Metode notranjih točk za LP

Rešujemo primarni linearni program

$$\begin{aligned} \min \quad & c^T x \\ & Ax = b \\ & x \geq 0 \end{aligned}$$

in dualni linearni program

$$\begin{aligned} \max \quad & b^T y \\ & A^T y \leq c \end{aligned}$$

oziroma

$$\begin{aligned} \max \quad & b^T y \\ & A^T y + s = c \\ & s \geq 0 \end{aligned}$$

po vpeljavi dopolnilnih spremenljivk.

KKT pogoji so potrebni in zadostni za LP. Rešitev sistema

$$\begin{aligned} Ax &= b \\ A^T y + s &= c \\ x \circ s &= 0 \\ x,\ s &\geq 0 \end{aligned} \tag{F}$$

določa optimalni rešitvi x^* za primarni in (y^*, s^*) za dualni program.

Osnovna ideja: Simpleksna metoda se sprehaja po robu poliedra in pregleduje njegova oglišča. *Metode notranjih točk (Interior point methods)* se sprehodijo čez notranjost množice dopustnih rešitev do optimalne rešitve. Pri tem se izogibajo robu.

Predpostavka: Primarni in dualni program sta strogo dopustna, tj. obstaja (x_0, y_0, s_0), da velja

$$Ax_0 = b, \quad A^T y_0 + s_0 = c, \quad x_0, s_0 > 0.$$

V odvisnosti od parametra $\mu \in [0, \infty)$ definiramo družino funkcij

$$f_\mu(x) = c^T x - \mu \sum_{i=1}^{n} \ln(x_i)$$

in

$$\widetilde{f}_\mu(y) = b^T y + \mu \sum_{i=1}^{n} \ln(c_i - a_i^T y).$$

Pri $\mu = 0$ dobimo kriterijski funkciji primarnega in dualnega programa. Za vsak μ je $f_\mu(x)$ konveksna, $\widetilde{f}_\mu(x)$ pa konkavna. Funkciji imenujemo *logaritmični odbojni* *funkciji*. f_μ je definirana tako, da njene vrednosti rastejo proti ∞, ko se x bliža robu poliedra, definiranega s sistemom enačb $Ax = b,\ x \geq 0$. Podobno velja za $\widetilde{f}_\mu$.

Vzemimo optimizacijska problema

$$\begin{array}{ll} \min & f_\mu(x) \\ & Ax = b \\ & x > 0 \end{array} \tag{P_μ}$$

in

$$\begin{array}{ll} \max & \widetilde{f}_\mu(y) \\ & A^T y < c \end{array} \tag{D_μ}$$

Na predavanjih ste pokazali nasledji izrek.

Izrek. Naj bo $\mu > 0$. Naslednje trditve so ekvivalentne.

- Primarni in dualni linearni program sta strogo dopustna.
- Obstaja enolična rešitev programa (P_μ).
- Obstaja enolična rešitev programa (D_μ).
- Sistem

$$F_\mu(x, y, s) = \begin{pmatrix} Ax - b \\ A^T y + s - c \\ x \circ s - \mu e \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad x \geq 0,\ \ s \geq 0 \tag{F_μ}$$

  ima enolično rešitev.

Pri tem sta optimalni rešitvi $x(\mu)$ za (P_μ) in $y(\mu)$ za (D_μ) ravno rešitvi sistema (F_μ). Pri tem velja $s(\mu) = c - A^T y(\mu)$.

Definicija. Primarno-dualna *središčna pot* je množica $\{(x(\mu), y(\mu), s(\mu) \colon \mu > 0\}$.

Opazimo, da je sistem (F_μ) ravno perturbiran sistem sistema (F). Pri tem smo pogoj $x \circ s = 0$ nadomestili z $x \circ s = \mu e$. Ideja je, da začnemo v okolici točke $(x(\mu), y(\mu), s(\mu)$ pri nekem $\mu > 0$ in z zmanjševanjem $\mu \to 0$ sledimo središčni poti vse do optimalne rešitve.

Algoritem.

Splošna Metoda notranjih točk

Naj bo $(\mathbf{x}^{(0)}, \mathbf{y}^{(0)}, \mathbf{s}^{(0)}) \in \mathcal{F}_P^0 \times \mathcal{F}_D^0$, $\varepsilon > 0$, $\sigma \in [0, 1]$

Inicializacija: $(\mathbf{x}, \mathbf{y}, \mathbf{s}) = (\mathbf{x}^{(0)}, \mathbf{y}^{(0)}, \mathbf{s}^{(0)})$, $\tau = \mathbf{x}^\top \mathbf{y}/n$.

Dokler $\tau > \varepsilon$

- REŠI

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^\top & I \\ \mathrm{Diag}(\mathbf{s}) & 0 & \mathrm{Diag}(\mathbf{x}) \end{pmatrix} \cdot \begin{pmatrix} \Delta \mathbf{x} \\ \Delta \mathbf{y} \\ \Delta \mathbf{s} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \mathbf{x} \circ \mathbf{s} - \sigma \tau \mathbf{e} \end{pmatrix},$$

- $(\mathbf{x}, \mathbf{y}, \mathbf{s}) = (\mathbf{x}, \mathbf{y}, \mathbf{s}) - \alpha(\Delta \mathbf{x}, \Delta \mathbf{y}, \Delta \mathbf{s})$, kjer $\alpha \in (0, 1]$ primerna dolžina koraka.
- $\tau = \mathbf{x}^\top \mathbf{y}/n$.

