

České vysoké učení technické v Praze  
Fakulta elektrotechnická

# Sbírka řešených příkladů

Optimalizace a teorie her

Jakub Adamec  
Praha, 2025

<https://github.com/kned11k/A8B010GT>



# Obsah

	Strana
<b>1 První týden</b>	<b>2</b>
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
<b>2 Druhý týden</b>	<b>12</b>
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
<b>3 Třetí týden</b>	<b>15</b>
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
3.10	Výpočet gradientu skalárního součinu . . . . .	19
3.11	Ověření konvexnosti množiny . . . . .	20
3.12	Práce s maticemi . . . . .	20
3.13	Proložení bodů pomocí MNČ . . . . .	21
3.14	Formulace úlohy MNČ . . . . .	22
<b>4</b>	<b>Čtvrtý týden</b>	<b>23</b>
4.1	Konvexní funkce . . . . .	23
4.2	Příklad konvexní funkce . . . . .	23
4.3	Příklad konvexní funkce . . . . .	23
4.4	Dolní úrovněová množina . . . . .	24
4.5	Použití dolní úrovněové množiny . . . . .	24
4.6	Součet a součin zachovávají konvexitu . . . . .	24
4.7	Příklad ověření konvexity . . . . .	25
4.8	Skládání zachovává konvexitu . . . . .	25
4.9	Věta o extrémech konvexních funkcí . . . . .	26
4.10	Věta o konvexitě a první derivaci . . . . .	26
4.11	Věta o konvexitě a druhé derivaci . . . . .	27
4.12	Příklad ověření konvexnosti pomocí derivace . . . . .	28
4.13	Příklad ověření konvexnosti pomocí derivace . . . . .	28
4.14	Příklad ověření konvexnosti funkce s parametrem . . . . .	29
4.15	Příklad ověření konvexity množiny . . . . .	29
<b>5</b>	<b>Pátý týden</b>	<b>31</b>
5.1	Kužel přípustných směrů . . . . .	31
5.2	Přípustné směry poklesu . . . . .	31
5.3	Kužel směrů poklesu . . . . .	32
5.4	Nutná geometrická podmínka lokálního extrému . . . . .	32
5.5	Silný směr poklesu - linearisace směru poklesu . . . . .	32
5.6	Tvrzení o souvislosti přípustných směrů poklesu a jejich linearisaci . . . . .	32
5.7	Fermatova věta - nutná podmínka optimality . . . . .	33
5.8	Věta o nutných a postačujících podmínkách pro konvexní úlohu . . . . .	33
5.9	Hledání bodu minima . . . . .	34
5.10	Věta o podmínkách optimality 2. řádu . . . . .	34
5.11	Příklad použití věty o podmínkách optimality 2. řádu . . . . .	34

5.12	Hledání bodu minima . . . . .	35
5.13	Omezení ve tvaru nerovnosti - aproximace $\mathcal{F}(M; \hat{x})$ . . . . .	35
5.14	Příklad výpočtu $\mathcal{G}$ a $\mathcal{F}$ . . . . .	36
5.15	Ukázka, že aproximací $\mathcal{F}$ lze zkazit prázdnotu průniku . . . . .	37
5.16	Věta o nutných KKT podmínkách . . . . .	38
5.17	Příklad použití KKT podmínek . . . . .	39
5.18	Příklad, že KKT podmínky vždy nenaleznou všechny body . . . . .	39
5.19	Věta o postačujících KKT podmínkách . . . . .	40
5.20	Afinní podmínka regularity . . . . .	40
5.21	Slaterova podmínka regularity . . . . .	40
5.22	Použití podmínek regularity k ověření KKT podmínek . . . . .	40
5.23	Určení nutných a postačujících podmínek optimality . . . . .	41
5.24	Určení KKT podmínek . . . . .	41
5.25	Určení KKT podmínek . . . . .	43
<b>6</b>	<b>Šestý týden</b>	<b>44</b>
6.1	Pomocný důkaz vlastnosti infima . . . . .	44
6.2	Dualita - motivační příklad . . . . .	44
<b>7</b>	<b>Sedmý týden</b>	<b>46</b>
<b>8</b>	<b>Osmý týden</b>	<b>47</b>
<b>9</b>	<b>Devátý týden</b>	<b>48</b>
<b>10</b>	<b>Desátý týden</b>	<b>49</b>
<b>11</b>	<b>Jedenáctý týden</b>	<b>50</b>
<b>12</b>	<b>Dvanáctý týden</b>	<b>51</b>
<b>13</b>	<b>Třináctý týden</b>	<b>52</b>
<b>14</b>	<b>Čtrnáctý týden</b>	<b>53</b>

# 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 oficiálním 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 verze textu bude na stránce <https://github.com/knedl1k/A8B010GT>.

**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  $\text{\LaTeX}$  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 zprehlednění sbírky řešení jednotlivých příkladů ihned pod zadáním.

# 1 První týden

## 1.1 Důkaz souvislosti minima a maxima

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

- (1)  $\hat{x} \in \operatorname{argmin}_{x \in M} f(x) \iff \hat{x} \in \operatorname{argmax}_{x \in M} (-f(x)),$   
(2) jestliže  $\hat{x} \in \operatorname{argmin}_{x \in M} f(x)$ , pak  $\min_{x \in M} f(x) = -\max_{x \in M} (-f(x)).$

Důkaz.

- (1)  $\hat{x} \in \operatorname{argmin}_{x \in M} f(x)$ , tj.  $f(\hat{x}) \leq f(x), \forall x \in M \iff -f(\hat{x}) \geq -f(x), \forall x \in M$ , tj.  $\hat{x} \in \operatorname{argmax}_{x \in M} (-f(x)). \quad \square$   
(2) Ať  $\hat{x} \in \operatorname{argmin}_{x \in M} f(x)$ , pak  $\min_{x \in M} f(x) = f(\hat{x}) = -(-f(\hat{x})) \stackrel{(1)}{=} -\max_{x \in M} (-f(x)). \quad \square$

## 1.2 Hledání přípustných množin

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

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

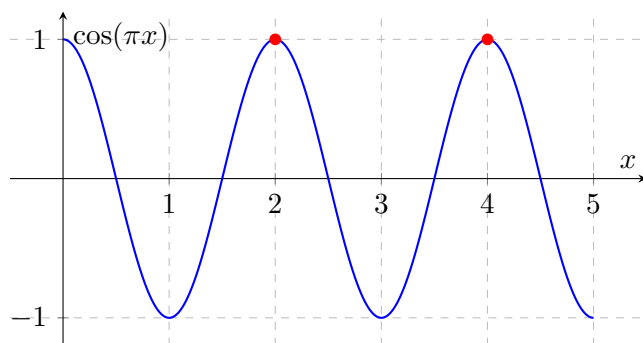
Úvahou pak lze uhodnout minimum - minimum leží v bodě  $x = 3$ .

## 1.3 Hledání přípustných množin

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

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

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



Očividně tedy  $M = \{2, 4\}$ .

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

## 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 investičního produktu (v tis. Kč) při investici  $y$  (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?

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

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

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

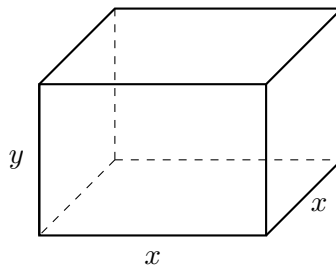
Zbavme se zlomků:

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

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

## 1.5 Minimalisační úloha

Ve firmě potřebují nalézt rozměry otevřené krabice (tj. krabice bez horní stěny) se čtvercovou podstavou o objemu  $10 \text{ dm}^3$  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.



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

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.

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

Zbavme se zlomků:

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

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

## 1.6 Optimalisační úloha s nadrovinami

V  $\mathbb{R}^n$  jsou dány množiny bodů  $A = \{a_1, \dots, a_k\}$  a  $B = \{b_1, \dots, b_l\}$ . 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  $\langle x, w \rangle + \lambda = -1$ .

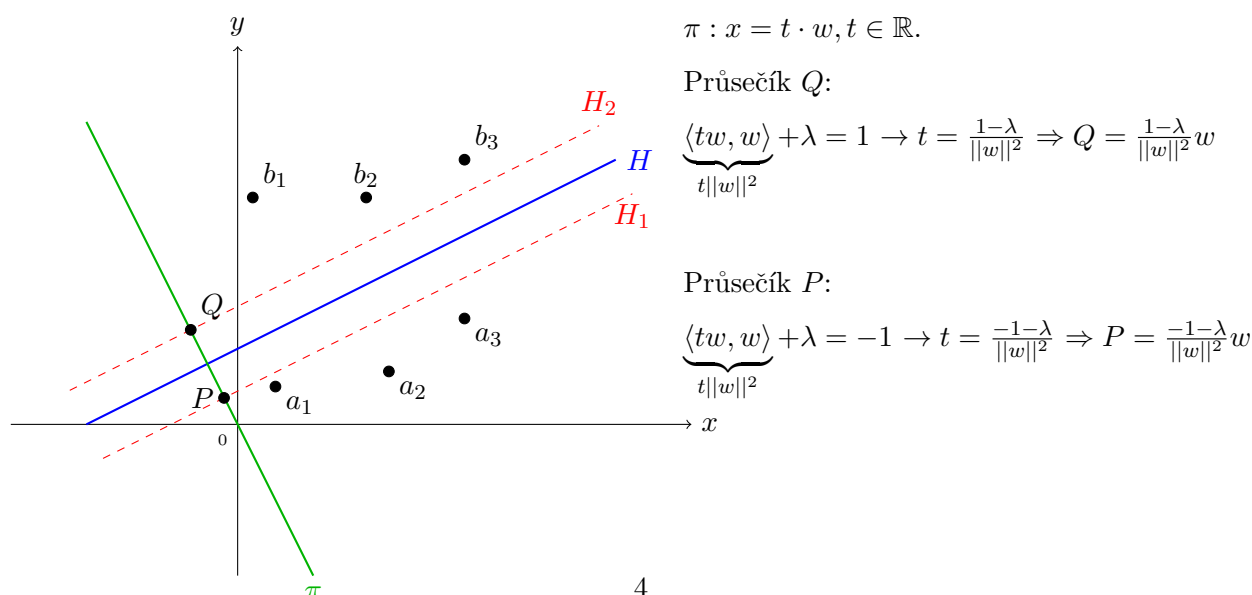
- (a) Ukažte, že vzdálenost mezi nadrovinami  $H_1$  a  $H_2$  je  $\frac{2}{\|w\|}$ . Dále ukažte, že  $\frac{1}{\|w\|}$  je vzdálenost  $H$  od  $H_1$  a také vzdálenost  $H$  od  $H_2$ .
- (b) Interpretujte optimalisační úlohu

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

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

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

(a)





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 prima, 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

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

se nazývá uzavřená úsečka s krajními body  $x$  a  $y$ .

### 1.8 Je nadrovina konvexní?

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

Důkaz.

Atť  $x, z \in H(y, \alpha)$ ,  $\lambda \in [0, 1]$ .

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

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

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

### 1.9 Je uzavřený poloprostor konvexní?

#### 1.10 Je uzavřená koule konvexní?

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

Důkaz.

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

Cíl:  $\|[\lambda x + (1 - \lambda)y] - a\| \leq r$ . 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**.

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

#### 1.11 Je okolí konvexní?

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

Důkaz.

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

Cíl:  $\|[\lambda x + (1 - \lambda)y] - a\| < r$ . Dle **definice**.

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

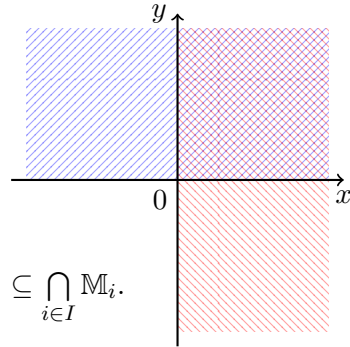
### 1.12 Je průnik množin konvexní?

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

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

Visuálně je průnik nekonvexní.



Důkaz.

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

□

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

### 1.14 Důkaz, že afinní zobrazení je konvexní

Tvrzení.

