6. přednáška 5. listopadu 2007

Souvislost diferenciálu a parciálních derivací. Diferenciál implikuje parciální derivace a spojité parciální derivace implikují diferenciál.

Tvrzení 2.3. *Když je funkce*

$$f:\ U \to \mathbf{R},\quad U \subset \mathbf{R}^m \text{ je okolí bodu } a,$$

diferencovatelná v a, pak má v a všechny parciální derivace a jejich hodnoty určují diferenciál:

$$\begin{aligned}\mathrm{D}f(a)(h) &= \frac{\partial f}{\partial x_1}(a)\cdot h_1 + \frac{\partial f}{\partial x_2}(a)\cdot h_2 + \cdots + \frac{\partial f}{\partial x_m}(a)\cdot h_m \\ &= \langle \nabla f(a), h\rangle.\end{aligned}$$

Také má v a všechny směrové derivace a platí $\mathrm{D}_v f(a) = \mathrm{D}f(a)(v)$.

Důkaz. Z linearity diferenciálu $L = \mathrm{D}f(a)$ máme

$$L(h) = L(h_1e_1 + h_2e_2 + \cdots + h_me_m) = L(e_1)h_1 + \cdots + L(e_m)h_m,$$

kde e_i je i-tý vektor kanonické báze. Ovšem

$$\begin{aligned}\frac{\partial f}{\partial x_i}(a) &= \lim_{t\to 0}\frac{f(a+te_i)-f(a)}{t} \\ &= \lim_{t\to 0}\frac{L(te_i)+o(\|te_i\|)}{t} \\ &= \lim_{t\to 0}\frac{tL(e_i)+o(t)}{t} \\ &= L(e_i),\end{aligned}$$

a tak $L(e_i) = \frac{\partial f}{\partial x_i}(a)$. Tvrzení o směrové derivaci plyne z definice a z právě dokázané formule pro diferenciál. □

Obecně je pro zobrazení $f:\ D \to \mathbf{R}^n$ diferenciál $L = \mathrm{D}f(a):\ \mathbf{R}^m \to \mathbf{R}^n$ reprezentován maticí tvaru $n \times m$ a L se na vektor h aplikuje maticovým násobením:

$$L(h) = \begin{pmatrix} L(h)_1 \\ L(h)_2 \\ \vdots \\ L(h)_n \end{pmatrix} = \begin{pmatrix} l_{1,1} & l_{1,2} & \dots & l_{1,m} \\ l_{2,1} & l_{2,2} & \dots & l_{2,m} \\ \vdots & \vdots & \dots & \vdots \\ l_{n,1} & l_{n,2} & \dots & l_{n,m} \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_m \end{pmatrix}.$$

Podle předešlého tvrzení a bodu 2 Tvrzení 2.1 má tato matice v i-tém řádku gradient souřadnicové funkce f_i v bodě a, takže

$$l_{i,j} = \frac{\partial f_i}{\partial x_j}(a).$$

Důsledek. *Diferenciál zobrazení $f:\ D\to\mathbf{R}^n$ v bodě a, kde $D\subset\mathbf{R}^m$ je okolí a a f má souřadnicové funkce $f=(f_1,f_2,\dots,f_n)$, je dán tzv.* Jacobiho maticí *zobrazení f v bodě a,*

$$\left(\frac{\partial f_i}{\partial x_j}(a)\right)_{i,j=1}^{n,m}=\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \dots & \frac{\partial f_1}{\partial x_m}(a)\\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \dots & \frac{\partial f_2}{\partial x_m}(a)\\ \vdots & \vdots & \dots & \vdots\\ \frac{\partial f_n}{\partial x_1}(a) & \frac{\partial f_n}{\partial x_2}(a) & \dots & \frac{\partial f_n}{\partial x_m}(a)\end{pmatrix}.$$

Je-li tato matice čtvercová, nazývá se její determinant jacobiánem.

