Nelineárne optimalizačné modely a metódy

Téma prednášky č. 7

Metódy pre riešenie úloh na voľný extrém

Prof. Ing. Michal Fendek, CSc.

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

Metódy minimalizácie funkcie jednej premennej

Hľadanie extrému funkcie jednej premennej je základom mnohých algoritmov pre riešenie úloh nelineárneho programovania. Obvyklá schéma algoritmov pre riešenie úlohy nelineárneho programovania je totiž nasledovná:

-zadá sa východiskový prípustný bod x^k, definuje sa smerový vektor d^k a optimálna dĺžka kroku λ^k.

-Potom sa určí nasledovný bod $x^{k+1}=x^k+\lambda^k d^k$. Tento proces sa opakuje. Nájdenie optimálnej dĺžky kroku λ^k sa realizuje prostredníctvom minimalizácie hodnoty funkcie $f(x^k+\lambda d^k)$ s premennou λ, čo je vlastne práve už spomenutá úloha lineárneho hľadania.

Skúmajme teda úlohu

$$\min \{ \theta(\lambda) \mid \lambda \in R \}$$

kde $\theta(\lambda)=f(x+\lambda d)$

Preskúmajme funkciu

$$f(x_1, x_2) = 2({x_1}^2 - 3) + {x_2}^2$$

Formulujme funkciu $\theta(\lambda)$ pre bod $\mathbf{x} = (2,3)$ s hodnotou funkcie $f = 11$ a vektor prípustného smeru $\mathbf{d} = (3, -1)$

Poznámka: $GLMIN(x_1, x_2) = (0,0)\ f = -6, -\nabla f(2,3) = -(4x_1, 2x_2) = (-8, -6)$

$$\theta(\lambda) = f(\mathbf{x} + \lambda\mathbf{d}) = f(2 + 3\lambda, 3 - \lambda) = 2\big((2 + 3\lambda)^2 - 3\big) + (3 - \lambda)^2 =$$
$$= 8 + 24\lambda + 18\lambda^2 - 6 + 9 - 6\lambda + \lambda^2 = 19\lambda^2 + 18\lambda + 11$$

NP extrému

$$\frac{d\theta(\lambda)}{d\lambda} = (19\lambda^2 + 18\lambda + 11)' = 38\lambda + 18 = 0 \Rightarrow \lambda^* = -18/38 = -0{,}47$$

PP extrému

$$\frac{d^2\big(\theta(\lambda)\big)}{d\lambda^2} = 38 > 0 \quad \text{Lokálne minimum}$$

$$\tilde{\mathbf{x}} = \mathbf{x} + \lambda\mathbf{d} = (2 - 1{,}41, 3 + 0{,}47) = (-0{,}59, 3{,}47)\ \ f = 6{,}31$$

Jedna z možností, ako riešiť túto úlohu, je pri splnení predpokladu o diferencovateľnosti funkcie $\theta(\lambda)$, riešiť rovnicu $\theta'(\lambda)=0$.

To však je nezriedka vzhľadom na možný zložitý analytický tvar funkcie $\theta(\lambda|$ a teda i komplikovaný tvar funkcie prvej derivácie rovnako náročné ako riešenie samotnej pôvodnej úlohy a teda dochádza už k spomenutému paradoxu nahradenia riešenia jednej komplikovanej úlohy riešením inej komplikovanej úlohy.

Bolo by totiž potrebné pre neznámu λ riešiť rovnicu

$$\theta'(\lambda) = \frac{d\mathrm{f}(\mathrm{x}+\lambda \mathrm{d})}{d\lambda} = \mathrm{d}^T \nabla \mathrm{f}(\mathrm{x}+\lambda \mathrm{d}) = 0.$$

Odhliadnuc od toho, že ak nie je dieferencovateľná funkcia $f(\mathrm{x})$, tak potom nie je diferencovateľná ani funkcia $\theta(\lambda)$. Preto sa takýto postup, s výnimkou niektorých špeciálnych prípadov nepoužíva.