Nechť  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ . 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  $A \in \mathbb{M}_{m,n}(\mathbb{R})$ ,  $b \in \mathbb{R}^m$ .

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

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

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

„ $\Leftarrow$ “: Cíl: Ukázat, že  $f$  je **afinní**, tedy  $f(x) = Ax + b$ .

Zvolme  $\varphi(x) = f(x) - f(0)$ .

Pokud je  $f$  **afinní**, pak zobrazení  $\varphi$  by mělo být dáno jako  $Ax$ , tedy být lineární.

Cíl:  $\varphi$  je lineární zobrazení.

Musíme ověřit uzavřenost na násobení a sčítání z definice.

(1) At  $x \in \mathbb{R}^n$ ,  $\alpha \in \mathbb{R}$ .

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) At  $x, y \in \mathbb{R}^n$ .

Cíl:  $\varphi(x + y) = \varphi(x) + \varphi(y)$ .

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

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

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

Tvrzení.

Je-li  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  **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.  $\implies [x, y] \subseteq C \implies \underbrace{f([x, y])}_{\subseteq f(C)} = \underbrace{[f(x), f(y)]}_a \subseteq f(C). \quad \square$

### 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.

„ $\implies$ “: 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**.

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

„ $\impliedby$ “: Definujme **afinní** zobrazení  $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$  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.  $g : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$  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 symetrická.  
 $Q=Q^T$

Pak platí:

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

Analogicky pro negativně semidefinitní, respektive definitní.

Matice je indefinitní pokud nesplňuje ani jednu možnost.

Pro symetrické matice také platí, že  $Q$  je negativně (semi)definitní, jestliže  $(-Q)$  je pozitivně (semi)definitní.

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

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

Hlavní minory jsou  $Q_{\{1\}}$  a  $Q_{\{1,2\}}$ .

Menší minory:  $Q_I$ , kde  $I \subseteq \{1, \dots, n\}$  neprázdná. Aby matice byla pozitivně semidefinitní, tak  $\det Q_I \geq 0$ .

Tedy:  $Q_{\{2\}} = [4]$ .  $\det Q_{\{2\}} = 4 > 0$ .

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

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

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

Pozorování: Matice je lineárně závislá, tedy  $\det Q = 0$ .

$$Q_{\{1\}} = 4 > 0,$$

$$Q_{\{2\}} = 1 > 0,$$

$$Q_{\{3\}} = 0 = 0.$$

Tedy matice je jediné pozitivně semidefinitní, nebo indefinitní.

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

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

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

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

Využijme tedy **vlastnosti** symetrických matic a určíme 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 pozitivně semidefinitní, nebo indefinitní.

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

$\implies (-Q)$  je pozitivně 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  $\langle Ax, x \rangle = \langle Bx, x \rangle$ .

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

(a) Využijme zmíněné **vlastnosti**.

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

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

$$\text{Zvolme: } \left. \begin{array}{l} B = \frac{1}{2}(A + A^T) \\ C = \frac{1}{2}(A - A^T) \end{array} \right\} B + C = A.$$

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

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

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

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

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

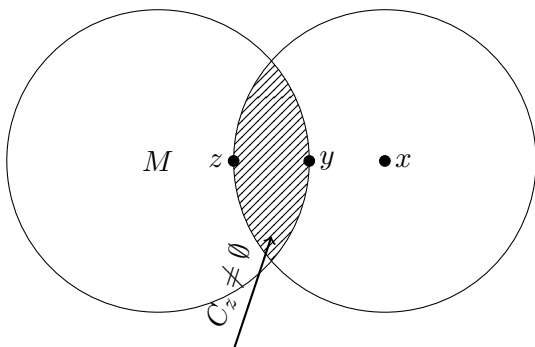
Je-li  $C \subseteq \mathbb{R}^n$  neprázdná uzavřená konvexní množina, pak pro každé  $x \in \mathbb{R}^n$  existuje právě jeden bod  $\hat{y} \in C$  tak, že  $\text{dist}(x; C) = \|x - \hat{y}\|$ .

Důkaz.

1. Existence

Cíl: Existuje bod minima

Úvaha:



$M$  je obecná konvexní množina.

$C \times R = \|x - z\|$ ,

$C_z = M \cap B(x, R) = M \cap \{a \in \mathbb{R}^n \mid \|z - a\| \leq R\}$ .

↑

uzavřená, omezená, neprázdná  
kompaktní

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

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

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

2. Jednoznačnost.

Cíl: Pokud  $a, b \in \mathbb{R}^n : \|x - a\| = \|x - b\| = \overbrace{\text{dist}(x, M)}^\delta$ , pak  $a = b$ .

Lemma, rovnoběžníkové pravidlo:  $u, v \in \mathbb{R}^n \Rightarrow \|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2)$ .

Důkaz lemma:

$$\begin{aligned} \|u + v\|^2 + \|u - v\|^2 &= \langle u + v, u + v \rangle + \langle u - v, u - v \rangle = \|u\|^2 + 2\langle u, v \rangle + \|v\|^2 + \|u\|^2 - 2\langle u, v \rangle + \|v\|^2 \\ &= 2(\|u\|^2 + \|v\|^2). \quad \square \end{aligned}$$

Důkaz jednoznačnosti:

At  $y = \frac{1}{2}a + \frac{1}{2}b$ .

Pak  $\delta^2 \leq \|x - y\|^2 = \|x - \frac{1}{2}a - \frac{1}{2}b\|^2 = \|\frac{1}{2}(x - a) + \frac{1}{2}(x - b)\|^2 = \frac{1}{4}\|\underbrace{(x - a)}_u + \underbrace{(x - b)}_v\|^2$

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

### 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

$$\begin{aligned} \|x - y\|^2 &\leq \|x - v_\lambda\|^2 = \|x - y - \lambda(z - y)\|^2 = \langle (x - y) - \lambda(z - y), (x - y) - \lambda(z - y) \rangle \\ \|x - y\|^2 &\leq \|x - y\|^2 - 2\lambda \langle x - y, z - y \rangle + \lambda^2 \|z - y\|^2 \\ \Rightarrow \langle x - y, z - y \rangle &\leq \frac{\lambda}{2} \|z - y\|^2 \rightarrow 0 \text{ pro } \lambda \rightarrow 0^+ \\ \Rightarrow \langle x - y, z - y \rangle &\leq 0. \quad \square \end{aligned}$$

(2)  $\Rightarrow$  (1):

Ať  $z \in C$ .

Pak

$$\begin{aligned} 0 &\geq \langle x - y, z - y \rangle = \langle x - y, (z - x) + (x - y) \rangle = \langle x - y, z - y \rangle + \|x - y\|^2 \\ \langle x - y, z - y \rangle + \|x - y\|^2 &\geq \|x - y\|^2 - \underbrace{|\langle x - y, z - y \rangle|}_{\text{odhad shora}} \geq \star \\ \star &= \|x - y\|^2 - \|x - y\| \cdot \|z - x\|. \end{aligned}$$

Je-li  $x \neq y$ , pak vydělíme:  $\|z - x\| \geq \|x - y\|$ .

Je-li  $x = y$ , pak  $y \in C : x \in C \dots$  triviální.  $\square$

## 2.3 Koule?

## 2.4 Věta o ortogonálním rozkladu

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

(a)  $P_L : \mathbb{R}^n \rightarrow \mathbb{R}^n$  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

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

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

2. : Ať  $z \in L$ .

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

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

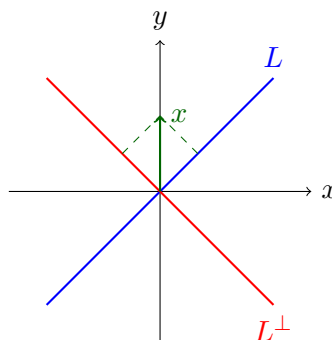
(b) Pro každé  $x \in \mathbb{R}^n$  je  $P_{L^\perp}(x) = x - P_L(x)$ .

$L$  ... 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



$$\begin{aligned} \langle x - (x - P_L(x)), z - (x - P_L(x)) \rangle &= \underbrace{\langle P_L(x), z - (x - P_L(x)) \rangle}_{\in L} \\ &= \underbrace{\langle P_L(x), z \rangle}_0 - \langle P_L(x), x - P_L(x) \rangle = \langle x - P_L(x), 0 - P_L(x) \rangle \leq 0. \quad \square \end{aligned}$$

(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  $x = a + b$ .

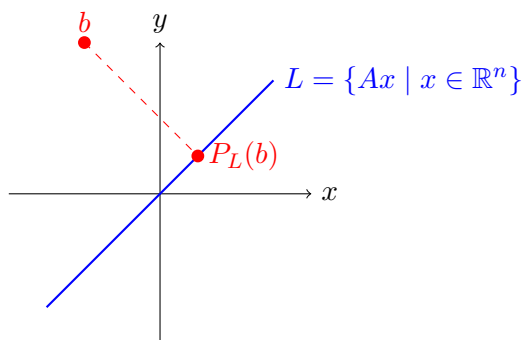
Cíl:  $a = P_L(x)$

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  $A \hat{x} = P_L(b) \stackrel{(2)}{=} b - P_{L^\perp}(b) \quad / \cdot A^T$

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

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

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

Ať  $x \in \mathbb{R}^n$ .

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

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

#### 3.2 Příklad výpočtu metody nejmenších čtverců

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

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

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

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

### 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 přímkou o rovnici  $y = kx + q$ , kde  $k, q \in \mathbb{R}$ .

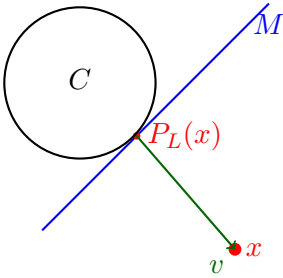
$$\left. \begin{aligned} 0k + q &= -\frac{1}{2} \\ 1k + q &= \frac{1}{3} \\ 2k + q &= \frac{2}{3} \end{aligned} \right\} A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}, b = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}$$

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

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

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

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

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

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

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

Důkaz sporem.

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

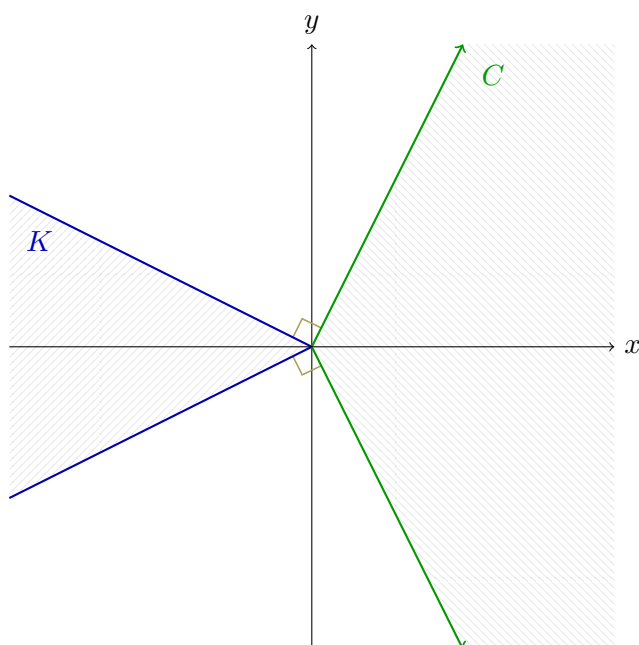
Pak  $x \in P$  tak, že  $x \notin C$ . Z věty o oddělitelnosti bodu a konvexní množiny existuje poloprostor  $M$  takový, že  $C \subseteq M$  a  $x \notin M$ . Ale to je ve sporu s tím, že  $x \in P$ .  $\square$

### 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

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

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



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  $\{Ax \mid x \in \mathbb{R}_+^n\}$  je neprázdna uzavřená konvexní množina.

Důkaz.

- neprázdna - 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)  $\implies \neg(b)$ “:

At'  $x \in \mathbb{R}_+^n$  a  $y \in \mathbb{R}^m$  tak, že  $Ax = b$  a  $A^T y \leq 0$ .

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

„ $\neg(a) \implies (b)$ “:

At'  $C = \{Ax \mid x \in \mathbb{R}_+^n\} \implies b \notin C, C \dots$  uzavřená neprázdná konvexní množina.

