4. Derivace vyšších řádů

4.1. Bilineární a kvadratická zobrazení Připomeňme si, že zobrazení $l: X_1 \times X_2 \to Y$, kde X_1, X_2 a Y jsou vektorové prostory, se nazývá *bilineární*, jsou-li pro každé $x_1 \in X_1$, $x_2 \in X_2$ zobrazení

$$\begin{aligned} X_1 &\to Y, x \to l(x, x_2), \\ X_2 &\to Y, x \to l(x_1, x) \end{aligned} \tag{4.1.1}$$

lineární. Vektorový prostor všech bilineárních zobrazení $l: X_1 \times X_2 \to Y$ označujeme symbolem $L(X_1, X_2; Y)$.

Libovolnému lineárnímu zobrazení $l \in L(X_1, L(X_2, Y))$ lze předpisem

$$\bar{l}(x_1, x_2) = l(x_1)(x_2) \tag{4.1.2}$$

přiřadit bilineární zobrazení $\bar{l} \in L(X_1, X_2; Y)$. Přiřazení $l \to \bar{l}$ je evidentně bijektivní.

Dokažte to.
Přesvědčte se o tom, že zobrazení $\bar{l}$ je skutečně bilineární!
Pro libovolné číslo x_1 definujme lineární zobrazení $l(x_1): \boldsymbol{R} \to \boldsymbol{R}$ předpisem

$$l(x_1)(x) = 2x_1 x. \tag{4.1.3}$$

Zobrazení l je zjevně lineární. Příslušné zobrazení $\bar{l}: \boldsymbol{R} \times \boldsymbol{R} \to \boldsymbol{R}$ je dáno předpisem

$$l(x_1, x_2) = 2x_1 x_2. \tag{4.1.4}$$

Bilineární zobrazení $l \in L(X, X; Y)$ se nazývá *symetrické*, jestliže pro libovolné dva vektory $x_1, x_2 \in X$ platí

$$l(x_1, x_2) = l(x_2, x_1). \tag{4.1.5}$$

Bilineární zobrazení l z (4.1.3) je symetrické.
Každé bilineární zobrazení z $\boldsymbol{R} \times \boldsymbol{R}$ do $\boldsymbol{R}$ je symetrické (lze dokázat pomocí vyjádření vektorů z $\boldsymbol{R}$ v bázi).

Symetrická a *antisymetrická* část bilineárního zobrazení $l \in L(X, X; Y)$ jsou bilineární zobrazení $\operatorname{sym} l$ a $\operatorname{alt} l$, definovaná předpisy

$$\begin{aligned} \operatorname{sym} l(x_1, x_2) &= \tfrac{1}{2}\big(l(x_1, x_2) + l(x_2, x_1)\big), \\ \operatorname{alt} l(x_1, x_2) &= \tfrac{1}{2}\big(l(x_1, x_2) - l(x_2, x_1)\big). \end{aligned} \tag{4.1.6}$$

Pro každé bilineární zobrazení l platí

$$\operatorname{sym} l + \operatorname{alt} l = l. \tag{4.1.7}$$

Zobrazení l je symetrické, právě když $\operatorname{sym} l = l$, což je ekvivalentní podmínce $\operatorname{alt} l = 0$.

Zobrazení $l: X \to Y$ se nazývá *kvadratické*, existuje-li bilineární zobrazení $p: X \times X \to Y$ takové, že pro každé $x \in X$ platí

$$l(x) = p(x, x). \tag{4.1.8}$$

Vektorový prostor všech kvadratických zobrazení $l{:}\,X \to Y$ označujeme $L_2(X,Y)$.

Zobrazení p, která splňují vztah (4.1.8) je více; pokud platí $\operatorname{sym} p_1 = \operatorname{sym} p_2$, pak samozřejmě pro každé x $p_1(x,x) = p_2(x,x)$. Důležité je, že platí také opačné tvrzení:

Lemma 4.1. *Jestliže pro bilineární zobrazení* $p_1, p_2{:}\,X \times X \to Y$ *a každé* $x \in X$ *platí* $p_1(x,x) = p_2(x,x)$, *pak* $\operatorname{sym} p_1 = \operatorname{sym} p_2$.

