# Matematika III – 3. týden
# Funkce více proměnných: Inverzní zobrazení, implicitně definovaná zobrazení a vázané extrémy

Jan Slovák

Masarykova univerzita
Fakulta informatiky

3. října – 9. října 2016

# Obsah přednášky

1. Literatura
2. Diferenciál složených zobrazení
3. Implicitně zadaná zobrazení
   - Tečny a normály k implicitně zadaným plochám
   - Gradient funkce
   - Tečné a normálové prostory
4. Vázané extrémy

# 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`

## Theorem

*Nechť* $F : E_n \to E_m$ *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 oboru* $F$ *kompozicí diferenciálů*

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

*Příslušná Jacobiho matice je dána součinem Jacobiho matic zobrazení* $G$ *a* $F$*.*

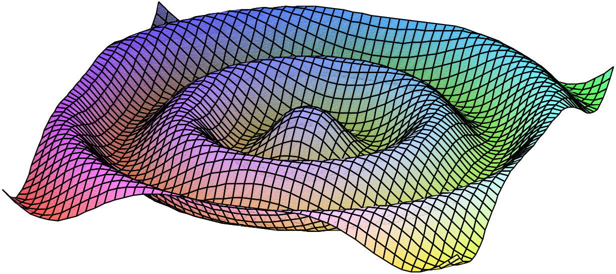
Polární souřadnice vzniknou z kartézských transformací $F : \mathbb{R}^2 \to \mathbb{R}^2$, kterou v souřadnicích $(x, y)$ a $(r, \varphi)$ zapíšeme:

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

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

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

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



Derivace $g(r,\varphi) = \sin(r - t)$ v kartézských souřadnicích:

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

a podobně

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

**Theorem**

*Nechť $F : E_n \to E_n$ je spojitě diferencovatelné zobrazení na nějakém okolí bodu $x_0 \in E_n$ a nechť 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 $x_0$.*

# Věta o implicitní funkci

Pro jednoduchost vyložíme ideu v rovině $E_2$:
Pro spojitě diferencovatelné zobrazení $F(x, y) : \mathbb{R}^2 \to \mathbb{R}$ hledejme body $(x, y)$, ve kterých platí $F(x, y) = 0$. Příkladem může být třeba obvyklá (implicitní) definice přímek a kružnic:

$$F(x, y) = ax + by + c = 0$$
$$F(x, y) = (x - s)^2 + (y - t)^2 - r^2 = 0, \;\; r > 0.$$

V prvém případě je (při $b \neq 0$) předpisem zadaná funkce

$$y = f(x) = -\frac{a}{b}x - \frac{c}{b}$$

pro všechna $x$, ve druhém umíme pouze pro $(a, b)$ splňující rovnici kružnice a $b \neq t$ najít okolí bodu $a$, na kterém nastane jedna z možností: $y = f(x) = t + \sqrt{(x - s)^2 - r}$,
$y = f(x) = t - \sqrt{(x - s)^2 - r}$.

Krajní body intervalu $[t - r, t + r]$ také vyhovují rovnici kružnice, platí v nich ale $F_y(s \pm r, t) = 0$, což vystihuje polohu tečny ke kružnici v těchto bodech rovnoběžné s osou $y$. V těchto bodech skutečně neumíme najít okolí, na němž by kružnice byla popsána jako funkce $y = f(x)$.
Navíc umíme i derivace:

$$f'(x) = \frac{1}{2} \frac{2(x - s)}{\sqrt{(x - s)^2 - r^2}} = \frac{x - s}{y - t} = -\frac{F_x}{F_y}.$$

Naopak, pokud budeme chtít najít závislost $x = f(y)$ takovou, aby $F(f(y), y) = 0$, pak v okolí bodů $(s \pm r, t)$ bez problémů uspějeme. Všimněme si, že v těchto bodech je parciální derivace $F_x$ nenulová.

