

Matematická analýza pro informatiky I.

12. přednáška

Extrémy funkcí více proměnných

Jan Tomeček
jan.tomecek@upol.cz
<http://aix-slx.upol.cz/~tomecek/index>

Univerzita Palackého v Olomouci

12. dubna 2011

Extrémy funkcí více proměnných

Na této přednášce se budeme bavit třemi typy extrémů funkcí více proměnných:

- lokální extrémy,
- vázané lokální extrémy,
- globální (absolutní) extrémy.

Lokální extrémy

Podobně jako u funkce jedné proměnné nás budou zajímat extrémy funkcí více proměnných. Při definování pojmů se budeme dost inspirovat právě analogiemi s funkcemi jedné proměnné. Nejprve se podíváme na tzv. lokální extrémy – půjde o extrémální hodnoty **na okolí vnitřních bodů definičního oboru** funkce.

Definice

Nechť $f : \mathbb{R}^N \rightarrow \mathbb{R}$, $a \in \text{int}(\mathcal{D}(f))$. Řekneme, že f **nabývá v bodě a lokální maximum (minimum)**, jestliže existuje $\mathcal{U}(a) \subset \mathcal{D}(f)$ tak, že

$$f(x) \leq f(a) \quad \forall x \in \mathcal{U}(a)$$

$$(f(x) \geq f(a) \quad \forall x \in \mathcal{U}(a)).$$

Funkční hodnotě $f(a)$ pak říkáme lokální maximum (minimum). Souhrnně jim říkáme lokální extrémy. Bodu a říkáme **bod lokálního maxima (minima)** – souhrnně bod extrému funkce f .

pokračování definice

Řekneme, že f **nabývá v bodě a ostré lokální maximum (minimum)**,
jestliže existuje $\mathcal{R}(a) \subset \mathcal{D}(f)$ tak, že

$$f(x) < f(a) \quad \forall x \in \mathcal{R}(a)$$

$$(f(x) > f(a) \quad \forall x \in \mathcal{R}(a)).$$

Funkční hodnotě $f(a)$ pak říkáme **ostré lokální maximum (minimum)**.
Souhrnně jim říkáme **ostré lokální extrémy**.

Hledání lokálních extrémů

Jak lokální extrémy hledat? Inspiraci opět najdeme u funkcí jedné proměnné. Lokální extrémy jsme hledali tak, že jsme nejprve našli tzv. stacionární body (body s nulovou derivací) a pomocí znaménka druhé derivace, popř. derivací vyššího řádu jsme určili zda jde o extrém a v kladném případě o jaký extrém šlo.

Věta

Nechť má funkce f v bodě $a \in \text{int}(\mathcal{D}(f))$ lokální extrém. Existuje-li $f_{x_i}(a)$ ($i \in \{1, \dots, N\}$), pak

$$\frac{\partial f}{\partial x_i}(a) = 0.$$

Důkaz

Položme

$$\varphi(t) = f(a + te^{[i]}) \quad \text{pro } t \in \mathcal{U}(0) \subset \{t \in \mathbb{R} : a + te^{[i]} \in \mathcal{D}(f)\}.$$

Z předpokladu plyne, že funkce φ má v bodě 0 lokální extrém. Tedy $\varphi'(0) = 0$. Z definice parciální derivace plyne

$$\frac{\partial f}{\partial x_i}(a) = \varphi'(0) = 0.$$

Asi nás proto nepřekvapí následující definice (vzpomeňte si na stacionární bod funkce jedné proměnné).

Definice

Mějme $f : \mathbb{R}^N \rightarrow \mathbb{R}$, $a \in \text{int}(\mathcal{D}(f))$. Řekneme, že a je stacionární bod funkce f , jestliže existují všechny parciální derivace funkce f v bodě a a platí

$$\frac{\partial f}{\partial x_i}(a) = 0 \quad \forall i = 1, \dots, N.$$

Nutná podmínka lokálního extrému

Důsledek

(*Nutná podmínka existence lokálního extrému*) Každý bod lokálního extrému, ve kterém existují všechny parciální derivace prvního řádu, je stacionární bod.

