Podmínky optimality

Fritze Johna + Karushovy-Kuhnovy-Tuckerovy

Nechť M je množina a

nechť $f: M \to \mathbb{R}$ a $g_1, \dots, g_m : M \to \mathbb{R}$ jsou funkce.

Uvažujme optimalizační úlohu (úlohu nelineárního programování)

$$(NP) \qquad \min f(x) \qquad \text{z.p. } g_i(x) \leqq 0 \quad \text{pro } i = 1, \dots, m$$

<u>Def.:</u> Říkáme, že bod $x^* \in M$ je globálním řešením úlohy (NP) $\Longleftrightarrow$ $g_i(x^*) \leqq 0$ pro $i = 1, \dots, m$

a

$$\forall x \in M : (g_i(x) \leqq 0 \text{ pro } i = 1, \dots, m) \implies f(x) \geqq f(x^*)$$

<u>Def.:</u> Říkáme, že bod $x^* \in M$ je lokálním řešením úlohy (NP) $\Longleftrightarrow$ $g_i(x^*) \leqq 0$ pro $i = 1, \dots, m$

a

$$\exists \delta > 0 \ \ \forall x \in M : (g_i(x) \leqq 0 \text{ pro } i = 1, \dots, m) \wedge \|x - x^*\| < \delta \implies f(x) \geqq f(x^*)$$

<u>☺</u>: jestliže bod $x^* \in M$ je globálním řešením úlohy (NP),

potom je jejím lokálním řešením.

<u>Motivace:</u> Chceme stanovit podmínky nutné (a postačující) pro to,

aby bod $\underline{x^*}$ byl globálním (popř. lokálním) řešením úlohy (NP).

<u>Poznámka:</u> úloha lineárního programování

$$(P) \qquad \max c^T x \qquad \text{z.p. } Ax \leqq b$$

je speciálním případem úlohy (NP):

- stačí položit $f(x) = -c^T x$

  a $g_i(x) = a_i x - b_i$ pro $i = 1, \dots, m$.

(klademe $M = \mathbb{R}^n$ ... celý prostor)

Lemma: (klíčové)

Nechť M je otevřená množina

a $f: M \to \mathbb{R}$ a $g_1, \dots, g_m : M \to \mathbb{R}$ nechť jsou libovolné funkce.

Bod $x^* \in M$ budiž lokálním řešením úlohy

$$(NP) \qquad \min f(x)$$
$$\text{s.p.} \quad g_i(x) \leqq 0 \qquad \text{pro } i = 1, \dots, m$$

Nechť

$I_0 = \{ i \in \{1, \dots, m\} \; ; \; g_i(x^*) = 0 \}$ ..... indexy podmínek aktivních v bodě $\underline{x}^*$

dále položme

$J = \{ i \in I_0 \; ; \; g_i \text{ ... je lokálně pseudokonkávní v } \underline{x}^* \}$

a $\quad I = I_0 \setminus J$ ..... ostatní indexy

Funkce g_i pro $i \in \{1, \dots, m\} \setminus I_0$ ..... neaktivní podmínky ..... budiž spojité v bodě $\underline{x}^*$

[stačí: g_i ... polospojité shora v $\underline{x}^*$ pro $i \in \{1, \dots, m\} \setminus I_0$]

Potom soustava

$$(*) \qquad \underline{D}_v^+ (-f)(x^*) > 0$$
$$\underline{D}_v^+ (-g_i)(x^*) > 0 \qquad \text{pro } i \in I$$
$$\underline{D}_v^+ (-g_i)(x^*) \geqq 0 \qquad \text{pro } i \in J$$

..... kde $\underline{v}$ je neznámá

..... nemá řešení.

Důkaz: (~~sporem~~ nepřímo)

– předpokládejme, že soustava

$$\underline{D}_v^+ -f(x^*) > 0$$
$$\underline{D}_v^+ -g_i(x^*) > 0 \qquad \text{pro } i \in I$$
$$\underline{D}_v^+ -g_i(x^*) > 0 \qquad \text{pro } i \in J$$

... má řešení

– protože funkce g_i jsou polospojité shora v bodě $\underline{x}^*$ pro $i \in \{1,\dots,m\} \setminus I_0$,

k $\varepsilon' = \frac{1}{2} \min \{ -g_i(x^*) \ ; \ i \in \{1,\dots,m\} \setminus I_0 \} > 0$ $\quad$ $\textcircled{\scriptsize ?}: \varepsilon' > 0$

$\exists \delta' > 0 \ \forall \lambda > 0 : \ 0 < \lambda < \delta' \Rightarrow g_i(x^* + \lambda v) < \underbrace{g_i(x^*) + \varepsilon'}_{<0}$ $\quad$ pro $i \in \{1,\dots,m\} \setminus I_0$

tudíž $g_i(x^* + \lambda v) < 0$ $\quad$ pro $i \in \{1,\dots,m\} \setminus I_0$

– protože $\underline{D}_v^+ (-f)(x^*) = \liminf\limits_{\lambda \to 0^+} \dfrac{-f(x^* + \lambda v) + f(x^*)}{\lambda} = A > 0$,

k $\varepsilon'' = \frac{A}{2} > 0$

$\exists \delta'' > 0 \ \forall \lambda > 0 : \ 0 < \lambda < \delta'' \Rightarrow \dfrac{-f(x^* + \lambda v) + f(x^*)}{\lambda} > \underbrace{A - \varepsilon''}_{>0}$

tudíž $f(x^* + \lambda v) < f(x^*)$

– protože $\underline{D}_v^+ (-g_i)(x^*) = \liminf\limits_{\lambda \to 0^+} \dfrac{-g_i(x^* + \lambda v) + g_i(x^*)}{\lambda} = A_i > 0$ $\quad$ pro $i \in I$

k $\varepsilon''' = \frac{1}{2} \min \{ A_i \ ; \ i \in I \} > 0$

$\exists \delta''' > 0 \ \forall \lambda > 0 : \ 0 < \lambda < \delta''' \Rightarrow \dfrac{-g_i(x^* + \lambda v) + g_i(x^*)}{\lambda} > A_i - \varepsilon'''$ $\quad$ pro $i \in I$

tudíž $g_i(x^* + \lambda v) < g_i(x^*) = 0$ $\quad$ pro $i \in I$

– protože g_i jsou pseudokonkávní v bodě $\underline{x}^*$ & $\underline{D}_v^+ (-g_i)(x^*) \geqq 0$ $\quad$ pro $i \in J$

neboli $(-g_i)$ jsou pseudokonvexní v bodě $\underline{x}^*$ & $\underline{D}_{\lambda v}^+ (-g_i)(x^*) \geqq 0$ $\quad$ pro libovolné $\lambda > 0$,

dostáváme $\quad -g_i(x^* + \lambda v) \geqq -g_i(x^*) = 0$

neboli $\quad g_i(x^* + \lambda v) \leqq g_i(x^*) = 0$

pro libovolné $\lambda > 0$ takové, že $x^* + \lambda v \in M$,
pro $i \in J$.

Dohromady dostáváme: bod $(x^* + \lambda v)$ je přípustný

& $f(x^* + \lambda v) < f(x^*)$ $\quad$ ~~pro všechna $\underline{\lambda}$, $0 < \lambda < \min\{\delta', \delta'', \delta'''\}$~~
kdykoliv $0 < \lambda < \min\{\delta', \delta'', \delta'''\}$,

tudíž bod $\underline{x}^*$ není lokálním řešením úlohy (NP), $\quad$ c.b.d.

Poznámka: Jestliže funkce f a g_i v bodě $\underline{x}^*$ mají jednostranné směrové derivace $\delta_v^+ f(x^*)$ a $\delta_v^+ g_i(x^*)$ pro každý směr $\underline{v}$, potom platí

$$\underline{D}_v^+ (-f)(x^*) = \delta_v^+ (-f)(x^*) = -\delta_v^+ f(x^*)$$

$$\underline{D}_v^+ (-g_i)(x^*) = \delta_v (-g_i)(x^*) = -\delta_v^+ g_i(x^*)$$

a soustava (*) získává tvar

$$\delta_v^+ f(x^*) < 0$$

$$\delta_v^+ g_i(x^*) < 0 \quad \text{pro } i \in I$$

$$\delta_v^+ g_i(x^*) \leqq 0 \quad \text{pro } i \in J$$

Je zajímavé, že klíčové lemma připouští částečné obrácení.

Lemma: (klíčové, částečně obrácené)

Nechť M je otevřená množina

$f: M \to \mathbb{R}$ a $g_1, \dots, g_m: M \to \mathbb{R}$ jsou funkce

$x^* \in M$ je bod takový, že $g_i(x^*) \leqq 0$ pro $i = 1, \dots, m$.

Označme

$$I_0 = \{ i \in \{1, \dots, m\} \; ; \; g_i(x^*) = 0 \}$$

Funkce f a g_i nechť v bodě $\underline{x}^*$ mají jednostranné směrové derivace $\delta_v^+ f(x^*)$ a $\delta_v^+ g_i(x^*)$ pro $i \in I_0$ pro každý směr $\underline{v}$.

