

Matematická analýza pro informatiky I.

10. přednáška

Diferenciální počet funkcí více proměnných (II)

Jan Tomeček
jan.tomecek@upol.cz
<http://aix-slx.upol.cz/~tomecek/index>

Univerzita Palackého v Olomouci

12. dubna 2011

Totální diferenciál

Na minulé přednášce jsme se seznámili s pojmy parciální derivace a derivace ve směru. Motivovali jsme je potřebou analogie derivace funkce jedné proměnné (třeba kvůli vyšetřování lokálních extrémů funkcí více proměnných).

Viděli jsme ovšem, že **derivace ve směru (a tedy i parciální derivace) nejsou plnohodnotnou analogií k derivaci funkce jedné proměnné.**

Například funkce (více proměnných) mající všechny směrové derivace nemusí být vůbec spojitá (zatímco u funkce jedné proměnné **existence vlastní derivace v bodě implikuje spojitost v tomto bodě**). Zavedeme si pojem totální diferenciál (nebo jen krátce diferenciál). Jde o zobecnění **diferenciálu** funkce jedné proměnné.

Nejprve trocha lineární algebry

Definice

Nechť V, W jsou reálné vektorové prostory. Zobrazení $\ell : V \rightarrow W$ nazveme *lineární zobrazení*, jestliže

$$\forall \alpha, \beta \in \mathbb{R} \quad \forall x, y \in V : \quad \ell(\alpha x + \beta y) = \alpha \ell(x) + \beta \ell(y).$$

Je-li $V = \mathbb{R}^N$, $W = \mathbb{R}$, nazýváme takové zobrazení ℓ **lineární formou**.
Měli bychom z algebry vědět, že **lineární formy jsou reprezentovány vektorem** (z $\mathbb{R}^N$):

Věta

Nechť $\ell : \mathbb{R}^N \rightarrow \mathbb{R}$ je lineární zobrazení. Pak existuje právě jedno $\alpha \in \mathbb{R}^N$ takové, že

$$\ell(x) = (\alpha, x) \quad \text{pro všechna } x \in \mathbb{R}^N$$

($(\cdot, \cdot)$ je standardní skalární součin – tak, jak jsme ho definovali v jedné z předchozích přednášek).

Důkaz

Zřejmě pro každé $x \in \mathbb{R}^N$ platí

$$x = \sum_{i=1}^N x_i e^{[i]},$$

kde $e^{[i]}$ jsou prvky standardní báze $\mathbb{R}^N$, $x_i \in \mathbb{R}$ jsou souřadnice vektoru x vzhledem k této bázi – jsou určeny jednoznačně. Pak z linearity plyne

$$\ell(x) = \ell\left(\sum_{i=1}^N x_i e^{[i]}\right) = \sum_{i=1}^N x_i \ell(e^{[i]}) = \sum_{i=1}^N x_i \alpha_i = (\alpha, x),$$

kde $\alpha_i = \ell(e^{[i]})$, $\alpha = (\alpha_1, \dots, \alpha_N)$.

Definice

Vektor $\alpha \in \mathbb{R}^N$ z předchozí věty nazýváme *reprezentantem* lineární formy.

Příklad

Uvažujme lineární formu $\ell : \mathbb{R}^2 \rightarrow \mathbb{R}$ definovanou předpisem

$$\ell(x) = x_1 - x_2.$$

Její reprezentant je zřejmě $\alpha = (1, -1)$.

Definice totálního diferenciálu

Již známe **diferenciál funkce jedné proměnné**:

Nechť $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$, $\mathcal{U}(a) \subset D$. Lineární zobrazení $df(a) : \mathbb{R} \rightarrow \mathbb{R}$ jsme nazývali diferenciálem funkce f v bodě a , jestliže platilo

$$f(a+h) - f(a) = df(a)(h) + o(h)$$

kde $\lim_{h \rightarrow 0} \frac{o(h)}{h} = 0$.