Poznámka

Naopak to již neplatí. Například bod $a = (0, 0)$ je stacionárním bodem funkce

$$f(x_1, x_2) = x_1 x_2$$

a přitom není bodem lokálního extrému této funkce. Nakreslete si hladiny funkce (nebo si nechte vykreslit graf nějakým programkem).

Jak určit, že jde o extrém?

Tedy podobně jako u funkce jedné proměnné musíme nějakým způsobem zjistit, zda je stacionární bod také hledaným lokálním extrémem – k tomu nám poslouží diferenciály vyšších řádů – pro jednoduchost si probereme podrobně pouze druhý diferenciál – což je kvadratická forma.

Lokální extrémy funkce jedné proměnné jsme určovali například takto: Našli jsme **stacionární body**. V těch jsme si spočítali **druhou derivaci**. Pokud vyšla **kladná**, šlo o ostré lokální **minimum**; pokud **záporná**, šlo o ostré lokální **maximum**.

Zde bude situace podobná – s tím rozdílem, že kladnost druhé derivace nahradí **pozitivní definitnost** druhého diferenciálu (zápornost nahradí **negativní definitnost** druhého diferenciálu).

Definitnost kvadratických forem

Kvadratické formy rozdělujeme na pozitivně (negativně) definitní (semidefinitní) a indefinitní.

Definice

Nechť $Q : \mathbb{R}^N \rightarrow \mathbb{R}$ je kvadratická forma. Řekneme, že Q je

- *pozitivně definitní, jestliže $Q(h) > 0$ pro všechna $h \in \mathbb{R}^N$, $h \neq 0$,*
- *negativně definitní, jestliže $Q(h) < 0$ pro všechna $h \in \mathbb{R}^N$, $h \neq 0$,*
- *pozitivně semidefinitní, jestliže $Q(h) \geq 0$ pro všechna $h \in \mathbb{R}^N$,*
- *negativně semidefinitní, jestliže $Q(h) \leq 0$ pro všechna $h \in \mathbb{R}^N$,*
- *indefinitní, jestliže existují $h_1, h_2 \in \mathbb{R}^N$ tak, že $Q(h_1) > 0$ a $Q(h_2) < 0$.*

Jaký je vztah mezi jednotlivými definitnostmi?

Postačující podmínka lokálního extrému

Věta

(Postačující podmínka existence lokálního extrému) Necht' $f : \mathbb{R}^N \rightarrow \mathbb{R}$, $a \in \text{int}(\mathcal{D}(f))$ je **stacionární** bod funkce f a funkce f má na nějakém okolí $\mathcal{U}(a)$ spojitě všechny parciální derivace druhého řádu. Pak platí

- je-li $d^2f(a)$ **pozitivně definitní**, pak f má v bodě a **ostré lokální minimum**,
- je-li $d^2f(a)$ **negativně definitní**, pak f má v bodě a **ostré lokální maximum**,
- je-li $d^2f(a)$ **indefinitní**, pak f **nemá** v bodě a **lokální extrém**.

Důkaz

Z Taylorova vzorce plyne existence $\tau \in (0, 1)$ tak, že pro všechna $h \in \mathbb{R}^N$ taková, že $a + h \in \mathcal{U}(a)$ platí

$$f(a + h) = f(a) + \frac{1}{2!} d^2 f(a + \tau h)(h).$$

Je-li $d^2 f(a)$ pozitivně definitní, dá se ukázat, že existuje $\mathcal{U}_\delta(a) \subset \mathcal{U}(a)$ takové, že $d^2 f(a + \tau h)$ je pozitivně definitní (pokuste se to dokázat, nebo se podívejte do literatury) pro každé $h \in \mathbb{R}^N$ takové, že $a + h \in \mathcal{U}_\delta(a)$. Pak pro $a + h \in \mathcal{R}_\delta(a)$ platí

$$f(a + h) - f(a) = \frac{1}{2!} d^2 f(a + \tau h)(h) > 0.$$

Tedy f má v bodě a ostré lokální minimum. Ostatní případy se dokáží podobně.

