Optimalizace

Konvexní funkce a konvexní optimalizace

Tomáš Kroupa

FEL ČVUT

Obsah

Konvexní funkce

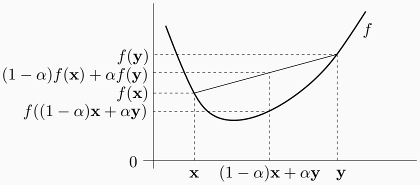
Nechť $X \subseteq \mathbb{R}^n$ je konvexní množina a $f\colon \mathbb{R}^n \to \mathbb{R}$ funkce.

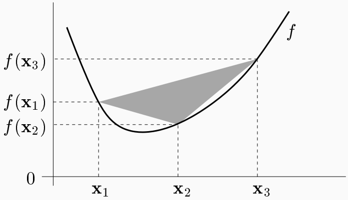
Definice

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

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

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



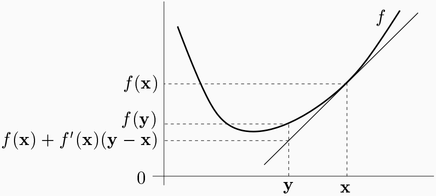


Konvexní diferencovatelné funkce

Podmínka prvního řádu

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

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



Podmínka druhého řádu

Nechť f 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

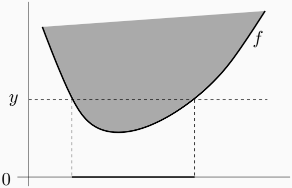
- $f(x) = e^{ax}$ pro $a \in \mathbb{R}$, $\quad f(x) = |x|^a$ pro $a \geq 1$
- $f(\mathbf{x}) = \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}$ pro $\mathbf{A}$ pozitivně semidefinitní
- $f(\mathbf{x}) = \mathbf{a}^\mathsf{T}\mathbf{x} + b$
- $f(\mathbf{x}) = \max\{x_1, \ldots, x_n\}$
- $f(\mathbf{x}) =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

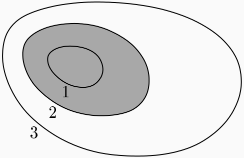
Vztah konvexní funkce a konvexní množiny

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

Věta

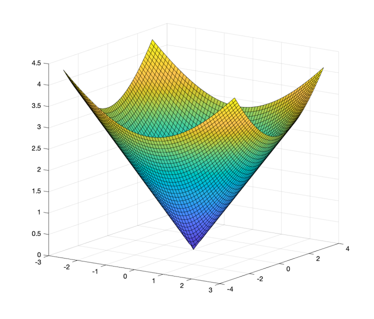
- f je konvexní, právě když její epigraf je konvexní množina.
- Každá subkontura konvexní funkce je konvexní množina.

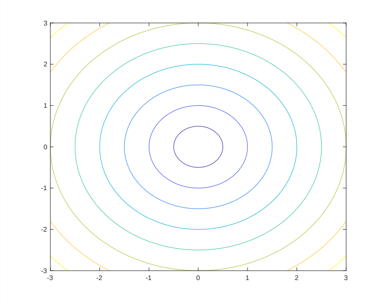




Epigraf a subkontury pro kužel druhého řádu v $\mathbb{R}^2$

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





Operace zachovávající konvexitu

Nezáporné lineární kombinace

Jsou-li $g_1, \dots, g_k\colon \mathbb{R}^n \to \mathbb{R}$ konvexní funkce a $\alpha_1, \dots, \alpha_k \geq 0$, pak je funkce $f = \alpha_1 g_1 + \cdots + \alpha_k g_k$ 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, existuje-li pro každé $\mathbf{x} \in \mathbb{R}^n$ maximum výše.

Příklady

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

Konvexní optimalizace

Definice

Nechť funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ je konvexní na konvexní množině $X \subseteq \mathbb{R}^n$. Úloha konvexní optimalizace je optimalizační úloha

$$\min \ f(\mathbf{x}) \qquad \text{za podmínky } \mathbf{x} \in X.$$

Konvexní úlohu vyřešíme nalezením lokálního minima!

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 X je globální.

