Optimalizace

12. Konvexní funkce a konvexní optimalizace

Tom Werner, Tom Kroupa

FEL ČVUT

Konvexní funkce

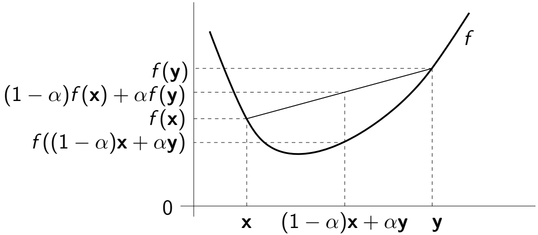
Definice

Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina.

Funkce $f : \mathbb{R}^n \to \mathbb{R}$ je **konvexní** na X, jestliže pro každá $\mathbf{x}, \mathbf{y} \in X$ a $\alpha \in [0, 1]$ je

$$f((1-\alpha)\mathbf{x} + \alpha\mathbf{y}) \leq (1-\alpha)f(\mathbf{x}) + \alpha f(\mathbf{y}).$$

Funkce f je **konkávní** na X, jestliže funkce $-f$ je konvexní na X.



Konvexní funkce

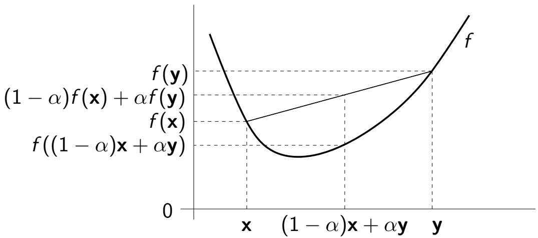
Definice

Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina.

Funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ je **konvexní** na X, jestliže pro každá $\mathbf{x}, \mathbf{y} \in X$ a $\alpha \in [0,1]$ je

$$f((1-\alpha)\mathbf{x} + \alpha\mathbf{y}) \leq (1-\alpha)f(\mathbf{x}) + \alpha f(\mathbf{y}).$$

Funkce f je **konkávní** na X, jestliže funkce $-f$ je konvexní na X.



Ekvivalentní podmínka (**Jensenova nerovnost**):

Pro každé $\mathbf{x}_1, \ldots, \mathbf{x}_k \in X$ a $\alpha_1, \ldots, \alpha_k \geq 0$ kde $\alpha_1 + \cdots + \alpha_k = 1$ je

$$f(\alpha_1\mathbf{x}_1 + \cdots + \alpha_k\mathbf{x}_k) \leq \alpha_1 f(\mathbf{x}_1) + \cdots + \alpha_k f(\mathbf{x}_k).$$

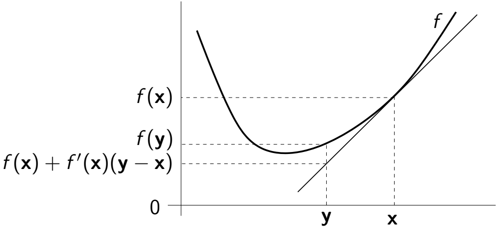
Podmínka konvexity pro diferencovatelné funkce

Věta (podmínka prvního řádu na konvexitu funkce)

Nechť funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ je diferencovatelná.

Funkce f je konvexní, právě když pro všechna $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ platí

$$f(\mathbf{y}) \geq f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{x} - \mathbf{y}).$$



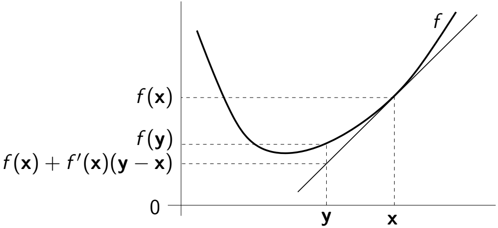
Podmínka konvexity pro diferencovatelné funkce

Věta (podmínka prvního řádu na konvexitu funkce)

Nechť funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ je diferencovatelná.

Funkce f je konvexní, právě když pro všechna $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ platí

$$f(\mathbf{y}) \geq f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{x} - \mathbf{y}).$$



Věta (podmínka druhého řádu na konvexitu funkce)

Nechť $f\colon \mathbb{R}^n \to \mathbb{R}$ je dvakrát diferencovatelná. Funkce f je konvexní, právě když je pro každé $\mathbf{x} \in \mathbb{R}^n$ Hessián $f''(\mathbf{x})$ pozitivně semidefinitní.