Namiesto toho sa využívajú numerické procedúry minimalizácie funkcie $\theta(\lambda)$ na uzavretom ohraničenom intervale, t.j. procedúry pre riešenie úlohy

$$\min \{ \theta(\lambda) \mid a \leq \lambda \leq b \} \qquad (7.2)$$

O funkcii $\theta(\lambda)$ predpokladáme, že je rýdzo kvázikonvexná, prípadne unimodálna a interval $\langle a,b\rangle$, ktorý obsahuje bližšie neidentifikovaný bod minima λo funkcie $\theta(\lambda)$ sa nazýva *intervalom neurčitosti L.* Ak v procese hľadania možno z intervalu neurčirosti eliminovať nejaké jeho časti, o ktorých je preukázané, že neobsahujú bod minima, tak interval neurčitosti sa skracuje. Dá sa ukázať, že pre rýdzo kvázikonvexnú funkciu možno interval neurčitosti redukovať pomocou výpočtu hodnôt funkcie $\theta(\lambda)$ v dvoch rôznych bodoch tohto intervalu.

Veta 7.1

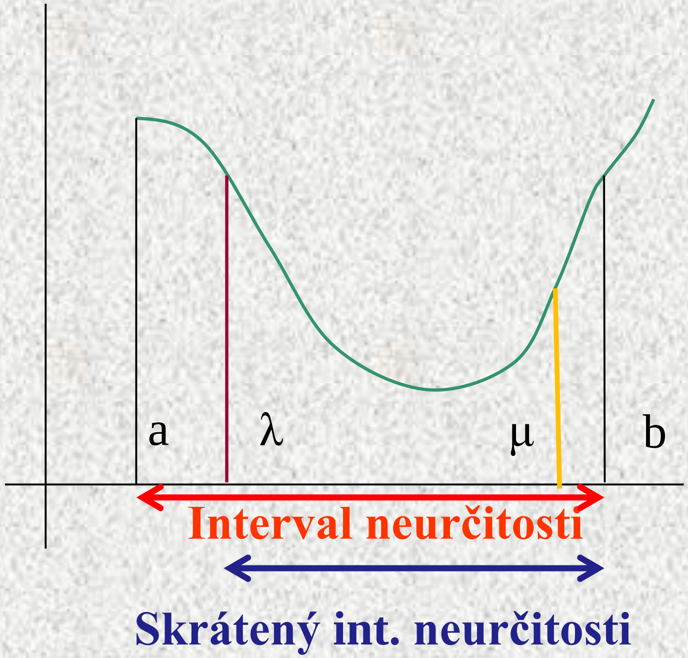
Nech $\theta : R \rightarrow R$ je rýdzo kvázikonvexná funkcia na intervale $\langle a,b\rangle$. Nech ďalej sú dané čísla $\lambda, \mu \in \langle a, b\rangle$ také, že platí $\lambda < \mu$. Potom:

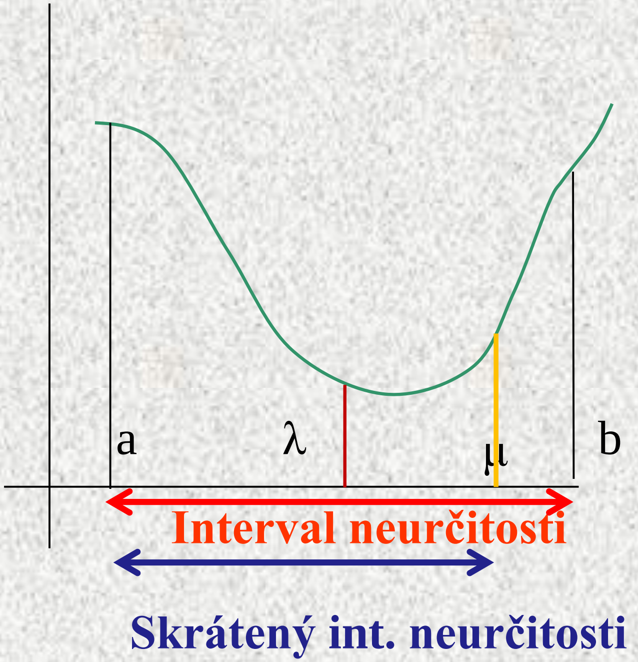
1. ak $\theta(\lambda) > \theta(\mu)$, potom $\theta(z) \geq \theta(\mu)$ pre $\forall z \in \langle a, \lambda)$

2. ak $\theta(\lambda) \leq \theta(\mu)$, potom $\theta(z) \geq \theta(\mu)$ pre $\forall z \in \langle \mu, b)$

Z vety 7.1 vyplýva, že pri platnosti predpokladu rýdzej kvázikonvexnosti funkcie θ z podmienky θ(λ) > θ(μ) vyplýva, že novým intervalom neurčitosti je interval ⟨λ,b⟩. Na druhej strane, ak platí θ(λ) ≤ θ(μ), potom novým intervalom neurčitosti je interval ⟨a,μ⟩. Závery vety 7.1 sú ilustrované na obr.č.7.1.

