbi állítások ekvivalensek:*

a) az $\mathbf{a}_1$, $\mathbf{a}_2, \dots, \mathbf{a}_k$ *vektorok lineárisan függetlenek;*

b) az $\mathbf{A}$ *együtthatómátrixú homogén lineáris egyenletrendszernek a triviálison kívül nincs más megoldása;*

c) az $\mathbf{A}$ *lépcsős alakjának minden oszlopában van főelem, azaz* $\mathrm{r}(\mathbf{A}) = k$.

3.23. példa (Vektorok lineáris függetlenségének eldöntése). *Mutassuk meg, hogy a 4-dimenziós* $(1, 2, 3, 4)$, $(0, 1, 0, 1)$ *és* $(1, 1, 1, 0)$ *vektorok lineárisan függetlenek.*

Megoldás. A vektorokból képzett mátrix és lépcsős alakja

$$\begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 3 & 0 & 1 \\ 4 & 1 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix},$$

ami azt mutatja, hogy a homogén lineáris egyenletrendszernek csak egyetlen megoldása van, azaz az oszlopvektorok lineárisan függetlenek. □

Feladatok

3.1.• IGAZ – HAMIS Melyek igazak, melyek hamisak az alábbi állítások közül?

a) Ha egy 10-ismeretlenes egyenletrendszer csak 6 egyenletből áll, azaz alulhatározott, akkor végtelen sok megoldása van.

b) Ha egy 10-ismeretlenes egyenletrendszer csak 6 egyenletből áll, azaz alulhatározott, akkor lehet, hogy végtelen sok megoldása van, de az is lehet, hogy csak egy.

c) Ha egy 10-ismeretlenes egyenletrendszer 20 egyenletből áll, azaz túlhatározott, akkor biztosan nem oldható meg!

d) Ha egy 10-ismeretlenes egyenletrendszer 20 egyenletből áll, azaz túlhatározott, nem lehet végtelen sok megoldása.

3.2.• ALTEREK TULAJDONSÁGAI: IGAZ – HAMIS

a) $\mathbb{R}^n$ bármely három alterének metszete altér.

b) Ha az $\mathcal{U}$ altér altere a $\mathcal{V}$ és a $\mathcal{W}$ altérnek is, akkor altere metszetüknek is.

c) Alterek egyesítése altér.

d) Alterek összege altér.

e) Minden altérnek eleme a zérusvektor.

f) Minden altérnek van legalább egy nemzérus vektora.

3.3.• ALTEREK: IGAZ – HAMIS

a) Egy lineáris egyenletrendszer megoldásai alteret alkotnak.

b) Rögzített $\mathbf{A}$ mátrix mellett azok a $\mathbf{b}$ vektorok, melyekre az $[\mathbf{A}|\mathbf{b}]$ egyenletrendszer megoldható, alteret alkotnak.

c) Egy egyenletrendszer megoldásvektorainak különbségeként kapott vektorok halmaza alteret alkot.

d)

e)

3.4.• MEGOLDHATÓSÁG: IGAZ – HAMIS

a) Az $[\mathbf{A}|\mathbf{b}]$ mátrixú egyenletrendszer pontosan akkor oldható meg, ha $\mathbf{b}$ előáll $\mathbf{A}$ oszlopainak lineáris kombinációjaként.

b) Az $[\mathbf{A}|\mathbf{b}]$ mátrixú egyenletrendszer bármely két megoldásának különbsége megoldása a homogén $[\mathbf{A}|\mathbf{0}]$ egyenletrendszernek.

c) Az $[\mathbf{A}|\mathbf{b}]$ mátrixú egyenletrendszer bármely megoldása előáll a homogén $[\mathbf{A}|\mathbf{0}]$ mátrixú egyenletrendszer két megoldásának különbségeként.

d) Az $[\mathbf{A}|\mathbf{b}]$ mátrixú egyenletrendszer pontosan akkor oldható meg, ha $\mathrm{r}(\mathbf{A}|\mathbf{b}) \le \mathrm{r}(\mathbf{A})$.

e) Az n-ismeretlenes $[\mathbf{A}|\mathbf{b}]$ mátrixú egyenletrendszer pontosan akkor oldható meg egyértelműen, ha $\mathrm{r}(\mathbf{A}) = n$.

3.5.• INHOMOGÉN LINEÁRIS EGYENLETRENDSZER MEGOLDÁSAI Egy négyismeretlenes lineáris egyenletrendszer megoldását számítógéppel próbálom ellenőrizni, de más jön ki. A saját eredményem ez:

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + s\begin{bmatrix} 1 \\ 0 \\ -2 \\ -1 \end{bmatrix} + t\begin{bmatrix} 0 \\ 1 \\ -3 \\ -2 \end{bmatrix},$$

a számítógépé ez:

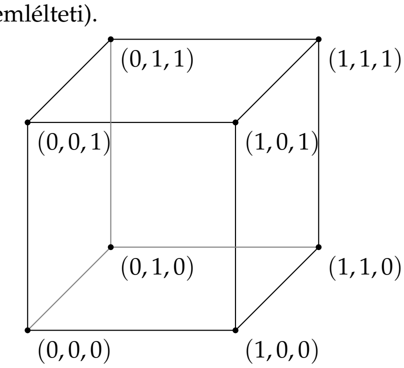
$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 2 \\ 2 \end{bmatrix} + s\begin{bmatrix} -2 \\ 1 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 3 \\ -2 \\ 0 \\ 1 \end{bmatrix}.$$

Lehet-e mindkét eredmény jó?

3.6. MEGOLDÁSOK LINEÁRIS KOMBINÁCIÓJA Adjunk új bizonyítást a 3.9. tételre a sormodellt használva.

3.7. HOMOGÉN ÉS INHOMOGÉN EGYENLETRENDSZER MEGOLDÁSAI Adjunk az oszlopmodellben megfogalmazott új bizonyítást a 3.18. tételre.

3.8. $\mathbb{F}_2^3$ ALTEREI Soroljuk fel $\mathbb{F}_2^3$ összes alterét (ehhez segítségül hívhatjuk az alábbi ábrát, mely az $\mathbb{F}_2^3$ vektortér vektorait szemlélteti).



Alterek tulajdonságai és az egyenletrendszerek

E szakaszban az alterek tulajdonságait, és az egyenletrendszerek kapcsán felmerülő alterek viszonyát vizsgáljuk. Különösen fontos az együtthatómátrixhoz tartozó négy kitüntetett altér kapcsolata.

Sor- és oszloptér Az előző feladatokban a homogén lineáris egyenletrendszer együtthatómátrixát lépcsős alakra hoztuk. Ebből az alakból azonban több minden leolvasható.

3.24. TÉTEL (ELEMI SORMŰVELETEK HATÁSA A SOR- ÉS OSZLOPVEKTOROKRA). *Elemi sorműveletek közben a sortér nem változik, az oszlopvektorok pedig olyan vektorokba transzformálódnak, melyek megőrzik az eredeti lineáris kapcsolatokat.*

BIZONYÍTÁS. Megmutatjuk, hogy egy mátrix sortere nem változik az elemi sorműveletek közben. A sorcserére ez nyilvánvaló. Legyenek $\mathbf{A}$ sorvektorai $\mathbf{v}_1$, $\mathbf{v}_2, \dots$, $\mathbf{v}_m$ és legyen $\mathbf{u}$ a sortér egy tetszőleges vektora, ami azt jelenti, hogy vannak olyan c_1, $c_2, \dots$, c_m skalárok, hogy

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_m\mathbf{v}_m.$$

Ha egy sort (mondjuk az elsőt) beszorozzuk egy $d \neq 0$ skalárral, akkor $\mathbf{u}$ az új mátrix sorterében is benne van, melyet a $d\mathbf{v}_1$, $\mathbf{v}_2, \dots$, $\mathbf{v}_m$ vektorok feszítenek ki, hisz

$$\mathbf{u} = \frac{c_1}{d}(d\mathbf{v}_1) + c_2\mathbf{v}_2 + \dots + c_m\mathbf{v}_m.$$

Ez azt jelenti, hogy a sortér nem csökken, azaz minden vektor, ami eddig benne volt a sortérben, benne lesz az új mátrix sorterében is. A hozzáadás műveleténél ugyanezt tapasztaljuk: az általánosság megszorítása nélkül feltehetjük, hogy az első sor d-szeresét adjuk a második sorhoz. Ekkor az új sorteret kifeszítő vektorok: $\mathbf{v}_1$, $\mathbf{v}_2 + d\mathbf{v}_1, \dots$, $\mathbf{v}_m$. Az $\mathbf{u}$ vektor e térben is benne van, ugyanis egy egyszerű átalakítás után

$$\mathbf{u} = (c_1 - c_2 d)\mathbf{v}_1 + c_2(\mathbf{v}_2 + d\mathbf{v}_1) + \dots + c_m\mathbf{v}_m.$$

Mivel minden sorművelet inverze is egy sorművelet, az inverz sorműveletben sem csökken a sortér, ami csak úgy lehet, ha a sortér az elemi sorműveletek során változatlan marad.

Megmutatjuk, hogy az elemi sorműveletek közben az oszlopvektorok közt nem változnak a lineáris kapcsolatok. Az egyszerűség kedvéért tegyük fel, hogy például $c\mathbf{v}_1 + d\mathbf{v}_2 = \mathbf{0}$. Mivel a skalárral való szorzás és a vektorösszeadás is koordinátánként végezhető, könnyen látható, hogy sorcsere, egy koordináta nemnulla skalárral való szorzása, vagy az i-edik koordináta konstansszorosának a j-edikhez adása után is fönn fog állni a két új vektor közt a fenti összefüggés. Tetszőleges számú vektor lineáris kombinációjára az állítás hasonlóan adódik. □

3.25. KÖVETKEZMÉNY (MÁTRIX LÉPCSŐS ALAKJÁNAK VEKTORAI). *Legyen* **B** *az* **A** *mátrix egy lépcsős alakja. Ekkor*

1. **A** *és* **B** *sortere megegyezik,*
2. *az* **A** *oszlopvektorai között lévő lineáris kapcsolatok azonosak a* **B** *ugyanolyan sorszámú oszlopai közti lineáris kapcsolatokkal,*
3. **B** *nemzérus sorvektorai lineárisan függetlenek,*
4. *a főelemek oszlopvektorai* **A**-*ban és* **B**-*ben is lineárisan függetlenek.*

BIZONYÍTÁS. Az első két állítás közvetlen következménye az előző tételnek.

A harmadik állítás bizonyításához megmutatjuk, hogy egy lépcsős alak egy nemzérus sorvektora nem fejezhető ki a többi sorvektor lineáris kombinációjaként. Tekintsük a lépcsős alak k-adik sorvektorát. Főeleme legyen a j-edik oszlopban. E főelem nem állítható elő a k-nál nagyobb indexű sorok lineáris kombinációjával, mert azokban a j-edik koordináta 0. A k-nál kisebb indexű sorvektorok pedig nem szerepelhetnek a lineáris kombinációban, mivel a legkisebb indexű vektor főelemét a többi vektor nem eliminálhatja, pedig a k-adik sorban azon a helyen 0 áll.

Annak bizonyítása, hogy a főelemek oszlopai **B**-ben lineárisan függetlenek, ugyanúgy megy, mint a sorvektorok esetén. Innen pedig az előző tétellel adódik, hogy az ilyen indexű oszlopok **A**-ban is lineárisan függetlenek. □

Fontos megjegyezni, hogy míg a lépcsős alak sortere megegyezik az eredeti mátrix sorterével, addig az oszloptér az elemi sorműveletek alatt megváltozik, tehát a mátrix és lépcsős alakjának oszloptere különbözik!

Bázis Az elemi sorműveleteket alkalmazva, egy mátrix sorterében és oszlopterében is találtunk olyan lineárisan független vektorokat, melyek kifeszítik az adott teret. Azt már az 1.9. tételben megmutattuk, hogy a háromdimenziós tér tetszőleges három lineárisan független vektorának lineáris kombinációjaként a tér minden vektora előáll. Más szavakkal ez azt jelenti, hogy a tér három lineárisan független vektora kifeszíti a teret. Az ilyen vektorhármasokat, melyeket egy koordinátarendszer alapvektorainak vettünk, bázisnak neveztük. Ezek vezetnek a következő definícióhoz.

3.26. DEFINÍCIÓ (BÁZIS). *Az* $\mathbb{R}^n$ *tér egy* alterének bázisán *vektorok olyan halmazát értjük, mely*

1. *lineárisan független vektorokból áll és*

2. *kifeszíti az alteret.*
Az $\mathbf{e}_1 = (1, 0, \ldots, 0)$, $\mathbf{e}_2 = (0, 1, \ldots, 0)$,..., $\mathbf{e}_n = (0, 0, \ldots, 0, 1)$ *vektorokból álló halmazt* $\mathbb{R}^n$ standard bázisának *nevezzük.*

► Világos, hogy a zérustérnek nincs bázisa, hisz abban nincs egyetlen lineárisan független vektor sem.
► A standard bázis $\mathbb{R}^n$ egy n-elemű bázisa.
► Hamarosan meg fogjuk mutatni, hogy $\mathbb{R}^n$ minden bázisa n-elemű, és hogy bármely alterének bázisa legfeljebb n-elemű.

3.27. PÉLDA (ALTÉR BÁZISÁNAK MEGHATÁROZÁSA). *Határozzuk meg az* $(1, 1, 0, -2)$, $(2, 3, 3, -2)$, $(1, 2, 3, 0)$ *és* $(1, 3, 6, 2)$ *vektorok által kifeszített altér egy bázisát!*

MEGOLDÁS. *Első megoldás:* A megadott vektorokból, mint sorvektorokból képzett mátrix valamely sorlépcsős alakjának nemnulla sorai az altér egy bázisát adják:

$$\begin{bmatrix} 1 & 1 & 0 & -2 \\ 2 & 3 & 3 & -2 \\ 1 & 2 & 3 & 0 \\ 1 & 3 & 6 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 0 & -2 \\ 0 & 1 & 3 & 2 \\ 0 & 1 & 3 & 2 \\ 0 & 2 & 6 & 4 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 0 & -2 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

A bázis vektorai $(1, 1, 0, -2)$, $(0, 1, 3, 2)$.

Második megoldás: Ha a bázist az adott vektorokból akarjuk kiválasztani, akkor képezzünk egy mátrixot e vektorokból, mint oszlopvektorokból. Lépcsős alakjában a főelemek oszlopai lineárisan független vektorok. A nekik megfelelő oszlopvektorok az eredeti mátrixban az oszloptér bázisát alkotják (ld. a 3.24. tételt és a 3.25. következmény 4. pontjának állítását).

$$\begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 3 & 2 & 3 \\ 0 & 3 & 3 & 6 \\ -2 & -2 & 0 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 3 & 3 & 6 \\ 0 & 2 & 2 & 4 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Tehát az adott négy vektor közül az első kettő, azaz az $(1, 1, 0, -2)$ és $(2, 3, 3, -2)$ vektorok bázist alkotnak.

Ha a megadott vektorokat más sorrendben írjuk a mátrixba, másik bázist kaphatunk. □

3.28. PÉLDA (VEKTOR FELÍRÁSA A BÁZISVEKTOROK LINEÁRIS KOMBINÁCIÓJAKÉNT). *Az előző feladatban megadott négy vektor mindegyikét fejezzük ki az általuk kifeszített altér bázisvektorainak lineáris kombinációjaként!*

MEGOLDÁS. Az előző feladat második megoldásában találtunk egy bázist a megadott vektorok közül. Mivel az oszlopvektorokkal dolgoztunk, a vektorok közti lineáris kapcsolat leolvasható bármelyik lépcsős alakból: legkényelmesebben a *redukált* lépcsős alakból. Folytatjuk tehát az előző példabeli eliminációs lépéseket:

$$\begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 3 & 2 & 3 \\ 0 & 3 & 3 & 6 \\ -2 & -2 & 0 & 2 \end{bmatrix} \Rightarrow \cdots \Rightarrow \begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & -1 & -3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{3.1}$$

A redukált lépcsős alakból látjuk, hogy például a harmadik oszlop a második és az első különbsége. Ezek alapján az eredeti vektoroknak a bázisvektorok lineáris kombinációiként való felírása:

$$\begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix} = -\begin{bmatrix} 1 \\ 1 \\ 0 \\ -2 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \\ 3 \\ -2 \end{bmatrix}, \qquad \begin{bmatrix} 1 \\ 3 \\ 6 \\ 2 \end{bmatrix} = -3\begin{bmatrix} 1 \\ 1 \\ 0 \\ -2 \end{bmatrix} + 2\begin{bmatrix} 2 \\ 3 \\ 3 \\ -2 \end{bmatrix}.$$

Ha a megadott vektorokat más sorrendben írjuk a mátrixba, más bázisra juthatunk. Például

$$\begin{bmatrix} 1 & 1 & 2 & 1 \\ 3 & 1 & 3 & 2 \\ 6 & 0 & 3 & 3 \\ 2 & -2 & -2 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & -2 & -3 & -1 \\ 0 & -6 & -9 & -3 \\ 0 & -4 & -6 & -2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & \frac{3}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{3.2}$$

Eszerint bázisvektoroknak választhatjuk az $(1, 3, 6, 2)$ és az $(1, 1, 0, -2)$ vektorokat is, a többi vektor pedig kifejezhető ezek lineáris kombinációjaként:

$$\begin{bmatrix} 2 \\ 3 \\ 3 \\ -2 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 \\ 3 \\ 6 \\ 2 \end{bmatrix} + \frac{3}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \\ -2 \end{bmatrix}, \qquad \begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 \\ 3 \\ 6 \\ 2 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \\ -2 \end{bmatrix}.$$

□

3.29. ÁLLÍTÁS (BÁZIS EKVIVALENS DEFINÍCIÓI). *Legyen $\mathcal{U}$ az $\mathbb{R}^n$ egy tetszőleges altere, és legyen $V = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\} \subseteq \mathcal{U}$ vektorok véges halmaza. A következő állítások ekvivalensek:*

1. *V lineárisan független vektorokból áll és kifeszíti az $\mathcal{U}$ alteret, azaz V az $\mathcal{U}$ altér bázisa;*
2. *V minimális méretű halmaz, mely kifeszíti $\mathcal{U}$-t;*
3. *V maximális méretű, független vektorokból álló halmaza $\mathcal{U}$-nak.*

Vektor egy bázisra vonatkozó koordinátás alakja A koordinátarendszer bevezetésénél ugyanazt tettük, mint itt az előző példában: minden vektor előállítható egy bázis elemeinek lineáris kombinációjaként, és e vektor koordinátás alakja erre a bázisra vonatkozóan a lineáris kombináció konstansaiból áll.

Egy altérben több bázist is vizsgálhatunk, és a vektorok koordinátás alakjai különbözhetnek a különböző bázisokban. Félreértések elkerülésére a bázis jelét a koordinátás alak indexében jelöljük. Például ha egy $\mathbf{v}$ vektor standard bázisbeli és $\mathcal{B}$ bázisbeli koordinátás alakjai $(4,3)$, illetve $(0,5)$, akkor azt írjuk, hogy

$$\mathbf{v} = (4,3) = (0,5)_{\mathcal{B}}, \text{ vagy mátrixjelöléssel } \mathbf{v} = \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}_{\mathcal{B}}.$$

Ha általában akarunk utalni – a konkrét koordináták nélkül – egy $\mathbf{v}$ vektor $\mathcal{B}$ bázisbeli koordinátás alakjára, akkor a $[\mathbf{v}]_{\mathcal{B}}$ vagy a $(\mathbf{v})_{\mathcal{B}}$ alakot használjuk. Így írhatjuk azt is, hogy

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}_{\mathcal{B}}, \text{ vagy egyszerűbben, hogy } [\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}.$$

3.30. PÉLDA (VEKTOR KOORDINÁTÁS ALAKJA A $\mathcal{B}$ BÁZISBAN). *Tekintsük a 3.27. és a 3.28. példákban is szereplő* $\mathbf{v}_1 = (1,1,0,-2)$, $\mathbf{v}_2 = (2,3,3,-2)$, $\mathbf{v}_3 = (1,2,3,0)$ *és* $\mathbf{v}_4 = (1,3,6,2)$ *vektorok által kifeszített alteret. Ennek* $\mathcal{A} = \{\, \mathbf{v}_1, \mathbf{v}_2 \,\}$ *és* $\mathcal{B} = \{\, \mathbf{v}_4, \mathbf{v}_1 \,\}$ *is bázisa. Írjuk fel a négy vektornak a két bázisra vonatkozó koordinátás alakjait!*

MEGOLDÁS. Az előző példában a (3.1) képletbeli redukált lépcsős alak nemzérus soraiból álló

$$\begin{bmatrix} 1 & 0 & -1 & -3 \\ 0 & 1 & 1 & 2 \end{bmatrix}$$

mátrix azt mutatja, hogy az $\mathcal{A}$ bázisban e négy vektor koordinátái rendre

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}_{\mathcal{A}}, \ \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}_{\mathcal{A}}, \ \mathbf{v}_3 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}_{\mathcal{A}}, \ \mathbf{v}_4 = \begin{bmatrix} -3 \\ 2 \end{bmatrix}_{\mathcal{A}}.$$

Ez a 3.25. állítás 2. pontjából következik, mely szerint a redukált lépcsős alak oszlopai közti lineáris kapcsolatok megegyeznek az eredeti mátrix oszlopai közti lineáris kapcsolatokkal. Hasonlóképp a (3.2) képletbeli redukált lépcsős alak nemzérus soraiból álló

$$\begin{bmatrix} 1 & 0 & 1/2 & 1/2 \\ 0 & 1 & 3/2 & 1/2 \end{bmatrix}$$

mátrixból kiolvasható, hogy a fenti altérnek $\mathcal{B} = \{\, \mathbf{v}_4, \mathbf{v}_1 \,\}$ is bázisa, és ebben a bázisban a $\mathbf{v}_4$, $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ koordinátás alakja rendre

$$\mathbf{v}_4 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}_{\mathcal{B}}, \ \mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}_{\mathcal{B}}, \ \mathbf{v}_2 = \begin{bmatrix} 1/2 \\ 3/2 \end{bmatrix}_{\mathcal{B}}, \ \mathbf{v}_3 = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}_{\mathcal{B}}.$$

Az eredmény úgy írható, hogy $(\mathbf{v}_1)_{\mathcal{A}} = (1,0)$, $(\mathbf{v}_1)_{\mathcal{B}} = (0,1)$. □

Dimenzió és rang Az előzőekben bázist kerestünk egy altérhez. Azt tapasztaltuk, hogy a bázis mindig ugyanannyi vektorból állt. Ez nem véletlen. Egy különösen fontos tétel következik.

3.31. TÉTEL (BÁZIS-TÉTEL). *Tekintsük az $\mathbb{R}^n$ vektortér egy tetszőleges alterét. Ennek bármely két bázisa azonos számú vektorból áll.*

BIZONYÍTÁS. Tegyük fel, hogy $\mathbb{R}^n$-nek van olyan $\mathcal{A}$ altere, és annak két olyan bázisa,

$$V = \{\,\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\,\}, \text{ és } W = \{\,\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_r\,\},$$

melyek nem ugyanannyi vektorból állnak, azaz például $k < r$. Mivel V bázis $\mathcal{A}$-ban, ezért a W bázis vektorai is kifejezhetők lineáris kombinációikként, azaz léteznek olyan a_{ij} skalárok, hogy

$$\mathbf{w}_i = a_{i1}\mathbf{v}_1 + a_{i2}\mathbf{v}_2 + \ldots + a_{ik}\mathbf{v}_k, \quad (i = 1, \dots, r). \tag{3.3}$$

Mivel a W bázis vektorai lineárisan függetlenek, ezért a

$$c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \ldots + c_r\mathbf{w}_r = \mathbf{0} \tag{3.4}$$

egyenlőség csak a $c_1 = c_2 = \ldots = c_k = 0$ konstansokra áll fenn. A (3.3) egyenlőségeit a (3.4) egyenletbe helyettesítve

$$\begin{aligned} c_1(a_{11}\mathbf{v}_1 + a_{12}\mathbf{v}_2 + \ldots + a_{1k}\mathbf{v}_k) + c_2(a_{21}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + \ldots + a_{2k}\mathbf{v}_k) + \\ \ldots + c_r(a_{r1}\mathbf{v}_1 + a_{r2}\mathbf{v}_2 + \ldots + a_{rk}\mathbf{v}_k) = \mathbf{0}, \end{aligned}$$

aminek V vektorai szerinti rendezése után kapjuk, hogy

$$\begin{aligned} (a_{11}c_1 + a_{21}c_2 + \ldots + a_{r1}c_r)\mathbf{v}_1 + (a_{12}c_1 + a_{22}c_2 + \ldots + a_{r2}c_r)\mathbf{v}_2 + \\ \ldots + (a_{1k}c_1 + a_{2k}c_2 + \ldots + a_{rk}c_r)\mathbf{v}_k = \mathbf{0}. \end{aligned}$$

Ez azt jelenti, hogy a homogén lineáris

$$\begin{aligned} a_{11}c_1 + a_{21}c_2 + \ldots + a_{r1}c_r &= 0 \\ a_{12}c_1 + a_{22}c_2 + \ldots + a_{r2}c_r &= 0 \\ \vdots \qquad \vdots \qquad\qquad \vdots \quad & \\ a_{1k}c_1 + a_{2k}c_2 + \ldots + a_{rk}c_r &= 0 \end{aligned}$$

egyenletrendszernek a $c_1 = c_2 = \ldots = c_k = 0$ az egyetlen megoldása. Ez viszont a 3.7. tétel szerint nem teljesülhet, mivel a fenti homogén egyenletrendszer egyenleteinek száma kisebb ismeretlenjei számánál ($k < r$). Ez az ellentmondás bizonyítja, hogy indirekt feltevésünk helytelen volt, tehát valóban, egy altér bármely két bázisa azonos számú vektorból áll. □

Ha a háromdimenziós térben tekintünk egy origón átmenő síkot, látjuk, hogy bármely két független vektora kifeszíti. Azaz minden bázisa kételemű. Ha egy origón átmenő egyenest tekintünk, azt minden nemnulla vektora, mint egyelemű bázisa, kifeszíti. A hétköznapi életben is használt fogalom, a dimenzió, megegyezik a bázis elemszámával. Miután egy altér minden bázisa azonos elemszámú, értelmes a következő definíció:

3.32. definíció (Dimenzió). *Az* $\mathbb{R}^n$ *tér egy* $\mathcal{A}$ *alterének* dimenzióján *egy bázisának elemszámát értjük, és e számot* dim $\mathcal{A}$*-val jelöljük.*

Az $\mathbb{R}^n$ altere saját magának, és standard bázisa épp n vektorból áll, így $\dim \mathbb{R}^n = n$. A nullvektorból álló altérben nincsenek független vektorok, így dimenziója 0.

Mielőtt a következő állítást megfogalmaznánk, definiálunk egy egyszerű fogalmat. Mivel a mátrixalkalmazásokban nincs olyan természetes elv, amely alapján el lehetne dönteni, hogy mely mennyiségeket rendeljünk a sorokhoz, melyeket az oszlopokhoz, ugyanakkor a számítások közben ez korántsem mindegy, ezért szükség lehet a következő egyszerű műveletre: az $m \times n$-es $\mathbf{A}$ mátrix *transzponáltján* azt az $\mathbf{A}^T$-vel jelölt $n \times m$-es mátrixot értjük, amelyet az $\mathbf{A}$ sorainak és oszlopainak felcserélésével kapunk. Azaz

$$\mathbf{A}^T = [a_{ij}]^T := [a_{ji}].$$

3.33. példa (Mátrix transzponáltja). *Az*

$$\mathbf{A} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 5 & 6 \end{bmatrix}$$

mátrixok transzponáltja

$$\mathbf{A}^T = \begin{bmatrix} 1 & 2 \end{bmatrix}, \quad \mathbf{B}^T = \begin{bmatrix} 0 & 2 \\ 1 & 3 \end{bmatrix}, \quad \mathbf{C}^T = \begin{bmatrix} 0 & 3 \\ 1 & 4 \\ 2 & 5 \end{bmatrix}, \quad \mathbf{D}^T = \begin{bmatrix} 5 \\ 6 \end{bmatrix}.$$

Adott véges sok $\mathbb{R}^n$-beli vektor által kifeszített altér dimenzióját úgy határozhatjuk meg, hogy meghatározzuk a vektorokból képzett mátrix rangját. Igaz ugyanis a következő állítás:

3.34. állítás (Dimenzió = rang). *Egy mátrix rangja, sorterének dimenziója és oszlopterének dimenziója megegyezik. Ebből következőleg* $\mathrm{r}(\mathbf{A}) = \mathrm{r}(\mathbf{A}^T)$.

Bizonyítás. A mátrix rangja megegyezik a lépcsős alakjában lévő nemzérus sorainak számával. A 3.25. tétel szerint viszont e sorok lineárisan függetlenek és kifeszítik a sorteret, tehát bázist alkotnak, így számuk a sortér dimenzióját adja. Az oszloptérről láttuk, hogy a főelemeknek megfelelő oszlopok az eredeti mátrixban lineárisan függetlenek és kifeszítik az oszlopteret, tehát e tér dimenziója is a mátrix rangjával egyezik meg. Az utolsó állítás abból következik, hogy $\mathbf{A}$ sortere megegyezik $\mathbf{A}^T$ oszlopterével. □

3.35. DEFINÍCIÓ (VEKTORRENDSZER RANGJA). *Egy $\mathbb{R}^n$-beli vektorokból álló vektorrendszer* rangján *a vektorokból képzett mátrix rangját, vagy ami ezzel egyenlő, az általuk kifeszített altér dimenzióját értjük.*

3.36. PÉLDA (DIMENZIÓ KISZÁMÍTÁSA). *Határozzuk meg az $\mathbf{A}$ mátrix sorterének és nullterének dimenzióját!*

$$\mathbf{A} = \begin{bmatrix} 3 & 3 & 3 & 3 & 3 \\ 3 & 4 & 5 & 4 & 3 \\ 3 & 2 & 1 & 2 & 3 \\ 3 & 3 & 3 & 3 & 3 \end{bmatrix}$$

MEGOLDÁS. Az $\mathbf{A}$ redukált lépcsős alakja

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & -1 & 0 & 1 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Innen leolvasható, hogy a mátrix rangja 2, így sorterének dimenziója is 2. A nulltér dimenziója megegyezik az egyenletrendszer megoldásterének dimenziójával, ami megegyezik a szabad változók számával, esetünkben ez 3. Vegyük észre, hogy a sortér és a nulltér dimenziójának összege megegyezik a változók számával, azaz a mátrix oszlopainak számával, jelen példában 5-tel. □

3.37. TÉTEL (DIMENZIÓTÉTEL). *Bármely $\mathbf{A} \in \mathbb{R}^{m \times n}$ mátrix esetén a sortér dimenziójának és a nulltér dimenziójának összege n. Képlettel:*

$$\dim(\mathcal{S}(\mathbf{A})) + \dim(\mathcal{N}(\mathbf{A})) = n.$$

BIZONYÍTÁS. A mátrix sorterének dimenziója megegyezik a mátrix rangjával, azaz az $[\mathbf{A}|\mathbf{0}]$ mátrixú egyenletrendszerben a kötött változók számával. Megmutatjuk, hogy a nulltér dimenziója megegyezik a szabad változók számával, így e két szám összege valóban n, ami bizonyítja az állítást (ld. még a 4.37. tételben).

Elég tehát megmutatnunk, hogy egy homogén lineáris egyenletrendszer redukált lépcsős alakkal előállított megoldásában a szabad változók száma megegyezik a nulltérből kiválasztható bázis elemszámával. Először lássunk egy ilyen megoldást konkrétan. Például a 2.43. példabeli homogén lineáris egyenletrendszer megoldása

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2s - \frac{3}{2}t - u \\ s \\ -\frac{1}{2}t \\ t \\ u \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -\frac{3}{2} \\ 0 \\ -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + u \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

ahol $x_2 = s$, $x_4 = t$ és $x_5 = u$ a három szabad változó. A nullteret kifeszítő három vektor közül az elsőben $x_2 = 1$, de az összes többiben $x_2 = 0$, így az első vektor független a többitől. Hasonlóképp általában is igaz, hogy a redukált lépcsős alakból való származtatás következtében a nullteret kifeszítő minden megoldásvektorban az összes szabad változóhoz tartozó koordináta 0, azt az egy koordinátát kivéve, amelyikhez a vektor tartozik. Így viszont mindegyik vektor független a többitől, vagyis e vektorok függetlenek, és mivel kifeszítik a nullteret, számuk megadja a nulltér dimenzióját. □

*Elemi bázistranszformáció** Az előző paragrafusokban azt láttuk, hogy az elemi sorműveletek eredményeként az eredeti mátrix oszlopainak egy másik bázisban felírt koordinátás alakját kapjuk meg. Ez adja az ötletet ahhoz, hogy az altérből választott bázisok segítségével szemléltessük, és értsük meg mi történik akkor, amikor a mátrix egy oszlopában főelemet (pivotelemet) választunk, és oszlopának többi elemét elimináljuk. Így egy másik megközelítéshez jutunk, melyet a továbbiakban nem használunk, ezért e paragrafus átugorható.

A folyamat lényege egy kétoszlopos mátrixon is jól szemléltethető. Az egyszerűség kedvéért a két oszlop legyen az $\mathbf{a}$ és a $\mathbf{b}$ vektor, a bázis, melyben e vektorok meg vannak adva, legyen a standard bázis. Tegyük fel, hogy $a_i \neq 0$. Ekkor az a_i pozícióját választva, a kiküszöbölés eredményeként a következőket kapjuk.

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_i & b_i \\ \vdots & \vdots \\ a_m & b_m \end{bmatrix} \Longrightarrow \begin{bmatrix} 0 & b_1 - \frac{b_i}{a_i}a_1 \\ 0 & b_2 - \frac{b_i}{a_i}a_2 \\ \vdots & \vdots \\ 1 & \frac{b_i}{a_i} \\ \vdots & \vdots \\ 0 & b_m - \frac{b_i}{a_i}a_m \end{bmatrix}$$

Tudjuk, az oszlopok az elemi sorműveletek után az eredeti vektorokat adják egy másik bázisban. Az $\mathbf{a}$ vektor nyilvánvalóan egy olyan bázisban lett felírva, amelyben szerepel az $\mathbf{a}$ vektor is, mégpedig ez az i-edik bázisvektor. Megmutatjuk, hogy mindkét oszlop az

$$\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{i-1}, \mathbf{a}, \mathbf{e}_{i+1}, \ldots, \mathbf{e}_m$$

bázisban lett felírva. Az $\mathbf{a}$ vektorra ez nyilván igaz. Nézzük a $\mathbf{b}$ vektort! Fejezzük ki az $\mathbf{e}_i$ vektort az $\mathbf{a} = a_1\mathbf{e}_1 + \ldots + a_i\mathbf{e}_i + \ldots + a_m\mathbf{e}_m$ felírásból:

$$\mathbf{e}_i = -\frac{1}{a_i}a_1\mathbf{e}_1 - \frac{1}{a_i}a_2\mathbf{e}_2 - \ldots + \frac{1}{a_i}\mathbf{a} - \ldots - \frac{1}{a_i}a_m\mathbf{e}_m.$$

Ezt behelyettesítjük a $\mathbf{b} = b_1\mathbf{e}_1 + \ldots + b_i\mathbf{e}_i + \ldots + b_m\mathbf{e}_m$ kifejezésbe:

$$\mathbf{b} = (b_1 - \frac{b_i}{a_i}a_1)\mathbf{e}_1 + (b_2 - \frac{b_i}{a_i}a_2)\mathbf{e}_2 + \ldots + \frac{b_i}{a_i}\mathbf{a} + \ldots + (b_m - \frac{b_i}{a_i}a_m)\mathbf{e}_m.$$

Tehát valóban, a $\mathbf{b}$ koordinátás alakja e módosított bázisban épp az, amit az eredeti mátrix eliminálása után kaptunk a második oszlopban. Az imént tárgyalt lépést elemi bázistranszformációnak nevezzük, mert egy másik bázisra való áttérés egy elemi lépésének tekintjük, amikor egyetlen bázisvektort cserélünk ki. A történtek szemléltetésére a mátrixot fejléccel együtt egy táblázatba írjuk, a sorok elé az $\mathbf{e}_1, \ldots, \mathbf{e}_m$ bázisvektorok, az oszlopok fölé az oszlopvektorok neve kerül.

$$\begin{array}{c|cc}
 & \mathbf{a} & \mathbf{b} \\
\hline
\mathbf{e}_1 & a_1 & b_1 \\
\mathbf{e}_2 & a_2 & b_2 \\
\vdots & \vdots & \vdots \\
\mathbf{e}_i & a_i & b_i \\
\vdots & \vdots & \vdots \\
\mathbf{e}_m & a_m & b_m
\end{array}
\quad\Longrightarrow\quad
\begin{array}{c|cc}
 & \mathbf{a} & \mathbf{b} \\
\hline
\mathbf{e}_1 & 0 & b_1 - \frac{b_i}{a_i}a_1 \\
\mathbf{e}_2 & 0 & b_2 - \frac{b_i}{a_i}a_2 \\
\vdots & \vdots & \vdots \\
\mathbf{a} & 1 & \frac{b_i}{a_i} \\
\vdots & \vdots & \vdots \\
\mathbf{e}_m & 0 & b_m - \frac{b_i}{a_i}a_m
\end{array}$$

Összefoglalva és egyúttal általánosabban megfogalmazva a fentieket:

3.38. TÉTEL (ELEMI BÁZISTRANSZFORMÁCIÓ). *Tegyük fel, hogy az* $\mathbf{a}$ *vektor* $E = \{\, \mathbf{e}_1, \ldots, \mathbf{e}_m \,\}$ *bázisra vonatkozó i-edik koordinátája* $a_i \neq 0$. *Ekkor az E által generált* $\mathcal{E}$ *altérnek az*

$$\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{i-1}, \mathbf{a}, \mathbf{e}_{i+1}, \ldots, \mathbf{e}_m$$

vektorok is bázisát alkotják. Az $\mathcal{E}$ *egy tetszőleges b vektorának koordinátás alakja megkapható e bázisban elemi sorműveletekkel, ha* a_i*-t választjuk főelemnek.*

Az elemi bázistranszformáció alkalmas arra, hogy a bázisok változásán keresztül egy más nézőpontból világítsa meg a redukált lépcsős alakra hozással megoldható feladatokat. Példaként vizsgáljuk meg, mi történik egy egyenletrendszer megoldásakor. Megjegyezzük, hogy itt nincs szükség sorcserére, mert egy oszlopból szabadon választhatunk olyan sort, amelynek fejlécében még az eredeti bázisvektor szerepel.

3.39. PÉLDA (EGYENLETRENDSZER MEGOLDÁSA ELEMI BÁZISTRANSZFORMÁCIÓVAL). *Oldjuk meg a 2.40. és a 2.48. példában megoldott egyenletrendszert elemi bázistranszformációval.*

MEGOLDÁS. A táblázatokat egybefűzzük, a sorok fejlécein mindig jelezzük az aktuális bázist, az oszlopok fejléceit a jobb érthetőség végett mindig kiírjuk, a kiválasztott főelemeket külön jelöljük:

	$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$	$\mathbf{b}$		$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$	$\mathbf{b}$		$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$	$\mathbf{b}$		$\mathbf{a}_1$	$\mathbf{a}_2$	$\mathbf{a}_3$	$\mathbf{b}$
$\mathbf{e}_1$	1	1	2	0	$\mathbf{a}_1$	1	1	2	0	$\mathbf{a}_1$	1	0	3	-5	$\mathbf{a}_1$	1	0	0	1
$\mathbf{e}_2$	2	2	3	2	$\mathbf{e}_2$	0	0	-1	2	$\mathbf{e}_2$	0	0	-1	2	$\mathbf{e}_2$	0	0	0	0
$\mathbf{e}_3$	1	3	3	4	$\mathbf{e}_3$	0	2	1	4	$\mathbf{e}_3$	0	0	3	-6	$\mathbf{a}_3$	0	0	1	-2
$\mathbf{e}_4$	1	2	1	5	$\mathbf{e}_4$	0	1	-1	5	$\mathbf{a}_2$	0	1	-1	5	$\mathbf{a}_2$	0	1	0	3

A táblázaton kicsit lehet egyszerűsíteni, azt az oszlopot, amelyben már csak egy standard egységvektor van, felesleges kiírni, az oszlopok és a sorok fejléceibe pedig elég csak azt a változót írni, amelyik a bázisba vett oszlopvektorhoz tartozik. Így a következőt kapjuk:

x	y	z	$\mathbf{b}$		y	z	$\mathbf{b}$		z	$\mathbf{b}$		$\mathbf{b}$
1	1	2	0	x	1	2	0	x	3	-5	x	1
2	2	3	2		0	-1	2		-1	2		0
1	3	3	4		2	1	4		3	-6	z	-2
1	2	1	5		1	-1	5	y	-1	5	y	3

Az egyenletrendszer megoldása tehát $x = 1$, $y = 3$, $z = -2$. □

Feladatok

3.9.• Bázis: igaz – hamis

a) A $\mathcal{V}$ altérben a $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ vektorrendszer bázis, ha tetszőleges $\mathbf{v} \in \mathcal{V}$ vektor egyértelműen felírható e vektorok lineáris kombinációjaként.

b) Van olyan altér, melynek bármely nemnulla vektora bázis.

c) Van olyan altér, melynek van kételemű bázisa, és van ettől különböző három lineárisan független vektora.

3.10. Bázisvektorok skaláris szorzatai Igazoljuk, hogy a

$$\begin{bmatrix} \mathbf{v}_1 \cdot \mathbf{v}_1 & \mathbf{v}_1 \cdot \mathbf{v}_2 & \ldots & \mathbf{v}_1 \cdot \mathbf{v}_k \\ \mathbf{v}_2 \cdot \mathbf{v}_1 & \mathbf{v}_2 \cdot \mathbf{v}_2 & \ldots & \mathbf{v}_2 \cdot \mathbf{v}_k \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{v}_k \cdot \mathbf{v}_1 & \mathbf{v}_k \cdot \mathbf{v}_2 & \ldots & \mathbf{v}_k \cdot \mathbf{v}_k \end{bmatrix}$$

mátrix rangja k, ha $\mathbf{v}_1$, $\mathbf{v}_2$,…, $\mathbf{v}_k$ vektorok lineárisan független $\mathbb{R}^n$-beli vektorok ($k \leq n$).

A lineáris algebra alaptétele

A lineáris algebra alaptétele szerint bármely $m \times n$-es valós $\mathbf{A}$ mátrix esetén $\mathbb{R}^n$ minden vektora egyértelműen felbomlik egy $\mathbf{A}$ sorterébe és egy nullterébe eső vektor összegére, melyek merőlegesek egymásra. Más szavakkal ezt úgy fogjuk kifejezni, hogy a sortér és a nulltér merőleges kiegészítő alterei egymásnak.

A sortér és a nulltér merőlegessége Az $\mathbf{A}$ mátrix sorvektorai érdekes tulajdonsággal rendelkeznek. Egy homogén lineáris egyenletrendszer minden egyenletének bal oldala az együtthatómátrix egy sorvektorának és a megoldásvektornak a skaláris szorzata. Mivel e szorzat értéke 0, ezért e két vektor merőleges egymásra.

3.40. PÉLDA (VEKTOROKRA MERŐLEGES ALTÉR). *Határozzuk meg az összes olyan vektort $\mathbb{R}^4$-ben, mely merőleges a $\mathbf{v}_1 = (1, 0, 1, 2)$ és $\mathbf{v}_2 = (-1, 2, -2, 1)$ vektorok mindegyikére!*

MEGOLDÁS. Olyan $\mathbf{x}$ vektort keresünk, melyre $\mathbf{v}_1 \cdot \mathbf{x} = 0$ és $\mathbf{v}_2 \cdot \mathbf{x} = 0$. Ezt koordinátákkal felírva két egyenletet kapunk, melynek együtthatómátrixa és annak egy lépcsős alakja a következő:

$$\begin{bmatrix} 1 & 0 & 1 & 2 \\ -1 & 2 & -2 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 2 & -1 & 3 \end{bmatrix},$$

amiből $\mathbf{x} = (-s - 2t, (s - 3t)/2, s, t)$, azaz

$$\mathbf{x} = s \begin{bmatrix} -1 \\ 1/2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ -3/2 \\ 0 \\ 1 \end{bmatrix}.$$

A megoldás tehát két vektor által kifeszített altér összes vektora.

Másként fogalmazva, e feladatban meghatároztuk az összes olyan vektort, mely egy mátrix sorvektorainak mindegyikére merőleges. □

Az $m \times n$-es $\mathbf{A}$ együtthatómátrixú homogén lineáris egyenletrendszer i-edik egyenletének alakja

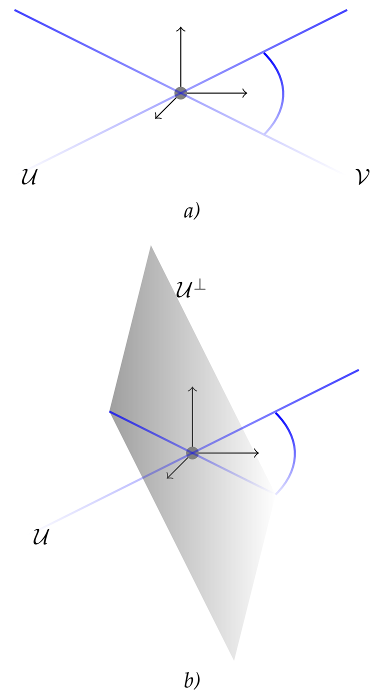
$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = 0, \text{ azaz } \mathbf{a}_{i*} \cdot \mathbf{x} = 0.$$

Eszerint a homogén lineáris egyenletrendszer minden megoldása merőleges az $\mathbf{A}$ mátrix minden sorvektorára. A merőlegesség azonban a sortér összes vektorára is fönnáll a következő állítás következtében.

3.41. ÁLLÍTÁS (A SORTÉR ÉS A NULLTÉR MERŐLEGESSÉGE). *Az* $\mathbf{A}$ *mátrix sorterének bármely* $\mathbf{s}$ *vektora és nullterének tetszőleges* $\mathbf{x}$ *vektora merőleges egymásra, azaz* $\mathbf{s} \cdot \mathbf{x} = \mathbf{0}$.

BIZONYÍTÁS. A sortér minden vektora az $\mathbf{A}$ sorvektorainak valamely $c_1, \ldots, c_m$ skalárokkal vett lineáris kombinációja. Ezt felhasználva

$$\begin{aligned}\mathbf{s} \cdot \mathbf{x} &= (c_1\mathbf{a}_{1*} + c_2\mathbf{a}_{2*} + \cdots + c_m\mathbf{a}_{m*}) \cdot \mathbf{x} \\ &= c_1\mathbf{a}_{1*} \cdot \mathbf{x} + c_2\mathbf{a}_{2*} \cdot \mathbf{x} + \cdots + c_m\mathbf{a}_{m*} \cdot \mathbf{x} \\ &= c_1 0 + c_2 0 + \cdots + c_m 0 = 0.\end{aligned}$$ □

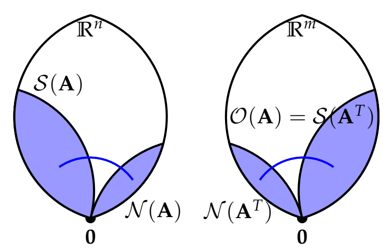


3.12. ábra: *a)* Két merőleges altér $\mathcal{U}$ és $\mathcal{V}$; *b)* Egy altér és merőleges kiegészítő altere: $\mathcal{U}$ és $\mathcal{U}^\perp$.

Ez az állítás a következő definíciókra vezet: egy vektortér két altere *merőleges*, ha bárhogy választva egy vektort az egyik altérből, és egy másikat a másik altérből, azok merőlegesek egymásra. Eszerint az előző állítás úgy fogalmazható meg, hogy bármely mátrix sortere és nulltere merőleges. Ennél azonban több is igaz, a nulltér nem csak merőleges a sortérre, de az összes olyan vektort tartalmazza, mely merőleges a sortérre. Az $\mathbb{R}^n$ egy $\mathcal{W}$ alterére merőleges vektorok alterét a $\mathcal{W}$ *merőleges kiegészítő alterének* nevezzük és $\mathcal{W}^\perp$-pel jelöljük (amit olvashatunk „W perp"-nek).

A két fogalom közti különbséget mutatja a 3.12. ábra, melynek *a)* része a 3-dimenziós tér két, egymásra merőleges 1-dimenziós $\mathcal{U}$ és $\mathcal{V}$ alterét szemlélteti, míg a *b)* rész az $\mathcal{U}$ alteret, és annak 2-dimenziós merőleges kiegészítő $\mathcal{U}^\perp$ alterét szemlélteti.

Vegyük az $\mathbf{A}$ mátrix transzponáltját! Az $\mathbf{A}^T$ együtthatómátrixú homogén lineáris egyenletrendszer megoldásai merőlegesek $\mathbf{A}^T$ sorvektoraira, azaz az $\mathbf{A}$ oszlopvektoraira. E két-két altér merőlegességét szemlélteti a 3.13. ábra. E négy altér igen fontos lesz a továbbiakban is, ezért ezeket a mátrixhoz tartozó *négy kitüntetett altérnek* vagy négy fundamentális altérnek nevezzük. Ezek tehát $\mathcal{S}(\mathbf{A})$, $\mathcal{N}(\mathbf{A})$, $\mathcal{O}(\mathbf{A}) = \mathcal{S}(\mathbf{A}^T)$, $\mathcal{N}(\mathbf{A}^T)$.



3.13. ábra: Az $\mathbf{A}$ mátrix sortere merőleges nullterére, oszloptere az $\mathbf{A}^T$ nullterére. A berajzolt két ív az alterek merőlegességét jelöli.

A nulltér az összes olyan vektort tartalmazza, mely merőleges a sorvektorokra, azaz a sortérre. Vajon a nulltér minden vektorára merőleges vektorok egybeesnek a sortér vektoraival? Más szóval mi a nulltér merőleges kiegészítő altere?

Kiegészítő altér Legyen $\mathcal{V}$ és $\mathcal{W}$ az $\mathbb{R}^n$ két tetszőleges altere. Azt mondjuk, hogy $\mathcal{W}$ a $\mathcal{V}$ *kiegészítő altere*, vagy hogy egymás kiegészítő alterei, ha

$$\mathcal{V} \cap \mathcal{W} = \{\mathbf{0}\}, \quad \mathcal{V} + \mathcal{W} = \mathbb{R}^n,$$

azaz a két altérnek a zérusvektoron kívül nincs közös eleme, és $\mathbb{R}^n$ minden vektora előáll $\mathcal{V}$- és $\mathcal{W}$-beli elemek összegeként!

E fogalom, mint látni fogjuk, a sík koordinátázására emlékeztető dolog: a síkban az origón átmenő két koordinátatengely vektorai a két alteret adják, melyekben csak a zérusvektor közös, és a sík minden vektora (egyértelműen) előáll az egyikből és a másikból vett vektor összegeként.

3.42. TÉTEL (KIEGÉSZÍTŐ ALTEREK TULAJDONSÁGAI). *Legyen $\mathcal{V}$ és $\mathcal{W}$ az $\mathbb{R}^n$ két altere. Az alábbi állítások ekvivalensek:*

a) $\mathcal{V} \cap \mathcal{W} = \{\mathbf{0}\}$ *és* $\mathcal{V} + \mathcal{W} = \mathbb{R}^n$, *azaz $\mathcal{V}$ és $\mathcal{W}$ kiegészítő alterek,*

b) $\mathbb{R}^n$ *minden vektora egyértelműen áll elő egy $\mathcal{V}$- és egy $\mathcal{W}$-beli vektor összegeként,*

c) $\mathcal{V} \cap \mathcal{W} = \{\mathbf{0}\}$ *és* $\dim \mathcal{V} + \dim \mathcal{W} = n$.

BIZONYÍTÁS. *a)* $\Rightarrow$ *b)* : Meg kell mutatnunk, hogy minden vektor egyértelműen áll elő egy $\mathcal{V}$- és egy $\mathcal{W}$-beli vektor összegeként. Legyen $\mathbf{u} \in \mathbb{R}^n$ olyan vektor, hogy $\mathbf{u} = \mathbf{v}_1 + \mathbf{w}_1 = \mathbf{v}_2 + \mathbf{w}_2$, ahol $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}$ és $\mathbf{w}_1, \mathbf{w}_2 \in \mathcal{W}$. Ekkor átrendezés után kapjuk, hogy $\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{w}_2 - \mathbf{w}_1$. Ennek bal oldalán $\mathcal{V}$-beli, jobb oldalán $\mathcal{W}$-beli vektor áll, amik csak a nullvektor esetén lehetnek azonosak, mivel $\mathcal{V} \cap \mathcal{W} = \{\mathbf{0}\}$. Így $\mathbf{v}_1 = \mathbf{v}_2$ és $\mathbf{W}_1 = \mathbf{w}_2$.

b) $\Rightarrow$ *c)* : Legyen $\mathcal{V}$ egy bázisa $\mathbf{v}_1$, $\mathbf{v}_2, \ldots$, $\mathbf{v}_r$, $\mathcal{W}$ egy bázisa $\mathbf{w}_1$, $\mathbf{w}_2, \ldots$, $\mathbf{w}_k$. Mivel bármely $\mathbf{u} \in \mathbb{R}^n$ vektor előáll $\mathbf{v} + \mathbf{w}$ alakban, ahol $\mathbf{v} \in \mathcal{V}$ és $\mathbf{w} \in \mathcal{W}$, e két vektor pedig előáll a bázisvektorok lineáris kombinációjaként, ezért a két bázis egyesítésével kapott vektorrendszer kifeszíti $\mathbb{R}^n$-t, tehát elemszáma legalább n, azaz $r + k \geq n$. Másrészt megmutatjuk, hogy a vektorrendszer független vektorokból áll. Tegyük fel, hogy

$$c_1\mathbf{v}_1 + \cdots + c_r\mathbf{v}_r + d_1\mathbf{w}_1 + \cdots + d_k\mathbf{w}_k = \mathbf{0}.$$

Mivel a $\mathbf{0}$ vektor *egyértelműen áll elő* $\mathbf{v} + \mathbf{w}$ alakban, és egy előállítása a $\mathbf{0} + \mathbf{0}$, ezért

$$c_1\mathbf{v}_1 + \cdots + c_r\mathbf{v}_r = \mathbf{0}, \quad \text{és} \quad d_1\mathbf{w}_1 + \cdots + d_k\mathbf{w}_k = \mathbf{0}.$$

Innen pedig a bázisvektorok lineáris függetlenségéből következik, hogy minden együttható 0. Tehát a vektorok függetlenek, vagyis $r + k \leq n$. Összevetve az előző egyenlőtlenséggel kapjuk, hogy $r + k = n$.

c) $\Rightarrow$ *a)* : Csak azt kell megmutatni, hogy ha $\mathcal{V} \cap \mathcal{W} = \{\mathbf{0}\}$ és $\dim \mathcal{V} + \dim \mathcal{W} = n$, akkor $\mathbb{R}^n = \mathcal{V} + \mathcal{W}$. Ehhez legyen $\mathcal{V}$ egy bázisa $\mathbf{v}_1$, $\mathbf{v}_2, \ldots$, $\mathbf{v}_r$, a $\mathcal{W}$ egy bázisa $\mathbf{w}_1$, $\mathbf{w}_2, \ldots$, $\mathbf{w}_{n-r}$. Ha egyesítésük bázis $\mathbb{R}^n$-ben, akkor kész vagyunk, hisz minden vektor e bázis vektorainak lineáris kombinációja, mely felbomlik $\mathcal{V}$-beli és $\mathcal{W}$-beli részre. Ezért tegyük fel, hogy e vektorok lineárisan összefüggők, azaz a nullvektor megkapható valamely nemtriviális lineáris kombinációval:

$$c_1\mathbf{v}_1 + \cdots + c_r\mathbf{v}_r + d_1\mathbf{w}_1 + \cdots + d_{n-r}\mathbf{w}_{n-r} = \mathbf{0}.$$

Átrendezve

$$c_1\mathbf{v}_1 + \cdots + c_r\mathbf{v}_r = -d_1\mathbf{w}_1 - \cdots - d_{n-r}\mathbf{w}_{n-r},$$

ami ellentmond a $\mathcal{V} \cap \mathcal{W} = \{\mathbf{0}\}$ feltételnek, mivel indirekt feltevésünk szerint nem minden együttható 0. □

Ha két altér a fenti tétel *a)* pontját kielégíti, akkor azt mondjuk, hogy $\mathbb{R}^n$ a $\mathcal{V}$ és $\mathcal{W}$ alterek *direkt összege*, amit az összegtől megkülönböztetendő $\mathcal{V} \oplus \mathcal{W}$ jelöl.

Már láttunk példát kiegészítő alterekre, hisz a sortér és a nulltér dimenziójának összege n, és a két altérnek a nullvektoron kívül nincs közös eleme, így $\mathcal{S}(\mathbf{A})$ és $\mathcal{N}(\mathbf{A})$ kiegészítő alterek, azaz $\mathbb{R}^n = \mathcal{S}(\mathbf{A}) \oplus \mathcal{N}(\mathbf{A})$.

Egy $\mathcal{W}$ altér esetén $\mathcal{W}^\perp$ jelölte a $\mathcal{W}$-re merőleges vektorok alterét. Ezt merőleges kiegészítő altérnek neveztük, de azt, hogy ez valóban kiegészítő altér-e, még nem mutattuk meg.

3.43. TÉTEL (A MERŐLEGES KIEGÉSZÍTŐ ALTÉR TULAJDONSÁGAI). *Legyen $\mathcal{W}$ az $\mathbb{R}^n$ egy altere. Ekkor*
a) $\mathcal{W} \cap \mathcal{W}^\perp = \{\mathbf{0}\}$,
b) $\mathcal{W} + \mathcal{W}^\perp = \mathbb{R}^n$,
c) $\mathbb{R}^n$ *minden vektora egyértelműen előáll egy $\mathcal{W}$- és egy $\mathcal{W}^\perp$-beli vektor összegeként,*
d) $(\mathcal{W}^\perp)^\perp = \mathcal{W}$.

BIZONYÍTÁS. *a)* igaz, hisz ha $\mathbf{x} \in \mathcal{W} \cap \mathcal{W}^\perp$, akkor $\mathbf{x} \cdot \mathbf{x} = 0$, ami csak a $\mathbf{0}$ vektorra áll fönn.

b) abból adódik, hogy az előző, a kiegészítő alterekről szóló tétel szerint, ha két altér dimenzióinak összege n, és a két altér metszete csak a zérusvektorból áll, akkor a két altér összege $\mathbb{R}^n$. Esetünkben a két altér $\mathcal{W}$ és $\mathcal{W}^\perp$. Ha $\mathcal{W}$ egy bázisának vektoraiból, mint sorvektorokból mátrixot képzünk, annak sortere épp $\mathcal{W}$, nulltere $\mathcal{W}^\perp$ lesz, és a sortér és nulltér dimenzióinak összege valóban n a dimenziótétel szerint.

Ugyancsak az előző tétel és az *a)* és *b)* állítások következménye, hogy a „merőleges kiegészítő alterek” valóban kiegészítő alterek, ami bizonyítja *c)*-t.

d) bizonyításához megmutatjuk, hogy $\mathcal{W} \subseteq (\mathcal{W}^\perp)^\perp$ és $\mathcal{W} \supseteq (\mathcal{W}^\perp)^\perp$, ami bizonyítja, hogy $\mathcal{W} = (\mathcal{W}^\perp)^\perp$.

Legyen $\mathbf{w}$ a $\mathcal{W}$ tér egy tetszőleges vektora. Mivel $\mathcal{W}^\perp$ épp azokból a vektorokból áll, melyek merőlegesek $\mathcal{W}$ minden vektorára, ezért $\mathbf{w}$ merőleges $\mathcal{W}^\perp$ minden vektorára. Ez viszont épp azt jelenti, hogy $\mathbf{w}$ benne van a $(\mathcal{W}^\perp)^\perp$ altérben, tehát $\mathcal{W} \subseteq (\mathcal{W}^\perp)^\perp$.

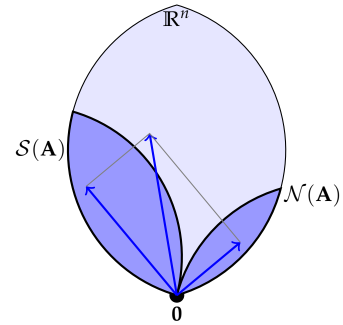
A fordított tartalmazás bizonyításához legyen $\mathbf{w} \in (\mathcal{W}^\perp)^\perp$. A *b)* pont szerint e vektor előáll $\mathbf{w} = \mathbf{v} + \mathbf{v}^\perp$ alakban, ahol $\mathbf{v} \in \mathcal{W}$ és $\mathbf{v}^\perp \in \mathcal{W}^\perp$. Elég lenne megmutatnunk, hogy $\mathbf{v}^\perp = \mathbf{0}$. A $(\mathcal{W}^\perp)^\perp$ és $\mathcal{W}^\perp$ merőlegessége miatt $\mathbf{w} \cdot \mathbf{v}^\perp = 0$, így

$$0 = \mathbf{w} \cdot \mathbf{v}^\perp = \mathbf{w} \cdot (\mathbf{v} + \mathbf{v}^\perp) = \mathbf{v} \cdot \mathbf{v}^\perp + \mathbf{v}^\perp \cdot \mathbf{v}^\perp = \mathbf{v}^\perp \cdot \mathbf{v}^\perp,$$

hisz $\mathbf{v} \cdot \mathbf{v}^\perp = 0$. A $\mathbf{v}^\perp \cdot \mathbf{v}^\perp = 0$ egyenlőség viszont csak $\mathbf{v}^\perp = \mathbf{0}$ esetén áll fönn. Így tehát $\mathbf{w} = \mathbf{v}$, azaz $\mathbf{w} \in \mathcal{W}$, ami bizonyítja az állítást. □

Mindezek alapján eljutottunk oda, hogy megfogalmazzuk a lineáris algebra egyik legfontosabb tételét.

3.44. TÉTEL (A LINEÁRIS ALGEBRA ALAPTÉTELE). *Minden valós mátrix sortere és nulltere merőleges kiegészítő alterei egymásnak.*



3.14. ábra: A lineáris algebra alaptétele: az $\mathbf{A}$ mátrix sortere és nulltere merőleges kiegészítő alterek. Eszerint a sortér bármely vektora merőleges a nulltér bármely vektorára, és $\mathbb{R}^n$ bármely vektora egyértelműen felbomlik egy sortérbe és egy nulltérbe eső vektor összegére.

E tételnek több fontos következménye van. Ezek azonnal adódnak az előző tételekből, valamint ha a lineáris algebra alaptételét az $\mathbf{A}$ mátrix transzponáltjára is alkalmazzuk.

3.45. TÉTEL (A NÉGY KITÜNTETETT ALTÉR). *Tekintsük az $\mathbf{A} \in \mathbb{R}^{m\times n}$ mátrixot. Akkor a következő állítások teljesülnek:*

a) $\mathcal{S}(\mathbf{A})^\perp = \mathcal{N}(\mathbf{A})$, $\mathcal{O}(\mathbf{A})^\perp = \mathcal{N}(\mathbf{A}^T)$.

b) $\mathbb{R}^n$ *minden vektora egyértelműen felbomlik egy $\mathcal{S}(\mathbf{A})$- és egy $\mathcal{N}(\mathbf{A})$-beli vektor összegére, azaz képlettel kifejezve:* $\mathcal{S}(\mathbf{A}) \oplus \mathcal{N}(\mathbf{A}) = \mathbb{R}^n$.

c) $\mathbb{R}^m$ *minden vektora egyértelműen felbomlik egy $\mathcal{O}(\mathbf{A})$- és egy $\mathcal{N}(\mathbf{A}^T)$-beli vektor összegére, azaz képlettel kifejezve:* $\mathcal{O}(\mathbf{A}) \oplus \mathcal{N}(\mathbf{A}^T) = \mathbb{R}^m$.

A lineáris egyenletrendszer megoldásainak jellemzése E fejezet végén eljutottunk oda, hogy nagyon szép leírását tudjuk adni a lineáris egyenletrendszerek megoldásainak.

3.46. TÉTEL (LINEÁRIS EGYENLETRENDSZER MEGOLDÁSAI). *Minden megoldható (konzisztens) lineáris egyenletrendszerre igazak a következő állítások:*

a) egyetlen megoldása esik az együtthatómátrix sorterébe;

b) a sortérbe eső megoldás az összes megoldás közül a legkisebb abszolút értékű;

c) az összes megoldás előáll úgy, hogy a sortérbe eső megoldáshoz hozzáadjuk a homogén rész összes megoldását.

BIZONYÍTÁS. A tétel a homogén lineáris egyenletrendszerekre semmitmondó, hisz ekkor a megoldások a nullteret adják, és mivel annak metszete a sortérrel csak a nullvektorból áll, ezért csak a nullvektor esik a sortérbe, mely természetesen a legkisebb abszolút értékű megoldás. Ráadásul a nullvektort hozzáadva a nulltérhez, valóban a nullteret kapjuk, vagyis az összes megoldások terét. Így ezután csak az inhomogén esettel foglalkozunk.

a) Tegyük fel, hogy $\mathbf{x}_1$ és $\mathbf{x}_2$ két megoldása az $[\mathbf{A}|\mathbf{b}]$ mátrixú egyenletrendszernek, és mindkettő a sortérbe esik. Az i-edik egyenlet alakja $\mathbf{a}_{i*} \cdot \mathbf{x} = b_i$, így $\mathbf{a}_{i*} \cdot \mathbf{x}_1 = b_i$ és $\mathbf{a}_{i*} \cdot \mathbf{x}_2 = b_i$ is fönnáll minden $i = 1, 2, \dots m$ értékre. A két megoldás különbsége is a sortérbe esik, hisz sortérbeli vektorok lineáris kombinációja a sortérbe esik. Ekkor viszont minden i esetén

$$\mathbf{a}_{i*} \cdot (\mathbf{x}_1 - \mathbf{x}_2) = b_i - b_i = 0,$$

vagyis $\mathbf{x}_1 - \mathbf{x}_2$ megoldása a homogén egyenletrendszernek, tehát a nulltérbe esik. Annak metszete a sortérrel csak a nullvektort tartalmazza, így $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}$, vagyis $\mathbf{x}_1 = \mathbf{x}_2$.

Meg kell még mutatnunk, hogy mindig van a sortérbe eső megoldás. Legyen $\mathbf{x}$ egy tetszőleges megoldás, és tekintsük az egyértelműen létező felbontását egy sortérbeli és egy nulltérbeli vektor összegére, azaz legyen

$$\mathbf{x} = \mathbf{x}_{\mathcal{S}} + \mathbf{x}_{\mathcal{N}}.$$

E megoldásvektort beírva az i-edik egyenletbe kapjuk, hogy

$$b_i = \mathbf{a}_{i*} \cdot \mathbf{x} = \mathbf{a}_{i*} \cdot (\mathbf{x}_{\mathcal{S}} + \mathbf{x}_{\mathcal{N}}) = \mathbf{a}_{i*} \cdot \mathbf{x}_{\mathcal{S}} + \mathbf{a}_{i*} \cdot \mathbf{x}_{\mathcal{N}} = \mathbf{a}_{i*} \cdot \mathbf{x}_{\mathcal{S}}.$$

Ez azt jelenti, hogy bármely megoldás sortérbeli összetevője is megoldása az egyenletrendszernek! Találtunk tehát egy megoldást a sortérben. Egyúttal azt is beláttuk, hogy az összes megoldás e sortérbeli megoldás és a homogén egy megoldásának összege. Az előző egyenlőségekből az is kiolvasható, hogy az $\mathbf{x}_{\mathcal{S}}$ megoldáshoz bármely nulltérbeli vektort adva az egyenletrendszer egy megoldását kapjuk, igazoltuk tehát a *c)* állítást is.

A sortér és a nulltér merőlegessége miatt az $\mathbf{x} = \mathbf{x}_{\mathcal{S}} + \mathbf{x}_{\mathcal{N}}$ felbontás vektorai merőlegesek, azaz $\mathbf{x}_{\mathcal{S}} \perp \mathbf{x}_{\mathcal{N}}$. Használhatjuk tehát Pithagorász-tételét:

$$\mathbf{x}^2 = \mathbf{x}_{\mathcal{S}}^2 + \mathbf{x}_{\mathcal{N}}^2 \geq \mathbf{x}_{\mathcal{S}}^2, \text{ azaz } |\mathbf{x}| \geq |\mathbf{x}_{\mathcal{S}}|.$$

Így tehát minden megoldás abszolút értéke nagyobb vagy egyenlő a sortérbeli megoldás abszolút értékénél, ami bizonyítja a *b)* állítást is. □

A sortérbe eső egyetlen megoldás létezése azt sugallja, hogy minden megoldható egyenletrendszer további egyenletek hozzávételével kiegészíthető olyan egyenletrendszerré, melynek már csak egyetlen megoldása van. Ez valóban igaz.

3.47. PÉLDA (LINEÁRIS EGYENLETRENDSZER SORTÉRBE ESŐ MEGOLDÁSA). *Határozzuk meg az*

$$\begin{aligned}
x + y + z + 3u + 2w &= 4\\
x + 2y + z + 5u + 2w &= 5\\
2x + 3y + z + 8u + 3w &= 7\\
2x + 3y + 2z + 8u + 4w &= 9
\end{aligned}$$

egyenletrendszer minimális abszolút értékű megoldását! Ehhez adjunk az egyenletrendszerhez olyan további egyenlet(ek)et, hogy az így kapott egyenletrendszernek már csak ez legyen az egyetlen megoldása!

MEGOLDÁS. Először oldjuk meg az egyenletrendszert! A bővített mátrixból annak redukált lépcsős alakja könnyen adódik:

$$\begin{bmatrix} 1 & 1 & 1 & 3 & 2 & 4 \\ 1 & 2 & 1 & 5 & 2 & 5 \\ 2 & 3 & 1 & 8 & 3 & 7 \\ 2 & 3 & 2 & 8 & 4 & 9 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 2 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 2 \end{bmatrix}$$

Így a megoldás:

$$(x, y, z, u, w) = (1, 1, 2, 0, 0) + (-1, -2, 0, 1, 0)u + (-1, 0, -1, 0, 1)w.$$

Mivel a sortér merőleges a nulltérre, és mi egy sortérbe eső megoldást keresünk, ezért e megoldásnak merőlegesnek kell lennie a nullteret kifeszítő vektorokra, vagyis a $(-1, -2, 0, 1, 0)$ és a $(-1, 0, -1, 0, 1)$ vektorra. Így a következő két egyenletet kell az eredeti egyenletrendszerhez, vagy az egyszerűség kedvéért inkább a redukált lépcsős alak szerinti egyenletrendszerhez adni:

$$\begin{aligned} -x - 2y \quad + u \quad &= 0 \\ -x \quad - z \quad + w &= 0 \end{aligned}$$

Így a kiegészített egyenletrendszer bővített mátrixa és annak redukált lépcsős alakja

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 2 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 2 \\ -1 & -2 & 0 & 1 & 0 & 0 \\ -1 & 0 & -1 & 0 & 1 & 0 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -4/17 \\ 0 & 1 & 0 & 0 & 0 & 5/17 \\ 0 & 0 & 1 & 0 & 0 & 19/17 \\ 0 & 0 & 0 & 1 & 0 & 6/17 \\ 0 & 0 & 0 & 0 & 1 & 15/17 \end{bmatrix},$$

tehát a keresett megoldás $(-4/17, 5/17, 19/17, 6/17, 15/17)$. □

Feladatok

A sortérbe eső megoldás meghatározása *Keressük meg az alábbi egyenletrendszerek sortérbe eső egyetlen megoldását, és annak segítségével írjuk fel összes megoldását!*

3.11.•
$$\begin{aligned} x + y + z &= 3 \\ 2x + y - z &= 2 \\ 3x + 2y &= 5 \end{aligned}$$

3.12. $x + 4y + 8z + 12w = 15$

3.13.•
$$\begin{aligned} x + y + z + w &= 3 \\ x + y - z - w &= 1 \end{aligned}$$

Bizonyítások

3.14. Mutassuk meg, hogy tetszőleges $H \subseteq \mathbb{R}^n$ vektorhalmazra a következő két állítás ekvivalens:

1. bármely véges sok H-beli vektor bármely lineáris kombinációja H-ban van;
2. bármely H-beli vektor tetszőleges skalárszorosa, és bármely két H-beli vektor összege H-ban van.

3.15. **Lineáris függetlenség egy szükséges és elégséges feltétele** A V vektorrendszer pontosan akkor lineárisan független, ha $\operatorname{span}(V)$ bármely vektora csak egyféleképp áll elő V lineáris kombinációjaként.

3.16. **Sortér és nulltér** A 2.43. példában megoldottuk a

$$\begin{aligned} x_1 + 2x_2 + x_3 + 2x_4 + x_5 &= 0 \\ x_1 + 2x_2 + 3x_3 + 3x_4 + x_5 &= 0 \\ 3x_1 + 6x_2 + 7x_3 + 8x_4 + 3x_5 &= 0 \end{aligned}$$

egyenletrendszert. A megoldása

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2s - \frac{3}{2}t - u \\ s \\ -\frac{1}{2}t \\ t \\ u \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} s + \begin{bmatrix} -\frac{3}{2} \\ 0 \\ -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} t + \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} u.$$

Ezt fölhasználva fejben számolva adjunk meg egy olyan vektorrendszert, amely kifeszíti az

$$\begin{aligned} -2x_1 + x_2 \qquad\qquad\quad &= 0 \\ -3x_1 \qquad - x_3 + 2x_4 \qquad &= 0 \\ -x_1 \qquad\qquad\qquad + x_5 &= 0 \end{aligned}$$

homogén lineáris egyenletrendszer megoldásai alterét.

3.17. Általánosítsuk az előző feladat eredményét tetszőleges homogén lineáris egyenletrendszerre!

Megoldások

1.10. Ha $|\overline{AP}| : |\overline{PB}| = m : n$, akkor $|\overline{AB}| : |\overline{PB}| = (m+n) : n$, amiből $\overrightarrow{BP} = \frac{n}{m+n}\overrightarrow{BA}$. De $\overrightarrow{OP} = \overrightarrow{OB} + \overrightarrow{BP}$ és $\overrightarrow{BA} = \overrightarrow{OA} - \overrightarrow{OB}$, így $\overrightarrow{OP} = \overrightarrow{OB} + \frac{n}{m+n}(\overrightarrow{OA} - \overrightarrow{OB})$, amiből azonnal következik a bizonyítandó formula. A felezőpontot az $m = n = 1$ esetben kapjuk, és ekkor valóban $\overrightarrow{OP} = \frac{1}{2}\overrightarrow{OA} + \frac{1}{2}\overrightarrow{OB}$.

1.12. Legyenek $\mathbf{a}$ és $\mathbf{c}$ független vektorok, $\mathbf{b}$ pedig tetszőleges. Ekkor az $(\mathbf{a}\cdot\mathbf{b})\mathbf{c}$ szorzat párhuzamos a $\mathbf{c}$ vektorral, míg az $\mathbf{a}(\mathbf{b}\cdot\mathbf{c})$ szorzat az $\mathbf{a}$ vektorral, tehát $(\mathbf{a}\cdot\mathbf{b})\mathbf{c} \neq \mathbf{a}(\mathbf{b}\cdot\mathbf{c})$.

1.13. Jelölje P szomszédait Q, R és S.

a) Ekkor két lapátló-vektor például a $\overrightarrow{PQ} + \overrightarrow{PR}$ és a $\overrightarrow{PR} + \overrightarrow{PS}$ vektorok. Ezek szorzata:

$$\left(\overrightarrow{PQ} + \overrightarrow{PR}\right) \cdot \left(\overrightarrow{PR} + \overrightarrow{PS}\right) =$$
$$\overrightarrow{PQ} \cdot \overrightarrow{PR} + \overrightarrow{PQ} \cdot \overrightarrow{PS} + \overrightarrow{PR} \cdot \overrightarrow{PR} + \overrightarrow{PR} \cdot \overrightarrow{PS} =$$
$$\overrightarrow{PR} \cdot \overrightarrow{PR} = 1.$$

Kihasználtuk, hogy merőleges vektorok skaláris szorzata 0.

b) Hasonlóan kapható meg egy lapátló-vektor és a testátló-vektor $(\overrightarrow{PQ} + \overrightarrow{PR} + \overrightarrow{PS})$ szorzata:

$$\left(\overrightarrow{PQ} + \overrightarrow{PR}\right) \cdot \left(\overrightarrow{PQ} + \overrightarrow{PR} + \overrightarrow{PS}\right) = \overrightarrow{PQ} \cdot \overrightarrow{PQ} + \overrightarrow{PR} \cdot \overrightarrow{PR} = 2.$$

c) A Q, R és S csúcsok olyan sorrendben legyenek megválasztva, hogy $\overrightarrow{PQ}$, $\overrightarrow{PR}$ és $\overrightarrow{PS}$ ebben a sorrendben jobbrendszert alkosson. Ki fogjuk használni, hogy ekkor $\overrightarrow{PQ} \times \overrightarrow{PR} = \overrightarrow{PS}$. Egy élvektor és egy szomszédos lapátló-vektor vektori szorzata:

$$\overrightarrow{PQ} \times \left(\overrightarrow{PQ} + \overrightarrow{PR}\right) =$$
$$\overrightarrow{PQ} \times \overrightarrow{PQ} + \overrightarrow{PQ} \times \overrightarrow{PR} =$$
$$\mathbf{0} + \overrightarrow{PS} = \overrightarrow{PS},$$

vagyis a szorzat a két vektor lapjára merőleges élvektor.

d) Legyen a lapvektor a $\overrightarrow{PR}$, a nem szomszédos lapátló-vektor $\overrightarrow{PR} + \overrightarrow{PS}$. Ezek szorzata:

$$\overrightarrow{PQ} \times \left(\overrightarrow{PR} + \overrightarrow{PS}\right) =$$
$$\overrightarrow{PQ} \times \overrightarrow{PR} + \overrightarrow{PQ} \times \overrightarrow{PS} = \overrightarrow{PS} - \overrightarrow{PR},$$

ami a lapátló-vektor síkjának másik lapátló-vektora.

1.15. Három különböző dolog (így három vektor is) hatféleképp rakható sorba. Ha az $\mathbf{a}$, $\mathbf{b}$ és $\mathbf{c}$ vektorok jobbrendszert alkotnak, akkor ugyancsak jobbrendszert alkotnak a $\mathbf{b}$, $\mathbf{c}$, $\mathbf{a}$ és a $\mathbf{c}$, $\mathbf{a}$, $\mathbf{b}$ vektorhármasok is. A további három esetben, azaz a $\mathbf{c}$, $\mathbf{b}$, $\mathbf{a}$, valamint a $\mathbf{b}$, $\mathbf{a}$, $\mathbf{c}$ és az $\mathbf{a}$, $\mathbf{c}$, $\mathbf{b}$ hármasok esetén balrendszert kapunk. A „bizonyítás” egyelőre a kezünk három ujjáról való leolvasással történik. Később matematikai bizonyítást is adunk, lásd ???.

1.16. Egyik lehetőség a megoldásra: $||\mathbf{b}|\mathbf{a}| = ||\mathbf{a}|\mathbf{b}| = |\mathbf{a}||\mathbf{b}|$, ezért a parallelogramma-módszert egy rombuszra kell alkalmazni. Egy másik lehetőség: az $\mathbf{a}/|\mathbf{a}|$ és $\mathbf{b}/|\mathbf{b}|$ két egységvektor, így összegük szögfelező, mivel a parallelogramma-módszer rombuszt ad. E vektor $|\mathbf{a}||\mathbf{b}|$-szerese ugyanúgy szögfelező, és épp ez a feladatbeli vektor.

1.18. Milyen irányokat cserél föl a tükör, és milyeneket nem? Nem cseréli föl a síkkal párhuzamos irányokat: minden, a tükör síkjával párhuzamos vektor tükörképe önmaga. Tehát, ha a tükör előtt állunk, és a tükör is függőleges, akkor a „fölfelé” irány a tük9örképen sem változik. Viszont a tükör fölcseréli a tükörre merőleges irányokat. Mielőtt megnézzük, hogy hogy cserélődik fel a jobb és a bal, definiálnunk kell mi az, hogy jobb és bal? ....

1.22. Minden olyan vektor, amelyben páros sok 1-es van, eleget tesz a feladatbeli feltételnek, a többi nem.

2.1. *a)* igen, *b)* igen, *c)* nem, *d)* igen, *e)* igen, *f)* nem.

2.2. Könnyen látható, hogy mindkét egyenletrendszer egyetlen megoldása: $x = 2$, $y = 1$, tehát a két egyenletrendszer ekvivalens.

2.3. Az első egyenletrendszer nem oldható meg a $0 = 3$ alakú egyenlet miatt, de a második sem, hisz nincs olyan x és y, melyre $x + y = 2$ és $x + y = 7$ lenne, hisz $2 \neq 7$.

2.4. Behelyettesítés után mindkét egyenlet $0 = 0$ alakú, amit tetszőleges x és y kielégít, így az összes (x, y) számpár megoldása az egyenletrendszernek.

2.5. $x = 1$, y tetszőleges, azaz az összes $(1, y)$ alakú számpár megoldás.

2.6. A második egyenlet behelyettesítés után $0 = 1$ alakú, így az egyenletrendszernek nincs megoldása.

2.7. $x = 1$, $y = 2$, azaz $(x, y) = (1, 2)$ az egyetlen megoldás.

2.9. *a)* A sormodellben két metsző egyenest kell megrajzolni ($y = 7/3x - 2/3x$, $y = -2 + 3/2x$), melyek a $(2, 1)$ pontban metszik egymást, míg az oszlopmodellben a $(2, 3)$, a $(3, -2)$ vektorokat és azok lineáris kombinációjaként előállított $(7, 4) = 2(2, 3) + (3, -2)$ vektort!

b) A sormodellben két párhuzamos egyenest kell megrajzolni, míg az oszlopmodellben a $(2, 3)$ és a $(4, 6)$ vektorokat, melyek egy egyenesbe esnek, és semmilyen lineáris kombinációjuk sem adja ki a $(3, 4)$ vektort!

2.10. A három sík közül semelyik kettő nem párhuzamos, másrészt a normálvektorai egy síkba esnek, ugyanis $2(1,1,2) + (1,2,4) = (3,4,8)$. Ez azt jelenti, hogy van olyan vektor, mely mindhárom síkkal párhuzamos. Az első esetben a három sík egy egyenesen megy át, mivel van a síkoknak közös pontjuk, pl. a $(3,0,0)$ pont, így végtelen sok megoldása is van, míg a második esetben a síkoknak nincs közös pontjuk.

Az egyenletrendszerek ekvivalensek a következő vektoregyenletekkel:

$$\begin{bmatrix}1\\1\\3\end{bmatrix}x + \begin{bmatrix}1\\2\\4\end{bmatrix}y + \begin{bmatrix}2\\4\\8\end{bmatrix}z = \begin{bmatrix}3\\3\\9\end{bmatrix}, \quad \begin{bmatrix}1\\1\\3\end{bmatrix}x + \begin{bmatrix}1\\2\\4\end{bmatrix}y + \begin{bmatrix}2\\4\\8\end{bmatrix}z = \begin{bmatrix}3\\3\\1\end{bmatrix}.$$

Itt a közös együtthatómátrix minden oszlopvektora benne van a

$$2x + y - z = 0$$

egyenletű síkban, (ez könnyen ellenőrizhető a vektorok koordinátáinak a sík egyenletébe való helyettesítésével), és ki is feszítik a síkot, mert a három vektor nem kollineáris. Másrészt a $(3,3,9)$ vektor is benne van e síkban, a $(3,3,1)$ vektor viszont nem. Tehát az első egyenletrendszer megoldható, a második nem.

2.11. A sormodell szerinti ábra az *a)* esetben 3 síkbeli egyenest tartalmaz, melyek közt van két párhuzamos, így az egyenletrendszer nem oldható meg. A *b)* esetben a három egyenes egy ponton megy át, ez a megoldás: $x = 2$, $y = 1$. A *c)* esetben ugyan nincsenek párhuzamos egyenesek, de nincs közös pontjuk sem, így az egyenletrendszer nem oldható meg. Az oszlopmodell szerint az *a)*

$$\begin{aligned} x + \ \ y &= 3\\ x + \ \ y &= 4\\ x + 2y &= 4 \end{aligned}$$

egyenletrendszer ekvivalens a következővel:

$$\begin{bmatrix}1\\1\\1\end{bmatrix}x + \begin{bmatrix}1\\1\\2\end{bmatrix}y = \begin{bmatrix}3\\4\\4\end{bmatrix}.$$

Az $(1,1,1)$ és az $(1,1,2)$ vektorok benne fekszenek az $x = y$ egyenletű síkban, mivel első két koordinátájuk megegyezik, ezért minden lineáris kombinációjuk is ebbe a síkba esik. A $(3,4,4)$ vektor viszont nem esik e síkba, így *független* az előbbi kettőtől, tehát nem áll elő azok lineáris kombinációjaként. Vagyis ez az egyenletrendszer nem oldható meg. A *b)* és *c)* egyenletrendszerekben a bal oldali két vektor, az $(1,1,1)$ és az $(1,2,3)$ az $x - 2y + z = 0$ egyenletű síkban van, melyben a $(3,4,5)$ vektor benne van, míg a $(3,3,5)$ vektor nincs benne, tehát *b)* megoldható, *c)* nem.

2.12. *a)* hamis, az állítás csak úgy igaz, ha a párhuzamos hipersíkok különbözőek is (két azonos hipersíkot párhuzamosnak tekintünk), *b)* hamis, például a 2.9. *(a)* ábrán látható esetben nincsenek párhuzamos síkok, és mégsincs megoldás, *c)* igaz, mert akkor a jobb oldalon álló bármely vektor kifejezhető e két kétdimenziós vektor lineáris kombinációjaként, tehát az egyenletrendszer megoldható.

2.13.

a) Egy két egyenletből álló háromismeretlenes egyenletrendszer sormodellje szerinti ábra a **három**dimenziós térben **két** darab **síkból** áll, melyek ha **párhuzamosak, de nem azonosak**, akkor az egyenletrendszernek nincs megoldása, egyébként megoldásainak száma **végtelen**. Oszlopmodellje a **két**dimenziós térben **négy** darab **vektorból** áll (három lineáris kombinációja a negyedik).

b) Egy három egyenletből álló kétismeretlenes egyenletrendszer sormodellje szerinti ábra a **két**dimenziós térben **három egyenesből** áll, míg az oszlopmodellje a **három**dimenziós térben **három** darab **vektorból**.

c) Egy négy egyenletből álló ötismeretlenes egyenletrendszer sormodellje szerinti ábra az **öt**dimenziós térben **négy** darab **hipersíkból** áll. Oszlopmodellje a **négy**dimenziós térben **hat** darab **vektorból** áll.

2.14. *a)* Igaz, ez a szám megegyezik a főelemek számával. *b)* Igaz, ez a szám megegyezik a főelemek számával. *c)* Hamis, van lépcsős alakja minden mátrixnak, de csak a redukált lépcsős alak egyértelmű. *d)* Hamis. *e)* Igaz.

2.15. *a)* Igaz. *b)* Hamis, az egyenletrendszer megoldhatósága nem függ az egyenletek számától. Az ismeretlenek számánál akár kevesebb, akár több egyenletből álló rendszer akár ellentmondásos, akár konzisztens is lehet. *c)* Igaz. *d)* Igaz, a nullvektor mindig megoldás.

2.16. $\begin{bmatrix}1&0&1&0&-1\\0&1&0&1&2\\0&0&0&0&0\end{bmatrix}$

2.17. $\begin{bmatrix}1&0&0&0&0\\0&1&0&1&2\\0&0&1&0&-1\end{bmatrix}$

2.18. Igen, mindkettőnek $(x,y,z) = (1,2,3)$ az egyetlen megoldása, azaz megoldáshalmazaik megegyeznek.

2.19. Igen, mindkettő bővített együtthatómátrixának redukált lépcsős alakja a zérussor nélkül

$$\begin{bmatrix}1&0&-0.8&1.8\\0&1&2.2&-1.2\end{bmatrix}$$

2.20. *a)*Igen. *b)*Nem.

2.21. Az $S_i \leftrightarrow S_j$ művelet eredményét önmaga visszaállítja, az cS_i műveletét $\frac{1}{c}S_i$, és az $S_i + cS_j$ műveletét $S_i - cS_j$.

2.22. Legyen $\mathbf{x}_0 = (0, 0)$. Az iteráció képletei:

$$x = \frac{y+8}{4}, \quad y = \frac{2x+5}{5}.$$

Az iteráció lépéseinek táblázata 3 értékes jeggyel számolva:

	$\mathbf{x}_0$	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_3$	$\mathbf{x}_4$	$\mathbf{x}_5$	$\mathbf{x}_6$
x	0	2	2.25	2.45	2.48	2.50	2.50
y	0	1	1.80	1.90	1.98	1.99	2.00

Az iteráció lépéseinek táblázata 4 értékes jeggyel számolva:

	$\mathbf{x}_0$	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_3$	$\mathbf{x}_4$	$\mathbf{x}_5$	$\mathbf{x}_6$	$\mathbf{x}_7$	$\mathbf{x}_8$
x	0	2	2.25	2.45	2.475	2.495	2.498	2.500	2.5
y	0	1	1.80	1.90	1.980	1.990	1.998	1.999	2.0

2.24.

2.25. A Jacobi-iteráció szerinti módon, a vonatok valamelyikének és a légynek a k-adik találkozásából kiszámítva a $k+1$-edik találkozásra jellemző távolságokat, az $x_{k+1} = ay_k + b$, $y_{k+1} = cy_k + d$ egyenletekre jutunk. Az első táblázat adatait behelyettesítve, és a, b, c és d értékre megoldva az

3.1. Mindegyik állítás hamis.

3.2. 1. Igaz. 2. Igaz. 3. Hamis, csak akkor igaz, ha egyik a másik altere. 4. Igaz. 5. Igaz. 6. Hamis, a zérustér az egyetlen zérusvektorból áll.

3.3. 1. Hamis, csak a homogén lineáris egyenletrendszer megoldásai alkotnak alteret, az inhomogéné eltolt alteret. 2. Igaz. Ez épp az oszloptér, ugyanis csak az oszloptérből való $\mathbf{b}$ vektorokra oldható meg az egyenletrendszer. 3. 4. 5.

3.4. 1. Igaz. 2. Igaz. 3. Hamis. 4. Igaz, ugyanis az állításbeli $\mathrm{r}(\mathbf{A}|\mathbf{b}) \le \mathrm{r}(\mathbf{A})$ feltétel pontosan akkor teljesül, ha $\mathrm{r}(\mathbf{A}|\mathbf{b}) = \mathrm{r}(\mathbf{A})$, és ez pontosan akkor teljesül, ha az egyenletrendszer megoldható. 5. Hamis, ha $\mathrm{r}(\mathbf{A}) = n$, és az egyenletrendszer több, mint n egyenletből áll, akkor előfordulhat, hogy $\mathrm{r}(\mathbf{A}|\mathbf{b}) = n+1$, és ekkor az egyenletrendszer nem oldható meg!

3.5. Első ránézésre csak annyi látszik, hogy mindkét megoldás egy kétdimenziós altér eltoltja. Először megvizsgáljuk, hogy a két altér – vagyis az egyenletrendszer homogén részére adott két megoldás – egybeesik-e. Elég megmutatni, hogy az egyik altérben benne van a másikat generáló két vektor. Ha igen, a két altér megegyezik. Ezesetben el kell dönteni, hogy az inhomogén két partikuláris megoldása az altérnek ugyanabban az eltoltjában van-e. Vagy egyszerűbben, hogy a két partikuláris megoldás különbsége benne van-e az altérben. E kérdéseket egyetlen mátrix lépcsős alakra hozásával is megoldhatjuk. Az első két oszlop az első, a második két oszlop a második altér generátorait tartalmazza, az ötödik oszlop a két partikuláris megoldás különbsége.

$$\left[\begin{array}{rr|rr|r} 1 & 0 & -2 & 3 & 1 \\ 0 & 1 & 1 & -2 & -1 \\ -2 & -3 & 1 & 0 & 1 \\ -1 & -2 & 0 & 1 & 1 \end{array}\right] \Longrightarrow \left[\begin{array}{rr|rr|r} 1 & 0 & -2 & 3 & 1 \\ 0 & 1 & 1 & -2 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Az eredményből látszik, hogy a két megoldás azonos.

3.6. Ha $\mathbf{a}_{i*}$ jelöli az együtthatómátrix i-edik sorát és $\mathbf{x}$, illetve $\mathbf{y}$ a homogén egyenletrendszer egy-egy megoldását, azaz $\mathbf{a}_{i*} \cdot \mathbf{x} = 0$, $\mathbf{a}_{i*} \cdot \mathbf{y} = 0$ $(i = 1, 2, \ldots, m)$, akkor

$$\mathbf{a}_{i*} \cdot (c\mathbf{x} = d\mathbf{y}) = c\mathbf{a}_{i*} \cdot \mathbf{x} + d\mathbf{a}_{i*}\mathbf{y} = 0 + 0 = 0,$$

tehát a két megoldásvektor bármely lineáris kombinációja is megoldás. Másként fogalmazva a homogén lineáris egyenletrendszerek megoldásainak bármely lineáris kombinációja is megoldás, tehát a megoldások alteret alkotnak.

3.7. Legyen $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ az inhomogén egy partikuláris megoldása, és jelölje $\mathbf{a}_1$, $\mathbf{a}_2$,…,$\mathbf{a}_n$ az $\mathbf{A}$ oszlopvektorait, H a homogén, I az inhomogén egyenletrendszer általános megoldását. Megmutatjuk, hogy $\mathbf{x} + H = I$, ahol a bal oldali összeadást elemenként értjük.

$\mathbf{x} + H \subseteq I$: Meg kell mutatnunk, hogy ha $\mathbf{x}$-hez adjuk a H egy tetszőleges $\mathbf{y} = (y_1, y_2, \ldots, y_n) \in H$ elemét, az inhomogén egyenletrendszer egy megoldását kapjuk. Valóban, $\mathbf{x}$, illetve $\mathbf{y}$ eleget tesz az

$$\begin{aligned} \mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \ldots + \mathbf{a}_n x_n &= \mathbf{b}, \text{ illetve} \\ \mathbf{a}_1 y_1 + \mathbf{a}_2 y_2 + \ldots + \mathbf{a}_n y_n &= \mathbf{0} \end{aligned}$$

egyenletnek. Ebből

$$\begin{aligned} &\mathbf{a}_1(x_1 + y_1) + \mathbf{a}_2(x_2 + y_1) + \ldots + \mathbf{a}_n(x_n + y_1) = \\ &(\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \ldots + \mathbf{a}_n x_n) + (\mathbf{a}_1 y_1 + \mathbf{a}_2 y_2 + \ldots + \mathbf{a}_n y_n) = \\ &\mathbf{b} + \mathbf{0} = \mathbf{b}, \end{aligned}$$

tehát $\mathbf{x} + \mathbf{y}$ megoldása az inhomogén egyenletrendszernek, azaz $\mathbf{x} + \mathbf{y} \in I$.

$\mathbf{x} + H \supseteq I$: Meg kell mutatnunk, hogy ha $\mathbf{z}$ az inhomogén egy tetszőleges megoldása, azaz $\mathbf{z} \in I$, akkor található olyan $\mathbf{y} \in H$, hogy $\mathbf{z} = \mathbf{x} + \mathbf{y}$. Valóban, az $\mathbf{y} = \mathbf{z} - \mathbf{x}$ megteszi, mert

$$\begin{aligned} &\mathbf{a}_1(z_1 - x_1) + \mathbf{a}_2(z_2 - x_1) + \ldots + \mathbf{a}_n(z_n - x_1) = \\ &(\mathbf{a}_1 z_1 + \mathbf{a}_2 z_2 + \ldots + \mathbf{a}_n z_n) - (\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \ldots + \mathbf{a}_n x_n) = \\ &\mathbf{b} - \mathbf{b} = \mathbf{0}, \end{aligned}$$

azaz $\mathbf{z} - \mathbf{x} \in H$. Ezzel kész a bizonyítás.

3.8. Összesen 16 altere van $\mathbb{F}_2^3$-nek. Van egy 0-dimenziós, a $\mathcal{Z} = \{\mathbf{0}\}$ tér. Az egydimenziós alterek a nullvektorból és egyetlen tőle különböző további vektorból állnak (7 ilyen altér van). A kétdimenziós alterek mindegyike a nullvektorból, két további egymástól is különböző vektorból és azok összegéből áll. Ezeket felsoroljuk:

$$\begin{gathered}\{(0,0,0),(1,0,0),(0,1,0),(1,1,0)\},\\ \{(0,0,0),(0,1,0),(0,0,1),(0,1,1)\},\\ \{(0,0,0),(0,0,1),(1,0,0),(1,0,1)\},\\ \{(0,0,0),(1,0,0),(0,1,1),(1,1,1)\},\\ \{(0,0,0),(0,1,0),(1,0,1),(1,1,1)\},\\ \{(0,0,0),(0,0,1),(1,1,0),(1,1,1)\},\\ \{(0,0,0),(0,1,1),(1,0,1),(1,1,0)\}.\end{gathered}$$

Végül altér maga $\mathbb{F}_2^3$ is.

3.9. 1. Igaz. 2. Igaz, bármely 1-dimenziós altér ilyen. 3. Hamis, ha van kételemű bázis, akkor a lineárisan független vektorrendszerek elemszáma legföljebb 2.

3.10. E mátrix rangja pontosan akkor k, ha oszlopvektorai lineárisan függetlenek, azaz ha az oszlopvektorok bármely lineáris kombinációja csak úgy lehet a nullvektor, ha minden együttható 0. Tekintsünk az oszlopvektorok egy $c_1,\dots,$ c_k skalárokkal vett, nullvektort adó lineáris kombinációját. Ennek i-edik koordinátája

$$\begin{aligned}0&=c_1\mathbf{v}_i\cdot\mathbf{v}_1+c_2\mathbf{v}_i\cdot\mathbf{v}_2+\dots+c_k\mathbf{v}_i\cdot\mathbf{v}_k\\ &=\mathbf{v}_i\cdot(c_1\mathbf{v}_1+c_2\mathbf{v}_2+\dots+c_k\mathbf{v}_k).\end{aligned}$$

Tehát azt kaptuk, hogy az $\mathbf{x}=c_1\mathbf{v}_1+c_2\mathbf{v}_2+\dots+c_k\mathbf{v}_k$ vektor olyan, hogy a $\mathbf{v}_1,\dots,$ $\mathbf{v}_k$ vektorok mindegyikével vett skaláris szorzata 0, így ezek bármelyik lineáris kombinációjával vett skaláris szorzata is 0, tehát például az $\mathbf{x}$ vektorral vett szorzat is 0, azaz $\mathbf{x}\cdot\mathbf{x}=0$. Ez viszont csak $\mathbf{x}=0$ esetén állhat fönn, és mivel a $\mathbf{v}_i$ vektorok lineárisan függetlenek, csak a $c_i=0$ konstansokkal vett lineáris kombinációjuk lehet 0, ahol $i=1,2,\dots,k$.

3.11. Az egyenletrendszer bővített mátrixának redukált lépcsős alakja

$$\begin{bmatrix}1&0&-2&-1\\0&1&3&4\end{bmatrix}$$

így a megoldása $(x,y,z)=(-1,4,0)+(2,-3,1)t$. A nullteret a $(2,-3,1)$ vektor feszíti ki, a sortérbe eső vektornak erre merőlegesnek kell lennie, tehát fönn kell állnia a

$$2x-3y+z=0$$

egyenletnek is. Ezt az egyenletet a redukált lépcsős alakból származó egyenletrendszerhez (vagy akár az eredetihez) adva egy egyetlen megoldást adó egyenletrendszert kapunk. Ennek bővített mátrixa és annak redukált lépcsős alakja:

$$\begin{bmatrix}1&0&-2&-1\\0&1&3&4\\2&-3&1&0\end{bmatrix}\Longrightarrow\begin{bmatrix}1&0&0&1\\0&1&0&1\\0&0&1&1\end{bmatrix}$$

Innen a sortérbe eső megoldás $(1,1,1)$.

3.12. A sortérbe eső megoldás meghatározása egyetlen egyenlet esetén egyszerű. Mivel a sorteret az $(1,4,8,12)$ vektor feszíti ki, ennek egy skalárszorosát keressük, mellyel vett skalárszorzata 225. Mivel $1^2+4^2+8^2+12^2=15^2=225$, ezért a sortérbe eső egyetlen megoldás $(x,y,z,w)=(1,4,8,12)$. A homogén egyenletrendszer összes megoldását meghatározva majd hozzáadva kapjuk, hogy

$$\begin{bmatrix}x\\y\\z\\w\end{bmatrix}=\begin{bmatrix}1\\4\\8\\12\end{bmatrix}+\begin{bmatrix}-4\\1\\0\\0\end{bmatrix}t+\begin{bmatrix}-8\\0\\1\\0\end{bmatrix}s+\begin{bmatrix}-12\\0\\0\\1\end{bmatrix}u$$

az összes megoldás.

3.13. A bővített mátrix és redukált lépcsős alakja:

$$\begin{bmatrix}1&1&1&1&3\\1&1&-1&-1&1\end{bmatrix}\Longrightarrow\begin{bmatrix}1&1&0&0&2\\0&0&1&1&1\end{bmatrix}$$

Az egyenletrendszer megoldása:

$$\begin{bmatrix}x\\y\\z\\w\end{bmatrix}=\begin{bmatrix}2-s\\s\\1-t\\t\end{bmatrix}=\begin{bmatrix}2\\0\\1\\0\end{bmatrix}+\begin{bmatrix}-1\\1\\0\\0\end{bmatrix}s+\begin{bmatrix}0\\0\\-1\\1\end{bmatrix}t.$$

Tehát a nullteret a $(-1,1,0,0)$ és a $(0,0,-1,1)$ vektork feszítik ki. A sortérbe eső megoldásvektor ezekre merőleges, tehát az eredeti egyenleteken kívül kielégíti a következő két egyenletet is:

$$\begin{aligned}-x+y\qquad&=0\\ -z+w&=0\end{aligned}$$

Ezek mátrixával kibővítve a redukált lépcsős alakot, majd azt redukált lépcsős alakra hozva kapjuk, hogy

$$\begin{bmatrix}1&1&0&0&2\\0&0&1&1&1\\-1&1&0&0&0\\0&0&-1&1&0\end{bmatrix}\Longrightarrow\begin{bmatrix}1&0&0&0&1\\0&1&0&0&1\\0&0&1&0&1/2\\0&0&0&1&1/2\end{bmatrix}$$

tehát a sortérbe eső megoldás $(1,1,1/2,1/2)$, az összes megoldás

$$\begin{bmatrix}x\\y\\z\\w\end{bmatrix}=\begin{bmatrix}1\\1\\1/2\\1/2\end{bmatrix}+\begin{bmatrix}-1\\1\\0\\0\end{bmatrix}s+\begin{bmatrix}0\\0\\-1\\1\end{bmatrix}t.$$

3.16. Az új egyenletrendszer sorterét az eredeti egyenletrendszer megoldásvektorai feszítik ki, ezért ennek nulltere megegyezik az eredeti sorterével, melyet a sorvektorok feszítenek ki. (Természetesen elég a sorvektorok közül a függetleneket kiválasztani. Esetünkben tehát a megadott egyenletrendszer nullterét kifeszítik az $(1,2,1,2,1)$ és az $(1,2,3,3,1)$ vektorok.)

II. rész

Mátrixok algebrája és geometriája

Eddig a mátrixokat csak egyszerű jelölésnek tekintettük, mely az egyenletrendszer együtthatóinak tárolására, és az egyenletrendszer megoldása közbeni számítások egyszerűsítésére való. E részt a számok közti műveletek számtáblázatokra való kiterjesztésével kezdjük, majd ezeket átültetjük mátrixokra, és megvizsgáljuk algebrai tulajdonságaikat. E műveletek segítségével újravizsgáljuk az egyenletrendszerek megoldhatóságának és a megoldások kiszámításának kérdését. A mátrixok „számtani” fejezetei után a „mértaniak” következnek: a determináns, mint a négyzetes mátrixhoz rendelt előjeles mérték, majd a mátrixleképezések geometriája lesz e rész tárgya.

4
Mátrixműveletek definíciói

A valós számok közti műveletek természetes módon kiterjeszthetők számtáblázatok közti műveletekre. Változók lineáris helyettesítésének kompozíciója is a táblázatok szorzásának műveletéhez vezet.

Táblázatok

Táblázattal nap, mint nap találkozunk. Fogalmát nem definiáljuk precízen. Számszerű adatok téglalap alakban elrendezett, áttekinthető összefoglalására való. Általában a sorok előtt és az oszlopok fölött van egy *fejléc*, melyben az adott sor, illetve oszlop adatait jellemző valamely információ áll. Például az oszlopok fejlécében gyakran szerepel az oszlop számadatainak közös mértékegysége. Használata kikerülhetetlen a gazdasági adatok kezelésében, így minden irodai szoftvercsomag tartalmaz táblázatkezelőt, de a mérnöki vagy természettudományos közleményekben is nélkülözhetetlen eszköz.

A mátrixra úgy is tekinthetünk, mint amelyet egy olyan absztrakció során kapunk a táblázatból, melyben azt megfosztjuk fejléceitől, az adatokból pedig csak a számokat őrizzük meg, azok jelentésétől, mértékegységétől eltekintünk.

Táblázatok összeadása 3 alma meg 2 alma az 5 alma. Az összeadás e művelete természetes módon kiterjeszthető egyszerre többféle gyümölccsel való számolásra. Ha az asztalon két gyümölcskosárban piros és zöld alma és szőlő van az alábbi táblázatok szerint, akkor összeöntésük után számukat az alábbi táblázat adja meg:

	alma (db)	*szőlő* (fürt)
piros	3	2
zöld	2	1

$+$

	alma (db)	*szőlő* (fürt)
piros	2	2
zöld	0	1

$=$

	alma (db)	*szőlő* (fürt)
piros	5	4
zöld	2	2

Azonos méretű, azonos fejlécű táblázatok összeadásának egy lehetséges módja az, ha az azonos pozícióiban lévő elemek összeadásával képezzük az összeget.

Táblázat szorzása számmal Az asztalon 2 alma van. Ha számukat megháromszorozzuk, összeszorzunk egy mértékegység nélküli számot (3) egy mértékegységgel rendelkezővel (2 darab), és az eredmény mértékegysége is ez. Ezt megtehetjük egy kosár egész tartalmával is:

$$3 \cdot \begin{array}{lcc} & \textit{alma} & \textit{szőlő} \\ & \text{(db)} & \text{(fürt)} \\ \hline \textit{piros} & 3 & 2 \\ \textit{zöld} & 2 & 1 \end{array} = \begin{array}{lcc} & \textit{alma} & \textit{szőlő} \\ & \text{(db)} & \text{(fürt)} \\ \hline \textit{piros} & 9 & 6 \\ \textit{zöld} & 6 & 3 \end{array}$$

Táblázatok szorzása Egy adag (e példában a továbbiakban mindig 10 dkg) alma energiatartalma 30 kcal. Mennyi 5 adag energiatartalma? A választ ismét szorzással kapjuk meg – most mindkét mennyiség mértékegységgel is rendelkezik:

$$5\,\text{adag} \cdot 30\,\frac{\text{kcal}}{\text{adag}} = 150\,\text{kcal}.$$

Több gyümölcsből (alma, banán, narancs) többféle (A, B, C) gyümölcssalátát készítünk, és a szénhidrát- és energiatartalmukat szeretnénk összehasonlítani. Két táblázatot készítünk, egyikbe a gyümölcssaláták összetételét, a másikba az összetevők szénhidrát- és energiatartalmát írjuk. Mindkét táblázatban a sorokba kerülnek azok a tételek, melyek összetételét/összetevőit részletezzük, az oszlopokba pedig az összetevők.

	Alma (adag)	*Banán* (adag)	*Narancs* (adag)
A	5	1	4
B	4	4	2
C	4	2	4

	Szénhidrát (g/adag)	*Energia* (kcal/adag)
Alma	7	30
Banán	24	105
Narancs	8	40

A következőképp tudjuk az A saláta energiatartalmát kiszámítani:

$$5\,\text{adag} \cdot 30\,\frac{\text{kcal}}{\text{adag}} + 1\,\text{adag} \cdot 105\,\frac{\text{kcal}}{\text{adag}} + 4\,\text{adag} \cdot 40\,\frac{\text{kcal}}{\text{adag}} = 415\,\text{kcal},$$

vagyis az első táblázat egy sorának és a második táblázat egy oszlopának kellett a skaláris szorzatát venni. Végezzük el e számításokat mindhárom gyümölcssaláta szénhidrát és energiatartalmára is, és az eredményt ismét egy olyan táblázatba tegyük, melynek soraiba a részletezendő tételek (A, B, C saláta), oszlopaiba a tartalmi összetevők (szénhidrát-, energiatartalom) kerüljenek.

	Szénhidrát (g)	*Energia* (kcal)
A	91	415
B	140	620
C	108	490

Az áttekinthetőség kedvéért a két összeszorzandó mátrixot és az eredményt úgy helyezzük el, hogy az elvégzett műveletek jobban követhetőek legyenek. Az A saláta energiatartalmát kiemeljük:

	Szénhidrát (g/adag)	***Energia*** **(kcal/adag)**
Alma	7	**30**
Banán	24	**105**
Narancs	8	**40**

	Alma (adag)	*Banán* (adag)	*Narancs* (adag)
A	**5**	**1**	**4**
B	4	4	2
C	4	2	4

	Szénhidrát (g)	*Energia* (kcal)
A	91	**415**
B	140	620
C	108	490

Érdemes megfigyelni, hogy ha csak az A és C gyümölcssalátákra vagyunk kiváncsiak, elég az első táblázat és a végeredmény második sorát elhagyni, hasonlóképp ha csak az energiatartalmat figyeljük, elég a második táblázat és a végeredmény második oszlopát megtartani. Az is látszik, hogy az első táblázat oszlopainak és a második táblázat sorainak száma megegyezik. Általában az igaz, hogy (a fejléceket nem számolva) egy $m \times n$-es táblázat csak olyan $p \times k$-as táblázattal szorozható össze, ahol $p = n$, és az eredmény $m \times k$-as lesz.

Lineáris helyettesítés A lineáris algebra alapvető fogalmai megfogalmazhatóak a lineáris helyettesítés nyelvén.

4.1. DEFINÍCIÓ (LINEÁRIS HELYETTESÍTÉS). *Lineáris helyettesítésről akkor beszélünk, ha változók egy halmazát más változók konstansszorosainak összegeként állítjuk elő, azaz változókat más változók lineáris kifejezéseivel helyettesítünk.*

Például

$$h = f + g, \qquad \begin{aligned} u &= x + y + z \\ v &= -9y + 2z \\ w &= 3z \end{aligned} \quad \text{és} \quad \begin{aligned} r &= 0.25p - 0.12q \\ s &= -0.45p + 2.11q \\ t &= -0.39q \end{aligned}$$

három lineáris helyettesítés. Tekintsük a következő két lineáris helyettesítést:

$$\begin{aligned} a &= 5x + y + 4z \\ b &= 4x + 4y + 2z \\ c &= 4x + 2y + 4z \end{aligned} \qquad \text{és} \qquad \begin{aligned} x &= 7s + 30k \\ y &= 24s + 105k \\ z &= 8s + 40k \end{aligned} \tag{4.1}$$

Egy néhány sor erejéig a lineáris helyettesítést is táblázatokkal írjuk le. Az előző két lineáris helyettesítés táblázatba foglalva:

	x	y	z
a	5	1	4
b	4	4	2
c	4	2	4

	s	k
x	7	30
y	24	105
z	8	40

(4.2)

4.2. PÉLDA (LINEÁRIS HELYETTESÍTÉSEK KOMPOZÍCIÓJA). *Tekintsük a (4.1)-ben megadott két lineáris helyettesítést! Az első majd a második helyettesítés egymás után való elvégzése – más szóval* kompozíciója *– milyen helyettesítéssel egyenértékű, és az lineáris-e?*

MEGOLDÁS. Az $a = 5x + y + 4z$ kifejezésben helyettesítsük x, y és z helyébe a második lineáris helyettesítés szerinti kifejezéseket, azaz

$$a = 5x + y + 4z = 5(7s + 30k) + (24s + 105k) + 4(8s + 40k) = 91s + 415k.$$

Emeljük ki pl. a k együtthatójának kiszámítását:

$$5 \cdot 30 + 1 \cdot 105 + 4 \cdot 40 = 415.$$

Mint látjuk ez az első helyettesítés táblázata első sorának és második táblázat második oszlopának skaláris szorzata. Hasonló módon b és c is kifejezhető s és k segítségével, így kapjuk, hogy a két lineáris helyettesítés egymásutánja ekvivalens az

$$\begin{aligned} a &= 91s + 415k \\ b &= 140s + 620k \\ c &= 108s + 490k \end{aligned}$$

helyettesítéssel. Ez lineáris, hisz a számolás közben lineáris kifejezéseket csak konstanssal szoroztunk, és ezeket adtuk össze. Ennek a helyettesítésnek a táblázata

	s	k
a	91	415
b	140	620
c	108	490

vagyis azt kaptuk, hogy a lineáris helyettesítések kompozíciójának táblázataik szorzata felel meg. □

	s	k
x	7	30
y	24	105
z	8	40

	x	y	z
a	5	1	4
b	4	4	2
c	4	2	4

	s	k
a	91	415
b	140	620
c	108	490

Feladatok

4.1. Anti, Bori, Cili almát, banánt és citromot vesz a piacon, a hipermarketben vagy a csarnokban. Ha csak az ár számít, melyikük hol vásároljon?

	alma (kg)	banán (kg)	citrom (kg)
Anti	2	2	1
Bori	3	2	0.5
Cili	2	1	1

	csarnok (Ft/kg)	hipermarket (Ft/kg)	piac (Ft/kg)
alma	180	100	130
banán	390	420	360
citrom	210	210	230

4.2. Egy $f(x, y)$ kifejezésben elvégezzük az

$$\begin{aligned} x &= 2a + b \\ y &= 3a + b \end{aligned}$$

helyettesítést, majd az így kapott $f(2a + b, 3a + b)$ kifejezésben az

$$\begin{aligned} a &= -3s + t \\ b &= 4s - t \end{aligned}$$

helyettesítést. Számítsuk ki a két helyettesítés kompozícióját a helyettesítések végrehajtásával, és a nekik megfelelő táblázatok szorzásával is, azaz írjuk fel azt a helyettesítést, mely e két helyettesítés kompozíciójával ekvivalens!

4.3. Tegyük fel, hogy egy kifejezésben elvégezzük a következő helyettesítést:

$$\begin{aligned} x &= 2a + b + 6c \\ y &= 4a + b + 7c \\ z &= 3a + b + 6c \end{aligned}$$

majd azt követően a következő helyettesítést:

$$\begin{aligned} a &= -s \phantom{{}-6t} + u \\ b &= -3s - 6t + 10u \\ c &= s + t - 2u \end{aligned}$$

Hogyan számíthatjuk ki a két helyettesítés kompozícióját? Írjuk fel azt a helyettesítést, mely e két helyettesítés kompozíciójával ekvivalens!

4.4. ét versengő kereskedelmi TV-csatorna valóságshow-műsora kezdetben fele-fele arányban vonzza a nézőket. Az első hét végére a tv1 nézőinek fele, míg a tv2 nézőnek negyede átpártol a másik csatornára.

1. Készítsük el az átpártolás 2×2-es táblázatát, és a
2. nézők megoszlásának 2×1-es vagy 1×2-es táblázatát!
3. Táblázatok szorzásának segítségével határozzuk meg, hogy mi a nézők megoszlása az első és a második hét végén, ha az átpártolók aránya az idővel nem változik.
4. Írjuk fel az átpártolók kéthetenkénti táblázatát, azaz azt, amelyből kiolvasható, hogy két hét elteltével az egyes csatornák nézőinek hányadrésze pártol át, és mennyi marad!

Elemenkénti mátrixműveletek

A táblázatok műveletei után a mátrixműveletek definíciói magától értetődőek.

Alapfogalmak, jelölések Az eddigiek alapján megfogalmazhatjuk: az m sorba és n oszlopba rendezett mn darab szám rendszerét $m \times n$-típusú *mátrixnak* nevezzük. A mátrixban szereplő számokat a mátrix elemeinek nevezzük. E megfogalmazás még mindig nem tekinthető precíz definíciónak, mert nincs tisztázva, mik is azok a számok, amik egy mátrix elemei lehetnek. Egyelőre tekintsük úgy, hogy mindig egy algebrailag jól definiálható számhalmaz elemeit írhatjuk egy mátrixba, ennek megfelelően fogunk egész elemű, racionális elemű vagy valós elemű, illetve az egész számok, a racionálisok, a valósok fölött definiált mátrixról beszélni. Később ezt is tágítani fogjuk, pl. vizsgálni fogunk függvényekből álló mátrixokat.

A matematikában elterjedt az a szokás, hogy a mátrixot jelölő nagy betűvel azonos kis betűk jelölik a mátrix elemeit, tehát $\mathbf{A}$ elemei a_{11}, $a_{12}\dots$ Az $m \times n$-típusú

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

mátrixra szokás még a tömörebb

$$\mathbf{A} = [a_{ij}]_{m\times n} \quad \text{vagy egyszerűen az} \quad \mathbf{A} = [a_{ij}]$$

jelölést használni. [1]

Mindig az első index jelöli a sor, a második az oszlop számát, tehát a_{23} a 2-dik sor 3-adik oszlop kereszteződésében álló elemet jelöli. Időnként, a félreérthetőség elkerülésére a_{ij} helyett $a_{i,j}$-t írunk (pl. $a_{n,n-1}$). Általában $\mathbf{a}_j$ jelöli az $\mathbf{A}$ mátrix j-edik oszlopvektorát, ha csak oszlopvektorokkal dolgozunk. Ha sor- és oszlopvektorok is együtt szerepelnek, az i-edik sorvektort $\mathbf{a}_{i*}$, a j-edik oszlopvektort $\mathbf{a}_{*j}$ jelöli összhangban az elemek indexelésével. Ehhez hasonló jelölést használnak a mátrix alapú nyelvek is (ld. a széljegyzetet). Az $\mathbf{A}$ mátrix i-edik sorára az $(\mathbf{A})_{i*}$, j-edik oszlopára az $(\mathbf{A})_{*j}$, elemére az $(\mathbf{A})_{ij}$ jelölés is használatos.

4.3. PÉLDA (MÁTRIXOK ÉS ELEMEIK). *Ha*

$$\mathbf{C} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 7 \end{bmatrix}, \textit{ akkor } c_{23} = 7,\ \mathbf{c}_2 = \mathbf{c}_{*2} = \begin{bmatrix} 2 \\ 5 \end{bmatrix},\ \mathbf{c}_{3*} = \begin{bmatrix} 6 & 8 & 9 \end{bmatrix}.$$

[1] A programnyelvekben – ellentétben a matematikával – a kisbetűvel/nagybetűvel való jelölésnek nincs a mátrixot az elemétől való megkülönböztető szerepe. A legtöbb magasszintű nyelvben az `A`-val jelölt mátrix (informatikai szóhasználattal *tömb*) i-edik sorának j-edik elemét `A[i,j]` vagy `A(i,j)` jelöli. Az alacsonyabb szintű C-típusú nyelvekben nincs 2-dimenziós tömb, a mátrixot egy olyan 1-dimenziós tömb reprezentálja, melynek minden eleme 1-dimenziós tömb, így `A[i]` az i-edik sort, `A[i][j]` az i-edik sor j-edik elemét jelöli. A mátrix alapú nyelvekben egy mátrix egy sorvektora vagy oszlopvektora könnyen kiemelhető, pl. az `A` mátrix 2. sorát az `A(2,:)`, 3. oszlopát a `A(:,3)` kóddal érhetjük el.

4.4. DEFINÍCIÓ (MÁTRIXOK EGYENLŐSÉGE). *Két mátrixot akkor tekintünk egyenlőnek, ha azonos típusúak, és az azonos indexű elemek egyenlőek.*

Például az

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & x \end{bmatrix}$$

egyenlőség pontosan akkor áll fönn, ha $x = 4$. Fontos felidézni, hogy minden vektornak megfeleltethetünk egy sor- vagy oszlopvektor alakba írt mátrixot, azok azonban nem egyenlők egymással. Például

$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \neq \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$

mert nem azonos típusúak, így el kell dönteni, hogy e két mátrix közül melyik reprezentálja a $(1, 2, 3)$ vektort. Mint említettük, a modern matematika legtöbb területén alapértelmezésben az oszlopvektorokat rendeljük a vektorokhoz. E könyvben mi is így teszünk, kivéve a kódelméleti alkalmazásokat, ahol a kódvektoroknak sorvektorokat feleltetünk meg.

Egy mátrix *négyzetes*, ha sorainak és oszlopainak száma megegyezik. Az $\mathbf{A}$ mátrix főátlójának elemei a_{11}, a_{22}, $a_{33}, \ldots$ Ez nem csak négyzetes mátrixra értelmezhető. Az olyan négyzetes mátrixot, melynek főátlón kívüli elemei mind nullák, *diagonális mátrixnak* nevezzük. Az ilyen mátrixok egyszerű megadására a diag függvényt használjuk, melynek argumentumába a főátló elemei vannak felsorolva. Például

$$\operatorname{diag}(1, 2, 3) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Gyakran fogunk azonos típusú mátrixokkal dolgozni. Rendszerint az is fontos, hogy a mátrixok elemei ugyanabból az algebrai struktúrából valók legyenek. Pl. vizsgálhatjuk a 3×2-es valós mátrixok vagy a 4×4-es egész elemű mátrixok halmazát.

4.5. DEFINÍCIÓ (ADOTT TÍPUSÚ MÁTRIXOK TERE). *Legyen S egy tetszőleges halmaz (általában S egy algebrai struktúra, pl. $S = \mathbb{R}$, $\mathbb{Q}$, $\mathbb{N}$, $\mathbb{Z}\ldots$). Az S elemeiből képzett összes $m \times n$-es mátrixok halmazát jelölje*

$$S^{m \times n}.$$

Azt mondjuk, hogy $S^{m \times n}$ az S fölötti $m \times n$ típusú mátrixok tere.

Például az $\left[\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right]$ mátrix eleme az $\mathbb{N}^{2\times 2}$, $\mathbb{Z}^{2\times 2}$, $\mathbb{Q}^{2\times 2}$, $\mathbb{R}^{2\times 2}$ terek mindegyikének.

```
OCTAVE a = [1 2 3
>           4 5 7]
a =
   1   2   3
   4   5   7

OCTAVE b = [1 2;3 4]
b =
   1   2
   3   4

OCTAVE diag([1,2,3])
ans =
   1   0   0
   0   2   0
   0   0   3

OCTAVE a(2,3)
ans =  7
OCTAVE a(2,:)
ans =
   4   5   7

OCTAVE a(:,3)
ans =
   3
   7

OCTAVE v = [1 2 3]
v =
   1   2   3

OCTAVE w = [1;2;3]
w =
   1
   2
   3

OCTAVE size(v)
ans =
   1   3

OCTAVE size(w)
ans =
   3   1
```

4.1. ábra: Mátrix megadása, elemeinek, sorainak és oszlopainak és azok számának lekérdezése mátrix alapú nyelvekben.

Elemenkénti mátrixműveletek Több olyan mátrixműveletet ismerünk, melyben a számok közt már értelmezett műveletet úgy általánosítjuk mátrixokra, hogy azt a mátrix vagy mátrixok minden elemén végrehajtjuk. Ilyen pl. a mátrixok összeadása és skalárral való szorzása.

A mátrixalapú nyelvekben mátrixok közötti elemenkénti művelet definiálható a műveleti jel elé tett ponttal. Így az `A` és `B` mátrixok elemenkénti szorzata az

```
A .* B
```

paranccsal kapható meg. Eszerint az `A .+ B` és `A + B` kódok az eredményt tekintve ekvivalensek.

4.6. DEFINÍCIÓ (MÁTRIXOK ÖSSZEGE, KÜLÖNBSÉGE). *Az $m \times n$-es típusú $\mathbf{A} = [a_{ij}]$ és $\mathbf{B} = [b_{ij}]$ mátrixok összegén azt az ugyancsak $m \times n$-es típusú, és $\mathbf{A} + \mathbf{B}$-vel jelölt mátrixot értjük, melynek i-edik sorában a j-edik elem $a_{ij} + b_{ij}$, ahol $i = 1, \ldots, m$, $j = 1, \ldots, n$. Képletben:*

$$\mathbf{A} + \mathbf{B} = [a_{ij}] + [b_{ij}] := [a_{ij} + b_{ij}].$$

Hasonlóan definiálható $\mathbf{A}$ és $\mathbf{B}$ különbsége is, azaz $\mathbf{A} - \mathbf{B} := [a_{ij} - b_{ij}]$.

4.7. PÉLDA (MÁTRIXOK ÖSSZEGE, KÜLÖNBSÉGE). *Ellenőrizzük az alábbi műveleteket!*

$$\begin{bmatrix} 0 & 2 & 4 \\ 1 & 3 & 5 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 4 & 5 \end{bmatrix}, \quad \begin{bmatrix} 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

4.8. DEFINÍCIÓ (ZÉRUSMÁTRIX). *A csupa nullából álló mátrixokat zérusmátrixoknak nevezzük. Az $m \times n$-es zérusmátrixot $\mathbf{O}_{m \times n}$, míg az $n \times n$-es négyzetes zérusmátrixot $\mathbf{O}_n$ jelöli.*

Tetszőleges $\mathbf{A}$ mátrixhoz egy azonos típusú zérusmátrixot adva $\mathbf{A}$-t kapunk, azaz $\mathbf{A} + \mathbf{O} = \mathbf{O} + \mathbf{A} = \mathbf{A}$.

4.9. DEFINÍCIÓ (MÁTRIX SZORZÁSA SKALÁRRAL). *Az $m \times n$-es típusú $\mathbf{A} = [a_{ij}]$ mátrix c számmal képzett szorzatán azt az ugyancsak $m \times n$-es típusú, és $c\mathbf{A}$-val jelölt mátrixot értjük, melyre*

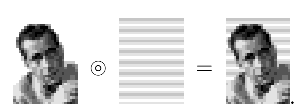
$$c\mathbf{A} = c[a_{ij}] := [ca_{ij}].$$

Az $\mathbf{A}$ mátrix *ellentettjének* azt a $-\mathbf{A}$-val jelölt mátrixot nevezzük, melyre $\mathbf{A} + (-\mathbf{A}) = \mathbf{O}$. Könnyen megmutatható, hogy ilyen mátrix csak egy van, nevezetesen $-\mathbf{A} = (-1)\mathbf{A}$.

Azonos méretű mátrixokon még számtalan elemenkénti művelet definiálható, ezek azonban az alkalmazásokban és a matematikán belül is lényegesen ritkábban fordulnak elő, ezért nem definiáljuk őket és nem vezetünk be rájuk jelölést. Érdekességként mutatunk egy példát egy ilyen műveletre.

A digitális képfeldolgozásban, ahol a képpontokra (pixelekre) bontott kép adatai mátrixokban vannak tárolva, sok művelet elemenként mátrixművelettel valósítható meg. A 4.2. ábra mátrixa az alatta lévő férfiarc 10 szürkeárnyalatos képe, melyen a háttér egy egyszerű elemenkénti művelettel megváltoztatható (részletek a 4.5. feladatban).

$$\begin{bmatrix}
9&9&9&9&9&9&9&9&9&9&9&9&9&9&9&9&9&9&9&9&9&9&9&9\\
9&9&9&9&9&9&9&9&9&9&9&9&9&9&9&9&9&9&9&9&9&9&9&9\\
9&9&9&9&9&6&3&4&5&3&2&2&5&6&9&9&9&9&9&9&9&9&9&9\\
9&9&9&9&4&1&0&0&2&2&0&0&0&4&9&9&9&9&9&9&9&9&9&9\\
9&9&9&4&1&0&0&1&1&1&1&0&1&1&3&9&9&9&9&9&9&9&9&9\\
9&9&4&1&0&1&2&1&0&0&0&1&2&1&0&4&9&9&9&9&9&9&9&9\\
9&9&2&2&2&2&1&1&0&1&1&3&3&2&1&0&5&9&9&9&9&9&9&9\\
9&5&1&2&3&3&2&1&2&2&4&4&3&2&1&0&3&9&9&9&9&9&9&9\\
9&3&1&3&7&8&7&6&6&5&6&4&3&2&2&0&2&9&9&9&9&9&9&9\\
9&4&1&5&8&8&8&7&6&5&5&4&3&2&2&1&0&6&9&9&9&9&9&9\\
9&5&2&5&8&8&7&6&5&5&5&4&4&3&2&2&1&3&7&9&9&9&9&9\\
9&9&2&3&6&7&7&6&5&5&4&4&4&0&0&0&3&3&5&9&9&9&9&9\\
9&9&5&1&4&6&7&7&7&5&4&3&1&1&1&1&2&3&4&9&9&9&9&9\\
9&9&9&2&2&6&7&5&1&0&2&3&1&2&3&2&2&3&3&9&9&9&9&9\\
9&9&9&7&1&6&6&1&0&1&1&6&3&2&2&3&3&2&3&9&9&9&9&9\\
9&9&9&7&3&6&4&2&3&3&5&8&4&3&3&3&2&2&4&9&9&9&9&9\\
9&9&9&9&5&6&6&7&5&3&7&8&6&3&3&4&3&2&3&7&7&7&7&9\\
9&9&9&9&7&4&7&8&7&5&7&8&8&4&3&4&3&2&3&7&3&3&4&5\\
9&9&9&9&9&7&7&6&6&7&7&6&6&2&2&3&3&3&3&4&3&3&3&5\\
9&9&9&9&9&9&8&6&6&7&5&7&7&2&3&3&2&3&3&3&3&3&3&4\\
9&9&9&9&9&5&3&5&6&7&5&7&7&4&2&1&3&3&2&3&3&3&3&3\\
9&9&9&8&6&1&3&7&6&6&5&4&3&2&3&3&3&3&3&3&3&4&4&3\\
9&9&9&7&5&3&2&7&8&6&5&4&6&4&2&3&3&2&3&3&3&4&4&3\\
9&9&7&6&5&4&2&7&8&7&6&4&4&4&6&4&3&1&4&4&6&6&6&6\\
9&8&7&6&4&4&3&7&8&8&6&6&3&5&3&1&0&1&6&8&7&6&6&5\\
8&8&7&5&4&3&4&7&8&8&6&2&2&0&1&2&3&8&7&7&7&7&7&6\\
8&8&7&5&4&3&4&6&8&8&6&5&2&0&1&1&6&8&8&7&7&6&5&3\\
8&8&8&5&4&4&4&4&7&8&5&5&3&3&1&3&8&8&8&6&6&5&6&4\\
8&8&8&7&5&4&4&4&6&8&7&5&2&2&5&8&8&8&8&7&6&3&3&2\\
6&8&7&5&6&5&4&3&4&7&7&7&4&4&7&8&8&8&8&7&6&7&6&4\\
4&7&6&4&4&6&5&4&4&5&6&7&7&7&7&8&8&8&8&8&6&4&2&3\\
6&4&7&5&3&4&6&4&4&4&5&7&8&6&8&8&8&8&7&4&2&1&2&3
\end{bmatrix}$$



4.2. ábra: Egy elemenkénti mátrixművelet a képfeldolgozásban

Mátrixok lineáris kombinációi A vektorokhoz hasonlóan, a skalárral való szorzás és az összeadás művelete lehetővé teszi, hogy mátrixokra is definiáljuk a *lineáris kombináció*, a *lineáris függetlenség* és a *kifeszített altér* fogalmát. A vektorokra korábban adott definíciók kis változtatással kimondhatók, ha a vektorok helyébe azonos típusú mátrixokat helyettesítünk, ezért ezt az olvasóra bízzuk, de gyakorlásként mutatunk két egyszerű példát.

4.10. PÉLDA (MÁTRIXOK LINEÁRIS KOMBINÁCIÓJA). *Számítsuk ki a*

$$2\begin{bmatrix}0 & 1\\ 2 & 1\\ 0 & -1\end{bmatrix} - 3\begin{bmatrix}1 & 0\\ -1 & -2\\ -1 & 0\end{bmatrix}.$$

lineáris kombinációt!

MEGOLDÁS. Először a skalárral való szorzásokat végezzük el:

$$2\begin{bmatrix}0 & 1\\ 2 & 1\\ 0 & -1\end{bmatrix} = \begin{bmatrix}0 & 2\\ 4 & 2\\ 0 & -2\end{bmatrix}, \quad -3\begin{bmatrix}1 & 0\\ -1 & -2\\ -1 & 0\end{bmatrix} = \begin{bmatrix}-3 & 0\\ 3 & 6\\ 3 & 0\end{bmatrix},$$

majd az összeadást:

$$\begin{bmatrix}0 & 2\\ 4 & 2\\ 0 & -2\end{bmatrix} + \begin{bmatrix}-3 & 0\\ 3 & 6\\ 3 & 0\end{bmatrix} = \begin{bmatrix}-3 & 2\\ 7 & 8\\ 3 & -2\end{bmatrix}.$$

A műveletek természetesen elemenként is elvégezhetők, pl. a második sor első eleme így is megkapható: $2 \cdot 2 - 3 \cdot (-1) = 7$. □

Látjuk, a mátrixok az összeadásra és a skalárral való szorzásra nézve a vektorokhoz hasonlóan viselkednek. Mondhatjuk, hogy az $\mathbb{R}^{m\times n}$-beli $m \times n$-es mátrixok e két műveletre nézve úgy viselkednek, mint $\mathbb{R}^{mn}$ vektorai (csak másként vannak elrendezve). Beszélhetünk ezért arról, hogy $\mathbb{R}^{m\times n}$ egy mn-dimenziós vektortér. Lásd erről pl. a 4.6. és a 4.7. feladatokat.

