MPE_VPAM: Vybrané partie aplikované matematiky

Petr Zemánek (Masarykova Univerzita, Brno)

Kapitola 7: Maximalizace s omezujícími podmínkami

(verze: 24. září 2020)

Omezení ve tvaru rovností

Určeme nyní maximum funkce $F : \mathbb{R}^N \to \mathbb{R}$, přičemž požadujeme, aby proměnná $x \in \mathbb{R}^N$ vyhovovala M rovnicím

$$\left.\begin{aligned} g_1(x) &= b_1 \\ &\vdots \\ g_M(x) &= b_M \end{aligned}\right\} \begin{array}{l} g_i : \mathbb{R}^N \to \mathbb{R} \\ b_1, \dots, b_M \in \mathbb{R} \end{array} \qquad \text{neboli} \qquad \begin{aligned} \tilde{g}_1(x) &= 0 \\ &\vdots \\ \tilde{g}_M(x) &= 0 \end{aligned}$$

Souhrnně lze psát tato omezení ve tvaru vektorové funkce

$$G(x) = b \quad \text{neboli} \quad \widetilde{G}(x) = 0,$$

kde $G : \mathbb{R}^N \to \mathbb{R}^M$, $b \in \mathbb{R}^M$ je daný vektor (pro $\widetilde{G}$ je $b = 0$), přičemž

$$G(x) = \begin{pmatrix} g_1(x) \\ \vdots \\ g_M(x) \end{pmatrix} \qquad b = \begin{pmatrix} b_1 \\ \vdots \\ b_M \end{pmatrix}$$

Množina vektorů $x \in \mathbb{R}^N$ splňujících omezení $\widetilde{G}(x) = 0$ se značí jako $\mathcal{C}$ (constrained set). Lokální maximum funkce F s omezujícími podmínkami $\widetilde{G}(x) = 0$ (tzv. vázané lokální maximum) je bod $x^* \in \mathcal{C}$ s vlastností, že existuje okolí A bodu x^* takové, že

$$F(x) \leqslant F(x^*) \quad \text{pro každé } x \in A \cap \mathcal{C}.$$

Bod x^* je globální maximum s omezujícími podmínkami $\widetilde{G}(x) = 0$ (tzv. vázané globální maximum), pokud

$$F(x) \leqslant F(x^*) \quad \text{pro každé } x \in \mathcal{C} \quad (\text{tj. lze zvolit } A = \mathbb{R}^N).$$

Jaký je vztah mezi počtem proměnných a počtem omezujících rovnic? Má smysl uvažovat zejména $N \geqslant (>) M$.

Věta 7.5(i)

Nechť funkce $F, G \in C^1$ a nechť bod $x^* \in \mathcal{C}$ je lokálním extrémem s omezujícími podmínkami $G(x) = 0$. Předpokládejme dále, že $M \times N$ matice $DG(x^*)$ má plnou hodnost. Potom existuje vektor multiplikátorů $\lambda \in \mathbb{R}^M$ takový, že bod x^* je stacionárním bodem funkce (tzv. Lagrangeova funkce)

$$L(x, \lambda) \coloneqq F(x) + \langle \lambda, G(x) \rangle = F(x) + \lambda^\top G(x), \quad \text{tj.} \quad DF(x^*) + \lambda^\top DG(x^*) = 0.$$

Vektor $\lambda \in \mathbb{R}^M$ se nazývá vektor Lagrangeových multiplikátorů (v ekonomii se λ odpovídající extrému nazývá stínovou cenou, viz později). Pokud $\lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_M \end{pmatrix}$, kde $\lambda_i \in \mathbb{R}$ pro $i = 1, \dots, M$, pak je podmínka na stacionární bod ve tvaru

$$\operatorname{grad} F(x^*) + \sum_{m=1}^{M} \lambda_m \operatorname{grad} G_m(x^*) = 0.$$

Podmínka na hodnost $h(DG(x^*)) = M$ znamená, že je splněn požadavek tzv. kvalifikovaného omezení (constraint qualification). Tato podmínka udává, že omezení v bodě x^* jsou nezávislá (přesněji – linearizace těchto omezení jsou lineárně nezávislé) $\rightsquigarrow$ $\operatorname{grad} g_1(x^*), \dots, \operatorname{grad} g_M(x^*)$.

Příklad

Požadavek kvalifikovaného omezení nelze vyloučit (tj. nelze vypustit nezávislost lineárních kombinací gradientů omezení)! Pro $N = 2$ a omezující množinu

$$\mathcal{C} = \{[x, y] \in \mathbb{R}^2 \mid g_1(x, y) = 0,\ \ g_2(x, y) = 0\}$$

najděte maximum lineární funkce $F(x, y)$.

