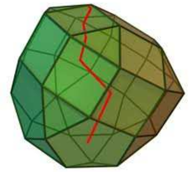
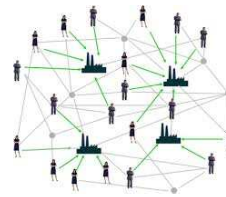
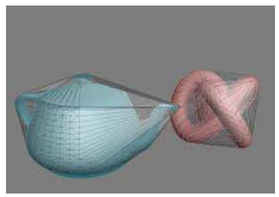
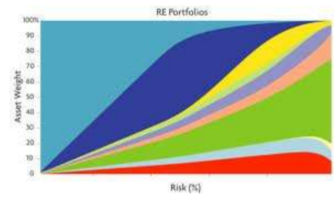
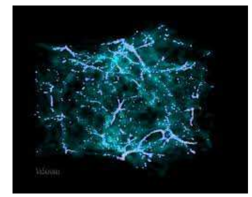
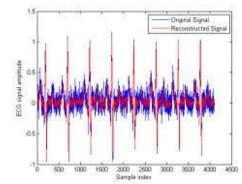
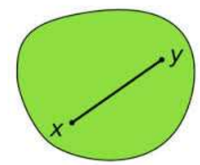


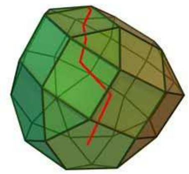




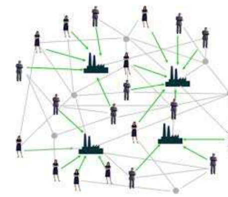
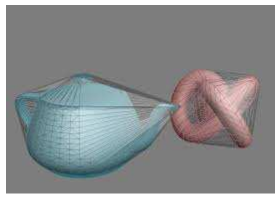
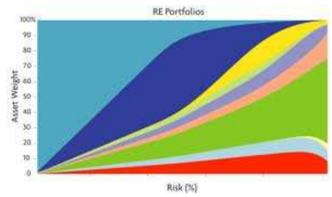



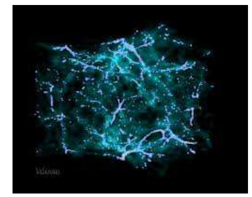
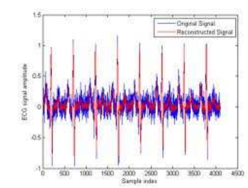



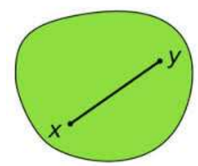
Konvexná optimalizácia - letný semester 2020

M. Trnovská, KAMŠ, FMFI UK





Téma 1: Klasické konvexné programovanie - opakovanie

Úlohou **konvexného programovania v štandardnom tvare** nazývame úlohu nelineárneho programovania

$$
\begin{array}{ll}
\min & f_0(x), \\
& f_i(x) \leq 0,\ i = 1, \ldots, m, \\
& Ax = b,
\end{array} \tag{1}
$$

kde

- funkcie $f_0, f_1, \ldots, f_m : \mathbb{R}^n \to \mathbb{R}$ sú konvexné,
- $A \in \mathbb{R}^{p \times n}$ a $b \in \mathbb{R}^p$.

Množina

$$\mathcal{P} = \{x \mid f_i(x) \leq 0,\ i = 1, \ldots, m,\ Ax = b\}.$$

sa nazýva **množina prípustných riešení** úlohy (1).

Zrejme $\mathcal{P}$ je konvexná množina

Ak definičný obor funkcií $f_i, i = 0, 1, .., m$ nie je $\mathbb{R}^n$ ale nejaká konvexná množina $\mathcal{D}_i, i = 0, 1, .., m$, tak sa zvykne definovať $\mathcal{D} = \cap_{i=0}^{m} \mathcal{D}_i$ a

$$\mathcal{P} = \{x \in \mathcal{D} \mid f_i(x) \leq 0,\ i = 1, \ldots, m,\ Ax = b\}$$

Optimálnu hodnotu úlohy (1) definujeme ako

$$p^* = \inf\{f_0(x) \mid x \in \mathcal{P}\},$$

teda p^* nadobúda hodnoty z $\mathbb{R} \cup \{\pm\infty\}$. Špecálne:

- definujeme $p = +\infty$, ak je úloha (1) neprípustná, t.j. $\mathcal{P} = \emptyset$;
- definujeme $p = -\infty$, ak je úloha (1) neohraničená, t.j. existuje postupnosť bodov $\{x_k\}_{k=1}^{\infty} \in \mathcal{P}$ taká, že $\lim_{k\to\infty} f_0(x_k) = -\infty$;
- ak je hodnota p^* konečná, tak buď existuje bod minima x^* taký, že $f_0(x^*) = p^*$ (minimum sa nadobúda), alebo nie (minimum sa nenadobúda).

Množinu optimálnych riešení úlohy (1) definujeme ako

$$\mathcal{P}^* = \{x \in \mathcal{P} \mid f_0(x) = p^*\}.$$

Dá sa ľahko ukázať, že $\mathcal{P}^*$ je konvexná množina.

Kritériá konvexnosti

Funkcia $f : \mathbb{R}^n \to \mathbb{R}$ sa nazýva **konvexná** ak $\forall x, y \in \mathbb{R}^n$ a $\forall \lambda \in [0,1]$ platí

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

Pre diferencovateľné funkcie sú známe **kritéria I. a II. rádu:** $f : \mathbb{R}^n \to \mathbb{R}$ je konvexná práve vtedy keď

- $f(x) \geq f(y) + \nabla f(y)^T (x-y), \ \forall x, y$ (Taylorova aproximácia I. rádu leží vždy pod grafom funkcie);
- $\nabla^2 f(x) \succeq 0, \forall x$ (kladná semidefinitnosť Hessovej matice).

