Přednáška 7, 3. dubna 2013

Jako cvičení zdůvodněte tyto vlastnosti otevřených množin v $\mathbb{R}^m$: (i) množiny $\emptyset$ a $\mathbb{R}^m$ jsou otevřené, (ii) sjednocení $\bigcup_{i\in I} A_i$ libovolného systému $\{A_i \mid i \in I\}$ otevřených množin A_i je otevřená množina a (iii) průnik dvou (a tedy i konečně mnoha) otevřených množin je otevřená množina. Průnik nekonečně mnoha otevřených množin už ale nemusí být otevřená množina. *Okolí bodu* $a \in \mathbb{R}^m$ je libovolná otevřená množina v $\mathbb{R}^m$ obsahující a.

Budeme pracovat s funkcemi $f:\ M \to \mathbb{R}$, $f = f(x_1, x_2, \ldots, x_m)$, definovanými na podmnožinách $M \subset \mathbb{R}^m$ (které budou většinou otevřené), nebo i obecněji se zobrazeními

$$f:\ M \to \mathbb{R}^n,\ M \subset \mathbb{R}^m,\ f = (f_1, f_2, \ldots, f_n)\ ,$$

kde $f_i = f_i(x_1, x_2, \ldots, x_m)$ jsou *souřadnicové funkce.* Naším cílem bude jednak zobecnit derivaci jako lineární aproximaci z funkcí jedné proměnné na funkce několika proměnných, a pak stejně zobecnit kritéria existence extrému.

Než se pustíme do derivací, zobecníme spojitost: je-li $U \subset \mathbb{R}^m$ okolí bodu a, funkci $f:\ U \to \mathbb{R}$ nazveme *spojitou v bodě* a, když

$$\forall\varepsilon > 0\ \exists\delta > 0:\ \|x - a\| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon\ .$$

Stejně se definuje spojitost pro zobrazení $f:\ U \to \mathbb{R}^n$, pouze absolutní hodnotu $|f(x) - f(a)|$ (což je norma v $\mathbb{R}^1$) nahradíme normou $\|f(x) - f(a)\|$ v $\mathbb{R}^n$.

Směrová derivace, parciální derivace, diferenciál. Nechť $U \subset \mathbb{R}^m$ je okolí bodu a a $f:\ U \to \mathbb{R}$ je funkce. Její *směrovou derivací* v bodu a ve směru vektoru $v \in \mathbb{R}^m \backslash \{\overline{0}\}$ rozumíme limitu

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

existuje-li. Představme si U jako oblast v třírozměrném euklidovském prostoru, kde funkce f měří teplotu a kterou prolétává nějaká částice. Směrová derivace $\mathrm{D}_v f(a)$ udává okamžitou změnu teploty částice ve chvíli, kdy se nalézá v bodu a a má vektor rychlosti v.

Parciální derivace funkce f v bodě a podle proměnné x_i je směrová derivace $\mathrm{D}_{e_i} f(a)$, kde e_i je i-tý vektor kanonické báze, tj. $e_i =$

$(0, 0, \ldots, 0, 1, 0, 0, \ldots, 0)$ má na i-tém místě 1 a jinde nuly. Značíme ji $\frac{\partial f}{\partial x_i}(a)$. Explicitně,

$$\frac{\partial f}{\partial x_i}(a) = \lim_{h \to 0} \frac{f(a_1, \ldots, a_{i-1}, a_i + h, a_{i+1}, \ldots, a_m) - f(a_1, a_2, \ldots, a_m)}{h} \ .$$

Vektor hodnot všech parciálních derivací funkce f v bodě a je *gradient funkce f v a*,

$$\nabla f(a) := (\tfrac{\partial f}{\partial x_1}(a), \tfrac{\partial f}{\partial x_2}(a), \ldots, \tfrac{\partial f}{\partial x_m}(a)) \ .$$

Funkce f má v bodě a *(totální) diferenciál*, jinými slovy f je v a *diferencovatelná*, když existuje takové lineární zobrazení $L: \mathbb{R}^m \to \mathbb{R}$, že

$$\lim_{\|h\| \to 0} \frac{f(a+h) - f(a) - L(h)}{\|h\|} = 0 \ .$$

Toto lineární zobrazení L nazýváme *diferenciálem* a značíme $\mathrm{D}f(a)$, jeho hodnota $L(h)$ na vektoru h pak je $\mathrm{D}f(a)(h)$. Obecněji, pokud $f: U \to \mathbb{R}^n$ je zobrazení, řekneme, že je v a *diferencovatelné*, když existuje takové lineární zobrazení $L: \mathbb{R}^m \to \mathbb{R}^n$, že

$$\lim_{\|h\| \to 0} \frac{\|f(a+h) - f(a) - L(h)\|}{\|h\|} = 0$$

(norma ve jmenovateli se bere v $\mathbb{R}^m$ a norma v čitateli v $\mathbb{R}^n$). L opět nazýváme *diferenciálem* a značíme $\mathrm{D}f(a)$. Podstatný rozdíl ve srovnání se směrovou a parciální derivací je ten, že ty jsou pouhá čísla, kdežto diferenciál je složitější věc, lineární zobrazení.

