

Optimalizálás: Példák

Peter Hajnal

Bolyai Intézet, Szegedi Tudományegyetem, Szeged

2021 tavasz

A tudományág



L.V. Kantorovics
(1912-1986)

- Az optimalizálás a matematika legkülönfélébb területeinek találkozási pontja, ezért például a folytonosság fogalmára épülő analitikus megfontolások és diszkrét matematikai módszerek egyaránt részei az optimalizálási apparátusnak.
- Heterogén jellegét jelzi, hogy a megannyi természettudományos, közgazdasági és informatikai alkalmazás, melyek alapját optimalizálási eredmények képezik.
- Az optimalizálás terültén elért áttörések – a széleskörű alkalmazásnak köszönhetően – komoly tudományos elismeréssel járnak.
- Kantorovics (szovjet matematikus) az optimális erőforrás-allokációval kapcsolatos eredményeiért közgazdasági Nobel-díjat kapott 1975-ben.

Az alapkérdés

- Legyen $c: \text{dom}(c) \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ egy adott n -változós valós függvény ($n \in \mathbb{N}$), melyet **célfüggvénynek** nevezünk.
- A $\text{dom}(c)$ értelmezési tartomány egy általános elemére az $x = (x_1, \dots, x_n)^T$ jelölést használjuk. FIGYELEM! Az előadás során mindvégig oszlopvektorokkal dolgozunk.
- Az optimalizálás alapfeladata a c célfüggvény minimalizálása, előre kiszabott feltételek mellett. Erre a továbbiakban az alábbi rövidített írásmódot használjuk:

Minimalizáljuk	$c(x)$ -et
feltéve, hogy	$x \in \mathcal{F}$,

ahol $x \in \text{dom}(c)$ és $\mathcal{F} \subseteq \mathbb{R}^n$ a feltételek által meghatározott tartomány.

Feltételek

- A kikötések eltérő eredetűek lehetnek: formalizálhatnak matematikai feltételeket vagy származhatnak fizikai, gazdasági kényszerekből
- A feltételek előírására is többféle lehetőség létezik. Itt kettő megadási módot említünk:
- **Implicit feltételek.** Adott $x \in \mathbb{R}^n$ elemről egy algoritmus/orákulum/szubrutin eldönti, hogy teljesíti-e a feltételeket:

$$x \Rightarrow \boxed{\text{ALGORITMUS}} \Rightarrow \text{jó/rossz} (\in \mathcal{F}/\notin \mathcal{F})$$

- **Explicit feltételek.** Véges sok egyenlet és/vagy egyenlőtlenség írja le a feltételt:

$$\begin{cases} f_i(x) \leq 0, & i \in [k] := \{1, 2, \dots, k\}, \\ g_j(x) = 0, & j \in [\ell], \end{cases} \quad (1.1)$$

ahol tehát f_i ($i \in [k]$) és g_j ($j \in [\ell]$) szintén n -változós valós függvények. Ekkor \mathcal{F} a fenti rendszer megoldáshalmaza.

Értelmezési tartomány, Lehetséges megoldások

- Az **optimalizálási feladat értelmezési tartománya** a célfüggvény és a feltételek értelmezési tartományának metszete. Ez az (1.1) explicit feltételek esetén a

$$\mathcal{D} := \text{dom}(c) \cap \left(\bigcap_{i=1}^k \text{dom}(f_i) \right) \cap \left(\bigcap_{j=1}^{\ell} \text{dom}(g_j) \right)$$

halmaz, amely tehát azon x -eket tartalmazza, amelyekre mind a célfüggvény, mind pedig a feltételek értelmezettek.

- **Lehetséges/megengedett megoldásoknak** azon $x \in \mathcal{D}$ vektorokat nevezzük, amelyek eleget tesznek a kritériumoknak is, azaz $x \in \mathcal{F}$. Ezen x -ek halmazát \mathcal{L} jelöli. Tehát

$$\mathcal{L} := \mathcal{D} \cap \mathcal{F},$$

amely (1.1) explicit feltételek esetén

$$\mathcal{L} = \{x \in \mathcal{D} : f_i(x) \leq 0, g_j(x) = 0, i \in [k], j \in [\ell]\}.$$

Optimális érték, optimális hely

- A feladathoz tartozó **optimális érték**

$$p^* = \inf_{x \in \mathcal{L}} c(x) \in \mathbb{R} \cup \{-\infty\} \cup \{\infty\},$$

ahol az infimum ∞ , ha a lehetséges megoldások halmaza üres és $-\infty$, ha a c célfüggvény a lehetséges megoldásokon tetszőlegesen kicsi értéket is felvesz.

- Egy $x^* \in \mathcal{L}$ vektor **optimális hely**, ha ott a célfüggvény felveszi az optimális értéket

$$c(x^*) = p^*.$$

- Az x_ℓ vektor **lokális optimum**, ha annak egy környezetének minden pontjában c legalább akkora, mint az x_ℓ helyen:

$$\exists \varepsilon > 0, \forall x : \|x - x_\ell\|_2 < \varepsilon \quad \text{esetén} \quad c(x_\ell) \leq c(x),$$

ahol $\|\cdot\|_2$ az \mathbb{R}^n -en értelmezett euklidészi norma,

$$\|\cdot\|_2 : \mathbb{R}^n \rightarrow [0, \infty), y \mapsto \|y\|_2 := \sqrt{y_1^2 + \cdots + y_n^2}.$$

Közelítő megoldások

- Azt mondjuk, hogy x_0 egy **ε -közelítő megoldás** ($\varepsilon > 0$), ha

$$c(x_0) \leq p^* + \varepsilon.$$

- Egy x_0 vektor **ε -approximációs megoldás** ($\varepsilon > 0$), ha

$$(0 < p^* \leq) c(x_0) \leq (1 + \varepsilon)p^*.$$

- Az ε -approximációs megoldás fenti definíciójába beleértjük, hogy az optimális érték pozitív, $p^* > 0$. A negatív optimális értékű feladatokban az $c(x_0) \leq (1 - \varepsilon)p^*$ feltételt kell tennünk.

Példa I

Minimalizáljuk	$\frac{1}{x}$ -et
feltéve, hogy	$x \geq 0$.