Nechť f je lokálně pseudokonvexní v bodě $\underline{x}^*$

$$I = \{ i \in I_0 \; ; \; g_i \text{ je lokálně striktně pseudokonvexní v bodě } \underline{x}^* \}$$

$$J = \{ i \in I_0 \; ; \; g_i \text{ je lokálně kvazikonvexní v bodě } \underline{x}^* \}$$

Nechť $I \cup J = I_0$. $J \subseteq I_0 \setminus I$

Jestliže soustava

$$\delta_v^+ f(x^*) < 0$$

$$\delta_v^+ g_i(x^*) < 0 \quad \text{pro } i \in I \setminus J$$

$$\delta_v^+ g_i(x^*) \leqq 0 \quad \text{pro } i \in J$$

kde $\underline{v}$ je neznámá

nemá řešení,

potom bod $\underline{x}^*$ je bodem globálního minima v úloze

$$\text{(NP)} \quad \begin{aligned} &\min f(x) \\ &\text{z.p. } g_i(x) \leqq 0 \quad \text{pro } i = 1, \dots, m. \end{aligned}$$

Důkaz: (nepřímo)

- předpokládejme, že bod $\underline{x}^*$ není globálním řešením úlohy (NP),

tedy $\exists y \in M$, $g_i(y) \leq 0$ pro $i = 1, \dots, m$, $f(y) < f(x^*)$

$\Downarrow$

$(y \neq x^*)$

Potom:

- protože funkce f je pseudokonvexní v bodě $\underline{x}^*$ a $f(y) < f(x^*)$,

$$\delta^+_{y-x^*} f(x^*) < 0$$

- protože funkce g_i jsou striktně pseudokonvexní v bodě $\underline{x}^*$ a $g_i(y) \leq g_i(x^*) = 0$ & $y \neq x^*$,

$$\delta^+_{y-x^*} g_i(x^*) < 0 \qquad \text{pro } i \in I \setminus J$$

- protože funkce g_i jsou lokálně kvazikonvexní v bodě $\underline{x}^*$ a $g_i(y) \leq g_i(x^*) = 0$,

$$\delta^+_{y-x^*} g_i(x^*) \leq 0 \qquad \text{pro } i \in J$$

Tedy soustava (*) má řešení,

c.b.d.

Složením obou lemmat dostáváme následující

Důsledek: Nechť M je otevřená množina

$$\left.\begin{array}{l} f : M \to \mathbb{R} \\ g_1, \dots, g_m : M \to \mathbb{R} \\ a_1, \dots, a_n : M \to \mathbb{R} \end{array}\right\}$$ funkce, které mají jednostranné směrové derivace $\delta_v^+ f(x)$, $\delta_v^+ g_i(x)$, $\delta_v^+ a_j(x)$ v každém bodě $x \in M$ pro každý směr $\underline{v}$

Nechť $f : M \to \mathbb{R}$ ...... je pseudokonvexní

$$\forall x \in M \ \forall y \in M : \quad f(y) < f(x) \implies \delta^+_{y-x} f(x) < 0$$

$g_1, \dots, g_m : M \to \mathbb{R}$ ...... jsou striktně pseudokonvexní

$$\forall x \in M \ \forall y \in M, \ y \neq x : \quad g_i(y) \leqq g_i(x) \implies \delta^+_{y-x} g_i(x) < 0$$

$a_1, \dots, a_n : M \to \mathbb{R}$ ...... jsou lokálně kvazikonvexní & pseudokonkávní

$$\forall x \in M \ \forall y \in M : \quad a_j(y) \leqq a_j(x) \iff \delta^+_{y-x} a_j(x) \leqq 0$$

Nechť $$\left.\begin{array}{l} g_1, \dots, g_m : M \to \mathbb{R} \\ a_1, \dots, a_n : M \to \mathbb{R} \end{array}\right\}$$ jsou polospojité shora.

Zvolme
Uvažujme bod $x^* \in M$ takový, že $g_i(x^*) \leqq 0$ pro $i = 1, \dots, m$ a $a_j(x^*) \leqq 0$ pro $j = 1, \dots, n$,

položme

$$I = \{ i \in \{1, \dots, m\} ; \quad g_i(x^*) = 0 \}$$

$$J = \{ j \in \{1, \dots, n\} ; \quad a_j(x^*) = 0 \}$$

Potom bod $\underline{x^*}$ je globálním řešením úlohy

$$\text{(NP)} \qquad \begin{array}{lll} \min & f(x) & \\ \text{s.p.} & g_i(x) \leqq 0 & \text{pro } i = 1, \dots, m \\ & a_j(x) \leqq 0 & \text{pro } j = 1, \dots, n \end{array}$$

právě tehdy, když

soustava

$$\delta_v^+ f(x^*) < 0$$ ...... kde $\underline{v}$ je neznámá

$$\delta_v^+ g_i(x^*) < 0 \qquad \text{pro } i \in I$$

$$\delta_v^+ a_j(x^*) \leqq 0 \qquad \text{pro } j \in J$$ ..... nemá řešení.

…známka: Jak vypadá uvedený důsledek v případě úloh klasického konvexního programování?

…ž víme: Jestliže funkce f je konvexní, potom je pseudokonvexní.

…ěta: (…g.:) Množina M nechť je otevřená a konvexní.
Jestliže funkce f je striktně konvexní, potom je striktně pseudo-
pseudokonvexní.

Důkaz: Již víme, že pokud funkce f je striktně konvexní, potom je konvexní, a tudíž má jednostrannou směrovou derivaci $\delta_v^+ f(x)$ pro každý směr $\underline{v}$ a je lokálně konvexní v každém bodě $x \in M$:

$$\forall x \in M \ \forall y \in M: \ f(y) \geqq f(x) + \delta^+_{y-x} f(x)$$

- zvol dva různé body $x, y \in M$, $x \neq y$
- předpokládej $f(y) \leqq f(x)$
- máme dokázat, že $\delta^+_{y-x} f(x) < 0$.

Jestliže $f(y) < f(x)$, potom $0 > f(y) - f(x) \geqq \delta^+_{y-x} f(x)$

Předpokládejme proto, že $f(y) = f(x)$.

- protože f je striktně konvexní, platí $f\left(\frac{x+y}{2}\right) < \frac{1}{2} f(x) + \frac{1}{2} f(y) = f(x)$
- potom, protože f je konvexní (striktně)

$$f\left(x + \frac{\lambda}{2}(y-x)\right) = f\left((1-\lambda)\cdot x + \lambda \cdot \frac{x+y}{2}\right) < (1-\lambda)\cdot f(x) + \lambda \cdot f\left(\frac{x+y}{2}\right)$$

$$= f(x) + \lambda \cdot \left(f(x) - f\left(\frac{x+y}{2}\right)\right) \quad \text{pro } \lambda \in (0,1)$$

- tudíž

$$\delta^+_{y-x} f(x) = \lim_{\lambda \to 0^+} \frac{f(x + \lambda(y-x)) - f(x)}{\lambda} =$$

$$= \lim_{\lambda \to 0^+} \frac{f\left(x + \frac{\lambda}{2}(y-x)\right) - f(x)}{\frac{\lambda}{2}} < \lim_{\lambda \to 0^+} \frac{2 \cdot \lambda \cdot \left(f(x) - f\left(\frac{x+y}{2}\right)\right)}{\lambda} < 0,$$

tedy f je striktně pseudokonvexní,

c.b.d.

…námka: O funkcích $a_1, \dots, a_n : M \to \mathbb{R}$ v naposledy uvedeném důsledku (viz list FJ+KKT 5) se předpokládá, že mají jednostranné směrové derivace $\delta_v^+ a_j(x)$, že jsou lokálně kvazikonvexní & pseudokonkávní a že jsou polospojité shora.

…íme: (cvičení)

Jestliže M je otevřená a $f : M \to \mathbb{R}$ je pseudokonvexní, potom je lokálně kvazikonvexní (v každém bodě $x \in M$).

…vaha: Podle definice je:

$$a_j \dots \text{pseudokonkávní} \iff -a_j \dots \text{pseudokonvexní},$$

tudíž funkce $-a_1, \dots, -a_n$ jsou též lokálně kvazikonvexní:

$$\forall x \in M \ \forall y \in M: \quad -a_j(y) \leqq -a_j(x) \implies \delta^+_{y-x} -a_j(x) \leqq 0$$

$$a_j(y) \geqq a_j(x) \implies \delta^+_{y-x} a_j(x) \geqq 0,$$

tudíž funkce $a_1, \dots, a_n$ jsou i lokálně kvazikonkávní.

Současně víme podle předpokladu:

$$\forall x \in M \ \forall y \in M: \quad a_j(y) \leqq a_j(x) \iff \delta^+_{y-x} a_j(x) \leqq 0$$

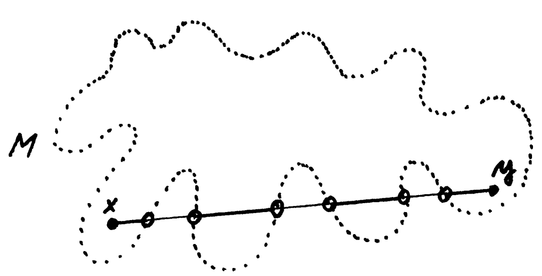
$$a_j(y) > a_j(x) \iff \delta^+_{y-x} a_j(x) > 0$$

Dohromady odvodíme:

$$\forall x \in M \ \forall y \in M: \quad a_j(x) = a_j(y) \implies \delta^+_{y-x} a_j(x) = 0$$

Nyní:

- zvol dva body $x, y \in M$ takové, že $a_j(x) = a_j(y)$
- tudíž $\delta^+_{y-x} a_j(x) = 0$
- tudíž $a_j(x + \lambda \cdot (y-x)) \leqq a_j(x) = a_j(y)$ pro všechna $\lambda \in \langle 0, 1 \rangle$ taková, že $x + \lambda \cdot (y-x) \in M$



POZOR: O M předpokládáme, že je otevřená, ale M nemusí být konvexní!

