7. přednáška 12. listopadu 2007

Geometrie diferenciálu a parciálních derivací. Nechť $U \subset \mathbf{R}^m$ je okolí bodu a a $f: U \to \mathbf{R}$ je funkce o m proměnných. Její okamžitý růst v bodu a ve směru v, kde v je jednotkový vektor (tj. $\|v\| = 1$) z $\mathbf{R}^m$, je dán směrovou derivací

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

V jakém směru roste funkce nejrychleji? Když je f v a diferencovatelná, pak podle tvrzení 2.3 a Cauchyovy–Schwarzovy nerovnosti (věta 2.1)

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

a rovnost se 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:

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

(Přesně řečeno, tohle je pravda, pokud je gradient $\nabla f(a)$ nenulový vektor. Je-li to nulový vektor, pak má f ve všech směrech nulový růst.)

Zobecníme pojem tečny ke grafu funkce jedné proměnné na (nad)rovinu tečnou ke grafu funkce více proměnných. Pro jednoduchost značení se omezíme 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 třírozměrném euklidovském prostoru. Na P leží 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. 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 $\mathrm{D}f(x_0, y_0)(x - x_0, y - y_0) = T(x, y) - z_0$. 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 z $\mathbf{R}^3$ 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*.

Parciální derivace vyšších řádů. Pokud má funkce $f : U \to \mathbf{R}$ definovaná na otevřené množině $U \subset \mathbf{R}^m$ v každém jejím bodě parciální derivaci $F = \partial_i f$ a ta má má v bodě $a \in U$ parciální derivaci $\partial_j F(a) = \partial_j \partial_i f(a)$, řekneme, že f má v bodě *a parciální derivaci druhého řádu podle proměnných x_i a x_j* a její hodnotu značíme

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

Induktivně definujeme parciální derivace vyšších řádů: má-li f v každém bodě $x \in U$ parciální derivaci

$$F = \frac{\partial^{k-1} f}{\partial x_{i_{k-1}} \partial x_{i_{k-2}} \dots \partial x_{i_1}}(x)$$

a ta má v bodě $a \in U$ parciální derivaci $\partial_j F(a)$, řekneme, že f má v bodě *a parciální derivaci k-tého řádu podle proměnných $x_{i_1}, \dots, x_{i_{k-1}}, x_j$* a její hodnotu značíme

$$\frac{\partial^k f}{\partial x_j \partial x_{i_{k-1}} \dots \partial x_{i_1}}(a).$$

Na pořadí proměnných pří parciálním derivování obecně záleží, jak ukazuje příklad funkce

$$f(x, y) = \begin{cases} \frac{xy(x^2 - y^2)}{x^2 + y^2} & \text{pro } x^2 + y^2 \neq 0 \\ 0 & \text{pro } x^2 + y^2 = 0, \end{cases}$$

která má v počátku obě smíšené parciální derivace druhého řádu, ale s různými hodnotami:

$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) = 1 \neq -1 = \frac{\partial^2 f}{\partial y \partial x}(0, 0)$$

(úloha 1). Nicméně při spojitých parciálních derivacích na pořadí proměnných nezáleží.

Tvrzení 2.9. *Nechť funkce $f : U \to \mathbf{R}$ má na okolí $U \subset \mathbf{R}^m$ bodu a parciální derivace druhého řádu $\partial_j \partial_i f$ a $\partial_i \partial_j f$ a ty jsou v a spojité. Potom*

$$\partial_j \partial_i f(a) = \partial_i \partial_j f(a).$$

Důkaz. Pro jednoduchost buď $m = 2$ a $a = (0,0)$. Díky spojitosti obou parciálních derivací v počátku stačí nalézt pro každé $h > 0$ ve čtverci $[0,h]^2$ dva body σ a τ, v nichž $\partial_x \partial_y f(\sigma) = \partial_y \partial_x f(\tau)$.

