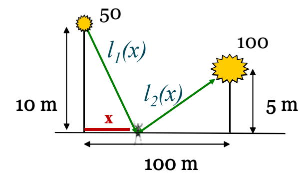
DISKRÉTNA OPTIMALIZÁCIA

Úlohy nelineárneho programovania

➢ **Nelineárne modely**

OPTIMALIZÁCIA NELINEÁRNYCH ÚLOH - PRÍKLAD



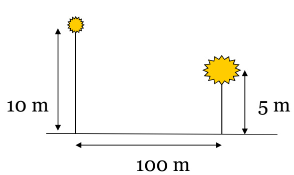
Intenzita osvetlenia

$$\frac{\text{Svietivosť}}{\text{Vzdialenosť}^2}$$

- **Intenzita osvetlenia** je priamo úmerná svietivosti zdroja a nepriamo úmerná druhej mocnine vzdialenosti od zdroja.
- **Intenzita osvetlenia v mieste, kde stojí panáčik:**
  - **Vysoká lampa:** $\frac{50}{(l_1(x))^2} = \frac{50}{\left(\sqrt{10^2+x^2}\right)^2} = \frac{50}{100+x^2}$
  - **Nízka lampa:** $\frac{100}{(l_2(x))^2} = \frac{100}{\left(\sqrt{5^2+(100-x)^2}\right)^2} = \frac{100}{25+(100-x)^2}$

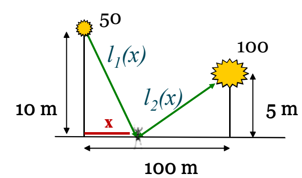
OPTIMALIZÁCIA NELINEÁRNYCH ÚLOH - PRÍKLAD

- Nájdite medzi dvomi lampami bod, ktorý je minimálne osvetlený. Lampy sú od seba vzdialené 100 m. Lampa 1 je na 10 m vysokom stožiari, lampa 2 je na 5 m vysokom stožiari. Svietivosť lampy 1 je 50 jednotiek a svietivosť lampy 2 je 100 jednotiek.
- Intenzita osvetlenia z rôznych zdrojov sa sčíta.



- Intenzita osvetlenia je priamo úmerná svietivosti zdroja a nepriamo úmerná druhej mocnine vzdialenosti od zdroja.

OPTIMALIZÁCIA NELINEÁRNYCH ÚLOH - PRÍKLAD



Intenzita osvetlenia

$$\frac{\text{Svietivosť}}{\text{Vzdialenosť}^2}$$

- **Intenzita osvetlenia v mieste, kde stojí panáčik:**
  - **Vysoká lampa:** $\frac{50}{100+x^2}$ **Nízka lampa:** $\frac{100}{25+(100-x)^2}$
- **Intenzita** osvetlenia z rôznych zdrojov **sa sčíta:**

$$\frac{50}{100+x^2} + \frac{100}{25+(100-x)^2}$$

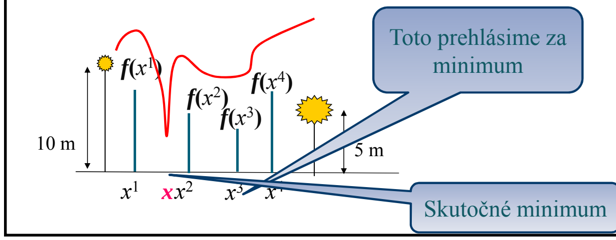
OPTIMALIZÁCIA NELINEÁRNYCH ÚLOH - PRÍKLAD

$$Min\ f(x) = \frac{50}{10^2 + x^2} + \frac{100}{5^2 + (100 - x)^2}$$
$$Subject\ to\ \ x \in <0,\ 100>$$

- Funkcia f(x) môže byť zadaná analyticky.
- Funkcia f(x) môže byť výsledok programu so vstupom x.
- Funkcia f(x) môže byť výsledok časovo náročného merania.

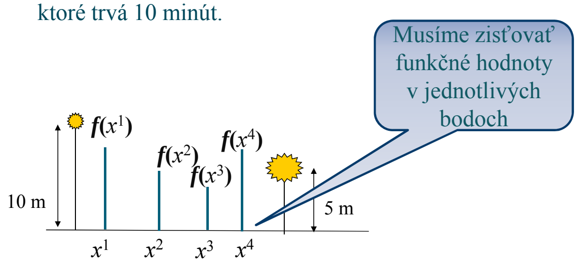
PRÍSTUPY K OPTIMALIZÁCII: NEMÁME INFORMÁCIU O FUNKCII

- Nemáme žiadnu informáciu o vlastnostiach minimalizovanej funkcie f(x) na danom intervale <a, b>, preto sa môžeme mýliť: skutočné minimum môže byť v bode *x*.



PRÍSTUPY K OPTIMALIZÁCII: NEMÁME INFORMÁCIU O FUNKCII

- Nájdite medzi dvomi lampami bod, ktorý je minimálne osvetlený. Lampy sú od seba vzdialené 100 m.
- Intenzitu osvetlenia v danom **mieste získate meraním**, ktoré trvá 10 minút.



PRÍSTUPY K OPTIMALIZÁCII: MÁME INFORMÁCIU O FUNKCII

- Záleží na tom, akú máme informáciu o vlastnostiach minimalizovanej funkcie f(x) na danom intervale <a, b>.
- Máme analytický zápis funkcie f(x):

$$Min\ f(x) = \frac{50}{10^2 + x^2} + \frac{100}{5^2 + (100 - x)^2}$$
$$Subject\ to\ \ x \in <0,\ 100>$$

- Vlastnosti tejto funkcie:
  - je spojitá na intervale <0, 100>,
  - na intervale <0, 100> má spojité derivácie,

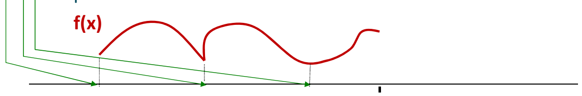
HĽADANIE MINIMA ANALYTICKY ZADANEJ FCIE JEDNEJ PREMENNEJ

$$Min\ f(x) = \frac{50}{10^2 + x^2} + \frac{100}{5^2 + (100 - x)^2}$$

$$Subject\ \ to\ \ x \in <0,\ 100>$$

- Ak je funkcia spojitá na intervale <0, 100> a má skoro všade aj spojité derivácie, potom sa minimum hľadá:

a) v krajných bodoch intervalu,
b) v bodoch, kde neexistuje prvá derivácia,
c) v bodoch, kde je prvá derivácia nulová (stacionárne body). V stacionárnom bode má funkcia minimum, ak je najnižšia párna derivácia kladná.



HĽADANIE MINIMA ANALYTICKY ZADANEJ FCIE JEDNEJ PREMENNEJ

$$f(x) = \frac{50}{100 + x^2} + \frac{100}{25 + (100 - x)^2}$$

Funkčná hodnota v krajných bodoch intervalu

- *f(0)*= 0,509975
- *f(100)*=4,00495

$$f'(x) = \frac{-50*2*x}{(100 + x^2)^2} + \frac{100*2*(100 - x)}{(25 + (100 - x)^2)^2}$$

HĽADANIE MINIMA ANALYTICKY ZADANEJ FCIE JEDNEJ PREMENNEJ

- Záleží na tom, akú máme informáciu o vlastnostiach minimalizovanej funkcie f(x) na danom intervale <a, b>.
- Máme analytický zápis funkcie f(x):

$$Min\ f(x) = \frac{50}{10^2 + x^2} + \frac{100}{5^2 + (100 - x)^2}$$

$$Subject\ \ to\ \ x \in <0,\ 100>$$

- Vlastnosti tejto funkcie:
  - je spojitá na intervale <0, 100>,
  - na intervale <0, 100> má spojité derivácie,

HĽADANIE MINIMA ANALYTICKY ZADANEJ FCIE JEDNEJ PREMENNEJ

$$f(x) = \frac{50}{100 + x^2} + \frac{100}{25 + (100 - x)^2}$$

Hľadanie stacionárneho bodu môže byť numericky náročné:

$$f'(x) = \frac{-50*2*x}{(100 + x^2)^2} + \frac{100*2*(100 - x)}{(25 + (100 - x)^2)^2} = 0$$

$$\frac{100*2*(100 - x)}{(25 + (100 - x)^2)^2} = \frac{50*2*x}{(100 + x^2)^2}$$

$$2(100 - x)(100 + x^2)^2 = x(25 + (100 - x)^2)^2$$

HĽADANIE MINIMA ANALYTICKY ZADANEJ FCIE JEDNEJ PREMENNEJ

$$f(x) = \frac{50}{100 + x^2} + \frac{100}{25 + (100 - x)^2}$$

$$f'(x) = \frac{-50*2*x}{(100 + x^2)^2} + \frac{100*2*(100 - x)}{(25 + (100 - x)^2)^2}$$

$$f''(x) = 200\frac{x^2 - 100}{(100 + x^2)^3} + 200\frac{3(100 - x)^2 - 25}{(25 + (100 - x)^2)^3}$$

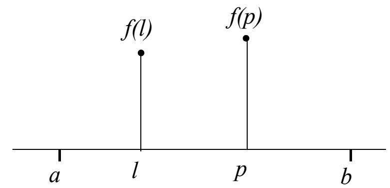
- **Druhá derivácia** je v okolí x=0 záporná, teda ostrá konvexnosť na intervale <0, 100> nie je zaistená.
- Teda bolo by potrebné vyšetriť všetky stacionárne body.

