7. přednáška 15. listopadu 2005

Viděli jsme, že diferencovatelnost implikuje existenci parciálních derivací. Nyní dokážeme, že naopak existence parciálních derivací v okolí bodu a jejich spojitost v něm implikují diferencovatelnost.

Tvrzení 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. (Na přednášce řečena jen hlavní myšlenka.) Můžeme předpokládat, že $a = \overline{0}$ je počátek souřadnic. Pro $h \in \mathbf{R}^m$, $\|h\| \to 0$, máme $\partial_i f(h) = \partial_i f(\overline{0}) + o(1)$ (spojitost $\partial_i f$ v počátku). Počátek a bod h spojuje jednoznačně určená lomená čára $s = s_1 \dots s_m$, jejíž i-tá úsečka $s_i = a_i b_i$ má směr i-té souřadnicové osy x_i:

$$a_i = (h_1, h_2, \dots, h_{i-1}, 0, 0, \dots, 0) \text{ a } b_i = (h_1, h_2, \dots, h_i, 0, 0, \dots, 0),$$

takže $a_1 = \overline{0}$, $b_i = a_{i+1}$ pro $1 \le i \le m-1$ a $b_m = h$. Na s_i se f chová jako funkce jedné proměnné x_i, s derivací rovnou $\partial_i f$. Podle Lagrangeovy věty o střední hodnotě máme

$$f(b_i) - f(a_i) = \partial_i f(\zeta_i) h_i,$$

kde ζ_i je nějaký bod ležící uvnitř s_i. Odtud

$$\begin{aligned} f(h) - f(\overline{0}) &= \sum_{i=1}^m f(b_i) - f(a_i) \\ &= \sum_{i=1}^m \partial_i f(\zeta_i) h_i \\ &= \sum_{i=1}^m (\partial_i f(\overline{0}) + o(1)) h_i \\ &= \sum_{i=1}^m \partial_i f(\overline{0}) h_i + o(\|h\|) \end{aligned}$$

($\|h\| \le |h_1| + \cdots + |h_m|$), pročež je f diferencovatelná v počátku. □

Tvrzení 5 zobecňuje Lagrangeovu větu o střední hodnotě na funkce více proměnných a důsledek 6 zobecňuje fakt, že nulovost derivace implikuje (za jistých předpokladů) konstantnost funkce.

Tvrzení 5. *Nechť $U \subset \mathbf{R}^m$ je otevřená množina, $u = ab$ je úsečka ležící v U a funkce $f : U \to \mathbf{R}$ je spojitá na úsečce u a diferencovatelná v každém jejím bodě, s možnou výjimkou krajních bodů a a b. Pak existuje takový vnitřní bod ζ úsečky u, 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 $t_0 \in (0,1)$ takové, 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 $\zeta = a + t_0 h$. □

Důsledek 6. *Pokud reálná funkce m proměnných má 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 $f : U \to \mathbf{R}$ je v každém bodě $a \in U$ diferencovatelná a $\mathrm{D}f(a)$ je vždy nulové (lineární) zobrazení. Vezmeme dva libovolné body $a, b \in U$ a spojíme je lomenou čarou $s = s_1 s_2 \dots s_r$ ležící v U (což je možné, viz věta 15 v kapitole 1 a následující cvičení). Pro libovolnou úsečku $s_i = a_i b_i$ z s máme podle předchozího tvrzení a předpokladu o f, že

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

(zde ζ je nějaký vnitřní bod s_i), tedy $f(a_i) = f(b_i)$. Hodnoty funkce f na koncích všech úseček s_i jsou proto všechny stejné a speciálně $f(a) = f(b)$. □

