

# Szemidefinit optimalizálás és az S-lemma

---

Pólik Imre  
SAS Institute, USA

BME Optimalizálás szeminárium  
2011. október 6.



THE  
POWER  
TO KNOW.

# Outline

- 1 Egyenlőtlenségrendszerek megoldhatósága
- 2 Az S-lemma
- 3 Szemidefinit kapcsolatok
- 4 Szemidefinit optimalizálás
- 5 Alkalmazások
- 6 Kutatási irányok

# Egyenlőtlenségrendszerek megoldhatósága

## Feladat

*Honnan tudjuk, hogy az*

$$f(x) < 0$$

$$g_i(x) \leq 0, \quad i = 1, \dots, m$$

*rendszernek nincs megoldása? ( $f, g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ )*

## Példa

$$x_1^2 - x_2^2 + 4x_1x_2 < 0$$

$$x_2^2 - x_1x_2 \leq 0$$

# Egyenlőtlenségrendszerek megoldhatósága

## Feladat

Honnan tudjuk, hogy az

$$f(x) < 0$$

$$g_i(x) \leq 0, \quad i = 1, \dots, m$$

rendszernek nincs megoldása? ( $f, g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ )

## Példa

$$x_1^2 - x_2^2 + 4x_1x_2 < 0$$

$$x_2^2 - x_1x_2 \leq 0$$

## Elégséges feltétel

$$(x_1^2 - x_2^2 + 4x_1x_2) + 2(x_2^2 - x_1x_2) = x_1^2 + 2x_1x_2 + x_2^2 = (x_1 + x_2)^2 \geq 0$$

# Egyenlőtlenségrendszerek megoldhatósága

## Elégséges feltétel

Tetszőleges  $f, g_i$  esetén

$$\begin{aligned} &\exists y \in \mathbb{R}^m, y \geq 0 \\ &f(x) + \sum_{i=1}^m y_i g_i(x) \geq 0, \forall x \in \mathbb{R}^n \end{aligned}$$

⇓

$$\begin{aligned} &\nexists x \in \mathbb{R}^n \\ &f(x) < 0 \\ &g_i(x) \leq 0, i = 1, \dots, m \end{aligned}$$

# Egyenlőtlenségrendszerek megoldhatósága

## Elégséges feltétel

Tetszőleges  $f, g_i$  esetén

$$\exists y \in \mathbb{R}^m, y \geq 0$$
$$f(x) + \sum_{i=1}^m y_i g_i(x) \geq 0, \forall x \in \mathbb{R}^n$$



$$\nexists x \in \mathbb{R}^n$$
$$f(x) < 0$$
$$g_i(x) \leq 0, i = 1, \dots, m$$

# Egyenlőtlenségrendszerek megoldhatósága

## Ekvivalens feltétel

Ha  $f$  és  $g_i$  konvex függvények és  $\exists x \in \mathbb{R}^n : g_i(x) < 0$ , akkor

$$\begin{aligned} &\exists y \in \mathbb{R}^m, y \geq 0 \\ &f(x) + \sum_{i=1}^m y_i g_i(x) \geq 0, \forall x \in \mathbb{R}^n \end{aligned}$$



$$\begin{aligned} &\nexists x \in \mathbb{R}^n \\ &f(x) < 0 \\ &g_i(x) \leq 0, i = 1, \dots, m \end{aligned}$$

# Az S-lemma

Yakubovich (1971)

Ha  $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$  kvadratikus függvények és  $\exists x \in \mathbb{R}^n : g(x) < 0$ ,  
akkor

$$\begin{aligned} &\exists y \geq 0 \\ &f(x) + y \cdot g(x) \geq 0, \forall x \in \mathbb{R}^n \end{aligned}$$



$$\begin{aligned} &\nexists x \in \mathbb{R}^n \\ &f(x) < 0 \\ &g(x) \leq 0 \end{aligned}$$

Konvexitás nélkül!

