Nelineárne optimalizačné modely a metódy

Téma prednášky č. 8

Metódy transformujúce úlohu naviazaný extrém na úlohu na voľný extrém

Prof. Ing. Michal Fendek, CSc.

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

Metódy transformujúce úlohu na viazaný extrém na úlohu na voľný extrém

Nemálo aplikačných problémov ekonomického rozhodovania vedie, k formulácii úlohy nelineárneho programovania na viazaný extrém. Úloha potom vyzerá tak, že sa hľadá extrém účelovej funkcie, ktorá reprezentuje kvantitatívne formalizovaný cieľ rozhodovania, napr. maximalizácia zisku firmy, pri splnení určitých ohraničujúcich podmienok, napr. podmienky neprekročenia disponibilných zásob výrobných faktorov.

Tieto podmienky potom formálne opisujú množinu prípustných riešení úlohy. V najvšeobecnejšom tvare možno takúto úlohu formulovať nasledovne:

$$\min \{ f(x) \mid x \in D \subset R^n \} \qquad (7.5)$$

V súlade s konkrétnymi vlastnosťami účelovej funkcie a so štruktúrou množiny prípustných riešení *D* potom skúmame určité typové triedy úloh nelineárneho programovania, pre ktoré sú v mnohých prípadoch k dispozícii aj špecifické algoritmy pre ich riešenie. Zaoberajme sa teraz skúmaním takýchto tried úloh na viazaný extrém a ukážeme si niektoré algoritmy použiteľné pre ich riešenie. Začneme pomerne jednoduchými algoritmami, ktoré sú založené na princípe transformácie úlohy na viazaný extrém na úlohu na voľný extrém.

Lagrangeova metóda

Metóda sa používa na riešenie úloh nelineárneho programovania s ohraničeniami v tvare rovníc

$$f(x) \rightarrow \min\ (\max)$$

pri ohraničeniach **(7.6)**

$$h_i(x) = 0, \qquad i=1,\dots,m$$

kde funkcie f: $R^n \rightarrow R$, h_i: $R^n \rightarrow R$ pre i=1,...,m sú spojité a dvakrát spojite diferencovateľné, pričom pre počet premenných a počet ohraničení platí vzťah *m<n*.

Postup riešenia spočíva v tom, že pre úlohu (7.6) sa skonštruuje *Lagrangeova funkcia* v tvare

(7.7)

$$\boxed{L(x,v) = f(x) - \sum_{i=1}^{m} v_i h_i(x)}$$

kde v_i pre *i=1,...,m* sú tzv. *Lagrangeove multiplikátory*.

Lagrangeova metóda

Viazaný extrém funkcie *f* je totožný s voľným extrémom funkcie *L*.

Totiž viazaný extrém úlohy (7.6) x^o musí vyhovovať sústave ohraničení $h_i(x^o)=0$ pre $i=1,...m$. Preto v bode x^o platí $L(x^o,u^o)=f(x^o)$.

Takže riešenie úlohy na viazaný extrém (7.6) môže byť nahradené riešením úlohy na voľný extrém (7.7), nakoľko na množine prípustných riešení je možné nahradiť funkciu *f* funkciou *L*.

Pre riešenie úlohy (7.7) využijeme známe poznatky o extrémoch funkcií, ktoré boli uvedené v prednáške č. 3. a 4.

Na základe *nutných podmienok pre extrém* určíme stacionárne body Lagrangeovej funkcie a na základe *postačujúcich podmienok* overíme charakter identifikovaných stacionárnych bodov.

Nutné podmienky optimálnosti

Nutnou podmienkou pre to, aby vektor $x^o \in R^n$ bol riešením úlohy (7.7) je existencia takého vektora Lagrangeových multiplikátorov v_i pre $i=1,...,m$, pre ktoré platí:

$$\frac{\partial L(x^o,v^o)}{\partial x_j} = \frac{\partial f(x^o)}{\partial x_j} - \sum_{i=1}^{m} v_i \frac{\partial h_i(x^o)}{\partial x_j} = 0, \quad j=1,\dots,n \tag{7.8}$$

$$\frac{\partial L(x^o,v^o)}{\partial v_i} = h_i(x^o) = 0, \quad i=1,\dots,m$$

Riešením sústavy rovníc (7.8) vypočítame stacionárne body Lagrangeovej funkcie (x^o,v^o).

Od stupňa zložitosti funkcií f, h_i závisí, samozrejme miera numerickej obtiažnosti riešenia sústavy rovníc nutných podmienok extrému (7.8).

