9. přednáška 29. listopadu 2005

Věta 13. *Nechť $f \in C^2(U)$, kde $U \subset \mathbf{R}^m$ je otevřená množina, a $a \in U$ je bod.*

1. *Pokud $\nabla f(a) \neq \overline{0}$, nemá f v a ani neostrý lokální extrém. (Zde samozřejmě stačí místo $f \in C^2(U)$ předpokládat pouze existenci gradientu $\nabla f(a)$.)*

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

3. *Pokud $\nabla f(a) = \overline{0}$ a Hessova matice $H_f(a)$ 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 $\delta > 0$ takové, 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. V dalším předpokládáme, že $\nabla f(a) = \overline{0}$. Kvadratickou formu $xH_f(a)x^T$ označíme jako $P(x)$ a f rozvineme v okolí a do Taylorova rozvoje řádu $n = 2$ (tvrzení 12). Protože $\nabla f(a) = \overline{0}$, sčítanec s $i = 1$ zmizí; $f(a)$ odpovídající $i = 0$ převedeme vlevo. Protože $P(x)$ je homogenní polynom stupně 2, dostáváme vyjádření přírůstku

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

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 $\mathbf{R}^m$ (věta 9 v 1. kapitole) a spojitá funkce $P(x)$ na ní nabývá svého minima a maxima (věta 10 v 1. kapitole):

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

Pozitivní (negativní) definitnost $H_f(a)$ je zřejmě ekvivalentní nerovnostem $0 < \mu \le M$ ($\mu \le M < 0$) a indefinitnost je ekvivalentní $\mu < 0 < M$.

Je-li $H_f(a)$ pozitivně definitní, máme $P(e) \ge \mu > 0$ 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. $\square$

Hledání extrémů funkcí více proměnných. Chceme nalézt lokální i globální extrémy funkce m proměnných $f : D \to \mathbf{R}$ na množině $D \subset \mathbf{R}^m$. Začneme lokálními extrémy a budeme nejprve předpokládat, že množina D je otevřená a $f \in C^2(D)$. Z části 1 věty 13 víme, že všechny lokální (a tedy i globální) extrémy jsou obsaženy v množině *stacionárních bodů*

$$S = \{a \in D \mid \nabla f(a) = \overline{0}\}.$$

Nejprve nalezneme S. U stacionárních bodů $a \in S$ s definitní nebo indefinitní maticí $H_f(a)$ známe díky částem 2 a 3 věty 13 povahu lokálního extrému v a. Je-li $H_f(a)$ semidefinitní, neříká věta 13 o chování f v okolí a nic.

O definitnosti, semidefinitnosti či indefinitnosti matice $H_f(a) = (b_{i,j})_{i,j=1}^m$ rozhodneme metodami lineární algebry. Připomeňme Sylvestrovo kritérium: pokud jsou všechny subdeterminanty $d_n = \det(b_{i,j})_{i,j=1}^n$, $1 \le n \le m$, nenulové, pak, jsou-li všechny kladné, je matice $H_f(a)$ pozitivně definitní, nastává-li $(-1)^{n+1}d_n > 0$, $1 \le n \le m$, je $H_f(a)$ negativně definitní a jinak je indefinitní; o případu, kdy $d_n = 0$ pro alespoň jedno n, Sylvestrovo kritérium netvrdí nic. Obecně vždy můžeme matici $H_f(a)$ změnou báze diagonalizovat: nalezneme regulární matici C takovou, že $B = C \cdot H_f(a) \cdot C^T$ má mimo hlavní diagonálu jen nuly. Má-li B na diagonále kladné i záporné prvky, je $H_f(a)$ indefinitní. Jsou-li všechny diagonální prvky B kladné (záporné), je $H_f(a)$ pozitivně (negativně) definitní. Ve zbývajících případech je $H_f(a)$ odpovídajícím způsobem semidefinitní. Pro malé m, např. $m = 2$, je možné přímo vzít kvadratickou formu a doplnit ji na čtverce, viz následující příklad.

