# 3. Taylorův polynom, lokální extrémy funkce více proměnných

*(Matematika 3)*

Petr Gurka

aktualizováno 16. října 2019

## Obsah



# 1 Aproximace funkce Taylorovým polynomem

## 1.1 Taylorův polynom funkce jedné proměnné

Připomeňme, že má-li funkce $f$ v bodě $a$ derivace do řádu $m$, $n \in \mathbb{N}$, pak *Taylorovým polynomem* $m$-tého stupně funkce $f$ v bodě $a$ je polynom

$$T^f_m(x) := f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(m)}(a)}{m!}(x-a)^m$$

Položíme-li $x = a + h$, dostaneme

$$T^f_m(a+h) := f(a) + \frac{f'(a)}{1!}h + \frac{f''(a)}{2!}h^2 + \cdots + \frac{f^{(m)}(a)}{m!}h^m.$$

Je-li

$$D^k_h f(a) = f^{(k)}(a)\, h^k, \quad k \in \mathbb{N},$$

*diferenciál* řádu $k$ v bodě $a$, můžeme psát

$$T^f_m(a+h) := f(a) + \frac{D_h f(a)}{1!} + \frac{D^2_h f(a)}{2!} + \cdots + \frac{D^m_h f(a)}{m!}.$$

*Jak velké chyby se dopustíme, když funkci $f$ nahradíme Taylorovým polynomem $T_m^f$?* Označme

$$R_m(x) = f(x) - T_m^f(x), \qquad \text{tj.} \quad f(x) = T_m^f(x) + R_m(x).$$

**Tvrzení 1.** *Nechť $x > a$. Nechť funkce $f$ má derivaci řádu $m+1$ ve všech bodech intervalu $\langle a, x\rangle$ (v krajních bodech příslušné jednostranné). Pak existuje bod $c \in (a, x)$ tak, že*

$$R_m(x) = \frac{f^{(m+1)}(c)}{(m+1)!}(x-a)^{m+1} \qquad \textit{(Lagrangeův tvar zbytku).}$$

*Pozor, ke každému $x$ může být jiné $c$!*

Vždy ale platí $\displaystyle\lim_{x\to a} \frac{R_m(x)}{(x-a)^m} = 0$ (Peanův odhad zbytku).

## 1.2 Taylorova věta pro funkci $n$ proměnných

Zvolíme

$$g(t) = f(\mathbf{a} + t\,\mathbf{h}), \quad t \in \langle 0, 1\rangle,$$

a tuto funkci aproximujeme Taylorovým polynomem v $t = 0$.

**Věta 2.** *Nechť funkce $f : \mathbb{R}^n \to \mathbb{R}$ má všechny parciální derivace řádu $m+1$ spojité na otevřené množině $G$. Nechť $\mathbf{a}, \mathbf{h} \in G$ a úsečka s krajními body $\mathbf{a}$, $\mathbf{a}+\mathbf{h}$ (tj. body $\mathbf{x} = \mathbf{a} + t\,\mathbf{h}$, $t \in \langle 0, 1\rangle$) leží v $G$. Potom existuje číslo $\tau \in (0, 1)$ takové, že*

$$f(\mathbf{a}+\mathbf{h}) = \sum_{k=0}^{m} \frac{(\mathbf{h}\cdot\nabla)^k f(\mathbf{a})}{k!} + \frac{(\mathbf{h}\cdot\nabla)^{m+1} f(\mathbf{a}+\tau\mathbf{h})}{(m+1)!}. \tag{1}$$

Je

$$(\mathbf{h}\cdot\nabla)f = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}, \quad (\mathbf{h}\cdot\nabla)^2 f = \sum_{i,j=1}^{n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}. \tag{2}$$

Vzhledem k tomu, že funkce $f$ je $(m+1)$ krát spojitě diferencovatelná, platí, že všechny smíšené parciální derivace do řádu $(m+1)$ (včetně) jsou záměnné. Tedy k vyjádření $(\mathbf{h}\cdot\nabla)^k f$, $k = 1, \ldots, m+1$, můžeme použít binomický vzorec, který má ve speciálním případě $n = 2$ tvar

$$(\mathbf{h}\cdot\nabla)^k f = \sum_{j=0}^{k} \binom{k}{j} h_1^{k-j} h_2^j \frac{\partial^k f}{\partial x_1^{k-j} \partial x_2^j}. \tag{3}$$

## 1.3 Totální diferenciály vyšších řádů

Snadno vidíme, že $(\mathbf{h}\cdot\nabla)f(\mathbf{a}) = D_{\mathbf{h}} f(\mathbf{a})$ je totální diferenciál funkce $f$ v bodě $\mathbf{a}$. Definujeme totální diferenciál řádu $k$ v bodě $\mathbf{a}$:

$$D_{\mathbf{h}}^k f(\mathbf{a}) = (\mathbf{h}\cdot\nabla)^k f(\mathbf{a})$$

jako totální diferenciál funkce $\mathbf{x} \mapsto D_{\mathbf{h}}^{k-1} f(\mathbf{x})$ (proměnné $\mathbf{x}$) v bodě $\mathbf{a}$.

