Konvexná optimalizácia

Podmienky optimality

Máme primárno-duálnu dvojicu úloh:

$$\left.\begin{array}{lll} Min & f_0(x) & \\ & f_i(x) \le 0, & i=1,\dots,m \\ & h_i(x) = 0, & i=1,\dots,p \end{array}\right\} \quad (P), \qquad \left.\begin{array}{ll} Max & G(u,v) \\ & u \ge 0 \end{array}\right\} \quad (D),$$

s príslušnou Lagrangeovou funkciou

$$L(x,u,v) = f_0(x) + \sum_{i=1}^{m} u_i f_i(x) + \sum_{i=1}^{p} v_i h_i(x),$$

pričom

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

KKT (Karush-Kuhn-Tucker) podmienky optimality:

$$f_i(x^*) \leq 0, \quad i = 1, \ldots, m,$$

$$h_i(x^*) = 0, \quad i = 1, \ldots, p,$$

$$u^* \geq 0, \quad i = 1, \ldots, m,$$

$$u_i f_i(x^*) = 0, \quad i = 1, \ldots, m$$

$$\nabla f_0(x^*) + \sum_{i=1}^{m} u_i^* \nabla f_i(x^*) + \sum_{i=1}^{p} v_i^* \nabla h_i(x^*) = 0.$$

Ak x^* je optimálne riešenie (P) a (u^*, v^*) je optimálne riešenie (D) a $p^* = d^*$, tak zrejme

$$f_0(x^*) = L(x^*, u^*, v^*) = G(u^*, v^*).$$

Dôsledky:

- Komplementarita - $u_i f_i(x^*) = 0, \ i = 1, \ldots, m$
- Ak sú funkcie $f_0, f_1, \ldots, f_m, h_1, \ldots, h_p$ diferencovateľné, tak z toho, že x^* minimalizuje $L(x, u^*, v^*)$ dostávame, že $\nabla_x L(x^*, u^*, v^*) = 0$.

Platí pre **ľubovolnú** úlohu matematického programovania so **silnou dualitou** a diferencovateľnými funkciami:

Ak x^* a (u^*, v^*) sú optimálne riešenia (P), (D), tak spĺňajú systém KKT podmienok.

Pre konvexné úlohy s diferencovatel'nými funkciami platí:
Ak body x^* a (u^*, v^*) spĺňajú systém KKT podmienok, tak x^* je optimálne pre (P) a (u^*, v^*) je optimálne pre (D).

Za predpokladu

- konvexnosti úlohy,
- diferencovatel'nosti funkcií $f_0, f_1, \ldots, f_m, h_1, \ldots, h_p$
- splnenia Slaterovej podmienky

KKT podmienky sú **nutnými aj postačujúcimi** podmienkami optimality.

Význam:

- V niektorých prípadoch je možné analyticky riešit' systém KKT podmienok a získat' optimálne riešenia.
- Mnohé algoritmy konvexnej optimalizácie sú založené na riešení systému KKT podmienok.

Príklad: "Water filling"

Uvažujme úlohu konvexného programovania:

$$\left.\begin{array}{ll} Min & -\sum_{i=1}^{n} \ln(\alpha_i + x_i) \\ & \mathbf{1}^T x = 1, \\ & x \geq 0, \end{array}\right\} \quad (WF)$$

kde $\alpha_i > 0$ sú dané.

- n komunikačných (prenosových) kanálov
- x_i - výkon pridelený i-temu kanálu (vo Wattoch)
- $\ln(\alpha_i + x_i)$ kapacita prenosového kanála
- Úloha: prerozdelit' celkový výkon vysielača medzi komunikačné kanály tak aby sa maximalizovala celková kapacita.

Lagrangeova funkcia:

$$L(x,u,v) = -\sum_{i=1}^{n} \ln(\alpha_i + x_i) - u^T x + v(\mathbf{1}^T x - 1)$$

KKT podmienky:

- Prípustnosť: $\mathbf{1}^T x = 1$, $x \geq 0$, $u \geq 0$;
- Komplementarita: $u_i x_i = 0,\ i = 1, \ldots, m$;
- $\nabla_x L(x,u,v) = -\frac{1}{\alpha_i + x_i} - u_i + v = 0$

Premennú u možno z podmienok **eliminovať**: $u = v - \frac{1}{\alpha_i + x_i}$

- Prípustnosť: $\mathbf{1}^T x = 1$, $x \geq 0$, $v \geq \frac{1}{\alpha_i + x_i}$;
- Komplementarita: $x_i \left(v - \frac{1}{\alpha_i + x_i} \right) = 0,\ i = 1, \ldots, m$;