– pro spor předpokládejme, že existuje bod $\hat{x} \in M$ takový, že

& $\hat{x} = x + \hat{\lambda}\cdot(y-x)$ .... pro vhodné $\hat{\lambda} \in \langle 0,1\rangle$

$a_j(\hat{x}) < a_j(x) = a_j(y)$ —— tudíž: $0 < \hat{\lambda} < 1$

– tudíž

$$\delta^+_{x-\hat{x}}\, a_j(\hat{x}) > 0 \qquad i \qquad \delta^+_{y-\hat{x}}\, a_j(\hat{x}) > 0$$

– tudíž:

– uvažujeme-li funkci

$$\varphi : D_\varphi \longrightarrow \mathbb{R}$$

$$D_\varphi = \{ \lambda \in (0,1) ;\ x + \lambda\cdot(y-x) \in M \}$$

$$\varphi(\lambda) = a_j(x + \lambda\cdot(y-x)) \quad \text{.... pro } \lambda \in D_\varphi$$

definiční obor je otevřená množina (sjednocení otevřených intervalů), protože množina M je otevřená.

– potom $\delta^+_{+1}\varphi(\hat{\lambda}) > 0$ i $\delta^+_{-1}\varphi(\hat{\lambda}) > 0$

– tudíž bod $\lambda = \hat{\lambda}$ je bodem ostrého lokálního minima

– navíc funkce φ je polospojitá shora, protože funkce a_j je polospojitá shora

– tudíž (protože v $\hat{\lambda}$ je lok. min.) funkce φ je v bodě $\hat{\lambda}$ spojitá.

– máme bod $\hat{\lambda} \in D_\varphi$ takový, že $\varphi(\hat{\lambda}) < a_j(x) = a_j(y)$ a funkce φ je v bodě $\hat{\lambda}$ spojitá (polospojitá shora)

– tudíž k $\varepsilon = (a_j(x) - \varphi(\hat{\lambda}))$

$$\exists\, \delta > 0\ \forall \lambda \in D_\varphi : |\lambda - \hat{\lambda}| < \delta \implies \varphi(\lambda) < \varphi(\hat{\lambda}) + \varepsilon = a_j(x)$$

bůno je $\delta > 0$ tak malé, že $(\hat{\lambda}-\delta, \hat{\lambda}+\delta) \subseteq D_\varphi$

– potom – na základě výše odvozeného (pro bod $\hat{\lambda}$) – dostáváme, že:

– funkce φ je spojitá v každém bodě intervalu $(\hat{\lambda}-\delta, \hat{\lambda}+\delta)$

– každý bod intervalu $(\hat{\lambda}-\delta, \hat{\lambda}+\delta)$ je bodem ostrého lokálního minima funkce φ

– což není možné (proč?)

– tedy $\varphi(\lambda) = a_j(x) = a_j(y)$ pro všechna $\lambda \in D_\varphi$

neboli $a_j(x + \lambda\cdot(y-x)) = a_j(x) = a_j(y)$ pro všechna $\lambda \in \langle 0,1\rangle$ taková, že $x + \lambda\cdot(y-x) \in M$

Poznámka: Tedy o funkcích $\underline{a}_j$ z posledního důsledku můžeme říci, že jsou lokálně kvazikonvexní v každém bodě & kvazikonkávní & pseudokonkávní.

Z předchozích výsledků ihned plyne:

Věta: Množina M nechť je otevřená a konvexní.
(ad a_j) Jestliže funkce $\underline{a}_j$ je konvexní & pseudokonkávní, potom je lokálně kvazikonvexní & pseudokonkávní.

Poznámka: Předpoklady věty splňují – tj. konvexní & pseudokonkávní jsou – například funkce lineární ...... $a_j x + b_j$, kde $\underline{a}_j$ je řádkový vektor a $b_j \in \mathbb{R}$ je číslo.

Třída takových funkcí je ale širší (větší)!

Tedy v případě konvexního programování má důsledek – složení obou lemmat – následující podobu:

ůsledek: Nechť M je otevřená a konvexní množina

$$f : M \to \mathbb{R} \quad \text{je konvexní funkce}$$

$$g_1, \dots, g_m : M \to \mathbb{R} \quad \text{jsou shora polospojité a striktně konvexní funkce}$$

$$a_1, \dots, a_n : M \to \mathbb{R} \quad \text{jsou shora polospojité a pseudokonkávní konvexní funkce.}$$

Zvolme bod $x^* \in M$ takový, že $g_i(x^*) \leqq 0$ pro $i = 1, \dots, m$ a
$a_j(x^*) \leqq 0$ pro $j = 1, \dots, n$ a

položme

$$I = \{ i \in \{1, \dots, m\} \; ; \; g_i(x^*) = 0 \}$$

$$J = \{ j \in \{1, \dots, n\} \; ; \; a_j(x^*) = 0 \}$$

Potom bod $\underline{x}^*$ je globálním řešením úlohy

$$\text{(NP)} \qquad \min f(x)$$
$$\text{s.t.} \quad g_i(x) \leqq 0 \quad \text{pro } i = 1, \dots, m$$
$$a_j(x) \leqq 0 \quad \text{pro } j = 1, \dots, n$$

právě tehdy, když

soustava

$$\delta_v^+ f(x^*) < 0 \qquad \text{— kde } \underline{v} \text{ je neznámá}$$

$$\delta_v^+ g_i(x^*) < 0 \quad \text{pro } i \in I$$

$$\underline{\delta_v^+ a_j(x^*) \leqq 0} \quad \text{pro } j \in J$$

... nemá řešení.

ičení: Porovnejte tento výsledek s podmínkou optimality pro úlohu lineárního programování!

Pointa: Připomeňme motivaci celého našeho snažení: vyřešit úlohu

$$(NP) \quad \min f(x) \quad \text{z.p.} \quad g_i(x) \leqq 0 \quad \text{pro } i = 1, \dots, m$$

Úkolem je najít bod globálního minima funkce f na množině přípustných řešení. Obvyklý postup je ten, že stanovíme podmínku nutnou pro to, aby bod $\underline{x}^*$ byl bodem globálního minima. (Jinými slovy: hledáme body $\underline{x}^*$ tzv. "podezřelé" z toho, že je v nich globální minimum.)

Proto se zaměřujeme na přípustné body $\underline{x}^*$ takové, že soustava

$$(*) \quad \begin{array}{ll} -\underline{D}^+_v - f(x^*) < 0 & \\ -\underline{D}^+_v - g_i(x^*) < 0 & \text{pro } i \in I \\ -\underline{D}^+_v - g_i(x^*) \leqq 0 & \text{pro } i \in J \end{array} \quad \dots \text{nemá řešení}$$

(via první klíčové lemma).

Takovýto přímý postup

- zvol přípustný bod $\underline{x}^*$
- ověř (asi výpočtem), zda soustava $(*)$ (ne)má řešení

je ovšem krajně nepraktický.

Bod $\underline{x}^*$ bychom raději nalezli řešením nějaké soustavy rovnic (popř. nerovnic). Za tímto účelem použijeme některou z tzv. <u>vět o alternativě</u>. Zde připomeneme dvě věty o alternativě pro homogenní soustavy (tj. soustavy s nulovou pravou stranou):

- Farkasovo lemma,
- Motzkinovu větu.

$\longrightarrow$

Farkasovo lemma: Nechť $A \in \mathbb{R}^{m\times n}$
$c^T \in \mathbb{R}^{1\times n}$

Potom
$$\forall x \in \mathbb{R}^n:\ Ax \leqq 0 \implies c^T x \leqq 0$$
$$\iff$$
$$\exists \mu^T \in \mathbb{R}^{1\times m},\ \mu^T \geqq 0^T:\ c^T = \mu^T A$$

Obecněji:
Nechť X ... vektorový prostor
$A: X \to \mathbb{R}^m$ ... lin. zobr.
$c^T: X \to \mathbb{R}^1$ ... lin. zobr.
Potom
$\forall x \in X:\ Ax \leqq 0 \implies c^T x \leqq 0$
$\iff$
$\exists \mu^T \in \mathbb{R}^{1\times m},\ \mu^T \geqq 0^T:\ c^T = \mu^T A$

← Např.: $X = \mathbb{R}^n$

Motzkinova věta: Nechť $A \in \mathbb{R}^{m\times N}$
$B \in \mathbb{R}^{n\times N}$

Potom
$$\nexists x \in \mathbb{R}^n:\ \left.\begin{array}{l} Ax < 0 \\ \& \ \underline{Bx \leqq 0} \end{array}\right\}$$ tj. soustava nemá řešení

$$\iff$$
$$\exists \lambda^T \in \mathbb{R}^{m\times 1},\ \lambda^T \geqq 0^T,\ \lambda^T \neq 0^T,$$
$$\exists \mu^T \in \mathbb{R}^{n\times 1},\ \mu^T \geqq 0^T:\ \lambda^T A + \mu^T B = 0^T$$

Obecněji:
Nechť X ... vektorový prostor
$A: X \to \mathbb{R}^m$ ... lin. zobr.
$B: X \to \mathbb{R}^n$ ... lin. zobr.
Potom
$\nexists x \in X:\ Ax < 0\ \&\ Bx \leqq 0$
$\iff$
$\exists \lambda^T \in \mathbb{R}^{1\times m},\ \lambda^T \geqq 0^T,\ \lambda^T \neq 0^T,$
$\exists \mu^T \in \mathbb{R}^{1\times n},\ \mu^T \geqq 0^T:$
$\lambda^T A + \mu^T B = 0^T$

← Např. $X = \mathbb{R}^N$

Nyní se vraťme k soustavě (*) a položme:
$$p(v) = -\underline{D}^+_v - f(x^*)$$
$$q_i(v) = -\underline{D}^+_v - g_i(x^*) \quad \text{pro } i \in I \cup J$$
pro každý směr $\underline{v}$.

