Optimalizace

12. Lagrangeova dualita

Tomáš Werner

FEL ČVUT

Minimaxní nerovnost

Věta

Pro každou funkci $L\colon X \times Y \to \mathbb{R}$ platí **minimaxní nerovnost**

$$\min_{x\in X} \underbrace{\max_{y\in Y} L(x,y)}_{F(x)} \;\geq\; \max_{y\in Y} \underbrace{\min_{x\in X} L(x,y)}_{G(y)}.$$

Rovnost nastává právě tehdy, existuje-li $(x^*, y^*) \in X \times Y$ (tzv. **sedlový bod**) splňující

$$L(x^*, y) \leq L(x^*, y^*) \leq L(x, y^*) \qquad \forall x \in X,\; y \in Y.$$

Příklady pro $X = Y = \{1, 2, 3, 4\}$:

$x \backslash y$	1	2	3	4	$F(x)$
1	-1	4	7	4	7
2	4	4	6	-2	6
3	1	5	3	3	5
4	3	5	3	2	5
$G(y)$	-1	4	3	-2	

$x \backslash y$	1	2	3	4	$F(x)$
1	-1	4	7	4	7
2	4	4	6	-2	6
3	1	0	$\boxed{3}$	3	3
4	3	3	$\boxed{3}$	2	3
$G(y)$	-1	0	3	-2	

Věta platí i pro nekonečné množiny X, Y, když zaměníme min/max za inf/sup.

Funkce s rozšířenou hodnotou

Úlohu

$$\min\{\, f(\mathbf{x}) \mid \mathbf{x} \in X,\ \mathbf{g}(\mathbf{x}) \le \mathbf{0} \,\}$$

můžeme ekvivalentně napsat jako

$$\min_{\mathbf{x} \in X} F(\mathbf{x})$$

kde

$$F(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{když } \mathbf{g}(\mathbf{x}) \le \mathbf{0} \\ +\infty & \text{jinak} \end{cases}$$

Omezení $\mathbf{g}(\mathbf{x}) \le \mathbf{0}$ je 'schováno' ve funkci F (tzv. **implicitní omezení**).

Máme $F\colon X \to \overline{\mathbb{R}}$ kde $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ značí **rozšířenou množinu reálných čísel**.

Lagrangeova dualita

Pro $f\colon \mathbb{R}^n \to \mathbb{R}$ a $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^m$ zavedeme **Lagrangeovu funkci**

$$L(\mathbf{x},\mathbf{y}) \;=\; f(\mathbf{x}) + \mathbf{y}^T\mathbf{g}(\mathbf{x}) \;=\; f(\mathbf{x}) + y_1 g_1(\mathbf{x}) + \cdots + y_m g_m(\mathbf{x})$$

Princip Lagrangeovy duality:

$$\underbrace{\min_{\mathbf{x}\in X} \underbrace{\max_{\mathbf{y}\in Y} L(\mathbf{x},\mathbf{y})}_{F(\mathbf{x})}}_{\text{primární úloha}} \;\geq\; \underbrace{\max_{\mathbf{y}\in Y} \underbrace{\min_{\mathbf{x}\in X} L(\mathbf{x},\mathbf{y})}_{G(\mathbf{y})}}_{\text{duální úloha}}$$

- F je primární účelová funkce (ve které jsou schována primární omezení). Zobrazení $f, \mathbf{g}$ a množiny $X \subseteq \mathbb{R}^n$, $Y \subseteq \mathbb{R}^m$ volíme tak, aby $\min_{\mathbf{x}\in X} F(\mathbf{x})$ byla kýžená primární úloha.
- G je duální účelová funkce (ve které jsou schována duální omezení). Je vždy **konkávní**, neboť je to minimum (přes $\mathbf{x} \in X$) afinních funkcí $L(\mathbf{x}, \cdot\,)$.
- Z minimaxní nerovnosti plyne **slabá dualita**: primární opt. hodnota $\geq$ duální opt. hodnota. Jejich rozdíl se nazývá **dualitní mezera**.
- **Silná dualita** platí (tj. dualitní mezera je nulová), právě když L má sedlový bod.