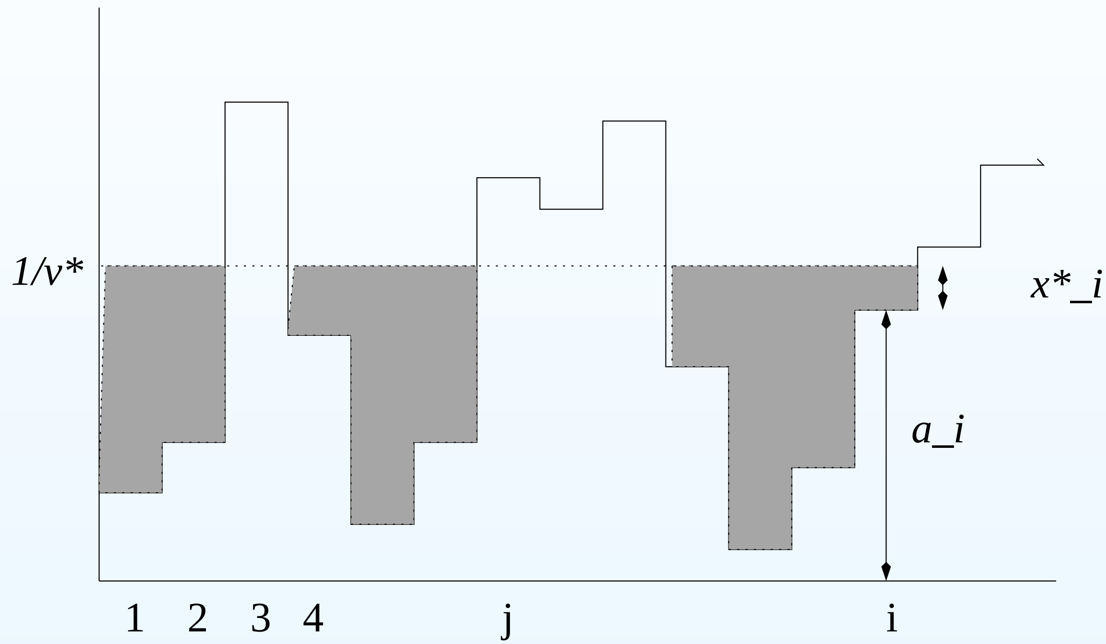
- Ak $v < \frac{1}{\alpha_i}$ tak $x_i = \frac{1}{v} - \alpha_i$.
- Ak $v \geq \frac{1}{\alpha_i}$, tak $x_i = 0$. Teda

$$x_i = \max\left\{0, \frac{1}{v}\right\}, \quad i = 1, \dots, m$$

a

$$\sum_{i=1}^{n} \max\left\{0, \frac{1}{v}\right\} = 1.$$

Funkcia vľavo je po častiach lineárna a rastúca v $\frac{1}{v}$ a teda rovnica má jediné riešenie.
Vyriešením rovnice vyššie nájdeme riešenie systému KKT podmienok a teda (keďže úloha je konvexná) aj optimálne riešenie úlohy (WF) a k nej duálnej úlohy.



"Water filling" metóda: máme nádrž s n rôznymi úrovňami, výška každej úrovne je daná hodnotou α_i. Nádrž sa naplní vodou, ktorej celkové množstvo je 1. Úroveň hladiny je potom rovná optimálnej hodnote $1/v^*$ a výška hladiny nad i-tou úrovňou je rovná optimálnej hodnote x_i.

PERTURBÁCIA A ANALÝZA CITLIVOSTI

$$\left.\begin{array}{ll} Min & f_0(x) \\ & f_i(x) \leq r_i, \quad i = 1, \ldots, m, \\ & h_i(x) = s_i, \quad i = 1, \ldots, p. \end{array}\right\} (P(r,s))$$

- $r_i > 0$ - relaxácia i-teho ohraničenia
- $r_i < 0$ - zúženie/stesnenie i-teho ohraničenia
- Optimálna hodnota perturbovanej úlohy (P(r,s)):

$$p^*(r,s) = \inf\{f_0(x) \mid f_i(x) \leq r_i, i = 1, \ldots, m,\ h_i(x) = s_i, i = 1, \ldots, p\}$$

- $p^*(0,0) = p^*$
- Ak je (P) konvexná úloha, tak $p^*(r,s)$ je konvexná funkcia (r,s)