Postačujúce podmienky

Vieme, že overenie postačujúcich podmienok sa realizuje na základe určenia typu definitnosti kvadratickej formy Hessovej matice skúmanej funkcie. Nie je ťažké ukázať, že Hessova matica Lagrangeovej funkcie (7.7) má nasledovný tvar

$$H(L(x,v)) = \begin{pmatrix} \dfrac{\partial^2 L(x,v)}{\partial x_i \partial x_j} & \dfrac{\partial^2 L(x,v)}{\partial x_i \partial v_l} \\ \\ \dfrac{\partial^2 L(x,v)}{\partial v_l \partial x_j} & 0 \end{pmatrix}, \quad pre\ i, j = 1,\dots,n;\ l = 1,\dots,m$$

Hessova matica Larangeovej funkcie má teda rozmer (n+m,n+m) a je to bloková matica pozostávajúca zo 4 blokov.
Ľavý horný blok je tvorený maticou druhých parciálnych derivácií Lagrangeovej funkcie podľa premenných x_j.
Ľavý dolný blok je tvorený maticou, ktorej riadky sú gradienty funkcií sústavy ohraničení s opačným znamienkom a pravý horný blok je transponovanou maticou matice ľavého dolného bloku.
Pravý dolný blok je nulovou maticou.

Lagrangeova metóda - postačujúce podmienky

Hessovu maticu Lagrengeovej funkcie teda môžeme zapísať aj nasledovne:

$$H(L(x,v)) = \begin{pmatrix} \dfrac{\partial^2 L(x,v)}{\partial x_i \partial x_j} & -\dfrac{\partial h_l(x)}{\partial x_i} \\ -\dfrac{\partial h_l(x)}{\partial x_j} & 0 \end{pmatrix}, \quad pre\ i, j = 1,\dots,n;\ l = 1,\dots,m$$

Označme ďalej jako

$$H_k(L(x,v))$$

submaticu o rozmere (m+k,m+k), ktorú dostaneme z Hessovej matice H(L(x,v)) tak, že z nej vyškrtneme všetky riadky počnúc (k+1)-vým a končiac n-tým a analogicky všetky stĺpce počnúc (k+1)-vým a končiac n-tým.

Potom

a) postačujúcou podmienkou pre to, aby úloha (7.6) mala viazané maximum v stacionárnom bode (x^o,v^o), je platnosť podmienky

$$(-1)^k \left| H_k(L(x^o,v^o)) \right| > 0, \quad \text{pre } k=m+1,\dots,n \qquad (7.9)$$

b) postačujúcou podmienkou pre to, aby úloha (7.6) mala viazané minimum v stacionárnom bode (x^o,v^o), je platnosť podmienky

$$(-1)^m \left| H_k(L(x^o,v^o)) \right| > 0, \quad \text{pre } k=m+1,\dots,n \qquad (7.10)$$

<u>Príklad č.7.5</u>

S použitím *Lagrangeovej metódy* riešme úlohu nelineárneho programovania

$$f(x_1,x_2) = 6 - 4x_1 - 3x_2 \rightarrow \min$$

pri ohraničení

$$x_1^2 + x_2^2 - 1 = 0$$

Riešenie:

Lagrangeova funkcia uvedenej úlohy má tvar

$$L(x_1,x_2,v) = 6 - 4x_1 - 3x_2 - v(x_1^2 + x_2^2 - 1)$$

Sformulujme teraz nutné podmienky existencie extrému Lagrangeovej funkcie v tvare (7.8)

Sformulujme teraz nutné podmienky existencie extrému Lagrangeovej funkcie v tvare (7.8)

$$1.\ \frac{\partial L(x,v)}{\partial x_1} = -4 - 2x_1 v = 0$$

$$2.\ \frac{\partial L(x,v)}{\partial x_2} = -3 - 2x_2 v = 0$$

$$3.\ \frac{\partial L(x,v)}{\partial v} = x_1^2 + x_2^2 - 1 = 0$$

Riešenín tejto sústavy rovníc sú dva stacionárne body Lagrangeovej funkcie,

a) $(x^{o1}, v^{o1}) = (-0.8, -0.6, 2.5)$
b) $(x^{o2}, v^{o2}) = (0.8, 0.6, -2.5)$