## Kvadratikus függvény (homogén)

$$f(x) = x^T A x, A \in \mathbb{R}^{n \times n} \text{ szimmetrikus}$$

## Konvex kvadratikus függvény

Ha  $A \succeq 0$  (pozitív szemidefinit), vagyis  $x^T A x \geq 0, \forall x \in \mathbb{R}^n$ .

## Mátrixok skalárszorzata

$$U \bullet V = \text{Tr}(UV) = \sum_{i,j=1}^n U_{ij} V_{ij}$$

$$x^T A x = \text{Tr}(x^T A x) = \text{Tr}(A x x^T) = \text{Tr}(A(x x^T)) = A \bullet x x^T$$

# Az S-lemma, homogén alak

Yakubovich (1971)

$A, B \in \mathbb{R}^{n \times n}$  és  $\exists x \in \mathbb{R}^n : x^T Bx < 0$ , akkor

$$\begin{aligned} \exists y \geq 0 \\ x^T Ax + y \cdot x^T Bx \geq 0, \forall x \in \mathbb{R}^n \end{aligned}$$



$$\begin{aligned} \nexists x \in \mathbb{R}^n \\ x^T Ax < 0 \\ x^T Bx \leq 0 \end{aligned}$$

# Az S-lemma, homogén alak

Yakubovich (1971)

$A, B \in \mathbb{R}^{n \times n}$  és  $\exists x \in \mathbb{R}^n : x^T B x < 0$ , akkor

$$\exists y \geq 0$$

$$A + yB \succeq 0 \text{ (PSD)}$$



$$\nexists x \in \mathbb{R}^n$$

$$x^T A x < 0$$

$$x^T B x \leq 0$$

# Az S-lemma

- Miért?
- Konvexitás nélkül!
  - » Rejtett konvexitás
- Alkalmazások
  - » Ljapunov-féle stabilitásvizsgálat
  - » Ellipszoidtartalmazás
  - » Számítógépes grafika

# Klasszikus bizonyítás

- A primál feladat nem megoldható

$$\begin{aligned} & \nexists x : x^T A x < 0, x^T B x \leq 0 \\ \Leftrightarrow \mathbb{R}_- \times \mathbb{R}_\ominus \cap \underbrace{\{(x^T A x, x^T B x) : x \in \mathbb{R}^n\}}_{\text{konvex! (Dines, 1941)}} = \emptyset \end{aligned}$$

# Klasszikus bizonyítás

- A primál feladat nem megoldható

$$\nexists x : x^T A x < 0, x^T B x \leq 0$$
$$\Leftrightarrow \mathbb{R}_- \times \mathbb{R}_\ominus \cap \underbrace{\{(x^T A x, x^T B x) : x \in \mathbb{R}^n\}}_{\text{konvex! (Dines, 1941)}} = \emptyset$$

- Kicsit általánosabb eredmény  
(Poljak, 1998)  $n \geq 3$ , az  $A, B_1, B_2$  mátrixoknak van PD lineáris kombinációjuk

$$\Rightarrow \{(x^T A x, x^T B_1 x, x^T B_2 x) : x \in \mathbb{R}^n\} \quad \text{konvex}$$

# Klasszikus bizonyítás

- A primál feladat nem megoldható

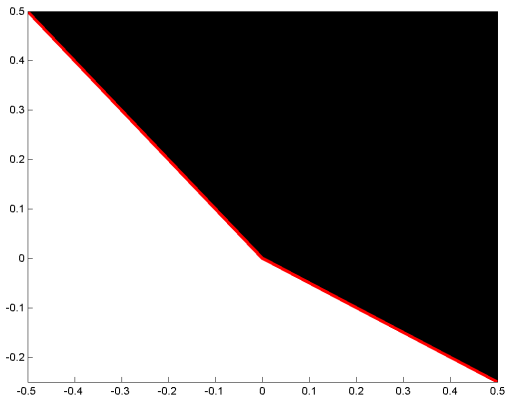
$$\nexists x : x^T A x < 0, x^T B x \leq 0$$
$$\Leftrightarrow \mathbb{R}_- \times \mathbb{R}_\ominus \cap \underbrace{\{(x^T A x, x^T B x) : x \in \mathbb{R}^n\}}_{\text{konvex! (Dines, 1941)}} = \emptyset$$

