Matematika III – 4. přednáška
Funkce více proměnných: Zobrazení mezi euklidovskými prostory, inverzní zobrazení a implicitně definované zobrazení

Michal Bulant

Masarykova univerzita
Fakulta informatiky

14. 10. 2009

Obsah přednášky

Plán přednášky

1. Literatura
2. Zobrazení mezi euklidovskými prostory
   - Zobrazení a transformace
   - "Chain Rule"
   - Věta o inverzním zobrazení
   - Implicitně zadaná zobrazení
3. Tečny a normály k implicitně zadaným plochám
   - Gradient funkce
   - Tečné a normálové prostory
4. Vázané extrémy
   - Metoda Lagrangeových multiplikátorů
   - Speciální optimalizační metody

Doporučené zdroje

- Martin Panák, Jan Slovák, Drsná matematika, e-text.

Doporučené zdroje

- Martin Panák, Jan Slovák, Drsná matematika, e-text.
- Zuzana Došlá, Ondřej Došlý, Diferenciální počet funkcí více proměnných, MU Brno, 2006, 150 s.

Doporučené zdroje

- Martin Panák, Jan Slovák, Drsná matematika, e-text.
- Zuzana Došlá, Ondřej Došlý, Diferenciální počet funkcí více proměnných, MU Brno, 2006, 150 s.
- Zuzana Došlá, Roman Plch, Petr Sojka, Diferenciální počet funkcí více proměnných s programem Maple, MU Brno, 1999, 273 s. (příp. `http://www.math.muni.cz/~plch/mapm`).

Doporučené zdroje

- Martin Panák, Jan Slovák, Drsná matematika, e-text.
- Zuzana Došlá, Ondřej Došlý, Diferenciální počet funkcí více proměnných, MU Brno, 2006, 150 s.
- Zuzana Došlá, Roman Plch, Petr Sojka, Diferenciální počet funkcí více proměnných s programem Maple, MU Brno, 1999, 273 s. (příp. `http://www.math.muni.cz/~plch/mapm`).
- *Předmětové záložky v IS MU*

Obsah první písemky

- určování limit, resp. důkaz neexistence

Obsah první písemky

- určování limit, resp. důkaz neexistence
- výpočty parciálních a směrových derivací

Obsah první písemky

- určování limit, resp. důkaz neexistence
- výpočty parciálních a směrových derivací
- použití diferenciálu – aproximace, tečná nadrovina

Obsah první písemky

- určování limit, resp. důkaz neexistence
- výpočty parciálních a směrových derivací
- použití diferenciálu – aproximace, tečná nadrovina
- využití Taylorovy věty a Hessiánu pro aproximaci

Obsah první písemky

- určování limit, resp. důkaz neexistence
- výpočty parciálních a směrových derivací
- použití diferenciálu – aproximace, tečná nadrovina
- využití Taylorovy věty a Hessiánu pro aproximaci
- Jacobián zobrazení a jeho inverze (včetně důkazu existence)

Obsah první písemky

- určování limit, resp. důkaz neexistence
- výpočty parciálních a směrových derivací
- použití diferenciálu – aproximace, tečná nadrovina
- využití Taylorovy věty a Hessiánu pro aproximaci
- Jacobián zobrazení a jeho inverze (včetně důkazu existence)
- určování lokálních a globálních extrémů (vázané extrémy nikoliv).

Plán přednášky

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 $f_1, \dots, f_m$.

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

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

funkcí $f_i : E_n \to \mathbb{R}$. Řekneme, že F je *diferencovatelné* nebo ***spojitě* diferencovatelné zobrazení**, jestliže tuto vlastnost mají všechny funkce $f_1, \ldots, f_m$.

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**.

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 v E_2 je přechod mezi polárními a kartézkými souřadnicemi:

$$[r, \theta] \mapsto [r \cos\theta, r \sin\theta]$$

s inverzí

$$[x, y] \mapsto [\sqrt{x^2 + y^2}, \operatorname{arctg} \frac{y}{x}],\ [0, y] \mapsto [y, \frac{\pi}{2} \operatorname{sgn} y].$$

Lineární zobrazení $\mathrm{d}f_i(x) : E^n \to \mathbb{R}$ lineárně aproximují přírůstky f_i.

