Podmínky optimality

Milan Hladík

28. června 2010

Reference

[1] Bonnans, J. F., Shapiro, A. (2000): Perturbation analysis of optimization problems, Springer, New York.

[2] Luenberger, D. G. (1969): Optimization by vector space methods, John Wiley & Sons, New York.

[3] Bazaraa, M. S., Sherali, H. D., Shetty, C. M. (1993): Nonlinear programming. Theory and algorithms. 2. ed., John Wiley & Sons, New York.

Obsah

1 **Lagrangeovy multiplikátory** 3

1.1 Karush-Kuhn-Tuckerovy podmínky 9

1.2 Zobecněné Lagrangeovy multiplikátory 11

1.3 Podmínky Fritze Johna 13

2 **Ekelandův variační princip** 14

3 **Postačující podmínky** 15

3.1 Postačující KKT-podmínky 17

3.2 Postačující podmínky Fritze Johna 17

4 **Podmínky druhého řádu** 18

1 Lagrangeovy multiplikátory

Uvažujeme úlohu

$$\min_{x \in \mathcal{Q}} f(x) \quad s.\ t. \quad G(x) \in K, \tag{P}$$

kde $f : X \to \mathbb{R}$, $G : X \to Y$ jsou spojité funkce, množiny $\mathcal{Q}$, $K \neq \emptyset$ uzavřené, konvexní, X, Y Banachovy prostory. Množina přípustných řešení

$$\Phi = \{x \in \mathcal{Q};\ G(x) \in K\} = \mathcal{Q} \cap G^{-1}(K). \tag{1}$$

Podmínky optimality mají tvar

$$x_0 \in \arg\min_{x \in \mathcal{Q}} L(x, \lambda), \quad \lambda \in N_K(G(x_0)), \tag{2}$$

kde lagrangian

$$L(x, \lambda) = f(x) + \langle \lambda, G(x_0) \rangle. \tag{3}$$

Speciálně:

A) K konvexní kužel, pak

$$\lambda \in N_K(G(x_0)) \;\Leftrightarrow G(x_0) \in K,\; \lambda \in K^-,\; \langle \lambda, G(x_0) \rangle = 0. \tag{4}$$

B) Úloha (P) konvexní, tj. f konvexní, $G(x)$ konvexní vzhledem k $-K$

$$x_0 \in \arg\min_{x \in \mathcal{Q}} L(x, \lambda) \;\Leftrightarrow\; 0 \in \partial_x L(x_0, \lambda) + N_{\mathcal{Q}}(x_0), \tag{5}$$

navíc jsou-li $f,\ G$ (Gâteaux) diferencovatelné, pak

$$x_0 \in \arg\min_{x \in \mathcal{Q}} L(x, \lambda) \;\Leftrightarrow -D_x L(x_0, \lambda) \in N_{\mathcal{Q}}(x_0), \tag{6}$$

a pokud ještě $\mathcal{Q} = X$, pak

$$x_0 \in \arg\min_{x \in \mathcal{Q}} L(x, \lambda) \;\Leftrightarrow -D_x L(x_0, \lambda) = 0. \tag{7}$$

Definice. Nechť jsou $f,\ G$ diferencovatelné. Pak λ je Langrangeův multiplikátor, splňuje-li

$$-D_x L(x_0, \lambda) \in N_{\mathcal{Q}}(x_0), \quad \lambda \in N_K(G(x_0)). \tag{8}$$

Označme:

$$\Lambda(x_0) \ldots \text{ množina všech Langrangeových multiplikátorů}$$

$$\Lambda_0 \ldots \text{ množina všech } \lambda, \text{ splňujících (2)}$$

Platí $\Lambda_0 \subseteq \Lambda(x_0)$.

Věta. Nechť úloha (P) je konvexní, x_0 optimum (P) a platí podmínka regularity

$$0 \in int\{G(\mathcal{Q}) - K\}. \tag{9}$$

Potom $\Lambda(x_0) = \Lambda_0 \neq \emptyset$, konvexní, omezená, w^*-kompaktní v Y^* a nezávisí na výběru x_0. Navíc je nulová „duality gap“.

