Szemidefinit programozás

Hajnal Péter

2021. tavasz

Szemidefinit programozás, SDP

A szemidefinit programozási alapfeladat általános alakja

Minimalizáljuk	$c^{\mathsf{T}}x$-t
Feltéve, hogy	$\sum_{i=1}^{n} x_i A_i \preceq B$
	$Dx = e$,

ahol $c, x \in \mathbb{R}^n$, $A_i, B \in \mathcal{S}^k = \{M \in \mathbb{R}^{k\times k} : M^{\mathsf{T}} = M\} \subset \mathbb{R}^{k\times k}$, $D \in \mathbb{R}^{\ell\times n}$ és $e \in \mathbb{R}^{\ell}$.

$\mathcal{S}^n$ a valós szimmetrikus $n \times n$ méretű mátrixok halmaza. Azaz $M \in \mathbb{R}^{n\times n}$ akkor és csak akkor eleme $\mathcal{S}^n$-nek, ha $M^{\mathsf{T}} = M$. Speciálisan $\mathcal{S}^n \subset \mathbb{R}^{n\times n}$.

Jelölés

$A \preceq B$ akkor és csak akkor, ha $A, B \in \mathcal{S}^n$, továbbá $0 \preceq B - A$, azaz $B - A$ pozitív szemidefinit, jelölésben $B - A \in \mathcal{S}^n_+$.

$\sum_{i=1}^{n} x_i A_i$

$$\begin{pmatrix} A^{(1)}_{1,1}x_1+\ldots+A^{(n)}_{1,1}x_n & A^{(1)}_{1,2}x_1+\ldots+A^{(n)}_{1,2}x_n & \ldots & A^{(1)}_{1,k}x_1+\ldots+A^{(n)}_{1,k}x_n \\ A^{(1)}_{2,1}x_1+\ldots+A^{(n)}_{2,1}x_n & A^{(1)}_{2,2}x_1+\ldots+A^{(n)}_{2,2}x_n & \ldots & A^{(1)}_{2,k}x_1+\ldots+A^{(n)}_{2,k}x_n \\ \vdots & \vdots & & \vdots \\ A^{(1)}_{k,1}x_1+\ldots+A^{(n)}_{k,1}x_n & A^{(1)}_{k,2}x_1+\ldots+A^{(n)}_{k,2}x_n & \ldots & A^{(1)}_{k,k}x_1+\ldots+A^{(n)}_{k,k}x_n \end{pmatrix}$$

Példa

$$\begin{pmatrix} 2x_1-3x_2+x_3 & 5x_1+2x_2-x_3 & x_1-x_2+8x_3 & 6x_1+5x_2+x_3 \\ 5x_1+2x_2-x_3 & -x_1+7x_2-2x_3 & 9x_1-3x_2+x_3 & -2x_1-x_2+4x_3 \\ x_1-x_2+8x_3 & 9x_1-3x_2+x_3 & 10x_1-2x_2+2x_3 & 8x_1+x_2+x_3 \\ 6x_1+5x_2+x_3 & -2x_1-x_2+4x_3 & 8x_1+x_2+x_3 & -11x_1+2x_2-3x_3 \end{pmatrix}$$

Lineáris egyenlőtlenségek mint SDP feltételek

Az előjelfeltételek felírhatók, mint alkalmas mátrix pozitív szemidefinitása:

$$x_1, \ldots x_n \geq 0 \iff \begin{pmatrix} x_1 & & 0 \\ & \ddots & \\ 0 & & x_n \end{pmatrix} \succeq 0 \iff \begin{pmatrix} -x_1 & & 0 \\ & \ddots & \\ 0 & & -x_n \end{pmatrix} \preceq 0$$

Észrevétel

Az LP alapfeladat egy speciális SDP probléma.

$\mathcal{S}^n$

$\mathcal{S}^n$ az $\mathbb{R}^{n\times n}/\mathbb{R}^{[n]\times[n]}/\mathbb{R}^{H\times H}$ egy lineáris altere. Dimenziója $\left(\!\binom{n}{2}\!\right)$.

