České vysoké učení technické v Praze

Fakulta elektrotechnická

Sbírka řešených příkladů

Optimalizace a teorie her

Jakub Adamec

Praha, 2025

https://github.com/knedl1k/A8B01OGT

Obsah

Strana

1 První týden **2**

1.1 Důkaz souvislosti minima a maxima 2
1.2 Hledání přípustných množin 2
1.3 Hledání přípustných množin 2
1.4 Maximalisační úloha 3
1.5 Minimalisační úloha 3
1.6 Optimalisační úloha s nadrovinami 4
1.7 Uzavřená úsečka 6
1.8 Je nadrovina konvexní? 6
1.9 Je uzavřený poloprostor konvexní? 6
1.10 Je uzavřená koule konvexní? 6
1.11 Je okolí konvexní? 6
1.12 Je průnik množin konvexní? 7
1.13 Důkaz, že rozdíl a sjednocení nezachovává konvexitu 7
1.14 Důkaz, že afinní zobrazení je konvexní 7
1.15 Důkaz, že obraz konvexní množiny při afinním zobrazení je konvexní 8
1.16 Důkaz, že kartézský součin je konvexní 8
1.17 Určení definitnosti matic 9
1.18 Existence matice 10

2 Druhý týden **12**

2.1 Věta o nejlepší aproximaci 12
2.2 Projekce bodu a variační nerovnost 12
2.3 Koule? 13
2.4 Věta o ortogonálním rozkladu 13

3 Třetí týden **15**

3.1 Metoda nejmenších čtverců 15
3.2 Příklad výpočtu metody nejmenších čtverců 15
3.3 Příklad výpočtu metody nejmenších čtverců 16
3.4 Věta o oddělitelnosti bodu a konvexní množiny 16
3.5 Příklad na použití věty o oddělitelnosti nadrovinou 17
3.6 Lemma neprázdné uzavřené konvexní 17
3.7 Farkasovo lemma 18

3.8 Krajní body konvexní množiny 18
3.9 Kreinova-Milmanova věta 19

4 Čtvrtý týden 20

4.1 Konvexní funkce 20
4.2 Příklad konvexní funkce 20
4.3 Příklad konvexní funkce 20
4.4 Dolní úrovňová množina 21
4.5 Použití dolní úrovňové množiny 21
4.6 Součet a součin zachovávají konvexitu 21
4.7 Příklad ověření konvexity 22
4.8 Skládání zachovává konvexitu 22
4.9 Věta o extrémech konvexních funkcí 23
4.10 Věta o konvexitě a první derivaci 23
4.11 Věta o konvexitě a druhé derivaci 24
4.12 Příklad ověření konvexnosti pomocí derivace 25
4.13 Příklad ověření konvexnosti pomocí derivace 25

5 Pátý týden 26

6 Šestý týden 27

7 Sedmý týden 28

8 Osmý týden 29

9 Devátý týden 30

10 Desátý týden 31

11 Jedenáctý týden 32

12 Dvanáctý týden 33

13 Třináctý týden 34

14 Čtrnáctý týden 35

Úvod

Tento text není psán jako učebnice, nýbrž jako soubor řešených příkladů, u kterých je vždy uveden celý korektní postup a případné moje poznámky, které často nebývají formální, a tedy by neměly být používány při oficálním řešení problémů, například při zkoušce. Jedná se pouze o pokus předat probíranou látku z různých úhlů pohledu, pokud by korektní matematický nebyl dostatečně výřečný.

Velmi ocením, pokud čtenáři zašlou své podněty, úpravy anebo připomínky k textu. Budu rád za všechnu konstruktivní kritiku a nápady na změny. Dejte mi také prosím vědět, pokud v textu objevíte překlepy, chyby a jiné.

Errata a aktuální verse textu bude na stránce `https://github.com/knedl1k/A8B01OGT`.

Poděkování. Rád bych poděkoval docentu Martinu Bohatovi nejen za zadání, okolo kterých je postavena celá sbírka, ale také za celý předmět Optimalizace a teorie her.

Text je vysázen makrem LaTeX Leslieho Lamporta s využitím balíků `hypperref` Sebastiana Rahtze a Heiko Oberdiek.

Stručné informace o textu

Všechny růžové texty jsou zároveň hypertextové odkazy. Často jsou použity u přednáškových příkladů, pomocí nichž lze vidět ukázkové řešení příkladu na přednášce.

U každého příkladu je pro ušetření místa a zpřehlednění sbírky řešení jednotlivých příkladů ihned pod zadáním.

1 První týden

1.1 Důkaz souvislosti minima a maxima

Tvrzení. Pro $f : D \to \mathbb{R}, M \subseteq D, \hat{x} \in M$ platí:

(1) $\hat{x} \in \underset{x\in M}{\operatorname{argmin}} f(x) \iff \hat{x} \in \underset{x\in M}{\operatorname{argmax}}(-f(x))$,

(2) jesliže $\hat{x} \in \underset{x\in M}{\operatorname{argmin}} f(x)$, pak $\min_{x\in M} f(x) = -\max_{x\in M}(-f(x))$.

Důkaz.

(1) $\hat{x} \in \underset{x\in M}{\operatorname{argmin}} f(x)$, tj. $f(\hat{x}) \leq f(x), \forall x \in M \underset{\cdot(-1)}{\iff} -f(\hat{x}) \geq -f(x), \forall x \in M$, tj. $\hat{x} \in \underset{x\in M}{\operatorname{argmax}}(-f(x))$. □

(2) Ať $\hat{x} \in \underset{x\in M}{\operatorname{argmin}} f(x)$, pak $\min_{x\in M} f(x) = f(\hat{x}) = -(-f(\hat{x})) \overset{(1)}{=} -\max_{x\in M}(-f(x))$. □

1.2 Hledání přípustných množin

$$\begin{aligned} \text{minimalizujte } & x^2 + 1 \\ \text{za podmínek } & \frac{3}{x} \leq 1, \\ & x \in \mathbb{N}. \end{aligned}$$

Upravíme podmínky a uděláme jejich průnik: $(x - 3 \geq 0) \land (x \in \mathbb{N}) \Rightarrow M = \mathbb{N} \setminus \{1, 2\}$.

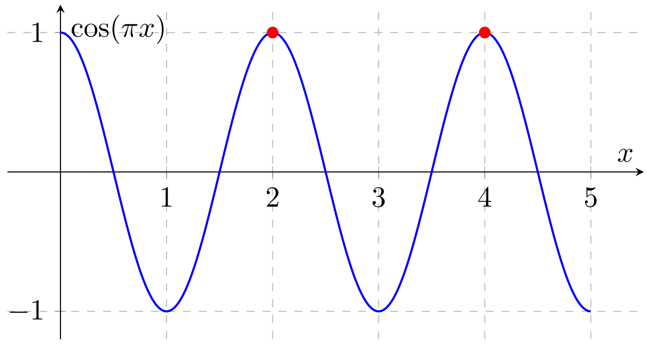
Úvahou pak lze uhodnout minimum - minimum leží v bodě $x = 3$.

1.3 Hledání přípustných množin

$$\begin{aligned} \text{maximalizujte } & \ln x \\ \text{za podmínek } & x \leq 5, \\ & \cos(\pi x) = 1. \end{aligned}$$

$D(f) = (0, \infty)$.

Udělejme průnik definičního oboru funkce a podmínek: $(x \in (0, \infty)) \land (x \leq 5) \land (\cos(\pi x) = 1)$.



Očividně tedy $M = \{2, 4\}$.

Úvahou pak lze uhodnout $\underset{x\in M}{\operatorname{argmax}} \ln x = \{4\}$.

1.4 Maximalisační úloha

Banka nabízí dva investiční produkty. Očekávaný měsíční výnos prvního investičního produktu (v tis. Kč) při investici x (v tis. Kč) je $\frac{2x}{4x+25}$ a očekávaný měsíční výnos druhého invetičního produktu (v tis. Kč) při investici x (v tis. Kč) je $\frac{x}{x+50}$. Jakým způsobem má investor rozdělit částku $c = 100000$ Kč mezi uvedené dva produkty tak, aby celkový očekávaný měsíční výnos byl co největší?

$$\begin{aligned} \text{maximalisujme } & \frac{x}{x+50} + \frac{2y}{4y+25} \\ \text{za podmínek } & x + y = 100, \\ & x, y \geq 0. \end{aligned}$$

Vyjádřeme si jednu proměnnou v závislosti na druhé, například $x = 100 - y$. Následně dosadíme do úlohy a vyšetříme stacionární body pomocí první derivace.

$$\frac{\mathrm{d}}{\mathrm{d}y}\left(\frac{100-y}{150-y} + \frac{2y}{4y+25}\right) = \frac{-50}{(150-y)^2} + \frac{50}{(4y+25)^2} \stackrel{!}{=} 0$$

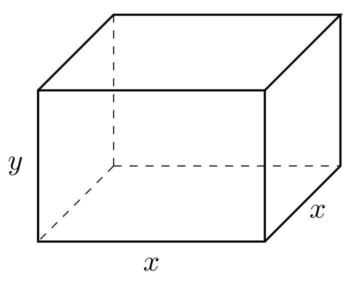
Zbavme se zlomků:

$$\begin{aligned} -50(4y+25)^2 + 50(150-y)^2 &= 0 \\ (150-y)^2 - (4y+25)^2 &= 0 \\ (150 - y - 4y - 25) - (150 - y + 4y + 25) &= 0 \\ (125 - 5y)(175 + 3y) &= 0 \\ y_1 = 25, y_2 \approx -58.3 \end{aligned}$$

Tedy aby byly splněny všechny podmínky je jediné možné řešení $y = 25 \to x = 75$.