Obr.č.7.1:Skracovanie intervalu neurčitosti L





Poznamenávame, že interval neurčitosti *L*, na ktorom sa hľadá minimum funkcie θ má obvykle jeden z nasledovných tvarov:

$$L = \{ \lambda \mid \lambda \in R \}$$
$$L = \{ \lambda \mid \lambda \geq 0 \}$$
$$L = \{ \lambda \mid a \leq \lambda \leq b \}$$

Je zrejmé, že trieda metód lineárneho hľadania, ktorá nevyužíva derivácie funkcie vychádza z horeuvedeného princípu skracovania intervalu neurčitosti *L*, a dosiahnutie jeho vopred stanovenej *konečnej dĺžky l*, je kritériom konca algoritmov.

Metódy sa od seba odlišujú *de facto* iba spôsobom, akým sa na *k-tom* kroku algoritmu vypočítavajú body λ^k, $\mu^k \in \langle a^k, b^k \rangle$.

Metóda dichotomického hľadania

Metóda dichotomického hľadania vyberá body λ^k, μ^k symetricky s odchyľkou $\varepsilon > 0$ voči stredu aktuálneho intervalu $\langle a^k , b^k \rangle$. číslo $\varepsilon > 0$ sa volí dostatočne malé, ale tak, aby hodnoty funkcií $\theta(\lambda^1)$ a $\theta(\mu^1)$ pri prvom skracovaní intervalu boli rôzne.

Algoritmus metódy dichotomického hľadania:

Skúmajme rýdzo kvázikonvexnú funkciu $\theta(\lambda)$, ktorej hodnotu minimalizujeme na intervale $<a^1,b^1>$.

Inicializačná fáza

Zvolíme

- rozlišovaciu konštantu $2\varepsilon>0$,

- prípustnú konečnú dĺžku intervalu neurčitosti λ.

Nech $<a^1,b^1>$ je východiskový interval neurčitosti. Položíme $k=1$ a prejdeme k výpočtovej fáze.

Výpočtová fáza

Krok 1°.

- Ak $a^k \rightarrow b^k < \lambda$, algoritmus končí a hľadaný bod minima patrí do intervalu $<a^k,b^k>$.

- V opačnom prípade vypočítame

$$\lambda^k=(a^k+b^k)/2 - \varepsilon, \qquad \mu^k=(a^k+b^k)/2 + \varepsilon$$

a pokračujem na krok 2°.

Krok 2°.

a) - Ak $\theta(\lambda^k)<\theta(\mu^k)$, položíme $a^{k+1}=a^k$, a $b^{k+1}=\mu^k$.

- V opačnom prípade položíme $a^{k+1}=\lambda^k$, a $b^{k+1}=b^k$.

b) Zmeníme

$k \rightarrow k+1$ a vrátime sa na krok 1°.

Príklad č.7.1

S použitím metódy dichotomického hľadania riešte úlohu f(x) → min, kde f: R → R, pričom analytický tvar funkcie *f(x)* je nasledovný

$$f(x) = \begin{cases} x^2 + 1, & x \in < -1,2 > \\ 3x - 1, & x \in < 2,3 > \end{cases}$$

Riešenie:

Zvoľme konečnú dĺžku intervalu neurčitosti λ = 0.05 a rozlišovaciu konštantu 2ε = 0.01. Je zrejmé, že funkcia je rýdzo kvázikonvexná. Nie je však diferencovateľná, nakoľko derivácia funkcie v bode *x* = 2 zľava sa nerovná derivácii funkcii v tomto bode sprava,

$$\frac{df(x)}{dx} = 2x \qquad [x \to 2_L], \qquad \frac{df(x)}{dx} = 3 \qquad [x \to 2_P]$$

Optimálne riešenie úlohy sa realizuje v bode x^o = 0, f(*x*) = 1, ľahko sa o tom môžeme presvedčiť napríklad na grafe funkcie (všimnime si tiež, že funkcia je skutočne rýdzo kvázikonvexná).

Vypočty realizované na jednotlivých iteráciách metódy dichotomického hľadania sú uvedené v tabuľke č.7.1.