- A célfüggvény a $c(x) = x^{-1}$ lineáris törtfüggvény, melynek értelmezési tartománya $\text{dom}(c) = \mathbb{R} \setminus \{0\}$. Az explicit feltételt egyetlen egyenlőtlenség adja, nevezetesen $f_1(x) = -x \leq 0$, ezért $k = 1$, $\ell = 0$. Mivel $\text{dom}(f_1) = \mathbb{R}$, így a feladat értelmezési tartománya

$$\mathcal{D} = \mathbb{R} \setminus \{0\}.$$

Ebből látható, hogy a lehetséges megoldások halmaza az

$$\mathcal{L} = \mathbb{R}_{>0} := (0, \infty) =]0, \infty[$$

nyílt intervallum.

Példa I (folytatás)

- Mivel az $x \in \mathcal{L}$ megengedett megoldásokon felvett függvényértékek infimuma zéró, ezért

$$p^* = 0.$$

- Azonban nem létezik olyan $x \in \mathcal{L}$, amelyen c felveszi ezt az értéket, tehát optimális hely nincs.
- Megjegyezzük, hogy bármely $\varepsilon > 0$ esetén az $x_0 \geq \varepsilon^{-1}$ számok ε -közelítő megoldások.
- Viszont ε -approximáló megoldások nem léteznek.

Példa II

Minimalizáljuk	$x \log x$ -et
----------------	----------------

- Most nincsenek feltételek kiszabva, azaz $\mathcal{F} = \mathbb{R}$.
- Ilyenkor azt mondjuk, hogy egy *globális optimalizálási problémát* tekintünk.
- A célfüggvény differenciálható értelmezési tartományán. Az optimalizálás a kalkulus kurzus standard része: A $c(x) = x \log x$ egyváltozós célfüggvény értelmezési tartománya $\text{dom}(c) = \mathbb{R}_{>0}$. Mivel feltétel nincs, ezért

$$\mathcal{L} = \mathcal{D} \cap \mathcal{F} = \mathcal{D} = \text{dom}(c) = \mathbb{R}_{>0}.$$

Az optimális értéket például elemi függvénydiskussziót elvégezve állapíthatjuk meg.

- Eredményül kapjuk, hogy

$$p^* = -\frac{1}{e}, \quad x^* = \frac{1}{e}.$$

Alakítsuk át optimalizálási problémánkat

- Egy megértett optimalizálási probléma esetén több lehetőség van annak formalizálására.
- Másképpen fogalmazva egy formalizált optimalizálási probléma gyakran átfogalmazható úgy, hogy ugyanazt a problémát írja le.
- Ezek az átírások az eredetivel ekvivalens problémához vezethetnek. Azonban ugyanazon probléma kissé eltérő alakjai közt lényeges különbség lehet.
- Ekvivalens átalakítás alatt olyan formális átlalakítást értünk, hogy bármelyik feladat optimális értéke/helye a másik feladat optimális értéke/helye alapján könnyen megadható.
- Most (a teljesség igénye nélkül) néhány lehetőségét megemlítünk.

Átalakítások: Min/max csere

Maximalizáljuk	$c(x)$ -et	
feltéve, hogy	$x \in \mathcal{F}$	\equiv

Minimalizáljuk	$-c(x)$ -et
feltéve, hogy	$x \in \mathcal{F}$

- A minimalizálási probléma egy optimális pontja egyben a maximalizálási problémának is optimális pontja.
- Ha a minimalizálási probléma optimális értékét ismerjük, akkor annak ellentettje lesz a maximalizálási probléma optimális értéke.

Átalakítások: A feltételek ekvivalens átírása

- A középiskolában már láttunk formulák/egyenlőtlenségek/egyenlőségek ekvivalens átalakításait.
- A feltételeinknél ezeket természetesen felhasználhatjuk.

$$\frac{x_1}{x_2^2 + 1} \leq 0 \iff x_1 \leq 0.$$

$$(x_1 + x_2)^2 \leq 0 \iff (x_1 + x_2)^2 = 0 \iff x_1 + x_2 = 0.$$

Átalakítások: egyenlőtlenségek helyettesítése előjelfeltételekkel:

Minimalizáljuk	$c(x)$ -et
feltéve, hogy	$f_i(x) \leq 0$
	$g_i(x) = 0$

 \equiv

Minimalizáljuk	$c(x)$ -et
feltéve, hogy	$f_i(x) + s_i = 0$
	$g_i(x) = 0$
	$s_i \geq 0$

- A bevezetett s_i változók neve „slack változók”.

Átalakítások: A lineáris egyenlőségek kiküszöbölése

- Az $Ax = b$ lineáris egyenlőségrendszer megoldása alap algebra tananyag. Az általános megoldás $x = x_0 + Fy$ alakban írható, ahol x_0 egy tetszőleges megoldás, míg F oszlopai az $Ax = 0$ egyenletrendszer által leírt lineáris altér egy generáló rendszere. Ha F -nek s oszlopa van, akkor $y \in \mathbb{R}^s$.
- x_0 és F meghatározása hatékonyan megtehető.

Minimalizáljuk	$c(x)$ -et	
feltéve, hogy	$f_i(x) \leq 0$	
	$Ax = b$	\equiv

Minimalizáljuk	$c(x_0 + Fy)$ -et
feltéve, hogy	$f_i(x_0 + Fy) \leq 0$

Átalakítások: A célfüggvény helyettesítése egy monoton függvénybe

- Legyen $m : \mathbb{R} \rightarrow \mathbb{R}$ egy szigorúan monoton függvény range c -n (a célfüggvény által felvett értékek halmazán).
- Ekkor

Minimalizáljuk	$c(x)$ -et
feltéve, hogy	$x \in \mathcal{F}$

 \equiv

Minimalizáljuk	$m(c(x))$ -et
feltéve, hogy	$x \in \mathcal{F}$

Átalakítások: A célfüggvény helyettesítése egy monoton függvénybe: Példa

Példa

Minimalizáljuk	$\ x\ _2$ -et	
feltéve, hogy	$x \in \mathcal{F}$	\equiv

Minimalizáljuk	$\ x\ _2^2$ -et
feltéve, hogy	$x \in \mathcal{F}$

- Fenti esetben $\text{range } c = \mathbb{R}_{\geq 0}$, ahol az $m(x) = x^2$ függvény monoton.
- Minimális változtatást eszközöltünk, de az új célfüggvény differenciálható. Látni fogjuk, hogy egy ilyen „kis” előny nagyon jelentős lehet.