Věta. Je-li $G(x)$ spojitě diferencovatelná funkce a konvexní vzhledem k $-K$, pak $\forall x_0 \in \Phi$ je podmínka regularity (9) ekvivalentní s „Robinson's constraint qualification“:

$$0 \in int\{G(x_0) + DG(x_0)(\mathcal{Q} - x_0) - K\} \tag{RCQ}$$

Věta. Buď x_0 lokální optimum (P), f, G spojitě diferencovatelné v x_0. Potom

(1) $h = 0$ je optimum pro úlohu

$$\min_{h \in X} Df(x_0)h \quad s.\ t. \quad h \in T_\Phi(x_0) \tag{10}$$

(2) Za předpokladu (RCQ) je $h = 0$ je optimum pro linearizovanou úlohu

$$\min_{h \in X} Df(x_0)h \quad s.\ t. \quad h \in T_{\mathcal{Q}}(x_0),\ DG(x_0)h \in T_K(G(x_0)) \tag{11}$$

Duál k (11) je

$$\max_\lambda 0 \ \ s.\,t.\ -D_xL(x_0,\lambda) \in N_{\mathcal{Q}}(x_0),\ \lambda \in N_K(G(x_0)). \tag{12}$$

a podmínka regularity pro linearizovanou úlohu (11):

$$0 \in int\{DG(x_0)T_{\mathcal{Q}}(x_0) - T_K(G(x_0))\} \tag{13}$$

Důsledek. Nechť x_0 je lokální optimum (P), $f,\ G$ jsou spojitě diferencovatelné a platí (RCQ). Potom $\Lambda(x_0) \neq \emptyset$, konvexní, omezená a w^*-kompaktní v Y^*.

Příklad. (variační počet z [2]) Úloha

$$\min \int_{t_0}^{t_1} f(x,\dot{x},t)dt \quad s.\ t. \quad G(x,t)=0, \tag{14}$$

kde $x \in D^n(t_0,t_1)$, pevné konce $x(t_0)$, $x(t_1)$, $f,\ G$ reálné funkce se spojitými parciálními derivacemi 2. řádu. Podmínka optimality má tvar diferenciální rovnice

$$f_x(x,\dot{x},t)+\lambda(t).\Phi_x(x,\dot{x},t)=\frac{d}{dt}f_{\dot{x}}(x,\dot{x},t). \tag{15}$$

<u>Další kvalifikační omezení</u>:

$$\bullet\ (RCQ) \Leftrightarrow DG(x_0)(\mathcal{R}_{\mathcal{Q}}(x_0))-\mathcal{R}_K(G(x_0))=Y \tag{16}$$

$intK \neq \emptyset$, pak:

$$\bullet\ (RCQ) \Leftrightarrow \exists h \in \mathcal{Q} : G(x_0)+DG(x_0)(h-x_0) \in intK \tag{17}$$

Kvalifikační omezení pro $\mathcal{Q} = X$:

$$\bullet\ (RCQ) \equiv 0 \in int\{G(x_0) + DG(x_0)X - K\} \tag{18}$$

$$\text{je-li } G(x_0) \in K : \tag{19}$$

$$\bullet\ (RCQ) \Leftrightarrow 0 \in core\{G(x_0) + DG(x_0)X - K\} \tag{20}$$

$$\bullet\ (RCQ) \Leftrightarrow DG(x_0)X \pm \mathcal{R}_K(G(x_0)) = Y\} \tag{21}$$

