Nelineárne optimalizačné modely a metódy

Téma prednášky č. 4

Podmienky optimálnosti Kuhna – Tuckera

(Časť 2)

Prof. Ing. Michal Fendek, CSc.

Katedra operačného výskumu a ekonometrie
Ekonomická univerzita
Dolnozemská 1
852 35 Bratislava

Pokračuj F 27

Podmienky optimálnosti

• Vidíme teda, že v niektorých prípadoch podmienky optimálnosti F.Johna zlyhávajú. Zaujímavejší z hľadiska riešenia úloh je teda prípad, keď $u_o \neq 0$. *Kuhn, H.W.* a *Tucker, A.W.* sformulovali analogické podmienky optimálnosti ako *F.John,* ale s dodatočnou vlastnosťou kladnej hodnoty multiplikátora u_o, a to $u_o = 1$. Ďalej formulované podmienky sa v odbornej literatúre obvykle uvádzajú ako *Kuhnove-Tuckerove podmienky optimálnosti.*

Podmienky optimálnosti Kuhna-Tuckera (Nutné podmienky optimálnosti *Kuhna-Tuckera*)

Nech X je neprázdna otvorená množina v R^n a $f:R^n \to R$, $g_i:R^n \to R$, $i=1,\dots,m$. Skúmajme úlohu

$$f(\mathbf{x}) \to \min$$

pri ohraničeniach $\quad g_i(\mathbf{x}) \le 0, \qquad i=1, \dots, m$

$$\mathbf{x} \in X$$

a jej prípustné riešenie $\mathbf{x}^{\mathbf{o}}$. Označme ako množinu I

$$I = \{\ i \mid g_i(\mathbf{x}) = 0\ \}$$

aktívnych ohraničení. Predpokladajme, že funkcie $f(\mathbf{x})$, $g_i(\mathbf{x})$, $i \in I$ sú diferencovateľné v $\mathbf{x}^{\mathbf{o}}$ a funkcie $g_i(\mathbf{x})$, $i \notin I$ sú spojité v $\mathbf{x}^{\mathbf{o}}$. Nech vektory $\nabla g_i(\mathbf{x})$, pre $i \in I$ sú navyše lineárne nezávislé. Ak $\mathbf{x}^{\mathbf{o}}$ je bodom lokálneho optima úlohy, potom existujú také čísla u_i, pre $i \in I$, pre ktoré platí

$$\boxed{\begin{gathered}\nabla f(\mathbf{x}^{\mathbf{0}}) + \sum_{i \in I} u_i \nabla g_i(\mathbf{x}^{\mathbf{0}}) = \mathbf{0} \\ u_i \ge 0, \qquad i \in I\end{gathered}}$$

Podmienky optimálnosti

• **Ilustrujme teraz podmienky optimálnosti *Kuhna-Tuckera* geometricky. Ľubovoľný vektor $y \in R^n$, ktorý sa dá vyjadriť ako nezáporná lineárna kombinácia gradientov funkcií aktívnych ohraničení v tvare**

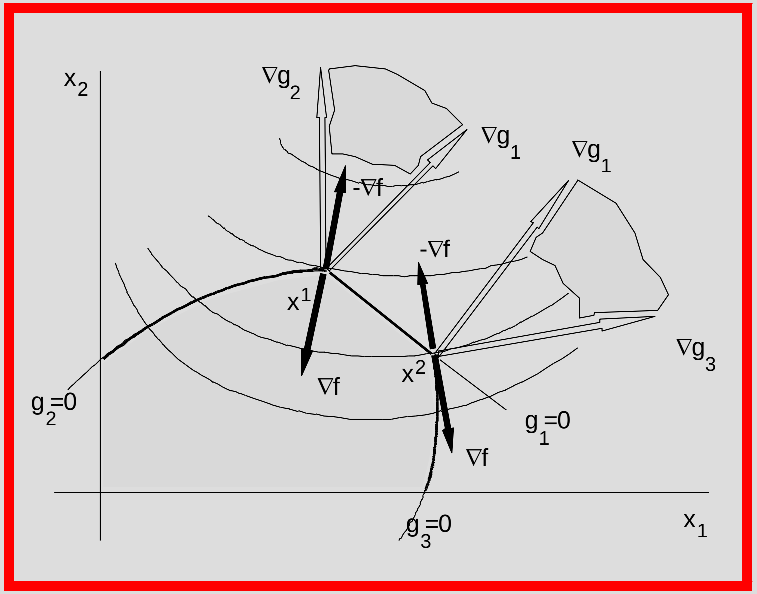
$$\mathbf{y} = \sum_{i=1}^{m} u_i \nabla g_i(\mathbf{x^0}) \qquad \text{pričom} \quad u_i \geq 0, \quad \text{pre } \forall i \in I$$

patrí do kónusu, resp. kužeľa, ktorý je vytvorený týmito gradientami v bode x^o. Z podmienok optimálnosti *Kuhna-Tuckera* vyplýva, že

$$-\nabla f(\mathbf{x}^0) = \sum_{i=1}^{m} u_i \nabla g_i(\mathbf{x^0}) \qquad \text{pričom} \quad u_i \geq 0, \quad \text{pre } \forall i \in I$$

t.j. vektor *- ∇f(x^o)* patrí do tohto kónusu. Skúmajme dva body x^1, x^2, ktoré sú prípustnými riešenia úlohy zobrazenej na obr. č.4.8.

Obr.č.4.8: Geometrická interpretácia podmienok optimálnosti Kuhna-Tuckera



Podmienky optimálnosti

Ak sú navyše funkcie g_i(x) diferencovateľné v bode x^o aj pre $i \notin I$, tak podmienky *Kuhna-Tuckera* možno zapísať v nasledovnej ekvivalentnej forme:

$$\nabla f(\mathbf{x}^0) + \sum_{i=1}^{m} u_i \nabla g_i(\mathbf{x}^0) = \mathbf{0}$$
$$u_i g_i(\mathbf{x}^0) = 0 \qquad \text{pre } i = 1,2,\ldots,m$$
$$u_i \geq 0, \qquad \text{pre } i = 1,2,\ldots,m$$

Analogicky, ako pri formulácii nutných podmienok F.Johna čísla u_i, $i=1,\ldots,m$ predstavujú *Lagrangeove multiplikátory* a rovnosti $u_i g_i(x^o)=0$ *podmienky komplementárnej rovnováhy*

Podmienky *Kuhna-Tuckera* môžeme formulovať aj vo vektorovom tvare nasledovne

$$u_0 \nabla f(\mathbf{x}^0) + \mathbf{G}\mathbf{u} = \mathbf{0}$$
$$\mathbf{u}^{\mathrm{T}} \mathbf{g}(\mathbf{x}^0) = \mathbf{0}$$
$$\mathbf{u} \geq \mathbf{0}$$

kde

- **g(x) - vektor funkcií sústavy ohraničení so zložkami g_i(x) pre i=1,...,m;**
- **G - matica typu (n,m) so stĺpcami $\nabla g_i(x^o)$;**
- **u - vektor *Lagrangeovych multiplikátorov*, $u \in R^m$**

• **Uvedené podmienky optimálnosti *Kohna-Tuckera* sú nutnými podmienkami, po zavedení určitých predpokladov o konvexnosti funkcií však predstavujú zároveň aj postačujúce podmienky.**Chyba! Záložka nie je definovaná.