**oddělitelnost**  $\implies$  existuje  $y \in \mathbb{R}^m \setminus \{0\}, \alpha \in \mathbb{R}$  tak, že:  $\langle Ax, y \rangle \leq \alpha < \langle b, y \rangle, \quad \forall x \in \mathbb{R}_+^n$ .

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

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

Ted' musíme dokázat, že  $y$  skutečně řeší zadanou soustavu nerovnic.

Víme tedy, že:

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

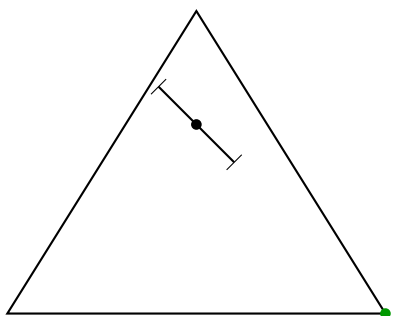
Odtud'  $\langle x, A^T y \rangle \leq 0, \quad \forall x \in \mathbb{R}_+^n$ , neboť:

At'  $\tilde{x} \in \mathbb{R}_+^n$  je takový, že  $\langle \tilde{x}, A^T y \rangle > 0$ .

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

$$\text{At' } e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, e_n = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix}. \text{ Pak } (A^T y)_i \leq 0, \quad \forall i \in \{1, \dots, n\}, \text{ neboť } (A^T y)_i = \langle e_i, A^T y \rangle. \quad \square$$

### 3.8 Krajiní 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 nákrese? V takovém případě nejsme schopni sestroji nedegenerativní úsečku, na jejímž středu by ležel tento bod.

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

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

$\text{ext}(C) \dots$  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  $C = \text{conv}(\text{ext}(C))$ .  
Důkaz vynecháme.

Kompaktnost je důležitá.

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

### 3.10 Výpočet gradientu skalárního součinu

Nalezněte  $\nabla f(x)$  a  $\nabla^2 f(x)$ , jestliže

(a)  $f(x) = \langle x, c \rangle$ , kde  $c \in \mathbb{R}^n$ ;

(b)  $f(x) = \langle Ax, x \rangle$ , kde  $A \in \mathbb{M}_n(\mathbb{R})$ . Určete také  $\nabla f(x)$  a  $\nabla^2 f(x)$  za dodatečného předpokladu, že  $A$  je symetrická matice.

(a)

$$\frac{\partial f}{\partial x_k} = \frac{\partial}{\partial x_k} \sum_{i=1}^n x_i c_i \stackrel{\text{limita}}{=} \sum_{i=1}^n c_i \frac{\partial x_i}{\partial x_k} = \sum_{i=1}^n c_i \delta_{ik} \stackrel{\text{rozvoj}}{=} c_k$$

$$\Rightarrow \nabla f(x) = \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix} = c; \Rightarrow \nabla^2 f(x) = 0, \text{ kde } \delta_{ik} = \begin{cases} 1, & \text{pokud } i = k, \\ 0, & \text{pokud } i \neq k. \end{cases}$$

(b)

$$\begin{aligned} \frac{\partial f}{\partial x_k} &= \frac{\partial}{\partial x_k} \sum_{i=1}^n \underbrace{\left[ \sum_{j=1}^n a_{ij} x_j \right]}_{(Ax)_i} x_i = \sum_{i,j=1}^n a_{ij} \underbrace{\left( \frac{\partial x_i}{\partial x_k} x_j + x_i \frac{\partial x_j}{\partial x_k} \right)}_{\text{derivace součinu}} = \sum_{i,j=1}^n a_{ij} (\delta_{ik} x_j + x_i \delta_{jk}) \\ &= \sum_{i,j=1}^n a_{ij} \delta_{ik} x_j + a_{ij} \delta_{jk} x_i = \underbrace{\sum_{j=1}^n a_{kj} x_j}_{(Ax)_k} + \underbrace{\sum_{i=1}^n a_{ik} x_i}_{(A^T x)_k} \end{aligned}$$

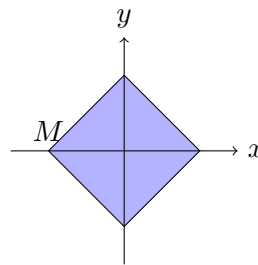
$$\Rightarrow \nabla f(x) = Ax + A^T x \text{ (Speciálně: } \nabla f(x) = 2Ax \text{ pro } A = A^T)$$

$$\frac{\partial^2 f}{\partial x_k \partial x_l} = \sum_{j=1}^n a_{kj} \delta_{jl} + \sum_{i=1}^n a_{ik} \delta_{il} = a_{kl} + a_{lk}$$

$$\Rightarrow \nabla^2 f(x) = A + A^T$$

### 3.11 Ověření konvexnosti množiny

Je množina  $M = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \mid |x| + |y| \leq 1 \right\}$  konvexní?



1. způsob - dle definice

$$\lambda \begin{bmatrix} x \\ y \end{bmatrix} + (1 - \lambda) \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \lambda x + (1 - \lambda)a \\ \lambda y + (1 - \lambda)b \end{bmatrix} \stackrel{?}{\in} M, \lambda \in [0, 1].$$

$$|\lambda x + (1 - \lambda)a| + |\lambda y + (1 - \lambda)b| \leq \underbrace{\lambda|x| + (1 - \lambda)|a| + \lambda|y| + (1 - \lambda)|b|}_{\lambda \underbrace{(|x| + |y|)}_{\leq 1} + (1 - \lambda) \underbrace{(|a| + |b|)}_{\leq 1}} \leq \lambda + 1 - \lambda = 1 \quad \square$$

$M$  je konvexní.

2. způsob - úvaha nad vlastnostmi

$|x|$  je konvexní,  $|y|$  je konvexní. Součet zachovává konvexitu, tedy i  $|x| + |y|$  je konvexní.

### 3.12 Práce s maticemi

Je dána matice  $A \in \mathbb{M}_{m,n}(\mathbb{R})$ . Ať  $L = \{Ax \mid x \in \mathbb{R}^n\}$ .

Ukažte, že  $A$  má lineárně nezávislé sloupce  $\iff A^T A$  je invertibilní.

Pomocný důkaz.

Ukažme, že:  $\ker(A) = \ker(A^T A)$

Chci:  $\ker(A) \subseteq \ker(A^T A)$

$$x \in \ker(A) \Rightarrow Ax = 0 \quad / \cdot A^T$$

$$A^T A = 0 \Rightarrow x \in \ker(A^T A) \quad \square$$

Chci:  $\ker(A^T A) \subseteq \ker(A)$

$$x \in \ker(A^T A) \Rightarrow A^T Ax = 0 \Rightarrow 0 = \langle A^T Ax, x \rangle$$

$$= \langle Ax, Ax \rangle$$

$$= \|Ax\|^2 \Rightarrow Ax = 0 \Rightarrow x \in \ker(A) \quad \square$$

Konec pomocného důkazu.

$A$  má lineárně nezávislé sloupce  $\iff \{0\} = \ker(A) = \ker(A^T A) \iff A^T A$  je invertibilní (protože  $A^T A$  je čtvercová a  $A^T A$  je prosté).



### 3.13 Proložení bodů pomocí MNČ

Jsou dány body  $a = \begin{bmatrix} -2 \\ -1 \end{bmatrix}, b = \begin{bmatrix} -1 \\ -2 \end{bmatrix}, c = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, d = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ . Metodou nejmenších čtverců proložte těmito body graf

(a) afinní funkce  $f(x) = \alpha x + \beta$ , kde  $\alpha, \beta \in \mathbb{R}$ ;

(b) funkce  $f(x) = \alpha x^2 + \beta x + \gamma$ , kde  $\alpha, \beta, \gamma \in \mathbb{R}$ .

(a)

$$\begin{aligned} -2\alpha + \beta &= -1 \\ -\alpha + \beta &= -2 \\ 0\alpha + \beta &= 0 \\ \alpha + \beta &= 2 \end{aligned} \iff A \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = b, \text{ kde } A = \begin{bmatrix} -2 & 1 \\ -1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}, b = \begin{bmatrix} -1 \\ -2 \\ 0 \\ 2 \end{bmatrix}.$$

$A^T A \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = A^T b$ .  $A$  má lineárně nezávislé sloupce  $\Rightarrow (A^T A)^{-1}$  existuje.

Pak:  $\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = (A^T A)^{-1} A^T b$ .

$$A^T A = \begin{bmatrix} -2 & -1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ -1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 6 & -2 \\ -2 & 4 \end{bmatrix} \Rightarrow (A^T A)^{-1} = \frac{1}{20} \begin{bmatrix} 4 & 2 \\ 2 & 6 \end{bmatrix}.$$

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \frac{1}{10} \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} -2 & -1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ -2 \\ 0 \\ 2 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 6 \\ -1 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} 11 \\ 3 \end{bmatrix} \Rightarrow \alpha = \frac{11}{10}; \beta = \frac{3}{10}.$$

(b)

$$\begin{aligned} 4\alpha - 2\beta + \gamma &= -1 \\ \alpha - \beta + \gamma &= -2 \\ 0\alpha + 0\beta + \gamma &= 0 \\ \alpha + \beta + \gamma &= 2 \end{aligned} \iff A \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = b, \text{ kde } A = \begin{bmatrix} 4 & -2 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}, b = \begin{bmatrix} -1 \\ -2 \\ 0 \\ 2 \end{bmatrix}.$$

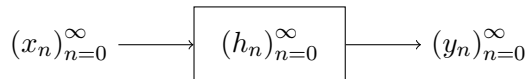
$A$  má lineárně nezávislé sloupce  $\Rightarrow A^T A$  je invertibilní.

$$A^T A = \begin{bmatrix} 4 & 1 & 0 & 1 \\ -2 & -1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & -2 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 18 & -8 & 6 \\ -8 & 6 & -2 \\ 6 & -2 & 4 \end{bmatrix} \Rightarrow (A^T A)^{-1} = \frac{1}{20} \begin{bmatrix} 5 & 5 & -5 \\ 5 & 9 & -3 \\ -5 & -3 & 11 \end{bmatrix}.$$

$$\begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \frac{1}{20} \begin{bmatrix} 5 & 5 & -5 \\ 5 & 9 & -3 \\ -5 & -3 & 11 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix} = \frac{1}{20} \begin{bmatrix} 25 \\ 35 \\ -15 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 5 \\ 7 \\ -3 \end{bmatrix} \Rightarrow \alpha = \frac{5}{4}; \beta = \frac{7}{4}; \gamma = \frac{-3}{4}.$$

### 3.14 Formulace úlohy MNČ

At' závislost výstupního signálu  $(y_n)_{n=0}^{\infty}$  systému na vstupním signálu  $(x_n)_{n=0}^{\infty}$  je dána konvolucí posloupností  $(x_n)_{n=0}^{\infty}$  s posloupností  $(h_n)_{n=0}^{\infty}$  ( $(h_n)_{n=0}^{\infty}$  popisuje odezvu systému na jednotkový impuls), tj.  $y_n = \sum_{i=0}^n h_i x_{n-i}$ . Předpokládejte dále, že  $h_n = 0$  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  $x_0, \dots, x_{20}$ . Formulujte úlohu nejmenších čtverců pro nalezení koeficientů  $h_0, h_1, h_2, h_3$ .



$$y_k = \sum_{l=0}^k h_l x_{k-l} = h_0 x_k + \dots + h_k x_0$$

předpokládejme:  $h_l = 0 \forall l \geq 4$ .

$$y_0 = h_0 x_0$$

$$y_1 = h_1 x_0 + h_0 x_1$$

$$y_2 = h_2 x_0 + h_1 x_1 + h_0 x_2$$

$$y_3 = h_3 x_0 + h_2 x_1 + h_1 x_2 + h_0 x_3$$

$$y_4 = h_3 x_1 + h_2 x_2 + h_1 x_3 + h_0 x_4$$

$$\vdots$$

$$y_{20} = h_3 x_{17} + h_2 x_{18} + h_1 x_{19} + h_0 x_{20}$$

Minimalisujme  $f(x) = \|Ax + b\|^2$ , kde

$$A = \begin{bmatrix} x_0 & 0 & 0 & 0 \\ x_1 & x_0 & 0 & 0 \\ x_2 & x_1 & x_0 & 0 \\ x_3 & x_2 & x_1 & x_0 \\ \vdots & \vdots & \vdots & \vdots \\ x_{20} & x_{19} & x_{18} & x_{17} \end{bmatrix}; \quad b = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{20} \end{bmatrix}.$$

## 4 Čtvrtý týden

### 4.1 Konvexní funkce

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

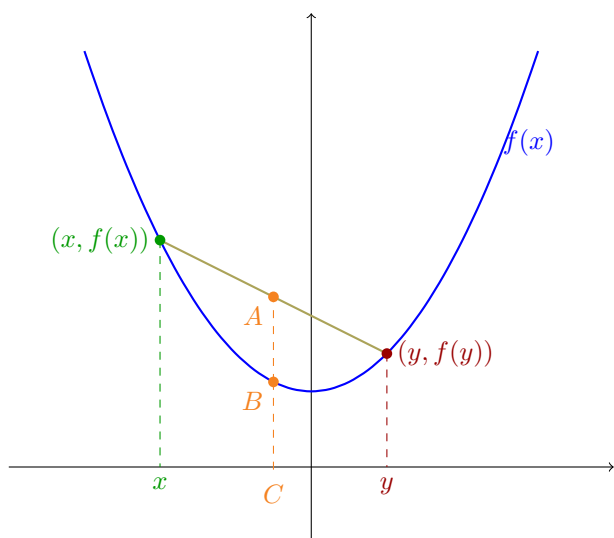
- (a) konvexní na  $C$ , jestliže pro každé  $x, y \in C$  a každé  $\lambda \in [0, 1]$  je

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

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

Důkaz.

