Kapitola 5

Symetrická úloha NLP

V této kapitole se budeme zabývat **symetrickou úlohou nelineárního programování**

$$
\begin{array}{lll}
\text{minimalizovat} & f(\mathbf{x}), & \\
\text{za podmínek} & g_k(\mathbf{x}) \leq 0 & \forall\ k = 1, 2, \ldots, \mathsf{m}, \\
 & x_i \geq 0 & \forall\ i = 1, 2, \ldots, \mathsf{n}.
\end{array}
\tag{5.1}
$$

Úloha je speciálním případem obecné úlohy NLP, nemá omezení ve tvaru rovnic (2.13) a podmínka nezápornosti je explicitně vyjádřena.

Úloze (5.1) přiřadíme **Lagrangeovu funkci**

$$
\mathsf{L}(\mathbf{x}; \mathbf{y}) := f(\mathbf{x}) + \sum_{k=1}^{\mathsf{m}} y_k g_k(\mathbf{x}). \tag{5.2}
$$

5.1 Globální podmínky optimality

Definice 5.1.: *Řekneme, že pro* $\mathbf{x}^* \in \mathbb{R}^{\mathsf{n}}$, $\mathbf{y}^* \in \mathbb{R}^{\mathsf{m}}$ *jsou splněny* **globální podmínky optimality** *(GPO) pro úlohu (5.1), jestliže*

$$
\mathbf{x}^* \geq \mathbf{0},\ \mathbf{y}^* \geq \mathbf{0} \tag{5.3}
$$

$$
\textit{a pro každé} \quad \mathbf{x} \geq \mathbf{0},\ \mathbf{y} \geq \mathbf{0} \quad \textit{platí} \quad \mathsf{L}(\mathbf{x}; \mathbf{y}^*) \geq \mathsf{L}(\mathbf{x}^*; \mathbf{y}^*) \geq \mathsf{L}(\mathbf{x}^*; \mathbf{y})\,. \tag{5.4}
$$

Říkáme též, že $(\mathbf{x}^*, \mathbf{y}^*)$ *je sedlový bod Lagrangeovy funkce (5.2) v oboru* $\mathbf{x} \geq \mathbf{0}, \mathbf{y} \geq \mathbf{0}$.

Věta 5.2.: *Nechť jsou pro* $\mathbf{x}^* \in \mathbb{R}^{\mathsf{n}}$, $\mathbf{y}^* \in \mathbb{R}^{\mathsf{m}}$ *splněny GPO pro úlohu NLP (5.1). Pak je* $\mathbf{x}^*$ *optimální řešení zmíněné úlohy a je splněna* **podmínka komplementarity**

$$
\textit{pro každé} \quad k = 1, 2, \ldots, \mathsf{m} \quad \textit{je buď} \quad y_k^* = 0 \quad \textit{nebo} \quad g_k(\mathbf{x}^*) = 0\,. \tag{5.5}
$$

Důkaz: Podle předpokladu je $\mathbf{x}^* \geq \mathbf{0}$. Dosazením tvaru Lagrangeovy funkce (5.2) do (5.4) dostáváme

$$
f(\mathbf{x}) + \sum_{k=1}^{\mathsf{m}} y_k^* g_k(\mathbf{x}) \geq f(\mathbf{x}^*) + \sum_{k=1}^{\mathsf{m}} y_k^* g_k(\mathbf{x}^*) \geq f(\mathbf{x}^*) + \sum_{k=1}^{\mathsf{m}} y_k g_k(\mathbf{x}^*)\,. \tag{5.6}
$$

Nerovnosti (5.6) podle předpokladu platí pro nějaké $\mathbf{y}^* \geq \mathbf{0}$ a pro všechna $\mathbf{x} \geq \mathbf{0}$, $\mathbf{y} \geq \mathbf{0}$. Použijeme nejprve pravou větev (5.6), tj. nerovnost

$$\sum_{k=1}^{\mathsf{m}} y_k^* g_k(\mathbf{x}^*) \geq \sum_{k=1}^{\mathsf{m}} y_k g_k(\mathbf{x}^*) \quad \forall \mathbf{y} \geq \mathbf{0}\,. \tag{5.7}$$