Veta 4.8 (Postačujúce podmienky optimálnosti *Kuhna-Tuckera*)

Nech X je neprázdna otvorená množina v R^n a $f:R^n \to R$, $g_i:R^n \to R$, $i=1,...,m$. Skúmajme úlohu

$$\min \{ f(\mathbf{x}) \mid \mathbf{x} \in X, g_i(\mathbf{x}) \le 0, i=1, ..., m \}$$

a jej prípustné riešenie $\mathbf{x}^o$. Označme ako množinu $I = \{ i \mid g_i(\mathbf{x}^o) = 0 \}$ množinu aktívnych ohraničení v bode $\mathbf{x}^o$.

Predpokladajme, že funkcia $f(x)$ je pseudokonvexná v bode $\mathbf{x}^o$ a funkcie $g_i(x)$, $i \in I$ sú kvázikonvexné a diferencovateľné v $\mathbf{x}^o$. Ak sú v bode $\mathbf{x}^o$ splnené podmienky optimálnosti Kuhna-Tuckera, čiže existujú také čísla u_i, pre $i \in I$, pre ktoré platí

$$\nabla f(\mathbf{x}^0) + \sum_{i=1}^{m} u_i \nabla g_i(\mathbf{x}^0) = \mathbf{0}$$
$$u_i g_i(\mathbf{x}^0) = 0 \quad \text{pre } i = 1,2,\dots,m$$
$$u_i \ge 0, \quad \text{pre } i = 1,2,\dots,m$$

potom $\mathbf{x}^o$ je bodom *globálneho minima úlohy*.

Poznámka 4.2

Vychádzajúc z vlastností konvexných funkcií poznamenávame, že ak funkcie $f(x)$, $g_i(x)$ sú konvexné, potom sú podmienky vety 4.8 splnené.

Príklad č.4.8

Je daná úloha

$$f(x_1,x_2) = (x_1 - 9/4)^2 + (x_2 - 2)^2 \to \min$$

pri ohraničeniach

$$-x_1^2 + x_2 \geq 0$$
$$x_1 + x_2 \leq 6$$
$$x_1, x_2 \geq 0$$

a) Zapíšte podmienky optimálnosti Kuhna-Tuckera,

b) Presvedčte sa či sú podmienky optimálnosti *Kuhna-Tuckera* splnené pre body

$$x^o=(3/2,\ 9/4)^T,\ x^1=(0,\ 6)^T$$

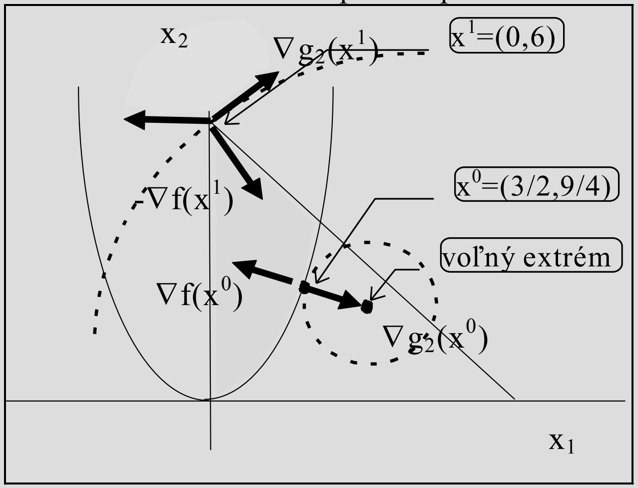
c) Geometricky interpretujte splnenie podmienok *Kuhna-Tuckera* v obidvoch bodoch z úlohy (b),

d) Ukážte, že bod x^o je bodom globálneho optima úlohy (podmienky *Kuhna-Tuckera* sa realizujú ako postačujúce)

Riešenie

Geometrická interpretácia úlohy je uvedená na obr. č.4.9.

Obr.č.4.9: Geometrická interpretácia príkladu č.4.8



$$\nabla f(\mathbf{x})=(2(x_1-9/4),2(x_2-2))^T,$$

$$\nabla g_1(\mathbf{x})=(2x_1,1)^T,$$

$$\nabla g_2(\mathbf{x})=(1,1)^T,$$

$$\nabla g_3(\mathbf{x})=(-1,0)^T,$$

$$\nabla g_4(\mathbf{x})=(0,-1)^T$$

a formulujme nutné podmienky optimálnosti Kuhna-Tuckera

$$\begin{pmatrix} 2(x_1-9/4) \\ 2(x_2-2) \end{pmatrix} + u_1 \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix} + u_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + u_3 \begin{pmatrix} -1 \\ 0 \end{pmatrix} + u_4 \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$u_1\,(x_1^2 - x_2) = 0$$

$$u_2\,(x_1 + x_2 - 6) = 0$$

$$u_3\,(-x_1) = 0$$

$$u_4\,(-x_2) = 0$$

$$u_1,\ u_2,\ u_3,\ u_4 \geq 0$$

b) Preverme platnosť podmienok optimálnosti pre bod $\mathbf{x}^o=(3/2,\ 9/4)^T$, ktorý je, ako vidíme na obr.č.4.9, optimálnym riešením úlohy. Overme, ktoré z ohraničení sú aktívne. Dostávame:

$$g_1(\mathbf{x}^o) = 0,\ g_2(\mathbf{x}^o) < 0,\ g_3(\mathbf{x}^o) < 0,\ g_4(\mathbf{x}^o) < 0$$

