

Optimalizace a teorie her

Konvexní množiny

Martin Bohata

Katedra matematiky
FEL ČVUT v Praze
bohata@math.feld.cvut.cz

Definice konvexní množiny

Definice

Nechť $x, y \in \mathbb{R}^n$. Množina

$$[x, y] := \{\lambda x + (1 - \lambda)y \mid 0 \leq \lambda \leq 1\}$$

se nazývá (**uzavřená**) **úsečka** s krajními body x, y .

Definice

Množina $C \subseteq \mathbb{R}^n$ se nazve **konvexní**, jestliže pro každé $x, y \in C$ je $[x, y] \subseteq C$.

Příklad

- 1 Každý interval v $\mathbb{R}$ je konvexní množina.
- 2 Každý lineární podprostor prostoru $\mathbb{R}^n$ je konvexní množina.
- 3 Každá úsečka a prázdná množina v $\mathbb{R}^n$ jsou konvexní množiny.
- 4 Nechť $\alpha \in \mathbb{R}$ a $y \in \mathbb{R}^n \setminus \{0\}$. **Nadrovina**

$$H(y; \alpha) := \{x \in \mathbb{R}^n \mid \langle x, y \rangle = \alpha\}.$$

je konvexní množina. Vektor y se nazývá **normálový vektor** nadroviny $H(y; \alpha)$.

Základní příklady

Příklad

- ⑥ Necht' $y \in \mathbb{R}^n \setminus \{0\}$ a $\alpha \in \mathbb{R}$. **Uzavřený poloprostor**

$$\{x \in \mathbb{R}^n \mid \langle x, y \rangle \leq \alpha\}$$

je konvexní množina.

- ⑦ **Uzavřená koule**

$$B(a; r) = \{x \in \mathbb{R}^n \mid \|x - a\| \leq r\}$$

o středu $a \in \mathbb{R}^n$ a poloměru $r > 0$ je konvexní množina.

- ⑧ **Okolí**

$$U(a; r) = \{x \in \mathbb{R}^n \mid \|x - a\| < r\}$$

bodů $a \in \mathbb{R}^n$ o poloměru $r > 0$ je konvexní množina.

Jsou množiny v ⑦ a ⑧ konvexní i v případě, že euklidovskou normu nahradíme jinou normou (například součtovou nebo maximovou)?

Množinové operace a konvexita

Tvrzení

Průnik libovolného počtu konvexních množin je konvexní množina.

Důkaz: Viz přednáška. ■

Značení:

- $\mathbb{M}_{m,n}(\mathbb{R})$... množina všech $m \times n$ reálných matic.
- $\mathbb{M}_n(\mathbb{R})$... množina všech $n \times n$ reálných matic.

Příklad

- ① Necht' $A \in \mathbb{M}_{m,n}(\mathbb{R})$ a $b \in \mathbb{R}^m$. **Afinní podprostor** (**lineární varieta**)

$$\{x \in \mathbb{R}^n \mid Ax = b\}.$$

je konvexní množina.

- ② **Nezáporný ortant**

$$\mathbb{R}_+^n := \left\{ (x_1, \dots, x_n)^T \in \mathbb{R}^n \mid x_1 \geq 0, \dots, x_n \geq 0 \right\}$$

je konvexní množina.

Příklad (Pokračování)

- 3 Necht' $A \in \mathbb{M}_{m,n}(\mathbb{R})$ a $b \in \mathbb{R}^m$. **Konvexní polyedrická množina**

$$\{x \in \mathbb{R}^n \mid Ax \leq b\},$$

kde $Ax \leq b$ znamená $b - Ax \in \mathbb{R}_+^m$, je konvexní množina.

Omezená konvexní polyedrická množina se nazývá **konvexní mnohostěn**. Například

$$\left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$$

je konvexní mnohostěn.

- Sjednocení a rozdíly množin konvexitu nezachovávají (viz např. $[0, 1] \setminus (0, 1) = \{0\} \cup \{1\}$).