Átalakítások: A célfüggvény helyettesítése egy monoton függvénybe: Egy összetettebb példa

$$\begin{aligned} \text{Maximalizáljuk} \quad & x_1x_2 + x_2x_3 + x_3x_1\text{-et.} \\ \text{feltéve, hogy} \quad & x_1 + x_2 + x_3 = 100, \\ & x_1, x_2, x_3 \geq 0. \end{aligned}$$

 \equiv

$$\begin{aligned} \text{Minimalizáljuk} \quad & \sqrt{\frac{x_1^2 + x_2^2 + x_3^2}{3}}\text{-et.} \\ \text{feltéve, hogy} \quad & \frac{x_1 + x_2 + x_3}{3} = \frac{100}{3}, \\ & x_1, x_2, x_3 \geq 0. \end{aligned}$$

- A két feltételrendszer nyilván ekvivalens.

Átalakítások: A célfüggvény helyettesítése egy monoton függvénybe: Egy összetettebb példa (folytatás)

- A két célfüggvény: $c_1(x_1, x_2, x_3) = x_1x_2 + x_2x_3 + x_3x_1$, illetve $c_2(x_1, x_2, x_3) = \sqrt{\frac{x_1^2 + x_2^2 + x_3^2}{3}}$.

- A kapcsolat nyilvánvaló (a feltételek teljesülése esetén):

$$3c_2^2 + 2c_1 = (x_1 + x_2 + x_3)^2 = 100^2.$$

- c_1 maximalizálása helyett áttérhetünk $100^2 - 2c_1 = 3c_2^2$ minimalizálására. Majd itt alkalmazhatjuk az $m(x) = \sqrt{\frac{x}{3}}$ szigorúan monoton függvényt (az új célfüggvény által felvett nemnegatív értékek halmazán) az aktuális célfüggvényre. Így a fenti második alakot kapjuk.

- Az új alak előnye nyilvánvaló. A négyzetes közeget kell minimalizálni adott számtani közép érték esetén. A számtani és négyzetes közép közötti egyenlőtlenség alapján elemi matematikával megválaszolható az optimalizálási kérdés.

Szünet



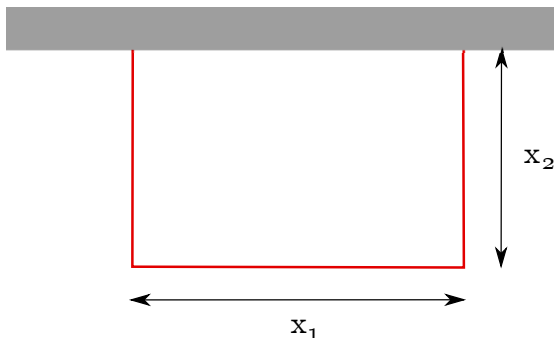
Szöveges példák vs az alapproblémánk

- Az előadássorozatban többször látni fogunk olyan feladatot, amely formalizálása okozza a legtöbb problémát.
- A gyakorlatban nem formalizált problémákat kapunk. Beszélünk kell az alkalmazóval. Meg kell értenünk nyelvezetét, gondolkozását. Eseteg tanulnunk kell fizikát, kémiát, biológiát ahhoz hogy egyenlőségeket, egyenlőtlenségeket lássunk magunk előtt.
- Gyakran a formális leíráshoz szükségesek a matematikai ötletek.
- Gyakran egy jó, szerencsés formalizálás után az optimalizálási algoritmus már készen vár.
- Az alábbiakban néhány bevezető példát mutatunk az alapfogalmakra, illetve elemi formalizálási trükkökre.

Példa I

Példa

Tegyük fel, hogy a lehető legnagyobb téglalap alakú tartományt szeretnék bekeríteni 100 m hosszú kerítéssel úgy, hogy a terület egyik oldalát egy fal képezze.



Példa I (folytatás)

- Ez a probléma a következőképpen írható le optimalizálási feladatként:

Maximalizáljuk	$x_1 x_2$ -t
feltéve, hogy	$x_1 + 2x_2 = 100,$
	$x_1, x_2 \geq 0.$

- Nyilvánvaló, hogy a $c(x_1, x_2) = x_1 x_2$ célfüggvény az egész \mathbb{R}^2 síkon értelmezett és
 $f_1(x_1, x_2) = -x_1$, $f_2(x_1, x_2) = -x_2$, $g_1(x_1, x_2) = x_1 + 2x_2 - 100$,
ezért $\mathcal{D} = \mathbb{R}^2$ és $\mathcal{L} = \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$.
- Alkalmazva a számtani és mértani közepek közötti egyenlőtlenséget és a feltételeket figyelembe véve adódik, hogy

$$50 = \frac{x_1 + 2x_2}{2} \geq \sqrt{x_1(2x_2)} \quad (\geq 0).$$

Példa I (folytatás)

- Ebből négyzetre emeléssel kapjuk, hogy

$$2500 \geq 2x_1x_2 = 2c(x_1, x_2),$$

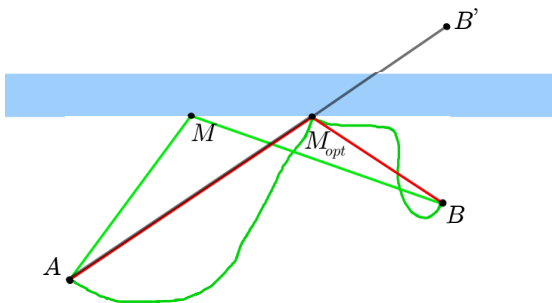
vagyis a célfüggvény felülről korlátos: $c(x_1, x_2) \leq 1250$.

- Mivel $x^* = (50, 25) \in \mathcal{L}$ egy lehetséges megoldás, ahol c eléri ezt a korlátot, így $p^* = 1250$.
- A számtani és mértani közép közötti egyenlőtlenségben az egyenlőség esetének analízise jól ismert. Ez alapján $(50, 25)$ nem csak egy optimális hely. $(50, 25)$ az EGYETLEN optimális hely.

Példa II

Példa

Tegyük fel, hogy egy lovas egy egyenesen haladó folyó egyik oldalán található A pontból az ugyanazon oldalon lévő B pontba szeretne eljutni, de közben a lovat is meg szeretné itatni. Melyik a legrövidebb ilyen út?