D ů k a z . Pro libovolné $x_1, x_2 \in X$ máme

$$\begin{aligned}\operatorname{sym} p_1(x_1,x_2) &= \tfrac{1}{2}\big(p_1(x_1,x_2) + p_1(x_2,x_1)\big)\\ &= \tfrac{1}{2}\big(p_1(x_1,x_2) + p_1(x_1,x_1) + p_1(x_2,x_1) + p_1(x_2,x_2) - p_1(x_1,x_1) - p_1(x_2,x_2)\big)\\ &= \tfrac{1}{2}\big(p_1(x_1+x_2,x_1+x_2) - p_1(x_1,x_1) - p_1(x_2,x_2)\big)\\ &= \tfrac{1}{2}\big(p_2(x_1+x_2,x_1+x_2) - p_2(x_1,x_1) - p_2(x_2,x_2)\big)\\ &= \operatorname{sym} p_2(x_1,x_2).\end{aligned}$$

Z uvedeného vyplývá

Věta 4.2. *K libovolnému kvadratickému zobrazení* $l \in L_2(X,Y)$ *existuje právě jedno symetrické bilineární zobrazení* $p{:}\,X \times X \to Y$, *splňující (4.1.8).*

Zobrazení l a p z uvedené věty se nazývají *asociovaná*.

Symetrické bilineární zobrazení, asociované s daným kvadratickým zabrazením l lze vypočítat podle vztahu

$$p(x_1,x_2) = \tfrac{1}{2}\big(l(x_1+x_2) - l(x_1) - l(x_2)\big). \tag{4.1.9}$$

Na prostoru $L(X, L(X,Y))$ je dána norma pomocí předpisu (1.7.1). Pro zobrazení p a l z (4.1.8) můžeme položit

$$\|l\| = \|p\|. \tag{4.1.10}$$

Dostaneme tak normu na vektorovém prostoru $L_2(X,Y)$.

Uvedená definice patří k oněm poněkud abstraktním a napoprvé těžko pochopitelným. Pro začátek jistě postačí, když si z ní odneseme poznatek, že jsme na prostoru $L_2(X,Y)$ definovali normu. Beztak jsou podle věty 1.7 všechny ekvivalentní a my se budeme zajímat pouze o topologii touto normou indukovanou. To také znamená, že jsme normu na $L_2(X,Y)$ mohli definovat jinak — indukovaná topologie vyjde stejně.

4.2. Derivace a diferenciál druhého řádu. Než přistoupíme k definici derivace druhého řádu, uvědomme si, že v definici Fréchetovy derivace v kapitole 2. jsme ze všech vlastností prostoru $\boldsymbol{R}^n$ (a $\boldsymbol{R}^m$) využili pouze jeho strukturu normovaného prostoru. Téměř celý odstavec 2.1 tedy zůstane v platnosti, pokud v něm všude místo o $\boldsymbol{R}^n$ a $\boldsymbol{R}^m$ budeme hovořit o libovolných normovaných prostorech. Výjimku tvoří jedině věty 2.8 a 2.10, které pojednávají o složkách zobrazení.

Mějme zobrazení $f{:}\,U \subset \boldsymbol{R}^n \to \boldsymbol{R}^m$, diferencovatelné v každém bodě nějakého okolí V bodu $x \in U$. Zobrazení Df je tedy definováno v každém bodě množiny V a jeho oborem hodnot je normovaný prostor $L(\boldsymbol{R}^n, \boldsymbol{R}^m)$. Fréchetova derivace $D(Df)(x)$ tohoto zobrazení v bodě x (pokud existuje) je tedy lineární zobrazení z $\boldsymbol{R}^n$ do prostoru $L(\boldsymbol{R}^n, \boldsymbol{R}^m)$. Příslušné bilineární zobrazení z $\boldsymbol{R}^n \times \boldsymbol{R}^n$ do $\boldsymbol{R}^m$ označujeme symbolem $D^2 f(x)$ a nazýváme *druhou* (*Fréchetovou*) *derivací zobrazení* f *v bodě* x.

Zobrazení, které má v bodě x druhou derivaci, se nazývá *dvakrát diferencovatelné v bodě* x.

Pro derivace podle vektoru a parciální derivace zavádíme tato označení: $D_{h_2}(D_{h_1} f)(x) = D_{h_1 h_2} f(x)$, $D_{i_2}(D_{i_1} f)(x) = D_{i_1 i_2} f(x)$.

