Drsná matematika III – 2. přednáška
Funkce více proměnných: Aproximace vyšších rádů, Taylorova věta, inverzní zobrazení

Masarykova univerzita
Fakulta informatiky

26. 9. 2011

Obsah přednášky

1. Literatura
2. Derivace vyšších řádů
   - Iterované parciální derivace
   - Hessián – aproximace 2. řádu
   - Lokální extrémy funkcí více proměnných
3. Zobrazení mezi euklidovskými prostory
   - Zobrazení a transformace
   - „Chain Rule“
   - Věta o inverzním zobrazení

Plán přednáky

Kde je dobré číst?

- Zuzana Došlá, Roman Plch, Petr Sojka, Diferenciální počet funkcí více proměnných s programem Maple, MU Brno, 1999, 273 s.
- Riley, K.F., Hobson, M.P., Bence, S.J. Mathematical Methods for Physics and Engineering, second edition, Cambridge University Press, Cambridge 2004, ISBN 0 521 89067 5, xxiii + 1232 pp.

Plán přednáky

1. Literatura

2. Derivace vyšších řádů
   - Iterované parciální derivace
   - Hessián – aproximace 2. řádu
   - Lokální extrémy funkcí více proměnných

3. Zobrazení mezi euklidovskými prostory
   - Zobrazení a transformace
   - „Chain Rule“
   - Věta o inverzním zobrazení

Definition

Funkce $f : \mathbb{R}^n \to \mathbb{R}$ je **diferencovatelná v bodě** x, jestliže

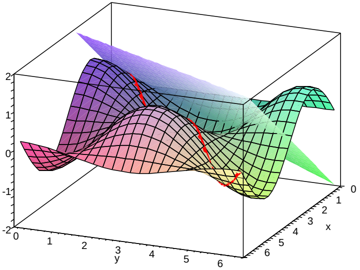
1. v bodě x existují všechny směrové derivace $d_v f(x)$, $v \in \mathbb{R}^n$,
2. $d_v f(x)$ je lineární v závislosti na přírůstku v a
3. $0 = \lim_{v \to 0} \frac{1}{\|v\|}\big(f(x+v) - f(x) - d_v f(x)\big)$.

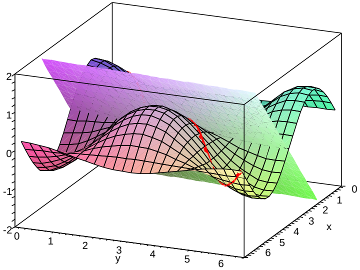
Theorem

Nech $f : E_n \to \mathbb{R}$ je funkce n proměnných, která má v okolí bodu $x \in E_n$ spojité parciální derivace. Pak existuje její diferenciál df v bodě x a jeho souřadné vyjádření je dáno rovnicí

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \cdots + \frac{\partial f}{\partial x_n} dx_n. \qquad (*)$$

Diferenciál zadává tečné (nad)roviny funkce n proměnných.





Graf funkce $f(x, y) = \sin(x)\cos(y)$, červená čára je obrazem křivky $c(t) = (t, t, f(t, t))$.

Diferencovatelná funkce f na E_n v bodě $x \in E_n$ má nulový diferenciál tehdy a jen tehdy, když její složení s libovolnou křivkou procházející tímto bodem zde má stacionární bod. To ovšem neznamená, že v takovém bodě musí mít f aspoň lokálně buď maximum nebo minimum. Stejně jako u funkcí jedné proměnné můžeme rozhodovat teprve podle derivací vyšších.

Pro pevný přírůstek $v \in \mathbb{R}^n$ je vyčíslení diferenciálů na tomto přírůstku opět (diferenciální) operace na funkcích $f : E_n \to \mathbb{R}$

$$f \mapsto d_v f = df(v).$$

Výsledkem je $df(v) : E_n \to \mathbb{R}$. Jestliže je tato funkce opět diferencovatelná, můžeme iterovat.
Pro **parciální derivace druhého řádu** píšeme

$$\left(\frac{\partial}{\partial x_j} \circ \frac{\partial}{\partial x_i}\right) f = \frac{\partial^2}{\partial x_i \partial x_j} f = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

v případě opakované volby $i = j$ píšeme také

$$\left(\frac{\partial}{\partial x_i} \circ \frac{\partial}{\partial x_i}\right) f = \frac{\partial^2}{\partial x_i^2} f = \frac{\partial^2 f}{\partial x_i^2}.$$

Úplně stejně postupujeme při dalších iteracích a hovoříme o **parciálních derivacích k-tého řádu**

$$\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}.$$

Theorem

Nech $f : E_n \to \mathbb{R}$ je k-krát diferencovatelná funkce se spojitými parciálními derivacemi až do řádu k včetně v okolí bodu $x \in \mathbb{R}^n$. Pak všechny parciální derivace nezávisí na pořadí derivování.