Směrová derivace, parciální derivace a diferenciál funkce f v bodu a dávají lokální aproximace f poblíž a lineární funkcí:

$$\begin{aligned} f(a + tv) &= f(a) + \mathrm{D}_v f(a) \cdot t + o(t), \ t \to 0 \ , \\ f(a + te_i) &= f(a) + \frac{\partial f}{\partial x_i}(a) \cdot t + o(t), \ t \to 0 \ , \\ f(a + h) &= f(a) + \mathrm{D}f(a)(h) + o(\|h\|), \ \|h\| \to 0 \ . \end{aligned}$$

V prvních dvou vztazích je t reálné číslo jdoucí k nule a aproximace platí pouze pro argumenty funkce na přímce jdoucí bodem a ve směru v, resp. ve směru i-té souřadnicové osy. Ve třetím vztahu h probíhá body $\mathbb{R}^m$ a aproximace platí pro všechny argumenty funkce v okolí bodu a. Diferencovatelnost je silnější vlastnost f než existence směrových nebo parciálních derivací, z nichž neplyne ani spojitost funkce v daném bodě. Například funkce

$f = f(x, y) : \mathbb{R}^2 \to \mathbb{R}$ definovaná jako $f(x, y) = 1$ pokud $xy = 0$ a $f(x, y) = 0$ jinde (tj. f je 1 na souřadnicových osách a jinak 0) má v počátku obě parciální derivace a jejich hodnota je 0, ale není v počátku spojitá. Jako cvičení vymyslete funkci z $\mathbb{R}^2$ do $\mathbb{R}$, která má v počátku všechny směrové derivace, ale přesto tam není spojitá.

Počítat parciální derivace už umíme, při výpočtu $\frac{\partial f}{\partial x_i}$ se proměnné různé od x_i berou jako konstanty a f tak derivujeme jako funkci jediné proměnné x_i. Například pro funkci $f = f(x, y, z) = x^3 y \sin(yz) + x \log z$ máme

$$\frac{\partial f}{\partial y} = x^3(\sin(yz) + zy\cos(yz)) \ .$$

Jednoduchý důkaz následujícího tvrzení ponecháváme jako cvičení.

Tvrzení (vlastnosti diferenciálu). *Nechť $f = (f_1, f_2, \ldots, f_n) : U \to \mathbb{R}^n$ je zobrazení, kde $U \subset \mathbb{R}^m$ je okolí bodu a.*

1. *Diferenciál zobrazení f v a je určený jednoznačně (když existuje).*
2. *Zobrazení f je diferencovatelné v a, právě když je jeho každá souřadnicová funkce f_i diferencovatelná v a.*
3. *Když je zobrazení f diferencovatelné v bodu a, potom je v bodu a spojité.*

Tvrzení (diferenciál $\Rightarrow \partial$). *Když je funkce $f : U \to \mathbb{R}$, $U \subset \mathbb{R}^m$ je okolí bodu a, diferencovatelná v a, pak má v a všechny parciální derivace a jejich hodnoty diferenciál určují:*

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

(tj. hodnota diferenciálu v h se dostane jako obvyklý skalární součin h a gradientu f v a). Funkce f pak má také 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

$$L(h) = L(h_1 e_1 + h_2 e_2 + \cdots + h_m e_m) = L(e_1)h_1 + \cdots + L(e_m)h_m \ ,$$

kde e_i je i-tý vektor kanonické báze. Ovšem ($f(a + te_i) = f(a) + L(te_i) + o(\|te_i\|)$ pro $t \to 0$ vzhledem k diferencovatelnosti f v a)

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

a tak $L(e_i) = \frac{\partial f}{\partial x_i}(a)$.

Nechť $v \in \mathbb{R}^m$ je nenulový vektor. Protože $v = v_1 e_1 + \cdots + v_m e_m$, kde e_i jsou vektory kanonické báze $\mathbb{R}^m$, a f má v a všechny parciální derivace, z definice směrové derivace plyne, že

$$\mathrm{D}_v f(a) = v_1 \cdot (\partial f/\partial x_1)(a) + v_2 \cdot (\partial f/\partial x_2)(a) + \cdots + v_m \cdot (\partial f/\partial x_m)(a)\ ,$$

což je právě hodnota diferenciálu $\mathrm{D}f(a)$ ve v. □

Obecné zobrazení $f : U \to \mathbb{R}^n$ má diferenciál $L = \mathrm{D}f(a) : \mathbb{R}^m \to \mathbb{R}^n$ popsaný maticí typu $n \times m$ a L se na vektor h aplikuje maticovým násobením:

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

Podle předešlých tvrzení má tato matice v i-tém řádku gradient souřadnicové funkce f_i v bodě a:

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

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

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

Je-li Jacobiho matice čtvercová, její determinant se nazývá *jacobián.*

Věta ($\partial \Rightarrow$ diferenciál). *Nechť $U \subset \mathbb{R}^m$ je okolí bodu $a \in \mathbb{R}^m$. Má-li funkce $f : U \to \mathbb{R}$ na U všechny parciální derivace a ty jsou v bodě a spojité, pak je f v bodě a diferencovatelná.*

Důkaz si řekneme příště.