

Optimizacijske metode

Napisal Jure Pustoslemšek po zapiskih predavanj prof. dr. Matjaža Konvalinke

Junij 2020

1 Optimizacijski problemi

Hocemo maksimizirati ali minimizirati realno funkcijo, definirano na neki množici.

Zgled 1.1. Minimum/maksimum funkcije $f(x) = x^2 - 2x + 4$ na $[-2, 2]$
Uporabimo odvod funkcije.

Zgled 1.2. Minimum/maksimum funkcije $f(x, y) = x^2 - y^2$ na $[-3, 1] \times [0, 2]$
Gledamo parcialne odvode in vrednosti funkcije na robu.

Zgled 1.3 (Problem kmetije). Na kmetiji s 50ha pridelovalne površine želimo maksimizirati dobiček po spodnji tabeli

pridelek	ure dela	stroški	dobiček
pšenica	60	400	240
koruza	80	600	400
krompir	100	480	320

Na voljo imamo 5.000 ur delovne sile in 24.000€ kapitala.

Kaj nas zanima? Zanima nas, kako bomo razporedili različne pridelke po pridelovalni površini za največji dobiček.

x_1 . . . površina pšenice v ha

x_2 . . . površina koruze v ha

x_3 . . . površina krompirja v ha

Problem kmetije lahko matematično izrazimo takole:

$$\begin{array}{ll} \max & 240x_1 + 400x_2 + 320x_3 \\ \text{p.p.} & x_1 + x_2 + x_3 \leq 50 \\ & 60x_1 + 80x_2 + 100x_3 \leq 5.000 \\ & 400x_1 + 600x_2 + 480x_3 \leq 24.000 \\ & x_1, x_2, x_3 \geq 0 \end{array}$$

To je linearni program (LP).

Zgled 1.4. Imamo $2n$ jabolk, ki tehtajo w_1, \dots, w_{2n}

Jabolka razvrstimo v dve košari tako, da je v vsaki n jabolk, in da sta teži kosar čim bolj podobni.

Jabolka v 1. košari ponazorimo z množico A . Ker je v množici A natanko polovica jabolk, lahko 2. košaro ponazorimo z množico A^C .

Torej iščemo

$$\min \left| \sum_{w_i \in A} w_i - \sum_{w_i \in A^C} w_i \right|$$

Poskusimo ta problem predstaviti kot LP:

$x_i; i \in [2n]$

$x_i = 1$, če damo i -to jabolko v levo košaro.

$x_i = -1$, če damo i -to jabolko v desno košaro.

$$\begin{aligned} \min \quad & \left| \sum_{i=1}^{2n} w_i x_i \right| \\ \text{p.p.} \quad & \sum_{i=1}^{2n} x_i = 0 \\ & x_i \in \{-1, 1\} \end{aligned}$$

To ni linearen program! Absolutna vrednost ni linearna.

Definicija 1.5 (Optimizacijski problem). Optimizacijski problem karakteriziramo kot $(\min/\max, f, \Omega)$, kjer je f kriterijska funkcija (tj. funkcija, ki jo želimo minimizirati oz. maksimizirati), Ω pa množica dopustnih rešitev, tj. vseh x , ki ustrezajo danim pogojem.

Zanima nas optimalna vrednost, tj. $x^* \in \Omega$, za katerega je $f(x^*)$ največji oz. najmanjši možen.

