1. přednáška 1. října 2007

Kapitola 1. Metrické prostory.

Definice MP, izometrie. *Metrický prostor* je struktura formalizující jev vzdálenosti. Je to dvojice (M, d) složená z množiny M a funkce dvou proměnných

$$d : M \times M \to \mathbf{R},$$

tak zvané *metriky*, splňující tři axiomy:

a) $d(x, y) \geq 0$ (nezápornost) a $d(x, y) = d(y, x)$ (symetrie),
b) $d(x, y) = 0 \iff x = y$ a
c) $d(x, y) \leq d(x, z) + d(z, y)$ (trojúhelníková nerovnost).

Nezápornost metriky v a) se nemusí požadovat, plyne z axiomů b) a c).

Izometrie metrických prostorů (M, d) a (N, e) je bijekce $f : M \to N$ zachovávající vzdálenosti: $d(x, y) = e(f(x), f(y))$ pro všechny $x, y \in M$. Existuje-li taková bijekce, jsou prostory (M, d) a (N, e) *izometrické*, prakticky nerozlišitelné, izomorfní.

Příklady metrických prostorů. V následujících příkladech se axiomy a) a b) ověří obvykle snadno (nicméně viz úlohu 7). Dokázat trojúhelníkovou nerovnost bývá často obtížnější, viz závěrečné úlohy.

Příklad 1. $M = \mathbf{R}^n$ a $p \geq 1$ je reálné číslo. Na M definujeme metriky $d_p(x, y)$ vztahem

$$d_p(x, y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{1/p}$$

$(x = (x_1, x_2, \ldots, x_n), y = (y_1, y_2, \ldots, y_n))$. Pro $n = 1$ dostáváme klasickou metriku $|x - y|$ na $\mathbf{R}$ a pro $p = 2, n \geq 2$ *euklidovskou metriku*

$$d_2(x, y) = \|x - y\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}.$$

Euklidovskými prostory rozumíme metrické prostory $(\mathbf{R}^n, d_2)$ s euklidovskou metrikou. Pro $p = 1, n \geq 2$ dostáváme *pošťáckou metriku*

$$d_1(x, y) = \sum_{i=1}^{n} |x_i - y_i|$$

a pro $p \to \infty$ *maximovou metriku*

$$d_\infty(x, y) = \max_{1 \leq i \leq n} |x_i - y_i|.$$

Příklad 2. Za M vezmeme množinu všech omezených funkcí $f : X \to \mathbf{R}$ definovaných na množině X. Na M pak máme *supremovou metriku*

$$d(f, g) = \sup_{x \in X} |f(x) - g(x)|.$$

Pokud $M = \mathcal{C}[a,b]$ (množina reálných funkcí definovaných a spojitých na intervalu $[a,b]$), supremum se nabývá a máme *maximovou metriku*

$$d(f,g) = \max_{x \in [a,b]} |f(x) - g(x)|.$$

Příklad 3. Vezmeme $M = \mathcal{C}[a,b]$ a reálné číslo $p \geq 1$. Podobně jako v prvním příkladu máme na M metriky

$$d_p(f,g) = \left(\int_a^b |f(x) - g(x)|^p \, dx \right)^{1/p}.$$

Hodnota $p = 1$ dává *integrální metriku* a $p \to \infty$ dává maximovou metriku z druhého příkladu. Důležitý je opět případ $p = 2$. Co je na exponentu $p = 2$ zvláštního? Ukazuje se, že metrika $d_2(\cdot,\cdot)$, v prvním i ve třetím příkladu, je odvozena ze skalárního součinu na vektorovém prostoru, a proto má řadu pěkných a důležitých vlastností. Vrátíme se k tomu v závěru první kapitoly.

Vezmeme-li širší třídu funkcí $M = \mathcal{R}[a,b]$ (funce mající na $[a,b]$ Riemannův integrál), je $d_p(f,g)$ definovaná, ale nesplňuje axiom b) a nedostáváme metriku. Změníme-li například hodnotu funkce $f \in \mathcal{R}[a,b]$ v jediném bodě, dostaneme odlišnou funkci $f_0 \in \mathcal{R}[a,b]$, ale $d_p(f,f_0) = 0$. (Tato potíž se odstraní tak, že místo s $\mathcal{R}[a,b]$ se pracuje s $\mathcal{R}[a,b]/\sim$ pro vhodnou relaci ekvivalence $\sim$.)

Příklad 4. Na souvislém grafu $G = (M,E)$ s množinou vrcholů M máme metriku

$$d(u,v) = \text{počet hran na nejkratší cestě v } G \text{ spojující } u \text{ a } v.$$

Příklad 5. Je-li A konečná množina (abeceda), máme na množině $M = A^m$ slov délky m nad abecedou A tak zvanou *Hammingovu metriku* ($u = a_1a_2\ldots a_m, v = b_1b_2\ldots b_m$)

$$d(u,v) = \text{počet souřadnic } i, \text{ pro něž } a_i \neq b_i.$$

Měří míru odlišnosti obou slov—jaký nejmenší počet změn v písmenech stačí k přeměně u ve v.

Příklad 6. Na dvourozměrné sféře

$$M = S_2 = \{(x_1,x_2,x_3) \in \mathbf{R}^3 \mid x_1^2 + x_2^2 + x_3^2 = 1\}$$

můžeme zavést metriku

$$d(x,y) = \text{délka nejkratší křivky ležící v } S_2 \text{ a spojující } x \text{ a } y.$$

Konkrétně je $d(x,y)$ rovna délce kratšího z obou oblouků, na něž body x a y dělí hlavní kružnici $K(x,y)$ jimi procházející. $K(x,y)$ je průnik S_2 s rovinou určenou počátkem souřadnic a body x a y. Leží-li tyto tři body na přímce (x a y jsou antipodální), není $K(x,y)$ určena jednoznačně, ale to nic nemění na tom, že pak $d(x, y) = \pi$. Tuto metriku nazveme *sférickou metrikou.* Můžeme ji uvažovat i na podmnožinách S_2, například na horní polosféře

$$S_2^+ = \{(x_1, x_2, x_3) \in \mathbf{R}^3 \mid x_1^2 + x_2^2 + x_3^2 = 1, x_3 \geq 0\}.$$

Dá se dokázat (úloha 9), že S_2^+ se sférickou metrikou není izometrická žádné množině $X \subset \mathbf{R}^n$ s euklidovskou metrikou. Totéž tedy platí i pro celou sféru. Sféra ani polosféra tak nejsou „ploché“, nedají se „splácnout“ do roviny ani žádného jiného euklidovského prostotu se zachováním vzdáleností.

Příklad 7. Položme $M = \mathbf{Z}$ (množina celých čísel) a vezměme nějaké prvočíslo p, například $p = 29$. Pro $z \in \mathbf{Z}$ jako $m_p(z)$ označíme největší celé číslo $e \geq 0$ takové, že p^e dělí z; $m_p(0) := \infty$. Na M definujeme tzv. *p-adickou metriku* $(2^{-\infty} = 0)$

$$d_p(x, y) = 2^{-m_p(x-y)}.$$

Dá se ukázat, že p-adická metrika splňuje toto zesílení trojúhelníkové nerovnosti:

$$d_p(x, y) \leq \max(d_p(x, z), d_p(z, y)).$$

Metrikám splňujícím tuto silnější verzi trojúhelníkové nerovnosti se říká *ultrametriky* (nebo *nearchimedovské metriky*). Pro zajímavé vlastnosti ultrametrik viz úlohu 12.

Koule, otevřené a uzavřené množiny. (M, d) buď metrický prostor. Pro bod $a \in M$ a reálné číslo $r > 0$ nazveme množinu

$$B(a, r) = \{x \in M \mid d(a, x) < r\}, \text{ resp. } \overline{B}(a, r) = \{x \in M \mid d(a, x) \leq r\},$$

(otevřenou) koulí se středem a a poloměrem r, resp. *uzavřenou koulí se středem a a poloměrem r*. Podmnožina $X \subset M$ je *otevřená*, pokud

$$\forall a \in X \; \exists r > 0 : B(a, r) \subset X,$$

jinak řečeno, X s každým bodem obsahuje i nějakou kouli kolem něj. $X \subset M$ je *uzavřená*, pokud je $M \backslash X$ otevřená množina. Každá koule je otevřená množina a každá uzavřená koule je uzavřená množina.