Příklad

Uvažte jednoduchou maximalizaci užitkové funkce

$$F(x, y) = xy$$

s omezením $x + 4y = 16$.
Video: ☞ 😎 🖵 https://mpevpam.page.link/2xbH
Zápisník: ☞ 😎 https://mpevpam.page.link/4UZW

Příklad

$\max F(x,y) = x^2 y$ v omezující množině $\mathcal{C} = \{[x,y] \in \mathbb{R}^2 \mid 2x^2 + y^2 = 3\}$.
Video: ☞ 😎 🖵 https://mpevpam.page.link/bGSr
Zápisník: ☞ 😎 https://mpevpam.page.link/5fHW

Příklad

$\max F(x,y,z) = xyz$ s omezeními $x^2 + y^2 = 1$ a $x + z = 1$.
Video: ☞ 😎 🖵 https://mpevpam.page.link/eomQ
Zápisník: ☞ 😎 https://mpevpam.page.link/vsjk

Příklad

$\max F(x,y,z) = x^2 + 4y^2 + 9z^2$ na množině $x^2 + y^2 + z^2 = 1$.
Video: ☞ 😎 🖵 https://mpevpam.page.link/nNG5
Zápisník: ☞ 😎 https://mpevpam.page.link/WbCG

Pro rozhodnutí o tom, zda je či není v daném bodě vázaný extrém máme podmínky 2. řádu. Nutná podmínka pro maximum:

$$D_x^2 L(x^*, \lambda) = \underbrace{\underbrace{D^2 F(x^*)}_{\mathbb{R}^{N\times N}} + \underbrace{\underbrace{\lambda^\top}_{\mathbb{R}^{1\times M}} \underbrace{D^2 G(x^*)}_{\mathbb{R}^{M\times N\times N}}}_{\mathbb{R}^{N\times N}}}_{\mathbb{R}^{N\times N}} \leqslant 0 \quad \text{na} \ \operatorname{Ker} DG(x^*).$$

Pozor!

Nalezení vázaného lokálního extrému neznamená nalezení lokálního extrému Lagrangeovy funkce!

A postačující podmínka?

Věta 7.9(i)

Nechť $F, g_1, \dots, g_M \in C^2$ jsou funkce definované na $\mathbb{R}^N$. Uvažujme problém

$$\max F(x)$$

na omezující množině

$$\mathcal{C} = \{x \in \mathbb{R}^N \mid g_1(x) = 0, \dots, g_M(x) = 0\}.$$

Uvažujme Lagrangeovu funkci $L(x, \lambda)$ z Věty 7.5(i) a nechť

- $x^* \in \mathcal{C}$, tj.

$$g_1(x^*) = 0 \quad \& \quad g_2(x^*) = 0 \quad \& \quad \cdots \quad \& \quad g_M(x^*) = 0$$

- existují $\lambda_1^*, \dots, \lambda_M^*$ taková, že

$$\frac{\partial L}{\partial x_1}(x^*, \lambda^*) = 0, \dots, \frac{\partial L}{\partial x_N}(x^*, \lambda^*) = 0$$

- Hessova matice $D_x^2 L(x^*, \lambda^*) \in \mathbb{R}^{N \times N}$ je negativně definitní na množině $\{v \in \mathbb{R}^N \mid DG(x^*)v = 0\}$, tj.

$$v \in \operatorname{Ker} DG(x^*),\ v \neq 0 \implies v^\top \big(D_x^2 L(x^*, \lambda^*)\big) v < 0.$$

Potom bod x^* je ostré lokální vázané maximum funkce F na množině $\mathcal{C}$.

Poznámka

Kdy je splněna druhá podmínka z Věty 7.9(i)?

Jak ověřit třetí podmínku z Věty 7.9(i)? Speciálně pro $N = 2$. A je-li funkce $L(x, \lambda^*)$ konkávní?

Poznámka

Hledání vázaného lokálního minima?

Příklad (pokrač.)

$\max F(x, y, z) = x^2 + 4y^2 + 9z^2$ na množině $x^2 + y^2 + z^2 = 1$ (různé postupy).

Příklad

$F(x, y) = (x - 3)^2 + (y - 3)^2 \to$ extr. při omezení $x^2 + y^2 = 2$.

Příklad (ekonomie)

Spotřebitel užívá dva statky a jeho užitková funkce je tvaru $u(x, y) = 100xy + x + 2y$, přičemž statek A stojí 2 Kč a statek B stojí 4 Kč. Spotřebitel má k dispozici 1000 Kč, které chce všechny utratit za oba tyto statky tak, aby maximalizoval svůj užitek.

Příklad (ekonomie)

Nákladová funkce firmy je tvaru

$$0,05\,Q_A^2 + 0,01\,Q_A\,Q_B + 0,03\,Q_B^2 + 10\,Q_A + 15\,Q_B + 12000,$$

kde Q_A a Q_B je množství výrobků firmy, které jsou vyráběny ve firemních odděleních A a B. Jestliže má firma smlouvu na produkci 2000 výrobků, jak má rozdělit svoji výrobu, aby minimalizovala náklady?

Příklad (ekonomie)

Roční produkční funkce firmy je tvaru $20\,u^{0,3}\,v^{0,2}\,w^{0,5}$, kde u, v a w jsou množství vstupů U, V a W za jeden rok, které firma spotřebuje k produkci výstupu. Ceny za jednotu vstupů U, V a W jsou po řadě $P_u = 1000$ Kč, $P_v = 2500$ Kč a $P_w = 4000$ Kč. Je-li roční rozpočet 1 milion Kč, jaké množství vstupů má firma nakoupit, aby maximalizovala svoji produkci?

