

Optimalizace

10. Konvexní množiny a mnohostěny

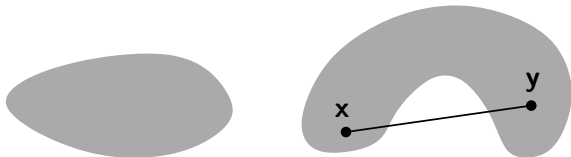
Tom Werner, Tom Kroupa

FEL ČVUT

Konvexní množiny

Množina $X \subseteq \mathbb{R}^n$ je **konvexní**, jestliže

$$\mathbf{x} \in X, \mathbf{y} \in X, 0 \leq \alpha \leq 1 \implies (1 - \alpha)\mathbf{x} + \alpha\mathbf{y} \in X.$$



- **Konvexní kombinace** bodů $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ je bod $\alpha_1\mathbf{x}_1 + \dots + \alpha_k\mathbf{x}_k$, kde

$$\alpha_1 + \dots + \alpha_k = 1, \quad \alpha_1, \dots, \alpha_k \geq 0.$$

Tvrzení: Množina $X \subseteq \mathbb{R}^n$ je konvexní, právě když je uzavřená na konvexní kombinace.

- **Konvexní obal** bodů $\mathbf{x}_1, \dots, \mathbf{x}_k$ je množina všech jejich konvexních kombinací:

$$\text{conv}\{\mathbf{x}_1, \dots, \mathbf{x}_k\} = \{\alpha_1\mathbf{x}_1 + \dots + \alpha_k\mathbf{x}_k \mid \alpha_1 + \dots + \alpha_k = 1, \alpha_1, \dots, \alpha_k \geq 0\}.$$

Obecněji: **konvexní obal množiny** $X \subseteq \mathbb{R}^n$ je průnik všech konvexních množin, které množinu X obsahují. Značíme $\text{conv } X$.

Příbuzné kombinace, obaly a množiny

Lineární kombinace $\alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k$ vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k$ se nazývá jejich

afinní kombinace, jestliže $\alpha_1 + \dots + \alpha_k = 1$.

nezáporná kombinace, jestliže $\alpha_1, \dots, \alpha_k \geq 0$.

konvexní kombinace, jestliže $\alpha_1 + \dots + \alpha_k = 1$, $\alpha_1, \dots, \alpha_k \geq 0$.

Množina, která je uzavřená na

lineární kombinace, se nazývá **lineární podprostor**.

afinní kombinace, se nazývá **afinní podprostor**.

nezáporné kombinace, se nazývá **konvexní kužel**.

konvexné kombinace, se nazývá **konvexní množina**.

Některé operace zachovávající konvexitu množin

- **Průnik.** Průnik konvexních množin je konvexní množina.
- **Afinní zobrazení.** Necht' $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ je afinní zobrazení, tj.

$$\mathbf{f}(\mathbf{x}) = \mathbf{Ax} + \mathbf{b}.$$

Jestliže $X \subseteq \mathbb{R}^n$ je konvexní množina, pak

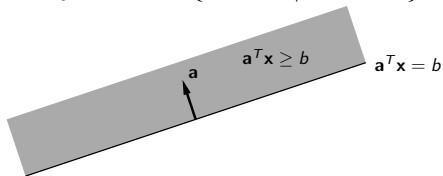
$$\mathbf{f}(X) = \{ \mathbf{f}(\mathbf{x}) \mid \mathbf{x} \in X \}$$

je také konvexní množina.

Důkazy obou tvrzení jsou snadné.

