

Optimalizace a teorie her

Dualita

Martin Bohata

Katedra matematiky
FEL ČVUT v Praze
martin.bohata@fel.cvut.cz

Infimum a supremum – připomenutí

- Řekneme, že $a \in \mathbb{R} \cup \{-\infty, +\infty\}$ je **infimum** množiny $M \subseteq \mathbb{R}$ (píšeme $\inf M = a$), jestliže
 - 1 $a \leq x$ pro každé $x \in M$;
 - 2 jestliže $b \leq x$ pro každé $x \in M$, pak $b \leq a$.
- Řekneme, že $a \in \mathbb{R} \cup \{-\infty, +\infty\}$ je **supremum** množiny $M \subseteq \mathbb{R}$ (píšeme $\sup M = a$), jestliže
 - 1 $x \leq a$ pro každé $x \in M$;
 - 2 jestliže $x \leq b$ pro každé $x \in M$, pak $a \leq b$.
- Je-li $f : D \rightarrow \mathbb{R}$ a $M \subseteq D$, pak místo $\inf f(M)$ (resp. $\sup f(M)$) píšeme často $\inf_{x \in M} f(x)$ (resp. $\sup_{x \in M} f(x)$).

Příklad

At' $f : D_f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ a $g : D_g \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$.

① Je-li $m = n$ a $M \subseteq D_f \cap D_g$, pak

$$\inf_{x \in M} f(x) + \inf_{x \in M} g(x) \leq \inf_{x \in M} f(x) + g(x).$$

② Jsou-li $M \subseteq D_f$ a $N \subseteq D_g$ neprázdné, pak

$$\inf_{x \in M} f(x) + \inf_{y \in N} g(y) = \inf_{(x,y)^T \in M \times N} f(x) + g(y).$$

Motivace

Jak můžeme „co nejlépe“ zdola odhadnout minimum cílové funkce?

Příklad

Je dána úloha

minimalizujte $2x_1 + 3x_2$

za podmínek $1 - x_1 - x_2 \leq 0$,

$$x_1, x_2 \in [0, 2].$$

Označme $f(x_1, x_2) = 2x_1 + 3x_2$,

$$M = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in [0, 2]^2 \mid 1 - x_1 - x_2 \leq 0 \right\}$$

a $\hat{f} = \min_{x \in M} f(x)$.

Příklad (pokračování)

Využitím omezení ve tvaru nerovnosti dostaneme

- dolní odhad \hat{f} : pro každé $(x_1, x_2)^T \in M$ je
 $x_1 + 2x_2 + 1 = f(x_1, x_2) + (1 - x_1 - x_2) \leq f(x_1, x_2)$, a proto $1 \leq \hat{f}$;
- lepší dolní odhad \hat{f} : pro každé $(x_1, x_2)^T \in M$ je
 $x_2 + 2 = f(x_1, x_2) + 2(1 - x_1 - x_2) \leq f(x_1, x_2)$, a proto $2 \leq \hat{f}$;

Jaký je „nejlepší“ možný dolní odhad \hat{f} ?

Položme

$$\begin{aligned}L(x_1, x_2, \mu) &= 2x_1 + 3x_2 + \mu(1 - x_1 - x_2), \\ \varphi(\mu) &= \min_{(x_1, x_2)^T \in [0, 2]^2} L(x_1, x_2, \mu).\end{aligned}$$

Zřejmě $\varphi(\mu) \leq \hat{f}$ pro každé $\mu \geq 0$.

Nejlepší dolní odhad \hat{f} pomocí funkce φ je maximum $\hat{\varphi}$ funkce φ na \mathbb{R}_+ .

Příklad (Pokračování)

Problém nalezení „co nejlepšího“ dolního odhadu hodnoty \hat{f} tak přirozeně vede k úloze

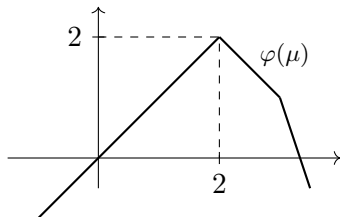
maximalizujte $\varphi(\mu)$
za podmínky $\mu \geq 0$.

Jedná se o tzv. duální úlohu k původně zadané úloze.

Zřejmě

$$\varphi(\mu) = \min_{(x_1, x_2)^T \in [0, 2]^2} L(x_1, x_2, \mu) = \begin{cases} \mu & \text{pro } \mu < 2, \\ 4 - \mu & \text{pro } \mu \in [2, 3), \\ 10 - 3\mu & \text{pro } \mu \geq 3. \end{cases}$$

Příklad (Pokračování)



Tedy

$$\hat{\varphi} = \max_{\mu \in \mathbb{R}_+} \varphi(\mu) = \varphi(2) = 2.$$

Odtud $\hat{f} \geq \hat{\varphi} = 2$. Navíc $f(1, 0) = 2$, a proto $\hat{f} = 2$.

