Prednáška 1

Základné matematické pojmy minimalizácie funkcií

Študujme **reálnu funkciu** n-premenných

$$f\colon R^n \to R$$

$$y = f(x_1, x_2, \ldots, x_n) = f(\boldsymbol{x})$$

Symbol $\boldsymbol{x}$ vyjadruje **vektor**, $\boldsymbol{x} \in R^n$,

$$\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T$$

Norma vektora $\boldsymbol{x} \in R^n$ je definovaná takto

$$|\boldsymbol{x}| = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}$$

Poznámka: Vo všeobecnosti pod normou vektora rozumieme zobrazenie, ktoré priradí každému vektoru nezáporné celé číslo, ktoré vyhovuje týmto podmienkam:

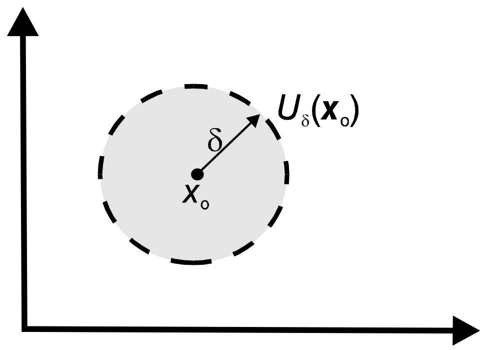
(1) $\forall \boldsymbol{x} \in R^n : |\boldsymbol{x}| \geq 0 \quad (|\boldsymbol{x}| = 0, \text{len pre } \boldsymbol{x} = 0)$

(2) $|\alpha\, \boldsymbol{x}| = |\alpha| \cdot |\boldsymbol{x}|$

(3) $|\boldsymbol{x} + \boldsymbol{y}| \leq |\boldsymbol{x}| + |\boldsymbol{y}|$ (trojuholníková nerovnosť)

Okolie bodu $\boldsymbol{x}_o$ je definované ako množina

$$U_\delta(\boldsymbol{x}_o) = \left\{ \boldsymbol{x} \in R^n ; \left| \boldsymbol{x} - \boldsymbol{x}_o \right| < \delta \right\}$$



Funkcia $f(\boldsymbol{x})$ má v bode $\boldsymbol{x}_o$ **lokálne minimum**, ak existuje také $\delta>0$, že pre každé $\boldsymbol{x} \in U_\delta(\boldsymbol{x}_o)$ platí

$$f(\boldsymbol{x}) \geq f(\boldsymbol{x}_o)$$

Formálne

$$\exists(\delta > 0)\forall\left(\boldsymbol{x} \in U_\delta(\boldsymbol{x}_o)\right): f(\boldsymbol{x}) \geq f(\boldsymbol{x}_o)$$

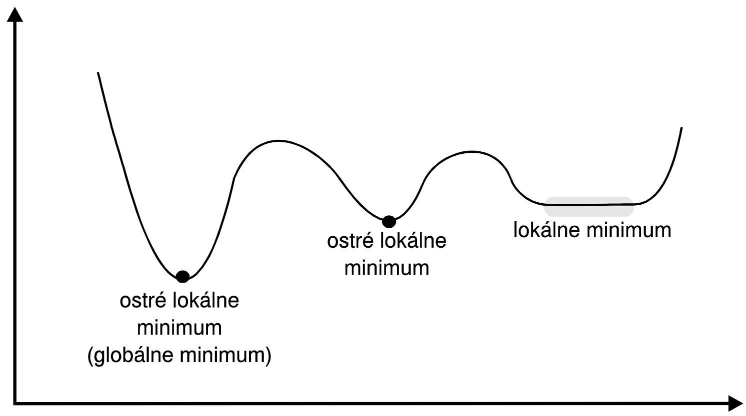
Ostré lokálne minimum je také lokálne minimum, kde

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_o) \quad \text{platí len pre } \boldsymbol{x} = \boldsymbol{x}_o$$

Bod $\boldsymbol{x}_o$ sa nazýva **globálne minimum**, ak pre $\forall \boldsymbol{x} \in \boldsymbol{R}^n$ platí

$$f(\boldsymbol{x}) \geq f(\boldsymbol{x}_o)$$

Grafické ilustračné príklady minimím:



Základné vety o minimách diferencovateľných funkcií

1. Funkcia 1-premennej (n=1)

Taylorov rozvoj funkcie 1-premennej

$$f(x_o + h) = f(x_o) + \frac{1}{1!} f'(x_o) h + \frac{1}{2!} f''(x_o) h^2 + \ldots$$

Nech x_o je minimum, potom

$$f(x_o + h) - f(x_o) = \frac{1}{1!} f'(x_o) h + \frac{1}{2!} f''(x_o) h^2 + \ldots \geq 0$$

Veta 1. Ak funkcia $f(x)$ ma v bode x_o minimum a je v tomto bode diferencovateľná, potom

$$f'(x_o)=0$$

Veta 2. Ak funkcia $f(x)$ v bode x_o vyhovuje podmienkam

$$f'(x_o)=0$$
$$f''(x_o)>0$$

potom má v bode x_o minimum.