Metódy na riešenie úloh na voľný extrém

k	a^k	b^k	λ^k	μ^k	$\theta(\lambda^k)$	$\theta(\mu^k)$
1	-1	3	0,9950	1,0050	1,9900	2,0100
2	-1	1,0050	-0,0025	0,0075	1,0000	1,0001
3	-1	0,0075	-0,5013	-0,4913	1,2513	1,2414
4	-0,5013	0,0075	-0,2519	-0,2419	1,0635	1,0585
5	-0,2519	0,0075	-0,1272	-0,1172	1,0162	1,0137
6	-0,1272	0,0075	-0,0649	-0,0549	1,0042	1,0030
7	-0,0649	0,0075	-0,0337	-0,0237	1,0011	1,0006
8	-0,0337	0,0075	-0,0181	-0,0081	1,0003	1,0001
9	-0,0181	0,0075	-0,0103	-0,0003	1,0001	1,0000
10	-0,0103	0,0075	-0,0064	0,0036	1,00004	1,00001
11	-0,0103	0,0036				

Vidíme, že na 11. iterácii je splnená podmienka ukončenia algoritmu

$$|a^{11} - b^{11}| = 0{,}0463 < l = 0{,}05$$

a minimum funkcie sa dosahuje v bode v rámci intervalu

$$x^o \in \langle -0{,}0103;\ 0{,}0036 \rangle$$

Metódy minimalizácie funkcie *n* premenných

sa používajú na riešenie úloh nelineárneho progamovania v tvare

$$\min \{ f(x) \mid x \in R^n \} \qquad (7.4)$$

Jeden z možných prístupov riešenia, ktorý nevyžaduje diferencovateľnosť funkcie *f*(x) je nasledovný. Zadá sa východiskový prípustný bod x^k a definuje sa smerový vektor d^k. Potom, vychádzajúc z bodu x^k sa hľadá minimum funkcie v smere d^k s použitím niektorej metódy minimalizácie funkcie *jednej* premennej.

Rieši sa teda úloha

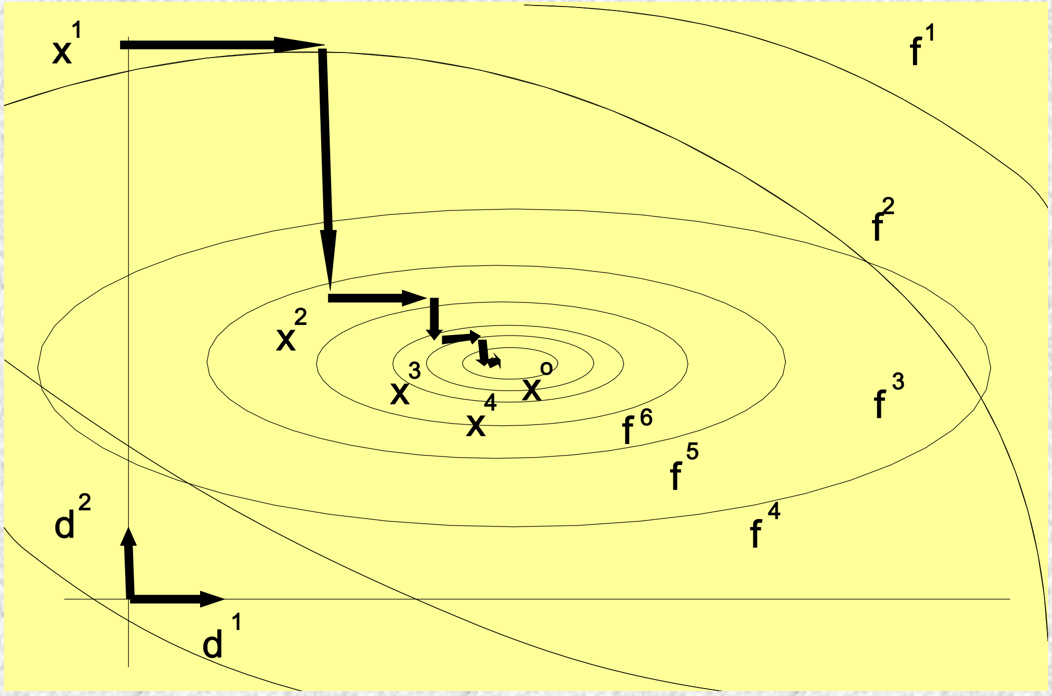
$$\min \{ f(x^k + \lambda d^k) \mid \lambda \in L \}$$