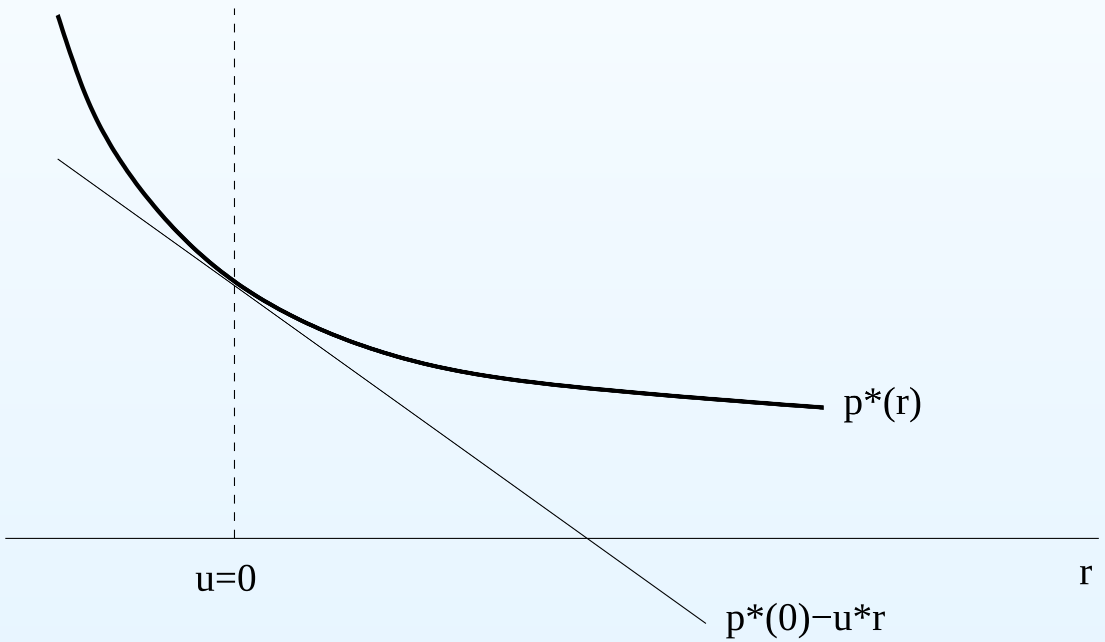
Predpokladajme, že pre neperturbovanú úlohu platí $d^* = p^*$ a duálne optimum sa nadobúda. Nech (u^*, v^*) je optimálne riešenie duálnej úlohy k neperturbovanej primárnej úlohe. Potom platí:

$$p^*(r,s) \geq p^*(0,0) - (u^*)^T r - (v^*)^T s$$

Dôsledky:

- Ak je u_i^* vel'ké a zvolíme $r_i < 0$, tak hodnota $p^*(r,s)$ prudko vzrastie.
- Ak je v_i^* vel'ké a kladné a zvolíme $s_i < 0$ alebo ak je v_i^* vel'ké a záporné a zvolíme $s_i > 0$, tak hodnota $p^*(r,s)$ prudko vzrastie.
- Ak je u_i^* malé a zvolíme $r_i > 0$, tak hodnota $p^*(r,s)$ príliš neklesne.
- Ak je v_i^* malé a kladné a zvolíme $s_i > 0$ alebo ak je v_i^* malé a záporné a zvolíme $s_i < 0$, tak hodnota $p^*(r,s)$ príliš neklesne.

$$\left.\begin{array}{ll} Min & f_0(x) \\ & f_1(x) \le r \end{array}\right\}$$



Optimálna hodnota $p^*(r)$ perturbovanej konvexnej úlohy s jediným ohraničením $f_1(x) \le r$ je konvexnou funkciou r. Afínna funkcia $p^*(0) - u^*r$ je dolným ohraničením na funkciu $p^*(r)$.

Lokálna analýza citlivosti

- Predpokladajme, že $p^*(r, s)$ je diferencovateľná v $r = 0, s = 0$.
- Položme $r = te_i,\ s = 0$. Potom

$$\lim_{t\to 0} \frac{p^*(te_i,0) - p^*(0,0)}{t} = \frac{\partial p^*(0,0)}{\partial r_i}$$

- Pre $t > 0$ máme

$$\frac{p^*(te_i,0) - p^*(0,0)}{t} \geq -u_i^*$$

- Pre $t < 0$ máme

$$\frac{p^*(te_i,0) - p^*(0,0)}{t} \leq -u_i^*$$

- Teda

$$\frac{\partial p^*(0,0)}{\partial r_i} = -u_i^*$$

- Analogicky dostaneme

$$\frac{\partial p^*(0,0)}{\partial s_i} = -v_i^*$$

Teda platí:

$$\frac{\partial p^*(0,0)}{\partial r_i} = -u_i^*, \quad \frac{\partial p^*(0,0)}{\partial s_i} = -v_i^*$$

Ak x^* je optimálne riešenie (P):

- $f_i(x^*) < 0$ - neaktívne ohraničenie, $u_i^* = 0$
- $f_i(x^*) = 0$ - aktívne ohraničenie -
    - u_i^* malé - uvol'nenie / zúženie i-teho ohraničenia nemá vel'ký vplyv na p^*.
    - u_i^* vel'ké - uvol'nenie / zúženie i-teho ohraničenia má vel'ký vplyv na p^*.