Optimalizace

Reálné funkce a zobrazení, lokální extrémy

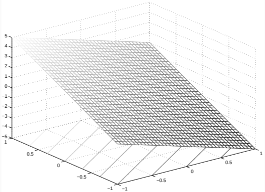
Tomáš Kroupa, Mirko Navara

FEL ČVUT

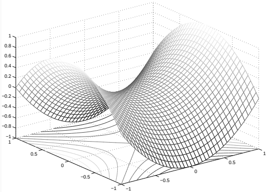
Funkce $f\colon \mathbb{R}^n \to \mathbb{R}$

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

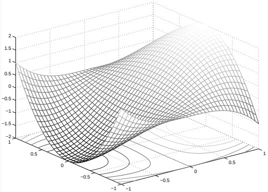
Příklady funkcí $f(x_1, x_2)$



$-2x_1 + 3x_2$



$x_1^2 - x_2^2$

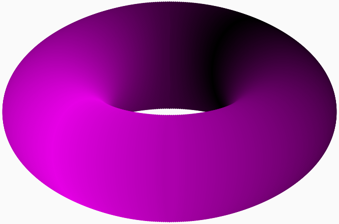


$3x_1 - x_1^3 - 3x_1x_2^2$

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

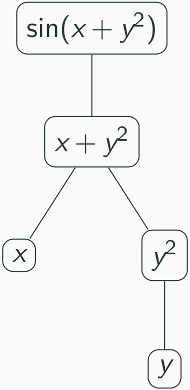
Příklady

- Afinní zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$, kde $\mathbf{A} \in \mathbb{R}^{m\times n}$ a $\mathbf{b} \in \mathbb{R}^m$
- Skalární pole $\mathbb{R}^3 \to \mathbb{R}$, vektorové pole $\mathbb{R}^3 \to \mathbb{R}^3$
- Parametrizace toru
  $\mathbf{f}(u, v) = \big((R + r\cos v)\cos u, (R + r\cos v)\sin u, r\sin v\big)$



Spojitost zobrazení $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ v bodě $\mathbf{x} \in \mathbb{R}^n$

- Definice pomocí limity
- Spojitost se zachovává skládáním funkcí, což vede na prakticky použitelnou postačující podmínku



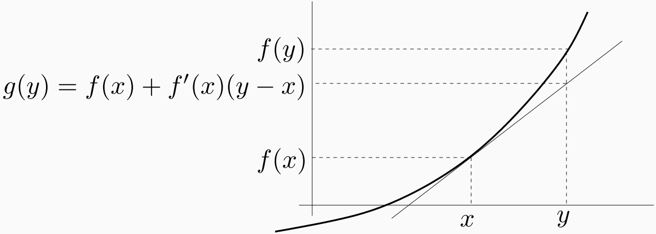
Derivace funkce $f\colon \mathbb{R} \to \mathbb{R}$

Pokud existuje

$$a = \lim_{y \to x} \frac{f(y) - f(x)}{y - x}, \quad \text{t.j.} \quad \lim_{y \to x} \frac{f(y) - f(x) - a\,(y - x)}{y - x} = 0,$$

pak a se nazývá derivace funkce f v bodě x, píšeme $f'(x) \coloneqq a$ a říkáme, že funkce f je v bodě x diferencovatelná nebo že tam má diferenciál, kterým je afinní zobrazení

$$g(y) = f(x) + a\,(y - x).$$



Značení derivací

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

značení	derivace	hodnota v x	hodnota v 1
Lagrangeovo	f'	$f'(x)$	$f'(1)$
Leibnizovo	$\frac{\mathrm{d}f}{\mathrm{d}x}$	$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{\mathrm{d}f}{\mathrm{d}x}(x)$	$\frac{\mathrm{d}f}{\mathrm{d}x}(1) = \frac{\mathrm{d}f(x)}{\mathrm{d}x}\Big\vert_{x=1}$
operátorové	$\mathrm{D}f$	$\mathrm{D}f(x)$	$\mathrm{D}f(1)$

