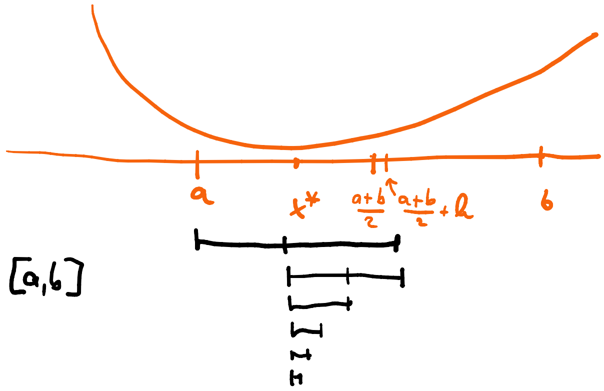
Minule: $\min_{x\in\mathbb{R}} f(x)$ ... **Met. bisekcie:**



(ii)

$f(\frac{a+b}{2}) < f(\frac{a+b}{2}+h) \Rightarrow$ rastie $\sim \frac{a+b}{2}$

$\Rightarrow$ vezmeme $\underbrace{[a, \frac{a+b}{2}+h]}$ → nové $[a,b]$

$f(\frac{a+b}{2}) > f(\frac{a+b}{2}+h) \Rightarrow$ klesá $\sim \frac{a+b}{2}$

$\Rightarrow$ vezmeme $\underbrace{[\frac{a+b}{2}, b]}_{\text{nové } [a,b]}$

- (ii) opakujeme až kým dĺžka intervalu $< \delta$ ← vopred zvolené

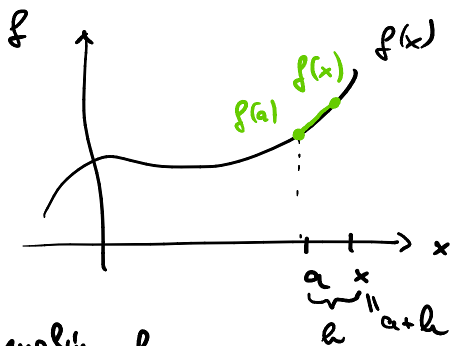
→ ∃ podobné metódy

Rko: optimize

Odbočka: numerické derivovanie

– napr. kvôli f' pre grad. metódu

→ **met. konečných diferencií:** (finite differences)



$$f'(a) = \lim_{x\to a} \frac{f(x)-f(a)}{x-a} = \lim_{h\to 0} \frac{f(a+h)-f(a)}{\cancel{a}+h-\cancel{a}}$$: aproximujeme malým h:

subst. $x = a+h$, $h = x-a$; $x \to a \Downarrow h = x-a \to a-a = 0$

$$f'(a) \approx \frac{f(a+h)-f(a)}{h}$$ (napr. $h = 10^{-6}$)

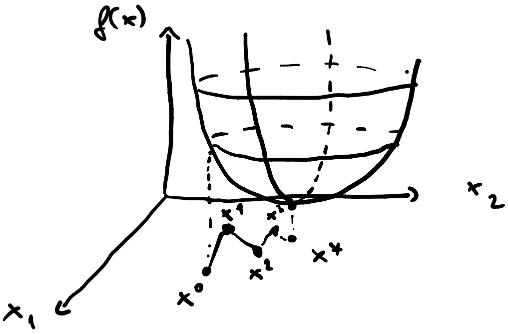
- a podobné: $\frac{f(a+h)-f(a-h)}{2h}$

Viacrozmerná optim. (stále voľná)

$\boxed{\min_{x\in\mathbb{R}^n} f(x)}$ pre $f: \mathbb{R}^n \to \mathbb{R}$

1.) gradientná met.

$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ iterácie → $x^{(0)}, x^{(1)}, \dots \in \mathbb{R}^n$; $x^0, x^1, \dots \in \mathbb{R}^n$; $x_0, x_1, \dots \in \mathbb{R}^n$



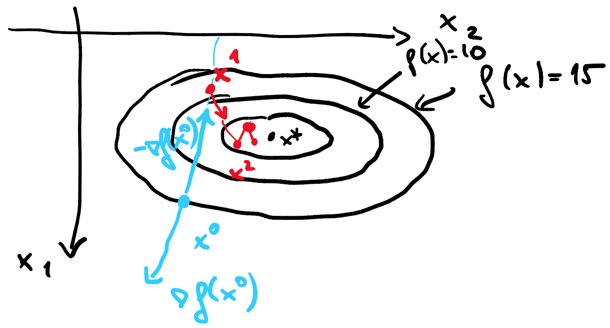
vrstevnice: $f(x) = c$

- smer najväčšieho poklesu

vieme: smer najväčšieho nárastu: **gradient**

$$\nabla_x f(x) = \nabla f(x) = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix} \in \mathbb{R}^n$$

vrstevnice: $f(x)=c$



$\left|\frac{\partial f(x)}{\partial x_n}\right|$

je kolmý na vrstevnicu v bode x

$\Rightarrow$ pohneme sa v smere $-\nabla f(x^k)$

$$\Rightarrow \boxed{x^{k+1} = x^k - \lambda_k \nabla f(x^k)} \quad \forall k = 0,1,2,\dots$$

