České vysoké učení technické v Praze

Fakulta elektrotechnická

Sbírka řešených příkladů

Optimalizace a teorie her

Jakub Adamec

Praha, 2025

https://github.com/knedl1k/A8B01OGT

Obsah

Úvod

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

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

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

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

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

Stručné informace o textu

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

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

1 Úvod do matematické optimalisace

1.1 Důkaz souvislosti minima a maxima

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

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

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

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

1.2 Hledání přípustných množin

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

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

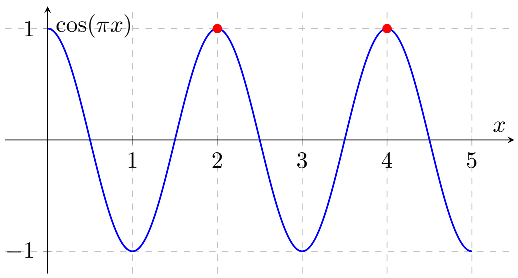
Úvahou pak lze uhodnout minimum - minimum leží v bodě $x = 3$.

1.3 Hledání přípustných množin

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

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

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



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 invetičního produktu (v tis. Kč) při investici x (v tis. Kč) je $\frac{x}{x+50}$. Jakým způsobem má investor rozdělit částku $c = 100000$ Kč mezi uvedené dva produkty tak, aby celkový očekávaný měsíční výnos byl co největší?

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

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

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

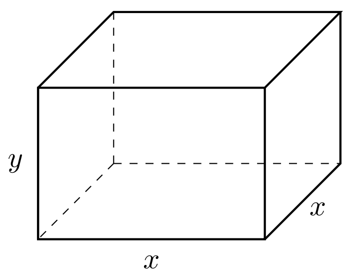
Zbavme se zlomků:

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

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

1.5 Minimalisační úloha

Ve firmě potřebují nalézt rozměry otevřené krabice (tj. krabice bez horní stěny) se čtvercovou podstavou o objemu 10 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.



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

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

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

Zbavme se zlomků:

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

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

1.6 Optimalisační úloha s nadrovinami

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

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

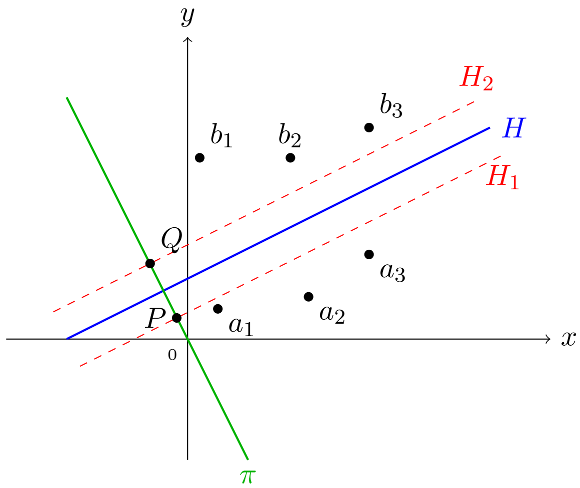
(b) Iterpretujte optimalisační úlohu

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

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

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

(a)



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

Průsečík Q:

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

Průsečík P:

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

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

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

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

(b)

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

2 Konvexní množiny

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

2.1 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.

2.2 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.

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

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

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

2.3 Je uzavřený poloprostor konvexní?

2.4 Je uzavřená koule konvexní?

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

Důkaz.

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

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

2.5 Je okolí konvexní?

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

Důkaz.

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

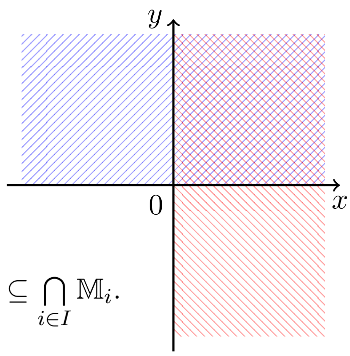
$$
\begin{aligned}
||[\lambda x + (1-\alpha)y] - a|| &= ||\lambda x - (1-\lambda)a + (1-\lambda)y - \lambda a|| = ||\lambda(x-a) + (1-\lambda)(y-a)|| \\
&\leq \lambda||\underbrace{x-a}_{<r}|| + (1-\lambda)||\underbrace{y-a}_{<r}|| < \lambda r + (1-\lambda)r = r. \quad \square
\end{aligned}
$$

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

Visuálně je průnik nekonvexní.



Důkaz.

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

2.7 Důkaz, že rozdíl a sjednocení nezachovává konvexitu

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

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

Afinní zobrazení

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

2.8 Důkaz, že afinní zobrazení je konvexní

Tvrzení.

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

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

Důkaz.

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

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

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

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

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

Musíme ověřit uzavřenost na násobení a sčítání z definice.

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

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

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

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

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

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

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

Tvrzení.

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

Důkaz.

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

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

2.10 Důkaz, že kartézský součin je konvexní

Tvrzení.

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

Důkaz.

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

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

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

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

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

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

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

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

2.11 Určení definitnosti matic

Určete definitnost matice A, jestliže

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

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

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

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

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

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

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

Pak platí:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

2.12 Existence matice

Ať $A \in \mathbb{M}_n(\mathbb{R})$.

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

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

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

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

(a)Využijme zmíněné vlastnosti.

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

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

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

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

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

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

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

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

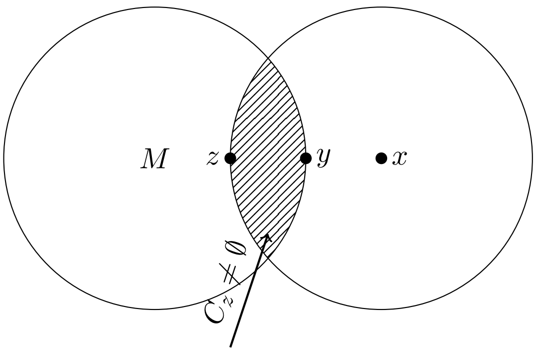
3 Projekce

3.1 Věta o nejlepší aproximaci

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

Důkaz.

1. Existence
Cíl: Existuje bod minima
Úvaha:



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

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

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

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

2. Jednoznačnost.

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

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

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

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

3.2 Projekce bodu a variační nerovnost

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

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

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

Důkaz.

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

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

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

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

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

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

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

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

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

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

3.3 Koule?

3.4 Věta o ortogonálním rozkladu

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

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

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

(c) Pro každé $x \in \mathbb{R}^n$ existují jednoznačně určené body $y \in L$ a $z \in L^\perp$ tak, že $x = y + z$. Navíc $y = P_L(x)$ a $z = P_{L^\perp}(x)$.

Důkaz.

(a)
Cíl: Dokázat vlastnosti lineárního zobrazení, tedy

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

1. : Ať $z \in L$. Pak

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

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

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

2. : Ať $z \in L$.

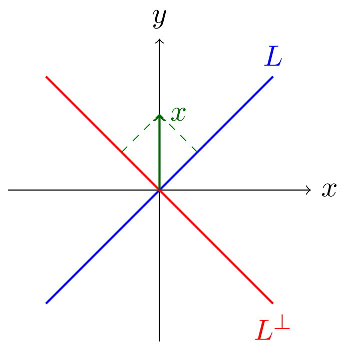
$$\langle \underbrace{x + y - (P_L(x) + P_L(y))}_{(x-P_L(x))+(y-P_L(y))}, z - (P_L(x) + P_L(y))\rangle$$

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

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

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

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



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

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

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

(c) Pro každé $x \in \mathbb{R}^n$ existují jednoznačně určené body $y \in L$ a $z \in L^\perp$ tak, že $x = y + z$. Navíc $y = P_L(x)$ a $z = P_{L^\perp}(x)$.

Ať $x \in \mathbb{R}^n$.

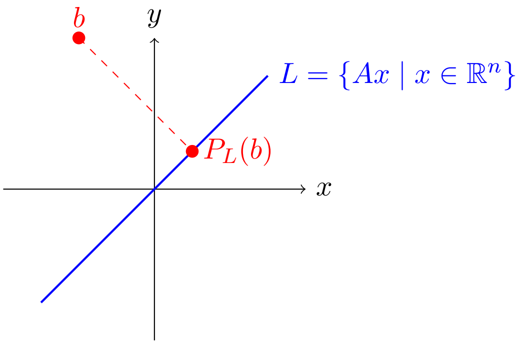
Důkaz existence.
Pak $x = \underbrace{P_L(x)}_{\in L} + \underbrace{(x - P_L(x))}_{\in L^\perp}$. $\square$

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

Ať $z \in L$.

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

4 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 $\hat{x} \in \operatorname*{argmin}_{x\in\mathbb{R}^n} \|Ax - b^2\| \iff A^TA\hat{x} = A^Tb$.

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

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

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

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

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

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

4.1 Příklad výpočtu metody nejmenších čtverců

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

$A^TA\hat{x} = A^Tb$

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

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

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

4.2 Příklad výpočtu metody nejmenších čtverců

V rovině jsou dány body $(0, -\frac{1}{2})^T, (1, \frac{1}{3})^T$ a $(2, \frac{2}{3})^T$. Pomocí metody nejmenších čtverců proložme těmito body přímku o rovnici $y = kx + q$, kde $k, q \in \mathbb{R}$.

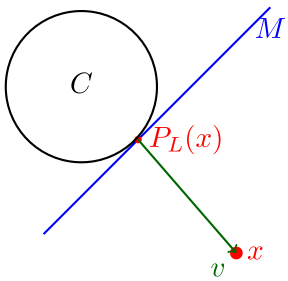
$$\left.\begin{aligned} 0k + q &= -\tfrac{1}{2} \\ 1k + q &= \tfrac{1}{3} \\ 2k + q &= \tfrac{2}{3} \end{aligned}\right\} A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}, b = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}$$

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

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

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

4.3 Věta o oddělitelnosti bodu a konvexní množiny



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

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

Důkaz.
$v = x - P_L(x) \neq 0$

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

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

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

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

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

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

Důkaz sporem.

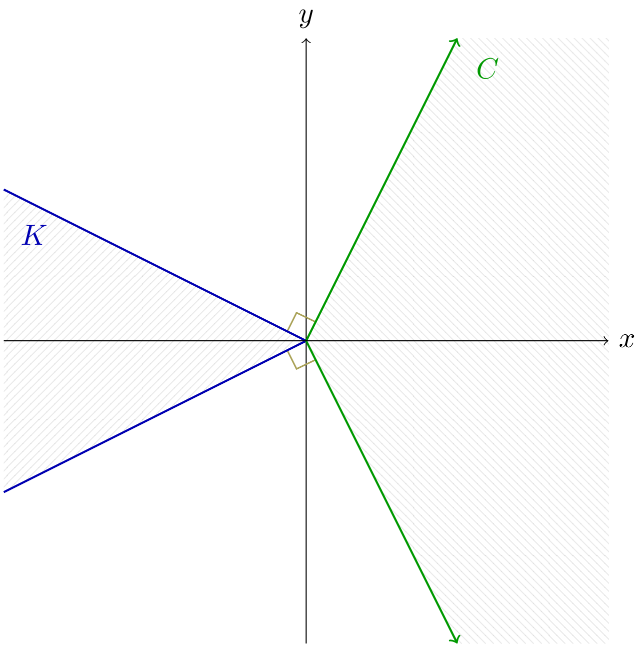
Ať neplatí: tj. existuje $C \in \mathbb{R}^n$ uzavřená konvexní množina tak, že není průnikem P všech poloprostorů obsahujících C.