1.5 Minimalisační úloha

Ve firmě potřebují nalézt rozměry otevřené krabice (tj. krabice bez horní stěny) se čtvercovou podstavou o objemu 10 dm^3 tak, aby obsah plochy jejího pláště byl co nejmenší. Formulujte odpovídající optimalisační úlohu za předpokladu, že krabice je vyrobena z materiálu, jehož tloušťka je zanedbatelná. Tuto úlohu poté vyřešte.



$$\begin{aligned} \text{minimalisujme } & 4xy + x^2 \\ \text{za podmínek } & x^2 y = 10, \\ & x, y > 0. \end{aligned}$$

Vyjádřeme si jednu proměnnou v závislosti na druhé, například $y = \frac{10}{x^2}$. Následně dosadíme do úlohy a vyšetříme stacionární body pomocí první derivace.

$$\frac{\mathrm{d}}{\mathrm{d}y}\left(4x\frac{10}{x^2}+x^2\right)=\frac{-40}{x^2}+2x\stackrel{!}{=}0$$

Zbavme se zlomků:

$$\begin{aligned}-40+2x^3&=0\\ x^3&=20\\ x&=\sqrt[3]{20}\end{aligned}$$

Tedy jediné možné řešení $x=\sqrt[3]{20}\to y=\frac{10}{(\sqrt[3]{20})^2}=\sqrt[3]{\frac{5}{2}}$.

1.6 Optimalisační úloha s nadrovinami

V $\mathbb{R}^n$ jsou dány množiny bodů $A=\{a_1,\dots,a_k\}$ a $B=\{b_1,\dots,b_t\}$. Ať $w\in\mathbb{R}^n$ a $\lambda\in\mathbb{R}$. Předpokládejme, že H je nadrovina o rovnici $\langle x,w\rangle+\lambda=0$, H_1 je nadrovina o rovnici $\langle x,w\rangle+\lambda=1$ a H_2 je nadrovina o rovnici $\langle x,w\rangle+\lambda=-1$.

(a) Ukažte, že vzdálenost mezi nadrovinami H_1 a H_2 je $\frac{2}{||w||}$. Dále ukažte, že $\frac{1}{||w||}$ je vzdálenost H od H_1 a také vzdálenost H od H_2.

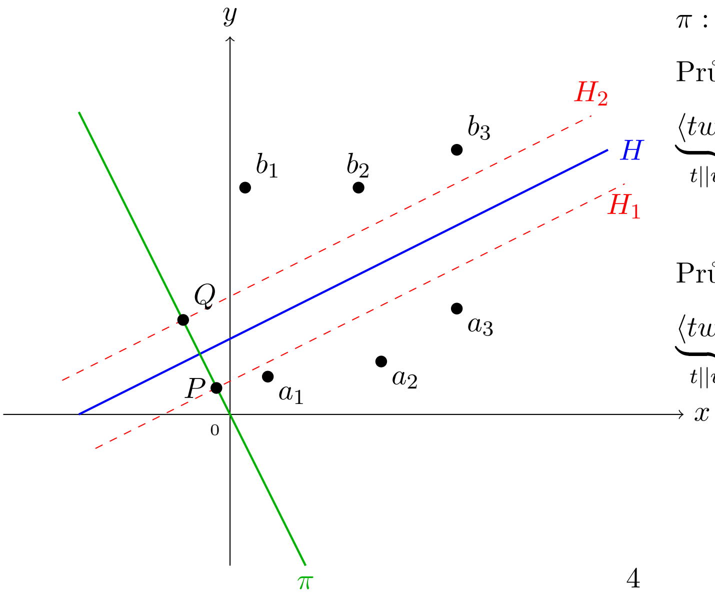
(b) Iterpretujte optimalisační úlohu

$$\begin{aligned}\text{maximalisujte } g(w,\lambda)&=\frac{2}{||w||}\\ \text{za podmínek } \langle a_i,w\rangle+\lambda&\geq 1 \text{ pro všechna } i=1,\dots,k,\\ \langle b_i,w\rangle+\lambda&\leq -1 \text{ pro všechna } j=1,\dots,l.\end{aligned}$$

(c) Ukažte, že $(\hat{w},\hat{\lambda})$ je řešením úlohy z předchozího bodu právě tehdy, když je řešením úlohy (kvadratického programování) ve tvaru

$$\begin{aligned}\text{minimalisujte } h(w,\lambda)&=\frac{1}{2}||w||^2\\ \text{za podmínek } \langle a_i,w\rangle+\lambda&\geq 1 \text{ pro všechna } i=1,\dots,k,\\ \langle b_i,w\rangle+\lambda&\leq -1 \text{ pro všechna } j=1,\dots,l.\end{aligned}$$

(a)



$\pi: x=t\cdot w, t\in\mathbb{R}$.

Průsečík Q:

$\underbrace{\langle tw,w\rangle}_{t||w||^2}+\lambda=1\to t=\frac{1-\lambda}{||w||^2}\Rightarrow Q=\frac{1-\lambda}{||w||^2}w$

Průsečík P:

$\underbrace{\langle tw,w\rangle}_{t||w||^2}+\lambda=-1\to t=\frac{-1-\lambda}{||w||^2}\Rightarrow P=\frac{-1-\lambda}{||w||^2}w$

Pak vzdálenost mezi nadrovinami H_1 a H_2 je dána rozdílem průsečíků P a Q v normě. Tedy:

$$||Q - P|| = \left\| \frac{1-\lambda}{||w||^2} w + \frac{1+\lambda}{||w||^2} w \right\| = \left\| \frac{2w}{||w||^2} \right\| = \frac{2}{||w||^2} ||w|| = \frac{2}{||w||}.$$

To je príma, to jsme přesně chtěli. $\square$

(b)

(c) V úloze (b) maximalisujeme zlomek, kde se proměnná nachází ve jmenovateli. Tedy snažíme se najít co nejmenší možný jmenovatel, aby zlomek měl co největší hodnotu. Můžeme úlohu převrátit a minimalisovat samotný jmenovatel. Protože násobení je lineární a zachovává nám všechny nerovnosti, můžeme různě modifikovat jakou konstantou násobíme námi minimalisovanou proměnnou. Zároveň si můžeme dovolit umocnit normu, protože i to nám zachová všechny nerovnosti. Zde si tedy chytře zvolíme násobení $\frac{1}{2}$, protože při následném hledání stacionárních bodů funkce nám vyskočí z kvadrátu dvojka, jenž pěkně pokrátíme. Podmínky nám zůstaly stejné, není co řešit.

Konvexní množiny

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

1.7 Uzavřená úsečka

Nechť $x, y \in \mathbb{R}^n$. Množina
$$[x, y] \coloneqq \{\lambda x + (1-\lambda)y \mid 0 \leq \lambda \leq 1\}$$
se nazývá uzavřená úsečka s krajními body x a y.

1.8 Je nadrovina konvexní?

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

Důkaz.

Ať $x, z \in H(y, \alpha), \lambda \in [0, 1]$.
Cíl: $\lambda x + (1-\lambda)z \in H(y, \alpha)$. Tedy dokazujeme podle definice.

$\langle \lambda x + (1-\lambda)z, y\rangle = \lambda \underbrace{\langle x, y\rangle}_{\alpha} + (1-\lambda) \underbrace{\langle z, y\rangle}_{\alpha} = \lambda\alpha + (1-\lambda)\alpha = \alpha.$

$\Rightarrow \lambda x + (1-\lambda)z \in H(y, \alpha)$. $\square$

1.9 Je uzavřený poloprostor konvexní?

1.10 Je uzavřená koule konvexní?

Definice uzavřené koule: $B(a, r) = \{a \in \mathbb{R}^n \mid ||x - a|| \leq r\}$, o středu $a \in \mathbb{R}^n$ a poloměru $r > 0$.

Důkaz.

Ať $x, y \in \mathbb{R}^n, \lambda \in [0, 1]$.
Cíl: $||[\lambda x + (1-\alpha)y] - a|| \leq r$. Tedy za x z definice dosadíme úsečku mezi body x a y, které jsme si vybrali a chceme ukázat, že i tato úsečka leží v uzavřené kouli, dle definice.

$$\begin{aligned}||[\lambda x + (1-\alpha)y] - a|| &= ||\lambda x - (1-\lambda)a + (1-\lambda)y - \lambda a|| = ||\lambda(x-a) + (1-\lambda)(y-a)|| \\ &\leq \lambda||\underbrace{x-a}_{\leq r}|| + (1-\lambda)||\underbrace{y-a}_{\leq r}|| \leq \lambda r + (1-\lambda)r = r. \quad \square\end{aligned}$$

1.11 Je okolí konvexní?

Definice okolí: $B(a, r) = \{a \in \mathbb{R}^n \mid ||x - a|| < r\}$, o středu $a \in \mathbb{R}^n$ a poloměru $r > 0$.