Uvědomíme si, že $g_k(\mathbf{x}^*)$ jsou pevné koeficienty a že $\mathbf{y}^* \geq \mathbf{0}$ je z hlediska uvažované nerovnosti také pevný vektor. Odtud přímo vyplývá, že (5.7) může platit jen když $g_k(\mathbf{x}^*) \leq 0$, $k = 1, 2, \ldots, \mathsf{m}$ a

$$\sum_{k=1}^{\mathsf{m}} y_k^* g_k(\mathbf{x}^*) = 0\,. \tag{5.8}$$

Tato podmínka je ekvivalentní s podmínkou komplementarity (5.5).

To znamená, že $\mathbf{x}^*$ je přípustné řešení úlohy (5.1) a že se levá větev nerovností (5.6) redukuje na

$$f(\mathbf{x}) + \sum_{k=1}^{\mathsf{m}} y_k^* g_k(\mathbf{x}) \geq f(\mathbf{x}^*) \quad \forall \mathbf{x} \geq \mathbf{0}\,. \tag{5.9}$$

Omezme se jen na přípustná $\mathbf{x}$. Pro ně je $g_k(\mathbf{x}) \leq 0$, $k = 1, 2, \ldots, \mathsf{m}$ a z nerovnosti (5.9) plyne $f(\mathbf{x}) \geq f(\mathbf{x}^*)$.

Tedy $\mathbf{x}^*$ je optimální řešení úlohy (5.1).

Q.E.D.

5.2 Podmínky regularity

Věta 5.2 platí bez jakýchkoli předpokladů na funkce f, g_k. Avšak pro ověřování optimality řešení nebo pro nalezení optimálního řešení je nepraktická. Jako vedlejší výsledek jsme dostali podmínku komplementarity (5.5). (Porovnejte s podmínkami komplementarity v LP (3.32), (3.33).)

Věta 5.2 dává postačující podmínku pro bod globálního minima funkce f na množině popsané nerovnostmi (5.1). Otázkou je, kdy bude globální podmínka (5.6) nutná. Snadno najdeme příklad, kdy k optimálnímu řešení $\mathbf{x}^*$ úlohy (5.1) neexistuje vektor Lagrangeových multiplikátorů $\mathbf{y}^*$ tak, aby platilo (5.6).

Příklad 5.3.: Řešme úlohu $\min\left\{-x \,:\, x^2 \leq 0,\ x \geq 0,\ x \in \mathbb{R}\right\}$, tj. $f(x) = -x$, $g(x) = x^2$.

Existuje jediné přípustné a tedy i optimální řešení úlohy. Je to $x^* = 0$.

Lagrangeova funkce má tvar $\mathsf{L}\,(x; y) = -x + yx^2$.

1. V bodě $(0, 0)$ nemá sedlový bod neboť funkce $x \mapsto \mathsf{L}\,(x; 0) = -x$ nemá na $x \geq 0$ minimum v $x^* = 0$.
2. Pro $y > 0$, v bodě $(0, y)$ také není sedlový bod Lagrangeovy funkce, neboť $x \mapsto \mathsf{L}\,(x; y) = -x + yx^2$ má na $x \geq 0$ minimum v $\bar{x} = \frac{1}{2y}$ a nikoli v $x^* = 0$.

K bodu $x^* = 0$ tedy neexistuje žádné $y^* \in \mathbb{R}$ tak, aby dvojice (x^*, y^*) splňovala GPO.

$\triangle$

Je tedy zapotřebí dalších předpokladů, tzv. **podmínek regularity**, které vyloučí podobné degenerované případy. Z příkladu vidíme, že samotný předpoklad konvexnosti funkcí $f, g_1, g_2, \ldots, g_{\mathsf{m}}$ nestačí.