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

4.4 Příklad na použití věty o oddělitelnosti nadrovinou

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

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



Vždy nastane jeden z případů:

(a) $b \in C$

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

4.5 Lemma neprázdné uzavřené konvexní

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

Důkaz.

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

4.6 Farkasovo lemma

Výslovnost [farkášovo].

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

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

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

Důkaz.

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

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

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

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

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

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

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

Víme tedy, že:

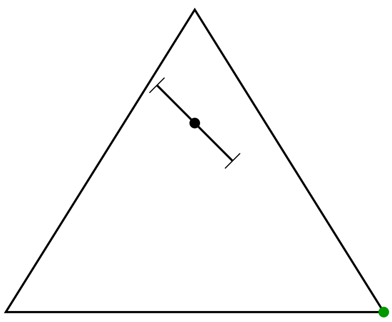
$$\begin{aligned} \langle Ax, y \rangle &\leq \alpha, \quad \forall x \in \mathbb{R}^n_+ \\ \langle x, A^T y \rangle &\leq \alpha, \quad \forall x \in \mathbb{R}^n_+ \end{aligned}$$

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

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

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

4.7 Krajní body konvexní množiny



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

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

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

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

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

4.8 Kreinova-Milmanova věta

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

Kompaktnost je důležitá.

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

4.9 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)

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

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

(b)

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

$$= \sum_{i,j=1}^n a_{ij}\delta_{ik}x_j + a_{ij}\delta_{jk}x_i = \underbrace{\sum_{j=1}^n a_{kj}x_j}_{(Ax)_k} + \underbrace{\sum_{i=1}^n a_{ik}x_i}_{(A^Tx)_k}$$

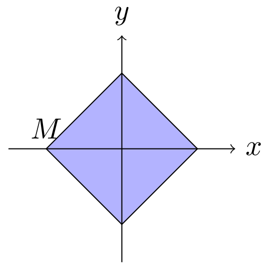
$\implies \nabla f(x) = Ax + A^T x$ (Speciálně: $\nabla f(x) = 2Ax$ pro $A = A^T$)

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

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

4.10 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

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

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

M je konvexní.

2. způsob - úvaha nad vlastnostmi

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

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

$$\begin{aligned} x \in \ker(A^T A) \Rightarrow A^T Ax = 0 \Rightarrow 0 &= \langle A^T Ax, x \rangle \\ &= \langle Ax, Ax \rangle \\ &= \|Ax\|^2 \Rightarrow Ax = 0 \Rightarrow x \in \ker(A) \quad \square \end{aligned}$$

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).

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

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

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

(b)

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

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

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

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

4.13 Formulace úlohy MNČ

Ať závislost výstupního signálu $(y_n)_{n=0}^\infty$ systému na vstupním signálu $(x_n)_{n=0}^\infty$ je dána konvolucí posloupnosti $(x_n)_{n=0}^\infty$ s posloupností $(h_n)_{n=0}^\infty$ ($(h_n)_{n=0}^\infty$ 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 $y_0, \ldots, y_{20}$ výstupního signálu, když na vstupu byl signál s počátečními koeficienty $x_0, \ldots, x_{20}$. Formulujte úlohu nejmenších čtverců pro nalezení koeficientů h_0, h_1, h_2, h_3.

$$(x_n)_{n=0}^\infty \longrightarrow \boxed{(h_n)_{n=0}^\infty} \longrightarrow (y_n)_{n=0}^\infty$$

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

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

$$\begin{aligned}
y_0 &= h_0 x_0 \\
y_1 &= h_1 x_0 + h_0 x_1 \\
y_2 &= h_2 x_0 + h_1 x_1 + h_0 x_2 \\
y_3 &= h_3 x_0 + h_2 x_1 + h_1 x_2 + h_0 x_3 \\
y_4 &= h_3 x_1 + h_2 x_2 + h_3 x_3 + h_0 x_4 \\
&\vdots \\
y_{20} &= h_3 x_{17} + h_2 x_{18} + h_1 x_{19} + h_0 x_{20}
\end{aligned}$$

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

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

5 Konvexní funkce

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

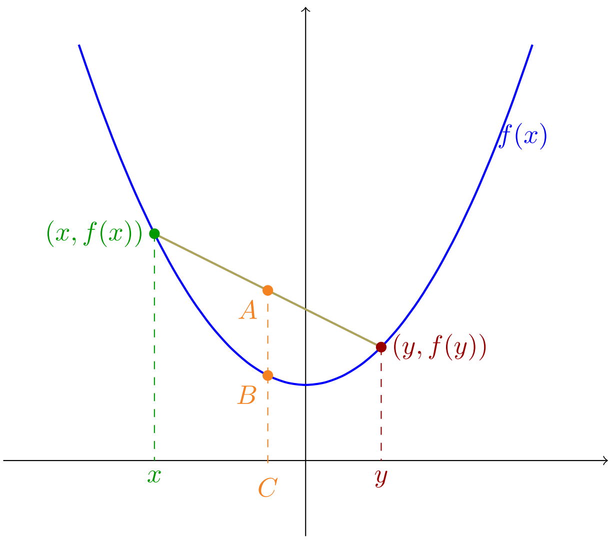
(a) konvexní na C, jestliže pro každé $x, y \in C$ a každé $\lambda \in [0, 1]$ je

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

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

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

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



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

Pozorování: úsečka vždy leží nad funkcí.

5.1 Příklad konvexní funkce

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

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

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

5.2 Příklad konvexní funkce

Je funkce $f(x) = \|x\|$ konvexní?

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

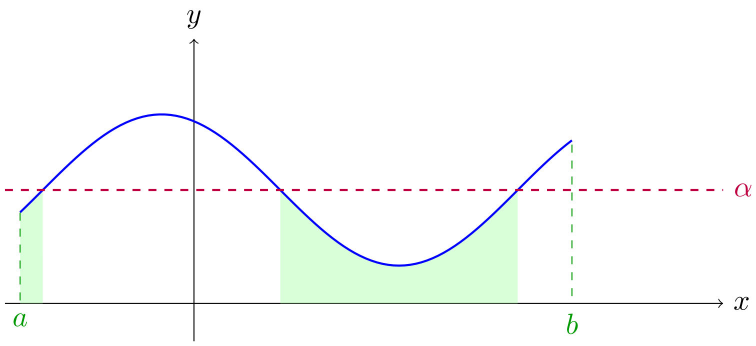
$$\begin{aligned} f(\lambda x + (1-\lambda)y) = \|\lambda x + (1-\lambda)y\| &\overset{\text{odhad}}{\leq} \|\lambda x\| + \|(1-\lambda)y\| = \lambda\|x\| + (1-\lambda)\|y\| \\ &= \lambda f(x) + (1-\lambda)f(y) \implies f \text{ je konvexní.} \quad \square \end{aligned}$$

5.3 Dolní úrovňová množina

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

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

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



Důkaz.

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

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

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

Poznámka.

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

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

5.4 Použití dolní úrovňové množiny

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

Důkaz.

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

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

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

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

5.5 Součet a součin zachovávají konvexitu

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

(a) $f + g$ je konvexní na C

(b) αf je konvexní na C

Důkaz.

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

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

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

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

5.6 Příklad ověření konvexity

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

Rozeberme si jednotlivé části funkce.

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

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

5.7 Skládání zachovává konvexitu

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

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

1. Mějme tedy tvrzení.

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

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

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

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

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

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

5.8 Věta o extrémech konvexních funkcí

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

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

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

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

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

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

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

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

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

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

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

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

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

5.9 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 $x, y \in C$, $\lambda \in (0, 1]$.

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

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

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

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

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

Pronásobením a sečtením dostaneme:

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

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

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

5.10 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 $\nabla^2 f(x)$ positivně semidefinitní matice, pak f je konvexní na C.

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

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

Důkaz (a).

Ať $x, y \in C$.

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

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

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

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

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

Ať $x \in C, y \in \mathbb{R}^n$.

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

Taylorův polynom:

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

kde w má nulovou limitu v 0.

Použijme fakt, že f je konvexní:

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

Když tedy dosadíme:

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

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

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

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

Poznámka. Nutnost otevřenosti C je velmi důležitá!

Důkaz (c). Podobně jako (a).

5.11 Příklad ověření konvexnosti pomocí derivace

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

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

5.12 Příklad ověření konvexnosti pomocí derivace

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

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

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

5.13 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 α je funkce f konvexní?

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

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

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:

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

Vyškrtněme 1. řádek a 1. sloupec:

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

A teď zbylé minoru po vyškrtání dvou řádků a sloupců:

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

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

5.14 Příklad ověření konvexity množiny

Mějme množinu

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

Je M konvexní?

Označme: $g_1(x, y) = x + 2e^{-x+y^2} \ldots M_1 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \;\middle|\; g_1(x, y) \leq 4 \right\}$

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

$M = M_1 \cap M_2 \implies$ 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 $\to$ 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.

6 Podmínky optimality

6.1 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 $x \in \mathrm{int}(M)$, pak $\mathcal{F}(M; x) = \mathbb{R}^n$.
Je-li M konečná (neprázdná), pak $\mathcal{F}(M; x) = \{0\}$ pro každé $x \in M$.

6.2 Přípustné směry poklesu

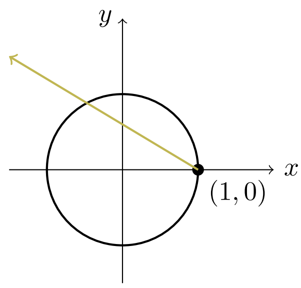
Mějme

(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

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

(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 $d \neq 0 \in \mathbb{R}^2$

$$1 = \|x + \alpha d\|^2 = \langle x + \alpha d, x + \alpha d \rangle = \underbrace{\|x\|^2}_{1} + 2\alpha \langle x, d \rangle + \alpha^2 \|d\|^2$$

$$\to 0 = \alpha(2\langle x, d \rangle + \alpha \|d\|^2) \implies \alpha = -\frac{2\langle x, d \rangle}{\|d\|^2} \implies \mathcal{F}(M; x) = \{0\}$$

(b) Uvažujme kouli

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

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

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

6.3 Kužel směrů poklesu

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

- 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 $[x, x + \delta d] \subseteq D$.

6.4 Nutná geometrická podmínka lokálního extrému

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

Důkaz. Sporem.
Ať ne, tj. existuje $d \in \mathcal{F}(M, x) \cap D(f, x)$.
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. $\square$

6.5 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.

6.6 Tvrzení 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:

(a) Je-li $d \in \mathcal{D}(f; x)$, potom $\langle \nabla f(x), d \rangle \leq 0$.

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

Důkaz.
(a) Ať $d \in D(f; x)$.

$$\frac{f(x + \alpha d) - f(x)}{\alpha} < 0 \text{ pro } \alpha > 0 \text{ dostatečně malé.}$$

$$\implies \underbrace{\lim_{x \to 0^+} \frac{f(x + \alpha d) - f(x)}{\alpha}}_{= \langle \nabla f(x), d \rangle} \leq 0 \quad \square$$

(b) Ať $\alpha > 0$.