Důkaz.

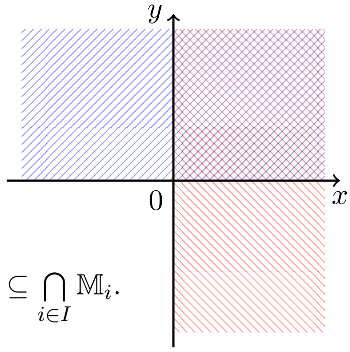
Ať $x, y \in \mathbb{R}^n, \lambda \in [0, 1]$.
Cíl: $||[\lambda x + (1-\alpha)y] - a|| < r$. Dle definice.

$$\begin{aligned}||[\lambda x + (1-\alpha)y] - a|| &= ||\lambda x - (1-\lambda)a + (1-\lambda)y - \lambda a|| = ||\lambda(x-a) + (1-\lambda)(y-a)|| \\ &\leq \lambda||\underbrace{x-a}_{<r}|| + (1-\lambda)||\underbrace{y-a}_{<r}|| < \lambda r + (1-\lambda)r = r. \quad \square\end{aligned}$$

1.12 Je průnik množin konvexní?

Úvaha pro 2 množiny ve $\mathbb{R}^2$:

Mějme jednu modrou ($y \geq 0$) a druhou červenou ($x \geq 0$) konvexní množinu. Jejich průnik je pak nezáporný ortant, tedy
$\mathbb{R}^n_+ = \{(x_1, \ldots, x_n)^T \in \mathbb{R}^n \mid x_1 \geq 0, \ldots, x_n \geq 0\}$.



Visuálně je průnik nekonvexní.

Důkaz.

Nechť $x, y \in \bigcap\limits_{i \in I} \mathbb{M}_i, \forall i \in I \implies [x, y] \in \mathbb{M}_i, \forall i \in I \implies [x, y] \subseteq \bigcap\limits_{i \in I} \mathbb{M}_i$.

□

1.13 Důkaz, že rozdíl a sjednocení nezachovává konvexitu

Mějme $[0, 1] \setminus (0, 1) = \{0, 1\} = \{0\} \cup \{1\}$.

$[0, 1]$ a $(0, 1)$ jsou konvexní množiny. Jejich rozdíl ale už konvexní není.
$\{0\}$ a $\{1\}$ jsou konvexní množiny. Jejich sjednocení ale už konvexní není.

Afinní zobrazení

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

1.14 Důkaz, že afinní zobrazení je konvexní

Tvrzení.

Nechť $f : \mathbb{R}^n \to \mathbb{R}^m$. Pak f je afinní $\iff$ 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.

„$\Rightarrow$“: Ať $f(x) = Ax + b$, kde $A \in \mathbb{M}_{m,n}(\mathbb{R})$, $b \in \mathbb{R}^n$.

Ať $x, y \in \mathbb{R}^n, \lambda \in \mathbb{R}$.

$$f(\lambda x + (1 - \lambda)y) = A[\lambda x + (1 - \lambda)y] + b = \lambda Ax + (1 - \lambda)Ay + \lambda b + (1 - \lambda)b =$$

$$\lambda \underbrace{(Ax + b)}_{f(x)} + (1 - \lambda) \underbrace{(Ay + b)}_{f(y)} = \lambda f(x) + (1 - \lambda)f(y). \quad \square$$

„$\Leftarrow$“: Cíl: Ukázat, že f je afinní, tedy $f(x) = Ax + b$.
Zvolme $\varphi(x) = f(x) - f(0)$.
Pokud je f afinní, pak zobrazení φ by mělo být dáno jako Ax, tedy být lineární.
Cíl: φ je lineární zobrazení.

Musíme ověřit uzavřenost na násobení a sčítání z definice.

(1) Ať $x \in \mathbb{R}^n$, $\alpha \in R$.
Cíl: $\varphi(\alpha x) = \alpha\varphi(x)$.

$$\varphi(\alpha x) = f(\alpha x) - f(0) = f(\alpha x + (1-\alpha)0) - f(0) = \alpha f(x) + (1-\alpha) f(0) - f(0) =$$

$$\alpha f(x) - \alpha f(0) = \alpha(f(x) - f(0)) = \alpha\varphi(x - 0). \quad \square$$

(2) Ať $x, y \in \mathbb{R}^n$.
Cíl: $\varphi(x+y) = \varphi(x) + \varphi(y)$.

$$\varphi(x+y) = \varphi\left(2\left(\frac{1}{2}(x+y)\right)\right) \overset{(1)}{=} 2\varphi\left(\frac{1}{2}(x+y)\right) = 2\left[f(\frac{1}{2}x + \frac{1}{2}y) - f(0)\right] = 2\left[\frac{1}{2}f(x) + \frac{1}{2}f(y) - f(0)\right] =$$

$$f(x) + f(y) - f(0) - f(0) = \underbrace{f(x) - f(0)}_{\varphi(x)} + \underbrace{f(y) - f(0)}_{\varphi(y)} = \varphi(x) + \varphi(y). \quad \square$$

1.15 Důkaz, že obraz konvexní množiny při afinním zobrazení je konvexní

Tvrzení.

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

Důkaz.

Mějme $a, b \in f(C) \implies \exists x, y \in C : f(x) = a, f(y) = b$.

Dle předpokladu je C konvexní. $\implies [x, y] \subseteq C \implies \underbrace{f([x, y])}_{\subseteq f(C)} = [\underbrace{f(x)}_{a}, \underbrace{f(y)}_{b}] \subseteq f(C)$. $\square$

1.16 Důkaz, že kartézský součin je konvexní

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.

„$\Rightarrow$“: Mějme $\begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} c \\ d \end{bmatrix} \in C_1 \times C_2, \lambda \in [0, 1]$

Cíl: $\lambda \begin{bmatrix} a \\ b \end{bmatrix} + (1-\lambda) \begin{bmatrix} c \\ d \end{bmatrix} \in C_1 \times C_2$. Dle definice.

$$\lambda \begin{bmatrix} a \\ b \end{bmatrix} + (1-\lambda) \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} \lambda a \\ \lambda b \end{bmatrix} + \begin{bmatrix} (1-\lambda)c \\ (1-\lambda)d \end{bmatrix} = \begin{bmatrix} \lambda a + (1-\lambda)c \\ \lambda b + (1-\lambda)d \end{bmatrix} \in C_1 \times C_2. \quad \square$$

„$\Leftarrow$“: Definujme afinní zobrazení $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ předpisem

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

Pak f je afinní. Navíc $f(C_1 \times C_2) = C_1$. $\implies$ C_1 je konvexní, protože afinní zobrazení zachovává konvexitu. A důkaz bude obdobný pro C_2, zde zadefinujme afinní zobr. $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ předpisem

$$g(x, y) = y.$$

Pak g je afinní. Navíc $g(C_1 \times C_2) = C_2$. $\implies$ C_2 je konvexní, protože afinní zobrazení zachovává konvexitu. $\square$

1.17 Určení definitnosti matic

Určete definitnost matice A, jestliže

(a) $\begin{bmatrix} 9 & 6 \\ 6 & 4 \end{bmatrix}$;

(b) $\begin{bmatrix} 15 & 3 & 2 \\ 3 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$;

(c) $\begin{bmatrix} 4 & 2 & 2 \\ 2 & 1 & 1 \\ 2 & 1 & 0 \end{bmatrix}$;

(d) $\begin{bmatrix} 3 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$;

(e) $\begin{bmatrix} -1 & 0 & 1 \\ 0 & -2 & 2 \\ 1 & 2 & -3 \end{bmatrix}$;

(f) $\begin{bmatrix} 1 & 2 & 0 \\ 2 & 5 & 1 \\ 0 & 1 & 1 \end{bmatrix}$.

Matice, u které chceme určovat definitnost, musí být $\underbrace{\text{symetrická}}_{Q=Q^T}$.

Pak platí:

$$\begin{aligned} \langle Qx, x\rangle \geq 0 \forall x \in \mathbb{R}^n &\iff Q \text{ je positivně semidefinitní.} \\ \langle Qx, x\rangle > 0 \forall x \in \mathbb{R}^n &\iff Q \text{ je positivně definitní.} \end{aligned}$$

Analogicky pro negativně semidefinitní, respektive definitní.
Matice je indefitní pokud nesplňuje ani jednu možnost.

Pro symetrické matice také platí, že Q je negativně (semi)defitní, jestliže $(-Q)$ je positivně (semi)defintní.

Pomocí Sylvesterova kritéria lze určit positivní, či negativní definitnost. Pro případy podezření na semidefinitnost je potřeba navíc prozkoumat menší minory matice.

(a) $\begin{bmatrix} 9 & 6 \\ 6 & 4 \end{bmatrix} \to |9| = 9 > 0, \begin{vmatrix} 9 & 6 \\ 6 & 4 \end{vmatrix} = 36 - 36 = 0. \to$ podezření na positivní semidefinitnost.

Hlavní minory jsou $Q_{\{1\}}$ a $Q_{\{1,2\}}$.
Menší minory: Q_I, kde $I \subseteq \{1, \ldots, n\}$ neprázdná. Aby matice byla positivně semidefinitní, tak $\det Q_I \geq 0$.

