Přednáška 10, 24. dubna 2015

Vzpomeňme si na klasickou postačující podmínku pro existenci lokálního extrému funkce $f(x)$ jedné proměnné v bodě $a \in \mathbb{R}$: z jejího Taylorova polynomu stupně 2 (se středem v a)

$$f(a+h) = f(a) + f'(a)h + \tfrac{1}{2}f''(a)h^2 + o(h^2),\ h \to 0\ ,$$

ihned vidíme, že

1. pokud $f'(a) \neq 0$, funkce f nemá v a lokální extrém;
2. pokud $f'(a) = 0$ a $f''(a) > 0$, funkce f má v a ostré lokální minimum a
3. pokud $f'(a) = 0$ a $f''(a) < 0$, funkce f má v a ostré lokální maximum.

Pokud $f'(a) = f''(a) = 0$, nelze bez další analýzy o existenci extrému v a říci nic. Pokud $f'(a) = 0$ (a je „podezřelý“ bod), nemůžeme tedy jen ze samotné hodnoty druhé derivace $f''(a)$ nikdy vydedukovat *neexistenci* lokálního extrému. Jak uvidíme, pro funkce více proměnných je situace jiná. Uvedenou postačující podmínku nyní zobecníme na tyto funkce.

Nejprve ale zavedeme značení a oživíme si pár věcí z lineární algebry. Nechť $A = (a_{i,j}) \in \mathbb{R}^{n\times n}$ je symetrická ($a_{i,j} = a_{j,i}$) reálná $n \times n$ matice. Přiřadíme jí kvadratickou formu (= homogenní polynom stupně 2) o n proměnných

$$P_A(x_1, x_2, \ldots, x_n) = xAx^T = \sum_{i,j=1}^{n} a_{i,j}x_ix_j :\ \mathbb{R}^n \to \mathbb{R}\ ,$$

kde x označuje řádkový vektor $(x_1, x_2, \ldots, x_n)$ a x^T je týž vektor psaný transponovaně ve sloupci. Zřejmě $P_A(\overline{0}) = 0$ a $P_A(tx) = t^2P_A(x)$ (díky homogenitě P) pro každou matici A, vektor $x \in \mathbb{R}^n$ a skalár $t \in \mathbb{R}$. Matice A se nazývá

- *pozitivně (resp. negativně) definitní*, když $P_A(x) > 0$ (resp. $P_A(x) < 0$) pro každý bod $x \in \mathbb{R}^n\backslash\{\overline{0}\}$;
- *pozitivně (resp. negativně) semidefinitní*, když $P_A(x) \geq 0$ (resp. $P_A(x) \leq 0$) pro každý bod $x \in \mathbb{R}^n$ a
- *indefinitní*, není-li ani pozitivně ani negativně semidefinitní, to jest $P_A(x) > 0$ a $P_A(y) < 0$ pro nějaké dva body $x, y \in \mathbb{R}^n$.

Vzhledem ke zmíněné rovnosti $P_A(tx) = t^2 P_A(x)$ určují definitnost A už hodnoty P_A na *jednotkové sféře*

$$S = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$$

(proč?). Jak poznat definitnost A se o něco podrobněji zmíníme později.

Připomeneme nomenklaturu extrémů. Funkce $f:\ U \to \mathbb{R}$, kde $U \subset \mathbb{R}^m$ je okolí bodu a, má v a *ostré lokální minimum*, existuje-li takové $\delta > 0$, že $0 < \|x - a\| < \delta \Rightarrow f(x) > f(a)$. *(Neostré) lokální minimum* v a znamená, že $\|x - a\| < \delta \Rightarrow f(x) \geq f(a)$. Podobně pro ostré lokální maximum a (neostré) lokální maximum. Funkce f *nemá v a lokální extrém*, nemá-li v a ani lokální minimum ani lokální maximum, to jest pro každé $\delta > 0$ existují takové dva body x, y, že $\|x - a\|, \|y - a\| < \delta$ a $f(y) < f(a) < f(x)$. Funkce $f:\ M \to \mathbb{R}$, kde $M \subset \mathbb{R}^m$, *nabývá na množině M maximum v bodě $a \in M$*, když $f(a) \geq f(b)$ pro každý bod $b \in M$. Podobně pro nabývání minima.

