

Optimalizace

Konvexní optimalizační úlohy

Tomáš Kroupa Tomáš Werner

2023 LS

Fakulta elektrotechnická
ČVUT v Praze

Definice

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

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

- Různé třídy úloh liší se tvarem účelové funkce a omezeními
- Mnoho numerických metod i specializovaných řešičů
- Výhodné matematické vlastnosti

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

Věta

Pro každou konvexní úlohu platí:

1. Každé lokální minimum funkce f na X je globální.
2. Množina všech optimálních argumentů funkce f je konvexní.
3. Pokud je f ryze konvexní, existuje nejvýše jedno optimum.

Konvexní úloha ve standardním tvaru

- $f: \mathbb{R}^n \rightarrow \mathbb{R}$ je konvexní
- $g_1, \dots, g_m: \mathbb{R}^n \rightarrow \mathbb{R}$ jsou konvexní
- $h_1, \dots, h_\ell: \mathbb{R}^n \rightarrow \mathbb{R}$ jsou afinní

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

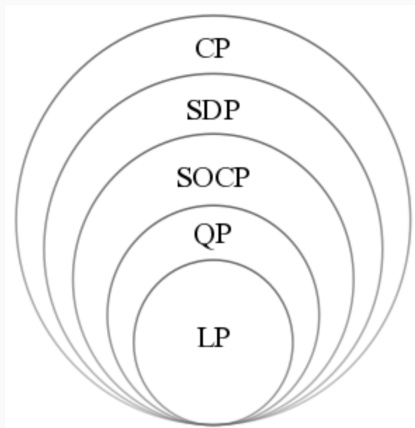
$$\min f(x_1, \dots, x_n)$$

$$\text{za podmíněk } g_i(x_1, \dots, x_n) \leq 0, \quad i = 1, \dots, m$$

$$h_i(x_1, \dots, x_n) = 0, \quad i = 1, \dots, \ell$$

$$x_1, \dots, x_n \in \mathbb{R}$$

Třídy konvexních úloh - schéma



Wikipedia

- Lineární programování
 f, g_i, h_i afinní
- Kvadratické programování
 f konvexní kvadratická a g_i, h_i afinní
- Kvadratické programování s kvadratickými omezeními
 f, g_i konvexní kvadratické a h_i afinní
- Programování na kuželu druhého řádu
- Semidefinitní programování
- Kónické programování

Kvadratické programování

- f je kvadratická
- g_i, h_i jsou afinní

Maticově

$$\begin{array}{ll} \min & \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} \\ \text{za podmíněk} & \mathbf{C} \mathbf{x} \leq \mathbf{d} \\ & \mathbf{E} \mathbf{x} = \mathbf{f} \end{array}$$

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

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

Řešení přeúčtené soustavy $\mathbf{Ax} = \mathbf{b}$

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

Řešíme převodem na lineární rovnice (normální rovnice).

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

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

Lze převést na řešení soustavy lineárních rovnic (Lagrange).

Řešení přeúčtené soustavy $\mathbf{Ax} = \mathbf{b}$ s intervalovými omezeními

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

Kvadratické programování – Support vector machines (SVM)

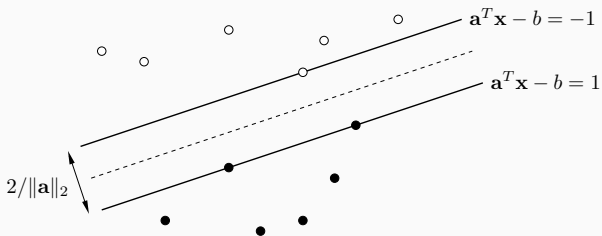
Pro m bodů $(\mathbf{x}_i, y_i) \in \mathbb{R}^n \times \{-1, 1\}$ hledáme **oddělující nadrovinu**.

Tedy hledáme $\mathbf{a} \in \mathbb{R}^n$ a $b \in \mathbb{R}$ tak, aby

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

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

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



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

Kvadratické programování s kvadratickými omezeními

- f, g_i jsou kvadratické
- h_i jsou afinní

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

Kam umístit záchrannou stanici?