Tedy: $Q_{\{2\}} = \begin{bmatrix} 4 \end{bmatrix}$. $\det Q_{\{2\}} = 4 > 0$.

Tedy matice $\begin{bmatrix} 9 & 6 \\ 6 & 4 \end{bmatrix}$ je positivně semidefinitní.

(b) $\begin{vmatrix} 15 & 3 & 2 \\ 3 & 1 & 0 \\ 2 & 0 & 1 \end{vmatrix} \begin{matrix} R_1 - 2R_3 \\ R_2 \\ R_3 \end{matrix} = \begin{vmatrix} 11 & 3 & 0 \\ 3 & 1 & 0 \\ 2 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 11 & 3 \\ 3 & 1 \end{vmatrix} = 11 - 9 = 2 > 0.$ Matice je positivně definitní.

(c) $Q = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 1 & 1 \\ 2 & 1 & 0 \end{bmatrix}$

Pozorování: Matice je lineárně závislá, tedy det$Q = 0$.
$Q_{\{1\}} = 4 > 0$,
$Q_{\{2\}} = 1 > 0$,
$Q_{\{3\}} = 0 = 0$.
Tedy matice je jedině positivně semidefinitní, nebo indefinitní.

Spočtěme tedy vedlejší minor, například vynechejme 1. řádek a 1. sloupec:

$\begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1 < 0$. Aby matice Q byla positivně semidefinitní, musely by i všechny vedlejší minory být ≥ 0. Protože jsme našli případ, kdy tomu tak není, matice Q je indefinitní.

(e) $\begin{bmatrix} -1 & 0 & 1 \\ 0 & -2 & 2 \\ 1 & 2 & -3 \end{bmatrix}$

Pozorování: matice může být negativně (semi)definitní, nebo indefinitní.

Využijme tedy vlastnosti symetrických matic a určeme definitnost pro matici $(-Q)$.

$-Q = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & -2 \\ -1 & -2 & 3 \end{bmatrix}$

$\det(-Q) = \begin{vmatrix} 1 & 0 & -1 \\ 0 & 2 & -2 \\ -1 & -2 & 3 \end{vmatrix} \begin{matrix} R_1 \\ R_2 \\ R_3 + R_1 + R_2 \end{matrix} = \begin{vmatrix} 1 & 0 & -1 \\ 0 & 2 & -2 \\ 0 & 0 & 0 \end{vmatrix} = 0.$

Tedy matice $(-Q)$ je positivně semidefinitní, nebo indefinitní.

$\begin{vmatrix} 1 & 0 \\ 0 & 2 \end{vmatrix} = 2 \geq 0$. $\begin{vmatrix} 2 & -2 \\ -2 & 3 \end{vmatrix} = 2 \geq 0$. $\begin{vmatrix} 1 & -1 \\ -1 & 3 \end{vmatrix} = 2 \geq 0$.

$\implies (-Q)$ je positivně semidefinitní $\iff$ Q je negativně semidefinitní.

1.18 Existence matice

Ať $A \in \mathbb{M}_n(\mathbb{R})$.

(a) Ukažte, že $\langle Ax, y \rangle = \langle x, A^T y \rangle$ pro všechna $x, y \in \mathbb{R}^n$.

(b) Ukažte, že existují matice $B, C \in \mathbb{M}_n(\mathbb{R})$ takové, že $B^T = B$, $C^T = -C$ a $A = B + C$. Jsou matice B a C určeny jednoznačně?

(c) Ukažte, že existuje symetrická matice $B \in \mathbb{M}_n(\mathbb{R})$ taková, že $\langle Ax, x \rangle = \langle Bx, x \rangle$.

Zadefinujme si vlastnost skalárního součinu: $\langle a, b \rangle = b^T a$, kde $b^T = \begin{pmatrix} b_1, \ldots, b_n \end{pmatrix}, a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$.

(a)Využijme zmíněné vlastnosti.

$\langle Ax, y \rangle = y^T A x = \underbrace{y^T (A^T)^T}_{(A^T y)^T} x = (A^T y)^T x = \langle x, A^T y \rangle. \quad \square$

(b) Pozorování: Matice B je symetrická a matice C je antisymetrická.

Zvolme: $\left.\begin{aligned} B &= \tfrac{1}{2}(A + A^T) \\ C &= \tfrac{1}{2}(A - A^T) \end{aligned}\right\} B + C = A.$

$C^T = \frac{1}{2}(A - A^T)^T = \frac{1}{2}(A^T - A) = -\frac{1}{2}(A - A^T) = -C.\checkmark$

$B^T = \frac{1}{2}(A + A^T)^T = \frac{1}{2}(A^T + A) = \frac{1}{2}(A + A^T) = B.\checkmark \quad \square$

(c) $\langle Cx, x\rangle \overset{?}{=} 0$

$$\langle Cx, x\rangle \overset{(a)}{=} \langle x, C^T x\rangle \overset{-C=C^T}{=} -\langle x, Cx\rangle = -\langle Cx, x\rangle = 0.$$

Matice C tedy nijak nepřispívá do výsledku. Takže platí $\langle Ax, x\rangle = \langle Bx, x\rangle$. $\square$

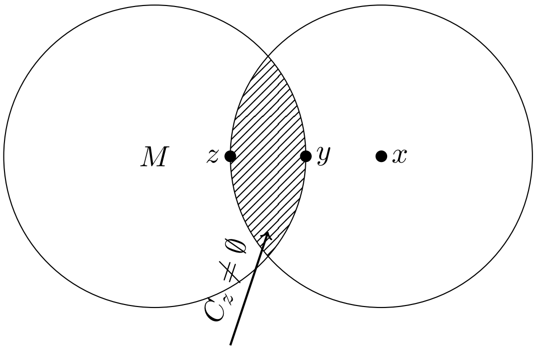
2 Druhý týden

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

Důkaz.

1. Existence
Cíl: Existuje bod minima
Úvaha:



M je obecná konvexní množina.
c x $R = ||x - z||$,
$Cz = M \cap B(x, R) = M \cap \{a \in \mathbb{R}^n \mid ||z - a|| \leq R\}$.
$\uparrow$
$\underbrace{\text{uzavřená, omezená}}_{\text{kompaktní}}$, neprázdná

Tedy $a \mapsto ||x - a||$ je spojitá.

$\Rightarrow$ Spojitost na kompaktní množině znamená, že f nabývá na C_z minima dle Weierstrassova kritéria.

Ať y je bod minima. Všechny body v M mají od x vzdálenost $\geq ||x - y||$. $\square$

2. Jednoznačnost.

Cíl: Pokud $a, b \in \mathbb{R}^n : ||x - a|| = ||x - b|| = \overbrace{\operatorname{dist}(x, M)}^{\delta}$, pak $a = b$.
Lemma, rovnoběžníkové pravidlo: $u, v \in \mathbb{R}^n \Rightarrow ||u + v||^2 + ||u - v||^2 = 2\left(||u||^2 + ||v||^2\right)$.
Důkaz lemma:

$$||u+v||^2 + ||u-v||^2 = \langle u+v, u+v\rangle + \langle u-v, u-v\rangle = ||u||^2 + 2\langle u, v\rangle + ||v||^2 + ||u||^2 - 2\langle u, v\rangle + ||v||^2$$

$$= 2\left(||u||^2 + ||v||^2\right). \quad \square$$

Důkaz jednoznačnosti:
Ať $y = \frac{1}{2}a + \frac{1}{2}b$.
Pak $\delta^2 \leq ||x - y||^2 = ||x - \frac{1}{2}a - \frac{1}{2}b||^2 = ||\frac{1}{2}(x - a) + \frac{1}{2}(x - b)||^2 = \frac{1}{4}||\underbrace{(x - a)}_{u} + \underbrace{(x - b)}_{v}||^2$

$$\overset{\text{lemma}}{=} \frac{1}{4}\left[2\left(\underbrace{||x - a||^2}_{\delta^2} + \underbrace{||x - b||^2}_{\delta^2}\right) - \underbrace{||(x - a) + (x - b)||^2}_{b-a}\right] = \delta^2 - \frac{1}{4}||b - a||^2 \Rightarrow \delta^2 \leq \delta^2 - \underbrace{\frac{1}{4}||b - a||^2}_{\leq 0 \Rightarrow a = b}.$$

2.2 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) $y = P_C(x)$, kde $P_C(x)$ je projekční operátor.

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

Důkaz.

(1) $\Rightarrow$ (2):
Ať $v_\lambda = y + \lambda(z-y)$, $\lambda \in (0,1]$.
Pak

$$||x-y||^2 \leq ||x-v_\lambda||^2 = ||x-y-\lambda(z-y)||^2 = \langle (x-y)-\lambda(z-y), (x-y)-\lambda(z-y)\rangle$$

$$||x-y||^2 \leq ||x-y||^2 - 2\lambda\langle x-y, z-y\rangle + \lambda^2||z-y||^2$$

$$\Rightarrow \langle x-y, z-y\rangle \leq \frac{\lambda}{2}||z-y||^2 \to 0 \text{ pro } \lambda \to 0^+$$

$$\Rightarrow \langle x-y, z-y\rangle \leq 0. \quad \square$$

(2) $\Rightarrow$ (1):
Ať $z \in C$.
Pak

$$0 \geq \langle x-y, z-y\rangle = \langle x-y, (z-x)+(x-y)\rangle = \langle x-y, z-y\rangle + ||x-y||^2$$

$$\langle x-y, z-y\rangle + ||x-y||^2 \geq ||x-y||^2 - \underbrace{|\langle x-y, z-y\rangle|}_{\text{odhad shora}} \geq \star$$

$$\star = ||x-y||^2 - ||x-y|| \cdot ||z-x||.$$

Je-li $x \neq y$, pak vydělíme: $||z-x|| \geq ||x-y||$.
Je-li $x = y$, pak $y \in C : x \in C \ldots$ triviální. $\square$

2.3 Koule?

2.4 Věta o ortogonálním rozkladu

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

(a) $P_L : \mathbb{R}^n \to \mathbb{R}^n$ je lineární zobrazení.

