3 Funkce více proměnných

Naším základním cílem bude zobecnit diferenciální počet funkcí jedné proměnné na funkce *více* proměnných. Budeme uvažovat pouze zobrazení mezi prostory $\mathbb{R}^n$ a $\mathbb{R}^m$, tj. reálné vektorové funkce konečně mnoha reálných proměnných, ačkoliv některé definice a věty je možné vyslovit i pro zobrazení mezi obecnějšími prostory bez větších změn. Odtud dále se budeme také držet konvence, že elementy prostoru $\mathbb{R}^n$ jsou **sloupcové vektory**.

3.1 Derivace funkce více proměnných

Připomeňme si definici derivace funkce jedné proměnné. Nechť $\Omega \subset \mathbb{R}$ je otevřený interval. Funkce $f : \Omega \to \mathbb{R}$ je *diferencovatelná* v bodě $a \in \Omega$, existuje-li konečná limita

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} =: f'(a).$$

Číslo $f'(a)$ se nazývá *derivace* f v bodě a a geometricky je směrnicí tečny ke grafu f v bodě $(a, f(a))$. Je jasné, že tuto definici nelze použít pro funkce $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$, neboť elementy prostorů $\mathbb{R}^m$ a $\mathbb{R}^n$ mezi sebou nelze dělit.

Definici derivace lze ale ekvivalentně vyjádřit ve tvaru

$$\lim_{x \to a} \frac{|f(x) - f(a) - f'(a)(x - a)|}{|x - a|} = 0$$

za předpokladu, že číslo $f'(a)$ existuje. Předpis $Tx := f'(a)x$, kde $x \in \mathbb{R}$, definuje lineární zobrazení $T : \mathbb{R} \to \mathbb{R}$ a funkce f je diferencovatelná v bodě a, pokud lze hodnoty funkce v okolí bodu a „dobře aproximovat“ hodnotami afinní funkce $x \mapsto f(a) + T(x - a)$. Totiž tak, že zbytek $R_a(x) := f(x) - f(a) - T(x - a)$ klesá k nule rychleji než $|x - a|$ pro $x \to a$, tj.

$$\lim_{x \to a} \frac{R_a(x)}{|x - a|} = 0.$$

Tyto úvahy jsou motivací pro následující definici. Připomeňme, že množinu lineárních zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^m$ značíme $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

Definice 3.1 (Diferencovatelná funkce, derivace)**:** Nechť $m, n \in \mathbb{N}$, $\emptyset \neq \Omega \subset \mathbb{R}^n$ je otevřená množina, $f : \Omega \to \mathbb{R}^m$ a $a \in \Omega$. Řekneme, že funkce f je *diferencovatelná v bodě* a, pokud existuje $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ tak, že

$$\lim_{x \to a} \frac{\|f(x) - f(a) - T(x - a)\|_{\mathbb{R}^m}}{\|x - a\|_{\mathbb{R}^n}} = 0.$$

Lineární zobrazení T nazýváme *derivací* f *v bodě* a a značíme $Df(a)$. Je-li f diferencovatelná v každém bodě množiny Ω, říkáme, že f je *diferencovatelná na* Ω.

Poznámka: 1. V Definici 3.1 je $\|\cdot\|_{\mathbb{R}^n}$ nějaká pevně zvolená norma na $\mathbb{R}^n$ a podobně $\|\cdot\|_{\mathbb{R}^m}$. Nicméně z Věty 2.102 víme, že všechny normy na $\mathbb{R}^n$ jsou ekvivalentní, z čehož jednoduše vyplývá, že ani diferencovatelnost ani derivace na volbě normy na $\mathbb{R}^n$ resp. $\mathbb{R}^m$ nezávisí (rozmyslete). Proto dále nebudeme normy indexovat a normou $\|\cdot\|$ budeme rozumět euklidovskou normu na $\mathbb{R}^n$ resp. $\mathbb{R}^m$, nebude-li řečeno jinak.
2. Derivace $Df(a)$ se někdy nazývá ***úplná*** nebo ***totální derivace*** f v bodě a. Definici 3.1 lze vyslovit i pro zobrazení mezi obecnými normovanými prostory. V takovém případě se $Df(a)$ nazývá *Fréchetova derivace.*
3. Limitu z Definice 3.1 lze ekvivalentně psát ve tvaru

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

4. Ještě jedna ekvivalentní formulace diferencovatelnosti je užitečná (dokažte jako Cvičení 3.1). Funkce f je diferencovatelná v a, právě když existuje $T\in\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$, okolí H_a bodu a a funkce $\epsilon_a : H_a\to\mathbb{R}^m$ tak, že

$$(\forall x\in H_a)\left(f(x)=f(a)+T(x-a)+\epsilon_a(x)\|x-a\|\right),$$

kde $\epsilon_a(x)\to 0$ pro $x\to a$.

Zdůrazněme, že derivace $Df(a)$ funkce více proměnných již není číslo, nýbrž lineární zobrazení! V dalším textu budeme pořád uvažovat funkce $f:\Omega\subset\mathbb{R}^n\to\mathbb{R}^m$ definované na nějaké neprázdné otevřené množině Ω, a proto **budeme mlčky otevřenost a neprázdnost** Ω **předpokládat**, aniž bychom to explicitně psali. Podobně m,n budou vždy indexy z množiny $\mathbb{N}$.

Je dobré si uvědomit, že pokud derivace funkce v bodě existuje, je určená jednoznačně.

Lemma 3.2: Derivace diferencovatelné funkce $f:\Omega\subset\mathbb{R}^n\to\mathbb{R}^m$ v bodě $a\in\Omega$ je určena jednoznačně.

Důkaz. Předpokládejme, že dvě zobrazení $T,S\in\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$ jsou derivace f v bodě a. Potom existují funkce $\epsilon_a^{(T)}$ a $\epsilon_a^{(S)}$ definované na okolí počátku H_0 tak, že

$$\lim_{h\to 0}\epsilon_a^{(T)}(h)=\lim_{h\to 0}\epsilon_a^{(S)}(h)=0$$

a

$$\begin{aligned}f(a+h)&=f(a)+Th+\epsilon_a^{(T)}(h)\|h\|,\\ f(a+h)&=f(a)+Sh+\epsilon_a^{(S)}(h)\|h\|\end{aligned}$$

pro všechna $h\in H_0$. Odečtením rovnic dostaneme pro lineární zobrazení $L:=T-S$ vztah

$$Lh=\left(\epsilon_a^{(S)}(h)-\epsilon_a^{(T)}(h)\right)\|h\|$$

pro všechna $h\in H_0$. Odtud plyne, že

$$\lim_{h\to 0}\frac{Lh}{\|h\|}=\lim_{h\to 0}\left(\epsilon_a^{(T)}(h)-\epsilon_a^{(S)}(h)\right)=0.$$

Zvolíme-li libovolné $x \in \mathbb{R}^n$, $x \neq 0$, potom pro $t > 0$ dostatečně malé můžeme položit $h := tx$ a dostaneme

$$0 = \lim_{h \to 0} \frac{Lh}{\|h\|} = \lim_{t \to 0} \frac{L(tx)}{\|tx\|} = \lim_{t \to 0} \frac{tLx}{t\|x\|} = \frac{Lx}{\|x\|}.$$

Tedy $Lx = 0$ pro všechna $x \in \mathbb{R}^n$, což znamená, že $L = 0$. Neboli $T = S$. $\square$

Stejně jako v případě funkcí jedné proměnné diferencovatelnost implikuje spojitost. Než si toto jednoduché, ale důležité tvrzení dokážeme, musíme si nejprve vyjasnit pojem norma lineárního zobrazení a omezenost lineárního zobrazení.

Definice 3.3 (Norma lineárního zobrazení, omezené lineární zobrazení)**:** Buďte $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ normované prostory a $T \in \mathcal{L}(X, Y)$. Číslo $\|T\| \in [0, \infty]$ definované vztahem

$$\|T\| := \sup_{\|x\|_X = 1} \|Tx\|_Y$$

nazýváme *normou lineárního zobrazení* T. Je-li $\|T\| < \infty$, nazýváme T *omezené* lineární zobrazení.

Norma lineárního zobrazení tedy nemusí být konečná, viz Cvičení 3.4. Nicméně množina lineárních zobrazení s konečnou normou je lineární prostor a norma lineárního zobrazení je skutečně normou na tomto prostoru, viz Cvičení 3.2. Norma lineárního zobrazení závisí také na normách $\|\cdot\|_X$ a $\|\cdot\|_Y$ na prostorech X a Y, ačkoliv to ve značení nezdůrazňujeme.

Lemma 3.4: Nechť $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ jsou normované prostory a $T \in \mathcal{L}(X, Y)$. Potom platí:

1. $(\forall x \in X)(\|Tx\|_Y \leq \|T\|\|x\|_X)$. (*V případě, že je $x = 0$ a $\|T\| = \infty$, klademe $\infty \cdot 0 := 0$.*)
2. T je omezené $\Leftrightarrow$ T je spojité $\Leftrightarrow$ T je spojité v 0.
3. Je-li $\dim X < \infty$, potom je T omezený.

Důkaz. 1. Pro $x = 0$ platí nerovnost triviálně. Jinak pro libovolné $x \in X$, $x \neq 0$, máme

$$\|Tx\|_Y = \|x\|_X \left\| T \frac{x}{\|x\|_X} \right\|_Y \leq \|x\|_X \sup_{\|u\|_X = 1} \|Tu\|_Y = \|x\|_X \|T\|.$$

2. Dokážeme řetězec implikací pro $0 \neq T \in \mathcal{L}(X, Y)$. Pro $T = 0$ je tvrzení triviální.

Nechť T je omezené. Buď $x_0 \in X$. Potom pro každé $x \in X$ je

$$\|Tx - Tx_0\|_Y = \|T(x - x_0)\|_Y \leq \|T\|\|x - x_0\|_X.$$

Jelikož je $\|T\| < \infty$, lze k libovolnému $\epsilon > 0$ zvolit číslo $0 < \delta < \epsilon/\|T\|$. S touto volbou pak pro všechna $x \in X$ taková, že $\|x - x_0\|_X < \delta$, platí $\|Tx - Tx_0\|_Y < \epsilon$. Neboli T je spojité v x_0. Protože x_0 bylo libovolné, je T spojité.

Zřejmě, je-li T spojité, je speciálně spojité i v bodě $x = 0$.

Předpokládejme nakonec, že T je spojité v 0. Položíme-li v definici spojitosti $\epsilon = 1$, dostaneme, že

$$(\exists\delta > 0)(\forall x \in X, \|x\|_X < \delta)(\|Tx\|_Y < 1).$$

Potom pro libovolné $u \in X$ s $\|u\|_X = 1$ můžeme odhadovat

$$\|Tu\|_Y = \frac{2}{\delta}\left\|T\left(\frac{\delta}{2}u\right)\right\|_Y < \frac{2}{\delta},$$

neboť $\|\delta u/2\|_X = \delta/2 < \delta$. Z toho plyne, že $\|T\| \leq 2/\delta$, a tudíž je T omezené.
3. Je-li $\dim X = 0$ je tvrzení triviální. Předpokládejme, že $\dim X = n \in \mathbb{N}$ a zvolme v X bázi $\mathcal{X} = (x_1, \ldots, x_n)$. Potom zobrazení $\|\cdot\|_{\mathcal{X},\infty} : X \to [0,\infty)$ definované vztahem

$$\|x\|_{\mathcal{X},\infty} := \max_{i\in\hat{n}} |x_i^\#(x)|$$

je norma v X. Protože je $\dim X < \infty$, jsou všechny normy na X ekvivalentní (Věta 2.102), a tudíž

$$(\exists K > 0)(\forall x \in X)(\|x\|_{\mathcal{X},\infty} \leq K\|x\|_X).$$

Odtud dále dostaneme pro $x \in X$ odhad

$$\begin{aligned}\|Tx\|_Y &= \left\|T\left(\sum_{i=1}^n x_i^\#(x)x_i\right)\right\|_Y = \left\|\sum_{i=1}^n x_i^\#(x)Tx_i\right\|_Y \leq \sum_{i=1}^n |x_i^\#(x)|\,\|Tx_i\|_Y \\ &\leq \|x\|_{\mathcal{X},\infty}\sum_{i=1}^n \|Tx_i\|_Y \leq K\|x\|_X\sum_{i=1}^n \|Tx_i\|_Y = C\|x\|_X,\end{aligned}$$

kde jsme označili

$$C := K\sum_{i=1}^n \|Tx_i\|_Y\,.$$

Z nerovnosti $\|Tx\|_Y \leq C\|x\|_X$, která platí pro všechna $x \in X$, plyne, že $\|T\| \leq C$, a tudíž je T omezené. □

Věta 3.5: Je-li $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ diferencovatelná v bodě $a \in \Omega$, potom je f spojitá v bodě a.

Důkaz. Protože je $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ diferencovatelná v a, platí pro všechna x z nějakého okolí H_a rovnost

$$f(x) = f(a) + Df(a)(x-a) + \epsilon_a(x)\|x-a\|,$$

kde $e_a(x) \to 0$ pro $x \to a$. Pošleme-li v poslední rovnosti $x \to a$, jde třetí člen napravo k nule. Totéž platí i o druhém členu, neboť z tvrzení 3. a 1. Lemma 3.4 plyne, že

$$\|Df(a)(x-a)\| \leq \underbrace{\|Df(a)\|}_{<\infty}\,\|x-a\|$$

pro všechna $x \in H_a$. Celkem tedy je $\lim_{x\to a} f(x) = f(a)$, neboli f je spojité v a. □

Další dvě tvrzení se týkají pravidel pro derivování funkcí. Důkaz prvního tvrzení je jednoduchou modifikací důkazu analogické věty pro funkce jedné proměnné. Jediný podstatný rozdíl je v tom, že místo absolutních hodnot píšeme příslušné normy. Důkaz je proto přenechán čtenáři jako cvičení.

Věta 3.6: Nechť $f, g : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ jsou diferencovatelné v bodě $a \in \Omega$. Potom platí:

1. Pro $c \in \mathbb{R}$ je funkce $f + cg$ diferencovatelná v a a platí
$$D(f+cg)(a) = Df(a) + cDg(a).$$
2. Je-li $m = 1$, je součin fg diferencovatelná funkce v a a platí
$$D(fg)(a) = g(a)Df(a) + f(a)Dg(a).$$
3. Je-li $m = 1$ a $g(a) \neq 0$, je podíl f/g diferencovatelná funkce v a a platí
$$D\left(\frac{f}{g}\right)(a) = \frac{g(a)Df(a) - f(a)Dg(a)}{g^2(a)}.$$

Důkaz. Proveďte jako Cvičení 3.5. □

Tvrzení následující věty o derivaci složené funkce se někdy nazývá *řetězové pravidlo* (chain rule), viz také Důsledek 3.19.

Věta 3.7 (Derivace složené funkce)**:** Nechť $\Omega \subset \mathbb{R}^n$ je otevřená, $f : \Omega \to \mathbb{R}^m$ a $g : U \to \mathbb{R}^p$, kde $U \subset \mathbb{R}^m$ je otevřená taková, že $f(\Omega) \subset U$. Je-li f diferencovatelná v bodě $a \in \Omega$ a g diferencovatelná v $f(a)$, potom je složená funkce $g \circ f$ diferencovatelná v a a platí
$$D(g \circ f)(a) = Dg\left(f(a)\right) Df(a).$$

Důkaz. Označme $b := f(a)$. Z předpokladu diferencovatelnosti funkcí f a g plyne, že
$$f(x) = f(a) + Df(a)(x-a) + \epsilon_1(x)\|x-a\|$$
pro x z nějakého okolí H_a bodu a a podobně
$$g(y) = g(b) + Dg(b)(y-b) + \epsilon_2(y)\|y-b\|$$
pro y z nějakého okolí H_b bodu b, kde $\epsilon_1(x) \to 0$ pro $x \to a$ a $\epsilon_2(y) \to 0$ pro $y \to b$. Jelikož je f spojitá v bodě a, viz Věta 3.5, můžeme bez újmy na obecnosti předpokládat, že $f(H_a) \subset H_b$. Položíme-li nyní $y = f(x)$ a využijeme-li obou předchozích rovnic, dostaneme
$$\begin{aligned} g(f(x)) &= g(f(a)) + Dg(f(a))\left(Df(a)(x-a) + \epsilon_1(x)\|x-a\|\right) \\ &\quad + \epsilon_2(f(x))\left\|Df(a)(x-a) + \epsilon_1(x)\|x-a\|\right\| \\ &= g(f(a)) + Dg(f(a))Df(a)(x-a) + \|x-a\|Dg(f(a))\epsilon_1(x) \\ &\quad + \epsilon_2(f(x))\left\|Df(a)\frac{x-a}{\|x-a\|} + \epsilon_1(x)\right\|\|x-a\| \end{aligned}$$

pro všechna $x \in H_a \setminus \{a\}$. Neboli

$$g(f(x)) = g(f(a)) + Dg(f(a))Df(a)(x-a) + \epsilon(x)\|x-a\|,$$

pro všechna $x \in H_a \setminus \{a\}$, kde jsme označili

$$\epsilon(x) := Dg(f(a))\epsilon_1(x) + \epsilon_2(f(x)) \left\| Df(a)\frac{x-a}{\|x-a\|} + \epsilon_1(x) \right\|.$$

Důkaz tvrzení bude kompletní, ukážeme-li, že $\epsilon(x) \to 0$ pro $x \to a$. K tomu stačí využít omezenosti derivací jakožto lineárních zobrazení, viz Lemma 3.4, a trojúhelníkové nerovnosti, odkud plyne, že

$$\|\epsilon(x)\| \leq \|Dg(f(a))\|\|\epsilon_1(x)\| + \|\epsilon_2(f(x))\| \left(\|Df(a)\| + \|\epsilon_1(x)\|\right)$$

pro všechna $x \in H_a \setminus \{a\}$. Jelikož $\|\epsilon_1(x)\| \to 0$ pro $x \to a$ a vzhledem ke spojitosti f v a také $\|\epsilon_2(f(x))\| \to 0$ pro $x \to a$, vidíme, že skutečně

$$\lim_{x \to a} \epsilon(x) = 0.$$

□

3.2 Parciální a směrové derivace

Nyní se budeme chvíli zabývat pouze **skalárními** funkcemi $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ definovaných na otevřené množině Ω a diferencovatelných v bodě $a \in \Omega$. V tomto případě je $Df(a)$ lineární funkcionál na $\mathbb{R}^n$, tj. $Df(a) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$. Z Rieszovy věty vyplývá, že existuje právě jeden vektor $w \in \mathbb{R}^n$ takový, že pro všechna $x \in \mathbb{R}^n$ platí

$$w^T x = Df(a)x, \tag{17}$$

neboť $w^T x = (w, x)$ a $(\cdot, \cdot)$ značí euklidovský skalární součin na $\mathbb{R}^n$.

Definice 3.8 (Gradient)**:** Buď $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ diferencovatelná v $a \in \Omega$. Řádkový vektor w^T splňující (17) pro každé $x \in \mathbb{R}^n$ nazýváme *gradient* funkce f v bodě a a značíme $\nabla f(a)$.

Tedy pro funkci $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ diferencovatelnou v bodě $a \in \Omega$ můžeme psát:

$$f(x) = f(a) + \nabla f(a)(x-a) + \epsilon_a(x)\|x-a\|$$

pro x z nějakého okolí H_a, kde $\epsilon_a(x) \to 0$ pro $x \to a$. Naším cílem bude najít vhodné explicitní vyjádření pro gradient f.

Definice 3.9 (Směr, směrová derivace)**:** Libovolný nenulový vektor $v \in \mathbb{R}^n$ nazýváme *směr* v $\mathbb{R}^n$. Buď $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$, $v \in \mathbb{R}^n \setminus \{0\}$ a $a \in \Omega$. Existuje-li konečná limita

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

nazýváme ji *(směrová) derivace f ve směru v a v bodě a.*

Všimněte si, že funkce $g(t) := f(a + tv)$, která je zúžením funkce f na přímku $t \to a + tv$ procházející bodem a se směrnicí v, splňuje

$$g'(0) = \lim_{t \to 0} \frac{g(t) - g(0)}{t} = \lim_{t \to 0} \frac{f(a + tv) - f(a)}{t} =: D_v f(a).$$

Proto směrová derivace $D_v f(a)$ udává „míru změny“ f ve směru v na okolí bodu a.

Věta 3.10: Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ je diferencovatelná v bodě $a \in \Omega$. Potom pro každý směr $v \in \mathbb{R}^n \setminus \{0\}$ existuje směrová derivace $D_v f(a)$ a platí:

$$D_v f(a) = \nabla f(a) v.$$

Důkaz. Z diferencovatelnosti f v bodě a máme

$$f(x) = f(a) + \nabla f(a)(x - a) + \epsilon_a(x)\|x - a\|$$

pro $x \in H_a$, kde H_a je okolí a a $\epsilon_a(x) \to 0$ pro $x \to a$. Položíme-li $x = a + tv$ pro $t \in \mathbb{R}$ dostatečně malé, aby $a + tv \in H_a$, dostaneme z poslední rovnosti

$$f(a + tv) = f(a) + t\nabla f(a)v + \epsilon_a(a + tv)|t|\|v\|,$$

neboli

$$\left| \frac{f(a + tv) - f(a)}{t} - \nabla f(a)v \right| = |\epsilon_a(a + tv)| \, \|v\|.$$

Jelikož výraz napravo jde k nule pro $t \to 0$, máme

$$\lim_{t \to 0} \frac{f(a + tv) - f(a)}{t} = \nabla f(a)v,$$

což implikuje tvrzení věty. □

Důsledek 3.11: Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ je diferencovatelná v bodě $a \in \Omega$, $u, v \in \mathbb{R}^n \setminus \{0\}$ a $c \in \mathbb{R}$ tak, že $u + cv \neq 0$. Potom

$$D_{u+cv} f(a) = D_u f(a) + c D_v f(a).$$

Ještě jeden důsledek Věty 3.10 stojí za pozornost, neboť ukazuje jeden **geometrický význam gradientu**. Z Cauchy–Schwarzovy nerovnosti plyne, že

$$|D_v f(a)| = |\nabla f(a) v| \leq \|\nabla f(a)\| \|v\|$$

a rovnost nastává, pokud je v skalárním násobkem $\nabla f(a)$. Odtud a z Věty 3.10 vidíme, že mezi všemi vektory v se zafixovanou normou rovnou $\|\nabla f(a)\|$ je hodnota směrové derivace $D_v f(a)$ maximální pro $v = \nabla f(a)$ a minimální pro $v = -\nabla f(a)$, pokud $\nabla f(a) \neq 0$. To lze slovy interpretovat tak, že gradient $\nabla f(a)$ udává směr, v němž funkční hodnoty f nejvíce rostou („kudy je to po grafu funkce f nejvíce do kopce, vyrážíme-li z bodu a“). Podobně vektor $-\nabla f(a)$ udává směr, v němž funkční hodnoty f nejvíce klesají.

Vezmeme-li v definici směrové derivace za směr speciálně vektory standardní báze e_i, $i \in \hat{n}$, prostoru $\mathbb{R}^n$ dostaneme se k tzv. *derivaci parciální.*

Definice 3.12 (Parciální derivace)**:** Buď $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$, $i \in \hat{n}$ a $a \in \Omega$. Existuje-li konečná limita
$$\lim_{t\to 0} \frac{f(a+te_i)-f(a)}{t} =: \frac{\partial f}{\partial x_i}(a) \equiv \partial_{x_i} f(a) \equiv \partial_i f(a),$$
kde e_i je i-tý vektor standardní báze $\mathbb{R}^n$, nazýváme ji *parciální derivace f podle i-té proměnné v bodě a.*

Poznámka: Všimněte si, že definice parciální derivace f podle i-té proměnné v bodě a splývá s obvyklou derivací funkce jedné proměnné $g_i : \mathbb{R} \to \mathbb{R} : s \mapsto f(a_1, \ldots, a_{i-1}, s, a_{i+1}, \ldots, a_n)$, tj. funkce, kterou získáme z f zafixováním všech proměnných kromě i-té na okolí a, protože
$$\frac{\partial f}{\partial x_i}(a) = \lim_{t\to 0} \frac{f(a+te_i)-f(a)}{t} = \lim_{t\to 0} \frac{g_i(a_i+t)-g_i(a)}{t} = g_i'(a_i).$$

Je zřejmé, že parciální derivace je speciálním případem derivace směrové, neboť $\partial_{x_i} f(a) = D_{e_i} f(a)$. Položíme-li $v = e_i$ ve Větě 3.10, zjistíme, že $\partial_{x_i} f(a)$ je i-tou souřadnicí gradientu $\nabla f(a)$. Dostáváme tak důležité pozorování, že pro $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ diferencovatelnou v bodě $a \in \Omega$ platí:
$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a) \right).$$

Navíc, protože $Df(a)x = \nabla f(a)x$ pro každé $x \in \mathbb{R}^n$, je vektor $\nabla f(a)$ současně maticí zobrazení $Df(a) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ vzhledem ke standardním bázím. Připomeňme, že mezi lineárním zobrazením $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ a jeho maticí vzhledem k bázím $\mathcal{X} = (x_1, \ldots x_n)$ a $\mathcal{Y} = (y_1, \ldots y_m)$ prostorů $\mathbb{R}^n$ a $\mathbb{R}^m$ platí vztah
$$\left({}^{\mathcal{X}}A^{\mathcal{Y}} \right)_{j,i} = y_j^{\#}(Ax_i)$$
pro všechna $i \in \hat{n}$ a $j \in \hat{m}$. Vyjádření matice derivace $Df(a)$ vzhledem ke standardním bázím pomocí derivací parciálních není speciální vlastností skalárních funkcí, ale platí obecně i pro funkce vektorové.

