Optimalizace

6. Nelineární funkce a zobrazení

Tom Kroupa, Tom Werner

FEL ČVUT

Funkce $f\colon \mathbb{R}^n \to \mathbb{R}$

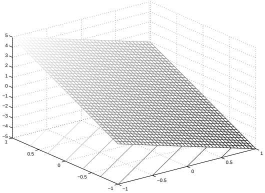
Příklady:

- $f(x,y) = x^2 - y^2$
- $f(\mathbf{x}) = \mathbf{a}^T\mathbf{x} + b$
- $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{b}^T\mathbf{x} + c$

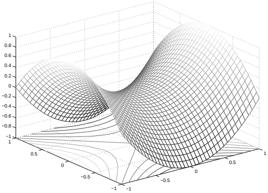
Definice

- **Graf** funkce f je množina $\{\, (\mathbf{x}, y) \in \mathbb{R}^{n+1} \mid \mathbf{x} \in \mathbb{R}^n,\ f(\mathbf{x}) = y \,\}$.
- **Vrstevnice** funkce f výšky y je množina $\{\, \mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) = y \,\}$.

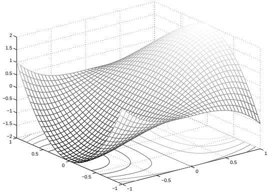
Příklady pro $n = 2$:



$f(x,y) = -2x + 3y$



$f(x,y) = x^2 - y^2$



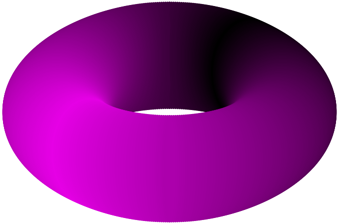
$f(x,y) = 3x - x^3 - 3xy^2$

Příklady pro $n = 1$:

- $\mathbf{f}(t) = \mathbf{x} + t\mathbf{v}$
- $\mathbf{f}(t) = (\cos t, \sin t)$
- $\mathbf{f}(t) = (\cos t, \sin t, t)$

Příklady pro $n > 1$:

- $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$
- vektorové pole $\mathbb{R}^3 \to \mathbb{R}^3$
- $\mathbf{f}(u, v) = ((R + r\cos v)\cos u, (R + r\cos v)\sin u, r\sin v)$



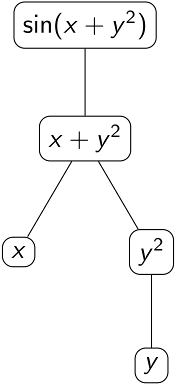
Spojitost

Definice

Zobrazení $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ je **spojité** v bodě $\mathbf{x} \in \mathbb{R}^n$, jestliže

$$\lim_{\mathbf{y}\to\mathbf{x}} \mathbf{f}(\mathbf{y}) = \mathbf{f}(\mathbf{x}).$$

Postačující podmínka: Spojitost se zachovává skládáním funkcí.



Derivace

Derivace funkce $f : \mathbb{R} \to \mathbb{R}$

Definice

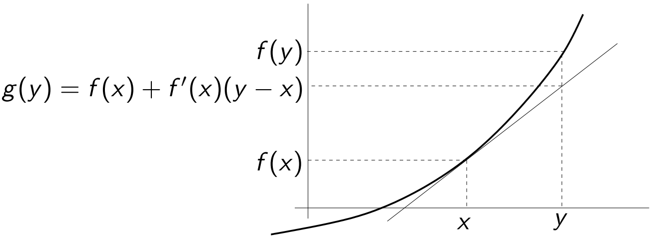
Funkce $f : \mathbb{R} \to \mathbb{R}$ je **diferencovatelná** v bodě x, jestliže existuje $a \in \mathbb{R}$ tak, že

$$\lim_{y \to x} \frac{f(y) - f(x) - a(y-x)}{y - x} = 0.$$

Pak se a nazývá **derivace** funkce f v bodě x a píšeme $f'(x) = \dfrac{\mathrm{d}f(x)}{\mathrm{d}x} = a$.