Příklady konstrukce primární úlohy

- Volbou $Y = \mathbb{R}^m$ dostaneme tuto primární účelovou funkci/úlohu:

$$F(\mathbf{x}) = \max_{\mathbf{y}\in\mathbb{R}^m}(f(\mathbf{x}) + \mathbf{y}^T\mathbf{g}(\mathbf{x})) = f(\mathbf{x}) + \max_{\mathbf{y}\in\mathbb{R}^m} \mathbf{y}^T\mathbf{g}(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{když } \mathbf{g}(\mathbf{x}) = \mathbf{0} \\ +\infty & \text{jinak} \end{cases}$$

$$\min_{\mathbf{x}\in X} F(\mathbf{x}) = \min\{\, f(\mathbf{x}) \mid \mathbf{x} \in X,\ \mathbf{g}(\mathbf{x}) = \mathbf{0} \,\}$$

- Volbou $Y = \mathbb{R}^m_+$ dostaneme tuto primární účelovou funkci/úlohu:

$$F(\mathbf{x}) = \max_{\mathbf{y}\geq\mathbf{0}}(f(\mathbf{x}) + \mathbf{y}^T\mathbf{g}(\mathbf{x})) = f(\mathbf{x}) + \max_{\mathbf{y}\geq\mathbf{0}} \mathbf{y}^T\mathbf{g}(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{když } \mathbf{g}(\mathbf{x}) \leq \mathbf{0} \\ +\infty & \text{jinak} \end{cases}$$

$$\min_{\mathbf{x}\in X} F(\mathbf{x}) = \min\{\, f(\mathbf{x}) \mid \mathbf{x} \in X,\ \mathbf{g}(\mathbf{x}) \leq \mathbf{0} \,\}$$

- Volbou $Y = \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}_+$ dostaneme kombinaci obou možností (omezení typu rovností i nerovností).

Příklad: LP

Napišme duální úlohu k úloze LP

$$\min\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \,\}$$

Zvolíme

$$f(\mathbf{x}) = \mathbf{c}^T\mathbf{x}, \quad \mathbf{g}(\mathbf{x}) = \mathbf{b} - \mathbf{A}\mathbf{x}, \quad X = \mathbb{R}^n_+, \quad Y = \mathbb{R}^m.$$

Lagrangeova funkce je

$$L(\mathbf{x}, \mathbf{y}) = \mathbf{c}^T\mathbf{x} + \mathbf{y}^T(\mathbf{b} - \mathbf{A}\mathbf{x}) = \mathbf{c}^T\mathbf{x} - \mathbf{y}^T\mathbf{A}\mathbf{x} + \mathbf{y}^T\mathbf{b}.$$

Ověříme, že

$$F(\mathbf{x}) = \max_{\mathbf{y}\in Y} L(\mathbf{x}, \mathbf{y}) = \begin{cases} \mathbf{c}^T\mathbf{x} & \text{když } \mathbf{A}\mathbf{x} = \mathbf{b} \\ +\infty & \text{jinak} \end{cases}$$

Platí

$$G(\mathbf{y}) = \min_{\mathbf{x}\in X} L(\mathbf{x}, \mathbf{y}) = \begin{cases} \mathbf{y}^T\mathbf{b} & \text{když } \mathbf{y}^T\mathbf{A} \leq \mathbf{c}^T \\ -\infty & \text{jinak} \end{cases}$$

Duální úloha tedy je

$$\max_{\mathbf{y}\in\mathbb{R}^m} G(\mathbf{y}) = \max\{\, \mathbf{b}^T\mathbf{y} \mid \mathbf{A}^T\mathbf{y} \leq \mathbf{c},\ \mathbf{y} \in \mathbb{R}^m \,\}$$

Příklad: LP s intervalovým omezením

Napišme duální úlohu k úloze LP

$$\min\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{x} \in [0,1]^n,\ \mathbf{A}\mathbf{x} = \mathbf{b} \,\}$$

Zvolíme

$$f(\mathbf{x}) = \mathbf{c}^T\mathbf{x}, \quad \mathbf{g}(\mathbf{x}) = \mathbf{b} - \mathbf{A}\mathbf{x}, \quad X = [0,1]^n, \quad Y = \mathbb{R}^m.$$

Lagrangeova funkce L a primární účelová funkce F jsou stejné jako minule.