Věta 3.13: Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ je diferencovatelná v bodě $a \in \Omega$. Označme komponenty $f = (f_1, \ldots f_m)$. Potom pro každé $i \in \hat{n}$ a $j \in \hat{m}$ existuje $\partial_{x_i} f_j$ v bodě a a platí:
$$ {}^{\mathcal{E}_n}(Df(a))^{\mathcal{E}_m} = \begin{pmatrix} \partial_{x_1} f_1(a) & \partial_{x_2} f_1(a) & \ldots & \partial_{x_n} f_1(a) \\ \partial_{x_1} f_2(a) & \partial_{x_2} f_2(a) & \ldots & \partial_{x_n} f_2(a) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{x_1} f_m(a) & \partial_{x_2} f_m(a) & \ldots & \partial_{x_n} f_m(a) \end{pmatrix}.$$

Poznámka: Je obvyklé ztotožnit lineární zobrazení z $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ a jejich matice z $\mathbb{R}^{m,n}$ vzhledem ke standardním bázím. I my **odtud dále ztotožníme** $Df(a)$ a ${}^{\mathcal{E}_n}(Df(a))^{\mathcal{E}_m}$ a budeme jednoduše psát $Df(a)$ v obou případech. To, zda $Df(a)$ představuje lineární zobrazení, či jeho matici vzhledem ke standardním bázím, bude vždy jasné z kontextu.

Důkaz Věty 3.13. Nejprve si uvědomíme, že konvergence v obvyklé topologii $\mathbb{R}^n$ je ekvivalentní konvergenci v jednotlivých složkách, tj.
$$\lim_{x\to a} \phi(x) = c \quad \Leftrightarrow \quad (\forall j \in \hat{m}) \left(\lim_{x\to a} \phi_j(x) = c_j \right)$$

kde $\phi : \mathbb{R}^n \to \mathbb{R}^m$, $a \in \mathbb{R}^n$ a $c \in \mathbb{R}^m$. Skutečně vzhledem k ekvivalenci norem na $\mathbb{R}^m$, viz Věta 2.66, je $\phi(x) \to c$ pro $x \to a$, právě když

$$\lim_{x \to a} \|\phi(x) - c\|_\infty = \lim_{x \to a} \max_{j \in \hat{m}} |\phi_j(x) - c_j| = 0,$$

z čehož tvrzení už přímo vyplývá. Tento fakt budeme často používat.

Vraťme se k důkazu věty. Předpoklad diferencovatelnosti f v bodě a lze psát jako limitu

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

Odtud ve složkách pro libovolné $j \in \hat{m}$ dostaneme

$$\lim_{h \to 0} \frac{|f_j(a+h) - f_j(a) - (Df(a)h)_j|}{\|h\|} = 0.$$

Položíme-li $h = te_i$, kde $i \in \hat{n}$ a $t \in \mathbb{R}$ dostatečně malé, vyplývá z předchozí rovnosti, že

$$\lim_{t \to 0} \frac{|f_j(a+te_i) - f_j(a) - t\,(Df(a)e_i)_j|}{|t|\|e_i\|} = 0.$$

Neboli

$$\lim_{t \to 0} \left| \frac{f_j(a+te_i) - f_j(a)}{t} - (Df(a))_{j,i} \right| = 0,$$

kde $(Df(a))_{j,i} = (Df(a)e_i)_j$ značí (j,i)-tý element matice ${}^{\mathcal{E}_n}(Df(a))^{\mathcal{E}_m}$. Zjistili jsme tedy, že $\partial_{x_i} f_j(a)$ existuje a platí pro ni rovnost

$$\frac{\partial f_j}{\partial x_i}(a) = (Df(a))_{j,i}\,.$$

$\square$

Definice 3.14 (Jacobiho matice zobrazení, jakobián)**:** Matici

$$\frac{\partial(f_1, \ldots, f_m)}{\partial(x_1, \ldots, x_n)}(a) := \begin{pmatrix} \partial_{x_1} f_1(a) & \partial_{x_2} f_1(a) & \ldots & \partial_{x_n} f_1(a) \\ \partial_{x_1} f_2(a) & \partial_{x_2} f_2(a) & \ldots & \partial_{x_n} f_2(a) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{x_1} f_m(a) & \partial_{x_2} f_m(a) & \ldots & \partial_{x_n} f_m(a) \end{pmatrix}$$

nazýváme *Jacobiho maticí zobrazení* $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ v bodě a za předpokladu, že parciální derivace $\partial_{x_i} f_j(a)$ existují pro všechna $i \in \hat{n}$ a $j \in \hat{m}$. Pokud je $m = n$, nazýváme determinant Jacobiho matice f *jakobián* f.

Tedy maticí derivace diferencovatelného zobrazení $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ v bodě $a \in \Omega$ (vzhledem ke standardním bázím) je jeho Jacobiho matice

$$Df(a) = \frac{\partial(f_1, \ldots, f_m)}{\partial(x_1, \ldots, x_n)}(a).$$

Jacobiho matice f je ovšem definovaná, existují-li všechny parciální derivace f, které tvoří elementy této matice. Nabízí se tedy otázka, zda existence všech parciálních derivací $\partial_{x_i} f_j(a)$ pro $i \in \hat{n}$ a $j \in \hat{m}$ implikuje diferencovatelnost f v bodě a. Následující příklad ukazuje, že tomu tak není. Zesílíme-li ovšem předpoklad tak, že budeme požadovat existenci všech parciálních derivací f na nějakém okolí a a jejich spojitost v a, bude již diferencovatelnost f v bodě a zaručena, viz Věta 3.16.

Příklad 3.15: Mějme funkci $f : \mathbb{R}^2 \to \mathbb{R}$ definovanou vztahem

$$f(x, y) = \begin{cases} 0, & \text{pokud } x = 0 \text{ nebo } y = 0, \\ 1, & \text{jinak.} \end{cases}$$

Potom

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{t \to 0} \frac{f(t, 0) - f(0, 0)}{t} = \lim_{t \to 0} \frac{0}{t} = 0.$$

Podobně vyjde $\partial_y f(0, 0) = 0$. Avšak funkce f není v bodě $(0, 0)$ spojitá (ověřte), a proto podle Věty 3.5 nemůže být f ani diferencovatelná v $(0, 0)$. Tento fakt je důsledkem toho, že parciální derivace funkce závisejí jen na chování funkce ve směrech kartézských os (tj. vektorů standardní báze) v okolí a, kdežto derivace f zohledňuje chování f na celém okolí a. Ani existence všech směrových derivací neimplikuje diferencovatelnost, jak ukazuje Cvičení 3.9.

Věta 3.16: Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ a $a \in \Omega$. Existují-li parciální derivace $\partial_{x_i} f_j$ na okolí a a jsou spojité v a pro všechna $i \in \hat{n}$ a $j \in \hat{m}$, potom je f diferencovatelné v a.

Důkaz. Má-li derivace f v bodě a existovat, musí být dána Jacobiho maticí f v a. Chceme dokázat, že

$$\lim_{x \to a} \frac{\|f(x) - f(a) - T(x - a)\|}{\|x - a\|} = 0.$$

kde jsme označili

$$T := \frac{\partial(f_1, \ldots, f_m)}{\partial(x_1, \ldots, x_n)}(a).$$

K tomu stačí ukázat konvergenci v jednotlivých složkách

$$\lim_{x \to a} \frac{|f_j(x) - f_j(a) - (T(x - a))_j|}{\|x - a\|} = 0.$$

pro všechna $j \in \hat{m}$ (viz začátek důkazu Věty 3.13). Odtud plyne, že větu stačí dokázat pro speciální případ $m = 1$, tj. pro skalární funkci $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ bez újmy na obecnosti. Chceme tedy ukázat, že

$$\lim_{x \to a} \frac{|f(x) - f(a) - \nabla f(a)(x - a)|}{\|x - a\|} = 0.$$

Nechť B_a označuje nějakou kouli se středem v a poloměrem dostatečně malým, aby na ní existovaly všechny parciální derivace $\partial_{x_i} f$, $i \in \hat{n}$. Nejprve si pro $x \in B_a$ zapíšeme rozdíl $f(x) - f(a)$ následujícím způsobem:

$$f(x) - f(a) = \sum_{i=1}^{n} \left(f(v_i) - f(v_{i-1}) \right), \tag{18}$$

kde $v_0 = a$, $v_i = (x_1, \ldots, x_i, a_{i+1}, \ldots, a_n)$ pro $i \in \widehat{n-1}$ a $v_n = x$. Funkce

$$g_i(t) := f(x_1, \ldots, x_{i-1}, t, a_{i+1}, \ldots, a_n)$$

zobrazuje uzavřený interval s koncovými body a_i a x_i do $\mathbb{R}$, je na něm spojitá a diferencovatelná na jeho vnitřku, neboť

$$g_i'(t) = \frac{\partial f}{\partial x_i}(x_1, \ldots, x_{i-1}, t, a_{i+1}, \ldots, a_n)$$

a parciální derivace vpravo existují podle předpokladů. Zde je třeba si uvědomit, že

$$(x_1, \ldots, x_{i-1}, t, a_{i+1}, \ldots, a_n) \in B_a$$

pro každé t z uzavřeného intervalu s koncovými body a_i a x_i a $i \in \hat{n}$ (ověřte). Můžeme proto na g_i aplikovat Lagrangeovu větu o přírůstku. Přihlédneme-li ještě k tomu, že $g_i(x_i) = f(v_i)$ a $g_i(a_i) = f(v_{i-1})$, dostaneme

$$f(v_i) - f(v_{i-1}) = (x_i - a_i) \frac{\partial f}{\partial x_i}(x_1, \ldots, x_{i-1}, \xi, a_{i+1}, \ldots, a_n), \tag{19}$$

kde ξ_i leží uvnitř intervalu s koncovými body a_i a x_i.

Označme $u_i := (x_1, \ldots, x_{i-1}, \xi, a_{i+1}, \ldots, a_n)$. Potom využijeme-li rovnic (18), (19) a nerovnosti $|x_i - a_i| \leq \|x - a\|$, která platí pro euklidovskou normu, dostaneme

$$\begin{aligned} \frac{|f(x) - f(a) - \sum_{i=1}^n \partial_{x_i} f(a)(x_i - a_i)|}{\|x - a\|} &= \frac{|\sum_{i=1}^n \left[\partial_{x_i} f(u_i) - \partial_{x_i} f(a)\right](x_i - a_i)|}{\|x - a\|} \\ &\leq \sum_{i=1}^{n} \left| \frac{\partial f}{\partial x_i}(u_i) - \frac{\partial f}{\partial x_i}(a) \right|. \end{aligned}$$

Poslední výraz jde k 0, pokud $x \to a$, neboť $u_i \to a$ pro $x \to a$ a funkce $\partial_{x_i} f$ jsou spojité v a pro všechna $i \in \hat{n}$ podle předpokladu. Odtud plyne rovnost

$$\lim_{x \to a} \frac{|f(x) - f(a) - \nabla f(a)(x - a)|}{\|x - a\|} = 0,$$

což jsme chtěli dokázat. □

Následující příklad ukazuje, že kritérium diferencovatelnosti z Věty 3.16 je pouze postačující podmínka.

Příklad 3.17: Vyšetříme diferencovatelnost funkce $f : \mathbb{R}^2 \to \mathbb{R}$ definované vztahem

$$f(x,y) = \begin{cases} (x^2+y^2)\sin\frac{1}{\sqrt{x^2+y^2}}, & \text{je-li } (x,y) \neq (0,0), \\ 0, & \text{je-li } (x,y) = (0,0). \end{cases}$$

V bodech $(x,y) \neq (0,0)$ spočítáme parciální derivace f přímým výpočtem:

$$\frac{\partial f}{\partial x}(x,y) = 2x\sin\frac{1}{\sqrt{x^2+y^2}} - \frac{x}{\sqrt{x^2+y^2}}\cos\frac{1}{\sqrt{x^2+y^2}},$$
$$\frac{\partial f}{\partial y}(x,y) = 2y\sin\frac{1}{\sqrt{x^2+y^2}} - \frac{y}{\sqrt{x^2+y^2}}\cos\frac{1}{\sqrt{x^2+y^2}}.$$

Jelikož jsou obě funkce $\partial_x f$ a $\partial_y f$ spojité na $\mathbb{R}^2 \setminus \{(0,0)\}$, je podle Věty 3.16 funkce f diferencovatelná na $\mathbb{R}^2 \setminus \{(0,0)\}$.

Pro bod $(x,y) = (0,0)$ máme

$$\frac{\partial f}{\partial x}(0,0) = \lim_{t\to 0}\frac{f(t,0)-f(0,0)}{t} = \lim_{t\to 0} t\sin\frac{1}{t} = 0$$

a podobně je také $\partial_y f(0,0) = 0$. Nicméně zúžíme-li $\partial_x f$ na osu x, dostaneme pro $x \neq 0$ z předchozího výpočtu

$$\frac{\partial f}{\partial x}(x,0) = 2x\sin\frac{1}{|x|} - \frac{x}{|x|}\cos\frac{1}{|x|}.$$

Tato funkce nemá limitu pro $x \to 0$, a proto parciální derivace $\partial_x f$ není v bodě $(0,0)$ spojitá. Stejný závěr dostaneme i pro $\partial_y f$. Ukážeme nicméně, že f je diferencovatelná i v bodě $(0,0)$, ačkoliv v tomto případě nemůžeme použít Větu 3.16. K tomu stačí ukázat, že funkce

$$\epsilon(x,y) := \frac{f(x,y)-f(0,0)-\partial_x f(0,0)x - \partial_y f(0,0)y}{\|(x,y)\|} = \frac{f(x,y)}{\sqrt{x^2+y^2}}$$

jde k 0 pro $(x,y) \to (0,0)$. To je ovšem pravda, neboť pro $(x,y) \neq (0,0)$ platí odhad

$$\frac{|f(x,y)|}{\sqrt{x^2+y^2}} \leq \sqrt{x^2+y^2}.$$

Příklad 3.18: Vyšetříme diferencovatelnost funkce

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

V bodech množiny $\mathbb{R}^2 \setminus \{(0,0)\}$ je f diferencovatelná, neboť podobně jako v předchozím příkladě obě parciální derivace

$$\frac{\partial f}{\partial x}(x,y) = \frac{y}{\sqrt{x^2+y^2}} - \frac{x^2y}{(x^2+y^2)^{3/2}},$$
$$\frac{\partial f}{\partial y}(x,y) = \frac{x}{\sqrt{x^2+y^2}} - \frac{xy^2}{(x^2+y^2)^{3/2}}$$

existují a jsou spojité na $\mathbb{R}^2 \setminus \{(0,0)\}$. Snadno také spočítáme, že

$$\frac{\partial f}{\partial x}(0,0) = \lim_{t\to 0} \frac{f(t,0) - f(0,0)}{t} = 0 \quad \text{a} \quad \frac{\partial f}{\partial y}(0,0) = \lim_{t\to 0} \frac{f(0,t) - f(0,0)}{t} = 0.$$

Můžeme se analogickým způsobem jako v předchozí úloze také přesvědčit, že ani nyní nejsou funkce $\partial_x f$ a $\partial_y f$ spojité v bodě $(0,0)$. Z toho nemůžeme udělat žádný závěr o diferencovatelnosti f v počátku a musíme opět analyzovat funkci

$$\epsilon(x,y) := \frac{f(x,y) - f(0,0) - \partial_x f(0,0)x - \partial_y f(0,0)y}{\|(x,y)\|} = \frac{xy}{x^2+y^2},$$

kde $(x,y) \neq (0,0)$. Funkce f je diferencovatelná v počátku, právě když $\epsilon(x,y) \to 0$ pro $(x,y) \to (0,0)$. To ale neplatí, neboť

$$\epsilon(x,x) = \frac{1}{2}, \quad \forall x \neq 0,$$

a proto f není diferencovatelná v bodě $(0,0)$. Všimněte si ale, že f je spojitá v bodě $(0,0)$. To vyplývá z definice spojitosti a odhadu

$$|f(x,y)| = \frac{|xy|}{\sqrt{x^2+y^2}} \leq \frac{1}{2}\sqrt{x^2+y^2}.$$

Věta 3.13 umožňuje vyjádřit řetězové pravidlo z Věty 3.7 pomocí parciálních derivací, neboť za předpokladů Věty 3.7 dostaneme přechodem k maticím vztah

$$D(g\circ f)(a) = \begin{pmatrix} \partial_{y_1} g_1(f(a)) & \dots & \partial_{y_m} g_1(f(a)) \\ \vdots & \ddots & \vdots \\ \partial_{y_1} g_p(f(a)) & \dots & \partial_{y_m} g_p(f(a)) \end{pmatrix} \cdot \begin{pmatrix} \partial_{x_1} f_1(a) & \dots & \partial_{x_n} f_1(a) \\ \vdots & \ddots & \vdots \\ \partial_{x_1} f_m(a) & \dots & \partial_{x_n} f_m(a) \end{pmatrix},$$

který lze ekvivalentně také napsat ve tvaru

$$\frac{\partial\left((g\circ f)_1, \dots, (g\circ f)_p\right)}{\partial(x_1,\dots,x_n)}(a) = \frac{\partial\left(g_1,\dots,g_p\right)}{\partial(y_1,\dots,y_m)}(f(a)) \cdot \frac{\partial\left(f_1,\dots,f_m\right)}{\partial(x_1,\dots,x_n)}(a).$$

Srovnáme-li (k,i)-té elementy matic na obou stranách, dostaneme následující důsledek.

Důsledek 3.19 (Řetězové pravidlo)**:** Za předpokladů Věty 3.7 platí:

$$\frac{\partial (g\circ f)_k}{\partial x_i}(a) = \sum_{j=1}^{m} \frac{\partial g_k}{\partial y_j}(f(a)) \frac{\partial f_j}{\partial x_i}(a)$$

pro každé $i \in \hat{n}$ a $k \in \hat{p}$.

Dva speciální případy se objevují častěji. První je případ skalární funkce $g : U \subset \mathbb{R}^m \to \mathbb{R}$. tj. $p = 1$. Potom

$$\frac{\partial (g\circ f)}{\partial x_i}(a) = \sum_{j=1}^{m} \frac{\partial g}{\partial y_j}(f(a)) \frac{\partial f_j}{\partial x_i}(a), \quad i \in \hat{n}.$$

V druhém ještě speciálnějším případě skládáme skalární funkci $f : U \subset \mathbb{R}^n \to \mathbb{R}$ s vektorovou funkcí jedné proměnné $\phi : \Omega \subset \mathbb{R} \to \mathbb{R}^n$. V takovém případě je $f \circ \phi$ skalární funkce jedné reálné proměnné, pro jejíž derivaci dostáváme

$$(f \circ \phi)'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\phi(t))\phi_i'(t) = \nabla f(\phi(t))\phi'(t),$$

kde jsme pro tentokrát použili značení $\phi'(t) \equiv D\phi(t) = (\phi_1'(t), \ldots, \phi_m'(t))^T$.

Vektorová funkce ϕ, je-li spojitá, je křivkou v $\mathbb{R}^m$. Je-li navíc diferencovatelná v nějakém bodě t, představuje vektor $\phi'(t)$ *tečný vektor* ke křivce ϕ v bodě t. Fyzikální interpretace stopy křivky ϕ může být trajektorie, po které se pohybuje částice. Vektor $\phi'(t)$ je potom vektor rychlosti částice v čase t (velocity) a euklidovská norma $\|\phi'(t)\|$ je rychlost částice v t (speed).

Dále se podíváme na vztah gradientu funkce $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ a tzv. *vrstevnic* funkce f, tj. množin $S_c := \{x \in \Omega \mid f(x) = c\}$, kde $c \in \mathbb{R}$. Předpokládejme, že f je diferencovatelná v $a \in \Omega$. Potom lze stručně říct, že gradient $\nabla f(a)$ je kolmý na vrstevnici $S_{f(a)}$. Přesný smysl předchozí věty je následující: *Je-li ϕ libovolná křivka s hodnotami v $S_{f(a)}$ definovaná na okolí* 0, *diferencovatelná v* 0 *a $\phi(0) = a$, potom je $\nabla f(a)$ kolmý na tečný vektor* $\phi'(0)$. Skutečně podle řetězového pravidla je

$$\nabla f(a)\phi'(0) = (f \circ \phi)'(0) = 0.$$

Nulovost derivace $f \circ \phi$ plyne z toho, že $f \circ \phi$ je konstantní funkce, neboť podle předpokladu $f(\phi(t)) = f(a)$ pro všechna t z okolí 0. Je-li $\nabla f(a) \neq 0$, představuje množina tečných vektorů v a ke křivkám z vrstevnice $S_{f(a)}$ nadrovinu v $\mathbb{R}^n$ s normálovým vektorem $\nabla f(a)$. Tato nadrovina je určena rovnicí

$$\nabla f(a)(x - a) = 0.$$

Tyto úvahy nás motivují k následující definici.

Definice 3.20 (Tečná nadrovina k vrstevnici, tečný prostor)**:** Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ je diferencovatelná funkce v $a \in \Omega$. Pokud $\nabla f(a) \neq 0$, nazýváme množinu

$$\{x \in \mathbb{R}^n \mid \nabla f(a)(x - a) = 0\}$$

tečnou nadrovinou k vrstevnici $S_{f(a)}$ v bodě a. Zaměření tečné nadroviny, tj. ortogonální doplněk k $\nabla f(a)$ nazýváme *tečný prostor k $S_{f(a)}$ v bodě* a a značíme $T_a(S_{f(a)})$.

Příklad 3.21 (Tečná nadrovina ke grafu funkce)**:** Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ je diferencovatelná funkce v $a \in \Omega$. Její graf $\{(x, f(x)) \in \mathbb{R}^{n+1} \mid x \in \Omega\}$ můžeme chápat také jako vrstevnici $S := \{(x, y) \in \Omega \times \mathbb{R} \mid F(x, y) = 0\}$ funkce $F(x, y) := f(x) - y$. Buď $a \in \Omega$. Protože

$$\nabla F(x, y) = \left(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x), -1\right)$$

lze rovnici tečné nadroviny k S v $(a, f(a))$

$$\nabla F(a, f(a))\,[(x, y) - (a, f(a))] = 0$$

vyjádřit jako rovnici

$$y = f(a) + \nabla f(a)(x - a),$$

kterou nazýváme rovnice *tečné nadroviny ke grafu funkce f v bodě* $(a, f(a))$. Rozepíšeme-li ještě poslední rovnici, dostaneme

$$y = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i).$$

Ve speciálním případě funkce jedné proměnné ($n = 1$), je rovnice tečné nadroviny ke grafu f v bodě $(x_0, f(x_0))$ známá rovnice tečné přímky

$$y = f(x_0) + f'(x_0)(x - x_0).$$

V případě funkce dvou proměnných ($n = 2$) má rovnice tečné (nad)roviny ke grafu f v bodě $(x_0, y_0, f(x_0, y_0))$ tvar

$$z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)\,(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)\,(y - y_0)\,.$$

Všimněte si dimenzionální rozdílu u pojmů tečná nadrovina k vrstevnici funkce v bodě a tečná nadrovina ke grafu funkce v bodě. Tečná nadrovina k vrstevnici funkce f v bodě a je lineární varieta v $\mathbb{R}^n$ dimenze $n-1$, pokud $\nabla f(a) \neq 0$, kdežto tečná nadrovina ke grafu funkce f v bodě $(a, f(a))$ je lineární varieta v $\mathbb{R}^{n+1}$ dimenze n.

Příklad 3.22: Uvažujme funkci

$$f(x_1, \ldots, x_n) = \|x\| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Pro $a \in \mathbb{R}^n \setminus \{0\}$ je

$$\frac{\partial f}{\partial x_i}(a) = \frac{a_i}{\sqrt{a_1^2 + \cdots + a_n^2}} = \frac{a_i}{\|a\|}, \quad \forall i \in \hat{n}.$$

Odtud máme

$$\nabla f(a)(x - a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) = \frac{a^T x}{\|a\|} - \|a\|.$$

Tudíž tečná nadrovina k vrstevnici $S_{f(a)}$ v a je dána rovnicí

$$a^T x = \|a\|^2$$

a tečná nadrovina ke grafu f v $(a, f(a))$ je dána rovnicí

$$y = \frac{a^T x}{\|a\|}.$$

V bodě $a = 0$ není f diferencovatelná, neboť neexistují ani parciální derivace f v 0, což vyplývá z toho, že neexistuje limita funkce

$$\frac{f(0 + te_i) - f(0)}{t} = \frac{|t|}{t}$$

pro $t \to 0$.

3.3 Věta o přírůstku

Už jsme několikrát použili Lagrangeovu větu o přírůstku pro funkci jedné proměnné. Nyní si větu zobecníme na funkce více proměnných.