Geometrie gradientu a diferenciálu. Mějme funkci $f : U \to \mathbf{R}$ o m proměnných, která je definovaná na nějakém okolí U bodu $a \in \mathbf{R}^m$. Zkoumáme její okamžité přírůstky

$$\mathrm{D}_v f(a) = \lim_{t \to 0} \frac{f(a+tv) - f(a)}{t}$$

v bodě a ve směrech jednotkových vektorů $v \in \mathbf{R}^m$, $\|v\| = 1$. Zajímá nás, pro který směr je přírůstek co největší. Když je f v a diferencovatelná, pak podle tvrzení 2 a věty 1 máme

$$|\mathrm{D}_v f(a)| = |\mathrm{D}f(a)(v)| = |\langle \nabla f(a), v \rangle| \le \|\nabla f(a)\| \cdot \|v\| = \|\nabla f(a)\|.$$

Podle věty 1 se rovnost nabývá, právě když v je skalárním násobkem $\nabla f(a)$, to jest právě pro dva vektory

$$v^+ = \frac{\nabla f(a)}{\|\nabla f(a)\|} \ \text{ a } \ v^- = -\frac{\nabla f(a)}{\|\nabla f(a)\|}.$$

Ve směru v^+ svého normovaného gradientu tedy f roste nejrychleji a v opačném směru v^- stejnou měrou nejrychleji klesá:

$$\mathrm{D}_{v^+} f(a) = \langle \nabla f(a), v^+ \rangle = \|\nabla f(a)\| \ \text{ a } \ \mathrm{D}_{v^-} f(a) = \langle \nabla f(a), v^- \rangle = -\|\nabla f(a)\|.$$

Nyní zobecníme pojem tečny ke grafu funkce jedné proměnné na (nad)rovinu tečnou ke grafu funkce více proměnných. Pro jednoduchost značení se omezíme jen na případ tečné roviny a dvou proměnných; obecná tečná nadrovina ke grafu funkce m proměnných se zavádí analogicky.

Nechť $(x_0, y_0) \in D \subset \mathbf{R}^2$, kde D je otevřená množina v rovině, a $f : D \to \mathbf{R}$ je funkce. Její graf

$$P = \{(x, y, z) \in \mathbf{R}^3 \mid (x, y) \in D, z = f(x, y)\}$$

je plocha v 3-rozměrném euklidovském prostoru; P obsahuje bod (x_0, y_0, z_0), kde $z_0 = f(x_0, y_0)$. Předpokládejme, že funkce f je v bodě (x_0, y_0) diferencovatelná. Tvrdíme, že potom mezi všemi rovinami $z = L(x, y)$ (L je afinní funkce dvou proměnných), které obsahují bod (x_0, y_0, z_0), je pouze jediná splňující (pro $(x, y) \to (x_0, y_0)$) aproximaci

$$f(x, y) = L(x, y) + o(\sqrt{(x - x_0)^2 + (y - y_0)^2}),$$

totiž rovina

$$T(x, y) = z_0 + \frac{\partial f}{\partial x}(x_0, y_0) \cdot (x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0) \cdot (y - y_0).$$

To plyne hned z existence diferenciálu a jeho jednoznačnosti, protože zřejmě $\mathrm{D}f(x_0, y_0)(x, y) = T(x, y) - z_0$. Graf funkce $T(x, y)$,

$$T = \{(x, y, z) \in \mathbf{R}^3 \mid (x, y) \in D, z = T(x, y)\}$$

se nazývá *tečnou rovinou ke grafu funkce* f *v bodě* (x_0, y_0, z_0).

Rovnici tečné roviny $z = T(x, y)$ přepíšeme ve tvaru

$$\begin{aligned} \frac{\partial f}{\partial x}(x_0, y_0) \cdot (x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0) \cdot (y - y_0) \ - (z - z_0) &= 0 \\ \langle V, (x - x_0, y - y_0, z - z_0) \rangle &= 0, \end{aligned}$$

kde $V \in \mathbf{R}^3$ je vektor

$$V = \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0), -1 \right).$$

Označíme-li $X = (x, y, z)$ a $X_0 = (x_0, y_0, z_0)$, můžeme tečnou rovinu T definovat jako

$$T = \{X \in \mathbf{R}^3 \mid \langle V, X - X_0 \rangle = 0\}.$$

Je tedy tvořena právě těmi body, jejichž směrové vektory k bodu X_0 jsou kolmé na V. Vektor V se nazývá *normálovým vektorem ke grafu funkce* f *v bodě* X_0.