V tomto případě

$$df(a) : \mathbb{R} \rightarrow \mathbb{R},$$

tedy reprezentant je reálné číslo $\alpha \in \mathbb{R}$ a platí

$$df(a)(h) = \alpha h \quad \text{pro } h \in \mathbb{R}.$$

Také jsme dokázali, že **funkce má v bodě a diferenciál právě tehdy, když existuje vlastní $f'(a)$** a v tomto případě platí

$$\alpha = f'(a).$$

Tedy **$f'(a)$** je **reprezentantem** diferenciálu funkce jedné proměnné.

Nyní se podívejme na funkci více proměnných.

Definice

Nechť $f : \mathbb{R}^N \rightarrow \mathbb{R}$, $a \in \mathcal{U}(a) \subset \mathcal{D}(f)$. Lineární formu $df(a)$ nazveme **totálním diferenciálem funkce f v bodě a** , jestliže

$$f(a+h) - f(a) = df(a)(h) + o(h) \quad \text{pro všechna } h \in \mathbb{R}^N,$$

taková, že $a+h \in \mathcal{U}(a)$, kde $o : \mathbb{R}^N \rightarrow \mathbb{R}$ splňuje

$$\lim_{h \rightarrow 0} \frac{o(h)}{\|h\|} = 0.$$

V tomto případě je diferenciálem lineární forma typu $\mathbb{R}^N \rightarrow \mathbb{R}$, tzn. je jednoznačně určena svým **reprezentantem** $\alpha = (\alpha_1, \dots, \alpha_N) \in \mathbb{R}^N$, takže platí

$$df(a)(h) = (\alpha, h) = \sum_{i=1}^N \alpha_i h_i.$$

Reprezentant totálního diferenciálu

Bude nás zajímat, jestli lze (podobně jako u funkce jedné proměnné) určit koeficienty α_i pomocí směrových derivací.

Příklad

Uvažujme funkci $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definovanou předpisem

$$f(x_1, x_2) = x_1^2 + x_2^3,$$

$a = (a_1, a_2) \in \mathbb{R}^2$, $h = (h_1, h_2) \in \mathbb{R}^2$. Pak platí

$$\begin{aligned} f(a + h) - f(a) &= f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2) \\ &= (a_1 + h_1)^2 + (a_2 + h_2)^3 - a_1^2 - a_2^3 \\ &= 2a_1h_1 + 3a_2^2h_2 + h_1^2 + 3a_2h_2^2 + h_2^3. \end{aligned}$$

Výraz $2a_1h_1 + 3a_2^2h_2$ by mohl být předpisem diferenciálu $df(a)$ v proměnné $h = (h_1, h_2)$.

...pokračování příkladu

Tedy reprezentant v tom případě by byl určen takto

$$\alpha_1 = 2a_1, \quad \alpha_2 = 3a_2^2.$$

V našem případě tedy platí

$$\alpha_1 = \frac{\partial f}{\partial x_1}(a), \quad \alpha_2 = \frac{\partial f}{\partial x_2}(a).$$

Musíme ale ještě dokázat, že

$$\lim_{h \rightarrow 0} \frac{h_1^2 + 3a_2 h_2^2 + h_2^3}{\|h\|} = 0.$$

...pokračování příkladu

Snadno odhadneme (použijeme nerovností $|h_i| \leq \|h\|$ pro $i = 1, 2$)

$$\begin{aligned} \left| \frac{h_1^2 + 3a_2h_2^2 + h_2^3}{\|h\|} - 0 \right| &\leq \frac{h_1^2 + 3|a_2|h_2^2 + |h_2|h_2^2}{\|h\|} \\ &\leq \frac{\|h\|^2 + 3|a_2|\|h\|^2 + \|h\|^3}{\|h\|} = \|h\| + 3|a_2|\|h\| + \|h\|^2 \end{aligned}$$

Limitním přechodem $h \rightarrow 0$ (nebo-li $\|h\| \rightarrow 0$) dostáváme to, co jsme chtěli...