Příklady konvexních funkcí

- $f(x) = e^{ax}$ pro $a \in \mathbb{R}$
- $f(x) = |x|^a$ pro $a \geq 1$
- $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ pro $\mathbf{A}$ pozitivně semidefinitní
- $f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} + b$
- $f(\mathbf{x}) = \max\{x_1, \ldots, x_n\}$
- libovolná norma

Vztah konvexní funkce a konvexní množiny

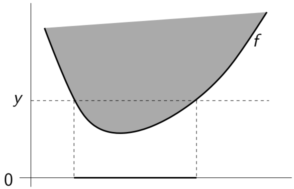
- **Epigraf** funkce $f\colon X \to \mathbb{R}$ je množina $\{\, (\mathbf{x}, y) \in X \times \mathbb{R} \mid f(\mathbf{x}) \leq y \,\}$.
- **Subkontura** výšky y funkce f je množina $\{\, \mathbf{x} \in X \mid f(\mathbf{x}) \leq y \,\}$.

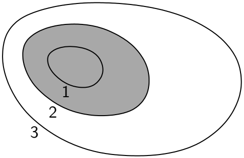
Vztah konvexní funkce a konvexní množiny

- **Epigraf** funkce $f: X \to \mathbb{R}$ je množina $\{\, (\mathbf{x}, y) \in X \times \mathbb{R} \mid f(\mathbf{x}) \le y \,\}$.
- **Subkontura** výšky y funkce f je množina $\{\, \mathbf{x} \in X \mid f(\mathbf{x}) \le y \,\}$.

Tvrzení

- Funkce je konvexní, právě když její epigraf je konvexní množina.
- Každá subkontura konvexní funkce je konvexní množina.





Pozor: Jsou-li všechny subkontury funkce konvexní množiny, funkce nemusí být konvexní.

Operace zachovávající konvexitu

Nezáporné lineární kombinace

Jsou-li $g_1, \ldots, g_k \colon \mathbb{R}^n \to \mathbb{R}$ konvexní funkce a $\alpha_1, \ldots, \alpha_k \geq 0$, pak funkce

$$f = \alpha_1 g_1 + \cdots + \alpha_k g_k$$

je konvexní.

Operace zachovávající konvexitu

Nezáporné lineární kombinace

Jsou-li $g_1, \ldots, g_k \colon \mathbb{R}^n \to \mathbb{R}$ konvexní funkce a $\alpha_1, \ldots, \alpha_k \geq 0$, pak funkce

$$f = \alpha_1 g_1 + \cdots + \alpha_k g_k$$

je konvexní.

Skládání funkcí

Následující funkce jsou konvexní:

- $h = g \circ f$, kde $f \colon \mathbb{R} \to \mathbb{R}$ je konvexní a $g \colon \mathbb{R} \to \mathbb{R}$ je konvexní a neklesající.
- $h(\mathbf{x}) = g(\mathbf{A}\mathbf{x} + \mathbf{b})$, kde $g \colon \mathbb{R}^m \to \mathbb{R}$ je konvexní.

Pozor: Složení konvexních funkcí obecně nemusí být konvexní funkce.

Maximum

Věta

Nechť $g_i \colon \mathbb{R}^n \to \mathbb{R}$ jsou konvexní funkce pro všechna $i \in I$. Pak

$$f(\mathbf{x}) = \max_{i \in I} g_i(\mathbf{x})$$

je konvexní funkce (předpokládáme, že pro každé $\mathbf{x} \in \mathbb{R}^n$ maximum existuje).

Důkaz: Ukážeme (není těžké), že epigraf funkce f je průnik epigrafů funkcí g_i.
Z toho plyne, že epigraf funkce f je konvexní množina. Tedy f je konvexní funkce.

Věta platí, i když nahradíme maximum supremem.

Maximum

Věta

Nechť $g_i : \mathbb{R}^n \to \mathbb{R}$ jsou konvexní funkce pro všechna $i \in I$. Pak

$$f(\mathbf{x}) = \max_{i \in I} g_i(\mathbf{x})$$

je konvexní funkce (předpokládáme, že pro každé $\mathbf{x} \in \mathbb{R}^n$ maximum existuje).

Důkaz: Ukážeme (není těžké), že epigraf funkce f je průnik epigrafů funkcí g_i.
Z toho plyne, že epigraf funkce f je konvexní množina. Tedy f je konvexní funkce.