Množinové operace a konvexita

Definice

Zobrazení $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ se nazývá **afinní**, existují-li $A \in \mathbb{M}_{m,n}(\mathbb{R})$ a $b \in \mathbb{R}^m$ tak, že $f(x) = Ax + b$.

Tvrzení

Nechť $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Pak f je afinní právě tehdy, když pro každé $x, y \in \mathbb{R}^n$ a každé $\lambda \in \mathbb{R}$ platí

$$f(\lambda x + (1 - \lambda)y) = \lambda f(x) + (1 - \lambda)f(y).$$

Důkaz: Viz přednáška. ■

- Z uvedeného tvrzení speciálně plyne, že afinní zobrazení f zachovává úsečky, tj. pro všechna $x, y \in \mathbb{R}^n$ je

$$f([x, y]) = [f(x), f(y)].$$

Množinové operace a konvexita

Tvrzení

Je-li $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ afinní a $C \subseteq \mathbb{R}^n$ konvexní, pak $f(C)$ je konvexní.

Důkaz: Viz přednáška. ■

Tvrzení

Nechť $C_1 \subseteq \mathbb{R}^n$ a $C_2 \subseteq \mathbb{R}^m$. Pak C_1 a C_2 jsou konvexní množiny právě tehdy, když $C_1 \times C_2$ je konvexní množina.

Důkaz: Viz přednáška. ■

Příklad

- 1 $[0, 1]^n$ je konvexní množina.
- 2 $B(0; 1) \times [0, 1]$ je konvexní množina.
- 3 **Kladný ortant** $\mathbb{R}_{++}^n = (0, \infty)^n$.

Konvexní obal

Definice

Konvexní obal množiny $M \subseteq \mathbb{R}^n$ je průnik všech konvexních množin v $\mathbb{R}^n$ obsahujících M . Konvexní obal množiny M se značí symbolem $\text{conv}(M)$.

- Konvexní obal libovolné množiny je konvexní množinou.
- Množina C je konvexní právě tehdy, když $C = \text{conv}(C)$.

Příklad

- 1 $\text{conv}(S(0; 1)) = B(0; 1)$.
- 2 Necht' $x = (0, 0)^T$, $y = (1, 0)^T$ a $z = (0, 1)^T$. Pak $\text{conv}(\{x, y, z\})$ je trojúhelník s vrcholy x, y, z .

3

$$\text{conv}\left(\text{⌒}\right) = \text{⌒}.$$

Konvexní kombinace

Definice

Řekneme, že $x \in \mathbb{R}^n$ je **konvexní kombinací** bodů $x_1, \dots, x_k \in \mathbb{R}^n$, jestliže existují nezáporná reálná čísla $\alpha_1, \dots, \alpha_k$ tak, že $x = \sum_{i=1}^k \alpha_i x_i$ a $\sum_{i=1}^k \alpha_i = 1$.

Věta

Je-li $M \subseteq \mathbb{R}^n$ neprázdná, pak $\text{conv}(M)$ je množina všech konvexních kombinací bodů z M .

Důkaz: Vynecháváme. ■

Věta (Carathéodoryho věta)

Je-li $M \subseteq \mathbb{R}^n$ neprázdná, pak každý prvek z $\text{conv}(M)$ lze vyjádřit jako konvexní kombinaci nejvýše $n + 1$ prvků z M .

Důkaz: Vynecháváme. ■

Kužely

Definice

Množina $K \subseteq \mathbb{R}^n$ se nazve **kužel**, jestliže pro každé $x \in K$ a každé $\alpha > 0$ je $\alpha x \in K$.

Je-li kužel navíc konvexní množinou, nazývá se **konvexní kužel**.

Příklad

- 1 Množiny $\emptyset$, $\{0\}$, $\mathbb{R}_+^n$, $\mathbb{R}_{++}^n$ a $\mathbb{R}^n$ jsou konvexní kužely.
- 2 Množina $\{(x, 0)^T \in \mathbb{R}^2 \mid x \geq 0\} \cup \{(0, x)^T \in \mathbb{R}^2 \mid x \geq 0\}$ je kužel, ale není konvexní.