Obecný problém lokálních extrémů je ovšem složitější: funkce $f:\ \Omega \to \mathbf{R}$ je sice obvykle definovaná na nějaké otevřené množině $\Omega \subset \mathbf{R}^m$ a $f \in C^2(\Omega)$, ale lokální extrémy hledáme na množině $D \subset \Omega$, která nemusí být otevřená. D může být například uzavřená koule $\{x \in \mathbf{R}^m \mid \|x\| = r\}$. V takovém případě D rozložíme jako $D = V \cup H$, kde V je vnitřek D a H je množina hraničních bodů ležících v D. Množina lokálních extrémů f na D je obsažena ve sjednocení množiny lokálních extrémů f na V a množiny lokálních extrémů f na H. Protože V je otevřená, na první množinu můžeme aplikovat postup podle věty 13 (ovšem v dimenzi $m > 1$ i pro složitou D může být $V = \emptyset$ a nijak si nepomůžeme). Hranice H se často dá definovat pomocí soustavy rovnic jako $H = \{x \in \mathbf{R}^m \mid F_1(x) = F_2(x) = \cdots = F_n(x) = 0\}$, kde F_i jsou "pěkné" funkce. Pro hledání lokálních extrémů na takových množinách H se používá *Lagrangeova metoda* (též *metoda Lagrangeových multiplikátorů*), kterou uvedeme na následující přednášce; jejím základem je věta strukturně analogická větě 13.

Po určení lokálních extrémů zbývá rozhodnout o globálních extrémech. Nejprve triviální ale užitečné pozorování: Protože globální extrém musí být i lokálním extrémem, nemá-li funkce například lokální minimum, nemá ani globální minimum. Pokud je množina D kompaktní, použijeme základní výsledek: spojitá funkce na kompaktní množině nabývá globální maximum i globální minimum. Když D kompaktní není, nemusí globální extrém existovat a musíme si pomoci jinak, třeba rozdělením D na "zvládnutelné" kusy. Pokud například lze D vyjádřit jako $D = D_1 \cup D_2$, kde (i) D_1 je kompaktní a (ii) existuje bod $b \in D_1$ takový, že pro každé $x \in D_2$ máme $f(x) \ge f(b)$, potom f nabývá na D svého minima a $\min_{x \in D} f(x) = \min_{x \in D_1} f(x)$. Tolik v obecnosti a nyní konkrétní příklad.

Příklad (z bonifikačního testu 25.11.2005). Pro funkci

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

nalezněte lokální a globální extrémy.

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 (odpovídá jí kvadratická forma $P(x,y) = -x^2 + 2xy + 2y^2 = -(x-y)^2 + 3y^2)$ 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. Definiční obor $\mathbf{R}^2$ není kompaktní. Funkce f je však 2π-periodická v x a pro vyšetření globálních minim stačí uvážit její hodnoty v pásu P daném nerovnostmi $0 \le x \le 2\pi$. Na jeho hranici máme hodnoty $f(0,y) = f(2\pi, y) = y^2 + y - 2 = (y + \frac{1}{2})^2 - \frac{9}{4} \ge -\frac{9}{4} > -3$.[1] Dále pro $|y| \ge 2$ a libovolné $x \in \mathbf{R}$ máme

[1] V této chvíli ještě 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 asymptotě pod -3; globální minimum by pak neexistovalo. Následující jednoduchý odhad ukazuje, že takto se f nechová.

$f(x,y) \geq y^2 - |y| - 3 = (y \pm \frac{1}{2})^2 - \frac{13}{4} \geq -1 > -3$. Takže, píšeme-li $P = P_1 \cup P_2$, kde P_1 je (kompaktní) obdélník $[0, 2\pi] \times [-2, 2]$ a P_2 je (nekompaktní) zbytek pásu P, pro každé $a \in P_2$ máme $f(a) \geq -1 > f(s_0) = -3$, kde $s_0 \in P_1$. 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 všechna neostrá globální minima funkce f na $\mathbf{R}^2$.

