LINEÁRIS ALGEBRA II.

matematika BSc, 2023. őszi félév

Waldhauser Tamás

A 4.3, 4.4, 5.1, 5.2, F.2, F.3 fejezetek a 2022 őszi kurzus alábbi hallgatóinak jegyzetei alapján készültek:
Csíkos Enikő, Domonics Dávid, Galántai-Fekete Csenge Lilla,
Gálfi Dávid, Glavosits Villő, Kollár Csenge, Pintér Sándor

Tartalomjegyzék

1. Gauss-elimináció (ismétlés)

1.1. Definíció. Azt mondjuk, hogy egy T test feletti mátrix redukált lépcsős alakú, ha

- minden sorban az első nemnulla elem egyes (főelem vagy vezéregyes);
- minden sorban a vezéregyes hátrébb (jobbra) helyezkedik el, mint az előző sor vezéregyese;
- a vezéregyesek oszlopaiban az összes többi elem nulla;
- ha vannak csupa nullákból álló sorok, akkor azok a mátrix alján vannak.

1.2. Példa. Íme egy redukált lépcsős alakú mátrix (a csillagok helyén bármi állhat):

$$\begin{pmatrix} \boxed{1} & * & 0 & 0 & * & * & 0 & * & * & * \\ 0 & 0 & \boxed{1} & 0 & * & * & 0 & * & * & * \\ 0 & 0 & 0 & \boxed{1} & * & * & 0 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

1.3. Megjegyzés. Ha egy redukált lépcsős alakú mátrixban az esetlegesen csupa nullákból álló sorokat elhagyjuk, és a vezéregyeseket tartalmazó oszlopokat egymás mellé írjuk, akkor egységmátrixot kapunk. Az 1.1. Definíció tartalmaz még megszorításokat a többi oszlopra is, de amikor redukált lépcsős alakot fogunk használni, akkor mindig csak arra lesz szükségünk, hogy van néhány oszlop, amelyek egységmátrixot alkotnak. A fenti példában ezt így szemléltethetjük:

$$\begin{pmatrix} \boxed{1} & * & 0 & 0 & * & * & 0 & * & * & * \\ 0 & * & \boxed{1} & 0 & * & * & 0 & * & * & * \\ 0 & * & 0 & \boxed{1} & * & * & 0 & * & * & * \\ 0 & * & 0 & 0 & * & * & \boxed{1} & * & * & * \end{pmatrix}.$$

1.4. Definíció. Egy T test feletti mátrix elemi sorátalakításain az alábbi átalakításokat értjük:

(1) egy sorhoz hozzáadjuk egy másik sor λ-szorosát, ahol $\lambda \in T$;

(2) egy sort megszorzunk λ-val, ahol $\lambda \in T \setminus \{0\}$;

(3) két sort felcserélünk.

Hasonlóan definiálhatóak az elemi oszlopátalakítások is.

1.5. Állítás. Az A mátrixon elemi sorátalakításokat végrehajtva, az $A\mathbf{x} = \mathbf{0}$ egyenletrendszer megoldáshalmaza nem változik.

1.6. Tétel. Minden T test feletti mátrix elemi sorátalakítássokkal redukált lépcsős alakra hozható.

2. Vektorterek

2.1. A vektortér fogalma, nevezetes példák

2.1. Definíció. Azt mondjuk, hogy a nemüres V halmaz vektortér a T test felett, ha értelmezettek rajta az alábbi műveletek:

$$\text{összeadás: } V \times V \to V,\ (\mathbf{u}, \mathbf{v}) \mapsto \mathbf{u} + \mathbf{v};$$

$$\lambda\text{-val való szorzás:} \qquad V \to V, \qquad \mathbf{v} \mapsto \lambda\mathbf{v} \qquad (\text{minden egyes } \lambda \in T \text{ elemre});$$

és ezek a műveletek rendelkeznek az alábbi tulajdonságokkal:

(1) $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V : (\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$;

(2) $\exists \mathbf{0} \in V\ \forall \mathbf{v} \in V : \mathbf{v} + \mathbf{0} = \mathbf{v} = \mathbf{0} + \mathbf{v}$;

(3) $\forall \mathbf{v} \in V\ \exists \mathbf{u} \in V : \mathbf{v} + \mathbf{u} = \mathbf{v} = \mathbf{u} + \mathbf{v}$;

(4) $\forall \mathbf{u}, \mathbf{v} \in V : \mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;

(5) $\forall \lambda \in T\ \forall \mathbf{u}, \mathbf{v} \in V : \lambda(\mathbf{u} + \mathbf{v}) = \lambda\mathbf{u} + \lambda\mathbf{v}$;

(6) $\forall \lambda, \mu \in T\ \forall \mathbf{v} \in V : (\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}$;

(7) $\forall \lambda, \mu \in T\ \forall \mathbf{v} \in V : (\lambda\mu)\mathbf{v} = \lambda(\mu\mathbf{v})$;

(8) $\forall \mathbf{v} \in V : 1\mathbf{v} = \mathbf{v}$.

2.2. Megjegyzés. A V halmaz elemeit vektoroknak, a T test elemeit skalároknak nevezzük. Az első négy tulajdonság szerint $(V; +)$ Abel-csoport, melynek egységeleme $\mathbf{0}$ (nullvektor). A harmadik tulajdonságbeli $\mathbf{u}$ vektor a $\mathbf{v}$ vektor additív inverze, és ennek szokásos jelölése $\mathbf{u} = -\mathbf{v}$.

2.3. Állítás. Ha V vektortér a T test felett, akkor tetszőleges $\mathbf{v} \in V$ és $\lambda \in T$ esetén

(i) $\lambda\mathbf{v} = \mathbf{0} \iff \lambda = 0$ vagy $\mathbf{v} = \mathbf{0}$;

(ii) $(-\lambda)\mathbf{v} = \lambda(-\mathbf{v}) = -\lambda\mathbf{v}$.

Bizonyítás. Lásd [SzL] 6.3. Tétel. □

2.4. Megjegyzés. A fenti állítás második részében három „mínusz” jel szerepel. Ezek közül kettő V-beli additív inverzet jelöl, egy pedig T-beli additív inverzet (melyik melyik?).

2.5. Példa. A $V = \{$☂, 🚗, 🎂, 🌲$\}$ halmaz vektorteret alkot a $\mathbb{Z}_2 = \{0, 1\}$ test felett az alábbi műveletekkel:

+	☂	🚗	🎂	🌲
☂	☂	🚗	🎂	🌲
🚗	🚗	☂	🌲	🎂
🎂	🎂	🌲	☂	🚗
🌲	🌲	🎂	🚗	☂

$0 \cdot$ ☂ $=$ ☂	$1 \cdot$ ☂ $=$ ☂
$0 \cdot$ 🚗 $=$ ☂	$1 \cdot$ 🚗 $=$ 🚗
$0 \cdot$ 🎂 $=$ ☂	$1 \cdot$ 🎂 $=$ 🎂
$0 \cdot$ 🌲 $=$ ☂	$1 \cdot$ 🌲 $=$ 🌲

2.6. Példa. Az alábbi halmazok vektorteret alkotnak a valós számtest felett a szokásos műveletekkel:

(i) a síkbeli, illetve térbeli vektorok halmaza;

(ii) a komplex számok halmaza;

(iii) az $n \times m$ méretű valós mátrixok halmaza;

(iv) a legfeljebb n-edfokú valós polinomok halmaza;

(v) a valós sorozatok halmaza;

(vi) a valós függvények halmaza.

2.7. Példa. Az alábbi halmazok vektorteret alkotnak a T test felett a szokásos műveletekkel:

(i) T^n, azaz a T feletti elem n-esek halmaza (speciális esetek: $\mathbb{R}^2$ és $\mathbb{R}^3$);

(ii) T bármely testbővítése (speciális esetek: $\mathbb{C}$ vektortér $T = \mathbb{R}$ felett, $\mathbb{R}$ vektortér $T = \mathbb{Q}$ felett, az algebrai számok halmaza vektortér $T = \mathbb{Q}$ felett);

(iii) $T^{n \times m}$, azaz a T feletti $n \times m$ méretű mátrixok halmaza;

(iv) a T feletti legfeljebb n-edfokú polinomok halmaza;

(v) $T^{\mathbb{N}}$, azaz a T elemeiből álló sorozatok halmaza;

(vi) az összes $X \to T$ leképezések halmaza (tetszőleges nemüres X halmaz esetén).

2.2. Altér, lineáris kombináció, generálás

2.8. Definíció. Legyen V vektortér a T test felett. Az $U \subseteq V$ nemüres részhalmazt V <u>alterének</u> nevezzük (jelölés: $U \leq V$), ha zárt az összeadásra és a skalárokkal való szorzásokra:

(1) $\forall \mathbf{u}_1, \mathbf{u}_2 \in U \colon \mathbf{u}_1 + \mathbf{u}_2 \in U$;

(2) $\forall \lambda \in T \; \forall \mathbf{u} \in U \colon \lambda \mathbf{u} \in U$.

2.9. Megjegyzés. Ha $U \subseteq V$ a fenti definíció értelmében altere V-nek, akkor U maga is vektorteret alkot a V-beli műveletek megszorításaival, ezért jogosan nevezzük altérnek.

2.10. Példa. Tetszőleges V vektortérben altér $\{\mathbf{0}\}$ (triviális altér) és persze maga V is altér (nem valódi altér).

2.11. Példa. A $V = \mathbb{Z}_2^3$ vektorérben altér az alábbi U halmaz (a teljesség kedvéért felsoroljuk V elemeit is):

$$V = \{000, 001, 010, 011, 100, 101, 110, 111\};$$
$$U = \{000, 011, 101, 110\}.$$

Hogy U valóban zárt az összeadásra, azt közvetlen számolással lehet ellenőrizni (a skalárok pedig $\mathbb{Z}_2$ esetén nem sok vizet zavarnak). Egyszerűbb (és tanulságosabb!) azonban, ha kitaláljuk, hogy a V halmazból milyen tulajdonságú elemeket vettünk be az U halmazba (HF).

2.12. Példa. Néhány altér a 2.6. Példában látott valós vektorterekben:

(i) minden origón átmenő egyenes altér $\mathbb{R}^2$-ben, minden origón átmenő egyenes és sík altér $\mathbb{R}^3$-ban;

(ii) a valós tengely és a képzetes tengely is altér $\mathbb{C}$-ben;

(iii) a szimmetrikus mátrixok alteret alkotnak $\mathbb{R}^{n\times n}$-ben;

(iv) azon polinomok, amelyeknek gyöke $\pi + \sqrt{2}$ alteret alkotnak a legfeljebb n-edfokú valós polinomok vektorterében;

(v) a konvergens sorozatok alteret alkotnak a valós sorozatok vektorterében;

(vi) a folytonos függvények alteret alkotnak a valós függvények vektorterében.

2.13. Definíció. Legyen V vektortér a T test felett, legyen $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ és $\lambda_1, \ldots, \lambda_k \in T$. Ekkor a

$$\mathbf{w} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k$$

vektort a $\mathbf{v}_1, \ldots, \mathbf{v}_k$ vektorok ($\lambda_1, \ldots, \lambda_k$ együtthatókkal vett) lineáris kombinációjának nevezzük.

2.14. Definíció. Legyen V vektortér a T test felett, és legyen $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$. A $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ vektorok összes lineáris kombinációinak halmazát a $\mathbf{v}_1, \ldots, \mathbf{v}_k$ vektorok által kifeszített altérnek (vagy generált altérnek) nevezzük. Erre az altérre a $[\mathbf{v}_1, \ldots, \mathbf{v}_k]$ jelölést használjuk:

$$[\mathbf{v}_1, \ldots, \mathbf{v}_k] = \big\{ \lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k : \lambda_1, \ldots, \lambda_k \in T \big\}.$$

2.15. Állítás. A $[\mathbf{v}_1, \ldots, \mathbf{v}_k]$ halmaz valóban altere V-nek, mégpedig ez a legszűkebb olyan altér, ami tartalmazza a $\mathbf{v}_1, \ldots, \mathbf{v}_k$ vektorok mindegyikét.

Bizonyítás. Lásd [SzL] 6.10. Tétel. □

2.16. Megjegyzés. Nemcsak véges, hanem végtelen sok vektor generátuma is értelmezhető, és a fenti állítás is érvényben marad. Arra kell csak figyelni, hogy ha végtelen sok vektorunk van is, egy lineáris kombinációban akkor is csak véges sokat (de akármilyen sokat) használhatunk közülük. Tehát tetszőleges $X \subseteq V$ halmaz generátuma így fest:

$$[X] = \big\{ \lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k : k \in \mathbb{N}_0,\ \mathbf{v}_1, \ldots, \mathbf{v}_k \in V,\ \lambda_1, \ldots, \lambda_k \in T \big\}.$$

2.3. Alterek kétféle megadási módja

2.17. Példa. A valós függvények vektorterében az $f'' = -f$ differenciálegyenletet kielégítő függvények alteret alkotnak; ez könnyen ellenőrizhető a jól ismert deriválási szabályok segítségével. Kevésbé triviális, hogy ezt az alteret a $\sin x$ és $\cos x$ függvény generálja:

$$\big\{ f \in \mathbb{R}^{\mathbb{R}} : f'' = -f \big\} = [\sin x, \cos x] = \big\{ \lambda \sin x + \mu \cos x : \lambda, \mu \in \mathbb{R} \big\}.$$

Kétféleképpen adhatunk meg egy vektortérben egy alteret: vagy megmondjuk, hogy milyen tulajdonságú vektorok tartoznak az altérbe (ahogy a 2.12. Példában is tettük), vagy megadjuk egy generátorrendszerét.[1] A fenti példa mutatja, hogy nem mindig egyszerű az „átváltás” egyik megadási módról a másikra. Jó hír: ha a T^n vektortérben dolgozunk, akkor az alteret definiáló tulajdonságo(ka)t mindig le lehet írni egy homogén lineáris egyenletrendszerrel, és a Gauss-elimináció segítségével könnyen válthatunk a „tulajdonságos” (azaz „egyenletrendszeres”) felírás és a „generátorrendszeres” felírás között.

2.18. Tétel. Tetszőleges $A \in T^{k\times n}$ mátrix esetén az $A\mathbf{x} = \mathbf{0}$ homogén lineáris egyenletrendszer megoldásai alteret alkotnak T^n-ben.

Bizonyítás. Lásd [SzL] 5.11. Tétel □

2.19. feladat. Tekintsük az $\mathbb{R}^5$ vektortérben az alábbi W alteret:

$$W = \big\{ (x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5 : x_1 - x_4 = 0,\ x_1 - x_2 - x_3 = 0,\ 3x_1 - x_5 = 0 \big\}.$$

Adjunk meg bázist (generátorrendszert) a W altérben.

Megoldás. Felírjuk az egyenletrendszer mátrixát (mivel homogén, a konstansok oszlopát nem kell kiírni), és

[1] Ha véges halmazról van szó, akkor van egy harmadik lehetőség is: egyszerűen felsoroljuk az altér elemeit, ahogy a 2.11. Példában is tettük.

elemi sorátalakításokkal redukált lépcsős alakra hozzuk:

$$\begin{pmatrix} 1 & 0 & 0 & -1 & 0 \\ 1 & -1 & -1 & 0 & 0 \\ 3 & 0 & 0 & 0 & -1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} \boxed{1} & 0 & 0 & 0 & -1/3 \\ 0 & \boxed{1} & 1 & 0 & -1/3 \\ 0 & 0 & 0 & \boxed{1} & -1/3 \end{pmatrix}.$$

A redukált lépcsős alakból látszik, hogy x_1, x_2, x_4 lehet kötött változó, x_3, x_5 pedig szabad változók:

$$x_1 = \tfrac{1}{3}x_5, \quad x_2 = -x_3 + \tfrac{1}{3}x_5, \quad x_4 = \tfrac{1}{3}x_5 \quad (x_3, x_5 \in \mathbb{R}).$$

Tehát a W altér paraméteres felírása (az $x_3 = a$, $x_5 = b$ paraméterekkel):

$$W = \left\{ \left(\tfrac{1}{3}b,\ -a + \tfrac{1}{3}b,\ a,\ \tfrac{1}{3}b,\ b\right) : a, b \in \mathbb{R} \right\}.$$

A paraméterek egyikének 1 értéket adunk, a többinek nullát, minden lehetséges módon. Így kapunk egy bázist W-ben:

$$\mathbf{w}_1 = (0,\ -1,\ 1,\ 0,\ 0), \quad \mathbf{w}_2 = \left(\tfrac{1}{3},\ \tfrac{1}{3},\ 0,\ \tfrac{1}{3},\ 1\right).$$

Megjegyzés: Az, hogy az elimináció során nem keletkezett csupa nulla sor, azt jelenti, hogy a feladatban megadott három egyenlet független; kevesebb egyenlettel nem lehet W-t leírni. ◇

2.20. Definíció. Legyen V vektortér a T test felett. Egy V-beli vektorrendszer elemi átalakításain az alábbi átalakításokat értjük:

(1) egyik vektorhoz hozzáadjuk egy másik vektor λ-szorosát, ahol $\lambda \in T$;

(2) egy vektort megszorzunk λ-val, ahol $\lambda \in T \setminus \{0\}$;

(3) két vektort felcserélünk.

Hasonlóan definiálhatóak az elemi oszlopátalakítások is. (Az 1.4. Definícióval való hasonlóság nem a véletlen műve.)

2.21. Tétel. Vektorrendszer elemi átalakításai nem változtatják meg a vektorrendszer által generált alteret.

Bizonyítás. Csak az első típusú átalakításra igazoljuk az állítást, a többi (szinte) triviális. A jelölések egyszerűsítése végett (és az általánosság megszorítása nélkül) tfh. a második vektorhoz adjuk hozzá az első vektor λ-szorosát:

$$\mathbf{v}_1,\ \mathbf{v}_2,\ \mathbf{v}_3, \ldots, \mathbf{v}_k \quad \rightsquigarrow \quad \mathbf{v}_1,\ \mathbf{v}_2 + \lambda\mathbf{v}_1,\ \mathbf{v}_3, \ldots, \mathbf{v}_k.$$

Az világos, hogy $[\mathbf{v}_1,\ \mathbf{v}_2,\ \mathbf{v}_3, \ldots, \mathbf{v}_k] \supseteq [\mathbf{v}_1,\ \mathbf{v}_2 + \lambda\mathbf{v}_1,\ \mathbf{v}_3, \ldots, \mathbf{v}_k]$ (ugye?). A másik irányú tartalmazáshoz azt kell megmutatnunk, hogy a „régi” vektorok mind előállnak az „új” vektorok lineáris kombinációiként. Elég csak a $\mathbf{v}_2$ vektorral foglalkozni, mert a többi vektor egyaránt szerepel a régi és az új vektorrendszerben is. A $\mathbf{v}_2$ vektor pedig így kombinálható ki az új vektorokból (csak az első két új vektort használjuk): $\mathbf{v}_2 = (\mathbf{v}_2 + \lambda\mathbf{v}_1) - \lambda\mathbf{v}_1$. □

2.22. feladat. Tekintsük az $\mathbb{R}^5$ vektortérben az alábbi három vektort:

$$\mathbf{v}_1 = (1, 4, -3, 1, 3), \quad \mathbf{v}_2 = (-2, 1, 0, 1, 0), \quad \mathbf{v}_3 = (3, -2, -1, 1, 1).$$

Írjuk le a $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ vektorok által kifeszített U alteret, vagyis adjuk meg, hogy milyen összefüggéseknek kell teljesülniük az x_1, x_2, x_3, x_4, x_5 számok között ahhoz, hogy az $(x_1, x_2, x_3, x_4, x_5)$ vektor előálljon $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ lineáris kombinációjaként. Másképp fogalmazva: adjunk meg olyan homogén lineáris egyenletrendszert, melynek megoldástere éppen U.

Megoldás. Írjuk egymás alá a három vektort, és hozzuk a kapott mátrixot redukált lépcsős alakra elemi sorátalakításokkal:

$$\begin{pmatrix} 1 & 4 & -3 & 1 & 3 \\ -2 & 1 & 0 & 1 & 0 \\ 3 & -2 & -1 & 1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} \boxed{1} & 0 & 0 & -1 & 0 \\ 0 & \boxed{1} & 0 & -1 & 0 \\ 0 & 0 & \boxed{1} & -2 & -1 \end{pmatrix}.$$

A 2.21. Tétel garantálja, hogy a kapott mátrix sorai szintén az U alteret feszítik ki. Jelölje a redukált lépcsős alakú mátrix sorait $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$:

$$\mathbf{u}_1 = (1, 0, 0, -1, 0), \quad \mathbf{u}_2 = (0, 1, 0, -1, 0), \quad \mathbf{u}_3 = (0, 0, 1, -2, -1).$$

Ezekkel a vektorokkal így írható fel az U altér:

$$\begin{aligned} U = [\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3] &= \left\{ \lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2 + \lambda_3\mathbf{u}_3 : \lambda_1, \lambda_2, \lambda_3 \in T \right\} \\ &= \left\{ (\lambda_1,\ \lambda_2,\ \lambda_3,\ -\lambda_1 - \lambda_2 - 2\lambda_3,\ -\lambda_3) : \lambda_1, \lambda_2, \lambda_3 \in T \right\}. \end{aligned}$$

Tehát U elemeinél az első három koordináta szabadon megválasztható, az utolsó két koordináta pedig egyértelműen kifejezhető az első háromból:

$$U = \{(x_1, x_2, x_3, x_4, x_5) : x_1, x_2, x_3 \in T,\ x_4 = -x_1 - x_2 - 2x_3,\ x_5 = -x_3\}.$$

Ez pedig azt jelenti, hogy U az alábbi homogén lineáris egyenletrendszer megoldáshalmaza:

$$\begin{pmatrix} 1 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Megjegyzés: A redukált lépcsős alakból világos, hogy $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ lineárisan független vektorrendszer, és így bázisa U-nak. Ebből következik, hogy $\dim U = 3$, és a feladatban megadott $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ generátorrendszer is bázis. (Másképp fogalmazva: abból, hogy az elimináció során nem keletkezett csupa nulla sor, következik, hogy $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ lineárisan független vektorrendszer.) ◇

2.23. Tétel. Tetszőleges T test esetén T^n minden altere előáll egy n-ismeretlenes homogén lineáris egyenletrendszer megoldástereként.

Bizonyítás. Legyen $U \le T^n$ és legyen $\mathbf{v}_1, \ldots, \mathbf{v}_k$ generátorrendszere U-nak. (Később megmutatjuk majd, hogy T^n minden altere véges dimenziós, és így végesen generált.) Írjuk sorvektorként egymás alá a $\mathbf{v}_1, \ldots, \mathbf{v}_k$ vektorokat, így egy $A \in T^{k \times n}$ mátrixot kapunk. Hozzuk elemi sorátalakításokkal redukált lépcsős alakra az A mátrixot (lásd az 1.6 Tételt), és hagyjuk el a csupa nulla sorokat (ha vannak). A 2.21. Tétel szerint így egy olyan $A' \in T^{r \times n}$ mátrixot kapunk, melynek sorai szintén az U alteret feszítik ki. (Mikor fordulhat elő, hogy $r < k$?) A jelölések egyszerűsítése végett (és az általánosság megszorítása nélkül) tfh. így fest az A' mátrix:

$$A' = \begin{pmatrix} \boxed{1} & 0 & 0 & a & b \\ 0 & \boxed{1} & 0 & c & d \\ 0 & 0 & \boxed{1} & e & f \end{pmatrix}.$$

Ekkor tehát U nem más, mint az A' mátrix sorai, vagyis az $\mathbf{u}_1 = (1, 0, 0, a, b)$, $\mathbf{u}_2 = (0, 1, 0, c, d)$, $\mathbf{u}_3 = (0, 0, 1, e, f)$ vektorok által generált altér:

$$\begin{aligned} U = [\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3] &= \{\lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \lambda_3 \mathbf{u}_3 : \lambda_1, \lambda_2, \lambda_3 \in T\} \\ &= \{(\lambda_1,\ \lambda_2,\ \lambda_3,\ a\lambda_1 + c\lambda_2 + e\lambda_3,\ b\lambda_1 + d\lambda_2 + f\lambda_3) : \lambda_1, \lambda_2, \lambda_3 \in T\}. \end{aligned}$$

Tehát U elemeinél az első három koordináta szabadon megválasztható, az utolsó két koordináta pedig egyértelműen kifejezhető az első háromból:

$$U = \{(x_1, x_2, x_3, x_4, x_5) : x_1, x_2, x_3 \in T,\ x_4 = ax_1 + cx_2 + ex_3,\ x_5 = bx_1 + dx_2 + fx_3\}.$$

Ez pedig azt jelenti, hogy U az alábbi homogén lineáris egyenletrendszer megoldáshalmaza:

$$\begin{pmatrix} -a & -c & -e & 1 & 0 \\ -b & -d & -f & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Tetszőleges redukált lépcsős alakú mátrixhoz hasonló módon lehet megkonstruálni olyan homogén lineáris egyenletrendszert, melynek megoldástere éppen a mátrix sorvektorai által kifeszített altér. □

2.4. Alterek metszete és összege

2.24. Állítás. Alterek metszete is altér.

Bizonyítás. Lásd [SzL] 6.7. Tétel. □

2.25. Megjegyzés. Nemcsak kettő, hanem több (akár végtelen sok!) altér metszete is altér. A bizonyítás igazából nem nehezebb, mint két altér esetén, csak macerásabb leírni.

2.26. Definíció. Az $U_1, U_2 \leq V$ alterek összege az alábbi altér:

$$U_1 + U_2 = \{\mathbf{u}_1 + \mathbf{u}_2 : \mathbf{u}_1 \in U_1,\ \mathbf{u}_2 \in U_2\}.$$

2.27. Állítás. Alterek összege valóban altér, mégpedig $U_1 + U_2$ a legszűkebb olyan altér, ami tartalmazza U_1-et és U_2-t is:

$$U_1 + U_2 = [U_1 \cup U_2].$$

Bizonyítás. Lásd [SzL] 6.12. Tétel. □

2.28. Következmény. Tetszőleges V vektortér összes alterei hálót alkotnak a tartalmazásra nézve. Az altérhálóbeli metszés és egyesítés:

$$U_1 \wedge U_2 = U_1 \cap U_2, \qquad U_1 \vee U_2 = U_1 + U_2.$$

2.29. feladat. Tekintsük a 2.19. és 2.22. feladatokban megadott W és U altereket $\mathbb{R}^5$-ben, és határozzuk meg az összegüket és a metszetüket. Mindkettőt adjuk meg bázissal és egyenletrendszer megoldástereként is.

Megoldás. A W altérnek egy bázisa $\mathbf{w}_1, 3\mathbf{w}_2$ (a 2.19. feladat megoldásában számoltuk ki ezeket a vektorokat; $\mathbf{w}_2$ helyett inkább $3\mathbf{w}_2$-t használjuk, hogy ne legyenek törtek). Az U altérnek egy bázisa $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ (a 2.22. feladat megoldásában számoltuk ki ezeket a vektorokat; vehetnénk a feladatban megadott $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ bázist is, de $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ szebb). Ha a két bázist egymás mellé (vagy inkább alá) tesszük, akkor máris megvan $W + U$ egy generátorrendszere. Hogy kiderüljön, lehet-e ötnél kevesebb vektorral generálni a $W + U$ alteret, illetve, hogy megkapjuk az „egyenletrendszeres" leírását, hozzuk a mátrixot redukált lépcsős alakra:

$$\begin{pmatrix} 0 & -1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 3 \\ 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & -2 & -1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Ebből látszik, hogy $W + U$ négydimenziós, a kapott mátrix első négy sora ad egy bázist, valamint – a 2.22. feladat megoldásához hasonlóan – kiolvasható a redukált lépcsős alakból $W + U$ „egyenletrendszeres" leírása is:

$$\begin{aligned} W + U &= [(1,0,0,0,1),\ (0,1,0,0,1),\ (0,0,1,0,1),\ (0,0,0,1,1)] \\ &= \{(x_1,x_2,x_3,x_4,x_5) \in \mathbb{R}^5 : x_1 + x_2 + x_3 + x_4 - x_5 = 0\}. \end{aligned}$$

A $W \cap U$ altérbe azok a vektorok tartoznak, amelyek kielégítik W és U definiáló egyenleteit is. Tehát itt az altereket leíró egyenleteket kell összerakni. A W altér eleve egyenletrendszerrel volt megadva, de inkább a 2.19. feladat megoldásában kiszámolt redukált lépcsős alakú egyenletrendszert vesszük. Ez alá írjuk le az U alteret leíró egyenletrendszert, amit a 2.22. feladat megoldásában kaptunk, majd redukált lépcsős alakra hozzuk a kapott mátrixot:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & -1/3 \\ 0 & 1 & 1 & 0 & -1/3 \\ 0 & 0 & 0 & 1 & -1/3 \\ 1 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 & -1/3 \\ 0 & 1 & 0 & 0 & -4/3 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1/3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Tehát nem kell öt egyenlet, elég négy is (a mátrix első négy sora adja az egyenletekbeli együtthatókat), és – a 2.19. feladat megoldásához hasonlóan – kapunk egy bázist is (beszoroztunk mindent 3-mal, hogy eltűnjenek a törtek):

$$\begin{aligned} W \cap U &= \{(x_1,x_2,x_3,x_4,x_5) \in \mathbb{R}^5 : 3x_1 - x_5 = 0,\ 3x_2 - 4x_5 = 0,\ x_3 + x_5 = 0,\ 3x_4 - x_5 = 0\} \\ &= [(1,4,-3,1,3)]. \end{aligned}$$

◇

2.30. feladat. Tekintsük most a 2.19. és a 2.22. feladatbeli W, U altereket $\mathbb{Z}_2$ felett, és határozzuk meg az összegüket és metszetüket. Mind a négy alteret adjuk meg bázissal és egyenletrendszerrel is.

Megoldás. A 2.19. feladatbeli mátrixot most $\mathbb{Z}_2$ felett tekintjük, és redukált lépcsős alakra hozzuk:

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} \boxed{1} & 0 & 0 & 0 & 1 \\ 0 & \boxed{1} & 1 & 0 & 1 \\ 0 & 0 & 0 & \boxed{1} & 1 \end{pmatrix}.$$

Ebből kiolvashatjuk W kétféle leírását:

$$\begin{aligned} W &= \left\{(x_1,x_2,x_3,x_4,x_5) \in \mathbb{Z}_2^5 : x_1+x_5=0,\ x_2+x_3+x_5=0,\ x_4+x_5=0\right\} \\ &= \left[(0,1,1,0,0),\ (1,1,0,1,1)\right]. \end{aligned}$$

A 2.22. feladatbeli mátrixot is redukált lépcsős alakra hozzuk $\mathbb{Z}_2$ felett:

$$\begin{pmatrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} \boxed{1} & 0 & 1 & 1 & 1 \\ 0 & \boxed{1} & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Ebből kiolvashatjuk U kétféle leírását:

$$\begin{aligned} U &= \left[(1,0,1,1,1),\ (0,1,0,1,0)\right] \\ &= \left\{(x_1,x_2,x_3,x_4,x_5) \in \mathbb{Z}_2^5 : x_1+x_3=0,\ x_1+x_2+x_4=0,\ x_1+x_5=0\right\}. \end{aligned}$$

Írjuk egymás alá W és U bázisát, és hozzuk a mátrixot redukált lépcsős alakra:

$$\begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} \boxed{1} & 0 & 0 & 0 & 1 \\ 0 & \boxed{1} & 0 & 1 & 0 \\ 0 & 0 & \boxed{1} & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Ebből kiolvashatjuk $W+U$ kétféle leírását:

$$\begin{aligned} W+U &= \left[(1,0,0,0,1),\ (0,1,0,1,0),\ (0,0,1,1,0)\right] \\ &= \left\{(x_1,x_2,x_3,x_4,x_5) \in \mathbb{Z}_2^5 : x_2+x_3+x_4=0,\ x_1+x_5=0\right\}. \end{aligned}$$

Írjuk egymás alá W és U egyenleteit, és hozzuk a mátrixot redukált lépcsős alakra:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} \boxed{1} & 0 & 0 & 0 & 1 \\ 0 & \boxed{1} & 0 & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 & 1 \\ 0 & 0 & 0 & \boxed{1} & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Ebből kiolvashatjuk $W \cap U$ kétféle leírását:

$$\begin{aligned} W \cap U &= \left\{(x_1,x_2,x_3,x_4,x_5) \in \mathbb{Z}_2^5 : x_1+x_5=0,\ x_2=0,\ x_3+x_5=0,\ x_4+x_5=0\right\} \\ &= \left[(1,0,1,1,1)\right]. \end{aligned}$$

Megjegyzések:

1. A W altérre ugyanazt kaptuk, mintha a 2.19. feladat végeredményét vettük volna modulo 2 (hogyan kell az $\frac{1}{3}$ törtet $\mathbb{Z}_2$-ben értelmezni?). Az U altérnél viszont nem ugyanazt kaptuk, mintha a 2.22. feladat végeredményét vettük volna modulo 2 (még a dimenzió se stimmel!). Mi lehet ennek az oka?

2. Tanulságos lehet felsorolni a négy altér elemeit, és ellenőrizni a metszetet meg az összeget:

$$\begin{aligned} W &= \{00000, 10111, 01100, 11011\}; \\ U &= \{00000, 10111, 01010, 11101\}; \\ W \cap U &= \{00000, 10111\}; \\ W+U &= \{00000, 10111, 01100, 11011, 01010, 11101, 00110, 10001\}. \end{aligned}$$

$\diamond$

2.5. Lineáris függetlenség

2.31. Definíció. Legyen V vektortér a T test felett és legyen $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$. Azt mondjuk, hogy a $\mathbf{v}_1, \ldots, \mathbf{v}_k$ vektorrendszer lineárisan független, ha a nullvektor csak triviális módon (csupa nulla együtthatókkal) kombinálható ki a $\mathbf{v}_1, \ldots, \mathbf{v}_k$ vektorokból:

$$\forall \lambda_1, \ldots, \lambda_k \in T \colon \lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k = \mathbf{0} \iff \lambda_1 = \cdots = \lambda_k = 0.$$

Ha a nullvektor nemtriviális lineáris kombinációval is megkapható, akkor azt mondjuk, hogy a vektorrendszer lineárisan függő.

2.32. Tétel. Legyen V vektortér a T test felett. Tetszőleges $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ vektorokra ekvivalensek az alábbiak:

(i) $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ lineárisan függő;

(ii) valamelyik vektor előáll a többiek lineáris kombinációjaként: $\exists i \colon \mathbf{v}_i \in [\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_k]$;

(iii) valamelyik vektor előáll a „korábbi" vektorok lineáris kombinációjaként: $\exists i \colon \mathbf{v}_i \in [\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}]$.

Bizonyítás. Lásd [SzL] 7.2. Tétel. □

2.33. Tétel. Legyen V vektortér a T test felett. Tetszőleges $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ vektorokra ekvivalensek az alábbiak:

(i) $\mathbf{v}_1, \ldots, \mathbf{v}_k$ lineárisan független;

(ii) minden $\mathbf{v} \in V$ vektor legfeljebb egyféleképpen áll elő a $\mathbf{v}_1, \ldots, \mathbf{v}_k$ vektorok lineáris kombinációjaként;

(iii) minden $\mathbf{v} \in [\mathbf{v}_1, \ldots, \mathbf{v}_k]$ vektor pontosan egyféleképpen áll elő a $\mathbf{v}_1, \ldots, \mathbf{v}_k$ vektorok lineáris kombinációjaként.

Bizonyítás. Az világos, hogy (ii) és (iii) ekvivalens, hiszen – a generátum definíciója szerint – $[\mathbf{v}_1, \ldots, \mathbf{v}_k]$ elemei előállnak $\mathbf{v}_1, \ldots, \mathbf{v}_k$ lineáris kombinációjaként, a többi vektor pedig nem áll elő. Ha (iii)-at a $\mathbf{v} = \mathbf{0}$ vektorra felírjuk, akkor látjuk, hogy (iii) $\Longrightarrow$ (i). Hátra van még (i) $\Longrightarrow$ (iii) igazolása. Ehhez tfh. $\mathbf{v}_1, \ldots, \mathbf{v}_k$ lineárisan független, és tekintsük egy $\mathbf{v} \in [\mathbf{v}_1, \ldots, \mathbf{v}_k]$ vektor két felírását a $\mathbf{v}_1, \ldots, \mathbf{v}_k$ vektorok lineáris kombinációjaként: $\mathbf{v} = \sum_i \lambda_i \mathbf{v}_i = \sum_i \mu_i \mathbf{v}_i$. Kivonva egymásból a kettőt, megkapjuk a nullvektort a $\mathbf{v}_1, \ldots, \mathbf{v}_k$ vektorok lineáris kombinációjaként: $\mathbf{0} = \sum_i (\lambda_i - \mu_i) \mathbf{v}_i$. Mivel feltettük, hogy $\mathbf{v}_1, \ldots, \mathbf{v}_k$ lineárisan független, itt minden együttható nulla, azaz $\lambda_i = \mu_i$ minden i-re. Ez pedig azt jelenti, hogy $\mathbf{v}$ tekintett két felírása szóról szóra (pontosabban: együtthatóról együtthatóra) megegyezik. □

2.6. Bázis, koordináták

2.34. Definíció. Egy vektortér bázisán lineárisan független generátorrendszert értünk.

2.35. Példa. Néhány nevezetes bázis néhány nevezetes vektortérben:

(i) A T^n vektortérben az alábbi $\mathbf{e}_1, \ldots, \mathbf{e}_n$ vektorok bázist alkotnak, ezt nevezzük standard bázisnak

$$\mathbf{e}_1 = (1, 0, 0, \ldots, 0, 0), \quad \mathbf{e}_2 = (0, 1, 0, \ldots, 0, 0), \; \ldots, \; \mathbf{e}_n = (0, 0, 0, \ldots, 0, 1).$$

(ii) A komplex számok $\mathbb{R}$ feletti vektorterében bázis $1, i$.

(iii) Legyen $E_{ij} \in T^{n \times m}$ az a mátrix, amelyben az i-edik sor j-edik eleme 1, az összes többi elem pedig nulla. A $T^{n \times m}$ vektortérben bázist alkotnak az E_{ij} $(i = 1, \ldots, n, \; j = 1, \ldots, m)$ mátrixok.

(iv) A T feletti legfeljebb n-edfokú polinomok vektorterében bázis $1, x, x^2, \ldots, x^n$.

2.36. Tétel. Legyen V vektortér a T test felett. Tetszőleges $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ vektorokra ekvivalensek az alábbiak:

(i) $\mathbf{v}_1, \ldots, \mathbf{v}_n$ bázisa V-nek;

(ii) V minden eleme pontosan egyféleképpen áll elő a $\mathbf{v}_1, \ldots, \mathbf{v}_n$ vektorok lineáris kombinációjaként.

Bizonyítás. Rögtön következik a 2.33. Tételből. □

2.37. Definíció. Legyen $\mathcal{B} \colon \mathbf{v}_1, \ldots, \mathbf{v}_n$ bázisa a T test feletti V vektortérnek, és legyen $\mathbf{v} \in V$ tetszőleges vektor. Ekkor léteznek olyan egyértelműen meghatározott $\lambda_1, \ldots, \lambda_n \in T$ skalárok, melyekre $\mathbf{v} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_n \mathbf{v}_n$ (lásd a 2.36. Tételt). A $\lambda_1, \ldots, \lambda_n$ skalárokat a $\mathbf{v}$ vektor $\mathcal{B}$ bázisra vonatkozó koordinátáinak nevezzük, a $\mathbf{v}$ vektor $\mathcal{B}$ bázisra vonatkozó koordinátasorán pedig a következőt értjük:

$$[\![\mathbf{v}]\!]_{\mathcal{B}} = (\lambda_1, \ldots, \lambda_n).$$

2.38. Példa. A 2.35. Példában megadott bázisokban nagyon egyszerű meghatározni a koordinátákat.

(i) A T^n vektortérben az $\mathcal{E}\colon \mathbf{e}_1, \dots, \mathbf{e}_n$ bázisban egy $\mathbf{v} = (x_1, \dots, x_n)$ vektor koordinátái maguk az x_i skalárok, azaz $[\![\mathbf{v}]\!]_{\mathcal{E}} = (x_1, \dots, x_n)$.

(ii) A komplex számok $\mathbb{R}$ feletti vektorterében az $1,\ i$ bázisra vonatkozóan egy z komplex szám koordinátái $\operatorname{Re} z$ és $\operatorname{Im} z$.

(iii) A $T^{n\times m}$ vektortérben az E_{ij} $(i = 1, \dots, n,\ j = 1, \dots, m)$ bázisra vonatkozóan egy tetszőleges mátrix koordinátái éppen a mátrix elemei.

(iv) A T feletti legfeljebb n-edfokú polinomok vektortérben az $1,\ x,\ x^2,\ \dots, x^n$ bázisra vonatkozóan egy tetszőleges polinom koordinátái éppen az együtthatói: $[\![a_0 + a_1x + \cdots + a_nx^n]\!] = (a_0, a_1, \dots, a_n)$.

2.39. feladat. Tekintsük a $\mathbb{Z}_3^4$ vektortérben az alábbi $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4$ bázist:

$$\mathbf{b}_1 = 1011,\ \mathbf{b}_2 = 1112,\ \mathbf{b}_3 = 2002,\ \mathbf{b}_4 = 1121.$$

Számítsuk ki a $\mathbf{v} = 1222$ vektor koordinátáit ebben a bázisban.

Megoldás. A $\lambda_1\mathbf{b}_1 + \lambda_2\mathbf{b}_2 + \lambda_3\mathbf{b}_3 + \lambda_4\mathbf{b}_4 = \mathbf{v}$ egyenletrendszert kell megoldanunk a $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \in \mathbb{Z}_3$ ismeretlenekre. Írjuk fel az egyenletrendszer mátrixát, és hozzuk elemi sorátalakításokkal redukált lépcsős alakra.

$$\begin{array}{cccc|c} \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 & \mathbf{b}_4 & \mathbf{v} \\ 1 & 1 & 2 & 1 & 1 \\ 0 & 1 & 0 & 1 & 2 \\ 1 & 1 & 0 & 2 & 2 \\ 1 & 2 & 2 & 1 & 2 \end{array} \quad \rightsquigarrow \quad \begin{array}{cccc|c} \boxed{1} & 0 & 0 & 0 & 2 \\ 0 & \boxed{1} & 0 & 0 & 1 \\ 0 & 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & 0 & \boxed{1} & 1 \end{array}$$

Tehát a koordináták: $\lambda_1 = 2,\ \lambda_2 = 1,\ \lambda_3 = 0,\ \lambda_4 = 1$, vagyis a koordinátasor: $[\![\mathbf{v}]\!] = (2, 1, 0, 1)$. ◇

2.40. Tétel. Legyen V vektortér a T test felett. Tetszőleges $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ vektorokra ekvivalensek az alábbiak:

(i) $\mathbf{v}_1, \dots, \mathbf{v}_k$ bázisa V-nek;

(ii) $\mathbf{v}_1, \dots, \mathbf{v}_k$ minimális generátorrendszere V-nek (vagyis a vektorok közül bármelyiket elhagyva a maradék már nem generálja V-t);

(iii) $\mathbf{v}_1, \dots, \mathbf{v}_k$ maximális független vektorrendszer (vagyis bármely V-beli vektort hozzávéve már függő vektorrendszert kapunk).

Bizonyítás. Lásd [SzL] 8.2. Tétel. □

2.41. Tétel. Legyen V egy véges dimenziós vektortér a T test felett, és legyen $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$.

(i) Ha $\mathbf{v}_1, \dots, \mathbf{v}_k$ generátorrendszere V-nek, akkor el lehet hagyni néhányat a $\mathbf{v}_1, \dots, \mathbf{v}_k$ vektorok közül úgy, hogy V egy bázisát kapjuk.

(ii) Ha $\mathbf{v}_1, \dots, \mathbf{v}_k$ lineárisan független vektorrendszer, akkor ki lehet egészíteni néhány vektorral úgy, hogy V egy bázisát kapjuk.

Bizonyítás. Előkészítő megjegyzés: a tételben szereplő véges dimenziós feltevés ekvivalens azzal, hogy V végesen generált, azaz van véges generátorrendszere (ezt majd később látjuk, lásd a 2.50. Megjegyzést).

(i) Tfh. $\mathbf{v}_1, \dots, \mathbf{v}_k$ generátorrendszer. Ha van olyan vektor a $\mathbf{v}_1, \dots, \mathbf{v}_k$ vektorok között, ami kifejezhető a többiek lineáris kombinációjaként, akkor ezt a vektort elhagyva, a maradék még mindig generálja V-t. Így egyenként kidobálva a felesleges vektorokat, végül olyan generátorrendszert kapunk, amelyben már egyik vektor sem áll elő a többiek lineáris kombinációjaként, tehát lineárisan független (lásd a 2.32. Tételt), vagyis bázis.

(ii) Tfh. $\mathbf{v}_1, \dots, \mathbf{v}_k$ lineárisan független. Legyen $\mathbf{u}_1, \dots, \mathbf{u}_n$ egy tetszőleges generátorrendszere V-nek; ekkor persze $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{u}_1, \dots, \mathbf{u}_n$ is generátorrendszer (itt most fontos lesz a vektorok sorrendje!). Az (i) állítás bizonyításában leírt módszerrel dobálunk ki felesleges vektorokat ebből a generátorrendszerből, de most csak olyan vektort dobunk ki, ami előáll az őt megelőző vektorok lineáris kombinációjaként. Így ismét lineárisan független generátorrendszert, azaz bázist kapunk. Mivel $\mathbf{v}_1, \dots, \mathbf{v}_k$ lineárisan független, közülük senki sem áll elő az őt megelőző vektorok lineáris kombinációjaként (lásd a 2.32. Tételt), így ezek közül a vektorok közül egyet se fogunk kidobni. Tehát olyan bázist kaptunk, amiben szerepel az eredeti $\mathbf{v}_1, \dots, \mathbf{v}_k$ vektorok mindegyike. □

Hogy a fenti tételt a gyakorlatban alkalmazni tudjuk, módszert kell adnunk annak eldöntésére, hogy egy vektorrendszer valamelyik tagja előáll-e a többiek (vagy az őt megelőző vektorok) lineáris kombinációjaként. Ebben segít a következő tétel.

2.42. Tétel. Legyen $A, A' \in T^{k\times n}$ olyan mátrixok, amelyek megkaphatóak egymásból elemi sorátalakításokkal. Ekkor A oszlopai között pontosan azok a lineáris összefüggések teljesülnek, mint A' oszlopai között. Precízebben: ha A oszlopai $\mathbf{v}_1, \dots, \mathbf{v}_n$, és A' oszlopai $\mathbf{v}'_1, \dots, \mathbf{v}'_n$, akkor minden $\lambda_1, \dots, \lambda_n \in T$ esetén

$$\lambda_1\mathbf{v}_1 + \dots + \lambda_n\mathbf{v}_n = \mathbf{0} \iff \lambda_1\mathbf{v}'_1 + \dots + \lambda_n\mathbf{v}'_n = \mathbf{0}. \tag{1}$$

Bizonyítás. A tétel egyszerű de macerás számolással igazolható, ha elemenként kiírjuk a mátrixokat. Ennél a földhözragadt bizonyításnál tanulságosabb lesz kicsit magasabb nézőpontból tekinteni a szituációt.
Csak az 1.4. Definícióbeli első fajta átalakítással foglalkozunk (a többi (szinte) triviális), és az egyszerűség kedvéért (és az általánosság megszorítása nélkül) feltesszük, hogy a második sorhoz adjuk hozzá az első sor μ-szörösét. Ez azt jelenti, hogy mindegyik oszlopvektorban a második komponenshez hozzáadjuk az első komponens μ-szörösét. Ezt az alábbi leképezéssel írhatjuk le:

$$\varphi\colon T^k \to T^k,\ (x_1,\ x_2,\ x_3,\ \dots,\ x_k) \mapsto (x_1,\ x_2 + \mu x_1,\ x_3,\ \dots,\ x_k).$$

Egyszerű számolás mutatja, hogy φ rendelkezik az alábbi tulajdonságokkal:

(1) $\forall \mathbf{x}, \mathbf{y} \in T^k\colon (\mathbf{x}+\mathbf{y})\varphi = \mathbf{x}\varphi + \mathbf{y}\varphi$;

(2) $\forall \mathbf{x} \in T^k\ \forall \lambda \in T\colon (\lambda\mathbf{x})\varphi = \lambda(\mathbf{x}\varphi)$;

(3) φ bijektív (mi az inverze?).