Parciální derivace funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ podle i-té proměnné

značení	derivace	hodnota v $\mathbf{x}$	hodnota v $\mathbf{1}$
Lagrangeovo	f_{x_i}	$f_{x_i}(\mathbf{x})$	$f_{x_i}(\mathbf{1})$
Leibnizovo	$\frac{\partial f}{\partial x_i}$	$\frac{\partial f(\mathbf{x})}{\partial x_i} = \frac{\partial f}{\partial x_i}(\mathbf{x})$	$\frac{\partial f}{\partial x_i}(\mathbf{1}) = \frac{\partial f(\mathbf{x})}{\partial x_i}\Big\vert_{\mathbf{x}=\mathbf{1}}$
operátorové	$\mathrm{D}_i f$	$\mathrm{D}_i f(\mathbf{x})$	$\mathrm{D}_i f(\mathbf{1})$

Parciální derivace funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ podle všech proměnných

značení	derivace	hodnota v $\mathbf{x}$	hodnota v $\mathbf{1}$
Lagrangeovo	f'	$f'(\mathbf{x})$	$f'(\mathbf{1})$
Leibnizovo	$\frac{\mathrm{d}f}{\mathrm{d}\mathbf{x}}$	$\frac{\mathrm{d}f(\mathbf{x})}{\mathrm{d}\mathbf{x}} = \frac{\mathrm{d}f}{\mathrm{d}\mathbf{x}}(\mathbf{x})$	$\frac{\mathrm{d}f}{\mathrm{d}\mathbf{x}}(\mathbf{1}) = \frac{\mathrm{d}f(\mathbf{x})}{\mathrm{d}\mathbf{x}}\Big\vert_{\mathbf{x}=\mathbf{1}}$
operátorové	$\nabla f^\top$	$\nabla f(\mathbf{x})^\top$	$\nabla f(\mathbf{1})^\top$

Konvergence vektorové funkce vektorového argumentu $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$

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

znamená

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

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

Definice

Pokud existuje matice $\mathbf{A} \in \mathbb{R}^{m\times n}$ taková, 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 $\mathbf{A}$ se nazývá derivace nebo Jacobiho matice zobrazení $\mathbf{f}$ v bodě $\mathbf{x}$, píšeme $\mathbf{f}'(\mathbf{x}) := \mathbf{A}$ a říkáme, že zobrazení $\mathbf{f}$ je v bodě $\mathbf{x}$ diferencovatelné nebo že tam má totální diferenciál, kterým je afinní zobrazení

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

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

Existuje-li derivace zobrazení $\mathbf{f} = (f_1, \dots, f_m)$ v bodě $\mathbf{x} \in \mathbb{R}^n$, platí

$$\mathbf{f}'(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial f_1(\mathbf{x})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial f_m(\mathbf{x})}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \mathrm{D}_1 f_1(\mathbf{x}) & \dots & \mathrm{D}_n f_1(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ \mathrm{D}_1 f_m(\mathbf{x}) & \dots & \mathrm{D}_n f_m(\mathbf{x}) \end{bmatrix} \in \mathbb{R}^{m \times n}.$$

Věta

Jestliže v bodě $\mathbf{x}$ všechny parciální derivace $\frac{\partial f_i}{\partial x_j} = \mathrm{D}_j f_i$ existují a jsou spojité, potom má $\mathbf{f}$ v $\mathbf{x}$ derivaci (totální diferenciál).

Speciální případy

$\mathbf{g}\colon \mathbb{R} \to \mathbb{R}^m$

$$\mathbf{g}'(x) = \begin{bmatrix} g_1'(x) \\ \vdots \\ g_m'(x) \end{bmatrix} .$$

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

$$f'(\mathbf{x}) = \begin{bmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix} = \nabla f(\mathbf{x})^\mathsf{T} ,$$

kde

$$\nabla f(\mathbf{x}) := \begin{bmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \mathrm{D}_1 f(\mathbf{x}) \\ \vdots \\ \mathrm{D}_n f(\mathbf{x}) \end{bmatrix} = f'(\mathbf{x})^\mathsf{T}$$

je gradient funkce f v bodě $\mathbf{x}$.