Feladatok

4.5. ELEMENKÉNTI MÁTRIXMŰVELET A DIGITÁLIS KÉPFELDOLGOZÁSBAN Az egészelemű $\mathbf{A}_{m\times n}$ mátrix reprezentáljon egy $m \times n$ képpontból álló szürkeárnyalatos képet. Minden mátrixelem egy képpont árnyalatát adja meg a $\{0, 1, \dots k\}$ tartományból, ahol 0 a fekete, k a fehér színnek felel meg. A háttér fehér, azaz képpontjaihoz a maximális k érték van rendelve és más fehér képpont a képen nincs. Legyen $\mathbf{B}_{m\times n}$ egy tetszőleges másik kép azonos módon reprezentált mátrixa. Konstruáljuk meg azt a $\circledcirc$ jellel jelölt műveletet, amellyel az elemenkénti

$$\mathbf{A} \circledcirc \mathbf{B} := [a_{ij} \circledcirc b_{ij}]$$

mátrixművelet az $\mathbf{A}$ kép hátterébe másolja a $\mathbf{B}$ képet. Képletben:

$$a_{ij} \circledcirc b_{ij} = \begin{cases} b_{ij}, & \text{ha } a_{ij} = k, \\ a_{ij}, & \text{egyébként.} \end{cases}$$

A megoldásban használhatjuk a $x \mapsto \lfloor x \rfloor$ függvényt, mely egy x számhoz annak alsó egész részét rendeli.

4.6. $\mathbb{R}^{m\times n}$ BÁZISA Adjuk meg az $\mathbb{R}^{m\times n}$ tér egy bázisát.

4.7. MÁTRIXOK ÁLTAL KIFESZÍTETT ALTÉR Jellemezzük az $\mathbb{R}^{2\times 2}$ térnek azt az alterét, melyet az alább megadott $\mathbf{A}$, $\mathbf{B}$ és $\mathbf{C}$ mátrixok feszítenek ki! Másként fogalmazva: milyen összefüggések állnak fönn azon 2×2-es valós mátrixok elemei között, melyek az alábbi mátrixok lineáris kombinációiként állnak elő?

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Mátrixszorzás

A táblázatok szorzásánál látott szabályt követjük a következő definícióban:

4.11. DEFINÍCIÓ (MÁTRIXOK SZORZÁSA). *Egy $m \times t$-s $\mathbf{A}$ és egy $t \times n$-es $\mathbf{B}$ mátrix szorzatán azt az $\mathbf{AB}$-vel jelölt $m \times n$-es $\mathbf{C}$ mátrixot értjük, amelynek i-edik sorában és j-edik oszlopában álló eleme*

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ik}b_{kj} + \ldots + a_{it}b_{tj}.$$

$$\begin{array}{cc} & \begin{bmatrix} & b_{1j} & \\ & b_{2j} & \\ & \vdots & \\ & b_{tj} & \end{bmatrix} \\ \begin{bmatrix} a_{i1} & a_{i2} & \ldots & a_{it} \end{bmatrix} & \begin{bmatrix} & & \\ & c_{ij} & \\ & & \end{bmatrix} \end{array}$$

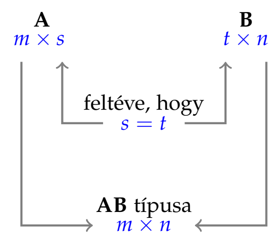
A definícióbeli összefüggés több módon is kifejezhető. Szummával fölírva:

$$c_{ij} = \sum_{k=1}^{t} a_{ik}b_{kj},$$

de mondhatjuk azt is, hogy c_{ij} az $\mathbf{A}$ mátrix i-edik sorának és a $\mathbf{B}$ mátrix j-edik oszlopának skaláris szorzata, azaz

$$c_{ij} = \mathbf{a}_{i*} \cdot \mathbf{b}_{*j}.$$

Fontos kiemelni, hogy egy $m \times s$-es $\mathbf{A}$ és egy $t \times n$-es $\mathbf{B}$ mátrix csak akkor szorozható össze, ha $s = t$, és ekkor a szorzat $m \times n$ típusú.



Nyilvánvaló, hogy a szorzás sorrendje fontos. Lehet, hogy az $\mathbf{AB}$ szorzás elvégezhető, de a $\mathbf{BA}$ nem, és lehet, hogy elvégezhető, de különböző eredményt kapunk.

4.12. PÉLDA (MÁTRIXOK SZORZÁSA). *Legyen*

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 2 & 3 \\ 2 & 1 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 2 & 1 \end{bmatrix}, \mathbf{C} = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix},$$

$$\mathbf{D} = \begin{bmatrix} 6 & 6 \\ -2 & -1 \end{bmatrix}, \mathbf{E} = \begin{bmatrix} -2 & -6 \\ 2 & 5 \end{bmatrix}.$$

Döntsük el, hogy fönnállnak-e az $\mathbf{AB} = \mathbf{BA}$, $\mathbf{BC} = \mathbf{CB}$, $\mathbf{CD} = \mathbf{DC}$ és $\mathbf{DE} = \mathbf{ED}$ egyenlőségek.

MEGOLDÁS. A méretek alapján a $\mathbf{BC}$ szorzat nincs értelmezve, a többi:

$$\mathbf{AB} = \begin{bmatrix} 3 & 2 & 1 \\ 9 & 8 & 7 \\ 3 & 4 & 5 \end{bmatrix}, \mathbf{BA} = \begin{bmatrix} 6 & 5 \\ 6 & 10 \end{bmatrix}, \mathbf{CB} = \begin{bmatrix} 3 & 2 & 1 \\ 3 & 4 & 5 \end{bmatrix},$$

$$\mathbf{CD} = \begin{bmatrix} -2 & -1 \\ 10 & 11 \end{bmatrix}, \mathbf{DC} = \begin{bmatrix} 12 & 12 \\ -2 & -3 \end{bmatrix}, \mathbf{DE} = \mathbf{ED} = \begin{bmatrix} 0 & -6 \\ 2 & 7 \end{bmatrix}.$$

Összefoglalva: $\mathbf{AB} \neq \mathbf{BA}$, mert különböző típusúak, $\mathbf{BC} \neq \mathbf{CB}$, mert az egyik oldal nincs értelmezve, $\mathbf{CD} \neq \mathbf{DC}$, bár mindkét oldal értelmezve van és azonos típusú. Az előzőekkel ellentétben viszont fennáll a $\mathbf{DE} = \mathbf{ED}$ egyenlőség. Azaz vannak felcserélhető mátrixok, de a mátrixszorzás nem felcserélhető művelet, tehát nem kommutatív! □

Mivel a mátrixszorzás nem felcserélhető, ha szükséges, az „**A**-t balról szorozzuk **B**-vel”, vagy az „**A**-t jobbról szorozzuk **B**-vel” kifejezésekkel teszünk különbséget a **BA** és az **AB** szorzatok közt.

Fontosak azok a mátrixszorzatok, amelyekben az egyik mátrixnak csak egy sora vagy oszlopa van, tehát sor- vagy oszlopvektor. Egy $1 \times m$-es mátrixot, azaz egy sorvektort jobbról, egy $n \times 1$-es mátrixot, azaz egy oszlopvektort balról lehet beszorozni egy $m \times n$-es mátrixszal. Például

$$\begin{bmatrix} 2 & 3 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 7 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

Skaláris szorzat és diadikus szorzat mátrixszorzatos alakja Két oszlopvektor nem szorozható össze, ha 1-nél nagyobb dimenziósak. Viszont az egyikük transzponálása után a szorzás elvégezhető.

Legyen $\mathbf{a}$ és $\mathbf{b}$ két $\mathbb{R}^n$-beli vektor. Az $\mathbf{a}^T\mathbf{b}$ szorzat a két vektor skaláris szorzatát adja, azaz

$$\mathbf{a}^T\mathbf{b} = \mathbf{a} \cdot \mathbf{b},$$

ugyanis

$$\mathbf{a}^T\mathbf{b} = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = a_1b_1 + a_2b_2 + \dots + a_nb_n = \mathbf{a} \cdot \mathbf{b}.$$

Ha a második vektort transzponáljuk, a két vektor lehet különböző dimenziós is.

4.13. DEFINÍCIÓ (DIADIKUS SZORZAT). *Legyen* $\mathbf{u} \in \mathbb{R}^m$, $\mathbf{v} \in \mathbb{R}^n$. *Az* $\mathbf{u}\mathbf{v}^T$ *szorzatot a két vektor* diadikus szorzatának, *röviden* diádnak *nevezzük. E szorzat egy* $m \times n$*-es mátrix:*

$$\mathbf{u}\mathbf{v}^T = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix} \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} = \begin{bmatrix} u_1v_1 & u_1v_2 & \dots & u_1v_n \\ u_2v_1 & u_2v_2 & \dots & u_2v_n \\ \vdots & \vdots & \ddots & \vdots \\ u_mv_1 & u_mv_2 & \dots & u_mv_n \end{bmatrix}.$$

Két vektor diadikus szorzatát $\mathbf{u} \otimes \mathbf{v}$ *jelöli.*

4.14. PÉLDA (SKALÁRIS ÉS DIADIKUS SZORZAT). *Legyen* $\mathbf{u} = (1, 0, 2)$, $\mathbf{v} = (3, 2, 1)$. *Írjuk fel mátrixszorzatos alakba skaláris és diadikus szorzatukat, és számítsuk ki!*

MEGOLDÁS.

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\mathbf{v} = \begin{bmatrix} 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = 5,$$

$$\mathbf{u} \otimes \mathbf{v} = \mathbf{u}\mathbf{v}^T = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \begin{bmatrix} 3 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 1 \\ 0 & 0 & 0 \\ 6 & 4 & 2 \end{bmatrix}$$ □

Lineáris egyenletrendszer mátrixszorzatos alakja A mátrixszorzást felhasználva a lineáris egyenletrendszerek egyszerű alakba írhatók.

4.1 (LINEÁRIS EGYENLETRENDSZER MÁTRIXSZORZATOS ALAKJA). *Ha* $\mathbf{A}$ *jelöli egy egyenletrendszer együtthatómátrixát, illetve* $\mathbf{b}$ *a konstans tagok és* $\mathbf{x}$ *az ismeretlenek oszlopvektorát, azaz*

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \text{és} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

akkor az

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \vdots \qquad \vdots \qquad\qquad \vdots \quad &\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned}$$

egyenletrendszer

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

alakba írható.

4.15. PÉLDA (EGYENLETRENDSZER MÁTRIXSZORZATOS ALAKJA). *Könnyen ellenőrizhető a mátrixszorzás elvégzésével, hogy a*

$$2x_1 + 3x_2 - x_3 = 5, \qquad \begin{aligned} ax \qquad &= u \\ by \quad &= v \\ cz &= w \end{aligned} \qquad \text{és} \qquad \begin{aligned} x + 2y &= 1 \\ y &= 1 \\ 0 &= 1 \end{aligned}$$

egyenletrendszerek mátrixszorzatos alakjai rendre:

$$\begin{bmatrix} 2 & 3 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 5, \quad \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} u \\ v \\ w \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

A szimultán egyenletrendszerek (ld. 104. oldal) ugyanígy fölírhatóak mátrixszorzatos alakba.

4.16. PÉLDA (SZIMULTÁN EGYENLETRENDSZER MÁTRIXSZORZATOS ALAKJA). *Írjuk az alábbi két egyenletrendszert egyetlen mátrixszorzatos alakba!*

$$\begin{aligned} 2x_{11} + 3x_{21} &= 7 \qquad & 2x_{12} + 3x_{22} &= 9 \\ 3x_{11} - 4x_{21} &= 2 \qquad & 3x_{12} - 4x_{22} &= 5 \end{aligned}$$

MEGOLDÁS. A két egyenletrendszer mátrixszorzatos alakjai külön-külön

$$\begin{bmatrix} 2 & 3 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix} = \begin{bmatrix} 7 \\ 2 \end{bmatrix}, \qquad \begin{bmatrix} 2 & 3 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} x_{12} \\ x_{22} \end{bmatrix} = \begin{bmatrix} 9 \\ 5 \end{bmatrix}.$$

Ezek egyetlen mátrixszorzattá olvaszthatók:

$$\begin{bmatrix} 2 & 3 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 7 & 9 \\ 2 & 5 \end{bmatrix}.$$

Általánosan a szimultán egyenletrendszerek $\mathbf{AX} = \mathbf{B}$ alakba írhatóak, ahol $\mathbf{X}$ az ismeretlenekből, $\mathbf{B}$ a jobb oldalakból alkotott mátrix. □

Lineáris helyettesítés mátrixszorzatos alakja Az egyenletrendszer mátrixszorzatos alakjához hasonló a lineáris helyettesítés mátrixszorzatos alakja.

4.2 (LINEÁRIS HELYETTESÍTÉS MÁTRIXSZORZATOS ALAKJA). *Az x_1, x_2,..., x_n változók lineáris kifejezéseinek az y_1, y_2,..., y_m változók helyébe való helyettesítését általánosan leíró*

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ &\vdots \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{aligned}$$

képletek mátrixszorzatos alakja

$$\mathbf{y} = \mathbf{A}\mathbf{x},$$

ahol

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}, \quad \textit{és} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Az $\mathbf{A}$ mátrixot nevezzük a *lineáris helyettesítés mátrixának*.

Például az

$$\begin{aligned} x &= 3a + 2b + 4c \\ y &= a - 3b + 2c \\ z &= 2a - b + 2c \end{aligned}$$

helyettesítés mátrixszorzatos alakja

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 & 2 & 4 \\ 1 & -3 & 2 \\ 2 & -1 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

Szorzás vektorral Egy $m \times n$-es mátrix vektorral kétféleképp szorozható: jobbról egy $n \times 1$-es oszlopvektorral, balról egy $1 \times m$-es sorvektorral.

Az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszer oszlopmodelljéből láttuk, hogy az egyenletrendszer bal oldala az $\mathbf{A}$ oszlopvektorainak az $\mathbf{x}$ koordinátáival vett lineáris kombinációja. Ez mátrixszorzással is könnyen ellenőrizhető. Hasonló állítás igaz a balról szorzásra is.

4.17. ÁLLÍTÁS (MÁTRIXSZORZÁS ÉS LINEÁRIS KOMBINÁCIÓ). *Legyen* $\mathbf{A}$ $m \times n$*-es mátrix,* $\mathbf{x}$ n*-dimenziós,* $\mathbf{y}$ m*-dimenziós vektor. Ekkor az* $\mathbf{Ax}$ *szorzat az* $\mathbf{A}$ *oszlopvektorainak*

$$\mathbf{a}_{*1}x_1 + \mathbf{a}_{*2}x_2 + \cdots + \mathbf{a}_{*n}x_n$$

lineáris kombinációját, míg az $\mathbf{y}^T\mathbf{A}$ *szorzat az* $\mathbf{A}$ *sorvektorainak*

$$\mathbf{a}_{1*}y_1 + \mathbf{a}_{2*}y_2 + \cdots + \mathbf{a}_{m*}y_m$$

lineáris kombinációját adja.

Szorzás standard egységvektorral Könnyen igazolhatók azok az összefüggések, melyeket a standard egységvektorokkal való szorzással kapunk. Jelölje $\mathbf{e}_i = (0, 0, \ldots, 1, \ldots, 0)$ azt a vektort, melynek i-edik koordinátája 1, a többi 0.

4.18. ÁLLÍTÁS (MÁTRIX ELEMEINEK, SOR- ÉS OSZLOPVEKTORAINAK ELŐÁLLÍTÁSA). *Legyen* $\mathbf{A}$ *egy* $m \times n$*-es mátrix,* $\mathbf{e}_i$ m*-dimenziós,* $\mathbf{e}_j$ n*-dimenziós standard egységvektor. Ekkor*

$$\mathbf{e}_i^T\mathbf{A} = \mathbf{a}_{i*} \quad \text{és} \quad \mathbf{A}\mathbf{e}_j = \mathbf{a}_{*j},$$

továbbá

$$\mathbf{e}_i^T(\mathbf{A}\mathbf{e}_j) = (\mathbf{e}_i^T\mathbf{A})\mathbf{e}_j = a_{ij}.$$

Az $\mathbf{e}_j\mathbf{e}_i^T$ diád egy olyan mátrix, melynek (i,j)-indexű eleme 1, az összes többi 0:

$$\mathbf{e}_j\mathbf{e}_i^T = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \begin{bmatrix} 0 & \dots & 1 & \dots & 0 \end{bmatrix} = \begin{bmatrix} 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{bmatrix}.$$

A báziscsere mátrixszorzatos alakja A 138. oldalon a vektor bázisra vonatkozó koordinátás alakjáról szóló paragrafusban láttuk, hogy e koordinátás alak hogy határozható meg. Kérdés, hogy egy *vektornak* egy altér két különböző bázisára vonatkozó koordinátás alakja közt mi a kapcsolat.

4.19. PÉLDA (ÁTTÉRÉS A STANDARD BÁZISRA). *Az $\mathbb{R}^3$ térnek $B = \{\,(1,1,1),(2,3,4),(6,6,7)\,\}$ egy bázisa. Az e bázisban megadott $[\mathbf{v}]_B$ vektornak írjuk fel a koordinátás alakját a standard bázisban egyetlen mátrixszorzással. Számítsuk ki konkrétan a $[\mathbf{v}]_B = (2,3,-1)$ vektor standard koordinátás alakját.*

MEGOLDÁS. $[\mathbf{v}]_B = (2,3,-1)$ azt jelenti, hogy

$$\mathbf{v} = 2\begin{bmatrix}1\\1\\1\end{bmatrix} + 3\begin{bmatrix}2\\3\\4\end{bmatrix} - \begin{bmatrix}6\\6\\7\end{bmatrix} = \begin{bmatrix}2\\5\\7\end{bmatrix},$$

ami mátrixszorzatos alakban

$$\mathbf{v} = \begin{bmatrix}1&2&6\\1&3&6\\1&4&7\end{bmatrix}\begin{bmatrix}2\\3\\-1\end{bmatrix} = \begin{bmatrix}2\\5\\7\end{bmatrix}.$$

Legyen $[\mathbf{v}]_B = (x,y,z)$. Ekkor

$$\mathbf{v} = x\begin{bmatrix}1\\1\\1\end{bmatrix} + y\begin{bmatrix}2\\3\\4\end{bmatrix} + z\begin{bmatrix}6\\6\\7\end{bmatrix} = \begin{bmatrix}1&2&6\\1&3&6\\1&4&7\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}.$$

Analóg képlet érvényes akkor is, ha két tetszőleges bázis közti áttérést számolunk. □

4.20. PÉLDA (BÁZISCSERE). *Legyen $B = \{\,(1,1,1),(2,3,4),(6,6,7)\,\}$ és $C = \{\,(1,2,3),(0,2,3),(3,5,8)\,\}$ az $\mathbb{R}^3$ két bázisa. Vektorok B bázisban megadott $[\mathbf{v}]_B$ koordinátás alakjához keressük a C bázisbeli $[\mathbf{v}]_C$ alakját. Hogyan számoljunk?*

MEGOLDÁS. Egyik lehetőség, hogy meghatározzuk $\mathbf{v}$ standard bázisbeli koordinátás alakját az előző feladat mintájára, majd az így kapott koordinátás alakból a C bázisra vonatkozó alakot a 138. oldalon leírtak szerint.

A másik lehetőség egyszerűbb, és hatékonyabb, ha több vektor koordinátás alakját kell meghatározni. A B bázis vektorainak a $[\mathbf{v}]_B$ koordinátáival vett lineáris kombinációja megadja a $\mathbf{v}$ vektort, így ha meghatározzuk a B bázis vektorainak C bázisbeli alakját, ezek lineáris kombinációja a $\mathbf{v}$ vektor C-beli alakját, azaz a $[\mathbf{v}]_C$ alakot adja.

A B bázis egy vektorának C-beli koordinátás alakjához egy egyenletrendszert kell megoldani, a B összes vektorának felírásához tehát egy szimultán egyenletrendszert:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 6 & 1 & 0 & 3 \\ 1 & 3 & 6 & 2 & 2 & 5 \\ 1 & 4 & 7 & 3 & 3 & 8 \end{array}\right] \Longrightarrow \left[\begin{array}{ccc|rrr} 1 & 0 & 0 & -1 & 2 & -7 \\ 0 & 1 & 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & 0 & -1 & 1 \end{array}\right]$$

Tehát a C bázis a B koordinátarendszerben:

$$C = \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}_B, \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}_B, \begin{bmatrix} -7 \\ 2 \\ 1 \end{bmatrix}_B \right\}.$$

Az ezekből képzett lineáris kombinációk mátrixszorzással is előállíthatók. Így egy $\mathbf{v}$ vektor B-beli koordinátás alakjának C bázisba való átírását a

$$[\mathbf{v}]_C = \begin{bmatrix} -1 & 2 & -7 \\ 1 & 2 & 2 \\ 0 & -1 & 1 \end{bmatrix} [\mathbf{v}]_B$$

képlet adja. Ha e mátrixot $\mathbf{A}_{C \leftarrow B}$ jelöli, akkor az előző képlet a

$$[\mathbf{v}]_C = \mathbf{A}_{C \leftarrow B} [\mathbf{v}]_B$$

alakot ölti. □

E példa a következő definícióhoz és állításhoz vezet:

4.21. DEFINÍCIÓ (ÁTTÉRÉS MÁTRIXA). *Legyen* $B = \{ \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n \}$ *és* $C = \{ \mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n \}$ *az* $\mathbf{R}^n$ *két bázisa. Írjuk fel* B *vektorait a* C *bázisban, és e vektorokból képezzünk mátrixot. Ezt nevezzük a* B *bázisról a* C*-re való áttérés mátrixának. Ha szükséges, a mátrix indexébe írjuk a két bázist* $C \leftarrow B$ *alakban. Tehát az áttérés mátrixa*

$$\mathbf{A}_{C \leftarrow B} = [\, [\mathbf{b}_1]_C \mid [\mathbf{b}_2]_C \mid \cdots \mid [\mathbf{b}_n]_C \,]$$

4.22. állítás (Koordináták változása a bázis cseréjénél). *Ha B és C az $\mathbf{R}^n$ két bázisa és $\mathbf{A}_{C \leftarrow B}$ az áttérés mátrixa, akkor bármely $\mathbf{v}$ vektor B-, illetve C-beli koordinátás alakja közt fennáll a*

$$[\mathbf{v}]_C = \mathbf{A}_{C \leftarrow B}[\mathbf{v}]_B$$

összefüggés.

*Bázisfelbontás** A 3.25. tétel második pontjának az előző feladatban is szemléltetett egyenes következménye, a következő állítás.

4.23. állítás (Bázisfelbontás). *Jelölje egy $m \times n$-es $\mathbf{A}$ mátrix redukált lépcsős alakjának nemzérus soraiból álló $r \times n$-es részmátrixát $\mathbf{R}$, az $\mathbf{R}$ főoszlopainak megfelelő $\mathbf{A}$-beli oszlopok alkotta $m \times r$-es részmátrixot $\mathbf{B}$, ahol $r = \mathrm{r}(\mathbf{A})$. Ekkor az $\mathbf{R}$ mátrix j-edik oszlopa megegyezik az $\mathbf{A}$ mátrix j-edik oszlopának a $\mathbf{B}$ oszlopai alkotta bázisban felírt koordinátás alakjával. Képletben ez azt jelenti, hogy*

$$\mathbf{A}_{*j} = \mathbf{B}\mathbf{R}_{*j}, \quad \textit{azaz} \quad \mathbf{A} = \mathbf{B}\mathbf{R}.$$

Egy mátrix fenti $\mathbf{A} = \mathbf{B}\mathbf{R}$ alakú felbontását *bázisfelbontásnak* nevezzük.

4.24. példa (Bázisfelbontás). *Határozzuk meg az alábbi mátrix bázisfelbontását, és magyarázzuk meg a két mátrix oszlopainak jelentését!*

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 8 & 6 & 2 \\ 1 & 2 & 7 & 0 & -11 \end{bmatrix}.$$

Megoldás. Az $\mathbf{A}$ mátrix redukált lépcsős alakja:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 8 & 6 & 2 \\ 1 & 2 & 7 & 0 & -11 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & 2 & 0 & 7 & 17 \\ 0 & 0 & 1 & -1 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

E mátrix első két sora alkotja az $\mathbf{R}$ mátrixot, az $\mathbf{A}$ mátrix első és harmadik oszlopa a $\mathbf{B}$ mátrixot, így a felbontás

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 8 & 6 & 2 \\ 1 & 2 & 7 & 0 & -11 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 2 & 8 \\ 1 & 7 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 & 7 & 17 \\ 0 & 0 & 1 & -1 & -4 \end{bmatrix} = \mathbf{B}\mathbf{R}.$$

Ellenőrizzük, hogy az $\mathbf{R}$ oszlopai az $\mathbf{A}$ oszlopvektorainak koordinátás alakjai a $\mathbf{B}$ oszlopai alkotta bázisban, azaz

$$[\mathbf{v}]_E = \mathbf{B}[\mathbf{v}]_B,$$

ahol az E indexszel a standard, B-vel a $\mathbf{B}$ mátrix oszlopai alkotta bázisbeli koordinátás alakját jelöltük ugyanannak a vektornak. Például

$$\begin{bmatrix}4\\6\\0\end{bmatrix} = \begin{bmatrix}1&3\\2&8\\1&7\end{bmatrix}\begin{bmatrix}7\\-1\end{bmatrix}, \quad \text{azaz} \quad [\mathbf{a}_4]_E = \begin{bmatrix}4\\6\\0\end{bmatrix}, [\mathbf{a}_4]_B = \begin{bmatrix}7\\-1\end{bmatrix}, \tag{4.3}$$

ahol $\mathbf{a}_4$ az $\mathbf{A}$ negyedik oszlopvektora. □

Egységmátrix, elemi mátrixok Egy adott $\mathbf{B}$ mátrixhoz találhatunk olyan $\mathbf{A}$-t, hogy az 1-gyel való szorzáshoz hasonlóan $\mathbf{AB} = \mathbf{B}$ legyen. Például

$$\mathbf{A} = \begin{bmatrix}3&-4\\-2&5\end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix}2&-2\\1&-1\end{bmatrix}$$

esetén

$$\begin{bmatrix}3&-4\\-2&5\end{bmatrix}\begin{bmatrix}2&-2\\1&-1\end{bmatrix} = \begin{bmatrix}2&-2\\1&-1\end{bmatrix}.$$

Az azonban már nem igaz, hogy $\mathbf{A}$-t bármely 2×2-es $\mathbf{B}$ mátrixszal szorozva $\mathbf{B}$ lesz az eredmény. Ilyen mátrix is létezik, némi próbálkozás után bárki rátalálhat.

4.25. DEFINÍCIÓ (EGYSÉGMÁTRIX). *Az $n \times n$-es*

$$\mathbf{I}_n := \operatorname{diag}(1, 1, \dots, 1) = \begin{bmatrix}1&0&\dots&0\\0&1&\dots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\dots&1\end{bmatrix}$$

mátrixot egységmátrixnak *nevezzük.*

Az egységmátrix elnevezés onnan származik, hogy tetszőleges $m \times n$-es $\mathbf{A}$ mátrixra igaz, hogy

$$\mathbf{I}_m \mathbf{A}_{m\times n} = \mathbf{A}_{m\times n}\mathbf{I}_n = \mathbf{A}_{m\times n},$$

azaz e mátrix hasonló tulajdonsággal rendelkezik, mint a számok közt az egy.

Az egységmátrixszal már találkoztunk: a Gauss–Jordan-módszernél egy n-ismeretlenes, n egyenletből álló egyértelműen megoldható egyenletrendszer együtthatómátrixa az elemi sorműveletek során egységmátrixszá transzformálódik!

Az egységmátrixon végrehajtott elemi sorműveletek olyan mátrixokat eredményeznek, melyek kapcsolatot létesítenek az elemi sorműveletek és a mátrixokkal való szorzás között. E mátrixoknak külön nevet adunk.

Az *egységmátrix* jelölésére használt $\mathbf{I}$ betű az angol *identity matrix* elnevezés első betűjéből származik. Az *azonosság* vagy *identitás* jelentésű *identity* szó az $\mathbf{IA} = \mathbf{A}$ összefüggésre utal (az $x \mapsto x$ függvényt ugyanezen okból hívjuk identikus függvénynek). Ráadásul az I betű hasonlít legjobban az 1-es számra.

4.26. definíció (Elemi mátrixok). *Az* $\mathbf{I}_n$ *egységmátrixon végrehajtott egyetlen elemi sorművelettel kapott mátrixot* elemi mátrixnak *nevezzük.*

4.27. példa (Elemi mátrixok). *Az alábbi mátrixok elemi mátrixok:*

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Ezt igazolja, hogy mindegyikük $\mathbf{I}_4$*-ből származik rendre a következő elemi sorműveletekkel:* $S_1 \leftrightarrow S_4$, $5S_2$, $S_1 + 2S_3$, $1S_1$. *Az utolsó mátrix az egységmátrix, amely maga is elemi mátrix, mert például egy sorának 1-gyel való szorzásával megkapható.*

4.28. példa (Mátrix balról szorzása elemi mátrixszal). *Az előbbi példa mátrixaival szorozzunk meg egy tetszőleges 4-soros mátrixot balról. Mit tapasztalunk?*

Megoldás. Legyen $\mathbf{A}$ egy 4- sorból, és az egyszerűség kedvéért csak 2 oszlopból álló mátrix. Végezzük el a fenti mátrixokkal való balról szorzást:

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} = \begin{bmatrix} a_{41} & a_{42} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{11} & a_{12} \end{bmatrix},$$

azaz a szorzás eredményeként fölcserélődött $\mathbf{A}$ első és negyedik sora. A következő szorzás eredménye

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ 5a_{21} & 5a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix},$$

azaz az $\mathbf{A}$ második sora be lett szorozva 5-tel. Végül

$$\begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} = \begin{bmatrix} a_{11} + 2a_{31} & a_{12} + 2a_{32} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix},$$

azaz az $\mathbf{A}$ a szorzás eredményeként az $\mathbf{A}$ első sorához hozzá lett adva harmadik sorának kétszerese. □

E példa eredménye kimondható tételként, melynek bizonyítása általánosan is úgy történik, mint az előző példában, ezért elhagyjuk:

4.29. TÉTEL (ELEMI SORMŰVELETEK MÁTRIXSZORZÁSSAL). *Legyen* $\mathbf{E}$ *az az elemi mátrix, melyet* $\mathbf{I}_m$*-ből egy elemi sorművelettel kapunk. Ha ugyanezt a sorműveletet egy tetszőleges* $m \times n$*-es* $\mathbf{A}$ *mátrixra alkalmazzuk, akkor eredményül az* $\mathbf{EA}$ *mátrixot kapjuk.*

Az elemi sorműveletek mátrixszorzással való elvégzésének nincs számítási praktikuma, annak célja az elemi sorműveletek – s ezzel az egyenletrendszerek megoldásának – algebraizálása.

Mátrixműveletek $\mathbb{Z}_m$*-ben** A mátrixműveletek

Blokkmátrixok

Műveletek blokkmátrixokkal A lineáris egyenletrendszerek bővített mátrixát tekinthetjük úgy, mint amelyet két mátrixból rakunk össze. De fordítva, mondhatjuk, hogy a bővített mátrixot két részmátrixra bontjuk.

Ha egy mátrixot a rajta végighaladó vízszintes és függőleges vonalakkal részmátrixokra bontunk, azt mondjuk, hogy e mátrix a részmátrixokból – más néven blokkokból – alkotott *blokkmátrix*, más néven *hipermátrix*.

Például egy 6-ismeretlenes, 5 egyenletből álló egyenletrendszer $\mathbf{B}$ bővített mátrixa lehet a következő, melynek blokkjait (részmátrixait) a mátrixokéhoz hasonló indexeléssel meg is jelöljük:

$$\left[\begin{array}{ccc|cc|c} 1 & 0 & 0 & 1 & 2 & 4 \\ 0 & 1 & 0 & 2 & 0 & 3 \\ 0 & 0 & 1 & 1 & 0 & 3 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \mathbf{B}_{13} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{B}_{23} \end{bmatrix}.$$

E felbontásban $\mathbf{B}_{11} = \mathbf{I}_3$, $\mathbf{B}_{21} = \mathbf{O}_{2\times 3}$, $\mathbf{B}_{22} = \mathbf{O}_2$.

Egy blokkmátrix sorait és oszlopait szokás a mátrix *blokksorainak* és *blokkoszlopainak* nevezni a közönséges soroktól és oszlopoktól való megkülönböztetés végett. Pl. a fenti blokkmátrix első blokksorának elemei $\mathbf{B}_{11}$, $\mathbf{B}_{12}$, $\mathbf{B}_{13}$.[2]

4.30. ÁLLÍTÁS (MŰVELETEK BLOKKMÁTRIXOKKAL). *Blokkmátrixok skalárral való szorzása és két azonos módon particionált blokkmátrix összeadása blokkonként is elvégezhető, azaz*

$$c[\mathbf{A}_{ij}] := [c\mathbf{A}_{ij}], \qquad [\mathbf{A}_{ij}] + [\mathbf{B}_{ij}] := [\mathbf{A}_{ij} + \mathbf{B}_{ij}].$$

Ha $\mathbf{A} = [\mathbf{A}_{ik}]_{m\times t}$, $\mathbf{B} = [\mathbf{B}_{kj}]_{t\times n}$ *két blokkmátrix, és minden* k*-ra az* $\mathbf{A}_{ik}$ *blokk oszlopainak száma megegyezik* $\mathbf{B}_{kj}$ *sorainak számával, akkor a*

[2] A BLOKKMÁTRIXOKRA a szakirodalomban a *hipermátrix* elnevezés is el van terjedve. Ekkor a blokksorokat hipersoroknak, a blokkoszlopokat hiperoszlopoknak nevezzük. Mi kerüljük e szóhasználatot a hipermátrix másik – többdimenziós tömb értelmű – jelentése miatt.

$\mathbf{C} = \mathbf{AB}$ *szorzat kiszámítható a szorzási szabály blokkokra való alkalmazásával is, azaz* $\mathbf{C}$ *olyan blokkmátrix, melynek i-edik blokksorában és j-edik blokkoszlopában álló blokk*

$$\mathbf{C}_{ij} = \sum_{k=1}^{t} \mathbf{A}_{ik}\mathbf{B}_{kj}.$$

4.31. példa (Műveletek blokkmátrixokkal). *Végezzük el az* $\mathbf{A} + 3\mathbf{C}$ *és az* $\mathbf{AB}$ *műveleteket közönséges mátrixműveletekkel és blokkmátrixként számolva is, ha*

$$\mathbf{A} = \left[\begin{array}{cc|c|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 2 \\ \hline 0 & 0 & 3 & 0 \end{array}\right], \quad \mathbf{B} = \left[\begin{array}{cc} 2 & 4 \\ 1 & 5 \\ \hline 2 & 2 \\ \hline 0 & 1 \end{array}\right], \quad \mathbf{C} = \left[\begin{array}{cc|c|c} 0 & 2 & 0 & 1 \\ 2 & 0 & 0 & 0 \\ \hline 1 & 1 & 1 & 1 \end{array}\right].$$

Megoldás. Számoljunk blokkmátrixként kezelve a mátrixokat:

$$\begin{aligned}
\mathbf{A} + 3\mathbf{C} &= \left[\begin{array}{cc|c|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 2 \\ \hline 0 & 0 & 3 & 0 \end{array}\right] + 3 \left[\begin{array}{cc|c|c} 0 & 2 & 0 & 1 \\ 2 & 0 & 0 & 0 \\ \hline 1 & 1 & 1 & 1 \end{array}\right] \\
&= \begin{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + 3 \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 2 \end{bmatrix} + 3 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 0 & 0 \end{bmatrix} + 3 \begin{bmatrix} 1 & 1 \end{bmatrix} & 3 + 3 \cdot 1 & 0 + 3 \cdot 1 \end{bmatrix} \\
&= \left[\begin{array}{cc|c|c} 1 & 6 & 1 & 3 \\ 6 & 1 & 1 & 2 \\ \hline 3 & 3 & 6 & 3 \end{array}\right]
\end{aligned}$$

Ellenőrizzük a számítást közönséges mátrixműveletekkel! Ezután tekintsük a blokkmátrixok szorzását!

$$\begin{aligned}
\mathbf{AB} &= \left[\begin{array}{cc|c|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 2 \\ \hline 0 & 0 & 3 & 0 \end{array}\right] \left[\begin{array}{cc} 2 & 4 \\ 1 & 5 \\ \hline 2 & 2 \\ \hline 0 & 1 \end{array}\right] \\
&= \begin{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 1 & 5 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 2 & 2 \end{bmatrix} + \begin{bmatrix} 0 \\ 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} \\ \begin{bmatrix} 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 1 & 5 \end{bmatrix} + 3 \begin{bmatrix} 2 & 2 \end{bmatrix} + 0 \begin{bmatrix} 0 & 1 \end{bmatrix} \end{bmatrix} \\
&= \begin{bmatrix} \begin{bmatrix} 2 & 4 \\ 1 & 5 \end{bmatrix} + \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \\ \begin{bmatrix} 0 & 0 \end{bmatrix} + \begin{bmatrix} 6 & 6 \end{bmatrix} + \begin{bmatrix} 0 & 0 \end{bmatrix} \end{bmatrix} = \left[\begin{array}{cc} 4 & 6 \\ 3 & 9 \\ \hline 6 & 6 \end{array}\right]
\end{aligned}$$

Ugyanezt az eredményt kapjuk, ha ellenőrzésül egyszerű mátrixszorzással is elvégezzük a műveletet! □

4.32. PÉLDA (2 × 2-ES BLOKKMÁTRIXOK). *Legyen* **A** *és* **B** *két* 2 × 2-*es blokkmátrix, azaz legyen*

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix}.$$

Írjuk fel szorzatukat a blokkok szorzatai segítségével.

MEGOLDÁS. Az **AB** szorzat felírható

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} \end{bmatrix}$$

alakban. A **BA** hasonlóképp írható fel! Ellenőrizzük, hogy a 4.30. állítás feltétele (minden k-ra az $\mathbf{A}_{ik}$ blokk oszlopainak száma megegyezik $\mathbf{B}_{kj}$ sorainak számával) valóban szükséges, de elégséges is. □

Vektorokra particionált mátrixok Fontosak azok a blokkmátrixok, amelyekben vagy oszlopvektor vagy sorvektor minden blokk.

Bontsuk fel az $\mathbf{A}_{m\times t}$ mátrixot sorvektoraira, és a $\mathbf{B}_{t\times n}$ mátrixot oszlopvektoraira, azaz tekintsük a következő két blokkmátrixot:

$$\mathbf{A} = \left[\begin{array}{c} \mathbf{a}_{1*} \\ \hline \mathbf{a}_{2*} \\ \hline \vdots \\ \hline \mathbf{a}_{m*} \end{array}\right], \qquad \mathbf{B} = \left[\begin{array}{c|c|c|c} \mathbf{b}_{*1} & \mathbf{b}_{*2} & \dots & \mathbf{b}_{*n} \end{array}\right].$$

Ekkor **AB** a mátrixszorzás definíciója alapján a következő alakba írható:

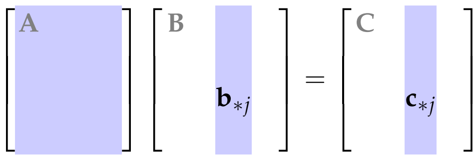
$$\mathbf{AB} = \begin{bmatrix} \mathbf{a}_{1*}\mathbf{b}_{*1} & \mathbf{a}_{1*}\mathbf{b}_{*2} & \dots & \mathbf{a}_{1*}\mathbf{b}_{*n} \\ \mathbf{a}_{2*}\mathbf{b}_{*1} & \mathbf{a}_{2*}\mathbf{b}_{*2} & \dots & \mathbf{a}_{2*}\mathbf{b}_{*n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_{m*}\mathbf{b}_{*1} & \mathbf{a}_{m*}\mathbf{b}_{*2} & \dots & \mathbf{a}_{m*}\mathbf{b}_{*n} \end{bmatrix},$$

ahol $\mathbf{a}_{i*}\mathbf{b}_{*j}$ az **A** mátrix i-edik sorának és a **B** mátrix j-edik oszlopának skaláris szorzata.

A blokkmátrixok szorzási szabálya akkor is alkalmazható, ha csak az egyik mátrixot particionáljuk (a másik mátrixot egyszerűen egyetlen blokkból álló mátrixnak tekintjük). Két eset lehetséges. Az első esetben a szorzatmátrixban „mátrixszor oszlopvektorok” szerepelnek:

$$\mathbf{C} = \mathbf{AB} = \mathbf{A}\left[\begin{array}{c|c|c|c} \mathbf{b}_{*1} & \mathbf{b}_{*2} & \dots & \mathbf{b}_{*n} \end{array}\right] = \left[\begin{array}{c|c|c|c} \mathbf{Ab}_{*1} & \mathbf{Ab}_{*2} & \dots & \mathbf{Ab}_{*n} \end{array}\right]$$

Itt tehát a **C** mátrix j-edik oszlopvektora az **A** mátrix és a **B** j-edik oszlopának szorzata, vagyis $\mathbf{c}_{*j} = \mathbf{A}\mathbf{b}_{*j}$. Sematikusan ábrázolva:

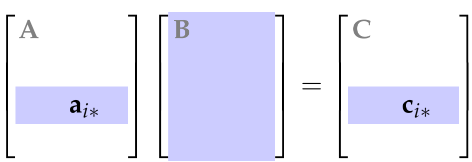


Ezzel az esettel már találkoztunk a szimultán egyenletrendszerek mátrixszorzatos alakjának fölírásánál (4.16. példa). Ha a fenti sematikus ábra egy szimultán egyenletrendszer mátrixszorzatos alakját reprezentálja, akkor a kékkel kiemelt rész a szimultán egyenletrendszer egyetlen egyenletrendszerének felel meg.

A másik esetben a szorzatmátrixban „sorvektorszor mátrixok" szerepelnek:

$$\mathbf{C} = \mathbf{AB} = \begin{bmatrix} \mathbf{a}_{1*} \\ \hline \mathbf{a}_{2*} \\ \hline \vdots \\ \hline \mathbf{a}_{m*} \end{bmatrix} \mathbf{B} = \begin{bmatrix} \mathbf{a}_{1*}\mathbf{B} \\ \hline \mathbf{a}_{2*}\mathbf{B} \\ \hline \vdots \\ \hline \mathbf{a}_{m*}\mathbf{B} \end{bmatrix}$$

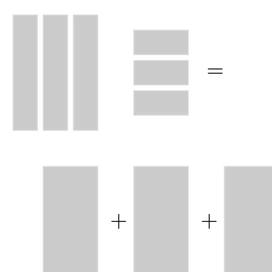
Azaz itt a $\mathbf{C} = \mathbf{AB}$ mátrix i-edik sora az $\mathbf{A}$ mátrix i-edik sorának $\mathbf{B}$-szerese. Másként írva $\mathbf{c}_{i*} = \mathbf{a}_{i*}\mathbf{B}$, sematikusan ábrázolva:



A mátrixszorzat egy felbontása megkapható a mátrixok másik partíciójából, azaz amikor $\mathbf{A}$-t oszlopokra, $\mathbf{B}$-t sorokra bontjuk, vagyis

$$\mathbf{A} = \left[\ \mathbf{a}_{*1} \mid \mathbf{a}_{*2} \mid \dots \mid \mathbf{a}_{*t}\ \right], \qquad \mathbf{B} = \begin{bmatrix} \mathbf{b}_{1*} \\ \hline \mathbf{b}_{2*} \\ \hline \dots \\ \hline \mathbf{b}_{t*} \end{bmatrix}.$$

Ekkor egyetlen blokksorból álló mátrixot szorzunk egy blokkoszlopból állóval, vagyis egy skaláris szorzatra emlékeztető összeget kapunk:



$$\mathbf{AB} = \mathbf{a}_{*1}\mathbf{b}_{1*} + \mathbf{a}_{*2}\mathbf{b}_{2*} + \dots + \mathbf{a}_{*t}\mathbf{b}_{t*}.$$

E felbontásban az $\mathbf{AB}$ mátrixot *diádok összegére* bontottuk! Mátrixok diádok összegére bontása olyan technika, amivel később még találkozunk.

4.33. PÉLDA (SZORZAT ELŐÁLLÍTÁSA DIÁDOK ÖSSZEGEKÉNT). *Bontsuk fel a*

$$\begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -2 & 0 \\ 1 & 1 \end{bmatrix}$$

szorzatot diádok összegére. Vajon felbontható-e a szorzat kevesebb diád összegére?

MEGOLDÁS.

$$\begin{bmatrix}0 & 1 & 2\\ 3 & 4 & 5\end{bmatrix}\begin{bmatrix}1 & 1\\ -2 & 0\\ 1 & 1\end{bmatrix} = \begin{bmatrix}0\\ 3\end{bmatrix}\begin{bmatrix}1 & 1\end{bmatrix} + \begin{bmatrix}1\\ 4\end{bmatrix}\begin{bmatrix}-2 & 0\end{bmatrix} + \begin{bmatrix}2\\ 5\end{bmatrix}\begin{bmatrix}1 & 1\end{bmatrix} =$$

$$\begin{bmatrix}0 & 0\\ 3 & 3\end{bmatrix} + \begin{bmatrix}-2 & 0\\ -8 & 0\end{bmatrix} + \begin{bmatrix}2 & 2\\ 5 & 5\end{bmatrix} = \begin{bmatrix}0 & 2\\ 0 & 8\end{bmatrix}.$$

Tehát a szorzatot három diád összegére bontottuk, azonban kevesebbre is lehet. Például

$$\begin{bmatrix}0 & 2\\ 0 & 8\end{bmatrix} = \begin{bmatrix}2\\ 8\end{bmatrix}\begin{bmatrix}0 & 1\end{bmatrix},$$

azaz a szorzat maga egy diád (egy diád összege). □

E felbontásnak fontos speciális esete az, amikor **A** egyetlen sorból, vagy **B** egyetlen oszlopból áll. Tekintsük az előző példában szereplő **A** mátrix első sorát! Ekkor a fenti felbontás a következő alakot ölti:

$$\begin{bmatrix}0 & 1 & 2\end{bmatrix}\begin{bmatrix}1 & 1\\ -2 & 0\\ 1 & 1\end{bmatrix} = 0\begin{bmatrix}1 & 1\end{bmatrix} + 1\begin{bmatrix}-2 & 0\end{bmatrix} + 2\begin{bmatrix}1 & 1\end{bmatrix} = \begin{bmatrix}0 & 2\end{bmatrix}.$$

Ez azt jelenti, hogy a szorzat, mely egy sorvektor, a **B** mátrix sorainak lineáris kombinációja. Hasonlóképp, ha **B** csak egyetlen oszlopból áll, a szorzat az **A** mátrix oszlopainak lineáris kombinációja. Például az előző példabeli **B** mátrix második oszlopát megtartva a következő felbontást kapjuk:

$$\begin{bmatrix}0 & 1 & 2\\ 3 & 4 & 5\end{bmatrix}\begin{bmatrix}1\\ 0\\ 1\end{bmatrix} = 1\begin{bmatrix}0\\ 3\end{bmatrix} + 0\begin{bmatrix}1\\ 4\end{bmatrix} + 1\begin{bmatrix}2\\ 5\end{bmatrix} = \begin{bmatrix}2\\ 8\end{bmatrix}.$$

Összefoglalva:

4.34. ÁLLÍTÁS (A SZORZAT OSZLOPAI ÉS SORAI). *Az* **AB** *mátrix minden oszlopa az* **A** *oszlopainak lineáris kombinációja, és minden sora a* **B** *sorainak lineáris kombinációja.*

BIZONYÍTÁS. Az **AB** mátrix j-edik oszlopa

$$(\mathbf{AB})_{*j} = \mathbf{Ab}_{*j} = \mathbf{a}_{*1}b_{1j} + \mathbf{a}_{*2}b_{2j} + \ldots + \mathbf{a}_{*t}b_{1t}$$

az i-edik sora pedig

$$(\mathbf{AB})_{i*} = \mathbf{a}_{i*}\mathbf{B} = a_{i1}\mathbf{b}_{1*} + a_{i2}\mathbf{b}_{2*} + \ldots + a_{it}\mathbf{b}_{t*},$$

ami bizonyítja az állítást. □

4.35. PÉLDA (NULLTÉR BÁZISA). *Határozzuk meg az alábbi mátrix nullterének bázisát!*

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 5 & 7 & 11 \\ 0 & 1 & 1 & 1 & -1 \end{bmatrix}$$

*Lineáris egyenletrendszer megoldásának blokkmátrix alakja**

MEGOLDÁS. A nulltér, azaz a mátrixhoz tartozó homogén lineáris egyenletrendszer megoldásainak tere könnyen leolvasható a redukált lépcsős alakból.

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 5 & 7 & 11 \\ 0 & 1 & 1 & 1 & -1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 1 & 2 & 7 \\ 0 & 1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

A szabad változókhoz rendelt paraméterek legyenek $x_3 = t_1$, $x_4 = t_2$, $x_5 = t_3$, amiből $x_1 = -t_1 - 2t_2 - 7t_3$ és $x_2 = -t_1 - t_2 + t_3$. Innen

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = t_1 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + t_3 \begin{bmatrix} -7 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & -2 & -7 \\ -1 & -1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix},$$

azaz a nullteret három vektor feszíti ki. Vegyük észre, hogy a redukált lépcsős alak blokkszerkezete

$$\begin{bmatrix} \mathbf{I}_2 & \mathbf{S} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}$$

és a megoldás

$$\mathbf{x} = \begin{bmatrix} -\mathbf{S} \\ \mathbf{I}_3 \end{bmatrix} \mathbf{t}$$

alakú, ahol $\mathbf{t}$ a paraméterek vektora. Itt csak annyi a kérdés, hogy a nullteret kifeszítő három vektor független-e. Ez jól látszik a három vektorból képzett mátrixon, mivel annak alsó blokkja $\mathbf{I}_3$, ami három független oszlopvektorból áll, és ezek akkor is függetlenek maradnak, ha eléjük további koordinátákat illesztünk. □

4.36. TÉTEL (A MEGOLDÁS FELÍRÁSA BLOKKMÁTRIXOKKAL). *Az egyszerűség kedvéért tegyük fel, hogy az r rangú $\mathbf{A}$ mátrix első r oszlopa lineárisan független – ez oszlopcserékkel mindig elérhető. Jelölje $\mathbf{B}_r$ az $\mathbf{A}$ első r oszlopából álló mátrixot, és legyen az $[\mathbf{A}|\mathbf{b}]$ bővített mátrix redukált lépcsős alakja*

$$\begin{bmatrix} \mathbf{I}_r & \mathbf{S} & \mathbf{d}_r \\ \mathbf{O} & \mathbf{O} & \mathbf{0} \end{bmatrix},$$

ahol $\mathbf{d}_r$ *egy* r*-dimenziós vektor. Ekkor az* $\mathbf{Ax} = \mathbf{b}$ *egyenletrendszer megoldható, és megoldása*

$$\mathbf{x} = \begin{bmatrix} \mathbf{d}_r \\ \mathbf{0}_s \end{bmatrix} + \begin{bmatrix} -\mathbf{S} \\ \mathbf{I}_s \end{bmatrix} \mathbf{t}_s,$$

ahol s *a szabad változók száma, azaz* $s = n - r$*, és* $\mathbf{t}_s$ *a szabad paraméterek vektora, ráadásul* $\mathbf{A} = \mathbf{B}_r[\mathbf{I}_r|\mathbf{S}]$ *és* $\mathbf{b} = \mathbf{B}_r\mathbf{d}_r$*.*

BIZONYÍTÁS. Mivel $[\mathbf{I}_r|\mathbf{S}]$ az $\mathbf{A}$ redukált lépcsős alakja, ezért ennek bármely oszlopa az $\mathbf{A}$ mátrix azonos sorszámú oszlopának koordinátás alakja az $\mathbf{B}_r$ oszlopvektoraiban, mint bázisban felírva. Ez épp azt jelenti, hogy $\mathbf{A} = \mathbf{B}_r[\mathbf{I}_r|\mathbf{S}]$. Ez az oszloptér bármely oszlopára, így $\mathbf{b}$-re is igaz, hisz $[\mathbf{A}|\mathbf{b}]$ redukált lépcsős alakja szerint az egyenletrendszer megoldható, így $\mathbf{b}$ eleme az oszloptérnek. Eszerint tehát $\mathbf{b} = \mathbf{B}_r\mathbf{d}_r$.

Az, hogy minden megoldás fölírható ilyen alakba, a Gauss–Jordan-módszerből következik. Meg kell még mutatni, hogy a tételben felírt $\mathbf{x}$ vektor valóban megoldás.

$$\begin{aligned}\mathbf{Ax} &= \mathbf{B}_r \begin{bmatrix} \mathbf{I}_r & \mathbf{S} \end{bmatrix} \left(\begin{bmatrix} \mathbf{d}_r \\ \mathbf{0}_s \end{bmatrix} + \begin{bmatrix} -\mathbf{S} \\ \mathbf{I}_s \end{bmatrix} \mathbf{t}_s \right) \\ &= \mathbf{B}_r \left(\mathbf{d}_r - \mathbf{S}\mathbf{t}_s + \mathbf{S}\mathbf{t}_s \right) = \mathbf{B}_r\mathbf{d}_r \\ &= \mathbf{b}.\end{aligned}$$

Ez bizonyítja az állítást. □

4.37. TÉTEL (A NULLTÉR BÁZISA). *Tegyük fel, hogy az* r *rangú* $\mathbf{A}$ *mátrix első* r *oszlopa lineárisan független – ez oszlopcserékkel mindig elérhető. Legyen az* $\mathbf{A}$ *mátrix redukált lépcsős alakja*

$$\begin{bmatrix} \mathbf{I}_r & \mathbf{S} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}.$$

Ekkor az $\mathbf{Ax} = \mathbf{0}$ *homogén lineáris egyenletrendszer megoldása*

$$\mathbf{x} = \begin{bmatrix} -\mathbf{S} \\ \mathbf{I}_s \end{bmatrix} \mathbf{t}_s,$$

ahol $s = n - r$ *a szabad változók száma, és a* $\left[\begin{smallmatrix} -\mathbf{S} \\ \mathbf{I}_s \end{smallmatrix}\right]$ *mátrix oszlopvektorai a nulltér bázisát alkotják.*

BIZONYÍTÁS. A bizonyítás közvetlen következménye a 4.36. tételnek, csak azt kell belátni, hogy $\left[\begin{smallmatrix} -\mathbf{S} \\ \mathbf{I}_s \end{smallmatrix}\right]$ oszlopvektorai a nulltér bázisát alkotják. Ez abból következik, hogy egyrészt kifeszítik a nullteret, másrészt lineárisan függetlenek, hisz az alsó blokkban lévő $\mathbf{I}_s$ mátrix oszlopai lineárisan függetlenek. □

Feladatok

Igaz – hamis

Döntsük el, igazak-e az alábbi állítások? Válaszunkat indokoljuk!

4.8. Ha az $\mathbf{AB}$ és a $\mathbf{BA}$ szorzat is értelmezve van, akkor mindkét szorzat négyzetes.

4.9. Ha az $(\mathbf{AB})\mathbf{C}$ szorzat elvégezhető, akkor biztosan elvégezhető az $\mathbf{A}(\mathbf{BC})$ szorzat is.

Mátrixműveletek

A következőkben legyen

$$\mathbf{A}=\begin{bmatrix}1&2\\4&2\end{bmatrix},\mathbf{B}=\begin{bmatrix}4&2\\4&5\end{bmatrix},\mathbf{C}=\begin{bmatrix}-3&2\\2&-1\end{bmatrix},\mathbf{D}=\begin{bmatrix}2&-1\\-1&1\end{bmatrix}.$$

Végezzük el az alábbi műveleteket!

4.10. $2\mathbf{A}-3\mathbf{B}^T$

4.11. $\mathbf{AB}-\mathbf{BA}+\mathbf{AC}-\mathbf{CA}$

4.12. $(\mathbf{CD}-\mathbf{DC})(\mathbf{ABC})$

4.13. $\mathbf{A}^2-\mathbf{C}^2$

4.14. $(\mathbf{C})_{2*}(\mathbf{D})_{*2}$

4.15. $(\mathbf{A})_{*1}(\mathbf{B})_{2*}$

4.16. A fenti jelölések mellett igazak-e a következő egyenlőségek?

$$(\mathbf{A}+\mathbf{B})^2=\mathbf{A}^2+2\mathbf{AB}+\mathbf{B}^2,\quad (\mathbf{A}+\mathbf{C})^2=\mathbf{A}^2+2\mathbf{AC}+\mathbf{C}^2.$$

4.17. A fenti jelölések mellett igazak-e a következő egyenlőségek?

$$(\mathbf{C}+\mathbf{D})(\mathbf{C}-\mathbf{D})=\mathbf{C}^2-\mathbf{D}^2,\quad (\mathbf{A}+\mathbf{D})(\mathbf{A}-\mathbf{D})=\mathbf{A}^2-\mathbf{D}^2.$$

Számítsuk ki az alábbi vektorok skaláris és diadikus szorzatát! Írjuk fel mindkét műveletet mátrixszorzatos alakban!

4.18. $\mathbf{a}=(1,2)$, $\mathbf{b}=(0,1)$

4.19. $\mathbf{u}=(1,2,0,1)$, $\mathbf{v}=(0,1,2,3)$

4.20. $\mathbf{a}=(1,2,0)$, $\mathbf{b}=(0,1,2,3)$

Mátrixszorzatos alakok

Írjuk fel az alábbi egyenletrendszerek mátrixszorzatos alakját!

4.21. $$\begin{aligned}x+y&=1\\x-z&=2\\z&=3\end{aligned}$$

4.22. $3x-2y+4z=5$

4.23. $$\begin{aligned}2x+z&=1\\x-y-w&=2\\y+z+w&=2\\0&=3\end{aligned}$$

Írjuk fel az alábbi lineáris helyettesítések mátrixszorzatos alakját!

4.24. $$\begin{aligned}u&=2x-4y\\v&=x+2y\end{aligned}$$

4.25. $$\begin{aligned}x&=3a-2b+c\\y&=2a-c\\z&=b+2c\end{aligned}$$

4.26. $$\begin{aligned}x&=3a+b\\y&=2a-b\\z&=b\end{aligned}$$

4.27. $$\begin{aligned}x&=3a-2b+c\\y&=2a-c\end{aligned}$$

Elemi mátrixok

Keressük meg azt az $\mathbf{E}$ mátrixot, mely megoldása az alábbi mátrixegyenletnek!

4.28. $\mathbf{E}\begin{bmatrix}a&b\\c&d\\e&f\end{bmatrix}=\begin{bmatrix}a&b\\e&f\\c&d\end{bmatrix}$

4.29. $\mathbf{E}\begin{bmatrix}a&b\\c&d\\e&f\end{bmatrix}=\begin{bmatrix}a&b\\3c&3d\\e&f\end{bmatrix}$

4.30. $\mathbf{E}\begin{bmatrix}a&b\\c&d\\e&f\end{bmatrix}=\begin{bmatrix}a&b\\c+2e&d+2f\\e&f\end{bmatrix}$

4.31. $\mathbf{E}\begin{bmatrix}a&b\\c&d\\e&f\end{bmatrix}=\begin{bmatrix}a-c&b-d\\c&d\\e&f\end{bmatrix}$

Elemi sorműveletekkel, mátrixszorzás nélkül határozzuk meg az alábbi mátrixszorzatok értékét!

4.32. $\begin{bmatrix}1&0&0\\2&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&2&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\-2&0&1\end{bmatrix}\begin{bmatrix}2&2&2\\3&3&3\\4&4&4\end{bmatrix}$

4.33. $\begin{bmatrix}1&0&0&0\\0&1&0&2\\0&0&1&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&-2&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&1&0&0\\3&0&1&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}a&b\\c&d\\e&f\\g&h\end{bmatrix}$

Blokkmátrixok

Számítsuk ki az alábbi feladatokban megadott mátrixszorzatokat a kijelölt blokkmátrixokat használva!

4.34. $\left[\begin{array}{cc|cc}1&0&1&1\\0&1&1&1\\\hline 0&0&2&2\\0&0&3&3\end{array}\right]\left[\begin{array}{cc|cc}2&3&1&1\\4&5&1&1\\\hline 0&0&1&0\\0&0&0&1\end{array}\right]$

4.35. $\left[\begin{array}{cc|cc} 2 & 3 & 1 & 1 \\ 4 & 5 & 1 & 1 \\ \hline 0 & 0 & 1 & 1 \end{array}\right] \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \\ \hline 0 & 0 & 2 \\ 0 & 0 & 3 \end{array}\right]$

4.36. $\left[\begin{array}{ccc|c|cc} 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ \hline 3 & 3 & 3 & 1 & 0 & 0 \end{array}\right] \left[\begin{array}{cc} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ \hline 3 & 3 \\ \hline 2 & 3 \\ 3 & 4 \end{array}\right]$

Vegyes feladatok

4.37. A *sudoku* egy olyan logikai játék, melyben egy olyan 9×9-es mátrixot kell megadni, melynek ismerjük néhány, de nem minden elemét. A feladat a nem ismert elemek meghatározása. A mátrix 9 darab 3×3-as blokkra van particionálva és eleget tesz annak a feltételnek, hogy minden sorában, minden oszlopában és minden blokkjában az 1-től 9-ig terjedő egészek mindegyike egyszer szerepel. Ez azt jelenti, hogy az egy sorban, egy oszlopban és egy blokkban lévő számok összege mindig 45. Fejezzük ki ezt mátrixműveletekkel, azaz írjunk fel a sudoku tábla $\mathbf{A}$ mátrixát is tartalmazó olyan mátrixegyenleteket, melyeket minden helyesen kitöltött sudoku tábla mátrixa kielégít!

4.38. Hány eleme van a $\mathbb{Z}_2^{2\times 2}$-nek, azaz a kételemű test fölötti 2×2-es mátrixok terének?

5
Mátrixműveletek tulajdonságai

Áttekintjük a mátrixműveletek legfontosabb algebrai tulajdonságait. Néhány dologban különböznek a számoknál megismert műveleti tulajdonságoktól.

Az alapműveletek algebrai tulajdonságai

Az összeadás és a skalárral való szorzás tulajdonságai Mivel a mátrixok összeadása és skalárral való szorzása elemenként végrehajtható műveletek, ezért műveleti tulajdonságaik természetes módon öröklik meg a számok műveleti tulajdonságait. A legfontosabb ilyen tulajdonságokat sorolja föl a következő tétel.

5.1. TÉTEL (ÖSSZEADÁS ÉS SKALÁRRAL SZORZÁS TULAJDONSÁGAI). *Legyen* $\mathbf{A}$, $\mathbf{B}$ *és* $\mathbf{C}$ *azonos típusú ($m \times n$-es) mátrix, c és d legyenek skalárok. Ekkor*

a)	$\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$	*felcserélhetőség, kommutativitás*
b)	$\mathbf{A}+(\mathbf{B}+\mathbf{C})=(\mathbf{A}+\mathbf{B})+\mathbf{C}$	*csoportosíthatóság, asszociativitás*
c)	$\mathbf{A}+\mathbf{O}_{m\times n}=\mathbf{A}$	*zérusmátrix (zéruselem) létezése*
d)	$\mathbf{A}+(-\mathbf{A})=\mathbf{O}_{m\times n}$	*ellentett (additív inverz) létezése*
e)	$c(d\mathbf{A})=(cd)\mathbf{A}$	*csoportosíthatóság*
f)	$0\mathbf{A}=\mathbf{O}_{m\times n}$	
g)	$1\mathbf{A}=\mathbf{A}$	
h)	$-1\mathbf{A}=-\mathbf{A}$	
i)	$(c+d)\mathbf{A}=c\mathbf{A}+d\mathbf{A}$	*disztributivitás*
j)	$c(\mathbf{A}+\mathbf{B})=c\mathbf{A}+c\mathbf{B}$	*disztributivitás*

► A csoportosíthatósági azonosságok következménye, hogy a zárójelek elhagyhatók, így $\mathbf{A}+(\mathbf{B}+\mathbf{C})$ helyett írható $\mathbf{A}+\mathbf{B}+\mathbf{C}$, $(cd)\mathbf{A}$ helyett $cd\mathbf{A}$.

► A két létezési állítás tartalma az, hogy *létezik* olyan mátrix, hogy *bármely* vele azonos típusú $\mathbf{A}$ mátrixhoz adva $\mathbf{A}$ lesz az eredmény, illetve minden $\mathbf{A}$ mátrixhoz *létezik* olyan mátrix, mellyel összeadva a zérusmátrixot kapjuk.

BIZONYÍTÁS. Mintaként bebizonyítjuk az *(a)* állítást.

$$\mathbf{A}+\mathbf{B}=[a_{ij}]+[b_{ij}]=[a_{ij}+b_{ij}]\stackrel{*}{=}[b_{ij}+a_{ij}]=[b_{ij}]+[a_{ij}]=\mathbf{B}+\mathbf{A}.$$

A *-gal jelzett egyenlőségnél használjuk a számok összeadásának kommutativitását.

A többi állítás hasonlóan bizonyítható. □

A szorzás tulajdonságai A számok szorzásának algebrai tulajdonságai már nem vihetők át olyan könnyen a mátrixműveletekre, mint az összeadáséi. Sőt, nem is teljesülnek mind, pl. a 4.12. példában láttuk, hogy a mátrixszorzás *nem kommutatív*. A következő példák óvatosságra intenek a mátrixkifejezésekkel való számolásokban.

5.2. PÉLDA (EGYSZERŰSÍTÉS MÁTRIXSZAL). *A valós számok közt igaz, hogy ha $a \neq 0$ és $ab = ac$, akkor a-val egyszerűsíthetünk, azaz akkor $b = c$. Mátrixoknál az* $\mathbf{AB} = \mathbf{AC}$ *egyszerűsíthetőséghez nem elég, hogy* $\mathbf{A}$ *ne legyen zérusmátrix! Például*

$$\begin{bmatrix}1&-2\\2&-4\\1&-2\end{bmatrix}\begin{bmatrix}1&-3\\2&-1\end{bmatrix}=\begin{bmatrix}1&-2\\2&-4\\1&-2\end{bmatrix}\begin{bmatrix}-1&3\\1&2\end{bmatrix},\ de\ \begin{bmatrix}1&-3\\2&-1\end{bmatrix}\neq\begin{bmatrix}-1&3\\1&2\end{bmatrix}.$$

5.3. PÉLDA (NULLOSZTÓ). *A valósok közt igaz, hogy ha $ab = 0$, akkor vagy $a = 0$, vagy $b = 0$. Mátrixok közt ilyen következtetés nem vonható le:*

$$\begin{bmatrix}1&2\\3&6\end{bmatrix}\begin{bmatrix}2&-2\\-1&1\end{bmatrix}=\begin{bmatrix}0&0\\0&0\end{bmatrix},$$

*(*Nullosztónak *nevezzük egy algebrai struktúra olyan nemzérus elemeit, melyek szorzata zérus.)*

► Megjegyezzük, hogy a $\mathbb{Z}_m$-ben való számolásnál is hasonlókat tapasztaltunk, ha m összetett. Például $\mathbb{Z}_6$-ban $2\cdot 3 = 0$, és bár $3\cdot 2 = 3\cdot 4$, $2 \neq 4$.

5.4. TÉTEL (MÁTRIXSZORZÁS ALGEBRAI TULAJDONSÁGAI). *Legyen* $\mathbf{A}$, $\mathbf{B}$ *és* $\mathbf{C}$ *olyan, hogy a kijelölt műveletek elvégezhetők legyenek, legyen továbbá c skalár. Ekkor*

a) $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$ *felcserélhetőség, asszociativitás*

b) $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$ *disztributivitás*

c) $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$ *disztributivitás*

d) $(c\mathbf{A})\mathbf{B} = c(\mathbf{AB}) = \mathbf{A}(c\mathbf{B})$

e) $\mathbf{A}_{m\times n}\mathbf{O}_{n\times t} = \mathbf{O}_{m\times t}$	*szorzás nullmátrixszal*
f) $\mathbf{I}_m\mathbf{A}_{m\times n} = \mathbf{A}_{m\times n}\mathbf{I}_n = \mathbf{A}_{m\times n}$	*szorzás egységmátrixszal*

BIZONYÍTÁS. A fenti tulajdonságok közül csak az elsőt és az utolsót bizonyítjuk, a többit feladatként tűzzük ki, mert vagy hasonlóan, vagy még egyszerűbben bizonyíthatóak.

a) Vizsgáljuk meg először, hogy három tetszőleges mátrix milyen feltételek mellett szorozható össze! Legyen $\mathbf{A}_{m\times s}$, $\mathbf{B}_{u\times v}$ és $\mathbf{C}_{t\times n}$ három tetszőleges mátrix. Az $(\mathbf{AB})\mathbf{C}$ szorzatban $\mathbf{AB}$ csak $s = u$ esetén végezhető el, és a szorzat típusa $m \times v$ lesz. Ez $\mathbf{C}$-vel csak akkor szorozható, ha $v = t$, és a szorzat $m \times n$-es. Tehát e szorzat csak akkor van értelmezve, ha $\mathbf{B}$ típusa $s \times t$. Hasonló érveléssel ugyanezt kapjuk az $\mathbf{A}(\mathbf{BC})$ szorzatról is.

Az indexek kezelésének könnyítésére elég lesz a bizonyítást sorvektor alakú $\mathbf{A}$ és oszlopvektor alakú $\mathbf{C}$ mátrixokra elvégezni, ugyanis az $(\mathbf{AB})\mathbf{C}$ szorzat i-edik sorában és j-edik oszlopában álló elem az $\mathbf{AB}$ i-edik sorának, azaz az $\mathbf{a}_{i*}\mathbf{B}$ sorvektornak és $\mathbf{C}$ j-edik oszlopának szorzata, azaz $(\mathbf{a}_{i*}\mathbf{B})\mathbf{c}_{*j}$. Hasonlóképp az $\mathbf{A}(\mathbf{BC})$ szorzat i-edik sorában és j-edik oszlopában álló elem $\mathbf{a}_{i*}(\mathbf{B}\mathbf{c}_{*j})$.

Legyen tehát

$$\mathbf{A} = \begin{bmatrix} a_1 & a_2 & \dots & a_m \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}.$$

Ekkor a szorzat 1×1-es. Először számoljuk ki az $\mathbf{AB}$ mátrixot, ami $1 \times m$-es: $\begin{bmatrix} \sum_{k=1}^m a_k b_{k1} & \sum_{k=1}^m a_k b_{k2} & \dots & \sum_{k=1}^m a_k b_{kn} \end{bmatrix}$. Innen számolva $(\mathbf{AB})\mathbf{C}$-t:

$$\begin{bmatrix} \sum_{k=1}^{m} a_k b_{k1} & \sum_{k=1}^{m} a_k b_{k2} & \dots & \sum_{k=1}^{m} a_k b_{kn} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \sum_{l=1}^{n} \sum_{k=1}^{m} a_k b_{kl} c_l.$$

Hasonlóan, először $\mathbf{BC}$-t fölírva, az $\mathbf{A}(\mathbf{BC})$ mátrixra kapjuk, hogy

$$\begin{bmatrix} a_1 & a_2 & \dots & a_m \end{bmatrix} \begin{bmatrix} \sum_{l=1}^n b_{1l} c_l \\ \sum_{k=1}^m b_{2l} c_l \\ \vdots \\ \sum_{k=1}^n b_{ml} c_l \end{bmatrix} = \sum_{k=1}^{m} a_k \left(\sum_{l=1}^{n} b_{kl} c_l \right) = \sum_{k=1}^{m} \sum_{l=1}^{n} a_k b_{kl} c_l.$$

Az utolsó lépésben a belső szumma minden tagját beszoroztuk a_k-val, a számok közti összeadás és szorzás közti disztributivitást használva. Vagyis mindkét oldalon olyan összeg áll, amely az összes $a_k b_{kl} c_l$ alakú szorzat összege, csak a tagok csoportosítása más. □

► Az asszociativitás következménye, hogy a többtényezős mátrixszorzatokat nem kell zárójelezni, hisz bármelyik zárójelezés ugyanazt az eredményt adja. Például

$$\mathbf{ABC} = (\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$$

Az állítás igaz többtényezős szorzatokra is, vagyis az $\mathbf{A}_1\mathbf{A}_2\ldots\mathbf{A}_k$ szorzat független a végrehajtás sorrendjétől, de a tényezők sorrendje nem változtatható!

► Indukcióval bizonyítható, hogy a *(g)*-beli összefüggés többtényezős szorzatokra is fönnáll, azaz

$$(\mathbf{A}_1\mathbf{A}_2\ldots\mathbf{A}_k)^T = \mathbf{A}_k^T\ldots\mathbf{A}_2^T\mathbf{A}_1^T.$$

► Megjegyezzük, hogy az asszociativitás imént leírt bizonyítása hasonlóan mondható el, ha az $\mathbf{A} = [a_{ik}]$ mátrix nem csak 1 sorból, és a $\mathbf{C} = [c_{lj}]$ mátrix nem csak egy oszlopból áll: ekkor a $\mathbf{D} = \mathbf{ABC}$ szorzat i-edik sorának j-edik elemére azt kapjuk, hogy az az összes $a_{ik}b_{kl}c_{lj}$ alakú szorzatok összege, azaz

$$d_{ij} = \sum_{k=1}^{m}\sum_{l=1}^{n} a_{ik}b_{kl}c_{lj}.$$

Ez, és az ehhez hasonló számtalan hasonló kifejezés vezette Einsteint arra a felismerésre, hogy az indexelt változók szorzatainak összegében a szumma jelek feleslegesek, hisz azokra az indexekre kell összegezni, amelyek legalább kétszer szerepelnek, míg az egyszer szereplőkre nem. Tehát az előző kettős szumma helyett írhatnánk azt is, hogy

$$d_{ij} = a_{ik}b_{kl}c_{lj},$$

hisz a jobb oldalon i és j csak egyszer szerepel, így k-ra és l-re kell összegezni, azt pedig tudjuk, hogy $k = 1,\ldots,m$ és $l = 1,\ldots,n$. Ezt a jelölésbeli egyszerűsítést *Einstein-konvenciónak* nevezik. Einstein ezt a relativitás általános elméletéről írt híres dolgozatában használta először 1916-ban. A konvenció használata főként a lineáris algebra fizikai alkalmazásaiban terjedt el, mi e könyvben nem fogjuk használni.

Mátrix hatványozása Csak a négyzetes mátrixok szorozhatók meg önmagukkal, hisz ha egy $m \times n$-es mátrix megszorozható egy $m \times n$-essel, akkor $m = n$. Ezt figyelembe véve természetes módon definiálható négyzetes mátrixok pozitív egész kitevős hatványa:

$$\mathbf{A}^k = \underbrace{\mathbf{AA}\ldots\mathbf{A}}_{k \text{ tényező}}$$

Kicsit elegánsabban – rekurzióval – is definiálhatjuk e fogalmat: $\mathbf{A}^1 := \mathbf{A}$ és $\mathbf{A}^{k+1} := \mathbf{A}^k\mathbf{A}$.

Mivel a mátrixszorzás asszociatív, mindegy, hogy milyen sorrendben végezzük el a hatványozást. Ezzel igazolható a következő két összefüggés is:

5.5. ÁLLÍTÁS (HATVÁNYOZÁS AZONOSSÁGAI). *Legyen* $\mathbf{A}$ *egy négyzetes mátrix! Ekkor*
a) $\mathbf{A}^k\mathbf{A}^m = \mathbf{A}^{k+m}$,
b) $(\mathbf{A}^k)^m = \mathbf{A}^{km}$.

Ha ki akarjuk terjeszteni a hatványozást 0 kitevőre is, kövessük a precedencia-elvet[1], azaz olyan értelmet adjunk $\mathbf{A}^0$-nak, hogy a fenti összefüggések érvényben maradjanak. Például tekintsük az *a)* azonosságot $m = 0$ esetén:

$$\mathbf{A}^k\mathbf{A}^0 = \mathbf{A}^{k+0} = \mathbf{A}^k.$$

Ez minden $\mathbf{A}$ mátrix esetén csak az egységmátrixra igaz, tehát

$$\mathbf{A}^0 = \mathbf{I}_n,$$

ahol n a négyzetes $\mathbf{A}$ mérete.

[1] A latin eredetű *precedencia* szó *előzményt* jelent (lásd még precedens). A *precedencia elv* a matematikában fogalmak jelentésének olyan kiterjesztését jelenti, melynek során a korábban megismert tulajdonságok, összefüggések érvényben maradnak.

► A valós számoknál tanult, különböző alapú hatványokra érvényes azonosság itt a kommutativitás hiánya miatt nem érvényes, azaz általában $(\mathbf{AB})^k \neq \mathbf{A}^k\mathbf{B}^k$.

5.6. PÉLDA (MÁTRIX HATVÁNYOZÁSA). *Számítsuk ki az*

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \textit{és a} \quad \mathbf{B} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

mátrixok k-adik hatványait!

MEGOLDÁS. Számoljuk ki $\mathbf{A}$ hatványait!

$$\mathbf{A}^2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

azaz $\mathbf{A}^2 = \mathbf{I}_2$, ebből pedig látjuk, hogy $\mathbf{A}^3 = \mathbf{I}_2\mathbf{A} = \mathbf{A}$, $\mathbf{A}^4 = \mathbf{A}^3\mathbf{A} = \mathbf{A}\mathbf{A} = \mathbf{I}_2, \dots$ Tehát általában $\mathbf{A}^{2k} = \mathbf{I}_2$ és $\mathbf{A}^{2k+1} = \mathbf{A}$.

A másik feladatot a hatványozás rekurzív definícióját használva indukcióval kényelmesen meg tudjuk oldani. Először számoljuk ki $\mathbf{A}$ néhány hatványát:

$$\mathbf{A}^2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \mathbf{A}^3 = \mathbf{A}^2\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}.$$

Ebből azt sejtjük, hogy $\mathbf{A}^k = \left[\begin{smallmatrix} 1 & k \\ 0 & 1 \end{smallmatrix}\right]$. Ha be tudjuk látni ennek az összefüggésnek az öröklődését k-ról $k+1$-re, akkor kész vagyunk. Más szóval meg kell mutatnunk, hogy ha $\mathbf{A}^k = \left[\begin{smallmatrix} 1 & k \\ 0 & 1 \end{smallmatrix}\right]$, akkor $\mathbf{A}^{k+1} = \left[\begin{smallmatrix} 1 & k+1 \\ 0 & 1 \end{smallmatrix}\right]$. Ezt a következő szorzás elvégzése igazolja:

$$\mathbf{A}^{k+1} = \mathbf{A}^k \mathbf{A} = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & k+1 \\ 0 & 1 \end{bmatrix} \qquad \square$$

Miután mátrixok lineáris kombinációja és négyzetes mátrixok egész kitevős hatványa értelmezve van, ezért négyzetes mátrixokra is definiálhatjuk skalár együtthatós polinom helyettesítési értékét. Legyen

$$p(x) = a_k x^k + a_{k-1} x^{k-1} + \ldots + a_2 x^2 + a_1 x + a_0$$

egy skalár együtthatós polinom. A p polinom $\mathbf{X} \in \mathbb{R}^{n \times n}$ helyen vett helyettesítési értékén a

$$p(\mathbf{X}) = a_k \mathbf{X}^k + \ldots + a_2 \mathbf{X}^2 + a_1 \mathbf{X} + a_0 \mathbf{I}_n$$

mátrixot értjük.

5.7. PÉLDA (POLINOM HELYETTESÍTÉSI ÉRTÉKE). *Legyen*

$$\mathbf{C} = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 3 & -4 \\ 3 & 4 & -6 \end{bmatrix}.$$

Mutassuk meg, hogy $p(\mathbf{C}) = \mathbf{O}$, *ha* $p(x) = x^3 + 2x^2 - 1$.

MEGOLDÁS. A $p(\mathbf{C}) = \mathbf{C}^3 + 2\mathbf{C}^2 - \mathbf{I}$ műveleteit elvégezve kapjuk, hogy

$$\begin{aligned} p(\mathbf{C}) &= \mathbf{C}^3 + 2\mathbf{C}^2 - \mathbf{I} \\ &= \begin{bmatrix} 9 & 8 & -14 \\ 8 & 7 & -12 \\ 14 & 12 & -21 \end{bmatrix} + 2 \begin{bmatrix} -4 & -4 & 7 \\ -4 & -3 & 6 \\ -7 & -6 & 11 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \qquad \square \end{aligned}$$

A transzponálás tulajdonságai A következő tétel a transzponálás és a többi művelet kapcsolatáról szól:

5.8. TÉTEL (TRANSZPONÁLÁS TULAJDONSÁGAI). $\mathbf{A}$ *és* $\mathbf{B}$ *legyenek azonos típusú mátrixok, c tetszőleges skalár. Ekkor*

a) $(\mathbf{A}^T)^T = \mathbf{A}$,

b) $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$,

c) $(c\mathbf{A})^T = c\mathbf{A}^T$,

d) $(\mathbf{A}\mathbf{B})^T = \mathbf{B}^T \mathbf{A}^T$.

BIZONYÍTÁS. Az első három összefüggés magától értetődő, csak az utolsót bizonyítjuk.

Először megmutatjuk, hogy ha $(\mathbf{AB})^T$ elvégezhető, akkor $\mathbf{B}^T\mathbf{A}^T$ is. Az $m \times t$ típusú $\mathbf{A}$ és a $t \times n$ típusú $\mathbf{B}$ szorzata $m \times n$-es, transzponáltja $n \times m$-es, így az $n \times t$ típusú $\mathbf{B}^T$ és a $t \times m$-es $\mathbf{A}^T$ összeszorozhatók, szorzatuk $n \times m$-es, így a tételbeli egyenlőség két oldalának típusa azonos.

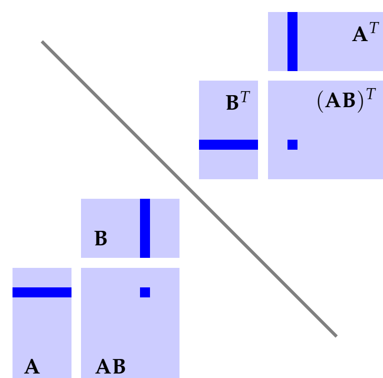
A tétel azon alapul, hogy két tetszőleges $\mathbf{u}$, $\mathbf{v}$ vektorra $\mathbf{u}^T\mathbf{v} = \mathbf{v}^T\mathbf{u}$. Ezt az összefüggést a *-gal jelölt egyenlőségnél fogjuk használni. Az $(\mathbf{AB})^T$ i-edik sorának j-edik eleme

$$\left((\mathbf{AB})^T\right)_{ij} = (\mathbf{AB})_{ji} = (\mathbf{A})_{j*}\,(\mathbf{B})_{*i}\,.$$

A $\mathbf{B}^T\mathbf{A}^T$ i-edik sorának j-edik eleme

$$\left(\mathbf{B}^T\mathbf{A}^T\right)_{ij} = \left(\mathbf{B}^T\right)_{i*}\left(\mathbf{A}^T\right)_{*j} \stackrel{*}{=} (\mathbf{A})_{j*}\,(\mathbf{B})_{*i}\,.$$

Tehát $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$. □



5.1. ábra: $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$ szemléletes bizonyítása

▶ A tétel *(b)* pontjának indukcióval könnyen bizonyítható következménye, hogy többtagú összeg transzponáltja megegyezik a transzponáltak összegével. A *(c)* pontot is figyelembe véve kapjuk, hogy mátrixok lineáris kombinációjának transzponáltja megegyezik a mátrixok transzponáltjainak azonos lineáris kombinációjával, azaz

$$(c_1\mathbf{A}_1 + c_2\mathbf{A}_2 + \ldots + c_k\mathbf{A}_k)^T = c_1\mathbf{A}_1^T + c_2\mathbf{A}_2^T + \ldots + c_k\mathbf{A}_k^T.$$

▶ A tétel *(d)* pontjára „szemléletes igazolás” is adható, ami leolvasható az 5.1. ábráról.

Mátrix inverze

Lehet-e mátrixszal osztani, és ha igen, meg tudjuk-e vele oldani az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszert úgy, ahogy az $ax = b$ egyenletet megoldjuk az a-val való osztással?

Az inverz Korábbi tanulmányainkban megtanultuk, hogy az összeadás és a szorzás invertálható műveletek, inverzeik a kivonás, illetve az osztás. De mit is jelent ez pontosan, és vajon a mátrixműveletek invertálhatóak-e?

Egy H halmazon értelmezett kétváltozós (más szóval bináris) *műveleten* olyan függvényt értünk, mely H-beli elempárokhoz H-beli elemet rendel. Például a valós számok összeadása esetén e függvény valós számpárhoz valós számot rendel, mondjuk a $(1.2, 0.1)$ számpárhoz a 1.3-at. E függvényt a $+$ jellel jelöljük, de a függvényeknél szokásos prefix „$+(a, b)$” jelölés helyett műveleteknél az ún. infix jelölést használjuk, azaz $a + b$-t írunk (lásd erről még a széljegyzetet).

Azon, hogy az összadás művelete invertálható, azt értjük, hogy bármely a és b valós esetén találunk olyan x valóst, hogy $a + x = b$. A szorzás is invertálható, de csak a nemzérus valósok halmazán. Ez azt jelenti, hogy bármely nemzérus valós a és b valóshoz található olyan x valós szám, hogy $ax = b$.

Elemi ismeret, hogy az $a + x = b$ egyenlet megoldásához elég ismerni *a ellentettjét*, és azt adni b-hez, az $ax = b$ egyenlet megoldásához elég ismerni *a reciprokát*, és azzal szorozni b-t. Az ellentett és a reciprok definíciójához mindkét esetben az adott művelet úgynevezett *semleges eleme* szükséges. Az összeadás semleges eleme a 0, mert bármely a elemhez adva a-t kapunk, hasonlóképp a szorzás semleges eleme az 1, mert bármely a elemet vele szorozva a-t kapunk. Az ellentetthez az $a + x = 0$ egyenlet megoldásával jutunk, a reciprokhoz az $ax = 1$ megoldásával. Az ellentettet, illetve a reciprokot additív, illetve multiplikatív *inverznek* is nevezzük. Keressük tehát a mátrixok additív, illetve multiplikatív inverzét.

Világos, hogy a mátrixösszeadás a számok összeadásánál tapasztaltakhoz hasonlóan invertálható: bármely azonos típusú (például $\mathbb{R}^{m \times n}$-beli) $\mathbf{A}$ és $\mathbf{B}$ mátrix esetén az $\mathbf{A} + \mathbf{X} = \mathbf{B}$ és az $\mathbf{Y} + \mathbf{A} = \mathbf{B}$ egyenlet megoldható, és mindkettő megoldása $\mathbf{B} + (-\mathbf{A})$, ahol $-\mathbf{A}$ az $\mathbf{A}$ ellentettje (additív inverze), vagyis az a mátrix, melyre $\mathbf{A} + (-\mathbf{A}) = \mathbf{O}$.

A mátrixok szorzásának invertálhatósága, és mátrixok (multiplikatív) inverzének definíciójához a következő szempontokat kell figyelembe venni:

- A mátrixszorzás nem kommutatív ezért a művelet invertálhatóságánál az $\mathbf{AX} = \mathbf{B}$ és az $\mathbf{YA} = \mathbf{B}$ egyenletre is gondolni kell, és ezeknek nem lesz szükségképp azonos a megoldásuk.
- Ha egy $\mathbf{A}$ mátrix inverzeként egyetlen mátrixot szeretnénk kapni, olyan $\mathbf{X}$ mátrixot kell keresni, melyre $\mathbf{AX} = \mathbf{I}$ és az $\mathbf{XA} = \mathbf{I}$ is fönnáll. Be lehet bizonyítani, hogy ez csak úgy lehetséges, ha $\mathbf{A}$ és $\mathbf{X}$ is négyzetes mátrixok.
- Vajon a számokhoz hasonlóan jelölhetjük-e az inverz elemet a -1-edik kitevőre emeléssel? A választ erre is a precedencia elv adja. Ha értelmezhető a negatív kitevőjű hatvány, akkor az 5.5. tétel érvényességét megtartva $\mathbf{A}^{-1}$ olyan mátrix kell hogy legyen, melyre $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}^{-1+1} = \mathbf{A}^0 = \mathbf{I}$ és ugyanígy $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{1-1} = \mathbf{A}^0 = \mathbf{I}$.
- Végül, minthogy a mátrixok közti műveleteket is a számok közti műveletek táblázatokra való kiterjesztésén keresztül vezettük be, elvárjuk, hogy az 1×1-es mátrixok inverze essen egybe a számok multiplikatív inverzével (reciprokával), azaz ha $\mathbf{A} = [a]$, akkor $\mathbf{A}^{-1} = [a^{-1}] = [1/a]$ legyen igaz.

E többfelől közelítő bevezetés után a definíció a következő:

A számítástechnikában gyakran találkozunk a műveletek *infix* jelölése mellett a *prefix* vagy *lengyel* és a *postfix* vagy *fordított lengyel jelölésével* is. A prefixnél a műveleti jel az argumentumai előtt, a postfixnél után van. Például a $(3 + 4) \cdot 2$ kifejezést a prefix jelölést használó Lisp nyelvcsalád nyelveiben a

```
(* (+ 3 4) 2)
```

kód, míg például a postfix jelölést használó PostScript nyelvben a

```
3 4 add 2 mul
```

kód számítja ki. (A PostScript nyelvvel találkozhatunk a PDF formátumú fájlokban is.) Ugyanez a formula a komputer algebra nyelvek közül a Mapleben prefix módon

```
`*`(`+`(3,4),2)
```

a Mathematicában

```
Times[Plus[3,4],2]
```

alakot ölt, míg a Sage a

```
prod([sum([3,4]),2])
```

és a

```
mul([add([3,4]),2])
```

formákat kínálja föl.

5.9. DEFINÍCIÓ (MÁTRIX INVERZE). *Legyen* $\mathbf{A}$ *egy* $n \times n$*-es mátrix. Azt mondjuk, hogy* $\mathbf{A}$ invertálható*, ha létezik olyan* $\mathbf{B}$ *mátrix, melyre*

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n.$$

A $\mathbf{B}$ *mátrixot* $\mathbf{A}$ inverzének *nevezzük, és* $\mathbf{A}^{-1}$*-nel jelöljük. A nem invertálható mátrixot* szingulárisnak *nevezzük.*

- Világos, hogy ha $\mathbf{A}$ inverze $\mathbf{B}$, akkor $\mathbf{B}$ inverze $\mathbf{A}$.
- A definícióból nem derül ki, hogy egy mátrixnak lehet-e több inverze, és hogy a mátrix milyen tulajdonsága befolyásolja invertálhatóságát. E kérdésekre hamarosan választ adunk.

5.10. PÉLDA (MÁTRIX INVERZE). *Tekintsük az*

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}, \quad \textit{és a} \quad \mathbf{B} = \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix}$$

mátrixokat. Számoljuk ki az $\mathbf{AB}$ *és a* $\mathbf{BA}$ *szorzatokat!*

$$\begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}_2, \qquad \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}_2.$$

Eszerint $\mathbf{A}$ *inverze* $\mathbf{B}$*, és ugyanakkor* $\mathbf{B}$ *inverze* $\mathbf{A}$*.*

5.11. PÉLDA (SZINGULÁRIS MÁTRIX). *Mutassuk meg, hogy az*

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 6 & 3 \end{bmatrix}$$

mátrix szinguláris.

MEGOLDÁS. Az $\mathbf{A}$ mátrix szinguláris, ha nem invertálható, azaz ha nincs olyan $\mathbf{X}$ mátrix, hogy $\mathbf{AX} = \mathbf{XA} = \mathbf{I}_2$. Már az is elég, ha megmutatjuk, hogy az $\mathbf{AX} = \mathbf{I}_2$ sosem állhat fönn! Vegyük észre, hogy $\mathbf{X}$-nek négy eleme van, így $\mathbf{X}$ meghatározása egy 4-ismeretlenes egyenletrendszerre vezet. Valóban, ha $\mathbf{X} = \left[\begin{smallmatrix} u & v \\ w & z \end{smallmatrix}\right]$, akkor az $\mathbf{AX} = \mathbf{I}_2$ egyenlet a következő szimultán egyenletrendszerrel ekvivalens:

$$\begin{aligned} 2u + w &= 1 \qquad & 2v + z &= 0 \\ 6u + 3w &= 0 \qquad & 6v + 3z &= 1 \end{aligned}$$

Ennek bővített mátrixát elemi sorműveletekkel lépcsős alakra hozzuk

$$\left[\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 6 & 3 & 0 & 1 \end{array}\right] \Longrightarrow \left[\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 0 & 0 & -3 & 1 \end{array}\right]$$

Ebből látjuk, hogy az egyenletrendszerek legalább egyike nem oldható meg, tehát nincs ilyen $\mathbf{X}$ mátrix, vagyis $\mathbf{A}$ nem invertálható! □

Egy négyzetes $\mathbf{A}$ mátrixot *nilpotensnek* nevezünk, ha van olyan k pozitív egész, hogy

$$\mathbf{A}^k = \mathbf{O}.$$

Például az $\left[\begin{smallmatrix}-2 & 4\\ -1 & 2\end{smallmatrix}\right]$ mátrix nilpotens, mert $\left[\begin{smallmatrix}-2 & 4\\ -1 & 2\end{smallmatrix}\right]^2 = \left[\begin{smallmatrix}0 & 0\\ 0 & 0\end{smallmatrix}\right]$.

5.12. PÉLDA ($\mathbf{I} - \mathbf{A}$ INVERZE NILPOTENS $\mathbf{A}$ ESETÉN). *Mutassuk meg, ha $\mathbf{A}$ nilpotens, azaz valamely pozitív k-ra $\mathbf{A}^k = \mathbf{O}$, akkor $\mathbf{I} - \mathbf{A}$ invertálható, és inverze $\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \ldots + \mathbf{A}^{k-1}$.*

MEGOLDÁS. Megmutatjuk, hogy $(\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \ldots + \mathbf{A}^{k-1}) = \mathbf{I}$.

$$\begin{aligned}
(\mathbf{I} - \mathbf{A})&(\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \ldots + \mathbf{A}^{k-1})\\
&= \mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \ldots + \mathbf{A}^{k-1} - \mathbf{A} - \mathbf{A}^2 - \ldots - \mathbf{A}^{k-1} - \mathbf{A}^k\\
&= \mathbf{I} - \mathbf{A}^k\\
&= \mathbf{I}
\end{aligned}$$

Az $(\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \ldots + \mathbf{A}^{k-1})(\mathbf{I} - \mathbf{A}) = \mathbf{I}$ összefüggés ugyanígy bizonyítható. □

Elemi mátrixok inverze Minden R elemi sorművelethez van egy olyan R' sorművelet, hogy az R sorművelettel átalakított mátrixot az R' visszaalakítja (ld. 2.21. feladat). Nevezzük ezt az R' sorműveletet az R sorművelet inverzének. Könnyen ellenőrizhető, hogy az $S_i \leftrightarrow S_j$ sorművelet inverze önmaga, a cS_i inverze $\frac{1}{c}S_i$, és $S_i + cS_j$ inverze $S_i - cS_j$.

5.13. ÁLLÍTÁS (SORMŰVELET INVERZÉNEK MÁTRIXA). *Minden elemi mátrix invertálható, nevezetesen egy sorművelet elemi mátrixának inverze megegyezik a sorművelet inverzének elemi mátrixával.*

A bizonyításhoz elég belátni, hogy egy sorművelet és az inverz sorművelet mátrixainak szorzata az egységmátrix. Az általános bizonyítás végiggondolását az Olvasóra hagyjuk, itt csak egy-egy konkrét esetet mutatunk meg, nevezetesen 3×3-as mátrixokon az $S_2 \leftrightarrow S_3$, a $3S_2$ és az $S_1 + 4S_3$ sorműveletek és inverzeik mátrixának szorzatát:

$$\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix} = \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix},$$

$$\begin{bmatrix}1&0&0\\0&3&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&\frac{1}{3}&0\\0&0&1\end{bmatrix} = \begin{bmatrix}1&0&0\\0&\frac{1}{3}&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&3&0\\0&0&1\end{bmatrix} = \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix},$$

$$\begin{bmatrix}1&0&4\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&-4\\0&1&0\\0&0&1\end{bmatrix} = \begin{bmatrix}1&0&-4\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&4\\0&1&0\\0&0&1\end{bmatrix} = \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}.$$

Az inverz kiszámítása A négyzetes $\mathbf{A}$ mátrix inverzének kiszámításához meg kell oldani az $\mathbf{AX} = \mathbf{I}$ mátrixegyenletet, ami egyúttal egy szimultán egyenletrendszer is, és az elemi sorműveletekkel megoldható.

Előbb azonban két kérdésre válaszolnunk kell: *(1)* nem fordulhat-e elő, hogy $\mathbf{A}$-nak több inverze is van, és *(2)* nem fordulhat-e elő, hogy az $\mathbf{AX} = \mathbf{I}$ mátrixegyenlet megoldható, de a megoldás nem tesz eleget az $\mathbf{XA} = \mathbf{I}$ mátrixegyenletnek? Négyzetes mátrixok esetén mindkét kérdésre *nem* a válasz, ami azt jelenti, hogy mátrix invertálásához elég az $\mathbf{AX} = \mathbf{I}$ mátrixegyenlet megoldása!

5.14. TÉTEL (AZ INVERZ EGYÉRTELMŰSÉGE). *Ha a négyzetes* $\mathbf{A}$ *mátrixhoz létezik olyan ugyancsak négyzetes* $\mathbf{B}$ *és* $\mathbf{C}$ *mátrix, hogy* $\mathbf{AB} = \mathbf{CA} = \mathbf{I}$, *akkor* $\mathbf{B} = \mathbf{C}$. *Ennek következménye, hogy egy mátrixnak legföljebb csak egy inverze lehet.*

BIZONYÍTÁS. $\mathbf{C} = \mathbf{CI} = \mathbf{C}(\mathbf{AB}) = (\mathbf{CA})\mathbf{B} = \mathbf{IB} = \mathbf{B}$, ami igazolja az első állítást. Ha $\mathbf{A}$-nak $\mathbf{B}$ és $\mathbf{C}$ is inverze volna, fönnállna az $\mathbf{AB} = \mathbf{I}$ és a $\mathbf{CA} = \mathbf{I}$ egyenlőség is, de ekkor $\mathbf{B} = \mathbf{C}$, tehát az inverz egyértelmű. □

5.15. TÉTEL (AZ INVERZ LÉTEZÉSÉHEZ ELÉG EGY FELTÉTEL). *A négyzetes* $\mathbf{A}$ *mátrix pontosan akkor invertálható, ha létezik olyan* $\mathbf{B}$ *mátrix, hogy az* $\mathbf{AB} = \mathbf{I}$ *és a* $\mathbf{BA} = \mathbf{I}$ *feltételek egyike teljesül. Másként szólva, a fenti két feltétel bármelyikének teljesülése maga után vonja a másik teljesülését is!*

BIZONYÍTÁS. Megmutatjuk, hogy ha a négyzetes $\mathbf{A}$ és $\mathbf{B}$ mátrixok kielégítik az $\mathbf{AB} = \mathbf{I}$ egyenletet, akkor $\mathbf{BA} = \mathbf{I}$ is fönnáll, azaz $\mathbf{A}$ és $\mathbf{B}$ inverzei egymásnak.

Tekintsük az $\mathbf{AX} = \mathbf{I}$ mátrixegyenletet. Ezt úgy oldjuk meg, hogy az $[\mathbf{A}|\mathbf{I}]$ mátrixot redukált lépcsős alakra hozzuk. Ha ez $[\mathbf{I}|\mathbf{B}]$ alakú, akkor $\mathbf{B}$ az $\mathbf{AX} = \mathbf{I}$ egyenlet megoldása, ezért $\mathbf{AB} = \mathbf{I}$ fennáll. A redukált lépcsős alakban zérus sor nem keletkezhet, mert a mátrix jobb oldalát az $\mathbf{I}$ mátrixból kaptuk, ami redukált lépcsős alak, s így egyértelmű. Ha elemi sorműveletekkel zérus sort kapnánk a jobb oldali félmátrixban, akkor volna olyan redukált lépcsős alakja is, mely zérus sort tartalmazna, ami ellentmondás. Ha csak a mátrix bal felén kapnánk zérus sort, akkor az $\mathbf{AX} = \mathbf{I}$ egyenletnek nem lenne megoldása, vagyis nem állhatna fenn az $\mathbf{AB} = \mathbf{I}$ egyenlőség sem.

Ezután megmutatjuk, hogy $\mathbf{BA} = \mathbf{I}$. Ehhez tekintsük a $\mathbf{BY} = \mathbf{I}$ mátrixegyenletet. Ennek megoldásához a $[\mathbf{B}|\mathbf{I}]$ mátrixot kell redukált lépcsős alakra hozni. A előzőekből tudjuk, hogy elemi sorműveletekkel az

$$[\mathbf{A}|\mathbf{I}] \Longrightarrow [\mathbf{I}|\mathbf{B}]$$

átalakítás megvalósítható. Az átalakítás lépéseinek inverzeit fordított sorrendben elvégezve az

$$[\mathbf{I}|\mathbf{B}] \Longrightarrow [\mathbf{A}|\mathbf{I}]$$

transzformációt kapjuk. Itt minden lépésben fölcserélve a két részmátrixot a kívánt

$$[\mathbf{B}|\mathbf{I}] \Longrightarrow [\mathbf{I}|\mathbf{A}]$$

átalakítást kapjuk. Ez azt jelenti, hogy a $\mathbf{BY} = \mathbf{I}$ mátrixegyenletetnek, az $\mathbf{Y} = \mathbf{A}$ megoldása, azaz $\mathbf{BA} = \mathbf{I}$. □

Összefoglalva:

5.1 (Inverz kiszámítása elemi sorműveletekkel). *A négyzetes* $\mathbf{A}$ *mátrix inverze kiszámítható, ha az* $[\mathbf{A}|\mathbf{I}]$ *mátrixot elemi sorműveletekkel* $[\mathbf{I}|\mathbf{B}]$ *alakra hozzuk, ekkor* $\mathbf{A}$ *inverze* $\mathbf{B}$ *lesz. Ha* $\mathbf{A}$ *redukált lépcsős alakja nem az* $\mathbf{I}$ *mátrix, akkor* $\mathbf{A}$ *nem invertálható.*

5.16. példa (Az inverz kiszámítása). *Számítsuk ki az*

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 6 \end{bmatrix} \quad \textit{és a} \quad \mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 \\ 4 & 3 & 2 & 1 \end{bmatrix}$$

mátrixok inverzét!

Megoldás. A kiküszöböléssel oszloponként haladva:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 3 & 4 & 0 & 1 & 0 \\ 3 & 4 & 6 & 0 & 0 & 1 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & -2 & -2 & 1 & 0 \\ 0 & -2 & -3 & -3 & 0 & 1 \end{array}\right] \Rightarrow$$
$$\left[\begin{array}{ccc|ccc} 1 & 0 & -1 & -3 & 2 & 0 \\ 0 & 1 & 2 & 2 & -1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & 0 & 1 \\ 0 & 1 & 0 & 0 & 3 & -2 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right]$$

Tehát

$$\mathbf{A}^{-1} = \begin{bmatrix} -2 & 0 & 1 \\ 0 & 3 & -2 \\ 1 & -2 & 1 \end{bmatrix}.$$

A $\mathbf{B}$ inverzének kiszámítása hasonló lépésekkel:

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 & 0 & 0 & 1 & 0 \\ 4 & 3 & 2 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \Rightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 & -3 & 0 & 1 & 0 \\ 0 & 3 & 2 & 1 & -4 & 0 & 0 & 1 \end{array}\right] \Rightarrow$$
$$\left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 2 & 1 & 2 & -3 & 0 & 1 \end{array}\right] \Rightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & -2 & 1 \end{array}\right]$$

Tehát

$$\mathbf{B}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{bmatrix}.$$

□

5.17. TÉTEL (2 × 2-ES MÁTRIX INVERZE). *Az* $\mathbf{A} = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ *mátrix pontosan akkor invertálható, ha* $ad - bc \neq 0$*, és ekkor*

$$\mathbf{A}^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

BIZONYÍTÁS. Azt, hogy az $\mathbf{A}$ mátrixnak valóban a fenti mátrix az inverze, egyszerű mátrixszorzással ellenőrizhetjük. Azt, hogy az $ad - bc \neq 0$ feltétel az invertálhatóságnak elégséges feltétele, a képlet bizonyítja. A feltétel szükségességének belátásához vegyük észre, hogy $ad - bc = 0$, azaz $ad = bc$ pontosan akkor áll fenn, ha $\mathbf{A}$ egyik sora a másik skalárszorosa. Ekkor viszont az egyik sor kinullázható, vagyis az $\mathbf{A}$ mátrix nem alakítható elemi sorműveletekkel egységmátrixszá. □

Az inverz tulajdonságai Megvizsgáljuk a mátrixinvertálás más műveletekkel való kapcsolatát.

5.18. TÉTEL (AZ INVERZ ALAPTULAJDONSÁGAI). *Tegyük fel, hogy* $\mathbf{A}$ *és* $\mathbf{B}$ *egyaránt* $n \times n$*-es invertálható mátrixok,* $c \neq 0$ *skalár és* k *pozitív egész. Ekkor igazak a következők:*

a) $\mathbf{A}^{-1}$ *invertálható, és inverze* $\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}$,
b) $c\mathbf{A}$ *invertálható, és inverze* $\frac{1}{c}\mathbf{A}^{-1}$,
c) $\mathbf{AB}$ *invertálható, és inverze* $\mathbf{B}^{-1}\mathbf{A}^{-1}$,
d) $\mathbf{A}^k$ *invertálható, és inverze* $\left(\mathbf{A}^k\right)^{-1} = \left(\mathbf{A}^{-1}\right)^k$*, definíció szerint ezt értjük* $\mathbf{A}^{-k}$*-n,*
e) $\mathbf{A}^T$ *invertálható, és* $\left(\mathbf{A}^T\right)^{-1} = \left(\mathbf{A}^{-1}\right)^T$.

BIZONYÍTÁS. Az állítások mindegyikét bizonyítjuk:

a) $\mathbf{A}^{-1}$ invertálható, ha van olyan $\mathbf{B}$ mátrix, hogy $\mathbf{A}^{-1}\mathbf{B} = \mathbf{B}\mathbf{A}^{-1} = \mathbf{I}$. Ilyen mátrix viszont van, hisz $\mathbf{A}$ invertálható, és inverzére $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$. Eszerint a $\mathbf{B} = \mathbf{A}$ választás megfelel, és $\mathbf{A}^{-1}$ inverze $\mathbf{A}$.
b) A $(c\mathbf{A})(\frac{1}{c}\mathbf{A}^{-1}) = (c \cdot \frac{1}{c})(\mathbf{A}\mathbf{A}^{-1}) = \mathbf{I}$ szorzat bizonyítja az állítást.
c) Az

$$(\mathbf{AB})(\mathbf{B}^{-1}\mathbf{A}^{-1}) = \mathbf{A}(\mathbf{B}\mathbf{B}^{-1})\mathbf{A}^{-1} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$$

szorzat bizonyítja, hogy $\mathbf{AB}$ invertálható, és inverze $\mathbf{B}^{-1}\mathbf{A}^{-1}$.

d) Az $\left(\mathbf{A}^k\right)^{-1} = \left(\mathbf{A}^{-1}\right)^k$ egyenlőség igaz volta a

$$\underbrace{\mathbf{A}\mathbf{A}\ldots\mathbf{A}}_{k \text{ tényező}}\underbrace{\mathbf{A}^{-1}\mathbf{A}^{-1}\ldots\mathbf{A}^{-1}}_{k \text{ tényező}} = \underbrace{\mathbf{A}^{-1}\mathbf{A}^{-1}\ldots\mathbf{A}^{-1}}_{k \text{ tényező}}\underbrace{\mathbf{A}\mathbf{A}\ldots\mathbf{A}}_{k \text{ tényező}} = \mathbf{I}$$

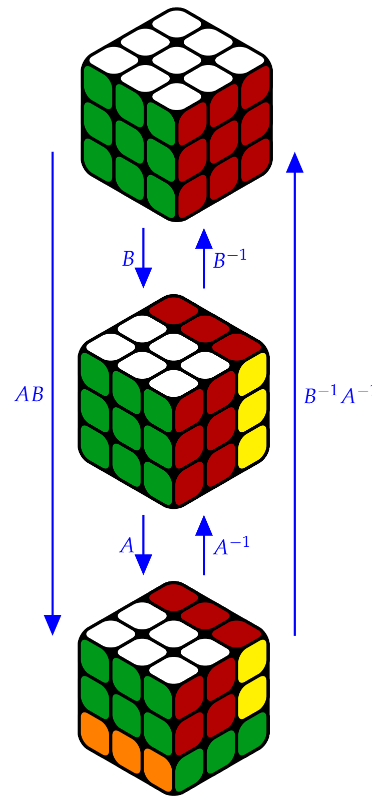
felírásból leolvasható, mert a szorzatok közepén lévő két mátrix szorzata mindig $\mathbf{I}$, ami elhagyható, és e lépést k-szor ismételve végül a kívánt eredményt kapjuk:

$$\underbrace{\mathbf{A}\mathbf{A}\ldots(\mathbf{A}}_{k \text{ tényező}}\underbrace{\mathbf{A}^{-1})\mathbf{A}^{-1}\ldots\mathbf{A}^{-1}}_{k \text{ tényező}} = \underbrace{\mathbf{A}\mathbf{A}\ldots\mathbf{A}}_{k-1 \text{ tényező}}\underbrace{\mathbf{A}^{-1}\ldots\mathbf{A}^{-1}}_{k-1 \text{ tényező}} = \cdots = \mathbf{I}.$$

e) Ha $\mathbf{A}$ invertálható, akkor $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$, és ennek transzponáltját véve az

$$\left(\mathbf{A}^{-1}\right)^T \mathbf{A}^T = \mathbf{A}^T \left(\mathbf{A}^{-1}\right)^T = \mathbf{I}^T = \mathbf{I}$$

egyenlőségből leolvasható, hogy $\mathbf{A}^T$ invertálható, és inverze $\left(\mathbf{A}^T\right)^{-1} = \left(\mathbf{A}^{-1}\right)^T$. □



5.2. ábra: Jelölje A a bűvös kocka alsó, B a jobb hátsó oldalának elforgatását, és jelölje AB a B, majd az A egymás után való elvégzésével kapott transzformációt. (Ahogy a függvények összetételénél, előbb a jobb oldali, majd a bal oldali függvényt értékeljük ki, hajtjuk végre.) Ennek inverze $(AB)^{-1}$ úgy kapható meg, ha előbb végrehajtjuk az A^{-1} majd a B^{-1} transzformációt. Ezek szorzata $B^{-1}A^{-1}$, tehát $(AB)^{-1} = B^{-1}A^{-1}$.

► A *(c)* állítás indukcióval általánosítható véges sok mátrix szorzatára: ha az azonos méretű négyzetes $\mathbf{A}_1$, $\mathbf{A}_2$,… $\mathbf{A}_m$ mátrixok mindegyike invertálható, akkor szorzatuk is, és

$$(\mathbf{A}_1\mathbf{A}_2\ldots\mathbf{A}_m)^{-1} = \mathbf{A}_m^{-1}\ldots\mathbf{A}_2^{-1}\mathbf{A}_1^{-1}.$$

► A *(c)* állításbeli összefüggéshez hasonlóval találkozhatunk a Rubik-kocka forgatása közben is. Jelölje A az egyik oldal derékszöggel való elforgatását, és B egy másik oldalét. A függvények kompozíciójához hasonlóan definiáljuk e két transzformáció szorzatát: a B majd az A forgatás egymás után való elvégzésével kapott transzformációt tekintsük a két forgatás szorzatának, és jelöljük AB-vel (ld. 5.2 ábra). E transzformációt invertálni akarjuk, azaz vissza akarjuk állítani az eredeti állapotot, azaz azt az $(AB)^{-1}$ transzformációt keressük, melyet AB után elvégezve az identikus (egy kockát sem mozgató) transzformációt kapjuk. Világos, előbb a A transzformáció inverzét kell végrehajtani, majd azután az B inverzét, azaz $(AB)^{-1} = B^{-1}A^{-1}$.

► Az $\mathbf{A}^{-k}$ *(d)* pontbeli definíciója is a precedencia elvből következik, például az $\mathbf{A}^m\mathbf{A}^n = \mathbf{A}^{m+n}$ összefüggés kiterjesztése negatív kitevőre az $\mathbf{A}^k\mathbf{A}^{-k} = \mathbf{A}^0$ formulához vezet, amiből azt kapjuk, hogy $\mathbf{A}^{-k} = (\mathbf{A}^k)^{-1}$.

5.19. PÉLDA (INVERZ TULAJDONSÁGAINAK ALKALMAZÁSA). *Az 5.16. példa mátrixainak inverzét használva számítsuk ki az alábbi* *mátrixok inverzét!*

$$\begin{bmatrix}3 & 6 & 9\\ 6 & 9 & 12\\ 9 & 12 & 18\end{bmatrix}, \quad \begin{bmatrix}1 & 0 & 0 & 0\\ 2 & 1 & 0 & 0\\ 3 & 2 & 1 & 0\\ 4 & 3 & 2 & 1\end{bmatrix}^2, \quad \begin{bmatrix}2 & 4 & 6 & 8\\ 0 & 2 & 4 & 6\\ 0 & 0 & 2 & 4\\ 0 & 0 & 0 & 2\end{bmatrix}.$$

MEGOLDÁS. A feladatban megadott mátrixok kifejezhetők az 5.16. példában megadott $\mathbf{A}$ és $\mathbf{B}$ mátrixokkal. A három invertálandó mátrix: $3\mathbf{A}$, $\mathbf{B}^2$, $2\mathbf{A}^T$. Ezek inverzei az előző tétel szerint rendre $\frac{1}{3}\mathbf{A}^{-1}$, $(\mathbf{B}^{-1})^2$, $\frac{1}{2}(\mathbf{B}^{-1})^T$. Tehát a három inverz:

$$(3\mathbf{A})^{-1} = \frac{1}{3}\mathbf{A}^{-1} = \frac{1}{3}\begin{bmatrix}-2 & 0 & 1\\ 0 & 3 & -2\\ 1 & -2 & 1\end{bmatrix} = \begin{bmatrix}-\frac{2}{3} & 0 & \frac{1}{3}\\ 0 & 1 & -\frac{2}{3}\\ \frac{1}{3} & -\frac{2}{3} & \frac{1}{3}\end{bmatrix}$$

$$(\mathbf{B}^2)^{-1} = (\mathbf{B}^{-1})^2 = \begin{bmatrix}1 & 0 & 0 & 0\\ -2 & 1 & 0 & 0\\ 1 & -2 & 1 & 0\\ 0 & 1 & -2 & 1\end{bmatrix}^2 = \begin{bmatrix}1 & 0 & 0 & 0\\ -4 & 1 & 0 & 0\\ 6 & -4 & 1 & 0\\ -4 & 6 & -4 & 1\end{bmatrix}$$

$$(2\mathbf{A}^T)^{-1} = \frac{1}{2}(\mathbf{B}^{-1})^T = \frac{1}{2}\begin{bmatrix}1 & 0 & 0 & 0\\ -2 & 1 & 0 & 0\\ 1 & -2 & 1 & 0\\ 0 & 1 & -2 & 1\end{bmatrix}^T = \begin{bmatrix}\frac{1}{2} & -1 & \frac{1}{2} & 0\\ 0 & \frac{1}{2} & -1 & \frac{1}{2}\\ 0 & 0 & \frac{1}{2} & -1\\ 0 & 0 & 0 & \frac{1}{2}\end{bmatrix}. \square$$

Az invertálhatóság és az egyenletrendszerek megoldhatósága A következő tétel a mátrixok invertálhatóságát, az egyenletrendszerek megoldásánál használt elemi sorműveleteket és az egyenletrendszerek megoldhatóságát kapcsolja össze.

5.20. TÉTEL (AZ INVERTÁLHATÓSÁG ÉS AZ EGYENLETRENDSZEREK). *Adva van egy $n \times n$-es $\mathbf{A}$ mátrix. Az alábbi állítások ekvivalensek:*

a) $\mathbf{A}$ *invertálható;*

b) az $\mathbf{AX} = \mathbf{B}$ mátrixegyenlet bármely $n \times t$-es $\mathbf{B}$ mátrixra egyértelműen megoldható;

c) az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszer bármely n dimenziós $\mathbf{b}$ vektorra egyértelműen megoldható;

d) a homogén lineáris $\mathbf{Ax} = \mathbf{0}$ egyenletrendszernek a triviális $\mathbf{x} = \mathbf{0}$ az egyetlen megoldása;

e) $\mathbf{A}$ *redukált lépcsős alakja* $\mathbf{I}$*;*

f) $\mathbf{A}$ *előáll elemi mátrixok szorzataként.*

BIZONYÍTÁS. Az állítások ekvivalenciáját az *(a)* $\Rightarrow$ *(b)* $\Rightarrow$ *(c)* $\Rightarrow$ *(d)* $\Rightarrow$ *(e)* $\Rightarrow$ *(f)* $\Rightarrow$ *(a)*, implikációk igazolásával bizonyítjuk.

(a) $\Rightarrow$ *(b)*: Legyen tehát $\mathbf{A}$ invertálható és legyen $\mathbf{B}$ egy tetszőleges $n \times t$ méretű mátrix. Ekkor az $\mathbf{AX} = \mathbf{B}$ egyenlet mindkét oldalát $\mathbf{A}^{-1}$-gyel balról szorozva kapjuk, hogy $\mathbf{A}^{-1}\mathbf{AX} = \mathbf{A}^{-1}\mathbf{B}$, azaz $\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$. Ez azt mutatja, hogy egyrészt a mátrixegyenletnek van megoldása, másrészt hogy más megoldása nincs, mivel így minden megoldás megkapható, és $\mathbf{A}$ inverze egyértelmű.

(b) $\Rightarrow$ *(c)*: Nyilvánvaló a $\mathbf{B} = \mathbf{b}$ választással.

(c) $\Rightarrow$ *(d)*: Nyilvánvaló a $\mathbf{b} = \mathbf{0}$ választással.

(d) $\Rightarrow$ *(e)*: Egy n-ismeretlenes, n egyenletből álló homogén lineáris egyenletrendszer pontosan akkor oldható meg egyértelműen, ha együtthatómátrixának redukált lépcsős alakja $\mathbf{I}_n$.

(e) $\Rightarrow$ *(f)*: Ha $\mathbf{A}$ redukált lépcsős alakja $\mathbf{I}_n$, akkor létezik elemi sorműveletek olyan sorozata, mely az $\mathbf{A} \Rightarrow \mathbf{I}_n$ transzformációt elvégzi. Jelölje az elemi sorműveletekhez tartozó elemi mátrixokat $\mathbf{E}_1, \ldots \mathbf{E}_k$. Ekkor tehát $\mathbf{E}_1\mathbf{E}_2 \ldots \mathbf{E}_k\mathbf{A} = \mathbf{I}_n$. Innen $\mathbf{A}$ kifejezhető az $\mathbf{E}_1^{-1}$-nel,… $\mathbf{E}_k^{-1}$-nel balról való beszorzás után:

$$\mathbf{A} = \mathbf{E}_k^{-1} \ldots \mathbf{E}_2^{-1}\mathbf{E}_1^{-1}.$$

Elemi mátrixok inverze elemi mátrix, tehát $\mathbf{A}$ előáll elemi mátrixok szorzataként.

(f) $\Rightarrow$ *(a)*: Az $\mathbf{A} = \mathbf{E}_k^{-1} \ldots \mathbf{E}_2^{-1}\mathbf{E}_1^{-1}$ mátrix minden tényezője invertálható, mivel mindegyik elemi mátrix, így szorzatuk is, és az inverz

$$\mathbf{A}^{-1} = \mathbf{E}_1\mathbf{E}_2 \ldots \mathbf{E}_k.$$ □

► A tétel sok pontjának ekvivalenciája azt jelenti, hogy közülük bármely kettőre igaz, hogy „az egyik pontosan akkor igaz, ha a másik”. Például „$\mathbf{A}$ pontosan akkor invertálható, ha az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszer minden $\mathbf{b}$ vektorra egyértelműen megoldható”.

► Később megmutatjuk azt is, hogy $\mathbf{A}$ pontosan akkor invertálható, ha az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszer minden $\mathbf{b}$ vektorra megoldható. Azaz az egyértelműség a feltételből kihagyható. Másként fogalmazva, ha $\mathbf{Ax} = \mathbf{b}$ minden $\mathbf{b}$ vektorra megoldható, akkor a megoldás minden $\mathbf{b}$-re egyértelmű.

5.21. PÉLDA (EGYENLETRENDSZER MEGOLDÁSA MÁTRIXINVERTÁLÁSSAL). *Oldjuk meg az*

$$\begin{aligned} 2x + y &= 2 \\ 5x + 3y &= 3 \end{aligned}$$

egyenletrendszert mátrixinvertálással.

MEGOLDÁS. Az együtthatómátrix és inverze az 5.17. tétel szerint

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}, \quad \mathbf{A}^{-1} = \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix},$$

így az ismeretlenek (x, y) vektorára

$$\begin{bmatrix} x \\ y \end{bmatrix} = \mathbf{A}^{-1} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}.$$

□

5.22. PÉLDA (MÁTRIXEGYENLET MEGOLDÁSA MÁTRIXINVERTÁLÁSSAL). *Oldjuk meg az* $\mathbf{AX} = \mathbf{B}$ *mátrixegyenletet, ahol*

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}, \quad \textit{és} \quad \mathbf{B} = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 3 & 1 \end{bmatrix}.$$

MEGOLDÁS. Az $\mathbf{A}$ mátrix megegyezik az előző feladatbeli mátrixszal, így tudjuk, hogy invertálható, és ismerjük az inverzét. Az $\mathbf{AX} = \mathbf{B}$ mátrixegyenlet megoldása:

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{B} = \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 1 & 3 & 2 \\ 4 & 3 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 6 & 5 \\ 3 & -9 & -8 \end{bmatrix}.$$

□

► Megjegyezzük, hogy lineáris egyenletrendszert mátrixinvertálással ritkán oldunk meg, mert műveleigénye valamivel nagyobb, mint az egyszerű kiküszöbölésnek.

5.23. PÉLDA (MÁTRIX ELEMI MÁTRIXOK SZORZATÁRA BONTÁSA). *Bontsuk fel az* $\mathbf{A} = \left[\begin{smallmatrix} 1 & 2 \\ 3 & 5 \end{smallmatrix}\right]$ *mátrixot elemi mátrixok szorzatára!*

MEGOLDÁS. Az 5.20. tétel bizonyításának *(e)* $\Rightarrow$ *(f)* lépése szerint ha egy $\mathbf{A}$ mátrixot elemi sorműveletekkel az egységmátrixba lehet transzformálni, akkor az elemi sorműveletek inverzei fordított sorrendben elvégezve az $\mathbf{I}$-t $\mathbf{A}$-ba transzformálják. Ez viszont azt jelenti, hogy a hozzájuk tartozó elemi mátrixok szorzata épp $\mathbf{A}$.

Elemi sorműveletek		Elemi mátrixok	Elemi mátrixok inverzei
$\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$			
$\Downarrow$	$S_2 - 3S_1$	$\mathbf{E}_1 = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}$	$\mathbf{E}_1^{-1} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}$
$\begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$			
$\Downarrow$	$-S_2$	$\mathbf{E}_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	$\mathbf{E}_2^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$			
$\Downarrow$	$S_1 - 2S_2$	$\mathbf{E}_3 = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$	$\mathbf{E}_3^{-1} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$
$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$			

A fenti átalakítás nyomán tehát $\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \mathbf{I}$, amiből $\mathbf{A} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\mathbf{E}_3^{-1}$, azaz

$$\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix},$$

így $\mathbf{A}$-t három elemi mátrix szorzatára bontottuk. □

Invertálhatóság, bázis, báziscsere Az 5.20. tétel szerint a négyzetes $\mathbf{A}$ mátrix invertálhatósága azzal ekvivalens, hogy a homogén lineáris $\mathbf{Ax} = \mathbf{0}$ egyenletrendszernek a triviális az egyetlen megoldása. Mivel $\mathbf{Ax}$ az $\mathbf{A}$ oszlopvektorainak egy lineáris kombinációja, ezért ez azt jelenti, hogy a nullvektor csak egyféleképp áll elő $\mathbf{A}$ oszlopvektorainak lineáris kombinációjaként, a triviális módon. Tehát $\mathbf{A}$ oszlopvektorai lineárisan függetlenek! Ez egyúttal azt is jelenti, hogy $\mathbf{A}$ oszlopvektorai bázist alkotnak, és hogy $\mathrm{r}(\mathbf{A}) = n$. Felhasználva a **??** tételt, mely szerint a sortér és az oszloptér dimenziója megegyezik a ranggal, a következő tételt kapjuk:

5.24. KÖVETKEZMÉNY (INVERTÁLHATÓSÁG ÉS BÁZIS). *Adva van egy valós $n \times n$-es $\mathbf{A}$ mátrix. Az alábbi állítások ekvivalensek:*

a) $\mathbf{A}$ *invertálható;*
b) $\mathbf{A}$ *oszlopvektorai lineárisan függetlenek;*
c) $\mathbf{A}$ *oszlopvektorai bázist alkotnak* $\mathbb{R}^n$*-ben;*
d) $\mathbf{A}$ *sorvektorai lineárisan függetlenek;*
e) $\mathbf{A}$ *sorvektorai bázist alkotnak* $\mathbb{R}^n$*-ben;*
f) $\mathrm{r}(\mathbf{A}) = n$.

Legyen $B = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ és $C = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$ az $\mathbf{R}^n$ két bázisa, és jelölje $\mathbf{X}_{C\leftarrow B}$ a B-ről C-re, $\mathbf{Y}_{B\leftarrow C}$ a C-ről B-re való áttérés mátrixát. Legyen továbbá $\mathbf{v}$ a tér egy tetszőleges vektora, a B bázisbeli alakja $[\mathbf{v}]_B$. A 4.22. tétel szerint

$$[\mathbf{v}]_C = \mathbf{X}_{C\leftarrow B}[\mathbf{v}]_B, \quad \text{és} \quad [\mathbf{v}]_B = \mathbf{Y}_{B\leftarrow C}[\mathbf{v}]_C.$$

A második egyenletbe helyettesítve az elsőt kapjuk, hogy

$$[\mathbf{v}]_B = \mathbf{Y}_{B\leftarrow C}\mathbf{X}_{C\leftarrow B}[\mathbf{v}]_B,$$

azaz $\mathbf{Y}_{B\leftarrow C}\mathbf{X}_{C\leftarrow B}$ minden vektort önmagába visz, tehát egyenlő az egységmátrixszal, vagyis $\mathbf{Y}_{B\leftarrow C}$ és $\mathbf{X}_{C\leftarrow B}$ inverzei egymásnak.

5.25. TÉTEL (AZ ÁTTÉRÉS MÁTRIXÁNAK INVERZE). *Ha* $B = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ *és* $C = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$ *az* $\mathbf{R}^n$ *két bázisa, akkor a B-ről C-re való áttérés* $\mathbf{X}_{C\leftarrow B}$ *mátrixa, valamint a C-ről B-re való áttérés* $\mathbf{Y}_{B\leftarrow C}$ *mátrixa is invertálható, és egymás inverzei, azaz* $\mathbf{X}_{C\leftarrow B}^{-1} = \mathbf{Y}_{B\leftarrow C}$ *vagy más alakban* $\mathbf{X}_{C\leftarrow B}\mathbf{Y}_{B\leftarrow C} = \mathbf{I}_n$.

5.26. PÉLDA (AZ ÁTTÉRÉS MÁTRIXÁNAK INVERZE). *Az* $\mathbb{R}^3$ *egy* $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ *bázisában felírtuk a standard egységvektorokat:*

$$\mathbf{i} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}_B, \quad \mathbf{j} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}_B, \quad \mathbf{k} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}_B.$$

Írjuk fel $\mathcal{B}$ *bázisvektorainak standard bázisbeli koordinátás alakját!*

MEGOLDÁS. Jelölje $\mathcal{E}$ a standard bázist. Ennek vektorait kifejeztük a $\mathcal{B}$ bázis elemeivel, az ezekből képzett mátrixszal tehát az $\mathcal{E}$-beli vektorok $\mathcal{B}$-beli koordinátás alakja fölírható, tehát ez a $\mathcal{B} \leftarrow \mathcal{E}$ áttérés mátrixa, azaz

$$\mathbf{X}_{\mathcal{B}\leftarrow\mathcal{E}} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 2 & 4 \end{bmatrix}.$$

Ennek inverze a keresett mátrix:

$$\mathbf{Y}_{\mathcal{E}\leftarrow\mathcal{B}} = \mathbf{X}^{-1}_{\mathcal{B}\leftarrow\mathcal{E}} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 2 & 4 \end{bmatrix}^{-1} = \begin{bmatrix} 2 & -2 & 1 \\ -1 & 3 & -2 \\ 0 & -1 & 1 \end{bmatrix}.$$

Ennek oszlopvektorai adják a $\mathcal{B}$ vektorainak $\mathcal{E}$-beli alakját. □

Feladatok

Igaz – hamis

5.1. A négyzetes $\mathbf{A}$ mátrix pontosan akkor invertálható, ha elemi sorműveletekkel megkapható az $\mathbf{I}$ mátrixból.

5.2. Ha elemi sorműveletek $\mathbf{A}$-t $\mathbf{B}$-be viszik, akkor az inverz sorműveletek $\mathbf{B}$-t $\mathbf{A}$-ba viszik.

5.3. Ha elemi sorműveletek $\mathbf{A}$-t $\mathbf{B}$-be viszik, akkor az inverz sorműveletek fordított sorrendben végrehajtva $\mathbf{B}$-t $\mathbf{A}$-ba viszik.

Műveleti azonosságok

5.4. Egy algebrai kifejezésben végrehajtjuk az alábbi helyettesítést:

$$\begin{aligned} u &= 3x_1 + 2x_2 + 4x_3 \\ v &= x_1 - 3x_2 + x_3 \\ w &= 2x_1 - x_2 - 3x_3 \end{aligned}$$

Írjuk fel e lineáris helyettesítést mátrixszorzatos alakban. Legyen $(u^2+v^2+w^2)(2u-v-w)$ az a kifejezés, melyben a helyettestést elvégezzük. Írjuk fel e kifejezést a helyettesítés előtt és után mátrixműveletek segítségével!

Számítási feladatok

Bontsuk fel a következő mátrixokat elemi mátrixok szorzatára!

5.5. $\begin{bmatrix} 1 & 3 \\ 2 & 8 \end{bmatrix}$

5.6. $\begin{bmatrix} 1 & 2 \\ -2 & -1 \end{bmatrix}$

5.7. $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

5.8. $\begin{bmatrix} 2 & 4 \\ 3 & 8 \end{bmatrix}$

5.9. $\begin{bmatrix} 2 & 0 & 4 \\ 0 & 2 & 0 \\ 3 & 2 & 7 \end{bmatrix}$

5.10. $\begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 2 \\ 2 & 4 & 5 \end{bmatrix}$

5.11. Határozzuk meg az összes olyan 2×2-es $\mathbf{A}$ mátrixot, melyre $\mathbf{A}^2 = \mathbf{O}$. Másként fogalmazva határozzuk meg a nullmátrix összes négyzetgyökét!

5.12. Számítsuk ki az

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

mátrix k-adik hatványait!

5.13. Írjuk fel a mátrixszorzás definícióját az Einstein-konvenciót használva.

Blokkmátrixok

5.14. Mutassuk meg, hogy ha $\mathbf{A}$ és $\mathbf{D}$ invertálható mátrixok, akkor a következő ún. blokkdiagonális mátrix invertálható, és inverze

$$\begin{bmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} & \mathbf{O} \\ \mathbf{O} & \mathbf{D}^{-1} \end{bmatrix}.$$

továbbá tetszőleges, de megfelelő típusú $\mathbf{B}$ mátrix esetén

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{D}^{-1} \\ \mathbf{O} & \mathbf{D}^{-1} \end{bmatrix}.$$

5.15. Mutassuk meg, hogy ha $\mathbf{A}$ és $\mathbf{D}$ négyzetes mátrixok, akkor

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{X} & -\mathbf{X}\mathbf{B}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}\mathbf{X} & \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}\mathbf{X}\mathbf{B}\mathbf{D}^{-1} \end{bmatrix},$$

ahol $\mathbf{X} = (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}$, és feltételezzük, hogy minden felírt mátrixinverz létezik.

Az előbbi két feladat valamelyikének felhasználásával számítsuk ki az alábbi mátrixok inverzét!

5.16. $\begin{bmatrix} 2 & 3 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 7 & 3 & 3 \\ 0 & 0 & 8 & 1 & 2 \\ 0 & 0 & 4 & 4 & 3 \end{bmatrix}$

5.17. $\begin{bmatrix} 2 & 3 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 & 1 \\ 0 & 0 & 7 & 3 & 3 \\ 0 & 0 & 8 & 1 & 2 \\ 0 & 0 & 4 & 4 & 3 \end{bmatrix}$

5.18. $\begin{bmatrix} 2 & 3 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix}$

Bizonyítások

5.19. Bizonyítsuk be, hogy ha $c\mathbf{A} = \mathbf{O}$, akkor vagy $c = 0$, vagy $\mathbf{A} = \mathbf{O}$.

5.20. Az $S_i \leftrightarrow S_j$ sorművelethez tartozó elemi mátrixot jelölje $\mathbf{E}_{ij}$, a cS_i-hez tartozót $\mathbf{E}_i(c)$ és a $S_i + cS_j$ sorművelethez tartozót $\mathbf{E}_{ij}(c)$. Mutassuk meg, hogy $\mathbf{E}_{ij}^{-1} = \mathbf{E}_{ij}$, $\mathbf{E}_i(c)^{-1} = \mathbf{E}_i(\frac{1}{c})$ és $\mathbf{E}_{ij}(c)^{-1} = \mathbf{E}_{ij}(-c)$.

5.21. Mutassuk meg, hogy ha $\mathbf{A}$ fölcserélhető $\mathbf{B}$-vel és $\mathbf{B}$ invertálható, akkor $\mathbf{A}$ fölcserélhető $\mathbf{B}^{-1}$-gyel is.

Absztrakció

5.22. INVERTÁLHATÓ MŰVELET Legyen $\circledcirc$ egy H-n értelmezett kétváltozós művelet, azaz egy $H^2 \to H$ függvény. Fogalmazzuk meg, mit értünk azon, hogy $\circledcirc$ invertálható egy $R \subseteq H$ részhalmazán. Hogyan változik a definíció, ha a művelet kommutatív?