Tvrzení 1.1. *V metrickém prostoru (M, d) jsou množiny $\emptyset$ a M otevřené i uzavřené. Sjednocení libovolně mnoha otevřených množin je otevřená množina a průnik konečně mnoha otevřených množin je otevřená množina. Sjednocení konečně mnoha uzavřených množin je uzavřená množina a průnik libovolně mnoha uzavřených množin je uzavřená množina.*

Důkaz. Množiny $\emptyset$ a M jsou zjevně otevřené a protože jedna je doplňkem druhé, jsou i uzavřené. Nechť $\{G_i \mid i \in I\}$ je systém otevřených množin a $a \in G = \bigcup_{i \in I} G_i$. Pak a leží v nějaké G_j a tedy, pro nějaké $r > 0$, $B(a, r) \subset G_j \subset G$ a G je otevřená. Nechť je indexová množina I konečná a $a \in G = \bigcap_{i \in I} G_i$. To znamená, že $a \in G_i$ pro každé $i \in I$, a tak $B(a, r_i) \subset G_i$ pro nějaká čísla $r_i > 0$. Protože jich je jen konečně mnoho, můžeme vzít $r > 0$, že $r < \min(r_i : i \in I)$, a máme $B(a,r) \subset B(a,r_i) \subset G_i$ pro všechna $i \in I$. Tedy $B(a,r) \subset G$ a G je otevřená. Tvrzení o uzavřených množinách vyplývá z tvrzení o otevřených množinách pomocí de Morganových identit přechodem k doplňkům. $\square$

Otevřené množiny jsou „robustní“ množiny—leží-li bod a v otevřené množině X, pak pro dostatečně malé $\varepsilon > 0$ žádná porucha či posunutí menší než ε nedostane a ven z X. Uzavřené množiny si zase můžeme představovat jako množiny s ostrými obrysy—zanedlouho dokážeme, že X je uzavřená, právě když s každou konvergentní posloupností bodů obsahuje i její limitu.

Okolí bodu, vnitřní, vnější a další body. (M,d) buď metrický prostor a $a \in M$ jeho bod. Každou otevřenou množinu $U \subset M$ splňující $a \in U$ nazveme *okolím* bodu a. Například každá koule $B(a,r)$ je okolím a a ovšem každé okolí bodu a obsahuje nějakou takovou kouli.

Nechť $a \in M$ je bod a $X \subset M$ je množina. V následujících definicích symbolem U rozumíme okolí bodu a. Řekneme, že

- a je *vnitřním bodem* X, když existuje U tak, že $U \subset X$;
- a je *vnějším bodem* X, když existuje U tak, že $U \subset M \backslash X$;
- a je *hraničním bodem* X, když každé U protíná X i $M \backslash X$;
- a je *limitním bodem* X, když je pro každé U průnik $U \cap X$ nekonečný;
- a je *izolovaným bodem* X, když existuje U tak, že $U \cap X = \{a\}$.

Vnitřní a izolované body X nutně leží v X a vnější body leží mimo X. Hraniční a limitní body X mohou ležet v X i mimo X.

Jako příklad vezmeme v euklidovské rovině $\mathbf{R}^2$ množinu

$$X = \{x \in \mathbf{R}^2 \mid 0 < \|x\| < 1\} \cup \{(0,2)\},$$

jednotkový kruh se středem v počátku, z něhož jsme počátek odstranili a k němuž jsme přidali bod $(0,2)$. Pak vnitřní body X tvoří množinu $\{x \in \mathbf{R}^2 \mid 0 < \|x\| < 1\}$, vnější body množinu $\{x \in \mathbf{R}^2 \mid \|x\| > 1, x \neq (0,2)\}$, hraniční body množinu $\{x \in \mathbf{R}^2 \mid \|x\| = 1\} \cup \{(0,0),(0,2)\}$, limitní body množinu $\{x \in \mathbf{R}^2 \mid \|x\| \leq 1\}$ a izolované body množinu $\{(0,2)\}$.

Podprostory a součiny metrických prostorů. Popíšeme dvě jednoduché konstrukce vyrábějící nové metrické prostory ze starých. Metrický prostor (M_1,d_1) je *podprostorem* metrického prostoru (M_2,d_2), pokud $M_1 \subset M_2$ a pro každé dva body x,y z M_1 máme $d_1(x,y) = d_2(x,y)$. Metrický prostor (M,d) a podmnožina $X \subset M$ dávají nový metrický prostor (X,d'), kde d' je metrika d zúžená na $X \times X$; je jasné, že (X,d') je podprostorem (M,d). Říká se také, že (X,d') je podprostor s *indukovanou metrikou.*

(M,d) je *součinem* metrických prostorů (M_1,d_1) a (M_2,d_2), když $M = M_1 \times M_2$ a d je metrika ($x = (x_1,x_2)$ a $y = (y_1,y_2)$ jsou z M):

$$d(x,y) = \sqrt{d_1(x_1,y_1)^2 + d_2(x_2,y_2)^2}.$$

Všimněte si, že součin euklidovských prostorů $\mathbf{R}^m$ a $\mathbf{R}^n$ je euklidovský prostor $\mathbf{R}^{m+n}$ (přesně řečeno, $\mathbf{R}^m \times \mathbf{R}^n$ je izometrický $\mathbf{R}^{m+n}$ prostřednictvím zobrazení $((x_1, \ldots, x_m), (y_1, \ldots, y_n)) \mapsto (x_1, \ldots, x_m, y_1, \ldots, y_n))$.

Vzorec pro součinovou metriku jsme zafixovali poněkud libovolně, hlavním důvodem bylo, aby pro euklidovské prostory platilo $\mathbf{R}^m \times \mathbf{R}^n = \mathbf{R}^{m+n}$. Ale například vzorce

$$d(x,y) = d_1(x_1,y_1) + d_2(x_2,y_2) \text{ nebo } d(x,y) = \max\{d_1(x_1,y_1), d_2(x_2,y_2)\}$$

by posloužily dobře také. Není totiž těžké vidět, že všechny tři definují stejné otevřené (a tedy i uzavřené) množiny v $M_1 \times M_2$, a proto se výsledné součinové prostory velmi podobají. Pro úvahy a pojmy, které vystačí jen s otevřenými množinami (což zahrnuje většinu úvah v této přednášce, s vyjímkou izometrie a cauchyovskosti posloupnosti), jsou tyto tři součinové metriky nerozlišitelné.

Úlohy

1. Ukažte, že nezápornost metriky plyne z axiomů b) a c).
2. Co se stane, když v definici metriky zapomeneme na symetrii? Plyne z ostatních axiomů?
3. Dokažte vzorec pro maximovou metriku: $\lim_{p\to+\infty} d_p(x,y) = d_\infty(x,y)$.
4. Odvoďte z Cauchyovy–Schwarzovy nerovnosti $((a_1b_1 + \ldots + a_nb_n)^2 \le (a_1^2 + \ldots + a_n^2)(b_1^2 + \ldots + b_n^2))$ trojúhelníkovou nerovnost pro euklidovskou metriku.
5. Ověřte trojúhelníkovou nerovnost pro supremovou metriku.
6. Dokažte vzorec pro maximovou metriku pro funkce: je-li f spojitá na $[a,b]$, pak
$$\lim_{p\to+\infty} \left(\int_a^b |f(x)|^p \, dx \right)^{1/p} = \max_{x\in[a,b]} |f(x)|.$$
7. Ověřte v příkladu 3 pro $M = \mathcal{C}[a,b]$ axiom b) metrického prostoru.
8. Ověřte trojúhelníkové nerovnosti v diskrétních příkladech 4 a 5.
9. Dokažte, že horní polosféra S_2^+ se sférickou metrikou není izometrická žádné podmnožině euklidovského prostoru s euklidovskou metrikou. Návod: uvažte čtyři body $(0,0,1)$, $(1,0,0)$, $(0,1,0)$ a $(1/\sqrt{2}, 1/\sqrt{2}, 0)$.
10. Jak by se totéž dokázalo pro sférický vrchlík $S_2^v = \{(x_1,x_2,x_3) \in S_2 \mid x_3 \ge v\}$, kde $0 \le v < 1$?
11. Ověřte, že metrika v příkladu 7 je ultrametrika.

12. Dokažte tyto neintuitivní vlastnosti ultrametrického prostoru: každý trojúhelník je rovnoramenný a v každé kouli je libovolný bod jejím středem.

13. Ukažte, že v metrickém prostoru jsou koule $B(a, r)$ otevřené množiny a uzavřené koule $\overline{B}(a, r)$ a jednobodovky $\{a\}$ jsou uzavřené množiny.

14. Je konečná podmnožina metrického prostoru vždy uzavřená?

15. Co lze říci o otevřených množinách metrického prostoru (M, d), jehož každý bod je izolovaný (jako bod množiny M)?

16. Ukažte, že bod množiny X v metrickém prostoru je limitním bodem X, právě když není izolovaným bodem X. A ukažte, že bod mimo množinu X je limitním bodem X, právě když je hraničním bodem X.

2. přednáška 8. října 2007

Konvergence v metrických prostorech. Posloupnost bodů $(a_n) \subset M$ v metrickém prostoru (M, d) *konverguje* (je *konvergentní*), když v M existuje takový bod a, že

$$\lim_{n\to\infty} d(a_n, a) = 0.$$

Pak říkáme, že bod a je *limitou* posloupnosti (a_n) a píšeme $\lim_{n\to\infty} a_n = a$ či $a_n \to a$ pro $n \to \infty$. Ekvivalentní formulace konvergence (a_n) k a jsou

$$\forall \varepsilon > 0 \; \exists n_0 : \; n > n_0 \Rightarrow d(a_n, a) < \varepsilon$$

a

$$\forall \text{okolí } U \text{ bodu } a \; \exists n_0 : \; n > n_0 \Rightarrow a_n \in U.$$

Posloupnost bodů $(a_n) \subset M$ je *cauchyovská*, když

$$\forall \varepsilon > 0 \; \exists n_0 : \; m > n > n_0 \Rightarrow d(a_m, a_n) < \varepsilon.$$

Platí triviální implikace, že konvergentní posloupnost je cauchyovská. Opačná implikace platí v *úplných* metrických prostorech, k nimž se později dostaneme.