Zhrňme pozorování (pro pouhé dva příklady):
Pro funkci $F(x, y)$ a bod $(a, b) \in E_2$ takový, že $F(a, b) = 0$, umíme najít funkci $y = f(x)$ splňující $F(x, f(x)) = 0$, pokud je $F_y(a, b) \neq 0$. V takovém případě umím i vypočíst $f'(x) = -F_x/F_y$.
Dokážeme, že takto to platí vždy, navíc rozšířené i na libovolné počty proměnných.
Poslední tvrzení o derivaci přitom je dobře zapamatovalné (a při pečlivém vnímání věcí i pochopitelné) z výrazu pro diferenciál:

$$0 = dF = F_x dx + F_y dy = (F_x + F_y f'(x))dx.$$

Obdobně pro implicitní výrazy $F(x, y, z) = 0$, kdy hledáme funkci $g(x, y)$ takovou, že $F(x, y, g(x, y)) = 0$. Např. graf funkce $f(x, y) = x^2 + y^2$ (rotační paraboloid) můžeme implicitně zadat rovnicí

$$0 = F(x, y, z) = z - x^2 - y^2.$$

Jaké dimenze se mohou/mají v problému vyskytovat obecně? Pokud bychom pro naší poslední $F$ chtěli najít křivku $c(x) = (c_1(x), c_2(x))$ v rovině takovou, že

$$F(x, c(x)) = F(x, c_1(x), c_2(x)) = 0,$$

pak to lze, ale výsledek nebude jednoznačný pro danou počáteční podmínku.

Obecný postup:
Jedna funkce $m+1$ proměnných zadává implicitně nadplochu v $\mathbb{R}^{m+1}$, kterou vyjadřujeme alespoň lokálně jako graf jedné funkce v $m$ proměnných.
Proto $n$ funkcí v $m+n$ proměnných bude zadávat průnik $n$ nadploch v $\mathbb{R}^{m+n}$, což je ve „většině" případů $m$–rozměrný objekt.
Příklad: dvě rovnice $f(x,y,z)=0$, $g(x,y,z)=0$ v $E_3$ zadávají (za nějaké podmínky na derivace) křivku v $E_3$.

Uvažujme proto spojitě diferencovatelné zobrazení

$$F = (f_1, \dots, f_n) : \mathbb{R}^{m+n} \to \mathbb{R}^n.$$

Jacobiho matice tohoto zobrazení bude mít $n$ řádků a $m+n$ sloupců a můžeme si ji symbolicky zapsat jako

$$D^1 F = (D^1_x F, D^1_y F) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_m} & \frac{\partial f_1}{\partial x_{m+1}} & \dots & \frac{\partial f_1}{\partial x_{m+n}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \dots & \frac{\partial f_n}{\partial x_m} & \frac{\partial f_n}{\partial x_{m+1}} & \dots & \frac{\partial f_n}{\partial x_{m+n}} \end{pmatrix},$$

kde $(x_1, \dots, x_{m+n}) \in \mathbb{R}^{m+n}$ zapisujeme jako $(x, y) \in \mathbb{R}^m \times \mathbb{R}^n$, $D^1_x F$ je matice s $n$ řádky a prvními $m$ sloupci v Jacobiho matici, zatímco $D^1_y F$ je čtvercová matice řádu $n$ se zbylými sloupci. Vícerozměrnou analogií k předchozí úvaze s nenulovou parciální derivací podle $y$ je požadavek, aby matice $D^1_y$ byla invertibilní.

