# Optimalizace

## Konvexní optimalizační úlohy

Tomáš Kroupa Tomáš Werner

2022 LS

Fakulta elektrotechnická
ČVUT v Praze

# 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

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

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

# Vlastnosti konvexních úloh

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

**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

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

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

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

# Třídy konvexních úloh

- 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

# Kvadratické programování

- $f$ je kvadratická
- $g_i, h_i$ jsou afinní

**Maticově**

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

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

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

# Kvadratické programování – příklady konvexních úloh

## Řešení přeurčené soustavy $\mathbf{A}\mathbf{x} = \mathbf{b}$

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

Řešíme převodem na lineární rovnice (normální rovnice).

## Řešení přeurčené soustavy $\mathbf{A}\mathbf{x} = \mathbf{b}$ s lineárními omezeními

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

Lze převést na řešení soustavy lineárních rovnic (Lagrange).

## Řešení přeurčené soustavy $\mathbf{A}\mathbf{x} = \mathbf{b}$ s intervalovými omezeními

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

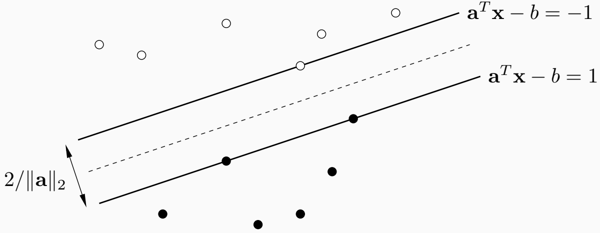
# Kvadratické programování – Support vector machines (SVM)

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

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

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

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




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

# 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 $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^2$ byla minimální:

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

Ekvivalentně – hledáme nejmenší kruh obsahující zadané body:

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

# Programování na kuželu druhého řádu

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

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

je epigraf eukleidovské normy (tzv. kužel druhého řádu)

**Maticově (pro $\ell = 0$)**

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

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

# Programování na kuželu druhého řádu – geometrický medián

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

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

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

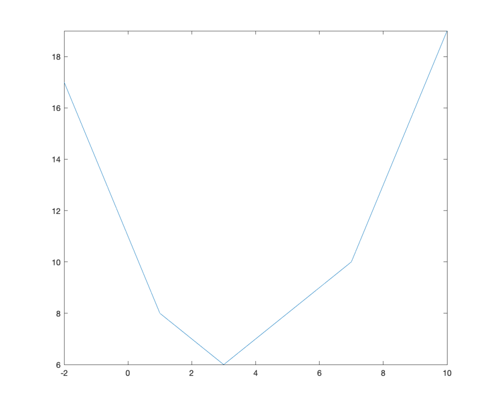
$$\begin{aligned} \text{min} \quad & z_1 + \cdots + z_m \\ \text{za podmínek} \quad & \|\mathbf{x} - \mathbf{a}_i\|_2 \leq z_i, \quad i = 1, \dots, m. \end{aligned}$$

# Případ $n = 1$

## Úloha

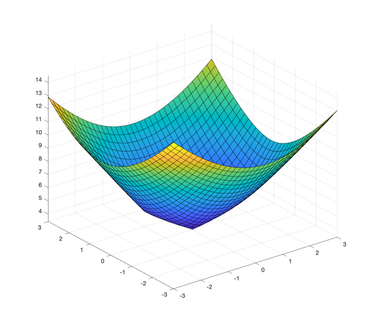
Pro zadaná $a_1, \ldots, a_m \in \mathbb{R}$ minimalizujeme $f(x) = \sum_{i=1}^{m} |x - a_i|$, $x \in \mathbb{R}$. Řešením je medián těch čísel.

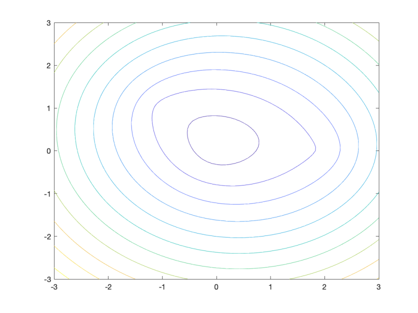
Funkce $f$ pro $a_1 = 1$, $a_2 = 3$, $a_3 = 7$



# Případ $n = 2$ a $m = 3$ (1)

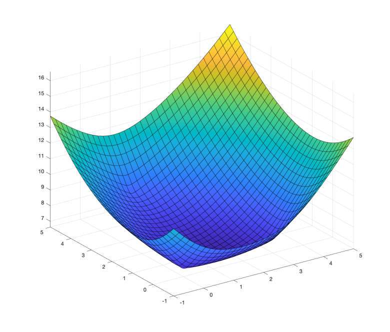
- Funkce $f$ pro $\mathbf{a}_1 = (0, 0)$, $\mathbf{a}_2 = (2, 0)$, $\mathbf{a}_3 = (-1, 1)$
- Minima se nabývá v bodě $\mathbf{a}_1$, kde funkce nemá derivaci
- V trojúhelníku $\triangle \mathbf{a}_1\mathbf{a}_2\mathbf{a}_3$ je úhel u vrcholu $\mathbf{a}_1$ větší než $120^\circ$

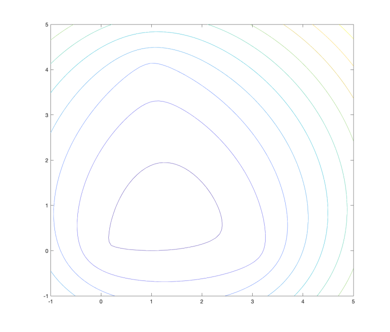


# Případ $n = 2$ a $m = 3$ (2)

- 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 $120°$

# Různé nekonvexní úlohy (1)

**Úloha vedoucí na spektrální rozklad**

Minimalizuj $\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}$ 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 $\mathbf{a}^\mathsf{T}\mathbf{x} = 0$ a $\|\mathbf{x}\|_\infty \leq 1$, kde $\mathbf{x} \in \mathbb{R}^n$.

**Úloha celočíselného lineárního programování**

Minimalizuj $\mathbf{a}^\mathsf{T}\mathbf{x}$ 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 $\mathbf{x}_j \in \mathbb{R}^n$,

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

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