Totální diferenciál 2. řádu je kvadratickou formou na $\mathbb{R}^n$, kterou můžeme zapsat (viz (2)) ve tvaru

$$D_{\mathbf{h}}^2 f(\mathbf{a}) = \mathbf{h}^T \boldsymbol{H}\, \mathbf{h},$$

kde matice $\boldsymbol{H}$ je Hessova matice, $\mathbf{h}$ je vektor zapsaný jako sloupec,

$$\boldsymbol{H} = \boldsymbol{H}(\mathbf{a}) = \begin{pmatrix} \frac{\partial^2}{\partial x_1^2} f(\mathbf{a}) & \frac{\partial^2}{\partial x_1 \partial x_2} f(\mathbf{a}) & \dots & \frac{\partial^2}{\partial x_1 \partial x_n} f(\mathbf{a}) \\ \frac{\partial^2}{\partial x_2 \partial x_1} f(\mathbf{a}) & \frac{\partial^2}{\partial x_2^2} f(\mathbf{a}) & \dots & \frac{\partial^2}{\partial x_2 \partial x_n} f(\mathbf{a}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2}{\partial x_n \partial x_2} f(\mathbf{a}) & \frac{\partial^2}{\partial x_n \partial x_2} f(\mathbf{a}) & \dots & \frac{\partial^2}{\partial x_n^2} f(\mathbf{a}) \end{pmatrix}.$$

Jsou-li všechny parciální derivace 2. řádu funkce $f$ spojité v bodě $\mathbf{a}$, potom je matice $\boldsymbol{H}(\mathbf{a})$ symetrická.

### 1.4 Taylorova věta pro $n = 2$, $m = 2$

Obecnou Taylorovu formuli můžeme psát ve tvaru

$$f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + D_{\mathbf{h}} f(\mathbf{a}) + \dots + \frac{D_{\mathbf{h}}^m f(\mathbf{a})}{m!} + R_m(\mathbf{a}, \mathbf{h}),$$

přičemž pro zbytek $R_m(\mathbf{a}, \mathbf{h})$ platí $\lim\limits_{\mathbf{h} \to \mathbf{o}} \dfrac{R_m(\mathbf{a}, \mathbf{h})}{\|\mathbf{h}\|^m} = 0.$

**Pro $n = 2$, $m = 2$, $\mathbf{a} = (a_1, a_2)$, $\mathbf{h} = (h_1, h_2)$ je**
(viz (1), (3))

$$f(a_1 + h_1, a_2 + h_2) = f(a_1, a_2) + \left(h_1 \tfrac{\partial}{\partial x_1} f(a_1, a_2) + h_2 \tfrac{\partial}{\partial x_2} f(a_1, a_2)\right)$$
$$+ \tfrac{1}{2}\left(h_1^2 \tfrac{\partial^2}{\partial x_1^2} f(a_1, a_2) + 2h_1 h_2 \tfrac{\partial^2}{\partial x_1 \partial x_2} f(a_1, a_2) + h_2^2 \tfrac{\partial^2}{\partial x_2^2} f(a_1, a_2)\right) + R_2,$$

$$\lim_{(h_1, h_2) \to (0,0)} \frac{R_2(\mathbf{a}, \mathbf{h})}{h_1^2 + h_2^2} = 0.$$

## 2 Lokální extrémy funkce více proměnných

### 2.1 Definice lokálního maxima a lokálního minima

**Definice 3.** Funkce $f$ má v bodě $\mathbf{a}$

- lokální minimum, jestliže existuje $\delta > 0$, že pro každé $\mathbf{x} \in U_\delta(\mathbf{a})$ je $f(\mathbf{x}) \geq f(\mathbf{a})$;
- lokální maximum, jestliže existuje $\delta > 0$, že pro každé $\mathbf{x} \in U_\delta(\mathbf{a})$ je $f(\mathbf{x}) \leq f(\mathbf{a})$.

Lokální maximum (resp. lokální minimum) je ostré, jestliže je nerovnost ostrá pro všechny body $\mathbf{x} \neq \mathbf{a}$, když $\delta > 0$ je dostatečně malé.

Z definice vyplývá, že lokální extrém (tj. lokální maximum nebo lokální minimum) může mít funkce pouze ve *vnitřním* bodě svého definičního oboru.