Pr.: $f(x) = x_1^2 + x_2^2 - x_1 x_2$

$$\nabla f(x) = \begin{pmatrix} 2x_1 - x_2 \\ 2x_2 - x_1 \end{pmatrix}$$

pravidlá zastavenia: ako v 1-D: $\|x^{k+1} - x^k\| < \varepsilon$

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

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

iter. / čas

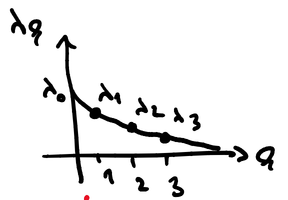
dĺžka kroku (i) $\lambda_k \equiv$ konšt. $\rightarrow$ nebezpečné

(napr. $\lambda_k \equiv 0{,}2$)

(ii) zmenšujúca sa: napr. $\lambda_k = \lambda_0 \cdot e^{-ck}$

$$\lambda_k = \frac{\lambda_0}{1+ck}$$

pre zvolené $c>0$, $\lambda_0>0$

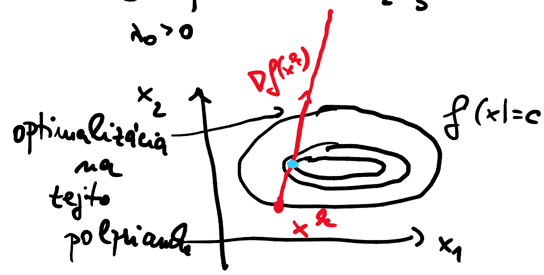


(iii) optimálna:

$$\lambda_k \in \arg\min_{\lambda>0} f(x^k - \lambda \nabla f(x^k))$$

$\uparrow$

1-D optimalizácia: napr. met. bisekcie



• zase iba lok. optimá

Pr.: $f(x) = \frac{x_1^4 + x_2^4}{4} - \frac{x_1^3 + x_2^3}{3} - x_1^2 - x_2^2 + 4$

$$\nabla f(x) = \begin{pmatrix} x_1^3 - x_1^2 - 2x_1 \\ x_2^3 - x_2^2 - 2x_2 \end{pmatrix}$$

2.) Newtonova met.

- aj 2. der.

1-D: $f(x) \doteq f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2} f''(x_0)(x-x_0)^2$

1-D: $f(x) \doteq f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2} f''(x_0)(x-x_0)^2$

$\uparrow$ Taylor v x_0

viacrozmer: $\underset{\substack{\uparrow \\ \text{gradient}}}{\nabla f(x)} = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix}$, $\underset{\substack{\uparrow \\ \text{Hessova mat.}}}{\nabla^2 f(x)} = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & & & \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}$: $n \times n$, sym.

$\frac{\partial^2 f(x)}{\partial x_n \partial x_1} = \frac{\partial^2 f(x)}{\partial x_1 \partial x_n}$

$$\Rightarrow \text{Taylor: } \underset{1\times1}{f(x)} \doteq \underbrace{\underset{1\times1}{f(x^0)} + \overbrace{\underbrace{\underset{1\times n}{\nabla f(x^0)^T} \cdot \underset{n\times1}{(x-x^0)}}_{1\times1}}^{c^T\cdot(x-x_0) \to c} + \frac{1}{2} \overbrace{\underset{1\times n}{(x-x^0)^T} \underbrace{\underset{n\times n}{\nabla^2 f(x^0)} \underset{n\times1}{(x-x^0)}}_{1\times1}}^{(x-x^0)^T A (x-x^0) \to 2A(x-x^0)}}_{h(x)}$$

- namiesto $f(x)$ minimalizujem $h(x)$

Deriváčie v n-D: • $\nabla_x (c^T x) = c$ (pre $c \in \mathbb{R}^n$) $\longleftrightarrow$ Pre $n=1$: $\frac{d}{dx}(c \cdot x) = c$

• $\nabla_x x^T A x = 2Ax$ (pre $A: n\times n$, sym) $\longleftrightarrow$ $\frac{d}{dx}(a \cdot x^2) = 2ax$

$\Rightarrow$ Riešime: $\boxed{\nabla h(x) = 0}$

$$\nabla h(x) = 0 + \nabla f(x^0) + \frac{1}{\cancel{2}} \cdot \cancel{2} \nabla^2 f(x^0)(x-x^0) = 0$$

$$\nabla^2 f(x^0)(x-x^0) = -\nabla f(x^0) \quad / \cdot (\nabla^2 \ )^{-1}$$

$$x - x^0 = -[\nabla^2 f(x^0)]^{-1} \nabla f(x^0)$$

$$x = x^0 \underbrace{- [\nabla^2 f(x^0)]^{-1} \nabla f(x^0)}$$

pohneme sa v tomto smere o λ-násobok

$$\Rightarrow x^1 = x^0 - \lambda_0 [\nabla^2 f(x^0)]^{-1} \nabla f(x^0)$$

N. met.: $\boxed{x^{k+1} = x^k - \lambda_k [\nabla^2 f(x^k)]^{-1} \nabla f(x^k), \forall k = 0, 1, 2, \ldots}$