Další:

$$\bullet\ (RCQ) \Rightarrow DG(x_0)X \pm T_K(G(x_0)) = Y \tag{22}$$

$$\bullet\ (22) \Rightarrow [DG(x_0)X]^{\perp} \cap N_K(G(x_0)) = \{0\} \tag{23}$$

$$\bullet\ (23) \Leftrightarrow cl[DG(x_0)X \pm T_K(G(x_0))] = Y \tag{24}$$

$$\bullet\ (23) \Leftrightarrow cl[DG(x_0)X \pm \mathcal{R}_K(G(x_0))] = Y \tag{25}$$

$$\bullet\ \text{pokud } ri\{DG(x_0)X - \mathcal{R}_K(G(x_0))\} \neq \emptyset, \text{ pak:}$$

$$(RCQ) \Leftrightarrow (22) \Leftrightarrow (23) \Leftrightarrow (24) \Leftrightarrow (25)$$

$$\bullet\ DG(x_0) \text{ je „na“ } \Rightarrow (RCQ) \tag{26}$$

1.1 Karush-Kuhn-Tuckerovy podmínky

Mějme úlohu

$$\min_{x \in \mathbb{R}^n} f(x) \quad s.\ t. \quad \boldsymbol{h}(x) = 0_m,\ \boldsymbol{g}(x) \leq 0_l, \tag{P_n}$$

kde $f,\ g,\ h$ jsou funkce dvakrát spojitě diferencovatelné. Lagrangian:

$$L(x, \nu, \lambda) = f(x) + \nu.\boldsymbol{h}(x) + \lambda.\boldsymbol{g}(x).$$

Podmínky optimality:

$$\begin{aligned} \nabla f(x_0) + \nu.\nabla \boldsymbol{h}(x_0) + \lambda.\nabla \boldsymbol{g}(x_0) = 0, \quad \lambda \geq 0, \\ \boldsymbol{h}(x_0) = 0, \quad \boldsymbol{g}(x_0) \leq 0, \quad \lambda.\boldsymbol{g}(x_0) = 0. \end{aligned} \tag{KKT}$$

Množina aktivních podmínek:

$$I(x_0) = \{i \in \{1, \ldots, l\} :\ g_i(x_0) = 0\}. \tag{27}$$

Kvalifikační omezení:

$$(28) \Rightarrow (RCQ) \Leftrightarrow (30) \Leftrightarrow (31) \Rightarrow (32) \Rightarrow (33)$$

$$(29) \Uparrow$$

kde

- lineární nezávislost: (28)

  $$\nabla g_i(x_0),\ \forall i \in I(x_0),\quad \nabla h_j(x_0)\ \forall j \text{ jsou l.n.}$$

- Slater: (29)

  $$g_i(x),\ i \in I(x_0),\ \text{pseudokonvexní },h_j(x) \text{ kvazilineární}$$

  $$\nabla h_j(x_0)\ \forall j \text{ jsou l.n. a } \exists x \in X : \boldsymbol{g}(x) < 0,\ \boldsymbol{h}(x) = 0$$

- Mangasarian-Fromovitz: (30)

  $$Dh_j(x_0)\ \forall j \text{ jsou lin. nezávislá a } \exists s \in X :$$

  $$Dh_j(x_0)s = 0\ \forall j,\ Dg_i(x_0)s < 0\ \forall i \in I(x_0)$$

- Cottle: (31)
- Karush-Kuhn-Tucker: (32)
- Abadie: (33)

1.2 Zobecněné Lagrangeovy multiplikátory

Zobecněný lagrangian:

$$L^g(x,\alpha,\lambda) = \alpha.f(x) + \langle \lambda, G(x_0)\rangle, \quad \alpha \geq 0. \tag{34}$$

Zobecněné podmínky optimality:

$$\begin{gathered} -D_x L^g(x_0,\alpha,\lambda) \in N_{\mathcal{Q}}(x_0), \\ \lambda \in N_K(G(x_0)), \quad \alpha \geq 0, \quad (\alpha,\lambda) \neq (0,0). \end{gathered} \tag{35}$$

Označme:

$$\Lambda^g(x_0) \text{ ... množina zobecněných Lagr. multiplikátorů} \tag{36}$$

$$\Lambda^s(x_0) = \{\lambda \in Y^* : \ (0,\lambda) \in \Lambda^g(x_0)\} \tag{37}$$

$$\text{... množina singulárních Lagr. multiplikátorů}$$

Věta. Buď $\Lambda(x_0) \neq \emptyset$. Pak $\Lambda^s(x_0) \cup \{0\}$ tvoří recesní kužel $\Lambda(x_0)$.