Hledáme lokaci $\mathbf{x} \in \mathbb{R}^2$ pro záchrannou službu tak, aby *nejdelší vzdálenost* k místům $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^2$ byla minimální:

$$\text{Minimalizuj } \max_{i=1}^m \|\mathbf{x} - \mathbf{a}_i\|_2 \quad \text{z.p. } \mathbf{x} \in \mathbb{R}^2$$

Ekvivalentně – hledáme nejmenší kruh obsahující zadané body:

$$\text{Minimalizuj } y \quad \text{z.p. } \|\mathbf{x} - \mathbf{a}_i\|_2^2 \leq y, \quad i = 1, \dots, m.$$

Programování na kuželu druhého řádu (SOCP)

- f a h_1, \dots, h_ℓ jsou afinní
- $g_i(\mathbf{x}) = \|\mathbf{A}_i\mathbf{x} + \mathbf{b}_i\|_2 - (\mathbf{c}_i^\top \mathbf{x} + d_i)$ pro $i = 1, \dots, m$
- Podmínka $g_i(\mathbf{x}) \leq 0$ znamená $(\mathbf{A}_i\mathbf{x} + \mathbf{b}_i, \mathbf{c}_i^\top \mathbf{x} + d_i) \in K_2^n$, kde

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

je epigraf eukleidovské normy (tzv. **kužel druhého řádu**)

Maticově (pro $\ell = 0$)

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

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

Který bod minimalizuje součet vzdáleností od daných bodů?

Fermat-Weberův problém

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

$$f(\mathbf{x}) = \sum_{i=1}^m \|\mathbf{x} - \mathbf{a}_i\|_2, \quad \mathbf{x} \in \mathbb{R}^n.$$

To je ekvivalentní úloze s pomocnými proměnnými $z_1, \dots, z_m \in \mathbb{R}$

$$\min \quad z_1 + \dots + z_m$$

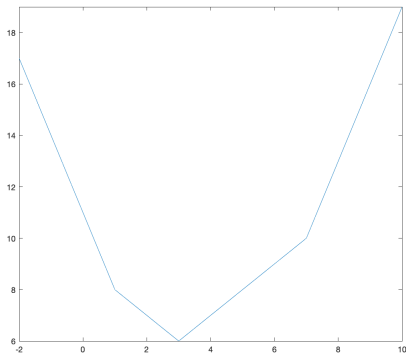
$$\text{za podmíněk} \quad \|\mathbf{x} - \mathbf{a}_i\|_2 \leq z_i, \quad i = 1, \dots, m.$$

Případ $n = 1$

Úloha

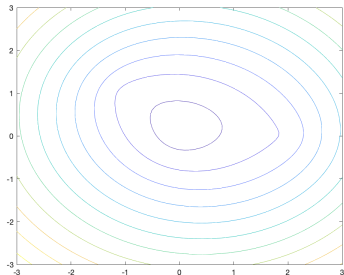
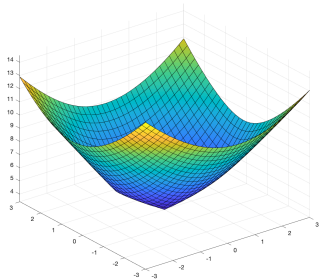
Pro zadaná $a_1, \dots, a_m \in \mathbb{R}$ minimalizujeme $f(x) = \sum_{i=1}^m |x - a_i|$, $x \in \mathbb{R}$. Řešením je **medián** těch čísel.

Funkce f pro $a_1 = 1, a_2 = 3, a_3 = 7$



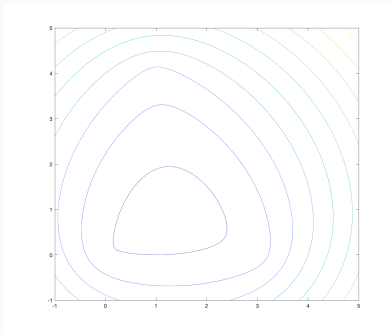
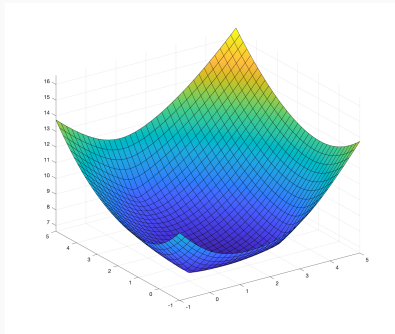
Případ $n = 2$ a $m = 3$ (1)

- Funkce f pro $\mathbf{a}_1 = (0, 0)$, $\mathbf{a}_2 = (2, 0)$, $\mathbf{a}_3 = (-1, 1)$
- Minima se nabývá v bodě \mathbf{a}_1 , kde funkce nemá derivaci
- V trojúhelníku $\triangle \mathbf{a}_1 \mathbf{a}_2 \mathbf{a}_3$ je úhel u vrcholu \mathbf{a}_1 větší než 120°



Případ $n = 2$ a $m = 3$ (2)

- Funkce f pro $\mathbf{a}_1 = (0, 0)$, $\mathbf{a}_2 = (3, 0)$, $\mathbf{a}_3 = (1, 4)$
- Minima se nabývá v bodě uvnitř trojúhelníka $\triangle \mathbf{a}_1 \mathbf{a}_2 \mathbf{a}_3$, který má všechny úhly ostré