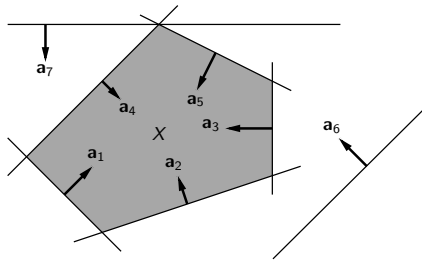
Konvexní mnohostěny

- **Nadrovina** je množina $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = b \}$, kde $\mathbf{0} \neq \mathbf{a} \in \mathbb{R}^n$, $b \in \mathbb{R}$.
- (Uzavřený) **poloprostor** je množina $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} \geq b \}$



- **Konvexní mnohostěn** je průnik konečně mnoha poloprosorů

$$X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_i^T \mathbf{x} \geq b_i \quad \forall i = 1, \dots, m \} = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A} \mathbf{x} \geq \mathbf{b} \}$$



Dimenze konvexního mnohostěnu X se definuje jako $\dim X = \dim \text{aff } X$. 5/16

Příklady jednoduchých konvexních mnohostěnů v \mathbb{R}^n :

- \emptyset, \mathbb{R}^n
- afinní podprostor (bod, přímka, rovina, nadrovina, ...)
- polopřímka $\{ \mathbf{x} + \alpha \mathbf{v} \mid \alpha \geq 0 \}$
- poloprostor
- hyperkrychle $[a, b]^n = \{ \mathbf{x} \in \mathbb{R}^n \mid a \leq x_i \leq b \forall i \}$
- hyperkvádr (box) $[a_1, b_1] \times \dots \times [a_n, b_n] = \{ \mathbf{x} \in \mathbb{R}^n \mid a_i \leq x_i \leq b_i \forall i \}$
- simplex: konvexní obal afinně nezávislé množiny bodů
 - jednotkový simplex $\{ \mathbf{x} \in \mathbb{R}^n \mid x_i \geq 0, \sum_{i=1}^n x_i \leq 1 \}$
 - standardní (= pravděpodobnostní) simplex $\{ \mathbf{x} \in \mathbb{R}^n \mid x_i \geq 0, \sum_{i=1}^n x_i = 1 \}$
- křížový polytop (*cross-polytope*) $\{ \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i| \leq 1 \}$
- polyhedrální konvexní kužel (je zároveň konvexní kužel a konvexní mnohostěn)
- nezáporný kužel \mathbb{R}_+^n (kde \mathbb{R}_+ je množina nezáporných reálných čísel)

Příklady konvexních množin, které nejsou mnohostěny:

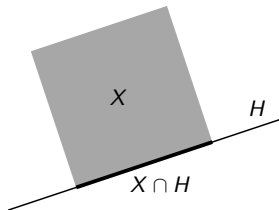
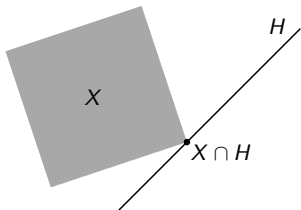
- Koule $\{ \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_2 \leq 1 \}$ pro $n \geq 2$ (průnik nekonečně mnoha poloprostorů)
- otevřený poloprostor $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} > b \}$

Opěrná nadrovina

Opěrná nadrovina konvexní množiny $X \subseteq \mathbb{R}^n$ je nadrovina

$$H = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T \mathbf{x} = b \}$$

taková, že $X \cap H \neq \emptyset$ a $\mathbf{a}^T \mathbf{x} \geq b$ pro všechna $\mathbf{x} \in X$.



Tvrzení: Každým hraničním bodem konvexního mnohostěnu prochází aspoň jedna opěrná nadrovina.

Stěny mnohostěnu

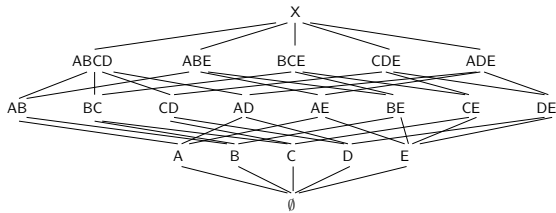
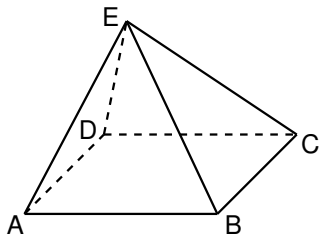
Stěna konvexního mnohostěnu X je množina $X \cap H$, kde H je nějaká opěrná nadrovina mnohostěnu X . Stěny jsou také \emptyset a X .