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

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

### 4.3 Příklad konvexní funkce

Je funkce  $f(x) = \|x\|$  konvexní?

Důkaz.

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

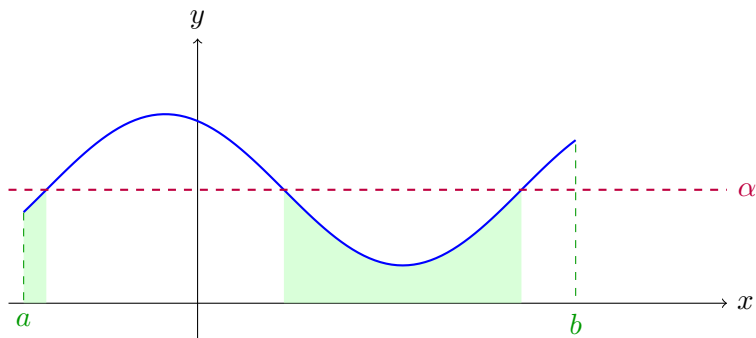
$$\begin{aligned} f(\lambda x + (1 - \lambda)y) &= \|\lambda x + (1 - \lambda)y\| \stackrel{\text{odhad}}{\leq} \|\lambda x\| + \|(1 - \lambda)y\| = \lambda \|x\| + (1 - \lambda)\|y\| \\ &= \lambda f(x) + (1 - \lambda)f(y) \implies f \text{ je konvexní. } \square \end{aligned}$$

#### 4.4 Dolní úrovněová množina

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

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

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



Důkaz.

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

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

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

Poznámka.

Opačná implikace neplatí. Tedy pomocí dolní úrovněové 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 úrovněová množina bude konvexní.

#### 4.5 Použití dolní úrovněové 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:

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

$M_2 = \left\{x \in \mathbb{R}^2 \mid \left\langle x, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\rangle \leq 1\right\} = \text{lev}_{\leq}\left(\left\langle x, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\rangle, 1\right) \rightarrow$  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)  $\alpha f$  je konvexní na  $C$

Důkaz.

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

$$\begin{aligned} (f+g)(\lambda x + (1-\lambda)y) &= \underbrace{f(\lambda x + (1-\lambda)y)}_{\leq \lambda f(x) + (1-\lambda)f(y)} + \underbrace{g(\lambda x + (1-\lambda)y)}_{\leq \lambda g(x) + (1-\lambda)g(y)} \\ &\leq \lambda f(x) + (1-\lambda)f(y) + \lambda g(x) + (1-\lambda)g(y) = \lambda(f+g)(x) + (1-\lambda)(f+g)(y). \quad \square \end{aligned}$$

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

## 4.7 Příklad ověření konvexity

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

Rozeberme si jednotlivé části funkce.

- $e^x \dots$  exponenciála je z grafu očividně konvexní.
- $-3 \ln x \dots$  logaritmus je konkávní, ale díky „-“ se celý výraz stane konvexní. Násobení konstatou konvexitu neovlivní, viz důkaz (b).
- $2x \dots$  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  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$  je afinní. Jestliže  $g(C) \subseteq K$  (tedy  $g$  „obtiskne“ množinu  $C$  do  $K$ ), pak  $f \circ g$  je konvexní na  $C$ .

Důkaz.

At  $x, y \in C, \lambda \in [0, 1]$ .

Pak

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

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.

At  $x, y \in C, \lambda \in [0, 1]$ .

Pak

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

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

## 4.9 Věta o extrémeh 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

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

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

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

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

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

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

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  $f \in C^1(\Omega)$ . 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).

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

$$\begin{aligned} f(x + \lambda(y - x)) &= f(\lambda y + (1 - \lambda)x) \stackrel{f \text{ je konvexní}}{\leq} \lambda f(y) + (1 - \lambda)f(x) = f(x) + \lambda[f(y) - f(x)] \\ &\Rightarrow \underbrace{\frac{f(x + \lambda(y - x)) - f(x)}{\lambda}}_{=\langle \nabla f(x), y - x \rangle \text{ pro } \lambda \rightarrow 0+ \text{ z definice směrové derivace}} \leq f(y) - f(x). \quad \square \end{aligned}$$

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

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

Z předpokladu:

$$f(z) + \langle \nabla f(z), x - z \rangle \leq f(x) \quad / \cdot \lambda \quad (1)$$

$$f(z) + \langle \nabla f(z), y - z \rangle \leq f(y) \quad / \cdot (-\lambda) \quad (2)$$

Pronásobením a sečtením dostaneme:

$$\begin{aligned} f(z) + \lambda \langle \nabla f(z), \underbrace{\lambda x + (1 - \lambda)y - z}_z \rangle &\leq \lambda f(x) + (1 - \lambda)f(y) \\ \Rightarrow f(z) &\leq \lambda f(x) + (1 - \lambda)f(y) \end{aligned}$$

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  $f \in C^2(\Omega)$ . Potom platí:

- (a) Jestliže pro každé  $x \in C$  je  $\nabla^2 f(x)$  pozitivně 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 pozitivně semidefinitní matice pro každé  $x \in C$ .
- (c) Jestliže pro každé  $x \in C$  je  $\nabla^2 f(x)$  pozitivně 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

$$\begin{aligned} f(y) &= f(x) + \langle \nabla f(x), y - x \rangle + \underbrace{\frac{1}{2} \langle \nabla^2 f(\xi)(y - x), y - x \rangle}_{\geq 0} \\ \Rightarrow f(y) &\geq f(x) + \langle \nabla f(x), y - x \rangle \end{aligned}$$

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  $\frac{1}{2} \alpha^2$  ( $\alpha > 0$ ).

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

V limitě  $\alpha \rightarrow 0_+$  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  $\mathbb{R} \times \{0\}$ . ( $\rightarrow$  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.

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

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



#### 4.14 Příklad ověření konvexnosti funkce s parametrem

Mějme funkci  $f(x) = \langle Ax, x \rangle$ , kde

$$A = \begin{bmatrix} 2 & 2 & 3 \\ 1 & 3 & 1 \\ 1 & 2 & \alpha \end{bmatrix}, \alpha \in \mathbb{R} \text{ je parametr.}$$

Pro jaké  $\alpha$  je funkce  $f$  konvexní?

$$\nabla^2 f(x) = \overbrace{A + A^T}^{\text{ze symetrie}} = \begin{bmatrix} 4 & 3 & 4 \\ 3 & 6 & 3 \\ 4 & 3 & 2\alpha \end{bmatrix}$$

$$\det \nabla^2 f(x) = \begin{vmatrix} 4 & 3 & 4 \\ 3 & 6 & 3 \\ 4 & 3 & 2\alpha \end{vmatrix} \stackrel{R_3 - R_1}{=} \begin{vmatrix} 4 & 3 & 4 \\ 3 & 6 & 3 \\ 0 & 0 & 2\alpha - 4 \end{vmatrix} = (2\alpha - 4) \begin{vmatrix} 4 & 3 \\ 3 & 6 \end{vmatrix} = 3(2\alpha - 4) \begin{vmatrix} 4 & 3 \\ 1 & 2 \end{vmatrix} = 30(\alpha - 2)$$

Tedy aby  $f$  byla konvexní funkce:  $30(\alpha - 2) \geq 0 \iff \alpha \geq 2$ .

Musíme vyšetřit menší minory matice.

Vyškrtněme 3. řádek a 3. sloupec:

$$\begin{vmatrix} 4 & 3 \\ 3 & 6 \end{vmatrix} = 15 > 0$$

Vyškrtněme 2. řádek a 2. sloupec:

$$\begin{vmatrix} 4 & 4 \\ 4 & 2\alpha \end{vmatrix} = 8 \begin{vmatrix} 1 & 1 \\ 2 & \alpha \end{vmatrix} = 8(\alpha - 2) \geq 0 \iff \alpha \geq 2 \dots \text{tuto podmínku již vyžadujeme.}$$

Vyškrtněme 1. řádek a 1. sloupec:

$$\begin{vmatrix} 6 & 3 \\ 3 & 2\alpha \end{vmatrix} = 3 \begin{vmatrix} 2 & 1 \\ 3 & 2\alpha \end{vmatrix} = 3(4\alpha - 3) \geq 0 \iff \alpha \geq \frac{3}{4} \dots \text{vyžadujeme již silnější podmínku.}$$

A teď zbylé minoru po vyškrtnutí dvou řádků a sloupců:

$$4 \geq 0, \quad 6 \geq 0, \quad 2\alpha \geq 0 \iff \alpha \geq 0 \dots \text{vyžadujeme již silnější podmínku.}$$

$\implies$  Pokud  $\alpha \geq 2$ , pak je funkce  $f$  konvexní. Při  $\alpha > 2$  je ryze konvexní.

#### 4.15 Příklad ověření konvexity množiny

Mějme množinu

$$M = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \mid x + 2e^{-x+y^2} \leq 4, -x^2 + 3xy - 3y^2 \geq -1 \right\}.$$

Je  $M$  konvexní?

$$\text{Označme: } g_1(x, y) = x + 2e^{-x+y^2} \dots M_1 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid g_1(x, y) \leq 4 \right\}$$

$$g_2(x, y) = x^2 - 3xy + 3y^2 \dots M_2 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid g_2(x, y) \leq 1 \right\}$$

$M = M_1 \cap M_2 \implies$  ukážeme konvexnost  $M_1$  a  $M_2$ , protože průnik zachovává konvexitu.

$\implies g_1$  a  $g_2$  musí být konvexní.

- $g_1$ :

- $x$  je afinní funkce  $\rightarrow$  konvexní.
- součet zachovává konvexitu.
- násobení zachovává konvexitu.
- exponenciála je konvexní funkce (dokonce striktně rostoucí).
- vnitřní funkce  $(-x + y^2)$  je také konvexní.

$\implies g_1$  je konvexní funkce  $\implies M_1$  je konvexní množina.

- $g_2$ :

- kvadrát je konvexní.
- je ale člen „ $xy$ “ konvexní? Musíme se podívat na Hessovu matici.

$$\nabla^2 g_2(x, y) = \begin{bmatrix} 2 & -3 \\ -3 & 6 \end{bmatrix}$$

$$\left. \begin{array}{l} \det \nabla^2 g_2(x, y) = 12 - 9 = 3 > 0 \\ 2 \geq 0 \end{array} \right\} g_2 \text{ je (ryze) konvexní funkce } \implies M_2 \text{ je konvexní množina.}$$

Protože  $M_1$  i  $M_2$  jsou konvexní množiny, pak nutně i  $M_1 \cap M_2 = M$  je konvexní množina.