Epigrafové kritérium: Funkcia $f : \mathbb{R}^n \to \mathbb{R}$ je konvexná práve vtedy keď jej epigraf

$$Epi(f) = \{(x, t) \mid f(x) \leq t\}$$

je konvexná množina.

Zúženie na úsečku (priamku): Funkcia $f : \mathbb{R}^n \to \mathbb{R}$ je konvexná práve vtedy keď $\forall x \neq y, \forall t \in [0, 1]$ $(t \in \mathbb{R})$ je funkcia

$$g_{x,y}(t) = f(tx + (1-t)y) = f(y + t(x - y))$$

konvexná.

- (**Suprémum z triedy konvexných funckií**)
  Nech $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, a $C \subseteq \mathbb{R}^m$ je ľubovoľná množina. Predpokladajme, že pre každé pevné $y \in C$ je $f(x, y)$ konvexná v x. Potom je funkcia

  $$g(x) = \sup_{y \in C} f(x, y)$$

  konvexná.

- (**Infímum z triedy konvexných funkcií**)
  Nech $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ je konvexná funkcia v (x, y) a $C \subseteq \mathbb{R}^m$ je konvexná množina. Potom je funkcia

  $$g(x) = \inf_{y \in C} f(x, y)$$

  konvexná.

K úlohe (1) definujeme Lagrangeovu funkciu $L : \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^p \to \mathbb{R}$ ako

$$L(x; u, v) = f_0(x) + \sum_{i=1}^{m} u_i f_i(x) + v^T (Ax - b).$$

Zrejme pre $x \in \mathcal{P}$ a $u \geq 0$ platí $L(x; u, v) \leq f_0(x)$.

Lagrangeova **duálna funkcia** je definovaná ako

$$G(u, v) = \inf_x L(x; u, v)$$

a Lagrangeova **duálna úloha** k úlohe (1) je potom

$$\begin{aligned} \max \quad & G(u, v), \\ & u \geq 0. \end{aligned} \tag{2}$$

Optimálnu hodnotu úlohy (2) definujeme ako

$$d^* = \sup\{G(u,v) \mid u \geq 0\},$$

teda d^* nadobúda hodnoty z $\mathbb{R} \cup \{\pm\infty\}$, pričom definujeme $d = -\infty$ ak je úloha (2) neprípustná a $d = +\infty$ ak je úloha (2) neohraničená.

Zrejme pre $\bar{x} \in \mathcal{P}$ a $u \geq 0$ platí

$$G(u,v) = \inf_x L(x; u, v) \leq L(\bar{x}; u, v) \leq f_0(\bar{x}),$$

a teda platí aj $d^* \leq p^*$, čo je vlastnosť **slabej duality.** Rozdiel $f_0(x) - G(u,v)$ pre prípustné x, u, v sa nazýva duálna medzera. Rozdiel $p^* - d^*$ sa nazýva optimálna duálna medzera. Ak $d^* = p^*$, t. j. optimálna duálna medzera je nulová, vtedy hovoríme, že nastala **silná dualita**.

Slaterova podmienka pre úlohu (1):

$$\exists \bar{x} : \; f_i(\bar{x}) < 0, \; i = 1, \ldots, m, \; A\bar{x} = b, \tag{3}$$

t.j. existuje relatívne vnútorný bod množiny $\mathcal{P}$ (vzhľadom na afínny podpriestor $\{x \mid Ax = b\}$.

Stačí uvažovať zoslabenú Slaterovu podmienku: pre funkcie f_i, ktoré sú afínne, stačí predpokladať $f_i(\bar{x}) \leq 0$.

(Slaterova veta)
Pre úlohu konvexného programovania (1) platí: ak je splnená Slaterova podmienka, tak $p^* = d^*$ (optimálna duálna medzera je nulová, nastala silná dualita). Navyše, ak je hodnota d^* konečná, tak existuje $u^* \geq 0, v^*$, také, že $G(u^*, v^*) = d^*$ (duálne optimum sa nadobúda).

Teda ak je splnená Slaterova podmienka, tak môžu nastať dva prípady:

- $p^* = d^* = -\infty$, t. j. duálna úloha je neprípustná a primárna je neohraničená;

- $-\infty < p^* = d^* < +\infty$, duálna úloha je prípustná a duálne optimum sa nadobúda (primárne optimum sa nadobúdať môže, ale nemusí).

Pre úlohu (1) máme systém **Kuhn-Tuckerovych (KT) podmienok optimality:**

KT1 $f_i(x) \leq 0, i = 1, \ldots, m,\ Ax = b$, (primárna prípustnosť),

KT2 $u \geq 0$, (duálna prípustnosť),

KT3 $u_i f_i(x) = 0,\ i = 1, \ldots, m$, (komplementarita),

KT4 $\nabla_x L(x; u, v) = \nabla f_0(x) + \Sigma_{i=1}^m u_i \nabla f_i(x) + A^T v = 0$, (spolu s KT1 stacionárny bod Lagrangeovej funkcie).

Pre konvexné úlohy platí: Ak bod (x^*, u^*, v^*) spĺňa systém KT podmienok optimality, tak x^* je optimálne riešenie (1) a (u^*, v^*) je optimálne riešenie (2).

Pre úlohy so silnou dualitou platí: Ak x^* je optimálne riešenie (1) a (u^*, v^*) je optimálne riešenie (2), tak bod (x^*, u^*, v^*) spĺňa systém KT podmienok optimality.

Teda pre konvexné úlohy, kde je splnená Slaterova podmienka je systém KT podmienok systémom **nutných a postačujúcich podmienok optimality.**