(Az ilyen leképezést izomorfizmusnak nevezzük; lásd a 2.53. Definíciót.) A fenti első két tulajdonságból következik, hogy φ felcserélhető a lineáris kombinációk képzésével, és így minden $\lambda_1, \dots, \lambda_n \in T$ esetén

$$(\lambda_1\mathbf{v}_1 + \dots + \lambda_n\mathbf{v}_n)\varphi = \lambda_1(\mathbf{v}_1\varphi) + \dots + \lambda_n(\mathbf{v}_n\varphi).$$

Mivel φ bijektív, a bal oldal akkor és csak akkor nullvektor, ha $\lambda_1\mathbf{v}_1 + \dots + \lambda_n\mathbf{v}_n = \mathbf{0}$. A jobb oldal pedig akkor és csak akkor nullvektor, ha $\lambda_1\mathbf{v}'_1 + \dots + \lambda_n\mathbf{v}'_n = \mathbf{0}$, hiszen $\mathbf{v}'_i = \mathbf{v}_i\varphi$ (pont így definiáltuk a φ leképezést). Ezzel igazoltuk, hogy (1) teljesül. □

Egy A mátrix elemei sorátalakításai nem változtatják meg

- **az $Ax = 0$ homogén lineáris egyenletrendszer megoldáshalmazát;**
- **a mátrix sorai átlal kifeszített alteret;**
- **a mátrix oszlopai közötti lineáris összefüggéseket.**

2.43. feladat. Szűkítsük a $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5, \mathbf{v}_6] \le \mathbb{R}^4$ altér bázisává a $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5, \mathbf{v}_6$ vektorrendszert.

$\mathbf{v}_1 = (1,-1,2,0)$, $\mathbf{v}_2 = (-1,2,-5,-1)$, $\mathbf{v}_3 = (0,1,-3,-1)$, $\mathbf{v}_4 = (2,-1,2,0)$, $\mathbf{v}_5 = (1,2,-6,-2)$, $\mathbf{v}_6 = (-3,3,-4,3)$

Megoldás. A vektorokat oszlopvektorokként egymás mellé írjuk, és a kapott mátrixot elemi sorátalakításokkal redukált lépcsős alakra hozzuk:

$$\begin{array}{rrrrrr} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4 & \mathbf{v}_5 & \mathbf{v}_6 \\ 1 & -1 & 0 & 2 & 1 & -3 \\ -1 & 2 & 1 & -1 & 2 & 3 \\ 2 & -5 & -3 & 2 & -6 & -4 \\ 0 & -1 & -1 & 0 & -2 & 3 \end{array} \quad \rightsquigarrow \quad \begin{array}{cccccc} \mathbf{v}'_1 & \mathbf{v}'_2 & \mathbf{v}'_3 & \mathbf{v}'_4 & \mathbf{v}'_5 & \mathbf{v}'_6 \\ \boxed{1} & 0 & 1 & 0 & 1 & 0 \\ 0 & \boxed{1} & 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & \boxed{1} & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \boxed{1} \end{array}$$

A kapott mátrixról látszik, hogy az 1., 2., 4., 6. oszlopa lineárisan független, a többi pedig kifejezhető ezek lineáris kombinációjaként. Mivel az elemi sorátalakítások nem változtatják meg az oszlopok közötti lineáris összefüggéseket, ugyanez érvényes az eredeti mátrixra is. Tehát $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4, \mathbf{v}_6$ lineárisan független, továbbá $\mathbf{v}_3 = \mathbf{v}_1 + \mathbf{v}_2$ és $\mathbf{v}_5 = \mathbf{v}_1 + 2\mathbf{v}_2 + \mathbf{v}_4$. Ezért $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4, \mathbf{v}_6$ bázisa a $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5, \mathbf{v}_6]$ altérnek (amiről így kiderült, hogy nem más, mint $\mathbb{R}^4$). ◇

2.44. feladat. Bővítsük $\mathbb{R}^4$ bázisává a $\mathbf{v}_1, \mathbf{v}_2$ vektorrendszert.

$$\mathbf{v}_1 = (3,2,2,3), \quad \mathbf{v}_2 = (6,1,2,3)$$

Megoldás. A vektorokat oszlopvektorokként egymás mellé írjuk, melléjük tesszük még $\mathbb{R}^4$ egy tetszőleges bázisát (pl. az $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4$ standard bázist) és a kapott mátrixot elemi sorátalakításokkal redukált lépcsős alakra hozzuk:

$$\begin{array}{cccccc} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 \\ 3 & 6 & 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 & 0 \\ 2 & 2 & 0 & 0 & 1 & 0 \\ 3 & 3 & 0 & 0 & 0 & 1 \end{array} \quad \rightsquigarrow \quad \begin{array}{cccccc} \mathbf{v}_1' & \mathbf{v}_2' & \mathbf{e}_1' & \mathbf{e}_2' & \mathbf{e}_3' & \mathbf{e}_4' \\ \boxed{1} & 0 & 0 & 1 & 0 & -1/3 \\ 0 & \boxed{1} & 0 & -1 & 0 & 2/3. \\ 0 & 0 & \boxed{1} & 3 & 0 & -3 \\ 0 & 0 & 0 & 0 & \boxed{1} & -2/3 \end{array}$$

Az előző feladathoz hasonlóan leolvasható, hogy $\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_1, \mathbf{e}_3$ bázisa a $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4] = \mathbb{R}^4$ (al)térnek. ⋄

2.7. Dimenzió

2.45. Tétel (kicserélési tétel). Legyen V vektortér a T test felett, legyen $\mathbf{u}_1, \ldots, \mathbf{u}_k$ lineárisan független vektorrendszer, és legyen $\mathbf{v}_1, \ldots, \mathbf{v}_\ell$ generátorrendszere V-nek. Ekkor bármelyik $\mathbf{u}_i$ vektorhoz van olyan $\mathbf{v}_j$ vektor, amelyre $\mathbf{u}_1, \ldots, \mathbf{u}_{i-1}, \mathbf{u}_{i+1}, \ldots, \mathbf{u}_k, \mathbf{v}_j$ is lineárisan független.

Bizonyítás. Indirekten bizonyítunk: tfh. van olyan $i \in \{1, \ldots, k\}$, hogy $\mathbf{u}_1, \ldots, \mathbf{u}_{i-1}, \mathbf{u}_{i+1}, \ldots, \mathbf{u}_k, \mathbf{v}_j$ lineárisan függő minden $j \in \{1, \ldots, \ell\}$ esetén. Ekkor az $\mathbf{u}_1, \ldots, \mathbf{u}_{i-1}, \mathbf{u}_{i+1}, \ldots, \mathbf{u}_k, \mathbf{v}_j$ vektorok valamelyike kifejezhető az előtte álló vektorok lineáris kombinációjaként (lásd a 2.32. Tételt). Mivel $\mathbf{u}_1, \ldots, \mathbf{u}_{i-1}, \mathbf{u}_{i+1}, \ldots, \mathbf{u}_k$ lineárisan független vektorrendszer, közülük egyik sem fejezhető ki a korábbi vektorok lineáris kombinációjaként (ismét a 2.32. Tétel), így $\mathbf{v}_j \in [\mathbf{u}_1, \ldots, \mathbf{u}_{i-1}, \mathbf{u}_{i+1}, \ldots, \mathbf{u}_k]$. Ez minden j-re teljesül, ezért $[\mathbf{v}_1, \ldots, \mathbf{v}_\ell] \subseteq [\mathbf{u}_1, \ldots, \mathbf{u}_{i-1}, \mathbf{u}_{i+1}, \ldots, \mathbf{u}_k]$. Ez pedig azt jelenti, hogy $[\mathbf{u}_1, \ldots, \mathbf{u}_{i-1}, \mathbf{u}_{i+1}, \ldots, \mathbf{u}_k] = V$, hiszen $\mathbf{v}_1, \ldots, \mathbf{v}_\ell$ generátorrendszere V-nek. Ebből következik, hogy $\mathbf{u}_i \in [\mathbf{u}_1, \ldots, \mathbf{u}_{i-1}, \mathbf{u}_{i+1}, \ldots, \mathbf{u}_k]$, ami ellentmond $\mathbf{u}_1, \ldots, \mathbf{u}_k$ lineáris függetlenségének (már megint a 2.32. Tétel). □

2.46. Következmény. Bármely lineárisan független vektorrendszer elemszáma legfeljebb akkora, mint bármely generátorrendszer elemszáma: ha $\mathbf{u}_1, \ldots, \mathbf{u}_k$ lineárisan független vektorrendszer és $\mathbf{v}_1, \ldots, \mathbf{v}_\ell$ generátorrendszer a V vektortérben, akkor $k \leq \ell$.

Bizonyítás. A kicserélési tételt alkalmazva, sorra kicserélhetjük az $\mathbf{u}_i$ vektorokat a $\mathbf{v}_j$ vektorokra úgy, hogy a lineáris függetlenség megmaradjon:

$$\begin{array}{llr l} & & \mathbf{u}_1,\ \mathbf{u}_2,\ \mathbf{u}_3,\ \ldots,\ \mathbf{u}_{k-1},\ \mathbf{u}_k & \text{lineárisan független} \\ \exists j_1 & \in \{1,\ldots,\ell\}: & \mathbf{v}_{j_1},\ \mathbf{u}_2,\ \mathbf{u}_3,\ \ldots,\ \mathbf{u}_{k-1},\ \mathbf{u}_k & \text{lineárisan független} \\ \exists j_2 & \in \{1,\ldots,\ell\}: & \mathbf{v}_{j_1},\ \mathbf{v}_{j_2},\ \mathbf{u}_3,\ \ldots,\ \mathbf{u}_{k-1},\ \mathbf{u}_k & \text{lineárisan független} \\ \exists j_3 & \in \{1,\ldots,\ell\}: & \mathbf{v}_{j_1},\ \mathbf{v}_{j_2},\ \mathbf{v}_{j_3},\ \ldots,\ \mathbf{u}_{k-1},\ \mathbf{u}_k & \text{lineárisan független} \\ & \vdots & & \\ \exists j_{k-1} & \in \{1,\ldots,\ell\}: & \mathbf{v}_{j_1},\ \mathbf{v}_{j_2},\ \mathbf{v}_{j_3},\ \ldots,\ \mathbf{v}_{j_{k-1}},\ \mathbf{u}_k & \text{lineárisan független} \\ \exists j_k & \in \{1,\ldots,\ell\}: & \mathbf{v}_{j_1},\ \mathbf{v}_{j_2},\ \mathbf{v}_{j_3},\ \ldots,\ \mathbf{v}_{j_{k-1}},\ \mathbf{v}_{j_k} & \text{lineárisan független} \end{array}$$

Lineárisan független vektorrendszerben nem lehetnek ismétlődő vektorok (miért?), így a $j_1, \ldots, j_k \in \{1, \ldots, \ell\}$ indexek páronként különbözőek. Ebből pedig már következik, hogy $k \leq \ell$. □

2.47. Következmény. Vektortér bármely két bázisának ugyanannyi eleme van: ha $\mathbf{u}_1, \ldots, \mathbf{u}_k$ és $\mathbf{v}_1, \ldots, \mathbf{v}_\ell$ is bázisa a V vektortérnek, akkor $k = \ell$.

2.48. Definíció. Vektortér bázisának elemszámát a vektortér dimenziójának nevezzük. Jelölés: $\dim V$. Ha nincs véges bázisa a vektortérnek, akkor azt mondjuk, hogy végtelen dimenziós.

2.49. Megjegyzés. A 2.47. Következmény szerint a dimenzió fogalma jól definiált (nem függ attól, hogy melyik bázist választjuk).

$$|\textbf{lin. fgtl. vrsz.}| \ \leq\ \textbf{dim} \ \leq\ |\textbf{gen. rsz.}|$$

2.50. Megjegyzés. Ha egy vektortérnek van véges generátorrendszere, akkor a 2.41. Tétel szerint van véges bázisa is. Tehát a véges dimenziós vektorterek ugyanazok, mint a végesen generált vektorterek.

2.51. Példa. A 2.35. Példában szereplő vektorterek dimenziója:

(i) A T^n vektortér dimenziója n.

(ii) A komplex számok $\mathbb{R}$ feletti vektorterének dimenziója 2.

(iii) A $T^{n\times m}$ vektortér dimenziója $n \cdot m$.

(iv) A T feletti legfeljebb n-edfokú polinomok vektorterének dimenziója $n+1$.

2.52. Példa. A valós számok teste, mint $\mathbb{Q}$ feletti vektortér végtelen dimenziós. Itt is lehet bázisról beszélni: bázison valós számok olyan (végtelen) halmazát értjük, amelynek elemeiből minden valós szám előállítható racionális együtthatós lineáris kombinációval, mégpedig egyértelműen. (Természetesen egy lineáris kombinációnak csak véges sok tagja lehet.) Az ilyen bázist <u>Hamel-bázisnak</u> nevezzük.

2.53. Definíció. Legyenek V és W vektorterek ugyanazon T test felett. A $\varphi\colon V \to W$ leképezést <u>izomorfizmusnak</u> nevezzük, ha rendelkezik az alábbi tulajdonságokkal:

(1) $\forall \mathbf{u}, \mathbf{v} \in V\colon (\mathbf{u}+\mathbf{v})\varphi = \mathbf{u}\varphi + \mathbf{v}\varphi$;

(2) $\forall \mathbf{v} \in V\ \forall \lambda \in T\colon (\lambda \mathbf{v})\varphi = \lambda(\mathbf{v}\varphi)$;

(3) φ bijektív.

Ha létezik $V \to W$ izomorfizmus, akkor azt mondjuk, hogy V és W <u>izomorfak</u>. Jelölés: $V \cong W$.

2.54. Tétel. Ha V egy n-dimenziós vektortér a T test felett, akkor $V \cong T^n$.

Bizonyítás. Legyen $\mathcal{E}\colon \mathbf{e}_1, \ldots, \mathbf{e}_n$ bázisa V-nek. Ekkor az alábbi leképezés izomorfozmus T^n-ről V-re:

$$\varphi\colon T^n \to V,\ (\lambda_1, \ldots, \lambda_n) \mapsto \lambda_1 \mathbf{e}_1 + \cdots + \lambda_n \mathbf{e}_n.$$

Az izomorfizmus definíciójában szereplő első két tulajdonság egyszerű számolással ellenőrizhető (lásd [SzL] 10.10. Tétel). A φ leképezés bijektivitása pedig következik a 2.36. Tételből. ☐

2.55. Példa. A 2.5. Példában szereplő vektortér izomorf a $\mathbb{Z}_2^2$ vektortérrel az alábbi izomorfizmus mellett:

☂ $\mapsto 00$, 🚗 $\mapsto 01$, 🎂 $\mapsto 10$, 🌲 $\mapsto 11$.

(Nem ez az egyetlen $V \to \mathbb{Z}_2^2$ izomorfizmus. Keressük meg az összeset!)

2.56. Tétel (kettőt fizet, hármat vihet!). Ha V egy n-dimenziós vektortér és $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$, akkor az alábbi három állítás közül bármelyik kettő maga után vonja a harmadikat:

(i) $\mathbf{v}_1, \ldots, \mathbf{v}_k$ lineárisan független;

(ii) $\mathbf{v}_1, \ldots, \mathbf{v}_k$ generátorrendszer;

(iii) $k = n$.

Bizonyítás.

- (i) & (ii) $\Longrightarrow$ (iii): Ha $\mathbf{v}_1, \ldots, \mathbf{v}_k$ lineárisan független generátorrendszer, akkor definíció szerint bázis, és így $k = \dim V = n$.

- (i) & (iii) $\Longrightarrow$ (ii): Ha $\mathbf{v}_1, \ldots, \mathbf{v}_k$ lineárisan független, akkor bázissá bővíthető. Mivel a bázis n-elemű és $k = n$, a bővítés igazából nem bővítés (egyetlen vektort sem veszünk hozzá), tehát $\mathbf{v}_1, \ldots, \mathbf{v}_k$ már bázis.

- (ii) & (iii) $\Longrightarrow$ (i): Ha $\mathbf{v}_1, \ldots, \mathbf{v}_k$ generátorrendszer, akkor bázissá szűkíthető. Mivel a bázis n-elemű és $k = n$, a szűkítés igazából nem szűkítés (egyetlen vektort sem hagyunk el), tehát $\mathbf{v}_1, \ldots, \mathbf{v}_k$ már bázis. ☐

2.8. Alterek dimenziótétele, direkt összeg

2.57. Tétel. Ha U altere az n-dimenziós V vektortérnek, akkor

(i) $\dim U \leq n$;

(ii) $\dim U = n \iff U = V$;

(iii) $\dim U = 0 \iff U = \{\mathbf{0}\}$.

Bizonyítás.

(i) A 2.46. Következmény szerint V-ben bármely lineárisan független vektorrendszernek legfeljebb n eleme lehet, és így persze n felső korlátja az U-beli lineárisan független vektorrendszerek méretének is. Ezért létezik U-ban véges maximális független vektorrendszer; ez a 2.40. Tétel szerint bázisa U-nak, és legfeljebb n eleme van.

(ii) Az világos, hogy $U = V \implies \dim U = n$. Fordítva, ha $\dim U = n$, akkor a 2.56. Tétel (*Kettőt fizet, hármat vihet!*) szerint U bármely bázisa egyúttal V-nek is bázisa (mert lineárisan független és n eleme van), és így $U = V$.

(iii) Az világos, hogy $U = \{\mathbf{0}\} \implies \dim U = 0$. Fordítva, ha $\dim U = 0$, akkor U-ban még az egyelemű vektorrendszerek sem lineárisan függetlenek, ezért U-ban nincs nemnulla vektor. □

2.58. Tétel. Ha U és V alterei egy véges dimenziós vektortérnek, akkor

$$\dim(U + V) = \dim U + \dim V - \dim(U \cap V).$$

Bizonyítás. Lásd [SzL] 8.7. Tétel. □

2.59. Tétel. Legyenek U_1 és U_2 alterei a V véges dimenziós vektortérnek. Ekkor az alábbi állítások ekvivalensek:

(i) V minden eleme egyértelműen előáll $\mathbf{u}_1 + \mathbf{u}_2$ ($\mathbf{u}_1 \in U_1$, $\mathbf{u}_2 \in U_2$) alakban;

(i+ε) $U_1 + U_2 = V$ és a nullvektor csak triviális módon áll elő $\mathbf{u}_1 + \mathbf{u}_2$ ($\mathbf{u}_1 \in U_1$, $\mathbf{u}_2 \in U_2$) alakban;

(ii) $U_1 + U_2 = V$ és $U_1 \cap U_2 = \{\mathbf{0}\}$;

(iii) $U_1 + U_2 = V$ és $\dim U_1 + \dim U_2 = \dim V$;

(iv) bárhogyan is veszünk egy-egy bázist az U_1 és U_2 alterekben, ezek egyesítése bázisa lesz V-nek;

(v) tudunk úgy venni egy-egy bázist az U_1 és U_2 alterekben, hogy ezek egyesítése bázisa legyen V-nek.

Bizonyítás. Először az első három állítás ekvivalenciáját igazoljuk.

- (i) $\implies$ (i+ε): Triviális.
- (i+ε) $\implies$ (ii): Az $U_1 + U_2 = V$ feltevés szerepel (i+ε)-ban, tehát csak a metszetre vonatkozó állítással kell foglalkozunk. Ha $\mathbf{v} \in U_1 \cap U_2$, akkor a nullvektort felírhatjuk így: $\mathbf{0} = \mathbf{v} + (-\mathbf{v})$, és itt az első tag U_1-ben, a második tag U_2-ben van (persze igazából mindkettő benne van mindkét altérben). Az (i+ε) állítás szerint ez csak $\mathbf{v} = \mathbf{0}$ esetén lehetséges.
- (ii) $\implies$ (i): Az világos, hogy V minden eleme előáll $\mathbf{u}_1 + \mathbf{u}_2$ ($\mathbf{u}_1 \in U_1$, $\mathbf{u}_2 \in U_2$) alakban. Az egyértelműség igazolásához tfh. $\mathbf{v} = \mathbf{u}_1 + \mathbf{u}_2 = \mathbf{u}_1' + \mathbf{u}_2'$, ahol $\mathbf{u}_1, \mathbf{u}_1' \in U_1$ és $\mathbf{u}_2, \mathbf{u}_2' \in U_2$. Ebből következik, hogy $\mathbf{u}_1 - \mathbf{u}_1' = \mathbf{u}_2' - \mathbf{u}_2$. Itt a bal oldal U_1-ben, a jobb oldal U_2-ben van, így mindkét oldal eleme az $U_1 \cap U_2$ altérnek. Mivel (ii) szerint $U_1 \cap U_2 = \{\mathbf{0}\}$, azt kapjuk, hogy $\mathbf{u}_1 - \mathbf{u}_1' = \mathbf{u}_2' - \mathbf{u}_2 = \mathbf{0}$, és ez épp azt jelenti, hogy a $\mathbf{v}$ vektor tekintett két felbontása „szőröstül-bőröstül” megegyezik.

A második és harmadik állítás ekvivalenciája rögtön következik az alterek dimenziótételéből, hiszen

$$\dim(U_1 \cap U_2) = \dim U_1 + \dim U_2 - \dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim V.$$

(Itt felhasználtuk, hogy $U_1 + U_2 = V$, node ez fel van téve a (ii) és a (iii) állításban is.)
Végül megmutatjuk, hogy az utolsó három állítás ekvivalens.

- (iii) $\implies$ (iv): Legyen $\mathcal{B}_i$ bázisa az U_i altérnek ($i = 1, 2$). Mivel $U_1 + U_2 = V$, a $\mathcal{B}_1 \cup \mathcal{B}_2$ vektorrendszer generálja V-t. A $\dim U_1 + \dim U_2 = \dim V$ feltevés szerint ennek a vektorrendszernek éppen $\dim V$ eleme van. A a 2.56. Tétel (Kettőt fizet, hármat vihet!) alapján ebből már következik, hogy $\mathcal{B}_1 \cup \mathcal{B}_2$ bázisa V-nek.
- (iv) $\implies$ (v): Ez majdnem triviális. (Miért csak majdnem?)
- (v) $\implies$ (iii): Legyen $\mathcal{B}_1$ és $\mathcal{B}_2$ az (v) állításban szereplő két bázis. Tudjuk, hogy $\mathcal{B}_1 \cup \mathcal{B}_2$ bázisa V-nek, ezért $\dim V = |\mathcal{B}_1 \cup \mathcal{B}_2| = \dim U_1 + \dim U_2$, továbbá $V = [\mathcal{B}_1 \cup \mathcal{B}_2] = U_1 + U_2$. □

2.60. Definíció. Ha az $U_1, U_2 \leq V$ alterek eleget tesznek a 2.59. Tételbeli ekvivalens feltételeknek, akkor azt mondjuk, hogy V az U_1 és U_2 alterek direkt összege. Jelölés: $V = U_1 \oplus U_2$.

A direkt összeg kettőnél több altérre is hasonlóan definiálható, de a 2.59. Tételbeli második állítás helyébe „kicsit" komplikáltabb feltétel lép.

2.61. Tétel. Legyenek $U_1, \ldots, U_k$ alterei a V véges dimenziós vektortérnek. Ekkor az alábbi állítások ekvivalensek:

(i) V minden eleme egyértelműen előáll $\mathbf{u}_1 + \cdots + \mathbf{u}_k$ $(\mathbf{u}_i \in U_i)$ alakban;

(i+ε) $U_1 + \cdots + U_k = V$ és a nullvektor csak triviális módon áll elő $\mathbf{u}_1 + \cdots + \mathbf{u}_k$ $(\mathbf{u}_i \in U_i)$ alakban;

(ii) $U_1 + \cdots + U_k = V$ és $U_j \cap (U_1 + \cdots + U_{j-1} + U_{j+1} + \cdots + U_k) = \{\mathbf{0}\}$ minden $j \in \{1, \ldots, n\}$ esetén;

(iii) $U_1 + \cdots + U_k = V$ és $\dim U_1 + \cdots + \dim U_k = \dim V$;

(iv) bárhogyan is veszünk egy-egy bázist az U_i alterekben, ezek egyesítése bázisa lesz V-nek;

(v) tudunk úgy venni egy-egy bázist az U_i alterekben, hogy ezek egyesítése bázisa legyen V-nek.

Bizonyítás. Először az első három állítás ekvivalenciáját igazoljuk.

- (i) $\Longrightarrow$ (i+ε): Triviális.

- (i+ε) $\Longrightarrow$ (ii): Az $U_1 + \cdots + U_k = V$ feltevés szerepel (i+ε)-ban, tehát csak a metszetre vonatkozó állítással kell foglalkozunk. Ha $\mathbf{v} \in U_j \cap (U_1 + \cdots + U_{j-1} + U_{j+1} + \cdots + U_k)$, akkor $\mathbf{v} = \mathbf{u}_j = \mathbf{u}_1 + \cdots + \mathbf{u}_{j-1} + \mathbf{u}_{j+1} + \cdots + \mathbf{u}_k$ teljesül alkalmas $\mathbf{u}_i \in U_i$ vektorokkal. Ebből megkapjuk a nullvektor egy felírását az U_i alterekbe tartozó vektorok összegeként: $\mathbf{0} = \mathbf{u}_1 + \cdots + \mathbf{u}_{j-1} + (-\mathbf{u}_j) + \mathbf{u}_{j+1} + \cdots + \mathbf{u}_k$. Az (i+$\varepsilon$) állítás szerint a nullvekor csak triviális módon állítható elő így, tehát $-\mathbf{u}_j = \mathbf{0}$, és így $\mathbf{v} = \mathbf{0}$ (persze a többi $\mathbf{u}_i$ is nullvektor).

- (ii) $\Longrightarrow$ (i): Az világos, hogy V minden eleme előáll $\mathbf{u}_1 + \cdots + \mathbf{u}_k$ $(\mathbf{u}_i \in U_i)$ alakban. Az egyértelműség igazolásához tegyük fel, hogy $\mathbf{u}_1 + \cdots + \mathbf{u}_k = \mathbf{u}'_1 + \cdots + \mathbf{u}'_k$, ahol $\mathbf{u}_i, \mathbf{u}'_i \in U_i$. Ebből következik, hogy $\mathbf{u}_1 - \mathbf{u}'_1 = (\mathbf{u}'_2 - \mathbf{u}_2) + \cdots + (\mathbf{u}'_k - \mathbf{u}_k)$. Itt a bal oldal U_1-ben, a jobb oldal pedig az $U_2 + \cdots + U_k$ altérben van, így mindkét oldal eleme az $U_1 \cap (U_2 + \cdots + U_k)$ altérnek. Mivel (ii) szerint $U_1 \cap (U_2 + \cdots + U_k) = \{\mathbf{0}\}$, azt kapjuk, hogy $\mathbf{u}_1 - \mathbf{u}'_1 = \mathbf{0}$. Hasonló módon igazolható minden i-re, hogy $\mathbf{u}_i = \mathbf{u}'_i$.

A második és harmadik állítás ekvivalenciája sajnos most nem következik egyszerűen az alterek dimenziótételéből. Helyette az (i) és (iv) állítások ekvivalenciáját mutatjuk meg.

- (i) $\Longrightarrow$ (iv): Legyen $\mathcal{B}_i\colon \mathbf{e}_i^{(1)}, \ldots, \mathbf{e}_i^{(d_i)}$ bázisa az U_i altérnek $(i = 1, \ldots, k)$. Az (i) állítás egzisztencia részéből következik, hogy $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ generátorrendszere V-nek. A lineáris függetlenséghez tfh.
$$\sum_{\substack{i=1,\ldots,k \\ h=1,\ldots,d_i}} \lambda_i^{(h)} \mathbf{e}_i^{(h)} = \mathbf{0},$$
ahol $\lambda_i^{(h)} \in T$ $(i = 1, \ldots, k,\ h = 1, \ldots, d_i)$. Tagoljuk az összeget (az i-edik csoportba az U_i-beli vektorok kerülnek):
$$\underbrace{\sum_{h=1}^{d_1} \lambda_1^{(h)} \mathbf{e}_1^{(h)}}_{\mathbf{u}_1} + \cdots + \underbrace{\sum_{h=1}^{d_k} \lambda_1^{(h)} \mathbf{e}_k^{(h)}}_{\mathbf{u}_k} = \mathbf{0}.$$
Az (i) állítás egyértelműségi része miatt $\mathbf{u}_i = \mathbf{0}$ minden $i \in \{1, \ldots, k\}$ esetén, ebből pedig $\mathcal{B}_i$ lineáris függetlensége miatt $\lambda_i^{(h)} = 0$ minden $h \in \{1, \ldots, d_i\}$ esetén. Ezzel beláttuk, hogy az összes $\lambda_i^{(h)}$ együttható nulla, tehát $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ valóban lineárisan független.

- (iv) $\Longrightarrow$ (i): Legyen $\mathcal{B}_i\colon \mathbf{e}_i^{(1)}, \ldots, \mathbf{e}_i^{(d_i)}$ bázisa az U_i altérnek $(i = 1, \ldots, k)$. Az (i) állítás egzisztencia része következik abból, hogy $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ generátorrendszere V-nek. Az egyértelműséghez tegyük fel, hogy $\mathbf{u}_1 + \cdots + \mathbf{u}_k = \mathbf{u}'_1 + \cdots + \mathbf{u}'_k$, ahol $\mathbf{u}_i, \mathbf{u}'_i \in U_i$. Ebből kapjuk, hogy $(\mathbf{u}_1 - \mathbf{u}'_1) + (\mathbf{u}_2 - \mathbf{u}'_2) + \cdots + (\mathbf{u}_k - \mathbf{u}'_k) = \mathbf{0}$. Minden i-re írjuk fel az $\mathbf{u}_i - \mathbf{u}'_i$ vektort a $\mathcal{B}_i$ bázisban:
$$\underbrace{\sum_{h=1}^{d_1} \lambda_1^{(h)} \mathbf{e}_1^{(h)}}_{\mathbf{u}_1 - \mathbf{u}'_1} + \cdots + \underbrace{\sum_{h=1}^{d_k} \lambda_k^{(h)} \mathbf{e}_k^{(h)}}_{\mathbf{u}_k - \mathbf{u}'_k} = \mathbf{0}.$$
Amit itt látunk, az a $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ vektorrendszer egy lineáris kombinációja. Ez egy lineárisan független vektorrendszer, ezért az összes $\lambda_i^{(h)}$ együttható nulla, és így $\mathbf{u}_i = \mathbf{u}'_i$ minden i-re.

Az utolsó három állítás ekvivalenciája a 2.59. Tétel bizonyításában látott gondolatmenettel igazolható.

- (iii) $\Longrightarrow$ (iv): Legyen $\mathcal{B}_i$ bázisa az U_i altérnek $(i = 1, \ldots, k)$. Mivel $U_1 + \cdots + U_k = V$, a $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ vektorrendszer generálja V-t. A $\dim U_1 + \cdots + \dim U_k = \dim V$ feltevés szerint ennek a vektorrendszernek éppen $\dim V$ eleme van. A 2.56. Tétel (Kettőt fizet, hármat vihet!) alapján ebből már következik, hogy $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ bázisa V-nek.

- (iv) $\implies$ (v): Ez majdnem triviális. (Miért csak majdnem?)
- (v) $\implies$ (iii): Legyenek $\mathcal{B}_i$ $(i = 1, \dots, k)$ az (v) állításban szereplő bázisok. Tudjuk, hogy $\mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ bázisa V-nek, ezért $\dim V = |\mathcal{B}_1 \cup \dots \cup \mathcal{B}_k| = \dim U_1 + \dots + \dim U_2$, továbbá $V = [\mathcal{B}_1 \cup \dots \cup \mathcal{B}_k] = U_1 + \dots + U_k$.

□

2.62. Definíció. Ha az $U_1, \dots, U_k \le V$ alterek eleget tesznek a 2.61. Tételbeli ekvivalens feltételeknek, akkor azt mondjuk, hogy V az $U_1, \dots, U_k$ alterek direkt összege. Jelölés: $V = U_1 \oplus \dots \oplus U_k$.

2.9. Rang

2.63. Definíció (a rang egyik definíciója). Vektorrendszer rangján a lineárisan független részrendszerei közül a legnagyobb(ak)nak az elemszámát értjük. Tehát a $\mathbf{v}_1, \dots, \mathbf{v}_k$ vektorrendszer rangja akkor r, ha ki lehet választani közülük r darab vektort úgy, hogy azok lineárisan függetlenek legyenek, de bárhogy választunk ki $r+1$ darab vektort, azok már lineárisan függők lesznek. Jelölés: $\mathrm{r}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

2.64. Definíció (a rang másik definíciója). Vektorrendszer rangján az általa kifeszített altér dimenzióját értjük. Tehát a $\mathbf{v}_1, \dots, \mathbf{v}_k$ vektorrendszer rangja:

$$\mathrm{r}(\mathbf{v}_1, \dots, \mathbf{v}_k) = \dim[\mathbf{v}_1, \dots, \mathbf{v}_k].$$

2.65. Tétel. A rang két definíciója ekvivalens.

Bizonyítás. Legyen $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$, és legyen $\mathrm{r}(\mathbf{v}_1, \dots, \mathbf{v}_k) = r$ a 2.63. Definíció szerint. Az egyszerűség kedvéért (de az általánosság megszorítása nélkül) tfh. $\mathbf{v}_1, \dots, \mathbf{v}_r$ lineárisan független. Ekkor persze $\mathbf{v}_1, \dots, \mathbf{v}_r$ bázisa a $[\mathbf{v}_1, \dots, \mathbf{v}_r]$ altérnek (ugye?). Tekintsük a $\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{v}_i$ vektorrendszert tetszőleges $i \in \{r+1, \dots, k\}$ esetén. A 2.63. Definíció szerint ez a vektorrendszer lineárisan összefüggő, így valamelyik vektora előáll a korábbi vektorok lineáris kombinációjaként. Ez a vektor csak $\mathbf{v}_i$ lehet, hiszen $\mathbf{v}_1, \dots, \mathbf{v}_r$ lineárisan független. Ez pedig azt jelenti, hogy $\mathbf{v}_i \in [\mathbf{v}_1, \dots, \mathbf{v}_r]$. Ez minden $i \in \{r+1, \dots, k\}$ indexre teljesül, ezért $[\mathbf{v}_1, \dots, \mathbf{v}_k] = [\mathbf{v}_1, \dots, \mathbf{v}_r]$. Ekkor persze a két altér dimenziója is megegyezik: $\dim[\mathbf{v}_1, \dots, \mathbf{v}_k] = \dim[\mathbf{v}_1, \dots, \mathbf{v}_r] = r$, tehát a $\mathbf{v}_1, \dots, \mathbf{v}_k$ vektorrendszer rangja a 2.64. Definíció szerint is r. □

2.66. feladat. Határozzuk meg az alábbi $A \in \mathbb{R}^{4\times 4}$ mátrix (oszlop)rangját, és adjunk meg $\mathrm{r}(A)$ darab oszlopvektort, amelyek lineárisan függetlenek.

$$A = \begin{pmatrix} 1 & 2 & -1 & 0 \\ 2 & 1 & 3 & 2 \\ -1 & -2 & 2 & 1 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$

Megoldás. A mátrixot elemi sorátalakításokkal redukált lépcsős alakra hozva a következő mátrixot kapjuk:

$$A' = \begin{pmatrix} \boxed{1} & 0 & 0 & -1 \\ 0 & \boxed{1} & 0 & 1 \\ 0 & 0 & \boxed{1} & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Szemlátomást $\mathrm{r}(A') = 3$, és az A' mátrix első három oszlopa lineárisan független. Mivel az elemi sorátalakítások nem változtatják meg az oszlopok közötti lineáris összefüggéseket, ugyanez érvényes az eredeti mátrixra is. Tehát $\mathrm{r}(A) = 3$, és az A mátrix első három oszlopa lineárisan független. ◇

2.67. feladat. Határozzuk meg az alábbi $A \in \mathbb{Z}_3^{5\times 4}$ mátrix (sor)rangját, és adjunk meg $\mathrm{r}(A)$ darab sorvektort, amelyek lineárisan függetlenek.

$$A = \begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 2 & 1 & 1 \\ 1 & 2 & 2 & 0 \\ 1 & 1 & 0 & 2 \\ 0 & 1 & 2 & 1 \end{pmatrix}$$

Megoldás. Elemi sorátalakítások nem változtatják meg a mátrix rangját, de a sorokat (a sorok közötti lineáris összefüggéseket) összekutyulhatják. Ezért most vagy elemi oszlopátalakításokat kell végeznünk, vagy a mátrix transzponáltján kell elemi sorátalakításokat végeznünk. Tegyük az utóbbit:

$$A^{\mathrm{T}} = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 2 & 2 & 1 & 1 \\ 1 & 1 & 2 & 0 & 2 \\ 2 & 1 & 0 & 2 & 1 \end{pmatrix} \qquad \rightsquigarrow \qquad A' = \begin{pmatrix} \boxed{1} & 0 & 1 & 0 & 1 \\ 0 & \boxed{1} & 1 & 0 & 1 \\ 0 & 0 & 0 & \boxed{1} & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Látható, hogy $\mathrm{r}(A') = 3$, és az A' mátrix első, második és negyedik oszlopa lineárisan független. Ugyanez érvényes az A^{T} mátrixra is, az eredeti A mátrixról pedig ennek alapján azt mondhatjuk, hogy $\mathrm{r}(A) = 3$, és A első, második és negyedik sora lineárisan független. ◇

2.68. Definíció. Egy T test feletti $A \in T^{k\times n}$ mátrix sorrangján sorvektorrendszerének (mint T^n-beli vektorrendszernek) a rangját értjük, oszloprangján pedig oszlopvektorrendszerének (mint T^k-beli vektorrendszernek) a rangját értjük. A 2.70. Tétel szerint ez a kettő mindig megegyezik, ezért beszélhetünk egyszerűen csak a mátrix rangjáról. Jelölés: $\mathrm{r}(A)$.

2.69. Tétel. Elemi sorátalakítások nem változtatják meg sem a mátrix sorrangját sem a mátrix oszloprangját.

Bizonyítás. A 2.21. Tétel szerint az elemi sorátalakítások nem változtatják meg a sorvektorok által kifeszített alteret, és így a sorrangot sem (itt a 2.64. Definíciót használjuk a rangra). A 2.42. Tétel szerint az elemi sorátalakítások nem változtatják meg az oszlopvektorok közötti lineáris összefüggéseket, és így az oszloprangot sem (itt a 2.63. Definíciót használjuk a rangra). □

2.70. Tétel (a mátrixok rangszámtétele). Tetszőleges T test feletti $A \in T^{k\times n}$ mátrix sorrangja és oszloprangja megegyezik.

Bizonyítás. Hozzuk az A mátrixot elemi sorátalakításokkal redukált lépcsős alakra (lásd az 1.6. Tételt). A 2.69. Tétel szerint a kapott A' mátrixnak a sorrangja és az oszloprangja is ugyanaz, mint az eredeti A mátrixé. Az A' mátrix valahogy így fest:

$$A' = \begin{pmatrix} \boxed{1} & * & 0 & 0 & * & * & 0 & * & * & * \\ 0 & * & \boxed{1} & 0 & * & * & 0 & * & * & * \\ 0 & * & 0 & \boxed{1} & * & * & 0 & * & * & * \\ 0 & * & 0 & 0 & * & * & \boxed{1} & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Legyen a vezéregyesek száma r. Ez ugyanannyi, mint a „szép” oszlopok száma, és ugyanannyi, mint a nem csupa nulla sorok száma. (A fenti példában $r = 4$.) Világos, hogy az A' mátrix sorrangja r, hiszen első r sora lineárisan független, a többi meg csupa nulla. Az oszloprang is r, hiszen az r darab „szép” oszlop lineárisan független, a többi oszlop meg előáll ezek lineáris kombinációjaként. □

2.71. Megjegyzés. A sorrang és az oszloprang mellett lehet még definiálni mátrix determinánsrangját is, ezen a legnagyobb nemnulla aldetermináns méretét értjük. Meg lehet mutatni, hogy a determinánsrang megegyezik a sor- és oszlopranggal. A rangszámtételt lehet úgy is bizonyítani, hogy igazoljuk, hogy a sorrang megegyezik a determinánsranggal, majd használjuk azt a tényt, hogy a determináns(rang) nem változik, ha a mátrixot transzponáljuk (lásd [SzL] 9.2. Tétel).

2.72. Tétel. Tetszőleges T test feletti $A \in T^{n\times n}$ négyzetes mátrixra ekvivalensek az alábbiak:

(i) a sorvektorok lineárisan függetlenek;

(ii) az oszlopvektorok lineárisan függetlenek;

(iii) $\mathrm{r}(A) = n$;

(iv) $\det(A) \neq 0$;

(v) A-nak van inverze.

Bizonyítás. Az első három állítás ekvivalenciája rögtön következik a rangszámtételből (ugye?), az utolsó kettő ekvivalenciája pedig volt linalg1-ből (ugye?). Ezért elég (iii) és (iv) ekvivalenciáját igazolni. Ehhez hozzuk megint redukált lépcsős alakra a mátrixot elemi sorátalakításokkal. Eközben a rang nem változik, a determináns pedig legföljebb előjelet vált. Mivel $n \times n$-es mátrixunk van, két eset lehetséges a redukált lépcsős alakra.

- A vezéregyesek száma n. Ekkor a redukált lépcsős alak nem más, mint az $n \times n$-es egységmátrix, aminek rangja n, determinánsa pedig nem nulla.

- A vezéregyesek száma kisebb, mint n. Ekkor a redukált lépcsős alaknak van csupa nulla sora, így rangja kisebb, mint n, determinánsa pedig nulla. □

2.73. Definíció. A 2.72. Tétel ekvivalens feltételeinek eleget tevő mátrixokat nemelfajuló mátrixoknak nevezzük.

2.74. Tétel (Kronecker–Capelli-tétel). A T test feletti $A\mathbf{x} = \mathbf{b}$ lineáris egyenletrendszernek akkor és csak akkor van megoldása, ha mátrixának és kiterjesztett mátrixának rangja megegyezik, azaz $\mathrm{r}(A) = \mathrm{r}(A\,|\,\mathbf{b})$.

Bizonyítás. Az egyenletrendszernek akkor és csak akkor van megoldása, ha $\mathbf{b}$ előáll az A mátrix oszlopvektorainak lineáris kombinációjaként, ez pedig nyilván ekvivalens azzal, hogy a $\mathbf{b}$ vektort A oszlopvektoraihoz hozzácsapva a rang nem változik. □

2.75. Tétel. Legyenek A és B olyan mátrixok a T test felett, amelyekre az AB szorzat értelmezett. Ekkor teljesülnek az alábbiak:

(i) $\mathrm{r}(AB) \leq \mathrm{r}(A)$ és $\mathrm{r}(AB) \leq \mathrm{r}(B)$;

(ii) ha A nemelfajuló, akkor $\mathrm{r}(AB) = \mathrm{r}(B)$;

(iii) ha B nemelfajuló, akkor $\mathrm{r}(AB) = \mathrm{r}(A)$.

Bizonyítás. Lásd [SzL] 9.6. Tétel. □

2.76. Következmény. Hasonló mátrixok rangja megegyezik.

Bizonyítás. Lásd [SzL] 9.6. Tétel. □

3. Lineáris leképezések

3.1. A lineáris leképezés fogalma, nevezetes példák

3.1. Definíció. Legyenek U és V vektorterek ugyanazon T test felett. A $\varphi\colon U \to V$ leképezést lineáris leképezésnek nevezzük, ha rendelkezik az alábbi tulajdonságokkal:

(1) $\forall \mathbf{u}_1, \mathbf{u}_2 \in U\colon (\mathbf{u}_1 + \mathbf{u}_2)\varphi = \mathbf{u}_1\varphi + \mathbf{u}_2\varphi$;

(2) $\forall \mathbf{u} \in U\ \forall \lambda \in T\colon (\lambda\mathbf{u})\varphi = \lambda(\mathbf{u}\varphi)$;

Az összes $U \to V$ lineáris leképezések halmazát $\mathrm{Hom}(U, V)$ jelöli.
Speciális lineáris leképezések:

- ha $\varphi\colon V \to V$ lineáris leképezés, akkor azt mondjuk, hogy φ lineáris transzformációja a V vektortérnek;
- ha φ bijektív lineáris leképezés, akkor izomorfizmusnak nevezzük (lásd a 2.53. Definíciót).
- ha $\varphi\colon V \to V$ izomorfizmus, akkor azt mondjuk, hogy φ automorfizmusa a V vektortérnek;

3.2. Példa. Néhány példa lineáris leképezésre (U és V tetszőleges vektorterek a T test felett):

(i) $\underline{\underline{\mathbf{0}}}\colon U \to V,\ \mathbf{v} \to \mathbf{0}_V$ lineáris leképezés;

(ii) $\mathrm{id}\colon V \to V,\ \mathbf{v} \to \mathbf{v}$ automorfizmus;

(iii) $\varphi\colon V \to V,\ \mathbf{v} \to \lambda\mathbf{v}$ lineáris transzformáció tetszőleges $\lambda \in T$ esetén;

(iv) $\varphi\colon C[0,1] \to \mathbb{R},\ f \to \int_0^1 f(x)\,dx$ lineáris leképezés;

(v) $\varphi\colon T^{n\times m} \to T^{m\times n},\ A \to A^{\mathrm{T}}$ izomorfizmus;

(vi) a síknak (mint $\mathbb{R}$ feletti kétdimenziós vektortérnek) lineáris transzformációi (melyik automorfizmus?) a következők:

- az origó körüli forgatások,
- az origón átmenő tengelyre való tükrözések,

- az origó középpontú nyújtások,
- az origón átmenő egyenesekre való merőleges vetítések;

(vii) tetszőleges $A \in T^{n\times m}$ mátrix esetén $\varphi\colon T^n \to T^m$, $\mathbf{x} \to \mathbf{x}A$ lineáris leképezés.

3.3. Példa. Tekintsük az alábbi leképezést

$$\varphi\colon \mathbb{Z}_3^4 \to \mathbb{Z}_3^3,\ (x_1,\, x_2,\, x_3,\, x_4) \mapsto\ (x_1 + x_2 + 2x_3 + x_4,\ x_1 + 2x_2 + x_3,\ 2x_2 + x_3 + x_4).$$

Nem nehéz ellenőrizni, hogy ez a leképezés lineáris, de megúszhatjuk a számolást, ha felírjuk a leképezést mátrixszorzással:

$$(x_1, x_2, x_3, x_4)\varphi \;=\; (x_1, x_2, x_3, x_4) \cdot \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 2 \\ 2 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

Ebből látszik, hogy φ a 3.2. Példa (vii) pontjában leírt leképezések családjába tartozik, és így linearitása a mátrixszorzás megfelelő tulajdonságaiból következik (pontosan mely tulajdonságokból?). Később látjuk majd, hogy lényegében minden lineáris leképezés felírható ilyen formában.
A „sorvektor · mátrix” alak helyett felírható a leképezés „mátrix · oszlopvektor” alakban is (mindkét írásmódnak megvannak az előnyei és a hátrányai is):

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}\varphi \;=\; \begin{pmatrix} 1 & 1 & 2 & 1 \\ 1 & 2 & 1 & 0 \\ 0 & 2 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}.$$

Figyeljük meg, hogy a két mátrix egymás transzponáltja.

3.2. Műveletek lineáris leképezésekkel, izomorfia

3.4. Definíció. Legyenek U és V vektorterek a T test felett. A $\varphi,\, \psi \in \mathrm{Hom}(U, V)$ lineáris leképezések összegét a következőképpen definiáljuk:

$$\varphi + \psi\colon U \to V,\ \mathbf{u} \mapsto \mathbf{u}\varphi + \mathbf{u}\psi.$$

Tetszőleges $\lambda \in T$ esetén pedig így definiáljuk a φ linearis leképezés λ-szorosát:

$$\lambda\varphi\colon U \to V,\ \mathbf{u} \mapsto \lambda(\mathbf{u}\varphi).$$

3.5. Állítás. Legyenek U és V vektorterek a T test felett. Ha $\varphi,\, \psi \in \mathrm{Hom}(U, V)$ és $\lambda \in T$, akkor

(i) $\varphi + \psi \in \mathrm{Hom}(U, V)$;

(ii) $\lambda\varphi \in \mathrm{Hom}(U, V)$;

(iii) $\mathrm{Hom}(U, V)$ ezekkel a műveletekkel vektorteret alkot.

Bizonyítás. Lásd [SzL] 11.2. Tétel. □

3.6. Tétel. Legyenek U, V és W vektorterek a T test felett.

(i) Ha $\varphi \in \mathrm{Hom}(U, V)$ és $\psi \in \mathrm{Hom}(V, W)$, akkor $\varphi\psi \in \mathrm{Hom}(U, W)$.

(ii) Ha $\varphi \in \mathrm{Hom}(U, V)$ bijektív, akkor $\varphi^{-1} \in \mathrm{Hom}(V, U)$.

(iii) Ha $\varphi \in \mathrm{Hom}(U, V)$ és $\psi \in \mathrm{Hom}(V, W)$, akkor $\lambda(\varphi\psi) = (\lambda\varphi)\psi = \varphi(\lambda\psi)$ teljesül minden $\lambda \in T$ skalárra.