Větu o alternativě – Motzkinovu větu – tedy aplikujeme na soustavu
$$p(v) < 0$$
$$q_i(v) < 0 \quad \text{pro } i \in I$$
$$\underline{q_i(v) \leqq 0} \quad \text{pro } i \in J$$

Motzkinovu větu máme dokázánu jen pro případ lineárních funkcionálů (zobrazení). [Poznámka: Studují se rozličná zobecnění Farkasova lemmatu i Motzkinovy věty – např. pro sublineární funkcionály nebo obecnější třídy funkcionálů, i pro nekonečné soustavy nerovnic – zde budeme uvažovat jen lineární případ a konečnou soustavu.]

Proto musíme požadovat, aby funkcionály $p(v) = -\underline{D}^+_v(-f)(x^*)$ a $q_i(v) = -\underline{D}^+_v(-g_i)(x^*)$ byly lineární. Jakmile budeme mít k dispozici obecnější Farkasovo lemma nebo Motzkinovu větu, můžeme formulovat podmínky optimality i pro obecnější úlohy nelineárního programování.

Spojením prvního klíčového lemmatu a Motzkinovy věty dostáváme následující výsledek: podmínku optimality Fritze Johna.
[V závorkách upozorňujeme na speciální případy přímo ve znění věty.]

ěta: (podmínka optimality Fritze Johna)

Nechť M je otevřená množina

$$\left.\begin{array}{r} f: M \to \mathbb{R} \\ g_1, \dots, g_m: M \to \mathbb{R} \end{array}\right\} \text{ libovolné funkce}$$

Nechť bod $x^* \in M$ je lokálním řešením úlohy

$$\text{(NP)} \quad \min f(x) \quad \text{z.p.} \quad g_i(x) \leqq 0 \quad \text{pro } i = 1, \dots, m$$

Položme

$$I_0 = \{ i \in \{1, \dots, m\} \ ; \ g_i(x^*) = 0 \}$$

a

$$J = \{ i \in I_0 \ ; \ g_i \text{ -- je lokálně pseudokonkávní v } \underline{x}^* \}$$

$$\forall y \in M: \ \underline{D}^+_{y-x^*}(-g_i)(x^*) \geqq 0 \implies g_i(y) \leqq g_i(x^*) = 0$$

$$I = I_0 \setminus J$$

Nechť
- funkcionál $p(v) = \underline{D}^+_v (-f)(x^*)$ — je lineární
- funkcionály $q_i(v) = \underline{D}^+_v (-g_i)(x^*)$ — jsou lineární pro $i \in I_0$

  [např. f má v bodě $\underline{x}^*$ Gâteauxovu derivaci
  a g_i mají v bodě $\underline{x}^*$ Gâteauxovu derivaci pro $i \in I_0$]
- funkce g_i jsou polospojité shora v bodě $\underline{x}^*$ pro $i \in \{1, \dots, m\} \setminus I_0$.

Potom

$$\exists \nu \in \mathbb{R}, \ \nu \geqq 0, \ \exists \mu_I \in \mathbb{R}^I, \ \mu_I \geqq 0, \ \nu \neq 0 \text{ nebo } \mu_I \neq \sigma,$$

$$\exists \lambda_J \in \mathbb{R}^J, \ \lambda_J \geqq 0: \quad \nu \cdot \underline{D}^+(-f)_{(x^*)} + \sum_{i \in I} \mu_i \cdot \underline{D}^+(-g_i)_{(x^*)} + \sum_{i \in J} \lambda_i \cdot \underline{D}^+(-g_i)_{(x^*)} = \sigma^T$$

mají-li funkce Gâteauxovy derivace:

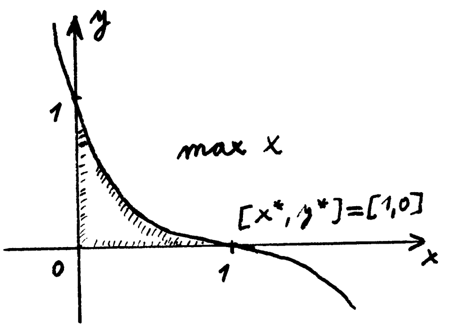
$$\nu \cdot Df(x^*) + \sum_{i \in I} \mu_i \cdot Dg_i(x^*) + \sum_{i \in J} \lambda_i \cdot Dg_i(x^*) = \sigma^T$$

zapsáno pomocí gradientu:

$$\nu \cdot \nabla f(x^*) + \sum_{i \in I} \mu_i \cdot \nabla g_i(x^*) + \sum_{i \in J} \lambda_i \cdot \nabla g_i(x^*) = \sigma^T$$

otivace: V podmínce optimality Fritze Johna může nastat případ, kdy $\nu = 0$. V tom případě vztah $\sum_{i \in I} \mu_i \cdot \underline{D}^+(-g_i)(x^*) + \sum_{i \in J} \lambda_i \cdot \underline{D}^+(-g_i)(x^*) = o^T$ vyjadřuje pouze lokální singularitu aktivních podmínek v bodě $\underline{x}^*$. Vůbec nezávisí na cílové funkci f, tj., cílová funkce f může být jiná, bod $\underline{x}^*$ už nemusí být lokálním optimem a lokální singularita je stále v platnosti.

lasický příklad je následující:



$$\min f(x,y) = -x$$

$$\text{z.p.} \quad g_1(x,y) = y - (1-x)^3 \leqq 0$$

$$g_2(x,y) = -y \leqq 0$$

$$g_3(x,y) = -x \leqq 0$$

Optimum je v bodě $[x^*, y^*] = [1, 0]$

Aktivní v bodě $[x^*, y^*]$ jsou podmínky $g_1 \leqq 0$, $g_2 \leqq 0$.

$$\nabla f(x^*, y^*) = (-1 \quad 0)$$

$$\nabla g_1(x^*, y^*) = (0 \quad 1)$$

$$\nabla g_2(x^*, y^*) = (0 \quad -1)$$

Řešíme rovnici

$$\nu \cdot \nabla f(x^*, y^*) + \mu_1 \cdot \nabla g_1(x^*, y^*) + \mu_2 \cdot \nabla g_2(x^*, y^*) = o^T$$

neboli

$$-\nu + 0 \cdot \mu_1 + 0 \cdot \mu_2 = 0 \qquad \nu \geqq 0$$

$$0 \cdot \nu + \mu_1 - \mu_2 = 0 \qquad \mu_1, \mu_2 \geqq 0$$

$$\nu + \mu_1 + \mu_2 > 0$$

Jediným možným řešením je $\nu = 0$ a libovolné $\mu_1 = \mu_2 > 0$.

Tedy bod $[x^*, y^*] = [0, 0]$ je "singulárním bodem" množiny přípustných řešení, ať už cílová funkce f je jakákoliv.

tivace: Případu, kdy $\nu = 0$, se snažíme zabránit, tj., případ $\nu = 0$ se snažíme vyloučit.

Toho dosahujeme přidáním dodatečného předpokladu, že podmínky aktivní v bodě $\underline{x}^*$ splňují v tomto bodě $\underline{x}^*$ podmínku kvalifikace omezujících podmínek (nebo kvalifikace omezení, angl. constraint qualification).

Některé podmínky kvalifikace omezení uvedeme později.

finice: (KKT podmínka resp. KKT-stacionární bod)

Nechť M je otevřená množina a nechť

$$\left.\begin{array}{r} f : M \longrightarrow \mathbb{R} \\ g_1, \dots, g_m : M \longrightarrow \mathbb{R} \end{array}\right\} \text{ jsou libovolné funkce}$$

Zvolme libovolný přípustný bod $\underline{x}^*$ úlohy

$$\text{(NP)} \qquad \begin{array}{l} \min f(x) \\ \text{z.p.} \quad g_i(x) \leqq 0 \quad \text{pro } i = 1, \dots, m \end{array}$$

neboli bod $\underline{x}^* \in M$ takový, že $g_i(x^*) \leqq 0$ pro $i = 1, \dots, m$.