Ekvivalentní definice: Stěna mnohostěnu X je množina

$$\operatorname{argmin}_{\mathbf{x} \in X} f(\mathbf{x})$$

kde f je nějaká lineární funkce. (Neboť H je optimální vrstevnice funkce f .)

- Stěna dimenze 0 se nazývá **vrchol**.
- Stěna dimenze 1 se nazývá **hrana**.
- Stěna dimenze $\dim X - 1$ se nazývá **faseta**.



Extremální body

Definice

Bod $\mathbf{x} \in X$ je **extremální bod** konvexní množiny X , jestliže

$$\mathbf{x}_1, \mathbf{x}_2 \in X, \mathbf{x} = \frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_2) \implies \mathbf{x}_1 = \mathbf{x}_2$$

Tvrzení: Bod $\mathbf{x} \in X$ je vrchol, právě když je to extremální bod.

Pozorování: Počet extremálních bodů některých mnohostěnů je exponenciální v m .

Charakterizace extrémálních bodů mnohostěnu

$$X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_i^T \mathbf{x} \geq b_i \quad \forall i = 1, \dots, m \} = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \geq \mathbf{b} \}$$

Pro $I \subseteq \{1, \dots, m\}$ označme jako \mathbf{A}_I řádky matice \mathbf{A} s indexy I . Podobně pro \mathbf{b}_I .

Věta

Nechť $\mathbf{x} \in X$ a $I = \{ i \mid \mathbf{a}_i^T \mathbf{x} = b_i \}$. Následující tvrzení jsou ekvivalentní:

- Bod \mathbf{x} je extrémální bod mnohostěnu X .
- Matice \mathbf{A}_I má lin. nezávislé sloupce (tj. soustava $\mathbf{A}_I \mathbf{x} = \mathbf{b}_I$ má jediné řešení).

Důkaz \Rightarrow :

Je $\mathbf{a}_i^T \mathbf{x} > b_i$ pro všechna $i \notin I$. Tedy existuje $\epsilon > 0$ tak, že $\mathbf{a}_i^T \mathbf{x} \pm \epsilon \geq b_i$ pro všechna $i \notin I$.

Předpokládej, že sloupce \mathbf{A}_I jsou lineárně závislé, tj. $\mathbf{A}_I \mathbf{v} = \mathbf{0}$ pro nějaké $\mathbf{v} \neq \mathbf{0}$.

Pak $\mathbf{A}(\mathbf{x} \pm \epsilon \mathbf{v}) \geq \mathbf{b}$ neboli $\mathbf{x} \pm \epsilon \mathbf{v} \in X$. Protože $\mathbf{x} = \frac{1}{2}((\mathbf{x} - \epsilon \mathbf{v}) + (\mathbf{x} + \epsilon \mathbf{v}))$, máme spor.

Důkaz \Leftarrow :

Nechť $\mathbf{x}_1, \mathbf{x}_2 \in X$ a $\mathbf{x} = \frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_2)$, tedy

$$\mathbf{A}_I \mathbf{x}_1 \geq \mathbf{b}_I$$

$$\mathbf{A}_I \mathbf{x}_2 \geq \mathbf{b}_I$$

$$\mathbf{A}_I \mathbf{x} = \frac{1}{2}(\mathbf{A}_I \mathbf{x}_1 + \mathbf{A}_I \mathbf{x}_2) = \mathbf{b}_I$$

Tedy $\mathbf{A}_I \mathbf{x}_1 \geq \frac{1}{2}(\mathbf{A}_I \mathbf{x}_1 + \mathbf{A}_I \mathbf{x}_2) = \mathbf{A}_I \mathbf{x}_2$. Z toho plyne $\mathbf{A}_I \mathbf{x}_1 = \mathbf{A}_I \mathbf{x}_2$.