(iv) Ha $\varphi,\, \psi \in \mathrm{Hom}(U, V)$ és $\tau \in \mathrm{Hom}(V, W)$, akkor $(\varphi + \psi)\tau = \varphi\tau + \psi\tau$.

(v) Ha $\varphi \in \mathrm{Hom}(U, V)$ és $\psi,\, \tau \in \mathrm{Hom}(V, W)$, akkor $\varphi(\psi + \tau) = \varphi\psi + \varphi\tau$.

Bizonyítás. Lásd [SzL] 10.7. Tétel és 11.3. Tétel. □

3.7. Következmény. Az izomorfia vektorterek bármely halmazán ekvivalenciareláció.

Bizonyítás. Lásd [SzL] 10.9. Tétel. □

3.8. Tétel. Ha U és V véges dimenziós vektorterek a T test felett, akkor

$$U \cong V \iff \dim U = \dim V.$$

Bizonyítás. Legyen $\dim U = n$ és $\dim V = m$.

- $\Longrightarrow$: Ha $\varphi\colon U \to V$ izomorfizmus és $\mathbf{e}_1, \ldots, \mathbf{e}_n$ bázisa U-nak, akkor $\mathbf{e}_1\varphi, \ldots, \mathbf{e}_n\varphi$ bázisa V-nek[a], ezért $\dim V = n$.
- $\Longleftarrow$: Láttuk korábban (2.54. Tétel), hogy $U \cong T^n$ és $V \cong T^m$. Ha $n = m$, akkor az izomorfia szimmetriája és tranzitivitása miatt következik, hogy $U \cong V$. □

[a]Ezt nem nehéz ellenőrizni (lásd pl. [SzL] 10.10. Tétel), de tulajdonképpen triviális, hogy az izomorfizmusok megőriznek minden algebrai tulajdonságot, mint pl. a lineáris függetlenség vagy a „generátorrendszerség”.

3.3. Képtér és magtér, projekciók

3.9. Definíció. Legyenek U és V vektorterek a T test felett, és legyen $\varphi\colon U \to V$ lineáris leképezés.

- Azon U-beli vektorok halmazát, melyekhez φ a nullvektort rendeli, φ magterének nevezzük. Jelölés: $\operatorname{Ker}\varphi$.
- Azon V-beli vektorok halmazát, melyek előállnak φ melletti képként, φ képterének nevezzük. Jelölés: $\operatorname{Im}\varphi$.

Formálisan:

$$\operatorname{Ker}\varphi = \{\mathbf{u} \in U : \mathbf{u}\varphi = \mathbf{0}_V\} \leq U, \qquad \operatorname{Im}\varphi = \{\mathbf{v} \in V : \exists \mathbf{u} \in U\colon \mathbf{u}\varphi = \mathbf{v}\} = \{\mathbf{u}\varphi : \mathbf{u} \in U\} \leq V.$$

3.10. Állítás. Legyenek U és V vektorterek a T test felett. Tetszőleges $\varphi \in \operatorname{Hom}(U, V)$ lineáris leképezésre teljesülnek az alábbiak:

(i) $\mathbf{0}_U\varphi = \mathbf{0}_V$;

(ii) $\operatorname{Ker}\varphi$ altér U-ban;

(iii) $\operatorname{Im}\varphi$ altér V-ben.

Bizonyítás. Lásd [SzL] 10.2. Tétel. □

3.11. Példa. Tekintsük az alábbi lineáris leképezést

$$\varphi\colon \mathbb{Z}_2^4 \to \mathbb{Z}_2^3,\ (x_1, x_2, x_3, x_4) \mapsto (x_1 + x_2 + x_3,\ x_2 + x_3 + x_4,\ x_1 + x_4).$$

Írjuk fel minden egyes vektor képét:

$$\begin{array}{lll lll lll lll}
0000 & \mapsto & 000 & 1000 & \mapsto & 101 & 0100 & \mapsto & 110 & 0001 & \mapsto & 011 \\
0110 & \mapsto & 000 & 1110 & \mapsto & 101 & 0010 & \mapsto & 110 & 0111 & \mapsto & 011 \\
1011 & \mapsto & 000 & 0011 & \mapsto & 101 & 1111 & \mapsto & 110 & 1010 & \mapsto & 011 \\
1101 & \mapsto & 000 & 0101 & \mapsto & 101 & 1001 & \mapsto & 110 & 1100 & \mapsto & 011
\end{array}$$

Innen kiolvasható, hogy

$$\operatorname{Ker}\varphi = \{0000,\ 0110,\ 1011,\ 1101\}, \qquad \operatorname{Im}\varphi = \{000,\ 101,\ 110,\ 011\}.$$

3.12. Állítás. Legyenek U és V vektorterek a T test felett. Tetszőleges $\varphi \in \operatorname{Hom}(U, V)$ lineáris leképezés és $\mathbf{u}_1, \mathbf{u}_2 \in U$ vektorok esetén

$$\mathbf{u}_1\varphi = \mathbf{u}_2\varphi \iff \mathbf{u}_1 - \mathbf{u}_2 \in \operatorname{Ker}\varphi.$$

Bizonyítás. A leképezés linearitásából és a mag definíciójából rögtön következik az állítás:

$$\mathbf{u}_1\varphi = \mathbf{u}_2\varphi \iff \mathbf{u}_1\varphi - \mathbf{u}_2\varphi = \mathbf{0} \iff (\mathbf{u}_1 - \mathbf{u}_2)\varphi = \mathbf{0} \iff \mathbf{u}_1 - \mathbf{u}_2 \in \operatorname{Ker}\varphi.$$ □

3.13. Következmény. Legyenek U és V vektorterek a T test felett. Tetszőleges $\varphi \in \operatorname{Hom}(U, V)$ lineáris leképezés és $\mathbf{u}_1, \mathbf{u}_2 \in U$ vektorok esetén

(i) φ akkor és csak akkor injektív, ha $\operatorname{Ker}\varphi = \{\mathbf{0}\}$;

(ii) φ akkor és csak akkor szürjektív, ha $\operatorname{Im}\varphi = V$.

3.14. Állítás. Legyenek U és V vektorterek a T test felett. Ha $\varphi \in \operatorname{Hom}(U, V)$ és $\mathbf{u}_1, \ldots, \mathbf{u}_k$ generátorrendszere V-nek, akkor $\mathbf{u}_1\varphi, \ldots, \mathbf{u}_k\varphi$ generátorrendszere $\operatorname{Im}\varphi$-nek.

Bizonyítás. Lásd [SzL] 10.2. Tétel. □

3.15. Tétel (a lineáris leképezések dimenziótétele). Legyenek U és V vektorterek a T test felett. Ha $\varphi \in \mathrm{Hom}(U,V)$ és U véges dimenziós, akkor
$$\dim U = \dim \mathrm{Ker}\,\varphi + \dim \mathrm{Im}\,\varphi.$$

Bizonyítás. Lásd [SzL] 10.3. Tétel. □

3.16. Definíció. A képtér dimenzióját a lineáris leképezés rangjának nevezzük, a magtér dimenzióját pedig a lineáris leképezés nullitásának. Ezért a fenti tételt szokás rang plusz nullitás tételnek is hívni.

3.17. Következmény. Ha V véges dimenziós vektortér a T test felett és $\varphi \in \mathrm{Hom}(V,V)$, akkor
$$\varphi \text{ injektív} \iff \varphi \text{ szürjektív}.$$

Bizonyítás. Lásd [SzL] 10.4. Következmény. □

3.18. Tétel. Tetszőleges T test és $A \in T^{m\times n}$ mátrix esetén az $A\mathbf{x} = \mathbf{0}$ homogén lineáris egyenletrendszer megoldásterének dimenziója $n - \mathrm{r}(A)$.

Bizonyítás. Tekintsük a $\varphi\colon T^n \to T^m$, $\mathbf{x} \to A\mathbf{x}$ lineáris leképezést (lásd a 3.2. Példa (vii) pontját és a 3.3. Példát). Ennek magja éppen az $A\mathbf{x} = \mathbf{0}$ homogén lineáris egyenletrendszer megoldástere, aminek dimenziója a rang plusz nullitás tétel (3.15. Tétel) szerint $n - \dim \mathrm{Im}\,\varphi$. Tehát elegendő azt igazolni, hogy $\dim \mathrm{Im}\,\varphi = \mathrm{r}(A)$. Ehhez pedig csak azt kell észrevenni, hogy $A\mathbf{x}$ nem más, mint az A mátrix oszlopainak lineáris kombinációja ($x_1, \dots, x_n$ együtthatókkal). Mivel $\mathrm{Im}\,\varphi$ éppen ezekből a vektorokból áll, $\mathrm{Im}\,\varphi$ megegyezik az A mátrix oszlopvektorai által generált altérrel, aminek dimenziója valóban A (oszlop)rangja. □

3.19. Következmény. Tetszőleges T test és $A \in T^{n\times n}$ négyzetes mátrix esetén az $A\mathbf{x} = \mathbf{0}$ homogén lineáris egyenletrendszernek akkor és csak akkor van nemtriviális megoldása, ha $\det(A) = 0$.

Bizonyítás. Akkor és csak akkor van nemtriviális megoldás, ha a megoldástér dimenziója nagyobb, mint nulla. A fenti tétek szerint ez ekvivalens azzal, hogy $\mathrm{r}(A) < n$, ez pedig pontosan akkor teljesül, ha $\det(A) = 0$ (lásd a 2.72. Tételt). □

3.20. Definíció. Legyen $V = U \oplus W$ egy direktösszeg-felbontása a V vektortérnek, és legyen π az alábbi lineáris transzformáció (felhasználjuk, hogy V minden eleme egyértelműen áll elő $\mathbf{u} + \mathbf{w}$ ($\mathbf{u} \in U$, $\mathbf{w} \in W$) alakban):
$$\pi\colon V \to V,\ \mathbf{u} + \mathbf{w} \mapsto \mathbf{u} \quad (\mathbf{u} \in U,\ \mathbf{w} \in W).$$
Ezt a π leképezést az U-ra történő, W-vel párhuzamos irányú projekciónak vagy vetítésnek nevezzük. (Ellenőrizzük, hogy π valóban lineáris leképezés, továbbá $\mathrm{Im}\,\pi = U$ és $\mathrm{Ker}\,\pi = W$.)

3.21. Tétel. Tetszőleges V vektortér és $\pi \in \mathrm{Hom}(V,V)$ lineáris transzformáció esetén
$$\pi \text{ projekció} \iff \pi^2 = \pi.$$

Bizonyítás. Előkészületként gondoljuk meg, hogy a $\pi^2 = \pi$ tulajdonság (idempotencia) azt jelenti, hogy π identikusan hat az értékkészletén: $\forall \mathbf{u} \in \mathrm{Im}\,\pi\colon \mathbf{u}\pi = \mathbf{u}$ (HF). Ez nemcsak vektorterek lineáris transzformációira igaz, hanem tetszőleges halmaz tetszőleges transzformációira is.

- $\implies$: Ha π projekció, akkor a 3.20. Definíció jelöléseit használva világos, hogy $\mathbf{u}\pi = \mathbf{u}$ teljesül minden $\mathbf{u} \in U = \mathrm{Im}\,\pi$ esetén. A fenti előkészítő észrevétel szerint ebből következik, hogy $\pi^2 = \pi$.

- $\impliedby$: Tfh. $\pi^2 = \pi$, és a 3.20. Definícióból „puskázva” vezessük be az $U := \mathrm{Im}\,\pi$ és $W := \mathrm{Ker}\,\pi$ jelöléseket. A fenti előkészítő észrevétel szerint ekkor $\mathbf{u}\pi = \mathbf{u}$ teljesül minden $\mathbf{u} \in U$ vektorra, és így
$$\forall \mathbf{u} \in U\ \forall \mathbf{w} \in W\colon\ (\mathbf{u} + \mathbf{w})\pi = \mathbf{u}\pi + \mathbf{w}\pi = \mathbf{u} + \mathbf{0} = \mathbf{u}.$$
Azt kell még belátnunk, hogy $V = U \oplus W$, ami a direkt összeg 2.59 Tételbeli (i) tulajdonsága szerint azt jelenti, hogy V minden eleme egyértelműen előáll $\mathbf{u}+\mathbf{w}$ ($\mathbf{u} \in U$, $\mathbf{w} \in W$) alakban. Először az egyértelműséget igazoljuk. Tfh. $\mathbf{v} = \mathbf{u} + \mathbf{w}$, ahol $\mathbf{u} \in U$ és $\mathbf{w} \in W$. Mindkét oldalra alkalmazva a π leképezést, azt kapjuk, hogy $\mathbf{v}\pi = \mathbf{u}$. Tehát a $\mathbf{v}$ vektor $\mathbf{u} + \mathbf{w}$ alakú felbontásában szükségképpen $\mathbf{u} = \mathbf{v}\pi$, és ekkor persze $\mathbf{w} = \mathbf{v} - \mathbf{v}\pi$. Ezzel bebizonyítottuk az unicitást, az egzisztenciához pedig csak azt kell ellenőrizni, hogy

a $\mathbf{v} = \underbrace{\mathbf{v}\pi}_{\in U?} + \underbrace{(\mathbf{v} - \mathbf{v}\pi)}_{\in W?}$ felbontás megfelelő. Az világos, hogy $\mathbf{v}\pi \in U = \operatorname{Im}\pi$, és π idempotenciájából következik az is, hogy $(\mathbf{v} - \mathbf{v}\pi)\pi \in W = \operatorname{Ker}\pi$:

$$(\mathbf{v} - \mathbf{v}\pi)\pi = \mathbf{v}\pi - \mathbf{v}\pi^2 = \mathbf{v}\pi - \mathbf{v}\pi = \mathbf{0}.$$

□

3.4. Lineáris leképezés mátrixa

3.22. Definíció. Legyenek U és V vektorterek a T test felett, legyen $\mathcal{E}\colon \mathbf{e}_1, \ldots, \mathbf{e}_n$ bázisa U-nak és $\mathcal{F}\colon \mathbf{f}_1, \ldots, \mathbf{f}_m$ bázisa V-nek. Egy $\varphi \in \operatorname{Hom}(U, V)$ lineáris leképezés mátrixán (az $\mathcal{E}$ és $\mathcal{F}$ bázisokban) azt a $[\![\varphi]\!]_{\mathcal{E},\mathcal{F}} \in T^{n\times m}$ mátrixot értjük, melyben az i-edik sor j-edik eleme az $\mathbf{e}_i\varphi$ vektor j-edik koordinátája az $\mathcal{F}$ bázisban. Formálisan: legyen $\mathbf{e}_i\varphi = \sum_{j=1}^{m} a_{ij}\mathbf{f}_j$; ekkor $[\![\varphi]\!]_{\mathcal{E},\mathcal{F}} = (a_{ij})_{n\times m}$. Ha $U = V$ és $\mathcal{E} = \mathcal{F}$, akkor $[\![\varphi]\!]_{\mathcal{E},\mathcal{F}}$ helyett használjuk az egyszerűsített $[\![\varphi]\!]_{\mathcal{E}}$ jelölést is.

3.23. Megjegyzés. A $[\![\varphi]\!]_{\mathcal{E},\mathcal{F}}$ mátrixot megkaphatjuk úgy, hogy egymás alá leírjuk az $[\![\mathbf{e}_1\varphi]\!]_{\mathcal{F}}, \ldots, [\![\mathbf{e}_n\varphi]\!]_{\mathcal{F}}$ vektorokat.

3.24. Tétel. Legyenek U és V vektorterek a T test felett, legyen $\mathcal{E}\colon \mathbf{e}_1, \ldots, \mathbf{e}_n$ bázisa U-nak és $\mathcal{F}\colon \mathbf{f}_1, \ldots, \mathbf{f}_m$ bázisa V-nek, és legyen $\varphi \in \operatorname{Hom}(U, V)$. Ekkor tetszőleges $\mathbf{u} \in U$ vektorra

$$[\![\mathbf{u}\varphi]\!]_{\mathcal{F}} = [\![\mathbf{u}]\!]_{\mathcal{E}} \cdot [\![\varphi]\!]_{\mathcal{E},\mathcal{F}}.$$

Bizonyítás. Legyen $[\![\mathbf{u}]\!]_{\mathcal{E}} = \mathbf{x} = (x_1, \ldots, x_n)$ és $[\![\mathbf{u}\varphi]\!]_{\mathcal{F}} = \mathbf{y} = (y_1, \ldots, y_m)$, és az egyszerűség kedvéért vezessük be a $[\![\varphi]\!]_{\mathcal{E},\mathcal{F}} = A = (a_{ij})$ jelölést is. A bizonyítandó állítás ezekkel a jelölésekkel így fest: $\mathbf{y} = \mathbf{x}A$. Írjuk fel az $\mathbf{u}$ vektort az $\mathcal{E}$ bázisban, alkalmazzuk rá a φ leképezést, majd írjuk fel a megjelenő $\mathbf{e}_i\varphi$ vektorokat az $\mathcal{F}$ bázisban:

$$\mathbf{u}\varphi = \left(\sum_{i=1}^{n} x_i\mathbf{e}_i\right)\varphi = \sum_{i=1}^{n} x_i(\mathbf{e}_i\varphi) = \sum_{i=1}^{n} x_i\left(\sum_{j=1}^{m} a_{ij}\mathbf{f}_j\right) = \sum_{i,j} x_i a_{ij}\mathbf{f}_j.$$

Ezzel megkaptuk az $\mathbf{u}\varphi$ vektor felírását az $\mathcal{F}$ bázisban. Az $\mathbf{f}_j$ bázisvektor együtthatója $y_j = \sum_i x_i a_{ij}$, ami a mátrixszorzás definíciója szerint valóban ugyanaz, mint az $\mathbf{x}A$ sorvektor j-edik komponense. □

3.25. Példa. Legyen $\mathcal{E}$ az $U = \mathbb{Z}_3^4$ vektortér standard bázisa és legyen $\mathcal{F}$ a $V = \mathbb{Z}_3^3$ vektortér standard bázisa. Ekkor a 3.3. Példában tekintett $\varphi \in \operatorname{Hom}(U, V)$ lineáris leképezés mátrixa ezekben a bázisokban:

$$[\![\varphi]\!]_{\mathcal{E},\mathcal{F}} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 2 \\ 2 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

3.26. feladat. Írjuk fel az $y = 2x$ egyenletű egyenesre való tükrözésnek (mint $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2$ lineáris transzformációnak) a mátrixát az $\mathcal{F}\colon \mathbf{f}_1, \mathbf{f}_2$ bázisban, majd az $\mathcal{E}\colon \mathbf{e}_1, \mathbf{e}_2$ bázisban is, ahol

$$\begin{aligned} \mathbf{f}_1 &= (1, 2), & \mathbf{f}_2 &= (-2, 1); \\ \mathbf{e}_1 &= (1, 0), & \mathbf{e}_2 &= (0, 1). \end{aligned}$$

Megoldás. Számítsuk ki az $\mathcal{F}$ bázis elemeinek képeit és a képek koordinátáit az $\mathcal{F}$ bázisban. Mivel $\mathbf{f}_1$ a tükrözés tengelyére esik, $\mathbf{f}_2$ pedig merőleges a tengelyre, nem nehéz meghatározni a tükörképeiket:

$$\begin{aligned} \mathbf{f}_1\varphi &= \mathbf{f}_1 = 1\,\mathbf{f}_1 + 0\,\mathbf{f}_2 \implies [\![\mathbf{f}_1\varphi]\!]_{\mathcal{F}} = (1, 0) \\ \mathbf{f}_2\varphi &= -\mathbf{f}_2 = 0\,\mathbf{f}_1 + (-1)\,\mathbf{f}_2 \implies [\![\mathbf{f}_2\varphi]\!]_{\mathcal{F}} = (0, -1) \end{aligned}$$

A koordinátasorokat egymás alá írva kapjuk φ mátrixát az $\mathcal{F}, \mathcal{F}$ bázispárban:

$$[\![\varphi]\!]_{\mathcal{F}} = [\![\varphi]\!]_{\mathcal{F},\mathcal{F}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

A $[\![\varphi]\!]_{\mathcal{E}}$ mátrix felírásához az $\mathbf{e}_1\varphi$ és $\mathbf{e}_2\varphi$ vektorokat kell meghatároznunk. Ehhez hasznos lesz kifejezni az $\mathbf{e}_1, \mathbf{e}_2$ vektorokat $\mathbf{f}_1, \mathbf{f}_2$ segítségével (hiszen az utóbbiak képeit már ismerjük). Ez két kis könnyű egyenletrendszer megoldását jelenti; a számolást mellőzzük, íme az eredmény:

$$\mathbf{e}_1 = \frac{1}{5}\mathbf{f}_1 - \frac{2}{5}\mathbf{f}_2, \quad \mathbf{e}_2 = \frac{2}{5}\mathbf{f}_1 + \frac{1}{5}\mathbf{f}_2.$$

Ennek alapján ki tudjuk számolni az $[\![\mathbf{e}_1\varphi]\!]_\mathcal{E}$ és $[\![\mathbf{e}_2\varphi]\!]_\mathcal{E}$ koordinátasorokat:

$$\begin{aligned}
\mathbf{e}_1\varphi &= \tfrac{1}{5}\mathbf{f}_1\varphi - \tfrac{2}{5}\mathbf{f}_2\varphi = \tfrac{1}{5}\mathbf{f}_1 + \tfrac{2}{5}\mathbf{f}_2 = (-3/5, 4/5) = -\tfrac{3}{5}\,\mathbf{e}_1 + \tfrac{4}{5}\,\mathbf{e}_2 \implies [\![\mathbf{e}_1\varphi]\!]_\mathcal{E} = (-3/5, 4/5)\\
\mathbf{e}_2\varphi &= \tfrac{2}{5}\mathbf{f}_1\varphi + \tfrac{1}{5}\mathbf{f}_2\varphi = \tfrac{2}{5}\mathbf{f}_1 - \tfrac{1}{5}\mathbf{f}_2 = (4/5, 3/5) = \tfrac{4}{5}\,\mathbf{e}_1 + \tfrac{3}{5}\,\mathbf{e}_2 \implies [\![\mathbf{e}_2\varphi]\!]_\mathcal{E} = (4/5, 3/5)
\end{aligned}$$

A koordinátasorokat egymás alá írva kapjuk φ mátrixát az $\mathcal{E},\mathcal{E}$ bázispárban:

$$[\![\varphi]\!]_\mathcal{E} = [\![\varphi]\!]_{\mathcal{E},\mathcal{E}} = \begin{pmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{pmatrix}.$$

Így nyertünk egy képletet tetszőleges (x,y) pont tükörképének koordinátáira:

$$(x,y)\varphi = (x,y)\cdot\begin{pmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{pmatrix} = \left(-\tfrac{3}{5}x + \tfrac{4}{5}y,\ \tfrac{4}{5}x + \tfrac{3}{5}y\right).$$

Megjegyzés: Ha oszlopvektorokkal szeretnénk dolgozni, akkor a koordinátasorokat oszlopokként kellene egymás mellé írni:

$$\begin{pmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{pmatrix}.$$

A két mátrix „véletlenül” ugyanaz, de a színek jelzik, hogy másképpen vannak elrendezve a bázisvektorok képeinek koordinátái. ◇

3.27. Tétel (bázis képei: kívánságműsor). Legyenek U és V vektorterek a T test felett, legyen $\mathcal{E}\colon \mathbf{e}_1,\dots,\mathbf{e}_n$ bázisa U-nak. Tetszőleges $\mathbf{v}_1,\dots,\mathbf{v}_n \in V$ vektorok esetén pontosan egy olyan $\varphi \in \operatorname{Hom}(U,V)$ lineáris leképezés létezik, amelyre $\mathbf{e}_i\varphi = \mathbf{v}_i$ $(i = 1,\dots,n)$.

Bizonyítás. Lásd [SzL] 12.1. Tétel. □

3.28. Megjegyzés (minden mtx linlekmtx). A 3.23. Megjegyzés tükrében a fenti tétel így is megfogalmazható: ha rögzítünk az U vektortérben egy $\mathcal{E}$ bázist és a V vektortérben egy $\mathcal{F}$ bázist, akkor minden $A \in T^{n\times m}$ mátrixhoz létezik pontosan egy $\varphi \in \operatorname{Hom}(U,V)$ lineáris leképezés, amelyre $[\![\varphi]\!]_{\mathcal{E},\mathcal{F}} = A$.

3.29. Tétel (összhang). Legyenek $\mathcal{E}$, $\mathcal{F}$ és $\mathcal{G}$ rendre bázisai az U, V és W ugyanazon T test feletti végesdimenziós vektortereknek. Tetszőleges $\varphi,\psi \in \operatorname{Hom}(U,V)$ és $\tau \in \operatorname{Hom}(V,W)$ lineáris leképezések és $\lambda \in T$ skalár esetén

$$[\![\varphi+\psi]\!]_{\mathcal{E},\mathcal{F}} = [\![\varphi]\!]_{\mathcal{E},\mathcal{F}} + [\![\psi]\!]_{\mathcal{E},\mathcal{F}}, \qquad [\![\lambda\varphi]\!]_{\mathcal{E},\mathcal{F}} = \lambda[\![\varphi]\!]_{\mathcal{E},\mathcal{F}}, \qquad [\![\varphi\tau]\!]_{\mathcal{E},\mathcal{G}} = [\![\varphi]\!]_{\mathcal{E},\mathcal{F}}\cdot[\![\tau]\!]_{\mathcal{F},\mathcal{G}}.$$

Bizonyítás. Lásd [SzL] 12.3. Tétel. □

3.30. Következmény. Legyenek $\mathcal{E}$ és $\mathcal{F}$ bázisai az U és V ugyanazon T test feletti végesdimenziós vektortereknek, és legyen $\dim U = n$, $\dim V = m$. Ekkor az alábbi Φ leképezés vektortér-izomorfizmus:

$$\Phi\colon \operatorname{Hom}(U,V) \to T^{n\times m},\ \varphi \mapsto [\![\varphi]\!]_{\mathcal{E},\mathcal{F}}.$$

Következésképp a $\operatorname{Hom}(U,V)$ halmaz nm-dimenziós vektorteret alkot T felett.

Bizonyítás. Lásd [SzL] 12.4. Következmény. □

3.31. Következmény. Legyen V egy n-dimenziós vektortér a T test felett és legyen $\mathcal{E}$ bázisa V-nek. Ekkor az alábbi Φ leképezés gyűrűizomorfizmus:

$$\Phi\colon \operatorname{Hom}(V,V) \to T^{n\times n},\ \varphi \mapsto [\![\varphi]\!]_\mathcal{E}.$$

Következésképp a két gyűrű egységcsoportjai is izomorfak:

$$\operatorname{Aut}(V) \cong \operatorname{GL}_n(T) = \left\{A \in T^{n\times n} : \det(A) \neq 0\right\},$$

és így egy lineáris transzformáció akkor és csak akkor bijektív, ha valamely/bármely bázisbeli mátrixa nemelfajuló.

3.32. Tétel. Legyenek U és V vektorterek a T test felett, legyen $\mathcal{E}\colon \mathbf{e}_1,\dots,\mathbf{e}_n$ bázisa U-nak és $\mathcal{F}\colon \mathbf{f}_1,\dots,\mathbf{f}_m$ bázisa V-nek. Tetszőleges $\varphi \in \operatorname{Hom}(U,V)$ lineáris leképezés rangja ugyanaz, mint a $[\![\varphi]\!]_{\mathcal{E},\mathcal{F}}$ mátrix rangja.

Bizonyítás. Mivel $V \cong T^m$ a $\mathbf{v} \mapsto [\![\mathbf{v}]\!]_\mathcal{F}$ izomorfizmus mellett, $\operatorname{Im}\varphi$ dimenziója ugyanaz, mint az $\operatorname{Im}\varphi$-beli vektorok ($\mathcal{F}$ bázisra vonatkozó) koordinátasorai alkotta T^m-beli altér dimenziója. Ezek a koordinátasorok a 3.24. Tétel szerint pontosan az $[\![\mathbf{u}]\!]_{\mathcal{E}} \cdot [\![\varphi]\!]_{\mathcal{E},\mathcal{F}}$ ($\mathbf{u} \in U$) alakú vektorok. Itt $[\![\mathbf{u}]\!]_{\mathcal{E}}$ tetszőleges T^n-beli vektor lehet, tehát az összes $\mathbf{x} \cdot [\![\varphi]\!]_{\mathcal{E},\mathcal{F}}$ ($\mathbf{x} \in T^n$) alakú vektorok terének dimenziója a kérdés. Ezek pedig éppen a $[\![\varphi]\!]_{\mathcal{E},\mathcal{F}}$ mátrix sorainak lineáris kombinációi. Ezzel beláttuk, hogy $\dim \operatorname{Im} \varphi$ nem más, mint a $[\![\varphi]\!]_{\mathcal{E},\mathcal{F}}$ mátrix (sor)rangja. □

3.5. Áttérés új bázisra, mátrixok hasonlósága

3.33. Definíció. Legyen V vektortér a T test felett és legyen $\mathcal{E}\colon \mathbf{e}_1, \dots, \mathbf{e}_n$ és $\mathcal{F}\colon \mathbf{f}_1, \dots, \mathbf{f}_n$ két bázisa V-nek. Legyen a_{ij} az $\mathbf{e}_i$ vektor j-edik koordinátája az $\mathcal{F}$ bázisban: $\mathbf{e}_i = \sum_{j=1}^{n} a_{ij}\mathbf{f}_j$. Ekkor az $[\![\mathcal{E} \to \mathcal{F}]\!] := (a_{ij})_{n\times m}$ mátrixot az $\mathcal{E}$ bázisról az $\mathcal{F}$ bázisra való bázisáttérés mátrixának nevezzük.

3.34. Megjegyzés. Az $[\![\mathcal{E} \to \mathcal{F}]\!]$ mátrixot megkaphatjuk úgy, hogy egymás alá leírjuk az $[\![\mathbf{e}_1]\!]_{\mathcal{F}}, \dots, [\![\mathbf{e}_n]\!]_{\mathcal{F}}$ vektorokat, azaz $[\![\mathcal{E} \to \mathcal{F}]\!] = [\![\mathrm{id}_V]\!]_{\mathcal{E},\mathcal{F}}$.

3.35. Állítás. Legyen V vektortér a T test felett és legyen $\mathcal{E}\colon \mathbf{e}_1, \dots, \mathbf{e}_n$, $\mathcal{F}\colon \mathbf{f}_1, \dots, \mathbf{f}_n$ és $\mathcal{G}\colon \mathbf{g}_1, \dots, \mathbf{g}_n$ három bázisa V-nek. Ekkor

(i) $\forall \mathbf{v} \in V\colon [\![\mathbf{v}]\!]_{\mathcal{F}} = [\![\mathbf{v}]\!]_{\mathcal{E}} \cdot [\![\mathcal{E} \to \mathcal{F}]\!]$;

(ii) $[\![\mathcal{E} \to \mathcal{F}]\!] \cdot [\![\mathcal{F} \to \mathcal{G}]\!] = [\![\mathcal{E} \to \mathcal{G}]\!]$;

(iii) $[\![\mathcal{F} \to \mathcal{E}]\!] = [\![\mathcal{E} \to \mathcal{F}]\!]^{-1}$.

Bizonyítás. Mindhárom állítás bizonyításának kulcsa az, hogy a 3.34. Megjegyzés segítségével visszavezetjük (identikus) lineáris leképezésekre.

(i) Alkalmazzuk a 3.24. Tételt a $\varphi = \mathrm{id}_V$ leképezésre: $[\![\mathbf{v}]\!]_{\mathcal{F}} = [\![\mathbf{v}\,\mathrm{id}_V]\!]_{\mathcal{F}} = [\![\mathbf{v}]\!]_{\mathcal{E}} \cdot [\![\mathrm{id}_V]\!]_{\mathcal{E},\mathcal{F}} = [\![\mathbf{v}]\!]_{\mathcal{E}} \cdot [\![\mathcal{E} \to \mathcal{F}]\!]$.

(ii) A 3.29. Tételt használjuk az $U = V = W$, $\varphi = \tau = \mathrm{id}_V$ szereposztással:

$$[\![\mathcal{E} \to \mathcal{F}]\!] \cdot [\![\mathcal{F} \to \mathcal{G}]\!] = [\![\mathrm{id}_V]\!]_{\mathcal{E},\mathcal{F}} \cdot [\![\mathrm{id}_V]\!]_{\mathcal{F},\mathcal{G}} = [\![\mathrm{id}_V\,\mathrm{id}_V]\!]_{\mathcal{E},\mathcal{G}} = [\![\mathrm{id}_V]\!]_{\mathcal{E},\mathcal{G}} = [\![\mathcal{E} \to \mathcal{G}]\!].$$

(iii) Ez következik a (ii) állításból a $\mathcal{G} = \mathcal{E}$ speciális esetben: $[\![\mathcal{E} \to \mathcal{F}]\!] \cdot [\![\mathcal{F} \to \mathcal{E}]\!] = [\![\mathcal{E} \to \mathcal{E}]\!] = E_n$. □

3.36. Tétel. Legyen V egy n-dimenziós vektortér a T test felett és legyen $Q \in T^{n\times n}$. Akkor és csak akkor léteznek olyan $\mathcal{E}$ és $\mathcal{F}$ bázisok V-ben, hogy $Q = [\![\mathcal{E} \to \mathcal{F}]\!]$, ha Q nemelfajuló.

Bizonyítás. Az előző állításból rögtön következik, hogy ha $Q = [\![\mathcal{E} \to \mathcal{F}]\!]$, akkor Q nemelfajuló. Fordítva, tfh. Q nemelfajuló és rögzítsünk egy tetszőleges $\mathcal{F}\colon \mathbf{f}_1, \dots, \mathbf{f}_n$ bázist V-ben. Legyen $\mathbf{e}_i$ az a vektor, melynek koordinátasora az $\mathcal{F}$ bázisban megegyezik a Q mátrix i-edik sorával ($i = 1, \dots, n$). Mivel Q nemelfajuló, sorai bázist alkotnak a T^n vektortérben, így a $V \to T^n$, $\mathbf{v} \mapsto [\![\mathbf{v}]\!]_{\mathcal{F}}$ izomorfizmus melletti ősképeik bázist alkotnak V-ben. Ezek az ősképek nem mások, mint az $\mathbf{e}_1, \dots, \mathbf{e}_n$ vektorok, így tehát $\mathcal{E}\colon \mathbf{e}_1, \dots, \mathbf{e}_n$ bázisa V-nek. Ezzel kész is a bizonyítás, hiszen a Q mátrixot pont úgy konstruáltuk, hogy $Q = [\![\mathcal{E} \to \mathcal{F}]\!]$ teljesüljön. □

3.37. feladat. Adjuk meg az $[\![\mathcal{E} \to \mathcal{F}]\!]$ és $[\![\mathcal{F} \to \mathcal{E}]\!]$ bázisáttérés-mátrixokat az $\mathbb{R}^2$ vektortér alábbi $\mathcal{F}\colon \mathbf{f}_1, \mathbf{f}_2$ és $\mathcal{E}\colon \mathbf{e}_1, \mathbf{e}_2$ bázisaira:

$$\begin{aligned} \mathbf{f}_1 &= (1,2), \quad \mathbf{f}_2 = (-2,1); \\ \mathbf{e}_1 &= (1,0), \quad \mathbf{e}_2 = (0,1). \end{aligned}$$

Megoldás. Az $[\![\mathcal{F} \to \mathcal{E}]\!]$ mátrixot könnyű felírni, mert ehhez csak az $\mathcal{F}$ bázis elemeit kell a standatd $\mathcal{E}$ bázisban felírni (és abban már fel is vannak írva a feladatban):

$$[\![\mathcal{F} \to \mathcal{E}]\!] = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}.$$

Az $[\![\mathcal{E} \to \mathcal{F}]\!]$ mátrixot a fenti mátrix inverzeként kapjuk:

$$[\![\mathcal{E} \to \mathcal{F}]\!] = [\![\mathcal{F} \to \mathcal{E}]\!]^{-1} = \begin{pmatrix} 1/5 & -2/5 \\ 2/5 & 1/5 \end{pmatrix}.$$

Megjegyzés: A 3.26. Feladatban kiszámoltuk, hogy $[\![\mathbf{e}_1]\!]_{\mathcal{F}} = (1/5, -2/5)$ és $[\![\mathbf{e}_2]\!]_{\mathcal{F}} = (2/5, 1/5)$. E két koordinátasort egymás alá írva is megkaphattuk volna az $[\![\mathcal{E} \to \mathcal{F}]\!]$ mátrixot. Ha már megvan a bázisáttérés mátrixa, akkor könnyen felírhatjuk egy tetszőleges $\mathbf{v} = (x, y)$ vektor koordinátáit az $\mathcal{F}$ bázisban:

$$[\![\mathbf{v}]\!]_{\mathcal{F}} = [\![\mathbf{v}]\!]_{\mathcal{E}} \cdot [\![\mathcal{E} \to \mathcal{F}]\!] = (x, y) \cdot \begin{pmatrix} 1/5 & -2/5 \\ 2/5 & 1/5 \end{pmatrix} = (x/5 + 2y/5,\ -2x/5 + y/5).$$ ◇

3.38. Tétel. Legyenek U és V vektorterek a T test felett, legyen $\mathcal{E}_1$ és $\mathcal{E}_2$ két bázisa U-nak és legyen $\mathcal{F}_1$ és $\mathcal{F}_2$ két bázisa V-nek. Ekkor tetszőleges $\varphi \in \mathrm{Hom}(U, V)$ lineáris leképezésre:

$$[\![\varphi]\!]_{\mathcal{E}_2,\mathcal{F}_2} = [\![\mathcal{E}_2 \to \mathcal{E}_1]\!] \cdot [\![\varphi]\!]_{\mathcal{E}_1,\mathcal{F}_1} \cdot [\![\mathcal{F}_1 \to \mathcal{F}_2]\!].$$

Bizonyítás. A 3.29. Tételt és a 3.34. Megjegyzést használjuk a φ leképezés $U \xrightarrow{\mathrm{id}_U} U \xrightarrow{\varphi} V \xrightarrow{\mathrm{id}_V} V$ felbontására:

$$[\![\varphi]\!]_{\mathcal{E}_2,\mathcal{F}_2} = [\![\mathrm{id}_U\, \varphi\, \mathrm{id}_V]\!]_{\mathcal{E}_2,\mathcal{F}_2} = [\![\mathrm{id}_U]\!]_{\mathcal{E}_2,\mathcal{E}_1} \cdot [\![\varphi]\!]_{\mathcal{E}_1,\mathcal{F}_1} \cdot [\![\mathrm{id}_V]\!]_{\mathcal{F}_1,\mathcal{F}_2} = [\![\mathcal{E}_2 \to \mathcal{E}_1]\!] \cdot [\![\varphi]\!]_{\mathcal{E}_1,\mathcal{F}_1} \cdot [\![\mathcal{F}_1 \to \mathcal{F}_2]\!].$$ □

3.39. Következmény. Legyen V vektortér a T test felett és legyen $\mathcal{E}\colon \mathbf{e}_1, \ldots, \mathbf{e}_n$ és $\mathcal{F}\colon \mathbf{f}_1, \ldots, \mathbf{f}_n$ két bázisa V-nek. Ekkor tetszőleges $\varphi \in \mathrm{Hom}(V, V)$ lineáris transzformációra:

$$[\![\varphi]\!]_{\mathcal{F}} = [\![\mathcal{F} \to \mathcal{E}]\!] \cdot [\![\varphi]\!]_{\mathcal{E}} \cdot [\![\mathcal{E} \to \mathcal{F}]\!] = [\![\mathcal{E} \to \mathcal{F}]\!]^{-1} \cdot [\![\varphi]\!]_{\mathcal{E}} \cdot [\![\mathcal{E} \to \mathcal{F}]\!].$$

3.40. Definíció. Azt mondjuk, hogy a T test feletti $A, B \in T^{n\times n}$ mátrixok hasonlóak, ha létezik olyan $Q \in T^{n\times n}$ nemelfajuló mátrix, amelyre $B = Q^{-1}AQ$. Jelölés: $A \sim B$.

3.41. Állítás. A hasonlóság ekvivalenciareláció a $T^{n\times n}$ halmazon bármely T test és n pozitív egész szám esetén.

Bizonyítás. Lásd [SzL] 4.5. Tétel. □

3.42. Tétel. Bármely T test és n pozitív egész szám esetén az $A, B \in T^{n\times n}$ mátrixok akkor és csak akkor hasonlóak, ha ugyanannak a lineáris transzformációnak különböző bázisokban felírt mátrixai:

$$A \sim B \iff \exists \varphi \in \mathrm{Hom}(V, V)\colon A = [\![\varphi]\!]_{\mathcal{E}} \text{ és } B = [\![\varphi]\!]_{\mathcal{F}} \text{ alkalmas } \mathcal{E}, \mathcal{F} \text{ bázisokra.}$$

Bizonyítás. Ha $A = [\![\varphi]\!]_{\mathcal{E}}$ és $B = [\![\varphi]\!]_{\mathcal{F}}$, akkor a 3.39. Következmény szerint A és B hasonló. Fordítva, tfh. $B = Q^{-1}AQ$ valamely nemelfajuló $Q \in T^{n\times n}$ mátrixra. A 3.36. Tétel szerint léteznek olyan $\mathcal{E}$ és $\mathcal{F}$ bázisok V-ben, melyekre $[\![\mathcal{E} \to \mathcal{F}]\!] = Q$, továbbá a 3.28. Megjegyzés szerint létezik olyan $\varphi \in \mathrm{Hom}(V, V)$ lineáris transzformáció, amelyre $[\![\varphi]\!]_{\mathcal{E}} = A$. Mindezekből a 3.39. Következmény alapján következik, hogy

$$B = Q^{-1}AQ = [\![\mathcal{E} \to \mathcal{F}]\!]^{-1} \cdot [\![\varphi]\!]_{\mathcal{E}} \cdot [\![\mathcal{E} \to \mathcal{F}]\!] = [\![\varphi]\!]_{\mathcal{F}}.$$ □

3.43. Következmény. Lineáris transzformáció mátrixának determinánsa nem függ a bázis megválasztásától.

Bizonyítás. A fenti tétel szerint elég azt igazolni, hogy hasonló mátrixok determinánsa megegyezik. Ez pedig könnyen adódik a determinánsok szorzástételéből:

$$|Q^{-1}AQ| = |Q^{-1}| \cdot |A| \cdot |Q| = |Q^{-1}| \cdot |Q| \cdot |A| = |Q^{-1}Q| \cdot |A| = |E| \cdot |A| = |A|.$$ □

3.44. feladat. Tekintsük a 3.26. Feladatban megadott φ lineáris transzformációt és az $\mathcal{E}$, $\mathcal{F}$ bázisokat. Számítsuk ki φ mátrixát az $\mathcal{E}$ bázisban a 3.39. Következményben szereplő képlettel.

Megoldás. Tudjuk, hogy

$$[\![\varphi]\!]_{\mathcal{F}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \text{(lásd a 3.26. Feladatot)} \quad \text{és} \quad [\![\mathcal{F} \to \mathcal{E}]\!] = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \text{(lásd a 3.37. Feladatot).}$$

Ezeket könnyű volt felírni, nem is kellett hozzájuk semmit számolni. A $[\![\varphi]\!]_{\mathcal{E}}$ mátrixot (amit a 3.26. Feladatban már meghatároztunk) így lehet megkapni a fentiekből a 3.39. Következmény képletével:

$$[\![\varphi]\!]_{\mathcal{E}} = [\![\mathcal{F} \to \mathcal{E}]\!]^{-1} \cdot [\![\varphi]\!]_{\mathcal{F}} \cdot [\![\mathcal{F} \to \mathcal{E}]\!] = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{pmatrix}.$$ ⋄

$$[\![\mathbf{u}\varphi]\!]_{\mathcal{F}} = [\![\mathbf{u}]\!]_{\mathcal{E}} \cdot [\![\varphi]\!]_{\mathcal{E},\mathcal{F}} \qquad [\![\mathbf{v}]\!]_{\mathcal{F}} = [\![\mathbf{v}]\!]_{\mathcal{E}} \cdot [\![\mathcal{E} \to \mathcal{F}]\!]$$

$$[\![\varphi]\!]_{\mathcal{E}_2,\mathcal{F}_2} = [\![\mathcal{E}_2 \to \mathcal{E}_1]\!] \cdot [\![\varphi]\!]_{\mathcal{E}_1,\mathcal{F}_1} \cdot [\![\mathcal{F}_1 \to \mathcal{F}_2]\!]$$

$$[\![\varphi]\!]_{\mathcal{F}} = [\![\mathcal{F} \to \mathcal{E}]\!] \cdot [\![\varphi]\!]_{\mathcal{E}} \cdot [\![\mathcal{E} \to \mathcal{F}]\!] = [\![\mathcal{E} \to \mathcal{F}]\!]^{-1} \cdot [\![\varphi]\!]_{\mathcal{E}} \cdot [\![\mathcal{E} \to \mathcal{F}]\!].$$

4. Lineáris transzformációk

Ebben a fejezetben T mindig egy testet jelöl, V egy T feletti n-dimenziós vektorteret, φ pedig egy $V \to V$ lineáris transzformációt.

4.1. Sajátérték, sajátvektor, karakterisztikus polinom

4.1. Definíció. Azt mondjuk, hogy az $U \leq V$ altér invariáns a φ transzformációra, ha $U\varphi \subseteq U$, azaz $\mathbf{u}\varphi \in U$ teljesül minden $\mathbf{u} \in U$ esetén.

4.2. Állítás. Tegyük fel, hogy $V = U_1 \oplus \cdots \oplus U_k$, ahol az $U_i \leq V$ alterek invariánsak φ-re. Legyen $\mathcal{B}_i$ bázis U_i-ben és legyen a $\varphi|_{U_i} \in \mathrm{Hom}(U_i, U_i)$ lineáris transzformáció mátrixa a $\mathcal{B}_i$ bázisban A_i $(i = 1, \dots, k)$. Ekkor φ mátrixa a $\mathcal{B} := \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ bázisban az alábbi blokkdiagonális mátrix:

$$[\![\varphi]\!]_{\mathcal{B}} = \begin{pmatrix} \boxed{A_1} & & & \\ & \boxed{A_2} & & \\ & & \ddots & \\ & & & \boxed{A_k} \end{pmatrix}.$$

Bizonyítás. Ha $\mathbf{b} \in \mathcal{B}_i$, akkor $\mathbf{b}\varphi \in U_i$, mivel U_i invariáns altér. Ebből következik, hogy a $[\![\varphi]\!]_{\mathcal{B}}$ mátrix $\mathbf{b}$-hez tartozó sorában (ami nem más, mint a $[\![\mathbf{b}\varphi]\!]_{\mathcal{B}}$ vektor), a $\mathcal{B}_i$-n kívüli bázisvektorokhoz tartozó oszlopokban nullák állnak, a $\mathcal{B}_i$-beli bázisvektorokhoz tartozó oszlopokban lévő komponensek pedig épp a $[\![\mathbf{b}\varphi]\!]_{\mathcal{B}_i}$ vektort, azaz az A_i mátrix $\mathbf{b}$-hez tartozó sorát alkotják. □

4.3. Definíció. Azt mondjuk, hogy a $\mathbf{v} \in V$ vektor a $\lambda \in T$ sajátértékhez tartozó sajátvektora a $\varphi \in \mathrm{Hom}(V, V)$ lineáris transzformációnak, ha $\mathbf{v}\varphi = \lambda\mathbf{v}$ és $\mathbf{v} \neq \mathbf{0}$. Hasonlóan, az $\mathbf{x} \in T^n$ vektor a $\lambda \in T$ sajátértékhez tartozó sajátvektora az $A \in T^{n\times n}$ mátrixnak, ha $\mathbf{x}A = \lambda\mathbf{x}$ és $\mathbf{x} \neq \mathbf{0}$.

4.4. Megjegyzés. Oszlopvektorokkal dolgozva a sajátvektort így definiálhatnánk: $A\mathbf{x} = \lambda\mathbf{x}$. Ilyenkor jobb oldali sajátvektorról beszélünk, a fenti definícióban pedig bal oldali sajátvektorról beszélhetnénk. Nyilván A jobb oldali sajátvektorai éppen A^{T} bal oldali sajátvektorai, és viszont. Megállapodunk abban, hogy sajátvektoron mindig bal oldali sajátvektort (sorvektort) értünk. A sajátértékeknél szerencsére nem merül fel ez a kétértelműség (lásd a 4.14. Megjegyzést).

4.5. Állítás. A V vektortér tetszőleges $\mathcal{B}$ bázisa és tetszőleges $\mathbf{v} \in V$ vektor és λ skalár esetén ekvivalensek az alábbiak:

(i) $\mathbf{v}$ a λ sajátértékhez tartozó sajátvektora a $\varphi \in \mathrm{Hom}(V, V)$ transzformációnak;

(ii) $[\![\mathbf{v}]\!]_{\mathcal{B}}$ a λ sajátértékhez tartozó sajátvektora a $[\![\varphi]\!]_{\mathcal{B}}$ mátrixnak.