**Theorem (Věta o implicitním zobrazení)**

*Nechť* $F : \mathbb{R}^{m+n} \to \mathbb{R}^n$ *je zobrazení třídy* $C^1$ *na otevřeném okolí bodu* $(a, b) \in \mathbb{R}^m \times \mathbb{R}^n = \mathbb{R}^{m+n}$*, ve kterém je* $F(a, b) = 0$ *a* $\det D^1_y F \neq 0$*. Potom existuje zobrazení třídy* $C^1$ $G : \mathbb{R}^m \to \mathbb{R}^n$ *definované na nějakém okolí* $U$ *bodu* $a \in \mathbb{R}^m$ *s obrazem* $G(U)$*, který obsahuje bod* $b$*, a takové, že* $F(x, G(x)) = 0$ *pro všechny* $x \in U$*. Navíc je Jacobiho matice* $D^1 G$ *zobrazení* $G$ *na okolí bodu* $a$ *zadána součinem matic*

$$D^1 G(x) = -(D^1_y F)^{-1}(x, G(x)) \cdot D^1_x F(x, G(x)).$$

Dokážeme pro nejjednodušší případ rovnice $F(x, y) = 0$ s funkcí $F$ dvou proměnných.
Rozšíříme funkci $F$ na

$$\tilde{F} : \mathbb{R}^2 \to \mathbb{R}^2, \quad (x, y) \mapsto (x, F(x, y)).$$

Jacobiho matice zobrazení $\tilde{F}$ je

$$D^1\tilde{F}(x, y) = \begin{pmatrix} 1 & 0 \\ F_x(x, y) & F_y(x, y) \end{pmatrix}.$$

Z předpokladu $F_y(a, b) \neq 0$ vyplývá, že totéž platí i na nějakém okolí bodu $(a, b)$ a tam je funkce $\tilde{F}$ invertibilní a $\tilde{F}^{-1}$ je jednoznačně definované a spojitě diferencovatelné.

Nechť $\pi : \mathbb{R}^2 \to \mathbb{R}$ je projekce na druhou souřadnici, pak

$$f(x) = \pi \circ \tilde{F}^{-1}(x, 0)$$

je spojitě diferencovatelná funkce.
Spočteme $F(x, f(x)) = F(x, \pi(\tilde{F}^{-1}(x, 0)))$.
Z definice $\tilde{F}(x, y) = (x, F(x, y))$ vidíme, že i její inverze má tvar $\tilde{F}^{-1}(x, y) = (x, \pi\tilde{F}^{-1}(x, y))$. Proto

$$F(x, f(x)) = \pi(\tilde{F}(x, \pi(\tilde{F}^{-1}(x, 0)))) = \pi(\tilde{F}(\tilde{F}^{-1}(x, 0))) = \pi(x, 0) = 0.$$

Tím máme dokázánu první část věty a zbývá spočíst derivaci funkce $f(x)$.

Tuto derivaci můžeme odečíst opět z věty o inverzním zobrazení pomocí matice $(D^1\tilde{F})^{-1}$.

$$\begin{pmatrix} 1 & 0 \\ F_x(x,y) & F_y(x,y) \end{pmatrix}^{-1} = (F_y(x,y))^{-1} \begin{pmatrix} F_y(x,y) & 0 \\ -F_x(x,y) & 1 \end{pmatrix}.$$

Dle definice $f(x) = \pi\tilde{F}^{-1}(x,0)$ nás z této matice zajímá první položka na druhém řádku, která je právě Jacobiho maticí $D^1 f$. V našem jednoduchém případě je to právě požadovaný skalár $-F_x(x,f(x))/F_y(x,f(x))$.
Důkaz věty je ukončen. Pro obecný případ je zcela stejný, jen pracujeme s násobením matic a vektorů místo skalárů.

# Gradient funkce

## Definition

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

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

nazývá **gradient funkce** $F$.
V technické a fyzikální literatuře se často zapisuje také jako grad $F$.

Rovnost $F(x_1, \dots, x_n) = b$ s pevnou hodnotou $b \in \mathbb{R}$ zadává podmnožinu $M \subset \mathbb{R}^n$, která mívá vlastnosti $(n-1)$–rozměrné nadplochy. Přesněji: pokud je vektor parciálních derivací nenulový, můžeme lokálně množinu $M$ popsat jako graf spojitě diferencovatelné funkce v $n-1$ proměnných.
Hovoříme v této souvislosti také o **úrovňových množinách** $M_b$.

