Reševanje semidefinitnih programov

Rešujemo primarni semidefinitni program

$$\begin{aligned} \min \quad & C \bullet X \\ & A(X) = b \\ & X \succeq 0 \end{aligned}$$

in pripadajoči dualni program

$$\begin{aligned} \max \quad & b^T y \\ & A^T y + Z = C \\ & Z \succeq 0. \end{aligned}$$

Pri tem sta za dane matrike A_i, $i = 1, 2 \ldots, m$ operator $A \colon \mathbb{R}^{n\times n} \to \mathbb{R}^m$ in pripadajoči adjungiran operator $A^T \colon \mathbb{R}^m \to \mathbb{R}^{n\times n}$ definirana kot

$$A(X) = \begin{pmatrix} A_1 \bullet X \\ \vdots \\ A_m \bullet X \end{pmatrix} \quad \text{in} \quad A^T(y) = \sum_{i=1}^{m} y_i A_i.$$

V prvem delu si bomo ogledali reševanja semidefinitnih programov s pomočjo programskega paketa SeDuMi, ki je dostopen na spodnji povezavi

`http://sedumi.ie.lehigh.edu/?page_id=58`

Prenesite datoteke pod `SeDuMi for Octave` in v Matlabu poženite skripto `install_sedumi`.

SeDuMi rešuje konične optimizacijske probleme nad sebidualnimi stožci:

- Nenegativen ortant $\mathbb{R}^n_+$,
- Lorentzov stožec $\mathcal{L}^n$,
- Stožec pozitivno semidefinitnih matrik $\mathcal{S}^n_+$.

Primarni SDP, ki je dan s podatki C, b in A_i, $i = 1, 2, \ldots, m$, rešimo s klicem funkcije `sedumi(A,b,c,K)`. Pri tem je $c = vec(C)$ vektorizacija matrike C, K polje, ki določa dimenzijo stožca (če je matrika X velikosti $n \times n$, je $K.s = n$) in A je matrika, katere i-ta vrstica je vektorizacija matrike A_i.

1. Dan je SDP oblike

$$\begin{aligned} \max \quad & b^T y \\ & \begin{pmatrix} 1 - y_1 & 2 - 2y_2 & 3 - y_1 - 8y_2 \\ 2 - 2y_2 & 9 - 3y_1 - 6y_2 & -7y_1 \\ 3 - y_1 - 8y_2 & -7y_1 & 7 - 5y_1 - 4y_2 \end{pmatrix} \succeq 0, \end{aligned}$$

kjer je $b = (11, 9)$. Izračunajte njegovo optimalno vrednost. *Rešitev: 9.5259.*

2. Dan je cikel C_5, tj. graf $G = (V, E)$, kjer je $V = \{1, 2, 3, 4, 5\}$ množica vozlišč in
$$E = \{(1,2), (2,3), (3,4), (4,5), (1,5)\}$$
množica povezav. Izračunajte optimalno vrednost semidefinitnega programa
$$\begin{aligned} \max \quad & J \bullet X \\ & \mathrm{sled}(X) = 1 \\ & X_{ij} = 0, \quad \text{če } (i,j) \in E \\ & X \succeq 0. \end{aligned}$$
Pri tem smo z J označili matriko samih enic. *Rešitev:* $\sqrt{5} = 2.2361$.

3. Naložite si podatke semidefinitnega programa z ukazom `podatki_max_cut_sdp.mat` in izračunajte njegovo optimalno vrednost. *Rešitev: -294.8725.*

4. Problem izračuna najmanjše lastne vrednosti matrike A lahko formuliramo kot SDP
$$\begin{aligned} \max \quad & t \\ & A - tI \succeq 0. \end{aligned}$$
Izračunajte najmanjšo lastno vrednost matrike `A = mat(c)` iz prejšnjega primera in rezultat preverite z vgrajeno funkcijo `eig`. *Rezultat: -8.6255.*

Metode notranjih točk

Predpostavka: Primarni in dualni program sta strogo dopustna, tj. obstaja trojica (X_0, y_0, Z_0), da velja
$$A(X_0) = b, \quad A^T y_0 + Z_0 = C, \quad X_0, Z_0 \succ 0.$$
V tem primeru so KKT pogoji potrebni in zadostni. Rešitev sistema
$$\begin{aligned} A(X) &= b \\ A^T y + Z &= c \\ ZX &= 0 \\ Z, X &\succeq 0 \end{aligned} \tag{F}$$
določa optimalni rešitvi X^* za primarni in (y^*, Z^*) za dualni program.