Példa II: A megoldás

- A feladattal már általános iskolában találkozhattunk geometria órán.
- Legyen t azon folyópart egyenese, amelyik oldalon lovasunk van.
- Legyen M az itatás pontja. Tetszőleges lehetséges megoldás esetén a lovas pályájának AM szakaszát meghagyva, majd a későbbi szakaszt t -re tükrözve egy olyan pályát kapunk, ami A -ból, B' -be vezet (B' a B pont tükörképe t -re). A módosított pálya hossza a tetszőlegesen választott lehetséges megoldás hossza/költsége.
- A módosított pályák között nyilván az AB' szakasz bejárása az optimális. Azaz t és az AB' szakasz M metszéspontja az „optimális itató hely”.
- Ide A -ból egyenes úton érkezve, majd B -be ugyancsak egyenes mentén haladva lesz legrövidebb a pályája a feladatbeli lovasnak.

Példa II: A formalizálás

- A formalizált optimalizálási feladatot többféleképpen is felírhatjuk.
- Egy lehetőség, ha derékszögű koordinátákat vezetünk be (például a lovas felőli folyópart az x tengely).
- Feladatunk azon $M(x^*, 0)$ pont keresése, amelyre az AM , MB szakaszok hosszának összege minimális.
- Hiszen nyilvánvaló, hogy ha A és M vagy M és B pontok között nem egyenes szakaszon mozog a lovas, akkor pályája nem optimális.
- Tehát adott $A(a_1, a_2)$ és $B(b_1, b_2)$ esetén a feladat az alábbi módon fogalmazható meg:

Minimalizáljuk	$\sqrt{(a_1 - x)^2 + a_2^2} + \sqrt{(b_1 - x)^2 + b_2^2}$ -et.
----------------	--

Példa III

Példa

Az $x_1 + x_2 + x_3 = 100$ egyenletű sík mely pontjának minimális az origótól mért távolsága?

- Vegyük észre, hogy a távolság helyett írhatjuk távolság $\sqrt{\frac{1}{3}}$ -szeresét is.
- A formalizálás (egy korábbi példával megegyezően) lehet a következő:

Minimalizáljuk

$$\sqrt{\frac{x_1^2 + x_2^2 + x_3^2}{3}}\text{-et}$$

feltéve, hogy

$$\frac{x_1 + x_2 + x_3}{3} = \frac{100}{3},$$

$$x_1, x_2, x_3 > 0.$$

Példa III (folytatás)

A számtani és a négyzetes közepek közötti egyenlőtlenséget használjuk fel az eredeti optimális érték becslésére:

$$\sqrt{\frac{1}{3}}c(x) = \sqrt{\frac{x_1^2 + x_2^2 + x_3^2}{3}} \geq \frac{x_1 + x_2 + x_3}{3} = \frac{100}{3}.$$

- Azaz

$$p^* \geq \sqrt{3} \frac{100}{3}.$$

- Továbbá a korlát elérhető, ha $x_1 = x_2 = x_3 = \frac{100}{3}$ (és csak ekkor). Tehát a feladat egyetlen optimális helye

$$x^* = \left(\frac{100}{3}, \frac{100}{3}, \frac{100}{3} \right).$$

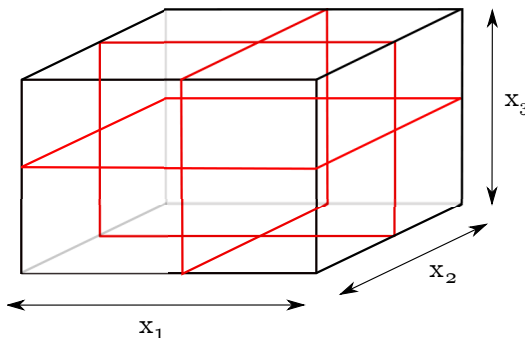
- Továbbá optimális értéke

$$p^* = c(x^*) = \sqrt{3} \frac{100}{3}.$$

Példa IV

Példa

Mekkora a legnagyobb felszínű (téglatest alakú) doboz, amelyet egy 400 cm-es madzaggal át tudunk kötni (a mellékelt ábrának megfelelő módon)?



Példa IV: Formalizálás

A formalizálás nyilvánvaló:

Maximalizáljuk	$2x_1x_2 + 2x_2x_3 + 2x_3x_1$ -et
feltéve, hogy	$4x_1 + 4x_2 + 4x_3 = 400,$
	$x_1, x_2, x_3 > 0.$

A korábbiak alapján ez ekvivalens feladat az előzővel. Az ekvivalencia újbóli meggondolása adja, hogy az optimális hely azonos az előző feladat optimális helyével.

$$x^* = \left(\frac{100}{3}, \frac{100}{3}, \frac{100}{3} \right).$$

Ebből az eredeti feladat optimális értéke

$$p^* = c(x^*) = \frac{20\,000}{3}.$$

Példa V

Példa

A 100 pontú 3-részes egyszerű gráfok közül melyeknek maximális az élszáma?

- Ezzel a feladattal Kombinatorika kurzuson a Turán-tétel kapcsán találkoztunk. Ha a három rész méretét rendre x_1 , x_2 és x_3 jelöli és élünk a természetes észrevétellel, hogy teljes 3-részes gráfok között keressük az optimálist, akkor a feladat a következő:

Maximalizáljuk	$x_1x_2 + x_2x_3 + x_3x_1$ -et
feltéve, hogy	$x_1 + x_2 + x_3 = 100,$
	$x_1, x_2, x_3 > 0,$
	$x_1, x_2, x_3 \in \mathbb{Z}.$

Példa V (folytatás)

- Mivel \mathcal{F} egy nemüres véges halmaz, így a lehetséges megoldások között biztosan létezik x^* optimális hely, $p^* \in \mathbb{R}$ optimális értékkel.
- A feltételből az is látható, hogy ha (x_1, x_2, x_3) egy lehetséges megoldás, akkor például $(x_1 - 1, x_2 + 1, x_3)$ is lehetséges megoldás (feltéve, hogy $x_1 > 1$).
- Ha $x_1 \geq x_2 + 2$, akkor az utóbbi helyen nagyobb a célfüggvény, mint az előbbin.
- Ez azt is jelenti, hogy az optimális helyen $|x_i^* - x_j^*| \leq 1$ minden $i, j \in \{1, 2, 3\}$ esetén.
- Három ilyen lehetséges megoldás van, ahol a célfüggvény közös értéket vesz fel. Így az optimális helyek a $(34, 33, 33)$, $(33, 34, 33)$, $(33, 33, 34)$ pontok.
- Az optimális érték $p^* = 3333$.