METÓDA POSTUPNÉHO VYHĽADÁVANIA MINIMA OSTRO KVÁZIKONVEXNEJ FUNKCIE $f(x)$ NA KONEČNOM INTERVALE

A. Sz.1

Predpoklady: Ostrá kvázikonvexnosť funkcie a uzavretý interval **<*a, b* >.**

Princíp: Vo vnútri intervalu vyberieme dva body *l* a *p* a určíme funkčné hodnoty *f*(*l*) a *f*(*p*).



INÁ MOŽNÁ VLASTNOSŤ FUNKCIE $f(x)$

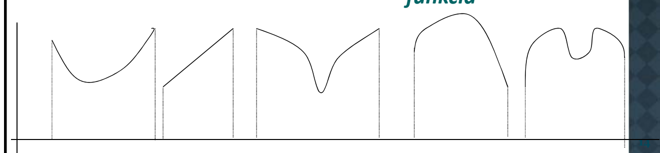
Ostrá kvázikonvexnosť funkcie

Vravíme, že funkcia *f(x)* je ostro kvázikonvexná na konvexnom definičnom obore $X \subseteq E^1$, ak pre každé $x, y \in X$ a pre každé $\alpha \in (0,1)$ platí

$$f(\alpha x + (1-\alpha) y) \leq \max\{f(x), f(y)\}.$$

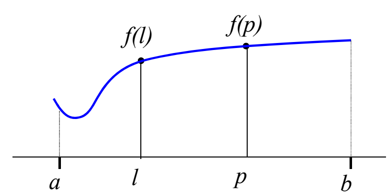
Kvázikonvexná funkcia

Nekvázikonvexná funkcia



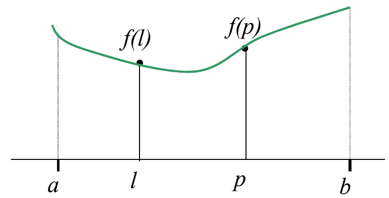
METÓDA POSTUPNÉHO VYHĽADÁVANIA MINIMA OSTRO KVÁZIKONVEXNEJ FUNKCIE $f(x)$ NA KONEČNOM INTERVALE

Ako **môže** vyzerať kvázikonvexná funkcia *f* za daných hodnôt *f*(*l*) a *f*(*p*) ?



METÓDA POSTUPNÉHO VYHĽADÁVANIA MINIMA OSTRO KVÁZIKONVEXNEJ FUNKCIE $f(x)$ NA KONEČNOM INTERVALE

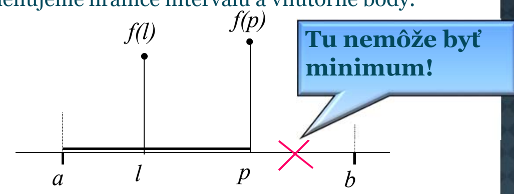
Ako **môže** vyzerať kvázikonvexná funkcia *f* za daných hodnôt $f(l)$ a $f(p)$?



METÓDA POSTUPNÉHO VYHĽADÁVANIA MINIMA OSTRO KVÁZIKONVEXNEJ FUNKCIE $f(x)$ NA KONEČNOM INTERVALE

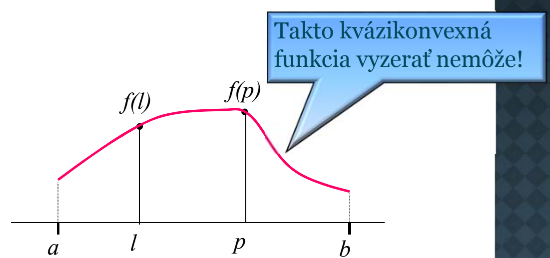
Princíp:

1. Vo vnútri intervalu vyberieme dva body l a p a určíme funkčné hodnoty $f(l)$ a $f(p)$.
2. Ak je $f(l) \leq f(p)$, môžeme vylúčiť interval $<p, b>$ a minimum ďalej hľadať v intervale $<a, p>$.
3. Premenujeme hranice intervalu a vnútorné body.



METÓDA POSTUPNÉHO VYHĽADÁVANIA MINIMA OSTRO KVÁZIKONVEXNEJ FUNKCIE $f(x)$ NA KONEČNOM INTERVALE

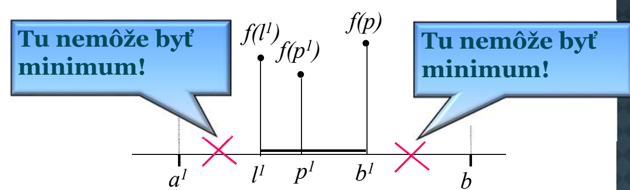
Ako **nemôže** vyzerať kvázikonvexná funkcia *f* za daných hodnôt $f(l)$ a $f(p)$?



METÓDA POSTUPNÉHO VYHĽADÁVANIA MINIMA OSTRO KVÁZIKONVEXNEJ FUNKCIE $f(x)$ NA KONEČNOM INTERVALE

Princíp:

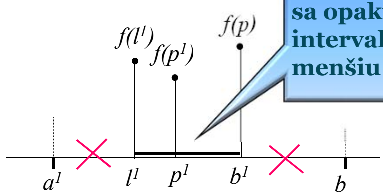
1. Ďalej vo vnútri intervalu $<a^1, b^1>$ vyberieme dva body l^1 a p^1, resp. doplníme chybujúci a určíme $f(l^1)$ a $f(p^1)$.
2. Ak napr. $f(l^1) \geq f(p^1)$, môžeme vylúčiť interval $< a^1, l^1 >$ a minimum ďalej hľadať v intervale $< l^1, b^1 >$.



METÓDA POSTUPNÉHO VYHĽADÁVANIA MINIMA OSTRO KVÁZIKONVEXNEJ FUNKCIE $f(x)$ NA KONEČNOM INTERVALE

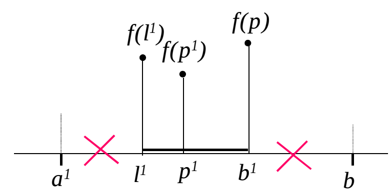
Postup končí, ak dĺžka intervalu, ktorý musí obsahovať bod minima, zaručuje, že bod vnútri intervalu x^i, pre ktorý sme vypočítali hodnotu $f(x^i)$, nie je od skutočného minima vzdialený o viac než vopred zvolené α.

Redukcia intervalu sa opakuje pokiaľ interval nemá dĺžku menšiu než α.



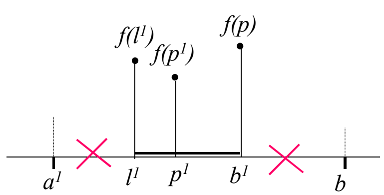
METÓDA ZLATÉHO REZU PRE VYHĽADÁVANIE MINIMA OSTRO KVÁZIKONVEXNEJ FUNKCIE $f(x)$ NA INTERVALE <a^1, b^1>

- Body l^i a p^i sa vyberajú tak, aby delili interval <a^i, b^i> v pomere **zlatého rezu**.
- **Zlatý rez**: Pomer kratšej časti intervalu k dlhšej musí byť rovnaký ako pomer dlhšej časti intervalu k jeho dĺžke.