(b) Pro každé $x \in \mathbb{R}^n$ je $P_{L^\perp}(x) = x - P_L(x)$.

(c) 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.

(a)
Cíl: Dokázat vlastnosti lineárního zobrazení, tedy

1. $P_L(\alpha x) = \alpha \cdot P_L(x), \forall \alpha \in \mathbb{R}, x \in \mathbb{R}^n$.
2. $P_L(x+y) = P_L(x) + P_L(y), \forall x, y \in \mathbb{R}^n$.

1. : Ať $z \in L$. Pak

$$\langle \alpha x - \alpha P_L(x), z - \alpha P_L(x)\rangle = \alpha\langle x - P_L(x), z - \alpha P_L(x)\rangle$$

$$\overset{\alpha\neq 0}{=} \underbrace{\alpha^2}_{>0}\langle x - P_L(x), \underbrace{\frac{1}{\alpha}\cdot z}_{\in L} - P_L(x)\rangle$$

Tedy $P_L(\alpha x) = \alpha P_L(x), \forall \alpha \neq 0$. Pro $\alpha = 0$ zřejmě plyne z lineárnosti zobrazení.

2. : Ať $z \in L$.

$$\langle \underbrace{x + y - (P_L(x) + P_L(y))}_{(x-P_L(x))+(y-P_L(y))}, z - (P_L(x) + P_L(y))\rangle$$

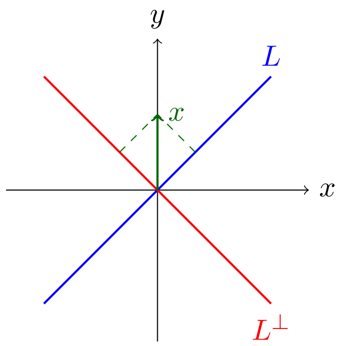
$$\underbrace{\langle x - P_L(x), \underbrace{(z - P_L(y))}_{\in L} - P_L(x)\rangle}_{\leq 0} + \underbrace{\langle y - P_L(y), \underbrace{(z - P_L(x))}_{\in L} - P_L(y)\rangle}_{\leq 0} \leq 0.$$

Z variační nerovnosti tedy plyne, že P_L je nutně lineární. $\square$

(b) Pro každé $x \in \mathbb{R}^n$ je $P_{L^\perp}(x) = x - P_L(x)$.

L . . . lineární podprostor $\mathbb{R}^n$,
$L^\perp = \{x \in \mathbb{R}^n \mid \langle x, y\rangle = 0, \forall y \in L\}$.

Důkaz.
Cíl: $P_{L^\perp}(x) = x - P_L(x)$.
Ať $x \in \mathbb{R}^n, z \in L^\perp$. Pak



$$\langle x - (x - P_L(x)), z - (x - P_L(x))\rangle = \langle \underbrace{P_L(x)}_{\in L}, z - (x - P_L(x))\rangle$$

$$= \underbrace{\langle P_L(x), z\rangle}_{0} - \langle P_L(x), x - P_L(x)\rangle = \langle x - P_L(x), 0 - P_L(x)\rangle \leq 0. \quad \square$$

(c) 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 $z = P_{L^\perp}(x)$.

Ať $x \in \mathbb{R}^n$.

Důkaz existence.
Pak $x = \underbrace{P_L(x)}_{\in L} + \underbrace{(x - P_L(x))}_{\in L^\perp}$. $\square$

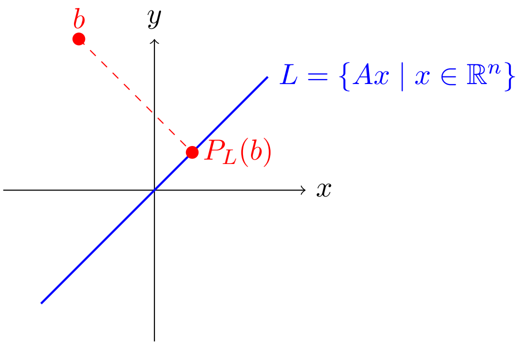
Důkaz jednoznačnosti.
Ať $a \in L, b \in L^\perp$ takové, že $x = a + b$.
Cíl: $a = P_L(x)$

Ať $z \in L$.

$$\langle x - a, z - a\rangle = \langle b, \underbrace{z - a}_{\in L}\rangle = 0 \leq 0 \implies a = P_L(x) \implies x - P_L(x) = b \stackrel{(2)}{\implies} P_{L^\perp}(x) = b. \quad \square$$

3 Třetí týden

3.1 Metoda nejmenších čtverců



Pokud $b \in L$, řešíme úlohu $Ax = b$.
Pokud $b \notin L$, řešíme $Ax = P_L(b)$.

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

Důkaz.

Chceme ukázat, že $\hat{x} \in \operatorname*{argmin}_{x\in\mathbb{R}^n} \|Ax - b^2\| \iff A^T A\hat{x} = A^T b$.

„$\Rightarrow$“: Ať $A\hat{x} = P_L(b) \overset{(2)}{=} b - P_{L^\perp}(b) \quad / \cdot A^T$

$$A^T A\hat{x} = A^T b - \underbrace{A^T P_{L^\perp}(b)}_{\overset{?}{=}0}$$

$$\to \|A^T P_{L^\perp}(b)\|^2 = \langle A^T P_{L^\perp}(b), A^T P_{L^\perp}(b)\rangle = \langle \underbrace{P_{L^\perp}(b)}_{\in L^\perp}, \underbrace{(A^T)^T(A^T P_{L^\perp}(b))}_{\in L}\rangle = 0. \quad \square$$

„$\Leftarrow$“: Ať $A^T A\hat{x} = A^T b$.
Ať $x \in \mathbb{R}^n$.

$$0 = \langle \underbrace{x, A^T A\hat{x} - A^T b}_{A^T(A\hat{x}-b)}\rangle = \langle \underbrace{(A^T)^T x}_{L}, A\hat{x} - b\rangle \implies A\hat{x} - b \in L^\perp$$

$$\to b = \underbrace{A\hat{x}}_{\in L} + \underbrace{(b - A\hat{x})}_{L^\perp} \overset{(c)}{\implies} A\hat{x} = P_L(b). \quad \square$$

3.2 Příklad výpočtu metody nejmenších čtverců

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

$A^T A\hat{x} = A^T b$

$$A^T A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \to \det = 3 \implies \text{existuje inverze.}$$

$$(A^T A)^{-1} = \tfrac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \implies \hat{x} = (A^T A)^{-1} A^T b = \tfrac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$= \tfrac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 2 \\ 2 \end{bmatrix} = \tfrac{1}{3}\begin{bmatrix} 2 \\ 2 \end{bmatrix}.$$

3.3 Příklad výpočtu metody nejmenších čtverců

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 proložme těmito body přímku o rovnici $y = kx + q$, kde $k, q \in \mathbb{R}$.

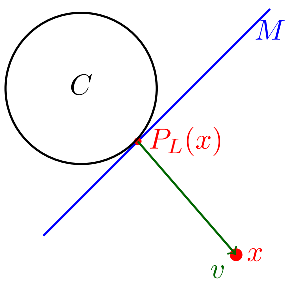
$$\left.\begin{aligned} 0k + q &= -\tfrac{1}{2} \\ 1k + q &= \tfrac{1}{3} \\ 2k + q &= \tfrac{2}{3} \end{aligned}\right\} A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}, b = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 3 \end{bmatrix}$$

$$(A^T A)^{-1} = \tfrac{1}{6} \begin{bmatrix} 3 & -3 \\ -3 & 5 \end{bmatrix}$$

$$\hat{x} = \frac{1}{6} \begin{bmatrix} 3 & -3 \\ -3 & 5 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{3} \\ \frac{2}{3} \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 3 & -3 \\ -3 & 5 \end{bmatrix} \begin{bmatrix} \frac{5}{3} \\ \frac{1}{2} \end{bmatrix} = \frac{1}{6} \begin{bmatrix} \frac{7}{2} \\ -\frac{5}{2} \end{bmatrix} = \frac{1}{12} \begin{bmatrix} 7 \\ -5 \end{bmatrix}.$$

3.4 Věta o oddělitelnosti bodu a konvexní množiny



$C \in \mathbb{R}^n$ je neprázdná uzavřená konvexní množina.
$x \in \mathbb{R}^n \setminus C \implies$ existuje $v \in \mathbb{R}^n \setminus \{0\}$ a $\alpha \in \mathbb{R}$ tak, že

$$\langle y, v \rangle \leq \alpha < \langle x, v \rangle, \quad \forall y \in C.$$

Důkaz.
$v = x - P_L(x) \neq 0$

$$\langle v, y \rangle = \langle v, P_L(x) \rangle \leq 0, \quad \forall y \in C.$$
$$\langle y, v \rangle \leq \langle v, P_L(x) \rangle, \quad \forall y \in C.$$

Položme $\alpha = \langle v, P_L(x) \rangle$.

$$\langle y, v \rangle \leq \alpha, \quad \forall y \in C.$$

$$\langle x, v \rangle - \overbrace{\underbrace{\langle v, P_L(x) \rangle}_{\langle P_L(x), v \rangle}}^{\alpha} = \langle \underbrace{x - P_L(x)}_{v}, v \rangle = \|v\|^2 > 0. \implies \alpha < \langle x, v \rangle. \quad \square$$

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

Důkaz sporem.