Definice

$$D^1F(x) = \begin{pmatrix} \mathrm{d}f_1(x) \\ \mathrm{d}f_2(x) \\ \vdots \\ \mathrm{d}f_m(x) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \dots & \frac{\partial f_m}{\partial x_n} \end{pmatrix} (x)$$

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

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

Důsledek Věty o existenci diferenciálu pro funkce n proměnných je:

Věta

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

Diferenciál složeného zobrazení

Věta ("Chain rule")

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 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 souřadnic transformací $F : E_2 \to E_2$, kterou v souřadnicích $[x, y]$ a $[r, \varphi]$ zapíšeme:

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

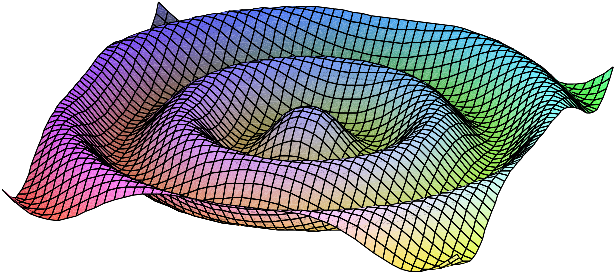
Polární souřadnice vzniknou z kartézských souřadnic transformací $F : E_2 \to E_2$, kterou v souřadnicích $[x, y]$ a $[r, \varphi]$ zapíšeme:

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

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:



Chceme-li vypočítat derivaci funkce zadané parametricky v kartézských souřadnicích, využijeme větu o derivaci složeného zobrazení $g \circ F : E_2 \to \mathbb{R}$:

$$D^1(g \circ F)(x) = D^1 g(F(x)) \circ D^1 F(x) =$$

$$= \begin{pmatrix} \frac{\partial g}{\partial r} & \frac{\partial g}{\partial \varphi} \end{pmatrix} \begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \varphi}{\partial x} & \frac{\partial \varphi}{\partial y} \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{\partial g}{\partial r}(r,\varphi)\frac{\partial r}{\partial x}(x,y) + \frac{\partial g}{\partial \varphi}(r,\varphi)\frac{\partial \varphi}{\partial x}(x,y) \\ \frac{\partial g}{\partial r}(r,\varphi)\frac{\partial r}{\partial y}(x,y) + \frac{\partial g}{\partial \varphi}(r,\varphi)\frac{\partial \varphi}{\partial y}(x,y) \end{pmatrix}.$$

Tedy

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

a podobně

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

Věta o inverzní funkci jedné proměnné – připomenutí

Pokud k dané funkci $f : \mathbb{R} \to \mathbb{R}$ inverzní funkce f^{-1} existuje (nezaměňujme značení s funkcí $x \mapsto (f(x))^{-1}$), pak je dána jednoznačně kterýmkoliv ze vztahů $f^{-1} \circ f = \mathrm{id}_{\mathbb{R}}, \quad f \circ f^{-1} = \mathrm{id}_{\mathbb{R}},$.

Věta o inverzní funkci jedné proměnné – připomenutí

Pokud k dané funkci $f : \mathbb{R} \to \mathbb{R}$ inverzní funkce f^{-1} existuje (nezaměňujme značení s funkcí $x \mapsto (f(x))^{-1}$), pak je dána jednoznačně kterýmkoliv ze vztahů $f^{-1} \circ f = \mathrm{id}_{\mathbb{R}}, \quad f \circ f^{-1} = \mathrm{id}_{\mathbb{R}},$. Pokud bychom věděli, že pro diferencovatelnou funkci f je i f^{-1} diferencovatelná, vztah pro derivaci složené funkce nám (pro $y = f(x)$) dává

$$1 = (\mathrm{id})'(x) = (f^{-1} \circ f)'(x) = (f^{-1})'(f(x)) \cdot f'(x)$$

a tedy přímo víme formuli (zjevně $f'(x)$ v takovém případě nemůže být nulové) $(f^{-1})'(f(x)) = \frac{1}{f'(x)}$.

Věta

Je-li f diferencovatelná funkce na okolí bodu x_0 a $f'(x_0) \neq 0$, pak existuje na nějakém okolí bodu $y_0 = f(x_0)$ funkce f^{-1} inverzní k f a platí vztah

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)}.$$

Věta o inverzním zobrazení

Věta

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

Věta o inverzním zobrazení

Věta

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

Princip důkazu: Z pravidla pro derivování složené funkce vyplývá, že pokud diferencovatelná inverze existuje, pak musí být její Jacobiho matice inverzí k původní Jacobiho matici (srovnejte s případem 1 proměnné). Důkaz poměrně komplikovaným způsobem vyvozuje, že díky invertovatelnosti Jacobiho matice existuje diferencovatelná inverze.