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

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

6.7 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 $\hat{x} \in \operatorname{int}(M)$, pak $\nabla f(\hat{x}) = 0$.

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

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

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

(b)

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

Ať $d = -\nabla f(\hat{x})$.

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

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

(b) Předpokládejme, že $\hat{x} \in \operatorname{int}(C)$. Pak $\hat{x} \in \operatorname{argmin}_{x\in C} f(x)$ 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 $y \in C : \overbrace{f(y) < f(\hat{x})}^{f(y)-f(\hat{x})<0}$.
Ať $d = y - \hat{x} (\neq 0) \in \mathcal{F}(C,\hat{x})$.
Cíl: $d \in \mathcal{F}(C,\hat{x}) \cap \mathcal{D}_0(f,\hat{x})$.

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

f 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.

„$\Leftarrow$“ Ať $\nabla f(\hat{x}) = 0$.
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 $\overset{(a)}{\Longrightarrow}$ $\hat{x} \in \operatorname{argmin}_{x \in C} f(x)$. $\square$

6.9 Hledání bodu minima

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

$\nabla^2 f(x, y) = \begin{bmatrix} 2 & -2 \\ -2 & 6 \end{bmatrix} \dots$ dle Sylvesterova kritéria je positivně definitní.
$\Longrightarrow$ f je nutně (ryze) konvexní.

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

6.10 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 positivně semidefinitní.

(b) Jestliže $\nabla f(\hat{x}) = 0$ a $\nabla^2 f(\hat{x})$ je positivně definitní, pak $\hat{x}$ je bod ostrého lokálního minima.

Důkaz vynecháme.

6.11 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.

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

Podezřelé body jsou:

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

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

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

6.12 Hledání bodu minima

Nalezněte, pokud existují, všechny body minima funkce

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

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

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

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 $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

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

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

- Množina $\mathcal{I}\left((g_i)_{i=1}^k; x\right) := \{i \in \{1,\dots,k\} \mid g_i(x) = 0\}$ se nazývá indexová množina aktivních omezení v bodě x.
- Jestliže $i \in \mathcal{I}\left((g_i)_{i=1}^k; x\right)$, pak $g_i(x) \le 0$ se nazve **aktivní** omezení (ve tvaru nerovnosti) v bodě x.
- Jestliže $i \notin \mathcal{I}\left((g_i)_{i=1}^k; x\right)$, pak $g_i(x) \le 0$ 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 $\left((g_i)_{i=1}^k; x\right)$, 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 $M = \{x \in \Omega \mid g_1(x) \le 0, \dots, g_k(x) \le 0\}$. Definujme množinu

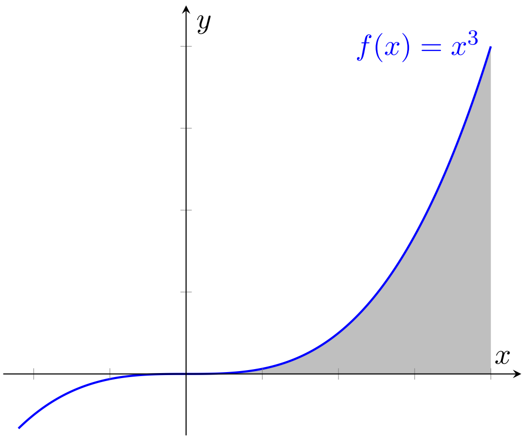
$$\begin{aligned} \mathcal{G}\left((g_i)_{i=1}^k; x\right) &:= \{d \in \mathbb{R}^n \mid \langle \nabla g_i(x), d\rangle \le 0 \quad \forall i \in \mathcal{I}(x)\} \\ &= \bigcap_{i \in \mathcal{I}(x)} \{d \in \mathbb{R}^n \mid \langle \nabla g_i(x), d\rangle \le 0\} \end{aligned}$$

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

6.14 Příklad výpočtu $\mathcal{G}$ a $\mathcal{F}$

Je dána množina
$$M = \left\{(x,y)^T \in \mathbb{R}^2 \mid y - x^3 \leq 0, -y \leq 0\right\}$$
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.

$$\alpha d_2 - \alpha^3 d_1^3 \leq 0 \qquad (3)$$
$$-\alpha d_2 \leq 0 \quad \forall \alpha > 0 \text{ dostatečně malé.} \qquad (4)$$

$(4) \implies d_2 \geq 0$

$(3) \implies d_2 \leq \alpha^2 d_1^3 \quad \forall \alpha > 0$ 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 libovolné malé, tedy i limitně blízké nule, α.

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

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

Označme si $g_1(x,y) = y - x^3$ a $g_2(x,y) = -y$.
Pak:
$$\left.\begin{aligned} \langle \nabla g_1(0), d\rangle &= \left\langle \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \right\rangle = d_2 \leq 0 \\ \langle \nabla g_1(0), d\rangle &= \left\langle \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \right\rangle = d_2 \geq 0 \end{aligned}\right\} d_2 = 0$$

$\implies \mathcal{G}((g_1,g_2),(0,0)) = \left\{ \begin{bmatrix} d_1 \\ 0 \end{bmatrix} \,\middle|\, d_1 \in \mathbb{R} \right\} \to \mathcal{G}$ 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.

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

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

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

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

6.15 Ukázka, že aproximací $\mathcal{F}$ lze zkazit prázdnost průniku

Mějme optimalisační úlohu

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

Pak

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

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

7 KKT podmínky

7.1 Věta o nutných KKT podmínkách

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

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

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

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

Důkaz.

- $\mathcal{I}(\hat{x}) = \emptyset \implies \hat{x} \in \operatorname{int}(M) \implies \nabla f(\hat{x}) = 0$ z Fermatovy věty.
  $\to$ 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 $(\hat{x}) \overset{\text{Fermatova}}{\underset{\text{věta}}{\implies}} \mathcal{F}(M;\hat{x}) \cap \mathcal{D}_0(f;\hat{x}) = \emptyset$,
  tj. $\langle \nabla f(\hat{x}), d \rangle \ge 0 \quad \forall d \in \mathcal{F}(M;\hat{x})$.

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

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

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

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

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

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

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

  A to jsou přesně KKT podmínky. $\square$

7.2 Příklad použití KKT podmínek

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

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:

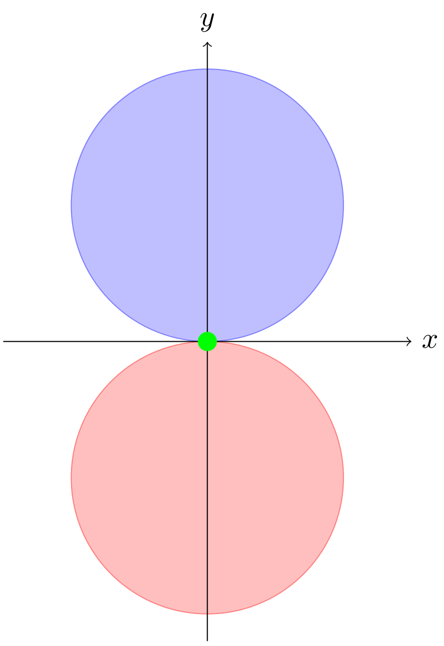
$$
\begin{aligned}
1 + \mu_1(-1) + \mu_2(0) &= 0 \qquad \rightarrow \mu_1 = 1 \\
1 + \mu_1(0) + \mu_2(-1) &= 0 \qquad \rightarrow \mu_2 = 1 \\
\mu_1(-x) &= 0 \\
\mu_2(-y) &= 0 \\
\mu_1, \mu_2 &\geq 0
\end{aligned}
$$

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

7.3 Příklad, že KKT podmínky vždy nenaleznou všechny body

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

Nákres.



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

KKT podmínky:

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

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

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

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

$$f(x) - f(\hat{x}) \underset{\substack{f \text{ je konvexní} \\ \text{na } C \subseteq \Omega}}{\geq} \langle \nabla f(\hat{x}), x - \hat{x} \rangle \overset{\text{podmínka}}{\underset{\text{stacionarity}}{=}} \langle -\sum_{i=1}^{k} \mu_i \nabla g_i(\hat{x}), x - \hat{x} \rangle$$

$$= \sum_{i=1}^{k} -\langle \nabla g_i(\hat{x}), x - \hat{x} \rangle \mu_i = \sum_{i=1}^{n} (g_i(\hat{x}) - g_i(x)) \mu_i \overset{\text{podmínka}}{\underset{\text{komplementarity}}{=}} -\sum_{i=1}^{n} \underbrace{\mu_i}_{\substack{\text{podmínka} \\ \text{nezápornosti}}} \overbrace{g_i(x)}^{\leq 0\, \forall x \in C} \geq 0. \quad \square$$

7.5 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.

7.6 Slaterova podmínka regularity

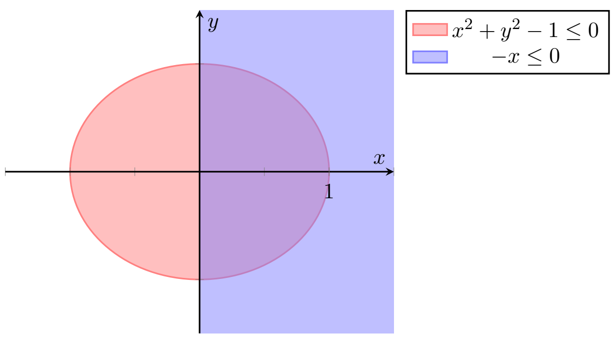
Nechť $\Omega \subseteq \mathbb{R}^n$ je otevřená množina, $g_1, \dots, g_k \in C^1(\Omega)$ a

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

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

7.7 Použití podmínek regularity k ověření KKT podmínek

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



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.

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

$\implies$ KKT podmínky:

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

$y = 0$:

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

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

7.8 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 nutnné a postačující podmínky optimality?

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

$\implies f(x) = \left\langle \left(A^T A + \lambda D^T D\right) x, x \right\rangle - 2 \left\langle x, A^T b \right\rangle + \|b\|^2$
Je f konvexní?
Ano, neboť $\nabla^2 f(x) = 2(A^T A + \lambda D^T D)$ je positivně semidefinitní, protože pro $x \in \mathbb{R}$:

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

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

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

$$\implies A^T b = (A^T A + \lambda D^T D)x$$

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

7.9 Určení KKT podmínek

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

(a) Napište KKT podmínky.

(b) Jsou nutné a postačující?

(c) Ukažte, že $(3,3)^T$ je jediný bod minima.

(a) Mějme

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

$\rightarrow$ použijeme afinní podmínku regularity $\rightarrow$ g_i jsou affiní.

KKT podmínky:

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

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

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

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

$\implies$ $h(x,y)$ 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.