Jelölés

$\langle .,. \rangle : \mathcal{S}^n \times \mathcal{S}^n \to \mathbb{R}$ az alábbi skalárszorzat: $M, N \in \mathcal{S}^n$ esetén

$$\langle M, N\rangle = \mathrm{Tr}(M^{\mathrm{T}}N) = \sum_j (M^{\mathrm{T}}N)_{jj} = \sum_{i,j} M^{\mathrm{T}}_{ji} N_{ij} = \sum_{i,j} M_{ij}N_{ij}.$$

Egy alternatíva: Legyen $vec : \mathbb{R}^{n\times n} \to \mathbb{R}^{n^2}$ egy mátrix oszlopfolytonos olvasata (azaz egy számtáblázatból egy számsort kézpünk). Ekkor $\langle M, N\rangle = vec(M)^{\mathsf{T}} vec(N)$.

$\langle .,.\rangle$

A továbbiakban összefoglaljuk a bevezetett skalárszorzás tulajdonságait. Ezeket az érdeklődő hallgató maga is igazolhatja.

Lemma

Legyen $M, N, P \in \mathcal{S}^n$

(i) $\langle M, N\rangle = \langle N, M\rangle$

(ii) $\langle MN, P\rangle = \langle M, PN\rangle$

(iii) $\langle M, M\rangle = \mathrm{Tr} M^2 = \sum_{i=1}^{n} \lambda_i^2 \geq 0$

Jelölés

$\|M\|_F = \sqrt{\langle M, M\rangle}$, az M szimmetrikus mátrix Frobenius normája.

$\mathcal{S}^n_+$, pozitív szemidefinit mátrixok

Lemma

Legyen $M \in \mathcal{S}^n$ tetszőleges. Ekkor a következők ekvivalensek:

(i) M pozitív szemidefinit, azaz $x^\mathsf{T} M x \geq 0$ minden $x \in \mathbb{R}^n$ esetén,

(ii) M sajátértékei nemnegatívok,

(iii) M egy Gram-mátrix, azaz alkalmas $V \in \mathbb{R}^{k \times n}$ mátrixra $M = V^\mathsf{T} V$,

(iv) M minden szimmetrikus részmátrixa (tetszőleges módon elhagyunk s sort és a nekik megfelelő s oszlopot) nemnegatív determinánsú.

Lemma, ami szintén egy ekvivalens leírás

Lemma

Ha $M \succeq 0$, akkor létezik olyan $M^{\frac{1}{2}} \succeq 0$, amelyre $M = M^{\frac{1}{2}} M^{\frac{1}{2}}$.

- M sajátértékei legyenek $\lambda_1, \ldots, \lambda_n \geq 0$.
- Továbbá legyen

$$\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \qquad \Lambda^{\frac{1}{2}} = \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix}.$$

- Legyen M szimmetrikus, ezért létezik olyan Q ortogonális mátrix, amelyre $Q^{\mathrm{T}} M Q = \Lambda$.
- Így

$$M = Q \Lambda Q^{\mathsf{T}} = Q \Lambda^{\frac{1}{2}} \Lambda^{\frac{1}{2}} Q^{\mathsf{T}} = Q \Lambda^{\frac{1}{2}} Q^{\mathsf{T}} Q \Lambda^{\frac{1}{2}} Q^{\mathsf{T}} = (Q \Lambda^{\frac{1}{2}} Q^{\mathsf{T}})(Q \Lambda^{\frac{1}{2}} Q^{\mathsf{T}}).$$

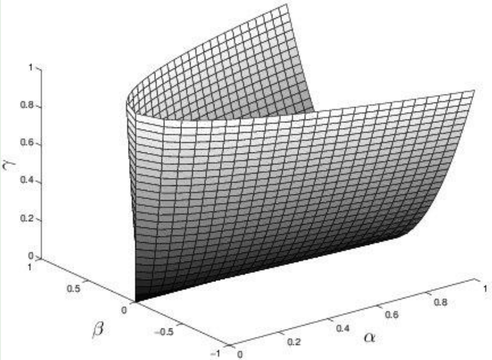
- $M^{\frac{1}{2}} = Q \Lambda^{\frac{1}{2}} Q^{\mathsf{T}}$ választás bizonyítja a Lemmát.

$\mathcal{S}_+^n$ geometriailag

Észrevétel

$\mathcal{S}_+^n$ egy kúp $\mathcal{S}^n \subset \mathbb{R}^{n\times n}$-ben.