Příklad (ekonomie)

Uvažte tři cenné papíry s hodnotami rizika $\sigma_1 = 1,2$, $\sigma_2 = 0,8$, $\sigma_3 = 1,1$ a hodnotami výnosností $r_1 = 0,8$, $r_1 = 0,3$ a $r_3 = 0,6$. Kovariance jsou $\sigma_{1,2} = -0,1$, $\sigma_{1,3} = -0,5$ a $\sigma_{2,3} = 0,3$. Určete skladbu portfolia obsahujícího pouze tyto cenné papíry tak, aby riziko bylo minimální s výnosem 35 %.

Omezení ve tvaru nerovností

Uvažujme úlohu

$$
\left.\begin{array}{c}
F(x) \to \max \\[1em]
\text{za podmínek } x \in \mathbb{R}^N \text{ a} \\[1em]
\left.\begin{array}{r} g_1(x) = 0, \dots, g_M(x) = 0 \\ h_1(x) \leqslant 0, \dots, h_K(x) \leqslant 0 \end{array}\right\} \begin{array}{l} G(x) = 0,\ G : \mathbb{R}^N \to \mathbb{R}^M \\ H(x) \leqslant 0,\ H : \mathbb{R}^N \to \mathbb{R}^K \end{array}
\end{array}\right\} \tag{1}
$$

(přičemž je povoleno $M = 0$ i/nebo $K = 0$; smysl dává především $N \geqslant M + K$). Opět označíme množinu „přípustných" vektorů jako

$$\mathcal{C} \coloneqq \{x \in \mathbb{R}^N \mid G(x) = 0 \ \&\ H(x) \leqslant 0\}.$$

Řekneme, že v bodě x^* je splněna podmínka kvalifikovaného omezení, pokud jsou gradienty těch omezení, pro které platí rovnost (tzv. aktivní omezení), lineárně nezávislé, tj. vektory

$$\operatorname{grad} g_1(x^*), \dots, \operatorname{grad} g_M(x^*), \operatorname{grad} h_k(x^*),\ k \in \{1, \dots, K\} \text{ taková, že } h_k(x^*) = 0,$$

jsou lineárně nezávislé.

Příklad

$\max F(x,y) = x^2 + xy - y^2$ za podmínek $x + y = 2$, $x - y \leqslant 4$.

Příklad (ekonomie)

Jedním z nejzákladnějších problémů v ekonomii je maximalizace užitku. Nechť x_i reprezentuje množství komodity i a $F(x_1, \dots, x_N) = U(x_1, \dots, x_N)$ určuje míru individuálního užitku z konzumace x_1 jednotek prvního zboží, x_2 jednotek druhého zboží atd. Nechť $p_1, \dots, p_N$ značí ceny jednotlivých komodit a nechť I je dostupný příjem. Potom vlastně požadujeme

$$\max U(x_1, \dots, x_N) \quad \& \quad p_1 x_1 + \dots + p_N x_N \leqslant I \quad \& \quad x_1 \geqslant 0, \dots, x_N \geqslant 0$$

(měli bychom spíše psát $-x_1 \leqslant 0, \dots, -x_N \leqslant 0$).

Příklad (ekonomie)

Nyní maximalizujme užitek v případě, že si uživatel může rozhodnout o tom, jak rozloží svůj čas na práci a volný čas. Použijeme stejné značení jako v předchozím příkladě a navíc w značí tarifní mzdu, I' netarifní složku mzdy, ℓ_0 počet hodin práce a ℓ_1 počet hodin volna. Zákazník má tedy obnos $I' + w\ell_0$ na utracení a požaduje

$$\begin{aligned} \max U(x_1, \dots, x_N) \quad &\& \quad p_1 x_1 + \dots + p_N x_N \leqslant I' + w\ell_0 \quad \& \quad \ell_0 + \ell_1 = 24 \\ &\& \quad x_1 \geqslant 0, \dots, x_N \geqslant 0, \ell_0 \geqslant 0, \ell_1 \geqslant 0. \end{aligned}$$

Příklad (ekonomie)

Maximalizace zisku firmy v konkurenčním prostředí: uvažujme firmu v konkurenčním průmyslovém odvětví, která používá N vstupů k výrobě svého produktu. Nechť y značí množství výstupů a $x_1, \ldots, x_N$ množství jednotlivých složek. Nechť $y = F(x_1, \ldots, x_N)$ značí produkční funkci firmy určující maximální množství výrobků, které lze získat z koše vstupů $(x_1, \ldots, x_N)$. Nechť p je jednotková cena výstupu a w_i jednotkové ceny vstupů. Cílem firmy je proto najít hodnoty $x_1, \ldots, x_N$ tak, aby

$$\max \Pi(x_1, \ldots, x_N) = \max \left[pF(x_1, \ldots, x_N) - \sum_{i=1}^{N} w_i x_i \right]$$

za podmínek

$$pF(x_1, \ldots, x_N) - \sum_{i=1}^{N} w_i x_i \geqslant 0$$

$$g_1(x) \leqslant b_1, \ldots, g_k(x) \leqslant b_k \quad \& \quad x_1 \geqslant 0, \ldots, x_N \geqslant 0,$$

kde druhé omezení určuje dostupnost jednotlivých komponent.