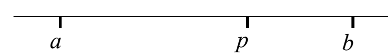
METÓDA POSTUPNÉHO VYHĽADÁVANIA MINIMA OSTRO KVÁZIKONVEXNEJ FUNKCIE $f(x)$ NA KONEČNOM INTERVALE

Podľa spôsobu, akým sa vyberajú zo súčasného intervalu body l_i a p_i rozlišujeme, rôzne metódy postupného vyhľadávania.



ZLATÝ REZ

Zlatý rez: Pomer kratšej časti intervalu k dlhšej musí byť rovnaký ako pomer dlhšej časti intervalu k jeho dĺžke.



$$\frac{b-a}{p-a}=\frac{p-a}{b-p}$$

$$(b-a)b-(b-a)p=(p-a)^2 \;\Rightarrow$$

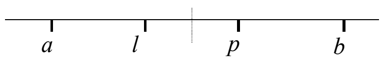
$$(b-a)b-(b-a)a+(b-a)a-(b-a)p=(p-a)^2 \Rightarrow$$

$$(p-a)^2+(b-a)(p-a)-(b-a)^2=0 \Rightarrow$$

$$(p-a)=\frac{-(b-a)\pm\sqrt{(b-a)^2+4(b-a)^2}}{2}=(b-a)\frac{-1\pm\sqrt{5}}{2}$$

ZLATÝ REZ

Zlatý rez: Pomer kratšej časti intervalu k dlhšej musí byť rovnaký ako pomer dlhšej časti intervalu k jeho dĺžke.



$$(p-a)=(b-a)\frac{-1+\sqrt{5}}{2} \Rightarrow \; p=a+(b-a)\frac{-1+\sqrt{5}}{2}$$

$$\frac{a+b}{2}-l=p-\frac{a+b}{2} \Rightarrow l=\frac{a+b}{2}+\frac{a+b}{2}-p=a+b-p$$

$$l=a+b-a-(b-a)\frac{-1+\sqrt{5}}{2}=a+(b-a)(\frac{3-\sqrt{5}}{2})$$

URČENIE POTREBNÉHO POČTU KROKOV

Stačí teda určiť najmenšie celé *i* tak, aby platilo:

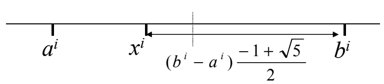
$$(b-a)\left(\frac{2}{1+\sqrt{5}}\right)^{i+1} \le \alpha \;\Rightarrow\; \left(\frac{2}{1+\sqrt{5}}\right)^{i+1} \le \frac{\alpha}{b-a} \Rightarrow$$

$$\log(\left(\frac{2}{1+\sqrt{5}}\right)^{i+1}) \le \log(\frac{\alpha}{b-a}) \Rightarrow (i+1)\log(\frac{2}{1+\sqrt{5}}) \le \log(\frac{\alpha}{b-a})$$

$$i \ge -1+\frac{\log(\frac{\alpha}{b-a})}{\log(\frac{2}{1+\sqrt{5}})}$$

URČENIE POTREBNÉHO POČTU KROKOV

Požadujeme, aby výsledný bod x^i, ktorý získame ako odhad bodu minima po *i* redukciách (*i*+1 vypočítaných hodnotách *f*), bol od bodu minima vzdialený najviac α.



Pri každom kroku sa interval zmenší (-1+√5)/2 krát. Teda:

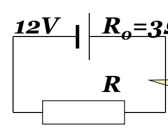
$$(b^i-a^i)=(b-a)\left(\frac{-1+\sqrt{5}}{2}\right)^{i} \;\Rightarrow\; (b-a)\left(\frac{-1+\sqrt{5}}{2}\right)^{i+1} \le \alpha$$

$$resp.\; (b-a)\left(\frac{2}{1+\sqrt{5}}\right)^{i+1} \le \alpha$$

PRÍKLAD

S presnosťou 0.01 mm určte optimálnu šírku 3 m dlhého drôtika rozmrazovača skla tak, aby ste dosiahli maximálny výkon daného zariadenia.

Vlastné zariadenie (viď schéma) je napájané 12 V autobatériou s vnútorným odporom 3Ω, drôtik obdĺžnikového prierezu má výšku 0.05mm a jeho materiál má merný odpor ρ=5*10^{-7} [Ωm].



Odpor drôtika sa vypočíta z merného odporu podľa vzorca R=ρd/S. kde d je dĺžka a S je plocha prierezu.

PRÍSTUPY K OPTIMALIZÁCII NELINEÁRNYCH ÚLOH VIAC PREMENNÝCH PRÍKLAD

- Navrhnite čo najvýnosnejší výrobný program firme vyrábajúcej dva produkty P1 a P2, ak na výrobu každej jednotky P1 je potrebujeme dve a na každé jednotky P2 tri jednotky výrobnej kapacity o veľkosti 24.
- Predajná cena jednotky P1 a P2 je 6 a 8. Výrobné **jednotkové náklady** sú u P1 0.2x_1 a u P2 0.4x_2, kde x_1 a x_2 sú množstvá vyrábaných produktov.

PRÍSTUPY K OPTIMALIZÁCII NELINEÁRNYCH ÚLOH VIAC PREMENNÝCH: PRÍKLAD

- Predajná cena jednotky P1 a P2 je 6 a 8. Výrobné **jednotkové náklady** sú u P1 0.2x_1 a u P2 0.4x_2, kde x_1 a x_2 sú množstvá vyrábaných produktov.

$$Min \quad 6x_1 + 8x_2 - (0.2x_1)x_1 - (0.4x_2)x_2 =$$

$$= -0.2x_1^2 - 0.4x_2^2 + 6x_1 + 8x_2$$

$$Subject \;\; to \;\; 2x_1 + 3x_2 \le 24$$

$$x_1,\; x_2 \ge 0$$

PRÍSTUPY K OPTIMALIZÁCII NELINEÁRNYCH ÚLOH VIAC PREMENNÝCH PRÍKLAD

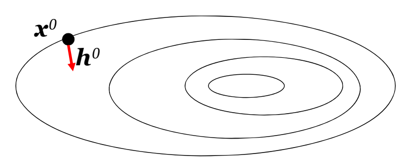
- Navrhnite čo najvýnosnejší výrobný program firme vyrábajúcej dva produkty P1 a P2, ak na výrobu každej jednotky P1 je potrebujeme dve a na každé jednotky P2 tri jednotky výrobnej kapacity o veľkosti 24.

$$Subject \;\; to \;\; 2x_1 + 3x_2 \le 24$$

$$x_1,\; x_2 \ge 0$$

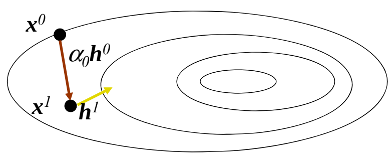
ITERAČNÉ METÓDY PRE RIEŠENIA NELINEÁRNYCH ÚLOH VIAC PREMENNÝCH

- Začínajú prácu vždy v nejakom prípustnom riešení $\boldsymbol{x}^0$, bodu z E^n.
- Nájdu smer $\boldsymbol{h}^i$, v ktorom od súčasného riešenia $\boldsymbol{x}^i$ funkcie f klesá.
- Vykonajú presun k ďalšiemu riešeniu podľa výrazu $\boldsymbol{x}^{i+1}=\boldsymbol{x}^i+\alpha_i\boldsymbol{h}^i$, kde α kladné reálne číslo.



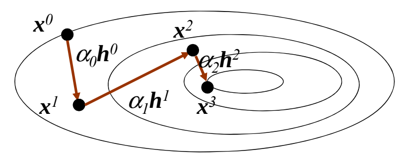
ITERAČNÉ METÓDY PRE RIEŠENIA NELINEÁRNYCH ÚLOH VIAC PREMENNÝCH

- Začínajú prácu vždy v nejakom prípustnom riešení $\boldsymbol{x}^0$, bodu z E^n.
- Nájdu smer $\boldsymbol{h}^i$, v ktorom od súčasného riešenia $\boldsymbol{x}^i$ funkcia f klesá.
- Vykonajú presun k ďalšiemu riešeniu podľa výrazu $\boldsymbol{x}^{i+1}=\boldsymbol{x}^i+\alpha_i\boldsymbol{h}^i$, kde α kladné reálne číslo.



