9. přednáška 26. listopadu 2007

Věta 2.11. *Nechť $f \in C^2(U)$, kde $U \subset \mathbf{R}^m$ je otevřená množina, a $a \in U$ je bod.*

- *Pokud $\nabla f(a) \neq \overline{0}$, nemá f v a ani neostrý lokální extrém.*
- *Pokud $\nabla f(a) = \overline{0}$ a $H_f(a)$ je pozitivně (negativně) definitní, potom má f v a ostré lokální minimum (maximum).*
- *Pokud $\nabla f(a) = \overline{0}$ a $H_f(a)$ je indefinitní, nemá f v a ani neostrý 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 $f(a_1+h, a_2, \ldots, a_m) = f(a)+\partial_{x_1} f(a)h+o(h)$. 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 $f(a_1 + h, a_2, \ldots, a_m) - f(a) > \frac{1}{2}\partial_{x_1} f(a)h > 0$. Funkce f nemá v a ani neostrý lokální extém.

2 a 3. Nyní $\nabla f(a) = \overline{0}$. Kvadratickou formu $xH_f(a)x^T$ označíme jako $P(x)$ a f rozvineme v okolí a do Taylorova rozvoje řádu $n = 2$ (tvrzení 2.10). Sčítanec $f(a)$ odpovídající $i = 0$ převedeme vlevo, sčítanec s $i = 1$ zmizí, protože $\nabla f(a) = \overline{0}$. $P(x)$ je homogenní polynom stupně 2, takže

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

kde vektor $e = e(h) = (h_1/\|h\|, h_2/\|h\|, \ldots, h_m/\|h\|)$ leží na jednotkové sféře $S = \{x \in \mathbf{R}^m \mid \|x\| = 1\}$. S je kompaktní podmnožina $\mathbf{R}^m$ (je uzavřená a omezená) a spojitá funkce $P(x)$ na ní proto nabývá minima a maxima:

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

pro nějaké vektory α a β z S. Pozitivní (negativní) definitnost $H_f(a)$ je ekvivalentní nerovnostem $0 < \mu \leq M$ ($\mu \leq M < 0$) a indefinitnost je ekvivalentní $\mu < 0 < M$.

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

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

—f má v a ostré lokální minimum. Analogicky pro negativně definitní $H_f(a)$ dostáváme ostré lokální maximum. Když je $H_f(a)$ indefinitní, pak existuje $\delta > 0$ takové, že pro každé $t \in (0, \delta)$ máme

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

—f nemá v a ani neostrý lokální extrém. □

Důležité 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 nebo nemusí. Konečně zdůrazněme, že se věta týká otevřených množin U, respektive vnitřních bodů a množiny U. Pokud je bod a v U ale není jejím vnitřním bodem, pak může f mít v a lokální extrém vzhledem k U, i když je $\nabla f(a)$ nenulový. Lokálními extrémy v hraničních bodech množin se budeme zabývat později (v partii o Lagrangeových multiplikátorech).

Příklad. Nalezněte lokální a globální extrémy funkce

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

Definiční obor $\mathbf{R}^2$ je otevřená množina, pro hledání lokálních extrémů můžeme bez problémů použít větu 2.11. Máme

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

a

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

Soustava rovnic $\nabla f(x,y) = (0,0)$ se snadno vyřeší a dává stacionární body

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

Tedy

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

a

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

První matice je indefinitní,

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

a druhá je pozitivně definitní,

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

Pro liché k v s_k není 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 minimum. Definiční obor $\mathbf{R}^2$ není kompaktní, nelze hned použít větu o extrémech spojitých funkcí na kompaktech. Funkce f je však 2π-periodická v x a pro vyšetření globálních minim stačí uvážit její hodnoty v pásu

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

Na jeho hranici máme

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

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

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

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 $f(a) \ge -1 > f(s_0) = -3$ a $s_0 \in P_1$. Na hranici obdélníka P_1 má f vždy hodnotu alespoň $-9/4 > -3$ a na jeho vnitřku má f jediné lokální minimum $f(s_0) = -3$. Proto má f na obdélníku P_1 a na celém pásu P jediné ostré globální minimum $f(s_0) = -3$. Z 2π-periodičnosti v proměnné x plyne, že hodnoty $f(s_{2k}) = -3$, $k \in \mathbf{Z}$, jsou právě všechna neostrá globální minima funkce f na $\mathbf{R}^2$.

Věta o implicitních funkcích. Uvažujme soustavu n rovnic o $m+n$ neznámých

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

F_i jsou reálné funkce definované na okolí bodu (x_0,y_0) v $\mathbf{R}^{m+n}$, kde x_0 je v $\mathbf{R}^m$ a y_0 v $\mathbf{R}^n$, který je řešením této soustavy, to jest $F_1(x_0,y_0) = F_2(x_0,y_0) = \ldots = F_n(x_0,y_0) = 0$. Nedaly by se neznámé $y_1,\dots,y_n$ ze soustavy eliminovat a nedaly by se vyjádřit, alespoň lokálně v okolí x_0, jako funkce $y_i = f_i(x_1,\dots,x_m)$ neznámých $x_1,\dots,x_m$? Následující věta ukazuje, že jistých předpokladů to možné je.

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 2.12. *Nechť*

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

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

1. *$F_i = F_i(x,y) \in C^1(W)$ 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 $U \subset \mathbf{R}^m$ a $V \subset \mathbf{R}^n$ 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 $1 \leq i \leq n$. 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 $C^1(U)$, 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

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

a z $f_i \in C^1(U)$ 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

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

To je soustava n rovnic s n neznámými $\partial_i f_j(x)$, $1 \leq j \leq n$, 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 $(x, f(x)) \in U \times V$). Podrobnosti viz úloha 1.

Úlohy

1. Rozmyslete si odvození vzorce pro Jacobiho matici implicitních funkcí ve tvaru součinu dvou matic a pro jejich parciální derivace ve tvaru podílu determinantů.