V ZS jsme dokázali větu pro funkce jedné proměnné: každá funkce f : $I \to \mathbb{R}$, spojitá na kompaktním intervalu I (tj. $I = [a, b]$, $-\infty < a \leq b < +\infty$), nabývá na I maximum i minimum. Budeme potřebovat ji zobecnit na více proměnných. Množina $M \subset \mathbb{R}^m$ je *omezená*, když existuje takový poloměr $R > 0$, že $M \subset B(\overline{0}, R)$, a je *uzavřená*, když její doplněk $\mathbb{R}^m \backslash M$ je otevřená množina (to jest každý bod b mimo M leží mimo M i s nějakou celou koulí se středem v b). Množina $M \subset \mathbb{R}^m$ je *kompaktní*, když je současně omezená a uzavřená.

Věta (nabývání extrémů na kompaktu). *Nechť je funkce $f:\ M \to \mathbb{R}$, kde $M \subset \mathbb{R}^m$ je neprázdná kompaktní množina, na M spojitá. Pak f nabývá na M minimumm i maximum.*

Důkaz. Zatím bez důkazu. Podáme ho pravděpodobně později v partii o metrických prostorech. $\square$

Například výše definovaná jednotková sféra S je kompaktní podmnožina $\mathbb{R}^n$, a proto každá spojitá funkce $f:\ S \to \mathbb{R}$ nabývá na S minimum i maximum.

Poslední definice před větou o lokálních extrémech: *Hessova matice $H_f(a)$ funkce f v bodě a*, kde f : $U \to \mathbb{R}$ a $U \subset \mathbb{R}^m$ je okolí bodu a a f má na U všechny derivace druhého řádu, je matice zaznamenávající hodnoty těchto derivací:

$$H_f(a) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right)_{i,j=1}^{m}.$$

Podle tvrzení o $\partial_x\partial_y = \partial_y\partial_x$ je pro funkci z $\mathcal{C}^2(U)$ (tj. se spojitými druhými derivacemi na U) její Hessova matice symetrická.

Věta (o lokálních extrémech). *Nechť $f \in \mathcal{C}^2(U)$, kde $U \subset \mathbb{R}^m$ je okolí bodu a. Připomeňme, že (gradient) $\nabla f(a)$ je vektor hodnot prvních derivací a (Hessova matice) $H_f(a)$ je matice hodnot druhých derivací.*

1. *Pokud $\nabla f(a) \neq \overline{0}$, nemá f v a ani neostrý lokální extrém. (Zde stačí předpokládat pouze existenci gradientu $\nabla f(a)$.)*

2. *Pokud $\nabla f(a) = \overline{0}$ a $H_f(a)$ je pozitivně (resp. negativně) definitní, potom má f v a ostré lokální minimum (resp. maximum).*

3. *Pokud $\nabla f(a) = \overline{0}$ a $H_f(a)$ je indefinitní, nemá f v a lokální extrém.*

Důkaz. 1. Pokud $\nabla f(a) \neq \overline{0}$, pak např. $\partial_{x_1} f(a) > 0$ (pro $\partial_{x_1} f(a) < 0$ postupujeme obdobně), a $f(a_1 + h, a_2, \ldots, a_m) = f(a) + \partial_{x_1} f(a)h + o(h)$ pro $h \to 0$. Existuje tedy takové $\delta > 0$, že pro $h \in (-\delta, 0)$ máme $f(a_1 + h, a_2, \ldots, a_m) - f(a) < \frac{1}{2}\partial_{x_1} f(a)h < 0$ a pro $h \in (0, \delta)$ máme $f(a_1 + h, a_2, \ldots, a_m) - f(a) > \frac{1}{2}\partial_{x_1} f(a)h > 0$. Proto funkce f nemá v a ani neostrý lokální extrém.