## 5 Pátý týden

### 5.1 Kužel přípustných směrů

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

- Vektor  $d \in \mathbb{R}^n$  se nazve přípustný směr množiny  $M$  v bodě  $x$ , jestliže existuje  $\delta > 0$  tak, že pro každé  $\alpha \in (0, \delta]$  je  $x + \alpha d \in M$ .
- Množina  $\mathcal{F}(M; x)$  všech přípustných směrů množiny  $M$  v bodě  $x$  se nazývá kužel přípustných směrů množiny  $M$  v bodě  $x$ .

$\mathcal{F}(M; x) \neq \emptyset$ .

Je-li  $x \in \text{int}(M)$ , pak  $\mathcal{F}(M; x) = \mathbb{R}^n$ .

Je-li  $M$  konečná (neprázdná), pak  $\mathcal{F}(M; x) = \{0\}$  pro každé  $x \in M$ .

### 5.2 Přípustné směry poklesu

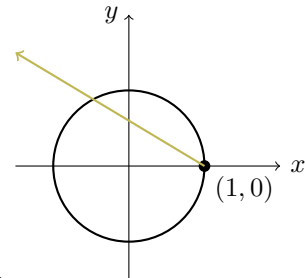
Mějme

(a) Je-li  $M = S(0; 1)$ , pak  $\mathcal{F}(M; x) = \{0\}$  pro každé  $x \in M$ .

(b) Je-li  $C = B(0; 1)$  a  $\hat{x} = (1, 0)^T$ , pak

$$\mathcal{F}(C; \hat{x}) = \left\{ \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \in \mathbb{R}^2 \mid d_1 < 0 \right\} \cup \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$$

(a)  $M = S(0; 1) = \{x \in \mathbb{R}^2 \mid \|x\| = 1\}$



Úvaha: Polopřímka z bodu  $(1, 0)$  projde maximálně  $2 \times$  skrz kružnici.

Ať  $d \neq 0 \in \mathbb{R}^2$

$$\begin{aligned} 1 &= \|x + \alpha d\|^2 = \langle x + \alpha d, x + \alpha d \rangle = \underbrace{\|x\|^2}_1 + 2\alpha \langle x, d \rangle + \alpha^2 \|d\|^2 \\ \rightarrow 0 &= \alpha(2\langle x, d \rangle + \alpha \|d\|^2) \implies \alpha = -\frac{2\langle x, d \rangle}{\|d\|^2} \implies \mathcal{F}(M; x) = \{0\} \end{aligned}$$

(b) Uvažujme kouli

$$M = S(0; 1) = \{x \in \mathbb{R}^2 \mid \|x\| \leq 1\}; \quad x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$1 \geq \left\| \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha d \right\|^2 = \left\langle \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha d, \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha d \right\rangle = \underbrace{\left\| \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\|^2}_1 + 2\alpha \left\langle \begin{bmatrix} 1 \\ 0 \end{bmatrix}, d \right\rangle + \alpha^2 \|d\|^2$$

$$\rightarrow 0 \geq \alpha(2d_1 + \alpha \|d\|^2) \implies \alpha = -\frac{2d_1}{\|d\|^2} \implies \mathcal{F}\left(M; \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \left\{ \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \in \mathbb{R}^2 \mid d_1 < 0 \right\}.$$

### 5.3 Kužel směrů poklesu

Nechť  $D \subseteq \mathbb{R}^n$ ,  $x \in D$  a  $f : D \rightarrow \mathbb{R}$ .

- Vektor  $d \in \mathbb{R}^n$  se nazve směr poklesu funkce  $f$  v bodě  $x$ , jestliže existuje  $\delta > 0$  tak, že pro každé  $\alpha \in (0, \delta]$  je  $f(x + \alpha d) < f(x)$ .
- Množina  $\mathcal{D}(f; x)$  všech směrů poklesu funkce  $f$  v bodě  $x$  se nazývá kužel směrů poklesu funkce  $f$  v bodě  $x$ .

Definice implicitně obsahuje podmínku  $[x, x + \delta d] \subseteq D$ .

### 5.4 Nutná geometrická podmínka lokálního extrému

Jestliže  $x$  je bod lokálního minima funkce  $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  na  $M \subseteq D$ , pak  $\mathcal{F}(M; x) \cap \mathcal{D}(f; x) = \emptyset$ .

Důkaz. Sporem.

At ne, tj. existuje  $d \in \mathcal{F}(M; x) \cap \mathcal{D}(f; x)$ .

Pak:  $f(x + \alpha d) < f(x)$  a  $x + \alpha d \in M$  pro všechna  $\alpha > 0$  dostatečně malá.

Tedy spor s tím, že  $x$  je bod lokálního minima  $f$  na  $M$ .  $\square$

### 5.5 Silný směr poklesu - linearisace směru poklesu

Nechť  $\Omega \subseteq \mathbb{R}^n$  je otevřená množina,  $x \in \Omega$  a  $f \in C^1(\Omega)$ .

- Vektor  $d \in \mathbb{R}^n$  se nazve silný směr poklesu funkce  $f$  v bodě  $x$ , jestliže  $\langle \nabla f(x), d \rangle < 0$ .
- Množina  $\mathcal{D}_0(f; x)$  všech silných směrů poklesu funkce  $f$  v bodě  $x$  se nazývá kužel silných směrů poklesu funkce  $f$  v bodě  $x$ .

Kužel  $\mathcal{D}_0(f; x)$  je množina všech řešení lineární nerovnice

$$\langle \nabla f(x), d \rangle < 0.$$

$\mathcal{D}_0(f; x)$  je konvexní kužel.

### 5.6 Tvzení o souvislosti přípustných směrů poklesu a jejich linearisaci

Nechť  $\Omega \subseteq \mathbb{R}^n$  je otevřená množina,  $x \in \Omega$  a  $f \in C^1(\Omega)$ . Potom platí:

- Je-li  $d \in \mathcal{D}(f; x)$ , potom  $\langle \nabla f(x), d \rangle \leq 0$ .
- $\mathcal{D}_0(f; x) \subseteq \mathcal{D}(f; x)$  (tj. jestliže  $\langle \nabla f(x), d \rangle < 0$ , pak  $d \in \mathcal{D}(f; x)$ ).

Důkaz.

(a) At  $d \in \mathcal{D}(f; x)$ .

$$\begin{aligned} \frac{f(x + \alpha d) - f(x)}{\alpha} &< 0 \text{ pro } \alpha > 0 \text{ dostatečně malé.} \\ \implies \lim_{\alpha \rightarrow 0^+} \underbrace{\frac{f(x + \alpha d) - f(x)}{\alpha}}_{= \langle \nabla f(x), d \rangle} &\leq 0 \quad \square \end{aligned}$$

(b) Ať  $\alpha > 0$ .

$$f(x + \alpha d) = f(x) + \alpha \langle \nabla f(x), d \rangle + \alpha \|d\| \overbrace{\omega(\alpha d)}^{\text{zbytek}}$$

$$\frac{f(x + \alpha d) - f(x)}{\alpha} = \underbrace{\langle \nabla f(x), d \rangle + \|d\| \omega(\alpha d)}_{\substack{\rightarrow \langle \nabla f(x), d \rangle \text{ pro } \alpha \rightarrow 0^+ \\ \text{a navíc } \langle \nabla f(x), d \rangle < 0}} \implies \frac{f(x + \alpha d) - f(x)}{\alpha} < 0 \text{ pro všechna } \alpha < 0 \text{ dostatečně malá.}$$

## 5.7 Fermatova věta - nutná podmínka optimality

Nechť  $\Omega \subseteq \mathbb{R}^n$  je otevřená množina,  $M \subseteq \Omega$  a  $\hat{x} \in M$  je bodem lokálního minima funkce  $f \in C^1(\Omega)$  na  $M$ . Potom platí:

(a)  $\mathcal{F}(M; \hat{x}) \cap \mathcal{D}_0(f; \hat{x}) = \emptyset$  (tj.  $\langle \nabla f(\hat{x}), d \rangle \geq 0$  pro všechny  $d \in \mathcal{F}(M; \hat{x})$ ).

(b) Jestliže  $\hat{x} \in \text{int}(M)$ , pak  $\nabla f(\hat{x}) = 0$ .

Důkaz.

(a) Víme, že  $\mathcal{F}(M; \hat{x}) \cap \mathcal{D}(f; \hat{x}) = \emptyset$ .

Pak:

$$\mathcal{D}_0(f, \hat{x}) \subseteq \mathcal{D}(f, \hat{x}) \implies \mathcal{F}(M; \hat{x}) \cap \mathcal{D}_0(f, \hat{x}) = \emptyset. \quad \square$$

(Tj.  $\langle \nabla f(\hat{x}), d \rangle \geq 0 \quad \forall d \in \mathcal{F}(M, \hat{x})$ )

(b)

$$\hat{x} \in \text{int}(M) \implies \mathcal{F}(M; \hat{x}) = \mathbb{R}^n \xrightarrow{(a)} \langle \nabla f(\hat{x}), d \rangle \geq 0 \quad \forall d \in \mathbb{R}^n$$

Ať  $d = -\nabla f(\hat{x})$ .

$$-\|\nabla f(\hat{x})\|^2 \geq 0 \implies \nabla f(\hat{x}) = 0. \quad \square$$

## 5.8 Věta o nutných a postačujících podmínkách pro konvexní úlohu

Nechť  $\Omega \subseteq \mathbb{R}^n$  je otevřená množina,  $f \in C^1(\Omega)$  je konvexní na  $C \subseteq \Omega$  a  $\hat{x} \in C$ . Potom platí:

(a)  $\hat{x} \in \text{argmin}_{x \in C} f(x)$  právě tehdy, když  $\mathcal{F}(C; \hat{x}) \cap \mathcal{D}_0(f; \hat{x}) = \emptyset$ .

(b) Předpokládejme, že  $\hat{x} \in \text{int}(C)$ . Pak  $\hat{x} \in \text{argmin}_{x \in C} f(x)$  právě tehdy, když  $\nabla f(\hat{x}) = 0$ .

Důkaz.

(a)

„ $\Rightarrow$ “: Víme. Když máme bod minima, je určitě bodem lokálního minima  $\implies$  průnik je prázdný.

„ $\Leftarrow$ “: Sporem.

Ať existuje  $y \in C : \overbrace{f(y) - f(\hat{x})}^{< 0} < 0$ .

Ať  $d = y - \hat{x} (\neq 0) \in \mathcal{F}(C, \hat{x})$ .

Cíl:  $d \in \mathcal{F}(C, \hat{x}) \cap \mathcal{D}_0(f, \hat{x})$ .

$$\underbrace{\hat{x} + \alpha d}_{\hat{x} + \alpha(y - \hat{x}) = \alpha y + (1 - \alpha)\hat{x}} \in C \quad \forall \alpha \in [0, 1] \text{ z konvexity } C.$$

$$f \text{ je konvexní na } C \iff f(\hat{x}) + \langle \nabla f(\hat{x}), \overbrace{y - \hat{x}}^d \rangle \leq f(y). \implies \langle \nabla f(\hat{x}), d \rangle \leq f(y) - f(\hat{x}) \underset{\text{z předp.}}{<} 0.$$

To je ale spor, protože byl předpoklad, že průnik je prázdný. My jsme ale ukázali, že není.  $\square$

(b)

„ $\Rightarrow$ “ Víme.

„ $\Leftarrow$ “ At  $\nabla f(\hat{x}) = 0$ .