Bizonyítás. Rögtön következik a $V \to T^n$, $\mathbf{v} \to [\![\mathbf{v}]\!]_{\mathcal{B}}$ izomorfizmusból:

$$\mathbf{v}\varphi = \lambda\mathbf{v} \iff [\![\mathbf{v}\varphi]\!]_{\mathcal{B}} = [\![\lambda\mathbf{v}]\!]_{\mathcal{B}} \iff [\![\mathbf{v}]\!]_{\mathcal{B}} \cdot [\![\varphi]\!]_{\mathcal{B}} = \lambda[\![\mathbf{v}]\!]_{\mathcal{B}},$$

és persze $\mathbf{v} \neq \mathbf{0}$ akkor és csak akkor, ha $[\![\mathbf{v}]\!]_{\mathcal{B}} \neq \mathbf{0}$. □

4.6. Definíció. A $\varphi \in \mathrm{Hom}(V, V)$ lineáris transzformáció λ sajátértékhez tartozó sajátalterén az alábbi S_λ halmazt értjük:

$$S_\lambda = \left\{\mathbf{v} \in V : \mathbf{v}\varphi = \lambda\mathbf{v}\right\}.$$

4.7. Megjegyzés. Az S_λ jelölést akkor is használhatjuk, ha λ nem sajátérték, ilyenkor $S_\lambda = \{\mathbf{0}\}$.

4.8. Állítás. Az S_λ sajátaltér valóban altere V-nek, és invariáns φ-re.

Bizonyítás. Vegyük észre, hogy S_λ nem más, mint a $\varphi - \lambda \cdot \mathrm{id}$ lineáris transzformáció magja, így valóban altér. Az invariancia világos, hiszen minden altér zárt a λ-val való szorzásra. □

4.9. Definíció. Az $A \in T^{n\times n}$ mátrix karakterisztikus polinomján a $p_A(x) = \det(A - xE)$ polinomot értjük:

$$p_A(x) = \begin{vmatrix} a_{11} - x & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - x & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - x \end{vmatrix}.$$

4.10. Megjegyzés. Figyeljük meg, hogy $p_A(x) \in T[x]$ egy n-edfokú polinom:

$$p_A(x) = (-1)^n \cdot x^n + (-1)^{n-1} \cdot \operatorname{tr}(A) \cdot x^{n-1} + \cdots + \det(A).$$

Itt $\operatorname{tr}(A) = a_{11} + \cdots + a_{nn}$ az A mátrix nyoma.

4.11. Állítás. Hasonló mátrixok karakterisztikus polinomja megegyezik.

Bizonyítás. [SzL: 14.3] □

4.12. Definíció. A fenti állítás szerint van értelme lineáris transzformáció karakterisztikus polinomjáról beszélni, hiszen a transzformáció különböző bázisokhoz tartozó mátrixainak ugyanaz a karakterisztikus polinomja: a φ lineáris transzformáció karakterisztikus polinomján valamely/bármely bázisbeli mátrixának karakterisztikus polinomját értjük.

4.13. Tétel. A $\lambda \in T$ skalár akkor és csak akkor sajátértéke az $A \in T^{n\times n}$ mátrixnak, ha gyöke A karakterisztikus polinomjának. (Ugyanez érvényes mátrixok helyett lineáris transzformációkra is.)

Bizonyítás. A λ sajátértékhez tartozó sajátaltér $S_\lambda = \{\mathbf{x} \in T^n : \mathbf{x}A = \lambda\mathbf{x}\}$, ami nem más, mint az $\mathbf{x}(A-\lambda E) = \mathbf{0}$ homogén lineáris egyenletrendszer megoldástere. (Itt kellemetlenséget okoznak a sorvektorok: az egyenletrendszer szokásos felírása ez lenne: $(A - \lambda E)^{\mathrm{T}}\mathbf{x}^{\mathrm{T}} = \mathbf{0}^{\mathrm{T}}$.) Akkor és csak akkor lesz λ sajátérték, ha van nemtriviális megoldása ennek az egyenletrendszernek, azaz ha $|A - \lambda E| = 0$, ez pedig épp azt jelenti, hogy λ gyöke A karakterisztikus polinomjának. □

4.14. Megjegyzés. Mivel a transzponálás nem változtatja meg a determinánst, A és A^{T} karakterisztikus polinomja ugyanaz, és így sajátértékeik is megegyeznek, ezért nem szükséges bal vagy jobb oldali sajátértékről beszélni (lásd a 4.4. Megjegyzést).

4.15. Lemma. A nullvektor nem áll elő különböző sajátértékekhez tartozó sajátvektorok összegeként.

Bizonyítás. Indirekten bizonyítunk: tfh. $\mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_k = \mathbf{0}$, ahol $\mathbf{u}_i$ a λ_i sajátértékhez tartozó sajátvektora a φ lineáris transzformációnak ($i = 1, \ldots, k$), és a λ_i sajátértékek páronként különbözőek. Tfh. itt k a lehető legkisebb, azaz a nullvektor nem áll elő k-nál kevesebb, különböző sajátértékekhez tartozó sajátvektor összegeként. Alkalmazzuk az $\mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_k$ összegre a φ transzformációt:

$$\mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_k = \mathbf{0} \implies \lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2 + \cdots + \lambda_k\mathbf{u}_k = \mathbf{0}.$$

Most pedig szorozzuk be az $\mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_k$ összeget λ_1-gyel:

$$\mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_k = \mathbf{0} \implies \lambda_1\mathbf{u}_1 + \lambda_1\mathbf{u}_2 + \cdots + \lambda_1\mathbf{u}_k = \mathbf{0}.$$

Kivonva a kettőt egymásból, azt kapjuk, hogy

$$(\lambda_2 - \lambda_1)\mathbf{u}_2 + \cdots + (\lambda_k - \lambda_1)\mathbf{u}_k = \mathbf{0}.$$

Mivel a λ_i sajátértékek páronként különbözőek, itt egyik tag sem nulla, tehát $(\lambda_i - \lambda_1)\mathbf{u}_i$ sajátvektor λ_i sajátértékkel ($i = 2, \ldots, k$). Ez pedig ellentmondás, mert feltettük, hogy a nullvektor nem áll elő k-nál kevesebb, különböző sajátértékekhez tartozó sajátvektor összegeként. □

4.16. Következmény. Ha a φ lineáris transzformációnak $\lambda_1, \ldots, \lambda_k$ páronként különböző sajátértékei, akkor a hozzájuk tartozó sajátalterek összege direkt összeg:

$$S_{\lambda_1} + \cdots + S_{\lambda_k} = S_{\lambda_1} \oplus \cdots \oplus S_{\lambda_k}.$$

Bizonyítás. Ez rögtön következik a 4.15. Lemmából (lásd a 2.61. Tétel (i+ε) pontját). □

4.17. Következmény. Különböző sajátértékekhez tartozó sajátvektorok lineárisan függetlenek.

Bizonyítás. Tfh. $\mathbf{v}_i$ a λ_i sajátértékhez tartozó sajátvektora a φ lineáris transzformációnak ($i = 1, \ldots, k$) és a λ_i sajátértékek páronként különbözőek. Mindegyik S_{λ_i} sajátaltérben tudunk egy olyan $\mathcal{B}_i$ bázist megadni, amely tartalmazza a $\mathbf{v}_i$ vektort (ugye?). A 4.16. Következmény szerint $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ bázisa az $S_{\lambda_1} \oplus \cdots \oplus S_{\lambda_k}$ altérnek (lásd a 2.61. Tétel (iv) pontját), így lineárisan független. Következésképp $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is lineárisan független vektorrendszer (hiszen lineárisan független vektorrendszer bármely részrendszere is lineárisan független). □

4.2. Diagonalizálhatóság

4.18. Definíció. Azt mondjuk, hogy...

- a $\varphi \in \operatorname{Hom}(V,V)$ lineáris transzformáció diagonalizálható, ha van V-nek olyan bázisa, melyben φ mátrixa diagonális (az ilyen bázis sajátvektorokból áll, ezért szokás sajátbázisnak nevezni);
- az $A \in T^{n\times n}$ mátrix diagonalizálható, ha hasonló egy diagonális mátrixhoz (van olyan Q nemelfajuló mátrix, hogy QAQ^{-1} diagonális).

4.19. Állítás. Egy lineáris transzformáció akkor és csak akkor diagonalizálható, ha valamely/bármely bázisbeli mátrixa diagonalizálható.

Bizonyítás. Következik a 3.39. Következményből. □

4.20. Definíció. Legyenek a $\varphi \in \operatorname{Hom}(V,V)$ lineáris transzformáció sajátértékei $\lambda_1,\ldots,\lambda_k$.

- A λ_i sajátérték geometriai multiplicitásán az S_{λ_i} sajátaltér dimenzióját értjük. Jelölés: g_i.
- A λ_i sajátérték algebrai multiplicitásán azt értjük, hogy hányszoros gyöke a karakterisztikus polinomnak. Jelölés: a_i.

4.21. Tétel. A geometriai multiplicitás nem nagyobb, mint az algebrai multiplicitás: ha a $\varphi \in \operatorname{Hom}(V,V)$ lineáris transzformáció sajátértékei $\lambda_1,\ldots,\lambda_k$ rendre $g_1,\ldots,g_k$ és $a_1,\ldots,a_k$ geometriai és algebrai multiplicitásokkal, akkor

$$g_i \leq a_i \qquad (i=1,\ldots,k).$$

Bizonyítás. Vegyünk egy $\mathcal{B}_i$ bázist az S_{λ_i} sajátaltérben ($i=1,\ldots,k$); ekkor a 4.15. Lemma és a 2.61. Tétel szerint $\mathcal{B}_1\cup\cdots\cup\mathcal{B}_k$ bázisa az $S_{\lambda_1}\oplus\cdots\oplus S_{\lambda_k}$ altérnek. Mivel $\mathcal{B}_1\cup\cdots\cup\mathcal{B}_k$ lineárisan független vektorrendszer, kiegészíthető a V vektortér egy $\mathcal{B}$ bázisává. Ebben a bázisban φ mátrixa így fest:

$$A := [\![\varphi]\!]_{\mathcal{B}} = \begin{pmatrix}
\lambda_1 &&&&&&&&&&&& \\
& \ddots &&&&&&&&&&& \\
&& \lambda_1 &&&&&&&&&& \\
&&& \lambda_2 &&&&&&&&& \\
&&&& \ddots &&&&&&&& \\
&&&&& \lambda_2 &&&&&&& \\
&&&&&& \ddots &&&&&& \\
&&&&&&& \lambda_k &&&&& \\
&&&&&&&& \ddots &&&& \\
&&&&&&&&& \lambda_k &&& \\
* & * & * & * & * & * & * & * & * & * & * & \cdots & * \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
* & * & * & * & * & * & * & * & * & * & * & \cdots & *
\end{pmatrix}$$

Itt az A mátrix főátlóján g_i db λ_i szerepel. Írjuk fel φ karakterisztikus polinomját az $A - xE$ mátrix determinánsaként (lásd a 4.12. Definíciót), és fejtsük ki a determinánst az első $g_1+\cdots+g_k$ sora szerint:

$$p_\varphi(x) = p_A(x) = |A - xE| = \pm(x-\lambda_1)^{g_1}\cdot\ldots\cdot(x-\lambda_k)^{g_k}\cdot\begin{vmatrix} *-x & \cdots & * \\ \vdots & \ddots & \vdots \\ * & \cdots & *-x \end{vmatrix}.$$

Innen látható, hogy a $p_A(x)$ polinomnak λ_i legalább g_i-szeres gyöke, tehát $a_i \geq g_i$. □

4.22. feladat. Határozzuk meg az alábbi A mátrix sajátértékeit, minden sajátérték algebrai és geometriai multiplicitását, adjunk meg bázisokat a sajátalterekben, és diagonalizáljuk a mátrixot, amennyire lehet:

$$A = \begin{pmatrix} 2 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 1 & 1 & 1 \end{pmatrix} \in \mathbb{Z}_3^{4\times 4}.$$

Megoldás. Először számítsuk ki a karakterisztikus polinomot (érdemes az első oszlop szerint kifejteni, mert akkor mindjárt ki lesz emelve egy gyöktényező):

$$p_A = \begin{vmatrix} 2-x & 0 & 2 & 1 \\ 0 & -x & 2 & 1 \\ 0 & 1 & -x & 2 \\ 0 & 1 & 1 & 1-x \end{vmatrix} = (2-x)(2x^3+x^2+2x) = (x-2)x(x^2+2x+1) = (x-2)x(x+1)^2 = x(x-2)^3.$$

Tehát két sajátérték van: $\lambda_1 = 0$ (algebrai multiplicitása: $a_1 = 1$), illetve $\lambda_2 = 2$ (algebrai multiplicitása: $a_2 = 3$). Most keressünk bázist mindkét sajátaltérben. Ehhez az $(A-\lambda E)^{\mathrm{T}}$ mátrixú homogén lineáris egyenletrendszert kell megoldanunk mindkét sajátértékre:

$$(A-\lambda_1 E)^{\mathrm{T}} = A^{\mathrm{T}} = \begin{pmatrix} 2&0&0&0 \\ 0&0&1&1 \\ 2&2&0&1 \\ 1&1&2&1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} \boxed{1}&0&0&0 \\ 0&\boxed{1}&0&2 \\ 0&0&\boxed{1}&1 \\ 0&0&0&0 \end{pmatrix} \Longrightarrow S_0 = [0121]$$

$$(A-\lambda_2 E)^{\mathrm{T}} = (A-2E)^{\mathrm{T}} = \begin{pmatrix} 0&0&0&0 \\ 0&1&1&1 \\ 2&2&1&1 \\ 1&1&2&2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} \boxed{1}&0&1&1 \\ 0&\boxed{1}&1&1 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix} \Longrightarrow S_2 = [2210, 2201]$$

Foglaljuk össze az eddigieket. Két sajátérték van:

- $\lambda_1 = 0$, algebrai multiplicitása $a_1 = 1$, geometriai multiplicitása $g_1 = 1$;
- $\lambda_2 = 2$, algebrai multiplicitása $a_2 = 3$, geometriai multiplicitása $g_2 = 2$.

Mivel a geometriai multiplicitások összege kisebb, mint 4, a mátrix nem diagonalizálható. A legtöbb, amit tehetünk, hogy a sajátalterekben talált bázisok egyesítését kibővítjük $\mathbb{Z}_3^4$ egy $\mathcal{B}$ bázisává (a 2.44. feladat módszerével):

$$\mathbf{b}_1 = 0121, \quad \mathbf{b}_2 = 2210, \quad \mathbf{b}_3 = 2201, \quad \mathbf{b}_4 = 1000.$$

Tekintsük a $\varphi\colon \mathbb{Z}_3^4 \to \mathbb{Z}_3^4$, $\mathbf{x} \mapsto \mathbf{x}A$ lineáris transzformációt. Ennek mátrixa a standard $\mathcal{E}$ bázisban éppen A. Térjünk át erről a $\mathcal{B}$ bázisra:

$$[\![\mathcal{B} \to \mathcal{E}]\!] = Q = \begin{pmatrix} 0&1&2&1 \\ 2&2&1&0 \\ 2&2&0&1 \\ 1&0&0&0 \end{pmatrix}, \qquad [\![\mathcal{E} \to \mathcal{B}]\!] = Q^{-1} = \begin{pmatrix} 0&0&0&1 \\ 1&1&2&0 \\ 1&2&2&1 \\ 1&1&0&1 \end{pmatrix}.$$

A φ lineáris transzformáció mátrixa a $\mathcal{B}$ bázisban „majdnem” diagonális:

$$[\![\varphi]\!]_{\mathcal{B}} = [\![\mathcal{B} \to \mathcal{E}]\!] \cdot [\![\varphi]\!]_{\mathcal{E}} \cdot [\![\mathcal{E} \to \mathcal{B}]\!] = QAQ^{-1} = \begin{pmatrix} 0&0&0&0 \\ 0&2&0&0 \\ 0&0&2&0 \\ 0&2&1&2 \end{pmatrix}.$$

A diagonális részben minden sajátérték annyiszor lép fel, amennyi a geometriai multiplicitása. Mivel a kapott mátrix trianguláris, könnyen leolvasható a karakterisztikus polinomja, a sajátértékei, és ezek algebrai multiplicitásai. Ezek természetesen ugyanazok, mint az eredeti A mátrixnál, hiszen a két mátrix hasonló.
Megjegyzés: Ha elrontottuk volna a sajátalterekben talált bázisok egyesítésének kibővítését az egész tér bázisává (azaz $\mathbf{b}_4$ nem lenne lineárisan független a $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ vektoroktól), akkor a $[\![\mathcal{B} \to \mathcal{E}]\!]$ mátrixnak nem lenne inverze. Tehát a 2.44. feladat (hosszadalmas) módszere helyett megtehetjük azt, hogy „érzésből” választunk egy $\mathbf{b}_4$ vektort, és a mátrixinvertálással mutatjuk meg, hogy jó volt a megérzésünk. Ha tényleg jó volt, akkor ezzel megtakarítottunk némi munkát (az inverz kiszámítását úgysem úszhatjuk meg). Ha nem volt jó a megérzésünk, akkor viszont egy másik $\mathbf{b}_4$ vektorral kell próbálkozni. ◇

4.23. Tétel. Legyen $\varphi \in \mathrm{Hom}(V,V)$, ahol $\dim V = n$, és tfh. a φ transzformációnak k különböző sajátértéke van rendre $g_1, \ldots, g_k$ geometriai és $a_1, \ldots, a_k$ algebrai multiplicitásokkal. Ekkor φ diagonalizálhatóságnak...

(i) szükséges és elegendő feltétele: $\sum g_i = n$;

(ii) szükséges és elegendő feltétele: $\sum a_i = n$ és $g_i = a_i$ (minden i-re);

(iii) szükséges feltétele: $\sum a_i = n$;

(iv) szükséges feltétele: $g_i = a_i$ (minden i-re);

(v) elegendő feltétele: $k = n$ és $a_1 = \cdots = a_n = 1$.

Bizonyítás. Akkor és csak akkor diagonalizálható a φ transzformáció, ha a sajátalterek összege kiadja az egész vektorteret: $V = S_{\lambda_1} \oplus \cdots \oplus S_{\lambda_k}$, ez pedig a 2.57. Tétel szerint ekvivalens azzal, hogy

$$\dim V = \dim(S_{\lambda_1} \oplus \cdots \oplus S_{\lambda_k}) = \dim S_{\lambda_1} + \cdots + \dim S_{\lambda_k} = g_1 + \cdots + g_k.$$

Ezzel igazoltuk az (i) állítást, és ezután (ii) már következik abból, hogy

$$\sum_{i=1}^{k} g_i \le \sum_{i=1}^{k} a_i \le n.$$

(Az első egyenlőtlenség a 4.21. Tételből következik, a második pedig abból, hogy egy T feletti polinomnak nem lehet több gyöke (multiplicitással számolva), mint amekkora a fokszáma.) A többi állítás egyszerű következménye (ii)-nek. □

4.24. Tétel. Ha $A \in T^{n\times n}$ diagonalizálható, akkor...

(i) A karakterisztikus polinom lineáris tényezőkre bomlik T felett:

$$p_A(x) = (-1)^n \cdot (x - \lambda_1)^{m_1} \cdot \ldots \cdot (x - \lambda_k)^{m_k}.$$

Itt $\lambda_1, \ldots, \lambda_k$ az A mátrix sajátértékei; a λ_i sajátérték algebrai és geometriai multiplicitása is m_i (és persze $m_1 + \cdots + m_k = n$).

(ii) A sajátértékek szorzata a determináns, összegük pedig a nyom:

$$\lambda_1^{m_1} \cdot \ldots \cdot \lambda_k^{m_k} = \det(A), \quad \text{és} \quad m_1 \cdot \lambda_1 + \ldots + m_k \cdot \lambda_k = \operatorname{tr}(A).$$

(iii) Van olyan $Q \in T^{n\times n}$ nemelfajuló mátrix, amelyre

$$QAQ^{-1} = \operatorname{diag}(\underbrace{\lambda_1, \ldots, \lambda_1}_{m_1}, \ \ldots, \ \underbrace{\lambda_k, \ldots, \lambda_k}_{m_k}) =: D.$$

Bizonyítás. Az első állítás következik a 4.23. Tétel (iii) pontjából, a második állítás pedig az elsőből, figyelembe véve a 4.10. Megjegyzést. A harmadik állítás csak a 4.18. Definíció megismétlése, csak annyival kell kiegészítenünk, hogy a diagonális alakban szükségképpen a sajátértékek állnak a főátlón, mindegyik annyiszor, amennyi a multiplicitása (miért?). □

4.25. Megjegyzés. A fenti tételbeli $QAQ^{-1} = D$ egyenlőség azzal ekvivalens, hogy $QA = DQ$, ami azt jelenti, hogy a Q mátrix sorai A-nak bal oldali sajátvektorai (és sajátbázist alkotnak). Ha az egyenlőséget $AQ^{-1} = Q^{-1}D$ alakba rendezzük, akkor pedig azt látjuk, hogy a Q^{-1} mátrix oszlopai A-nak jobb oldali sajátvektorai.

4.3. Minimálpolinom

4.26. Definíció. Legyen $f(x) = c_k x^k + c_{k-1}x^{k-1} + \cdots + c_1 x + c_0$ egy T feletti polinom. Ebbe a polinomba a következő módon helyettesíthetünk be egy $A \in T^{n\times n}$ mátrixot vagy egy $\varphi \in \operatorname{Hom}(T^n, T^n)$ lineáris transzformációt:

$$f(A) = c_k A^k + c_{k-1}A^{k-1} + \cdots + c_1 A + c_0 E = \textstyle\sum_{i=0}^{k} c_i A^i \in T^{n\times n};$$
$$f(\varphi) = c_k \varphi^k + c_{k-1}\varphi^{k-1} + \cdots + c_1 \varphi + c_0 \operatorname{id}_V = \textstyle\sum_{i=0}^{k} c_i \varphi^i \in \operatorname{Hom}(T^n, T^n).$$

Amint az (el)várható, a mátrixok és lineáris transzformációk közötti összhang a polinomokba való behelyettesítés során megmarad; ezt mondja ki az alábbi állítás.

4.27. Állítás. A T^n vektortér bármely $\mathcal{B}$ bázisára, bármely $\varphi \in \operatorname{Hom}(T^n, T^n)$ lineáris transzformációra és bármely $f \in T[x]$ polinomra $[\![f(\varphi)]\!]_{\mathcal{B}} = f([\![\varphi]\!]_{\mathcal{B}})$.

Bizonyítás. HF. □

A mátrixszorzás nem kommutatív, de egy mátrix hatványai felcserélhetőek egymással, ezért $f(A)$ és $g(A)$ is felcserélhető tetszőleges $f, g \in T[x]$ polinomok esetén. Ezt a tényt gyakran fogjuk használni, ezért külön kimondjuk a következő állításban (az összeadásra vonatkozó analóg ténnyel együtt).

4.28. Állítás. Tetszőleges $A \in T^{n\times n}$ mátrix és $f, g \in T[x]$ polinomok esetén

(i) $f(A) + g(A) = (f + g)(A)$;

(ii) $f(A) \cdot g(A) = (fg)(A) = g(A) \cdot f(A)$.

Bizonyítás. Az összeadásra vonatkozó állítás triviális (ugye?). A második állításhoz legyen $f = \sum c_i x^i$ és $g = \sum d_j x^j$. (Nem lesz szükségünk a polinomok fokszámaira, ezért nem vezetünk be rájuk külön jelölést, elég annyit tudni, hogy mindkét „szumma" véges számú tagból álló összeget jelöl. Később is gyakran fogunk élni ezzel az egyszerűsített jelöléssel.) Számítsuk ki az $f(A) \cdot g(A)$ szorzatot:

$$f(A) \cdot g(A) = \Big(\sum_i c_i A^i\Big) \cdot \Big(\sum_j d_j A^j\Big) = \sum_{i,j} c_i A^i d_j A^j = \sum_{i,j} c_i d_j A^{i+j} = \sum_t \Big(\sum_{i+j=t} c_i d_j\Big) A^t = (fg)(A).$$

Hasonlóan belátható, hogy $g(A) \cdot f(A) = (gf)(A)$, ami persze ugyanaz, mint $(fg)(A)$, hiszen a polinomok szorzása kommutatív művelet. □

Célunk olyan polinomo(ka)t találni, amely(ek)nek gyöke az A mátrix. A konstans nulla polinom persze ilyen, de a következő tétel szerint van nemnulla polinom is, amely annullálja (kinullázza) az A mátrixot.

4.29. Tétel. Bármely $A \in T^{n \times n}$ mátrixhoz létezik nemnulla $f \in T[x]$ polinom, amelynek ez a mátrix gyöke, azaz $f(A) = \underline{\underline{\mathbf{0}}}$.

Bizonyítás. Az $n \times n$-es mátrixok n^2 dimenziós vektorteret alkotnak T felett. Írjuk fel A első $n^2 + 1$ hatványát a nulladik hatványtól kezdve: $A^0, A^1, A^2, \ldots, A^{n^2}$. Tudjuk, hogy a dimenziószámnál eggyel több vektor (azaz most mátrix) már biztosan lineárisan függő, így ennek az $n^2 + 1$ mátrixnak van olyan nemtriviális lineáris kombinációja, ami a nullmátrixot adja. Tehát léteznek $c_0, c_1, \ldots, c_{n^2-1}, c_{n^2} \in T$ skalárok, amelyek nem mind nullák és $\sum_{i=0}^{n^2} c_i A^i = \underline{\underline{\mathbf{0}}}$. Ez pedig azt jelenti, hogy az $f = \sum_{i=0}^{n^2} c_i x^i$ nemnulla polinomnak gyöke az A mátrix. □

4.30. feladat. Adjunk meg az alábbi $A \in \mathbb{R}^{2\times 2}$ mátrixhoz egy minél kisebb fokszámú $f \in \mathbb{R}[x]$ polinomot, amelyre $f(A) = \underline{\underline{\mathbf{0}}}$:

$$A = \begin{pmatrix} 0,9 & 0,1 \\ 0,5 & 0,5 \end{pmatrix}.$$

Megoldás. Mivel a 2×2-es mátrixok tere 4-dimenziós, az E, A, A^2, A^3, A^4 mátrixok között biztosan teljesül valamilyen nemtriviális lineáris összefüggés. Számítsuk ki a mátrixokat:

$$E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A = \begin{pmatrix} 0,9 & 0,1 \\ 0,5 & 0,5 \end{pmatrix}, \quad A^2 = \begin{pmatrix} 0,86 & 0,14 \\ 0,7 & 0,3 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 0,844 & 0,156 \\ 0,78 & 0,22 \end{pmatrix}, \quad A^4 = \begin{pmatrix} 0,8376 & 0,1624 \\ 0,812 & 0,188 \end{pmatrix}.$$

Írjuk oszlopvektorokként egymás mellé a mátrixokat, és végezzünk elemi sorátalakításokat, hogy redukált lépcsős alakot kapjunk:

$$\begin{array}{ccccc} E & A & A^2 & A^3 & A^4 \\ \hline 1 & 0,9 & 0,86 & 0,844 & 0,8376 \\ 0 & 0,1 & 0,14 & 0,156 & 0,1624 \\ 0 & 0,5 & 0,7 & 0,78 & 0,812 \\ 1 & 0,5 & 0,3 & 0,22 & 0,188 \end{array} \quad \rightsquigarrow \quad \begin{array}{ccccc} \boxed{1} & 0 & -0,4 & -0,56 & -0,624 \\ 0 & \boxed{1} & 1,4 & 1,56 & 1,624 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}$$

Látható, hogy az első két oszlop (azaz E és A) lineárisan független, de a harmadik oszlop (azaz A^2) már kikombinálható ezekből: $A^2 = -0{,}4E + 1{,}4A$, vagyis $A^2 - 1{,}4A + 0{,}4E = \underline{\underline{\mathbf{0}}}$. Tehát a legkisebb fokú polinom, aminek az A mátrix gyöke: $m_A = x^2 - 1{,}4x + 0{,}4$. Ezt a polinomot nevezzük az A mátrix minimálpolinomjának (lásd a 4.31. Definíciót).
Megjegyzés. Vegyük észre, hogy $m_A = x^2 - \mathrm{tr}(A)x + \det(A) = p_A$, vagyis itt a minimálpolinom megegyezik a karakterisztikus polinommal (ez persze nem igaz minden mátrixra!). ◇

Jelölés. Tekintsük mindazon polinomok halmazát, amelyeknek gyöke az A mátrix:

$$I_A := \{f \in T[x] : f(A) = \underline{\underline{\mathbf{0}}}\} \subseteq T[x].$$

A konstans nulla polinom nyilván eleme ennek a halmaznak, de a fenti tétel szerint vannak benne nemnulla polinomok is. Ezek közül a legkisebb fokút nevezzük az A mátrix minimálpolinomjának.

4.31. Definíció. Az $A \in T^{n\times n}$ mátrix minimálpolinomján az $I_A \setminus \{0\}$ halmaz legkisebb fokú elemét értjük, vagyis az A-t annulláló legkisebb fokú nemnulla polinomot. Jelölés: m_A.

4.32. Megjegyzés. Hasonló módon definiálható lineáris transzformáció minimálpolinomja is, és egy transzformáció minimálpolinomja megegyezik bármely bázisbeli mátrixának minimálpolinomjával.

A fenti definícióban a határozott névelő nem igazán helyes, hiszen több minimális fokú eleme is lehet az $I_A \setminus \{0\}$ halmaznak. Meg fogjuk mutatni, hogy ezek a minimális fokú elemek egymás asszociáltjai, tehát a minimálpolinom asszociáltság erejéig egyértelmű. (Szokás a minimálpolinomot az A-t annulláló legkisebb fokú főpolinomként definiálni, így már egyértelmű lenne a minimálpolinom.) Ehhez szükségünk lesz az I_A halmaz alábbi két egyszerű tulajdonságára.

4.33. Állítás. Legyen $A \in T^{n\times n}$ tetszőleges mátrix.

(i) Ha $f, g \in I_A$, akkor $f \pm g \in I_A$. (Tehát az I_A halmaz zárt az összeadásra és a kivonásra.)

(ii) Ha $f \in I_A$ és $g \in T[x]$, akkor $fg \in I_A$. (Ez a tulajdonság erősebb, mint a szorzásra való zártság: az I_A halmaz „magába szippantja" a szorzatokat. Ez az ún. szívó tulajdonság.)

Bizonyítás. Mindkét állítás egyszerűen levezethető az I_A halmaz definíciójából és a 4.28. Állításból.

(i) Ha $f, g \in I_A$, akkor $f(A) = \underline{\underline{0}}$ és $g(A) = \underline{\underline{0}}$, tehát $(f \pm g)(A) = f(A) \pm g(A) = \underline{\underline{0}} \pm \underline{\underline{0}} = \underline{\underline{0}}$.

(ii) Ha $f \in I_A$, akkor $f(A) = \underline{\underline{0}}$, tehát $(fg)(A) = f(A) \cdot g(A) = \underline{\underline{0}} \cdot g(A) = \underline{\underline{0}}$. □

4.34. Tétel. Tetszőleges $A \in T^{n\times n}$ mátrix és $f \in T[x]$ polinom esetén $f(A) = \underline{\underline{0}}$ pontosan akkor teljesül, ha $m_A \mid f$, azaz

$$f \in I_A \iff m_A \mid f.$$

Bizonyítás. Először azt lássuk be, hogy ha $m_A \mid f$, akkor $f(A) = \underline{\underline{0}}$. Tegyük fel, hogy $m_A \mid f$, azaz $f = m_A \cdot q$ teljesül alkalmas $q \in T[x]$ polinommal. Ekkor $f(A) = m_A(A) \cdot q(A) = \underline{\underline{0}} \cdot q(A) = \underline{\underline{0}}$, hiszen $m_A(A) = \underline{\underline{0}}$ a minimálpolinom definíciója szerint.
Most pedig azt bizonyítjuk, hogy ha $f(A) = \underline{\underline{0}}$, akkor $m_A \mid f$. Tegyük fel tehát, hogy $f(A) = \underline{\underline{0}}$, azaz $f \in I_A$. Osszuk el az f polinomot maradékosan a minimálpolinommal (azt szeretnénk belátni, hogy nulla lesz a maradék): $f = q \cdot m_A + r$. Itt q a hányados, r a maradék, és utóbbiról tudjuk, hogy $\deg r < \deg m_A$. Ekkor rendezéssel adódik, hogy $r = f - q \cdot m_A$. Az $f(A) = \underline{\underline{0}}$ feltevés szerint $f \in I_A$, továbbá $m_A \in I_A$ is teljesül a minimálpolinom definíciója szerint. Ezekből $q \cdot m_A \in I_A$ majd $f - q \cdot m_A \in I_A$ következik a 4.33 Állítás alapján. Azt kaptuk tehát, hogy $r \in I_A$. Mivel $\deg r < \deg m_A$, szükségképpen $r = 0$, hiszen az I_A-beli nemnulla polinomok között m_A a legkisebb fokszámú. Így $m_A \mid f$ teljesül, és ezt akartuk belátni. □

4.35. Következmény. Test feletti négyzetes mátrix minimálpolinomja asszociáltság erejéig egyértelmű.

Bizonyítás. Tegyük fel, hogy m_A és m'_A is minimálpolinomja az $A \in T^{n\times n}$ mátrixnak. Az előző tétel szerint $m_A \mid m'_A$ (mert m_A minimálpolinom és $m'_A \in I_A$) és $m'_A \mid m_A$ (mert m'_A minimálpolinom és $m_A \in I_A$), tehát m_A és m'_A asszociáltak. □

4.36. Megjegyzés. A $T[x]$ polinomgyűrűben (vagy bármely más gyűrűben) az olyan részhalmazokat, melyek rendelkeznek a 4.33. Állításban megfogalmazott két tulajdonsággal, ideáloknak nevezzük. (Ezért használtuk az I_A jelölést.) A 4.34. Tételben gyakorlatilag azt láttuk be, hogy a $T[x]$ gyűrű minden ideálja egy polinom összes többszöröseiből áll. Az ilyen ideálokat főideáloknak nevezzük. Általában egy gyűrűben létez(het)nek olyan ideálok is, amelyek nem főideálok, de test feletti polinomgyűrűben a 4.34. Tétel szerint minden ideál főideál, azaz $T[x]$ úgynevezett főideálgyűrű. (Fun fact: $\mathbb{Z}[x]$ nem főideálgyűrű, pl. az $\{f \in \mathbb{Z}[x] : f(0) \text{ páros}\}$ halmaz ideál a $\mathbb{Z}[x]$ gyűrűben, de nem főideál.)

4.4. A Cayley–Hamilton-tétel

A Cayley–Hamilton-tétel azt állítja, hogy minden négyzetes mátrix gyöke a saját karakterisztikus polinomjának: tetszőleges $M \in T^{n\times n}$ mátrix esetén $p_M(M) = \underline{\underline{0}}$. Mivel $p_M(x) = \det(M - xE)$, azt gondolhatnánk, hogy a tétel triviális: $p_M(M) = \det(M - ME) = \det(\underline{\underline{0}}) = 0$. Ez valóban triviális gondolatmenet, csakhogy SZAMÁRSÁG! Nézzük meg egy példán, hogy mi a baj ezzel.

4.37. Példa. Az $M = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right) \in \mathbb{R}^{2\times 2}$ mátrix karakterisztikus polinomja:

$$p_M(x) = \det(M - xE) = \begin{vmatrix} 1-x & 2 \\ 3 & 4-x \end{vmatrix} = x^2 - 5x - 2 \in \mathbb{R}[x].$$

Behelyettesítve az M mátrixot, valóban nullát kapunk (pontosabban nullmátrixot):

$$p_M(M) = M^2 - 5M - 2E = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix} - \begin{pmatrix} 5 & 10 \\ 15 & 20 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Figyeljük meg, hogy a karakterisztikus polinom definíciójában az $M - xE = \left(\begin{smallmatrix} 1-x & 2 \\ 3 & 4-x \end{smallmatrix}\right) \in \mathbb{R}[x]^{2\times 2}$ polinommmátrix (azaz polinomokból álló mátrix) szerepel. Ha ebben próbálnánk az $x = M$ helyettesítést elvégezni, akkor az alábbi furcsa dolgot kapnánk:

$$\begin{pmatrix} 1-M & 2 \\ 3 & 4-M \end{pmatrix} = \begin{pmatrix} \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) - \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right) & \left(\begin{smallmatrix} 2 & 0 \\ 0 & 2 \end{smallmatrix}\right) \\ \left(\begin{smallmatrix} 3 & 0 \\ 0 & 3 \end{smallmatrix}\right) & \left(\begin{smallmatrix} 4 & 0 \\ 0 & 4 \end{smallmatrix}\right) - \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right) \end{pmatrix} = \begin{pmatrix} \left(\begin{smallmatrix} 0 & -2 \\ -3 & -3 \end{smallmatrix}\right) & \left(\begin{smallmatrix} 2 & 0 \\ 0 & 2 \end{smallmatrix}\right) \\ \left(\begin{smallmatrix} 3 & 0 \\ 0 & 3 \end{smallmatrix}\right) & \left(\begin{smallmatrix} 3 & -2 \\ -3 & 0 \end{smallmatrix}\right) \end{pmatrix}.$$

Az $M - xE$ kifejezést nem polinommátrixként, hanem mátrixpolinomként (azaz mátrixegyütthatós polinomként) értelmezve ezt kapjuk:

$$-Ex + M = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} x + \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in \mathbb{R}^{2\times 2}[x].$$

Itt minden probléma nélkül elvégezhető az $x = M$ helyettesítés, és eredményül persze a nullmátrixot kapjuk. De a Cayley–Hamilton-tétel nem az $M - xE$ elsőfokú mátrixpolinomról szól, hanem az $M - xE$ polinommátrix determinánsáról, ami az $x^2 - 5x - 2$ másodfokú polinom. Tehát a bizonyítást sajnos nem ússzuk meg annyival, hogy $\det(M - ME) = \det(M - M) = \det(\underline{\mathbf{0}}) = 0$.

A fenti példában kétféleképpen „házasítottunk össze” mátrixokat és polinomokat. Mindkettőre, és a kettő kapcsolatára is szükségünk lesz a Cayley–Hamilton-tétel bizonyításához, ezért formálisan is definiáljuk őket. A polinommátrixok nem mások, mint a $T[x]$ polinomgyűrű feletti mátrixok. Eddig csak test feletti mátrixokkal foglalkoztunk, de a tanultak jó része (amikor nincs szükség multiplikatív inverzekre) érvényes kommutatív egységelemes gyűrűk feletti mátrixokra is, így speciálisan $T[x]$ feletti mátrixokra is.

4.38. Definíció. A T test feletti $n \times n$-es polinommátrixon a $T[x]$ polinomgyűrű feletti $n \times n$-es mátrixot értünk, azaz $A = (f_{ij})$ alakú mátrixot, ahol $f_{ij} \in T[x]$ $(i, j = 1, \ldots, n)$. Ezek a mátrixok a szokásos módon értelmezett mátrixműveletekkel egységelemes gyűrűt alkotnak, melyet $T[x]^{n\times n}$ jelöl ($n \geq 2$ esetén nemkommutatív gyűrű).

A mátrixpolinomok a $T^{n\times n}$ gyűrű feletti polinomok. Mivel $T^{n\times n}$ általában nemkommutatív gyűrű, óvatosnak kell lennünk: a polinomokról korábban tanultak egy része nem marad érvényben.

4.39. Definíció. Egy T test feletti mátrixpolinomon $f(x) = \sum_{i=0}^{k} A_i x^i = A_k x^k + \cdots + A_1 x + A_0$ alakú formális kifejezést értünk, ahol az együtthatók azonos méretű T feletti négyzetes mátrixok: $A_i \in T^{n\times n}$ $(i = 0, \ldots, k)$. Ezen polinomok halmazát $T^{n\times n}[x]$ jelöli. Legyenek $f, g \in T^{n\times n}[x]$ mátrixpolinomok:

$$f = \sum_i A_i x^i, \quad g = \sum_i B_i x^i.$$

Ekkor f és g összegét és szorzatát a következőképpen definiáljuk:

$$(f + g) = \sum_i (A_i + B_i) x^i, \quad fg = \sum_t \sum_{i+j=t} A_i B_j x^t.$$

Ezekkel a műveletekkel $T^{n\times n}[x]$ egységelemes gyűrűt alkot ($n \geq 2$ esetén nemkommutatív gyűrű).

4.40. Példa. Tekintsük az alábbi A polinommátrixot és $f(x)$ mátrixpolinomot:

$$A = \begin{pmatrix} x^2 + 2x + 3 & 5x - 2 \\ x^2 + 5 & 6x^2 - 4x \end{pmatrix} \in \mathbb{R}[x]^{2\times 2}, \qquad f(x) = \begin{pmatrix} 1 & 0 \\ 1 & 6 \end{pmatrix} x^2 + \begin{pmatrix} 2 & 5 \\ 0 & -4 \end{pmatrix} x + \begin{pmatrix} 3 & -2 \\ 5 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}[x].$$

Amint látható, a kettő „majdnem ugyanaz”: x helyébe tetszőleges valós számot helyettesítve, és elvégezve a műveleteket, ugyanazt a mátrixot kapjuk. Szigorúbb értelemben, formális kifejezésként tekintve viszont A és f nem ugyanaz.[2] A példa sugall egy természetes megfeleltetést a polinommátrixok és a mátrixpolinomok között, amiről meg lehet mutatni, hogy izomorfizmus a $T[x]^{n\times n}$ és $T^{n\times n}[x]$ gyűrűk között. Ezt nem fogjuk formalizálni, csak azt a tényt mondjuk ki, hogy két polinommátrix szorzatához tartozó mátrixpolinom ugyanaz, mint a megfelelő mátrixpolinomok szorzata.

4.41. Lemma. Legyen $A, B \in T[x]^{n\times n}$ polinommátrixok, és legyenek $f, g \in T^{n\times n}[x]$ a nekik megfelelő mátrixpolinomok. Ekkor az AB polinommátrixnak az fg mátrixpolinom felel meg.

Bizonyítás. Lásd az F.1. fejezetet a függelékben. □

4.42. Példa. A 4.41. Lemma illusztrálásaként tekintsük a 4.40. Példában szereplő A polinommátrixot és $f(x)$ mátrixpolinomot, valamint az alábbi B polinommátrixot és a neki megfelelő $g(x)$ mátrixpolinomot:

$$B = \begin{pmatrix} x - 3 & 4 \\ 3x + 5 & 2x \end{pmatrix} \in \mathbb{R}[x]^{2\times 2}, \qquad g(x) = \begin{pmatrix} 1 & 0 \\ 3 & 2 \end{pmatrix} x + \begin{pmatrix} -3 & 4 \\ 5 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}[x].$$

[2] A polinomok formális kifejezések. Például $x \in \mathbb{Z}_2[x]$ és $x^2 \in \mathbb{Z}_2[x]$ két különböző polinom, még a fokszámuk sem egyezik, pedig minden behelyettesítésre ugyanazt az eredményt adják, hiszen $\mathbb{Z}_2$ mindkét elemének önmaga a négyzete.

Szorozzuk össze az A és B mátrixokat úgy, ahogy mátrixokat szorozni szoktunk:

$$A \cdot B = \begin{pmatrix} x^2+2x+3 & 5x-2 \\ x^2+5 & 6x^2-4x \end{pmatrix} \cdot \begin{pmatrix} x-3 & 4 \\ 3x+5 & 2x \end{pmatrix} = \begin{pmatrix} x^3+14x^2+16x-19 & 14x^2+4x+12 \\ 19x^3+15x^2-15x-15 & 12x^3-4x^2+20 \end{pmatrix} \in \mathbb{R}[x]^{2\times 2}.$$

Szorozzuk most össze az f és g polinomokat a 4.39. Definíció szerint:

$$\begin{aligned} f \cdot g &= \left(\begin{pmatrix} 1 & 0 \\ 1 & 6 \end{pmatrix} x^2 + \begin{pmatrix} 2 & 5 \\ 0 & -4 \end{pmatrix} x + \begin{pmatrix} 3 & -2 \\ 5 & 0 \end{pmatrix} \right) \cdot \left(\begin{pmatrix} 1 & 0 \\ 3 & 2 \end{pmatrix} x + \begin{pmatrix} -3 & 4 \\ 5 & 0 \end{pmatrix} \right) = \\ &= \begin{pmatrix} 1 & 0 \\ 19 & 12 \end{pmatrix} x^3 + \begin{pmatrix} 14 & 14 \\ 15 & -4 \end{pmatrix} x^2 + \begin{pmatrix} 16 & 4 \\ -15 & 0 \end{pmatrix} x + \begin{pmatrix} -19 & 12 \\ -15 & 20 \end{pmatrix} \in \mathbb{R}^{2\times 2}[x]. \end{aligned}$$

Látjuk, hogy a két szorzás összhangban van egymással: az AB polinommátrixnak épp az fg mátrixpolinom felel meg.

Amíg x helyébe csak skalárokat helyettesítünk, addig a polinommátrix és a mátrixpolinom lényegében ekvivalensnek tekinthető. Amikor viszont mátrixot akarunk behelyettesíteni, akkor már lényeges különbség van a kettő között: polinommátrix esetén olyan mátrixot kapunk, amelynek elemei maguk is mátrixok, míg mátrixpolinom esetén „rendes” mátrix lesz az eredmény (lásd a 4.37. Példát). Ezért a Cayley–Hamilton-tétel bizonyításában a polinommátrixot előbb átalakítjuk majd mátrixpolinommá, és csak utána végezzük el az $x = M$ behelyettesítést. Mivel a mátrixok szorzása nem kommutatív, még ezzel is óvatosnak kell lenni, mert kétféleképpen is elvégezhető a behelyettesítés. Megállapodunk abban, hogy mindig „jobbról” helyettesítjük be az M mátrixot; lásd az alábbi a 4.43. Definíciót. (Lehetne „balról” is behelyettesíteni, így a $\sum_i M^i A_i$ mátrixot kapnánk, ami általában nem ugyanaz, mint $\sum_i A_i M^i$.)

4.43. Definíció. Legyen $f(x) = \sum_{i=0}^{k} A_i x^i = A_k x^k + \cdots + A_1 x + A_0 \in T^{n\times n}[x]$ egy T test feletti mátrixpolinom. Tetszőleges $M \in T^{n\times n}$ mátrix esetén az $f(x)$ mátrixpolinom M helyen vett helyettesítési értékét így definiáljuk:

$$f(M) = \sum_{i=0}^{k} A_i M^i = A_k M^k + \cdots + A_1 M + A_0.$$

A mátrixszorzás nemkommutativitása miatt sajnos a 4.28. Állítás második része nem marad érvényben mátrixpolinomokra (az első persze igen). A következő lemma viszont ad egy elegendő feltételt arra, hogy két mátrixpolinom szorzatát tényezőnként lehessen kiértékelni.

4.44. Lemma. Legyenek $f(x), g(x) \in T^{n\times n}[x]$ mátrixpolinomok, és legyen $M \in T^{n\times n}$ négyzetes mátrix. Ha M felcserélhető $g(x)$ együtthatóival, akkor $f(M) \cdot g(M) = (fg)(M)$.

Bizonyítás. A 4.39. Definícióban használt jelölésekkel írjuk fel az $f(M) \cdot g(M)$ szorzatot:

$$f(M) \cdot g(M) = \Big(\sum_i A_i M^i\Big)\Big(\sum_j B_j M^j\Big) = \sum_{i,j} A_i M^i B_j M^j.$$

Feltevésünk szerint minden j-re B_j felcserélhető M-mel, és így M^i-vel is, ezért a fenti összeget átalakíthatjuk így:

$$\sum_{i,j} A_i M^i B_j M^j = \sum_{i,j} A_i B_j M^i M^j = \sum_{i,j} (A_i B_j) M^{i+j} = \sum_t \sum_{i+j=t} (A_i B_j)\, M^t = (fg)(M). \qquad \square$$

4.45. Definíció. Az $M \in T^{n\times n}$ mátrix adjungált mátrixát a következő módon képezzük:

1. Felírjuk az M mátrix aldeterminánsaiból álló mátrixot. Ebben az i-edik sor j-edik eleme annak a mátrixnak a determinánsa, melyet M-ből az i-edik sor és j-edik oszlop elhagyásával kapunk.
2. Ezután az aldeterminánsmátrix (i, j)-edik elemét megszorozzuk $(-1)^{i+j}$-vel minden i és j esetén. Így kapjuk az előjeles aldeterminánsmátrixot.
3. Végső lépésként a kapott mátrixot transzponáljuk.

Az M mátrix adjungált mátrixát $\mathrm{adj}(M)$-mel jelöljük.