Iba prvé ohraničenie je aktívne, to znamená $I=\{1\}$. Potom ale na základe komplementárnej podmienky platí, že *Lagrangeove multiplikátory* zodpovedajúce neaktívnym ohraničeniam sú nulové, a teda platí

$$u_2 = u_3 = u_4 = 0.$$

Vypočítajme hodnoty gradientov účelovej funkcie a funkcií ohraničení

$$\nabla f(\mathbf{x}^o)=(-3/2),1/2)^T,\ \ \nabla g_1(\mathbf{x}^o)=(3,-1)^T,\ \ \nabla g_2(\mathbf{x}^o)=(1,1)^T,\ \ \nabla g_3(\mathbf{x}^o)=(-1,0)^T,\ \ \nabla g_4(\mathbf{x}^o)=(0,-1)^T$$

Preverme platnosť nutných podmienok optimálnosti (na základe podmienky komplementárnosti v nich bude figurovať iba multiplikátor u_1). Dostávame:

$$\begin{pmatrix} -3/2 \\ 1/2 \end{pmatrix} + u_1 \begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$u_1 \geq 0$$

$$\Rightarrow\ u_1 = 1/2 > 0$$

Hodnoty Lagrangeových multiplikátorov sú teda nasledovné

$$(u_1, u_2, u_3, u_4) = (1/2, 0, 0, 0) \geq \mathbf{0}$$

a vidíme, že multiplikátory vyhovujú nutným podmienkam optimálnosti Kuhna-Tuckera.

Overme ďalej platnosť Kuhnových-Tuckerových podmienok v bode $\mathbf{x}^1=(0,\ 6)^T$, ktorý evidentne nie je optimálnym riešením úlohy. ľahko sa presvedčíme o tom, že množina indexov aktívnych ohraničení je $I = \{2,3\}$, a teda $u_1 = u_4 = 0$. Hodnoty gradientov účelovej funkcie a funkcií aktívnych ohraničení sú nasledovné

$$\nabla f(\mathbf{x}^1)=(-9/2),8)^T,\ \nabla g_2(\mathbf{x}^1)=(1,1)^T,\ \nabla g_3(\mathbf{x}^1)=(-1,0)^T$$

Hodnoty multiplikátorov u_2, u_3 získame riešením sústavy rovníc zodpovedajúcich prvej podmienke optimálnosti Kuhna-Tuckera:

$$\begin{pmatrix} -9/2 \\ 8 \end{pmatrix} + u_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + u_3 \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow$$

$$u_2 = -25/2,\ u_3 = -8$$

Hodnoty Lagrangeovych multiplikátorov sú teda nasledovné

$$(u_1, u_2, u_3, u_4) = (0, -8, -25/2, 0) \leq \mathbf{0}$$

a vidíme, že bod $\mathbf{x}^1$ **nevyhovuje** nutným podmienkam optimálnosti Kuhna-Tuckera.

d) Ukážme, že pre bod $\mathbf{x}^o$ sú nutné podmienky optimálnosti Kuhna-Tuckera zároveň aj postačujúcimi. Vypočítajme Hessovu maticu pre účelovú funkciu $f(\mathbf{x})$ a funkciu aktívneho ohraničenia $g_1(\mathbf{x})$:

$$\mathbf{H}_\mathbf{f}(\mathbf{x}) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \qquad \mathbf{H}_{g1}(\mathbf{x}) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

Vidíme, že Hessova matica funkcie $f(\mathbf{x})$ je kladne definitná nezávisle na hodnotách premenných x_1, x_2, preto funkcia $f(\mathbf{x})$ je rýdzokonvexná, a teda aj pseudokonvexná na celom obore definície. Hessova matica funkcie $g_1(\mathbf{x})$ je kladne semidefinitná nezávisle na hodnotách premenných x_1, x_2, preto funkcia $g_1(\mathbf{x})$ je konvexná, a teda aj rýdzo kvázikonvexná na celom obore definície. Bod $\mathbf{x}^o$ je teda bodom globálneho optima úlohy.

Úlohy so zmiešanými typmi ohraničení

Budeme skúmať nasledovnú úlohu nelineárneho programovania

$$f(\mathrm{x}) \to \min$$

pri ohraničeniach

$$g_i(\mathrm{x}) \le 0, \qquad i=1, \ldots, m \qquad (4.4)$$

$$h_k(\mathrm{x}) = 0, \qquad k=1, \ldots, l$$

$$\mathrm{x} \in X$$

kde m **je počet ohraničení úlohy typu** ***menší alebo rovný,***

l **je počet ohraničení úlohy typu** ***rovný,***

$f(\mathrm{x}), g_i(\mathrm{x}), h_k(\mathrm{x}): R^n \to R$ **pre** $i=1,\ldots,m,\ k=1,\ldots,l$ **a**

X **je neprázdna otvorená množina v** R^n**.**

Úlohy so zmiešanými typmi ohraničení

Budeme skúmať nasledovnú úlohu nelineárneho programovania

$$f(\mathrm{x}) \to \min$$

pri ohraničeniach

$$g_i(\mathrm{x}) \le 0, \qquad i=1, ..., m \quad (4.4)$$

$$h_k(\mathrm{x}) = 0, \qquad k=1, ..., l$$

$$\mathrm{x} \in X$$

kde m je počet ohraničení úlohy typu *menší alebo rovný*,

l je počet ohraničení úlohy typu *rovný*,

$f(\mathrm{x}), g_i(\mathrm{x}), h_k(\mathrm{x}): R^n \to R$ pre $i=1,...,m$, $k=1,...,l$ a

X je neprázdna otvorená množina v R^n.

Geometrická interpretácia podmienok optimálnosti

Zovšeobecnením podmienky optimálnosti formulovanej pre úlohy s ohraničeniami v tvare nerovníc ako

$$F^o \cap G^o = \varnothing$$

pre túto triedu úloh.

Ak bod x^o je bodom lokálneho minima úlohy (4.4), potom platí

$$F^o \cap G^o \cap H^o = \varnothing$$

kde

$F^o = \{ d \mid \nabla f(x^o)^T d < 0 \}$

$G^o = \{ d \mid \nabla g_i(x^o)^T d < 0,\ i \in I \}$

$H^o = \{ d \mid \nabla h_k(x^o)^T d = 0,\ k = 1, \ldots, l \}$.