Tvrzení

Průnik libovolného počtu kuželů je kužel.

Důkaz: Domácí úkol. ■

- Průnik libovolného počtu konvexních kuželů je konvexní kužel.

Vzdálenost bodu od množiny

Jaký bod z neprázdné množiny $M \subseteq \mathbb{R}^n$ je nejbližší bodu $x \in \mathbb{R}^n$?

Definice

Nechť $M \subseteq \mathbb{R}^n$ je neprázdna a $x \in \mathbb{R}^n$.

- Reálné číslo

$$\text{dist}(x; M) = \inf_{y \in M} \|x - y\|$$

se nazve **vzdálenost bodu x od množiny M** .

- Bod $y \in M$ se nazývá **projekce bodu x na M** , jestliže

$$y \in \operatorname{argmin}_{z \in M} \|x - z\|$$

(tj. $\|x - y\| \leq \|x - z\|$ pro všechna $z \in M$).

- Existuje-li právě jedna projekce x na M , pak ji budeme značit symbolem $P_M(x)$.

Vzdálenost bodu od množiny

Příklad

- ❶ $\operatorname{argmin}_{y \in (1,2]} |0 - y| = \emptyset.$
- ❷ $\operatorname{argmin}_{y \in (1,2]} |3 - y| = \{2\}.$
- ❸ Je-li $K = \{x \in \mathbb{R}^2 \mid \|x\| = 1\}$, pak

$$\operatorname{argmin}_{y \in K} \|0 - y\| = K.$$

Připomeňme důležitou Weierstrassova větu o existenci bodů minima a maxima.

Věta (Weierstrassova věta)

Je-li neprázdná množina $M \subseteq \mathbb{R}^n$ kompaktní (tj. omezená a uzavřená) a $f \in C(M)$, pak existuje bod minima a bod maxima funkce f na M .

Věta (Věta o nejlepší aproximaci)

Je-li $C \subseteq \mathbb{R}^n$ neprázdná uzavřená konvexní množina, pak pro každé $x \in \mathbb{R}^n$ existuje právě jeden bod $\hat{y} \in C$ tak, že

$$\text{dist}(x; C) = \|x - \hat{y}\|.$$

Důkaz: Viz přednáška. ■

Definice

Nechť $C \subseteq \mathbb{R}^n$ je neprázdná uzavřená konvexní množina. Zobrazení $P_C : x \mapsto P_C(x)$, $x \in \mathbb{R}^n$, se nazývá **projekční operátor** na C .

- $P_C(x) = x$ pro každé $x \in C$.
- $P_C \circ P_C = P_C$.
- Lze ukázat, že projekční operátor je spojitě zobrazení.

Vzdálenost bodu od množiny

Příklad

$$P_{[0,1]}(x) = \begin{cases} x, & \text{pro } x \in [0, 1], \\ 0, & \text{pro } x < 0, \\ 1, & \text{pro } x > 1. \end{cases}$$

Připomeňme Cauchyovu-Schwarzovu nerovnost dobře známou z kurzu lineární algebry.

Tvrzení (Cauchyova-Schwarzova nerovnost)

Pro každé $x, y \in \mathbb{R}^n$ platí

$$|\langle x, y \rangle| \leq \|x\| \|y\|.$$

Vzdálenost bodu od množiny

Věta (Projekce bodu a variační nerovnost)

Nechť $C \subseteq \mathbb{R}^n$ je neprázdná uzavřená konvexní množina, $x \in \mathbb{R}^n$ a $y \in C$. Pak následující tvrzení jsou ekvivalentní:

- 1 $y = P_C(x)$.
- 2 Pro každé $z \in C$ je $\langle x - y, z - y \rangle \leq 0$.