Položme

$$I_0 = \{ i \in \{1, \dots, m\} \; ; \; g_i(x^*) = 0 \}$$

Funkcionál $p(v) = \underline{D}_v^+ (-f)(x^*)$ budiž lineární a funkcionály $q_i(v) = \underline{D}_v (-g_i)(x^*)$ budiž lineární pro $i \in I_0$.

Říkáme, že bod $\underline{x}^*$ je KKT-stacionární (neboli splňuje KKT podmínku)

$\Longleftrightarrow$

$$\exists\, u^T \in \mathbb{R}^{1 \times m}, \; u^T \geqq o^T : \quad \underline{D}^+(-f)(x^*) + \sum_{i=1}^m u_i \cdot \underline{D}^+(-g_i)(x^*) = o^T \quad \&$$

$$\underline{D}^+(-f)(x^*) + u^T \cdot \underline{D}^+(-g)(x^*) = o^T \qquad \& \quad u^T \cdot g(x^*) = 0$$

zapsáno pomocí Gâteauxových derivací (mají-li je funkce)

$$Df(x^*) + \sum_{i=1}^m u_i \cdot Dg_i(x^*) = o^T$$

$$Df(x^*) + u^T \cdot Dg(x^*) = o^T \qquad \& \quad u^T \cdot g(x^*) = 0$$

pomocí gradientu:

$$\nabla f(x^*) + \sum_{i=1}^m u_i \cdot \nabla g_i(x^*) = o^T$$

$$\nabla f(x^*) + u^T \cdot \nabla g(x^*) = o^T \qquad \& \quad u^T \cdot g(x^*) = 0$$

Poznámka: KKT = Karush – Kuhn – Tucker

Poznámka: Podmínka $\mu^T \cdot g(x^*) = 0$ je podmínkou komplementarity:

$$\mu^T \cdot g(x^*) = \sum_{i=1}^{m} \underbrace{\underbrace{\mu_i}_{\geqq 0} \cdot \underbrace{g_i(x^*)}_{\leqq 0}}_{\leqq 0} = 0$$

Tudíž:

$$g_i(x^*) < 0 \implies \mu_i = 0$$

popř.

$$\mu_i > 0 \implies g_i(x^*) = 0$$

Pozorování ⓪: Bod $\underline{x}^*$ je KKT stacionární $\iff$ v bodě $\underline{x}^*$ je splněna podmínka Fritze Johna s $\nu > 0$.

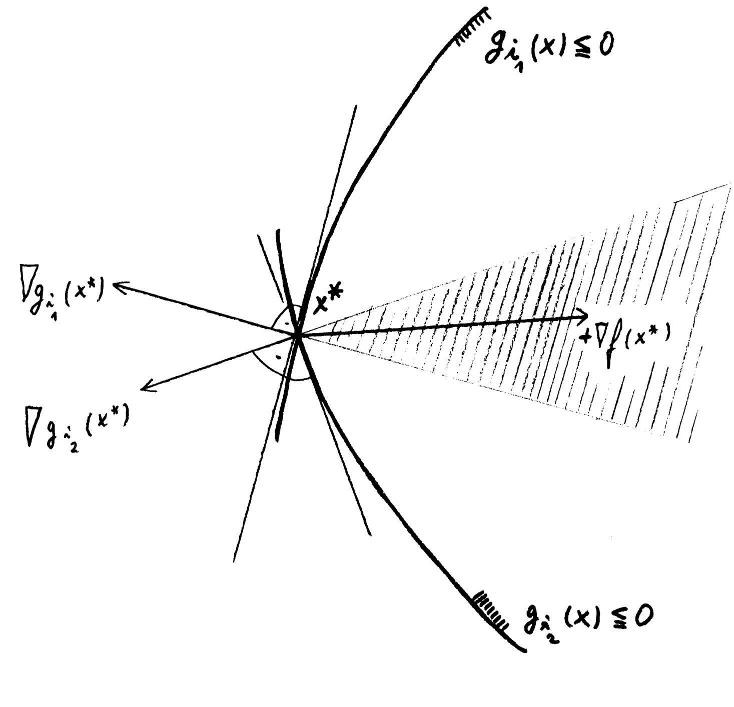
Důkaz: – zřejmý

Geometrická interpretace KKT podmínky:

$$\nabla f(x^*) + \mu^T \cdot \nabla g_i(x^*) = o^T \qquad \& \qquad \mu^T \cdot g(x^*) = 0$$

neboli

$$\nabla f(x^*) = -\sum_{i \in I_0} \mu_i \cdot \nabla g_i(x^*)$$



$$-\nabla f(x^*) = \sum_{i \in I_0} \mu_i \cdot \nabla g_i(x^*)$$

...... záporný gradient $-\nabla f(x^*)$ cílové funkce v bodě $\underline{x}^*$ leží v kuželovém obalu gradientů $\nabla g_i(x^*)$ podmínek aktivních v bodě $\underline{x}^*$.

Resp.: gradient $\nabla f(x^*)$ cílové funkce v bodě $\underline{x}^*$ leží v kuželovém obalu záporných gradientů $-\nabla g_i(x^*)$ podmínek aktivních v bodě $\underline{x}^*$.

Definice: (Mangasarianova-Fromovitzova podmínka kvalifikace omezení, MFCQ = Mangasarian-Fromovitz constraint qualification)

Nechť M je otevřená množina

a $g_1, \dots, g_m : M \to \mathbb{R}$ nechť jsou libovolné funkce.

Bod $x^* \in M$ budiž přípustný, tj. $g_i(x^*) \leqq 0$ pro $i = 1, \dots, m$.

Nechť
$$I_0 = \{ i \in \{1, \dots, m\} \; ; \; g_i(x^*) = 0 \}$$

a položme
$$J = \{ i \in I_0 \; ; \; g_i \text{ ... je lokálně pseudokonkávní v } \underline{x}^* \}$$
$$\forall y \in M: \; g_i(y) > g_i(x^*) = 0 \implies \underline{D}^+_{y-x^*}(-g_i)(x^*) < 0$$
$$I = I_0 \setminus J$$

Říkáme, že funkce $g_1, \dots, g_m$ splňují v bodě $\underline{x}^*$ Mangasarianovu-Fromovitzovu kvalifikaci omezujících podmínek (MFCQ) $\iff$

$\iff$ soustava
$$\underline{D}^+_v (-g_i)(x^*) > 0 \quad \text{pro } i \in I$$
$$(**) \qquad \underline{D}^+_v (-g_i)(x^*) \geqq 0 \quad \text{pro } i \in J$$
..... má řešení

Pozn.: jestliže funkce g_i pro $i \in I \cup J$ mají Gâteauxovy derivace, potom soustava má tvar
$$D_v g_i(x^*) < 0 \quad \text{pro } i \in I$$
$$D_v g_i(x^*) \leqq 0 \quad \text{pro } i \in J$$

⊙ : Jestliže $I = \emptyset$, potom soustava $(**)$ má řešení (řešení je např. $v = 0$), tudíž (MFCQ) je jistě splněna.

Nyní složením prvního klíčového lemmatu a Motzkinovy věty (tedy podmínky optimality Fritze Johna) a MFCQ dostáváme KKT podmínku optimality:

<u>Věta:</u> (Karushova - Kuhnova - Tuckerova podmínka optimality)

Nechť M je otevřená množina a nechť

$f : M \to \mathbb{R}$ a $g_1, \dots, g_m : M \to \mathbb{R}$ jsou libovolné funkce.

Bod $x^* \in M$ budiž lokálním řešením úlohy

$$(NP) \quad \min f(x) \quad \text{s.p.} \quad g_i(x) \leqq 0 \quad \text{pro } i = 1, \dots, m$$

Nechť $I_0 = \{ i \in \{1, \dots, m\} \; ; \; g_i(x^*) = 0 \}$ ...... jsou indexy aktivních podmínek

a $J = \{ i \in I_0 \; ; \; g_i$ — je lokálně pseudokonkávní v $\underline{x}^* \}$

$\forall y \in M : \; \underline{D}^+_{y - x^*}(-g_i)(x^*) \geqq 0 \implies g_i(y) \leqq g_i(x^*) = 0$

$I = I_0 \setminus J$

ať

- funkcionál $p(v) = \underline{D}^+_v(-f)(x^*)$ je lineární [např. f má v bodě $\underline{x}^*$ Gâteauxovu derivaci]
- funkcionály $q_i(v) = \underline{D}^+_v(-g_i)(x^*)$ jsou lineární [např. g_i mají — // —] pro $i \in I_0 = I \cup J$
- funkce g_i jsou v bodě $\underline{x}^*$ ~~polos~~ shora polospojité pro $i \in \{1, \dots, m\} \setminus I_0$

*) – v bodě $\underline{x}^*$ platí (MFCQ)

*) Potom bod $\underline{x}^*$ je KKT - stacionární

neboli

$$\exists u^T \in \mathbb{R}^{1 \times m}, \; u^T \geqq o^T : \; \underline{D}^+(-f)(x^*) + u^T \cdot D(-g)(x^*) = o^T \;\; \& \;\; u^T \cdot g(x^*) = 0$$

[mají-li funkce Gâteauxovy derivace: $\quad Df(x^*) + u^T \cdot Dg(x^*) = o^T \quad \& \quad u^T \cdot g(x^*) = 0$]

<u>Poznámka:</u> jestliže $I = \emptyset$, neboli $I_0 = J$, potom MFCQ je splněna.