Příklad

Rozhodněte, zda zobrazení $F = (f, g) : \mathbb{R}^2 \to \mathbb{R}^2$ definované po souřadnicích

$$f(x, y) = xy, g(x, y) = \frac{x}{y}$$

je prosté v okolí bodu $[2, 1]$. V kladném případě určete Jacobiho matici inverzního zobrazení v bodě $F(2, 1)$.

Věta o implicitní funkci – neformální připomenutí

Pokud máme zadánu funkci $f(x)$ vzorcem $y = f(x)$, hovoříme o jejím explicitním zadání. Obecnějším zadáním funkce je rovnice $F(x, y) = 0$, kde závislá proměnná y představuje *neznámou* funkci. Pokud tuto rovnici nelze (nebo to nepotřebujeme) vyřešit vzhledem k y, pak hovoříme o funkci zadané implicitně. Avšak i v tomto obecnějším případě budeme schopni vypočítat $y'(x)$ (aniž bychom znali explicitní vzorec pro $y(x)$), a to pomocí pravidla pro derivaci složené funkce.

Příklad

Rovnice $y^2 = x$ definuje dvě diferencovatelné funkce

$$y_1 = \sqrt{x}, \qquad y_2 = -\sqrt{x}$$

Věta o implicitní funkci – neformální připomenutí

Pokud máme zadánu funkci $f(x)$ vzorcem $y = f(x)$, hovoříme o jejím explicitním zadání. Obecnějším zadáním funkce je rovnice $F(x, y) = 0$, kde závislá proměnná y představuje *neznámou* funkci. Pokud tuto rovnici nelze (nebo to nepotřebujeme) vyřešit vzhledem k y, pak hovoříme o funkci zadané implicitně. Avšak i v tomto obecnějším případě budeme schopni vypočítat $y'(x)$ (aniž bychom znali explicitní vzorec pro $y(x)$), a to pomocí pravidla pro derivaci složené funkce.

Příklad

Rovnice $y^2 = x$ definuje dvě diferencovatelné funkce

$$y_1 = \sqrt{x}, \qquad y_2 = -\sqrt{x}$$

Při derivování implicitně zadaných funkcí obsahuje výsledná derivace y' jak proměnnou x tak proměnnou y (na rozdíl od *běžného* derivování funkce, kdy je ve výsledku pouze proměnná x).

Věta o implicitní funkci

Pro jednoduchost vyložíme ideu v rovině E_2:

Věta o implicitní funkci

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

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

Věta o implicitní funkci

Pro jednoduchost vyložíme ideu v rovině E_2:
Pro spojitě diferencovatelnou funkci $F(x, y) : E_2 \to \mathbb{R}$ 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 prvním případě je (při $b \neq 0$) předpisem zadaná funkce

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

pro všechna x.

Věta o implicitní funkci

Pro jednoduchost vyložíme ideu v rovině E_2:
Pro spojitě diferencovatelnou funkci $F(x, y) : E_2 \to \mathbb{R}$ 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 prvním případě je (při $b \neq 0$) předpisem zadaná funkce

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

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

Body $[s \pm r, t]$ také vyhovují rovnici kružnice, platí v nich ale $F_y(s \pm r, t) = 0$, což znamená, že tečna ke kružnici v těchto bodech je rovnoběžná s osou y. V těchto bodech neumíme najít okolí, na němž by kružnice byla popsána jako funkce $y = f(x)$.

Body $[s \pm r, t]$ také vyhovují rovnici kružnice, platí v nich ale $F_y(s \pm r, t) = 0$, což znamená, že tečna ke kružnici v těchto bodech je rovnoběžná s osou y. V těchto bodech neumíme najít okolí, na němž by kružnice byla popsána jako funkce $y = f(x)$.

Navíc umíme spočítat i **derivace**:

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

Body $[s \pm r, t]$ také vyhovují rovnici kružnice, platí v nich ale $F_y(s \pm r, t) = 0$, což znamená, že tečna ke kružnici v těchto bodech je rovnoběžná s osou y. V těchto bodech neumíme najít okolí, na němž by kružnice byla popsána jako funkce $y = f(x)$. Navíc umíme spočítat 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.

Shrňme pozorování (pro pouhé dva příklady):

Shrňme 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 umíme i vypočíst $f'(x) = -F_x/F_y$. Z následující věty plyne, že takto to platí vždy, navíc rozšířené i na libovolné počty proměnných.