A Cayley–Hamilton-tétel bizonyításához kulcsfontosságú az adjungált mátrix következő lemmában kimondott tulajdonsága, ami lényegében a determinánsok kifejtésének és ferde kifejtésének összefoglalása. Ez a lemma szolgáltatja nemelfajuló mátrixok esetén az $M^{-1} = \frac{1}{\det(M)} \mathrm{adj}(M)$ formulát az inverzre, de az alábbi megfogalmazásban elfajuló mátrixokra is érvényes. Ebben a formában nem kell a bizonyításban osztást vagy multiplikatív inverzet használni, ezért nemcsak test feletti mátrixokra, hanem kommutatív egységelemes gyűrű feletti mátrixokra is érvényes, speciálisan $T[x]$ feletti mátrixokra, azaz polinommátrixokra is.

4.46. Lemma. Tetszőleges kommutatív egységelemes gyűrű feletti négyzetes M mátrixra

$$M \cdot \mathrm{adj}(M) = \mathrm{adj}(M) \cdot M = \begin{pmatrix} |M| & 0 & \cdots & 0 & 0 \\ 0 & |M| & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & |M| & 0 \\ 0 & 0 & \cdots & 0 & |M| \end{pmatrix} = \det(M) \cdot E.$$

Bizonyítás. Lásd linalg1. $\square$

4.47. Tétel Cayley–Hamilton-tétel. Tetszőleges T test feletti négyzetes mátrix gyöke saját karakterisztikus polinomjának: bármely $M \in T^{n\times n}$ esetén $p_M(M) = \underline{\underline{\mathbf{0}}}$.

Bizonyítás. Legyen $M \in T^{n\times n}$ négyzetes mátrix; ennek karakterisztikus polinomja $p_M = \det(M - Ex) \in T[x]$. Itt $M - xE$ egy polinommátrix, azaz $T[x]$ feletti mátrix. Mivel $T[x]$ kommutatív egységelemes gyűrű, használhatjuk a 4.46. Lemmát:

$$\operatorname{adj}(M - Ex) \cdot (M - Ex) = \det(M - Ex) \cdot E = p_M(x) \cdot E \in T[x]^{n\times n}.$$

A 4.41. Tétel szerint ugyanez érvényes akkor is, ha polinommátrixok helyett mátrixpolinomokkal dolgozunk. Legyen $f(x) = \operatorname{adj}(M - Ex)$ és $g(x) = M - Ex$, most már mindkettőt mátrixpolinomként tekintve. Ezekkel a jelölésekkel tehát

$$f(x) \cdot g(x) = p_M(x) \cdot E \in T^{n\times n}[x].$$

A g mátrixpolinomot írjuk ki részletesen: $g(x) = M - Ex = (-E)x + M = B_1x + B_0$. ahol $B_1 = -E$ és $B_0 = M$. Mivel M felcserélhető a B_1 és B_0 együtthatókkal, a 4.44. Lemma szerint $f(M) \cdot g(M) = (fg)(M)$. Itt $g(M) = M - EM = \underline{\underline{\mathbf{0}}}$, ezért $(fg)(M) = \underline{\underline{\mathbf{0}}}$. Node, amint fent megfigyeltük, $(fg)(x) = p_M(x) \cdot E$, tehát azt kaptuk, hogy $p_M(M) \cdot E = \underline{\underline{\mathbf{0}}}$, és így $p_M(M) = \underline{\underline{\mathbf{0}}}$. $\square$

4.48. Következmény. A minimálpolinom osztója a karakterisztikus polinomnak: minden $A \in T^{n\times n}$ esetén $m_A \mid p_A$.

Bizonyítás. A Cayley-Hamilton tétel szerint $p_A \in I_A$, és így a 4.34. Tétel következtében $m_A \mid p_A$. $\square$

Nézzük meg a Cayley-Hamilton tétel bizonyításának gondolatmenetét a 4.37. Példában szereplő mátrixon.

4.49. Példa. Az $M - xE$ mátrix adjungált mátrixa:

$$\operatorname{adj}(M - xE) = \begin{pmatrix} 4-x & -2 \\ -3 & 1-x \end{pmatrix} \in \mathbb{R}[x]^{2\times 2}.$$

Az $\operatorname{adj}(M - xE)$ és $M - xE$ polinommátrixok szorzata a 4.46. Lemmának megfelelően így alakul:

$$\operatorname{adj}(M - xE) \cdot (M - xE) = \begin{pmatrix} 4-x & -2 \\ -3 & 1-x \end{pmatrix} \cdot \begin{pmatrix} 1-x & 2 \\ 3 & 4-x \end{pmatrix} = \begin{pmatrix} x^2 - 5x - 2 & 0 \\ 0 & x^2 - 5x - 2 \end{pmatrix} \in \mathbb{R}[x]^{2\times 2}.$$

Ebben a formában bajos lenne elvégezni az $x = M$ behelyettesítést, ezért polinommátrixokról áttérünk mátrixpolinomokra. A tétel bizonyításában használt $f(x) = \operatorname{adj}(M - Ex)$ és $g(x) = M - Ex$ jelölésekkel

$$f(x) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} x + \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}[x] \quad \text{és} \quad g(x) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} x + \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in \mathbb{R}^{2\times 2}[x].$$

A két polinom szorzatát kiszámítva „ugyanazt” kapjuk, mint fent a polinommátrixok szorzatára:

$$f(x) \cdot g(x) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} x^2 + \begin{pmatrix} -5 & 0 \\ 0 & -5 \end{pmatrix} x + \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} = Ex^2 - 5Ex - 2E \in \mathbb{R}^{2\times 2}[x].$$

Ide már behelyettesíthetjük az M mátrixot, és az eredmény $f(M) \cdot g(M) = EM^2 - 5EM - 2E = (M^2 - 5M - 2E) \cdot E = p_M(M) \cdot E$ lesz, ami nem más, mint a nullmátrix, hiszen $g(M) = \underline{\underline{\mathbf{0}}}$. Ebből következik, hogy $p_M(M) = \underline{\underline{\mathbf{0}}}$.

5. A Jordan-normálalak

Ebben a fejezetben a komplex számtest feletti mátrixoknak adjuk meg egy kanonikus alakját: megmutatjuk, hogy minden $\mathbb{C}$ feletti négyzetes mátrix hasonló egy nagyon szép, „majdnem” diagonális mátrixhoz. Ez az úgynevezett Jordan-normálalak lényegében egyértelmű, ezért segítségével két mátrixról könnyen eldönthetjük, hogy hasonlóak-e. Legyen $A \in \mathbb{C}^{n\times n}$ egy $\mathbb{C}$ feletti négyzetes mátrix, és legyen φ a $V = \mathbb{C}^n$ vektortér A-hoz tartozó lineáris transzformációja: $\varphi : V \to V$, $\mathbf{v} \mapsto \mathbf{v}A$. Ekkor tehát $[\![\varphi]\!]_{\mathcal{E}} = A$, ahol $\mathcal{E}$ a standard bázis V-ben. Legyenek a φ transzformáció (vagy, ha úgy tetszik az A mátrix) karakterisztikus polinomjának gyökei $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$, rendre $a_1, \ldots, a_k$ algebrai multiplicitásokkal:

$$p_\varphi \sim p_A \sim (x - \lambda_1)^{a_1} \cdot \ldots \cdot (x - \lambda_k)^{a_k}.$$

Itt fontos, hogy a komplex számtest felett dolgozunk: az algebra alaptételének köszönhető, hogy a karakterisztikus polinomot a fenti módon gyöktényezők szorzatára tudjuk bontani. A most bevezetett jelöléseket az egész fejezetben használni fogjuk.

5.1. Visszavezetés nilpotens transzformációkra

Célunk a $V = \mathbb{C}^n$ vektorteret φ-re nézve invariáns „kicsi” alterek direkt összegére bontani. Az invariancia azért fontos, hogy legyen értelme a φ transzformációt megszorítani ezekre az alterekre. A „kicsiség” pedig azért fontos, hogy könnyebben megérthessük ezeket a megszorításokat. Így a φ transzformáció globális szerkezetét feltárhatjuk a (remélhetőleg elég egyszerű) lokális viselkedések alapján.

Először idézzünk fel két tényt polinomokról, amelyekre a későbbiek során szükségünk lesz.

5.1. Lemma. Két polinom legnagyobb közös osztója kifejezhető a két polinom „lineáris kombinációjaként”:

$$\forall f, g \in \mathbb{C}[x]\ \exists s, t \in \mathbb{C}[x] :\ \operatorname{lnko}(f, g) = fs + gt.$$

5.2. Lemma. A komplex számtest felett két polinom akkor és csak akkor relatív prím, ha nincs közös gyökük:

$$\forall f, g \in \mathbb{C}[x] :\ \operatorname{lnko}(f, g) \sim 1 \iff \nexists \lambda \in \mathbb{C} : f(\lambda) = g(\lambda) = 0.$$

A következő két lemma azt mutatja, hogy hogyan lehet az A mátrix polinomjai segítségével olyan altereket konstruálni, amelyek invariánsak az A-hoz tartozó φ lineáris transzformációra. Itt, és a továbbiakban, az egyszerűség kedvéért egy mátrix magján a standard bázisban hozzá tartozó lineáris transzformáció magját értjük.

5.3. Lemma. Bármely $f \in \mathbb{C}[x]$ polinomra a $\operatorname{Ker}(f(A)) \leq V$ altér invariáns φ-re, azaz minden $\mathbf{v} \in \operatorname{Ker}(f(A))$ esetén $\mathbf{v}A \in \operatorname{Ker}(f(A))$.

Bizonyítás. Tudjuk, hogy lineáris transzformáció magja mindig altér, ezért $\operatorname{Ker}(f(A)) \leq V$. Az invarianciához tegyük fel, hogy $\mathbf{v} \in \operatorname{Ker}(f(A))$, azaz $\mathbf{v}f(A) = \mathbf{0}$. Ekkor a 4.28. Állítást használva

$$(\mathbf{v}A)f(A) = \mathbf{v}(A \cdot f(A)) = \mathbf{v}(f(A) \cdot A) = (\mathbf{v}f(A))A = \mathbf{0}A = \mathbf{0},$$

hiszen $\mathbf{v} \in \operatorname{Ker}(f(A))$. Ezzel beláttuk, hogy $\mathbf{v}A \in \operatorname{Ker}(f(A))$, azaz $\operatorname{Ker}(f(A))$ valóban invariáns A-ra (vagyis az A-hoz tartozó φ transzformációra). □

5.4. Lemma. Tetszőleges relatív prím $f, g \in \mathbb{C}[x]$ polinomok esetén $\operatorname{Ker}((fg)(A)) = \operatorname{Ker}(f(A)) \oplus \operatorname{Ker}(g(A))$.

Bizonyítás. Az világos, hogy $\operatorname{Ker}(f(A)) \subseteq \operatorname{Ker}((fg)(A))$, mert ha $\mathbf{v}f(A) = \mathbf{0}$, akkor $\mathbf{v}((fg)(A)) = \mathbf{v}f(A)g(A) = \mathbf{0}g(A) = \mathbf{0}$ a 4.28. Állítás szerint. Hasonlóan belátható, hogy $\operatorname{Ker}(g(A)) \subseteq \operatorname{Ker}((fg)(A))$, következésképp $\operatorname{Ker}(f(A)) + \operatorname{Ker}(g(A)) \subseteq \operatorname{Ker}((fg)(A))$. Ehhez még nem is használtuk azt a feltételt, hogy f és g relatív prímek.

A másik irányú tartalmazáshoz és az összeg „direktségéhez” viszont már szükségünk lesz arra, hogy $\operatorname{lnko}(f, g) \sim 1$, amiből az 5.1. Lemma alapján következik, hogy alkalmas $s, t \in \mathbb{C}[x]$ polinomokkal $fs + gt = 1$ teljesül. Az $x = A$ behelyettesítéssel azt kapjuk, hogy $(fs)(A) + (gt)(A) = E$, és így a 4.28. Állítást használva $f(A)s(A) + g(A)t(A) = E$. Ebből tetszőleges $\mathbf{v} \in V$ vektorral balról szorozva nyerjük a következő egyenlőséget, ami a bizonyítás kulcsa:

$$\mathbf{v}f(A)s(A) + \mathbf{v}g(A)t(A) = \mathbf{v}. \tag{2}$$

Igazoljuk, hogy $\operatorname{Ker}((fg)(A)) \subseteq \operatorname{Ker}(f(A)) + \operatorname{Ker}(g(A))$. Ehhez elég megmutatni, hogy amennyiben $\mathbf{v} \in \operatorname{Ker}((fg)(A))$, akkor a fenti $\mathbf{v} = \mathbf{v}f(A)s(A) + \mathbf{v}g(A)t(A)$ felbontásban az első tag eleme a $\operatorname{Ker}(g(A))$ altérnek, a második tag pedig eleme a $\operatorname{Ker}(f(A))$ altérnek. A 4.28. Állítás segítségével ellenőrizzük, hogy a $g(A)$ mátrix valóban annullálja a $\mathbf{v}f(A)s(A)$ vektort:

$$\mathbf{v}f(A)s(A)g(A) = \mathbf{v}f(A)g(A)s(A) = \mathbf{v}((fg)(A))s(A) = \mathbf{0}s(A) = \mathbf{0},$$

mivel $\mathbf{v} \in \operatorname{Ker}((fg)(A))$. Ez azt jelenti, hogy $\mathbf{v}f(A)s(A) \in \operatorname{Ker} g(A)$. Hasonlóan belátható, hogy a (2) bal oldalán álló összeg második tagját annullálja $f(A)$. Ezzel megkaptuk a $\operatorname{Ker}((fg)(A)) = \operatorname{Ker}(f(A)) + \operatorname{Ker}(g(A))$ felbontást. A 2.59. Tétel (ii) pontja szerint már csak annak igazolása van hátra, hogy $\operatorname{Ker}(f(A)) \cap \operatorname{Ker}(g(A)) = \{\mathbf{0}\}$. Tegyük fel, hogy $\mathbf{v} \in \operatorname{Ker}(f(A)) \cap \operatorname{Ker}(g(A))$, vagyis $\mathbf{v}f(A) = \mathbf{v}g(A) = \mathbf{0}$. Ekkor (2) szerint $\mathbf{v} = \mathbf{v}f(A)s(A) + \mathbf{v}g(A)t(A) = \mathbf{0}s(A) + \mathbf{0}t(A) = \mathbf{0}$, tehát $\mathbf{v} = \mathbf{0}$. □

A fenti lemmát az A mátrix karakterisztikus polinomjának gyöktényezős felbontására alkalmazva a Cayley–Hamilton-tétel segítségével megkapjuk a V vektortér felbontását „kicsi” invariáns alterek direkt összegére.

5.5. Tétel. Tetszőleges $A \in \mathbb{C}^{n \times n}$ négyzetes mátrix esetén a $V = \mathbb{C}^n$ vektortér felbomlik a következő direkt összegre:

$$V = \operatorname{Ker}(A - \lambda_1 E)^{a_1} \oplus \cdots \oplus \operatorname{Ker}(A - \lambda_k E)^{a_k},$$

ahol $\lambda_1, \ldots, \lambda_k$ az A mátrix sajátértékei rendre $a_1, \ldots, a_k$ algebrai multiplicitásokkal.

Bizonyítás. A Cayley–Hamilton-tétel szerint $p_A(A) = \underline{\underline{0}}$, azaz $\operatorname{Ker}(p_A(A)) = V$. A karakterisztikus polinom gyöktényezős felbontását felírva ezt így is megfogalmazhatjuk:

$$V = \operatorname{Ker}\Big((A-\lambda_1 E)^{a_1} \cdot (A-\lambda_2\ E)^{a_2} \cdot \ldots \cdot (A-\lambda_k E)^{a_k}\Big).$$

Az $f = (x-\lambda_1)^{a_1}$ és $g = (x-\lambda_2)^{a_2} \cdot \ldots \cdot (x-\lambda_k)^{a_k}$ polinomok relatív prímek, hiszen nincs közös gyökük (lásd az 5.2. Lemmát). Ezért az 5.4. Lemma alapján

$$V = \operatorname{Ker}(A-\lambda_1 E)^{a_1} \oplus \operatorname{Ker}\Big((A-\lambda_2\ E)^{a_2} \cdot \ldots \cdot (A-\lambda_k E)^{a_k}\Big).$$

Ismét az 5.4. Lemmát alkalmazva, ezúttal az $f = (x-\lambda_2)^{a_2}$ és $g = (x-\lambda_3)^{a_3} \cdot \ldots \cdot (x-\lambda_k)^{a_k}$ „szereposztással”, kapjuk, hogy

$$V = \operatorname{Ker}(A-\lambda_1 E)^{a_1} \oplus \operatorname{Ker}(A-\lambda_2 E)^{a_2} \oplus \operatorname{Ker}\Big((A-\lambda_3 E)^{a_3} \cdot \ldots \cdot (A-\lambda_k E)^{a_k}\Big).$$

Hasonlóan folytatva, az 5.4. Lemma ismételt alkalmazásával végül megkapjuk a kívánt direkt felbontást:

$$V = \operatorname{Ker}(A-\lambda_1 E)^{a_1} \oplus \cdots \oplus \operatorname{Ker}(A-\lambda_k E)^{a_k}.$$ □

5.6. Definíció. Azt mondjuk, hogy egy $\mathbf{v} \in V \setminus \{\mathbf{0}\}$ vektor a λ sajátértékhez tartozó általánosított sajátvektora az A mátrixnak (illetve a megfelelő lineáris transzformációnak), ha létezik olyan ℓ természetes szám, melyre $\mathbf{v}(A-\lambda E)^{\ell} = \mathbf{0}$.

5.7. Megjegyzés. Az $\ell = 1$ esetben éppen a „rendes” sajátvektor fogalmát kapjuk.

Figyeljük meg, hogy a $\operatorname{Ker}(A-\lambda_i E)^{a_i}$ altér elemei (a nullvektor kivételével) mind általánosított sajátvektorok ($\ell = a_i$ választással). Így az 5.5. Tétel szerint minden vektor előáll általánosított sajátvektorok összegeként, azaz létezik általánosított sajátvektorokból álló bázisa V-nek.

5.8. Következmény. Tetszőleges $A \in \mathbb{C}^{n\times n}$ négyzetes mátrix esetén a $V = \mathbb{C}^n$ vektortérnek létezik A általánosított sajátvektoraiból álló bázisa.

Emlékeztető: „rendes” sajátvektorokból álló bázis akkor és csak akkor létezik, ha a mátrix diagonalizálható. Így a fenti következmény alapján bízhatunk abban, hogy minden mátrix „majdnem” diagonalizálható. Ez a „majdnem diagonális” alak lesz a Jordan-normálalak.

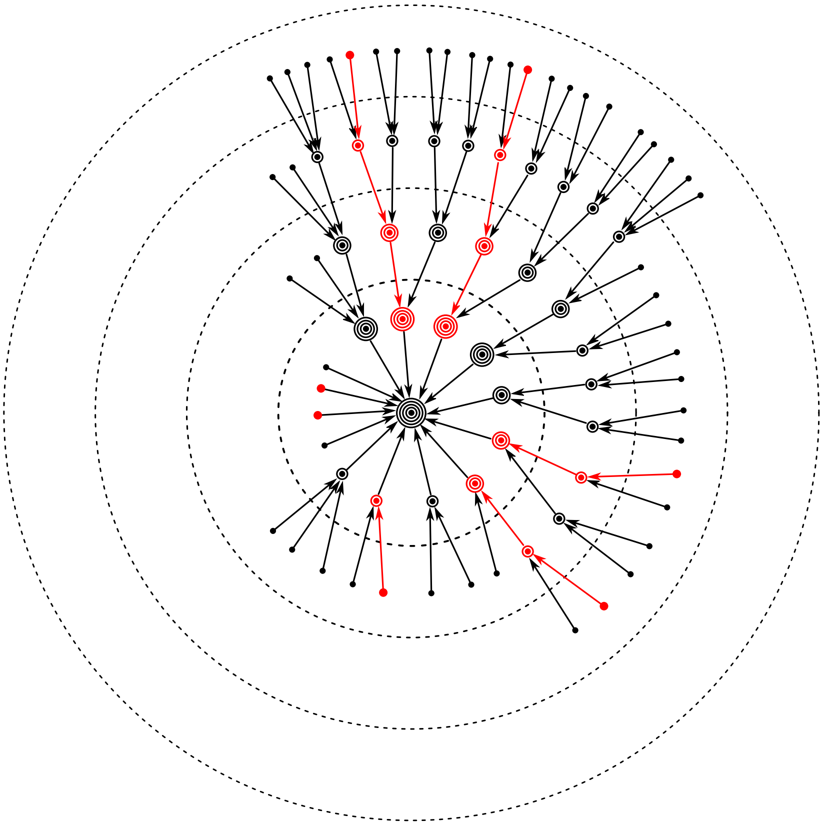
Legyen $W_i = \operatorname{Ker}(A-\lambda_i E)^{a_i}$, legyen $\mathcal{B}_i$ bázis ebben az altérben, és legyen $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$. Ekkor az 5.5. Tétel szerint $\mathcal{B}$ bázis a $V = W_1 \oplus \cdots \oplus W_k$ vektortérben. A 4.2. Állítás mutatja, hogy a φ lineáris transzformáció mátrixa ebben a bázisban blokkdiagonális lesz, ahol a blokkok a $\varphi_i := \varphi|_{W_i}$ transzformációk mátrixai a $\mathcal{B}_i$ bázisokban (itt az egyes nullák különböző méretű, általában nem is négyzetes nullmátrixokat jelölnek):

$$[\![\varphi]\!]_{\mathcal{B}} = \begin{pmatrix} [\![\varphi_1]\!]_{\mathcal{B}_1} & \underline{\underline{0}} & \cdots & \underline{\underline{0}} \\ \underline{\underline{0}} & [\![\varphi_2]\!]_{\mathcal{B}_2} & \cdots & \underline{\underline{0}} \\ \vdots & \vdots & \ddots & \vdots \\ \underline{\underline{0}} & \underline{\underline{0}} & \cdots & [\![\varphi_k]\!]_{\mathcal{B}_k} \end{pmatrix}.$$

Ahhoz, hogy megértsük a φ transzformáció szerkezetét, elég megérteni, hogy az egyes blokkokban mi történik. Legyen ψ_i a $\varphi - \lambda_i \operatorname{id}_V$ transzformáció megszorítása a W_i altérre: $\psi_i = (\varphi - \lambda_i \operatorname{id}_V)|_{W_i} = \varphi_i - \lambda_i \operatorname{id}_{W_i}$. Célunk az eddig tetszőlegesen megválasztott $\mathcal{B}_i$ bázis helyett olyan $\mathcal{F}_i$ bázist találni W_i-ben, amelyben ψ_i mátrixa minél egyszerűbb szerkezetű. Ha M_i jelöli ezt az „egyszerű” mátrixot, azaz $[\![\psi_i]\!]_{\mathcal{F}_i} = M_i$, akkor $[\![\varphi_i]\!]_{\mathcal{F}_i} = M_i + \lambda_i E_i$, hiszen $\varphi_i = \psi_i + \lambda_i \operatorname{id}_{W_i}$. Tehát φ mátrixa az $\mathcal{F} = \mathcal{F}_1 \cup \cdots \cup \mathcal{F}_k$ bázisban az $M_i + \lambda_i E_i$ blokkokból álló blokkdiagonális mátrix lesz (ahol E_i a megfelelő, $\dim W_i$ méretű egységmátrixot jelöli):

$$[\![\varphi]\!]_{\mathcal{F}} = \begin{pmatrix} M_1 + \lambda_1 E_1 & \underline{\underline{0}} & \cdots & \underline{\underline{0}} \\ \underline{\underline{0}} & M_2 + \lambda_2 E_2 & \cdots & \underline{\underline{0}} \\ \vdots & \vdots & \ddots & \vdots \\ \underline{\underline{0}} & \underline{\underline{0}} & \cdots & M_k + \lambda_k E_k \end{pmatrix}.$$

Figyeljük meg, hogy mivel $W_i = \operatorname{Ker}(A-\lambda_i E)^{a_i}$, minden $\mathbf{w} \in W_i$ vektorra $\mathbf{w}(A-\lambda_i E)^{a_i} = \mathbf{0}$ teljesül, azaz $\mathbf{w}\psi_i^{a_i} = \mathbf{0}$. Tehát $\psi_i^{a_i}$ az azonosan nulla transzformációja a W_i vektortérnek. Az ilyen transzformációt, amelynek valamely hatványa nulla, nilpotens transzformációnak nevezzük. (Előfordulhat, hogy van a_i-nál kisebb m kitevő is, amelyre $\psi_i^m = 0$. A legkisebb ilyen m kitevőt a ψ_i transzformáció nilpotenciaindexének nevezzük.) Épp ez a tulajdonság az, ami miatt áttértünk a φ_i transzformációról a $\psi_i = \varphi_i - \lambda_i \operatorname{id}_{W_i}$ transzformációra. A következőkben látni fogjuk, hogy az ilyen nilpotens transzformációk nem túl bonyolult szerkezetűek, fel fogjuk tárni titkaikat, majd ennek segítségével „összerakjuk” φ blokkdiagonális mátrixát, a Jordan-normálalakot.



1. ábra. Nilpotens transzformáció gráfja

5.2. Nilpotens transzformáció szerkezete

Figyelem: ezen fejezet tanulmányozása előtt célszerű elolvasni az F.2. fejezetet a függelékben!

Legyen W egy véges dimenziós vektortér a komplex számtest (vagy akár tetszőleges test) felett. Legyen $\psi : W \to W$ egy nilpotens transzformáció, és legyen m a legkisebb kitevő, amelyre $\psi^m = \underline{\underline{\mathbf{0}}}$, ekkor tehát $\psi^{m-1} \neq \underline{\underline{\mathbf{0}}}$. Egy tetszőleges $\mathbf{v} \in W$ vektorból kiindulva, a $\mathbf{v}, \mathbf{v}\psi, \mathbf{v}\psi^2, \mathbf{v}\psi^3, \ldots$ sorozat előbb-utóbb (legkésőbb az m-edik lépésben) eléri a nullvektort, ami természetesen fixpontja a ψ transzformációnak. Az 1. ábra mutatja ψ gráfjának vázlatos szerkezetét. Minden pont W egy elemét jelképezi (mivel $\mathbb{C}$ feletti vektortérben dolgozunk, igazából végtelen sok pontot kellene a gráfon ábrázolni), és minden pontból a ψ melletti képébe mutat a nyíl. Az ábra közepén található a nullvektor, belőle önmagába mutatna a nyíl (hurokél), de ezt nem rajzoltuk be. A legalább egyszer bekarikázott pontok alkotják az $\operatorname{Im}\psi$ halmazt, a legalább kétszer bekarikázott pontok az $\operatorname{Im}\psi^2$ halmazt, és így tovább. A szaggatott vonallal rajzolt körök pedig a magtereket mutatják: a legkisebb kör a $\operatorname{Ker}\psi$ halmaz, a második legkisebb kör a $\operatorname{Ker}(\psi^2)$ halmaz, és így tovább. Az ábrán látható példában $m = 4$, azaz $\psi^4 = \underline{\underline{\mathbf{0}}}$, ezért $\operatorname{Im}(\psi^4) = \{\mathbf{0}\}$ és $\operatorname{Ker}(\psi^4) = W$. Az ábráról leolvashatóak a képterek és magterek közötti alábbi összefüggések.

5.9. Lemma. Ha $\psi : W \to W$ nilpotens lineáris transzformációja a W véges dimenziós vektortérnek, amelyre $\psi^m = \underline{\underline{\mathbf{0}}}$ de $\psi^{m-1} \neq \underline{\underline{\mathbf{0}}}$, akkor

$$\begin{array}{ccccccccc} \operatorname{Im}(\psi) \supset \operatorname{Im}(\psi^2) \supset \cdots \supset \operatorname{Im}(\psi^m) & = & \operatorname{Im}(\psi^{m+1}) & = & \operatorname{Im}(\psi^{m+2}) & = \cdots = & \{\mathbf{0}\}, \\ \operatorname{Ker}(\psi) \subset \operatorname{Ker}(\psi^2) \subset \cdots \subset \operatorname{Ker}(\psi^m) & = & \operatorname{Ker}(\psi^{m+1}) & = & \operatorname{Ker}(\psi^{m+2}) & = \cdots = & W. \end{array}$$

Bizonyítás. Az F.1. Lemmát kell alkalmazni, felhasználva, hogy $\psi^m = \underline{\underline{\mathbf{0}}}$ és $\psi^{m-1} \neq \underline{\underline{\mathbf{0}}}$. □

5.10. Megjegyzés. A nilpotenciának köszönhetően további összefüggések is fennállnak a képterek és a magterek között: $\operatorname{Im}\psi^i \subseteq \operatorname{Ker}\psi^{m-i}$ minden $i = 1, 2, \ldots, m-1$ esetén. Valóban, ha $\mathbf{v} \in \operatorname{Im}\psi^i$, akkor $\mathbf{v}$ előáll $\mathbf{v} = \mathbf{u}\psi^i$ alakban, és így $\mathbf{v}\psi^{m-i} = (\mathbf{u}\psi^i)\psi^{m-i} = \mathbf{u}\psi^m = \mathbf{0}$. Figyeljük meg ezt az 1. ábrán!

5.11. Definíció. Azt mondjuk, hogy a $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{\ell-1} \in W$ nemnulla vektorok Jordan-láncot alkotnak, ha $\mathbf{v}_i\psi = \mathbf{v}_{i+1}$ teljesül minden $i = 0, 1, \ldots, \ell-2$ esetén.

A $\mathbf{v}_0$ ból induló Jordan-lánc tehát nem más, mint $\mathbf{v}_0$ képeinek sorozata a ψ transzformáció iterálása során:

$$\mathbf{v}_0 \xrightarrow{\psi} \mathbf{v}_1 \xrightarrow{\psi} \mathbf{v}_2 \xrightarrow{\psi} \cdots \xrightarrow{\psi} \mathbf{v}_{\ell-1} \xrightarrow{\psi} \mathbf{0},$$

de a nullvektort már nem tekintjük a Jordan-lánc tagjának. Az 1. ábrán hét Jordan-lánc látható pirossal kiemelve (rendre 4, 4, 3, 3, 2, 1, 1 hosszúságúak).

Célunk diszjunkt Jordan-láncokból álló bázist találni a W vektortérben. Ehhez persze szükséges, hogy az egyes Jordan-láncok lineárisan független vektorrendszerek legyenek. Ezt bizonyítjuk be a következő tételben, és azt is megmutatjuk, hogy milyen szép ψ egy Jordan-lánc által kifeszített altérre való megszorításának mátrixa.

5.12. Tétel. Legyen $\psi : W \to W$ nilpotens lineáris transzformációja a W véges dimenziós vektortérnek, és tegyük fel, hogy a $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{\ell-1} \in W$ vektorok egy Jordan-láncot alkotnak.

(i) A $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{\ell-1}$ vektorok lineárisan függetlenek.

(ii) Az $U = [\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{\ell-1}]$ altér invariáns ψ-re.

(iii) A ψ transzformáció U altérre való megszorításának mátrixa a $\mathcal{J} = \{\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{\ell-1}\}$ bázisban így fest:

$$[\![\psi|_U]\!]_{\mathcal{J}} = \begin{pmatrix} 0 & 1 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & \ldots & 0 & 0 \end{pmatrix}.$$

Bizonyítás.

(i) Indirekten bizonyítunk: tegyük fel, hogy $\sum_{i=0}^{\ell-1} \lambda_i \mathbf{v}_i = 0$, de az együtthatók nem mind nullák. Legyen λ_s az első nemnulla együttható, ekkor tehát $\lambda_s \neq 0$, de minden $i < s$ indexre $\lambda_i = 0$. Alkalmazzuk a $\psi^{\ell-s-1}$ transzformációt a lineáris kombinációra, és használjuk a linearitást (az s-nél kisebb indexű tagokat elhagyjuk, mert azok 0 együtthatóval szerepelnek):

$$\left(\sum_{i=0}^{\ell-1} \lambda_i \mathbf{v}_i\right)\psi^{\ell-s-1} = \sum_{i=s}^{\ell-1} \lambda_i \mathbf{v}_i \psi^{\ell-s-1} = \lambda_s \mathbf{v}_s \psi^{\ell-s-1} + \lambda_{s+1}\mathbf{v}_{s+1}\psi^{\ell-s-1} + \cdots + \lambda_{\ell-1}\mathbf{v}_{\ell-1}\psi^{\ell-s-1} = \mathbf{0}.$$

A Jordan-lánc definíciója szerint a ψ leképezés hatására minden vektor eggyel „jobbra” lép a láncban: $\mathbf{v}_i\psi = \mathbf{v}_{i+1}$, de amikor az index elérné vagy meghaladná az ℓ értéket, akkor a nullvektort kapjuk (hiszen $\mathbf{v}_{\ell-1}\psi = \mathbf{0}$). A $\psi^{\ell-s-1}$ transzformáció hatására pedig minden vektor $\ell - s - 1$ lépést tesz jobbra, így $\mathbf{v}_s\psi^{\ell-s-1} = \mathbf{v}_{\ell-1}$, továbbá a $\mathbf{v}_{s+1}, \ldots, \mathbf{v}_{\ell-1}$ képe már $\mathbf{0}$ lesz. Tehát a fenti összegből csak egy tag marad: $\lambda_s\mathbf{v}_s\psi^{\ell-s-1} = \lambda_s\mathbf{v}_{\ell-1} = \mathbf{0}$. Ez ellentmondás, hiszen feltettük, hogy $\lambda_s \neq 0$, és a Jordan-lánc definíciója miatt $\mathbf{v}_{\ell-1}$ nem a nullvektor.

(ii) Ha $0 \leq i \leq \ell - 1$, akkor $\mathbf{v}_i\psi = \mathbf{v}_{i+1} \in U$, ha pedig $i = \ell - 1$, akkor $\mathbf{v}_i\psi = \mathbf{0} \in U$. Ebből következik, hogy az U altér valóban invariáns ψ-re.

(iii) A Jordan-lánc definíciója alapján a $\mathbf{v}_i\psi$ vektorok koordinátái a $\mathcal{J}$ bázisban a következőképpen alakulnak:

	$\mathbf{v}_0$	$\mathbf{v}_1$	$\mathbf{v}_2$	$\ldots$	$\mathbf{v}_{\ell-2}$	$\mathbf{v}_{\ell-1}$
$\mathbf{v}_0\psi$	0	1	0	$\ldots$	0	0
$\mathbf{v}_1\psi$	0	0	1	$\ldots$	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$\mathbf{v}_{\ell-3}\psi$	0	0	0	$\ldots$	1	0
$\mathbf{v}_{\ell-2}\psi$	0	0	0	$\ldots$	0	1
$\mathbf{v}_{\ell-1}\psi$	0	0	0	$\ldots$	0	0

□

Következzék a fenti tétel első állításának „nagy testvére”.

5.13. Tétel. Ha $\psi : W \to W$ nilpotens lineáris transzformációja a W véges dimenziós vektortérnek, akkor létezik W-nek ψ diszjunkt Jordan-láncaiból álló bázisa.

Ennek a tételnek a bizonyítása hosszú, ezért a fejezet végére „száműzzük” (lásd az F.3. fejezetet), és előbb learatjuk a gyümölcseit.

5.3. A Jordan-normálalak

Idézzük fel az 5.1. fejezet végén bevezetett jelöléseket: ψ_i a $\varphi - \lambda_i\,\mathrm{id}_V$ transzformáció megszorítása a $W_i = \mathrm{Ker}(A - \lambda_i E)^{a_i}$ altérre. Mivel ψ_i nilpotens transzformációja a W_i altérnek, az 5.13. Tétel szerint létezik W_i-nek olyan $\mathcal{F}_i$ bázisa, ami diszjunkt Jordan-láncokból áll. Ebben a bázisban ψ_i mátrixa blokkdiagonális lesz, ahol a blokkok az 5.12. Tétel harmadik állításában leírt módon néznek ki (annyi blokk van, ahány Jordan-lánc alkotja az $\mathcal{F}_i$ bázist, a blokkok méretei pedig megegyeznek a Jordan-láncok hosszaival). Például így fest a mátrix, ha a bázist három Jordan-lánc alkotja, melyek rendre $4, 3, 3$ vektorból állnak:

$$[\![\psi_i]\!]_{\mathcal{F}_i} = M_i = \left(\begin{array}{cccc|ccc|ccc} 0 & \boxed{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \boxed{1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & \boxed{1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Figyeljük meg, hogy csak a főátló felett vannak egyesek (minden más elem nulla), és az egyesek sorát a blokkok határán egy-egy nulla szakítja meg. A φ transzformáció W_i-re való megszorításának mátrixa csak annyiban különbözik a fentitől, hogy a főátlón mindenütt λ_i áll:

$$[\![\varphi_i]\!]_{\mathcal{F}_i} = M_i + \lambda_i E_i = \left(\begin{array}{cccc|ccc|ccc} \lambda_i & \boxed{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_i & \boxed{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_i & \boxed{1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_i & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \lambda_i & \boxed{1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_i & \boxed{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_i & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_i & \boxed{1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_i & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_i \end{array}\right).$$

Hogy könnyebben le tudjuk írni az ilyen szerkezetű mátrixokat, bevezetünk egy jelölést.

5.14. Definíció. Tetszőleges $\lambda \in \mathbb{C}$ és $t \in \mathbb{N}$ esetén a λ-hoz tartozó t méretű Jordan-blokkon a következő mátrixot értjük:

$$J_{\lambda,t} = \begin{pmatrix} \lambda & 1 & \dots & 0 & 0 \\ 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & \dots & 0 & \lambda \end{pmatrix} \in \mathbb{C}^{t\times t}.$$

(Figyeljük meg, hogy a $J_{\lambda,t}$ mátrix egyetlen sajátértéke λ. Keresztkérdés: mennyi ennek a sajátértéknek az algebrai, illetve geometriai multiplicitása?) Ezzel a jelöléssel a fenti példában szereplő 10×10-es mátrix így írható le tömörebben:

$$[\![\varphi_i]\!]_{\mathcal{F}_i} = M_i + \lambda_i E_i = \begin{pmatrix} J_{\lambda_i,4} & 0 & 0 \\ 0 & J_{\lambda_i,3} & 0 \\ 0 & 0 & J_{\lambda_i,3} \end{pmatrix} = \operatorname{diag}(J_{\lambda_i,4}, J_{\lambda_i,3}, J_{\lambda_i,3}).$$

Mivel $V = W_1 \oplus \dots \oplus W_k$ (lásd az 5.5. Tételt), az $\mathcal{F}_i$ bázisok egyesítése bázisa V-nek, és ebben a bázisban φ mátrixa a fentihez hasonló mátrixokból álló blokkdiagonális mátrix (egy $M_i + \lambda_i E_i$ mátrix több Jordan-blokkból állhat!):

$$[\![\varphi]\!]_{\mathcal{F}} = \begin{pmatrix} [\![\varphi_1]\!]_{\mathcal{F}_1} & \underline{\underline{0}} & \dots & \underline{\underline{0}} \\ \underline{\underline{0}} & [\![\varphi_2]\!]_{\mathcal{F}_2} & \dots & \underline{\underline{0}} \\ \vdots & \vdots & \ddots & \vdots \\ \underline{\underline{0}} & \underline{\underline{0}} & \dots & [\![\varphi_k]\!]_{\mathcal{F}_k} \end{pmatrix} = \begin{pmatrix} M_1 + \lambda_1 E_1 & \underline{\underline{0}} & \dots & \underline{\underline{0}} \\ \underline{\underline{0}} & M_2 + \lambda_2 E_2 & \dots & \underline{\underline{0}} \\ \vdots & \vdots & \ddots & \vdots \\ \underline{\underline{0}} & \underline{\underline{0}} & \dots & M_k + \lambda_k E_k \end{pmatrix}.$$

Azt kaptuk tehát, hogy a komplex számtest felett minden lineáris transzformációhoz van olyan bázis, melyben a transzformáció mátrixa a fenti alakú, azaz Jordan-blokkokból álló blokkdiagonális mátrix. Figyelembe véve, hogy két mátrix akkor és csak akkor hasonló, ha ugyanannak a transzformációnak különböző bázisokban felírt mátrixai, kimondhatjuk a következő tételt.

5.15. Tétel. Minden $A \in \mathbb{C}^{n\times n}$ mátrix hasonló egy blokkdiagonális alakú $J \in \mathbb{C}^{n\times n}$ mátrixhoz, melynek blokkjai mind Jordan-blokkok:

$$J = \operatorname{diag}(J_{\lambda_1,t_{1,1}}, \dots, J_{\lambda_1,t_{1,g_1}}, \ \dots, \ J_{\lambda_k,t_{k,1}}, \dots, J_{\lambda_k,t_{k,g_k}}).$$

Itt $\lambda_1, \dots, \lambda_k$ az A mátrix sajátértékékei. A J mátrixot (amely a benne szereplő Jordan-blokkok sorrendjétől eltekintve egyértelműen meghatározott), az A mátrix Jordan-normálalakjának nevezzük.

5.16. Megjegyzés. Meg lehet mutatni, hogy két Jordan-normálalak akkor és csak akkor hasonló, ha legfeljebb csak a bennük szereplő Jordan-blokkok sorrendjében térnek el. Így a Jordan-normálalakok ismeretében nagyon könnyű eldönteni, hogy a komplex számtest felett hasonló-e két mátrix: $A \sim B$ akkor és csak akkor teljesül, ha A-nak és B-nek ugyanazok a sajátértékei, és minden egyes λ sajátérték esetén λ-hoz ugyanannyi és ugyanolyan méretű Jordan-blokk tartozik A-nál, mint B-nél.

5.17. Megjegyzés. Nem nehéz belátni, hogy a λ_i sajátértékhez tartozó Jordan-blokkok száma nem más, mint λ_i geometriai multiplicitása, ezért használtuk erre a fenti tételben a g_i jelölést. Az is világos, hogy λ_i algebrai multiplicitása a hozzá tartozó Jordan-blokkok méreteinek összege: $a_i = t_{i,1} + \cdots + t_{i,g_i}$. (Keresztkérdés: hogyan lehet a Jordan-normálalakból kiolvasni a minimálpolinomot?)

6. Euklideszi terek

6.1. Bilineáris leképezések

6.1. Definíció. Legyenek U és V vektorterek a T test felett. Azt mondjuk, hogy $\ell\colon U \times V \to T$ bilineáris leképezés, ha rendelkezik az alábbi tulajdonságokkal:

(1) $\forall \mathbf{u}_1, \mathbf{u}_2 \in U \ \forall \mathbf{v} \in V\colon \ell(\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}) = \ell(\mathbf{u}_1, \mathbf{v}) + \ell(\mathbf{u}_2, \mathbf{v})$;

(2) $\forall \mathbf{u} \in U \ \forall \mathbf{v} \in V \ \forall \lambda \in T\colon \ell(\lambda\mathbf{u}, \mathbf{v}) = \lambda\ell(\mathbf{u}, \mathbf{v})$;

(3) $\forall \mathbf{u} \in U \ \forall \mathbf{v}_1, \mathbf{v}_2 \in V\colon \ell(\mathbf{u}, \mathbf{v}_1 + \mathbf{v}_2) = \ell(\mathbf{u}, \mathbf{v}_1) + \ell(\mathbf{u}, \mathbf{v}_2)$;

(4) $\forall \mathbf{u} \in U \ \forall \mathbf{v} \in V \ \forall \lambda \in T\colon \ell(\mathbf{u}, \lambda\mathbf{v}) = \lambda\ell(\mathbf{u}, \mathbf{v})$.

6.2. Tétel. Legyen $\ell\colon U \times V \to T$ bilineáris leképezés, legyen $\mathcal{E}\colon \mathbf{e}_1, \ldots, \mathbf{e}_n$ bázisa U-nak és $\mathcal{F}\colon \mathbf{f}_1, \ldots, \mathbf{f}_m$ bázisa V-nek. Ekkor létezik egy egyértelműen meghatározott $A \in T^{n\times m}$ mátrix, amelyre minden $\mathbf{u} \in U$ és $\mathbf{v} \in V$ esetén

$$\ell(\mathbf{u}, \mathbf{v}) = [\![\mathbf{u}]\!]_{\mathcal{E}} \cdot A \cdot [\![\mathbf{v}]\!]_{\mathcal{F}}^{\mathrm{T}}.$$

Bizonyítás. Az egzisztencia bizonyításához tekintsünk tetszőleges $\mathbf{u} \in U$ és $\mathbf{v} \in V$ vektorokat, és legyen $[\![\mathbf{u}]\!]_{\mathcal{E}} = \mathbf{x} = (x_1, \ldots, x_n)$ és $[\![\mathbf{v}]\!]_{\mathcal{F}} = \mathbf{y} = (y_1, \ldots, y_m)$. A bilinearitást kihasználva így számolhatjuk ki $\ell(\mathbf{u}, \mathbf{v})$ értékét:

$$\ell(\mathbf{u}, \mathbf{v}) = \ell\left(\sum_{i=1}^{n} x_i\mathbf{e}_i,\ \sum_{j=1}^{m} y_j\mathbf{f}_j\right) = \sum_{i,j} x_i y_j \ell(\mathbf{e}_i, \mathbf{f}_j).$$

Ez épp azt jelenti, hogy $a_{ij} = \ell(\mathbf{e}_i, \mathbf{f}_j)$ esetén az $A = (a_{ij}) \in T^{n\times m}$ mátrixra $\ell(\mathbf{u}, \mathbf{v}) = \mathbf{x}A\mathbf{y}^{\mathrm{T}} = [\![\mathbf{u}]\!]_{\mathcal{E}} \cdot A \cdot [\![\mathbf{v}]\!]_{\mathcal{F}}^{\mathrm{T}}$.
Az egyértelműség igazolásához tfh. az $A = a_{ij}$ mátrix rendelkezik a kívánt tulajdonsággal. Az $\mathbf{u} = \mathbf{e}_i$, $\mathbf{v} = \mathbf{f}_j$ szereposztással azt kapjuk, hogy

$$\ell(\mathbf{e}_i, \mathbf{f}_j) = [\![\mathbf{e}_i]\!]_{\mathcal{E}} \cdot A \cdot [\![\mathbf{f}_j]\!]_{\mathcal{F}}^{\mathrm{T}} = (0, \ldots, 0, \overset{i}{\smile}{1}, 0, \ldots, 0) \cdot A \cdot (0, \ldots, 0, \overset{j}{\smile}{1}, 0, \ldots, 0)^{\mathrm{T}} = a_{ij}.$$

Tehát A nem lehet más, mint az $\ell(\mathbf{e}_i, \mathbf{f}_j)$ számokból alkotott mátrix. □

6.3. Definíció. A fenti tételben szereplő A mátrixot az ℓ bilineáris leképezés mátrixának nevezzük (az $\mathcal{E}$ és $\mathcal{F}$ bázisokban). Jelölés: $A = [\![\ell]\!]_{\mathcal{E},\mathcal{F}}$. Ha $U = V$ és $\mathcal{E} = \mathcal{F}$, akkor $[\![\ell]\!]_{\mathcal{E},\mathcal{E}}$ helyett használjuk az $[\![\ell]\!]_{\mathcal{E}}$ egyszerűsített jelölést. Az $\ell(\mathbf{u}, \mathbf{v}) = \sum_{i,j} a_{ij} x_i y_j$ felírást a bilineáris leképezés koordinátás alakjának hívjuk.

6.4. Példa. A $V = \mathbb{R}^2$ vektortéren a skaláris szorzat bilineáris leképezés: $\ell(\mathbf{u}, \mathbf{v}) = |\mathbf{u}| \cdot |\mathbf{v}| \cdot \cos\alpha$, ahol α az $\mathbf{u}$ és $\mathbf{v}$ vektorok által bezárt szög. A koordinátás alak: $x_1y_1 + x_2y_2$.

6.5. Példa. Tekintsük a $V = \mathbb{R}^3$ vektortéren az alábbi bilineáris leképezést:

$$\ell(\mathbf{x}, \mathbf{y}) = \ell((x_1, x_2, x_3),\ (y_1, y_2, y_3)) = x_1y_1 + x_2y_2 + 3x_3y_3 + 2x_1y_2 + 2x_2y_1 + x_1y_3 + x_3y_1 + x_2y_3 + x_3y_2.$$

Ez már eleve koordinátás alakban van megadva, ezért a standard bázisbeli mátrixát egyszerűen leolvashatjuk az együtthatókból:

$$A = [\![\ell]\!]_{\mathcal{E}} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 3 \end{pmatrix}.$$

(Az a tény, hogy ℓ felírható $\ell(\mathbf{x}, \mathbf{y}) = \mathbf{x}A\mathbf{y}^{\mathrm{T}}$ alakban egyúttal igazolja is, hogy ℓ bilineáris.)

