10. přednáška 6. prosince 2005

Příklad (z bonifikačního testu 25.11.2005). Rozhodněte, zda soustava rovnic
$$x + y - \sin z = 0 \quad \text{a} \quad -x - y^3 + \mathrm{e}^z - 1 = 0$$
definuje v okolí 0 funkce $y = y(x)$ a $z = z(x)$ splňující $y(0) = z(0) = 0$, které jsou třídy C^1. Pokud ano, spočítejte hodnoty derivací $y'(0)$ a $z'(0)$.

Pro $F_1(x, y, z) = x+y-\sin z$, $F_2(x, y, z) = -x-y^3+\mathrm{e}^z-1$ a $F = (F_1, F_2)$ máme skutečně $F(0, 0, 0) = (0, 0)$ a

$$\begin{aligned}
\det(\partial_y F(0,0,0), \partial_z F(0,0,0)) &= \det\begin{pmatrix} 1 & , & -\cos z \\ -3y^2 & , & \mathrm{e}^z \end{pmatrix}(0,0,0) \\
&= \det\begin{pmatrix} 1 & , & -1 \\ 0 & , & 1 \end{pmatrix} \\
&= 1 \neq 0.
\end{aligned}$$

Předpoklady věty o implicitních funkcích jsou tedy splněny a uvedené funkce $y(x)$ a $z(x)$ jsou na okolí nuly definovány. Protože $\partial_x F(0,0,0) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, podle vztahů uvedených na konci předešlé přednášky máme

$$y'(0) = -\frac{\det\begin{pmatrix} 1 & , & -1 \\ -1 & , & 1 \end{pmatrix}}{1} = 0 \quad \text{a} \quad z'(0) = -\frac{\det\begin{pmatrix} 1 & , & 1 \\ 0 & , & -1 \end{pmatrix}}{1} = 1.$$

Důsledek 15. *Nechť $f : U \to \mathbf{R}^m$, kde $U \subset \mathbf{R}^m$ je okolí bodu x_0, je zobrazení z $C^1(U)$, které má v x_0 nenulový jacobián. Potom existují okolí $U_1 \subset U$ a $V \subset \mathbf{R}^m$ bodů x_0 a $y_0 = f(x_0)$ taková, že $f : U_1 \to V$ je bijekce, inverzní zobrazení $f^{-1} : V \to U_1$ je z $C^1(V)$ a pro každé $x \in U_1$ v bodě $y = f(x) \in V$ máme*
$$\mathrm{D}f^{-1}(y) = (\mathrm{D}f(x))^{-1}.$$
Jacobiho matice zobrazení f^{-1} v bodě y je tedy inverzní k Jacobiho matici zobrazení f v bodě x.

Důkaz. Uvažme zobrazení o $2m$ proměnných
$$F(x, y) = f(x) - y : U \times \mathbf{R}^m \to \mathbf{R}^m,$$

tj. $F = (F_1, \dots, F_m)$ a $F_i(x, y) = f_i(x) - y_i$. Patrně $F_i(x_0, y_0) = 0$, F_i jsou třídy C^1 a Jacobiho matice zobrazení F vzhledem k x-ovým proměnným $x_1, x_2, \dots, x_m$ v bodě x_0 je právě Jacobiho matice f v x_0. Podle věty 14 tedy existují taková okolí $U_2 \subset U$ a $V \subset \mathbf{R}^m$ bodů x_0 a y_0 a takové zobrazení $g = (g_1, \dots, g_m) :\ V \to U_2$, že $g_i \in C^1(V)$ a

$$\forall\ (x, y) \in U_2 \times V :\ F(x, y) = f(x) - y = \overline{0} \iff x = g(y)$$

(speciálně $g(y_0) = x_0$). Takže pro všechny $y \in V$ máme $f(g(y)) = y$ a zobrazení f a g jsou navzájem inverzní. Označme $U_1 = g(V)$. Pak $f :\ U_1 \to V$ je bijekce s inverzem g. Množina U_1 je otevřená, protože je vzorem otevřené množiny V ve spojitém zobrazení f, a je to okolí x_0. Vzorec pro diferenciál inverzního zobrazení plyne ze vztahu ve větě 14 nebo diferencováním složeného zobrazení $f^{-1} \circ f = id$ (věta 8). $\square$

Bijektivní zobrazení mezi dvěma otevřenými podmnožinami $\mathbf{R}^m$, které je třídy C^1 a jeho inverz rovněž, se nazývá *difeomorfismus*. Důsledek 15 tedy praví, že zobrazení třídy C^1 definované na okolí bodu x_0, které má v x_0 nenulový jacobián, je na okolí x_0 lokální difeomorfismus.