2. Nyní $\nabla f(a) = \overline{0}$. Kvadratickou formu $xH_f(a)x^T$ označíme jako $P(x)$ a f aproximujeme v okolí a Taylorovým polynomem stupně $n = 2$ (podle věty o Taylorově polynomu). Sčítanec $f(a)$ odpovídající $i = 0$ převedeme vlevo, sčítanec s $i = 1$ zmizí, protože $\nabla f(a) = \overline{0}$. Dále je $P(x)$ homogenní polynom stupně 2. Pro $\|h\| \to 0$ tak máme vyjádření

$$
\begin{aligned}
f(a+h) - f(a) &= \sum_{i=1}^{2} \frac{1}{i!}(h_1\partial_1 + h_2\partial_2 + \cdots + h_m\partial_m)^i f(a) + o(\|h\|^2) \\
&= \frac{1}{2}\sum_{i,j=1}^{m} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)h_i h_j + o(\|h\|^2) \\
&= \frac{1}{2} hH_f(a)h^T + o(\|h\|^2) \\
&= \frac{1}{2} P(h_1, h_2, \ldots, h_m) + o(\|h\|^2) \\
&= \frac{1}{2}\|h\|^2 \Big(P(h_1/\|h\|, h_2/\|h\|, \ldots, h_m/\|h\|) + o(1) \Big) \\
&= \frac{1}{2}\|h\|^2 (P(e(h)) + o(1)) \ ,
\end{aligned}
$$

kde vektor $e(h) = (h_1/\|h\|, h_2/\|h\|, \ldots, h_m/\|h\|)$ leží na jednotkové sféře S. Jak jsme se již zmínili, S je kompaktní podmnožina $\mathbb{R}^m$ a funkce $P(x) : \mathbb{R}^m \to \mathbb{R}$, která je jistě spojitá na celém $\mathbb{R}^m$, na S nabývá minimum a maximum:

$$\mu = P(\alpha) = \min_{\|x\|=1} P(x) \text{ a } M = P(\beta) = \max_{\|x\|=1} P(x)$$

pro nějaké dva vektory $\alpha, \beta \in S$. Pozitivní (resp. negativní) definitnost $H_f(a)$ je ekvivalentní nerovnostem $0 < \mu \leq M$ (resp. $\mu \leq M < 0$) a indefinitnost je ekvivalentní nerovnostem $\mu < 0 < M$.

Je-li $H_f(a)$ pozitivně definitní, máme $P(e) \geq \mu > 0$ pro každý vektor $e \in S$, a tak existuje takové $\delta > 0$, že pro každé h splňující $0 < \|h\| < \delta$ platí

$$f(a+h) - f(a) = \frac{1}{2}\|h\|^2(P(e) + o(1)) > \frac{\|h\|^2}{2} \cdot \frac{\mu}{2} > 0\,.$$

Takže f má v a ostré lokální minimum. Podobně pro negativně definitní $H_f(a)$ dostáváme v a ostré lokální maximum.

3. Když je $H_f(a)$ indefinitní, pak existuje takové $\delta > 0$, že pro každé $t \in (0, \delta)$ máme

$$\begin{aligned} f(a+t\alpha) - f(a) &= \frac{t^2}{2}(P(\alpha) + o(1)) < \frac{t^2}{2} \cdot \frac{\mu}{2} < 0 \text{ a} \\ f(a+t\beta) - f(a) &= \frac{t^2}{2}(P(\beta) + o(1)) > \frac{t^2}{2} \cdot \frac{M}{2} > 0\,. \end{aligned}$$

Takže f nemá v a lokální extrém. □

Poznámky. Podle této věty funkce, která má v každém bodu otevřené množiny U gradient, může mít lokální extrém pouze v bodech, v nichž je gradient nulový. Těmto bodům se říká *stacionární body*. Dostaneme je jako řešení rovnice $\nabla f(a) = 0$. Když je matice $H_f(a)$ semidefinitní, neříká věta nic, funkce může mít v a extrém a nemusí. Konečně zdůrazněme, že se věta týká otevřených množin U, respektive vnitřních bodů množin. Pokud $f : M \to \mathbb{R}$, kde $M \subset \mathbb{R}^m$, a bod $a \in M$ není vnitřním bodem M, to jest neleží v M spolu s nějakým svým okolím, pak stále může f mít v a lokální extrém vzhledem k M, i když je gradient $\nabla f(a)$ nenulový vektor. Lokálními extrémy v hraničních bodech množin se budeme zabývat později, v partii o Lagrangeových multiplikátorech.

