3. Inverzní a implicitní zobrazení

V této kapitole uvádíme dvě důležité věty, které nacházejí aplikace v mnoha oblastech matematiky: Větu o inverzním a větu o implicitním zobrazení.

3.1. Inverzní zobrazení. V tomto odstavci formulujeme větu o inverzním zobrazení.

Věta 3.1 (o inverzním zobrazení). *Nechť $f : U \subset \boldsymbol{R}^n \to \boldsymbol{R}^n$ je spojité zobrazení, spojitě diferencovatelné v bodě $x_0 \in U$. Jestliže lineární zobrazení $Df(x_0) : \boldsymbol{R}^n \to \boldsymbol{R}^n$ je izomorfismus, pak existuje okolí V bodu x_0 tak, že množina $W = f(V)$ je okolí bodu $y_0 = f(x_0)$ a zúžení $f : V \to W$ je homeomorfismus. Inverzní zobrazení $f^{-1} : W \to V$ je diferencovatelné v bodě y_0. Platí*

$$Df^{-1}(y_0) = \left(Df(x_0)\right)^{-1}. \tag{3.1.1}$$

Lemma 3.2. *Nechť $f : U \subset \boldsymbol{R}^n \to \boldsymbol{R}^n$ je spojité zobrazení, spojitě diferencovatelné v bodě $x_0 \in U$ a takové, že $Df(x_0) = \mathrm{id}_{\boldsymbol{R}^n}$. Pak existuje okolí V bodu x_0 tak, že množina $W = f(V)$ je okolí bodu $y_0 = f(x_0)$ a zúžení $f : V \to W$ je homeomorfismus.*

D ů k a z . Bez újmy na obecnosti předpokládejme, že zobrazení f má Gâteauxovu derivaci v každém bodě množiny U. Položme $g = f - \mathrm{id}_{\boldsymbol{R}^n}$ (čili $g(x) = f(x) - x$ pro každé $x \in U$). Máme

$$Dg(x_0) = Df(x_0) - D\,\mathrm{id}_{\boldsymbol{R}^n}(x_0) = \mathrm{id}_{\boldsymbol{R}^n} - \mathrm{id}_{\boldsymbol{R}^n} = 0 \tag{3.1.2}$$

(viz věty 2.12 a 2.7). Máme tedy

$$\|Dg(x_0)\| = 0. \tag{3.1.3}$$

Podle předpokladu je zobrazení f spojitě diferencovatelné v bodě x_0. Zobrazení g je tedy rovněž spojitě diferencovatelné v bodě x_0, což znamená, že zobrazení $D_G g : U \to L(\boldsymbol{R}^n, \boldsymbol{R}^n)$ je spojité v bodě x_0. Existuje tedy uzavřená koule B se středem v bodě x_0 taková, že pro všechny její prvky x platí

$$\begin{aligned} &\|D_G g(x)\| \le \frac{1}{2}, \\ &\det D_j f^i(x) \ne 0. \end{aligned} \tag{3.1.4}$$

Zvolme nyní dva libovolné body $x_1, x_2 \in B$. Podle důsledku 1 věty o střední hodnotě 2.22 je

$$\|g(x_2) - g(x_1)\| \le \sup_{y \in [x_1, x_2]} \|D_G g(y)\| \cdot \|x_2 - x_1\|, \tag{3.1.5}$$

což společně s (3.1.4) dává

$$\|g(x_2) - g(x_1)\| \le \frac{1}{2}\|x_2 - x_1\|. \tag{3.1.6}$$

Máme tedy

$$\frac{1}{2}\|x_2 - x_1\| \geq \|g(x_2) - g(x_1)\| = \|f(x_2) - x_2 - f(x_1) + x_1\| =$$
$$= \|(x_2 - x_1) - (f(x_2) - f(x_1))\| \geq \|x_2 - x_1\| - \|f(x_2) - f(x_1)\|,$$

neboli

$$2\|f(x_2) - f(x_1)\| \geq \|x_2 - x_1\|, \tag{3.1.7}$$

což dokazuje, že zobrazení f je na množině B injektivní.