Vázané extrémy. Nechť $f, F_1, \dots, F_n :\ U \to \mathbf{R}$, kde $U \subset \mathbf{R}^m$ je otevřená množina, jsou funkce z $C^1(U)$ a $n < m$. Budeme hledat lokální extrémy funkce f na množině

$$H = \{x \in \mathbf{R}^m \mid F_1(x) = F_2(x) = \cdots = F_n(x) = 0\}.$$

Následující důsledek věty o implicitních funkcích udává nutnou podmínku, aby v bodě $a \in H$ funkce f měla lokální *vázaný* extrém, tj. lokální extrém vzhledem k množině H.

Důsledek 16 (Lagrangeovy multiplikátory). *Pokud v popsané situaci má Jacobiho matice zobrazení $F = (F_1, \dots, F_n)$ v bodě $a \in H$ největší možnou hodnost n (to jest, $\nabla F_1(a), \dots, \nabla F_n(a)$ jsou lineárně nezávislé vektory v $\mathbf{R}^m$) a v bodě a má funkce f (ostrý nebo neostrý) lokální extrém vzhledem k množině H, potom existují taková čísla $\lambda_1, \dots, \lambda_n \in \mathbf{R}$ (tzv.* Lagrangeovy multiplikátory*), že*

$$\nabla f(a) - \sum_{i=1}^{n} \lambda_i \nabla F_i(a) = \overline{0},$$

to jest, pro $1 \le j \le m$, $\partial_{x_j} f(a) - \lambda_1 \partial_{x_j} F_1(a) - \cdots - \lambda_n \partial_{x_j} F_n(a) = 0$.

Důkaz. Že uvedená Jacobiho matice má maximální hodnost n také ekvivalentně znamená, že pro nějakých n sloupců má odpovídající čtvercová $n \times n$ podmatice nenulový determinant. Můžeme předpokládat, že to nastává pro posledních n sloupců. Označíme-li tedy

$$x_1, x_2, \ldots, x_m = y_1, y_2, \ldots, y_{m-n}, z_1, z_2, \ldots, z_n,$$

pak $\det(\partial_{z_1} F(a), \ldots, \partial_{z_n} F(a)) \neq 0$. Podle věty o implicitních funkcích existují taková okolí U_1 a V_1 bodů $y_0 = (a_1, \ldots, a_{m-n})$ a $z_0 = (a_{m-n+1}, \ldots, a_m)$ a takové zobrazení $g = (g_1, \ldots, g_n) : U_1 \to V_1$, že pro (y, z) probíhající $U_1 \times V_1$ máme

$$F_i(y, z) = 0 \ \text{ pro } 1 \leq i \leq n \quad \Longleftrightarrow \quad z = g(y)$$

(speciálně $g(y_0) = z_0$). Uvažme nyní funkci

$$h(y) = f(y, g_1(y), \ldots, g_n(y)),$$

která je definovaná na U_1. Protože má v y_0 lokální extrém (nyní už bez vazby), $\nabla h(y_0) = \overline{0}$. Pro $1 \leq i \leq m - n$ to znamená, že

$$\partial_{y_i} f(y_0, g(y_0)) + \sum_{j=1}^{n} \partial_{z_j} f(y_0, g(y_0)) \cdot \partial_{y_i} g_j(y_0) = 0.$$

V zápisu pomocí Jacobiho matic:

$$\begin{aligned} f'_y + f'_z g' &= \overline{0} \\ f'_y - f'_z (F'_z)^{-1} F'_y &= \overline{0} \\ f'_y - \lambda F'_y &= \overline{0} \end{aligned}$$