Pak  $\langle \nabla f(\hat{x}), d \rangle = 0 \quad \forall d \in \mathbb{R}^n = \mathcal{F}(C; \hat{x})$ . Nemáme tedy žádný směr poklesu  $\xRightarrow{(a)} \hat{x} \in \operatorname{argmin}_{x \in C} f(x)$ .  $\square$

## 5.9 Hledání bodu minima

$$f(x, y) = x^2 + 3y^2 - 2xy + x - 2y$$

$$\nabla^2 f(x, y) = \begin{bmatrix} 2 & -2 \\ -2 & 6 \end{bmatrix} \dots \text{dle Sylvesterova kritéria je pozitivně definitní.}$$

$\Rightarrow f$  je nutně (ryze) konvexní.

$$0 = \nabla f(x, y) = \begin{bmatrix} 2x - 2y + 1 \\ 6y - 2x - 2 \end{bmatrix} \rightarrow \begin{matrix} 2x - 2y = -1 \\ -2x + 6y = 2 \end{matrix} \rightarrow y = \frac{1}{4} \Rightarrow x = -\frac{1}{4}.$$

## 5.10 Věta o podmínkách optimality 2. řádu

Nechť  $\Omega \subseteq \mathbb{R}^n$  je otevřená množina,  $M \subseteq \Omega$ ,  $\hat{x} \in \operatorname{int}(M)$  a  $f \in C^2(\Omega)$ . Potom platí:

- (a) Jestliže  $\hat{x}$  je bod lokálního minima funkce  $f$  na  $M$ , pak  $\nabla^2 f(\hat{x})$  je pozitivně semidefinitní.
- (b) Jestliže  $\nabla f(\hat{x}) = 0$  a  $\nabla^2 f(\hat{x})$  je pozitivně definitní, pak  $\hat{x}$  je bod ostrého lokálního minima.

Důkaz vynecháme.

## 5.11 Příklad použití věty o podmínkách optimality 2. řádu

Je dána funkce

$$f(x, y) = \frac{1}{3}x^3 + \frac{1}{2}y^2 + xy + 2y.$$

Určete lokální extrémy funkce.

$$0 = \nabla f(x, y) = \begin{bmatrix} x^2 + y \\ y + x + 2 \end{bmatrix} \rightarrow \begin{matrix} x^2 + y = 0 \\ y + x + 2 = 0 \end{matrix} \rightarrow x^2 - x - 2 = 0 \rightarrow (x + 1)(x - 2) = 0$$

Podezřelé body jsou:

- $x = -1 \Rightarrow y = -1$
- $x = 2 \Rightarrow y = -4$

$$\nabla^2 f(x, y) = \begin{bmatrix} 2x & 1 \\ 1 & 1 \end{bmatrix}$$

$\nabla^2 f(-1, -1) = \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix} \dots$  dle Sylvesterova kritéria není pozitivně semidefinitní, není ani negativně semidefinitní, je indefinitní. Dle věty o podmínkách optimality 2. řádu není lokálním minimem ani maximem.

$\nabla^2 f(2, -4) = \begin{bmatrix} 4 & 1 \\ 1 & 1 \end{bmatrix} \dots$  dle Sylvesterova kritéria je pozitivně definitní. V bodě  $(2, -4)$  se tedy nachází (ostré) lokální minimum, nikoliv však globální.

## 5.12 Hledání bodu minima

Nalezněte, pokud existují, všechny body minima funkce

$$f(x, y, z) = 2x^2 + y^2 + z^2 + xy - 2xz$$

$$\nabla^2 f(x, y, z) = \begin{bmatrix} 4 & 1 & -2 \\ 1 & 2 & 0 \\ -2 & 0 & 2 \end{bmatrix}$$

$$\left. \begin{aligned} \det \nabla^2 f(x, y, z) &= \begin{vmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ -2 & 0 & 2 \end{vmatrix} = 2 \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 6 > 0 \\ \begin{vmatrix} 4 & 1 \\ 1 & 2 \end{vmatrix} &= 7 > 0 \\ |4| &= 4 > 0 \end{aligned} \right\} \text{ pozitivně definitní} \implies f \text{ je ryze konvexní.}$$

Protože  $f$  je konvexní, body minima budou přesně stacionární body. A protože  $f$  je ryze konvexní, tak bude mít právě jeden bod minima.

$$\begin{aligned} 4x + y - 2z &= 0 &\Rightarrow 2z + y &= 0 \Rightarrow z = -2y \\ x + 2y &= 0 &\Rightarrow x &= -2y \\ -2x &+ 2z &= 0 &\Rightarrow x = z \end{aligned}$$

Jediný bod minima je tedy očividně  $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ .

## 5.13 Omezení ve tvaru nerovnosti - aproximace $\mathcal{F}(M; \hat{x})$

Ať  $g_1, \dots, g_k$  jsou reálné funkce definované na množině  $\Omega \subseteq \mathbb{R}^n$ ,  $M = \{x \in \Omega \mid g_1(x) \leq 0, \dots, g_k(x) \leq 0\}$  a  $x \in M$ . Označme si:

- Množina  $\mathcal{I}((g_i)_{i=1}^k; x) := \{i \in \{1, \dots, k\} \mid g_i(x) = 0\}$  se nazývá indexová množina aktivních omezení v bodě  $x$ .
- Jestliže  $i \in \mathcal{I}((g_i)_{i=1}^k; x)$ , pak  $g_i(x) \leq 0$  se nazve **aktivní** omezení (ve tvaru nerovnosti) v bodě  $x$ .
- Jestliže  $i \notin \mathcal{I}((g_i)_{i=1}^k; x)$ , pak  $g_i(x) < 0$  se nazve **neaktivní** omezení (ve tvaru nerovnosti) v bodě  $x$ .

Poznámka. V textu dále se obvykle bude uvádět pouze  $\mathcal{I}(x) = \{i \in \{1, \dots, k\} \mid g_i(x) = 0\}$ . Když přeindexujeme funkce  $g_i(x)$ , znamenalo by to něco jiného, proto se u  $\mathcal{I}$  uvádí  $((g_i)_{i=1}^k; x)$ , ale my většinou přeindexovávat nebudeme.

Definice.

Nechť  $\Omega \subseteq \mathbb{R}^n$  je otevřená množina,  $g_1, \dots, g_k \in C^1(\Omega)$ ,  $x \in M$  a  $M = \{x \in \Omega \mid g_1(x) \leq 0, \dots, g_k(x) \leq 0\}$ . Definujme množinu

$$\begin{aligned} \mathcal{G}((g_i)_{i=1}^k; x) &:= \{d \in \mathbb{R}^n \mid \langle \nabla g_i(x), d \rangle \leq 0 \quad \forall i \in \mathcal{I}(x)\} \\ &= \bigcap_{i \in \mathcal{I}(x)} \{d \in \mathbb{R}^n \mid \langle \nabla g_i(x), d \rangle \leq 0\} \end{aligned}$$

jako aproximaci  $\mathcal{F}(M; \hat{x})$ .

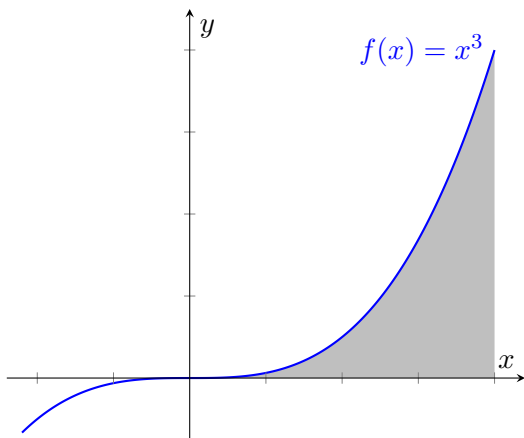
## 5.14 Příklad výpočtu $\mathcal{G}$ a $\mathcal{F}$

Je dána množina

$$M = \{(x, y)^T \in \mathbb{R}^2 \mid y - x^3 \leq 0, -y \leq 0\}$$

a bod  $\hat{x} = (0, 0)^T$ . Určete množiny  $\mathcal{F}(M; \hat{x})$  a  $\mathcal{G}((g_i)_{i=1}^k; \hat{x})$ .

Nákres množiny.



Výpočet  $\mathcal{F}(M; \hat{x})$ .

?  $0 + \alpha d \in M \quad \forall \alpha > 0$  dostatečně malé.

$$\alpha d_2 - \alpha^3 d_1^3 \leq 0 \quad (3)$$

$$-\alpha d_2 \leq 0 \quad \forall \alpha > 0 \text{ dostatečně malé.} \quad (4)$$

$$(4) \implies d_2 \geq 0$$

$$(3) \implies d_2 \leq \alpha^2 d_1^3 \quad \forall \alpha > 0 \text{ dostatečně malé.}$$

$d_2 \geq 0 \implies d_1 \geq 0$  a protože to platí  $\forall \alpha > 0$  dostatečně malá, pak  $d_2 = 0$ , protože si můžu vzít libovolně malé, tedy i limitně blízké nule,  $\alpha$ .

$$\implies \mathcal{F}(M; (0, 0)) = \left\{ \begin{bmatrix} d_1 \\ 0 \end{bmatrix} \in \mathbb{R}^2 \mid d_1 \geq 0 \right\}.$$

Výpočet  $\mathcal{G}((g_i)_{i=1}^k; \hat{x})$ .

Označme si  $g_1(x, y) = y - x^3$  a  $g_2(x, y) = -y$ .

Pak:

$$\left. \begin{aligned} \langle \nabla g_1(0), d \rangle &= \left\langle \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \right\rangle = d_2 \leq 0 \\ \langle \nabla g_2(0), d \rangle &= \left\langle \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \right\rangle = d_2 \geq 0 \end{aligned} \right\} d_2 = 0$$

$$\implies \mathcal{G}((g_1, g_2), (0, 0)) = \left\{ \begin{bmatrix} d_1 \\ 0 \end{bmatrix} \mid d_1 \in \mathbb{R} \right\} \rightarrow \mathcal{G} \text{ je větší jak } \mathcal{F}.$$

Protože  $\mathcal{G}$  je pouze aproximací  $\mathcal{F}$ , může a bude se stávat, že  $\mathcal{G}$  bude větší jak  $\mathcal{F}$ .

Přidejme si další, fakticky zbytečnou, podmínku navíc.

$$M = \{(x, y)^T \in \mathbb{R}^2 \mid \underbrace{y - x^3}_{g_1(x, y)} \leq 0, \underbrace{-y}_{g_2(x, y)} \leq 0, \underbrace{-x - y}_{g_3(x, y)} \leq 0\}$$

$$\langle \nabla g_3(0), d \rangle = \left\langle \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \right\rangle = -d_1 - \underbrace{d_2}_{=0} \leq 0 \implies -d_1 \leq 0$$

$$\implies \mathcal{G}((g_1, g_2, g_3), (0, 0)) = \left\{ \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \mid d_1 \geq 0 \right\}. \text{ Což odpovídá přesně množině } \mathcal{F}.$$

Je tedy očividné, že  $\mathcal{G}$  závisí na popisu množiny.



### 5.15 Ukázka, že aproximací $\mathcal{F}$ lze zkazit prázdnotu průniku

Mějme optimalisační úlohu

$$\begin{aligned} &\text{minimalisujte } x + y \\ &\text{za podmínek } y - x^3 \leq 0, \\ &\quad -y \leq 0. \end{aligned}$$

Pak

$$\begin{aligned} \mathcal{D}_0(f; 0) &= \{d \in \mathbb{R}^n \mid \langle \nabla f(0), d \rangle < 0\} \\ &\stackrel{=}{=}_{\nabla f(0)=(1,1)} \left\{ \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \mid d_1 + d_2 < 0 \right\} \dots \text{například } \begin{bmatrix} -1 \\ 0 \end{bmatrix} \in \mathcal{D}_0(f; 0), \text{ ale } \begin{bmatrix} -1 \\ 0 \end{bmatrix} \in \mathcal{G}(x)! \end{aligned}$$

Tedy  $\mathcal{G}(\hat{x}) \cap \mathcal{D}_0(f; \hat{x}) \neq \emptyset \implies$  nahrazením podmínek optimality můžeme zkazit prázdnotu průniku, protože  $\mathcal{G}$  může být větší jak  $\mathcal{F}$ .