Z tohoto příkladu bychom mohli usoudit, že diferenciál funkce f v bodě a je lineární forma reprezentovaná vektorem

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_r}(a) \right).$$

Skutečně tomu tak je. Nazýváme ho **gradientem funkce f v bodě a** . To ale musíme dokázat obecně.

Věta o reprezentaci diferenciálu

Věta

Nechť má funkce f v bodě a totální diferenciál. Pak je funkce f v bodě a spojitá, existují $f_{x_i}(a)$ pro $i = 1, \dots, N$ a platí

$$df(a)(h) = (\nabla f(a), h) \quad \text{pro } h \in \mathbb{R}^N.$$

Důkaz

(a) spojitost: Předpokládejme, že funkce f má v bodě a diferenciál, tzn. existuje $\alpha = (\alpha_1, \dots, \alpha_N)$ tak, že

$$f(a + h) - f(a) = df(a)(h) + o(h) = (\alpha, h) + o(h)$$

kde $\lim_{h \rightarrow 0} \frac{o(h)}{\|h\|} = 0$ (z čehož plyne

$$\lim_{h \rightarrow 0} o(h) = \lim_{h \rightarrow 0} \frac{o(h)}{\|h\|} \|h\| = 0).$$

...pokračování důkazu

Pak podle Cauchy–Schwarzovy nerovnosti platí

$$|f(a+h) - f(a)| \leq |(\alpha, h)| + |o(h)| \leq \|\alpha\| \|h\| + |o(h)|.$$

Pravá strana jde k nule pro $h \rightarrow 0$. Tedy f je spojitá v bodě a .

...pokračování důkazu

(b) *reprezentace diferenciálu*: V definici diferenciálu položíme $h = te^{[i]}$, kde $t \in \mathbb{R} \setminus \{0\}$. Pak platí

$$f(a + te^{[i]}) - f(a) = (\alpha, te^{[i]}) + o(te^{[i]}) = t\alpha_i + o(te^{[i]}).$$

Podělíme t a přejdeme $t \rightarrow 0$ a dostáváme

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t \rightarrow 0} \frac{f(a + te^{[i]}) - f(a)}{t} = \alpha_i + \lim_{t \rightarrow 0} \frac{o(te^{[i]})}{t}.$$

Z faktu

$$\left| \frac{o(te^{[i]})}{t} \right| = \left| \frac{o(te^{[i]})}{\|te^{[i]}\|} \right| \rightarrow 0 \quad \text{pro } t \rightarrow 0$$

plyne existence parciálních derivací i reprezentace diferenciálu.

Věta

Má-li funkce f v bodě a diferenciál, pak pro každé $\nu \neq 0$ existuje $f_\nu(a)$ a platí

$$\frac{\partial f}{\partial \nu}(a) = (\nabla f(a), \nu).$$

Důkaz

Pro $\nu \neq 0$ platí

$$\begin{aligned} \frac{\partial f}{\partial \nu}(a) &= \lim_{t \rightarrow 0} \frac{f(a + t\nu) - f(a)}{t} \\ &= \lim_{t \rightarrow 0} \frac{(\nabla f(a), t\nu) + o(t\nu)}{t} = (\nabla f(a), \nu) + \lim_{t \rightarrow 0} \frac{o(t\nu)}{t}. \end{aligned}$$

Podobně jako v důkazu předešlé věty dokážeme, že limita napravo je rovna nule.

Uvědomíme—li si geometrický význam skalárního součinu, můžeme odvodit následující.

Poznámka

(*Geometrický význam gradientu*) Předpokládejme, že f je definována na okolí $\mathcal{U}(a)$ bodu a a má v něm totální diferenciál. Je-li $\nabla f(a) \neq 0$, pak udává směr největšího růstu funkce f (podobně vektor $-\nabla f(a)$ udává směr největšího úbytku/poklesu).

Vyplývá to z následujících dvou faktů:

- (i) Uvažujme vektor $\|\nu\| = 1$ (protože jen pro takový má f_ν geometrický význam – rychlosti růstu funkce). Pak platí

$$\frac{\partial f}{\partial \nu}(a) = (\nabla f(a), \nu) \leq \|\nabla f(a)\| \|\nu\| = \|\nabla f(a)\|.$$

Tedy funkce v žádném směru nemá růst vyšší než $\|\nabla f(a)\|$.