Optimizacijski problem je:

- Nedopusten, če $\Omega = \emptyset$,
- Dopusten, če je $\Omega \neq \emptyset$

Dopustne probleme delimo na:

- Neomejen, če je Ω neomejena, torej lahko vedno najdemo boljšo rešitev,
- Omejen, če je Ω omejena

Omejene probleme pa še delimo na:

- Optimalen, če obstaja optimalna rešitev,
- Neoptimalen, če lahko vedno najdemo boljšo rešitev

2 Linearno programiranje

Definicija 2.1 (Linearni program). *Linearni program (LP) je optimizacijski problem, v katerem je kriterijska funkcija linearna, dopustna množica pa je podana z linearnimi enačbami in neenačbami.*

Zgled 2.2.

$$\begin{array}{ll} \min & 2x_1 - 3x_2 + x_3 \\ \text{p.p.} & x_1 + x_2 \leq 4 \\ & 3x_1 - x_2 - x_3 \geq 1 \\ & 2x_2 - 5x_3 = -1 \\ & x_1 \geq 0 \\ & x_3 \leq 0 \end{array}$$

Definicija 2.3 (Standardna oblika LP). *LP je v standardni obliki, če:*

1. je tipa \max ,
2. vse omejitve so oblike $g_j(x) \leq b_j$ in
3. za vse spremenljivke so omejene z $x_i \geq 0$

$$\begin{array}{ll} \max & c_1x_1 + \dots + c_nx_n \\ \text{p.p.} & a_{11}x_1 + \dots + a_{1n}x_n \leq b_1 \\ & a_{21}x_1 + \dots + a_{2n}x_n \leq b_2 \\ & \cdot \\ & \cdot \\ & \cdot \\ & a_{m1}x_1 + \dots + a_{mn}x_n \leq b_m \\ & x_1, x_2, \dots, x_n \geq 0 \end{array}$$

Slika 1: LP v standardni obliki

LP v standardni obliki lahko izrazimo tudi v matričnem zapisu.

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ x_n \end{bmatrix} \quad c = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ c_n \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ b_m \end{bmatrix} \quad A = [a_{ij}] \in \mathbb{R}^{m \times n}$$

S temi oznakami lahko vsak LP (tudi takšne, ki niso v matrični obliki) zapišemo v matričnem zapisu:

$$\begin{array}{ll} \max & c^T x \\ \text{p.p.} & Ax \leq b \\ & x \geq 0 \end{array}$$

Slika 2: LP v standardni obliki (matrični zapis)

Izrek 2.4. Vsak LP lahko pretvorimo v ekvivalentni LP v standardni obliki.

Dokaz. Pretvorbo poljubnega LP v standardno obliko opravimo v več korakih:

1. Minimizacijski problem je ekvivalenten maksimizacijskemu problemu negativne kriterijske funkcije z negativnimi pogoji.

$$\min f(x) \rightarrow \max -f(x)$$

$$a_{ij}x_i = b_j \rightarrow -a_{ij}x_i = -b_j$$

$$a_{ij}x_i \leq b_j \rightarrow -a_{ij}x_i \leq -b_j$$

$$a_{ij}x_i \geq b_j \rightarrow -a_{ij}x_i \geq -b_j$$

2. Omejitve spremenljivk popravimo z vpeljavo novih spremenljivk:

2.1. $x_i \geq b_j$ zamenjamo z novo spremenljivko $x'_i = x_i - b_j$, da dobimo omejitev $x'_i \geq 0$

2.2. $x_i \leq b_j$ zamenjamo z novo spremenljivko $x'_i = b_j - x_i$, da dobimo omejitev $x'_i \geq 0$

2.3. Neomejene spremenljivke zamenjamo z dvema novima spremenljivkama, podanima z enačbo $x_i = x'_i - x''_i$, novi spremenljivki pa sta omejeni z omejitvijo $x'_i, x''_i \geq 0$

3. Vsak pogoj oblike $g_j(x) = b_j$ pretvorimo v dva pogoja:

$$g_j(x) \leq b_j \text{ in } g_j(x) \geq b_j$$

4. Pogoje oblike $g_j(x) \geq b_j$ pomnožimo z -1 . S tem ga pretvorimo v $-g_j(x) \leq -b_j$

□

TODO: Konveksne množice

2.1 Simpleksna metoda

Simpleksna metoda zahteva LP v standardni obliki:

$$\begin{array}{ll} \max & c^T x \\ \text{p.p.} & Ax \leq b \\ & x \geq 0 \end{array}$$

