

Nutné a postačující podmínky optimality

```
$$ \xdef\scal#1#2{\langle #1, #2 \rangle} \xdef\norm#1{\left\| \mathrm{#1} \right\|} \xdef\dist{\rho} \xdef\and{\&} \xdef\brackets#1{\left\{ #1 \right\}} \xdef\parc#1#2{\frac {\partial #1}{\partial #2}} \xdef\mtr#1{\begin{pmatrix} #1 \end{pmatrix}} \xdef\bm#1{\boldsymbol{#1}} \xdef\mcal#1{\mathcal{#1}} \xdef\vv#1{\mathbf{#1}} \xdef\vvp#1{\pmb{#1}} \xdef\ve{\varepsilon} \xdef\l{\lambda} \xdef\th{\vartheta} \xdef\a{\alpha} \xdef\Tagged#1{(\text{#1})} \xdef>tagged*#1{\text{#1}} \xdef>tagEqHere#1#2{\href{#2#eq-#1}{(\text{#1})}} \xdef>tagDeHere#1#2{\href{#2#de-#1}{\text{#1}}} \xdef>tagEq#1{\href{#eq-#1}{(\text{#1})}} \xdef>tagDe#1{\href{#de-#1}{\text{#1}}} \xdef\T#1{\htmlId{eq-#1}{#1}} \xdef\D#1{\htmlId{de-#1}{#1}} \xdef\conv#1{\mathrm{conv}}, #1 \xdef\cone#1{\mathrm{cone}}, #1 \xdef\aff#1{\mathrm{aff}}, #1 \xdef\lin#1{\mathrm{Lin}}, #1 \xdef\span#1{\mathrm{span}}, #1 \xdef\O{\mathcal O} \xdef\ri#1{\mathrm{ri}}, #1 \xdef\rd#1{\mathrm{r}} \partial, #1 \xdef\interior#1{\mathrm{int}}, #1 \xdef\proj{\Pi} \xdef\epi#1{\mathrm{epi}}, #1 \xdef\grad#1{\mathrm{grad}}, #1 \xdef\gradT#1{\mathrm{grad}}^T #1 \xdef\gradx#1{\mathrm{grad}}_x #1 \xdef\hess#1{\nabla^2, #1} \xdef\hessx#1{\nabla^2_x #1} \xdef\subdif#1{\partial #1} \xdef\co#1{\mathrm{co}}, #1 \xdef\iter#1{\wedge^{#1}} \xdef\str{\wedge^*} \xdef\spv{\mathrm{V}} \xdef\civ{\mathrm{U}} $$
```

Obecný úvod

Úlohou matematického programování nazveme $f(x) \rightarrow \min, \quad x \in X$ $\tag{T{4.1}}$ kde **přípustná množina** X je zadána systém rovností a nerovností $X := \{x \in P \mid g_i(x) \leq 0, \quad i = 1, \dots, k; \quad g_j(x) = 0, \quad j = k+1, \dots, m\}$ $\tag{T{4.2}}$

“ Zde je důležité podotknout, že vždy chceme nerovnostní omezení **pouze tvaru** $\mathbf{g}_i(x) \leq 0$ ”

“ Je dobré si zapatovat, že

- m - počet omezení

- k - počet nerovností (tzn. g_1, \dots, g_k jsou **nerovnosti**, zbytek rovnosti)

Omezení zakomponovaná v P se nazývají **přímá**, naopak omezení ve formě g_i se nazývají **funkcionální**. Dále definujeme

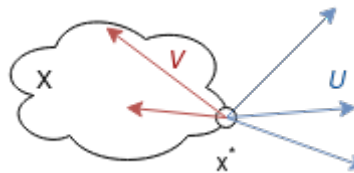
- **množinu přípustných vektorů**

$$\text{spv}(x^*, X) := \{h \in \mathbb{R}^n \mid \exists \lambda, \lambda_0 > 0 : x^* + th \in X \text{ pro } t \in (0, \lambda_0)\}$$

- je to **kužel**

- **množinu spádových vektorů (kužel zlepšujících vektorů)**

$$\text{civ}(x^*, f) := \{h \in \mathbb{R}^n \mid \exists \lambda, \lambda_0 > 0 : x^* + th \in D(f); \text{ a } f(x^* + th) < f(x^*) \text{ pro } t \in (0, \lambda_0)\}$$



Lemma 4.1.1

Nechť $X \subseteq \mathbb{R}^n$ a $f: X \rightarrow \mathbb{R}$ jsou dány. Je-li bod $x^* \in X$ lokálním řešením úlohy 4.1, potom $\text{spv}(x^*, X) \cap \text{civ}(x^*, f) = \emptyset$

Definice 4.1.2 (Stacionární bod)

Nechť množina $X \subseteq \mathbb{R}^n$ je **konvexní** a funkce $f: X \rightarrow \mathbb{R}$ je **diferencovatelná** na (nějaké otevřené množině obsahující) X . Řekneme, že bod $x^* \in X$ je **stacionárním bodem** úlohy 4.1 (nebo také **stacionárním bodem funkce** f na množině X), jestliže $\nabla f(x^*)^T (x - x^*) \geq 0$ pro každé $x \in X$.