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

$$\begin{aligned} 4 \cdot 27 + 24 \cdot 3 - 4 - \mu_1 + \mu_2 - \mu_3 &= 0 \quad \text{(I.)} \\ 4 \cdot +12 \cdot 3 - 4 - \mu_1 - 2\mu_2 - \mu_4 &= 0 \quad \text{(II.)} \\ \mu_1 \cdot 0 &= 0 \\ \mu_2 \cdot 0 &= 0 \\ 3\mu_3 &= 0 \implies \mu_3 = 0 \\ 3\mu_4 &= 0 \implies \mu_4 = 0 \\ \mu_1, \mu_2 &\geq 0 \end{aligned}$$

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

7.10 Určení KKT podmínek

$$\begin{aligned} &\text{minimalisujte } \alpha x + y, \alpha \in \mathbb{R} \text{ je parametr.} \\ &\text{za podmínek } x^2 + y^2 - 25 \leq 0, \\ &\qquad\qquad\qquad x - y - 1 \leq 0. \end{aligned}$$

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

KKT podmínky:

$$\begin{aligned} \alpha + \mu_1(2x) + \mu_2 \cdot 1 &= 0 \\ 1 + \mu_1(2y) - \mu_2 &= 0 \\ \mu_1(x^2 + y^2 - 25) &= 0, \\ \mu_2(x - y - 1) &= 0, \\ \mu_1, \mu_2 &\geq 0. \end{aligned}$$

g_i jsou konvexní, f je konvexní $\implies$ KKT podmínky jsou postačující.
Slaterova podmínka optimality je splněna $\implies$ KKT podmínky jsou nutné.

$x = 4, y = 3$:

$$\begin{aligned} \alpha + 8\mu_1 + \mu_2 &= 0 \quad \text{(I.)} \\ 1 + 6\mu_1 - \mu_2 &= 0, \quad \text{(II.)} \\ \mu_1, \mu_2 &\geq 0. \end{aligned}$$

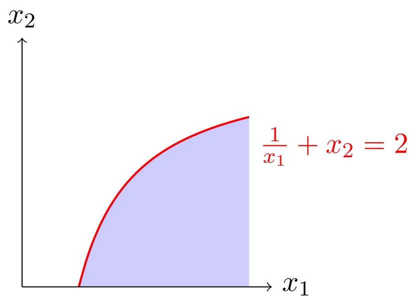
I.+II.: $\alpha + 1 + 14\mu_1 = 0$
$\mu_1 = \frac{-\alpha-1}{14} \overset{!}{\geq} 0 \implies -1 \geq \alpha$. A tedy $\mu_2 = 1 + 6\mu_1 \geq 0$.

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

7.11 Určení KKT podmínek s trikem

Mějme zadání

$$\begin{aligned} \text{minimalisujte } & \frac{x_1}{x_2} \\ \text{za podmínek } & \frac{1}{x_1} + x_2 \leq 2, \\ & x_1, x_2 > 0. \end{aligned}$$



Z nákresu množina vypadá konvexní, co ale minimalisovaná funkce?

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 0 & -\frac{1}{x_2^2} \\ -\frac{1}{x_2^2} & \frac{2x_1}{x_2^3} \end{bmatrix}$$

$$\det \nabla^2 f(x_1, x_2) = -\frac{1}{x_2^4} < 0 \ldots \text{indefinitní}$$

$\implies$ KKT podmínky jsou jen nutné, nikoliv postačující.

Využijeme trik, uděláme substituci: $x_1 = e^{y_1}$, $x_2 = e^{y_2}$ ... $\varphi(y_1, y_2) = (e^{y_1}, e^{y_2})$, $\varphi(\hat{y}_1, \hat{y}_2) = (\hat{x}_1, \hat{x}_2)$. A úlohu převedeme na:

$$\begin{aligned} \text{minimalisujte } & e^{y_1} - e^{y_2} \\ \text{za podmínek } & e^{-y_1} + e^{y_2} \leq 2. \end{aligned}$$

$$e^{\hat{y}_1 - \hat{y}_2} \leq e^{y_1 - y_2}$$

$$\underbrace{\frac{e^{\hat{y}_1}}{e^{\hat{y}_2}}}_{f(\varphi(\hat{y}_1, \hat{y}_2))} \leq \underbrace{\frac{e^{y_1}}{e^{y_2}}}_{f(\varphi(y_1, y_2))}$$

$$f(\hat{x}_1, \hat{x}_2) \leq f(x_1, x_2) \quad \forall (x_1, x_2) \in M.$$

Slaterova podmínka je splněna $\rightarrow (y_1, y_2) = (1, 0)$.
$\implies$ KKT podmínky jsou nutné a postačující.

$$e^{y_1 - y_2} + \mu(-e^{-y_1}) = 0 \tag{I}$$

$$-e^{y_1 - y_2} + \mu e^{y_2} = 0 \rightarrow \mu = \frac{e^{y_1 - y_2}}{e^{y_2}} = e^{y_1 - 2y_2} \tag{II}$$

$$\mu(e^{-y_1} + e^{y_2} - 2) = 0 \tag{III}$$

$$\mu \geq 0 \tag{IV}$$

Očividně $\mu \neq 0 \implies e^{-y_1} + e^{y_2} - 2 = 0$ (III).

$$\begin{aligned} \text{Dosazení (II) do (I): } e^{y_1 - y_2} - e^{-2y_2} &= 0. \\ e^{y_1 - y_2} &= e^{-2y_2} \\ y_1 - y_2 &= -2y_2 \\ y_1 &= -y_2 \end{aligned}$$

Dosazením do (III) získáme $2e^{y_2} - 2 = 0 \Rightarrow e^{y_2} = 1 \Rightarrow y_2 = 0 = y_1$.
Jediný bod minima je $[0, 0]^T$.

Teď zpětný chod na původní úlohu: $x_1 = e^0 = 1$, $x_2 = e^0 = 1$.

Původní úloha má řešení $[1, 1]^T$.

8 Dualita

8.1 Pomocný důkaz vlastnosti infima

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

Důkaz.
„≥“:

$$\inf_{x\in M} h_1(x) \le h_1(t) \quad \forall t \in M$$
$$\inf_{y\in N} h_2(y) \le h_2(t) \quad \forall t \in N$$
$$\implies \inf_{x\in M} h_1(x) + \inf_{y\in N} h_2(y) \le \inf_{(x,y)^T \in M\times N} (h_1(x) + h_2(y)) \quad \square$$

„≤“:

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

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

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

A to samé lze ukázat i pro h_2:

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

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

Teď sečtěme tyto dvě nerovnice:

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

Tedy stačí dokázat, že

$$\inf_{(x,y)^T \in M\times N} (h_1(x) + h_2(y)) \le -h_1(t) - h_2(s) + 2 \inf_{(x,y)^T \in M\times N} (h_1(x) + h_2(y)) \quad \forall t \in M, \forall s \in N.$$
$$0 \le -h_1(t) - h_2(s) + \inf_{(x,y)^T \in M\times N} (h_1(x) + h_2(y)) \quad \forall t \in M, \forall s \in N.$$
$$h_1(t) + h_2(s) \le \inf_{(x,y)^T \in M\times N} (h_1(x) + h_2(y)) \quad \forall t \in M, \forall s \in N.$$

8.2 Dualita - motivační příklad

Je dána úloha

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

Označme $f(x, y) = 2x + 3y$, $M = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in [0, 2]^2 \,\middle|\, 1 - x - y \le 0 \right\}$ 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

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

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 φ vede na úlohu

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

Kde

$$\begin{aligned} \varphi(\mu) &= \min_{(x,y)^T \in [0,2]^T} [(2-\mu)x + (3-\mu)y + \mu] \\ &= \mu + \min_{x \in [0,2]} (2-\mu)x + \min_{y \in [0,2]} (3-\mu)y \\ &= \begin{cases} \mu & \mu < 2 \\ \mu + 4 - 2\mu & \mu \in [2,3) \\ 10 - 3\mu & \mu \in [3,\infty) \end{cases} \end{aligned}$$

Tu budeme nazývat duální úlohou.

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

8.3 Tvrzení o konkávnosti duální úlohy

Jestliže $D_\varphi \neq \emptyset$, pak φ je konkávní.

Důkaz.
Mějme $\mu, \nu \in D_\varphi$, $\lambda \in [0,1]$.

$$\begin{aligned} \varphi(\lambda\mu + (1-\lambda)\nu) &\stackrel{?}{=} \inf_{x \in \Omega} \overbrace{f(x)}^{\lambda f(x) + (1-\lambda) f(x)} + \overbrace{\langle g(x), \lambda\mu + (1-\lambda)\nu \rangle}^{\lambda \langle g(x),\mu\rangle + (1-\lambda)\langle g(x),\nu\rangle} \\ &= \inf_{x \in \Omega} \lambda(f(x) + \langle g(x), \mu \rangle) + (1-\lambda)(f(x) + \langle g(x), \nu \rangle) \\ &\underset{\text{infima}}{\overset{\text{vlastnost}}{\geq}} \lambda \inf_{x \in \Omega} (f(x) + \langle g(x), \mu \rangle) + (1-\lambda) \inf_{x \in \Omega} (f(x) + \langle g(x), \nu \rangle) \\ &= \lambda\varphi(\mu) + (1-\lambda)\varphi(\nu) > -\infty \implies \lambda\mu + (1-\lambda)\nu \in D_\varphi. \quad \square \end{aligned}$$

8.4 Věta o slabé dualitě

(a) Pro každé $x \in M$ a $\mu \in N$ je $\varphi(\mu) \leq f(x)$.

(b) $\hat{\varphi} \leq \hat{f}$.

(a) Důkaz.
Víme: $L(x, \mu) \leq f(x) \quad \forall x \in M, \forall \mu \geq 0$.

$$\varphi(\mu) = \inf_{y \in \Omega} L(y, \mu) \leq \inf_{y \in M} L(y, \mu) \leq L(x, \mu) \leq f(x) \quad \forall x \in M, \forall \mu \in N. \quad \square$$

(b) Důkaz.

Z (a) máme $\overbrace{\sup_{\mu \in N} \varphi(\mu)}^{\hat{\varphi}} \leq f(x) \quad \forall x \in M$.

$$\implies \hat{\varphi} \leq \inf_{x \in M} f(x) = \hat{f}. \quad \square$$

8.5 Důsledek věty o slabé dualitě

(a) Jestliže existují $\hat{x} \in M$ a $\hat{\mu} \in N$ splňující $\varphi(\hat{\mu}) = f(\hat{x})$, pak

$$\hat{\mu} \in \operatorname*{argmax}_{\mu \in N} \varphi(\mu) \quad \text{a} \quad \hat{x} \in \operatorname*{argmin}_{x \in M} f(x).$$

(b) Je-li $\hat{f} = -\infty$, pak $N = \emptyset$.

(c) Je-li $\hat{\varphi} = +\infty$, pak $M = \emptyset$.