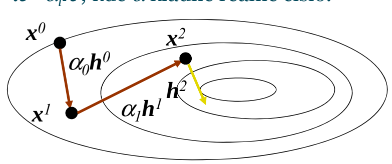
ITERAČNÉ METÓDY PRE RIEŠENIA NELINEÁRNYCH ÚLOH VIAC PREMENNÝCH

- Začínajú prácu vždy v nejakom prípustnom riešení $\boldsymbol{x}^0$, bodu z E^n.
- Nájdu smer $\boldsymbol{h}^i$, v ktorom od súčasného riešenia $\boldsymbol{x}^i$ funkcia f klesá.
- Vykonajú presun k ďalšiemu riešeniu podľa výrazu $\boldsymbol{x}^{i+1}=\boldsymbol{x}^i+\alpha_i\boldsymbol{h}^i$, kde α kladné reálne číslo.



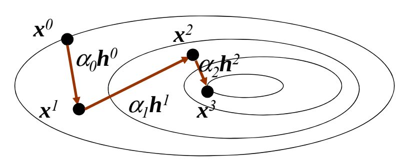
ITERAČNÉ METÓDY PRE RIEŠENIA NELINEÁRNYCH ÚLOH VIAC PREMENNÝCH

- Začínajú prácu vždy v nejakom prípustnom riešení $\boldsymbol{x}^0$, bodu z E^n.
- Nájdu smer $\boldsymbol{h}^i$, v ktorom od súčasného riešenia $\boldsymbol{x}^i$ funkcia f klesá.
- Vykonajú presun k ďalšiemu riešeniu podľa výrazu $\boldsymbol{x}^{i+1}=\boldsymbol{x}^i+\alpha_i\boldsymbol{h}^i$, kde α kladné reálne číslo.



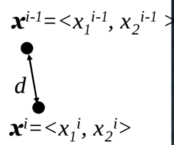
ITERAČNÉ METÓDY PRE RIEŠENIA NELINEÁRNYCH ÚLOH VIAC PREMENNÝCH

- Iteračné metódy vytvárajú tzv. **minimalizujúcu postupnosť** $\boldsymbol{x}^0, \boldsymbol{x}^1, \boldsymbol{x}^2, \boldsymbol{x}^3$ ….., kde $f(\boldsymbol{x}^0)> f(\boldsymbol{x}^1)> f(\boldsymbol{x}^2)>f(\boldsymbol{x}^3)> \dots$
- Otázka je: **„Kedy iteračné metódy končia?“**



PRAVIDLÁ ZAKONČENIA ITERAČNÝCH METÓD

„Kedy iteračné metódy končia?“



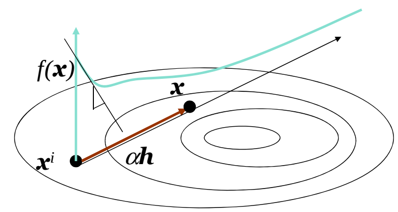
1. Po zadanom počte krokov.
2. Ak majú po sebe idúce riešenia x^{i-1} a x^i medzi sebou vzdialenosť menšiu než vopred zvolené ε.

$$d=\sqrt{(x_1^{i-1}-x_1^i)^2+(x_2^{i-1}-x_2^i)^2}$$

3. Ak sa líšia funkčné hodnoty po sebe idúcich riešení x^{i-1} a x^i o menej než vopred zvolené ε.
   T.j. $|\, f(x^{i-1}) - f(x^i)\,| < \varepsilon$.
4. Ak je splnené kombinované pravidlo

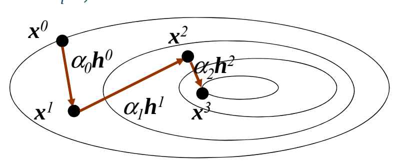
ITERAČNÉ METÓDY PRE RIEŠENIE NELI-NEÁRNYCH ÚLOH VIAC PREMENNÝCH

- Podľa spôsobu, akým sa určuje smer $\boldsymbol{h}$ a koeficient α, rozlišujeme rôzne iteračné metódy.
- Súvisí s tým otázka: „**V ktorom smere v blízkom okolí bodu x^i klesá hodnota účelovej funkcie najrýchlejšie**?“



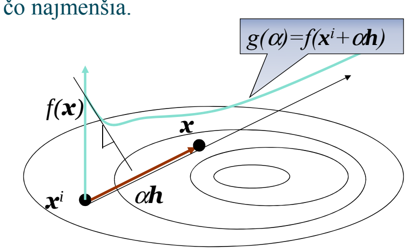
ITERAČNÉ METÓDY PRE RIEŠENIA NELINEÁRNYCH ÚLOH VIAC PREMENNÝCH

- Začínajú prácu vždy v nejakom prípustnom riešení $\boldsymbol{x}^0$, bodu z E^n.
- Nájdu smer $\boldsymbol{h}^i$, v ktorom od súčasného riešenia $\boldsymbol{x}^i$ funkcia f klesá.
- Vykonajú presun k ďalšiemu riešeniu podľa výrazu $\boldsymbol{x}^{i+1}=\boldsymbol{x}^i+\alpha_i\boldsymbol{h}^i$, kde α kladné reálne číslo.



SMER NAJVÄČŠIEHO POKLESU FUNKCIE V OKOLÍ BODU x^i

- **„V ktorom smere v blízkom okolí bodu x^i klesá hodnota účelovej funkcie najrýchlejšie?“**
- Hľadáme smer $\boldsymbol{h}$ taký, aby derivácia funkcie $g(\alpha)$ v $\alpha=0$ bola čo najmenšia.



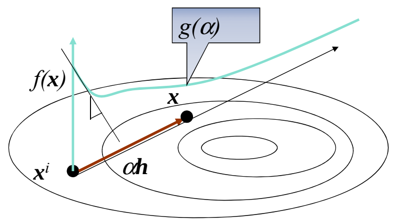
SMER NAJVÄČŠIEHO POKLESU

$$\frac{dg(\alpha)}{d\alpha} = \sum_{j=1}^{n} \frac{\partial f(\mathbf{x}^i + \alpha \mathbf{h})}{\partial x_j} h_j \quad pro \quad \alpha = 0 \quad dos\ tan\ eme$$

$$\frac{dg(0)}{d\alpha} = \sum_{j=1}^{n} \frac{\partial f(\mathbf{x}^i)}{\partial x_j} h_j = f'(\mathbf{x}).\mathbf{h} = |f'(\mathbf{x})|.|\mathbf{h}|.\cos\ \omega$$

$$g(\alpha) = f(\mathbf{x}^i + \alpha \mathbf{h})$$

Hľadáme smer ***h*** taký, aby derivácia funkcie $g(\alpha)$ v α=0 bola čo najmenšia.

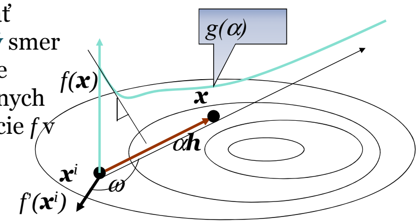


SMER NAJVÄČŠIEHO POKLESU

$$\frac{dg(0)}{d\alpha} = \sum_{j=1}^{n} \frac{\partial f(\mathbf{x}^i)}{\partial x_j} h_j = f'(\mathbf{x}^i).\mathbf{h} = |f'(\mathbf{x}^i)|.|\mathbf{h}|.\cos\ \omega$$

Pre ***h*** konštantnej dĺžky bude mať výraz najmenšiu hodnotu pre ω=π, keď cos(π)= -1.

Teda ***h*** má mať práve **opačný** smer než *f'*($\mathbf{x}^i$), čo je vektor parciálnych derivácii funkcie *f* v bode $\mathbf{x}^i$.