Protože \mathbf{A}_I má l.n. sloupce, plyne z toho $\mathbf{x}_1 = \mathbf{x}_2$.

Nalezení všech extrémálních bodů mnohostěnu

Z Věty plyne **algoritmus**: Pro každou neprázdnou $I \subseteq \{1, \dots, m\}$ otestuj:

- Má soustava $\mathbf{A}_I \mathbf{x} = \mathbf{b}_I$ právě jedno řešení?
- Splňuje toto řešení \mathbf{x} soustavu $\mathbf{Ax} \geq \mathbf{b}$?

Pokud oboje platí, pak \mathbf{x} je extrémální bod.

Příklad: $X = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0, x + y \geq 1\} = \{\mathbf{x} \in \mathbb{R}^2 \mid \mathbf{Ax} \geq \mathbf{b}\}$ kde

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Zkoumejme neprázdné podmnožiny $I \subseteq \{1, 2, 3\}$:

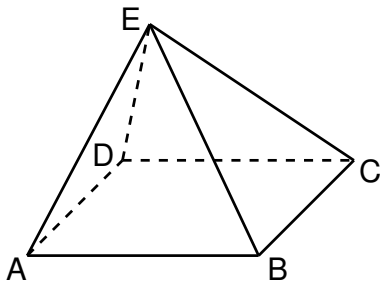
- $I = \{1\}$ a $I = \{2\}$: soustava $\mathbf{A}_I \mathbf{x} = \mathbf{b}_I$ má nekonečně mnoho řešení.
- $I = \{1, 2\}$: soustava má jediné řešení $\mathbf{x} = (0, 0)$, které nesplňuje $\mathbf{Ax} \geq \mathbf{b}$.
- $I = \{1, 3\}$: soustava má jediné řešení $\mathbf{x} = (0, 1)$, které splňuje $\mathbf{Ax} \geq \mathbf{b}$.
- $I = \{2, 3\}$: soustava má jediné řešení $\mathbf{x} = (1, 0)$, které splňuje $\mathbf{Ax} \geq \mathbf{b}$.
- $I = \{1, 2, 3\}$: soustava nemá řešení.

Závěr: Mnohostěn má dva extrémální body $(0, 1)$ a $(1, 0)$.

Geometrický pohled

$$X = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \geq \mathbf{b} \}$$

- Řešení soustavy $\mathbf{A}_I \mathbf{x} = \mathbf{b}_I$ je průnik nadrovin $\mathbf{a}_i^T \mathbf{x} = b_i$ s indexy $i \in I$.
- Soustava má jediné řešení, právě když se nadroviny I protínají v jediném bodě.
- Tento bod je extrémální bod mnohostěnu, jestliže v něm leží (splňuje $\mathbf{Ax} \geq \mathbf{b}$).



Extremální body stěny

Lema

Nechť H je opěrná nadrovina konvexní množiny X .

Pak každý extrémální bod množiny $X \cap H$ je extrémální bod množiny X .

Důkaz:

Nechť $x \in X \cap H$ není extrémální bod X . Tedy $x = \frac{1}{2}(x_1 + x_2)$ pro $x_1, x_2 \in X$ kde $x_1 \neq x_2$.

Označíme-li $H = \{x \in \mathbb{R}^n \mid a^T x = b\}$, pak

$$a^T x_1 \geq b$$

$$a^T x_2 \geq b$$

$$a^T x = \frac{1}{2}(a^T x_1 + a^T x_2) = b$$

Tedy $a^T x_1 \geq \frac{1}{2}(a^T x_1 + a^T x_2) \leq a^T x_2$, z toho $a^T x_1 = a^T x_2 = b$, tedy $x_1, x_2 \in H$.

Tedy x není extrémální bod množiny $X \cap H$.

Kdy má mnohostěn extrémální bod?