Slika 3: LP v standardni obliki (matrični zapis)

Označimo z n število spremenljivk in z m število neenakosti, tj. $c, x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ in $A \in \mathbb{R}^{m \times n}$.

$$\begin{array}{ll} \max & 240x_1 + 400x_2 + 320x_3 \\ \text{p.p.} & x_1 + x_2 + x_3 \leq 50 \\ & 60x_1 + 80x_2 + 100x_3 \leq 5000 \\ & 400x_1 + 600x_2 + 480x_3 \leq 2400 \\ & x_1, x_2, x_3 \geq 0 \end{array}$$

Slika 4: LP problema kmetije v standardni obliki

Zgled 2.5 (Problem kmetije). *Koeficiente v kriterijski funkciji in vsaki omejitvi posebej bomo delili z največjim skupnim deliteljem, da bomo dobili ekvivalenten LP z "lepšimi" številkami.*

$$\begin{array}{ll} \max & 3x_1 + 5x_2 + 4x_3 \\ \text{p.p.} & x_1 + x_2 + x_3 \leq 50 \\ & 3x_1 + 4x_2 + 5x_3 \leq 250 \\ & 10x_1 + 15x_2 + 12x_3 \leq 600 \\ & x_1, x_2, x_3 \geq 0 \end{array}$$

Slika 5: LP problema kmetije z "lepšimi" koeficienti

Vsako neenakost spremenimo v enakost z novo spremenljivko.

$$\begin{array}{rcl} x_4 & = & 50 - x_1 - x_2 - x_3 \\ x_5 & = & 250 - 3x_1 - 4x_2 - 5x_3 \\ x_6 & = & 600 - 10x_1 - 15x_2 - 12x_3 \\ \hline z & = & 3x_1 + 5x_2 + 4x_3 \end{array}$$

Slika 6: Prvi slovar simpleksne metode na problemu kmetije

To je prvi slovar simpleksne metode.

V splošnem ima slovar $m+1$ linearnih enačb in $n+m+1$ spremenljivk. m spremenljivk izmed x_1, x_2, \dots, x_{n+m} in z je izraženih z ostalimi n spremenljivkami.

Teh m spremenljivk, ki so izražene z ostalimi n spremenljivkami, imenujemo bazne spremenljivke, ostale spremenljivke pa so nebazne (spremenljivka z predstavlja vrednost kriterijske funkcije in ne spada med bazne ali nebazne spremenljivke).

x_1, x_2, \dots, x_n so prvotne spremenljivke.

$x_{n+1}, x_{n+2}, \dots, x_{n+m}$ so dodatne spremenljivke.

V prvem slovarju so bazne spremenljivke natanko dodatne spremenljivke, nebazne pa prvotne.

Rečemo, da je slovar dopusten, če so vsi konstatni koeficienti baznih spremenljivk nenegativni. Prvi slovar je torej dopusten natanko takrat, ko je $b \geq 0$. Če je $b \not\geq 0$, uporabimo *dvofazno simpleksno metodo* (kasneje). Zato zaenkrat predpostavimo $b \geq 0$. V tem primeru je LP zagotovo dopusten, saj je $x_1 = x_2 = \dots = x_n = 0$ dopustna rešitev.

Opomba. Vse spremenljivke (tako prvotne kot dodatne) so nenegativne.

Z dopustnim slovarjem imamo *bazno dopustno rešitev*: nebazne spremenljivke nastavimo na 0. $f(x)$ je vrednost kriterijske funkcije, x pa so vrednosti prvotnih spremenljivk (nebazne smo nastavili na 0, bazne pa imajo vrednosti konstatnih koeficientov).

TODO: opis algoritma simpleksne metode

Izrek 2.6 (Osnovni izrek linearnega programiranja). 1. Če ima LP dopustno rešitev, ima tudi bazno dopustno rešitev.