kde *L* je vopred definovaný interval neurčitosti. Vo formuláciách algoritmov sa predpokladá, že optimálny bod λ^o existuje. Avšak v reálnych úlohách tento predpoklad nemusí byť splnený. Optimálna hodnota účelovej funkcie v úlohe lineárneho hľadania môže byť totiž neohraničená, resp. aj keď konečná je, táto hodnota sa nedosahuje pri žiadnej hodnote λ. V prvom prípade je účelová funkcia východiskovej úlohy neohraničená a výpočet prerušujeme. V druhom prípade je možné vybrať takú hodnotu λ^o, že $f(x^k + \lambda^o d^k)$ bude dostatočne blízke k hodnote $\inf \{f(x^k + \lambda d^k) \mid \lambda \in L \}$.

Do tejto triedy metód patrí aj *metóda cyklickej posúradnicovej optimalizácie*, ktorú teraz podrobnejšie popíšeme.

Metóda cyklickej súradnicovej redukcie

V tejto metóde budeme ako smerové vektory pre hľadaný smer presunu od aktuálneho bodu trajektórie k nasledujúcemu používať *súradnicové vektory*.



Presnejšie, algoritmus realizuje hľadania pozdĺž smerov $d^1, d^2, \ldots, d^n$, kde d^j je smerový vektor, pre ktorého zložky platí:

$$d_i^j = 1, \text{ ak } i = j$$

$$d_i^j = 0, \text{ ak } i \neq j.$$

Týmto spôsobom sa pri hľadaní v smere vektora d^j mení iba súradnica x_j aktuálneho prípustného bodu a ostatné zložky ostávajú fixované. Schematicky je tento proces znázornený na obr. č.7.2.

Ďalej uvedieme algoritmus metódy cyklickej posúradnicovej optimalizácie, ktorý umožňuje nájsť minimum fukcie viacerých premenných bez použitia gradientov funkcie.

Čo sa týka konvergencie algoritmu, dá sa dokázať, že ak je funkcia *diferencovateľná,* tak algoritmus *konverguje k stacionárnemu bodu* funkcie.

V 6. kapitole sme uiedli niekoľko možných kritérií pre ukončenie algoritmov. V ďalej uvedenej metóde sa iteračný procs hľadania zastavuje, ak je splnená podmienka

$$\| x^{k+1} - x^k \| < \varepsilon$$

Možno však, pravdaže, použiť aj iné kritérium pre zastavenie procesu hľadania.

Algoritmus metódy cyklickej súradnicovej redukcie

Inicializačná fáza

- zvolíme číslo $\varepsilon > 0$, ktoré budeme používať pre test ukončenia algoritmu,
- ako smerové vektory zvolíme vektory súradnicových smerov

$d^1, d^2, \ldots, d^n$ so zložkami

$$d_i^j = \begin{cases} 0 & \text{ak } i \neq j, \qquad i, j = 1, 2, \ldots, n \\ 1 & \text{ak } i = j, \qquad i, j = 1, 2, \ldots, n \end{cases}$$

- zvolíme východiskový bod x^1
- položíme $y^1 = x^1$, $k = j = 1$ a prejdeme k výpočtovej fáze

Výpočtová fáza

Krok 1^o^.

- optimálnu veľkosť kroku λj získame riešením úlohy jednorozmernej minimalizácie

$$\min \{ f(y^j + \lambda d^j) \mid \lambda \in R^1 \}$$

- položíme

$$y^{j+1} = y^j + \lambda^j d^j$$

potom

a) ak $j < n$, tak zameníme $j \to j+1$ a vrátime sa na krok 1^o,

b) ak $j = n$, tak pokračujeme na krok 2^o.

Krok 2^o.

Položíme $x^{k+1} = y^{n+1}$.

a) ak platí $\| x^{k+1} - x^k \| < \varepsilon$, algoritmus končí,

b) v opačnom prípade

- položíme $y^1 = x^{k+1}$,
- položíme $j = 1$, zameníme $k \rightarrow k+1$,
- vrátime sa na krok 1^o algoritmu.

Príklad č.7.3 [Bazaraa]

Je daná nasledovná úloha dvojrozmernej minimalizácie:

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

kde $(x_1, x_2) \in R^2$.

Vychádzajúc z bodu $x^1 = (0,3)$ **určíme s použitím metódy cyklickej súradnicovej redukcie stacionárny bod funkcie.**

Riešenie:

Je zrejmé, že bodom globálneho minima funkcie je bod $x^o=(2,1)$.