- Kicsit általánosabb eredmény  
(Poljak, 1998)  $n \geq 3$ , az  $A, B_1, B_2$  mátrixoknak van PD lineáris kombinációjuk

$$\Rightarrow \{(x^T A x, x^T B_1 x, x^T B_2 x) : x \in \mathbb{R}^n\} \quad \text{konvex}$$

- Szeparációs bizonyítás
- Norma-feltétel

Figure:  $\{(x^T Ax, x^T Bx) : x \in \mathbb{R}^n\}$



$$A = \begin{pmatrix} -2 & 0 \\ 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix}$$

Vissza a konvexitáshoz!

## Szemidefinit relaxáció

$$x^T A x < 0$$

$$x^T B x \leq 0$$

$$x \in \mathbb{R}^n$$

## Szemidefinit relaxáció

$$\begin{aligned}x^T A x &< 0 & A \bullet x x^T &< 0 \\x^T B x &\leq 0 & B \bullet x x^T &\leq 0 \\x &\in \mathbb{R}^n\end{aligned}$$

## Szemidefinit relaxáció

$$\begin{array}{llll} x^T A x < 0 & & A \bullet x x^T < 0 & & A \bullet X < 0 \\ x^T B x \leq 0 & \Leftrightarrow & B \bullet x x^T \leq 0 & \Leftrightarrow & B \bullet X \leq 0 \\ x \in \mathbb{R}^n & & & & \text{rank}(X) = 1 \\ & & & & X \succeq 0 \end{array}$$

# Modern bizonyítás

## Szemidefinit relaxáció

$$\begin{array}{l} x^T A x < 0 \\ x^T B x \leq 0 \\ x \in \mathbb{R}^n \end{array} \Leftrightarrow \begin{array}{l} A \bullet x x^T < 0 \\ B \bullet x x^T \leq 0 \end{array} \Leftrightarrow \begin{array}{l} A \bullet X < 0 \\ B \bullet X \leq 0 \\ \text{rank}(X) = 1 \\ X \succeq 0 \end{array}$$

# Modern bizonyítás

## Szemidefinit relaxáció

$$\begin{array}{llll} x^T A x < 0 & A \bullet x x^T < 0 & A \bullet X < 0 \\ x^T B x \leq 0 & B \bullet x x^T \leq 0 & B \bullet X \leq 0 \\ x \in \mathbb{R}^n & & \text{rank}(X) = 1 \\ & & X \succeq 0 \end{array}$$

## Pataki, 1998

$\mathcal{A} \subseteq \mathbb{S}^n$  affin altér,  $\dim \mathcal{A} \geq \binom{n}{2} - \binom{r+2}{2} + 1$ ,  $\mathbb{P}\mathbb{S}^n \cap \mathcal{A} \neq \emptyset$   
 $\Rightarrow \exists X \in \mathbb{P}\mathbb{S}^n \cap \mathcal{A}$ , amelyre  $\text{rank}(X) \leq r$ .

# Modern bizonyítás

## Szemidefinit relaxáció

$$\begin{array}{llll} x^T A x < 0 & A \bullet x x^T < 0 & A \bullet X < 0 \\ x^T B x \leq 0 & B \bullet x x^T \leq 0 & B \bullet X \leq 0 \\ x \in \mathbb{R}^n & & \text{rank}(X) = 1 \\ & & X \succeq 0 \end{array}$$

## Pataki, 1998

$\mathcal{A} \subseteq \mathbb{S}^n$  affin altér,  $\dim \mathcal{A} \geq \binom{n}{2} - \binom{r+2}{2} + 1$ ,  $\mathbb{P}\mathbb{S}^n \cap \mathcal{A} \neq \emptyset$   
 $\Rightarrow \exists X \in \mathbb{P}\mathbb{S}^n \cap \mathcal{A}$ , amelyre  $\text{rank}(X) \leq r$ .