2.3. Počítání s diferenciály a parciálními derivacemi. 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.

Tvrzení 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* $\kappa f + \lambda g$ *diferencovatelná v* a *a její diferenciál v* a *se rovná*

$$\kappa \mathrm{D} f(a) + \lambda \mathrm{D} g(a).$$

2. *Součinová funkce fg je diferencovatelná v a a její diferenciál v a se rovná*

$$g(a)\mathrm{D}f(a) + f(a)\mathrm{D}g(a).$$

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

$$\frac{1}{g(a)^2}\Big(g(a)\mathrm{D}f(a) - f(a)\mathrm{D}g(a)\Big).$$

Důkaz. Tyto vzorce plynou z analogických vzorců pro parciální derivace a z 3 tvrzení 2. $\square$

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

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

Věta 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. Předpokládejme, že zobrazení f je diferencovatelné v bodě $a \in U$ a g je diferencovatelné v bodě $b = f(a) \in V$. Potom složené zobrazení

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

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

$$\mathrm{D}(g \circ f)(a) = \mathrm{D}g(b) \circ \mathrm{D}f(a).$$

Speciální případ této věty ($m = 1$ a $k = 1$) jsme už vlastně dokázali na 6. přednášce v příkladu 3 o částici, takže víme, jak na to. Udělejme teď ale pořádný důkaz. (Následující pasáž až po str. 7, důkaz věty 8, z časových důvodů nebyla na přednášce.)

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

$\|z(x)\| = O(\|x\|)$; zde bereme vždy $x \to \overline{0}$. Značení $z(x) = o(x)$ je tedy zkratka pro

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

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

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

Připomeneme několik jednoduchých vlastností symbolů o a O, které ponecháme čtenáři jako cvičení. Pokud $z(x)$ je lineární zobrazení, pak $z(x) = O(x)$. (Tato samozřejmost obecně neplatí ve vektorových prostorech s nekonečně mnoha rozměry!) Pokud $z(x) = o(x)$, pak i $z(x) = O(x)$. Pokud máme dvě taková zobrazení $z_1, z_2 :\ U \to \mathbf{R}^n$ a $z = z_1 + z_2$, potom $(i = 1, 2)$

$$z_i(x) = o(x) \Rightarrow z(x) = o(x) \ \text{ a } \ z_i(x) = O(x) \Rightarrow z(x) = O(x).$$

Speciálně, $z_1(x) = o(x)$ a $z_2(x) = O(x)$ implikují $z(x) = O(x)$. Obdobné vlastnosti symbolů o a O při skládání zobrazení se dokazují rovněž snadno, ale stojí za zdůraznění—jsou klíčové v důkazu věty 8.

Lemma. *Nechť $u :\ U \to V$ a $v :\ V \to \mathbf{R}^k$ jsou zobrazení, přičemž $U \subset \mathbf{R}^m$ a $V \subset \mathbf{R}^n$ jsou okolí počátků souřadnic.*
1. *Pokud $u(x) = o(x)$ a $v = O(x)$, pak $v(u(x)) = o(x)$.*
2. *Pokud $u(x) = O(x)$ a $v(x) = o(x)$, pak $v(u(x)) = o(x)$.*

Důkaz. Cvičení. □

Důkaz věty 8. V okolí počátků souřadnic máme aproximace

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

kde $\gamma(h)$ i $\beta(h)$ je $o(h)$. 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

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

kde

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

První sčítanec definující $\alpha(h)$ je $o(h)$ podle 1 lemmatu a druhý je $o(h)$ podle 2 lemmatu. Celkem $\alpha(h) = o(h)$ a $g \circ f$ 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 (důsledek 3). Na úrovni popisu Jacobiho maticemi tak věta 8 dává následující důsledek.

Důsledek 9. *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, z nichž každá má m proměnných, 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 $\partial_i f = (\partial_i f_1, \partial_i f_2, \ldots, \partial_i f_n)$.