Uvažujme kužel:

$$Z(x) \equiv DG(x)[\mathcal{R}_{\mathcal{Q}}(x)] - \mathcal{R}_K(G(x)). \tag{38}$$

Věta. Buď x_0 lokální optimum (P), pak

(1) $\Lambda^s(x_0) \neq \emptyset \Leftrightarrow \quad cl[Z(x_0)] \neq Y$,

(2) $ri\{Z(x_0)\} \neq \emptyset$, pak platí (RCQ) nebo $\Lambda^s(x_0) \neq \emptyset$.

Důsledek. Buď x_0 lokální optimum (P), $ri\{Z(x_0)\} \neq \emptyset$, pak $\Lambda^g(x_0) \neq \emptyset$.

Věta. Buď $\mathcal{Q} = X$. Pak $riZ(x_0) \neq \emptyset$, pokud platí alespoň jedno z následujících tvrzení:

(1) $intK \neq \emptyset$,

(2) $DG(x_0)$ je „na“,

(3) Y je konečněrozměrný,

(4) $DG(x_0)X$ jako podprostor Y má konečnou kodimenzi.

1.3 Podmínky Fritze Johna

Uvažujme úlohu (P_n). Zobecněný lagrangian:

$$L^g(x, \alpha, \nu, \lambda) = \alpha.f(x) + \nu.\boldsymbol{h}(x) + \lambda.\boldsymbol{g}(x).$$

Podmínky optimality:

$$\begin{aligned} &\alpha.\nabla f(x_0) + \nu.\nabla \boldsymbol{h}(x_0) + \lambda.\nabla \boldsymbol{g}(x_0) = 0, \quad (\alpha, \lambda) \geq (0, 0), \\ &(\alpha, \nu, \lambda) \neq (0, 0, 0), \; \boldsymbol{h}(x_0) = 0, \quad \boldsymbol{g}(x_0) \leq 0, \quad \lambda.\boldsymbol{g}(x_0) = 0. \end{aligned} \tag{FJ}$$

Věta. Buď x_0 lokální optimum (P_n). Potom $\Lambda^g(x_0) \neq \emptyset$.

2 Ekelandův variační princip

Věta. Buď (V, ρ) úplný metrický prostor, $f : V \mapsto \mathbb{R} \cup \{+\infty\}$ zdola polospojitá funkce. Předpokládejme, že $\inf_{x \in V} f(x) \in \mathbb{R}$ a pro dané $\varepsilon > 0$ buď $x_\varepsilon \in V$ ε-minimizer (tj. $f(x_\varepsilon) \leq \inf_{x \in V} f(x) + \varepsilon$). Potom $\forall k > 0\ \exists x_k \in V$:

(1) $\rho(x_\varepsilon, x_k) \leq k^{-1}$,

(2) $f(x_k) \leq f(x_\varepsilon) - \varepsilon.k.\rho(x_\varepsilon, x_k)$,

(3) $f(x) > f(x_k) - \varepsilon.k.\rho(x_k, x)\ \forall x \in V,\ x \neq x_k$.

Věta. Buď x_ε ε-minimizer úlohy (P) pro $\varepsilon \geq 0$, f, G spojitě diferencovatelné, (RCQ) platí $\forall x \in \overline{B}(x_\varepsilon, \sqrt{\varepsilon})$. Pak existuje další $\hat{x}$ ε-minimizer úlohy (P), vyhovující Ekelandově principu ($k = \varepsilon^{-\frac{1}{2}}$), a $\lambda \in Y^*$ takové, že:

$$dist\{-D_x L(\hat{x}, \lambda), N_{\mathcal{Q}}(\hat{x})\} \leq \sqrt{\varepsilon}, \quad \lambda \in N_K(G(\hat{x})). \qquad (39)$$

3 Postačující podmínky

Funkce $f,\ G$ spojitě diferencovatelné.