## Barvinok, 2001

$\mathcal{A} \subseteq \mathbb{S}^n$  affin altér,  $\dim \mathcal{A} = \binom{n}{2} - \binom{r+2}{2}$ ,  $\mathbb{P}\mathbb{S}^n \cap \mathcal{A} \neq \emptyset$  és korlátos  
 $\Rightarrow \exists X \in \mathbb{P}\mathbb{S}^n \cap \mathcal{A}$ , amelyre  $\text{rank}(X) \leq r$ .

# A rangfeltétel és a konvexitás ekvivalenciája

- Az  $\{(x^T Ax, x^T Bx) : x \in \mathbb{R}^n\}$  halmaz konvexitása
  - »  $y, z \in \mathbb{R}^n, \lambda \in [0, 1]$
  - » Kell:  $x \in \mathbb{R}^n$

$$x^T Ax = \lambda y^T Ay + (1 - \lambda) z^T Az$$

$$x^T Bx = \lambda y^T By + (1 - \lambda) z^T Bz$$

# A rangfeltétel és a konvexitás ekvivalenciája

- Az  $\{(x^T Ax, x^T Bx) : x \in \mathbb{R}^n\}$  halmaz konvexitása
  - »  $y, z \in \mathbb{R}^n, \lambda \in [0, 1]$
  - » Kell:  $x \in \mathbb{R}^n$

$$\begin{aligned}x^T Ax &= \lambda y^T Ay + (1 - \lambda) z^T Az \\x^T Bx &= \lambda y^T By + (1 - \lambda) z^T Bz\end{aligned}$$

- $X = xx^T$  a következő rendszer 1-rangú megoldása

$$\begin{aligned}A \bullet X &= \lambda y^T Ay + (1 - \lambda) z^T Az \\B \bullet X &= \lambda y^T By + (1 - \lambda) z^T Bz\end{aligned}$$

- Pataki: létezik 1-rangú megoldás

# Bizonyítás Helly-tétellel

$$H_x = \{y \geq 0 : x^T Ax + y \cdot x^T Bx \geq 0\} \subseteq \mathbb{R}$$

## $H_x$ tulajdonságai

- konvex
- zárt
- bármelyik kettő metszete nemüres

$$\implies \bigcap_x H_x \neq \emptyset, \text{ vagyis}$$

$$\exists y \geq 0 : x^T Ax + y \cdot x^T Bx \geq 0, \forall x \in \mathbb{R}^n$$

# Bizonyítás Helly-tétellel

$$H_x = \{y \geq 0 : x^T Ax + y \cdot x^T Bx \geq 0\} \subseteq \mathbb{R}$$

## $H_x$ tulajdonságai

- konvex
- zárt
- bármelyik kettő metszete nemüres
- van köztük korlátos! (Slater-feltétel)

$$\implies \bigcap_x H_x \neq \emptyset, \text{ vagyis}$$

$$\exists y \geq 0 : x^T Ax + y \cdot x^T Bx \geq 0, \forall x \in \mathbb{R}^n$$

Yuan, 1990

$A, B \in \mathbb{R}^{n \times n}$  szimmetrikus mátrixok,  $\mathcal{F}, \mathcal{G} \subseteq \mathbb{R}^n$  zárt halmazok,  $\mathcal{F} \cup \mathcal{G} = \mathbb{R}^n$ . Ha

$$x^T A x \geq 0, x \in \mathcal{F}$$

$$x^T B x \geq 0, x \in \mathcal{G},$$

akkor  $\exists \lambda \in [0, 1]$ , amelyre  $\lambda x^T A x + (1 - \lambda)x^T B x \geq 0, \forall x$ , vagyis  $\lambda A + (1 - \lambda)B \succeq 0$ .