Príklad č.4.9

Preskúmajme úlohu

$$f(x_1,x_2) = x_1^2 + x_2^2 \rightarrow min$$

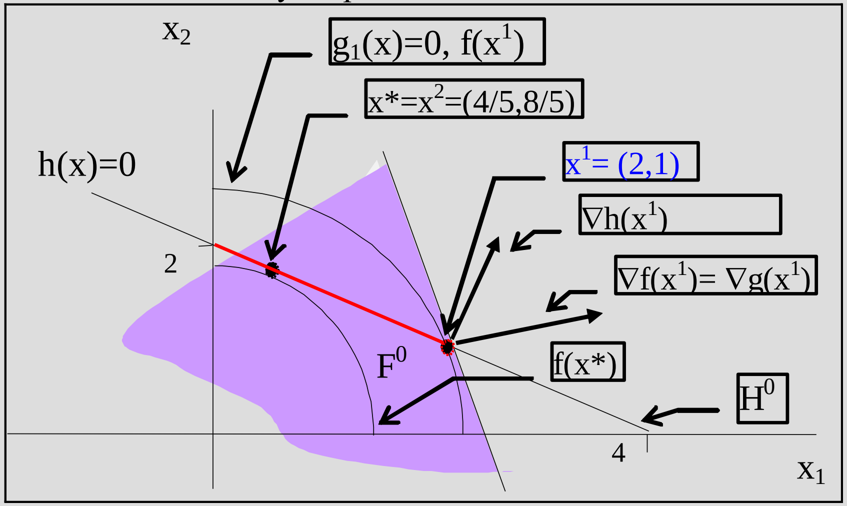
pri ohraničeniach

$$g_1(x) = x_1^2 + x_2^2 - 5 \leq 0$$
$$g_2(x) = -x_1 \leq 0$$
$$g_3(x) = -x_2 \leq 0$$
$$h(x) = x_1 + 2x_2 - 4 = 0$$

Úloha:

Preverte platnosť geometrických podmienok optimálnosti v bodoch $x^1=(2,1)^T$, $x^2=(4/5,8/5)^T$.

0br.č.4.12: Podmienky nesplnené: $F^o \cap G^o \cap H^o \neq \emptyset$



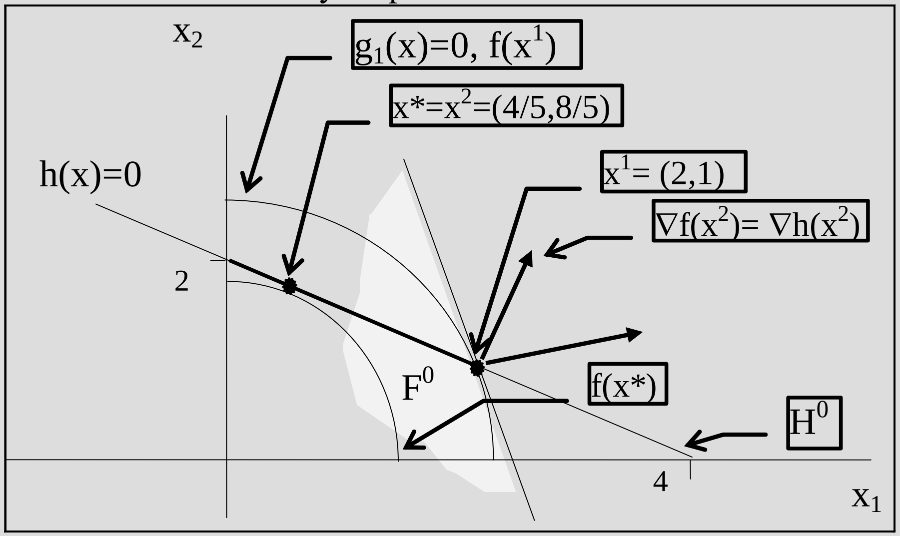
Podmienky optimálnosti

a) Preskúmajme bod $\mathbf{x}^1=(2,1)^T$. V tomto bode sa ako aktívne realizuje ohraničene $g_1(\mathbf{x})$, takže I={1}. Vypočítajme gradient učelovej funkcie a funkcie aktívneho ohraničenia a funkcie $h(\mathbf{x})$ v bode $\mathbf{x}^1$

$$\nabla f(\mathbf{x}^1) = (4,2)^T,\ \nabla g_1(\mathbf{x}^1) = (4,2)^T,\ \nabla h(\mathbf{x}^1) = (1,2)^T$$

Na obr. je zobrazená množina progresívnych smerov F^o a množiny prípustných smerov G^o a H^o. Vidíme, že prienik množín prípustných a množiny progresívnych smerov nie je prázdna množina, takže v bod $\mathbf{x}^1=(2,1)$ sa nesplňajú geometrické podmienky optimálnosti a bod preto nie je optimálnym riešením úlohy.

Podmienky nesplnené: $F^o \cap G^o \cap H^o \neq \emptyset$

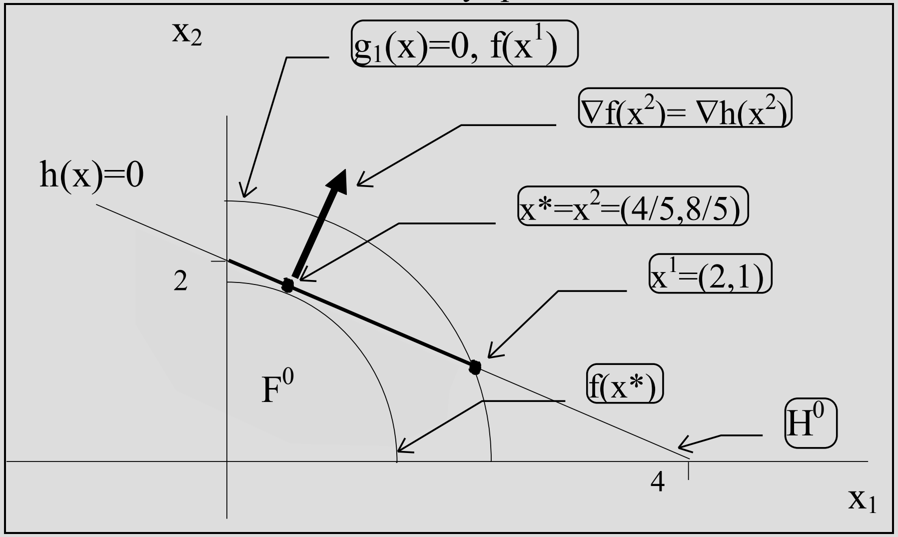


b) Preskúmajme bod $\mathbf{x}^2=(4/5,8/5)^T$. Pre tento bod neexistuje aktívne ohraničenie úlohy, takže I={∅}. Vypočítajme gradienty funkcií *f*(**x**) a *h*(**x**) bode $\mathbf{x}^2$:

$$\nabla f(\mathbf{x}^2) = (8/5,16/5)^T, \qquad \nabla h(\mathbf{x}^2) = (1,2)^T$$

Na obr. sú zobrazené množiny F^o, H^o a vidíme, že pre ich prienik platí $F^o \cap H^o$=Ć, čím sú pre bod x^2 splnené podmienky vety 4.9. Tieto podmienky sú však len nutné, a preto splnenie podmienky $F^o \cap H^o = \emptyset$ vo všeobecnosti negarantuje, že bod $\mathbf{x}^*$ je bodom optima, je iba bodom, ktorý môže ale aj nemusí byť optimálnym riešením úlohy.