## 5.16 Věta o nutných KKT podmínkách

Nechť  $\Omega \subseteq \mathbb{R}^n$  je otevřená množina,  $f, g_1, \dots, g_k \in C^1(\Omega)$ ,

$$M = \{x \in \Omega \mid g_1(x) \leq 0, \dots, g_k(x) \leq 0\}$$

a  $\hat{x} \in M$ . Jestliže  $\overline{\mathcal{F}(M; \hat{x})} = \mathcal{G}(\hat{x})$  a  $\hat{x}$  je bod lokálního minima na  $f$  na  $M$ , pak existuje  $(\mu_1, \dots, \mu_k)^T \in \mathbb{R}^k$  tak, že

$$\begin{aligned} \nabla f(\hat{x}) + \sum_{i=1}^k \mu_i \nabla g_i(\hat{x}) &= 0, \\ \mu_i g_i(\hat{x}) &= 0 \text{ pro všechna } i \in \{1, \dots, k\}, \\ \mu_i &\geq 0 \text{ pro všechna } i \in \{1, \dots, k\}. \end{aligned}$$

Důkaz.

- $\mathcal{I}(\hat{x}) = \emptyset \implies \hat{x} \in \text{int}(M) \implies \nabla f(\hat{x}) = 0$  z **Fermatovy věty**.  
 $\rightarrow$  volba  $\mu_1 = \dots = \mu_k = 0$ . Pak KKT podmínky splněny.

- $\emptyset \neq \mathcal{I}(\hat{x}) = \{1, \dots, l\}$

Víme, že máme bod lokálního minima  $(\hat{x}) \xRightarrow{\text{Fermatova věta}} \mathcal{F}(M; \hat{x}) \cap \mathcal{D}_0(f; \hat{x}) = \emptyset$ ,

tj.  $\langle \nabla f(\hat{x}), d \rangle \geq 0 \quad \forall d \in \mathcal{F}(M; \hat{x})$ .

Tedy chceme dokázat, že platí  $\langle \nabla f(\hat{x}), d \rangle \geq 0 \quad \forall d \in \mathcal{G}(\hat{x})$ .

Protože  $\overline{\mathcal{F}(M; \hat{x})}$  koinciduje s  $\mathcal{G}(\hat{x})$  a ze spojitosti skalárního součinu plyne, že  
 $\langle \nabla f(\hat{x}), d \rangle \geq 0 \quad \forall d \in \underbrace{\overline{\mathcal{F}(M; \hat{x})}}_{\mathcal{G}(\hat{x})}$ .

To tedy znamená  $\langle \nabla f(\hat{x}), d \rangle \geq 0 \quad \forall d \in \mathcal{G}(\hat{x})$ . Z toho plyne, že neexistuje  $d \in \mathbb{R}^n$ , pro který platí:

$$\left. \begin{aligned} \langle \nabla f(\hat{x}), d \rangle &< 0 \dots \text{ tj. } \langle -\nabla f(\hat{x}), d \rangle > 0 \\ \langle \nabla g_1(\hat{x}), d \rangle &\leq 0 \\ &\vdots \\ \langle \nabla g_l(\hat{x}), d \rangle &\leq 0 \end{aligned} \right\} A^T d \leq 0, \text{ kde } A = (\nabla g_1(\hat{x}), \dots, \nabla g_l(\hat{x}))$$

No a z **Farkasova lemma** tedy nutně platí: ex.  $\mu = (\mu_1, \dots, \mu_l)^T \in \mathbb{R}_+^l : \underbrace{A\mu}_{\sum_{i=1}^l \mu_i \nabla g_i} = -\nabla f(\hat{x})$ .

$\rightarrow$  volme dále  $\mu_{l+1}, \dots, \mu_k = 0$ . Pak

$$\begin{aligned} -\nabla f(\hat{x}) &= \sum_{i=1}^k \mu_i \nabla g_i(\hat{x}), \\ \mu_i \nabla g_i(\hat{x}) &= 0 \text{ pro všechna } i \in \{1, \dots, k\}, \\ \mu_i &\geq 0 \text{ pro všechna } i \in \{1, \dots, k\}. \end{aligned}$$

A to jsou přesně KKT podmínky.  $\square$

## 5.17 Příklad použití KKT podmínek

$$\begin{aligned} &\text{minimalisujte } \underbrace{x+y}_{f(x,y)} \\ &\text{za podmínek } \underbrace{x}_{g_1(x,y)} \geq 0, \\ &\quad \underbrace{y}_{g_2(x,y)} \geq 0. \end{aligned}$$

Určete KKT body.

$$\nabla f(x, y) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \nabla g_1(x, y) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \nabla g_2(x, y) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

KKT podmínky:

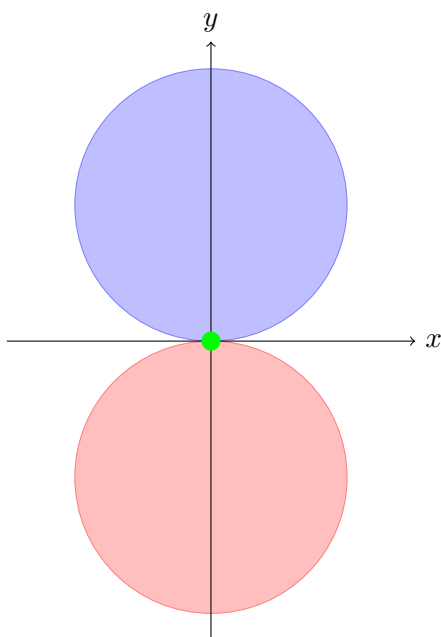
$$\begin{aligned} 1 + \mu_1(-1) + \mu_2(0) &= 0 &\rightarrow \mu_1 &= 1 \\ 1 + \mu_1(0) + \mu_2(-1) &= 0 &\rightarrow \mu_2 &= 1 \\ \mu_1(-x) &= 0 \\ \mu_2(-y) &= 0 \\ \mu_1, \mu_2 &\geq 0 \end{aligned}$$

Jediný KKT bod je tedy  $\hat{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$  a jedná se o bod minima.

## 5.18 Příklad, že KKT podmínky vždy nenaleznou všechny body

$$\begin{aligned} &\text{minimalisujte } x \\ &\text{za podmínek } x^2 + (y-1)^2 \leq 1, \\ &\quad x^2 + (y+1)^2 \leq 1. \end{aligned}$$

Nákres.



**Přípustná** množina:  $M = \{0\} \rightarrow$  určitě konvexní množina.

KKT podmínky:

$$\begin{aligned} 1 + \mu_1(2 \cdot 0) + \mu_2(2 \cdot 0) &= 0 \times \\ &\vdots \end{aligned}$$

$\Rightarrow (0, 0)$  není KKT bod i když je úloha konvexní a bod  $(0, 0)$  je očividně bodem minima.

## 5.19 Věta o postačujících KKT podmínkách

Nechť  $\Omega \subseteq \mathbb{R}^n$  je otevřená množina,  $f, g_1, \dots, g_k \in C^1(\Omega)$  jsou konvexní funkce na  $C = \{x \in \Omega \mid g_1(x) \leq 0, \dots, g_k(x) \leq 0\}$ . Jestliže  $\hat{x} \in C$  je KKT bod, pak  $\hat{x}$  je bod minima funkce  $f$  na  $C$ .

Důkaz. Ať  $x \in C$ .

Cíl:  $f(x) - f(\hat{x}) \geq 0$  ( $= \hat{x}$  je minimum)

Charakterisace pomocí tečné nadroviny:  $f(\hat{x}) + \langle \nabla f(\hat{x}), x - \hat{x} \rangle \leq f(x) \quad x, \hat{x} \in C$

$$\begin{aligned} f(x) - f(\hat{x}) &\stackrel{\substack{f \text{ je konvexní} \\ \text{na } C \subseteq \Omega}}{\geq} \langle \nabla f(\hat{x}), x - \hat{x} \rangle \stackrel{\substack{\text{podmínka} \\ \text{stacionarity}}}{=} \left\langle - \sum_{i=1}^k \mu_i \nabla g_i(\hat{x}), x - \hat{x} \right\rangle \\ &= \sum_{i=1}^k -\langle \nabla g_i(\hat{x}), x - \hat{x} \rangle \mu_i = \sum_{i=1}^n (g_i(\hat{x}) - g_i(x)) \mu_i \stackrel{\substack{\text{podmínka} \\ \text{komplementarity}}}{=} - \sum_{i=1}^n \underbrace{\mu_i}_{\substack{\text{podmínka} \\ \text{nezápornosti}}} \overbrace{g_i(x)}^{\leq 0 \forall x \in C} \geq 0. \quad \square \end{aligned}$$

## 5.20 Afinní podmínka regularity

Nechť  $\Omega \subseteq \mathbb{R}^n$  je otevřená množina,  $g_1, \dots, g_k \in C^1(\Omega)$  a

$$M = \{x \in \Omega \mid g_1(x) \leq 0, \dots, g_k(x) \leq 0\}.$$

Řekněme, že  $(g_i)_{i=1}^k$  splňuje afinní podmínku regularity, jestliže  $g_1, \dots, g_k$  jsou afinní.

## 5.21 Slaterova podmínka regularity

Nechť  $\Omega \subseteq \mathbb{R}^n$  je otevřená množina,  $g_1, \dots, g_k \in C^1(\Omega)$  a

$$M = \{x \in \Omega \mid g_1(x) \leq 0, \dots, g_k(x) \leq 0\}.$$

Řekněme, že  $(g_i)_{i=1}^k$  splňuje Slaterovu podmínku regularity, jestliže  $g_1, \dots, g_k$  jsou konvexní na  $\Omega$  a existuje  $x \in \Omega$  tak, že pro každé  $i \in \{1, \dots, k\}$  je  $g_i(x) < 0$ .

## 5.22 Použití podmínek regularity k ověření KKT podmínek

$$\begin{aligned} &\text{minimalisujte } 2x^2 + y^2 \\ &\text{za podmínek } x^2 + y^2 - 1 \leq 0, \\ &\quad \quad \quad -x \leq 0. \end{aligned}$$

Afinní podmínka splněna není, ověříme Slaterovu.

Množina je očividně konvexní a zároveň zvolme  $x = \begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix} \in \Omega$ . Pak  $g_i(x) < 0$ , Slaterova podmínka je tedy očividně splněna.

$$\nabla f(x) = \begin{bmatrix} 4x \\ 2y \end{bmatrix}, \nabla g_1(x) = \begin{bmatrix} 2x \\ 2y \end{bmatrix}, \nabla g_2(x) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

$\Rightarrow$  KKT podmínky:

$$\begin{aligned} 4x + \mu_1 2x + \mu_2(-1) &= 0 \Leftrightarrow 2x(2 + \mu_1) - \mu_2 = 0 \\ 2y + \mu_1 2y + \mu_2 0 &= 0 \Leftrightarrow 2y(1 + \mu_1) = 0 \xrightarrow{\mu_1 \geq 0} y = 0 \\ \mu_1(x^2 + y^2 - 1) &= 0 \\ \mu_2(-x) &= 0 \\ \mu_1, \mu_2 &\geq 0 \end{aligned}$$

$y = 0$ :

$$\left. \begin{aligned} 2x(2 + \mu_1) &= \mu_2 \\ \mu_1(x^2 + 1) &= 0 \\ \mu_2 x &= 0 \\ \mu_1, \mu_2 &\geq 0 \end{aligned} \right\} \begin{aligned} x \neq 0 &\Rightarrow \mu_2 = 0 \Rightarrow 2 + \mu_1 = 0 \dots \text{spor s } \mu_1 \geq 0. \\ x = 0 &\Rightarrow \mu_1 = 0 \Rightarrow \mu_2 = 0 \quad \checkmark \end{aligned}$$

Existuje bod  $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ , pro který jsou splněny nutné a postačující KKT podmínky.

### 5.23 Určení nutných a postačujících podmínek optimality

Ať  $A \in \mathbb{M}_{m,n}(\mathbb{R})$ ,  $D \in \mathbb{M}_{r,n}(\mathbb{R})$ ,  $b \in \mathbb{R}^m$  a  $\lambda > 0$ . Je dána úloha

$$\text{minimalisujte } f(x) = \|Ax - b\|^2 + \lambda \|Dx\|^2 \text{ na } \mathbb{R}^n.$$

Jaké jsou nutné a postačující podmínky optimality?

$$\begin{aligned} f(x) &= \langle Ax - b, Ax - b \rangle + \lambda \langle Dx, Dx \rangle \\ &= \underbrace{\langle Ax, Ax \rangle}_{A^T A x, x} - 2 \langle Ax, b \rangle + \|b\|^2 + \lambda \underbrace{\langle Dx, Dx \rangle}_{D^T D x, x} \end{aligned}$$

$$\Rightarrow f(x) = \langle (A^T A + \lambda D^T D) x, x \rangle - 2 \langle x, A^T b \rangle + \|b\|^2$$

Je  $f$  konvexní?