*) jestliže v bodě $\underline{x}^*$ platí (MFCQ), potom bod $\underline{x}^*$ je KKT - stacionární

Důkaz: Podle prvního klíčového lemmatu platí, že soustava

$$(*) \qquad \begin{array}{lll} \underline{D}^+_v(-f)(x^*) > 0 & & \\ \underline{D}^+_v(-g_i)(x^*) > 0 & \text{pro } i \in I & \\ \underline{D}^+_v(-g_i)(x^*) \geqq 0 & \text{pro } i \in J & \text{.... nemá řešení} \end{array}$$

Podle Motzkinovy věty

$\exists \nu \geqq 0,\ \exists \mu^T \in \mathbb{R}^{1\times I},\ \mu^T \geqq o^T,\ \nu \neq 0$ nebo $\mu^T \neq o^T$,

$\exists \lambda^T \in \mathbb{R}^{1\times J},\ \lambda^T \geqq o^T:\quad \nu \cdot \underline{D}^+(-f)(x^*) + \sum_{i\in I} \mu_i \cdot \underline{D}^+(-g_i)(x^*) + \sum_{i\in J} \lambda_i \cdot \underline{D}^+(-g_i)(x^*) = o^T$

Kdyby $\nu = 0$, potom by ani soustava

$$(**) \qquad \begin{array}{lll} \underline{D}^+_v(-g_i)(x^*) > 0 & \text{pro } i \in I & \\ \underline{D}^+_v(-g_i)(x^*) \geqq 0 & \text{pro } i \in J & \text{..... neměla řešení} \end{array}$$

V bodě $\underline{x}^*$ ovšem předpokládáme platnost MFCQ neboli, že soustava (**) má řešení. Proto musí být $\nu > 0$.

Položíme
$$\begin{array}{ll} u_i = \dfrac{\mu_i}{\nu} & \text{pro } i \in I \\ u_i = \dfrac{\lambda_i}{\nu} & \text{pro } i \in J \\ u_i = 0 & \text{pro } i \in \{1,\dots,m\} \setminus I_0 \end{array}$$

c.b.d.

Dosud jsme se zabývali podmínkou nutnou pro globální minimum úlohy (NP).

Složením druhého (částečně obráceného) klíčového lemmatu a Motzkinovy věty dostáváme následující podmínku postačující:

Věta: (podmínka postačující pro globální minimum)

Nechť M je otevřená množina

$f : M \to \mathbb{R}$ a $g_1, \dots, g_m : M \to \mathbb{R}$ jsou funkce

$x^* \in M$ je bod takový, že $g_i(x^*) \leqq 0$ pro $i = 1, \dots, m$.

Označme

$$I_0 = \{ i \in \{1, \dots, m\} \, ; \; g_i(x^*) = 0 \}$$

Nechť f a g_i mají v bodě $\underline{x}^*$ Gâteauxovy derivace pro $i \in I_0$.

Nechť f .... je .... pseudokonvexní v bodě $\underline{x}^*$

$\forall y \in M: \; f(y) < f(x^*) \implies D_{y-x^*} f(x^*) < 0$

$I = \{ i \in I_0 \, ; \; g_i$ je striktně pseudokonvexní v bodě $\underline{x}^* \}$

$\forall y \in M, \, y \neq x^*: \; g_i(y) \leqq g_i(x^*) \implies D_{y-x^*} g_i(x^*) < 0$

$J = \{ i \in I_0 \, ; \; g_i$ je lokálně kvazikonvexní v bodě $x^* \}$

$\forall y \in M: \; g_i(y) \leqq g_i(x^*) \implies D_{y-x^*} g_i(x^*) \leqq 0$

Předpokládejme, že $I_0 = I \cup J$.

$J \subseteq I_0 \setminus I$

Jestliže v bodě $\underline{x}^*$ platí podmínka optimality Fritze Johna

[speciálně: jestliže bod $\underline{x}^*$ je KKT-stacionární]

$$\exists \nu \geqq 0 \;\; \exists \lambda \in \mathbb{R}^{I_0}, \lambda \geqq 0 : \quad \nu + \sum_{i \in I} \lambda_i > 0 \quad \&$$

$$\& \quad \nu \cdot Df(x^*) + \sum_{i \in I_0} \lambda_i \cdot Dg_i(x^*) = o^T$$

potom bod $\underline{x}^*$ je bodem globálního minima v úloze

$$\text{(NP)} \qquad \min f(x) \quad \text{z.p.} \quad g_i(x) \leqq 0 \quad \text{pro } i = 1, \dots, m$$

Na závěr je naším cílem obě podmínky (nutnou i postačující) sloučit do jediného výsledku, tj. vyslovit předpoklady, za jejichž splnění je bod $\underline{x}^*$ globálním řešením úlohy (NP) $\Longleftrightarrow$ je KKT-stacionární.

<u>efinice</u>: (zobecněná Slaterova podmínka)

Nechť M je otevřená množina

a $g_1, \dots, g_m : M \to \mathbb{R}$ jsou libovolné funkce

Bod $x^* \in M$ budiž přípustný, tj. $g_i(x^*) \leqq 0$ pro $i = 1, \dots, m$.

Položme $I_0 = \{ i \in \{1, \dots, m\} \; ; \; g_i(x^*) = 0 \}$

nechť g_i mají Gâteauxovy derivace $Dg_i(x^*)$ v bodě $\underline{x}^*$ pro $i \in I_0$ a

$I = \{ i \in I_0 \; ; \; g_i$ je lokálně pseudokonvexní v bodě $\underline{x}^* \}$

$\forall y \in M: \; g_i(y) < g_i(x^*) = 0 \implies D_{y-x^*} g_i(x^*) < 0$

$J = \{ i \in I_0 \; ; \; g_i$ je lokálně kvazikonvexní v bodě $\underline{x}^* \}$

$\forall y \in M: \; g_i(y) \leqq g_i(x^*) = 0 \implies D_{y-x^*} g_i(x^*) \leqq 0$

Nechť $I_0 = I \cup J$ ..... $J \subseteq I_0 \setminus I$

Říkáme, že v bodě $\underline{x}^*$ platí zobecněná Slaterova podmínka $\Longleftrightarrow$

$$\exists x_0 \in M : \quad \begin{aligned} g_i(x_0) &< 0 \quad \text{pro } i \in I \setminus J \\ g_i(x_0) &\leqq 0 \quad \text{pro } i \in J \end{aligned}$$

Věta: Jestliže v bodě $\underline{x}^*$ je splněna zobecněná Slaterova podmínka

a navíc $\forall i \in J$: g_i je navíc pseudokonkávní v bodě $\underline{x}^*$

$$\forall y \in M: \quad g_i(y) > g_i(x^*) = 0 \quad \Longrightarrow \quad D_{y-x^*}\, g_i(x^*) > 0$$

potom v bodě $\underline{x}^*$ je splněna i (MFCQ).

Důkaz:

- máme bod $\underline{x}^*$ a máme bod $\underline{x}_0$
- pro $i \in I$ platí

$$g_i(x_0) < 0 = g_i(x^*), \quad \text{tudíž} \quad D_{x_0-x^*}\, g_i(x^*) < 0,$$

protože g_i je pseudokonvexní v $\underline{x}^*$

- pro $i \in J$ platí

$$g_i(x_0) \leqq 0 = g_i(x^*), \quad \text{tudíž} \quad D_{x_0-x^*}\, g_i(x^*) \leqq 0,$$

protože g_i je kvazikonvexní v $\underline{x}^*$

- tudíž soustava

$$D_v\, g_i(x^*) < 0 \quad \text{pro } i \in I$$
$$\underline{D_v\, g_i(x^*) \leqq 0} \quad \text{pro } i \in J$$

...... má řešení, např. $v = x_0 - x^*$

c.b.d.

efinice: (Slaterova podmínka)

Nechť M je množina a

$g_1, \dots, g_m : M \to \mathbb{R}$ jsou funkce.

Říkáme, že funkce $g_1, \dots, g_m$ splňují Slaterovu podmínku na $M \Leftrightarrow$

$\exists x_0 \in M : \quad g_i(x_0) < 0 \quad$ pro $i = 1, \dots, m$

námka: Uvedli jsme klasickou Slaterovu podmínku pro úlohu

(NP) $\min f(x)$

z.p. $g_i(x) \leqq 0$ pro $i = 1, \dots, m$

Bod $\underline{x_0}$ se nazývá <u>Slaterův bod</u>.

ta: Nechť M je otevřená množina,

$g_1, \dots, g_m : M \to \mathbb{R}$ jsou libovolné funkce,

$x^* \in M$ je přípustný bod, tj. $g_i(x^*) \leqq 0$ pro $i = 1, \dots, m$

Nechť g_i mají Gâteauxovy derivace $Dg_i(x^*)$ v bodě $\underline{x}^*$ pro $i \in I_0$, kde

$$I_0 = \{ i \in \{1, \dots, m\} \; ; \; g_i(x^*) = 0 \}$$

$$I = \{ i \in I_0 \; ; \; g_i \text{ je lokálně pseudokonvexní v bodě } \underline{x}^* \}$$

$$J = \{ i \in I_0 \; ; \; g_i \text{ je lokálně kvazikonvexní v bodě } x^* \}$$

a nechť $I_0 = I \cup J$

$J \subseteq I_0 \setminus I$

Jestliže je splněna Slaterova podmínka,

potom v bodě $\underline{x}^*$ platí i zobecněná Slaterova podmínka.