Tvrzení 1.2. *Podmnožina $X \subset M$ je uzavřená, právě když limita každé konvergentní posloupnosti (a_n) obsažené v X také leží v X.*

Důkaz. Nechť X není uzavřená. Pak $M \backslash X$ není otevřená, a tak existuje takový bod $a \in M \backslash X$, že pro každé n koule $B(a, 1/n)$ protíná X. Z průniku $B(a, 1/n) \cap X$ vybereme bod a_n. Pak $a_n \to a$ pro $n \to \infty$, $(a_n) \subset X$, ale $a \notin X$.

Nechť je X uzavřená a posloupnost $(a_n) \subset X$ je konvergentní, $a_n \to a$ pro $n \to \infty$. Kdyby limita a neležela v X, ležela by v otevřené množině $M \backslash X$ a existovalo by $r > 0$, že $B(a, r) \subset M \backslash X$. Pro nějaké n_0 bychom pak měli, že $n > n_0 \Rightarrow a_n \in B(a, r) \subset M \backslash X$. To ale je ve sporu s tím, že $a_n \in X$ pro každé n. Takže $a \in X$. □

Připomínáme, že konvergence posloupnosti je relativní pojem: posloupnost $(a_n) \subset X \subset M$, která je konvergentní v celém metrickém prostoru (M, d), nemusí být konvergentní v podprostoru (X, d) s indukovanou metrikou, protože $\lim a_n$ nemusí ležet v X. Takže třeba $(1/n)$ je konvergentní v $\mathbf{R}$, ale ne v podprostoru $(0, 1]$. Pro uzavřené množiny X tato obtíž nenastává.

Uzávěr množiny $X \subset M$ je množina

$$\overline{X} = \{a \in M \mid a = \lim_{n\to\infty} a_n \text{ pro nějakou } (a_n) \subset X\}.$$

Protože každý bod $a \in X$ je limitou konstantní posloupnosti $(a, a, a, \ldots)$, máme $X \subset \overline{X}$. K X tedy přidáme všechny limitní body X (stačí přidat jen ty ležící mimo X):

$$\overline{X} = X \cup \{\text{limitní body } X\}.$$

Další ekvivalentní definice uzávěru množiny je tato:

$$\overline{X} = \bigcap_{V \supset X,\ V \text{ je uzavřená}} V.$$

$\overline{X}$ je tedy ve smyslu inkluze nejmenší uzavřená množina obsahující X. Je jasné, že X je uzavřená, právě když $X = \overline{X}$.

Jako příklady uvažme podmnožiny $\mathbf{Q}$ a $(0,1)$ v euklidovském prostoru $\mathbf{R}$. Pak $\overline{\mathbf{Q}} = \mathbf{R}$ a $\overline{(0,1)} = [0,1]$. Množina

$$X = \{(x, \sin(1/x)) \in \mathbf{R}^2 \mid x \in (0,1]\},$$

graf funkce $\sin(1/x)$ na intervalu $(0,1]$, má v euklidovské rovině $\mathbf{R}^2$ uzávěr

$$\overline{X} = X \cup (\{0\} \times [-1,1]).$$

Spojitá zobrazení mezi metrickými prostory. (M_1, d_1) a (M_2, d_2) buďte dva metrické prostory a $f: M_1 \to M_2$ buď zobrazení mezi nimi. Řekneme, že f je spojité v bodu $a \in M_1$, když

$$\forall \varepsilon > 0\ \exists \delta > 0: \ d_1(x,a) < \delta \Rightarrow d_2(f(x), f(a)) < \varepsilon.$$

Ekvivalentní definice spojitosti f v a pomocí okolí bodů je, že

$$\forall \text{okolí } V \text{ bodu } f(a)\ \exists \text{okolí } U \text{ bodu } a: \ x \in U \Rightarrow f(x) \in V.$$

Další ekvivaletní definice spojitosti f v a je *Heineho definice*: f je spojité v a, právě když pro každou posloupnost $(a_n) \subset M_1$ platí

$$(a_n \to a, n \to \infty) \Rightarrow (f(a_n) \to f(a), n \to \infty).$$

Zobrazení f je *spojité*, když je spojité v každém bodu prostoru M_1. Spojitost zobrazení lze popsat jen s použitím otevřených množin.

Tvrzení 1.3. *Zobrazení $f: M_1 \to M_2$ mezi metrickými prostory je spojité, právě když vzor každé otevřené množiny v M_2 je otevřená množina v M_1:*

$$V \subset M_2 \text{ je otevřená} \Rightarrow f^{-1}(V) = \{x \in M_1 \mid f(x) \in V\} \text{ je otevřená v } M_1.$$

Analogická ekvivalence platí i pro vzory uzavřených množin.

Důkaz. Nechť je $f: M_1 \to M_2$ spojité zobrazení mezi metrickými prostory (M_1, d_1) a (M_2, d_2), $V \subset M_2$ je otevřená množina a $a \in f^{-1}(V)$ je libovolný bod. Máme $f(a) \in V$, takže existuje takové $r > 0$, že $B_2(f(a), r) \subset V$ (index 2 zde odkazuje k metrice d_2). Díky spojitosti existuje takové $s > 0$, že $f(B_1(a,s)) \subset B_2(f(a), r)$. Tedy $B_1(a,s) \subset f^{-1}(V)$. Množina $f^{-1}(V)$ obsahuje libovolný bod s nějakou koulí kolem něj a je tedy otevřená.

Nechť f splňuje podmínku pro vzory otevřených množin a $a \in M_1$ je libovolny bod. Pro dané $\varepsilon > 0$ uvážíme kouli $B_2(f(a), \varepsilon)$. Je to otevřená množina, takže její vzor $f^{-1}(B_2(f(a), \varepsilon))$ je otevřená množina v M_1. Protože bod a v ní leží, existuje takové $\delta > 0$, že $B_1(a, \delta) \subset f^{-1}(B_2(f(a), \varepsilon))$. Takže $f(B_1(a, \delta)) \subset B_2(f(a), \varepsilon)$ a f je spojité v a.

Ekvivalence spojitosti s podmínkou pro vzory uzavřených množin plyne z právě dokázaného přechodem k doplňkům a s pomocí identity $f^{-1}(M_2 \backslash X) = M_1 \backslash f^{-1}(X)$. □

Skládání zobrazení zachovává spojitost. Jsou-li zobrazení $f: M_1 \to M_2$ a $g: M_2 \to M_3$ mezi metrickými prostory spojitá, je i složené zobrazení $h = g(f)$ spojité. Je-li f spojité v bodu $a \in M_1$ a g je spojité v bodu $f(a) \in M_2$, je h spojité v bodu $a \in M_1$.

Homeomorfismus. Bijekce $f: M_1 \to M_2$ mezi dvěma metrickými prostory je *homeomorfismus*, když zobrazení f i inverzní zobrazení f^{-1} je spojité. Existuje-li taková bijekce mezi M_1 a M_2, jsou oba metrické prostory *homeomorfní*. Homeomorfismus je druh izomorfismu metrických prostorů, který je slabší než izometrie. Každá izometrie je homeomorfismem, ale obecně ne naopak. Homeomorfismus je izomorfismus struktur otevřených množin obou prostorů. Homeomorfní metrické prostory se nedají odlišit jen pomocí otevřených množin.

Jako příklad uvažme zobrazení $x \mapsto \tan x$ mezi euklidovskými prostory $(-\pi/2, \pi/2)$ a $\mathbf{R}$. Toto zobrazení je homeomorfismus (je bijektivní a $\tan x$ i inverz $\arctan x$ je spojité zobrazení). Tyto prostoty zjevně nejsou izometrické, protože první je omezený, ale druhý ne. (Množina $X \subset M$ v metrickém prostoru (M, d) je *omezená*, když existuje bod $a \in M$ a poloměr $r > 0$ tak, že $X \subset B(a, r)$.)

Na druhou stranu zobrazení $f(\varphi) = (\cos\varphi, \sin\varphi)$ mezi euklidovskými prostory

$$[0, 2\pi) \quad \text{a} \quad K = \{x \in \mathbf{R}^2 \mid \|x\| = 1\},$$

což je interval v $\mathbf{R}$ a jednotková kružnice v $\mathbf{R}^2$, homeomorfismem není. Je sice bijektivní a spojité, ale inverzní zobrazení spojité není (není spojité v bodu $(1, 0)$). Jak uvidíme, metrické prostory $[0, 2\pi)$ a K ani homeomorfní nejsou, protože první z nich není kompaktní, ale druhý je.

Kompaktní metrické prostory. Metrický prostor (M, d) je *kompaktní*, když má každá posloupnost bodů $(a_n) \subset M$ konvergentní podposloupnost. Podmnožina $X \subset M$ je kompaktní, když je podprostor (X, d) s indukovanou metrikou kompaktní. Z MAI dobře víme, že intervaly $[a, b]$ jsou kompaktní podmnožiny v euklidovském prostoru $\mathbf{R}$. Důležitost kompaktních množin spočívá v tom, že spojité reálné funkce na nich nabývají maxima a minima a také v tom, že spojitá zobrazení definovaná na kompaktních prostorech jsou stejnoměrně spojitá.