Poznámka

K určení lokálního extrému nám tedy stačí zjistit definitnost druhého diferenciálu ve stacionárním bodě. Všimněte si, že věta nám nic neřekla o případech, kdy $d^2f(a)$ byla semidefinitní. V takovém případě si musíme poradit jinak – tento případ zde nebudeme řešit.

Nabízí se otázka co nastane, je-li $d^2f(a) = 0$ (popř. jsou takové diferenciály i vyšších řádů). Podobně jako u funkce jedné proměnné (srovnejte! – 9. přednáška, 19. slajd) platí následující věta.

Věta

(O lokálním extrému – nebude vyžadována u zkoušky) Necht' $f : \mathbb{R}^N \rightarrow \mathbb{R}$, $a \in \text{int}(\mathcal{D}(f))$ je **stacionární bod** funkce f takový, že f má na $\mathcal{U}(a)$ spojitě všechny parciální derivace až do m -tého řádu a

$$d^2 f(a) = 0, d^3 f(a) = 0, \dots, d^{m-1} f(a) = 0.$$

Platí

- je-li m **liché**, pak f **nemá** v bodě a **lokální extrém**,
- je-li m **sudé**, pak
 - ▶ je-li $d^m f(a)$ **pozitivně definitní**, pak f má v bodě a **ostré lokální minimum**,
 - ▶ je-li $d^m f(a)$ **negativně definitní**, pak f má v bodě a **ostré lokální maximum**,
 - ▶ je-li $d^m f(a)$ **indefinitní**, pak f **nemá** v bodě a **lokální extrém**.

Nastudujte si v literatuře co znamená pojem definitnosti pro homogenní formu sudého stupně (např. viz RACHŮNKOVÁ, RACHŮNEK).

Definitnost kvadratických forem a její určování

Nyní se podíváme, jak se dá efektivně zjišťovat definitnost kvadratické formy.

Na minulé přednášce zazněl **pojem kvadratické formy a jejího reprezentanta**. Kvadratickou formou rozumíme zobrazení $Q(x) : \mathbb{R}^N \rightarrow \mathbb{R}$, které můžeme definovat takto

$$Q(x) = xBx^T,$$

kde B je symetrická matice typu $N \times N$. Důležitý je fakt, že

každá kvadratická forma je svým reprezentantem určena **jednoznačně**

(tzn. kvadratická forma má jediného reprezentanta a naopak každá symetrická matice určuje jedinou kvadratickou formu).

Definitnost symetrické matice

Definice

Řekneme, že matice B je

- pozitivně definitní, jestliže $hBh^T > 0$ pro každé $h \in \mathbb{R}^N \setminus \{0\}$,
- negativně definitní, jestliže $hBh^T < 0$ pro každé $h \in \mathbb{R}^N \setminus \{0\}$,
- pozitivně semidefinitní, jestliže $hBh^T \geq 0$ pro každé $h \in \mathbb{R}^N$,
- negativně semidefinitní, jestliže $hBh^T \leq 0$ pro každé $h \in \mathbb{R}^N$,
- indefinitní, jestliže existují $h_1, h_2 \in \mathbb{R}^N$ tak, že $h_1 B h_1^T > 0$ a $h_2 B h_2^T < 0$,

Přímo z definice definitnosti kvadratické formy a symetrické matice plyne následující:

Věta

Kvadratická forma je pozitivně definitní (negativně definitní; pozitivně semidefinitní; negativně semidefinitní; indefinitní) **právě tehdy, když** je takový její reprezentant.

Existuje jednoduché kritérium pro určení definitnosti symetrické matice.

Věta

(Sylvesterovo kritérium) Necht' $B = (b_{ij})_{i,j=1}^{N,N}$ je symetrická čtvercová matice typu $N \times N$. Označme

$$\Delta_i = \begin{vmatrix} b_{11} & b_{12} & \cdots & b_{1i} \\ b_{21} & b_{22} & \cdots & b_{2i} \\ \vdots & \vdots & \ddots & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{ii} \end{vmatrix} \quad \text{pro } i = 1, \dots, N.$$

Platí

- je-li $\Delta_i > 0$ pro každé $i = 1, \dots, N$, pak je matice B **pozitivně definitní**,
- je-li $(-1)^i \Delta_i > 0$ pro každé $i = 1, \dots, N$, pak je matice B **negativně definitní**,
- je-li $\Delta_i \neq 0$ pro každé $i = 1, \dots, N$ a B není ani pozitivně definitní ani negativně definitní, pak je **indefinitní**.