Lagrangeova metóda

Preverme platnosť postačujúcich podmienok optimálnosti pre tieto stacionárne body. Vypočítajme Hessovu maticu Lagrangeovej funkcie úlohy

a) (x^{o1},v^{o1}) = (-0.8, -0.6, 2.5)
b) (x^{o2},v^{o2}) = (0.8, 0.6, -2.5)

$$H(L(x,v))=\begin{pmatrix} -2v & 0 & -2x_1 \\ 0 & -2v & -2x_2 \\ -2x_1 & -2x_2 & 0 \end{pmatrix}$$

a napokon Hessove matice úlohy v jej stacionárnych bodoch.

$$H(L(x^o,v^o)^1)=\begin{pmatrix} -5 & 0 & 1.6 \\ 0 & -5 & 1.2 \\ 1.6 & 1.2 & 0 \end{pmatrix}$$

$$H(L(x^o,v^o)^2)=\begin{pmatrix} 5 & 0 & -1.6 \\ 0 & 5 & -1.2 \\ -1.6 & -1.2 & 0 \end{pmatrix}$$

V súlade s postačujúcimi podmienkami optimálnosti (7.9),(7.10) preskúmame znamienka determinantov matíc $H_k(L(x^o,v^o))$ pre oba stacionárne body pre všetky $k=m+1,\ldots,n$. V našom príklade však $m=1$, $n=2$ a preto $k=2=n$, takže stačí preskúmať iba determinanty pôvodných Hessových matíc, nakoľko $H_n=H$.

Lagrangeova metóda

Dostávame

$\det H(L(x^{o1},v^{o1}) = 12.8$ $\qquad$ $\det H(L(x^{o2},v^{o2}) = -12.8$

Vidíme, že:

- determinant Hesovej matice v bode $(x^{o1},v^{o1}) = (-0.8, -0.6, 2.5)$ vyhovuje postačujúcej podmienke (7.9), pretože

$(-1)^k\det H(L(x^{o1},v^{o1}) = (-1)^2(12.8) > 0$, pre $k=m+1=2$

a bod (x^{o1},v^{o1}) je teda bodom lokálneho maxima úlohy s hodnotou účelovej funkcie $f_{max}=11$;

- determinent Hessovej matice v bode $(x^{o2},v^{o2}) = (0.8, 0.6, -2.5)$ vyhovuje postačujúcej podmienke (7.10), pretože

$(-1)^m\det H(L(x^{o2},v^{o2}) = (-1)^1(-12.8) > 0$,

a bod (x^{o2},v^{o2}) je bodom lokálneho minima úlohy s hodnotou účelovej funkcie $f_{min}=1$.

Metóda penalizačnej funkcie

Pre objasnenie podstaty tejto metódy skúmajme najprv úlohu nelineárneho programovania s jedným ohraničením v tvare

$$f(x) \rightarrow min$$

pri ohraničení

$$h(x) = 0$$

Túto úlohu budeme transformovať na úlohu voľnej optimalizácie

$$f(x) + \mu h^2(x) \rightarrow min$$

pri ohraničení

$$x \in R^n$$

kde μ>0 je určité veľké číslo. Intuitívne je zrejmé, že pri optimálnom riešení x^o transformovanej úlohy musí byť hodnota $h^2(x^o)$ blízka k nule, pretože v opačnom prípade je vždy možné nájsť nejaký iný bod x^*, v ktorom prírastok f(x) bude pri dostatočne veľkom μ menší ako μh^2(x).

Metóda penalizačnej funkcie

Príklad č.7.6 [Bunday]

S použitím penalizačnej funkcie riešme nasledovnú úlohu nelineárneho programovania.

$f(x) = x^2 \to \min$

pri ohraničení $h(x) = x - 2 = 0$.

Riešenie:

Optimálnym riešením úlohy je evidentne bod $x^o=2$ a hodnota účelovej funkcie $f(x^o)=4$. Skonštruujme penalizačnú funkciu úlohy a riešme zodpovedajúcu úlohu na voľný extrém

$$h^2(x) = (x-2)^2$$

$$F(x,\mu) = f(x) + \mu h^2(x) \to \min$$

$$F(x,\mu) = x^2 + \mu(x-2)^2 \to \min$$

kde $x \in R^1$, $\mu > 0$. Preverme nutnú podmienku optimálnosti a určme stacionárny bod funkcie $F(x,\mu)=f(x)+\mu h(x)$:

$$\frac{dF(x,\mu)}{dx} = 2x + \mu(2x-4) = 0 \to x + \mu x - 2\mu = 0 \to x(1+\mu) = 2\mu$$