5.23. ELEM INVERZE Legyen $\circledcirc$ egy H-n értelmezett kétváltozós művelet.

1. Mit értünk azon, hogy $e \in H$ e művelet semleges eleme?
2. Mit értünk azon, hogy $b \in H$ az $a \in H$ elem inverze?

Műveletek speciális mátrixokkal

A gyakorlatban gyakran találkozunk olyan speciális mátrixokkal, melyekkel a műveletek egyszerűbben végezhetők el, és olyanokkal, melyek mátrixműveletek segítségével definiálhatók.

Diagonális mátrixok A mátrixműveletek definíciói alapján magától értetődő, hogy diagonális mátrixokkal hogyan végezhetők el a mátrixműveletek. Először egyszerű példákat mutatunk.

5.27. PÉLDA (MŰVELETEK DIAGONÁLIS MÁTRIXOKKAL). *Legyen* $\mathbf{A} = \operatorname{diag}(1,2,3)$, $\mathbf{B} = \operatorname{diag}(5,4,3)$. *Ellenőrizzük az alábbi számításokat, és fogalmazzunk meg általános összefüggéseket diagonális mátrixokkal végzett műveletekről.*

$$\mathbf{AB} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

$$\mathbf{A}^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

$$\mathbf{A}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}$$

$$\mathbf{A}^k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}^k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2^k & 0 \\ 0 & 0 & 3^k \end{bmatrix}, \text{ ahol } k \text{ egész szám.}$$

5.28. TÉTEL (MŰVELETEK DIAGONÁLIS MÁTRIXOKKAL). *Legyen* $\mathbf{A} = \operatorname{diag}(a_1, a_2, \dots, a_n)$, $\mathbf{B} = \operatorname{diag}(b_1, b_2, \dots, b_n)$, *és legyen k egész szám. Ekkor*

a) $\mathbf{AB} = \operatorname{diag}(a_1 b_1, a_2 b_2, \dots, a_n b_n)$,
b) $\mathbf{A}^k = \operatorname{diag}(a_1^k, a_2^k, \dots, a_n^k)$, *speciálisan*
c) $\mathbf{A}^{-1} = \operatorname{diag}(a_1^{-1}, a_2^{-1}, \dots, a_n^{-1})$.

Permutációs mátrixok és kígyók Könnyen kezelhetők a diagonális mátrixok sorainak permutációjával kapott mátrixok.

5.29. PÉLDA (SOROK PERMUTÁCIÓJA MÁTRIXSZORZÁSSAL). *Alkalmazzunk több sorcserét az egységmátrixon. Az így kapott mátrixszal balról való szorzás milyen hatással van a beszorzott mátrixra? Szemléltessük ezt* $\mathbf{I}_4$*-en az* $S_1 \leftrightarrow S_2$, $S_2 \leftrightarrow S_4$ *sorcserékkel.*

MEGOLDÁS. Legyen $\mathbf{P}$ az $\mathbf{I}_4$-ből a fent megadott két sorcserével kapott mátrix, és legyen $\mathbf{A}$ egy tetszőleges $4 \times n$-es mátrix. Ekkor

$$\mathbf{I}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \overset{S_1 \leftrightarrow S_2}{\Longrightarrow} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \overset{S_2 \leftrightarrow S_4}{\Longrightarrow} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} = \mathbf{P}$$

$$\mathbf{PA} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} = \begin{bmatrix} a_{21} & a_{22} \\ a_{41} & a_{42} \\ a_{31} & a_{32} \\ a_{11} & a_{12} \end{bmatrix}$$

Azt tapasztaljuk, hogy a $\mathbf{PA}$ az $\mathbf{A}$-ból a soroknak épp azzal a permutációjával kapható, amely permutációval $\mathbf{I}$-ből a $\mathbf{P}$-t kaptuk. □

5.30. DEFINÍCIÓ (PERMUTÁCIÓS MÁTRIX, KÍGYÓ). *A diagonális mátrixok sorainak permutációjával kapott mátrixot* kígyónak *nevezzük, speciálisan az egységmátrixból ugyanígy kapott mátrixot* permutációs mátrixnak *hívjuk.*

► Könnyen látható, hogy a permutációs mátrix olyan négyzetes mátrix, melynek minden sorában és minden oszlopában *pontosan* egy 1-es van, az összes többi elem 0. A kígyó olyan négyzetes mátrix, melynek minden sorában és minden oszlopában *legföljebb* egy nemnulla elem van.

► Minden kígyó megkapható egy diagonális mátrixból oszlopcserékkel is. Egy diagonális mátrixból akkor is kígyót kapunk, ha a sorok permutációja mellett az oszlopokat is permutáljuk.

5.31. PÉLDA (KÍGYÓK). *Az alábbi mátrixok mindegyike kígyó, az utolsó kettő egyúttal permutációs mátrix is:*

$$\begin{bmatrix} 0 & 5 & 0 \\ 0 & 0 & 9 \\ 3 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & \alpha & 0 & 0 \\ \gamma & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

5.32. TÉTEL (MŰVELETEK PERMUTÁCIÓS MÁTRIXOKKAL). *Bármely két azonos méretű permutációs mátrix szorzata és egy permutációs mátrix bármely egész kitevős hatványa permutációs mátrix. Permutációs mátrix inverze megegyezik a transzponáltjával, azaz ha* $\mathbf{P}$ *permutációs mátrix, akkor*

$$\mathbf{P}^{-1} = \mathbf{P}^T.$$

Bizonyítás. Legyen $\mathbf{A}$ és $\mathbf{B}$ két permutációs mátrix. Szorzatuk sorvektorai $\mathbf{a}_{i*}\mathbf{B}$ alakúak, ahol $\mathbf{a}_{i*}$ megegyezik valamelyik standard egységvektorral, pl. $\mathbf{a}_{i*} = \mathbf{e}_k$. Ekkor a szorzatvektornak csak az az eleme 1, amelyik oszlop $\mathbf{e}_k$-val megegyezik, és ilyen oszlop pontosan egy van. Tehát a szorzatmátrix minden sorában pontosan egy elem 1, a többi 0. Oszlopokra az állítás hasonlóan bizonyítható. A szorzatra vonatkozó állítás természetes következménye a pozitív egész kitevős hatványokra vonatkozó állítás. A negatív egész kitevőkre is igaz az állítás. Ennek igazolásához elég az inverzre belátni.

Tekintsük a $\mathbf{PP}^T$ szorzatot. A $(\mathbf{PP}^T)_{ii}$ elem a $\mathbf{P}_{i*}$ vektornak és a $(\mathbf{P}^T)_{*i} = \mathbf{P}_{i*}$ vektornak a szorzata, vagyis 1, míg

$$(\mathbf{PP}^T)_{ij} = (\mathbf{P})_{i*}(\mathbf{P}^T)_{*j} = (\mathbf{P})_{i*} \cdot (\mathbf{P})_{j*},$$

azaz a szorzat i-edik sorának j-edik eleme a $\mathbf{P}$ i-edik és j-edik sorvektorának skalárszorzata, ami 0, mivel két különböző sorban az 1-es különböző helyen van. □

5.33. példa (Permutációs mátrix inverze). *Az alábbi példa szemlélteti a tételben kimondott egyszerű állítást:*

$$\mathbf{PP}^T = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Háromszögmátrixok A Gauss-kiküszöbölés végrehajtásakor az együtthatómátrixot lépcsős alakra transzformáltuk, melyben a főátló alatt mindig csak nullák szerepelnek. Az ilyen mátrixok nem csak a Gauss-kiküszöbölésnél fontosak.

5.34. definíció (Háromszögmátrix). *Azokat a mátrixokat, melyek főátlója alatt csak* 0*-elemek szerepelnek* felső háromszögmátrixnak, *azokat, melyek főátlója fölött csak* 0*-elemek vannak* alsó háromszögmátrixnak *nevezzük. Ha egy háromszögmátrix főátlójában csupa* 1*-es áll,* egység háromszögmátrixról *beszélünk.*

A Gauss-kiküszöbölésnél kapott felső háromszögmátrixhoz hasonlóan azok az egyenletrendszerek is megoldhatók csak behelyettesítésekkel, amelyek együtthatómátrixa alsó háromszögmátrix. A különbség kizárólag annyi, hogy ekkor az első egyenlettel kezdjük, és az első változóval. Például az

$$\begin{aligned} x \phantom{{}+3y+2z} &= 3 \\ 2x + 3y \phantom{{}+2z} &= 3 \\ 2x + y + 2z &= 3 \end{aligned}$$

egyenletrendszer első egyenletéből $x = 3$, a másodikba való behelyettesítés után $y = -1$, végül a harmadikba való behelyettesítés után $z = -1$.[2]

[2] Az angol nyelvű lineáris algebra tankönyvek különbséget tesznek a felső és az alsó háromszögmátrixú egyenletrendszerek megoldása között. *Forward substitution*, illetve *backward substitution* a neve a behelyettesítésnek ha alsó, illetve ha felső háromszögmátrix az együttható mátrix. Ez arra utal, hogy a változókat előre vagy hátra haladva számoljuk ki. Mi nem fogjuk használni e finom különbségtételt.

5.35. TÉTEL (MŰVELETEK HÁROMSZÖGMÁTRIXOKKAL). *Felső háromszögmátrixok összege, szorzata, és invertálható felső háromszögmátrix inverze felső háromszögmátrix. Analóg tétel igaz az alsó háromszögmátrixokra is. Egy háromszögmátrix pontosan akkor invertálható, ha főátlóbeli elemeinek egyike sem zérus.*

A bizonyítást feladatként az Olvasóra hagyjuk.

Szimmetrikus és ferdén szimmetrikus mátrixok Gyakran használunk olyan mátrixokat, melyekben az elemek egyenlők vagy ellentettjei a főátlóra nézve szimmetrikusan elhelyezkedő párjuknak. E tulajdonság a transzponálttal könnyen kifejezhető.

5.36. DEFINÍCIÓ (SZIMMETRIKUS ÉS FERDÉN SZIMMETRIKUS MÁTRIXOK). *A négyzetes* $\mathbf{A}$ *mátrixot* szimmetrikusnak *nevezzük, ha* $\mathbf{A}^T = \mathbf{A}$, *és* ferdén szimmetrikusnak *nevezzük, ha* $\mathbf{A}^T = -\mathbf{A}$.

5.37. PÉLDA (SZIMMETRIKUS ÉS FERDÉN SZIMMETRIKUS MÁTRIXOK). *Az alábbi mátrixok közül az* $\mathbf{A}$ *szimmetrikus, a* $\mathbf{B}$ *ferdén szimmetrikus, a* $\mathbf{C}$ *egyik osztályba sem tartozik.*

$$\mathbf{A} = \begin{bmatrix} 5 & 6 & 1 \\ 6 & 2 & 0 \\ 1 & 0 & 3 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 1 & -2 \\ -1 & 0 & 3 \\ 2 & -3 & 0 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 9 & 9 \\ -9 & 2 & 9 \\ -9 & -9 & 3 \end{bmatrix}.$$

Ha $\mathbf{A}$ ferdén szimmetrikus, akkor minden elemére $a_{ij} = -a_{ji}$, azaz $i = j$ esetén $a_{ii} = -a_{ii}$. Ez csak $a_{ii} = 0$ esetén áll fönn, azaz a ferdén szimmetrikus mátrixok főátlójában csupa 0 áll.

5.38. ÁLLÍTÁS (MŰVELETEK (FERDÉN) SZIMMETRIKUS MÁTRIXOKKAL). *Szimmetrikus mátrixok összege, skalárszorosa, szorzata, inverze szimmetrikus. Hasonlóan ferdén szimmetrikus mátrixok összege, skalárszorosa, szorzata, inverze ferdén szimmetrikus.*

Az állítás bizonyítását feladatként az olvasóra hagyjuk.

5.39. TÉTEL (FELBONTÁS SZIMMETRIKUS ÉS FERDÉN SZIMMETRIKUS MÁTRIX ÖSSZEGÉRE). *Minden négyzetes mátrix előáll egy szimmetrikus és egy ferdén szimmetrikus mátrix összegeként, nevezetesen minden* $\mathbf{A}$ *négy-*

zetes mátrixra

$$\mathbf{A} = \underbrace{\frac{1}{2}\left(\mathbf{A} + \mathbf{A}^T\right)}_{\text{szimmetrikus}} + \underbrace{\frac{1}{2}\left(\mathbf{A} - \mathbf{A}^T\right)}_{\text{ferdén szimm.}}.$$

BIZONYÍTÁS. Ha egy mátrix szimmetrikus, konstansszorosa is, így elég megmutatni, hogy az $\mathbf{A} + \mathbf{A}^T$ mátrix szimmetrikus:

$$\left(\mathbf{A} + \mathbf{A}^T\right)^T = \mathbf{A}^T + \left(\mathbf{A}^T\right)^T = \mathbf{A}^T + \mathbf{A} = \mathbf{A} + \mathbf{A}^T$$

Hasonlóképp $\mathbf{A} - \mathbf{A}^T$ ferdén szimmetrikus, hiszen

$$\left(\mathbf{A} - \mathbf{A}^T\right)^T = \mathbf{A}^T - \left(\mathbf{A}^T\right)^T = \mathbf{A}^T - \mathbf{A} = -\left(\mathbf{A} - \mathbf{A}^T\right)$$

Végül a két mátrix összege valóban $\mathbf{A}$:

$$\frac{1}{2}\left(\mathbf{A} + \mathbf{A}^T\right) + \frac{1}{2}\left(\mathbf{A} - \mathbf{A}^T\right) = \frac{1}{2}\mathbf{A} + \frac{1}{2}\mathbf{A}^T + \frac{1}{2}\mathbf{A} - \frac{1}{2}\mathbf{A}^T = \mathbf{A}.$$ □

Fontos következményei lesznek az alábbi egyszerű állításnak.

5.40. TÉTEL ($\mathbf{A}^T\mathbf{A}$ ÉS $\mathbf{AA}^T$ SZIMMETRIKUS). *Az $\mathbf{A}^T\mathbf{A}$ és az $\mathbf{AA}^T$ mátrixok tetszőleges $\mathbf{A}$ mátrix esetén szimmetrikusak.*

BIZONYÍTÁS. $(\mathbf{AA}^T)^T = (\mathbf{A}^T)^T\mathbf{A}^T = \mathbf{AA}^T$. Az állítás másik fele ugyanígy bizonyítható. □

*Mátrix és diád összegének inverze** Összegmátrix inverzére – a valósok összegének inverzéhez hasonlóan – nincs egyszerű képlet, de speciális mátrixokra nagyon hasznos eredmények vannak. Ilyen a következő tétel is.

5.41. TÉTEL (SHERMAN–MORRISON-FORMULA). *Tegyük fel, hogy az $\mathbf{A} \in \mathbb{R}^{n\times n}$ mátrix invertálható, és $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ két olyan vektor, hogy $1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u} \neq 0$. Ekkor $\mathbf{A} + \mathbf{uv}^T$ invertálható, és*

$$\left(\mathbf{A} + \mathbf{uv}^T\right)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{uv}^T\mathbf{A}^{-1}}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}}.$$

BIZONYÍTÁS. Elég megmutatni, hogy

$$\left(\mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{uv}^T\mathbf{A}^{-1}}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}}\right)\left(\mathbf{A} + \mathbf{uv}^T\right) = \mathbf{I},$$

mert ez a formula igazolása mellett azt is bizonyítja, hogy $\mathbf{A} + \mathbf{uv}^T$

invertálható.

$$\begin{aligned}
&\left(\mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}}{1+\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}}\right)(\mathbf{A}+\mathbf{u}\mathbf{v}^T)\\
&\quad = \mathbf{A}^{-1}\mathbf{A} + \mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}\mathbf{A}}{1+\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T}{1+\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}}\\
&\quad = \mathbf{I} + \mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T - \frac{\mathbf{A}^{-1}\mathbf{u}1\mathbf{v}^T}{1+\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}} - \frac{\mathbf{A}^{-1}\mathbf{u}\left(\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}\right)\mathbf{v}^T}{1+\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}}\\
&\quad = \mathbf{I} + \mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T - \frac{\mathbf{A}^{-1}\mathbf{u}\left(1+\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}\right)\mathbf{v}^T}{1+\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}}\\
&\quad \stackrel{*}{=} \mathbf{I} + \mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T - \mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\\
&\quad = \mathbf{I}.
\end{aligned}$$

A $*$-gal jelzett egyenlőségnél azt használtuk ki, hogy 1×1-es mátrixszal való szorzás egybeesik a skalárral való szorzással, a skalár tényező pedig egy mátrixszorzatban áttehető más helyre, így az adott törtkifejezésben egyszerűsíthettünk vele. □

A Sherman – Morrison-formula sokhelyütt használható, mi itt egyet emelünk ki: megvizsgáljuk, hogy hogyan változik egy mátrix inverze, ha a mátrixnak csak egyetlen elemén változtatunk.

5.42. PÉLDA (INVERZ VÁLTOZÁSA). *Legyen* $\mathbf{A}$ *invertálható mátrix, és változtassunk meg az* a_{ij} *elemet* $a_{ij} + \varepsilon$*-ra. Fejezzük ki az így kapott mátrix inverzét* $\mathbf{A}^{-1}$ *segítségével.*

MEGOLDÁS. Első lépésként kifejezzük az új mátrixot $\mathbf{A}$-ból mátrixműveletekkel. Legyen $\mathbf{e}_i$ és $\mathbf{e}_j$ az i-edik és j-edik standard egységvektor. Ekkor a módosított mátrix

$$\mathbf{B} = \mathbf{A} + \varepsilon\mathbf{e}_i\mathbf{e}_j^T.$$

Erre alkalmazható a Sherman – Morrison-formula az $\mathbf{u} = \mathbf{e}_i$ és $\mathbf{v} = \varepsilon\mathbf{e}_j$ választással.

$$\begin{aligned}
\mathbf{B}^{-1} &= \left(\mathbf{A} + \varepsilon\mathbf{e}_i\mathbf{e}_j^T\right)^{-1}\\
&= \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{e}_i(\varepsilon\mathbf{e}_j)^T\mathbf{A}^{-1}}{1+\varepsilon\mathbf{e}_j^T\mathbf{A}^{-1}\mathbf{e}_i}\\
&= \mathbf{A}^{-1} - \varepsilon\frac{(\mathbf{A}^{-1})_{*i}(\mathbf{A}^{-1})_{j*}}{1+\varepsilon(\mathbf{A}^{-1})_{ji}}
\end{aligned}$$

□

5.43. PÉLDA (INVERZ VÁLTOZÁSA SZÁMPÉLDÁN). *Adva van egy* $\mathbf{A}$ *mátrix és annak inverze:*

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4\\ 2 & 0 & 0 & 3\\ 3 & 0 & 0 & 2\\ 4 & 3 & 2 & 1 \end{bmatrix} \qquad \mathbf{A}^{-1} = \begin{bmatrix} 0 & -2/5 & 3/5 & 0\\ -2/5 & 7/5 & -8/5 & 3/5\\ 3/5 & -8/5, & 7/5 & -2/5\\ 0 & 3/5 & -2/5 & 0 \end{bmatrix}.$$

Változtassuk meg a_{11} értékét 1-ről 11/10-re. Az így kapott mátrixot jelölje **B**. *Határozzuk meg inverzét!*

MEGOLDÁS. Az előző példa alkalmazásával

$$\mathbf{B}^{-1} = \mathbf{A}^{-1} - \frac{1}{10}\frac{(\mathbf{A}^{-1})_{*1}(\mathbf{A}^{-1})_{1*}}{1+\frac{1}{10}(\mathbf{A}^{-1})_{11}}$$

$$= \begin{bmatrix} 0 & -2/5 & 3/5 & 0 \\ -2/5 & 7/5 & -8/5 & 3/5 \\ 3/5 & -8/5 & 7/5 & -2/5 \\ 0 & 3/5 & -2/5 & 0 \end{bmatrix} - \frac{1}{10}\frac{\begin{bmatrix} 0 \\ -2/5 \\ 3/5 \\ 0 \end{bmatrix}\begin{bmatrix} 0 & -2/5 & 3/5 & 0 \end{bmatrix}}{1+\frac{1}{10}\cdot 0}$$

$$= \begin{bmatrix} 0 & -2/5 & 3/5 & 0 \\ -2/5 & 7/5 & -8/5 & 3/5 \\ 3/5 & -8/5 & 7/5 & -2/5 \\ 0 & 3/5 & -2/5 & 0 \end{bmatrix} - \frac{1}{10}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 4/25 & -6/25 & 0 \\ 0 & -6/25 & 9/25 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & -2/5 & 3/5 & 0 \\ -2/5 & 173/125 & -197/125 & 3/5 \\ 3/5 & -197/125 & 341/250 & -2/5 \\ 0 & 3/5 & -2/5 & 0 \end{bmatrix}.$$

Tizedestörtekkel számolva:

$$\mathbf{A}^{-1} = \begin{bmatrix} 0.0 & -0.4 & 0.6 & 0.0 \\ -0.4 & 1.4 & -1.6 & 0.6 \\ 0.6 & -1.6 & 1.4 & -0.4 \\ 0.0 & 0.6 & -0.4 & 0.0 \end{bmatrix},$$

$$\mathbf{B}^{-1} = \begin{bmatrix} 0.000 & -0.400 & 0.600 & 0.000 \\ -0.400 & 1.384 & -1.576 & 0.600 \\ 0.600 & -1.576 & 1.364 & -0.400 \\ 0.000 & 0.600 & -0.400 & 0.000 \end{bmatrix}.$$ □

*Gyorsszorzás** Két 2×2-es mátrix szokásos módon való összeszorzásához 8 szorzásra és 4 összeadásra van szükség. Strassen 1969-ben egy olyan módszert talált, mellyel e mátrixszorzást 7 szorzással is el lehet végezni, igaz azon az áron, hogy az összeadások száma 16-ra nő.

5.2 (STRASSEN-FORMULÁK). *Legyen* $\mathbf{A}$, $\mathbf{B}$ *és* $\mathbf{C}$ *is* 2×2*-es. A* $\mathbf{C} = \mathbf{AB}$ *szorzás elvégezhető a következő formulákkal:*

$$
\begin{aligned}
d_1 &= (a_{11} + a_{22})(b_{11} + b_{22}) & c_{11} &= d_1 + d_4 - d_5 + d_7 \\
d_2 &= (a_{21} + a_{22})b_{11} & c_{21} &= d_2 + d_4 \\
d_3 &= a_{11}(b_{12} - b_{22}) & c_{12} &= d_3 + d_5 \\
d_4 &= a_{22}(-b_{11} + b_{21}) & c_{22} &= d_1 + d_3 - d_2 + d_6 \\
d_5 &= (a_{11} + a_{12})b_{22} \\
d_6 &= (-a_{11} + a_{21})(b_{11} + b_{12}) \\
d_7 &= (a_{12} - a_{22})(b_{21} + b_{22})
\end{aligned}
$$

Ha olyan számítógépet használunk, melyen a számok szorzása sokkal több időt igényel, mint az összeadása, akkor már ez is gyorsíthatja a műveletet. A módszer azonban kiterjeszthető tetszőleges méretű négyzetes mátrixokra is, és elegendően nagy n-ekre a műveletek összszáma is csökken. A standard mátrixszorzás műveletigénye $2n^3 - n^2$ (ebből n^3 szorzás és $n^3 - n^2$ összeadás – gondoljunk utána!), a Strassen-formulákkal való szorzásé pedig $n = 2^k$ esetén legföljebb $7 \cdot 7^k - 6 \cdot 4^k$). Ez $n = 2^{10}$ esetén már kevesebb műveletet ad. Az általánosítás lényege, hogy a Strassen-formulák 2×2-es blokkmátrixokra is használhatók, mert a szorzás kommutativitását nem használják, így ha $M(n)$ jelöli két $n \times n$-es mátrix összeszorzásához szükséges szorzások, és $S(n)$ a szükséges összeadások számát, akkor $M(2n) \leq 7M(n)$ és $S(2n) \leq 18n^2 + 7S(n)$. Az $M(1) = 1$, $S(1) = 0$ kezdeti feltételeket is használva megmutatható, hogy $M(2^k) \leq 7^k$, $S(2^k) \leq 6(7^k - 4^k)$. E képletekből a felső egészrész jelét használva és a $k = \lceil \log_2 n \rceil$ jelöléssel az műveletek összámára a $cn^{\log_2 7} \leq cn^{2.81}$ felső becslést kapjuk, ami a $2n^3 - n^2$ értéknél jobb, függetlenül a c konstans konkrét értékétől. Mivel a két összeszorzandó mátrix mindegyikének mind az n^2 elemét használni kell, ezért a szükséges műveletek számának alsó becslése cn^2. A $cn^{2.81}$ felső becslés mára $cn^{2.376}$-ra lett javítva (Coppersmith és Winograd, 1990), de az a sejtés, hogy a kitevő 2-re, de legalább $2 + \varepsilon$-ra lenyomható, ahol ε tetszőlegesen kis pozitív szám.

A módszer gyengéje numerikus instabilitása, így a gyakorlatban csak bizonyos mátrixokra érdemes használni, például nagyméretű egészelemű mátrixokra tetszőleges pontosságú aritmetika használata esetén.

Feladatok

Döntsük el, igazak-e az alábbi állítások? Válaszunkat indokoljuk!

5.24. Szimmetrikus mátrixok összege és skalárszorosa is szimmetrikus, így szimmetrikus mátrixok tetszőleges lineáris kombinációja is szimmetrikus.

5.25. Ferdén szimmetrikus mátrixok összege és skalárszorosa is ferdén szimmetrikus, így ferdén szimmetrikus mátrixok tetszőleges lineáris kombinációja is ferdén szimmetrikus.

5.26. Számítsuk ki az alábbi mátrixok inverzeit, négyzetét és köbét!

$$\begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 4 \\ 3 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 2 \\ 0 & 4 & 0 \\ 3 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 3 & 0 \\ 4 & 0 & 0 & 0 \end{bmatrix}.$$

5.27. Hogyan oldanánk meg a következő egyenletrendszert a lehető legkevesebb lépésben?

$$\begin{aligned} x + 4y + 3z + 5w &= 3 \\ 6x + 3y \phantom{{}+3z+5w} &= 3 \\ 2x + 3y + 2z \phantom{{}+5w} &= 3 \\ 2x + 4y + 3z + 5w &= 4 \end{aligned}$$

Bizonyítások

5.28. Mutassuk meg, hogy minden permutációs mátrix oszlopcserékkel is megkapható az egységmátrixból, és hogy permutációs mátrixszal jobbról való szorzás a beszorzott mátrix oszlopain ugyanazt a permutációt hajtja végre, mint amellyel a permutációs mátrix az egységmátrixból megkapható.

5.29. Bizonyítsuk be, hogy bármely két azonos méretű kígyó szorzata és egy kígyó bármely pozitív egész kitevős hatványa kígyó.

5.30.• Mutassuk meg, hogy egy **K** kígyó pontosan akkor invertálható, ha minden sorában pontosan egy elem nem 0, és ekkor inverze megkapható úgy, hogy minden nemnulla elem helyébe annak reciprokát írjuk, majd az így kapott mátrixot transzponáljuk.

5.31.* GYORSINVERTÁLÁS Legyen $\mathbf{B} = \mathbf{A}^{-1}$, mindketten 2×2-es mátrixok. Mutassuk meg, hogy az alábbi eljárással definiált mátrixinvertálás segítségével $n \times n$-es mátrixokra olyan algoritmus készíthető, melynek műveletigénye legföljebb $cn^{2.81}$.

$$\begin{aligned} c_1 &= a_{11}^{-1} & b_{12} &= c_3 c_6 \\ c_2 &= a_{21} c_1 & b_{21} &= c_6 c_2 \\ c_3 &= c_1 a_{12} & c_7 &= c_3 b_{21} \\ c_4 &= a_{21} c_3 & b_{11} &= c_1 - c_7 \\ c_5 &= c_4 - a_{22} & b_{22} &= -c_6 \\ c_6 &= c_5^{-1} & & \end{aligned}$$

Az LU-felbontás

Mátrixok szorzatalakba írásával (faktorizációjával) már találkoztunk, amikor a Gauss – Jordan-kiküszöbölést használva bizonyítottuk, hogy minden invertálható mátrix előáll elemi mátrixok szorzataként. E szakaszban a Gauss-kiküszöbölésből indulva megmutatjuk, hogy egy mátrix felbontható egy alsó és egy felső háromszögmátrix szorzatára. E mátrixfaktorizáció lineáris algebrai feladatok számítógépes megoldásának igen gyakran használt eszköze, sokszor előnyösebb, mint maga a Gauss-kiküszöbölés.

5.44. PÉLDA (GAUSS-KIKÜSZÖBÖLÉS MÁTRIXSZORZÁSSAL). *Elemi sorműveletekkel hozzuk az*

$$\mathbf{A} = \begin{bmatrix} 4 & 1 & 2 \\ 2 & 4 & 1 \\ 1 & 2 & 4 \end{bmatrix}$$

mátrixot lépcsős alakra (azaz felső háromszögmátrix alakra), amit jelöljön $\mathbf{U}$. *Az elemi sorműveleteket helyettesítsük mátrixokkal való szorzással. Végül az így kapott mátrixszorzatos alak segítségével állítsuk elő* $\mathbf{A}$*-t az* $\mathbf{U}$ *és egy másik mátrix szorzataként.*

MEGOLDÁS. Oszloponként haladva végezzük el a Gauss-kiküszöbölést. Minden elemi sorművelet mellett (zárójelben) megadjuk a hozzá tartozó elemi mátrixot:

$$\mathbf{A} = \begin{bmatrix} 4 & 1 & 2 \\ 2 & 4 & 1 \\ 1 & 2 & 4 \end{bmatrix} \overset{S_2 - 1/2 S_1}{\Longrightarrow} \qquad \left(\mathbf{E}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right)$$

$$\mathbf{E}_1\mathbf{A} = \begin{bmatrix} 4 & 1 & 2 \\ 0 & 7/2 & 0 \\ 1 & 2 & 4 \end{bmatrix} \overset{S_3 - 1/4 S_1}{\Longrightarrow} \qquad \left(\mathbf{E}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1/4 & 0 & 1 \end{bmatrix} \right)$$

$$\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \begin{bmatrix} 4 & 1 & 2 \\ 0 & 7/2 & 0 \\ 0 & 7/4 & 7/2 \end{bmatrix} \overset{S_3 - 1/2 S_2}{\Longrightarrow} \qquad \left(\mathbf{E}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1/2 & 1 \end{bmatrix} \right)$$

$$\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \begin{bmatrix} 4 & 1 & 2 \\ 0 & 7/2 & 0 \\ 0 & 0 & 7/2 \end{bmatrix} = \mathbf{U}.$$

Tehát $\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \mathbf{U}$, amiből az $(\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1)^{-1}$ mátrixszal való beszorzás után $\mathbf{A} = (\mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\mathbf{E}_3^{-1})\mathbf{U}$. Kiszámoljuk az elemi mátrixok inverzeinek szorzatát, azaz az $\mathbf{L} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\mathbf{E}_3^{-1}$ mátrixot. Fölhasználjuk a 204. oldalon mondottakat, miszerint $S_i + cS_j$ mátrixának inverze $S_i - cS_j$ mátrixával egyenlő:

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1/2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/4 & 1/2 & 1 \end{bmatrix}.$$

Meglepő módon ezeknek az elemi mátrixoknak a szorzata a főátló alatti számok átmásolásával megkapható. Az eredmény egy alsó egység háromszögmátrix. □

5.45. DEFINÍCIÓ (LU-FELBONTÁS). *Azt mondjuk, hogy a négyzetes* $\mathbf{A}$ *mátrix egy* $\mathbf{A} = \mathbf{LU}$ *alakú tényezőkre bontása* LU-felbontás *(*LU-faktorizáció *vagy* LU-dekompozíció*), ha* $\mathbf{L}$ *alsó,* $\mathbf{U}$ *felső háromszögmátrix.*

Az 5.44. példában konstruált felbontás LU-felbontás: [3]

$$\begin{bmatrix} 4 & 1 & 2 \\ 2 & 4 & 1 \\ 1 & 2 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/4 & 1/2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 1 & 2 \\ 0 & 7/2 & 0 \\ 0 & 0 & 7/2 \end{bmatrix}. \tag{5.1}$$

[3] Az *LU-felbontásban* az L és U betűk az *alsó* és *felső* jelentésű angol *lower* és *upper* szavak kezdőbetűi.

Az LU-felbontás használata egyenletrendszer megoldására Ha ismerjük az $n \times n$-es invertálható $\mathbf{A}$ mátrix LU-felbontását, akkor az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszer könnyen megoldható. Az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszer megoldása az $\mathbf{Ly} = \mathbf{b}$, $\mathbf{Ux} = \mathbf{y}$ egyenletrendszerek megoldásával ekvivalens. Ha ugyanis $\mathbf{x}$ megoldása $\mathbf{Ax} = \mathbf{b}$ egyenletrendszernek, akkor $\mathbf{LUx} = \mathbf{b}$, és $\mathbf{y} = \mathbf{Ux}$ jelöléssel $\mathbf{Ly} = \mathbf{b}$. Másrészt, ha $\mathbf{y}$ megoldása az $\mathbf{Ly} = \mathbf{b}$ egyenletrendszernek, és $\mathbf{x}$ az $\mathbf{Ux} = \mathbf{y}$ egyenletrendszernek, akkor $\mathbf{y}$-t behelyettesítve $\mathbf{L}(\mathbf{Ux}) = \mathbf{b}$, azaz $\mathbf{Ax} = \mathbf{b}$. Tömören:

$$\mathbf{Ax} = \mathbf{b} \text{ megoldható} \iff \mathbf{Ly} = \mathbf{b},\ \mathbf{Ux} = \mathbf{y} \text{ megoldható.}$$

Mivel $\mathbf{L}$ és $\mathbf{U}$ is háromszögmátrix, ezért az $\mathbf{Ly} = \mathbf{b}$, $\mathbf{Ux} = \mathbf{y}$ egyenletrendszerek egyszerű behelyettesítésekkel megoldhatók. A megoldás minkét esetben egyértelmű, mert $\mathbf{L}$ és $\mathbf{U}$ rangja is n.

5.46. PÉLDA (EGYENLETRENDSZER MEGOLDÁSA LU-FELBONTÁSSAL). *Oldjuk meg a következő egyenletrendszert!*

$$\begin{aligned} 4x_1 + x_2 + 2x_3 &= 8 \\ 2x_1 + 4x_2 + x_3 &= 4 \\ x_1 + 2x_2 + 4x_3 &= 9 \end{aligned}$$

MEGOLDÁS. Mivel ismerjük az együtthatómátrix LU-felbontását – az épp az (5.1)-beli felbontás –, ezért ezt használjuk, és először megoldjuk az $\mathbf{Ly} = \mathbf{b}$ egyenletrendszert:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/4 & 1/2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \\ 9 \end{bmatrix}$$

Ebből $y_1 = 8$, ezt a második egyenletbe helyettesítve kapjuk, hogy $y_2 = 0$, majd ezeket a harmadikba helyettesítve kapjuk, hogy $y_3 = 7$. Ezután megoldjuk az $\mathbf{Ux} = \mathbf{y}$ egyenletrendszert, aminek alakja

$$\begin{bmatrix} 4 & 1 & 2 \\ 0 & 7/2 & 0 \\ 0 & 0 & 7/2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 8 \\ 0 \\ 7 \end{bmatrix}.$$

Ismét egyszerű behelyettesítésekkel kapjuk, hogy $x_3 = 2$, $x_2 = 0$ és $x_1 = 1$. A megoldás $\mathbf{x} = (1, 0, 2)$. □

Mátrix invertálása LU-felbontással Mátrix invertálásához elég megoldanunk az $\mathbf{AX} = \mathbf{I}$ egyenletrendszert. Ha $\mathbf{A} = \mathbf{LU}$ egy LU-felbontása $\mathbf{A}$-nak, akkor az $\mathbf{LUX} = \mathbf{I}$ megoldása a vele ekvivalens két mátrixegyenlet megoldásával megkapható:

$$\mathbf{AX} = \mathbf{I} \iff \mathbf{LY} = \mathbf{I},\ \mathbf{UX} = \mathbf{Y}.$$

E két utóbbi egyenletrendszer viszont megoldható kizárólag helyettesítésekkel is!

5.47. PÉLDA (MÁTRIX INVERTÁLÁSA LU-FELBONTÁSSAL). *Invertáljuk az 5.44. példában megadott* $\mathbf{A}$ *mátrixot!*

MEGOLDÁS. Megadtuk az adott mátrix LU-felbontását az (5.1) egyenletben. Ezt használva először megoldjuk az $\mathbf{LY} = \mathbf{I}$ mátrixegyenletet:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/4 & 1/2 & 1 \end{bmatrix} \begin{bmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Az $\mathbf{L}$ első sorával való szorzásból: $[y_{11}\ y_{12}\ y_{13}] = [1\ 0\ 0]$. A második sorral való szorzásból $\frac{1}{2}[y_{11}\ y_{12}\ y_{13}] + [y_{21}\ y_{22}\ y_{23}] = [0\ 1\ 0]$. Behelyettesítés után $[y_{21}\ y_{22}\ y_{23}] = [-\frac{1}{2}\ 1\ 0]$. Végül a harmadik sorral való szorzásból:

$$\frac{1}{4}\begin{bmatrix} y_{11} & y_{12} & y_{13} \end{bmatrix} + \frac{1}{2}\begin{bmatrix} y_{21} & y_{22} & y_{23} \end{bmatrix} + \begin{bmatrix} y_{31} & y_{32} & y_{33} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix},$$

amiből behelyettesítés után kifejezve $\mathbf{Y}$ harmadik sorát kapjuk, hogy $[y_{31}\ y_{32}\ y_{33}] = [0\ -\frac{1}{2}\ 1]$. Azaz

$$\mathbf{Y} = \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & -1/2 & 1 \end{bmatrix}.$$

Ezután ugyanígy, egyszerű helyettesítésekkel megoldható az $\mathbf{UX} = \mathbf{Y}$, azaz a

$$\begin{bmatrix} 4 & 1 & 2 \\ 0 & 7/2 & 0 \\ 0 & 0 & 7/2 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & -1/2 & 1 \end{bmatrix}$$

mátrixegyenlet is, melynek megoldása

$$\mathbf{X} = \frac{1}{7} \begin{bmatrix} 2 & 0 & -1 \\ -1 & 2 & 0 \\ 0 & -1 & 2 \end{bmatrix}.$$

□

Az LU-felbontás kiszámítása Az 5.44. példában követett eljárás egyszerűen általánosítható tetszőleges méretű négyzetes mátrixra.

5.48. ALGORITMUS (EGY LU-FELBONTÁS ELŐÁLLÍTÁSA). *Tegyük fel, hogy az $n \times n$-es $\mathbf{A}$ mátrix kizárólag csak a hozzáadás elemi sorműveletével felső háromszög alakra hozható, azaz a kiküszöbölés során a főátlóba sosem kerül 0. Ekkor a következő lépésekkel állítsunk elő egy $\mathbf{U}$ és egy $\mathbf{L}$ mátrixot.*

1. *A Gauss-kiküszöbölést az első oszloppal kezdjük, és az első sor konstansszorosainak kivonásával, azaz az $S_2 - l_{21}S_1$, $S_3 - l_{31}S_1, \dots S_n - l_{n1}S_1$ sorműveletekkel elimináljuk az első oszlop elemeit (természetesen $l_{k1} = a_{k1}/a_{11}$).*
2. *Folytassuk az eliminációt a második oszloppal, azaz hajtsuk végre az $S_3 - l_{32}S_2, \dots S_n - l_{n2}S_2$ sorműveleteket, majd elimináljuk sorban a többi oszlop főátló alatti elemeit is.*
3. *Az eredményül kapott lépcsős alak lesz $\mathbf{U}$.*
4. *A kiküszöbölés konstans l_{ij} elemeit írjuk egy mátrix index szerinti helyébe, melynek főátlójába 1-eket, fölé 0-kat írunk. Ez lesz az $\mathbf{L}$ mátrix, azaz*

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ l_{21} & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \dots & 1 \end{bmatrix}. \tag{5.2}$$

Ismét végigszámoljuk az 5.44. példabeli mátrix felbontását. Először írjunk le egy egységmátrixot, de a főátló alatti helyeket üresen hagyva, ebből lesz $\mathbf{L}$. Írjuk mellé az $\mathbf{A}$ mátrixot, és amikor elvégzünk egy $S_i - l_{ji}S_j$ sorműveletet rajta, akkor az l_{ji} értéket bejegyezzük az $\mathbf{L}$ mátrix j-edik sorának i-edik oszlopába. Az alábbi számítások bal hasábjában látjuk a fentiek szerinti lépéseket.

$$\begin{bmatrix} 1 & 0 & 0 \\ & 1 & 0 \\ & & 1 \end{bmatrix} \begin{bmatrix} 4 & 1 & 2 \\ 2 & 4 & 1 \\ 1 & 2 & 4 \end{bmatrix} \qquad \begin{bmatrix} 4.00 & 1.00 & 2.00 \\ 2.00 & 4.00 & 1.00 \\ 1.00 & 2.00 & 4.00 \end{bmatrix}$$

$$\Downarrow \qquad\qquad \Downarrow$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ & & 1 \end{bmatrix} \begin{bmatrix} 4 & 1 & 2 \\ 0 & 7/2 & 0 \\ 1 & 2 & 4 \end{bmatrix} \qquad \begin{bmatrix} 4.00 & 1.00 & 2.00 \\ 0.50 & 3.50 & 0.00 \\ 1.00 & 2.00 & 4.00 \end{bmatrix}$$

$$\Downarrow \qquad\qquad \Downarrow$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/4 & & 1 \end{bmatrix} \begin{bmatrix} 4 & 1 & 2 \\ 0 & 7/2 & 0 \\ 0 & 7/4 & 7/2 \end{bmatrix} \qquad \begin{bmatrix} 4.00 & 1.00 & 2.00 \\ 0.50 & 3.50 & 0.00 \\ 0.25 & 1.75 & 3.50 \end{bmatrix}$$

$$\Downarrow \qquad\qquad \Downarrow$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/4 & 1/2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 1 & 2 \\ 0 & 7/2 & 0 \\ 0 & 0 & 7/2 \end{bmatrix} \qquad \begin{bmatrix} 4.00 & 1.00 & 2.00 \\ 0.50 & 3.50 & 0.00 \\ 0.25 & 0.50 & 3.50 \end{bmatrix}$$

Vegyük észre, hogy az **A** mátrixon folytatott elemi átalakítások eredménye és az **L** már kiszámolt elemei egyetlen mátrixban is „elférnek”, ugyanis **L**-ben épp akkor és oda kerül egy elem, amikor és ahova **A**-ban 0. Ezt a számítógépprogramok kihasználják, ha igen nagy méretű **A** mátrixot kell felbontani, és az **L** és **U** mátrixot is az **A** helyében konstruálják meg. A fenti számítások jobb hasábjában ezt a számítógépes technikát alkalmazzuk. Színes háttérrel jelöljük az **L**-beli elemeket. A számítógépes jelleg erősítésére tizedestört alakot használunk.

Igazolnunk kell még, hogy a fenti algoritmus valóban LU-felbontást ad.

5.49. TÉTEL (AZ LU-FELBONTÁS LÉTEZÉSE ÉS EGYÉRTELMŰSÉGE). *Legyen* **A** *invertálható* $n \times n$*-es mátrix.*

1. *Ha létezik* **A***-nak LU-felbontása, akkor ezek egyike megkapható az 5.48. algoritmussal.*
2. *Az* **A***-nak olyan LU-felbontása, melyben* **L** *főátlójában csak 1-esek vannak egyértelmű.*

BIZONYÍTÁS. *(a)* Először belátjuk, hogy az 5.48. algoritmusban megkonstruált **L** és **U** mátrixok LU-felbontást adnak. Ehhez elég belátni, hogy $\mathbf{A} = \mathbf{LU}$. Az algoritmus szerint $S_i - l_{ij}S_j$ $(i > j)$ alakú elemi műveletekkel transzformáljuk **A**-t **U**-ba. E transzformációk inverze $S_i + l_{ij}S_j$, melynek mátrixát jelölje $\mathbf{L}_{ij}$. Ez felírható $\mathbf{I} + l_{ij}\mathbf{e}_i\mathbf{e}_j^T$ alakban, azaz

$$\mathbf{L}_{ij} = \begin{bmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & l_{ij} & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix} = \mathbf{I} + l_{ij} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} [0 \dots 1 \dots 0 \dots 0].$$

Mivel $\mathbf{A} = \mathbf{L}_{21}\mathbf{L}_{31}\dots\mathbf{L}_{n1}\mathbf{L}_{32}\dots\mathbf{L}_{n2}\dots\mathbf{L}_{n-1,n}$, ezért az $\mathbf{L}_{ij} = \mathbf{I} + l_{ij}\mathbf{e}_i\mathbf{e}_j^T$ helyettesítések után csak azt kell belátni, hogy az (5.2)-beli $\mathbf{L}$ mátrixra

$$\mathbf{L} = (\mathbf{I} + l_{21}\mathbf{e}_2\mathbf{e}_1^T)(\mathbf{I} + l_{31}\mathbf{e}_3\mathbf{e}_1^T)\dots(\mathbf{I} + l_{n1}\mathbf{e}_n\mathbf{e}_1^T)\dots(\mathbf{I} + l_{n-1,n}\mathbf{e}_{n-1}\mathbf{e}_n^T).$$

Azt kell bizonyítanunk, hogy

$$\mathbf{L} = \mathbf{I} + l_{21}\mathbf{e}_2\mathbf{e}_1^T + l_{31}\mathbf{e}_3\mathbf{e}_1^T + \dots + l_{n1}\mathbf{e}_n\mathbf{e}_1^T + \dots + l_{n-1,n}\mathbf{e}_{n-1}\mathbf{e}_n^T,$$

vagyis elég belátni, hogy a fenti szorzatban a zárójelek felbontása után keletkező

$$\left(l_{ij}\mathbf{e}_i\mathbf{e}_j^T\right)(l_{st}\mathbf{e}_s\mathbf{e}_t^T) = l_{ij}l_{st}\mathbf{e}_i\mathbf{e}_j^T\mathbf{e}_s\mathbf{e}_t^T$$

alakú szorzatok mindegyike nullmátrix. Egyrészt $j < i$, $t < s$, másrészt mivel az elimináció oszloponként balról jobbra, oszlopon belül fentről lefelé haladt, ezért vagy $j < t$ vagy $j = t$ és akkor $i < s$. E feltételek esetén viszont $j \neq s$, tehát $\mathbf{e}_j^T\mathbf{e}_s = 0$, vagyis $\mathbf{e}_i\mathbf{e}_j^T\mathbf{e}_s\mathbf{e}_t^T = \mathbf{O}$, amit bizonyítani akartunk.

(b) Tegyük fel, hogy létezik $\mathbf{A}$-nak két LU-felbontása is, azaz $\mathbf{A} = \mathbf{L}_1\mathbf{U}_1 = \mathbf{L}_2\mathbf{U}_2$. Mivel $\mathbf{A}$ invertálható, ezért a **??** tétel szerint $\mathbf{L}_1$, $\mathbf{U}_1$, $\mathbf{L}_2$ és $\mathbf{U}_2$ is. Balról $\mathbf{L}_1$, jobbról $\mathbf{U}_2$ inverzével szorozva kapjuk, hogy

$$\mathbf{U}_1\mathbf{U}_2^{-1} = \mathbf{L}_1^{-1}\mathbf{L}_2.$$

A bal oldalon két felső háromszögmátrix szorzataként egy felső háromszögmátrix van, míg a jobb oldalon két alsó háromszögmátrix szorzata, ami alsó háromszögmátrix (**??** feladat). Ráadásul a jobb oldal egység főátlójú (**??** feladat). Ez csak akkor állhat fönn, ha $\mathbf{U}_1\mathbf{U}_2^{-1} = \mathbf{L}_1^{-1}\mathbf{L}_2 = \mathbf{I}$, azaz ha $\mathbf{L}_1 = \mathbf{L}_2$ és $\mathbf{U}_1 = \mathbf{U}_2$. □

PLU-felbontás Világos, hogy az

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

mátrixnak nincs LU-felbontása, hisz bármely LU-felbontás bal felső elemére $(\mathbf{LU})_{11} = l_{11}u_{11} \neq 0$, ugyanakkor $a_{11} = 0$. Az első és második sorok cseréje után kapott

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

mátrixnak hasonló okok miatt ugyancsak nincs LU-felbontása: ez levezethető az $a_{22} = 0$ összefüggésből (ld. 5.47. feladat). A második és harmadik sorok felcserélése után viszont már van LU-felbontás. Jelölje $\mathbf{P}$ a fenti sorcseréket megvalósító permutációs mátrixot. Ekkor

$$\mathbf{PA} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

ami egy LU-felbontás. Így a $\mathbf{PA} = \mathbf{LU}$ alakra jutottunk, amiből a $\mathbf{P}^{-1} = \mathbf{P}^T$ mátrixszal való szorzás után az $\mathbf{A} = \mathbf{P}^T\mathbf{LU}$ felbontást kapjuk.

Az LU-felbontást csak négyzetes mátrixokra definiáltuk, az egyenletrendszer megoldásában és a mátrix invertálásában való fontos szerepére és egy unicitástétel kimondhatóságára gondolva. Most több szempontból is általánosabb definíciót adunk, olyat, mely tetszőleges mátrixokra egy egzisztenciatétel kimondását is lehetővé teszi.

5.50. DEFINÍCIÓ (PLU-FELBONTÁS). *Egy tetszőleges $m \times n$-es $\mathbf{A}$ mátrixnak egy permutációs, egy négyzetes alsó háromszög- és egy $m \times n$-es felső háromszögmátrixra való bontását* PLU-felbontásnak *nevezzük.*

Az 5.48. algoritmus kis változtatással alkalmassá tehető e felbontás elvégzésére is. A módosított algoritmus ráadásul a négyzetes *szinguláris mátrixokra is* használható lesz.

1. Az elemi sorműveletek közben sorcserékre is szükség lehet, ha a főátlóban 0, de alatta valahol nem 0 áll. Az a sorcsere, ami ilyenkor elvégzendő, elvégezhető a kiküszöbölési eljárás előtt is (ezt az állítást nem igazoljuk). E sorcserékhez tartozó elemi mátrixok szorzata egy $\mathbf{P}$ permutációs mátrix. Így az algoritmust nem az $\mathbf{A}$-n, hanem $\mathbf{PA}$-n hajtjuk végre. Eredményül egy

$$\mathbf{PA} = \mathbf{LU}$$

felbontást kapunk, mely a $\mathbf{P}^T = \mathbf{P}^{-1}$ mátrixszal való beszorzás után a kívánalmaknak megfelelő $\mathbf{A} = \mathbf{P}^T\mathbf{LU}$ felbontást adja.
2. Ha sorcserékkel sem lehet elérni, hogy a főátlóban ne legyen zérus, akkor az algoritmusban tovább lépünk a következő oszlopra.

3. Nem kell lépcsős alakra hozni **A**-t, a felső háromszögmátrixszá való alakítás kevesebbet kér. Ez azt is jelenti, hogy itt a főelemet mindig a főátlóról választjuk, akkor is, ha a lépcsős alakra hozásnál tudnánk fölötte is főelemet választani.
4. A gyakorlatban – a hatékony számítógépprogramok – részleges főelemkiválasztással dolgoznak, így olyankor is alkalmaznak sorcseréket, ha a főátlóban nem zérus van.

5.51. PÉLDA (PLU-FELBONTÁS ELŐÁLLÍTÁSA). *Keressük meg az*

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 & 2 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 4 & 1 \end{bmatrix}$$

*mátrix egy PLU-felbontását! Az **A** felső háromszög-alakra hozása közben alkalmazzunk részleges főelemkiválasztást! Ügyeljünk arra, hogy főelemet itt csak a főátlóról válasszunk!*

MEGOLDÁS. Az egyszerűség kedvéért kis többletmunkát vállalva először csak azért hajtjuk végre a felső háromszög-alakra hozás lépéseit, hogy megállapítsuk, melyik lépésben melyik sorban lesz a pivotelem. Ezután előállítjuk e sorpermutációnak megfelelő **P** permutációs mátrixot, és végrehajtjuk az LU-felbontást a **PA** mátrixon.

$$\mathbf{A} \overset{\substack{S_1 \leftrightarrow S_2 \\ S_3 - S_1 \\ S_4 - S_1}}{\Longrightarrow} \begin{bmatrix} 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 2 & 4 & 0 \end{bmatrix} \overset{\substack{S_3 \leftrightarrow S_4 \\ S_4 - 1/2 S_3}}{\Longrightarrow} \begin{bmatrix} 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 2 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

A két sorcsere alapján a permutációs mátrix:

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Ezután elvégezve a sorműveleteket a megpermutált sorú **PA** mátrixon és az 5.48. algoritmus szerint megkreálva az **L** mátrixot kapjuk, hogy

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & \frac{1}{2} & 1 \end{bmatrix} \quad \mathbf{U} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 2 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

P transzponáltja önmaga, így a fenti jelölésekkel $\mathbf{A} = \mathbf{PLU}$ egy PLU-felbontás. □

Az LU-felbontás a gyakorlatban Az LU-felbontás műveletigénye megegyezik a Gauss-kiküszöbölésével, azaz nagyságrendileg $2n^3/3$ (utóbbiban a kiküszöbölést a jobb oldallal is meg kell csinálni, az LU-felbontásnál viszont az alsó háromszögmátrixhoz tartozó egyenletrendszert is meg kell oldani: mindkettő $n(n-1)/2$ összeadás/kivonás és ugyanennyi szorzás/osztás). Az LU-felbontásnak viszont több olyan előnyös tulajdonsága van, ami miatt használata meghatározó az egyenletrendszerek megoldásában és amellett több más feladatban is. Néhány a legfontosabbak közül:

1. Mivel az egyenletrendszer együtthatómátrixának LU-felbontásához nincs szükség a jobb oldalra, ezért használható olyan esetekben, amikor a jobb oldal még nem ismeretes, vagy több különböző jobb oldallal is dolgozni kell.

2. Az LU-felbontás ismeretében több mátrixokkal kapcsolatos számítás gyorsabban elvégezhető mint egyébként, pl. ilyen a mátrix inverzének, vagy a később tanulandó determinánsának meghatározása.

3. Korábban említettük, hogy az LU-felbontás igen memóriatakarékos, ráadásul vannak olyan speciális mátrixosztályok (pl. a szalagmátrixok), melyekre létezik a kiküszöbölésnél gyorsabb algoritmus az LU-felbontásra.

Feladatok

Adjuk meg az alábbi mátrixok egy LU-felbontását!

5.32. $\begin{bmatrix} 4 & 4 & 4 \\ 2 & 5 & 5 \\ 1 & 2 & 4 \end{bmatrix}$

5.33. $\begin{bmatrix} 4 & 8 & 4 \\ 2 & 7 & 8 \\ 1 & 3 & 4 \end{bmatrix}$

5.34. $\begin{bmatrix} 5 & -4 & -2 \\ 4 & -5 & -5 \\ -3 & 1 & -4 \end{bmatrix}$

5.35. $\begin{bmatrix} -3 & 1 & -3 & 0 \\ -2 & 4 & 3 & -4 \\ 1 & 3 & 3 & 0 \\ -3 & 0 & -3 & -1 \end{bmatrix}$

5.36. $\begin{bmatrix} -1 & 0 & 2 & -2 \\ 0 & 2 & -2 & -1 \\ -2 & -2 & 0 & 3 \\ 1 & 0 & 2 & 1 \end{bmatrix}$

5.37. $\begin{bmatrix} 2.0 & 2.0 & -2.0 \\ -0.5 & 0.0 & -1.0 \\ 1.0 & 1.5 & 1.0 \end{bmatrix}$

Az előző feladatokban megkonstruált LU-felbontásokat használva oldjuk meg az alábbi egyenletrendszereket, azaz oldjuk meg előbb az $\mathbf{Ly} = \mathbf{b}$, majd az $\mathbf{Ux} = \mathbf{y}$ egyenletrendszereket!

5.38. $\begin{bmatrix} 4 & 4 & 4 \\ 2 & 5 & 5 \\ 1 & 2 & 4 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 0 \\ 3 \\ 5 \end{bmatrix}$

5.39. $\begin{bmatrix} 4 & 8 & 4 \\ 2 & 7 & 8 \\ 1 & 3 & 4 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix}$

5.40. $\begin{bmatrix} 5 & -4 & -2 \\ 4 & -5 & -5 \\ -3 & 1 & -4 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 3 \\ -1 \\ -7 \end{bmatrix}$

5.41. $\begin{bmatrix} 2.0 & 2.0 & -2.0 \\ -0.5 & 0.0 & -1.0 \\ 1.0 & 1.5 & 1.0 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 5.6 \\ -1.0 \\ 4.6 \end{bmatrix}$

Határozzuk meg az alábbi mátrixok inverzét az LU-felbontásuk ismeretében, azaz oldjuk meg az $\mathbf{LY} = \mathbf{I}$ és az $\mathbf{UX} = \mathbf{Y}$ mátrixegyenleteket!

5.42. $\mathbf{A} = \begin{bmatrix} 4 & 4 & 4 \\ 2 & 5 & 5 \\ 1 & 2 & 4 \end{bmatrix}$

5.43. $\begin{bmatrix} 4 & 8 & 4 \\ 2 & 7 & 8 \\ 1 & 3 & 4 \end{bmatrix}$

Adjuk meg az alábbi mátrixok egy PLU-felbontását! Alkalmazzunk részleges főelemkiválasztást!

5.44. $\begin{bmatrix} 1 & 1 & 2 \\ 3 & 3 & 3 \\ 2 & 2 & 3 \end{bmatrix}$

5.45. $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 2 & 3 & 3 & 4 \\ 3 & 4 & 6 & 7 & 9 \end{bmatrix}$

5.46. $\begin{bmatrix} 0.0 & -1.0 & 1.5 \\ 0.5 & -2.0 & 2.0 \\ 0.0 & 2.0 & 2.0 \end{bmatrix}$

5.47. Igazoljuk, hogy az

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

mátrixnak nincs LU-felbontása.

Megoldások

4.1. A két táblázat szorzata:

	csarnok	hipermarket	piac
Anti	1350	1250	1210
Bori	1425	1245	1225
Cili	960	830	850

Tehát Antinak és Borinak a piacon, Cilinek a hipermarketben érdemes vásárolnia.

4.2. A két helyettesítést elvégezve:

$$x = 2a + b = 2(-3s + t) + (4s - t) = -2s + t,$$
$$y = 3a + b = 3(-3s + t) + (4s - t) = -5s + 2t.$$

A két helyettesítés kompozíciója a két helyettesítés táblázatának szorzatával megkapható:

$$\begin{array}{c|cc} & a & b \\ \hline x & 2 & 1 \\ y & 3 & 1 \end{array} \times \begin{array}{c|cc} & s & t \\ \hline a & -3 & 1 \\ b & 4 & -1 \end{array} = \begin{array}{c|cc} & s & t \\ \hline x & -2 & 1 \\ y & -5 & 2 \end{array}$$

4.3. A két helyettesítés kompozíciója a két helyettesítés táblázatának szorzatával megkapható. E szorzatból az olvasható le, hogy a kompozícióval kapott helyettesítés: $x = s$, $y = t$, $z = u$. Ez azt jelenti, hogy a két helyettesítés valamilyen értelemben egymás inverze.

4.4.

1. Kétféleképp adhatjuk meg a táblázatot, ha az első sor és oszlop a tv1-é:

$$\begin{array}{c|cc} \textit{-re} & \textit{tv1} & \textit{tv2} \\ \hline \textit{tv1} & 1/2 & 1/4 \\ \textit{tv2} & 1/2 & 3/4 \end{array} \qquad \begin{array}{c|cc} \textit{-ről} & \textit{tv1} & \textit{tv2} \\ \hline \textit{tv1} & 1/2 & 1/2 \\ \textit{tv2} & 1/4 & 3/4 \end{array}$$

2. A nézők kezdeti eloszlásának táblázatára a két lehetőség:

$$\begin{array}{c|c} & \textit{arány} \\ \hline \textit{tv1} & 1/2 \\ \textit{tv2} & 1/2 \end{array} \qquad \begin{array}{c|cc} & \textit{tv1} & \textit{tv2} \\ \hline \textit{arány} & 1/2 & 1/2 \end{array}$$

3. Először válaszoljuk meg a kérdést a tv1-re: a saját nézőinek fele marad ($\frac{1}{2} \cdot \frac{1}{2}$), ehhez jön a tv2 nézőinek negyede ($\frac{1}{2} \cdot \frac{1}{4}$), ez összesen $\frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{4} = \frac{3}{8}$. A tv2-re a számítás: $\frac{1}{2} \cdot \frac{3}{4} + \frac{1}{2} \cdot \frac{1}{2} = \frac{5}{8}$. Ez táblázatok szorzásával az előző 2-2 felírást használva:

$$\begin{array}{c|cc} \textit{-re} & \textit{tv1} & \textit{tv2} \\ \hline \textit{tv1} & 1/2 & 1/4 \\ \textit{tv2} & 1/2 & 3/4 \end{array} \times \begin{array}{c|c} & \textit{arány} \\ \hline \textit{tv1} & 1/2 \\ \textit{tv2} & 1/2 \end{array} = \begin{array}{c|c} & \textit{arány} \\ \hline \textit{tv1} & 3/8 \\ \textit{tv2} & 5/8 \end{array}$$

$$\begin{array}{c|cc} & \textit{tv1} & \textit{tv2} \\ \hline \textit{arány} & 1/2 & 1/2 \end{array} \times \begin{array}{c|cc} \textit{-ről} & \textit{tv1} & \textit{tv2} \\ \hline \textit{tv1} & 1/2 & 1/2 \\ \textit{tv2} & 1/4 & 3/4 \end{array} = \begin{array}{c|cc} & \textit{tv1} & \textit{tv2} \\ \hline \textit{arány} & 3/8 & 5/8 \end{array}$$

4. Csak az átpártolás táblázatát nézve, a második hét végére a tv1 nézői első héten megmaradt felének csak a fele marad meg, míg a tv2-től átpártolt negyednyi közönségnek is a fele, tehát a tv1 → tv1 „mozgás” a nézők 3/8-adát érinti, mert $\frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{4} = \frac{3}{8}$. Hasonló számításokkal a többi érték is megkapható, melyet az alábbi, két, táblázatok közti szorzással is meg lehet adni:

$$\begin{array}{c|cc} \textit{-re} & \textit{tv1} & \textit{tv2} \\ \hline \textit{tv1} & 1/2 & 1/4 \\ \textit{tv2} & 1/2 & 3/4 \end{array} \times \begin{array}{c|cc} \textit{-re} & \textit{tv1} & \textit{tv2} \\ \hline \textit{tv1} & 1/2 & 1/4 \\ \textit{tv2} & 1/2 & 3/4 \end{array} = \begin{array}{c|cc} \textit{-re} & \textit{tv1} & \textit{tv2} \\ \hline \textit{tv1} & 3/8 & 5/16 \\ \textit{tv2} & 5/8 & 11/16 \end{array}$$

$$\begin{array}{c|cc} \textit{-ről} & \textit{tv1} & \textit{tv2} \\ \hline \textit{tv1} & 1/2 & 1/2 \\ \textit{tv2} & 1/4 & 3/4 \end{array} \times \begin{array}{c|cc} \textit{-ről} & \textit{tv1} & \textit{tv2} \\ \hline \textit{tv1} & 1/2 & 1/2 \\ \textit{tv2} & 1/4 & 3/4 \end{array} = \begin{array}{c|cc} \textit{-ről} & \textit{tv1} & \textit{tv2} \\ \hline \textit{tv1} & 3/8 & 5/8 \\ \textit{tv2} & 5/16 & 11/16 \end{array}$$

4.5. Ha a a $[0, k]$ intervallumba eső szám, akkor $0 \leq a/k \leq 1$, így a/k egész része 0 vagy 1. Részletezve $\lfloor a/k \rfloor$ pontosan akkor 1, ha $a = k$, egyébként 0. Másrészt $1 - \lfloor a/k \rfloor$ pontosan akkor 0, ha $a = k$, egyébként 1. Ezt kihasználva könnyen definiálható a kívánt művelet:

$$a \circledcirc b = \left\lfloor \frac{a}{k} \right\rfloor b + \left(1 - \left\lfloor \frac{a}{k} \right\rfloor\right) a.$$

Így e művelettel elemenként definiált $\mathbf{A} \circledcirc \mathbf{B}$ művelet a kívánt eredményt adja. A 4.2. ábrán három képet 32×24-es mátrixszal szemléltetünk, a férfiarc mátrixát is megadtuk, a másik a háttér. A művelet eredménye a harmadik kép.

4.6. A standard bázisba azon mátrixok tartoznak, amelyekben egyetlen elem 1, a többi 0.

4.7. E mátrixok összes lineáris kombinációja

$$a\mathbf{A} + b\mathbf{B} + c\mathbf{C} = \begin{bmatrix} a+b+c & a+c \\ a & b+c \end{bmatrix}$$

alakú. Ha egy tetszőleges $\left[\begin{smallmatrix} u & v \\ w & z \end{smallmatrix}\right]$ mátrixról el akarjuk dönteni, hogy a fenti alakú-e, azaz fönnáll-e valamely a, b, c ismeretlenekre az

$$\begin{bmatrix} a+b+c & a+c \\ a & b+c \end{bmatrix} = \begin{bmatrix} u & v \\ w & z \end{bmatrix}$$

egyenlőség, akkor meg kell oldani a mátrixok négy elemére vonatkozó négy egyenletből álló 3-ismeretlenes egyenletrendszert:

$$\begin{aligned} a+b+c &= u \\ a \quad\;\; +c &= v \\ a \qquad\quad\; &= w \\ b+c &= z \end{aligned}$$

Ha ennek van megoldása, akkor létezik a megfelelő lineáris kombináció, tehát az adott $\left[\begin{smallmatrix} u & v \\ w & z \end{smallmatrix}\right]$ mátrix a kifeszített térbe esik. Ennek az egyenletrendszernek a bővített mátrixát fölírva, majd elemi sorműveletekkel megoldva a következőt kapjuk:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & u \\ 1 & 0 & 1 & v \\ 1 & 0 & 0 & w \\ 0 & 1 & 1 & z \end{array}\right] \Longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & w \\ 0 & 1 & 1 & u-w \\ 0 & 0 & 1 & v-w \\ 0 & 1 & 1 & z \end{array}\right] \Longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & w \\ 0 & 1 & 1 & u-w \\ 0 & 0 & 1 & v-w \\ 0 & 0 & 0 & w+z-u \end{array}\right]$$

A lépcsős alakból leolvasható, hogy ez az egyenletrendszer pontosan akkor oldható meg, ha $w+z-u=0$. Például az $\left[\begin{smallmatrix} 5 & 4 \\ 3 & 2 \end{smallmatrix}\right]$ mátrix ebbe az altérbe esik. A fenti egyenletrendszer megoldásával az is megkapható, hogy mik a lineáris kombináció együtthatói. Azt kapjuk, hogy $a=3$, $b=1$ és $c=1$.

4.18. $\mathbf{a}\cdot\mathbf{b} = \mathbf{a}^T\mathbf{b} = \begin{bmatrix} 1 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = 2$, $\mathbf{a}\otimes\mathbf{b} = \mathbf{a}\mathbf{b}^T = \begin{bmatrix} 1 \\ 2 \end{bmatrix}\begin{bmatrix} 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}$.

4.19. $\mathbf{u}\cdot\mathbf{v} = \mathbf{u}^T\mathbf{v} = \begin{bmatrix} 1 & 2 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix} = 5,$

$$\mathbf{u}\otimes\mathbf{v} = \mathbf{u}\mathbf{v}^T = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}\begin{bmatrix} 0 & 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 0 & 2 & 4 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 \end{bmatrix}.$$

4.20. A skaláris szorzat nem végezhető el, a diadikus szorzat

$$\mathbf{a}\otimes\mathbf{b} = \mathbf{a}\mathbf{b}^T = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}\begin{bmatrix} 0 & 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 0 & 2 & 4 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

4.21. $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

4.27. $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 & -2 & 1 \\ 2 & 0 & -1 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix}$

4.28. $\mathbf{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

4.29. $\mathbf{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

4.30. $\mathbf{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$

4.31. $\mathbf{E} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

4.34. $\left[\begin{array}{cc|cc} 2 & 3 & 2 & 2 \\ 4 & 5 & 2 & 2 \\ \hline 0 & 0 & 2 & 2 \\ 0 & 0 & 3 & 3 \end{array}\right]$

4.35. $\left[\begin{array}{cc|c} 2 & 3 & 10 \\ 4 & 5 & 14 \\ \hline 0 & 0 & 5 \end{array}\right]$

4.36. $\left[\begin{array}{cc} 5 & 6 \\ 6 & 7 \\ \hline 12 & 12 \end{array}\right]$

4.37. Jelölje $\mathbf{j}$ a csupa 1-esből álló 9-dimenziós vektort, $\mathbf{j}_{456}$ azt, amelynek 4, 5, 6 indexű eleme 1-es, a többi 0. Ekkor a „minden sorösszeg 45” és a „minden oszlopösszeg 45” feltételek ekvivalensek az $\mathbf{A}\mathbf{j} = 45\mathbf{j}$, $\mathbf{j}^T\mathbf{A} = 45\mathbf{j}^T$ egyenletekkel, míg pl. az „első blokkoszlop, második blokksor metszetében álló blokk elemeinek összege 45” feltételnek a $\mathbf{j}_{456}^T\mathbf{A}\mathbf{j}_{123} = 45$ egyenlet felel meg.

4.38. $\mathbb{Z}_2^{2\times 2}$-be $2^4 = 16$ mátrix tartozik:

$$\mathbb{Z}_2^{2\times 2} = \left\{\left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right], \left[\begin{smallmatrix} 1 & 1 \\ 0 & 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} 1 & 0 \\ 1 & 0 \end{smallmatrix}\right], \dots, \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]\right\}.$$

5.4. Helyettesítés előtt:

$$\mathbf{u}^T\mathbf{u}\begin{bmatrix} 2 & -1 & -1 \end{bmatrix}\mathbf{u}.$$

Az $\mathbf{u} = \mathbf{A}\mathbf{x}$ helyettesítés elvégzése után

$$\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x}\begin{bmatrix} 2 & -1 & -1 \end{bmatrix}\mathbf{A}\mathbf{x},$$

ahol

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \ \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}, \ \mathbf{A} = \begin{bmatrix} 3 & 2 & 4 \\ 1 & -3 & 1 \\ 2 & -1 & -3 \end{bmatrix}.$$

5.5. $\begin{bmatrix}1&3\\2&8\end{bmatrix}=\begin{bmatrix}1&0\\2&1\end{bmatrix}\begin{bmatrix}1&0\\0&2\end{bmatrix}\begin{bmatrix}1&3\\0&1\end{bmatrix}$

5.6. $\begin{bmatrix}1&2\\-2&-1\end{bmatrix}=\begin{bmatrix}1&0\\-2&1\end{bmatrix}\begin{bmatrix}1&0\\0&3\end{bmatrix}\begin{bmatrix}1&2\\0&1\end{bmatrix}$

5.7. $\begin{bmatrix}1&-1\\1&1\end{bmatrix}=\begin{bmatrix}1&0\\1&1\end{bmatrix}\begin{bmatrix}1&0\\0&2\end{bmatrix}\begin{bmatrix}1&-1\\0&1\end{bmatrix}$

5.8. $\begin{bmatrix}2&4\\3&8\end{bmatrix}=\begin{bmatrix}2&0\\0&1\end{bmatrix}\begin{bmatrix}1&0\\3&1\end{bmatrix}\begin{bmatrix}1&0\\0&2\end{bmatrix}\begin{bmatrix}1&2\\0&1\end{bmatrix}$

5.9. $\begin{bmatrix}2&0&4\\0&2&0\\3&2&7\end{bmatrix}=\begin{bmatrix}2&0&0\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\3&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&2&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\0&2&1\end{bmatrix}\begin{bmatrix}1&0&2\\0&1&0\\0&0&1\end{bmatrix}$

5.10. $\begin{bmatrix}1&1&2\\1&2&2\\2&4&5\end{bmatrix}=\begin{bmatrix}1&0&0\\1&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\2&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\0&2&1\end{bmatrix}\begin{bmatrix}1&0&2\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&1&0\\0&1&0\\0&0&1\end{bmatrix}$

5.11. Legyen $\mathbf{A}=\left[\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right]$. Négyzete a zérusmátrix, azaz

$$\mathbf{A}^2=\begin{bmatrix}a&b\\c&d\end{bmatrix}\begin{bmatrix}a&b\\c&d\end{bmatrix}=\begin{bmatrix}a^2+bc&b(a+d)\\c(a+d)&bc+d^2\end{bmatrix}=\begin{bmatrix}0&0\\0&0\end{bmatrix}.$$

Innen vagy $a=d=0$ és c vagy d legalább egyike 0, vagy $a\neq 0$, $c\neq 0$ és $b=-a^2/c$, $d=-a$.

5.12. A feladat érdekes, abban a Fibonacci sorozat elemei bukkannak föl. Ez az $f_0=0$, $f_1=1$, $f_{k+1}=f_k+f_{k-1}$ egyenlőségekkel definiált sorozat, melynek első néhány tagja: 0, 1, 1, 2, 3, 5, 8, 13, 21,.... Tekintsük $\mathbf{B}$ néhány hatványát:

$$\mathbf{A}^2=\begin{bmatrix}0&1\\1&1\end{bmatrix}\begin{bmatrix}0&1\\1&1\end{bmatrix}=\begin{bmatrix}1&1\\1&2\end{bmatrix}=\begin{bmatrix}f_1&f_2\\f_2&f_3\end{bmatrix},$$
$$\mathbf{A}^3=\mathbf{A}^2\mathbf{A}=\begin{bmatrix}1&1\\1&2\end{bmatrix}\begin{bmatrix}0&1\\1&1\end{bmatrix}=\begin{bmatrix}1&2\\2&3\end{bmatrix}=\begin{bmatrix}f_2&f_3\\f_3&f_4\end{bmatrix}.$$

Ennek alapján azt sejtjük, hogy

$$\mathbf{A}^n=\begin{bmatrix}f_{n-1}&f_n\\f_n&f_{n+1}\end{bmatrix}.$$

Az állítás $n=1,2,3$ esetén igaz, és n-ről öröklődik $n+1$-re, ugyanis

$$\begin{aligned}\mathbf{A}^{n+1}=\mathbf{A}^n\mathbf{A}&=\begin{bmatrix}f_{n-1}&f_n\\f_n&f_{n+1}\end{bmatrix}\begin{bmatrix}0&1\\1&1\end{bmatrix}\\&=\begin{bmatrix}f_n&f_{n+1}\\f_{n-1}+f_n&f_n+f_{n+1}\end{bmatrix}=\begin{bmatrix}f_n&f_{n+1}\\f_{n+1}&f_{n+2}\end{bmatrix}.\end{aligned}$$

5.13. $\mathbf{C}=\mathbf{AB}$ Einstein-konvencióval: $c_{ij}=a_{ik}b_{kj}$.

5.16. $\begin{bmatrix}2&-3&0&0&0\\-1&2&0&0&0\\0&0&-5&3&3\\0&0&-16&9&10\\0&0&28&-16&-17\end{bmatrix}$

5.17. $\begin{bmatrix}2&-3&7&-4&-4\\-1&2&-7&4&4\\0&0&-5&3&3\\0&0&-16&9&10\\0&0&28&-16&-17\end{bmatrix}$

5.18. $\begin{bmatrix}-1&0&1&1&1\\2&-1&-1&-1&-1\\-1&1&1&0&0\\-1&1&0&1&0\\-1&0&0&0&1\end{bmatrix}$

5.21. A fölcserélhetőségre vonatkozó $\mathbf{AB}=\mathbf{BA}$ egyenletet szorozzuk meg mindkét oldalról $\mathbf{B}^{-1}$-gyel:

$$\mathbf{B}^{-1}(\mathbf{AB})\mathbf{B}^{-1}=\mathbf{B}^{-1}(\mathbf{BA})\mathbf{B}^{-1}.$$

Az asszociativitást használva

$$(\mathbf{B}^{-1}\mathbf{A})(\mathbf{BB}^{-1})=(\mathbf{B}^{-1}\mathbf{B})(\mathbf{AB}^{-1}),$$

amiből a $\mathbf{BB}^{-1}=\mathbf{I}$ azonosság fölhasználásával kapjuk, hogy

$$\mathbf{B}^{-1}\mathbf{A}=\mathbf{AB}^{-1}.$$

5.22. Azt mondjuk, hogy a $\circledcirc$ művelet invertálható a H egy R részhalmazán, ha bármely $a,b,c\in R$ elem esetén az

$$a\circledcirc x=b,\qquad y\circledcirc a=c$$

egyenletek mindegyike megoldható, azaz vannak olyan $x,y\in H$ elemek, melyek kielégítik a fenti egyenleteket. Ha a definícióbeli $\circledcirc$ kommutatív művelet, akkor elég a fenti két egyenlet egyikét tekinteni.

5.23. *a)* Az $e\in H$ semleges elem, ha minden $a\in H$ elemre $a\circledcirc e=e\circledcirc a=a$. *b)* Azt mondjuk, hogy a $\circledcirc$ műveletre nézve a inverze b, ha $a\circledcirc b=b\circledcirc a=e$.

5.27. Az első egyenletet kivonjuk az utolsóból, innen $x=1$, ezután visszahelyettesítés a második, harmadik, majd az első egyenletbe.

5.32. $\mathbf{L}=\begin{bmatrix}1&0&0\\1/2&1&0\\1/4&1/3&1\end{bmatrix}$, $\mathbf{U}=\begin{bmatrix}4&4&4\\0&3&3\\0&0&2\end{bmatrix}$.

5.33. $\mathbf{L}=\begin{bmatrix}1&0&0\\1/2&1&0\\1/4&1/3&1\end{bmatrix}$, $\mathbf{U}=\begin{bmatrix}4&8&4\\0&3&6\\0&0&1\end{bmatrix}$

5.34. $\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 4/5 & 1 & 0 \\ -3/5 & 7/9 & 1 \end{bmatrix}$, $\mathbf{U} = \begin{bmatrix} 5 & -4 & -2 \\ 0 & -9/5 & -17/5 \\ 0 & 0 & -23/9 \end{bmatrix}$.

5.35. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{2}{3} & 1 & 0 & 0 \\ -\frac{1}{3} & 1 & 1 & 0 \\ 1 & -\frac{3}{10} & -\frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} -3 & 1 & -3 & 0 \\ 0 & \frac{10}{3} & 5 & -4 \\ 0 & 0 & -3 & 4 \\ 0 & 0 & 0 & -\frac{2}{10} \end{bmatrix}$

5.36. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{3} & 1 \end{bmatrix} \begin{bmatrix} -2 & -2 & 0 & 3 \\ 0 & 2 & -2 & -1 \\ 0 & 0 & 3 & -3 \\ 0 & 0 & 0 & 3 \end{bmatrix}$

5.37. $\begin{bmatrix} 1.00 & 0.00 & 0.00 \\ 0.25 & 1.00 & 0.00 \\ 0.50 & 1.00 & 1.00 \end{bmatrix} \begin{bmatrix} 2.00 & 2.00 & -2.00 \\ 0.00 & 0.50 & -1.50 \\ 0.00 & 0.00 & 3.50 \end{bmatrix}$.

5.38. $\mathbf{y} = (0, 3, 4)$, $\mathbf{x} = (-1, -1, 2)$.

5.39. $\mathbf{y} = (0, 3, 1)$, $\mathbf{x} = (1, -1, 1)$.

5.40. $\mathbf{y} = (3, -17/5, -23/9)$, $\mathbf{x} = (1, 0, 1)$.

5.41. $\mathbf{y} = (5.6, 0.4, 1.4)$, $\mathbf{x} = (1.2, 2.0, 0.4)$.

5.42. Az $\mathbf{LY} = \mathbf{I}$ egyenletből

$$\mathbf{Y} = \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ -1/12 & -1/3 & 1 \end{bmatrix},$$

míg az $\mathbf{UX} = \mathbf{Y}$ egyenlet megoldása, egyúttal $\mathbf{A}$ inverze

$$\mathbf{X} = \begin{bmatrix} 5/12 & -1/3 & 0 \\ -1/8 & 1/2 & -1/2 \\ -1/24 & -1/6 & 1/2 \end{bmatrix}.$$

5.43. $\mathbf{Y} = \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ -1/12 & -1/3 & 1 \end{bmatrix}$, $\mathbf{X} = \begin{bmatrix} 1/3 & -5/3 & 3 \\ 0 & 1 & -2 \\ -1/12 & -1/3 & 1 \end{bmatrix}$.

5.44. $\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{2}{3} & 1 & 0 \\ \frac{1}{3} & \frac{1}{2} & 1 \end{bmatrix}$, $\mathbf{U} = \begin{bmatrix} 3 & 3 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}$, $\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$.

5.45. $\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{2}{3} & -1 & 1 \end{bmatrix}$, $\mathbf{U} = \begin{bmatrix} 3 & 4 & 6 & 7 & 9 \\ 0 & \frac{2}{3} & 1 & \frac{5}{3} & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$, $\mathbf{P} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$.

5.46. $\mathbf{L} = \begin{bmatrix} 1.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 \\ 0.0 & -0.5 & 1.0 \end{bmatrix}$, $\mathbf{U} = \begin{bmatrix} 0.5 & -2.0 & 2.0 \\ 0.0 & 2.0 & 2.0 \\ 0.0 & 0.0 & 2.5 \end{bmatrix}$, $\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$.

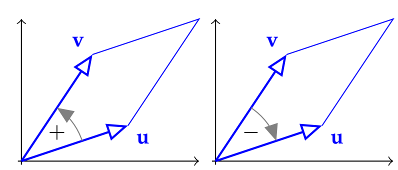
6

Determináns

Egy valós négyzetes mátrix sorvektorai által kifeszített parallelepipedon térfogata jó jellemezője lehet a mátrixnak. Ehhez közel áll a determináns fogalma, mely egy négyzetes mátrixokon értelmezett skalárértékű függvény. E skalár könnyen kiszámolható az elemi sorműveletekkel. A determináns fontos tulajdonságainak egyike, hogy pontosan akkor nulla, ha a sorvektorok lineárisan összefüggők.

A determináns a mátrix egy igen fontos, és sokhelyütt használt jellemzője. Fogalmát egy egyszerű és szemléletes fogalommal, a parallelepipedon előjeles térfogatának segítségével fogjuk bevezetni.

Parallelogramma előjeles területe A parallelogramma területéről szóló példában láttuk, hogy az (a, b) és a (c, d) vektorok által kifeszített parallelogramma területe $|ad - bc|$, és hogy $ad - bc$ pontosan akkor pozitív, ha az (a, b) és a (c, d) vektorok jobbrendszert alkotnak, és pontosan akkor negatív, ha az (a, b) és a (c, d) vektorok balrendszert alkotnak. Ez vezet a következő definícióhoz: Két síkbeli vektor által kifeszített parallelogramma *előjeles területén* a területét értjük, ha a két vektor jobbrendszert alkot, és a terület -1-szeresét, ha a két vektor balrendszert alkot.



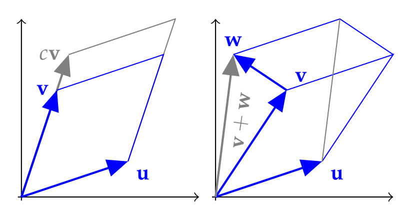
6.1. ábra: Két vektor sorrendjének cseréje megváltoztatja a parallelogramma körüljárását.

Az előzőek szerint az $\mathbf{u} = (a, b)$ és a $\mathbf{v} = (c, d)$ vektorok által kifeszített parallelogramma előjeles területe $ad - bc$. Az előjeles terület tehát egy vektorpárokon értelmezett valós értékű függvény – jelölje most f –, mely eleget tesz a következő tulajdonságoknak.

1. $f(\mathbf{u}, \mathbf{v}) = -f(\mathbf{v}, \mathbf{u})$. Ez nyilvánvaló, hisz a két vektor sorrendjét megcserélve megváltozik a parallelogramma irányítása (6.1. ábra).

2. $f(\mathbf{u}, \mathbf{u}) = 0$, ugyanis az elfajuló parallelogramma területe 0.

3. $f(c\mathbf{u}, \mathbf{v}) = cf(\mathbf{u}, \mathbf{v})$, és $f(\mathbf{u}, c\mathbf{v}) = cf(\mathbf{u}, \mathbf{v})$, azaz f homogén mindkét változójában. Ez igaz, mert egy parallelogramma egyik oldalának c-szeresére növelése c-szerezi a területét is (6.2. ábra).

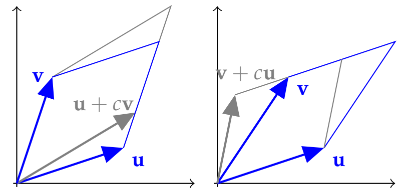


6.2. ábra: Az $f(\mathbf{u}, c\mathbf{v}) = cf(\mathbf{u}, \mathbf{v})$ és az $f(\mathbf{u}, \mathbf{v}) + f(\mathbf{u}, \mathbf{w}) = f(\mathbf{u}, \mathbf{v} + \mathbf{w})$ összefüggések szemléltetése.

4. $f(\mathbf{u},\mathbf{v}) + f(\mathbf{u},\mathbf{w}) = f(\mathbf{u},\mathbf{v}+\mathbf{w})$ és $f(\mathbf{u},\mathbf{v}) + f(\mathbf{w},\mathbf{v}) = f(\mathbf{u}+\mathbf{w},\mathbf{v})$, azaz f additív mindkét változójában. Az állítás igaz volta leolvasható a 6.3. ábráról.

5. $f(\mathbf{u},\mathbf{v}) = f(\mathbf{u}+c\mathbf{v},\mathbf{v}) = f(\mathbf{u},\mathbf{v}+c\mathbf{u})$, azaz az $\left[\begin{smallmatrix}\mathbf{u}\\ \mathbf{v}\end{smallmatrix}\right] = \left[\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right]$ mátrix sorvektorai által kifeszített parallelogramma területe megegyezik a hozzáadás sorművelete után kapott mátrix sorvektoraihoz tartozó területtel.

6. $f(\mathbf{i},\mathbf{j}) = 1$, azaz a standard bázis által kifeszített egységnégyzet területe 1.



6.3. ábra: Az $f(\mathbf{u},\mathbf{v}) = f(\mathbf{u}+c\mathbf{v},\mathbf{v})$ és az $f(\mathbf{u},\mathbf{v}) = f(\mathbf{u},\mathbf{v}+c\mathbf{u})$ összefüggések szemléltetése.

Az állítások a fenti ábrákkal szemléltetett egyszerű geometriai érvelések mellett az $f((a,b),(c,d)) = ad - bc$ formulával is bizonyíthatóak. E tulajdonságok segítségével általánosítani lehet az előjeles terület fogalmát, és bevezetni az előjeles térfogat fogalmát az n-dimenziós valós tér parallelepipedonjaira.

Parallelepipedon előjeles térfogata A vegyes szorzat tárgyalásakor láttuk, hogy a valós háromdimenziós térben három vektor vegyes szorzatának abszolút értéke a vektorok által kifeszített parallelepipedon térfogata lesz, előjele pedig aszerint pozitív vagy negatív, hogy a három vektor jobb- vagy balrendszert alkot. A háromdimenziós tér parallelepipedonjaira a parallelogrammához hasonló tulajdonságok igazolhatók.

1. Bármely két argumentum felcserélése megváltoztatja a függvényérték előjelét, pl. $f(\mathbf{u},\mathbf{v},\mathbf{w}) = -f(\mathbf{w},\mathbf{v},\mathbf{u})$.

2. Ha f bármely két argumentuma megegyezik, a függvényérték 0, pl. $f(\mathbf{u},\mathbf{v},\mathbf{u}) = 0$.

3. f homogén mindhárom argumentumában, pl. $f(c\mathbf{u},\mathbf{v},\mathbf{w}) = cf(\mathbf{u},\mathbf{v},\mathbf{w})$.

4. f additív mindhárom argumentumában, pl. $f(\mathbf{u},\mathbf{v},\mathbf{w}) + f(\mathbf{u},\mathbf{v},\mathbf{z}) = f(\mathbf{u},\mathbf{v},\mathbf{w}+\mathbf{z})$.

5. f bármely argumentumához hozzáadva egy másik konstansszorosát, a függvényérték nem változik, pl. $f(\mathbf{u},\mathbf{v},\mathbf{w}) = f(\mathbf{u}+c\mathbf{w},\mathbf{v},\mathbf{w})$.

6. $f(\mathbf{i},\mathbf{j},\mathbf{k}) = 1$, azaz az egységkocka térfogata 1.

A háromdimenziós parallelepipedon térfogatát úgy számoltuk ki, hogy két vektor által kifeszített parallelogramma területét szoroztuk a harmadik vektor csúcsának a parallelogramma síkjától való távolságával. Ha pedig a parallelogramma irányítását is figyelembe vettük, e távolság előjeles távolsággá vált, hisz az ott használt skaláris szorzat a sík egyik oldalán pozitív, másik oldalán negatív eredményt ad. Ugyanez az eljárás megismételhető magasabb dimenzióban is. Például, ha a négydimenziós tér egy parallelepipedonjának előjeles térfogatát akarjuk kiszámolni, akkor egy háromdimenziós parallelepipedon térfogatát szorozzuk a negyedik vektor végpontjának a másik három terétől való előjeles távolságával. Ezután bebizonyíthatnánk, hogy az így definiált előjeles térfogat is rendelkezik a korábban felsorolt tulajdonságok négydimenziós megfelelőivel. E helyett egy más, egyszerűbb és szebb utat választunk.

A determináns, mint sorvektorainak függvénye

A determináns definíciója A parallelepipedon fogalma helyett – mely nem értelmezhető bármely számtest esetén – egyszerűen csak az azt kifeszítő vektorokat, illetve az azokból képzett mátrixot fogjuk használni. Az előjeles térfogat helyett olyan fogalmat fogunk definiálni, mely speciális esetként ezt is tartalmazza. Ez lesz a determináns. A definícióban csak az előjeles térfogat vizsgálatában megismert függvénytulajdonságokat használjuk, azok közül is csak annyit, amennyi már egyértelműen definiálja azt.

A determinánsokat először Takakazu Shinsuke Seki (1642–1708) vizsgálta, eredményei 1683-ban jelentek meg. Seki 2-től 5-ödrendűek értékét tudta kiszámolni, egyenletek megoldására használta, de lineáris egyenletrendszerek megoldására nem. Európában is 1683-ban jelenik meg e fogalom először Leibniz egy l'Hôpitalnak írt levelében, melyet később rezultánsnak hív. A determináns név Gausstól származik. A mátrixok és determinánsok történetének szép összefoglalója olvasható a MacTutor History of Mathematics weboldalon.

6.1. DEFINÍCIÓ (DETERMINÁNS). *A* determináns *a négyzetes mátrixokon értelmezett és* det*-tel jelölt olyan skalár értékű függvény, mely eleget tesz az alábbi tulajdonságoknak:*
D1. *lineáris a mátrix minden sorára nézve,*
D2. *két azonos sort tartalmazó mátrixhoz* 0*-t,*
D3. *az egységmátrixhoz* 1*-et rendel.*
Az $\mathbf{A} = [a_{ij}]_n$ *mátrix determinánsát* $\det(\mathbf{A})$, $|\mathbf{A}|$ *vagy* $|a_{ij}|_n$ *jelöli.*

► Az $n \times n$-es mátrixok determinánsát szokás *n-edrendű* determinánsnak is nevezni.

► Részletezve az általános jelölést az $\mathbf{A}$ mátrixra és determinánsára:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, \qquad \det(\mathbf{A}) = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

E jelölésnek megfelelően a $\det(\mathbf{A})$ determináns sorain, oszlopain, elemein az $\mathbf{A}$ mátrix sorait, oszlopait, elemeit értjük.

► Az 1×1-es $[a]$ mátrix determinánsa $\det([a]) = a$, ugyanis a determináns definíciója szerint $\det([1]) = 1$, és $\det([a]) = \det([a \cdot 1]) = a\det([1]) = a$. Ez a függvény pedig additív is, így a definíció minden feltételét kielégíti. A jelölésbeli zavarok elkerülésére az 1×1-es $[a]$ mátrix determinánsára csak a $\det([a])$ vagy $\det(a)$ jelölést használjuk, mert $|a|$ az a abszolút értékét jelöli!

► Részletezzük a definíció három feltételét. A determináns lineáris a mátrix minden sorában, tehát ha két determináns sorvektorai megegyeznek, kivéve esetleg az i-ediket ($1 \le i \le n$), akkor tetszőleges c és d skalárokkal

$$c\begin{vmatrix} \mathbf{a}_{1*} \\ \vdots \\ \mathbf{x}_{i*} \\ \vdots \\ \mathbf{a}_{n*} \end{vmatrix} + d\begin{vmatrix} \mathbf{a}_{1*} \\ \vdots \\ \mathbf{y}_{i*} \\ \vdots \\ \mathbf{a}_{n*} \end{vmatrix} = \begin{vmatrix} \mathbf{a}_{1*} \\ \vdots \\ c\mathbf{x}_{i*} + d\mathbf{y}_{i*} \\ \vdots \\ \mathbf{a}_{n*} \end{vmatrix}.$$

Azonos sorokat tartalmazó determináns értéke 0, azaz ha valamely $i \neq j$ esetén $\mathbf{a}_{i*} = \mathbf{a}_{j*} = \mathbf{b}$, akkor

$$\begin{vmatrix} \vdots \\ \mathbf{b} \\ \vdots \\ \mathbf{b} \\ \vdots \end{vmatrix} = 0.$$

Végül az egységmátrix determinánsa 1, azaz

$$\begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{vmatrix} = 1.$$

► Hamarosan igazolni fogjuk, hogy a fejezet elején említett

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

képlet illeszkedik e definícióhoz.

► Láttunk determinánst, például a 2×2-es mátrixokét. Azt azonban még nem tudjuk, hogy létezik-e minden n-re és egyértelmű-e. A definíció alapján ez bizonyítható. A bizonyítást későbbre hagyjuk, addig feltételezzük, hogy létezik és egyértelmű.

► A determináns tekinthető olyan n-változós függvénynek, melynek n argumentumába a mátrix n sorvektora kerül. Nem okoz félreértést, ha ezt a függvényt is det jelöli. Ha tehát $\mathbf{A}$ sorvektorai $\mathbf{a}_{1*}, \mathbf{a}_{2*}, \dots, \mathbf{a}_{n*}$, akkor $\det(\mathbf{A})$ megegyezik a $\det(\mathbf{a}_{1*}, \mathbf{a}_{2*}, \dots, \mathbf{a}_{n*})$ függvényértékkel. Például a 3×3-as egységmátrix determinánsa az alábbi alakokba írható:

$$\det(I_3) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = \det([1\ 0\ 0],\ [0\ 1\ 0],\ [0\ 0\ 1]) = \det(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$$

ahol $\mathbf{e}_1$ $\mathbf{e}_2$, $\mathbf{e}_3$ a standard egység-sorvektorokat jelöli. A determináns fenti definíciója könnyen fölírható e jelöléssel is (ld. 6.10. feladat).

```
sage: M = matrix(3,range(9))
sage: M[2,2]=9
sage: M
[0 1 2]
[3 4 5]
[6 7 9]
sage: M.det()
-3
sage: det(M)
-3
```

6.4. kód: Determináns kiszámítása

A determináns értékének kiszámítása A determináns kiszámításához az elemi sorműveleteket fogjuk használni. A háttérben lényegében ezt teszik a számítógépek is (ld. 6.4. kód és a program).

Két kérdésre kell válaszolnunk: (1) hogyan változik a determináns értéke elemi sorműveletek közben, (2) mennyi a determinánsa a háromszög alakra hozott mátrixoknak?

6.2. ÁLLÍTÁS (NULLVEKTORT TARTALMAZÓ DETERMINÁNS). *Egy determináns értéke* 0*, ha valamely sorában minden elem* 0*.*

BIZONYÍTÁS. Ha egy determináns egy 0-sorát bármely c számmal beszorozzuk, az 0-sor marad, így a determináns értéke nem változik, másrészt a c-vel való szorzás miatt c-szeresére módosul. Mivel csak a 0 egyezik meg tetszőleges c-re saját c-szeresével, így a determináns értéke csak 0 lehet. □

6.3. TÉTEL (ELEMI SORMŰVELETEK DETERMINÁNSON). *Az elemi sorműveletek eredményeként a determináns értéke az alábbiak szerint változik:*

a) sorcsere *közben előjelet vált;*
b) a c skalárral való beszorzás *után értéke c-szeresére változik;*
c) egy sor konstansszorosának egy másikhoz való hozzáadása *után értéke nem változik.*

BIZONYÍTÁS. A tétel második pontja definíció szerint igaz. A harmadik bizonyítása:

$$\begin{vmatrix} \vdots \\ \mathbf{a}_{i*} \\ \vdots \\ \mathbf{a}_{j*}+c\mathbf{a}_{i*} \\ \vdots \end{vmatrix} = \begin{vmatrix} \vdots \\ \mathbf{a}_{i*} \\ \vdots \\ \mathbf{a}_{j*} \\ \vdots \end{vmatrix} + \begin{vmatrix} \vdots \\ \mathbf{a}_{i*} \\ \vdots \\ c\mathbf{a}_{i*} \\ \vdots \end{vmatrix} = \begin{vmatrix} \vdots \\ \mathbf{a}_{i*} \\ \vdots \\ \mathbf{a}_{j*} \\ \vdots \end{vmatrix} + c \begin{vmatrix} \vdots \\ \mathbf{a}_{i*} \\ \vdots \\ \mathbf{a}_{i*} \\ \vdots \end{vmatrix} = \begin{vmatrix} \vdots \\ \mathbf{a}_{i*} \\ \vdots \\ \mathbf{a}_{j*} \\ \vdots \end{vmatrix}.$$

Az első állítás igazolásához a harmadikat használjuk:

$$\begin{vmatrix} \vdots \\ \mathbf{a}_{i*} \\ \vdots \\ \mathbf{a}_{j*} \\ \vdots \end{vmatrix} = \begin{vmatrix} \vdots \\ \mathbf{a}_{i*} \\ \vdots \\ \mathbf{a}_{j*}-\mathbf{a}_{i*} \\ \vdots \end{vmatrix} = \begin{vmatrix} \vdots \\ \mathbf{a}_{j*} \\ \vdots \\ \mathbf{a}_{j*}-\mathbf{a}_{i*} \\ \vdots \end{vmatrix} = \begin{vmatrix} \vdots \\ \mathbf{a}_{j*} \\ \vdots \\ -\mathbf{a}_{i*} \\ \vdots \end{vmatrix} \overset{D2}{=} - \begin{vmatrix} \vdots \\ \mathbf{a}_{j*} \\ \vdots \\ \mathbf{a}_{i*} \\ \vdots \end{vmatrix}$$

Először az i-edik sort kivontuk a j-edikből, majd az így kapott j-ediket hozzáadtuk az i-edikhez, végül az i-ediket ismét kivontuk a j-edikből. □

Az elemi mátrixok egyetlen sorművelettel kaphatók az egységmátrixból, így ezek determinánsa könnyen számolható. Hasonlóan könnyen számolható egy elemi mátrix és egy tetszőleges mátrix szorzatának determinánsa.

6.4. KÖVETKEZMÉNY (ELEMI MÁTRIXOK DETERMINÁNSA).
a) *A hozzáadás sorműveletével kapott elemi mátrix determinánsa* 1, *a sorcserével kapotté* -1, *egy sor* c*-vel való sorzásával kapotté* c.
b) *Egy* $\mathbf{E}$ *elemi mátrix és egy tetszőleges négyzetes* $\mathbf{A}$ *mátrix szorzatának determinánsa megegyezik determinánsaik szorzatával, azaz*

$$\det(\mathbf{EA}) = \det(\mathbf{E})\det(\mathbf{A}).$$

BIZONYÍTÁS. Az *(a)* állítás abból következik, hogy az elemi mátrixok az 1 determinánsú egységmátrixból kaphatók egyetlen sorművelettel. Hasonlóképp adódik *(b)* abból, hogy az $\mathbf{EA}$ egyetlen sorművelettel kapható $\mathbf{A}$-ból. □

Például:

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 4 & 0 & 1 \end{vmatrix} = 1, \quad \begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{vmatrix} = -1, \quad \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{vmatrix} = 3.$$

6.5. ÁLLÍTÁS (PERMUTÁCIÓS MÁTRIX DETERMINÁNSA). *Permutációs mátrix determinánsa* 1 *vagy* -1.

BIZONYÍTÁS. Mivel a permutációs mátrix csak elemi sorcserékkel megkapható az egységmátrixból, és a sorcsere csak a determináns előjelét változtatja meg, ezért permutációs mátrix determinánsa 1, ha páros sok sorcserére volt szükség, -1, ha páratlan sokra. Például az alábbi determinánsok közül az első determináns két sorcserével, a második három sorcserével kapható meg az egységmátrixból, tehát

$$\begin{vmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{vmatrix} = 1, \quad \begin{vmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{vmatrix} = -1.$$

□

6.6. TÉTEL (HÁROMSZÖGMÁTRIX DETERMINÁNSA). *Az alsó vagy felső háromszögmátrix, s így a diagonális mátrix determinánsa megegyezik a főátlóbeli elemek szorzatával.*

BIZONYÍTÁS. Ha egy háromszögmátrix főátlójában van 0, akkor a redukált lépcsős alakra hozás után a főelemek száma kevesebb lesz, mint a sorok száma, azaz a mátrixban lesz egy zérussor, így determinánsának értéke 0. Ha nincs 0-elem a főátlóban, mind az alsó, mind a felső háromszögmátrix csak a hozzáadás sorműveletével – azaz a determináns értékének megváltoztatása nélkül – diagonálissá alakítható a főátlón kívüli elemek kiküszöbölésével, azaz

$$\begin{vmatrix} a_{11} & 0 & \dots & 0 \\ ? & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ ? & ? & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & ? & \dots & ? \\ 0 & a_{22} & \dots & ? \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix}$$

Egy diagonális mátrix determinánsában minden sorból kiemelve a főátlóban szereplő számot kapjuk, hogy

$$\begin{vmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix} = a_{11}a_{22}\dots a_{nn} \begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{vmatrix} = a_{11}a_{22}\dots a_{nn},$$

tehát a determináns értéke valóban a főátlóbeli elemek szorzata. □

Például az alábbi determináns értéke egyetlen sorcsere után azonnal leolvasható:

$$\begin{vmatrix} 3 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 2 & 0 \\ 3 & 0 & 2 & 0 & 0 \\ 3 & 2 & 0 & 0 & 0 \\ 3 & 3 & 3 & 3 & 3 \end{vmatrix} = - \begin{vmatrix} 3 & 0 & 0 & 0 & 0 \\ 3 & 2 & 0 & 0 & 0 \\ 3 & 0 & 2 & 0 & 0 \\ 3 & 0 & 0 & 2 & 0 \\ 3 & 3 & 3 & 3 & 3 \end{vmatrix} = -3 \cdot 2 \cdot 2 \cdot 2 \cdot 3 = -72$$

6.7. PÉLDA (DETERMINÁNS KISZÁMÍTÁSA HÁROMSZÖG ALAKRA HOZÁSSAL). *Számítsuk ki a*

$$\begin{vmatrix} 2 & 2 & -3 \\ 2 & 2 & -4 \\ 4 & 5 & -6 \end{vmatrix} \quad \textit{és a} \quad \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{vmatrix}$$

determinánsok értékét!

MEGOLDÁS. Elemi sorműveletekkel kapjuk, hogy

$$\begin{vmatrix} 2 & 2 & -3 \\ 2 & 2 & -4 \\ 4 & 5 & -6 \end{vmatrix} \overset{\substack{S_2-S_1 \\ S_3-2S_1}}{=} \begin{vmatrix} 2 & 2 & -3 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{vmatrix} \overset{S_2 \leftrightarrow S_3}{=} - \begin{vmatrix} 2 & 2 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{vmatrix} = -(-2) = 2.$$

A következő determinánsnál sorcsere nélkül eliminálhatók a főátló alatti elemek, ezért a sorműveleteket nem is jelezzük.

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 2 & 5 & 9 \\ 0 & 3 & 9 & 19 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 3 & 10 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 1.$$

Egy érdekes észrevétel: a fenti determinánsban és sorlécsős alakjában is a Pascal-háromszög számai találhatók. Ezt kihasználva a fenti eredmény általánosítható, ahhoz azonban egy másik megoldás jobb esélyt ad: az első oszlop főátló alatti elemeit nullázzuk ki úgy, hogy először vonjuk ki az utolsó előtti sort az utolsóból, majd a második sort a harmadikból, végül az elsőt a másodikból, majd kövessük e módszert a többi főátló alatti elemre is:

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{vmatrix} \overset{S_4-S_3}{=} \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 0 & 1 & 4 & 10 \end{vmatrix} \overset{S_3-S_2}{=} \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 0 & 1 & 3 & 6 \\ 0 & 1 & 4 & 10 \end{vmatrix}$$

$$\overset{S_2-S_1}{=} \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 3 & 6 \\ 0 & 1 & 4 & 10 \end{vmatrix} \overset{\substack{S_4-S_3 \\ S_3-S_2}}{=} \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 \end{vmatrix} \overset{S_4-S_3}{=} \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 1.$$

Az általánosítás a 6.14. feladatban található. □

Mátrixműveletek és determináns Kérdés, hogy milyen kapcsolat van a mátrixműveletek és a determináns között. Fontos megjegyezni, hogy a determinánsfüggvénynek *nincs* a mátrixösszeadásra és a skalárral való szorzásra nézve művelettartó tulajdonsága, azaz általában $\det(\mathbf{A}+\mathbf{B}) \neq \det(\mathbf{A}) + \det(\mathbf{B})$, és $\det(c\mathbf{A}) \neq c\det(\mathbf{A})$. A determináns definíciója szerint lineáris tulajdonsága csak a sorvektorokra vonatkozóan van.

A skalárral való szorzás esetén viszont itt is mondható valami: mivel egy mátrix c-szeresének determinánsa minden sorából kiemelhető c, ez annyi kiemelést jelent, ahány sora van a mátrixnak. Így tetszőleges $n \times n$-es $\mathbf{A}$ mátrixra és tetszőleges c skalárra $\det(c\mathbf{A}) = c^n \det(\mathbf{A})$. Ez világos, ha $\mathbb{R}^2$- vagy $\mathbb{R}^3$-beli vektorokra gondolunk, hisz például egy parallelogramma előjeles területe 4-szeresére, egy parallelepipedon előjeles térfogata 8-szorosára nő, ha minden élét 2-szeresére növeljük.

A determináns művelettartó a négyzetes mátrixok szorzására nézve. Ezt mondja ki a következő állítás.

6.8. ÁLLÍTÁS (DETERMINÁNSOK SZORZÁSSZABÁLYA). *Ha* $\mathbf{A}$ *és* $\mathbf{B}$ *azonos méretű négyzetes mátrixok, akkor* $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$.

BIZONYÍTÁS. Tudjuk, hogy ha $\mathbf{A}$ szinguláris, akkor $\mathbf{AB}$ is, azaz ha $\det(\mathbf{A}) = 0$, akkor $\det(\mathbf{AB})$ is 0, tehát $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$. Ha $\mathbf{A}$ nem szinguláris, akkor felbontható elemi mátrixok szorzatára: $\mathbf{A} = \mathbf{E}_1\mathbf{E}_2\dots\mathbf{E}_k$, így $\mathbf{AB} = \mathbf{E}_1\mathbf{E}_2\dots\mathbf{E}_k\mathbf{B}$. A 6.4. állítás szerint tetszőleges $\mathbf{E}$ elemi mátrixra $\det(\mathbf{EB}) = \det(\mathbf{E})\det(\mathbf{B})$. Ezt az összefüggést $\mathbf{E}_1\mathbf{E}_2\dots\mathbf{E}_k$-ra és $\mathbf{E}_1\mathbf{E}_2\dots\mathbf{E}_k\mathbf{B}$-re is használva kapjuk, hogy

$$\begin{aligned}
\det(\mathbf{A})\det(\mathbf{B}) &= \det(\mathbf{E}_1\mathbf{E}_2\dots\mathbf{E}_k)\det(\mathbf{B}) = \det(\mathbf{E}_1)\det(\mathbf{E}_2\dots\mathbf{E}_k)\det(\mathbf{B}) \\
&= \det(\mathbf{E}_1)\det(\mathbf{E}_2)\det(\mathbf{E}_3\dots\mathbf{E}_k)\det(\mathbf{B}) = \dots = \\
&= \det(\mathbf{E}_1)\det(\mathbf{E}_2)\dots\det(\mathbf{E}_k)\det(\mathbf{B}), \text{ másrészt} \\
\det(\mathbf{AB}) &= \det(\mathbf{E}_1\mathbf{E}_2\dots\mathbf{E}_k\mathbf{B}) = \det(\mathbf{E}_1)\det(\mathbf{E}_2\dots\mathbf{E}_k\mathbf{B}) \\
&= \det(\mathbf{E}_1)\det(\mathbf{E}_2)\det(\mathbf{E}_3\dots\mathbf{E}_k\mathbf{B}) = \dots = \\
&= \det(\mathbf{E}_1)\det(\mathbf{E}_2)\dots\det(\mathbf{E}_k)\det(\mathbf{B}),
\end{aligned}$$

ami bizonyítja az állítást. Egy másik, nagyon szép bizonyítás található a 6.11. feladatban. □

A determinánsok szorzásszabályának egy fontos alkalmazása a determináns értékének kiszámítása PLU-felbontással (ld. 6.5. kód és a program).

```
sage: M = matrix(3,range(9))
sage: M[2,2]=9
sage: N=M.change_ring(RDF)
sage: N
[0.0 1.0 2.0]
[3.0 4.0 5.0]
[6.0 7.0 9.0]
sage: N.det()
-3.0
sage: P,L,U = N.LU()
sage: P
[0.0 0.0 1.0]
[1.0 0.0 0.0]
[0.0 1.0 0.0]
sage: U
[ 6.0  7.0  9.0]
[ 0.0  1.0  2.0]
[ 0.0  0.0 -0.5]
sage: P.det()
1.0
sage: U.det()
-3.0
```

6.5. kód: Determináns kiszámítása a PLU-felbontásból. A felbontás az egészek gyűrűjében nem működik, ezért gyűrűt váltunk és dupla pontosságú lebegőpontos számokkal számolunk (RDF).

6.9. PÉLDA (DETERMINÁNS KISZÁMOLÁSA PLU-FELBONTÁSBÓL). *Hogyan határozzuk meg egy* $\mathbf{A}$ *mátrix determinánsát, ha ismerjük PLU-felbontását? Konkrétan mennyi a következő mátrix determinánsa?*

$$\begin{bmatrix} 0 & 1 & 2 \\ 3 & 5 & 6 \\ 4 & 7 & 9 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3/4 & -1/4 & 1 \end{bmatrix} \begin{bmatrix} 4 & 7 & 9 \\ 0 & 1 & 2 \\ 0 & 0 & -1/4 \end{bmatrix}$$

MEGOLDÁS. Egy PLU-felbontásban szereplő mindegyik mátrix determinánsa könnyen meghatározható. $\mathbf{P}$ két sorcserével egységmátrixszá válik, tehát $\det \mathbf{P} = 1$. $\mathbf{L}$ és $\mathbf{U}$ háromszögmátrixok, amelyek determinánsa a főátlóbeli elemek szorzata, ami $\mathbf{L}$ esetén mindig 1. A megadott konkrét esetben tehát $\det \mathbf{A} = 4 \cdot 1 \cdot (-1/4) = -1$. □

Mátrix determinánsa és transzponáltjának determinánsa megegyezik. Ez lehetővé teszi, hogy a determináns kiszámításához nem csak az elemi sor-, de az elemi oszlopműveleteket is használjuk, hisz egy mátrixon végzett oszlopművelet a transzponált sorművelete.

6.10. ÁLLÍTÁS (TRANSZPONÁLT DETERMINÁNSA). *Mátrix determinánsa megegyezik transzponáltjának determinánsával, azaz bármely négyzetes* $\mathbf{A}$ *mátrixra* $\det(\mathbf{A}) = \det(\mathbf{A}^T)$.

BIZONYÍTÁS. Az $\mathbf{A}$ mátrix redukált lépcsős alakra hozásának mátrixszorzatos alakja legyen $\mathbf{A} = \mathbf{E}_1\mathbf{E}_2 \ldots \mathbf{E}_k\mathbf{R}$, ahol $\mathbf{E}_i$ elemi mátrix, $\mathbf{R}$ az $\mathbf{A}$ redukált lépcsős alakja. A transzponált determinánsa

$$|\mathbf{A}^T| = |\mathbf{R}^T\mathbf{E}_k^T \ldots \mathbf{E}_2^T\mathbf{E}_1^T| = |\mathbf{R}^T||\mathbf{E}_k^T| \ldots |\mathbf{E}_2^T||\mathbf{E}_1^T|.$$

Könnyen ellenőrizhető, hogy minden elemi mátrix determinánsa megegyezik transzponáltjának determinánsával (ellenőrizzük!). Mivel $\mathbf{R}$ redukált lépcsős alak, ezért $\mathbf{R} = \mathbf{I}$, vagy $\mathbf{R}$-nek van egy zérus sora. Ha $\mathbf{R} = \mathbf{I}$, akkor $|\mathbf{R}^T| = |\mathbf{R}| = |\mathbf{I}| = 1$, ha pedig $\mathbf{R}$-nek van zérus sora, akkor $\mathbf{R}^T$-nak zérus oszlopa, és egy ilyen mátrix nem alakítható elemi sorműveletekkel egységmátrixszá, tehát determinánsa 0. Azaz $|\mathbf{R}| = |\mathbf{R}^T|$ ekkor is fönnáll. Ekkor pedig

$$\begin{aligned}&|\mathbf{R}^T||\mathbf{E}_k^T| \ldots |\mathbf{E}_2^T||\mathbf{E}_1^T| = |\mathbf{R}||\mathbf{E}_k| \ldots |\mathbf{E}_2||\mathbf{E}_1| = \\ &|\mathbf{E}_1||\mathbf{E}_2| \ldots |\mathbf{E}_k||\mathbf{R}| = |\mathbf{E}_1\mathbf{E}_2 \ldots \mathbf{E}_k\mathbf{R}| = |\mathbf{A}|.\end{aligned}$$

Tehát $|\mathbf{A}^T| = |\mathbf{A}|$. □

6.11. PÉLDA (DETERMINÁNS KISZÁMÍTÁSA ELEMI OSZLOPMŰVELETEKKEL). *Az alábbi determinánst elemi sor- és oszlopműveletek alkalmazásával 2 lépésben is kiszámíthatjuk:*

$$\begin{vmatrix}1&0&0&1&0\\2&2&1&3&1\\1&1&1&1&0\\1&1&1&2&0\\1&1&1&2&1\end{vmatrix} \overset{S_2-S_5}{=} \begin{vmatrix}1&0&0&1&0\\1&1&0&1&0\\1&1&1&1&0\\1&1&1&2&0\\1&1&1&2&1\end{vmatrix} \overset{O_4-O_1}{=} \begin{vmatrix}1&0&0&0&0\\1&1&0&0&0\\1&1&1&0&0\\1&1&1&1&0\\1&1&1&1&1\end{vmatrix} = 1$$

Mikor 0 a determináns értéke Gyakran vízválasztó, hogy egy determináns értéke zérus-e. A determináns definíciója és a 6.2. állítás alapján eddig annyit tudunk, hogy a determináns 0, ha van két azonos sora, vagy egy zérussora. Most szükséges és elégséges feltételeket adunk.

6.12. TÉTEL (ZÉRUS ÉRTÉKŰ DETERMINÁNS). *Legyen* $\mathbf{A}$ *négyzetes mátrix. A következő állítások ekvivalensek:*

1. $\det(\mathbf{A}) = 0$,
2. $\mathbf{A}$ *sorvektorai lineárisan összefüggők.*
3. $\mathbf{A}$ *szinguláris,*
4. *a homogén lineáris* $\mathbf{A}\mathbf{x} = \mathbf{0}$ *egyenletrendszernek van nemtriviális megoldása.*

BIZONYÍTÁS. A ??? tételben láttuk, hogy négyzetes mátrix sorvektorai pontosan akkor lineárisan összefüggők, ha a mátrix szinguláris, azaz ha a lépcsős alakra hozás során keletkezik egy 0-sor, ez pedig azzal ekvivalens, hogy a determináns értéke 0. Az utolsó állítás ekvivalenciája a mátrix invertálhatóságáról szóló 5.20. tétel közvetlen következménye. □

6.13. PÉLDA (ZÉRUS ÉRTÉKŰ DETERMINÁNSOK). *A sorvektorok lineáris összefüggőségének igazolásával mutassuk meg, hogy*

$$\begin{vmatrix} 5 & 6 & 8 \\ 2 & 1 & 2 \\ 3 & 5 & 6 \end{vmatrix} = 0, \qquad \begin{vmatrix} 2 & -1 & 0 & -1 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ -1 & 0 & -1 & 2 \end{vmatrix} = 0.$$

MEGOLDÁS. Az első determináns első sora a második és a harmadik összege. De fogalmazhatunk úgy is, hogy az első sorból kivonva a másodikat és a harmadikat, a nullvektort kapjuk. Tehát az első mátrix sorvektorai lineárisan összefüggők, így determinánsa 0.

A második determináns sorvektorainak összege a nullvektor, tehát ennek is 0 az értéke.

A legegyszerűbb eseteket leszámítva a sorvektorok lineáris összefüggősége „ránézésre” nem látható, de az összefüggőséget bizonyító skalárok – ha szükségünk van rá – megkaphatók az $\mathbf{A}^T\mathbf{x} = \mathbf{0}$ egyenletrendszer nemtriviális megoldásaiból. □

Az előző tétel, valamit az 5.20. tétel fontos következménye a determinánsnak az egyenletrendszerek megoldhatóságával való kapcsolatáról szól:

6.14. TÉTEL (EGYENLETRENDSZER MEGOLDHATÓSÁGA ÉS A DETERMINÁNS). *Legyen* $\mathbf{A}$ *négyzetes mátrix. Ekkor az alábbi állítások ekvivalensek:*

1. $\det \mathbf{A} \neq 0$,
2. *az* $\mathbf{Ax} = \mathbf{b}$ *egyenletrendszer egyértelműen megoldható,*
3. *az* $\mathbf{Ax} = \mathbf{0}$ *egyenletrendszernek csak triviális megoldása van.*

A gyakorlatban – például mért adatok esetén – az, hogy egy determináns nulla-e, nehezen dönthető el! Fontos tudni, hogy az, hogy a determináns értéke „közel van a nullához”, nem jelenti azt, hogy a determináns „közel szinguláris”. Például az

$$\begin{vmatrix} \frac{1}{n} & 0 \\ 0 & n \end{vmatrix} = 1, \quad \text{és az} \quad \begin{vmatrix} \frac{1}{2} & 0 & \dots & 0 \\ 0 & \frac{1}{2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{2} \end{vmatrix} = \frac{1}{2^n}$$

determinánsok közül az első értéke tetszőlegesen nagy n-re is 1, pedig $\frac{1}{n}$ tetszőlegesen közel lehet 0-hoz, és az $\left[\begin{smallmatrix} 0 & 0 \\ 0 & n \end{smallmatrix}\right]$ mátrix már szinguláris. A második determinánsbeli $\frac{1}{2}\mathbf{I}_n$ mátrix nem szinguláris, pedig determinánsának értéke tetszőlegesen közel lehet 0-hoz, igaz, csak elegendően nagy n esetén.

A determináns minden sorában (sorvektorában) lineáris, ami lehetővé teszi a determináns előállítását determinánsok lineáris kombinációjaként. Két ilyen módszert ismertetünk a következő szakaszban. Ezek igen fontosak, gyakran ezek segítségével definiálják a determináns fogalmát.

Feladatok

6.1.• Melyek igazak az alábbi állítások közül? (Az $\mathbf{A}$ itt mindig négyzetes mátrixot jelöl.)

1. Ha egy determináns értéke 0, akkor van két azonos sora.
2. Ha egy determináns értéke nem 0, akkor oszlopvektorai lineárisan függetlenek.
3. Ha az $\mathbf{Ax} = \mathbf{0}$ egyenletrendszernek van nemtriviális megoldása, akkor $|\mathbf{A}| \neq 0$.
4. $|\mathbf{A}| \neq 0$ pontosan akkor igaz, ha az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszer nem oldható meg.
5. $|\mathbf{A}| = 0$ pontosan akkor igaz, ha az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszer egyértelműen megoldható.

6.2.♥ Számítsuk ki az alábbi determinánsok értékét fejben!

a) $\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$ *b)* $\begin{vmatrix} 0 & 0 & 0 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{vmatrix}$ *c)* $\begin{vmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{vmatrix}$

d) $\begin{vmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{vmatrix}$ *e)* $\begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix}$ *f)* $\begin{vmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 3 & 3 & 3 \end{vmatrix}$

g) $\begin{vmatrix} 1 & -2 & 1 & 2 \\ 4 & 3 & 4 & -1 \\ 5 & 1 & 5 & 4 \\ 6 & 5 & 6 & 0 \end{vmatrix}$

6.3.♥ Mutassuk meg – lineáris összefüggőséget keresve a sorok közt –, hogy az alábbi determinánsok értéke 0.

a) $\begin{vmatrix} 1 & 1 & 1 \\ 2 & 3 & 5 \\ 3 & 4 & 6 \end{vmatrix}$ *b)* $\begin{vmatrix} 1 & -2 & 3 \\ -2 & 4 & -6 \\ 3 & 6 & 9 \end{vmatrix}$ *c)* $\begin{vmatrix} 2 & 1 & 0 \\ 3 & 2 & 1 \\ 5 & 3 & 1 \end{vmatrix}$

d) $\begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{vmatrix}$ *e)* $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$ *f)* $\begin{vmatrix} 1 & 1 & -2 \\ 1 & -2 & 1 \\ -2 & 1 & 1 \end{vmatrix}$

g) $\begin{vmatrix} \sin\alpha & \cos\alpha & \sin(\alpha+\delta) \\ \sin\beta & \cos\beta & \sin(\beta+\delta) \\ \sin\gamma & \cos\gamma & \sin(\gamma+\delta) \end{vmatrix}$ *h)* $\begin{vmatrix} \ln 10 & \ln 4 & \ln 40 \\ \ln 5 & \ln 4 & \ln 20 \\ \ln 2 & 0 & \ln 2 \end{vmatrix}$

6.4.♥ Fölhasználva, hogy

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = -2.$$

számítsuk ki az alábbi determinánsok értékét:

a) $\begin{vmatrix} a & b & c \\ g & h & i \\ d & e & f \end{vmatrix}$ *b)* $\begin{vmatrix} a & b & c \\ 2d & 2e & 2f \\ g & h & i \end{vmatrix}$

c) $\begin{vmatrix} a & b & c \\ a+d & b+e & c+f \\ g & h & i \end{vmatrix}$ *d)* $\begin{vmatrix} a & d & g \\ b & e & h \\ c & f & i \end{vmatrix}$

e) $\begin{vmatrix} a & 3b & c \\ d & 3e & f \\ g & 3h & i \end{vmatrix}$ *f)* $\begin{vmatrix} a & b & c+a \\ d & e & f+d \\ g & h & i+g \end{vmatrix}$

g) $\begin{vmatrix} 2a & 3b & c+a \\ 2d & 3e & f+d \\ 2g & 3h & i+g \end{vmatrix}$ *h)* $\begin{vmatrix} 2a & 2b & 2c \\ 3d & 3e & 3f \\ g+4a & h+4b & i+4c \end{vmatrix}$

6.5.♥ Legyen $\mathbf{A}$ és $\mathbf{B}$ két 3×3-as mátrix, és legyen $\det(\mathbf{A}) = 5$, $\det(\mathbf{B}) = 4$. Számítsuk ki a következő determinánsok értékét!

a) $\det(\mathbf{A}^2)$ *b)* $\det(2\mathbf{A})$ *c)* $\det((2\mathbf{A})^2)$
d) $\det(\mathbf{A}^{-1})$ *e)* $\det(5\mathbf{A}^{-1})$ *f)* $\det((5\mathbf{A})^{-1})$
g) $\det(\mathbf{AB}^{-1})$ *h)* $\det(\mathbf{A}^T\mathbf{B})$ *i)* $|\mathbf{A}^{-1}||\mathbf{B}^{-1}\mathbf{A}||\mathbf{B}|$

6.6.♥ Csak sorcserék segítségével hozzuk egyszerűbb alakra (például háromszögalakra) az alábbi determinánsokat, és így számítsuk ki értéküket:

a) $\begin{vmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 1 & 0 & 0 \end{vmatrix}$ *b)* $\begin{vmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 4 & 0 \\ 5 & 0 & 0 & 0 & 0 \end{vmatrix}$

c) $\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}$, $\begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix}$, $\begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{vmatrix}$

d) $\begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 5 \\ 0 & 3 & 6 & 8 \\ 4 & 7 & 9 & 2 \end{vmatrix}$ *e)* $\begin{vmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 0 \\ 3 & 3 & 0 & 0 \\ 4 & 0 & 0 & 0 \end{vmatrix}$

f) $\begin{vmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 1 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{vmatrix}$ *g)* $\begin{vmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 1 & \dots & 1 & 1 \\ 1 & 1 & \dots & 1 & 1 \end{vmatrix}$

6.7.♥ Számítsuk ki elemi sorműveletekkel az alábbi determinánsokat!

a) $\begin{vmatrix} 1 & 2 & 3 \\ 1 & 3 & 5 \\ 1 & 3 & 6 \end{vmatrix}$ *b)* $\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 5 & 7 \\ 1 & 4 & 7 & 10 \end{vmatrix}$

c) $\begin{vmatrix} 3 & 8 & 6 & 3 \\ 1 & 2 & 0 & 1 \\ 1 & 1 & -1 & 2 \\ 2 & 5 & 1 & 5 \end{vmatrix}$ *d)* $\begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 0 & 0 & 0 & 4 \\ 3 & 0 & 1 & 0 & 3 \\ 4 & 0 & 0 & 0 & 2 \\ 5 & 4 & 3 & 2 & 1 \end{vmatrix}$

6.8.☆ Számítsuk ki elemi sorműveletekkel az alábbi n-edrendű determinánsokat!

a) $\begin{vmatrix} 1+x_1y_1 & 1+x_1y_2 & \dots & 1+x_1y_n \\ 1+x_2y_1 & 1+x_2y_2 & \dots & 1+x_2y_n \\ \vdots & \vdots & \ddots & \vdots \\ 1+x_ny_1 & 1+x_ny_2 & \dots & 1+x_ny_n \end{vmatrix}$

b) $\begin{vmatrix} 1 & a & a^2 & \dots & a^{n-1} \\ a^{n-1} & 1 & a & \dots & a^{n-2} \\ a^{n-2} & a^{n-1} & 1 & \dots & a^{n-3} \\ \vdots & \vdots & & \ddots & \vdots \\ a & a^2 & a^3 & \dots & 1 \end{vmatrix}$

c) $\begin{vmatrix} a & b & b & \dots & b \\ b & a & b & \dots & b \\ b & b & a & \dots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \dots & a \end{vmatrix}$ d) $\begin{vmatrix} a & b & b & \dots & b \\ c & a & b & \dots & b \\ c & c & a & \dots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c & c & c & \dots & a \end{vmatrix}$

6.9.■ Számítsuk ki a Petersen-gráf szomszédsági mátrixának (ld. **??** feladat) determinánsát!

6.10.• Írjuk fel a determináns definícióját oly módon, hogy det egy n-változós, n-dimenziós vektorokon értelmezett skalár értékű függvény legyen.

6.11. Adjunk új bizonyítást a determinánsok szorzásszabályára azt igazolva, hogy az $\mathbf{A} \mapsto \det(\mathbf{AB})/\det(\mathbf{B})$ leképezés eleget tesz a determináns definíciójában kirótt feltételeknek.

6.12.⋆ Bizonyítsuk be az LU-felbontás fölhasználásával a transzponált determinánsára vonatkozó 6.10. állítást.

6.13. Fejezzük ki az elemi mátrixokra használt jelöléseket használva ($\mathbf{E}_{S_i+cS_j}$, $\mathbf{E}_{S_i\leftrightarrow S_j}$, $\mathbf{E}_{cS_i}$) azok determinánsát!

6.14.⋆ Számítsuk ki az

$$\left|\binom{i+j-2}{j-1}\right|_{n\times n} = \begin{vmatrix} \binom{0}{0} & \binom{1}{1} & \dots & \binom{n-1}{n-1} \\ \binom{1}{0} & \binom{2}{1} & \dots & \binom{n}{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \binom{n-1}{0} & \binom{n}{1} & \dots & \binom{2n-2}{n-1} \end{vmatrix}$$

determináns értékét! (Útmutatás: az utolsó sorral kezdve mindegyik sorból vonjuk ki az előzőt, majd mindegyik oszlopból az előzőt!)

6.15. Mutassuk meg, hogy egy legalább 3-adrendű determináns értéke 0, ha elemei sorfolytonosan olvasva számtani sorozatot adnak. Például

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 0.$$

6.16. Mutassuk meg, hogy egy legalább 3-adrendű determináns értéke 0, ha minden sora számtani sorozat, például

$$\begin{vmatrix} 1 & 2 & 3 \\ 1 & 4 & 7 \\ 1 & 3 & 5 \end{vmatrix} = 0.$$

6.17. Mutassuk meg, hogy ha egy determináns elemei sorfolytonosan olvasva mértani sorozatot alkotnak, akkor értéke 0. Például

$$\begin{vmatrix} \frac18 & \frac14 & \frac12 \\ 1 & 2 & 4 \\ 8 & 16 & 32 \end{vmatrix} = 0.$$

6.18. Mutassuk meg, hogy tetszőleges a, b, c és d valósokra

$$\begin{vmatrix} a^2 & (a+1)^2 & (a+2)^2 & (a+3)^2 \\ b^2 & (b+1)^2 & (b+2)^2 & (b+3)^2 \\ c^2 & (c+1)^2 & (c+2)^2 & (c+3)^2 \\ d^2 & (d+1)^2 & (d+2)^2 & (d+3)^2 \end{vmatrix} = 0.$$

6.19.• Mutassuk meg, hogy ha $\mathbf{C}$ invertálható, akkor $\det(\mathbf{CAC}^{-1}) = \det(\mathbf{A})$ tetszőleges azonos méretű $\mathbf{A}$ mátrixra fennáll.

6.20.• VEKTOROK DETERMINÁNSA MÁSIK BÁZISBAN Igazoljuk, hogy ha $\mathbf{A}_{\mathcal{C}\leftarrow\mathcal{B}}$ az áttérés mátrixa, akkor a $\mathbf{v}_1$, $\mathbf{v}_2,\dots$, $\mathbf{v}_n$ vektorok $\mathcal{B}$- és $\mathcal{C}$-beli koordinátás alakjaiból képzett $\mathbf{V}_\mathcal{B}$ és $\mathbf{V}_\mathcal{C}$ mátrixok determinánsára $|\mathbf{V}_\mathcal{C}| = |\mathbf{A}_{\mathcal{C}\leftarrow\mathcal{B}}||\mathbf{V}_\mathcal{B}|$.

6.21.• Igazoljuk, hogy páratlan rendű ferdén szimmetrikus mátrix determinánsa 0.

6.22.• MÁTRIX NÉGYZETÉNEK DETERMINÁNSA Igazoljuk, hogy bármely négyzetes $\mathbf{A}$ mátrixra $|\mathbf{A}^2| = |\mathbf{AA}^T|$.

6.23. A determináns négyzetének kiszámításával (6.22. feladat) és a determinánsok szorzástételének alkalmazásával számítsuk ki az alábbi determinánsok értékét:

$$\begin{vmatrix} a & b \\ -b & a \end{vmatrix}, \quad \begin{vmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{vmatrix},$$

$$\begin{vmatrix} a & b & c & d & e & f & g & h \\ b & -a & -d & c & f & -e & h & -g \\ c & d & -a & -b & g & -h & -e & f \\ d & -c & b & -a & h & g & -f & -e \\ e & -f & -g & -h & -a & b & c & d \\ f & e & h & -g & -b & -a & d & -c \\ g & -h & e & f & -c & -d & -a & b \\ h & g & -f & e & -d & c & -b & -a \end{vmatrix}.$$

6.24. Mutassuk meg, hogy az $(x_1^2+x_2^2)(y_1^2+y_2^2)$ szorzat előállítható két szám négyzetének összegeként, azaz

$$(x_1^2+x_2^2)(y_1^2+y_2^2) = (z_1^2+z_2^2),$$

ahol z_1 és z_2 mindegyike külön az x_i és külön az y_i változóknak is lineáris kifejezése ($i = 1, 2$). (Hasonló összefüggések bizonyíthatóak négy illetve nyolc négyzetszám összegéről is. Például a négy szám négyzetösszegére vonatkozó képlet

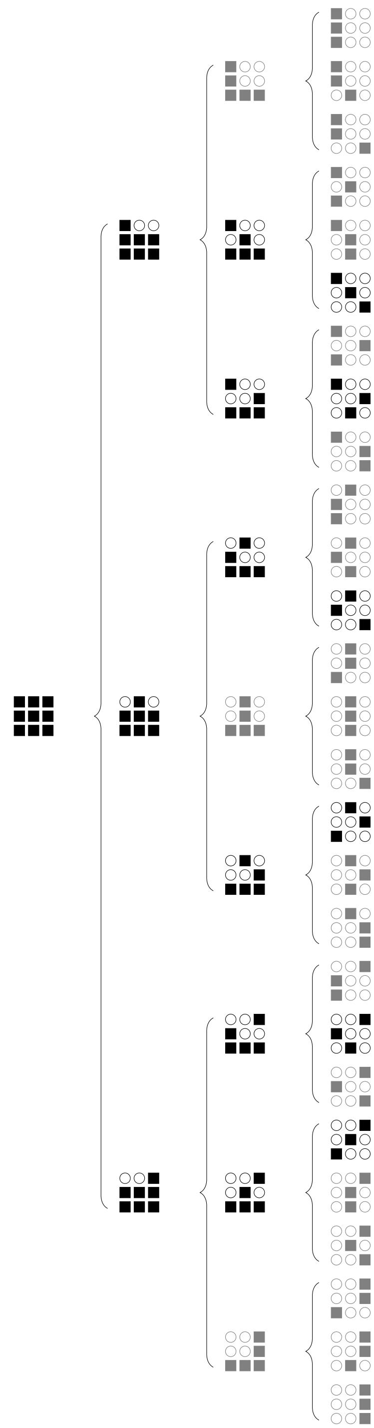
$$(x_1^2 + x_2^2 + x_3^2 + x_4^2)(y_1^2 + y_2^2 + y_3^2 + y_4^2) = (z_1^2 + z_2^2 + z_3^2 + z_4^2),$$

ahol z_i az x_i és az y_i $(i = 1, 2, 3, 4)$ változókban lineáris. A megoldáshoz használjuk fel az előző feladat állítását.)