Věta 3.23 (O přírůstku)**:** Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ je diferencovatelná funkce na otevřené a *konvexní* množině Ω. Jsou-li $a, b \in \Omega$, potom existuje c na úsečce spojující body a a b tak, že

$$f(b) - f(a) = \nabla f(c)(b - a).$$

Důkaz. Označme $\phi : [0, 1] \to \mathbb{R}^n$ křivku, jejíž stopou je úsečka spojující a a b, tj.

$$\phi(t) = a + t(b - a), \quad t \in [0, 1].$$

Potom funkce jedné proměnné $g := f \circ \phi$ je spojitá na $[0, 1]$ a diferencovatelná na $(0, 1)$ podle Věty 3.7. Můžeme proto na g aplikovat Lagrangeovu větu o přírůstku, podle které existuje $t_0 \in (0, 1)$ tak, že

$$g(1) - g(0) = g'(t_0).$$

Tuto rovnost lze přepsat s použitím řetězového pravidla do tvaru

$$f(b) - f(a) = \nabla f(\phi(t_0))(b - a).$$

Nyní stačí označit $c := \phi(t_0)$ a dostáváme tvrzení věty. $\square$

Důsledek 3.24: Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ je diferencovatelná na Ω. Nechť dále $K \subset \Omega$ je konvexní množina a $(\exists M > 0)(\forall x \in K)(\|\nabla f(x)\| \leq M)$. Potom

$$|f(x) - f(y)| \leq M\|x - y\|$$

pro všechna $x, y \in K$.

Důkaz. Stačí použít Větu 3.23 a Cauchyho–Schwarzovu nerovnost. $\square$

Na tomto místě je dobré upozornit, že Větu o přírůstku nelze zobecnit na vektorové funkce $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$, kde $m \geq 2$. To znamená, že pro vektorovou funkci f diferencovatelnou na konvexní množině Ω **neplatí**

$$f(b) - f(a) = Df(c)(b - a)$$

pro nějaká $a, b, c \in \Omega$, jak by člověk mohl čekat v analogii s Lagrangeovou větou o přírůstku funkce jedné proměnné. Tento fakt ilustruje následující příklad.

Příklad 3.25: Uvažujme funkci $f : \mathbb{R} \to \mathbb{R}^2$ definovanou předpisem

$$f(t) = \begin{pmatrix} t^2 \\ t^3 \end{pmatrix}.$$

Zkoumejme, zda pro nějaké $c \in \mathbb{R}$ platí

$$f(1) - f(0) = Df(c)(1 - 0),$$

což je ekvivalentní rovnosti

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2c \\ 3c^2 \end{pmatrix}.$$

Žádné $c \in \mathbb{R}$ ovšem nevyhovuje současně rovnicím $2c = 1$ a $3c^2 = 1$.

Vektorová analogie Důsledku 3.24 už ale platí.

Věta 3.26: Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ je diferencovatelná na Ω. Nechť dále $K \subset \Omega$ je konvexní množina a

$$(\exists M > 0)(\forall x \in K)(\|Df(x)\| \leq M).$$

Potom

$$\|f(x) - f(y)\| \leq M\|x - y\|$$

pro všechna $x, y \in K$.

Důkaz. Buďte $x, y \in K$ a označme $h := y - x$. Dále pro $j \in \hat{m}$ definujme funkci $g_j : [0,1] \to \mathbb{R}$ vztahem

$$g_j(t) := f_j(x + th).$$

Všimněte si, že $x + th$ pro $t \in [0,1]$ jsou body úsečky spojující x a y, a proto $x + th \in K$ pro všechna $t \in [0,1]$ z konvexnosti K. Dále z diferencovatelnosti f plyne, že jsou funkce g_j spojité na $[0,1]$ a diferencovatelné na $(0,1)$ pro všechna $j \in \hat{m}$. Potom s využitím Základní věty integrálního počtu a řetězového pravidla dostaneme

$$f_j(x+h) - f_j(x) = g_j(1) - g_j(0) = \int_0^1 g_j'(t)\mathrm{d}t = \int_0^1 \nabla f_j(x+th)h\mathrm{d}t, \quad \forall j \in \hat{m},$$

což můžeme zapsat vektorově ve tvaru

$$f(x+h) - f(x) = \int_0^1 Df(x+th)h\mathrm{d}t, \tag{20}$$

kde Df chápeme jako Jacobiho matici zobrazení f a integrál s vektorovým integrandem aplikujeme po jednotlivých složkách.

Lemma 3.27: Je-li $g : [a,b] \to \mathbb{R}^m$ spojité na intervalu $[a,b]$, potom platí

$$\left\| \int_a^b g(t)\mathrm{d}t \right\| \leq \int_a^b \|g(t)\|\mathrm{d}t.$$

Důkaz Lemma 3.27. Označme

$$u := \int_a^b g(t)\mathrm{d}t.$$

Potom z Cauchyho–Schwarzovy nerovnosti plyne, že

$$\|u\|^2 = u^T u = u^T \int_a^b g(t)\mathrm{d}t = \int_a^b u^T g(t)\mathrm{d}t \leq \int_a^b \|u\|\|g(t)\|\mathrm{d}t = \|u\| \int_a^b \|g(t)\|\mathrm{d}t.$$

Nyní stačí na obou stranách zkrátit $\|u\|$, pokud $u \neq 0$, a dostáváme kýženou nerovnost. Je-li $u = 0$, platí lemma triviálně. □

Aplikujeme-li nyní Lemma 3.27 v rovnici (20) a použijeme-li také nerovnost z 1. trzení Lemma 3.4, dostaneme

$$\|f(x+h)-f(x)\| \leq \int_0^1 \|Df(x+th)h\| \mathrm{d}t \leq \int_0^1 \|Df(x+th)\|\|h\| \mathrm{d}t \leq M\|h\|,$$

neboli

$$\|f(y)-f(x)\| \leq M\|y-x\|.$$

□

Dalším důsledkem Věty o přírůstku je, že za jistých předpokladů funkce s nulovou derivací jsou funkce konstantní.

Důsledek 3.28: Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ je diferencovatelná funkce na otevřené a konvexní množině Ω. Je-li $\nabla f = 0$ na Ω, potom je f konstantní na Ω.

Důkaz. Zvolme libovolně $a, b \in \Omega$. Z Věty 3.23 a nulovosti gradientu f na Ω plyne, že $f(a) - f(b) = 0$. Tzn., že $f(a) = f(b)$ pro všechna $a, b \in \Omega$, neboli f je konstantní. □

Poslední důsledek lze zobecnit. Jednak funkce f nemusí být pouze skalární. Za druhé množina Ω nemusí být konvexní, stačí aby byla souvislá. Předpoklad souvislosti Ω je už docela přirozený, jak ukazuje následující jednoduchý příklad.

Příklad 3.29: Buď $\Omega = (-\infty, 0) \cup (0, \infty)$ a $f : \Omega \to \mathbb{R}$ definovaná vztahem

$$f(x) = \begin{cases} 1, & \text{je-li } x > 0, \\ -1, & \text{je-li } x < 0. \end{cases}$$

Potom $(\forall x \in \Omega)(f'(x) = 0)$, ale f není konstantní na Ω.

Věta 3.30: Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ je diferencovatelná na *oblasti* Ω. Pokud

$$(\forall x \in \Omega)(Df(x) = 0),$$

potom je f konstantní na Ω.

Důkaz. Nejprve větu dokážeme pro případ skalární funkce $m = 1$. Ukážeme, že množina

$$S_c := \{x \in \Omega \mid f(x) = c\},$$

je obojetná pro každé $c \in \mathbb{R}$. Potom zvolíme-li libovolně $a \in \Omega$, bude $S_{f(a)} \neq \emptyset$, neboť $a \in S_{f(a)}$. Ze souvislosti Ω vyvodíme, že $S_{f(a)} = \Omega$, neboli $f(x) = f(a)$ pro všechna $x \in \Omega$, což znamená, že f je na Ω konstantní.

Zbývá tedy ověřit obojetnost množin S_c. Jelikož $S_c = f^{-1}(\{c\})$ a f je spojitá na Ω, viz Věta 3.5, je podle Věty 2.61 množina S_c uzavřená (tj. v $c\tau_\Omega$). Ukážeme, že je S_c také otevřená. To platí triviálně, pokud $S_c = \emptyset$. Předpokládejme proto, že $S_c \neq \emptyset$ a zvolme libovolně $x \in S_c$. Z otevřenosti Ω plyne existence $r > 0$ tak, že $B_x(r) \subset \Omega$. Podle předpokladu je $\nabla f = 0$ na

$B_x(r)$, a tudíž podle Důsledku 3.28 je $f = c$ na $B_x(r)$, neboť koule je konvexní množina, viz Cvičení 2.5. Tedy $B_x(r) \subset S_c$ a S_c je proto otevřená.

V případě vektorové funkce f, kdy $m \in \mathbb{N}$, si stačí uvědomit, že z předpokladu $Df(x) = 0$ plyne $\nabla f_j(x) = 0$, $\forall j \in \hat{m}$, neboť tyto gradienty jsou řádky matice $Df(x)$. Z předchozí části potom plyne, že komponenty f_j jsou konstantní na Ω pro všechna $j \in \hat{m}$, a tudíž i vektorová funkce $f = (f_1, \ldots, f_m)$ je konstantní na Ω. $\square$

3.4 Derivace vyšších řádů

Nejprve se zaměříme na derivace parciální. Předpokládejme, že funkce $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ má parciální derivaci $\partial_{x_i} f$ na okolí bodu $a \in \Omega$ pro nějaké $i \in \hat{n}$. Potom můžeme opět aplikovat definici parciální derivace např. podle j-té proměnné tentokrát na funkci $\partial_{x_i} f$ a dostaneme *parciální derivaci f vyššího (druhého) řádu*:

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(a) \equiv \partial^2_{x_j, x_i} f(a) := \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(a) = \lim_{t \to 0} \frac{\partial_{x_i} f(a + te_j) - \partial_{x_i} f(a)}{t}$$

za předpokladu, že limita vpravo existuje. Je-li $i = j$, používáme stručnější značení

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

Opakováním téhož postupu definujeme parciální derivace třetího a vyšších řádů, např.

$$\frac{\partial^3 f}{\partial x_k \partial x_j \partial x_i}(a) := \frac{\partial}{\partial x_k}\left(\frac{\partial^2 f}{\partial x_j \partial x_i}\right)(a)$$

nebo

$$\frac{\partial^3 f}{\partial x_k \partial x_i^2}(a) := \frac{\partial}{\partial x_k}\left(\frac{\partial^2 f}{\partial x_i^2}\right)(a)$$

a analogicky pro další případy. Všimněte si, že použité značení ukazuje v jakém pořadí funkci parciálně derivujeme. Obecně totiž na tomto pořadí záleží.

Příklad 3.31: Spočítáme *smíšené* druhé parciální derivace

$$\frac{\partial^2 f}{\partial x \partial y} \quad \text{a} \quad \frac{\partial^2 f}{\partial y \partial x}$$

v bodě $(0, 0)$ funkce

$$f(x, y) = \begin{cases} \frac{x^3 y}{x^2 + y^2}, & \text{je-li } (x, y) \neq (0, 0), \\ 0, & \text{je-li } (x, y) = (0, 0). \end{cases}$$

Nejprve spočítáme parciální derivace prvního řádu:

$$\frac{\partial f}{\partial x}(x, y) = \begin{cases} \frac{3x^2 y(x^2 + y^2) - 2x^4 y}{(x^2 + y^2)^2} = \frac{x^2 y(x^2 + 3y^2)}{(x^2 + y^2)^2}, & \text{je-li } (x, y) \neq (0, 0), \\ \lim_{t \to 0} \frac{f(t, 0) - f(0, 0)}{t} = 0, & \text{je-li } (x, y) = (0, 0), \end{cases}$$

$$\frac{\partial f}{\partial y}(x,y) = \begin{cases} \frac{x^3(x^2+y^2)-2x^3y^2}{(x^2+y^2)^2} = \frac{x^3(x^2-y^2)}{(x^2+y^2)^2}, & \text{je-li } (x,y) \neq (0,0), \\ \lim_{t\to 0} \frac{f(0,t)-f(0,0)}{t} = 0, & \text{je-li } (x,y) = (0,0). \end{cases}$$

Odtud dále dostaneme

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = \lim_{t\to 0} \frac{1}{t}\left(\frac{\partial f}{\partial y}(t,0) - \frac{\partial f}{\partial y}(0,0)\right) = 1$$

a

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) = \lim_{t\to 0} \frac{1}{t}\left(\frac{\partial f}{\partial x}(0,t) - \frac{\partial f}{\partial x}(0,0)\right) = 0.$$

Pozorujeme tedy, že

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

Pro $(x,y) \neq (0,0)$ ověříme přímým výpočtem, že

$$\frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y) = \frac{x^6 + 6x^4y^2 - 3x^2y^4}{(x^2+y^2)^3}.$$

Všimněte si, že tato funkce není spojitá v bodě $(0,0)$.

Pokud jsou ale smíšené derivace spojité, na pořadí parciálního derivování už nezáleží. Toto tvrzení lze vyslovit i za mírně slabších předpokladů.

Věta 3.32: Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$, $a \in \Omega$ a $i, j \in \hat{n}$. Pokud $\partial_{x_i} f$ a $\partial^2_{x_i,x_j} f$ existují na okolí a a $\partial^2_{x_i,x_j} f$ je spojitá v a, potom existuje také $\partial^2_{x_j,x_i} f(a)$ a platí

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

Důkaz. Jelikož se ve větě objevují jen proměnné x_i a x_j, stačí tvrzení dokázat pro funkci dvou proměnných. Předpokládejme tedy, že $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ taková, že existují $\partial_y f$ a $\partial^2_{y,x} f$ na okolí bodu $(a,b) \in \Omega$ a $\partial^2_{y,x} f$ je spojitá v bodě (a,b).

Pro $s \in \mathbb{R}$ dostatečně malé, máme

$$\frac{\partial f}{\partial y}(a+s,b) - \frac{\partial f}{\partial y}(a,b) = \lim_{t\to 0} \frac{g_t(s) - g_t(0)}{t},$$

kde jsme označili

$$g_t(\xi) := f(a+\xi, b+t) - f(a+\xi, b).$$

Podle Lagrangeovy věty o přírůstku je

$$\frac{g_t(s) - g_t(0)}{s} = g_t'(\tilde{s}) = \frac{\partial f}{\partial x}(a+\tilde{s}, b+t) - \frac{\partial f}{\partial x}(a+\tilde{s}, b),$$

kde $\tilde{s}$ leží mezi 0 a s. Odtud pak dostaneme

$$\frac{1}{s}\left(\frac{\partial f}{\partial y}(a+s,b) - \frac{\partial f}{\partial y}(a,b)\right) = \lim_{t\to 0} \frac{1}{t}\left(\frac{\partial f}{\partial x}(a+\tilde{s}, b+t) - \frac{\partial f}{\partial x}(a+\tilde{s}, b)\right) = \frac{\partial^2 f}{\partial y \partial x}(a+\tilde{s}, b).$$

Pošleme-li v poslední rovnosti $s \to 0$, potom také $\tilde{s} \to 0$ a vzhledem ke spojitosti funkce $\partial^2_{y,x} f$ v bodě (a,b), dostaneme

$$\lim_{s \to 0} \frac{1}{s} \left(\frac{\partial f}{\partial y}(a+s,b) - \frac{\partial f}{\partial y}(a,b) \right) = \frac{\partial^2 f}{\partial y \partial x}(a,b),$$

z čehož plyne tvrzení věty. □

Definice 3.33 (Třídy C^k, hladká funkce)**:** Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$, kde Ω je otevřená, a $k \in \mathbb{N}$. Pokud všechny parciální derivace f řádu k, tj. funkce

$$\frac{\partial^k f}{\partial x_1^{j_1} \dots \partial x_n^{j_n}},$$

kde $j_1, \dots, j_n \in \mathbb{N}_0$, $j_1 + \dots + j_n = k$, existují a jsou spojité na Ω, nazýváme f funkcí *třídy* C^k *na* Ω a píšeme $f \in C^k(\Omega)$. Pokud $f \in C^k(\Omega)$ pro každé $k \in \mathbb{N}$, nazýváme f *hladká* funkce na Ω a píšeme $f \in C^\infty(\Omega)$. Množinu spojitých funkcí na Ω značíme $C^0(\Omega) \equiv C(\Omega)$.

Je-li $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ vektorová funkce $f = (f_1, \dots f_m)^T$ a $k \in \mathbb{N}_0 \cup \{\infty\}$, řekneme, že f je *třídy* C^k *na* Ω a opět píšeme $f \in C^k(\Omega)$, právě když $f_j \in C^k(\Omega)$ pro všechna $j \in \hat{m}$.

Z Věty 3.16 plyne, že funkce $f \in C^1(\Omega)$ je diferencovatelná na Ω a Věta 3.5 implikuje inkluzi

$$C^1(\Omega) \subset C^0(\Omega).$$

Obecněji máme řetězec inkluzí

$$C^\infty(\Omega) \subset \dots \subset C^k(\Omega) \subset C^{k-1}(\Omega) \subset \dots \subset C^1(\Omega) \subset C^0(\Omega).$$

Dále z Věty 3.32 vyplývá, že při výpočtu libovolné parciální derivace k-tého řádu funkce $f \in C^k(\Omega)$ nezáleží na pořadí jednotlivých parciálních derivací.

Zamysleme se dále nad tím, jak definovat (totální) derivace vyšších řádů. Předpokládejme, že je dána funkce $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$, která je diferencovatelná na okolí H_a bodu $a \in \Omega$. Její derivace Df je tedy zobrazení z $\mathbb{R}^n$ s hodnotami v $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, přesněji $Df : H_a \subset \mathbb{R}^n \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Použijeme-li zobecnění Definice 3.1 na zobrazení mezi obecnými normovanými prostory, dospějeme k následující definici: Řekneme, že funkce f je *dvakrát diferencovatelná v bodě* a, právě když existuje lineární zobrazení $T \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$ takové, že

$$\lim_{x \to a} \frac{\|Df(x) - Df(a) - T(x-a)\|}{\|x-a\|} = 0.$$

Druhu derivaci f v bodě a značíme $D^2 f(a) := T$.

Analogicky bychom zavedli i derivace vyšších řádů a dostali tak zobrazení:

$$\begin{aligned}
&D^3 f(a) \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))), \\
&D^4 f(a) \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)))), \\
&D^5 f(a) \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))))), \text{ atd.}
\end{aligned}$$

Prostor $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ je izomorfní s prostorem matic $\mathbb{R}^{m,n}$, a proto bylo možné první derivaci $Df(a)$ reprezentovat (Jacobiho) maticí a používat maticový počet pro práci s první derivací. U vyšších derivací roli matic přebírají tzv. *tenzory*, jež lze chápat jako vícerozměrné analogie matic. Např. druhá derivace $D^2 f(a)$ je element prostoru $\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$, který je izomorfní prostoru $\mathbb{R}^{m,n,n}$ objektů (tenzorů) indexovanými třemi indexy. Existuje partie lineární algebry věnovaná tenzorovému počtu, která se ale v základním kurzu nepřednáší.

My se zde práci s tenzory vyhneme, protože v dalším výkladu vystačíme s druhou derivací skalární funkce $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$. V tomto případě ztotožňujeme první derivaci Df diferencovatelné funkce f s gradientem ∇f a můžeme proto Df reprezentovat jako zobrazení

$$Df : H_a \subset \mathbb{R}^n \to \mathbb{R}^n : x \mapsto \begin{pmatrix} \partial_{x_1} f(x) \\ \vdots \\ \partial_{x_n} f(x) \end{pmatrix}.$$

Zde výjimečně interpretujeme gradient jako sloupcový vektor, abychom se drželi konvence, že elementy $\mathbb{R}^n$ jsou sloupcové vektory. Toto zobrazení Df můžeme znovu derivovat podle Definice 3.1 a dostáváme se tak k následující definici.

Definice 3.34 (Druhá derivace skalární funkce)**:** Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ je diferencovatelná funkce na okolí $a \in \Omega$. Řekneme, že f je *dvakrát diferencovatelná v bodě* a, právě když existuje $B \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ tak, že

$$\lim_{x \to a} \frac{\|Df(x) - Df(a) - B(x - a)\|}{\|x - a\|} = 0.$$

Zobrazení B nazýváme *druhá derivace* f *v bodě* a a značíme $D^2 f(a)$.

Podobně jako tomu bylo u první derivace můžeme elementy matice zobrazení $Df(a)$ vyjádřit pomocí parciálních derivací f tentokrát druhého řádu.

Věta 3.35: Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ je dvakrát diferencovatelná funkce v bodě $a \in \Omega$. Potom parciální derivace druhého řádu $\partial^2_{x_i, x_j} f(a)$ existují pro všechna $i, j \in \hat{n}$ a platí

$$\mathcal{E}_n\left(D^2 f(a)\right) = \begin{pmatrix} \partial^2_{x_1,x_1} f(a) & \partial^2_{x_2,x_1} f(a) & \dots & \partial^2_{x_n,x_1} f(a) \\ \partial^2_{x_1,x_2} f(a) & \partial^2_{x_2,x_2} f(a) & \dots & \partial^2_{x_n,x_2} f(a) \\ \vdots & \vdots & \ddots & \vdots \\ \partial^2_{x_1,x_n} f(a) & \partial^2_{x_2,x_n} f(a) & \dots & \partial^2_{x_n,x_n} f(a) \end{pmatrix}.$$

Důkaz. Vyplývá přímo z Věty 3.13. □

Definice 3.36 (Hessova matice)**:** Má-li funkce $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ všechny parciální derivace druhého řádu v bodě $a \in \Omega$, nazýváme matici $\nabla^2 f(a) \in \mathbb{R}^{n,n}$ s elementy

$$\left(\nabla^2 f(a)\right)_{i,j} = \frac{\partial^2 f}{\partial x_j \partial x_i}(a), \quad i, j \in \hat{n},$$

Hessova matice funkce f *v bodě* a.

Poznámka (Hessián)**:** Matice $\nabla^2 f(a)$ se také často v literatuře nazývá *Hessián* f v bodě a, ačkoliv by logika tohoto názvosloví spíše naznačovala, že Hessián je determinant Hessovy matice podobně jako tomu je u Jacobiho matice zobrazení a Jacobiánu.

Podle Věty 3.35 je Hessova matice maticí druhé derivace (vzhledem ke standardním bázím) dvakrát diferencovatelného zobrazení. Jak jsme viděli v Příkladu 3.31, Hessova matice nemusí být symetrická. Avšak z Věty 3.32 dostáváme speciálně následující důsledek.

Důsledek 3.37: Je-li $f \in C^2(\Omega)$, potom je matice $\nabla^2 f(x)$ symetrická pro všechna $x \in \Omega$.

3.5 Taylorova věta

Zobecnit Taylorovu větu na funkce více proměnných lze jednoduše na základě Taylorovy věty pro funkce jedné proměnné. Pro pochopení postupu si nejprve spočítáme první dva členy Taylorova rozvoje a zbytek (v Lagrangeově tvaru).

Uvažujme funkci $f \in C^3(\Omega)$, kde $\Omega \subset \mathbb{R}^n$ je konvexní otevřená množina (např. koule). Zvolme body $a, x \in \Omega$ a nechť $\phi : [0,1] \to \Omega$ je úsečka spojující a a x, tj. $\phi(t) = a + t(x-a)$. Podle Taylorovy věty (s Langrangeovým tvarem zbytku) aplikované na funkci jedné proměnné $g(t) := f(\phi(t)) = f(a + t(x-a))$ dostaneme rozvoj

$$g(1) = g(0) + g'(0) + \frac{1}{2!}g''(0) + \frac{1}{3!}g'''(\xi), \tag{21}$$

kde $\xi \in (0,1)$. Zřejmě $g(1) = f(x)$ a $g(0) = f(a)$. Pomocí řetězového pravidla dále spočítáme, že

$$g'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a + t(x-a))(x_i - a_i),$$

$$g''(t) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(a + t(x-a))(x_j - a_j)(x_i - a_i)$$

a

$$g'''(t) = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \frac{\partial^3 f}{\partial x_k \partial x_j \partial x_i}(a + t(x-a))(x_k - a_k)(x_j - a_j)(x_i - a_i).$$

Dosazením spočítaných výrazů do (21) dostáváme výsledek

$$f(x) = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) + \frac{1}{2!} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(a)(x_j - a_j)(x_i - a_i) + R_3(x; a),$$

kde

$$R_3(x; a) := \frac{1}{3!} \sum_{i,j,k=1}^{n} \frac{\partial^3 f}{\partial x_k \partial x_j \partial x_i}(a + \xi(x-a))(x_k - a_k)(x_j - a_j)(x_i - a_i).$$

Všimněte si, že bod $a + \xi(x-a)$ je bod uvnitř úsečky spojující body a a x. Dále také platí, že

$$\lim_{x \to a} \frac{R_3(x; a)}{\|x - a\|^2} = 0,$$

neboť

$$|R_3(x;a)| \leq \frac{1}{3!}\sum_{i,j,k=1}^{n}\left|\frac{\partial^3 f}{\partial x_k \partial x_j \partial x_i}(a+\xi(x-a))\right||x_k-a_k||x_j-a_j||x_i-a_i| \leq \frac{Cn^3}{3!}\|x-a\|^3$$

kde

$$C := \max_{i,j,k\in\hat{n}}\max_{t\in[0,1]}\left|\frac{\partial^3 f}{\partial x_k \partial x_j \partial x_i}(a+t(x-a))\right|$$

je konečné, protože funkce $\partial^3_{x_k,x_j,x_i} f$ jsou podle předpokladu spojité na kompaktní množině $[\phi]$.