Věta platí, i když nahradíme maximum supremem.

Příklady (funkce jsou konvexní na $\mathbb{R}^n$):

- $f(\mathbf{x}) = \max\{x_1, \ldots, x_n\}$
- $f(\mathbf{x}) = \max_{i=1}^{k}(\mathbf{a}_i^T \mathbf{x} + b_i)$
- $f(\mathbf{y}) = \max_{\mathbf{x} \in X} \|\mathbf{x} - \mathbf{y}\|$ (kde $X \subseteq \mathbb{R}^n$ je libovolná kompaktní množina)
- $f(\mathbf{c}) = \max_{\mathbf{x} \in X} \mathbf{c}^T \mathbf{x}$ (kde $X \subseteq \mathbb{R}^n$ je libovolná kompaktní množina)

Konvexní optimalizační úlohy

Konvexní optimalizační úloha

Konvexní optimalizační úloha je

$$\min_{\mathbf{x} \in X} f(\mathbf{x})$$

kde $X \subseteq \mathbb{R}^n$ je konvexní množina a $f\colon \mathbb{R}^n \to \mathbb{R}$ je konvexní funkce na množině X.

Konvexní optimalizační úloha

Konvexní optimalizační úloha je

$$\min_{\mathbf{x} \in X} f(\mathbf{x})$$

kde $X \subseteq \mathbb{R}^n$ je konvexní množina a $f\colon \mathbb{R}^n \to \mathbb{R}$ je konvexní funkce na množině X.

Věta

Nechť funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ je konvexní na konvexní množině $X \subseteq \mathbb{R}^n$.
Pak každé lokální minimum funkce f na množině X je globální.

Tedy konvexní úlohu vyřešíme nalezením libovolného lokálního minima!

Konvexní optimalizační úloha ve standardním tvaru

Tvrzení

Úloha

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{za podmínek} \quad & g_i(\mathbf{x}) \le 0, \quad i = 1, \ldots, m \\ & h_i(\mathbf{x}) = 0, \quad i = 1, \ldots, \ell \\ & \mathbf{x} \in \mathbb{R}^n \end{aligned}$$

je konvexní, jestliže funkce $f, g_1, \ldots, g_m$ jsou konvexní a $h_1, \ldots, h_\ell$ jsou afinní.

Důkaz: Množina přípustných řešení je průnik konvexních množin:

$$X = \bigcap_{i=1}^{m} \underbrace{\{\, \mathbf{x} \in \mathbb{R}^n \mid g_i(\mathbf{x}) \le 0 \,\}}_{\text{subkontura konv. fce}} \cap \bigcap_{i=1}^{\ell} \underbrace{\{\, \mathbf{x} \in \mathbb{R}^n \mid h_i(\mathbf{x}) = \mathbf{0} \,\}}_{\text{nadrovina}}$$

Tato podmínka je postačující ale ne nutná pro konvexitu úlohy.

Příklad: Množina

$$X = \{\, (x, y) \in \mathbb{R}^2 \mid (x + y)^2 = 0 \,\}$$

je konvexní, ale $h(x, y) = (x + y)^2$ není afinní.

Třídy konvexních optimalizačních úloh

- Lineární programování (LP):
  f, g_i, h_i afinní

- Kvadratické programování (QP):
  f kvadratická konvexní, g_i, h_i afinní

- Kvadratické programování s kvadratickými omezeními (QCQP):
  f, g_i kvadratické konvexní, h_i afinní

- Programování na kuželu druhého řádu (SOCP)

- Semidefinitní progamování (SDP)

Kvadratické programování (QP)

Funkce f je kvadratická, funkce g_i, h_i jsou afinní.

$$\begin{aligned} \min \quad & \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} \\ \text{za podm.} \quad & \mathbf{C}\mathbf{x} \leq \mathbf{d} \\ & \mathbf{E}\mathbf{x} = \mathbf{f} \end{aligned}$$

Je to konvexní úloha, právě když $\mathbf{A}$ je pozitivně semidefinitní.

Kvadratické programování (QP)

Funkce f je kvadratická, funkce g_i, h_i jsou afinní.

$$\begin{aligned} \min \quad & \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} \\ \text{za podm.} \quad & \mathbf{C}\mathbf{x} \leq \mathbf{d} \\ & \mathbf{E}\mathbf{x} = \mathbf{f} \end{aligned}$$

Je to konvexní úloha, právě když $\mathbf{A}$ je pozitivně semidefinitní.