- Általánosítás
  - » több egyenlet
  - » speciális mátrixok
  - » speciális egyenlőtlenségek
- Alkalmazások

# Szemidefinit optimalizálás

## Mátrixváltozó

$$\min \operatorname{Tr}(CX)$$

$$\operatorname{Tr}(A_i X) = b_i, i = 1, \dots, m$$

$$X \succeq 0$$

$$\max b^T y$$

$$\sum_{i=1}^m A_i y_i + S = C$$

$$S \succeq 0$$

$C, X, S, A_i$   $n \times n$ -es szimmetrikus mátrixok,  $b, y \in \mathbb{R}^m$

Speciális struktúra:  $A_i, C$  lehet ritka, vagy alacsony rangú

- Általában belsőpontos módszerek
- Iterációk:  $\mathcal{O}(\sqrt{n})$ , valójában 50 – 100
- Egy iteráció költsége:  $\mathcal{O}(mn^3 + m^2n^2 + m^3)$
- Megoldható feladatok:  $m \leq 10000$ ,  $n \leq 10000$   
(ritka mátrixokkal több)
- Nagy pontosság

# Implementáció

- Kezdőpont
  - » beágyazás
  - » nem-megengedett módszerek
- Műveletek ritka mátrixokkal
  - » tárolás, szimmetria
  - »  $UV + VU, U + uu^T$
- $Ux = r$  megoldása
  - » Cholesky-faktorizáció:  $U = LDL^T$
  - » Iteratív módszerek
- Speciális struktúrák
  - » általános decompozíció (Kojima et al.)
  - » egyedi módszerek adott feladatra

# Bináris változók relaxációja

Bináris változók:  $x_i \in \{0, 1\}$

Lineáris relaxáció:  $x_i \in [0, 1]$

Bináris feltétel ekvivalens alakja:

$$z_i = 2x_i - 1$$

$$z_i^2 = 1 (\Leftrightarrow z_i = \pm 1)$$

Matrixokkal:

$$Z \succeq 0$$

$$\text{diag}(Z) = 1$$

$$\text{rank}(Z) = 1 (\Leftrightarrow Z = zz^T)$$

# Gráfpartícionálás

Egy  $2m$  csúcsú élsúlyozott gráf csúcsait osszuk fel két **egyenlő** részre úgy, hogy a két partíció között futó élek összszúlya minimális legyen.

$A$ : incidencia mátrix,  $A_{kl}$ : a  $kl$  él súlya

$y_{ij} = 1$ : az  $i$  csúcs a  $j$  partícióban van ( $j = 1, 2$ )

$y_j$ : a  $j$  partíció indikátorvektora

$y_j^T A y_j$ :  $2 \times$  a  $j$  partícióban lévő élek összszúlya

$\text{Tr}(Y^T A Y)$ :  $2 \times$  az elvágatlan élek összszúlya

$e^T A e$ :  $2 \times$  az élek összszúlya

$$\min e^T A e - \text{Tr}(Y^T A Y)$$

$Y$  partíciómátrix

$$\min e^T A e - \text{Tr}(A X)$$

$$\text{diag}(X) = 1$$

$$X e = m$$

$$X \succeq 0$$

$$X \preceq 0$$

$$\text{rank}(X) = 2$$

SDP relaxáció ( $X = Y Y^T$ )  $\Rightarrow$

# Gráfpartícionálás

Egy  $2m$  csúcsú élsúlyozott gráf csúcsait osszuk fel két **egyenlő** részre úgy, hogy a két partíció között futó élek összszúlya minimális legyen.