V odvisnosti od parametra $\mu \in [0, \infty)$ definiramo družino funkcij $f_\mu \colon S^n_+ \to \mathbb{R}$ s predpisom
$$f_\mu(X) = C \bullet X - \mu \log \det(X).$$
Za vsak μ je $f_\mu(X)$ konveksna funkcija in na vajah smo pokazali, da velja
$$\nabla f_\mu(X) = C - \mu X^{-1}.$$

Logaritmična odbojna funkcija f_μ je definirana tako, da njene vrednosti rastejo proti ∞, ko se X bliža robu stožca pozitivno semidefinitnih matrik. Rob sestavljajo vse matrike $X \succeq 0$, ki imajo vsaj eno lastno vrednost enako 0.

Vzemimo optimizacijski problem

$$\begin{aligned} \min \quad & f_\mu(x) \\ & A(X) = b \\ & X \succ 0. \end{aligned} \tag{P_μ}$$

Na predavanjih ste pokazali naslednji izrek.

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

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

  $$F_\mu(x,y,s) = \begin{pmatrix} A(X) - b \\ A^T y + Z - C \\ ZX - \mu I \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad X \succeq 0, \;\; Z \succeq 0 \tag{F_μ}$$

  ima enolično rešitev.

Pri tem rešitev sistema (F_μ) določa točko na središčni poti.

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

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

Metode prediktor-korektor

V nadaljevanju si bomo ogledali prediktor-korektor metodo, ki potrebuje začetni matriki z lastnostjo $X_0, Z_0 \succ 0$. Ker točke (X_k, y_k, Z_k), dobljene tekom algoritma, ne bodo nujno dopustne, moramo zraven dualnega razmika preverjati tudi primarno in dualno dopustnost .

V prediktor koraku iščemo popravke $(\Delta X_1, \Delta y_1, \Delta Z_1)$, da bo veljalo

$$\begin{aligned} A(X + \Delta X_1) &= b \\ A^T(y + \Delta y_1) + Z + \Delta Z_1 &= C \\ (Z + \Delta Z_1)(X + \Delta X_1) &= 0 \end{aligned}$$

Če zanemarimo člen $\Delta Z_1 \Delta X_1$, iščemo rešitev sistema

$$\begin{aligned} A(\Delta X_1) &= b - A(X) = r_p \\ A^T \Delta y_1 + \Delta Z_1 &= C - A^T y - Z = r_d \\ \Delta Z_1 X + Z \Delta X_1 &= -ZX = r_c \end{aligned}$$

Iz druge enačbe izrazimo ΔZ_1 in vstavimo v tretjo enačbo. Iz te izrazimo ΔX_1 in uporabimo prvo enačbo. Dobimo enačbe

$$\begin{aligned} M\Delta y_1 &= b + A(Z^{-1} r_d X) \\ \Delta Z_1 &= r_d - A^T \Delta y_1 \\ Z\Delta x_1 &= r_c - \Delta Z_1 X \end{aligned}$$

pri čemer velja $M\Delta y = A(Z^{-1} A^T \Delta y X)$ in je M pozitivno definitna matrika.

5. Dokažite, da velja

$$m_{ij} = \text{sled}(A_i Z^{-1} A_j X).$$

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 $(\Delta X_2, \Delta y_2, \Delta Z_2)$, da bo veljalo

$$\begin{aligned} A(X + \Delta X_2) &= b \\ A^T(y + \Delta y_2) + Z + \Delta Z_2 &= C \\ (Z + \Delta Z_2)(X + \Delta X_2) &= \mu I \end{aligned}$$

Če člen $\Delta Z_2 \Delta X_2$ zamenjamo s produktom $\Delta Z_1 \Delta X_1$ iz prediktor koraka, rešujemo sistem

$$\begin{aligned} A(\Delta X_2) &= b - A(X) = r_p \\ A^T \Delta y_2 + \Delta Z_2 &= C - A^T y - Z = r_d \\ \Delta Z_2 X + Z \Delta X_2 &= \mu I - ZX - \Delta Z_1 \Delta X_1 = r_c' \end{aligned}$$

Podobno kot zgoraj se rešitev izraža z enačbami

$$\begin{aligned} M\Delta y_2 &= b + A(Z^{-1} r_d X) - \mu A(Z^{-1}) + A(Z^{-1} \Delta Z_1 \Delta X_1) \\ \Delta Z_2 &= r_d - A^T \Delta y_2 \\ Z\Delta X_2 &= r_c' - \Delta Z_2 X \end{aligned}$$

Opomba: Matriki ΔX_1 in ΔX_2 v splošnem nista simetrični. Ena izmed rešitev, da po vsakem izračunu ΔX matriko simetriziramo

$$\Delta X = \frac{\Delta X + \Delta X^T}{2}.$$

Nadaljujemo z iskanjem α_p in α_d kot pri dolgokoračni metodi. Pri tem namesto pozitivnosti vektorja preverjamo pozitivno definitnost s pomočjo razcepa Choleskega.

6. Delovanje vašega programa preverite na primeru iz točke 2. Za začetni par vzemite `X0 = eye(5)` in `y0 = (0,0,0,0,0,-6)`.