Věta 2.4. *Nechť $U\subset\mathbf{R}^m$ je okolí bodu $a\in\mathbf{R}^m$. Pokud má funkce $f:\ U\to\mathbf{R}$ na U všechny parciální derivace a ty jsou v bodě a spojité, pak je f v bodě a diferencovatelná.*

Důkaz. Pro jednoduchost nechť $m=2$ a $a=\overline{0}=(0,0)$. (Viz úlohu 1.) Označíme $h=(h_1,h_2)$ a $h'=(h_1,0)$. Přírůstek $f(h)-f(\overline{0})$ napíšeme pomocí přírůstků ve směrech souřadnicových os:

$$f(h)-f(\overline{0})=(f(h)-f(h'))+(f(h')-f(\overline{0})).$$

Na úsečkách $h'h$ a $\overline{0}h'$ funkce f závisí pouze na proměnné x_2, resp. na x_1. Použijeme Lagrangeovu větu o střední hodnotě:

$$f(h)-f(\overline{0})=\frac{\partial f}{\partial x_2}(\zeta_2)\cdot h_2+\frac{\partial f}{\partial x_1}(\zeta_1)\cdot h_1,$$

kde ζ_2 (resp. ζ_1) je jistý vnitřní bod úsečky $h'h$ (resp. $\overline{0}h'$). Oba body leží v otevřené kouli $B(\overline{0},\|h\|)$. Díky spojitosti v počátku máme

$$\frac{\partial f}{\partial x_2}(\zeta_2)=\frac{\partial f}{\partial x_2}(\overline{0})+\alpha(\zeta_2)\ \text{ a }\ \frac{\partial f}{\partial x_1}(\zeta_1)=\frac{\partial f}{\partial x_1}(\overline{0})+\beta(\zeta_1),$$

kde $\alpha(h),\beta(h)=o(1)$ pro $h\to\overline{0}$ (tj. pro každé $\varepsilon>0$ máme $\delta>0$, že $\|h\|<\delta\Rightarrow|\alpha(h)|<\varepsilon$ a podobně pro $\beta(h)$). Tedy

$$f(h)-f(\overline{0})=\frac{\partial f}{\partial x_2}(\overline{0})\cdot h_2+\frac{\partial f}{\partial x_1}(\overline{0})\cdot h_1+\alpha(\zeta_2)h_2+\beta(h_1)h_1.$$

Díky nerovnostem $0<\|\zeta_1\|,\|\zeta_2\|<\|h\|$ a $|h_1|,|h_2|\le\|h\|$ je jasné, že $\alpha(\zeta_2)h_2+\beta(\zeta_1)h_1=o(h)$ pro $h\to\overline{0}$. Funkce f je diferencovatelná v počátku. □

Lagrangeova věta o střední hodnotě pro funkce více proměnných. Následující dvě tvrzení zobecňují Lagrangeovu větu o střední hodnotě a fakt, že nulovost derivace implikuje konstantnost funkce.

Tvrzení 2.5. *Nechť $U\subset\mathbf{R}^m$ je otevřená množina a $u=ab$ je úsečka s koncovými body a a b ležící v U. Nechť je funkce $f:\ U\to\mathbf{R}$ na u spojitá a má* *v každém vnitřním bodě u diferenciál. Pak existuje vnitřní bod ζ úsečky u s vlastností, že*

$$f(b) - f(a) = \mathrm{D}f(\zeta)(b-a).$$

Důkaz. Položíme $F(t) = f(a+th)$, kde $h = b - a$ a reálné číslo t probíhá interval $[0,1]$. Funkce F je patrně spojitá na $[0,1]$ a v $t \in (0,1)$ má derivaci

$$\begin{aligned} F'(t) &= \lim_{\Delta \to 0} \frac{f(a+th+\Delta h) - f(a+th)}{\Delta} \\ &= \lim_{\Delta \to 0} \frac{\mathrm{D}f(a+th)(\Delta h) + o(\|\Delta h\|)}{\Delta} \\ &= \mathrm{D}f(a+th)(h). \end{aligned}$$