Duální účelová funkce je

$$\begin{aligned}
G(\mathbf{y}) = \min_{\mathbf{x}\in X} L(\mathbf{x},\mathbf{y}) &= \min_{\mathbf{x}\in[0,1]^n} (\mathbf{c}^T\mathbf{x} - \mathbf{y}^T\mathbf{A}\mathbf{x} + \mathbf{y}^T\mathbf{b}) \\
&= \mathbf{y}^T\mathbf{b} + \min_{\mathbf{x}\in[0,1]^n} (\mathbf{c}^T - \mathbf{y}^T\mathbf{A})\mathbf{x} \\
&= \mathbf{y}^T\mathbf{b} + \min_{\mathbf{x}\in[0,1]^n} \sum_j (c_j - \mathbf{y}^T\mathbf{a}_j)x_j \\
&= \mathbf{y}^T\mathbf{b} + \sum_j \min_{x\in[0,1]} (c_j - \mathbf{y}^T\mathbf{a}_j)x \\
&= \mathbf{y}^T\mathbf{b} + \sum_j \min\{c_j - \mathbf{y}^T\mathbf{a}_j, 0\} \quad \left(\text{protože } \min_{x\in[0,1]} zx = \min\{z,0\}\right)
\end{aligned}$$

Duální úloha je $\max\limits_{\mathbf{y}\in\mathbb{R}^m} G(\mathbf{y})$.

Příklad: 0-1 LP

Napišme duální úlohu k úloze celočíselného LP

$$\min\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{x} \in \{0,1\}^n,\ \mathbf{A}\mathbf{x} = \mathbf{b} \,\}$$

Zvolíme

$$f(\mathbf{x}) = \mathbf{c}^T\mathbf{x}, \quad \mathbf{g}(\mathbf{x}) = \mathbf{b} - \mathbf{A}\mathbf{x}, \quad X = \{0,1\}^n, \quad Y = \mathbb{R}^m.$$

Duální úloha vyjde stejná jako minule, protože

$$\min_{x\in\{0,1\}} zx = \min_{x\in[0,1]} zx = \min\{z, 0\}.$$

Příklad: Řešení lineární soustavy s nejmenší euklid. normou

Napišme duální úlohu k úloze

$$\min\{\, \tfrac{1}{2}\|\mathbf{x}\|^2 \mid \mathbf{x} \in \mathbb{R}^n,\ \mathbf{A}\mathbf{x} = \mathbf{b} \,\}.$$

Zvolíme

$$f(\mathbf{x}) = \tfrac{1}{2}\|\mathbf{x}\|^2, \quad \mathbf{g}(\mathbf{x}) = \mathbf{b} - \mathbf{A}\mathbf{x}, \quad X = \mathbb{R}^n, \quad Y = \mathbb{R}^m.$$

Lagrangeova funkce je

$$L(\mathbf{x}, \mathbf{y}) = \tfrac{1}{2}\|\mathbf{x}\|^2 + \mathbf{y}^T(\mathbf{b} - \mathbf{A}\mathbf{x}) = \tfrac{1}{2}\|\mathbf{x}\|^2 - \mathbf{y}^T\mathbf{A}\mathbf{x} + \mathbf{y}^T\mathbf{b}.$$

Duální účelová funkce je

$$G(\mathbf{y}) = \min_{\mathbf{x}\in X} L(\mathbf{x}, \mathbf{y}) = \min_{\mathbf{x}\in\mathbb{R}^n}(\tfrac{1}{2}\|\mathbf{x}\|^2 - \mathbf{y}^T\mathbf{A}\mathbf{x} + \mathbf{y}^T\mathbf{b}) = \mathbf{y}^T\mathbf{b} - \tfrac{1}{2}\|\mathbf{y}^T\mathbf{A}\|^2$$

protože pro každé $\mathbf{c}$ platí

$$\min_{\mathbf{x}\in\mathbb{R}^n}(\tfrac{1}{2}\|\mathbf{x}\|^2 - \mathbf{c}^T\mathbf{x}) = -\tfrac{1}{2}\|\mathbf{c}\|^2$$

Duální úloha je

$$\max_{\mathbf{y}\in\mathbb{R}^m} G(\mathbf{y}) = \max_{\mathbf{y}\in\mathbb{R}^m}(\mathbf{y}^T\mathbf{b} - \tfrac{1}{2}\|\mathbf{y}^T\mathbf{A}\|^2)$$