Poznámky o definitnosti matic. Nechť $A = (a_{i,j}) \in \mathbb{R}^{n \times n}$ je symetrická matice a $P = P(x_1, x_2, \ldots, x_n) = \sum_{i,j=1}^{n} a_{i,j} x_i x_j$ je jí odpovídající kvadratická forma. Z lineární algebry víme, že existuje taková regulární matice $B = (b_{i,j}) \in \mathbb{R}^{n\times n}$, že po změně proměnných $x^T = By^T$ přejde P do tvaru

$$P(x_1, x_2, \ldots, x_n) = P(b_{1,1}y_1+\cdots+b_{1,n}y_n, \ldots, b_{n,1}y_1+\cdots+b_{n,n}y_n) = \sum_{i=1}^{n} d_i y_i^2 \ ,$$

kde $d_i \in \{-1, 0, 1\}$. Ekvivalentně,

$$P(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} d_i(c_{i,1}x_1 + c_{i,2}x_2 + \cdots + c_{i,n}x_n)^2 \ ,$$

kde $C = (c_{i,j}) \in \mathbb{R}^{n\times n}$ je inverzní k B. Počty koeficientů d_i rovných $-1, 0$ a 1 v tomto vyjádření jsou určené jednoznačně maticí A (nezávisejí na změně proměnných B) a udávají tzv. signaturu kvadratické formy P. Jsou-li všechny d_i rovny 1 (resp. -1), je A pozitivně (resp. negativně) definitní. Je-li některé d_i rovno 1 a jiné -1, je A indefinitní. Jsou-li všechny d_i rovny 1 nebo 0, je A pozitivně semidefinitní, a ve zbývajícím případě -1 a 0 je A negativně semidefinitní. Do tohoto tvaru $P =$ součet $\pm$ čtverců lze P při malém počtu proměnných $n = 2$ či $n = 3$ transformovat snadno ručně, viz počítání v následujícím příkladu.

Připomeňme ještě *Sylvestrovo kritérium definitnosti* z lineární algebry: pokud jsou všechny subdeterminanty $d_m = \det(a_{i,j})_{i,j=1}^{m}$, $1 \le m \le n$, nenulové, pak, jsou-li všechny kladné, je matice A pozitivně definitní, nastává-li $(-1)^m d_m > 0$, $1 \le m \le n$, je A negativně definitní, a jinak je indefinitní; o případu, kdy $d_m = 0$ pro alespoň jedno m, Sylvestrovo kritérium neříká nic.

Příklad. Nalezněte lokální a globální extrémy funkce

$$f:\ \mathbb{R}^2 \to \mathbb{R},\ f(x, y) = y^2 + y\cos x - \sin x - 2\ .$$

Řešení (nebylo uvedeno na přednášce). Definiční obor $\mathbb{R}^2$ je otevřená množina, a pro hledání lokálních extrémů tak můžeme bez problémů použít větu o lokálních extrémech. Máme

$$\nabla f(x, y) = (\partial_x f, \partial_y f) = (-y\sin x - \cos x, 2y + \cos x)$$

a

$$H_f(x, y) = \begin{pmatrix} \partial^2_{xx} f & \partial^2_{xy} f \\ \partial^2_{yx} f & \partial^2_{yy} f \end{pmatrix} = \begin{pmatrix} -y\cos x + \sin x & -\sin x \\ -\sin x & 2 \end{pmatrix} .$$

Vyřešíme soustavu rovnic $\nabla f(x,y) = (0,0)$:

$$-y\sin x - \cos x = 0,\ 2y + \cos x = 0\ .$$

Sečtením rovnic obdržíme $y(2 - \sin x) = 0$. Nutně (neboť $|\sin x| \le 1$) $y = 0$, a tedy $\cos x = 0$. Dostáváme stacionární body

$$s_k = (\pi/2 + k\pi, 0),\ k \in \mathbb{Z}\ .$$

Podle věty o lokálních extrémech má f lokální extrémy pouze v těchto bodech.