6.6. Tétel. Legyen $\ell\colon V \times V \to T$ bilineáris leképezés, és legyen $\mathcal{E}$, $\mathcal{F}$ két bázisa V-nek. Ekkor

$$[\![\ell]\!]_{\mathcal{F}} = [\![\mathcal{F} \to \mathcal{E}]\!] \cdot [\![\ell]\!]_{\mathcal{E}} \cdot [\![\mathcal{F} \to \mathcal{E}]\!]^{\mathrm{T}}.$$

Bizonyítás. Az egyszerűség kedvéért használjuk az $A = [\![\ell]\!]_{\mathcal{E}}$ és $S = [\![\mathcal{F} \to \mathcal{E}]\!]$ jelöléseket, továbbá legyen $\mathbf{x} = [\![\mathbf{u}]\!]_{\mathcal{E}}$, $\mathbf{x}' = [\![\mathbf{u}]\!]_{\mathcal{F}}$ és $\mathbf{y} = [\![\mathbf{v}]\!]_{\mathcal{E}}$, $\mathbf{y}' = [\![\mathbf{v}]\!]_{\mathcal{F}}$, ahol $\mathbf{u}, \mathbf{v} \in V$ tetszőleges vektorok. A 3.35. Állítás első pontja szerint $\mathbf{x} = \mathbf{x}'S$ és hasonlóan $\mathbf{y} = \mathbf{y}'S$. Helyettesítsük be ezeket a bilineáris leképezés $\mathcal{E}$ bázisbeli mátrixát definiáló $\ell(\mathbf{u}, \mathbf{v}) = \mathbf{x}A\mathbf{y}^{\mathrm{T}}$ összefüggésbe:

$$\ell(\mathbf{u}, \mathbf{v}) = \mathbf{x}A\mathbf{y}^{\mathrm{T}} = (\mathbf{x}'S)A(\mathbf{y}'S)^{\mathrm{T}} = \mathbf{x}'(SAS^{\mathrm{T}})\mathbf{y}'^{\mathrm{T}}.$$

A 6.2. Tétel egyértelműségi része szerint ebből következik, hogy $[\![\ell]\!]_{\mathcal{F}} = SAS^{\mathrm{T}}$. □

$$\ell(\mathbf{u}, \mathbf{v}) = [\![\mathbf{u}]\!]_{\mathcal{E}} \cdot [\![\ell]\!]_{\mathcal{E},\mathcal{F}} \cdot [\![\mathbf{v}]\!]_{\mathcal{F}}^{\mathrm{T}} = \mathbf{x}A\mathbf{y}^{\mathrm{T}} = \sum_{i,j} a_{ij}x_iy_j$$

$$[\![\ell]\!]_{\mathcal{F}} = [\![\mathcal{F} \to \mathcal{E}]\!] \cdot [\![\ell]\!]_{\mathcal{E}} \cdot [\![\mathcal{F} \to \mathcal{E}]\!]^{\mathrm{T}} = SAS^{\mathrm{T}}.$$

6.7. Definíció. Azt mondjuk, hogy az $\ell\colon V \times V \to T$ bilineáris leképezés szimmetrikus, ha bármely $\mathbf{u}, \mathbf{v} \in V$ vektorokra $\ell(\mathbf{u}, \mathbf{v}) = \ell(\mathbf{v}, \mathbf{u})$ teljesül.

6.8. Tétel. Tetszőleges $\ell\colon V \times V \to T$ bilineáris leképezés esetén ekvivalensek az alábbiak:

(i) ℓ szimmetrikus;

(ii) ℓ mátrixa V bármely bázisában szimmetrikus;

(iii) ℓ mátrixa V valamely bázisában szimmetrikus.

Bizonyítás. Használjuk a 6.6. Tétel bizonyításában bevezetett jelöléseket. Világos, hogy (i) $\Longrightarrow$ (ii), hiszen ha ℓ szimmetrikus, akkor $a_{ij} = \ell(\mathbf{e}_i, \mathbf{e}_j) = \ell(\mathbf{e}_j, \mathbf{e}_i) = a_{ji}$. A (ii) $\Longrightarrow$ (iii) következtetés (majdnem) triviális. Végül (iii) $\Longrightarrow$ (i) bizonyításához tfh. a V vektortér valamely $\mathcal{E}$ bázisára $A = [\![\ell]\!]_{\mathcal{E}}$ szimmetrikus mátrix. Tetszőleges $\mathbf{u}, \mathbf{v} \in V$ vektorokra $\ell(\mathbf{u}, \mathbf{v}) \in T$ egy 1×1-es mátrix, így megegyezik a saját transzponáltjával: $\ell(\mathbf{u}, \mathbf{v})^{\mathrm{T}} = \ell(\mathbf{u}, \mathbf{v})$. Ezt, valamint A szimmetriáját kihasználva belátjuk, hogy $\ell(\mathbf{u}, \mathbf{v}) = \ell(\mathbf{v}, \mathbf{u})$:

$$\ell(\mathbf{u}, \mathbf{v}) = \ell(\mathbf{u}, \mathbf{v})^{\mathrm{T}} = (\mathbf{x}A\mathbf{y}^{\mathrm{T}})^{\mathrm{T}} = \mathbf{y}A^{\mathrm{T}}\mathbf{x}^{\mathrm{T}} = \mathbf{y}A\mathbf{x}^{\mathrm{T}} = \ell(\mathbf{v}, \mathbf{u}).$$ □

6.2. Kvadratikus alakok

6.9. Definíció. Legyen V vektortér a T test felett. Azt mondjuk, hogy $q\colon V \to T$ kvadratikus alak, ha létezik olyan $\ell\colon V \times V \to T$ szimmetrikus bilineáris leképezés, amelyre $q(\mathbf{v}) = \ell(\mathbf{v}, \mathbf{v})$ minden $\mathbf{v} \in V$ vektorra.

6.10. Tétel. Ha T számtest, akkor bármely T feletti kvadratikus alak egyértelműen meghatározza a hozzá tartozó szimmetrikus bilineáris leképezést.

Bizonyítás. Tfh. a q kvadratikus alak az ℓ szimmetrikus bilineáris leképezésből származik, azaz $q(\mathbf{v}) = \ell(\mathbf{v}, \mathbf{v})$ minden $\mathbf{v} \in V$ vektorra. Számítsuk ki $q(\mathbf{u} + \mathbf{v})$ értékét, felhasználva ℓ bilinearitását és szimmetriáját:

$$q(\mathbf{u} + \mathbf{v}) = \ell(\mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v}) = \ell(\mathbf{u}, \mathbf{u}) + \ell(\mathbf{u}, \mathbf{v}) + \ell(\mathbf{v}, \mathbf{u}) + \ell(\mathbf{v}, \mathbf{v}) = q(\mathbf{u}) + 2\ell(\mathbf{u}, \mathbf{v}) + q(\mathbf{v}).$$

Innen már könnyen kifejezhetjük az ℓ bilineáris leképezést a q kvadratikus alakból:

$$\ell(\mathbf{u}, \mathbf{v}) = \frac{1}{2} \cdot (q(\mathbf{u} + \mathbf{v}) - q(\mathbf{u}) - q(\mathbf{v})). \tag{3}$$ □

6.11. Megjegyzés. A fenti tétel nemcsak számtestekre igaz, hanem minden olyan testre, ahol $1 + 1 \neq 0$ (itt 0 és 1 a test additív és multiplikatív egységelemét jelöli). Például a $T = \mathbb{Z}_p$ testre $p > 2$ esetén igaz a tétel (keressünk példát, ami mutatja, hogy $p = 2$ esetén nem igaz!). A (3) összefüggést szokás polarizációs azonosságnak nevezni.

6.12. Megjegyzés. Ha az $\ell\colon V \times V \to T$ szimmetrikus bilineáris leképezés mátrixa az $\mathcal{E}$ bázisban $A = [\![\ell]\!]_{\mathcal{E}}$, akkor a megfelelő $q(\mathbf{v}) = \ell(\mathbf{v}, \mathbf{v})$ kvadratikus alak koordinátás alakja egy T feletti homogén másodfokú polinom:

$$q(\mathbf{v}) = \ell(\mathbf{v}, \mathbf{v}) = \mathbf{x}A\mathbf{x}^{\mathrm{T}} = \sum_{i,j} a_{ij}x_ix_j \qquad (\text{ahol } \mathbf{x} = [\![\mathbf{v}]\!]_{\mathcal{E}}).$$

6.13. Példa. A 6.5. Példában szereplő bilineáris leképezés szimmetrikus; a neki megfelelő kvadratikus alak:

$$q(\mathbf{x}) = \ell(\mathbf{x}, \mathbf{x}) = x_1^2 + x_2^2 + 3x_3^2 + 4x_1x_2 + 2x_1x_3 + 2x_2y_3 = \mathbf{x}A\mathbf{x}^{\mathrm{T}}.$$

Több olyan $B \in \mathbb{R}^{3\times 3}$ mátrix is van, amelyre $q(\mathbf{x}) = \mathbf{x}B\mathbf{x}^{\mathrm{T}}$, például

$$q(\mathbf{x}) \;=\; \mathbf{x} \cdot \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 3 \end{pmatrix} \cdot \mathbf{x}^{\mathrm{T}} \;=\; \mathbf{x} \cdot \begin{pmatrix} 1 & 3 & 0 \\ 1 & 1 & 2 \\ 2 & 0 & 3 \end{pmatrix} \cdot \mathbf{x}^{\mathrm{T}}.$$

Megfelelő szimmetrikus mátrix azonban csak egy van: a kvadratikus alakban x_ix_j együtthatóját „igazságosan” kell elosztani a mátrix a_{ij} és a_{ji} eleme között minden $i \neq j$ esetén.

6.14. Definíció. A $q\colon V \to T$ kvadratikus alak $\mathcal{E}$ bázisbeli $[\![q]\!]_{\mathcal{E}}$ mátrixán a neki megfelelő szimmetrikus bilineáris leképezés mátrixát értjük (itt fel kell tennünk, hogy T számtest vagy legalábbis olyan test, amelyben $1 + 1 \neq 0$; lásd a 6.11. Megjegyzést). Tehát $[\![q]\!]_{\mathcal{E}}$ nem más, mint az az egyértelműen meghatározott $A \in T^{n\times n}$ szimmetrikus mátrix, amelyre $q(\mathbf{v}) = \mathbf{x}A\mathbf{x}^{\mathrm{T}}$ minden $\mathbf{v} \in V$ vektorra (szokás szerint az $\mathbf{x} = [\![\mathbf{v}]\!]_{\mathcal{E}}$ jelölést használva).

6.15. Definíció. Legyen $q(\mathbf{x}) = \mathbf{x}A\mathbf{x}^{\mathrm{T}} = \sum a_{ij}x_ix_j$ kvadratikus alak a T^n vektortéren (vagy, ha úgy tetszik, $T[x_1, \ldots, x_n]$-beli homogén másodfokú polinom). Vezessük be az új $y_1, \ldots, y_n$ változókat, amelyek lineáris függvényei az $x_1, \ldots, x_n$ változóknak:

$$\begin{aligned} y_1 &= c_{11}x_1 + c_{21}x_2 + \ldots c_{n1}x_n; \\ y_2 &= c_{12}x_1 + c_{22}x_2 + \ldots c_{n2}x_n; \\ &\vdots \\ y_n &= c_{1n}x_1 + c_{2n}x_2 + \ldots c_{nn}x_n. \end{aligned}$$

A $C = (c_{ij}) \in T^{n\times n}$ mátrix segítségével tömörebben is leírhatjuk a változótranszfomációt: $\mathbf{y} = \mathbf{x}C$. Feltesszük, hogy C nemelfajuló mátrix; ekkor a változócsere „visszacsinálható”: $\mathbf{x} = \mathbf{y}S$, ahol $S = C^{-1}$. A q kvadratikus alakot az $y_1, \ldots, y_n$ változók függvényeként felírva egy új q' kvadratikus alakot kapunk, melynek mátrixa SAS^{T}:

$$q(\mathbf{x}) = q(\mathbf{y}S) = \big(\mathbf{y}S\big)A\big(\mathbf{y}S^{\mathrm{T}}\big) = \mathbf{y}SAS^{\mathrm{T}}\mathbf{y}^{\mathrm{T}} =: q'(\mathbf{y}).$$

Ilyenkor azt mondjuk, hogy a q' kvadratikus alakot nemelfajuló lineáris helyettesítéssel kaptuk a q kvadratikus alakból.

6.16. Megjegyzés. Nemelfajuló lineáris helyettesítések egymásutánja is nemelfajuló lineáris helyettesítés: ha az első helyettesítés mátrixa C_1, a másodiké pedig C_2, akkor a két helyettesítés egymás utáni elvégzését a C_1C_2 mátrix írja le, és persze $\det(C_1C_2) = \det(C_1) \cdot \det(C_2) \neq 0$.

6.17. Megjegyzés. A 6.6. Tétel szerint ha a $q\colon V \to T$ kvadratikus alak mátrixa az $\mathcal{E}$ bázisban $A = [\![q]\!]_{\mathcal{E}}$, az $\mathcal{F}$ bázisban pedig $B = [\![q]\!]_{\mathcal{F}}$, akkor $B = SAS^{\mathrm{T}}$, ahol $S = [\![\mathcal{F} \to \mathcal{E}]\!]$. Ezt összevetve a 6.15. Definícióval azt látjuk, hogy a nemelfajuló lineáris helyettesítés lényegében ugyanaz, mint a más bázisra való áttérés, hiszen a bázisáttérések mátrixai pontosan a nemelfajuló mátrixok (lásd a 3.36. Tételt).

6.18. Definíció. A $q\colon V \to T$ kvadratikus alak rangján mátrixának rangját értjük. Ez nem függ a bázis megválasztásától, mert ha A és B két különböző bázisbeli mátrixa a q kvadratikus alaknak, akkor $B = SAS^{\mathrm{T}}$, ahol S a bázisátmenet mátrixa. Mivel S nemelfajuló mátrix, $\mathrm{r}(A) = \mathrm{r}(SAS^{\mathrm{T}}) = \mathrm{r}(B)$ a 2.75. Tétel szerint.

6.19. Definíció. Azt mondjuk, hogy a $q\colon V \to T$ kvadratikus alak az $\mathcal{E}$ bázisban kanonikus alakú, ha $[\![q]\!]_{\mathcal{E}}$ diagonális mátrix. A koordinátás alak ebben az esetben így fest (az $A = (a_{ij}) = [\![q]\!]_{\mathcal{E}}$ jelöléssel):

$$q(\mathbf{v}) = \mathbf{x}A\mathbf{x}^{\mathrm{T}} = \sum_{i,j} a_{ij}x_ix_j = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 \qquad (\text{ahol } \mathbf{x} = [\![\mathbf{v}]\!]_{\mathcal{E}}).$$

6.20. feladat. Hozzuk nemelfajuló lineáris helyettesítéssel kanonikus alakra az alábbi q kvadratikus alakot:

$$q = x_1^2 + x_2^2 + 3x_3^2 + 4x_1x_2 + 2x_1x_3 + 2x_2x_3.$$

Megoldás. Alakítsuk teljes négyzetté az x_1-et tartalmazó tagokat:

$$q = (x_1 + 2x_2 + x_3)^2 - 3x_2^2 + 2x_3^2 - 2x_2x_3.$$

Vezessük be x_1 helyett az $y_1 = x_1 + 2x_2 + x_3$ új változót.:

$$q = y_1^2 - 3x_2^2 + 2x_3^2 - 2x_2x_3.$$

Alakítsuk teljes négyzetté az x_2-t tartalmazó tagokat:

$$q = y_1^2 - 3\left(x_2 + \frac{1}{3}x_3\right)^2 + \frac{7}{3}x_3^2.$$

Vezessük be x_2 helyett az $y_2 = x_2 + \frac{1}{3}x_3$ új változót, és – csak a szépség kedvéért – nevezzük át x_3-at y_3-ra. Ezzel meg is kapjuk q egy kanonikus alakját:

$$q = y_1^2 - 3y_2^2 + \frac{7}{3}y_3^2.$$

Ellenőrzés. Írjuk fel a helyettesítés C mátrixát:

$$(y_1, y_2, y_3) = \mathbf{y} = \mathbf{x}C = (x_1, x_2, x_3) \cdot \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 1/3 & 1 \end{pmatrix}.$$

A C mátrix determinánsa 1, ezért a helyettesítés valóban nemelfajuló. Ellenőrizzük magát a kanonikus alakot is (lehet számítógéppel is):

$$y_1^2 - 3y_2^2 + \frac{7}{3}y_3^2 = (x_1 + 2x_2 + x_3)^2 - 3\left(x_2 + \frac{1}{3}x_3\right)^2 + \frac{7}{3}x_3^2 = x_1^2 + x_2^2 + 3x_3^2 + 4x_1x_2 + 2x_1x_3 + 2x_2x_3. \checkmark$$

Ellenőrizhetjük a megoldást mátrixokkal is. Legyen A az eredeti kvadratikus alak mátrixa, és legyen $S = C^{-1}$. Ekkor SAS^{T} diagonális mátrix lesz, és a főátlójában épp a kanonikus alak együtthatói szerepelnek:

$$SAS^{\mathrm{T}} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1/3 & -1/3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & -2 & -1/3 \\ 0 & 1 & -1/3 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 7/3 \end{pmatrix}.$$ ◇

6.21. Tétel. Legyen $q\colon V \to T$ kvadratikus alak a T számtest feletti V végesdimenziós vektortéren. Ekkor létezik V-nek olyan $\mathcal{E}$ bázisa, amelyben q kanonikus alakú, azaz $[\![q]\!]_{\mathcal{E}}$ diagonális mátrix.

Bizonyítás. Lásd [SzL] 15.7. Tétel. □

6.22. Megjegyzés. A 6.17. Megjegyzés szerint a fenti tétel így is megfogalmazható: minden számtest feletti kvadratikus alak nemelfajuló lineáris helyettesítéssel kanonikus alakra hozható. Egy másik megfogalmazás: ha T számtest, akkor minden $A \in T^{n\times n}$ szimmetrikus mátrixhoz van olyan $S \in T^{n\times n}$ nemelfajuló mátrix, amelyre SAS^{T} diagonális mátrix.

6.3. Valós kvadratikus alakok

6.23. Definíció. Azt mondjuk, hogy a $q\colon V \to \mathbb{R}$ valós kvadratikus alak az $\mathcal{E}$ bázisban normálalakú, ha $[\![q]\!]_{\mathcal{E}}$ olyan diagonális mátrix, amelyben a főátlón csak 1, −1 és 0 szerepel (nem kell mindhárom számnak megjelennie). A koordinátás alak ebben az esetben a változók sorrendjének erejéig így fest (az $A = (a_{ij}) = [\![q]\!]_{\mathcal{E}}$ jelöléssel):

$$q(\mathbf{v}) = \mathbf{x}A\mathbf{x}^{\mathrm{T}} = \sum_{i,j} a_{ij}x_ix_j = x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_r^2 \qquad (\text{ahol } \mathbf{x} = [\![\mathbf{v}]\!]_{\mathcal{E}}).$$

6.24. Tétel. Legyen $q\colon V \to \mathbb{R}$ kvadratikus alak a valós számtest feletti V végesdimenziós vektortéren. Ekkor létezik V-nek olyan bázisa, amelyben q normálalakú.

Bizonyítás. Lásd [SzL] 16.2. Tétel. □

6.25. feladat. Hozzuk nemelfajuló lineáris helyettesítéssel normálalakra az alábbi q valós kvadratikus alakot:

$$q = x_1^2 + x_2^2 + 3x_3^2 + 4x_1x_2 + 2x_1x_3 + 2x_2x_3.$$

Megoldás. A 6.20. feladat megoldásában már kanonikus alakra hoztuk a kvadratikus alakot:

$$q = y_1^2 - 3y_2^2 + \frac{7}{3}y_3^2.$$

Vezessük be a $z_1 = y_1$, $z_2 = \sqrt{3}y_2$, $z_3 = \frac{\sqrt{7}}{\sqrt{3}}y_3$ új változókat; ez nyilván nemelfajuló lineáris helyettesítés. Ezzel q máris normálalakot ölt:

$$q = z_1^2 - z_2^2 + z_3^2.$$ ◇

6.26. Tétel Sylvester-féle tehetetlenségi tétel. Valós kvadratikus alak normálalakjában a pozitív, negatív és nulla tagok száma nem függ a bázis megválasztásától, azaz a normálalak a tagok sorrendjétől eltekintve egyértelműen meghatározott.

Bizonyítás. Lásd [SzL] 16.3. Tétel. □

6.27. Definíció. Legyen $q\colon V \to \mathbb{R}$ kvadratikus alak a valós számtest feletti V végesdimenziós vektortéren. Azt mondjuk, hogy

(i) q pozitív definit, ha $q(\mathbf{v}) \geq 0$ minden $\mathbf{v} \in V$ vektorra, és $q(\mathbf{v}) = 0 \iff \mathbf{v} = \mathbf{0}$;

(ii) q pozitív szemidefinit, ha $q(\mathbf{v}) \geq 0$ minden $\mathbf{v} \in V$ vektorra, és van olyan $\mathbf{w} \neq \mathbf{0}$ vektor, amelyre $q(\mathbf{w}) = 0$;

(iii) q negatív definit, ha $q(\mathbf{v}) \leq 0$ minden $\mathbf{v} \in V$ vektorra, és $q(\mathbf{v}) = 0 \iff \mathbf{v} = \mathbf{0}$;

(iv) q negatív szemidefinit, ha $q(\mathbf{v}) \leq 0$ minden $\mathbf{v} \in V$ vektorra, és van olyan $\mathbf{w} \neq \mathbf{0}$ vektor, amelyre $q(\mathbf{w}) = 0$;

(v) q indefinit, ha felvesz pozitív és negatív értéket is.

6.28. Tétel. Legyen $q\colon V \to \mathbb{R}$ kvadratikus alak a valós számtest feletti V végesdimenziós vektortéren, és legyen $\dim V = n$. Ekkor

(i) q akkor és csak akkor pozitív definit, ha normálalakja $x_1^2 + \cdots + x_n^2$;

(ii) q akkor és csak akkor pozitív szemidefinit, ha normálalakja $x_1^2 + \cdots + x_r^2$, ahol $r < n$;

(iii) q akkor és csak akkor negatív definit, ha normálalakja $-x_1^2 - \cdots - x_n^2$;

(iv) q akkor és csak akkor negatív szemidefinit, ha normálalakja $-x_1^2 - \cdots - x_r^2$, ahol $r < n$;

(v) q akkor és csak akkor indefinit, ha normálalakjában fellép pozitív és negatív előjelű tag is.

Bizonyítás. Triviális. □

6.29. Definíció. Azt mondjuk, hogy az $A \in \mathbb{R}^{n\times n}$ szimmetrikus mátrix pozitív definit, ha az $\mathbf{x}A\mathbf{x}^{\mathrm{T}}$ valós kvadratikus alak pozitív definit.

6.30. Tétel. Az $A \in \mathbb{R}^{n\times n}$ mátrix akkor és csak akkor pozitív definit, ha létezik olyan $P \in \mathbb{R}^{n\times n}$ nemelfajuló mátrix, amelyre $A = PP^{\mathrm{T}}$.

Bizonyítás. A feltétel elegendőségének bizonyításához tfh. $A = PP^{\mathrm{T}}$, ahol $P \in \mathbb{R}^{n\times n}$ nemelfajuló mátrix. Világos, hogy ekkor A szimmetrikus. Legyen $\mathbf{x} \in \mathbb{R}^n$ tetszőleges vektor és legyen $\mathbf{y} = \mathbf{x}P$. Ekkor $\mathbf{x}A\mathbf{x}^{\mathrm{T}} = \big(\mathbf{x}P\big)\big(P^{\mathrm{T}}\mathbf{x}^{\mathrm{T}}\big) = \mathbf{y}\mathbf{y}^{\mathrm{T}} = y_1^2 + \cdots + y_n^2$. Ez nyilván mindig nemnegatív, és nullát csak $\mathbf{y} = \mathbf{0}$ esetén kapunk, utóbbi pedig csak $\mathbf{x} = \mathbf{0}$ esetén fordulhat elő, hiszen P nemelfajuló. Tehát A valóban pozitív definit mátrix.
A szükségesség bizonyításához tfh. A pozitív definit, és tekintsük a $q = \mathbf{x}A\mathbf{x}^{\mathrm{T}}$ valós kvadratikus alakot; ekkor persze $A = [\![q]\!]_{\mathcal{E}}$. Mivel q pozitív definit, normálalakja $y_1^2 + \cdots + y_n^2$, azaz van $\mathbb{R}^n$-nek olyan $\mathcal{F}$ bázisa, amelyben $[\![q]\!]_{\mathcal{F}}$ az egységmátrix. Az $[\![\mathcal{F} \to \mathcal{E}]\!]$ bázisátmenet-mátrixot S-sel, ennek inverzét pedig P-vel jelölve $[\![q]\!]_{\mathcal{F}} = E = SAS^{\mathrm{T}}$, amiből $A = PP^{\mathrm{T}}$ következik. □

6.31. Definíció. Az $A \in T^{n\times n}$ mátrix k-adik főminorán a bal felső $k \times k$-as aldeterminánsát értjük:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{vmatrix}$$

6.32. Tétel. Egy $A \in T^{n\times n}$ szimmetrikus valós mátrix akkor és csak akkor pozitív definit, ha minden főminora pozitív.

6.4. Az euklideszi tér fogalma, norma és szög

6.33. Definíció. Euklideszi térnek nevezünk egy V valós vektorteret, ha értelmezett rajta egy $\ell\colon V\times V\to\mathbb{R}$ szimmetrikus bilineáris leképezés, amelyhez pozitív definit kvadratikus alak tartozik. Az ℓ bilineáris leképezést belső szorzatnak nevezzük, és a továbbiakban az $\langle\mathbf{u},\mathbf{v}\rangle=\ell(\mathbf{u},\mathbf{v})$ jelölést használjuk. A $\mathbf{v}\in V$ vektor normáján a $\|\mathbf{v}\|=\sqrt{\langle\mathbf{v},\mathbf{v}\rangle}$ nemnegatív valós számot értjük, az $\|\mathbf{u}-\mathbf{v}\|$ nemnegatív valós számot pedig az $\mathbf{u}$ és $\mathbf{v}$ vektorok távolságának nevezzük.

6.34. Példa. A $V=\mathbb{R}^n$ vektortér euklideszi tér az $\langle\mathbf{x},\mathbf{y}\rangle=x_1y_1+\dots+x_ny_n=\mathbf{x}\cdot\mathbf{y}^{\mathrm{T}}$ belső szorzattal. Az $\mathbf{x}\in\mathbb{R}^n$ vektor normája ebben a térben $\|\mathbf{x}\|=\sqrt{x_1^2+\dots+x_n^2}$. Általánosabban, tetszőleges $A\in\mathbb{R}^{n\times n}$ pozitív definit mátrix esetén $\mathbf{x}A\mathbf{y}^{\mathrm{T}}$ belső szorzatot definiál az $\mathbb{R}^n$ vektortéren, és minden belső szorzat felírható ilyen alakban.

6.35. Példa. Legyen ℓ^2 a négyzetesen összegezhető valós sorozatok vektortere: $(x_1,x_2,\dots)\in\ell^2\iff\sum x_n^2<\infty$. (A jelölésben szereplő ℓ betűnek semmi köze a bilineáris leképezésekre használt ℓ betűhöz.) Ekkor ℓ^2 euklideszi teret alkot az $\langle\mathbf{x},\mathbf{y}\rangle=\sum_{n=1}^{\infty}x_ny_n$ belső szorzattal.

6.36. Példa. Legyen L^2 a négyzetesen integrálható valós függvények vektortere: $f\in L^2\iff\int f^2(x)\,dx<\infty$. Ekkor L^2 euklideszi teret alkot az $\langle f,g\rangle=\int f(x)g(x)\,dx$ belső szorzattal.

6.37. Tétel (Cauchy–Schwarz–Bunyakovszkij-egyenlőtlenség). Tetszőleges V euklideszi tér és $\mathbf{u},\mathbf{v}\in V$ vektorok esetén

$$|\langle\mathbf{u},\mathbf{v}\rangle|\le\|\mathbf{u}\|\cdot\|\mathbf{v}\|.$$

Bizonyítás. Rögzített (de tetszőleges) $\mathbf{u}$ és $\mathbf{v}$ vektorokra számítsuk ki $\lambda\mathbf{u}+\mathbf{v}$ normájának négyzetét, mint a λ valós paraméter függvényét, felhasználva a belső szorzat bilinearitását és szimmetriáját:

$$\begin{aligned}\|\lambda\mathbf{u}+\mathbf{v}\|^2&=\langle\lambda\mathbf{u}+\mathbf{v},\lambda\mathbf{u}+\mathbf{v}\rangle=\langle\lambda\mathbf{u},\lambda\mathbf{u}\rangle+\langle\lambda\mathbf{u},\mathbf{v}\rangle+\langle\mathbf{v},\lambda\mathbf{u}\rangle+\langle\mathbf{v},\mathbf{v}\rangle\\&=\lambda^2\langle\mathbf{u},\mathbf{u}\rangle+2\lambda\langle\mathbf{u},\mathbf{v}\rangle+\langle\mathbf{v},\mathbf{v}\rangle\\&=\lambda^2\|\mathbf{u}\|^2+2\lambda\langle\mathbf{u},\mathbf{v}\rangle+\|\mathbf{v}\|^2.\end{aligned}$$

Mivel $\|\lambda\mathbf{u}+\mathbf{v}\|^2$ mindig nemnegatív, a kapott másodfokú függvénynek legfeljebb egy valós zérushelye van, így diszkriminánsa nem pozitív:

$$\big(2\langle\mathbf{u},\mathbf{v}\rangle\big)^2-4\cdot\|\mathbf{u}\|^2\cdot\|\mathbf{v}\|^2\le 0.$$

Ebből már egyszerű átrendezéssel megkapható a bizonyítandó egyenlőtlenség. □

6.38. Következmény (háromszög-egyenlőtlenség). Tetszőleges V euklideszi tér és $\mathbf{u},\mathbf{v}\in V$ vektorok esetén

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

Bizonyítás. A belső szorzat bilinearitását és szimmetriáját, valamint a norma(négyzet) definícióját használva könnyen levezethető a háromszög-egyenlőtlenség a CSÉB-egyenlőtlenségből:

$$\begin{aligned}\|\mathbf{u}+\mathbf{v}\|^2=\langle\mathbf{u}+\mathbf{v},\mathbf{u}+\mathbf{v}\rangle=\langle\mathbf{u},\mathbf{u}\rangle+2\langle\mathbf{u},\mathbf{v}\rangle+\langle\mathbf{v},\mathbf{v}\rangle=\\=\|\mathbf{u}\|^2+2\langle\mathbf{u},\mathbf{v}\rangle+\|\mathbf{v}\|^2\le\|\mathbf{u}\|^2+2\cdot\|\mathbf{u}\|\cdot\|\mathbf{v}\|+\|\mathbf{v}\|^2=\big(\|\mathbf{u}\|+\|\mathbf{v}\|\big)^2.\end{aligned}$$ □

6.39. Definíció. Legyenek $\mathbf{u}$ és $\mathbf{v}$ nemnulla vektorok egy uklideszi térben. A CSÉB-egyenlőtlenség szerint létezik egy egyértelműen meghatározott $\alpha\in[0,\pi]$ szög, amelyre $\langle\mathbf{u},\mathbf{v}\rangle=\|\mathbf{u}\|\cdot\|\mathbf{v}\|\cdot\cos\alpha$. Ezt az α szöget az $\mathbf{u}$ és $\mathbf{v}$ vektorok által bezárt szögnek nevezzük.

6.40. feladat. Határozzuk meg az $\mathbb{R}^4$ euklideszi térben az $\mathbf{u}=(1,1,1,1)$ és $\mathbf{v}=(3,-1,3,-1)$ vektorok hosszát és az általuk bezárt szöget.

Megoldás.

$$\|\mathbf{u}\|=\sqrt{4}=2,\qquad\|\mathbf{v}\|=\sqrt{20}=2\sqrt{5},\qquad\langle\mathbf{u},\mathbf{v}\rangle=4,\qquad\measuredangle(\mathbf{u},\mathbf{v})=\arccos\frac{\langle\mathbf{u},\mathbf{v}\rangle}{\|\mathbf{u}\|\cdot\|\mathbf{v}\|}=\arccos\frac{1}{\sqrt{5}}\approx 63^\circ$$ ⋄

6.5. Ortogonalitás

6.41. Definíció. Legyen V egy euklideszi tér.

(i) Azt mondjuk, hogy $\mathbf{u},\mathbf{v}\in V$ ortogonálisak, ha derékszöget zárnak be, azaz $\langle\mathbf{u},\mathbf{v}\rangle=0$. Jelölés: $\mathbf{u}\perp\mathbf{v}$.

(ii) A $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ vektorrendszert ortogonális vektorrendszernek nevezzük, ha a vektorok páronként merőlegesek, azaz $\mathbf{v}_i \perp \mathbf{v}_j$ minden $i \neq j$ esetén.

(iii) Ha egy ortogonális vektorrendszer minden elemének 1 a normája, akkor ortonormált vektorrendszernek nevezzük. A Kronecker-delta szimbólum segítségével ezt így írhatjuk le tömören: $\mathbf{v}_1, \ldots, \mathbf{v}_k$ akkor és csak akkor ortonormált vektorrendszer, ha $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{ij}$ minden $i \neq j$ esetén.

6.42. Tétel. Minden nemnulla vektorokból álló ortogonális vektorrendszer lineárisan független.

Bizonyítás. Legyen $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ nemnulla vektorokból álló ortogonális vektorrendszer egy V euklideszi térben. Tfh. $\mathbf{0} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k$ valamely $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ együtthatókra. Számítsuk ki mindkét oldal belső szorzatát a $\mathbf{v}_j$ vektorral, a jobb oldalon a belső szorzat első változó szerinti linearitását használva:

$$0 = \langle \mathbf{0}, \mathbf{v}_j \rangle = \left\langle \sum_{i=1}^{k} \lambda_i \mathbf{v}_i,\, \mathbf{v}_j \right\rangle = \sum_{i=1}^{k} \langle \lambda_i \mathbf{v}_i, \mathbf{v}_j \rangle = \sum_{i=1}^{k} \lambda_i \langle \mathbf{v}_i, \mathbf{v}_j \rangle.$$

Az ortogonalitási feltevés miatt $i \neq j$ esetén $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$, ezért az összegből csak a j-edik tag marad meg: $0 = \langle \mathbf{0}, \mathbf{v}_j \rangle = \lambda_j \langle \mathbf{v}_j, \mathbf{v}_j \rangle = \lambda_j \|\mathbf{v}_j\|^2$. Mivel $\mathbf{v}_j \neq \mathbf{0}$, a normája sem nulla, tehát szükségképpen $\lambda_j = 0$. Ez minden $j \in \{1, \ldots, k\}$ indexre érvényes, tehát a vektorrendszerünk valóban lineárisan független. □

6.43. Definíció. Az $\mathbf{e}_1, \ldots, \mathbf{e}_n$ vektorrendszert ortonormált bázisnak nevezzük a V euklideszi térben, ha ortonormált és bázisa V-nek.

6.44. Tétel (ONB-ben szép az élet). Legyen V végesdimenziós euklideszi tér, és legyen $\mathcal{E} \colon \mathbf{e}_1, \ldots, \mathbf{e}_n$ ortonormált bázisna V-nek. Ha $[\![\mathbf{u}]\!]_{\mathcal{E}} = (x_1, \ldots, x_n)$ és $[\![\mathbf{v}]\!]_{\mathcal{E}} = (y_1, \ldots, y_n)$, akkor

(i) $x_j = \langle \mathbf{u}, \mathbf{e}_j \rangle \quad (j = 1, \ldots, n)$;

(ii) $\langle \mathbf{u}, \mathbf{v} \rangle = x_1 y_1 + \cdots + x_n y_n = [\![\mathbf{u}]\!]_{\mathcal{E}} \cdot [\![\mathbf{v}]\!]_{\mathcal{E}}^{\mathrm{T}}$;

(iii) $\|\mathbf{u}\| = \sqrt{x_1^2 + \cdots + x_n^2}$.

Bizonyítás. A koordinátasor definíciója szerint $\mathbf{u} = \sum_{i=1}^{n} x_i \mathbf{e}_i$.

(i) Számítsuk ki az $\langle \mathbf{u}, \mathbf{e}_j \rangle$ belső szorzatot a belső szorzat első változó szerinti linearitását és a bázis ortonormáltságát használva:

$$\langle \mathbf{u}, \mathbf{e}_j \rangle = \left\langle \sum_{i=1}^{n} x_i \mathbf{e}_i,\, \mathbf{e}_j \right\rangle = \sum_{i=1}^{k} \langle x_i \mathbf{e}_i, \mathbf{e}_j \rangle = \sum_{i=1}^{k} x_i \langle \mathbf{e}_i, \mathbf{e}_j \rangle = \sum_{i=1}^{k} x_i \delta_{ij} = x_j.$$

(ii) A fentihez hasonlóan számolunk, de most két „szummát” is fel kell bontanunk a belső szorzat bilinearitását használva:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \left\langle \sum_{i=1}^{n} x_i \mathbf{e}_i,\, \sum_{j=1}^{n} y_j \mathbf{e}_j \right\rangle = \sum_{i,j=1}^{n} \langle x_i \mathbf{e}_i,\, y_j \mathbf{e}_j \rangle = \sum_{i,j=1}^{n} x_i y_j \langle \mathbf{e}_i, \mathbf{e}_j \rangle = \sum_{i,j=1}^{n} x_i y_j \delta_{ij} = \sum_{i=1}^{n} x_i y_i.$$

(iii) Ez tüstént következik a második állításból.

□

6.6. Ortogonális vetület, ortogonalizáció

6.45. Lemma. Legyen V euklideszi tér és $\mathbf{u}, \mathbf{v}_1, \ldots, \mathbf{v}_k \in V$. Ha $\mathbf{u}$ ortogonális a $\mathbf{v}_i$ vektorok mindegyikére, akkor ortogonális ezek minden lineáris kombinációjára, azaz az $S = [\mathbf{v}_1, \ldots, \mathbf{v}_k]$ altér minden elemére.

Bizonyítás. Tfh. $\langle \mathbf{u}, \mathbf{v}_i \rangle = 0$ minden i-re, és legyen $\mathbf{v} = \sum \lambda_i \mathbf{v}_i$ tetszőleges eleme az S altérnek . Számítsuk ki a $\langle \mathbf{u}, \mathbf{v} \rangle$ belső szorzatot, kihasználva a második változó szerinti linearitást:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \left\langle \mathbf{u},\, \sum_{i=1}^{k} \lambda_i \mathbf{v}_i \right\rangle = \sum_{i=1}^{k} \langle \mathbf{u},\, \lambda_i \mathbf{v}_i \rangle = \sum_{i=1}^{k} \lambda_i \langle \mathbf{u}, \mathbf{v}_i \rangle = \sum_{i=1}^{k} \lambda_i \cdot 0 = 0.$$

Tehát $\mathbf{u}$ valóban ortogonális minden $\mathbf{v} \in S$ vektorra. □

6.46. feladat. Bontsuk az $\mathbf{u}$ vektort egy $\mathbf{v}$-vel párhuzamos $\mathbf{u}_{\|}$ és egy $\mathbf{v}$-re merőleges $\mathbf{u}_{\perp}$ vektor összegére, ahol

$$\mathbf{u} = (3, 5, 5), \qquad \mathbf{v} = (1, 2, 3).$$

Megoldás. Az $\mathbf{u}_{\|}$ vektort $\mathbf{u}_{\|} = \lambda\mathbf{v}$ alakban keressük, és a λ együtthatót úgy kell megválasztanunk, hogy az $\mathbf{u}_{\perp} := \mathbf{u} - \mathbf{u}_{\|}$ vektor merőleges legyen $\mathbf{v}$-re:

$$\langle \mathbf{u}_{\perp}, \mathbf{v} \rangle = \langle \mathbf{u} - \mathbf{u}_{\|}, \mathbf{v} \rangle = \langle \mathbf{u} - \lambda\mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle - \lambda \langle \mathbf{v}, \mathbf{v} \rangle = 0.$$

Innen ki tudjuk fejezni a λ együtthatót, és ezzel megkapjuk az $\mathbf{u}_{\|}$ és $\mathbf{u}_{\perp}$ vektorokat:

$$\lambda = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} = \frac{28}{14} = 2, \quad \mathbf{u}_{\|} = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \cdot \mathbf{v} = 2\mathbf{v} = (2,4,6), \quad \mathbf{u}_{\perp} = \mathbf{u} - \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \cdot \mathbf{v} = \mathbf{u} - 2\mathbf{v} = (1,1,-1).$$

Interaktív ábra: http://www.math.u-szeged.hu/~twaldha/tanitas/linalg2bsc_2023osz/gs-sage.html. ◇

6.47. Lemma. Legyen V euklideszi tér, legyen $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ nemnulla vektorokból álló ortogonális vektorrendszer, és legyen $S = [\mathbf{v}_1, \ldots, \mathbf{v}_k] \leq V$. Tetszőleges $\mathbf{u} \in V$ vektor egyértelműen felbontható $\mathbf{u} = \mathbf{u}_{\|} + \mathbf{u}_{\perp}$ alakban, ahol $\mathbf{u}_{\|} \in S$ és $\mathbf{u}_{\perp}$ ortogonális az S altérre (azaz ortogonális S minden elemére):

$$\mathbf{u}_{\|} = \sum_{i=1}^{k} \frac{\langle \mathbf{u}, \mathbf{v}_i \rangle}{\langle \mathbf{v}_i, \mathbf{v}_i \rangle} \mathbf{v}_i, \qquad \mathbf{u}_{\perp} = \mathbf{u} - \sum_{i=1}^{k} \frac{\langle \mathbf{u}, \mathbf{v}_i \rangle}{\langle \mathbf{v}_i, \mathbf{v}_i \rangle} \mathbf{v}_i.$$

Bizonyítás. A 6.42. Tétel szerint $\mathbf{v}_1, \ldots, \mathbf{v}_k$ bázisa az S altérnek, ezért minden eleme egyértelműen felírható $\sum \lambda_i \mathbf{v}_i$ alakban. Keressük tehát az $\mathbf{u}_{\|}$ vektort is $\mathbf{u}_{\|} = \sum \lambda_i \mathbf{v}_i$ alakban, és vizsgáljuk meg, hogy mikor lesz $\mathbf{u} - \mathbf{u}_{\|}$ ortogonális a $\mathbf{v}_j$ vektorra:

$$(\mathbf{u} - \mathbf{u}_{\|}) \perp \mathbf{v}_j \iff \langle \mathbf{u} - \mathbf{u}_{\|}, \mathbf{v}_j \rangle = 0 \iff \langle \mathbf{u}, \mathbf{v}_j \rangle - \langle \mathbf{u}_{\|}, \mathbf{v}_j \rangle = 0 \iff \langle \mathbf{u}, \mathbf{v}_j \rangle = \langle \mathbf{u}_{\|}, \mathbf{v}_j \rangle.$$

Számítsuk ki az $\langle \mathbf{u}_{\|}, \mathbf{v}_j \rangle$ belső szorzatot a belső szorzat első változó szerinti linearitását használva:

$$\langle \mathbf{u}_{\|}, \mathbf{v}_j \rangle = \left\langle \sum_{i=1}^{k} \lambda_i \mathbf{v}_i,\ \mathbf{v}_j \right\rangle = \sum_{i=1}^{k} \langle \lambda_i \mathbf{v}_i, \mathbf{v}_j \rangle = \sum_{i=1}^{k} \lambda_i \langle \mathbf{v}_i, \mathbf{v}_j \rangle.$$

Az ortogonalitási feltevés miatt $i \neq j$ esetén $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$, ezért az összegből csak a j-edik tag marad meg: $\langle \mathbf{u}_{\|}, \mathbf{v}_j \rangle = \lambda_j \langle \mathbf{v}_j, \mathbf{v}_j \rangle$. Azt kaptuk tehát, hogy $\mathbf{u}_{\perp} := \mathbf{u} - \mathbf{u}_{\|}$ akkor és csak akkor lesz ortogonális mindegyik $\mathbf{v}_j$ vektorra, ha $\langle \mathbf{u}, \mathbf{v}_j \rangle = \lambda_j \langle \mathbf{v}_j, \mathbf{v}_j \rangle$, azaz $\lambda_j = \frac{\langle \mathbf{u}, \mathbf{v}_j \rangle}{\langle \mathbf{v}_j, \mathbf{v}_j \rangle}$. Ezekkel az együtthatókkal $\mathbf{u}_{\perp}$ merőleges lesz S minden elemére (lásd a 6.45. Lemmát). □

6.48. Megjegyzés. Meg lehet mutatni, hogy $\mathbf{u}_{\|}$ az S altérnek az $\mathbf{u}$ vektorhoz legközelebbi pontja, így $\|\mathbf{u}_{\perp}\|$ nem más, mint az $\mathbf{u}$ vektornak az S altértől való távolsága (HF).

6.49. feladat. Bontsuk az $\mathbf{u}$ vektort egy S-sel párhuzamos $\mathbf{u}_{\|}$ és egy S-re merőleges $\mathbf{u}_{\perp}$ vektor összegére, ahol

$$\mathbf{u} = (9,1,1), \qquad S = \big[\, (1,2,3),\ (3,5,5) \,\big].$$

Megoldás. A 6.46. feladatban már meghatároztunk egy $\mathbf{v}_1$, $\mathbf{v}_2$ ortogonális bázist az S altérben: $\mathbf{v}_1 = (1,2,3)$, $\mathbf{v}_2 = (1,1,-1)$. A keresett $\mathbf{u}_{\|}$ vektor nem más lesz, mint az $\mathbf{u}$ vektor $\mathbf{v}_1$-gyel párhuzamos komponensének és $\mathbf{v}_2$-vel párhuzamos komponensének összege, $\mathbf{u}_{\perp}$ pedig a „maradék” lesz:

$$\begin{aligned} \mathbf{u}_{\|} &= \frac{\langle \mathbf{u}, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \cdot \mathbf{v}_1 + \frac{\langle \mathbf{u}, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} \cdot \mathbf{v}_2 = \mathbf{v}_1 + 3\mathbf{v}_2 = (4,5,0), \\ \mathbf{u}_{\perp} &= \mathbf{u} - \frac{\langle \mathbf{u}, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \cdot \mathbf{v}_1 - \frac{\langle \mathbf{u}, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} \cdot \mathbf{v}_2 = \mathbf{u} - \mathbf{v}_1 - 3\mathbf{v}_2 = (5,-4,1). \end{aligned}$$

Interaktív ábra: http://www.math.u-szeged.hu/~twaldha/tanitas/linalg2bsc_2023osz/gs-sage.html. ◇

6.50. Tétel (Gram–Schmidt-ortogonalizáció). Legyenek $\mathbf{u}_1, \ldots, \mathbf{u}_k$ lineárisan független vektorrendszer a V euklideszi térben. Ekkor létezik olyan $\mathbf{v}_1, \ldots, \mathbf{v}_k$ ortogonális vektorrendszer, amelyre $[\mathbf{u}_1, \ldots, \mathbf{u}_j] = [\mathbf{v}_1, \ldots, \mathbf{v}_j]$ teljesül minden $j \in \{1, \ldots, k\}$ indexre.

Bizonyítás. A vektorok száma szerinti teljes indukcióval bizonyítunk. A $k = 1$ esetben nyilván $\mathbf{v}_1 = \mathbf{u}_1$ megfelelő lesz. Az indukciós lépéshez tfh. $(k-1)$-elemű vektorrendszerekre igaz az állítás (indukciós hipotézis), és tekintsünk egy k-elemű $\mathbf{u}_1, \ldots, \mathbf{u}_k$ lineárisan független vektorrendszert. Az indukciós hipotézist az $\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}$ lineárisan független vektorrendszerre alkalmazva kapjuk, hogy léteznek olyan páronként ortogonális $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1} \in V$ vektorok, amelyre $[\mathbf{u}_1, \ldots, \mathbf{u}_j] = [\mathbf{v}_1, \ldots, \mathbf{v}_j]$ teljesül minden $j \in \{1, \ldots, k-1\}$ indexre.