Řetízkové pravidlo

Věta o derivaci složeného zobrazení

Pro diferencovatelná zobrazení

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "\mathbf{f}"] \arrow[rr, bend right, "\mathbf{h}:=\mathbf{g}\circ\mathbf{f}"'] & \mathbb{R}^m \arrow[r, "\mathbf{g}"] & \mathbb{R}^l \end{tikzcd}$$

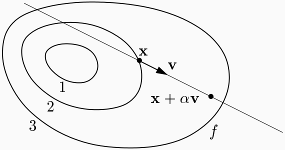
platí

$$\mathbf{h}'(\mathbf{x}) = \mathbf{g}'\big(\mathbf{f}(\mathbf{x})\big)\, \mathbf{f}'(\mathbf{x}).$$

Směrová derivace

Směrová derivace 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 vektor

$$\lim_{\alpha \to 0} \frac{\mathbf{f}(\mathbf{x} + \alpha \mathbf{v}) - \mathbf{f}(\mathbf{x})}{\alpha} \in \mathbb{R}^m .$$



Tvrzení

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

Speciální případ: funkce $f\colon \mathbb{R}^n \to \mathbb{R}$

Tvrzení

Je-li funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ v bodě $\mathbf{x}$ diferencovatelná, pak její směrová derivace v bodě $\mathbf{x}$ ve směru $\mathbf{v}$ je rovna skalárnímu součinu $\nabla f(\mathbf{x})^\mathsf{T}\,\mathbf{v} = f'(\mathbf{x})\,\mathbf{v}$ a je pro jednotkový vektor $\mathbf{v}$

- maximální, je-li $\mathbf{v} = \frac{\nabla f(\mathbf{x})}{\|\nabla f(\mathbf{x})\|}$;
- nulová, je-li $\mathbf{v} \perp \nabla f(\mathbf{x})$.

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

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_i\,\partial x_j} = \mathrm{D}_j\mathrm{D}_i f(\mathbf{x})\,, \qquad \frac{\partial^2 f(\mathbf{x})}{\partial x_j\,\partial x_i} = \mathrm{D}_i\mathrm{D}_j f(\mathbf{x})$$

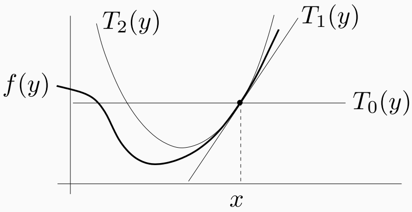
v bodě $\mathbf{x}$ existují a jsou v bodě $\mathbf{x}$ spojité, potom jsou si rovny.

Hessova matice je

$$f''(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1\,\partial x_1} & \dots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1\,\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n\,\partial x_1} & \dots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n\,\partial x_n} \end{bmatrix} = \begin{bmatrix} \mathrm{D}_1\mathrm{D}_1 f(\mathbf{x}) & \dots & \mathrm{D}_n\mathrm{D}_1 f(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ \mathrm{D}_1\mathrm{D}_n f(\mathbf{x}) & \dots & \mathrm{D}_n\mathrm{D}_n f(\mathbf{x}) \end{bmatrix} \in \mathbb{R}^{n\times n}$$

Taylorův polynom pro funkci $f\colon \mathbb{R}^n \to \mathbb{R}$

Hledáme polynom k-tého stupně $T_k\colon \mathbb{R}^n \to \mathbb{R}$, který má v bodě $\mathbf{x}$ všechny parciální derivace až do řádu k stejné jako funkce f.



Taylorovy polynomy v bodě x do stupně 2

$$T_0(\mathbf{y}) = f(\mathbf{x})$$

$$T_1(\mathbf{y}) = f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{y} - \mathbf{x})$$

$$T_2(\mathbf{y}) = f(\mathbf{x}) + f'(\mathbf{x})(\mathbf{y} - \mathbf{x}) + \tfrac{1}{2}(\mathbf{y} - \mathbf{x})^{\mathsf{T}} f''(\mathbf{x})(\mathbf{y} - \mathbf{x})$$

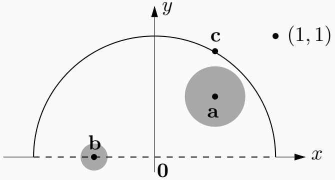
Vnitřek a hranice množiny

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

- vnitřní bod, pokud existuje $\varepsilon > 0$ takové, že $B_\varepsilon(\mathbf{x}) \subseteq X$,
- hraniční bod, pokud pro každé $\varepsilon > 0$ je $B_\varepsilon(\mathbf{x}) \cap X \neq \emptyset$ a $B_\varepsilon(\mathbf{x}) \cap (\mathbb{R}^n \setminus X) \neq \emptyset$.

