

# Az optimalizálás alapfogalmai

Hajnal Péter

2021. tavasz

- Az optimalizálás a matematika, fizika, gazdaság (és sok más emberi cselekvés) természetes alapproblémája.
- A klasszikus Fermat-elv azt mondja, hogy a „fény mozgása során optimalizál/minimális idejű utat választva mozog”.
- A modern logisztikai, pénzügyi problémák során optimalizálással óriási összegek takaríthatók meg.
- Az optimalizálás a középiskolai szélsőérték feladatok jobban hangzó neve. Kétféle szélsőérték létezik: minimum, illetve maximum.
- Ebben a pontban megfogalmazzuk az optimalizálás alapfeladatát és bevezetjük az alapvető fogalmakat, elnevezéseket. Most az alapfogalmakat minimalizálási feladatra alapozva mondjuk el. Maximalizálási kérdések teljesen hasonlóan kezelhetők (ha van különbség az természetes, józan ésszel kitalálható).

# Alapfogalmak: Az alapfeladat

- 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.
- Az optimalizálás alapfeladata a  $c$  célfüggvény minimalizálása, előre kiszabott feltételek mellett. Erre a továbbiakban az alábbi rövidített írásmódot használjuk:

Minimalizáljuk	$c(x)$ -et-t
Feltéve, hogy	$x \in \mathcal{F} \cap \text{dom}(c)$ ,

ahol  $\mathcal{F} \subseteq \mathbb{R}^n$  a feltételek által meghatározott tartomány.

- **Megjegyzés:** A feltételek eltérő eredetűek lehetnek. Formalizálhatnak matematikai feltételeket vagy származhatnak fizikai, gazdasági kényszerekből.

# Alapfogalmak: A feltételek

- A feltételek előírására is többféle lehetőség létezik.
- Itt mi egy esetet vizsgálunk: Véges sok egyenlet és/vagy egyenlőtlenség írja le a feltételrendszert.

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

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.

- Az alapfeladat újra leírva a véges sok, explicit formulákkal leírt feltétel esetére:

Minimalizáljuk	$c(x)$ -t
Feltéve, hogy	$f_i(x) \leq 0, \quad i \in [k] := \{1, 2, \dots, k\},$ $g_j(x) = 0, \quad i \in [\ell] := \{1, 2, \dots, \ell\}.$

# Alapfogalmak: Az értelmezési tartomány, lehetséges megoldások halmaza

- Az **optimalizálási feladat értelmezési tartománya** a célfüggvény és a feltételek értelmezési tartományának metszete:

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

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

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

azaz

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

- A feladathoz tartozó **optimális érték**

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

ahol az infimum  $\infty$ , 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

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

- Megjegyezzük, hogy a fenti formalizmus alapján  $p^*$  mindig definiált.  $x^*$  nem mindig definiált és amikor  $x^*$  létezik akkor sem biztos, hogy egyértelmű.

# Alapfogalmak: Relaxált optimumok

- Az  $x_\ell$  vektor **lokális optimum/minimum**, ha annak egy környezetének minden pontjában  $c$  legalább akkora, mint az  $x_\ell$  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,

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

- Azt mondjuk, hogy  $x_0 \in \mathcal{L}$  egy  **$\varepsilon$ -közelítő megoldás** ( $\varepsilon > 0$ ), ha

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

- Egy  $x_0 \in \mathcal{L}$  vektor  **$\varepsilon$ -approximációs megoldás** ( $\varepsilon > 0$ ), ha

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

## Alapfogalmak: Relaxált optimumok (folytatás)

- Az  $\varepsilon$ -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.
- Azt is megjegyezzük, hogy a fenti fogalmak bevezetését egy minimalizálási feladatra alapoztuk. Minden definíciónk (természetes módon megváltoztatva) elmondható maximalizálási feladatra is.