$$\to \quad x^o = \frac{2\mu}{1+\mu}$$

Metóda penalizačnej funkcie

Stacionárny bod je však funkciou parametra μ, ktorý ale v bode minima transformovanej úlohy nadobúda hodnotu $\mu \to \infty$ a platí

$$x^o = x^o(\mu) = \frac{2\mu}{1+\mu}, \quad \mu > 0 \quad \to$$

$$x^o = \lim_{\mu \to \infty} \frac{2\mu}{1+\mu} = 2$$

Overme teraz postačujúcu podmienku existencie minima funkcie $F(x,\mu)$ v bode x^o

$$\frac{d^2 F(x,\mu)}{dx^2} = 2 + 2\mu > 0, \quad pre\ \mu > 0$$

Funkcia $F(x,\mu)$ je teda rýdzokonvexná na celom obore definície pre $\mu>0$, takže bod $x^o=2$ je bodom globálneho minima funkcie $F(x,\mu)$ a je teda aj riešením pôvodnej úlohy na viazaný extrém.

Metóda penalizačnej funkcie

Uvedený príklad bol, pravdaže, iba ilustratívny a jeho cieľom bolo vysvetliť ídeu penalizačnej funkcie. Úlohy so zložitejšou štruktúrou ohraničení a s väčším počtom premenných sa nerieši jednoduchým overením nutnej a postačujúcej podmienky extrému. Používané algoritmy však z tejto idey vychádzajú.

Preskúmajme teraz úlohu s ohraničením-nerovnicou v tvare

$$f(x) \to min$$

pri ohraničení

$$g(x) \leq 0$$

Je zrejmé, že v takomto prípade nie je tvar penalizačnej funkcie $f(x)+\mu h^2(x)$ vhodný, pretože pri výbere bodu, pre ktorý $g(x) \neq 0$ sa uplatní automaticky pokuta bez ohľadu na to, aké je znamienko hodnoty $g(x)$. Pokuta je však želateľná len v prípadoch neprípustnosti bodu x, t.j. ak $g(x)>0$. Prijateľnou bude teda v tomto prípade nasledujúca transformovaná úloha voľnej minimalizácie

$$F(x,\mu) = f(x) + \mu \times max\{0, g(x)\} \to min$$

pri ohraničení

$$x \in R^n$$

Vidíme, že v prípade, ak skúmame prípustný bod x, a teda platí $g(x) \le 0$, tak sa penalizačná funkcia neuplatní, naopak, ak skúmame neprípustný bod x, a teda platí $g(x) > 0$, tak sa uplatní pokuta vo výške $\mu g(x)$.

V prípade úlohy nelineárneho programovania so zmiešanými typmi ohraničení v tvare

$$f(x) \rightarrow \min$$

pri ohraničeniach

$$g_i(x) \le 0, \quad \text{pre } i = 1,\dots,m \qquad (7.11)$$
$$h_i(x) = 0, \quad \text{pre } i=1,\dots,l$$
$$x \in X \in R^n$$

sa obvykle konštruuje penalizačná funkcia v tvare

$$\boxed{\alpha(x) = \sum_{i=1}^{m} [\max\{0, g_i(x)\}]^p + \sum_{i=1}^{l} |h_i(x)|^p} \qquad \textbf{(7.12)}$$

kde $p > 0$, $p \in Z$.

Metóda penalizačnej funkcie

Potom sa konštruuje tzv. *pomocná úloha* v tvare

$\theta(\mu) \rightarrow \min$
pri ohraničení

$$\mu \geq 0$$

kde

$$\theta(\mu) = \inf \{ f(x) + \mu\alpha(x) \mid x \in X \}$$

pričom $\alpha(x)$ je spojitá penalizačná funkcia v tvare (7.12).
Dá sa ukázať, že platí

$$\inf \{ f(x) \mid x \in X, g(x) \leq 0, h(x) = 0 \} =$$

$$= \sup_{\mu \geq 0} \theta(\mu) = \lim_{\mu \to \infty} \theta(\mu)$$

Z uvedeného vzťahu vyplýva, že kvantifikáciou funkcie $\theta(\mu)$ pre dostatočne veľkú hodnotu μ sa možno dosť "priblížiť" k optimálnej hodnote účelovej funkcie východiskovej úlohy. Môžu ale vzniknúť určité problémy ak zvolíme dosť veľkú hodnotu μ a pokúsime sa riešiť pomocnú úlohu.

Metóda penalizačnej funkcie

Pri veľkej hodnote μ sa totiž najväčšia pozornosť venuje prípustnosti aktuálneho skúmaného riešenia a väčšina algoritmov voľnej optimalizácie má potom tendenciu pohybom v smere gradientu pomocnej funkcie identifikovať čo najskôr prípustné riešenie a predčasne zastaviť výpočty.