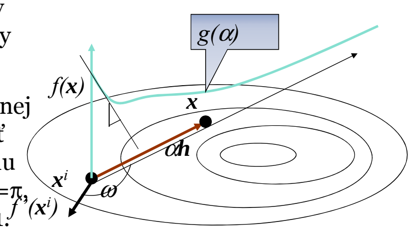
SMER NAJVÄČŠIEHO POKLESU

$$\frac{dg(\alpha)}{d\alpha} = \sum_{j=1}^{n} \frac{\partial f(\mathbf{x}^i + \alpha \mathbf{h})}{\partial x_j} h_j \quad pro \quad \alpha = 0 \quad dos\ tan\ eme$$

$$\frac{dg(0)}{d\alpha} = \sum_{j=1}^{n} \frac{\partial f(\mathbf{x}^i)}{\partial x_j} h_j = f'(\mathbf{x}).\mathbf{h} = |f'(\mathbf{x})|.|\mathbf{h}|.\cos\ \omega$$

ω je uhol, ktorý zvierajú vektory *f'*($\mathbf{x}^i$) a ***h***.

Pre ***h*** konštantnej dĺžky bude mať výraz najmenšiu hodnotu pre ω=π, kedy cos(π)= -1.

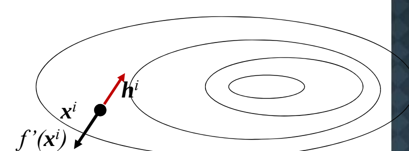


SMER NAJVÄČŠIEHO POKLESU FUNKCIE V OKOLÍ BODU x^i

- Hľadáme smer ***h*** taký, aby derivácia fcie $g(\alpha) = f(\mathbf{x}^i + \alpha \mathbf{h})$ v α =0 bola čo najmenšia.
- Nech ***f'*** ($\mathbf{x}^i$) je **gradient funkcie *f*** v bode $\mathbf{x}^i$, čo je vektor parciálnych derivácii **funkcie *f*** v bode $\mathbf{x}^i$.

$$f'(\mathbf{x}^i) = \begin{pmatrix} \frac{\partial f(\mathbf{x}^i)}{\partial x_1} \\ \frac{\partial f(\mathbf{x}^i)}{\partial x_2} \\ \vdots \\ \frac{\partial f(\mathbf{x}^i)}{\partial x_n} \end{pmatrix}$$

- Vektor $\mathbf{h}^i$ má mať práve **opačný** smer než *f'*($\mathbf{x}^i$)



PRÍKLAD: GRADIENTOVÁ METÓDA S KONŠTANTNÝM KROKOM

Minimalizujte $f(\mathbf{x}) = (x_1)^2+4(x_2)^2-4x_1-8x_2$ pre $\mathbf{x}\in E^2$
Zvoľme $\mathbf{x}^0=<0, 2>^T$, α=0,5.

Postup: Určíme gradient $f'(\mathbf{x}^0)$ a smer $\mathbf{h}^0=-f'(\mathbf{x}^0)$.

$$f'(\mathbf{x}^i) = \begin{pmatrix} \frac{\partial f(\mathbf{x}^i)}{\partial x_1} \\ \frac{\partial f(\mathbf{x}^i)}{\partial x_2} \\ \vdots \\ \frac{\partial f(\mathbf{x}^i)}{\partial x_n} \end{pmatrix}$$

$$\frac{\partial f(\mathbf{x}^0)}{\partial x_1} = 2x_1 - 4$$

$$\frac{\partial f(\mathbf{x}^0)}{\partial x_2} = 8x_2 - 8$$

$$f'(\begin{bmatrix}0\\2\end{bmatrix}) = \begin{bmatrix}-4\\8\end{bmatrix}, \quad \mathbf{h}^0 = \begin{bmatrix}4\\-8\end{bmatrix},$$

$$dlzka\ |\mathbf{h}^0| = \sqrt{(4)^2+(-8)^2} = \sqrt{80} = 8{,}94$$

PRÍKLAD: GRADIENTOVÁ METÓDA S KONŠTANTNÝM KROKOM

Overíme, či funkcia f má v bode $\mathbf{x}$ menšiu HÚF než v $\mathbf{x}^0$:

$\mathbf{x}^0=<0, 2>^T$, $\mathbf{x}=<0.22, 1.55>^T$.

Minimalizujte $f(\mathbf{x}) = (x_1)^2+4(x_2)^2-4x_1-8x_2$ pre $\mathbf{x}\in E^2$

$$f(\mathbf{x}^0) = f(\begin{bmatrix}0\\2\end{bmatrix}) = (0)^2+4(2)^2-4*0-8*2 = 0$$

$$f(\mathbf{x}) = f(\begin{bmatrix}0.22\\1.55\end{bmatrix}) = (0.22)^2+4(1.55)^2-4*0.22-8*1.55 = -3.62$$

Pokiaľ áno, definujeme $\mathbf{x}^1=\mathbf{x}$, inak postup opakujeme so zmenšeným krokom α.
V tomto prípade $\mathbf{x}^1=< 0.22, 1.55 >^T$.

PRÍKLAD: GRADIENTOVÁ METÓDA S KONŠTANTNÝM KROKOM

Minimalizujte $f(\mathbf{x}) = (x_1)^2+4(x_2)^2-4x_1-8x_2$ pre $\mathbf{x}\in E^2$
Zvoľme $\mathbf{x}^0=<0, 2>^T$, α=0,5.

Postup: Určíme gradient $f'(\mathbf{x}^0)$ a smer $\mathbf{h}^0=-f'(\mathbf{x}^0)$.

$$f'(\begin{bmatrix}0\\2\end{bmatrix}) = \begin{bmatrix}-4\\8\end{bmatrix}, \quad \mathbf{h}^0 = \begin{bmatrix}4\\-8\end{bmatrix}, \quad dlzka\ |\mathbf{h}^0| = \sqrt{(4)^2+(-8)^2} = \sqrt{80} = 8{,}94$$

Vykonáme presun v smere $\mathbf{h}^0$ o veľkosti kroku α=0.5

$$\mathbf{x} = \mathbf{x}^0 + \alpha(\frac{1}{|\mathbf{h}^0|})\mathbf{h}^0 = \begin{bmatrix}0\\2\end{bmatrix} + 0.5*(\frac{1}{8.94})\begin{bmatrix}4\\-8\end{bmatrix} = \begin{bmatrix}0\\2\end{bmatrix} + \begin{bmatrix}0.22\\-0.45\end{bmatrix} = \begin{bmatrix}0.22\\1.55\end{bmatrix}$$

PRÍKLAD: GRADIENTOVÁ METÓDA S KONŠTANTNÝM KROKOM

$\mathbf{x}^0=<0, 2>^T$, $\mathbf{x}^1=<0.22, 1.55>^T$, α=0,5.

Pokračujeme: Určíme gradient $f'(\mathbf{x}^1)$ a smer $\mathbf{h}^1=-f'(\mathbf{x}^1)$.

$$f'(\begin{bmatrix}0.22\\1.55\end{bmatrix}) = \begin{bmatrix}-3.56\\4.40\end{bmatrix}, \quad \mathbf{h}^1 = \begin{bmatrix}3.56\\-4.40\end{bmatrix}, \quad |\mathbf{h}^1| = \sqrt{(3.56)^2+(-4.40)^2} = 5.66$$

Vykonáme presun v smere $\mathbf{h}^1$ o veľkosti kroku α=0.5

$$\mathbf{x} = \mathbf{x}^1 + \alpha(\frac{1}{|\mathbf{h}^1|})\mathbf{h}^1 = \begin{bmatrix}0.22\\1.55\end{bmatrix} + 0.5*(\frac{1}{5.66})\begin{bmatrix}3.56\\-4.4\end{bmatrix} = \begin{bmatrix}0.22\\1.55\end{bmatrix} + \begin{bmatrix}0.31\\-0.39\end{bmatrix} = \begin{bmatrix}0.53\\1.16\end{bmatrix}$$

PRÍKLAD: GRADIENTOVÁ METÓDA S KONŠTANTNÝM KROKOM

Overíme, či funkcia f má v bode $\boldsymbol{x}$ menšiu HÚF než v $\boldsymbol{x}^1$:

$\boldsymbol{x}^1$=<0.22, 1.55>^{T}, $\boldsymbol{x}$=<0.53, 1.16>^{T}.

Minimalizujte $f(\boldsymbol{x})= (x_1)^2+4(x_2)^2-4x_1-8x_2$ pre $\boldsymbol{x}\in E^2$

$$f(\mathbf{x}) = f(\begin{bmatrix}0.22\\1.55\end{bmatrix}) = (0.22)^2 + 4(1.55)^2 - 4*0.22 - 8*1.55 = -3.62$$

$$f(\mathbf{x}) = f(\begin{bmatrix}0.53\\1.16\end{bmatrix}) = (0.53)^2 + 4(1.16)^2 - 4*0.53 - 8*1.16 = -5.74$$

Pokiaľ áno, definujeme $\boldsymbol{x}^2=\boldsymbol{x}$, inak postup opakujeme so zmenšeným krokom α.

