8. přednáška 22. listopadu 2005

Věta 8 je dalekosáhlým zobecněním formule pro derivaci složené funkce jedné proměnné. V tomto okamžiku se nabízí otázka, jak to je se zobecněním formule pro derivaci inverzní funkce. Dostaneme se k němu později, vyplyne jako důsledek věty o implicitní funkci.

Příklad. Ukážeme, jak v jisté situaci řetízkové pravidlo umožňuje odvodit zákon zachování energie. Uvažme otevřenou množinu $U \subset \mathbf{R}^m$ a zobrazení $F:\ U \to \mathbf{R}^m$. To přiřazuje každému bodu z U vektor z $\mathbf{R}^m$ a představujeme si ho jako vektorové pole na U, které v každém bodě $a \in U$ udává směr a velikost síly působící na hmotný bod nacházející se v a. Toto silové pole může být například polem gravitační síly nějakého tělesa. Oblastí U prolétá částice po dráze $\gamma:\ [0,1] \to U$ plně určené polem F. Pole F určuje její pohyb prostřednictvím Newtonova zákona síly (síla = hmotnost × zrychlení):

$$F(\gamma(t)) = m\gamma''(t).$$

Zde $m > 0$ je hmotnost částice, tedy konstanta, a $\gamma''(t) = (\gamma_1''(t), \ldots, \gamma_m''(t))$ je vektor zrychlení částice (zrychlení je okamžitá velikost změny rychlosti částice, čili druhá derivace její polohy). Předpokládejme dále, že existuje funkce $f:\ U \to \mathbf{R}$ taková, že má v každém bodě množiny U gradient a pro každé $a \in U$ platí

$$F(a) = -\nabla f(a).$$

Taková silová pole F se vyskytují často a říká se jim *konzervativní pole*; například pole gravitační síly je konzervativní. Fyzikové definují *potenciál* konzervativního pole F v bodě a jako $f(a)$, je to také *potenciálová energie* částice vzhledem k poli F, když se nachází v a. Dále definují *kinetickou energii* částice letící po dráze $\Gamma:\ [0,1] \to U$ (ne nutně určené polem F) jako

$$\tfrac{1}{2}mv(t)^2 = \tfrac{1}{2}m\Gamma'(t)^2 = \tfrac{1}{2}m\langle \Gamma'(t), \Gamma'(t)\rangle.$$

Zde $m > 0$ je hmotnost částice, $v(t)$ označuje její vektor rychlosti a kinetickou energii částice počítáme v okamžiku $t \in [0,1]$. Vyslovíme zákon zachování energie.

> V této situaci—částice o hmotnosti m se pohybuje v konzervativním silovém poli $F = -\nabla f$ po dráze $\gamma(t)$ určené Newtonovým zákonem síly—je součet potenciálové a kinetické energie částice konstantní, nezávislý na čase.

Abychom to dokázali, zderivujeme součet obou energií

$$S(t) = f(\gamma(t)) + \tfrac{1}{2}m\gamma'(t)^2$$

podle času t. S využitím řetízkového pravidla, Newtonova zákona síly a konzervativnosti silového pole F dostaneme, že

$$\begin{aligned} S'(t) &= \langle \nabla f(\gamma(t)), \gamma'(t)\rangle + m\langle \gamma'(t), \gamma''(t)\rangle \\ &= \langle \nabla f(\gamma(t)), \gamma'(t)\rangle + \langle \gamma'(t), m\gamma''(t)\rangle \\ &= \langle \nabla f(\gamma(t)), \gamma'(t)\rangle + \langle \gamma'(t), F(\gamma(t))\rangle \\ &= \langle \nabla f(\gamma(t)), \gamma'(t)\rangle + \langle \gamma'(t), -\nabla f(\gamma(t))\rangle \\ &= \langle \nabla f(\gamma(t)), \gamma'(t)\rangle - \langle \nabla f(\gamma(t)), \gamma'(t)\rangle \\ &= 0. \end{aligned}$$

Tudíž $S(t) = f(\gamma(t)) + \frac{1}{2}m\gamma'(t)^2 = const.$ pro t probíhající $[0,1]$.

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 parciálních derivací 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:

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

(spočítejte podrobně jako cvičení).