(v bodech $(y_0, g(y_0)) = (y_0, z_0) = a$ a y_0), kde za g' jsme nejprve dosadili podle vzorce ve větě 14 a pak jsme označili $\lambda = (\lambda_1, \ldots, \lambda_n) = f'_z(y_0, g(y_0)) \cdot F'_z(y_0, g(y_0))^{-1}$. Ovšem z $\lambda = f'_z (F'_z)^{-1}$ plyne, že stejný vztah platí i v z-ových proměnných: $f'_z - \lambda F'_z = \overline{0}$. Celkem

$$f' - \lambda F' = \overline{0},$$

což jsme chtěli dokázat. $\square$

Ekvivalentní formulace podmínky Lagrangeových multiplikátorů je, že $\nabla f(a)$ leží v lineárním obalu vektorů $\mathcal{F} = \{\nabla F_1(a), \ldots, \nabla F_n(a)\}$. Všimněme si, že podmínka je triviálně splněna pokud $\nabla f(a) = \overline{0}$. Uvedeme ještě další ekvivalentní formulaci. Uvažme vektorové podprostory $\mathbf{R}^m$ složené z vektorů kolmých na $\nabla f(a)$, resp. z vektorů kolmých na všechny vektory v $\mathcal{F}$ (předpokládáme, že $\nabla f(a) \neq \overline{0}$ a že vektory v $\mathcal{F}$ jsou lineárně nezávislé):

$$\begin{aligned} TN_a &= \{x \in \mathbf{R}^m \mid \langle \nabla f(a), x \rangle = 0\} \\ TH_a &= \{x \in \mathbf{R}^m \mid \langle \nabla F_1(a), x \rangle = \cdots = \langle \nabla F_n(a), x \rangle = 0\}. \end{aligned}$$

Podprostor TN_a má dimenzi $m - 1$ a TH_a má dimenzi $m - n$. Podmínka Lagrangeových multiplikátorů ekvivalentně praví, že $TH_a \subset TN_a$. (Proč přesně platí ekvivalence $\nabla f(a) \in \mathrm{Lin}(\mathcal{F}) \iff TH_a \subset TN_a$? Vzpomeňte si na ortogonální doplněk v Lineární algebře.) Pomocí implicitních funkcí se dá ukázat, že $a + TH_a$ je tečným afinním podprostorem k ploše $H = \{x \in \mathbf{R}^m \mid F_1(x) = \cdots = F_n(x) = 0\}$ v bodě a. Podobně $a + TN_a$ je tečnou afinní nadrovinou k "vrstevnicové" ploše

$$N = \{x \in \mathbf{R}^m \mid f(x) = f(a)\}$$

v bodě a. Podprostorům TN_a a TH_a říkáme *tečné prostory* (k odpovídajícím plochám v bodě a). Nutná podmínka lokálního vázaného extrému funkce f v bodě $a \in H$ se tedy dá zformulovat takto:

> Tečný prostor TH_a k ploše H v bodě a musí být obsažen v tečném prostoru TN_a k vrstevnicové ploše N funkce f v bodě a, $TH_a \subset TN_a$.

Příklad či spíše ilustrace. Podíváme se na situaci $m = 2$ a $n = 1$. Funkce $f : \mathbf{R}^2 \to \mathbf{R}$ například udává nadmořskou výšku, tj. $f(x)$ je nadmořská výška bodu v terénu se zeměpisnými souřadnicemi x, a křivka $H = \{x \in \mathbf{R}^2 \mid F(x) = 0\}$ je třeba silnice. Vrstevnice $N = \{x \in \mathbf{R}^2 \mid f(x) = f(a) = b\}$, kde $a \in H$, je též rovinná křivka. Nechť $\nabla f(a), \nabla F(a) \neq \overline{0}$. Tečné prostory TH_a a TN_a pak mají dimenzi 1. Předpokládejme, že $TH_a \not\subset TN_a$. Křivky H a N potom lokálně v okolí svého průsečíku a vypadají jako dvě různé přímky p_H a p_N procházející bodem a. Zvětšíme-li trochu nadmořskou výšku na $b' = b + \delta$, $\delta > 0$, vrstevnice se trochu posune ve směru kolmém na N a stane se z ní vrstevnice N'; lokálně přímka $p_{N'}$ vznikne z p_N malým posunem ve směru kolmém na p_N a případným malým pootočením.[1] Každopádně se, pro každé dostatečně malé $\delta > 0$, přímky p_H a $p_{N'}$ a tedy i křivky H a N' opět protnou, jenom se průsečík obou křivek trochu posune po H z a do a'. Totéž nastane, když b trochu zmenšíme na $b' = b - \delta$, k posunům však dojde na opačnou stranu, speciálně nový průsečík a' křivek H a N' se dostane posunutím a po H na opačnou stranu. Funkce f tedy lokálně na H na jedné straně od bodu a nabývá hodnot menších než $b = f(a)$, zatímco na straně druhé nabývá hodnot větších. Takže f nemá v a vzhledem k H lokální extrém a pokud na silnici H zastaví v bodě a auto, v neutrálu se bez ruční brzdy určitě rozjede!