V tomto prípade $\boldsymbol{x}^2$=< 0.53, 1.16 >^{T}.

PRÍKLAD: GRADIENTOVÁ METÓDA S KONŠTANTNÝM KROKOM

Overíme, či funkcia f má v bode $\boldsymbol{x}$ menšiu HÚF než v $\boldsymbol{x}^2$:

$\boldsymbol{x}^2$=<0.53, 1.16>, $\boldsymbol{x}$=<0.99, 0.96>^{T}.

Minimalizujte $f(\boldsymbol{x})= (x_1)^2+4(x_2)^2-4x_1-8x_2$ pre $\boldsymbol{x}\in E^2$

$$f(\mathbf{x}) = f(\begin{bmatrix}0.53\\1.16\end{bmatrix}) = (0.53)^2 + 4(1.16)^2 - 4*0.53 - 8*1.16 = -5.74$$

$$f(\mathbf{x}) = f(\begin{bmatrix}0.99\\0.96\end{bmatrix}) = (0.99)^2 + 4(0.96)^2 - 4*0.99 - 8*0.96 = -6.97$$

Pokiaľ áno, definujeme $\boldsymbol{x}^3=\boldsymbol{x}$, inak postup opakujeme so zmenšeným krokom α.

V tomto prípade $\boldsymbol{x}^3$=< 0.99, 0.96>^{T}.

PRÍKLAD: GRADIENTOVÁ METÓDA S KONŠTANTNÝM KROKOM

$\boldsymbol{x}^0$=<0, 2>^{T}, $\boldsymbol{x}^1$=<0.22, 1.55>^{T} , $\boldsymbol{x}^2$=<0.53, 1.16>^{T} α=0,5.

Pokračujeme: Určíme gradient $f'(\boldsymbol{x}^2)$ a smer $\boldsymbol{h}^2=-f'(\boldsymbol{x}^2)$.

$$f'(\begin{bmatrix}0.53\\1.16\end{bmatrix}) = \begin{bmatrix}-2.94\\1.28\end{bmatrix},\quad \mathbf{h}^2 = \begin{bmatrix}2.94\\-1.28\end{bmatrix},\quad |\mathbf{h}^2| = \sqrt{(2.94)^2 + (-1.28)^2} = 3.20$$

Vykonáme presun v smere $\boldsymbol{h}^2$ o veľkosti kroku α=0.5

$$\mathbf{x} = \mathbf{x}^2 + \alpha(\frac{1}{|\mathbf{h}^2|})\mathbf{h}^2 = \begin{bmatrix}0.53\\1.16\end{bmatrix} + 0.5*(\frac{1}{3.20})\begin{bmatrix}2.94\\-1.28\end{bmatrix} = \begin{bmatrix}0.53\\1.16\end{bmatrix} + \begin{bmatrix}0.46\\-0.20\end{bmatrix} = \begin{bmatrix}0.99\\0.96\end{bmatrix}$$

PRÍKLAD: GRADIENTOVÁ METÓDA S KONŠTANTNÝM KROKOM

$\boldsymbol{x}^1$=<0.22,1.55>^{T}, $\boldsymbol{x}^2$=<0.53,1.16>^{T}, $\boldsymbol{x}^3$=<0.99,0.96>^{T}, α=0,5

Pokračujeme: Určíme gradient $f'(\boldsymbol{x}^3)$ a smer $\boldsymbol{h}^3=-f'(\boldsymbol{x}^3)$.

$$f'(\begin{bmatrix}0.99\\0.96\end{bmatrix}) = \begin{bmatrix}-2.02\\-0.32\end{bmatrix},\quad \mathbf{h}^3 = \begin{bmatrix}2.02\\0.32\end{bmatrix},\quad |\mathbf{h}^3| = \sqrt{(2.02)^2 + (0.32)^2} = 2.05$$

Vykonáme presun v smere $\boldsymbol{h}^3$ o veľkosti kroku α=0.5

$$\mathbf{x} = \mathbf{x}^3 + \alpha(\frac{1}{|\mathbf{h}^3|})\mathbf{h}^3 = \begin{bmatrix}0.99\\0.96\end{bmatrix} + 0.5*(\frac{1}{2.05})\begin{bmatrix}2.02\\0.32\end{bmatrix} = \begin{bmatrix}0.99\\0.96\end{bmatrix} + \begin{bmatrix}0.49\\0.08\end{bmatrix} = \begin{bmatrix}1.48\\1.04\end{bmatrix}$$

PRÍKLAD: GRADIENTOVÁ METÓDA S KONŠTANTNÝM KROKOM

Overíme, či funkcia f má v bode $\boldsymbol{x}$ menšiu HÚF než v $\boldsymbol{x}^3$:

$\boldsymbol{x}^3$=<0.99, 0.96>, $\boldsymbol{x}$=<1.48, 1.04>^{T}.

Minimalizujte $f(\boldsymbol{x}) = (x_1)^2+4(x_2)^2-4x_1-8x_2$ pre $\boldsymbol{x} \in E^2$

$$f(\mathbf{x}) = f(\begin{bmatrix} 0.99 \\ 0.96 \end{bmatrix}) = (0.99)^2 + 4(0.96)^2 - 4*0.99 - 8*0.96 = -6.97$$

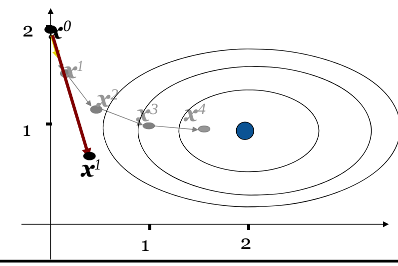
$$f(\mathbf{x}) = f(\begin{bmatrix} 1.48 \\ 1.04 \end{bmatrix}) = (1.48)^2 + 4(1.04)^2 - 4*1.48 - 8*1.04 = -7.72$$

Pokiaľ áno, definujeme $\boldsymbol{x}^4 = \boldsymbol{x}$, inak postup opakujeme so zmenšeným krokom α.

V tomto prípade $\boldsymbol{x}^4$=<1.48, 1.04>^{T}.

PRÍKLAD: GRADIENTOVÁ METÓDA NAJVÄČŠIEHO POKLESU

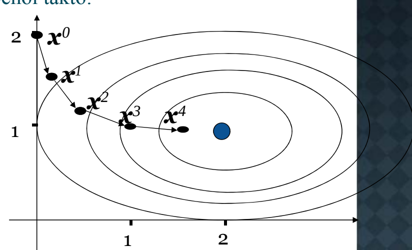
Líši sa od gradientovej metódy s konštantným krokom spôsobom výpočtu α_i, ktoré je všeobecne v každom kroku iné a to také, aby v **danom smere $\boldsymbol{h}^i$** nadobudla funkcia f v bode $\boldsymbol{x}^{i+1} = \boldsymbol{x}^i + \alpha_i \boldsymbol{h}^i$ svoje minimum.



PRÍKLAD: GRADIENTOVÁ METÓDA S KONŠTANTNÝM KROKOM

Náš postup od $\boldsymbol{x}^0$ k $\boldsymbol{x}^4$ prebehol takto:

$\boldsymbol{x}^i$	$x_1{}^i$	$x_2{}^i$	$f(\boldsymbol{x}^i)$
$\boldsymbol{x}^0$	0	2	0
$\boldsymbol{x}^1$	0.22	1.55	-3.62
$\boldsymbol{x}^2$	0.53	1.16	-5.74
$\boldsymbol{x}^3$	0.99	0.96	-6.97
$\boldsymbol{x}^4$	1.48	1.04	-7.72



To bola ukážka **gradientovej metódy s konštantným krokom**.

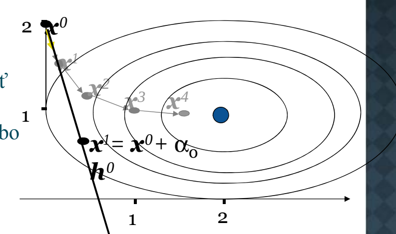
PRÍKLAD: GRADIENTOVÁ METÓDA NAJVÄČŠIEHO POKLESU

AKO SA URČÍ α_i ?