Důkaz (a).
Z věty o slabé dualitě platí:

$$\varphi(\mu) \leq f(\hat{x}) \overset{\text{předpoklad}}{=} \varphi(\hat{\mu}) \quad \forall \mu \in N \iff \hat{\mu} \in \operatorname*{argmax}_{\mu \in N} \varphi(\mu).$$

Analogicky:

$$f(\hat{x}) \overset{\text{předpoklad}}{=} \varphi(\hat{\mu}) \leq f(x) \quad \forall x \in M \iff \hat{x} \in \operatorname*{argmin}_{x \in M} f(x). \quad \square$$

Důkaz (b).
Sporem. Ať $N \neq \emptyset$. Volme $\mu \in N$.

$$\text{Pak } \underbrace{\varphi(\mu)}_{\in \mathbb{R}} \leq \hat{\varphi} \leq \hat{f} = -\infty \ldots \text{ spor.} \quad \square$$

Důkaz (C).
Sporem. Ať $M \neq \emptyset$. Volme $x \in M, \mu \in N$.

$$\text{Pak } \varphi(\mu) \leq \hat{\varphi} = +\infty \leq \underbrace{f(x)}_{\in \mathbb{R}} \ldots \text{ spor.} \quad \square$$

8.6 Ukázkový příklad na slabou dualitu

Je dána úloha

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

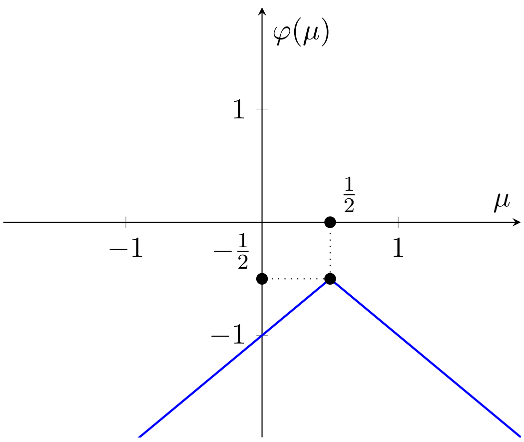
Tedy:

$$L(x,\mu) = -x^2 + \mu(2x-1) = (-x^2 + 2x\mu) - \mu$$

$$\varphi(\mu) = \left[\min_{x\in[0,1]} (-x^2 + 2x\mu)\right] - \mu$$

Pozorování. Minimalisovaná funkce je (ryze) konkávní. Nemůže tedy v žádném vnitřním bodě nabývat minima. Dosazení krajních bodů intervalu:

$$\varphi(\mu) = \min\{0, 2\mu-1\} - \mu = \begin{cases} \mu - 1 & \text{pro } \mu < \frac{1}{2}, \\ -\mu & \text{pro } \mu \geq \frac{1}{2}. \end{cases}$$



Z grafu vyčteme: $\hat{\varphi} = -\frac{1}{2}$. A to samé uděláme pro f, kde výsledek bude $\hat{f} = -\frac{1}{4}$.

Tedy $\hat{\varphi} < \hat{f}$.

8.7 Věta o silné dualitě

Nechť $\hat{f} < \infty$ a cílová funkce $f : \mathbb{R}^n \to \mathbb{R}$ je konvexní. Předpokládejme, že platí alespoň jedna z následujících podmínek:

(a) Komponenty $g_1, \ldots, g_k$ zobrazení g splňují Slaterovu podmínku regularity.

(b) Zobrazení g je afinní a Ω je konvexní polyedrická množina.

Potom $\hat{f} = \hat{\varphi}$. Je-li navíc $\hat{f} \in \mathbb{R}$, pak existuje řešení úlohy (D).

Důkaz vynecháme.

9 Lineární programování

Úlohy lineárního programování jsou optimalisační úlohy, ve kterých je

(a) cílová funkce afinní (bez újmy na obecnosti se můžeme omezit na lineární funkce)

(b) přípustná množina je konvexní polyedrická množina (tj. lze popsat pomocí konečné soustavy lineárních rovnic a nerovnic)

Příklad.

Firma vyrábí 2 druhy výrobků A a B. V tabulce je uvedeno množství materiálu (ve vhodných jednotkách) potřebný k výrobě jednotkového množství daného druhu výrobku a také jeho prodejní cena.

	Materiál X	Materiál Y	Cena
Výrobek A	2	3	6000 Kč
Výrobek B	4	4	10000 Kč

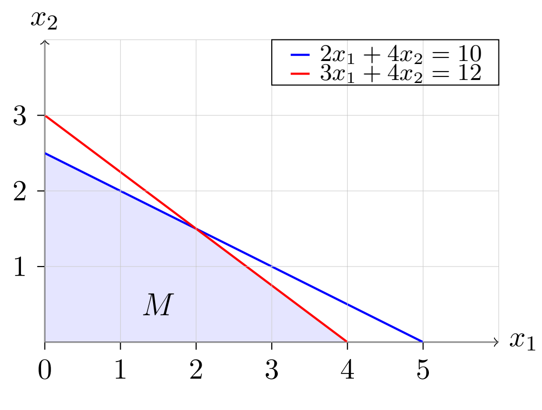
Na skladu je jen 10 jednotek materiálu X a 12 jednotek materiálu Y. Jak mají ve firmě nastavit výrobni proces, aby celková cena za vyrobené množství výrobků byla co největší?

Odpověď je přímo v zadání.

$x_1 \dots$ množství výrobku A

$x_2 \dots$ množství výrobku B

$$\begin{aligned} \text{maximalisujte } & 6x_1 + 10x_2 \\ \text{za podmínek } & 2x_1 + 4x_2 \leq 10, \\ & 3x_1 + 4x_2 \leq 12, \\ & x_1, x_2 \geq 0. \end{aligned}$$



Graficky lze nalézt, že maximum se nabývá v bodě $(2, \frac{3}{2})^T$. Maximum je $f(2, \frac{3}{2}) = 27$.

Pokračování příkladu.

Obchodník chce od firmy koupit veškerý materiál ze skladu. Jaké ceny za materiál X a Y by měl firmě nabídnout, aby zaplatil co nejmenší částku a firmě se přesto vyplatilo materiál prodat namísto výroby výrobků?

Tato otázka vede na úlohu:

$y_1 \ldots$ cena za jednotkové množství materiálu X
$y_2 \ldots$ cena za jednotkové množství materiálu Y

$$\begin{aligned} \text{minimalisujte } & 10y_1 + 12y_2 \\ \text{za podmínek } & 2y_1 + 3y_2 > 6, \\ & 4y_1 + 4y_2 > 10, \\ & y_1, y_2 \geq 0. \end{aligned}$$

Pozorování. Tyto dvě úlohy jsou navzájem duální.

9.1 Zápis úlohy lineárního programování

Je dána úloha

$$\begin{aligned} \text{minimalisujte } & x_1 - x_2 \\ \text{za podmínek } & 2x_1 - 3x_2 = 5, \\ & -2 \leq x_2 \leq 3, \\ & x_1 \leq 0. \end{aligned}$$

Zapišme úlohu v kanonickém tvaru.
Pomocné substituce: $y_1 = -x_1$, $x_2 = y_2 - y_3$, $y_2, y_3 \geq 0$.

$$\begin{aligned} \text{minimalisujte } & -y_1 - y_2 + y_3 \\ \text{za podmínek } & -2y_1 - 3y_2 + 3y_3 \geq 5, \\ & 2y_1 + 3y_2 - 3y_3 \geq -5, \\ & -y_2 + y_3 \geq -3, \\ & y_2 - y_3 \geq -2, \\ & y_1, y_2, y_3 \geq 0. \end{aligned}$$

Zapišme úlohu ve standardním tvaru.

$$\begin{aligned} \text{minimalisujte } & -y_1 - y_2 + y_3 \\ \text{za podmínek } & -2y_1 - 3y_2 + 3y_3 = 5, \\ & y_2 - y_3 - y_4 = -2, \\ & y_2 - y_3 + y_5 = 3, \\ & y_1, y_2, y_3, y_4, y_5 \geq 0. \end{aligned}$$

9.2 Basický přípustný bod

Bod $x \in M$ se nazve basický přípustný bod (BPB) úlohy lineárního programování, pokud existuje m-prvková množina $B \subseteq \{1, \ldots, n\}$ taková, že

(a) A_B je regulární,

(b) $x_j = 0$ pro každé $j \in \mathbb{N}$.

Množina B z definice BPB se nazývá přípustná báse.

9.3 Příklad BPB

Nechť $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \end{bmatrix}$ a $b = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$. Jaké jsou BPB?

- $B = \{1,2\} \dots A_B = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$. Evidentně invertibilní.

  $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \dots \underbrace{Ax}_{A_B x_B + A_N \underbrace{x_N}_{=0}} = b$. Tedy $\left[\begin{array}{cc|c} 1 & 2 & 5 \\ 0 & 2 & 3 \end{array}\right] \to x = \frac{1}{2}\begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} \in M$ je BPB.

- $B = \{1,3\} \dots A_B = \begin{bmatrix} 1 & 3 \\ 0 & 3 \end{bmatrix}$. Evidentně invertibilní.

  Tedy $\left[\begin{array}{cc|c} 1 & 3 & 5 \\ 0 & 3 & 3 \end{array}\right] \to x = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \in M$ je BPB.

- $B = \{2,3\} \dots A_B = \begin{bmatrix} 2 & 3 \\ 2 & 3 \end{bmatrix}$. Evidentně není regulární. Žádný bod nemůže být BPB.

9.4 Tvrzení o charakterisaci BPB

Nechť $x \in M$. Pak x je BPB právě tehdy, když $\{a_j \mid j \in J(x)\}$ je lineárně nezávislá množina.

Důkaz.
„$\Rightarrow$“: x je BPB $\implies$ existuje $B \subseteq \{1, \dots, n\}$ m-prvková tak, že $\{a_j \mid j \in B\}$ je lineárně nezávislá. Navíc $J(x) \subseteq B$, protože $J(x)$ obsahuje ty indexy, které odpovídají kladným komponentám a všechny komponenty indexované mimo indexy z B jsou nulové.

Tedy $\{a_j \mid j \in J(x)\}$ je lineárně nezávislá.

„$\Leftarrow$“: Je-li $|J(x) = m|$, pak jasné $(B = J(x))$.
Ať $|J(x)| < m$. Z předpokladu víme $\text{rank}(A) = m$. Pak lze $J(x)$ doplnit do m-prvkové množiny $B \subseteq \{1, \dots, n\}$ tak, že $\{a_j \mid j \in B\}$ je lineárně nezávislá. $\implies$ x je BPB. $\square$

9.5 Tvrzení, že dva různé PBP musí mít různé množiny B

Pro každou m-prvkovou množinu $B \subseteq \{1, \dots, n\}$ takovou, že A_B je regulární, existuje nejvýše jedno $x \in M$ splňující $x_j = 0$ pro každé $j \in N$.

Důkaz. Sporem.

Ať $x, y \in M$ jsou různé a splňují $x_j = y_j = 0$ pro každé $j \in N$.

$$\left.\begin{array}{l} b = Ax = \sum_{j=1}^{n} x_j a_j = \sum_{j \in B} x_j a_j = A_B x_B \\ b = Ay = A_B y_B \end{array}\right\} A_B x_B = A_B y_B$$

A protože A je dle předpokladu regulární, tak dostaneme:

$$x_B = y_B \implies x = y$$

Což je ale spor, protože x a y mají být různé. $\square$

Horní hranice počtu BPB úlohy LP je tedy $\binom{n}{m}$.

9.6 Příklad na degenerované BPB

Nechť
$$A = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{bmatrix} \quad \text{a} \quad b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$
Určete všechny basické přípustné body.