Uvažme nyní množinu $f(\mathrm{fr}\,B)$. Tato množina je kompaktní (množina $\mathrm{fr}\,B$ je kompaktní a zobrazení f spojité) a neobsahuje bod y_0 (zobrazení f je na kouli B prosté a $x_0 \notin \mathrm{fr}\,B$). Existuje tedy číslo $d > 0$ takové, že pro každý prvek $y \in f(\mathrm{fr}\,B)$ platí

$$\|y - y_0\| \geq d. \tag{3.1.8}$$

Položme

$$W = B^{d/2}(y_0) \tag{3.1.9}$$

a dokažme, že $W \subset f(\mathrm{int}\,B)$.

Nechť $y_1 \in W$ je libovolný bod. Definujme funkci $h : B \to \boldsymbol{R}$ předpisem

$$h(x) = \|f(x) - y_1\|^2 = \sum_{i=1}^{n}\left(f^i(x) - y_1^i\right)^2. \tag{3.1.10}$$

Tato funkce je spojitá a v každém bodě uvnitř množiny B má Gâteauxovu derivaci. Ze spojitosti této funkce a z kompaktnosti množiny B plyne, že existuje bod $x_1 \in B$, ve kterém funkce f má minimum. Bod x_1 leží určitě uvnitř množiny B: kdyby totiž ležel na její hranici, bylo by $\|f(x_1) - y_1\| \geq d$ (to plyne z (3.1.8)) a $\|f(x_0) - y_1\| < d/2$ (viz (3.1.9)), což by znamenalo, že $h(x_0) < h(x_1)$ a v x_1 funkce h nemá minimum. A jelikož $x_1 \in \mathrm{int}\,B$ a funkce h má v x_1 minimum, musí v tomto bodě mít všechny parciální derivace nulové:[1)] $D_i h(x_1) = 0$ pro každé $i \in \{1, \ldots, n\}$. Jelikož

$$D_j h(x_1) = \sum_{i=1}^{n} 2\left(f^i(x_1) - y_1^i\right) D_j f^i(x_1) \tag{3.1.11}$$

a $\det D_j f^i(x_1) \neq 0$ (3.1.4), dostáváme, že $f(x_1) = y_1$.

Položme $V = f^{-1}(W) \cap \mathrm{int}\,B$. Právě jsme ukázali, že zúžení $f : V \to W$ je bijekce. Zbývá tedy dokázat, že zobrazení $f^{-1} : W \to V$ je spojité. To ovšem snadno plyne ze vztahu

$$\|f^{-1}(y_2) - f^{-1}(y_1)\| \leq 2\|y_2 - y_1\|, \tag{3.1.12}$$

který pro libovolné $y_1, y_2 \in W$ plyne z (3.1.7).

Tím je lemma dokázáno.

Lemma 3.3. *Nechť $f : V \to W$ je bijekce mezi otevřenými množinami v $\boldsymbol{R}^n$, diferencovatelná v bodě $x_0 \in V$ a taková, že $Df(x_0) = \mathrm{id}_{\boldsymbol{R}^n}$. Dále předpokládejme, že zobrazení f^{-1} je spojité v bodě $y_0 = f(x_0)$. Pak zobrazení f^{-1} je diferencovatelné v bodě y_0 a platí*

$$Df^{-1}(y_0) = \mathrm{id}_{\boldsymbol{R}^n}. \tag{3.1.13}$$

D ů k a z . Pro libovolný bod $y \in W$, $y = f(x)$, máme

$$\left\|f^{-1}(y) - f^{-1}(y_0) - \mathrm{id}_{\boldsymbol{R}^n}(y - y_0)\right\| = \|x - x_0 - f(x) + f(x_0)\| =$$
$$= \left\|f(x) - f(x_0) - \mathrm{id}_{\boldsymbol{R}^n}(x - x_0)\right\| = \|x - x_0\| \cdot \varepsilon(x - x_0),$$

kde $\lim_{x \to x_0} \varepsilon(x - x_0) = 0$ (viz větu 2.2 (o diferenciálu)). Dále,

[1)] Toto tvrzení je vcelku zřejmé; dokážeme je však až v kapitole 5.

$$\|x-x_0\|\cdot\varepsilon(x-x_0)=\|f^{-1}(y)-f^{-1}(y_0)\|\cdot\varepsilon\big(f^{-1}(y)-f^{-1}(y_0)\big)=$$
$$=\|y-y_0\|\cdot\frac{\|f^{-1}(y)-f^{-1}(y_0)\|}{\|y-y_0\|}\cdot\varepsilon\big(f^{-1}(y)-f^{-1}(y_0)\big).$$

Jelikož $\lim_{y\to y_0}\varepsilon\big(f^{-1}(y)-f^{-1}(y_0)\big)=0$ (zobrazení f^{-1} je spojité v y_0), stačí ukázat, že výraz $\|f^{-1}(y)-f^{-1}(y_0)\|/\|y-y_0\|$ je na nějakém okolí bodu y_0 ohraničený. Tvrzení pak bude vyplývat z věty 2.2 (o diferenciálu).