Důkaz: Viz přednáška

Příklad

Nechť $C = B(0; 1)$. Pak

$$P_C(x) = \begin{cases} x, & \text{jestliže } \|x\| \leq 1, \\ \frac{x}{\|x\|}, & \text{jestliže } \|x\| > 1. \end{cases}$$

Vzdálenost bodu od množiny

Tvrzení

Nechť $L \subseteq \mathbb{R}^n$ je lineární podprostor. Potom platí:

- 1 $P_L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ je lineární zobrazení.
- 2 Pro každé $x \in \mathbb{R}^n$ je $P_{L^\perp}(x) = x - P_L(x)$.
- 3 Pro každé $x \in \mathbb{R}^n$ existují jednoznačně určené body $y \in L$ a $z \in L^\perp$ tak, že $x = y + z$. Navíc $y = P_L(x)$ a $z = P_{L^\perp}(x)$.

Důkaz: Viz přednáška. ■

Příklad

Nechť $y \in \mathbb{R}^n$ a $L = \text{span}(\{y\})$. Pak

$$P_L(x) = \frac{\langle x, y \rangle}{\langle y, y \rangle} y.$$

Vzdálenost bodu od množiny

Příklad

Ať $C_1 \subseteq \mathbb{R}^n$ a $C_2 \subseteq \mathbb{R}^m$ jsou neprázdné, konvexní a uzavřené množiny. Potom

$$P_{C_1 \times C_2}(x) = \begin{pmatrix} P_{C_1}(a) \\ P_{C_2}(b) \end{pmatrix},$$

kde $x = \begin{pmatrix} a \\ b \end{pmatrix}$, $a \in \mathbb{R}^n$, $b \in \mathbb{R}^m$.

Příklad

Ať $C = \mathbb{R}_+^n$. Potom

$$P_C(x) = \begin{pmatrix} \max\{0, x_1\} \\ \vdots \\ \max\{0, x_n\} \end{pmatrix}$$

Metoda nejmenších čtverců

- Necht' $A \in \mathbb{M}_{m,n}(\mathbb{R})$ a $b \in \mathbb{R}^m$. Jaké prvky obsahuje množina

$$\operatorname{argmin}_{x \in \mathbb{R}^n} \|Ax - b\| = \operatorname{argmin}_{x \in \mathbb{R}^n} \|Ax - b\|^2 ?$$

- Nejen pro důkaz následujícího tvrzení bude užitečné následující jednoduché pozorování: je-li $A \in \mathbb{M}_{m,n}(\mathbb{R})$, $x \in \mathbb{R}^n$ a $y \in \mathbb{R}^m$, pak $\langle Ax, y \rangle = \langle x, A^T y \rangle$.

Tvrzení

Necht' $A \in \mathbb{M}_{m,n}(\mathbb{R})$, $b \in \mathbb{R}^m$ a $\hat{x} \in \mathbb{R}^n$. Potom

$$\hat{x} \in \operatorname{argmin}_{x \in \mathbb{R}^n} \|Ax - b\|^2 \text{ právě tehdy, když } A^T A \hat{x} = A^T b.$$

Důkaz: Viz přednáška. ■

- $A^T A x = A^T b \dots$ **soustava normálních rovnic**.

Metoda nejmenších čtverců

Ať $b \in \mathbb{R}^n$ a $L = \{Ax \mid x \in \mathbb{R}^n\}$, kde $A \in \mathbb{M}_{m,n}(\mathbb{R})$.

- A má lineárně nezávislé sloupce právě tehdy, když $A^T A$ je invertibilní.
- Má-li A lineárně nezávislé sloupce, pak $P_L(b) = A(A^T A)^{-1} A^T b$.
- Má-li matice A sloupce $a_1, \dots, a_n$, které jsou nenulové a vzájemně ortogonální, pak

$$P_L(b) = \sum_{i=1}^n \frac{\langle b, a_i \rangle}{\|a_i\|^2} a_i.$$

Příklad

Ať $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$ a $b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Pak jediné řešení soustavy $Ax = b$ ve smyslu nejmenších čtverců je

$$\frac{2}{3} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Metoda nejmenších čtverců