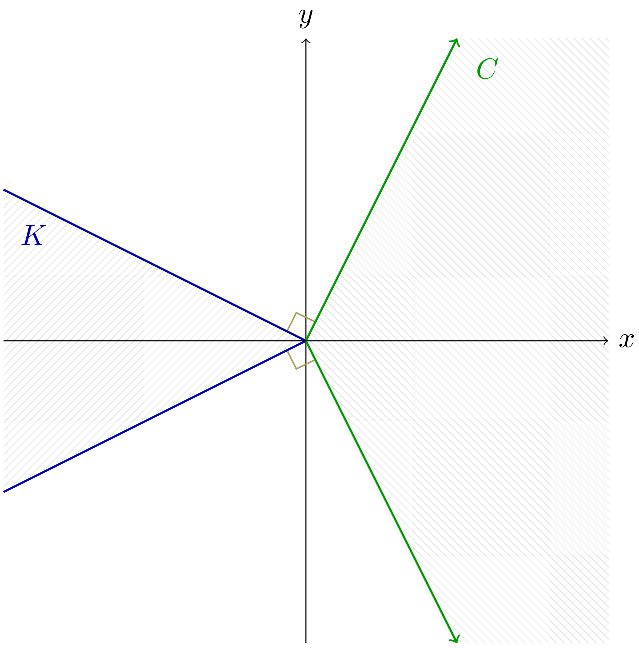
Ať neplatí: tj. existuje $C \in \mathbb{R}^n$ uzavřená konvexní množina tak, že není průnikem P všech poloprostorů obsahujících C.

Pak $x \in P$ tak, že $x \notin C$. Z věty o oddělitelnosti bodu a konvexní množiny existuje poloprostor M takový, že $C \subseteq M$ a $x \neq M$. Ale to je ve sporu s tím, že $x \in P$. $\square$

3.5 Příklad na použití věty o oddělitelnosti nadrovinou

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

$$\begin{aligned} C &= \left\{ Ax \middle| x \in \mathbb{R}^2_+ \right\} = \left\{ \alpha \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ -1 \end{bmatrix} \mid \alpha, \beta \geq 0 \right\} \\ K &= \left\{ y \in \mathbb{R}^2 \mid A^T y \leq 0 \right\} \\ &= \left\{ y \in \mathbb{R}^2 \middle| \left\langle \begin{bmatrix} 1 \\ 2 \end{bmatrix}, y \right\rangle \leq 0, \left\langle \begin{bmatrix} 1 \\ -1 \end{bmatrix}, y \right\rangle \leq 0 \right\}. \end{aligned}$$



Vždy nastane jeden z případů:

(a) $b \in C$

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

3.6 Lemma neprázdné uzavřené konvexní

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

Důkaz.

- neprázdná - vždy obsahuje alespoň 0,
- konvexní - lineární zobrazení (matice) zachovává konvexitu,
- uzavřenost dokazovat nebudeme.

3.7 Farkasovo lemma

Výslovnost [farkášovo].

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

(a) Existuje $x \in \mathbb{R}^n$ tak, že $Ax = b$ a $x \geq 0$.

(b) Existuje $y \in \mathbb{R}^m$ tak, že $A^T y \leq 0$ a $\langle y, b\rangle > 0$.

Důkaz.

„$(a) \implies \neg(b)$“:
Ať $x \in \mathbb{R}^n_+$ a $y \in \mathbb{R}^m$ tak, že $Ax = b$ a $A^T y \leq 0$.

$$\langle y, b\rangle \overset{b=Ax}{=} \langle y, Ax\rangle = \langle \underbrace{A^T y}_{\leq 0}, \underbrace{x}_{\geq 0}\rangle \leq 0. \quad \square$$

„$\neg(a) \implies (b)$“:
Ať $C = \left\{Ax \mid x \in \mathbb{R}^n_+\right\} \implies b \notin C$, $C \ldots$ uzavřená neprázdná konvexní množina.

$\overset{\text{oddělitelnost}}{\Longrightarrow}$ existuje $y \in \mathbb{R}^m \setminus \{0\}, \alpha \in \mathbb{R}$ tak, že: $\langle Ax, y\rangle \leq \alpha < \langle b, y\rangle, \quad \forall x \in \mathbb{R}^n_+$.

Začněme s $\alpha < \langle b, y\rangle$. Chceme, aby $\langle b, y\rangle$ byl kladný. Pak nám y bude svírat ostrý úhel s b.

Protože v $0 \in C$, je $0 \leq \alpha < \langle b, y\rangle$ (za Ax dosadíme 0, takže budeme mít $\langle 0, y\rangle$).

Teď musíme dokázat, že y skutečně řeší zadanou soustavu nerovnic.

Víme tedy, že:

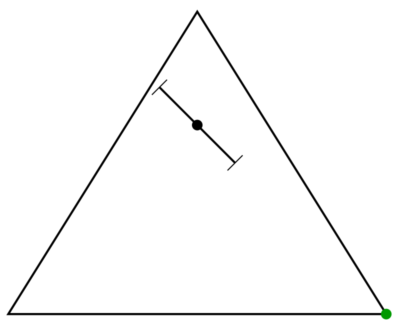
$$\begin{aligned} \langle Ax, y\rangle &\leq \alpha, \quad \forall x \in \mathbb{R}^n_+ \\ \langle x, A^T y\rangle &\leq \alpha, \quad \forall x \in \mathbb{R}^n_+ \end{aligned}$$

Odtuď $\langle x, A^T y\rangle \leq 0, \quad \forall x \in \mathbb{R}^n_+$, neboť:

Ať $\tilde{x} \in \mathbb{R}^n_+$ je takový, že $\langle \tilde{x}, A^T y\rangle > 0$.
Pak $\langle \underbrace{\lambda\tilde{x}}_{\lambda>0,\text{ tedy } \lambda\tilde{x}\in\mathbb{R}^n_+}, A^T y\rangle = \lambda \underbrace{\langle \tilde{x}, A^T y\rangle}_{>0} \to +\infty$, pro $\lambda \to +\infty$. Což je spor s $\langle x, A^T y\rangle \leq \alpha, \forall x \in \mathbb{R}^n_+$.

Ať $e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, e_n = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix}$. Pak $(A^T y)_i \leq 0, \quad \forall i \in \{1, \ldots, n\}$, neboť $(A^T y)_i = \langle e_i, A^T y\rangle$. $\square$

3.8 Krajní body konvexní množiny



Mějme konvexní množinu. Když sestrojíme libovolnou nedegenerativní (tzn. netriviální = není to pouze bod) úsečku, vždy nalezneme bod, který bude ležet přesně uprostřed této úsečky.

Co když ale vezmeme například zelený bod vyznačený na nákresu? V takovém případě nejsme schopni sestroji nedegenerativní úsečku, na jejímž středu by ležel tento bod.

Definujme: Krajní bod $x \in C$ konvexní množiny $C \subseteq \mathbb{R}^n$ je takový bod, pro který neexistují dva <u>různé</u> body y, z tak, že

$$x = \frac{1}{2}y + \frac{1}{2}z.$$

$\mathrm{ext}(C) \ldots$ množina všech krajních (extremálních) bodů

3.9 Kreinova-Milmanova věta

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

Kompaktnost je důležitá.

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

4 Čtvrtý týden

4.1 Konvexní funkce

Nechť $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ a $C \subseteq D$ je neprázdná konvexní množina.
Řekněme, že f je

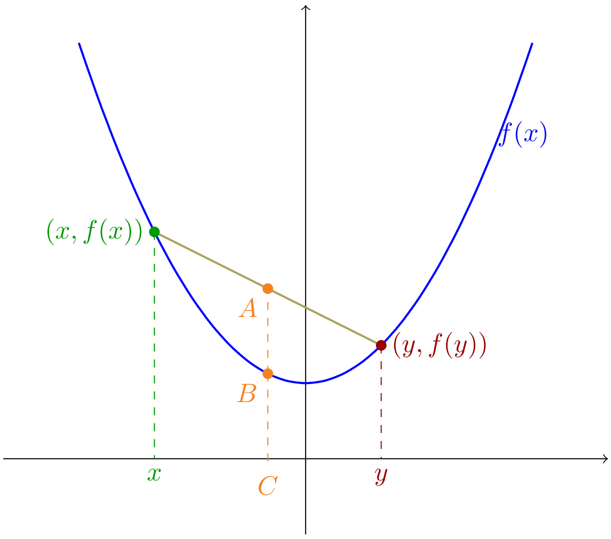
(a) konvexní na C, jestliže pro každé $x, y \in C$ a každé $\lambda \in [0, 1]$ je

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

(b) ryze konvexní na C, jestliže pro každé dva různé body $x, y \in C$ a $\lambda \in (0, 1)$ je

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

(c) konkávní (resp. ryze konkávní) na C, jestliže $(-f)$ je konvexní (resp. ryze konvexní) na C.