Avšak takto nájdené riešenie môže byť dosť "vzdialené" od optimálneho riešenia východiskovej úlohy.

Výpočtové problémy spojené s aplikáciou penalizačného parametra μ veľkej hodnoty sa vo väčšine konkrétnych implementácií algoritmov penalizačných funkcií riešia postupnosťou rastúcich hodnôt penalizačného parametra μ_k pre $k=1,2,...$ na jednotlivých iteráciách algoritmu.

Algoritmus metódy penalizačnej funkcie

Inicializačná fáza

Zvolíme

- **konštantu testu ukončenia algoritmu $\epsilon>0$,**
- **východiskový bod x^1,**
- **východiskový penalizačný parameter $\mu_1>0$,**
- **koeficient transformácie penalizačného parametra $b>0$.**

Položíme $k=1$ a prejdeme k výpočtovej fáze.

Výpočtová fáza

Krok 1º.

Pri zadanom východiskovom bode x^1 riešime úlohu

$$F(x) = f(x) + \mu_k \alpha(x)$$

pri ohraničení

$$x \in X$$

Položíme

$$x^{k+1} = x^{\mu k}$$

kde $x^{\mu k}$ je optimálne riešenie úlohy v kroku 1º.
Prejdeme na *Krok 2º*.

Krok 2º.

- ak $\mu_k \alpha(x^{k+1}) < \varepsilon$ algoritmus končí a bod x^{k+1} je optimálnym riešením východiskovej úlohy,

- v opačnom prípade

- položíme $\mu_{k+1} = b\mu_k$

- zameníme $k \to k+1$ a vrátime sa na *Krok 1º* výpočtovej fázy.

Poznámka

Algoritmus je formulovaný pre úlohy v tvare (7.11) a penalizačnú funkciu $\alpha(x)$ v tvare (7.12). Táto metóda nevyžaduje žiadne špeciálne vlastnosti funkcií f, g_i, h_i, okrem ich spojitosti. Efektívne použitie algoritmu však predpokladá efektívnu procedúru pre riešenie úlohy voľnej minimalizácie na kroku 1º *výpočtovej fázy* algoritmu.

Príklad č.7.7 [Bazaraa]

S použitím algoritmu penalizačnej funkcie riešme úlohu

$$f(x_1, x_2) = (x_1 - 2)^4 + (x_1 - 2x_2)^2 \rightarrow \min$$

pri ohraničeniach

$$x_1^2 - x_2 = 0$$

$$x \in X \subset R^2$$

Metóda penalizačnej funkcie

Riešenie

Na každej k-tej iterácii musíme v snahe získať $x^{\mu k}$ riešiť pri zadanej hodnote μ^k úlohu voľnej optimalizácie v tvare

$$F(x_1, x_2) = (x_1 - 2)^4 + (x_1 - 2x_2)^2 + \mu(x_1^2 - x_2)^2 \rightarrow \min$$