“ Výraz v 4.1.2 je **směrová derivace** x^* do libovolného bodu v X - v těchto směrech musí být f **neklesající**

Pro $X = \mathbb{R}^n$ je podmínka 4.1.2 splněna **pouze** v případě $\nabla f(x^*) = 0$

Dále ukažme, že **stacionární bod** ve smyslu Definice 4.1.2 má vlastnosti, které od něj očekáváme.

Věta 4.1.3 (Vlastnosti stacionárního bodu)

Nechť $f: X \rightarrow \mathbb{R}$ je **diferencovatelná** na (nějaké otevřené množině obsahující **konvexní** množinu) $X \subseteq \mathbb{R}^n$.

1. Je-li $x^* \in X$ **lokálním extrémem** funkce f na X (tj. *lokálním* řešením úlohy $\tag{4.1}$), pak x^* je **stacionárním bodem** funkce f na X
2. Naopak, je-li f (*ostře*) **konvexní** na X a $x^* \in X$ je **stacionárním bodem** f na X , pak x^* je (*jediným*) řešením úlohy $\tag{4.1}$, tj. (*jediným*) **globálním minimem** f na X .

Pokud avšak f **není konvexní**, potřebujeme o rozhodnutí "extrémnosti" stacionárního bodu další nástroje

Nutná podmínka pro $\tag{4.1}$

Je-li $x^* \in X$ **lokálním minimem** funkce $f: X \rightarrow \mathbb{R}$, $f \in C^2$, na **konvexní** množině $X \subseteq \mathbb{R}^n$, pak $(x - x^*)^T \nabla f(x^*) (x - x^*) \geq 0$ pro všechna $x \in X$ taková, že $\nabla f(x^*) (x - x^*) = 0$, tj. pro vektory $(x - x^*) \in \ker \nabla f(x^*)$

Postačující podmínka pro $\tag{4.1}$

Bod x^* je **lokálním minimem** funkce $f: X \rightarrow \mathbb{R}$, $f \in C^2$, na **konvexní** množině $X \subseteq \mathbb{R}^n$, jestliže $\nabla f(x^*) (x - x^*) \geq 0$, $\forall x \in X$, (tj. je to **stacionární bod** ve smyslu Definice $\tag{4.1.2}$), množina X je **polyedr** a platí $(x - x^*)^T \nabla f(x^*) (x - x^*) > 0$ pro všechna $x \in X$ taková, že $x \neq x^*$ a $(x - x^*) \in \ker \nabla f(x^*)$

“ Je-li X **polyedr**, pak je $\text{spv}(x^*, X)$ **uzavřená**

Nutné a postačující podmínky optimality

Uvažujme přidruženou **Lagrangeovu funkci** $L: \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}$ k úloze $\tag{4.1}$ \wedge $\tag{4.2}$ definovanou jako $L(x, y_0, y) := y_0 f(x) + \sum_{i=1}^m y_i g_i(x)$, $\tag{4.1.2}$ ∇ přičemž v případě $y_0 = 1$ bude $L(x, 1, y) := L(x, y)$. Navíc čísla y_0, \dots, y_m nazýváme **Lagrangeovými multiplikátory**.

“ Omezení nazveme **aktivní**, pokud se realizuje jako **rovnost**

Dále ještě zavedme následující $Q := \{y = (y_1, \dots, y_m)^T \mid y_1, \dots, y_k \geq 0\}$ a dvě další množiny:

- **Množinu aktivních omezení** v bodě $x \in X$
 $I(x) := \{i \in \{1, \dots, k\} \mid g_i(x) = 0\}$, $\forall x \in X$

- **Množinu indexů** všech funkcí, které se v bodě x realizují jako rovnost $S(x) := \bigcup_{k=1, \dots, m} \{x \in X \mid g_k(x) = 0\}$

Věta D{4.2.1} (Lagrangeův princip)

Nechť množina $P \subseteq \mathbb{R}^n$ je **konvexní**, funkce $f, g_1, \dots, g_k : P \rightarrow \mathbb{R}$ jsou **diferencovatelné** v bodě $x^* \in X$ a g_{k+1}, \dots, g_m jsou **spojitě diferencovatelné** na nějakém okolí bodu x^* . Je-li bod $x^* \in X$ **lokálním řešením** úlohy $\tag{Eq{4.1}} \quad \text{minimize } f(x) \quad \text{subject to } g_i(x) \leq 0, \quad i = 1, \dots, m$, pak existují **Lagrangeovy multiplikátory** $y_0 > 0$ a $y \in Q$ takové, že **ne všechna** y_0, \dots, y_m jsou **nulová** a platí $\text{grad}_x L(x, y_0, y) = 0$ a $y_i g_i(x) = 0, \quad i = 1, \dots, m$ $\tag{T{4.2.3}}$ $\tag{T{4.2.4}}$

Podmínka $\tag{Eq{4.2.3}}$ znamená, že x musí být **stacionárním bodem** funkce $L(x, y_0, y)$ (*podmínka stacionarity*). Dále podmínka $\tag{Eq{4.2.4}}$ se nazývá **podmínka komplementarity** a požadavek $y_1, \dots, y_k > 0$ jako **podmínka duality**.