Definice. Buď $\emptyset \neq S \subseteq \Phi$ tak, že $f(x)/_S = f_0 \in \mathbb{R}$ a buď $\gamma > 0$. Pak je splněna „γ-order growth condition“ na S, jestliže existuje $c > 0$ a okolí N množiny S tak, že $\forall x \in \Phi \cap N$ platí:

$$f(x) \geq f_0 + c.[dist(x, S)]^{\gamma}. \tag{40}$$

Definujme množinu

$$\Upsilon_{\eta}(x_0) = \{h \in T_{\mathcal{Q}}(x_0) :\ dist(DG(x_0)h,\ T_K(G(x_0))) \leq \eta.\|h\|\}, \tag{41}$$

pro $\eta = 0$:

$$\Upsilon_0(x_0) = \{h \in T_{\mathcal{Q}}(x_0) :\ DG(x_0)h \in T_K(G(x_0))\}. \tag{42}$$

Věta. Buď $x_0 \in \Phi$.

(1) Jestliže existují $\alpha > 0,\ \eta > 0$ tak, že

$$Df(x_0)h \geq \alpha\|h\|, \quad \forall h \in \Upsilon_{\eta}(x_0), \tag{43}$$

pak v x_0 platí „first order growth condition“.

(2) Jestliže je splněna podmínka regularity

$$DG(x_0)(T_{\mathcal{Q}}(x_0)) - T_K(G(x_0)) = Y \tag{44}$$

a (43) pro nějaké $\alpha > 0$ a $\eta = 0$, pak platí „first order growth condition“.

(3) Nechť platí (RCQ) pro x_0. Pak existuje $\alpha > 0$ tak, že (43) je splněno pro $\eta = 0$ právě tehdy, když platí „first order growth condition“.

Věta. Pro $\eta = 0$ podmínka (43) je ekvivalední s

$$-Df(x_0) \in int\{\Upsilon_0(x_0)^-\}. \tag{45}$$

Za platnosti podmínky regularity (44) je (43) ekvivalentní s

$$-Df(x_0) \in int\{[DG(x_0)]^*(N_K(G(x_0))) + N_{\mathcal{Q}}(x_0)\}. \tag{46}$$

Pro X konečněrozměrné je postačující podmínka tvaru

$$Df(x_0)h > 0 \; \forall h \in T_{\mathcal{Q}}(x_0) \setminus \{0\} \; s.t. \; DG(x_0)h \in T_K(G(x_0)), \tag{47}$$

tj. $h = 0$ je jedinečné řešení linearizovaného problému (11).

3.1 Postačující KKT-podmínky

Věta. Nechť $x_0 \in X$ vyhovuje podmínkám optimality (KKT). Označme $J_+ = \{j: \ \nu_j > 0\}$, $J_- = \{j: \ \nu_j < 0\}$. Předpokládejme, že $f(x)$ je pseudokonvexní v x_0, $g_i(x)$, $i \in I(x_0)$, kvazikonvexní v x_0, $h_j(x)$, $j \in J_+$, kvazikonvexní v x_0 a $h_j(x)$, $j \in J_-$, kvazikonkávní v x_0, potom x_0 je globální optimum úlohy (P_n).

3.2 Postačující podmínky Fritze Johna

Věta. Nechť $x_0 \in X$ vyhovuje podmínkám optimality (FJ). Definujme $S = \{x: \ g_i(x) \leq 0 \ \forall i \in I(x_0), \ h_j(x) = 0 \ \forall j\}$. Jestliže $h_j(x)$ jsou afinní, $\nabla h_j(x_0)$ jsou lineárně nezávislé a existuje ε-okolí $N_\varepsilon(x_0)$, $\varepsilon > 0$, tak, že $f(x)$ je pseudokonvexní na $S \cap N_\varepsilon(x_0)$ a $g_i(x)$, $i \in I(x_0)$, jsou striktně pseudokonvexní na $S \cap N_\varepsilon(x_0)$, potom x_0 je lokální minimum úlohy (P_n).