Vegyük észre, hogy a $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$ vektorok egyike sem lehet nullvektor, mert $\dim[\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}] = \dim[\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}] = k-1$ az $\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}$ vektorok lineáris függetlensége miatt. Legyen $\mathbf{v}_k$ az $\mathbf{u}_k$ vektor $[\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}]$ altérre merőleges komponense (lásd a 6.47. Lemmát):

$$\mathbf{v}_k = \mathbf{u}_k - \frac{\langle \mathbf{u}_k, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \cdot \mathbf{v}_1 - \ldots - \frac{\langle \mathbf{u}_k, \mathbf{v}_{k-1} \rangle}{\langle \mathbf{v}_{k-1}, \mathbf{v}_{k-1} \rangle} \cdot \mathbf{v}_{k-1}. \tag{4}$$

Ekkor $\mathbf{v}_k$ ortogonális a $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$ vektorok mindegyikére, ezért $\mathbf{v}_1, \ldots, \mathbf{v}_k$ ortogonális vektorrendszer. Meg kell még mutatnunk, hogy $[\mathbf{u}_1, \ldots, \mathbf{u}_k] = [\mathbf{v}_1, \ldots, \mathbf{v}_k]$. Az indukciós hipotézisből tudjuk, hogy $[\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}] = [\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}]$, következésképp $[\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}, \mathbf{u}_k] = [\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}, \mathbf{u}_k]$. A (4) összefüggés szerint $\mathbf{v}_k$ előáll az $\mathbf{u}_k, \mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$ vektorok lineáris kombinációjaként, és átrendezve az egyenlőséget $\mathbf{u}_k$ kifejezhető a $\mathbf{v}_k, \mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$ vektorok lineáris kombinációjaként, tehát $[\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}, \mathbf{u}_k] = [\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}, \mathbf{v}_k]$, és ezzel kész a bizonyítás. □

6.51. feladat. Hajtsuk végre a Gram–Schmidt-féle ortogonalizációs eljárást az alábbi $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3 \in \mathbb{R}^3$ vektorokra:

$$\mathbf{u}_1 = (1,2,3), \qquad \mathbf{u}_2 = (3,5,5), \qquad \mathbf{u}_3 = (9,1,1).$$

Megoldás. A 6.46. és 6.49 feladatok lényegében tartalmazzák a megoldást, de azért foglaljuk össze a lépéseket a 6.50. Tétel jelöléseit használva:

$$\mathbf{v}_1 = \mathbf{u}_1 = (1,2,3);$$

$$\mathbf{v}_2 = \mathbf{u}_2 - \frac{\langle \mathbf{u}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \cdot \mathbf{v}_1 = \mathbf{u}_2 - 2\mathbf{v}_1 = (1,1,-1);$$

$$\mathbf{v}_3 = \mathbf{u}_3 - \frac{\langle \mathbf{u}_3, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \cdot \mathbf{v}_1 - \frac{\langle \mathbf{u}_3, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} \cdot \mathbf{v}_2 = \mathbf{u}_3 - \mathbf{v}_1 - 3\mathbf{v}_2 = (5,-4,1).$$ ◇

Interaktív ábra: `http://www.math.u-szeged.hu/~twaldha/tanitas/linalg2bsc_2023osz/gs-sage.html`.

6.52. feladat. Legyen $S \leq \mathbb{R}^4$ az alábbi $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ vektorok által kifeszített altér:

$$\mathbf{u}_1 = (1,1,1,1), \quad \mathbf{u}_2 = (3,-1,3,-1), \quad \mathbf{u}_3 = (6,2,2,-2).$$

Adjunk meg egy ortonormált bázist az S altérben.

Megoldás. Hajtsuk végre a Gram–Schmidt-féle ortogonalizációs eljárást:

1. $\mathbf{v}_1 = \mathbf{u}_1 = (1,1,1,1)$;
2. $\mathbf{v}_2 = \mathbf{u}_2 - \frac{\langle \mathbf{u}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \cdot \mathbf{v}_1 = \mathbf{u}_2 - \frac{4}{4} \cdot \mathbf{v}_1 = \mathbf{u}_2 - \mathbf{v}_1 = (2,-2,2,-2)$;
3. $\mathbf{v}_3 = \mathbf{u}_3 - \frac{\langle \mathbf{u}_3, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \cdot \mathbf{v}_1 - \frac{\langle \mathbf{u}_3, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} \cdot \mathbf{v}_2 = \mathbf{u}_3 - \frac{8}{4} \cdot \mathbf{v}_1 - \frac{16}{16} \cdot \mathbf{v}_2 = \mathbf{u}_3 - 2\mathbf{v}_1 - \mathbf{v}_2 = (2,2,-2,-2)$.

A kapott $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ vektorrendszer ortogonális és kifeszíti az S alteret. Tehát $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ ortogonális bázisa S-nek. Hogy ortonormált bázist kapjunk, „normáljuk le” a vektorokat:

1. $\mathbf{e}_1 = \frac{1}{\|\mathbf{v}_1\|} \cdot \mathbf{v}_1 = \frac{1}{2} \cdot \mathbf{v}_1 = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$;
2. $\mathbf{e}_2 = \frac{1}{\|\mathbf{v}_2\|} \cdot \mathbf{v}_2 = \frac{1}{4} \cdot \mathbf{v}_2 = \left(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\right)$;
3. $\mathbf{e}_3 = \frac{1}{\|\mathbf{v}_3\|} \cdot \mathbf{v}_3 = \frac{1}{4} \cdot \mathbf{v}_3 = \left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right)$.

Az $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ vektorrendszer ortonormált bázisa S-nek (az ortonormáltság fejben is könnyen ellenőrizhető). ◇

6.53. Megjegyzés. A fenti feladat megoldása mátrixokkal is felírható. Legyen $A \in \mathbb{R}^{4\times 3}$ az $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ vektorok oszloponkénti egymás mellé írásával keletkező mátrix, és hasonlóan legyen $Q_1 \in \mathbb{R}^{4\times 3}$ az $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ vektorok oszloponkénti egymás mellé írásával keletkező mátrix:

$$A = \begin{pmatrix} 1 & 3 & 6 \\ 1 & -1 & 2 \\ 1 & 3 & 2 \\ 1 & -1 & 2 \end{pmatrix}, \qquad Q_1 = \begin{pmatrix} 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 \\ 1/2 & -1/2 & -1/2 \end{pmatrix}.$$

A feladat megoldását végigkövetve mindegyik $\mathbf{u}_i$ vektor kifejezhető az $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ vektorok lineáris kombinációjaként:

1. $\mathbf{u}_1 = 2\mathbf{e}_1$;
2. $\mathbf{u}_2 = 2\mathbf{e}_1 + 4\mathbf{e}_2$;
3. $\mathbf{u}_3 = 4\mathbf{e}_1 + 4\mathbf{e}_2 + 4\mathbf{e}_3$.

Ezekből az együtthatókból készítjük az $R_1 \in \mathbb{R}^{3\times 3}$ mátrixot:

$$R_1 = \begin{pmatrix} 2 & 2 & 4 \\ 0 & 4 & 4 \\ 0 & 0 & 4 \end{pmatrix}.$$

Az $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ vektorok fenti kikombinálása az $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ vektorokból az $A = Q_1 R_1$ mátrixszorzással írható le:

$$A = \begin{pmatrix} 1 & 3 & 6 \\ 1 & -1 & 2 \\ 1 & 3 & 2 \\ 1 & -1 & 2 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 \\ 1/2 & -1/2 & -1/2 \end{pmatrix} \cdot \begin{pmatrix} 2 & 2 & 4 \\ 0 & 4 & 4 \\ 0 & 0 & 4 \end{pmatrix} = Q_1 R_1.$$

Itt Q_1 úgynevezett szemiortogonális mátrix: oszlopai ortonormált vektorrendszert alkotnak. Az R_1 mátrix pedig felső trianguláris mátrix, amelynek főátlóján pozitív számok állnak (ezek nem mások, mint a $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ vektorok normái). Az $A = Q_1 R_1$ felbontást az A mátrix (redukált) QR-felbontásának nevezzük. Az $\mathbf{e}_4 = (1/2, -1/2, -1/2, 1/2)$ vektorral $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4$ ortonormált bázisa $\mathbb{R}^4$-nek, ezért a belőlük, mint oszlopvektorokból alkotott Q mátrix egy 4×4-es ortogonális mátrix. Egészítsük ki az R mátrixot egy csupa nulla sorral, így az A mátrix alábbi felbontását kapjuk:

$$A = \begin{pmatrix} 1 & 3 & 6 \\ 1 & -1 & 2 \\ 1 & 3 & 2 \\ 1 & -1 & 2 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \end{pmatrix} \cdot \begin{pmatrix} 2 & 2 & 4 \\ 0 & 4 & 4 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix} = QR.$$

Figyeljük meg, hogy az R mátrix legalsó, csupa nulla sorának köszönhetően a Q mátrix utolsó oszlopa (az $\mathbf{e}_4$ vektor) semmilyen szerepet nem játszik a mátrixszorzatban. Az $A = QR$ felbontást az A mátrix (teljes) QR-felbontásának nevezzük. Ezzel a módszerrel (vagyis lényegében a Gram–Schmidt-ortogonalizációval) bármely teljes oszloprangú mátrixnak megkapható a redukált/teljes QR-felbontása.

6.7. Ortonormált bázis létezése, ortogonális komplementum

6.54. Tétel. Végesdimenziós euklideszi térben minden ortonormált vektorrendszer kibővíthető ortonormált bázissá.

Bizonyítás. Legyen V egy n-dimenziós euklideszi tér, és legyen $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ ortonormált vektorrendszer. Ekkor $\mathbf{v}_1, \ldots, \mathbf{v}_k$ lineárisan független vektorrendszer (6.42. Tétel), ezért kibővíthető bázissá (2.41. Tétel). Léteznek tehát olyan $\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n \in V$ vektorok, amelyekkel $\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n$ bázisa V-nek. Ha erre a vektorrendszerre végrehajtjuk a Gram–Schmidt-ortogonalizációt, akkor az első k vektor nem változik, mert azok már ortogonálisak voltak. Tehát az eljárás végeredménye egy $\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n$ ortogonális bázis, amelyből normálással ortonormált bázist kapunk. □

6.55. Következmény. Minden végesdimenziós euklideszi térben létezik ortonormált bázis.

Bizonyítás. Az üres halmaz ortonormált vektorrendszer, és ez a fenti tétel szerint ortonormált bázissá bővíthető. □

6.56. Definíció. Legyenek V és W euklideszi terek. Azt mondjuk, hogy a $\varphi\colon V \to W$ leképezés izomorfizmus, ha egyrészt vektortér-izomorfizmus (azaz bijektív lineáris leképezés), másrészt megőrzi a belső szorzatot:

$$\forall \mathbf{u}, \mathbf{v} \in V \colon \langle \mathbf{u}, \mathbf{v} \rangle_V = \langle \mathbf{u}\varphi, \mathbf{v}\varphi \rangle_W.$$

6.57. Tétel. Minden n-dimenziós euklideszi tér izomorf az $\mathbb{R}^n$ euklideszi térrel.

Bizonyítás. Legyen $\mathcal{E}\colon \mathbf{e}_1, \ldots, \mathbf{e}_n$ ortonormált bázis a V euklideszi térben, és tekintsük az alábbi ψ leképezést:

$$\psi\colon V \to \mathbb{R}^n, \mathbf{v} \mapsto [\![\mathbf{v}]\!]_{\mathcal{E}}.$$

Tudjuk, hogy ψ vektortér-izomorfizmus (lásd a 2.54. Tételt), a belső szorzat megőrzése pedig következik a 6.44. Tételből):

$$\forall \mathbf{u}, \mathbf{v} \in V\colon \langle \mathbf{u}, \mathbf{v} \rangle_V = [\![\mathbf{u}]\!]_{\mathcal{E}} \cdot [\![\mathbf{v}]\!]_{\mathcal{E}}^{\mathrm{T}} = \langle \mathbf{u}\psi, \mathbf{v}\psi \rangle_{\mathbb{R}^n}. \qquad \square$$

6.58. Definíció. Legyen U altér a V euklideszi térben. Az U altér ortogonális komplementumán az U-ra merőleges vektorok alkotta $U^\perp \leq V$ alteret értjük:

$$U^\perp = \{\mathbf{v} \in V : \mathbf{u} \perp \mathbf{v} \text{ minden } \mathbf{u} \in U \text{ esetén}\}.$$

6.59. Tétel. Ha U altér a V végesdimenziós euklideszi térben, akkor

(i) $U^\perp$ valóban altér;

(ii) $V = U \oplus U^\perp$;

(iii) $(U^\perp)^\perp = U$.

Bizonyítás.

(i) Következik a 6.45. Lemmából ($U^\perp$ még akkor is altér lenne, ha U nem lenne altér, hanem egy akármilyen részhalmaza V-nek).

(ii) Tetszőleges V-beli vektor felbontható egy U-beli és egy U-ra merőleges vektor összegére (lásd a 6.47. Lemmát), tehát $V = U + U^\perp$. Az összeg „direktségéhez” azt kell igazolnunk, hogy $U \cap U^\perp = \{\mathbf{0}\}$. Valóban, ha $\mathbf{u} \in U \cap U^\perp$, akkor $\mathbf{u}$ merőleges saját magára, így $\|\mathbf{u}\|^2 = \langle \mathbf{u}, \mathbf{u} \rangle = \mathbf{0}$, tehát $\mathbf{u} = \mathbf{0}$.

(iii) Az világos, hogy $(U^\perp)^\perp \supseteq U$. Az előző pontbeli direktösszeg-felbontást az $U^\perp$ altérre is felírhatjuk: $V = U^\perp \oplus (U^\perp)^\perp$. A 2.59. Tételt alkalmazva mindkét felbontásra azt kapjuk, hogy $\dim V = \dim U + \dim U^\perp = \dim U^\perp + \dim (U^\perp)^\perp$. Ebből következik, hogy $\dim U = \dim (U^\perp)^\perp$, ez pedig az $(U^\perp)^\perp \supseteq U$ tartalmazással együtt már igazolja, hogy $(U^\perp)^\perp = U$ (lásd a 2.57. Tételt). $\square$

6.60. Megjegyzés. Amikor egy $A\mathbf{x} = \mathbf{0}$ homogén lineáris egyenletrendszert oldunk meg, akkor valójában az A mátrix sorai által kifeszített altér ortogonális komplementumát határozzuk meg. Hasonlítsuk össze a fenti tételt a 3.18. Tétellel!

6.61. Megjegyzés. Az $U = (U^\perp)^\perp$ összefüggés így is értelmezhető: minden U altérhez van olyan W altér, melyre $U = W^\perp$ (mégpedig $W = U^\perp$). Ha $\mathbf{w}_1, \dots, \mathbf{w}_t$ bázisa W-nek, akkor a 6.45. Lemma szerint $U = W^\perp = [w_1]^\perp \cap \cdots \cap [w_t]^\perp$. Itt mindegyik $[\mathbf{w}_i]^\perp$ úgynevezett 1-kodimenziós altér (azaz $\dim[\mathbf{w}_i] = \dim V - 1$); az ilyen alteret szokás hipersíknak nevezni. Azt kaptuk tehát, hogy végesdimenziós euklideszi térben minden altér előáll véges sok hipersík metszeteként.

6.62. Megjegyzés. Az ortogonális komplementum megkapható a következő módon: megadunk egy bázist U-ban, ezt kibővítjük az egész tér bázisává, majd lefuttatjuk a Gram–Schmidt-ortogonalizációt. Így olyan ortonormált bázist kapunk V-ben, amelynek első néhány (pontosabban: $\dim U$) vektora ortonormált bázisa U-nak, a többi vektor pedig $U^\perp$ ortonormált bázisát adja.

6.8. Lineáris transzformáció adjungáltja

6.63. Tétel. Tetszőleges V végesdimenziós euklideszi tér és $\varphi, \psi \in \mathrm{Hom}(V, V)$ lineáris transzformációk esetén ekvivalensek az alábbiak:

(i) $\forall \mathbf{u}, \mathbf{v} \in V\colon \langle \mathbf{u}\varphi, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v}\psi \rangle$;

(ii) bármely $\mathcal{E}$ ortonormált bázisra $[\![\psi]\!]_{\mathcal{E}} = [\![\varphi]\!]_{\mathcal{E}}^{\mathrm{T}}$;

(iii) van olyan $\mathcal{E}$ ortonormált bázis, amelyre $[\![\psi]\!]_{\mathcal{E}} = [\![\varphi]\!]_{\mathcal{E}}^{\mathrm{T}}$.

Bizonyítás. Vezessünk be néhány jelölést: legyen $\mathcal{E}$ tetszőleges ortonormált bázisa V-nek, legyen $A = [\![\varphi]\!]_{\mathcal{E}}$ és $B = [\![\psi]\!]_{\mathcal{E}}$. Legyen továbbá $\mathbf{x} = [\![\mathbf{u}]\!]_{\mathcal{E}}$ és $\mathbf{y} = [\![\mathbf{v}]\!]_{\mathcal{E}}$, ahol $\mathbf{u}, \mathbf{v} \in V$ tetszőleges vektorok. A 3.24. Tétel szerint ekkor $[\![\mathbf{u}\varphi]\!]_{\mathcal{E}} = \mathbf{x}A$ és $[\![\mathbf{v}\psi]\!]_{\mathcal{E}} = \mathbf{y}B$.

- (i) $\Longrightarrow$ (ii): A 6.44. Tétel szerint $\langle \mathbf{u}\varphi, \mathbf{v} \rangle = (\mathbf{x}A)\mathbf{y}^{\mathrm{T}}$ és $\langle \mathbf{u}, \mathbf{v}\psi \rangle = \mathbf{x}(\mathbf{y}B)^{\mathrm{T}}$. Ezekből az (i) feltétel alapján következik, hogy $\mathbf{x}A\mathbf{y}^{\mathrm{T}} = \mathbf{x}B^{\mathrm{T}}\mathbf{y}^{\mathrm{T}}$. Az $\mathbf{x} = (0, \dots, 0, \overset{i}{\overbrace{1}}, 0, \dots, 0)$, $\mathbf{y} = (0, \dots, 0, \overset{j}{\overbrace{1}}, 0, \dots, 0)$ szereposztással (azaz $\mathbf{u} = \mathbf{e}_i$, $\mathbf{v} = \mathbf{e}_j$ esetén) azt kapjuk, hogy az A mátrix i-edik sorának j-edik eleme ugyanaz, mint a B^{T} mátrix i-edik sorának j-edik eleme. Ez minden $i, j \in \{1, \dots, n\}$ esetén igaz, ezért $A = B^{\mathrm{T}}$.
- (ii) $\Longrightarrow$ (iii): Ez majdnem triviális.

- (iii) $\Longrightarrow$ (i): Tfh. A és B egymás transzponáltja. A fent levezetett $\langle \mathbf{u}\varphi, \mathbf{v}\rangle = (\mathbf{x}A)\mathbf{y}^{\mathrm{T}}$ és $\langle \mathbf{u}, \mathbf{v}\psi\rangle = \mathbf{x}(\mathbf{y}B)^{\mathrm{T}}$ összefüggések szerint ekkor $\langle \mathbf{u}\varphi, \mathbf{v}\rangle = \mathbf{x}A\mathbf{y}^{\mathrm{T}} = \mathbf{x}B^{\mathrm{T}}\mathbf{y}^{\mathrm{T}} = \mathbf{x}(\mathbf{y}B)^{\mathrm{T}} = \langle \mathbf{u}, \mathbf{v}\psi\rangle$, tehát (i) valóban teljesül tetszőleges $\mathbf{u}, \mathbf{v} \in V$ vektorokra. □

6.64. Definíció. A fenti tétel szerint minden $\varphi \in \mathrm{Hom}(V, V)$ transzformációhoz pontosan egy olyan $\psi \in \mathrm{Hom}(V, V)$ transzformáció létezik, amelyre az (i) feltétel teljesül. Ezt a ψ transzformációt φ adjungáltjának nevezzük. Jelölés $\psi = \varphi^*$ (és ekkor persze $\varphi = \psi^*$ is fennáll).

6.65. Definíció. Legyen V végesdimenziós euklideszi tér és $\varphi \in \mathrm{Hom}(V, V)$ lineáris transzformáció. Ha $\varphi^* = \varphi$, akkor azt mondjuk, hogy φ önadjungált transzformáció. Ha φ bijektív és $\varphi^* = \varphi^{-1}$, akkor azt mondjuk, hogy φ ortogonális transzformáció.

6.66. Definíció. Legyen $A \in \mathbb{R}^{n\times n}$ valós négyzetes mátrix. Ha $A^{\mathrm{T}} = A$, akkor azt mondjuk, hogy A szimmetrikus mátrix. Ha A nemelfajuló és $A^{\mathrm{T}} = A^{-1}$, akkor azt mondjuk, hogy A ortogonális mátrix.

6.67. Tétel. Tetszőleges V végesdimenziós euklideszi tér és $\varphi \in \mathrm{Hom}(V, V)$ lineáris transzformáció esetén ekvivalensek az alábbiak:

(i) φ önadjungált transzformáció;

(ii) bármely $\mathcal{E}$ ortonormált bázisra $[\![\varphi]\!]_{\mathcal{E}}$ szimmetrikus mátrix;

(iii) van olyan $\mathcal{E}$ ortonormált bázis, amelyre $[\![\varphi]\!]_{\mathcal{E}}$ szimmetrikus mátrix.

Bizonyítás. Tüstént következik a 6.63. Tételből. □

6.68. Tétel. Tetszőleges V végesdimenziós euklideszi tér és $\varphi \in \mathrm{Hom}(V, V)$ lineáris transzformáció esetén ekvivalensek az alábbiak:

(i) φ ortogonális transzformáció;

(ii) bármely $\mathcal{E}$ ortonormált bázisra $[\![\varphi]\!]_{\mathcal{E}}$ ortogonális mátrix;

(iii) van olyan $\mathcal{E}$ ortonormált bázis, amelyre $[\![\varphi]\!]_{\mathcal{E}}$ ortogonális mátrix.

Bizonyítás. Izibe következik a 6.63. Tételből. □

6.69. Állítás. Legyen V végesdimenziós euklideszi tér.

(i) Tetszőleges $\varphi \in \mathrm{Hom}(V, V)$ esetén $\mathrm{Ker}\,\varphi^* = (\mathrm{Im}\,\varphi)^{\perp}$ és $\mathrm{Im}\,\varphi^* = (\mathrm{Ker}\,\varphi)^{\perp}$.

(ii) Ha $\pi \in \mathrm{Hom}(V, V)$ az U altérre történő, W-vel párhuzamos irányú projekció (ahol $V = U \oplus W$), akkor π^* nem más, mint a $W^{\perp}$ altérre történő, $U^{\perp}$-sel párhuzamos irányú projekció.

(iii) A fenti $\pi \in \mathrm{Hom}(V, V)$ projekció akkor és csak akkor önadjungált, ha $U \perp W$ (ilyenkor ortogonális projekcióról beszélünk).

(iv) Ha $\mathbf{u}$ sajátvektora φ-nek λ sajátértékkel, $\mathbf{v}$ sajátvektora φ^*-nak μ sajátértékkel és $\lambda \neq \mu$, akkor $\mathbf{u} \perp \mathbf{v}$. Speciálisan, önadjungált transzformáció különböző sajátértékekhez tartozó sajátvektorai mindig ortogonálisak. (Hasonlítsuk össze ezt a 4.17. Következménnyel!)

Bizonyítás. HF. □

6.70. Állítás. Legyen V végesdimenziós euklideszi tér és $\varphi \in \mathrm{Hom}(V, V)$. Ha az $U \leq V$ altér invariáns φ-re, akkor $U^{\perp}$ invariáns φ^*-ra.

Bizonyítás. Tfh. U invariáns φ-re, és legyen $\mathbf{v} \in U^{\perp}$. Igazolnunk kell, hogy $\mathbf{v}\varphi^* \in U^{\perp}$, azaz $\mathbf{v}\varphi^*$ merőleges minden $\mathbf{u} \in U$ vektorra. Használjuk az adjungált definícióját: $\langle \mathbf{u}, \mathbf{v}\varphi^*\rangle = \langle \mathbf{u}\varphi, \mathbf{v}\rangle$. Ez a belső szorzat valóban nulla, hiszen $\mathbf{u}\varphi \in U$ és $\mathbf{v} \in U^{\perp}$. □

6.9. Ortogonális transzformációk és ortogonális mátrixok

6.71. Tétel. Tetszőleges V végesdimenziós euklideszi tér és $\varphi \in \mathrm{Hom}(V, V)$ lineáris transzformáció esetén ekvivalensek az alábbiak:

(i) φ ortogonális transzformáció (azaz $\varphi^{-1} = \varphi^*$);

(ii) $\forall \mathbf{u}, \mathbf{v} \in V\colon \langle \mathbf{u}\varphi, \mathbf{v}\varphi\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$;

(iii) $\forall \mathbf{v} \in V\colon \|\mathbf{v}\varphi\| = \|\mathbf{v}\|$.

Bizonyítás.

- (i) $\Longrightarrow$ (ii): A bizonyítandó $\langle \mathbf{u}\varphi, \mathbf{v}\varphi\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$ egyenlőség az adjungált definíciója alapján így fogalmazható át: $\langle \mathbf{u}, \mathbf{v}\varphi\varphi^*\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$. Ha $\varphi^* = \varphi^{-1}$, akkor ez nyilván teljesül.

- (ii) $\Longrightarrow$ (i): A (ii) feltétel fenti $\langle \mathbf{u}, \mathbf{v}\varphi\varphi^*\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$ átfogalmazását a belső szorzat linearitása alapján nullára rendezhetjük: $\langle \mathbf{u}, \mathbf{v}\varphi\varphi^* - \mathbf{v}\rangle = 0$. Ez azt jelenti, hogy a $\mathbf{v}\varphi\varphi^* - \mathbf{v}$ vektor merőleges minden vektorra, így saját magára is. Ez pedig csak úgy lehetséges, hogy $\mathbf{v}\varphi\varphi^* - \mathbf{v} = \mathbf{0}$, azaz $\mathbf{v}\varphi\varphi^* = \mathbf{v}$. Ebből következik, hogy $\varphi\varphi^* = \mathrm{id}_V$, tehát $\varphi^* = \varphi^{-1}$.

- (ii) $\Longleftrightarrow$ (iii): Tudjuk, hogy a belső szorzat és a norma kölcsönösen meghatározzák egymást: $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}$ és $\langle \mathbf{u}, \mathbf{v}\rangle = \frac{1}{2} \cdot \left(\|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2\right)$ (lásd a (3) polarizációs azonosságot). Ebből következik, hogy egy leképezés akkor és csak akkor őrzi meg a belső szorzatot, ha megőrzi a normát. □

A fenti tételt az $\mathbb{R}^n$ euklideszi tér $\mathbf{x} \to \mathbf{x}A$ ($A \in \mathbb{R}^{n\times n}$) alakú transzformációira alkalmazva ortogonális mátrixok ekvivalens jellemzéseit kapjuk. A következő tételben nem ezeket írjuk le, hanem ortogonális mátrixok néhány más ekvivalens leírását.

6.72. Tétel. Tetszőleges $A \in \mathbb{R}^{n\times n}$ mátrixra ekvivalensek az alábbiak:

(i) A ortogonális mátrix (azaz $A^{-1} = A^{\mathrm{T}}$);

(ii) A sorai ortonormált bázist alkotnak $\mathbb{R}^n$-ben;

(iii) A oszlopai ortonormált bázist alkotnak $\mathbb{R}^n$-ben;

(iv) van olyan n-dimenziós euklideszi tér és abban olyan $\mathcal{E}, \mathcal{F}$ ortonormált bázisok, hogy $A = [\![\mathcal{F} \to \mathcal{E}]\!]$.

Bizonyítás.

- (i) $\Longleftrightarrow$ (ii): Ha $\mathbf{s}_i$ jelöli az A mátrix i-edik sorát, akkor az AA^{T} mátrix i-edik sorának j-edik eleme nem más, mint $\mathbf{s}_i\mathbf{s}_j^{\mathrm{T}} = \langle \mathbf{s}_i, \mathbf{s}_j\rangle_{\mathbb{R}^n}$. Tehát A sorai akkor és csak akkor alkotnak ortonormált bázist $\mathbb{R}^n$-ben, ha $\mathbf{s}_i\mathbf{s}_j^{\mathrm{T}} = \delta_{ij}$ minden i és j esetén, ez pedig épp azt jelenti, hogy $AA^{\mathrm{T}} = E$.

- (i) $\Longleftrightarrow$ (iii): A fentihez hasonlóan belátható, hogy A oszlopai akkor és csak akkor alkotnak ortonormált bázist $\mathbb{R}^n$-ben, ha $A^{\mathrm{T}}A = E$.

- (ii) $\Longleftrightarrow$ (iv): Legyen V tetszőleges euklideszi tér, legyen $\mathcal{E}$ tetszőleges ortonormált bázisa V-nek, és legyen $\mathcal{F}\colon \mathbf{f}_1, \dots, \mathbf{f}_n$ az az egyértelműen meghatározott bázisa V-nek, amelyre $A = [\![\mathcal{F} \to \mathcal{E}]\!]$. Ekkor tehát az A mátrix i-edik sora: $\mathbf{s}_i = [\![\mathbf{f}_i]\!]_{\mathcal{E}}$ (lásd a 3.34. Megjegyzést). Mivel $\mathcal{E}$ ortonormált bázis, a 6.44. Tétel szerint $\langle \mathbf{f}_i, \mathbf{f}_j\rangle = \mathbf{s}_i\mathbf{s}_j^{\mathrm{T}}$. Ebből következik, hogy $\mathcal{F}$ akkor és csak akkor ortonormált bázisa V-nek, ha A sorai ortonormált bázist alkotnak $\mathbb{R}^n$-ben. □

6.73. Megjegyzés. A 6.71. Tételbeli (i) $\Longleftrightarrow$ (ii) ekvivalencia szerint egy $\varphi\colon V \to V$ lineáris transzformáció akkor és csak akkor ortogonális, ha izomorfizmus, azaz automorfizmusa V-nek. Ebből következik, hogy az ortogonális transzformációk csoportot alkotnak, és ezt a csoportot izomorfia erejéig meghatározza $n := \dim V$ (a 6.57. Tétel szerint feltehetjük, hogy $V = \mathbb{R}^n$), ezért n-dimenziós ortogonális csoportnak nevezzük, és $\mathrm{O}(n)$-nel jelöljük. A transzformációkat egy rögzített ortonormált báziseli mátrixukkal reprezentálva látható, hogy az $\mathrm{O}(n)$ csoport izomorf az $n \times n$-es ortogonális mátrixok csoportjával (lásd a 6.68. Tételt).

A 6.71. Tételbeli (i) $\Longleftrightarrow$ (iii) ekvivalencia szerint az ortogonális transzformációk éppen a távolságtartó lineáris transzformációk, így $\mathrm{O}(n)$ izomorf az $\mathbb{R}^n$ euklideszi tér origót fixáló egybevágósági transzformációinak csoporjával. Az irányítástartó egybevágóságok (vagyis az 1 determinánsú ortogonális mátrixok) alkotják az $\mathrm{SO}(n) \leq \mathrm{O}(n)$ speciális ortogonális csoportot. Két dimenzióban ezek nem mások, mint az origó körüli forgatások, és ezek mátrixai a standard bázisban (vagy bármely ortonormált bázisban) ilyen alakúak:

$$R_\alpha = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}.$$

Minden $\mathrm{SO}(3)$-beli transzformáció valamilyen tengely körüli forgatás. Ha α a forgatás szöge, $\mathbf{b}_1$ a forgástengely irányába mutató egységvektor és $\mathbf{b}_2, \mathbf{b}_3$ a tengelyre és egymásra merőleges egységvektorok, akkor a $\mathcal{B}\colon \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ bázisban így fest a transzformáció mátrixa:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{pmatrix}.$$

Általában minden $\mathrm{SO}(n)$-beli transzformációhoz van olyan ortonormált bázis amelyben mátrixa blokkdiagonális, ahol minden blokk vagy 1×1-es mátrix (1 vagy -1) vagy pedig 2×2-es R_α alakú mátrix (nem feltétlenül egyforma α szögekkel).

6.10. Spektráltétel, főtengelytétel

6.74. Lemma. Végesdimenziós euklideszi tér minden önadjungált lineáris transzformációjának van valós sajátértéke.

Bizonyítás. Legyen V egy n-dimenziós euklideszi tér, és legyen $\varphi \in \mathrm{Hom}(V, V)$ önadjungált lineáris transzformáció. Az algebra alaptétele szerint a p_φ karakterisztikus polinomnak van $\lambda \in \mathbb{C}$ komplex gyöke. Tekintsük az $A = [\![\varphi]\!]_{\mathcal{E}}$ mátrixot, ahol $\mathcal{E}$ egy tetszőleges ortonormált bázisa V-nek. Az A mátrix indukál egy Φ lineáris transzformációt a $\mathbb{C}^n$ komplex vektortéren: $\Phi\colon \mathbb{C}^n \to \mathbb{C}^n$, $\mathbf{x} \to \mathbf{x}A$. Mivel λ gyöke a $p_\varphi = p_A$ polinomnak, van olyan $\mathbf{x} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$ nemnulla vektor, amelyre $\mathbf{x}A = \lambda\mathbf{x}$. A konjugálás felcserélhető a komplex számok összeadásával és szorzásával (automorfizmusa a komplex számtestnek), ezért az $\mathbf{x}A = \lambda\mathbf{x}$ egyenlőség mindkét oldalát konjugálva azt kapjuk, hogy $\overline{\mathbf{x}}\overline{A} = \overline{\lambda}\overline{\mathbf{x}}$ (itt a vektorokat és mátrixokat elemenként konjugáljuk). Ezt a két összefüggést felhasználva számítsuk ki az alábbi két szorzatot:

$$\mathbf{x}A\overline{\mathbf{x}}^{\mathrm{T}} = \lambda\mathbf{x}\overline{\mathbf{x}}^{\mathrm{T}} = \lambda \cdot \sum_{i=1}^{n} x_i\overline{x_i} = \lambda \cdot \sum_{i=1}^{n} |x_i|^2 = \lambda \cdot ☺;$$

$$\overline{\mathbf{x}}\overline{A}\mathbf{x}^{\mathrm{T}} = \overline{\lambda}\overline{\mathbf{x}}\mathbf{x}^{\mathrm{T}} = \overline{\lambda} \cdot \sum_{i=1}^{n} \overline{x_i}x_i = \overline{\lambda} \cdot \sum_{i=1}^{n} |x_i|^2 = \overline{\lambda} \cdot ☺.$$

Vegyük észre, hogy ugyanazt számoltuk ki kétszer, hiszen egyrészt $\mathbf{x}A\overline{\mathbf{x}}^{\mathrm{T}}$ egy 1×1-es mátrix (így megegyezik a saját transzponáltjával), másrészt A valós szimmetrikus mátrix (így $A^{\mathrm{T}} = A = \overline{A}$):

$$\mathbf{x}A\overline{\mathbf{x}}^{\mathrm{T}} = \left(\mathbf{x}A\overline{\mathbf{x}}^{\mathrm{T}}\right)^{\mathrm{T}} = \overline{\mathbf{x}}A^{\mathrm{T}}\mathbf{x}^{\mathrm{T}} = \overline{\mathbf{x}}A\mathbf{x}^{\mathrm{T}} = \overline{\mathbf{x}}\overline{A}\mathbf{x}^{\mathrm{T}}.$$

Azt kaptuk tehát, hogy $\lambda \cdot ☺ = \overline{\lambda} \cdot ☺$, amiből $\lambda = \overline{\lambda}$ következik, hiszen $☺ > 0$. Ezzel beláttuk, hogy $\lambda \in \mathbb{R}$. □

6.75. Tétel (spektráltétel). Legyen $\varphi \in \mathrm{Hom}(V, V)$ lineáris transzformációja a V végesdimenziós euklideszi térnek. Akkor és csak akkor létezik V-nek olyan $\mathcal{F}$ ortonormált bázisa, amely φ sajátvektoraiból áll (azaz amelyre $[\![\varphi]\!]_{\mathcal{F}}$ diagonális mátrix), ha φ önadjungált.

Bizonyítás. A „csak akkor" rész a 6.67. Tétel alapján világos, hiszen a diagonális mátrixok szimmetrikusak. Az „akkor" részt $n := \dim V$ szerinti teljes indukcióval bizonyítjuk. Az $n = 1$ eset triviális, hiszen minden 1×1-es mátrix diagonális. Indukciós hipotézisként tfh. minden $(n-1)$-dimenziós euklideszi tér minden önadjungált lineáris transzformációjához létezik ortonormált sajátbázis. Legyen $\dim V = n$ és legyen $\varphi \in \mathrm{Hom}(V, V)$ önadjungált lineáris transzformáció. A 6.74. Lemma szerint létezik sajátértéke φ-nek, és így persze létezik $\mathbf{v} \in V$ sajátvektora is. A sajátvektort normálva kapunk egy $\mathbf{f}_1 := \frac{1}{\|\mathbf{v}\|}\mathbf{v}$ egységnyi normájú sajátvektort. Legyen $W = [\mathbf{f}_1]^\perp$; ekkor a 6.70. Állítás következtében W invariáns altere a $\varphi^* = \varphi$ transzformációnak (hiszen $[\mathbf{f}_1]$ invariáns φ-re). Ezért van értelme megszorítani a φ transzformációt a W altérre; így kapjuk a $\varphi|_W \in \mathrm{Hom}(W, W)$ transzformációt, ami persze szintén önadjungált. Alkalmazzuk az indukciós hipotézist erre a transzformációra (megtehetjük, mert a 6.59. Tétel szerint $\dim W = n - 1$), így kapunk egy $\mathbf{f}_2, \ldots, \mathbf{f}_n$ ortonormált bázist W-ben, ami $\varphi|_W$ sajátvektoraiból áll. Ezek természetesen φ-nek is sajátvektorai, így az $\mathbf{f}_1$ vektorral együtt kapjuk V-nek egy $\mathcal{F}\colon \mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ ortonormált bázisát, ami φ sajátvektoraiból áll. Ekkor $[\![\varphi]\!]_{\mathcal{F}}$ diagonális mátrix; a főátlóban az $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ sajátvektorokhoz tartozó sajátértékek állnak. □

6.76. Következmény (spektrálfelbontás). Legyen V egy n-dimenziós euklideszi tér, és legyen $\varphi \in \mathrm{Hom}(V, V)$ önadjungált lineáris transzformáció. Ekkor φ előáll $\varphi = \lambda_1\pi_1 + \cdots + \lambda_n\pi_n$ alakban, ahol $\lambda_1, \ldots, \lambda_n$ a φ transzformáció sajátértékei (mind valós, és mindegyik annyiszor szerepel, amennyi az algebra/geometriai multiplicitása), a $\pi_1, \ldots, \pi_n \in \mathrm{Hom}(V, V)$ transzformációk pedig páronként ortogonális egydimenziós (saját)alterekre való ortogonális projekciók.

Bizonyítás. A spektráltételből tudjuk, hogy létezik $\mathcal{F}\colon \mathbf{f}_1, \ldots, \mathbf{f}_n$ ortonormált sajátbázis. Legyenek $\lambda_1, \ldots, \lambda_n$ a megfelelő sajátértékek, és legyen π_i az $[\mathbf{f}_i]$ altérre való merőleges vetítés. Ekkor $\mathbf{f}_i\pi_j = \delta_{ij}\mathbf{f}_i$, és így $\mathbf{f}_i \sum \lambda_j\pi_j = \lambda_i\mathbf{f}_i$. Tehát az $\mathbf{f}_i$ vektorokon a $\sum \lambda_j\pi_j$ transzformáció megegyezik φ-vel. Mivel ezek a vektorok kifeszítik a V vektorteret, a linearitásból következően $\sum \lambda_j\pi_j$ mindenütt megegyezik φ-vel. □

6.77. Definíció. Azt mondjuk, hogy az $A \in \mathbb{R}^{n\times n}$ mátrix <u>ortogonálisan diagonalizálható</u>, ha létezik olyan $Q \in \mathbb{R}^{n\times n}$ ortogonális mátrix, amelyre QAQ^{-1} diagonális.

6.78. Következmény (spektráltétel). Egy valós négyzetes mátrix akkor és csak akkor diagonalizálható ortogonálisan, ha szimmetrikus.

Bizonyítás. Legyen $A \in \mathbb{R}^{n\times n}$ négyzetes mátrix, és tekintsük a $V = \mathbb{R}^n$ euklideszi téren a $\varphi\colon V \to V$, $\mathbf{x} \to \mathbf{x}A$ lineáris transzformációt. Ennek mátrixa a standard $\mathcal{E}$ bázisban $[\![\varphi]\!]_{\mathcal{E}} = A$, és ha $\mathcal{F}$ egy tetszőleges bázisa V-nek, akkor $[\![\varphi]\!]_{\mathcal{F}} = QAQ^{-1}$, ahol $Q = [\![\mathcal{F} \to \mathcal{E}]\!]$ (lásd a 3.39. Következményt). A 6.72. Tétel szerint Q akkor és csak akkor ortogonális mátrix, ha $\mathcal{F}$ ortonormált bázis. Tehát az A mátrix ortogonális diagonalizálhatósága ekvivalens azzal, hogy a φ transzformációhoz létezik ortonormált sajátbázis. Utóbbi a 6.75. Tétel alapján akkor és csak akkor teljesül, ha φ önadjungált, azaz A szimmetrikus mátrix (lásd a 6.67. Tételt). □

6.79. Következmény (főtengelytétel). Végesdimenziós euklideszi téren értelmezett kvadratikus alakhoz mindig van olyan ortonormált bázis, amelyben a kvadratikus alak mátrixa diagonális. Másképp fogalmazva: minden valós kvadratikus alak kanonikus alakra hozható olyan lineáris helyettesítéssel, melynek mátrixa ortogonális.

Bizonyítás. Legyen V egy n-dimenziós euklideszi tér, és legyen $q\colon V \to \mathbb{R}$ kvadratikus alak. Ha $\mathcal{E}$ ortonormált bázisa V-nek, akkor $A := [\![q]\!]_{\mathcal{E}} \in \mathbb{R}^{n\times n}$ szimmetrikus mátrix. A 6.78. Következmény szerint van olyan $S \in \mathbb{R}^{n\times n}$ ortogonális mátrix, amelyre SAS^{-1} diagonális. Mivel S ortogonális, $S^{-1} = S^{\mathrm{T}}$, tehát SAS^{T} diagonális mátrix. Ez pedig azt jelenti, hogy az S mátrix által leírt lineáris helyettesítés kanonikus alakra hozza q-t (lásd a 6.17. Megjegyzést), vagyis $[\![q]\!]_{\mathcal{F}}$ diagonális mátrix, ahol $\mathcal{F}$ az az egyértelműen meghatározott bázisa V-nek, amelyre $[\![\mathcal{F} \to \mathcal{E}]\!] = S$ (ez a bázis a 6.72. Tétel szerint ortonormált). □

6.80. feladat. Adjunk meg $\mathbb{R}^2$-ben az alábbi A szimmetrikus mátrix sajátvektoraiból álló ortonormált bázist.

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix} \in \mathbb{R}^{2\times 2}.$$

Megoldás. Először számítsuk ki a karakterisztikus polinomot:

$$p_A = \begin{vmatrix} 4-x & 2 \\ 2 & 4-x \end{vmatrix} = (x-4)^2 - 4 = (x-6)(x-2).$$

Tehát két sajátérték van: $\lambda_1 = 6$ (algebrai multiplicitása: $a_1 = 1$), illetve $\lambda_2 = 2$ (algebrai multiplicitása: $a_2 = 1$). Most keressünk bázist mindkét sajátaltérben (már lehet tudni, hogy mindkettő egydimenziós lesz). Ehhez az $(A - \lambda E)^{\mathrm{T}}$ mátrixú homogén lineáris egyenletrendszert kell megoldanunk mindkét sajátértékre (a transzponálást elhagyhatjuk, mert a mátrix szimmetrikus):

$$(A - \lambda_1 E)^{\mathrm{T}} = A - 6E = \begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \implies S_6 = [(1,1)];$$

$$(A - \lambda_2 E)^{\mathrm{T}} = A - 2E = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \implies S_2 = [(-1,1)].$$

Foglaljuk össze az eddigieket. Két sajátérték van:

- $\lambda_1 = 6$, algebrai multiplicitása $a_1 = 1$, geometriai multiplicitása $g_1 = 1$, a sajátaltér egy bázisvektora $(-1, 1)$.
- $\lambda_2 = 2$, algebrai multiplicitása $a_2 = 1$, geometriai multiplicitása $g_2 = 1$, a sajátaltér egy bázisvektora $(1, 1)$;

Már meg is van egy sajátbázis, és „szerencsére" a két sajátvektor ortogonális (nem szerencse, hanem a 6.69. Állítás utolsó pontjának következménye, de a spektráltétel bizonyításából is látszik, hogy szimmetrikus mátrix esetén a különböző sajátértékekhez tartozó sajátvektorok merőlegesek). Így már csak normálnunk kell, hogy megkapjuk az ortonormált $\mathcal{F}$ sajátbázist:

$$\mathbf{f}_1 = \big(1/\sqrt{2},\, 1/\sqrt{2}\big), \qquad \mathbf{f}_2 = \big(-1/\sqrt{2},\, 1/\sqrt{2}\big).$$

Írjuk fel az $[\![\mathcal{F} \to \mathcal{E}]\!]$ bázisáttérés mátrixát ($\mathcal{E}$ a standard bázis):

$$Q := [\![\mathcal{F} \to \mathcal{E}]\!] = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}.$$

Ellenőrizzük a diagonalizálást:

$$QAQ^{-1} = QAQ^{\mathrm{T}} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \cdot \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 4 \end{pmatrix}$$

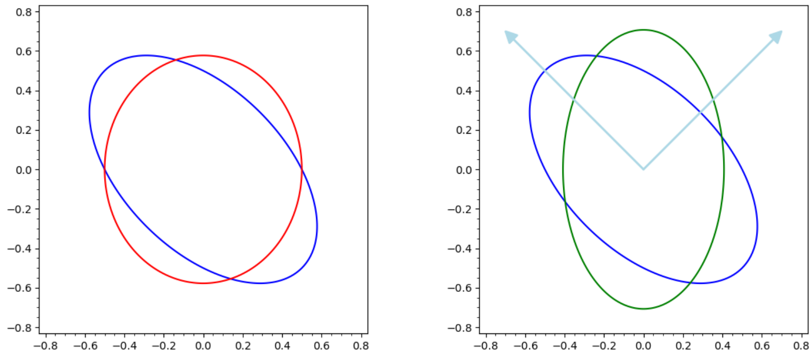
◇

6.81. feladat. Ábrázoljuk a $4x_1^2 + 4x_1x_2 + 4x_2^2 = 1$ másodrendű görbét.

Megoldás. A megadott görbe nem más, mint az előző feladatbeli A mátrixhoz tartozó q kvadratikus alak egy szintvonala: $q(x_1, x_2) = \mathbf{x}A\mathbf{x}^{\mathrm{T}} = 4x_1^2 + 4x_1x_2 + 4x_2^2$. Hozzuk kanonikus alakra ezt a kvadratikus alakot a 6.21. Tétel bizonyításának módszerével (lényegében teljes négyzetté alakítással, lásd még a 6.20. feladatot):

$$q(x_1, x_2) = 4x_1^2 + 4x_1x_2 + 4x_2^2 = 4\Big(x_1 + \frac{x_2}{2}\Big)^2 + 3x_2^2 = 4y_1^2 + 3y_2^2.$$

A bal oldali ábrán kék szín mutatja az $4x_1^2 + 4x_1x_2 + 4x_2^2 = 1$ egyenletű ellipszist (az eredeti kvadratikus alak), piros szín pedig a $4x_1^2 + 3x_2^2 = 1$ egyenletű ellipszist (a fenti kanonikus alak).



Látható, hogy a két ellipszis nem ugyanolyan alakú (nem ugyanaz a nagy- és kistengely aránya). A kanonikus alak nem is egyértelmű, további helyettesítéssel megkaphattuk volna például a $(2y_1)^2 + (\sqrt{3}y_2)^2 = z_1^2 + z_2^2$ normálalakot is, ami egy kört rajzolna ki. Egy tetszőleges nemelfajuló helyettesítés tehát eltorzíthatja az ellipszis alakját (de azért ellipszis marad). Ha viszont a spektráltételben szereplő ortogonális helyettesítést hajtjuk végre, akkor a sajátértékek jelennek meg a kanonikus alakban: $6y_1^2 + 2y_2^2$. A jobb oldali ábrán zöld szín mutatja ezt a $6x_1^2 + 2x_2^2 = 1$ egyenletű ellipszist. Ez már egybevágó az eredeti kék ellipszissel, csak el van forgatva (kanonikus alak esetén az ellipszis tengelyei mindig párhuzamosak a koordinátatengelyekkel). A világoskék színnel berajzolt sajátvektorok mutatják a kék ellipszis tengelyeinek irányát. Tehát a spektráltételből pontosan megkapjuk az ellipszis „alakját” (sajátértékek) és „állását” (sajátvektorok, avagy a Q mátrix által leírt forgatás). Innen ered a főtengelytétel elnevezés, és kvadratikus alak ilyen „főtengelyes” kanonikus alakra hozását nevezik főtengely-transzformációnak. Persze nem csak ellipszist kaphatunk, hanem más másodrendű görbét vagy felületet is, de a sajátvektorok mindig a szimmetriatengelyeket mutatják. Itt található egy háromváltozós kvadratikus alak főtengelytranszformációja „háromdimenziós” interaktív ábrával (SageMath):
http://www.math.u-szeged.hu/~twaldha/tanitas/linalg2bsc_2023osz/kvadr-sage.html. ◇

6.82. feladat. Írjuk fel a 6.80. feladatbeli A mátrix spektrálfelbontását (lásd a 6.76. Következményt).