Ještě další ekvivalentní formulace podmínky Lagrangeových multiplikátorů užívá *Lagrangeovu funkci.* V situaci popsané v důsledku 16 tuto funkci $m + n$ proměnných definujeme jako

$$L(x, \lambda) = L(x_1, \dots, x_m, \lambda_1, \dots, \lambda_n) = f(x) - \sum_{i=1}^{n} \lambda_i F_i(x).$$

Protože

$$\nabla L = (\partial_{x_1} f - \textstyle\sum_1^n \lambda_i \partial_{x_1} F_i,\ \dots,\ \partial_{x_m} f - \sum_1^n \lambda_i \partial_{x_m} F_i,\ -F_1,\ \dots,\ -F_n)$$

(v bodech (x, λ) a x), je $\nabla L(a, \lambda) = \overline{0}$ přesně ekvivalentní tomu, že bod a leží na ploše H (posledních n souřadnic gradientu) a že koeficienty $\lambda = (\lambda_1, \dots, \lambda_n)$ jsou Lagrangeovy multiplikátory (prvních m souřadnic gradientu). Nutnou podmínku lokálního extrému funkce f v bodě a vzhledem k H tedy můžeme zformulovat i takto:

Existuje bod $\lambda \in \mathbf{R}^n$ takový, že $\nabla L(a, \lambda) = \overline{0}$.

Zde se o náležení a do H nemusíme starat, protože je v podmínce $\nabla L(a, \lambda) = \overline{0}$ automaticky zahrnuto. Závěrem partie o vázaných extrémech uvedeme bez důkazu větu analogickou větě 13 a pak ji ilustrujeme příkladem.

[1] Je to ale opravdu tak, že malá změna nadmořské výšky jen málo změní vrstevnici? Vždycky to pravda není—představte si, že se nacházíte na rovném horském hřbetu a vrstevnice je hřbetová čára. Jakkoli malé zvětšení nadmořské výšky pak vede k naprosto radikální změně vrstevnice, protože ta úplně zmizí. Jako cvičení vysvětlete, proč se díky předpokladu $\nabla f(a) \neq \overline{0}$ vrstevnice funkce f v okolí a takto nechovají a malá změna b jen (popsaným způsobem) málo změní N.

Věta 17. *Nechť $f, F_1, \dots, F_n \in C^2(U)$, kde $U \subset \mathbf{R}^m$ je otevřená množina a $n < m$, $H = \{x \in \mathbf{R}^m \mid F_1(x) = F_2(x) = \cdots = F_n(x) = 0\}$ a nechť $a \in H$ je bod. Předpokládejme dále, že Jacobiho matice zobrazení $F = (F_1, \dots, F_n)$ má v každém bodě $x \in U$ největší možnou hodnost n (to jest, $\nabla F_1(x), \dots, \nabla F_n(x)$ jsou lineárně nezávislé vektory v $\mathbf{R}^m$).*