Példa V: Megjegyzések

- A fenti feladat NEM ekvivalens a korábbi példáink egyikével sem.
- Ezt jelzi az optimális helyek és a lehetséges megoldások halmazának eltérése, melynek oka a feltételek „lényeges” különbözősége: most kizárólag egész koordinátájú helyek jöhetnek szóba.
- A „folytonos” feladatban (előző két példa) az optimális érték $3333\frac{1}{3}$, nagyobb a mostaninál.
- Ez természetes: a folytonos „versenyben” több „résztevő” van, a maximum értéke legalább annyi mint ahol csak egész koordinátájú „versenyzők” indultak.
- Érdekes, de talán a diszkrét probléma első pillantásra könnyebben tűnik: Véges sok lehetőség közül kell a optimális(ak)at megtalálni. Mégis a diszkrét esethez mintha több ötlet kellene.

Példa VI

Példa

Legyen $d, n \in \mathbb{N}$ és $\ell_1(x), \dots, \ell_n(x): \mathbb{R}^d \rightarrow \mathbb{R}$ adott lineáris függvények. Határozzuk meg a $c(x) = \max_{1 \leq i \leq n} \ell_i(x)$ függvény minimumát.

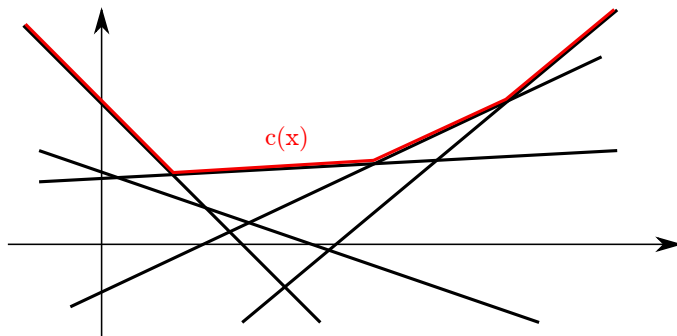
Tehát a formalizált feladat:

Minimalizáljuk	$c(x)$ -et
----------------	------------

A feladat globális (nincsenek feltételek).

Példa VI: Ábra

- A mellékelt ábra a $d = 1$, $n = 5$ eset egy általános konfigurációját szemlélteti.



- Már ez a speciális választás is hű képet ad a célfüggvényről. c lineáris függvények maximuma „szakaszonként” lineáris függvény. Ez $d > 1$ esetén azt jelenti, hogy c értelmezési tartománya felbontható olyan összefüggő részekre, melyeken c lineáris.

Példa VI: Átfogalmazás

- Az előzővel ekvivalens optimalizálási feladatban az m számot minimalizáljuk úgy, hogy az ottani célfüggvény definícióját (maximalitását) beépítjük a feltételekbe.

Minimalizáljuk	m -et
feltéve, hogy	$l_1(x) \leq m,$
	\vdots
	$l_n(x) \leq m,$

ahol $(x, m) \in \mathbb{R}^d \times \mathbb{R}$.

Példa VII: Egy probléma osztály

- Egy optimalizálási feladat ezen formája ismerős lehet az operációkutatás kurzusról.

Lineáris programozás, LP

Legyen $A \in \mathbb{R}^{m \times n}$ adott mátrix, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ rögzített vektorok és $x \in \mathbb{R}^n$ az ismeretlen vektor. A feladat:

Minimalizáljuk	$c^T x$ -et
feltéve, hogy	$Ax \preceq b$,

ahol $Ax \preceq b$ az Ax és b \mathbb{R}^m -beli vektorok „minden komponensben kisebb egyenlő” viszonyát jelöli.

Példa VII: Egy probléma osztály

- Az LP feladatosztályt az Operációkutatás kurzuson részletesen tárgyaljuk.
- Megjegyezzük, hogy az LP feladat nagyon központi probléma. Sok gyakorlatban és elméletben is jónak tartott algoritmus létezik rá.
- Ha egy problémát LP feladatként fogalmazunk meg, akkor a „nehezen túl vagyunk” (akár középiskolában, ha egyenletmegoldás során másodfokú egyenletre redukáltuk munkánkat).
- Valamelyik általános LP algoritmussal fejezzük be munkánkat (ilyenek nyilvános forráskóddal is könnyen elérhetőek).

Példa VIII

Példa

Az LP problémát tekintjük, amelyben a lineáris célfüggvény $(c^T x)$ c vektora valószínűségi változó. Feltesszük, hogy várható értéke, $\bar{c} = \mathbb{E}[c]$ ismert, továbbá a feltételek már nem függenek a véletlentől.

Minimalizáljuk a célfüggvény értékének várható értékét.

- A várható érték linearitása alapján $\mathbb{E}[c^T x] = (\mathbb{E}[c])^T x$.
- Egy LP feladat marad optimalizálási kérdésünk:

Minimalizáljuk	$(\mathbb{E}[c])^T x$
feltéve, hogy	$Ax \preceq b.$

Példa IX: Csebisev középpont

Definíció

A $\mathcal{P} = \{x \in \mathbb{R}^n : a_i^T x \leq b, i = 1, \dots, k\}$ alakú ponthalmazokat poliédereknek nevezzük. Ha a \mathcal{P} poliéder korlátos (amikor is kompakt: korlátos és zárt), akkor politópnak nevezzük.

Példa: Politópok Csebisev-középpontja

Adott \mathcal{P} poliéder. Melyek \mathcal{P} legbelsőbb pontjai?

- Igazából az is központi probléma, hogy \mathcal{P} -ről döntsük el, hogy üres-e.
- Eseteünkben azt kell mérnünk, hogy \mathcal{P} egy pontja milyen mélyen van \mathcal{P} belsejében.
- Sokféle megoldás/válasz van. Mi egy, Csebisev nevéhez fűzött, megoldásról beszélünk.

Példa IX: Csebisev-középpont (folytatás)

Definíció

$B(c, r) = \{x \in \mathbb{R}^n : |x - c|^2 \leq r^2\}$ a c középpontú r sugarú gömb.