Označme jako $\mathcal{K}^*$ množinu indexů aktivních omezení z nerovností, tj.

$$\mathcal{K}^* := \{k \in \{1, \dots, K\} \mid h_k(x^*) = 0\}.$$

Věta 7.16(i) "KKT"

Nechť funkce $F : \mathbb{R}^N \to \mathbb{R}$, $G : \mathbb{R}^N \to \mathbb{R}^M$, $H : \mathbb{R}^N \to \mathbb{R}^K$ jsou třídy C^1 a nechť je bod $x^* \in \mathbb{R}^N$ lokálním extrémem úlohy

$$F(x) \to \text{extr.} \quad \& \quad G(x) = 0 \quad \& \quad H(x) \leqslant 0$$

a platí podmínka kvalifikovaného omezení v bodě x^*, tj. gradienty

$$\operatorname{grad} g_1(x^*), \dots, \operatorname{grad} g_M(x^*), \operatorname{grad} h_k(x^*), \ k \in \mathcal{K}^*$$

jsou lineárně nezávislé. Pak existují multiplikátory $\lambda \in \mathbb{R}^M$ a $\mu \in \mathbb{R}^K$ (přičemž $\mu \leqslant 0$ pro maximum a $\mu \geqslant 0$ pro minimum) takové, že

- bod x^* je stacionárním bodem *Lagrangeovy funkce*, tj.

$$DF(x^*) + \lambda^\top DG(x^*) + \mu^\top DH(x^*) = 0,$$

- a platí *podmínka komplementarity*

$$\mu_k h_k(x^*) = 0 \quad \text{pro každé } k \in \{1, \dots, K\},$$

(zejména tedy $\mu_k = 0$ pro $k \notin \mathcal{K}^*$).

Poznámka

Pro $\lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_M \end{pmatrix}$ a $\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_M \end{pmatrix}$ máme $\mu_1 \leqslant 0, \dots, \mu_K \leqslant 0$ pro maximum a $\mu_1 \geqslant 0, \dots, \mu_K \geqslant 0$ pro minimum, přičemž $\mu_k = 0$ pro každé $k \notin \mathcal{K}^*$ (tj. $\mu_k = 0$ pro indexy těch omezení, které se v bodě x^* realizují jako skutečné nerovnosti, tzv. neaktivní omezení), a rovnice (16) je pak tvaru

$$\operatorname{grad} F(x^*) + \sum_{m=1}^{M} \lambda_m \operatorname{grad} g_m(x^*) + \sum_{k=1}^{K} \mu_k \operatorname{grad} h_k(x^*) = 0$$

neboli

$$F_{x_n}(x^*) + \sum_{m=1}^{M} \lambda_m \, (g_m)_{x_n}(x^*) + \sum_{k=1}^{K} \mu_k \, (h_k)_{x_n}(x^*) = 0$$

pro každé $n = 1, \dots, N$. Tedy (16) znamená, že bod $[x^*, \lambda, \mu]$ je stacionárním bodem Lagrangeovy funkce

$$L(x, \lambda, \mu) \coloneqq F(x) + \lambda^\top G(x) + \mu^\top H(x), \quad L : \mathbb{R}^N \times \mathbb{R}^M \times \mathbb{R}^K \to \mathbb{R},$$

přičemž z (16) je vidět, že do Lagrangeovy funkce vstupují ve skutečnosti pouze ta omezení, která se nakonec realizují jako rovnosti.

Poznámka

Podmínky na znaménko multiplikátorů $\mu \leqslant 0$ pro maximum a $\mu \geqslant 0$ pro minimum si lze snadno zapamatovat: je to stejné jako podmínka na znaménko pro druhou derivaci ($\leqslant$ pro maximum a $\geqslant$ pro minimum) — toto ovšem platí, pokud je Lagrangeova funkce ve tvaru výše, tj. s $\oplus$ u členů s omezeními

$$F(x) \oplus \lambda^{\top} G(x) \oplus \mu^{\top} H(x) = 0.$$

Příklad

$F(x, y) = (x-3)^2 + (y-3)^2 \to \min$ při omezení $x^2 + y^2 \leqslant 2$.

Poznámka

V aplikacích bývá často uvažováno omezení proměnných na znaménka, tj. např. $x_n \geqslant 0$ pro nějaké $n = 1, \dots, N$. Uvažme tedy úlohu

$$F(x) \to \text{extr.} \quad G(x) = 0 \quad \& \quad H(x) \leqslant 0 \quad \& \quad x \geqslant 0 \rightsquigarrow -x \leqslant 0,$$

Je-li bod $x^* \in \mathbb{R}^N$ řešením a platí-li požadavek kvalifikovaného omezení, pak podle Věty 7.16(i) existují multiplikátory $\lambda \in \mathbb{R}^M$, $\mu \in \mathbb{R}^K$, $\nu \in \mathbb{R}^N$, přičemž $\mu \leqslant 0, \nu \leqslant 0$ pro maximum a $\mu \geqslant 0$, $\nu \geqslant 0$ pro minimum, tak, že

$$DF(x^*) + \lambda^\top DG(x^*) + \mu^\top DH(x^*) + \nu^\top D(-x)\Big|_{x=x^*} = 0$$

tedy

$$DF(x^*) + \lambda^\top DG(x^*) + \mu^\top DH(x^*) - \nu^\top = 0,$$

kde $\mu_k = 0$ pro omezení $h_k(x^*) < 0$, $k \in \{1, \dots, K\}$, a $\nu_n = 0$ pro omezení $(x^*)_n > 0$, $n \in \{1, \dots, N\}$. Celkem tedy máme podmínku prvního řádu

$$DF(x^*) + \lambda^\top DG(x^*) + \mu^\top DH(x^*) \leqslant 0 \quad \text{pro maximum,}$$
$$DF(x^*) + \lambda^\top DG(x^*) + \mu^\top DH(x^*) \geqslant 0 \quad \text{pro minimum}$$

a rovnost platí pro ty souřadnice, kde je $(x^*)_n > 0$.