Podle Lagrangeovy věty o střední hodnotě existuje takové $t_0 \in (0,1)$, že $F(1) - F(0) = F'(t_0)$. Odtud

$$f(b) - f(a) = F(1) - F(0) = F'(t_0) = \mathrm{D}f(a+t_0h)(h) = \mathrm{D}f(\zeta)(h),$$

kde $\zeta = a + t_0 h$. □

Řekneme, že otevřená množina D v $\mathbf{R}^m$ je *souvislá*, když lze každé její dva body spojit lomenou čarou, která celá leží v D. Například koule s jednotkovým poloměrem v $\mathbf{R}^m$, celé $\mathbf{R}^m$ a $\mathbf{R}^3 \setminus L$, kde L je sjednocení konečně mnoha přímek, jsou souvislé otevřené množiny. Na druhou stranu množina $B \setminus R$, kde B je otevřená koule v $\mathbf{R}^3$ a R rovina protínající B, není souvislá.

Tvrzení 2.6. *Má-li reálná funkce m proměnných v každém bodě otevřené a souvislé množiny nulový diferenciál, je na této množině konstantní.*

Důkaz. Nechť $U \subset \mathbf{R}^m$ je otevřená a souvislá množina a funkce $f : U \to \mathbf{R}$ má na U nulový diferenciál. Vezmeme dva libovolné body $a, b \in U$ a spojíme je lomenou čarou $s = s_1 s_2 \dots s_r$ ležící v U. Pro libovolnou úsečku $s_i = a_i b_i$ z s máme podle předchozího tvrzení a předpokladu o f, že

$$f(a_i) - f(b_i) = \mathrm{D}f(\zeta)(a_i - b_i) = 0$$

(ζ je nějaký vnitřní bod s_i), tedy $f(a_i) = f(b_i)$. Hodnoty funkce f na koncích všech úseček s_i se rovnají a tedy $f(a) = f(b)$. □

Tvrzení 2.3, 2.4 a 2.6 dávají následující důsledek.

Důsledek. *Má-li reálná funkce m proměnných v každém bodě otevřené a souvislé množiny každou parciální derivaci nulovou, je na této množině konstantní.*

Počítání s parciálními derivacemi a diferenciály. Pro dvě funkce $f, g : U \to \mathbf{R}$, které jsou definované na okolí $U \subset \mathbf{R}^m$ bodu $a \in U$ a mají v bodě a

i-tou parciální derivaci, máme pro i-tou parciální derivaci jejich lineární kombinace, součinu a podílu stejné vzorce jako v případě funkcí jedné proměnné (místo $\frac{\partial f}{\partial x_i}$ píšeme $\partial_i f$):

$$\begin{aligned}\partial_i(\kappa f + \lambda g)(a) &= \kappa \partial_i f(a) + \lambda \partial_i g(a) \\ \partial_i(fg)(a) &= g(a)\partial_i f(a) + f(a)\partial_i g(a) \\ \partial_i(f/g)(a) &= \frac{g(a)\partial_i f(a) - f(a)\partial_i g(a)}{g(a)^2} \quad \text{(pokud } g(a) \neq 0\text{).}\end{aligned}$$

Tyto vzorce fakticky jsou vzorce pro funkce jedné proměnné, protože ∂_i se počítá z funkce závisející jen na x_i. Podobně pro diferenciály.

Tvrzení 2.7. *Nechť $U \subset \mathbf{R}^m$ je otevřená množina, $a \in U$ a $f, g : U \to \mathbf{R}$ jsou dvě funkce, obě diferencovatelné v bodě a.*
1. *Pro všechny $\kappa, \lambda \in \mathbf{R}$ je i funkce $\kappa f + \lambda g$ v bodu a diferencovatelná a*

$$D(\kappa f + \lambda g)(a) = \kappa Df(a) + \lambda Dg(a).$$

2. *Součinová funkce fg je diferencovatelná v a a*

$$D(fg)(a) = g(a)Df(a) + f(a)Dg(a).$$

3. *Pokud $g(a) \neq 0$, je podílová funkce f/g diferencovatelná v a a*

$$D(f/g)(a) = \frac{1}{g(a)^2}\Big(g(a)Df(a) - f(a)Dg(a)\Big).$$

Důkaz. Tyto vzorce plynou z analogických vzorců pro parciální derivace a z Tvrzení 2.3. (Viz úlohu 2.) □

Vzorec pro diferenciál lineární kombinace v části 1 platí obecněji i pro zobrazení $f, g : U \to \mathbf{R}^n$.