Věta

Jestliže neprázdný konvexní mnohostěn neobsahuje přímku, pak má (aspoň jeden) extrémální bod.

Důkaz: Indukcí podle dimenze mnohostěnu.

Věta triviálně platí pro mnohostěny dimenze 0 (body).

Důkaz, že když věta platí pro mnohostěny dimenze $< n$, pak platí pro mnohostěny dimenze n :

- Necht' $P \subseteq \text{aff } X$ je přímka protínající mnohostěn X .
Protože X neobsahuje žádnou přímku, P protíná hranici X v aspoň jednom bodě \mathbf{x} .
- V bodě \mathbf{x} existuje opěrná nadrovina H neobsahující X . Tedy $\dim(X \cap H) < \dim X$.
- Mnohostěn $X \cap H$ neobsahuje přímku (protože X ji neobsahuje).
- Dle předpokladu má tedy $X \cap H$ extrémální bod. Ten je dle Lematu extrémální bod X .

Příklady mnohostěnu obsahujících přímku:

- 'panel' $\{ \mathbf{x} \in \mathbb{R}^n \mid b \leq \mathbf{a}^T \mathbf{x} \leq c \}$
- $\{ \mathbf{x} + t\mathbf{v} \mid \mathbf{x} \in X, t \in \mathbb{R} \}$ kde $X \subseteq \mathbb{R}^n$ je mnohostěn a $\mathbf{0} \neq \mathbf{v} \in \mathbb{R}^n$

Minimum lineární funkce na mnohostěnu

Věta

Nechť konvexní mnohostěn X neobsahuje přímku.

Nechť lineární funkce f nabývá na X minima.

Potom existuje extrémální bod X , v němž f nabývá na X minima.

Důkaz: Nechť H je optimální vrstevnice funkce f na X . Tedy H je opěrná nadrovina X .
 X neobsahuje přímku, tedy $X \cap H$ neobsahuje přímku, tedy dle Věty má $X \cap H$ extrémální bod.
Dle Lematu je tento bod také extrémální bod mnohostěnu X .

Tedy minimum f na X lze najít 'hrubou silou':

Projdi všechny extrémální body mnohostěnu a vyber ten s nejmenší hodnotou f .

Ale potíž: Extrémálních bodů může být příliš mnoho.

Simplexový algoritmus na řešení LP (Dantzig, 1947):

- Začne z nějakého extrémálního bodu mnohostěnu.
- V každé iteraci přejde (po hraně) k jinému extrémálnímu bodu s lepší (nebo aspoň stejnou) hodnotou f .

Algoritmy vnitřního bodu (Interior Point Algorithms, IPM)

Druhá (kromě simplexové metody) široce užívaná třída algoritmů na řešení LP. Pro postupně se zmenšující $\mu > 0$ řešíme posloupnost úloh

$$\min \{ \mathbf{c}^T \mathbf{x} - \mu \sum_j \log x_j \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} \}$$

kde $-\sum_j \log x_j$ je **bariérová funkce**.

- Pro $\mu \rightarrow 0$ je úloha ekvivalentní původnímu LP.
- Pro $\mu \rightarrow \infty$ je člen $\mathbf{c}^T \mathbf{x}$ zanedbatelný.

Optimální řešení \mathbf{x} se nazývá **analytický střed** soustavy $\mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}$.

Na řešení úlohy použijme metodu Lagrangeových multiplikátorů:

$$L(\mathbf{x}, \mathbf{y}) = \mathbf{c}^T \mathbf{x} - \mu \sum_j \log x_j + \mathbf{y}^T (\mathbf{b} - \mathbf{Ax})$$

Podmínka stacionarity:

$$\begin{aligned} \mu/x_j + \mathbf{a}_j^T \mathbf{y} &= c_j \quad \forall j \\ \mathbf{Ax} &= \mathbf{b} \end{aligned}$$

Tuto soustavu rovnic řešíme Newtonovou metodou.

Podmínky $x_j > 0$ jsou zajištěny implicitně.