(ii) Položme

$$\nu = \frac{\nabla f(a)}{\|\nabla f(a)\|}$$

(jde o normovaný vektor – ověřte!, který určuje stejný směr jako gradient $\nabla f(a)$). Pak

$$\frac{\partial f}{\partial \nu}(a) = (\nabla f(a), \frac{\nabla f(a)}{\|\nabla f(a)\|}) = \frac{1}{\|\nabla f(a)\|} (\nabla f(a), \nabla f(a)) = \|\nabla f(a)\|.$$

Z toho plyne, že směrová derivace ve směru normovaného vektoru ν majícího stejný směr jako $\nabla f(a)$ nabývá maximální hodnoty.

Poznámka

Existence všech parciálních derivací $f_{x_i}(a)$ pro $i = 1, \dots, N$ nestačí k existenci totálního diferenciálu $df(a)$.

Uvažujme funkci

$$f(x_1, x_2) = \begin{cases} x_1 & 0 \leq x_1 \leq x_2, \\ x_2 & 0 \leq x_2 \leq x_1, \\ 0 & \text{jinak.} \end{cases}$$

Dokažte, že neexistuje $df(0, 0)$, přitom

$$\frac{\partial f}{\partial x_1}(0, 0) = \frac{\partial f}{\partial x_2}(0, 0) = 0$$

Zatím jsme se bavili o tom, co všechno platí pro funkci mající v bodě totální diferenciál, ale neukázali jsme si, **jak efektivně zjistit zda diferenciál vůbec existuje**. Napravíme to v následující větě:

Věta

Má-li funkce f v bodě a spojité parciální derivace podle všech proměnných, pak má f v bodě a diferenciál.

Důkaz

viz Kopáček, str. 95

Poznámka

(*geometrický význam diferenciálu*) Vzpomeneme si opět na funkci jedné proměnné. Existence diferenciálu funkce f v bodě $x_0 \in \mathbb{R}$ (označme $df(a)(h) = \alpha h$) byla ekvivalentní s existencí vlastní derivace $\alpha = f'(x_0)$. Ta měla geometrický význam ten, že přímka

$$y - f(x_0) = \alpha(x - x_0)$$

udávala rovnici tečny ke grafu funkce f v bodě $(x_0, f(x_0))$, nebo-li zapsáno v řeči diferenciálu: přímka

$$y - f(x_0) = df(x_0)(x - x_0)$$

udává rovnici tečny ke grafu funkce f v bodě $(x_0, f(x_0))$.

... pokračování poznámky

Vraťme se opět k funkcím více proměnných – konkrétně k funkci f dvou proměnných x, y . Dá dokázat, že existence diferenciálu funkce f v bodě (x_0, y_0)

$$df(x_0, y_0)(h, k) = \alpha h + \beta k = \frac{\partial f}{\partial x}(x_0, y_0)h + \frac{\partial f}{\partial y}(x_0, y_0)k$$

je ekvivalentní existenci tečné roviny ke grafu funkce f v bodě $(x_0, y_0, f(x_0, y_0))$. Tečná rovina je pak daná rovnicí

$$z - f(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0).$$

nebo-li zapsáno v řeči diferenciálu:

$$z - f(x_0, y_0) = df(x_0, y_0)(x - x_0, y - y_0).$$

Příklad

Uvažujme funkci

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

a bod $(x_0, y_0) = (1, 2)$. Najdeme rovnici tečné roviny ke grafu funkce f v bodě $(x_0, y_0, f(x_0, y_0))$.