Konvexní optimalizační úloha ve standardním tvaru

- $f\colon \mathbb{R}^n \to \mathbb{R}$ je konvexní
- $g_1, \dots, g_m\colon \mathbb{R}^n \to \mathbb{R}$ jsou konvexní
- $h_1, \dots, h_l\colon \mathbb{R}^n \to \mathbb{R}$ jsou afinní

Maticově: $\min \{ f(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0} \}$

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

Třídy konvexních optimalizačních úloh

- Lineární programování (LP)
- Kvadratické programování (QP)
- Kvadratické programování s kvadrat. omezeními (QCQP)
- Programování na kuželu druhého řádu (SOCP)

Poznámky

- V úlohách LP jsou všechny zúčastněné funkce afinní.
- Jen speciální třídy úloh QP a QCQP jsou konvexní!

Kvadratické programování

- $f\colon \mathbb{R}^n \to \mathbb{R}$ je kvadratická
- $g_1, \ldots, g_m\colon \mathbb{R}^n \to \mathbb{R}$ jsou afinní
- $h_1, \ldots, h_l\colon \mathbb{R}^n \to \mathbb{R}$ jsou afinní

Maticově

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

- $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{b} \in \mathbb{R}^n$
- $\mathbf{C} \in \mathbb{R}^{m\times n}$, $\mathbf{d} \in \mathbb{R}^m$
- $\mathbf{E} \in \mathbb{R}^{l\times n}$, $\mathbf{f} \in \mathbb{R}^l$

Je to konvexní úloha, právě když **A** je pozitivně semidefinitní.

Kvadratické programování – příklady konvexních úloh

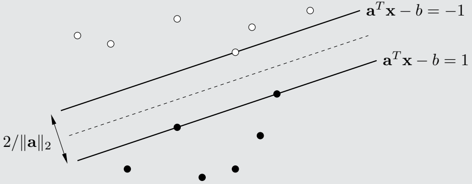
Řešení přeurčené soustavy $\mathbf{Ax} = \mathbf{b}$ s lineárními omezeními

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

Support vector machines

Pro zadaných m bodů $(\mathbf{x}_i, y_i) \in \mathbb{R}^n \times \{-1, 1\}$ hledáme oddělující nadrovinu popsanou pomocí $\mathbf{a} \in \mathbb{R}^n$ a $b \in \mathbb{R}$. Řešíme úlohu

$$\min\,\{\|\mathbf{a}\|_2^2 \mid y_i(\mathbf{a}^\mathsf{T}\mathbf{x}_i - b) \geq 1,\ i = 1, \ldots, m\}.$$



Kvadratické programování s kvadratickými omezeními

- $f\colon \mathbb{R}^n \to \mathbb{R}$ je kvadratická
- $g_1, \ldots, g_m\colon \mathbb{R}^n \to \mathbb{R}$ jsou kvadratické
- $h_1, \ldots, h_l\colon \mathbb{R}^n \to \mathbb{R}$ jsou afinní

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

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

Úloha $\min \{\max_{i=1}^{m} \|\mathbf{x} - \mathbf{a}_i\|_2 \mid \mathbf{x} \in \mathbb{R}^2\}$ je ekvivalentní QCQP úloze

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

Programování na kuželu druhého řádu

- $f\colon \mathbb{R}^n \to \mathbb{R}$ je afinní
- $g_i(\mathbf{x}) = \|\mathbf{A}_i\mathbf{x} + \mathbf{b}_i\|_2 - (\mathbf{c}_i^\mathsf{T}\mathbf{x} + d_i)$ pro $i = 1, \ldots, m$
- $h_1, \ldots, h_l\colon \mathbb{R}^n \to \mathbb{R}$ jsou afinní

Maticově (klademe $h_i(\mathbf{x}) = 0$)

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

- $\mathbf{e} \in \mathbb{R}^n$
- $\mathbf{A}_i \in \mathbb{R}^{n_i \times n}$, $\mathbf{b}_i \in \mathbb{R}^{n_i}$
- $\mathbf{c}_i \in \mathbb{R}^n$, $d_i \in \mathbb{R}$

Programování na kuželu druhého řádu – geometrický medián

Fermat-Weberův problém

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

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

Fermat-Weberův problém – SOCP formulace

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

Reference

- T. Werner. *Optimalizace* (kapitola 15–16). Elektronická skripta. FEL ČVUT, 2020.
- F. J. Aragón, M. A. Goberna, M. A. López, M. M. Rodríguez. *Nonlinear optimization*. Springer, 2019.