## 2.2 Podmínky pro existenci lokálních extrémů

**Věta 4** (Nutné podmínky)**.** *Nechť funkce* $f : \mathbb{R}^n \to \mathbb{R}$ *má v bodě* $\mathbf{a}$ *lokální extrém a má v tomto bodě slabý diferenciál* $D_{\mathbf{v}} f(\mathbf{a})$*. Potom* $D_{\mathbf{v}} f(\mathbf{a}) = 0$ *pro každý vektor* $\mathbf{v} \in \mathbb{R}^n$*. To také znamená, že*

$$\frac{\partial}{\partial x_i} f(\mathbf{a}) = 0, \quad i = 1, \ldots, n. \tag{4}$$

**Věta 5** (Postačující podmínky)**.** *Nechť funkce* $f : \mathbb{R}^n \to \mathbb{R}$ *má v bodě* $\mathbf{a}$ *spojité všechny parciální derivace 2. řádu a diferenciál* $\mathbf{v} \mapsto D_{\mathbf{v}} f(\mathbf{a})$ *je nulová lineární forma na* $\mathbb{R}^n$ *(což je ekvivalentní s podmínkou* (4)*). Je-li kvadratická forma* $\mathbf{v} \mapsto D^2_{\mathbf{v}} f(\mathbf{a})$

- *negativně definitní, má* $f$ *v* $\mathbf{a}$ *ostré lokální maximum;*
- *pozitivně definitní, má* $f$ *v* $\mathbf{a}$ *ostré lokální minimum;*
- *indefinitní, nemá* $f$ *v* $\mathbf{a}$ *lokální extrém.*

Všimněme si, že díky předpokladu spojitosti parciálních derivací 2. řádu má funkce $f$ v bodě $\mathbf{a}$ totální diferenciály 1. a 2. řádu.

*Argument důkazu Věty 5.* Za silnějších předpokladů na hladkost funkce $f$ platí pomocí Taylorovy věty

$$f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + \tfrac{1}{2} D^2_{\mathbf{h}} f(\mathbf{a}) + R_2(\mathbf{a}, \mathbf{h}), \tag{5}$$

přičemž pro „malé“ nenulové vektory $\mathbf{h}$ (tj. norma $\|\mathbf{h}\|$ je malá) můžeme zbytek $R_2(\mathbf{a}, \mathbf{h})$ na pravé straně rovnosti (5) ignorovat. Dále stačí použít definici definitnosti příslušné kvadratické formy. □

- Z důkazu vidíme, že semidefinitnost nestačí k existenci extrému.
- Vzhledem k faktu, že kvadratická forma $D^2_{\mathbf{h}} f(\mathbf{a})$ je reprezentovaná Hessovou maticí $\boldsymbol{H}(\mathbf{a})$, je možné postačující podmínky pro existenci lokálních extrémů formulovat pomocí Sylvestrova kritéria.

## 2.3 Postačující podmínky pro lokální extrém funkce dvou proměnných

**Věta 6** (Postačující podmínky pro lokální extrém funkce 2 proměnných)**.** *Nechť* $f : \mathbb{R}^2 \to \mathbb{R}$ *má spojité parciální derivace 2. řádu v bodě* $\mathbf{a}$ *a nechť platí* (4)*, tj.* $\frac{\partial f}{\partial x}(\mathbf{a}) = 0, \quad \frac{\partial f}{\partial y}(\mathbf{a}) = 0.$ *Označme* $D(\mathbf{a}) = \det \boldsymbol{H}(\mathbf{a})$*, tj.*

$$D(\mathbf{a}) = \begin{vmatrix} \frac{\partial^2 f}{\partial x^2}(\mathbf{a}) & \frac{\partial^2 f}{\partial x \partial y}(\mathbf{a}) \\ \frac{\partial^2 f}{\partial y \partial x}(\mathbf{a}) & \frac{\partial^2 f}{\partial y^2}(\mathbf{a}) \end{vmatrix} = \frac{\partial^2 f}{\partial x^2}(\mathbf{a})\, \frac{\partial^2 f}{\partial y^2}(\mathbf{a}) - \frac{\partial^2 f}{\partial y \partial x}(\mathbf{a})\, \frac{\partial^2 f}{\partial x \partial y}(\mathbf{a}).$$

*Pak*

1. *Je-li* $D(\mathbf{a}) < 0$*, funkce* $f$ *v bodě* $\mathbf{a}$ *nemá lokální extrém (má tam sedlový bod).*

2. *Je-li* $D(\mathbf{a}) > 0$ *a* $\frac{\partial^2 f}{\partial x^2}(\mathbf{a}) < 0$, *má funkce* $f$ *v bodě* $\mathbf{a}$ *ostré lokální maximum.*

3. *Je-li* $D(\mathbf{a}) > 0$ *a* $\frac{\partial^2 f}{\partial x^2}(\mathbf{a}) > 0$, *má funkce* $f$ *v bodě* $\mathbf{a}$ *ostré lokální minimum.*

*Poznámka* 7.

- Ze spojitosti parciálních derivací druhého řádu v bodě $\mathbf{a}$ plyne, že

$$\frac{\partial^2}{\partial y \partial x} f(\mathbf{a}) = \frac{\partial^2}{\partial x \partial y} f(\mathbf{a}).$$

  Odtud je vidět, že pokud $D(\mathbf{a}) > 0$, pak $\frac{\partial^2 f}{\partial x^2}(\mathbf{a}) \neq 0$.

- Pokud je $D(\mathbf{a}) = 0$, může, ale nemusí být v bodě $\mathbf{a}$ lokální extrém.

Na následujícím obrázku vlevo je sedlový bod, na obrázku vpravo je lokální maximum.