Tvrzení 1.4. *Kompaktnost se zachovává následujícími operacemi.*

1. *Přechodem k uzavřenému podprostoru.*

2. Obrazem spojitým zobrazením.

3. Kartézským součinem.

Důkaz. 1. Nechť je (M, d) kompaktní, podmnožina $X \subset M$ je uzavřená a $(a_n) \subset X$ je libovolná posloupnost. Díky kompaktnosti celého prostoru má konvergentní podposloupnost (a_{k_n}) s limitou $a \in M$. Díky uzavřenosti X ale a leží v X, takže (a_{k_n}) je konvergentní i v podprostoru (X, d). Proto je X též kompaktní.

2. Nechť $f : M_1 \to M_2$ je spojité zobrazení mezi metrickými prostory (M_1, d_1) a (M_2, d_2), přičemž (M_1, d_1) je kompaktní. Tvrdíme, že $f(M_1)$ je kompaktní podmnožina M_2. Nechť $(b_n) \subset f(M_1)$ je libovolná posloupnost. Pro každé n zvolíme $a_n \in M_1$ tak, že $f(a_n) = b_n$. Z (a_n) vybereme konvergentní podposloupnost (a_{k_n}) s limitou $a \in M_1$. Protože $a_{k_n} \to a$ pro $n \to \infty$ a f je spojité v a, podle Heineho definice spojitosti máme i $b_{k_n} = f(a_{k_n}) \to f(a) = b$ pro $n \to \infty$. Takže (b_n) má konvergentní podposloupnost a $f(M_1)$ je kompaktní.

3. Přenecháváme pilnému čtenáři jako cvičení (úloha 9). $\square$

Z části 2 plyne, že homeomorfní metrické prostory buď současně jsou nebo současně nejsou kompaktní. V druhém příkladu na homeomorfismus $[0, 2\pi)$ není kompaktní ($(2\pi - 1/n)$ nemá konvergentní podposloupnost), ale jednotková kružnice K kompaktní je (jako spojitý obraz kompaktní množiny, $K = f([0, 2\pi])$). Tudíž $[0, 2\pi)$ a K nejsou homeomorfní. A co $[0, 2\pi]$ a K (oba prostory jsou teď kompaktní)? Viz úloha 7.

V příští přednášce dokážeme další výsledky o kompaktních prostorech:

Tvrzení 1.5. *Kompaktní podmnožiny v metrickém prostoru jsou uzavřené a omezené.*

Věta 1.6. *Kompaktní podmnožiny euklidovského prostoru $\mathbf{R}^n$ jsou právě a jen uzavřené a omezené množiny.*

Ukážeme si také příklady omezených a uzavřených množin, které nejsou kompaktní.

Úlohy

1. Dokažte, že konvergentní posloupnost je cauchyovská a že množina členů cauchyovské posloupnosti je omezená.

2. Nechť X je konečná množina v metrickém prostoru, jejíž každý bod je izolovaný. Co lze říci o uzávěru X?

3. Co lze říci o uzávěru množiny, která nemá hraniční body?

4. Nechť $N = (0,1) \cup (2,3]$ a (N,d) je metrika indukovaná z euklidovského prostoru $\mathbf{R}$. Jaké jsou uzávěry množin $X = (0,1)$ a $X = (2,3)$ v prostoru N?

5. Nechť (N,d) je jako v předchozí úloze a $f : N \to \mathbf{R}$ je na $(0,1)$ rovna konstantě a a na $(2,3]$ je rovna konstantě b. Pro jaké hodnoty a a b je f spojitá funkce?

6. Nechť $N = \{1, 1/2, 1/3, 1/4, \ldots\} \cup \{0\}$ a (N,d) je opět metrika indukovaná z euklidovského prostoru $\mathbf{R}$. Popište spojité funkce $f : N \to \mathbf{R}$.

7. Ukažte, že euklidovské prostory $I = [0, 2\pi]$ a jednotková kružnice $K = \{x \in \mathbf{R}^2 \mid \|x\| = 1\}$ nejsou homeomorfní. Návod: Dá se $I \backslash \{1\}$ vyjádřit jako sjednocení dvou neprázdných a disjunktních otevřených množin? Dá se tak vyjádřit K po vyhození jednoho bodu?

8. Ukažte, že euklidovské prostory $\mathbf{R}$ a $\mathbf{R}^2$ nejsou homeomorfní. Návod: stejný argument, jako v předchozí úloze.

9. Dokažte část 3 Tvrzení 1.4: když jsou (M_1, d_1) a (M_1, d_1) kompaktní, je součinový prostor $(M_1 \times M_2, d)$ (pro definici viz konec 1. přednášky) rovněž kompaktní.

3. přednáška 15. října 2007

Kompaktnost a uzavřené a omezené množiny. Kompaktní množiny jsou vždy uzavřené a omezené, a v euklidovských prostorech to platí i naopak. Obecně to ale naopak neplatí.

Tvrzení 1.5. *Kompaktní podmnožiny v metrickém prostoru jsou uzavřené a omezené.*

Důkaz. Nechť $X \subset M$ je podmnožina v metrickém prostoru (M, d), která není uzavřená. Existuje tedy konvergentní posloupnost $(a_n) \subset X$, jejíž limita a leží mimo X (Tvrzení 1.2). Každá její podposloupnost má ale a jako limitu a není proto konvergentní v (X, d). Podprostor (X, d) není kompaktní, neboť posloupnost (a_n) nemá konvergentní podposloupnost.

Nechť X není omezená. Inkluze $X \subset B(a, r)$ tedy neplatí pro žádnou kouli $B(a, r)$ a díky tomu lehce sestrojíme posloupnost $(a_n) \subset X$ splňující $d(a_m, a_n) \geq 1$ pro každé dva indexy $1 \leq m < n$. (Když už máme v X body $a_1, a_2, \ldots, a_k$, z nichž každé dva mají vzdálenost ≥ 1, zvolíme $a_{k+1} \in X$ mimo kouli $B(a_1, r)$, kde $r = 2 + \max_{1 \leq i \leq k} d(a_1, a_i)$. Pak $d(a_{k+1}, a_i) \geq 1$ pro $1 \leq i \leq k$.) Tuto vlastnost má zřejmě i každá podposloupnost a žádná proto není konvergentní. Podprostor (X, d) tedy není kompaktní. □

Jak jsme už poznamenali, konvergentnost je relativní vlastnost, která závisí na obklopujícím podprostoru—přechodem k podprostoru může posloupnost přestat být konvergentní. Je-li ovšem posloupnost konvergentní v podprostoru, je nutně konvergentní i v celém prostoru. S otevřeností a uzavřeností množin se to má opačně. Pokud $X \subset Y \subset M$ a množina X je otevřená v celém prostoru (M, d), pak je X otevřená i v podprostoru (Y, d), a totéž platí pro uzavřenost (úloha 1). Přechodem k nadprostoru se ale otevřenost a uzavřenost může ztratit—například Y je vždy otevřená i uzavřená v (Y, d), ale už to tak nemusí být v (M, d).

Kompaktnost je absolutní vlastnost, úplně nezávislá na obklopujícím podprostoru. Z definice je jasné, že pokud $X \subset Y \subset M$, pak X je kompaktní v podprostoru (Y, d), právě když je kompaktní v celém prostoru (M, d).

Věta 1.6. *Kompaktní podmnožiny euklidovského prostoru* $\mathbf{R}^n$ *jsou právě a jen uzavřené a omezené množiny.*

Důkaz. Kompaktní množina $X \subset \mathbf{R}^n$ je uzavřená a omezená podle Tvrzení 1.5. Naopak, nechť je podmnožina $X \subset \mathbf{R}^n$ uzavřená a omezená. Pak, díky omezenosti, existuje $c > 0$ tak, že $X \subset [-c, c]^n$. Krychle $[-c, c]^n$ je kompaktní množina (protože interval $[-c, c]$ je kompaktní v $\mathbf{R}$ a kompaktnost se zachovává kartézskými součiny, část 3 Tvrzení 1.4). Protože X je uzavřená v $\mathbf{R}^n$, je uzavřená i v podprostoru $[-c, c]^n$. Takže X je kompaktní v podprostoru $[-c, c]^n$ (část 1 Tvrzení 1.4) a je tedy kompaktní i v celém prostoru $\mathbf{R}^n$ (díky absolutnosti kompaktnosti). □

Obecně omezenost a uzavřenost množiny ještě nezaručují kompaktnost, což potvrdíme dvěma příklady. M buď libovolná nekonečná množina a (M, d) *diskrétní metrický prostor*, tedy $d(x, x) = 0$ a $d(x, y) = 1$ pro $x \neq y$. Celý prostor M je uzavřený a omezený ($M \subset B(a, 2)$ pro každý bod $a \in M$). Každá posloupnost $(a_n) \subset M$, jejíž členy jsou vzájemně různé, splňuje $d(a_m, a_n) = 1$ pro každé dva indexy $1 \leq m < n$ a není proto konvergentní. Totéž platí i pro každou její podposloupnost.