Analogickým způsobem dokážeme následující obecný tvar Taylorovy věty.

Věta 3.38 (Taylor)**:** Buď $f \in C^{s+1}(\Omega)$, kde $s \in \mathbb{N}_0$ a $\Omega \subset \mathbb{R}^n$ konvexní otevřená množina. Potom pro $a, x \in \Omega$ platí:

$$f(x) = f(a) + \sum_{k=1}^{s}\frac{1}{k!}\sum_{i_1,\ldots,i_k=1}^{n}\frac{\partial^k f}{\partial x_{i_1}\ldots\partial x_{i_k}}(a)(x_{i_1}-a_{i_1})\ldots(x_{i_k}-a_{i_k}) + R_{s+1}(x;a),$$

kde

$$R_{s+1}(x;a) := \frac{1}{(s+1)!}\sum_{i_1,\ldots,i_{s+1}=1}^{n}\frac{\partial^{s+1} f}{\partial x_{i_1}\ldots\partial x_{i_{s+1}}}(c)(x_{i_1}-a_{i_1})\ldots(x_{i_{s+1}}-a_{i_{s+1}})$$

a bod c leží uvnitř úsečky spojující body a a x. Navíc

$$\lim_{x\to a}\frac{R_{s+1}(x;a)}{\|x-a\|^s} = 0.$$

Definice 3.39 (Taylorův polynom, zbytek)**:** V Taylorově Větě 3.38 se objevující polynom více proměnných

$$P_s(x;a) := f(a) + \sum_{k=1}^{s}\frac{1}{k!}\sum_{i_1,\ldots,i_k=1}^{n}\frac{\partial^k f}{\partial x_{i_1}\ldots\partial x_{i_k}}(a)(x_{i_1}-a_{i_1})\ldots(x_{i_k}-a_{i_k})$$

se nazývá *s-tý Taylorův polynom funkce f se středem v bodě a*. Dále $R_{s+1}(x;a)$ nazýváme *zbytkem po s-tém Taylorově polynomu funkce f se středem v a.*

Pro $f \in C^\infty(\Omega)$ je dle Taylorovy věty

$$f(x) = P_s(x;a) + R_{s+1}(x;a)$$

pro každé $s \in \mathbb{N}_0$. Nicméně z hladkosti f ještě **neplyne**, že

$$\lim_{s\to\infty} R_{s+1}(x;a) = 0,$$

jak jsme viděli už v Příkladu 1.50. Podobně jako ve Větě 1.52 zajistí limitu $R_{s+1}(x;a) \to 0$ pro $s \to \infty$ určitá kontrola všech parciálních derivací f.

Důsledek 3.40: Nechť $f \in C^\infty(\Omega)$ a $a \in \Omega$. Pokud

$$(\exists C > 0)(\exists s_0 \in \mathbb{N})(\forall s \geq s_0)(\forall i_1, \dots i_s \in \hat{n}) \left(\max_{x \in \Omega} \left| \frac{\partial^s f}{\partial x_{i_1} \dots \partial x_{i_s}}(x) \right| \leq C^s s! \right),$$

potom existuje $r > 0$ tak, že

$$\lim_{s \to \infty} R_s(x; a) = 0$$

pro všechna $x \in B_a(r)$.

Důkaz. Předpoklady věty umožňují odhadnout zbytek R_s následovně:

$$|R_s(x; a)| \leq \frac{1}{s!} \sum_{i_1, \dots i_s = 1}^{n} \max_{x \in \Omega} \left| \frac{\partial^s f}{\partial x_{i_1} \dots \partial x_{i_s}}(x) \right| |x_{i_1} - a_{i_1}| \dots |x_{i_s} - a_{i_s}| \leq (nC\|x - a\|)^s.$$

Tudíž, zvolíme-li $r < 1/(nC)$ také dostatečně malé, aby $B_a(r) \subset \Omega$, bude $\|x - a\| < 1/(nC)$ pro každé $x \in B_a(r)$, a proto pravá strana horního odhadu zbytku $R_s(x; a)$ jde k nule pro $s \to \infty$. □

Lineární a kvadratický člen v Taylorově polynomu lze vyjádřit elegantně pomocí gradientu a Hessovy matice f, neboť

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) = \nabla f(a)(x - a)$$

a

$$\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)(x_j - a_j)(x_i - a_i) = (x - a)^T \nabla^2 f(a)(x - a).$$

Taylorova věta pro funkci $f \in C^3(\Omega)$ nám tedy dává vzorec

$$f(x) = f(a) + \nabla f(a)(x - a) + \frac{1}{2}(x - a)^T \nabla^2 f(a)(x - a) + R_3(x; a).$$

Všimněte si, že první dva členy rozvoje určují tečnou nadrovinu ke grafu funkce f v bodě a:

$$y = f(a) + \nabla f(a)(x - a),$$

což představuje nejlepší lineární aproximaci hodnot funkce f v okolí bodu a. Přidáme-li i kvadratický člen, je

$$y = f(a) + \nabla f(a)(x - a) + \frac{1}{2}(x - a)^T \nabla^2 f(a)(x - a)$$

nejlepší kvadratickou aproximací hodnot funkce f v okolí bodu a a odpovídající plochu bychom mohli nazývat tečnou kvadratickou plochou ke grafu funkce f v bodě a (pokud $\nabla^2 f(a) \neq 0$).

Příklad 3.41: Uvažujme funkci $f : \mathbb{R}^2 \to \mathbb{R}$ definovanou vztahem

$$f(x,y) = e^{x+y}\sin(xy).$$

Snadným výpočet parciálních derivací f v bodě $(0,0)$ dostaneme

$$\nabla f(0,0) = \left(\frac{\partial f}{\partial x}(0,0), \frac{\partial f}{\partial y}(0,0)\right) = (0,0)$$

a

$$\nabla^2 f(0,0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(0,0) & \frac{\partial^2 f}{\partial x \partial y}(0,0) \\ \frac{\partial^2 f}{\partial x \partial y}(0,0) & \frac{\partial^2 f}{\partial y^2}(0,0) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

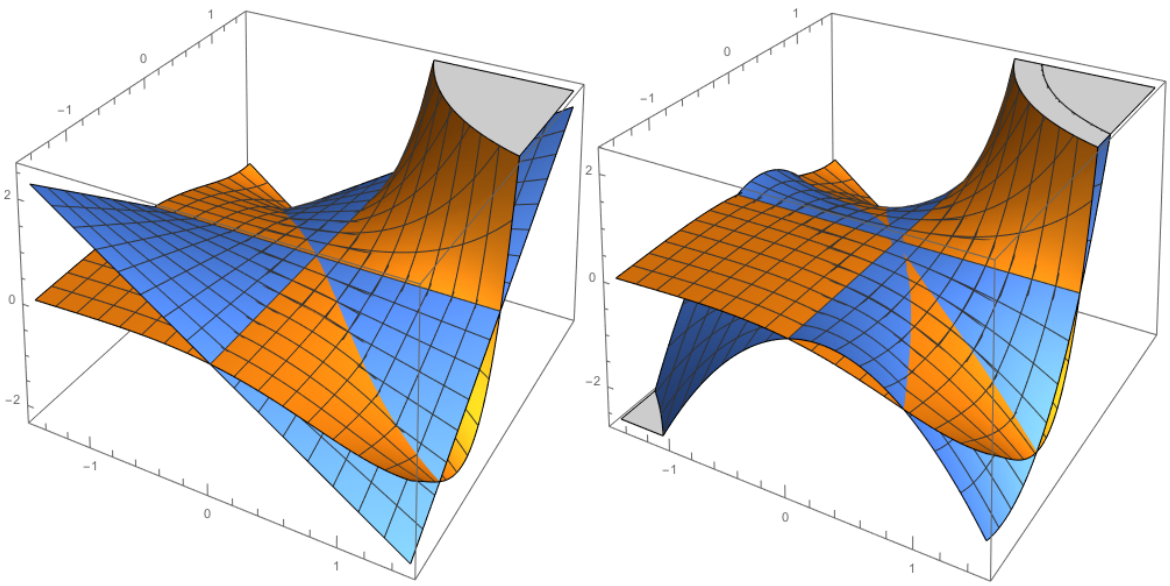
A proto je druhý Taylorův polynom f v bodě $(0,0)$ funkce

$$P_2(x,y;0,0) = f(0,0) + \nabla f(0,0)\begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2}(x,y)\nabla^2 f(0,0)\begin{pmatrix} x \\ y \end{pmatrix} = xy.$$

Třetí Taylorův polynom také snadno spočítáme a vyjde nám

$$P_3(x,y;0,0) = xy + x^2y + xy^2.$$

Jak aproximují funkce P_2 a P_3 funkční hodnoty f na okolí počátku ilustruje Obrázek 8.



Obrázek 8: Graf funkce f (oranžová) a Taylorových polynomů P_2 (modrá vlevo) a P_3 (modrá vpravo) z Příkladu 3.41.

3.6 Lokální extrémy funkcí více proměnných

Známý postup hledání lokálních extrémů funkce jedné proměnné využívající diferenciálního počtu nyní zobecníme na funkce více proměnných.

Definice 3.42 ((Ostré) lokální minimum/maximum, lokální extrém)**:** Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ a $a \in \Omega$. Řekneme, že funkce f má v bodě a:

1. *lokální maximum* $\Leftrightarrow$ $(\exists H_a \subset \Omega)(\forall x \in H_a)(f(x) \leq f(a))$,
2. *ostré lokální maximum* $\Leftrightarrow$ $(\exists H_a \subset \Omega)(\forall x \in H_a \setminus \{a\})(f(x) < f(a))$,
3. *lokální minimum* $\Leftrightarrow$ $(\exists H_a \subset \Omega)(\forall x \in H_a)(f(x) \geq f(a))$,
4. *ostré lokální minimum* $\Leftrightarrow$ $(\exists H_a \subset \Omega)(\forall x \in H_a \setminus \{a\})(f(x) > f(a))$.

Funkce f má v bodě a *lokální extrém*, právě když f má v a lokální minimum nebo lokální maximum.

Věta 3.43 (Nutná podmínka extrému)**:** Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ má v bodě $a \in \Omega$ lokální extrém. Potom existuje-li $\partial_{x_i} f(a)$ pro $i \in \hat{n}$, je $\partial_{x_i} f(a) = 0$. Speciálně, je-li f diferencovatelná v bodě a, potom $\nabla f(a) = 0$.

Důkaz. Předpokládejme, že pro $i \in \hat{n}$ parciální derivace $\partial_{x_i} f(a)$ existuje. Potom funkce

$$g_i(t) := f(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_n)$$

je funkce jedné proměnné, která má v bodě $t = a_i$ extrém a je v tomto bodě také diferencovatelná. Odtud a z nutné podmínky extrému funkce jedné proměnné, dostaneme

$$0 = g_i'(a_i) = \frac{\partial f}{\partial x_i}(a).$$

□

Definice 3.44 (Kritický bod)**:** Buď $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ diferencovatelná v bodě $a \in \Omega$. Pokud $\nabla f(a) = 0$, nazýváme a *kritický bod* (nebo také *stacionární bod*) funkce f.

Podle Věty 3.43 má-li diferencovatelná funkce extrém v bodě a, je to její kritický bod. Samozřejmě naopak to neplatí, tzn., že v kritickém bodě f ještě nemusí mít f extrém (např. $f(x) = x^3$ v $x = 0$). V případě dvakrát diferencovatelné funkce jedné proměnné má f v kritickém bodě a ostré lokální minimum resp. maximum, pokud $f''(a) > 0$ resp. $f''(a) < 0$. Pro funkce více proměnných funguje obdobná postačující podmínka, jen roli znaménka hraje definitnost Hessovy matice f.

Připomeňme si definici definitnosti symetrické matice z lineární algebry.

Definice 3.45 (PD, PSD, ND, NSD, IND)**:** Symetrickou matici $A \in \mathbb{R}^{n,n}$ nazveme:

1. *pozitivně definitní (PD)* $\Leftrightarrow$ $(\forall x \in \mathbb{R}^n \setminus \{0\})(x^T A x > 0)$,
2. *pozitivně semidefinitní (PSD)* $\Leftrightarrow$ $(\forall x \in \mathbb{R}^n)(x^T A x \geq 0)$ a A není PD,

3. *negativně definitní (ND)* $\Leftrightarrow$ $(\forall x \in \mathbb{R}^n \setminus \{0\})(x^T A x < 0)$,

4. *negativně semidefinitní (NSD)* $\Leftrightarrow$ $(\forall x \in \mathbb{R}^n)(x^T A x \leq 0)$ a A není ND,

5. *indefinitní (IND)* $\Leftrightarrow$ $(\exists x, y \in \mathbb{R}^n)(x^T A x > 0 \ \wedge \ y^T A y < 0)$.

Poznámka: V anglické literatuře se definice PSD/NSD většinou uvádí bez dodatku „A není PD/ND", tzn., že PD matice tvoří podmnožinu PSD matic podobně jako čísla kladná jsou také nezáporná. My se zde ovšem přidržíme uvedené definice používané v kurzu lineární algebry na FJFI.

Čtenář by měl být obeznámen z kurzu lineární algebry s tím, jak určit definitnost symetrické matice tzv. metodou „doplňování na čtverce" kvadratické formy $q_A(x) := x^T A x$. Připomeňme si ještě užitečné Sylvestrovo kritérium striktní definitnosti matice.

Věta 3.46 (Sylvestrovo kritérium)**:** Nechť $A \in \mathbb{R}^{n,n}$ a $A = A^T$. Pro $k \in \hat{n}$ označme $\Delta_k := \det A[k]$ (***hlavní minory** A*), kde $A[k] \in \mathbb{R}^{k,k}$ je $k \times k$ podmatice matice A, tj. $(A[k])_{i,j} = A_{i,j}$, $\forall i, j \in \hat{k}$. Potom platí:

1. A je PD $\Leftrightarrow$ $(\forall k \in \hat{n})(\Delta_k > 0)$,

2. A je ND $\Leftrightarrow$ $(\forall k \in \hat{n})((-1)^k \Delta_k > 0)$.

Pro důkaz postačující podmínky pro extrém funkce více proměnných budeme ještě potřebovat následující pomocné tvrzení.

Lemma 3.47: Nechť $A : \Omega \subset \mathbb{R}^n \to \mathbb{R}^{n,n}_{\text{sym}}$ je zobrazení s hodnotami v prostoru symetrických matic $\mathbb{R}^{n,n}_{\text{sym}}$, které je spojité v bodě $a \in \Omega$.

1. Je-li $A(a)$ PD, resp. ND, potom existuje H_a tak, že pro všechna $x \in H_a$ je $A(x)$ PD, resp. ND.

2. Je-li $A(a)$ IND, potom existuje H_a a vektory $u, v \in \mathbb{R}^n$ tak, že pro všechna $x \in H_a$ je $u^T A(x) u > 0$ a $v^T A(x) v < 0$.

Důkaz. 1. Nechť $A(a)$ je PD. Všechny hlavní minory $\Delta_k := \det A[k] : \Omega \to \mathbb{R}$ jsou spojité funkce v bodě a, neboť determinant je spojitou funkcí maticových elementů $A_{i,j}$. A maticové elementy $A_{i,j} : \Omega \to \mathbb{R}$ jsou také spojité funkce v bodě a, jak plyne z předpokladu spojitosti maticové funkce A v bodě a. Tedy hlavní minory Δ_k jsou složením spojitých funkcí, a tudíž spojité v a pro každé $k \in \hat{n}$.

Podle Sylvestrova kritéria je $\Delta_k(a) > 0$ pro všechna $k \in \hat{n}$. Díky spojitosti funkce Δ_k v bodě a najdeme okolí $H_a^{(k)}$ tak, že $\Delta_k(x) > 0$ pro všechna $x \in H_a^{(k)}$. Položíme-li

$$H_a := \bigcap_{k=1}^{n} H_a^{(k)},$$

potom $(\forall x \in H_a)(\forall k \in \hat{n})(\Delta_k(x) > 0)$, a proto jsou matice $A(x)$ PD pro všechna $x \in H_a$ podle Sylvestrova kritéria.

Pro ND variantu tvrzení stačí využít již dokázanou PD variantu a to, že matice A je PD, právě když $-A$ je ND.

2. Nakonec předpokládejme, že $A(a)$ je IND. Potom existují vektory $u, v \in \mathbb{R}^n$ takové, že $u^T A(a)u > 0$ a $v^T A(a)v < 0$. Jelikož jsou funkce $x \mapsto u^T A(x)u$ a $x \mapsto v^T A(x)v$ spojité v bodě a (rozmyslete), musí existovat okolí $H_a^{(u)}$ a $H_a^{(v)}$ bodu a tak, že

$$(\forall x \in H_a^{(u)})(u^T A(x)u > 0) \quad \text{a} \quad (\forall x \in H_a^{(v)})(v^T A(x)v < 0).$$

Tudíž pro důkaz tvrzení stačí položit $H_a := H_a^{(u)} \cap H_a^{(v)}$. □

Věta 3.48 (Postačující podmínka pro extrém)**:** Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$, $f \in C^2(\Omega)$ a bod $a \in \Omega$ je kritický bod f. Potom platí:

1. Je-li $\nabla^2 f(a)$ PD, potom má f v bodě a ostré lokání minimum.
2. Je-li $\nabla^2 f(a)$ ND, potom má f v bodě a ostré lokání maximum.
3. Je-li $\nabla^2 f(a)$ IND, potom f nemá v bodě a lokální extrém f a bod a se nazývá *sedlový bod f*.

Důkaz. Podle předpokladu jsou $\partial_{x_i,x_j} f$ spojité funkce na Ω, je i Hessova matice $\nabla^2 f$ spojité zobrazení na Ω s hodnotami v symetrických maticích, viz Důsledek 3.37. Zde požíváme argument analogický tomu z důkazu Věty 3.13.

1. Předpokládejme nejprve, že $\nabla^2 f(a)$ je PD. Podle Lemma 3.47 existuje $\delta > 0$ tak, že $\nabla^2 f(x)$ je PD pro všechna $x \in B_a(\delta)$. Vezměme $x \in B_a(\delta) \setminus \{a\}$ libovolné ale pevné. Podle Taylorovy věty 3.38 máme

$$f(x) = f(a) + \nabla f(a)(x-a) + \frac{1}{2}(x-a)^T \nabla^2 f(c)(x-a),$$

kde c leží na úsečce spojující a a x, a tudíž je $c \in B_a(\delta)$. Tedy matice $\nabla^2 f(c)$ je PD. Přihlédneme-li ještě k tomu, že podle předpokladu je a kritický bod f, tzn., že $\nabla f(a) = 0$, dostáváme

$$f(x) - f(a) = \frac{1}{2}(x-a)^T \nabla^2 f(c)(x-a) > 0,$$

neboť $x \neq a$ a $\nabla^2 f(c)$ je PD. Odtud plyne, že $f(x) > f(a)$ pro každé $x \in B_a(\delta) \setminus \{a\}$, a tedy f má v bodě a ostré lokální minimum.

2. Je-li $\nabla^2 f(a)$ ND, je postup zcela analogický jako v bodě 1. Jediný rozdíl je, že kvadratická forma $(x-a)^T \nabla^2 f(c)(x-a)$ z pravé strany poslední rovnice je tentokrát záporná, neboť $\nabla^2 f(c)$ je ND. Odtud vyvodíme, že $f(x) < f(a)$ pro každé $x \in B_a(\delta) \setminus \{a\}$, a proto f má v bodě a ostré lokální maximum.

3. Je-li $\nabla^2 f(a)$ IND, potom opět podle Lemma 3.47 existuje $\delta > 0$ a vektory $u, v \in \mathbb{R}^n$ tak, že

$$u^T \nabla^2 f(x)u > 0 \quad \text{a} \quad v^T \nabla^2 f(x)v < 0$$

pro všechna $x \in B_a(\delta)$. Dále z Taylorovy věty a předpokladu $\nabla f(a) = 0$ plyne pro každé $x \in B_a(\delta)$ rovnost

$$f(x) - f(a) = \frac{1}{2}(x-a)^T \nabla^2 f(c_x)(x-a),$$

kde $c_x \in B_a(\delta)$; zde pro jistotu zdůrazňujeme indexem závislost c_x na x. Položme $x := a + tu$, kde $t \in \mathbb{R}$. Je-li $|t| < \epsilon := \delta/\|u\|$ (uvědomte si, že $u \neq 0$), potom $x = a + tu \in B_a(\delta)$, a proto máme

$$f(a+tu) - f(a) = \frac{1}{2}(tu)^T \nabla^2 f(c_t)(tu) = \frac{t^2}{2} u^T \nabla^2 f(c_t) u > 0$$

pro všechna $t \in (-\epsilon, \epsilon) \setminus \{0\}$, neboť $c_t \in B_a(\delta)$. Dokázali jsme tedy, že existuje $\epsilon > 0$ tak, že

$$f(a+tu) > f(a)$$

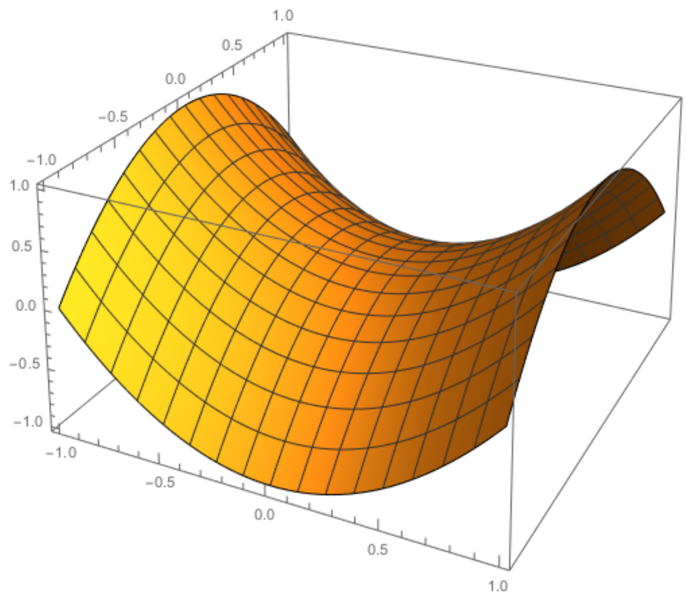
pro všechna $t \in (-\epsilon, \epsilon) \setminus \{0\}$. Analogicky bychom ukázali, že

$$f(a+tv) < f(a)$$

pro všechna $t \in (-\epsilon, \epsilon) \setminus \{0\}$. Tudíž f nemá v bodě a lokální extrém. □

Poznámka: Za silnějšího předpokladu $f \in C^3(\Omega)$, je možné důkaz Věty 3.48 zjednodušit, viz např. [14, Věta 9.23].

Poznámka: Důkaz Věty 3.48 ukazuje, že pokud je $\nabla f(a) = 0$ a matice $\nabla^2 f(a)$ je IND, potom existují dva směry $u, v \in \mathbb{R}^n$ tak, že funkce f zúžená na přímku $t \mapsto a + tu$ má v bodě a ostré lokální minimum, kdežto funkce f zúžená na přímku $t \mapsto a + tv$ má v bodě a ostré lokální maximum. Geometricky tato situace odpovídá představě *sedla* v bodě a, viz Obrázek 9.



Obrázek 9: Sedlový bod (funkce $f(x,y) = x^2 - y^2$ na okolí bodu $(0,0)$).

Všimněte si, že kritérium, které poskytuje Věta 3.48, není vyčerpávající. V případě, že $\nabla f^2(a)$ je PSD nebo NSD, nemůžeme udělat žádný závěr, neboť v kritickém bodě a může i nemusí mít funkce f lokální extrém, viz následující příklady. V tomto případě je matice $\nabla f^2(a)$ singulární a kritický bod a se nazývá *degenerovaný*.

Příklad 3.49: Funkce $f(x, y) = x^4 + y^4$ má v bodě $(0, 0)$ ostré dokonce globální minimum, ale Hessova matice v kritickém bodě $(0, 0)$ funkce f je nulová, tedy PSD (i NSD).

Příklad 3.50: Funkce

$$f(x, y) = x^2 + y^3$$

nemá v bodě $(0, 0)$ lokální extrém, neboť hodnoty $f(0, t) = t^3$ jsou kladné pro $t > 0$ a záporné pro $t < 0$. Nicméně bod $(0, 0)$ je kritický bod f a Hessova matice f v počátku

$$\nabla^2 f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

je PSD.

Platí ale následující opačná implikace.

Věta 3.51: Nechť $f \in C^2(\Omega)$ má v bodě $a \in \Omega$ lokální minimum, resp. lokální maximum, potom

$$(\forall h \in \mathbb{R}^n)(h^T \nabla^2 f(a) h \geq 0), \qquad \text{resp.} \qquad (\forall h \in \mathbb{R}^n)(h^T \nabla^2 f(a) h \leq 0).$$

(Tzn. $\nabla^2 f(a)$ je PSD, nebo PD, resp. NSD, nebo ND.)