Definition

Je-li $f : \mathbb{R}^n \to \mathbb{R}$ libovolná dvakrát diferencovatelná funkce, nazýváme symetrickou matici funkcí

$$Hf(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n}(x) \end{pmatrix}$$

Hessián funkce f v bodě x.

Pro křivku $c(t) = (x(t), y(t)) = (x_0 + \xi t, y_0 + \eta t)$ mají funkce

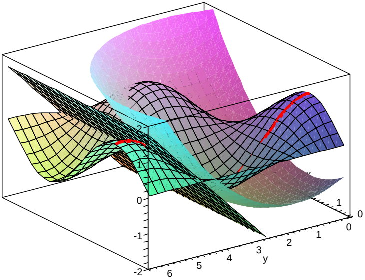
$$\alpha(t) = f(x(t), y(t))$$
$$\beta(t) = f(x_0, y_0) + t \cdot \left(\frac{\partial f}{\partial x}(x_0, y_0)\xi + \frac{\partial f}{\partial y}(x_0, y_0)\eta \right) + \frac{1}{2}t^2 \left(f_{xx}(x_0, y_0)\xi^2 + 2f_{xy}(x_0, y_0)\xi\eta + f_{yy}(x_0, y_0)\eta^2 \right)$$

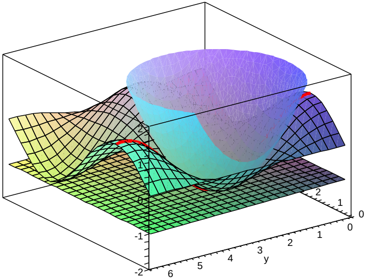
stejné derivace do druhého řádu včetně. Funkci β píšeme vektorově:

$$\beta(t) = f(x_0, y_0) + t df(x_0, y_0) \cdot \begin{pmatrix} \xi \\ \eta \end{pmatrix} + \frac{1}{2}t^2 (\xi \;\; \eta) \cdot Hf(x_0, y_0) \cdot \begin{pmatrix} \xi \\ \eta \end{pmatrix}$$

nebo $\beta(t) = f(x_0, y_0) + t df(x_0, y_0)(v) + \frac{1}{2}t^2 Hf(x_0, y_0)(v, v)$, kde $v = (\xi, \eta) = c'(t)$ je přírůstek zadaný derivací křivky $c(t)$ a Hessián symetrická 2–forma.

Použití Hessiánu připomíná Taylorovu větu funkcí jedné proměnné!





Tečná rovina je vynesena spolu s kvadratickým přiblížením pro funkci $f(x, y) = \sin(x)\cos(y)$.
Obecně pro funkce $f : E_n \to \mathbb{R}$, body $x = (x_1, \dots, x_2) \in E_n$ a přírůstky $v = (\xi_1, \dots, \xi_n)$ klademe

$$D^k f(x)(v) = \sum_{1 \le i_1, \dots, i_k \le n} \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(x_1, \dots, x_n) \cdot \xi_{i_1} \cdots \xi_{i_k}.$$

Ukažme ve dvou proměnných:
Tečná rovina: $f(x_0, y_0) + D^1(x_0, y_0)(x - x_0, y - y_0)$
Aproximace pomocí hesiánu:
$f(x_0, y_0) + D^1(x_0, y_0)(x - x_0, y - y_0) + \frac{1}{2} D^2(x_0, y_0) f(x - x_0, y - y_0)$
výraz třetího řádu

$$D^3 f(x,y)(\xi,\eta) = \frac{\partial^3 f}{\partial x^3}\xi^3 + 3\frac{\partial^3 f}{\partial x^2 \partial y}\xi^2\eta + 3\frac{\partial^3 f}{\partial x \partial y^2}\xi\eta^2 + \frac{\partial^3 f}{\partial y^3}\eta^3$$

a obecně

$$D^k f(x,y)(\xi,\eta) = \sum_{\ell=0}^{k} \binom{k}{\ell} \frac{\partial^k f}{\partial x^{k-\ell} \partial y^{\ell}} \xi^{k-\ell}\eta^{\ell}.$$