Ano, neboť  $\nabla^2 f(x) = 2(A^T A + \lambda D^T D)$  je pozitivně semidefinitní, protože pro  $x \in \mathbb{R}^n$ :

$$\begin{aligned} \langle 2(A^T A + \lambda D^T D) x, x \rangle &= 2[\langle Ax, Ax \rangle + \lambda \langle Dx, Dx \rangle] \\ &= 2[\|Ax\|^2 + \lambda \|Dx\|^2] \geq 0 \end{aligned}$$

Tedy  $f$  je konvexní  $\Rightarrow$  stačí najít stacionární body.

$$\begin{aligned} 0 = \nabla^2 f(x) &= 2(A^T A + \lambda D^T D)x - 2(A^T b) + 0 \\ &= (A^T A + \lambda D^T D)x - A^T b \\ \Rightarrow A^T b &= (A^T A + \lambda D^T D)x \end{aligned}$$

A to je nutná a postačující podmínka pro  $x$ , aby byl bodem minima  $f$  na  $\mathbb{R}^n$ .

### 5.24 Určení KKT podmínek

$$\begin{aligned} &\text{minimalisujte } x^4 + y^4 + 12x^2 + 6y^2 - xy - x - y \\ &\text{za podmínkou } x + y \geq 6, \\ &\quad 2x - y \geq 3, \\ &\quad x, y \geq 0. \end{aligned}$$

- a) Napište KKT podmínky.  
 b) Jsou nutné a postačující?  
 c) Ukažte, že  $(3, 3)^T$  je jediný bod minima.

a) Mějme

$$\begin{aligned}g_1(x, y) &= -x - y + 6, \\g_2(x, y) &= 2x - y + 3, \\g_3(x, y) &= -x, \\g_4(x, y) &= -y, \\f(x, y) &= x^4 + y^4 + 12x^2 + 6y^2 - xy - x - y.\end{aligned}$$

→ použijeme **afinní podmínku regularity** →  $g_i$  jsou affinní.

KKT podmínky:

$$\begin{aligned}\nabla f(x, y) + \mu_1 \nabla g_1(x, y) + \mu_2 \nabla g_2(x, y) + \mu_3 \nabla g_3(x, y) + \mu_4 \nabla g_4(x, y) &= 0 \\ \mu_i g_i(x, y) &= 0, i = 1, 2, \dots, \\ \mu_i &\geq 0, i = 1, 2, \dots\end{aligned}$$

Tedy:

$$\begin{aligned}4x^3 + 24x - y - 1 - \mu_1 - 2\mu_2 - \mu_3 &= 0 \\ 4y^3 + 12y - x - 1 - \mu_1 + \mu_2 - \mu_4 &= 0 \\ \mu_1(-x - y + 6) &= 0, \\ \mu_2(x - 2y + 3) &= 0, \\ x\mu_3 &= 0, \\ y\mu_4 &= 0, \\ \mu_1, \mu_2, \mu_3, \mu_4 &\geq 0.\end{aligned}$$

Jsou postačující? Máme konvexní úlohu? Musíme ověřit konvexitu u  $g_i$  a  $f$ .

- $g_i$  jsou afinní  $\implies$  jsou konvexní.
- $f$  :
  - kvadráty jsou ryze konvexní
  - součet ryzích konvexních je ryzí konvexní

$$h(x, y) = 12x^2 + 6y^2 - xy - x - y$$

$$\nabla^2 h(x, y) = \begin{bmatrix} 24 & -1 \\ -1 & 12 \end{bmatrix} = 24 \cdot 12 - 1 > 0; \quad 24 > 0 \implies h(x, y) \text{ je pozitivně definitní.}$$

$\implies h(x, y)$  je ryze konvexní.

A proto je i  $f(x, y)$  ryze konvexní, protože součet ryze konvexních dává ryze konvexní  $\implies$  existuje právě jeden bod minima.

Overíme  $\begin{bmatrix} 3 \\ 3 \end{bmatrix}$ . Ať  $x = y = 3$ . Pak

$$4 \cdot 27 + 24 \cdot 3 - 4 - \mu_1 + \mu_2 - \mu_3 = 0 \quad (\text{I.})$$

$$4 \cdot 12 + 12 \cdot 3 - 4 - \mu_1 - 2\mu_2 - \mu_4 = 0 \quad (\text{II.})$$

$$\mu_1 \cdot 0 = 0$$

$$\mu_2 \cdot 0 = 0$$

$$3\mu_3 = 0 \implies \mu_3 = 0$$

$$3\mu_4 = 0 \implies \mu_4 = 0$$

$$\mu_1, \mu_2 \geq 0$$

$$\begin{aligned} \text{I.} - \text{II.}: 24 \cdot 3 - 12 \cdot 3 - 3\mu_2 = 0 &\implies \mu_2 = \frac{1}{3}(24 \cdot 3 - 12 \cdot 3) > 0. \\ \mu_1 = 4 \cdot 27 + 24 \cdot 3 - 4 - \frac{2}{3}(24 \cdot 3 - 36) &> 0. \end{aligned}$$

## 5.25 Určení KKT podmínek

minimalisujte  $\alpha x + y$ ,  $\alpha \in \mathbb{R}$  je parametr.

za podmínek  $x^2 + y^2 - 25 \leq 0$ ,

$$x - y - 1 \leq 0.$$

Určete  $\alpha$  tak, aby  $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$  bylo řešení.

KKT podmínky:

$$\alpha + \mu_1(2x) + \mu_2 \cdot 1 = 0$$

$$1 + \mu_1(2y) - \mu_2 = 0$$

$$\mu_1(x^2 + y^2 - 25) = 0,$$

$$\mu_2(x - y - 1) = 0,$$

$$\mu_1, \mu_2 \geq 0.$$

$g_i$  jsou konvexní,  $f$  je konvexní  $\implies$  KKT podmínky jsou postačující.

**Slaterova podmínka optimality** je splněna  $\implies$  KKT podmínky jsou nutné.

$x = 4, y = 3$ :

$$\alpha + 8\mu_1 + \mu_2 = 0 \quad (\text{I.})$$

$$1 + 6\mu_1 - \mu_2 = 0, \quad (\text{II.})$$

$$\mu_1, \mu_2 \geq 0.$$

$$\text{I.} + \text{II.}: \alpha + 1 + 14\mu_1 = 0$$

$$\mu_1 = \frac{-\alpha-1}{14} \stackrel{!}{\geq} 0 \implies -1 \geq \alpha. \text{ A tedy } \mu_2 = 1 + 6\mu_1 \geq 0.$$

Tedy aby  $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$  bylo řešení této úlohy, musí platit  $\alpha \leq -1$ .

## 6 Šestý týden

### 6.1 Pomocný důkaz vlastnosti infima

$$\inf_{(x,y) \in M \times N} (h_1(x) + h_2(y)) = \inf_{x \in M} h_1(x) + \inf_{y \in N} h_2(y)$$

Důkaz.

„ $\geq$ “:

$$\inf_{x \in M} h_1(x) \leq h_1(t) \quad \forall t \in M$$

$$\inf_{y \in N} h_2(y) \leq h_2(s) \quad \forall s \in N$$

$$\implies \inf_{x \in M} h_1(x) + \inf_{y \in N} h_2(y) \leq \inf_{(x,y) \in M \times N} (h_1(x) + h_2(y)) \quad \square$$

„ $\leq$ “:

$$\inf_{(x,y) \in M \times N} (h_1(x) + h_2(y)) \leq h_1(t) + h_2(s) \quad \forall \begin{bmatrix} t \\ s \end{bmatrix} \in M \times N$$

$$\text{což lze upravit: } -h_2(s) + \inf_{(x,y) \in M \times N} (h_1(x) + h_2(y)) \leq h_1(t) \quad \forall t \in M, \forall s \in N.$$

$$\implies -h_2(s) + \inf_{(x,y) \in M \times N} (h_1(x) + h_2(y)) \leq \inf_{x \in M} h_1(x) \quad \forall s \in N.$$

A to samé lze ukázat i pro  $h_2$ :

$$-h_1(t) + \inf_{(x,y) \in M \times N} (h_1(x) + h_2(y)) \leq h_2(s) \quad \forall t \in M, \forall s \in N.$$

$$\implies -h_1(t) + \inf_{(x,y) \in M \times N} (h_1(x) + h_2(y)) \leq \inf_{y \in N} h_2(y) \quad \forall t \in M.$$

Ted' sečteme tyto dvě nerovnice:

$$-h_1(t) - h_2(s) + 2 \inf_{(x,y) \in M \times N} (h_1(x) + h_2(y)) \leq \inf_{x \in M} h_1(x) + \inf_{y \in N} h_2(y) \quad \forall t \in M, \forall s \in N.$$

Tedy stačí dokázat, že

$$\inf_{(x,y) \in M \times N} (h_1(x) + h_2(y)) \leq -h_1(t) - h_2(s) + 2 \inf_{(x,y) \in M \times N} (h_1(x) + h_2(y)) \quad \forall t \in M, \forall s \in N.$$

$$0 \leq -h_1(t) - h_2(s) + \inf_{(x,y) \in M \times N} (h_1(x) + h_2(y)) \quad \forall t \in M, \forall s \in N.$$

$$h_1(t) + h_2(s) \leq \inf_{(x,y) \in M \times N} (h_1(x) + h_2(y)) \quad \forall t \in M, \forall s \in N.$$

### 6.2 Dualita - motivační příklad

Je dána úloha

$$\begin{aligned} &\text{minimalisujte } 2x + 3y \\ &\text{za podmínek } 1 - x - y \leq 0, \\ &x, y \in [0, 2]. \end{aligned}$$

Označme  $f(x, y) = 2x + 3y$ ,  $M = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in [0, 2]^2 \mid 1 - x - y \leq 0 \right\}$  a  $\hat{f} = \min_{x \in M} f(x)$ .

Odhadněme min funkce ze spoda.



Pro  $(x, y)^T \in M$ :

$$f(x, y) \geq f(x, y) + g_1(x, y) = 2x + 3y + (1 - x - y) = x + 2y + 1 \geq 1.$$

A protože  $\hat{f} = \min f(x)$ , nutně musí platit  $\hat{f} \geq 1$ .

Zkusme teď jiný odhad.

$$f(x, y) \geq f(x, y) + 2g_1(x, y) = 2x + 3y + 2(1 - x - y) = 2y + 1 \geq 2.$$

Nalezli jsme lepší odhad:  $\hat{f} \geq 2$ . Jak tedy správně určit „nejlepší“ možný dolní odhad  $\hat{f}$ ?

Definujme si

$$L(x, y, \mu) = 2x + 3y + \mu(1 - x - y),$$
$$\varphi(\mu) = \min_{(x, y)^T \in [0, 2]^2} L(x, y, \mu).$$

Pro každé  $\mu \geq 0$  pak platí:

$$\varphi(\mu) = \min_{(x, y)^T \in \Omega} L(x, y, \mu) \leq \min_{(x, y)^T \in M} L(x, y, \mu) \leq \hat{f}$$

„Optimální“ dolní odhad  $\hat{f}$  pomocí  $\varphi$  vede na úlohu

$$\begin{aligned} &\text{maximalisujte } \varphi(\mu), \\ &\text{za podmínek } \mu \geq 0. \end{aligned}$$

Tu budeme nazývat duální úlohou.

Hodnota  $\max \varphi(\mu)$  na  $[0, +\infty]$  je  $\hat{\varphi} \implies \hat{\varphi} \leq \hat{f}$ .

## 7 Sedmý týden

## 8 Osmý týden

## 9 Devátý týden

## 10 Desátý týden

## 11 Jedenáctý týden

## 12 Dvanáctý týden

## 13 Třináctý týden



## 14 Čtrnáctý týden