$A = (\lambda x + (1-\lambda)y, \lambda f(x) + (1-\lambda)f(y))$
$B = (\lambda x + (1-\lambda)y, f(\lambda x + (1-\lambda)y))$
$C = \lambda x + (1-\lambda)y$

Pozorování: úsečka vždy leží nad funkcí.

4.2 Příklad konvexní funkce

Je afinní zobrazení $f : \mathbb{R}^n \to \mathbb{R}^n$ (tj. $f(x) = \langle x, a\rangle + b, b \in \mathbb{R}$) konvexní?

Důkaz.
Ať $x, y \in \mathbb{R}^n, \lambda \in [0, 1]$.

$$\begin{aligned} f(\lambda x + (1-\lambda)y) &= \langle \lambda x + (1-\lambda)y, a\rangle + b \\ &= \lambda\langle x, a\rangle + (1-\lambda)\langle y, a\rangle + \lambda b + (1-\lambda)b \\ &= \lambda f(x) + (1-\lambda)f(y) \implies f \text{ je konvexní i konkávní.} \quad \square \end{aligned}$$

4.3 Příklad konvexní funkce

Je funkce $f(x) = \|x\|$ konvexní?

Důkaz.
Ať Ať $x, y \in \mathbb{R}^n, \lambda \in [0, 1]$.

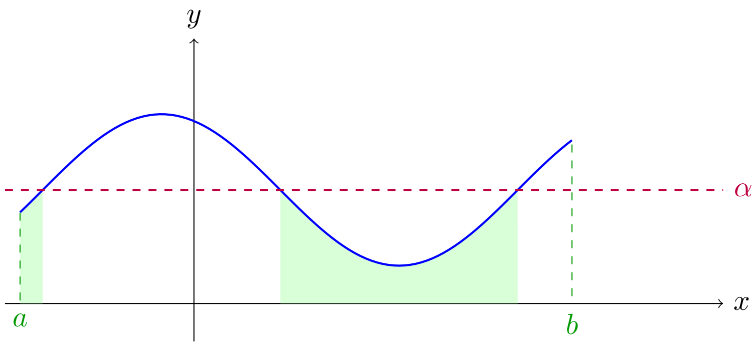
$$\begin{aligned} f(\lambda x + (1-\lambda)y) = \|\lambda x + (1-\lambda)y\| &\overset{\text{odhad}}{\leq} \|\lambda x\| + \|(1-\lambda)y\| = \lambda\|x\| + (1-\lambda)\|y\| \\ &= \lambda f(x) + (1-\lambda)f(y) \implies f \text{ je konvexní.} \quad \square \end{aligned}$$

4.4 Dolní úrovňová množina

Dolní úrovňování množina funkce $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ hladiny $\alpha \in \mathbb{R}$ je množina

$$\operatorname{lev}_{\leq}(f; \alpha) := \{x \in D \mid f(x) \leq \alpha\} .$$

Je-li f konvexní na $C \subseteq \mathbb{R}^n$, pak $\operatorname{lev}_{\leq}(f\big|_C; \alpha)$ je konvexní pro $\forall \alpha \in \mathbb{R}$.



Důkaz.

Ať $x, y \in \operatorname{lev}_{\leq}(f\big|_C; \alpha), \lambda \in [0,1]$.

Cíl: $\lambda x + (1-y)\lambda \overset{?}{\in} \operatorname{lev}_{\leq}(f\big|_C; y)$.

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \leq \lambda\alpha + (1-\lambda)\alpha = \alpha. \quad \square$$

Poznámka.

Opačná implikace neplatí. Tedy pomocí dolní úrovňové množiny **nelze** určit, jestli původní funkce je konvexní.

Například $f = x^3$ není konvexní funkce na intervalu $x = [-2, 2]$, ale když zvolíme $\alpha = 8$, tak dolní úrovňová množina bude konvexní.

4.5 Použití dolní úrovňové množiny

Je množina $M = \left\{x \in \mathbb{R}^2 \mid \|x\| \leq 1, \left\langle x, \begin{pmatrix}2\\1\end{pmatrix}\right\rangle \leq 1\right\}$ konvexní?

Důkaz.

Rozdělme si množinu M na dvě podmnožiny M_1 a M_2, kde:

$M_1 = \left\{x \in \mathbb{R}^2 \mid \|x\| \leq 1\right\} = \operatorname{lev}_{\leq}(\|x\|, 1) \to$ konvexní, protože norma je konvexní funkce.

$M_2 = \left\{x \in \mathbb{R}^2 \mid \left\langle x, \begin{pmatrix}2\\1\end{pmatrix}\right\rangle \leq 1\right\} = \operatorname{lev}_{\leq}\left(\left\langle x, \begin{pmatrix}2\\1\end{pmatrix}\right\rangle, 1\right) \to$ konvexní, protože skalární součin je konvexní.

To nám ale dává průnik dvou konvexních množin, tedy $M = M_1 \cap M_2$ je také konvexní. $\square$

4.6 Součet a součin zachovávají konvexitu

Mějme funkce f, g, které jsou konvexní na C, $\alpha \geq 0$. Pak:

(a) $f + g$ je konvexní na C

(b) αf je konvexní na C

Důkaz.

(a) Ať $\lambda \in [0,1], x, y \in C$.

$$(f+g)(\lambda x + (1-\lambda)y) = \underbrace{f(\lambda x + (1-\lambda)y)}_{\le \lambda f(x) + (1-\lambda) f(y)} + \underbrace{g(\lambda x + (1-\lambda)y)}_{\le \lambda g(x) + (1-\lambda) g(y)}$$

$$\le \lambda f(x) + (1-\lambda) f(y) + \lambda g(x) + (1-\lambda) g(y) = \lambda (f+g)(x) + (1-\lambda)(f+g)(y). \quad \square$$

(b) Ať $\lambda \in [0,1], x, y \in C, \alpha \ge 0$.

4.7 Příklad ověření konvexity

Je funkce $f(x) = e^x - 3\ln x + 2x$ konvexní?

Rozeberme si jednotlivé části funkce.

- e^x … exponenciála je z grafu očividně konvexní.
- $-3\ln x$ … logaritmus je konkávní, ale díky „$-$“ se celý výraz stane konvexní. Násobení konstatou konvexitu neovlivní, viz důkaz (b).
- $2x$ … lineární funkce je konvexní.

Protože všechny komponenty funkce f jsou konvexní, pak je i funkce f nutně konvexní.

4.8 Skládání zachovává konvexitu

Skládání konvexních funkcí není obecně konvexní funkce. Například: $f(x) = x^2$ a $g(x) = x^2 - 1$ jsou konvexní, ale

$$(f \circ g)(x) = (f(g(x))) = (x^2 - 1)^2 \text{ z grafu očividně není konvexní.}$$

1. Mějme tedy tvrzení.

Nechť f je konvexní na $K \subseteq \mathbb{R}^m$, $C \subseteq \mathbb{R}^n$ je neprázdná konvexní a $g : \mathbb{R}^n \to \mathbb{R}^m$ je afinní. Jestliže $g(C) \subseteq K$ (tedy g „obtiskne“ množinu C do K), pak $f \circ g$ je konvexní na C.

Důkaz.
Ať $x, y \in C$, $\lambda \in [0,1]$.
Pak

$$f(g(\lambda x + (1-\lambda)y)) \overset{g \text{ je afinní}}{=} f(\lambda \overbrace{g(x)}^{\in K} + (1-\lambda) \overbrace{g(y)}^{\in K}) \overset{f \text{ je konvexní}}{\le} \lambda f((g(x))) + (1-\lambda) f(g(y))$$

A to přesně dle definice konvexní funkce dává, že $f \circ g$ je konvexní funkce. $\square$

2. Mějme ještě druhé tvrzení.

Jestliže f je konvexní a **neklesající** na intervalu I, g je konvexní na $C \subseteq \mathbb{R}^n$ a $g(C) \subseteq I$, pak $f \circ g$ je konvexní na C.

Důkaz.
Ať $x, y \in C$, $\lambda \in [0,1]$.
Pak

$$f(\underbrace{g(\lambda x + (1-\lambda)y)}_{\substack{\le \lambda g(x) + (1-\lambda) g(y) \\ \text{odhad, díky konvexitě } g}}) \underset{g \text{ je konvexní}}{\overset{f \text{ je neklesající}}{\le}} f(\lambda g(x) + (1-\lambda) g(y)) \overset{f \text{ je konvexní}}{\le} \lambda f(g(x)) + (1-\lambda) f(g(y))$$

A to přesně dle definice konvexní funkce dává, že $f \circ g$ je konvexní funkce. $\square$

4.9 Věta o extrémech konvexních funkcí

Nechť f je konvexní na $C \subseteq \mathbb{R}^n$. Potom platí:

(a) Každý bod lokálního minima f na C je bodem minima f na C.