Definuje sa funkcia g premennej α≥ 0 takto: $g(\alpha)=f(\boldsymbol{x}^i+\alpha\,\boldsymbol{h}^i)$.

Hodnota α_i sa nájde ako bod minima funkcie $g(\alpha)$ pre $\alpha \in <0, \infty)$.

Minimalizácia g môže byť vykonaná prostriedkami matematickej analýzy alebo numericky.



PRÍKLAD: GRADIENTOVÁ METÓDA NAJVÄČŠIEHO POKLESU

Minimalizujte $f(\mathbf{x}) = (x_1)^2 + 4(x_2)^2 - 4x_1 - 8x_2$ pre $\mathbf{x} \in E^2$

Zvoľme $\mathbf{x}^0 = <0, 2>^T$.

Postup: Určíme gradient $f'(\mathbf{x}^0)$ a smer $\mathbf{h}^0 = -f'(\mathbf{x}^0)$.

$$f'(\mathbf{x}^i) = \begin{pmatrix} \frac{\partial f(\mathbf{x}^i)}{\partial x_1} \\ \frac{\partial f(\mathbf{x}^i)}{\partial x_2} \\ \vdots \\ \frac{\partial f(\mathbf{x}^i)}{\partial x_n} \end{pmatrix}$$

$$\frac{\partial f(\mathbf{x}^0)}{\partial x_1} = 2x_1 - 4$$

$$\frac{\partial f(\mathbf{x}^0)}{\partial x_2} = 8x_2 - 8$$

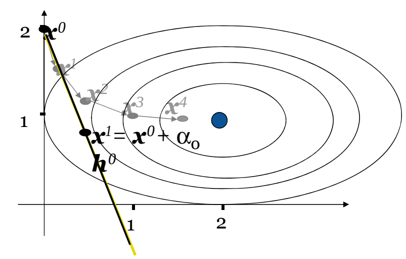
$$f'(\begin{bmatrix} 0 \\ 2 \end{bmatrix}) = \begin{bmatrix} -4 \\ 8 \end{bmatrix}, \quad \mathbf{h}^0 = \begin{bmatrix} 4 \\ -8 \end{bmatrix},$$

$$dlzka\ |\mathbf{h}^0| = \sqrt{(4)^2 + (-8)^2} = \sqrt{80} = 8{,}94$$

PRÍKLAD: GRADIENTOVÁ METÓDA NAJVÄČŠIEHO POKLESU

Vykonáme presun v smere $\mathbf{h}^0$ o veľkosti kroku $\alpha = 0.15$.

$$\mathbf{x} = \mathbf{x}^0 + \alpha \mathbf{h}^0 = \begin{bmatrix} 0 \\ 2 \end{bmatrix} + \alpha \begin{bmatrix} 4 \\ -8 \end{bmatrix} = \begin{bmatrix} 4\alpha \\ 2 - 8\alpha \end{bmatrix} = \begin{bmatrix} 0.6 \\ 0.8 \end{bmatrix}$$



PRÍKLAD: GRADIENTOVÁ METÓDA NAJVÄČŠIEHO POKLESU

Minimalizujte $f(\mathbf{x}) = (x_1)^2 + 4(x_2)^2 - 4x_1 - 8x_2$ pre $\mathbf{x} \in E^2$

Pre smer $\mathbf{h}^0$ definujeme $g(\alpha)$.

$$g(\alpha) = f(\mathbf{x}^0 + \alpha \mathbf{h}^0) = f(\begin{bmatrix} 0 \\ 2 \end{bmatrix} + \alpha \begin{bmatrix} 4 \\ -8 \end{bmatrix}) = f(\begin{bmatrix} 4\alpha \\ 2 - 8\alpha \end{bmatrix}) =$$

$$= (4\alpha)^2 + 4(2 - 8\alpha)^2 - 4(4\alpha) - 8(2 - 8\alpha)$$

Nájdeme minimum g na intervale $\alpha \in <0, \infty)$.

$$g'(\alpha) = 2(4\alpha)4 + 4 * 2(2 - 8\alpha)(-8) - 4(4) - 8(-8) = 0$$

$$(2\alpha) + (2 - 8\alpha)(-4) - (1) - (-4) = 0$$

$$-5 + 34\alpha = 0 \Rightarrow \alpha = 5/34 = 0.15$$

$$g''(\alpha) = 32 + 4 * 2(64) > 0$$

PRÍKLAD: GRADIENTOVÁ METÓDA NAJVÄČŠIEHO POKLESU

Pre $\mathbf{x}^1 = <0.6, 0.8>^T$ určme gradient $f'(\mathbf{x}^1)$ a smer $\mathbf{h}^1 = -f'(\mathbf{x}^1)$

$$f'(\mathbf{x}^i) = \begin{pmatrix} \frac{\partial f(\mathbf{x}^i)}{\partial x_1} \\ \frac{\partial f(\mathbf{x}^i)}{\partial x_2} \\ \vdots \\ \frac{\partial f(\mathbf{x}^i)}{\partial x_n} \end{pmatrix}$$

$$\frac{\partial f(\mathbf{x}^1)}{\partial x_1} = 2x_1 - 4 = -2.8,$$

$$\frac{\partial f(\mathbf{x}^1)}{\partial x_2} = 8x_2 - 8 = -1.6,$$

$$f'(\begin{bmatrix} 0.6 \\ 0.8 \end{bmatrix}) = \begin{bmatrix} -2.8 \\ -1.6 \end{bmatrix}, \quad \mathbf{h}^1 = \begin{bmatrix} 2.8 \\ 1.6 \end{bmatrix},$$

PRÍKLAD: GRADIENTOVÁ METÓDA NAJVÄČŠIEHO POKLESU

Minimalizujte $f(\mathbf{x}) = (x_1)^2 + 4(x_2)^2 - 4x_1 - 8x_2$ pre $\mathbf{x} \in E^2$

Pre smer $\boldsymbol{h}^1$ definujeme $g(\alpha)$.

$$g(\alpha) = f(\mathbf{x}^1 + \alpha\mathbf{h}^1) = f(\begin{bmatrix}0.6\\0.8\end{bmatrix} + \alpha\begin{bmatrix}2.8\\1.6\end{bmatrix}) = f(\begin{bmatrix}0.6+2.8\alpha\\0.8+1.6\alpha\end{bmatrix}) =$$

$$= (0.6+2.8\alpha)^2 + 4(0.8+1.6\alpha)^2 - 4(0.6+2.8\alpha) - 8(0.8+1.6\alpha)$$

Nájdeme minimum g na intervale $\alpha \in <0, \infty)$.

$$g'(\alpha) = 2(0.6+2.8\alpha)*2.8 + 4*2(0.8+1.6\alpha)(1.6) - 4(2.8) - 8(1.6) = 0$$

$$(0.6+2.8\alpha)*0.7 + (0.8+1.6\alpha)(1.6) - (1.4) - (1.6) = 0$$

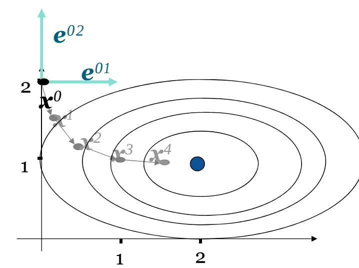
$$4.52\alpha - 1.3 = 0 \Rightarrow \alpha = 1.3/4.52 = 0.29$$

$$g''(\alpha) = 2*2.8*2.8 + 4*2(1.6)*1.6 > 0$$

POWELLOVA METÓDA

Líši sa od gradientových metód spôsobom určovania smeru minimalizácie $\boldsymbol{h}^i$, k čomu **nepotrebujeme** výpočet parciálnych derivácií.

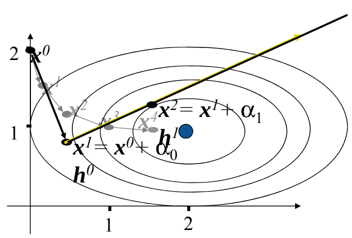
Smer $\boldsymbol{h}^i$ je tu určený na základe množiny elementárnych smerov $\{\boldsymbol{e}^{i1}, \boldsymbol{e}^{i2}, \ldots, \boldsymbol{e}^{in}\}$, ktoré je po každom kroku i aktualizovaná smerom minimalizácie $\boldsymbol{h}^i$.