Jako druhý příklad vezmeme spojité funkce $M = \mathcal{C}[0, 1]$ s maximovou metrikou a podmnožinu $X = \{f \in M \mid \max_{[0,1]} |f| \leq 1\}$. Lehce se sestrojí taková posloupnost funkcí $(f_n) \subset X$, že pro každé dva indexy $1 \leq m < n$ platí

$$d(f_m, f_n) = \max_{0 \leq x \leq 1} |f_m(x) - f_n(x)| = 1.$$

Tato posloupnost pak zřejmě nemá konvergentní podposloupnost, i když je množina X omezená a uzavřená (podrobnosti viz úloha 3).

Spojitá funkce nabývá na kompaktu extrémy. Užitečnost kompaktních množin popisuje následující věta.

Věta 1.7. *Nechť $f : M_1 \to M_2$ je spojité zobrazení mezi metrickými prostory (M_1, d_1) a (M_2, d_2), přičemž M_1 je kompaktní.*

1. *Je-li $M_2 = \mathbf{R}$ jednorozměrný euklidovský prostor, nabývá f na M_1 maxima i minima.*
2. *Je-li f navíc bijekce, pak je inverzní zobrazení f^{-1} nutně spojité.*
3. *Zobrazení f je dokonce stejnoměrně spojité.*

Důkaz. 1. Podle části 2 Tvrzení 1.4 je $f(M_1)$ kompaktní podmnožina v $\mathbf{R}$. Což podle Tvrzení 1.5 znamená, že je uzavřená a omezená. Takže $f(M_1)$ má konečné supremum, které je rovno maximu (supremum $f(M_1)$ je totiž limitou jisté posloupnosti bodů z $f(M_1)$) a množina $f(M_1)$ proto má maximum. Podobně pro minimum.

2. Podle Tvrzení 1.3 stačí ověřit, že pro každou uzavřenou množinu $X \subset M_1$ je její vzor $(f^{-1})^{-1}(X) \subset M_2$ uzavřená množina. Protože f je bijekce, máme $(f^{-1})^{-1} = f$. Nechť $X \subset M_1$ je uzavřená. Podle části 1 Tvrzení 1.4 je X kompaktní. Podle části 2 je $f(M_1)$ kompaktní množina v M_2. Podle Tvrzení 1.5 je $f(M_1)$ uzavřená.

3. Přenecháváme čtenářce jako cvičení (úloha 4). □

Jako aplikaci části 1 této věty nyní dokážeme, že každý nekonstantní komplexní polynom má alespoň jeden kořen.

Základní věta algebry. *Nechť $p(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0$ je polynom s komplexními koeficienty stupně $n \geq 1$ (takže $a_i \in \mathbf{C}$ a $a_n \neq 0$). Pak existuje takové číslo $\alpha \in \mathbf{C}$, že $p(\alpha) = 0$.*

Důkaz. Komplexní rovina $\mathbf{C}$ se standardní metrikou

$$d(a+bi, c+di) = |a+bi-(c+di)| = \sqrt{(a-c)^2+(b-d)^2}$$

je izometrická euklidovské rovině $\mathbf{R}^2$. Bereme ji tedy jako euklidovský prostor $\mathbf{R}^2$. Existence kořene α polynomu $p(z)$ plyne okamžitě z následujících dvou kroků.

Krok 1. *Pro každý komplexní polynom $p(z)$ nabývá funkce $f(z) = |p(z)|$, $f:\ \mathbf{C} \to \mathbf{R}_{\geq 0}$, na $\mathbf{C}$ svého minima.*

Krok 2. *Nechť nekonstantní komplexní polynom $p(z)$ splňuje v bodu $\alpha \in \mathbf{C}$ nerovnost $|p(\alpha)| > 0$ (tj. $p(\alpha) \neq 0$). Pak existují $\delta > 0$ a polopřímka $\ell \subset \mathbf{C}$ vycházející z α tak, že*

$$z \in \ell \ \&\ 0 < |z-\alpha| < \delta \Rightarrow |p(z)| < |p(\alpha)|.$$

(Analogický výsledek platí i pro opačnou nerovnost $|p(z)| > |p(\alpha)|$.)

Skutečně, pro nekonstantní komplexní polynom $p(z)$ vezmeme α, v němž se podle kroku 1 nabývá nejmenší hodnota modulu $|p(z)|$. Podle kroku 2 musí platit $|p(\alpha)| = 0$, takže $p(\alpha) = 0$ a α je kořenem $p(z)$.

Základní myšlenky důkazů obou kroků budeme ilustrovat na konkrétním polynomu

$$p(z) = z^5 + (3i)z^3 + (-1+i).$$

Důkaz kroku 1. Funkce $f(z) = |p(z)|$ je spojitá (úloha 5), nemůžeme však hned použít Větu 1.7, protože $\mathbf{C}$ není kompaktní. Máme

$$|p(0)| = |-1+i| = \sqrt{2} < 2.$$

Na druhou stranu, vytknutím nejvyšší mocniny

$$p(z) = z^5(1 + 3i/z^2 + (-1+i)/z^5)$$

dostáváme odhad

$$|z| \geq 2 \Rightarrow |p(z)| \geq 2^5(1 - |3i|/2^2 - |-1+i|/2^5) = 2^5 - 3 \times 2^3 - \sqrt{2} > 6.$$

(Použili jsme nerovnost $|a+b| \geq |a| - |b|$.) Mimo kruh

$$K = \{z \in \mathbf{C} \mid |z| \leq 2\}$$

má tedy $|p(z)|$ všechny hodnoty větší než v nule, což je bod K, a pokud nabývá na $\mathbf{C}$ minima, musí to být někde na K. Nabývání minima na K je už ale zaručeno Větou 1.7 (část 1), protože K je kompaktní (podle Věty 1.6, K je uzavřená a omezená množina). Takže $|p(z)|$ nabývá na $\mathbf{C}$ minimum. Podobné odhady fungují pro obecný polynom.

Důkaz kroku 2. Můžeme předpokládat, že $\alpha = 0$. Pro obecné α uděláme substituci $z = (z - \alpha) + \alpha = t + \alpha$, kterou přejdeme k polynomu

$$q(t) = p(t + \alpha)$$

v okolí bodu $t = 0$. Podívejme se tedy na náš konkrétní polynom $p(z) = z^5 + (3i)z^3 + (-1+i)$ v okolí bodu $z = \alpha = 0$. Jak už víme, $|p(0)| = |-1+i| = \sqrt{2} > 0$. Nyní nám pomůže vytknutí nejnižší nekonstantní mocniny,

$$p(z) = -1 + i + (3i)z^3 + (3i)z^3 \cdot z^2/3i.$$

Idea je zvolit z blízko u nuly a s vhodným argumentem, aby se k $-1 + i$ přičetlo číslo, které je k nule blíže a leží na opačné straně. Výsledek pak je číslo s modulem menším než $|-1+i|$.

Zvolíme tedy $\delta > 0$ dostatečně malé, že

$$|z| < \delta \Rightarrow |(3i)z^3| < |-1+i| \;\text{ a }\; |z^2/3i| < 1.$$

Pokud navíc

$$\arg(z) = \frac{\arg(-1+i) + \pi - \arg(3i)}{3} = \frac{3\pi/4 + \pi - \pi/2}{3} = 5\pi/12,$$

máme $\arg((3i)z^3) - \arg(-1+i) = \pi$ (takže $(3i)z^3$ a $-1+i$ leží na opačné straně od nuly) a

$$|-1+i+(3i)z^3| = |-1+i| - |(3i)z^3|.$$

Celkem pro $z \in \mathbf{C}$ splňující $0 < |z| < \delta$ a $\arg(z) = 5\pi/12$ máme

$$\begin{aligned} |-1+i+(3i)z^3+(3i)z^3 \cdot z^2/3i| &\leq |-1+i+(3i)z^3| + |(3i)z^3| \cdot |z^2/3i| \\ &= |-1+i| - |(3i)z^3| + |(3i)z^3| \cdot |z^2/3i| \\ &< |-1+i|. \end{aligned}$$

Pro $0 < |z| < \delta$ a $\arg(z) = 5\pi/12$ tak platí, že $|p(z)| < |p(0)|$.

Pro obecný polynom se podobné odhady použijí pro $p(z)$ ve tvaru $p(z) = \kappa + \lambda z^k + q(z)$, kde $\kappa, \lambda \in \mathbf{C}$ jsou nenulové konstanty, $k \geq 1$ (zde využíváme, že $p(z)$ je nekonstantní) a v $q(z)$ jsou mocniny z s exponenty vyššími než k. Pro malé $|z|$ je $q(z)$ zanedbatelná porucha a $p(z)$ se chová víceméně jako $\kappa + \lambda z^k$. Vhodným nastavením $\arg(z)$ pak dosáhneme, že $|p(z)| \doteq |\kappa+\lambda z^k| = |\kappa| - |\lambda z^k| < |\kappa| = |p(0)|$. $\square$

Všimněte si, že výsledek v kroku 2 říká hodně o tom, kde může funkce $z \mapsto |p(z)|$ nabývat na kompaktní množině $X \subset \mathbf{C}$ lokální extrém. Ve vnitřním bodě X to je možné, jen když jde o lokální minimum s nulovou hodnotou. Ostatní lokální extrémy se musejí nabývat v hraničních bodech množiny X (které v ní leží, X je uzavřená). V komplexní analýze se tento výsledek, tzv. princip maxima modulu, dokazuje pro daleko širší třídu funkcí, než jsou polynomy.