Definice 5.4.: *Pro úlohu (5.1) zavedeme* **množinu indexů aktivních omezení** *v bodě* $\mathbf{x} \geq \mathbf{0}$

$$\mathsf{B}_g\,(\mathbf{x}) \;=\; \{k = 1, 2, \ldots, \mathsf{m} \,:\, g_k(\mathbf{x}) = 0\}\,, \tag{5.10}$$

$$\mathsf{B}_x\,(\mathbf{x}) \;=\; \{i = 1, 2, \ldots, \mathsf{n} \,:\, \mathbf{x}_i = 0\}\,. \tag{5.11}$$

Uvedeme si tři podmínky, které spolu s konvexností funkcí $f, g_1, g_2, \ldots, g_{\mathsf{m}}$ zaručují ekvivalenci mezi splněním (5.4) a existencí optimálního řešení úlohy (5.1).

Definice 5.5.: *Řekneme, že pro přípustné řešení* $\mathbf{x} \in \mathbb{R}^{\mathsf{n}}$ *úlohy (5.1) je splněna* **podmínka lineární nezávislosti**, *jsou-li funkce* g_k, $k = 1, 2, \ldots, \mathsf{m}$ *v bodě* $\mathbf{x}$ *diferencovatelné a gradienty aktivních omezení jsou lineárně nezávislé, tj. matice*

$$(\nabla_x g_k(\mathbf{x}) \;\; k \in \mathsf{B}_g(\mathbf{x}), \;\; \mathbf{e}_i \;\; i \in \mathsf{B}_x(\mathbf{x}))$$

má plnou sloupcovou hodnost.

Definice 5.6.: *Řekneme, že úloha (5.1) splňuje* **Slaterovu podmínku**, *jestliže existuje* $\tilde{\mathbf{x}} \geq 0$, *pro které platí* $g_k(\tilde{\mathbf{x}}) < 0$ *pro každé* $k = 1, 2, \ldots, \mathsf{m}$.

Platí následující věty.

Věta 5.7.: *Nechť* f, g_k, $k = 1, 2, \ldots, \mathsf{m}$ *jsou konvexní funkce a v bodě* $\mathbf{x}^* \in \mathbb{R}^{\mathsf{n}}$ *je splněna podmínka lineární nezávislosti. Pak* $\mathbf{x}^*$ *je optimální řešení úlohy (5.1) tehdy a jen tehdy, když existuje* $\mathbf{y}^* \in \mathbb{R}^{\mathsf{m}}$ *tak, že dvojice* $(\mathbf{x}^*, \mathbf{y}^*)$ *splňuje GPO.*

Důkaz: Důkaz je uveden například v [1] a v [4].

Q.E.D.

Věta 5.8.: *Nechť* f, g_k, $k = 1, 2, \ldots, \mathsf{m}$ *jsou konvexní funkce a je splněna Slaterova podmínka. Pak* $\mathbf{x}^* \in \mathbb{R}^{\mathsf{n}}$ *je optimální řešení úlohy (5.1) tehdy a jen tehdy, když existuje* $\mathbf{y}^* \in \mathbb{R}^{\mathsf{m}}$ *tak, že dvojice* $(\mathbf{x}^*, \mathbf{y}^*)$ *splňuje GPO.*

Důkaz: Důkaz je uveden například v [1], [4] a v [5].

Q.E.D.

Věta 5.9.: *Nechť* f *je konvexní funkce a* g_k, $k = 1, 2, \ldots, \mathsf{m}$ *jsou lineární funkce. Pak* $\mathbf{x}^* \in \mathbb{R}^{\mathsf{n}}$ *je optimální řešení úlohy (5.1) tehdy a jen tehdy, když existuje* $\mathbf{y}^* \in \mathbb{R}^{\mathsf{m}}$ *tak, že dvojice* $(\mathbf{x}^*, \mathbf{y}^*)$ *splňuje GPO.*

Důkaz: Důkaz je uveden například v [1], [4] a v [5].

Q.E.D.

Povšimněte si, že podmínka lineární nezávislosti je obdobou podmínky používané ve větě o vázaném extrému z matematické analýzy.

5.3 Lokální podmínky optimality

V této podkapitole představíme nutné a postačující podmínky k tomu, aby dvojice $\mathbf{x}^*, \mathbf{y}^*$ splňovala GPO. K tomu cíli budeme většinou předpokládat diferencovatelnost a konvexnost funkcí $f, g_1, g_2, \ldots, g_{\mathsf{m}}$.