Nechť V_0 je okolí bodu x_0 takové, že pro každé $x\in V_0$ platí

$$\frac{\|f(x)-f(x_0)-x+x_0\|}{\|x-x_0\|}<\frac{1}{2}. \tag{3.1.14}$$

Pak

$$\frac{1}{2}>\frac{\|f(x)-f(x_0)-x+x_0\|}{\|x-x_0\|}\geq\frac{\|x-x_0\|}{\|x-x_0\|}-\frac{\|f(x)-f(x_0)\|}{\|x-x_0\|}=1-\frac{\|f(x)-f(x_0)\|}{\|x-x_0\|},$$

neboli

$$\frac{\|f(x)-f(x_0)\|}{\|x-x_0\|}>\frac{1}{2}. \tag{3.1.15}$$

Označme W_0 okolí bodu y_0, pro jehož všechny prvky y platí $f^{-1}(y)\in V_0$. Pro libovolný prvek $y\in W_0$ a $x=f^{-1}(y)$ pak máme

$$\frac{\|f^{-1}(y)-f^{-1}(y_0)\|}{\|y-y_0\|}=\frac{\|x-x_0\|}{\|f(x)-f(x_0)\|}<2. \tag{3.1.16}$$

Tím je lemma dokázáno.

Nyní dokážeme dvě lemmata podobná lemmatům 3.2 a 3.3, avšak s oslabeným předpokladem o derivaci zobrazení f.

Lemma 3.4. *Nechť $f:U\subset \boldsymbol{R}^n\to\boldsymbol{R}^n$ je spojité zobrazení, spojitě diferencovatelné v bodě $x_0\in U$. Jestliže lineární zobrazení $Df(x_0):\boldsymbol{R}^n\to\boldsymbol{R}^n$ je izomorfismus, pak existuje okolí V bodu x_0 tak, že množina $W=f(V)$ je okolím bodu $y_0=f(x_0)$ a zúžení $f:V\to W$ je homeomorfismus.*

D ů k a z . Označme $l=Df(x_0)$, $y_0=f(x_0)$ a položme $\bar f=l^{-1}\circ f$. Máme

$$D\bar f(x_0)=D(l^{-1}\circ f)(x_0)=Dl^{-1}(y_0)\circ Df(x_0)=l^{-1}\circ l=\mathrm{id}_{\boldsymbol{R}^n}.$$

Na zobrazení $\bar f$ tedy lze aplikovat lemma 3.2. Platí $f^{-1}=\bar f^{-1}\circ l^{-1}$.

Lemma 3.5. *Nechť $f:V\to W$ je bijekce mezi otevřenými množinami v $\boldsymbol{R}^n$, diferencovatelná v bodě $x_0\in V$ a taková, že $Df(x_0):\boldsymbol{R}^n\to\boldsymbol{R}^n$ je izomorfismus. Dále předpokládejme, že zobrazení f^{-1} je spojité v bodě $y_0=f(x_0)$. Pak zobrazení f^{-1} je diferencovatelné v bodě y_0 a platí*

$$Df^{-1}(y_0)=\big(Df(x_0)\big)^{-1}. \tag{3.1.17}$$

D ů k a z . Stejně jako v důkazu předchozího lemmatu označme $l=Df(x_0)$ a položme $\bar f=l^{-1}\circ f$. Máme $D\bar f(x_0)=\mathrm{id}_{\boldsymbol{R}^n}$ a podle lemmatu 3.3 $D\bar f^{-1}(y_0)=\mathrm{id}_{\boldsymbol{R}^n}$. Nyní $f^{-1}=\bar f^{-1}\circ l^{-1}$ a

$$Df^{-1}(y_0)=D\big(\bar f^{-1}\circ l^{-1}\big)(y_0)=D\bar f^{-1}(x_0)\circ Dl^{-1}(y_0)=\mathrm{id}_{\boldsymbol{R}^n}\circ l^{-1}=l^{-1}.$$

Tím je lemma dokázáno.

D ů k a z v ě t y 3 . 1. Plyne z předchozích lemmat.