Topologická kompaktnost. Podobně jako spojitost zobrazení se kompaktnost dá také ekvivalentně vyjádřit topologicky, jen pomocí otevřených množin. Pro množinu X v metrickém prostoru (M, d) nazveme systém množin $\{O_i \mid i \in I\}$ v M jejím *otevřeným pokrytím*, když jsou všechny množiny O_i otevřené a

$$X \subset \bigcup_{i \in I} O_i.$$

Konečné podpokrytí pak je konečný podsystém $\{O_j \mid j \in J\}$, $J \subset I$ je konečná, který stále pokrývá X. Množina X je *topologicky kompaktní*, když každé její otevřené pokrytí má konečné podpokrytí.

Věta 1.8. *Množina v metrickém prostoru je kompaktní, právě když je topologicky kompaktní.*

Důkaz. *Tento důkaz nebyl na přednášce a nebude se zkoušet.* Stačí se omezit na případ celého prostoru $X = M$ (úloha 6).

Kompaktnost ⇒ topologická kompaktnost. Předpokládáme, že prostor (M, d) je kompaktní a dokážeme, že je i topologicky kompaktní. Nejprve ukážeme, že pro každé $r > 0$ existuje konečná množina S, že

$$M = \bigcup_{a \in S} B(a, r).$$

Množině S se říká r-*síť*. Každý bod má od nějakého prvku r-sítě vzdálenost menší než r. Řekněme, že pro nějaké $s > 0$ žádná konečná množina v M není s-síť. Vezmeme $a_1 \in M$ libovolně. Protože $\{a_1\}$ není s-síť, existuje $a_2 \in M$ tak, že $d(a_1, a_2) \geq s$. Protože $\{a_1, a_2\}$ není s-síť, existuje $a_3 \in M$, že $d(a_1, a_3) \geq s$ a $d(a_2, a_3) \geq s$. Takto postupujeme dále a sestrojíme posloupnost $(a_n) \subset M$ s vlastností, že $d(a_m, a_n) \geq s$ pro každé dva indexy $1 \leq m < n$. Tato posloupnost nemá konvergentní podposloupnost, což je spor s předpokladem kompaktnosti.

Předpokládejme pro spor, že systém množin $\{O_i \mid i \in I\}$ je otevřené pokrytí M, které nemá konečné podpokrytí. Jako S_n si označíme konečnou $1/n$-síť. Kdyby pro každý bod $a \in S_n$ koule $B(a, 1/n)$ celá ležela v nějaké množině $O_{i(a)}$, podsystém $\{O_{i(a)} \mid a \in S_n\}$ by byl konečným podpokrytím: $B(a, 1/n) \subset O_{i(a)}$ pro každé $a \in S_n$, takže

$$M = \bigcup_{a \in S_n} B(a, 1/n) \subset \bigcup_{a \in S_n} O_{i(a)}.$$

Pro každé $n = 1, 2, \ldots$ tedy můžeme vybrat bod a_n z S_n, že koule $B(a_n, 1/n)$ není obsažena v žádné množině O_i. Posloupnost (a_n) má konvergentní podposloupnost (a_{k_n}) s limitou a. Protože systém množin pokrývá M, existuje $j \in I$, že $a \in O_j$. Množina O_j je otevřená, a tak $B(a, r) \subset O_j$ pro nějaké $r > 0$. Vezmeme tak velké $N \in \mathbf{N}$, že $d(a_{k_N}, a) < r/2$ a $1/k_N < r/2$. Pak, díky trojúhelníkové nerovnosti,

$$B(a_{k_N}, 1/k_N) \subset B(a, r) \subset O_j,$$

což je spor s definicí bodů a_n.

Topologická kompaktnost ⇒ kompaktnost. Předpokládáme, že prostor (M, d) je topologicky kompaktní a dokážeme, že je i kompaktní. Nechť $(a_n) \subset M$ je libovolná posloupnost. Ukážeme, že existuje takový bod a, že pro každé $r > 0$ je množina indexů $\{n \in \mathbf{N} \mid a_n \in B(a, r)\}$ nekonečná. Je lehké vidět, že takový bod už je limitou nějaké podposloupnosti vybrané z (a_n).

Kdyby to tak nebylo, tak pro každý bod $a \in M$ existuje poloměr $r(a) > 0$, že množina

$$I(a) = \{n \in \mathbf{N} \mid a_n \in B(a, r(a))\}$$

je konečná. Systém $\{B(a, r(a)) \mid a \in M\}$ je otevřené pokrytí M. Podle předpokladu má konečné podpokrytí určené konečnou množinou $X \subset M$. Uvažme množinu indexů

$$I = \bigcup_{a \in X} I(a).$$

Protože je konečná (je konečným sjednocením konečných množin), mohu vybrat index $N \in \mathbf{N} \backslash I$. Pak (podsystém $\{B(a, r(a)) \mid a \in X\}$ je pokrytí M)

$$a_N \in M = \bigcup_{a \in X} B(a, r(a)),$$

takže $a_N \in B(b, r(b))$ pro nějaké $b \in X$ a $N \in I(b) \subset I$, což je spor s výběrem indexu N. $\square$

Úlohy

1. Dokažte, že otevřenost množiny se zachová přechodem k podprostoru, a totéž platí pro uzavřenost.

2. Rozhodněte, zda je konečný diskrétní metrický prostor kompaktní.

3. Doplňte detaily v druhém příkladu za Větou 1.6: ukažte, že X je omezená a uzavřená a definujte v X posloupnost funkcí, v níž každé dvě funkce mají v maximové metrice vzdálenost 1.

4. Dokažte, že spojité zobrazení z kompaktního metrického prostoru (M_1, d_1) do jiného metrického prostoru (M_2, d_2) je nutně stejnoměrně spojité, to jest

$$\forall \varepsilon > 0\ \exists \delta > 0 :\ x, y \in M_1, d_1(x, y) < \delta \Rightarrow d_2(f(x), f(y)) < \varepsilon.$$

5. Dokažte, že pro komplexní polynom $p(z)$ je $z \mapsto |p(z)|$ spojité zobrazení z $\mathbf{C}$ do $\mathbf{R}$.

6. Ukažte, že pro $X \subset Y \subset M$ je X topologicky kompaktní v podprostoru (Y, d), právě když je topologicky kompaktní v celém prostoru (M, d).

4. přednáška 22. října 2007

Úplné metrické prostory. Metrický prostor (M, d) je *úplný*, když každá cauchyovská posloupnost bodů v M konverguje.

Příklady. 1. Euklidovský prostor $\mathbf{R}$ je úplný, každá cauchyovská posloupnost reálných čísel má limitu. Úplné jsou i podprostory $[2, 3]$ a $[-5, +\infty)$. Naopak podprostory $\mathbf{Q}$ a $(0, 1]$ úplné nejsou. Obecněji i euklidovské prostory $\mathbf{R}^n$ jsou úplné.

2. Z Matematické analýzy II víme, že prostor $\mathcal{C}[a, b]$ funkcí spojitých na $[a, b]$ s maximovou metrikou je úplný. Je-li totiž posloupnost (f_n) cauchyovská, splňuje stejnoměrnou Bolzanovu-Cauchyovu podmínku a tedy na $[a, b]$ konverguje stejnoměrně k jisté funkci f. Funkce f je na M spojitá, protože je stejnoměrnou limitou spojitých funkcí. Tedy $f \in \mathcal{C}(M)$ a v supremové metrice máme

$$\lim_{n\to\infty} f_n = f.$$

3. Vezmeme znovu spojité funkce $\mathcal{C}[a, b]$, ale teď $\mathcal{C}[a, b]$ vybavíme integrální metrikou. Vzniklý metrický prostor není úplný. Sestrojíme cauchyovskou posloupnost, která nemá limitu. Položíme $a = -1, b = 1$ a uvážíme funkce

$$f_n(x) = \begin{cases} -1 & \text{pro } -1 \le x \le -n^{-1} \\ nx & \text{pro } -n^{-1} \le x \le n^{-1} \\ 1 & \text{pro } n^{-1} \le x \le 1. \end{cases}$$

Pak $(f_n) \subset \mathcal{C}[-1, 1]$ a (f_n) je cauchyovská, protože pro $m \le n$ máme

$$d(f_m, f_n) = \int_{-1}^{1} |f_m(x) - f_n(x)|\, dx \le \int_{-1/m}^{1/m} 1\, dx = 2/m.$$

Neexistuje však funkce $f \in \mathcal{C}[-1, 1]$, pro níž by $f_n \to f$ pro $n \to \infty$. Taková funkce f by podle definice f_n musela být na intervalu $[-1, 0)$ identicky rovna -1 a na intervalu $(0, 1]$ identicky rovna 1, což je pro funkci spojitou na $[-1, 1]$ nemožné.

4. Kompaktní metrický prostor je vždy úplný (úloha 1). Naopak to obecně neplatí, $\mathbf{R}$ je úplný a nekompaktní metrický prostor.