Theorem

Nech $f : E_n \to \mathbb{R}$ je k–krát diferencovatelná funkce v okolí $\mathcal{O}_\delta(x)$ bodu $x \in E_n$. Pro každý přírůstek $v \in \mathbb{R}^n$ s velikostí $\|v\| < \delta$ pak existuje číslo $0 \le \theta \le 1$ takové, že

$$f(x+v) = f(x) + D^1 f(x)(v) + \cdots + \frac{1}{(k-1)!} D^{k-1} f(x)(v) + \frac{1}{k!} D^k f(x + \theta \cdot v)(v).$$

Náznak důkazu: Pro přírůstek $v \in \mathbb{R}^n$ volíme $c(t) = x + tv$ v E_n a zkoumáme $\varphi : \mathbb{R} \to \mathbb{R}$ definovanou složením $\varphi(t) = f \circ c(t)$. Taylorova věta pro funkce jedné proměnné říká:

$$\varphi(t) = \varphi(0) + \varphi'(0)t + \cdots + \frac{1}{(k-1)!}\varphi^{(k-1)}(0)t^{k-1} + \frac{1}{k!}\varphi^{(k)}(\theta)t^k.$$

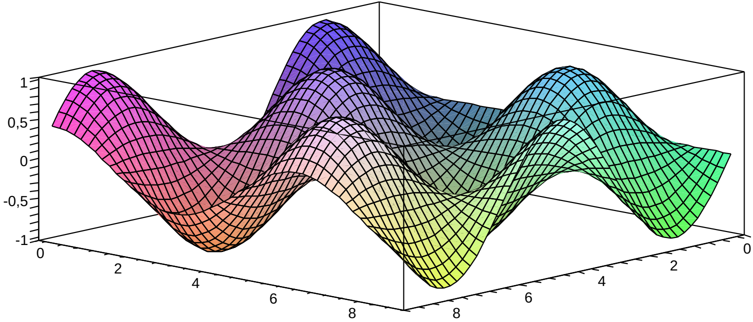
Zbývá nám tedy jen ověřit, že postupným derivováním složené funkce φ dostaneme právě požadovaný vztah. To lze provést indukcí přes řád k.

Definition

Vnitřní bod $x_0 \in E_n$ definičního oboru funkce f je (lokálním) **maximem** nebo **minimem**, jestliže existuje jeho okolí U takové, že pro všechny body $x \in U$ splňuje funkční hodnota $f(x) \leq f(x_0)$ nebo $f(x) \geq f(x_0)$. Pokud nastává v předchozích nerovnostech ostrá nerovnost pro všechny $x \neq x_0$, hovoříme o **ostrém extrému**. Vnitřní bod $x \in E_n$ definičního oboru funkce f, ve kterém je diferenciál $df(x)$ nulový nazýváme **stacionární bod funkce** f.

Nutnou podmínkou pro existenci maxima nebo minima v bodě x_0 je vymizení diferenciálu v tomto bodě, tj. $df(x_0) = 0$. Skutečně, pokud je $df(x_0) \neq 0$, pak existuje směr v, ve kterém je $d_v f(x_0) \neq 0$. Pak ovšem nutně je podél přímky $x_0 + tv$ na jednu stranu od bodu x_0 hodnota funkce roste a na druhou klesá.

Uvažme funkci $f(x, y) = \sin(x)\cos(y)$, která připomíná známá kartonová plata na vajíčka



Spočtěme si první a poté druhé derivace:

$$f_x(x,y) = \cos(x)\cos(y),\;\; f_y(x,y) = -\sin(x)\sin(y),$$

takže obě derivace budou nulové pro dvě sady bodů

1. $\cos(x) = 0$, $\sin(y) = 0$, to je $(x,y) = (\frac{2k+1}{2}\pi, \ell\pi)$, pro libovolné $k,\ell \in \mathbb{Z}$
2. $\cos(x) = 0$, $\sin(y) = 0$, to je $(x,y) = (k\pi, \frac{2\ell+1}{2}\pi)$, pro libovolné $k,\ell \in \mathbb{Z}$.