Řešení. Platí

$$\frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = 2y.$$

Pak rovnice tečné roviny je

$$z - f(1, 2) = \frac{\partial f}{\partial x}(1, 2)(x - 1) + \frac{\partial f}{\partial y}(1, 2)(y - 2),$$

tedy

$$z - 5 = 2(x - 1) + 4(y - 2).$$

Diferenciály vyšších řádů

Podobně jako se daly definovat parciální a směrové derivace libovolného řádu, můžeme to samé provést s diferenciálem. Například diferenciál 2. řádu bychom mohli nadefinovat podobným způsobem jako totální diferenciál a posléze odvodit jeho reprezentanta. My zvolíme opačný a pohodlnější postup (nebude ovšem z něj úplně jasná motivace – což nám až tak moc vadit nebude – diferenciály vyšších řádů jsou stejně těžko uchopitelné – podobně jako derivace vyšších řádů – umíme s nimi jednoduše pracovat, ale jejich geometrický význam už není vůbec zřejmý).

Následující úvahy provedeme nejprve pro diferenciál 2. řádu, protože je nejjednodušší a budeme ho používat při vyšetřování lokálních extrémů funkcí více proměnných.

Podobně jako totální diferenciál byl lineární forma, bude diferenciál 2. řádu tzv. kvadratická forma. Pro úplnost si zopakujeme některé pojmy z algebry.

Opět trocha algebry

Nechť V a W jsou reálné vektorové prostory. Zobrazení $b : V \times V \rightarrow W$ nazýváme **bilineární**, jestliže je lineární v obou proměnných, tzn. pro každé $x \in V$ je $b(x, \cdot)$ lineární zobrazení a pro každé $y \in V$ je $b(\cdot, y)$ lineární zobrazení. Speciálně pro $V = \mathbb{R}^N$ a $W = \mathbb{R}$ nazýváme b **bilineární formou**.

Bilineární formy mají také svého reprezentanta – tentokrát to je matice:

Věta

Nechť $b : \mathbb{R}^N \times \mathbb{R}^N \rightarrow \mathbb{R}$ je bilineární forma. Pak existuje $B \in M_{r \times r}(\mathbb{R})$ tak, že

$$b(x, y) = xBy^T \quad \text{pro } x, y \in \mathbb{R}^N.$$

Důkaz

Doma...

Definice

Bilineární formu nazveme symetrickou, jestliže její reprezentant je symetrická matice.

Zobrazení $Q : \mathbb{R}^N \rightarrow \mathbb{R}$ nazveme **kvadratická forma**, jestliže existuje symetrická bilineární forma $b : \mathbb{R}^N \times \mathbb{R}^N \rightarrow \mathbb{R}$ tak, že

$$Q(x) = b(x, x) \quad \text{pro všechna } x \in \mathbb{R}^N.$$

Poznámka

Je-li B reprezentantem f , pak

$$Q(x) = xBx^T \quad \text{pro všechna } x \in \mathbb{R}^N.$$

Definice diferenciálu druhého řádu

Definice

Necht' má funkce f v bodě a spojitě všechny parciální derivace až do druhého řádu včetně. Diferenciálem 2. řádu funkce f v bodě a rozumíme kvadratickou formu

$$d^2f(a)(h) = \sum_{i=1}^N \sum_{j=1}^N \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j (= h \nabla^2 f(a) h^T).$$

Poznámka

Reprezentantem diferenciálu druhého řádu je čtvercová symetrická matice

$$\nabla^2 f(\mathbf{a}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{a}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{a}) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_N}(\mathbf{a}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{a}) & \frac{\partial^2 f}{\partial x_2^2}(\mathbf{a}) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_N}(\mathbf{a}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_N \partial x_1}(\mathbf{a}) & \frac{\partial^2 f}{\partial x_N \partial x_2}(\mathbf{a}) & \cdots & \frac{\partial^2 f}{\partial x_N^2}(\mathbf{a}) \end{pmatrix}$$

tzv. **Hessova** matice (determinant této matice se nazývá hessián).

Příklad

Uvažujme funkci $f(x, y) = x^2 + y^2$. Pak

$$\frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = 2y, \quad \frac{\partial^2 f}{\partial x^2} = 2, \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = 0, \quad \frac{\partial^2 f}{\partial y^2} = 2.$$

Hessova matice funkce f v bodě (x, y) je

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Ještě jednou algebra...