## Példa optimalizálási feladatra: A lineáris programozás

- A lineáris programozás az optimalizálási feladatok egyik (talán a legfontosabb) osztálya. Természetesen itt is két lehetőségünk van: minimalizálhatunk/maximalizálhatunk. Mi most a maximalizálást vesszük alapértelmezett célként.
- Ekkor a célfüggvény egy lineáris függvény. A feltételek pedig lineáris egyenletek és egyenlőtlenégek:

Maximalizáljuk	$c^T x$
Feltéve, hogy	$Ax \preceq b$
	$Dx = e,$

ahol  $x = (x_1, x_2, \dots, x_n)^T$ ,  $c \in \mathbb{R}^n$ ,  $A \in \mathbb{R}^{k \times n}$ ,  $b \in \mathbb{R}^k$ ,  $D \in \mathbb{R}^{\ell \times n}$ ,  $e \in \mathbb{R}^\ell$ .

- A maximalizálandó célfüggvény egy (homogén) lineáris függvény.
- Két fajta feltételünk van. Számuk  $k + \ell$ .
- Az első  $k$  darab feltétel egyenlőtlenség. Mindegyik úgy van rendezve, hogy a konstansok a jobb oldali, nagyobb oldalon legyenek, a homogén lineáris függvények („betűs rész”) pedig a bal oldali, kisebb oldalon.
- A további  $\ell$  darab feltétel egyenlőség. Mindegyik úgy van rendezve, hogy a konstansok a jobb oldalon legyenek, a homogén lineáris függvények („betűs rész”) pedig a bal oldalon.

# LP: Értelmezési tartomány, lehetséges megoldások

- $\mathcal{D} = \mathbb{R}^n$ .  $\mathcal{F}$  egy poliéder és egy affin altér metszete, ami egy poliéder.
- Tudjuk, hogy  $\mathcal{L} (= \mathcal{F})$  felírható

$$\langle u_1, u_2, \dots, u_K \rangle_{\text{conv}} + \langle v_1, v_2, \dots, v_L \rangle_{\text{kúp}}$$

alakban.

- Ha  $x = x' + x''$ , akkor  $c^T x = c^T x' + c^T x''$ . Így könnyű látni, hogy a fenti felírás miatt a következő két optimalizálási problémát kell megoldanunk:

Maximalizáljuk	$c^T x$ -t
Feltéve, hogy	$x \in \langle u_1, u_2, \dots, u_K \rangle_{\text{conv}}$

Maximalizáljuk	$c^T x$ -t
Feltéve, hogy	$x \in \langle v_1, v_2, \dots, v_L \rangle_{\text{kúp}}$

- A politóp feletti lineáris függvény optimalizálása egyszerűnek tűnik. Könnyű ellenőrizni, hogy az egyik  $u_i$ -ben felvevődik az optimum.
- A kép csalóka. Az hogy „csak véges sok” érték közül kell kiválasztani az optimálisat, az nem jelenti azt hogy egyszerű a feladat. Nagyon nagy halmaz lehet a generátor elemek halmaza. Végignézni mindet nem egy hatékony eljárás.
- A kúp feletti optimalizáció is egyszerűnek tűnik. Ha  $c^T v_i$  pozitív valamely  $i$ -re akkor a célfüggvény tetszőleges nagy lehet. Ha mindegyik  $c^T v_i$  nempozitív, akkor a célfüggvény nem vesz fel pozitív értéket. A maximális érték 0 lesz, ami az origóban felvevődik.



- 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 ekvivalensek. Mégis ugyanazon probléma kissé eltérő alakjai közt lényeges különbség lehet.
- Ekvivalens átalakítás alatt olyan formális átalakítást értünk, hogy bármelyik feladat optimális értéke/helye a másik feladat optimális értéke/helye alapján könnyen megadható.
- Most (a teljesség igénye nélkül) néhány lehetőségét megemlítünk.

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.