Důkaz. Nechť má f v bodě a lokální minimum. Potom $\nabla f(a) = 0$ podle Věty 3.43. Nechť $r > 0$ je takové, že $B_a(r) \subset \Omega$ a $f(a) \leq f(x)$ pro všechna $x \in B_a(r)$. Zvolme pevně $h \in \mathbb{R}^n$, $h \neq 0$. Potom pro všechna $t \in \mathbb{R}$ taková, že $0 < |t| < r/\|h\|$, je $a + th \in B_a(r)$ a podle Taylorovy Věty 3.38 platí

$$0 \leq \frac{f(a + th) - f(a)}{t^2} = \frac{1}{t} \nabla f(a) h + \frac{1}{2} h^T \nabla^2 f(c_t) h = \frac{1}{2} h^T \nabla^2 f(c_t) h. \tag{22}$$

Vektor c_t leží na úsečce mezi a a $a+th$, a proto $c_t \to a$ pro $t \to 0$. Navíc zobrazení $\nabla^2 f$ je spojité v a, jak plyne z předpokladu $f \in C^2(\Omega)$. Proto limitním přechodem $t \to 0$ v nerovnosti (22) dostaneme

$$h^T \nabla^2 f(a) h \geq 0.$$

Tato nerovnost platí pro libovolné $h \in \mathbb{R}^n$, z čehož plyne, že matice $\nabla^2 f(a)$ je PSD, nebo PD.

Má-li f v bodě a lokální maximum, stačí aplikovat právě dokázané tvrzení na funkci $-f$. $\square$

Tedy v případech, kdy je a degenerovaný kritický bod f, tj. $\nabla^2 f(a)$ je PSD nebo NSD, nemáme postačující podmínku pro extrém. I v těchto případech se někdy dá rozhodnout na základě vyšších derivací, pokud existují. Situace je podobná jako v případě funkce jedné proměnné (uvažte např. $f(x) = x^4$), ale značně komplikovanější. Pojem definitnosti je třeba rozšířit z matic na tenzory, což není problém, ale ukázat, že daný tenzor je pozitivně definitní může být kombinatoricky komplikovaná úloha. V případě funkcí dvou proměnných ještě existují poměrně rozumné metody, viz např. [18, Sekce 3.7.2]. My se zde případy degenerovaných kritických bodů zabývat nebudeme.

Metody diferenciálního počtu umožňují najít pouze *lokální* extrémy, ačkoliv v praxi bychom často rádi uměli najít *globální* extrémy nějaké funkce. Někdy lze z lokálních odpovědí udělat odpovědi globální např. za předpokladu konvexnosti funkce f, viz Cvičení 3.11 a 3.12. Pokud nás otázka globálních extrémů zajímá pro spojité funkce omezené na kompaktní množinu, víme že globální extrémy musí existovat z Věty 2.95. Jednou možností, jak je hledat, je vyšetřit funkci na vnitřku kompaktní množiny a zvlášť na její hranici Lagrangeovou metodou pro vázané extrémy (viz Sekce 3.9). Potom stačí srovnat hodnoty funkce v nalezených extrémech. Obecně ale žádný univerzální postup pro důkaz existence a lokalizaci globálních extrémů neexistuje a jednotlivé optimalizační problémy musíme řešit případ od případu.

Jednoduchou ale užitečnou aplikací vyložené metody hledání extrémů, která se objevuje v numerické matematice či statistice, ukazuje následující příklad.

Příklad 3.52 (Metoda nejmenších čtverců)**:** Předpokládejme, že máme k dispozici sadu dat $\{(x_i, y_i)\}_{i=1}^n \subset \mathbb{R}^2$ a chceme nalézt lineární kombinaci daných funkcí $f_1, f_2, \ldots, f_m$, tj. funkci

$$f = \sum_{i=1}^{m} c_i f_i \,,$$

tak, aby funkční hodnoty funkce f v bodech x_i „co nejlépe" odpovídaly hodnotám y_i pro každé $i \in \hat{n}$. Cílem je tedy určit neznámé koeficienty lineární kombinace $c_1, c_2, \ldots, c_m$. Budeme předpokládat, že $m \leq n$, což není na závadu, neboť typicky je množství dat n mnohem větší než počet funkcí m. Dále budeme předpokládat, že soubor funkcí $f_1, \ldots, f_m$ je lineárně nezávislý, což je také přirozený předpoklad, který nepředstavuje zásadní omezení.

Funkce f_i mohou být např. monomy $f_i(x) = x^i$. V takovém případě prokládáme data polynomiální křivkou.

Metoda nejmenších čtverců spočívá v myšlence minimalizovat tzv. střední kvadratickou odchylku celkové chyby, tzn., že hledáme koeficienty $c_1, c_2, \ldots, c_m$ tak, aby hodnota

$$F(c) := \sum_{i=1}^{n} \big(y_i - f(x_i)\big)^2 = \sum_{i=1}^{n} \big(y_i - (Ac)_i\big)^2 = \|y - Ac\|^2$$

byla co nejmenší. Zde jsme označili $c := (c_1, \ldots c_m)^T$ a

$$A := \begin{pmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_m(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_m(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_n) & f_2(x_n) & \cdots & f_m(x_n) \end{pmatrix} .$$

Aplikujme analytický postup pro hledání extrémů funkcí více proměnných na funkci F. Nejprve spočítáme gradient funkce F. Pro $j \in \hat{m}$ máme

$$\frac{\partial F}{\partial c_j}(c) = -2 \sum_{i=1}^{n} \big(y_i - (Ac)_i\big) A_{i,j} = -2(A^T y)_j + 2(A^T Ac)_j \,.$$

A tedy $\nabla F(c) = 0$, právě když
$$A^T y - A^T A c = 0. \tag{23}$$
Z lineární nezávislosti souboru funkcí $f_1, \dots, f_m$ plyne, že matice A má plnou hodnost, tzn. $h(A) = m$, neboť $m \leq n$. V takovém případě je matice $A^T A$ regulární, viz Cvičení 3.14. Rovnice (23) má proto jediné řešení
$$c = (A^T A)^{-1} A^T y.$$

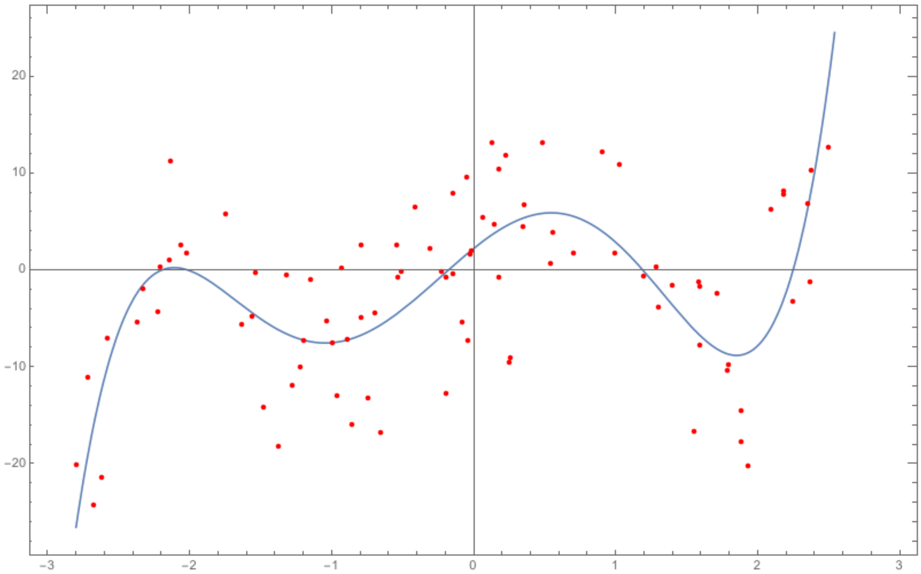
Dále ověříme, že F má v nalezeném vektoru c ostré lokální minimum. Parciální derivace F druhého řádu jsou
$$\frac{\partial^2 F}{\partial c_i \partial c_j}(c) = 2(A^T A)_{j,i} \quad \forall i, j \in \hat{m},$$
a tedy
$$\nabla^2 F(c) = 2A^T A.$$
Matice $A^T A$ je pozitivně definitní, neboť pro $x \in \mathbb{R}^m$ je
$$x^T A^T A x = (Ax)^T (Ax) = \|Ax\|^2 \geq 0$$
a rovnost nastává pouze pro $x = 0$, protože je $A^T A$ je regulární.

Nakonec ověříme, že je funkce F ryze konvexní na $\mathbb{R}^m$. Potom bude vektor c jediným globálním minimem F podle tvrzení ze Cvičení 3.11 a 3.12. Uvažujme tedy nějaká $c_1, c_2 \in \mathbb{R}^m$, $c_1 \neq c_2$ a $t \in (0, 1)$. Potom
$$\begin{aligned} F(tc_1 + (1-t)c_2) &= \|y - A\,(tc_1 + (1-t)c_2)\,\|^2 = \|t\,(y - Ac_1) + (1-t)\,(y - Ac_2)\,\|^2 \\ &\leq (t\|y - Ac_1\| + (1-t)\|y - Ac_2\|)^2 < tF(c_1) + (1-t)F(c_2), \end{aligned}$$
kde poslední nerovnost vyplývá z ryzí konvexnosti kvadratické funkce $x \mapsto x^2$.

3.7 Věta o inverzní funkci

Naším dalším cílem je dokázat Větu o inverzní funkci a s její pomocí také Větu o implicitní funkci. Obě věty mají mnoho aplikací.

Připomeňme, že derivaci $Df(a)$ chápeme podle potřeby buď jako lineární zobrazení, nebo jako jeho matici vzhledem ke standardním bázím a ve značení to nijak nerozlišujeme. Ztotožňujeme tak derivaci $Df(a)$ a Jacobiho matici
$$\frac{\partial(f_1, \dots, f_n)}{\partial(x_1, \dots, x_n)}(a)$$
pro $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ diferencovatelnou v bodě $a \in \Omega$. Použití symbolu $Df(a)$ ve smyslu Jacobiho matice f v bodě a je výhodné z důvodu úspory místa, a proto ho budeme nadále preferovat.



Obrázek 10: Ilustrace metody nejmenších čtverců. Data proložená polynomiální křivkou stupně 5.

Představme si, že chceme řešit rovnici tvaru $f(x) = y$, kde $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ je daná vektorová funkce. Rozepíšeme-li rovnici po komponentách, hledáme řešení obecného systému

$$\begin{aligned} f_1(x_1, \dots, x_n) &= y_1, \\ f_2(x_1, \dots, x_n) &= y_2, \\ &\vdots \\ f_n(x_1, \dots, x_n) &= y_n. \end{aligned}$$

To znamená, že chceme vědět, zda lze vyjádřit neznámé $x_1, \dots, x_n$ jako funkce $y_1, \dots y_n$. Takový systém téměř nikdy nelze řešit explicitně i v případě, že řešení existuje. Důležitou odpověď na otázku existence řešení nám dává právě Věta o inverzní funkci. Nicméně má pouze **lokální** charakter. Přesněji věta říká, že je-li $a \in \Omega$ a $f \in C^1(\Omega)$, vyplývá z podmínky $\det Df(a) \neq 0$, že f je prostá na nějakém okolí bodu a, a tudíž invertibilní.

I v případě, že je $\det Df(x) \neq 0$ ve všech bodech $x \in \Omega$, neplyne z toho, že je f prostá na množině Ω, viz Příklad 3.55. Zde je situace odlišná od speciálního případu funkce jedné proměnné, kde platí, že spojitě diferencovatelná funkce f na intervalu (a, b) s nenulovou derivací na (a, b) je (globálně) prostá na (a, b). Rozhodnout o (globální) prostotě funkce více proměnných je obecně obtížný problém.

Pro potřeby důkazu Věty o inverzní funkci si nejprve dokážeme jedno pomocné tvrzení týkající se invertibility lineárních zobrazení. Toto tvrzení nachází aplikace i na mnoha jiných místech. Připomeňme, že na prostoru $\mathcal{L}(\mathbb{R}^n)$ zavádíme normu podle Definice 3.3.

Lemma 3.53: Nechť $A, B \in \mathcal{L}(\mathbb{R}^n)$ a A je bijekce. Pokud platí nerovnost

$$\|B - A\| \|A^{-1}\| < 1,$$

potom je také B bijekce.

Důkaz. Operátor B lze zapsat ve tvaru

$$B = A + (B - A) = A(I - C),$$

kde $C := -A^{-1}(B-A)$. Odtud je jasné, že pokud je $I-C$ bijekce, musí být i B bijekcí. Důležité je, že

$$\|C\| \leq \|A^{-1}\| \|B - A\| < 1,$$

kde jsme využili nerovnost z Lemma 3.4 a předpoklad dokazovaného tvrzení. Stačí tedy dokázat implikaci

$$\|C\| < 1 \quad \Rightarrow \quad I - C \text{ je bijekce},$$

což provedeme v další části důkazu.

Důkaz provedeme tak, že inverzi k $I - C$ přímo zkonstruujeme. Definujme operátor $X := \lim_{n\to\infty} X_n$, kde

$$X_n := \sum_{k=0}^{n} C^k = I + C + \cdots + C^n.$$

Nejprve ověříme, že limita v definici X existuje. Protože má prostor $\mathcal{L}(\mathbb{R}^n)$ konečnou dimenzi (rovnu n^2), je $\mathcal{L}(\mathbb{R}^n)$ s normou $\|\cdot\|$ úplný prostor, viz Příklad 2.110. Stačí proto ukázat, že je posloupnost $\{X_n\}_{n=1}^{\infty}$ cauchyovská. Jelikož číselná řada

$$\sum_{k=0}^{\infty} \|C\|^k$$

konverguje, neboť podle předpokladu je $\|C\| < 1$, dostaneme z Bolzanovy–Cauchyho podmínky pro číselné řady

$$(\forall \epsilon > 0)(\exists n_0 \in \mathbb{N})(\forall n \geq n_0)(\forall p \in \mathbb{N}) \left(\sum_{k=n+1}^{n+p} \|C\|^k < \epsilon \right).$$

Potom pro $n \geq n_0$ a $p \in \mathbb{N}$ máme

$$\|X_{n+p} - X_n\| = \left\| \sum_{k=n+1}^{n+p} C^k \right\| \leq \sum_{k=n+1}^{n+p} \|C^k\| \leq \sum_{k=n+1}^{n+p} \|C\|^k < \epsilon,$$

kde jsme použili nerovnost $\|C^k\| \leq \|C\|^k$, která plyne z 1. tvrzení Lemma 3.4. Tedy $\{X_n\}_{n=1}^{\infty}$ je cauchyovská.

Nakonec ověříme, že $(I - C)X = I$, z čehož už vyplývá, že $I - C$ je bijekce (a navíc $(I - C)^{-1} = X$) díky konečnosti dimenze $\mathbb{R}^n$. Protože jsou I i C omezené operátory, neboť $\dim \mathbb{R}^n = n < \infty$, jsou podle Lemma 3.4 I i C spojité. Odtud poté máme

$$\begin{aligned}(I - C)X &= (I - C) \lim_{n\to\infty} X_n = \lim_{n\to\infty} (I - C) \sum_{k=0}^{n} C^k = \lim_{n\to\infty} \sum_{k=0}^{n} C^k - \sum_{k=0}^{n} C^{k+1} \\ &= \lim_{n\to\infty} (I - C^{n+1}) = I,\end{aligned}$$

kde jsme využili toho, že $C^n \to 0$, což plyne opět z nerovnosti $\|C^n\| \leq \|C\|^n$ a předpokladu $\|C\| < 1$. $\square$

Nyní již můžeme dokázat Větu o inverzní funkci.

Věta 3.54 (O inverzní funkci)**:** Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ je třídy C^1 na Ω a $a \in \Omega$. Pokud $\det Df(a) \neq 0$, potom existuje okolí H_a takové, že:

1. f je prosté na H_a,
2. $f(H_a)$ je otevřená,
3. inverzní zobrazení f^{-1} je třídy C^1 na $f(H_a)$, $\det Df(x) \neq 0$ pro všechna $x \in H_a$ a platí
$$Df^{-1}(f(x)) = (Df(x))^{-1}$$
pro všechna $x \in H_a$.

Důkaz. Označme $A := Df(a)$. Podle předpokladu je matice A regulární, a tudíž invertibilní. Dále z předpokladu $f \in C^1(\Omega)$ plyne, že je Df spojité zobrazení na Ω. Proto existuje $r > 0$ takové, že koule $B_a(r) \subset \Omega$ (z otevřenosti Ω) a pro všechny $x \in B_a(r)$ je

$$\|Df(x) - A\| < \frac{1}{2\|A^{-1}\|}.$$

Ukážeme, že $H_a := B_a(r)$ je hledaným okolím z tvrzení věty.

1\. Nejprve dokážeme, že f je prosté na H_a. Pro $y \in \mathbb{R}^n$ si definujeme pomocnou funkci $\phi_y : \Omega \to \mathbb{R}^n$ vztahem

$$\phi_y(x) := x + A^{-1}\left(y - f(x)\right).$$

Všimněte si, že

$$f(x) = y \quad \Leftrightarrow \quad x \textbf{ je pevný bod } \phi_y, \text{ tj. } \phi_y(x) = x.$$

Pro derivaci ϕ_y máme

$$D\phi_y(x) = I - A^{-1}Df(x) = A^{-1}\left(A - Df(x)\right)$$

pro každé $x \in \Omega$. Odtud pro libovolné $x \in H_a$ dostaneme

$$\|D\phi_y(x)\| \leq \|A^{-1}\|\|A - Df(x)\| < \|A^{-1}\|\frac{1}{2\|A^{-1}\|} = \frac{1}{2},$$

kde jsme využili nerovnost z Lemma 3.4. Potom z Věty 3.26 vyplývá nerovnost

$$\|\phi_y(x_1) - \phi_y(x_2)\| \leq \frac{1}{2}\|x_1 - x_2\| \tag{24}$$

pro všechna $x_1, x_2 \in H_a$, neboť okolí $H_a \equiv B_a(r)$ je konvexní množina.

Z nerovnosti (24) plyne, že zobrazení ϕ_y má nejvýše jeden pevný bod v množině H_a. Jinak by z předpokladu existence dvou různých bodů $x_1, x_2 \in H_a$ takových, že $\phi_y(x_1) = x_1$ a $\phi_y(x_2) =$

x_2, nerovnost (24) implikovala logický spor $1 \leq 1/2$. Jinými slovy k libovolnému $y \in \mathbb{R}^n$ existuje nejvýše jedno $x \in H_a$ takové, že $f(x) = y$, což znamená, že f je na H_a prosté.

2. Dokážeme, že $f(H_a)$ je otevřená. Nechť $y_0 \in f(H_a)$. Potom existuje $x_0 \in H_a$ tak, že $f(x_0) = y_0$. Jelikož je H_a otevřená, existuje $\epsilon > 0$ tak, že $\overline{B_{x_0}(\epsilon)} \subset H_a$. Ukážeme inkluzi

$$B_{y_0}\left(\frac{\epsilon}{2\|A^{-1}\|}\right) \subset f(H_a),$$

z čehož plyne otevřenost množiny $f(H_a)$.

Zvolme tedy $y \in B_{y_0}\left(\epsilon/(2\|A^{-1}\|)\right)$ libovolné ale pevné. Potom

$$\|\phi_y(x_0) - x_0\| = \|A^{-1}(y - y_0)\| \leq \|A^{-1}\|\|y - y_0\| < \|A^{-1}\|\frac{\epsilon}{2\|A^{-1}\|} = \frac{\epsilon}{2}.$$

Odtud s využitím nerovnosti (24) dostaneme pro každé $x \in \overline{B_{x_0}(\epsilon)}$ odhad

$$\|\phi_y(x) - x_0\| \leq \|\phi_y(x) - \phi_y(x_0)\| + \|\phi_y(x_0) - x_0\| < \frac{1}{2}\|x - x_0\| + \frac{\epsilon}{2} \leq \epsilon.$$

Tedy $\phi_y(x) \in B_{x_0}(\epsilon)$ pro $x \in \overline{B_{x_0}(\epsilon)}$. Z toho vidíme, že $\phi_y : \overline{B_{x_0}(\epsilon)} \to \overline{B_{x_0}(\epsilon)}$ je zobrazení mezi úplnými prostory $\overline{B_{x_0}(\epsilon)}$, neboť množina $\overline{B_{x_0}(\epsilon)}$ je uzavřená podmnožina úplného prostoru $\mathbb{R}^n$, viz Věta 2.115. Podle nerovnosti (24) je navíc ϕ_y kontrahující zobrazení na $\overline{B_{x_0}(\epsilon)}$. Potom podle Banachovy věty o pevném bodě (Věta 2.119) existuje $x \in \overline{B_{x_0}(\epsilon)}$ takové, že $\phi_y(x) = x$, neboli $f(x) = y$. Odtud plyne, že

$$y \in f\left(\overline{B_{x_0}(\epsilon)}\right) \subset f(H_a),$$

což jsme chtěli dokázat.

3. Zvolme $y \in f(H_a)$ a $k \in \mathbb{R}^n$ tak, že $y + k \in f(H_a)$. Potom existují $x, x + h \in H_a$ tak, že

$$f(x) = y \quad \text{a} \quad f(x + h) = y + k.$$

Nejprve odvodíme pomocnou nerovnost. Pro libovolné $z \in \mathbb{R}^n$ je

$$\phi_z(x + h) - \phi_z(x) = h + A^{-1}\left(f(x) - f(x + h)\right) = h - A^{-1}k.$$

Odtud a z nerovnosti (24) dostaneme

$$\|h - A^{-1}k\| \leq \frac{1}{2}\|h\|.$$

Přihlédneme-li ještě k nerovnosti, viz Cvičení 2.1,

$$\|h - A^{-1}k\| \geq \|h\| - \|A^{-1}k\|,$$

vidíme, že

$$\|h\| \leq 2\|A^{-1}k\| \leq 2\|A^{-1}\|\|k\|. \tag{25}$$

Na začátku důkazu jsme okolí H_a zkonstruovali tak, že platilo $\|A^{-1}\|\|Df(x) - A\| < 1/2$ pro všechna $x \in H_a$. Odtud a z Lemma 3.53 plyne, že je $Df(x)$ invertibilní pro všechna $x \in H_a$, neboli

$$(\forall x \in H_a)(\det Df(x) \neq 0).$$

Z toho dále vyvodíme, že je zobrazení $x \mapsto (Df(x))^{-1}$ spojité na H_a. To je vidět např. ze známého vzorečku pro inverzní matici z lineární algebry

$$(Df(x))^{-1} = \frac{1}{\det Df(x)} (Df(x))^{\mathrm{adj}},$$

neboť determinant $\det Df(x) \neq 0$ a je spojitou funkcí maticových elementů $Df(x)$, což jsou parciální derivace f, které jsou podle předpokladu $f \in C^1(\Omega)$ spojité. Také elementy adjugované matice $(Df(x))^{\mathrm{adj}}$ jsou spojitými funkcemi parciálních derivací f, neboť jsou opět definovány jako jisté determinanty.

Označíme-li si nyní inverzní funkci $g := f^{-1}$ na $f(H_a)$ (jejíž existence již byla dokázána), potom k důkazu 3. tvrzení nyní už stačí ukázat, že je g diferencovatelná na $f(H_a)$ a pro její derivaci platí

$$Dg(f(x)) = (Df(x))^{-1}$$

pro všechna $x \in H_a$. Potom bude $g \in C^1(f(H_a))$ a platí vzorec pro derivaci inverzní funkce.

Protože

$$\begin{aligned} g(y+k) - g(y) - (Df(x))^{-1} k &= h - (Df(x))^{-1} k \\ &= -(Df(x))^{-1} (f(x+h) - f(x) - Df(x)h), \end{aligned}$$

dostaneme s využitím (25) nerovnost

$$\frac{\|g(y+k) - g(y) - (Df(x))^{-1} k\|}{\|k\|} \leq 2\|A^{-1}\|\| (Df(x))^{-1} \| \frac{\|f(x+h) - f(x) - Df(x)h\|}{\|h\|}$$

Z nerovnosti (25) plyne, že $h \to 0$, pokud $k \to 0$. Protože výraz napravo v poslední nerovnosti jde k nule pro $h \to 0$, neboť f je diferencovatelná v $x \in H_a$, dostáváme konečně, že

$$\lim_{k \to 0} \frac{\|g(y+k) - g(y) - (Df(x))^{-1} k\|}{\|k\|} = 0,$$

což jsme chtěli dokázat. □

Poznámka: Všimněte si, že předpoklad $f \in C^1(\Omega)$ byl plně využit až v důkazu 3. tvrzení Věty 3.54. Pro důkaz tvrzení 1. a 2. Věty 3.54 stačilo předpokládat, že f je diferencovatelná na Ω, Df spojitá v bodě a a $\det Df(a) \neq 0$. Spojitost derivace f v bodě a už ale vypustit nelze, viz Cvičení 3.17.