Lineární úloha nejmenších čtverců

Úloha

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2$$

nemá omezení, proto její duál je triviální. Netriviální duál ale dostaneme, přepíšeme-li úlohu jako

$$\min\{\, \tfrac{1}{2}\|\mathbf{r}\|^2 \mid \mathbf{A}\mathbf{x}-\mathbf{b}=\mathbf{r},\ \mathbf{x}\in\mathbb{R}^n,\ \mathbf{r}\in\mathbb{R}^m \,\}$$

Lagrangeova funkce:

$$L(\mathbf{x},\mathbf{r},\mathbf{y}) = \tfrac{1}{2}\|\mathbf{r}\|^2 + \mathbf{y}^T(\mathbf{r}-\mathbf{A}\mathbf{x}+\mathbf{b})$$

Duální účelová funkce je

$$G(\mathbf{y}) = \min_{\substack{\mathbf{x}\in\mathbb{R}^n\\ \mathbf{r}\in\mathbb{R}^m}} L(\mathbf{x},\mathbf{r},\mathbf{y}) = \mathbf{y}^T\mathbf{b} + \min_{\substack{\mathbf{x}\in\mathbb{R}^n\\ \mathbf{r}\in\mathbb{R}^m}} (\tfrac{1}{2}\|\mathbf{r}\|^2 + \mathbf{y}^T\mathbf{r} - \mathbf{y}^T\mathbf{A}\mathbf{x}) = \begin{cases} \mathbf{y}^T\mathbf{b} - \tfrac{1}{2}\|\mathbf{y}\|^2 & \text{když } \mathbf{y}^T\mathbf{A}=\mathbf{0} \\ -\infty & \text{když } \mathbf{y}^T\mathbf{A}\neq\mathbf{0} \end{cases}$$

Duální úloha je tedy

$$\max\{\, \mathbf{y}^T\mathbf{b} - \tfrac{1}{2}\|\mathbf{y}\|^2 \mid \mathbf{y}^T\mathbf{A}=\mathbf{0},\ \mathbf{y}\in\mathbb{R}^m \,\}$$

Extrémy kvadratické funkce na sféře

Úloha

$$\min\{\, \mathbf{x}^T\mathbf{A}\mathbf{x} \mid \mathbf{x}^T\mathbf{x} = 1,\ \mathbf{x} \in \mathbb{R}^n \,\}$$

Lagrangeova funkce

$$L(\mathbf{x}, y) = \mathbf{x}^T\mathbf{A}\mathbf{x} + y(1 - \mathbf{x}^T\mathbf{x})$$

Duální účelová funkce

$$G(y) = \min_{\mathbf{x}\in\mathbb{R}^n} L(\mathbf{x}, y) = y + \min_{\mathbf{x}\in\mathbb{R}^n} \mathbf{x}^T(\mathbf{A} - y\mathbf{I})\mathbf{x} = \begin{cases} y & \text{když } \mathbf{A} - y\mathbf{I} \succeq \mathbf{0} \\ -\infty & \text{jinak} \end{cases}$$

neboť $\min_{\mathbf{x}\in\mathbb{R}^n} \mathbf{x}^T\mathbf{B}\mathbf{x}$ je rovno 0 když $\mathbf{B} \succeq \mathbf{0}$ a $-\infty$ jinak.

Tedy duální úloha je

$$\max\{\, y \mid \mathbf{A} - y\mathbf{I} \succeq \mathbf{0},\ y \in \mathbb{R} \,\}$$

tedy hledáme největší číslo y tak, aby matice $\mathbf{A} - y\mathbf{I}$ byla ještě positivně semidefinitní.

Podmínky komplementarity

Mějme primární úlohu

$$\min\{\, f(\mathbf{x}) \mid \mathbf{x} \in X,\ \mathbf{g}(\mathbf{x}) \le \mathbf{0} \,\}$$

a k ní Lagrangeovu funkci

$$L\colon X \times \mathbb{R}^m_+ \to \mathbb{R}, \qquad L(\mathbf{x},\mathbf{y}) = f(\mathbf{x}) + \mathbf{y}^T\mathbf{g}(\mathbf{x}).$$

Věta (o komplementaritě)

Nechť $(\mathbf{x}^*, \mathbf{y}^*)$ je sedlový bod Lagrangeovy funkce (tj. platí silná dualita).
Pak $\mathbf{y}^{*T}\mathbf{g}(\mathbf{x}^*) = 0$.