Poznámka

Zobrazení

$$F : (\mathbb{R}^N)^m = \mathbb{R}^N \times \mathbb{R}^N \times \dots \times \mathbb{R}^N \rightarrow \mathbb{R}$$

nazveme **m –lineární formou** jestliže je F lineární v každém svém argumentu.

Například lineární forma je 1–lineární forma a bilineární forma je 2–lineární forma – dokažte.

Je-li F m –lineární forma, pak zobrazení

$$Q : \mathbb{R}^N \rightarrow \mathbb{R},$$

definované předpisem

$$Q(x) = F(x, x, \dots, x) \quad \text{pro } x \in \mathbb{R}^N$$

nazýváme **formou stupně m** . Například kvadratická forma je homogenní forma stupně 2 – dokažte.

Nyní zadefinujeme diferenciál m -tého řádu.

Definice

Nechť $m \in \mathbb{N}$ a funkce f má v bodě a spojitě všechny parciální derivace až do m -tého řádu včetně. **Diferenciálem m -tého řádu** funkce f v bodě a rozumíme homogenní formu stupně m definovanou předpisem

$$d^m f(a)(h) = \sum_{i_1, i_2, \dots, i_m=1}^N \frac{\partial^m f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_m}}(a) h_{i_1} h_{i_2} \cdot h_{i_m}.$$

pro všechna $h = (h_1, \dots, h_N) \in \mathbb{R}^N$.

Taylorův vzorec

U funkce více proměnných jsme schopni dokázat analogii Taylorova vzorce pro funkci jedné proměnné.

Věta

Nechť má funkce $f : \mathbb{R}^N \rightarrow \mathbb{R}$ spojité parciální derivace až do řádu $m + 1$ ($m \in \mathbb{N} \cup \{0\}$). Necht' $a \in \mathbb{R}^N$, $h \in \mathbb{R}^N$ jsou takové, že úsečka

$$\{x \in \mathbb{R}^N : x = a + th : t \in \langle 0, 1 \rangle\}$$

je obsažena v $\text{int}(\mathcal{D}(f))$. Pak existuje $\tau \in (0, 1)$ tak, že

$$\begin{aligned} f(a + h) = & f(a) + df(a)(h) + \frac{1}{2!} d^2 f(a)(h) + \dots + \frac{1}{m!} d^m f(a)(h) \\ & + \frac{1}{(m+1)!} d^{m+1} f(a + \tau h)(h). \end{aligned}$$

Z tvrzení zhruba plyne, že **funkce**, která má na okolí nějakého bodu spojitě parciální derivace vyšších řádů, pak ji **lze** v okolí tohoto bodu poměrně dobře **aproximovat m -lineárními formami** (což jsou polynomy v $\mathbb{R}^N$ = jednoduché funkce).

Tuto větu budeme používat zejména při vyšetřování extrémů funkcí více proměnných.

Doporučená literatura

KOPÁČEK J. Matematická analýza pro fyziky II. Matfyzpress, Praha, 2005.

DOŠLÁ, Z, DOŠLÝ O.: Diferenciální počet funkcí více proměnných, Masarykova univerzita, Brno, 2003.

RACHŮNKOVÁ I. RACHŮNEK L.: Diferenciální počet funkcí více proměnných, Univerzita Palackého, Olomouc, 2004.

FAQ aneb často kladené otázky u zkoušky

- Co je totální diferenciál a jaký má vztah ke směrovým a parciálním derivacím. Za jakých podmínek existuje. Co plyne z jeho existence?
- Jak se dají reprezentovat lineární, bilineární, kvadratické formy? (popř, jak lze reprezentovat homogenní formy stupně $m \in \mathbb{N}$)
- Co znamená $o(h)$?
- Jaký je geometrický význam gradientu (diferenciálu)?
- Co je Hessova matice – jaký je její vztah k diferenciálu 2. řádu?
- Z čeho plyne symetrie Hessovy matice?