Poznámka

Podmínky 2. řádu jsou analogické jako na str. 8 a ve Větě 7.9(i), ale musí být splněna tzv. ostrá podmínka komplementarity, tj. $\mu_k > 0$ pro $k \in \mathcal{K}^*$. Pak stačí uvažovat tečný prostoru ke všem omezením ve tvaru rovností, tj. k omezením $G(x^*) = 0$ (tj. $g_m(x^*) = 0$ pro každé $m = 1, \dots, M$), a současně k omezením $h_k(x^*) = 0$ pro $k \in \mathcal{K}^*$. (Bez ostré podmínky komplementarity bude uvažovaný prostor mnohem komplikovanější!) Pokud máme v úloze také omezení na znaménko, tak je nemusíme zahrnovat mezi nerovnostní omezení (tj. zvětšovat rozměr úlohy), jestliže jsou v uvažovaném stacionárním bodě tyto nerovnosti ostré (tj. tato omezení jsou neaktivní).

Poznámka

Zde pak vyšetřujeme negativní definitnost (pro maximum) nebo pozitivní definitnost (pro minimum) pro $D_x^2 L(x^*, \lambda^*, \mu^*)$. Např. pomocí rozšířené Hessovy matice $H_{M+|\mathcal{K}^*|+N}(x^*, \lambda^*, \mu^*)$

$$\left(\begin{array}{cccccc|ccc}
0 & \cdots & 0 & 0 & \cdots & 0 & \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & 0 & \cdots & 0 & \frac{\partial g_M}{\partial x_1} & \cdots & \frac{\partial g_M}{\partial x_n} \\
0 & \cdots & 0 & 0 & \cdots & 0 & \frac{\partial h_{k_1}}{\partial x_1} & \cdots & \frac{\partial h_{k_1}}{\partial x_n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & 0 & \cdots & 0 & \frac{\partial h_{k_e}}{\partial x_1} & \cdots & \frac{\partial h_{k_e}}{\partial x_n} \\
\hline
\frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_M}{\partial x_1} & \frac{\partial h_{k_1}}{\partial x_1} & \cdots & \frac{\partial h_{k_e}}{\partial x_1} & \frac{\partial^2 L}{\partial x_1^2} & \cdots & \frac{\partial^2 L}{\partial x_n x_1} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\frac{\partial g_1}{\partial x_n} & \cdots & \frac{\partial g_M}{\partial x_n} & \frac{\partial h_{k_1}}{\partial x_n} & \cdots & \frac{\partial h_{k_e}}{\partial x_n} & \frac{\partial^2 L}{\partial x_1 x_n} & \cdots & \frac{\partial^2 L}{\partial x_n^2}
\end{array}\right)$$

Pak pro maximum musí $N - (|\mathcal{K}^*| + M)$ VHM střídat znaménko a číslo $\det H_{M+|\mathcal{K}^*|+N}(x^*, \lambda^*, \mu^*)$ musí mít stejné znaménko jako $(-1)^N$.

V případě $M = 0$ (tj. žádné omezení ve tvaru rovností) lze Větu 7.16(i) přeformulovat (variace na základní větu konvexního programování).

Věta 7.22(i)

Nechť jsou funkce $h_k \in C^1$, $k = 1, \dots, K$ kvazikonvexní a funkce $F \in C^1$ je konkávní (kvazikonkávní). Jestliže bod $x^* \in \mathbb{R}^N$ splňuje KKT podmínky

$$DF(x^*) + \mu^\top DH(x^*) = 0 \quad \& \quad \mu \leqslant 0 \quad \& \quad \mu_k h_k(x^*) = 0$$

pro každé $k = 1, \dots, K$ a jestliže je splněn požadavek kvalifikovaného omezení, tj. vektory $\operatorname{\mathbf{grad}} h_k(x^*)$ pro $k \in \mathcal{K}^*$ jsou lineárně nezávislé, (a navíc $\operatorname{\mathbf{grad}} F(x^*) \neq 0$), potom bod x^* je globálním maximem funkce F.

Poznámka

- Omezení $h_k(x) \leqslant 0$ pomocí kvazikonvexních funkcí udávají konvexní množinu $\mathcal{C}$, do které patří body $x \in \mathbb{R}^N$.
- Předchozí tvrzení je možné upravit i pro případ, že omezení ve tvaru rovnosti jsou zadána pomocí funkcí $g_i(x) = a_i x + b_i$ pro $a_i \in \mathbb{R}^N$ a $b_i \in \mathbb{R}$.

Ještě jeden výsledek podtrhující význam konvexnosti množiny $\mathcal{C}$ a ostré kvazikonkávnosti funkce F.