Věta (O implicitní funkci)

Nechť $F : E_{n+1} \to \mathbb{R}$ *je spojitě diferencovatelná funkce na otevřeném okolí bodu* $[x^*, y^*] \in E_n \times \mathbb{R}$, *ve kterém je* $F(x^*, y^*) = 0$ *a* $\frac{\partial F}{\partial y}(x^*, y^*) \neq 0$. *Potom existuje spojitá funkce* $f : E_n \to \mathbb{R}$ *definovaná na nějakém okolí* U *bodu* $x^* \in E_n$ *taková, že* $F(x, f(x)) = 0$ *pro všechny* $x \in U$.
Navíc má funkce f *v okolí bodu* x^* *parciální derivace splňující*

$$\frac{\partial f}{\partial x_i}(x) = -\frac{\frac{\partial F}{\partial x_i}(x, f(x))}{\frac{\partial F}{\partial y}(x, f(x))}.$$

Věta (O implicitní funkci)

Nechť $F : E_{n+1} \to \mathbb{R}$ *je spojitě diferencovatelná funkce na otevřeném okolí bodu* $[x^*, y^*] \in E_n \times \mathbb{R}$*, ve kterém je* $F(x^*, y^*) = 0$ *a* $\frac{\partial F}{\partial y}(x^*, y^*) \neq 0$*. Potom existuje spojitá funkce* $f : E_n \to \mathbb{R}$ *definovaná na nějakém okolí* U *bodu* $x^* \in E_n$ *taková, že* $F(x, f(x)) = 0$ *pro všechny* $x \in U$*.*
Navíc má funkce f *v okolí bodu* x^* *parciální derivace splňující*

$$\frac{\partial f}{\partial x_i}(x) = -\frac{\frac{\partial F}{\partial x_i}(x, f(x))}{\frac{\partial F}{\partial y}(x, f(x))}.$$

Poslední tvrzení o derivaci přitom je dobře zapamatovalné (i pochopitelné) z výrazu pro diferenciál:

$$0 = \mathrm{d}F = F_x \,\mathrm{d}x + F_y \,\mathrm{d}y = (F_x + F_y f'(x))\,\mathrm{d}x.$$

Příklad

Určete lokální extrémy funkce $z = f(x, y)$, která je určena implicitně rovnicí $F(x, y, z) = x^2 + y^2 + z^2 - xz - \sqrt{2}yz = 1$.

Řešení

Derivováním rovnosti podle x a y dostáváme:

$$2x + 2z \cdot z_x - z - x \cdot z_x - \sqrt{2}yz_x = 0$$
$$2y + 2z \cdot z_y - x \cdot z_y - \sqrt{2}z - \sqrt{2}yz_y = 0,$$

odkud vyjádříme

$$z_x = \frac{z - 2x}{2z - x - \sqrt{2}y}, \qquad z_y = \frac{\sqrt{2}z - 2y}{2z - x - \sqrt{2}y}.$$

Řešení (pokr.)

Stacionární body musí splňovat: $z_x = 0$, $z_y = 0$, tj. $z = 2x = \sqrt{2}y$, a tedy $y = \sqrt{2}x$. Dosazením do původní rovnice dostáváme stacionární body $[1, \sqrt{2}, 2]$ a $[-1, -\sqrt{2}, -2]$. V těchto bodech je $F_z \neq 0$ (je to zároveň jmenovatel všech zde vystupujících zlomků), proto je v jejich okolí implicitně určena jistá funkce $z = f(x, y)$. Dalším derivováním implicitní rovnice vypočteme parciální derivace f 2. řádu:

$$z_{xx} = -\frac{2}{2z - x - \sqrt{2}y}, \quad z_{xy} = 0, \quad z_{yy} = -\frac{2}{2z - x - \sqrt{2}y}.$$

Řešení (pokr.)

Stacionární body musí splňovat: $z_x = 0$, $z_y = 0$, tj. $z = 2x = \sqrt{2}y$, a tedy $y = \sqrt{2}x$. Dosazením do původní rovnice dostáváme stacionární body $[1, \sqrt{2}, 2]$ a $[-1, -\sqrt{2}, -2]$. V těchto bodech je $F_z \neq 0$ (je to zároveň jmenovatel všech zde vystupujících zlomků), proto je v jejich okolí implicitně určena jistá funkce $z = f(x, y)$. Dalším derivováním implicitní rovnice vypočteme parciální derivace f 2. řádu:

$$z_{xx} = -\frac{2}{2z - x - \sqrt{2}y}, \quad z_{xy} = 0, \quad z_{yy} = -\frac{2}{2z - x - \sqrt{2}y}.$$

Ve stacionárních bodech je Hf negativně, resp. pozitivně definitní, proto zde nastávají lokální maximum, resp. minimum funkce f.