Lemma 4.3. *Nechť zobrazení* $f{:}\,U \subset \boldsymbol{R}^n \to \boldsymbol{R}^m$ *má v bodě* $x \in U$ *druhou derivaci. Pak pro libovolný vektor* $h \in \boldsymbol{R}^n$ *je derivace zobrazení* $D_h f$ *v bodě* x *rovna lineárnímu zobrazení* $\bar{h} \to D^2 f(x)(\bar{h}, h)$.

D ů k a z . Uvažme zobrazení $\mathrm{ev}_h\colon L(\boldsymbol{R}^n,\boldsymbol{R}^m)\to\boldsymbol{R}^m$, $\mathrm{ev}_h(l)=l(h)$. Platí

$$D_h f(x)=Df(x)(h)=\mathrm{ev}_h(Df(x)),$$

neboli

$$D_h f=\mathrm{ev}_h\circ Df. \tag{4.2.1}$$

Zobrazení ev_h je lineární, je tedy diferencovatelné v libovolném bodě $l\in L(\boldsymbol{R}^n,\boldsymbol{R}^m)$ a platí $D\mathrm{ev}_h(l)=\mathrm{ev}_h$. Zobrazení $D_h f$ je tedy (jako kompozice diferencovatelných zobrazení) diferencovatelné v bodě x. Dostáváme

$$\begin{aligned}D(D_h f)(x)(\bar h)&=D(\mathrm{ev}_h\circ Df)(x)(\bar h)=\big(\mathrm{ev}_h\circ D(Df)(x)\big)(\bar h)\\&=\mathrm{ev}_h\big(D(Df)(x)(\bar h)\big)=\big(D(Df)(x)(\bar h)\big)(h)=D^2f(x)(\bar h,h).\end{aligned} \tag{4.2.2}$$

Podívejme se podrobněji na zobrazení ev_h, použité v minulém důkazu. Tak například pro $n=2$ a $m=1$ je každé lineární zobrazení $l\in L(\boldsymbol{R}^2,\boldsymbol{R})$ dáno maticí typu 1×2 (l_1,l_2) tak, že pro každé $h\in\boldsymbol{R}^2$ platí

$$l(h)=l_1h^1+l_2h^2. \tag{4.2.3}$$

Pokud tedy například $h=(2,3)$, pak zobrazení ev_h je určeno předpisem

$$\mathrm{ev}_h(l)=2l_1+3l_2. \tag{4.2.4}$$

Nyní tedy snad již není divu, že se jedná o lineární (a tedy spojité a diferencovatelné) zobrazení.

Jiný důkaz lemmatu 4.3 pomocí definice derivace lze provést takto:

$$\begin{aligned}\lim_{\bar h\to 0}\frac{D_hf(x+\bar h)-D_hf(x)-D^2f(x)(\bar h,h)}{\|\bar h\|}&=\lim_{\bar h\to 0}\frac{Df(x+\bar h)(h)-D_hf(x)(h)-D(Df)(x)(\bar h)(h)}{\|\bar h\|}\\&=\lim_{\bar h\to 0}\frac{\big(Df(x+\bar h)-D_hf(x)-D(Df)(x)(\bar h)\big)(h)}{\|\bar h\|}\\&=\lim_{\bar h\to 0}\frac{\mathrm{ev}_h\big(Df(x+\bar h)-D_hf(x)-D(Df)(x)(\bar h)\big)}{\|\bar h\|}\\&=\mathrm{ev}_h\left(\lim_{\bar h\to 0}\frac{Df(x+\bar h)-D_hf(x)-D(Df)(x)(\bar h)}{\|\bar h\|}\right)=\mathrm{ev}_h(0)=0.\end{aligned}$$

Lemma 4.4. *Nechť zobrazení $f\colon U\subset\boldsymbol{R}^n\to\boldsymbol{R}^m$ má v bodě $x\in U$ druhou derivaci. Pak pro libovolné dva vektory $h_1,h_2\in\boldsymbol{R}^n$ platí*

$$D_{h_1h_2}f(x)=D^2f(x)(h_2,h_1) \tag{4.2.5}$$

D ů k a z . Plyne přímo z lemmatu 4.3.