Máme

$$H_f(s_k) = \begin{pmatrix} (-1)^k & (-1)^{k+1} \\ (-1)^{k+1} & 2 \end{pmatrix}\ ,$$

to jest

$$H_f(s_k) = \begin{pmatrix} -1 & 1 \\ 1 & 2 \end{pmatrix} \text{ pro liché } k \text{ a } H_f(s_k) = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \text{ pro sudé } k\ .$$

První matice je indefinitní, protože

$$P(x,y) = -x^2 + 2xy + 2y^2 = -(x-y)^2 + 3y^2,$$

a druhá je pozitivně definitní, protože

$$P(x,y) = x^2 - 2xy + 2y^2 = (x-y)^2 + y^2\ .$$

Pro liché k není v s_k lokální extrém a pro sudé k je v s_k ostré lokální minimum, vždy s hodnotou

$$f(s_{2k}) = -3\ .$$

Jediné lokální extrémy funkce f tedy jsou tato ostrá lokální minima.

Globální maximum neexistuje, protože f je shora neomezená: $f(\pi/2, y) = y^2 - 3$. Jiný důvod je ten, že f nemá žádné lokální maximum (a globální maximum by muselo být i lokálním maximem). Nalezneme globální minima. Definiční obor $\mathbb{R}^2$ není kompaktní (není omezený), nelze hned použít větu o extrémech spojitých funkcí na kompaktech. Funkce f je však 2π-periodická v x (tj. $f(x \pm 2\pi, y) = f(x,y)$ pro každé $x, y \in \mathbb{R}$) a pro vyšetření globálních minim stačí uvážit její hodnoty ve svislém nekonečném pásu

$$P = \{(x,y) \in \mathbb{R}^2 \mid 0 \le x \le 2\pi, y \in \mathbb{R}\}\ .$$

Na jeho hranici máme

$$f(0, y) = f(2\pi, y) = y^2 + y - 2 = \left(y + \frac{1}{2}\right)^2 - \frac{9}{4} \geq -\frac{9}{4} > -3 \, .$$

Ještě ale nejsme hotovi. I když hodnoty f na hranici pásu P jsou větší než -3, pás sám je nekompaktní a pro $y \to \pm\infty$ by někde uprostřed něj mohla f klesat do hodnot menších než -3, třeba do $-\infty$, a globální minimum by nemuselo existovat. Jednoduchý odhad však ukazuje, že f se tak nechová. Pro $|y| \geq 2$ a libovolné $x \in \mathbb{R}$ máme

$$f(x, y) \geq y^2 - |y| - 3 = \left(y \pm \frac{1}{2}\right)^2 - \frac{13}{4} \geq -1 > -3 \, .$$

Když tedy pás P rozložíme na disjunktní sjednocení

$$P = P_1 \cup P_2 \, ,$$

kde $P_1 = [0, 2\pi] \times [-2, 2]$ je kompaktní obdélník a P_2 je nekompaktní zbytek, pro každé $a \in P_2$ platí $f(a) \geq -1 > f(s_0) = -3$, přičemž $s_0 \in P_1$. Všechny hodnoty f na P_2 jsou větší než hodnota f v bodu s_0. Na hranici obdélníka P_1 má f vždy hodnotu alespoň $\min(-9/4, -1) = -9/4 > -3$ a na jeho vnitřku, což je otevřená množina, má f jediné lokální minimum $f(s_0) = -3$. Proto má f na obdélníku P_1 i na celém pásu P jediné ostré globální minimum $f(s_0) = -3$. Z 2π-periodičnosti v proměnné x plyne, že hodnoty $f(s_{2k}) = -3$, $k \in \mathbb{Z}$, jsou právě všechna neostrá globální minima funkce f na $\mathbb{R}^2$.