A determináns, mint elemeinek függvénye

A determinánst eddig sorvektorainak függvényeként kezeltük, a következőkben elemeinek függvényeként fogjuk.

Eddig nagyvonalúan bántunk a determináns elemeinek mibenlétével. Annyit feltételeztünk róluk kimondatlanul, hogy azonos algebrai struktúrából valók, és az összeadás, kivonás, szorzás és osztás elvégezhető köztük. E szakaszban ki fog derülni, hogy a determináns kiszámolható osztás nélkül is. Tehát nem csak testek (pl. a valós $\mathbb{R}$, a racionális $\mathbb{Q}$, a komplex $\mathbb{C}$ számtestek vagy a véges $\mathbb{F}_q$ testek) elemeiből álló determinánst számolhatunk ki az adott struktúrán belül, hanem pl. az egészek $\mathbb{Z}$ gyűrűjének vagy a legfeljebb k-adfokú polinomok gyűrűjének elemeiből képzett determinánsokat is. (További részletekért lásd a függelék *A* szakaszát.)



6.6. ábra: Egy 3×3-as determináns felbontása 3^3 determináns összegére, melyek közül $3! = 6$ darabot kivéve mindegyikben van egy zérusoszlop – ezek sematikus ábráját szürke szín jelöli.

Kígyók determinánsa A 2×2-es determináns kiszámítására ismerjük azt a formulát, amely a determináns értékét a determináns elemeinek függvényében írja fel: $\det\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] = ad - bc$. Itt tehát csak az összeadásra, kivonásra és a szorzásra van szükség. Hasonló formulát keresünk az n-edrendű determinánsokra. Ehhez a kígyókat használjuk.

Minden kígyó megkapható egy diagonális mátrix sorainak permutációjával, azaz minden $\mathbf{K}$ kígyó felírható $\mathbf{K} = \mathbf{P}\,\mathrm{diag}(a_1, a_2, \ldots, a_n)$ alakban, ahol $\mathbf{P}$ egy permutációs mátrix. Ezt a *kígyóhoz tartozó permutációs mátrixnak* fogjuk nevezni. Mivel $\mathbf{P}$ determinánsa 1 vagy -1, ezért $|\mathbf{K}| = a_1 a_2 \ldots a_n$ vagy $|\mathbf{K}| = -a_1 a_2 \ldots a_n$.

A determinánsok soronkénti linearitását használva érdekes felbontását kapjuk a determinánsnak. Tekintsük példaként az

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$$

determinánst. Első sorvektorának

$$(a, b, c) = (a, 0, 0) + (0, b, 0) + (0, 0, c)$$

felbontását fölhasználva bontsuk fel a determinánst három determináns összegére:

$$\begin{vmatrix} a+0+0 & 0+b+0 & 0+0+c \\ d & e & f \\ g & h & i \end{vmatrix} = \begin{vmatrix} a & 0 & 0 \\ d & e & f \\ g & h & i \end{vmatrix} + \begin{vmatrix} 0 & b & 0 \\ d & e & f \\ g & h & i \end{vmatrix} + \begin{vmatrix} 0 & 0 & c \\ d & e & f \\ g & h & i \end{vmatrix}$$

Ezután folytassuk e felbontást a második sorvektorral, így már az eredeti determinánst 9 determináns összegére bontottuk. Végül tegyük ugyanezt az utolsó sorral is. Az így kapott 27 determinánst nem írjuk föl, de szemléltetésül egy sematikus ábrán megmutatjuk a felbontás lépéseit (6.6 ábra). Tömör négyzet jelöli azokat a helyeket, ahol megtartjuk a determináns eredeti elemét, üres kör azokat, ahová zérust írunk. A 27 determináns mindegyikének minden sorában egy elem az eredeti determinánsból való, a többi zérus. Közöttük azonban csak 6 kígyó van. A többinek van zérus oszlopa, így azok értéke 0, vagyis az eredeti determinánst 6 kígyó összegére bontottuk (a 0 értékű determinánsokat szürke színnel jeleztük). Hasonló módon bármely n-edrendű determináns fölbomlik n^n olyan determináns összegére, melynek minden sorában egyetlen elem az eredeti determinánsból való, a többi 0, de ezek közül csak azok lesznek kígyók determinánsai, melyek minden oszlopában is van egy elem az eredetiből. (Ezeket nevezzük a mátrixból/determinánsból kiválasztható kígyóknak.) Ezek száma $n!$, mert az első sorból n-féleképp választhatunk egy elemet, a második sorból minden esetben már csak $n-1$-féleképp,..., és ez összesen $n(n-1)\dots 3\cdot 2\cdot 1 = n!$ eset. Igaz tehát a következő állítás:

6.15. ÁLLÍTÁS (FELBONTÁS KÍGYÓK DETERMINÁNSAINAK ÖSSZEGÉRE). *Minden n-edrendű determináns fölbomlik az összes belőle kiválasztható kígyó determinánsának összegére. Ha az $a_{1j_1}, a_{2j_2}, \dots, a_{nj_n}$ elemeket tartalmazó kígyóhoz tartozó permutációs mátrix determinánsát $d_{j_1j_2\dots j_n}$ jelöli (ennek értéke $+1$ vagy -1), akkor*

$$\det([a_{ij}]) = \sum d_{j_1j_2\dots j_n} a_{1j_1} a_{2j_2} \dots a_{nj_n},$$

ahol az összegzés az $\{1,2,\dots,n\}$ halmaz összes lehetséges $\{j_1, j_2, \dots, j_n\}$ permutációján végigfut.

Az $n!$ az n növekedtével rendkívül gyorsan nő (pl. $10! = 3628800$), determináns ilyen módon való számítása viszonylag kis rend esetén már számítógéppel sem lehetséges emberi idő alatt. E felbontást a determinánsok tulajdonságainak vizsgálatában használjuk. Számításhoz csak az $n = 2$ és $n = 3$ esetekben használjuk, igaz, azokra gyakran. $n = 2$ esetén az előző állítás szerint

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} = ad - bc,$$

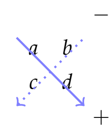
mivel a második determináns egyetlen sorcserével hozható diagonális alakra. $n = 3$ esetén – felhasználva a 6.6 ábrát is – kapjuk, hogy

$$\begin{aligned}
\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} &= \begin{vmatrix} a & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & i \end{vmatrix} + \begin{vmatrix} a & 0 & 0 \\ 0 & 0 & f \\ 0 & h & 0 \end{vmatrix} + \begin{vmatrix} 0 & b & 0 \\ d & 0 & 0 \\ 0 & 0 & i \end{vmatrix} \\
&\quad + \begin{vmatrix} 0 & b & 0 \\ 0 & 0 & f \\ g & 0 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 & c \\ d & 0 & 0 \\ 0 & h & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 & c \\ 0 & e & 0 \\ g & 0 & 0 \end{vmatrix} \\
&= aei - afh - bdi + bfg + cdh - ceg \\
&= aei + bfg + cdh - afh - bdi - ceg
\end{aligned}$$

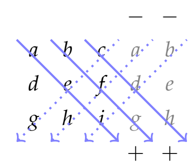
E két formula könnyen megjegyezhető egy egyszerű szabállyal, amelyet az $n = 2$ és $n = 3$ esetben *Sarrus-szabálynak* is neveznek: a főátló irányú szorzatok összegéből vonjuk ki a mellékátló irányú szorzatokat. (Hogy mit értsünk főátló és mellékátló irányú szorzaton, a mellékelt ábrákról megérthető.) Fontos, hogy hasonló szabály $n > 3$ esetén már nem érvényes (ld. a 6.29. feladatot).

A determináns 6.15. tételbeli felbontása a determináns értékét a determináns elemeinek függvényeként állítja elő. Ennek sok szép és fontos következménye van. Íme kettő:

- ▶ Egy algebrai következmény: a determináns kiszámolásához elég csak az összeadás és szorzás művelete, az osztásra, melyet az elemi sorműveletek során használhatunk, nincs szükség. Eszerint egész számokból álló determináns értéke egész szám.
- ▶ Egy függvényanalízis körébe tartozó következmény: a determináns értéke folytonos, sőt differenciálható függvénye elemeinek. Eszerint bármely kis pozitív ε-hoz van olyan $\delta > 0$ szám, hogy ha a determináns bármely eleme legföljebb δ értékkel megváltozik, akkor a determináns értéke legföljebb ε-nyit változik.

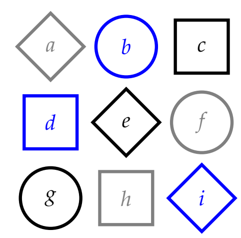


(a) $ad - bc$



(b) $aei + bfg + cdh - afh - bdi - ceg$

6.7. ábra: Az *(a)* másod- és a *(b)* harmadrendű determináns kiszámítása: a főátló irányú szorzatok összegéből vonjuk ki a mellékátló irányú szorzatokat. Harmadrendű esetben kezdetben könnyíthetünk magunknak a determináns első két oszlopának a determináns utáni megismétlésével.



6.8. ábra: A harmadrendű determináns kiszámítására egy – IQ-tesztek típuskérdésére emlékeztető – másik módszer: az egyforma alakúak szorzatának összegéből ki kell vonni az egyforma színűek szorzatait.

*Permutációs mátrix determinánsa** Kígyó determinánsának kiszámításában egyetlen bizonytalan pont maradt, a hozzá tartozó permutációs mátrix értékének kiszámítása. Vajon nem fordulhat-e elő, hogy páros és páratlan sok sorcserével is eljuthatunk egy permutációs mátrixból az identikusba.

Azt mondjuk, hogy egy permutációs mátrix két sora inverzióban áll, ha az előbb álló sorbeli 1-es hátrébb van, mint a másik sorbeli. Például a

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

mátrix inverzióinak száma 4, mert az első-második, első-negyedik, második-negyedik, harmadik-negyedik sorpárok inverzióban vannak.

6.16. TÉTEL (PERMUTÁCIÓS MÁTRIX DETERMINÁNSA). *A permutációs mátrix aszerint* $+1$ *vagy* -1*, hogy inverzióban álló sorpárjainak száma páros vagy páratlan.*

BIZONYÍTÁS. Elég megmutatni, hogy egy sorcsere mindig megváltoztatja az inverziók számának paritását, vagyis azok száma párosból páratlanra, páratlanból párosra változik. Így ha egy permutációs mátrix inverzióinak száma páros, akkor csak páros sok sorcserével vihető az identikus mátrixba. Hasonlóan, ha az inverziók száma páratlan, akkor csak páratlan sokkal.

Ha a két megcserélendő sor szomszédos, akkor a sorcsere megváltoztatja e két sor viszonyát: ha inverzióban álltak, akkor ezután nem fognak, és fordítva. Az előttük és mögöttük álló sorokhoz való viszonyuk nem változott. Eszerint az inverziók száma eggyel nőtt vagy eggyel csökkent, azaz paritása megváltozott.

Ezután cseréljük fel az i-edik és j-edik sorokat (legyen $i < j$). Az inverziók számának nyomon követése érdekében ezt szomszédos sorok cseréjével valósítjuk meg. Cseréljük ki az i-ediket az $(i+1)$-edikkel, majd azt az $(i+2)$-edikkel,..., míg az eredetileg i-edik sor a j-edik helyére nem kerül. Ehhez $j-i$ sorcserére van szükség. Ezután az eredetileg j-edik sort $j-i-1$ sorcserével az i-edik helyre visszük. Ez összesen $2(j-i)-1$, azaz páratlan sok sorcsere, ami a paritást valóban ellenkezőjére változtatja. □

6.17. PÉLDA (INVERZIÓK SZÁMA ÉS A DETERMINÁNS). *Hány sor áll inverzióban abban a mátrixban, melynek mellékátlójában egyesek, egyebütt nullák állnak, és mennyi ennek determinánsa?*

MEGOLDÁS. E mátrixban bármely két sor inverzióban áll egymással, így ha a sorok száma n, a sorpároké $n(n-1)/2$. Eszerint e mátrix determinánsa $(-1)^{n(n-1)/2}$. (Az egységmátrix $\lfloor \frac{n}{2} \rfloor$ sorcserével is megkapható e mátrixból, így determinánsát $(-1)^{\lfloor \frac{n}{2} \rfloor}$ alakban is ki lehet fejezni, ld. még a 6.6. feladatban). □

Végül egy fontos következmény:

6.18. KÖVETKEZMÉNY (DETERMINÁNSFÜGGVÉNY LÉTEZÉSE). *A determinánsfüggvény létezik, és egyértelmű.*

BIZONYÍTÁS. Mivel minden determináns egyenlő a belőle kiválasztható kígyók determinánsainak összegével, és minden kígyó determinánsa egyértelműen meghatározható, ezért minden négyzetes mátrix determinánsa egyértelmű. A kérdés már csak az, hogy az így meghatározott függvény valóban kielégíti-e a determináns definíciójának mindegyik pontját. Ez azonnal látszik, hisz ezeket elég csak kígyókra ellenőrizni, azokra pedig a linearitás és az azonos sorok feltétele is fönnáll. □

Előjeles aldetermináns Az előző paragrafushoz hasonlóan bontsuk az

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$$

determinánst, az első sorvektorának felbontásával három determináns összegére, de egyúttal emeljük is ki az első sor elemét, majd oszlopcserékkel vigyük az 1-est tartalmazó oszlopot az első oszlop helyére:

$$\begin{aligned}
\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} &= a\begin{vmatrix} 1 & 0 & 0 \\ d & e & f \\ g & h & i \end{vmatrix} + b\begin{vmatrix} 0 & 1 & 0 \\ d & e & f \\ g & h & i \end{vmatrix} + c\begin{vmatrix} 0 & 0 & 1 \\ d & e & f \\ g & h & i \end{vmatrix} \\
&= a\begin{vmatrix} 1 & 0 & 0 \\ d & e & f \\ g & h & i \end{vmatrix} - b\begin{vmatrix} 1 & 0 & 0 \\ e & d & f \\ h & g & i \end{vmatrix} + c\begin{vmatrix} 1 & 0 & 0 \\ f & d & e \\ i & g & h \end{vmatrix}
\end{aligned}$$

Ha ezt összevetjük a Sarrus-szabályban kapott képlettel, igen érdekes sejtést fogalmazhatunk meg:

$$\begin{aligned}
\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} &= aei + bfg + cdh - afh - bdi - ceg \\
&= a(ei - fh) - b(fg - di) + c(dh - eg) \\
&= a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}.
\end{aligned}$$

Mielőtt ezt megtennénk, némi előkészítés következik.

6.19. DEFINÍCIÓ (ELŐJELES ALDETERMINÁNS). *Az n-edrendű $|\mathbf{A}|$ determináns i-edik sorának és j-edik oszlopának elhagyásával kapott $(n-1)$-edrendű determináns $(-1)^{i+j}$-szeresét az $|\mathbf{A}|$ determináns a_{ij} eleméhez tartozó* előjeles aldeterminánsának *nevezzük.*

Az előjeles aldeterminánshoz kiszámítandó előjel a mátrixon sakktáblaszerűen változik, azaz a bal felső sarokban $+$, és két egymás melletti vagy alatti mezőben ellenkező előjelű. Ezt nevezik *sakktáblaszabálynak*.

+	−	+	−	+	−	+	−
−	+	−	+	−	+	−	+
+	−	+	−	+	−	+	−
−	+	−	+	−	+	−	+
+	−	+	−	+	−	+	−
−	+	−	+	−	+	−	+
+	−	+	−	+	−	+	−
−	+	−	+	−	+	−	+

6.9. ábra: Sakktáblaszabály: a $(-1)^{i+j}$ előjele a bal felső sarokban, vagyis az első sor első oszlopában $+$, él mentén szomszédos mezőkben pedig ellentétes.

6.20. PÉLDA (ELŐJELES ALDETERMINÁNS). *Számítsuk ki az*

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 9 & 1 \\ 2 & 2 & 2 & 2 \\ 0 & 1 & 2 & 0 \end{vmatrix}$$

determináns második sor harmadik eleméhez tartozó előjeles aldeterminánsát!

MEGOLDÁS. A determináns második sorát és harmadik oszlopát kiemeltük

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 9 & 1 \\ 2 & 2 & 2 & 2 \\ 0 & 1 & 3 & 0 \end{vmatrix}$$

Az ezek elhagyása után megmaradó aldetermináns és -1 megfelelő hatványának szorzata, vagyis a kért előjeles aldetermináns

$$(-1)^{2+3}\begin{vmatrix} 1 & 2 & 4 \\ 2 & 2 & 2 \\ 0 & 1 & 0 \end{vmatrix} = -1 \cdot (-6) = 6.$$

Tehát a determináns második sor harmadik eleméhez tartozó előjeles aldeterminánsa -2. □

6.21. ÁLLÍTÁS (DETERMINÁNS RENDJÉNEK CSÖKKENTÉSE). *Tegyük fel, hogy az n-edrendű $|\mathbf{A}|$ determináns a_{ij} elemének sorában vagy oszlopában minden további elem 0. Jelölje A_{ij} az a_{ij} elemhez tartozó előjeles aldeterminánst. Ekkor*

$$|\mathbf{A}| = a_{ij}A_{ij}.$$

BIZONYÍTÁS. Legyen az $|\mathbf{A}|$ determináns i-edik sorában az a_{ij}-n kívül minden elem 0 (hasonlóan tárgyalható, ha a j-edik oszlopban vannak nullák). Cseréljük ki a j-edik oszlopot a $(j-1)$-edikkel, majd ezt a $(j-2)$-edikkel..., addig, míg az $\mathbf{A}_{*j}$ oszlop az első oszlopba nem kerül. Ez $j-1$ oszlopcserét jelent, azaz a determináns értéke $(-1)^{j-1}$-szeresére változik. Ezután hasonlóképp vigyük az i-edik sort szomszédos sorok cseréjével az első sorba. Ehhez $i-1$ csere szükséges, miközben a determináns értéke $(-1)^{i-1}$-szeresére változik.

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & a_{ij} & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{nn} \end{vmatrix} = (-1)^{j-1} \begin{vmatrix} a_{1j} & a_{11} & a_{12} & \dots & a_{1n} \\ a_{2j} & a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{ij} & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ a_{nj} & a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

$$= (-1)^{i-1}(-1)^{j-1} \begin{vmatrix} a_{ij} & 0 & 0 & \dots & 0 \\ a_{1j} & a_{11} & a_{12} & \dots & a_{1n} \\ a_{2j} & a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{nj} & a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

$$\stackrel{*}{=} (-1)^{i+j} a_{ij} \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ a_{1j} & a_{11} & a_{12} & \dots & a_{1n} \\ a_{2j} & a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{nj} & a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

$$\stackrel{**}{=} (-1)^{i+j} a_{ij} \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

$$= a_{ij} A_{ij}.$$

Az *-os egyenlőségnél kihasználtuk, hogy $i + j - 2$ és $i + j$ paritása azonos, tehát -1 kitevőjeként is azonos eredményt adnak, továbbá kiemeltük a_{ij}-t az első sorból. A **-os egyenlőség előtt álló determináns kiszámításához csak a másodiktól lefelé lévő sorokat kell használni, a végeredményt az első oszlop elemei nem befolyásolják, így az első sor és első oszlop elhagyásával kapott determináns értéke ugyanaz. Végül az így kapott determináns az előjellel együtt épp A_{ij}, és ezzel bizonyítottuk az állítást. □

6.22. PÉLDA (DETERMINÁNS RENDJÉNEK CSÖKKENTÉSE). *A determináns rendjének csökkentésével számítsuk ki az alábbi determináns értékét!*

$$\begin{vmatrix} 1 & 2 & 0 & 3 & 4 \\ 1 & 2 & 0 & 8 & 4 \\ 6 & 0 & 0 & 7 & 0 \\ 8 & 9 & 8 & 7 & 6 \\ 5 & 4 & 0 & 3 & 2 \end{vmatrix}.$$

MEGOLDÁS. Minden lépésben – esetleg egy apró átalakítás után – találunk egy sort vagy oszlopot, melyben csak egy nemnulla szám áll, így a determináns könnyen számolható:

$$
\begin{aligned}
\begin{vmatrix} 1 & 2 & 0 & 3 & 4 \\ 1 & 2 & 0 & 8 & 4 \\ 6 & 0 & 0 & 7 & 0 \\ 8 & 9 & \boxed{8} & 7 & 6 \\ 5 & 4 & 0 & 3 & 2 \end{vmatrix} &= (-1)^{4+3} \cdot 8 \begin{vmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 8 & 4 \\ 6 & 0 & 7 & 0 \\ 5 & 4 & 3 & 2 \end{vmatrix} \\
(S_2 - S_1) \qquad &= (-8) \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & \boxed{5} & 0 \\ 6 & 0 & 7 & 0 \\ 5 & 4 & 3 & 2 \end{vmatrix} \\
&= (-8) \cdot (-1)^{2+3} \cdot 5 \begin{vmatrix} 1 & 2 & 4 \\ \boxed{6} & 0 & 0 \\ 5 & 4 & 2 \end{vmatrix} \\
&= (-8) \cdot (-5) \cdot (-1)^{2+1} \cdot 6 \begin{vmatrix} 2 & 4 \\ 4 & 2 \end{vmatrix} \\
&= (-8) \cdot (-5) \cdot (-6) \cdot (-12) \\
&= 2880.
\end{aligned}
$$

□

Determináns kifejtése Ritkán adódik, hogy a determináns rendje az előző (6.21.) állítás segítségével csökkenthető, viszont fölhasználásával a determinánsok egy gyönyörű kifejtési tételét kapjuk.

E kifejtési tételt egyes könyvek *Laplace-féle kifejtési tételnek* nevezik, míg más könyvek csak ennek egy – a feladatok közt megtalálható – általánosítását hívják így, sok könyv pedig e tételbeli összefüggéssel definiálja a determinánst.

6.23. TÉTEL (DETERMINÁNSOK KIFEJTÉSI TÉTELE). *Egy determináns értéke megkapható úgy, hogy egy tetszőleges sorának vagy oszlopának minden elemét beszorozzuk a hozzá tartozó előjeles aldeterminánssal, és e szorzatokat összeadjuk. Képletben, az n-edrendű $|\mathbf{A}|$ determináns értéke i-edik sora szerint kifejtve*

$$|\mathbf{A}| = \sum_{k=1}^{n} a_{ik} A_{ik},$$

és j-edik oszlopa szerint kifejtve

$$|\mathbf{A}| = \sum_{k=1}^{n} a_{kj} A_{kj}.$$

BIZONYÍTÁS. Hasonlóan a korábbiakban látottakhoz, az i-edik sorvektor felbontásával a determinánst n olyan determináns összegére bontjuk, amelyek i-edik sorában csak egy elem származik az eredeti determinánsból, a többi 0. Az egyszerűség kedvéért e felbontást csak $n = 3$ és $i = 2$ esetére írjuk fel, de tetszőleges n-re ugyanígy megy. Ezután a 6.21. állítást alkalmazzuk mindegyik új determinánsra:

$$\begin{aligned}
|\mathbf{A}| &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & 0 & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\
&= a_{21}A_{21} + a_{22}A_{22} + a_{23}A_{23} \\
&= \sum_{k=1}^{3} a_{2k}A_{2k}.
\end{aligned}$$

A bizonyítás ugyanígy megy az oszlopokra is, amit példaként az $n = 3$, $j = 3$ esettel szemléltetünk:

$$\begin{aligned}
|\mathbf{A}| &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & 0 \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & 0 \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\
&= a_{13}A_{13} + a_{23}A_{23} + a_{33}A_{33} \\
&= \sum_{k=1}^{3} a_{k3}A_{k3}.
\end{aligned}$$

□

6.24. PÉLDA (KIFEJTÉSI TÉTEL). *Számítsuk ki az alábbi determináns értékét a kifejtési tételt használva!*

$$\begin{vmatrix} 3 & 2 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 2 \end{vmatrix}.$$

MEGOLDÁS. Érdemes e determinánst a harmadik oszlopa szerint kifejteni, mert ott két 0 is van, így a velük megszorzott aldeterminánsokat le sem kell írni.

$$\begin{vmatrix} 3 & 2 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 2 \end{vmatrix} = 1 \cdot \begin{vmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{vmatrix} - 1 \cdot \begin{vmatrix} 3 & 2 & 2 \\ 2 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 1 - 0 = 1.$$

□

Cramer-szabály és a mátrix inverze Eddig akár az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszer megoldására, akár az $\mathbf{A}$ mátrix inverzének kiszámítására olyan módszert használtunk, mely az elemi sorműveletek használatával csak egy algoritmust ad a számításokra, de nem adja meg a kapcsolatot (képletet) az adatok és a kiszámítandók közt. E paragrafusban ezt pótoljuk!

Jelölje $\mathbf{A}_{i,\mathbf{b}}$ azt a mátrixot, melyet akkor kapunk, ha az $\mathbf{A}$ mátrix i-edik oszlopának helyére a $\mathbf{b}$ vektort írjuk. Kifejtve

$$\mathbf{A}_{i,\mathbf{b}} = [\mathbf{a}_{*1} \ \ldots \ \mathbf{a}_{*,i-1} \ \mathbf{b} \ \mathbf{a}_{*,i+1} \ \ldots \ \mathbf{a}_{*n}].$$

E jelöléssel $\mathbf{I}_{i,\mathbf{x}}$ mátrixon az $[\mathbf{e}_{*1}\ \dots\ \mathbf{e}_{*,i-1}\ \mathbf{x}\ \mathbf{e}_{*,i+1}\ \dots\ \mathbf{e}_{*n}]$ mátrixot értjük.

6.25. TÉTEL (CRAMER-SZABÁLY). *Legyen* $\mathbf{A}$ *egy* $n \times n$*-es mátrix. Az* $\mathbf{Ax} = \mathbf{b}$ *egyenletrendszer pontosan akkor oldható meg egyértelműen, ha* $\det \mathbf{A} \neq 0$. *Ekkor a megoldás előáll*

$$x_i = \frac{\det \mathbf{A}_{i,\mathbf{b}}}{\det \mathbf{A}}, \qquad (i = 1, 2, \dots, n)$$

alakban.

Gabriel Cramer (1704–1752) genfi születésű svájci matematikus, akinek az algebrai görbékről szóló „Introduction à l'analyse des lignes courbes algébraique" című, 1750-ben publikált munkájában szerepelt a ma Cramer-szabály néven ismert tétel. A szabályt korábban már mások is ismerték.

BIZONYÍTÁS. Az állítás első felét már bizonyítottuk a 6.14. tételben. Ebből felhasználjuk, hogy mivel az egyenletrendszer megoldható, $\det \mathbf{A} \neq 0$. Kihasználva, hogy $\mathbf{Ax} = \mathbf{b}$, továbbá hogy $\mathbf{Ae}_i = \mathbf{a}_{*i}$, kapjuk, hogy

$$\begin{aligned}
\mathbf{AI}_{i,\mathbf{x}} &= \mathbf{A}[\mathbf{e}_{*1}\ \dots\ \mathbf{e}_{*,i-1}\ \mathbf{x}\ \mathbf{e}_{*,i+1}\ \dots\ \mathbf{e}_{*n}] \\
&= [\mathbf{Ae}_{*1}\ \dots\ \mathbf{Ae}_{*,i-1}\ \mathbf{Ax}\ \mathbf{Ae}_{*,i+1}\ \dots\ \mathbf{Ae}_{*n}] \\
&= [\mathbf{a}_{*1}\ \dots\ \mathbf{a}_{*,i-1}\ \mathbf{b}\ \mathbf{a}_{*,i+1}\ \dots\ \mathbf{a}_{*n}] \\
&= \mathbf{A}_{i,\mathbf{b}}
\end{aligned}$$

Mivel az $\mathbf{I}_{i,\mathbf{x}}$ mátrix i-edik sorának és oszlopának elhagyása után egy identikus mátrix marad, ezért az i-edik sora szerint kifejtve

$$\det \mathbf{I}_{i,\mathbf{x}} = \begin{vmatrix}
1 & 0 & \dots & x_1 & \dots & 0 \\
0 & 1 & \dots & x_2 & \dots & 0 \\
\vdots & \vdots & \ddots & \vdots & & \vdots \\
0 & 0 & \dots & x_i & \dots & 0 \\
\vdots & \vdots & & \vdots & \ddots & \vdots \\
0 & 0 & \dots & x_n & \dots & 1
\end{vmatrix} = (-1)^{i+i} x_i = x_i.$$

Így a determinánsok szorzási szabályát is használva $\det(\mathbf{AI}_{i,\mathbf{x}}) = \det \mathbf{A}_{i,\mathbf{b}}$, amiből $x_i \det \mathbf{A} = \det \mathbf{A}_{i,\mathbf{b}}$, azaz $x_i = \det \mathbf{A}_{i,\mathbf{b}} / \det \mathbf{A}$. □

6.26. PÉLDA (CRAMER-SZABÁLY). *Oldjuk meg az*

$$\begin{aligned}
2x + 5y &= 4 \\
5x + 3y &= 6
\end{aligned}$$

egyenletrendszert a Cramer-szabállyal!

MEGOLDÁS. A kiszámolandó determinánsok a $\mathbf{b} = \left[\begin{smallmatrix} 4 \\ 6 \end{smallmatrix}\right]$ jelöléssel:

$$\mathbf{A} = \begin{vmatrix} 2 & 5 \\ 5 & 3 \end{vmatrix} = -19, \quad \mathbf{A}_{1,\mathbf{b}} = \begin{vmatrix} 4 & 5 \\ 6 & 3 \end{vmatrix} = -18, \quad \mathbf{A}_{2,\mathbf{b}} = \begin{vmatrix} 2 & 4 \\ 5 & 6 \end{vmatrix} = -8.$$

Innen $x = \frac{-18}{-19} = \frac{18}{19}$, $y = \frac{-8}{-19} = \frac{8}{19}$. □

Ha egyenletrendszert meg tudunk oldani, akkor szimultán egyenletrendszert is, és így pl. az $\mathbf{AX} = \mathbf{I}$ megoldásával a mátrix inverzét is ki tudjuk számítani. Az $\mathbf{x}_{ij}$ elem kiszámításához az $\mathbf{Ax}_{*j} = \mathbf{e}_j$ egyenletrendszert kell megoldani. A megoldás i-edik koordinátája az $\mathbf{x}_{ij}$ elem. A Cramer-szabály szerint

$$\mathbf{x}_{ij} = \frac{\det \mathbf{A}_{i,\mathbf{e}_j}}{\det \mathbf{A}}$$

Mivel az $\mathbf{A}_{i,\mathbf{e}_j}$ mátrix i-edik oszlopában csak egy elem nem 0, a kifejtési tétel szerint

$$\det \mathbf{A}_{i,\mathbf{e}_j} = \begin{vmatrix} a_{11} & a_{12} & \dots & 0 & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & 0 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \dots & 1 & \dots & a_{jn} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & 0 & \dots & a_{nn} \end{vmatrix} = A_{ji},$$

vagyis e determináns megegyezik az $\mathbf{A}$ egy előjeles aldeterminánsával, tehát

$$\mathbf{x}_{ij} = \frac{\det \mathbf{A}_{i,\mathbf{e}_j}}{\det \mathbf{A}} = \frac{\det \mathbf{A}_{ji}}{\det \mathbf{A}}.$$

Mint látjuk, az $\mathbf{X} = \mathbf{A}^{-1}$ előállításához az $\mathbf{A}$ előjeles aldeterminánsai mátrixának transzponáltjára van szükség. E mátrixot az $\mathbf{A}$ *klasszikus adjungáltjának* nevezzük és $\operatorname{adj}(\mathbf{A})$-val jelöljük. A klasszikus jelzőre azért van szükség, mert az adjungált szót komplex elemű mátrix konjugált transzponáltjára is használjuk, és ez félreértésekhez vezethet. Képletben tehát

$$\operatorname{adj} \mathbf{A} = [A_{ij}]^T = [A_{ji}]. \tag{6.1}$$

Így a következő tételt kapjuk:

6.27. TÉTEL (MÁTRIX INVERZÉNEK ELEMEI). *Tegyük fel, hogy* $\mathbf{A}$ *egy invertálható mátrix. Ekkor inverzének ij indexű eleme az* a_{ji} *elemhez tartozó előjeles aldetermináns és az* $\mathbf{A}$ *mátrix determinánsának hányadosa, azaz*

$$[\mathbf{A}^{-1}]_{ij} = \frac{A_{ji}}{\det \mathbf{A}}.$$

Így az inverz mátrix az

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}}[A_{ij}]^T = \frac{1}{\det \mathbf{A}} \operatorname{adj} \mathbf{A}. \tag{6.2}$$

alakba írható.

► Könnyen ellenőrizhető, hogy az $\mathbf{A} = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ mátrix klasszikus adjungáltja

$$\begin{bmatrix} d & -c \\ -b & a \end{bmatrix}^T = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix},$$

így inverze

$$\mathbf{A}^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{\det \mathbf{A}} \operatorname{adj} \mathbf{A} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

► A mátrix inverzének e kifejezése azt mutatja, hogy az inverz mátrix minden eleme folytonos függvénye a mátrix minden elemének minden olyan helyen, ahol a determináns nem 0, azaz minden olyan helyen, ahol az inverz egyáltalán létezik.

► Az előző megjegyzésből az is következik, hogy egy n-ismeretlenes n egyenletből álló egyenletrendszer megoldásvektorának minden koordinátája folytonos függvénye az egyenletrendszer együtthatóinak és a jobb oldalán álló vektor koordinátáinak, hisz a megoldás az inverzzel való szorzással megkapható.

► Egészelemű mátrix inverze pontosan akkor egészelemű, ha determinánsa 1 vagy -1. Ez abból adódik, hogy $\det(\mathbf{A})\det(\mathbf{A}^{-1}) = \det \mathbf{I} = 1$, tehát ha $|\det \mathbf{A}| \neq 1$, akkor $\det(\mathbf{A}^{-1})$ nem egész szám, tehát $\mathbf{A}^{-1}$ nem lehet egészelemű, ha pedig $|\det \mathbf{A}| = 1$, akkor a (6.2) képlet szerint $\mathbf{A}^{-1}$ minden eleme egész szám.

6.28. PÉLDA (MÁTRIX INVERZE). *Számítsuk ki a szemléltetés céljából csupa különböző elemet tartalmazó*

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 5 & 6 \\ 4 & 7 & 9 \end{bmatrix}$$

mátrix inverzét!

MEGOLDÁS. Az $\operatorname{adj} \mathbf{A}$ determinánst olyan alakba írjuk föl, ahonnan látszik minden elem kiszámításának módja. Szürke színnel szedjük az elhagyandó elemeket:

$$\operatorname{adj}\mathbf{A} = \begin{bmatrix} +\begin{vmatrix} 0 & 1 & 2\\ 3 & 5 & 6\\ 4 & 7 & 9 \end{vmatrix} & -\begin{vmatrix} 0 & 1 & 2\\ 3 & 5 & 6\\ 4 & 7 & 9 \end{vmatrix} & +\begin{vmatrix} 0 & 1 & 2\\ 3 & 5 & 6\\ 4 & 7 & 9 \end{vmatrix}\\ -\begin{vmatrix} 0 & 1 & 2\\ 3 & 5 & 6\\ 4 & 7 & 9 \end{vmatrix} & +\begin{vmatrix} 0 & 1 & 2\\ 3 & 5 & 6\\ 4 & 7 & 9 \end{vmatrix} & -\begin{vmatrix} 0 & 1 & 2\\ 3 & 5 & 6\\ 4 & 7 & 9 \end{vmatrix}\\ +\begin{vmatrix} 0 & 1 & 2\\ 3 & 5 & 6\\ 4 & 7 & 9 \end{vmatrix} & -\begin{vmatrix} 0 & 1 & 2\\ 3 & 5 & 6\\ 4 & 7 & 9 \end{vmatrix} & +\begin{vmatrix} 0 & 1 & 2\\ 3 & 5 & 6\\ 4 & 7 & 9 \end{vmatrix} \end{bmatrix}^T$$

$$= \begin{bmatrix} +\begin{vmatrix} 5 & 6\\ 7 & 9 \end{vmatrix} & -\begin{vmatrix} 3 & 6\\ 4 & 9 \end{vmatrix} & +\begin{vmatrix} 3 & 5\\ 4 & 7 \end{vmatrix}\\ -\begin{vmatrix} 1 & 2\\ 7 & 9 \end{vmatrix} & +\begin{vmatrix} 0 & 2\\ 4 & 9 \end{vmatrix} & -\begin{vmatrix} 0 & 1\\ 4 & 7 \end{vmatrix}\\ +\begin{vmatrix} 1 & 2\\ 5 & 6 \end{vmatrix} & -\begin{vmatrix} 0 & 2\\ 3 & 6 \end{vmatrix} & +\begin{vmatrix} 0 & 1\\ 3 & 5 \end{vmatrix} \end{bmatrix}^T$$

$$= \begin{bmatrix} 3 & -3 & 1\\ 5 & -8 & 4\\ -4 & 6 & -3 \end{bmatrix}^T$$

Mivel $\det\mathbf{A} = -1$, ezért

$$\mathbf{A}^{-1} = \frac{1}{\det\mathbf{A}}\operatorname{adj}\mathbf{A} = -\begin{bmatrix} 3 & -3 & 1\\ 5 & -8 & 4\\ -4 & 6 & -3 \end{bmatrix}^T = \begin{bmatrix} -3 & -5 & 4\\ 3 & 8 & -6\\ -1 & -4 & 3 \end{bmatrix} \qquad \square$$

Már ezekből az egyszerű példákból is látszik, hogy mátrix invertálása e módszerrel igen műveletigényes. Valóban, gyakorlati számításokhoz nem használjuk, elméleti okfejtésekben vesszük nagy hasznát.

*Blokkmátrixok determinánsa** Az $\mathbf{M} = \left[\begin{smallmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{smallmatrix}\right]$ mátrix általában még négyzetes részmátrixok esetén sem számítható az $\mathbf{AD} - \mathbf{BC}$ képlettel (ld. a **??** feladatban)! Először egy speciális, de fontos esettel kezdjük.

6.29. TÉTEL (DETERMINÁNSOK SZORZATA BLOKKMÁTRIXBAN). *Legyenek* $\mathbf{A}$ *és* $\mathbf{D}$ *négyzetes mátrixok. Ekkor*

$$\begin{vmatrix} \mathbf{A} & \mathbf{B}\\ \mathbf{O} & \mathbf{D} \end{vmatrix} = \begin{vmatrix} \mathbf{A} & \mathbf{O}\\ \mathbf{C} & \mathbf{D} \end{vmatrix} = |\mathbf{A}||\mathbf{D}|.$$

BIZONYÍTÁS. Megmutatjuk, hogy minden olyan kígyó, melynek nincs eleme a $\mathbf{O}$-mátrixból, egy $\mathbf{A}$-beli és egy $\mathbf{D}$-beli kígyó szorzata. Ehhez elég megmutatni, hogy ha egy kígyónak van eleme a $\mathbf{B}$, illetve a $\mathbf{C}$ mátrixból, akkor az $\mathbf{O}$-ból is. Valóban, ha pl. $\mathbf{B}$ egy eleme benne van egy kígyóban, akkor oszlopában nincs elem $\mathbf{D}$-ben, így $\mathbf{D}$-ben marad egy sor is üresen, amelyet csak egy $\mathbf{O}$-beli elem foghat le. Ellenőrizni kell még, hogy az $\mathbf{A}$- és $\mathbf{D}$-beli kígyók előjeleinek szorzata megegyezik-e az egyesítésükkel kapott kígyó előjelével. Ez nyilván igaz, hisz egy $\mathbf{A}$-t és egy $\mathbf{D}$-t metsző sor nem lehet inverzióban, így az egyesített kígyó inverzióinak száma megegyezik a két kígyó inverzióinak összegével, az előjelet pedig a -1-nek az inverziók számára emelt hatványa adja. □

6.30. TÉTEL (2×2-ES BLOKKMÁTRIX DETERMINÁNSA). *Legyen*

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix},$$

ahol $\mathbf{A}$ *és* $\mathbf{D}$ *négyzetes mátrixok.*

1. *Ha* $|\mathbf{A}| \neq 0$, *akkor* $|\mathbf{M}| = |\mathbf{A}||\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}|$.
2. *Ha* $|\mathbf{D}| \neq 0$, *akkor* $|\mathbf{M}| = |\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}||\mathbf{D}|$.

BIZONYÍTÁS. Ha $\mathbf{A}$ invertálható, akkor $\mathbf{M}$ alábbi alsó és felső blokkháromszögmátrix szorzatára való bontása segít:

$$\begin{aligned}\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} &= \begin{bmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{C} & \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{O} & \mathbf{I} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{C}\mathbf{A}^{-1} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{O} & \mathbf{I} \end{bmatrix}\end{aligned}$$

Az utóbbi három mátrix közül a szélsők determinánsa 1, a középsőé pedig a bizonyítandó kifejezés. Az

$$\mathbf{M} = \begin{bmatrix} \mathbf{I} & \mathbf{B}\mathbf{D}^{-1} \\ \mathbf{O} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C} & \mathbf{O} \\ \mathbf{O} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{D}^{-1}\mathbf{C} & \mathbf{I} \end{bmatrix}$$

felbontás bizonyítja a második összefüggést. □

Vandermonde-determináns Bemutatunk egy fontos determinánst. Számtalan alkalmazása van, melyek egyike a polinominterpoláció.

6.31. PÉLDA (INTERPOLÁCIÓ MÁSODFOKÚ POLINOMOKRA). *Legyen x, y és z három különböző valós, a, b és c három tetszőleges valós. Mutassuk meg, hogy egyetlen olyan legföljebb másodfokú f polinom létezik, melyre $f(x) = a$, $f(y) = b$ és $f(z) = c$.*

MEGOLDÁS. Legyen $f : x \mapsto p + qx + rx^2$, ahol p, q és r a polinom ismeretlen együtthatói. Az $f(x) = a$, $f(y) = b$ és $f(z) = c$ egyenlőségek a következő egyenletrendszerre vezetnek:

$$\begin{bmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{bmatrix} \begin{bmatrix} p \\ q \\ r \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Ez az egyenletrendszer a 6.14. tétel szerint pontosan akkor oldható meg egyértelműen, ha az együtthatómátrix determinánsa nem 0. Oszlopműveletekkel kezdjük az átalakítást:

$$\begin{aligned} \begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix} &\overset{O_3-xO_2}{=} \begin{vmatrix} 1 & x & 0 \\ 1 & y & y^2-xy \\ 1 & z & z^2-xz \end{vmatrix} \overset{O_2-xO_1}{=} \begin{vmatrix} 1 & 0 & 0 \\ 1 & y-x & y^2-xy \\ 1 & z-x & z^2-xz \end{vmatrix} \\ = \begin{vmatrix} y-x & y^2-xy \\ z-x & z^2-xz \end{vmatrix} &= (y-x) \begin{vmatrix} 1 & y \\ z-x & z^2-xz \end{vmatrix} = (y-x)(z-x) \begin{vmatrix} 1 & y \\ 1 & z \end{vmatrix} \\ =&(y-x)(z-x)(z-y) \end{aligned}$$

Mivel x, y és z három különböző valós, ezért a determináns értéke nem 0, tehát az egyenletrendszer egyértelműen megoldható, vagyis egyetlen olyan polinom létezik, mely a feltételeket teljesíti. □

E probléma, és a benne szereplő determináns általánosítása a következő definícióhoz vezet:

6.32. DEFINÍCIÓ (VANDERMONDE-DETERMINÁNS). *Az $x_1, x_2, \dots x_n$ számokhoz tartozó* Vandermonde-determinánson *a*

$$V_n(x_1, x_2, \dots, x_n) = \begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix} \tag{6.3}$$

determinánst vagy ennek

$$\begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{vmatrix}$$

transzponáltját értjük. A hozzá tartozó mátrixot Vandermonde-mátrixnak *nevezzük.*

Mivel egy determináns értéke megegyezik transzponáltjának értékével, ezért a definícióbeli két determináns értéke is azonos, így mindegy melyik alakot használjuk.

6.33. TÉTEL (VANDERMONDE-DETERMINÁNS ÉRTÉKE). *Az $x_1, x_2, \ldots x_n$ $(n > 1)$ számokhoz tartozó Vandermonde-determináns értéke megegyezik az olyan $(x_j - x_i)$ alakú különbségek szorzatával, ahol $i < j$, azaz*

$$V_n(x_1, x_2, \ldots, x_n) = \prod_{i<j}(x_j - x_i).$$

BIZONYÍTÁS. A determináns utolsó oszlopával kezdve minden oszlopból vonjuk ki az előző oszlop x_1-szeresét.

$$\begin{aligned}
V_n(x_1, x_2, \ldots, x_n) &= \begin{vmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} \end{vmatrix} \\
&= \begin{vmatrix} 1 & 0 & 0 & \ldots & 0 \\ 1 & x_2 - x_1 & x_2^2 - x_1x_2 & \ldots & x_2^{n-1} - x_1x_2^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n - x_1 & x_n^2 - x_1x_2 & \ldots & x_n^{n-1} - x_1x_n^{n-2} \end{vmatrix}
\end{aligned}$$

ami az első sora szerinti kifejtés, majd minden sorból kiemelve az első oszlopbeli elemet, a következő alakra vezet:

$$\begin{aligned}
&= \begin{vmatrix} x_2 - x_1 & x_2^2 - x_1x_2 & \ldots & x_2^{n-1} - x_1x_2^{n-2} \\ \vdots & \vdots & & \vdots \\ x_n - x_1 & x_n^2 - x_1x_2 & \ldots & x_n^{n-1} - x_1x_n^{n-2} \end{vmatrix} \\
&= (x_2 - x_1)(x_3 - x_1)\ldots(x_n - x_1) \begin{vmatrix} 1 & x_2 & x_2^2 & \ldots & x_2^{n-2} \\ 1 & x_3 & x_3^2 & \ldots & x_3^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-2} \end{vmatrix} \\
&= (x_2 - x_1)(x_3 - x_1)\ldots(x_n - x_1)V_{n-1}(x_2, \ldots, x_n) \\
&= \prod_{j>1}(x_j - x_1) \cdot V_{n-1}(x_2, \ldots, x_n).
\end{aligned}$$

Eredményül egy rekurzív képletet kaptunk, melyet önmagába helyettesítve, és a $V_2(x_{n-1}, x_n) = x_n - x_{n-1}$ képletet is fölhasználva a tételbeli összefüggésre jutunk. $\square$

Feladatok

6.25. Melyek igazak az alábbi állítások közül? (Az **A** itt mindig négyzetes mátrixot jelöl.)

1. A determináns folytonos függvénye minden elemének.
2. A determináns differenciálható függvénye minden elemének.
3. Ha egy determináns minden eleme racionális szám, akkor értéke is racionális.
4. Ha egy determináns minden sorában és minden oszlopában pontosan egy elem nem 0, akkor a determináns értéke nem 0.
5. Ha egy mátrix két kígyó összege, akkor determinánsa is két kígyó determinánsának összege.
6. Ha $i+j$ páratlan szám, akkor az előjeles A_{ij} aldetermináns negatív.
7. Ha egy determináns minden eleme pozitív, akkor értéke nem lehet negatív.
8. Mátrix inverze folytonos függvénye minden elemének.

Felbontás kígyók determinánsainak összegére

6.26. Válasszuk ki az alábbi determinánsokból az összes nemnulla determinánsú kígyót, és ezek segítségével számítsuk ki a determináns értékét!

a) $\begin{bmatrix} 0 & 1 & 0 \\ 2 & 3 & 4 \\ 5 & 0 & 6 \end{bmatrix}$

b) $\begin{vmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 0 \\ 0 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{vmatrix}$

c) $\begin{vmatrix} 1 & 0 & 2 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 & 0 \\ 2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 2 \end{vmatrix}$

6.27. A determináns értékének kiszámítása nélkül mutassuk meg, hogy osztható 30-cal:

$$\begin{vmatrix} 24 & 40 & 68 \\ 27 & 15 & 31 \\ 51 & 55 & 53 \end{vmatrix}$$

6.28. Az alábbi – lottótippekből álló – determináns elemeinek csak a paritását vizsgálva minden számolás nélkül igazoljuk, hogy

$$\begin{vmatrix} 12 & 25 & 28 & 44 & 56 \\ 21 & 34 & 54 & 68 & 80 \\ 10 & 40 & 52 & 69 & 72 \\ 24 & 36 & 53 & 56 & 84 \\ 18 & 24 & 28 & 58 & 87 \end{vmatrix} \neq 0.$$

6.29. A 4-edrendű determinánsok $4! = 24$ kígyó determinánsának összegére bonthatók. Soroljuk fel közülük azt a 12 darabot, melyet elemei szorzata után -1-gyel kell szorozni! (A Sarrus-szabály 4-edrendű determinánsra csak 8 kígyóból állna, ezért nem használható!)

Kifejtési tétel

6.30. Tudjuk, hogy 504, 747 és 855 egyaránt oszthatók 9-cel. Ezt fölhasználva, a determináns értékének kiszámítása nélkül mutassuk meg, hogy az alábbi determináns osztható 9-cel:

$$\begin{vmatrix} 5 & 0 & 4 \\ 7 & 4 & 7 \\ 8 & 5 & 5 \end{vmatrix}.$$

6.31. Konstruáljunk olyan nemnulla értékű determinánst, melynek van olyan eleme, amelyet tetszőlegesen változtatva a determináns értéke nem változik.

Blokkdeterminánsok

6.32. Számítsuk ki az alábbi determinánsok értékét kihasználva blokkstruktúrájukat!

a) $\begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & 3 & 3 & 3 \end{vmatrix}$ *b)* $\begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 3 & 3 & 0 & 0 \\ 0 & 2 & 2 & 2 & 0 \\ 0 & 1 & 1 & 1 & 1 \end{vmatrix}$

Speciális mátrixok determinánsa

6.33. Számítsuk ki az alábbi determinánsok értékét!

a) $\begin{vmatrix} 1 & 1 & 1 & 1 \\ 2 & -1 & -2 & 1 \\ 4 & 1 & 4 & 1 \\ 8 & -1 & -8 & 1 \end{vmatrix}$

b) $\begin{vmatrix} 1 & -3 & 9 & -27 & 81 \\ 1 & 2 & 4 & 8 & 16 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & -2 & 4 & -8 & 16 \\ 1 & -1 & 1 & -1 & 1 \end{vmatrix}$

c) $\begin{vmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d+e & d^2+e^2 & d^3+e^3 \end{vmatrix}$

6.34. Bizonyítsuk be, hogy

$$D = \begin{vmatrix} p^2 & p & 1 & qrs \\ q^2 & q & 1 & prs \\ r^2 & r & 1 & pqs \\ s^2 & s & 1 & pqr \end{vmatrix} = (p-q)(p-r)(p-s)(q-r)(q-s)(r-s).$$

6.35. Igazoljuk, hogy az $a_1 = 1$, $a_2 = 2$, $a_n = a_{n-1} + a_{n-2}$ képletekkel definiált Fibonacci-sorozat n-edik eleme egyenlő az alábbi $n \times n$-es tridiagonális determinánssal:

$$a_n = \begin{vmatrix} 1 & -1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 1 & -1 & 0 & \dots & 0 & 0 \\ 0 & 1 & 1 & -1 & \dots & 0 & 0 \\ 0 & 0 & 1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & -1 \\ 0 & 0 & 0 & 0 & \dots & 1 & 1 \end{vmatrix}$$

6.36. Legyen

$$P_n = \begin{vmatrix} a_n & -1 & 0 & 0 & \dots & 0 & 0 \\ 1 & a_{n-1} & -1 & 0 & \dots & 0 & 0 \\ 0 & 1 & a_{n-2} & -1 & \dots & 0 & 0 \\ 0 & 0 & 1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & a_2 & -1 \\ 0 & 0 & 0 & 0 & \dots & 1 & a_1 \end{vmatrix}$$

Mutassuk meg, hogy

$$\frac{P_k}{P_{k-1}} = a_k + \cfrac{1}{a_{k-1} + \cfrac{1}{a_{k-2} + \cfrac{1}{\ddots \cfrac{}{a_2 + \cfrac{1}{a_1}}}}}.$$

Vegyes feladatok

6.37. Elérhető-e egyetlen elem megváltoztatásával, hogy egy tetszőleges $n \times n$-es nem szinguláris mátrix determinánsa 0-vá váljon?

6.38. Ferde kifejtés Vegyük egy determináns egy sorának elemeit, és szorozzuk meg mindegyiket egy másik sor azonos oszlopbeli eleméhez tartozó előjeles aldeterminánsával, majd képezzük ezek összegét. Ez mindig 0. Hasonló állítás igaz a determináns minden oszlopára is. Tehát az i-edik és u-adik sorra ($i \neq u$) és a j-edik és v-edik oszlopra ($j \neq v$):

$$\sum_{k=1}^{n} a_{ik} A_{uk} = 0, \qquad \sum_{k=1}^{n} a_{kj} A_{kv} = 0.$$

6.39. Foglaljuk egyetlen állításba a kifejtési és a ferde kifejtési tételeket!

6.40. Mátrix inverze a kifejtési tételekkel A kifejtési és a ferde kifejtési (ld. az előző és a 6.38. feladatokat) segítségével adjunk új bizonyítást a mátrix inverzére vonatkozó (6.2) formulátra!

6.41. Legyen $x_0, x_1, x_2, \dots, x_n$ $n+1$ darab különböző valós, $y_0, y_1, \dots, y_n$ ugyanannyi tetszőleges valós. Mutassuk meg, hogy egyetlen olyan legföljebb n-edfokú p polinom van, melyre $p(x_i) = y_i$ minden $i = 0, \dots, n$ esetén.

Cramer-szabály és mátrix inverze

6.42.♥ Oldjuk meg Cramer-szabállyal az alábbi egyenletrendszereket!

a) $\begin{aligned} x + y &= 1 \\ x - 2y &= 4 \end{aligned}$

b) $\begin{aligned} 2x - y - z &= 4 \\ 3x + 4y - 2z &= 11 \\ 3x - 2y + 4z &= 11 \end{aligned}$

c) $\begin{aligned} x + 2y + 4z &= 31 \\ 5x + y + 2z &= 29 \\ 3x - y + z &= 10 \end{aligned}$

d) $\begin{aligned} x + y &= 1 \\ x + 2y + z &= 2 \\ y + 2z + w &= 3 \\ z + 2w &= 4 \end{aligned}$

6.43.♥ Határozzuk meg a megadott mátrixok inverzének megadott indexű elemét!

a) $\begin{bmatrix} 1 & 4 & 7 \\ 2 & 4 & 6 \\ 3 & 2 & 3 \end{bmatrix}$, $a_{23} = ?$

b) $\begin{bmatrix} 1 & 2 & 5 & 7 \\ 0 & 1 & 3 & 6 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ $a_{24} = ?$

6.44.♥ Határozzuk meg a megadott mátrixok inverzét a klasszikus adjungált kiszámolásával:

a) $\begin{bmatrix} 3 & 1 & 4 \\ -7 & 2 & 7 \\ 2 & 1 & 4 \end{bmatrix}$

b) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & 2 \\ 3 & 2 & 1 \end{bmatrix}$

c) $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

d) $\begin{bmatrix} 0 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{bmatrix}$

e) $\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$ $(abc \neq 0)$ *f)* $\begin{bmatrix} 1+\mathrm{i} & \mathrm{i} \\ \mathrm{i} & \mathrm{i} \end{bmatrix}$

g) $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 3 & 0 \\ 4 & 0 & 0 & 0 \end{bmatrix}$ *h)* $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

6.45. Igazoljuk, hogy tetszőleges négyzetes mátrixra $\mathbf{A}\,\mathrm{adj}(\mathbf{A}) = \det(\mathbf{A})\mathbf{I}$.

*Véges testek fölötti mátrixok determinánsa**

6.46. A determináns kiszámításának megismert technikái véges testek fölött is működnek. Számítsuk ki az alábbi – a megadott test fölött értelmezett – mátrixok determinánsát!

a) $\begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$, $\mathbb{F}_2$, $\mathbb{F}_3$, $\mathbb{F}_5$

b) $\begin{bmatrix} 3 & 2 & 3 \\ 5 & 7 & 6 \\ 2 & 7 & 2 \end{bmatrix}$, $\mathbb{F}_{11}$

6.47.■ VÉLETLEN BITMÁTRIX DETERMINÁNSA Számítsuk ki $\mathbb{F}_2$ fölötti véletlen mátrixok determinánsát! Egy $\mathbb{F}_2^{5\times 5}$-beli mátrix determinánsa mekkora valószínűséggel 0? Kísérletezzünk számítógéppel, majd válaszoljuk meg a kérdést pontosan.

Projekt: a vektori szorzás általánosítása

6.48. Bizonyított tény, hogy nem lehet olyan bináris vektorműveletet definiálni az n-dimenziós tér vektorain, mely eredményül ugyanannak a térnek egy vektorát adja és rendelkezik a vektori szorzás műveleti tulajdonságaival. E feladatsorban egy másik irányú általánosítást dolgozunk fel, mely nem a bináris műveleti tulajdonságokat, hanem az eredménynek a vektorokra való merőlegességét tarja meg.

1. Fogalmazzuk meg, hogy mit kapunk eredményül, ha a vektori szorzásra vonatkozó formális

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

összefüggést 2×2-es vagy 4×4-es formális determinánsokra írjuk föl, vagyis mit ad eredményül az

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} \\ a_1 & a_2 \end{vmatrix} \quad \text{és az} \quad \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 \\ a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{vmatrix}$$

kifejezés?
2. Igazoljuk, hogy az n-dimenziós

$$\begin{aligned} \mathbf{a}_1 &= (a_{11}, a_{12}, \dots, a_{1n}), \\ \mathbf{a}_2 &= (a_{21}, a_{22}, \dots, a_{2n}), \\ &\vdots \\ \mathbf{a}_{n-1} &= (a_{n-1,1}, a_{n-1,2}, \dots, a_{n-1,n}) \end{aligned}$$

vektorok által kifeszített $n-1$-dimenziós paralelepipedon térfogata megegyezik az

$$\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_n \\ a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \dots & a_{n-1,n} \end{vmatrix}$$

determináns abszolút értékével.
3. Ha a fentiek alapján általánosított képlettel $n-1$ darab n-dimenziós vektorhoz egy n-ediket rendelünk, akkor mit mondhatunk az így kapott n vektor körüljárásáról?
4. Keressünk olyan formulát, mely a lineárisan független n-dimenziós $\mathbf{a}_1, \dots, \mathbf{a}_{n-1}$ vektorokhoz olyan $\mathbf{a}_n$ vektort rendel, mely merőleges mindegyikükre, és amely azokkal az index szerinti sorrendben jobbrendszert alkot, és nincs a formulában -1-hatvány!

6.49.♥ Határozzuk meg azt a vektort, mely merőleges a $(1,1,1,1)$, $(1,2,2,2)$, $(1,2,3,3)$ vektorokra, hossza megegyezik a három vektor által kifeszített parallelepipedon térfogatával, és e három vektor mellé negyediknek véve velük jobbrendszert alkot.

Megoldások

6.1. 1. Hamis. 2. Igaz. 3. Hamis. 4. Hamis. Az $|\mathbf{A}| \neq 0$ azzal ekvivalens, hogy az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszer nem oldható meg egyértelműen, vagyis vagy nem oldható meg, vagy több megoldása is van. 5. Hamis.

6.2.
a) -2.
b) 0, mert van két azonos sora.
c) 0, mert a második sor az első konstansszorosa.
d) 1, mert háromszögmátrix determinánsa a főátlóbeli elemek szorzata.
e) 6, mert háromszögmátrix determinánsa a főátlóbeli elemek szorzata.
f) 0, mert van két azonos oszlopa.

6.3.
a) a második sor az első -1-szerese.
b) a harmadik sor egyenlő az első kettő összegével.
c) a harmadik sor egyenlő az első kettő összegével.
d) a második sor az első és a harmadik számtani közepe (másként: a harmadik sorból kivonva a másodikat, majd a másodikból az elsőt, mindkétszer az $(1, 1, 1)$ vektort kapjuk, azaz így van két azonos sor).
e) a második sor az első és a harmadik számtani közepe.
f) a három sorvektor összege a zérusvektor.
g) $\sin(\xi + \delta) = \sin\xi\cos\delta + \cos\xi\sin\delta$, így a harmadik oszlop az első és a második oszlop lineáris kombinációja, vagyis az oszlopvektorok lineárisan összefüggőek, tehát a determináns értéke 0.
h) Az első és második oszlop összege a harmadik oszlop (ill. az első és a második sor különbsége a harmadik sor), tehát az oszlopvektorok (ill. sorvektorok) lineárisan összefüggőek.

6.5. *a)* 125, *b)* 40, *c)* 1600, *d)* $1/5$, *e)* 25, *f)* $1/625$, *g)* $5/4$, *h)* 20, *i)* 1.

6.6.
a)
$$\begin{vmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 1 & 0 & 0 \end{vmatrix} = -\begin{vmatrix} 0 & 2 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{vmatrix} = 6$$
b) Az 1. és 2., azután az 1. és 3., végül az 1. és 5. sorokat felcserélve:
$$\begin{vmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 4 & 0 \\ 5 & 0 & 0 & 0 & 0 \end{vmatrix} = -\begin{vmatrix} 0 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 4 & 0 \\ 5 & 0 & 0 & 0 & 0 \end{vmatrix} =$$
$$\begin{vmatrix} 0 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 5 & 0 & 0 & 0 & 0 \end{vmatrix} = -\begin{vmatrix} 5 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{vmatrix} = -120.$$
c) -1, -1, 1.
d) 24.
e) 24.
f) Az első sort cseréljük fel az utolsóval, a másodikat az utolsó előttivel, ..., így $\lfloor \frac{n}{2} \rfloor$ sorcserét hajtottunk végre, tehát a determináns értéke $(-1)^{\lfloor n/2 \rfloor}$ (itt $\lfloor . \rfloor$ az egészrész-függvényt jelöli). Ennek értéke 1, ha $n = 4k$ vagy $n = 4k + 1$, és -1, ha $n = 4k + 2$ vagy $n = 4k + 3$ valamilyen k természetes számra. Más alakban kapjuk meg az eredményt, ha csak szomszédos sorokat cserélünk: először az első sort visszük (szomszédos sorok cseréjével) az utolsóba, majd az eredeti determináns második sorát az utolsó előttibe, ..., azaz az alábbi sorpárok cseréjét hajtjuk végre:
$(1, 2)$, $(2, 3)$, $(3, 4)$,..., $(n-1, n)$,
$(1, 2)$, $(2, 3)$,..., $(n-2, n-1)$,
...
$(1, 2)$, $(2, 3)$,
$(1, 2)$.
Ez összesen $(n-1) + (n-2) + \cdots + 2 + 1 = \frac{n(n-1)}{2}$ sorcsere. Minden sorcserével (-1)-szeresére változik a determináns értéke, így a végeredmény $(-1)^{n(n-1)/2}$. Természetesen e hatvány értéke is akkor 1, ha $n = 4k$ vagy $4k + 1$, és akkor -1, ha $n = 4k + 2$ vagy $4k + 3$. (Ugyanilyen gondolatmenettel kimutatható, hogy ha egy determináns mellékátlója felett csupa 0 áll, akkor a determináns értéke a mellékátlóbeli elemek szorzatának $(-1)^{\lfloor n/2 \rfloor}$-szerese vagy más alakban $(-1)^{n(n-1)/2}$-szerese.)
g) ld. az előző pontot.

6.7.
a) Az első sort kivonjuk a másodikból és a harmadikból, majd a másodikat a harmadikból:
$$\begin{vmatrix} 1 & 2 & 3 \\ 1 & 3 & 5 \\ 1 & 3 & 6 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 1 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{vmatrix} = 1.$$

b) Az első sort kivonva a többi sorból, majd a második sor kétszeresét kivonva a harmadikból, kapjuk, hogy

$$\begin{vmatrix}1&1&1&1\\0&1&2&3\\0&2&4&6\\0&3&6&9\end{vmatrix}=\begin{vmatrix}1&1&1&1\\0&1&2&3\\0&0&0&0\\0&3&6&9\end{vmatrix}=0.$$

c) $\begin{vmatrix}3&8&6&3\\1&2&0&1\\1&1&-1&2\\2&5&1&5\end{vmatrix}=-\begin{vmatrix}1&2&0&1\\3&8&6&3\\1&1&-1&2\\2&5&1&5\end{vmatrix}=$
$-\begin{vmatrix}1&2&0&1\\0&2&6&0\\0&-1&-1&1\\0&1&1&3\end{vmatrix}=-2\begin{vmatrix}1&2&0&1\\0&1&3&0\\0&-1&-1&1\\0&1&1&3\end{vmatrix}=$
$-2\begin{vmatrix}1&2&0&1\\0&1&3&0\\0&0&2&1\\0&0&-2&3\end{vmatrix}=-2\begin{vmatrix}1&2&0&1\\0&1&3&0\\0&0&2&1\\0&0&0&4\end{vmatrix}=-16.$ Részletezzük a megoldás lépéseit:

1. lépés: Cseréljük ki az első és második sort, hogy az első sor első eleme 1 legyen, s így ne kelljen törtekkel számolni. A determináns értéke (-1)-szeresére változik.

2. lépés: Az első sor (-3)-, (-1)- ill. (-2)-szeresét adjuk a második, harmadik ill. negyedik sorhoz.

3. lépés: Hogy a második sor második eleme is 1 legyen, emeljünk ki 2-t a második sorból.

4. lépés: A második sort ill. (-1)-szeresét adjuk a harmadik ill. negyedik sorhoz.

5. lépés: Adjuk a harmadik sort a negyedikhez. A determináns értéke -16.

d) 144.

6.8.

a) $n=1$ esetén $1+x_1y_1$, $n=2$ esetén $x_1y_1+x_2y_2-x_1y_2-x_2y_1$ a determináns értéke. Ha $n\ge 3$, akkor a determináns értéke 0. Ezt úgy bizonyítjuk, hogy először a determinánst két determináns összegére bontjuk, majd mindkettőről belátjuk, hogy értékük 0. Az első determináns csupa 1-esből álló első sorát kivonjuk az összes többi sorból, az így kapott determináns értéke pedig valóban 0, hisz ha $x_2=0$, akkor a második sor csupa 0-ból áll, ha pedig $x_2\ne 0$, akkor a második sorának x_3/x_2-szerese egyenlő a harmadik sorral. A második determináns értéke is 0, hiszen ha $x_1=0$, akkor az első sor csupa 0-ból áll, ha pedig $x_1\ne 0$, akkor az első sor x_i/x_1-szeresét kivonva az i-edik sorból egy olyan determinánst kapunk, amelyben a második sortól kezdve minden sor 1-esekből áll, tehát a determinánsnak van két azonos sora.

$$\begin{vmatrix}1&1&\dots&1\\1+x_2y_1&1+x_2y_2&\dots&1+x_2y_n\\\vdots&\vdots&\ddots&\vdots\\1+x_ny_1&1+x_ny_2&\dots&1+x_ny_n\end{vmatrix}$$
$$+\begin{vmatrix}x_1y_1&x_1y_2&\dots&x_1y_n\\1+x_2y_1&1+x_2y_2&\dots&1+x_2y_n\\\vdots&\vdots&\ddots&\vdots\\1+x_ny_1&1+x_ny_2&\dots&1+x_ny_n\end{vmatrix}$$
$$=\begin{vmatrix}1&1&\dots&1\\x_2y_1&x_2y_2&\dots&x_2y_n\\\vdots&\vdots&\ddots&\vdots\\x_ny_1&x_ny_2&\dots&x_ny_n\end{vmatrix}$$
$$+\begin{vmatrix}x_1y_1&x_1y_2&\dots&x_1y_n\\1&1&\dots&1\\\vdots&\vdots&\ddots&\vdots\\1&1&\dots&1\end{vmatrix}$$
$$=0.$$

b) $(1-a^n)^{n-1}$. Vonjuk ki az első sor a^{n-1}-szeresét a második sorból, a^{n-2}-szeresét a harmadik sorból, …, a-szorosát az utolsó sorból: így a főátló alatt csak nullák lesznek, a determináns értéke.

c) $(a+(n-1)b)(a-b)^{n-1}$. Első megoldás: adjunk minden sort az elsőhöz, emeljük ki a közös $a+(n-1)b$ értéket, majd e sor b-szeresét vonjuk ki minden sorból. Másik megoldás: az utolsó sorral kezdve mindegyik sorból vonjuk ki a fölötte lévőt, majd jobbról kezdve mindegyik oszlopot adjuk a megelőzőhöz.

6.9. Az eredmény 48. Megoldás Sage-ben:

```
g = graphs.PetersenGraph()
G = matrix(g)
G.det()
```

6.12. Az $\mathbf{A}$ előáll $\mathbf{PLU}$ alakban, ahol $\mathbf{P}$ permutációs mátrix, $\mathbf{L}$ alsó, $\mathbf{U}$ felső háromszögmátrix. Az $\mathbf{L}$ és az $\mathbf{U}$ háromszögmátrixok, így determinánsuk megegyezik transzponáltjuk determinánsával, hisz a főátlóbeli elemek helyben maradnak a transzponálás során. A $\mathbf{P}$ permutációs mátrix determinánsa 1 vagy -1, transzponáltja pedig megegyezik inverzével, így $\det(\mathbf{I})=\det(\mathbf{PP}^T)=\det(\mathbf{P})\det(\mathbf{P}^T)=1$, azaz $\mathbf{P}$ és $\mathbf{P}^{-1}$ egyszerre 1 vagy -1, tehát megegyeznek. Végül $\det(\mathbf{A})=\det(\mathbf{PLU})=\det(\mathbf{P})\det(\mathbf{L})\det(\mathbf{U})$, és $\det(\mathbf{A}^T)=\det((\mathbf{PLU})^T)=\det(\mathbf{U}^T\mathbf{L}^T\mathbf{P}^T)=\det(\mathbf{U})\det(\mathbf{L})\det(\mathbf{P})$ összevetése bizonyítja az állítást.

6.13. $\det(\mathbf{E}_{S_i+cS_j})=1$, $\det(\mathbf{E}_{S_i\leftrightarrow S_j})=-1$, $\det(\mathbf{E}_{cS_i})=c$.

6.14. Felhasználva, hogy $\binom{n}{k} - \binom{n-1}{k} = \binom{n-1}{k-1}$, elvégezve az ajánlott sor-, majd oszlopműveleteket, majd azt megismételve az egyre kisebb bal alsó részdeterminánssal kapjuk, hogy

$$D = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & \binom{1}{0} & \binom{2}{0} & \dots & \binom{n-1}{0} \\ 1 & \binom{2}{1} & \binom{3}{1} & \dots & \binom{n}{1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \binom{n-1}{n-2} & \binom{n}{n-2} & \dots & \binom{2n-3}{n-2} \end{vmatrix}$$
$$= \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & \binom{0}{0} & \binom{1}{0} & \dots & \binom{n-2}{0} \\ 0 & \binom{1}{1} & \binom{2}{1} & \dots & \binom{n-1}{1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \binom{n-2}{n-2} & \binom{n-1}{n-2} & \dots & \binom{2n-4}{n-2} \end{vmatrix}$$
$$= \dots = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \binom{0}{0} \end{vmatrix} = 1.$$

6.15. Az első sort kivonva a másodikból és a harmadikból két konstans sort kapunk, melyek egymás konstansszorosai, tehát a determináns értéke 0.

6.16. Vonjuk ki az első oszlopot a másodikból és a harmadikból. Az így kapott harmadik oszlop kétszerese a másodiknak, tehát a determináns értéke 0.

6.18. Első megoldás: vonjuk ki az első oszlopot a többiből, ezzel eltüntetve azokból a négyzetes tagot, majd vonjuk a második oszlop megfelelő skalárszorosát a harmadik és negyedik oszlopból, hogy elimináljuk azok lineáris tagját, végül a harmadik oszlop konstansszorosát vonjuk ki a negyedikből, hogy ott csak 0-k maradjanak.

Második megoldás: Elég megmutatnunk, hogy a determináns oszlopai lineárisan összefüggőek. Az $a^2x + (a+1)^2y + (a+2)^2z + (a+3)^2w = 0$ egyenlet a homogén

$$\begin{aligned} x + \ y + \ z + \ w &= 0 \\ 2y + 4z + 6w &= 0 \\ y + 4z + 9w &= 0 \end{aligned}$$

egyenletrendszerre vezet, aminek biztosan van nemtriviális megoldása, hisz 4 ismeretlenhez csak 3 egyenlet van adva. (Megoldani már szükségtelen, elég a megoldás létezését igazolni, de például az $(x, y, z, w) = (1, -3, 3, -1)$ egy megoldás).

6.20. Mivel a koordináták báziscserében való változásáról szóló 4.22. állításban láttuk, hogy a koordinátás alakokat a $[\mathbf{v}_i]_{\mathcal{C}} = \mathbf{A}_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{v}_i]_{\mathcal{B}}$ képlet kapcsolja össze, ezért a $\mathbf{v}$ vektorok koordinátás alakjaiból, mint oszlopvektorokból képzett mátrixokra $\mathbf{V}_{\mathcal{C}} = \mathbf{A}_{\mathcal{C}\leftarrow\mathcal{B}}\mathbf{V}_{\mathcal{B}}$, így determinánsaikra $|\mathbf{V}_{\mathcal{C}}| = |\mathbf{A}_{\mathcal{C}\leftarrow\mathcal{B}}||\mathbf{V}_{\mathcal{B}}|$.

6.21. Egyrészt $\det(\mathbf{A}) = \det(\mathbf{A}^T)$, másrészt mivel $\mathbf{A}^T = -\mathbf{A}$, ezért $\det(\mathbf{A}^T) = (-1)^n \det(\mathbf{A})$, azaz $\det(\mathbf{A}) = -\det(\mathbf{A})$, amiből $\det(\mathbf{A}) = 0$.

6.22. $|\mathbf{A}^2| = |\mathbf{A}|^2 = |\mathbf{A}||\mathbf{A}| = |\mathbf{A}||\mathbf{A}^T| = |\mathbf{A}\mathbf{A}^T|$.

6.23. Mindhárom determinánst a következőképpen számítjuk ki. Legyen $\mathbf{A}$ a determinánshoz tartozó mátrix. Tekintsük az $|\mathbf{A}\mathbf{A}^T|$ determinánst. Ezt könnyű kiszámítani (hisz a főátlón kívül csak nullák állnak), s ennek négyzetgyöke lesz a determináns értéke. Ezek alapján a három determináns értéke: $a^2 + b^2$, $(a^2 + b^2 + c^2 + d^2)^2$, $(a^2 + b^2 + c^2 + d^2 + e^2 + f^2 + g^2 + h^2)^4$.

6.24. A determinánsok szorzási szabályát is felhasználva:

$$\begin{aligned} (x_1^2 + x_2^2)(y_1^2 + y_2^2) &= \begin{vmatrix} x_1 & x_2 \\ -x_2 & x_1 \end{vmatrix} \begin{vmatrix} y_1 & y_2 \\ -y_2 & y_1 \end{vmatrix} \\ &= \begin{vmatrix} x_1y_1 - x_2y_2 & x_1y_2 + x_2y_1 \\ -x_2y_1 - x_1y_2 & -x_2y_2 + x_1y_1 \end{vmatrix} \\ &= (x_1y_1 - x_2y_2)^2 + (x_1y_2 + x_2y_1)^2. \end{aligned}$$

A négy illetve a nyolc négyzet összegére vonatkozó analóg összefüggések hasonlóan bizonyíthatóak. (Hurwitz bebizonyította, hogy ha n négyzetszám összegére igaz a feladatbelivel analóg összefüggés, akkor $n = 1, 2, 4$ vagy 8.)

6.25. 1. Igaz. 2. Igaz. 3. Igaz. 4. Igaz. 5. Hamis. Mátrixok összegének determinánsa általában nem egyenlő determinánsaik összegével (ld. a 6.26. feladatot). 6. Hamis. Egy aldetermináns értéke bármilyen előjelű lehet, az előjeles aldeterminánst belőle úgy kapjuk, hogy páratlan $i + j$ esetén megszorozzuk -1-gyel. 7. Hamis. 8. Hamis. Csak azokon a helyeken folytonos függvénye a mátrix elemeinek, ahol a determinánsa nem 0.

6.26.

a) $$\begin{bmatrix} 0 & 1 & 0 \\ 2 & 3 & 4 \\ 5 & 0 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 4 \\ 5 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 6 \end{bmatrix} = 8$$

b) $$\begin{vmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 0 \\ 0 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} 0 & 0 & 0 & 2 \\ 0 & 0 & 2 & 0 \\ 0 & 2 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{vmatrix} +$$
$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 \end{vmatrix} = 1 + 16 - 4 - 4 = 9$$

c) $$\begin{vmatrix}1&0&2&0&0\\0&2&0&0&1\\0&1&0&2&0\\2&0&0&1&0\\0&0&1&0&2\end{vmatrix}=\begin{vmatrix}1&0&0&0&0\\0&0&0&0&1\\0&1&0&0&0\\0&0&0&1&0\\0&0&1&0&0\end{vmatrix}+\begin{vmatrix}0&0&2&0&0\\0&2&0&0&0\\0&0&0&2&0\\2&0&0&0&0\\0&0&0&0&2\end{vmatrix}$$

6.27. Az első sor minden eleme páros, az első oszlop minden eleme osztható 3-mal, a második oszlop minden eleme osztható 5-tel, tehát minden kígyó osztható $2\cdot 3\cdot 5=30$-cal, így az összegük is.

6.28. Csak egyetlen kígyó áll csupa páratlan számból, így a kígyók összegére bontásnál csak annak determinánsa páratlan, a többié páros, összegük tehát páratlan, vagyis nem lehet 0.

6.29. Megadjuk, hogy melyik sorban hányadik elem lesz a kígyóba választva. A 12 kígyó: 1243, 1324, 1432, 2134, 2341, 2413, 3142, 3214, 3421, 4123, 4231, 4312. Ezek alapján a 12 determináns – a kígyó elemeit négyzettel jelölve:

$$\begin{vmatrix}\blacksquare&0&0&0\\0&\blacksquare&0&0\\0&0&0&\blacksquare\\0&0&\blacksquare&0\end{vmatrix}\quad\begin{vmatrix}\blacksquare&0&0&0\\0&0&\blacksquare&0\\0&\blacksquare&0&0\\0&0&0&\blacksquare\end{vmatrix}\quad\begin{vmatrix}\blacksquare&0&0&0\\0&0&0&\blacksquare\\0&0&\blacksquare&0\\0&\blacksquare&0&0\end{vmatrix}$$

$$\begin{vmatrix}0&\blacksquare&0&0\\\blacksquare&0&0&0\\0&0&\blacksquare&0\\0&0&0&\blacksquare\end{vmatrix}\quad\begin{vmatrix}0&\blacksquare&0&0\\0&0&\blacksquare&0\\0&0&0&\blacksquare\\\blacksquare&0&0&0\end{vmatrix}\quad\begin{vmatrix}0&\blacksquare&0&0\\0&0&0&\blacksquare\\\blacksquare&0&0&0\\0&0&\blacksquare&0\end{vmatrix}$$

$$\begin{vmatrix}0&0&\blacksquare&0\\\blacksquare&0&0&0\\0&0&0&\blacksquare\\0&\blacksquare&0&0\end{vmatrix}\quad\begin{vmatrix}0&0&\blacksquare&0\\0&\blacksquare&0&0\\\blacksquare&0&0&0\\0&0&0&\blacksquare\end{vmatrix}\quad\begin{vmatrix}0&0&\blacksquare&0\\0&0&0&\blacksquare\\0&\blacksquare&0&0\\\blacksquare&0&0&0\end{vmatrix}$$

$$\begin{vmatrix}0&0&0&\blacksquare\\\blacksquare&0&0&0\\0&\blacksquare&0&0\\0&0&\blacksquare&0\end{vmatrix}\quad\begin{vmatrix}0&0&0&\blacksquare\\0&\blacksquare&0&0\\0&0&\blacksquare&0\\\blacksquare&0&0&0\end{vmatrix}\quad\begin{vmatrix}0&0&0&\blacksquare\\0&0&\blacksquare&0\\\blacksquare&0&0&0\\0&\blacksquare&0&0\end{vmatrix}$$

6.30. Adjuk a harmadik oszlophoz az első 100-szorosát és a második 10-szeresét. Így az utolsó sorban a megadott, 9-cel osztható számok szerepelnek. Ha e sor szerint fejtjük ki a determinánst, akkor minden összeadandó osztható lesz 9-tel, tehát a determináns is.

6.31. Egy olyan determinánst kell konstruálni, melynek van egy nulla értékű aldeterminánsa. például a

$$\begin{vmatrix}1&2&1\\1&9&2\\1&1&1\end{vmatrix}$$

determináns értéke 1, de a 9-hez tartozó aldetermináns értéke 0, így a második sor vagy oszlop szerinti kifejtésben e szám 0-val szorzódik, vagyis nem befolyásolja a determináns értékét.

6.32.

a) $-6\cdot 6=-36$, mert a blokkmátrixok determinánsára vonatkozó tétel szerint a bal felső 2×2-es és a jobb alsó 3×3-as determinánsok szorzata adja az eredményt.

b) 24, mert a bal felső 1×1-es és a jobb alsó 4×4-es determinánsok értéke 1, illetve 24, és ezek szorzata 24. Másik megoldáshoz jutunk, ha a determinánst az első oszlopa, az egyetlen kiszámítandó aldeterminánst az első sora... szerint fejtjük ki.

6.33.

a) A determináns a 2, -1, -2, 1 számokból képezett Vandermonde-determináns, így értéke: $(-1-2)(-2-2)(1-2)(-2-(-1))(1-(-1))(1-(-2))=72$.

b) Vandermonde-determináns; értéke -2880.

c) A determináns két Vandermonde-determináns összegére bomlik:

$$\begin{vmatrix}1&a&a^2&a^3\\1&b&b^2&b^3\\1&c&c^2&c^3\\1&d&d^2&d^3\end{vmatrix}+\begin{vmatrix}1&a&a^2&a^3\\1&b&b^2&b^3\\1&c&c^2&c^3\\1&e&e^2&e^3\end{vmatrix}$$
$$=(b-a)(c-a)(c-b)$$
$$\times\left[(d-a)(d-b)(d-c)+(e-a)(e-b)(e-c)\right].$$

6.34. Ha $pqrs\neq 0$, akkor szorozzuk be az első sort p-vel, a másodikat q-val, a harmadikat r-rel, a negyediket s-sel, majd a negyedik oszlopból emeljünk ki $pqrs$-t; így egy Vandermonde-determinánst kapunk:

$$D=\frac{pqrs}{pqrs}\begin{vmatrix}p^3&p^2&p&1\\q^3&q^2&q&1\\r^3&r^2&r&1\\s^3&s^2&s&1\end{vmatrix}$$
$$=(q-p)(r-p)(s-p)(r-q)(s-q)(s-r).$$

Ha $pqrs=0$, például $s=0$, akkor az eredeti determináns negyedik oszlopa szerinti kifejtéssel kapjuk, hogy

$$D=pqr\begin{vmatrix}p^2&p&1\\q^2&q&1\\r^2&r&1\end{vmatrix}.$$

Ezekből rövid átalakítás után látható, hogy az összefüggés ebben az esetben is fennáll.

6.35. $a_1=\det[1]=1$, $a_2=\left[\begin{smallmatrix}1&-1\\1&1\end{smallmatrix}\right]=2$, az $(n\times n)$-es determinánst első sora szerint kifejtve kapjuk, hogy $a_n=a_{n-1}+a_{n-2}$.

6.37. Igen. Tekintsük a determináns első sor szerinti kifejtését! Ha mindegyik elemhez tartozó előjeles aldetermináns 0 lenne, akkor a mátrix szinguláris lenne, így valamelyik elemhez tartozó aldetermináns nem 0. Legyen pl. $A_{1j}\neq 0$. Ekkor a kifejtés összes többi tagját összevonva kapjuk, hogy $\det \mathbf{A} = a_{1j}A_{1j} + c$. Mivel $A_{1j} \neq 0$, ezért az $a_{1j}A_{1j} + c = 0$ egyenlet megoldható a_{ij}-re, tehát ennek az elemnek a megváltoztatása 0-vá teszi a determinánst.

6.38. Ha az i-edik sor elemeit az u-adik sorhoz tartozó előjeles aldeterminánsokkal szorozzuk, akkor az u-adik sor elemeit nem használjuk, tehát szabadon megváltoztathatjuk. Másoljuk az i-edik sort az u-adik helyére, tehát minden k-ra $a_{ik} = a_{uk}$. Ekkor egyrészt $\sum_{k=1}^n a_{ik}A_{uk} = \sum_{k=1}^n a_{uk}A_{uk}$, azaz e determináns u-adik sor szerinti kifejtését kaptuk, másrészt e determinánsnak van két azonos sora, tehát determinánsa 0. Az oszlopokra vonatkozó állítás egy transzponálással visszavezethető erre.

6.39. A két tétel képletei közös képletbe foglalhatók. Sorokra:

$$\sum_{k=1}^{n} a_{ik}A_{uk} = \begin{cases} \det \mathbf{A}, & \text{ha } i = u, \\ 0, & \text{ha } i \neq u, \end{cases} \tag{6.4}$$

oszlopokra:

$$\sum_{k=1}^{n} a_{kj}A_{kv} = \begin{cases} \det \mathbf{A}, & \text{ha } j = v, \\ 0, & \text{ha } j \neq v. \end{cases} \tag{6.5}$$

6.40. A két kifejtési tételből adódik, hogy

$$[a_{ij}][A_{ij}]^T = \det(\mathbf{A})\mathbf{I},$$

ugyanis $[a_{ij}]$ i-edik sorának és $[A_{ij}]^T$ u-adik oszlopának, azaz $[A_{ij}]$ u-adik sorának skaláris szorzata a (6.4) képlet szerint $\det(\mathbf{A})$, ha $i = u$, azaz a szorzat főátlójában, egyebütt pedig 0. Ebből pedig mindkét képlet adódik.

6.44.

a) Az előjeles aldeterminánsok mátrixának transzponáltja:

$$\begin{bmatrix} \begin{vmatrix} 2 & 7 \\ 1 & 4 \end{vmatrix} & -\begin{vmatrix} -7 & 7 \\ 2 & 4 \end{vmatrix} & \begin{vmatrix} -7 & 2 \\ 2 & 1 \end{vmatrix} \\ -\begin{vmatrix} 1 & 4 \\ 1 & 4 \end{vmatrix} & \begin{vmatrix} 3 & 4 \\ 2 & 4 \end{vmatrix} & -\begin{vmatrix} 3 & 1 \\ 2 & 1 \end{vmatrix} \\ \begin{vmatrix} 1 & 4 \\ 2 & 7 \end{vmatrix} & -\begin{vmatrix} 3 & 4 \\ -7 & 7 \end{vmatrix} & \begin{vmatrix} 3 & 1 \\ -7 & 2 \end{vmatrix} \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & -1 \\ 42 & 4 & -49 \\ -11 & -1 & 13 \end{bmatrix}$$

Mivel a mátrix determinánsa 1, ezért inverze megegyezik az előjeles aldeterminánsok előbb kiszámolt mátrixával.

b) Az előjeles aldeterminánsok mátrixának transzponáltja:

$$\begin{bmatrix} \begin{vmatrix} 0 & 2 \\ 2 & 1 \end{vmatrix} & -\begin{vmatrix} 2 & 2 \\ 3 & 1 \end{vmatrix} & \begin{vmatrix} 2 & 0 \\ 3 & 2 \end{vmatrix} \\ -\begin{vmatrix} 2 & 3 \\ 2 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 3 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 3 & 2 \end{vmatrix} \\ \begin{vmatrix} 2 & 3 \\ 0 & 2 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 2 & 2 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 2 & 0 \end{vmatrix} \end{bmatrix}^T = \begin{bmatrix} -4 & 4 & 4 \\ 4 & -8 & 4 \\ 4 & 4 & -4 \end{bmatrix}.$$

Mivel a mátrix determinánsa 16, ezért az inverz mátrix

$$\frac{1}{4}\begin{bmatrix} -1 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -1 \end{bmatrix}.$$

c) Mivel $\det(\mathbf{A}) = 1$, ezért $\mathbf{A}^{-1}$ megegyezik az előjeles aldeterminánsok mátrixának transzponáltjával. Ennek mind a 16 elemét nem kell kiszámolni, mert felső háromszögmátrix inverze felső háromszögmátrix. Hasonlóan könnyen látható, hogy a főátlóbeli elemekhez tartozó előjeles aldeterminánsok értéke 1. Tehát csak a főátló alatti elemek előjeles aldeterminánsait kell kiszámolni. Példaként egyet mutatunk:

$$A_{32} = (-1)^{3+2}\begin{vmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & 1 \end{vmatrix} = -2.$$

Hasonlóan kiszámolva a többit is kapjuk, hogy

$$\mathbf{A}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ -1 & 3 & -3 & 1 \end{bmatrix}^T = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Mi lehet e feladat általánosítása, és mi a válasz?

d) A mátrixból csak egy nemnulla kígyó választható ki, így determinánsa könnyen számolható: $\det \mathbf{B} = 16$. Az inverz kiszámításához nem kell sok aldeterminánst számolni, mert nagy részük láthatóan 0 értékű. Vegyük figyelembe a számolásnál azt is, hogy $\mathbf{B}$ szimmetrikus, így egyrészt a szimmetrikusan elhelyezkedő elemek közül csak az egyiket kell kiszámolni, másrészt a szimmetria miatt a végén szükségtelen a transzponálás.

$$\mathbf{A}^{-1} = \frac{1}{16}\begin{bmatrix} 0 & 8 & 0 & -8 \\ 8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 \\ -8 & 0 & 8 & 0 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2} & 0 & -\frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 & \frac{1}{2} & 0 \end{bmatrix}.$$

e) Az inverz

$$\frac{1}{abc}\begin{bmatrix} bc & 0 & 0 \\ 0 & ac & 0 \\ 0 & 0 & ab \end{bmatrix}^T = \begin{bmatrix} \frac{1}{a} & 0 & 0 \\ 0 & \frac{1}{b} & 0 \\ 0 & 0 & \frac{1}{c} \end{bmatrix},$$

ha $abc \neq 0$. Az $abc = 0$ esetben a mátrix nem invertálható.

f) $\begin{bmatrix} 1 & -1 \\ -1 & 1-i \end{bmatrix}$.

g) Az inverz $\begin{bmatrix} 0 & 0 & 0 & 1/4 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1/3 & 0 \\ 0 & 1/2 & 0 & 0 \end{bmatrix}$

h) $\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

6.46. *a)* A három eredmény: 1, 2, 4. Mivel mindhárom test esetén ugyanazokat a számolásokat kell elvégezni, csak más modulus szerinti maradék lesz az eredmény, legegyszerűbb, ha az egészek fölött számolunk, és annak maradékait tekintjük. Valóban, az egészek fölött -1 a determináns, és $-1 \bmod 2 = 1$, $-1 \bmod 3 = 2$, $-1 \bmod 5 = 4$. *b)* 5. Legegyszerűbb, ha az első sor 2-szeresét hozzáadjuk a második, és 3-szorosát a harmadik sorhoz.

6.47. Sage-kód egy $\mathbb{F}_2$ fölötti véletlen mátrix kiírására:

```
sage: random_matrix(GF(2), 5)
[1 0 0 1 1]
[1 1 1 0 1]
[1 1 1 0 0]
[1 0 0 0 0]
[0 0 1 0 0]
sage: _.det()
1
```

Hány olyan $\mathbb{F}_2^{5\times 5}$-beli mátrix van, amelynek determinánsa nem 0? Az első sora bármelyik vektor lehet, kivéve a $\mathbf{0}$-vektort, így $2^5 - 1$ lehetőség van. A második sor nem lehet ez a vektor és a $\mathbf{0}$-vektor, ez $2^5 - 2$ lehetőség. A harmadik vektor nem lehet az előző két vektor által kifeszített altér, melynek a $\mathbf{0}$-vektorral együtt $2^2 = 4$ eleme van, e vektor kiválasztására tehát $2^5 - 2^2$ lehetőség adódik. Hasonlóan folytatva kapjuk, hogy az összes független vektorötösök – azaz a nemnulla értékű determinánsok – száma $(2^5 - 2^0)(2^5 - 2^1)(2^5 - 2^2)(2^5 - 2^3)(2^5 - 2^4)$. Ha ezt elosztjuk az összes $\mathbb{F}_2^{5\times 5}$-beli mátrixok számával, 0.2980-t kapunk, így a determináns 0.7020 valószínűséggel lesz 0.

6.48. 3. Ha a megadott $n-1$ darab vektort $\mathbf{a}_1 = (a_{11}, a_{12}, \ldots, a_{1n}), \ldots, \mathbf{a}_{n-1} = (a_{n-1,1}, a_{n-1,2}, \ldots, a_{n-1,n})$ jelöli, akkor az a formula, amiben van -1-hatvány, az eddigiek alapján a következő:

$$(-1)^{n-1}\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \ldots & \mathbf{e}_n \\ a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \ldots & a_{n-1,n} \end{vmatrix}$$

Ennek könnyű olyan változatát kitalálni, melyben eltűnik a -1-hatvány, a következő kettő bármelyike jó:

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{e}_1 & \mathbf{e}_2 & \ldots & \mathbf{e}_n \end{vmatrix} \quad \text{vagy} \quad \begin{vmatrix} a_{11} & a_{21} & \ldots & \mathbf{e}_1 \\ a_{12} & a_{22} & \ldots & \mathbf{e}_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \ldots & \mathbf{e}_n \end{vmatrix}.$$

6.49. Az előző feladat szerint a kért vektort a következőképp kaphatjuk meg:

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 \end{vmatrix} = \mathbf{e}_4 - \mathbf{e}_3$$

azaz a negyedik vektor $(0, 0, -1, 1)$.

7 Mátrixleképezések és geometriájuk

Minden $\mathbf{A}$ mátrixhoz tartozik egy $\mathbf{x} \mapsto \mathbf{Ax}$ leképezés. E leképezések épp egybeesnek a lineáris kombinációt megtartó leképezésekkel, melyeket lineáris leképezéseknek nevezünk. Lineáris leképezés végtelen dimenziós terek között is értelmezhető, így e fogalom a mátrix általánosításának is tekinthető. A legfontosabb lineáris leképezések – vetítés, forgatás – szemléletes, geometriai megközelítést kínálnak.

Mátrixleképezés, lineáris leképezés

A mátrixleképezés fogalma *Mátrixhoz tartozó leképezésen*, *mátrix által indukált leképezésen* vagy egyszerűen *mátrixleképezésen* az $\mathbf{x} \mapsto \mathbf{Ax}$ leképezést értjük, ahol $\mathbf{A}$ egy mátrix. Például egy $m \times n$-es $\mathbf{A} \in \mathbb{R}^{m\times n}$ mátrixhoz így egy $\mathbb{R}^n \to \mathbb{R}^m$ leképezés tartozik, ugyanis ha $\mathbf{x} \in \mathbb{R}^n$ és $\mathbf{y} = \mathbf{Ax}$, akkor $\mathbf{y} \in \mathbb{R}^m$. A mátrixok jelölésére félkövér betűket használunk, a leképezésekére dőlt betűket. A továbbiakban azt a konvenciót követjük, hogy egy mátrixhoz tartozó mátrixleképezést ugyanannak a betűnek a dőlt változatával jelöljük, például az $\mathbf{A}$ mátrixhoz tartozó mátrixleképezést A jelöli, azaz

$$A : \mathbf{x} \mapsto A(\mathbf{x}) = \mathbf{Ax}.$$

Az $A(\mathbf{x})$ mellett az $A\mathbf{x}$ jelölés is használatos.

Az A leképezés értékkészletét $\mathrm{Im}(A)$ jelöli, mely az $\mathbb{R}^m$ altere. Ezt szokás *képtérnek* is nevezni, minthogy ez az $\mathbb{R}^n$ tér képe. Ez megegyezik az $\mathbf{A}$ mátrix oszlopterével, azaz $\mathcal{O}(\mathbf{A})$-val. Azoknak a vektoroknak az alterét, melyet A a nullvektorba visz, az A leképezés *magterének* nevezzük. Magtérre a *kernel* szó is használatos. $\mathrm{Ker}(A)$-val jelöljük. Ez megegyezik a hozzá tartozó $\mathbf{A}$ mátrix nullterével. Tehát

Az Im rövidítés a kép jelentésű *image*, a Ker a mag jelentésű *kernel* szóból származik.

$$\mathrm{Im}(A) = \mathcal{O}(\mathbf{A}), \qquad \mathrm{Ker}(A) = \mathcal{N}(\mathbf{A}).$$

Az új szóhasználat hamarosan értelmet fog kapni, amikor e leképezések fogalmát kiterjesztjük.

7.1. példa (Vektori szorzással definiált mátrixleképezés). *Legyen $\mathbf{a} = (a_1, a_2, a_3)$ egy adott $\mathbb{R}^3$-beli vektor. Legyen A az a transzformáció, mely a tér tetszőleges $\mathbf{x}$ vektorához az $\mathbf{a} \times \mathbf{x}$ vektort rendeli. Tehát*

$$A : \mathbb{R}^3 \to \mathbb{R}^3 : \mathbf{x} \mapsto \mathbf{a} \times \mathbf{x}.$$

Mutassuk meg, hogy az A függvény egy mátrixleképezés, azaz létezik egy olyan $\mathbf{A}$ mátrix, hogy $A(\mathbf{x}) = \mathbf{A}\mathbf{x}$.

Megoldás. Az $\mathbf{a} \times \mathbf{x}$ vektori szorzat koordinátás alakban:

$$\mathbf{y} = \mathbf{a} \times \mathbf{x} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} a_2x_3 - a_3x_2 \\ a_3x_1 - a_1x_3 \\ a_1x_2 - a_2x_1 \end{bmatrix}.$$

Az eredményből azonnal látszik, hogy e transzformáció mátrixleképezés, hisz $\mathbf{y}$ minden koordinátája $\mathbf{x}$ koordinátáinak lineáris kifejezése. A szorzatot $\mathbf{x}$ koordinátái szerint rendezzük, ahonnan azonnal leolvasható a transzformáció mátrixa, amit a továbbiakban $[\mathbf{a}]_\times$ jelöl. Segítségével fölírható a transzformáció mátrixszorzatos alakja:

$$\mathbf{a} \times \mathbf{x} = \begin{bmatrix} & - a_3x_2 & + a_2x_3 \\ a_3x_1 & & - a_1x_3 \\ - a_2x_1 & + a_1x_2 & \end{bmatrix} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Tehát

$$[\mathbf{a}]_\times = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}. \tag{7.1}$$

E feladatra később még többször visszatérünk. □

Műveletek mátrixleképezések között A mátrixleképezések $\mathbb{R}^n$-ből $\mathbb{R}^m$-be képző függvények (a későbbiekben $\mathbb{C}^n \to \mathbb{C}^m$ és $\mathbb{F}_p^n \to \mathbb{F}_p^m$ függvényekkel is foglalkozunk). Kérdés, hogy mi a kapcsolat e mátrixleképezések közti függvényműveletek és a mátrixműveletek között! Például ha az $\mathbf{A}$ és $\mathbf{B}$ mátrixhoz tartozó leképezés A és B (azaz $A : \mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ és $B : \mathbf{x} \mapsto \mathbf{B}\mathbf{x}$), akkor igaz-e, hogy $(A + B)(\mathbf{x}) = (\mathbf{A} + \mathbf{B})\mathbf{x}$, azaz hogy $\mathbf{A} + \mathbf{B}$-hez az $A + B$ leképezés tartozik.

7.2. tétel (Mátrixleképezések alapműveletei). *Legyen $\mathbf{A}$, $\mathbf{B}$ és $\mathbf{C}$ három $m \times n$-es mátrix, legyen A, B és C a hozzájuk tartozó három mátrixleképezés és legyen c egy skalár. Ekkor*

a) $\mathbf{A} + \mathbf{B} = \mathbf{C}$ *pontosan akkor igaz, ha $A + B = C$, és*

b) $c\mathbf{A} = \mathbf{C}$ *pontosan akkor igaz, ha $cA = C$.*

Ha $\mathbf{X}$, $\mathbf{Y}$ és $\mathbf{Z}$ típusa rendre $m \times k$, $k \times n$, illetve $m \times n$, és X, Y és Z a hozzájuk tartozó három mátrixleképezés, akkor *c)* $\mathbf{XY} = \mathbf{Z}$ *pontosan akkor igaz, ha* $X \circ Y = Z$*, azaz mátrixok szorzásának a függvények kompozíciója felel meg.*

Bizonyítás. A bizonyítások a mátrixműveletek tulajdonságaiból következnek. Ott, ahol valamelyik mátrixazonosságot használjuk, az ekvivalenciát kimondó nyíl fölé M-betűt írunk, ahol pedig a függvények közti műveleti tulajdonságokat használjuk, ott egy F-betűt:

a) $(A+B)(\mathbf{x}) = C(\mathbf{x}) \overset{F}{\iff} A(\mathbf{x}) + B(\mathbf{x}) = C(\mathbf{x}) \iff \mathbf{Ax} + \mathbf{Bx} = \mathbf{Cx}$ $\overset{M}{\iff} (\mathbf{A}+\mathbf{B})\mathbf{x} = \mathbf{Cx}$.

b) $(cA)(\mathbf{x}) = C(\mathbf{x}) \overset{F}{\iff} cA(\mathbf{x}) = C(\mathbf{x}) \iff c\mathbf{Ax} = \mathbf{Cx} \overset{M}{\iff}$ $(c\mathbf{A})\mathbf{x} = \mathbf{Cx}$

c) $(X \circ Y)(\mathbf{x}) = Z(\mathbf{x}) \overset{F}{\iff} X(Y(\mathbf{x})) = Z(\mathbf{x}) \iff \mathbf{X}(\mathbf{Yx}) = \mathbf{Zx} \overset{M}{\iff}$ $(\mathbf{XY})\mathbf{x} = \mathbf{Zx}$. □

Az $f : \mathbb{R}^n \to \mathbb{R}^n$ függvénynek a $g : \mathbb{R}^n \to \mathbb{R}^n$ függvény inverze, ha minden $\mathbf{x} \in \mathbb{R}^n$ helyen $(f(g(\mathbf{x})) = \mathbf{x}$ és $(g(f(\mathbf{x})) = \mathbf{x}$, azaz ha kompozícióik az $f \circ g$ és a $g \circ f$ függvények megegyeznek az identikus leképezéssel.

7.3. ÁLLÍTÁS (INVERZ MÁTRIXLEKÉPEZÉSEK). *Legyenek* A *és* B *az* $n \times n$*-es* $\mathbf{A}$ *és* $\mathbf{B}$ *mátrixokhoz tartozó mátrixleképezések. Ekkor az* $\mathbf{A}$ *mátrix inverze pontosan akkor a* $\mathbf{B}$ *mátrix, ha az* A *leképezés inverze a* B *leképezés.*

Bizonyítás. A 7.2. tételből tudjuk, hogy $(A \circ B)(\mathbf{x}) = \mathbf{ABx}$ és $(B \circ A)(\mathbf{x}) = \mathbf{BAx}$. Így ha $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$, azaz $\mathbf{A}$ inverze $\mathbf{B}$, akkor $(A \circ B)(\mathbf{x}) = \mathbf{ABx} = \mathbf{Ix} = \mathbf{x}$, és hasonlóan $(B \circ A)(\mathbf{x}) = \mathbf{BAx} = \mathbf{Ix} = \mathbf{x}$, azaz $A \circ B$ és $B \circ A$ az identikus leképezés. Hasonlóképp, ha $A \circ B$ és $B \circ A$ az identikus leképezés, akkor $(\mathbf{AB})\mathbf{x} = (A \circ B)(\mathbf{x}) = \mathbf{x}$ és $(\mathbf{BA})\mathbf{x} = (B \circ A)(\mathbf{x}) = \mathbf{x}$, így $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$, vagyis $\mathbf{A}$ és $\mathbf{B}$ egymás inverzei. □

Mátrixleképezések tulajdonságai E paragrafus tartalma az, hogy a mátrixleképezések megőrzik a lineáris kombinációkat, azaz vektorok egy lineáris kombinációjának képe a vektorok képének azonos lineáris kombinációjával egyenlő.

7.4. TÉTEL (A LINEÁRIS KOMBINÁCIÓT MEGŐRZŐ LEKÉPEZÉSEK). *Legyen* $A : \mathbb{R}^n \to \mathbb{R}^m$ *egy tetszőleges függvény. A következő három állítás ekvivalens.*

a) Létezik egy olyan $m \times n$*-es* $\mathbf{A}$ *mátrix, hogy az* A *függvény megegyezik az* $\mathbf{x} \mapsto \mathbf{Ax}$ *leképezéssel, azaz* A *az* $\mathbf{A}$ *mátrixhoz tartozó mátrixleképezés.*

b) Az A *függvény* homogén *és* additív*, azaz tetszőleges* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *vektorra és* $c \in \mathbb{R}$ *skalárra*

1. $A(c\mathbf{x}) = cA(\mathbf{x})$ *(A homogén),*
2. $A(\mathbf{x}+\mathbf{y}) = A(\mathbf{x}) + A(\mathbf{y})$ *(A additív).*

c) Az A függvény megőrzi a lineáris kombinációt, azaz tetszőleges $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ vektorra és $c, d \in \mathbb{R}$ skalárra $A(c\mathbf{x} + d\mathbf{y}) = cA(\mathbf{x}) + dA(\mathbf{y})$.

BIZONYÍTÁS. *a)* $\Rightarrow$ *b)*: Ha van olyan $\mathbf{A}$ mátrix, hogy minden $\mathbf{x}$ vektorra $A\mathbf{x} = \mathbf{A}\mathbf{x}$, akkor

$$\begin{aligned} A(c\mathbf{x}) &= \mathbf{A}(c\mathbf{x}) = c\mathbf{A}\mathbf{x} = cA\mathbf{x}, \text{ és} \\ A(\mathbf{x} + \mathbf{y}) &= \mathbf{A}(\mathbf{x} + \mathbf{y}) = \mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{y} = A\mathbf{x} + A\mathbf{y}. \end{aligned}$$

b) $\Rightarrow$ *c)*: Ha A homogén és additív, akkor

$$A(c\mathbf{x} + d\mathbf{y}) = A(c\mathbf{x}) + A(d\mathbf{y}) = cA\mathbf{x} + dA\mathbf{y},$$

ahol az első egyenlőségnél az A additivitását, a másodiknál a homogenitását használtuk.
c) $\Rightarrow$ *a)*: Tekintsük $\mathbb{R}^n$ standard bázisát és az $A\mathbf{e}_i$ vektorokból képzett

$$\mathbf{A} = [A\mathbf{e}_1 | A\mathbf{e}_2 | \dots | A\mathbf{e}_n] \tag{7.2}$$

mátrixot, valamint legyen az $\mathbf{x} \in \mathbb{R}^n$ egy tetszőleges vektor. Ha A olyan leképezés, mely megőrzi a lineáris kombinációkat, akkor

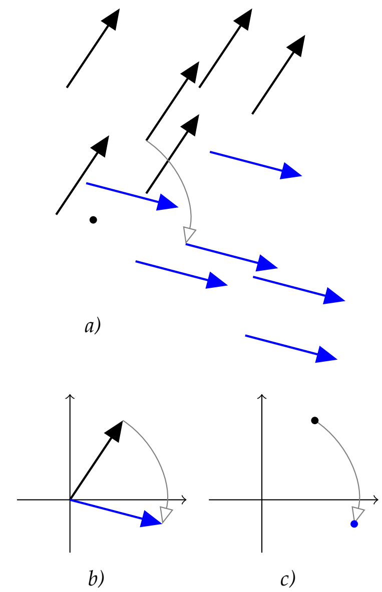
$$\begin{aligned} A\mathbf{x} &= A(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \dots + x_n\mathbf{e}_n) \\ &= x_1 A\mathbf{e}_1 + x_2 A\mathbf{e}_2 + \dots + x_n A\mathbf{e}_n \\ &= \begin{bmatrix} A\mathbf{e}_1 & A\mathbf{e}_2 & \dots & A\mathbf{e}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= \mathbf{A}\mathbf{x} \end{aligned}$$

Tehát valóban létezik olyan $\mathbf{A}$ mátrix, hogy $A\mathbf{x} = \mathbf{A}\mathbf{x}$. Ráadásul ilyen mátrix csak ez az egy van, mert bármely $\mathbf{e}_i$ bázisvektorra és bármely $\mathbf{A}$ mátrixra $\mathbf{A}\mathbf{e}_i = \mathbf{A}_{*i}$, tehát az $\mathbf{A}_{*i}$ oszlopvektor csak $A\mathbf{e}_i$ lehet. □

► A *c)* pont csak két vektor lineáris kombinációjáról szól, de ebből indukcióval adódik bármely lineáris kombinációra az állítás. Ezt a bizonyításban ki is használtuk.
► A tétel bizonyításában a (7.2) képlettel azt is megkaptuk, hogy hogyan írható föl egy lineáris kombinációt megőrző leképezés mátrixa!
► A tétel *b)* és *c)* pontja lehetővé teszi a mátrixleképezés általánosítását. Ezt tesszük a lineáris leképezésekről szóló paragrafusban.

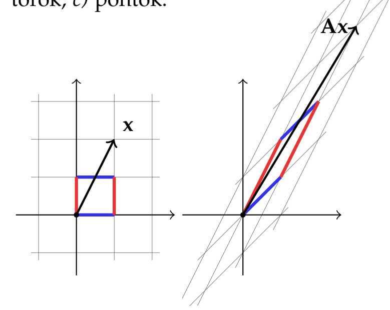
A mátrixleképezés hatásának szemléltetései Egy mátrixszal való szorzásnak egy adott alkalmazásban konkrét jelentése van (pl. gazdasági, fizikai,...). Ilyen esetekben a mátrixot használónak elég lehet a tárgyban való ismerete ahhoz, hogy „értse”, mi történik. De még ilyenkor is segíthet a matematikai eszközök megértésében, ha absztraktabb, vagy vizuálisan is megjeleníthető képet is tud alkotni magának. Most a mátrixleképezések vizuális megjelenítésének lehetőségeit tekintjük át.

Egy mátrixszal való szorzás hatása lehet például az, hogy a sík minden vektorát elforgatja β szöggel. Ez szemléltethető szabad vektorokkal, helyvektorokkal, és pontokkal is. Ha szabad vektorokként ábrázoljuk a vektorokat, akkor – mivel az azokat ábrázoló irányított szakaszok szabadon eltolhatók – a forgatás középpontja is szabadon megválasztható az ábrázolásban. Ha helyvektorokkal ábrázoljuk a vektorokat, akkor a forgatás origó körüli – csak e pont körüli forgatás visz helyvektort helyvektorba. Ugyanez a helyzet, ha a vektorokat pontokkal, a helyvektorok végpontjaival ábrázoljuk. A 7.1 ábra e három ábrázolást mutatja. Ezzel a módszerrel általában csak egy-egy vektor képe ábrázolható, az egész leképezés ritkán.



7.1. ábra: Vektorok elforgatásának szemléltetései a vektor különböző ábrázolásai szerint: *a)* szabad vektorok, *b)* helyvektorok, *c)* pontok.

$\mathbb{R}^2 \to \mathbb{R}^2$ leképezések esetén a legegyszerűbb szemléltetéshez elég csak az egységnégyzet képének megrajzolása, amint azt a 7.2 ábra mutatja. A kép mindig egy parallelogramma (esetleg elfajuló), melynek körüljárását is jelölni kell valahogy. Ezt az ábrán az oldalak különböző színezése megteszi. A parallelogramma területe a determináns abszolút értékét adja. Segíti az ábrázolást, ha nem csak az egységnégyzetet, hanem az egységnégyzetrács egy részének képe is meg van rajzolva. Ez egy parallelogrammarács. Ennek segítségével egy tetszőleges vektor képének megszerkesztése egyszerű, hisz a lineáris leképezés megtartja a lineáris kombinációt (ld. 7.2 ábra).



7.2. ábra: Az $\mathbf{A} = \left[\begin{smallmatrix}1 & 1\\ 1 & 2\end{smallmatrix}\right]$ mátrix hatása az egységnégyzetrácson, és az $\mathbf{x} = (1, 2)$ vektoron. Az $\mathbf{Ax}$ vektor végpontja a parallelogrammarácson 1 lépés az x-tengely képének irányába, és 2 lépés az y-tengely képének irányába.

7.5. példa (Mátrixleképezés ábrázolása az egységrács képével). *Ábrázoljuk az egységnégyzet és az egységnégyzetrács képét az alábbi mátrixok esetében:*

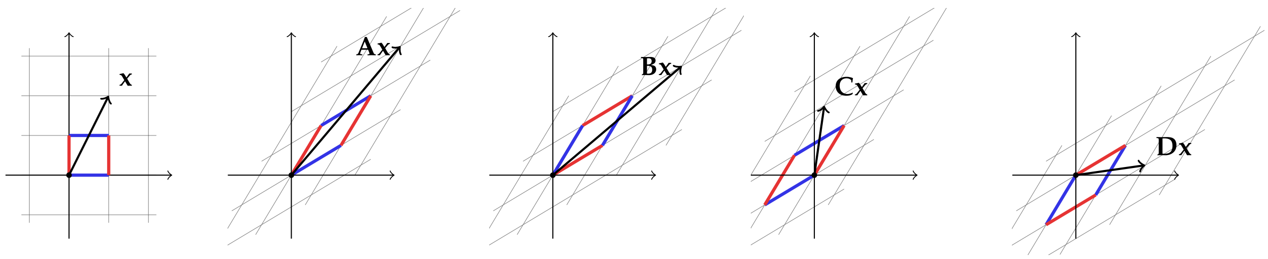
$$\mathbf{A} = \begin{bmatrix} \frac{5}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{5}{4} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \frac{3}{4} & \frac{5}{4} \\ \frac{5}{4} & \frac{3}{4} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} -\frac{5}{4} & \frac{3}{4} \\ -\frac{3}{4} & \frac{5}{4} \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} -\frac{3}{4} & \frac{5}{4} \\ -\frac{5}{4} & \frac{3}{4} \end{bmatrix}.$$

A rács képét használva ábrázoljuk az $(1, 2)$ vektor képét mind a négy esetben és határozzuk meg a determináns előjelét.

Megoldás. Világos, hogy itt négy különböző mátrixról, így négy különböző mátrixleképezésről van szó, a rácsok mégis azonosak, hisz mindegyiket az $(\frac{5}{4}, \frac{3}{4})$ és a $(\frac{3}{4}, \frac{5}{4})$ vektor (vagy -1-szerese) feszíti ki. A különbség abban van, hogy az egységnégyzet képe és annak körüljárása nem azonos. Az alábbi ábráról leolvashatók a különbségek.

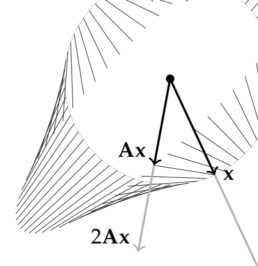
Az egységnégyzet körüljárásából leolvasható a determináns előjele is. Mivel a determinánsok abszolút értéke mindegyik esetben 1, így az előjel ismeretében a determináns is megadható: $\det(\mathbf{A}) = \det(\mathbf{D}) = 1$, $\det(\mathbf{B}) = \det(\mathbf{C}) = -1$. □

Egy másik ábrázolási lehetőség még fontosabb lehet ennél. Mivel a



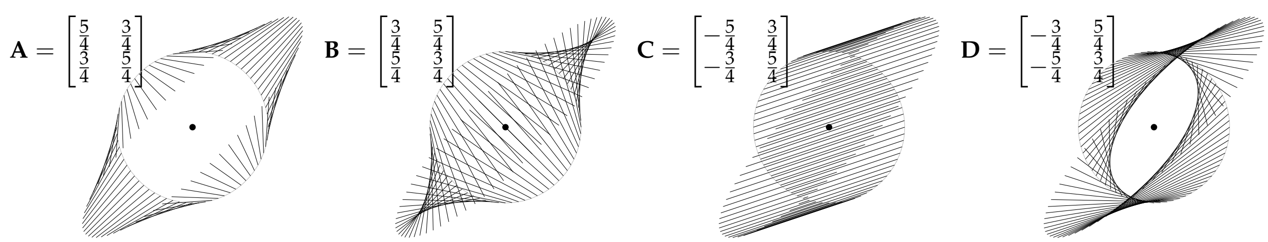
7.3. ábra: Az egységnégyzet és az egységnégyzetrács képe a megadott négy mátrix által, és az $(1,2)$ vektor képe.

mátrixleképezés homogén, azaz egy vektor c-szereséhez a vektor képének c-szeresét rendeli, ezért elég minden irányból egyetlen vektor, például az egységvektor képének megszerkesztése. Hogy a kép áttekinthető legyen, helyvektorok helyett csak az őket reprezentáló pontokat tekintjük, és az egységkör összes pontja helyett csak néhányat (pl. 50–100-at). Mivel a kör lineáris leképezés általi képe mindig egy ellipszis (esetleg elfajuló), ezért a leképezés szemléltetéséhez elég összekötni az egységkörön kiválasztott pontot a képével ahhoz, hogy nagyjából a sík bármelyik vektorának „lássuk”, hogy mi a képe. Az így kialakuló ábra sokat elmond a leképezésről. Ezt mutatja a 7.4 ábra, ahol a $-65°$-os irányhoz tartozó $\mathbf{x}$ egységvektort és $\mathbf{Ax}$ képét külön berajzoltuk, és azt is megmutattuk, hogy pl. hogyan kapható meg $2\mathbf{x}$ képe.



7.4. ábra: Az $\mathbf{A} = \left[\begin{smallmatrix} 5/4 & 3/4 \\ 3/4 & 5/4 \end{smallmatrix}\right]$ mátrix hatását szemlélteti oly módon, hogy az egységkör néhány pontját összeköti a képükkel. Az ábrán egy egységnyi $\mathbf{x}$ vektort és $\mathbf{Ax}$ képét, valamint a $2\mathbf{x}$ vektort és képét a $\mathbf{A}(2\mathbf{x}) = 2\mathbf{Ax}$ vektort kiemeltük.

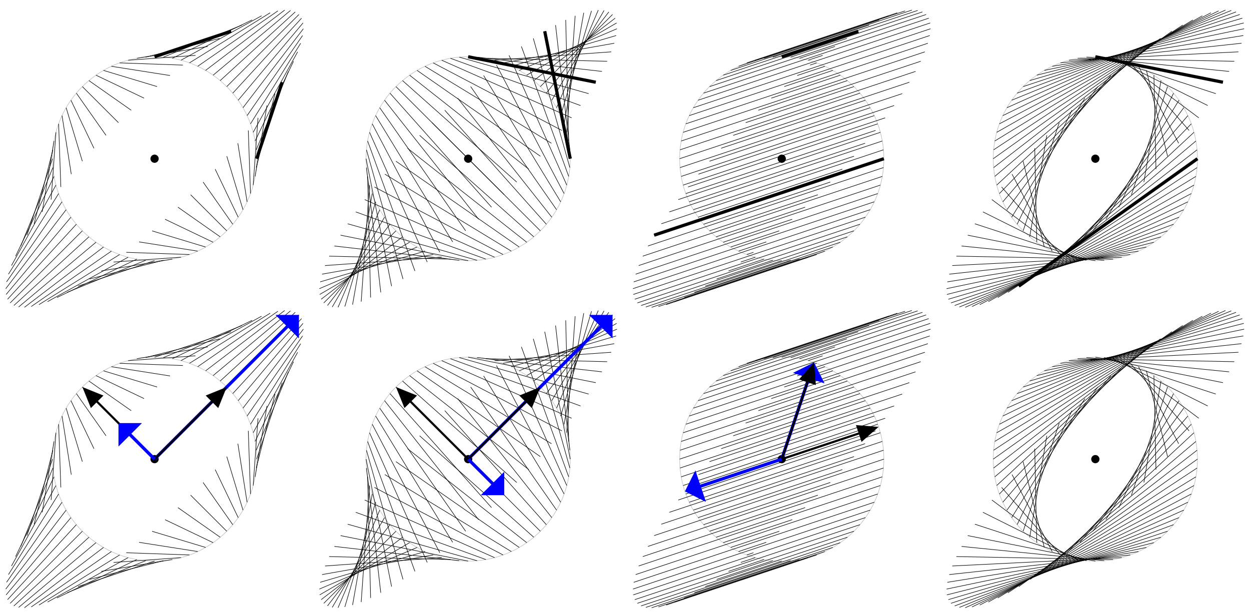
7.6. PÉLDA (MÁTRIXLEKÉPEZÉS ÁBRÁZOLÁSA AZ EGYSÉGKÖR KÉPÉVEL). *Az alábbi 7.5 ábra az előző példabeli mátrixleképezések egységkör-ábráját mutatja. Olvassuk le az ábráról az $(1,0)$ és a $(0,1)$ vektorok képét! Keressünk olyan egységvektorokat mind a négy ábrán, melyek képe saját maguk skalárszorosa!*



7.5. ábra: Négy leképezés egységkör-ábrája.

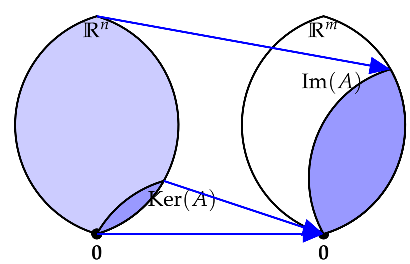
MEGOLDÁS. A megoldás leolvasható a 7.6 ábráról. Az ábra első sorában látható képeken az $(1,0)$ és $(0,1)$ vektorok végpontjából induló és a leképezés általi képükbe vezető vonalak kiemelésével jeleztük az első kérdésre adandó választ.

A második kérdésben olyan vonalakat keresünk, melyek a kör középpontja felé mutatnak (miért?). Ide tartozik az az eset is, amikor egy egységvektor képe önmaga, mert az saját maga 1-szerese. Ekkor azonban a két pontot összekötő vonal nem látszik. Hogy a második kérdésre adandó válasz világos legyen, a vektorokat is berajzoltuk, az egységvektort fekete, a képét kék színnel jelölve. A második sor első két ábráján a $\frac{\pi}{4}$, $\frac{3\pi}{4}$, $\frac{5\pi}{4}$, $\frac{7\pi}{4}$ irányok ilyenek. A négy vektor közül csak kettőt és ezek képét rajzoltuk be, a másik kettő a tükörképük (-1-szeresük). Az első ábrán a $\pi/4$ irányú egységvektor a 2-szeresébe, a $3\pi/4$ irányú egységvektor a 0.5-szeresébe megy, a harmadik ábráról az irányok nehezebben olvashatók le, a negyedik ábrán nem találni ilyen irányt. □



7.6. ábra: Négy leképezés egységkör-ábrája a bázisvektorokat és képüket összekötő vonalak kiemelésével, és a saját maguk konstansszorosába képzett vektorokkal.

Nehezebb a helyzet, ha egy általános $\mathbb{R}^n \to \mathbb{R}^m$ leképezést szeretnénk szemléltetni. Ehhez a levéldiagramot hívjuk segítségül. Itt elsőként a képteret és a magteret tudjuk szemléltetni, amint azt a 7.7. ábra mutatja.



7.7. ábra: Egy $A : \mathbb{R}^n \to \mathbb{R}^m, \mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ mátrixleképezés levéldiagrammja. Az ábrán három altér színezéssel ki van emelve – az értelmezési tartomány ($\mathbb{R}^n$), az értékkészlet (Im$(A) = \mathcal{O}(\mathbf{A})$) és a magtér (Ker$(A) = \mathcal{N}(\mathbf{A})$).

Lineáris leképezés A lineáris leképezést a mátrixleképezés tulajdonságaiból általánosítjuk:

7.7. DEFINÍCIÓ (LINEÁRIS LEKÉPEZÉS). *Legyen H_1 és H_2 mindegyike olyan halmaz, melynek elemein értelmezve van egy összeadás és egy „$\mathbb{R}$-beli skalárral való szorzás" művelet. Azt mondjuk, hogy egy $A : H_1 \to H_2$ leképezés* lineáris, *ha homogén és additív, azaz ha tetszőleges $\mathbf{x}, \mathbf{y} \in H_1$ elemekre és $c \in \mathbb{R}$ skalárra*

$$A(c\mathbf{x}) = cA(\mathbf{x}), \qquad \text{(homogén), és}$$
$$A(\mathbf{x}+\mathbf{y}) = A(\mathbf{x}) + A(\mathbf{y}), \qquad \text{(additív).}$$

$H_1 = H_2$ esetén a lineáris leképezéseket lineáris transzformációknak *is nevezzük.*

7.8. PÉLDA (A DERIVÁLÁS ÉS AZ INTEGRÁLÁS LINEÁRIS LEKÉPEZÉS). *Legyen H_1 az egyváltozós valós, és minden valós helyen differenciálható függvények halmaza, H_2 pedig az egyváltozós valós függvények halmaza. Világos, hogy H_1 és H_2 is olyan halmaz, melynek elemei közt értelmezve van az összeadás és a skalárral való szorzás művelete. A deriválás, azaz a $D : H_1 \to H_2 : f \mapsto D(f) = f'$ leképezés lineáris. Fogalmazzunk meg hasonló állítást az integrálra is.*

MEGOLDÁS. Tetszőleges $c \in \mathbb{R}$ skalárra és $f, g \in H_1$ függvényre

$$D(cf) = (cf)' = cf' = cD(f), \text{ és}$$
$$D(f+g) = (f+g)' = f' + g' = D(f) + D(g).$$

Hasonló összefüggések állnak fönn az integrálokra is, például legyen H_1 a $[0,1]$ intervallumon Riemann-integrálható függvények halmaza, és legyen $H_2 = \mathbb{R}$. Ekkor az $f \mapsto \int_0^1 f$ leképezés lineáris, ugyanis tetszőleges $c \in \mathbb{R}$ skalárra és tetszőleges $f, g \in H_1$ függvényre

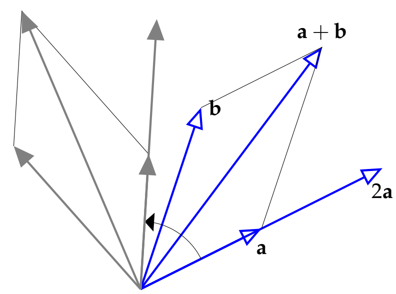
$$\int_0^1 cf = c\int_0^1 f, \text{ és } \int_0^1 (f+g) = \int_0^1 f + \int_0^1 g. \qquad \square$$

7.9. ÁLLÍTÁS (SÍKBELI FORGATÁS, TÜKRÖZÉS, VETÍTÉS). *A síkbeli vektorok egy rögzített O pont körüli forgatása, egy egyenesre való tükrözése és merőleges vetítése lineáris leképezés.*

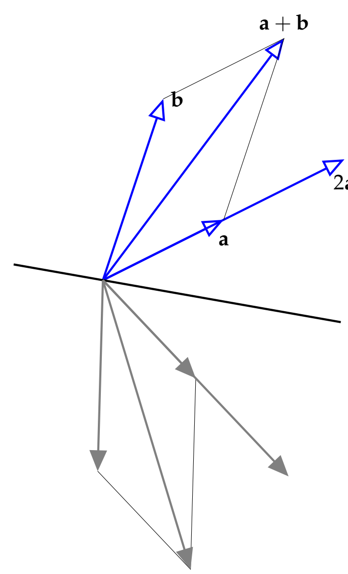
BIZONYÍTÁS. A precíz bizonyításokat mellőzük, csak a tényt szemléltetjük. A síkbeli vektorok pont körüli forgatása lineáris leképezés, ugyanis könnyen látható, hogy egy vektor c-szeresének ($c \in \mathbb{R}$) elforgatottja megegyezik a vektor elforgatottjának c-szeresével, valamint hogy két vektor összegének elforgatottja megegyezik a vektorok elforgatottjainak összegével (ld. 7.8. ábra).

Hasonlóan egyszerűen látszik, hogy egy egyenesre való tükrözés egy vektor c-szeresét a vektor tükörképének c-szeresébe viszi, és két vektor összegét a két vektor tükörképének összegébe (ld. 7.9 ábra).

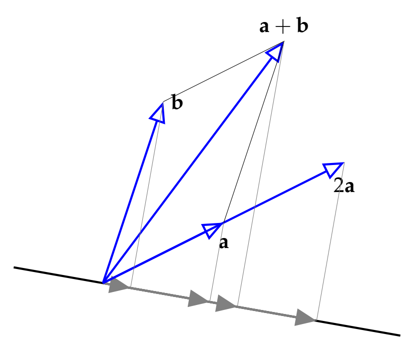
Végül ugyanígy megmutatható, hogy egy egyenesre való merőleges vetítés egy vektor c-szeresét vetületének c-szeresébe viszi, és két vektor összegét a két vektor vetületének összegébe (ld. 7.10 ábra). $\square$



7.8. ábra: A pont körüli elforgatás lineáris leképezés



7.9. ábra: Az egyenesre való tükrözés lineáris leképezés



7.10. ábra: Az egyenesre való merőleges vetítés lineáris leképezés

Most visszatérünk a a 7.4. tételhez, melynek bizonyításában megkonstruáltuk egy lineáris $\mathbb{R}^n \to \mathbb{R}^m$ leképezés mátrixát.

7.10. KÖVETKEZMÉNY (LINEÁRIS LEKÉPEZÉS MÁTRIXA). *Jelölje $\mathbf{L}$ az $L : \mathbb{R}^n \to \mathbb{R}^m$ leképezés standard bázisbeli mátrixát. Ekkor*

$$\mathbf{L} = [L\mathbf{e}_1 | L\mathbf{e}_2 | \dots | L\mathbf{e}_n],$$

ahol $\mathbf{e}_1$, $\mathbf{e}_2$,..., $\mathbf{e}_n$ a standard bázis vektorai. Eszerint egy tetszőleges $\mathbf{x} \in \mathbb{R}^n$ vektorra $L(\mathbf{x}) = \mathbf{L}\mathbf{x}$.

A bizonyítás kiolvasható a 7.4. tétel bizonyításának *c)* $\Longrightarrow$ *a)* részéből.

► A lineáris leképezés és a mátrixleképezés közti különbség tisztázása fontos. Mint azt a 7.8. példa mutatja, lineáris leképezésekről olyan esetben is beszélhetünk, amikor a leképezésnek nincs mátrixa, azaz a lineáris leképezés általánosabb fogalom.
► Különbség van a lineáris leképezés és a mátrixleképezés közt $\mathbb{R}^n \to \mathbb{R}^m$ függvények esetén is. Minden mátrixleképezés lineáris leképezés, és minden lineáris leképezéshez tartozik egyetlen mátrixleképezés a standard bázisban. A különbség lényege az, hogy a lineáris leképezés független a bázistól, az csak maga a függvény, míg a mátrixleképezés mindig valamely bázisra vonatkozik. Egy lineáris leképezéshez minden bázisban tartozik egy mátrixleképezés, de ezt a leképezést megadó mátrix függ a bázistól.
► Érdemes jellemezni a lineáris $\mathbb{R} \to \mathbb{R}$ transzformációkat. Ehhez először lássuk tisztán, hogy $\mathbb{R}$ elemei az 1-dimenziós vektorok, azaz a számok. E térben az $e = 1$ vektor (szám) a bázis. Az előző tétel szerint egy lineáris $L : \mathbb{R} \to \mathbb{R}$ transzformáció mátrixa $[Le] = [L(1)]$, ami megintcsak egyetlen szám, jelölje ezt $c := L(1)$. Így $L(x) = L(1x) = cx$, azaz a lineáris $\mathbb{R} \to \mathbb{R}$ transzformációk azonosak az $x \mapsto cx$ függvényekkel, ahol c egy tetszőleges konstans. E leképezések grafikonja egy origón átmenő egyenes.
► Az $\mathbb{R} \to \mathbb{R}$ lineáris leképezések tehát nem azonosak az lineáris $\mathbb{R} \to \mathbb{R}$ függvényekkel, melyek általános alakja $f(x) = cx + b$, ahol $c, d \in \mathbb{R}$. Egy ilyen függvény pontosan akkor lesz lineáris leképezés, ha $b = 0$.

7.11. PÉLDA. *Mutassuk meg, hogy az*

$$L : \mathbb{R}^2 \to \mathbb{R}^3; (x, y) \mapsto (x - y, 2x + y, -x)$$

leképezés lineáris leképezés, és írjuk fel a mátrixát, majd ellenőrizzük tetszőleges $\mathbf{x}$ vektorra az $L(\mathbf{x}) = \mathbf{L}\mathbf{x}$ összefüggést!

MEGOLDÁS. *1. megoldás:* Alakítsuk át a függvény értékéül kapott vektort:

$$L\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x-y \\ 2x+y \\ -x \end{bmatrix} = x\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + y\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 2 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix},$$

ami igazolja, hogy L mátrixleképezés, és mátrixa

$$\mathbf{L} = \begin{bmatrix} 1 & -1 \\ 2 & 1 \\ -1 & 0 \end{bmatrix}.$$

A mátrixleképezések pedig lineáris leképezések, tehát L is. Az $L(\mathbf{x}) = \mathbf{Lx}$ összefüggésen e megoldásban nincs mit ellenőrizni, hisz $\mathbf{L}$ épp úgy született, hogy ez igaz legyen.