Příklad

Ať $A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \\ 1 & 2 & -3 \end{pmatrix}$ a $b = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$. Pak množina všech řešení soustavy $Ax = b$ ve smyslu nejmenších čtverců je

$$\left\{ \begin{pmatrix} \frac{2}{7} - t \\ -\frac{3}{7} + 2t \\ t \end{pmatrix} \mid t \in \mathbb{R} \right\}$$

Příklad

V rovině jsou dány body $(0, -\frac{1}{2})^T$, $(1, \frac{1}{3})^T$ a $(2, \frac{2}{3})^T$. Pomocí metody nejmenších čtverců proložíme těmito body přímkou o rovnici $y = kx + q$, kde $k, q \in \mathbb{R}$. Hledané hodnoty koeficientů k a q jsou $k = \frac{7}{12}$ a $q = -\frac{5}{12}$.

Metoda nejmenších čtverců

Příklad

Metodou nejmenších čtverců proložme body $(-2, -1)^T$, $(-1, -2)^T$, $(1, 3)^T$ a $(2, 2)^T$ graf funkce $f(x) = \alpha \frac{x}{1+x^2} + \beta \sin\left(\frac{\pi x}{2}\right) + \gamma$, kde $\alpha, \beta, \gamma \in \mathbb{R}$. Hledané koeficienty jsou $\alpha = \frac{15}{4}$, $\beta = \frac{5}{8}$ a $\gamma = \frac{1}{2}$.

Příklad

Ať $w_1, \dots, w_m$ jsou kladná reálná čísla, $W = \text{diag}(w_1, \dots, w_m)$ a $A \in \mathbb{M}_{m,n}(\mathbb{R})$ má řádky $a_1, \dots, a_m$. Je dána úloha (váhovaných nejmenších čtverců) minimalizujte funkci $f(x) = \sum_{i=1}^m w_i \left(\langle a_i^T, x \rangle - b_i \right)^2$ na $\mathbb{R}^n$.

- 1 Úlohu lze formulovat jako obyčejnou úlohu nejmenších čtverců ve tvaru: minimalizujte $g(x) = \left\| W^{\frac{1}{2}} Ax - W^{\frac{1}{2}} b \right\|^2$ na $\mathbb{R}^n$.
- 2 Má-li A lineárně nezávislé sloupce, potom jediné řešení uvedené optimalizační úlohy je $\hat{x} = (A^T W A)^{-1} A^T W b$.

Metoda nejmenších čtverců

Příklad

- Závislost výstupního signálu $(y_n)_{n=0}^{\infty}$ systému na vstupním signálu $(x_n)_{n=0}^{\infty}$ je dána konvolucí posloupnosti $(x_n)_{n=0}^{\infty}$ s posloupností $(h_n)_{n=0}^{\infty}$ (tj. $y_n = \sum_{i=0}^n h_i x_{n-i}$).
- Ať $h_n = 0$ pro všechna $n \geq 4$.
- Měřením byla zjištěna hodnota koeficientů $y_0, \dots, y_{20}$ výstupního signálu, když na vstupu byl signál s počátečními koeficienty $x_0, \dots, x_{20}$.

Úloha nejmenších čtverců pro nalezení koeficientů h_0, h_1, h_2, h_3 je:

Minimalizujte $f(x) = \|Ax - b\|^2$ na $\mathbb{R}^4$, kde A je matice s řádky $a_1 = (x_0, 0, 0, 0)$, $a_2 = (x_1, x_0, 0, 0)$, $\dots$, $a_{20} = (x_{20}, x_{19}, x_{18}, x_{17})$ a $b = (y_0, \dots, y_{20})^T$.

