10. přednáška 3. prosince 2007

Věta 2.12 tedy říká, že soustavu

$$F_1(x_1,\ldots,x_m,y_1,\ldots,y_n)=\ldots=F_n(x_1,\ldots,x_m,y_1,\ldots,y_n)=0$$

lze lokálně vyřešit pomocí implicitních funkcí $y_1=f_1(x_1,\ldots,x_m),\ldots,y_n=f_n(x_1,\ldots,x_m)$, pokud F_i mají lokálné spojité první parciální derivace a v daném bodě je determinant matice derivací funkcí F_i podle y_j nenulový. Navíc pak f_i také mají lokálné spojité první parciální derivace a ty se spočtou jako minus podíl dvou determinantů.

Příklad. Ukažte, že soustava rovnic

$$x+y-\sin z=0 \ \text{ a } \ -x^3-y^3+\mathrm{e}^z-1=0$$

definuje v okolí bodu $x=0$ funkce $y=y(x)$ a $z=z(x)$ třídy C^1 splňující $y(0)=z(0)=0$ a spočítejte hodnoty derivací $y'(0)$ a $z'(0)$.

Pro $F_1(x,y,z)=x+y-\sin z$, $F_2(x,y,z)=-x^3-y^3+\mathrm{e}^z-1$ 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 splněny a uvedené funkce $y(x)$ a $z(x)$ jsou na okolí nuly definovány. Protože

$$\partial_x F(0,0,0)=\begin{pmatrix}1\\ -3x^2\end{pmatrix}(0,0,0)=\begin{pmatrix}1\\ 0\end{pmatrix},$$

podle vztahů uvedených na konci předešlé přednášky máme

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

Diferenciál inverzního zobrazení. Důsledkem věty o implicitních funkcích je zobecnění formule pro derivaci inverzní funkce pro více proměnných.

Důsledek 2.13. *Nechť $U\subset\mathbf{R}^m$ je okolí bodu $a\in\mathbf{R}^m$ a*

$$f:\ U\to\mathbf{R}^m$$

je zobrazení z $C^1(U)$ *splňující* $\det(f'(a)) \neq 0$. *Potom existují okolí* $U_1 \subset U$, *respektive* $V \subset \mathbf{R}^m$ *bodů* a, *respektive* $b = f(a)$ *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 inverzního zobrazení f^{-1} *v bodě* y *je tedy inverzní k Jacobiho matici zobrazení* f *v bodě* x.

Důkaz. Uvažme zobrazení o $2m$ proměnných

$$F(x, y) = f(x) - y : U \times \mathbf{R}^m \to \mathbf{R}^m,$$

kde $F = (F_1, \dots, F_m)$ a $F_i(x, y) = f_i(x) - y_i$. Pak $F_i(a, b) = 0$ pro $1 \leq i \leq n$. Na soustavu $F(x, y) = \overline{0}$ aplikujeme větu 2.12. Doplňte další podrobnosti důkazu jako cvičení (úloha 1). □

Vázané extrémy. Dalším důsledkem věty o implicitních funkcích je zobecnění první části věty 2.11 (nutnou podmínkou lokálního extrému funkce v bodě otevřené množiny je nulovost všech parciálních derivací) na extrémy na množině zadané soustavou rovnic. Nechť $U \subset \mathbf{R}^m$ je otevřená množina a

$$f, F_1, \dots, F_n : U \to \mathbf{R}$$

jsou funkce z $C^1(U)$, přičemž $n < m$. Hledáme lokální extrémy funkce f na množině

$$H = \{x \in U \mid F_1(x) = F_2(x) = \cdots = F_n(x) = 0\}.$$

Typicky tato množina nemá žádný vnitřní bod a nelze použít větu 2.11. Příkladem je jednotková sféra v $\mathbf{R}^m$:

$$H = \{x \in \mathbf{R}^m \mid x_1^2 + x_2^2 + \ldots + x_m^2 - 1 = 0\}.$$

Důsledek 2.14 (Lagrangeovy multiplikátory). *Nechť* a *je bod z* H. *Když jsou vektory* $\nabla F_1(a), \dots, \nabla F_n(a)$ *z* $\mathbf{R}^m$ *lineárně nezávislé a vektor* $\nabla f(a)$ *není jejich lineární kombinací, pak* f *nemá v bodu* a *na množině* H *ani neostrý lokální extrém.*