Megoldás. A 6.80. feladat megoldásában elvégeztük az A mátrix ortogonális diagonalizálását. Ezt átalakítva megkapjuk a spektrálfelbontást:

$$\begin{aligned} A \;&=\; Q^{-1} \cdot \begin{pmatrix} 6 & 0 \\ 0 & 2 \end{pmatrix} \cdot Q \;=\; Q^{\mathrm{T}} \cdot \begin{pmatrix} 6 & 0 \\ 0 & 2 \end{pmatrix} \cdot Q \;=\; 6 \cdot Q^{\mathrm{T}} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot Q \;+\; 2 \cdot Q^{\mathrm{T}} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \cdot Q \\ &=\; 6 \cdot \mathbf{f}_1^{\mathrm{T}}\mathbf{f}_1 + 2 \cdot \mathbf{f}_2^{\mathrm{T}}\mathbf{f}_2 \;=\; 6 \cdot P_1 + 2 \cdot P_2. \end{aligned}$$

Itt $P_i := \mathbf{f}_i^{\mathrm{T}}\mathbf{f}_i \in \mathbb{R}^{2\times 2}$ az $\mathbf{f}_i$ sajátvektorra való merőleges vetítés mátrixa a standard bázisban (ellenőrizzük!). ◇

6.11. Szingulárisérték-felbontás

A 6.76. Következményben láttuk, hogy minden önadjungált transzformáció felírható páronként ortogonális egydimenziós alterekre való ortogonális projekciók lineáris kombinációjaként. Itt található egy interaktív illusztráció (SageMath) arra, hogy hogyan lehet ezt használni szimmetrikus mátrixok közelítésére:
http://www.math.u-szeged.hu/~twaldha/tanitas/linalg2bsc_2023osz/svd-sage.html.
A fenti illusztráció mutatja azt is, hogyan lehet a spektrálfelbontást képtömörítésre használni. Ez azonban csak akkor működik, ha a kép négyzet alakú, és szimmetrikus a főátlóra, ami nem túl realisztikus elvárás... Szerencsére létezik a spektrálfelbontáshoz hasonló felbontás tetszőleges, akár nem is négyzetes mátrixokra is; ezt fogjuk a következőkben levezetni.

Legyen $M \in \mathbb{R}^{k\times n}$ tetszőleges valós mátrix. Ekkor $MM^{\mathrm{T}} \in \mathbb{R}^{k\times k}$ szimmetrikus mátrix, tehát rá már tudjuk alkalmazni a spektráltételt.

6.83. Állítás. Az MM^{T} mátrix pozitív definit vagy szemidefinit, és így a sajátértékei mind nemnegatív valós számok.

Bizonyítás. HF. ☐

Rendezzük a sajátértékeket csökkenő (pontosabban nemnövekvő) sorrendbe, így az esetleges nulla sajátértékek a sor végére kerülnek: $\lambda_1 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_k = 0$.

6.84. Állítás. A fenti r szám, azaz MM^{T} pozitív sajátértékeinek száma nem más, mint MM^{T} rangja, és ez megegyezik M és $M^{\mathrm{T}}M$ rangjával is.

Bizonyítás. HF. ☐

Most már jöhet a spektráltétel: létezik olyan $\mathbf{u}_1, \dots, \mathbf{u}_k$ ortonormált bázis $\mathbb{R}^k$-ban, ami MM^{T} sajátvektoraiból áll:

$$\mathbf{u}_i MM^{\mathrm{T}} = \lambda_i \mathbf{u}_i \ (i = 1, \dots, k).$$

Az $\mathbf{u}_i M \in \mathbb{R}^n$ vektorokat fogjuk normálni, majd ezekből konstruálunk $\mathbb{R}^n$-ben is egy ortonormált bázist, ami $M^{\mathrm{T}}M$ sajátvektoraiból áll. Ehhez először ellenőrizzük, hogy ezek a vektorok páronként ortogonálisak, és számítsuk ki a normájukat (az $\mathbb{R}^n$-beli standard $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}\mathbf{y}^{\mathrm{T}}$ belső szorzatot használjuk).

6.85. Állítás. Az $\mathbf{u}_i M \in \mathbb{R}^n$ $(i = 1, \dots, k)$ vektorok páronként ortogonálisak és $\|\mathbf{u}_i M\| = \sqrt{\lambda_i}$.

Bizonyítás. HF. ☐

Az $\mathbf{u}_{r+1}M, \dots, \mathbf{u}_k M$ vektorok tehát nullvektorok, de az $\mathbf{u}_1 M, \dots, \mathbf{u}_r M$ vektorokat normálva egy $\mathbf{v}_1, \dots, \mathbf{v}_r$ ortonormált vektorrendszert kapunk $\mathbb{R}^n$-ben:

$$\mathbf{v}_i := \frac{1}{\sqrt{\lambda_i}} \mathbf{u}_i M \ (i = 1, \dots, r).$$

6.86. Állítás. Minden $i \in \{1, \dots, r\}$ esetén a $\mathbf{v}_i$ vektor sajátvektora az $M^{\mathrm{T}}M$ mátrixnak λ_i sajátértékkel.

Bizonyítás. HF. ☐

Jöhet a spektráltétel az $M^{\mathrm{T}}M \in \mathbb{R}^{n \times n}$ szimmetrikus mátrixra is. A 6.83. Állítást M helyett az M^{T} mátrixra alkalmazva azt kapjuk, hogy $M^{\mathrm{T}}M$ sajátértékei is nemnegatívak, és a 6.84. Állítás szerint $M^{\mathrm{T}}M$ sajátértékei közül (multiplicitással számolva) r pozitív, a maradék $n - r$ sajátérték pedig nulla. A fenti 6.86. Állításban már meg is találtuk az r pozitív sajátértéket $(\lambda_1, \dots, \lambda_r)$ és hozzájuk tartozó sajátvektorokat $(\mathbf{v}_1, \dots, \mathbf{v}_r)$. Tehát az általánosság megszorítása nélkül feltehetjük, hogy a spektráltétel által szolgáltatott ortonormált sajátbázis első r eleme $\mathbf{v}_1, \dots, \mathbf{v}_r$, a többi pedig legyen $\mathbf{v}_{r+1}, \dots, \mathbf{v}_n$ (utóbbiak a 0 sajátértékhez tartoznak).

Foglaljuk össze, hogy mit tudunk eddig az $\mathbf{u}_i$ és $\mathbf{v}_i$ vektorokról:

$\mathbf{u}_1, \dots, \mathbf{u}_k$ ONB $\mathbb{R}^k$-ban,

$$i \in \{1, \dots, r\} \text{ esetén } \mathbf{u}_i MM^{\mathrm{T}} = \lambda_i \mathbf{u}_i,$$
$$i \in \{r+1, \dots, k\} \text{ esetén } \mathbf{u}_i MM^{\mathrm{T}} = \mathbf{0},$$

$\mathbf{v}_1, \dots, \mathbf{v}_n$ ONB $\mathbb{R}^n$-ben,

$$i \in \{1, \dots, r\} \text{ esetén } \mathbf{v}_i M^{\mathrm{T}}M = \lambda_i \mathbf{v}_i,$$
$$i \in \{r+1, \dots, n\} \text{ esetén } \mathbf{v}_i M^{\mathrm{T}}M = \mathbf{0}.$$

Azért az MM^{T} és $M^{\mathrm{T}}M$ mátrixokkal foglalkoztunk eddig, mert ezek szimmetrikusak, így tudtuk rájuk alkalmazni a spektráltételt. Node minket elsősorban nem ezek a mátrixok érdekelnek, hanem az M mátrix. A következő állítás az M mátrix (és a szimmetria kedvéért az M^{T} mátrix) hatását írja le az $\mathbf{u}_i$, $\mathbf{v}_i$ vektorokon.

6.87. Állítás. Legyen $\sigma_i = \sqrt{\lambda_i}$ $(i = 1, \dots, r)$. Ekkor

$$i \in \{1, \dots, r\} \text{ esetén } \mathbf{u}_i M = \sigma_i \mathbf{v}_i, \qquad i \in \{1, \dots, r\} \text{ esetén } \mathbf{v}_i M^{\mathrm{T}} = \sigma_i \mathbf{u}_i,$$
$$i \in \{r+1, \dots, k\} \text{ esetén } \mathbf{u}_i M = \mathbf{0}, \qquad i \in \{r+1, \dots, n\} \text{ esetén } \mathbf{v}_i M^{\mathrm{T}} = \mathbf{0}.$$

Bizonyítás. HF. ☐

Legyen $\varphi\colon \mathbb{R}^k \to \mathbb{R}^n$, $\mathbf{x} \mapsto \mathbf{x}M$ az M mátrix által indukált lineáris leképezés és $\varphi^*\colon \mathbb{R}^n \to \mathbb{R}^k$, $\mathbf{y} \mapsto \mathbf{y}M^{\mathrm{T}}$ az M^{T} mátrix által indukált lineáris leképezés. (A jelölés nem véletlenül ugyanaz, mint az adjungáltnál. Mi ugyan csak $V \to V$ lineáris transzformációk adjungáltját definiáltuk, de a definíció kiterjeszthető $U \to V$ lineáris leképezésekre is.)

6.88. Állítás. A φ és φ^* leképezések mag- és képtere így alakul (hasonlítsuk össze a 6.69. Állítás első pontjával!):

$$[\mathbf{u}_1, \dots, \mathbf{u}_r] = (\operatorname{Ker} \varphi)^{\perp} = \operatorname{Im} \varphi^*, \qquad [\mathbf{v}_1, \dots, \mathbf{v}_r] = (\operatorname{Ker} \varphi^*)^{\perp} = \operatorname{Im} \varphi,$$
$$[\mathbf{u}_{r+1}, \dots, \mathbf{u}_k] = \operatorname{Ker} \varphi = (\operatorname{Im} \varphi^*)^{\perp}, \qquad [\mathbf{v}_{r+1}, \dots, \mathbf{v}_n] = \operatorname{Ker} \varphi^* = (\operatorname{Im} \varphi)^{\perp}.$$

Bizonyítás. HF. □

A kinullázott vektorok nem túl érdekesek, a φ leképezés „java” az $[\mathbf{u}_1,\dots,\mathbf{u}_r] \leq \mathbb{R}^k$ altéren történik, amit a $[\mathbf{v}_1,\dots,\mathbf{v}_r] \leq \mathbb{R}^n$ altérre képez. Ezekben az alterekben az $\mathbf{u}_1,\dots,\mathbf{u}_r$ és $\mathbf{v}_1,\dots,\mathbf{v}_r$ ortonormált bázisok kicsit olyasmik, mintha sajátbázisok lennének, csak itt a φ leképezés az $\mathbf{u}_i$ vektort nem $\mathbf{u}_i$ skalárszorosába képezi, hanem $\mathbf{v}_i$ skalárszorosába: $\mathbf{u}_i\varphi = \sigma_i\mathbf{v}_i$ $(i = 1,\dots,r)$. (A φ^* leképezés pedig ugyanezt csinálja „visszafelé”: $\mathbf{v}_i\varphi^* = \sigma_i\mathbf{u}_i$ teljesül $i = 1,\dots,r$ esetén.) A σ_i számokat szinguláris értékeknek nevezzük, az $\mathbf{u}_i$ és $\mathbf{v}_i$ vektorokat pedig szinguláris vektoroknak nevezzük. (Szokás a nullát is megengedni szinguláris értéknek, ekkor összesen nem r, hanem $\min(k,n)$ darab szinguláris érték van, amelyek közül az első r pozitív, a többi nulla.)

A 6.87. Állítást mátrixokkal megfogalmazva kapjuk a következő tételt.

6.89. Tétel (szingulárisérték-felbontás). Legyen $U \in \mathbb{R}^{k\times k}$ az a mátrix, amit az $\mathbf{u}_1,\dots,\mathbf{u}_k$ vektorok egymás alá írásával kapunk, hasonlóan legyen $V \in \mathbb{R}^{n\times n}$ a $\mathbf{v}_1,\dots,\mathbf{v}_n$ vektorokból, mint sorvektorokból alkotott mátrix, és legyen $\Sigma \in \mathbb{R}^{k\times n}$ az a mátrix, amelyben a főátló első r eleme $\sigma_1,\dots,\sigma_r$, a többi eleme pedig nulla (a főátlón és azon kívül is). Ekkor $U \in \mathrm{O}(k)$ és $V \in \mathrm{O}(n)$, továbbá fennáll az $UM = \Sigma V$ összefüggés. Utóbbit átrendezve kapjuk az M mátrix szingulárisérték-felbontását:

$$M = U^{\mathrm{T}}\Sigma V = \sigma_1\cdot\mathbf{u}_1^{\mathrm{T}}\mathbf{v}_1 + \dots + \sigma_r\cdot\mathbf{u}_r^{\mathrm{T}}\mathbf{v}_r$$

Bizonyítás. HF. □

A szingulárisérték-felbontás (angolul SVD = Singular Value Decomposition) a spektrálfelbontáshoz hasonlóan 1 rangú mátrixok összegére bontja az M mátrixot. Ha ezek közül csak az első néhányat (a legnagyobb szinguláris értékekhez tartozókat) tartjuk meg, akkor egy közelítését kapjuk az M mátrixnak. Ez már akármilyen (nem szimmetrikus, nem is négyzet alakú) képek tömörítésére használható:
http://www.math.u-szeged.hu/~twaldha/tanitas/linalg2bsc_2023osz/svd-sage.html.
A szingulárisérték-felbontás másik fontos alkalmazása az úgynevezett főkomponens-analízis (PCA = Principal Component Analysis), ami egy statisztikai eljárás nagy, sokdimenziós adathalmazok elemzésére.

6.90. feladat. Határozzuk meg az alábbi M mátrix szingulárisérték-felbontását:

$$M = \begin{pmatrix} 1 & \sqrt{2}/2-2 & \sqrt{2}/2+2 \\ 1 & \sqrt{2}/2+2 & \sqrt{2}/2-2 \end{pmatrix} \in \mathbb{R}^{2\times 3}.$$

Megoldás. Először számítsuk ki az MM^{T} mátrixot:

$$MM^{\mathrm{T}} = \begin{pmatrix} 10 & -6 \\ -6 & 10 \end{pmatrix} \in \mathbb{R}^{2\times 2}.$$

Írjuk fel MM^{T} karakterisztikus polinomját:

$$p_{MM^{\mathrm{T}}} = \begin{pmatrix} 10-x & -6 \\ -6 & 10-x \end{pmatrix} = (x-10)^2 - 36 = (x-16)(x-4).$$

Tehát MM^{T} sajátértékei $\lambda_1 = 16$ és $\lambda_2 = 4$; mindkettő egyszeres algebrai multiplicitású, ezért a sajátalterek egydimenziósak lesznek. Egyszerű számolás adja a sajátaltereket:

$$S_{16} = \{(x,-x) : x\in\mathbb{R}\}, \qquad S_4 = \{(x,x) : x\in\mathbb{R}\}.$$

Legyen $\mathbf{u}_1$ normált sajátvektor az S_{16} altérből, $\mathbf{u}_2$ pedig normált sajátvektor az S_4 altérből:

$$\mathbf{u}_1 = \big(\sqrt{2}/2,\, -\sqrt{2}/2\big), \qquad \mathbf{u}_2 = \big(\sqrt{2}/2,\, \sqrt{2}/2\big).$$

A következő lépés az $\mathbf{u}_iM$ vektorok kiszámítása:

$$\mathbf{u}_1M = \big(0,\, -2\sqrt{2},\, 2\sqrt{2}\big), \qquad \mathbf{u}_2M = \big(\sqrt{2},\, 1,\, 1\big).$$

Normálva ezeket a vektorokat megkapjuk a $\mathbf{v}_1$ és $\mathbf{v}_2$ vektorokat (ellenőrizzük, hogy $\|\mathbf{u}_1\| = \sqrt{\lambda_1} = 4$ és $\|\mathbf{u}_2\| = \sqrt{\lambda_2} = 2$):

$$\mathbf{v}_1 = \big(0,\, -\sqrt{2}/2,\, \sqrt{2}/2\big), \qquad \mathbf{v}_2 = \big(\sqrt{2}/2,\, 1/2,\, 1/2\big).$$

Ez a két vektor merőleges egymásra (ellenőrizzük!) és sajátvektorai az $M^{\mathrm{T}}M \in \mathbb{R}^{3\times 3}$ szimmetrikus mátrixnak rendre $\lambda_1 = 16$ és $\lambda_2 = 4$ sajátértékekkel (ellenőrizzük!). A $\mathbf{v}_3$ vektor kiszámítására két lehetőségünk is van: vagy kiszámítjuk az $M^{\mathrm{T}}M$ mátrix 0 sajátértékhez tartozó sajátalterét (egydimenziós lesz), és ebből veszünk egy normált sajátvektort, vagy pedig a $\mathbf{v}_1,\mathbf{v}_2$ ortonormált vektorrendszert „csak úgy” kibővítjük $\mathbb{R}^3$ ortonormált bázisává. Akár így, akár úgy, a $\mathbf{v}_3$ vektor egy (-1)-es szorzótól eltekintve egyértelműen meghatározott: $\mathbf{v}_3 = \big(-\sqrt{2}/2,\, 1/2,\, 1/2\big)$. A szinguláris értékek: $\sigma_1 = \sqrt{\lambda_1} = 4$ és $\sigma_2 = \sqrt{\lambda_2} = 2$, az $\mathbf{u}_i$ és $\mathbf{v}_i$ vektorok „szingularitása” pedig a következőket jelenti (ellenőrizzük!):

$$\begin{array}{llll} \mathbf{u}_1, \mathbf{u}_2 & \text{ONB } \mathbb{R}^2\text{-ben}, & \mathbf{u}_1 M = \sigma_1 \mathbf{v}_1, & \mathbf{u}_2 M = \sigma_2 \mathbf{v}_2, \\ \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 & \text{ONB } \mathbb{R}^3\text{-ban}, & \mathbf{v}_1 M^{\mathrm{T}} = \sigma_1 \mathbf{u}_1, & \mathbf{v}_2 M^{\mathrm{T}} = \sigma_2 \mathbf{u}_2, \quad \mathbf{v}_3 M^{\mathrm{T}} = \mathbf{0}. \end{array}$$

A szingulárisérték-felbontást adó mátrixok:

$$U = \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix} \in \mathrm{O}(2), \quad V = \begin{pmatrix} 0 & -\sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & 1/2 & 1/2 \\ -\sqrt{2}/2 & 1/2 & 1/2 \end{pmatrix} \in \mathrm{O}(3), \quad \Sigma = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}.$$

Maga a szingulárisérték-felbontás pedig így fest (ellenőrizzük!):

$$M = \begin{pmatrix} 1 & \sqrt{2}/2 - 2 & \sqrt{2}/2 + 2 \\ 1 & \sqrt{2}/2 + 2 & \sqrt{2}/2 - 2 \end{pmatrix} = U^{\mathrm{T}} \Sigma V = \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix} \cdot \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & -\sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & 1/2 & 1/2 \\ -\sqrt{2}/2 & 1/2 & 1/2 \end{pmatrix}$$

$$= \sigma_1 \cdot \mathbf{u}_1^{\mathrm{T}} \mathbf{v}_1 + \sigma_2 \cdot \mathbf{u}_2^{\mathrm{T}} \mathbf{v}_2 = \begin{pmatrix} 0 & -2 & 2 \\ 0 & 2 & -2 \end{pmatrix} + \begin{pmatrix} 1 & \sqrt{2}/2 & \sqrt{2}/2 \\ 1 & \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}. \qquad \diamond$$

Függelék

F.1. Polinommátrixok és mátrixpolinomok szorzatának kapcsolata

Legyenek $f_{ij} \in T[x]$ polinomok ($i, j = 1, \ldots, n$), és alkossunk ezekből egy $A = (f_{ij})_{i,j=1}^n \in T[x]^{n\times n}$ polinommátrixot:

$$A = \begin{pmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1} & f_{n2} & \cdots & f_{nn} \end{pmatrix}.$$

Szükségünk lesz arra, hogy kiírjuk a polinomok együtthatóit, és ehhez sajnos tripla indexeket kell használnunk: legyen $a_{ij,k}$ az f_{ij} polinomban x^k együtthatója, azaz $f_{ij} = \sum_k a_{ij,k}x^k$. A polinomok fokszámai lehetnek különbözőek, ezt az egyszerűség kedvéért nem is jelezzük (így is elég bonyolult a jelölés). Teljes pompájában tehát így fest az A polinommátrix:

$$A = \begin{pmatrix} \sum_k a_{11,k}x^k & \sum_k a_{12,k}x^k & \cdots & \sum_k a_{1n,k}x^k \\ \sum_k a_{21,k}x^k & \sum_k a_{22,k}x^k & \cdots & \sum_k a_{2n,k}x^k \\ \vdots & \vdots & \ddots & \vdots \\ \sum_k a_{n1,k}x^k & \sum_k a_{n2,k}x^k & \cdots & \sum_k a_{nn,k}x^k \end{pmatrix}.$$

Legyen A_k a k-adfokú tagok együtthatóiból képezett mátrix: $A_k = (a_{ij,k})_{i,j=1}^n$. Ezek lesznek az A mátrixnak megfelelő f mátrixpolinom együtthatói: $f = \sum_k A_k x^k \in T^{n\times n}[x]$. Teljes pompájában tehát így fest az f mátrixpolinom:

$$f = \sum_k \begin{pmatrix} a_{11,k} & a_{12,k} & \cdots & a_{1n,k} \\ a_{21,k} & a_{22,k} & \cdots & a_{2n,k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1,k} & a_{n2,k} & \cdots & a_{nn,k} \end{pmatrix} x^k.$$

Vegyünk egy B polinommátrixot és a neki megfelelő g mátrixpolinomot, ezekben legyen $b_{ij,k}$ az (i, j)-edik elemben a k-adfokú tag együtthatója, illetve a k-adfokú tag együtthatójában az (i, j)-edik elem:

$$B = (g_{ij})_{i,j=1}^n \in T[x]^{n\times n}, \text{ ahol } g_{ij} = \sum_k b_{ij,k}x^k \in T[x], \text{ és}$$
$$g = \sum_k B_k x^k \in T^{n\times n}[x], \text{ ahol } B_k = (b_{ij,k})_{i,j=1}^n.$$

Azt szeretnénk megmutatni, hogy a polinommátrixok szorzása összhangban van a mátrixpolinomok szorzásával, vagyis az $A \cdot B$ polinommátrixnak megfelelő mátrixpolinom éppen $f \cdot g$. A mátrixok szorzásának definíciója alapján az $A \cdot B$ polinommátrix (p, q)-adik eleme

$$\sum_r f_{pr}g_{rq} = \sum_r \Big(\sum_k a_{pr,k}x^k\Big)\Big(\sum_\ell b_{rq,\ell}x^\ell\Big) \in T[x].$$

Ebből kiolvasható, hogy

(5) $\qquad$ az $A \cdot B$ mátrix (p, q)-adik elemében x^t együtthatója $\displaystyle\sum_r \sum_{k+\ell=t} a_{pr,k}b_{rq,\ell}$.

A polinomok szorzásának definíciója alapján az $f \cdot g$ mátrixpolinomban x^t együtthatója

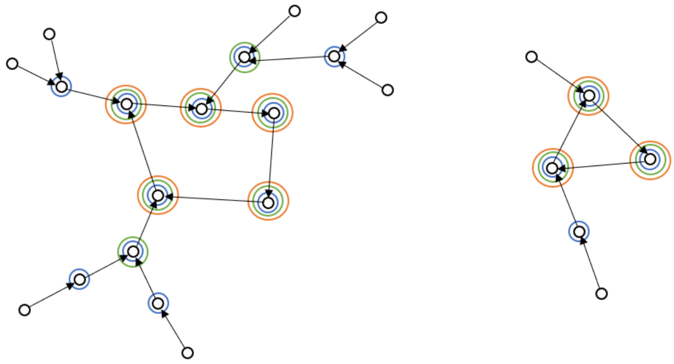
$$\sum_{k+\ell=t} A_k B_\ell \in T^{n\times n}.$$

Ebből kiolvasható, hogy

(6) $\qquad$ az $f \cdot g$ polinomban x^t együtthatójának (p, q)-adik eleme $\displaystyle\sum_{k+\ell=t} \sum_r a_{pr,k}b_{rq,\ell}$.

Összehasonlítva az (5) és (6) megfigyeléseket, látjuk, hogy az $A \cdot B$ polinommátrix „ugyanaz”, mint az $f \cdot g$ mátrixpolinom. Ezzel bebizonyítottuk a 4.41. Lemmát.

4.41. Lemma. Legyen $A, B \in T[x]^{n\times n}$ polinommátrixok, és legyenek $f, g \in T^{n\times n}[x]$ a nekik megfelelő mátrixpolinomok. Ekkor az AB polinommátrixnak az fg mátrixpolinom felel meg.



2. ábra. Véges halmaz transzformációjának gráfja

F.2. Lineáris transzformáció iterálása

Néhány általános megfigyelést teszünk lineáris transzformációk iteráltjairól. (Ebben a fejezetben nem fontos, hogy milyen test felett dolgozunk.)

Lineáris transzformáció helyett először tekintsünk egy akármilyen $\eta : H \to H$ leképezést egy tetszőleges H véges halmazon. Ha egy $h \in H$ elemre ismételten alkalmazzuk az η leképezést, akkor a $h, h\eta, h\eta^2, h\eta^3, \ldots$ sorozatot kapjuk. Mivel H véges halmaz, ebben a sorozatban előbb-utóbb ismétlődni fognak az elemek, azaz a sorozat valahonnan kezdve periodikus lesz. A leképezés gráfjában ez azt jelenti, hogy bármelyik pontból is indulunk ki, a nyilakat követve előbb-utóbb egy ciklusba (irányított körbe) jutunk. Ha η bijektív (vagyis permutáció), akkor a gráf csak ezekből a ciklusokból áll. Általános esetben viszont minden összefüggő komponens egy ciklusból és a ciklusokba futó (és akár egymással is összefutó) utakból áll.

A 2. ábrán egy olyan példa látható, amelyben két ciklus (és így két összefüggő komponens) van. A kékkel bekarikázott pontok alkotják η értékkészletét, a zölddel (is) bekarikázott pontok alkotják η^2 értékkészletét, a narancssárgával (is) bekarikázott pontok pedig η^3 értékkészletét alkotják. Megfigyelhető, hogy $\mathrm{Im}(\eta) \supset \mathrm{Im}(\eta^2) \supset \mathrm{Im}(\eta^3) = \mathrm{Im}(\eta^4) = \mathrm{Im}(\eta^5) = \cdots$ (a ciklusokon lévő pontok halmaza éppen $\mathrm{Im}(\eta^3)$). Nem nehéz belátni, hogy véges halmaz tetszőleges transzformációjánál hasonló a helyzet: a transzformációt iterálva az értékkészlet egy darabig szűkül, majd stabilizálódik.

Tekintsük most már egy véges dimenziós W vektortér η lineáris transzformációját. Ugyan itt W már végtelen halmaz (hacsak nem véges test feletti vektortérről van szó), mivel W dimenziója véges, és η nem akármilyen leképezés, hanem lineáris transzformáció, a véges halmazoknál megfigyeltek egy része érvényben marad, sőt, az értékkészletek mellett a magokra is hasonló állítás fogalmazható meg. (Emlékeztető: a skatulya-elvnek is van lineáris algebrai megfelelője. Mi az?)

F.1. Lemma. Ha $\eta : W \to W$ lineáris transzformációja a W véges dimenziós vektortérnek, akkor létezik olyan p természetes szám, amelyre

$$\begin{aligned} \mathrm{Im}(\eta) \supset \mathrm{Im}(\eta^2) \supset \cdots \supset \mathrm{Im}(\eta^p) &= \mathrm{Im}(\eta^{p+1}) = \mathrm{Im}(\eta^{p+2}) = \cdots, \\ \mathrm{Ker}(\eta) \subset \mathrm{Ker}(\eta^2) \subset \cdots \subset \mathrm{Ker}(\eta^p) &= \mathrm{Ker}(\eta^{p+1}) = \mathrm{Ker}(\eta^{p+2}) = \cdots. \end{aligned}$$

Bizonyítás. Az világos, hogy $\mathrm{Im}(\eta^i) \supseteq \mathrm{Im}(\eta^{i+1})$. Mivel W véges dimenziós, az $\mathrm{Im}(\eta^i)$ sorozat egy idő után nem lehet szigorúan csökkenő, ezért létezik olyan p, amelyre $\mathrm{Im}(\eta^p) = \mathrm{Im}(\eta^{p+1})$. Bebizonyítjuk, hogy ekkor $\mathrm{Im}(\eta^p)$-től kezdve a sorozat stabilizálódik, azaz $\mathrm{Im}(\eta^{p+m-1}) = \mathrm{Im}(\eta^{p+m})$ minden m természetes számra. Elég azt igazolni, hogy $\mathrm{Im}(\eta^{p+m-1}) \subseteq \mathrm{Im}(\eta^{p+m})$, hiszen a másik irányú tartalmazást már láttuk.
Teljes indukcióval bizonyítunk. Az $m = 1$ eset azonnal következik a p szám definíciójából. Az indukciós lépéshez tegyük fel, hogy $\mathrm{Im}(\eta^{p+m-1}) \subseteq \mathrm{Im}(\eta^{p+m})$, ez az indukciós hipotézis; azt kell ebből levezetnünk, hogy $\mathrm{Im}(\eta^{p+m}) \subseteq \mathrm{Im}(\eta^{p+m+1})$. Tekintsünk egy tetszőleges $\mathbf{v} \in \mathrm{Im}(\eta^{p+m})$ vektort, ekkor $\mathbf{v}$ előáll $\mathbf{v} = \mathbf{u}\eta^{p+m} = (\mathbf{u}\eta^{p+m-1})\eta$ alakban alkalmas $\mathbf{u} \in W$ vektorral. Itt az $\mathbf{u}\eta^{p+m-1}$ vektor benne van az $\mathrm{Im}(\eta^{p+m-1})$ altérben, így az indukciós hipotézis szerint eleme az $\mathrm{Im}(\eta^{p+m})$ altérnek is, tehát létezik olyan $\mathbf{w} \in W$ vektor, amelyre $\mathbf{u}\eta^{p+m-1} = \mathbf{w}\eta^{p+m}$.
Következésképp

$$\mathbf{v} = \mathbf{u}\eta^{p+m} = (\mathbf{u}\eta^{p+m-1})\eta = (\mathbf{w}\eta^{p+m})\eta = \mathbf{w}\eta^{p+m+1} \in \mathrm{Im}(\eta^{p+m+1}).$$

Ezzel beláttuk, hogy $\mathrm{Im}(\eta^{p+m}) \subseteq \mathrm{Im}(\eta^{p+m+1})$.
A magokra vonatkozóan az világos, hogy $\mathrm{Ker}(\eta^i) \subseteq \mathrm{Ker}(\eta^{i+1})$. A lineáris leképezések dimenziótétele (3.15. Tétel) szerint $\dim \mathrm{Im}(\eta^i) + \dim \mathrm{Ker}(\eta^i) = \dim W$, ami egy i-től független konstans. Tehát, ha az η^i-ről η^{i+1}-re való áttérésnél a képtér dimenziója csökken, akkor a magtér dimenziója szükségképpen növekszik, mégpedig ugyanannyival. Ezért a magterek sorozata is stabilizálódni fog, mégpedig ugyanakkor, amikor a képterek sorozata stabilizálódik. (A magterekre vonatkozó állítást is lehetett volna képterekhez hasonlóan teljes indukcióval bizonyítani (gyakorlásképp érdemes megcsinálni!). De az egyrészt hosszabb lett volna, másrészt nem derült volna ki belőle, hogy a két sorozat pontosan ugyanott stabilizálódik.) □

F.3. Jordan-bázis létezésének bizonyítása

Ebben a fejezetben az 5.13. Tételt bizonyítjuk be. Legyen tehát $\psi\colon W \to W$ nilpotens lineáris tranaszformációja a W véges dimenziós vektortérnek. Legyen m a legkisebb természetes szám, amelyre $\psi^m = \underline{\underline{\mathbf{0}}}$, ekkor persze $\psi^{m-1} \neq \underline{\underline{\mathbf{0}}}$. Célunk diszjunkt Jordan-láncokból álló bázist (röviden: Jordan-bázist) konstruálni a W vektortérben. Gyakran fogjuk használni ψ magját és ψ hatványainak értékkészleteit, ezért ezekre külön jelöléseket vezetünk be: legyen $K = \mathrm{Ker}(\psi)$ és legyen $I_k = \mathrm{Im}(\psi^k)$. Technikai okokból célszerű lesz a $k = 0$ esetet is megengedni: $I_0 = \mathrm{Im}(\psi^0) = W$, hiszen $\psi^0 = \mathrm{id}_W$.

Az 5.9. Lemma szerint az I_k $(k = 0, 1, \ldots, m)$ alterek szigorúan csökkenő sorozatot alkotnak:

$$W = I_0 \supset I_1 \supset I_2 \supset \cdots \supset I_{m-2} \supset I_{m-1} \supset I_m = \{\mathbf{0}\}.$$

Mindegyik alteret K-val metszve, továbbra is monoton (de nem biztos, hogy szigorúan) csökkenő sorozatot kapunk:

$$K = K \cap I_0 \supseteq K \cap I_1 \supseteq K \cap I_2 \supseteq \cdots \supseteq K \cap I_{m-2} \supseteq K \cap I_{m-1} \supseteq K \cap I_m = \{\mathbf{0}\}. \tag{7}$$

A Jordan-bázis konstrukciójának első lépéseként K-ban adunk meg egy bázist, mégpedig úgy, hogy ennek megfelelő részhalmazai bázisokat alkossanak a $K \cap I_k$ alterekben. Először tekintsük az altérsorozat legkisebb nemtriviális tagját[3], a $K \cap I_{m-1}$ alteret. Legyen $d_{m-1} = \dim(K \cap I_{m-1})$, és legyen $\mathbf{v}_1, \ldots, \mathbf{v}_{d_{m-1}}$ egy tetszőleges bázis ebben az altérben. Ez egy lineárisan független vektorrendszer a $K \cap I_{m-2}$ altérben (hiszen $K \cap I_{m-1} \subseteq K \cap I_{m-2}$), így kibővíthető $K \cap I_{m-2}$ bázisává. Legyen ez a bázis $\mathbf{v}_1, \ldots, \mathbf{v}_{d_{m-1}},\ \mathbf{v}_{d_{m-1}+1}, \ldots, \mathbf{v}_{d_{m-1}+d_{m-2}}$, ahol $d_{m-2} = \dim(K \cap I_{m-2}) - \dim(K \cap I_{m-1})$. Ez a lineárisan független vektorrendszer kibővíthető $K \cap I_{m-3}$ bázisává, ami $K \cap I_{m-4}$ bázisává bővíthető, és így tovább. Az eljárás végén kapunk egy bázist a $K = K \cap I_0$ altérben:

$$\mathbf{v}_1, \ldots, \mathbf{v}_{d_{m-1}},\ \mathbf{v}_{d_{m-1}+1}, \ldots, \mathbf{v}_{d_{m-1}+d_{m-2}},\ \ldots,\ \mathbf{v}_{d_{m-1}+d_{m-2}+\cdots+d_1+1}, \ldots, \mathbf{v}_{d_{m-1}+d_{m-2}+\cdots+d_0}. \tag{8}$$

Itt $d_k = \dim(K \cap I_k) - \dim(K \cap I_{k+1})$, és $\mathbf{v}_1, \ldots, \mathbf{v}_{d_{m-1}+d_{m-2}+\cdots+d_k}$ bázisa a $K \cap I_k$ altérnek. Később még szükségünk lesz ezekre a dimenziókra, ezért rögzítsük, hogy

$$\dim(K \cap I_k) = d_{m-1} + d_{m-2} + \cdots + d_k. \tag{9}$$

A konstrukció következő lépése az, hogy Jordan-láncokat „növesztünk” K fenti bázisvektoraiból. Egy adott $\mathbf{v} \in K$ vektorból úgy növesztünk Jordan-láncot, hogy ψ gráfjában $\mathbf{v}$-ből indulva a nyilak mentén visszafelé lépkedünk (azaz ősképeket veszünk), ameddig csak lehetséges. Egy elemnek több ősképe is lehet, és nem mindegy, melyiket választjuk, mert ettől függhet, hogy hány lépés után akadunk el, vagyis milyen hosszú Jordan-láncot tudunk növeszteni. (Figyeljük ezt meg az 1. ábrán!) Amennyiben a (7) altérsorozatban $K \cap I_k$ a legutolsó tag, amely még tartalmazza $\mathbf{v}$-t (vagyis $\mathbf{v}$ az ábrán a belső körben van és pontosan k-szor van bekarikázva), akkor van olyan $\mathbf{u} \in W$ vektor, amelyre $\mathbf{v} = \mathbf{u}\psi^k$. Ekkor a $\mathbf{v}$-ből növeszthető leghosszabb Jordan-lánc hossza $k+1$ lesz, és egy ilyen maximális hosszúságú Jordan-lánc például a következő:

$$\mathbf{u} \xrightarrow{\psi} \mathbf{u}\psi \xrightarrow{\psi} \mathbf{u}\psi^2 \xrightarrow{\psi} \cdots \xrightarrow{\psi} \mathbf{u}\psi^{k-2} \xrightarrow{\psi} \mathbf{u}\psi^{k-1} \xrightarrow{\psi} \mathbf{u}\psi^k = \mathbf{v} \xrightarrow{\psi} 0.$$

(Halványan jeleztük, hogy $\mathbf{v}\psi = \mathbf{0}$, hiszen $\mathbf{v} \in K$, de ne feledjük, hogy a nullvektor már nem tagja a Jordan-láncnak!) A későbbiek kedvéért vezessük be a $\mathbf{v}^{(j)}$ jelölést a Jordan-lánc tagjaira:

$$\mathbf{v}^{(k+1)} \xrightarrow{\psi} \mathbf{v}^{(k)} \xrightarrow{\psi} \mathbf{v}^{(k-1)} \xrightarrow{\psi} \cdots \xrightarrow{\psi} \mathbf{v}^{(3)} \xrightarrow{\psi} \mathbf{v}^{(2)} \xrightarrow{\psi} \mathbf{v}^{(1)} = \mathbf{v} \xrightarrow{\psi} 0.$$

A (8)-beli vektorrendszer első d_{m-1} tagjából a fentiek szerint m hosszúságú Jordan-lánc növeszthető, a következő d_{m-2} vektorból $m-1$ hosszúságú Jordan-lánc növeszthető, …, végül az utolsó d_0 tagból már csak 1 hosszúságú Jordan-lánc növeszthető. Így festenek ezek a láncok $m = 4$ esetén (az alsó index mindig azt mutatja, hogy melyik

[3] Pontosabban az utolsó olyan tagját, ami nem biztos, hogy triviális. Ha esetleg $K \cap I_{m-1} = \{\mathbf{0}\}$, akkor $d_{m-1} = 0$, és ekkor üres vektorrendszer lesz a bázis.

vektorból növesztettük a láncot, a felső pedig azt, hogy a lánc hányadik tagjáról van szó):

$$\begin{array}{cccccccc}
\mathbf{v}_1^{(4)} & \xrightarrow{\psi} & \mathbf{v}_1^{(3)} & \xrightarrow{\psi} & \mathbf{v}_1^{(2)} & \xrightarrow{\psi} & \mathbf{v}_1^{(1)} & = \mathbf{v}_1 \\
\vdots & & \vdots & & \vdots & & \vdots & \\
\mathbf{v}_{d_3}^{(4)} & \xrightarrow{\psi} & \mathbf{v}_{d_3}^{(3)} & \xrightarrow{\psi} & \mathbf{v}_{d_3}^{(2)} & \xrightarrow{\psi} & \mathbf{v}_{d_3}^{(1)} & = \mathbf{v}_{d_3} \\
 & & & & & & & \\
 & & \mathbf{v}_{d_3+1}^{(3)} & \xrightarrow{\psi} & \mathbf{v}_{d_3+1}^{(2)} & \xrightarrow{\psi} & \mathbf{v}_{d_3+1}^{(1)} & = \mathbf{v}_{d_3+1} \\
 & & \vdots & & \vdots & & \vdots & \\
 & & \mathbf{v}_{d_3+d_2}^{(3)} & \xrightarrow{\psi} & \mathbf{v}_{d_3+d_2}^{(2)} & \xrightarrow{\psi} & \mathbf{v}_{d_3+d_2}^{(1)} & = \mathbf{v}_{d_3+d_2} \\
 & & & & & & & \\
 & & & & \mathbf{v}_{d_3+d_2+1}^{(2)} & \xrightarrow{\psi} & \mathbf{v}_{d_3+d_2+1}^{(1)} & = \mathbf{v}_{d_3+d_2+1} \\
 & & & & \vdots & & \vdots & \\
 & & & & \mathbf{v}_{d_3+d_2+d_1}^{(2)} & \xrightarrow{\psi} & \mathbf{v}_{d_3+d_2+d_1}^{(1)} & = \mathbf{v}_{d_3+d_2+d_1} \\
 & & & & & & & \\
 & & & & & & \mathbf{v}_{d_3+d_2+d_1+1}^{(1)} & = \mathbf{v}_{d_3+d_2+d_1+1} \\
 & & & & & & \vdots & \\
 & & & & & & \mathbf{v}_{d_3+d_2+d_1+d_0}^{(1)} & = \mathbf{v}_{d_3+d_2+d_1+d_0}
\end{array}$$

(Lásd még az 1. ábrát, ahol pirossal vannak jelezve a $\mathbf{v}_i^{(j)}$ vektorok. Ebben a példában $m = 4$ és $d_3 = 2$, $d_2 = 2$, $d_1 = 1$, $d_0 = 2$.) Azt állítjuk, hogy az így kapott $\mathbf{v}_i^{(j)}$ vektorrendszer bázisa W-nek.

A függetlenség bizonyítása nagyon hasonló az 5.12. Tétel első állításának bizonyításához, csak a jelölések bonyolultabbak. Előkészületként figyeljük meg, hogy

$$\mathbf{v}_i^{(j)}\psi^t = \begin{cases} \mathbf{v}_i^{(j-t)}, & \text{ha } t < j; \\ \mathbf{0}, & \text{ha } t \geq j. \end{cases} \tag{10}$$

Tegyük fel, hogy $\sum_{i,j} \lambda_i^{(j)} \mathbf{v}_i^{(j)} = \mathbf{0}$, de nem minden együttható nulla. Vegyünk a nemnulla együtthatók közül egy olyat, amelyben a felső index a lehető legnagyobb. Ha ez az együttható $\lambda_k^{(s)} \neq 0$, akkor $\lambda_i^{(s+1)} = \lambda_i^{(s+2)} = \cdots = 0$ minden i-re. Alkalmazzuk a ψ^{s-1} leképezést a lineáris kombinációra és használjuk a linearitást:

$$\left(\sum_{i,j} \lambda_i^{(j)} \mathbf{v}_i^{(j)} \right) \psi^{s-1} = \sum_{i,j} \lambda_i^{(j)} \mathbf{v}_i^{(j)} \psi^{s-1} = \mathbf{0}.$$

Elhagyhatjuk azokat a tagokat, ahol $j > s$, hiszen ilyenkor a $\lambda_i^{(j)}$ együttható nulla. Másrészt, $j < s$ esetén $\mathbf{v}_i^{(j)}\psi^{s-1} = \mathbf{0}$ a (10) megfigyelés szerint. Ezeket a tagokat is elhagyva, az összegből csak azok a tagok maradnak meg, ahol $j = s$. Mivel (10) szerint $\mathbf{v}_i^{(s)}\psi^{s-1} = \mathbf{v}_i^{(1)} = \mathbf{v}_i$, azt kaptuk, hogy $\sum_i \lambda_i^{(s)} \mathbf{v}_i = \mathbf{0}$. Ez már a K altér (8)-ben felírt bázisvektorainak lineáris kombinációja, ezért minden i-re $\lambda_i^{(s)} = 0$, ami ellentmondás, mert feltettük, hogy van olyan k, amelyre $\lambda_k^{(s)} \neq 0$. Ezzel beláttuk, hogy a $\mathbf{v}_i^{(j)}$ vektorrendszer lineárisan független.

A bizonyítás befejezéséhez elég igazolni, hogy a vektorrendszer pontosan $\dim W$ vektorból áll. A konstrukció alapján látható, hogy d_{m-1} darab m hosszúságú, d_{m-2} darab $m-1$ hosszúságú, $\ldots$, d_1 darab 2 hosszúságú és d_0 darab 1 hosszúságú Jordan-láncunk van, a vektorrendszerünk elemszáma tehát

$$m \cdot d_{m-1} + (m-1) \cdot d_{m-2} + \cdots + 2 \cdot d_1 + 1 \cdot d_0. \tag{11}$$

Megmutatjuk, hogy W dimenziója is éppen ennyi.

Ehhez vizsgáljuk meg a ψ transzformáció I_k-ra való megszorításának kép- és magterét. Egy $\mathbf{v}$ vektor akkor és csak akkor tartozik $\psi|_{I_k}$ magjába, ha $\mathbf{v} \in I_k$ és $\mathbf{v}\psi = \mathbf{0}$, tehát $\operatorname{Ker}(\psi|_{I_k}) = K \cap I_k$. Egy $\mathbf{v}$ vektor akkor és csak akkor tartozik $\psi|_{I_k}$ képterébe, ha $\mathbf{v} \in I_k$ és létezik olyan $\mathbf{u} \in I_k$, amelyre $\mathbf{v} = \mathbf{u}\psi$. Itt az $\mathbf{u} \in I_k$ feltétel azt jelenti, hogy $\mathbf{u} = \mathbf{w}\psi^k$ alkalmas $\mathbf{w} \in W$ vektorral, ezért, $\operatorname{Im}(\psi|_{I_k})$ nem más, mint a $\mathbf{v} = \mathbf{u}\psi = \mathbf{w}\psi^{k+1}$ ($\mathbf{w} \in W$) alakú vektorok halmaza (ezek persze automatikusan I_k-ban vannak). Nyertük, hogy $\operatorname{Im}(\psi|_{I_k}) = I_{k+1}$. Ezek után már fel tudjuk írni a lineáris leképezések dimenziótételét (3.15. Tétel) a $\psi|_{I_k} \colon I_k \to I_k$ transzformációra:

$$\dim \operatorname{Ker}(\psi|_{I_k}) + \dim \operatorname{Im}(\psi|_{I_k}) = \dim(K \cap I_k) + \dim I_{k+1} = \dim I_k.$$

Ezt az összefüggést alkalmazzuk rendre a $k = 0, 1, \ldots, m-1$ értékekre $\dim W$ kiszámításához:

$$\begin{aligned}
\dim W = \dim I_0 &= \dim(K \cap I_0) + \dim I_1 = \\
&= \dim(K \cap I_0) + \dim(K \cap I_1) + \dim I_2 = \\
&= \dim(K \cap I_0) + \dim(K \cap I_1) + \dim(K \cap I_2) + \dim I_3 = \\
&\ \ \vdots \\
&= \dim(K \cap I_0) + \dim(K \cap I_1) + \dim(K \cap I_2) + \cdots + \dim(K \cap I_{m-1}) + \dim I_m.
\end{aligned}$$

Összefoglalva, a következő összefüggést kaptuk W dimenziójára (mivel $I_m = \{\mathbf{0}\}$, a $\dim I_m$ tagot elhagyjuk):

$$\dim W = \dim(K \cap I_0) + \dim(K \cap I_1) + \dim(K \cap I_2) + \cdots + \dim(K \cap I_{m-1}).$$

Mindegyik $\dim(K \cap I_k)$ dimenziót (9) alapján kifejezhetjük a d_i számokkal:

$$\begin{aligned}
\dim W = \quad & d_{m-1} + d_{m-2} + \cdots + d_1 + d_0 + \\
+ \; & d_{m-1} + d_{m-2} + \cdots + d_1 + \\
& \vdots \\
+ \; & d_{m-1} + d_{m-2} + \\
+ \; & d_{m-1}.
\end{aligned}$$

Oszloponként összegezve azt kapjuk, hogy

$$\dim W = m \cdot d_{m-1} + (m-1) \cdot d_{m-2} + \cdots + 2 \cdot d_1 + 1 \cdot d_0,$$

és ez épp ugyanaz, mint (11). Ezzel ellenőriztük, hogy a $\mathbf{v}_i^{(j)}$ vektorrendszer bázisa W-nek, és így bebizonyítottuk az 5.13. Tételt.