Závěr. Jediné lokální extrémy f jsou ostrá lokální minima $f(s_{2k}) = -3$ v (nekonečně mnoha) bodech $s_{2k} = (\frac{\pi}{2} + 2k\pi, 0), k \in \mathbb{Z}$. Tyto body jsou i body neostrého globálního minima funkce f. Globální maximum f nemá. □

Implicitní funkce. Jak víme z lineární algebry, soustava n lineárních rovnic o n neznámých $a_{i,1}y_1 + a_{i,2}y_2 + \cdots + a_{i,n}y_n + b_i = 0$, $i = 1, 2, \ldots, n$, kde $a_{i,j}, b_i \in \mathbb{R}$ jsou dané a $\det(a_{i,j})_{i,j=1}^n \neq 0$, má pro každou volbu n konstant b_i jednoznačné řešení $y_1, y_2, \ldots, y_n$. Navíc toto řešení y_j je jakožto funkce volených konstant b_i homogenní lineární funkce: $y_j(b_1, b_2, \ldots, b_n) = c_{j,1}b_1 + c_{j,2}b_2 + \cdots + c_{j,n}b_n$, $j = 1, 2, \ldots, n$, pro jisté konstanty $c_{j,i} \in \mathbb{R}$ (to plyne z Cramerova vzorce vyjadřujícího řešení nehomogenní lineární soustavy ve tvaru podílu dvou determinantů).

Tento výsledek nyní zobecníme na situaci, kdy jsou lineární funkce nahrazeny obecnými funkcemi a kdy v každé rovnici lze předem zvolit více než jeden parametr. Vezmeme soustavu n rovnic o $m+n$ neznámých

$$\begin{aligned} F_1(x_1,\dots,x_m,y_1,\dots,y_n) &= 0 \\ F_2(x_1,\dots,x_m,y_1,\dots,y_n) &= 0 \\ &\vdots \\ F_n(x_1,\dots,x_m,y_1,\dots,y_n) &= 0 \,, \end{aligned}$$

kde F_i jsou reálné funkce definované na okolí bodu (x_0,y_0) v $\mathbb{R}^{m+n}$, kde $x_0 \in \mathbb{R}^m$ a $y_0 \in \mathbb{R}^n$, který je řešením této soustavy, to jest $F_1(x_0,y_0) = F_2(x_0,y_0) = \cdots = F_n(x_0,y_0) = 0$. Jak uvidíme, za jistých předpokladů lze neznámé $y_1, y_2, \dots, y_n$ ze soustavy eliminovat a vyjádřit je, lokálně v okolí x_0, jako funkce $y_i = f_i(x_1, x_2, \dots, x_m)$ neznámých $x_1, x_2, \dots, x_m$. Nejprve ale zavedeme značení. Pro zobrazení $F = (F_1, F_2, \dots, F_n)$ a $f = (f_1, f_2, \dots, f_n)$, přičemž $F_i = F_i(x_1,\dots,x_m,y_1,\dots,y_n)$ a $f_j = f_j(x_1,\dots,x_m)$, označíme $x = (x_1, x_2, \dots, x_m)$, $y = (y_1, y_2, \dots, y_n)$ a

$$F'_x(x,y) = \left(\frac{\partial F_i}{\partial x_j}\right)_{i,j=1}^{n,m}(x,y) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots & \frac{\partial F_1}{\partial x_m} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial F_n}{\partial x_1} & \frac{\partial F_n}{\partial x_2} & \cdots & \frac{\partial F_n}{\partial x_m} \end{pmatrix}(x,y)$$

$$F'_y(x,y) = \left(\frac{\partial F_i}{\partial y_j}\right)_{i,j=1}^{n}(x,y) = \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \frac{\partial F_1}{\partial y_2} & \cdots & \frac{\partial F_1}{\partial y_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial F_n}{\partial y_1} & \frac{\partial F_n}{\partial y_2} & \cdots & \frac{\partial F_n}{\partial y_n} \end{pmatrix}(x,y)$$

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

První a třetí matice mají rozměr $n \times m$, druhá matice je čtvercová s rozměrem $n \times n$.