# A feltételek ekvivalens átírása

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

Példa

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

Példa

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

Példa

$$\begin{cases} -x_1 + x_2 = 1 \\ x_1 - 2x_2 = 3 \end{cases} \iff \begin{cases} -x_1 + x_2 = 1 \\ -x_2 = 4 \end{cases} \iff \begin{cases} -x_1 = 5 \\ -x_2 = 4 \end{cases}$$



# 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 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	$f_i(x_0 + Fy) \leq 0$

# Slack változók, 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 célfüggvény helyettesítése egy monoton függvénybe

Legyen  $m : \mathbb{R} \rightarrow \mathbb{R}$  egy szigorúan monoton növekvő 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	$x \in \mathcal{F}$	

## Példa

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

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

# LP: Poliedrikus normálforma

- Az első normálformában nem engedünk meg egyenlőségeket:

Maximalizáljuk	$c^T x$
Feltéve, hogy	$Ax \preceq b$

- Azaz a lehetséges megoldások halmaza egy poliéder. Ezen halmaz felett maximalizálunk egy lineáris függvényt. Ezt a normálformát **poliedrikus normálformának** nevezem.

## Példa

Maximalizáljuk

$$x_1 + x_2 - 7x_3 - x_4 + 2x_5 - t$$

Feltéve, hogy

$$\begin{cases} -x_1 + x_2 & & -x_4 + 3x_5 \leq 1 \\ x_1 & -5x_3 & + 7x_5 \leq 3 \\ & x_2 + x_3 + x_4 & \leq 2 \end{cases}$$

## LP: Előjeles poliedrikus normálforma

- A második normálforma esetén az egyenlőség feltételek mellőzése mellett minden változóról feltesszük, hogy nemnegatív:

Maximalizáljuk	$c^T x$
Feltéve, hogy	$Ax \preceq b$
	$x \succeq 0$

- Tehát ez is egy speciális poliedrikus forma, ahol  $n$  változó esetén szükségszerűen  $n$  előjelfeltétel szerepel. Geometriailag a lehetséges megoldások poliédere a nemnegatív koordinátájú pontok tér  $2^n$ -ed részében van. Ezt a normálformát **előjeles poliedrikus normálformának** nevezem.

## Példa

Maximalizáljuk	$x_1 + x_2 - t$
Feltéve, hogy	$\begin{cases} -x_1 + x_2 \leq 1 \\ x_1 \leq 3 \\ x_2 \leq 2 \\ x_1, x_2 \geq 0 \end{cases}$

# Szimplex normálforma

- A harmadik normálforma a szimplex algoritmushoz legjobban „illő” normálforma. A változókról feltesszük, hogy nemnegatívak, de most lineáris egyenlőségek a további feltételek:

Maximalizáljuk	$c^T x$
Feltéve, hogy	$Ax = b$
	$x \succeq 0$

Ezt a formát **szimplex normálformának** vagy egyszerűen **szimplexalaknak** nevezem.



## Példa

Maximalizáljuk

$$x_1 + x_2 - t$$

Feltéve, hogy

$$\begin{cases} -x_1 + x_2 + x_3 & = 1 \\ x_1 & + x_4 = 3 \\ x_2 & + x_5 = 2 \end{cases}$$

$$x_1, x_2, x_3, x_4, x_5 \geq 0$$

# Megjegyzések a szimplexalakhhoz

- Feltesszük, hogy  $A$  sorai ( $k$  darab vektor) lineárisan független vektorok.
- Ez nem megszorító feltétel: Az  $Ax = b$  egyenletrendszer Gauss-eliminációval ekvivalens módon átalakítható. Ha  $A$  sorai nem lineárisan függetlenek, akkor az elimináció során valamelyik egyenlet bal oldala lenullázódik. Egyenletrendszerünk vagy ellentmondásos, vagy valamelyik egyenlet elhagyható.
- $A$  sorok függetlensége azt jelenti, hogy  $A \in \mathbb{R}^{k \times n}$ -nak van olyan  $k \times k$  méretű részmátrixa, ami invertálható.
- Sőt, azt is feltesszük, hogy  $A$ -nak  $I_k$  részmátrixa. Ez sem megszorító feltételezés.
- Ha a második poliedrikus formában lévő LP feladatot szimplex formára hozzuk (slack változók segítségével), akkor feltevésünk automatikusan teljesül.

Vége van!

Köszönöm a figyelmet!