Inicializačná fáza:

Ako rozlišovaciu konštantu ukončenia algoritmu zvolíme $\varepsilon = 0.05$.

Ako smerové vektory budeme používať vektory $d^1 = (1,0)^T$, $d^2 = (0,1)^T$.

Položíme $k = j = 1$ a $y^1 = x^1 = (0,3)^T$.

Výpočtová fáza:

Krok 1º.

Tento 1. krok v rámci 1. iterácie prepočítame podrobne.

Riešme úlohu jednorozmernej minimalizácie

$$\min \{ f(y^1 + \lambda d^1 \mid \lambda \in R \}$$

kde

$$f(y^1 + \lambda d^1) = f((0,3)^T+\lambda(1,0]^T) = f(\lambda,3) = (\lambda - 2)^4 + (\lambda - 6)^2$$

Pre riešenie tejto úlohy jednorozmernej minimalizácie sme použili metódu dichotomického hľadania s východiskovým intervalom $\langle a^1, b^2 \rangle = \langle 3,4 \rangle$ s rozlišovacou konštantou $2\varepsilon = 0.002$ a intervalom neurčitosti $l = 0{,}005$. Vypočty na jednotlivých iteráciách metódy dichotomického hľadania sú uvedené v tabuľke

k	a^k	b^k	λ^k	μ^k	$\theta(\lambda^k)$	$\theta(\mu^k)$
1	3.000	4.000	3.499	3.501	11.30401	11.32101
2	3.000	3.501	3.249	3.251	10.00275	10.00739
3	3.000	3.251	3.124	3.126	9.86744	9.86736
4	3.124	3.251	3.187	3.189	9.89829	9.90046
5	3.124	3.181	3.152	3.154	9.87244	9.87333
6	3.124	3.154	3.138	3.140	9.86827	9.86867
7	3.124	3.140	3.131	3.133	9.86744	9.86760
8	3.124	3.133	3.128	3.130	9.86734	9.86737
9	3.124	3.130	3.126	3.128	9.86737	9.86734
10	3.130	3.130	3.129	3.131	9.86735	9.86742

Vidíme, že po 10. iteráciách platí $|a^{10} - b^{10}| < l = 0.005$, algoritmus dichotomického hľadania končí a optimálna dĺžka kroku je potom λ^1=3.13.

Vypočítame

$$y^2 = y^1 + \lambda^1 d^1 = (0, 3) + 3.13(1, 0) = (3.13, 3)$$

Nakoľko *j* =1 < *n* =2, položíme *j* = *j* + *1* = *2* a vrátime sa na krok 1°.

Ďalej pokračujeme v riešení analogicky ako v demonštrovanom kroku algoritmu. Spolu sa vykonalo 7 iterácií, ktorých výsledky sú uvedené v tabuľke č.7.7.

k	x^k $f(x^k)$	j	d^j	y^j	λ^j	y^{j+1}
1	(0.00,3.00) 52.00	1 2	(1,0) (0,1)	(0.00,3.00) (3.13,3.00)	3.13 -1.44	(3.13,3.00) (3.13,1.56)
2	(3.13,1.56) 1.63	1 2	(1,0) (0,1)	(3.13,1.56) (2.63,1.56)	-0.50 -0.25	(2.63,1.56) (2.63,1.31)
3	(2.63,1.31) 0.16	1 2	(1,0) (0,1)	(2.63,1.31) (2.44,1.31)	-0.19 -0.09	(2.44,1.31) (2.44,1.22)
4	(2.44,1.22) 0.04	1 2	(1,0) (0,1)	(2.44,1.22) (2.35,1.22)	-0.09 -0.05	(2.35,1.22) 23353,1.17)
5	(2.35,1.17) 0.015	1 2	(1,0) (0,1)	(2.35,1.17) (2.29,1.17)	-0.06 -0.03	(2.29,1.17) (2.29,1.14)
6	(2.29,1.14) 0.007	1 2	(1,0) (0,1)	(2.29,1.14) (2.25,1.14)	-0.04 -0.02	(2.25,1.14) (2.25,1.12)
7	(2.25,1.12) 0.004	1 2	(1,0) (0,1)	(2.25,1.12) (2.22,1.12)	-0.03 -0.01	(2.22,1.12) (2.22,1.11)

Ďalej pokračujeme v riešení analogicky ako v demonštrovanom kroku algoritmu. Spolu sa vykonalo 7 iterácií, ktorých výsledky sú uvedené v tabuľke č.7.7.