2. *megoldás:* L linearitása a definíció alapján is bizonyítható, bár itt kicsit körülményesebb. L homogén, ugyanis

$$\begin{aligned} L(c(x,y)) &= L(cx,cy) = (cx-cy, 2cx+cy, -cx) = c(x-y, 2x+y, -x) \\ &= cL(x,y). \end{aligned}$$

L additív, ugyanis

$$\begin{aligned} L((x_1,y_1)+(x_2,y_2)) &= L(x_1+x_2, y_1+y_2) \\ &= (x_1+x_2-y_1-y_2, 2x_1+2x_2+y_1+y_2, -x_1-x_2) \\ &= (x_1-y_1, 2x_1+y_1, -x_1) + (x_2-y_2, 2x_2+y_2, -x_2) \\ &= L(x_1,y_1) + L(x_2,y_2) \end{aligned}$$

Az L mátrixa pedig felírható a bázisvektorok képe segítségével:

$$L(1,0) = (1,2,-1), \quad L(0,1) = (-1,1,0),$$

így a leképezés mátrixa e két vektorból, mint oszlopvektorból képzett mátrix lesz, azaz

$$\mathbf{L} = \begin{bmatrix} 1 & -1 \\ 2 & 1 \\ -1 & 0 \end{bmatrix}.$$

Az $L(\mathbf{x}) = \mathbf{Lx}$ ellenőrzése:

$$\mathbf{Lx} = \begin{bmatrix} 1 & -1 \\ 2 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x-y \\ 2x+y \\ -x \end{bmatrix}.$$

□

Lineáris leképezések alaptulajdonságai A lineáris leképezések legfontosabb tulajdonsága, hogy alteret altérbe visz, így speciálisan a zérusvektort a zérusvektorba.

7.12. TÉTEL (LINEÁRIS LEKÉPEZÉSEK ALAPTULAJDONSÁGAI). *Legyen $A : \mathbb{R}^n \to \mathbb{R}^m$ egy tetszőleges lineáris leképezés. Ekkor*

a) $A\mathbf{0} = \mathbf{0}$,

b) tetszőleges altér képe altér,

c) tetszőleges eltolt altér képe eltolt altér.

BIZONYÍTÁS. A bizonyításokat a lineáris leképezés definíciójából vezetjük le. *a)* igaz, mert bármely $\mathbf{x}$ vektorra $A\mathbf{0} = A(0\mathbf{x}) = 0A(\mathbf{x}) = \mathbf{0}$. *b)* abból következik, hogy ha $\mathbf{b}_1, \dots, \mathbf{b}_k$ egy $\mathcal{U}$ altér bázisa, akkor ezek összes lineáris kombinációjának, vagyis az altér vektorainak képe

$$A(c_1\mathbf{b}_1 + \ldots + c_k\mathbf{b}_k) = c_1A(\mathbf{b}_1) + \ldots + c_kA(\mathbf{b}_k).$$

Világos, hogy e vektorok kiadják az $A\mathbf{b}_1, \ldots, A\mathbf{b}_k$ vektorok által kifeszített altér minden vektorát, tehát $A(\mathcal{U})$ altér. Hasonlóan *c)*-ben, ha $\mathbf{u} \in \mathbb{R}^n$ egy tetszőleges vektor és $\mathcal{U}$ a fenti altér, akkor

$$\begin{aligned} A(\mathbf{u} + \mathcal{U}) &= A(u + c_1\mathbf{b}_1 + \ldots + c_k\mathbf{b}_k) \\ &= A(\mathbf{u}) + c_1 A(\mathbf{b}_1) + \ldots + c_k A(\mathbf{b}_k) \\ &= A(\mathbf{u}) + A(\mathcal{U}) \end{aligned}$$

ami egy altér eltoltja. □

Lineáris leképezés mátrixa különböző bázisokban Tegyük fel, hogy az L lineáris leképezés mátrixa az $\mathcal{A}$ bázisban $\mathbf{L}_\mathcal{A}$, a $\mathcal{B}$ bázisban $\mathbf{L}_\mathcal{B}$, és az $\mathcal{A}$ bázisról a $\mathcal{B}$-re való áttérés mátrixa $\mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}$. Kérdés, hogy mi a kapcsolat e három mátrix között.

A válasz egyszerűen megadható, ha megvizsgáljuk egy tetszőleges $\mathbf{x}$ vektornak és $L\mathbf{x}$ képének koordinátás alakját. Ezeket jelölje $[\mathbf{x}]_\mathcal{A}$, $[\mathbf{x}]_\mathcal{B}$, $[L\mathbf{x}]_\mathcal{A}$, $[L\mathbf{x}]_\mathcal{B}$. Az áttérés mátrixa köztük a következő kapcsolatokat létesíti:

$$[\mathbf{x}]_\mathcal{B} = \mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}[\mathbf{x}]_\mathcal{A}, \quad [L\mathbf{x}]_\mathcal{B} = \mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}[L\mathbf{x}]_\mathcal{A},$$

az $\mathbf{L}_\mathcal{A}$ és $\mathbf{L}_\mathcal{B}$ mátrixok pedig a következőket:

$$\mathbf{L}_\mathcal{A}[\mathbf{x}]_\mathcal{A} = [L\mathbf{x}]_\mathcal{A}, \quad \mathbf{L}_\mathcal{B}[\mathbf{x}]_\mathcal{B} = [L\mathbf{x}]_\mathcal{B}.$$

Ezeket összevetve kapjuk, hogy

$$\mathbf{L}_\mathcal{B}[\mathbf{x}]_\mathcal{B} = [L\mathbf{x}]_\mathcal{B} = \mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}[L\mathbf{x}]_\mathcal{A} = \mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}\mathbf{L}_\mathcal{A}[\mathbf{x}]_\mathcal{A},$$

azaz

$$\mathbf{L}_\mathcal{B}\mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}[\mathbf{x}]_\mathcal{A} = \mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}\mathbf{L}_\mathcal{A}[\mathbf{x}]_\mathcal{A}$$

vagyis csak mátrixokra fölírva:

$$\mathbf{L}_\mathcal{B}\mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}} = \mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}\mathbf{L}_\mathcal{A} \text{ vagy átrendezve } \mathbf{L}_\mathcal{A} = \mathbf{C}^{-1}_{\mathcal{B}\leftarrow\mathcal{A}}\mathbf{L}_\mathcal{B}\mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}.$$

A bizonyítás – és általában e mátrixok közti összefüggés – egy diagrammon is szemléltethető. A diagramm csúcsaiban az $\mathbf{x}$ és az $L\mathbf{x}$ vektorok szerepelnek. A függőleges nyilak az $\mathcal{A}$ bázisról a $\mathcal{B}$-re való áttérés irányát mutatják. Ebbe az irányba lépve az eredmény a $\mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}$ mátrixszal való szorzással kapható meg. A vízszintes nyilak az L leképezés hatását mutatják. Ezirányba lépve az $\mathbf{L}_\mathcal{A}$, illetve $\mathbf{L}_\mathcal{B}$ mátrixszal való szorzás adja meg az eredményt. A bal alsó sarokban lévő $[\mathbf{x}]_\mathcal{A}$ vektorból a jobb felsőben lévő $[L\mathbf{x}]_\mathcal{B}$ vektorba kétféleképp juthatunk: vagy először hat az L leképezés, aztán áttérünk a $\mathcal{B}$ bázisra, vagy előbb áttérünk a $\mathcal{B}$ bázisra, és azután hat L. Tehát

$$\begin{aligned} [\mathbf{x}]_\mathcal{A} &\longrightarrow \mathbf{L}_\mathcal{A}[\mathbf{x}]_\mathcal{A} \longrightarrow \mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}\mathbf{L}_\mathcal{A}[\mathbf{x}]_\mathcal{A}, \\ [\mathbf{x}]_\mathcal{A} &\longrightarrow \mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}[\mathbf{x}]_\mathcal{A} \longrightarrow \mathbf{L}_\mathcal{B}\mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}[\mathbf{x}]_\mathcal{A}. \end{aligned}$$

$$\begin{tikzcd} {[\mathbf{x}]_\mathcal{B}} \arrow[r, "\mathbf{L}_\mathcal{B}"] & {[L\mathbf{x}]_\mathcal{B}} \\ {[\mathbf{x}]_\mathcal{A}} \arrow[u, "\mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}"] \arrow[r, "\mathbf{L}_\mathcal{A}"'] & {[L\mathbf{x}]_\mathcal{A}} \arrow[u, "\mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}"'] \end{tikzcd}$$

7.11. ábra: A vízszintes nyilak az L leképezés hatását mutatják. E hatás elérhető az $\mathbf{L}_\mathcal{A}$, illetve az $\mathbf{L}_\mathcal{B}$ mátrixszal való szorzással. A függőleges nyilak a báziscsere irányát mutatják. A $\mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}$ mátrixszal való szorzással megvalósítható. Az ábráról leolvasható a $\mathbf{L}_\mathcal{B}\mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}} = \mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}\mathbf{L}_\mathcal{A}$ összefüggés.

$$\begin{tikzcd} {[\mathbf{x}]_\mathcal{B}} \arrow[r, "\mathbf{L}_\mathcal{B}"] & {[L\mathbf{x}]_\mathcal{B}} \arrow[d, "\mathbf{C}_{\mathcal{A}\leftarrow\mathcal{B}} = \mathbf{C}^{-1}_{\mathcal{B}\leftarrow\mathcal{A}}"] \\ {[\mathbf{x}]_\mathcal{A}} \arrow[u, "\mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}"] \arrow[r, "\mathbf{L}_\mathcal{A}"'] & {[L\mathbf{x}]_\mathcal{A}} \end{tikzcd}$$

7.12. ábra: Ezt az ábrát az előzőből egyetlen nyil és felirata megváltoztatásával kaptuk. Erről közvetlenül leolvasható az $\mathbf{L}_\mathcal{A} = \mathbf{C}_{\mathcal{A}\leftarrow\mathcal{B}}\mathbf{L}_\mathcal{B}\mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}$, illetve az $\mathbf{L}_\mathcal{A} = \mathbf{C}^{-1}_{\mathcal{B}\leftarrow\mathcal{A}}\mathbf{L}_\mathcal{B}\mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}$ összefüggés. Ehhez az $[\mathbf{x}]_\mathcal{A}$-ból az $[L\mathbf{x}]_\mathcal{A}$-ba vezető két utat kell bejárnunk, és közben a megfelelő mátrixokat összeszoroznunk.

A két végeredménynek meg kell egyeznie. Így ugyanazt kaptuk, amit behelyettesítésekkel. Összefoglalva tehát bizonyítottuk az alábbi tételt:

7.13. TÉTEL (LINEÁRIS LEKÉPEZÉS MÁTRIXAI KÖZTI KAPCSOLAT). *Legyen az L lineáris leképezés mátrixa az $\mathcal{A}$ bázisban $\mathbf{L}_{\mathcal{A}}$, a $\mathcal{B}$ bázisban $\mathbf{L}_{\mathcal{B}}$, és az $\mathcal{A}$ bázisról a $\mathcal{B}$-re való áttérés mátrixa $\mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}$. Ekkor*

$$\mathbf{L}_{\mathcal{A}} = \mathbf{C}_{\mathcal{A}\leftarrow\mathcal{B}}\mathbf{L}_{\mathcal{B}}\mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}, \text{ azaz } \mathbf{L}_{\mathcal{A}} = \mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}^{-1}\mathbf{L}_{\mathcal{B}}\mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}.$$

7.14. PÉLDA (LINEÁRIS LEKÉPEZÉS MÁTRIXA MÁSIK BÁZISBAN). *Az L lineáris leképezés mátrixa $\mathbf{L} = \left[\begin{smallmatrix}-1 & 6\\ -2 & 6\end{smallmatrix}\right]$. Írjuk fel mátrixát az $\mathcal{B} = \{(-2,-1),(3,2)\}$ bázisban!*

MEGOLDÁS. A megadott $\mathcal{B}$ bázisról a standard $\mathcal{E}$ bázisra való áttérés mátrixa a bázisvektorokból, mint oszlopvektorokból áll, azaz

$$\mathbf{C}_{\mathcal{E}\leftarrow\mathcal{B}} = \begin{bmatrix}-2 & 3\\ -1 & 2\end{bmatrix}.$$

Ekkor az $\mathbf{L} = \mathbf{L}_{\mathcal{E}}$ jelölést használva

$$\mathbf{L}_{\mathcal{B}} = \mathbf{C}_{\mathcal{E}\leftarrow\mathcal{B}}^{-1}\mathbf{L}_{\mathcal{E}}\mathbf{C}_{\mathcal{E}\leftarrow\mathcal{B}} = \begin{bmatrix}-2 & 3\\ -1 & 2\end{bmatrix}^{-1}\begin{bmatrix}-1 & 6\\ -2 & 6\end{bmatrix}\begin{bmatrix}-2 & 3\\ -1 & 2\end{bmatrix} = \begin{bmatrix}2 & 0\\ 0 & 3\end{bmatrix}.$$ □

Hasonlóság Tudjuk, hogy ha egy L lineáris transzformációnak egy $\mathcal{A}$ bázisban $\mathbf{A}$ a mátrixa, egy $\mathcal{B}$ bázisban $\mathbf{B}$, akkor a két mátrix kapcsolata

$$\mathbf{B} = \mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}}\mathbf{A}\mathbf{C}_{\mathcal{A}\leftarrow\mathcal{B}},$$

ahol $\mathbf{C}_{\mathcal{B}\leftarrow\mathcal{A}} = \mathbf{C}_{\mathcal{A}\leftarrow\mathcal{B}}^{-1}$ az áttérés mátrixa. E tény motiválja a következő definíciót:

7.15. DEFINÍCIÓ (HASONLÓSÁG). *Azt mondjuk, hogy az $n \times n$-es $\mathbf{A}$ mátrix* hasonló *a $\mathbf{B}$ mátrixhoz, ha létezik olyan invertálható $\mathbf{C}$ mátrix, hogy*

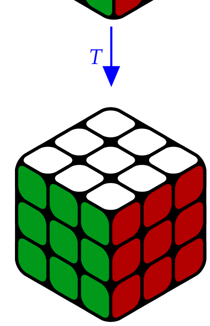
$$\mathbf{B} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}. \tag{7.3}$$

Jelölés: $\mathbf{A} \sim \mathbf{B}$.

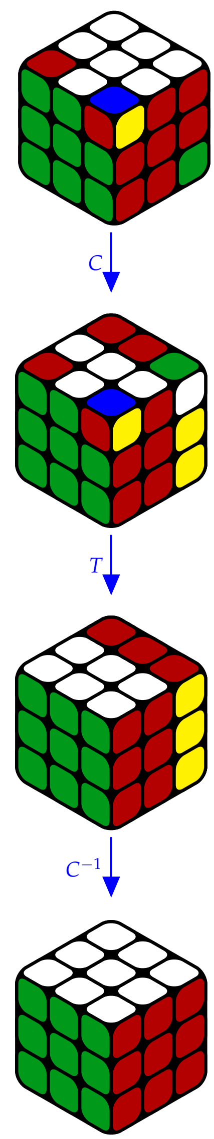
► Ha $\mathbf{A}$ hasonló $\mathbf{B}$-hez, akkor $\mathbf{B}$ is hasonló $\mathbf{A}$-hoz. Legyen ugyanis $\hat{\mathbf{C}} = \mathbf{C}^{-1}$. Akkor

$$\mathbf{A} = \mathbf{C}\mathbf{B}\mathbf{C}^{-1} = (\mathbf{C}^{-1})^{-1}\mathbf{B}\mathbf{C}^{-1} = \hat{\mathbf{C}}^{-1}\mathbf{B}\hat{\mathbf{C}}.$$

Így tehát mondható az, hogy $\mathbf{A}$ és $\mathbf{B}$ hasonlóak, mivel a hasonlóság szimmetrikus reláció.



7.13. ábra: A T transzformáció a bűvös kocka egy lapon lévő 3 csúcsát ciklikusan fölcseréli egymással (a jobb felsőt sarokban lévőt a bal sarokba, azt a középsőbe viszi), az összes többit helyben hagyja. Mit tegyünk ha 3 nem egy lapon lévő csúcsot kell ciklikusan fölcserélni?



7.14. ábra: A fölső kockán látható három csúcskockát kell kicserélni. Először a három csúcsot egy síkba mozgatjuk (C transzformáció), majd az előző ábrán látható T transzformációval kicseréljük őket, végül C inverzének alkalmazása után minden kocka a helyére kerül. Így a transzformációk egymás utáni elvégzését szorzásnak tekintve a megol-

▶ Például $\left[\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right] \sim \left[\begin{smallmatrix}-6&4\\-9&6\end{smallmatrix}\right]$, ugyanis

$$\begin{bmatrix}-6&4\\-9&6\end{bmatrix}=\begin{bmatrix}-2&-1\\-3&-2\end{bmatrix}\begin{bmatrix}0&1\\0&0\end{bmatrix}\begin{bmatrix}-2&1\\3&-2\end{bmatrix},\qquad \left(\mathbf{C}=\begin{bmatrix}-2&1\\3&-2\end{bmatrix}\right).$$

▶ A $\mathbf{B}=\mathbf{C}^{-1}\mathbf{A}\mathbf{C}$ összefüggés ekvivalens az $\mathbf{C}\mathbf{B}=\mathbf{A}\mathbf{C}$ összefüggéssel, amit még egyszerűbb lehet ellenőrizni. Példánk esetében

$$\begin{bmatrix}-2&1\\3&-2\end{bmatrix}\begin{bmatrix}-6&4\\-9&6\end{bmatrix}=\begin{bmatrix}0&1\\0&0\end{bmatrix}\begin{bmatrix}-2&1\\3&-2\end{bmatrix}\quad\left(=\begin{bmatrix}3&-2\\0&0\end{bmatrix}\right).$$

▶ A $\mathbf{C}^{-1}\mathbf{A}\mathbf{C}$ alakú kifejezést az $\mathbf{A}$ mátrix $\mathbf{C}$-vel való *konjugáltjának* nevezik. A konjugált más algebrai struktúrákban is fontos szerepet kap. Példaként véges halmazok permutációinak struktúráját említjük. Ekkor a permutációk közti művelet a permutációk egymás után való elvégzése, aminek eredményeként megintcsak a halmaz egy permutációját kapjuk. Konkrétan a Rubik kockán mutatjuk meg a konjugált szerepét a 7.13 és a 7.14 ábrák segítségével.

7.16. TÉTEL (HASONLÓ MÁTRIXOK HATÁSA). *Két mátrix pontosan akkor hasonló, ha van két olyan bázis, melyekben e két mátrix ugyanannak a lineáris leképezésnek a mátrixa.*

BIZONYÍTÁS. Ha $\mathbf{A}$ és $\mathbf{B}$ hasonlóak, azaz $\mathbf{B}=\mathbf{C}^{-1}\mathbf{A}\mathbf{C}$, akkor $\mathbf{C}$-t, mint a $\mathcal{C}=\{\mathbf{c}_1,\mathbf{c}_2,\dots,\mathbf{c}_n\}$ bázisról az $\mathcal{E}$ standard bázisra való áttérés mátrixát tekintve azt kapjuk, hogy

$$\mathbf{B}=\mathbf{C}_{\mathcal{E}\leftarrow\mathcal{C}}^{-1}\mathbf{A}\mathbf{C}_{\mathcal{E}\leftarrow\mathcal{C}}.$$

Eszerint ha L az $\mathbf{A}$ mátrixhoz tartozó mátrixleképezés, azaz $\mathbf{A}$ az L mátrixa a standard bázisban, akkor $\mathbf{B}$ az L mátrixa a $\mathcal{C}$ bázisban. A fordított állítást a bevezetőben igazoltuk. □

7.17. TÉTEL (HASONLÓSÁGRA INVARIÁNS TULAJDONSÁGOK). *Ha* $\mathbf{A}$ *és* $\mathbf{B}$ *hasonló mátrixok, azaz* $\mathbf{A}\sim\mathbf{B}$*, akkor*

a) $\mathrm{r}(\mathbf{A})=\mathrm{r}(\mathbf{B})$,

b) $\dim(\mathcal{N}(\mathbf{A}))=\dim(\mathcal{N}(\mathbf{B}))$,

c) $\det(\mathbf{A})=\det(\mathbf{B})$.

A latin eredetű *invariáns* szó jelentése: *átalakulás közben változatlanul maradó*. Matematikában valamilyen művelet, átalakítás, leképezés során változatlanul maradó kifejezés, mennyiség, érték. Esetünkben a bázis megváltoztatása után is változatlanul maradó mennyiségeket jelenti.

BIZONYÍTÁS. A bizonyítás során föltesszük, hogy valamely invertálható $\mathbf{C}$ mátrixszal $\mathbf{A}=\mathbf{C}^{-1}\mathbf{B}\mathbf{C}$.

a) Mivel mátrix rangja megegyezik az oszloptér dimenziójával, az oszloptér pedig megegyezik a mátrixleképezés képterével, ami a két mátrixra azonos, ezért a rangok is megegyeznek.

b) Hasonlóan bizonyítható az is, hogy a két mátrix nullterének dimenziója megegyezik, hisz a nulltér megegyezik a lineáris leképezés magterével, ami pedig közös.

c) $\det(\mathbf{A}) = \det(\mathbf{C}^{-1}\mathbf{B}\mathbf{C}) = \det(\mathbf{C}^{-1})\det(\mathbf{B})\det(\mathbf{C}) = \det(\mathbf{B})$, mivel $\det(\mathbf{C})\det(\mathbf{C}^{-1}) = 1$. □

Egy fontos következménye e tételnek, hogy a rang és a determináns fogalma természetes módon átvihető lineáris leképezésekre, legalábbis $\mathbb{R}^n \to \mathbb{R}^m$ leképezések esetén. A lineáris $L : \mathbb{R}^n \to \mathbb{R}^m$ leképezés *rangján* képterének dimenzióját értjük, azaz $\mathrm{r}(L) = \dim(\mathrm{Im}(L))$. A lineáris $L : \mathbb{R}^n \to \mathbb{R}^n$ transzformáció *determinánsán* az L leképezés bármely bázisban fölírt mátrixának determinánsát értjük. A definíció értelmes, hisz ez az érték a bázis választásától független.

Tartományok képe és mértékük változása Ha különböző alakzatok lineáris leképezések általi képét vizsgáljuk, felvetődik a kérdés: hogyan változik a területük/térfogatuk a leképezések során? Pl. az $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ mátrixszal reprezentált lineáris leképezés egy T területű síkidomot mekkora területűbe visz?

Nem nyilvánvaló, hogy a kérdésre van-e egyáltalán válasz, és nem fordulhat-e elő, hogy a fenti leképezés egyik síkidomot például a 2-szeresébe, másikat a 3-szorosába viszi.

Akár a területmérték Jordan-, akár Lebesgue-féle definícióját tekintjük, elég a kérdést olyan téglalapokra megválaszolni, amelyek oldalai párhuzamosak a tengelyekkel, de ezt az állítást nem igazoljuk (lásd a széljegyzetet). Legyen egy ilyen téglalap négy csúcsa: (p, q), $(p + x, q)$, $(p, q + y)$, $(p + x, q + y)$, ahol $x, y > 0$. Tehát a téglalap oldalhossza x és y, területe xy. A csúcsok képe kiszámolható egyetlen mátrixszorzással:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} p & p+x & p & p+x \\ q & q & q+y & q+y \end{bmatrix} = \begin{bmatrix} ap+bq & ap+ax+bq & ap+bq+by & ap+ax+bq+by \\ cp+dq & cp+cx+dq & cp+dq+dy & cp+cx+dq+dy \end{bmatrix}$$

Innen leolvasható, hogy a téglalap képeként kapott parallelogramma oldalvektorai (ax, cx) és (by, dy), és így területe

$$|(ax)(dy) - (cx)(by)| = |ad - bc|xy.$$

Eszerint tehát a téglalap képének területe független a téglalap helyzetétől, és mindig a téglalap területének $|ad - bc|$-szerese, s így minden területmértékkel rendelkező síkbeli tartomány képének területe az eredeti $|ad - bc|$-szerese.

Gondoljuk meg, mi a különbség T két képe közt az $ad - bc > 0$, és az $ad - bc < 0$ esetben?

Az imént vázolt gondolatmenettel bizonyítható a következő állítás, mely a determináns egy fontos tulajdonságát írja le:

Az $\mathbb{R}^2$ TARTOMÁNYAI TERÜLETÉNEK definiálására a matematika több megközelítést is ismer. Kettőt vázolunk. Mindkettő közös kiindulási pontja, hogy az $[a, b] \times [c, d]$ téglalap területe $(b - a)(d - c)$, és mindkét esetben csak a tengelyekkel párhuzamos helyzetű téglalapokat használunk.

Egy síkbeli T halmazhoz vegyünk véges sok közös belső pont nélküli téglalapot, melyek egyesítése lefedi T-t. E fedő téglalaprendszer területe legyen a téglalapok területeinek összege. Azt mondjuk, hogy T külső Jordan-mértéke a tartományt lefedő téglalaprendszerek területeinek infimuma. Hasonlóan, a T-be írt közös belső pont nélküli téglalaprendszerek területeinek szuprémumát T belső Jordan-mértékének nevezzük. T *Jordan-mérhető*, ha belső és külső Jordan-mértéke megegyezik, és e közös érték T *Jordan-mértéke*.

A LEBESGUE-MÉRTÉK definiálásához olyan fedő téglalaprendszereket veszünk, melyek megszámlálható (tehát akár végtelen) elemszámúak, és ezekhez a téglalapjaik területösszegét rendeljük. A T halmazt fedő összes téglalaprendszerhez ilyen módon rendelt számok infimumát a T külső mértékének nevezzük és $\lambda^*(T)$-vel jelöljük. A T-t *Lebesgue-mérhetőnek nevezzük*, ha bármely síkbeli H halmazra

$$\lambda^*(H) = \lambda^*(H \cap T) + \lambda^*(H \setminus T).$$

Egy Jordan-mérhető T tartomány Lebesgue-mérhető is, és e két mérték ekkor megegyezik. Viszont például az egységnégyzet racionális koordinátájú pontjainak halmaza nem Jordan-mérhető, mert belső mértéke 0, külső mértéke 1, viszont Lebesgue-mérhető, mértéke 0.

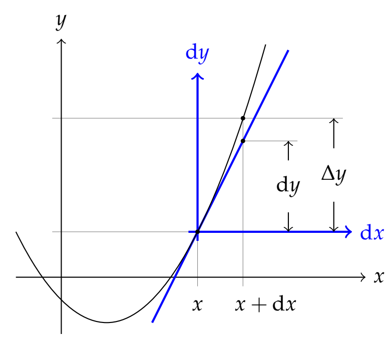
7.18. TÉTEL (TARTOMÁNY MÉRTÉKÉNEK VÁLTOZÁSA LINEÁRIS TRANSZFORMÁCIÓBAN). *Ha $T \subseteq \mathbb{R}^n$ egy tartomány, melynek $m(T)$ a mértéke, és $L : \mathbb{R}^n \to \mathbb{R}^n$ egy lineáris leképezés, akkor az $L(T)$ tartománynak van mértéke, és az* $m(T)\det(L)$.

*Többváltozós függvények differenciálása** Az $\mathbb{R}^n$-ből $\mathbb{R}^m$-be képző lineáris leképezések egy igen fontos alkalmazása a vektor-vektor függvények differenciálhatóságának fogalma.

A differenciálhatóság szokásos definíciója a következő: azt mondjuk, hogy az $f : \mathbb{R} \to \mathbb{R}$ függvény *differenciálható* az x helyen, ha létezik a

$$D = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \tag{7.4}$$

határérték. A D számnak fontos jelentése van: az f függvény x körüli megváltozása jól közelíthető a $\mathrm{d}x \mapsto D\mathrm{d}x$ függvény 0 körüli megváltozásával. Szemléltetve ez azt jelenti, hogy ha az f grafikonján az $(x, f(x))$ pontra helyezünk egy $\mathrm{d}x$ és $\mathrm{d}y$ változójú koordinátarendszert, akkor a $\mathrm{d}x \mapsto \mathrm{d}y = D\mathrm{d}x$ grafikonja az f függvény grafikonjának érintője (ld. a 7.15 ábrát). Eszerint, kicsit leegyszerűsítve a megfogalmazást, a differenciálhatóság azt jelenti, hogy a függvény „jól közelíthető” egy $\mathbb{R} \to \mathbb{R}$ lineáris leképezéssel, hisz a $\mathrm{d}x \mapsto D\mathrm{d}x$ leképezés ilyen (ld. a 7.10. tétel utáni megjegyzést az $\mathbb{R} \to \mathbb{R}$ lineáris transzformációkról).



7.15. ábra: A $\mathrm{d}x$ és $\mathrm{d}y$ koordinátatengelyeket és a $\mathrm{d}y = D\mathrm{d}x$ függvény grafikonját színezéssel kiemeltük. Az ábra egyúttal a $\Delta y \approx \mathrm{d}y$ kapcsolatot is szemlélteti.

A „jól közelítés” szemléletesen azt jelenti, hogy az f grafikonjára „zoomolva”, azaz azt folyamatosan nagyítva a grafikon kiegyenesedni látszik. Ez az az egyenes, melyet a grafikon érintőjének nevezünk, és amelynek $\mathrm{d}y = D\mathrm{d}x$ az egyenlete az új koordinátarendszerben.

Ez a definíció ekvivalens módon átfogalmazható: azt mondjuk, hogy az $f : \mathbb{R} \to \mathbb{R}$ függvény *differenciálható* az x helyen, ha van olyan D szám, hogy

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - Dh}{h} = 0. \tag{7.5}$$

Ez utóbbi alak azzal az előnnyel is jár, hogy könnyen általánosítható. Az általánosítás legfőbb nehézsége az, hogy a vektorral való osztás nem definiálható megfelelően, ezért e formulán még egy apró, de még mindig ekvivalens változtatást teszünk: nem h-val, hanem annak abszolút értékével osztunk.

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - Dh}{|h|} = 0.$$

Mindezek a következő definícióra vezetnek:

7.19. DEFINÍCIÓ (DIFFERENCIÁLHATÓSÁG). *Azt mondjuk, hogy az* $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ *függvény differenciálható az* $\mathbf{x}$ *helyen, ha létezik olyan* $D_{\mathbf{f},\mathbf{x}} : \mathbb{R}^n \to \mathbb{R}^m$ *lineáris leképezés, melyre*

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) - D_{\mathbf{f},\mathbf{x}}\mathbf{h}}{|\mathbf{h}|} = \mathbf{0}.$$

A $D_{\mathbf{f},\mathbf{x}}$ *leképezést az* $\mathbf{f}$ *függvény* $\mathbf{x}$ *ponthoz tartozó* deriváltleképezésének *nevezzük.*

► A $D_{\mathbf{f},\mathbf{x}}$ jelölés arra utal, hogy a deriváltleképezés az $\mathbf{f}$ függvénytől és az $\mathbf{x}$ helytől is függ, maga viszont mint leképezés egy $\mathbf{h}$ vektorhoz a $D_{\mathbf{f},\mathbf{x}}\mathbf{h}$ vektort rendeli.
► Elterjedtebb a $D_{\mathbf{x}}(\mathbf{f})$ jelölés, itt didaktikai okból választottunk olyat, mely jobban világossá teszi, hogy ez egy lineáris leképezés, mely majd hat valamely $\mathbf{h}$ vektoron, és annak képe $D_{\mathbf{x}}(\mathbf{f})\mathbf{h}$ vagy $D_{\mathbf{x}}(\mathbf{f})(\mathbf{h})$ – az általunk használt jelölésben $D_{\mathbf{f},\mathbf{x}}\mathbf{h}$.

7.20. TÉTEL (JACOBI-MÁTRIX). *Ha az* $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m; (x_1, x_2, \dots, x_n) \mapsto (f_1, f_2, \dots, f_m)$ *függvény differenciálható az* $\mathbf{x}$ *helyen, akkor a lineáris* $D_{\mathbf{f},\mathbf{x}}$ *deriváltleképezés mátrixa a következő ún.* Jacobi-mátrix*:*

$$\mathbf{D}_{\mathbf{f},\mathbf{x}} = \frac{\partial(f_1, f_2, \dots, f_m)}{\partial(x_1, x_2, \dots, x_n)}(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}) & \dots & \frac{\partial f_2}{\partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}) & \frac{\partial f_m}{\partial x_2}(\mathbf{x}) & \dots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}) \end{bmatrix}$$

BIZONYÍTÁS. Ha $\mathbf{f}$ differenciálható, akkor a definícióbeli határérték akkor is fönnáll, ha $\mathbf{h}$ speciális módon tart a nullvektorhoz, például ha $\mathbf{h} = t\mathbf{e}_j$, és $t \to 0$. Ekkor

$$\lim_{t\to 0} \frac{\mathbf{f}(\mathbf{x}+t\mathbf{e}_j) - \mathbf{f}(\mathbf{x}) - D_{\mathbf{f},\mathbf{x}}(t\mathbf{e}_j)}{|t|} = \mathbf{0}.$$

Az $\mathbf{f}$ függvény i-edik koordinátafügvénye f_i, a $D_{\mathbf{f},\mathbf{x}}(t\mathbf{e}_j)$ vektor i-edik koordinátája $\mathbf{e}_i' D_{\mathbf{f},\mathbf{x}}(t\mathbf{e}_j)$. Ennek alapján

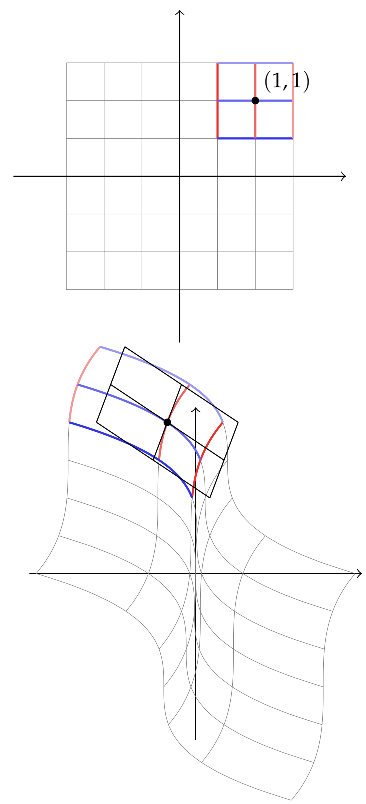
$$\lim_{t\to 0} \frac{f_i(\mathbf{x}+t\mathbf{e}_j) - f_i(\mathbf{x}) - \mathbf{e}_i' D_{\mathbf{f},\mathbf{x}}(t\mathbf{e}_j)}{|t|} = 0.$$

Ez a határérték viszont már egy egyváltozós függvény deriváltja, ami nem más, mint az f_i függvény j-edik parciális deriváltja, ugyanis átrendezve az egyenlőséget és t előjelével is osztva kapjuk, hogy

$$\lim_{t\to 0} \frac{f_i(\mathbf{x}+t\mathbf{e}_j) - f_i(\mathbf{x})}{t} = \mathbf{e}_i' D_{\mathbf{f},\mathbf{x}}\mathbf{e}_j, \text{ azaz } \mathbf{e}_i' D_{\mathbf{f},\mathbf{x}}\mathbf{e}_j = \frac{\partial f_i}{\partial x_j}(\mathbf{x}).$$

Ez bizonyítja állításunkat. □

► A gyakorlatban az $\mathbb{R}^n \to \mathbb{R}$ függvények, vagyis az n-változós skalárértékű függvények esetén az egyetlen sorból álló Jacobi-mátrix helyett annak vektoralakját használják, melyet *gradiensvektornak* neveznek, és ∇f-fel jelölnek.
► Hasonlóképp, mivel az $\mathbb{R} \to \mathbb{R}^n$ függvények Jacobi-mátrixa egyetlen oszlopból áll, gyakran használják annak vektoralakját. Ha például egy $\mathbf{r} : \mathbb{R} \to \mathbb{R}^3; t \mapsto \mathbf{r}(t)$ függvény a térben mozgó tárgy mozgását az idő függvényében írja le, e vektor épp a mozgás sebességvektora.
► Ha az $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ függvény differenciálható, a deriváltleképezés lineáris transzformáció, melynek van determinánsa. A $\det(D_{\mathbf{f},\mathbf{x}})$ determináns egyik fontos – és a többváltozós integrálok helyettesítésében kihasznált – jelentése az, hogy az $\mathbf{x}$ hely kis környezetében egy „kicsiny" mértékű T tartomány képének mértéke közelítőleg $\det(D_{\mathbf{f},\mathbf{x}})$-szerese a T mértékének. A $\det(D_{\mathbf{f},\mathbf{x}})$ determinánst, illetve mátrixának, a Jacobi-mátrixnak determinánsát *Jacobi-determinánsnak* nevezzük.



7.16. ábra: A felső ábra az $\mathbf{f}(x,y) = (-x^3/2 + y^3/8, x+y)$ függvény értelmezési tartományán megadott rácsot, és annak egy kis 2×2-es részét mutatja, melynek középpontja az $(1,1)$ pont. Az alsó ábra egyrészt halványan jelöli e rács és színesen a kiemelt rács képét, valamint az $(1,1)$ ponthoz tartozó deriváltleképezés hatását e kiemelt rácson.

7.21. PÉLDA (JACOBI-MÁTRIX KISZÁMÍTÁSA). *Határozzuk meg az alábbi függvények egy általános ponthoz és a megadott ponthoz tartozó Jacobi-mátrixát, és ahol lehet, a Jacobi-determinánsát!*
a) $f(x,y) = x^2y - xy^3 + 1$, $(x,y) = (0,1)$.
b) $\mathbf{f}(x,y) = (-x^3/2 + y^3/8, x+y)$, $(x,y) = (1,1)$.
c) $\mathbf{r}(t) = (t^3, t^2, t)$, $t = 2$.
d) $\mathbf{f}(x_1,x_2,x_3) = (2x_1 + 3x_2, x_1 - x_2 - x_3)$, $(x_1,x_2,x_3) = (1,2,0)$.

MEGOLDÁS. *a)* $f(x,y) = x^2y - xy^3$, parciális deriváltjai $\frac{\partial}{\partial x}f(x,y) = 2xy - y^3$, $\frac{\partial}{\partial y}f(x,y) = x^2 - 3xy^2$. A deriváltleképezés mátrixa, azaz a Jacobi-mátrix itt

$$\begin{bmatrix} 2xy - y^3 & x^2 - 3xy^2 \end{bmatrix}$$

E mátrix vektor alakja, azaz a gradiensvektor

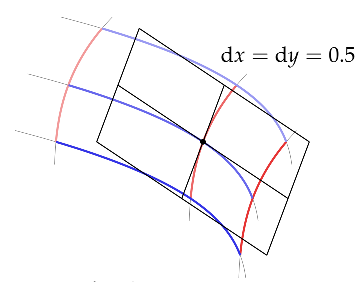
$$\nabla f(x,y) = (2xy - y^3, x^2 - 3xy^2).$$

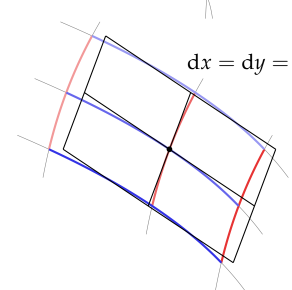
Ennek értéke a $(0,1)$ helyen $\nabla f(0,1) = (-1,0)$, illetve a Jacobi-mátrix e helyen $[-1\ 0]$. E mátrix nem négyzetes, így a Jacobi-determináns nincs értelmezve.

b) Az $\mathbf{f}(x,y) = (-x^3/2 + y^3/8, x+y)$ függvény Jacobi-mátrixa és annak értéke a megadott $(x,y) = (1,1)$ pontban

$$\begin{bmatrix} -\frac{3}{2}x^2 & \frac{3}{8}y^2 \\ 1 & 1 \end{bmatrix}, \text{ illetve } \begin{bmatrix} -\frac{3}{2} & \frac{3}{8} \\ 1 & 1 \end{bmatrix}.$$

Például az első sor első eleme $\frac{\partial}{\partial x}(-x^3/2 + y^3/8) = -\frac{3}{2}x^2$. Az $\mathbf{f}$ függvény deriváltleképezésének, vagyis Jacobi-mátrixának hatását szemlélteti a 7.16 és a 7.17 ábra. Az (x,y) ponthoz tartozó Jacobi-determináns értéke $-\frac{3}{2}x^2 - \frac{3}{8}y^2$, ami az origót kivéve mindenütt negatív. Ez azt





jelenti, hogy e leképezés minden pont környezetében megfordítja a körüljárást. Az $(1,1)$ pontban a Jacobi-determináns értéke $-\frac{15}{8}$. Tehát e pont kis környezetében tartomány területét a függvény közel megkétszerezi.

c) Az $\mathbf{r}(t) = (t^3, t^2, t)$ függvény Jacobi-mátrixa

$$\begin{bmatrix} 3t^2 \\ 2t \\ 1 \end{bmatrix}, \text{ ami a } t = 2 \text{ helyen } \begin{bmatrix} 12 \\ 4 \\ 1 \end{bmatrix}.$$

A térben mozgó pont (test) mozgsának leírására is $\mathbb{R} \to \mathbb{R}^3$ függvényt használunk. Ha e függvény egy ilyen mozgást ír le, akkor sebességvektora egy tetszőleges pontban

$$\dot{\mathbf{r}}(t) = (3t^2, 2t, 1),$$

a $t = 2$ paraméterhez tartozó pontban $\dot{\mathbf{r}}(2) = (12, 4, 1)$.

d) Az utolsó példa fontos állítást szemléltet, nevezetesen azt, hogy egy lineáris leképezés deriváltja minden $\mathbf{x}$ helyen megegyezik magával a leképezéssel, azaz a deriváltja önmaga. Világos, hogy a megadott leképezés egy lineáris leképezés, mátrixszorzatos alakja, nevezetesen

$$\mathbf{f}(x_1, x_2, x_3) = \begin{bmatrix} 2 & 3 & 0 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Ennek Jacobi-mátrixa valóban bármely (x_1, x_2, x_3) helyen

$$\begin{bmatrix} 2 & 3 & 0 \\ 1 & -1 & -1 \end{bmatrix},$$

ugyanis az i-edik koordinátafüggvény j-edik parciális deriváltja épp az együtthatómátrix i-edik sor-, j-edik oszlopbeli eleme, azaz egy konstans. Így minden helyen e mátrix lesz a Jacobi-mátrix, speciálisan az $(x_1, x_2, x_3) = (1, 2, 0)$ helyen is. □

7.22. PÉLDA (FÜGGVÉNYÉRTÉK BECSLÉSE JACOBI-MÁTRIXSZAL). *Ismerjük egy differenciálható függvény értelmezési tartományának egy pontjához tartozó Jacobi-mátrixát és a függvényértéket ugyan ebben a pontban! Becsüljük meg a függvény értékét egy e ponthoz közeli helyen az alábbi adatok ismeretében!*

a) $f(0,1) = 1$, $\mathbf{D}_{f,(0,1)} = [-1 \;\; 0]$, $(x, y) = (-0.05, 1.1)$,

b) $\mathbf{f}(1,1) = (-\frac{3}{8}, 2)$, $\mathbf{D}_{\mathbf{f},(1,1)} = \left[\begin{smallmatrix} -3/2 & 3/8 \\ 1 & 1 \end{smallmatrix}\right]$, $(x, y) = (0.8, 1.1)$.

Mennyire lennének jók e becslések, ha a függvények az előző feladatbeli a) és b) függvényei lennének?

MEGOLDÁS. A függvény megváltozásának becsléséhez az $\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x})$ értéket kell megbecsülni. A differenciálhatóság definíciója szerint erre a $\mathbf{D}_{\mathbf{f},\mathbf{x}}\mathbf{h}$ mennyiség alkalmas, ha a függvény differenciálható az $\mathbf{x}$ pontban. Eszerint tehát

$$\mathbf{f}(\mathbf{x}+\mathbf{h}) \approx \mathbf{f}(\mathbf{x}) + \mathbf{D}_{\mathbf{f},\mathbf{x}}\mathbf{h}. \tag{7.6}$$

E képletet felhasználva az alábbi megoldásokra jutunk:

a) E feladatban $\mathbf{h} = (-0.05, 0.1)$, így a függvény megváltozása a

$$\mathbf{D}_{f,(0,1)}\mathbf{h} = \begin{bmatrix} -1 & 0 \end{bmatrix} \begin{bmatrix} -0.05 \\ 0.1 \end{bmatrix} = 0.05$$

értékkel becsülhető, tehát a függvény értéke

$$f(\mathbf{x}+\mathbf{h}) = f(-0.05, 1.1) \approx f(0,1) + \mathbf{D}_{f,(0,1)} \begin{bmatrix} -0.05 \\ 0.1 \end{bmatrix} = 1.05,$$

azaz $f(-0.05, 1.1) \approx 1.05$. Ha f az előző *a)* feladatbeli függvény, azaz $f(x,y) = x^2y - xy^3 + 1$, akkor a pontos érték $f(-0.05, 0.1) = 1.0693$.

b) Itt $\mathbf{h} = (-0.2, 0.1)$, így a függvény megváltozása a

$$\mathbf{D}_{\mathbf{f},(1,1)}\mathbf{h} = \begin{bmatrix} -\frac{3}{2} & \frac{3}{8} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -0.2 \\ 0.1 \end{bmatrix} = \begin{bmatrix} \frac{3}{2}\cdot\frac{2}{10} + \frac{3}{8}\cdot\frac{1}{10} \\ -\frac{2}{10} + \frac{1}{10} \end{bmatrix} = \begin{bmatrix} 0.3375 \\ -0.1 \end{bmatrix}$$

értékkel becsülhető, tehát a függvény értéke $\mathbf{f}(0.8, 1.1) \approx \mathbf{f}(1,1) + (0.3375, -0.1) = (-0.0375, 1.9)$. Ha $\mathbf{f}$ az előző *b)* feladatbeli függvény, azaz $\mathbf{f}(x,y) = (-x^3/2 + y^3/8, x+y)$, akkor a pontos érték $f(0.8, 1.1) = (-0.089625, 1.9)$. □

E paragrafusnak nem célja a függvényanalízis területére tartozó témák feldolgozása, de a többváltozós függvények kompozíciójának deriváltleképezése az egyváltozós függvények láncszabályához hasonló módon számolható, és erre érdemes egy pillantást vetnünk, mert a megoldást a deriváltleképezések kompozíciója, azaz a Jacobi-mátrixok szorzata adja. Bizonyítás nélkül közöljük a következő tételt.

7.23. tétel (Láncszabály). *Legyen $\mathbf{f} : \mathbb{R}^k \to \mathbb{R}^m$, $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^k$ két függvény. Ha $\mathbf{g}$ differenciálható az $\mathbf{x}$ helyen, és $\mathbf{f}$ az $\mathbf{u} = \mathbf{g}(\mathbf{x})$ helyen, akkor $\mathbf{f}\circ\mathbf{g}$ differenciálható az $\mathbf{x}$ helyen, és deriváltleképezése, illetve annak mátrixa kifejezhető az alábbi alakban:*

$$D_{\mathbf{f}\circ\mathbf{g},\mathbf{x}} = D_{\mathbf{f},\mathbf{g}(\mathbf{x})} \circ D_{\mathbf{g},\mathbf{x}}, \quad \textit{mátrixjelöléssel} \quad \mathbf{D}_{\mathbf{f}\circ\mathbf{g},\mathbf{x}} = \mathbf{D}_{\mathbf{f},\mathbf{g}(\mathbf{x})}\mathbf{D}_{\mathbf{g},\mathbf{x}}, \quad \textit{illetve}$$

$$\frac{\partial(f_1, f_2, \dots, f_m)}{\partial(x_1, x_2, \dots, x_n)}(\mathbf{x}) = \frac{\partial(f_1, f_2, \dots, f_m)}{\partial(u_1, u_2, \dots, u_k)}(\mathbf{g}(\mathbf{x})) \frac{\partial(g_1, g_2, \dots, g_k)}{\partial(x_1, x_2, \dots, x_n)}(\mathbf{x})$$

7.24. példa (Láncszabály). *Írjuk fel a láncszabály általános képleteit a megadott függvénytípusokra, az összetett függvény deriváltját pedig a láncszabállyal és behelyettesítéssel is számítsuk ki!*

a) $f : (x, y) \mapsto x^2 - y$, $\mathbf{g} : u \mapsto (u^2 + u, u - 1)$, $u = 1$.
b) $\mathbf{f} : \mathbb{R} \to \mathbb{R}^2; x \mapsto (x^2, x - 1)$, $g : \mathbb{R}^2 \to \mathbb{R}; (u, v) \mapsto x = u^2 v$, $(u, v) = (1, 2)$.
c) $\mathbf{f}(x, y) = (xy^2 - 1, x - y)$, $\mathbf{g}(u, v) = (u + 1, u - v)$, $(u, v) = (0, 1)$.

Megoldás. Az *a)* esetben az f-hez, illetve $\mathbf{g}$-hez tartozó láncszabály általános alakja

$$\frac{\mathrm{d}f}{\mathrm{d}u} = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\mathrm{d}g_1}{\mathrm{d}u} \\ \frac{\mathrm{d}g_2}{\mathrm{d}u} \end{bmatrix} = \frac{\partial f}{\partial x}\frac{\mathrm{d}g_1}{\mathrm{d}u} + \frac{\partial f}{\partial y}\frac{\mathrm{d}g_2}{\mathrm{d}u},$$

a függvények parciális deriváltjait kiszámolva és a helyet megadva

$$\frac{\mathrm{d}f}{\mathrm{d}u}(1) = \begin{bmatrix} 2x & -1 \end{bmatrix}_{\mathbf{g}(1)=(2,0)} \begin{bmatrix} 2u + 1 \\ 1 \end{bmatrix}_{u=1},$$

végül a helyettesítést is elvégezve

$$\begin{bmatrix} 4 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 11.$$

Ugyanezt az eredményt kapjuk, ha a deriválás előtt elvégezzük a helyettesítést: $(f \circ \mathbf{g})(u) = (u^2 + u)^2 - (u - 1) = u^4 + 2u^3 + u^2 - u + 1$, ennek u szerinti deriváltja $4u^3 + 6u^2 + 2u - 1$, és ennek értéke az $u = 1$ helyen 11.

A *b)* esetben $\mathbf{f} : \mathbb{R} \to \mathbb{R}^2$, $g : \mathbb{R}^2 \to \mathbb{R}$, így $\mathbf{f} \circ g : \mathbb{R}^2 \to \mathbb{R}^2$, és

$$\begin{bmatrix} \frac{\partial f_1}{\partial u} & \frac{\partial f_1}{\partial v} \\ \frac{\partial f_2}{\partial u} & \frac{\partial f_2}{\partial v} \end{bmatrix} = \begin{bmatrix} \frac{\mathrm{d}f_1}{\mathrm{d}x} \\ \frac{\mathrm{d}f_2}{\mathrm{d}x} \end{bmatrix} \begin{bmatrix} \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v} \end{bmatrix} = \begin{bmatrix} \frac{\mathrm{d}f_1}{\mathrm{d}x}\frac{\partial g}{\partial u} & \frac{\mathrm{d}f_1}{\mathrm{d}x}\frac{\partial g}{\partial v} \\ \frac{\mathrm{d}f_2}{\mathrm{d}x}\frac{\partial g}{\partial u} & \frac{\mathrm{d}f_2}{\mathrm{d}x}\frac{\partial g}{\partial v} \end{bmatrix}$$

A megadott függvényekre és a helyettesítendő értékeket is megadva

$$\begin{bmatrix} 2x \\ 1 \end{bmatrix}_{x=g(1,2)=2} \begin{bmatrix} 2uv & u^2 \end{bmatrix}_{u=1,v=2} = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \begin{bmatrix} 4 & 1 \end{bmatrix} = \begin{bmatrix} 16 & 4 \\ 4 & 1 \end{bmatrix}$$

Behelyettesítés után a függvény $(u, v) \mapsto (u^4 v^2, u^2 v - 1)$, aminek deriváltja az $(u, v) = (1, 2)$ helyen

$$\begin{bmatrix} 4u^3 v^2 & 2u^4 v \\ 2uv & u^2 \end{bmatrix}_{(1,2)} = \begin{bmatrix} 16 & 4 \\ 4 & 1 \end{bmatrix},$$

ami természetesen megegyezik az előző eredménnyel.

Végül a *c)* esetben az általános alak

$$\begin{bmatrix} \frac{\partial f_1}{\partial u} & \frac{\partial f_1}{\partial v} \\ \frac{\partial f_2}{\partial u} & \frac{\partial f_2}{\partial v} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\partial g_1}{\partial u} & \frac{\partial g_1}{\partial v} \\ \frac{\partial g_2}{\partial u} & \frac{\partial g_2}{\partial v} \end{bmatrix}.$$

A parciális deriváltakat kiszámolva és a helyettesítési értékeket is megadva kapjuk, hogy

$$\begin{bmatrix} \frac{\partial f_1}{\partial u} & \frac{\partial f_1}{\partial v} \\ \frac{\partial f_2}{\partial u} & \frac{\partial f_2}{\partial v} \end{bmatrix}_{(0,1)} = \begin{bmatrix} y^2 & 2xy \\ 1 & -1 \end{bmatrix}_{(1,-1)} \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}_{(0,1)} = \begin{bmatrix} 1 & -2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ 0 & 1 \end{bmatrix}.$$

Itt fölhasználtuk, hogy $\mathbf{g}(0,1) = (-1,1)$. Ha a deriválás előtt elvégezzük a függvények kompozícióját, akkor ugyanerre az eredményre jutunk, ugyanis

$$(\mathbf{f}(\mathbf{g}(u,v)) = \Big((u+1)(u-v)^2 - 1, v+1)\Big),$$

aminek a deriváltmátrixa

$$\begin{bmatrix} (u-v)^2 + 2(u+1)(u-v) & -2(u+1)(u-v) \\ 0 & 1 \end{bmatrix}_{(0,1)} = \begin{bmatrix} -1 & 2 \\ 0 & 1 \end{bmatrix}. \qquad \square$$

Feladatok

Mátrixleképezések

Döntsük el, hogy az alábbi leképezések mátrixleképezések-e! Amelyik igen, annak írjuk fel a mátrixát! Legyen $\mathbf{a} = (a_1, a_2, a_3)$ egy tetszőleges vektor.

7.1. $A : \mathbf{x} \mapsto \mathbf{a} \cdot \mathbf{x}$,

7.2. $A : \mathbf{x} \mapsto \mathbf{a} + \mathbf{x}$,

7.3. $A : \mathbf{x} \mapsto \frac{1}{2}(\mathbf{a} \times \mathbf{x})$,

7.4. $A : \mathbf{x} \mapsto \mathbf{a}(\mathbf{a} \cdot \mathbf{x})$.

Lineáris leképezések

Döntsük el, hogy az alábbi leképezések lineáris leképezések-e!

7.5. $A : (x, y) \mapsto (x + 2y, x - y)$.

7.6. Legyen $\mathcal{P}_3$ a legföljebb 3-adfokú polinomok halmaza, és legyen $D : \mathcal{P}_3 \to \mathcal{P}_3 : p(x) \mapsto p'(x)$.

7.7. Legyen $\mathcal{D}_{[0,1]}$ a $[0, 1]$ intervallumon differenciálható függvények halmaza, és $\mathcal{F}_{[0,1]}$ a $[0, 1]$ intervallumon értelmezett függvények halmaza. Legyen továbbá $A : \mathcal{D}_{[0,1]} \to \mathcal{F}_{[0,1]}; f(x) \mapsto xf'(x)$.

Lineáris leképezés mátrixa

2- *és 3-dimenziós geometriai transzformációk mátrixa*

A 7.4. tétel bizonyításában megkonstruáltuk egy lineáris $A : \mathbb{R}^n \to \mathbb{R}^m$ leképezéshez azt a mátrixot, mely ezt a leképezést generálja, nevezetesen bizonyítottuk, hogy minden $\mathbf{x}$ vektorra

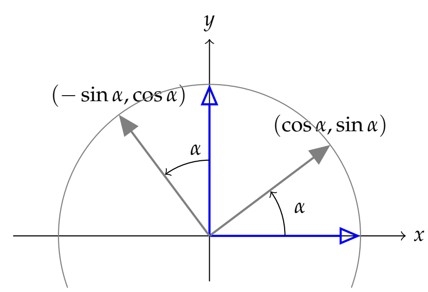
$$A\mathbf{x} = [A\mathbf{e}_1|A\mathbf{e}_2|\dots|A\mathbf{e}_n]\,\mathbf{x}.$$

Először vizsgáljunk meg néhány geometriailag leírható $\mathbb{R}^2 \to \mathbb{R}^2$ és $\mathbb{R}^3 \to \mathbb{R}^3$ leképezést.

Forgatás Vizsgáljuk meg a síkbeli pont körüli és a térbeli egyenes körüli forgatások mátrixát!

7.25. ÁLLÍTÁS (A FORGATÁS MÁTRIXA). *A sík vektorait egy pont körül α szöggel elforgató leképezés mátrixa*

$$\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}.$$



7.18. ábra: Az $\mathbf{i}$ és $\mathbf{j}$ vektorok α szöggel való elforgatottjai

BIZONYÍTÁS. A 7.9. állítás szerint a forgatás lineáris leképezés, így van mátrixa, melynek alakja $[A\mathbf{i}\ A\mathbf{j}]$, ahol $\mathbf{i}$ és $\mathbf{j}$ jelöli az $\mathbb{R}^2$ standard bázisának elemeit. E vektorokat szemlélteti a 7.18 ábra.

Az $A\mathbf{i}$ vektor megegyezik $\mathbf{i}$ elforgatottjával, amelynek ismerjük koordinátáit: $A\mathbf{i} = \left[\begin{smallmatrix}\cos\alpha\\ \sin\alpha\end{smallmatrix}\right]$. A $\mathbf{j}$ vektor α szöggel való elforgatottja megegyezik az $A\mathbf{i}$ vektor $\pi/2$ szöggel való elforgatottjával, azaz $A\mathbf{j} = \left[\begin{smallmatrix}-\sin\alpha\\ \cos\alpha\end{smallmatrix}\right]$. Így az A-hoz tartozó mátrix

$$\mathbf{A} = \begin{bmatrix} A\mathbf{i} & A\mathbf{j} \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

Tehát egy $\mathbf{x}$ vektor α szöggel való elforgatottja $\mathbf{A}\mathbf{x} = \left[\begin{smallmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{smallmatrix}\right]\mathbf{x}$. □

7.26. PÉLDA (FORGATÁS EGY TETSZŐLEGES PONT KÖRÜL). *Határozzuk meg a koordinátáit a $(4,3)$ pont $(2,1)$ körül $\pi/3$ radiánnal való elforgatásával kapott pontnak!*

MEGOLDÁS. A forgatás középpontját toljuk az origóba, így a $(4,3)$ pont a $(4,3)-(2,1)=(2,2)$ pontba kerül. E pontot, illetve az oda mutató helyvektort forgassuk el $\pi/3$ radiánnal, azaz $60°$-kal. Ez a forgatás mátrixával való beszorzással megkapható:

$$\begin{bmatrix} \cos\frac{\pi}{3} & -\sin\frac{\pi}{3} \\ \sin\frac{\pi}{3} & \cos\frac{\pi}{3} \end{bmatrix}\begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1-\sqrt{3} \\ 1+\sqrt{3} \end{bmatrix}$$

E pontot a $(2,1)$ vektorral eltoljuk, hogy ne az origó, hanem a $(2,1)$

pont körüli elforgatottat kapjuk meg:

$$\begin{bmatrix} 1-\sqrt{3} \\ 1+\sqrt{3} \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3-\sqrt{3} \\ 2+\sqrt{3} \end{bmatrix}$$

□

7.27. PÉLDA (KOORDINÁTATENGELY KÖRÜLI FORGATÁS A TÉRBEN). *Írjuk fel a koordinátatengelyek körüli α szöggel való forgatás mátrixát.*

MEGOLDÁS. Tekintsük először a z-tengely körüli forgatást. Ekkor az $\mathbf{i}$ és $\mathbf{j}$ vektorok úgy transzformálódnak, mint a sík elforgatásánál, míg a $\mathbf{k}$ vektor helyben marad, tehát a bázisvektorok így transzformálódnak:

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} \cos\alpha \\ \sin\alpha \\ 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} -\sin\alpha \\ \cos\alpha \\ 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Így a z-tengely körüli forgatás mátrixa:

$$\begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Hasonlóképp kapjuk az x- és az y-tengely körüli forgatás mátrixát is:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}, \qquad \begin{bmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{bmatrix}.$$

Ez utóbbi mátrix előjelhibásnak tűnhet, de nem az, ha itt is a forgatás tengelyiránya felől nézve pozitív a forgásirány, azaz $\mathbf{k}$-t forgatjuk $\mathbf{i}$-be, és nem fordítva. □

7.28. TÉTEL (TENGELY KÖRÜLI FORGATÁS – RODRIGUES-FORMULA). *Ha $\mathbf{e} \in \mathbb{R}^3$ egységvektor, akkor az $\mathbf{e}$ körüli α szögű forgatás tetszőleges $\mathbf{x}$ vektort az*

$$\mathbf{x}\cos\alpha + (\mathbf{e}\times\mathbf{x})\sin\alpha + \mathbf{e}(\mathbf{e}\cdot\mathbf{x})(1-\cos\alpha) \tag{7.7}$$

vektorba visz. E leképezés mátrixa

$$\begin{aligned} \mathbf{R} &= \mathbf{I} + \sin\alpha[\mathbf{e}]_\times + (1-\cos\alpha)[\mathbf{e}]_\times^2 \\ &= \mathbf{I} + \sin\alpha[\mathbf{e}]_\times + (1-\cos\alpha)(\mathbf{e}\mathbf{e}' - \mathbf{I}) \end{aligned} \tag{7.8}$$

BIZONYÍTÁS. Ha $\mathbf{x}$ párhuzamos $\mathbf{e}$-vel, akkor elforgatottja önmaga, és valóban, ekkor $(\mathbf{e}\times\mathbf{x}) = \mathbf{0}$ és $\mathbf{e}(\mathbf{e}\cdot\mathbf{x}) = \mathbf{x}$, így a (7.7) képlet $\mathbf{x}$-et ad eredményül.

A továbbiakban legyen tehát $\mathbf{x}$ az $\mathbf{e}$-val nem párhuzamos vektor. Jelölje $\mathbf{x}$-nek az $\mathbf{e}$-ra eső merőleges vetületét $\mathbf{x_e}$, azaz legyen

$$\mathbf{x_e} = (\mathbf{e} \cdot \mathbf{x})\mathbf{e}.$$

Jelölje továbbá az $\mathbf{x}$ vektornak az $\mathbf{e}$-re merőleges síkra eső merőleges vetületét $\mathbf{x}_1$, azaz

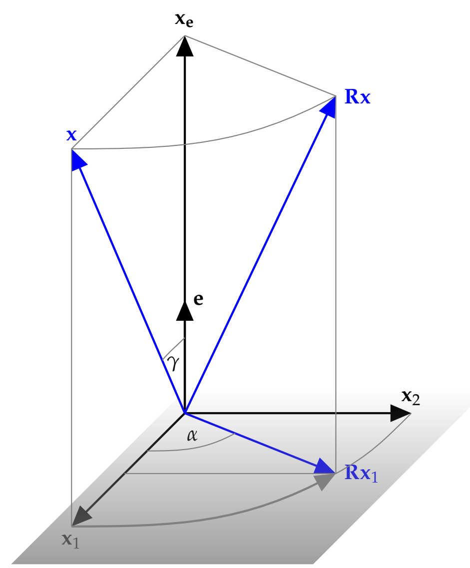
$$\mathbf{x}_1 = \mathbf{x} - (\mathbf{e} \cdot \mathbf{x})\mathbf{e}.$$

Végül legyen $\mathbf{x}_2 = \mathbf{e} \times \mathbf{x}$. Világos, hogy $\mathbf{x}_1 \perp \mathbf{x}_2$. E két vektor hossza:

$$\begin{aligned} |\mathbf{x}_1| &= |\mathbf{x}| \sin\gamma, \\ |\mathbf{x}_2| &= |\mathbf{e}||\mathbf{x}| \sin\gamma = |\mathbf{x}| \sin\gamma, \end{aligned}$$

tehát $|\mathbf{x}_1| = |\mathbf{x}_2|$. Ha R jelöli a forgató leképezést, akkor

$$\begin{aligned} R\mathbf{x}_1 &= \mathbf{x}_1 \cos\alpha + \mathbf{x}_2 \sin\alpha \\ &= (\mathbf{x} - (\mathbf{e} \cdot \mathbf{x})\mathbf{e}) \cos\alpha + (\mathbf{e} \times \mathbf{x}) \sin\alpha. \end{aligned}$$



Mivel $R\mathbf{x}_u = \mathbf{x}_u$, és $\mathbf{x} = \mathbf{x}_u + \mathbf{x}_1$, ezért

$$\begin{aligned} R\mathbf{x} &= R\mathbf{x}_u + R\mathbf{x}_1 \\ &= (\mathbf{e} \cdot \mathbf{x})\mathbf{e} + (\mathbf{x} - (\mathbf{e} \cdot \mathbf{x})\mathbf{e}) \cos\alpha + (\mathbf{e} \times \mathbf{x}) \sin\alpha \\ &= \mathbf{x} \cos\alpha + (\mathbf{e} \times \mathbf{x}) \sin\alpha + \mathbf{e}(\mathbf{e} \cdot \mathbf{x})(1 - \cos\alpha). \end{aligned}$$

Ezzel igazoltuk a (7.7) formulát. A leképezés könnyen átírható mátrixszorzat alakba:

$$\cos\alpha \mathbf{I}\mathbf{x} + [\mathbf{e}]_\times \mathbf{x} \sin\alpha + (1 - \cos\alpha)(\mathbf{e}\mathbf{e}')\mathbf{x},$$

így a forgatás $\mathbf{R}$ mátrixa

$$\mathbf{R} = \cos\alpha \mathbf{I} + \sin\alpha [\mathbf{e}]_\times + (1 - \cos\alpha)(\mathbf{e}\mathbf{e}').$$

Egyszerű számolással igazolható, hogy $\mathbf{e}\mathbf{e}' - \mathbf{I} = [\mathbf{e}]_\times^2$ (ld. **??** feladat), amiből azonnal adódnak a (7.8) képletei. □

7.29. PÉLDA (FORGATÁS MÁTRIXA). *Írjuk fel annak a leképezésnek a mátrixát, mely az $(2, 0, 1)$ vektor egyenese körül α szöggel forgat, ahol $\cos\alpha = \frac{2}{3}$. Határozzuk meg a $(3, 2, -1)$ vektor elforgatottját! Más eredményt kapnánk-e, ha a $(-2, 0, -1)$ vektor egyenese körül kéne forgatnunk α szöggel?*

MEGOLDÁS. Az $\mathbf{e} = \frac{1}{\sqrt{5}}(1, 0, 2)$ egységvektorral

$$[\mathbf{e}]_\times = \frac{1}{\sqrt{5}} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & -2 \\ 0 & 2 & 0 \end{bmatrix}, \quad [\mathbf{e}]_\times^2 = \frac{1}{5} \begin{bmatrix} -1 & 0 & 2 \\ 0 & -5 & 0 \\ 2 & 0 & -4 \end{bmatrix}$$

Így rövid számolás után a forgatás mátrixa

$$\begin{aligned}\mathbf{R} &= \mathbf{I} + \sin\alpha[\mathbf{e}]_\times + (1-\cos\alpha)[\mathbf{e}]_\times^2 \\ &= \begin{bmatrix} 14/15 & -1/3 & 2/15 \\ 1/3 & 2/3 & -2/3 \\ 2/15 & 2/3 & 11/15 \end{bmatrix} = \frac{1}{15}\begin{bmatrix} 14 & -5 & 2 \\ 5 & 10 & -10 \\ 2 & 10 & 11 \end{bmatrix}\end{aligned}$$

és $\mathbf{R}\cdot(3,2,-1) = (2,3,1)$.

A $(-1,0,-2)$ vektor körüli forgatással más eredményt kapnánk, hisz a forgatás iránya a vektor irányától is függ, és mivel az az ellenkezőjére változott, így a forgásirány is ellenkező irányú lesz. □

7.30. PÉLDA (A FORGATÁS MÁTRIXÁNAK INVERZE). *Határozzuk meg a síkot α szöggel elforgató mátrix inverzét!*

MEGOLDÁS. Először megállapítjuk, hogy a forgatás mátrixa invertálható, ugyanis determinánsa nem 0, hiszen $\begin{vmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{vmatrix} = \cos^2\alpha + \sin^2\alpha = 1$. Az egyik lehetséges megoldás, hogy egyszerűen a 2×2-es mátrixok 5.17. tételben leírt invertálási technikáját alkalmazva határozzuk meg az inverzet, mely szerint

$$\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}^{-1} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}.$$

Egy másik megoldás arra épül, hogy a 7.3. állítás szerint két mátrix pontosan akkor inverze egymásnak, ha a hozzájuk tartozó lineáris leképezések is inverzei egymásnak. Az α szöggel való elforgatásnak, mint leképezésnek az inverze nyilvánvalóan a $-\alpha$ szöggel való elforgatás, tehát mátrixaik is egymás inverzei. Eszerint

$$\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}^{-1} = \begin{bmatrix} \cos(-\alpha) & -\sin(-\alpha) \\ \sin(-\alpha) & \cos(-\alpha) \end{bmatrix} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}. \quad □$$

Merőleges vetítés A merőlegesség mind az elméleti matematika, mind az alkalmazások fontos fogalma.

7.31. ÁLLÍTÁS (EGYENESRE VALÓ MERŐLEGES VETÍTÉS MÁTRIXA). *A sík vagy a tér vektorait egy* $\mathbf{b}$ *irányvektorú egyenesre merőlegesen vetítő leképezés mátrixa*

$$\mathbf{P} = \frac{1}{\mathbf{b}'\mathbf{b}}\mathbf{b}\mathbf{b}'. \tag{7.9}$$

Speciálisan e mátrix alakja

$$\mathbf{P} = \mathbf{e}\mathbf{e}', \tag{7.10}$$

ha az egyenes irányvektora az $\mathbf{e}$ *egységvektor.*

BIZONYÍTÁS. A 7.9. állítás szerint a merőleges vetítés lineáris leképezés, van tehát mátrixa. Az 1.24. tétel szerint ha $\mathbf{x}$ egy tetszőleges vektor és $\mathbf{e}$ egy egységvektor, akkor $\mathbf{x}$-nek a $\mathbf{e}$ egyenesére eső merőleges vetülete

$$\operatorname{proj}_{\mathbf{e}} \mathbf{x} = (\mathbf{x} \cdot \mathbf{e}) \cdot \mathbf{e}.$$

Ennek mátrixszorzással való átírása:

$$(\mathbf{x} \cdot \mathbf{e})\mathbf{e} = \mathbf{e}(\mathbf{e} \cdot \mathbf{x}) = \mathbf{e}(\mathbf{e}'\mathbf{x}) = (\mathbf{e}\mathbf{e}')\mathbf{x},$$

tehát

$$\operatorname{proj}_{\mathbf{e}} \mathbf{x} = (\mathbf{e}\mathbf{e}')\mathbf{x}.$$

Ebből kiolvasható, hogy az $\mathbf{e}$ egységvektor-irányú egyenesre való merőleges vetítés mátrixa

$$\mathbf{P} = \mathbf{e}\mathbf{e}'.$$

Ha $\mathbf{b}$ egy tetszőleges zérustól különböző vektor, akkor az $\mathbf{e} = \mathbf{b}/|\mathbf{b}|$ jelölés mellett $\mathbf{P} = \mathbf{e}\mathbf{e}' = \mathbf{b}\mathbf{b}'/|\mathbf{b}|^2$, ami $|\mathbf{b}|^2 = \mathbf{b}'\mathbf{b}$ behelyettesítésével bizonyítja a tételt. □

► A tételből következik, hogy a sík vektorait az x-tengellyel α szöget bezáró egyenesre merőlegesen vetítő lineáris leképezés mátrixa

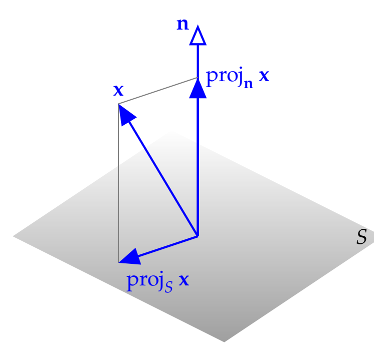
$$\mathbf{P} = \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix} \begin{bmatrix} \cos\alpha & \sin\alpha \end{bmatrix} = \begin{bmatrix} \cos^2\alpha & \sin\alpha\cos\alpha \\ \sin\alpha\cos\alpha & \sin^2\alpha \end{bmatrix}, \tag{7.11}$$

mivel ekkor $\mathbf{e} = (\cos\alpha, \sin\alpha)$.

7.32. ÁLLÍTÁS (SÍKRA VALÓ MERŐLEGES VETÍTÉS MÁTRIXA). *A tér vektorait az $\mathbf{n}$ normálvektorú síkra merőlegesen vetítő leképezés mátrixa*

$$\mathbf{P} = \mathbf{I} - \mathbf{n}\mathbf{n}'.$$

BIZONYÍTÁS. Egy tetszőleges $\mathbf{x}$ vektornak a normálvektor egyenesére eső merőleges vetülete a 7.31. állítás szerint $\operatorname{proj}_{\mathbf{n}} \mathbf{x} = \mathbf{n}\mathbf{n}'$. Az $\mathbf{n}$ normálvektorú S síkra eső merőleges vetületre $\operatorname{proj}_{S} \mathbf{x} = \mathbf{x} - \operatorname{proj}_{\mathbf{n}} \mathbf{x} = \mathbf{x} - (\mathbf{n}\mathbf{n}')\mathbf{x}$ (lásd a 7.19. ábrát). Ebből következik, hogy a síkra való merőleges vetítés mátrixa $\mathbf{I} - \mathbf{n}\mathbf{n}'$.



7.19. ábra: Vektor vetülete egy síkra

7.33. PÉLDA (SÍKRA ESŐ MERŐLEGES VETÜLET KISZÁMÍTÁSA). *Határozzuk meg a $(-2, 1, 3)$ vektornak a $2x + y - 2z = 0$ egyenletű síkra eső merőleges vetületét! (ld. később a 7.38. példát)*

MEGOLDÁS. Az egységnyi hosszú normálvektor $\mathbf{n} = (2/3, 1/3, -2/3)$. A $\mathbf{P}$ vetítő mátrix

$$\mathbf{P} = \mathbf{I}_3 - \mathbf{n}\mathbf{n}' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{1}{9} \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} \begin{bmatrix} 2 & 1 & -2 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{bmatrix}.$$

Így a $(-2, 1, 3)$ vektor merőleges vetülete

$$\frac{1}{9}\begin{bmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{bmatrix}\begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}.$$

□

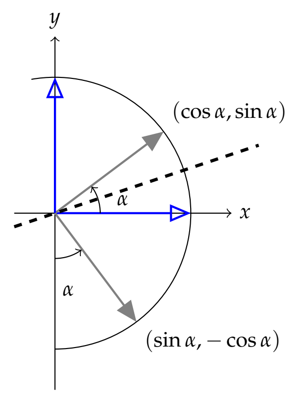
Tükrözés Vizsgáljuk meg síkban az egyenesre való és térben a síkra való tükrözés mátrixát!

7.34. ÁLLÍTÁS (SÍKBELI TÜKRÖZÉS MÁTRIXA). *A sík vektorait az x-tengellyel $\alpha/2$ szöget bezáró egyenesre tükröző lineáris leképezés mátrixa*

$$\begin{bmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{bmatrix}.$$

BIZONYÍTÁS. a 7.9. állítás szerint a tükrözés lineáris leképezés. A vektorok tükörképe csak a tükrözés tengelyének állásától függ, ami most $\alpha/2$. Helyvektorokban gondolkodva a tükrözés tengelyének át kell mennie az origón.

A mellékelt ábráról leolvasható, hogy $\mathbf{i}$ tükörképe $A\mathbf{i} = \left[\begin{smallmatrix} \cos\alpha \\ \sin\alpha \end{smallmatrix}\right]$, míg a $\mathbf{j}$ vektoré $A\mathbf{j} = \left[\begin{smallmatrix} \sin\alpha \\ -\cos\alpha \end{smallmatrix}\right]$. Így a sík vektorait az első tengellyel $\alpha/2$ szöget bezáró egyenesre tükröző leképezés mátrixa $\left[\begin{smallmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{smallmatrix}\right]$. □



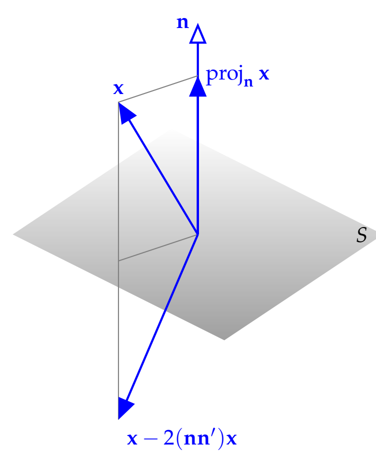
7.20. ábra: Az $\mathbf{i}$ és $\mathbf{j}$ vektorok egy egyenesre való tükörképe

A térben egy síkra való tükrözés feladata a síkra való vetítéshez hasonlóan adódik:

7.35. ÁLLÍTÁS (SÍKRA VALÓ TÜKRÖZÉS MÁTRIXA). *Igazoljuk, hogy a tér vektorait az $\mathbf{n}$ normálvektorú síkra tükröző leképezés mátrixa*

$$\mathbf{P} = \mathbf{I} - 2\mathbf{n}\mathbf{n}'.$$

BIZONYÍTÁS. A 7.32. állításhoz hasonlóan minden leolvasható a mellékelt 7.21. ábráról: ha $\mathbf{x}$-ből kivonjuk a $\operatorname{proj}_{\mathbf{n}} \mathbf{x}$ vektort, akkor a síkra eső vetületet kapjuk, így ha a kétszeresét vonjuk ki, a tükörképhez jutunk. E leképezés mátrixa az $\mathbf{x} - 2(\mathbf{n}\mathbf{n}')\mathbf{x} = (\mathbf{I} - 2\mathbf{n}\mathbf{n}')\mathbf{x}$ összefüggésből adódik.



7.21. ábra: Vektor tükörképe egy síkra

Vetítés Tárgyaltuk a merőleges vetítést. Vetíteni azonban másként is lehet.

7.36. PÉLDA. *Határozzuk meg annak a lineáris leképezésnek a mátrixát, mely a tér összes pontját az $(1, -2, 1)$ vektorral párhuzamos irányban az $x + y + 2z = 0$ egyenletű síkra vetíti.*

MEGOLDÁS. Elemi geometriai eszközökkel könnyen látható, hogy e leképezés valóban lineáris. Világos, hogy a képtér az $x + y + 2z = 0$

egyenletű sík összes vektora lesz. E teret megkapjuk a sík egyenletéből, ha azt mint egyenletrendszert megoldjuk. A megoldás $(-s-2t, s, t)$, azaz e tér egy bázisa a $(-1,1,0)$, és a $(-2,0,1)$ vektorokból áll. Könnyen látható az is, hogy a nulltérbe pontosan azok a vektorok tartoznak, amelyek párhuzamosak a vetítő vektorral, azaz az $(1,-2,1)$ vektorral. A vetítés $\mathbf{P}$ mátrixa tehát eleget kell hogy tegyen az alábbi feltételeknek:

$$\mathbf{P}\begin{bmatrix}-1\\1\\0\end{bmatrix}=\begin{bmatrix}-1\\1\\0\end{bmatrix},\quad \mathbf{P}\begin{bmatrix}-2\\0\\1\end{bmatrix}=\begin{bmatrix}-2\\0\\1\end{bmatrix},\quad \mathbf{P}\begin{bmatrix}1\\-2\\1\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}.$$

E három feltétel egyetlen mátrixszorzásba foglalható:

$$\mathbf{P}\begin{bmatrix}-1&-2&1\\1&0&-2\\0&1&1\end{bmatrix}=\begin{bmatrix}-1&-2&0\\1&0&0\\0&1&0\end{bmatrix}, \text{ ahonnan } \mathbf{P}=\begin{bmatrix}0&-1&-2\\2&3&4\\-1&-1&-1\end{bmatrix}. \square$$

► Az előző feladatban kapott $\mathbf{P}$ mátrix eleget tesz a $\mathbf{P}^2=\mathbf{P}$ összefüggésnek. Ez szemléletesen világos is, hisz ha P jelöli a lineáris transzformációt, akkor arra is igaz, hogy $P^2=P$. Ez abból következik, hogy a P vetítés az $x+y+2z=0$ egyenletű sík minden vektorát helyben hagyja, másrész bármely $\mathbf{x}$ vektor esetén $P\mathbf{x}$ ebben a síkban van, így a második vetítés már minden vektort helyben hagy.

Eltolás Az eltolás nem lineáris leképezés, hisz minden vektorhoz egy konstans vektort ad, tehát a nullvektort nem a nullvektorba képzi. Egy szellemes ötlettel mégis megvalósítható lineáris leképezéssel.

El szeretnénk tolni a síkot egy (a,b) vektorral. Az ötlet az, hogy beágyazzuk a síkot a térbe, és ott keresünk egy olyan térbeli lineáris leképezést, amely ezt a síkot eltolja (hogy másutt meg mit csinál, nem is érdekes). Legyen tehát a vizsgált sík a $z=1$ egyenletű sík, és keressük azt a lineáris T leképezést, melyre

$$T\begin{bmatrix}x\\y\\1\end{bmatrix}=\begin{bmatrix}x+a\\y+b\\1\end{bmatrix}.$$

Ez ugyan még mindig nem tűnik lineárisnak, de mivel $z=1$, ezért a

$$T\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}x+az\\y+bz\\z\end{bmatrix}$$

leképezés már minden tekintetben megfelel. E leképezés mátrixa

$$\mathbf{T}=T\begin{bmatrix}\mathbf{i}&\mathbf{j}&\mathbf{k}\end{bmatrix}=\begin{bmatrix}1&0&a\\0&1&b\\0&0&1\end{bmatrix}.$$

Hasonló ötlettel a tér eltolása is megvalósítható. A tér tetszőleges $(x, y, z) \mapsto (x+a, y+b, z+c)$ eltolása megvalósítható a következő mátrixleképezéssel:

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & a \\ 0 & 1 & 0 & b \\ 0 & 0 & 1 & c \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{T}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & a \\ 0 & 1 & 0 & b \\ 0 & 0 & 1 & c \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x+a \\ y+b \\ z+c \\ 1 \end{bmatrix}.$$

Merőleges vetítés és a legjobb közelítés

A legjobb közelítés, a legkisebb négyzetek elve, vagy a lineáris regresszió az alkalmazásokban igen gyakran előforduló fontos fogalmak. Lényegük az $\mathbb{R}^n$ egy alterére való merőleges vetítésének fogalmával jól megvilágítható.

Merőleges vetítés $\mathbb{R}^n$ egy alterére Azt mondjuk, hogy $\mathbb{R}^n$ egy $\mathbf{v}$ vektorának a $\mathcal{W}$ altérre eső *merőleges vetülete* a $\mathbf{w}$ vektor, ha $\mathbf{w} \in \mathcal{W}$, és $\mathbf{v} - \mathbf{w}$ merőleges a $\mathcal{W}$ altérre, azaz $\mathbf{v} - \mathbf{w} \in \mathcal{W}^\perp$. A $\mathbf{v} - \mathbf{w}$ vektort a $\mathbf{v}$ vektor $\mathcal{A}$ altérre *merőleges összetevőjének* nevezzük.

Kérdés, hogy létezik-e minden vektornak egy altérre eső merőleges vetülete, és hogy egyértelmű-e. A 3.43. tétel *c)* pontja szerint ha $\mathcal{W}$ az $\mathbb{R}^n$ egy altere, akkor minden $\mathbf{v} \in \mathbb{R}^n$ vektor egyértelműen felbomlik egy $\mathcal{W}$-beli $\mathbf{w}$ és egy $\mathcal{W}^\perp$-beli $\mathbf{w}^\perp$ vektor összegére. Ez azt jelenti, hogy e $\mathbf{w}$ vektor épp a $\mathbf{v}$ vektor $\mathcal{W}$ altérre eső merőleges vetülete. Ezt az egyértelműen létező vektort – összhangban korábbi jelölésünkkel – jelölje $\operatorname{proj}_{\mathcal{W}} \mathbf{v}$.

7.37. TÉTEL (ALTÉRRE VALÓ VETÍTÉS MÁTRIXA). *Ha $\mathcal{W}$ az $\mathbb{R}^n$ egy altere, és az $\mathbf{A}$ mátrix oszlopvektorai a $\mathcal{W}$ egy bázisát alkotják (tehát $\mathbf{A}$ teljes oszloprangú), akkor a $\mathcal{W}$ altérre való merőleges vetítés, azaz a $\operatorname{proj}_{\mathcal{W}}$ leképezés mátrixa*

$$\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'.$$

BIZONYÍTÁS. Legyen a $\mathbf{v} \in \mathbb{R}^n$ vektor $\mathcal{W}$-re eső merőleges vetülete $\mathbf{w}$. Mivel $\mathbf{A}$ definíciója szerint $\mathbf{A}$ oszloptere $\mathcal{W}$, ezért létezik olyan $\mathbf{x}$ vektor, hogy $\mathbf{Ax} = \mathbf{w}$. Másrészt $\mathcal{W} = \mathcal{O}(\mathbf{A})$ miatt $\mathcal{W}^\perp = \mathcal{N}(\mathbf{A}')$, így $\mathbf{v} - \mathbf{w}$ benne van $\mathbf{A}'$ nullterében, mivel a merőleges vetület definíciója szerint $\mathbf{v} - \mathbf{w}$ merőleges $\mathcal{W}$-re. Eszerint $\mathbf{A}'(\mathbf{v} - \mathbf{w}) = \mathbf{0}$, azaz $\mathbf{A}'(\mathbf{v} - \mathbf{Ax}) = \mathbf{0}$. Átrendezve kapjuk, hogy

$$\mathbf{A}'\mathbf{Ax} = \mathbf{A}'\mathbf{v}.$$

Az $\mathbf{A}$ mátrix teljes oszloprangú, így a **??** tétel szerint $\mathbf{A}'\mathbf{A}$ invertálható, azaz $\mathbf{x} = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{v}$, amiből kapjuk, hogy $\operatorname{proj}_{\mathcal{W}} \mathbf{v} = \mathbf{w} = \mathbf{A}\mathbf{x} = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{v}$, ami bizonyítja az állítást. □

► A tételbeli képlet könnyen megjegyezhető, hisz összhangban van az egyenesre való merőleges vetítés (7.9) képletével. Ha ugyanis az $\mathbf{A}$ mátrix egyetlen oszlopból áll, $(\mathbf{A}'\mathbf{A})^{-1}$ egyetlen szám, ami kiemelhető, azaz az $\mathbf{A} = \mathbf{b}$ jelöléssel $\mathbf{b}(\mathbf{b}'\mathbf{b})^{-1}\mathbf{b}' = \frac{1}{\mathbf{b}'\mathbf{b}}\mathbf{b}\mathbf{b}'$.

7.38. példa (Merőleges vetület kiszámítása). *Határozzuk meg a $(-2, 1, 3)$ vektornak az $(1, 0, 1)$ és a $(-1, 2, 0)$ vektorok által kifeszített síkra eső merőleges vetületét! (ld. még a 7.33. példát)*

Megoldás. Az altér bázisvektoraiból képzett mátrix

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 0 & 2 \\ 1 & 0 \end{bmatrix}, \text{ amiből } \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}' = \frac{1}{9}\begin{bmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{bmatrix}.$$

Így a $(-2, 1, 3)$ vektor merőleges vetülete

$$\frac{1}{9}\begin{bmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{bmatrix}\begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}.$$

Ez a feladat megegyezik a 7.33. példabelivel, mivel ennek a síknak is $2x + y - 2z = 0$ az egyenlete, ugyanis $(1, 0, 1) \times (-1, 2, 0) = (-2, -1, 2)$. □

Melyik mátrix merőleges vetítés mátrixa? Olyan – könnyen ellenőrizhető – feltételeket keresünk egy lineáris leképezés mátrixára, melyek segítségével azonnal megállapítható, hogy a mátrixleképezés merőleges vetítés-e.

7.39. tétel (Merőleges vetítés mátrixai). *Egy $\mathbf{P}$ mátrix pontosan akkor merőleges vetítés mátrixa, ha $\mathbf{P} = \mathbf{P}' = \mathbf{P}^2$.*

Bizonyítás. A $\mathbf{P} = \mathbf{P}^2$ feltétel szükségessége szemléletesen világos, hisz minden P lineáris leképezés, mely az egész $\mathbb{R}^n$ teret egy altérre – nevezetesen $\operatorname{Im} P$-re – vetíti, az altér vektorait helyben hagyja. Tehát $P^2\mathbf{x} = P\mathbf{x}$ minden $\mathbf{x}$-re fennáll, így ennek az összefüggésnek P minden mátrixára is igaznak kell lennie.

($\Longrightarrow$) Tegyük fel, hogy $\mathbf{P}$ egy P merőleges vetítés mátrixa $\mathbb{R}^n$ standard bázisában. Tekintsük $\operatorname{Im}(P) = \mathcal{O}(\mathbf{P})$ egy tetszőleges bázisát, és legyen $\mathbf{A}$ az a mátrix, melynek e bázis elemei az oszlopai. A 7.37. tétel szerint ekkor $\mathbf{P} = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$. Erre viszont könnyen ellenőrizhető, a tételbeli feltétel.

$$\begin{aligned}\mathbf{P}^2 &= \left(\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\right)^2 = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}' \\ &= \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}' = \mathbf{P},\end{aligned}$$

másrészt

$$\begin{aligned}\mathbf{P}' &= \left(\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\right)' = \mathbf{A}\left((\mathbf{A}'\mathbf{A})^{-1}\right)'\mathbf{A}' \\ &= \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}' = \mathbf{P}.\end{aligned}$$

($\Longleftarrow$) Tegyük fel, hogy $\mathbf{P} = \mathbf{P}' = \mathbf{P}^2$. Megmutatjuk, hogy $\mathbf{P}$ az $\mathcal{O}(\mathbf{P})$-re való merőleges vetítés mátrixa. Ehhez elég megmutatnunk, hogy az $\mathbf{x} - \mathbf{P}\mathbf{x}$ vektor merőleges $\mathcal{O}(\mathbf{P})$-re bármely $\mathbf{x}$ vektor esetén. A $\mathbf{P}^2 = \mathbf{P}$ feltétel miatt $\mathbf{P}(\mathbf{x} - \mathbf{P}\mathbf{x}) = \mathbf{P}\mathbf{x} - \mathbf{P}^2\mathbf{x} = \mathbf{0}$, tehát $\mathbf{x} - \mathbf{P}\mathbf{x} \in \mathcal{N}(\mathbf{P})$, de $\mathbf{P} = \mathbf{P}'$, így $\mathbf{x} - \mathbf{P}\mathbf{x} \in \mathcal{N}(\mathbf{P}')$. Ez épp azt jelenti, hogy $\mathbf{x} - \mathbf{P}\mathbf{x}$ merőleges $\mathcal{O}(\mathbf{P})$-re, és ezt akartuk belátni. □

► A $\mathbf{P} = \mathbf{P}'$ összefüggés azt jelenti, hogy $\mathbf{P}$ szimmetrikus. A $\mathbf{P}^2 = \mathbf{P}$ tulajdonságnak eleget tevő mátrixokat *idempotensnek* nevezzük. A tétel tehát úgy is fogalmazható, hogy egy mátrix pontosan akkor egy merőleges vetítés mátrixa, ha idempotens és szimmetrikus.

► Később látni fogjuk, hogy a – később definiálandó – nem feltétlenül merőleges vetítés mátrixai egybe esnek az idempotens mátrixokkal, tehát a vetítő lineáris leképezések az idempotens lineáris leképezésekkel.

► Azt, hogy egy vetítés hány dimenziós térre vetít, annak rangja mondja meg, hisz az megegyezik a képtér dimenziójával.

7.40. PÉLDA. *Igazoljuk, hogy az*

$$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix},\quad \frac{1}{2}\begin{bmatrix}1&0&0&1\\0&1&1&0\\0&1&1&0\\1&0&0&1\end{bmatrix},\quad \frac{1}{4}\begin{bmatrix}3&-1&-1&-1\\-1&3&-1&-1\\-1&-1&3&-1\\-1&-1&-1&3\end{bmatrix}$$

mátrixok merőleges vetítés mátrixai! Hány dimenziós térre vetítenek?

MEGOLDÁS. Könnyen ellenőrizhető, hogy mindegyik mátrix szimmetrikus és idempotens, azaz kielégíti a $\mathbf{P}' = \mathbf{P}$ és a $\mathbf{P}^2 = \mathbf{P}$ egyenlőségeket. Az első két mátrixról átalakítás nélkül is leolvasható, hogy rangjuk 2. A harmadik mátrix rangja 3, ugyanis egyrészt legalább 3, hisz ha kivonjuk az utolsó sort az első háromból, egy 3×3-as egységmátrixot kapunk benne, másrészt nem lehet 4, mert a négy sorvektor összege a zérusvektor, azaz lineárisan összefüggők. □

Altértől való távolság Adva van az $\mathbb{R}^n$ tér egy $\mathbf{x}$ vektora és egy $\mathcal{W}$ altere. $\mathbf{x}$-nek a $\mathcal{W}$ *altértől való távolságán* a $\mathcal{W}$ altér $\mathbf{x}$-hez legközelebbi $\mathbf{w}$ vektorának tőle való távolságát értjük. Kérdés azonban, hogy létezik-e ilyen vektor egyáltalán! Meg fogjuk mutatni, hogy ilyen $\mathbf{w}$ vektor létezik és egyértelmű. E vektort az $\mathbf{x}$ vektor $\mathcal{W}$-beli *legjobb közelítésének* nevezzük.

7.41. TÉTEL (LEGJOBB KÖZELÍTÉS TÉTELE). *Adva van az* $\mathbb{R}^n$ *tér egy* $\mathbf{x}$ *vektora és egy* $\mathcal{W}$ *altere. Az* $\mathbf{x}$ *vektornak egyetlen* $\mathcal{W}$*-beli legjobb* $\hat{\mathbf{x}}$ *közelítése van, nevezetesen* $\hat{\mathbf{x}} = \operatorname{proj}_{\mathcal{W}} \mathbf{x}$.

BIZONYÍTÁS. Legyen $\mathbf{w}$ a $\mathcal{W}$ egy tetszőleges vektora. Ekkor

$$\mathbf{x} - \mathbf{w} = (\mathbf{x} - \operatorname{proj}_{\mathcal{W}} \mathbf{x}) + (\operatorname{proj}_{\mathcal{W}} \mathbf{x} - \mathbf{w}).$$

A merőleges vetítés definíciója miatt az egyenlőség jobb oldalán álló első kifejezés $\mathcal{W}^{\perp}$ eleme, míg a második $\mathcal{W}$ eleme. Tehát az $\mathbf{x} - \operatorname{proj}_{\mathcal{W}} \mathbf{x}$ és a $\operatorname{proj}_{\mathcal{W}} \mathbf{x} - \mathbf{w}$ vektorok merőlegesek egymásra, így alkalmazható rájuk Pithagorász tétele:

$$|\mathbf{x} - \mathbf{w}|^2 = |\mathbf{x} - \operatorname{proj}_{\mathcal{W}} \mathbf{x}|^2 + |\operatorname{proj}_{\mathcal{W}} \mathbf{x} - \mathbf{w}|^2.$$

Ebből világos, hogy

$$|\mathbf{x} - \mathbf{w}|^2 \geq |\mathbf{x} - \operatorname{proj}_{\mathcal{W}} \mathbf{x}|^2,$$

és egyenlőség csak akkor állhat fönn, ha $\mathbf{w} = \hat{\mathbf{x}} = \operatorname{proj}_{\mathcal{W}} \mathbf{x}$, ami egyúttal a legjobb közelítés egyértelműségét is bizonyítja. □

► E tétel egyik következménye, hogy $\mathbb{R}^n$ minden vektora felbontható egy $\mathcal{W}$-beli és egy rá merőleges vektor összegére, ugyanis

$$\mathbf{x} = \operatorname{proj}_{\mathcal{W}} \mathbf{x} + \mathbf{w}^{\perp}, \quad \text{ahol} \quad \mathbf{w}^{\perp} = \mathbf{x} - \operatorname{proj}_{\mathcal{W}} \mathbf{x}.$$

Ennél azonban több is igaz, nevezetesen az, hogy e felbontás egyértelmű.

7.42. TÉTEL (VEKTOR FELBONTÁSA ÖSSZETEVŐKRE). *Adva van az* $\mathbb{R}^n$ *tér egy* $\mathbf{x}$ *vektora és egy* $\mathcal{W}$ *altere. Az* $\mathbf{x}$ *vektor egyértelműen felbomlik egy* $\mathcal{W}$*-beli* $\mathbf{w}$ *és egy* $\mathcal{W}$*-re merőleges* $\mathbf{w}^{\perp}$ *vektor összegére, nevezetesen* $\mathbf{w} = \operatorname{proj}_{\mathcal{W}} \mathbf{x}$ *és* $\mathbf{w}^{\perp} = \mathbf{x} - \mathbf{w}$.

BIZONYÍTÁS. Tegyük fel, hogy létezik $\mathbf{x}$-nek egy másik ilyen tulajdonságú felbontása is, tehát $\mathbf{x} = \mathbf{w} + \mathbf{w}^{\perp}$ és $\mathbf{x} = \mathbf{v} + \mathbf{v}^{\perp}$. A második egyenletet az elsőből kivonva, majd átrendezve kapjuk, hogy

$$\mathbf{v} - \mathbf{w} = \mathbf{w}^{\perp} - \mathbf{v}^{\perp}.$$

A bal oldal eleme $\mathcal{W}$-nek a jobb oldali vektor viszont merőleges rá, hisz mindkét vektor eleme a $\mathcal{W}^{\perp}$ altérnek. Ez viszont csak akkor állhat fenn, ha mindkét oldal egyenlő a zérusvektorral, tehát $\mathbf{v} = \mathbf{w}$. □

7.43. PÉLDA. *Tekintsük az* $\mathbb{R}^4$ *tér* $(1, -1, 1, 0)$ *és* $(0, 1, -1, 0)$ *vektorai által kifeszített* $\mathcal{W}$ *alterét és legyen* $\mathbf{x} = (8, 4, 2, 1)$. *Bontsuk fel az* $\mathbf{x}$ *vektort* $\mathcal{W}$*-be eső és* $\mathcal{W}$*-re merőleges vektorok összegére.*

MEGOLDÁS. A $\mathcal{W}$-re való merőleges vetítés mátrixa $\mathbf{P} = \mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'$, ahol $\mathbf{W}$ két oszlopa a megadott két bázisvektor, tehát

$$\mathbf{W} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 1 & -1 \\ 0 & 0 \end{bmatrix}, \quad \text{amiből} \quad \mathbf{P}\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & -1/2 & 0 \\ 0 & -1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 8 \\ 4 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 1 \\ -1 \\ 0 \end{bmatrix}.$$

Így $\operatorname{proj}_{\mathcal{W}} \mathbf{x} = \mathbf{P}\mathbf{x} = (8, 1, -1, 0)$ és $\mathbf{x} - \operatorname{proj}_{\mathcal{W}} \mathbf{x} = (0, 3, 3, 1)$. Egyszerű számítással ellenőrizhető, hogy a $(8, 1, -1, 0) \in \mathcal{W}$ és hogy $(0, 3, 3, 1) \perp \mathcal{W}$, azaz merőleges a $\mathcal{W}$-t kifeszítő bázisvektorok mindegyikére. □

Egyenletrendszer optimális megoldása Az altérre való merőleges vetítés és a legjobb közelítés fogalmával olyan eszközhöz jutottunk, amellyel a lineáris egyenletrendszerek elmélete e szinten teljessé tehető. A gyakorlatban rendkívül gyakran előfordul, hogy az ismeretlen mennyiségek meghatározására méréseket végzünk, de az elkerülhetetlen mérési hibák ellenmondó egyenletrendszerre vezetnek. Hogyan határozható meg ekkor a valóságban bizonyosan létező megoldás, egy ellentmondásos, tehát nem megoldható egyenletrendszerből?

Tudjuk, hogy az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszer pontosan akkor oldható meg, ha $\mathbf{b}$ benne van az oszloptérben, azaz $\mathcal{O}(\mathbf{A})$-ban. Természetes ötlet, hogy $\mathbf{b}$ helyett az azt legjobban közelítő oszloptérbeli $\hat{\mathbf{b}} = \operatorname{proj}_{\mathcal{O}(\mathbf{A})} \mathbf{b}$ vektorral oldjuk meg az egyenletrendszert. Ez már biztosan megoldható lesz, és olyan megoldásokat szolgáltat, melyekre $\mathbf{Ax}$ ugyan nem lesz egyenlő $\mathbf{b}$-vel, de attól a lehető legkisebb távolságra van. Az ilyen megoldásokat az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszer *optimális megoldásainak* vagy a *legkisebb négyzetek elve szerinti megoldásainak* nevezzük. Világos, hogy ha egy egyenletrendszer konzisztens, akkor optimális megoldásai megegyeznek a megoldásaival. E definícióból azt is látjuk, mit tegyünk, ha egy egyenletrendszer ellentmondásos (azaz inkonzisztens): határozzuk meg a $\hat{\mathbf{b}} = \operatorname{proj}_{\mathcal{O}(\mathbf{A})} \mathbf{b}$ vektort, és az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszer helyett oldjuk meg az $\mathbf{A}\hat{\mathbf{x}} = \hat{\mathbf{b}}$ egyenletrendszert. Ez kiindulásul jó, de adódik egy egyszerűbb módszer is.

Az természetes ötlet, hogy $\mathbf{b}$ helyett $\hat{\mathbf{b}}$-vel oldjuk meg az egyenletrendszert, de vajon nincs-e jobb ötlet, végülis miért épp e merőleges vetület adja számunkra a „legjobb megoldást” és mit is jelent itt a „legjobb”. Az igazi választ a valószínűségszámítási előismeretet igénylő Gauss–Markov-tétel adja meg, melyet a feladatok közt ismertetünk.

7.44. TÉTEL (EGYENLETRENDSZER OPTIMÁLIS MEGOLDÁSA). *Az* $\mathbf{Ax} = \mathbf{b}$ *egyenletrendszer optimális megoldásai megegyeznek az*

$$\mathbf{A}'\mathbf{A}\hat{\mathbf{x}} = \mathbf{A}'\mathbf{b} \tag{7.12}$$

egyenletrendszer megoldásaival. Ezek közül egyetlen egy esik az $\mathbf{A}$ *mátrix sorterébe, a legkisebb abszolút értékű.*

► A (7.12) egyenletrendszert az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszerhez tartozó *normálegyenlet-rendszernek* nevezzük. (A *normálegyenlet* kifejezés is helyes, ha a (7.12) kifejezésre, mint mátrixegyenletre gondolunk.)

Bizonyítás. Az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszer optimális megoldásai megegyeznek az $\mathbf{A\hat{x}} = \operatorname{proj}_{\mathcal{O}(\mathbf{A})} \mathbf{b}$ egyenletrendszer megoldásaival. Ezeket fogjuk tehát keresni.

Először megmutatjuk, hogy ha $\mathbf{\hat{x}}$ egy optimális megoldás, akkor $\mathbf{\hat{x}}$ kielégíti a (7.12) egyenletet. Mivel $\mathbf{b} - \operatorname{proj}_{\mathcal{O}(\mathbf{A})} \mathbf{b}$ a vetítés definíciója miatt merőleges $\mathcal{O}(\mathbf{A})$-ra, ezért $\mathbf{A}'$ nullterében van, tehát

$$\mathbf{A}'(\mathbf{b} - \operatorname{proj}_{\mathcal{O}(\mathbf{A})} \mathbf{b}) = \mathbf{0}.$$

Másrészt felhasználva, hogy $\mathbf{A\hat{x}} = \operatorname{proj}_{\mathcal{O}(\mathbf{A})} \mathbf{b}$, kapjuk, hogy

$$\mathbf{A}'(\mathbf{b} - \mathbf{A\hat{x}}) = \mathbf{0},$$

azaz átrendezés után

$$\mathbf{A}'\mathbf{A\hat{x}} = \mathbf{A}'\mathbf{b}.$$

Ezután megmutatjuk, hogy a (7.12) egyenletet kielégítő minden $\mathbf{\hat{x}}$ vektor optimális megoldás. Ha (7.12) teljesül, akkor

$$\mathbf{A}'(\mathbf{b} - \mathbf{A\hat{x}}) = \mathbf{0},$$

tehát $\mathbf{b} - \mathbf{A\hat{x}}$ benne van $\mathbf{A}'$ nullterében, így merőleges $\mathbf{A}$ oszlopterére. Ezért a

$$\mathbf{b} = \mathbf{A\hat{x}} + (\mathbf{b} - \mathbf{A\hat{x}})$$

felbontás két merőleges kiegészítő altérbe eső megoldás, hisz $\mathbf{A\hat{x}}$ az $\mathbf{A}$ oszlopterébe esik. Így a merőleges vetület definíciója szerint

$$\mathbf{A\hat{x}} = \operatorname{proj}_{\mathcal{O}(\mathbf{A})} \mathbf{b},$$

azaz $\mathbf{\hat{x}}$ optimális megoldás.

Végül meg kell mutatnunk, hogy a megoldások közt egyetlen van, mely $\mathbf{A}$ sorterébe esik. Ez azonnal következik abból, hogy a normálegyenlet megoldásai közt egyetlen egy van, mely $\mathbf{A}'\mathbf{A}$ sorterébe esik, az $\mathbf{A}'\mathbf{A}$ és $\mathbf{A}$ sorterei pedig megegyeznek. □

*A pszeudoinverz fogalma** A mátrix inverzének olyan – pszeudoinverznek nevezett – általánosítását keressük, mely képes lesz bármely $\mathbf{Ax} = \mathbf{b}$ egyenletrendszerből a minimális abszolút értékű optimális megoldást a mátrix inverzéhez hasonló módon megadni. Az $A : \mathbf{x} \mapsto \mathbf{Ax}$ mátrixleképezés a sorteret az oszloptérbe viszi kölcsönösen egyértelmű módon, merőleges kiegészítő alterét pedig a nullvektorba. Természetes gondolat, hogy $\mathbf{A}$ pszeudoinverzén annak a leképezésnek a mátrixát értsük, mely az oszlopteret inverz módon a sortérbe képzi, merőleges kiegészítő alterét pedig a nullvektorba.

7.45. DEFINÍCIÓ (A MOORE–PENROSE-FÉLE PSZEUDOINVERZ). *Legyen* $\mathbf{A}$ *egy* $m \times n$*-es valós mátrix.* Pszeudoinverzén *vagy* Moore–Penrose-féle pszeudoinverzén *azt az* $\mathbf{A}^+$*-szal jelölt mátrixot értjük, amellyel a sortér minden* $\mathbf{x}$ *vektorára* $\mathbf{A}^+(\mathbf{A}\mathbf{x}) = \mathbf{x}$, *továbbá az oszloptérre merőleges minden* $\mathbf{z}$ *vektorra* $\mathbf{A}^+\mathbf{z} = \mathbf{0}$.

► Azonnal látható, hogy $m \times n$-es mátrix pszeudoinverze $n \times m$-es.
► Mivel $\mathbf{A}\mathbf{x} \in \mathcal{O}(\mathbf{A})$, és definíció szerint $\mathbf{z} \in \mathcal{O}(\mathbf{A})^\perp$, ezért az $\mathbf{A}^+$-hoz tartozó leképezés hatását ismerjük az $\mathcal{O}(\mathbf{A})$ altéren és merőleges kiegészítő alterén, így ismerjük az egész téren is. Be fogjuk látni, hogy a definícióbeli feltételeket kielégítő mátrix létezik és egyértelmű, tehát a definíció értelmes.

7.46. PÉLDA (NÉHÁNY PSZEUDOINVERZ). *A definíció alapján igazoljuk az alábbi összefüggéseket!*
a) $\mathbf{A}^+ = \mathbf{A}^{-1}$, *ha* $\mathbf{A}$ *invertálható,*
b) $\mathbf{O}^+_{m\times n} = \mathbf{O}_{n\times m}$,
c) $[a]^+ = [1/a]$, *ha* $a \neq 0$, *és* $[0]^+ = [0]$,
d) $(\mathbf{A}^+)^+ = \mathbf{A}$,
e) *ha* $a_{ii} \neq 0$ $(i = 1, 2, \ldots, r)$, *akkor*

$$\left[\begin{array}{cccc|c} a_{11} & 0 & \dots & 0 & \\ 0 & a_{22} & \dots & 0 & \\ \vdots & \vdots & \ddots & \vdots & \mathbf{O} \\ 0 & 0 & \dots & a_{rr} & \\ \hline & & \mathbf{O} & & \mathbf{O} \end{array}\right]^+_{m\times n} = \left[\begin{array}{cccc|c} \frac{1}{a_{11}} & 0 & \dots & 0 & \\ 0 & \frac{1}{a_{22}} & \dots & 0 & \\ \vdots & \vdots & \ddots & \vdots & \mathbf{O} \\ 0 & 0 & \dots & \frac{1}{a_{rr}} & \\ \hline & & \mathbf{O} & & \mathbf{O} \end{array}\right]_{n\times m}$$

MEGOLDÁS. *a)* Ha $\mathbf{A}$ egy $n \times n$-es méretű invertálható mátrix, akkor sortere és oszloptere is a teljes n-dimenziós tér, és tetszőleges $\mathbf{x}$ vektorra $\mathbf{A}^{-1}\mathbf{y} = \mathbf{x}$ pontosan akkor áll fenn, ha $\mathbf{A}\mathbf{x} = \mathbf{y}$, azaz $\mathbf{A}^+ = \mathbf{A}^{-1}$.

b) Zérusmátrix oszloptere a zérusvektorból áll, így pszeudoinverze annak merőleges kiegészítő alterét, vagyis az egész teret a nullvektorba viszi, tehát $\mathbf{O}^+_{m\times n} = \mathbf{O}_{n\times m}$.

c) Az előző két eredményből azonnal következik.

d) Ha $\mathbf{x} \in \mathcal{S}(\mathbf{A})$ és $\mathbf{y} = \mathbf{A}\mathbf{x}$, akkor $\mathbf{A}^+\mathbf{y} = \mathbf{x}$, és mivel $\mathbf{y} \in \mathcal{O}(\mathbf{A}) = \mathcal{S}(\mathbf{A}^+)$, ezért $(\mathbf{A}^+)^+\mathbf{x} = \mathbf{y}$. Ha pedig $\mathbf{z} \perp \mathcal{O}(\mathbf{A}^+)$, akkor $(\mathbf{A}^+)^+\mathbf{z} = \mathbf{0}$, azaz az $\mathbf{A}$ és az $(\mathbf{A}^+)^+$ mátrixokhoz tartozó leképezések megegyeznek az $\mathcal{S}(\mathbf{A})$ és az $\mathcal{N}(\mathbf{A})$ altereken, így az általuk kifeszített téren, azaz $\mathbb{R}^n$-en is. Tehát $\mathbf{A} = (\mathbf{A}^+)^+$.

e) Jelölje $\mathbb{R}^n$ standard bázisának elemeit $\mathbf{e}_i$ $(i = 1, 2, \ldots, n)$, $\mathbb{R}^m$ standard bázisának elemeit $\mathbf{f}_i$ $(i = 1, 2, \ldots, m)$. Ha $a_{ii} \neq 0$, akkor $\mathbf{A}\mathbf{e}_i = a_{ii}\mathbf{f}_i$, és $\mathbf{e}_i$ a sortérben, $\mathbf{f}_i$ az oszloptérben van. Ha $\mathbf{a}_{ii} = 0$, akkor $\mathbf{A}\mathbf{e}_i = \mathbf{0}$. Így $\mathbf{A}^+\mathbf{f}_i = 1/a_{ii}\mathbf{e}_i$ ha $a_{ii} \neq 0$ és $\mathbf{A}^+\mathbf{f}_i = \mathbf{0}$ egyébként. Ennek

E példa *a)* pontja mutatja, hogy a pszeudoinverz név nem igazán jó, hisz itt nem *álinverzről*, nem *hamis inverzről* van szó, hanem az inverz általánosításáról, tehát az *általánosított inverz* helyesebb kifejezés. Szokás ezt is használni, de a Moore–Penrose pszeudoinverz kifejezés sokkal elterjedtebb (angol nyelvű művekben is leginkább a pseudoinverse szót használják).

alapján $\mathbf{A}^+$ főátlójában $1/a_{ii}$ áll, ha $a_{ii} \neq 0$, egyebütt 0. Gondoljuk meg, miért elég megadni $\mathbf{A}^+$ hatását csak a báziselemeken? □

7.47. TÉTEL (PSZEUDOINVERZ LÉTEZÉSE ÉS EGYÉRTELMŰSÉGE). *Tetszőleges $\mathbf{A} \in \mathbb{R}^{m\times n}$ mátrixnak létezik pszeudoinverze és az egyértelmű.*

BIZONYÍTÁS. A létezés bizonyítható a pszeudoinverz megkonstruálása nélkül is, de mi inkább konstruktív bizonyítást adunk. Ha $\mathbf{A} \in \mathbb{R}^{m\times n}$, akkor megadunk egy olyan $\mathbf{A}^+ \in \mathbb{R}^{n\times m}$ mátrixot, melyre $\mathbf{A}^+(\mathbf{A}\mathbf{x}) = \mathbf{x}$, ha $\mathbf{x} \in \mathcal{S}(\mathbf{A})$, és $\mathbf{A}^+\mathbf{z} = \mathbf{0}$, ha $\mathbf{z} \perp \mathcal{O}(\mathbf{A})$ ($\mathbf{z} \in \mathcal{N}(\mathbf{A}')$).

Legyen $\mathbf{A} = \mathbf{B}\mathbf{R}$ az $\mathbf{A}$ bázisfelbontása, ahol $\mathbf{R} = \operatorname{rref}(\mathbf{A})$ egy $r \times n$-es mátrix, $\mathbf{B}$ az $\mathbf{A}$ főoszlopaiból álló $m \times r$-es mátrix, és $r = \operatorname{r}(\mathbf{A})$. $\mathbf{R}$ sorvektorai a sortér egy $\mathcal{R}$ bázisát adják. Jelölje egy tetszőleges $\mathbf{x}$ sortérbeli vektor e bázisra vonatkozó koordinátás alakját $[\mathbf{x}]_{\mathcal{R}}$, így $\mathbf{x} = \mathbf{R}'[\mathbf{x}]_{\mathcal{R}}$ és $\mathbf{y} = \mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{R}'[\mathbf{x}]_{\mathcal{R}}$. Olyan mátrixot keresünk tehát, melyre $\mathbf{A}^+\mathbf{y} = \mathbf{x}$, azaz $\mathbf{A}^+\mathbf{A}\mathbf{R}'[\mathbf{x}]_{\mathcal{R}} = \mathbf{R}'[\mathbf{x}]_{\mathcal{R}}$. Kihasználjuk, hogy $\mathbf{B}$ és $\mathbf{R}'$ teljes oszloprangú, így $\mathbf{B}'\mathbf{B}$ és $\mathbf{R}\mathbf{R}'$ invertálható. Kiindulva az $\mathbf{y} = \mathbf{A}\mathbf{x}$ egyenletből, annak mindkét oldalát $(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'$-tal, majd $\mathbf{R}'(\mathbf{R}\mathbf{R}')^{-1}$-zel szorozva kapjuk, hogy

$$
\begin{aligned}
\mathbf{y} &= \mathbf{A}\mathbf{R}'[\mathbf{x}]_{\mathcal{R}} = \mathbf{B}\mathbf{R}\mathbf{R}'[\mathbf{x}]_{\mathcal{R}}\\
(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\mathbf{y} &= (\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\mathbf{B}\mathbf{R}\mathbf{R}'[\mathbf{x}]_{\mathcal{R}} = \mathbf{R}\mathbf{R}'[\mathbf{x}]_{\mathcal{R}}\\
\mathbf{R}'(\mathbf{R}\mathbf{R}')^{-1}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\mathbf{y} &= \mathbf{R}'[\mathbf{x}]_{\mathcal{R}} = \mathbf{x}.
\end{aligned}
$$

Eszerint $\mathbf{A}^+ = \mathbf{R}'(\mathbf{R}\mathbf{R}')^{-1}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'$ jó jelölt. Megmutatjuk, hogy az így definiált $\mathbf{A}^+$ mátrix eleget tesz a pszeudoinverz definíciójának. Ha $\mathbf{x} \in \mathcal{S}(\mathbf{A})$, és $\mathbf{y} = \mathbf{A}\mathbf{x}$, akkor épp $\mathbf{A}^+$ származtatásából adódóan $\mathbf{A}^+\mathbf{y} = \mathbf{x}$. Legyen tehát $\mathbf{z} \perp \mathcal{O}(\mathbf{A})$, azaz $\mathbf{z} \in \mathcal{N}(\mathbf{A}')$. Ekkor, mivel $\mathbf{B}$ oszlopai kifeszítik az $\mathcal{O}(\mathbf{A})$ alteret, $\mathbf{B}'\mathbf{z} = \mathbf{0}$. Így

$$
\mathbf{A}^+\mathbf{z} = \mathbf{R}'(\mathbf{R}\mathbf{R}')^{-1}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\mathbf{z} = \mathbf{0}.
$$

Képezzük az $m \times m$-es $\mathbf{C} = [\mathbf{A}\mathbf{R}'|\mathbf{Z}]$ mátrixot, ahol $\mathbf{Z}$ oszlopvektorai alkossák $\mathcal{N}(\mathbf{A}')$ egy bázisát. Mivel $\mathbf{A}\mathbf{R}'$ oszlopai az $\mathcal{O}(\mathbf{A})$-ban bázist alkotnak, ezért $\mathbf{C}$ oszlopai függetlenek, így $\mathbf{C}$ invertálható. A pszeudoinverz definíciójának megfelelő bármely $\mathbf{X}$ mátrixra $\mathbf{X}\mathbf{C} = [\mathbf{R}'|\mathbf{O}]$, így $\mathbf{A}^+\mathbf{C} = \mathbf{X}\mathbf{C}$, amiből $\mathbf{X} = \mathbf{A}^+$. Ezzel az egyértelműséget is igazoltuk. □

7.48. TÉTEL (A PSZEUDOINVERZ KISZÁMÍTÁSA). *Legyen $\mathbf{A}$ egy valós mátrix. Ekkor*

$$
\begin{aligned}
\mathbf{A}^+ &= \mathbf{R}'(\mathbf{R}\mathbf{R}')^{-1}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}' && (7.13)\\
&= \mathbf{R}'(\mathbf{B}'\mathbf{A}\mathbf{R}')^{-1}\mathbf{B}' && (7.14)
\end{aligned}
$$

ha $\mathbf{A} = \mathbf{B}\mathbf{R}$ az $\mathbf{A}$ mátrix bázisfelbontása. Ha $\mathbf{A}$ teljes oszloprangú, akkor

$$
\mathbf{A}^+ = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}', \qquad (7.15)
$$

ha $\mathbf{A}$ *teljes sorrangú, akkor*

$$\mathbf{A}^+ = \mathbf{A}'(\mathbf{A}\mathbf{A}')^{-1}. \tag{7.16}$$

BIZONYÍTÁS. Az $\mathbf{A}^+$ előállítását az előző tétel bizonyításában megadtuk. A kifejezés másik alakja a szorzat inverzére vonatkozó $(\mathbf{XY})^{-1} = \mathbf{Y}^{-1}\mathbf{X}^{-1}$ képletből adódik, mivel

$$(\mathbf{RR}')^{-1}(\mathbf{B}'\mathbf{B})^{-1} = (\mathbf{B}'\mathbf{BRR}')^{-1} = (\mathbf{B}'\mathbf{AR}')^{-1}.$$

Ha $\mathbf{A}$ teljes oszloprangú, akkor a bázisfelbontásban $\mathbf{B} = \mathbf{A}$ és $\mathbf{R} = \mathbf{I}_r$, így $\mathbf{A}^+ = \mathbf{I}_r'(\mathbf{I}_r\mathbf{I}_r')^{-1}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}' = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$. A teljes sorrangú esetre vonatkozó bizonyítást lásd a 7.18. feladatban. □

► A (7.15) képlet tökéletes összhangban van az egyenletrendszer optimális megoldásáról szóló 7.44. tétel állításával. Ott arra jutottunk, hogy az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszer optimális megoldásai megegyeznek az $\mathbf{A}'\mathbf{A}\hat{\mathbf{x}} = \mathbf{A}'\mathbf{b}$ egyenletrendszer megoldásaival. Ha pedig $\mathbf{A}$ teljes oszloprangú, akkor $\mathbf{A}'\mathbf{A}$ invertálható, így az optimális megoldás $\hat{\mathbf{x}} = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{b}$, azaz a (7.15) képlet szerint $\hat{\mathbf{x}} = \mathbf{A}^+\mathbf{b}$, ahogy azt célul tűztük ki e paragrafus elején.

► A (7.13) képlet nagyon bonyolultnak néz ki, de a (7.15) és a (7.16) képletek könnyen megjegyezhetővé teszik. Ha $\mathbf{A} = \mathbf{BR}$ a bázisfelbontás, akkor $\mathbf{B}$ teljes oszloprangú, $\mathbf{R}$ teljes sorrangú, így a (7.15) és a (7.16) képletek szerint $\mathbf{R}^+ = \mathbf{R}'(\mathbf{RR}')^{-1}$, $\mathbf{B}^+ = (\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'$, tehát $\mathbf{A}^+ = \mathbf{R}^+\mathbf{B}^+$. Általában nem igaz a pszeudoinverzre az $(\mathbf{XY})^+ = \mathbf{Y}^+\mathbf{X}^+$ összefüggés, de ha $\mathbf{X}$ teljes oszloprangú, és $\mathbf{Y}$ teljes sorrangú, akkor igen (ld. a 7.24. feladatot).

► A bizonyításban a bázisfelbontásból csak annyit használtunk, hogy $\mathbf{B}$ teljes oszloprangú, $\mathbf{R}$ teljes sorrangú. Valóban, a tétel tetszőleges olyan felbontásra igaz, mely e feltételt teljesíti (ld. a 7.18.–7.21. feladatokat).

7.49. PÉLDA (A PSZEUDOINVERZ KISZÁMÍTÁSA). *Számítsuk ki az*

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix} \quad \textit{és a} \quad \mathbf{C} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}$$

mátrixok pszeudoinverzét.

MEGOLDÁS. Az $\mathbf{A}$ mátrix pszeudoinverze a 7.48. tétel *a)* pontja szerinti képlettel számolható. Mivel $\mathbf{R} = \mathrm{rref}(\mathbf{A}) = \left[\begin{smallmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{smallmatrix}\right]$, és $\mathbf{B}$ az $\mathbf{A}$ első két oszlopát tartalmazza, ezért

$$\begin{aligned}\mathbf{A}^+ &= \mathbf{R}'(\mathbf{B}'\mathbf{A}\mathbf{R}')^{-1}\mathbf{B}'\\ &= \begin{bmatrix}1&0\\0&1\\1&1\end{bmatrix}\left(\begin{bmatrix}0&1&1\\1&1&0\end{bmatrix}\begin{bmatrix}0&1&1\\1&1&2\\1&0&1\end{bmatrix}\begin{bmatrix}1&0\\0&1\\1&1\end{bmatrix}\right)^{-1}\begin{bmatrix}0&1&1\\1&1&0\end{bmatrix}\\ &= \frac{1}{9}\begin{bmatrix}-4&1&5\\5&1&-4\\1&2&1\end{bmatrix}\end{aligned}$$

Mivel $\mathbf{C}$ teljes oszloprangú, ezért a 7.48. tétel *b)* pontja szerint:

$$\begin{aligned}\mathbf{C}^+ = (\mathbf{C}'\mathbf{C})^{-1}\mathbf{C}' &= \begin{bmatrix}2&1\\1&2\end{bmatrix}^{-1}\begin{bmatrix}0&1&1\\1&1&0\end{bmatrix}\\ &= \begin{bmatrix}2/3&-1/3\\-1/3&2/3\end{bmatrix}\begin{bmatrix}0&1&1\\1&1&0\end{bmatrix} = \begin{bmatrix}-1/3&1/3&2/3\\2/3&1/3&-1/3\end{bmatrix}\end{aligned}$$

□

*A pszeudoinverz tulajdonságai** Az $\mathbf{A}$ mátrix inverzét az $\mathbf{AX} = \mathbf{I}$ egyenlőséggel definiáltuk. Hasonló egyenlőségeket keresünk a pszeudoinverzhez is. Közben azt a fontos tényt is fölfedezzük, hogy $\mathbf{A}^+\mathbf{A}$ és $\mathbf{AA}^+$ is egy-egy merőleges vetítés mátrixa.

Azt nem tudjuk garantálni, hogy az $\mathbf{AA}^+$ és az $\mathbf{A}^+\mathbf{A}$ mátrixok az egységmátrixszal legyenek egyenlők, de legalább szimmetrikusak, és az $\mathbf{A}$-val, illetve az $\mathbf{A}^+$-szal való szorzásra nézve egységmátrixként viselkednek, azaz $\mathbf{AA}^+\mathbf{A} = \mathbf{A}$ és $\mathbf{A}^+\mathbf{AA}^+ = \mathbf{A}^+$. E feltételek már elegendők lesznek a pszeudoinverz algebrai leírásához.

7.50. tétel (Penrose-tétel). *A valós $\mathbf{A}$ mátrixnak $\mathbf{X}$ pontosan akkor pszeudoinverze, ha az alábbi négy feltétel mindegyike fennáll:*

$$a)\ \mathbf{AXA} = \mathbf{A},\quad b)\ \mathbf{XAX} = \mathbf{X},\quad c)\ (\mathbf{AX})' = \mathbf{AX},\quad d)\ (\mathbf{XA})' = \mathbf{XA}.$$

Bizonyítás. Azt, hogy $\mathbf{A}$ pszeudoinverze teljesíti e négy feltételt, egyszerű behelyettesítéssel ellenőrizhetjük.

$$\begin{aligned}\mathbf{AA}^+\mathbf{A} &= \mathbf{AR}'(\mathbf{RR}')^{-1}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\mathbf{A}\\ &= \mathbf{BRR}'(\mathbf{RR}')^{-1}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\mathbf{BR} = \mathbf{BR} = \mathbf{A}\\ \mathbf{A}^+\mathbf{AA}^+ &= \mathbf{R}'(\mathbf{RR}')^{-1}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\mathbf{BRR}'(\mathbf{RR}')^{-1}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\\ &= \mathbf{R}'(\mathbf{RR}')^{-1}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}' = \mathbf{A}^+\end{aligned}$$

A *c)* és az *d)* ellenőrzéséhez egyszerűsítsük az $\mathbf{A}^+\mathbf{A}$ és $\mathbf{AA}^+$ kifejezéseket:

$$\mathbf{A}^+\mathbf{A} = (\mathbf{R}'(\mathbf{RR}')^{-1}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}')(\mathbf{BR}) = \mathbf{R}'(\mathbf{RR}')^{-1}\mathbf{R}, \tag{7.17}$$

$$\mathbf{AA}^+ = (\mathbf{BR})(\mathbf{R}'(\mathbf{RR}')^{-1}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}') = \mathbf{B}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'. \tag{7.18}$$

Ezeket fölhasználva kapjuk, hogy

$$\begin{aligned}(\mathbf{A}^+\mathbf{A})' &= (\mathbf{R}'(\mathbf{R}\mathbf{R}')^{-1}\mathbf{R})' = \mathbf{R}'(\mathbf{R}\mathbf{R}')^{-1}\mathbf{R} = \mathbf{A}^+\mathbf{A}\\ (\mathbf{A}\mathbf{A}^+)' &= (\mathbf{B}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}')' = \mathbf{B}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}' = \mathbf{A}\mathbf{A}^+,\end{aligned}$$

ami bizonyítja az *c)* és az *d)* egyenlőségeket. Már csak azt kell bizonyítani, hogy ezeket az összefüggéseket legföljebb csak egy mátrix teljesíti. Tegyük fel, hogy $\mathbf{X}$ és $\mathbf{Y}$ is teljesíti a négy feltételt. Ekkor

$$\begin{aligned}\mathbf{AY} &\overset{a)}{=} \mathbf{AXAY} \overset{c)}{=} (\mathbf{AX})'(\mathbf{AY})' = \mathbf{X}'\mathbf{A}'\mathbf{Y}'\mathbf{A}'\\ &= \mathbf{X}'(\mathbf{AYA})' \overset{a)}{=} \mathbf{X}'\mathbf{A}' = (\mathbf{AX})' \overset{c)}{=} \mathbf{AX} \qquad (7.19)\\ \mathbf{YA} &\overset{a)}{=} \mathbf{YAXA} \overset{d)}{=} (\mathbf{YA})'(\mathbf{XA})' = \mathbf{A}'\mathbf{Y}'\mathbf{A}'\mathbf{X}'\\ &= (\mathbf{AYA})'\mathbf{X}' \overset{a)}{=} \mathbf{A}'\mathbf{X}' = (\mathbf{XA})' \overset{d)}{=} \mathbf{XA} \qquad (7.20)\\ \mathbf{Y} &\overset{b)}{=} \mathbf{YAY} \overset{(7.19)}{=} \mathbf{YAX} \overset{(7.20)}{=} \mathbf{XAX} \overset{b)}{=} \mathbf{X}.\end{aligned}$$

Ezzel bizonyítottuk a tételt. □

7.51. KÖVETKEZMÉNY ($\mathbf{A}^+\mathbf{A}$ ÉS $\mathbf{A}\mathbf{A}^+$ MERŐLEGES VETÍTÉS). *Tetszőleges $\mathbf{A} \in \mathbb{R}^{m\times n}$ mátrix esetén*

$$\mathbf{A}^+\mathbf{A} = \operatorname{proj}_{\mathcal{S}(\mathbf{A})} \quad \textit{és} \quad \mathbf{A}\mathbf{A}^+ = \operatorname{proj}_{\mathcal{O}(\mathbf{A})}.$$

Tehát $\mathbf{A}^+\mathbf{A}$ az $\mathbb{R}^n$ teret merőlegesen vetíti $\mathbf{A}$ sorterére, míg $\mathbf{A}\mathbf{A}^+$ az $\mathbb{R}^m$ teret merőleges vetíti $\mathbf{A}$ oszlopterére.

BIZONYÍTÁS. A (7.17) egyenlőség szerint $\mathbf{A}^+\mathbf{A} = \mathbf{R}'(\mathbf{R}\mathbf{R}')^{-1}\mathbf{R}$, ami az altérre való merőleges vetítés mátrixáról szóló 7.37. tétel szerint az $\mathbf{R}'$ oszlopvektorai által kifeszített térre – azaz a sortérre – való merőleges vetítés mátrixa. Hasonlóképp a (7.18) egyenlet szerint $\mathbf{A}\mathbf{A}^+ = \mathbf{B}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'$, ami a $\mathbf{B}$ oszlopvektorai által kifeszített térre – azaz az oszloptérre – való merőleges vetítés mátrixa. □

*A pszeudoinverz és a minimális abszolút értékű optimális megoldás** Megmutatjuk, hogy $\mathbf{A}^+$ úgy használható egy tetszőleges együtthatómátrixú $\mathbf{Ax} = \mathbf{b}$ egyenletrendszer optimális megoldásának meghatározására, ahogy $\mathbf{A}^{-1}$ használható akkor, ha $\mathbf{A}$ invertálható.

7.52. TÉTEL (OPTIMÁLIS MEGOLDÁS PSZEUDOINVERZZEL). *Legyen $\mathbf{A}$ egy valós mátrix. Az $\mathbf{Ax} = \mathbf{b}$ egyenletrendszernek $\hat{\mathbf{x}} = \mathbf{A}^+\mathbf{b}$ a minimális abszolút értékű optimális megoldása.*

BIZONYÍTÁS. Először megmutatjuk, hogy $\mathbf{A}^+\mathbf{b}$ optimális megoldás, azaz megoldása az $\mathbf{A}'\mathbf{Ax} = \mathbf{A}'\mathbf{b}$ normálegyenlet-rendszernek. Tehát igazolni kell, hogy $\mathbf{A}'\mathbf{A}\mathbf{A}^+\mathbf{b} = \mathbf{A}'\mathbf{b}$. Ehhez elég belátni, hogy $\mathbf{A}'\mathbf{A}\mathbf{A}^+ =$

$\mathbf{A}'$. Legyen $\mathbf{A} = \mathbf{BR}$ az $\mathbf{A}$ mátrix bázisfelbontása. Ekkor

$$\begin{aligned}\mathbf{A}'\mathbf{A}\mathbf{A}^+ &= (\mathbf{R}'\mathbf{B}')(\mathbf{B}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}')\\ &= \mathbf{R}'(\mathbf{B}'\mathbf{B})(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}' = \mathbf{R}'\mathbf{B}' = \mathbf{A}'\end{aligned}$$

Mivel $\mathbf{A}^+\mathbf{b}$ a definícióból következőleg a sortérben van, és a sortérbe csak egyetlen optimális megoldás – a minimális abszolút értékű megoldás – esik, ezért $\mathbf{A}^+\mathbf{b}$ valóban a minimális abszolút értékű optimális megoldás. □

A gyakorlatban gyakran előfordul, hogy mért adatokból kell bizonyos változók értékét meghatározni. Ha az n ismeretlen értékre n mérést végzünk, az egyenletrendszer egyértelműen megoldható, ha azonban a mérési hibákat kiküszöbölendő több mérést végzünk, az egyenletrendszer már ellentmondásossá válik. Ilyen esetet mutat a következő példa.

7.53. PÉLDA (EGYENLETRENDSZER OPTIMÁLIS MEGOLDÁSA). *Az alábbi háromismeretlenes egyenletrendszer négy egyenletből áll:*

$$\begin{aligned} x + 3y + 6z &= 8\\ x - y + 2z &= 2\\ x + 3y + 2z &= 2\\ x - y - 2z &= 0\end{aligned}$$

Bármelyik három egyértelműen megoldható egyenletrendszert ad, de a négy együtt ellentmondásos. Határozzuk meg az optimális megoldását!

