Optimalizace

12. Konvexní funkce a konvexní optimalizace

Tom Kroupa, Tom Werner

FEL ČVUT

Konvexní funkce

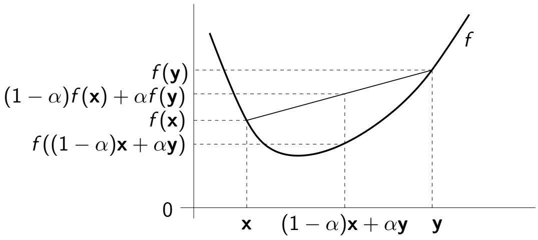
Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina a $f : \mathbb{R}^n \to \mathbb{R}$ funkce.

Definice

Funkce f je **konvexní** na X, jestliže pro všechna $\mathbf{x}, \mathbf{y} \in X$ a každé $\alpha \in [0, 1]$ platí

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

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



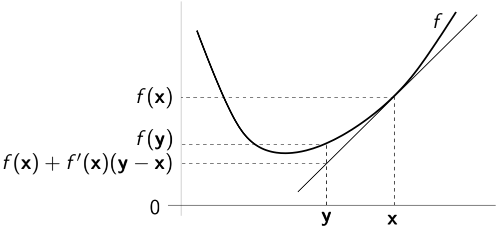
Konvexní diferencovatelné funkce

Podmínka prvního řádu

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 druhého řádu

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

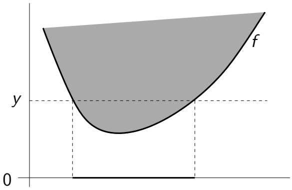
- $f(x) = e^{ax}$ pro $a \in \mathbb{R}$, $\quad f(x) = |x|^a$ pro $a \geq 1$
- $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ pro $\mathbf{A}$ pozitivně semidefinitní
- $f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} + b$
- $f(\mathbf{x}) = \max \{x_1, \ldots, x_n\}$
- $f(\mathbf{x}) = \|\mathbf{x}\|$,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

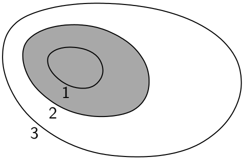
Vztah konvexní funkce a konvexní množiny

- **Epigraf** funkce f je množina $\{\,(\mathbf{x},y) \in \mathbb{R}^{n+1} \mid f(\mathbf{x}) \le y\,\}$.
- **Subkontura** výšky y je množina $\{\,\mathbf{x} \in \mathbb{R}^{n} \mid f(\mathbf{x}) \le y\,\}$.

Věta

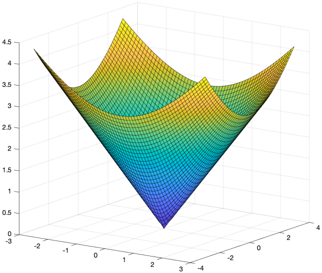
- f je konvexní, právě když její epigraf je konvexní množina.
- Každá subkontura konvexní funkce je konvexní množina.

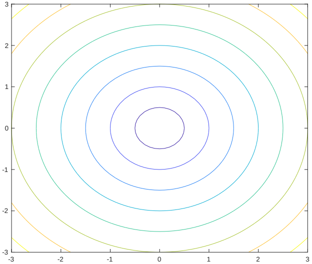




Příklad: Epigraf a subkontury pro kužel druhého řádu v $\mathbb{R}^2$

- Epigraf eukleidovské normy v $\mathbb{R}^2$ je $K_2^2 = \{ (\mathbf{x}, y) \in \mathbb{R}^3 \mid \|\mathbf{x}\|_2 \leq y \}$
- Subkontury jsou kruhy





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.

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).

Příklady:

- $f(\mathbf{x}) = \max_{i=1}^{k}(\mathbf{a}_i^T \mathbf{x} + b_i)$
- $f(\mathbf{x}) = \max_{\mathbf{y} \in C} \|\mathbf{x} - \mathbf{y}\|$ pro libovolnou kompaktní množinu $C \subseteq \mathbb{R}^n$
- $f(\mathbf{c}) = \max\{\, \mathbf{c}^T \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n,\ \mathbf{A}\mathbf{x} \geq \mathbf{b} \,\}$

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 $f : \mathbb{R}^n \to \mathbb{R}$ je konvexní funkce na množině X.

Věta

Nechť funkce $f : \mathbb{R}^n \to \mathbb{R}$ 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 lokálního minima!

Konvexní optimalizační úloha ve standardním tvaru

Tvrzení

Úloha

$$\begin{aligned} \min \quad & f(x_1,\ldots,x_n) \\ \text{za podmínek} \quad & g_i(x_1,\ldots,x_n) \leq 0, \quad i=1,\ldots,m \\ & h_i(x_1,\ldots,x_n) = 0, \quad i=1,\ldots,\ell \\ & x_1,\ldots,x_n \in \mathbb{R} \end{aligned}$$

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

Důkaz:

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

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.

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

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

Příklad: Řešení přeurčené soustavy s lineárními omezeními

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

Příklad: Řešení přeurčené soustavy s intervalovými omezeními

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

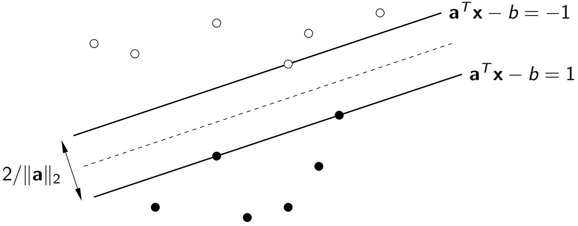
Příklad na QP: Support Vector Machine (SVM)

Pro zadaných m bodů $(\mathbf{x}_i, y_i) \in \mathbb{R}^n \times \{-1, 1\}$ hledáme **oddělující nadrovinu**.
Tedy hledáme $\mathbf{a} \in \mathbb{R}^n$ a $b \in \mathbb{R}$ tak, aby

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

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

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



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

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 kruh obsahující zadané body $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^2$.
Úloha

$$\min_{\mathbf{x}\in\mathbb{R}^2} \max_{i=1}^{m} \|\mathbf{x} - \mathbf{a}_i\|_2$$

je ekvivalentní QCQP úloze

$$\begin{aligned} \min \quad & y \\ \text{za podmínek} \quad & \|\mathbf{x} - \mathbf{a}_i\|_2^2 \le y, \quad i = 1, \ldots, m \end{aligned}$$

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 $g_i(\mathbf{x}) \leq 0$ znamená

$$(\mathbf{A}_i\mathbf{x} + \mathbf{b}_i, \mathbf{c}_i^T\mathbf{x} + d_i) \in K_2^n$$

kde

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

je epigraf eukleidovské normy $\|\cdot\|_2$ (**kužel druhého řádu**)

Maticově (pro $\ell = 0$):

$$\begin{aligned} \text{min} \quad & \mathbf{e}^T\mathbf{x} \\ \text{za podm.} \quad & \|\mathbf{A}_i\mathbf{x} + \mathbf{b}_i\|_2 \leq \mathbf{c}_i^T\mathbf{x} + d_i \end{aligned}$$

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:

$$\begin{aligned} \min \quad & z_1 + \cdots + z_m \\ \text{za podmínek} \quad & \|\mathbf{x} - \mathbf{a}_i\|_2 \leq z_i, \quad i = 1, \ldots, m. \end{aligned}$$

Semidefinitní programování (SDP)

Věta

Pro každé $n \in \mathbb{N}$ je množina všech positivně semidefinitních matic rozměru $n \times n$ 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:

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