Druhé parciální derivace jsou

$$Hf(x,y) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix}(x,y) = \begin{pmatrix} -\sin(x)\cos(y) & -\cos(x)\sin(y) \\ -\cos(x)\sin(y) & -\sin(x)\cos(y) \end{pmatrix}$$

V našich dvou sadách bodů tedy dostáváme následující hessiány:

1. $Hf(k\pi + \frac{\pi}{2}, \ell\pi) = \pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, přičemž znaménko $+$ nastává, když parity k a ℓ jsou stejné a naopak pro $-$,
2. $Hf(k\pi, \ell\pi + \frac{\pi}{2}) = \pm \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, přičemž znaménko $+$ nastává, když parity k a ℓ jsou stejné a naopak pro $-$.

Taylorova věta pro řád $k = 2$ dává okolí stacionárních bodů (x_0, y_0)

$$f(x,y) = f(x_0,y_0) + \frac{1}{2} Hf(x_0+\theta(x-x_0), y_0+\theta(y-y_0))(x-x_0, y-y_0),$$

kde Hf nyní vnímáme jako kvadratickou formu vyčíslenou na přírůstku $(x - x_0, y - y_0)$. nastane lokální maximum tehdy a jen tehdy, když náš bod (x_0, y_0) patří do první skupiny se stejnými paritami k a ℓ. Když budou parity opačné, pak bod z první skupiny bude naopak bodem lokálního minima. Naopak, hessián u druhé skupiny bodů se vyčíslí kladně na některých přírůstcích a záporně na jiných. Stejně se proto bude chovat i celá funkce f v okolí.

Definition

Kvadratická forma $h : E_n \to \mathbb{R}$ je

- **positivně definitní**, je-li $h(u) > 0$ pro všechny $u \neq 0$
- **positivně semidefinitní**, je-li $h(u) \geq 0$ pro všechny $u \in V$
- **negativně definitní**, je-li $h(u) < 0$ pro všechny $u \neq 0$
- **negativně semidefinitní**, je-li $h(u) \leq 0$ pro všechny $u \in V$
- **indefinitní**, je-li $h(u) > 0$ a $f(v) < 0$ pro vhodné $u, v \in V$.

1. semestr $\to$ metody, které umožňují přímo zjistit, zda daná forma má některou z těchto vlastností.

Theorem

Nech $f : E_n \to \mathbb{R}$ je dvakrát spojitě diferencovatelná funkce a $x \in E_n$ nech je stacionární bod funkce f. Potom

1. *f má v x ostré lokální minimum, je-li $Hf(x)$ positivně definitní,*
2. *f má v x ostré lokální maximum, je-li $Hf(x)$ negativně definitní,*
3. *f nemá v bodě x lokální extrém je-li $Hf(x)$ indefinitní.*

Všimněme si, že věta nedává žádný výsledek, pokud je hessián funkce ve zkoumaném bodě degenerovaný a přitom není indefinitní. Důvod je opět stejný jako u funkcí jedné proměnné. V takových případech totiž existují směry, ve kterých první i druhá derivace zmizí a my proto v tomto řádu přiblížení neumíme poznat, zda se funkce bude chovat jako t^3 nebo jako $\pm t^4$ dokud nespočteme alespoň v potřebných směrech derivace vyšší.

Plán přednáky

1. Literatura

2. Derivace vyšších řádů
   - Iterované parciální derivace
   - Hessián – aproximace 2. řádu
   - Lokální extrémy funkcí více proměnných

3. Zobrazení mezi euklidovskými prostory
   - Zobrazení a transformace
   - „Chain Rule“
   - Věta o inverzním zobrazení

Zobrazení $F : E_n \to E_m$ je při zvolených kartézských souřadnicích na obou stranách obyčejná m–tice

$$F(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n))$$

funkcí $f_i : E_n \to \mathbb{R}$. Řekneme, že F je *diferencovatelné* nebo *spojitě* **diferencovatelné zobrazení**, jestliže tuto vlastnost mají všechny funkce $f_1, \dots, f_m$.
Diferencovatelná zobrazení $F : E_n \to E_n$, která mají inverzní zobrazení $G : E_n \to E_n$ definované na celém svém obrazu, se nazývají **(diferencovatelné) transformace**. Příkladem transformace byl přechod mezi kartézkými a polárními souřadnicemi.
Lineární zobrazení $D^1 f_i(x) : \mathbb{R}^n \to \mathbb{R}$ lineárně aproximují přírůstky f_i.