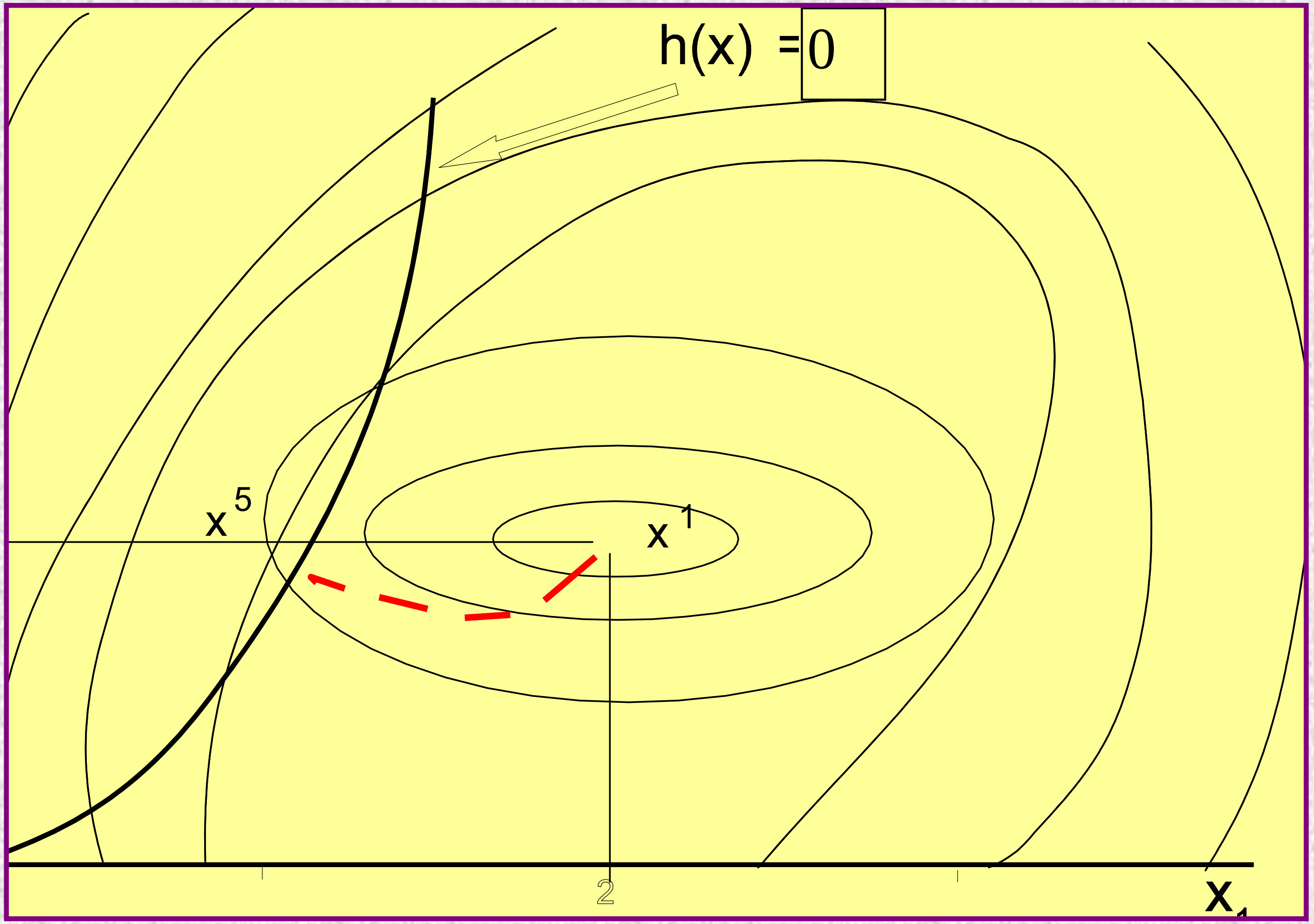
pri ohraničení

$$x \in X \subset R^2$$

V tabuľke č.7.10 sú uvedené výsledky výpočtov pri riešení úlohy metódou penalizačnej funkcie na jednotlivých iteráciách. Na obr. 7.3 je postup výpočtu znázornený graficky.

Metóda penalizačnej funkcie

Obr.č.7.3: Geometrická interpretácia metódy penalizačnej funkcie



Metóda penalizačnej funkcie

Inicializačná fáza

Zvolíme

- východiskový bod $x^1=(2,1)^T$, t.j. neprípustný bod globálneho voľného minima účelovej funkcie úlohy,
- penalizačný východiskový parameter $\mu_1=0.1$,
- transformačný koeficient b=10.0.
- parameter testu konca algoritmu $\varepsilon = 0{,}05$

Položíme *k=1* a prejdeme na *výpočtovú fázu*

Výpočtová fáza

Krok 1°.

Riešime úlohu voľnej minimalizácie

$$F(x_1, x_2) = (x_1 - 2)^4 + (x_1 - 2x_2)^2 + 0.1(x_1^2 - x_2)^2 \to \min$$

pri ohraničení

$$x \in X \subset R^2$$

Metóda penalizačnej funkcie

Optimálne riešenie úlohy je

$$\mathbf{x}^{\mu 1} = (1.4539,\ 0.7608)^T$$

a ďalšie vypočítané hodnoty sú nasledovné

- aktuálny bod po prvej iterácii $x^2 = x^{\mu 1}$,

- hodnota účelovej funkcie východiskovej úlohy: $f(x^2) = 0.0935$

<u>Tab. č.7.10:</u> *Riešenie úlohy s metódou penalizačnej funkcie*

k	μ_k	$x^{k+1}=x^{\mu k}$	$f(x^{k+1})$	$\alpha(x^{\mu k})=h^2(x^{\mu k})$	$\theta(\mu_k)=F(x^{\mu k})$	$\mu_k\alpha(x^{\mu k})$
1	0.100	(1.45,0.76)	0.0935	1.8307	0.2766	0.1831
2	1.000	(1.17,0.74)	0.5753	0.3908	0.9661	0.3908
3	10.00	(0.99,0.84)	0.5203	0.0192	1.7129	0.1926
4	100.0	(0.95,0.88)	1.8917	0.000267	1.9184	0.0267
5	10000.	(0.94,0.89)	1.9405	0.0000028	1.9433	0.0028

Krok 2º.