Důkaz:

- zřejmý

Složením dosud uvedených výsledků dostáváme následující

<u>ísledek:</u> Nechť M je otevřená množina

$$\left.\begin{array}{l} f: M \to \mathbb{R} \\ g_1, \dots, g_m : M \to \mathbb{R} \\ a_1, \dots, a_n : M \to \mathbb{R} \end{array}\right\}$$ funkce, které mají Gâteauxovy derivace $Df(x)$, $Dg_i(x)$, $Da_j(x)$ v každém bodě $x \in M$

Nechť $\left.\begin{array}{l} g_1, \dots, g_m : M \to \mathbb{R} \\ a_1, \dots, a_n : M \to \mathbb{R} \end{array}\right\}$ polospojité shora

Nechť $f: M \to \mathbb{R}$ .... pseudokonvexní

$\forall x \in M \ \forall y \in M:$ $\quad f(y) < f(x) \implies D_{y-x} f(x) < 0$

$g_1, \dots, g_m : M \to \mathbb{R}$ .... striktně pseudokonvexní

$\forall x \in M \ \forall y \in M, \ y \neq x:$ $\quad g_i(y) \leqq g_i(x) \implies D_{y-x} g_i(x) < 0$

$a_1, \dots, a_n : M \to \mathbb{R}$ .... lokálně kvazikonvexní & pseudokonkávní

$\forall x \in M \ \forall y \in M:$ $\quad a_j(y) \leqq a_j(x) \iff D_{y-x} a_j(x) \leqq 0$

Nechť $\exists x_0 \in M$: $\quad g_i(x^*) < 0$ pro $i = 1, \dots, m$

$\underline{a_j(x^*) \leqq 0}$ pro $j = 1, \dots, n$

Zvolme libovolný přípustný bod $x^* \in M$, tj. bod $x^* \in M$ takový, že $\quad g_i(x^*) \leqq 0$ pro $i = 1, \dots, m$
$a_j(x^*) \leqq 0$ pro $j = 1, \dots, n$

Potom: bod $\underline{x}^*$ je globálním řešením úlohy

$$\begin{array}{lll} \text{(NP)} & \min f(x) & \\ & \text{s.p.} \quad g_i(x) \leqq 0 & \text{pro } i = 1, \dots, m \\ & \phantom{\text{s.p.}} \quad a_j(x) \leqq 0 & \text{pro } j = 1, \dots, n \end{array}$$

právě tehdy, když

bod $\underline{x}^*$ je KKT stacionární

$\exists u^T \in \mathbb{R}^{1 \times m}, \ u^T \geqq o^T, \ \exists v^T \in \mathbb{R}^{1 \times n}, \ v^T \geqq o^T:$

$$Df(x^*) + u^T \cdot Dg(x^*) + v^T \cdot Da(x^*) = o^T \quad \& \quad \begin{array}{l} u^T \cdot g(x^*) = 0 \\ v^T \cdot a(x^*) = 0 \end{array} \ \&$$

Důkaz: Položme

$$I = \{ i \in \{1,\dots,m\} \; ; \; g_i(x^*) = 0 \}$$

$$J = \{ j \in \{1,\dots,n\} \; ; \; a_j(x^*) = 0 \}$$

$\Leftarrow$ Protože bod $\underline{x}^*$ je KKT - stacionární, dostáváme, že soustava

$$\begin{aligned} &D_v f(x^*) < 0 \\ &D_v g_i(x^*) \leqq 0 \quad \dots \text{ pro } i \in I \\ &D_v a_j(x^*) \leqq 0 \quad \dots \text{ pro } j \in J \end{aligned} \qquad \dots\dots \text{ nemá řešení}$$

tudíž ani soustava

$$\begin{aligned} &D_v f(x^*) < 0 \\ &D_v g_i(x^*) < 0 \quad \dots \text{ pro } i \in I \\ &D_v a_j(x^*) \leqq 0 \quad \dots \text{ pro } j \in J \end{aligned} \qquad \dots\dots \text{ nemá řešení}$$

Druhé (částečně obrácené) klíčové lemma dává výsledek, že bod $\underline{x}^*$ je globálním řešením úlohy (NP).

$\Rightarrow$ Protože bod $\underline{x}^*$ je optimálním řešením úlohy (NP), je splněna podmínka optimality Fritze Johna.

Protože předpokládáme existenci Slaterova bodu … $g_i(x_0) < 0$ pro $i = 1,\dots,m$; $a_j(x_0) \leqq 0$ pro $j = 1,\dots,n$

v bodě $\underline{x}^*$ platí zobecněná Slaterova podmínka,

$$g_i(x_0^*) < 0 \quad \rightsquigarrow \quad D_{x_0 - x^*} g_i(x^*) < 0 \quad \text{pro } i \in I \quad \dots \text{ pseudokonvexita}$$

$$a_j(x_0^*) \leqq 0 \quad \rightsquigarrow \quad D_{x_0 - x^*} a_j(x^*) \leqq 0 \quad \text{pro } j \in J \quad \dots \text{ kvazikonvexita}$$

tudíž v bodě $\underline{x}^*$ platí (MFCQ).

Odtud plyne, že bod $\underline{x}^*$ je KKT - stacionární.

c. b. d.

vičení: Lze poslední velký důsledek použít na úlohy

– lineárního programování ?

$$(LP) \quad \min \; -c^T x \qquad \text{kde } A \in \mathbb{R}^{m \times n}$$
$$\text{z.p.} \quad Ax - b \leqq 0 \qquad b \in \mathbb{R}^m$$
$$c^T \in \mathbb{R}^{1 \times n}$$
$$x \in \mathbb{R}^n \text{ .... neznámá}$$

– kvadratického programování ?

$$(QP) \quad \min \; x^T Q x + p^T x \qquad \text{kde } Q \in \mathbb{R}^{n \times n} \text{ ... pozitivně semidefinitní}$$
$$\text{z.p.} \quad Ax - b \leqq 0 \qquad p^T \in \mathbb{R}^{1 \times n}$$

– konvexního programování ?

$$(CP) \quad \min \; f(x) \qquad \text{kde } \left.\begin{array}{l} f: M \to \mathbb{R} \\ g_{1,\dots,m}: M \to \mathbb{R} \end{array}\right\} \text{konvexní}$$
$$\text{z.p.} \quad g_i(x) \leqq 0 \text{ pro } i = 1, \dots, m \qquad x \in M \text{ .... neznámá}$$

známka: Poslední velký důsledek naznačuje výpočetní postup, s jehož pomocí lze řešit úlohy nelineární optimalizace. Úkolem je vyřešit soustavu rovnic a nerovnic

$$\begin{array}{lll} g(x) \leqq 0 & \& & Df(x) + u^T \cdot Dg(x) + v^T \cdot Da(x) = o^T \\ a(x) \leqq 0 & & \\ & \& & u^T \cdot g(x) + v^T \cdot a(x) = 0 \\ & & u^T \geqq o^T, \; v^T \geqq o^T \end{array}$$

$$\left.\begin{array}{l} x \in M \\ u^T \in \mathbb{R}^{1 \times m} \\ v^T \in \mathbb{R}^{1 \times n} \end{array}\right\} \text{neznámé}$$

ičení: Jak postupujeme při řešení úloh LP (jak pracuje simplexová metoda)?

Jak postupujeme při řešení úloh QP (jak pracuje Wolfeho metoda)? (Jak pracuje Murtyho algoritmus popř. Lemkeho algoritmus ?)

Jak postupujeme při řešení úloh nelineární optimalizace ?

Dodatek: KKT podmínka optimality odvozená pomocí jiné kvalifikace omezení

Nechť M je otevřená množina

$$\left.\begin{array}{r} f : M \to \mathbb{R} \\ g_1, \dots, g_m : M \to \mathbb{R} \end{array}\right\} \text{ jsou funkce}$$

Uvažujme úlohu

$$\text{(NP)} \quad \begin{array}{ll} \min & f(x) \\ \text{z.p.} & g_i(x) \leqq 0 \quad \text{pro } i = 1, \dots, m \end{array}$$

finice: (kvalifikace omezení CQ ... constraint qualification)

Zvolme libovolný přípustný bod $\underline{x}^*$ úlohy (NP), tj. bod $x^* \in M$ takový, že $g_i(x^*) \leqq 0$ pro $i=1,\dots,m$

Nechť

$I_0 = \{ i \in \{1,\dots,m\} ;\ g_i(x^*) = 0 \}$ .... označuje množinu indexů podmínek aktivních v bodě $\underline{x}^*$.

Předpokládejme, že g_i mají v bodě $\underline{x}^*$ jednostranné směrové derivace $\delta_v^+ g_i(x^*)$ pro každý směr $\underline{v} \in \mathbb{R}^N$ pro každé $i \in I_0$.