Věta 7.22(ii)

Předpokládejme, že pro problém (1) je množina $\mathcal{C}$ konvexní a funkce F je (ostře) kvazikonkávní. Potom množina řešení (globálních maxim) je konvexní (nejvýše jednoprvková).

KKT v diferenciálním tvaru

Uvažme úlohu

$$F(x) \to \max, \quad x \in X, \tag{2}$$

kde *přípustná množina* X je zadána systémem rovností a nerovností

$$X := \{x \in P \mid g_i(x) = 0,\ h_j(x) \leqslant 0,\ i = 1, \dots, M,\ j = 1, \dots, K\}, \tag{3}$$

přičemž množina $P \subseteq \mathbb{R}^N$ je konvexní, funkce $g_1, \dots, g_M$ *afinní* na P, funkce F konkávní a funkce $h_1, \dots, h_K$ konvexní na P a diferencovatelné na (nějaké otevřené množině obsahující) P. Nechť navíc platí (alespoň) jedna z následujících podmínek (regularity):

(i) (LNZ) množina P je otevřená, vektory $\operatorname{grad} g_i(x)$, $\operatorname{grad} h_j(x)$ pro $i = 1, \dots, M$ a $j \in \mathcal{K}^*(x)$ jsou lineárně nezávislé pro každé $x \in X$;

(ii) (Slaterova) $M = 0$ a existuje bod $\overline{x} \in P$ takový, že $h_j(\overline{x}) < 0$ pro $j = 1, \dots, K$;

(iii) (lineární) množina P je *polyedr* ("průnik poloprostorů") a funkce $h_1, \dots, h_K$ jsou afinní.

Pak x^* je řešením úlohy (2) & (3) právě tehdy, když existují multiplikátory $\lambda \in \mathbb{R}^M$ a $\mu \in \mathbb{R}^K_-$ splňující

$$\langle \operatorname{grad}_x L(x^*, \lambda, \mu),\ x - x^* \rangle \geqslant 0 \quad \text{pro všechna } x \in P$$
$$\mu_j h_j(x^*) = 0, \quad j = 1, \dots, K.$$

Příklad (pokrač.)

$F(x,y) = (x-3)^2 + (y-3)^2 \to \min$ při omezení $x^2 + y^2 \leqslant 2$.

Příklad

Uvažte úlohu

$$F(x,y) = x^2 + 2xy + 2y^2 - 10x - 12y \to \text{extr.}$$
$$x + 3y \leqslant 8 \quad \& \quad x^2 + y^2 + 2x - 2y \leqslant 3.$$

Ověřte, zda platí v bodě $[1,2]$ kvalifikovanost omezení. Ukažte, že bod $[x^*, y^*] = [1,2]$ splňuje KKT podmínky 1. řádu a rozhodněte podle testu druhých derivací, zda se jedná o vázané lokální minimum.

Příklad (ekonomie)

Užitek z konzumace x jednotek výrobku A a y jednotek výrobku B je dán funkcí

$$U(x,y) = \ln x + \ln y.$$

Cena za jednotku výrobku A je 10,- a výrobku B je 5,-. Máme nejvýše 350,- k útratě. Současně platí, že konzumace 1 jednotky výrobku A zabere 0,1 hod. a 0,2 hod. pro výrobek B. Máte maximálně 8 hodin. Jak maximalizovat užitek?

Příklad (ekonomie)

Cena ve špičkách: uvažme producenta elektrické energie, který používá uhlí nebo zemní plyn. Poptávka po elektřině kolísá mezi špičkami, během nichž je využita veškerá produkční kapacita, a mimošpičkovými obdobími, kdy je k dispozici volná výrobní kapacita. Mějme nějaký časový interval (např. 1 rok), který rozdělíme na N stejně dlouhých období. Předpokládejme, že množství prodané elektrické energie v těchto N obdobích je $x_1, \dots, x_N$. Nechť také existuje regulátor, který stanoví ceny v jednotlivých obdobích na $p_1, \dots, p_N$. Celkové provozní náklady v průběhu všech N období jsou dány funkcí $C(x_1, \dots, x_N)$ a k je celková výrobní kapacita. Nechť dále $D(k)$ značí cenu udržení produkce na úrovni k. Celkový zisk producenta je potom

$$\Pi(x_1, \dots, x_N) = \sum_{i=1}^{N} p_i\, x_i - C(x) - D(k).$$

Protože nelze překročit kapacitu k v žádném okamžiku, platí omezení

$$x_1 \leqslant k, \dots, x_N \leqslant k. \tag{4}$$

Uvažme problém nalezení $x_1 \geqslant 0, \dots, x_N \geqslant 0, k \geqslant 0$ tak, že maximalizujeme zisk za omezení (4).

Příklad

Určete $F(x,y) = 2x^2+2xy+y^2-20x-14y \to$ extr. za omezení $x+3y \leqslant 5$ a $2x-y \leqslant 4$.
Video: ☞ 😎 🖵 https://mpevpam.page.link/VE8U
Zápisník: ☞ 😎 https://mpevpam.page.link/AhQK

Příklad

Určete $F(x,y) = x^2/2 - xy + y^2 - 2x + y \to \min$ za omezení $x+y \leqslant 3$, $2x-y \leqslant 4$, $x \geqslant 0$ a $y \geqslant 0$.
Video: ☞ 😎 🖵 https://mpevpam.page.link/tuHo
Zápisník: ☞ 😎 https://mpevpam.page.link/zL9s

Příklad

Určete $F(x,y) = xy \to \max$ za omezení $x+y \geqslant 6$, $x \geqslant 0$, $y \geqslant 0$.