4 Podmínky druhého řádu

Předpokládejme $\mathcal{Q} = X$. Definujme pro $x \in \Phi$:

$$C(x) \equiv \{h \in X : Df(x)h \leq 0,\ DG(x_0)h \in T_K(G(x_0))\}. \tag{48}$$

Věta. Buď x_0 lokální optimum (P) a nechť v x_0 platí (RCQ). Pak $\forall h \in C(x_0)$ a pro každou konvexní množinu $\mathcal{T}(h) \subset T_K^2(G(x_0), DG(x_0)h)$ platí:

$$\sup_{\lambda \in \Lambda(x_0)} \{D_{xx}^2 L(x_0, \lambda)(h, h) - \sigma(\lambda, \mathcal{T}(h))\} \geq 0, \tag{49}$$

kde „outer second order tangent set“

$$T_S^2(x_0, h) = \limsup_{t \downarrow 0} \frac{S - x_0 - th}{\frac{1}{2}t^2}. \tag{50}$$

Příklad. Mějme $S = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 \geq |x_1|^{\frac{3}{2}}\}$, $x_0 = (0, 0)$, $h = (1, 0)$. Pak

$$h \in T_S(x_0) = \{x_2 \geq 0\},$$

$$T_S^2(x_0, h) = \emptyset.$$

Příklad. Mějme $S = \{x \in X \mid g(x) \le 0\}$, funkce $g(x)$ je konvexní, spojitá, $g'(x_0, h) = 0$, $g(x_0) = 0$ a $\exists \hat{x} : g(\hat{x}) < 0$. Pak

$$T_S^2(x_0, h) = \{w \in X \mid g''_-(x_0, h, w) \le 0\},$$

kde

$$g''_-(x_0, h, w) = \liminf_{t \downarrow 0} \frac{g(x + th + \frac{1}{2}t^2 w) - g(x) - t.g'(x, h)}{\frac{1}{2}t^2}. \tag{51}$$

Příklad. Mějme $S = \{x \in X \mid h_j(x) = 0 \ j = 1, \ldots, m, \ g_i(x) \le 0 \ i = 1, \ldots . l\}$, funkce $h_j(x)$, $g_i(x)$ jsou dvakrát spojitě diferencovatelné a pro $x_0 \in S$ předpokládejme platnost „Mangasarian-Fromovitz constraint qualification“ (30). Pak

$$T_S(x_0) = \{h \in X \mid Dh_j(x_0)h = 0 \ \forall j, \ Dg_i(x_0)h \le 0 \ \forall i \in I(x_0)\},$$

$$\begin{aligned} T_S^2(x_0, h) = \{w \in X \mid\ & Dh_j(x_0)w + D^2 h_j(x_0)(h, h) = 0 \ \forall j, \\ & Dg_i(x_0)w + D^2 g_i(x_0)(h, h) \le 0 \ \forall i \in I_1(x_0, h)\}, \end{aligned}$$

kde

$$I_1(x_0, h) = \{i \in I(x_0) \mid Dg_i(x_0)h = 0\}.$$

Pro úlohu (P_n) je pro $x \in \Phi$:

$$\begin{aligned} C(x) \equiv \{h \in X : Df(x)h \le 0,\ Dh_j(x)h = 0\ \forall j, \\ Dg_i(x)h \le 0\ \forall i \in I(x)\}. \end{aligned} \tag{52}$$

Věta. Buď $x_0 \in \Phi$.

(1) Je-li x_0 lokální optimum úlohy (P_n), pak $\forall h \in C(x_0)\ \exists(\alpha, \lambda) \in \Lambda^g(x_0)$ tak, že

$$D^2_{xx} L^g(x_0, \alpha, \lambda)(h, h) \ge 0. \tag{53}$$

(2) Jestliže $\forall h \in C(x_0) \setminus \{0\}\ \exists(\alpha, \lambda) \in \Lambda^g(x_0)$ tak, že

$$D^2_{xx} L^g(x_0, \alpha, \lambda)(h, h) > 0, \tag{54}$$

pak x_0 je lokální optimum úlohy (P_n) splňující „second order growth condition“.