PRÍKLAD: GRADIENTOVÁ METÓDA NAJVÄČŠIEHO POKLESU

Vykonáme presun v smere $\boldsymbol{h}^1$ o veľkosti kroku α=0.29.

$$\mathbf{x} = \mathbf{x}^1 + \alpha\mathbf{h}^1 = \begin{bmatrix}0.6\\0.8\end{bmatrix} + \alpha\begin{bmatrix}2.8\\1.6\end{bmatrix} = \begin{bmatrix}0.6+2.8\alpha\\0.8+1.6\alpha\end{bmatrix} = \begin{bmatrix}1.41\\1.26\end{bmatrix}$$

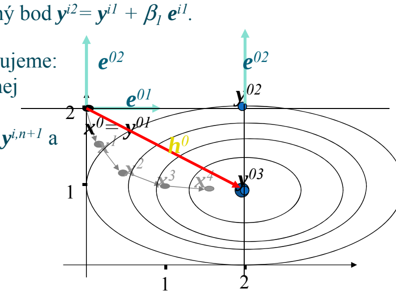


POWELLOVA METÓDA

Určenie smeru minimalizácie $\boldsymbol{h}^i$.
Ideme z bodu $\boldsymbol{x}^i$ a položíme $\boldsymbol{y}^{i1} = \boldsymbol{x}^i$.

Na **priamke** $\boldsymbol{y}(\beta) = \boldsymbol{y}^{i1} + \beta\,\boldsymbol{e}^{i1}$ vykonáme minimalizáciu f pre $\beta \in R$. Získame β_1 a príslušný bod $\boldsymbol{y}^{i2} = \boldsymbol{y}^{i1} + \beta_1\,\boldsymbol{e}^{i1}$.

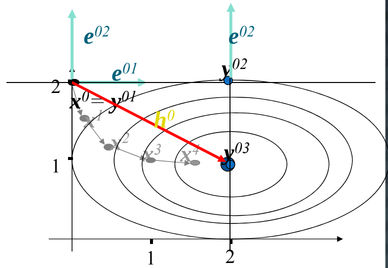
Ďalšiu priamku definujeme: $\boldsymbol{y}(\beta) = \boldsymbol{y}^{i2} + \beta\,\boldsymbol{e}^{i2}$ a z nej získame $\boldsymbol{y}^{i3}$.
Nakoniec dostaneme $\boldsymbol{y}^{i,n+1}$ a potom
$\boldsymbol{h}^i = \boldsymbol{y}^{i,n+1} - \boldsymbol{x}^i$.



POWELLOVA METÓDA

Určenie smeru minimalizácie h^i.
Bod x^{i+1} získame minimalizáciou funkcie g premennej $\alpha \geq 0$ takto:
$g(\alpha)=f(x^i+\alpha\, h^i)$. $x^{i+1}=x^i+\alpha_i\, h^i$.

Množinu elementárnych smerov $\{e^{i1}, e^{i2}, \ldots, e^{in}\}$ aktualizujeme pre krok $i+1$ takto: $\{e^{i2}, \ldots, e^{in}, h^i\}$.



PRÍKLAD: POWELLOVA METÓDA

Minimalizujte $f(\mathbf{x})=(x_1)^2+4(x_2)^2-4x_1-8x_2$ pre $\mathbf{x}\in E^2$

$$\mathbf{y}(\beta)=y^{02}+\beta\,\mathbf{e}^{02}=\begin{bmatrix}2\\2\end{bmatrix}+\beta\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}2\\2+\beta\end{bmatrix}$$

$$g^2(\beta)=f(\mathbf{y}(\beta))=(y_1(\beta))^2+4(y_2(\beta))^2-4y_1(\beta)-8y_2(\beta)=$$
$$=(2)^2+4(2+\beta)^2-4*2-8*(2+\beta)$$

$$g^2{}'(\beta)=8(2+\beta)-8=0\Rightarrow\beta_1=-1\Rightarrow\mathbf{y}^{03}=\begin{bmatrix}2\\1\end{bmatrix}$$

PRÍKLAD: POWELLOVA METÓDA

Minimalizujte $f(\mathbf{x})=(x_1)^2+4(x_2)^2-4x_1-8x_2$ pre $\mathbf{x}\in E^2$

Zvoľme $\mathbf{x}^0=\langle 0, 2\rangle^T$ a množinu elementárnych smerov ako $\{\langle 1, 0\rangle^T, \langle 0, 1\rangle^T\}$.

$$\mathbf{y}(\beta)=\mathbf{x}^0+\beta\,\mathbf{e}^{01}=\begin{bmatrix}0\\2\end{bmatrix}+\beta\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}\beta\\2\end{bmatrix}$$

$$g^1(\beta)=f(\mathbf{y}(\beta))=(y_1(\beta))^2+4(y_2(\beta))^2-4y_1(\beta)-8y_2(\beta)=$$
$$=(\beta)^2+4(2)^2-4\beta-8*2$$

$$g^1{}'(\beta)=2\beta-4=0\Rightarrow\beta_1=2\Rightarrow\mathbf{y}^{02}=\begin{bmatrix}2\\2\end{bmatrix}$$

PRÍKLAD: POWELLOVA METÓDA

Minimalizujte $f(\mathbf{x})=(x_1)^2+4(x_2)^2-4x_1-8x_2$ pre $\mathbf{x}\in E^2$

$$\mathbf{h}^0=\mathbf{y}^{03}-\mathbf{x}^0=\begin{bmatrix}2\\1\end{bmatrix}-\begin{bmatrix}0\\2\end{bmatrix}=\begin{bmatrix}2\\-1\end{bmatrix}$$

$$\mathbf{x}(\alpha)=\mathbf{x}^0+\alpha\,\mathbf{h}^0=\begin{bmatrix}0\\2\end{bmatrix}+\alpha\begin{bmatrix}2\\-1\end{bmatrix}=\begin{bmatrix}2\alpha\\2-\alpha\end{bmatrix}$$

$$g(\alpha)=f(\mathbf{x}(\alpha))=(x_1(\alpha))^2+4(x_2(\alpha))^2-4x_1(\alpha)-8x_2(\alpha)=$$
$$=(2\alpha)^2+4(2-\alpha)^2-4(2\alpha)-8*(2-\alpha)$$
$$g'(\alpha)=2(2\alpha)2+4*2(2-\alpha)(-1)-4*2-8(-1)=0$$
$$2\alpha-2=0\Rightarrow\alpha=1\Rightarrow x^1=\begin{bmatrix}2\\1\end{bmatrix}.$$

V kroku $i=1$ bude východiskový bod $\mathbf{x}^1=\langle 2, 1\rangle^T$ a množina elementárnych smerov bude $\{\mathbf{e}^{02}, \mathbf{h}^0\}=\{\langle 0, 1\rangle^T, \langle 2, -1\rangle^T\}$.

ITERAČNÉ METÓDY PRVÉHO RÁDU

- Uvedené iteračné metódy patria do skupiny iteračných metód **prvého rádu**, podľa toho, že v okolí $\boldsymbol{x}^i$ aproximujú minimalizovanú funkciu nadrovinou a postupujú v smere $\boldsymbol{h}^i$, v ktorom táto nadrovina najrýchlejšie klesá.
- Ďalšou metódou prvého rádu je napr. „Metóda združených smerov“ (Fletcher-Reevesova metóda).
- **Metódy druhého rádu** aproximujú minimalizovanú funkciu kvadratickou formou a postupujú od súčasného riešenia k bodu minima kvadratické formy.

ITERAČNÉ METÓDY PRVÉHO RÁDU

- Ak sú iteračné metódy prvého rádu používané ako heuristiky, bez nároku na konvergenciu ku globálnemu minimu, stačí, ak **existujú spojité prvé parciálne derivácie minimalizovanej funkcie**.
- Predpoklady pre konvergenciu minimalizujúcich postupností ku globálnemu sú podstatne prísnejšie: **konvexnosť funkcie *f*, platnosť Lipschicovej podmienky pre *f'* a obmedzenosť množiny** $\{\boldsymbol{x}: \boldsymbol{x} \in E^n, f(\boldsymbol{x}) \leq f(\boldsymbol{x}^0)\}$.