Nejobecnější případ, kdy definujeme implicitně zadané zobrazení, popisuje následující věta (v případě $n=1$ kopíruje větu o implicitní funkci).

Nejobecnější případ, kdy definujeme implicitně zadané zobrazení, popisuje následující věta (v případě n=1 kopíruje větu o implicitní funkci).

Věta (O implicitním zobrazení)

Nechť $F : E_{m+n} \to E_n$ je spojitě diferencovatelné zobrazení na otevřeném okolí bodu $[x^*, y^*] \in E_m \times E_n = E_{m+n}$, v němž platí $F(x^*, y^*) = \mathbf{0}$ a $\det D^1_y F \neq 0$. Potom existuje spojitě diferencovatelné zobrazení $G : E_m \to E_n$ definované na nějakém okolí U bodu $x^* \in E_m$ s obrazem $G(U)$, který obsahuje bod y^*, 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 x^* zadána součinem matic

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

Plán přednášky

Gradient funkce

Definice

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

$$\mathrm{d}f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\right)$$

nazývá **gradient funkce** f.

V technické a fyzikální literatuře se často zapisuje také jako grad f.

Gradient funkce

Definice

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

$$\mathrm{d}f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right)$$

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

Rovnost $f(x_1, \ldots, x_n) = b$ s pevnou hodnotou $b \in \mathbb{R}$ zadává podmnožinu $M \subset E_n$, která mívá vlastnosti $(n-1)$–rozměrné nadplochy.

Gradient funkce

Definice

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

$$df = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)$$

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 $M \subset E_n$, 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.

Gradient funkce

Definice

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

$$\mathrm{d}f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)$$

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 $M \subset E_n$, 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 (analogie vrstevnic v př. $n = 2$).

Na derivacích křivek ležících v úrovňové množině M_b se bude diferenciál df vždy vyčíslovat nulově:

$f(c(t)) = b$ pro všechna t, proto

$$\frac{d}{dt} f(c(t)) = \mathrm{d}f(c'(t)) = 0.$$

Na derivacích křivek ležících v úrovňové množině M_b se bude diferenciál $\mathrm{d}f$ vždy vyčíslovat nulově:
$f(c(t)) = b$ pro všechna t, proto

$$\frac{d}{dt} f(c(t)) = \mathrm{d}f(c'(t)) = 0.$$

Pro obecný vektor $v = (v_1, \dots, v_n) \in E_n$ je velikost příslušné směrové derivace funkce f:

$$|\,\mathrm{d}_v f| = \left| \frac{\partial f}{\partial x_1} v_1 + \dots + \frac{\partial f}{\partial x_n} v_n \right| = \langle \mathrm{d}f, v \rangle = \cos\varphi \|\,\mathrm{d}f\| \|v\|$$

kde φ je odchylka vektoru v od gradientu funkce f. Dokázali jsme:

Na derivacích křivek ležících v úrovňové množině M_b se bude diferenciál $\mathrm{d}f$ vždy vyčíslovat nulově:
$f(c(t)) = b$ pro všechna t, proto

$$\frac{d}{dt} f(c(t)) = \mathrm{d}f(c'(t)) = 0.$$

Pro obecný vektor $v = (v_1, \dots, v_n) \in E_n$ je velikost příslušné směrové derivace funkce f:

$$|\,\mathrm{d}_v f| = \left| \frac{\partial f}{\partial x_1} v_1 + \cdots + \frac{\partial f}{\partial x_n} v_n \right| = \langle \mathrm{d}f, v \rangle = \cos\varphi \|\,\mathrm{d}f\| \|v\|$$

kde φ je odchylka vektoru v od gradientu funkce f. Dokázali jsme:

Věta

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 $\mathrm{d}f$ je určena ortogonálním doplňkem ke gradientu.

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

Věta

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

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

Příklad (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 $I_0 \cos\varphi$, kde φ je úhel mezi normálou zadanou gradientem a vektorem opačným ke směru světla. (Znaménko říká, kterou stranu plochy osvětlujeme, I_0 je tzv. svítivost.)

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 $F = (f_1, \dots, f_n) : E_{m+n} \to E_n$ 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 $G : E_m \to E_n$. 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:

Tečné a normálové prostory

Obecné dimenze: funkce $F = (f_1, \dots, f_n) : E_{m+n} \to E_n$ 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 $G : E_m \to E_n$. 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 E_{m+n} obsahující právě všechny tečny k M bodem P dán rovnicemi:

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

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.

