6. přednáška 8. listopadu 2005

Vektorový prostor se skalárním součinem (VPSS) je vektorový prostor X nad tělesem $\mathbf{R}$, který je vybaven zobrazením $\langle \cdot, \cdot \rangle : X \times X \to \mathbf{R}$, zvaným *skalární součin*, splňujícím tři axiomy (pro všechny $x, x', y \in X$ a $\kappa, \lambda \in \mathbf{R}$):

a) $\langle x, x \rangle \geq 0$, $\langle x, x \rangle = 0 \iff x = 0$ (positivní definitnost),
b) $\langle \kappa x + \lambda x', y \rangle = \kappa \langle x, y \rangle + \lambda \langle x', y \rangle$ (bilinearita) a
c) $\langle x, y \rangle = \langle y, x \rangle$ (symetrie).

Na rozdíl od metriky a normy může skalární součin nabývat záporných hodnot. Příkladem VPSS je euklidovský prostor $\mathbf{R}^m$ se skalárním součinem

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_m y_m$$

nebo prostor $C[a, b]$ se skalárním součinem

$$\langle f, g \rangle = \int_a^b f(x) g(x)\, dx.$$

Věta 1 (Cauchyova-Schwarzova nerovnost). *Ve VPSS X pro každé dva vektory $x, y \in X$ platí nerovnost*

$$\langle x, y \rangle^2 \leq \langle x, x \rangle \langle y, y \rangle.$$

Rovnost nastává, právě když je jeden z vektorů skalárním násobkem druhého: $x = \lambda y$ pro $\lambda \in \mathbf{R}$.

Důkaz. Byl v Lineární algebře, proto ho zde neuvádíme. $\square$

Na každém VPSS máme i normu: uvažme zobrazení

$$\|x\| := \sqrt{\langle x, x \rangle}.$$

První dva axiomy normy jsou zřejmě splněny. Trojúhelníková nerovnost plyne z věty 1:

$$\begin{aligned}
\langle x, y \rangle^2 \leq \langle x, x \rangle \langle y, y \rangle \quad &\Longrightarrow \quad \langle x, y \rangle \leq \sqrt{\langle x, x \rangle \langle y, y \rangle} \\
&\iff \quad \langle x, x \rangle + \langle y, y \rangle + 2\langle x, y \rangle \leq \left(\sqrt{\langle x, x \rangle} + \sqrt{\langle y, y \rangle} \right)^2 \\
&\iff \quad \langle x + y, x + y \rangle \leq \left(\sqrt{\langle x, x \rangle} + \sqrt{\langle y, y \rangle} \right)^2 \\
&\iff \quad \|x + y\| \leq \|x\| + \|y\|.
\end{aligned}$$

Hilbertův prostor[1] je úplný VPSS, tj. odvozená metrika

$$d(x, y) := \sqrt{\langle x - y, x - y \rangle}$$

je úplná.

Každý vektorový prostor se skalárním součinem $\langle \cdot, \cdot \rangle$ je normovaným vektorovým prostorem ($\|x\| = \sqrt{\langle x, x \rangle}$), tedy i metrickým prostorem ($d(x, y) = \|x - y\|$) a tedy i topologickým prostorem.

2.2. Lokální lineární aproximace funkcí více proměnných. Budeme pracovat v euklidovském VPSS $\mathbf{R}^m$ s normou

$$|x| = \|x\| = \|x\|_2 = \sqrt{\sum_{i=1}^{m} x_i^2}.$$

Nechť $D \subset \mathbf{R}^m$ je otevřená množina, $f : D \to \mathbf{R}$ je funkce a $a \in D$ je bod. *Derivace funkce f v bodě a ve směru* $v \in \mathbf{R}^m$, $v \neq 0$, nebo také *směrová derivace*, je limita

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

(pokud existuje). Představte si D jako oblast v třírozměrném euklidovském prostoru, jejíž teplotu měří funkce f a kterou prolétá bodem a po přímočaré dráze, s vektorem rychlosti v, nějaká částice. Směrová derivace $\mathrm{D}_v f(a)$ pak udává okamžitou změnu teploty částice v okamžiku, kdy se nachází v bodu a.