Všimněme si, že v tomto konkrétním příkladu je $\hat{f} = \hat{\varphi}$.

Terminologie a značení

Nebude-li řečeno jinak, budeme v této kapitole používat níže uvedenou terminologii a značení.

- $f : D_f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ a $g : D_g \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^k$.
- $X = D_f \cap D_g$ a $\Omega \subseteq X$ neprázdná.
- **Primární úloha:**

$$\left. \begin{array}{l} \text{minimalizujte } f(x) \\ \text{za podmínky } g(x) \leq 0, \\ x \in \Omega. \end{array} \right\} \text{(P)}$$

- **Lagrangeova funkce** $L : X \times \mathbb{R}^k \rightarrow \mathbb{R}$ pro úlohu (P) je daná předpisem

$$L(x, \mu) = f(x) + \langle g(x), \mu \rangle.$$

- $D_\varphi = \left\{ \mu \in \mathbb{R}^k \mid \inf_{x \in \Omega} L(x, \mu) > -\infty \right\}$.

Terminologie a značení

- **Lagrangeova duální funkce** $\varphi : D_\varphi \rightarrow \mathbb{R}$ je dána předpisem

$$\varphi(\mu) = \inf_{x \in \Omega} L(x, \mu).$$

- **Duální úloha** k úloze (P):

$$\left. \begin{array}{l} \text{maximalizujte } \varphi(\mu) \\ \text{za podmínky } \mu \geq 0. \end{array} \right\} \text{(D)}$$

- $M = \{x \in \Omega \mid g(x) \leq 0\}$... přípustná množina úlohy (P).
- $N = \{\mu \in D_\varphi \mid \mu \geq 0\}$... přípustná množina úlohy (D).
- $\hat{f} = \inf_{x \in M} f(x)$... **hodnota primární úlohy** (P).
- $\hat{\varphi} = \sup_{\mu \in N} \varphi(\mu)$... **hodnota duální úlohy** (D).

Konstrukce duální úlohy

Příklad

Je dána úloha

$$\begin{aligned} &\text{minimalizujte } x_1 + 2x_2 \\ &\text{za podmínek } x_1 + x_2 \geq 1, \\ & \quad \quad \quad x_1, x_2 \geq 0. \end{aligned}$$

- Považujeme-li $x_1, x_2 \geq 0$ za přímé omezení, pak duální úloha je

$$\begin{aligned} &\text{maximalizujte } \mu \\ &\text{za podmínky } \mu \leq 1 \\ & \quad \quad \quad \mu \geq 0. \end{aligned}$$

Příklad (Pokračování)

- Považujeme-li $x \in \mathbb{R}^2$ za přímé omezení, pak duální úloha je

maximalizujte μ_1

za podmínek $1 - \mu_1 - \mu_2 = 0,$

$2 - \mu_1 - \mu_3 = 0$

$\mu_1, \mu_2, \mu_3 \geq 0.$

Věty o slabé a silné dualitě

Tvrzení

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

Důkaz: Viz přednáška. ■

- V netriviálním případě $N \neq \emptyset$ je duální úloha úlohou konvexní optimalizace!
- Hlavní otázka teorie duality je vztah \hat{f} a $\hat{\varphi}$.

Věta (O slabé dualitě)

- 1 Pro každé $x \in M$ a $\mu \in N$ je $\varphi(\mu) \leq f(x)$.
- 2 $\hat{\varphi} \leq \hat{f}$.

Důkaz: Viz přednáška. ■

Věty o slabé a silné dualitě

Důsledek

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

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

- ❷ Je-li $\hat{f} = -\infty$, pak $N = \emptyset$.
❸ Je-li $\hat{\varphi} = \infty$, pak $M = \emptyset$.

Důkaz: Viz přednáška. ■

- Rozdíl $\hat{f} - \hat{\varphi}$ se nazývá **duální mezera**.
- Pokud platí předpoklady z ❶, pak $\hat{\varphi} = \hat{f}$.

Věty o slabé a silné dualitě

Příklad

Je dána úloha

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

Snadno nalezneme, že

$$\begin{aligned} L(x, \mu) &= -x^2 + \mu(2x - 1), \\ \varphi(\mu) &= \begin{cases} \mu - 1 & \text{pro } \mu < \frac{1}{2}, \\ -\mu & \text{pro } \mu \geq \frac{1}{2}. \end{cases} \end{aligned}$$

Zřejmě $\hat{\varphi} = -\frac{1}{2}$ a $\hat{f} = -\frac{1}{4}$. Tedy $\hat{\varphi} < \hat{f}$!

Věty o slabé a silné dualitě

- Za jakých předpokladů nastává rovnost $\hat{\varphi} = \hat{f}$?

Věta (O silné dualitě)

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

- 1 Komponenty g_1, \dots, g_k zobrazení g splňují Slaterovu podmínku regularity.
- 2 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: Vynecháváme. ■

- Předpoklad $\hat{f} < \infty$ znamená, že přípustná množina M je neprázdná.
- Pokud $\hat{f} = -\infty$, pak díky větě O slabé dualitě musí být $\hat{\varphi} = -\infty$.