Poznámka

Pokud je $\Delta_i = 0$ pro nějaké $i \in \{1, \dots, N\}$, pak musíme o definitnosti kvadratické formy rozhodnout jinak.

Příklad

Jsou dány matice

$$B_1 = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}, \quad B_2 = \begin{pmatrix} -1 & -1 \\ -1 & -5 \end{pmatrix}, \quad B_3 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Pro matici B_1 platí $\Delta_1 = 1 > 0$, $\Delta_2 = 1 > 0$, matice je tedy pozitivně definitní.

Pro matici B_2 platí $\Delta_1 = -1 < 0$, $\Delta_2 = 4 > 0$, matice je tedy negativně definitní.

Pro matici B_3 platí $\Delta_1 = 0$, takže podle Sylvestrova kritéria nemůžeme rozhodnout. Zvolíme-li $h_1 = (1, -1/2)$, pak $h_1 B_3 h_1^T = -3/4 < 0$, a pro $h_2 = (0, 1)$ platí $h_2 B_3 h_2^T = 1 > 0$. Tedy matice je indefinitní.

Příklad

Nalezněte lokální extrémy funkce

$$f(x, y, z) = x^3 + y^2 + z^2 + 12xy + z.$$

Vázané extrémy

Příklad

Budete stavět nádrž ve tvaru kváдру. Máte k dispozici $56m^2$ dlaždiček. Jak zvolit rozměry nádrže, aby měla co největší objem?

Řešení. Označme a, b, c – rozměry nádrže. Zřejmě budeme hledat maximum funkce třech proměnných

$$V(a, b, c) = abc,$$

ale toto maximum nás zajímá pouze v takových bodech $(a, b, c) \in \mathbb{R}^3$ pro které

$$a, b, c > 0 \quad \text{a} \quad ab + 2(ac + bc) = 56.$$

Definice

Nechť $f : \mathbb{R}^N \rightarrow \mathbb{R}$, $\emptyset \neq M \subset \mathcal{D}(f)$, $a \in M$. Řekneme, že f má v bodě a **lokální maximum (minimum) vzhledem k množině M** (nebo-li vázané lok. max/min), jestliže existuje $\mathcal{U}(a)$ takové, že

$$f(x) \leq f(a) \quad \text{pro každé } x \in \mathcal{U}(a) \cap M$$

$$(f(x) \geq f(a) \quad \text{pro každé } x \in \mathcal{U}(a) \cap M).$$

Podobně definujeme ostrá lokální max/min vzhledem k množině.

Množina M bývá zadána:

- parametricky,
- implicitně.

Nejprve se podívejme na **parametrické zadání množiny M** .

Příklad

Určete vázané lokální extrémy funkce

$$f(x, y) = x^2 + 2y^2$$

vzhledem k množině

$$M = \{(x, y) \in \mathbb{R}^2 : x = 3t, y = t + 2, t \in \mathbb{R}\}.$$

Řešení. Uvažujme funkci

$$g(t) = f(3t, t + 2) = 11t^2 + 8t + 8,$$

což je funkce jedné reálné proměnné t . Najdeme její lokální extrémy. Platí

$$g'(t) = 22t + 8, \quad g''(t) = 22.$$

Funkce g má v bodě $t = -\frac{4}{11}$ lokální minimum. Tedy funkce f má vzhledem k množině M lokální minimum v bodě $(-12/11, 18/11)$.

Příklad

Určete vázané lokální extrémy funkce

$$f(x, y, z) = 3x^2 + 2y^2 + z^2$$

vzhledem k množině

$$M = \{(x, y, z) \in \mathbb{R}^3 : z = x - y\}.$$

Řešení. Úlohu převedeme na hledání lokálního extrému funkce $g(x, y) = f(x, y, x - y)$, tedy funkce dvou proměnných.