Vrcholy čtverce označíme $a = (0,0)$, $b = (0,h)$, $c = (h,0)$, $d = (h,h)$ a uvážíme číslo $f(d) - f(b) - f(c) + f(a)$. Lze ho dvěma způsoby napsat jako rozdíl rozdílů:

$$\begin{aligned} f(d) - f(b) - f(c) + f(a) &= (f(d) - f(b)) - (f(c) - f(a)) = \psi(h) - \psi(0) \\ &= (f(d) - f(c)) - (f(b) - f(a)) = \phi(h) - \phi(0), \end{aligned}$$

kde

$$\psi(t) = f(h,t) - f(0,t) \ \text{ a } \ \phi(t) = f(t,h) - f(t,0).$$

Máme $\psi'(t) = \partial_y f(h,t) - \partial_y f(0,t)$ a $\phi'(t) = \partial_x f(t,h) - \partial_x f(t,0)$. Lagrangeova věta o střední hodnotě dává vyjádření

$$\begin{aligned} f(d) - f(b) - f(c) + f(a) &= \psi'(t_0)h = (\partial_y f(h,t_0) - \partial_y f(0,t_0))h \\ &= \phi'(s_0)h = (\partial_x f(s_0,h) - \partial_x f(s_0,0))h, \end{aligned}$$

kde $0 < s_0, t_0 < h$. Použijeme ji ještě jednou na rozdíly parciálních derivací f a máme

$$f(d) - f(b) - f(c) + f(a) = \partial_x \partial_y f(s_1,t_0)h^2 = \partial_y \partial_x f(s_0,t_1)h^2, \ s_1, t_1 \in (0,h).$$

Body $\sigma = (s_1,t_0)$ a $\tau = (s_0,t_1)$ leží ve čtverci $[0,h]^2$ a $\partial_x \partial_y f(\sigma) = \partial_y \partial_x f(\tau)$. □

Rovnost hodnot obou derivací lze dokázat i za slabšího předpokladu: existuje-li $\partial_x \partial_y f$ v okolí bodu a a je v něm spojitá, potom existuje i $\partial_y \partial_x f(a)$ a $\partial_y \partial_x f(a) = \partial_x \partial_y f(a)$.

Pro otevřenou množinu $U \subset \mathbf{R}^m$ označíme symbolem $\mathcal{C}^k(U)$ množinu funkcí $f : U \to \mathbf{R}$, jejichž parciální derivace až do řádu k včetně jsou na U definované a spojité.

Důsledek. *Pro každou funkci $f = f(x_1, x_2, \ldots, x_m)$ z $\mathcal{C}^k(U)$ hodnoty jejích parciálních derivací až do řádu k nezávisí na pořadí proměnných—pro $l \le k$ a $a \in U$ platí*

$$\frac{\partial^l f}{\partial x_{i_l} \partial x_{i_{l-1}} \ldots \partial x_{i_1}}(a) = \frac{\partial^l f}{\partial x_{j_l} \partial x_{j_{l-1}} \ldots \partial x_{j_1}}(a),$$

jakmile je posloupnost $(i_1, \ldots, i_l)$ permutací posloupnosti $(j_1, \ldots, j_l)$.

Důkaz. Když je posloupnost $v = (j_1, \ldots, j_l)$ pouze permutací posloupnosti $u = (i_1, \ldots, i_l)$, dokážeme u transformovat ve v prohazováním dvojic členů v u, dokonce vystačíme s prohazováním sousedních členů: v u nalezneme člen j_1 a necháme ho „propadnout“ až dolů, pak necháme propadnout na správné místo j_2 atd. Rovnost hodnot parciálních derivací tak plyne z tvrzení 2.9. □

V případě spojitých parciálních derivací tak záleží jen na multimnožině proměnných, podle kterých se derivuje, ale ne na jejich pořadí. Místo $\partial_x \partial_x$ píšeme stručně ∂x^2 apod. Například, pro f z $C^5(U)$ máme

$$\frac{\partial^5 f}{\partial y\,\partial x\,\partial y\,\partial y\,\partial z} = \frac{\partial^5 f}{\partial y^2\,\partial x\,\partial z\,\partial y} = \frac{\partial^5 f}{\partial x\,\partial z\,\partial y^3} = \cdots$$