Obecný případ pro $n = 2$ a $m = 3$

Předpoklad: body $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \in \mathbb{R}^2$ neleží na přímce

Trojúhelník $\triangle \mathbf{a}_1 \mathbf{a}_2 \mathbf{a}_3$ má úhel větší než 120°

Řešením je vrchol \mathbf{a}_i , u něhož je úhel větší než 120°

Trojúhelník $\triangle \mathbf{a}_1 \mathbf{a}_2 \mathbf{a}_3$ má úhly menší než 120°

Řešením je **Torricelliho bod** trojúhelníka $\triangle \mathbf{a}_1 \mathbf{a}_2 \mathbf{a}_3$, tj. bod ze kterého jsou “vidět” všechny strany pod úhlem 120°

Semidefinitní programování

- $\mathbf{C}, \mathbf{A}_1, \dots, \mathbf{A}_m \in \mathbb{R}^{n \times n}$, $b_i \in \mathbb{R}$
- Neznámá je matice $\mathbf{X} \in \mathbb{R}^{n \times n}$

SDP

$$\begin{array}{ll} \min & \langle \mathbf{C}, \mathbf{X} \rangle \\ \text{za podmínek} & \mathbf{X} \text{ je PSD} \\ & \langle \mathbf{A}_i, \mathbf{X} \rangle = b_i \\ & i = 1, \dots, m \end{array}$$

- Množina PSD matic je konvexní kužel
- LP, konvexní QCQP a SOCP jsou speciální případy

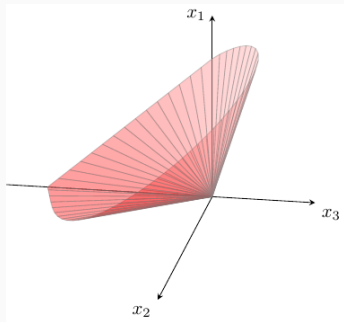
Programování s exponenciálním kuželem

- $\mathbf{c}^T \in \mathbb{R}^n$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$
- K je konvexní kužel, např. exponenciální kužel

$$K_e = \left\{ \mathbf{x} \in \mathbb{R}^3 \mid x_1 \geq x_2 e^{\frac{x_3}{x_2}}, x_2 > 0 \right\} \cup \{(x_1, 0, x_3) \mid x_1 \geq 0, x_3 \leq 0\}$$

Definice úlohy

$$\begin{aligned} \min \quad & \mathbf{c}^T \mathbf{x} \\ \text{za podmíněk} \quad & \mathbf{A} \mathbf{x} = \mathbf{b} \\ & x \in K \\ & x \in \mathbb{R}^n \end{aligned}$$



Různé nekonvexní úlohy (1)

Úloha vedoucí na spektrální rozklad

Minimalizuj $\mathbf{x}^T \mathbf{A} \mathbf{x}$ za podmínky $\|\mathbf{x}\|_2^2 = 1$, kde $\mathbf{x} \in \mathbb{R}^n$.

Lehká úloha s mnoha lokálními maximy

Maximalizuj $\|\mathbf{x}\|_2^2$ za podmínky $\|\mathbf{x}\|_\infty \leq 1$, kde $\mathbf{x} \in \mathbb{R}^n$.

Těžká nekonvexní úloha pro $\mathbf{a} \in \mathbb{N}^n$

Maximalizuj $\|\mathbf{x}\|_2^2$ za podmínek $\mathbf{a}^T \mathbf{x} = 0$ a $\|\mathbf{x}\|_\infty \leq 1$, kde $\mathbf{x} \in \mathbb{R}^n$.

Úloha celočíselného lineárního programování

Minimalizuj $\mathbf{a}^T \mathbf{x}$ za podmínek $\mathbf{A} \mathbf{x} \geq \mathbf{b}$, $\mathbf{x} \geq \mathbf{0}$, kde $\mathbf{x} \in \mathbb{Z}^n$.

Různé nekonvexní úlohy (2)

Shlukování

- Pro zadaných m bodů $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$ hledáme k shluků \mathcal{C}_j popsaných prototypem $\mathbf{x}_j \in \mathbb{R}^n$,

$$\mathcal{C}_j = \{\mathbf{a}_i \mid \|\mathbf{a}_i - \mathbf{x}_j\| \leq \|\mathbf{a}_i - \mathbf{x}_\ell\| \quad \forall i, \ell\},$$

tak, aby součet vzdáleností k prototypům byl minimální

- Minimalizujeme tak funkci

$$f(\mathbf{x}_1, \dots, \mathbf{x}_k) = \sum_{i=1}^m \min_{j=1}^k \|\mathbf{a}_i - \mathbf{x}_j\|$$

na množině vektorů $(\mathbf{x}_1, \dots, \mathbf{x}_k) \in \mathbb{R}^{kn}$