Speciální případ: chybí podmínky typu nerovnosti ($\mathbf{C}, \mathbf{d}$ nulové):
podmínka optimality prvního řádu je soustava lineárních rovnic.

Kvadratické programování (QP)

Funkce f je kvadratická, funkce g_i, h_i jsou afinní.

$$\begin{aligned} \min \quad & \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} \\ \text{za podm.} \quad & \mathbf{C}\mathbf{x} \leq \mathbf{d} \\ & \mathbf{E}\mathbf{x} = \mathbf{f} \end{aligned}$$

Je to konvexní úloha, právě když $\mathbf{A}$ je pozitivně semidefinitní.

Speciální případ: chybí podmínky typu nerovnosti ($\mathbf{C}, \mathbf{d}$ nulové): podmínka optimality prvního řádu je soustava lineárních rovnic.

Příklad: Řešení přeurčené soustavy s omezeními typu lineárních rovností:

$$\min\{\, \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 \mid \mathbf{C}\mathbf{x} = \mathbf{d},\, \mathbf{x} \in \mathbb{R}^n \,\}$$

Kvadratické programování (QP)

Funkce f je kvadratická, funkce g_i, h_i jsou afinní.

$$\begin{aligned} \min \quad & \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} \\ \text{za podm.} \quad & \mathbf{C}\mathbf{x} \leq \mathbf{d} \\ & \mathbf{E}\mathbf{x} = \mathbf{f} \end{aligned}$$

Je to konvexní úloha, právě když $\mathbf{A}$ je pozitivně semidefinitní.

Speciální případ: chybí podmínky typu nerovnosti ($\mathbf{C}, \mathbf{d}$ nulové): podmínka optimality prvního řádu je soustava lineárních rovnic.

Příklad: Řešení přeurčené soustavy s omezeními typu lineárních rovností:

$$\min\{\, \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 \mid \mathbf{C}\mathbf{x} = \mathbf{d},\, \mathbf{x} \in \mathbb{R}^n \,\}$$

Příklad pro nenulové $\mathbf{C}, \mathbf{d}$: Řešení přeurčené soustavy s intervalovými omezeními

$$\min\{\, \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 \mid \mathbf{c} \leq \mathbf{x} \leq \mathbf{d},\, \mathbf{x} \in \mathbb{R}^n \,\}$$

Příklad na QP: Support Vector Machine (SVM)

Pro dané $(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_m, y_m) \in \mathbb{R}^n \times \{-1, +1\}$ hledáme **oddělující nadrovinu**, tedy hledáme $\mathbf{a} \in \mathbb{R}^n$ a $b \in \mathbb{R}$ tak, aby

$$
\begin{aligned}
\mathbf{a}^T\mathbf{x}_i > b &\quad \text{pro } y_i = +1 \\
\mathbf{a}^T\mathbf{x}_i < b &\quad \text{pro } y_i = -1
\end{aligned}
$$

neboli

$$
y_i(\mathbf{a}^T\mathbf{x}_i - b) > 0 \quad i = 1, \ldots, m
$$

Příklad na QP: Support Vector Machine (SVM)

Pro dané $(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_m, y_m) \in \mathbb{R}^n \times \{-1, +1\}$ hledáme **oddělující nadrovinu**, tedy hledáme $\mathbf{a} \in \mathbb{R}^n$ a $b \in \mathbb{R}$ tak, aby

$$\begin{aligned} \mathbf{a}^T\mathbf{x}_i &> b \quad \text{pro } y_i = +1 \\ \mathbf{a}^T\mathbf{x}_i &< b \quad \text{pro } y_i = -1 \end{aligned}$$

neboli

$$y_i(\mathbf{a}^T\mathbf{x}_i - b) > 0 \quad i = 1, \ldots, m$$

Vydělením $(\mathbf{a}, b)$ vhodným kladným číslem je toto ekvivalentní

$$y_i(\mathbf{a}^T\mathbf{x}_i - b) \geq 1, \quad i = 1, \ldots, m \tag{$\star$}$$

Příklad na QP: Support Vector Machine (SVM)