Důležitým nástrojem při studiu vlastností funkcí je Taylorův rozvoj, jehož verzi pro více proměnných nyní odvodíme. Jak rozumět použitému symbolickému zápisu diferenciálního operátoru vysvětlíme na příkladu, v němž $f = f(x,y,z)$ je funkce a $a \in \mathbf{R}^3$, $\alpha, \beta \in \mathbf{R}$ jsou konstanty. Zápisem

$$(\alpha\partial_y + \beta\partial_z)^3 f(a)$$

se rozumí

$$\begin{aligned}
& (\alpha^3(\partial_y)^3 + 3\alpha^2\beta(\partial_y)^2\partial_z + 3\alpha\beta^2\partial_y(\partial_z)^2 + \beta^3(\partial_z)^3)f(a) \\
= \; & \alpha^3\frac{\partial^3 f}{\partial y^3}(a) + 3\alpha^2\beta\frac{\partial^3 f}{\partial y^2 \partial z}(a) + 3\alpha\beta^2\frac{\partial^3 f}{\partial y \partial z^2}(a) + \beta^3\frac{\partial^3 f}{\partial z^3}(a).
\end{aligned}$$

Tvrzení 2.10. *Nechť $U \subset \mathbf{R}^m$ je otevřená množina, $a \in U$ je bod a $f : U \to \mathbf{R}$ je funkce z $\mathcal{C}^n(U)$. Potom v okolí bodu a máme Taylorův rozvoj*

$$\begin{aligned}
f(a+h) &= \sum_{i=0}^{n} \frac{1}{i!}(h_1\partial_1 + h_2\partial_2 + \cdots + h_m\partial_m)^i f(a) + o(\|h\|^n) \\
&= \sum \frac{1}{i_1!\ldots i_m!} \cdot \frac{\partial^{i_1+\cdots+i_m} f}{\partial x_1^{i_1}\ldots\partial x_m^{i_m}}(a) \cdot h_1^{i_1}\ldots h_m^{i_m} + o(\|h\|^n) \\
&= f(a) + \sum_{i=1}^{m}\partial_{x_i} f(a) h_i + \sum_{i<j}\partial_{x_i}\partial_{x_j} f(a) h_i h_j + \frac{1}{2}\sum_{i=1}^{m}\partial^2_{x_i} f(a) h_i^2 + \\
&\quad + \cdots + o(\|h\|^n).
\end{aligned}$$

V první sumě mocninu chápeme symbolicky (ve smyslu operátorového počtu) a ve druhé sumě sčítáme přes všechny m-tice nezáporných celých čísel $i_1, i_2, \ldots, i_m$, jejichž součet je nanejvýš n.

Důkaz. Vezmeme Taylorův rozvoj až do řádu n pomocné funkce jedné proměnné $F(t) = f(a+th)$, kde $t \in [0,1]$. Opakovaným použitím řetízkového pravidla ($F = f \circ l$, kde l je lineární zobrazení, přímka $l(t) = a + th$) pro $k \le n$ dostáváme

$$F^{(k)}(t) = \sum_{i_1, i_2, \ldots, i_k} \frac{\partial^k f}{\partial x_{i_1}\,\partial x_{i_2}\ldots\partial x_{i_k}}(a+th) \cdot h_{i_1} h_{i_2} \ldots h_{i_k},$$

kde $i_1, \ldots, i_k$ probíhají nezávisle na sobě všechny indexy $1, 2, \ldots, m$. Dosazením do Taylorova rozvoje funkce F se zbytkem v Lagrangeově tvaru

$$f(a+h) = F(1) = \sum_{i=0}^{n-1} \frac{1}{i!} F^{(i)}(0) + \frac{F^{(n)}(\theta)}{n!},\ 0 < \theta < 1,$$

dostáváme, s využitím kompaktního symbolického zápisu parciálních derivací, první formuli pro $f(a+h)$. Druhá formule vyplývá z první pomocí multinomické věty:

$$(h_1\partial_1 + h_2\partial_2 + \cdots + h_m\partial_m)^i = \sum_{i_1, i_2, \ldots, i_m} \binom{i}{i_1, i_2, \ldots, i_m} \prod_{j=1}^{m} (h_j\partial_j)^{i_j},$$

kde $i_1, i_2, \ldots, i_m$ probíhají nezáporná celá čísla se součtem i a

$$\binom{i}{i_1, i_2, \ldots, i_m} = \frac{i!}{i_1! \cdot i_2! \cdot \ldots \cdot i_m!}$$

je multinomický koeficient. □

Sčítance odpovídající $i = 0$ a 1 jsou $f(a)$ a $\mathrm{D}f(a)(h)$. Taylorova formule zobecňuje lokální aproximaci pomocí diferenciálu, kterou dostáváme pro $n = 1$.

Extrémy funkcí více proměnných. Symetrická (tj. $a_{i,j} = a_{j,i}$) reálná $n \times n$ matice $A \in \mathbf{R}^{n \times n}$ definuje kvadratickou formu

$$P(x_1, x_2, \ldots, x_n) = xAx^T = \sum_{i,j=1}^{n} a_{i,j} x_i x_j :\ \mathbf{R}^n \to \mathbf{R}$$

(x je řádkový vektor $(x_1, x_2, \ldots, x_n)$). Připomeňme si, že A se nazývá

- *pozitivně (negativně) definitní*, když $P(x) > 0$ ($P(x) < 0$) pro všechny $x \in \mathbf{R}^n \backslash \{\overline{0}\}$;
- *pozitivně (negativně) semidefinitní*, když $P(x) \geq 0$ ($P(x) \leq 0$) pro všechny $x \in \mathbf{R}^n$;
- *indefinitní*, není-li ani pozitivně ani negativně semidefinitní, tj. $P(x) > 0$ a $P(y) < 0$ pro nějaké dva vektory $x, y \in \mathbf{R}^n$.

Hessova matice funkce f v bodě a, kde $f :\ U \to \mathbf{R}$ je definovaná na okolí $U \subset \mathbf{R}^m$ bodu a a má na U všechny derivace druhého řádu, je matice zaznamenávající hodnoty těchto derivací:

$$H_f(a) := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)_{i,j=1}^{m}.$$

Podle tvrzení 2.9 je Hessova matice funkcí z $\mathcal{C}^2(U)$ symetrická.

Odvodíme kritérium existence lokálních extrémů funkcí m proměnných, které zobecňuje výsledek pro funkce jedné proměnné. Roli hodnoty druhé derivace přebírá Hessova matice. Připomeňme si, že funkce $f : U \to \mathbf{R}$, kde $U \subset \mathbf{R}^m$ je otevřená množina, má v bodě $a \in U$ ostré lokální minimum, pokud existuje $\delta > 0$ takové, že $0 < \|x - a\| < \delta$ implikuje $f(x) > f(a)$. (Neostré) lokální minimum znamená, že $\|x - a\| < \delta$ implikuje $f(x) \geq f(a)$. Podobně pro ostré a neostré lokální maximum. Funkce f nemá v a ani neostrý lokální extrém, nemá-li v tomto bodě ani lokální neostré minimum ani lokální neostré maximum, to jest pro každé $\delta > 0$ existují body x, y takové, že $\|x - a\|, \|y - a\| < \delta$ a $f(x) > f(a)$, $f(y) < f(a)$.

Věta 2.11. *Nechť $f \in C^2(U)$, kde $U \subset \mathbf{R}^m$ je otevřená množina, a $a \in U$ je bod.*

- *Pokud $\nabla f(a) \neq \overline{0}$, nemá f v a ani neostrý lokální extrém.*
- *Pokud $\nabla f(a) = \overline{0}$ a $H_f(a)$ je pozitivně (negativně) definitní, potom má f v a ostré lokální minimum (maximum).*
- *Pokud $\nabla f(a) = \overline{0}$ a $H_f(a)$ je indefinitní, nemá f v a ani neostrý lokální extrém.*

Úlohy

1. Ověřte, že uvedená funkce $f(x, y)$ má v počátku různé smíšené parciální derivace druhého řádu.