$A$ : incidencia mátrix,  $A_{kl}$ : a  $kl$  él súlya

$y_{ij} = 1$ : az  $i$  csúcs a  $j$  partícióban van ( $j = 1, 2$ )

$y_j$ : a  $j$  partíció indikátorvektora

$y_j^T A y_j$ :  $2 \times$  a  $j$  partícióban lévő élek összszúlya

$\text{Tr}(Y^T A Y)$ :  $2 \times$  az elvágatlan élek összszúlya

$e^T A e$ :  $2 \times$  az élek összszúlya

$$\min e^T A e - \text{Tr}(Y^T A Y)$$

$Y$  partíciómátrix

$$\min e^T A e - \text{Tr}(A X)$$

$$\text{diag}(X) = 1$$

$$X e = m$$

$$X \succeq 0$$

$$X \preceq 0$$

$$\text{rank}(X) = 2$$

SDP relaxáció ( $X = Y Y^T$ )  $\Rightarrow$

Komplexitás:  $\mathcal{O}(m^{6.5})!$

# Polinomoptimalizálás I

## Tétel

Ha  $p(x) : \mathbb{R} \rightarrow \mathbb{R}$  egyváltozós polinom, akkor  
 $p(x) \geq 0, \forall x \Leftrightarrow p(x)$  négyzetösszeg (SOS)

## Példa

$$p(x) = x^6 - 5x^4 + 6x^3 + 8x^2 - 14x + 5$$

# Polinomoptimalizálás I

## Tétel

Ha  $p(x) : \mathbb{R} \rightarrow \mathbb{R}$  egyváltozós polinom, akkor  
 $p(x) \geq 0, \forall x \Leftrightarrow p(x)$  négyzetösszeg (SOS)

## Példa

$$p(x) = x^6 - 5x^4 + 6x^3 + 8x^2 - 14x + 5 = (x^2 + x - 1)^2 + (x^3 - 3x + 2)^2$$

# Polinomoptimalizálás I

## Tétel

Ha  $p(x) : \mathbb{R} \rightarrow \mathbb{R}$  egyváltozós polinom, akkor  
 $p(x) \geq 0, \forall x \Leftrightarrow p(x)$  négyzetösszeg (SOS)

## Példa

$p(x) = x^6 - 5x^4 + 6x^3 + 8x^2 - 14x + 5 = (x^2 + x - 1)^2 + (x^3 - 3x + 2)^2$   
Általában nem igaz:  $z^6 + x^4y^2 + x^2y^4 - 3x^2y^2z^2 \geq 0$ , de nem SOS



## Polinomoptimalizálás II

$$q = (1, x, x^2, \dots, x^n)$$

Négyzet:

$$p(x) = \left( \sum_{i=0}^n u_i x^i \right)^2 = \left( \sum_{i=0}^n u_i q_i \right)^2 = (u^T q)^2 = q^T (u u^T) q$$

SOS:  $q^T U q$ , ahol  $U \succeq 0$

$$u_{44} = 1$$

$$u_{34} + u_{43} = 0$$

$$u_{24} + u_{33} + u_{42} = -5$$

$$u_{14} + u_{23} + u_{32} + u_{41} = 6$$

$$u_{13} + u_{22} + u_{31} = 8$$

$$u_{21} + u_{12} = -14$$

$$u_{11} = 5$$

$$U \succeq 0$$

$$U = \begin{pmatrix} 5 & -7 & -1 & 2 \\ -7 & 10 & 1 & -3 \\ -1 & 1 & 1 & 0 \\ 2 & -3 & 0 & 1 \end{pmatrix}$$

$$u^{(1)} = \begin{pmatrix} -1 & 1 & 1 & 0 \end{pmatrix}$$

$$u^{(2)} = \begin{pmatrix} 2 & -3 & 0 & 1 \end{pmatrix}$$

Software: (Gloptipoly, SOSTools), Yalmip

# Kutatási irányok

- 1994 óta nincs lényeges eredmény
- Speciális struktúrák
- Új algoritmusok
  - » szimplex, perceptron, gravity, megengedett irányok, row-by-row, ...
- Kapcsolódó kutatások
  - » SOCP
  - » Kopozitív optimalizálás ( $x^T U x \geq 0, \forall x \geq 0$ )
- Egészértékű és bináris változók