Příklad

Napište Kuhnovy–Tuckerovy podmínky 1. řádu pro úlohu $F(x,y,z) = -x^3+y^3-2xz^2 \to \min$ za omezení $2x+y^2+z=5$, $5x^2-y^2-z \geqslant 2$, $x \geqslant 0$ a $z \geqslant 0$. Ukažte také, že bod $[1,0,3]$ je bodem lokálního minima.

Příklad

Určete $F(x,y) = y^2+x+1 \to \min$ za omezení $x^3-y^3 \leqslant 1$ a $x \geqslant 0$.

Příklad (ekonomie)

Firma vyrábějící letadla je může provozovat ve dvou různých zemích. Náklady na provoz $x \geqslant 0$ kusů letadel v zemi A jsou $C_A(x) = \ln(1 + 3x/100)$. Náklady na provoz $y \geqslant 0$ kusů letadel v zemi B jsou $C_B(y) = 2\ln(1 + y/100)$. Firma chce rozdělit svoji produkci mezi x a y tak, aby minimalizovala své celkové náklady na provoz při výrobě nejméně q kusů letadel, kde $q > 0$.

(i) Ukažte, že optimální volba x a y musí být řešení jistého optimalizačního problému s „omezením na znaménko“.

(ii) Ukažte, že v závislosti na hodnotě q existuje jeden, dva nebo tři kandidáti na řešení tohoto problému.

(iii) Ukažte, že v optimálním případě má firma dodávat letadla pouze do země A, je-li q pro jistou kritickou hodnotou q^*, a pouze do země B, je-li $q > q^*$.

(iv) Určete minimální náklady v závislosti na q a ukažte, že tato funkce není diferencovatelná v bodě q^*.

Nyní budeme hledat maximum funkce na množině určené pomocí omezení ve tvaru rovností, přičemž účelová funkce i funkce zadávající omezení bude záviset na „nějakých“ parametrech, tj.

$$\left.\begin{aligned} F(x;p) &\to \max, \quad F: \mathbb{R}^N \times \mathbb{R}^S \to \mathbb{R} \\ \left.\begin{aligned} g_1(x;p) &= 0 \\ &\vdots \\ g_M(x;p) &= 0 \end{aligned}\right\} & \begin{aligned} &G(x;p) = 0, \quad G: \mathbb{R}^N \times \mathbb{R}^S \to \mathbb{R}^M \\ &p = (p_1, \dots, p_S)^\top \in \mathbb{R}^S \dots \text{parametry} \end{aligned} \end{aligned}\right\} \tag{5}$$

Je-li $x^* = x^*(p)$ řešení úlohy (5) (to také závisí na hodnotě parametru p), pak dosazením tohoto řešení $x^*(p)$ do účelové funkce F dostaneme její optimální (max.) hodnotu $F(x^*(p); p)$ na dané přípustné množině. Tato optimální hodnota opět závisí na parametrech p a nazývá se <u>funkcí optimální hodnoty</u> (value function)

$$v(p) \coloneqq F(x^*(p); p) \quad (\text{též } v(p) \rightsquigarrow F^*(p))$$

Otázka: jak se změní hodnota $v(p)$ při změně parametrů p? Jinými slovy, jaká je závislost funkce $v(p)$ na parametru p? A proč by nás to vlastně mělo zajímat?

Nejdříve uvažme úlohu bez omezení, tj. pouze

$$F(x;p) \to \max .$$

Pokud pro jednoduchost připustíme, že všechny níže potřebné derivace existují, pak s pomocí pravidla pro derivování složené funkce dostaneme

$$D_p v(p) = \underbrace{F_x(x^*(p);p)}_{=0 \text{ pro max. (i min.)}} D_p x^*(p) + D_p F(x^*(p);p).$$

Dostáváme tedy

$$D_p v(p) = D_p F(x^*(p);p), \quad \text{tj.} \quad \frac{\partial}{\partial p_i} v(p) = \frac{\partial}{\partial p_i} F(x^*(p);p),$$

což je speciální případ tzv. „Věty o obálce“: změna funkce optimální hodnoty vzhledem k parametru p je tudíž dána pouze přímou změnou funkce F v parametru p (vyhodnocenou v optimálním řešení) a již ne změnou funkce F v proměnné x. Zejména tak v případě jednoho parametru $p \in \mathbb{R}$ platí, že

$$v(p+\Delta p) \approx v(p) + \frac{d}{dp} F(x^*(p);p)\, \Delta p.$$

Obecně pro $p \in \mathbb{R}^s$ máme přibližný odhad

$$v(p+\Delta p) \approx v(p) + \frac{\partial}{\partial p_1} F(x^*(p);p)\, \Delta p_1 + \cdots + \frac{\partial}{\partial p_s} F(x^*(p);p)\, \Delta p_s .$$

Příklad (ekonomie)