0br.č.4.11: Podmienky splnené: $F^o \cap G^o \cap H^o = \emptyset$



<u>Veta 4.11</u> (Nutné podmienky optimálnosti *Kuhna-Tuckera*)

Nech X je neprázdna otvorená množina v R^n a $f:R^n \to R$, $g_i:R^n \to R$, $i=1,\dots,m$ a $h_k:R^n \to R$. Skúmajme úlohu

$$f(x) \to min$$

pri ohraničeniach

$$g_i(x) \le 0, \quad i=1, \dots, m$$
$$h_k(x) = 0, \quad k=1, \dots, l$$
$$x \in X$$

a jej prípustné riešenie x^o. Označme množinu aktívnych ohraničení ako množinu

$$I = \{ i \, / \, g_i(x) = 0 \}$$

Predpokladajme, že v bode x^o sú funkcie $g_i(x)$, $i \notin I$ spojité, funkcie $f(x)$, $g_i(x)$, $i \in I$ diferencovateľné a funkcie $h_k(x)$ spojito diferencovateľné. Okrem toho nech vektory $\nabla g_i(x^o)$ pre $i \notin I$ a vektory $\nabla h_k(x^o)$ pre $k=1,\dots,l$ sú lineárne nezávislé. Ak x^o je bodom lokálneho optima úlohy, potom existujú také čísla u_i, $i \in I$, a v_k pre $k=1,\dots,l$, pre ktoré platí

$$\nabla f(\mathbf{x^0}) + \sum_{i \in I} u_i \nabla g_i(\mathbf{x^0}) + \sum_{k=1}^{l} v_k \nabla h_k(\mathbf{x^0}) = \mathbf{0}$$
$$u_i \ge 0, \qquad \mathrm{i} \in \mathrm{I}$$

Podmienky optimálnosti

Ak sú navyše funkcie g_i diferencovateľné v bode $\mathbf{x}^0$ aj pre $i \notin I$, tak podmienky *Kuhna-Tuckera* možno zapísať v nasledovnej ekvivalentnej forme:

$$\boxed{\begin{gathered}\nabla f(\mathbf{x}^0) + \sum_{i=1}^{m} u_i \nabla g_i(\mathbf{x}^0) + \sum_{k=1}^{l} h_k \nabla h_k(\mathbf{x}^0) = \mathbf{0} \\ u_i g_i(\mathbf{x}^0) = 0 \qquad \text{pre } i = 1,2,\dots,m \\ u_i \geq 0, \qquad \text{pre } i = 1,2,\dots,m\end{gathered}}$$

Podmienky *Kuhna-Tuckera* môžeme formulovať aj vo vektorovom tvare následovně

$$\begin{gathered}\nabla f(\mathbf{x}^0) + \mathbf{G}\mathbf{u} + \mathbf{H}\mathbf{v} = \mathbf{0} \\ \mathbf{u}^{\mathrm{T}}\, \mathbf{g}(\mathbf{x}^0) = \mathbf{0} \\ \mathbf{u} \geq \mathbf{0}\end{gathered}$$

kde

- **g**(**x**) - vektorová funkcií sústavy ohraničení so zložkami $g_i(\boldsymbol{x})$ pre i=1,...,m;
- **G** - matica typu (n,m) so stĺpcami $\nabla g_i(\boldsymbol{x})$;
- **u** - vektor *Lagrangeovych multiplikátorov*, $\mathbf{u} \in \mathrm{R}^m$;
- **H** - matica typu (n,l) so stĺpcami $\nabla h_k(\boldsymbol{x})$;
- **v** - vektor *Lagrangeovych multiplikátorov*, $\mathbf{v} \in \mathrm{R}^L$

Podmienky optimálnosti

Veta 4.12 (Postačujúce podmienky optimálnosti *Kuhna-Tuckera*)

Nech X je neprázdna otvorená množina v R^n a $f(x):R^n \to R$, $g_i(x):R^n \to R$, $i=1,\dots,m$ a $h_k(x):R^n \to R$, $k=1,\dots,l$. Skúmajme úlohu

$$f(x) \to min$$

pri ohraničeniach

$$g_i(x) \le 0, \quad i=1, \dots, m$$
$$h_k(x) = 0 \quad k=1, \dots, l$$
$$x \in X$$

a jej prípustné riešenie x^o.
Označme ako množinu

$$I = \{ i \mid g_i(x) = 0 \}$$

Predpokladajme, že v bode x^o sú splnené podmienky optimálnosti *Kuhna-Tuckera*, t.j. existujú také čísla u_i, $i \in I$, a čísla v_k, $k=1,\dots,l$, pre ktoré platí

$$\nabla f(\mathbf{x}^0) + \sum_{i \in I} u_i \nabla g_i(\mathbf{x}^0) + \sum_{k=1}^{l} v_k \nabla h_k(\mathbf{x}^0) = \mathbf{0}$$

Zaveďme označenie $J=\{k \,/\, v_k > 0\}$, $K=\{k \,/\, v_k < 0\}$ a predpokladajme, že v bode x^o je funkcia $f(x)$ pseudokonvexná, funkcie $g_i(x)$ pre $i \in I$ sú kvázikonvexné a funkcie $h_k(x)$ pre $k \in J$ sú kvázikonvexné a pre $k \in K$ kvázikonkávne. Potom bod x^o je globálnym optimálnym riešením úlohy.