Tvrzení o derivaci inverzního zobrazení (vztah (3.1.1)) lze také dokázat přímo, pomocí věty o derivaci složeného zobrazení 2.5. Platí totiž

$$\mathrm{id}_{\boldsymbol{R}^n} = D\mathrm{id}_{\boldsymbol{R}^n}(x_0) = D(f^{-1} \circ f)(x_0) = Df^{-1}(y_0) \circ Df(x_0).$$

Odtud už vztah (3.1.1) plyne.

Za poznámku jistě také stojí, že dimenze definičního oboru a oboru hodnot zobrazení f ve větě 3.1 musí být stejné; jinak by totiž nemohl platit vztah (3.1.1) (to víme z lineární algebry). Tento závěr lze dalekosáhle zobecnit: snadno lze například ukázat, že žádná spojitě diferencovatelná funkce $f : \boldsymbol{R}^2 \to \boldsymbol{R}$ není bijekce (pokud je $D_1 f(x,y) \neq 0$ pro každé (x,y) z nějaké otevřené množiny V, pak stačí uvážit zobrazení $g(x,y) = (f(x,y), y)$).

Předpokládáme-li ve větě o inverzním zobrazení navíc, že zobrazení f je na množině U diferencovatelné, lze okolí V a W nalézt tak, aby inverzní zobrazení f^{-1} bylo diferencovatelné na okolí W a spojitě diferencovatelné v bodě y_0. Pokuste se to dokázat.

3.2. Věta o implicitním zobrazení. Mějme zobrazení $f : U \subset \boldsymbol{R}^n \times \boldsymbol{R}^m \to \boldsymbol{R}^m$, diferencovatelné v bodě $(x_0, y_0) \in U$ a označme $l_1 : \boldsymbol{R}^n \to \boldsymbol{R}^m$ a $l_2 : \boldsymbol{R}^m \to \boldsymbol{R}^m$ lineární zobrazení, definovaná předpisy

$$\begin{aligned} l_1(h_1) &= Df(x_0, y_0)(h_1, 0), \\ l_2(h_2) &= Df(x_0, y_0)(0, h_2). \end{aligned} \tag{3.2.1}$$

Evidentně platí

$$Df(x_0, y_0) = l_1 + l_2. \tag{3.2.2}$$

Zobrazením l_1 a l_2 se někdy říká *parciální derivace zobrazení f v bodě (x_0, y_0)* (zobecněného typu). Přestože my jsme si tento termín vyhradili pro jiný objekt, vidíme, že s ním velmi úzce souvisí.

Věta 3.6 (o implicitním zobrazení). *Nechť $f : U \subset \boldsymbol{R}^n \times \boldsymbol{R}^m \to \boldsymbol{R}^m$ je spojité zobrazení, spojitě diferencovatelné v bodě $(x_0, y_0) \in U$ takovém, že $f(x_0, y_0)$. Uvažme lineární zobrazení l_1 a l_2 z (3.2.1) a množinu $M = \{(x,y) \in U \mid f(x,y) = 0\}$. Jestliže zobrazení l_2 je izomorfismus, pak existují okolí V_1 bodu x_0, okolí V_2 bodu y_0 a zobrazení $g : V_1 \to V_2$ tak, že množina $M \cap (V_1 \times V_2)$ je rovna grafu zobrazení g.*

Zobrazení g je v bodě x_0 diferencovatelné a platí

$$Dg(x_0) = -l_2^{-1} \circ l_1 \tag{3.2.3}$$

D ů k a z . Uvažme zobrazení $F : U \to \boldsymbol{R}^n \times \boldsymbol{R}^m$, $F(x,y) = (x, f(x,y))$. Toto zobrazení je spojité na U a spojitě diferencovatelné v bodě (x_0, y_0). Pro lineární zobrazení $DF(x_0, y_0)$ a libovolný vektor $(h_1, h_2) \in \boldsymbol{R}^n \times \boldsymbol{R}^m$ platí

$$DF(x_0, y_0)(h_1, h_2) = (h_1, Df(x_0, y_0)(h_1, h_2)) = (h_1, l_1(h_1) + l_2(h_2)),$$

Toto zobrazení je tedy izomorfismus: inverzní zobrazení je totiž dáno předpisem

$$\left(DF(x_0, y_0)\right)^{-1}(u_1, u_2) = \left(u_1, l_2^{-1}\left(u_2 - l_1(u_1)\right)\right). \tag{3.2.4}$$

Podle věty 3.1 tedy existují okolí V bodu (x_0, y_0), okolí W bodu $(x_0, 0) \in \boldsymbol{R}^n \times \boldsymbol{R}^m$ a inverzní zobrazení $F^{-1} : W \to V$. Bez újmy na obecnosti můžeme předpokládat, že množina V je rovna součinu dvou otevřených množin: $V = V_1 \times V_2$.