1. *Pokud pro každé $\lambda \in \mathbf{R}^n$ platí $\nabla L(a, \lambda) \neq \overline{0}$, potom f nemá v a vzhledem k H ani neostrý lokální extrém.*
2. *Pokud $\lambda \in \mathbf{R}^n$ splňuje $\nabla L(a, \lambda) = \overline{0}$ a kvadratická forma*
$$P(h_1, \dots, h_m) = \sum_{i,j=1}^{m} \partial^2_{x_i x_j} L(a, \lambda) h_i h_j$$
je pozitivně (negativně) definitní na vektorech $h \in TH_a$, potom má funkce f v a vzhledem k množině H ostré lokální minimum (maximum).
3. *Pokud za stejných předpokladů jako ve 2 je $P(h_1, \dots, h_m)$ indefinitní na vektorech $h \in TH_a$, nemá f v a vzhledem k H ani neostrý lokální extrém.*

Podmínka 1 je samozřejmě už bůhví kolikátou reformulací důsledku 16. Nepřehlédněte, že definitnost či indefinitnost kvadratické formy P ve 2 a 3 se požaduje na tečném prostoru TH_a. To je ale jen malá obtíž—relace $h \in TH_a$ dává pro vektor h n lineárních rovnic. Můžeme tedy eliminovat n závislých proměnných h_i a dostaneme kvadratickou formu v $m-n$ nezávislých proměnných, jejíž definitnost či indefinitnost už vyšetřujeme na celém $\mathbf{R}^{m-n}$.

Příklad (nebyl na přednášce). Chceme nalézt lokální a globální extrémy funkce
$$f: \mathbf{R}^3 \to \mathbf{R},\ f(x, y, z) = x^2 - y^2 + z^2$$
na rovině H dané rovnicí
$$F(x, y, z) = 2x - y - 3 = 0.$$

Gradient Lagrangeovy funkce $L(x, y, z, \lambda) = x^2 - y^2 + z^2 - \lambda(2x - y - 3)$ je
$$\nabla L = (2x - 2\lambda,\ -2y + \lambda,\ 2z,\ -2x + y + 3).$$

Soustava lineárních rovnic $\nabla L(x, y, z, \lambda) = (0, 0, 0, 0)$ má řešení $(x, y, z, \lambda) = (2, 1, 0, 2)$. Prozkoumáme tedy (jediný) bod $a = (2, 1, 0)$, který leží na H a splňuje podmínku Lagrangeových multiplikátorů. Matice druhých derivací funkce L podle proměnných x, y, z je

$$(\partial^2_{xx,xy,\ldots,zz} L(x, y, z, \lambda)) = \begin{pmatrix} 2 & , & 0 & , & 0 \\ 0 & , & -2 & , & 0 \\ 0 & , & 0 & , & 2 \end{pmatrix}$$

(vůbec nezávisí na hodnotách proměnných) a odpovídá jí kvadratická forma $P(x, y, z) = 2x^2 - 2y^2 + 2z^2$. [2] Tečný prostor TH_a je dán rovnicí

$$\langle \nabla F(a), (x, y, z) \rangle = 2x - y = 0$$

(je to pochopitelně opět H, jen posunutá do počátku). Odtud vyjádříme $y = 2x$ a dosadíme do P: $P = -6x^2 + 2z^2$. Tato kvadratická forma je indefinitní na $\mathbf{R}^2$ a bod $a = (2, 1, 0)$ tak podle 3 předchozí věty není bodem lokálního extrému. Funkce f tedy na rovině H nemá žádný lokální a tedy ani žádný globální extrém.

Pro úplnost po tomto "profesorském" řešení uvedeme ještě řešení, řekněme, "studentské". Z rovnice definující H vyjádříme $y = 2x - 3$ a dosadíme do funkce f: $f = x^2 - (2x - 3)^2 + z^2 = -3x^2 + 12x - 9 + z^2 = -3(x - 2)^2 + z^2 + 3$, kde x, z už probíhají bez vazby celé $\mathbf{R}^2$. To je indefinitní kvadratická forma a proto f vskutku nemá na H ani lokální ani globální extrém.

Jako cvičení si zkuste metodou Lagrangeovy funkce nalézt lokální a globální extrémy funkce $f(x, y) = x^2 + y^2$ na elipse H dané rovnicí

$$F(x, y) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0.$$

[2] Forma P je sice indefinitní, ale musíme počítat dál, protože její zúžení na TH_a by mohlo být definitní nebo semidefinitní.