Príklad č.4.13

Skúmajme nasledovnú ulohu nelineárneho programovania

$$f(x) = -10/3x_1 - 2x_2 + x_3^2 \rightarrow min$$

pri ohraničeniach

$$g_1(x) = x_1 + x_2 + x_3 \leq 0$$

$$h_1(x) = -x_1 - 2x_2 + x_3^2 = 0$$

Úlohy:

a) formulujte nutné podmienbky optimálnosti *Kuhna-Tuckera,*

b) na základe podmienok optimálnosti určte bod x^o,

c) preverte, či bod x^o je globálnym optimálnym riešením úlohy.

Riešenie

a) Gradienty funkcií úlohy sú nasledovné

$$\nabla f(\mathbf{x}) = (-10/3, -2, 2x_3)^T, \ \nabla g_1(\mathbf{x}) = (1,1,1)^T, \ \nabla h_1(\mathbf{x}) = (-1, -2, 2x_3)^T$$

$$\begin{pmatrix} -10/3 \\ -2 \\ 2x_3 \end{pmatrix} + u_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + v_1 \begin{pmatrix} -1 \\ -2 \\ 2x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$u_1 (x_1 + x_2 + x_3) = 0$$

$$u_1 \geq 0$$

b) Na základe formulovaných podmienok optimálnosti riešime sústavu lineárnych rovníc

$$-10/3 + u_1 - v_1 = 0$$

$$-2 + u_1 - 2v_1 = 0$$

$$2x_3 + u_1 + 2v_1 x_3 = 0$$

Riešením sústavy je vektor

$$(u_1, v_1, x_3) = (14/3, 4/3, -1)^T$$

Keďže $u_1 \neq 0$, tak podľa komplementárnej podmienky *Kuhna-Tuckera* platí

$$g_1(x)=0.$$

Po dosadení hodnoty premennej $x_3=-1$ *do obidvoch ohraničení úlohy* dostávame sústavu dvoch rovníc o dvoch neznámych,

$g_1(\mathbf{x}) = x_1 + x_2 + x_3 = 0$	$\mathbf{x_1 + x_2 = 1}$
$h_1(\mathbf{x}) = -x_1 - 2x_2 + x_3^2 = 0$	$\mathbf{-x_1 - 2x_2 = -1}$

ktorej riešením získame hodnoty zostávajúcich premenných $x_2=0$, $x_1=1$. Bodom, ktorý vyhovuje podmienkam optimálnosti *Kuhna-Tuckera* je bod

$$\mathbf{x^o = (1, 0, -1)^T}$$

c) Na základe postačujúcich podmienok optimálnosti preveríme, či bod x^o je globálnym riešením úlohy.

Preskúmajme dodatočné požiadavky konvexnosti kladené na funkcie úlohy:

- pseudokonvexnosť funkcie *f*(x):

$$f(x) = 10/3\, x_1 - 2x_2 + x_3^2 \;\Rightarrow\; H(x) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Hessova matica je kladne semidefinitná, preto funkcia *f(x)* je konvexná, a teda aj pseudokonvexná;

- kvázikonvexnosť funkcie $g_1(x)$, keďže $I=\{1\}$:

$$g_1(x) = x_1 + x_2 + x_3$$

Vidíme, že funkcia je lineárna, preto je konvexná a zároveň aj kvázikonvexná;

- **kvázikonvkávnosť funkcie h_1,**

keďže $v_1 > 0$, a teda $K=\{1\}$:

$$h_1(\mathrm{x}) = -x_1 + 2x_2 + x_3^2 \Rightarrow \mathrm{H}(\mathrm{x}) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Hessova matica je kladne semidefinitná, preto funkcia $h_1(x)$ je konvexná, a preto bod x^o je globálnym optimálnym riešením úlohy.

V bode x^o je hodnota účelovej funkcie

$$f(x^o) = -7/3$$

4.3 Alternatívne formy podmienok Kuhna-Tuckera pre rôzne typy úloh

Doposiaľ sme sa zaoberali skúmaním podmienok optimálnosti v najvšeobecnejšom prípade pre úlohy so zmiešanými typmi ohraničení v tvare

$$\min \{f(\mathbf{x}) \mid g_i(x) \le 0,\ i=1, \ldots, m;\ h_k(x) = 0,\ k=1, \ldots, l;\ \mathbf{x} \in X\} \qquad (4.5)$$

pričom podmienky optimálnosti Kuhna-Tuckera v súlade so znením vety 4.10 predpokladali pre prípustné riešenie $\mathbf{x}^0$ existenciu multiplikátorov u_i, $i=1,\ldots,m$, v_k, $k=1,\ldots,l$, pre ktoré platí

$$\boxed{\begin{array}{c} \nabla f(\mathbf{x}^0) + \sum_{i=1}^{m} u_i \nabla g_i(\mathbf{x}^0) + \sum_{k=1}^{l} h_k \nabla h_k(\mathbf{x}^0) = \mathbf{0} \\ u_i g_i(\mathbf{x}^0) = 0 \qquad \text{pre } i = 1,2,\ldots,m \\ u_i \ge 0, \qquad \text{pre } i = 1,2,\ldots,m \end{array}}$$

- **I keď vo všeobecnosti možno ľubovoľnú úlohu upraviť na horeuvedený tvar (4.5), je predsa len užitočné ukázať formuláciu Kuhnovych-Tuckerovych podmienok pre rôzne typy úloh.**

- **Veľmi často sa požiadavka prípustnosti riešenia úlohy stáva priamo súčasťou formulácie podmienok optimálnosti, ktoré majú potom pre úlohu (4.5) nasledovný tvar**

$$\boxed{\begin{array}{c}\nabla f(\mathbf{x})+\sum_{i=1}^{m}u_i\nabla g_i(\mathbf{x})+\sum_{k=1}^{l}h_k\nabla h_k(\mathbf{x})=\mathbf{0} \quad \mathrm{g}_i(\mathbf{x})\le 0 \;\; i=1,\ldots,m \quad h_{ki}(\mathbf{x})=0 \;\; i=1,\ldots,m \\ u_i g_i(\mathbf{x})=0 \quad i=1,\ldots,m \\ u_i\ge 0, \quad i=1,\ldots,m\end{array}}$$