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

$$e_i = (0, 0, \ldots, 0, 1, 0, 0, \ldots, 0)$$

s 1 na i-tém místě; značíme ji $\frac{\partial f}{\partial x_i}(a)$. Explicitně,

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

[1]Nazvaný podle německého matematika Davida Hilberta (1862–1943).

Když má f parciální derivaci podle x_i v každém bodě D, dostáváme funkci $\frac{\partial f}{\partial x_i} : D \to \mathbf{R}$. Vektor hodnot všech parciálních derivací funkce f v bodě a je *gradient* funkce f v a,

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

Počítat parciální derivace umíme, při výpočtu $\frac{\partial f}{\partial x_i}$ jsou všechny proměnné kromě x_i konstanty a f derivujeme jako funkci jediné proměnné x_i. Například

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

Řekneme, že funkce f má v bodě a *(totální) diferenciál* nebo že f je v a *diferencovatelná*, pokud existuje lineární zobrazení $L : \mathbf{R}^m \to \mathbf{R}$ takové, 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 $\mathrm{D}f(a)(h)$.

Přepis definic směrové derivace, parciální derivace a diferenciálu dává lokální aproximace přírůstku funkční hodnoty lineárními výrazy v přírůstku argumentu:

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

V prvních dvou rovnicích je t reálné číslo, $t \to 0$, a aproximace platí pouze pro argumenty funkce f ležící na přímce jdoucí bodem a ve směru v, resp. ve směru i-té souřadnicové osy. Ve třetí rovnici h probíhá $\mathbf{R}^m$, $\|h\| \to 0$, a aproximace platí pro všechny argumenty funkce f (rozumnou aproximaci ovšem dostaneme jen v nějakém malém okolí bodu a). Diferencovatelnost je daleko silnější požadavek na f než požadavek existence směrových nebo parciálních derivací.

Příklady. 1. Funkce $f = f(x,y) : \mathbf{R}^2 \to \mathbf{R}$ definovaná jako 1 na množině $\{(x,y) \in \mathbf{R}^2 : y = x^2, x \neq 0\}$ a jako 0 pro všechny zbylé body roviny má v počátku všechny směrové derivace (jsou rovné nule), ale není tam spojitá.

2. Podobně, definujeme-li f jako 1 na souřadnicových osách, tj. na množině $\{(x,y) \in \mathbf{R}^2 : \ xy = 0\}$, a jako 0 pro všechny zbylé body roviny, má f v počátku obě parciální derivace (jsou rovné nule), ale kromě nich už žádnou další směrovou derivaci.

3. Vraťme se k příkladu s částicí z definice směrové derivace. Předpokládejme, že funkce $f : \ D \to \mathbf{R}$ ($D \subset \mathbf{R}^m$ je otevřená množina a třeba $m = 3$) je diferencovatelná v $a \in D$ a představujme si ji zase jako teplotu bodů oblasti prostoru D. Bodem a prolétá částice po křivočaré dráze $\gamma : \ [0,1] \to D$, kde $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_m)$, $\gamma_i : \ D \to \mathbf{R}$, $\gamma(u) = a$ (částice je v a v čase $u \in (0,1)$) a všechny funkce γ_i mají v u vlastní derivaci. V okamžiku, kdy se částice nachází v bodě a, má vektor rychlosti

$$v = (\gamma_1'(u), \gamma_2'(u), \ldots, \gamma_m'(u)).$$

Jaká je v této chvíli její okamžitá změna teploty? Pro $t \to 0$ a $1 \le i \le m$ máme

$$\gamma_i(u+t) = \gamma_i(u) + \gamma_i'(u)t + o(t),$$

což dává $\gamma(u+t) = \gamma(u) + tv + \alpha(t) = a + tv + \alpha(t)$, kde $\|\alpha(t)\| = o(t)$. Pro $h \in \mathbf{R}^m$, $\|h\| \to 0$, máme $f(a+h) = f(a) + \mathrm{D}f(a)(h) + \beta(h)$, kde $\beta(h) = o(\|h\|)$. Takže, pro $t \to 0$,

$$\begin{aligned}
\frac{f(\gamma(u+t)) - f(\gamma(u))}{t} &= \frac{f(a) + \mathrm{D}f(a)(tv + \alpha(t)) + \beta(tv + \alpha(t)) - f(a)}{t} \\
&= \frac{t\mathrm{D}f(a)(v) + \mathrm{D}f(a)(\alpha(t)) + \beta(tv + \alpha(t))}{t} \\
&= \mathrm{D}f(a)(v) + \mathrm{D}f(a)(\tfrac{1}{t}\alpha(t)) + \frac{\beta(tv + \alpha(t))}{t} \\
&= \mathrm{D}f(a)(v) + o(1).
\end{aligned}$$

Okamžitá změna teploty se tedy rovná $\mathrm{D}f(a)(v)$ a je stejná pro všechny dráhy částice, na nichž má v bodě a vektor rychlosti v.