5. Uvažme euklidovské metrické prostory $\mathbf{R}$ a $(-\pi/2, \pi/2)$. Bijekce

$$f(x) = \arctan(x) : \ \mathbf{R} \to (-\pi/2, \pi/2).$$

je homeomorfismus, f i $f^{-1}(x) = \tan(x) : \ (-\pi/2, \pi/2) \to \mathbf{R}$ jsou spojitá zobrazení. Ovšem $\mathbf{R}$ je úplný metrický prostor, ale $(-\pi/2, \pi/2)$ nikoli. Úplnost metrického prostoru není, na rozdíl od kompaktnosti, topologická vlastnost, není určena pouze otevřenými množinami, závisí i na metrice. Nicméně se úplnost se zachovává homeomorfismem, který je v obou směrech stejnoměrně spojitý (funkce $\tan x$ není na $(-\pi/2, \pi/2)$ stejnoměrně spojitá).

Tvrzení 1.9. *Úplnost metrického prostoru se zachovává následujícími operacemi.*

1. *Přechodem k uzavřenému podprostoru.*

2. *Obrazem stejnoměrně spojitým prostým zobrazením, pokud je i inverzní zobrazení stejnoměrně spojité.*

3. *Kartézským součinem.*

Důkaz. 1. Úloha 2. 2. Úloha 3. 3. Úloha 4. □

Banachova věta o pevném bodu. Pomocí úplnosti se dá o mnohých rovnicích dokázat, že v úplném metrickém prostoru mají řešení. Typickým příkladem je rovnice $x^2 = 2$, která sice nemá řešení v oboru racionálních čísel, ale v širším oboru reálných čísel se díky úplnosti dokáže existence řešení. Popíšeme obecný postup, který zaručuje existenci řešení jisté třídy rovnic v úplných metrických prostorech.

Zobrazení $f: M \to M$ metrického prostoru (M, d) do sebe je *kontrahující*, když pro nějaké číslo $q \in \mathbf{R}$ splňující $0 < q < 1$ pro každé dva body x, y v M platí

$$d(f(x), f(y)) \le qd(x, y).$$

Kontrahující zobrazení tedy kontrahuje, zmenšuje vzdálenost každých dvou bodů alespoň o pevný faktor q menší než 1. Je jasné, že kontrahující zobrazení je stejnoměrně spojité. *Pevným bodem* zobrazení f množiny X do sebe rozumíme bod a z X splňující $f(a) = a$. Posloupnost $(x_n) \subset X$ je *posloupností iterací* zobrazení $f: X \to X$, když pro $n = 1, 2, \ldots$ platí $x_{n+1} = f(x_n)$ ($x_1 \in X$ je libovolný startovací bod posloupnosti iterací).

Věta 1.10 (Banachova věta o pevném bodu). *Kontrahující zobrazení f úplného metrického prostoru (M, d) do sebe má právě jeden pevný bod a každá posloupnost iterací $(x_n) \subset M$ zobrazení f k němu konverguje.*

Důkaz. Uvažme libovolnou posloupnost iterací (x_n) kontrahujícího zobrazení f. Protože $x_n = f(x_{n-1})$ a f je kontrahující s konstantou q, pro každé $n \in \mathbf{N}$ máme odhad

$$d(x_{n+1}, x_n) \le qd(x_n, x_{n-1}) \le q^2 d(x_{n-1}, x_{n-2}) \le \ldots \le q^{n-1} d(x_2, x_1).$$

Pomocí trojúhelníkové nerovnosti pak pro každé $k, n \in \mathbf{N}$ máme

$$\begin{aligned} d(x_{n+k}, x_n) &\le d(x_{n+k}, x_{n+k-1}) + d(x_{n+k-1}, x_{n+k-2}) + \ldots + d(x_{n+1}, x_n) \\ &\le d(x_2, x_1)(q^{n+k-2} + q^{n+k-3} + \ldots + q^{n-1}) \\ &< d(x_2, x_1)(q^{n-1} + q^n + q^{n+1} + \ldots) \\ &= d(x_2, x_1) q^{n-1}/(1-q) \\ &\to 0 \text{ pro } n \to \infty \text{ (neboť } 0 < q < 1). \end{aligned}$$

Posloupnost (x_n) je tedy cauchyovská. Díky úplnosti prostoru M má limitu a. Ze spojitosti f pak plyne, že a je pevným bodem f:

$$a = \lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n) = f(a).$$

Nechť a, b jsou dva pevné body f. Pak

$$d(a,b) = d(f(a), f(b)) \leq q d(a,b),$$

což vynucuje $d(a,b) = 0$ a $a = b$. Pevný bod je jediný. $\square$

Dá se ukázat (úloha 5), že věta platí i za zdánlivě slabšího předpokladu, že kontrahující je jen nějaká iterace $f^{(n)}(x) = f(f(\ldots(f(x))))$ zobrazení f.

Ukážeme použití Věty 1.10 při řešení diferenciálních rovnic. Začneme jednoduchou rovnicí $y'(x) = y(x)$, kdy chceme najít funkci rovnou své derivaci. Řešením této rovnice je exponenciála $y(x) = \exp(x)$ a spousta dalších funkcí, jako třeba $-3\exp(x+10)$. Pro každou dvojici reálných čísel a, b dokonce existuje takové řešení, že $y(a) = b$, sice $y(x) = b\exp(x-a)$. Jak uvidíme, s tímto požadavkem je řešení (lokálně) jednoznačné. Pomocí Banachovy věty o pevném bodu se dá lokální existence a jednoznačnost řešení dokázat pro širokou třídu diferenciálních rovnic

$$(*) \left\{ \begin{array}{lcl} y(a) & = & b \\ y'(x) & = & f(x, y(x)). \end{array} \right.$$

Zde $f:\ \mathbf{R}^2 \to \mathbf{R}$ je zadaná funkce (pravá strana rovnice) a $a, b \in \mathbf{R}$ jsou zadaná čísla. Hledáme reálnou funkci $y(x)$ a otevřený interval I obsahující a, že $y(x)$ je na I definovaná, $y(a) = b$ (říkáme, že $y(x)$ splňuje *počáteční podmínku* $y(a) = b$) a $y(x)$ má I derivaci splňující pro každé $x \in I$ druhý vztah v (*), tj. vlastní diferenciální rovnici.

Věta 1.11 (Picardova). *Pokud je $f:\ \mathbf{R}^2 \to \mathbf{R}$ spojitá a existuje konstanta $M > 0$ taková, že pro každá tři čísla $u, v, w \in \mathbf{R}$ platí*

$$|f(u,v) - f(u,w)| \leq M|v - w|,$$

pak každý bod $a \in \mathbf{R}$ má okolí $I = (a-\delta, a+\delta)$, na němž má úloha (*) jednoznačné řešení $y(x)$.

Důkaz. Budeme pracovat na intervalu $I = (a-\delta, a+\delta)$ pro nějaké $\delta > 0$ a na jeho uzávěru $J = [a-\delta, a+\delta]$. Z vlastností Riemannova integrálu (výpočet Riemannova integrálu Newtonovým integrálem, Riemannův integrál jako funkce horní integrační meze) plyne, že pro spojitou funkci f je úloha (*) ekvivalentní rovnici

$$y(x) = b + \int_a^x f(t, y(t))\, dt,\ \ x \in I$$

—je-li $y(x)$ na I řešením úlohy (*), je řešením rovnice a naopak. Ukážeme, že pro dostatečně malé δ má na intervalu I poslední rovnice—a tedy i úloha (*)—jednoznačné řešení $y(x)$. Pravá strana poslední rovnice definuje zobrazení A, které funkci $y(x)$ spojité na J přiřadí funkci $z(x)$,

$$z(x) = A(y(x)) = b + \int_a^x f(t, y(t))\, dt.$$

Integrál je spojitou funkcí své horní integrační meze, takže $z(x)$ je na J rovněž spojitá (dokonce má na J spojitou první derivaci: $z'(x) = f(x, y(x))$). Máme zobrazení

$$A:\ \mathcal{C}[a-\delta, a+\delta] \to \mathcal{C}[a-\delta, a+\delta].$$

Odvodíme, že A má pro dostatečně malé δ jednoznačný pevný bod y.

Pro tento účel vybavíme $\mathcal{C}[a-\delta, a+\delta]$ maximovou metrikou $d(\cdot,\cdot)$, čímž dostaneme úplný metrický prostor (viz příklad 2), a použijeme Větu 1.10. Uvidíme, že pro dostatečně malé δ je A kontrahující. Nechť $y(x)$ a $z(x)$ jsou dvě funkce z $\mathcal{C}[a-\delta, a+\delta]$. Pak

$$\begin{aligned}
d(A(y), A(z)) &= \max_{x\in J} |A(y)(x) - A(z)(x)| \\
&= \max_{x\in J} \left| \int_a^x f(t, y(t))\, dt - \int_a^x f(t, z(t))\, dt \right| \\
&= \max_{x\in J} \left| \int_a^x (f(t, y(t)) - f(t, z(t)))\, dt \right| \\
&\leq \max_{x\in J} \left| \int_a^x |f(t, y(t)) - f(t, z(t))|\, dt \right| \\
&\leq \max_{x\in J} \left| \int_a^x M|y(t) - z(t)|\, dt \right| \\
&\leq \max_{x\in J} \left| \int_a^x M \max_{t\in J} |y(t) - z(t)|\, dt \right| \\
&= \max_{x\in J} \left| \int_a^x M d(y, z)\, dt \right| \\
&= M\delta \cdot d(y, z).
\end{aligned}$$

Zvolíme-li $\delta \leq \frac{1}{2M}$, máme $d(A(y), A(z)) \leq \frac{1}{2} d(y, z)$ pro libovolné dvě funkce z $\mathcal{C}[a-\delta, a+\delta]$ a zobrazení A je kontrahující. Podle Věty 1.10 má jednoznačný pevný bod a Věta 1.11 je dokázána. □

Když reálná funkce dvou proměnných $f(u, v)$ splňuje pro nějakou konstantu $M > 0$ na množině $D \subset \mathbf{R}^2$ podmínku Věty 1.11, to jest

$$\forall\ (u, v), (u, w) \in D:\ |f(u, v) - f(u, w)| \leq M|v - w|,$$

řekneme, že f je na D *lipschitzovská* nebo že na D splňuje *Lipschitzovu podmínku* (v druhé proměnné). Funkce $f(u, v) = v$ z úvodního příkladu je lipschitzovská na celém $\mathbf{R}^2$, třeba s konstantou $M = 1$. Funkce $b\exp(x - a)$ je proto pro každé dvě čísla $a, b \in \mathbf{R}$ jednoznačným lokálním řešením diferenciální rovnice $y(a) = b$, $y'(x) = y(x)$.