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

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

kde F_i jsou reálné funkce definované na okolí bodu $(x_0, y_0) \in \mathbf{R}^{m+n}$, $x_0 \in \mathbf{R}^m$ a $y_0 \in \mathbf{R}^n$, který je řešením soustavy, tj. $F_i(x_0, y_0) = 0$ pro $1 \leq i \leq n$. Nedaly by se neznámé $y_1, \ldots, y_n$ pomocí soustavy eliminovat a nedaly by se vyjádřit, alespoň lokálně v okolí x_0, jako funkce $y_i = f_i(x_1, \ldots, x_m)$ prvních m neznámých? Následující věta ukazuje, že za určitých předpokladů o funkcích F_i—jsou z třídy C^1 a lineární aproximace jejich y-ových částí v bodě (x_0, y_0) jsou lineárně nezávislé—to možné je. Takto *implicitně* definované funkce f_i jsou také z třídy C^1 a jejich parciální derivace se snadno vypočítají z parciálních derivací funkcí F_i.

Nejprve zavedeme značení. Pro zobrazení $F = (F_1, F_2, \ldots, F_n)$ a $f = (f_1, f_2, \ldots, f_n)$, kde $F_i = F_i(x_1, \ldots, x_m, y_1, \ldots, y_n)$ a $f_j = f_j(x_1, \ldots, x_m)$, označíme $x = (x_1, x_2, \ldots, x_m)$, $y = (y_1, y_2, \ldots, y_n)$ a

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

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

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

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

Věta 14 (o implicitních funkcích). *Nechť*

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

je zobrazení definované na okolí $W \subset \mathbf{R}^{m+n}$ bodu (x_0, y_0), kde $x_0 \in \mathbf{R}^m$ a $y_0 \in \mathbf{R}^n$, a splňující 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. $\det(F'_y(x_0, y_0)) \ne 0$.

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 $1 \le i \le n$. 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 $C^1(U)$, takže zobrazení f je diferencovatelné na U a jeho Jacobiho matice $f'(x)$ v bodě $x \in U$ splňuje

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

Důkaz věty 14 v této přednášce pomineme. (Lze ho začít nejprve případem jedné rovnice $n = 1$ a pak postupovat indukcí podle n nebo je možné hned dokázat obecnou verzi pomocí Banachovy–Picardovy věty o kontrahujícím zobrazení, viz V. A. Zorich, *Mathematical Analysis*, Springer 2004, podkapitola 8.5 ve sv. 1 a podkapitola 10.7 ve sv. 2.) Ukážeme alespoň, 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 $f_i \in C^1(U)$ plyne hořejší formule pro $f'(x)$ a také praktičtější explicitní formule pro $\partial_i f_j(x)$. Parciálním derivováním těchto n rovnic podle x_i, kde $i \in \{1, 2, \dots, m\}$ je pevné, dostáváme n vztahů

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

Máme soustavu n rovnic s n neznámými $\partial_i f_j(x)$, $1 \le j \le n$, kterou zapíšeme maticově v kanonické podobě jako

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

kde $F_y' = F_y'(x, f(x))$, $\partial_i F$ je sloupcový vektor $(\partial_{x_i} F_1, \partial_{x_i} F_2, \dots, \partial_{x_i} F_n)^T$, $\partial_i f$ je analogický sloupcový vektor pro f a argumenty parciálních derivací $x, f(x)$ a x pro stručnost vynecháváme. Vynásobíme-li tento vztah zleva inverzní maticí $(F_y')^{-1}$ a výsledných m rovností odpovídajících $1 \le i \le m$ sloučíme do jedné, dostaneme

$$(\partial_1 f, \dots, \partial_m f) = -(F_y)^{-1} \cdot (\partial_1 F, \dots, \partial_m F),$$

což je přesně rovnice $f'(x) = -(F_y'(x, f(x)))^{-1} F_x'(x, f(x))$. Podle Cramerova pravidla (známého z Lineární algebry) se ale $\partial_i f_j = \partial_i f_j(x)$ také rovná $\det A' / \det A$, kde $A = F_y' = F_y'(x, f(x))$ a A' je modifikovaná matice soustavy, která z A vznikne nahrazením j-tého sloupce matice A sloupcem pravé (rodné) strany $-\partial_i F$. Takže

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

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