Které body jsou vnitřní a které hraniční?

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



Extrémy funkce na množině

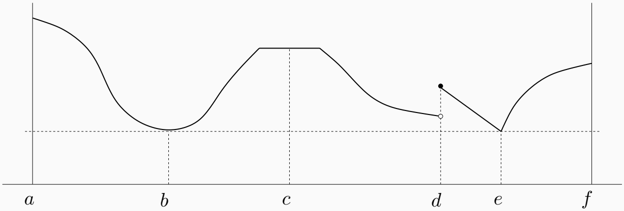
Funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ nabývá na množině $X \subseteq \mathbb{R}^n$ v bodě $\mathbf{x} \in X$

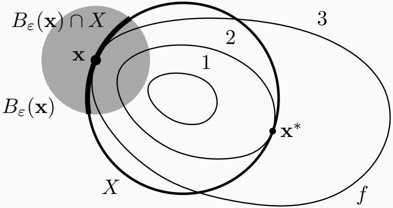
- minima, pokud $f(\mathbf{x}) \leq f(\mathbf{y})$ pro všechna $\mathbf{y} \in X$,
- lokálního minima, pokud existuje $\varepsilon > 0$ tak, že $f(\mathbf{x}) \leq f(\mathbf{y})$ pro všechna $\mathbf{y} \in X \cap B_\varepsilon(\mathbf{x})$.

(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.

Extrémy funkce na množině – příklady





Podmínka prvního řádu

Věta

Nechť $X \subseteq \mathbb{R}^n$, $\mathbf{x} \in X$ je vnitřní bod množiny X a $f\colon \mathbb{R}^n \to \mathbb{R}$ je diferencovatelná v $\mathbf{x}$. Je-li $\mathbf{x}$ lokální extrém funkce f na X, platí

$$f'(\mathbf{x}) = \mathbf{0}.$$

Stacionární bod $\mathbf{x} \in \mathbb{R}^n$ funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ splňuje $f'(\mathbf{x}) = \mathbf{0}$.

Podmínky druhého řádu

Věta

Nechť $X \subseteq \mathbb{R}^n$, $\mathbf{x} \in X$ je vnitřní bod množiny X a $f\colon \mathbb{R}^n \to \mathbb{R}$ je dvakrát diferencovatelná v $\mathbf{x}$. Je-li $\mathbf{x}$ stacionární bod, platí:

- Pokud je $\mathbf{x}$ lokální minimum, pak je Hessova matice $f''(\mathbf{x})$ pozitivně semidefinitní.
- Pokud je $f''(\mathbf{x})$ pozitivně definitní, pak je $\mathbf{x}$ ostré lokální minimum.
- Pokud je $f''(\mathbf{x})$ indefinitní, pak $\mathbf{x}$ není lokální extrém.

Sedlový bod $\mathbf{x} \in \mathbb{R}^n$ funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ je stacionární bod takový, že $f''(\mathbf{x})$ je indefinitní.

Reference

- T. Werner. *Optimalizace* (kapitola 8). Elektronická skripta. FEL ČVUT, 17. 10. 2022. https://cw.fel.cvut.cz/wiki/_media/courses/b0b33opt/opt.pdf
- J. Hamhalter a J. Tišer. *Diferenciální počet funkcí více proměnných*. FEL ČVUT, 2005. https://math.fel.cvut.cz/en/people/tiser/dipocet.pdf
- *Matlab Activities for Multivariable Calculus.* https://mse.redwoods.edu/darnold/math50c/matlab/
- M. Navara. *Matematika spojitého světa*. Elektronická skripta. FEL ČVUT, 2011. https://cmp.felk.cvut.cz/~navara/MSS/MSS_print.pdf