Pro dané $(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_m, y_m) \in \mathbb{R}^n \times \{-1, +1\}$ hledáme **oddělující nadrovinu**, tedy hledáme $\mathbf{a} \in \mathbb{R}^n$ a $b \in \mathbb{R}$ tak, aby

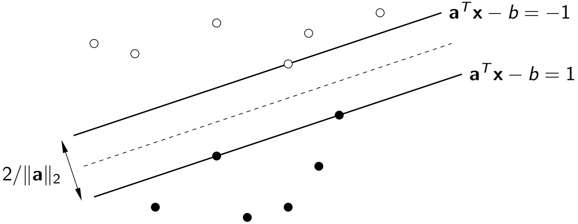
$$\begin{aligned} \mathbf{a}^T\mathbf{x}_i > b \quad & \text{pro } y_i = +1 \\ \mathbf{a}^T\mathbf{x}_i < b \quad & \text{pro } y_i = -1 \end{aligned}$$

neboli

$$y_i(\mathbf{a}^T\mathbf{x}_i - b) > 0 \quad i = 1, \ldots, m$$

Vydělením $(\mathbf{a}, b)$ vhodným kladným číslem je toto ekvivalentní

$$y_i(\mathbf{a}^T\mathbf{x}_i - b) \geq 1, \quad i = 1, \ldots, m \qquad (\star)$$



Aby šířka pásu byla co největší, minimalizujeme $\|\mathbf{a}\|_2^2 = \mathbf{a}^T\mathbf{a}$ za podmínek $(\star)$.

Kvadratické programování s kvadratickými omezeními (QCQP)

Funkce f, g_i jsou kvadratické, h_i jsou afinní.

Je to konvexní úloha, právě když funkce f, g_i jsou konvexní.

Kvadratické programování s kvadratickými omezeními (QCQP)

Funkce f, g_i jsou kvadratické, h_i jsou afinní.
Je to konvexní úloha, právě když funkce f, g_i jsou konvexní.

Příklad: Nejmenší koule obsahující zadané body $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$:

$$\min_{\mathbf{x} \in \mathbb{R}^n} \max_{i=1}^{m} \|\mathbf{x} - \mathbf{a}_i\|_2$$

Je ekvivalentní QCQP úloze

$$\begin{aligned} \min \quad & y \\ \text{za podm.} \quad & \|\mathbf{x} - \mathbf{a}_i\|_2^2 \leq y \end{aligned}$$

Programování na kuželu druhého řádu (SOCP)

$f, h_1, \ldots, h_\ell$ afinní a

$$g_i(\mathbf{x}) = \|\mathbf{A}_i\mathbf{x} + \mathbf{b}_i\|_2 - (\mathbf{c}_i^T\mathbf{x} + d_i)$$

Podmínka $g_i(\mathbf{x}) \leq 0$ se dá psát jako

$$\begin{bmatrix} \mathbf{A}_i \\ \mathbf{c}_i \end{bmatrix} \mathbf{x} + \begin{bmatrix} \mathbf{b}_i \\ d_i \end{bmatrix} \in K_2^{m_i}$$

kde $\mathbf{A}_i \in \mathbb{R}^{m_i \times n}$ a

$$K_2^{m_i} = \{\, (\mathbf{x}, y) \in \mathbb{R}^{m_i+1} \mid \|\mathbf{x}\|_2 \leq y \,\}$$

je epigraf eukleidovské normy $\|\cdot\|_2$ (**kužel druhého řádu**) na $\mathbb{R}^{m_i}$.

Programování na kuželu druhého řádu (SOCP)

$f, h_1, \ldots, h_\ell$ afinní a

$$g_i(\mathbf{x}) = \|\mathbf{A}_i\mathbf{x} + \mathbf{b}_i\|_2 - (\mathbf{c}_i^T\mathbf{x} + d_i)$$

Podmínka $g_i(\mathbf{x}) \leq 0$ se dá psát jako

$$\begin{bmatrix} \mathbf{A}_i \\ \mathbf{c}_i \end{bmatrix} \mathbf{x} + \begin{bmatrix} \mathbf{b}_i \\ d_i \end{bmatrix} \in K_2^{m_i}$$

kde $\mathbf{A}_i \in \mathbb{R}^{m_i \times n}$ a

$$K_2^{m_i} = \{\, (\mathbf{x}, y) \in \mathbb{R}^{m_i+1} \mid \|\mathbf{x}\|_2 \leq y \,\}$$

je epigraf eukleidovské normy $\|\cdot\|_2$ (**kužel druhého řádu**) na $\mathbb{R}^{m_i}$.