Na derivacích křivek ležících v úrovňové množině $M_b$ se bude diferenciál $dF$ vždy vyčíslovat nulově:
$F(c(t)) = b$ pro všechna $t$, proto

$$\frac{d}{dt}F(c(t)) = dF(c'(t)) = 0.$$

Pro obecný vektor $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ je velikost příslušné směrové derivace funkce $F$:

$$|d_v F| = \left| \frac{\partial f}{\partial x_1} v_1 + \cdots + \frac{\partial f}{\partial x_n} v_n \right| = \cos\varphi \|D^1 F\| \|v\|$$

kde $\varphi$ je odchylka vektoru $v$ od gradientu $F$. Dokázali jsme:

**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 $D^1 F$ 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*

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

## Example (Model osvětlení 3D objektu)

Pro 2D povrch známe směr $v$ dopadu světla, tj. máme množinu $M$ zadanou implicitně rovnicí $F(x, y, z) = 0$ a vektor $v$. Intenzitu osvětlení bodu $P \in M$ pak definujme jako $I \cos \varphi$, kde $\varphi$ je úhel mezi normálou zadanou gradientem a vektorem opačným ke směru světla. (Znaménko říká, kterou stranu plochy osvětlujeme.)

Např. $v = (1, 1, -1)$ (tj. „šikmo dolů“) a $F(x, y, z) = x^2 + y^2 + z^2 - 1$. Pro bod $P = (x, y, z) \in M$

$$I(P) = \frac{\operatorname{grad} F \cdot v}{\|\operatorname{grad} F\| \|v\|} I_0 = \frac{-2x - 2y + 2z}{2\sqrt{3}} I_0.$$

Dle očekávání je plnou intenzitou $I_0$ osvětlen bod $P = \frac{1}{\sqrt{3}}(-1, -1, 1)$ na povrchu koule.

# Tečné a normálové prostory

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:
Afinní podprostor v $\mathbb{R}^{m+n}$ obsahující právě všechny tečny k $M$ bodem $P$ dán rovnicemi:

$$0 = \frac{\partial f_1}{\partial x_1}(P) \cdot (x_1 - a_1) + \dots + \frac{\partial f_1}{\partial x_n}(P) \cdot (x_{m+n} - a_{m+n})$$

$$\vdots$$

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

Tento podprostor se nazývá **tečný prostor** k (implicitně zadané) ploše $M$ v bodě $P$. **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 $D^1F$.
Spočtěme tečnu a normálový prostor ke kuželosečce v $\mathbb{R}^3$.
Uvažujme rovnici

$$0 = f(x, y, z) = z - \sqrt{x^2 + y^2}$$

kuželu s vrcholem v počátku a rovinu zadanou

$$0 = g(x, y, z) = z - 2x + y + 1.$$

Bod $P = (1, 0, 1)$ patří jak kuželu tak rovině a průnik $M$ těchto dvou ploch je křivka.

Její tečnou v bodě $P$ bude přímka zadaná rovnicemi

$$\begin{aligned}0 &= -\left.\frac{1}{2\sqrt{x^2+y^2}}2x\right|_{x=1,y=0} \cdot (x-1) - \left.\frac{1}{2\sqrt{x^2+y^2}}2y\right|_{x=1,y=0} \cdot y \\ &\quad + 1\cdot(z-1) \\ &= -x+z \\ 0 &= -2(x-1)+y+(z-1) = -2x+y+z+1\end{aligned}$$

zatímco rovina kolmá k naší křivce bodem $P$ bude parametricky dána výrazem

$$(1,0,1)+\tau(-1,0,1)+\sigma(-2,1,1)$$

s parametry $\tau$ a $\sigma$.

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 $F = (f_1, \dots, f_n) : \mathbb{R}^{m+n} \to \mathbb{R}^n$ 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 $\lambda_1, \dots, \lambda_n$ 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 $\lambda_i$ 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).