Věta (o implicitních funkcích). *Nechť*

$$F = (F_1, F_2, \dots, F_n) : \; W \to \mathbb{R}^n$$

je zobrazení definované na okolí $W \subset \mathbb{R}^{m+n}$ bodu (x_0, y_0), kde $x_0 \in \mathbb{R}^m$ a $y_0 \in \mathbb{R}^n$, které splňuje následující podmínky.

1. $F_i = F_i(x, y) \in \mathcal{C}^1(W)$ *pro* $1 \le i \le n$.

2. $F_i(x_0, y_0) = 0$ *pro* $1 \le i \le n$.

3. $\det(F'_y(x_0, y_0)) \neq 0$.

Potom existují okolí $U \subset \mathbb{R}^m$ *a* $V \subset \mathbb{R}^n$ *bodů* x_0 *a* y_0 *taková, že* $U \times V \subset W$ *a pro každý bod* $x \in U$ *existuje právě jeden bod* $y \in V$ *splňující* $F_i(x, y) = 0$ *pro* $1 \le i \le n$. *Jinak řečeno, existuje zobrazení* $f = (f_1, f_2, \ldots, f_n) : U \to V$ *takové, že*

$$\forall (x, y) \in U \times V : \ F(x, y) = \overline{0} \iff y = f(x) \ .$$

Navíc každá funkce f_i *je v* $\mathcal{C}^1(U)$, *takže zobrazení* f *je diferencovatelné na* U *a jeho Jacobiho matice* $f'(x)$ *v bodě* $x \in U$ *splňuje*

$$f'(x) = -(F'_y(x, f(x)))^{-1} \cdot F'_x(x, f(x)) \ .$$

Důkaz této věty dělat nebudeme. Naznačíme ale, jak ze vztahů

$$F_k(x, f_1(x), \ldots, f_n(x)) = 0, \ 1 \le k \le n \ \text{ a } \ x \in U \ ,$$

a z $f_i \in \mathcal{C}^1(U)$ plyne hořejší formule pro $f'(x)$ a také praktičtější explicitní formule pro $\partial_i f_j(x)$. Parciálním derivováním těchto n rovnic podle proměnné x_i dostáváme n vztahů

$$\frac{\partial F_k}{\partial x_i}(x, f(x)) + \sum_{j=1}^{n} \frac{\partial F_k}{\partial y_j}(x, f(x)) \cdot \frac{\partial f_j}{\partial x_i}(x) = 0, \ 1 \le k \le n \ .$$

To je soustava n rovnic s n neznámými $\partial_i f_j(x)$, $1 \le j \le n$, kterou zapíšeme maticově jako

$$F'_y \cdot \partial_i f = -\partial_i F \ ,$$

kde $F'_y = F'_y(x, f(x))$, $\partial_i F$ je sloupcový vektor $(\partial_{x_i} F_1, \partial_{x_i} F_2, \ldots, \partial_{x_i} F_n)^T$, $\partial_i f$ je analogický sloupcový vektor pro f a argumenty parciálních derivací $x, f(x)$ a x pro stručnost vynecháváme. Odtud už pomocí lineární algebry plynou vztahy

$$f'(x) = -(F'_y(x, f(x)))^{-1} \cdot F'_x(x, f(x))$$

a

$$\frac{\partial f_j}{\partial x_i} = -\frac{\det(\partial_{y_1} F, \ldots, \partial_{y_{j-1}} F, \partial_{x_i} F, \partial_{y_{j+1}} F, \ldots, \partial_{y_n} F)}{\det(\partial_{y_1} F, \partial_{y_2} F, \ldots, \partial_{y_n} F)}$$

(v bodech $x \in U$ a $(x, f(x)) \in U \times V$).