

OPERÁCIÓKUTATÁS ELŐADÁS

matematika BSc hallgatók számára

Az optimalizálás alapjai

Előadó: Hajnal Péter

2020.

1. Optimalizálás alapfogalmai

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).

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	$x \in \mathcal{F}$,

ahol $x \in \text{dom}(c)$ és $\mathcal{F} \subseteq \mathbb{R}^n$ a feltételek által meghatározott tartomány. Maguk a kikötések eltérő eredetűek lehetnek: formalizálhatnak matematikai feltételeket vagy származhatnak fizikai, gazdasági kényszerekből. 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

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

Megjegyzés. A gyakorlati alkalmazásokban megjelenő célfüggvények „nagyon sok” változótól függhetnek: n értéke az ezres, sőt a milliós nagyságrendet is elérheti.

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 ∞ , 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.

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

$$\exists \varepsilon > 0, \forall x : \|x - x_\ell\|_2 < \varepsilon \quad \text{esetén}^1 \quad c(x_\ell) \leq c(x).$$

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

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

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

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

Megjegyzés. 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. Azt is megjegyezzük, hogy a fenti fogalmak bevezetését egy minimalizálási feladatra alapoztuk. Minden definíciónk (természetes módon megváltoztatva) elmondható maximalizálási feladatra is.

¹ $\|\cdot\|_2$ az $\mathbb{R}^n$ -en értelmezett euklidészi norma,

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

2. Lineáris programozás (LP)

A lineáris programozás az optimalizálási feladatok egyik (talán a) legfontosabb osztálya. Ekkor a célfüggvény egy lineáris függvény. A feltételek pedig lineáris egyenletek és egyenlőtlenségek:

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

ahol $x = (x_1, x_2, \dots, x_n)^T$, $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{k \times n}$, $b \in \mathbb{R}^k$, $D \in \mathbb{R}^{\ell \times n}$, $e \in \mathbb{R}^\ell$. Azaz k darab egyenlőtlenségünk van, mindegyik úgy rendezve, hogy a konstansok a jobb oldali, nagyobb oldalon legyenek, a homogén lineáris függvények („betűs rész”) pedig a bal oldali, kisebb oldalon. Továbbá van ℓ darab egyenlőség feltételünk, mindegyik úgy rendezve, hogy a konstansok a jobb oldalon legyenek, a homogén lineáris függvények („betűs rész”) pedig a bal oldalon. $\mathcal{D} = \mathbb{R}^n$. $\mathcal{F}$ egy poliéder és egy affin altér metszete.

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

Az hogy „csak véges sok” érték közül kell kiválasztani az optimálisat, az nem jelenti azt hogy egyszerű a feladat.

3. Optimalizálási problémák átfogalmazása

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. Sokféle helyzetben ezek az átírások az eredetivel ekvivalensek. Mászor azonban ugyaznazon 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.

1. Min/max csere:

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

2. A feltételek ekvivalens átírása: A középiskolában már láttunk formulák/egyenlőtlenségek/egyenlőségek ekvivalens átalakításait. A feltételeinknél ezeket természetesen felhasználhatjuk.

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

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

3. Slack változók, egyenlőtlenségek helyettesítése előjelfeltételekkel:

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

4. 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		Minimalizáljuk	$c(x_0 + Fy)$ -et
feltéve, hogy	$f_i(x) \leq 0$		feltéve, hogy	$f_i(x_0 + Fy) \leq 0$
	$Ax = b$	$\equiv$		

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

Példa.

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

Fenti esetben 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.

Egy összetettebb példa.

Példa.

Maximalizáljuk $x_1x_2 + x_2x_3 + x_3x_1$ -et. feltéve, hogy $x_1 + x_2 + x_3 = 100$, $x_1, x_2, x_3 > 0$.	$\equiv$	Minimalizáljuk $\sqrt{\frac{x_1^2 + x_2^2 + x_3^2}{3}}$ -et. feltéve, hogy $\frac{x_1 + x_2 + x_3}{3} = \frac{100}{3}$, $x_1, x_2, x_3 > 0$.
--	----------	--

A két feltételrendszer nyilván ekvivalens. 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 közepet kell minimalizálni adott számtani közép érték esetén. A számtani és négyzetes közép közötti egyenlőtlenség alapján elemi matematikával megválaszolható az optimalizálási kérdés.

4. LP normálformái

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.

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 nem-negatív koordinátájú pontok tér 2^n -ed részében van. Ezt a normálformát **előjeles poliedrikus normálformának** nevezem.

Példa.

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

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.

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 -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, akkor feltevésünk automatikusan teljesül.

Példa.

Maximalizáljuk	$x_1 + x_2$
Feltéve, hogy	$\begin{cases} -x_1 + x_2 + x_3 & = 1 \\ x_1 & + x_4 = 3 \\ x_2 & + x_5 = 2 \end{cases}$
	$x_1, x_2, x_3, x_4, x_5 \geq 0$