MEGOLDÁS. A 7.52. tétel szerint az optimális megoldás $\mathbf{A}^+\mathbf{b}$, ahol

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 6\\ 1 & -1 & 2\\ 1 & 3 & 2\\ 1 & -1 & -2\end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 8\\ 2\\ 2\\ 0\end{bmatrix}$$

Mivel $\mathbf{A}$ teljes oszloprangú, ezért csak egyetlen optimális megoldása van, másrészt pszeudoinverze a (7.15) képlettel számolható, így

$$\mathbf{A}^+ = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}' = \begin{bmatrix} 0 & 1/4 & 1/4 & 1/2\\ 0 & -1/4 & 1/4 & 0\\ 1/8 & 1/8 & -1/8 & -1/8\end{bmatrix} \quad \mathbf{A}^+\mathbf{b} = \begin{bmatrix} 1\\ 0\\ 1\end{bmatrix}$$

Így az egyenletrendszer optimális megoldása $(1, 0, 1)$. □

A következő példa olyan egyenletrendszert vizsgál, melyben az együtthatómátrix rangja kisebb mind az egyenletek, mind az ismeretlenek számánál.

7.54. PÉLDA (EGYENLETRENDSZER OPTIMÁLIS MEGOLDÁSA). *Oldjuk meg az*

$$\begin{aligned} y + z &= 3 \\ x + y + 2z &= 2 \\ x + y &= 2 \end{aligned}$$

egyenletrendszert. Ha nem oldható meg, akkor határozzuk meg minimális abszolút értékű optimális megoldását!

MEGOLDÁS. Az egyenletrendszer nem oldható meg, ami bővített mátrixának redukált lépcsős alakjából leolvasható:

$$\left[\begin{array}{ccc|c} 0 & 1 & 1 & 3 \\ 1 & 1 & 2 & 2 \\ 1 & 1 & 0 & 2 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

Az együtthatómátrix pszeudoinverzét meghatároztuk a 7.49. példában. Ezt felhasználva a minimális abszolút értékű optimális megoldás

$$\hat{\mathbf{x}} = \mathbf{A}^+ \mathbf{b} = \frac{1}{9} \begin{bmatrix} -4 & 1 & 5 \\ 5 & 1 & -4 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

□

Lineáris és polinomiális regresszió Az egyenletrendszerek optimális megoldásainak egyik fontos alkalmazása a lineáris regresszió. Tegyük fel, hogy két változó mennyiség, az x és az y között az $y = a + bx$ kapcsolat van. Méréseket végzünk, melyek eredménye az (x_i, y_i) $(i = 1, 2, \dots n)$ párok sorozata. Keressük az a és b értékét, mely kielégíti az $y_i = a + bx_i$ $(i = 1, 2, \dots n)$ egyenletek mindegyikét! Ez egy kétismeretlenes lineáris egyenletrendszer, melynek mátrixalakja:

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

A hozzá tartozó normálegyenlet-rendszer

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},$$

amely a mátrixműveletek elvégzése után a következő alakra vezet:

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

Ennek $\hat{a}$ és $\hat{b}$ megoldása adja az eredeti egyenletrendszer optimális megoldását! Az ilyen módon kapott $y = \hat{a} + \hat{b}x$ egyenest *regressziós egyenesnek* nevezzük, mely a megadott adatokra a legkisebb négyzetek elve szerinti legjobban illeszkedő egyenes.

Összefoglalva:

7.55. ÁLLÍTÁS (LINEÁRIS REGRESSZIÓ). *Az (x_i, y_i) $(i = 1, 2, \ldots n)$ párokhoz tartozó, $y = \hat{a} + \hat{b}x$ egyenletű regressziós egyenes paraméterei kielégítik az*

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

egyenletet. Ez egyértelműen megoldható, ha van legalább két különböző x_i érték.

BIZONYÍTÁS. Az összefüggést már fent igazoltuk, csak a egyértelmű megoldhatóság igazolása maradt hátra. A számtani és négyzetes közép közti összefüggés szerint bármely x_i $(i = 1, 2 \ldots, n)$ valósokra

$$\frac{x_1 + \cdots + x_n}{n} \leq \sqrt{\frac{x_1^2 + \cdots + x_n^2}{n}},$$

és egyenlőség csak akkor állhat fenn, ha $x_1 = \cdots = x_n$. Mivel az együtthatómátrix determinánsa $n \sum x_i^2 - (\sum x_i)^2$, ezért a számtani és négyzetes közép közti összefüggés miatt ez csak akkor lehet 0, ha az x_i értékek mind azonosak. □

A lineáris regresszió gyakran egyéb függvénykapcsolat esetén is alkalmazható:

7.56. ÁLLÍTÁS (LINEARIZÁLHATÓ REGRESSZIÓS MODELLEK). *Ha az x és y mennyiségek között az alábbi táblázat szerinti függvénykapcsolatok valamelyike áll, akkor a táblázatban megadott helyettesítéssel a kapcsolat $Y = a + bX$ alakúvá, azaz lineárissá válik, így lineáris regresszió végezhető.*

Modell	*Függvénykapcsolat*	*Helyettesítés*		
hatványfüggvény	$y = cx^b$	$X = \ln x$	$Y = \ln y$	$a = \ln c$
exponenciális	$y = ce^{bx}$	$X = x$	$Y = \ln y$	$a = \ln c$
logaritmikus	$y = a + b \ln x$	$X = \ln x$	$Y = y$	

BIZONYÍTÁS. Az $y = cx^b$ egyenlőség mindkét oldalának logaritmusát véve az $\ln y = \ln c + b \ln x$ egyenlőséget kapjuk, ami a megadott helyettesítésekkel az $Y = a + bX$ kifejezést adja. Ugyanígy, az $y = ce^{bx}$ egyenlet logaritmusát véve az $\ln y = \ln c + bx$ egyenletet kapjuk. A szükséges helyettesítés a harmadik esetben még nyilvánvalóbb. □

A regresszió hasonló módon más függvényekkel is végezhető, ezek közül a polinomiálist emeljük ki:

7.57. PÉLDA. *Keressünk az $a_0 + a_1 x + \cdots + a_k x^k$ polinom együtthatóira optimális becslést a legkisebb négyzetek módszerével, ha az (x_i, y_i) ($i = 1, 2, \dots n$) párok sorozatát ismerjük.*

MEGOLDÁS. Keresendő az n egyenletből álló $k+1$-ismeretlenes

$$\begin{aligned} a_0 + a_1 x_1 + \ldots + a_k x_1^k &= y_1 \\ a_0 + a_1 x_2 + \ldots + a_k x_2^k &= y_2 \\ \vdots \quad\quad \vdots \quad\quad\quad\quad \vdots \quad &\quad \vdots \\ a_0 + a_1 x_n + \ldots + a_k x_n^k &= y_n \end{aligned}$$

egyenletrendszer megoldása az a_0, a_1,..., a_m ismeretlenekre. Mátrix-alakja

$$\begin{bmatrix} 1 & x_1 & \dots & x_1^k \\ 1 & x_2 & \dots & x_2^k \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \dots & x_n^k \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Ha az együtthatómátrixot $\mathbf{X}$ jelöli, az ismeretlenek vektorát $\mathbf{a}$, az y_i értékek vektorát $\mathbf{y}$, akkor az egyenletrendszer az $\mathbf{Xa} = \mathbf{y}$ alakba írható. Ez biztosan megoldható, mégpedig egyértelműen, ha az x_i értékek különbözőek, és $n = k+1$, ekkor ugyanis az együtthatómátrix négyzetes, determinánsa Vandermonde-determináns, melynek értéke nem 0. Egyéb esetekben a normálegyenlet-rendszert kell felírni, melynek mátrixszorzatos alakja

$$\mathbf{X}'\mathbf{Xa} = \mathbf{X}'\mathbf{y}.$$

Ez egyértelműen megoldható, ha $\mathbf{X}$ teljes oszloprangú, mert akkor $\mathbf{X}'\mathbf{X}$ invertálható, így a megoldás

$$\mathbf{a} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}.$$

Ez pontosan akkor áll fenn, ha van legalább $k+1$ különböző x_i érték, ekkor ugyanis $\mathbf{X}$-ben van egy $(k+1) \times (k+1)$-es nemnulla determinánsú részmátrix, nevezetesen a különböző x_i értékekhez tartozó sorokból álló Vandermonde-mátrix. □

Feladatok

7.8. Igazoljuk a síkbeli egyenesre merőlegesen vetítő mátrix (7.11) képletét az **i** és **j** vektorok vetületének meghatározásával! Kétféleképp is számolhatunk, *a)* legyen az egyenes hajlásszöge az x-tengellyel α, *b)* legyen az egyenes irányvektora (b_1, b_2). Vessük össze a két eredményt, és igazoljuk azonosságukat.

Pszeudoinverz

Számítsuk ki az alábbi mátrixok pszeudoinverzét!

7.9.• $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$

7.10.• $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

7.11. $\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$

7.12. $\begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix}$

7.13. $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$

7.14. $\begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 0 & 0 \end{bmatrix}$

7.15. $\begin{bmatrix} 1 & 0 \\ 2 & 2 \\ 0 & 1 \end{bmatrix}$

7.16. $\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$

7.17. $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$

7.18. Legyen **A** egy tetszőleges r-rangú valós $m \times n$-es mátrix, és legyen $\mathbf{A} = \mathbf{XY}$ egy olyan felbontás, melyben $\mathbf{X} \in \mathbb{R}^{m\times r}$, $\mathbf{Y} \in \mathbb{R}^{r\times n}$, másként fogalmazva legyen **X** teljes oszloprangú, **Y** teljes sorrangú. Mutassuk meg, hogy ekkor

$$\mathbf{A}^+ = \mathbf{Y}'(\mathbf{YY}')^{-1}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \tag{7.21}$$

$$= \mathbf{Y}'(\mathbf{X}'\mathbf{AY}')^{-1}\mathbf{X}' \tag{7.22}$$

Ezt felhasználva igazoljuk, hogy ha **A** teljes oszloprangú, akkor

$$\mathbf{A}^+ = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}',$$

ha **A** teljes sorrangú, akkor

$$\mathbf{A}^+ = \mathbf{A}'(\mathbf{AA}')^{-1}.$$

Az alábbi feladatokban határozzuk meg a felbontásaikkal megadott mátrixok pszeudoinverzét!

7.19. $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$

7.20. $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$

7.21. $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}$

7.22. Mutassuk meg, hogy ha **A** egy merőleges vetítés mátrixa, azaz $\mathbf{A}' = \mathbf{A} = \mathbf{A}^2$, akkor $\mathbf{A}^+ = \mathbf{A}$. Igaz-e az állítás megfordítása?

7.23.• Mutassuk meg, hogy ha $\mathbf{A} \in \mathbb{R}^{m\times n}$ és **A** teljes oszloprangú, akkor $\mathbf{A}^+\mathbf{A} = \mathbf{I}_m$, ha pedig **A** teljes sorrangú, akkor $\mathbf{AA}^+ = \mathbf{I}_n$.

7.24. Tegyük fel, hogy **A** teljes oszloprangú, **C** teljes sorrangú. Mutassuk meg, hogy ekkor $(\mathbf{AC})^+ = \mathbf{C}^+\mathbf{A}^+$.

7.25. 1-RANGÚ MÁTRIXOK PSZEUDOINVERZE Mutassuk meg, hogy ha $\mathrm{r}(\mathbf{A}) = 1$, akkor

$$\mathbf{A}^+ = \frac{1}{\operatorname{trace}(\mathbf{A}'\mathbf{A})}\mathbf{A}',$$

ahol $\operatorname{trace}(\mathbf{A}^+\mathbf{A})$ az **A** elemeinek négyzetösszege. Eszerint ha $\mathbf{a} \neq \mathbf{0}$, akkor

$$\mathbf{a}^+ = \frac{1}{\mathbf{a}'\mathbf{a}}\mathbf{a}' = \frac{1}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}'.$$

E feladat eredményét fölhasználva ellenőrizzük a 7.9., 7.10., 7.11., 7.12., 7.14., 7.16., 7.17. feladatok eredményeit!

7.26. BLOKKDIAGONÁLIS MÁTRIX PSZEUDOINVERZE Igazoljuk, hogy blokkdiagonális mátrix esetén

$$\begin{bmatrix} \mathbf{A}_1 & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \mathbf{A}_2 & \dots & \mathbf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \dots & \mathbf{A}_k \end{bmatrix}^+ = \begin{bmatrix} \mathbf{A}_1^+ & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \mathbf{A}_2^+ & \dots & \mathbf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \dots & \mathbf{A}_k^+ \end{bmatrix}.$$

7.27. Számítsuk ki a

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

mátrix pszeudoinverzét!

Ortonormált bázis, ortogonális mátrixok

Nem kell indokolni a merőlegesség fontosságát bizonyos természeti jelenségek leírásában. A lineáris algebrában is nélkülözhetetlen a fogalma. Egy altér ortonormált bázisának megkonstruálása, és azoknak a leképezéseknek az áttekintése, melyek ortonormált bázist ortonormáltba visznek alapvetően fontosak.

Ortogonális és ortonormált bázis Segíti az alterek vizsgálatát, ha a bázisvektorok páronként merőlegesek egymásra, ekkor ugyanis a különböző bázisvektorok skaláris szorzata 0. További könnyítést jelenthet, ha a bázisvektorok egységvektorok, mert ekkor egy vektor velük vett skaláris szorzata a merőleges vetület hosszát adja.

Páronként merőleges vektorok egy rendszerét *ortogonális rendszernek* nevezzük. Ortogonális rendszernek lehetnek 0-vektor tagjai. Páronként merőleges egységvektorok egy rendszerét *ortonormált rendszernek* nevezzük. Ortonormált rendszerben *nincsenek* 0-vektorok. A következő tételből azonnal adódik, hogy zérusvektort nem tartalmazó ortogonális rendszer vagy egy tetszőleges ortonormált rendszer mindig bázisa az általa kifeszített altérnek. Ezt az altér *ortogonális bázisának* (rövidítve OB), illetve *ortonormált bázisának* (rövidítve ONB) nevezzük. Ortogonális bázisból mindig kaphatunk egy ortonormáltat, ha az OB minden bázisvektorát elosztjuk a hosszával. Ezt a vektor *normálásának* nevezzük.

7.58. TÉTEL (ORTOGONÁLIS VEKTOROK FÜGGETLENSÉGE). *Ha a nullvektortól különböző $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ vektorok páronként ortogonálisak, akkor függetlenek is.*

BIZONYÍTÁS. Tekintsük a $c_1\mathbf{a}_1 + \cdots + c_k\mathbf{a}_k = \mathbf{0}$ egyenletet. Be kell látnunk, hogy ez csak a $c_1 = \cdots = c_k = 0$ esetben áll fönn. Szorozzuk be az egyenlőség mindkét oldalát az $\mathbf{a}_i$ vektorral ($i = 1, 2, \ldots, k$). Ekkor a jobb oldal 0, a bal oldalon pedig egy tag kivételével mindegyik 0 lesz:

$$\begin{aligned}(c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_k\mathbf{a}_k) \cdot \mathbf{a}_i &= \mathbf{0} \cdot \mathbf{a}_i \\ c_i\mathbf{a}_i \cdot \mathbf{a}_i &= 0.\end{aligned}$$

Mivel $\mathbf{a}_i \cdot \mathbf{a}_i \neq 0$, ezért $c_i = 0$, és ez igaz minden i-re. □

Tudjuk, hogy a háromdimenziós térben bármely $\mathbf{v} = (x, y, z)$ vektor koordinátáira igaz, hogy

$$x = \mathbf{v} \cdot \mathbf{i}, \quad y = \mathbf{v} \cdot \mathbf{j}, \quad z = \mathbf{v} \cdot \mathbf{k}.$$

Az is igaz, hogy az $\mathbf{i}$ és $\mathbf{j}$ által kifeszített síknak, azaz az xy-síknak a $\mathbf{v}$ vektorhoz legközelebb fekvő pontja, illetve az oda mutató helyvektor

$\hat{\mathbf{v}} = (x, y, 0)$, azaz
$$\hat{\mathbf{v}} = (\mathbf{v} \cdot \mathbf{i})\mathbf{i} + (\mathbf{v} \cdot \mathbf{j})\mathbf{j}$$
Azt is tudjuk, hogy a $\mathbf{v}$-hez legközelebbi pont épp $\mathbf{v}$-nek a síkra való merőleges vetülete.

A fenti nyilvánvaló összefüggések tetszőleges ONB esetén is használhatók, így igen értékesek.

7.59. TÉTEL (LEGJOBB KÖZELÍTÉS ONB ESETÉN). *Adva van az $\mathbb{R}^n$ térben egy $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k\}$ ortonormált rendszer által kifeszített $\mathcal{A}$ altér, valamint egy $\mathbf{v}$ vektor. Ekkor a*
$$\hat{\mathbf{v}} = (\mathbf{v} \cdot \mathbf{e}_1)\mathbf{e}_1 + (\mathbf{v} \cdot \mathbf{e}_2)\mathbf{e}_2 + \cdots + (\mathbf{v} \cdot \mathbf{e}_k)\mathbf{e}_k \tag{7.23}$$
vektor az $\mathcal{A}$ altér $\mathbf{v}$-hez legközelebb fekvő pontja, azaz $\hat{\mathbf{v}} = \operatorname{proj}_{\mathcal{A}} \mathbf{v}$.

BIZONYÍTÁS. Először megmutatjuk, hogy a (7.23) képlet szerinti pont van legközelebb $\mathbf{v}$-hez. $\mathbf{v}$ és $\hat{\mathbf{v}}$ távolságának négyzete
$$\begin{aligned}(\mathbf{v} - \hat{\mathbf{v}})^2 &= \left(\mathbf{v} - \sum_{i=1}^{k}(\mathbf{v} \cdot \mathbf{e}_i)\mathbf{e}_i\right)^2 \\ &= \mathbf{v}^2 - 2\sum_{i=1}^{k}(\mathbf{v} \cdot \mathbf{e}_i)^2 + \sum_{i=1}^{k}(\mathbf{v} \cdot \mathbf{e}_i)^2 \\ &= \mathbf{v}^2 - \sum_{i=1}^{k}(\mathbf{v} \cdot \mathbf{e}_i)^2.\end{aligned}$$
$\mathbf{v}$ és az altér egy tetszőleges $\mathbf{u}$ vektorának távolságnégyzete:
$$\begin{aligned}(\mathbf{v} - \mathbf{u})^2 &= \left(\mathbf{v} - \sum_{i=1}^{k} c_i\mathbf{e}_i\right)^2 \\ &= \mathbf{v}^2 - 2\sum_{i=1}^{k} c_i(\mathbf{v} \cdot \mathbf{e}_i) + \sum_{i=1}^{k} c_i^2.\end{aligned}$$
Ha az utóbbiból kivonva az előbbit pozitív értéket kapunk, ez azt jelenti, hogy valóban $\hat{\mathbf{v}}$ van $\mathbf{v}$-hez legközelebb:
$$\begin{aligned}&(\mathbf{v} - \mathbf{u})^2 - (\mathbf{v} - \hat{\mathbf{v}})^2 \\ &\quad= \left(\mathbf{v}^2 - 2\sum_{i=1}^{k} c_i(\mathbf{v} \cdot \mathbf{e}_i) + \sum_{i=1}^{k} c_i^2\right) - \left(\mathbf{v}^2 - \sum_{i=1}^{k}(\mathbf{v} \cdot \mathbf{e}_i)^2\right) \\ &\quad= \sum_{i=1}^{k} c_i^2 - 2\sum_{i=1}^{k} c_i(\mathbf{v} \cdot \mathbf{e}_i) + \sum_{i=1}^{k}(\mathbf{v} \cdot \mathbf{e}_i)^2 \\ &\quad= \sum_{i=1}^{k}(c_i - \mathbf{v} \cdot \mathbf{e}_i)^2 \geq 0.\end{aligned}$$
□

Ebből a legjobb közelítés tétele szerint következik, hogy $\hat{\mathbf{v}} = \operatorname{proj}_{\mathcal{A}} \mathbf{v}$.

7.60. PÉLDA (EGY PONT SÍKRA VALÓ MERŐLEGES VETÜLETE). *Határozzuk meg a $(3,1,2)$ pontnak az egymásra merőleges $\mathbf{a} = \frac{1}{7}(2,3,6)$ és $\mathbf{b} = \frac{1}{7}(3,-6,2)$ vektorok által kifeszített síkra való merőleges vetületét!*

MEGOLDÁS. Mivel $\mathbf{a}$ és $\mathbf{b}$ ortonormált bázisa az általuk kifeszített altérnek, ezért a $\mathbf{v} = (3,1,2)$ vektornak e síkra eső merőleges vetülete

$$\begin{aligned}
\hat{\mathbf{v}} &= (\mathbf{v}\cdot\mathbf{a})\mathbf{a} + (\mathbf{v}\cdot\mathbf{b})\mathbf{b}\\
&= \left((3,1,2)\cdot\left(\tfrac{2}{7},\tfrac{3}{7},\tfrac{6}{7}\right)\right)\left(\tfrac{2}{7},\tfrac{3}{7},\tfrac{6}{7}\right) + \left((3,1,2)\cdot\left(\tfrac{3}{7},\tfrac{-6}{7},\tfrac{2}{7}\right)\right)\left(\tfrac{3}{7},\tfrac{-6}{7},\tfrac{2}{7}\right)\\
&= 3\left(\tfrac{2}{7},\tfrac{3}{7},\tfrac{6}{7}\right) + 1\left(\tfrac{3}{7},\tfrac{-6}{7},\tfrac{2}{7}\right)\\
&= \left(\tfrac{9}{7},\tfrac{3}{7},\tfrac{20}{7}\right).
\end{aligned}$$

Összehasonlításul: a standard elemi módszer az $\mathbf{a}\times\mathbf{b}$ irányvektorú, $(3,1,2)$ ponton átmenő egyenes és a sík metszéspontjának meghatározása lenne. □

Ortogonális mátrixok Egy ortonormált vektorrendszerből képzett mátrixnak gyönyörű algebrai és geometriai tulajdonságai vannak.

7.61. DEFINÍCIÓ (ORTOGONÁLIS ÉS SZEMIORTOGONÁLIS MÁTRIX). *Egy valós négyzetes mátrixot* ortogonálisnak *nevezünk, ha oszlopvektorai ortonormált rendszert alkotnak. Ha nem kötjük ki, hogy a mátrix négyzetes legyen,* szemiortogonális *mátrixról beszélünk.*

► Ha egy $n\times k$ méretű $\mathbf{Q}$ mátrix szemiortogonális, akkor $k \le n$, ugyanis $\mathbf{Q}$ oszlopai ortonormált rendszert alkotnak, ezért lineárisan függetlenek, azaz $\mathbf{Q}$ teljes oszloprangú, így $\mathbf{Q}$ rangja k, tehát $k \le n$.
► Nagyon szerencsétlen az ortogonális mátrix elnevezés, de annyira el van terjedve, hogy nem lehet eltérni tőle. Nyilván jobb lenne az ortonormált mátrix elnevezés. A szemiortogonális az ortogonálisból származtatott név.

7.62. PÉLDA (ORTOGONÁLIS MÁTRIXOK). *Melyek ortogonálisak és melyek szemiortogonálisak az alábbi mátrixok közül?*

$$\mathbf{A} = \begin{bmatrix} 0&1&0&0\\ 0&0&0&1\\ 0&0&1&0\\ 1&0&0&0 \end{bmatrix},\ \mathbf{B} = \begin{bmatrix} \cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\\ 0&0\\ 0&0 \end{bmatrix},\ \mathbf{C} = \frac{1}{7}\begin{bmatrix} 2&3&6\\ 6&2&-3\\ 3&-6&2 \end{bmatrix}.$$

MEGOLDÁS. Mindhárom mátrix szemiortogonális, hisz oszlopvektorai ortonormált rendszert alkotnak (az $\mathbf{A}$ mátrix oszlopai $\mathbb{R}^4$ standard egységvektorai, a $\mathbf{B}$ mátrix oszlopvektorai az $\mathbb{R}^4$ első két standard egységvektorának xy-síkban α szöggel való elforgatásával kaphatók, a $\mathbf{C}$

mátrix esetén az oszlopvektorok skaláris szorzatainak elvégzésével ellenőrizhetjük ortonormalitásukat). A három mátrix közül csak $\mathbf{A}$ és $\mathbf{C}$ négyzetes, így csak ezek a mátrixok ortogonálisak. □

► E példa alapján könnyen látható, hogy minden permutációmátrix, így az egységmátrix is, ortogonális.

7.63. TÉTEL (SZEMIORTOGONÁLIS MÁTRIXOK EKVIVALENS DEFINÍCIÓI). *Legyen* $\mathbf{Q} \in \mathbb{R}^{m\times n}$. *Az alábbi állítások ekvivalensek:*
a) $\mathbf{Q}$ *szemiortogonális,*
b) $\mathbf{Q}'\mathbf{Q} = \mathbf{I}_n$.

BIZONYÍTÁS. *a)* $\Rightarrow$ *b)*: Legyen $\mathbf{Q} = [\mathbf{q}_1\ \mathbf{q}_2\ \dots\ \mathbf{q}_n]$. Ekkor $[\mathbf{Q}'\mathbf{Q}]_{ij} = \mathbf{q}_i'\mathbf{q}_j = \mathbf{q}_i \cdot \mathbf{q}_j$, de mivel a $\{\mathbf{q}_i\}$ vektorrendszer ortonormált, ezért $\mathbf{q}_i^2 = 1$ és $\mathbf{q}_i \cdot \mathbf{q}_j = 0$, ha $i \neq j$. Eszerint $[\mathbf{Q}'\mathbf{Q}]_{ii} = 1$, és $[\mathbf{Q}'\mathbf{Q}]_{ij} = 0$, ha $i \neq j$ és $i, j \leq k$, vagyis $\mathbf{Q}'\mathbf{Q} = \mathbf{I}_k$.

b) $\Rightarrow$ *a)*: A $\mathbf{Q}'\mathbf{Q} = \mathbf{I}_k$ összefüggésbeli mátrixszorzást sorvektorszor oszlopvektorként tekintve épp azt kapjuk, hogy $\mathbf{q}_i^2 = 1$ és $\mathbf{q}_i \cdot \mathbf{q}_j = 0$, ha $i \neq j$, azaz a $\{\mathbf{q}_i\}$ vektorrendszer ortonormált. □

► A *b)* állítás algebrai nyelven azt mondja, hogy $\mathbf{Q}$ pontosan akkor szemiortogonális, ha transzponáltja a bal oldali inverze. Mint mindjárt látni fogjuk, ortogonális mátrixokra ez egyúttal a jobb oldali inverz is, azaz a transzponált megegyezik $\mathbf{Q}$ inverzével.

7.64. TÉTEL (ORTOGONÁLIS MÁTRIXOK EKVIVALENS DEFINÍCIÓI). *Legyen* $\mathbf{Q} \in \mathbb{R}^{n\times n}$. *Az alábbi állítások ekvivalensek:*
a) $\mathbf{Q}$ *ortogonális, azaz oszlopvektorai ortonormált rendszert alkotnak.*
b) $\mathbf{Q}'\mathbf{Q} = \mathbf{I}_n$.
c) $\mathbf{Q}^{-1} = \mathbf{Q}'$.
d) $\mathbf{Q}\mathbf{Q}' = \mathbf{I}_n$.
e) $\mathbf{Q}'$ *ortogonális, azaz* $\mathbf{Q}$ *sorvektorai ortonormált rendszert alkotnak.*

BIZONYÍTÁS. Az *a)* $\Leftrightarrow$ *b)* ekvivalenciát az előző állításban bizonyítottuk.

b) $\Rightarrow$ *c)*: Mivel $\mathbf{Q}$ négyzetes, ezért a $\mathbf{Q}'\mathbf{Q} = \mathbf{I}$ összefüggés egyúttal azt is jelenti, hogy $\mathbf{Q}$ invertálható, tehát $\mathbf{Q}$ és $\mathbf{Q}'$ egymás inverzei, azaz $\mathbf{Q}^{-1} = \mathbf{Q}'$.

c) $\Rightarrow$ *d)*: Mivel $\mathbf{Q}^{-1} = \mathbf{Q}'$, ezért $\mathbf{Q}\mathbf{Q}' = \mathbf{I}_n$.

d) $\Rightarrow$ *e)*: A $\mathbf{Q}\mathbf{Q}' = \mathbf{I}_n$ egyenletben a mátrixszorzásra sovektorszoroszlopvektorként tekintve épp azt kapjuk, hogy $\mathbf{Q}'$ oszlopvektorai – és így $\mathbf{Q}$ soravektorai – ONB-t alkotnak.

e) $\Rightarrow$ *a)*: Az előzőekben beláttuk, hogy ha $\mathbf{Q}$ ortogonális, akkor $\mathbf{Q}'$ is, hisz bizonyítottuk, hogy *a)*-ból következik *e)*. Ezt $\mathbf{Q}'$-ra alkalmazva azt kapjuk, hogy ha $\mathbf{Q}'$ ortogonális, akkor $(\mathbf{Q}')' = \mathbf{Q}$ is, ami bizonyítja állításunkat. □

7.65. PÉLDA (ORTOGONÁLIS MÁTRIXOK INVERZE). *Számítsuk ki a 7.65. példabeli* **A** *és* **C** *mátrixok inverzét!*

MEGOLDÁS. Az **A** és **C** mátrixok ortogonalitását a 7.65. példában a definíció alapján ellenőriztük. Az **A** és **C** inverze az előző tétel szerint megegyezik transzponáltjukkal, tehát

$$\mathbf{A}^{-1} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad \mathbf{C}^{-1} = \frac{1}{7}\begin{bmatrix} 2 & 6 & 3 \\ 3 & 2 & -6 \\ 6 & -3 & 2 \end{bmatrix}. \qquad \square$$

Ortogonális mátrixok geometriája Ortogonális mátrixhoz tartozó mátrixleképezés ONB-t ONB-ba visz úgy, ahogy a síkban vagy térben a forgatás és a tükrözés.

7.66. TÉTEL (ORTOGONÁLIS MÁTRIXHOZ TARTOZÓ MÁTRIXLEKÉPEZÉS). *Legyen* $\mathbf{Q} \in \mathbb{R}^{n\times n}$. *Az alábbi állítások ekvivalensek:*

a) $\mathbf{Q}$ *ortogonális.*

b) $|\mathbf{Qx}| = |\mathbf{x}|$ *minden* $\mathbf{x} \in \mathbb{R}^n$ *vektorra.*

c) $\mathbf{Qx} \cdot \mathbf{Qy} = \mathbf{x} \cdot \mathbf{y}$ *minden* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *vektorra.*

BIZONYÍTÁS. *a)* $\Rightarrow$ *b)*: Ha $\mathbf{Q}$ ortogonális, akkor $\mathbf{Q}'\mathbf{Q} = \mathbf{I}$, így tetszőleges $\mathbf{x} \in \mathbb{R}^n$ vektorra

$$|\mathbf{Qx}|^2 = \mathbf{Qx} \cdot \mathbf{Qx} = (\mathbf{Qx})'(\mathbf{Qx}) = \mathbf{x}'\mathbf{Q}'\mathbf{Qx} = \mathbf{x}'\mathbf{x} = |\mathbf{x}|^2.$$

b) $\Rightarrow$ *c)*: A *b)*-ből következik, hogy

$$|\mathbf{Q}(\mathbf{x}+\mathbf{y})| = |\mathbf{x}+\mathbf{y}| \ \text{ és } \ |\mathbf{Q}(\mathbf{x}-\mathbf{y})| = |\mathbf{x}-\mathbf{y}|.$$

Ezt, és a skalárszorzás és az abszolút érték közti kapcsolatot megadó (1.8) egyenletet fölhasználva kapjuk, hogy minden $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ vektorra

$$\begin{aligned} \mathbf{Qx} \cdot \mathbf{Qy} &= \frac{1}{4}\left(|\mathbf{Qx}+\mathbf{Qy}|^2 - |\mathbf{Qx}-\mathbf{Qy}|^2\right) \\ &= \frac{1}{4}\left(|\mathbf{Q}(\mathbf{x}+\mathbf{y})|^2 - |\mathbf{Q}(\mathbf{x}-\mathbf{y})|^2\right) \\ &= \frac{1}{4}\left(|\mathbf{x}+\mathbf{y}|^2 - |\mathbf{x}-\mathbf{y}|^2\right) \\ &= \mathbf{x} \cdot \mathbf{y} \end{aligned}$$

c) $\Rightarrow$ *a)*: A $\mathbf{Q}$ mátrix i-edik oszlopát jelölje $\mathbf{q}_i$, azaz $\mathbf{q}_i = \mathbf{Qe}_i$, ahol $\mathbf{e}_i$ a standard bázis i-edik vektora. Ekkor

$$\mathbf{q}_i \cdot \mathbf{q}_j = \mathbf{Qe}_i \cdot \mathbf{Qe}_j = \mathbf{e}_i \cdot \mathbf{e}_j = \begin{cases} 0, & \text{ha } i \neq j, \\ 1, & \text{ha } i = j. \end{cases}$$

Tehát $\mathbf{Q}$ oszlopvektorai ortonormált rendszert alkotnak, azaz $\mathbf{Q}$ ortogonális. $\square$

► A tétel egyik állítása úgy is kimondható, hogy egy $\mathbf{Q}$ négyzetes mátrix pontosan akkor ortogonális, ha a $Q : \mathbf{x} \mapsto \mathbf{Qx}$ mátrixleképezés távolságtartó. A tétel másik állítása azt mondja, hogy $\mathbf{Q}$ pontosan akkor ortogonális, ha a Q megtartja a skaláris szorzatot.

► Hasonló állítás igaz szemiortogonális mátrixokra is (ld. **??**. feladat).

7.67. tétel (Ortogonális mátrixok tulajdonságai).
a) *Ha $\mathbf{Q}$ valós ortogonális mátrix, akkor $|\det(Q)| = 1$.*
b) *Az $n \times n$-es valós ortogonális mátrixok $O(n)$ halmazából nem vezet ki a mátrixszorzás és invertálás művelete.*

Bizonyítás. *a)* Mivel $\mathbf{Q}'\mathbf{Q} = \mathbf{I}$, ezért $\det(\mathbf{Q}')\det(\mathbf{Q}) = \det(\mathbf{I}) = 1$, de $\det(\mathbf{Q}') = \det(\mathbf{Q})$, így $\det(\mathbf{Q}) = 1$ vagy $\det(\mathbf{Q}) = -1$.

b) Ortogonális mátrix inverze megegyezik transzponáltjával, ami ugyancsak ortogonális, tehát inverze is az. Be kell még látni, hogy két ortogonális mátrix szorzata is ortogonális. Legyen $\mathbf{Q}_1$ és $\mathbf{Q}_2$ ortogonális. Ekkor

$$(\mathbf{Q}_1\mathbf{Q}_2)'\mathbf{Q}_1\mathbf{Q}_2 = \mathbf{Q}_2'\mathbf{Q}_1'\mathbf{Q}_1\mathbf{Q}_2 = \mathbf{Q}_2'\mathbf{Q}_2 = \mathbf{I},$$

tehát $\mathbf{Q}_1\mathbf{Q}_2$ valóban ortogonális. □

► Az is azonnal látható, hogy az $n \times n$-es 1 determinánsú valós ortogonális mátrixok halmazából sem vezet ki a mátrixszorzás és invertálás művelete. E mátrixhalmazt $SO(n)$ jelöli.

A valós ortogonális mátrixok $O(n)$ halmaza a mátrixszorzás műveletével *csoportot* alkot. Ezt ortogonális csoportnak nevezik. A csoportokról a függelékben írunk. Az 1 determinánsú ortogonális mátrixok $SO(n)$ csoportját speciális ortogonális csoportnak nevezik. $O(10)$ fontos szerepet játszik a modern fizika húrelméletében, mint a 10-dimenziós tér-idő szimmetriacsoportja.

A 2- és 3-dimenziós tér ortogonális transzformációi Forgatások és tükrözések segítségével leírhatók az ortogonális mátrixok.

7.68. tétel. *Minden $O(2)$-be eső ortogonális mátrix vagy egy α szögű forgatás, vagy egy $\alpha/2$ szögű egyenesre való tükrözés mátrixa.*

Bizonyítás. Legyen $\mathbf{Q} = \left[\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}\right]$. Ha e mátrix ortogonális, akkor oszlopvektorai ortonormált rendszert alkotnak, azaz

$$\begin{aligned} a^2 + b^2 &= 1 \\ c^2 + d^2 &= 1 \\ ac + bd &= 0. \end{aligned}$$

Az utolsó egyenlet szerint $a^2c^2 = b^2d^2$, azaz $a^2(1-d^2) = (1-a^2)d^2$, amiből $a^2 = d^2$, és $b^2 = c^2$ adódik. Végül kapjuk, hogy vagy $d = a$ és $c = -b$, vagy $d = -a$ és $c = b$. Az első esetben $\det(\mathbf{Q}) = ad - bc = 1$, a másodikban $\det(\mathbf{Q}) = -1$. Vegyük észre, hogy bármely megoldáshoz egyértelműen találunk egy olyan $\alpha \in [0, 2\pi)$ valóst, hogy $a = \cos\alpha$ és $b = \sin\alpha$. Vagyis az összes másodrendű ortogonális mátrix

$$\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \quad \text{vagy} \quad \begin{bmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{bmatrix}$$

alakba írható. Ha a determinánsa 1, akkor egy α szögű forgatás, ha determinánsa -1, akkor egy $\alpha/2$ szögű egyenesre való tükrözés mátrixa (ld. a 7.29. és a 7.34. pontokat). □

A 3-dimenziós eset kissé bonyolultabb. Számtalan klasszikus műszaki alkalmazásban – mindenek előtt a merev testek mozgásának leírásában – fontos szerepet játszanak $SO(3)$ elemei, azaz az 1 determinánsú ortogonális mátrixok. Ezek itt és nagyobb dimenzió esetén is a forgások mátrixai. A forgás azonban csak a 3-dimenzióban írható le úgy, mint egy tengely körül α szöggel való elfordulás. E tételt csak a sajátvektorok elméletének ismeretében fogjuk tudni bizonyítani (ld. **??** tétel).

7.69. PÉLDA (FORGATÁS TENGELYE ÉS SZÖGE). *Az* 1 *determinánsú ortogonális*

$$\frac{1}{15}\begin{bmatrix} 14 & -5 & 2 \\ 5 & 10 & -10 \\ 2 & 10 & 11 \end{bmatrix}$$

mátrix milyen tengely körüli és mekkora szöggel való forgatás mátrixa?

MEGOLDÁS. Ha $\mathbf{A}$ egy 0-tól különböző szöggel való forgatás mátrixa, és $\mathbf{v}$ a tengely egy irányvektora, akkor csak $\mathbf{v}$ skalárszorosai fogják kielégíteni az $\mathbf{Ax} = \mathbf{x}$ egyenletet. Ez ekvivalens a homogén lineáris

$$(\mathbf{A} - \mathbf{I})\mathbf{x} = \mathbf{0}$$

egyenletrendszerrel, melynek alakja és megoldása esetünkben

$$\frac{1}{15}\begin{bmatrix} -1 & -5 & 2 \\ 5 & -5 & -10 \\ 2 & 10 & -4 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \qquad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = t\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}.$$

Tehát a forgástengely egy irányvektora a $\mathbf{v} = (2, 0, 1)$ vektor. A forgásszög, illetve a forgásszög koszinuszának meghatározásához elég egy olyan $\mathbf{w}$ vektort találni, mely a tengelyre merőleges síkban van. Ilyen például a $\mathbf{w} = (0, 1, 0)$ vektor. Ennek képe a forgatásnál

$$\frac{1}{15}\begin{bmatrix} 14 & -5 & 2 \\ 5 & 10 & -10 \\ 2 & 10 & 11 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/3 \\ 2/3 \\ 2/3 \end{bmatrix}.$$

A forgásszög megegyezik e két vektor szögével, tehát

$$\cos\alpha = \frac{\mathbf{w} \cdot \mathbf{Aw}}{|\mathbf{w}||\mathbf{Aw}|} = \frac{2/3}{1 \cdot 1} = \frac{2}{3}.$$

Ez összhangban van a 7.29. feladat eredményével, ahol e forgatást a tengely és a szög ismeretében kellett megkonstruálni. □

A harmadrendű -1 determinánsú ortogonális mátrixok, azaz $O(3) - SO(3)$ elemei ugyan nem mind tükrözések, de egy tükrözés és egy forgatás egymás utáni alkalmazásával megkaphatók (ld. **??** tétel)!

*Givens-forgatás, Householder-tükrözés** Az n-dimenziós tér forgatásai és tükrözései közül kiválaszthatunk olyan egyszerű, ún. primitív ortogonális transzformációkat, melyek mátrixai szorzataként az összes ortogonális mátrix előállítható. E transzformációkat több hatékony numerikus matematikai módszer is használja.

Azt a forgatást, mely egy koordinátasík vektorain kívül minden más vektort helyben hagy, *Givens-forgatásnak* nevezzük. Az i-edik és j-edik koordinátatengely síkját érintő forgatás mátrixa

$$\mathbf{G} = \begin{bmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \dots & \cos\alpha & \dots & -\sin\alpha & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & \sin\alpha & \dots & \cos\alpha & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix}$$

amit úgy kapunk meg, hogy az egységmátrix i-edik és j-edik sorának és oszlopának metszetében lévő négy helyre az α szögű forgatás mátrixát tesszük.

E forgatással elérhető például, hogy egy $\mathbf{x}$ vektort egy olyan vektorba forgassunk, melynek j-edik koordinátája 0. Csak az i-edik és j-edik sorokat és oszlopokat kiemelve

$$\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}.$$

Ebből látható, hogy az

$$\begin{aligned} r &= \sqrt{a^2 + b^2} \\ \cos\alpha &= {}^{a}/_{r} \\ \sin\alpha &= -{}^{b}/_{r} \end{aligned} \tag{7.24}$$

egyenletek segítségével fölírható a forgatómátrix az a és b ismeretében. Ez használható mátrix háromszögalakra hozásában, például a QR-felbontás Givens-forgatások segítségével is elvégezhető. Ennek előnyei a ritka mátrixok esetén mutatkoznak, és a számítások párhuzamosíthatóak is.

Egy adott $\mathbf{a} \neq \mathbf{0}$ vektorra, vagy az $\mathbf{e} = {}^{\mathbf{a}}/_{|\mathbf{a}|}$ egységvektorra merőleges hipersíkra való tükrözést *Householder-tükrözésnek* nevezzük. Mátrixa

$$\mathbf{H} = \mathbf{I} - \frac{2}{\mathbf{a}'\mathbf{a}}\mathbf{a}\mathbf{a}'$$

Feladatként az Olvasóra hagyjuk annak bizonyítását, hogy e transzformáció valóban helyben hagyja az $\mathbf{e}^{\perp}$ tér összes vektorát és $-\mathbf{e}$-be viszi az $\mathbf{e}$ vektort (ld. **??**. feladat). E tükrözés is használható egy mátrix háromszögalakra hozásához, QR-felbontásának megkonstruálásához. Ehhez szükség lesz az alábbi állításra.

7.70. állítás (Egy vektor tükrözése egy másikba). *Ha $\mathbf{a}$ és $\mathbf{b}$ két különböző, de azonos hosszúságú vektor $\mathbb{R}^n$-ben, akkor az $(\mathbf{a}-\mathbf{b})^\perp$ hipersíkra való Householder-tükrözés az $\mathbf{a}$ vektort $\mathbf{b}$-be viszi és viszont.*

Bizonyítás. Meg kell mutatnunk, hogy $\mathbf{Ha} = \mathbf{b}$ és $\mathbf{Hb} = \mathbf{a}$, ahol

$$\mathbf{H} = \mathbf{I} - \frac{2}{(\mathbf{a}-\mathbf{b})'(\mathbf{a}-\mathbf{b})}(\mathbf{a}-\mathbf{b})(\mathbf{a}-\mathbf{b})'.$$

Kihasználjuk, hogy $\mathbf{a}$ és $\mathbf{b}$ azonos hosszúságúak, így $\mathbf{a}'\mathbf{a} = \mathbf{b}'\mathbf{b}$, és hogy a skaláris szorzás felcserélhető, azaz $\mathbf{a}'\mathbf{b} = \mathbf{b}'\mathbf{a}$. Így

$$(\mathbf{a}-\mathbf{b})'(\mathbf{a}-\mathbf{b}) = \mathbf{a}'\mathbf{a} - \mathbf{a}'\mathbf{b} - \mathbf{b}'\mathbf{a} + \mathbf{b}'\mathbf{b} = 2(\mathbf{a}'\mathbf{a} - \mathbf{b}'\mathbf{a}) = 2(\mathbf{a}-\mathbf{b})'\mathbf{a}.$$

Eszerint

$$\begin{aligned}
\mathbf{Ha} &= \mathbf{a} - \frac{2}{(\mathbf{a}-\mathbf{b})'(\mathbf{a}-\mathbf{b})}(\mathbf{a}-\mathbf{b})(\mathbf{a}-\mathbf{b})'\mathbf{a} \\
&= \mathbf{a} - \frac{1}{(\mathbf{a}-\mathbf{b})'\mathbf{a}}(\mathbf{a}-\mathbf{b})'\mathbf{a}(\mathbf{a}-\mathbf{b}) \\
&= \mathbf{a} - (\mathbf{a}-\mathbf{b}) = \mathbf{b}.
\end{aligned}$$

Mivel $\mathbf{H}^{-1} = \mathbf{H}$, ezért $\mathbf{Hb} = \mathbf{H}^{-1}\mathbf{b} = \mathbf{a}$. □

7.71. példa (Householder-tükrözés). *Határozzuk meg azt a $\mathbf{H}$ mátrixot, mely az $(1,-1,-1,1)$ vektort olyan vektorba viszi, melynek az elsőt kivéve minden koordinátája 0.*

Megoldás. $|(1,-1,-1,1)| = 2$, ezért a képvektor csak a $\pm(2,0,0,0)$ vektorok valamelyike lehet. Válasszuk a pozitív koordinátájút. Az $(1,-1,-1,1)-(2,0,0,0) = (-1,-1,-1,1)$ vektorra merőleges hipersíkra való tükrözés mátrixa

$$\mathbf{H} = \mathbf{I} - \frac{2}{\mathbf{a}'\mathbf{a}}\mathbf{a}\mathbf{a}' = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} -1 \\ -1 \\ -1 \\ 1 \end{bmatrix}\begin{bmatrix} -1 & -1 & -1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Fejben számolva is könnyen ellenőrizhetjük, hogy $\mathbf{H}\cdot(1,-1,-1,1) = (2,0,0,0)$. □

*Gram–Schmidt-ortogonalizáció** Nagy előnyökkel jár, ha egy altérnek nem csak egy bázisát, hanem egy ortogonális bázisát ismerjük. E paragrafusban megmutatjuk, hogy ilyen bázis létezik, és eljárást adunk a megkonstruálására. Ezt az eljárást Gram–Schmidt-ortogonalizációnak nevezzük.

7.72. TÉTEL (GRAM–SCHMIDT-ORTOGONALIZÁCIÓ). *Ha* $\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k\}$ *egy független vektorrendszer, akkor létezik olyan ortogonális* $\mathcal{V} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ *vektorrendszer, hogy minden* $i = 1, 2, \dots, k$ *esetén*

$$\operatorname{span}(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_i) = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_i). \tag{7.25}$$

Az ortogonális $\mathcal{V}$ *rendszerből a vektorok normálásával kapott*

$$\left\{ \frac{\mathbf{v}_1}{|\mathbf{v}_1|}, \frac{\mathbf{v}_2}{|\mathbf{v}_2|}, \dots, \frac{\mathbf{v}_k}{|\mathbf{v}_k|} \right\}$$

rendszer ortonormált.

BIZONYÍTÁS. A $\operatorname{span}(\mathbf{a}_1) = \operatorname{span}(\mathbf{v}_1)$ összefüggés teljesül, ha

$$\mathbf{v}_1 = \mathbf{a}_1.$$

A $\operatorname{span}(\mathbf{a}_1, \mathbf{a}_2) = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2)$ teljesülése érdekében olyan $\mathbf{v}_2$ vektort kell választani, mely az $\mathbf{a}_1$ és $\mathbf{a}_2$ síkjában van, másrészt $\mathbf{v}_2$-nek merőlegesnek kell lennie $\mathbf{v}_1$-re. E feltételeket teljesíti az $\mathbf{a}_2$-nek a $\mathbf{v}_1$ által kifeszített altérre merőleges összetevője, azaz a

$$\mathbf{v}_2 = \mathbf{a}_2 - \left(\mathbf{a}_2 \cdot \frac{\mathbf{v}_1}{|\mathbf{v}_1|} \right) \frac{\mathbf{v}_1}{|\mathbf{v}_1|} = \mathbf{a}_2 - \frac{\mathbf{a}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1$$

vektor. Látható, hogy e vektor nem lehet a 0-vektor, hisz $\mathbf{v}_2 = 0$ esetén $\mathbf{a}_2 = \frac{\mathbf{a}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 = \frac{\mathbf{a}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{a}_1$ lenne, azaz $\mathbf{a}_1$ és $\mathbf{a}_2$ nem lenne független, ami ellentmond annak, hogy $\mathcal{A}$ független. Az előző képletekből látható, hogy $\mathbf{v}_1$ és $\mathbf{v}_2$ előállítható az $\mathbf{a}_1$ és $\mathbf{a}_2$ lineáris kombinációjaként, és viszont, így $\operatorname{span}(\mathbf{a}_1, \mathbf{a}_2) = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2)$ fönnáll. Az eljárás hasonlóképp folytatható. Ha már megkonstruáltuk $\mathbf{v}_i$-t, akkor a 7.59. tétel szerint kiszámoljuk az $\mathbf{a}_{i+1}$ vektornak az $\operatorname{span}(\frac{\mathbf{v}_1}{|\mathbf{v}_1|}, \frac{\mathbf{v}_2}{|\mathbf{v}_2|}, \dots, \frac{\mathbf{v}_i}{|\mathbf{v}_i|})$ altér szerinti merőleges összetevőjét, és ezt választjuk $\mathbf{v}_{i+1}$-nek, azaz

$$\mathbf{v}_{i+1} = \mathbf{a}_{i+1} - \frac{\mathbf{a}_{i+1} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 - \frac{\mathbf{a}_{i+1} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2 - \dots - \frac{\mathbf{a}_{i+1} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i} \mathbf{v}_i$$

Könnyen látható, hogy $\mathbf{v}_{i+1} \neq \mathbf{0}$, mert ellenkező esetben $\mathcal{A}$ nem volna független. Látható az is, hogy $\mathbf{v}_{i+1}$ kifejezhető az $\mathbf{a}_1$, $\mathbf{a}_2$,..., $\mathbf{a}_{i+1}$ vektorok lineáris kombinációjaként, és $\mathbf{a}_{i+1}$ kifejezhető az $\mathbf{v}_1$, $\mathbf{v}_2$,..., $\mathbf{v}_{i+1}$ vektorok lineáris kombinációjaként, tehát a tétel kifeszített alterekre vonatkozó állítása is fennáll. □

7.73. PÉLDA (GRAM–SCHMIDT-ORTOGONALIZÁCIÓ). *Keressünk ortonormált bázist az* $(1,1,1,1)$, $(3,-1,3,-1)$, $(6,2,2,-2)$ *vektorok által kifeszített altérben.*

MEGOLDÁS. Először keressünk egy ortogonális bázist:

$$\begin{aligned}
\mathbf{v}_1 &= (1,1,1,1)\\
\mathbf{v}_2 &= (3,-1,3,-1) - \frac{(3,-1,3,-1)\cdot(1,1,1,1)}{(1,1,1,1)\cdot(1,1,1,1)}(1,1,1,1) = (2,-2,2,-2)\\
\mathbf{v}_3 &= (6,2,2,-2) - \frac{(6,2,2,-2)\cdot(1,1,1,1)}{(1,1,1,1)\cdot(1,1,1,1)}(1,1,1,1)\\
&\qquad - \frac{(6,2,2,-2)\cdot(2,-2,2,-2)}{(2,-2,2,-2)\cdot(2,-2,2,-2)}(2,-2,2,-2) = (2,2,-2,-2)
\end{aligned}$$

Végül az ortonormált bázis:

$$\left\{\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right),\left(\frac{1}{2},-\frac{1}{2},\frac{1}{2},-\frac{1}{2}\right),\left(\frac{1}{2},\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\right)\right\}$$ □

► Könnyen igazolható, hogy a Gram–Schmidt-ortogonalizáció működik nem független vektorokból álló vektorrendszerre is, annyi változással, hogy pontosan akkor lesz $\mathbf{v}_i = \mathbf{0}$, ha $\mathbf{a}_i$ nem független a kisebb indexű vektoroktól, azaz $\mathbf{a}_i$ benne van a $\operatorname{span}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{i-1})$ altérben.

*A QR-felbontás** Ahogyan egy mátrix elemi sorműveletekkel való háromszögalakra hozását tömör formában megőrizte az LU-felbontás, ugyanígy a QR-felbontás őrzi az ortogonalizációs eljárás eredményét. E felbontás mind a legkisebb négyzetek módszerében, mind a később tárgyalandó sajátértékprobléma megoldásában fontos szerephez jut.

Legyen $\mathbf{A}$ egy teljes oszloprangú mátrix. Az $\mathbf{A} = \mathbf{QR}$ felbontást QR-felbontásnak nevezzük, ha $\mathbf{Q}$ az $\mathbf{A}$-val azonos méretű szemiortogonális mátrix, és $\mathbf{R}$ felső háromszögmátrix.

A Gram–Schmidt-ortogonalizációs eljárásból könnyen előállítható egy mátrix QR-felbontása. Legyen $\mathbf{A} = [\mathbf{a}_1\ \mathbf{a}_2\ \ldots\ \mathbf{a}_k] \in \mathbb{R}^{n\times k}$. Mivel $\mathbf{A}$ teljes oszloprangú, azaz oszlopai függetlenek, ezért $k \le n$. Az ortogonalizációs eljárás végén kapott egységvektorokat jelölje $\mathbf{q}_i$, azaz $\mathbf{q}_i = \frac{\mathbf{v}_i}{|\mathbf{v}_i|}$ $(i = 1,2,\ldots,k)$. Mivel a Gram–Schmidt-ortogonalizációs tétel szerint $\operatorname{span}(\mathbf{a}_1,\ldots,\mathbf{a}_i) = \operatorname{span}(\mathbf{q}_1,\ldots,\mathbf{q}_i)$ minden $i = 1,2,\ldots,k$ értékre, ezért léteznek olyan r_{ij} skalárok, hogy

$$\begin{aligned}
\mathbf{a}_1 &= r_{11}\mathbf{q}_1\\
\mathbf{a}_2 &= r_{12}\mathbf{q}_1 + r_{22}\mathbf{q}_2\\
&\ \vdots\\
\mathbf{a}_k &= r_{1k}\mathbf{q}_1 + r_{2k}\mathbf{q}_2 + \cdots + r_{kk}\mathbf{q}_k.
\end{aligned}\tag{7.26}$$

Ezt mátrixszorzat-alakba írva épp a kívánt felbontást kapjuk:

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_k \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \dots & \mathbf{q}_k \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1k} \\ 0 & r_{22} & \dots & r_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_{kk} \end{bmatrix} = \mathbf{QR}.$$

Ezzel bizonyítottuk a QR-felbontás létezését. A Gram-Schmidt-eljárásból megkaphatjuk a $\mathbf{Q}$ mátrixot, azonban kérdés, hogy az $\mathbf{R}$ hogyan számítható ki egyszerűen. Az előző **??** állítás egyszerű megoldást ad. Ha $\mathbf{A} = \mathbf{QR}$, akkor az egyenlőség mindkét oldalát $\mathbf{Q}'$-tal szorozva kapjuk, hogy $\mathbf{Q}'\mathbf{A} = \mathbf{Q}'\mathbf{QR} = \mathbf{I}_k\mathbf{R} = \mathbf{R}$, tehát

$$\mathbf{R} = \mathbf{Q}'\mathbf{A}.$$

7.74. PÉLDA (QR-FELBONTÁS KISZÁMÍTÁSA). *Határozzuk meg az*

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 6 \\ 1 & -1 & 2 \\ 1 & 3 & 2 \\ 1 & -1 & -2 \end{bmatrix}$$

mátrix QR-felbontását.

MEGOLDÁS. A 7.73. példában épp az $\mathbf{A}$ mátrix oszlopvektoraiból álló vektorrendszert ortogonalizáltuk. Mivel a három vektor lineárisan független, ezért $\mathbf{A}$ teljes oszloprangú. A 7.73. példa megoldása alapján az $\mathbf{A}$ oszlopvektorainak ortogonalizálásával kapott vektorokkal fölírható a $\mathbf{Q}$ mátrix:

$$\mathbf{Q} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \end{bmatrix}$$

Innen

$$\mathbf{R} = \mathbf{Q}'\mathbf{A} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 6 \\ 1 & -1 & 2 \\ 1 & 3 & 2 \\ 1 & -1 & -2 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 4 \\ 0 & 4 & 4 \\ 0 & 0 & 4 \end{bmatrix}$$

Valóban,

$$\begin{bmatrix} 1 & 3 & 6 \\ 1 & -1 & 2 \\ 1 & 3 & 2 \\ 1 & -1 & -2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 2 & 4 \\ 0 & 4 & 4 \\ 0 & 0 & 4 \end{bmatrix}.$$

□

7.75. TÉTEL (QR-FELBONTÁS). *Bármely teljes oszloprangú* $\mathbf{A}$ *mátrixnak létezik QR-felbontása, azaz létezik egy szemiortogonális* $\mathbf{Q}$ *mátrix és egy* $\mathbf{R}$ *felső háromszögmátrix, hogy* $\mathbf{A} = \mathbf{QR}$*. Az* $\mathbf{R}$ *mátrix invertálható, és elérhető, hogy főátlójának minden eleme pozitív legyen. Az így kapott felbontás egyértelmű.*

BIZONYÍTÁS. A felbontás létezését a Gram–Schmidt-ortogonalizációra alapozva az előzőekben megmutattuk. Mivel $\mathbf{A} = \mathbf{QR}$, és $\mathbf{A}$ sorvektorai az $\mathbf{R}$ sorvektorainak lineáris kombinációi, ezért ha $\mathbf{R}$ rangja kisebb lenne k-nál, az $\mathbf{A}$ rangja is az lenne, de $\mathbf{A}$ rangja k. Beláttuk tehát, hogy $\mathbf{R}$ invertálható, és mivel felső háromszögmátrix, ezért főátlójában nem lehetnek 0-elemek. Ha valamely i-re $r_{ii} < 0$ lenne, akkor szorozzuk be az $\mathbf{R}$ mátrix i-edik sorvektorát és a $\mathbf{Q}$ mátrix i-edik oszlopvektorát, azaz a $\mathbf{q}_i$ vektort -1-gyel. Ez a szorzaton nem változtat. Így elérhetjük, hogy $\mathbf{R}$ minden főátlón lévő eleme pozitív legyen.

Az egyértelműség bizonyításához elég belátni, hogy a (7.26) egyenletek mindegyikéből – előjel erejéig egyértelműen – meghatározhatók az ismeretlen együtthatók és az ismeretlen $\mathbf{q}$-vektor, így az $r_{ii} > 0$ választás mellett az már egyértelmű lesz. Ez abból következik, hogy egyrészt az első egyenletből $r_{11} = |\mathbf{a}_1|$, ami nagyobb 0-nál és $\mathbf{q}_1 = \mathbf{a}_1/r_{11}$ és ezek egyértelműek. Másrészt minden további

$$\mathbf{a}_i = r_{1i}\mathbf{q}_1 + r_{2i}\mathbf{q}_2 + \cdots + r_{ii}\mathbf{q}_i, \quad (1 < i \leq k)$$

egyenletben $\mathbf{a}_i$ benne van az $\mathcal{A}_i = \operatorname{span}(\mathbf{a}_1, \ldots, \mathbf{a}_i) = \operatorname{span}(\mathbf{q}_1, \ldots, \mathbf{q}_i)$ altérben, de nincs benne az $\mathcal{A}_{i-1}$-ben, így ennek merőleges kiegészítő altere 1-dimenziós, melynek egységnyi bázisvektora kétféleképp választható meg, melyek egyikére lesz r_{ii} pozitív. □

7.76. PÉLDA (QR-FELBONTÁS GIVENS-FORGATÁSOKKAL). *Határozzuk meg az*

$$\mathbf{A} = \begin{bmatrix} 4 & 5 & 8 \\ 3 & 10 & 6 \\ 0 & 12 & 13 \end{bmatrix}$$

mátrix QR-felbontását Givens-forgatások segítségével!

MEGOLDÁS. Először az első és második sorokat és oszlopokat figyelve elimináljuk a második sor harmadik elemét. Itt a (7.24) egyenletekben is használt jelölésekkel $a = 4$, $b = 3$, tehát $r = \sqrt{3^2 + 4^2} = 5$, $\cos\alpha = 4/5$, $\sin\alpha = -3/5$. Így első lépésben a következő mátrixszorzással eliminálhatunk:

$$\mathbf{Q}_1 = \begin{bmatrix} 4/5 & 3/5 & 0 \\ -3/5 & 4/5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{Q}_1\mathbf{A} = \begin{bmatrix} 5 & 10 & 10 \\ 0 & 5 & 0 \\ 0 & 12 & 13 \end{bmatrix}.$$

Következő lépésben a $\mathbf{Q}_1\mathbf{A}$ mátrix harmadik sorának második elemét elimináljuk:

$$\mathbf{Q}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5/13 & 12/13 \\ 0 & -12/13 & 5/13 \end{bmatrix}. \qquad \mathbf{R} = \mathbf{Q}_2\mathbf{Q}_1\mathbf{A} = \begin{bmatrix} 5 & 10 & 10 \\ 0 & 13 & 12 \\ 0 & 0 & 5 \end{bmatrix}.$$

és innen

$$\mathbf{Q} = (\mathbf{Q}_2\mathbf{Q}_1)^{-1} = \mathbf{Q}_1'\mathbf{Q}_2' = \begin{bmatrix} 4/5 & -3/13 & 36/65 \\ 3/5 & 4/13 & -48/65 \\ 0 & 12/13 & 5/13 \end{bmatrix},$$

amely mátrixokkal $\mathbf{A} = \mathbf{QR}$ valóban fennáll. □

A Householder-tükrözést alkalmazva a QR-felbontásra egy további módszert kaphatunk. Az ötlet lényege, hogy a 7.71. példában látott módon először az első oszlopban elimináljuk az első elem alattiakat, majd olyan transzformációt választunk, mely az első sort és oszlopot nem változtatja, de a második sor második eleme alattiakat eliminálja, és így tovább. Egy 4×4-es mátrixon szemléltetjük az eljárást.

$$\mathbf{A} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \to \mathbf{Q}_1\mathbf{A} = \begin{bmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{bmatrix} \to \mathbf{Q}_2\mathbf{Q}_1\mathbf{A} = \begin{bmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{bmatrix} \to \mathbf{Q}_3\mathbf{Q}_2\mathbf{Q}_1\mathbf{A} = \begin{bmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \end{bmatrix}$$

$$\mathbf{Q}_1 = \mathbf{H}_1 \qquad \mathbf{Q}_2 = \left[\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & & & \\ 0 & & \mathbf{H}_2 & \\ 0 & & & \end{array}\right] \qquad \mathbf{Q}_3 = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & \multicolumn{2}{c}{\multirow{2}{*}{$\mathbf{H}_3$}} \\ 0 & 0 & & \end{array}\right]$$

Az első lépésben az $\mathbf{A}$ mátrix első oszlopához ($\mathbf{a}_1$) keresünk egy $\mathbf{b}_1$ vektort, mely vele egyenlő hosszú, és csak az első koordinátája nem 0. Ezután az $\mathbf{a}_1 - \mathbf{b}_1$ vektorhoz megkonstruáljuk a $\mathbf{Q}_1 = \mathbf{H}_1$ Hauseholder-mátrixot. Így $\mathbf{Q}_1\mathbf{A}$ első oszlopában kinulláztuk az első sor alatti elemeket. Ezután elhagyjuk az első sort és oszlopot, és az így kapott mátrix első oszlopvektorával ($\mathbf{a}_2$) és a vele egyenlő hosszú, és az első koordinátát kivéve 0 koordinátájú $\mathbf{b}_2$ vektorral megkonstruáljuk a $\mathbf{H}_2$ Hauseholder-mátrixot, melyet kiegészítünk egy sorral és oszloppal úgy, hogy a $\mathbf{Q}_1\mathbf{A}$ mátrixszal szorozva annak első sorát és oszlopát ne változtassa. Ez lesz a $\mathbf{Q}_2$ mátrix. Hasonlóan folytatva végül egy $\mathbf{R} = \mathbf{Q}_{n-1} \ldots \mathbf{Q}_2\mathbf{Q}_1\mathbf{A}$ felső háromszögmátrixhoz jutunk (a fenti mintán $n = 4$). Mivel a $\mathbf{Q}_i$ mátrixok mindegyike ortogonális, ezért szorzatuk inverze is az lesz. Így a QR-felbontásban $\mathbf{Q} = \mathbf{Q}_1'\mathbf{Q}_2' \ldots \mathbf{Q}_{n-1}'$. A QR-felbontás ilyen módon való meghatározását nevezzük *Hauseholder-módszernek*.

7.77. PÉLDA (QR-FELBONTÁS HAUSEHOLDER-TÜKRÖZÉSSEL). *Határozzuk meg az*

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & -3 \\ -2 & 5 & -7 \end{bmatrix}$$

mátrix QR-felbontását Hauseholder-módszerrel!

MEGOLDÁS. Az $(1,2,-2) \mapsto (3,0,0)$ transzformációhoz az

$$\mathbf{a} = (1,2,-2) - (3,0,0) = (-2,2,-2)$$

vektorral Hauseholder-tükrözést végzünk:

$$\mathbf{Q}_1 = \mathbf{I}_3 - \frac{2}{\mathbf{a}'\mathbf{a}}\mathbf{a}\mathbf{a}' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{1}{6}\begin{bmatrix} 4 & -4 & 4 \\ -4 & 4 & -4 \\ 4 & -4 & 4 \end{bmatrix}$$

$$= \frac{1}{3}\begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix}$$

$$\mathbf{Q}_1\mathbf{A} = \frac{1}{3}\begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & -3 \\ -2 & 5 & -7 \end{bmatrix} = \begin{bmatrix} 3 & -2 & 3 \\ 0 & 4 & -5 \\ 0 & 3 & -5 \end{bmatrix}$$

Ezután a $\mathbf{Q}_1\mathbf{A}$ mátrixból képzeletben elhagyva az első sort és oszlopot a $(4,3) \mapsto (5,0)$ transzformációhoz kell az $\mathbf{a} = (4,3) - (5,0) = (-1,3)$ vektorral Hauseholder-tükrözést végezni:

$$\mathbf{H}_2 = \mathbf{I}_2 - \frac{2}{\mathbf{a}'\mathbf{a}}\mathbf{a}\mathbf{a}' = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{1}{5}\begin{bmatrix} 1 & -3 \\ -3 & 9 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 4 & 3 \\ 3 & -4 \end{bmatrix}$$

$$\mathbf{Q}_2 = \left[\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & 4/5 & 3/5 \\ 0 & 3/5 & -4/5 \end{array}\right], \qquad \mathbf{R} = \mathbf{Q}_2\mathbf{Q}_1\mathbf{A} = \begin{bmatrix} 3 & -2 & 3 \\ 0 & 5 & -7 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{Q} = (\mathbf{Q}_2\mathbf{Q}_1)^{-1} = \mathbf{Q}_1'\mathbf{Q}_2' = \frac{1}{15}\begin{bmatrix} 5 & 2 & 14 \\ 10 & 10 & -5 \\ -10 & 11 & 2 \end{bmatrix}.$$

Az $\mathbf{A} = \mathbf{QR}$ egyenlőség fennállásának ellenőrzését az Olvasóra hagyjuk. □

*Egyenletrendszer optimális megoldása QR-felbontással** Ha egy egyenletrendszer ellentmondásos, az optimális megoldás megtalálásához fölírt normálegyenlet gyakran rosszul kondicionált, ezért érdemes olyan megoldási technikát keresni, mely hatékonyabb a számítási hibák kezelésében. Egy ilyen technikát ismertetünk.

7.78. TÉTEL (LEGKISEBB NÉGYZETEK QR-FELBONTÁSSAL). *Legyen* $\mathbf{A}$ *egy teljes oszloprangú* $m \times n$*-es valós mátrix,* $\mathbf{A} = \mathbf{QR}$ *egy QR-felbontása, és legyen* $\mathbf{b}$ *egy* $\mathbb{R}^m$*-beli vektor. Ekkor az* $\mathbf{Ax} = \mathbf{b}$ *egyenletrendszer egyetlen optimális megoldása* $\hat{\mathbf{x}} = \mathbf{R}^{-1}\mathbf{Q}'\mathbf{b}$*, ami megkapható az*

$$\mathbf{R}\hat{\mathbf{x}} = \mathbf{Q}'\mathbf{b}$$

egyenletrendszerből egyszerű visszahelyettesítéssel is.

BIZONYÍTÁS. Az egyenletrendszer optimális megoldásáról szóló 7.44. tétel szerint az optimális megoldás a normálegyenletből megkapható. Eszerint

$$\begin{aligned}
\mathbf{A}'\mathbf{A}\hat{\mathbf{x}} &= \mathbf{A}'\mathbf{b} && \mathbf{A} = \mathbf{QR} \text{ behelyettesítése után}\\
(\mathbf{QR})'\mathbf{QR}\hat{\mathbf{x}} &= (\mathbf{QR})'\mathbf{b} &&\\
\mathbf{R}'\mathbf{Q}'\mathbf{QR}\hat{\mathbf{x}} &= \mathbf{R}'\mathbf{Q}'\mathbf{b} && \mathbf{Q}'\mathbf{Q} = \mathbf{I}\\
\mathbf{R}'\mathbf{R}\hat{\mathbf{x}} &= \mathbf{R}'\mathbf{Q}'\mathbf{b} && \text{balról szorzás az } (\mathbf{R}')^{-1} \text{ mátrixszal}\\
\mathbf{R}\hat{\mathbf{x}} &= \mathbf{Q}'\mathbf{b}.
\end{aligned}$$

Az utolsó egyenlet visszahelyettesítésekkel is megoldható, mivel $\mathbf{R}$ felső háromszögmátrix. Mivel $\mathbf{R}$ főátlójában nincsenek zéruselemek, ezért $\mathbf{R}$ invertálható (ezt kihasználtuk, amikor $(\mathbf{R}')^{-1}$-gyel szoroztunk), tehát az egyenletből $\hat{\mathbf{x}}$ kifejezhető: $\hat{\mathbf{x}} = \mathbf{R}^{-1}\mathbf{Q}'\mathbf{b}$. □

7.79. PÉLDA (EGYENLETRENDSZER OPTIMÁLIS MEGOLDÁSA). *Az alábbi háromismeretlenes egyenletrendszert megoldottuk a 7.53. példában:*

$$\begin{aligned}
x + 3y + 6z &= 8\\
x - y + 2z &= 2\\
x + 3y + 2z &= 2\\
x - y - 2z &= 0
\end{aligned}$$

Adjunk rá új, a QR-felbontást használó megoldást!

MEGOLDÁS. Az egyenletrendszer együtthatómátrixának QR-felbontását meghatároztuk a 7.74. példában. Eszerint

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 6\\ 1 & -1 & 2\\ 1 & 3 & 2\\ 1 & -1 & -2 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1\\ 1 & -1 & 1\\ 1 & 1 & -1\\ 1 & -1 & -1 \end{bmatrix}\begin{bmatrix} 2 & 2 & 4\\ 0 & 4 & 4\\ 0 & 0 & 4 \end{bmatrix}.$$

Egyik lehetőség, hogy fölírjuk az $\mathbf{R}\hat{\mathbf{x}} = \mathbf{Q}'\mathbf{b}$ mátrixegyenletet:

$$\begin{bmatrix} 2 & 2 & 4\\ 0 & 4 & 4\\ 0 & 0 & 4 \end{bmatrix}\begin{bmatrix} \hat{x}_1\\ \hat{x}_2\\ \hat{x}_3 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1\\ 1 & -1 & 1 & -1\\ 1 & 1 & -1 & -1 \end{bmatrix}\begin{bmatrix} 8\\ 2\\ 2\\ 0 \end{bmatrix} = \begin{bmatrix} 6\\ 4\\ 4 \end{bmatrix}$$

Ezt az egyenletrendszert fejben is meg tudjuk oldani visszahelyettesítéssel: $(\hat{x}_1, \hat{x}_2, \hat{x}_3) = (1, 0, 1)$. Természetesen ha már kiszámoltuk az $\mathbf{R}^{-1}$ mátrixot, akkor segítségével is megkapható az optimális megoldás:

$$\hat{\mathbf{x}} = \mathbf{R}^{-1}\mathbf{Q}'\mathbf{b} = \frac{1}{4}\begin{bmatrix} 2 & -1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 8 \\ 2 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

□

Feladatok

7.28. Igazoljuk, hogy ha $\mathbf{e}$ az $\mathbb{R}^n$ egy egységvektora, akkor a $\mathbf{H} = \mathbf{I} - \frac{2}{\mathbf{a}'\mathbf{a}}\mathbf{a}\mathbf{a}'$ Householder-transzformáció helyben hagyja az $\mathbf{e}^{\perp}$ altér összes vektorát és $\mathbf{H}\mathbf{e} = -\mathbf{e}$.

*Komplex és véges test feletti terek**

A következőkben egyre többször szembesülünk azzal, hogy valós számokkal megfogalmazható problémák megválaszolásához is szükség van a komplex számokra. E fejezetben a geometriai szemléletmódot is kiterjesztjük a komplex terekre. Meglepő módon a geometriai analógiák még a véges test feletti terek esetén is jól használhatók.

Komplex vektorok és terek

Komplex vektorok skaláris szorzata Ha $\mathbb{C}^n$-beli vektorok skaláris szorzatát úgy értelmeznénk, mint a valós vektorok esetén, fura dolgok történnének.

Vegyük például a $(1, \mathrm{i})$ és az (i, i) vektorokat. Önmagával vett skaláris szorzata e két vektornak ez lenne:

$$(1, \mathrm{i}) \cdot (1, \mathrm{i}) \stackrel{?}{=} 1 - 1 = 0$$
$$(\mathrm{i}, \mathrm{i}) \cdot (\mathrm{i}, \mathrm{i}) \stackrel{?}{=} -1 - 1 = -2$$

Ez azt mutatja, hogy ha a komplex vektorok abszolút értékét (hosszát), a valósban használt skaláris szorzattal definiálnánk, az abszolút érték legfontosabb tulajdonságai nem maradnának igazak! Kérdés, kiterjeszthető-e a valós vektorok skaláris szorzatának definíciója a komplex vektorokra úgy, hogy a fontosabb tulajdonságok érvényben maradjanak? Az ötletet a komplex számok – mint egydimenziós vektorok – abszolút értéke adja. A $z = a + \mathrm{i}b$ szám abszolút értékének négyzete $z\bar{z}$, és nem z^2! Eszerint az egydimenziós z vektor önmagával vett skaláris szorzatának $z\bar{z}$-t vagy $\bar{z}z$-t kell adnia. Ennek megfelelően a $\mathbf{z} = (z_1, z_2, \dots, z_n)$ és a $\mathbf{w} = (w_1, w_2, \dots, w_n)$ vektorok skaláris szorzatának egy lehetséges definíciója

$$\mathbf{z} \cdot \mathbf{w} = z_1\overline{w_1} + z_2\overline{w_2} + \cdots + z_n\overline{w_n}, \text{ vagy}$$
$$\mathbf{z} \cdot \mathbf{w} = \overline{z_1}w_1 + \overline{z_2}w_2 + \cdots + \overline{z_n}w_n.$$

Mindkét fenti képlet használható, ízlés kérdése melyiket választjuk (könyvenként változik). Mi az utóbbit fogjuk használni a skaláris szorzat mátrixszorzatos alakjának egyszerűbb volta miatt (ld. majd a 7.81. definícióban). Mindenek előtt egy elnevezés:

7.80. DEFINÍCIÓ (KOMPLEX MÁTRIX ADJUNGÁLTJA). *Az* $\mathbf{A}$ *komplex mátrix* adjungáltján *(vagy* Hermite-féle transzponáltján*) elemenkénti konjugáltjának transzponáltját értjük. Az* $\mathbf{A}$ *adjungáltját* $\mathbf{A}^*$*, vagy Hermite neve után* $\mathbf{A}^{\mathsf{H}}$ *jelöli, tehát* $\mathbf{A}^* = \overline{\mathbf{A}}'$.

Legyünk óvatosak az *adjungált* kifejezéssel: a könyvünk determinánsokról szóló fejezetében *klasszikus adjungáltnak* nevezett fogalmat ne keverjük össze ezzel az adjungálttal, nincs közük egymáshoz!

Például $\begin{bmatrix} \mathrm{i} & 1+\mathrm{i} \\ -\mathrm{i} & 2 \end{bmatrix}^* = \begin{bmatrix} -\mathrm{i} & \mathrm{i} \\ 1-\mathrm{i} & 2 \end{bmatrix}$, míg $[1-\mathrm{i} \ \ \mathrm{i}]^* = \begin{bmatrix} 1+\mathrm{i} \\ -\mathrm{i} \end{bmatrix}$.

7.81. DEFINÍCIÓ (KOMPLEX VEKTOROK SKALÁRIS SZORZATA). *A $\mathbb{C}^n$-beli $\mathbf{z} = (z_1, z_2, \dots, z_n)$ és $\mathbf{w} = (w_1, w_2, \dots, w_n)$ vektorok skaláris szorzatán a*

$$\mathbf{z} \cdot \mathbf{w} = \overline{z_1} w_1 + \overline{z_2} w_2 + \dots + \overline{z_n} w_n$$

komplex skalárt értjük. Ennek mátrixszorzatos alakja $\mathbf{z} \cdot \mathbf{w} = \mathbf{z}^* \mathbf{w}$.

Így a fent említett $(1, \mathrm{i})$ és az (i, i) önmagukkal és egymással vett skaláris szorzatai:

$$\begin{aligned}
(1,\mathrm{i}) \cdot (1,\mathrm{i}) &= \begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix}^* \begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix} = \begin{bmatrix} 1 & -\mathrm{i} \end{bmatrix} \begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix} = 1 - \mathrm{i}^2 = 2, \\
(\mathrm{i},\mathrm{i}) \cdot (\mathrm{i},\mathrm{i}) &= \begin{bmatrix} \mathrm{i} \\ \mathrm{i} \end{bmatrix}^* \begin{bmatrix} \mathrm{i} \\ \mathrm{i} \end{bmatrix} = \begin{bmatrix} -\mathrm{i} & -\mathrm{i} \end{bmatrix} \begin{bmatrix} \mathrm{i} \\ \mathrm{i} \end{bmatrix} = -\mathrm{i}^2 - \mathrm{i}^2 = 2, \\
(1,\mathrm{i}) \cdot (\mathrm{i},\mathrm{i}) &= \begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix}^* \begin{bmatrix} \mathrm{i} \\ \mathrm{i} \end{bmatrix} = \begin{bmatrix} 1 & -\mathrm{i} \end{bmatrix} \begin{bmatrix} \mathrm{i} \\ \mathrm{i} \end{bmatrix} = \mathrm{i} - \mathrm{i}^2 = 1 + \mathrm{i}, \\
(\mathrm{i},\mathrm{i}) \cdot (1,\mathrm{i}) &= \begin{bmatrix} \mathrm{i} \\ \mathrm{i} \end{bmatrix}^* \begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix} = \begin{bmatrix} -\mathrm{i} & -\mathrm{i} \end{bmatrix} \begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix} = -\mathrm{i} - \mathrm{i}^2 = 1 - \mathrm{i}.
\end{aligned}$$

► Világos, hogy két valós vektor skaláris szorzata az eredeti és e definíció szerint is ugyanazt az eredményt adja, ugyanis minden valós r számra $\bar{r} = r$, tehát valós $\mathbf{u}$ és $\mathbf{v}$ vektorok esetén $\mathbf{u}^* = \mathbf{u}'$, így $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^* \mathbf{v} = \mathbf{u}' \mathbf{v}$. Tehát a fenti definíció kiterjesztése a valósban használt definíciónak.

► E definícióval a vektorok hosszára vonatkozó tulajdonságok is érvényben maradnak, amit hamarosan belátunk (ld. ??? tétel).

Az adjungált tulajdonságai kiterjesztései a valós mátrixok transzponáltja tulajdonságainak, hisz valós mátrix konjugáltja megegyezik önmagával. Ez azonnal bizonyítja az alábbi tételt:

7.82. TÉTEL (AZ ADJUNGÁLT TULAJDONSÁGAI). *Legyenek $\mathbf{A}$ és $\mathbf{B}$ komplex mátrixok, c komplex szám. Ekkor*

a) $(\mathbf{A}^*)^* = \mathbf{A}$,
b) $(\mathbf{A} + \mathbf{B})^* = \mathbf{A}^* + \mathbf{B}^*$,
c) $(c\mathbf{A})^* = \bar{c}\mathbf{A}^*$
d) $(\mathbf{AB})^* = \mathbf{B}^* \mathbf{A}^*$.

Az adjungált tulajdonságaiból azonnal következik a következő tétel:

7.83. TÉTEL (A KOMPLEX SKALÁRIS SZORZÁS TULAJDONSÁGAI). *Legyen $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{C}^n$, és legyen $c \in \mathbb{C}$. Ekkor*

a) $\mathbf{u} \cdot \mathbf{v} = \overline{\mathbf{v} \cdot \mathbf{u}}$
b) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$,

c) $(c\mathbf{u}) \cdot \mathbf{v} = \bar{c}(\mathbf{u} \cdot \mathbf{v})$ *és* $\mathbf{u} \cdot (c\mathbf{v}) = c(\mathbf{u} \cdot \mathbf{v})$,
d) $\mathbf{u} \cdot \mathbf{u} > 0$, *ha* $\mathbf{u} \neq \mathbf{0}$, *és* $\mathbf{u} \cdot \mathbf{u} = 0$, *ha* $\mathbf{u} = \mathbf{0}$.

► Könnyen látható, hogy e tétel kiterjesztése a valós térbeli vektorokra kimondott 1.18. tételnek, bár első pillanatra úgy tűnhet, hogy ellentmond neki. Például valósban a skaláris szorzás kommutatív, itt nem, de a most kimondott változat érvényes valós vektorokra is, hisz valós vektor konjugáltja megegyezik önmagával. Hasonló állítható a *c)* tulajdonságról is.
► A *c)*-beli két tulajdonság bármelyike következik a másikból az *a)* alkalmazásával. Ha a skaláris szorzatot az $\mathbf{u} \cdot \mathbf{v} = \mathbf{v}^* \cdot \mathbf{u}$ képlettel definiáltuk volna, akkor a természetesebbnek ható $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v})$ összefüggés volna igaz.
► *d)*-ben az is az állítás része, hogy egy komplex vektor önmagával vett skaláris szorzata egyáltalán valós szám.
► A *d)* úgy is megfogalmazható, hogy $\mathbf{u} \cdot \mathbf{u} \geq 0$, és $\mathbf{u} \cdot \mathbf{u} = 0$ pontosan akkor áll fönn, ha $\mathbf{u} = \mathbf{0}$.

BIZONYÍTÁS. A bizonyítás a skaláris szorzás mátrixszorzatos alakjából, és a konjugált tulajdonságaiból azonnal adódik. Példaként megmutatjuk az *a)* bizonyítását:

$$\begin{aligned}\overline{\mathbf{v} \cdot \mathbf{u}} &= \overline{\mathbf{v}^* \mathbf{u}} = \overline{\overline{\mathbf{v}}' \mathbf{u}} = \mathbf{v}' \overline{\mathbf{u}} = \overline{\mathbf{u}}' \mathbf{v} = \mathbf{u}^* \mathbf{v} \\ &= \mathbf{u} \cdot \mathbf{v}\end{aligned}$$

A többi állítás hasonlóan bizonyítható. □

Önadjungált mátrixok Ahogy a transzponált fogalmának a – komplex skaláris szorzatot figyelembe vevő – kiterjesztése az adjungált, ugyanúgy a szimmetrikus mátrix fogalmának kiterjesztése az önadjungált mátrix. Szimmetrikus mátrix az, amelyik megegyezik saját transzponáltjával, önadjungált az, amelyik megegyezik saját adjungáltjával.

Az $\mathbf{A}$ komplex mátrix *önadjungált*, ha

$$\mathbf{A}^* = \mathbf{A}. \tag{7.27}$$

► Az önadjungált mátrixokat *Hermite-féle mátrixnak* is nevezik.
► Világos, hogy önadjungált mátrix főátlójában csak valósok állhatnak, mert csak azok egyeznek meg saját konjugáltjukkal.
► Minden valós szimmetrikus mátrix önadjungált, hisz a valós számok megegyeznek saját konjugáltjukkal. Sőt, mivel a nem valós komplex számok nem egyeznek meg saját konjugáltjukkal, ezért a komplex szimmetrikus mátrixok pontosan akkor önadjungáltak, ha minden elemük valós.

7.84. PÉLDA (ÖNADJUNGÁLT MÁTRIXOK). *Az*

$$\begin{bmatrix} 1 & \mathrm{i} & 1+\mathrm{i} \\ -\mathrm{i} & 2 & 2-3\mathrm{i} \\ 1-\mathrm{i} & 2+3\mathrm{i} & 3 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, \quad \begin{bmatrix} \mathrm{i} & 1+\mathrm{i} \\ 1-\mathrm{i} & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1+\mathrm{i} \\ 1+\mathrm{i} & 2 \end{bmatrix}$$

mátrixok közül az első kettő önadjungált, a harmadik nem, mert főátlójában nem minden szám valós, a negyedik sem, az viszont komplex szimmetrikus mátrix!

Távolság és a merőleges vetítés komplex terekben A komplex skaláris szorzás segítségével – a valós esethez hasonlóan – definiálható a komplex vektorok távolsága és szöge, és így a merőlegessége is.

A komplex $\mathbf{u} \in \mathbb{C}^n$ vektor *hossza*, vagy *abszolút értéke* $|\mathbf{u}| = \sqrt{\mathbf{u} \cdot \mathbf{u}}$, *két vektor távolsága* megegyezik különbségük hosszával, azaz $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ vektorok esetén $d(\mathbf{u}, \mathbf{v}) = |\mathbf{u} - \mathbf{v}|$. *Két vektor szögének* koszinusza, ahogy azt az (1.4) képlettel definiáltuk, $\cos(\mathbf{u}, \mathbf{v})_\angle := \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}||\mathbf{v}|}$. Így *két vektor merőleges*, ha skaláris szorzatuk 0.

Unitér mátrixok

Diszkrét Fourier-transzformált

Fourier-mátrixok Az N-edik komplex egységgyök hatványaiból képzett Vandermonde-mátrix kiemelkedően fontos szerepet kapott a modern műszaki alkalmazásokban. E mátrix alaptulajdonságainak megismeréséhez a Fourier-összegek együtthatói és helyettesítési értékei közti kapcsolaton keresztül közelítünk.

A Fourier-sorok komplex

$$\sum_{n=-\infty}^{\infty} c_n e^{nit}$$

alakja, és ezek

$$\sum_{n=0}^{N-1} c_k e^{nit} = c_0 + c_1 e^{\mathrm{i}t} + c_1 e^{2\mathrm{i}t} + \cdots + c_{N-1} e^{(N-1)it} \qquad (7.28)$$

alakú részletösszegei kulcsszerepet játszanak a periodikus, illetve a korlátos tartományon értelmezett függvények leírásában. A (7.28) összeget (diszkrét) *Fourier-összegnek* nevezzük.

7.85. ÁLLÍTÁS (FOURIER-ÖSSZEG HELYETTESÍTÉSI ÉRTÉKEI). *A (7.28) Fourier-összeg együtthatóihoz a Fourier-összegnek a $[0, 2\pi]$ intervallumot N egyenlő részre osztó 0, $\frac{2\pi}{N}$, $\frac{4\pi}{N}, \ldots, \frac{2(N-1)\pi}{N}$ pontokban vett helyettesítési értékeit rendelő leképezés lineáris, melynek mátrixa $[\mathrm{e}^{\frac{2\pi \mathrm{i}}{N} mn}]$ $(0 \le m, n < N)$.*

MEGOLDÁS. Először vizsgáljuk meg az $N = 3$ esetet. Az osztópontok: $t_0 = 0$, $t_1 = 2\pi/3$, $t_2 = 4\pi/3$. A Fourier-összeg $c_0 + c_1 e^{it} + c_2 e^{2it}$, ennek t_k-beli helyettesítési értékét jelölje y_k. Tehát

$$\begin{aligned}
y_0 &= c_0 + c_1 e^{\mathrm{i}0} + c_2 e^{2\mathrm{i}0} = c_0 + c_1 + c_2 \\
y_1 &= c_0 + c_1 e^{\frac{2\pi \mathrm{i}}{3}} + c_2 e^{\frac{4\pi \mathrm{i}}{3}} = c_0 + c_1 \varepsilon + c_2 \varepsilon^2 \\
y_2 &= c_0 + c_1 e^{\frac{4\pi \mathrm{i}}{3}} + c_2 e^{\frac{8\pi \mathrm{i}}{3}} = c_0 + c_1 \varepsilon^2 + c_2 \varepsilon^4
\end{aligned}$$

ahol $\varepsilon = e^{\frac{2\pi \mathrm{i}}{3}}$ a legkisebb pozitív argumentumú harmadik komplex egységgyököt jelöli. Világos, hogy a $(c_0, c_1, c_2) \mapsto (y_0, y_1, y_2)$ leképezés lineáris, melynek mátrixszorzat-alakja

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \varepsilon & \varepsilon^2 \\ 1 & \varepsilon^2 & \varepsilon^4 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}$$

Hasonlóan egyszerű az általános eset is, azonban még tekintsük át az $N = 2$ és az $N = 4$ eset is. $N = 2$ esetén $\varepsilon = e^{\frac{2\pi \mathrm{i}}{2}} = -1$ a primitív egységgyök, így

$$\begin{bmatrix} y_0 \\ y_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & \varepsilon \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix},$$

míg $N = 4$ esetén $\varepsilon = e^{\frac{2\pi \mathrm{i}}{4}} = \mathrm{i}$, tehát

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \varepsilon & \varepsilon^2 & \varepsilon^3 \\ 1 & \varepsilon^2 & \varepsilon^4 & \varepsilon^6 \\ 1 & \varepsilon^3 & \varepsilon^6 & \varepsilon^9 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \mathrm{i} & -1 & -\mathrm{i} \\ 1 & -1 & 1 & -1 \\ 1 & -\mathrm{i} & -1 & \mathrm{i} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}.$$

Általános esetben az n-edik osztópont $\frac{2n\pi}{N}$ $(n = 0, 1, 2, \ldots, N-1)$, a Fourier-összeg e pontbeli helyettesítési értékét y_n-nel jelölve

$$\begin{aligned}
y_0 &= c_0 + c_1 e^{\mathrm{i}0} + c_2 e^{2\mathrm{i}0} + \cdots + c_{N-1} e^{(N-1)\mathrm{i}0} = c_0 + c_1 + \cdots + c_{N-1} \\
y_1 &= c_0 + c_1 e^{\frac{2\pi \mathrm{i}}{N}} + c_2 e^{\frac{4\pi \mathrm{i}}{N}} + \cdots + c_{N-1} e^{\frac{2(N-1)\pi \mathrm{i}}{N}} \\
y_2 &= c_0 + c_1 e^{\frac{4\pi \mathrm{i}}{N}} + c_2 e^{\frac{8\pi \mathrm{i}}{N}} + \cdots + c_{N-1} e^{\frac{4(N-1)\pi \mathrm{i}}{N}} \\
&\;\;\vdots \\
y_{N-1} &= c_0 + c_1 e^{\frac{2\pi \mathrm{i}(N-1)}{N}} + c_2 e^{\frac{4\pi \mathrm{i}(N-1)}{N}} + \cdots + c_{N-1} e^{\frac{2\pi \mathrm{i}(N-1)^2}{N}}
\end{aligned}$$

Az $\varepsilon = \mathrm{e}^{2\pi\mathrm{i}/N}$ jelöléssel mátrixszorzat-alakban

$$\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{N-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & \varepsilon & \varepsilon^2 & \varepsilon^3 & \dots & \varepsilon^{N-1} \\ 1 & \varepsilon^2 & \varepsilon^4 & \varepsilon^6 & \dots & \varepsilon^{2(N-1)} \\ 1 & \varepsilon^3 & \varepsilon^6 & \varepsilon^9 & \dots & \varepsilon^{3(N-1)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \varepsilon^{N-1} & \varepsilon^{2(N-1)} & \varepsilon^{3(N-1)} & \dots & \varepsilon^{(N-1)^2} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{N-1} \end{bmatrix} \qquad \square$$

E példában szereplő együtthatómátrix egy Vandermonde-mátrix, mégpedig a $\mathbf{V}_N(1, \varepsilon, \varepsilon^2, \dots, \varepsilon^{N-1})$ mátrix, melyet $\mathbf{\Phi}_{N,\varepsilon}$-nal jelölünk. Fontos lesz még e mátrix konjugáltja, mely az $\omega = \bar{\varepsilon} = \mathrm{e}^{-2\pi\mathrm{i}/N}$ egységgyökhöz tartozó Vandermonde-mátrix. E két mátrixot *Fourier-mátrixnak* is nevezik. Tehát $[\mathbf{\Phi}_{N,\varepsilon}]_{kn} = \varepsilon^{kn}$, $[\mathbf{\Phi}_{N,\omega}]_{kn} = \omega^{kn}$ $(0 \le k, n < N)$, azaz részletesen

$$\mathbf{\Phi}_{N,\varepsilon} = \mathbf{V}_N(1, \varepsilon, \varepsilon^2, \dots, \varepsilon^{N-1}) = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & \varepsilon & \dots & \varepsilon^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \varepsilon^{N-1} & \dots & \varepsilon^{(N-1)^2} \end{bmatrix} \tag{7.29}$$

$$\mathbf{\Phi}_{N,\omega} = \mathbf{V}_N(1, \omega, \dots, \omega^{N-1}) = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & \omega & \dots & \omega^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \dots & \omega^{(N-1)^2} \end{bmatrix} \tag{7.30}$$

Az előző példában a Fourier-együtthatók ismeretében meghatároztuk a függvény helyettesítési értékeit. A gyakorlati alkalmazásokban főként a fordított sorrend érdekes, vannak y_k mért adataink, és keressük a c_k együtthatókat. Ehhez nyújt alapismereteket a következő tétel.

7.86. TÉTEL (A FOURIER-MÁTRIXOK TULAJDONSÁGAI). *Legyen N pozitív egész szám, $\varepsilon = \mathrm{e}^{2\pi\mathrm{i}/N}$, $\omega = \bar{\varepsilon} = \mathrm{e}^{-2\pi\mathrm{i}/N}$. Az $\mathbf{\Phi}_{N,\varepsilon}$ és $\mathbf{\Phi}_{N,\omega}$ Fourier-mátrixok a következő tulajdonságokkal rendelkeznek:*

a) Bármelyik Fourier-mátrix k-adik és $N-k$-adik sora egymás konjugáltja, páros N esetén pedig az $N/2$-edik sorvektor $(1, -1, 1, -1, \dots)$.

b) A két Fourier-mátrix egymás konjugáltja és egyúttal egymás adjungáltja is, azaz $\mathbf{\Phi}_{N,\omega} = \overline{\mathbf{\Phi}}_{N,\varepsilon} = \mathbf{\Phi}^*_{N,\varepsilon}$ és $\mathbf{\Phi}_{N,\varepsilon} = \overline{\mathbf{\Phi}}_{N,\omega} = \mathbf{\Phi}^*_{N,\omega}$

c) $\mathbf{\Phi}_{N,\varepsilon}\mathbf{\Phi}_{N,\omega} = N\mathbf{I}_N$, így $\mathbf{\Phi}_{N,\varepsilon}$ és $\mathbf{\Phi}_{N,\omega}$ invertálható,

$$\mathbf{\Phi}^{-1}_{N,\varepsilon} = \frac{1}{N}\mathbf{\Phi}_{N,\omega}, \quad \mathbf{\Phi}^{-1}_{N,\omega} = \frac{1}{N}\mathbf{\Phi}_{N,\varepsilon},$$

továbbá $\frac{1}{\sqrt{N}}\mathbf{\Phi}_{N,\varepsilon}$ és $\frac{1}{\sqrt{N}}\mathbf{\Phi}_{N,\omega}$ unitér.

BIZONYÍTÁS. *a)* Az $\mathbf{\Phi}_{N,\varepsilon}$ mátrix k-adik, illetve $N-k$-adik sorának n-edik eleme ε^{kn}, illetve $\varepsilon^{(N-k)n}$. Ez utóbbit átalakítva kapjuk, hogy

$$\varepsilon^{(N-k)n} = \varepsilon^{Nn}\varepsilon^{-kn} = (\varepsilon^{-1})^{kn} = \bar{\varepsilon}^{kn}.$$

Mivel minden pozitív páros N-re $\varepsilon^{\frac{N}{2}} = -1$, ezért az $N/2$-edik sorban -1 hatványai szerepelnek.

b) Mivel $\omega = \bar{\varepsilon}$, ezért $\omega^s = \bar{\varepsilon}^s$, tehát $\mathbf{\Phi}_{N,\omega} = \overline{\mathbf{\Phi}}_{N,\varepsilon}$. Másrészt $\mathbf{\Phi}_{N,\varepsilon}$ szimmetrikus, következésképp $\overline{\mathbf{\Phi}}_{N,\varepsilon} = \overline{\mathbf{\Phi}}'_{N,\varepsilon} = \mathbf{\Phi}^*_{N,\varepsilon}$.

c) Számítsuk ki az $\mathbf{\Phi}_{N,\varepsilon}\mathbf{\Phi}_{N,\omega}$ mátrixot! A szorzat k-adik sorának n-edik oszlopában a

$$\sum_{m=0}^{N-1} \varepsilon^{km}\omega^{mn} = \sum_{m=0}^{N-1} \varepsilon^{m(k-n)} = \sum_{m=0}^{N-1} (\varepsilon^{k-n})^m$$

összeg szerepel. Ha $k = n$, azaz $\varepsilon^{k-n} = 1$, akkor ez az összeg N, minden más esetben 0 (ld. a **??** állítást a komplex számokról szóló függelékben). Mindezek egyik következménye, hogy $\mathbf{\Phi}_{N,\varepsilon}$ és $\mathbf{\Phi}_{N,\omega}$ invertálhatók, $\mathbf{\Phi}_{N,\varepsilon}$ inverze $\frac{1}{N}\mathbf{\Phi}_{N,\omega}$, $\mathbf{\Phi}_{N,\omega}$ inverze $\frac{1}{N}\mathbf{\Phi}_{N,\varepsilon}$. A másik következmény, hogy

$$\left(\frac{1}{\sqrt{N}}\mathbf{\Phi}_{N,\varepsilon}\right)\left(\frac{1}{\sqrt{N}}\mathbf{\Phi}_{N,\omega}\right) = \mathbf{I}_N,$$

ami a *b)*-t is figyelembe véve épp azt jelenti, hogy

$$\left(\frac{1}{\sqrt{N}}\mathbf{\Phi}_{N,\varepsilon}\right)\left(\frac{1}{\sqrt{N}}\mathbf{\Phi}_{N,\varepsilon}\right)^* = \mathbf{I}_N,$$

azaz hogy $\frac{1}{\sqrt{N}}\mathbf{\Phi}_{N,\varepsilon}$ unitér. □

Diszkrét Fourier-transzformáció A diszkrét Fourier-transzformációra úgy gondolhatunk, mint egy – általában komplex – függvény helyettesítési értékeinek vektorához a függvény trigonometrikus összetevői együtthatóinak vektorát rendelő lineáris $\mathbb{C}^N \to \mathbb{C}^N$ leképezésre.

A 7.85. példában egy Fourier-összeg együtthatóival kifejeztük a függvény megadott helyeken fölvett értékeit. A fordított irány sokkal érdekesebb: ismerjük egy f függvény N különböző megadott helyen fölvett értékét, és meg van adva N lineárisan független függvény. Olyan lineáris kombinációjuk együtthatóit keressük e függvényeknek, mely lineáris kombináció a megadott helyeken megegyezik f-fel. Mi a következőkben definiálandó diszkrét Fourier-transzformáció esetén az

$$f(t) = \frac{1}{N}\sum_{n=0}^{N-1} c_n e^{nit}$$

függvényből indulunk ki, a megadott helyek a $[0, 2\pi]$ intervallumot N részre osztó $2k\pi/N$ ($k = 0, 1, \ldots, N-1$) pontok. A $(c_0, c_1, \ldots, c_{N-1}) \mapsto (y_0, y_1, \ldots, y_{N-1})$ leképezés inverzét fogjuk diszkrét Fourier-transzformáltnak nevezni. Ennek mátrixa $\mathbf{\Phi}_{N,\omega}$, amelyre a továbbiakban az $\mathbf{F}_N$ jelölést is használjuk. E megközelítésből az f függvény teljesen elhagyható, hisz a lényeg az, hogy egy szám-N-eshez hozzárendelünk egy másikat!

7.87. definíció (Diszkrét Fourier-transzformáció (DFT)). *Az* $F_N : \mathbb{C}^N \to \mathbb{C}^N : \mathbf{x} \mapsto \mathbf{X} = \mathbf{F}_N\mathbf{x}$ *leképezést* diszkrét Fourier-transzformációnak *nevezzük.*

► A diszkrét Fourier-transzformáció tehát a (7.30) képlettel megadott $\mathbf{F}_N = \mathbf{\Phi}_{N,\omega}$ mátrixhoz tartozó mátrixleképezés.

► A leképezést kifejtve koordinátánként:

$$X_k = \sum_{n=0}^{N-1} x_n \mathrm{e}^{\frac{-2\pi\mathrm{i}}{N}kn} = \sum_{n=0}^{N-1} x_n\omega^{kn} \qquad (\omega = \mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}}). \tag{7.31}$$

► Az F_N transzformáció mátrixszorzatos alakja

$$F_N : \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{bmatrix} \mapsto \begin{bmatrix} X_0 \\ X_1 \\ \vdots \\ X_{N-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & \omega & \dots & \omega^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \dots & \omega^{(N-1)^2} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{bmatrix}.$$

► E témában elterjedt jelölések: a transzformálandó vektor dimenzióját nagy N jelöli, a képvektort a transzformálandó vektor nagybetűs változata jelöli, azaz $\mathbf{x}$ képe $\mathbf{X}$, $\mathbf{y}$ képe $\mathbf{Y}$, stb., a vektorok koordinátái 0-tól $N-1$-ig vannak indexelve.

► A diszkrét Fourier-transzformációt gyakran a Fourier-mátrixok valamelyikének egy másik konstansszorosával definiálják. Előfordul az unitér $\frac{1}{\sqrt{N}}\mathbf{F}_N$, az $\frac{1}{N}\mathbf{F}_N$ vagy a $\mathbf{\Phi}_{N,\varepsilon}$ mátrix is a transzformáció mátrixaként, sőt van aki minden olyan $\mathbf{\Phi}_{N,\hat{\varepsilon}}$ mátrixot egy DFT mátrixának tekint, ahol $\hat{\varepsilon}$ primitív N-edik egységgyök.

► Az általunk adott definíció a legelterjedtebb, a legtöbb ismert szoftver is ezt használja. Ennek oka e definíciónak a folytonos Fourier-transzformálttal való szorosabb kapcsolata, és a jelfeldolgozásban is ezt használják leginkább. Más alkalmazásokhoz viszont megfelelőbb lehet valamelyik fent említett másik definíció.

► Konkrétan az $\mathbf{F}_1$, $\mathbf{F}_2$, $\mathbf{F}_4$ és $\mathbf{F}_8$ mátrixok:

$$\mathbf{F}_1 = [1], \quad \mathbf{F}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad \mathbf{F}_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -\mathrm{i} & -1 & \mathrm{i} \\ 1 & -1 & 1 & -1 \\ 1 & \mathrm{i} & -1 & -\mathrm{i} \end{bmatrix},$$

$$\mathbf{F}_8 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \frac{1-\mathrm{i}}{\sqrt{2}} & -\mathrm{i} & \frac{-1-\mathrm{i}}{\sqrt{2}} & -1 & \frac{-1+\mathrm{i}}{\sqrt{2}} & \mathrm{i} & \frac{1+\mathrm{i}}{\sqrt{2}} \\ 1 & -\mathrm{i} & -1 & \mathrm{i} & 1 & -\mathrm{i} & -1 & \mathrm{i} \\ 1 & \frac{-1-\mathrm{i}}{\sqrt{2}} & \mathrm{i} & \frac{1-\mathrm{i}}{\sqrt{2}} & -1 & \frac{1+\mathrm{i}}{\sqrt{2}} & -\mathrm{i} & \frac{-1+\mathrm{i}}{\sqrt{2}} \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & \frac{-1+\mathrm{i}}{\sqrt{2}} & -\mathrm{i} & \frac{1+\mathrm{i}}{\sqrt{2}} & -1 & \frac{1-\mathrm{i}}{\sqrt{2}} & \mathrm{i} & \frac{-1-\mathrm{i}}{\sqrt{2}} \\ 1 & \mathrm{i} & -1 & \mathrm{i} & 1 & \mathrm{i} & -1 & \mathrm{i} \\ 1 & \frac{1+\mathrm{i}}{\sqrt{2}} & \mathrm{i} & \frac{-1+\mathrm{i}}{\sqrt{2}} & -1 & \frac{-1-\mathrm{i}}{\sqrt{2}} & -\mathrm{i} & \frac{1-\mathrm{i}}{\sqrt{2}} \end{bmatrix}$$

7.88. TÉTEL (A DFT TULAJDONSÁGAI). *Tekintsük a diszkrét F_N Fourier-transzformációt, és legyen az $\mathbf{x} = (x_0, x_1, \dots, x_{N-1})$ vektor képe $\mathbf{X} = (X_0, X_1, \dots, X_{N-1})$. Ekkor a következők igazak:*

a) *Konstans vektor képe inpulzusvektor (melynek a nulladikat kivéve mindegyik koordinátája 0), és fordítva, konkrétan*

$$F_N(c, c, \dots, c) = (Nc, 0, \dots, 0), \quad F_N(c, 0, \dots, 0) = (c, c, \dots, c).$$

ahol $c \in \mathbb{C}$ tetszőleges konstans.

b) *Ha $\mathbf{x}$ valós vektor, akkor $X_{N-k} = \overline{X}_k$.*

c) *Az F_N transzformáció invertálható, inverze (IDFT) többféle felírásban:*

$$\mathbf{x} = F_N^{-1}\mathbf{X} = \frac{1}{N}\mathbf{\Phi}_{N,\varepsilon}\mathbf{X}, \quad x_k = \frac{1}{N}\sum_{n=0}^{N-1} X_n \varepsilon^{kn} = \frac{1}{N}\sum_{n=0}^{N-1} X_n \mathrm{e}^{\frac{2\pi \mathrm{i}}{N}kn}.$$

BIZONYÍTÁS. *a)* Az állítás első részének bizonyítása közvetlenül leolvasható az $\mathbf{\Phi}_{N,\varepsilon}(c\mathbf{\Phi}_{N,\omega}) = cN\mathbf{I}_N$ szorzat első oszlopából, második része a $c\mathbf{\Phi}_{N,\varepsilon}$ mátrix első oszlopából. De a 7.86. tétel bizonyításában is használt **??** állításra hivatkozva közvetlenül is azonnal adódik.

b) A (7.31) képletet használva

$$\begin{aligned} X_{N-k} &= \sum_{n=0}^{N-1} x_n \omega^{(N-k)n} = \sum_{n=0}^{N-1} x_n \omega^{-kn} \\ &= \sum_{n=0}^{N-1} x_n \overline{\omega}^{kn} = \overline{\sum_{n=0}^{N-1} x_n \omega^{kn}} = \overline{X}_k \end{aligned}$$

c) Az invertálhatóság azonnal adódik abból, hogy a Fourier-mátrixok Vandermonde-mátrixok is egyúttal, melynek determinánsa nem 0. A tételbeli mindegyik összefüggés azonnali következménye az $\mathbf{F}_N^{-1} = \mathbf{\Phi}_{N,\omega}^{-1} = \frac{1}{N}\mathbf{\Phi}_{N,\varepsilon}$ képletnek. □

7.89. PÉLDA (DFT KISZÁMÍTÁSA). *Határozzuk meg az $\mathbf{x} = (1, \mathrm{i}, \mathrm{i}, 2)$ vektor diszkrét Fourier-transzformáltját!*

MEGOLDÁS. $N = 4$, így

$$\mathbf{X} = F_4\mathbf{x} = \mathbf{F}_4\mathbf{x} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -\mathrm{i} & -1 & \mathrm{i} \\ 1 & -1 & 1 & -1 \\ 1 & \mathrm{i} & -1 & -\mathrm{i} \end{bmatrix} \begin{bmatrix} 1 \\ \mathrm{i} \\ \mathrm{i} \\ 2 \end{bmatrix} = \begin{bmatrix} 3+2\mathrm{i} \\ 2+\mathrm{i} \\ -1 \\ -3\mathrm{i} \end{bmatrix}.$$

□

Periodikus összetevők szűrése Műszaki alkalmazásokban gyakran előfordul, hogy egy periodikus függvénnyel leírható jelhez magasabb frekvenciájú zaj adódik, amit utólag ki szeretnénk „szűrni”. Ez egy DFT-IDFT párral könnyen elvégezhető.

A szűrés általános modellje három lépésből áll, melyet az alábbi séma szemléltet:

$$\mathbf{x} \xrightarrow{\quad\text{DFT}\quad} \mathbf{X} \xrightarrow{\quad\text{„szűrés”}\quad} \hat{\mathbf{X}} \xrightarrow{\quad\text{IDFT}\quad} \hat{\mathbf{x}}$$

A műszaki gyakorlatban „szűrésen” sokféle transzformációt értenek, mely az $\mathbf{X}$ vektort az $\hat{\mathbf{X}}$-ba képzi. Mi csak a legegyszerűbb esettel foglalkozunk, $\mathbf{X}$ bizonyos koordinátáinak elhagyásával (kiszűrésével).

A következőkben egy mesterkélten leegyszerűsített, fejben számolva is követhető példát mutatunk a DFT e tipikus alkalmazására.

7.90. példa (Magas frekvenciájú összetevők szűrése). *Adva van egy p szerint periodikus függvény $t_k = kp/6$ ($k = 0, 1, \dots, 5$) helyeken fölvett függvényértékeinek $\mathbf{x} = (4, 1, -2, -2, -2, 1)$ vektora. Bontsuk fel e függvényt egy p szerint periodikus trigonometrikus függvény, és p/m periódusú függvények (zaj) összegére ($m > 1$ egész).*

Megoldás. Az $\mathbf{x}$ vektor diszkrét Fourier-transzformáltja

$$\mathbf{X} = F_6\mathbf{x} = \mathbf{F}_{6,\omega}\mathbf{x} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \omega & \omega^2 & -1 & \omega^4 & \omega^5 \\ 1 & \omega^2 & \omega^4 & 1 & \omega^2 & \omega^4 \\ 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & \omega^4 & \omega^2 & 1 & \omega^4 & \omega^2 \\ 1 & \omega^5 & \omega^4 & -1 & \omega^2 & \omega \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ -2 \\ -2 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 9 \\ 3 \\ 0 \\ 3 \\ 9 \end{bmatrix}$$

ahol kihasználtuk, hogy $\omega^6 = 1$, $\omega + \omega^5 = 1$, $\omega^2 + \omega^4 = -1$. Például

$$X_2 = [\mathbf{F}_{6,\omega}]_{2*}\,\mathbf{x} = 4 + \omega^2 - 2\omega^4 - 2 - 2\omega^2 + \omega^4 = 2 - \omega^2 - \omega^4 = 3.$$

Mivel $\mathbf{x}$ valós, ezért a 7.88. tétel *b)* pontja szerint $X_{N-k} = \overline{X}_k$, amit azonnal ellenőrizhetünk is. Az $\mathbf{x}$ vektor „mögött” lévő p szerint periodikus függvény e modellben

$$x(t) = \frac{1}{6}\sum_{n=0}^{5} X_n e^{n\mathrm{i}\frac{2\pi}{p}t} = \frac{1}{6}\left(9e^{\mathrm{i}\frac{2\pi}{p}t} + 3e^{2\mathrm{i}\frac{2\pi}{p}t} + 3e^{4\mathrm{i}\frac{2\pi}{p}t} + 9e^{5\mathrm{i}\frac{2\pi}{p}t}\right)$$

A p értékének valójában semmi szerepe, mert e függvényt csak a $k\frac{p}{n}$ ($k = 0, 1, \dots, 5$) pontokban értékeljük ki, így a fenti összegben csak a

$$e^{\mathrm{i}\frac{2\pi}{p}\frac{pn}{6}} = (e^{\mathrm{i}\frac{2\pi}{6}})^n = \varepsilon^n$$

értékek szerepelnek. Ezekre használható a

$$\begin{aligned} \varepsilon^n + \varepsilon^{6-n} &= (e^{\mathrm{i}\frac{2\pi}{6}})^n + (e^{-\mathrm{i}\frac{2\pi}{6}})^n = (e^{\mathrm{i}\frac{2\pi n}{6}}) + (e^{-\mathrm{i}\frac{2\pi n}{6}}) \\ &= 2\cos n\frac{\pi}{3} = 2\cos\frac{2\pi}{p}\frac{pn}{6} \end{aligned}$$

összefüggés, így

$$x(t) = \frac{1}{6}\left(18\cos\frac{2\pi}{p}t + 6\cos 2\frac{2\pi}{p}t\right) = 3\cos\frac{2\pi}{p}t + \cos 2\frac{2\pi}{p}t.$$

A függvény tehát $3\cos\frac{2\pi}{p}t$, a „zaj” $\cos 2\frac{2\pi}{p}t$. □

Gyors Fourier-transzformáció A diszkrét Fourier-transzformáció gyors kiszámítására konstruált algoritmusoknak döntő szerepük van a mai kultúránk alapját jelentő digitális technika fejlődésében.

A diszkrét Fourier-transzformált kiszámításához, azaz az n-edrendű Fourier-mátrixszal való szorzás kiszámításához n^2 szorzás elvégzésére van szükség. Bármely olyan algoritmust, mely e transzformáció eredményét $O(n\log n)$, azaz konstansszor $n\log n$ lépésben elvégzi, *gyors Fourier-transzformációnak* nevezzük. Sok változata létezik, mi csak az elsőként publikált, legismertebbet ismertetjük.

A transzformáció gyorsaságának becsléséhez most minden aritmetikai művelet elvégzésének idejét tekintsünk azonosnak. A DFT kiszámítására, azaz a Fourier-mátrixszal való szorzáshoz minden sorban N szorzás, $N-1$ összeadás kell, és N sor van, így a szükséges műveletek száma $N(2N-1)$.

Az egyszerűség kedvéért legyen a továbbiakban N kettőhatvány, és csoportosítsuk az X_k-t megadó összeget az indexek paritása szerint, azaz külön adjuk össze a páros és külön a páratlan indexűeket. Vegyük észre, hogy ez az összeg két fele akkora méretű Fourier-transzformációból megkapható:

$$\begin{aligned}
X_k &= \sum_{n=0}^{N-1} x_n \mathrm{e}^{\frac{-2\pi\mathrm{i}}{N}kn} = \sum_{n=0}^{N-1} x_n\omega_N^{kn}\\
&= \sum_{n=0}^{N/2-1} x_{2n}\mathrm{e}^{\frac{-2\pi\mathrm{i}}{N}2nk} + \sum_{n=0}^{N/2-1} x_{2n+1}\mathrm{e}^{\frac{-2\pi\mathrm{i}}{N}(2n+1)k}\\
&= \sum_{n=0}^{N/2-1} x_{2n}\mathrm{e}^{\frac{-2\pi\mathrm{i}}{N/2}nk} + \mathrm{e}^{\frac{-2\pi\mathrm{i}}{N}k}\sum_{n=0}^{N/2-1} x_{2n+1}\mathrm{e}^{\frac{-2\pi\mathrm{i}}{N/2}nk}\\
&= \sum_{n=0}^{N/2-1} x_{2n}\omega_{N/2}^{nk} + \omega_N^k\sum_{n=0}^{N/2-1} x_{2n+1}\omega_{N/2}^{nk}\\
&= E_k + \omega_N^k O_k.
\end{aligned}$$

Hogy különbséget tegyünk az N és $N/2$ dimenziós vektorok transzformációi közt, az N-edik és $N/2$-edik egységgyököt jelölje

$$\omega_N = \mathrm{e}^{\frac{-2\pi\mathrm{i}}{N}}, \text{ és } \omega_{N/2} = \mathrm{e}^{\frac{-2\pi\mathrm{i}}{N/2}}.$$

És ezen a ponton azzal tudjuk csökkenteni a számításokat, hogy mivel az E_k és O_k összegek $N/2$ szerint periodikusak, így $k \geq N/2$ esetén E_k és O_k értékét már nem kell újra számolni, ugyanis

$$E_{k+N/2} = E_k, \quad O_{k+N/2} = O_k,$$

és az ω_N^k együttható is újrahasznosítható:

$$\omega_N^{k+N/2} = \mathrm{e}^{\frac{-2\pi\mathrm{i}}{N}(k+N/2)} = \mathrm{e}^{\frac{-2\pi\mathrm{i}}{N}(k)}\mathrm{e}^{\frac{-2\pi\mathrm{i}}{N}\frac{N}{2}} = -\mathrm{e}^{\frac{-2\pi\mathrm{i}}{N}(k)} = -\omega_N^k.$$

Ezeket összevetve tehát $k < N/2$ esetén

$$\begin{aligned} X_k &= E_k + \omega_N^k O_k, \\ X_{k+N/2} &= E_k - \omega_N^k O_k. \end{aligned}$$

Így, ha E_k és O_k már ki van számolva, X_k és $X_{k+N/2}$ kiszámításához csak egy szorzásra, egy összeadásra és egy kivonásra van szükség, azaz az X_k $(0 \le k < N)$ együtthatók $3N/2$ művelettel megkaphatók. Ezután rekurzív módon E_k és O_k kiszámítását is ugyanígy végezzük: mivel fele akkora a vektor, de kettő van belőle, itt is $3N/2$ műveletre lesz szükség. Mivel N kettőhatvány, például $N = 2^s$, így $s = \log_2 N$-szer kell megismételnünk ezt a lépést, vagyis a teljes transzformáció műveletigénye $\frac{3}{2}N\log_2 N$. Konkrétan néhány N esetén:

N	$2^4 = 16$	$2^8 = 256$	$2^{10} = 1024$	$2^{16} = 65536$
DFT	496	130816	2096128	8589869056
FFT	96	3072	15360	1572864
hányados	> 5	> 42	> 136	> 5461

A műveletigény fenti számításában nem vettük figyelembe ω_N^k kiszámításának költségeit. Ha e hatvány kiszámítása C aritmetikai művelettel egyenértékű, akkor is csak $\frac{C+3}{2}N\log_2 N$ műveletre van szükségünk. Ezzel tehát bizonyítottuk a következő tételt:

A tételbeli eljárást Gauss már ismerte és 1805-ben használta a másodiknak fölfedezett Pallas és a harmadiknak fölfedezett Juno nevű kisbolygó pályájának kiszámításához. A felező eljárást Danielson és Lánczos 1942-ben újra fölfedezték, de ők sem vizsgálták az algoritmus sebességét. Az FFT ismertté és népszerűvé Cooley és Tukey 1965-ben megjelent cikke után vált.

7.91. TÉTEL (GYORS FOURIER-TRANSZFORMÁCIÓ). *Létezik olyan algoritmus, mely egy N-dimenziós vektor diszkrét Fourier-transzformáltját legföljebb $O(N\log_2 N)$ aritmetikai művelet elvégzésével kiszámolja.*

A fenti bizonyításban szereplő algoritmus pszeudokódja:

Mivel e transzformáció is lineáris leképezésekből áll, a gyors Fourier-transzformáció mátrixszorzat-alakba is fölírható:

$$\mathbf{F}_N = \mathbf{\Delta}_N \begin{bmatrix} \mathbf{F}_{N/2} & \mathbf{O} \\ \mathbf{O} & \mathbf{F}_{N/2} \end{bmatrix} \mathbf{\Pi}_N,$$

ahol $\mathbf{\Pi}_N$ az a permutációs mátrix, mely előre veszi a páros indexű elemeket, $\mathbf{\Delta}_N$ pedig a „fél” transzformáltakat összeadó, és a páratlan indexűeket egy ω-hatvánnyal beszorzó mátrix. Ezek kisebb indexű

```
function FFT(x)
    N ← dim(x)
    X legyen N-dimenziós vektor
    if N = 1 then
        X_0 ← x_0
    else
        y ← x páros indexű elemei
        z ← x páratlan indexű elemei
        Y ← FFT(y)
        Z ← FFT(z)
        for k ← 0 to N/2 − 1 do
            E ← Y_k
            O ← e^(−2πi/N k) Z_k
            X_k ← E + O
            X_{k+N/2} ← E − O
    return X
```

7.22. ábra: FFT algoritmus. A rekurzív függvény bemenete egy tetszőleges komplex $\mathbf{x}$ vektor, kimenete a diszkrét Fourier-transzformált $\mathbf{X}$ vektor.

példányai:

$$\mathbf{\Pi}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{\Pi}_8 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{\Delta}_4 = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -\mathrm{i} \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & \mathrm{i} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_2 & \mathbf{D}_2 \\ \mathbf{I}_2 & -\mathbf{D}_2 \end{bmatrix} \quad \mathbf{\Delta}_8 = \begin{bmatrix} \mathbf{I}_4 & \mathbf{D}_4 \\ \mathbf{I}_4 & -\mathbf{D}_4 \end{bmatrix}$$

A $\mathbf{\Delta}$ mátrixokban szereplő diagonális mátrixok az egységmátrixok, és az ω hatványait tartalmazó $\mathbf{D}$ mátrixok, ahol $\mathbf{D}_k = \operatorname{diag}(1, \omega, \omega^2, \dots, \omega^{k-1})$. Tehát például

$$\begin{aligned} \mathbf{F}_8 &= \mathbf{\Delta}_8 \begin{bmatrix} \mathbf{F}_4 & \mathbf{O} \\ \mathbf{O} & \mathbf{F}_4 \end{bmatrix} \mathbf{\Pi}_8 \\ &= \mathbf{\Delta}_8 \begin{bmatrix} \mathbf{\Delta}_4 & \mathbf{O} \\ \mathbf{O} & \mathbf{\Delta}_4 \end{bmatrix} \begin{bmatrix} \mathbf{F}_2 & \mathbf{O} & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{F}_2 & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{F}_2 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{O} & \mathbf{F}_2 \end{bmatrix} \begin{bmatrix} \mathbf{\Pi}_4 & \mathbf{O} \\ \mathbf{O} & \mathbf{\Pi}_4 \end{bmatrix} \mathbf{\Pi}_8. \end{aligned}$$

Látjuk, hogy a rekurzió következtében a transzformálandó vektort először a $\mathbf{\Pi}$-mátrixokból álló blokkmátrixokkal kell szorozni. E mátrixok szorzata is permutációmátrix. Hatását e konkrét esetben kiszámoljuk a fent megadott $\mathbf{\Pi}_4$ és $\mathbf{\Pi}_8$ mátrixok behelyettesítésével:

$$\begin{bmatrix} \mathbf{\Pi}_4 & \mathbf{O} \\ \mathbf{O} & \mathbf{\Pi}_4 \end{bmatrix} \mathbf{\Pi}_8 \mathbf{x} = \begin{bmatrix} 1&0&0&0&0&0&0&0 \\ 0&0&0&0&1&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&0&0&0&0&0&1&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&0&0&0&0&0&1 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = \begin{bmatrix} x_0 \\ x_4 \\ x_2 \\ x_6 \\ x_1 \\ x_5 \\ x_3 \\ x_7 \end{bmatrix}.$$

Ez első pillanatban áttekinthetetlen permutációnak tűnik, de valójában egy igen egyszerűen leírható transzformációt kapunk: a transzformálandó $\mathbf{x}$ vektor k-adik koordinátáját ($k = 0, 1, \ldots, N-1$) a j-edik helyébe viszi, ha j bináris alakja éppen fordítottja k bináris alakjának. Például ha $N = 8$, és $k = 6$, akkor x_6 a harmadik koordináta helyére kerül a permutáció során, mivel $6 = 110_2$, és ennek fordítottja $011_2 = 3$. Ennek igazolása rendkívül egyszerű, ha észrevesszük, hogy az i-edik $\mathbf{\Delta}$-mátrixszal való szorzás épp jobbról az első i koordináta szerinti lexikografikus sorrendbe rendezi az elemeket. Ennek szemléltetésére elég az a **??** ábrán bemutatott $N = 8$ eset vizsgálata.

Vektorok konvolúciója Vektorok konvolúciója igen sok helyen felmerül: a polinomok szorzásától kezdve az olyan transzformációkig, ahol egy koordinátát szomszédainak egy rögzített lineáris kombinációjával kell helyettesíteni. A gyors Fourier-transzformációval hatékonyan számolható.

Megoldások

7.1. Igen, $A : \mathbb{R}^3 \to \mathbb{R}$, mátrixa $\mathbf{A} = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}$.

7.2. Nem, az $A : \mathbf{x} \mapsto \mathbf{a} + \mathbf{x}$ leképezés a $\mathbf{0}$ vektort nem a $\mathbf{0}$-ba képzi, így nem lehet mátrixleképezés!

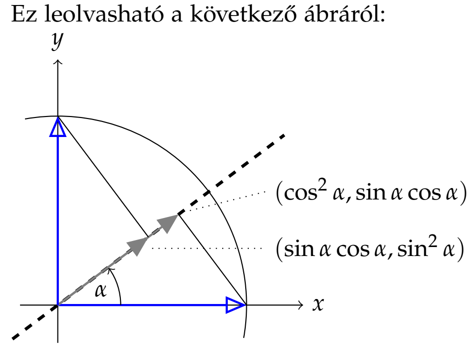
7.4. Igen, a mátrix

$$\mathbf{A} = \mathbf{a}(\mathbf{a} \cdot \mathbf{x}) = \mathbf{a}(\mathbf{a}'\mathbf{x}) = (\mathbf{a}\mathbf{a}')\mathbf{x} = \begin{bmatrix} a_1^2 & a_1a_2 & a_1a_3 \\ a_2a_1 & a_2^2 & a_2a_3 \\ a_1a_3 & a_2a_3 & a_3^2 \end{bmatrix}$$

7.8. *a)* A bizonyítandó képlet:

$$\mathbf{P} = \begin{bmatrix} \cos^2\alpha & \sin\alpha\cos\alpha \\ \sin\alpha\cos\alpha & \sin^2\alpha \end{bmatrix}.$$

Ez leolvasható a következő ábráról:



Az itt látható két derékszögű háromszög befogóinak hossza $\cos\alpha$, illetve $\sin\alpha$, és így például a $\cos\alpha$ hosszú szakasz két tengelyvetülete $\cos^2\alpha$ és $\cos\alpha\sin\alpha$ hosszú, tehát $\mathbf{i}$ képe $(\cos^2\alpha, \cos\alpha\sin\alpha)$. Hasonlóan $\mathbf{j}$ képe $(\sin\alpha\cos\alpha, \sin^2\alpha)$. E két oszlopvektorból álló mátrix pedig valóban megegyezik a fent megadottal.

b) Az $\mathbf{i}$ és $\mathbf{j}$ vetülete az egyenesre

$$\operatorname{proj}_{\mathbf{b}} \mathbf{i} = \frac{\mathbf{i}\cdot\mathbf{b}}{\mathbf{b}\cdot\mathbf{b}}\mathbf{b} = \frac{1}{b_1^2+b_2^2}\begin{bmatrix} b_1^2 \\ b_1b_2 \end{bmatrix} \quad \text{és}$$

$$\operatorname{proj}_{\mathbf{b}} \mathbf{j} = \frac{\mathbf{j}\cdot\mathbf{b}}{\mathbf{b}\cdot\mathbf{b}}\mathbf{b} = \frac{1}{b_1^2+b_2^2}\begin{bmatrix} b_1b_2 \\ b_2^2 \end{bmatrix}.$$

E két vektor egymás mellé írásával kapott mátrix lesz a leképezés mátrixa:

$$\frac{1}{b_1^2+b_2^2}\begin{bmatrix} b_1^2 & b_1b_2 \\ b_1b_2 & b_2^2 \end{bmatrix}.$$

Ez megegyezik a (7.11) képlettel, azaz

$$\frac{1}{b_1^2+b_2^2}\begin{bmatrix} b_1^2 & b_1b_2 \\ b_1b_2 & b_2^2 \end{bmatrix} = \begin{bmatrix} \cos^2\alpha & \sin\alpha\cos\alpha \\ \sin\alpha\cos\alpha & \sin^2\alpha \end{bmatrix},$$

ugyanis ha a $\mathbf{b}$ vektor x-tengellyel bezárt szöge α, akkor $\cos\alpha = b_1/\sqrt{b_1^2+b_2^2}$, és $\sin\alpha = b_2/\sqrt{b_1^2+b_2^2}$.

7.9. $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \end{bmatrix}$, így a (7.14) képlet szerint

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}^+ = \begin{bmatrix} 1 \\ 1 \end{bmatrix}\left(\begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)^{-1}\begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & 0 \end{bmatrix}$$

7.10. $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \end{bmatrix}$, így a (7.14) képlet szerint

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}^+ = \begin{bmatrix} 1 \\ 1 \end{bmatrix}\left(\begin{bmatrix} 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)^{-1}\begin{bmatrix} 1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{bmatrix}$$

7.11. $\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}^+ = \begin{bmatrix} 1/10 & 2/10 \\ 1/10 & 2/10 \end{bmatrix}$

7.12. $\begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix}^+ = \begin{bmatrix} 0 & 0 \\ 1/4 & 1/4 \end{bmatrix}$

7.13. $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}^+ = \begin{bmatrix} 2/3 & -1/3 \\ 1/3 & 1/3 \\ -1/3 & 2/3 \end{bmatrix}$

7.14. $\begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 0 & 0 \end{bmatrix}^+ = \begin{bmatrix} 1/10 & 2/10 & 0 \\ 1/10 & 2/10 & 0 \end{bmatrix}$

7.15. $\begin{bmatrix} 1 & 0 \\ 2 & 2 \\ 0 & 1 \end{bmatrix}^+ = \begin{bmatrix} 5/9 & 2/9 & -4/9 \\ -4/9 & 2/9 & 5/9 \end{bmatrix}$

7.16. $\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}^+ = \begin{bmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}$

7.17. $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}^+ = 1/30\begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}$

7.18. A bizonyítás lényegében megegyezik a 7.48. tétel bizonyításával, csak most az utolsó képlet is könnyen igazolhatóvá válik, hisz ha $\mathbf{A}$ teljes sorrangú, akkor e feladat szerinti egyik felbontása $\mathbf{A} = \mathbf{I}_r\mathbf{A}$, így

$$\begin{aligned}\mathbf{A}^+ &= \mathbf{A}'(\mathbf{A}\mathbf{A}')^{-1}(\mathbf{I}_r'\mathbf{I}_r)^{-1}\mathbf{I}_r' \\ &= \mathbf{A}'(\mathbf{A}\mathbf{A}')^{-1}.\end{aligned}$$

7.19. Mivel az $\mathbf{A}$ mátrix egy olyan $\mathbf{XY}$ felbontásával van megadva, melyben $\mathbf{X}$ teljes oszlop-, $\mathbf{Y}$ teljes sorrangú, ezért használható a 7.18. feladat eredménye. Így

$$\mathbf{A}^+ = \mathbf{Y}'(\mathbf{Y}\mathbf{Y}')^{-1}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \frac{1}{3}\begin{bmatrix} -1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 2 & -1 & 0 & 1 \end{bmatrix}$$

7.20. Az előző feladathoz hasonlóan a 7.18. feladat szerint

$$\mathbf{A}^+ = \mathbf{Y}'(\mathbf{Y}\mathbf{Y}')^{-1}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \frac{1}{15}\begin{bmatrix} -4 & 5 & 1 \\ -4 & 5 & 1 \\ 3 & 0 & 3 \\ 7 & -5 & 2 \end{bmatrix}$$

7.21. A 7.18. feladat szerint

$$\mathbf{A}^+ = \mathbf{Y}'(\mathbf{Y}\mathbf{Y}')^{-1}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \frac{1}{132}\begin{bmatrix} -32 & 34 & 2 & 4 \\ 1 & 1 & 2 & 4 \\ 34 & -32 & 2 & 4 \end{bmatrix}$$

7.22. *1. megoldás:* Ha $\mathbf{A}$ merőleges vetítés mátrixa, akkor $\mathbf{A}^2 = \mathbf{A}$ okán $\mathcal{S}(\mathbf{A}) = \mathcal{O}(\mathbf{A})$, és a sortér minden $\mathbf{x}$ vektorára $\mathbf{A}\mathbf{x} = \mathbf{x}$, így $\mathbf{A}^3\mathbf{x} = \mathbf{x}$ is igaz lesz. Másrészt $\mathbf{A}' = \mathbf{A}$ miatt $\mathcal{N}(\mathbf{A}') = \mathcal{N}(\mathbf{A})$, tehát minden $\mathbf{z} \in \mathcal{N}(\mathbf{A}')$ esetén $\mathbf{A}\mathbf{z} = \mathbf{0}$, és $\mathbf{A}^+\mathbf{z} = \mathbf{0}$ is fönnáll, tehát $\mathbf{A}$ és $\mathbf{A}^+$ hatása az $\mathcal{O}(\mathbf{A})$ és annak merőleges kiegészítő alterén is megegyezik, így a két mátrix azonos.

1. megoldás: A Penrose-tétel alapján azt kell ellenőriznünk, hogy $\mathbf{A}$ a négy feltétel mindegyikét teljesíti, mint pszeudoinverz. Ez igaz, hisz $\mathbf{A}' = \mathbf{A} = \mathbf{A}^2$ miatt $\mathbf{A}^3 = \mathbf{A}$, és $(\mathbf{A}^2)' = \mathbf{A}^2$.

Az állítás megfordítása nem igaz. Például az $\mathbf{A} = -\mathbf{I}$ mátrixra $\mathbf{A}^+ = \mathbf{A}$, de a $-\mathbf{I}$ mátrixra $(-\mathbf{I})^2 \neq -\mathbf{I}$.

7.23. Az első állítás azonnal következik a (7.15) egyenlőségből, ugyanis ha $\mathbf{A}$ teljes oszloprangú, akkor

$$\begin{aligned}\mathbf{A}^+\mathbf{A} &= \left((\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\right)\mathbf{A}\\ &= (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{A} = \mathbf{I}\end{aligned}$$

A másik állítás hasonlóan adódik a (7.16) képletből.