Metode notranjih točk, ki sledijo centralni poti delimo na

- kratkokoračne
- dolgokoračne
- metode prediktor-korektor

Pri kratkokoračnih metodah počasi zmanjšujemo τ. Za faktor σ velikokrat izberemo

$$\sigma = 1 - \frac{0.4}{\sqrt{n}}.$$

Zaradi manjšega koraka po dopustni množici lahko napravimo polni korak pri Newtonovi metodi, tj.

$$\begin{aligned} x_{k+1} &= x_k + \Delta x \\ y_{k+1} &= y_k + \Delta y \\ s_{k+1} &= s_k + \Delta s \end{aligned}$$

Pri dolgokoračnih metoda hitreje zmanjšujemo τ. Za faktor σ velikokrat izberemo $\sigma = 0.5$. Zaradi večjega koraka po dopustni množici lahko postane par (x_{k+1}, s_{k+1}) nepozitiven. Zato napravimo delni Newton korak. Poiščemo tak α_p, da velja $x_{k+1} = x_k + \alpha_p \Delta x > 0$. Podobno za s_{k+1}. Tako dobimo novo točko

$$\begin{aligned} x_{k+1} &= x_k + \alpha_p \Delta x \\ y_{k+1} &= y_k + \alpha_d \Delta y \\ s_{k+1} &= s_k + \alpha_d \Delta s \end{aligned}$$

Dolgokoračne metode so ponavadi v praksi učinkovitejše.

1. Sestavite algoritem, ki izvaja kratkokoračno metodo notranjih točk. Delovanje metode preverite na problemu s podatki $c = (-1, -3, -4)$, $A = (1, 1, 1)$ in $b = 1$. Optimalna vrednost je enaka -4.

2. Sestavite algoritem, ki izvaja dolgokoračno metodo notranjih točk in njegovo delovanje preverite na podatkih iz prejšnjega primera.

Metode prediktor-korektor

Pri obeh zgornjih algoritmih smo potrebovali začetno strogo dopustno rešitev (x_0, y_0, s_0). Za veliko programov je tako točko težko najti. Ogledali si bomo prediktor-korektor metodo, ki potrebuje začetno točko z lastnostjo $(x_0, s_0) > 0$. Ker točke (x_k, y_k, s_k), dobljene tekom algoritma, ne bodo nujno dopustne, moramo zraven dualnega razmika preverjati tudi primarno in dualno dopustnost .

V prediktor koraku iščemo popravke $(\Delta x_1, \Delta y_1, \Delta s_1)$, da bo veljalo

$$\begin{aligned} A(x + \Delta x_1) &= b \\ A^T(y + \Delta y_1) + s + \Delta s_1 &= c \\ (x + \Delta x_1) \circ (s + \Delta s_1) &= 0 \end{aligned}$$

Če zanemarimo člen $\Delta x_1 \circ \Delta s_1$, iščemo rešitev sistema

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta y_1 \\ \Delta s_1 \end{pmatrix} = \begin{pmatrix} b - Ax \\ c - A^T y - s \\ -x \circ s \end{pmatrix} = \begin{pmatrix} r_p \\ r_d \\ r_c \end{pmatrix}.$$

Iz druge enačbe izrazimo Δs_1 in vstavimo v tretjo enačbo. Iz te izrazimo Δx_1 in uporabimo prvo enačbo. Dobimo

$$\begin{aligned} AS^{-1}XA^T\Delta y_1 &= r_p - AS^{-1}(r_c - x \circ r_d) \\ \Delta s_1 &= r_d - A^T\Delta y_1 \\ S\Delta x_1 &= r_c - X\Delta s_1 \end{aligned}$$

Z rešitvijo je tako določena smer do optimalne rešitve. V drugem koraku (korektor) smer popravimo tako, da se premaknemo nazaj k središčni poti. Iščemo popravke $(\Delta x_2, \Delta y_2, \Delta s_2)$, da bo veljalo

$$\begin{aligned} A(x + \Delta x_2) &= b \\ A^T(y + \Delta y_2) + s + \Delta s_2 &= c \\ (x + \Delta x_2) \circ (s + \Delta s_2) &= \mu e \end{aligned}$$

Če člen $\Delta x_2 \circ \Delta s_2$ zamenjamo s produktom členov $\Delta x_1 \circ \Delta s_1$ iz prediktor koraka, rešujemo sistem

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x_2 \\ \Delta y_2 \\ \Delta s_2 \end{pmatrix} = \begin{pmatrix} b - Ax \\ c - A^T y - s \\ \mu e - x \circ s - \Delta x_1 \circ \Delta s_1 \end{pmatrix} = \begin{pmatrix} r_p \\ r_d \\ r_c' \end{pmatrix}.$$

Podobno kot zgoraj se rešitev izraža z enačbami

$$\begin{aligned} AS^{-1}XA^T\Delta y_2 &= r_p - AS^{-1}(r_c' - x \circ r_d) \\ \Delta s_2 &= r_d - A^T\Delta y_2 \\ S\Delta x_2 &= r_c' - X\Delta s_2 \end{aligned}$$

Nadaljujemo z iskanjem α_p in α_d kot pri dolgokoračni metodi.

3. Sestavite algoritem, ki izvaja prediktor-korektor metodo notranjih točk. Delovanje metode preverite na problemu

```
randn('state', 0);
rand('state', 0);

n = 500;
m = 400;

c  = rand(n,1) + 0.5;
x0 = abs(randn(n,1));

A = abs(randn(m,n));
b = A*x0;
```

Za začetna strogo pozitivna vektorja izberite `x0 = ones(n,1)` in `y0 = zeros(m,1)`. Optimalna vrednost je enaka 371.089.