- Když $\mathbf{A}_i = \mathbf{0}$ pro všechna i, SOCP se redukuje na LP.
- Když $\mathbf{c}_i = \mathbf{0}$ pro všechna i, SOCP se redukuje na QCQP.

Příklad: Geometrický medián (Fermat-Weberův problém)

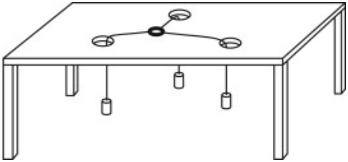
Pro zadané body $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ najděte minimum funkce

$$f(\mathbf{x}) = \sum_{i=1}^{m} \|\mathbf{x} - \mathbf{a}_i\|_2$$

SOCP formulace:

$$\begin{aligned} \min \quad & y_1 + \cdots + y_m \\ \text{za podmínek} \quad & \|\mathbf{x} - \mathbf{a}_i\|_2 \leq y_i, \quad i = 1, \ldots, m. \end{aligned}$$

Existuje na to speciální algoritmus (Weiszfeldův).



Varignon frame

Semidefinitní programování (SDP)

Věta

Množina všech positivně semidefinitních matic pevného rozměru $n \times n$ je konvexní kužel.

Důkaz: Pro každé $\mathbf{x} \in \mathbb{R}^n$ a $\alpha, \beta \geq 0$ platí $\mathbf{x}^T(\alpha\mathbf{A} + \beta\mathbf{B})\mathbf{x} = \alpha\mathbf{x}^T\mathbf{A}\mathbf{x} + \beta\mathbf{x}^T\mathbf{B}\mathbf{x} \geq 0$.

Semidefinitní programování (SDP)

Věta

Množina všech positivně semidefinitních matic pevného rozměru $n \times n$ je konvexní kužel.

Důkaz: Pro každé $\mathbf{x} \in \mathbb{R}^n$ a $\alpha, \beta \geq 0$ platí $\mathbf{x}^T(\alpha\mathbf{A} + \beta\mathbf{B})\mathbf{x} = \alpha\mathbf{x}^T\mathbf{A}\mathbf{x} + \beta\mathbf{x}^T\mathbf{B}\mathbf{x} \geq 0$.

Úloha SDP:

$$\begin{aligned} \text{min} \quad & \langle \mathbf{C}, \mathbf{X} \rangle \\ \text{za podmínek} \quad & \langle \mathbf{A}_i, \mathbf{X} \rangle = b_i, \quad i = 1, \ldots, m \\ & \mathbf{X} \in \mathbb{R}^{n \times n} \text{ je positivně semidefinitní} \end{aligned}$$

Tedy minimalizujeme lineární funkci p.s.d. matice za podmínek lineárních rovností.

Semidefinitní programování (SDP)

Věta

Množina všech positivně semidefinitních matic pevného rozměru $n \times n$ je konvexní kužel.

Důkaz: Pro každé $\mathbf{x} \in \mathbb{R}^n$ a $\alpha, \beta \geq 0$ platí $\mathbf{x}^T(\alpha\mathbf{A} + \beta\mathbf{B})\mathbf{x} = \alpha\mathbf{x}^T\mathbf{A}\mathbf{x} + \beta\mathbf{x}^T\mathbf{B}\mathbf{x} \geq 0$.

Úloha SDP:

$$\begin{aligned} \text{min} \quad & \langle \mathbf{C}, \mathbf{X} \rangle \\ \text{za podmínek} \quad & \langle \mathbf{A}_i, \mathbf{X} \rangle = b_i, \quad i = 1, \ldots, m \\ & \mathbf{X} \in \mathbb{R}^{n \times n} \text{ je positivně semidefinitní} \end{aligned}$$

Tedy minimalizujeme lineární funkci p.s.d. matice za podmínek lineárních rovností.

- Pro diagonální matice $\mathbf{A}_i, \mathbf{C}$ se SDP redukuje na LP.
- SOCP lze reprezentovat pomocí SDP (neuvádíme).

Semidefinitní programování (SDP)

Věta

Množina všech positivně semidefinitních matic pevného rozměru $n \times n$ je konvexní kužel.