Nyní se podívejme na složitější případ, ve kterém je množina M zadána implicitně, tzn. ve tvaru

$$M = \{x \in \mathbb{R}^N : g_j(x) = 0 \quad j = 1, \dots, m\}, \quad (1)$$

kde $g_1, \dots, g_m : \mathbb{R}^N \rightarrow \mathbb{R}$ jsou funkce, $m \leq N$. Množina M je tedy průnikem 0–hladin funkcí $g_1, \dots, g_m$.

Příklad

V prvním příkladu této části je množina M daná takto

$$M = \{(a, b, c) \in \mathbb{R}^3 : ab + 2ac + 2bc - 56 = 0\}.$$

Následující věta nám dává návod, jak hledat vázané lokální extrémy – **metodou Lagrangeových multiplikátorů**. Ačkoliv je tvrzení věty na první pohled trochu složitější, základní myšlenka je jednoduchá – body vázaných extrémů budou takové body z množiny M , ve kterých bude gradient funkce f (tedy funkce jejíž extrém hledáme) lineárně závislý se všemi gradienty funkcí $g_1, \dots, g_m$ v tomto bodě – stačí si nakreslit obrázek.

Věta

Nechť mají funkce $f, g_1, \dots, g_m : \mathbb{R}^N \rightarrow \mathbb{R}$ ($1 \leq m \leq N$) spojité parciální derivace prvního řádu a v každém bodě z množiny

$M = \{x \in \mathbb{R}^N : g_1(x) = \dots = g_m(x) = 0\}$ má matice

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_N} \\ \frac{\partial g_m}{\partial x_1} & \dots & \frac{\partial g_m}{\partial x_N} \end{pmatrix} \quad (2)$$

hodnotu rovnu m (tedy maximální řádkovou hodnotu). Má-li funkce f v bodě $a \in M$ lokální extrém vzhledem k množině M , pak existují $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ tak, že

$$\nabla f(a) = \sum_{j=1}^m \lambda_j \nabla g_j(a). \quad (3)$$

Náznak důkazu

Provedeme jen pro $N = 2$, $m = 1$. Pak f , g_1 jsou funkce dvou proměnných a M je 0–hladina funkce g_1 . Máme dokázat, že je-li $a \in M$ vázaný extrém funkce f vzhledem k M pak existuje $\lambda_1 \in \mathbb{R}$ tak, že

$$\nabla f(a) = \lambda_1 \nabla g_1(a).$$

To znamená, že vektory $\nabla f(a)$ a $\nabla g_1(a)$ jsou kolineární, z čehož plyne, že křivky

$$M \quad \text{a} \quad \{x \in D : f(x) = f(a)\}$$

mají v a společnou tečnu.

Poznámka

Definujeme takzvanou **Lagrangeovu funkci**

$$L(x, \lambda) = f(x) - \sum_{j=1}^m \lambda_j g_j(x),$$

kde $\lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$. Pak podmínka (3) se dá napsat jako

$$\frac{\partial L}{\partial x_i}(a, \lambda) = 0 \quad \text{pro každé } i = 1, \dots, N.$$

Definice

Nechť M je dána v (1). Řekneme, že $a \in M$ je stacionární bod funkce f na M , jestliže existují **Lagrangeovy multiplikátory** $\lambda_1, \dots, \lambda_m$ tak, že platí (3).

Důsledek

Funkce f může mít lokální extrém vzhledem k M jen ve stacionárních bodech na M . Opět musíme nějakým způsobem ověřit, jestli je stacionární bod bodem extrému vzhledem k M . K tomu nám opět poslouží druhý diferenciál.