Řekneme, že podmínky $g_1(x) \leqq 0, \dots, g_m(x) \leqq 0$ splňují v bodě $\underline{x}^*$ kvalifikaci omezení CQ .... constraint qualification .... právě tehdy, když

pro každý směr $\vec{v} \in \mathbb{R}^N$ platí:

- jestliže $\vec{v}$ vyhovuje soustavě $\delta_{\vec{v}}^+ g_i(x^*) \leqq 0$ pro $i \in I_0$
- potom existuje oblouk přípustných bodů tvaru $x^* + \lambda \cdot \vec{v} + \vec{\sigma}(\lambda)$

Formálně:

$$\forall \vec{v} \in \mathbb{R}^N : \left(\forall i \in I_0 : \delta_{\vec{v}}^+ g_i(x^*) \leqq 0 \right) \implies$$

$$\implies \Big(\exists \lambda_0 > 0 \quad \exists \vec{\sigma} : (0, \lambda_0) \to \mathbb{R}^N \text{ ... zobrazení ... } \lim_{\lambda \to 0^+} \frac{\| \vec{\sigma}(\lambda) \|}{\lambda} = 0 :$$

$$\forall \lambda \in (0, \lambda_0) \quad \forall i = 1,\dots,m : \quad g_i(x^* + \lambda \cdot \vec{v} + \vec{\sigma}(\lambda)) \leqq 0 \Big)$$

ěta: Nechť M je otevřená množina

$g_1, \dots, g_m : M \to \mathbb{R}$ jsou funkce

$x^* \in M$ je přípustný bod ..... $g_i(x^*) \leqq 0$ pro $i = 1, \dots, m$

$I_0 = \{ i \in \{1, \dots, m\} \; ; \; g_i(x^*) = 0 \}$ ..... indexy aktivních podmínek

Nechť

g_i ..... mají v bodě $\underline{x}^*$ jednostranné směrové derivace $\delta^+_v g_i(x^*)$ pro každý směr $\underline{v}$ &

& jsou lokálně pseudokonkávní v bodě $\underline{x}^*$ ..... pro $i \in I_0$

$$\forall y \in M: \; \delta^+_{y - x^*} g_i(x^*) \leqq 0 \implies g_i(y) \leqq g_i(x^*) = 0$$

g_i ..... jsou polospojité shora v bodě $\underline{x}^*$ ..... pro $i \in \{1, \dots, m\} \setminus I_0$

Potom: v bodě $\underline{x}^*$ je splněna podmínka (CQ)

námka: Za uvedených předpokladů je v bodě $\underline{x}^*$ splněna také (MFCQ).

Důkaz: Uvažujme směr $\underline{v}$, který je řešením soustavy $\delta^+_v g_i(x^*) \leqq 0$ pro $i \in I_0$.

a) $v = 0$ ..... pak uvažujeme "nulový" oblouk ..... zřejmé

b) $v \neq 0$: → pro $i \in \{1, \dots, m\} \setminus I_0$ je $g_i(x^*) < 0$ a funkce g_i jsou shora polospojité v x^*,

tudíž k $\varepsilon = \frac{1}{2} \cdot \min \{ -g_i(x^*) \; ; \; i \in \{1, \dots, m\} \setminus I_0 \}$

$$\exists \delta > 0 \;\; \forall x \in M: \; \|x - x^*\| < \delta \implies g_i(x) < g_i(x^*) + \varepsilon$$

- položme $t_0 = \dfrac{\delta}{\|v\|}$

pak $\forall t \in (0, t_0)$: $\quad g_i(x^* + t \cdot v) < g_i(x^*) + \varepsilon < 0$ ..... pro $i \in \{1, \dots, m\} \setminus I_0$

→ pro $i \in I_0$ je g_i pseudokonkávní a

$$\delta^+_{(x^* + t \cdot v) - x^*} g_i(x^*) \leqq 0 \rightsquigarrow \text{tudíž} \quad g_i(x^* + t \cdot v) \leqq g_i(x^*) = 0 \quad \text{pro } i \in I_0$$

Tedy, klademe-li $\vec{\sigma}(t) = 0$, pak všechny body oblouku

$$x^* + t \cdot v + \vec{\sigma}(t)$$

jsou přípustné pro $t \in (0, t_0)$,

c. b. d.

ěta: (KKT podmínka optimality pro úlohu (NP) .... pomocí (CQ) a Farkasova lemmatu)

Nechť M je otevřená množina

$$\left.\begin{array}{l} f: M \to \mathbb{R} \\ g_1, \dots, g_m : M \to \mathbb{R} \end{array}\right\} \text{ jsou funkce}$$

$x^* \in M$ je přípustný bod .... platí $g_i(x^*) \leqq 0$ pro $i = 1, \dots, m$

Nechť f má v bodě $\underline{x}^*$ jednostrannou směrovou derivaci $\delta_v^+ f(x^*)$

pro každý směr $\underline{v}$ à la Fréchet .....

..... existuje funkce η taková, že

$$\forall \vec{v}: \ f(x^* + \vec{v}) = f(x^*) + \delta_{\vec{v}}^+ f(x^*) + \eta(\vec{v}) \quad \& \quad \lim_{v \to 0} \frac{\eta(v)}{\|v\|} = 0$$

a nechť funkcionál $p(v) = \delta_v^+ f(x^*)$ je v bodě $v = 0$ spojitý ..... $\lim_{v \to 0} \delta_v^+ f(x^*) = 0$

resp. ... je spojitý v každém bodě $\underline{v}$

Označme

$$I_0 = \{ i \in \{1, \dots, m\} \, ; \ g_i(x^*) = 0 \}$$

Nechť g_i mají v bodě $\underline{x}^*$ jednostranné směrové derivace $\delta_v^+ g_i(x^*)$

pro každý směr $\underline{v}$ pro $i \in I_0$.

Funkcionály $p(v) = \delta_v^+ f(x^*)$ a $q_i(v) = \delta_v^+ g_i(x^*)$ pro $i \in I_0$ buďtež lineární

[tedy f má v bodě $\underline{x}^*$ totální diferenciál (Fréchetovu derivaci) a g_i mají v bodě $\underline{x}^*$ Gâteauxovy derivace pro $i \in I_0$].

V bodě $\underline{x}^*$ buďtež splněna podmínka (CQ).

Jestliže bod $\underline{x}^*$ je lokálním řešením úlohy

$$(NP) \quad \min f(x) \quad \text{z.p.} \quad g_i(x) \leqq 0 \ \text{pro} \ i = 1, \dots, m$$

potom bod $\underline{x}^*$ je KKT-stacionární.

Důkaz:

- zvolme libovolný směr $\vec{v}$
- předpokládejme, že je splněna soustava nerovnic

$$q_i(\vec{v}) = \delta^+_{\vec{v}}\, g_i(x^*) \leqq 0 \qquad \text{pro } i \in I_0$$

- potom podle (CQ) existuje oblouk přípustných bodů tvaru

$$x^* + t\cdot\vec{v} + \vec{o}(t)$$

- protože bod $\underline{x}^*$ je bodem lokálního minima funkce f, platí

$$f(x^* + t\cdot\vec{v} + \vec{o}(t)) \geqq f(x^*) \qquad \text{...... je-li } t>0 \text{ dostatečně malé}$$

neboli

$$f(x^*) + \delta^+_{t\cdot\vec{v}+\vec{o}(t)} f(x^*) + \eta(t\cdot\vec{v}+\vec{o}(t)) \geqq f(x^*)$$

$$\delta^+_{t\cdot\vec{v}+\vec{o}(t)} f(x^*) + \frac{\|t\cdot\vec{v}+\vec{o}(t)\|}{\|t\cdot\vec{v}+\vec{o}(t)\|}\cdot\eta(t\cdot\vec{v}+\vec{o}(t)) \geqq 0 \qquad \Big| : t>0$$

$$\delta^+_{\vec{v}+\underbrace{\frac{\vec{o}(t)}{t}}_{\to 0 \,+\, \text{spojitost}}} f(x^*) + \left\|\vec{v}+\frac{\vec{o}(t)}{t}\right\| \cdot \underbrace{\frac{\eta(t\cdot\vec{v}+\vec{o}(t))}{\|t\cdot\vec{v}+\vec{o}(t)\|}}_{\to 0} \geqq 0 \qquad \Big| \lim_{t\to 0^+}()$$

$$p(\vec{v}) = \delta^+_{\vec{v}} f(x^*) \geqq 0$$

Tedy jsme odvodili:

$$\forall \vec{v}: \ (\, q_i(v) \leqq 0 \ \text{pro } i \in I_0) \implies -p(\vec{v}) \leqq 0$$

- podle Farkasova lemmatu

$$\exists\, u^T \in \mathbb{R}^{1\times I_0},\ u^T \geqq o^T: \quad -p = u^T\cdot q$$

- neboli

$$o^T = df(x^*) + \sum_{i=1}^{m} u_i \cdot Dg_i(x^*)$$

kde jsme položili $u_i = 0$ pro $i \in \{1,\dots,m\} \setminus I_0$

což znamená, že bod $\underline{x}^*$ je KKT-stacionární, c.b.d.

Poznámka: Kvůli použití Farkasova lemmatu jsme museli předpokládat, že funkcionály $p(v) = \delta^+_v f(x^*)$ a $q_i(v) = \delta^+_v g_i(x^*)$ jsou lineární. Kdybychom znali obecnější Farkasovo lemma, mohli bychom předpoklád[at], že funkcionály $p(v)$ a $q_i(v)$ jsou obecnější, např. sublineární.