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.

Příklad

Spočtěme tečnu a normálový prostor ke kuželosečce v E_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.

Příklad (pokr.)

Její tečnou v bodě P bude přímka zadaná rovnicemi

Příklad (pokr.)

Její tečnou v bodě P bude přímka zadaná rovnicemi

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

Příklad (pokr.)

Její tečnou v bodě P bude přímka zadaná rovnicemi

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

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 τ a σ.

Plán přednášky

Již dříve jsme se zabývali úlohou nalézt absolutní extrém dané funkce na (uzavřené) množině, což vedlo na vyšetření lokálních extrémů funkce na hranici této množiny. Jinými slovy, na hledání extrémů funkce v bodech, které jsou *vázány* nějakou další podmínkou.

Již dříve jsme se zabývali úlohou nalézt absolutní extrém dané funkce na (uzavřené) množině, což vedlo na vyšetření lokálních extrémů funkce na hranici této množiny. Jinými slovy, na hledání extrémů funkce v bodech, které jsou *vázány* nějakou další podmínkou.

Ukážeme nejprve názorně graficky na případu funkcí dvou proměnných obecnou metodu.

Příklad

Určete lokální extrémy funkce $f(x, y) = x^2 y$ na množině M dané implicitně rovnicí $5x^2 + 2y^2 = 14$.

Metoda Lagrangeových multiplikátorů

V předchozím příkladu jsme viděli, že normálový vektor (tj. gradient) funkce, k níž hledáme extrém, musí být ve vyšetřovaném bodě prvkem normálového prostoru k ploše (v témže bodě). Toto samozřejmě platí i obecně.

Metoda Lagrangeových multiplikátorů

V předchozím příkladu jsme viděli, že normálový vektor (tj. gradient) funkce, k níž hledáme extrém, musí být ve vyšetřovaném bodě prvkem normálového prostoru k ploše (v témže bodě). Toto samozřejmě platí i obecně.

Pokud je M ve všech svých bodech grafem hladkého zobrazení, 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

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

Metoda Lagrangeových multiplikátorů

V předchozím příkladu jsme viděli, že normálový vektor (tj. gradient) funkce, k níž hledáme extrém, musí být ve vyšetřovaném bodě prvkem normálového prostoru k ploše (v témže bodě). Toto samozřejmě platí i obecně.
Pokud je M ve všech svých bodech grafem hladkého zobrazení, 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

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

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.

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 $F(x_1, \dots, x_{m+n}) = \mathbf{0}$ ($F : E_{m+n} \to E_n$).

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**:

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 $F(x_1, \dots, x_{m+n}) = \mathbf{0}$ ($F : E_{m+n} \to E_n$).
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**:

Věta

Nechť $F = (f_1, \dots, f_n) : E_{m+n} \to E_n$ je spojitě diferencovatelná v okolí bodu P, $F(P) = 0$ a M je zadána implicitně rovnicí $F(x_1, \dots, x_{m+n}) = 0$, přičemž hodnost matice D^1F v bodě P je n. Pak P je stacionárním bodem spojitě diferencovatelné funkce $h : E_{m+n} \to \mathbb{R}$ právě, když existují reálné parametry $\lambda_1, \dots, \lambda_n$ takové, že

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

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 neznámé 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 bude izolovaným bodem).

Absolutní extrémy

Výklad o vázaných extrémech jsme začali tím, že pro nalezení absolutních extrémů funkce na kompaktní množině často potřebujeme vyšetření extrémů na množině bodů vázaných nějakou podmínkou.

Absolutní extrémy

Výklad o vázaných extrémech jsme začali tím, že pro nalezení absolutních extrémů funkce na kompaktní množině často potřebujeme vyšetření extrémů na množině bodů vázaných nějakou podmínkou.
Ilustrujme si to na příkladu:

Příklad

Maximalizujte $f(x, y) = 2x + y$ za podmínky $\frac{x^2}{4} + y^2 \leq 1$.

Řešení

Množina určená vazební podmínkou je uzavřená a ohraničená, proto zde nabývá jakákoliv spojitá funkce svých extrémů, a to buď ve stacionárních bodech nebo na hranici. Snadno se ale přesvědčíme ($df(x,y) = (2,1)$), že uvnitř množiny extrémy nejsou. Proto maximalizujeme funkci f za podmínky $\frac{x^2}{4} + y^2 = 1$.