$\begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$ pozitív szemidefinit mátrixok halmaza



SDP normálformák

SDP: I. normálforma

Minimalizáljuk	$\langle C, X\rangle$-t
Feltéve, hogy	$\langle A_i, X\rangle = b_i, \quad i = 1, \dots k$
	$X \succeq 0,$

ahol $C, X, A_i \in \mathcal{S}^n$ és $b \in \mathbb{R}^k$.

SDP: II. normálforma

Minimalizáljuk	$c^{\mathsf{T}} x$-t
Feltéve, hogy	$\sum_{i=1}^{n} x_i A_i \preceq B,$

ahol $c, x \in \mathbb{R}^n$ és $A_i, B \in \mathcal{S}^k$.

Szünet

Dualizálás (Lagrange módszer)

Az első normálformában vizsgáljuk a primál feladatot:

Minimalizáljuk	$\langle C, X\rangle$-t
Feltéve, hogy	$\langle A_i, X\rangle = b_i, \quad i = 1, \dots k,$
	$X \succeq 0.$

A szokásos ötlettel élünk. A feltételeket „beépítjük" a Lagrange-függvénybe.

A véges sok lineáris feltétellel nincs probléma.

A pozitív szemidefinitás feltételét viszont nem tudjuk L-be beépíteni, L értelmezési tartományát „terheljük" ezen feltétellel.

$$L(X, \mu) = \langle C, X\rangle + \sum \mu_i(\langle A_i, X\rangle - b_i), \quad \operatorname{dom} L = \{X :\ X \succeq 0\}.$$

Dualizálás (Lagrange módszer)

Ennek a függvénynek a minimalizálása vezet a duális probléma célfüggvényéhez.

$$\widehat{c}(\mu) = \inf_{X:\ X \succeq 0} L(X, y) = -\sum b_i \mu_i + \inf \left\langle C + \sum \mu_i A_i, X \right\rangle .$$

Az $\inf_{X \succeq 0} \langle M, X \rangle =?$ optimalizálási feladatot kell megoldanunk.

Dualizálás (Lagrange módszer)

Lemma

M akkor és csak akkor pozitív szemidefinit, ha minden $X \succeq 0$ esetén $\langle M, X\rangle \geq 0$.

A bizonyítás előtt kiemelünk egy következményt.

Következmény

Ha $M, N \succeq 0$, akkor $\langle M, N\rangle \geq 0$.

Szükségünk lesz egy egyszerűen ellenőrizhető egyenlőségre:

Segédlemma

Legyen $A \in \mathbb{R}^{k\times\ell}$ és $B \in \mathbb{R}^{\ell\times k}$. Ekkor

$$\operatorname{Tr}(AB) = \operatorname{Tr}(BA).$$

Lemma bizonyítása

- Legyen $M, X \succeq 0$. Ekkor

$$\begin{aligned}
\langle M, X \rangle &= \operatorname{Tr}(MX) = \operatorname{Tr}(M^{1/2}M^{1/2}X^{1/2}X^{1/2}) \\
&= \operatorname{Tr}((M^{1/2}M^{1/2}X^{1/2})X^{1/2}) = \operatorname{Tr}(X^{1/2}(M^{1/2}M^{1/2}X^{1/2})) \\
&= \operatorname{Tr}((X^{1/2}M^{1/2})(M^{1/2}X^{1/2})) \\
&= \operatorname{Tr}((M^{1/2}X^{1/2})^{\mathsf{T}}(M^{1/2}X^{1/2})) \geq 0,
\end{aligned}$$

hiszen $M^{1/2}$ és $X^{1/2}$ is szimmetrikus.

- Fordítva tegyük fel, hogy $\langle M, X \rangle \geq 0$ minden $X \succeq 0$ mátrixra. Legyen $x \in \mathbb{R}^n$ tetszőleges. Feltevésünket alkalmazzuk az $X = xx^{\mathsf{T}}$ pozitív szemidefinit mátrixra:

$$\begin{aligned}
0 &\leq \langle M, X \rangle = \operatorname{Tr}(M(xx^{\mathsf{T}})) = \operatorname{Tr}((Mx)x^{\mathsf{T}}) = \operatorname{Tr}(x^{\mathsf{T}}(Mx)) \\
&= \operatorname{Tr}(x^{\mathsf{T}}Mx) = x^{\mathsf{T}}Mx.
\end{aligned}$$

Azaz M pozitív szemidefinit.