Ekvivalentně řečeno, když jsou $\nabla F_1(a), \dots, \nabla F_n(a)$ *lineárně nezávislé a funkce* f *má v bodě* a *na množině* H *(ostrý či neostrý) lokální extrém, potom existují taková reálná čísla* $\lambda_1, \dots, \lambda_n$ *(tzv.* Lagrangeovy multiplikátory*), že*

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

neboli $\partial_{x_j} f(a) - \lambda_1 \partial_{x_j} F_1(a) - \cdots - \lambda_n \partial_{x_j} F_n(a) = 0$ *pro* $1 \leq j \leq m$.

Důkaz. V důkazu použijeme druhou formulaci a předpokládáme, že f má v a na H lokální extrém. Lineární nezávislost uvedených vektorů znamená, že když je složíme jako řádky do matice (Jacobiho matice zobrazení $F = (F_1, \ldots, F_n)$), má tato $n \times m$ matice takových n sloupců, že determinant odpovídající čtvercové podmatice je nenulový. Pro jednoduchost značení předpokládáme, že to je posledních n sloupců. Přeznačíme-li tedy proměnné jako

$$\begin{aligned} x_1, x_2, \ldots, x_m &= y_1, y_2, \ldots, y_{m-n}, z_1, z_2, \ldots, z_n \\ &= y, z, \end{aligned}$$

máme $\det(\partial_{z_1} F(a), \ldots, \partial_{z_n} F(a)) \neq 0$. Podle věty 2.12 existují taková okolí U_1 a V_1 bodů

$$y_0 = (a_1, \ldots, a_{m-n}) \ \text{ a } \ z_0 = (a_{m-n+1}, \ldots, a_m)$$

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

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

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

$$h(y) = f(y, g_1(y), \ldots, g_n(y)) : \ U_1 \to \mathbf{R}.$$

Protože má v y_0 lokální extrém (nyní už bez vazby), první část věty 2.11 dává $\nabla h(y_0) = \overline{0}$. Parciálním derivováním složené funkce máme

$$\frac{\partial f}{\partial y_i}(y_0, g(y_0)) + \sum_{j=1}^{n} \frac{\partial f}{\partial z_j}(y_0, g(y_0)) \cdot \frac{\partial g_j}{\partial y_i}(y_0) = 0, \ 1 \leq i \leq m - n.$$

V řeči Jacobiho matic,

$$f'_y(y_0, z_0) + f'_z(y_0, z_0) \cdot g'(y_0) = \overline{0}.$$

Za $g'(y_0)$ dosadíme vyjádření podle vzorce ve větě 2.12:

$$f'_y(y_0, z_0) - f'_z(y_0, z_0) \cdot F'_z(y_0, z_0)^{-1} \cdot F'_y(y_0, z_0) = \overline{0}.$$

Označíme-li

$$\lambda = (\lambda_1, \ldots, \lambda_n) = f'_z(y_0, z_0) \cdot F'_z(y_0, z_0)^{-1},$$

dostáváme

$$f'_y(y_0, z_0) - \lambda F'_y(y_0, z_0) = \overline{0}.$$

Ale z $\lambda = f'_z(y_0, z_0) \cdot (F'_z)^{-1}(y_0, z_0)$ úpravou plyne, že stejný vztah platí i v z-ových proměnných: $f'_z(y_0, z_0) - \lambda F'_z(y_0, z_0) = \overline{0}$. Celkem v y-ových i z-ových proměnných máme

$$f'(y_0, z_0) - \lambda F'(y_0, z_0) = \overline{0},$$

takže $\nabla f(a) = \lambda_1 \nabla F_1(a) + \ldots + \lambda_n \nabla F_n(a)$. □