Řešení

Množina určená vazební podmínkou je uzavřená a ohraničená, proto zde nabývá jakákoliv spojitá funkce svých extrémů, a to buď ve stacionárních bodech nebo na hranici. Snadno se ale přesvědčíme ($df(x,y) = (2,1)$), že uvnitř množiny extrémy nejsou. Proto maximalizujeme funkci f za podmínky $\frac{x^2}{4} + y^2 = 1$.
Sestrojíme Lagrangeovu funkci
$L(x, y, \lambda) = 2x + y - \lambda(\frac{x^2}{4} + y^2 - 1)$. Pak dostáváme:

$$0 = L_x = 2 - \lambda\frac{x}{2}$$
$$0 = L_y = 1 - 2\lambda y$$
$$0 = \frac{x^2}{4} + y^2 - 1.$$

Řešení

Množina určená vazební podmínkou je uzavřená a ohraničená, proto zde nabývá jakákoliv spojitá funkce svých extrémů, a to buď ve stacionárních bodech nebo na hranici. Snadno se ale přesvědčíme ($\mathrm{d}f(x,y) = (2,1)$), že uvnitř množiny extrémy nejsou. Proto maximalizujeme funkci f za podmínky $\frac{x^2}{4} + y^2 = 1$.
Sestrojíme Lagrangeovu funkci
$L(x,y,\lambda) = 2x + y - \lambda(\frac{x^2}{4} + y^2 - 1)$. Pak dostáváme:

$$\begin{aligned} 0 &= L_x = 2 - \lambda\frac{x}{2} \\ 0 &= L_y = 1 - 2\lambda y \\ 0 &= \frac{x^2}{4} + y^2 - 1. \end{aligned}$$

Odtud snadno $x = \frac{4}{\lambda}$, $y = \frac{1}{2\lambda}$, a tedy $\lambda = \frac{\sqrt{17}}{2}, x = \frac{8}{\sqrt{17}}, y = \frac{1}{\sqrt{17}}$ (resp. $\lambda = -\frac{\sqrt{17}}{2}, x = -\frac{8}{\sqrt{17}}, y = -\frac{1}{\sqrt{17}}$ pro minimum).

Speciální optimalizační metody

Zmiňme se jen ve stručnosti o speciálních optimalizačních technikách, které se v dnešní praxi používají. Zájemce o bližší seznámení s nimi můžeme odkázat na další předměty MU, např.:

- Optimalizace – `PřF: M0160` (jaro)
- Optimalizace – `PV027` (jaro)
- Lineární programování – `PřF: M4110` (jaro)
- Matematické programování – `PřF: M5170` (podzim)

Metoda gradientu

Již dříve jsme zmínili, že funkce nejrychleji roste ve směru gradientu (a nejrychleji klesá ve směru opačném) – proto je přirozené se při hledání maxima vydat z daného bodu ve směru gradientu (analogie *chození do kopce nejprudším svahem*). Otázka je, jak dlouho "jít" a jak často gradient počítat.
Iterace:

$$x_{n+1} = x_n + \gamma_n \operatorname{grad} f(x_n),$$

pro dostatečně malé γ_n, aby $f(x_{n+1}) > f(x_n)$.

Metoda gradientu

Již dříve jsme zmínili, že funkce nejrychleji roste ve směru gradientu (a nejrychleji klesá ve směru opačném) – proto je přirozené se při hledání maxima vydat z daného bodu ve směru gradientu (analogie *chození do kopce nejprudším svahem*). Otázka je, jak dlouho "jít" a jak často gradient počítat.
Iterace:

$$x_{n+1} = x_n + \gamma_n \operatorname{grad} f(x_n),$$

pro dostatečně malé γ_n, aby $f(x_{n+1}) > f(x_n)$.
Problémy:

- náročný opakovaný výpočet γ,
- velký počet iterací v případě velmi různorodé křivosti v různých směrech; např *Rosenbrockova banánová funkce* – $f(x, y) = (1 - x)^2 + 100(y - x^2)^2$.

Newtonova optimalizační metoda

Newtonova metoda je dobře známý numerický postup pro nalezení kořenů dané reálné funkce f . Známe-li bod x_0 "rozumně" blízko kořene, zkonstruujeme v bodě $[x_0, f(x_0)]$ tečnu ke grafu funkce f a za bod x_1 zvolíme průsečík tečny s osou x. Tento postup opakujeme. Snadno je vidět, že platí rekurentní vztah

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Newtonova optimalizační metoda

Newtonova metoda je dobře známý numerický postup pro nalezení kořenů dané reálné funkce f . Známe-li bod x_0 "rozumně" blízko kořene, zkonstruujeme v bodě $[x_0, f(x_0)]$ tečnu ke grafu funkce f a za bod x_1 zvolíme průsečík tečny s osou x. Tento postup opakujeme. Snadno je vidět, že platí rekurentní vztah

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Tento postup např. poskytuje efektivní postup pro výpočet $\sqrt{2}$ s libovolnou přesností; pokud bychom ale chtěli hledat řešení rovnice $x^{1/3} = 0$, tak snadno vidíme, že metoda diverguje, ať začneme jakkoli blízko 0.