Poznámka: Pokud je funkce f ve Větě 3.54 o inverzní funkci třídy C^k na Ω, kde $k \geq 2$, potom je také f^{-1} třídy C^k na $f(H_a)$. To lze dokázat indukcí a derivováním vztahu pro Df^{-1} z 3. tvrzení Věty 3.54. Postup jenom naznačíme. Využijeme-li vyjádření inverzní matice pomocí matice adjugované, můžeme vzorec pro parciální derivaci $\partial\left(f^{-1}\right)_i/\partial y_j$ v bodě $y \in f(H_a)$ napsat ve tvaru

$$\frac{\partial\left(f^{-1}\right)_i}{\partial y_j}(y) = \left(Df^{-1}(y)\right)_{i,j} = \left(Df(f^{-1}(y))\right)^{-1}_{i,j} = \frac{(Df(f^{-1}(y)))^{adj}_{i,j}}{\det Df(f^{-1}(y))},$$

kde $i, j \in \hat{n}$. Z tohoto vyjádření je vidět, že jak čitatel tak jmenovatel výrazu vpravo jsou polynomiální funkce parciálních derivací $\partial f_r/\partial x_s$, $r, s \in \hat{n}$, v bodech $f^{-1}(y)$, tj.

$$\frac{\partial\left(f^{-1}\right)_i}{\partial y_j}(y) = \frac{P_{i,j}\left(\left(\frac{\partial f_r}{\partial x_s}(f^{-1}(y))\right)_{r,s\in\hat{n}}\right)}{Q\left(\left(\frac{\partial f_r}{\partial x_s}(f^{-1}(y))\right)_{r,s\in\hat{n}}\right)} \tag{26}$$

pro nějaké polynomy $P_{i,j}$ a Q v n^2 proměnných. Navíc nenulovost polynomu Q ve jmenovateli je zaručena Větou 3.54. Potom je-li např. $f \in C^2(\Omega)$, je také $f \in C^1(\Omega)$ a podle indukčního předpokladu je i $f^{-1} \in C^1(f(H_a))$. Potom bychom parciálním derivováním výrazu vpravo v rovnosti (26) podle y_k dostali napravo racionální funkci v proměnných $\partial^2 f_r/\partial x_p\partial x_s$ a $\partial f_l^{-1}/\partial y_k$, kde $p, r, s, l \in \hat{n}$, s nenulovým jmenovatelem. Z toho vyplývá, že existují druhé parciální derivace $\partial^2(f^{-1})_i/\partial y_k\partial y_j$ a jsou spojité na $f(H_a)$.

Příklad 3.55: Uvažujme vektorovou funkci $f : \mathbb{R}^2 \to \mathbb{R}^2$ definovanou vztahem

$$f(x,y) = \begin{pmatrix} e^x \cos y \\ e^x \sin y \end{pmatrix}.$$

Zřejmě je f hladká na $\mathbb{R}^2$, a tudíž také $f \in C^1(\mathbb{R}^2)$. Dále

$$\det Df(x,y) = \det\begin{pmatrix} e^x\cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix} = e^{2x}\left(\cos^2 y + \sin^2 y\right) = e^{2x} \neq 0$$

pro všechna $(x,y) \in \mathbb{R}^2$. Podle Věty 3.54 existuje pro každý bod $(x,y) \in \mathbb{R}^2$ nějaké okolí (x,y), na němž je funkce f prostá. Jinými slovy f je lokálně prostá na celém $\mathbb{R}^2$. Avšak f není prostá na $\mathbb{R}^2$, protože

$$f(x,y) = f(x, y+2\pi)$$

pro všechna $(x,y) \in \mathbb{R}^2$.

Přímým důsledkem Věty 3.54 je tzv. Věta o otevřeném zobrazení (Open mapping theorem). Zobrazení se nazývá *otevřené*, pokud zobrazuje otevřené množiny na otevřené množiny (pozor, neplést se spojitostí zobrazení, pro které jsou naopak **vzory** otevřených množin otevřené). V matematice existuje víc vět o otevřeném zobrazení. Zde uvedená se týká spojitě diferencovatelných funkcí. Další věta o otevřeném zobrazení platí např. pro holomorfní funkce.

Věta 3.56: (O otevřeném zobrazení) Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$, $f \in C^1(\Omega)$ a $\det Df(x) \neq 0$ pro všechna $x \in \Omega$. Potom je f otevřené, tj. pro každou otevřenou množinu $U \subset \Omega$ je $f(U)$ otevřená množina.

Důkaz. Buď $U \subset \Omega$ otevřená množina. Je-li $U = \emptyset$, pak $f(U) = \emptyset$. Nechť tedy $U \neq \emptyset$. Zvolme $b \in f(U)$. Potom existuje $a \in U$ tak, že $f(a) = b$. Podle předpokladů je $\det Df(a) \neq 0$ a jsou splněny předpoklady Věty 3.54. Potom existuje okolí H_a takové, že $f(H_a)$ je otevřená. Navíc můžeme bez újmy na obecnosti předpokládat, že $H_a \subset U$, jinak bychom vzali $H_a \cap U$. Množina $f(H_a)$ je tedy otevřená a navíc $b \in f(H_a)$, tzn., že $f(H_a)$ je okolí b. Dále z inkluze $H_a \subset U$ plyne, že $f(H_a) \subset f(U)$. Tudíž bod b leží v množině $f(U)$ i se svým okolím $f(H_a)$, a proto je $f(U)$ otevřená. $\square$

Tvrzení Věty 3.54 nelze obrátit v následujícím smyslu: Z prostoty funkce $f \in C^1(\Omega)$ **neplyne**, že $\det Df(x) \neq 0$ pro všechna $x \in \Omega$. To ilustruje už jednoduchý příklad funkce jedné proměnné $f(x) = x^3$, která je bijekcí $\mathbb{R}$ na $\mathbb{R}$, avšak $f'(0) = 0$. Přidáme-li ovšem předpoklad diferencovatelnosti inverze f^{-1} na Ω, potom již $\det Df(x) \neq 0$ pro všechna $x \in \Omega$. V takovém případě můžeme totiž aplikovat Větu 3.7 při derivaci identity

$$\left(f^{-1} \circ f\right)(x) = x,$$

ze které plyne, že

$$Df^{-1}\left(f(x)\right) Df(x) = I,$$

a tudíž $\det Df(x) \neq 0$ pro všechna $x \in \Omega$.

Definice 3.57 (Difeomorfismus)**:** Buďte $U, V \subset \mathbb{R}^n$ otevřené množiny. Zobrazení $f : U \to V$ nazýváme *difeomorfismus*, právě když f je bijekce U na V, $f \in C^1(U)$ a $f^{-1} \in C^1(V)$.

Poznámka: Požadavek $f^{-1} \in C^1(V)$ v definici difeomorfismu není nadbytečný, jak opět ilustruje např. funkce $f(x) = x^3$, která je hladkou bijekcí $\mathbb{R}$ na $\mathbb{R}$, ale inverzní funkce $f^{-1}(y) = \sqrt[3]{y}$ není diferencovatelná v bodě $y = 0$.

Zdůrazněme ještě jednou, že funkce splňující předpoklady Věty o inverzní funkci ještě nemusí být difeomorfismus (je to pouze lokální difeomorfismus). Dále z Věty 3.5 plyne, že každý difeomorfismus je homeomorfismus.

Příklad 3.58: Buď $f : \mathbb{R}^2 \to \mathbb{R}^2$ definována vztahem

$$f(x, y) = \begin{pmatrix} x^2 + y^2 \\ 2xy \end{pmatrix}.$$

Rozhodněte, na jaké množině $U \subset \mathbb{R}^2$ je f lokálně prostá a spočítejte $Df^{-1}\left(f(x, y)\right)$ pro $(x, y) \in U$.

Protože

$$\det Df(x, y) = \det \begin{pmatrix} 2x & 2y \\ 2y & 2x \end{pmatrix} = 4(x^2 - y^2),$$

je f lokálně invertibilní na množině $U = \{(x, y) \mid |x| \neq |y|\}$ podle Věty o inverzní funkci. Množina U rozděluje rovinu $\mathbb{R}^2$ na čtyři kvadranty. Můžete si rozmyslet, že f je prostá na každém z těchto kvadrantů a z rovnice $f(x, y) = (u, v)$ spočítat inverzní funkci. Např. pro (x, y) z kvadrantu $U_1 := \{(x, y) \mid |x| < y \ \wedge \ y > 0\}$ dostaneme

$$\begin{pmatrix} x \\ y \end{pmatrix} = f^{-1}(u, v) = \frac{1}{2} \begin{pmatrix} \sqrt{u+v} - \sqrt{u-v} \\ \sqrt{u+v} + \sqrt{u-v} \end{pmatrix}$$

a $f(U_1) = \{(u,v) \mid |v| < u \ \wedge \ u > 0\}$ (ověřte). Derivaci f^{-1} můžeme buďto hledat přímo, známe-li funkci f^{-1} explicitně, nebo opět podle Věty o inverzní funkci. V druhém případě máme

$$Df^{-1}(f(x,y)) = (Df(x,y))^{-1} = \frac{1}{4(x^2-y^2)}\begin{pmatrix} 2x & -2y \\ -2y & 2x \end{pmatrix} = \frac{1}{2(x^2-y^2)}\begin{pmatrix} x & -y \\ -y & x \end{pmatrix}$$

pro $(x,y) \in U$.

3.8 Věta o implicitní funkci

Problematiku implicitní funkce si nejprve vysvětlíme na jednoduchém příkladu s kružnicí. Zajímá nás, kdy a pokud vůbec rovnice

$$F(x,y) := x^2 + y^2 - 1 = 0 \tag{27}$$

definuje y jako funkci x. Samozřejmě víme, že řešení rovnice (27) existují, pokud $x \in [-1,1]$, a jsou $y = \pm\sqrt{1-x^2}$. Z definice funkce musí být její hodnoty určeny jednoznačně, a proto bez dalšího omezení na y (např. $y > 0$), nezadává rovnice (27) funkci $y = y(x)$ jednoznačně.

Řekněme, že nás zajímá, zda rovnice (27) zadává řešení $y = y(x)$ alespoň *lokálně*. Tzn., že předpokládáme, že je dáno nějaké (a,b) tak, že $F(a,b) = 0$ (bod na kružnici) a zajímá nás, jestli existují nějaká okolí H_a a H_b a funkce $f : H_a \to H_b$ tak, že $y = f(x)$ vyhovuje rovnici $F(x, f(x)) = 0$ pro všechna $x \in H_a$ a $f(a) = b$. Na příkladu kružnice odpověď najdeme snadno. Je-li $a \in (-1,1)$ a $b > 0$, potom je $f(x) = \sqrt{1-x^2}$, okolí H_a lze volit libovolně tak, aby $a \in H_a \subset (-1,1)$ a $H_b := f(H_a)$. Podobně je-li $a \in (-1,1)$ a $b < 0$, potom je $f(x) = -\sqrt{1-x^2}$ a volbu okolí lze provést stejně jako předtím. Na druhou stranu na okolí bodů $(a,b) = (\pm 1, 0)$ rovnice (27) jednoznačně funkci $y = f(x)$ *nezadává*. Tyto výjimečné body lze najít jako řešení rovnice

$$0 = \frac{\partial F}{\partial y}(a,b) = 2b$$

spolu s $F(a,b) = 0$. Můžeme se tedy domnívat, že podmínka $\partial_y F(a,b) \neq 0$ garantuje existenci okolí H_a a H_b a funkce $f : H_a \to H_b$ tak, že $F(x, f(x)) = 0$ pro všechna $x \in H_a$ a $f(a) = b$. Tato domněnka se ukáže jako pravdivá a takovouto funkci f nazýváme *implicitní funkcí* zadanou rovnicí $F(x,y) = 0$. Navíc jakmile víme, že implicitní funkce existuje, můžeme z rovnice $F(x, f(x)) = 0$ spočítat derivaci f, i když neznáme f explicitně. V případě s kružnicí (27) dostaneme derivováním rovnosti $x^2 + y^2 - 1 = 0$ podle x, kde $y = y(x)$, vztah

$$2x + 2yy' = 0.$$

Neboli

$$y' = -\frac{x}{y},$$

což je stejný výsledek, jaký bychom dostali derivováním explicitní formule $y = \pm\sqrt{1-x^2}$:

$$y' = -\frac{x}{\pm\sqrt{1-x^2}} = -\frac{x}{y}$$

bez ohledu na to, jakou funkci $y = \pm\sqrt{1-x^2}$ bereme.

Větu o implicitní funkci vyslovíme pro vektorovou funkci $F : \Omega_n \times \Omega_m \to \mathbb{R}^m$, kde $\Omega_n \subset \mathbb{R}^n$ a $\Omega_m \subset \mathbb{R}^m$ jsou otevřené množiny. Věta nám potom říká, za jakých předpokladů sytém m rovnic o $m+n$ neznámých

$$F(x, y) = 0$$

určuje lokálně řešení $y = y(x)$ na okolí nějakého bodu (a, b), který vyhovuje rovnici $F(a, b) = 0$. Klíčovým předpokladem bude nenulovost Jacobiánu zobrazení F jako funkce y v bodě (a, b). Pro zjednodušení zápisu použijeme stručnější značení

$$\frac{\partial F}{\partial y} := \frac{\partial\left(F_1, \ldots, F_m\right)}{\partial\left(y_1, \ldots, y_m\right)} = \begin{pmatrix} \frac{F_1}{\partial y_1} & \cdots & \frac{F_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{F_m}{\partial y_1} & \cdots & \frac{F_m}{\partial y_m} \end{pmatrix}$$

a

$$\frac{\partial F}{\partial x} := \frac{\partial\left(F_1, \ldots, F_m\right)}{\partial\left(x_1, \ldots, x_n\right)} = \begin{pmatrix} \frac{F_1}{\partial x_1} & \cdots & \frac{F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{F_m}{\partial x_1} & \cdots & \frac{F_m}{\partial x_n} \end{pmatrix}.$$

Věta 3.59 (O implicitní funkci)**:** Nechť $\Omega_n \subset \mathbb{R}^n$ a $\Omega_m \subset \mathbb{R}^m$ jsou otevřené množiny a $F : \Omega_n \times \Omega_m \to \mathbb{R}^m$ je třídy C^1 na $\Omega_n \times \Omega_m$. Dále předpokládejme, že pro nějaké $(a, b) \in \Omega_n \times \Omega_m$, platí:

$$F(a, b) = 0$$

a

$$\det \frac{\partial F}{\partial y}(a, b) \neq 0.$$

Potom:

1. Existují okolí $H_a \subset \Omega_n$ a $H_b \subset \Omega_m$ bodů a a b a jediná funkce $f : H_a \to H_b$ třídy C^1 na H_a tak, že
$$(\forall x \in H_a)\,(F(x, f(x)) = 0) \quad \text{a} \quad f(a) = b.$$

2. Dále $\det \frac{\partial F}{\partial y}(x, f(x)) \neq 0$ pro všechna $x \in H_a$ a pro derivaci f platí vzorec
$$Df(x) = -\left(\frac{\partial F}{\partial y}(x, f(x))\right)^{-1} \frac{\partial F}{\partial x}(x, f(x))$$
pro všechna $x \in H_a$.

Důkaz. 1. K důkazu použijeme Větu 3.54 o inverzní funkci. Definujme $G : \Omega_n \times \Omega_m \to \mathbb{R}^{n+m}$ vztahem $G(x, y) := (x, F(x, y))$. Podle předpokladu je $F \in C^1\left(\Omega_n \times \Omega_m\right)$, a proto také $G \in C^1\left(\Omega_n \times \Omega_m\right)$. Jacobiho matice G má blokový tvar

$$DG = \begin{pmatrix} I & 0 \\ \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \end{pmatrix},$$

kde $I \in \mathbb{R}^{n,n}$ je jednotková matice a $0 \in \mathbb{R}^{n,m}$ nulová matice. Odtud pro Jacobián G v bodě (a, b) dostaneme

$$\det DG(a,b) = \det \frac{\partial F}{\partial y}(a,b) \neq 0.$$

Tudíž můžeme aplikovat Větu 3.54 o inverzní funkci na funkci G a bod (a, b).

Podle Věty 3.54 existuje okolí $H_{(a,b)}$ bodu (a, b), které můžeme bez újmy na obecnosti uvažovat ve tvaru $H_{(a,b)} = H_a \times H_b$ (jinak vezmeme menší okolí H_a a H_b tak, aby $H_a \times H_b \subset H_{(a,b)}$), a okolí $V := G(H_a \times H_b)$ bodu $G(a, b) = (a, 0)$ tak, že G je prosté zobrazení $H_a \times H_b$ na otevřenou množinu V a $G^{-1} \in C^1(V)$. Jelikož zobrazení G je identické v první komponentě, je V tvaru $V = H_a \times F(H_a \times H_b)$ a také inverzní zobrazení G^{-1} bude v první komponentě fungovat jako identické zobrazení. Tedy G^{-1} je tvaru $G^{-1}(x, y) = (x, \Phi(x, y))$, kde $\Phi : V \to H_b$ je bijekce třídy C^1 na V. Pro libovolné $(x, y) \in V$ tedy máme

$$(x,y) = G \circ G^{-1}(x,y) = G\left(x, \Phi(x,y)\right) = \left(x, F\left(x, \Phi(x,y)\right)\right).$$

a odtud speciálně $y = F(x, \Phi(x, y))$. Položíme-li $y = 0$ a definujeme $f(x) := \Phi(x, 0)$ dostáváme z poslední rovnosti, že $F(x, f(x)) = 0$ pro všechna $x \in H_a$ Funkce $f \in C^1(H_a)$, protože $\Phi \in C^1(V)$. Navíc $(a, b) = G^{-1}(a, 0) = (a, \Phi(a, 0)) = (a, f(a))$, a tudíž $b = f(a)$.

V tvrzení 1. zbývá ověřit jednoznačnost. K libovolnému $x \in H_a$ jsme našli $y := f(x) \in H_b$ tak, že $F(x, y) = 0$. Předpokládejme, že k témuž $x \in H_a$ existuje nějaké další $\tilde{y} \in H_b$, $\tilde{y} \neq y$ tak, že $F(x, \tilde{y}) = 0$. Potom

$$G(x,y) = (x, F(x,y)) = (x,0) = (x, F(x,\tilde{y})) = G(x,\tilde{y}),$$

což je spor s tím, že G je prostá na $H_a \times H_b$. Ke každému $x \in H_a$ tedy existuje jediné $y \in H_b$ tak, že $F(x, y) = 0$ a funkce f je tak určena jednoznačně.

2. Už víme, že $f \in C^1(H_a)$, a tudíž můžeme derivovat rovnost $F(x, f(x)) = 0$ na H_a, kterou můžeme přepsat do tvaru $F(g(x)) = 0$, kde $g(x) := (x, f(x))^T$. Aplikací řetězového pravidla, viz Věta 3.7, dostaneme

$$0 = D(F \circ g)(x) = DF(g(x))Dg(x) = \begin{pmatrix} \frac{\partial F}{\partial x}(g(x)) & \frac{\partial F}{\partial y}(g(x)) \end{pmatrix} \begin{pmatrix} I \\ Df(x) \end{pmatrix},$$

neboli

$$0 = \frac{\partial F}{\partial x}(g(x)) + \frac{\partial F}{\partial y}(g(x))Df(x) \tag{28}$$

pro všechna $x \in H_a$. Protože je podle 3. tvrzení Věty 3.54 o inverzní funkci

$$\det \frac{\partial F}{\partial y}(g(x)) = \det G(g(x)) \neq 0$$

pro všechna $x \in H_a$, je matice $\frac{\partial F}{\partial y}(g(x))$ regulární a z rovnice (28) tak plyne, že

$$Df(x) = -\left(\frac{\partial F}{\partial y}(g(x))\right)^{-1} \frac{\partial F}{\partial x}(g(x))$$

pro všechna $x \in H_a$, což jsme chtěli dokázat. □

Poznámka: Ve smyslu poznámky za Větou 3.54 můžeme také předpoklad ve Větě 3.59 mírně zeslabit. Konkrétně platí věta: Nechť jsou $\Omega_n \subset \mathbb{R}^n$ a $\Omega_m \subset \mathbb{R}^m$ otevřené množiny, $F : \Omega_n \times \Omega_m \to \mathbb{R}^m$ spojité, existuje $\partial F/\partial y$ na $\Omega_n \times \Omega_m$ a je spojité v bodě $(a, b) \in \Omega_n \times \Omega_m$. Potom je-li

$$F(a, b) = 0 \quad \text{a} \quad \det \frac{\partial F}{\partial y}(a, b) \neq 0,$$

existují okolí $H_a \subset \Omega_n$ a $H_b \subset \Omega_m$ bodů a a b a právě jedna spojitá funkce $f : H_a \to H_b$ tak, že

$$(\forall x \in H_a)\,(F(x, f(x)) = 0) \quad \text{a} \quad f(a) = b.$$

Příklad 3.60: Jedno řešení systému rovnic

$$\begin{aligned} xy^2 + xzu + yv^2 &= 3, \\ yzu^3 + 2xv - u^2v^2 &= 2, \end{aligned}$$

je $(x, y, z, u, v) = (1, 1, 1, 1, 1)$. Rozhodněte, zda na okolí bodu $(x, y, z) = (1, 1, 1)$ určují rovnice implicitní funkce $u = u(x, y, z)$ a $v = v(x, y, z)$ takové, že $u(1, 1, 1) = 1$ a $v(1, 1, 1) = 1$ a poté spočítejte $\partial_y u(1, 1, 1)$ a $\partial_y v(1, 1, 1)$.

Do Věty o implicitní funkci položíme $n = 3, m = 2$, funkci $F : \mathbb{R}^3 \times \mathbb{R}^2 \to \mathbb{R}^2$ definovanou vztahem

$$F(x, y, z, u, v) := \begin{pmatrix} xy^2 + xzu + yv^2 - 3 \\ yzu^3 + 2xv - u^2v^2 - 2 \end{pmatrix}$$

a $a = (1, 1, 1)$ a $b = (1, 1)$. Zřejmě je F hladká. Dále

$$\frac{\partial(F_1, F_2)}{\partial(u, v)} = \begin{pmatrix} xz & 2yv \\ 3yzu^2 - 2uv^2 & 2x - 2u^2v \end{pmatrix},$$

a proto

$$\det \frac{\partial(F_1, F_2)}{\partial(u, v)}(1, 1, 1, 1, 1) = \det \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} = -2 \neq 0.$$

Proto podle Věty o implicitní funkci existuje funkce $f(x, y, z) = (f_1(x, y, z), f_2(x, y, z))$ třídy C^1 definovaná na okolí bodu $(1, 1, 1)$ taková, že

$$u = f_1(x, y, z), \quad v = f_2(x, y, z) \quad \text{a} \quad f(1, 1, 1) = (1, 1).$$

Abychom spočítali $\partial_y u(1, 1, 1)$ a $\partial_y v(1, 1, 1)$, můžeme parciálně derivovat podle y zadaný systém rovnic, kde u a v chápeme jako funkce x, y, z. Dostaneme rovnice

$$\begin{aligned} 2xy + xz\partial_y u + v^2 + 2yv\partial_y v &= 0, \\ 3yzu^2\partial_y u + zu^3 + 2x\partial_y v - 2uv^2\partial_y u - 2u^2v\partial_y v &= 0. \end{aligned}$$

Dosadíme-li $(x, z, y, u, v) = (1, 1, 1, 1, 1)$, máme systém

$$\begin{aligned} \partial_y u(1, 1, 1) + 2\partial_y v(1, 1, 1) &= -3, \\ \partial_y u(1, 1, 1) &= -1, \end{aligned}$$

jehož řešením je $\partial_y u(1, 1, 1) = -1$ a $\partial_y v(1, 1, 1) = -1$. Podobně bychom mohli spočítat další parciální derivace u a v v bodě $(1, 1, 1)$.

3.9 Vázané extrémy funkce více proměnných

V předchozí části jsme si ukázali metodu, jak hledat lokální extrémy funkce více proměnných. Nyní se budeme zabývat podobným problémem jen navíc zúžíme definiční obor účelové funkce přidáním tzv. *rovnostních vazeb*. Tzn., že budeme hledat lokální extrémy funkce $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ v množině tzv. *přípustných řešení*

$$M = \{x \in \Omega \mid g(x) = 0\},$$

kde $g : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ je daná vektorová funkce reprezentující *vazby*. V praktických problémech bývá počet proměnných n větší než počet vazeb m. Tato úloha je speciálním případem tzv. obecné optimalizační úlohy nelineárního programování, v níž navíc vystupují také nerovnostní vazby. Naším cílem bude dokázat si nutnou a postačující podmínku pro extrémy funkce s rovnostními vazbami podobné těm, které jsme si odvodili pro (volné) extrémy bez vazeb, viz Věty 3.43 a 3.48.