“Jistě $y_0, y_1, \dots, y_k \geq 0, y_{k+1}, \dots, y_m \in \mathbb{R}$ *iff* $y_0 > 0$ a $y \in Q$ ”

Jelikož situace s $y_0 = 0$ je problematická, existují podmínky na zaručení $y_0 \neq 0$, což je ekvivalentní s $y_0 = 1$. Tyto podmínky se nazývají **podmínky kvalifikovaného omezení**:

- **Regulárnost bodu** x , tj. bod x je **regulární**, pokud jsou $\text{grad } g_i(x)$ pro $i \in S(x)$ **lineárně nezávislé** (tj. *gradienty aktivních omezení jsou LNŽ*)
- **afinní omezení** - funkce g_1, \dots, g_m jsou **afinní**
- **Slaterova podmínka** - g_1, \dots, g_k jsou **konvexní**, g_{k+1}, \dots, g_m jsou **afinní**, **konstantní** vektory $\text{grad } g_i$ jsou **lineárně nezávislé** pro $i \in \{k+1, \dots, m\}$ a existuje $\bar{x} \in P$ takové, že $g_i(\bar{x}) < 0$ pro $i \in \{1, \dots, k\}$ a $g_j(\bar{x}) = 0$ pro $j \in \{k+1, \dots, m\}$

Důsledek D{4.2.2}

Nechť $P \subseteq \mathbb{R}^n$ je **konvexní** množina, funkce f, g_1, \dots, g_m jsou **diferencovatelné** na (nějaké otevřené množině obsahující) P a pro $x \in X$ existují multiplikátory $y \in Q$ takové, že platí $\tag{Eq{4.2.3}}$ & $\tag{Eq{4.2.4}}$ s $y_0 = 1$. Nechť je dále splněn (*alespoň*) jeden z následujících předpokladů:

1. funkce $L(x, y)$ je **konvexní** na množině P
2. úloha $\tag{Eq{4.1}}$ & $\tag{Eq{4.2}}$ je úlohou **konvexního programování**, tj. na **konvexní** množině P jsou funkce f, g_1, \dots, g_k **konvexní** a g_{k+1}, \dots, g_m **afinní**

Pak bod x je **globálním řešením** úlohy $\tag{Eq{4.1}}$ & $\tag{Eq{4.2}}$.

Věta D{4.2.3} (Karushova-Kuhnova-Tuckerova v diferenciálním tvaru)

Nechť $P \subseteq \mathbb{R}^n$ je **konvexní** množina, funkce f, g_1, \dots, g_k **konvexní** a **diferencovatelné** na (nějaké otevřené množině obsahující) P , funkce g_{k+1}, \dots, g_m **afinní** na P a necht' platí (alespoň) jedna z následujících podmínek:

1. **(LNZ)** množina P je **otevřená**, vektory $\nabla g_i(x)$, $i \in S(x)$ jsou **lineárně nezávislé** pro každé $x \in X$.
2. **(Slaterova)** funkcionální omezení jsou pouze tvaru **nerovností**, tj. $k = m$, a **existuje** bod $\bar{x} \in P$ takový, že $g_i(\bar{x}) < 0$ pro $i \in \{1, \dots, k\}$
3. **(lineární)** množina P je **polyedr** a funkce g_1, \dots, g_k jsou **afinní**

Pak x^* je řešením úlohy D{4.1} & D{4.2} právě tehdy, když existuje $y^* \in Q$ takové, že platí D{4.2.3} & D{4.2.4} s $y_0^* = 1$.

Věta D{4.2.4}

Nechť funkce f, g_1, \dots, g_m jsou **dvakrát spojitě** diferencovatelné v bodě x^* a $x^* \in \text{interior } P$ je takový, že **existují** multiplikátory $y^* \in Q$ splňující D{4.2.3} & D{4.2.4} s $y_0^* = 1$ a současně $y_i^* > 0$ pro $i \in I(x^*)$ (tzv. **podmínka ostré komplementarity**), tj. $\nabla_x L(x^*, y^*) = 0$, $g_i(x^*) \leq 0$ pro $i \in \{1, \dots, k\}$, $g_i(x^*) = 0$ pro $i \in \{k+1, \dots, m\}$, $y_i^* > 0$ pro $i \in I(x^*)$, $y_i^* = 0$ pro $i \in \{1, \dots, k\} \setminus I(x^*)$, $y_i^* \in \mathbb{R}$ pro $i \in \{k+1, \dots, m\}$. Jestliže $\nabla^2_x L(x^*, y^*) > 0$ na $\ker(\nabla g_i(x^*))_{i \in S(x^*)}$, tj. $h^T \nabla^2_x L(x^*, y^*) h > 0$ pro všechna $h \in \mathbb{R}^n \setminus \{0\}$ taková, že $\nabla g_i(x^*) h = 0$ pro $i \in S(x^*)$, pak bod x^* je **ostré lokální minimum** funkce f na množině X .