(b) Množina $\operatorname{argmin}_{x\in C} f(x)$ je konvexní. Je-li navíc f ryze konvexní na C, pak existuje nejvýše jeden bod minima funkce f na C.

Důkaz (a).
Sporem. Ať $\hat{x} \in C$ je bod lokálního minima f na C a ať existuje $\hat{y} \in C$ tak, že $f(\hat{y}) < f(\hat{x})$. $\lambda \in [0,1)$.
Pak

$$f(\lambda\hat{x} + (1-\lambda)\hat{y}) \overset{f \text{ je konvexní}}{\leq} \lambda f(\hat{x}) + (1-\lambda)\overbrace{f(\hat{y})}^{<f(\hat{x})} \underset{\text{odhad}}{<} \lambda f(\hat{x}) + (1-\lambda)f(\hat{x}) = f(\hat{x})$$

Což je ale spor s naším předpokladem, protože kdykoliv si vezmu bod na úsečce mezi $\hat{x}$ a $\hat{y}$, tak je v něm hodnota ostře menší než funkční hodnota v bodě $f(\hat{x})$. $\square$

Důkaz (b).
Ať $\hat{x}, \hat{y} \in \operatorname{argmin}_{x\in C} f(x)$, $\lambda \in [0,1]$.
Pak

$$f(\lambda\hat{x} + (1-\lambda)\hat{y}) \overset{f \text{ je konvexní}}{\leq} \lambda f(\hat{x}) + (1-\lambda)\overbrace{f(\hat{y})}^{=f(\hat{x})} = f(\hat{x})$$

$\implies \lambda\hat{x} + (1-\lambda)\hat{y} \in \operatorname{argmin}_{x\in C} f(x)$. $\square$

Ať f je navíc ryze konvexní na C.
Cíl: $\operatorname{argmin}_{x\in C} f(x)$ má nejvýše jeden prvek.

Důkaz.
Sporem. Ať $\hat{x}, \hat{y} \in \operatorname{argmin}_{x\in C} f(x)$, $\hat{x} \neq \hat{y}$. $\lambda \in (0,1)$.
Pak

$$f(\lambda\hat{x} + (1-\lambda)\hat{y}) \overset{f \text{ je ryze konv.}}{<} \lambda f(\hat{x}) + (1-\lambda)\underbrace{f(\hat{y})}_{=f(\hat{x})} = f(\hat{x})$$

Což je ale spor, protože mám nějakou funkční hodnotu bodu úsečky mezi $\hat{x}$ a $\hat{y}$ ostře menší jak funkční hodnotu bodu $\hat{x}$. To ale nemůže nastat, protože jako body minima funkce f na C musí mít stejnou hodnotu. Body $\hat{x}$ a $\hat{y}$ musí tedy nutně být stejné body. $\square$

4.10 Věta o konvexitě a první derivaci

Nechť $\Omega \subseteq \mathbb{R}^n$ je otevřená, $C \subseteq \Omega$ neprázdná konvexní a $f \in C^{'}(\Omega)$. Potom platí:

(a) f je konvexní na C právě tehdy, když pro každé $x, y \in C$ je

$$f(x) + \langle \nabla f(x), y - x \rangle \leq f(y).$$

(b) f je ryze konvexní na C právě tehdy, když pro každé dva různé body $x, y \in C$ je

$$f(x) + \langle \nabla f(x), y - x \rangle < f(y).$$

Důkaz (b) vynecháme.

Důkaz (a).

„⇒“: Ať $x, y \in C$, $\lambda \in (0, 1]$.

$$f(x + \lambda(y - x)) = f(\lambda y + (1 - \lambda)x) \overset{f \text{ je konvexní}}{\leq} \lambda f(x) + (1 - \lambda)f(x) = f(x) + \lambda[f(y) - f(x)]$$

$$\Rightarrow \underbrace{\frac{f(x + \lambda(y - x)) - f(x)}{\lambda}}_{\substack{=\langle \nabla f(x), y-x \rangle \text{ pro } \lambda \to 0_+ \\ \text{z definice směrové derivace}}} \leq f(y) - f(x). \quad \square$$

„⇐“: Ať $x, y \in C$, $\lambda \in [0, 1]$.

$z := \lambda x + (1 - \lambda)y \in C$
Z předpokladu:

$$f(z) + \langle \nabla f(z), x - z \rangle \leq f(x) \quad / \cdot \lambda \tag{1}$$
$$f(z) + \langle \nabla f(z), y - z \rangle \leq f(y) \quad / \cdot (-\lambda) \tag{2}$$

Pronásobením a sečtením dostaneme:

$$f(z) + \lambda \langle \nabla f(z), \underbrace{\lambda x + (1 - \lambda)y}_{z} - z \rangle \leq \lambda f(x) + (1 - \lambda)f(y)$$

$$\Rightarrow f(z) \leq \lambda f(x) + (1 - \lambda)f(y)$$

Což ale po dosazení za z je přesně ta nerovnost, která říká, že f je konvexní. $\square$

4.11 Věta o konvexitě a druhé derivaci

Nechť $\Omega \subseteq \mathbb{R}^n$ je otevřená, $C \subseteq \Omega$ neprázdná konvexní a $f \in C''(\Omega)$. Potom platí:

(a) Jestliže pro každé $x \in C$ je $\nabla^2 f(x)$ positivně semidefinitní matice, pak f je konvexní na C.

(b) Jestliže f je konvexní na C a C je otevřená, potom $\nabla^2 f(x)$ je positivně semidefinitní matice pro každé $x \in C$.

(c) Jestliže pro každé $x \in C$ je $\nabla^2 f(x)$ positivně definitní matice, pak f je ryze konvexní na C.

Důkaz (a).

Ať $x, y \in C$.

Taylorův polynom: existuje $\xi \in \{\lambda x + (1 - \lambda)y \mid \lambda \in (0, 1)\} \subseteq C$ tak, že

$$f(y) = f(x) + \langle \nabla f(x), y - x \rangle + \underbrace{\frac{1}{2} \left\langle \nabla^2 f(\xi)(y - x), y - x \right\rangle}_{\geq 0}$$

$$\Rightarrow f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$$

Což je přesné znění věty o konvexitě a první derivaci. Tedy f je nutně konvexní na C.

Důkaz (b).

Cíl: $\langle \nabla^2 f(x)y, y\rangle \geq 0$, $\forall y \in \mathbb{R}^n$

Ať $x \in C, y \in \mathbb{R}^n$.

Pak C otevřená $\Rightarrow$ existuje $\delta > 0$ tak, že $x + \alpha y \in C$ $\forall \alpha \in (0, \delta]$.

Taylorův polynom:

$$f(x+\alpha y) = f(x) + \alpha\langle \nabla f(x), y\rangle + \frac{1}{2}\alpha^2 \langle \nabla^2 f(x)y, y\rangle + \alpha^2 \|y\|^2 \omega(\alpha y),$$

kde w má nulovou limitu v 0.

Použijme fakt, že f je konvexní:

$$f(x+\alpha y) \geq f(x) + \langle \nabla f(x), \alpha y\rangle$$

Když tedy dosadíme:

$$f(x) + \alpha\langle \nabla f(x), y\rangle + \frac{1}{2}\alpha^2 \langle \nabla^2 f(x)y, y\rangle + \alpha^2 \|y\|^2 \omega(\alpha y) \geq f(x) + \langle \nabla f(x), \alpha y\rangle$$

Upravíme a podělíme výrazem $\frac{1}{2}\alpha^2$ $(\alpha > 0)$.

$$\langle \nabla^2 f(x)y, y\rangle + \underbrace{2\|y\|^2 \omega(\alpha y)}_{\to 0 \text{ pro } \alpha \to 0_+} \geq 0$$

V limitě $\alpha \to 0_+$ tedy máme $\langle \nabla^2 f(x)y, y\rangle \geq 0$, což je přesně to, co jsme chtěli. $\square$

Poznámka. Nutnost otevřenosti C je velmi důležitá!

Důkaz (c). Podobně jako (a).

4.12 Příklad ověření konvexnosti pomocí derivace

$f(x,y) = x^2 - y^2$ je konvexní na $\mathbb{R} \times \{0\}$. ($\to$ množina $\mathbb{R} \times \{0\}$ není otevřená, jedná se o přímku)

$\nabla^2 f(x,y) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$ je indefinitní, tedy funkce $f(x,y)$ není konvexní.

4.13 Příklad ověření konvexnosti pomocí derivace

$f(x,y) = x^2 + xy + y^2$ je ryze konvexní.

$\nabla^2 f(x,y) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \to 2 > 0$, $\det \nabla^2 f(x,y) = 4 - 1 > 0 \implies$ dle Sylvesterova kritéria je $\nabla^2 f(x,y)$ positivně definitní.

A podle bodu (c) věty o konvexitě a druhé derivaci můžeme říct, že funkce f je ryze konvexní.

5 Pátý týden

6 Šestý týden

7 Sedmý týden

8 Osmý týden

9 Devátý týden

10 Desátý týden

11 Jedenáctý týden

12 Dvanáctý týden

13 Třináctý týden

14 Čtrnáctý týden