Věta o oddělování nadrovinou

Věta (O oddělitelnosti bodu a konvexní množiny)

Je-li $C \subseteq \mathbb{R}^n$ neprázdná uzavřená konvexní množina a $x \in \mathbb{R}^n \setminus C$, pak existuje $v \in \mathbb{R}^n \setminus \{0\}$ a $\alpha \in \mathbb{R}$ tak, že pro každé $y \in C$ platí

$$\langle y, v \rangle \leq \alpha < \langle x, v \rangle.$$

Důkaz: Viz přednáška. ■

Důsledek

Každá uzavřená konvexní množina v $\mathbb{R}^n$ je průnikem všech poloprostorů, které ji obsahují.

Důkaz: Viz přednáška. ■

Věta o oddělování nadrovinou

- Existuje nezáporné řešení soustavy lineárních rovnic? Obecně ne, stačí uvážit rovnici $x_1 + x_2 = -1$. Kdy takové řešení existuje?

Příklad

Nechť $A = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$ a $b \in \mathbb{R}^2$. Označme

$$C = \left\{ Ax \mid x \in \mathbb{R}_+^2 \right\} = \left\{ \alpha \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ -1 \end{pmatrix} \mid \alpha, \beta \geq 0 \right\}$$

$$\begin{aligned} K &= \left\{ y \in \mathbb{R}^2 \mid A^T y \leq 0 \right\} \\ &= \left\{ y \in \mathbb{R}^2 \mid \left\langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, y \right\rangle \leq 0, \left\langle \begin{pmatrix} 1 \\ -1 \end{pmatrix}, y \right\rangle \leq 0 \right\}. \end{aligned}$$

Nastává právě jedna z možností:

- $b \in C$.
- Existuje nenulový vektor $y \in K$ svírající s b úhel $\varphi \in [0, \frac{\pi}{2})$.

Věta o oddělování nadrovinou

Lemma

Jestliže $A \in \mathbb{M}_{m,n}(\mathbb{R})$, pak $\{Ax \mid x \in \mathbb{R}_+^n\}$ je neprázdná uzavřená konvexní množina.

Důkaz: Viz přednáška (důkaz uzavřenosti vynecháváme). ■

Věta (Farkasovo lemma)

Je-li $A \in \mathbb{M}_{m,n}(\mathbb{R})$ a $b \in \mathbb{R}^m$, pak platí právě jedno z následujících tvrzení:

- ❶ *Existuje $x \in \mathbb{R}^n$ tak, že $Ax = b$ a $x \geq 0$.*
- ❷ *Existuje $y \in \mathbb{R}^m$ tak, že $A^T y \leq 0$ a $\langle y, b \rangle > 0$.*

Důkaz: Viz přednáška. ■

Krajní body konvexní množiny

Úsečku, trojúhelník a čtverec lze rekonstruovat z jejich vrcholů. Jaké další konvexní množiny lze rekonstruovat ze znalosti „vrcholů“?

Definice

Nechť $C \subseteq \mathbb{R}^n$ je konvexní množina. Bod $x \in C$ nazveme **krajní bod** množiny C , jestliže neexistují dva různé body $y, z \in C$ tak, že $x = \frac{y+z}{2}$. Množinu všech krajních bodů množiny C budeme značit symbolem $\text{ext}(C)$.

Příklad

- 1 $\text{ext}([0, 1]) = \{0, 1\}$.
- 2 Je-li $L \subseteq \mathbb{R}^n$ netriviální lineární podprostor, pak $\text{ext}(L) = \emptyset$.
- 3 $\text{ext}(B(0; 1)) = S(0; 1)$.

Věta (Kreinova-Milmanova)

Jestliže $C \subseteq \mathbb{R}^n$ je kompaktní (tj. omezená a uzavřená) konvexní množina, pak $C = \text{conv}(\text{ext}(C))$.

Důkaz: Vynecháváme. ■

Kompaktnost je důležitá:

- Interval $(0, 1)$ není uzavřený a $\text{ext}((0, 1)) = \emptyset$.
- Množina $\mathbb{R}_+^2$ není omezená a $\text{ext}(\mathbb{R}_+^2) = \{0\}$.

Příklad

$$B(0; 1) = \text{conv}(S(0; 1)).$$