A $p \in \mathcal{P}$ pont Csebisev-mélysége

$$M(p) = \sup\{r : B(p, r) \subset \mathcal{P}\}.$$

c a \mathcal{P} politóp egy Csebisev-középpontja, ha

$$M(c) = \sup_{p \in \mathcal{P}} M(p).$$

- Az alapproblémánk: Adott \mathcal{P} esetén keressünk egy Csebisev-középpontot.

Példa IX: Csebisev-középpont (folytatás)

- A feladat első ránézésre nem lineáris. Némi geometriai ismeretre lesz szükségünk.

Definíció

Egy hipersík \mathbb{R}^n -ben: $H = \{x : a^T x = b\}$ alakú ponthalmaz.

r pont pontosan akkor esik H -ra, ha $a^T r - b = 0$.

Egy p pont előjeles távolsága H -tól

$$\frac{a^T p - b}{|a|}$$

Az előjeles távolság abszolútértéke a távolság, előjele a hipersík azon oldalát írja le, ahová pontunk esik.

Kétféle előjeles távolság létezik. A fenti az, amelyik az $\{x : a^T x > b\}$ féltérben pozitív, a komplementer (nyílt) féltérben negatív.

Példa IX: Csebisev középpont (folytatás)

- A Csebisev-középpont problémája ekvivalens a következővel:

Maximalizáljuk	r
feltéve, hogy	$a_i^T x + a_i r \leq b_i, \quad i = 1, \dots, k$
	$r \geq 0$

- Első típusú feltételünk ekvivalens azzal, hogy $\frac{a_i^T}{|a_i|} x + r \leq \frac{b}{|a_i|}$, azaz

$$r \leq \frac{b}{|a_i|} - \frac{a_i^T}{|a_i|} x.$$

- A jobb oldalon egy előjeles távolság szerepel, amit úgy választottunk, hogy az $a_i x < b$ félterekben legyen pozitív.
- Az átfogalmazott optimalizálási feladat egy LP feladat.

Szünet



Példa X: Legkisebb négyzetek problémája

- A következő problémával és megoldási módszereivel a numerikus analízis tárgykörében találkozhattunk.

Példa: A legkisebb négyzetek problémája

A probléma

$$\text{Minimalizáljuk} \quad \|c - Ax\|$$

illetve

$$\text{Minimalizáljuk} \quad \|c - Ax\|^2$$

ahol $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{k \times n}$ valós mátrix és $c \in \mathbb{R}^k$.

- Ez egy feltétel nélküli optimalizálási feladat. Alap lineáris algebrai, geometriai ismeretek alapján egyszerűen kezelhető.

Példa X: Legkisebb négyzetek problémája

Példa

Adott egy mérésorozat (pl. egy kísérleti laboratórium különböző időpontokban vett mérései, vagy egy meteorológiai állomás különböző helyeken mért adatai), ahol t_1, t_2, \dots, t_N és p_1, p_2, \dots, p_N jelöli rendre a mérési időket/helyeket illetve a mért paramétereket. Keresünk egy "alacsony", d -fokú polinomot, ami jó "hipotézis" p változásaira (vagyis jól "illeszthető" a diszkrét idő-mért paraméter grafikonra).

- Legyen

$$p(x) = x_d x^d + \dots + x_1 x + x_0$$

azaz $p = (p_1, \dots, p_N)^T$ mért vektor esetén az

$$x_d (t_1^d, \dots, t_N^d)^T + x_{d-1} (t_1^{d-1}, \dots, t_N^{d-1})^T + \dots$$

vektor L_2 -ben mért p -től vett távolságát kell minimalizálni.

- A legkisebb négyzetek problémájában kvadratikus forma a

Példa XI: Kvadratikus programozás

Kvadratikus programozás, QP

Adottak az $A \in \mathbb{R}^{n \times n}$ szimmetrikus és $C \in \mathbb{R}^{m \times n}$ mátrixok, valamint a $b \in \mathbb{R}^n$, $d \in \mathbb{R}^m$ vektorok.

Minimalizáljuk	$x^T Ax + b^T x$ -et
feltéve, hogy	$Cx \preceq d$.

- A QP alapfeladat feltételrendszere lineáris. A célfüggvény azonban jóval általánosabb.
- Megjegyezzük, hogy a QP alapfeladata is kezelhető (hatékony algoritmus ismert rá) ha a célfüggvény konvex (azaz A pozitív szemidefinit). Ha egy problémát ilyen (konvex QP) alakra hozunk, akkor „készen vagyunk”.

Példa XII: Sztochasztikus lineáris programozás

- Az LP problémát tekintjük, amelyben a lineáris célfüggvény $(c^T x)$ c vektora valószínűségi változó. Feltesszük, hogy várható értéke, $\bar{c} = \mathbb{E}[c]$, kovarianciamátrixa $\Sigma = \mathbb{E}[(c - \bar{c})(c - \bar{c})^T]$ ismert. Továbbá a feltételek már nem függnek a véletlentől.
- Láttuk, ha csak a célfüggvény értékének várható értékét $(\mathbb{E}[c^T x] = (\mathbb{E}[c])^T x)$ minimalizáljuk, akkor egy LP feladathoz jutunk.
- Ekkor azonban a „kockázatot” nem vettük figyelembe.

Példa XII: Sztochasztikus lineáris programozás (folytatás)

- Megszokott a következő változat vizsgálata:

Sztochasztikus LP

Minimalizáljuk	$\mathbb{E}[c^T x] + \gamma \text{Var}[c^T x]$
feltéve, hogy	$Ax \preceq b$
	$Dx = e$

ahol $\gamma \in \mathbb{R}_{>0}$ egy az alkalmazáshoz választott paraméter.

- Egyszerű átlakításokkal

$$\begin{aligned} \text{Var}[c^T x] &= \mathbb{E}[(c^T x - \mathbb{E}[c^T x])^2] = \mathbb{E}[(c^T x - (\mathbb{E}[c])^T x)^2] \\ &= \mathbb{E}[((c^T - \mathbb{E}[c]^T)x)^2] = x^T \mathbb{E}[(c - \mathbb{E}[c])(c^T - \mathbb{E}[c]^T)]x = x^T \Sigma x. \end{aligned}$$

- A kérdés konvex QP alakja már látható.

Példa XIII: Poliéderek távolsága

Példa

Határozzuk meg az n -dimenziós euklidészi tér két adott poliéderjének távolságát!