Definition

$$D^1F(x) = \begin{pmatrix} df_1(x) \\ df_2(x) \\ \vdots \\ df_m(x) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}(x)$$

se nazývá **Jacobiho matice zobrazení** F v bodě x. Lineární zobrazení $D^1F(x)$ definované na přírůstcích $v = (v_1, \dots, v_n)$ pomocí stejně značené Jacobiho matice nazýváme **diferenciál zobrazení** F v bodě x z definičního oboru, jestliže

$$\lim_{v \to 0} \frac{1}{\|v\|}\big(F(x+v) - F(x) - D^1F(x)(v)\big) = 0.$$

Důsledek Věty o existenci diferenciálu pro funkce n proměnných je:

Theorem

Nech $F : E_n \to E_m$ je zobrazení, jehož všechny souřadné funkce mají spojité parciální derivace v okolí bodu $x \in E_n$. Pak existuje diferenciál $D^1F(x)$ zadaný Jacobiho maticí.

Theorem

Nech $F : E_n \to E_m$ a $G : E_m \to E_r$ jsou dvě diferencovatelná zobrazení, přičemž definiční obor G obsahuje celý obor hodnot F. Pak také složené zobrazení $G \circ F$ je diferencovatelné a jeho diferenciál je v každém bodě z definičního obodu F kompozicí diferenciálů

$$D^1(G \circ F)(x) = D^1G(F(x)) \circ D^1F(x).$$

Příslušná Jacobiho matice je dána součinem příslušných Jacobiho matic.

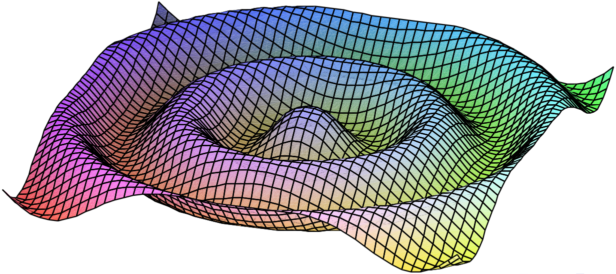
Polární souřadnice vzniknou z kartézských transformací $F : \mathbb{R}^2 \to \mathbb{R}^2$, kterou v souřadnicích (x, y) a (r, φ) zapíšeme:

$$r = \sqrt{x^2 + y^2}, \ \varphi = \arctan \frac{y}{x}.$$

Funkci $g_t : E_2 \to \mathbb{R}$ v polárních souřadnicích

$$g(r, \varphi, t) = \sin(\text{r-t})\,.$$

Funkce nám docela dobře přibližuje vlnění povrchu hladiny po bodovém vzruchu v počátku v čase t:



Derivace v kartézských souřadnicích:

$$\frac{\partial g}{\partial x}(x,y,t) = \frac{\partial g}{\partial r}(r,\varphi)\frac{\partial r}{\partial x}(x,y) + \frac{\partial g}{\partial \varphi}(r,\varphi)\frac{\partial \varphi}{\partial x}(x,y)$$
$$= \cos(\sqrt{x^2+y^2}-t)\frac{x}{\sqrt{x^2+y^2}} + 0$$

a podobně

$$\frac{\partial g}{\partial y}(x,y,t) = \frac{\partial g}{\partial r}(r,\varphi)\frac{\partial r}{\partial y}(x,y) + \frac{\partial g}{\partial \varphi}(r,\varphi)\frac{\partial \varphi}{\partial y}(x,y)$$
$$= \cos(\sqrt{x^2+y^2}-t)\frac{y}{\sqrt{x^2+y^2}}.$$

Theorem

Nech $F : E_n \to E_n$ je spojitě diferencovatelné zobrazení na nějakém okolí bodu $x_0 \in E_n$ a nech je Jacobiho matice $D^1 f(x_0)$ invertibilní. Pak na nějakém okolí bodu x_0 existuje inverzní zobrazení F^{-1} a jeho diferenciál v bodě $F(x_0)$ je inverzním zobrazením k $D^1 F(x_0)$, tzn. je zadán inverzní maticí k Jacobiho matici zobrazení F v bodě x_0.