Vidíme, že po 7. iterácii je splnená podmienka ukončenia algoritmu

$$\left\|\mathbf{x}^{k+1} - \mathbf{x}^{k}\right\| = \sqrt{(2.25 - 2.22)^2 + (1.12 - 1.11)^2} =$$

$$= \sqrt{0.03^2 + 0.01^2} = \sqrt{0.001} = 0.03 < 0.05$$

a teda bod $\mathbf{x}^8=(2.22,1.11)^T$ s hodnotou účelovej funkcie $f(\mathbf{x})=0.0016$ môžeme považovať za "dostatočne blízky" bodu globálneho minima fnkcie $\mathbf{x}^o=(2,1)$.

Gradientná metóda

Táto metóda je jednou z fundamentálnych procedúr pre riešenie úlohy minimalizácie diferencovateľnej funkcie *n* premenných. Metóda je v odbornej literatúre uvádzaná aj pod názvom metóda najrýchlejšieho spádu.

Vektor d budeme nazývať *progresívnym smerom* funkcie *f* v bode, ak existuje také $\delta > 0$, pre ktoré platí

$$f(\mathbf{x} + \lambda \mathbf{d}) < f(\mathbf{x}), \quad \text{pre všetky } \lambda \in (0,\delta)$$

Ak teda platí, že

$$\boxed{\lim_{\lambda \to 0^+} \frac{f(\mathbf{x} + \lambda \mathbf{d}) - f(\mathbf{x})}{\lambda} < 0}$$

potom vektor d reprezentuje progresívny smer. V metóde najrýchlejšieho spádu sa pohybujeme v smere d, pre ktorý $\| \mathbf{d} \| = 1$ a ktorý minimalizuje hodnotu vyššie uvedenej limity.

Dá sa ukázať (pozri napr. [Bazaraa, 1979]), že ak je funkcia *f*: $R^n \to R$ diferencovateľná v bode x a $\nabla f(x) \neq 0$, potom vektor

$$d^o = -\nabla f(x) / \| \nabla f(x) \|$$

je vektorom najrýchlejšieho spádu hodnoty funkcie *f* (x) v bode x.

Pri zadanom bode x teda metóda najrýchlejšieho spádu spočíva v realizácii jednorozmernej minimalizácie v smere $-\nabla f(x)/\| \nabla f(x) \|$ alebo, a to je vlastne to isté, pozdĺž vektora $-\nabla f(x)$.

Algoritmus gradientnej metódy

Inicializačná fáza

- zvolíme číslo $\varepsilon > 0$ ako konštantu pre kritérium ukončenia algoritmu,

- zvolíme východiskový bod x^1, položíme $k = 1$ a prejdeme k výpočtovej fáze

Výpočtová fáza

a) Ak $\|\nabla f(xk)\| < \varepsilon$, algoritmus končí.

b) V opačnom prípade:

- položíme $d^k = -\nabla f(x^k)$ a optimálnu veľkosť kroku λ^k nájdeme riešením úlohy minimalizácie jednej premennej

$$\min \{ f(x^k + \lambda d^k) \mid \lambda \geq 0 \}$$

- vypočítame nasledujúci bod trajektórie

$$x^{k+1} = x^k + \lambda^k d^k$$

- zameníme $k \to k+1$ a opakujeme výpočtovú fázu.

Príklad č.7.4

Skúmajme úlohu z

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

kde $(x_1, x_2)^T \in R^2$.

Budeme vychádzať z bodu $x^1=(0,3)T$ **a s použitím gradientnej metódy nájdeme stacionárny bod funkcie.**

Riešenie:

Inicializačná fáza:

Ako rozlišovaciu konštantu ukončenia algoritmu zvolíme $\varepsilon = 0.5$.
Položíme $k = 1$ **a** $x^1 = (0, 3)T$.