Intuice: Funkce f musí být v okolí bodu x podobná nějaké afinní funkci

$$g(y) = f(x) + a(y - x)$$



Derivace zobrazení $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$

Definice

Zobrazení $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ je **diferencovatelné** v bodě $\mathbf{x}$, jestliže existuje matice $\mathbf{A} \in \mathbb{R}^{m\times n}$ tak, že

$$\lim_{\mathbf{y}\to\mathbf{x}} \frac{\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) - \mathbf{A}(\mathbf{y}-\mathbf{x})}{\|\mathbf{y}-\mathbf{x}\|} = \mathbf{0}.$$

Pak se $\mathbf{A}$ nazývá **derivace** zobrazení $\mathbf{f}$ v bodě $\mathbf{x}$ a píšeme $\mathbf{f}'(\mathbf{x}) = \dfrac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}} = \mathbf{A}$.

Intuice: Zobrazení $\mathbf{f}$ musí být v okolí bodu $\mathbf{x}$ podobné afinnímu zobrazení

$$\mathbf{g}(\mathbf{y}) = \mathbf{f}(\mathbf{x}) + \mathbf{A}(\mathbf{y}-\mathbf{x})$$

Věta (Postačující podmínka pro diferencovatelnost)

Existují-li v bodě $\mathbf{x}$ všechny parciální derivace $\frac{\partial f_i(\mathbf{x})}{\partial x_j}$ a funkce $\frac{\partial f_i}{\partial x_j}$ jsou v bodě $\mathbf{x}$ spojité, pak je $\mathbf{f}$ v bodě $\mathbf{x}$ diferencovatelné.

Tvrzení

Je-li $\mathbf{f}$ v bodě $\mathbf{x}$ diferencovatelné, pak existují všechny parciální derivace $\frac{\partial f_i(\mathbf{x})}{\partial x_j}$ a

$$\mathbf{f}'(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial f_1(\mathbf{x})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial f_m(\mathbf{x})}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{m \times n}.$$

Speciální případy:

- Pro $f\colon \mathbb{R} \to \mathbb{R}$ je $f'(x)$ skalár
- Pro $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^m$ je $\mathbf{f}'(x) = \begin{bmatrix} f_1'(x) \\ \vdots \\ f_m'(x) \end{bmatrix}$ sloupcový vektor
- Pro $f\colon \mathbb{R}^n \to \mathbb{R}$ je $f'(\mathbf{x}) = \begin{bmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix}$ řádkový vektor

Derivace složeného zobrazení

Věta (Řetízkové pravidlo)

Pro diferencovatelná zobrazení

$$\mathbb{R}^n \xrightarrow{\mathbf{f}} \mathbb{R}^m \xrightarrow{\mathbf{g}} \mathbb{R}^l$$

platí

$$(\mathbf{g} \circ \mathbf{f})'(\mathbf{x}) = \frac{\mathrm{d}\mathbf{g}(\mathbf{f}(\mathbf{x}))}{\mathrm{d}\mathbf{x}} = (\mathbf{g}' \circ \mathbf{f})(\mathbf{x})\, \mathbf{f}'(\mathbf{x}) = \mathbf{g}'(\mathbf{f}(\mathbf{x}))\, \mathbf{f}'(\mathbf{x}).$$

V Leibnizově značení: Je-li $\mathbf{y} = \mathbf{g}(\mathbf{u})$ a $\mathbf{u} = \mathbf{f}(\mathbf{x})$, pak

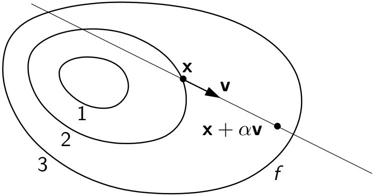
$$\frac{\mathrm{d}\mathbf{y}}{\mathrm{d}\mathbf{x}} = \frac{\mathrm{d}\mathbf{y}}{\mathrm{d}\mathbf{u}} \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}\mathbf{x}}$$

Přirozeně se zobecní pro složení více zobrazení.