Podíváme se na diferenciál složeného zobrazení. V následující větě budeme skládání funkcí a zobrazení zapisovat v pořadí zprava doleva podle pořadí aplikace: $(g \circ f)(x) = g(f(x))$.

Věta 2.8. *Nechť*

$$f : U \to V, \ g : V \to \mathbf{R}^k$$

jsou dvě zobrazení, kde $U \subset \mathbf{R}^m$ a $V \subset \mathbf{R}^n$ jsou otevřené množiny. Je-li zobrazení f diferencovatelné v bodě a z U a g je diferencovatelné v bodě $b = f(a)$ z V, je složené zobrazení

$$g \circ f = g(f) : U \to \mathbf{R}^k$$

diferencovatelné v bodě a a jeho diferenciál se rovná složenině diferenciálů zobrazení f a g:

$$D(g \circ f)(a) = Dg(b) \circ Df(a).$$

Než se pustíme do důkazu věty, připomeneme význam symbolů $o(h)$ a $O(h)$ a explicitně uvedeme jejich jednoduché vlastnosti, které v důkazu využijeme.

Pro zobrazení $z : U \to \mathbf{R}^n$ definované v okolí počátku $U \subset \mathbf{R}^m$ budeme psát stručně $z(x) = o(x)$ místo $\|z(x)\| = o(\|x\|)$ a $z(x) = O(x)$ místo $\|z(x)\| = O(\|x\|)$, bereme vždy $x \to \overline{0}$. Značení $z(x) = o(x)$ je zkratka pro

$$\forall \varepsilon > 0 \; \exists \delta > 0 : \; \|x\| < \delta \Rightarrow \|z(x)\| < \varepsilon \|x\|$$

a $z(x) = O(x)$ pro

$$\exists c > 0 \; \exists \delta > 0 : \; \|x\| < \delta \Rightarrow \|z(x)\| < c\|x\|.$$

Lemma. *Nechť $z_1, z_2 : U \to \mathbf{R}^n$, kde $U \subset \mathbf{R}^m$ je okolí počátku, jsou dvě zobrazení. Nechť $u : U \to V$ a $v : V \to \mathbf{R}^k$ jsou dvě zobrazení, přičemž $U \subset \mathbf{R}^m$ a $V \subset \mathbf{R}^n$ jsou okolí počátku. V následujících tvrzeních $x \to \overline{0}$.*

1. *Když je z_1 lineární, potom $z_1(x) = O(x)$.*
2. *Když je $z_1(x) = o(x)$ a $z_2(x) = o(x)$, potom $z_1(x) + z_2(x) = o(x)$.*
3. *Když je $z_1(x) = o(x)$ a $z_2(x) = O(x)$, potom $z_1(x) + z_2(x) = O(x)$.*
4. *Pokud $u(x) = o(x)$ a $v = O(x)$, pak $v(u(x)) = o(x)$.*
5. *Pokud $u(x) = O(x)$ a $v(x) = o(x)$, pak $v(u(x)) = o(x)$.*

Důkaz. Úlohy 3 a 4. □

Důkaz věty 2.8. V okolí počátků souřadnic máme aproximace

$$g(b+h) = g(b) + \mathrm{D}g(b)(h) + \gamma(h) \quad \text{a} \quad f(a+h) = f(a) + \mathrm{D}f(a)(h) + \beta(h),$$

kde $\gamma(h)$ a $\beta(h)$ jsou $o(h)$. Rozdíl $f(a+h) - f(a)$ si označíme jako $\Delta(h)$. Pak $f(a+h) = f(a) + \Delta(h) = b + \Delta(h)$ a $\Delta(h) = \mathrm{D}f(a)(h) + \beta(h)$. Takže