2. Če ima LP optimalno rešitev, ima tudi bazno optimalno rešitev.

3. Velja natanko ena od možnosti za dani LP:

- LP je nedopusten.
- LP je neomejen.
- LP je optimalen.

2.2 Dualnost pri linearnem programiranju

Ideja: Problem kmetije P

$$\begin{array}{ll} \max & 3x_1 + 5x_2 + 4x_3 \\ \text{p.p.} & x_1 + x_2 + x_3 \leq 50 \\ & 3x_1 + 4x_2 + 5x_3 \leq 250 \\ & 10x_1 + 15x_2 + 12x_3 \leq 600 \\ & x_1, x_2, x_3 \geq 0 \end{array}$$

Slika 7: LP problema kmetije

Zanimajo nas zgornje meje za kriterijsko funkcijo.

Neenakosti pomnožimo z $y_1, y_2, \dots, y_m \geq 0$ in seštejemo:

$$(x_1 + x_2 + x_3)y_1 + (3x_1 + 4x_2 + 5x_3)y_2 + (10x_1 + 15x_2 + 12x_3)y_3 \leq 50y_1 + 250y_2 + 600y_3$$

$$(y_1 + 3y_2 + 10y_3)x_1 + (y_1 + 4y_2 + 15y_3)x_2 + (y_1 + 5y_2 + 12y_3)x_3 \leq 50y_1 + 250y_2 + 600y_3$$

Če velja:

$$y_1 + 3y_2 + 10y_3 \geq 3$$

$$y_1 + 4y_2 + 15y_3 \geq 5$$

$$y_1 + 5y_2 + 12y_3 \geq 4$$

je $2x_1 + 5x_2 + 4x_3 \leq 50y_1 + 250y_2 + 600y_3$.

Dualni problem kmetije P' :

$$\begin{array}{ll} \min & 50y_1 + 250y_2 + 600y_3 \\ \text{p.p.} & y_1 + 3y_2 + 10y_3 \geq 3 \\ & y_1 + 4y_2 + 15y_3 \geq 5 \\ & y_1 + 5y_2 + 12y_3 \geq 4 \\ & y_1, y_2, y_3 \geq 0 \end{array}$$

Slika 8: Dualni problem kmetije

V splošnem:

Definicija 2.7 (Dualni LP). *Za dani LP P v standardni obliki:*

$$\begin{array}{ll} \max & c^T x \\ \text{p.p.} & Ax \leq b \\ & x \geq 0 \end{array}$$

je dualni problem LP P'

$$\begin{array}{ll} \min & b^T y \\ \text{p.p.} & A^T y \geq c \\ & y \geq 0 \end{array}$$

Trditev 2.8. Za poljuben LP P je $P'' = P$.

Dokaz. Naj bo LP P v standardni obliki. Potem je P' v standardni obliki

$$\begin{array}{ll} \max & -b^T y \\ \text{p.p.} & (-A^T)y \leq -c \\ & y \geq 0 \end{array}$$

P'' je potem:

$$\begin{array}{ll} \min & -c^T x \\ \text{p.p.} & (-A^T)^T x \geq -b \\ & x \geq 0 \end{array}$$

kar po pretvorbi v standardno obliko postane enako kot P :

$$\begin{array}{ll} \max & c^T x \\ \text{p.p.} & Ax \leq b \\ & x \geq 0 \end{array}$$

□

Opomba. Če ima LP P n spremenljivk in m neenačb, ima njegov dualni LP P' m spremenljivk in n neenačb.

Izrek 2.9 (Šibki izrek o dualnosti (ŠID)). Naj bo P linearni program, P' njegov dualni LP ter x in y dopustni rešitvi za P in P' . Ob uporabi oznak od prej je potem $c^T x \leq b^T y$.