- $B = \{1, 2\}$. Tedy $A_B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ je určitě regulární.
  Pak očividně $\left[1, 0, 0, 0\right]^T$ je BPB s přípustnou básí B.
- $B = \{1, 3\}$. Tedy $A_B = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ je určitě regulární.
  Pak $\left[1, 0, 0, 0\right]^T$ je BPB s přípustnou básí B.
- $B = \{1, 4\}$. $A_B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ je singulární, tedy není přípustnou básí BPB.
- $B = \{2, 3\}$. Tedy $A_B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ je určitě regulární.
  Pak $\left[0, 1, 1, 0\right]^T$ je BPB s přípustnou básí B.
- $B = \{2, 4\}$. Tedy $A_B = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ je určitě regulární.
  Pak $\left[0, 0, 0, 1\right]^T$ je BPB s přípustnou básí B.
- $B = \{3, 4\}$. Tedy $A_B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ je určitě regulární.
  Pak $\left[0, 0, 0, 1\right]^T$ je BPB s přípustnou básí B.

9.7 Příklad na souvislost BPB a krajních bodů

Nechť
$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \end{bmatrix} \quad \text{a} \quad b = \begin{bmatrix} 5 \\ 3 \end{bmatrix}.$$
$M = \left\{x \in \mathbb{R}^3_+ \mid Ax = b\right\}$.

Již víme, že $x = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$ a $y = \begin{bmatrix} 2 \\ \frac{3}{2} \\ 0 \end{bmatrix}$ jsou BPB.

$$Ax = b \ldots \left[\begin{array}{ccc|c} 1 & 2 & 3 & 5 \\ 0 & 2 & 3 & 3 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 2 & 3 & 3 \end{array}\right]$$

Tedy řešení soustavy $z = \begin{bmatrix} 2 \\ \frac{1}{2}(3 - 3t) \\ t \end{bmatrix}, t \in \mathbb{R}$. Kdy je $z \in \mathbb{R}^3_+$?

Právě tehdy, když $t \geq 0$ a $1 \geq t$, tedy $t \in [0, 1]$.

$$z \in [x, y] \iff z = tx + (1 - t)y = \begin{bmatrix} 2 \\ \frac{3}{2} \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ -\frac{3}{2} \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ \frac{3}{2}(1 - t) \\ t \end{bmatrix}, t \in [0, 1].$$

Tedy $M = [x, y]$.

9.8 Věta o souvislosti BPB a krajních bodů

(a) Nechť $x \in M$. Pak x je BPB úlohy LP právě tehdy, když x je krajní bod množiny M.

(b) M je neprázdná právě tehdy, když existuje BPB úlohy LP.

Důkaz (a).

„$\Rightarrow$“: Sporem.

Existují dva různé body $y, z \in M$ tak, že $x = \frac{y+z}{2}$. Ať B je přípustná báse BPB x.

Pak $y_j = z_j = 0$ pro každé $j \in N$. Navíc A_B je regulární dle definice BPB. Ale dle této stejné definice platí, že y a z jsou BPB s přípustnou básí B. Ale dle tvrzení nemohou mít dva různé BPB stejnou přípustnou bási. $\implies y = z$, což je spor.

„$\Leftarrow$“: Sporem.

Ať x není BPB. Pak z charakterisace BPB plyne, že $\{a_j \mid j \in J(x)\}$ je lineárně závislá množina.

$\to$ existují $d_j \in \mathbb{R}, j \in J(x)$, ne všechny nulové tak, že

$$\sum_{j \in J(x)} d_j a_j = 0.$$

Definujme $d_j = 0$ pro každé $j \in \{1, \ldots, n\} \setminus J(x)$.

Pak $Ad = 0$. Odtud $A(x \pm \alpha d) = b \pm \alpha \underbrace{Ad}_{=0} = b$ pro všechna $\alpha \in \mathbb{R}$.

Pro dostatečně malé $\alpha > 0$ je $x \pm \alpha d \geq 0$. Pro takové α je $x \pm \alpha d \in M$. Pak $x + \alpha d \neq x - \alpha d$ a navíc evidentně platí $x = \frac{(x+\alpha d)+(x-\alpha d)}{2}$. To je spor s tvrzením, že máme krajní bod. $\square$

Důkaz (b). Vynecháme.

9.9 Základní věta lineárního programování

(a) Úloha LP má řešení právě tehdy, když M je neprázdná a $\langle x, c\rangle$ je zdola omezená na M.

(b) Má-li LP řešení, pak existuje řešení úlohy LP, které je BPB.

Důkaz (a).
„$\Rightarrow$“: Když máme řešení, pak určitě leží v M a je určitě zdola omezená, protože to je právě to ono řešení.

„$\Leftarrow$“: Weierstrassova věta. $\square$

Důkaz (b). Ať $\hat{x} \in \operatorname{argmax}_{x \in M} \langle x, c\rangle$. Protože M je kompaktní a konvexní, tak víme $\hat{x} \in \operatorname{conv}(\operatorname{ext}(M))$.

$\operatorname{ext}(M) \ldots$ konečná množina, tj. $\operatorname{ext}(M) = \{e_1, \ldots, e_k\}$.

$$\underset{\text{obal}}{\overset{\text{konvexní}}{\Longrightarrow}} \hat{x} = \sum_{i=1}^{l} \lambda_i e_i \text{ pro nějaké } \lambda_1, \ldots, \lambda_l \geq 0 \text{ a } \sum \lambda_i = 1.$$

Alespoň jeden krajní bod musí být mezi e_i.
Ať $e_N \in \operatorname{ext}(M)$ splňuje $\langle e_N, c\rangle = \min_{i \in \{1,\ldots,l\}} \langle e_i, c\rangle$.

$$\langle \hat{x}, c\rangle = \sum_{i=1}^{l} \lambda_i \langle e_i, c\rangle \geq \left(\sum_{i=1}^{l} \lambda_i\right) \langle e_N, c\rangle = \langle e_N, c\rangle \implies e_N \in \operatorname*{argmax}_{x \in M} \langle x, c\rangle. \quad \square$$

9.10 Příklad na hledání duální úlohy

Mějme úlohu

$$\begin{aligned} \text{minimalisujte } & x_1 + 2x_2 \\ \text{za podmínek } & x_1 + x_2 \geq 1 \ldots - x_1 - x_2 + 1 \leq 0, \\ & x_1, x_2 \geq 0. \end{aligned}$$

(a) Najděte duální úlohu, jestliže $x_1, x_2 \geq 0$ je přímé omezení.

(b) Najděte duální úlohu, jestliže $x_1, x_2 \in \mathbb{R}$ je přímé omezení.

(a) $L(x_1, x_2, \mu) = x_1 + 2x_2 + \mu(-x_1 - x_2 + 1)$

$$\varphi(\mu) = \inf_{(x_1,x_2)\in\mathbb{R}^2_+} L(x_1, x_2, \mu) = \inf_{(x_1,x_2)\in\mathbb{R}^2_+} (1-\mu)x_1 + \mu$$

$$\varphi(\mu) = \underbrace{\left[\inf_{x_1\in\mathbb{R}_+} (1-\mu)x_1\right]}_{=\begin{cases} 0 & \text{pro } 1 \geq \mu, \\ -\infty & \text{pro } 1 < \mu. \end{cases}} + \underbrace{\left[\inf_{x_2\in\mathbb{R}_+} (2-\mu)x_2\right]}_{=\begin{cases} 0 & \text{pro } 2 \geq \mu, \\ -\infty & \text{pro } 2 < \mu. \end{cases}} + \mu$$

$$\implies \varphi(\mu) = \begin{cases} \mu & \text{pro } \mu \in [0,1], \\ -\infty & \text{pro } \mu \notin [0,1]. \end{cases}$$

A tedy duální úloha je

$$\begin{aligned} \text{maximalisujte } & \mu \\ \text{za podmínek } & \mu \in [0,1]. \end{aligned}$$

(b) $L(x_1, x_2, \mu_1, \mu_2, \mu_3) = x_1 + 2x_2 + \mu_1(-x_1 - x_2 + 1) + \mu_2(-x_1) + \mu_3(-x_2)$

$$\varphi(\mu) = \inf_{(x_1,x_2)\in\mathbb{R}^2} L(x_1, x_2, \mu_1, \mu_2, \mu_3) = \inf_{(x_1,x_2)\in\mathbb{R}^2} (1-\mu_1-\mu_2)x_1 + (2-\mu_1-\mu_3)x_2 + \mu_1$$

$$\varphi(\mu) = \underbrace{\left[\inf_{x_1\in\mathbb{R}} (1-\mu_1-\mu_2)x_1\right]}_{=\begin{cases} 0 & \text{pro } 1-\mu_1-\mu_2 = 0, \\ -\infty & \text{pro } 1-\mu_1-\mu_2 \neq 0. \end{cases}} + \underbrace{\left[\inf_{x_2\in\mathbb{R}} (2-\mu_1-\mu_3)x_2\right]}_{\begin{cases} 0 & \text{pro } 2-\mu_1-\mu_3 = 0, \\ -\infty & \text{pro } 2-\mu_1-\mu_3 \neq 0. \end{cases}} + \mu_1$$

$$\varphi(\mu_1, \mu_2, \mu_3) = \begin{cases} \mu_1 & \text{pro } D_\varphi = \left\{(\mu_1, \mu_2, \mu_3) \in \mathbb{R}^3_+ \;\middle|\; 1-\mu_1-\mu_2 = 0, 2-\mu_1-\mu_3 = 0\right\}, \\ -\infty & \text{jinak.} \end{cases}$$

A tedy duální úloha je

$$\begin{aligned} \text{maximalisujte } & \mu_1 \\ \text{za podmínek } & 1-\mu_1-\mu_2 = 0, \\ & 2-\mu_1-\mu_3 = 0, \\ & \mu_1, \mu_2, \mu_3 \geq 0. \end{aligned}$$

9.11 Příklad na hledání duální úlohy

Mějme úlohu

$$\begin{array}{ll} \text{minimalisujte} & x_1^2 + x_2^2 \\ \text{za podmínek} & -x_1 - x_2 + 4 \leq 0, \\ & x_1, x_2 \qquad\quad \geq 0. \end{array}$$

(a) Najděte duální úlohu, jestliže $x_1, x_2 \geq 0$ je přímé omezení.

(b) Najděte duální úlohu, jestliže $x_1, x_2 \in \mathbb{R}$ je přímé omezení.

9.12 Tvrzení o množině všech řešení úlohy LP

Množina všech řešení úlohy LP je konvexní polyedrická množina.