Derivace maticových výrazů

zobrazení	derivace
$\mathbf{f}(\mathbf{x}) = \mathbf{x}$	$\mathbf{f}'(\mathbf{x}) = \mathbf{I}$
$\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$	$\mathbf{f}'(\mathbf{x}) = \mathbf{A}$
$f(\mathbf{x}) = \mathbf{x}^T\mathbf{x} = \|\mathbf{x}\|^2$	$f'(\mathbf{x}) = 2\mathbf{x}^T$
$f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$	$f'(\mathbf{x}) = \mathbf{x}^T(\mathbf{A} + \mathbf{A}^T)$
$f(\mathbf{x}) = \|\mathbf{x}\|$	$f'(\mathbf{x}) = \mathbf{x}^T/\|\mathbf{x}\|$
$\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{g}(\mathbf{x})$	$\mathbf{f}'(\mathbf{x}) = \mathbf{A}\mathbf{g}'(\mathbf{x})$
$\mathbf{f}(\mathbf{x}) = \mathbf{g}(\mathbf{A}\mathbf{x} + \mathbf{b})$	$\mathbf{f}'(\mathbf{x}) = \mathbf{g}'(\mathbf{A}\mathbf{x} + \mathbf{b})\mathbf{A}$
$f(\mathbf{x}) = \mathbf{g}(\mathbf{x})^T\mathbf{g}(\mathbf{x})$	$f'(\mathbf{x}) = 2\mathbf{g}(\mathbf{x})^T\mathbf{g}'(\mathbf{x})$
$f(\mathbf{x}) = \mathbf{g}(\mathbf{x})^T\mathbf{h}(\mathbf{x})$	$f'(\mathbf{x}) = \mathbf{g}(\mathbf{x})^T\mathbf{h}'(\mathbf{x}) + \mathbf{h}(\mathbf{x})^T\mathbf{g}'(\mathbf{x})$

Směrová derivace



Řez zobrazení $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ v bodě $\mathbf{x} \in \mathbb{R}^n$ ve směru $\mathbf{v} \in \mathbb{R}^n$ je zobrazení

$$\varphi\colon \mathbb{R} \to \mathbb{R}^m, \quad \varphi(\alpha) = \mathbf{f}(\mathbf{x} + \alpha\mathbf{v})$$

Směrová derivace zobrazení $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ v bodě $\mathbf{x}$ ve směru $\mathbf{v}$ je vektor

$$\varphi'(0) = \lim_{\alpha\to 0} \frac{\varphi(\alpha) - \varphi(0)}{\alpha} = \lim_{\alpha\to 0} \frac{\mathbf{f}(\mathbf{x} + \alpha\mathbf{v}) - \mathbf{f}(\mathbf{x})}{\alpha}$$

Věta

Je-li zobrazení $\mathbf{f}$ v bodě $\mathbf{x}$ diferencovatelné, pak jeho směrová derivace v bodě $\mathbf{x}$ ve směru $\mathbf{v}$ je rovna $\mathbf{f}'(\mathbf{x})\mathbf{v}$.

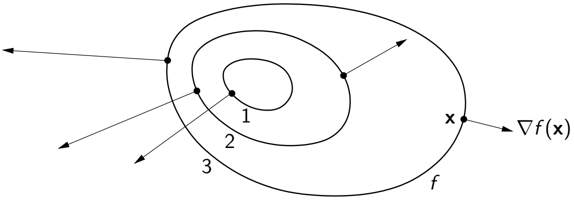
Myšlenka důkazu: Zobrazení $\varphi(\alpha) = \mathbf{f}(\mathbf{x} + \alpha\mathbf{v})$ je složením zobrazení $\mathbf{f}$ a zobrazení $g(\alpha) = \mathbf{x} + \alpha\mathbf{v}$. Věta pak plyne z řetízkového pravidla.

Gradient funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ je sloupcový vektor

$$\nabla f(\mathbf{x}) = f'(\mathbf{x})^T = \begin{bmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix} \in \mathbb{R}^n$$

Směrová derivace $f'(\mathbf{x})\mathbf{v} = \nabla f(\mathbf{x})^T \mathbf{v}$ je

- maximální ve směru $\dfrac{\nabla f(\mathbf{x})}{\|\nabla f(\mathbf{x})\|}$
- nulová ve směru ortogonálním na $\nabla f(\mathbf{x})$



Parciální derivace druhého řádu

Věta

Nechť $f\colon \mathbb{R}^n \to \mathbb{R}$ a $\mathbf{x} \in \mathbb{R}^n$. Jestliže druhé parciální derivace

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}, \qquad \frac{\partial^2 f(\mathbf{x})}{\partial x_j \partial x_i}$$

v bodě $\mathbf{x}$ existují a jsou v bodě $\mathbf{x}$ spojité, potom jsou si rovny.