Věnujme se nyní chvíli pojmu spojité diferencovatelnosti druhého řádu. Budeme postupovat podobně, jako v kapitole 2. Uvažme zobrazení $f\colon U\subset\boldsymbol{R}^n\to\boldsymbol{R}^m$ a předpokládejme, že v okolí V bodu $x\in U$ existují derivace Df a $D_G(Df)$. Druhá ze zmíněných derivací je zobrazením z V do prostoru $L(\boldsymbol{R}^n,L(\boldsymbol{R}^n,\boldsymbol{R}^m))$, na kterém je podle (1.7.1) definovaná norma. Můžeme tedy uvažovat spojitost zobrazení $D_G(Df)$ v bodě x: Je-li zobrazení $D_G(Df)$ v bodě x spojité, pak se o zobrazení f říká, že je v bodě x *dvakrát spojitě diferencovatelné*. O spojité diferencovatelnosti druhého řádu platí následující věty, které jsou (včetně důkazů!) velmi podobné větám 2.23 a 2.26.

Věta 4.5. *Zobrazení dvakrát spojitě diferencovatelné v bodě x je v tomto bodě dvakrát diferencovatelné.*

Věta 4.6. *Nechť pro zobrazení $f:U\subset\boldsymbol{R}^n\to\boldsymbol{R}^m$ existují všechny funkce $D_jf^i:U\to\boldsymbol{R}$*

a $D_{jk}f^i : U \to \boldsymbol{R}$, *kde* $i \in \{1,\ldots,m\}$, $j,k \in \{1,\ldots,n\}$, *a jsou na množině* U *spojité*. *Pak zobrazení* f *je dvakrát spojitě diferencovatelné v každém bodě množiny* U.

Jestliže zobrazení je dvakrát spojitě diferencovatelné v každém bodě množiny U, říkáme také, že je *třídy* C^2.

Posledním úkolem tohoto odstavce je ukázat Schwartzovu větu o symetrii druhé derivace. Důkaz této věty je v obecném případě poměrně složitý, vybereme si proto pouze speciální (ale základní a nejobvyklejší) případ pro dvakrát spojitě diferencovatelná zobrazení.

Lemma 4.7. *Mějme zobrazení* $f : U \subset \boldsymbol{R}^n \to \boldsymbol{R}^m$ *a vektory* $h_1, h_2 \in \boldsymbol{R}^n$ *takové, že zobrazení* $D^2_{h_1h_2}f$ *je definováno na množině* U *a spojité v bodě* $x \in U$. *Pak platí*

$$D_{h_1h_2}f(x) = \lim_{s\to 0}\frac{f(x+sh_1+sh_2)-f(x+sh_1)-f(x+sh_2)+f(x)}{s^2}. \tag{4.2.6}$$

D ů k a z . Bez újmy na obecnosti předpokládejme, že $n=1$ a pro pevné $s \in \boldsymbol{R}$ definujme zobrazení g předpisem

$$g(x) = f(x+sh_2) - f(x) \tag{4.2.7}$$

Podle věty o střední hodnotě pro funkce nyní platí

$$\begin{aligned} f(x+sh_1+sh_2)-f(x+sh_1)-f(x+sh_2)+f(x) &= g(x+sh_1)-g(x) = \\ = D_{sh_1}g(y_1) = D_{sh_1}f(y_1+sh_2)-D_{sh_1}f(y_1) &= D_{sh_2}D_{sh_1}f(y) = s^2D_{h_1h_2}f(y), \end{aligned}$$

kde bod y_1 leží na úsečce $[x, x+sh_1]$ a bod y na úsečce $[y_1, y_1+sh_2]$. Určitě tedy

$$\|y-x\| \le \|sh_1+sh_2\| = |s|\|h_1+h_2\|, \tag{4.2.8}$$

což znamená, že $\lim_{s\to 0} y = x$ a $\lim_{s\to 0} D_{h_1h_2}f(y) = D_{h_1h_2}f(x)$.[1)] Tím je lemma dokázáno.