- Mivel a poliéderek lineáris egyenlőtlenségrendszerek megoldáshalmazaként állnak elő, ezért mindkét politópnak megfelel egy egyenlőtlenségrendszer ($C_1 \in \mathbb{R}^{k_1 \times n}$, $C_2 \in \mathbb{R}^{k_2 \times n}$):

$$\mathcal{P}_1 = \{x_1 \in \mathbb{R}^n : C_1 x_1 \preceq d_1\} \quad \text{és} \quad \mathcal{P}_2 = \{x_2 \in \mathbb{R}^n : C_2 x_2 \preceq d_2\}.$$

- Ezek távolsága

$$d(\mathcal{P}_1, \mathcal{P}_2) = \inf\{d(x_1, x_2) : x_1 \in \mathcal{P}_1, x_2 \in \mathcal{P}_2\},$$

ahol $d(x_1, x_2) = \|x_1 - x_2\|_2$.

Példa XIII: Poliéderek távolsága (folytatás)

- A távolságnégyzetet tekintve ekvivalens optimalizálási feladatot nyerünk:

Minimalizáljuk	$\ x_1 - x_2\ _2^2$ -et
feltéve, hogy	$C_1 x_1 \preceq d_1,$
	$C_2 x_2 \preceq d_2.$

- Ehhez tekintsük a $d_1 \in \mathbb{R}^{k_1}, d_2 \in \mathbb{R}^{k_2}, x_1, x_2 \in \mathbb{R}^n$ vektorokból és $C_1, C_2 \in \mathbb{R}^{(k_1+k_2) \times n}$ mátrixokból képzett

$$d = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^{2n}, \quad C = \begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix} \in \mathbb{R}^{2n \times 2n}$$

vektorokat és mátrixot, ahol tehát 0 a megfelelő méretű nullmátrixokat jelöli. Feltételrendszerünk: $Cx \preceq d$.

Példa XIII: Poliéderek távolsága (folytatás)

- Ha az $n \times n$ -es egységmátrixból képezzük még, az

$$A = \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} \in \mathbb{R}^{2n \times 2n}$$

mátrixot, akkor feladatunk a célfüggvény $x^T A x$ alakba írható.

- A fenti írásmódot használva látjuk, hogy ez egy konvex QP feladat. A fentiek alapján két politóp távolságának meghatározása hatékonyan elvégezhető.

Példa XIV: Klikk probléma

Példa

Adott egy G egyszerű gráf. Határozzuk meg $\omega(G) = \max\{|K| : K \subset V(G) \text{ klikk}\}$ értékét.

- Egy tetszőleges U csúcshalmaz leírható $\{0, 1\}^V \subset \mathbb{R}^V$ -beli χ_U karakterisztikus vektorával. U elemszáma éppen $1^T \chi_U$, ahol $1 \in \mathbb{R}^V$ azon vektor, amely minden komponense 1.
- $\{0, 1\}^V$ elemeit a $0 \leq x_v \leq 1$, $x_v \in \mathbb{Z}$ tetszőleges v csúcsra tett feltételekkel írhatjuk le. Egy 0-1 vektor akkor lesz klikk karakterisztikus vektora, ha tetszőleges össze nem kötött u és v csúcsra legfeljebb egyikük esik bele. Azaz minden $uv \notin E(G)$, $u \neq v$ csúcsok esetén $x_u + x_v \leq 1$.

Példa XIV: Klikk probléma (folytatás)

- A jól ismert \mathcal{NP} -nehéz klikk probléma megfogalmazható a következő módon:

Maximalizáljuk	$1^T x$ -et,
feltéve, hogy	$0 \leq x_v \leq 1, x_v \in \mathbb{Z}$ minden v csúcsra
	$x_u + x_v \leq 1$
	tetszőleges $uv \notin E(G), u \neq v$ csúcsokra.

Példa XV: Egész értékű programozás

- Az előző példa egy általános séma eleme: Ismét „nyúljunk” bele az LP alapfeladatba. Most a feltételrendszert módosítjuk. A feltételek közé tesszük be azt, hogy „versenyző” vektoraink koordinátái legyenek egészek.

Egész értékű programozás, IP

Az alapfeladat:

Minimalizáljuk	$c^T x$ -et
feltéve, hogy	$Ax \preceq b,$
	$x \in \mathbb{Z}^d.$

- LP esetén a lehetséges megoldások halmaza egy poliéder volt. Most egy poliéder és \mathbb{Z}^d metszete (politópok esetén ez egy véges halmaz). Lehet, hogy ez a módosítás „ártatlanabbnak” tűnik, mint a célfüggvény módosítása, mégis a kapott probléma nehéz. Ez az általános forma képes \mathcal{NP} -teljes problémák formalizálására. Általános, hatékony megoldása nem várható.

Példa XVI: Párosítási probléma

Példa

Adott G gráf élei közötti maximális párosítást keressük.

- A feladat formalizálása:

Maximalizáljuk	$1^T x$ -et (ahol $1^T = (1, \dots, 1)$, $x \in \mathbb{R}^{E(G)}$)
feltéve, hogy	x párosítás karakterisztikus vektora.

- A feltételt algebraizálva az alábbi ekvivalens optimalizálási feladatot kapjuk:

Maximalizáljuk	$1^T x$ -et (ahol $x \in \mathbb{R}^{E(G)}$)
feltéve, hogy	$0 \leq x_e \leq 1$, $x_e \in \mathbb{Z}$
	$\sum_{e: v \in e} x_e \leq 1$ minden v csúcsra.

Példa XVI: Párosítási probléma: Relaxáció

- Azaz egy IP feladatot látunk. Az $x_e \in \mathbb{Z}$ feltétel elhagyását **LP relaxációnak** nevezzük. Ezzel a versenyt „megnyitjuk” a nem egész koordinátájú „versenyzők” felé is. Természetesen a maximalizálási feladat optimális értéke megnőhet.
- A relaxált feladat egy kezelhető probléma/LP feladat:

Maximalizáljuk	$1^T x$ -et
feltéve, hogy	$0 \leq x_e \leq 1,$
	$\sum_{e:v \in e} x_e \leq 1$ minden v csúcsra.

Jelölés

$$\nu^*(G)$$

jelölje a fenti LP feladat optimális értékét.

Példa XVI: Párosítási probléma (folytatás)

Tétel

Ha G páros, akkor $\nu^*(G) = \nu(G)$.

Azaz LP relaxációval nyert optimalizálási feladat ekvivalens az eredetivel.