Podmienky optimálnosti

Ak formulujeme pre úlohu (4.5) ***Lagrangeovu funkciu*** **v tvare**

$$L(\mathbf{x},\mathbf{u},\mathbf{v}) = f(\mathbf{x}) + \sum_{i=1}^{m} u_i g_i(\mathbf{x}) + \sum_{k=1}^{l} v_k h_k(\mathbf{x})$$

potom Kuhnove-Tuckerove podmienky môžeme na základe Lagrangeovej funkcie formulovať nasledovne

$$\frac{\partial L(x,\ u,\ v)}{\partial x_j} = 0,\quad j = 1,\dots,n \qquad \frac{\partial L(x,\ u,\ v)}{\partial u_i} \le 0,\quad i = 1,\dots,m \qquad \frac{\partial L(x,\ u,\ v)}{\partial v_k} = 0,\quad k = 1,\dots,l$$

$$u_i \frac{\partial L(x,u,v)}{\partial u_i} = 0,\quad i=1,\dots,m$$

$$u_i \ge 0,\quad i=1,\dots,m$$

kde

$$\frac{\partial L(x,u,v)}{\partial x_j} = \frac{\partial f(x)}{\partial x_j} + \sum_{i=1}^{m} u_i \frac{\partial g_i(x)}{\partial u_i} + \sum_{k=1}^{l} v_k \frac{\partial h_k(x)}{\partial v_k}$$

$$\frac{\partial L(x,u,v)}{\partial u_i} = g_i(x),\quad i=1,\dots,m$$

$$\frac{\partial L(x,u,v)}{\partial v_k} = h_k(x),\quad k=1,\dots,l$$

resp. s využitím vektorovej formulácie nasledovne

$$\frac{\partial \mathbf{L}}{\partial \mathbf{x}} = 0 \qquad \frac{\partial \mathbf{L}}{\partial \mathrm{u}} \le 0 \qquad \frac{\partial \mathbf{L}}{\partial \mathbf{v}} = 0$$

$$\mathbf{u}\frac{\partial \mathbf{L}}{\partial \mathrm{u}} = 0$$

$$\mathbf{u} \ge \mathbf{0}$$

kde

$$\frac{\partial L}{\partial x} = \left(\frac{\partial L(x,\ u,\ v)}{\partial x_1}\ ,\ \dots\ ,\frac{\partial L(x,\ u,\ v)}{\partial x_n}\right)^T;$$

$$\frac{\partial L}{\partial u} = \left(\frac{\partial L(x,\ u,\ v)}{\partial u_1}\ ,\ \dots\ ,\ \frac{\partial L(x,\ u,\ v)}{\partial u_m)}\right)^T$$

$$\frac{\partial L}{\partial v} = \left(\frac{\partial L(x,\ u,\ v)}{\partial v_1}\ ,\ \dots\ ,\ \frac{\partial L(x\ ,u,\ v)}{\partial v_l}\right)^T$$

Poznámka 4.3

Zvláštnu pozornosť, najmä pri analýze ekonomických úloh, je užitočné venovať vlastnostiam úloh, v ktorých sú premenné definované ako nezáporné, čo zodpovedá charakteru väčšiny ekonomických problémov.

Skúmajme úlohu na viazaný extrém, ktorej jedinými ohraničeniami sú podmienky nezápornosti premenných.

$$\min \{f(\mathrm{x}) \;/\; x_j \geq 0,\ j=1,\ldots,n;\ \mathrm{x} \in X\} \qquad (4.6)$$

Lagrangeova funkcia má pre ohraničenia $g_i(\mathrm{x})=-x_j \leq 0$ pre $i=1,\ldots,n$ a $i=j$ tvar

$$L(\mathrm{x}, u) = f(\mathrm{x}) + \sum_{i=1}^{n} u_i(-x_i)$$

Podmienky optimálnosti KuhnaTuckera majú nasledovný tvar

$$\frac{\partial L(x,u)}{\partial x_j} = \frac{\partial f(x)}{\partial x_j} + \sum_{i=1}^{n} u_i \frac{\partial(-x_i)}{\partial x_j} =$$

$$= \frac{\partial f(x)}{\partial x_j} + \sum_{i=1}^{n} u_i(-\delta_{ij}) = 0 \qquad j = 1,\dots,n$$

$$\frac{\partial L(x,u)}{\partial u_i} = -x_i \leq 0, i = 1,\dots,n$$

$$u_i \frac{\partial L(x,u)}{\partial u_i} = u_i x_i = 0, i = 1,\dots,n$$

$$u_i \geq 0, i = 1,\dots,n$$

kde

$$\delta_{ij} = \begin{cases} 1, & ak\ i = j \\ 0, & ak\ i \neq j \end{cases}$$

potom ale

$$\frac{\partial f(x)}{\partial x_j} - u_j = 0, \ resp. \ \frac{\partial f(x)}{\partial x_j} = u_j, \ j = 1,\dots,n$$

Podmienky optimálnosti

Vidíme, že na základe substitúcie na ľavej strane Kuhnovych-Tuckerovych podmienok môžeme v podmienkach eliminovať Lagrangeove multiplikátory pre ohraničenia nezápornosti premenných.

Ak ďalej nazveme *zovšeobecnenou Lagrangeovou funkciou* funkciu neobsahujúcu výrazy zodpovedajúce ohraničeniam nezápornosti premenných a Lagrangeovu funkciu teda zapíšeme v tvare

$$L(\mathbf{x}) = f(\mathbf{x}) \qquad \frac{\partial L(\mathbf{x})}{\partial x_i} = u_i$$

tak podmienky optimálnosti Kuhna-Tuckera pre úlohu (4.6) môžeme formulovať následovně

$$\begin{aligned} x_i &\geq 0, \quad i = 1,\ldots,n \\ x_i \frac{\partial L(\mathrm{x})}{\partial x_i} &= 0, \quad i = 1,\ldots,n \\ \frac{\partial L(\mathrm{x})}{\partial x_i} &\geq 0, \quad i = 1,\ldots,n \end{aligned}$$

a na základe vektorovej formulácie nasledovne

$$\begin{aligned} \frac{\partial \mathrm{L}}{\partial \mathrm{x}} &\geq 0 \\ \mathrm{x}\frac{\partial \mathrm{L}}{\partial \mathrm{x}} &= 0 \\ \mathrm{x} &\geq 0 \end{aligned}$$

Táto modifikácia podmienok optimálnosti *Kuhna-Tuckera* umožní zjednodušenú formuláciu podmienok pre rôzne typy úloh.

Podmienky optimálnosti

Na základe predchádzajúcich záverov sformulujme teraz podmienky optimálnosti pre nasledovnú úlohu nelineárneho programovania so zmiešanými typmi ohraničení

$$f(\mathbf{x},\mathbf{y}) \rightarrow min$$

pri ohraničeniach

$$g_i(\mathbf{x},\mathbf{y}) \leq 0,\ i=1,...,m \qquad (4.7)$$
$$h_k(\mathbf{x},\mathbf{y}) = 0,\ k=1,...,l$$
$$\mathbf{x} \geq 0$$

kde x je vektor nezáporných premenných a y je vektor voľných premenných.