Věta 4.8 (Schwartzova o symetrii druhé derivace). *Nechť zobrazení* $f : U \subset \boldsymbol{R}^n \to \boldsymbol{R}^m$ *je dvakrát diferencovatelné v bodě* $x \in U$. *Pak bilineární zobrazení* $D^2f(x)$ *je symetrické.*

D ů k a z s p e c i á l n í h o p ř í p a d u . Je-li zobrazení f v bodě x dvakrát spojitě diferencovatelné, pak můžeme použít předchozí lemma, z něhož plyne

$$D_{h_1h_2}f(x) = D_{h_2h_1}f(x). \tag{4.2.9}$$

Tvrzení tedy plyne z (4.2.5).

4.3. Multilineární a homogenní zobrazení.

4.4. Derivace vyššího řádu.

4.5. Diferenciál a Taylorův polynom.

Lemma 4.9. *Nechť* $l \in L_r(\boldsymbol{R}^n, \boldsymbol{R}^m)$. *Pak*
a) pro libovolné $p < r$ *platí* $d^p l(0) = 0$,
b) $d^r l(0) = r! \cdot l$,
c) pro libovolné $p > r$, $x \in \boldsymbol{R}^n$ *platí* $d^p l(x) = 0$.

Lemma 4.10. *Nechť pro zobrazení* $f : U \subset \boldsymbol{R}^n \to \boldsymbol{R}^m$ *a bod* $x \in U$ *platí*

[1)] V těchto limitách jsme poněkud nepřesní; necháme na čtenáři, aby si v nich udělal jasno sám (je třeba místo y psát jistou funkci proměnné s).

$$\begin{aligned} Df(x) &= 0 \\ D^2 f(x) &= 0 \\ &\vdots \\ D^k f(x) &= 0. \end{aligned} \tag{4.5.1}$$

Pak

$$\lim_{h\to 0} \frac{\|f(x+h)-f(x)\|}{\|h\|^k} = 0. \tag{4.5.2}$$

D ů k a z . Pro $k=1$ tvrzení plyne přímo z definice derivace. Předpokládejme nyní platnost tvrzení pro $k=r-1$ a položme $g=Df$. Máme $g(x)=0$ a

$$\lim_{h\to 0} \frac{\|g(x+h)-g(x)\|}{\|h\|^{r-1}} = \lim_{h\to 0} \frac{\|g(x+h)\|}{\|h\|^{r-1}} = 0, \tag{4.5.3}$$

což znamená, že k libovolnému $\varepsilon > 0$ existuje otevřená koule $B \subset \boldsymbol{R}^n$ se středem v nule taková, že pro každé $h \in B$ platí

$$\|g(x+h)\| = \|g(x+h)-g(x)\| < \varepsilon \|h\|^{r-1}. \tag{4.5.4}$$

Pro libovolný vektor $h \in B$ nyní podle důsledku věty o střední hodnotě platí

$$\|f(x+h)-f(x)\| \le \sup_{y\in[x,x+h]} \|Df(y)\| \cdot \|h\| = \sup_{y\in[x,x+h]} \|g(y)\| \cdot \|h\| < \varepsilon \|h\|^{r-1} \|h\| = \varepsilon \|h\|^r.$$

Tím je tvrzení dokázáno i pro $k=r$.

Věta 4.11 (o diferenciálu vyššího řádu). *Nechť zobrazení $f : U \subset \boldsymbol{R}^n \to \boldsymbol{R}^m$ má derivaci řádu k v bodě $x \in U$. Pak*

$$f(x+h)-f(x) = \mathrm{d}f(x)(h) + \frac{1}{2!}\mathrm{d}^2 f(x)(h) + \ldots + \frac{1}{k!}\mathrm{d}^k f(x)(h) + \varepsilon(h)\|h\|^k, \tag{4.5.5}$$

kde

$$\lim_{h\to 0} \varepsilon(h) = 0. \tag{4.5.6}$$

D ů k a z . Položme

$$g(h) = f(x+h) - f(x) - \mathrm{d}f(x)(h) - \frac{1}{2!}\mathrm{d}^2 f(x)(h) - \ldots - \frac{1}{k!}\mathrm{d}^k f(x)(h). \tag{4.5.7}$$

Podle lemmatu 4.9 máme pro $r \in \{1,\ldots,k\}$ $\mathrm{d}^r g(0) = 0$ a podle lemmatu 4.10 pro

$$\varepsilon(h) = \frac{g(h)}{\|h\|^k} \tag{4.5.8}$$

platí (4.5.6).