Věta

Nechť mají funkce $f, g_1, \dots, g_m : \mathbb{R}^N \rightarrow \mathbb{R}$ ($1 \leq m \leq r$) spojité parciální derivace druhého řádu v bodě a , který je stacionárním bodem funkce f na M , $\lambda_1, \dots, \lambda_m$ jsou jeho Lagrangeovy multiplikátory, ($\lambda = (\lambda_1, \dots, \lambda_m)$) tzn. platí

$$\frac{\partial L}{\partial x_j}(a, \lambda) = 0 \quad \text{pro každé } j = 1, \dots, N.$$

Nechť má matice (2) v bodě a hodnost rovnu m . Jestliže pro **každé** $h \neq 0$ **splňující**

$$(\nabla g_j(a), h) = 0 \quad \text{pro } j = 1, \dots, m,$$

platí

$$d^2L(a, \lambda)(h) > 0 (< 0)$$

(druhý diferenciál je vzhledem k x), pak má funkce f v bodě a ostré lokální minimum (maximum) vzhledem k M .

pokračování tvrzení věty

Jestliže existují $h_1, h_2 \in \mathbb{R}^N$ takové, že

$$(\nabla g_j(a), h_i) = 0 \quad \text{pro } j = 1, \dots, m, \quad i = 1, 2$$

a

$$d^2L(a, \lambda)(h_1) > 0 \quad \text{a} \quad d^2L(a, \lambda)(h_2) < 0$$

pak f nemá v a lokální extrém vzhledem k M .

Tato věta nám již tedy říká, jak zjistit, že stacionární bod funkce vzhledem k množině je také bodem lokálního extrému vzhledem k této množině.

Příklad

Vyřešíme příklad ze začátku kapitoly. Abychom zachovali značení z tvrzení vět, provedeme přejmenování našich funkcí. Budeme tedy hledat vázaný extrém funkce

$$f(x, y, z) = xyz$$

vzhledem k množině

$$M = \{(x, y, z) \in \mathbb{R}^3 : xy + 2xz + 2yz - 56 = 0\}$$

Definujeme Lagrangeovu funkci

$$L(x, y, z, \lambda) = xyz + \lambda(xy + 2xz + 2yz - 56).$$

Nalezneme stacionární body... NA CVIČENÍ

Globální (absolutní) extrémy

Nyní nás budou zajímat nejmenší a největší hodnoty funkce na dané množině.

Definice

Nechť $f : \mathbb{R}^N \rightarrow \mathbb{R}$, $M \subset \mathcal{D}(f)$. Řekneme, že bod $a \in M$ je bodem **absolutního maxima (minima)**, jestliže

$$f(x) \leq f(a) \quad \text{pro každé } x \in M,$$

$$(f(x) \geq f(a) \quad \text{pro každé } x \in M).$$

Souhrnně je nazýváme **absolutními extrémy**. Obdobně definujeme **ostré** absolutní extrémy.

Věta

Nechť $f : \mathbb{R}^N \rightarrow \mathbb{R}$, $M \subset \mathcal{D}(f)$ je kompaktní, f je spojitá na M . Pak f nabývá na M svého absolutního maxima i minima. Přitom je nabývá buď v $\text{int}(M)$ a to v bodě lokálního extrému nebo na ∂M .

Podobnou větu jsme měli pro funkci jedné proměnné (9. přednáška, 35. slajd – najděte a porovnejte!).

Poznámka

Tvrzení věty nám dává praktický návod k určení absolutního extrému. Stačí vyšetřit body ležící v

- $\text{int}(M)$: určíme stacionární body funkce f v $\text{int}(M)$ a body z $\text{int}(M)$, ve kterých funkce f nemá některou z parciálních derivací,
- ∂M : určíme stacionární body funkce f vzhledem k množině ∂M (a body ve kterých Lagrangeova funkce nemá některou z parciálních derivací).

Porovnáme funkční hodnoty v těchto bodech a určíme největší a nejmenší z nich. To bude absolutní maximum a absolutní minimum

funkce f na množině M

Příklad

Vypočtěte maximum a minimum funkce

$$f(x, y) = xy$$

na množině

$$M = \{(x, y) \in \mathbb{R}^2 : x^2 + 2y^2 \leq 1\}.$$

OPĚT ŘEŠENÍ NA CVIČENÍ

Doporučená literatura

KOPÁČEK J. Matematická analýza pro fyziky II. Matfyzpress, Praha, 2005.

DOŠLÁ, Z, DOŠLÝ O.: Diferenciální počet funkcí více proměnných, Masarykova univerzita, Brno, 2003.