Hessova matice je

$$f''(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_n} \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

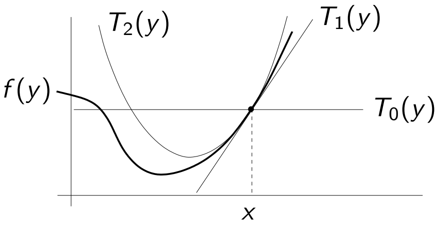
Taylorův polynom funkce

Definice

Taylorův polynom funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ stupně k v bodě $\mathbf{x}$ je polynom

$$T_k\colon \mathbb{R}^n \to \mathbb{R}$$

stupně k, který má v bodě $\mathbf{x}$ parciální derivace řádu $0, 1, \ldots, k$ stejné jako f.



Taylorovy polynomy funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ v bodě $\mathbf{x}$ do stupně dva:

$$\begin{aligned}
T_0(\mathbf{y}) &= f(\mathbf{x}) \\
T_1(\mathbf{y}) &= f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{y}-\mathbf{x}) \\
T_2(\mathbf{y}) &= f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{y}-\mathbf{x}) + \tfrac{1}{2}(\mathbf{y}-\mathbf{x})^T f''(\mathbf{x})(\mathbf{y}-\mathbf{x})
\end{aligned}$$

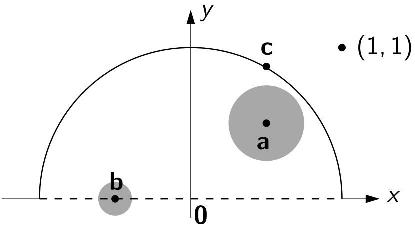
Lokální extrémy

Vnitřek a hranice množiny

Nechť $X \subseteq \mathbb{R}^n$. Bod $\mathbf{x} \in \mathbb{R}^n$ je její

- **vnitřní bod**, pokud $B_\epsilon(\mathbf{x}) \subseteq X$ pro nějaké $\epsilon > 0$,
- **hraniční bod**, pokud $B_\epsilon(\mathbf{x}) \cap X \neq \emptyset$ a $B_\epsilon(\mathbf{x}) \cap (\mathbb{R}^n \setminus X) \neq \emptyset$ pro všechna $\epsilon > 0$.

Příklad: $X = \{\, (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1,\ y > 0 \,\} \ \cup \ \{\, (1, 1) \,\}$



Extrémy funkce na množině

Definice

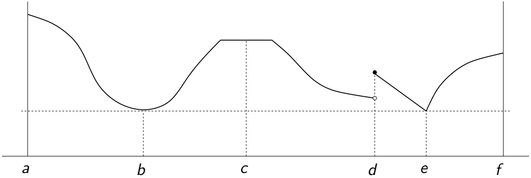
Funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ nabývá na množině $X \subseteq \mathbb{R}^n$ v bodě $\mathbf{x} \in X$

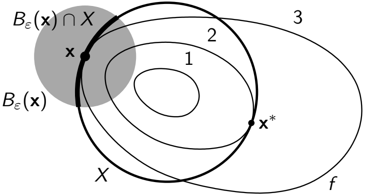
- **(globálního) minima**, jestliže $f(\mathbf{x}) \leq f(\mathbf{x}')$ pro všechna $\mathbf{x}' \in X$
- **lokálního minima**, jestliže f v bodě $\mathbf{x}$ nabývá (globálního) minima na množině $X \cap B_\epsilon(\mathbf{x})$ pro nějaké $\epsilon > 0$

(Globální, lokální) minimum v bodě $\mathbf{x}$ je

- **volné**, pokud $\mathbf{x}$ je vnitřním bodem množiny X
- **vázané**, pokud $\mathbf{x}$ je hraničním bodem množiny X

Příklady:





Shlukování: Pro dané $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ minimalizuj funkci

$$f(\mathbf{x}_1, \ldots, \mathbf{x}_k) = \sum_{i=1}^{m} \min_{j=1}^{k} \|\mathbf{a}_i - \mathbf{x}_j\|^2$$

přes $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{R}^n$.