9.13 Příkad na Simplexovu metodu

Je dána úloha

$$\begin{array}{ll} \text{minimalisujte} & -x_1 - 3x_2 \\ \text{za podmínek} & 2x_1 + 3x_2 + x_3 = 6, \\ & -x_1 + x_2 + x_4 = 1, \\ & x_1, x_2, x_3, x_4 \quad \geq 0. \end{array}$$

$$\begin{array}{c} \\ z \\ x_3 \\ x_4 \end{array} \left[\begin{array}{rrrr|r} x_1 & x_2 & x_3 & x_4 & b \\ \hline -1 & -3 & 0 & 0 & 0 \\ \hline 2 & 3 & 1 & 0 & 6 \\ -1 & 1 & 0 & 1 & 1 \end{array}\right] ; \text{BPB je } \begin{bmatrix} 0 \\ 0 \\ 6 \\ 1 \end{bmatrix}.$$

Vyměníme x_2 a x_4 v BPB.
$x_2 = 1 + x_1 - x_4$
$\rightarrow z = -x_1 - 3(1 + x_1 - x_4) = -4x_1 + 3x_4 - 3$
$\rightarrow x_3 = 6 - 2x_1 - 3(1 + x_1 - x_4) = 3 - 5x_1 + 3x_4$

$$\begin{array}{c} \\ z \\ x_3 \\ x_2 \end{array} \left[\begin{array}{rrrr|r} x_1 & x_2 & x_3 & x_4 & b \\ \hline -4 & 0 & 0 & 3 & -3 \\ \hline -5 & 0 & 0 & 3 & 3 \\ 1 & 0 & 0 & -1 & 1 \end{array}\right] ; \text{BPB je } \begin{bmatrix} 0 \\ 1 \\ 3 \\ 0 \end{bmatrix}.$$

9.14 Tvrzení o vztahu přípustné množiny a hodnoty cílové funkce

Přípustná množina M úlohy LP je neprázdná právě tehdy, když v bodě minima $\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} \in \Omega$ úlohy F_1 tak, že $g(\hat{x}, \hat{y}) = 0$. V tomto případě je $\hat{y} = 0$.

Důkaz.

„$\Rightarrow$“: Ať $\hat{x} \in M$. Pak $v = \begin{bmatrix} \hat{x} \\ 0 \end{bmatrix}$ leží v Ω a $g(v) = 0$ (tj. v je řešení úlohy (F_1) splňující $g(v) = 0$). A to jsme přesně chtěli dokázat. $\square$

„$\Leftarrow$“: Ať $\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix}$ je řešení úlohy (F_1), splňující $g(\hat{x}, \hat{y}) = 0$. Pak $\hat{y} = 0$, a tedy $b = A\hat{x} + \hat{y} = A\hat{x}$. A protože $\hat{x} \geq 0$, tak $\hat{x} \in M$. $\square$

9.15 Příklad dvoufázové Simplexové metody

Je dána úloha

$$\begin{array}{ll} \text{minimalisujte} & x_2 \\ \text{za podmínek} & x_1 = 1, \\ & x_1 - x_2 = 2, \\ & x_1,\ x_2 \geq 0. \end{array}$$

Sloupeček pravých stran je větší roven nule, takže můžeme použít dvoufázovou Simplexovu metodu.

1. fáze:

$$\begin{array}{ll} \text{minimalisujte} & y_1 + y_2 \\ \text{za podmínek} & x_1 + y_1 = 1, \\ & x_1 - x_2 + y_2 = 2, \\ & x_1, x_2, y_1, y_2 \geq 0. \end{array}$$

9.16 Tvrzení o primární a duální úloze

Nechť $A \in \mathbb{M}_{m,n}(\mathbb{R})$, $c \in \mathbb{R}^n$ a $b \in \mathbb{R}^m$. Duální úloha k úloze

$$\left.\begin{array}{ll} \text{minimalizujte} & \langle x, c \rangle \\ \text{za podmínky} & Ax \geq b, \\ & x \geq 0. \end{array}\right\} (P)$$

je

$$\left.\begin{array}{ll} \text{maximalisujte} & \langle y, b \rangle \\ \text{za podmínky} & A^T y \geq c, \\ & y \geq 0. \end{array}\right\} (D)$$

Důkaz.

$$L(x, y) = \langle x, c \rangle + \langle y, b - Ax \rangle = \langle x, c \rangle + \langle y, b \rangle - \underbrace{\langle y, Ax \rangle}_{\langle A^T y, x \rangle} = \langle x, c - A^T y \rangle + \langle y, b \rangle$$

$$\inf_{x \geq 0} L(x, y) = \langle y, b \rangle + \inf_{x \geq 0} \langle x, c - A^T y \rangle = \begin{cases} \langle y, b \rangle & \text{je-li } c - A^T y \geq 0, \\ -\infty & \text{jinak.} \end{cases}$$

Tedy duální úloha k (P) je

$$\begin{array}{ll} \text{maximalisujte} & \langle y, b \rangle \\ \text{za podmínky} & c - A^T y \geq 0, \rightarrow A^T y \leq c \\ & y \geq 0. \end{array}$$

$\square$

9.17 Hledání duální úlohy k duální úloze

Mějme

$$\left.\begin{array}{ll} \text{maximalisujte} & \langle y, b\rangle \\ \text{za podmínky} & A^T y \geq c, \\ & y \geq 0. \end{array}\right\} (D)$$

Přepíšeme:

$$\begin{array}{rl} \text{minimalisujte} & -\langle y, b\rangle \\ \text{za podmínky} & A^T y \leq c, \\ & y \geq 0. \end{array} \quad \text{tj.} \quad \begin{array}{rl} \text{minimalisujte} & \langle y, -b\rangle \\ \text{za podmínky} & (-A^T) y \geq -c, \\ & y \geq 0. \end{array}$$

Duální úloha po tom je:

$$\begin{array}{rl} \text{maximalisujte} & \langle x, -c\rangle \\ \text{za podmínky} & (A^T)^T x \leq -b, \\ & x \geq 0. \end{array} \quad \text{tj.} \quad \begin{array}{rl} \text{minimalisujte} & \langle x, c\rangle \\ \text{za podmínky} & Ax \geq b, \\ & x \geq 0. \end{array}$$

9.18 Věta o silné dualitě pro LP

Úloha (P) má řešení právě tehdy, když má řešení úloha (D). V takovém případě jsou hodnoty obou úloh stejné.

Nechť $\hat{f} < \infty$ a cílová funkce $f : \mathbb{R}^n \to \mathbb{R}$ je konvexní. Předpokládejme, že platí alespoň jedna z následujících podmínek

(a) Komponenty $g_1, \ldots, g_k$ zobrazení g splňují Slaterovu podmínku regularity.

(b) Zobrazení g je afinní a Ω je konvexní polyedrická množina.

Potom $\hat{f} = \hat{\varphi}$. Je-li navíc $\hat{f} \in \mathbb{R}$, pak existuje řešení úlohy (D).

Důkaz vynecháme.

9.19 Simplexová metoda a řešení duální úlohy

Do úlohy (P) zaveďme doplňkové proměnné $y = (y_1, \ldots, y_m)^T$. Tím dostaneme úlohu

$$\left.\begin{array}{ll} \text{minimalizujte} & \langle x, c\rangle \\ \text{za podmínky} & Ax - y \geq b, \\ & x, y \geq 0. \end{array}\right\} (\widetilde{P})$$

Je-li výsledná simplexová tabulka

$$\begin{array}{ccccccc|c} \widetilde{c}_1 & \cdots & \widetilde{c}_n & \widetilde{c}_{n+1} & \cdots & \widetilde{c}_{n+m} & & \cdots \\ \hline \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & & \vdots \end{array}$$

kde $\widetilde{c}_1, \ldots, \widetilde{c}_{m+n} \geq 0$ a sloupce na lévé straně odpovídají postupně proměnným $x_1, \ldots, x_n, y_1, \ldots, y_m$. Pak $\hat{y} = (\widetilde{c}_{n+1}, \ldots, \widetilde{c}_{m+n})^T$ je řešením úlohy (D).

9.20 Příklad řešení duální úlohy

Je dána dvojice vzájemně duálních úloh

$$\begin{array}{rlcrl} \text{minimalisujte} & \langle x, c\rangle & & \text{maximalisujte} & \langle y, b\rangle \\ \text{za podmínky} & Ax \geq b, & \text{tj.} & \text{za podmínky} & A^T y \leq c, \\ & x \geq 0. & & & y \geq 0. \end{array}$$

kde $A = \begin{bmatrix} 0 & 4 & 1 \\ 1 & 1 & 0 \end{bmatrix}$, $b = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $c = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$.

Zaměřme se na primární úlohu. Tedy doplníme doplňkové proměnné dle předchozí poznámky.

$$\text{minimalisujte} \quad \left\langle \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} \right\rangle$$

$$\text{za podmínek} \quad \begin{bmatrix} 0 & 4 & 1 & -1 & 0 \\ 1 & 1 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ y_1 \\ y_2 \end{bmatrix} \geq \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$x_1, \dots, y_2 \geq 0$$

Simplexová tabulka:

	x_1	x_2	x_3	y_1	y_2	
	1	2	4	0	0	0
x_3	0	4	1	-1	0	2
x_1	1	1	0	0	-1	1
	0	-15	0	4	1	-9
x_3	0	4	1	-1	0	2
x_1	1	1	0	0	-1	1

Zafixujme si sloupec x_2, protože má v sobě záporný koeficient. Teď vhodně vybrat řádek → vezmeme pravou stranou a podělíme ji koeficientem, tedy $\frac{2}{4} = \frac{1}{2}$ a $\frac{1}{1} = 1$. Vybereme menší podíl.

	x_1	x_2	x_3	y_1	y_2	
	0	0	$\frac{15}{4}$	$\frac{1}{4}$	1	$-9 + \frac{15}{2}$
x_2	0	1	$\frac{1}{4}$	$-\frac{1}{4}$	0	$\frac{1}{2}$
x_1	1	0	$-\frac{1}{4}$	$\frac{1}{4}$	-1	$\frac{1}{2}$

Protože v levé části prvního řádku není žádný záporný koeficient, algoritmus končí.

Úloha má řešení $\left(\frac{1}{2}, \frac{1}{2}, 0, 0, 0\right)^T$. Tedy řešení původní primární úlohy je $\left(\frac{1}{2}, \frac{1}{2}, 0\right)^T$. A když dosadíme, tak zjistíme, že toto je řešení i duální úlohy.

10 Kvadratické programování

Úlohy kvadratického programování jsou optimalisační úlohy, ve kterých je

(a) cílová funkce f kvadratická, tj.

$$f(x) = \langle Qx, x\rangle + \langle x, c\rangle + d,$$

kde $Q \in \mathbb{M}_n(\mathbb{R})$, $c \in \mathbb{R}^n$, $d \in \mathbb{R}$ (budeme předpokládat, že Q je symetrická a $d = 0$);

(b) přípustná množina je konvexní polyedrická množina.

Úloha kvadratického programování není obecně konvexní!

- Pokud ale minimalisujeme kvadratickou funkci f, ve které je Q positivně semidefitní matice, pak se jedná o konvexní úlohu.

Dále už budeme uvažovat jen úlohu kvadratického programování ve tvaru

$$\left.\begin{array}{ll} \text{minimalisujte} & \frac{1}{2}\langle Qx, x\rangle + \langle x, c\rangle \\ \text{za podmínky} & Ax \leq b, \end{array}\right\} (QP)$$

kde $Q \in \mathbb{M}_n(\mathbb{R})$ je positivně <u>definitní</u>, $A \in \mathbb{M}_{m,n}(\mathbb{R})$, $b \in \mathbb{R}^m$ a $c \in \mathbb{R}^n$.

Poznámka. $\frac{1}{2}$ v zápisu se nám zde zrovna hodí. Samozřejmě lze schovat přímo do matice Q, proto v původní definici není vidět.

Cílová funkce v (QP) je ryze konvexní. Úloha tak má nejvýše jedno řešení. KKT podmínky

$$\begin{aligned} Qx + c + A^T\mu &= 0 \\ \langle Ax - b, \mu\rangle &= 0 \\ \mu &\geq 0 \end{aligned}$$

jsou nutné a postačující.