Definice 5.10.: *Řekneme, že pro* $\mathbf{x}^* \in \mathbb{R}^{\mathsf{n}}$, $\mathbf{y}^* \in \mathbb{R}^{\mathsf{m}}$ *jsou splněny* **lokální podmínky optimality** *(LPO) pro úlohu (5.1), jestliže platí*

$$\nabla_x \mathsf{L}(\mathbf{x}^*; \mathbf{y}^*) \geq \mathbf{0}, \qquad (\mathbf{x}^*)^\top \nabla_x \mathsf{L}(\mathbf{x}^*; \mathbf{y}^*) = 0, \quad \mathbf{x}^* \geq \mathbf{0}, \tag{5.12}$$

$$\nabla_y \mathsf{L}(\mathbf{x}^*; \mathbf{y}^*) \leq \mathbf{0}, \qquad (\mathbf{y}^*)^\top \nabla_y \mathsf{L}(\mathbf{x}^*; \mathbf{y}^*) = 0, \quad \mathbf{y}^* \geq \mathbf{0}. \tag{5.13}$$

Poznamenejme, že pro každé $k = 1, 2, \ldots, \mathsf{m}$ platí $\frac{\partial \mathsf{L}}{\partial y_k}(\mathbf{x}^*; \mathbf{y}^*) = g_k(\mathbf{x}^*)$. Tudíž podmínka (5.13) je shodná s podmínkou

$$g_k(\mathbf{x}^*) \leq 0 \quad \text{pro každé } k = 1, 2, \ldots, \mathsf{m} \quad a \quad (\mathbf{y}^*)^\top g(\mathbf{x}^*) = \sum_{k=1}^{\mathsf{m}} y_k^* g_k(\mathbf{x}^*) = 0\,. \tag{5.14}$$

Věta 5.11.: *Nechť $f, g_1, g_2, \ldots, g_{\mathsf{m}}$ jsou diferencovatelné. Pak podmínky LPO, tj. (5.12), (5.13), jsou nutné k tomu, aby dvojice $\mathbf{x}^*, \mathbf{y}^*$ splňovala GPO pro úlohu (5.1).*

Důkaz: Nechť $\mathbf{x}^*, \mathbf{y}^*$ splňují GPO.

1. Pak pro každé $i = 1, 2, \ldots, \mathsf{n}$ platí

$$\frac{\partial \mathsf{L}}{\partial x_i}(\mathbf{x}^*; \mathbf{y}^*) = \lim_{t \to 0+} \frac{\mathsf{L}\left(\mathbf{x}^* + t\,\mathbf{e}_i; \mathbf{y}^*\right) - \mathsf{L}\left(\mathbf{x}^*; \mathbf{y}^*\right)}{t} \geq 0\,,$$

   neboť pro každé kladné t je

$$\mathsf{L}\left(\mathbf{x}^* + t\,\mathbf{e}_i; \mathbf{y}^*\right) \geq \mathsf{L}\left(\mathbf{x}^*; \mathbf{y}^*\right)\,.$$

   ($\mathbf{e}_i$ označuje vektor, jehož i-tá složka je 1 a ostatní složky jsou nulové.)

   Pokud $x_i^* > 0$, pak platí ještě

$$\frac{\partial \mathsf{L}}{\partial x_i}(\mathbf{x}^*; \mathbf{y}^*) = \lim_{t \to 0-} \frac{\mathsf{L}\left(\mathbf{x}^* + t\,\mathbf{e}_i; \mathbf{y}^*\right) - \mathsf{L}\left(\mathbf{x}^*; \mathbf{y}^*\right)}{t} \leq 0\,.$$

   Tím je splnění podmínky (5.12) dokázáno.