A lemma alkalmazása

A lemmából következik, hogy

$$\inf\langle M, X\rangle = \begin{cases} 0, & M \succeq 0 \\ -\infty, & \text{különben} \end{cases}.$$

• Valóban, ha $M \succeq 0$, akkor a lemma alapján $\langle M, X\rangle$ nem lehet negatív, de $M = 0$ mutatja, hogy 0 lehet.

• Szintén a lemma alapján, ha $M \not\succeq 0$, akkor $\langle M, X\rangle$ lehet negatív. X pozitív számmal való szorzásával elérhetjük, hogy tetszőlegesen nagy abszolútértékű negatív szám legyen.

A lemma geometriai jelentése

A következő definíció lehetőséget ad a kitérőbeli lemmánk egy hasznos újrafogalmazására:

Definíció

A $K \subseteq \mathbb{R}^N$ kúp duálisa:

$$K^* = \{x \in \mathbb{R}^N : \ x^{\mathrm{T}} k \geq 0 \ \forall k \in K\}.$$

Ezek után a bebizonyított lemmánk a következő formában is leírható:

Lemma

(i) $\mathcal{S}_+^n$ kúp.

(ii) S_+^n önduális, $(S_+^n)^* = S_+^n$.

Duális feladat

A kitérő után kapjuk, hogy

$$\inf_{X\succeq 0} L(X,\mu) = \begin{cases} -\sum b_i\mu_i, & C+\sum \mu_i A_i \succeq 0 \\ -\infty, & \text{különben} \end{cases}.$$

Azaz a kiinduló SDP (primál) feladat duálisa.

Maximalizáljuk	$-\sum b_i\mu_i$-t
Feltéve, hogy	$-\sum \mu_i A_i \preceq C$

Ekvivalens módon

Minimalizáljuk	$\sum b_i\mu_i$-t
Feltéve, hogy	$C+\sum \mu_i A_i = S$,
	$S \succeq 0$.

Gyenge dualitás tétel

Gyenge dualitás

Ha p^* a primál SDP optimuma és d^* a duális SDP optimuma, akkor $d^* \leq p^*$.

Feltehető, hogy $p^* < \infty$, $d^* > -\infty$.

Legyen $X \in \mathcal{L}_P$, egy lehetséges megoldása a primál feladatnak és $(\mu, S) \in \mathcal{L}_D$, a duális egy lehetséges megoldása.

Állítás

$\langle C, X \rangle \geq -b^\mathsf{T}\mu$.

Az állítás azt mondja, hogy egy tetszőleges primál célfüggvényérték legalább akkora mint egy tetszőleges duál célfüggvényérték. Ebből a tétel nyilván adódik.

A gyenge dualitás bizonyítása

Láttuk, hogy két pozitív szemidefinit mátrix skalárszorzata nemnegatív. Így

$$\begin{aligned} 0 \leq & \langle X, S\rangle = \langle X, C + \sum \mu_i A_i\rangle = \langle X, C\rangle + \sum \mu_i \langle A_i, X\rangle \\ = & \langle C, X\rangle + \sum \mu_i b_i = \langle C, X\rangle + b_i^{\mathsf{T}} \mu_i. \end{aligned}$$

Innen rendezéssel adódik az állítás.

Példa

Minimalizáljuk	x_{12}-t
Feltéve, hogy	$\begin{pmatrix} 0 & x_{12} & 0 \\ x_{12} & x_{22} & 0 \\ 0 & 0 & 1+x_{12} \end{pmatrix} \succeq 0$

Példa: Normálalakra hozás, bevezetés

A duálizáláshoz a feladatot a szemidefinit programozási feladatok I. standard alakjára hozzuk:

Bevezetjük a következő változó mátrixot:

$$X = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{12} & x_{22} & x_{23} \\ x_{13} & x_{23} & x_{33} \end{pmatrix}$$

Az eredeti alak mátrixa $\begin{pmatrix} 0 & x_{12} & 0 \\ x_{12} & x_{22} & 0 \\ 0 & 0 & 1 + x_{12} \end{pmatrix}$.

Ez ekvivalens azzal, hogy $x_{11} = 0$, $x_{13} = 0$, $x_{23} = 0$, $x_{33} = 1 + x_{12}$.