Firma určuje množství zboží x na vstupu tak, aby maximalizovala svůj zisk daný funkcí

$$ph(x) - wx,$$

kde h je výrobní funkce (určuje počet výrobků/zboží na výstupu při vstupu x), p je prodejní cena za jeden výrobek a w jsou náklady na výrobu jednoho výrobku. Chceme tedy

$$\max F(x; p) := ph(x) - wx$$

bez dalších omezení. Je-li $x^*(p)$ optimální řešení (tj. maximum), pak má funkce optimální hodnoty funkce tvar

$$v(p) = F(x^*(p); p) = ph(x^*(p)) - wx^*(p).$$

V tomto případě lze maximum najít standardním způsobem z podmínek prvního a druhého řádu, tj.

$$0 = D_x F(x; p) = ph'(x) - w \quad \Longrightarrow \quad h'(x) = \frac{w}{p},$$

což by po vyřešení dalo hledané $x^* = x^*(p)$.

Příklad (ekonomie; pokrač.)

Při značení $\dot{} := \frac{d}{dp}$ pak v souladu s předchozími výpočty dostáváme

$$\dot{v}(p) = \frac{d}{dp}[p\, h(x^*(p)) - w\, x^*(p)] = h(x^*(p)).$$

Je-li tedy $x^*(p)$ optimální počet výrobků na vstupu maximalizující zisk při ceně p za výrobek na výstupu, pak se aktuální maximální zisk mění přesně stejně jako počet výrobků $h(x^*(p))$, který se má z těchto $x^*(p)$ výrobků vyrobit.

Příklad Jaký bude mít vliv zvýšení hodnoty p na hodnotu maxima funkce $F(x;p) = -x^2 + 2px + 4p^2$?

Příklad Uvažte $F(x;p) = -p^3x^4 + 15x^3 - e^p x^2 + 17 \to \max$ bez dalších omezení a pro p v okolí 1.

Příklad (ekonomie) Hotellingovo lemma.

Příklad (ekonomie) Brněnská firma vyrábí mikročipy. Nákladová funkce při výrobě y mikročipů je $c(y)$, přičemž $c'(y) > 0$ a $c''(y) > 0$. Část produkce v poměru daném $1-\alpha$ ($0 \leqslant \alpha \leqslant 1$) z vyrobených mikročipů je vadná a nelze ji prodat. Zbytek je prodán za pevnou cenu p, neboť tržní prostřední je velmi konkurenční. Jak se projeví zvýšení kvality produkce na zisku firmy?

Poznámka Proč „Věta o obálce“ (geometrická interpretace)?

Věta 7.34(i)

Nechť $v(p)$ je funkce optimální hodnoty pro problém (5). Předpokládejme, že funkce $v : \mathbb{R}^S \to \mathbb{R}$ je diferencovatelná v nějaké hodnotě parametru $p_0 \in \mathbb{R}^S$ a že $\lambda \in \mathbb{R}^M$ je Lagrangeův multiplikátor příslušný optimálnímu řešení $x^*(p_0)$ pro hodnotu parametru $p = p_0$, tj.

$$D_x F(x^*(p_0); p_0) + \lambda^\top D_x G(x^*(p_0); p_0) = 0.$$

Potom platí rovnost

$$D_p v(p_0) = D_p F(x^*(p_0); p_0) + \lambda^\top D_p G(x^*(p_0); p_0), \tag{6}$$

tj.

$$\frac{\partial v}{\partial p_i}(p_0) = \frac{\partial F}{\partial p_i}(x^*(p_0); p_0) + \sum_{m=1}^{M} \lambda_m \frac{\partial g_m}{\partial p_i}(x^*(p_0); p_0)$$

pro každé $i = 1, \dots, S$.

Poznámka

Vztah (6) vlastně znamená, že

$$D_p v(p_0) = D_p L(x^*(p_0); p_0),$$

kde $L(x; p) = F(x; p) + \lambda^\top G(x; p)$ je Lagrangeova funkce, tj. derivujeme ji vzhledem k proměnné p (současně proměnnou x považujeme za konstantní) a pak dosadíme za $x = x^*(p_0)$ a $p = p_0$.

Příklad

Ekonomická interpretace Lagrangeových multiplikátorů $\rightsquigarrow$ stínová cena.

Příklad (ekonomie)

Uvažte maximalizaci užitkové funkce

$$u(x) \to \max$$

s omezením $\langle p, x\rangle = m$, kde $x \in \mathbb{R}_+^N$ je vektor zboží/komodit, p je vektor cen a m je příjem/bohatství zákazníka.

Příklad (ekonomie)

Uvažte firmu užívající K jednotek kapitálu a L jednotek práce k produkci $F(K, L)$ jednotek zboží. Cena kapitálu je r a cena práce je w. Uvažujme problém minimalizace nákladů

$$\min C = rK + wL$$

za podmínky $F(K, L) = Q$, kde chceme najít hodnoty K a L, které minimalizují náklady na produkci Q jednotek. Nechť $C^* = C(r, w; Q)$ je funkce optimální hodnoty pro tento problém. Určeme $\frac{\partial C^*}{\partial r}$, $\frac{\partial C^*}{\partial w}$ a $\frac{\partial C^*}{\partial Q}$.

Konec.