Pojem diferenciálu rozšíříme na obecnější situaci, kdy $f : \ D \to \mathbf{R}^n$ ($D \subset \mathbf{R}^m$ je otevřená množina) je zobrazení dané n-ticí reálných funkcí: $f = (f_1, f_2, \ldots, f_n)$ a $f_i : \ D \to \mathbf{R}$. Řekneme, že zobrazení f má v bodě a *diferenciál* nebo že tam je *diferencovatelné*, existuje-li lineární zobrazení $L : \ \mathbf{R}^m \to \mathbf{R}^n$ takové, že

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

Lineární zobrazení L značíme $\mathrm{D}f(a)$. Z aproximačního pohledu to opět znamená, že $f(a+h) = f(a)+\mathrm{D}f(a)(h)+\alpha(h)$, kde pro $\|h\| \to 0$ máme $\|\alpha(h)\| = o(\|h\|)$. Diferenciál je určen jednoznačně: kdyby K a L byly dva různé diferenciály zobrazení f v bodě a, pak $K(v) - L(v) = c$ je nenulový vektor pro nějaký nenulový vektor $v \in \mathbf{R}^m$ a tedy $\|K(tv) - L(tv)\| = \|tc\| = |t| \cdot \|c\|$ pro každé $t > 0$, což je ve sporu s $\|K(h) - L(h)\| = o(\|h\|)$ pro $\|h\| \to 0$.

Tvrzení 2. *Nechť $D \subset \mathbf{R}^m$ je otevřená množina, $a \in D$ je bod a $f : D \to \mathbf{R}^n$ je zobrazení.*
1. *Zobrazení f je diferencovatelné v a, právě když každá z n souřadnicových funkcí f_i je diferencovatelná v a. Navíc*

$$\mathrm{D}f(a) = (\mathrm{D}f_1(a), \mathrm{D}f_2(a), \ldots, \mathrm{D}f_n(a)).$$

2. *Je-li zobrazení f diferencovatelné v a, je v a spojité.*
3. *Nechť $n = 1$, tj. f je reálná funkce o m proměnných, a f je diferencovatelná v a. Pak f má v a všechny parciální derivace a jejich hodnoty určují diferenciál:*

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

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

Důkaz. 1. Cvičení na definice. Pro odhad chyby v lineárních aproximacích diferenciály si uvědomte, že pro $v \in \mathbf{R}^n$ platí $|v_i| \le \|v\| \le |v_1|+|v_2|+\cdots+|v_n|$.
2. Plyne hned z definice diferenciálu.
3. Z linearity diferenciálu $L = \mathrm{D}f(a)$ máme

$$L(h) = L(h_1e_1 + h_2e_2 + \cdots + h_me_m) = \alpha_1h_1 + \cdots + \alpha_mh_m,$$

kde $\alpha_i = L(e_i)$ (e_i je i-tý vektor kanonické báze). Ovšem, pro $t \to 0$,

$$\frac{f(a + te_i) - f(a)}{t} = \frac{L(te_i) + o(\|te_i\|)}{t} = \alpha_i + o(1),$$

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

V obecném případu zobrazení $f:\ D \to \mathbf{R}^n$ je diferenciál $L = \mathrm{D}f(a):\ \mathbf{R}^m \to \mathbf{R}^n$ reprezentován $n \times m$ maticí:

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

Podle bodů 1 a 3 posledního tvrzení 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 3. *Diferenciál zobrazení $f:\ D \to \mathbf{R}^n$ v bodě $a \in D$, kde $D \subset \mathbf{R}^m$ je otevřená množina a f má souřadnicové funkce $f = (f_1, f_2, \dots, f_n)$, je dán tzv.* Jacobiho maticí[2] *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.*

[2]Nazvaná podle německého matematika Carla Jacobiho (1804–1851).