2. Funkcia n-premenných (n>1)

Taylorov rozvoj funkcie n-premenných

$$f(\boldsymbol{x}_o+\boldsymbol{h})=f(\boldsymbol{x}_o)+\frac{1}{1!}\sum_{i=1}^{n}\frac{\partial f(\boldsymbol{x}_o)}{\partial x_i}h_i+\frac{1}{2!}\sum_{i,j=1}^{n}\frac{\partial^2 f(\boldsymbol{x}_o)}{\partial x_i \partial x_j}h_i h_j+\ldots$$

kde $\boldsymbol{h}=(h_1,h_2,...,h_n)^T$ je vektor "výchyliek".

Zavedieme nasledujúcu symboliku:

(1) gradient funkcie $f(\boldsymbol{x})$ v bode $\boldsymbol{x}_o$

$$grad\, f(\boldsymbol{x}_o)=\left(\frac{\partial f(\boldsymbol{x}_o)}{\partial x_1},\frac{\partial f(\boldsymbol{x}_o)}{\partial x_2},\ldots,\frac{\partial f(\boldsymbol{x}_o)}{\partial x_n}\right)^T$$

(2) Hessian funkcie $f(\boldsymbol{x})$ v bode $\boldsymbol{x}_o$

$$\boldsymbol{H}(\boldsymbol{x}_o)=\begin{pmatrix} \frac{\partial^2 f(\boldsymbol{x}_o)}{\partial x_1^2}, \frac{\partial^2 f(\boldsymbol{x}_o)}{\partial x_1 \partial x_2}, \dots, \frac{\partial^2 f(x_o)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\boldsymbol{x}_o)}{\partial x_2 \partial x_1}, \frac{\partial^2 f(\boldsymbol{x}_o)}{\partial x_2^2}, \dots, \frac{\partial^2 f(x_o)}{\partial x_2 \partial x_n} \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \frac{\partial^2 f(\boldsymbol{x}_o)}{\partial x_n \partial x_1}, \frac{\partial^2 f(\boldsymbol{x}_o)}{\partial x_n \partial x_2}, \dots, \frac{\partial^2 f(x_o)}{\partial x_n^2} \end{pmatrix}$$

Hessian je symetrická matica typu $n \times n$

$$\boldsymbol{H}(\boldsymbol{x}_o)=\boldsymbol{H}^T(\boldsymbol{x}_o)$$

Táto vlastnosť je dôsledkom toho, že pre 2. parciálne derivácie platí

$$\frac{\partial^2 f(\boldsymbol{x}_o)}{\partial x_i \partial x_j}=\frac{\partial^2 f(\boldsymbol{x}_o)}{\partial x_j \partial x_i}$$

Ak $\boldsymbol{x}_o$ je minimum, potom Taylorov rozvoj poskytuje nasledujúcu kompaktnú podmienku

$$f(\boldsymbol{x}_o+\boldsymbol{h})-f(\boldsymbol{x}_o)=grad\, f(\boldsymbol{x}_o)+\frac{1}{2}\boldsymbol{h}\cdot\boldsymbol{H}(\boldsymbol{x}_o)\boldsymbol{h}+\ldots\geq 0$$

Definícia. Symetrická matica $\boldsymbol{A}=(a_{ij})$ sa nazýva **pozitívne definitná** vtedy, ak pre ľubovolný vektor "výchyliek" $\boldsymbol{h}=(h_1,h_2,\ldots,h_n)^T$ platí

$$\sum_{i,j=1}^{n} h_i a_{ij} h_j \geq 0$$

pričom rovnosť je splnená len pre $\boldsymbol{h}=(0,0,\ldots,0)^T$.

Veta. Symetrická matica $\boldsymbol{A}=(a_{ij})$ je pozitívne definitná (formálne $\boldsymbol{A}>0$) vtedy a len vtedy, ak jej hlavné minory (determinanty na diagonále) sú kladné

$$a_{ij}>0,\quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}>0\quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{21} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}>0,\ldots$$

Veta 1'. Ak funkcia $f(\boldsymbol{x})$ ma v bode $\boldsymbol{x}_o$ minimum a je v tomto bode diferencovateľná, potom

$$grad\, f(\boldsymbol{x}_o) = \boldsymbol{0}$$

Veta 2'. Ak funkcia $f(\boldsymbol{x})$ v bode $\boldsymbol{x}_o$ vyhovuje podmienkam

$$grad\, f(\boldsymbol{x}_o) = \boldsymbol{0}$$

$$\boldsymbol{H}(x_o)>0$$

potom má v bode $\boldsymbol{x}_o$ minimum.

Špecifikácia vety 2' pre funkciu o 2-premenných: Ak funkcia $f(x,y)$ v bode $A=(x_o,y_o)$ vyhovuje podmienkam

$$\frac{\partial f(x_o,y_o)}{\partial x} = 0, \quad \frac{\partial f(x_o,y_o)}{\partial y} = 0$$

$$\frac{\partial^2 f(x_o,y_o)}{\partial x^2} > 0, \quad \frac{\partial^2 f(x_o,y_o)}{\partial x^2} \cdot \frac{\partial^2 f(x_o,y_o)}{\partial y^2} - \left(\frac{\partial^2 f(x_o,y_o)}{\partial x\, \partial y} \right)^2 > 0$$

potom má v bode $A=(x_o,y_o)$ minimum.