Výpočtová fáza:

Postup výpočtov v rámci prvej iterácie *k = 1* uvedieme podrobne.
Vypočítame gradient funkcie v bode x^1:

$$\nabla f(x) = (4(x_1-2)^3 + 2(x_1-2x_2), -4(x_1-2x_2))^T$$
$$\nabla f(0, 3) = (4(-2)3 + 2(-6), -4(-6))^T = (-44, 24)^T$$

Overme platnosť kritéria konca algoritmu:

$\| \nabla f(0,3) \| = \| (-44, 24) \| = 50,12 > 0,5 \Rightarrow$ **podmienka ukončenia algoritmu nie je splnená.**

Ďalej riešme úlohu jednorozmernej minimalizcie

$$\min \{ f(x^1 + \lambda d^1 \mid \lambda \in \langle 0, \infty) \}$$

kde

$$d^1 = -\nabla f(0, 3) = (44, -24)^T$$
$$f(x^1 + \lambda d^1) = f((0,3)^T+\lambda(44,-24)^T) = f(44\lambda, 3-24\lambda) = (44\lambda-2)^4 + (92\lambda-6)^2$$

Pre riešenie tejto úlohy jednorozmerného hľadania sme použili metódu dichotomického hľadania s východiskovým intervalom $\langle a^1, b^2\rangle = \langle 0,1\rangle$, s rozlišovacou konštantou $2\varepsilon = 0.002$ a intervalom neurčitosti L=0,005. Vypočty realizované na jednotlivých iteráciách metódy dichotomického hľadania sú uvedené v nasledujúcej tabuľke

k	a^k	b^k	λ^k	μ^k	$\theta(\lambda^k)$	$\theta(\mu^k)$
1	0.0000	1.0000	0.4990	0.5010	147905,1	163020,01
2	0.0000	0.5010	0.2495	0.2515	6784,52	7049,21
3	0.0000	0.2515	0.1248	0.1268	178,60	196,21
4	0.0000	0.1267	0.0624	0.0644	0,368	0,4566
5	0.0000	0.0643	0.0312	0.0332	10,07	8,87
6	0.0311	0.0643	0.0467	0.0487	2,902	2,309
7	0.0467	0.0643	0.0545	0.0565	1,08	0,76
8	0.0487	0.0643	0.0555	0.0575	2.3096	0.4799
9	0.0555	0.0643	0.0589	0.0609	0.8374	0.4799
10	0.0589	0.0643	0.0606	0.0626	0.4603	0.4799
11	0.0589	0.0626	0.0598	0.0618	0.4603	0.3819

Vidíme, že po 11. iteráciách platí $\| a^{11} - b^{11} \| < L = 0.005$, algoritmus dichotomického hľadania končí a optimálne dĺžka kroku je potom $\lambda^1 = 0.062$. Vypočítame

$$x^2 = x^1 + \lambda^1 d^1 = (0, 3)^T + 0.062\times(44, -24)^T = (2.70, 1.51)^T$$

Položíme *k = k + 1 = 2* a opakujeme *výpočtovú fázu*.

Metódy na riešenie úloh na voľný extrém

Ďalej pokračujeme v riešení analogicky ako v demonštrovanom kroku algoritmu. Spolu sa vykonalo 8 iterácií, ktorých výsledky sú uvedené v nasledujúcej tabuľke.

k	**x^k $f(x^k)$**	**$\nabla f(x^k)$**	**$\lvert\nabla f(x^k)\rvert$**	**$d^k = -\nabla f(x^k)$**	**λ^k**	**x^{k+1}**
1	(0.00,3.00) 52.00	(-44.0,24.00)	50.12	(44.0,-24.00)	0.062	(2.70,1.51)
2	(2.70,1.51) 0.34	(0.73,1.28)	1.47	(-0.73,-1.28)	0.24	(2.52,1.20)
3	(2.52,1.20) 0.09	(0.80,-0.48)	0.93	(-0.80,0.48)	0.11	(2.43,1.25)
4	(2.43,1.25) 0.04	(0.18,0.28)	0.33	(-0.18,-0.28)	0.31	(2.37,1.16)
5	(2.37,1.16) 52.00	(0.30,-0.20)	0.36	(-0.30,0.20)	0.12	(2.33,1.18)
6	(2.33,1.18) 0.01	(0.08,0.12)	0.14	(-0.08,-0.12)	0.36	(2.30,1.14)
7	(2.30,1.14) 0.009	(0.15,-0.08)	0.17	(-0.15,0.08)	0.13	(2.28,1.15)
8	(2.28,1.15) 0.007	(0.05,0.08)	0.09			

Vidíme, že po 8. iterácii je splnená podmienka ukončenia algoritmu

$$\|\nabla f(x^k)\| = 0.09 < \varepsilon$$

a teda bod x8 = $(2.28, 1.15)^T$ s hodnotou účelovej funkcie f(x) = 0.007 môžeme považovať za "dostatočne blízky" bodu globálneho minima fnkcie $x^o = (2,1)^T$.