Nyní pro $x \in V_1$ položme

$$g(x) = \mathrm{pr}_2\left(F^{-1}(x, 0)\right). \tag{3.2.5}$$

Nyní dokážeme, že zobrazení g má požadované vlastnosti. Nechť $(x,y) \in M \cap (V_1 \times V_2)$. Máme $f(x,y) = 0$, neboli $F(x,y) = (x, 0)$, což znamená, že

$$g(x) = \mathrm{pr}_2\left(F^{-1}(x, 0)\right) = \mathrm{pr}_2(x, y) = y.$$

Naopak, pro libovolné $x \in V_1$ máme

$$(x, f(x, g(x))) = F(x, g(x)) = F\big(x, \mathrm{pr}_2\big(F^{-1}(x,0)\big)\big).$$

Přitom $\mathrm{pr}_1\big(F^{-1}(x,0)\big) = x$, což znamená

$$(x, f(x, g(x))) = F\big(F^{-1}(x,0)\big) = (x,0)$$

a celkově

$$f(x, g(x)) = 0\,.$$

Zbývá tedy dokázat vztah (3.2.3). Podle (3.2.4) a (3.2.5) máme

$$Dg(x_0)(u_1) = \mathrm{pr}_2\big(DF^{-1}(x_0,0)(u_1,0)\big) = \mathrm{pr}_2\big(u_1, -l_2^{-1} \circ l_1(u_1)\big) = -l_2^{-1} \circ l_1(u_1).$$

Tím je celé tvrzení dokázáno.

O zobrazení g z předchozí věty se někdy říká, že je *implicitně definováno* zobrazením f.

Podobně jako u věty o inverzním zobrazení, vztah (3.2.3) můžeme dokázat přímo, pomocí vztahu (3.2.1): Víme, že pro každé $x \in V_1$ platí

$$f(x, g(x)) = 0\,. \tag{3.2.6}$$

Derivací této rovnice v bodě x_0 dostaneme

$$l_1 + l_2 \circ Dg(x_0) = 0\,, \tag{3.2.7}$$

což je ekvivalentní s (3.2.3). Tento způsob odvozování derivace implicitního zobrazení se často používá při praktických výpočtech.

Pro $n = m = 1$ lze vztah (3.2.3) přepsat do tvaru

$$g'(x_0) = -\frac{D_1 f(x_0, y_0)}{D_2 f(x_0, y_0)}. \tag{3.2.8}$$

Větu o implicitním zobrazení jsme vlastně dokázali jako poměrně jednoduchý důsledek věty o inverzním zobrazení. Kdybychom ji dokázali bez použití této věty, mohli bychom pak naopak větu o inverzním zobrazení dokázat pomocí věty o implicitním zobrazení, a to pomocí vztahu

$$f(f^{-1}(x)) - x = 0\,. \tag{3.2.9}$$

Dalším důsledkem věty o inverzním zobrazení je následující tvrzení: *Nechť $f : U \subset \boldsymbol{R}^n \times \boldsymbol{R}^m \to \boldsymbol{R}^m$ je spojité zobrazení, spojitě diferencovatelné v bodě $(x_0, y_0) \in U$ a takové, že lineární zobrazení $Df(x_0, y_0)$ je surjektivní. Pak existuje okolí W bodu $f(x_0, y_0)$ a zobrazení $g : W \to \boldsymbol{R}^n \times \boldsymbol{R}^m$, diferencovatelné v bodě $f(x_0, y_0)$ tak, že*

$$f \circ g = \mathrm{id}_W\,. \tag{3.2.10}$$

V důkazu tohoto tvrzení se postupuje stejně jako v důkazu věty o implicitním zobrazení a pak se položí

$$g(y) = F^{-1}(0, y)\,. \tag{3.2.11}$$

Uvedené tvrzení vlastně říká, že zobrazení f má na nějakém okolí bodu $f(x_0, y_0)$ pravou inverzi. Podobné tvrzení lze vyslovit i o inverzi levé (jak?).