Zapíšme zovšeobecnenú Lagrangeovu funkciu pre úlohu (4.7), t.j. funkciu bez explicitného zohľadnenia podmienok nezápornosti premenných:

$$L(\mathbf{x},\mathbf{y},\mathbf{u},\mathbf{v}) = f(\mathbf{x},\mathbf{y}) + \sum_{i=1}^{m} u_i g_i(\mathbf{x},\mathbf{y}) + \sum_{k=1}^{l} v_k h_k(\mathbf{x},\mathbf{y})$$

a *Kuhnove-Tuckerove* podmienky optimálnosti v tvare vektorovej formulácie

$$\frac{\partial \mathbf{L}}{\partial \mathbf{x}} \geq \mathbf{0} \qquad \frac{\partial \mathbf{L}}{\partial \mathbf{y}} = 0 \qquad \frac{\partial \mathbf{L}}{\partial \mathrm{u}} \leq 0 \qquad \frac{\partial \mathbf{L}}{\partial \mathbf{v}} = 0$$

$$\mathbf{x}\frac{\partial \mathbf{L}}{\partial \mathbf{x}} = \mathbf{0} \qquad \mathbf{u}\frac{\partial \mathbf{L}}{\partial \mathrm{u}} = 0$$

$$\mathbf{x} \geq \mathbf{0} \qquad \mathbf{u} \geq \mathbf{0}$$

Podmienky optimálnosti

Na základe formálnej analýzy *Kuhnovych-Tuckerovych* podmienok úlohy (4.7) môžeme formulovať nasledovné pravidlá pre konštrukciu *Kuhnových-Tuckerových* podmienok optimálnosti.

Nech x je vektor nezáporných premenných, y je vektor voľných premenných a *L* je zovšeobecnená Lagrangeova funkcia. Potom parciálne derivácie funkcie *L* sú:

a) podľa *nezáporných* premenných *nezáporné* a platí podmienka komplementárnej rovnováhy $x(\partial L/\partial x)=0$;

b) podľa *voľných* premenných *nulové*;

c) podľa *nezáporných* Lagrangeovych multiplikátorov zodpovedajúcich nerovniciam *nekladné* a platí podmienka komplementárnej rovnováhy $u(\partial L/\partial u)=0$;

d) podľa *voľných* Lagrangeovych multiplikátorov zodpovedajúcich rovniciam *nulové*.

Podmienky optimálnosti

Na záver tejto časti ešte odvodíme *Kuhnove-Tuckerove* podmienky optimálnosti pre dvojicu *nesymetrických duálnych úloh lineárneho programovania*.

PU: $f(\mathbf{x}) = \mathbf{c}^T\mathbf{x} \to \min$

$$\mathbf{A}\mathbf{x} \geq \mathbf{b}$$

DU: $g(\mathbf{u}) = \mathbf{u}^T\mathbf{b} \to \max$

$$\mathbf{u}^T\mathbf{A} = \mathbf{c}^T$$

$$\mathbf{u} \geq \mathbf{0}$$

$\mathbf{x} \geq \mathbf{0}$

kde

A - matica sústavy ohraničení o rozmere *(m,n);*
x - vektor rozhodovacích premenných, $\mathbf{x} \in \mathbf{R}^n$;
c - vektor koeficientov účelovej funkcie, $\mathbf{c} \in \mathbf{R}^n$;
b - vektor koeficientov pravej strany, $\mathbf{b} \in \mathbf{R}^m$;
u - vektor duálnych premenných, $\mathbf{u} \in \mathbf{R}^m$;
$f(\mathbf{x})$ - účelová funkcia primárnej úlohy, $f: R^n \to R$;
$g(\mathbf{u})$ - účelová funkcia duálnej úlohy, $g: R^m \to R$.
Lagrangeova funkcia úlohy pre upravenú sústavu ohraničení $\mathbf{g}(\mathbf{x}) = \mathbf{b} - \mathbf{A}\mathbf{x} \leq \mathbf{0}$ je následovní

$$L(\mathbf{x},\mathbf{u}) = \mathbf{c}^T\mathbf{x} + \mathbf{u}^T(\mathbf{b} - \mathbf{A}\mathbf{x})$$

Podmienky optimálnosti *Kuhna-Tuckera* na základe prv uvedených pravidiel zapíšeme nasledovne

$$\frac{\partial L}{\partial \mathbf{x}} = \mathbf{c}^T - \mathbf{u}^T\mathbf{A} = \mathbf{0}$$

$$\frac{\partial L}{\partial \mathbf{u}} = \mathbf{b} - \mathbf{A}\mathbf{x} \leq \mathbf{0}$$

$$\mathbf{u}\frac{\partial L}{\partial \mathbf{u}} = \mathbf{u}^T(\mathbf{b} - \mathbf{A}\mathbf{x}) = \mathbf{0}$$

$$\mathbf{u} \geq \mathbf{0}$$

Ukážme si ešte určitú analógiu medzi podmienkami optimálnosti *Kuhna-Tuckera* a vlastnosťami riešení dvojice duálnych úloh lineárneho programovania.

Podmienky optimálnosti

Pripomeňme si známy poznatok z tórie riešenia úloh lineárneho programovania, a to, že vektory x, u sú optimálnymi riešeniami dvojice duálnych úloh lineárneho programovania, vtedy a len vtedy, ak sú prípustnými riešeniami tejto dvojice úloh a vyhovujú podmienkam vety o rovnováhe (resp. vety o komplementárnosti doplnkových premenných).

Ukážme, že podmienky optimálnosti *Kuhna-Tuckera* pre skúmanú dvojicu úloh sú v súlade s touto vlastnosťou a garantujú optimálnosť riešení dvojice duálnych úloh:

- **platnosť podmienky** $\frac{\partial L}{\partial x} = 0$ **garantuje prípustnosť riešenia duálnej úlohy;**
- **platnosť podmienok** $\frac{\partial L}{\partial u} \leq 0, u \geq 0$ **garantuje prípustnosť riešenia primárnej úlohy;**
- **podmienka** $u \frac{\partial L}{\partial u} = 0$ **reperezentuje podmienky vety o rovnováhe.**