Példa: A feltételek átírása

Ezen feltételek átírása az I. alak formájára:

- $x_{11} = 0$ átírása: $\left\langle \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, X \right\rangle = 0$ (ezen egyenlőség feltételhez tartozzon a μ_1 duális változó).
- $x_{13} = 0$ átírása $\left\langle \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, X \right\rangle = 0$ (ezen egyenlőség feltételhez tartozzon a μ_2 duális változó).
- $x_{23} = 0$ feltétel átírása $\left\langle \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, X \right\rangle = 0$ (ezen egyenlőség feltételhez tartozzon a μ_3 duális változó).
- $x_{33} = 1 + x_{12}$ feltétel átírása $\left\langle \begin{pmatrix} 0 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, X \right\rangle = 1$ (ezen egyenlőség feltételhez tartozzon a μ_4 duális változó).

A példa normálalakban

Feladatunk új alakja

Minimalizáljuk $\left\langle \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, X \right\rangle$-t

Feltéve, hogy $\left\langle \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, X \right\rangle = 0, \left\langle \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, X \right\rangle = 0,$

$$\left\langle \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, X \right\rangle = 0, \left\langle \begin{pmatrix} 0 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, X \right\rangle = 1.$$

$X \succeq 0$

Példa: dualizálás

Maximalizáljuk $-\mu_4$-t

Feltéve, hogy
$$\begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \mu_1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \mu_2 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} +$$
$$\mu_3 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} + \mu_4 \begin{pmatrix} 0 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \succeq 0.$$

Azaz

Maximalizáljuk $-\mu_4$-t

Feltéve, hogy
$$\begin{pmatrix} \mu_1 & \frac{1-\mu_4}{2} & \mu_2 \\ \frac{1-\mu_4}{2} & 0 & \mu_3 \\ \mu_2 & \mu_3 & \mu_4 \end{pmatrix} \succeq 0.$$

p^* elemi módszerekkel

Minimalizáljuk	x_{12}-t
Feltéve, hogy	$\begin{pmatrix} 0 & x_{12} & 0 \\ x_{12} & x_{22} & 0 \\ 0 & 0 & 1+x_{12} \end{pmatrix} \succeq 0$

Ha létezik egy $X \in \mathcal{L}$, akkor a példabeli feltétel mátrixának balfelső 2×2-es főminorának pozitív szemidefinitnek kell lennie. Így a sajátértékeinek szorzata (determinánsa) legyen nagyobb, mint nulla.

Esetünkben $-x_{12}^2 \geqslant 0$, azaz $x_{12} = 0$. Tehát $p^* = 0$.

d^* elemi módszerekkel

Maximalizáljuk	$-\mu_4$-t
Feltéve, hogy	$\begin{pmatrix} \mu_1 & \frac{1-\mu_4}{2} & \mu_2 \\ \frac{1-\mu_4}{2} & 0 & \mu_3 \\ \mu_2 & \mu_3 & \mu_4 \end{pmatrix} \succeq 0.$

d^*: Az előző meggondolások alapján teljesülni kell, hogy
$\det \begin{pmatrix} \mu_1 & \frac{1-\mu_4}{2} \\ \frac{1-\mu_4}{2} & 0 \end{pmatrix} \geq 0$

Ez csak $\mu_4 = 1$ esetén teljesül. Így $d^* = -1$.

Azaz valóban teljesül a gyenge dualitás tétele, tehát $p^* \geqslant d^*$ ($0 \geqslant -1$).

Azonban erős dualitás nincs.

Erős dualitás

Jelölés

$X \in \mathcal{L}_P^+$ akkor és csak akkor teljesül, ha $\langle A_i, X\rangle = b_i, i = 1, \ldots, \ell$ és $X \succ 0$ (azaz X pozitív definit).

Hasonlóan definiálható $\mathcal{L}_D^+$.

Erős dualitás

$\mathcal{L}_P^+, \mathcal{L}_D^+ \neq \emptyset$ esetén $p^* = d^*$.

Sőt az optimumhelyek halmaza nem üres, kompakt

Az erős dualitás egy kissé gyengített feltételek mellett is igaz.

Erős dualitás

$\mathcal{L}_P^+, \mathcal{L}_D \neq \emptyset$ esetén $p^* = d^*$.

A tételeket most nem bizonyítjuk.

Vége van!

Köszönöm a figyelmet!