Důkaz: Pro každé $\mathbf{x} \in \mathbb{R}^n$ a $\alpha, \beta \geq 0$ platí $\mathbf{x}^T(\alpha\mathbf{A} + \beta\mathbf{B})\mathbf{x} = \alpha\mathbf{x}^T\mathbf{A}\mathbf{x} + \beta\mathbf{x}^T\mathbf{B}\mathbf{x} \geq 0$.

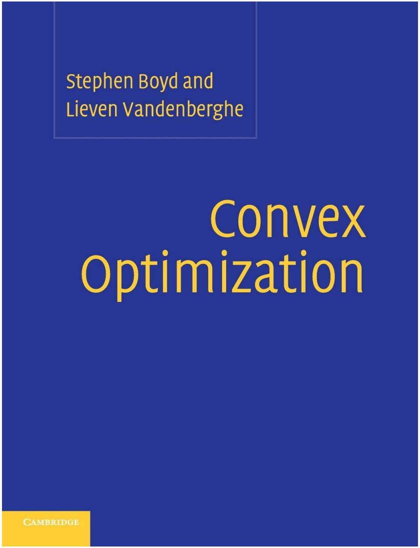
Úloha SDP:

$$\begin{aligned} \text{min} \quad & \langle \mathbf{C}, \mathbf{X} \rangle \\ \text{za podmínek} \quad & \langle \mathbf{A}_i, \mathbf{X} \rangle = b_i, \quad i = 1, \ldots, m \\ & \mathbf{X} \in \mathbb{R}^{n \times n} \text{ je positivně semidefinitní} \end{aligned}$$

Tedy minimalizujeme lineární funkci p.s.d. matice za podmínek lineárních rovností.

- Pro diagonální matice $\mathbf{A}_i, \mathbf{C}$ se SDP redukuje na LP.
- SOCP lze reprezentovat pomocí SDP (neuvádíme).
- Souhrn:

$$\text{LP} \subset \text{QP} \subset \text{QCQP} \subset \text{SOCP} \subset \text{SDP}$$



Zdarma ke stažení, spolu se slajdy atd.!

Příklad: Interface solveru SEDUMI na řešení LP, SOCP, SDP

```
% X = SEDUMI(A,b,c,K) minimizes c'*x subject to A*x = b and x is restricted to
%   a self-dual cone, defined by structure K. The fields of K can be
%   K.f, K.l, K.q, K.s (Free, Linear, Quadratic, Semi-definite).
%
%   (1) K.f is the number of FREE, i.e. UNRESTRICTED primal components.
%       These are ALWAYS the first components in x.
%
%   (2) K.l is the number of NONNEGATIVE components. E.g. if K.f=2, K.l=8
%       then x(3:10) >=0.
%
%   (3) K.q lists the dimensions of LORENTZ (quadratic, second-order cone)
%       constraints. E.g. if K.l=10 and K.q = [3 7] then
%           x(11) >= norm(x(12:13)),
%           x(14) >= norm(x(15:20)).
%       These components ALWAYS immediately follow the K.l nonnegative ones.
%
%   (4) K.s lists the dimensions of POSITIVE SEMI-DEFINITE (PSD) constraints
%       E.g. if K.l=10, K.q = [3 7] and K.s = [4 3], then
%           mat( x(21:36),4 ) is PSD,
%           mat( x(37:45),3 ) is PSD.
%       These components are ALWAYS the last entries in x.
```

Co s nekonvexními úlohami?

Některé nekonvexní úlohy jsou ‘snadné’ ...

Příklad: Minimalizace (ne nutně p.s.d.) kvadratické formy na sféře:

$$\begin{aligned} \min \quad & \mathbf{x}^T \mathbf{A} \mathbf{x} \\ \text{za podmínky} \quad & \mathbf{x}^T \mathbf{x} = 1 \end{aligned}$$

Některé nekonvexní úlohy jsou ‘snadné’ ...

Příklad: Minimalizace (ne nutně p.s.d.) kvadratické formy na sféře:

$$\begin{aligned} \min \quad & \mathbf{x}^T \mathbf{A} \mathbf{x} \\ \text{za podmínky} \quad & \mathbf{x}^T \mathbf{x} = 1 \end{aligned}$$

Příklad: QP s jediným omezením:

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{za podmínky} \quad & g(\mathbf{x}) \leq 0 \end{aligned}$$

kde $f, g \colon \mathbb{R}^n \to \mathbb{R}$ jsou dané (ne nutně konvexní) kvadratické funkce.
Tato úloha jde převést na SDP (neuvádíme).

