Matematika III – 4. týden
Funkce více proměnných: Vázané extrémy

Jan Slovák

Masarykova univerzita
Fakulta informatiky

6. 10. – 10. 10. 2014

Obsah přednášky

Kde je dobré číst?

- Zuzana Došlá, Roman Plch, Petr Sojka, Diferenciální počet funkcí více proměnných s programem Maple, MU Brno, 1999, 273 s.
- J. Slovák, M. Panák, M. Bulant, Matematika drsně a svižně, Muni Press, Brno 2013, v+773 s., elektronická edice `www.math.muni.cz/Matematika_drsne_svizne`

Gradient funkce

Definition

Pro spojitě diferencovatelnou funkci $F(x_1, \ldots, x_n) : \mathbb{R}^n \to \mathbb{R}$ se vektor

$$D^1 F = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right)$$

nazývá **gradient funkce** F.

Rovnost $F(x_1, \ldots, x_n) = b$ s pevnou hodnotou $b \in \mathbb{R}$ zadává podmnožinu $M \subset \mathbb{R}^n$, která mívá vlastnosti $(n-1)$–rozměrné nadplochy. Hovoříme v této souvislosti také o **úrovňových množinách** M_b.

Theorem

Směr zadaný gradientem v bodě $x = (x_1, \dots, x_n)$ je právě ten směr, ve kterém funkce F nejrychleji roste.
Tečná rovina k neprázdné úrovňové množině M_b v okolí jejího bodu s nenulovým gradientem D^1F je určena ortogonálním doplňkem ke gradientu.

Násobkům gradientu v tomto případě říkáme **normálový vektor** nadplochy M_b.

Theorem

Pro funkci F n proměnných a bod $P = (a_1, \dots, a_n) \in M_b$ v jehož okolí je M_b grafem funkce $(n-1)$ proměnných je implicitní rovnice pro tečnou nadrovinu

$$0 = \frac{\partial f}{\partial x_1}(P) \cdot (x_1 - a_1) + \cdots + \frac{\partial f}{\partial x_n}(P) \cdot (x_n - a_n).$$

Obecné dimenze: funkce $F = (f_1, \dots, f_n) : \mathbb{R}^{m+n} \to \mathbb{R}^n$ a n rovnic

$$f_i(x_1, \dots, x_{m+n}) = b_i, \quad i = 1, \dots, n.$$

dle věty o implicitní funkci je „většinou“ množina všech řešení $(x_1, \dots, x_{m+n})$ grafem zobrazení $G : \mathbb{R}^m \to \mathbb{R}^n$. Pro pevnou volbu $b = (b_1, \dots, b_n)$ je samozřejmě množinou M všech řešení průnik nadploch $M(b_i, f_i)$ příslušejících jednotlivým rovnicím $f_i = b_i$. Totéž platí pro tečné směry a normálové směry:
Tečný prostor k M bodem P je dán rovnicemi:

$$\begin{aligned}
0 &= \frac{\partial f_1}{\partial x_1}(P)\cdot(x_1 - a_1) + \dots + \frac{\partial f_1}{\partial x_n}(P)\cdot(x_{m+n} - a_{m+n})\\
&\vdots\\
0 &= \frac{\partial f_n}{\partial x_1}(P)\cdot(x_1 - a_1) + \dots + \frac{\partial f_n}{\partial x_n}(P)\cdot(x_{m+n} - a_{m+n}).
\end{aligned}$$

Normálový prostor v bodě P je afinní podprostor generovaný bodem P a gradienty všech funkcí $f_1, \dots, f_n$ v bodě P, tj. řádky Jacobiho matice D^1F.

V praxi mívají optimalizační úlohy často $m+n$ parametrů, které jsou vázány n podmínkami. V našem jazyce diferenciálního počtu tedy hledáme extrémy spojitě diferencovatelné funkce h na množině bodů M zadaných implicitně rovnicí $F(x_1, \dots, x_{m+n}) = 0$.
Pokud je M ve všech svých bodech grafem hladkého zobrazení v m proměnných, musí být každý extrém $P \in M$ stacionárním bodem, tj. pro každou křivku $c(t) \subset M$ procházející přes $P = c(0)$ musí být $h(c(t))$ extrémem pro tuto funkci jedné proměnné. Proto musí platit

$$\frac{d}{dt} h(c(t))_{|t=0} = d_{c'(0)} h(P) = dh(P)(c'(0)) = 0.$$

Tato vlastnost je ekvivalentní tvrzení, že gradient h leží v normálovém podprostoru (přesněji v jeho zaměření). Takové body $P \in M$ budeme nazývat **stacionární body** funkce H vzhledem k vazbám F.

Normálový prostor k naší množině M je generován řádky Jacobiho matice zobrazení F a stacionární body jsou proto ekvivalentně určeny následujícím tvrzením, kterému se říká **metoda Lagrangeových multiplikátorů**:

Theorem

Nechť $F = (f_1, \ldots, f_n) : \mathbb{R}^{m+n} \to \mathbb{R}^n$ je spojitě diferencovatelná v okolí bodu P, $F(P) = 0$ a M je zadána implicitně rovnicí $F(x, y) = 0$ a hodnost matice D^1F v bodě P je n. Pak P je stacionárním bodem spojitě diferencovatelné funkce $h : \mathbb{R}^{m+n} \to \mathbb{R}$ právě, když existují reálné parametry $\lambda_1, \ldots, \lambda_n$ takové, že

$$\operatorname{grad} h = \lambda_1 \operatorname{grad} f_1 + \cdots + \lambda_n \operatorname{grad} f_n.$$

Všimněme si počtu neznámých a rovnic v tomto algoritmu: gradienty jsou vektory o $m+n$ souřadnicích, tedy požadavek z věty dává $m+n$ rovnic. Jako proměnné máme jednak souřadnice $x_1, \dots, x_{m+n}$ hledaných stacionárních bodů P, ale navíc také n parametrů λ_i v hledané lineární kombinaci. Zbývá však požadavek, že hledaný bod P patří implicitně zadané množině M, což představuje dalších n rovnic. Celkem tedy máme $2n+m$ rovnic pro $2n+m$ proměnných a proto lze očekávat, že řešením bude diskrétní množina bodů P (tj. každý z nich zpravidla bude izolovaným bodem).