$$\begin{aligned}
(g \circ f)(a+h) - (g \circ f)(a) &= g(f(a+h)) - g(f(a)) \\
&= g(b + \Delta(h)) - g(b) \\
&= \mathrm{D}g(b)(\Delta(h)) + \gamma(\Delta(h)) \\
&= \mathrm{D}g(b)(\mathrm{D}f(a)(h)) + \mathrm{D}g(b)(\beta(h)) + \gamma(\Delta(h)) \\
&= (\mathrm{D}g(b) \circ \mathrm{D}f(a))(h) + \alpha(h),
\end{aligned}$$

kde

$$\alpha(h) = \mathrm{D}g(b)(\beta(h)) + \gamma(\Delta(h)) = \mathrm{D}g(b)(\beta(h)) + \gamma(\mathrm{D}f(a)(h) + \beta(h)).$$

První sčítanec definující $\alpha(h)$ je $o(h)$ podle částí 1 a 4 lemmatu a druhý je také $o(h)$ podle částí 1, 3 a 5. Celkem $\alpha(h) = o(h)$ podle části 2. Vidíme, že $g \circ f$ má v a diferenciál rovný lineárnímu zobrazení $\mathrm{D}g(b) \circ \mathrm{D}f(a)$. $\square$

Z lineární algebry víme, že matice lineárního zobrazení $g \circ f$ složeného z lineárních zobrazení f a g se dostane jako součin matice zobrazení g a matice zobrazení f (v tomto pořadí). Jacobiho matice zobrazení f v bodě a je matice lineárního zobrazení $\mathrm{D}f(a)$ vzhledem ke kanonické bázi a její prvky jsou hodnoty parciálních derivací souřadnicových funkcí v bodě a. Pomocí matic tak větu 2.8 vyjádříme následovně.

Důsledek. *Za situace popsané v předchozí větě je Jacobiho matice složeného zobrazení $h = g \circ f$ v bodě a rovna součinu Jacobiho matice zobrazení g v bodě $b = f(a)$ a Jacobiho matice zobrazení f v bodě a:*

$$\begin{aligned}\left(\frac{\partial h_i}{\partial x_j}(a)\right)_{i,j=1}^{k,m} &= \left(\frac{\partial g_i}{\partial x_j}(b)\right)_{i,j=1}^{k,n} \left(\frac{\partial f_i}{\partial x_j}(a)\right)_{i,j=1}^{n,m} \\ &= \left(\sum_{r=1}^{n} \frac{\partial g_i}{\partial x_r}(b) \cdot \frac{\partial f_r}{\partial x_j}(a)\right)_{i,j=1}^{k,m}.\end{aligned}$$

Speciálně pro $k = 1$, kdy funkce h o m proměnných je složeninou

$$h = g(f_1, f_2, \ldots, f_n)$$

funkce g o n proměnných a n funkcí $f_i = f_i(x_1, x_2, \ldots, x_m)$, dostáváme *řetízkové pravidlo* pro parciální derivaci složené funkce:

$$\begin{aligned}\frac{\partial h}{\partial x_i}(a) &= \sum_{j=1}^{n} \frac{\partial g}{\partial x_j}(f(a)) \cdot \frac{\partial f_j}{\partial x_i}(a) \\ &= \langle \nabla g(f(a)), \partial_i f(a) \rangle,\end{aligned}$$

kde $i = 1, 2, \ldots, m$, $f = (f_1, f_2, \ldots, f_n)$ a $\partial_i f = (\partial_i f_1, \partial_i f_2, \ldots, \partial_i f_n)$.

Úlohy

1. Zobecněte důkaz Tvrzení 2.4 na více než dvě proměnné.
2. Rozmyslete si důkaz Tvrzení 2.7.
3. Dokažte části 1–3 lemmatu.
4. Dokažte části 4 a 5 lemmatu.
5. Lemma můžeme zobecnit na zobrazení mezi normovanými prostory (které jako vektorové prostory nemusejí už mít konečnou dimenzi). Pak ale jedna z částí 1–5 obecně přestane platit. Která?