Definice 3.61 ((Ostré) lokální maximum/minimum vzhledem k množině)**:** Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ a $a \in M \subset \Omega$. Řekneme, že funkce f má v bodě a:

1. *lokální maximum vzhledem k* $M \;\Leftrightarrow\; (\exists H_a \subset \Omega)(\forall x \in H_a \cap M)(f(x) \leq f(a))$,
2. *ostré lokální maximum vzhledem k* $M \;\Leftrightarrow\; (\exists H_a \subset \Omega)(\forall x \in (H_a \cap M) \setminus \{a\})(f(x) < f(a))$,
3. *lokální minimum vzhledem k* $M \;\Leftrightarrow\; (\exists H_a \subset \Omega)(\forall x \in H_a \cap M)(f(x) \geq f(a))$,
4. *ostré lokální minimum vzhledem k* $M \;\Leftrightarrow\; (\exists H_a \subset \Omega)(\forall x \in (H_a \cap M) \setminus \{a\})(f(x) > f(a))$.

Bod a se nazývá *lokální extrém* funkce f vzhledem k M, právě když má f v a lokální minimum nebo lokální maximum vzhledem k M.

Pro formulaci nutné a postačující podmínky vázaného extrému použijeme pomocnou funkci pojmenovanou po J.–L. Lagrangeovi.

Definice 3.62 (Lagrangeova funkce, Lagrangeovy multiplikátory)**:** Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ a $g : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$. Funkci $L : \Omega \times \mathbb{R}^m \to \mathbb{R}$ definovanou vztahem

$$L(x; \lambda) := f(x) - \sum_{j=1}^{m} \lambda_j g_j(x)$$

nazýváme *Lagrangeova funkce* a parametry $\lambda = (\lambda_1, \ldots, \lambda_m)$ *Lagrangeovy multiplikátory*.

Uvažujme na chvíli, že je dána pouze jediná vazba, tj. $m = 1$ a skalární funkce $g : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ je třídy C^1. Připomeňme, že potom tečný prostor k množině $M = \{x \in \Omega \mid g(x) = 0\}$ v bodě $a \in M$ je ortogonální doplněk gradientu $\nabla g(a)$, tj.

$$T_a(M) = (\nabla g(a))^{\perp}.$$

Je-li $m \geq 1$, potom je $M = \cap_{j=1}^{m} M_j$, kde

$$M_j := \{x \in \Omega \mid g_j(x) = 0\}.$$

Rozšíříme definici tečného prostoru k $M = \{x \in \Omega \mid g(x) = 0\}$ v bodě $a \in M$ pro případ $m \geq 1$ následovně:

$$T_a(M) := \bigcap_{j=1}^{m} T_a(M_j).$$

Nejprve si dokážeme pomocné tvrzení o tečném prostoru $T_a(M)$.

Lemma 3.63: Buď $g : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ třídy C^1 na Ω, $M = \{x \in \Omega \mid g(x) = 0\}$ a $a \in M$. Potom platí:

1. $T_a(M) = \ker Dg(a)$.
2. Je-li $h(Dg(a)) = m < n$, potom ke každému $v \in T_a(M)$ existuje okolí počátku $H_0 \subset \mathbb{R}$ a křivka $\phi : H_0 \to M$ třídy C^1 tak, že $\phi(0) = a$ a $D\phi(0) = v$.

Důkaz. 1. Stačí si uvědomit, že

$$x \in T_a(M) = \bigcap_{j=1}^{m} T_a(M_j) = \bigcap_{j=1}^{m} (\nabla g_j(a))^{\perp},$$

právě když $\nabla g_j(a)x = 0$ pro všechna $j \in \hat{m}$. Jelikož gradienty $\nabla g_j(a)$, $j \in \hat{m}$, jsou řádky matice $Dg(a)$, je $x \in T_a(M)$, právě když $Dg(a)x = 0$, neboli $x \in \ker Dg(a)$.

2. Jelikož $h(Dg(a)) = m < n$, existuje m sloupců matice $Dg(a)$, které tvoří lineárně nezávislý soubor v $\mathbb{R}^n$. Bez újmy na obecnosti můžeme předpokládat, že je prvních m sloupců $Dg(a)$ lineárně nezávislých, jinak přejmenujeme proměnné $x_1, \ldots, x_n$. Definujme zobrazení $F : \Omega \to \mathbb{R}^n$ vztahem

$$F_j(x) := \begin{cases} g_j(x), & \text{pro} \quad j \in \hat{m}, \\ x_j, & \text{pro} \quad j \in \{m+1, \ldots, n\}. \end{cases}$$

Všimněte si, že $x \in M$, právě když $F_1(x) = \cdots = F_m(x) = 0$.

Podle předpokladu je $g \in C^1(\Omega)$, a proto také $F \in C^1(\Omega)$. Matice derivace F má blokový tvar

$$DF = \begin{pmatrix} \frac{\partial(g_1,\ldots,g_m)}{\partial(x_1,\ldots,x_m)} & \frac{\partial(g_1,\ldots,g_m)}{\partial(x_{m+1},\ldots,x_n)} \\ 0 & I \end{pmatrix}, \tag{29}$$

kde $0 \in \mathbb{R}^{n-m,m}$ je nulová matice a $I \in \mathbb{R}^{n-m,n-m}$ je jednotková matice. Tudíž matice $DF(a)$ je regulární, neboť

$$\det DF(a) = \det \frac{\partial(g_1, \ldots, g_m)}{\partial(x_1, \ldots, x_m)}(a) \neq 0.$$

Proto podle Věty 3.54 o inverzní funkci existuje $H_a \subset \Omega$, na němž je zobrazení F prosté a

$$F(a) = (\underbrace{0, \ldots, 0}_{m \text{ krát}}, a_{m+1}, \ldots, a_n)^T,$$

protože $g(a) = 0$, což plyne z předpokladu $a \in M$.

Označíme-li $G := F^{-1}$, potom $G : F(H_a) =: H_{F(a)} \to H_a$ je bijekce. Zvolme $v \in T_a(M)$ a okolí $H_0 \subset \mathbb{R}$ tak, aby

$$(0, \dots, 0, a_{m+1} + tv_{m+1}, \dots, a_n + tv_n)^T \in H_{F(a)}$$

pro všechna $t \in H_0$, což lze, neboť $H_{F(a)}$ je otevřená množina a $F(a) = (0, \dots, 0, a_{m+1}, \dots, a_n)^T$. Definujme $\phi : H_0 \to H_a$ vztahem

$$\phi(t) := G(0, \dots, 0, a_{m+1} + tv_{m+1}, \dots, a_n + tv_n).$$

Ve zbývající části důkazu ukážeme, že ϕ má všechny tři vlastnosti z tvrzení věty, tj. $\phi(H_0) \subset M$, $\phi(0) = a$ a $D\phi(0) = v$.

Z definice ϕ plyne, že $F_1(\phi(t)) = \dots = F_m(\phi(t)) = 0$ pro všechna $t \in H_0$. Tudíž $\phi(H_0) \subset M$. Dále

$$\phi(0) = G(0, \dots, 0, a_{m+1}, \dots, a_n) = G(F(a)) = a.$$

Nakonec dokážeme rovnost $DF(a)D\phi(0) = DF(a)v$, což implikuje $D\phi(0) = v$ díky regularitě $DF(a)$. Nejprve si všimneme, že pro $t \in H_0$ máme

$$(F \circ \phi)(t) = (0, \dots, 0, a_{m+1} + tv_{m+1}, \dots, a_n + tv_n)^T.$$

Odtud dostaneme

$$DF(a)D\phi(0) = D(F \circ \phi)(0) = (0, \dots, 0, v_{m+1}, \dots, v_n)^T = DF(a)v.$$

Poslední rovnost platí díky blokovému tvaru (29) matice $DF(a)$ a toho, že $Dg(a)v = 0$ podle již dokázaného bodu 1. □

Poznámka: Je-li $g : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ třídy C^1 a $h(Dg(x)) = m < n$ pro všechna $x \in \Omega$, potom je množina $M = \{x \in \Omega \mid g(x) = 0\}$ speciálním případem tzv. *diferencovatelné variety* dimenze $n - m$.

Nyní si již můžeme dokázat nutnou podmínku vázaného extrému.

Věta 3.64 (Nutná podmínka vázaného extrému)**:** Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ a $g : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ jsou funkce třídy C^1 na Ω a $m < n$. Předpokládejme dále, že má funkce f lokální extrém vzhledem k množině $M := \{x \in \Omega \mid g(x) = 0\}$ v bodě $a \in M$ a že $h(Dg(a)) = m$. Potom existují čísla $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ taková, že

$$\nabla_x L(a; \lambda) = \nabla f(a) - \sum_{j=1}^{m} \lambda_j \nabla g_j(a) = 0.$$

(Zde index x v ∇_x naznačuje, že se v gradientu ∇_x objevují pouze parciální derivace podle proměnných $x_1, \dots, x_n$, nikoliv podle $\lambda_1, \dots, \lambda_m$.)

Důkaz. Ukážeme, že $\nabla f(a) \in (T_a(M))^\perp$. Zvolme nějaké $v \in T_a(M)$. Podle předpokladu je $h(Dg(a)) = m < n$, a tedy podle Lemma 3.63 existuje $H_0 \subset \mathbb{R}$ a zobrazení $\phi : H_0 \to M$ třídy C^1 tak, že $\phi(0) = a$ a $D\phi(0) = v$. Protože má f lokální extrém v bodě a vzhledem k M, má složená funkce $f \circ \phi : H_0 \to \mathbb{R}$ lokální extrém v bodě 0. Tudíž podle nutné podmínky (volného) extrému je $(f \circ \phi)'(0) = 0$. S využitím Věty 3.7 o derivaci složené funkce tak dostaneme, že

$$0 = (f \circ \phi)'(0) = \nabla f\left(\phi(0)\right) D\phi(0) = \nabla f(a)v,$$

a tedy $\nabla f(a)$ je kolmé na v.

Použijeme-li jednu vlastnost ortogonálního doplňku, viz Cvičení 3.21, dostaneme

$$(T_a(M))^\perp = \left(\bigcap_{j=1}^{m} T_a(M_j)\right)^\perp = \sum_{j=1}^{m} (T_a(M_j))^\perp = \sum_{j=1}^{m} [\nabla g_j(a)]_\lambda.$$

A protože $\nabla f(a) \in (T_a(M))^\perp$, existují čísla $\lambda_1, \dots \lambda_m \in \mathbb{R}$ tak, že

$$\nabla f(a) = \sum_{j=1}^{m} \lambda_j \nabla g_j(a),$$

neboli

$$\nabla_x L(a; \lambda) = \nabla f(a) - \sum_{j=1}^{m} \lambda_j \nabla g_j(a) = 0.$$

□

Pokud dopředu víme, že funkce $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ má v množině přípustných řešení M lokální extrém, tj. např. tehdy, když je M kompaktní množina a f spojitá funkce (Věta 2.95), můžeme tyto lokální extrémy v M hledat jen na základě nutné podmínky vázaného extrému. Podle Věty 3.64 najdeme tyto extrémy mezi řešeními rovnic

$$\begin{aligned} g(x) &= 0, \\ \nabla_x L(x; \lambda) &= 0, \end{aligned} \tag{30}$$

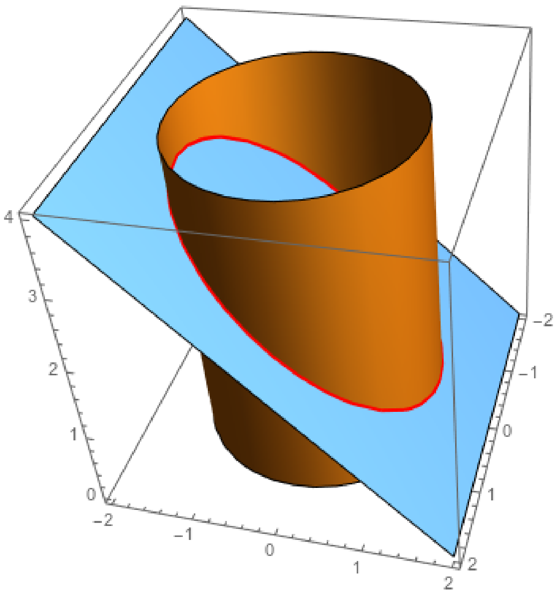
což je $m + n$ rovnic pro $m + n$ neznámých $x_1, \dots, x_n$ a $\lambda_1, \dots, \lambda_m$ za předpokladu, že $m < n$ a matice Dg má v nalezených řešení plnou hodnost. Samozřejmě najít řešení rovnic (30) nemusí být jednoduché a většinou ani explicitně možné. Následující jednoduchý příklad ukazuje, kdy je tento postup možné aplikovat.

Příklad 3.65: Najdeme vázané extrémy funkce $f(x, y, z) = xy + yz$ vzhledem k množině určené vazbami

$$x^2 + y^2 = 2 \quad \text{a} \quad y + z = 2.$$

Použijeme-li naše značení je $g : \mathbb{R}^3 \to \mathbb{R}^2$, kde

$$g_1(x, y, z) := x^2 + y^2 - 2 \quad \text{a} \quad g_2(x, y, z) := y + z - 2.$$



Obrázek 11: Ilustrace množiny M z Příkladu 3.65

Množina $M := \{(x, y, z) \in \mathbb{R}^3 \mid g(x, y, z) = 0\}$, která představuje průnik válce a roviny, viz Obr. 11, je kompaktní, protože je uzavřená a omezená. Proto dopředu víme, že spojitá funkce f má v množině M dokonce globální minimum i maximum.

Všimněte si, že pro všechna $(x, y, z) \in M$ je

$$h\left(Dg(x, y, z)\right) = h\left(\begin{pmatrix} 2x & 2y & 0 \\ 0 & 1 & 1 \end{pmatrix}\right) = 2,$$

neboť body $(0, 0, z) \notin M$ pro všechna $z \in \mathbb{R}$. Hledané extrémy proto musí ležet mezi řešeními rovnic (30). Lagrangeova funkce je

$$\begin{aligned} L(x, y, z; \lambda, \mu) &:= f(x, y, z) - \lambda g_1(x, y, z) - \mu g_2(x, y, z) \\ &= xy + yz - \lambda(x^2 + y^2 - 2) - \mu(y + z - 2), \end{aligned}$$

a proto rovnice (30) mají tvar

$$\begin{aligned} x^2 + y^2 &= 2, \\ y + z &= 2, \\ \frac{\partial L}{\partial x}(x, y, z; \lambda, \mu) &= y - 2x\lambda = 0, \\ \frac{\partial L}{\partial y}(x, y, z; \lambda, \mu) &= x + z - 2y\lambda - \mu = 0, \\ \frac{\partial L}{\partial z}(x, y, z; \lambda, \mu) &= y - \mu = 0. \end{aligned}$$

Tyto rovnice mají 4 řešení (detaily výpočtu jsou přenechány čtenáři):

$$
\begin{aligned}
(x_1, y_1, z_1, \lambda_1, \mu_1) &= \left(-1, 1, 1, -\frac{1}{2}, 1\right), \\
(x_2, y_2, z_2, \lambda_2, \mu_2) &= \left(1, 1, 1, \frac{1}{2}, 1\right), \\
(x_3, y_3, z_3, \lambda_3, \mu_3) &= \frac{1}{2}\left(-1+\sqrt{3}, -1-\sqrt{3}, 5+\sqrt{3}, -2-\sqrt{3}, -1-\sqrt{3}\right), \\
(x_4, y_4, z_4, \lambda_4, \mu_4) &= \frac{1}{2}\left(-1-\sqrt{3}, -1+\sqrt{3}, 5-\sqrt{3}, -2+\sqrt{3}, -1+\sqrt{3}\right).
\end{aligned}
$$

Nakonec stačí vybrat největší a nejmenší ze čtyř funkčních hodnot

$$
f(x_1, y_1, z_1) = 0, \quad f(x_2, y_2, z_2) = 2,
$$

$$
f(x_3, y_3, z_3) = -\frac{5+3\sqrt{3}}{2}, \quad f(x_4, y_4, z_4) = -\frac{5-3\sqrt{3}}{2},
$$

abychom uzavřeli, že f má maximum vzhledem k M v bodě (x_2, y_2, z_2) s hodnotou 2 a minimum vzhledem k M v bodě (x_3, y_3, z_3) s hodnotou $-(5+3\sqrt{3})/2$.

Pokud ovšem apriori nevíme, jestli vázaný extrém existuje, např. když M není kompaktní, potřebujeme nějakou postačující podmínku vázaného extrému. Jednu takovou podmínku si dokážeme. Budeme k tomu potřebovat mírně rozšířit pojem definitnost matice na *definitnost matice vzhledem k množině.*

Definice 3.66 (PD, ND, PSD, NSD, IND vzhledem k množině)**:** Nechť $\emptyset \neq P \subset \mathbb{R}^n$. Symetrickou matici $A \in \mathbb{R}^{n,n}$ nazveme:

1. *pozitivně definitní (PD) vzhledem k* $P \quad \Leftrightarrow \quad (\forall x \in P \setminus \{0\})(x^T A x > 0)$,
2. *pozitivně semidefinitní (PSD) vzhledem k* $P \quad \Leftrightarrow \quad (\forall x \in P)(x^T A x \geq 0)$ a A není PD vzhledem k P,
3. *negativně definitní (ND) vzhledem k* $P \quad \Leftrightarrow \quad (\forall x \in P \setminus \{0\})(x^T A x < 0)$,
4. *negativně semidefinitní (NSD) vzhledem k* $P \quad \Leftrightarrow \quad (\forall x \in P)(x^T A x \leq 0)$ a A není ND vzhledem k P,
5. *indefinitní (IND) vzhledem k* $P \quad \Leftrightarrow \quad (\exists x, y \in P)(x^T A x > 0 \ \wedge\ y^T A y < 0)$.

V následující větě používáme značení $\nabla_x^2 L(a; \lambda)$ pro Hessovu matici funkce L jakožto funkce proměnných $x_1, \ldots, x_n$, tj.

$$
\nabla_x^2 L(a; \lambda) = \begin{pmatrix}
\frac{\partial^2 L}{\partial x_1^2}(a; \lambda) & \frac{\partial^2 L}{\partial x_2 \partial x_1}(a; \lambda) & \cdots & \frac{\partial^2 L}{\partial x_n \partial x_1}(a; \lambda) \\
\frac{\partial^2 L}{\partial x_1 \partial x_2}(a; \lambda) & \frac{\partial^2 L}{\partial x_2^2}(a; \lambda) & \cdots & \frac{\partial^2 L}{\partial x_n \partial x_2}(a; \lambda) \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial^2 L}{\partial x_1 \partial x_n}(a; \lambda) & \frac{\partial^2 L}{\partial x_2 \partial x_n}(a; \lambda) & \cdots & \frac{\partial^2 L}{\partial x_n^2}(a; \lambda)
\end{pmatrix}.
$$

Věta 3.67 (Postačující podmínka pro vázaný extrém)**:** Předpokládejme, že funkce $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ a $g : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ jsou třídy C^2 na Ω a $M = \{x \in \Omega \mid g(x) = 0\}$. Nechť dále existuje bod $(a, \lambda) \in M \times \mathbb{R}^m$ tak, že $\nabla_x L(a; \lambda) = 0$. Potom platí:

1. Je-li $\nabla_x^2 L(a; \lambda)$ PD vzhledem k $T_a(M)$, potom má f v bodě a ostré lokální minimum vzhledem k M.

2. Je-li $\nabla_x^2 L(a; \lambda)$ ND vzhledem k $T_a(M)$, potom má f v bodě a ostré lokální maximum vzhledem k M.

Důkaz. Stačí dokázat 1. tvrzení, neboť f má v bodě a ostré lokální maximum vzhledem k M, právě když $-f$ má v bodě a ostré lokální minimum vzhledem k M, $\nabla_x^2 L(a; \lambda)$ je ND vzhledem k $T_a(M)$, právě když $\nabla_x^2(-L)(a; \lambda)$ je PD vzhledem k $T_a(M)$ a funkce g a $-g$ určují stejnou množinu přípustných řešení M.

Důkaz tentokrát provedeme sporem. Je-li a izolovaným bodem M, plyne z definice, že f má v a ostré lokální minimum vzhledem k M. Proto budeme nadále předpokládat, že a je hromadný bod M, ve kterém f nemá ostré lokální minimum vzhledem k M. Potom existuje posloupnost $\{x_n\}_{n=1}^{\infty}$ bodů z M taková, že $(\forall n \in \mathbb{N})(x_n \neq a)$, $x_n \to a$ a

$$(\forall n \in \mathbb{N})\,(f(x_n) \leq f(a))\,. \tag{31}$$

Podle Bolzanovy–Weirstrassovy věty 2.105 existuje podposloupnost $\{x_{n_k}\}_{k=1}^{\infty}$ taková, že existuje limita

$$\lim_{k\to\infty} \frac{x_{n_k} - a}{\|x_{n_k} - a\|} =: \eta.$$

Všimněte si, že $\eta \neq 0$, neboť $\|\eta\| = 1$. Bez újmy na obecnosti můžeme navíc předpokládat, že $\{x_{n_k}\}_{k=1}^{\infty} \subset B_a(r) \subset \Omega$ pro nějaké $r > 0$, protože je množina Ω otevřená.

Ukážeme, že $\eta \in T_a(M)$. Podle Věty 3.23 o přírůstku aplikované na funkce g_j existují body $\xi_k^{(g_j)}$ na spojnici bodů a a x_{n_k} tak, že

$$0 = g_j(x_{n_k}) - g_j(a) = \nabla g_j(\xi_k^{(g_j)})(x_{n_k} - a)$$

pro všechna $j \in \hat{m}$ a $k \in \mathbb{N}$. Výraz nalevo je roven nule, protože $a \in M$ a také $x_{n_k} \in M$ pro všechna $k \in \mathbb{N}$. Vydělíme-li poslední rovnost $\|x_{n_k} - a\|$ a pošleme $k \to \infty$, dostaneme

$$0 = \nabla g_j(a)\eta$$

pro každé $j \in \hat{m}$. Neboli η je kolmé na $\nabla g_j(a)$ pro každé $j \in \hat{m}$, a tudíž

$$\eta \in \bigcap_{j=1}^{m} (\nabla g_j(a))^{\perp} = T_a(M).$$

Jelikož $L(y, \lambda) = f(y)$, kdykoliv je $y \in M$ a vzhledem k předpokladu $f, g \in C^2(\Omega)$ dostaneme aplikací Taylorovy věty 3.38 rovnost

$$f(x_{n_k}) - f(a) = L(x_{n_k}; \lambda) - L(a; \lambda) = \nabla_x L(a; \lambda)(x_{n_k} - a) + \frac{1}{2}(x_{n_k} - a)^T \nabla_x^2 L(\xi_k^{(L)}; \lambda)(x_{n_k} - a)$$

pro všechna $k \in \mathbb{N}$, kde $\xi_k^{(L)}$ je bod ležící na spojnici a a x_{n_k}. Využijeme-li ještě předpokladu $\nabla_x L(a;\lambda) = 0$ a nerovnosti (31), vyvodíme, že

$$0 \geq \frac{1}{2}(x_{n_k} - a)^T \nabla_x^2 L(\xi_k^{(L)};\lambda)(x_{n_k} - a)$$

pro všechna $k \in \mathbb{N}$. Nyní opět stačí vydělit obě strany $\|x_{n_k} - a\|^2$ a poslat $k \to \infty$ a dostaneme nerovnost

$$0 \geq \eta^T \nabla_x^2 L(a;\lambda)\eta,$$

což je ovšem spor s tím, že $\nabla_x^2 L(a;\lambda)$ je PD vzhledem k $T_a(M)$. □

Postačující podmínka z Věty 3.67 nám dává návod, jak hledat lokální vázaný extrém dostatečně hladké funkce f za podmínky dané vazbami $g(x) = 0$. Nejprve najdeme všechna řešení (x,λ) rovnic (30), což jsou „body podezřelé z extrému“. Poté testujeme, zda kvadratická forma matice $\nabla_x^2 L(x,\lambda)$ v bodech podezřelých z extrému (x,λ) **zúžená na podprostor** $\ker Dg(x)$, viz Lemma 3.63, je PD, nebo ND. Ukažme si to na jednoduchém příkladě.