Všimněte si, že $\nabla f(a) = \overline{0}$ je lineární kombinací gradientů $\nabla F_i(a)$ vždy a pro nulový gradient f v a tedy důsledek 2.14 (stejně jako první část věty 2.11) nic neříká. Pomocí *Lagrangeovy funkce*

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

se důsledek 2.14 dá hezky přeformulovat. Protože

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

(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 $\lambda = (\lambda_1,\ldots,\lambda_n)$ jsou Lagrangeovy multiplikátory (prvních m souřadnic gradientu). Nutnou podmínku lokálního extrému funkce f v bodě a vzhledem k H tedy můžeme zformulovat i takto:

> Když má funkce f v bodě $a \in H$ lokální vázaný extrém, existuje takový bod $\lambda \in \mathbf{R}^n$, 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.

Uvedeme ještě jednu ekvivalentní formulaci důsledku 2.14. Nechť $\nabla f(a) \neq \overline{0}$ a vektory $\nabla F_1(a),\ldots,\nabla F_n(a)$ jsou lineárně nezávislé. Uvažme dva vektorové podprostory $\mathbf{R}^m$ složené z vektorů kolmých na $\nabla f(a)$, respektive z vektorů kolmých na každý z vektorů $\nabla F_1(a),\ldots,\nabla F_n(a)$:

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

Podprostor TN_a má dimenzi $m-1$ a TH_a má dimenzi $m-n$. 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 a podobně je $a + TN_a$ tečnou afinní nadrovinou v bodě a k „vrstevnicové" ploše

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

Podprostorům TN_a a TH_a se říká *tečné prostory* (k odpovídajícím plochám v bodě a). Z lineární algebry (teorie ortogonálních doplňků) víme, že $\nabla f(a)$ je lineární kombinací vektorů $\nabla F_1(a),\ldots,\nabla F_n(a)$, právě když $TH_a \subset TN_a$. Nutná podmínka lokálního vázaného extrému má tedy i tuto geometrickou formulaci.

> Když má funkce f v bodě $a \in H$ lokální vázaný extrém, je tečný prostor TH_a k ploše H v bodě a obsažený v tečném prostoru TN_a k vrstevnicové ploše N funkce f v bodě a,
>
> $$TH_a \subset TN_a.$$

Příklad: auto na horské silnici. Podíváme se na situaci $m = 2$ a $n = 1$. Funkce $f : \mathbf{R}^2 \to \mathbf{R}$ například udává nadmořskou výšku $f(x)$ bodu v terénu se zeměpisnými souřadnicemi x a křivka $H = \{x \in \mathbf{R}^2 \mid F(x) = 0\}$ je silnice. Vrstevnice $N = \{x \in \mathbf{R}^2 \mid f(x) = f(a) = b\}$, kde $a \in H$, je též rovinná křivka. Nechť $\nabla f(a)$ a $\nabla F(a)$ jsou nenulové vektory. Tečné prostory TH_a a TN_a pak mají dimenzi 1 a přímky $p = a + TH_a$ a $q = a + TN_a$ jsou tečny ke křivkám H a N v jejich průsečíku a. Předpokládejme, že $TH_a \not\subset TN_a$. Pak p a q jsou dvě různé přímky procházející společným bodem a. Vrstevnice N, která poblíž a úzce sleduje q, musí v průsečíku a přecházet z jedné strany silnice H na její druhou stranu, protože H zase úzce sleduje p. Pokud nadmořskou výšku vrstevnice b málo změníme, zmenšíme nebo zvětšíme, vrstevnice N se též změní jen málo a stále musí přecházet z jedné strany H na druhou a musí tak H protínat. Na silnici se tedy v okolí a nacházejí body jak s menší tak s větší nadmořskou výškou, než má a, a proto f nemá v a na H lokální extrém. Pokud na silnici H zastaví v bodě a auto, v neutrálu se bez ruční brzdy určitě rozjede!

Úlohy

1. (budou doplněny)