10.1 Tvrzení o duální úloze kvadratického programování

Duální úloha k úloze (QP) je

$$\left.\begin{array}{ll} \text{maximalisujte} & -\frac{1}{2}\langle By, y\rangle - \langle y, v\rangle - \frac{1}{2}\langle Q^{-1}c, c\rangle \\ \text{za podmínky} & y \geq 0, \end{array}\right\} (DQP)$$

kde $B = AQ^{-1}A^T$ a $v = AQ^{-1}c + b$.

Důkaz.

$$\begin{aligned} L(x, y) &= \frac{1}{2}\langle Qx, x\rangle + \langle x, c\rangle + \langle y, Ax - b\rangle \\ &= \frac{1}{2}\langle Qx, x\rangle + \langle x, c\rangle + \left\langle A^T y, x\right\rangle - \langle y, b\rangle \end{aligned}$$

Ať $y \geq 0$. Pak funkce $x \mapsto L(x, y)$ je určitě (ryze) konvexní díky předpokladu na Q. Tedy $\hat{x}$ je bodem minima funkce $x \mapsto L(x, y) \iff \nabla_x L(x, y) = 0$.

$$\varphi(y) = \inf_{x \in \mathbb{R}^n} L(x, y) = \min_{x \in \mathbb{R}^n} L(x, y) = L(\hat{x}, y)$$

$$\nabla_x L(x,y) = Q\hat{x} + c + A^T y \stackrel{!}{=} 0$$
$$\text{Tedy: } \hat{x} = -Q^{-1}(c + A^T y)$$

Dosaďme:

$$\begin{aligned}
\varphi(y) &= \frac{1}{2}\left\langle QQ^{-1}(c + A^T y), Q^{-1}(c + A^T y)\right\rangle - \left\langle Q^{-1}(c + A^T y), c + A^T y\right\rangle - \langle y, b\rangle \\
&= -\frac{1}{2}\left[\left\langle c, Q^{-1}c\right\rangle + 2\left\langle c, Q^{-1}A^T y\right\rangle + \left\langle A^T y, Q^{-1}A^T y\right\rangle\right] - \langle y, b\rangle \\
&= -\frac{1}{2}\left\langle y, AQ^{-1}A^T y\right\rangle - \left\langle y, AQ^{-1}c + b\right\rangle - \frac{1}{2}\left\langle c, Q^{-1}c\right\rangle \\
&= -\frac{1}{2}\left\langle AQ^{-1}A^T y, y\right\rangle - \left\langle y, AQ^{-1}c + b\right\rangle - \frac{1}{2}\left\langle Q^{-1}c, c\right\rangle
\end{aligned}$$

Což je přesně duální úloha (DQP). □

Poznámka. Úlohy kvadratického programování nejsou vzájemně duální.

10.2 Věta o silné dualitě pro kvadratické programování

Úloha (QP) má řešení právě tehdy, když (DQP) má řešení. Má-li (QP) řešení, pak se hodnoty obou úloh rovnají.

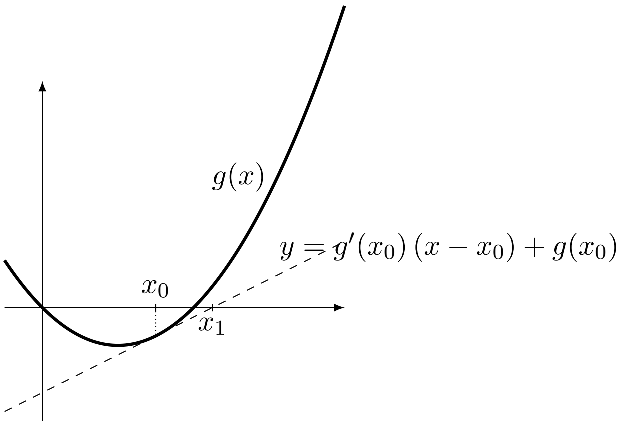
Důkaz vynecháme.

11 Numerické metody optimalisace

11.1 Newtonova metoda v jednorozměrné optimalisaci

Je dána rovnice $g(x) = 0$, kde $g \in C^1(\mathbb{R})$. Nechť $x_0 \in \mathbb{R}$. Položme

$$x_{k-1} = x_k - \frac{g(x_k)}{g'(x_k)},\ k \in \mathbb{N}_0.$$



- Předpokládejme, že $g'(x_k) \neq 0$ pro každé $k \in \mathbb{N}_0$.
- Pokud x_0 je dostatečně blízko řešení $\hat{x}$ rovnice $g(x)$, pak $x_k \to \hat{x}$.

Je dána funkce $f \in C^2(\mathbb{R})$. Hledejme stacionární body funkce f, tj. řešme rovnici $f'(x) = 0$. Z Newtonovy metody pro řešení rovnic plyne, že

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}$$

Algoritmus

(a) Zvolíme $\varepsilon > 0$ a $x_0 \in \mathbb{R}$. Položíme $k = 0$.

(b) Vypočítáme $f'(x_k)$ a $f''(x_k)$.

(c) Je-li $|f'(x_k)| < \varepsilon$, pak algoritmus končí a x_k je hledaná aproximace stacionárního bodu. V opačném případě přejdeme na další krok.

(d) Položíme

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}$$

(nutno ověřit, že $f''(x_k) \neq 0$) hodnotu k zvýšíme o 1 a jdeme na krok (b).

11.2 Omezení na minimalisační úlohy

Chceme nalézt alespoň přibližně bod minima (alespoň lokálního) funkce $f : \mathbb{R}^n \to \mathbb{R}$.
Postup:

- Zvolíme x_0 a konstruujeme posloupnost, jejíž členy jsou dány

  $$x_{k+1} = x_k + \alpha_k d_k,$$

  kde $\alpha_k \geq 0$ je <u>délka</u> k-tého kroku a $d_k \in \mathbb{R}^n$ je <u>směr</u> k-tého kroku.
- Vhodnou volbou délky kroku a směru se snažíme dosáhnout toho, aby $f(x_{k+1}) < f(x_k)$.

11.3 Nepodmíněná optimalisace - Metoda největšího spádu

V metodě největšího spádu předpokládáme, že $f \in C^1(\mathbb{R}^n)$.
Volba směru d_k:

- Chceme, aby $f(x_{k+1}) < f(x_k)$, a proto za směr d_k budeme volit směr poklesu, tj. prvek z $\mathcal{D}(f; x_k)$.
- Konkrétně volme $d_k = -\nabla f(x_k)$.
- Jestliže $\nabla f(x_k) \neq 0$, pak

  $$\langle d_k, \nabla f(x) \rangle = \langle -\nabla f(x_k), \nabla f(x_k) \rangle = -\|\nabla f(x_k)\|^2 < 0,$$

  tj. $d_k \in \mathcal{D}_0(f; x_k) \subseteq \mathcal{D}(f; x_k)$.
- Směr $d_k = -\nabla f(x_k)$ je směr největšího poklesu v bodě x_k.

Volba délky kroku α_k:

- Buď pevná volba kroku $\alpha_k = \alpha$ pro každé $k \in \mathbb{N}_0$. Příliš velké α může zkazit konvergenci.
- Nebo například $\alpha_k \in \operatorname{argmin}_{\alpha \geq 0} f(x_k - \alpha \nabla f(x_k))$.

Příklad špatně zvoleného α.

Mějme $f(x) = \frac{1}{2}\|x\|^2, \alpha = 11, x_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Pak $\nabla f(x) = \frac{1}{2} 2x = x$.

$$\begin{aligned} x_1 &= x_0 + \alpha(-x_0) = -10x_0 \\ x_2 &= x_1 - \alpha x_1 = -10x_1 = (-10)^2 x_0 \\ &\vdots \\ x_k &= (-10)^k x_0 \ldots \text{tedy očividně nekonverguje.} \end{aligned}$$

Volba kritéria zastaveni:

- $\|\nabla f(x_k)\| < \varepsilon$.
- Další možnosti jsou $\|x_{k+1} - x_k\| < \varepsilon$, $|f(x_{k+1} - f(x))| < \varepsilon$, ...
- Možná je i kombinace více kritérií.

Algoritmus

1. Zvolme $x_0 \in \mathbb{R}^n, \varepsilon > 0$. Položme $k = 0$.
2. Je-li $\|\nabla f(x_k)\| < \varepsilon$, pak algoritmus končí a x_k je hledaná aproximace. V opačném případě přejdeme na další krok.
3. Položíme $d_k = -\nabla f(x_k)$.
4. Nalezneme $\alpha_k \in \operatorname{argmin}_{\alpha \geq 0} f(x_k - \alpha \nabla f(x_k))$.
5. Položíme $x_{k+1} = x_k - \alpha_k \nabla f(x_k)$. Zvýšíme hodnotu k o 1 a jdeme na krok (b).

11.4 Podmíněná optimalisace - Metoda projekce gradientu

Modifikuje metodu největšího spádu.
Předpokládejme $f \in C^1(\mathbb{R}^n)$ a $C \subseteq \mathbb{R}^n$ je neprázdná, uzavřená a konvexní.
Nulovost gradientu již není vhodným kritériem pro zastavení.

Algoritmus

(a) Zvolme $x_0 \in C$ a $\varepsilon > 0$. Položmě $k = 0$.

(b) Vypočteme $d_k = -\nabla f(x_k)$.

(c) Nalezneme $\alpha_k \in \operatorname{argmin}_{\alpha \geq 0} f(x_k - \alpha \nabla f(x_k))$.

(d) Položíme $x_{k+1} = P_C(x_k - \alpha_k \nabla f(x_k))$.

(e) Je-li $|f(x_{k+1}) - f(x_k)| < \varepsilon$, pak algoritmus končí a x_k je hledaná aproximace. V opačném případě zvýšíme hodnotu k o 1 a jdeme na krok (b).

11.5 Podmíněná optimalisace - Metoda penalisačních funkcí

Nechť $f, g_1, \ldots, g_m \in C^1(\mathbb{R}^n)$ a je dána úloha

$$\left.\begin{array}{ll} \text{minimalisujte} & f(x) \\ \text{za podmínky} & g_1(x) \leq 0, \\ & \quad\vdots \\ & g_m(x) \leq 0. \end{array}\right\} (U)$$

- Chceme nahradit (U) úlohami nepodmíněné optimalisace.
- $p(x) = \sum_{i=1}^{m} \left[\max\{0, g_i(x)\}\right]^2 \ldots$ penalisační funkce.

Penalisační funkce zařídí, že čím dál budeme od přípustné množiny C, tím více budeme takové body penalisovat.

Algoritmus

(a) Zvolme $\varepsilon > 0, c_0 > 0$ a $\alpha > 1$. Položmě $k = 0$.

(b) Nalezněme bod minima x_k funkce $f(x) + c_k p(x)$ na $\mathbb{R}^n$.

(c) Je-li $c_k p(x) < \varepsilon$, pak algoritmus končí a x_k je hledaná aproximace. V opačném případě zvýšíme hodnotu k o 1 a jdeme na krok (b).

12 Úvod do strategických her

13 Maticové hry

14

15 Třináctý týden

16 Čtrnáctý týden