Při hledání extrémů funkcí (i více proměnných) může být Newtona metoda využita pro nalezení stacionárních bodů – v nich musí být derivace nulová, proto jde vlastně o nalezení kořenů derivace iterativním postupem

$$x_{n+1} = x_n - (Hf(x_n))^{-1} \cdot \operatorname{grad} f(x_n).$$

Při hledání extrémů funkcí (i více proměnných) může být Newtona metoda využita pro nalezení stacionárních bodů – v nich musí být derivace nulová, proto jde vlastně o nalezení kořenů derivace iterativním postupem

$$x_{n+1} = x_n - (Hf(x_n))^{-1} \cdot \operatorname{grad} f(x_n).$$

Výpočet inverze Hessiánu je časově náročná operace, proto se často místo toho využívá

- metoda sdružených gradientů pro řešení příslušné soustavy,
- různých tzv. *kvazi*-newtonovských metod, využívajících pouze přibližného Hessiánu (např. BFGS).

Lineární programování

Úloha lineárního programování

Pro daná $c \in \mathbb{R}^n$ řeší lineární programování úlohu optimalizovat (tj. maximalizovat nebo minimalizovat) lineární *účelovou funkci*

$$f(x) = c \cdot x = c_1 x_1 + \cdots + c_n x_n$$

Lineární programování

Úloha lineárního programování
Pro daná $c \in \mathbb{R}^n$ řeší lineární programování úlohu optimalizovat (tj. maximalizovat nebo minimalizovat) lineární *účelovou funkci*

$$f(x) = c \cdot x = c_1 x_1 + \cdots + c_n x_n$$

za daných (lineárních) omezení

$$\begin{aligned} a_1 \cdot x &\leq b_1 \\ &\dots \\ a_k \cdot x &\leq b_k \\ a_{k+1} \cdot x &= b_{k+1} \\ &\dots \\ a_\ell \cdot x &= b_\ell \end{aligned}$$

Lineární programování

Lze ukázat, že každou (rozumnou) úlohu lineárního programování lze převést na tzv. *standardní úlohu* tvaru

$$\text{maximalizovat} \quad f(x) = c \cdot x$$

za podmínek

$$a_1 \cdot x \le b_1$$
$$\dots$$
$$a_k \cdot x \le b_k,$$

kde $x_1 \ge 0, \dots, x_n \ge 0$.

Simplexová metoda

Standardní úlohu řeší klasická Simplexová metoda (George Dantzig, 1947).

Simplexová metoda

Standardní úlohu řeší klasická Simplexová metoda (George Dantzig, 1947).
Úvodní fáze spočívá v nalezení nějakého vrcholu na polytopu (zobecnění polyedru na více dimenzí), který je tvořen body vyhovujícími podmínkám. V dalších krocích postupuje po hranách do vrcholů s vyšší hodnotou účelové funkce.

Simplexová metoda

Standardní úlohu řeší klasická Simplexová metoda (George Dantzig, 1947).
Úvodní fáze spočívá v nalezení nějakého vrcholu na polytopu (zobecnění polyedru na více dimenzí), který je tvořen body vyhovujícími podmínkám. V dalších krocích postupuje po hranách do vrcholů s vyšší hodnotou účelové funkce.
Sice je ukázán příklad podmínek, kdy simplexová metoda projde nešikovně všech 2^n vrcholů (jde o příklad zborcené n-rozměrné krychle), a tedy metoda je v nejhorším případě exponenciální, ale v praxi je obvykle pozoruhodně úspěšná (kolem roku 2000 bylo dokázáno, že očekávaný čas běhu na náhodném vstupu je polynomiální).

Příklad

Maximalizujte $f = 2x - 3y + 4z$ za podmínek

$$\begin{aligned} 4x - 3y + z &\leq 3 \\ x + y + z &\leq 10 \\ 2x + y - z &\leq 10 \\ x \geq 0, y \geq 0, z &\geq 0. \end{aligned}$$