Overíme podmienku ukončenia algoritmu:

$$\mu^1 \alpha(x^{\mu 1}) = 0.1831 > \varepsilon$$

Vidíme, že podmienka kritéria nie je splnená, takže bod x^2 nie je optimálnym riešením úlohy.

Preto položíme

$$k=k+1=2,$$

vypočítame nový penalizačný parameter

$$\mu_2 = b\mu_1 = 10.0x0.1 = 1.0$$

a vrátime sa na krok 1º.

Všimnime si, že po vykonaní štyroch iterácií je podmienka konca algoritmu splnená, platí totiž

$$\mu_4 \alpha(x^{\mu 4}) = 0.0267 < 0.1 = \varepsilon.$$

Optimálnym riešením východiskovej úlohy je teda bod

$$x^{\mu 4} = (0.9507,\ 0.8875)^T.$$

Piata iterácia bola vykonaná iba preto, aby sme ukázali, že proces riešenia skutočne konverguje k nulovej hodnote

$$\mu_5 \alpha(x^{\mu 5}) = 0.0028 \to 0.$$

Môžeme sa ľahko presvedčiť o tom, že v bode $x^{\mu 5}=(0.9461,\ 0.8934)^T$ sú splnené podmienky optimálnosti Kuhna-Tuckera pre Lagrangeov multiplikátor s hodnotou $v=3.3631$.

Metódy bariérových funkcií

Základná myšlienka metódy bariérových funkcií spočíva v transformácii úlohy hľadania viazaného minima funkcie f(x), t.j. úlohy

$$\boxed{\min\left\{\; f(\mathbf{x}) \,\middle|\, \mathbf{x} \in D \subset R^n \;\right\}}$$

na úlohu nájdenia voľného minima funkcie F(x,μ), t.j. úlohy

$$\boxed{\min\left\{\; F(\mathbf{x},u) = f(\mathbf{x}) + \mu B(\mathbf{x}) \,\middle|\, \mathbf{x} \in R^n, \mu > 0 \;\right\}}$$

kde B(**x**) je tzv. *bariérová funkcia*. Požaduje sa, aby táto funkcia nadobudla vysokú hodnotu, ak skúmané riešenie je z hľadiska východiskovej úlohy neprípustné, inými slovami sú porušené niektoré ohraničenia úlohy.

Metódy bariérových funkcií

Skúmajme úlohu nelineárneho programovania s ohraničeniami v tvare nerovníc:

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

pri ohraničeniach

$$g_i(x) \leq 0, i = 1, \ldots, m$$

Typická bariérová funkcia má tvar

$$B(x) = \sum_{i=1}^{m} \frac{-1}{g_i(x)} \tag{7.13}$$

a transformovaná úloha voľnej optimalizácie je potom následovná

$$\min\{\ F(x,\mu) \mid x \in R^n, g_i(x) \leq 0, i = 1, \ldots, m; \mu > 0\ \}$$

kde

$$F(x,\mu) = f(x) + \mu B(x) = f(x) + \mu \sum_{i=1}^{m} \frac{-1}{g_i(x)} \rightarrow \min$$

Metódy bariérových funkcií

Skúmajme teraz jednoduchý ilustratívny príklad, ktorý však umožní ozrejmiť efekt pôsobenia bariérovej funkcie.

Príklad

S použitím bariérovej funkcie definovanej riešme nasledovnú úlohu nelineárneho programovania.

$$f(x_1, x_2) = 1/3(x_1 + 1)^3 + x_2 \to \min$$

pri ohraničeniach

$$1 - x_1 \leq 0$$
$$-x_2 \leq 0$$

Riešenie:

Bariérová funkcia úlohy je nasledovná

$$B(x) = \frac{-1}{-(x_1 - 1)} + \frac{1}{x_2} = \frac{1}{x_1 - 1} + \frac{1}{x_2}, \quad pre\ x_1, x_2 \neq 0$$

Metódy bariérových funkcií

Riešme ďalej zodpovedajúcu úlohu optimalizácie na voľný extrém

$$F(x,\mu) = f(x) + \mu B(x) =$$

$$= \frac{1}{3}(x_1 + 1)^3 + x_2 + \frac{\mu}{x_1 - 1} + \frac{\mu}{x_2} \rightarrow \min$$

kde $\mu > 0$.