Tvrzení 10. *Nechť funkce $f : U \to \mathbf{R}$ má na okolí $U \subset \mathbf{R}^m$ bodu $a \in U$ parciální derivace druhého řádu $\partial_j \partial_i f$ a $\partial_i \partial_j f$ a obě jsou v a spojité. Potom*

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

Důkaz. Můžeme předpokládat, že $m = 2$ a $a = (0,0)$. Z předpokladu spojitosti $\partial_y \partial_x f$ a $\partial_x \partial_y f$ v $(0,0)$ chceme odvodit, že

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

Díky spojitosti stačí ukázat, že pro každé $h > 0$ čtverec $K = [0,h]^2$ obsahuje body σ a τ takové, že $\partial_y \partial_x f(\tau) = \partial_x \partial_y f(\sigma)$.

Pro orientovanou úsečku $u = \alpha\beta \subset U$ a funkci $g : U \to \mathbf{R}$ označíme $g(u) = g(\alpha\beta) := g(\beta) - g(\alpha)$. Vrcholy čtverce K označíme $a = (0,0)$, $b = (0,h)$, $c = (h,0)$, $d = (h,h)$ a uvážíme číslo

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

kde $\phi(t) = f(u_t)$, $\psi(t) = f(v_t)$ a u_t, v_t jsou pro $0 \leq t \leq h$ úsečky $u_t = (t,0)(t,h)$ a $v_t = (0,t)(h,t)$. Podle Lagrangeovy věty o střední hodnotě

$$\phi(h) - \phi(0) = \phi'(t_0)h = \partial_x f(u_{t_0})h$$

pro nějaké $0 < t_0 < h$. Ovšem, znovu podle Lagrangeovy věty o střední hodnotě,

$$\partial_x f(u_{t_0}) = \partial_x f(t_0, h) - \partial_x f(t_0, 0) = \partial_y \partial_x f(t_0, t_1)h$$

pro nějaké $0 < t_1 < h$. Pro $\tau = (t_0, t_1)$ tedy máme $w = \partial_y \partial_x f(\tau) h^2$. Stejná úvaha aplikovaná na vyjádření $w = \psi(h) - \psi(0)$ dává $w = \partial_x \partial_y f(\sigma) h^2$ pro nějaké $\sigma = (s_0, s_1)$, kde $0 < s_0, s_1 < h$. Tedy $\partial_y \partial_x f(\tau) = \partial_x \partial_y f(\sigma)$ a oba body σ a τ leží uvnitř čtverce K. □

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

Důsledek 11. *Pro každou funkci $f \in C^k(U)$ hodnoty jejích parciálních derivací až do řádu k včetně nezávisí na pořadí proměnných, podle nichž se derivuje, tj. pro všechna $l \leq k$ a $a \in U$ platí*

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

jakmile se posloupnosti $(i_1, \dots, i_l)$ a $(j_1, \dots, j_l)$ liší pouze permutací svých členů.

Důkaz. Když je posloupnost $v = (j_1, \dots, j_l)$ pouze permutací posloupnosti $u = (i_1, \dots, i_l)$, 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 10. □

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

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

Velmi užitečným 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. Například 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 12. *Nechť $U \subset \mathbf{R}^m$ je otevřená množina, $a \in U$ je bod a $f : U \to \mathbf{R}$ je funkce, která je na U n krát spojitě diferencovatelná, tj. $f \in C^n(U)$. Potom v okolí bodu a máme Taylorův rozvoj*

$$\begin{aligned} f(a+h) &= \sum_{i=0}^{n} \frac{1}{i!}(h_1\partial_1 + h_2\partial_2 + \cdots + h_m\partial_m)^i f(a) + o(\|h\|^n) \\ &= \sum \frac{1}{i_1!\dots i_m!} \cdot \frac{\partial^{i_1+\cdots+i_m} f}{\partial x_1^{i_1}\dots\partial x_m^{i_m}}(a) \cdot h_1^{i_1}\dots h_m^{i_m} + o(\|h\|^n), \end{aligned}$$

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

Důkaz. Na přednášce zazněla jen hlavní myšlenka: 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, fakticky přímka $l(t) = a + th$) pro $k \le n$ dostáváme

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

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

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

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

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

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

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

je multinomický koeficient. □

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

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

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

(x je řádkový vektor $(x_1, x_2, \dots, x_n)$). 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 10 mají funkce z $C^2(U)$ v každém bodě z U symetrickou Hessovu matici.

Odvodíme kritérium existence lokálních extrémů funkcí m proměnných, které zobecňuje výsledek pro funkce jedné proměnné. Roli hodnoty druhé derivace převezme 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 13. *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 Hessova matice $H_f(a)$ funkce f v bodě a je pozitivně (negativně) definitní, potom má f v a ostré lokální minimum (maximum).*
- *Pokud $\nabla f(a) = \overline{0}$ a Hessova matice $H_f(a)$ je indefinitní, nemá f v a ani neostrý lokální extrém.*

Důkaz. Bude příště. $\square$