2. Pak pro každé $k = 1, 2, \ldots, \mathsf{m}$ platí

$$\frac{\partial \mathsf{L}}{\partial y_k}(\mathbf{x}^*; \mathbf{y}^*) = \lim_{t \to 0+} \frac{\mathsf{L}\left(\mathbf{x}^*; \mathbf{y}^* + t\,\mathbf{e}_k\right) - \mathsf{L}\left(\mathbf{x}^*; \mathbf{y}^*\right)}{t} \leq 0\,,$$

   neboť pro každé kladné t je

$$\mathsf{L}\left(\mathbf{x}^*; \mathbf{y}^* + t\,\mathbf{e}_k\right) \leq \mathsf{L}\left(\mathbf{x}^*; \mathbf{y}^*\right)\,.$$

   Pokud $y_k^* > 0$, pak platí ještě

$$\frac{\partial \mathsf{L}}{\partial y_k}(\mathbf{x}^*; \mathbf{y}^*) = \lim_{t \to 0-} \frac{\mathsf{L}\left(\mathbf{x}^*; \mathbf{y}^* + t\,\mathbf{e}_k\right) - \mathsf{L}\left(\mathbf{x}^*; \mathbf{y}^*\right)}{t} \geq 0\,.$$

   Tím je splnění podmínky (5.13) dokázáno.

Q.E.D.

Věta 5.12.: *Nechť $f, g_1, g_2, \ldots, g_{\mathsf{m}}$ jsou konvexní a diferencovatelné. Pak jsou podmínky LPO nutné a postačující k tomu, aby dvojice $\mathbf{x}^*, \mathbf{y}^*$ splňovala GPO pro úlohu (5.1).*

Důkaz: Nutnost podmínek již byla ukázána ve větě 5.11. Stačí proto ukázat pouze postačitelnost.

Pro $\mathbf{y}^* \geq \mathbf{0}$ je funkce $\mathsf{L}\left(\bullet; \mathbf{y}^*\right)$ konvexní na $\mathbb{R}^{\mathsf{n}}$. Podle (2.9) a (5.12) upravujeme postupně

$$\begin{aligned} \mathsf{L}\left(\mathbf{x}; \mathbf{y}^*\right) &\geq \mathsf{L}\left(\mathbf{x}^*; \mathbf{y}^*\right) + \left(\mathbf{x} - \mathbf{x}^*\right)^\top \nabla_x \mathsf{L}\left(\mathbf{x}^*; \mathbf{y}^*\right) \\ &\geq \mathsf{L}\left(\mathbf{x}^*; \mathbf{y}^*\right) + \mathbf{x}^\top \nabla_x \mathsf{L}\left(\mathbf{x}^*; \mathbf{y}^*\right) \geq \mathsf{L}\left(\mathbf{x}^*; \mathbf{y}^*\right) \quad \forall\, \mathbf{x} \geq \mathbf{0} \end{aligned}$$

Jako funkce $\mathbf{y}$ je $\mathsf{L}(\mathbf{x}^*;\bullet)$ lineární, tedy

$$\begin{aligned}\mathsf{L}(\mathbf{x}^*;\mathbf{y}) &= \mathsf{L}(\mathbf{x}^*;\mathbf{y}^*) + (\mathbf{y}-\mathbf{y}^*)^\top \nabla_y \mathsf{L}(\mathbf{x}^*;\mathbf{y}^*) \\ &= \mathsf{L}(\mathbf{x}^*;\mathbf{y}^*) + \mathbf{y}^\top \nabla_y \mathsf{L}(\mathbf{x}^*;\mathbf{y}^*) \le \mathsf{L}(\mathbf{x}^*;\mathbf{y}^*) \quad \forall \mathbf{y} \ge \mathbf{0}\end{aligned}$$

podle (5.13).

Tím je splnění GPO ukázáno.

Q.E.D.

Pro symetrickou úlohu NLP používáme Lagrangeovu funkci ve tvaru (5.2), GPO vyjádřené jako (5.4) a LPO zapsané v podmínkách (5.12), (5.13). Lagrangeovu funkci, GPO a LPO lze využít i pro pozměněný tvar úlohy NLP. Avšak Lagrangeova funkce a následně i podmínky budou mít pozměněný tvar. Budou podstatně záviset na tvaru úlohy NLP. Uveďme si v následující tabulce některé z používaných možností:

omezení	Lagrangeova funkce	lokální podmínky optimality
$g_k(\mathbf{x}) \le 0 \quad \forall k$	$f(\mathbf{x}) + \sum_k y_k g_k(\mathbf{x})$	$\nabla_x \mathsf{L}(\mathbf{x}^*;\mathbf{y}^*) = \mathbf{0}$ a (5.13)
$\mathbf{x} \ge \mathbf{0}$	$f(\mathbf{x})$	$\nabla_x f(\mathbf{x}^*) \ge \mathbf{0}, \mathbf{x}^* \ge \mathbf{0}, (\mathbf{x}^*)^\top \nabla_x f(\mathbf{x}^*) = 0$
$g_k(\mathbf{x}) = 0 \quad \forall k, \mathbf{x} \ge \mathbf{0}$	$f(\mathbf{x}) + \sum_k y_k g_k(\mathbf{x})$	(5.12) a $g_k(\mathbf{x}^*) = 0 \quad \forall k$

(5.15)

5.4 Citlivost úlohy NLP

Obdoba aproximace (3.42) pro úlohu LP v kapitole 3.8.2 platí také pro symetrickou úlohu NLP. Avšak roli duálních proměnných zde přebírají Lagrangeovy multiplikátory.

Věta 5.13.: *Uvažujme symetrickou úlohu NLP*

$$\max\{f(\mathbf{x}) \,:\, g_k(\mathbf{x}) \le b_k,\ k = 1,2,\ldots,\mathsf{m},\ \mathbf{x} \ge \mathbf{0}\} \tag{5.16}$$

a úlohu, která z ní vznikne po změně pravých stran nerovností o vektor Δ

$$\max\{f(\mathbf{x}) \,:\, g_k(\mathbf{x}) \le b_k + \Delta_k,\ k = 1,2,\ldots,\mathsf{m},\ \mathbf{x} \ge \mathbf{0}\}\,. \tag{5.17}$$

Abychom mohli použít vyloženou teorii, převedeme úlohy na minimalizaci

$$\min\{-f(\mathbf{x}) \,:\, g_k(\mathbf{x}) \le b_k,\ k = 1,2,\ldots,\mathsf{m},\ \mathbf{x} \ge \mathbf{0}\}\,, \tag{5.18}$$

$$\min\{-f(\mathbf{x}) \,:\, g_k(\mathbf{x}) \le b_k + \Delta_k,\ k = 1,2,\ldots,\mathsf{m},\ \mathbf{x} \ge \mathbf{0}\}\,. \tag{5.19}$$

Nechť $\widehat{\mathbf{x}}$ *je optimální řešení úlohy (5.18) splňující GPO s Lagrangeovými multiplikátory* $\widehat{\mathbf{y}}$ *a* $\widehat{\boldsymbol{\xi}}$ *je optimální řešení úlohy (5.19) splňující GPO s Lagrangeovými multiplikátory* $\widehat{\boldsymbol{\eta}}$*. Pak platí*

$$\sum_{k=1}^{\mathsf{m}} \Delta_k \widehat{\mathbf{y}}_k \ge f(\widehat{\boldsymbol{\xi}}) - f(\widehat{\mathbf{x}}) \ge \sum_{k=1}^{\mathsf{m}} \Delta_k \widehat{\boldsymbol{\eta}}_k\,. \tag{5.20}$$

Důkaz: Stačí si pouze napsat GPO pro obě úlohy, tj. pro všechna $\mathbf{x} \ge \mathbf{0}$, $\mathbf{y} \ge \mathbf{0}$ platí

$$-f(\mathbf{x}) + \sum_{k=1}^{\mathsf{m}} \widehat{\mathbf{y}}_k (g_k(\mathbf{x}) - b_k) \ge -f(\widehat{\mathbf{x}}) \ge -f(\widehat{\mathbf{x}}) + \sum_{k=1}^{\mathsf{m}} y_k (g_k(\widehat{\mathbf{x}}) - b_k)\,,$$

$$-f(\mathbf{x}) + \sum_{k=1}^{\mathsf{m}} \widehat{\boldsymbol{\eta}}_k (g_k(\mathbf{x}) - b_k - \Delta_k) \ge -f(\widehat{\boldsymbol{\xi}}) \ge -f(\widehat{\boldsymbol{\xi}}) + \sum_{k=1}^{\mathsf{m}} y_k (g_k(\widehat{\boldsymbol{\xi}}) - b_k - \Delta_k)\,.$$