Preverme nutné podmienky optimálnosti a určme stacionárny bod funkcie $F(x,\mu) = f(x) + \mu B(x)$:

$$\frac{\partial dF(x,\mu)}{\partial x_1} = (x_1 + 1)^2 - \frac{\mu}{(x_1 - 1)^2} = 0 \rightarrow (x_1 + 1)^2 (x_1 - 1)^2 - \mu = 0 \rightarrow$$

$$\left(x_1^2 - 1\right)^2 = \mu \rightarrow \left(x_1^2 - 1\right) = \sqrt{\mu} \rightarrow$$

$$x_1^o = \sqrt{1 + \sqrt{\mu}}$$

$$\frac{\partial dF(x,\mu)}{\partial x_2} = 1 - \frac{\mu}{x_2^2} = 0 \rightarrow x_2^2 = \mu \rightarrow x_2^o = \sqrt{\mu}$$

Metódy bariérových funkcií

Podobne ako v príklade z časti o „penalizačnej funkcii, je aj tu stacionárny bod funkciou parametra μ, ktorý v bode minima transformovanej úlohy nadobúda pri metóde bariérovej funkcie hodnotu $\mu \to 0$, takže platí

$$x_1^o = x_1^o(\mu) = \sqrt{1+\sqrt{\mu}}\,,\ \mu > 0 \ \rightarrow$$

$$x_1^o = \lim_{\mu \to 0} \sqrt{1+\sqrt{\mu}} = 1$$

$$x_2^o = x_2^o(\mu) = \sqrt{\mu}\,,\ \mu > 0 \ \rightarrow$$

$$x_2^o = \lim_{\mu \to 0} \sqrt{\mu} = 0$$

$$f(x_1, x_2) = 1/3(x_1+1)^3 + x_2 = 8/3$$

Metódy bariérových funkcií

Overme postačujúce podmienky existencie minima v bode $(x_1,x_2)^o$. Vypočítajme Hessovu maticu funkcie $F(\mathbf{x},\mu)$.

Dostávame:

$$H(x) = \begin{pmatrix} 2(x_1+1) + \dfrac{2\mu}{(x_1-1)^3} & 0 \\ 0 & \dfrac{2\mu}{x_2^3} \end{pmatrix}$$

Určme definitnosť Hessovej matice v stacionárnom bode funkcie $F(\mathbf{x},\mu)$, tj. v bode $(x_1,x_2)^o=(1,0)^T$.

Metódy bariérových funkcií

Vidíme, že druhé parciálne derivácie funkcie $F(\mathbf{x},\mu)$ nie sú v stacionárnom bode spojité. Pre hodnoty súradníc blížiace sa limitne k hodnotám komponent stacionárneho bodu sprava je však Hessova matica

$$H(x_1 \to 1^+,\ x_2 \to 0^+,\ \mu \to 0^+) = \begin{pmatrix} 2(x_1+1)+\dfrac{2\mu}{(x_1-1)^3} & 0 \\ 0 & \dfrac{2\mu}{x_2^3} \end{pmatrix} = \begin{pmatrix} a>0 & 0 \\ 0 & b>0 \end{pmatrix}$$

kladne definitná vo vnútorných bodoch množiny prípustných riešení úlohy. Takže funkcia $F(\mathbf{x},\mu)$ je rýdzokonvexná vo vnútri množiny prípustných riešení pre $\mu \to 0^+$ a bod

$(x_1^o \to 1^+,\ x_2^o \to 0^+)$

je bodom globálneho minima funkcie F(x,μ) s hodnotou funkcie

$$F(\mathbf{x},\mu) \to 8/3, \text{ pre } \mu \to 0^+$$

a je teda aj riešením pôvodnej úlohy na viazaný extrém.

Metódy bariérových funkcií

Poznámka:

Samozrejme, reálne úlohy so zložitejšou štruktúrou ohraničení a s väčším počtom premenných, sa nedajú riešiť jednoduchým overením nutnej a postačujúcej podmienky extrému. Algoritmus metódy bariérovej funkcie však využíva práve transformačnú funkciu v tvare (7.13) a na hľadanie jej minima sa používa určitá iteračná procedúra (bližšie pozri napr. [Bazaraa, 2009], [Elster, 1975]).