... ale většina ne

Příklad: Minimalizace (ne nutně p.s.d.) kvadratické formy na hyperkrychli:

$$\begin{aligned} \min \quad & \mathbf{x}^T \mathbf{A} \mathbf{x} \\ \text{za podmínky} \quad & \mathbf{x} \in [-1, 1]^n \end{aligned}$$

Pro $\mathbf{A} = -\mathbf{I}$ má úloha 2^n lokálních minim (vrcholy hyperkrychle).

Pro obecné $\mathbf{A}$ je úloha NP-těžká (důkaz neuvádíme).

... ale většina ne

Příklad: Minimalizace (ne nutně p.s.d.) kvadratické formy na hyperkrychli:

$$\begin{aligned} \min \quad & \mathbf{x}^T \mathbf{A} \mathbf{x} \\ \text{za podmínky} \quad & \mathbf{x} \in [-1, 1]^n \end{aligned}$$

Pro $\mathbf{A} = -\mathbf{I}$ má úloha 2^n lokálních minim (vrcholy hyperkrychle).
Pro obecné $\mathbf{A}$ je úloha NP-těžká (důkaz neuvádíme).

Příklad: Celočíselné lineární programování

$$\begin{aligned} \min \quad & \mathbf{c}^T \mathbf{x} \\ \text{za podmínek} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b} \\ & \mathbf{x} \in \{0, 1\}^n \end{aligned}$$

Podmínky $x_i \in \{0, 1\}$ lze napsat jako nekonvexní kvadratické rovnosti $x_i(1 - x_i) = 0$.

Konvexní relaxace nekonvexní úlohy

Je-li $X \subseteq Y \subseteq \mathbb{R}^n$, pak

$$\underbrace{\min\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{x} \in X \,\}}_{\text{původní úloha}} \quad \geq \quad \underbrace{\min\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{x} \in Y \,\}}_{\text{relaxovaná úloha}}$$

Nahradíme původní (složitou, nekonvexní) množinu X její (jednoduchou, konvexní) nadmnožinou Y.

Viděli jsme dříve pro 0-1 LP, kde

$$X = \{\, \mathbf{x} \in \{0,1\}^n \mid \mathbf{Ax} \geq \mathbf{b} \,\}$$
$$Y = \{\, \mathbf{x} \in [0,1]^n \mid \mathbf{Ax} \geq \mathbf{b} \,\}$$

Užitečnost:

- Poskytne dolní mez na hodnotu globálního minima nerelaxované úlohy.
- Projekce opt. argumentu relaxované úlohy na množinu X může být 'dobrou' aproximací opt. argumentu původní úlohy. (Viděli jsme pro nejmenší vrcholové pokrytí, kde projekce byla zaokrouhlení.)

Co s nekonvexními úlohami?

Obecně dvě cesty:

- Nalezení globálního optima, ale obecně (worst-case) v čase exponenciálním ve velikosti úlohy (**globální optimalizace**).
  - hierarchie stále těsnějších konv. relaxací (Sherali-Adams hierarchy pro 0-1 LP)
  - metoda sečných nadrovin (cutting-plane method)
  - metoda větví a mezí (branch-and-bound)

- Nalezení přibližného optima (tzv. sub-optima) v krátkém čase.
  Heuristiky (každá použitelná jen na určitý typ úloh):
  - najde lokální (v nějakém smyslu) optimum (lokální optimalizace, alternující optimalizace / optimalizace po blocích souřadnic, majorize-minimize methods)
  - konv. relaxace + projekce na původní přípustnou množinu
  - přírodou inspirované heuristiky (evoluční/genetické algoritmy, simulované žíhání, kolonie mravenců, ...)
  - tabu search
  - continuation/homotopy methods
  - stochastické metody

Taxonomie optimalizačních algoritmů

- iterační (opakují stejnou iteraci) – neiterační
- deterministické – stochastické
- přesné – přibližné
- lokální – globální
- centralizované – distribuované (např. výpočty v počítačových sítích, rozvodných elektrických sítích, na mnoha jádrech procesoru, ...)
- sekvenční – paralelní (překryv s centralizované – distribuované)
- heuristické:
    - lokální hledání (hill climbing)
    - hladové (greedy) algoritmy
    - tabu search
    - motivované přírodou: genetické/evoluční, simulované žíhání, kolonie mravenců