Dokaz. Brez izgube splošnosti predpostavimo, da je LP P v standardni obliki in opazimo, da je $A^T y \geq c$ in $Ax \leq b$. Ker velja tudi $x \geq 0$ in $y \geq 0$:

$$c^T x \leq (A^T y)^T x = y^T A x \leq y^T b = b^T y$$

□

Posledica 2.10. Če velja $c^T x^* = b^T y^*$ za dopustni rešitvi x^*, y^* , je x^* optimalna rešitev za P in y^* optimalna rešitev za P' .

Dokaz. Ali lahko velja $c^T x > c^T x^*$ za dopustno rešitev x ? Potem velja $c^T x > b^T y^*$, ampak po ŠID velja $c^T x \leq b^T y^*$. Protislovje.

Kaj pa $b^T y < b^T y^*$ za neko dopustno rešitev y ? Potem velja $b^T y < c^T x^*$, ampak po ŠID je $b^T y \geq c^T x^*$. Znova smo prišli v protislovje. Torej x^* in y^* morata biti optimalni rešitvi za P in P' . □

Posledica 2.11. Če je P neomejen, je P' nedopusten.

Izrek 2.12 (Krepki izrek o dualnosti (KID)). Če ima P optimalno rešitev x^* , ima tudi P' optimalno rešitev y^* in velja $c^T x^* = b^T y^*$.

Dokaz. Ogledamo si zadnji slovar:

$$z = v^* + \sum_{i=1}^{n+m} c_i^* x_i$$

$$v^* = \sum_{i=1}^n c_i x_i^*$$

$c_i^* = 0$, če je x_i bazna spremenljivka in $c_i^* \leq 0$, če je x_i nebazna spremenljivka.

$$y_i^* = -c_{i+n} \text{ za } i = 1, \dots, m$$

Trditev 2.13. *naj bo (y_1^*, \dots, y_m^*) dopustna rešitev za P' . Potem je*

$$\sum_{i=1}^n c_i c_i^* = \sum_{i=1}^m$$

Če dokažemo to trditev, sta po posledici 2.10 x^* in y^* optimalni.

$$z = v^* + \sum_{i=1}^n c_i^* x_i + \sum_{i=n+1}^{n+m} c_i^* x_i$$

$$z = v^* + \sum_{i=1}^n c_i^* x_i + \sum_{i=1}^m c_{i+n} x_{i+n}$$

$$c_{i+n}^* = -y_i^* \text{ in } x_{i+n} = b_i - \sum_{j=1}^n a_{ij} x_j$$

$$z = v^* + \sum_{i=1}^n c_i^* x_i + \sum_{j=1}^m (-y_j^*) (b_j - \sum_{i=1}^n a_{ji} x_i)$$

$$z = (v^* - \sum_{j=1}^m y_j^* b_j) + \sum_{i=1}^n x_i (c_i^* + \sum_{j=1}^m a_{ji} y_j^*)$$

Po drugi strani pa je $z = \sum_{i=1}^n c_i x_i$. Iz tega sledi:

$$v^* - \sum_{j=1}^m y_j^* b_j = 0$$

$$c_i^* + \sum_{j=1}^m a_{ij} y_j^* = c_i$$

Ker je $c_i^* \leq 0$ in $y_i^* = -c_{i+n}^*$, je $y_i^* = -c_{i+n}^* \geq 0$. Iz tega sledi, da je y^* dopustna rešitev:

$$\sum_{j=1}^m a_{ji} y_j^* \geq c_i$$

$$y_1^*, y_2^*, \dots, y_m^* \geq 0$$

Iz $v^* - \sum_{j=1}^m y_j^* b_j = 0$ sledi tudi, da je

$$v^* = \sum_{j=1}^m y_j^* b_j$$

Ker pa je $v^* = \sum_{i=1}^n c_i x_i^*$, sledi

$$\sum_{j=1}^m a_{ji} y_j^* = \sum_{j=1}^m b_j y_i^*$$

.

□