- Problémánk kezdeti alakját ekvivalens módod is átírhatjuk LP alakba. Eredetileg véges sok „versenyzőnk” volt, a párosítások karakterisztikus vektorai. Ezt a versenyzőhalmazt helyettesítve a konvex burkukkal egy ekvivalens problémát kapunk (célfüggvényünk lineáris!):

Maximalizáljuk	$1^T x$
feltéve, hogy	$x \in \text{conv}\{\chi_M : M \text{ párosítás}\}$

Példa XVI: Párosítási probléma (folytatás)

- A feltétel egy politópot ír le. Ez leírható véges sok féltér metszeteként, azaz algebrailag egy véges lineáris egyenlőtlenségrendszer megoldáshalmaza.

Tétel*

Legyen G páros gráf. Ekkor

$$\begin{aligned} \text{conv}\{\chi_M : M \text{ párosítás}\} = \\ = \{x \in \mathbb{R}^E : 0 \leq x_e \leq 1, \sum_{e: v \in e} x_e \leq 1 \text{ minden } v \text{ csúcsra.}\} \end{aligned}$$

- A nem páros gráfok esetét a későbbiekben vizsgáljuk meg jobban.

Példa XIV: Klikk probléma újra

- Végül egy \mathcal{NP} -nehéz problémát formalizálunk/algebraizálunk. Az eredmény alakja egy QP alapfeladat. Természetesen nem konvex feladat lesz.

Példa

Adott egy G gráf. Határozzuk meg a klikk paraméterét ($\omega(G)$ -t), azaz a legnagyobb klikk méretét.

- Láttuk, hogy a kérdés egy IP problémaként is megfogalmazható. Az alábbi tétel egy távolról sem nyilvánvaló, másik formalizálását adja $\omega(G)$ meghatározásának.

Példa XIV: Klikk probléma újra: A Tétel

Tétel

Legyen $A \in \mathbb{R}^{V \times V}$ a G gráf szomszédsági mátrixa és tekintsük az alábbi feladatot ($x, 0, 1 \in \mathbb{R}^V$):

Maximalizáljuk	$x^T A x$ -et
feltéve, hogy	$x \succeq 0,$
	$1^T x = 1.$

Ekkor

$$p^* = 1 - \frac{1}{\omega(G)}.$$

Példa XIV: Klickr probléma újra: A Bizonyítás

$$(1): p^* \geq 1 - \frac{1}{\omega(G)}.$$

- Legyen K egy optimális (maximális elemszámú) klikk. Ekkor az

$$x_K = (0, \dots, 0, \underbrace{\frac{1}{\omega}, \dots, \frac{1}{\omega}}_K, 0, \dots, 0)$$

lehetséges helyen c értéke

$$c(x_K) = x_K^T A x_K = \frac{1}{\omega^2} \omega(\omega - 1) = 1 - \frac{1}{\omega}.$$

Példa XIV: Klikk probléma újra: A Bizonyítás

$$(2): p^* \leq 1 - \frac{1}{\omega(G)}.$$

- Tekintsük az

$$x = \left(\frac{v_1}{N}, \dots, \frac{v_k}{N} \right) \in \mathcal{L} \cap \mathbb{Q}^{V(G)}$$

vektort, ahol tehát $k = |V(G)|$ és az $n_1, \dots, n_k \in \mathbb{N}$ számok az $N \in \mathbb{N}$ szám egy partícióját adják.

- Ezek után a G gráfhoz rendeljünk egy \tilde{G} -mal jelölt gráfot a következő módon:
 - A G gráf v_i pontjának egy n_i pontból álló élmentes V_i független ponthalmaz felel meg \tilde{G} -ban ($i = 1, \dots, k$).
 - Ha a $v_i, v_j \in V(G)$ pontok szomszédosak, akkor a megfelelő V_i, V_j csúcshalmazok közt egy K_{n_i, n_j} teljes páros gráfot teszünk (minden lehetséges keresztélt behúzunk).
 - Ha a $v_i, v_j \in V(G)$ pontokat nem köti össze él, akkor a megfelelő V_i, V_j csúcshalmazok között nincs él.

Példa XIV: Klikk probléma újra: A Bizonyítás

- A definíció alapján \tilde{G} pontjainak száma

$$|V(\tilde{G})| = \sum_{i=1}^k v_i = N \sum_{i=1}^k \frac{n_i}{N} = N$$

- A konstrukció alapján \tilde{G} élszáma

$$|E(\tilde{G})| = \sum_{\{v_i, v_j\} \in E(G)} n_i n_j = \frac{1}{2} N^2 \sum_{v_i, v_j \in E(G)} \frac{n_i}{N} \frac{n_j}{N} = \frac{N^2}{2} x^T A x = \frac{N^2}{2} c(x).$$

- \tilde{G} legnagyobb klikkje $\omega(G)$ méretű. A Turán-tételt alkalmazva élszáma legfeljebb az $\omega(G)$ osztályú/ N pontú Turán-gráf élszáma:

$$|E(\tilde{G})| \leq |E(T_{N, \omega(G)})| \leq \binom{\omega(G)}{2} \left(\frac{N}{\omega(G)} \right)^2 = \left(1 - \frac{1}{\omega(G)} \right) \frac{N^2}{2}.$$

Példa XIV: Klickr probléma újra: A Bizonyítás

- A fentiek összegzése után kapjuk, hogy

$$c(x) = \frac{2}{N^2} |E(\tilde{G})| \leq 1 - \frac{1}{\omega(G)},$$

ha $x \in \mathcal{L} \cap \mathbb{Q}^{V(G)}$.

- $\mathbb{Q}^{V(G)}$ egy sűrű halmaz \mathcal{L} -ben, a célfüggvény folytonos.
- Ez bizonyítja a hiányzó egyenlőtlenséget. Vagyis $p^* = 1 - \frac{1}{\omega(G)}$.

Példa XIV: Klikk probléma újra: Megjegyzések

- A tétel által adott formalizálásban egy kvadratikus alakot kell maximalizálni.
- A konvex QP alapfeladat feltételei közül „csak” a kvadratikus alak mátrixának pozitív szemidefinitisége hiányzik.
- Ezen feltétel elhagyása olyan alakot eredményez, ami képes reménytelen optimalizálási feladatok formalizálására.

Vége van!

Köszönöm a figyelmet!