Příklad 3.68: Rozhodněte, jaký trojúhelník s daným obvodem $\ell > 0$ má největší obsah. Délky stran trojúhelníka si označme x, y, z. Obsah trojúhelníka S jako funkci délek jeho stran x, y, z lze vyjádřit tzv. *Heronovým vzorcem*:

$$S(x,y,z) = \sqrt{\frac{\ell}{2}\left(\frac{\ell}{2} - x\right)\left(\frac{\ell}{2} - y\right)\left(\frac{\ell}{2} - z\right)}.$$

Chceme tedy najít taková $x, y, z > 0$, pro něž je $S(x,y,z)$ maximální za podmínky, že obvod trojúhelníku je fixní, tj.

$$x + y + z = \ell.$$

Výpočet si můžeme zjednodušit tak, že nebudeme maximalizovat funkci S, nýbrž funkci

$$f(x,y,z) := \frac{2}{\ell} S^2(x,y,z).$$

Tato transformace nás zbaví odmocniny a nepodstatného konstantního faktoru. Navíc protože $f = \phi \circ S$, kde $\phi(t) := 2t^2/\ell$ je rostoucí funkce na $[0,\infty)$, funkce f má extrémy ve stejných bodech jako S (existují-li), jak plyne přímo z definice. Dostáváme tedy optimalizační problém pro objektivní funkci

$$f(x,y,z) = \left(\frac{\ell}{2} - x\right)\left(\frac{\ell}{2} - y\right)\left(\frac{\ell}{2} - z\right)$$

s množinou přípustných řešení

$$M = \{(x,y,z) \in (0,\infty)^3 \mid g(x,y,z) := x + y + z - \ell = 0\}.$$

Lagrangeova funkce má tvar

$$L(x,y,z;\lambda) = \left(\frac{\ell}{2} - x\right)\left(\frac{\ell}{2} - y\right)\left(\frac{\ell}{2} - z\right) - \lambda(x + y + z - \ell).$$

Vázané extrémy musí ležet mezi řešeními systému rovnic

$$\begin{aligned}
\partial_x L(x,y,z;\lambda) &= -\left(\frac{\ell}{2}-y\right)\left(\frac{\ell}{2}-z\right)-\lambda=0,\\
\partial_y L(x,y,z;\lambda) &= -\left(\frac{\ell}{2}-x\right)\left(\frac{\ell}{2}-z\right)-\lambda=0,\\
\partial_z L(x,y,z;\lambda) &= -\left(\frac{\ell}{2}-x\right)\left(\frac{\ell}{2}-y\right)-\lambda=0,\\
x+y+z&=\ell.
\end{aligned}$$

Tento systém má jediné řešení

$$(x,y,z;\lambda)=\left(\frac{\ell}{3},\frac{\ell}{3},\frac{\ell}{3};-\frac{\ell^2}{36}\right).$$

Hessova matice Lagrangeovy funkce v tomto bodě je tvaru

$$\nabla^2_{(x,y,z)}L\left(\frac{\ell}{3},\frac{\ell}{3},\frac{\ell}{3};-\frac{\ell^2}{36}\right)=\frac{\ell}{6}\begin{pmatrix}0&1&1\\1&0&1\\1&1&0\end{pmatrix}.$$

Všimněte si, že tato matice je IND. Nás ale zajímá definitnost této matice vzhledem k podprostoru

$$\begin{aligned}
T_{(\ell/3,\ell/3,\ell/3)}(M)&=\left(\nabla g\left(\frac{\ell}{3},\frac{\ell}{3},\frac{\ell}{3}\right)\right)^{\perp}=((1,1,1))^{\perp}=\left[\begin{pmatrix}1\\-1\\0\end{pmatrix},\begin{pmatrix}1\\0\\-1\end{pmatrix}\right]_{\lambda}\\
&=\left\{\begin{pmatrix}s+t\\-s\\-t\end{pmatrix}\;\middle|\; s,t\in\mathbb{R}\right\}.
\end{aligned}$$

Proto zkoumáme definitnost kvadratické formy

$$(s+t,-s,-t)\nabla^2_{(x,y,z)}L\left(\frac{\ell}{3},\frac{\ell}{3},\frac{\ell}{3};-\frac{\ell^2}{36}\right)\begin{pmatrix}s+t\\-s\\-t\end{pmatrix}=-\frac{\ell}{3}\left(s^2+st+t^2\right)$$

v proměnných s a t. Doplněním formy na čtverce zjistíme, že

$$(s+t,-s,-t)\nabla^2_{(x,y,z)}L\left(\frac{\ell}{3},\frac{\ell}{3},\frac{\ell}{3};-\frac{\ell^2}{36}\right)\begin{pmatrix}s+t\\-s\\-t\end{pmatrix}=-\frac{\ell}{3}\left(\left(s+\frac{t}{2}\right)^2+\frac{3t^2}{4}\right)<0$$

pro všechna $(s,t)\in\mathbb{R}^2\setminus\{(0,0)\}$, a tudíž podle Věty 3.67 je bod $(\ell/3,\ell/3,\ell/3)$ ostré lokální maximum f vzhledem k M.

Otázka ze zadání úlohy se ovšem týká extrému globálního. Argument ukazující, že nalezené lokální maximum je též globálním maximem f vzhledem k M, lze postavit na tom, že $\overline{M}$ je kompaktní (rozmyslete). Uzavíráme tedy, že *mezi všemi trojúhelníky s fixním obvodem ℓ má největší obsah trojúhelník rovnostranný, a to $\ell^2/(12\sqrt{3})$.*

Nakonec si ještě dokážeme analogii nutné podmínky z Věty 3.51 pro případ vázaných extrémů.

Věta 3.69: Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ a $g : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ jsou třídy C^2 na Ω a $M = \{x \in \Omega \mid g(x) = 0\}$. Předpokládejme dále, že $a \in M$ a $h(Dg(a)) = m < n$. Potom platí:

1. Je-li a lokální minimum f vzhledem k M, potom je $\nabla_x^2 L(a; \lambda)$ PD, nebo PSD vzhledem k $T_a(M)$.

2. Je-li a lokální maximum f vzhledem k M, potom je $\nabla_x^2 L(a; \lambda)$ ND, nebo NSD vzhledem k $T_a(M)$.

Zde λ jsou Lagrangeovy multiplikátory z Věty 3.64.

Důkaz. Dokážeme 1. tvrzení. Buď $v \in T_a(M)$. Potom podle Lemma 3.63 existuje $H_0 \subset \mathbb{R}$ a křivka $\phi : H_0 \to M$ taková, že $\phi(0) = a$ a $D\phi(0) = v$. Vzhledem k druhé poznámce za Větou 3.54 o inverzní funkci a zavedení křivky ϕ z důkazu Lemma 3.63 vyplývá z předpokladu $g \in C^2(\Omega)$, že také $\phi \in C^2(H_0)$.

Jelikož má f v bodě a lokální minimum, je bod 0 lokálním minimem funkce $f \circ \phi : H_0 \to \mathbb{R}$. Potom podle Věty 3.51 je $(f \circ \phi)''(0) \geq 0$. Funkci $f \circ \phi$ můžeme derivovat jako složenou funkci a podle řetězového pravidla dostaneme

$$(f \circ \phi)'(t) = \nabla f(\phi(t)) D\phi(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\phi(t))\phi_i'(t)$$

pro všechna $t \in H_0$. Pro druhou derivaci máme

$$(f \circ \phi)''(t) = \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(\phi(t))\phi_i'(t)\phi_j'(t) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\phi(t))\phi_i''(t) \tag{32}$$

pro všechna $t \in H_0$. Využijeme-li toho, že $f(\phi(t)) = L(\phi(t); \lambda)$ pro libovolné $\lambda \in \mathbb{R}^m$, neboť $\phi(t) \in M$ pro všechna $t \in H_0$, a dosadíme $t = 0$ do rovnice (32), dostaneme

$$0 \leq (f \circ \phi)''(0) = \sum_{i,j=1}^{n} \frac{\partial^2 L}{\partial x_j \partial x_i}(a; \lambda) v_i v_j + \sum_{i=1}^{n} \frac{\partial L}{\partial x_i}(a; \lambda)\phi_i''(0). \tag{33}$$

Podle Věty 3.64 existují Lagrangeovy multiplikátory $\lambda \in \mathbb{R}^m$ tak, že $\nabla_x L(a; \lambda) = 0$, a proto je pro tato λ druhý člen v (33) nulový. Nakonec tedy zjišťujeme, že

$$0 \leq \sum_{i,j=1}^{n} \frac{\partial^2 L}{\partial x_j \partial x_i}(a; \lambda) v_i v_j = v^T \nabla_x^2 L(a; \lambda) v,$$

z čehož vyplývá, že matice $\nabla_x^2 L(a; \lambda)$ je PD, nebo PSD vzhledem k $T_a(M)$, neboť $v \in T_a(M)$ bylo voleno libovolně. $\square$

Věta 3.69 má následující okamžitý důsledek pro případ indefinitní Hessovy matice Lagrangeovy funkce vzhledem k tečnému prostoru. U vázaných extrému termín sedlový bod nepoužíváme.

Důsledek 3.70: Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ a $g : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ jsou třídy C^2 na Ω a $M = \{x \in \Omega \mid g(x) = 0\}$. Pokud existuje bod $(a, \lambda) \in M \times \mathbb{R}^m$ takový, že $\nabla_x L(a; \lambda) = 0$, matice $\nabla_x^2 L(a; \lambda)$ je IND vzhledem k $T_a(M)$ a platí $h(Dg(a)) = m < n$, potom a není extrém f vzhledem k M.

3.10 Cvičení

Cvičení 3.1: Dokažte, že funkce $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ je diferencovatelná v $a \in \Omega$, právě když existuje $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, okolí H_a bodu a a funkce $\epsilon_a : H_a \to \mathbb{R}^m$ tak, že

$$(\forall x \in H_a)(f(x) = f(a) + T(x-a) + \epsilon_a(x)\|x-a\|)$$

a $\epsilon_a(x) \to 0$ pro $x \to a$.

Cvičení 3.2: Nechť $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ jsou normované prostory. Ukažte, že množina

$$\mathcal{B}(X,Y) := \{T \in \mathcal{L}(X,Y) \mid \|T\| < \infty\}$$

je lineární prostor a zobrazení $\|\cdot\| : \mathcal{B}(X,Y) \to [0,\infty)$ je norma na $\mathcal{B}(X,Y)$.

Cvičení 3.3: Nechť $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ jsou normované prostory a $T \in \mathcal{L}(X,Y)$. Dokažte rovnosti

$$\sup_{\|x\|_X=1} \|Tx\|_Y = \sup_{\|x\|_X \le 1} \|Tx\|_Y = \sup_{x \ne 0} \frac{\|Tx\|_Y}{\|x\|_X}.$$

Cvičení 3.4: Uvažujme prostor reálných polynomů $\mathcal{P}$ s normou definovanou vztahem $\|p\| := \max_{0 \le j \le n} |\alpha_j|$ pro polynom $p(x) = \sum_{j=0}^n \alpha_j x^j$. Dokažte, že lineární operátor $T \in \mathcal{L}(\mathcal{P})$, který zobrazí polynom na jeho derivaci, $Tp := p'$, není omezený, tj. $\|T\| = \infty$.
(Hint: Uvažujte posloupnost monomů $p_n(x) = x^n$.)

Cvičení 3.5: Dokažte Větu 3.6.

Cvičení 3.6: Ukažte, že funkce $f(x,y) = \sqrt{|xy|}$ není diferencovatelná v počátku.

Cvičení 3.7: Ukažte, že funkce

$$f(x,y) = \begin{cases} \frac{x^3-y^3}{x^2+y^2}, & \text{je-li } (x,y) \ne (0,0), \\ 0, & \text{je-li } (x,y) = (0,0), \end{cases}$$

je spojitá v $(0,0)$, ale není diferencovatelná v $(0,0)$.

Cvičení 3.8: Ukažte, že funkce

$$f(x,y,z) = \begin{cases} \frac{xyz}{(x^2+y^2+z^2)^\alpha}, & \text{je-li } (x,y,z) \ne (0,0,0), \\ 0, & \text{je-li } (x,y,z) = (0,0,0), \end{cases}$$

je diferencovatelná v $(0,0,0)$, právě když $\alpha < 1$.

Cvičení 3.9: Ukažte, že funkce

$$f(x,y) = \begin{cases} \frac{x^3}{x^2+y^2}, & \text{je-li } (x,y) \ne (0,0), \\ 0, & \text{je-li } (x,y) = (0,0), \end{cases}$$

má všechny směrové derivace v bodě $(0,0)$, ale není v bodě $(0,0)$ diferencovatelná.

Cvičení 3.10: Dokažte, že pokud všechny první parciální derivace funkce $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ existují a jsou omezené na okolí bodu $a \in \Omega$, potom je f spojitá v a.

Cvičení 3.11: Funkci $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ nazýváme *konvexní* na Ω, právě když je Ω konvexní množina a platí

$$(\forall x, y \in \Omega, x \neq y)(\forall t \in (0,1))(f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)).$$

Nahradíme-li nerovnost výše ostrou nerovností, nazývá se f *ryze konvexní* na Ω. Dokažte, že každé lokální minimum konvexní funkce f na Ω je globálním minimem f na Ω.

Cvičení 3.12: Dokažte, že má-li ryze konvexní funkce lokální minimum, je toto lokální minimum jejím jediným globálním minimem.

Cvičení 3.13: Nechť $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ a Ω je konvexní. Dokažte následující tvrzení:

1. Je-li $f \in C^1(\Omega)$, potom

$$f \text{ je konvexní} \quad \Leftrightarrow \quad (\forall x, y \in \Omega)\,(f(y) \geq f(x) + \nabla f(x)(y-x))\,.$$

2. Je-li $f \in C^2(\Omega)$, potom

$$f \text{ je konvexní} \quad \Leftrightarrow \quad (\forall x \in \Omega)\,\big(\nabla^2 f(x) \text{ je PD nebo PSD}\big)\,.$$

Rozmyslete varianty těchto ekvivalencí pro ryze konvexní funkci f.

Cvičení 3.14: Nechť $A \in \mathbb{R}^{n,m}$, $m \leq n$ a $h(A) = m$. Dokažte, že matice $A^T A \in \mathbb{R}^{m,m}$ je regulární.

Cvičení 3.15 (Rolleova věta v $\mathbb{R}^n$)**:** Nechť $\Omega \subset \mathbb{R}^n$ je otevřená a omezená množina a $f : \overline{\Omega} \to \mathbb{R}$ je spojitá na $\overline{\Omega}$ a diferencovatelná na Ω. Dokažte, že je-li f konstantní na $\partial\Omega$, potom existuje $\xi \in \Omega$ tak, že $\nabla f(\xi) = 0$.
(Hint: Nutná podmínka extrému.)

Cvičení 3.16: Nechť $f : \mathbb{R}^n \to \mathbb{R}$ je diferencovatelná (na celém $\mathbb{R}^n$). Předpokládejme, že f má jediný kritický bod a a v tomto bodě má f lokální minimum.

1. Je-li $n = 1$, potom dokažte, že a je globální minimum f.

2. Je-li $n > 1$, potom ukažte, že předchozí tvrzení není pravdivé.

(Hint: 1. Využijte Rolleovu větu. 2. Uvažujte funkci $f(x,y) = x^2 + y^2(1-x)^3$ a ukažte, že f má v bodě $(0,0)$ lokální minimum, které ale není globální.)

Cvičení 3.17: Ukažte, že funkce

$$f(x) = \begin{cases} x + 2x^2 \sin \frac{1}{x}, & \text{pro } x \neq 0, \\ 0, & \text{pro } x = 0, \end{cases}$$

je diferencovatelná na $\mathbb{R}$, $f'(0) \neq 0$ a že f není prostá na žádném okolí bodu $x = 0$. Všimněte si, že f' není spojitá v počátku.

Cvičení 3.18: Nechť $n \geq 2$, $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ je třídy C^1, $a \in \Omega$ a $\nabla f(a) \neq 0$. Ozn. $M := \{x \in \Omega \mid f(x) = 0\}$. Dokažte, že existuje okolí $H_a \subset \Omega$ tak, že množina $H_a \cap M$ je grafem funkce $n-1$ proměnných třídy C^1.
(Hint: Aplikujte Větu o implicitní funkci.)

Cvičení 3.19: Nechť $\phi : (a,b) \to \mathbb{R}^2$ je třídy C^1, $t_0 \in (a,b)$ a $\phi'(t_0) \neq 0$. Dokažte, že existuje otevřený interval $I \subset (a,b)$, $t_0 \in I$ tak, že množina $\phi(I)$ je grafem C^1 funkce, tzn., že buď $\exists f : (\alpha,\beta) \to \mathbb{R}$ třídy C^1 tak, že $\phi(I) = \{(x, f(x)) \mid x \in (\alpha,\beta)\}$, nebo $\exists g : (\gamma,\delta) \to \mathbb{R}$ třídy C^1 tak, že $\phi(I) = \{(g(y), y) \mid y \in (\gamma,\delta)\}$.
(Hint: Za předpokladu, že $\phi_1'(t_0) \neq 0$, definujte $F(x,t) := x - \phi_1(t)$. Protože $\partial_t F(x_0,t_0) = -\phi_1'(t_0) \neq 0$, kde $x_0 := \phi_1(t_0)$, existují podle Věty o implicitní funkci otevřené intervaly (α,β), I a C^1 bijekce $\omega : (\alpha,\beta) \to I$ tak, že $x = \phi_1(t)$, kde $t = \omega(x)$, pro všechna $x \in (\alpha,\beta)$. Tudíž pro $t \in I$ máme $\phi(t) = (\phi_1(t), \phi_2(t)) = (x, \phi_2(\omega(x))$, a tedy $f := \phi_2 \circ \omega$. Je-li $\phi_2'(t_0) \neq 0$, postupuje se analogicky.)

Cvičení 3.20: Nechť $f : \mathbb{R}^n \to \mathbb{R}^n$ je diferencovatelná v bodě a a prostá na okolí a. Dokažte, že pokud $\det Df(a) = 0$, potom f^{-1} není diferencovatelná v bodě $f(a)$.

Cvičení 3.21: Buď V lineární prostor konečné dimenze se skalárním součinem a E a F podprostory V. Dokažte rovnosti:

1. $(E+F)^\perp = E^\perp \cap F^\perp$,
2. $(E \cap F)^\perp = E^\perp + F^\perp$.

Cvičení* 3.22: Pokuste se zobecnit Příklad 3.68 a dokázat, že mezi všemi jednoduchými (tj. bez samoprůseků) n-úhelníky s fixním obvodem má největší obsah pravidelný n-úhelník. V limitě $n \to \infty$ nám výsledek tohoto cvičení naznačí řešení tvz. *izoperimetrického problému*: Obsah vnitřku jednoduché uzavřené křivky s fixní délkou je maximální v případě kružnice.
(Hint: Označte si souřadnice vrcholů jednoduchého n-úhelníku

$$(x_0,y_0), (x_1,y_0), \ldots, (x_{n-1},y_{n-1}), (x_n,y_n) \equiv (x_0,y_0)$$

a hledejte maximum funkce pro plochu n-úhelníku

$$S = \frac{1}{2}\sum_{i=1}^{n-1}(x_i y_{i+1} - x_{i+1} y_i)$$

za podmínky konstantního obvodu

$$\sum_{i=1}^{n}\sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} = \sum_{i=1}^{n}\ell_i = \ell > 0.$$

Pro zjednodušení řešení je vhodná komplexní formulace získaných rovnic pomocí substituce $z_i := x_i - x_{i-1} + \mathrm{i}(y_i - y_{i-1})$, $i \in \hat{n}$.)

Cvičení 3.23: Určete vzdálenost elipsy a přímky v $\mathbb{R}^2$ určené rovnicemi:

$$x^2 + 4y^2 = 1 \quad \text{a} \quad x + y = 4.$$

Cvičení 3.24: Nechť $a \in \mathbb{R}^n$, $a \neq 0$. Dokažte, že

1. $\|a\|^{-2} = \min\left\{\|x\|^2 \mid a^T x = 1\right\}$,
2. $\|a\| = \max\left\{a^T x \mid \|x\| = 1\right\}$.

Cvičení 3.25:

1. Ukažte, že maximální hodnota funkce $f(x_1, \dots, x_n) = \prod_{i=1}^n x_i^2$ na množině $S = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$ je n^{-n}.
2. Poté dokažte, že pro každé $x \in \mathbb{R}^n$ platí nerovnost:
$$\prod_{i=1}^n |x_i| \leq n^{-n/2}\|x\|^n.$$
3. Nakonec ověřte, že *geometrický průměr* n nezáporných čísel $a_1, \dots, a_n$ je menší nebo roven jejich *aritmetickému průměru*, tj.
$$\left(\prod_{i=1}^n a_i\right)^{1/n} \leq \frac{1}{n}\sum_{i=1}^n a_i.$$

Cvičení 3.26: Nechť $p, q > 1$ a $\frac{1}{p} + \frac{1}{q} = 1$.

1. Ukažte, že pro funkci
$$f(x, y) = \frac{x^p}{p} + \frac{y^q}{q}$$
platí $\min\{f(x, y) \mid x, y > 0 \ \wedge \ xy = 1\} = 1$.
2. Odtud dále odvoďte *Youngovu nerovnost*:
$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$
pro libovolné $a, b \geq 0$.
3. Použijte předchozí bod k důkazu *Hölderovy nerovnosti*:
$$\sum_{i=1}^n a_i b_i \leq \left(\sum_{i=1}^n a_i^p\right)^{1/p} \left(\sum_{i=1}^n b_i^q\right)^{1/q}$$
pro libovolné $a_1, \dots, a_n, b_1, \dots, b_n \geq 0$.

4. Nakonec z Hölderovy nerovnosti odvoďte *Minkowského nerovnost*:

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{1/p}$$

pro libovolné $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}$.

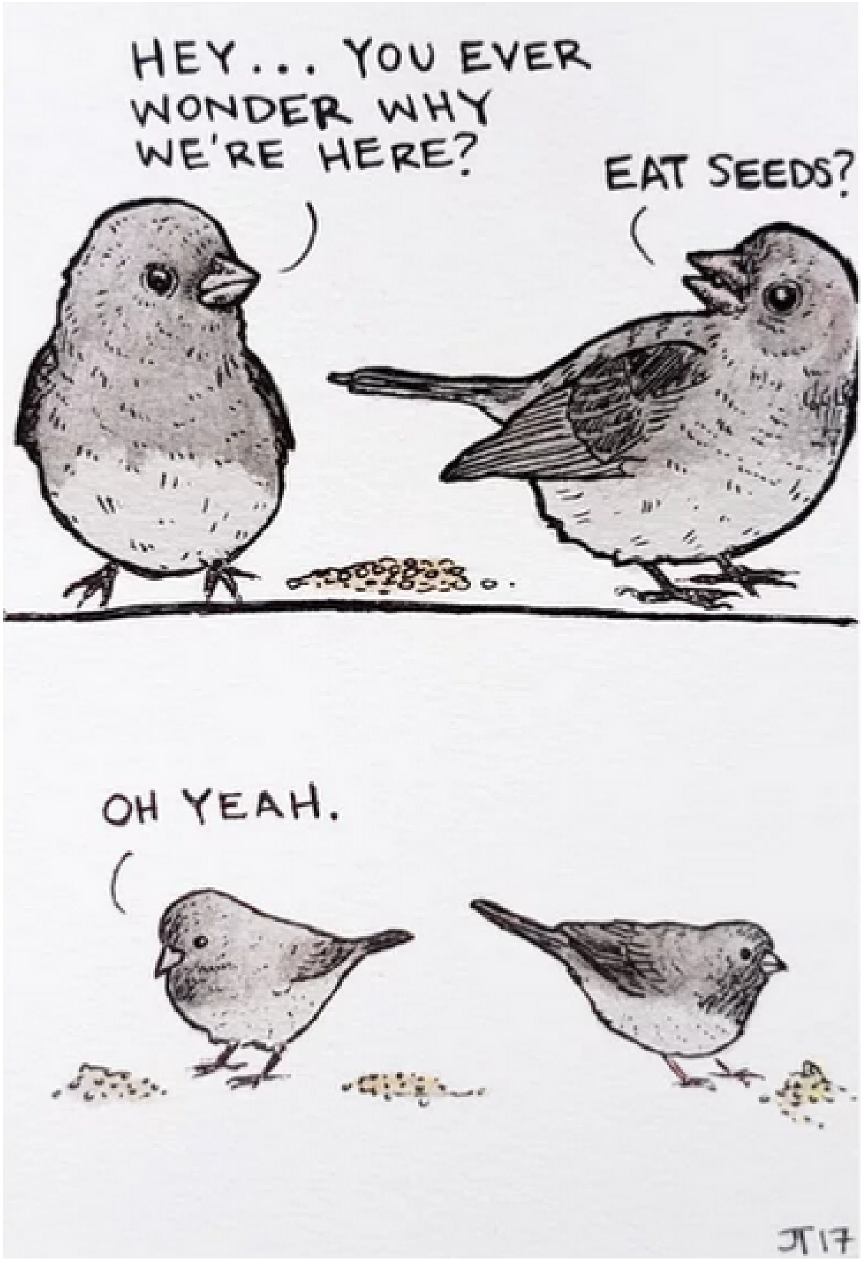
(Hint: Bod 3.: Položte $A := (\sum_{i=1}^n a_i^p)^{1/p}$ a $B := (\sum_{i=1}^n b_i^q)^{1/q}$ a aplikujte 2. bod s $a = a_i/A$ a $b = b_i/B$. Bod 4.: Využijte toho, že $|x+y|^p \leq |x+y||x+y|^{p/q} \leq |x||x+y|^{p/q} + |y||x+y|^{p/q}$.)

Cvičení* 3.27: Nechť $A \in \mathbb{R}^{n,n}$ je PD matice a $a \in \mathbb{R}^n$. Najděte extrémy funkce

$$f(x) = \frac{1}{2} x^T A x - a^T x.$$

Cvičení* 3.28: Maximalizujte funkci $f(X) = (\det X)^2$ na množině matic $X \in \mathbb{R}^{n,n} \simeq \mathbb{R}^{n^2}$ splňující podmínku $\operatorname{Tr} X^T X = 1$.

Cvičení* 3.29: Maximalizujte funkci $f(X) = \operatorname{Tr} X^T X$ na množině matic $X \in \mathbb{R}^{n,n} \simeq \mathbb{R}^{n^2}$ splňující podmínku $\det X = 1$.


HEY... YOU EVER WONDER WHY WE'RE HERE?
EAT SEEDS?
OH YEAH.
JT 17