Příklady

Příklad

Je dána úloha

$$\begin{aligned} &\text{minimalizujte } x_1^2 + x_2^2 \\ &\text{za podmíněk } x_1 + x_2 \geq 4, \\ & \quad \quad \quad x_1, x_2 \geq 0. \end{aligned}$$

- Je-li $x_1, x_2 \geq 0$ přímé omezení, pak duální úloha je

$$\begin{aligned} &\text{maximalizujte } -\frac{\mu^2}{2} + 4\mu \\ &\text{za podmínky } \mu \geq 0. \end{aligned}$$

Bod maxima je $\hat{\mu} = 4$. Z věty O silné dualitě máme $8 = \hat{\varphi} = \hat{f}$. Odtud

$$\begin{aligned} \operatorname{argmin}_{x \in M} f(x) &= \left\{ (x_1, x_2) \in \mathbb{R}_+^2 \mid x_1^2 + x_2^2 = 8, x_1 + x_2 \geq 4 \right\} \\ &= \{(2, 2)\}. \end{aligned}$$

Příklad (pokračování)

- Je-li přímé omezení ve tvaru $x \in \mathbb{R}^2$, pak duální úloha je

$$\text{maximalizujte } 4\mu_1 - \frac{(\mu_1 + \mu_2)^2}{4} - \frac{(\mu_1 + \mu_3)^2}{4}$$

$$\text{za podmíněk } \mu_1, \mu_2, \mu_3 \geq 0.$$

Příklad

V \mathbb{R}^n jsou dány množiny bodů $A = \{a_1, \dots, a_k\}$ a $B = \{b_1, \dots, b_l\}$. Ať $w \in \mathbb{R}^n$ a $\lambda \in \mathbb{R}$. K úloze

$$\text{minimalizujte } h(w, \lambda) = \frac{1}{2} \|w\|^2$$

$$\text{za podmíněk } \langle a_i, w \rangle + \lambda \geq 1 \text{ pro všechna } i = 1, \dots, k,$$
$$\langle b_j, w \rangle + \lambda \leq -1 \text{ pro všechna } j = 1, \dots, l.$$

zkonstruuje úlohu duální (přímé omezení je $w \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$).

Příklady

Příklad (pokračování)

Duální úloha je

maximalizujte $\varphi(y_1, \dots, y_k, z_1, \dots, z_l)$

za podmínek
$$\sum_{i=1}^k y_i - \sum_{i=1}^l z_i = 0,$$

$$y_1, \dots, y_k, z_1, \dots, z_l \geq 0,$$

kde

$$\begin{aligned} \varphi(y_1, \dots, y_k, z_1, \dots, z_l) = & -\frac{1}{2} \sum_{i=1}^k \sum_{j=1}^k \langle a_i, a_j \rangle y_i y_j + \sum_{i=1}^k \sum_{j=1}^l \langle a_i, b_j \rangle y_i z_j \\ & - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \langle b_i, b_j \rangle z_i z_j + \sum_{i=1}^k y_i + \sum_{i=1}^l z_i. \end{aligned}$$

Primární úlohy s omezeními ve tvaru rovnosti i nerovnosti

Analogicky lze konstruovat duální úlohu i pro úlohu tvaru

$$\begin{aligned} &\text{minimalizujte} && f(x) \\ &\text{za podmínky} && g(x) \leq 0, \\ & && h(x) = 0, \\ & && x \in \Omega, \end{aligned}$$

kde

- $f : D_f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$,
- $g : D_g \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^k$, $h : D_h \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^l$,
- $X = D_f \cap D_g \cap D_h$ a $\Omega \subseteq X$ neprázdná.

Lagrangeovu funkci $L : X \times \mathbb{R}^k \times \mathbb{R}^l \rightarrow \mathbb{R}$ této úlohy definujeme předpisem

$$L(x, \mu, \lambda) = f(x) + \langle g(x), \mu \rangle + \langle h(x), \lambda \rangle .$$

Primární úlohy s omezeními ve tvaru rovnosti i nerovnosti

Duální úloha k uvedené (primární) úloze je

$$\begin{array}{ll} \text{maximalizujte} & \varphi(\mu, \lambda) \\ \text{za podmínky} & \mu \geq 0, \end{array}$$

kde $\varphi : D_\varphi \rightarrow \mathbb{R}$,

$$D_\varphi = \left\{ \begin{pmatrix} \mu \\ \lambda \end{pmatrix} \in \mathbb{R}^k \times \mathbb{R}^l \mid \inf_{x \in \Omega} L(x, \mu, \lambda) > -\infty \right\},$$
$$\varphi(\mu, \lambda) = \inf_{x \in \Omega} L(x, \mu, \lambda).$$

- I v tomto případě lze dokázat analogie vět O slabé dualitě a O silné dualitě.