7.24. A feltételekből következik, hogy $\mathbf{A}^+\mathbf{A} = \mathbf{I}$, $\mathbf{C}\mathbf{C}^+ = \mathbf{I}$ Ellenőrizzük a 7.50. tétel feltételeit:

$$\begin{aligned}\mathbf{ACC}^+\mathbf{A}^+\mathbf{AC} &= \mathbf{AC},\\ \mathbf{C}^+\mathbf{A}^+\mathbf{ACC}^+\mathbf{A}^+ &= \mathbf{C}^+\mathbf{A}^+,\\ \mathbf{ACC}^+\mathbf{A}^+ &= \mathbf{AA}^+, \textit{amiszimmetrikus},\\ \mathbf{C}^+\mathbf{A}^+\mathbf{AC} &= \mathbf{C}^+\mathbf{C}, \textit{amiszimmetrikus}.\end{aligned}$$

A feltételek teljesülnek, tehát $\mathbf{AC}$ inverze valóban $\mathbf{C}^+\mathbf{A}^+$.

7.25. Ha $\mathrm{r}(\mathbf{A}) = 1$, akkor létezik olyan $\mathbf{a}$ és $\mathbf{b}$ vektor, hogy $\mathbf{A} = \mathbf{ab}'$. Ekkor a pszeudoinverz kiszámításáról szóló 7.18. feladat (7.21) képlete szerint

$$\begin{aligned}\mathbf{A}^+ &= \mathbf{b}(\mathbf{b}'\mathbf{b})^{-1}(\mathbf{a}'\mathbf{a})^{-1}\mathbf{a}'\\ &= \frac{1}{(\mathbf{a}'\mathbf{ab}'\mathbf{b})^{-1}}\mathbf{ba}'\\ &= \frac{1}{\mathrm{trace}(\mathbf{A}'\mathbf{A})}\mathbf{A}'.\end{aligned}$$

Az utóbbi egyenlőség azon múlik, hogy mind $\mathbf{a}'\mathbf{ab}'\mathbf{b}$, mind $\mathrm{trace}(\mathbf{A}'\mathbf{A})$ az összes $a_i^2 b_j^2$ alakú elem összege, ahol $\mathbf{a} = (a_1, \ldots, a_m)$, $\mathbf{b} = (b_1, \ldots, b_n)$. A vektorokra vonatkozó állítás ennek speciális esete.

7.26. Az állítás igazolásához elég csak a Penrose-tétel négy feltételét ellenőrizni.

7.27. A 7.26. feladat alapján az alábbi blokkosítással és 1-rangú mátrixok pszeudoinverzére vonatkozó állításból azonnal adódik a válasz:

$$\left[\begin{array}{ccc|c|cc} 1 & 1 & 1 & 0 & 0 & 0\\ \hline 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 0\\ \hline 0 & 0 & 0 & 0 & 1 & 1\end{array}\right]^+ = \left[\begin{array}{c|cc|c} 1/3 & 0 & 0 & 0\\ 1/3 & 0 & 0 & 0\\ 1/3 & 0 & 0 & 0\\ \hline 0 & 1/2 & 1/2 & 0\\ \hline 0 & 0 & 0 & 1/2\\ 0 & 0 & 0 & 1/2\end{array}\right]$$

7.28. ....

III. rész

Mátrixok sajátságai

E rész a mátrixok legfontosabb tulajdonságait, sajátságait vizsgálja. A *sajátságokra* való utalás egyúttal egy szójáték, mely a mátrix *sajátértékének*, *sajátvektorának* és *sajátalterének* igen fontos fogalmára, valamint a születésekor ugyancsak *sajátértéknek* nevezett szinguláris értékre utal. Vizsgálatainkat a négyzetes mátrixok lehető legegyszerűbb – diagonális – alakra hozásával kezdjük, amihez a sajátértékek meghatározása vezet. Az alkalmazásokban az utóbbi időben különösen fontossá vált, és minden mátrixra működő másik diagonalizáló technika a szinguláris értékekhez kapcsolódik. Ennek tárgyalását a diagonalizálhatósághoz kapcsolódó kérdések tisztázása, a „majdnem diagonális alak”, a Jordan-féle normálalak leírása követi, végül e részt az alkalmazásokban különösen fontos nemnegatív mátrixok vizsgálata zárja.

8

Sajátérték, diagonalizálás

Egy mátrix jellemzésének különösen hatékony eszköze azoknak a nullvektortól különböző $\mathbf{x}$ vektoroknak a meghatározása, amelyeket a mátrixszal való szorzás önmagukkal párhuzamos vektorba visz, azaz amelyekre $\mathbf{Ax} = \lambda\mathbf{x}$. E vektorok ismerete olyan bázis megtalálásához is hozzásegít, amelyben e mátrix lényegesen egyszerűbb – például diagonális – alakot ölt.

Sajátérték, sajátvektor, sajátaltér

A sajátérték és a sajátvektor fogalma Kezdjük egy egyszerű feladattal, melyből kiolvasható annak lényege, amiről e fejezetben szó lesz.

8.1. PÉLDA (JÓ BÁZIS TÜKRÖZÉSHEZ). *Tükrözzük a 3-dimenziós tér vektorait a tér egy megadott síkjára! Geometriai szemléletünkre hagyatkozva válasszunk e lineáris leképezés leírásához egy megfelelő bázist, majd írjuk fel a tükrözés e bázisra vonatkozó mátrixát!*

MEGOLDÁS. A síkra való tükrözés a síkra merőleges vektorokat ellentettjükbe viszi, míg a sík vektorait helyben hagyja. A tér minden vektora egyértelműen előáll egy síkba eső és egy rá merőleges vektor összegeként. A mellékelt ábra szemlélteti, hogy a tér mindegyik vektorának tükörképe úgy kapható meg, hogy a síkba eső összetevőjét meghagyjuk, és ahhoz adjuk a síkra merőleges összetevő ellentettjét. Természetes módon adódik az ötlet, hogy válasszuk ki a sík egy tetszőleges bázisát (álljon ez az $\mathbf{a}$ és $\mathbf{b}$ vektorokból), és e két vektorhoz vegyünk hozzá egy síkra merőleges $\mathbf{c}$ vektort harmadik bázisvektornak. Ekkor a tükröző T leképezés hatása e vektorokon: $T\mathbf{a} = \mathbf{a}$, $T\mathbf{b} = \mathbf{b}$ és $T\mathbf{c} = -\mathbf{c}$. Az $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ bázisban e három vektor koordinátás alakja a három standard egységvektor, így T mátrixa

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Így e bázisban egy tetszőleges (x, y, z) vektor tükörképe $(x, y, -z)$. □

E példában úgy választottunk bázist, hogy olyan vektorokat kerestünk, melyek önmaguk skalárszorosába mennek, azaz amelyek kielégítenek egy $T\mathbf{x} = \lambda\mathbf{x}$ alakú egyenletet. Ez a következő definícióhoz vezet, melyet először csak mátrixokra mondunk ki.

8.2. DEFINÍCIÓ (SAJÁTÉRTÉK, SAJÁTVEKTOR). *Azt mondjuk, hogy a λ szám az $\mathbf{A}$ mátrix* sajátértéke, *ha létezik olyan nemnulla $\mathbf{x}$ vektor, melyre $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$. Az ilyen $\mathbf{x}$ vektorokat az $\mathbf{A}$ mátrix λ sajátértékhez tartozó* sajátvektorainak, *a $(\lambda, \mathbf{x})$ párokat pedig az $\mathbf{A}$* sajátpárjainak *nevezzük.*

8.3. PÉLDA (SAJÁTÉRTÉK, SAJÁTVEKTOR). *Igazoljuk, hogy az $\mathbf{A} = \left[\begin{smallmatrix}-2 & 2\\ -2 & 3\end{smallmatrix}\right]$ mátrixnak -1 egy sajátértéke, és $(2,1)$ az egyik hozzátartozó sajátvektora, azaz $(-1,(2,1))$ egy sajátpár. Mutassuk meg, hogy a $(2,(1,2))$ pár egy másik sajátpár!*

MEGOLDÁS. Valóban,

$$\begin{bmatrix}-2 & 2\\ -2 & 3\end{bmatrix}\begin{bmatrix}2\\ 1\end{bmatrix} = \begin{bmatrix}-2\\ -1\end{bmatrix}, \quad \text{azaz} \quad \begin{bmatrix}-2 & 2\\ -2 & 3\end{bmatrix}\begin{bmatrix}2\\ 1\end{bmatrix} = (-1)\begin{bmatrix}2\\ 1\end{bmatrix}.$$

E mátrix egy másik sajátértéke 2, ugyanis

$$\begin{bmatrix}-2 & 2\\ -2 & 3\end{bmatrix}\begin{bmatrix}1\\ 2\end{bmatrix} = \begin{bmatrix}2\\ 4\end{bmatrix}, \quad \text{azaz} \quad \begin{bmatrix}-2 & 2\\ -2 & 3\end{bmatrix}\begin{bmatrix}1\\ 2\end{bmatrix} = (2)\begin{bmatrix}1\\ 2\end{bmatrix}.$$ □

Ha $\mathbf{x}$ egy sajátvektor, akkor minden nemnulla konstansszorosa is az, ugyanis

$$\mathbf{A}(c\mathbf{x}) = c\mathbf{A}\mathbf{x} = c\lambda\mathbf{x} = \lambda(c\mathbf{x}),$$

azaz $\mathbf{A}(c\mathbf{x}) = \lambda(c\mathbf{x})$. Ennél több is igaz:

8.4. ÁLLÍTÁS (A SAJÁTVEKTOROK ALTEREI). *Ha az $\mathbf{A}$ mátrixnak λ egy sajátértéke, akkor a λ-hoz tartozó sajátvektorok a nullvektorral együtt alteret alkotnak, mely megegyezik $\mathbf{A} - \lambda\mathbf{I}$ nullterével.*

BIZONYÍTÁS. A nem nullvektor $\mathbf{x}$ pontosan akkor egy λ sajátértékhez tartozó sajátvektor, ha kielégíti az $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ egyenletet, azaz az $\mathbf{A}\mathbf{x} - \lambda\mathbf{x} = \mathbf{0}$ egyenletet, vagyis ha megoldása a homogén lineáris $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ egyenletnek. Ez pedig épp azt jelenti, hogy $\mathbf{x}$ eleme $\mathbf{A} - \lambda\mathbf{I}$ nullterének. □

8.5. DEFINÍCIÓ (SAJÁTALTÉR). *A négyzetes $\mathbf{A}$ mátrix λ sajátértékhez tartozó sajátvektorai és a nullvektor által alkotott alteret a λ sajátértékhez tartozó* sajátaltérnek *nevezzük.*

8.6. PÉLDA (SAJÁTALTÉR BÁZISÁNAK MEGHATÁROZÁSA). *Adjuk meg az*

$$\mathbf{A} = \begin{bmatrix} 3 & 6 & 1 \\ 1 & 8 & 1 \\ 1 & 6 & 3 \end{bmatrix}$$

mátrix 2*-höz, mint sajátértékhez tartozó sajátalterét úgy, hogy megadjuk egy bázisát! Tegyük meg ugyanezt a* 10*-hez tartozó sajátaltérrel is.*

MEGOLDÁS. Először ellenőrizzük, hogy a 2 sajátérték! Ehhez meg kell mutatni, hogy az $(\mathbf{A} - 2\mathbf{I})\mathbf{x} = \mathbf{0}$ egyenletrendszernek van nemtriviális megoldása. Hozzuk az együtthatómátrixot redukált lépcsős alakra:

$$\mathbf{A} - 2\mathbf{I} = \begin{bmatrix} 1 & 6 & 1 \\ 1 & 6 & 1 \\ 1 & 6 & 1 \end{bmatrix} \quad \Longrightarrow \quad \begin{bmatrix} 1 & 6 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Mivel $\mathrm{r}(\mathbf{A} - 2\mathbf{I}) = 1$, ezért az $(\mathbf{A} - 2\mathbf{I})\mathbf{x} = \mathbf{0}$ egyenletrendszer szabad változóinak száma 2, és megoldása

$$\mathbf{x} = \begin{bmatrix} -6s - t \\ s \\ t \end{bmatrix} = s \begin{bmatrix} -6 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Tehát a sajátaltér egy bázisa a $(-6, 1, 0)$ és $(-1, 0, 1)$ vektorokból áll.

A 10 is sajátérték, mivel az $(\mathbf{A} - 10\mathbf{I})\mathbf{x} = \mathbf{0}$ egyenletrendszernek van nemtriviális megoldása, ugyanis

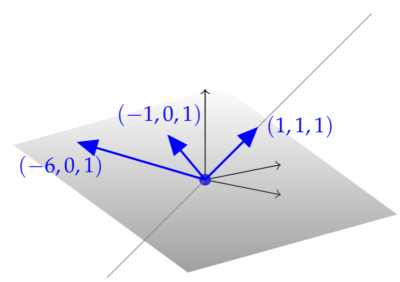
$$\mathbf{A} - 2\mathbf{I} = \begin{bmatrix} -7 & 6 & 1 \\ 1 & -2 & 1 \\ 1 & 6 & -7 \end{bmatrix} \quad \Longrightarrow \quad \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix},$$

tehát a megoldás

$$\mathbf{x} = \begin{bmatrix} t \\ t \\ t \end{bmatrix} = t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Így a sajátalteret az $(1, 1, 1)$ vektor feszíti ki.

A két sajátaltér egyike 2-, másika 1-dimenziós altér. Ezt szemlélteti a 8.1. ábra. □



8.1. ábra: A 8.6. feladatbeli **A** mátrix sajátalterei: a 2-höz tartozó 2-dimenziós altér, melyet a $(-6, 1, 0)$ és $(-1, 0, 1)$ vektorok feszítenek ki, és a 10-hez tartozó 1-dimenziós altér, melyet a $(1, 1, 1)$ vektor feszít ki.

Karakterisztikus polinom Láttuk, hogy az $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ egyenletnek pontosan akkor van a zérusvektortól különböző megoldása, ha a homogén lineáris $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ egyenletrendszernek van nemtriviális megoldása. Ez a 6.12. tétel szerint pontosan akkor igaz, ha

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0. \tag{8.1}$$

Ez tehát azt jelenti, hogy λ pontosan akkor sajátérték, ha kielégíti a (8.1) egyenletet. Ezt az egyenletet az **A** mátrix *karakterisztikus egyenletének* nevezzük. Ha **A** egy $n \times n$-es mátrix, akkor az egyenlet bal oldala a determináns kifejtése után egy n-edfokú polinom, melyet *karakterisztikus polinomnak* nevezünk.

A karakterisztikus polinomot a

$$\det(\lambda\mathbf{I} - \mathbf{A})$$

determinánssal is szokás definiálni. Előnye, hogy ekkor a polinom főegyütthatója mindig 1, míg az általunk használt definíció szerint a páratlan rendű mátrixok karakterisztikus polinomjának -1 a főegyütthatója. Hátránya viszont az, hogy a konstans tag nem mindig a determináns, másrészt a kézzel való számolás is nehézkesebb, ezért az elemi feladatok egyszerűbb számolhatósága érdekében is hasznosabb a $\det(\mathbf{A} - \lambda\mathbf{I})$ alakot választani.

8.7. PÉLDA (KARAKTERISZTIKUS POLINOM FELÍRÁSA). *Határozzuk meg az*

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} a & b & c \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

mátrixok karakterisztikus polinomját és ahol lehet, próbáljunk meg általánosabb érvényű állításokat megsejteni az eredmény alapján!

MEGOLDÁS. Ki kell számítanunk a $\det(\mathbf{A} - \lambda\mathbf{I})$ determináns értékét:

$$\begin{aligned}
\det(\mathbf{A} - \lambda\mathbf{I}) &= \begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = (a-\lambda)(d-\lambda) - bc \\
&= \lambda^2 - (a+d)\lambda + (ad - bc) \\
&= \lambda^2 - \operatorname{trace}(\mathbf{A})\lambda + \det \mathbf{A}. \qquad (8.2)
\end{aligned}$$

Kimondható a következtetés: 2×2-es mátrixok karakterisztikus polinomját a mátrix nyomával és determinánsával is ki tudjuk fejezni.

A **B** mátrix karakterisztikus polinomja

$$\det(\mathbf{B} - \lambda\mathbf{I}) = \begin{vmatrix} 1-\lambda & a & b \\ 0 & 1-\lambda & c \\ 0 & 0 & 1-\lambda \end{vmatrix} = (1-\lambda)^3.$$

Ebből leolvasható, hogy a háromszögmátrixok karakterisztikus polinomjának alakját nem befolyásolják a főátlón kívüli elemek (ld. a 8.8. tételt).

A **C** mátrix karakterisztikus polinomja

$$\begin{aligned}
\det(\mathbf{C} - \lambda\mathbf{I}) &= \begin{vmatrix} a-\lambda & b & c \\ 1 & -\lambda & 0 \\ 0 & 1 & -\lambda \end{vmatrix} \\
&= (a-\lambda)\lambda^2 + c + b\lambda \\
&= -\lambda^3 + a\lambda^2 + b\lambda + c.
\end{aligned}$$

Ez azt sejteti, hogy minden polinomhoz könnyen konstruálható olyan mátrix, melynek az a karakterisztikus polinomja (ld. a **??** feladatot). □

Az előző feladat tanulságait külön állításokban is megfogalmazzuk:

8.8. ÁLLÍTÁS (HÁROMSZÖGMÁTRIXOK SAJÁTÉRTÉKEI). *A háromszögmátrixok és így a diagonális mátrixok sajátértékei megegyeznek a főátló elemeivel.*

BIZONYÍTÁS. Ha $\mathbf{A}$ háromszögmátrix, akkor $\mathbf{A} - \lambda\mathbf{I}$ is, és egy háromszögmátrix determinánsa megegyezik főátlóbeli elemeinek szorzatával. Eszerint az $\mathbf{A} = [a_{ij}]$ mátrix karakterisztikus egyenlete

$$(a_{11} - \lambda)(a_{22} - \lambda)\dots(a_{nn} - \lambda) = 0$$

aminek a gyökei a_{ii} $(i = 1, \dots, n)$. Így ezek az $\mathbf{A}$ sajátértékei. □

8.9. ÁLLÍTÁS (DETERMINÁNS, NYOM ÉS A SAJÁTÉRTÉKEK). *Ha az n-edrendű $\mathbf{A}$ mátrix sajátértékei $\lambda_1, \dots, \lambda_n$, akkor*

$$\det(\mathbf{A}) = \lambda_1\lambda_2\dots\lambda_n$$
$$\operatorname{trace}(\mathbf{A}) = \lambda_1 + \lambda_2 + \dots + \lambda_n$$

Ezek az értékek megjelennek a karakterisztikus polinomban: a determináns a konstans tag, a nyom a $(-\lambda)^{n-1}$ együtthatója.

BIZONYÍTÁS. A karakterisztikus polinom gyöktényezős alakja:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\dots(\lambda_n - \lambda)$$

$\lambda = 0$ behelyettesítése után kapjuk, hogy

$$\det(\mathbf{A}) = \lambda_1\lambda_2\dots\lambda_n.$$

Az állítás nyomra vonatkozó részének bizonyítását feladatként tűzzük ki. □

A valós 2×2-es mátrixok sajátaltereinek jellemzése Olyan eredményekkel fogunk megismerkedni e paragrafusban, melyek általánosíthatók lesznek magasabb dimenzióra, de 2-dimenzió esetén egyszerűbb a szemléltetésük.

Láttuk, hogy ha $\mathbf{x}$ sajátvektor, akkor bármely konstansszorosa is az. Így egy egyenessel párhuzamos vektorok közül elég csak egy vektor képét vizsgálni, mondjuk az egységvektorét. Hasznos lesz tehát a lineáris leképezések korábban megismert egységkör-ábrázolása (ld. 7.4 ábra).

8.10. PÉLDA (2×2-ES MÁTRIXOK SAJÁTVEKTORAINAK ÁBRÁZOLÁSA). *Határozzuk meg a 7.5. és a 7.6. példákban is szereplő*

$$\mathbf{A} = \begin{bmatrix} \frac{5}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{5}{4} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \frac{3}{4} & \frac{5}{4} \\ \frac{5}{4} & \frac{3}{4} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} -\frac{5}{4} & \frac{3}{4} \\ -\frac{3}{4} & \frac{5}{4} \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} -\frac{3}{4} & \frac{5}{4} \\ -\frac{5}{4} & \frac{3}{4} \end{bmatrix}.$$

mátrixok sajátértékeit és sajátvektorait. Szemléltessük ezeket az egységkör-ábrákban.

MEGOLDÁS. Egyszerű számolással meghatározható mind a négy mátrix karakterisztikus egyenlete, sajátértékei és sajátvektorai, bár a $\mathbf{D}$ mátrix esetén ezek komplex számokat is tartalmaznak. A karakterisztikus polinomot jelölje p, indexében a mátrix jelével. Ezután megadjuk a sajátértékeket, majd a sajátvektorokat:

$$p_{\mathbf{A}}(\lambda) = \lambda^2 - \frac{5}{2}\lambda + 1, \quad \lambda_1 = 2,\ \lambda_2 = \frac{1}{2}, \quad \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix},\ \mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

$$p_{\mathbf{B}}(\lambda) = \lambda^2 - \frac{3}{2}\lambda - 1, \quad \lambda_1 = 2,\ \lambda_2 = -\frac{1}{2}, \quad \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix},\ \mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

$$p_{\mathbf{C}}(\lambda) = \lambda^2 - 1, \quad \lambda_1 = 1,\ \lambda_2 = -1, \quad \mathbf{x}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix},\ \mathbf{x}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

$$p_{\mathbf{D}}(\lambda) = \lambda^2 + 1, \quad \lambda_1 = \mathrm{i},\ \lambda_2 = -\mathrm{i}, \quad \mathbf{x}_1 = \begin{bmatrix} \frac{3}{5} - \frac{4}{5}\mathrm{i} \\ 1 \end{bmatrix},\ \mathbf{x}_2 = \begin{bmatrix} \frac{3}{5} + \frac{3}{4}\mathrm{i} \\ 1 \end{bmatrix}.$$

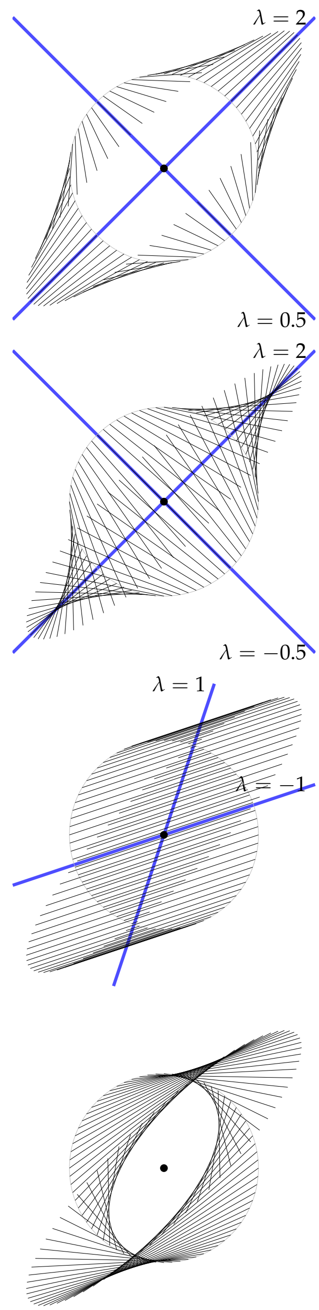
Némi nehézséget a $\mathbf{D}$ mátrix okoz, ezért az ahhoz tartozó számításokat részletezzük:

$$|\mathbf{D} - \lambda\mathbf{I}| = \begin{vmatrix} -\frac{3}{4} - \lambda & \frac{5}{4} \\ -\frac{5}{4} & -\frac{3}{4} - \lambda \end{vmatrix} = \lambda^2 + 1$$

$$\lambda_1 = \mathrm{i}: \quad \begin{bmatrix} -\frac{3}{4} - \mathrm{i} & \frac{5}{4} \\ -\frac{5}{4} & -\frac{3}{4} - \mathrm{i} \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & -\frac{3}{4} + \frac{4}{5}\mathrm{i} \\ 0 & 0 \end{bmatrix} \text{ amiből } \mathbf{x}_1 = \begin{bmatrix} \frac{3}{5} - \frac{4}{5}\mathrm{i} \\ 1 \end{bmatrix},$$

$$\lambda_2 = -\mathrm{i}: \quad \begin{bmatrix} -\frac{3}{4} + \mathrm{i} & \frac{5}{4} \\ -\frac{5}{4} & -\frac{3}{4} + \mathrm{i} \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & -\frac{3}{4} - \frac{4}{5}\mathrm{i} \\ 0 & 0 \end{bmatrix} \text{ amiből } \mathbf{x}_2 = \begin{bmatrix} \frac{3}{5} + \frac{4}{5}\mathrm{i} \\ 1 \end{bmatrix}.$$

A négy mátrixhoz tartozó egységkörábra a 8.2 ábrán látható. □



8.2. ábra: A négy leképezés sajátirányai.

8.11. TÉTEL (2 × 2-ES SZIMMETRIKUS MÁTRIXOK SAJÁTALTEREI). *Legyen* $\mathbf{A}$ *egy* 2×2*-es szimmetrikus mátrix. Ekkor*

a) $\mathbf{A}$ *minden sajátértéke valós,*

b) $\mathbf{A}$*-nak pontosan akkor van két azonos sajátértéke, ha* $a\mathbf{I}$ *alakú, ekkor a sík összes vektora sajátvektor,*

c) minden egyéb esetben $\mathbf{A}$*-nak két különböző sajátértéke van, és ekkor sajátalterei merőlegesek egymásra.*

BIZONYÍTÁS. A 2×2-es szimmetrikus valós mátrix általános alakja $\mathbf{A} = \left[\begin{smallmatrix} a & b \\ b & d \end{smallmatrix}\right]$, ahol $a, b, d \in \mathbb{R}$. Ennek karakterisztikus egyenlete a (8.2) szerint $\lambda^2 - (a+d)\lambda + (ad - b^2)$. Az egyenlet diszkriminánsa $D = (a+d)^2 - 4(ad - b^2) = (a-d)^2 + 4b^2 \geq 0$. Tehát a gyökök, vagyis a sajátértékek valósak. Ez bizonyítja *a)*-t. A két sajátérték csak akkor egyezik meg, ha $D = 0$, ez viszont csak $a = d$ és $b = 0$ esetén lehetséges, ami bizonyítja *b)*-t. A *c)* állítás igazolását a feladatok közt tűzzük ki. □

Mátrix összes sajátértékének és sajátvektorának meghatározása Az előző paragrafusokban leírtak alapján egy mátrix sajátértékeinek és sajátvektorainak meghatározása két lépésben elvégezhető:

1. megoldjuk a $\det(\mathbf{A}-\lambda\mathbf{I})=0$ karakterisztikus egyenletet, ennek gyökei a sajátértékek,
2. minden λ sajátértékhez meghatározzuk az $\mathbf{A}-\lambda\mathbf{I}$ nullterének egy bázisát, az általa kifeszített altér nemzérus vektorai a λ-hoz tartozó sajátvektorok.

8.12. példa (Az összes sajátérték és sajátvektor meghatározása). *Határozzuk meg a*

$$\begin{bmatrix}0&1&1\\0&2&0\\0&0&2\end{bmatrix}$$

mátrix sajátértékeit és sajátvektorait!

Megoldás. Az első lépés a karakterisztikus egyenletet felírása és megoldása. A kiszámítandó determináns háromszögalakú, így értéke a főátlóbeli elemek szorzata:

$$\det(\mathbf{A}-\lambda\mathbf{I})=\begin{vmatrix}0-\lambda&1&1\\0&2-\lambda&0\\0&0&2-\lambda\end{vmatrix}=-\lambda(2-\lambda)^2$$

A karakterisztikus egyenlet gyökei, és így az $\mathbf{A}$ mátrix sajátértékei $\lambda_1=0$, $\lambda_2=\lambda_3=2$.

Tekintsük először a $\lambda_1=0$ esetet. $\mathbf{A}-\lambda_1\mathbf{I}$ nullterének meghatározásához redukált lépcsős alakra hozzuk az $\mathbf{A}-\lambda_1\mathbf{I}$ mátrixot:

$$\begin{bmatrix}0&1&1\\0&2&0\\0&0&2\end{bmatrix}\Longrightarrow\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}\Longrightarrow\begin{aligned}x_2&=0\\x_3&=0\end{aligned}$$

Ennek megoldása $x_1=t$, azaz az összes megoldás

$$\begin{bmatrix}t\\0\\0\end{bmatrix}=t\begin{bmatrix}1\\0\\0\end{bmatrix}$$

Tehát a $\lambda_1=0$ sajátértékhez tartozó sajátaltér az $(1,0,0)$ vektor által kifeszített altér.

Tekintsük ezután a $\lambda_2=\lambda_3=2$ esetet. Meghatározzuk az $\mathbf{A}-2\mathbf{I}$ mátrix nullterét.

$$\begin{bmatrix}-2&1&1\\0&0&0\\0&0&0\end{bmatrix}\Longrightarrow\begin{bmatrix}2&-1&-1\\0&0&0\\0&0&0\end{bmatrix}\Longrightarrow 2x_1-x_2-x_3=0$$

Ennek az (egy egyenletből álló) egyenletrendszernek a megoldása: $x_2 = s$, $x_3 = t$, $x_1 = (s+t)/2$, azaz

$$\begin{bmatrix}(s+t)/2\\ s\\ t\end{bmatrix} = s\begin{bmatrix}1/2\\ 1\\ 0\end{bmatrix} + t\begin{bmatrix}1/2\\ 0\\ 1\end{bmatrix}$$

Tehát a $\lambda_2 = \lambda_3 = 2$ sajátértékhez tartozó sajátaltér az $(\frac{1}{2}, 1, 0)$ és az $(\frac{1}{2}, 0, 1)$ vektorok által kifeszített altér. □

Az $n \times n$-es mátrixok karakterisztikus egyenlete n-edfokú. Egy ilyen egyenlet megoldására $n \leq 4$ esetén van megoldóképlet, ezért ezeket az egyenleteket – például egy komputer algebra program segítségével – meg tudjuk oldani. Egyébként vagy szerencsénk van, és az egyenlet olyan alakú, amilyenhez vannak gyors megoldási lehetőségek, vagy csak közelítő megoldás megtalálására van esély.

8.13. PÉLDA (MAGASABBFOKÚ KARAKTERISZTIKUS EGYENLET). *Határozzuk meg az*

$$\mathbf{A} = \begin{bmatrix}1 & 2 & 2\\ 2 & 1 & 2\\ 3 & 3 & 2\end{bmatrix}$$

mátrix sajátértékeit és sajátvektorait!

MEGOLDÁS. A karakterisztikus egyenlet:

$$\begin{aligned}\mathbf{A} - \lambda\mathbf{I} &= \begin{bmatrix}1-\lambda & 2 & 2\\ 2 & 1-\lambda & 2\\ 3 & 3 & 2-\lambda\end{bmatrix}\\ &= (1-\lambda)^2(2-\lambda) + 24 - 12(1-\lambda) - 4(2-\lambda)\\ &= -(\lambda^3 - 4\lambda^2 - 11\lambda - 6)\end{aligned}$$

E harmadfokú egyenlet megoldására használhatunk számítógépet, vagy például a függelékben megtalálható Rolle-féle gyöktételt. Eszerint a karakterisztikus egyenlet $-(\lambda+1)^2(\lambda-6) = 0$, így gyökei $\lambda_1 = \lambda_2 = -1$ és $\lambda_3 = 6$.

A $\lambda_1 = \lambda_2 = -1$ esetben

$$\mathbf{A} + \mathbf{I} = \begin{bmatrix}2 & 2 & 2\\ 2 & 2 & 2\\ 3 & 3 & 3\end{bmatrix} \implies \begin{bmatrix}1 & 1 & 1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix} \implies x_1 + x_2 + x_3 = 0.$$

Ennek megoldása

$$\begin{bmatrix}-s-t\\ s\\ t\end{bmatrix} = s\begin{bmatrix}-1\\ 1\\ 0\end{bmatrix} + t\begin{bmatrix}-1\\ 0\\ 1\end{bmatrix},$$

azaz a -1 sajátértékhez tartozó sajátalteret a $(-1,1,0)$ és a $(-1,0,1)$ vektorok feszítik ki.

A $\lambda_3 = 6$ esetben

$$\mathbf{A} - 6\mathbf{I} = \begin{bmatrix} -5 & 2 & 2 \\ 2 & -5 & 2 \\ 3 & 3 & -4 \end{bmatrix} \implies \begin{bmatrix} 1 & 0 & -2/3 \\ 0 & 1 & -2/3 \\ 0 & 0 & 0 \end{bmatrix} \implies \begin{aligned} x_1 \qquad - \frac{2}{3}x_3 &= 0 \\ x_2 - \frac{2}{3}x_3 &= 0. \end{aligned}$$

Ennek megoldása a törtek alkalmazását elkerülő $x_3 = 3t$ paraméterválasztással

$$\begin{bmatrix} 2t \\ 2t \\ 3t \end{bmatrix} = t \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}.$$

Tehát a $\lambda_3 = 6$ sajátértékhez tartozó sajátalteret a $(2,2,3)$ vektor feszíti ki. □

A karakterisztikus egyenlet komplex gyökei Ha valóselemű mátrixot vizsgálunk, megeshet, hogy a karakterisztikus egyenletnek vannak komplex gyökei. Mivel a valós számok egyúttal komplexek is, a valós elemű mátrixot tekinthetjük komplex eleműnek is, ekkor viszont a karakterisztikus egyenlet komplex gyökeit is sajátértéknek tekinthetjük. Ebben az esetben a komplex sajátértékhez komplex elemű sajátvektor fog tartozni.

8.14. PÉLDA (KOMPLEX SAJÁTÉRTÉKEK ÉS KOMPLEX ELEMŰ SAJÁTVEKTOROK). *Határozzuk meg a komplex elemű*

$$\mathbf{A} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

mátrix sajátértékeit és sajátvektorait!

MEGOLDÁS. A karakterisztikus egyenlet

$$\begin{vmatrix} \frac{1}{2} - \lambda & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} - \lambda \end{vmatrix} = \left(\frac{1}{2} - \lambda\right)^2 + \left(\frac{\sqrt{3}}{2}\right)^2 = \lambda^2 - \lambda + 1.$$

A $\lambda^2 - \lambda + 1 = 0$ egyenlet gyökei $\frac{1}{2} \pm \frac{\sqrt{3}}{2}i$.

Először vizsgáljuk a $\frac{1}{2} + \frac{\sqrt{3}}{2}i$ sajátértéket:

$$\mathbf{A} - \left(\frac{1}{2} + \frac{\sqrt{3}}{2}i\right)\mathbf{I} = \begin{bmatrix} -\frac{\sqrt{3}}{2}i & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2}i \end{bmatrix} \implies \begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix} \implies x - iy = 0.$$

Ennek az egyenlet(rendszer)nek a megoldása az $y = t$ paraméterválasztással

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} it \\ t \end{bmatrix} = t \begin{bmatrix} i \\ 1 \end{bmatrix}$$

Tehát a $\frac{1}{2} + \frac{\sqrt{3}}{2}i$ sajátértékhez tartozó sajátaltér egy bázisa az $(i, 1)$ vektorból áll.

A $\frac{1}{2} - \frac{\sqrt{3}}{2}i$ sajátérték esetén

$$\mathbf{A} - \left(\frac{1}{2} - \frac{\sqrt{3}}{2}i\right)\mathbf{I} = \begin{bmatrix} \frac{\sqrt{3}}{2}i & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2}i \end{bmatrix} \implies \begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix} \implies x + iy = 0.$$

Ennek az egyenlet(rendszer)nek a megoldása az $y = t$ paraméterválasztással

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -it \\ t \end{bmatrix} = t\begin{bmatrix} -i \\ 1 \end{bmatrix}$$

Tehát a $\frac{1}{2} - \frac{\sqrt{3}}{2}i$ sajátértékhez tartozó sajátalteret a $(-i, 1)$ sajátvektor feszíti ki. □

A karakterisztikus egyenlet többszörös gyökei: az algebrai és a geometriai multiplicitás Ha λ a karakterisztikus egyenlet k-szoros gyöke, vagy más szóval λ multiplicitása vagy *algebrai multiplicitása* k, akkor a λ-hoz tartozó sajátaltér d dimenziójára $1 \le d \le k$. Ezt az állítást később bebizonyítjuk. A sajátaltér dimenzióját szokták a λ *geometriai multiplicitásának* nevezni. A 8.12. és a 8.13. példák olyan eseteket mutattak, amikor a sajátértékek algebrai és geometriai multiplicitása azonos, azaz minden sajátaltér épp annyi dimenziós, amennyi a gyök (algebrai) multiplicitása. A következő feladat azt mutatja, hogy a sajátaltér dimenziója kisebb is lehet.

8.15. PÉLDA (SAJÁTÉRTÉK ALGEBRAI ÉS GEOMETRIAI MULTIPLICITÁSA). *Határozzuk meg az*

$$\mathbf{A} = \begin{bmatrix} 4 & 1 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{bmatrix} \text{ és a } \mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

mátrix sajátértékeit és azok algebrai és geometriai multiplicitását?

MEGOLDÁS. Mivel **A** háromszögmátrix, ezért karakterisztikus polinomja $(4 - \lambda)^3$, a 4 tehát háromszoros gyök, azaz algebrai multiplicitása 3. Mivel

$$\mathbf{A} - 4\mathbf{I} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

ezért az $(\mathbf{A} - 4\mathbf{I})\mathbf{x} = \mathbf{0}$ egyenletrendszer az $y = 0$, $z = 0$ alakot ölti, aminek megoldása

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} t \\ 0 \\ 0 \end{bmatrix} = t\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Eszerint **A** sajátaltere 1-dimenziós, melyet az $(1, 0, 0)$ vektor feszít ki. A $\lambda = 4$ sajátérték geometriai multiplicitása tehát 1.

A **B** mátrix karakterisztikus polinomja $(1-\lambda)^2(2-\lambda)^2$, ennek gyökei 1 és 2, és mindegyiknek kettő az algebrai multiplicitása. Meghatározzuk sajátaltereiket. $\lambda = 1$ esetén

$$\mathbf{B} - \lambda\mathbf{I} = \mathbf{B} - \mathbf{I} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Az ehhez tartozó homogén egyenletrendszer megoldása:

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} s \\ t \\ 0 \\ 0 \end{bmatrix} = s \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

Tehát az altér dimenziója 2, vagyis a geometriai multiplicitás megegyezik az algebraival. Ha $\lambda = 2$, akkor

$$\mathbf{B} - \lambda\mathbf{I} = \mathbf{B} - 2\mathbf{I} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

ahonnan az ehhez tartozó homogén egyenletrendszer megoldása:

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ t \\ 0 \end{bmatrix} = t \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

Tehát az altér dimenziója most 1, vagyis a geometriai multiplicitás kisebb, mint az algebrai. □

Speciális mátrixok sajátértékei A mátrixok egyes különleges tulajdonságai a sajátértékek bizonyos tulajdonságát is befolyásolják.

8.16. tétel (Speciális mátrixok sajátértéke). *Legyen* **A** *egy* n*-edrendű valós mátrix. Ekkor*

a) ha **A** *szimmetrikus, akkor minden sajátértéke valós,*

b) ha **A** *ferdén szimmetrikus, akkor minden sajátértéke imaginárius,*

c) ha **A** *ortogonális, akkor minden sajátértékének* 1 *az abszolút értéke,*

d) **A** *pontosan akkor nilpotens, ha minden sajátértéke* 0*, azaz karakterisztikus polinomja* λ^n *alakú.*

BIZONYÍTÁS. *a), b)* Írjuk fel a sajátpár elemeit komplex algebrai alakban, azaz legyen $\lambda = a + bi$ a valós $\mathbf{A}$ mátrix sajátértéke, és legyen $\mathbf{x} = \mathbf{y} + \mathbf{z}\mathrm{i}$ a hozzá tartozó sajátvektor. Tehát $\mathbf{A}\mathbf{x} = \mathbf{A}(\mathbf{y} + \mathbf{z}\mathrm{i}) = (a + b\mathrm{i})(\mathbf{y} + \mathbf{z}\mathrm{i})$. Ha ennek mindkét oldalát balról beszorozzuk az $\mathbf{x}$ konjugáltjának transzponáltjával, azaz az $\overline{\mathbf{x}}^T = (\mathbf{y} - \mathbf{z}\mathrm{i})^T$ vektorral, akkor kapjuk, hogy

$$\overline{\mathbf{x}}^T\mathbf{A}\mathbf{x} = (\mathbf{y} - \mathbf{z}\mathrm{i})^T(a + b\mathrm{i})(\mathbf{y} + \mathbf{z}\mathrm{i}) = (a + b\mathrm{i})(|\mathbf{y}|^2 + |\mathbf{z}|^2).$$

Itt kihasználtuk, hogy $\mathbf{y}^T\mathbf{y} = |\mathbf{y}|^2$ és $\mathbf{z}^T\mathbf{z} = |\mathbf{z}|^2$. Véve mindkét oldal konjugáltját és kihasználva hogy $\mathbf{A}$ valós, tehát $\overline{\mathbf{A}} = \mathbf{A}$, kapjuk, hogy

$$\mathbf{x}^T\mathbf{A}\overline{\mathbf{x}} = (a - b\mathrm{i})(|\mathbf{y}|^2 + |\mathbf{z}|^2).$$

Ezután vegyük mindkét oldal transzponáltját. A jobb oldal nem változik, hisz ott egy szám, azaz egy 1×1-es mátrix áll:

$$\overline{\mathbf{x}}^T\mathbf{A}^T\mathbf{x} = (a - b\mathrm{i})(|\mathbf{y}|^2 + |\mathbf{z}|^2).$$

Ezt összevetve a (115) egyenlőség két oldalán álló kifejezéssel, és kihasználva, hogy $\mathbf{x} \neq \mathbf{0}$ miatt $|\mathbf{y}|^2 + |\mathbf{z}|^2 \neq 0$, kapjuk, hogy

$$a + b\mathrm{i} = \begin{cases} a - b\mathrm{i}, & \text{ha } \mathbf{A} \text{ szimmetrikus,} \\ -a + b\mathrm{i}, & \text{ha } \mathbf{A} \text{ ferdén szimmetrikus.} \end{cases} \qquad \square$$

Eszerint $b = 0$, ha $\mathbf{A}$ szimmetrikus, és $a = 0$, ha $\mathbf{A}$ ferdén szimmetrikus, ami bizonyítja az első két állítást.

c) Legyen $\mathbf{A}$ ortogonális, ekkor $\mathbf{A}^T\mathbf{A} = \mathbf{I}$.

Sajátértékek és a mátrix hatványai A mátrixok függvényeinek számolása szoros kapcsolatban van a sajátértékekkel. E témában első lépés a mátrixokhatványok sajátértékeinek és sajátvektoranak meghatározása.

8.17. TÉTEL (MÁTRIX INVERTÁLHATÓSÁGA ÉS A 0 SAJÁTÉRTÉK). *Az $\mathbf{A}$ mátrix pontosan akkor invertálható, ha a 0 nem sajátértéke.*

BIZONYÍTÁS. $\mathbf{A}$ pontosan akkor invertálható, ha $\det(\mathbf{A}) \neq 0$, de ez ekvivalens azzal, hogy $\det(\mathbf{A} - 0\mathbf{I}) \neq 0$, azaz 0 nem sajátértéke $\mathbf{A}$-nak. $\square$

8.18. TÉTEL (MÁTRIX HATVÁNYAINAK SAJÁTÉRTÉKEI ÉS SAJÁTVEKTORAI). *Ha λ az $\mathbf{A}$ mátrix egy sajátértéke és $\mathbf{x}$ egy hozzá tartozó sajátvektor, akkor bármely egész n esetén λ^n sajátértéke az $\mathbf{A}^n$ mátrixnak és $\mathbf{x}$ egy hozzá tartozó sajátvektor, amennyiben λ^n és $\mathbf{A}^n$ is értelmezve van.*

BIZONYÍTÁS. $n = 0$ esetén $\lambda^0 = 1$ és $\mathbf{A}^0 = \mathbf{I}$, és ekkor minden vektor sajátvektor, tehát ekkor az állítás igaz.

Pozitív n-re indukcióval igazoljuk az állítást: $n = 1$ esetén nyilván igaz, $n = 2$ esetén:

$$\mathbf{A}^2\mathbf{x} = \mathbf{A}(\mathbf{A}\mathbf{x}) = \mathbf{A}(\lambda\mathbf{x}) = \lambda(\mathbf{A}\mathbf{x}) = \lambda(\lambda\mathbf{x}) = \lambda^2\mathbf{x}.$$

Hasonlóan kapjuk, hogy ha $n = k - 1$ esetén már igaz az állítás, akkor $n = k$ esetén is:

$$\mathbf{A}^k\mathbf{x} = \mathbf{A}(\mathbf{A}^{k-1}\mathbf{x}) = \mathbf{A}(\lambda^{k-1}\mathbf{x}) = \lambda^{k-1}(\mathbf{A}\mathbf{x}) = \lambda^{k-1}(\lambda\mathbf{x}) = \lambda^k\mathbf{x}.$$

Ha $\mathbf{A}$ invertálható, akkor

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}, \text{ amiből } \frac{1}{\lambda}\mathbf{x} = \mathbf{A}^{-1}\mathbf{x}, \text{ azaz } \lambda^{-1}\mathbf{x} = \mathbf{A}^{-1}\mathbf{x}.$$

Végül negatív kitevők esetén:

$$\mathbf{A}^k\mathbf{x} = \lambda^k\mathbf{x}, \text{ amiből } \lambda^{-k}\mathbf{x} = \mathbf{A}^{-k}\mathbf{x}.$$ □

8.19. TÉTEL (MÁTRIX HATVÁNYAINAK HATÁSA). *Tegyük fel, hogy λ_1, $\lambda_2, \ldots\ \lambda_k$ sajátértékei az $n \times n$-es $\mathbf{A}$ mátrixnak, és hogy $\mathbf{x}_1, \ldots\ \mathbf{x}_k$ a hozzájuk tartozó sajátvektorok. Ha egy n-dimenziós $\mathbf{v}$ vektor előáll e sajátvektorok lineáris kombinációjaként, azaz*

$$\mathbf{v} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \ldots + c_k\mathbf{x}_k,$$

akkor bármely egész m esetén

$$\mathbf{A}^m\mathbf{v} = c_1\lambda_1^m\mathbf{x}_1 + c_2\lambda_2^m\mathbf{x}_2 + \ldots + c_k\lambda_k^m\mathbf{x}_k.$$

BIZONYÍTÁS. A bizonyítás magától értetődő, hisz

$$\begin{aligned}\mathbf{A}^m\mathbf{v} &= \mathbf{A}^m(c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \ldots + c_k\mathbf{x}_k)\\ &= c_1\mathbf{A}^m\mathbf{x}_1 + c_2\mathbf{A}^m\mathbf{x}_2 + \ldots + c_k\mathbf{A}^m\mathbf{x}_k\\ &= c_1\lambda_1^m\mathbf{x}_1 + c_2\lambda_2^m\mathbf{x}_2 + \ldots + c_k\lambda_k^m\mathbf{x}_k\end{aligned}$$ □

► Sajnos az nem igaz, hogy minden mátrixhoz találunk n független sajátvektort, amelyek lineáris kombinációjaként minden vektor felírható, így e tétel csak a sajátvektorok lineáris kombinációjaként előálló vektorokról szól! E tétel azt mutatja, fontos kérdés annak eldöntése, hogy egy mátrix sajátvektoraiból mikor alkotható bázis.

Feladatok

8.1. Igazoljuk a **??** tétel *c)* állítását, mely szerint ha egy 2×2-es szimmetrikus valós mátrixnak két különböző sajátértéke van, akkor sajátalterei merőlegesek egymásra.

Hasonlóság, diagonalizálhatóság

Egy lineáris transzformáció sajátértékei és karakterisztikus polinomja megőrződnek a különféle bázisokban fölírt mátrixaira is. Olyan bázist keresünk, melyben mátrixa a legegyszerűbb alakú.

Lineáris transzformációk sajátértékei A sajátérték, sajátvektor, sajátaltér fogalma természetes módon átvihető lineáris leképezésekre is.

8.20. definíció (Lineáris transzformáció sajátértéke, sajátvektora). *Azt mondjuk, hogy a λ szám az L lineáris transzformáció* sajátértéke, *ha létezik olyan nemnulla $\mathbf{x}$ vektor, melyre $L\mathbf{x} = \lambda\mathbf{x}$. Az ilyen $\mathbf{x}$ vektorokat az L lineáris transzformáció λ sajátértékhez tartozó* sajátvektorainak *nevezzük.*

Ha a lineáris transzformáció $\mathbb{R}^2 \to \mathbb{R}^2$ vagy $\mathbb{R}^3 \to \mathbb{R}^3$ leképezés, mely valamilyen egyszerű geometriai transzformációt valósít meg, akkor néha a transzformáció mátrixának ismerete nélkül is könnyen meghatározhatjuk a sajátértékeket és sajátvektorokat.

8.21. példa (Lineáris transzformáció sajátértéke, sajátaltere). *Adjuk meg – pusztán geometriai szemléletünkre hagyatkozva – az alábbi lineáris leképezések sajátértékeit és a hozzájuk tartozó sajátaltereket.*

a) a sík vektorainak tükrözése egy egyenesre (vagy pontjainak tükrözése egy origón átmenő egyenesre);

b) a sík vektorainak merőleges vetítése egy egyenesre (vagy pontjainak merőleges vetítése egy origón átmenő egyenesre);

c) a tér vektorainak elforgatása egy egyenes körül a π egész számú többszörösétől különböző szöggel;

d) a tér vektorainak merőleges vetítése egy síkra;

e) a tér vektorainak tükrözése egy síkra.

Megoldás. Az előző fejezetben, így a **??** példában bizonyítottakhoz hasonlóan látható, hogy mindegyik feladatbeli transzformáció lineáris.

a) Egy egyenesre való tükrözés esetén csak az egyenessel párhuzamos és rá merőleges vektorok mennek saját konstansszorosukba, mégpedig az egyenessel párhuzamos vektorok saját magukba, a rá merőlegesek a saját ellentettjükbe. Tehát e transzformációnak az 1 sajátértékhez tartozó sajátaltere az egyenessel párhuzamos vektorokból, a -1-hez tartozó sajátaltere a rá merőleges vektorokból áll. A pontokra vonatkozó állítás a pontokba mutató helyvektorokkal adódik.

b) A sík merőleges vetítése egy egyenesre – hasonlóan az előző esethez – helyben hagyja az egyenessel párhuzamos vektorokat, és a $\mathbf{0}$-vektorba viszi a rá merőlegeseket. Tehát az 1 sajátértékhez tartozó sajátaltér az egyenessel párhuzamos vektorokból, a 0-hoz tartozó sajátaltere a rá merőleges vektorokból áll.

c) A tér π egész számú többszörösétől különböző szöggel való elforgatása egy egyenes körül a forgástengellyel párhuzamos vektorokat önmagukba viszi, és semelyik másikat sem viszi a saját skalárszorosába, így az egyetlen sajátérték az 1, amelyhez tartozó sajátaltér a forgástengellyes párhuzamos vektorokból áll.

d) A tér vektorainak merőleges vetítése egy síkra helyben hagyja a sík összes vektorát, míg a síkra merőleges vektorokat a $\mathbf{0}$ vektorba viszi, tehát a két sajátérték 1 és 0. Az 1-hez tartozó sajátaltér a sík vektoraiból, a 0-hoz tartozó sajátaltér a rá merőleges vektorokból áll.

e) E feladatot megoldottuk a 8.1. példában. A két sajátérték 1 és -1, az 1 sajátértékhez tartozó sajátaltér a sík vektoraiból, a -1-hoz tartozó sajátaltér a rá merőleges vektorokból áll. □

Egy lineáris leképezéshez bázisonként más-más mátrix tartozhat, de a sajátértékeik mégis ugyanazok, hisz ha egy leképezés egy vektort a λ-szorosába viszi, akkor azt λ-szorosába viszi minden bázisban.

Hasonló mátrixok sajátértékei A 7.16. és a 7.17. tételekből tudjuk, hogy egy lineáris leképezéshez különböző bázisokban tartozó mátrixok hasonlóak, másrészt tudjuk azt is, hogy fontos mátrixtulajdonságok invariánsak a hasonlóságra. E paragrafusban e tulajdonságok körét fogjuk bővíteni.

8.22. TÉTEL (SAJÁTÉRÉKHEZ KAPCSOLÓDÓ INVARIÁNSOK). *Ha $\mathbf{A} \sim \mathbf{B}$, akkor $\mathbf{A}$ és $\mathbf{B}$ karakterisztikus polinomja azonos, így sajátértékei, azok algebrai és geometriai multiplicitásai is megegyeznek.*

BIZONYÍTÁS. A bizonyítás során föltesszük, hogy valamely invertálható $\mathbf{C}$ mátrixszal $\mathbf{A} = \mathbf{C}^{-1}\mathbf{B}\mathbf{C}$. Ekkor

$$\begin{aligned}\mathbf{A} - \lambda\mathbf{I} &= \mathbf{C}^{-1}\mathbf{B}\mathbf{C} - \lambda\mathbf{C}^{-1}\mathbf{I}\mathbf{C})\\ &= \mathbf{C}^{-1}(\mathbf{B}\mathbf{C} - \lambda\mathbf{I}\mathbf{C})\\ &= \mathbf{C}^{-1}(\mathbf{B} - \lambda\mathbf{I})\mathbf{C},\end{aligned}$$

azaz $\mathbf{A} - \lambda\mathbf{I}$ és $\mathbf{B} - \lambda\mathbf{I}$ is hasonlóak. A 7.17. tétel szerint hasonló mátrixok determinánsa megegyezik, így $\det(\mathbf{A} - \lambda\mathbf{I}) = \det(\mathbf{B} - \lambda\mathbf{I})$, azaz megegyeznek $\mathbf{A}$ és $\mathbf{B}$ karakterisztikus polinomjai is. Ez maga után vonja, hogy megegyeznek sajátértékeik, és azok (algebrai) multiplicitásai. A geometriai multiplicitások egyenlőségéhez elég belátni, hogy $\mathbf{A} - \lambda\mathbf{I}$ és $\mathbf{B} - \lambda\mathbf{I}$ nullterének dimenziója megegyezik, azt viszont ugyancsak a 7.17. tételben igazoltuk. □

► E tétel egyik következménye, hogy hasonló mátrixok determinánsa megegyezik, hisz ezek a sajátértékek szorzatával egyenlők. Ezt már bizonyítottuk a 7.17. tételben. Hasonlóan igaz, hogy hasonló mátrixok nyoma is megegyezik, hisz ezek a sajátértékek összegével egyenlők.

► Fontos következménye tehát e tételnek, hogy van értelme *lineáris transzformáció karakterisztikus polinomjáról, sajátértékeiről, sajátaltereiről* beszélni. Legalábbis véges dimenziós esetben, pl. $\mathbb{R}^n \to \mathbb{R}^n$ vagy $\mathbb{C}^n \to \mathbb{C}^n$ transzformációk esetén ezt látjuk.

Mátrixok diagonalizálása Korábban láttuk, hogy egy mátrix hatványainak hatását milyen könnyű a sajátvektorok lineáris kombinációin kiszámítani. Ez akkor lenne igazán hatékony eszköz, ha sajátvektorokból egy bázist tudnánk választani. Ebben a bázisban ugyanis a mátrix – mint azt bizonyítani fogjuk – diagonális alakot ölt.

8.23. DEFINÍCIÓ (DIAGONALIZÁLHATÓSÁG). *Az $n \times n$-es $\mathbf{A}$ mátrix* diagonalizálható*, ha hasonló egy diagonális mátrixhoz, azaz ha létezik egy olyan diagonális $\mathbf{\Lambda}$ és egy invertálható $\mathbf{C}$ mátrix, hogy*

$$\mathbf{\Lambda} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}.$$

8.24. TÉTEL (DIAGONALIZÁLHATÓSÁG SZÜKSÉGES ÉS ELÉGSÉGES FELTÉTELE). *Az $n \times n$-es $\mathbf{A}$ mátrix pontosan akkor diagonalizálható, ha van n lineárisan független sajátvektora. Ekkor a diagonális mátrix az $\mathbf{A}$ sajátértékeiből, $\mathbf{C}$ a sajátvektoraiból áll.*

BIZONYÍTÁS. Ha $\mathbf{A}$ hasonló egy diagonális mátrixhoz, azaz van olyan $\mathbf{C}$ mátrix, hogy $\mathbf{\Lambda} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$ diagonális, akkor $\mathbf{C}$-vel balról szorozva a $\mathbf{C}\mathbf{\Lambda} = \mathbf{A}\mathbf{C}$ egyenlőséget kapjuk. Ha $\mathbf{C} = [\mathbf{c}_1\ \mathbf{c}_2\ \dots\ \mathbf{c}_n]$ és $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, akkor

$$[\mathbf{c}_1\ \mathbf{c}_2\ \dots\ \mathbf{c}_n] \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} = \mathbf{A}[\mathbf{c}_1\ \mathbf{c}_2\ \dots\ \mathbf{c}_n], \tag{8.3}$$

és itt a bal oldali mátrix i-edik oszlopa $\lambda_i\mathbf{c}_i$, a jobb oldali mátrixé $\mathbf{A}\mathbf{c}_i$, amik megegyeznek, azaz $\mathbf{A}\mathbf{c}_i = \lambda_i\mathbf{c}_i$, tehát $\mathbf{c}_i$ a λ_i sajátértékhez tartozó sajátvektor. Mivel $\mathbf{C}$ invertálható, ezért oszlopvektorai függetlenek, ami bizonyítja az állításunk egyik felét. Tegyük most fel, hogy van $\mathbf{A}$-nak n független sajátvektora. Képezzünk a sajátértékekből egy $\mathbf{\Lambda}$ diagonális mátrixot, a sajátvektorokból pedig egy $\mathbf{C}$ mátrixot úgy, hogy a $\mathbf{\Lambda}$ i-edik oszlopába kerülő λ_i sajátértékhez tartozó sajátvektor a $\mathbf{C}$ mátrix i-edik oszlopába kerüljön. Mivel $\lambda_i\mathbf{c}_i = \mathbf{A}\mathbf{c}_i$, ezért fönnáll a (8.3) összefüggés, azaz $\mathbf{\Lambda}$ hasonló $\mathbf{A}$-hoz. □

8.25. PÉLDA (MÁTRIX DIAGONALIZÁLÁSA). *Diagonalizálható-e a 8.12. példabeli*

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

mátrix?

MEGOLDÁS. Az $\mathbf{A}$ mátrix sajátértékeit és sajátvektorait meghatároztuk a 8.12. példában. Mivel $\lambda_1 = 0$, $\lambda_2 = \lambda_3 = 2$, a hozzájuk tartozó sajátvektorok $(1,0,0)$, $(1/2,1,0)$ és $(1/2,0,1)$, és ezek a vektorok lineárisan függetlenek, ezért $\mathbf{A}$ hasonló a $\mathbf{\Lambda}$ diagonális mátrixhoz, ahol

$$\mathbf{\Lambda} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad \text{és} \quad \mathbf{C} = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Ez könnyen igazolható a $\mathbf{C\Lambda} = \mathbf{AC}$ összefüggés ellenőrzésével:

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

□

A diagonalizálhatóság fontos voltára tekintettel érdemes további feltételeket gyűjteni, melyek könnyen ellenőrizhetők. Több elégséges feltétel származtatható a következő tételből:

8.26. TÉTEL (KÜLÖNBÖZŐ SAJÁTÉRTÉKEK SAJÁTVEKTORAI). *Ha λ_1, $\lambda_2, \dots \lambda_k$ különböző sajátértékei az $n \times n$-es $\mathbf{A}$ mátrixnak, akkor a hozzájuk tartozó $\mathbf{x}_1$, $\mathbf{x}_2, \dots \mathbf{x}_k$ sajátvektorok lineárisan függetlenek.*

BIZONYÍTÁS. Indirekt módon tegyük fel, hogy e vektorok lineárisan összefüggők. Ekkor van a vektorok közt olyan, amely csak a kisebb indexűek lineáris függvénye. Legyen ezek közül a legkisebb indexű $\mathbf{x}_i$, azaz

$$\mathbf{x}_i = c_1\mathbf{x}_1 + \dots + c_{i-1}\mathbf{x}_{i-1}, \tag{8.4}$$

de az i-nél kisebb indexű vektorok már lineárisan függetlenek. Szorozzuk meg az egyenlőség mindkét oldalát balról az $\mathbf{A}$ mátrixszal:

$$\mathbf{A}\mathbf{x}_i = \mathbf{A}(c_1\mathbf{x}_1 + \dots + c_{i-1}\mathbf{x}_{i-1}) = c_1\mathbf{A}\mathbf{x}_1 + \dots + c_{i-1}\mathbf{A}\mathbf{x}_{i-1},$$

majd használjuk ki, hogy e vektorok sajátvektorok:

$$\lambda_i\mathbf{x}_i = c_1\lambda_1\mathbf{x}_1 + \dots + c_{i-1}\lambda_{i-1}\mathbf{x}_{i-1}. \tag{8.5}$$

Ezután a (8.4) egyenlet mindkét oldalát λ_i-vel szorozva kapjuk, hogy

$$\lambda_i\mathbf{x}_i = c_1\lambda_i\mathbf{x}_1 + \dots + c_{i-1}\lambda_i\mathbf{x}_{i-1}. \tag{8.6}$$

Végül a (8.6) egyenletből a (8.5) egyenletet kivonva kapjuk, hogy

$$\mathbf{0} = c_1(\lambda_i - \lambda_1)\mathbf{x}_1 + \ldots + c_{i-1}(\lambda_i - \lambda_{i-1})\mathbf{x}_{i-1},$$

Mivel az $\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}$ vektorok már lineárisan függetlenek, és a $\lambda_1, \ldots, \lambda_i$ értékek különbözőek, így $c_1 = \cdots = c_{i-1} = 0$. Eszerint

$$\mathbf{x}_i = 0\mathbf{x}_1 + \cdots + 0\mathbf{x}_{i-1} = \mathbf{0},$$

ami ellentmondás, hisz $\mathbf{x}_i$ sajátvektor, tehát nem lehet a $\mathbf{0}$. Ez bizonyítja az indirekt feltevés helytelen voltát, azaz igazolja állításunkat. □

► Szokás úgy fogalmazni, hogy a különböző sajátértékekhez tartozó sajátalterek lineárisan függetlenek, hisz bárhogy választunk közülük egy-egy nemzérus vektort, azok lineárisan függetlenek lesznek.
► Másik fontos következménye e tételnek, hogy ha különböző sajátértékekhez tartozó sajátalterek mindegyikéből lineárisan független vektorokat választunk, akkor még ezek egyesítése is lineárisan független lesz. Ha ugyanis lineárisan összefüggnének, akkor az egy altérbe eső vektorok lineáris kombinációit összevonva egyetlen vektorrá, minden sajátértékhez egy-egy sajátvektort kapnánk, melyek összefüggők lennének, ami a fenti tételnek ellentmond.
► Speciálisan az is igaz, hogy különböző sajátértékekhez tartozó sajátalterek mindegyikéből egy bázist választva, azok egyesítése is lineárisan független vektorrendszert ad.

8.27. KÖVETKEZMÉNY (KÜLÖNBÖZŐ SAJÁTÉRTÉKEK ÉS DIAGONALIZÁLHATÓSÁG). *Ha az n-edrendű* $\mathbf{A}$ *mátrixnak n darab különböző sajátértéke van, akkor diagonalizálható.*

BIZONYÍTÁS. A 8.24. és a 8.26. tételek szerint n különböző sajátértékhez n független sajátvektor tartozik, ami épp azt jelenti, hogy a mátrix diagonalizálható. □

Végezetül összefoglalásul fölsorolunk néhány mátrixosztályt, melyekbe tartozó mátrixok mindegyike egyformán viselkedik a diagonalizálhatóságra nézve:

a) Egy valós n-edrendű mátrix nem diagonalizálható a valós mátrixok körében, ha karakterisztikus egyenletének vannak komplex gyökei, a komplex sajátértékekhez ugyanis nem található valós sajátvektor, így nem találhatunk n független sajátvektort, hisz a valós sajátértékek száma kisebb n-nél. Például a

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

mátrix a valósok fölött nem diagonalizálható, de a komplexek fölött igen (ld. a **??** feladatot).

b) Nem diagonalizálhatók a nilpotens mátrixok, például a

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

mátrix (**??** feladat).

c) Diagonalizálható minden szimmetrikus mátrix, sőt, sajátvektoraiból ortonormált bázis is kiválasztható. Ezt hamarosan belátjuk (8.34. tétel).

8.28. PÉLDA (DIAGONALIZÁLHATÓSÁG MEGÁLLAPÍTÁSA). *Döntsük el, hogy az alábbi mátrixok közül melyik diagonalizálható valós mátrixként!*

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} 1 & -4 \\ 4 & 1 \end{bmatrix}, \ \mathbf{C} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}, \ \mathbf{D} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

MEGOLDÁS. Az $\mathbf{A}$ mátrix nilpotens, mert $\mathbf{A}^2 = \mathbf{O}$, így nem diagonalizálható. A $\mathbf{B}$ mátrixnak vannak komplex sajátértékei, így a valósok fölött nem diagonalizálható, de a komplexek fölött igen, mert két különböző sajátértéke van. A $\mathbf{C}$ mátrixnak különbözőek a sajátértékei, és mind valósak (1, 4, 6), így diagonalizálható. A $\mathbf{D}$ mátrix csak a figyelmet ellenőrző kérdés, e mátrix diagonális, tehát diagonalizálva van, azaz biztosan diagonalizálható. □

*Sajátértékek multiplicitása és a diagonalizálhatóság** A sajátértékek algebrai és geometriai multiplicitása, valamint a diagonalizálhatóság közt egyszerű, de fontos összefüggés van. Nevezetesen a geometriai multiplicitás sosem nagyobb az algebrainál, másrészt a diagonalizálhatóság ekvivalens azzal, hogy a geometriai és algebrai multiplicitások minden sajátérték esetén megegyeznek.

8.29. TÉTEL (ALGEBRAI ÉS GEOMETRIAI MULTIPLICITÁS KAPCSOLATA). *Egy mátrix valamely sajátértékének geometriai multiplicitása nem lehet nagyobb az algebrai multiplicitásánál.*

BIZONYÍTÁS. Az $\mathbf{A}$ mátrix egy μ sajátértékének geometriai multiplicitását jelölje g. Ez azt jelenti, hogy $\mathbf{A} - \mu\mathbf{I}$ nullterének van g vektorból álló bázisa. Legyen ez $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_g$. Egészítsük ki e bázist az egész tér bázisává a $\mathbf{x}_{g+1}, \ldots, \mathbf{x}_n$ vektorokkal. E független vektorokból képzett $\mathbf{C} = [\mathbf{x}_1 \ \ldots \ \mathbf{x}_g | \mathbf{x}_{g+1} \ \ldots \ \mathbf{x}_n]$ mátrix invertálható. Írjuk $\mathbf{C}$-t blokkmátrix alakba: $\mathbf{X}$ legyen az első g oszlopból álló blokk, azaz $\mathbf{C} = [\mathbf{X}|\mathbf{Y}]$. Mivel $\mathbf{X}$ oszlopai a μ-höz tartozó sajátvektorok, ezért $\mathbf{AX} = \mu\mathbf{X}$. A $\mathbf{C}^{-1}$

inverzet első g sora után bontsuk blokkokra:

$$\mathbf{C}^{-1} = \left[\begin{array}{c}\mathbf{Z}\\ \hline \mathbf{W}\end{array}\right].$$

Írjuk fel az $\mathbf{I} = \mathbf{C}^{-1}\mathbf{C}$ összefüggést blokkmátrix alakban:

$$\begin{bmatrix}\mathbf{I}_g & \mathbf{O}\\ \mathbf{O} & \mathbf{I}_{n-g}\end{bmatrix} = \begin{bmatrix}\mathbf{Z}\\ \mathbf{W}\end{bmatrix}\begin{bmatrix}\mathbf{X} & \mathbf{Y}\end{bmatrix} = \begin{bmatrix}\mathbf{ZX} & \mathbf{ZY}\\ \mathbf{WX} & \mathbf{WY}\end{bmatrix}.$$

Innen leolvasható, hogy $\mathbf{WX} = \mathbf{O}$, $\mathbf{ZY} = \mathbf{O}$, $\mathbf{ZX} = \mathbf{I}_g$, $\mathbf{WY} = \mathbf{I}_{n-g}$. Ezeket fölhasználva kapjuk, hogy

$$\mathbf{C}^{-1}\mathbf{AC} = \begin{bmatrix}\mathbf{Z}\\ \mathbf{W}\end{bmatrix}\mathbf{A}\begin{bmatrix}\mathbf{X} & \mathbf{Y}\end{bmatrix} = \begin{bmatrix}\mathbf{ZAX} & \mathbf{ZAY}\\ \mathbf{WAX} & \mathbf{WAY}\end{bmatrix} = \begin{bmatrix}\mu\mathbf{I}_g & \mathbf{ZAY}\\ \mathbf{O} & \mathbf{WAY}\end{bmatrix},$$

ugyanis $\mathbf{ZAX} = \mathbf{Z}\mu\mathbf{X} = \mu\mathbf{ZX} = \mu\mathbf{I}_g$, és $\mathbf{WAX} = \mu\mathbf{WX} = \mathbf{O}$. Az így kapott mátrix karakterisztikus polinomja

$$\begin{vmatrix}\mu\mathbf{I}_g - \lambda\mathbf{I}_g & \mathbf{ZAY}\\ \mathbf{O} & \mathbf{WAY} - \lambda\mathbf{I}_{n-g}\end{vmatrix},$$

ami a 6.29. tétel szerint $(\mu - \lambda)^g \det(\mathbf{WAY} - \lambda\mathbf{I}_{n-g})$. Ez pedig azt jelenti, hogy $\mathbf{C}^{-1}\mathbf{AC}$ és ezzel együtt $\mathbf{A}$ karakterisztikus polinomjának μ legalább g-szeres algebrai multiplicitású gyöke. □

8.30. TÉTEL (DIAGONALIZÁLHATÓSÁG ÉS A GEOMETRIAI MULTIPLICITÁS). *Egy n-edrendű négyzetes mátrix pontosan akkor diagonalizálható, ha a sajátértékeihez tartozó geometriai multiplicitások összege n.*

BIZONYÍTÁS. ($\Rightarrow$) Ha a mátrix diagonalizálható, akkor az azonos sajátértékhez tartozó vektorok által kifeszített sajátaltér dimenziója megegyezik e sajátérték geometriai multiplicitásával. A geometriai multiplicitások összege tehát épp n, hisz egyetlen sajátvektor sem lehet két sajátaltérben.

($\Leftarrow$) A geometriai multiplicitás nem nagyobb az algebrainál, az algebraiak összege pedig legföljebb n (komplex esetben pontosan n, valós esetben lehet n-nél kisebb is, ha a karakterisztikus polinomnak vannak komplex gyökei). Így ha a geometriai multiplicitások összege n, akkor minden sajátaltérből kiválasztva egy bázist, és véve ezek egyesítését, egy n sajátvektorból álló független vektorrendszert kapunk (ld. a 8.26. tétel utáni megjegyzéseket). Így tehát a mátrix diagonalizálható. □

► E tétel elegáns megfogalmazása így hangzik: egy $\mathbb{T}$ test fölötti n-edrendű mátrix pontosan akkor diagonalizálható, ha a $\mathbb{T}^n$ tér előáll sajátaltereinek direkt összegeként.

► A tétel lineáris leképezésekre is kimondható: az $A : \mathbb{T}^n \to \mathbb{T}^n$ lineáris leképezés pontosan akkor diagonalizálható, ha sajátalterei dimenziójának összege n.

8.31. PÉLDA (LINEÁRIS TRANSZFORMÁCIÓ DIAGONALIZÁLÁSA). *Az alábbi lineáris leképezésekhez keressünk – pusztán geometriai szemléletünkre hagyatkozva – olyan bázist, melyben a mátrixa diagonális. Használjuk fel a 8.21. példa eredményeit.*

a) a sík vektorainak tükrözése egy egyenesre (vagy pontjainak tükrözése egy origón átmenő egyenesre);

b) a sík vektorainak merőleges vetítése egy egyenesre (vagy pontjainak merőleges vetítése egy origón átmenő egyenesre);

c) a tér vektorainak elforgatása egy egyenes körül a π egész számú többszörösétől különböző szöggel;

d) a tér vektorainak merőleges vetítése egy síkra;

e) a tér vektorainak tükrözése egy síkra.

MEGOLDÁS. A 8.21. példában meghatároztuk e leképezések sajátaltereit. Ezeket használjuk a következőkben.

a) Az egyenes – melyre tükrözünk – egyik irányvektora legyen $\mathbf{a}$, egy rá merőleges nemnulla vektor legyen $\mathbf{b}$. Ekkor $T\mathbf{a} = \mathbf{a}$ és $T\mathbf{b} = -\mathbf{b}$, ahol T a tükröző lineáris leképezés. Ennek az $\{\mathbf{a}, \mathbf{b}\}$ bázisban a mátrixa

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

b) Az egyenes – melyre vetítünk – egyik irányvektora legyen $\mathbf{a}$, egy rá merőleges nemnulla vektor legyen $\mathbf{b}$. Ekkor $P\mathbf{a} = \mathbf{a}$ és $P\mathbf{b} = -\mathbf{b}$, ahol P a vetítő lineáris leképezés. Ennek az $\{\mathbf{a}, \mathbf{b}\}$ bázisban a mátrixa

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

c) E leképezésnek nincs valós diagonális mátrixa, mert csak egyetlen valós sajáteltere van, és az csak 1-dimenziós: ez a tengely irányvektora által kifeszített altér. A forgástengelyre merőleges sík ugyan nem sajátaltér, de a forgatás önmagába viszi (az ilyet nevezik *invariáns altérnek*), így ennek bázisával egy „diagonálishoz közeli” alakot kaphatunk. Ha a forgás tengelyének egy irányvektora $\mathbf{a}$, a rá merőleges sík egy ortonormált bázisa $\{\mathbf{b}, \mathbf{c}\}$, ahol a $\mathbf{b}$ vektor $\pi/2$ radiánnal való elforgatottja épp $\mathbf{c}$, akkor az $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ bázisban a forgató F leképezés mátrixa

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix},$$

ugyanis $F\mathbf{a} = \mathbf{a}$, $F\mathbf{b} = \cos\alpha\mathbf{b} + \sin\alpha\mathbf{c}$, $F\mathbf{c} = -\sin\alpha\mathbf{b} + \cos\alpha\mathbf{c}$.

d) A sík, melyre vetítünk az 1 sajátértékhez tartozik. Ha ebben választva egy $\{\mathbf{a}, \mathbf{b}\}$ bázist, és $\mathbf{c}$ egy a síkra merőleges nemzérus vektor, akkor $T\mathbf{a} = \mathbf{a}$, $T\mathbf{b} = \mathbf{b}$, $T\mathbf{c} = \mathbf{0}$, így T mátrixa

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

e) A sík, melyre tükrözünk az 1 sajátértékhez tartozik. Ha ebben választva egy $\{\mathbf{a}, \mathbf{b}\}$ bázist, és $\mathbf{c}$ egy a síkra merőleges nemzérus vektor, akkor $T\mathbf{a} = \mathbf{a}$, $T\mathbf{b} = \mathbf{b}$, $T\mathbf{c} = -\mathbf{c}$, így T mátrixa

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

□

Mátrixok hatványai és egyéb függvényei Ha egy folyamat egy $\mathbf{x}_k$ állapotát a következővel egy lineáris $\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k$ kapcsolat fűzi össze, akkor az $\mathbf{x}_k = \mathbf{A}^k\mathbf{x}_0$ összefüggés miatt a folyamatot az $\mathbf{A}$ mátrix hatványai jellemzik. Kérdés lehet például a mátrixhatványok aszimptotikus viselkedése, vagy a nagy kitevőjű hatványok gyors kiszámításának módja.

Diagonális mátrix hatványai igen könnyen számolhatók, csak a főátló elemeit kell hatványozni. Másrészt $(\mathbf{C}^{-1}\mathbf{M}\mathbf{C})^k = \mathbf{C}^{-1}\mathbf{M}^k\mathbf{C}$, ezért a diagonalizálható mátrixok igen könnyen hatványozhatók.

8.32. példa (Mátrixok nagy kitevős hatványai). *Tekintsük az alábbi két „majdnem egyenlő" mátrixot:*

$$\mathbf{A} = \begin{bmatrix} -0.3 & 1.8 \\ -0.6 & 1.8 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -0.3 & 1.8 \\ -0.5 & 1.8 \end{bmatrix}$$

Vizsgáljuk meg hatványaik határértékét, ha a kitevő tart a végtelenhez!

Megoldás. Mindkét mátrixot diagonalizáljuk:

$$\mathbf{\Lambda}_1 = \mathbf{C}^{-1}\mathbf{A}\mathbf{C} = \begin{bmatrix} 0.6 & 0.0 \\ 0.0 & 0.9 \end{bmatrix}, \quad \text{ahol} \quad \mathbf{C} = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{3}{\sqrt{13}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{13}} \end{bmatrix}$$

valamint

$$\mathbf{\Lambda}_2 = \mathbf{D}^{-1}\mathbf{B}\mathbf{D} = \begin{bmatrix} 1.2 & 0.0 \\ 0.0 & 0.3 \end{bmatrix}, \quad \text{ahol} \quad \mathbf{D} = \begin{bmatrix} 1 & 1 \\ \frac{5}{6} & \frac{1}{3} \end{bmatrix}, \quad \mathbf{D}^{-1} = \begin{bmatrix} -\frac{2}{3} & 2 \\ \frac{5}{3} & -2 \end{bmatrix}.$$

Így a k-adik hatvány könnyen számolható:

$$\mathbf{A}^k = \mathbf{C}\begin{bmatrix} 0.6 & 0.0 \\ 0.0 & 0.9 \end{bmatrix}^k \mathbf{C}^{-1} = \mathbf{C}\begin{bmatrix} 0.6^k & 0.0 \\ 0.0 & 0.9^k \end{bmatrix}\mathbf{C}^{-1}$$

Mivel mindkét sajátérték abszolút értéke kisebb 1-nél, ezért $\mathbf{\Lambda}_1^k \to \mathbf{O}$ és így $\mathbf{A}^k \to \mathbf{O}$, ha $k \to \infty$. A $\mathbf{B}$ mátrix esetén

$$\mathbf{B}^k = \mathbf{D}\begin{bmatrix} 1.2 & 0.0 \\ 0.0 & 0.3 \end{bmatrix}^k \mathbf{D}^{-1} = \mathbf{D}\begin{bmatrix} 1.2^k & 0.0 \\ 0.0 & 0.3^k \end{bmatrix}\mathbf{D}^{-1},$$

ami arra vezet, hogy $\mathbf{\Lambda}_2^k \to \left[\begin{smallmatrix} \infty & 0 \\ 0 & 0 \end{smallmatrix}\right]$ és a $\mathbf{D}$ és a $\mathbf{D}^{-1}$ elemeinek előjelét is figyelembe véve így $\mathbf{A}^k \to \left[\begin{smallmatrix} -\infty & \infty \\ -\infty & \infty \end{smallmatrix}\right]$, ha $k \to \infty$. □

Mátrixok ortogonális diagonalizálása Diagonalizálni egy mátrixot azzal ekvivalens, hogy a hozzá tartozó mátrixleképezéshez találni egy olyan bázist, melyben mátrixa diagonális. A gyakorlati alkalmazásokban különösen szerencsés, ha e bázis még ortonormált is! Valós mátrixok közt pontosan a szimmetrikusak azok, amelyek így diagonalizálhatók.

A 8.24. tétel szerint a diagonalizálhatóság szükséges és elégséges feltétele, hogy létezzék a mátrix rendjével egyező számú független sajátvektora. Ha e vektorok ortonormált rendszert alkotnak, akkor a belőlük alkotott mátrix ortogonális mátrix.

8.33. DEFINÍCIÓ (ORTOGONÁLIS DIAGONALIZÁLHATÓSÁG). *Az* $\mathbf{A}$ *mátrix* ortogonálisan diagonalizálható*, ha találunk egy ortogonális* $\mathbf{Q}$ *és egy diagonális* $\mathbf{\Lambda}$ *mátrixot, hogy* $\mathbf{Q}^T\mathbf{A}\mathbf{Q} = \mathbf{\Lambda}$.

Először megmutatjuk, hogy szimmetrikus mátrix esetén a sajátalterek nem csak függetlenek egymástól, de merőlegesek is egymásra.

8.34. TÉTEL (SZIMMETRIKUS MÁTRIX SAJÁTALTEREI). *Szimmetrikus mátrix bármely két különböző sajátaltere merőleges egymásra.*

BIZONYÍTÁS. Két különböző sajátaltér két különböző sajátértékhez tartozik. Megmutatjuk, hogy az egyik altér bármelyik vektora merőleges a másik altér bármely vektorára. Legyen tehát $(\lambda, \mathbf{x})$ és $(\mu, \mathbf{y})$ két sajátpár, ahol $\lambda \neq \mu$ két különböző sajátértéke $\mathbf{A}$-nak. Így $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ és $\mathbf{A}\mathbf{y} = \mu\mathbf{y}$. Ebből adódik, hogy

$$\lambda(\mathbf{x}^T\mathbf{y}) = (\lambda\mathbf{x})^T\mathbf{y} = (\mathbf{A}\mathbf{x})^T\mathbf{y} = \mathbf{x}^T\mathbf{A}^T\mathbf{y} = \mathbf{x}^T\mathbf{A}\mathbf{y} = \mathbf{x}^T\mu\mathbf{y} = \mu(\mathbf{x}^T\mathbf{y}).$$

Eszerint $(\lambda - \mu)(\mathbf{x}^T\mathbf{y}) = 0$, de $\lambda - \mu \neq 0$, ezért $\mathbf{x}^T\mathbf{y} = \mathbf{x} \cdot \mathbf{y} = 0$, azaz a két vektor merőleges egymásra. □

8.35. TÉTEL (VALÓS SPEKTRÁLTÉTEL). *A valós* $\mathbf{A}$ *mátrix pontosan akkor diagonalizálható ortogonálisan, ha szimmetrikus.*

BIZONYÍTÁS. $(\Rightarrow)$ Tegyük fel, hogy $\mathbf{A}$ ortogonálisan diagonalizálható, azaz létezik olyan ortogonális $\mathbf{Q}$ és diagonális $\mathbf{\Lambda}$ mátrix, hogy $\mathbf{Q}^T\mathbf{A}\mathbf{Q} = \mathbf{\Lambda}$. Ezt balról a $\mathbf{Q}$, jobbról a $\mathbf{Q}^T$ mátrixszal szorozva kapjuk, hogy $\mathbf{Q}\mathbf{Q}^T\mathbf{A}\mathbf{Q}\mathbf{Q}^T = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$. Mivel $\mathbf{Q}$ ortogonális, ezért $\mathbf{Q}\mathbf{Q}^T = \mathbf{I}$, így $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$. Ekkor

$$\mathbf{A}^T = (\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T)^T = (\mathbf{Q}^T)^T\mathbf{\Lambda}^T\mathbf{Q}^T = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T = \mathbf{A}.$$

Tehát $\mathbf{A}$ szimmetrikus.

($\Leftarrow$) Tegyük fel, hogy $\mathbf{A}$ szimmetrikus, azaz $\mathbf{A}^T = \mathbf{A}$. Ekkor minden sajátértéke valós, sajátértékei algebrai multiplicitásának összege tehát n. Először megmutatjuk, hogy minden sajátérték geometriai multiplicitása megegyezik algebrai multiplicitásával, tehát a mátrix diagonalizálható. Legyen μ egy tetszőleges sajátérték, geometriai multiplicitása legyen g. Ekkor sajátalteréből kiválaszthatunk egy $\mathbf{b}_1, \dots, \mathbf{b}_g$ bázist, amelyet kiegészítünk a teljes tér $\mathbf{b}_1, \dots, \mathbf{b}_g, \dots, \mathbf{b}_n$ bázisává. Az e vektorokból képzett $\mathbf{B}$ mátrixra tehát igaz az

$$\mathbf{AB} = \mathbf{B}\begin{bmatrix} \mu\mathbf{I}_g & \mathbf{X} \\ \mathbf{O} & \mathbf{Y} \end{bmatrix}$$

összefüggés. Mivel $\mathbf{B}$ oszlopai lineárisan függetlenek, ezért $\mathbf{B}$ invertálható, és így $\mathbf{B}^{-1}\mathbf{AB} = \left[\begin{smallmatrix} \mu\mathbf{I}_g & \mathbf{X} \\ \mathbf{O} & \mathbf{Y} \end{smallmatrix}\right]$. Eszerint $\mathbf{A} \sim \left[\begin{smallmatrix} \mu\mathbf{I}_g & \mathbf{X} \\ \mathbf{O} & \mathbf{Y} \end{smallmatrix}\right]$, tehát a két mátrixnak megegyeznek a sajátértékei. Mivel $\mathbf{Y}$-nak már nem lehet μ a sajátértéke, és a karakterisztikus polinom

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{bmatrix} (\mu - \lambda)\mathbf{I}_g & \mathbf{X} \\ \mathbf{O} & \mathbf{Y} - \lambda\mathbf{I} \end{bmatrix} = (\mu - \lambda)^g \det(\mathbf{Y} - \lambda\mathbf{I}),$$

ezért μ algebrai multiplicitása is g, amivel bizonyítottuk, hogy $\mathbf{A}$ diagonalizálható. Már csak annak bizonyítása maradt, hogy a sajátvektorokból ortonormált bázis is kiválasztható, ez viszont abból következik, hogy szimmetrikus mátrix sajátalterei merőlegesek egymásra, így ha minden sajátaltérből kiválasztunk egy ortonormált bázist, akkor azok egyesítése az egész tér ortonormált bázisa lesz (ld. még erről a 8.26. tételt és az azt követő második megjegyzést). □

8.36. példa. *Diagonalizáljuk az alábbi mátrixot ortogonálisan!*

$$\begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}.$$

Határozzuk meg az áttérés mátrixát a standardról arra a bázisra, melyben e mátrix diagonális alakot vesz fel!

Megoldás. A karakterisztikus polinom:

$$\begin{vmatrix} 3-\lambda & 1 & 1 \\ 1 & 3-\lambda & 1 \\ 1 & 1 & 3-\lambda \end{vmatrix} = -\lambda^3 + 9\lambda^2 - 24\lambda + 20,$$

melynek gyökei 2, 2 és 5. Tehát a diagonális alak $\operatorname{diag}(2,2,5)$. Az áttérés mátrixához szükségünk lesz a sajátvektorokra. $\lambda = 2$ esetén:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

az $x+y+z=0$ egyenletrendszer megoldása pedig $(x,y,z) = (-1,1,0)s + (-1,0,1)t$. Ennek az altérnek egy bázisa tehát a $(-1,1,0)$ és a $(-1,0,1)$ vektorokból áll. A $\lambda = 5$ esetén:

$$\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix},$$

amely egyenletrendszer megoldása $(x,y,z) = (1,1,1)t$. A két különböző sajátaltérből való vektorok merőlegesek egymásra, de a $\lambda = 2$-höz tartozó sajátaltér két sajátvektora nem alkot ortogonális rendszert, ezért az általuk kifeszített térben új bázist keresünk, egy ortonormáltat. Legyen az $\mathbf{a} = (-1,1,0)/\sqrt{2}$ vektor az egyik, ekkor

$$(-1,0,1) - \left((-1,0,1)\cdot\frac{(-1,1,0)}{\sqrt{2}}\right)\frac{(-1,1,0)}{\sqrt{2}} = \left(-\frac{1}{2},-\frac{1}{2},1\right).$$

Ezt normálva kapjuk a $\mathbf{b} = (-\frac{1}{\sqrt{6}},-\frac{1}{\sqrt{6}},-\frac{\sqrt{2}}{\sqrt{3}})$ vektort, végül a $\lambda = 5$-höz tartozó normált vektor $\mathbf{c} = \frac{1}{\sqrt{3}}(1,1,1)$. A standard bázisra való áttérés mátrixa tehát az $[\mathbf{a}|\mathbf{b}|\mathbf{c}]$ mátrix, azaz

$$\begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}.$$

Ennek inverze lesz a standard bázisról való áttérés mátrixa, mely ortogonális mátrixról lévén szó, a transzponáltja, azaz

$$\begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{\sqrt{2}}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}.$$

□

Kvadratikus formák

A csupa másodfokú tagot tartalmazó többváltozós polinomok mátrixok sajátértékeinek és sajátvektorainak ismeretében egyszerűbb alakra hozhatók, így könnyebben vizsgálhatók. E témának számtalan lineáris algebrán kívüli matematikai és matematikán kívüli alkalmazása is van.

Homogén másodfokú polinomok mátrixszorzatos alakja Egy polinom egy tagja *másodfokú*, ha abban az ismeretlenek fokszámainak összege 2. Például az x, y és z változókban másodfokú tagok az alábbiak: $3x^2$, axy, $2b^3xz$, $-\pi^2z^2$. Az olyan többváltozós polinomot, melyben csak másodfokú tagok vannak, többváltozós *homogén másodfokú polinomnak* nevezzük. Például $2x^2+4xy-y^2$ egy 2-változós homogén másodfokú polinom. Mivel $xy = yx$, ezért ezt a polinomot ekvivalens módon több más alakba is át lehet írni, például: $2x^2+3xy+yx-y^2$, $2x^2-5xy+9yx-y^2, \dots, 2x^2+2xy+2yx-y^2$. Az utolsó formula egy mátrixszorzatos fölírás lehetőségét sejteti:

$$2x^2+2xy+2yx-y^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Valóban, általában is igaz, hogy

$$ax^2+2bxy+cy^2 = ax^2+bxy+byx+cy^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Hasonlóképp a háromváltozós homogén másodfokú polinomok is mátrixszorzatos alakba írhatók:

$$\begin{aligned} & ax^2+2bxy+2cxz+dy^2+2eyz+fz^2 \\ = {} & ax^2+bxy+cxz+byx+dy^2+eyz+czx+ezy+fz^2 \\ = {} & \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \end{aligned}$$

8.37. PÉLDA (MÁSODFOKÚ POLINOM MÁTRIXSZORZATOS ALAKJA). *Írjuk fel az $x_1^2+2x_2^2+2x_3^2-5x_1x_2-3x_2x_1+5x_1x_3-x_3x_1$ kifejezést mátrixszorzatos alakban!*

MEGOLDÁS. A vegyes tagokat először összevonva, majd két egyenlő együtthatójú részre bontva kapjuk, hogy

$$\begin{aligned} & x_1^2+2x_2^2+2x_3^2-8x_1x_2+4x_1x_3 \\ = {} & x_1^2+2x_2^2+2x_3^2-4x_1x_2-4x_2x_1+2x_1x_3+2x_3x_1 \\ = {} & x_1^2-4x_1x_2+2x_1x_3-4x_2x_1+2x_2^2+0x_2x_3+2x_3x_1+0x_3x_2+2x_3^2 \\ = {} & \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 1 & -4 & 2 \\ -4 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \end{aligned}$$ □

A fentieket követve az $\mathbf{x} = (x_1, x_2, \dots, x_n)$ vektor koordinátáitól függő homogén másodfokú polinomok mindegyike

$$\mathbf{x}^T\mathbf{A}\mathbf{x}$$

alakra hozható, ahol $\mathbf{A}$ szimmetrikus mátrix. Ezt nevezzük kvadratikus alaknak, vagy kvadratikus formának. Tipikus alkalmazásokban a mátrix valós elemű. Ennek megfelelően e szakaszban a továbbiakban kvadratikus formán mindig valós kvadratikus formát értünk, tehát *kvadratikus formának* nevezzük azt az

$$\mathbb{R}^n \to \mathbb{R}; \mathbf{x} \mapsto \mathbf{x}^T \mathbf{A}\mathbf{x}$$

függvényt, ahol $\mathbf{A}$ szimmetrikus mátrix.

Főtengelytétel Egy kvadratikus formához tartozó szimmetrikus mátrix diagonalizálásával a kvadratikus forma is egyszerű alakra hozható.

A spektráltétel szerint minden valós szimmetrikus mátrix ortogonálisan diagonalizálható, azaz létezik egy olyan ortogonális $\mathbf{Q}$ mátrix, és egy diagonális $\mathbf{\Lambda}$ mátrix, melyre $\mathbf{Q}^T\mathbf{A}\mathbf{Q} = \mathbf{\Lambda}$. Tudjuk, hogy az $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ mátrixleképezés mátrixa a $\mathbf{Q}$ oszlopvektorai alkotta $\mathcal{Q}$ ortonormált bázisban $\mathbf{\Lambda}$. Ha egy tetszőleges $\mathbf{x}$ vektor alakja e bázisban $\mathbf{y}$, akkor $\mathbf{x} = \mathbf{Q}\mathbf{y}$. E helyettesítést elvégezve ugyanennek a függvénynek a $\mathcal{Q}$ bázisban fölírt alakját kapjuk:

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = (\mathbf{Q}\mathbf{y})^T\mathbf{A}(\mathbf{Q}\mathbf{y}) = \mathbf{y}^T\mathbf{Q}^T\mathbf{A}\mathbf{Q}\mathbf{y} = \mathbf{y}^T\mathbf{\Lambda}\mathbf{y}.$$

Eszerint, a kvadratikus forma e bázisban nagyon egyszerűvé válik, csak négyzetes tagokat tartalmaz: $\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$, ahol $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Ezzel bizonyítottuk az alábbi tételt:

8.38. TÉTEL (FŐTENGELYTÉTEL). *Ha $\mathbf{A}$ egy n-edrendű valós szimmetrikus mátrix, és a $\mathbf{Q}$ mátrix ortogonálisan diagonalizálja, azaz $\mathbf{Q}^T\mathbf{A}\mathbf{Q} = \mathbf{\Lambda}$ diagonális, akkor az $\mathbf{x} = \mathbf{Q}\mathbf{y}$ helyettesítés az $\mathbf{x}^T\mathbf{A}\mathbf{x}$ kvadratikus formát a $\mathbf{y}^T\mathbf{\Lambda}\mathbf{y}$ kvadratikus formába transzformálja, mely kifejtve csak négyzetes tagokat tartalmaz, azaz*

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{y}^T\mathbf{\Lambda}\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2, \tag{8.7}$$

ahol λ_1, λ_2,..., λ_n az $\mathbf{A}$ mátrix sajátértékei.

► A tétel nevét később fogjuk részletesen megmagyarázni, most csak annyit, hogy az $\mathbf{x}^T\mathbf{A}\mathbf{x} = c$ egyenletű felületnek a $\mathcal{Q}$ bázis vektorai mind szimmetriatengelyei, melyeket főtengelyeknek is nevezünk.

► Mivel $\mathbf{Q}$ ortogonális mátrix, ezért $\det \mathbf{Q} = 1$ vagy $\det \mathbf{Q} = -1$. Gyakorlati (például bizonyos 3-dimenziós) alkalmazásokban fontos lehet, hogy a $\mathcal{Q}$ bázis is jobbsodrású legyen, azaz hogy $\det \mathbf{Q} = 1$ legyen. Így a standard bázis beleforgatható az új bázisba. Ez elérhető, ha $\det \mathbf{Q} = -1$ esetén $\mathbf{Q}$ bármelyik oszlopát -1-szeresére változtatjuk. Ez a kvadratikus formán nem változtat, hisz abban csak a sajátértékek szerepelnek.

► A főtengelytétel alkalmazását egy kvadratikus formán *főtengely-transzformációnak* nevezzük.

8.39. példa (Főtengely-transzformáció). *Végezzük el a főtengely-transzformációt az*

$$f(x,y,z) = x^2 + 2y^2 + 2z^2 - 8xy + 4xz$$

kvadratikus formán. Keressünk olyan jobbsodrású ortonormált bázist, melyben épp ez a kvadratikus forma alakja. Mi az áttérés mátrixa?

Megoldás. A kvadratikus forma mátrixszorzat-alakja

$$\begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 1 & -4 & 2 \\ -4 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Mátrixának karakterisztikus polinomja $-\lambda^3 + 5\lambda^2 + 12\lambda - 36$, ennek gyökei, azaz a sajátértékek 6, -3, 2, a hozzájuk tartozó sajátvektorok rendre $(2,-2,1)$, $(-5,-4,2)$, $(0,1,2)$. Így a keresett kvadratikus forma

$$\begin{bmatrix} \xi & \eta & \zeta \end{bmatrix} \begin{bmatrix} 6 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} \xi \\ \eta \\ \zeta \end{bmatrix} = 6\xi^2 - 3\eta^2 + 2\zeta^2.$$

A sajátvektorokat normálva megkapjuk az ortonormált bázist, melynek vektoraiból képzett determináns

$$\begin{vmatrix} 2/9 & -5/3\sqrt{5} & 0 \\ -2/9 & -4/3\sqrt{5} & 1/\sqrt{5} \\ 1/9 & 2/3\sqrt{5} & 2/\sqrt{5} \end{vmatrix} = -1,$$

tehát egy megfelelő ortonormált bázis: $(\frac{2}{9}, -\frac{2}{9}, \frac{1}{9})$, $(\frac{5}{3\sqrt{5}}, \frac{4}{3\sqrt{5}}, -\frac{2}{3\sqrt{5}})$, $(0, \frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}})$. □

Kvadratikus formák és mátrixok definitsége A főtengelytétel könnyen áttekinthetővé teszi a kvadratikus forma által fölvett értékek lehetséges előjelét. Ez lehetővé teszi a kvadratikus formák egy fontos osztályozását.

8.40. definíció (Kvadratikus formák és mátrixok definitsége). *Azt mondjuk, hogy az $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ kvadratikus forma*

a) pozitív definit, *ha $f(\mathbf{x}) > 0$,*
b) pozitív szemidefinit, *ha $f(\mathbf{x}) \geq 0$,*
c) negatív definit, *ha $f(\mathbf{x}) < 0$,*
d) negatív szemidefinit, *ha $f(\mathbf{x}) \leq 0$*

bármely $\mathbf{x} \neq \mathbf{0}$ vektor esetén, és azt mondjuk, hogy f

e) indefinit, *ha pozitív és negatív értékeket is fölvesz.*
Azt mondjuk, hogy a szimmetrikus **A** *mátrix pozitív/negatív definit/szemidefinit, illetve indefinit, ha a hozzá tartozó kvadratikus forma az.*

► Ha **A** negatív definit, akkor $-\mathbf{A}$ pozitív definit. Hasonló állítás igaz a szemidefiniségre is.
► Ha $\mathbf{A} = [a]$, azaz $\mathbf{A}$ 1×1-es, akkor **A** pontosan akkor pozitív definit, ha $a > 0$.
► ....
► Világos, hogy ha egy kvadratikus formában csak négyzetes tagok szerepelnek, akkor azonnal leolvasható definitségének típusa. Például az $f(x,y) = x^2 + 2y^2$, $g(x,y) = x^2 - 2y^2$, $h(x,y) = -x^2 - 2y^2$, $k(x,y,z) = x^2 + 2y^2$ formákról látható, hogy f pozitív definit, hisz az $(x,y) \neq (0,0)$ esetén értéke mindig pozitív, g indefinit, h negatív definit, és k pozitív szemidefinit, hisz értéke $(x,y,z) \neq (0,0,0)$ esetén is lehet 0 (ha $x = y = 0$, de $z \neq 0$). Miután a főtengelytétel szerint minden kvadratikus forma egyenlő változók négyzeteinek a sajátértékekkel vett lineáris kombinációjával, ezért a definitség típusa pusztán csak a sajátértékek előjeleinek ismeretében eldönthető.

8.41. PÉLDA (DEFINITSÉG MEGHATÁROZÁSA A SAJÁTÉRTÉKEKBŐL). *Határozzuk meg az*

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \qquad \mathbf{C} = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}$$

mátrixok definitségének típusát!

MEGOLDÁS. Az **A** mátrix sajátértékei 1, 1 és 4. Így a főtengely-transzformáció után kapott

$$\begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \xi & \eta & \zeta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} \xi & \eta & \zeta \end{bmatrix} = \xi^2 + \eta^2 + 4\zeta^2$$

alakból látható, hogy e kvadratikus forma minden értéke pozitív, ha a változók nem mindegyike 0. Tehát e kvadratikus forma pozitív definit. Hasonlóképp a **B** sajátértékei -1, -1 és 2, a főtengely-transzformáció után kapott alak $-\xi^2 - \eta^2 + 4\zeta^2$. E forma negatív értéket vesz fel például az $(1,0,0)$ helyen és pozitívat a $(0,0,1)$ helyen, tehát indefinit. Végül **C** sajátértékei -3, -3 és 0, így a főtengely-transzformáció után kapott alak $-3\xi^2 - 3\eta^2 + 0\zeta^2 = -3\xi^2 - 3\eta^2$. Ennek értéke a $(0,0,1)$ helyen 0, és pozitív értéket nem vesz fel, tehát negatív szemidefinit. □

8.42. tétel (Definitség meghatározása a sajátértékekből). *A valós szimmetrikus* $\mathbf{A}$ *mátrix, illetve az* $\mathbf{x}^T\mathbf{A}\mathbf{x}$ *kvadratikus forma pontosan akkor*

a) pozitív definit, ha $\mathbf{A}$ *minden sajátértéke pozitív;*

b) pozitív szemidefinit, ha $\mathbf{A}$ *minden sajátértéke nemnegatív;*

c) negatív definit, ha $\mathbf{A}$ *minden sajátértéke negatív;*

d) negatív szemidefinit, ha $\mathbf{A}$ *minden sajátértéke nempozitív;*

e) indefinit, ha $\mathbf{A}$*-nak van pozitív és negatív sajátértéke is.*

Kúpszeletek osztályozása

Definitség és sajátértékek

Szélsőérték

Szélsőérték az egységgömbön

9 Szinguláris érték

A szimmetrikus mátrix ortogonális diagonalizálását fogjuk általánosítani tetszőleges mátrixra, egy helyett két ortonormált bázis megkeresésével. Az egyik bázis a mátrixleképezés értemezési tartományának, a másik az értékkészletnek lesz bázisa. Ebben az általánosításban a sajátértékek szerepét a szinguláris értékek veszik át, a spektrálfelbontásét a szinguláris érték szerinti felbontás (SVD). Az alkalmazások közül kiemelkednek az információtömörítéssel kapcsolatosak, de az egyenletrendszerek megoldásához használt leghatékonyabbak közé tartozó algoritmusok is ide sorolhatók.

Szinguláris érték, szinguláris vektor, SVD

Azt tudjuk, hogy egy mátrixleképezés kölcsönösen egyértelmű a sortér és az oszloptér között. Olyan ortonormált bázisát keressük a sortérnek, és az oszloptérnek, melyek közt a mátrixleképezés természetes kapcsolatot létesít.

Szinguláris érték Az, hogy az $n \times n$-es $\mathbf{A}$ mátrix ortogonálisan diagonalizálható, azt jelenti, hogy létezik egy olyan $\{\mathbf{v}_1, \mathbf{v}_2 \ldots, \mathbf{v}_n\}$ ortonormált bázis, és léteznek olyan λ_i $(i = 1, 2, \ldots, n)$ számok, hogy $\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i$ minden i indexre. Ha a mátrix téglalapalakú, például $m \times n$-es, azaz az értelmezési tartomány és az értékkészlet különböző tér is lehet, akkor mindkettőben választanunk kell egy bázist. Legyen $\{\mathbf{v}_1, \mathbf{v}_2 \ldots, \mathbf{v}_n\}$ az értelmezési tartomány ONB-a, az értékkészleté pedig $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$. Az analógia $\mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i$ alakú feltételek kikötését kívánja. Az $m > n$ és $m < n$ esetek nehézségeit leküzdendő szorítkozzunk a sortérre, azt ugyanis tudjuk, hogy az $A : \mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ leképezés kölcsönösen egyértelmű a sortér és az oszloptér között. Így először csak e két altérben keressünk megfelelő bázist. Közös dimenziójuk a ranggal egyenlő, jelölje ezt r.

9.1. definíció (Szinguláris érték). *Azt mondjuk, hogy a pozitív $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ számok az r-rangú valós vagy komplex* $\mathbf{A}$ *mátrix* szinguláris értékei, *ha van olyan $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ ortonormált bázisa a sortérnek, és $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ ortonormált bázisa az oszloptérnek, hogy*

$$\mathbf{A}\mathbf{v}_i = \sigma\mathbf{u}_i, \quad i = 1, \ldots, r.$$

A *szinguláris érték* fogalmát Erhard Schmidt vezette be 1907-ben, de ő még *sajátértéknek* nevezte. Mai nevét 1937-ben kapta, mert különösen akkor tűnt hasznos eszköznek – például az egyenletrendszerek megoldásában –, amikor az együtthatómátrix szinguláris.

► E definíció fontos következménye, hogy mivel $|\mathbf{u}_i| = 1$ $(i = 1, 2, \ldots, r)$, ezért $|\mathbf{A}\mathbf{v}_i| = \sigma_i$.

► A szinguláris értékekre adható olyan definíció is, melyben a két bázis létezése nem szerepel, csak az $\mathbf{A}$ mátrix – ezt később megemlítjük –, de didaktikai okból nem használjuk.

► A szinguláris értékeket úgy is definiálhatók, hogy ha $k = \min(m, n)$, akkor a szinguláris értékek száma k, és csak annyit kötünk ki róluk, hogy nemnegatívak. Látni fogjuk, hogy ha $k > r$, akkor e definíció mellett $\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_k = 0$.

9.2. példa (Szinguláris értékek). *Igazoljuk, hogy a*

$$\{\mathbf{v}_1, \mathbf{v}_2\} = \{(4/5, -3/5), (3/5, 4/5)\},$$
$$\{\mathbf{u}_1, \mathbf{u}_2\} = \{(-5/13, 12/13), (12/13, 5/13)\}$$

ONB-ok választása mellett az

$$\mathbf{A} = \begin{bmatrix} -4/13 & 6 \\ 111/13 & -4 \end{bmatrix}$$

mátrix szinguláris értékei $\sigma_1 = 10$ és $\sigma_2 = 5$.

Megoldás. Ez nyilvánvalóan igaz, hisz

$$\begin{bmatrix} -4/13 & 6 \\ 111/13 & -4 \end{bmatrix} \begin{bmatrix} 4/5 \\ -3/5 \end{bmatrix} = 10 \begin{bmatrix} -5/13 \\ 12/13 \end{bmatrix}, \quad \begin{bmatrix} -4/13 & 6 \\ 111/13 & -4 \end{bmatrix} \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix} = 5 \begin{bmatrix} 12/13 \\ 5/13 \end{bmatrix},$$

azaz fennállnak az $\mathbf{A}\mathbf{v}_1 = \sigma_1\mathbf{u}_1$ és az $\mathbf{A}\mathbf{v}_2 = \sigma_2\mathbf{u}_2$ összefüggések. □

Egyelőre nem tudjuk sem azt, hogy egy tetszőleges valós vagy komplex mátrixnak vannak-e szinguláris értékei, és ha igen, egyértelműek-e, azaz például függnek-e a bázis megválasztásától. Ha léteznek, akkor az $\mathbf{A}$ mátrix egy – a spektrálfelbontásra emlékeztető – felbontását kapjuk.

Szinguláris felbontás Képezzük a szinguláris értékekből a diagonális

$$\mathbf{\Sigma}_1 = \operatorname{diag}(\sigma_1, \ldots, \sigma_r) = \begin{bmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_r \end{bmatrix}$$

mátrixot, valamint a bázisvektorokból az $\mathbf{U}_1 = \{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ és a $\mathbf{V}_1 = \{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ mátrixokat. Ekkor az $\mathbf{A}\mathbf{v}_i = \sigma \mathbf{u}_i$ egyenlőségek az

$$\mathbf{A}\mathbf{V}_1 = \mathbf{U}_1\mathbf{\Sigma}_1, \tag{9.1}$$

azaz az

$$\mathbf{A}\begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_r\end{bmatrix} = \begin{bmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_r\end{bmatrix}\begin{bmatrix}\sigma_1 & 0 & \dots & 0\\ 0 & \sigma_2 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \sigma_r\end{bmatrix} \tag{9.2}$$

alakot öltik. Egészítsük ki a $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ vektorrendszert a teljes n-dimenziós tér ONB-ává, és jelöljük a belőlük képzett $n \times n$-es ortogonális mátrixot $\mathbf{V}$-vel, $n > r$ esetén az új vektorokból képzett mátrixot $\mathbf{V}_2$-vel, azaz $\mathbf{V}_2 = \begin{bmatrix}\mathbf{v}_{r+1} & \dots & \mathbf{v}_n\end{bmatrix}$. Mivel $\mathbf{V}_2$ oszlopvektorai merőlegesek a sortérre, ezért a nulltérben vannak, tehát $r < i \le n$ esetén $\mathbf{A}\mathbf{v}_i = \mathbf{0}$.

Hasonlóképp az előzőkhöz egészítsük ki az $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ vektorrendszert a teljes m-dimenziós tér ONB-ává, és jelöljük e vektorokból képzett $m \times m$-es mátrixot $\mathbf{U}$-val, és $m > r$ esetén legyen $\mathbf{U}_2 = \begin{bmatrix}\mathbf{u}_{r+1} & \dots & \mathbf{u}_m\end{bmatrix}$. Végül a $\mathbf{\Sigma}_1 = \operatorname{diag}(\sigma_1, \dots, \sigma_r)$ mátrixot egészítsük ki egy $m \times n$-es mátrixszá nullblokkok hozzávételével, jelölje e mátrixot $\mathbf{\Sigma}$, tehát $\mathbf{\Sigma} = \left[\begin{smallmatrix}\mathbf{\Sigma}_1 & \mathbf{O}\\ \mathbf{O} & \mathbf{O}\end{smallmatrix}\right]$. Ekkor a (9.2) egyenlőség a következőképp módosítható:

$$\begin{aligned}
\mathbf{A}\mathbf{V} &= \left[\begin{array}{ccc|ccc}\mathbf{A}\mathbf{v}_1 & \dots & \mathbf{A}\mathbf{v}_r & \mathbf{A}\mathbf{v}_{r+1} & \dots & \mathbf{A}\mathbf{v}_n\end{array}\right]\\
&= \left[\begin{array}{ccc|ccc}\sigma_1\mathbf{u}_1 & \dots & \sigma_r\mathbf{u}_r & \mathbf{0} & \dots & \mathbf{0}\end{array}\right]\\
&= \left[\begin{array}{ccc|ccc}\mathbf{u}_1 & \dots & \mathbf{u}_r & \mathbf{u}_{r+1} & \dots & \mathbf{u}_m\end{array}\right]\left[\begin{array}{cccc|ccc}\sigma_1 & 0 & \dots & 0 & 0 & \dots & 0\\ 0 & \sigma_2 & \dots & 0 & 0 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \dots & \vdots\\ 0 & 0 & \dots & \sigma_r & 0 & \dots & 0\\ \hline 0 & 0 & \dots & 0 & 0 & \dots & 0\\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots\\ 0 & 0 & \dots & 0 & 0 & \dots & 0\end{array}\right]\\
&= \mathbf{U}\mathbf{\Sigma}
\end{aligned}$$

A mátrixok méreteit is kiírva $\mathbf{A}_{m\times n}\mathbf{V}_{n\times n} = \mathbf{U}_{m\times m}\mathbf{\Sigma}_{m\times n}$, blokkmátrix alakba átírva

$$\mathbf{A}\left[\begin{array}{c|c}\mathbf{V}_1 & \mathbf{V}_2\end{array}\right] = \left[\begin{array}{c|c}\mathbf{U}_1 & \mathbf{U}_2\end{array}\right]\left[\begin{array}{c|c}\mathbf{\Sigma}_1 & \mathbf{O}\\ \hline \mathbf{O} & \mathbf{O}\end{array}\right]$$

Ha $r = n$, illetve $r = m$, akkor $\mathbf{V}_2$, illetve $\mathbf{U}_2$ üresek, azaz 0 számú oszlopból állnak, ami értelemszerűen változtat e képleten. Mivel a négyzetes $\mathbf{V}$ mátrix oszlopvektorai ONB-t alkotnak, ezért valós esetben $\mathbf{V}$ ortogonális, így $\mathbf{V}^{-1} = \mathbf{V}^T$, komplex esetben unitér, és $\mathbf{V}^{-1} = \mathbf{V}^H$. Ezt fölhasználva, az $\mathbf{AV} = \mathbf{U\Sigma}$ egyenlőségből kapjuk, hogy $\mathbf{A} = \mathbf{U\Sigma V}^T$ ($\mathbf{A} = \mathbf{U\Sigma V}^H$). Ezt nevezzük a valós (komplex) $\mathbf{A}$ mátrix *szinguláris érték szerinti felbontásának*, vagy röviden *szinguláris felbontásának*.

Tekintsük e felbontás blokkmátrixalakját. A műveleteket blokkonként elvégezve, a $\mathbf{\Sigma}$-ban lévő nullmátrixok miatt a következőt kapjuk:

$$\mathbf{A} = \left[\mathbf{U}_1 \;\middle|\; \mathbf{U}_2\right] \left[\begin{array}{c|c} \mathbf{\Sigma}_1 & \mathbf{O} \\ \hline \mathbf{O} & \mathbf{O} \end{array}\right] \left[\begin{array}{c} \mathbf{V}_1^T \\ \hline \mathbf{V}_2^T \end{array}\right] = \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^T.$$

Az $\mathbf{A} = \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^T$ (komplex esetben $\mathbf{A} = \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^H$) felbontást *redukált szinguláris felbontásnak* nevezzük. Ha ezt a szorzatot az $\mathbf{U}_1\mathbf{\Sigma}_1$ oszlopvektorokra, és a $\mathbf{V}_1^T$ sorvektorokra blokkosított alakján végezzük, az $\mathbf{A}$ mátrix egy diadikus felbontását kapjuk, melyet *szinguláris érték szerinti diadikus felbontásnak* vagy a *szinguláris felbontás diadikus alakjának* nevezünk:

$$\mathbf{A} = \sigma_1\mathbf{u}_1\mathbf{v}_1^T + \sigma_2\mathbf{u}_2\mathbf{v}_2^T + \cdots + \sigma_r\mathbf{u}_r\mathbf{v}_r^T.$$

(E felbontások mindegyikének komplex alakjában transzponált helyett adjungált áll.)

9.3. PÉLDA (SZINGULÁRIS FELBONTÁSOK). *Igazoljuk, hogy az*

$$\begin{bmatrix} 0 & -2 & 2 \\ -2 & 3 & -2 \\ 4 & -2 & 0 \end{bmatrix} = \begin{bmatrix} 1/3 & -2/3 & -2/3 \\ -2/3 & 1/3 & -2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix} \begin{bmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2/3 & -2/3 & 1/3 \\ 2/3 & 1/3 & -2/3 \\ 1/3 & 2/3 & 2/3 \end{bmatrix}$$

felbontás szinguláris felbontás. Ennek alapján írjuk fel az egyenlőség bal oldalán álló mátrix redukált szinguláris felbontását, és a szinguláris felbontás diadikus alakját!

MEGOLDÁS. Először ellenőriznünk kell, hogy valóban egy $\mathbf{A} = \mathbf{U\Sigma V}^T$ alakú egyenlőségről van szó, azaz hogy

a) az egyenlőség fennáll,

b) az $\mathbf{U}$ és $\mathbf{V}$ mátrixok ortogonálisak, $\mathbf{\Sigma}$ diagonális.

Könnyen ellenőrizhető, hogy igen. Ebből már következik, hogy az $\mathbf{U}$ első két oszlopát $\mathbf{u}_1$-gyel, illetve $\mathbf{u}_2$-vel, a $\mathbf{V}^T$ első két sorát $\mathbf{v}_1^T$-vel, illetve $\mathbf{v}_2^T$-vel jelölve fennállnak az $\mathbf{Av}_1 = 6\mathbf{u}_1$ és az $\mathbf{Av}_2 = 3\mathbf{u}_2$ egyenlőségek. Mivel $\mathbf{\Sigma}$-ról látható, hogy $\mathrm{r}(\mathbf{A}) = 2$, ezért $\mathbf{U}$ első 2 oszlopát, $\mathbf{V}^T$ első 2 sorát, $\mathbf{\Sigma}$ bal felső 2×2-es blokkját megtartva a redukált alakot kapjuk:

$$\begin{bmatrix} 0 & -2 & 2 \\ -2 & 3 & -2 \\ 4 & -2 & 0 \end{bmatrix} = \begin{bmatrix} 1/3 & -2/3 \\ -2/3 & 1/3 \\ 2/3 & 2/3 \end{bmatrix} \begin{bmatrix} 6 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2/3 & -2/3 & 1/3 \\ 2/3 & 1/3 & -2/3 \end{bmatrix}$$

Végül ennek alapján a szinguláris érték szerinti diadikus felbontás:

$$\begin{aligned}\begin{bmatrix} 0 & -2 & 2 \\ -2 & 3 & -2 \\ 4 & -2 & 0 \end{bmatrix} &= 6\begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \end{bmatrix}\begin{bmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix} + 3\begin{bmatrix} -\frac{2}{3} \\ \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}\begin{bmatrix} \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \end{bmatrix} \\ &= \begin{bmatrix} 4/3 & -4/3 & 2/3 \\ -8/3 & 8/3 & -4/3 \\ 8/3 & -8/3 & 4/3 \end{bmatrix} + \begin{bmatrix} -4/3 & -2/3 & 4/3 \\ 2/3 & 1/3 & -2/3 \\ 4/3 & 2/3 & -4/3 \end{bmatrix}\end{aligned}$$

Egy 2-rangú mátrix két 1-rangú összegére való bontása többféleképp is lehetséges. E felbontás tulajdonságaira hamarosan visszatérünk. □

A szinguláris értékek és a szinguláris felbontás meghatározása Kérdés, hogyan számolható ki a szinguláris felbontás? Egy egyszerű ötlettel visszavezethető egy sajátértékszámítási feladatra. Tekintsük az $\mathbf{A}^T\mathbf{A}$ (komplex esetben az $\mathbf{A}^H\mathbf{A}$) mátrixot! Mivel $\mathbf{U}^T\mathbf{U} = \mathbf{I}$ és $\mathbf{\Sigma}^T\mathbf{\Sigma} = \mathbf{\Sigma}^2$, ezért

$$\mathbf{A}^T\mathbf{A} = (\mathbf{U\Sigma V}^T)^T\mathbf{U\Sigma V}^T = \mathbf{V\Sigma}^T\mathbf{U}^T\mathbf{U\Sigma V}^T = \mathbf{V\Sigma}^2\mathbf{V}^T.$$

Ez épp $\mathbf{A}^T\mathbf{A}$ spektrálfelbontása, hisz $\mathbf{V}$ ortogonális (komplex esetben unitér), és $\mathbf{\Sigma}^2$ diagonális. Mivel $\mathbf{\Sigma}^T\mathbf{\Sigma} = \mathbf{\Sigma}^2 = \operatorname{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2, 0, \dots, 0)$, ezért $\mathbf{A}^T\mathbf{A}$ sajátértékei épp a szinguláris értékek négyzetei.

9.4. PÉLDA (SZINGULÁRIS ÉRTÉKEK MEGHATÁROZÁSA). *Számítsuk ki az*

$$\mathbf{A} = \begin{bmatrix} 0 & -2 & 2 \\ -2 & 3 & -2 \\ 4 & -2 & 0 \end{bmatrix}$$

mátrix szinguláris értékeit, és írjuk fel szinguláris felbontását!

MEGOLDÁS. A szinguláris értékek megegyeznek $\mathbf{A}^T\mathbf{A}$ nemnulla sajátértékeinek gyökeivel.

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 20 & -14 & 4 \\ -14 & 17 & -10 \\ 4 & -10 & 8 \end{bmatrix}$$

Ennek karakterisztikus polinomja $x^3 - 45x^2 + 324x$, melynek gyökei 36, 9 és 0. Tehát a szinguláris értékek 6 és 3. Az $\mathbf{A}^T\mathbf{A}$ mátrix egységnyi sajátvektorai:

$$\begin{aligned} \lambda_1 &= 36 & \mathbf{v}_1 &= (2/3, -2/3, 1/3) \\ \lambda_2 &= 9 & \mathbf{v}_2 &= (2/3, 1/3, -2/3) \\ \lambda_3 &= 0 & \mathbf{v}_3 &= (1/3, 2/3, 2/3). \end{aligned}$$

Mivel $\mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i$, ezért $\mathbf{u}_i = \mathbf{A}\mathbf{v}_i/\sigma_i$. Így kiszámolható $\mathbf{u}_1$ és $\mathbf{u}_2$ is:

$$\mathbf{u}_1 = \frac{\mathbf{A}\mathbf{v}_1}{\sigma_1} = \frac{(2,-4,4)}{6} = \left(\frac{1}{3}, -\frac{2}{3}, \frac{2}{3}\right)$$
$$\mathbf{u}_2 = \frac{\mathbf{A}\mathbf{v}_2}{\sigma_2} = \frac{(-2,1,2)}{3} = \left(-\frac{2}{3}, \frac{1}{3}, \frac{2}{3}\right)$$

A harmadik vektor ilyen módon nem számolható ki, mivel $\mathbf{A}\mathbf{v}_3 = \mathbf{0}$, így az $\{\mathbf{u}_1, \mathbf{u}_2\}$ rendszert nekünk kell kiegészítenünk $\mathbb{R}^3$ bázisává. Ezt több módon is megtehetjük. Mivel $\{\mathbf{u}_1, \mathbf{u}_2\}$ az oszloptér bázisa, ezért a merőleges kiegészítő altér – vagyis $\mathbf{A}^T$ nullterének – bázisát keressük. E példában legegyszerűbb megoldás vektori szorzással számolni: $\mathbf{u}_3 = \mathbf{u}_1 \times \mathbf{u}_2 = (-2/3, -2/3, -1/3)$. A szinguláris és a redukált szinguláris felbontás tehát:

$$\mathbf{A} = \frac{1}{3}\begin{bmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \frac{1}{3}\begin{bmatrix} 2 & -2 & 1 \\ 2 & 1 & -2 \\ 1 & 2 & 2 \end{bmatrix},$$
$$\mathbf{A} = \frac{1}{3}\begin{bmatrix} 1 & -2 \\ -2 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 6 & 0 \\ 0 & 3 \end{bmatrix} \frac{1}{3}\begin{bmatrix} 2 & -2 & 1 \\ 2 & 1 & -2 \end{bmatrix}$$

A felírásban az $\mathbf{U}$ és $\mathbf{V}$ mátrixból is kiemeltünk $\frac{1}{3}$-ot, de ez is a mátrixhoz tartozik – egyébként nem lenne ortogonális. □

Az $\mathbf{U}$ mátrix meghatározására egy további módszer is adódik. Az $\mathbf{A}^T\mathbf{A}$ helyett vizsgáljuk meg az $\mathbf{A}\mathbf{A}^T$ mátrixot.

$$\mathbf{A}\mathbf{A}^T = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T(\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T)^T = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T\mathbf{V}\boldsymbol{\Sigma}^T\mathbf{U}^T = \mathbf{U}\boldsymbol{\Sigma}^2\mathbf{U}^T.$$

Eszerint a szinguláris értékek az $\mathbf{A}\mathbf{A}^T$ mátrixból is meghatározhatók. Ennek sajátvektorai az $\mathbf{U}$ mátrix első r oszlopát adják. Mivel az $\mathbf{A}^T\mathbf{A}$ mátrix σ_i^2 értékhez tartozó sajátvektora $\mathbf{v}_i$, ezért $\mathbf{A}^T\mathbf{A}\mathbf{v}_i = \sigma_i^2\mathbf{v}_i$, másrészt $\mathbf{A}\mathbf{v}_i = \sigma_i\mathbf{u}_i$, így e két összefüggést összevetve kapjuk, hogy $\mathbf{A}^T\mathbf{A}\mathbf{v}_i = \mathbf{A}^T(\sigma_i\mathbf{u}_i) = \sigma_i^2\mathbf{v}_i$, azaz

$$\mathbf{A}^T\mathbf{u}_i = \sigma_i\mathbf{v}_i, \quad \text{azaz} \quad \mathbf{v}_i = \frac{\mathbf{A}^T\mathbf{u}_i}{\sigma_i}.$$

Érdemes lehet az $\mathbf{A}\mathbf{A}^T$ pozitív sajátértékekhez tartozó sajátvektorait keresni, ha $m < n$, mert ekkor csak m-dimenziós vektorokkal kell számolni (ld. a 9.2. feladatot).

9.5. PÉLDA (SZINGULÁRIS FELBONTÁS). *Számítsuk ki az*

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}$$

mátrix szinguláris érték szerinti felbontását!

MEGOLDÁS. $\mathbf{A}^T\mathbf{A} = \left[\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right]$, melynek karakterisztikus polinomja $\lambda^2 - 4\lambda + 3 = (\lambda - 3)(\lambda - 1)$. Az $\mathbf{A}^T\mathbf{A}$ sajátértékei 3 és 1, tehát $\mathbf{A}$ szinguláris értékei $\sqrt{3}$ és 1. A hozzájuk tartozó egységnyi hosszú sajátvektorok $\mathbf{v}_1 = (1/\sqrt{2}, 1/\sqrt{2})$, $\mathbf{v}_2 = (-1/\sqrt{2}, 1/\sqrt{2})$. Így

$$\mathbf{V} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}, \qquad \mathbf{\Sigma} = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Az $\mathbf{u}_i = \mathbf{A}\mathbf{v}_i/\sigma_i$ összefüggés alapján $\mathbf{u}_1 = \frac{1}{\sqrt{6}}(1, 2, 1)$, $\mathbf{u}_2 = \frac{1}{\sqrt{2}}(1, 0, -1)$. Az előző példához hasonlóan $\mathbf{u}_3$ kiszámítható az $\mathbf{u}_3 = \mathbf{u}_1 \times \mathbf{u}_2$ képlettel is, de most inkább számoljunk úgy, hogy keressük $\mathbf{A}^T$ nullterének bázisát. A nulltér meghatározásához meg kell oldani a $\mathbf{A}^T = \left[\begin{smallmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{smallmatrix}\right]$ együtthatómátrixú homogén lineáris egyenletrendszert. Innen is az adódik, hogy $\mathbf{u}_3 = \frac{1}{\sqrt{3}}(1, -1, 1)$. Itt választhatjuk e vektor ellentettjét is, mert $\mathbf{A}\mathbf{u}_3 = \mathbf{0}$, vagyis az előjelnek itt nincs szerepe. Így

$$\mathbf{U} = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ 2/\sqrt{6} & 0 & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & -1/\sqrt{3} \end{bmatrix}.$$

Tehát a szinguláris felbontás

$$\begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ 2/\sqrt{6} & 0 & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & -1/\sqrt{3} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}.$$

□

Szinguláris érték szerinti felbontás létezése Kérdés még, hogy bármely valós vagy komplex mátrixnak léteznek-e szinguláris értékei, azok egyértelműek-e, és ha igen, a szinguláris felbontás egyértelmű-e.

9.6. TÉTEL (A SZINGULÁRIS ÉRTÉKEK TULAJDONSÁGAI). *Minden r-rangú valós vagy komplex $\mathbf{A}$ mátrixnak létezik r szinguláris értéke. Ezek valós esetben megegyeznek az $\mathbf{A}^T\mathbf{A}$, illetve az $\mathbf{A}\mathbf{A}^T$ (komplex esetben az $\mathbf{A}^H\mathbf{A}$, illetve az $\mathbf{A}\mathbf{A}^H$) pozitív sajátértékeinek négyzetgyökével. A szinguláris értékek monoton csökkenő sorozata egyértelmű.*

BIZONYÍTÁS. A bizonyítást valós esetre írjuk le, komplexre lényegében azonos. Az $\mathbf{A}^T\mathbf{A}$ mátrix szimmetrikus ($\mathbf{A}^H\mathbf{A}$ önadjungált), mert $(\mathbf{A}^T\mathbf{A})^T = \mathbf{A}^T(\mathbf{A}^T)^T = \mathbf{A}^T\mathbf{A}$. Ennek következtében minden sajátértéke valós, és ortogonálisan diagonalizálható. Másrészt minden sajátértéke nemnegatív, ugyanis

$$|\mathbf{A}\mathbf{x}|^2 = (\mathbf{A}\mathbf{x})^T(\mathbf{A}\mathbf{x}) = \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{x}^T(\mathbf{A}^T\mathbf{A}\mathbf{x}) = \mathbf{x}^T(\lambda\mathbf{x}) = \lambda|\mathbf{x}|^2.$$

(Másként fogalmazva $\mathbf{A}^T\mathbf{A}$ pozitív szemidefinit.) A 0-tól különböző sajátértékek száma megegyezik $\mathbf{A}^T\mathbf{A}$ rangjával, hisz az megegyezik diagonális alakja nemnulla elemeinek számával. Másrészt ???? szerint $\mathrm{r}(\mathbf{A}^T\mathbf{A}) = \mathrm{r}(\mathbf{A}) = r$. Tehát, ha nagyság szerinti sorba rendezzük a sajátértékeket ($\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$), akkor $\lambda_i > 0$, ha $1 \leq i \leq r$, és $\lambda_i = 0$, ha $r < i \leq n$. Eszerint $\sigma_i = \sqrt{\lambda_i} > 0$, ha $1 \leq i \leq r$. Végül, mivel $\mathbf{A}^T\mathbf{A}$ sajátértékei egyértelműek, ezért $\mathbf{A}$ szinguláris értékei is. Azt, hogy $\mathbf{A}^T\mathbf{A}$ és $\mathbf{A}\mathbf{A}^T$ 0-tól különböző sajátértékei megegyeznek, korábban beláttuk. $\square$

Bal és jobb szinguláris vektorok

Szimmetrikus és önadjungált mátrixok szinguláris felbontása

Polárfelbontás

Pszeudoinverz

Információtömörítés

Feladatok

9.1. [Szinguláris felbontások] Az $\mathbf{A} = \left[\begin{smallmatrix}1&1&0\\1&1&0\end{smallmatrix}\right]$ mátrixnak egyetlen szinguláris értéke van, $\sigma_1 = 2$. Igazoljuk, hogy az

$$\begin{bmatrix}1&1&0\\1&1&0\end{bmatrix} = \left[\begin{array}{c|c}1/\sqrt{2}&-1/\sqrt{2}\\1/\sqrt{2}&1/\sqrt{2}\end{array}\right]\left[\begin{array}{c|cc}2&0&0\\\hline 0&0&0\end{array}\right]\left[\begin{array}{ccc}1/\sqrt{2}&1/\sqrt{2}&0\\\hline -1/\sqrt{2}&1/\sqrt{2}&0\\\hline 0&0&1\end{array}\right]$$

$$\begin{bmatrix}1&1&0\\1&1&0\end{bmatrix} = \begin{bmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}[2]\begin{bmatrix}1/\sqrt{2}&1/\sqrt{2}&0\end{bmatrix}$$

felbontások az **A** mátrix szinguláris és redukált szinguláris felbontásai. (Segítségül a szinguláris felbontásban a blokkstruktúrát is jelöltük.)

9.2. Számítsuk ki a

$$\mathbf{B} = \begin{bmatrix}1&0&1\\1&1&0\end{bmatrix}$$

mátrix szinguláris érték szerinti felbontását!

10

Jordan-féle normálalak

Mind bizonyos problémák megértésében, mind a gyakorlati alkalmazásokban fontos lehet, hogy egy lineáris leképezés mátrixát milyen bázisban írjuk fel. Láttuk milyen sokat segít, ha a bázis sajátvektorokból áll, sőt, ha még ortonormált is. Ekkor a mátrix diagonális alakot ölt. Kérdés azonban, milyen mátrixok diagonalizálhatók, és amelyek nem, azokkal mit tudunk kezdeni.

Schur-felbontás

Általánosított sajátvektorok és a Jordan-blokk Tekintsük a 8.15.. példában szereplő

$$\mathbf{A} = \begin{bmatrix} 4 & 1 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{bmatrix}$$

mátrix hatását a standard bázis vektorain:

$$\begin{aligned} \mathbf{A}\mathbf{e}_1 &= 4\mathbf{e}_1 \\ \mathbf{A}\mathbf{e}_2 &= \mathbf{e}_1 + 4\mathbf{e}_2 \\ \mathbf{A}\mathbf{e}_3 &= \mathbf{e}_2 + 4\mathbf{e}_3 \end{aligned}$$

Átrendezés után kapjuk, hogy:

$$\begin{aligned} (\mathbf{A} - 4\mathbf{I})\mathbf{e}_1 &= \mathbf{0} \\ (\mathbf{A} - 4\mathbf{I})\mathbf{e}_2 &= \mathbf{e}_1 \\ (\mathbf{A} - 4\mathbf{I})\mathbf{e}_3 &= \mathbf{e}_2 \end{aligned}$$

Az $\mathbf{A} - 4\mathbf{I}$ mátrix fenti hatását a következő diagrammal fogjuk szemléltetni:

$$\mathbf{0} \xleftarrow{\mathbf{A}-4\mathbf{I}} \mathbf{e}_1 \xleftarrow{\mathbf{A}-4\mathbf{I}} \mathbf{e}_2 \xleftarrow{\mathbf{A}-4\mathbf{I}} \mathbf{e}_3$$

Eszerint $\mathbf{e}_1$ sajátvektor, és mint a 8.15.. példában láttuk, $\mathbf{A}$-nak más sajátvektora nincs. Viszont az előző egyenletekből a következő összefüggés adódik:

$$(\mathbf{A}-4\mathbf{I})^2\mathbf{e}_2 = \mathbf{0}$$
$$(\mathbf{A}-4\mathbf{I})^3\mathbf{e}_3 = \mathbf{0}.$$

Az ilyen vektorokra külön elnevezést fogunk alkalmazni:

10.1. DEFINÍCIÓ (ÁLTALÁNOSÍTOTT SAJÁTVEKTOR). *Az $\mathbf{x} \neq \mathbf{0}$ vektort a négyzetes $\mathbf{A}$ mátrix λ sajátértékéhez tartozó* általánosított sajátvektorának *nevezzük, ha valamilyen k természetes számra $(\mathbf{A}-\lambda\mathbf{I})^k\mathbf{x} = \mathbf{0}$. $k = 1$ esetén $\mathbf{x}$ sajátvektor. Az általánosított sajátvektorokból álló $\mathbf{x}_i$ ($i = 1, 2, \ldots, k$) sorozatot* Jordan-láncnak *nevezzük, ha $(\mathbf{A}-\lambda\mathbf{I})\mathbf{x}_i = \mathbf{x}_{i-1}$ és $(\mathbf{A}-\lambda\mathbf{I})\mathbf{x}_1 = \mathbf{0}$.*

Az előzőekben vizsgált konkrét példában tehát a térnek sajátvektorokból álló bázisa nincs, hisz a sajátaltér csak 1-dimenziós, de általánosított sajátvektorokból álló bázisa van.