Důkaz. Stačí použít půlku sedlové podmínky: pro všechna $\mathbf{y} \in \mathbb{R}^m_+$ je $L(\mathbf{x}^*, \mathbf{y}) \le L(\mathbf{x}^*, \mathbf{y}^*)$, tj. $\mathbf{y}^T\mathbf{g}(\mathbf{x}^*) \le \mathbf{y}^{*T}\mathbf{g}(\mathbf{x}^*)$.
Protože $\mathbf{y}^* \in \mathbb{R}^m_+$, plyne z toho $\mathbf{g}(\mathbf{x}^*) \le \mathbf{0}$ a $\mathbf{y}^{*T}\mathbf{g}(\mathbf{x}^*) = 0$.

Podmínka $\mathbf{y}^{*T}\mathbf{g}(\mathbf{x}^*) = 0$ je ekvivalentní **podmínkám komplementarity**

$$y_i^* g_i(\mathbf{x}^*) = 0 \qquad \forall i = 1, \ldots, m.$$

Postačující podmínka pro silnou dualitu

Mějme primární úlohu

$$\min\{\, f(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n,\ \mathbf{g}(\mathbf{x}) \leq \mathbf{0}\,\}$$

a k ní Lagrangeovu funkci

$$L\colon \mathbb{R}^n \times \mathbb{R}^m_+ \to \mathbb{R}, \qquad L(\mathbf{x},\mathbf{y}) = f(\mathbf{x}) + \mathbf{y}^T\mathbf{g}(\mathbf{x}).$$

Věta (o silné dualitě)

Nechť

- funkce $f, g_1, \ldots, g_m$ jsou konvexní,
- funkce $g_1, \ldots, g_m$ splňují některé z podmínek zvaných **constraint qualifications**.

Pak Lagrangeova funkce má sedlový bod (tj. platí silná dualita).

Důkaz neuvádíme, je obtížný.

Často užívané constraint qualifications:

- **Afinní omezení:** funkce $g_1, \ldots, g_m$ jsou afinní
- **Slaterova podmínka:** existuje přípustné $\mathbf{x}$ tak, že $g_i(\mathbf{x}) < 0$ pro všechna i pro které g_i není afinní.

KKT (Karush-Kuhn-Tucker) podmínky optimality

Mějme primární úlohu a Lagrangeovu funkci jako minule, kde $f, \mathbf{g}$ jsou diferencovatelné.

KKT podmínky pro tuto úlohu:

$$L_{\mathbf{x}}(\mathbf{x}, \mathbf{y}) = f'(\mathbf{x}) + \mathbf{y}^T \mathbf{g}'(\mathbf{x}) = \mathbf{0} \tag{1}$$

$$\mathbf{g}(\mathbf{x}) \leq \mathbf{0} \tag{2}$$

$$\mathbf{y} \geq \mathbf{0} \tag{3}$$

$$\mathbf{y}^T \mathbf{g}(\mathbf{x}) = 0 \tag{4}$$

Věta (nutná podmínka)

Jestliže $(\mathbf{x}^*, \mathbf{y}^*)$ je sedlový bod Lagrangeovy funkce, pak splňuje KKT podmínky.

Důkaz. Podmínka (1) plyne z druhé půlky sedlové podmínky ($\mathbf{x}^*$ je minimum funkce $L(\,\cdot\,, \mathbf{y}^*)$ na $\mathbb{R}^n$).
Podmínka (3) je daná definičním oborem funkce L.
Podmínky (2) a (4) plynou z první půlky sedlové podmínky, viz důkaz věty o komplementaritě.

Věta (postačující podmínka pro konvexní úlohy)

Jestliže funkce $f, g_1, \ldots, g_m$ jsou konvexní a bod $(\mathbf{x}^*, \mathbf{y}^*)$ splňuje KKT podmínky, pak je to sedlový bod Lagrangeovy funkce.

Důkaz. Protože $\mathbf{y}^* \geq \mathbf{0}$, funkce $L(\,\cdot\,, \mathbf{y}^*)$ je konvexní. Z podmínky (1) tedy plyne, že funkce $L(\,\cdot\,, \mathbf{y}^*)$ na $\mathbb{R}^n$ nabývá minima v bodě $\mathbf{x}^*$ (druhá půlka sedlové podmínky).
Z podmínek (2) a (4) plyne, že funkce $L(\mathbf{x}^*, \,\cdot\,)$ na $\mathbb{R}^m_+$ nabývá maxima v bodě $\mathbf{y}^*$ (první půlka sedlové podmínky).