10.2. PÉLDA (JORDAN-LÁNC KONSTRUKCIÓJA). *Keressünk az alábbi mátrixok minden sajátalterének minden bázisvektorához egy abban végződő Jordan-láncot:*

$$\mathbf{A} = \begin{bmatrix} 6 & -1 & -3 \\ -1 & 5 & 2 \\ 2 & -1 & 1 \end{bmatrix} \text{ és a } \mathbf{B} = \begin{bmatrix} 3 & 0 & 1 \\ 0 & 4 & 0 \\ -1 & 0 & 5 \end{bmatrix}$$

MEGOLDÁS. A karakterisztikus polinom mindkét mátrix esetén $-\lambda^3 + 12\lambda^2 - 48\lambda + 64 = (4-\lambda)^3$, ezért $\lambda = 4$ háromszoros algebrai multiplicitású sajátérték.

Az $\mathbf{A}$ mátrix esetén a sajátaltér 1-dimenziós, melyet az $\mathbf{x}_1 = (1, -1, 1)$ vektor feszít ki.

Mivel $(\mathbf{A}-4\mathbf{I})^3 = \mathbf{O}$, ezért legföljebb három hosszú sorozatra számíthatunk: olyan $\mathbf{x}_2$ és $\mathbf{x}_3$ vektort keresünk, melyre $(\mathbf{A}-4\mathbf{I})\mathbf{x}_2 = \mathbf{x}_1$ és $(\mathbf{A}-4\mathbf{I})\mathbf{x}_3 = \mathbf{x}_2$. Ez két egyenletrendszer megoldását jelenti. Elég a megoldások közül egyet kiválasztani:

$$\mathbf{A}-4\mathbf{I} = \begin{bmatrix} 2 & -1 & -3 \\ -1 & 1 & 2 \\ 2 & -1 & -3 \end{bmatrix}, \text{ egy-egy megoldás: } \mathbf{x}_2 = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} -1 \\ -2 \\ 0 \end{bmatrix}.$$

Ezek a következő láncot adják:

$$\mathbf{0} \xleftarrow{\mathbf{A}-4\mathbf{I}} \mathbf{x}_1 = (1, -1, 1) \xleftarrow{\mathbf{A}-4\mathbf{I}} \mathbf{x}_2 = (0, -1, 0) \xleftarrow{\mathbf{A}-4\mathbf{I}} \mathbf{x}_3 = (-1, -2, 0)$$

A $\mathbf{B}$ mátrix esetén a sajátaltér 2-dimenziós, melyet az $\mathbf{x}_1 = (1, 0, 1)$ és az $\mathbf{y}_1(0, 1, 0)$ vektorok feszítenek ki.

Mivel $(\mathbf{B}-4\mathbf{I})^2 = \mathbf{O}$, ezért legföljebb kettő hosszú láncra számíthatunk: olyan $\mathbf{x}_2$ és $\mathbf{y}_2$ vektort keresünk, melyre $(\mathbf{B}-4\mathbf{I})\mathbf{x}_2 = \mathbf{x}_1$ és

$(\mathbf{B}-4\mathbf{I})\mathbf{y}_2 = \mathbf{y}_1$. Ezek közül az első megoldható, és megoldásainak egyike $(-1,0,0)$, míg a második nem oldható meg. Ez két láncra vezet:

$$\mathbf{0} \xleftarrow{\mathbf{B}-4\mathbf{I}} \mathbf{x}_1 = (1,0,1) \xleftarrow{\mathbf{B}-4\mathbf{I}} \mathbf{x}_2 = (-1,0,0)$$
$$\mathbf{0} \xleftarrow{\mathbf{B}-4\mathbf{I}} \mathbf{y}_1 = (0,1,0)$$

Az egyik lánc kettő hosszú, a másik egy hosszú. □

A bevezetőben mutatott mátrix-típus fontos szerepet játszik a továbbiakban:

10.3. DEFINÍCIÓ (JORDAN-BLOKK). *Azt a négyzetes mátrixot, melynek főátlójában azonos λ értékek, fölötte 1-esek, egyebütt 0-k állnak, azaz melynek alakja*

$$\mathbf{J}_\lambda = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{bmatrix} \tag{10.1}$$

Jordan-blokknak *nevezzük.*

A fentiek mintájára könnyen látható, hogy az ilyen mátrixoknak a standard bázis minden vektora általánosított sajátvektora, tehát kimondható a következő megállapítás: minden Jordan-blokknak van általánosított sajátvektorokból álló bázisa.

10.4. PÉLDA (JORDAN-LÁNCHOZ TARTOZÓ JORDAN-BLOKK). *Írjuk fel a 10.2.. példában szereplő $\mathbf{A}$ mátrixot az ott konstruált Jordan-lánc vektoraiból álló bázisban!*

MEGOLDÁS. Mivel $\mathbf{A}\mathbf{x}_3 = 4\mathbf{x}_3 + \mathbf{x}_2$, $\mathbf{A}\mathbf{x}_2 = 4\mathbf{x}_2 + \mathbf{x}_1$ és $\mathbf{A}\mathbf{x}_1 = 4\mathbf{x}_1$, ezért az $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ bázisban a leképezés mátrixa

$$\mathbf{J} = \begin{bmatrix} 4 & 1 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{bmatrix}$$

Ezt az eredményt az áttérés $\mathbf{C} = [\mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_3]$ mátrixával való közvetlen számolással is megkaphatjuk, ha elvégezzük az alábbi mátrixszorzásokat:

$$\mathbf{J} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C} = \begin{bmatrix} -2 & -3 & 0 \\ -1 & -1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 6 & -1 & -3 \\ -1 & 5 & 2 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -3 & 0 \\ -1 & 2 & 0 \\ 1 & -3 & 1 \end{bmatrix}$$ □

10.5. PÉLDA (JORDAN-LÁNCOK ÉS JORDAN-BLOKKOK KAPCSOLATA). *Tudjuk, hogy az* $\mathbf{A}$ *mátrixnak két különböző sajátértéke van,* $\lambda_1 = 2$ *és* $\lambda_2 = 4$*, valamint hogy a* $\mathbf{C}$ *mátrix oszlopvektorai az* $\mathbf{A}$ *egy Jordan-bázisát alkotják, ahol*

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 1 & -1 & 0 & 1 & -1 \\ 0 & 2 & 1 & -1 & 0 & 1 & -1 \\ 2 & -2 & 2 & 2 & 0 & -1 & 1 \\ 1 & -1 & 0 & 3 & 1 & 1 & -1 \\ -1 & 1 & 0 & -1 & 5 & 1 & -1 \\ -1 & 1 & 0 & -1 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}, \mathbf{C} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

Rajzoljuk fel a Jordan-láncok diagrammját, és határozzuk meg a $\mathbf{J} = \mathbf{C}^{-1}\mathbf{AC}$ *mátrixot a szorzások elvégzése nélkül!*

MEGOLDÁS. A $\mathbf{C}$ oszlopai Jordan-bázist alkotnak, azaz minden oszlopvektor egy Jordan-lánc eleme. Mivel az $\mathbf{A}$ mátrixnak csak két különböző sajátértéke van, ez azt jelenti, hogy minden oszlopvektort vagy az $\mathbf{A} - 2\mathbf{I}$ vagy az $\mathbf{A} - 4\mathbf{I}$ mátrix vagy a zérusvektorba, vagy egy másik oszlopvektorba visz (előbbi esetben az oszlopvektor sajátvektor, utóbbi esetben csak általánosított sajátvektor). E hatást az $(\mathbf{A} - 2\mathbf{I})\mathbf{C}$ és az $(\mathbf{A} - 4\mathbf{I})\mathbf{C}$ szorzatok kiszámításával könnyen megkaphatjuk:

$$\mathbf{A} - 2\mathbf{I} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 3 & 1 & 0 \\ 0 & 0 & 0 & 2 & 3 & 3 & 0 \\ 0 & 0 & 0 & 0 & 2 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 & 2 & 2 \end{bmatrix} \mathbf{A} - 4\mathbf{I} = \begin{bmatrix} -2 & -1 & 0 & 0 & 0 & 0 & 0 \\ -2 & -1 & -2 & 0 & 0 & 0 & 0 \\ 0 & -2 & -2 & 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Az első szorzatból látszik, hogy $(\mathbf{A} - 2\mathbf{I})[\mathbf{c}_1\ \mathbf{c}_2\ \mathbf{c}_3] = [\mathbf{0}\ \mathbf{c}_1\ \mathbf{0}]$, míg a másodikból, hogy $(\mathbf{A} - 4\mathbf{I})[\mathbf{c}_4\ \mathbf{c}_5\ \mathbf{c}_6\ \mathbf{c}_7] = [\mathbf{0}\ \mathbf{c}_4\ \mathbf{c}_5\ \mathbf{0}]$ (a második szorzatban már nem is kellett volna a $\mathbf{c}_1$, $\mathbf{c}_2$, $\mathbf{c}_3$ vektorokkal szorozni látva az előző szorzás eredményét). Ebből fölrajzolható a diagram:

$$\begin{array}{llll}
\mathbf{0} \xleftarrow{\mathbf{A}-2\mathbf{I}} & \mathbf{c}_1 \xleftarrow{\mathbf{A}-2\mathbf{I}} & \mathbf{c}_2 & \\
\mathbf{0} \xleftarrow{\mathbf{A}-2\mathbf{I}} & \mathbf{c}_3 & & \\
\mathbf{0} \xleftarrow{\mathbf{A}-4\mathbf{I}} & \mathbf{c}_4 \xleftarrow{\mathbf{A}-4\mathbf{I}} & \mathbf{c}_5 \xleftarrow{\mathbf{A}-4\mathbf{I}} & \mathbf{c}_6 \\
\mathbf{0} \xleftarrow{\mathbf{A}-4\mathbf{I}} & \mathbf{c}_7 & &
\end{array}$$

A diagramból kiolvasható az $\mathbf{A}$ hatása a $\mathbf{c}_i$ vektorokra: $\mathbf{Ac}_1 = 2\mathbf{c}_1$, $\mathbf{Ac}_2 = 2\mathbf{c}_2 + \mathbf{c}_1$, $\mathbf{Ac}_3 = 2\mathbf{c}_3$, $\mathbf{Ac}_4 = 4\mathbf{c}_4$, $\mathbf{Ac}_5 = 4\mathbf{c}_5 + \mathbf{c}_4$, $\mathbf{Ac}_6 = 4\mathbf{c}_6 + \mathbf{c}_5$,

$\mathbf{A}\mathbf{c}_7 = 4\mathbf{c}_7$. Ebből felírható e leképezés mátrixa:

$$\mathbf{J} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C} = \left[\begin{array}{cc|c|ccc|c} 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 4 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 4 \end{array}\right]$$ □

Jordan normálalak Nem hozható minden mátrix diagonális alakra, de egy ahhoz közeli alakra – ahogy azt a 10.5.. példa mutatja – igen. Ebben csak közvetlenül a főátló fölött lehetnek nemnulla elemek, és azok is csak 1-esek. Ráadásul az új bázis ismeretében ez az alak jól leírja a mátrixleképezés hatását is.

10.6. tétel (Jordan normálalak). *Bármely $\mathbb{C}^{n\times n}$-beli mátrix hasonló egy Jordan-blokkokból álló blokkdiagonális mátrixhoz. Részletesen megfogalmazva: minden $\mathbf{A} \in \mathbb{C}^{n\times n}$ mátrixhoz létezik olyan $\mathbf{C}$ mátrix, hogy a $\mathbf{J} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$ mátrix alakja*

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \mathbf{J}_2 & \dots & \mathbf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \dots & \mathbf{J}_k \end{bmatrix} \tag{10.2}$$

ahol k az $\mathbf{A}$ független sajátvektorainak maximális száma, és $\mathbf{J}_i$ az i-edik sajátvektorhoz tartozó Jordan-blokk.

► A tételbeli (10.2) mátrixot az $\mathbf{A}$ mátrix *Jordan-féle normálalakjának* nevezzük.

► A különböző Jordan-blokkok különböző sajátvektorokhoz tartoznak, de mivel több sajátvektor is tartozhat ugyanahhoz a sajátértékhez, ezért egy sajátérték több Jordan-blokkban is szerepelhet.

Bizonyítás. Megmutatjuk, hogy ha találunk k független sajátvektort, akkor találunk k Jordan-láncot is, melyek vektorai a tér bázisát adják, és ebben a bázisban a leképezés mátrixa a tételbeli Jordan-alakot ölti.

A mátrix méretére vonatkozó teljes indukcióval bizonyítunk. $n = 1$ esetén az állítás nyilván igaz. Tegyük fel, hogy igaz az állítás minden n-nél kisebb méretű mátrixra.

Legyen λ az $n \times n$-es $\mathbf{A}$ mátrix egy sajátértéke, és legyen $\mathbf{x}$ egy hozzá tartozó sajátvektor. Az $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ megoldásainak terét, vagyis a λ-hoz tartozó sajátalteret jelölje $\mathcal{N}_\lambda$. Ennek dimenzióját, vagyis $(\mathbf{A} - \lambda\mathbf{I})$ nullitását jelölje r.

Mivel $r > 0$, ezért a dimenziótétel miatt $(\mathbf{A} - \lambda\mathbf{I})$ oszlopterének dimenziója $n - r < n$. Jelölje az oszlopteret $\mathcal{O}_\lambda$. Az $\mathcal{O}_\lambda$ invariáns altere $\mathbf{A}$-nak, azaz $\mathbf{A}(\mathcal{O}_\lambda) \subseteq \mathcal{O}_\lambda$, ugyanis $\mathcal{O}_\lambda$ elemei $(\mathbf{A} - \lambda\mathbf{I})\mathbf{v}$ alakúak, ahol $\mathbf{v}$ tetszőleges, és $\mathbf{A}(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = (\mathbf{A}^2 - \lambda\mathbf{A})\mathbf{v} = (\mathbf{A} - \lambda\mathbf{I})(\mathbf{A}\mathbf{v})$ ugyancsak eleme $\mathcal{O}_\lambda$-nak. Az $\mathbf{A}$ mátrixleképezést szűkítsük le $\mathcal{O}_\lambda$-ra, jelölje ennek mátrixát $\hat{\mathbf{A}}$.

Az $\hat{\mathbf{A}}$ mátrix $n - r$ méretű, ezért az indukciós feltevés szerint van Jordan-láncokból álló bázisa. A következő diagram e láncokat szemlélteti, és azt, hogy $\hat{\mathbf{A}} - \lambda\mathbf{I}$ hogy hat rajtuk. Mivel azonban e halmazon ez megegyezik $\mathbf{A} - \lambda\mathbf{I}$ hatásával, a nyilak fölé is ezt írjuk:

$$
\begin{array}{ccccccc}
\mathbf{0} & \xleftarrow{\mathbf{A}-\lambda_1\mathbf{I}} & \mathbf{x}_1^1 & \xleftarrow{\mathbf{A}-\lambda_1\mathbf{I}} & \dots & \xleftarrow{\mathbf{A}-\lambda_1\mathbf{I}} & \mathbf{x}_{s_1}^1 \\
\mathbf{0} & \xleftarrow{\mathbf{A}-\lambda_2\mathbf{I}} & \mathbf{x}_1^2 & \xleftarrow{\mathbf{A}-\lambda_2\mathbf{I}} & \dots & \xleftarrow{\mathbf{A}-\lambda_2\mathbf{I}} & \mathbf{x}_{s_2}^2 \\
\vdots & & \vdots & & \vdots & & \vdots \\
\mathbf{0} & \xleftarrow{\mathbf{A}-\lambda_p\mathbf{I}} & \mathbf{x}_1^p & \xleftarrow{\mathbf{A}-\lambda_p\mathbf{I}} & \dots & \xleftarrow{\mathbf{A}-\lambda_p\mathbf{I}} & \mathbf{x}_{s_p}^p
\end{array}
$$

Itt az $\mathbf{x}_k^j$ vektor felső indexe a lánc sorszámát jelöli.

Az $\mathbf{A} - \lambda\mathbf{I}$ oszloptere és magtere metszetének dimenzióját jelölje q, azaz legyen $q = \dim(\mathcal{O}_\lambda \cap \mathcal{N}_\lambda)$. Mivel $\mathcal{N}_\lambda$ elemei az $\mathbf{A}$ sajátvektorai, azok $\mathcal{O}_\lambda$-ba eső báziselemei pedig az $\mathbf{x}_1^j$ vektorok, ezért q közülük a λ sajátvektora, azaz q sajátérték megegyezik λ-val, pl. az első q, azaz $\lambda_k = \lambda$ $(k = 1, 2, \dots, q)$. A láncok végén lévő $\mathbf{x}_{s_k}^k$ vektorok elemei $\mathcal{O}_\lambda$-nak, tehát mindegyikhez van olyan $\mathbf{y}_k$ vektor, hogy $(\mathbf{A} - \lambda\mathbf{I})\mathbf{y}_k = \mathbf{x}_{s_k}^k$ $(k = 1, 2, \dots, q)$.

Az $\mathcal{N}_\lambda$ altér r-dimenziós, annak q-dimenziós alteréhez találunk olyan $\mathcal{Z}$ alteret, mely $(r - q)$-dimenziós, és minden vektora sajátvektor. Így a Jordan-láncok így alakulnak:

$$
\begin{array}{ccccccccc}
\mathbf{0} & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_1^1 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \dots & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_{s_1}^1 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{y}_1 \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
\mathbf{0} & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_1^q & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \dots & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_{s_q}^q & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{y}_q \\
\mathbf{0} & \xleftarrow{\mathbf{A}-\lambda_{q+1}\mathbf{I}} & \mathbf{x}_1^{q+1} & \xleftarrow{\mathbf{A}-\lambda_{q+1}\mathbf{I}} & \dots & \xleftarrow{\mathbf{A}-\lambda_{q+1}\mathbf{I}} & \mathbf{x}_{s_{q+1}}^{q+1} & & \\
\vdots & & \vdots & & \vdots & & \vdots & & \\
\mathbf{0} & \xleftarrow{\mathbf{A}-\lambda_p\mathbf{I}} & \mathbf{x}_1^p & \xleftarrow{\mathbf{A}-\lambda_p\mathbf{I}} & \dots & \xleftarrow{\mathbf{A}-\lambda_p\mathbf{I}} & \mathbf{x}_{s_p}^p & & \\
\mathbf{0} & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{z}_1 & & & & & & \\
\vdots & & \vdots & & & & & & \\
\mathbf{0} & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{z}_{r-q} & & & & & &
\end{array}
\tag{10.3}
$$

Az $\mathbf{x}$-vektorok száma $n - r$, az $\mathbf{y}$-vektorok száma q, a $\mathbf{z}$-vektorok száma $r - q$, ezek összege pedig $(n - r) + q + (r - q) = n$, tehát van elég vektor egy bázishoz. Már csak a függetlenségüket kell belátni. A konstrukció olyan volt, hogy az $\mathbf{x}$- és $\mathbf{z}$-vektorok mind függetlenek egymástól, csak az $\mathbf{y}$-vektorok tőlük való függetlenségét kell bizonyítani. Tegyük fel, hogy

$$\sum_{j,k} \xi_{j,k} \mathbf{x}_k^j + \sum_t \eta_t \mathbf{y}_t + \sum_r \zeta_r \mathbf{z}_r = \mathbf{0},$$

ahol nem minden η_t nulla. Szorozzuk meg az egyenlőséget (balról) az $\mathbf{A} - \lambda\mathbf{I}$ mátrixszal. Mivel az első q lánc végén lévő $\mathbf{x}_{s_t}^t$-vektorok kisebb indexűekbe mennek, másrészt $(\mathbf{A} - \lambda\mathbf{I})\mathbf{y}_t = \mathbf{x}_{s_t}^t$, ezért olyan összefüggéshez jutunk, amelyben már csak $\mathbf{x}$-vektorok lesznek, de nem minden együttható nulla, hisz van nemnulla η együttható, és ez ellentmondás. □

10.7. PÉLDA (NORMÁLALAKOK). *Soroljuk fel az összes lehetséges Jordan normálalakját annak a mátrixnak, melyről csak annyit tudunk, hogy* $(1-\lambda)^4$ *a karakterisztikus polinomja. Ne tekintsünk különbözőnek két normálalkot, ha azok csak a Jordan-blokkok sorrendjében különböznek egymástól!*

MEGOLDÁS. A karakterisztikus polinom negyedfokú, így a mátrix 4×4-es. Mivel minden sajátérték 1, ezért a Jordan-alak főátlójában csupa 1-es szerepel. A lehetséges öt alak elemi leszámlálással megkapható:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$
$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Nem tekintjük különbözőnek a blokkok cseréjével egymásból megkapható alakokat. Például az

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

mátrix a negyedik alakból a két blokk cseréjével megkapható. □

A Jordan-alak egyértelműsége A $\mathbb{C}^{n\times n}$-beli mátrixok a Jordan normálalakjuk szerint osztályokba sorolhatók. Egy osztályba azok a mátrixok kerülnek, melyek normálalakja azonos. Egy ilyen osztályozás azonban csak akkor létezik, ha minden mátrixnak csak egyetlen normálalakja van, azaz a normálalak egyértelmű. Ezt biztosítja a következő tétel.

10.8. TÉTEL (A JORDAN-ALAK EGYÉRTELMŰSÉGE). *Egy mátrix Jordan normálalakja a Jordan-blokkok sorrendjétől eltekintve egyértelmű.*

BIZONYÍTÁS. A felbontás egyértelműségének bizonyításához elég belátni, hogy bármely két hasonló mátrix Jordan-alakjának meghatárázó adatai a hasonlóságra nézve invariánsak.

A Jordan-blokkok, és így a Jordan-láncok száma megegyezik a független sajátvektorok maximális számával – ez invariáns.

Az egyszerűség kedvéért először tegyük fel, hogy $\mathbf{A}$ minden sajátértéke azonos, jelölje λ. A továbbiak könnyebb megértésére lássunk egy konkrét példát. Legyen az $\mathbf{A}$ mátrix karakterisztikus polinomja $(\lambda - x)^{13}$, és tegyük fel, hogy Jordan-bázisa a következőképp néz ki:

$$\begin{array}{ccccccccc}
\mathbf{0} & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_1^1 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_2^1 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_3^1 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_4^1 \\
\mathbf{0} & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_1^2 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_2^2 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_3^2 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_4^2 \\
\mathbf{0} & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_1^3 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_2^3 & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_3^3 & & \\
\mathbf{0} & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_1^4 & & & & & & \\
\mathbf{0} & \xleftarrow{\mathbf{A}-\lambda\mathbf{I}} & \mathbf{x}_1^5 & & & & & &
\end{array} \tag{10.4}$$

A leghosszabb blokk mérete 4, ami $\mathbf{A} - \lambda\mathbf{I}$ ismeretében úgy kapható meg, hogy 4 az a legkisebb s kitevő, melyre $(\mathbf{A} - \lambda\mathbf{I})^s = \mathbf{O}$. Általában is igaz, a legnagyobb blokk mérete az a legkisebb s, melyre $(\mathbf{A} - \lambda\mathbf{I})^s = \mathbf{O}$. A mátrix hatványának zérus volta is invariáns, így hasonló mátrixokra a leghosszab lánc hossza is azonos.

Legyen a λ sajátértékhez tartozó i-hosszú Jordan-láncok száma m_i. A 10.4. diagramon $m_1 = 2$, $m_2 = 0$, $m_3 = 1$, $m_4 = 2$. Látható, hogy $(\mathbf{A} - \lambda\mathbf{I})$ hatványainak rangjából megmondható, hogy vektor nem futott még a nullvektorba, innen pedig az m_i értékek is kiszámolhatók. Esetünkben

$$\begin{aligned}
m_4 &= \mathrm{r}\left((\mathbf{A} - \lambda\mathbf{I})^3\right) = 2 \\
m_3 + 2m_4 &= \mathrm{r}\left((\mathbf{A} - \lambda\mathbf{I})^2\right) = 5 \\
m_2 + 2m_3 + 3m_4 &= \mathrm{r}\left(\mathbf{A} - \lambda\mathbf{I}\right) = 8 \\
m_1 + m_2 + m_3 + m_4 &= n - \mathrm{r}(\mathbf{A} - \lambda\mathbf{I}) = 13 - 8 = 5
\end{aligned}$$

és ez az egyenletrendszer egyértelműen megoldható. Általában

$$\begin{aligned}
m_s &= \mathrm{r}\left((\mathbf{A}-\lambda\mathbf{I})^{s-1}\right)\\
m_{s-1}+2m_s &= \mathrm{r}\left((\mathbf{A}-\lambda\mathbf{I})^{s-2}\right)\\
m_{s-2}+2m_{s-1}+3m_s &= \mathrm{r}\left((\mathbf{A}-\lambda\mathbf{I})^{s-2}\right)\\
&\vdots\\
m_2+2m_3+\cdots+(s-1)m_s &= \mathrm{r}(\mathbf{A}-\lambda\mathbf{I})\\
m_1+m_2+m_3+\cdots+m_{s-1}+m_s &= n-\mathrm{r}(\mathbf{A}-\lambda\mathbf{I})
\end{aligned}$$

Ennek az egyenletrendszernek a jobb oldalán a hasonlóságra nézve invariáns értékek vannak, az együtthatómátrix háromszög alakú, így egyszerű visszahelyettesítéssel megoldható, főátlójában csupa egyes van, melynek következtében a megoldás egyértelmű, másrészt a megoldások mindegyike egész szám.

Ha a mátrixnak nem csak egy sajátértéke van, akkor sajátértékenként egy ilyen egyenletrendszert kapunk, mely annyiban változik az előzőhöz képest, hogy ha λ multiplicitása $m(\lambda)$, akkor $A-\lambda\mathbf{I}$ hatványainak rangjához az összes λ-tól különböző sor eggyel hozzájárul, így az egyenletrendszerek jobb oldalából $(n-m(\lambda)$-t ki kell vonni. A legnagyobb blokk mérete ekkor $\mathbf{A}-\lambda\mathbf{I}$-nek az a legkisebb s hatványa, melynek rangja először éri el $n-m(\lambda)$-t. Így az összes esetre általánosan érvényes egyenletrendszer:

$$\begin{aligned}
m_s &= m(\lambda)-n+\mathrm{r}\left((\mathbf{A}-\lambda\mathbf{I})^{s-1}\right)\\
m_{s-1}+2m_s &= m(\lambda)-n+\mathrm{r}\left((\mathbf{A}-\lambda\mathbf{I})^{s-2}\right)\\
m_{s-2}+2m_{s-1}+3m_s &= m(\lambda)-n+\mathrm{r}\left((\mathbf{A}-\lambda\mathbf{I})^{s-2}\right)\\
&\vdots\\
m_2+2m_3+\cdots+(s-1)m_s &= m(\lambda)-n+\mathrm{r}(\mathbf{A}-\lambda\mathbf{I})\\
m_1+m_2+m_3+\cdots+m_{s-1}+m_s &= n-\mathrm{r}(\mathbf{A}-\lambda\mathbf{I})
\end{aligned}$$

Ennek egyértelmű megoldhatósága bizonyítja állításunkat. □

10.9. PÉLDA (JORDAN-BLOKKOK MÉRETE). *Egy* 10×10*-es* $\mathbf{A}$ *mátrixnak* λ 10*-szeres algebrai multiplicitású sajátértéke.* $\mathbf{A}-\lambda\mathbf{I}$ *hatványainak rangja rendre* $5, 2, 1, 0$. *Írjuk fel a Jordan normálalakját!*

MEGOLDÁS. A blokkok száma, ami megegyezik a Jordan-láncok számával 5, mivel $n-\mathrm{r}(\mathbf{A}) = 10-5 = 5$. A leghosszabb lánc hossza 4, mivel $\mathbf{A}-\lambda\mathbf{I}$ legkisebb zérusmátrixot adó hatványa a 4-dik. Az egyenletrendszer és megoldása, valamint a $\mathbf{J}$ Jordan-mátrix:

$$\begin{aligned} m_4 &= 1 \\ m_3 + 2m_4 &= 2 \\ m_2 + 2m_3 + 3m_4 &= 5 \\ m_1 + m_2 + m_3 + m_4 &= 5 \end{aligned} \quad \Rightarrow \quad \begin{aligned} m_4 &= 1 \\ m_3 &= 0 \\ m_2 &= 2 \\ m_1 &= 2 \end{aligned} \quad \Rightarrow \quad \mathbf{J} = \begin{bmatrix} \lambda & 1 &&&&&&&& \\ & \lambda & 1 &&&&&&& \\ && \lambda & 1 &&&&&& \\ &&& \lambda &&&&&& \\ &&&& \lambda & 1 &&&& \\ &&&&& \lambda &&&& \\ &&&&&& \lambda & 1 && \\ &&&&&&& \lambda && \\ &&&&&&&& \lambda & \\ &&&&&&&&& \lambda \end{bmatrix}$$

□

10.10. PÉLDA (JORDAN-BLOKKOK MÉRETE). *Egy* 14×14*-es* $\mathbf{A}$ *mátrix karakterisztikus polinomja* $(3-\lambda)^5(2-\lambda)^5(1-\lambda)^4$.

$\mathbf{A} - 3\mathbf{I}$ *hatványainak rangja rendre:*	*12, 11, 10, 9;*
$\mathbf{A} - 2\mathbf{I}$ *hatványainak rangja rendre:*	*12, 10, 9;*
$\mathbf{A} - \mathbf{I}$ *hatványainak rangja rendre:*	*11, 10.*

Írjuk fel a Jordan normálalakját!

MEGOLDÁS. Most $n = 14$, $m(3) = 5$, $m(2) = 5$, $m(1) = 4$.

$\lambda = 3$ esetén a blokkok (Jordan-láncok) száma $n - \mathrm{r}(\mathbf{A} - 3\mathbf{I}) = 14 - 12 = 2$. A leghosszabb lánc hossza $s = 4$, ugyanis ez a legkisebb s, melyre $\mathrm{r}((\mathbf{A} - 3\mathbf{I})^s) = 14 - 5 = 9$. Az egyenletrendszer és megoldása

$$\begin{aligned} m_4 &= 5 - 14 + 10 = 1 \\ m_3 + 2m_4 &= 5 - 14 + 11 = 2 \\ m_2 + 2m_3 + 3m_4 &= 5 - 14 + 12 = 3 \\ m_1 + m_2 + m_3 + m_4 &= 14 - 12 = 2 \end{aligned} \quad \Rightarrow \quad \begin{aligned} m_4 &= 1 \\ m_3 &= 0 \\ m_2 &= 0 \\ m_1 &= 1 \end{aligned}$$

Hasonlóan járunk el a többi sajátérték esetén is (gondoljuk végig):

$$\begin{aligned} m_3 &= 5 - 14 + 10 = 1 \\ m_2 + 2m_3 &= 5 - 14 + 12 = 3 \\ m_1 + m_2 + m_3 &= 14 - 12 = 2 \end{aligned} \quad \Rightarrow \quad \begin{aligned} m_3 &= 1 \\ m_2 &= 1 \\ m_1 &= 0 \end{aligned}$$

$$\begin{aligned} m_2 &= 4 - 14 + 11 = 1 \\ m_1 + m_2 &= 14 - 11 = 3 \end{aligned} \quad \Rightarrow \quad \begin{aligned} m_2 &= 1 \\ m_1 &= 2 \end{aligned}$$

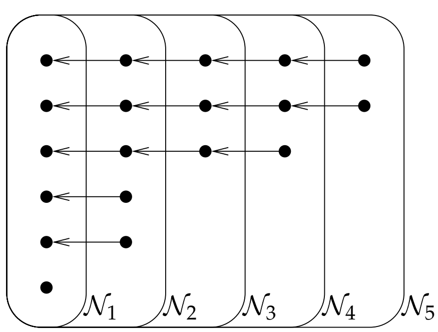
Összefoglalva:

$$\mathbf{J} = \left[\begin{array}{cccc|c|ccc|cc|cc|c|c}
3 & 1 & & & & & & & & & & & & \\
 & 3 & 1 & & & & & & & & & & & \\
 & & 3 & 1 & & & & & & & & & & \\
 & & & 3 & & & & & & & & & & \\
\hline
 & & & & 3 & & & & & & & & & \\
\hline
 & & & & & 2 & 1 & & & & & & & \\
 & & & & & & 2 & 1 & & & & & & \\
 & & & & & & & 2 & & & & & & \\
\hline
 & & & & & & & & 2 & 1 & & & & \\
 & & & & & & & & & 2 & & & & \\
\hline
 & & & & & & & & & & 1 & 1 & & \\
 & & & & & & & & & & & 1 & & \\
\hline
 & & & & & & & & & & & & 1 & \\
\hline
 & & & & & & & & & & & & & 1
\end{array}\right]$$

az $\mathbf{A}$ mátrix Jordan normálalakja. □

A Jordan-bázis konstrukciója A Jordan-tétel bizonyításában megkonstruált bázist kis mátrixokra kézzel is ki lehet számolni. Az itt ismertetendő egyszerű naív algoritmus nem számítógépes megvalósításra való, ennél hatékonyabb is létezik.

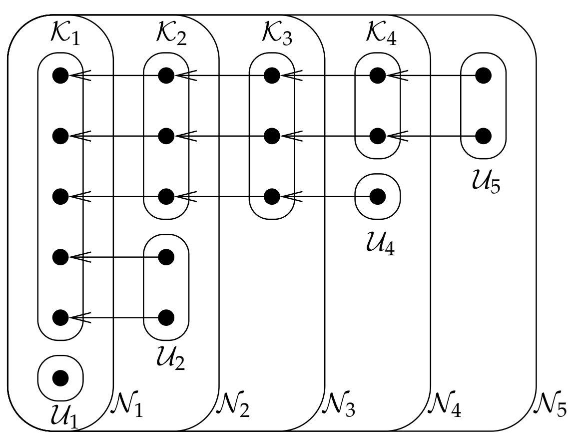
Példaként tekintsünk egy 18-adrendű $\mathbf{A}$ mátrixot, melynek λ 18-szoros sajátértéke, és amelyre a Jordan-láncok hossza rendre $m_5 = 2$, $m_4 = 1$, $m_3 = 0$, $m_2 = 1$, $m_1 = 2$. Ennek alapján az $\mathbf{A} - \lambda\mathbf{I}$ hatása a Jordan-bázison meghatározható, amit a 10.1. ábra szemléltet. Jelölje $(\mathbf{A} - \lambda\mathbf{I})^k$ nullterét $\mathcal{N}_k$. $\mathcal{N}_5$ a λ-hoz tartozó teljes általánosított sajátaltér – e példában $\mathbb{C}^{18}$. A nulltér meghatározása egy homogén lineáris egyenletrendszer megoldását jelenti, ami nem okoz nehézséget, de az már nem teljesen mindegy, hogy a bázist hogy választjuk benne.



10.1. ábra: Az $\mathbf{A} - \lambda\mathbf{I}$ hatványainak nullterei, és hatása az $\mathbf{A}$ mátrix Jordan-bázisán

Az algoritmus – melyet a 10.2. ábra is szemléltet – lényege, hogy a nagyobb indexű terek felől indulva minden lépésben kiterjesztjük $\mathcal{N}_{i-1}$ bázisát $\mathcal{N}_i$ bázisává, majd a kiterjesztésnek vesszük az $\mathbf{A} - \lambda\mathbf{I}$ általi képét, ami már $\mathcal{N}_{i-1}$-be esik, és ezt $\mathcal{N}_{i-2}$ bázisához adva azt kiterjesztjük $\mathcal{N}_{i-1}$ új bázisává. . . . A konvenció az lesz, hogy a vektortér egy generátorának vektorait egyetlen mátrixba tesszük, és azt azonos, de félkövér betűvel jelöljük, tehát pl. $\mathcal{N}$ generátormátrixát $\mathbf{N}$ jelöli. Részletezve:

- Meghatározzuk $\mathcal{N}_i$ egy tetszőleges $\mathbf{N}_i$ bázismátrixát ($i = 1, 2, 3, 4, 5$).
- Kiegészítjük $\mathbf{N}_4$-et $\mathcal{N}_5$ bázisává, az új báziselemek mátrixát jelölje $\mathbf{U}_5$, tehát $\mathcal{N}_5$ bázisa most $[\mathbf{N}_4|\mathbf{U}_5]$.
- Legyen $\mathbf{K}_4$ az új elemek $\mathbf{A} - \lambda\mathbf{I}$ általi képe, azaz $\mathbf{K}_4 = (\mathbf{A} - \lambda\mathbf{I})\mathbf{U}_5$. Mivel $\mathcal{K}_4 \subset \mathcal{N}_4$, de független az $\mathcal{N}_3$ altértől, ezért alkalmas arra, hogy bázisvektorait a Jordan-bázis $\mathcal{N}_4 \setminus \mathcal{N}_3$-ba eső elemei közé vegyük.
- Ha szükséges, egészítsük ki a $[\mathbf{N}_3|\mathbf{K}_4]$-et $\mathcal{N}_4$ bázisává új elemek hozzávételével, így $\mathcal{N}_4$ bázisa most $[\mathbf{N}_3|\mathbf{K}_4|\mathbf{U}_4]$.
- Vegyük a $[\mathbf{K}_4|\mathbf{U}_4]$ mátrix képét, azaz legyen $\mathbf{K}_3 = (\mathbf{A} - \lambda\mathbf{I})[\mathbf{K}_4|\mathbf{U}_4]$, és ha szükséges, egészítsük ki $[\mathbf{N}_2|\mathbf{K}_3]$-at $\mathcal{N}_3$ bázisává új elemek hozzávételével (azaz $\mathcal{N}_3$ bázisa most $[\mathbf{N}_2|\mathbf{K}_3|\mathbf{U}_3]$ – most $\mathcal{U}_3 = \emptyset$). Hasonlóan folytatjuk $\mathcal{N}_i$ indexét csökkentve: $\mathbf{K}_2 = (\mathbf{A} - \lambda\mathbf{I})[\mathbf{K}_3|\mathbf{U}_3]$, új elemek hozzávételével előállítjuk $\mathcal{N}_2$ bázisát: $[\mathbf{N}_1|\mathbf{K}_2|\mathbf{U}_2]$.
- Végül legyen $\mathbf{K}_1 = (\mathbf{A} - \lambda\mathbf{I})[\mathbf{K}_2|\mathbf{U}_2]$, amit $\mathbf{U}_1$-gyel kibővítünk $\mathcal{N}_1$ bázisává. A tér Jordan-bázisa a kép- és az új elemek egyesítése:
$$[\mathbf{K}_1|\mathbf{K}_2|\mathbf{K}_3|\mathbf{K}_4|\mathbf{U}_1|\mathbf{U}_2|\mathbf{U}_3|\mathbf{U}_4|\mathbf{U}_5]\,.$$
- Végül a Jordan-bázis elemeit úgy rendezzük, hogy a láncokat egymás után, minden láncot a sajátvektorból indulva felsoroljuk.



10.2. ábra: A Jordan-bázist megkonstruáló algoritmus

Ezek után lássunk egy konkrét példát, majd az algoritmust általánosan. A számolás kivitelezéséhez két megjegyzés:

► Ha $\mathcal{U} \subset \mathcal{V}$ két altér, és $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$, illetve $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ $(m < n)$ a bázisuk, akkor elemi sorműveletekkel konstruálhatunk $\mathcal{V}$-nek egy olyan $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ bázist, hogy $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ az $\mathcal{U}$ bázisa legyen. Ehhez írjuk $\mathcal{U}$, majd $\mathcal{V}$ bázisvektorait egyetlen

$$[\mathbf{u}_1 \ \ldots \ \mathbf{u}_m \mid \mathbf{v}_1 \ \ldots \ \mathbf{v}_n]$$

mátrixba. Ennek lépcsős alakjában vezéregyesek lesznek az első m oszlopban, és $n - m$ további oszlopban. A nekik megfelelő vektorok az eredeti bázisokban (tehát $\mathcal{V}$ összes vektora és $\mathcal{U}$-nak $n - m$ vektora) adják az új bázist.

► Emlékeztetünk rá, hogy ha egy mátrix redukált lépcsős alakja $[\mathbf{I}|\mathbf{S}]$ alakú, akkor $\left[\begin{smallmatrix}-\mathbf{S}\\ \mathbf{I}\end{smallmatrix}\right]$ vagy $\left[\begin{smallmatrix}\mathbf{S}\\ -\mathbf{I}\end{smallmatrix}\right]$ oszlopvektorai a mátrix nulltere bázisát adják. Vigyázzunk, ha $\mathbf{I}$ nem az első oszlopokban van, akkor $\left[\begin{smallmatrix}-\mathbf{S}\\ \mathbf{I}\end{smallmatrix}\right]$ sorait megfelelően permutálni kell.

10.11. PÉLDA (JORDAN-BÁZIS ELŐÁLLÍTÁSA). *Határozzuk meg az*

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & -2 & 1 & -1 \\ 3 & -3 & 6 & -2 & 4 \\ 4 & -5 & 9 & -3 & 5 \\ 4 & -5 & 8 & -2 & 5 \\ -1 & 1 & -1 & 0 & 0 \end{bmatrix}$$

mátrix Jordan normálalakját és Jordan-bázisát!

MEGOLDÁS. A karakterisztikus polinom

$$\det(\mathbf{A} - \lambda\mathbf{I}) = -\lambda^5 + 4\lambda^4 - 6\lambda^3 + 4\lambda^2 - \lambda = -\lambda(1-\lambda)^4.$$

A 0-hoz tartozó sajátvektor a redukált lépcsős alakból kiszámítva:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1 \\ -3 \\ -4 \\ -4 \\ 1 \end{bmatrix}$$

Mivel itt az algebrai és geometriai multiplicitás egyaránt 1, ez a Jordan-lánc egyelemű. A $\lambda = 1$ esetén a geometriai multiplicitás 2, ugyanis $\mathbf{A} - \mathbf{I}$ redukált lépcsős alakja és abból a nulltér bázisa:

$$\operatorname{rref}(\mathbf{A} - \mathbf{I}) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \mathbf{N}_1 = \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$(\mathbf{A} - 1 \cdot \mathbf{I})^2$ nullterét gyorsabb úgy számolni, ha nem $\mathbf{A} - \mathbf{I}$-t szorozzuk önmagával, hanem lépcsős alakját jobbról, és annak a szorzatnak vesszük a lépcsős alakját. A lépcsős alak kiszámolása ugyanis csak elemi mátrixokkal való balról szorzást jelent, így

$$(\operatorname{rref}(\mathbf{A} - \mathbf{I}))(\mathbf{A} - \mathbf{I}) = \mathbf{E}(\mathbf{A} - \mathbf{I})(\mathbf{A} - \mathbf{I}) = \mathbf{E}(\mathbf{A} - \mathbf{I})^2$$

vagyis a szorzat az $(\mathbf{A} - \mathbf{I})^2$-en végrehajtott elemi sorműveletek eredménye. Sokkal kevesebb viszont a számolnivaló, mivel a 0-sorok elhagyhatók:

$$\left\{ \begin{matrix} (\operatorname{rref}(\mathbf{A} - \mathbf{I}))\,(\mathbf{A} - \mathbf{I}) \\ \text{a 0-sorok nélkül} \end{matrix} \right\} = \begin{bmatrix} -1 & 1 & -2 & 1 & -1 \\ 0 & 0 & -1 & 1 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -1 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 & 0 \end{bmatrix}$$

ahonnan a bázis vektorai:

$$\mathbf{N}_2 = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Hasonlóan határozzuk meg $(\mathbf{A} - \mathbf{I})^3$ bázisát!

$$\begin{bmatrix} 1 & -1 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 & 0 \end{bmatrix} (\mathbf{A} - \mathbf{I}) = \begin{bmatrix} 1 & -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

ami már lépcsős alakú, és ahonnan a bázismátrix

$$\mathbf{N}_3 = \begin{bmatrix} 1 & -1 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Ezután határozzuk meg $\mathbf{U}_3$ vektorait, vagyis azokat, amelyek $\mathcal{N}_2$ bázisát ($\mathbf{N}_2$-t) $\mathcal{N}_3$ bázisává egészítik ki. Ehhez az $[\mathbf{N}_2|\mathbf{N}_3]$ mátrixot kell redukált lépcsős alakra hozni, $\mathcal{U}_3$ elemei a $\mathbf{N}_3$ azon oszlopai lesznek, melyek függetlenek $\mathbf{N}_2$-től, azaz melyek redukált lépcsős alakjában vezéregyes van.

$$[\mathbf{N}_2|\mathbf{N}_3] = \left[\begin{array}{ccc|cccc} 1 & -1 & -1 & 1 & \mathbf{-1} & 0 & -1 \\ 1 & 0 & 0 & 1 & \mathbf{0} & 0 & 0 \\ 0 & 1 & 0 & 0 & \mathbf{1} & 0 & 0 \\ 0 & 1 & 0 & 0 & \mathbf{0} & 1 & 0 \\ 0 & 0 & 1 & 0 & \mathbf{0} & 0 & 1 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|cccc} 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & \mathbf{1} & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Tehát $\mathbf{U}_3 = [-1\ 0\ 1\ 0\ 0]^T$, és ebből $\mathbf{K}_2 = (\mathbf{A} - \mathbf{I})\mathbf{U}_3 = [-1\ 3\ 4\ 4\ 0]^T$. Mivel $\mathbf{K}_2$ egyetlen vektorból áll, és $\mathcal{N}_3$ és $\mathcal{N}_2$ dimenzióinak különbsége is 1, ezért itt nem kell számolnunk semmit, $\mathcal{U}_2 = \{\mathbf{0}\}$, azaz $\mathbf{U}_2$ üres. $\mathbf{K}_1 = (\mathbf{A} - \mathbf{I})\mathbf{K}_2 = [0\ 1\ 1\ 1\ 0]^T$, és mivel e vektor benn van $\mathcal{N}_1$ bázisában, $\mathbf{U}_1 = [0\ 1\ 0\ 0\ 1]^T$ teret a másik bázisvektor generálja. A Jordan normálalak felírásához a Jordan-láncok vektorait egymás után fel kell sorolni, a belőlük képzett $\mathbf{P}$ mátrixszal lesz $\mathbf{J} = \mathbf{P}^{-1}\mathbf{AP}$. E két mátrix

$$\mathbf{P} = \left[\begin{array}{c|ccc|r} 0 & 0 & -1 & -1 & 1 \\ 1 & 1 & 3 & 0 & -3 \\ 0 & 1 & 4 & 1 & -4 \\ 0 & 1 & 4 & 0 & -4 \\ 1 & 0 & 0 & 0 & 1 \end{array}\right] \qquad \mathbf{J} = \left[\begin{array}{c|ccc|c} 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right] \qquad \square$$

Az algoritmus általánosan:

- Input: $\mathbf{A}$, $p(x)$ karakterisztikus polinom lineáris tényezőkre bontva,
- Minden λ sajátértékre
  - Határozzuk meg a leghosszabb lánc s hosszát, és $(\mathbf{A} - \lambda\mathbf{I})^i$ nullterét ($\mathcal{N}_i$, $i = 1, 2, \dots, s$). Legyen $\mathcal{U}_{s+1} = \mathcal{K}_{s+1} = \mathcal{N}_0 = \{\mathbf{0}\}$, azaz e terek bázisa az üreshalmaz.
  - Minden i-re s-től 1-ig haladva:
    * Legyen $\mathbf{K}_i = (\mathbf{A} - \lambda\mathbf{I})[\mathbf{K}_{i+1}|\mathbf{U}_{i+1}]$.
    * Határozzuk meg $\mathbf{U}_i$-t úgy, hogy $[\mathbf{N}_{i-1}|\mathbf{K}_i|\mathbf{U}_i]$ bázisa legyen $\mathcal{N}_i$-nek. Ehhez az $[\mathbf{N}_{i-1}|\mathbf{K}_i|\mathbf{N}_i]$ mátrix redukált lépcsős alakja alapján válasszuk a vezéroszlopokat $\mathbf{N}_i$-ből az $\mathbf{U}_i$-be.
- Tegyük a Jordan-láncok vektorait balról jobbra egymás mellé minden láncot a sajátvektorral kezdve, az így kapott $\mathbf{P}$ mátrixszal $\mathbf{J} = \mathbf{P}^{-1}\mathbf{AP}$.

10.12. DEFINÍCIÓ (MÁTRIX EXPONENCIÁLIS FÜGGVÉNYE). *Az* $\mathbf{A}$ *négyzetes mátrix exponenciális függvényét a következő sor definiálja:*

$$e^{\mathbf{A}} = \mathbf{I} + \mathbf{A} + \frac{\mathbf{A}^2}{2!} + \dots + \frac{\mathbf{A}^n}{n!} + \dots$$

10.13. ÁLLÍTÁS (EXPONENCIÁLIS FÜGGVÉNY KISZÁMÍTÁSA). *Ha* $\mathbf{J} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ *az* $\mathbf{A}$ *mátrix Jordan normálalakja, akkor*

$$e^{\mathbf{A}} = \mathbf{P}e^{\mathbf{J}}\mathbf{P}^{-1},$$

ahol $e^{\mathbf{J}}$ *blokkonként számolható, és*

$$\mathbf{J}_\lambda = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{bmatrix} \text{ esetén } e^{\mathbf{J}_\lambda} = e^\lambda \begin{bmatrix} 1 & 1 & \frac{1}{2!} & \dots & \frac{1}{(n-1)!} \\ 0 & 1 & 1 & \dots & \frac{1}{(n-2)!} \\ 0 & 0 & 1 & \dots & \frac{1}{(n-3)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

Mátrixfüggvények Általában mondható, hogy

$$f\left(\begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \lambda \end{bmatrix}\right) = \begin{bmatrix} \frac{f(\lambda)}{0!} & \frac{f'(\lambda)}{1!} & \frac{f''(\lambda)}{2!} & \dots & \frac{f^{(n-1)}(\lambda)}{(n-1)!} \\ 0 & \frac{f(\lambda)}{0!} & \frac{f'(\lambda)}{1!} & \dots & \frac{f^{(n-2)}(\lambda)}{(n-2)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \frac{f(\lambda)}{0!} \end{bmatrix}.$$

10.14. PÉLDA (MÁTRIX EXPONENCIÁLIS FÜGGVÉNYE). *Legyen*

$$\begin{pmatrix} -3 & 2 & 1 \\ 1 & -2 & 1 \\ -1 & -2 & -5 \end{pmatrix}$$

Határozzuk meg az $e^{\mathbf{A}}$ *mátrixot!*

MEGOLDÁS. $\mathbf{A}$ karakterisztikus polinomja

$$x^3 + 10x^2 + 32x + 32 = (x+2)(x+4)^2,$$

így

$$\mathbf{J} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -4 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & -1 & -2 \end{bmatrix},$$

$$e^{\mathbf{A}} = \mathbf{P}e^{\mathbf{J}}\mathbf{P}^{-1} = \begin{bmatrix} \frac{1}{2}\ \frac{e^2+1}{e^{4)}} & \frac{e^2-1}{e^4} & \frac{1}{2}\ \frac{e^2-1}{e^4} \\ \frac{1}{2}\ \frac{e^2-1}{e^4} & e^2 & \frac{1}{2}\ \frac{e^2-1}{e^4} \\ -\frac{1}{2}\ \frac{e^2-1}{e^4} & -\frac{e^2-1}{e^4} & -\frac{1}{2}\ \frac{e^2-3}{e^4} \end{bmatrix} \qquad \square$$

A Jordan normálalak használata a differenciálegyenletrendszerek megoldásában

11
Nemnegatív mátrixok

Különösen sok alkalmazása van azoknak a mátrixoknak, melyek elemei nem negatív számok. Ilyen mátrixok például azok, melyek elemei mérési eredmények, gazdasági adatok, valószínűségek,...

Mátrixok összehasonlítása

Mátrixok elemenkénti összehasonlítására a szokásos relációjeleket fogjuk használni. $\mathbf{A} > \mathbf{B}$ azt jelenti, hogy mindkét mátrix azonos méretű, és $a_{ij} > b_{ij}$. Hasonlóan $\mathbf{A} \geq \mathbf{B}$, ha $a_{ij} \geq b_{ij}$. Egy $\mathbf{A}$ mátrixot *pozitívnak* (*nemnegatívnak*) nevezünk, ha $\mathbf{A} > \mathbf{O}$ ($\mathbf{A} \geq \mathbf{O}$), azaz ha $a_{ij} > 0$ ($a_{ij} \geq 0$). Itt $\mathbf{O}$ a nullmátrixot jelöli. E fogalmakat és jelöléseket vektorokra is használjuk: az $\mathbf{x}$ vektor pozitív, azaz $\mathbf{x} > \mathbf{0}$, ha $\mathbf{x}$ minden koordinátája pozitív.

Néhány könnyen igazolható észrevétel:

$$\mathbf{A} \geq \mathbf{O} \Leftrightarrow \mathbf{A}\mathbf{x} \geq \mathbf{0} \text{ minden } \mathbf{x} \geq 0 \text{ vektorra,} \tag{11.1}$$

$$\mathbf{A} > \mathbf{O} \Leftrightarrow \mathbf{A}\mathbf{x} > \mathbf{0} \text{ minden } \mathbf{x} \geq 0,\ \mathbf{x} \neq \mathbf{0} \text{ vektorra,} \tag{11.2}$$

$$\mathbf{A} \geq \mathbf{O}, \text{ és } \mathbf{x} \geq \mathbf{y} \geq \mathbf{0} \Rightarrow \mathbf{A}\mathbf{x} \geq \mathbf{A}\mathbf{y}, \tag{11.3}$$

$$\mathbf{A} > \mathbf{O},\ \mathbf{B} \geq \mathbf{O} \Rightarrow \mathbf{A}\mathbf{B} > \mathbf{O}. \tag{11.4}$$

Később fontos szerepet kapnak a következő mátrixok is: egy mátrix *primitív*, ha nemnegatív, de valamely pozitív egészkitevős hatványa pozitív. Például a $\left[\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right]$ mátrix pozitív, így primitív is, hisz első hatványa pozitív, az $\left[\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ mátrix nemnegatív, de primitív, mert $\left[\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right]^2 = \left[\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right] > \mathbf{O}$, míg az $\left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right]$ mátrix nemnegatív, de nem primitív, mivel tetszőleges pozitív egész n-re $\left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right]^n = \left[\begin{smallmatrix} 1 & 0 \\ n & 1 \end{smallmatrix}\right]$, ami nem pozitív.

A valós vagy komplex elemű $\mathbf{A}$ mátrix $\varrho(\mathbf{A})$ *spektrálsugarán* a legnagyobb abszolút értékű sajátértékének abszolút értékét értjük. Másként fogalmazva a spektrálsugár a komplex számsík legkisebb olyan origó középű körének a sugara, amely tartalmazza az összes sajátértéket.

Pozitív mátrixok

E szakaszban csak pozitív mátrixokat vizsgálunk. Az itt ismertetendő elmélet Perrontól származik, melyet két tételbe foglalunk össze.

11.1. TÉTEL (PERRON-TÉTEL: POZITÍV SAJÁTÉRTÉK ÉS SAJÁTVEKTOR). *Ha* $\mathbf{A}$ *pozitív mátrix, és* $r = \varrho(\mathbf{A})$ *jelöli a spektrálsugarát, akkor*

1. $r > 0$,
2. *r sajátérték egy pozitív sajátvektorral,*
3. $\mathbf{A}$*-nak e pozitív sajátvektor skalárszorosain kívül nincs más nemnegatív sajátvektora.*

BIZONYÍTÁS. 1. Ha $r = 0$, akkor $\mathbf{A}$ minden sajátértéke 0, azaz $\mathbf{A}$ nilpotens a **??** tétel szerint. Ez viszont pozitív mátrixra lehetetlen, hisz annak minden hatványa pozitív, tehát semelyik sem $\mathbf{O}$.

2. Legyen $\lambda \in \mathbb{C}$ egyike a legnagyobb abszolút értékű sajátértékeknek, azaz $|\lambda| = r$, és legyen az $\mathbf{x}$ sajátvektorral $\mathbf{Ax} = \lambda\mathbf{x}$. Legyen $\mathbf{p}$ az $\mathbf{x}$ koordinátáinak abszolút értékéből álló vektor, azaz $\mathbf{p} = (|x_1|, |x_2|, \dots, |x_n|)$. Írjuk fel az $\mathbf{Ax} = \lambda\mathbf{x}$ mindkét oldalának i-edik koordinátáját, majd vegyük annak abszolút értékét:

$$\left|\sum_{j=1}^{n} a_{ij}x_j\right| = |\lambda||x_i|.$$

Ebből, a háromszög-egyenlőtlenséget fölhasználva kapjuk, hogy

$$rp_i = |\lambda||x_i| = \left|\sum_{j=1}^{n} a_{ij}x_j\right| \le \sum_{j=1}^{n} a_{ij}p_j, \quad \text{azaz } r\mathbf{p} \le \mathbf{Ap}.$$

Ha itt egyenlőség áll, kész vagyunk, hisz $r\mathbf{p} = \mathbf{Ap}$ esetén r valóban sajátérték. Ha nem, akkor az $\mathbf{u} = \mathbf{Ap} - r\mathbf{p}$ vektorra $\mathbf{u} \ge \mathbf{0}$ és $\mathbf{u}$ legalább egyik koordinátája határozottan pozitív. A (11.2) szerint ekkor $\mathbf{Au} > \mathbf{0}$, azaz $\mathbf{A}(\mathbf{Ap}) - r\mathbf{Ap} > \mathbf{0}$. A $\mathbf{v} = \mathbf{Ap}$ jelöléssel eszerint $\mathbf{Av} > r\mathbf{v}$, ahol $\mathbf{v} > \mathbf{0}$. Megmutatjuk, hogy ez ellentmondásra vezet, azaz megmutatjuk, hogy nincs olyan $\mathbf{v}$ vektor, hogy $\mathbf{Av} > r\mathbf{v}$. Legyen $\varepsilon > 0$ egy olyan szám, melyre még fennáll az $\mathbf{Av} \ge (r+\varepsilon)\mathbf{v}$ egyenlőtlenség. A $\mathbf{B} = \frac{1}{r+\varepsilon}\mathbf{A}$ mátrixra tehát egyrészt $\mathbf{Bv} \ge \mathbf{v}$, másrészt $\varrho(\mathbf{B}) = \varrho(\frac{1}{r+\varepsilon}\mathbf{A}) = \frac{r}{r+\varepsilon} < 1$, azaz $\mathbf{B}$ spektrálsugara 1-nél kisebb. Ez azt jelenti, hogy $\lim_{k\to\infty} \mathbf{B}^k = \mathbf{O}$. Így a $\mathbf{v} < \mathbf{Bv} < \mathbf{B}^2\mathbf{v} < \cdots < \mathbf{B}^k\mathbf{v}$ vektorsorozat a $\mathbf{0}$ vektorhoz tart, vagyis $\mathbf{v} < \mathbf{0}$, ami ellentmond korábbi feltevésünknek. Ezzel bizonyítottuk, hogy $\mathbf{Ap} = r\mathbf{p}$.
Még meg kell mutatnunk, hogy $\mathbf{p} > \mathbf{0}$. Tudjuk, hogy $\mathbf{p} \ge \mathbf{0}$, így a (11.2) összefüggés miatt $\mathbf{Ap} > \mathbf{0}$, de $\mathbf{Ap} = r\mathbf{p}$, tehát $r\mathbf{p} > 0$, azaz $\mathbf{p} > \mathbf{0}$.

3. Indirekt feltevés szerint legyen $(\lambda, \mathbf{x})$ egy sajátpár, azaz $\mathbf{Ax} = \lambda\mathbf{x}$ és ahol $\mathbf{x} \ge \mathbf{0}$, és $\lambda \ne r$, legyen továbbá $\mathbf{q} > \mathbf{0}$ az ugyancsak pozitív, és azonos spektrumú $\mathbf{A}^T$ mátrix r-hez tartozó pozitív sajátvektora, azaz

$\mathbf{q}^T\mathbf{A} = r\mathbf{q}^T$. Ekkor

$$r\mathbf{q}^T\mathbf{x} = (\mathbf{q}^T\mathbf{A})\mathbf{x} = \mathbf{q}^T(\mathbf{A}\mathbf{x}) = \lambda\mathbf{q}^T\mathbf{x},$$

amiből $\mathbf{q}^T\mathbf{x} > 0$ miatt $r = \lambda$ adódik, ellentmondás. □

Az alkalmazásokban gyakori szerepe indokolja a következő kitüntető elnevezést. Pozitív mátrix spektrálsugarához, mint sajátértékhez tartozó pozitív $\mathbf{p}$ sajátvektorát *Perron-vektornak* nevezzük, ha koordinátáinak összege 1. A hasonló módon definiált bal sajátvektort *bal Perron-vektornak* nevezzük. Ez megegyezik az $\mathbf{A}^T$ Perron-vektorával. Összefoglalva: a $\mathbf{p}$ Perron-vektort és a $\mathbf{q}$ bal Perron-vektort a

$$\mathbf{A}\mathbf{p} = r\mathbf{p},\ \sum_{i=1}^{n} p_i = 1, \qquad \mathbf{q}^T\mathbf{A} = r\mathbf{q}^T,\ \sum_{i=1}^{n} q_i = 1$$

képletek definiálják.

11.2. TÉTEL (PERRON-TÉTEL: EGYSZERES ÉS DOMINÁNS SAJÁTÉRTÉK). *Ha $\mathbf{A}$ pozitív mátrix, és $r = \varrho(\mathbf{A})$, akkor*

1. *r egyszeres sajátértéke $\mathbf{A}$-nak,*
2. *r domináns, azaz minden további λ sajátértékre $|\lambda| < r$.*

BIZONYÍTÁS. 1. Megmutatjuk, hogy r egyszeres sajátérték. Tegyük fel először, hogy van r-hez tartozó, de $\mathbf{p}$-től független $\mathbf{s}$ sajátvektor. Ekkor megfelelő c konstanssal elérhető, hogy a $\mathbf{p} + c\mathbf{s} \geq \mathbf{0}$ vektornak legyen 0 koordinátája. Erről azonban az imént láttuk be, hogy ellentmondásra vezet, tehát minden r-hez tartozó sajátvektor $\mathbf{p}$ többszöröse. Be kell még látnunk, hogy nincs olyan $\mathbf{v} > \mathbf{0}$ általánosított sajátvektor, hogy $(\mathbf{A} - r\mathbf{I})\mathbf{v} = \mathbf{p}$, azaz $\mathbf{A}\mathbf{v} = r\mathbf{v} + \mathbf{p}$ legyen. Ha ugyanis volna, akkor $\mathbf{A}\mathbf{v} = r\mathbf{v} + \mathbf{p} \geq r\mathbf{v}$ lenne, ami az előző pontban bizonyítottak szerint ellentmondásra vezet. Tegyük fel tehát, hogy létezik Könnyen elérhető a $\mathbf{p}$ egy megfelelően nagy d konstansszorosának hozzáadásával, hogy $\mathbf{v} > \mathbf{0}$ legyen, ugyanis $(\mathbf{A} - r\mathbf{I})(\mathbf{v} + d\mathbf{p}) = \mathbf{p} + d(\mathbf{A} - r\mathbf{I})\mathbf{p} = \mathbf{p}$ miatt ha $\mathbf{v}$ általánosított sajátvektor, akkor $\mathbf{v} + d\mathbf{p}$ is.

2. Belátjuk, hogy ha λ az $\mathbf{A}$ egy sajátértéke, akkor $|\lambda| < r$. Indirekt módon legyen $|\lambda| = r$. E bizonyítás 2. pontjában leírtakat ismételve a komplex számok összegére vonatkozó háromszögegyenlőtlenséget alkalmazva kapjuk, hogy

$$\sum_{j=1}^{n} a_{ij}|x_j| \leq \left|\sum_{j=1}^{n} a_{ij}x_j\right| = |\lambda||x_i|. \tag{11.5}$$

Mint azt beláttuk, ekkor $(|x_1|, |x_2|, \ldots, |x_n|)$ sajátvektor, a hozzá tartozó sajátérték $r = |\lambda|$, és a (11.5) egyenlőtlenségben egyenlőségnek kell állnia. A komplex számokra vonatkozó $|z_1 + \cdots + z_k| = |z_1| + \cdots + |z_k|$ egyenlőség csak akkor áll fenn, ha mindegyik komplex szám azonos argumentumú. Ez esetünkben azt jelenti, hogy van olyan φ szög, hogy minden i-re $x_i = \mathrm{e}^{\mathrm{i}\varphi}|x_i|$. Eszerint $\mathbf{x} = \mathrm{e}^{\mathrm{i}\varphi}\mathbf{p}$, tehát $\lambda = r$. □

Tipográfiai különbség van az imaginárius egység álló i-je és a változó index dőlt *i*-je között!

Feladatok

11.1. Legyen

$$\mathbf{A} = \begin{bmatrix} 6 & 1 & 1 \\ 5 & 6 & 1 \\ 6 & 4 & 4 \end{bmatrix}$$

Számítsuk ki a két Perron-vektort, és ellenőrizzük Perron tételét.

11.2. Egy pozitív elemű 4-edrendű mátrix három sajátértéke 1, $2i$, $-2i$. A -3, 2, 3, $4i$, 4 számok közül válasszuk ki mindegyik olyat, amelyik a negyedik sajátérték lehet!

11.3. Mutassuk meg, hogy ha az $\mathbf{A} > \mathbf{O}$ mátrix minden oszlopában vagy minden sorában c az elemek összege, akkor c a spektrálsugár.

Nemnegatív mátrixok

A pozitív mátrixok Perron tételeiben kimondott tulajdonságai közül változtatás nélkül egyik sem marad érvényben nemnegatív mátrixokra. Például

- a $\left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$ mátrix nemnegatív, de mivel mindkét sajátértéke 0, ezért spektrálsugara is 0,
- az $\left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ mátrix spektrálsugara 1, de az 1 kétszeres sajátérték, és több lineárisan független pozitív sajátvektor is tartozik hozzá,
- a $\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ mátrix sajátértékei 1 és -1, így spektrálsugara ugyancsak 1, de a spektrálkörön több különböző sajátértéke is van,
- az $\left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right]$ mátrixnak nincs pozitív sajátvektora.

Ugyanakkor az sem mondható, hogy ha egy nemnegatív mátrixnak vannak 0 elemei, akkor nem teljesülnek a Perron-tételek állításai. Például

- az $\left[\begin{smallmatrix} 1 & 1 \\ 2 & 0 \end{smallmatrix}\right]$ mátrix nemnegatív, sajátértékei 2, -1, spektrálsugara tehát 2, ami egyszeres sajátérték, és a spektrálkörön az egyetlen sajátérték, a hozzá tartozó $(1,1)$ sajátvektor pozitív.

A Perron-tételek állításaiból némi gyengítés után, de még az összes nemnegatív mátrixra érvényes módon a következő marad:

11.3. TÉTEL (PERRON–FROBENIUS-TÉTEL – GYENGE VÁLTOZAT).*Ha* $\mathbf{A}$ *nemnegatív mátrix, akkor az* $r = \varrho(\mathbf{A})$ *spektrálsugár sajátértéke* $\mathbf{A}$*-nak, melyhez tartozik nemnegatív sajátvektor.*

BIZONYÍTÁS. A bizonyítás alapötlete, hogy az $\mathbf{A}$ nemnegatív mátrixot pozitív mátrixokkal közelítjük, melyekre használhatók Perron-tételei. Legyen

$$\mathbf{A}_k = \mathbf{A} + \frac{1}{k}\begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{bmatrix}, \quad k \in \mathbb{N}.$$

Jelölje $\mathbf{A}_k$ spektrálsugarát r_k, Perron-vektorát $\mathbf{p}_k$, az $\mathbf{A}$ mátrix spektrálsugarát r. A $\mathbf{p}_k$ vektorok korlátos halmazt alkotnak $\mathbb{R}^n$-ben, mivel mindegyik koordinátájuk 0 és 1 közé esik, így benne vannak az egységkockában. A Bolzano–Weierstrass-tétel szerint kiválasztható közülük egy konvergens $\mathbf{p}_{k_m}$ részsorozat. A határértéket jelölje $\mathbf{p}$. Megmutatjuk, hogy $\mathbf{p}$ az $\mathbf{A}$-nak r-hez tartozó sajátvektora és hogy $\mathbf{p} \geq \mathbf{0}$, de $\mathbf{p} \neq \mathbf{0}$. Mivel $\mathbf{p}_k > \mathbf{0}$, ezért a határértékéről azt tudjuk, hogy $\mathbf{p} \geq \mathbf{0}$. Tekintsük a folytonos $f(\mathbf{x}) = \sum_{i=1}^{n} x_i$ függvényt. Mivel $f(\mathbf{p}_{k_m}) = 1$, ezért $f(\mathbf{p}) = 1$ is fennáll, így $\mathbf{p} \neq \mathbf{0}$.

Tekintsük ezután az r_k sorozatot. A ?? tétel szerint, $\mathbf{A}_1 > \mathbf{A}_2 > \dots \mathbf{A}_k > \dots > \mathbf{A}$, így $r_1 \geq r_2 \geq \dots \geq r_k \geq r$, azaz az r_k sorozat monoton csökkenő, és alulról korlátos, tehát konvergens. Határértékét jelölje $\hat{r}$. Ez egyúttal az r_{k_m} részsorozatnak is határértéke. A fentiek szerint $\hat{r} \geq r$. Másrészt

$$\mathbf{Ap} = \mathbf{A}(\lim_{m\to\infty} \mathbf{p}_{k_m}) = \lim_{m\to\infty} \mathbf{Ap}_{k_m} = \lim_{m\to\infty} r_{k_m}\mathbf{p}_{k_m} = \hat{r}\mathbf{p}. \qquad \square$$

Tehát $\hat{r}$ sajátérték, akkor viszont $\hat{r} \leq r$. Kaptuk tehát, hogy $\hat{r} = r$ és $\mathbf{Ap} = r\mathbf{p}$.

A következőkben két olyan tételt mondunk ki, melyek a nemnegatív – és így a pozitív – mátrixokra korlátozás nélkül érvényesek.

11.4. TÉTEL (COLLATZ–WIELANDT-TÉTEL). *Az $\mathbf{A} \geq \mathbf{O}$ mátrix r spektrálsugarára*

$$r = \max_{\substack{\mathbf{x} \\ \mathbf{0}\neq\mathbf{x}\geq\mathbf{0}}} \min_{\substack{1\leq i\leq n \\ x_i\neq 0}} \frac{[\mathbf{Ax}]_i}{x_i}. \tag{11.6}$$

Másként megfogalmazva:

$$r = \max_{\substack{\mathbf{x} \\ \mathbf{0}\neq\mathbf{x}\geq\mathbf{0}}} \max_{c\mathbf{x}\leq\mathbf{Ax}} c \tag{11.7}$$

A képletek úgy értendők, hogy minden $\mathbf{x}$ vektorra kiszámítjuk az $[\mathbf{Ax}]_i/x_i$ törtek minimumát, és ezen értékek maximumát vesszük, ha $\mathbf{x}$ végigfut a nemnegatív, de nullvektortól különböző vektorokon. Az $x_i = 0$ esetet a keresésből kizártuk, de mondhattuk volna azt is, hogy a tört ekkor legyen ∞, így nem változna a minimum. A második képletben minden $\mathbf{x}$ vektorra meghatározzuk azt a legnagyobb c számot, melyre $c\mathbf{x} \leq \mathbf{Ax}$, majd vesszük az így kapott c értékek maximumát.

BIZONYÍTÁS. A két megfogalmazás nyilván ekvivalens, hisz ha egy adott $\mathbf{x} \geq \mathbf{0}$ vektorra c az $\frac{[\mathbf{Ax}]_i}{x_i}$ törtek minimuma, akkor c egyúttal a legnagyobb olyan szám, melyre $c\mathbf{x} \leq \mathbf{Ax}$.

Először pozitív $\mathbf{A}$ mátrixra bizonyítunk. Legyen $\mathbf{q}$ a bal Perron-vektor, r a spektrálsugár. Ekkor a $\mathbf{q}^T\mathbf{x} > 0$ számmal való osztás lehetőségét is használva

$$c\mathbf{x} \leq \mathbf{Ax} \quad \rightsquigarrow \quad c\mathbf{q}^T\mathbf{x} \leq \mathbf{q}^T\mathbf{Ax} = r\mathbf{q}^T\mathbf{x} \quad \rightsquigarrow \quad c \leq r.$$

Másrészt az $\mathbf{x} = \mathbf{p}$ vektorra $r\mathbf{p} = \mathbf{Ap}$, tehát a lehetséges c értékek maximuma r.

Ezután marad az $\mathbf{A} \geq \mathbf{O}$ eset. Az előző tétel bizonyításában használt ötletet és az ott definiált $\mathbf{A}_k$ mátrixot használva kapjuk, és bal Perron-vektorát $\mathbf{q}_k$-val jelölve, hogy egy rögzített $\mathbf{x} \geq \mathbf{0}$, $\mathbf{x} \neq \mathbf{0}$ vektorra

$$0 \leq c\mathbf{x} \leq \mathbf{Ax} \leq \mathbf{A}_k\mathbf{x} \quad \rightsquigarrow \quad c\mathbf{q}_k^T\mathbf{x} \leq \mathbf{q}_k^T\mathbf{A}_k\mathbf{x} = r_k\mathbf{q}_k^T \quad \rightsquigarrow \quad c \leq r_k \leq r.$$

Ekkor az előzőekhez hasonlóan az $\mathbf{x} = \mathbf{p}$ vektorra $r\mathbf{p} = \mathbf{Ap}$, tehát a lehetséges c értékek maximuma r. $\square$

11.5. TÉTEL (Nemnegatív mátrixok spektrálsugarának becslése). *Ha $\mathbf{A} \geq \mathbf{O}$, akkor a spektrálsugár a sorösszegek minimuma és maximuma, illetve az oszlopösszegek minimuma és maximuma közé esik, azaz*

$$\min_i \left\{ \sum_{j=1}^n a_{ij} \right\} \leq \varrho(\mathbf{A}) \leq \max_i \left\{ \sum_{j=1}^n a_{ij} \right\}$$
$$\min_j \left\{ \sum_{i=1}^n a_{ij} \right\} \leq \varrho(\mathbf{A}) \leq \max_j \left\{ \sum_{i=1}^n a_{ij} \right\}$$

BIZONYÍTÁS. Az első egyenlőtlenség felső korlátjának bizonyításával kezdjük. Legyen $(\lambda, \mathbf{x})$ egy sajátpár, azaz $\lambda\mathbf{x} = \mathbf{A}\mathbf{x}$. Válasszuk úgy $\mathbf{x}$-et, hogy legnagyobb koordinátája 1 legyen, például $x_i = 1$, és tetszőleges koordinátára $x_j \leq 1$. Ekkor

$$|\lambda| = |\lambda||x_i| = |\lambda x_i| = \left| \sum_{j=1}^n a_{ij} x_j \right| \leq \sum_{j=1}^n a_{ij} |x_j| \leq \sum_{j=1}^n a_{ij}.$$

Ez bizonyítja az első egyenlőtlenség jobb oldalát.

A bal oldali egyenlőtlenség a Collatz–Wielandt-tételből következik, ha ugyanis $\mathbf{x} = \mathbf{1}$, akkor akkor az $[\mathbf{A}\mathbf{x}]_i / x_i$ hányados épp a sorösszeg, tehát annak minimuma kisebb vagy egyenlő r-nél.

A második egyenlőtlenségeket megkapjuk, ha az elsőt $\mathbf{A}^T$-ra alkalmazzuk, melynek spektruma és így spektrálsugara is azonos $\mathbf{A}$-éval. □

Feladatok

11.4. Egy nemnegatív mátrix az $(4,6,5)$ vektort az $(5,6,7)$ vektorba viszi. Mutassuk meg, hogy spektrálsugara legalább 1.

11.5. Legyen $\mathbf{x} > \mathbf{0}$ tetszőleges, és $\mathbf{A} \geq \mathbf{O}$. Igazoljuk az alábbi egyenlőtlenségeket!

$$\min_i \left\{ \sum_{j=1}^{n} a_{ij} \frac{x_j}{x_i} \right\} \leq \varrho(\mathbf{A}) \leq \max_i \left\{ \sum_{j=1}^{n} a_{ij} \frac{x_j}{x_i} \right\}$$

$$\min_j \left\{ \sum_{i=1}^{n} a_{ij} \frac{x_i}{x_j} \right\} \leq \varrho(\mathbf{A}) \leq \max_j \left\{ \sum_{i=1}^{n} a_{ij} \frac{x_i}{x_j} \right\}$$

(Ötlet: ha $\mathbf{D} = \operatorname{diag}(x_1, \ldots, x_n)$, akkor $\mathbf{B} = \mathbf{D}^{-1}\mathbf{A}\mathbf{D}$ hasonló $\mathbf{A}$-hoz, így $\varrho(\mathbf{B}) = \varrho(\mathbf{A})$. Alkalmazzuk a 11.5. tételt.)

11.6. Az előző feladat eredményét használva becsüljük meg az

$$\begin{bmatrix} 6 & 8 & 0 \\ 1 & 2 & 3 \\ 6 & 4 & 2 \end{bmatrix}, \quad \text{és a} \quad \begin{bmatrix} 6 & 3 & 7 \\ 6 & 2 & 2 \\ 1 & 1 & 2 \end{bmatrix}$$

mátrix spektrumát a $\mathbf{x} = (2,1,2)$ vektorral. Az eredmény alapján mit mondhatunk $\mathbf{x}$-ről?

Irreducibilis mátrixok

Az előző szakaszban láttuk, hogy Perron tételei nem maradnak érvényben általában, de vannak mátrixok, amelyekre igen. Frobenius talált rá arra a könnyen ellenőrizhető feltételre, mely alapján eldönthető, hogy egy nemnegatív mátrix melyik csoportba tartozik.

11.6. DEFINÍCIÓ (REDUCIBILIS ÉS IRREDUCIBILIS MÁTRIXOK). *Az* $\mathbf{R}$ *mátrixot* reducibilisnek *nevezzük, ha a sorok és oszlopok azonos permutációjával*

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{C} \end{bmatrix}$$

alakra hozható, ahol $\mathbf{A}$ *és* $\mathbf{C}$ *négyzetes mátrixok. Azaz létezik olyan* $\mathbf{P}$ *permutációmátrix, hogy* $\mathbf{PRP}^T$ *a fenti alakú. Az olyan mátrixot, amely nem hozható ilyen alakra,* irreducibilisnek *nevezzük.*

11.7. PÉLDA. *Döntsük el, hogy az alábbi mátrixok közül melyik reducibilis, melyik irreducibilis! (Segítségül a nemnulla mátrixelemekről leolvasható a sor- és oszlopindex.)*

$$\mathbf{A} = \begin{bmatrix} 11 & 0 & 13 & 14 & 0 \\ 21 & 22 & 23 & 24 & 25 \\ 31 & 0 & 33 & 34 & 0 \\ 41 & 0 & 43 & 44 & 0 \\ 51 & 52 & 53 & 54 & 55 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 12 & 0 & 0 & 0 \\ 0 & 0 & 0 & 24 & 25 \\ 31 & 0 & 0 & 0 & 0 \\ 0 & 0 & 43 & 0 & 0 \\ 0 & 52 & 0 & 54 & 0 \end{bmatrix}.$$

MEGOLDÁS. Az $\mathbf{A}$ mátrixon könnyű észrevenni, hogy reducibilis, mert az első és utolsó sorok és oszlopok cseréjével, vagyis a következő $\mathbf{P}$ permutációmátrixszal a kívánt alakra hozható:

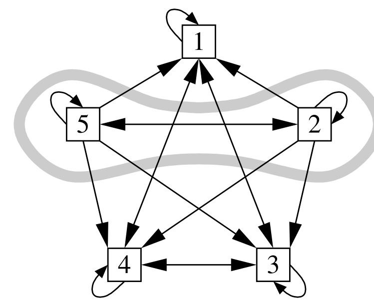
$$\mathbf{P} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{PAP}^T = \left[\begin{array}{cc|ccc} 55 & 52 & 53 & 54 & 51 \\ 25 & 22 & 23 & 24 & 21 \\ \hline 0 & 0 & 33 & 34 & 31 \\ 0 & 0 & 43 & 44 & 41 \\ 0 & 0 & 13 & 14 & 11 \end{array}\right].$$

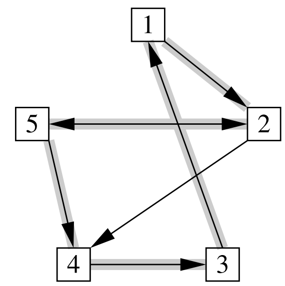
Nem ez az egyetlen permutáció, pl. az $1 \to 3 \to 4 \to 5 \to 2 \to 1$ csere is megteszi:

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{PAP}^T = \left[\begin{array}{cc|ccc} 22 & 25 & 21 & 23 & 24 \\ 52 & 55 & 51 & 53 & 54 \\ \hline 0 & 0 & 11 & 13 & 14 \\ 0 & 0 & 31 & 33 & 34 \\ 0 & 0 & 41 & 43 & 44 \end{array}\right].$$

A második mátrixban több 0 van, azt hinnénk, ez inkább lesz reducibilis, mégsem találunk megfelelő $\mathbf{P}$ permutációmátrixot. De az hogy igazolható, hogy nincs? Segít a következő ötlet. Tekintsük azt az n-csúcsú gráfot, amelyben az i-edik csúcsból a j-edikbe pontosan akkor fut irányított él, ha $a_{ij} \neq 0$.

A mátrix sorain és oszlopain végrehajtott azonos permutáció így megfelel a gráf csúcsai átszámozásának. Például a fenti **A** mátrixhoz rendelt gráf a 11.1. ábrán látható. Vegyük észre, hogy a $\{2,5\}$ ponthalmazból nem fut ki él. Ez azt jelenti, hogy ha a csúcsokat átszámozzuk az 5-ös és 1-es sorszám fölcserélésével, akkor az $\{1,2\}$ halmazból nem fut él a $\{3,4,5\}$ halmazba. Ez épp azt jelenti, hogy bármely mátrixban, melynek ez a gráfja, a bal alsó 3×2-es részmátrixa zérusmátrix. Tehát a mátrix reducibilis. Általánosan megfogalmazva: egy mátrix pontosan akkor reducibilis, ha gráfjának csúcsaiból kiválasztható egy olyan részhalmaz, melyből nem indul ki él a komplementer csúcshalmazba. Ebből az is következik, hogy egy mátrix pontosan akkor irreducibilis, ha ilyen részhalmaz nincs, azaz bármely csúcsból bármely csúcs elérhető irányított úton. Az ilyen gráfot *erősen összefüggőnek* nevezzük. A **B** mátrix gráfja például ilyen, hisz az 1-2-5-4-3-1 útvonalon bármely pontból bármely másik elérhető. Tehát **B** irreducibilis. □





11.1. ábra: Az **A** és **B** mátrixokhoz rendelt két gráf.

► Fontos megjegyezni, hogy a 11.6. definíció a *sorok és oszlopok azonos permutációjáról* szól, tehát nem elég a mátrixot elemi sorműveletekkel $\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{C} \end{bmatrix}$ alakra hozni. Ugyanazokat a műveleteket az oszlopokra is alkalmazni kell. Például a

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

mátrix irreducibilis, hisz egy 3-hosszú irányított kör szomszédsági mátrixa, de az első két sor cseréje a kívánt alakra hozza. Az első két oszlopot is kicserélve viszont már nem az $\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{C} \end{bmatrix}$ alakot kapjuk!

Frobenius vette észre és bizonyította, hogy az irreducibilitás az a feltétel, melynek fennállása esetén a nemnegatív mátrixokra is kiterjeszthetők a 11.1. tétel állításai.

11.8. TÉTEL (PERRON–FROBENIUS-TÉTEL – ERŐS VÁLTOZAT). *Ha az* **A** *nemnegatív mátrix irreducibilis, és $r = \varrho(\mathbf{A})$ jelöli a spektrálsugarát, akkor*

1. $r > 0$,
2. *r sajátértéke* **A***-nak, melyhez tartozik pozitív sajátvektor,*
3. **A***-nak e pozitív sajátvektor skalárszorosain kívül nincs más nemnegatív sajátvektora,*
4. *r egyszeres sajátérték.*

BIZONYÍTÁS.

Feladatok

11.7. Melyik irreducibilis az alábbi mátrixok közül? Amelyik nem, azt melyik permutációs mátrix viszi $\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{C} \end{bmatrix}$ alakba? Amelyik irreducibilis, annak mennyi a spektrálsugara és Perron-vektora?

$$\mathbf{R}_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{R}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

11.8. Keressünk egy-egy permutációs mátrixot az

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

mátrixok mindegyikéhez, mely bizonyítja reducibilitásukat!

Megoldások

9.1. Ellenőriznünk kell, hogy
a) az egyenlőség fennáll,
b) az $\mathbf{U}$ és $\mathbf{V}$ ortogonális mátrixok, $\mathbf{\Sigma}$ diagonális,
c) az $\mathbf{u}_1 = (1/\sqrt{2}, 1/\sqrt{2})$ és a $\mathbf{v}_1 = (1/\sqrt{2}, 1/\sqrt{2}, 0)$ vektorokra $\mathbf{A}\mathbf{v}_1 = 2\mathbf{u}_1$.

Ezek mind nyilvánvalóak, vagyis az első felbontás szinguláris. Mivel $\mathrm{r}(\mathbf{A}) = 1$, ezért $\mathbf{U}$ első oszlopát és $\mathbf{V}^T$ első sorát, valamint $\mathbf{\Sigma}$ bal felső elemét meghagyva valóban a második alakot kapjuk.

9.2. A $\mathbf{B}^T\mathbf{B}$ mátrix karakterisztikus polinomja $-\lambda^3 + 4\lambda^2 - 3\lambda = -(\lambda - 3)(\lambda - 1)\lambda$, így sajátértékei 3, 1 és 0. A szinguláris értékek $\sqrt{3}$ és 1. A sajátvektorok számításához 3-dimenziós vektorokkal kell számolnunk. Talán jobban járunk, ha inkább a $\mathbf{B}\mathbf{B}^T$ mátrixszal próbálkozunk. Mivel $\mathbf{B}\mathbf{B}^T = \left[\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right]$, ezért a karakterisztikus polinom $\lambda^2 - 4\lambda + 3 = (\lambda - 3)(\lambda - 1)$. A szinguláris értékek tehát $\sqrt{3}$ és 1, összhangban az előbbi számítással. A hozzájuk tartozó egységnyi hosszú sajátvektorok most nem a $\mathbf{V}_1$, hanem $\mathbf{U}_1$ oszlopait adják: $\mathbf{u}_1 = (1/\sqrt{2}, 1/\sqrt{2})$, $\mathbf{u}_2 = (-1/\sqrt{2}, 1/\sqrt{2})$. A $\mathbf{v}_i = \mathbf{A}^T\mathbf{u}_i/\sigma_i$ képletet használva a $\mathbf{V}$ mátrix is meghatározható. Tehát

$$\mathbf{B} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ 2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \\ -1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}$$

a $\mathbf{B}$ mátrix szinguláris felbontása.

11.1. Sajátértékek: 10, 3, 3, jobb sajátvektor: $\mathbf{u} = (5, 9, 11)$, bal sajátvektor: $\mathbf{v} = (4, 2, 1)$, a két Perron-vektor: $\mathbf{p} = \frac{1}{25}(5, 9, 11)$, bal sajátvektor: $\mathbf{v} = \frac{1}{7}(4, 2, 1)$.

11.2. A spektrálsugár még nincs a sajátértékek közt, így Perron tétele miatt csak a 3 és a 4 lehet sajátérték.

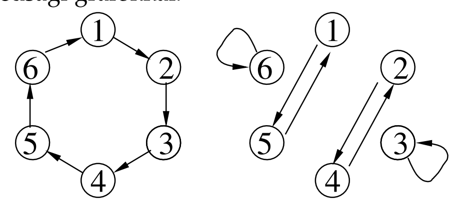
11.3. Ha $\mathbf{A}$ minden sorösszege c, akkor az $\mathbf{1}$ vektor sajátvektor, c sajátértékkel. Mivel $\mathbf{1} > \mathbf{0}$, ezért ez csak a Perron-vektor n-szerese lehet, és akkor c a hozzá tartozó sajátérték, így c a spektrálsugár. Hasonlóképp a bal Perron-vektor a másik állítást is igazolja.

11.4. Mivel a $\min\{5/4, 6/6, 7/5\} = 1$, ezért a Collatz–Wielandt-tétel szerint spektrálsugara is legalább ennyi. (Vagy a tételbeli másik képlettel: mivel a $c(4, 6, 5) \leq \mathbf{A} \cdot (4, 6, 5) = (5, 6, 7)$ egyenlőtlenségben c lehetséges maximuma 1, ezért a spektrálsugár legalább 1.)

11.5. Mivel $\mathbf{D}^{-1} = \mathrm{diag}(1/x_1, \ldots, 1/x_n)$, ezért követve az ötletben leírtakat, a 11.5. tétel első képlete a feladat első képletét adja. A $\mathbf{D}\mathbf{A}\mathbf{D}^{-1}$ mátrixból a második képletet kapjuk.

11.6. A 11.5. feladat első képlete az első, a második képlete a második mátrixról azt adja, hogy a minimum és a maximum is 10, így a spektrálsugár 10, tehát 10 a domináns sajátérték mindkét esetben, és $\mathbf{x}$ a hozzá tartozó sajátvektor – az első esetben a jobb, a másodikban a bal. (Gondoljuk meg!)

11.7. Az irreducibilitás eldönthető a mátrixokhoz rendelt szomszédsági gráfokkal:



$\mathbf{R}_1$ irreducibilis, mert a gráf erősen összefüggő, azaz bármely csúcsból bármely másikba el lehet jutni irányított úton. $\mathbf{R}_2$ reducibilis, hisz például nem indul irányított él a következő halmazokból a komplementerükbe: $\{6\}$, $\{3\}$, $\{1,5\}$, $\{2,4\}$, $\{1,5,6\}$, $\{2,3,4\}$,.... Így igen sok olyan $\mathbf{P}$ permutációs mátrix van, amelyik $\mathbf{R}_2$-t a kívánt alakba viszi. Közülük legegyszerűbb az identikus mátrix, hisz $\mathbf{R}_2$ már a kívánt alakú:

$$\mathbf{I}\mathbf{R}_2\mathbf{I}^T = \mathbf{R}_2 = \left[\begin{array}{ccccc|c} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right].$$

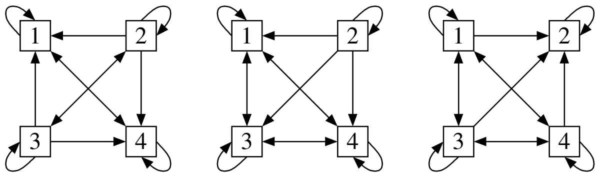
Az $\mathbf{R}_1$ mátrixnak nyilvánvalóan sajátvektora a $\mathbf{p} = (1, 1, 1, 1, 1, 1)$ vektor az 1 sajátértékkel. Mivel $\mathbf{R}_1$ nemnegatív és irreducibilis, ezért a Frobenius–Perron-tétel szerint a spektrálsugárhoz, mint sajátértékhez tartozó sajátvektor az egyetlen sajátvektor, mely pozitív elemű. Ebből következik, hogy a spektrálsugár 1.

Másik megoldás a feladat második részére:

$$\det(\mathbf{R}_1 - \lambda\mathbf{I}) = \begin{vmatrix} -\lambda & 1 & 0 & 0 & 0 & 0 \\ 0 & -\lambda & 1 & 0 & 0 & 0 \\ 0 & 0 & -\lambda & 1 & 0 & 0 \\ 0 & 0 & 0 & -\lambda & 1 & 0 \\ 0 & 0 & 0 & 0 & -\lambda & 1 \\ 1 & 0 & 0 & 0 & 0 & -\lambda \end{vmatrix} = \lambda^6 - 1$$

A karakterisztikus polinom gyökei a hatodik egységgyökök, melyek az 1-sugarú körön vannak, tehát 1 a spektrálsugár. A spektrálsugár valóban sajátérték, és a $\lambda = 1$-hez tartozó sajátvektor $\mathbf{p} = (1, 1, 1, 1, 1, 1)$.

11.8. A három mátrixhoz az alábbi gráfok tartoznak:



Ennek alapján az első gráfban az $\{1,4\}$, a másodikban az $\{1,3,4\}$, a harmadikban a $\{2\}$ halmazból nem érhető el a többi pont. A pontoknak egy olyan átsorszámozását keressük, melyben e pontok a többi után következnek, ugyanis általában, ha az $\{1,2,\dots,k,k+1,\dots,n\}$ csúcshalmazban az első k pontba nem fut él a $\{k+1,\dots,n\}$ halmazból, akkor a szomszédsági mátrix $\left[\begin{smallmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{O} & \mathbf{C}\end{smallmatrix}\right]$ alakú lesz. Az első gráfban például a 3-2-1-4 sorrend jó, hisz az $\{1,4\}$ halmaz elemei vannak hátul, amit a $\left(\begin{smallmatrix}1&2&3&4\\3&2&1&4\end{smallmatrix}\right)$ permutáció megvalósít:

$$\begin{bmatrix}0&0&1&0\\0&1&0&0\\1&0&0&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}1&0&0&1\\1&1&1&1\\1&1&1&1\\1&0&0&1\end{bmatrix}\begin{bmatrix}0&0&1&0\\0&1&0&0\\1&0&0&0\\0&0&0&1\end{bmatrix}=\left[\begin{array}{cc|cc}1&1&1&1\\1&1&1&1\\\hline 0&0&1&1\\0&0&1&1\end{array}\right]$$

A második esetben például jó a 2-1-3-4 sorrend, hisz az $\{1,3,4\}$ halmaz elemei vannak hátul, amit a $\left(\begin{smallmatrix}1&2&3&4\\2&1&3&4\end{smallmatrix}\right)$ permutáció megvalósít:

$$\begin{bmatrix}0&1&0&0\\1&0&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}1&0&1&1\\1&1&1&1\\1&0&1&1\\1&0&1&1\end{bmatrix}\begin{bmatrix}0&1&0&0\\1&0&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}=\left[\begin{array}{c|ccc}1&1&1&1\\\hline 0&1&1&1\\0&1&1&1\\0&1&1&1\end{array}\right]$$

Végül a harmadik mátrixnál jó az 1-4-3-2 sorrend, így a 2 elem van hátul, amit az $\left(\begin{smallmatrix}1&2&3&4\\1&4&3&2\end{smallmatrix}\right)$ permutáció megvalósít:

$$\begin{bmatrix}1&0&0&0\\0&0&0&1\\0&0&1&0\\0&1&0&0\end{bmatrix}\begin{bmatrix}1&1&1&1\\0&1&0&0\\1&1&1&1\\1&1&1&1\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&0&0&1\\0&0&1&0\\0&1&0&0\end{bmatrix}=\left[\begin{array}{ccc|c}1&1&1&1\\1&1&1&1\\1&1&1&1\\\hline 0&0&0&1\end{array}\right]$$

A balról szorzó permutációmátrix az egységmátrixból a megadott permutáció sorokra való alkalmazásával, míg a jobbról szorzó mátrix az oszlopokra való alkalmazásával lett meghatározva.

A
Függelék

Lebegőpontos számábrázolás

A lebegőpontos számábrázolás Tudjuk, hogy minden valós szám fölírható tizedestört alakban. Ha a szám racionális, azaz fölírható két egész szám hányadosaként, akkor a tizedes tört alak véges vagy végtelen periodikus. Az irracionális számok tizedes tört alakja mindig végtelen sok számjegyből áll. Véges tizedes tört alakkal – márpedig a számítógépek csak ilyenek kezelésére képesek – a végtelen tizedestörteket csak közelíteni tudjuk. A lebegőpontos számábrázolás alapötlete az, hogy egy tízhatvánnyal való szorzással minden szám „értékes jegyeit” a tizedes ponthoz képest azonos helyre „toljuk” (oda*lebeg*tetjük). A számok tudományos vagy exponenciális jelölésének normált alakjában ez a hely az első értékes (nem nulla) szám után van. Például $0.123 = 1.23 \cdot 10^{-1}$, $0.000321 = 3.21 \cdot 10^{-4}$, $123000 = 1.23 \cdot 10^{5}$, $.321 \cdot 10^{4} = 3.21 \cdot 10^{3}$.

A számrendszer alapja nem csak 10 lehet. Legyen b egy 1-nél nagyobb egész. Megmutatható, hogy minden valós c szám felírható b alapú számrendszerben, azaz

$$\pm a_0 a_1 \dots a_k . a_{k+1} \dots a_n \dots$$

alakban, ahol a_i egész, $a_0 \neq 0$ és $0 \leq a_i \leq b-1$. Ha $b = 2$, akkor a_i lehetséges értékei 0 és 1, ha $b = 10$, akkor a_i csak a $0, 1, \dots, 9$ jegyek valamelyike lehet. A félreértés elkerülésére a számrendszer alapját a szám után alsó indexbe lehet írni. Pl. 101101_2 egy kettes, míg 101101_3 egy hármas számrendszerbeli számot jelöl. A c szám b alapú fenti felírása azt jelenti, hogy

$$c = \pm \left(a_0 b^k + a_1 b^{k-1} + \cdots + a_k b^0 + a_{k+1} \frac{1}{b} + \cdots + a_{k+n} \frac{1}{b^n} + \dots \right)$$

Például a kettes számrendszerbeli 1001.101_2 szám értéke

$$1001.101_2 = 1 \cdot 2^3 + 1 + 1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2^3} = 8 + 1 + \frac{1}{2} + \frac{1}{8} = 9\frac{5}{8} = 9.625.$$

1.1. DEFINÍCIÓ (LEBEGŐPONTOS SZÁMOK). *A b alapú, p-jegyű (vagy p értékes jegyű, p pontosságú)* lebegőpontos számok *közé tartozik a* 0, *valamint a*

$$(-1)^s a_0.a_1a_2\dots a_{p-1}\cdot b^e,$$

alakú számok, ahol s értéke 0 vagy 1 aszerint, hogy pozitív vagy negatív számról van szó, $a_0 \neq 0$, $0 \le a_i \le b-1$ minden $i = 0, 1, \dots, p-1$ indexre és e egész szám. Egy ilyen alakú szám értéke

$$(-1)^s\left(a_0 + a_1\frac{1}{b} + a_2\frac{1}{b^2} + \dots + a_{p-1}\frac{1}{b^{p-1}}\right)\cdot b^e.$$

A reprezentáció teljes leírásához b és p értéke mellett ismernünk kell e maximális és minimális értékét (e_{max}, e_{min}).

1.2. PÉLDA (LEBEGŐPONTOS SZÁMOK ÉRTÉKE). *Adjuk meg a lebegőpontos bináris* $1.001\cdot 2^1$, $1.11\cdot 2^3$, $1.1\cdot 2^{-3}$ *számok értékét tízes számrendszerben:*

MEGOLDÁS. $1.001_2\cdot 2^1 = 10.01_2 = 2+\frac14 = 2.25$, $1.11_2\cdot 2^3 = 1110_2 = 14$, $1.1_2\cdot 2^{-3} = 0.0011_2 = \frac18+\frac1{16} = \frac3{16} = 0.1875$. □

1.3. PÉLDA (LEBEGŐPONTOS SZÁMOK HALMAZA). *Mely számok ábrázolhatók a $b = 2$, $p = 3$, $-1 \le e \le 1$ paraméterű lebegőpontos számokkal?*

MEGOLDÁS. Az adott paraméterű lebegőpontos számok halmaza – kihasználva, hogy a definíció szerint a_0 nem lehet 0, tehát értéke csak 1 lehet – a következő:

$$\left\{\pm\left(1+\frac{a_1}{2}+\frac{a_2}{2^2}\right)\cdot 2^e : a_1, a_2 \text{ értéke 0 vagy 1, } -1\le e\le 1\right\}\cup\{0\}.$$

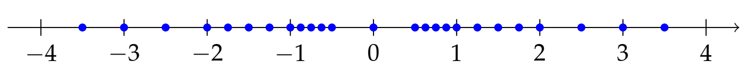
Tehát a halmaz pozitív elemei – jelezve e értékét is:

$$\underbrace{\frac12,\frac58,\frac34,\frac78}_{e=-1},\underbrace{1,\frac54,\frac32,\frac74}_{e=0},\underbrace{2,\frac52,3,\frac72}_{e=1}.$$

E számok tizedestört alakban fölírva:

$$.5, .625, .75, .875, 1, 1.25, 1.5, 1.75, 2, 2.5, 3, 3.5.$$

A számok eloszlását mutatja az A.1. ábra. □



A.1. ábra: Lebegőpontos számok $b = 2$, $p = 3$, $-1 \le e \le 1$ esetén.

A lebegőpontos és a valós számok között az fl függvény létesít kapcsolatot (fl a floating point kifejezésből). Legyen $x = a_0.a_1a_2\dots a_{p-1}a_p\dots\cdot 10^e$ egy tetszőleges pozitív valós szám. Az x szám p értékes jegyre való közelítése

$$\mathrm{fl}(x) = \begin{cases} a_0.a_1a_2\dots a_{p-1}\cdot 10^e & \text{ha } 0\le a_p < 5,\\ \left(a_0.a_1a_2\dots a_{p-1}+\frac{1}{10^{p-1}}\right)\cdot 10^e & \text{ha } 5 \le a_p \le 9,\end{cases}$$

vagyis ha az a_p 5-nél kisebb, lefelé, egyébként fölfelé kerekítünk. Az $x \mapsto \mathrm{fl}(x)$ leképezést *kerekítésnek* is nevezik. A valósok kerekítési szabálya más, 10-től különböző alap esetén is hasonlóan definiálható. A számítógépek a kerekítésnek más módját is ismerik és használják.

A modern számítógépek nagy részében a processzor az IEEE 754 szabvány szerint kezeli a lebegőpontos számokat (ld. a széljegyzetet).

IEEE 754 a lebegőpontos aritmetika nemzetközi szabványa, melynek számtalan hardver- és szoftverimplementációja van. 1985-ös változata még csak bináris számábrázolást definiált, 2008-as változata már a decimális számábrázolást is. Az IEEE 754-2008 öt alapformátumot ismer: 32, 64 és 128 bites bináris (más néven single, double, quad), valamint 64 és 128 bites decimális (double, quad) ábrázolást. Ezeken kívül ábrázolja a -0-t (pl. a 0-hoz való balról tartás jelzésére), szubnormális (azaz bizonyos, az alulcsordulás tartományába eső számokat), a két végtelent ($+\infty$, $-\infty$), és két nem-számot (NaN = not a number). A szabvány előírja több aritmetikai művelet, ötféle kerekítési szabály és öt kivétel kezelés megvalósítását (érvénytelen művelet, pl. negatív számból gyökvonás, 0-val osztás, túlcsordulás, alulcsordulás, nem egzakt eredmény jelzése).

Műveletek lebegőpontos számokkal Jelölje $\circ$ a négy alapművelet valamelyikét, és $\mathcal{F}$ a lebegőpontos számok egy halmazát. A lebegőpontos számok közti műveletek eredménye négy csoportba sorolható:

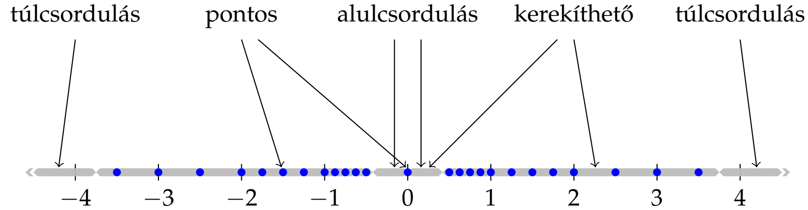
1. $x \circ y \in \mathcal{F}$, azaz az eredmény *pontos*;
2. $x \circ y$ nem $\mathcal{F}$-beli, de $\mathcal{F}$-belire *kerekíthető*;
3. $|x \circ y|$ nagyobb minden $\mathcal{F}$-belire kerekíthető számnál: ezt nevezzük *túlcsordulásnak*;
4. $|x \circ y|$ kisebb minden $\mathcal{F}$-belire kerekíthető pozitív számnál: ezt nevezzük *alulcsordulásnak*;

Amikor a számítógép kiírja a számokat tízes számrendszerbeli kerekítést végez, ha viszont kettes számrendszerben számol, az eredményt kettes számrendszerben kerekíti. Erre mutatunk egy leegyszerűsített példát.

1.4. PÉLDA (ALAPMŰVELETEK LEBEGŐPONTOS SZÁMOKKAL). *A fenti négy eset mindegyikére keressünk példát az* 1.3.. *példabeli lebegőpontos ábrázolás számaival. Vigyázzunk, itt nem a tízes számrendszerben kell kerekíteni!*

MEGOLDÁS. Pontos az eredménye például az $1 + \frac{3}{2} = \frac{7}{2}$ vagy az $1 \cdot \frac{3}{2} = \frac{3}{2}$ műveleteknek, kerekíthető az eredménye a $\frac{7}{8} \cdot \frac{3}{4} = \frac{21}{32} = .65625 \approx .625 = \frac{5}{8}$, vagy $\frac{3}{2} : \frac{7}{4} = \frac{3}{2} \cdot \frac{4}{7} = \frac{6}{7} \approx .85714 \approx .875 = \frac{7}{8}$. A kerekítéseket úgy végezzük, hogy mindig a legközelebb eső lebegőpontos számra kerekítünk. Túlcsordul például a $\frac{7}{2} + \frac{5}{2} = 6$ és a $-\frac{7}{2} \cdot \frac{5}{2} = -\frac{35}{4}$ művelet, míg alulcsordul a $-\frac{1}{2} \cdot \frac{1}{2} = -\frac{1}{4}$ művelet. Ez utóbbi 0-ra kerekíthető e számábrázolásban. □

Szemléltetésül ábrázoljuk az előbbi példa számtartományait:



A.2. ábra: A kerekíthető, a pontos, az alul- és túlcsorduló számhalmazok lebegőpontos számábrázolás esetén $b = 2$, $t = 3$, $-1 \le e \le 1$ paraméterek mellett. A 0 pontos érték, az alulcsorduló számok – feladattól függően – kerekíthetők 0-ra.

Mint láttuk, általában

$$\mathrm{fl}(x+y) \neq \mathrm{fl}(x) + \mathrm{fl}(y), \text{ és } \mathrm{fl}(xy) \neq \mathrm{fl}(x)\,\mathrm{fl}(y).$$

A lebegőpontos aritmetikában több azonosság sem áll fönn, például az asszociativitás, azaz ha x, y és z egy lebegőpontos számábrázolás számai, általában

$$\mathrm{fl}(\mathrm{fl}(x+y)+z) \neq \mathrm{fl}(x+\mathrm{fl}(y+z)).$$

Egy valódi példa látható a széljegyzetben.

```
> a=1; b=c=1.5*10^(-16);
> a+(b+c)==(a+b)+c
ans = 0
```

A.3. kód: Ha $a = 1$, $b = c = 1.5 \cdot 10^{-16}$, akkor a processzor aritmetika szerint $a + (b+c) \neq (a+b)+c$.

Algoritmusok műveletigénye: flop és flops Szokás az algoritmusok műveletigényét *flopban* mérni (flop = **fl**oating point **op**eration), ahol 1 flop egy lebegőpontos alapműveletet jelent ($+$, $-$, $*$ vagy $/$). A processzorok sebességét gyakran az egy másodperc alatt elvégzett flopok számával flops-ban mérik (itt a *flops* a **fl**oating point **op**eration per **s**econd rövidítése). Mivel a szorzás (osztás) processzorideje bizonyos gépeken hosszabb lehet az összeadásénál (kivonásénál), ezért szokás külön számolni e műveletek számát.

Ha egy algoritmus műveletigénye a feladat méretének egy polinomja, általában elég a legmagasabb fokú tagot figyelni, ugyanis egy algoritmus műveletigénye csak nagyobb méret esetén kezd érdekes lenni, a méret növekedtével pedig az alacsonyabb fokú tagok részesedése a polinom értékéből elhanyagolható. Precízen fogalmazva ennek oka, hogy tetszőleges k-nál kisebb fokú $p(x)$ polinomra

$$\lim_{n\to\infty} \frac{n^k + p(n)}{n^k} = 1.$$

1.5. PÉLDA (FLOP ÉS FLOPS). *Egy algoritmus műveletigénye n méretű bemenet esetén $n^3 + 1000n$, egy másiké $n^3 - 100n^2$ flop. Milyen gyorsan futnak le e programok egy 10 gigaflops sebességű gépen, ha $n = 10000$?*

MEGOLDÁS. $10\text{gigaflops} = 10 \cdot 10^9\text{flops}$, $n^3 + 1000n = (10^4)^3 + 1000(10^4) = 10^{12} + 10^7$, $n^3 - 100n^2 = (10^4)^3 - 100(10^4)^2 = 10^{12} - 10^{10}$, vagyis a két futási idő $(10^{12} + 10^7)/10^{10} = 100.001$, illetve $(10^{12} - 10^{10})/10^{10} = 99$ másodperc. □

Feladatok

1.1. Fogalmazzuk meg a 2-es számrendszerre vonatkozó kerekítési szabályt! Számítsuk ki az 1.4.. példa megoldásában szereplő műveleteket a kettes számrendszerben és kerekítsük az eredményeket e szabály szerint.

1.2. Az alábbi két interaktív programkód kiszámítja a $\sum_{n=1}^{9999} \frac{1}{n^2+n}$ összeget két különböző sorrendben, de a két eredmény különbözik, és csak a második jó. Vajon miért? (Elemi függvénytanból tudjuk, hogy a fenti összeg ún. teleszkópösszeg, mert átalakítás után a szélső tagokat kivéve minden tag kiesik: $\sum_{n=1}^{9999} \frac{1}{n^2+n} = \sum_{n=1}^{9999} \left(\frac{1}{n} - \frac{1}{n+1} \right) = 1 - \frac{1}{10000}$.)

```
> format long
> s1 = 0;
> for n = 1:9999
> s1 = s1 + (1/(n^2+n));
> end
> s1
s1 =  0.999900000000001
> s2 = 0;
> for n = 9999:-1:1
> s2 = s2 + (1/(n^2+n));
> end
> s2
s2 =  0.999900000000000
```

Komplex számok

Testek, gyűrűk

Az algebrai absztrakció lényege, hogy ha egy állítás igazolásához csak bizonyos műveleti tulajdonságokat használunk, akkor az állítás minden olyan struktúrában is igaz lesz, amely rendelkezik e műveleti tulajdonságokkal. Így született a racionális, valós, és komplex számok, valamint a prímmel való osztás maradékaival való számolás közös tulajdonságaiból az algebrai test fogalma.

1.6. DEFINÍCIÓ (TEST). *Egy legalább kételemű* $\mathbb{T}$ *halmazt* testnek *vagy* algebrai testnek *nevezünk, ha*

1. *értelmezve van* $\mathbb{T}$ *elempárjain egy összeadás és egy szorzás nevű* bináris művelet,
2. *az* összeadás *kommutatív, asszociatív, létezik nullelem és minden elemnek létezik ellentettje (additív inverze),*
3. *a* szorzás *kommutatív, asszociatív, létezik egységelem és a nullelemen kívül minden elemnek létezik multiplikatív inverze (reciproka),*
4. *az összeadás a szorzásra nézve disztributív.*

Formalizálva a fenti definíciót, egy test a következő tulajdonságokkal rendelkezik:

a) Bármely $a, b \in \mathbb{T}$ elemre $a + b \in \mathbb{T}$ ($\mathbb{T}$ zárt az összeadásra).
b) Bármely $a, b \in \mathbb{T}$ elemre $a + b = b + a$ (kommutatív).
c) Bármely $a, b, c \in \mathbb{T}$ elemre $(a + b) + c = a + (b + c)$ (asszociatív).
d) Van olyan $\mathbb{T}$-beli elem, jelölje 0, hogy bármely $a \in \mathbb{T}$-re $0 + a = a$ (zéruselem).
e) Bármely $a \in \mathbb{T}$ elemhez van olyan $b \in \mathbb{T}$, hogy $a + b = 0$ (additív inverz).
f) Bármely $a, b \in \mathbb{T}$ elemre $ab \in \mathbb{T}$ ($\mathbb{T}$ zárt az szorzásra).
g) Bármely $a, b \in \mathbb{T}$ elemre $ab = ba$ (kommutatív).
h) Bármely $a, b, c \in \mathbb{T}$ elemre $(ab)c = a(bc)$ (asszociatív).
i) Van olyan $\mathbb{T}$-beli elem, jelölje 1, hogy bármely $a \in \mathbb{T}$ elemre $1a = a$ (egységelem).
j) Bármely $a \in \mathbb{T} \setminus \{0\}$ (tehát bármely nem nulla) elemhez van olyan $b \in \mathbb{T}$, hogy $ab = 1$ (multiplikatív inverz).
k) Bármely $a, b, c \in \mathbb{T}$ elemre $(a + b)c = ac + bc$ (disztributivitás).

Az összeadás tulajdonságai az *a)-e)*, a szorzáséi az *f)-j)* pontban vannak, *k)* a két művelet közös tulajdonsága.

Néhány megjegyzés és néhány példa:

- Meg lehet mutatni, hogy a nullelem és az egységelem szükségképpen különböző, tehát jogosan jelöltük őket különböző jellel.
- Minden test tetszőleges a elemére igaz, hogy $0a = a0 = 0$.
- Testre példa a valós számok $\mathbb{R}$, a racionális számok $\mathbb{Q}$, a komplex

számok $\mathbb{C}$, a prím modulusú maradékosztályok $\mathbb{Z}_p$ struktúrája, ami egy véges, p-elemű test (szokásos jelölései még: $\mathbb{F}_p$, $\mathrm{GF}(p)$).

Bár leggyakrabban test elemeiből képzett determinánsokkal számolunk, látni fogjuk, hogy a determináns kiszámolásához nincs szükség testre, gyengébb struktúra is elég.

A test axiómái közül néhány elhagyásával gyakran használt és fontos struktúrákhoz jutunk. Ha a szorzás csak asszociatív, *gyűrűről* beszélünk, ha kommutatív is, *kommutatív gyűrűről*, ha az asszociativitás mellett van egységeleme is, *egységelemes gyűrűről* beszélünk.

Ha a testet definiáló kikötések közül elhagyjuk a *j)*-t, egységelemes gyűrűt kapunk, ha az *i)*-t is, kommutatív gyűrűt, ha a *g)*-t is, gyűrűt kapunk.

► Minden gyűrűben is igaz, hogy a nullelem és az egységelem szükségképpen különböző, és hogy bármely a elemre $0a = a0 = 0$.
► A természetes számok $\mathbb{N}$ halmaza nem gyűrű a szokásos műveletekkel, mert nincs minden elemnek ellentettje (a -1 nem természetes szám).
► A $\mathbb{Z}$ egységelemes kommutatív gyűrű, de nem test, mert nincs minden nemnulla elemnek multiplikatív inverze (az $1/5$ nem egész).
► A $\mathbb{Z}_m$ egységelemes kommutatív gyűrű, és pontosan akkor test, ha m prím.
► A valós együtthatós polinomok egységelemes kommutatív gyűrűt alkotnak.
► Az $\mathbb{R} \to \mathbb{R}$ függvények egységelemes kommutatív gyűrűt alkotnak. Ugyanígy egységelemes kommutatív gyűrűt alkotnak a $[0, 1]$ intervallumon értelmezett függvények, vagy a $[0, 1]$ intervallumon értelmezett és folytonos függvények. A műveletek minden esetben a függvények szokásos összeadása és szorzása.
► A végtelen sorozatok az elemenkénti összeadás és szorzás műveletére nézve egységelemes kommutatív gyűrűt alkotnak.

A valós számokhoz hasonlóan egy $\mathbb{T}$ test elemeiből képzett rendezett n-esek halmazán, azaz a $\mathbb{T}$ test feletti n-dimenziós vektorok $\mathbb{T}^n$ halmazán bevezethetünk két műveletet, a vektorösszeadás és a $\mathbb{T}$ elemeivel való szorzás műveletét. Az így kapott struktúrára, mint a $\mathbb{T}^n$ vektortérre fogunk hivatkozni.

A $\mathbb{Z}$ gyűrű elemeiből képzett n-dimenziós vektorok $\mathbb{Z}^n$ halmazából a vektorok összeadásának és $\mathbb{Z}$-beli skalárral való szorzásának művelete nem vezet ki. E struktúra neve *modulus*.

Feladatok

Testek, gyűrűk

1.3. PÉLDA EGY TESTRE Jelölje $\mathbb{Q}(\sqrt{2})$ az összes $a + b\sqrt{2}$ alakú számok halmazát, ahol a és b racionális számok, azaz $a, b \in \mathbb{Q}$. Mutassuk meg, hogy $\mathbb{Q}(\sqrt{2})$ a szokásos összeadás és szorzás műveletekkel testet alkot.

1.4.* NÉGYZETES MÁTRIXOK GYŰRŰJE Mutassuk meg, hogy ha $\mathbb{T}$ test, akkor $\mathbb{T}^{n\times n}$ elemei, azaz a $\mathbb{T}$ fölötti $n \times n$-es mátrixok a szokásos mátrixműveletekkel egységelemes gyűrűt alkotnak.

Prímelemű testek

Aritmetika véges halmazon Olyan műveleteket keresünk a véges ábécé elemei közt, melyek eredménye is az ábécébe esik. Több ilyen példát is ismertetünk:

1.7. PÉLDA (MŰVELETEK PARITÁSOKKA). *Tudjuk, hogy egész számokkal végzett összeadás és szorzás eredményének paritása csak az összeadandók, illetve szorzandók paritásától függ (művelettábláit lásd az A.4. a) ábrán). Például egy páros és egy páratlan szám összege mindig páratlan.*

Az előző példabelihez hasonló művelettáblákat kapunk két logikai művelettel is, melyeket az informatikában gyakran használunk. Egyik a logikai „és” művelete (AND), a másik a „kizáró vagy” (XOR = eXclusive OR).

1.8. PÉLDA (XOR ÉS AND). *A p és q – igaz vagy hamis logikai értékű – állítást az „és” szóval összekapcsolva olyan állítást kapunk, mely pontosan akkor igaz, ha p és q is igaz. A „kizáró vagy” művelete, melyet legtöbbször a „vagy... vagy...” kötőszavakkal fejezünk ki pontosan akkor ad igaz eredményt, ha p és q közül csak az egyik igaz.*

a)

+	prs	ptl
prs	prs	ptl
ptl	ptl	prs

×	prs	ptl
prs	prs	prs
ptl	prs	ptl

b)

XOR	h	i
h	h	i
i	i	h

AND	h	i
h	h	h
i	h	i

c)

+	0	1
0	0	1
1	1	0

×	0	1
0	0	0
1	0	1

A.4. ábra: Műveletek kételemű ábécéken: *a)* egészek közti műveletek paritása (prs = páros, ptl = páratlan), *b)* a logikai AND és XOR műveletek (h = hamis, i = igaz), *c)* e műveletek a $\{0, 1\}$ ábécén.

Az „és” műveletet szokás logikai szorzásnak is nevezni, hisz ha a hamis állítás logikai értékét 0, az igazét 1 jelöli, akkor e művelet eredménye megegyezik e két számmal végzett szorzás eredményével (A.4. *(b)* és *(c)* pont). Például a „ma este moziba megyünk” és a „beülünk valahová vacsorázni” állítások összekapcsolásával kapott „ma este moziba megyünk, *és* beülünk valahová vacsorázni” állítás pontosan akkor igaz, ha moziban is megyünk és vacsorázni is.

A „ma este *vagy* moziba megyünk, *vagy* beülünk valahová vacsorázni” állítás pontosan akkor igaz, ha az egyik helyre elmegyünk, de mindkettőre nem. E két művelet művelettábláit lásd az A.4. *(b)* ábrán. Ha a páros számokhoz és a hamis állításhoz a 0 számot, a páratlan számokhoz és az igaz állításhoz az 1-et rendeljük, akkor mindkét előző táblából az A.4. *(c)* ábrán látható művelettáblát kapjuk.

A számok paritása úgy is kifejezhető, hogy egy egész szám páros, ha 2-vel való osztási maradéka 0, és páratlan, ha osztási maradéka 1, vagyis amikor paritásokkal számolunk, mondhatjuk, hogy a 2-vel való osztási maradékokkal számolunk. A maradékokkal való hasonló számolással a hétköznapokban másutt is találkozunk, pl. az időpont kiszámításakor, de ott nem 2-vel, hanem 12-vel, 24-gyel vagy 60-nal kell osztani.

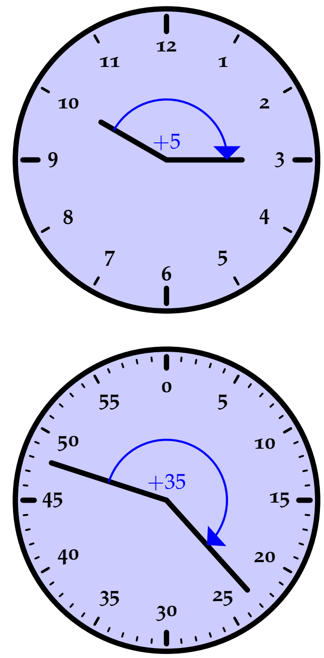
1.9. PÉLDA (SZÁMOLÁS AZ ÓRÁN). *Ha most 10 óra van, 5 óra múlva 3 óra lesz, azaz 10-hez 5-öt adva 3-at kapunk, ha órákban számolunk. Hasonlóképp 21 órához 4-et adva 1-et kapunk. Ha a percmutató most 48 percet* *mutat, akkor 35 perc múlva 23-at fog.*

Ha $m > 1$ egész szám és a tetszőleges egész, akkor a maradékos osztás tétele szerint egyetlen olyan q és r egész szám létezik, melyekre

$$a = mq + r,$$

és ahol $0 \leq r < m$. Ezt az r számot az a szám m-mel való *osztási maradékának* nevezzük, és az $a \bmod m$ kifejezéssel jelöljük. Kimondva: „a modulo m”. Például $20 \bmod 3 = 2$, hisz $20 = 6 \cdot 3 + 2$. Hasonlóképp kapjuk, hogy $-12 \bmod 10 = 8$, ugyanis $-12 = -2 \cdot 10 + 8$. Azokat az egészeket, amelyek ugyanazt a maradékot adják modulo m, egy osztályba soroljuk. Ezeket az osztályokat *maradékosztályoknak* nevezzük. Például a 13, 33, 103 ugyanabba a maradékosztályba tartoznak modulo 10.

Könnyen megmutatható, hogy bármely két egész szám összegének, különbségének és szorzatának m-mel való osztási maradéka csak a két szám osztási maradékától függ, a két szám nagyságától nem. Például a $11 + 5 = 16$, $2 + 2 = 4$, $-1 + 8 = 7$ összadások mindegyikében két olyan számot adtunk össze, melyek 3-mal osztva 2-t adnak maradékul, az eredmény pedig mindig 1-et.



A.5. ábra: Számolás az órán

Modulo: a mérték jelentésű latin modulus szó származéka. Itt a maradékosztályok számát jelenti.

1.10. DEFINÍCIÓ ($\mathbb{Z}_m$). *Legyen $m > 1$ egész szám, és jelölje $\mathbb{Z}_m$ a $\{0, 1, \ldots, m-1\}$ halmazt, melyen két műveletet definiálunk. Ha $a, b \in \mathbb{Z}_m$, akkor jelölje $a + b$, illetve ab azt a $\mathbb{Z}_m$-beli elemet, melyet a és b egész számként való összegének, illetve szorzatának m-mel való osztási maradékaként kapunk.*

Mivel sem az összeadás sem a szorzás műveletén, sem a számokon nem látszik, hogy egész számok közti műveletről, vagy valamilyen m-re $\mathbb{Z}_m$-beli műveletről van-e szó, ezért a fenti definícióbeli műveletek csak úgy egyértelműek, ha az elemekről tudjuk, hogy mely $\mathbb{Z}_m$ elemei. Pl. $\mathbb{Z}_2$-ben $1 + 1 = 0$, de $\mathbb{Z}_3$-ban $1 + 1 = 2$ (ld. az A.6. ábrát).

$+$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$\times$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

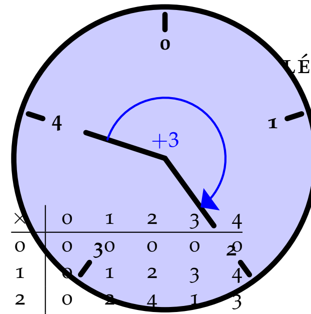
A.6. ábra: $\mathbb{Z}_3$ művelettáblái

1.11. PÉLDA (SZÁMOLÁS $\mathbb{Z}_m$-BEN). *Számítsuk ki $\mathbb{Z}_{60}$-ban a $48 + 35$, $\mathbb{Z}_5$-ben a $4 + 3$ és $\mathbb{Z}_{11}$-ben a $3 \cdot 7 + 10$ kifejezések értékét!*

MEGOLDÁS. Ha $\mathbb{Z}_m$ művelettáblái rendelkezésre állnak, használhatjuk, egyébként úgy számolhatunk, hogy egészeknek tekintjük a számokat, és vesszük az eredmények maradékát modulo m:

$$48 + 35 \bmod 60 = 83 \bmod 60 = 23,$$
$$4 + 3 \bmod 5 = 7 \bmod 5 = 2,$$
$$3 \cdot 7 + 10 \bmod 11 = 31 \bmod 11 = 9.$$

$\mathbb{Z}_m$ művelettábláival, vagy pl. $\mathbb{Z}_{12}$ és $\mathbb{Z}_{60}$ esetén egy óra mutatójával is számolhatunk, de más m-re is használhatunk „órát”: $\mathbb{Z}_{60}$-ban: $48 + 35 = 23$, $\mathbb{Z}_5$-ben: $4 + 3 = 2$ és $\mathbb{Z}_{11}$-ben: $3 \cdot 7 + 10 = 10 + 10 = 9$. □



A.7. ábra: „Óra” a $\mathbb{Z}_5$-beli számoláshoz

$\times$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

$\times$	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

A.8. ábra: $\mathbb{Z}_5$ és $\mathbb{Z}_6$ szorzástáblája

1.12. PÉLDA (MŰVELETTÁBLA). *Készítsünk művelettáblát a $\mathbb{Z}_5$- és a $\mathbb{Z}_6$-beli szorzáshoz!*

MEGOLDÁS. A művelettáblák elkészítése az 5-tel, illetve 6-tal vett maradékokkal számolva könnyen megkapható. Pl. $3 \cdot 4 \bmod 5 = 12 \bmod 5 = 2$, tehát $3 \cdot 4 = 2$ a $\mathbb{Z}_5$-ben. □

A két művelettábla közt két fontos különbség vehető észre. Az egyik, hogy míg $\mathbb{Z}_5$-nél az első sort és oszlopot kivéve minden sorban és oszlopban minden szám pontosan egyszer szerepel, addig $\mathbb{Z}_6$-nál nem. Ez azt jelenti, hogy $\mathbb{Z}_5$ szorzásművelete a 0 elemet elhagyva invertálható, azaz bármely két $a, b \in \mathbb{Z}_5 \setminus 0$ elemre az $ax = b$ egyenlet megoldható. Más szóval $\mathbb{Z}_5$-ben van osztás, és minden nemzérus elemnek van reciproka, azaz multiplikatív inverze. Épp úgy, mint $\mathbb{R}$-ben. Az is látható, hogy $\mathbb{Z}_6$-ban két nemzérus érték szorzata lehet 0 is.

1.13. PÉLDA (OSZTÁS, RECIPROK). *Oldjuk meg a $3 \cdot x = 2$ egyenletet és határozzuk meg a 3 reciprokát $\mathbb{Z}_5$-ben és $\mathbb{Z}_6$-ban!*

MEGOLDÁS. Mivel $\mathbb{Z}_5$-ben $3 \cdot 4 = 2$, és $3 \cdot 2 = 1$, ezért az egyenlet megoldása $x = 2$, és 3 reciproka 2, azaz $1/3 = 2$. $\mathbb{Z}_6$-ban az egyenletnek nincs megoldása, és 3-nak nincs reciproka. □

Megmutatható, hogy $\mathbb{Z}_m$-ben a szorzás pontosan akkor invertálható, ha m prím. Ezesetben $\mathbb{Z}_m$-et $\mathbb{F}_m$-mel is szokás jelölni annak kihangsúlyozására, hogy itt az osztás művelete is elvégezhető.

A valósok $\mathbb{R}$, a racionálisok $\mathbb{Q}$ és az imént tárgyalt $\mathbb{F}_p$ (p prím) struktúrák a következő algebrai tulajdonságokkal rendelkeznek: Az összeadás és a szorzás is kommutatív, asszociatív művelet, tetszőleges a elemre $1a = a$ és $0a = 0$, az összeadás invertálható művelet, a szorzás invertálható a 0 elhagyása mellett, és végül a szorzás disztributív az összeadásra nézve. Az ilyen algebrai struktúrákat (algebrai) *testnek* nevezik. Vannak egyéb testek is, de ebben a könyvben a fent felsoroltakon kívül csak a komplex számok $\mathbb{C}$-vel jelölt testével fogunk foglalkozni.

Polinomok

B

Lineáris algebra dióhéjban

Ebben a fejezetben egybe gyűjtünk olyan eredményeket, melyek a könyvben más-más fejezetekben elszórva szerepelnek.

2.1. TÉTEL (MÁTRIX RANGJA). *Legyen* $\mathbf{A}$ *egy* $m \times n$*-es mátrix, és jelölje* A *a hozzá tartozó* $\mathbf{x} \mapsto \mathbf{Ax}$ *leképezést. A következő számok mind megegyeznek.*

a) $\mathrm{r}(\mathbf{A})$,

b) *a főelemek száma az* $\mathbf{A}$ *bármelyik lépcsős alakjában,*

c) *a nemnulla sorok száma az* $\mathbf{A}$ *bármelyik lépcsős alakjában,*

d) *az* $\mathbf{A}$ *bázisoszlopainak (főoszlopainak) száma,*

e) *az* $\mathbf{A}$ *oszlopaiból kiválasztható maximális lineárisan független vektorrendszer elemszáma,*

f) *az* $\mathbf{A}$ *soraiból kiválasztható maximális lineárisan független vektorrendszer elemszáma,*

g) *az* $\mathbf{A}$ *oszlopterének dimenziója, azaz* $\dim(\mathcal{O}(\mathbf{A}))$,

h) *az* $\mathbf{A}$ *sorterének dimenziója, azaz* $\dim(\mathcal{S}(\mathbf{A}))$,

i) $n - \dim(\mathcal{N}(\mathbf{A}))$,

j) $m - \dim(\mathcal{N}(\mathbf{A}^T))$,

k) *az* A *leképezés képterének dimenziója, azaz* $\dim(\mathrm{Im}(A))$,

l) $n - \dim(\mathrm{Ker}(A))$,

m) *az* $\mathbf{A}$*-ból kiválasztható legnagyobb méretű nemnulla értékű aldetermináns rendje,*

n) *az* $\mathbf{A}$*-ból kiválasztható legnagyobb méretű nemszinguláris mátrix mérete,*

o) *az* $\mathbf{A}$ *(pozitív) szinguláris értékeinek száma.*

2.2. TÉTEL (INVERTÁLHATÓ NÉGYZETES MÁTRIXOK). *Legyen* $\mathbf{A}$ *egy* $n \times n$*-es mátrix, és jelölje* A *a hozzá tartozó* $\mathbf{x} \mapsto \mathbf{Ax}$ *leképezést. A következő állítások ekvivalensek:*

a) *Az* $\mathbf{A}$ *mátrix redukált lépcsős alakja* $\mathbf{I}_n$*, azaz* $\operatorname{rref}(\mathbf{A}) = \mathbf{I}_n$.
b) $\mathbf{A}$ *előáll elemi mátrixok szorzataként.*
c) $\mathrm{r}(\mathbf{A}) = n$.
d) $N(\mathbf{A}) = 0$.
e) *Az* $\mathbf{Ax} = \mathbf{b}$ *egyenletrendszer minden* $\mathbf{b} \in \mathbb{R}^n$ *vektorra megoldható.*
f) *Az* $\mathbf{Ax} = \mathbf{b}$ *egyenletrendszer minden* $\mathbf{b} \in \mathbb{R}^n$ *vektorra egyértelműen megoldható.*
g) *Az* $\mathbf{Ax} = \mathbf{0}$ *egyenletrendszernek a triviális az egyetlen megoldása.*
h) $\mathbf{A}$ *sorvektorai lineárisan függetlenek.*
i) $\mathbf{A}$ *oszlopvektorai lineárisan függetlenek.*
j) $\mathbf{A}$ *sorvektorai kifeszítik* $\mathbb{R}^n$*-et.*
k) $\mathbf{A}$ *oszlopvektorai kifeszítik* $\mathbb{R}^n$*-et.*
l) $\mathbf{A}$ *sorvektorai* $\mathbb{R}^n$ *bázisát alkotják.*
m) $\mathbf{A}$ *oszlopvektorai* $\mathbb{R}^n$ *bázisát alkotják.*
n) *Az* $\mathbf{A}$ *mátrix invertálható.*
o) $\det(\mathbf{A}) \neq 0$.
p) *A* 0 *nem sajátértéke* $\mathbf{A}$*-nak.*
q) *A* 0 *nem szinguláris értéke* $\mathbf{A}$*-nak.*
r) *Az* $\mathbf{A}$ *legkevesebb tagból álló diadikus felbontásának* n *tagja van.*
s) *Az* A *lineáris leképezés képtere* $\mathbb{R}^n$.
t) *Az* A *lineáris leképezés magtere* $\{\mathbf{0}\}$.
u) *Az* A *leképezés kölcsönösen egyértelmű.*
v) *Az* A *leképezés invertálható.*

Irodalomjegyzék

Wolf Holzmann. Uniqueness of reduced row echelon form. http://www.cs.uleth.ca/~holzmann/notes/reduceduniq.pdf, 2002.

Tárgymutató