Podmínka lipschitzovskosti na celém $\mathbf{R}^2$ je zbytečně silná a v praxi často není splněna. Stačí však její lokální splnění. Dokažte si (úloha 7), že Věta 1.11 platí i za slabšího předpokladu lokální lipschitzovskosti.

Úlohy

1. Dokažte, že kompaktní metrický prostor je úplný.
2. Dokažte, že podmnožina úplného metrického prostoru indukuje úplný podprostor, právě když je uzavřená.
3. Dokažte, že když $f : M \to N$ je bijekce mezi metrickými prostory, přičemž f i f^{-1} je stejnoměrně spojité zobrazení, pak je prostor M úplný, právě když je prostor N úplný.
4. Kartézský součin dvou úplných metrických prostorů je úplný.
5. Ukažte, že zobrazení úplného metrického prostoru do sebe, jehož nějaká iterace je kontrahující, má jediný pevný bod.
6. Dokažte pomocí Banachovy věty o pevném bodu, že polynom $x^2 - 2$ má v $\mathbf{R}$ kořen. Jak bude vypadat posloupnost iterací konvergující k $\sqrt{2}$? Návod: graf funkce $x^2 - 2$ aproximujte tečnou.
7. Nechť $D \subset \mathbf{R}^2$ je otevřená množina, $(a, b) \in D$ a $f : D \to \mathbf{R}$ je spojitá funkce, která je na D lipschitzovská ve druhé proměnné. Pak má diferenciální rovnice (*) lokálně jednoznačné řešení.

5. přednáška 29. října 2007

Závěrem první kapitoly o metrických prostorech se zmíníme o třech důležitých typech souvisejících matematických struktur.

Topologické prostory. Topologické prostory jsou chudší než metrické prostory: zapomeneme na metriku a necháme si jen otevřené množiny. *Topologický prostor*, stručněji *topologie*, je dvojice $(X, \mathcal{T})$, kde $\mathcal{T}$ je systém (ne nutně všech) podmnožin množiny X, který obsahuje množiny $\emptyset$ a X a je uzavřený na libovolná sjednocení a na konečné průniky. Explicitně,

(a) $\emptyset \in \mathcal{T},\ X \in \mathcal{T}$,
(b) $\{O_i \mid i \in I\} \subset \mathcal{T} \Rightarrow \bigcup_{i \in I} \in \mathcal{T}$ a
(c) $\{O_i \mid i \in I\} \subset \mathcal{T}$, I konečná $\Rightarrow \bigcap_{i \in I} \in \mathcal{T}$.

Prvkům systému $\mathcal{T}$ se říká *otevřené množiny* topologie $\mathcal{T}$. Jak víme (Tvrzení 1.1), systém otevřených množin metrického prostoru tvoří topologický prostor. Je ale mnoho topologií, které nelze vytvořit z metrického prostoru (úlohy 1 a 2). Na topologické prostory lze přenést mnohé z metrických prostorů (viz spojitost—Tvrzení 1.3 a kompaktnost—Věta 1.8).

Normované prostory. Normované prostory jsou bohatší mež metrické prostory, kromě metriky nesou další strukturu. *Normovaný prostor* je vektorový prostor X nad tělesem $\mathbf{R}$ vybavený zobrazením $\|\cdot\|: X \to \mathbf{R}$, zvaným *norma*, splňujícím tři axiomy (pro všechny $x, y \in X$ a $\lambda \in \mathbf{R}$):

(a) $\|x\| \ge 0$, $\|x\| = 0 \iff x = 0$,
(b) $\|\lambda x\| = |\lambda| \cdot \|x\|$ a
(c) $\|x + y\| \le \|x\| + \|y\|$ (trojúhelníková nerovnost).

Normovaný vektorový prostor je metrickým prostorem, funkce

$$d(x, y) := \|x - y\|$$

je metrika (úloha 3). Protože se zrodila z normy, je translačně invariantní (česky: nemění se při posunutí), pro každé tři vektory x, y, z z X máme $d(x+z, y+z) = d(x, y)$. *Banachův prostor* je úplný normovaný prostor, tj. odvozená metrika je úplná.

Metriky $d_p(\cdot, \cdot)$ na $\mathbf{R}^n$, pro $p \ge 1$ a $p = \infty$ (viz 1. přednáška), jsou odvozeny z norem

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}, \text{ resp. } \|x\|_\infty = \max_{1 \le i \le n} |x_i|.$$

Podobně i pro analogické metriky na prostoru spojitých funkcí $C[a, b]$. Všechny tyto prostory jsou Banachovy.

Prostory se skalárním součinem jsou ještě bohatší. *Prostor se skalárním součinem* (PSS) je vektorový prostor X nad tělesem $\mathbf{R}$, který je vybaven zobrazením $\langle \cdot, \cdot \rangle : X \times X \to \mathbf{R}$, zvaným *skalární součin*, splňujícím tři axiomy (pro všechny $x, x', y \in X$ a $\kappa, \lambda \in \mathbf{R}$):

a) $\langle x, x \rangle \geq 0$, $\langle x, x \rangle = 0 \iff x = 0$,
b) $\langle x, y \rangle = \langle y, x \rangle$ a
c) $\langle \kappa x + \lambda x', y \rangle = \kappa \langle x, y \rangle + \lambda \langle x', y \rangle$.

Symetrie v (b) a linearita v prvním argumentu v (c) dávají, že skalární součin je lineární i ve druhém argumentu, je to bilineární zobrazení. Na rozdíl od metriky a normy může skalární součin nabývat záporných hodnot, ale na diagonále $x = y$ musí být nezáporný. Měří úhel mezi vektory a lze z něj odvodit normu a tedy i metriku, jak hned ukážeme. Příkladem PSS je euklidovský prostor $\mathbf{R}^m$ se skalárním součinem

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_m y_m.$$

Dalším příkladem je prostor spojitých funkcí $C[a, b]$ se skalárním součinem

$$\langle f, g \rangle = \int_a^b f(x) g(x) \, dx.$$

Následující nerovnost je jedna z nejdůležitějších v matematice.

Věta 2.1 (Cauchyova–Schwarzova nerovnost). *V prostoru se skalárním součinem $(X, \langle \cdot, \cdot \rangle)$ pro každé dva vektory x a y platí, že*

$$\langle x, y \rangle^2 \leq \langle x, x \rangle \langle y, y \rangle.$$

Rovnost nastává, právě když je jeden z vektorů skalárním násobkem druhého, $x = \lambda y$ pro $\lambda \in \mathbf{R}$.

Důkaz. Byl v Lineární algebře, proto ho zde neuvádíme. □

Uvažme zobrazení $\| \cdot \| : X \to \mathbf{R}$ definované jako

$$\|x\| := \sqrt{\langle x, x \rangle}.$$

Dá se ukázat (úloha 4), že toto zobrazení je norma. PSS je tedy také normovaný prostor (a tedy i metrický prostor a topologický prostor). Cauchyovu–Schwarzovu nerovnost můžeme pomocí značení pro normu ekvivalentně zapsat ve tvaru

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|.$$

Hilbertův prostor je úplný PSS, tj. odvozená metrika

$$d(x, y) = \|x - y\| = \sqrt{\langle x - y, x - y \rangle}$$

je úplná. Euklidovský prostor $\mathbf{R}^n$ je Hilbertův, ale prostor spojitých funkcí $\mathcal{C}[a, b]$ Hilbertův není (úloha 5).

Kapitola 2. Diferenciální počet funkcí více proměnných.

Jak víme z MAI, za určitých předpokladů se funkce jedné proměnné dají lokálně aproximovat pomocí lineárních funkcí, s nimiž se lépe počítá. Konkrétně, má-li funkce $f:\ (a-\delta, a+\delta) \to \mathbf{R}$ v bodu a vlastní derivaci, máme v okolí a lineární aproximaci

$$f(a+h) = f(a) + f'(a) \cdot h + o(h),\ h \to 0.$$

Ve druhé kapitole ji zobecníme pro funkce s více proměnnými, a pak i pro zobrazení složená z několika takových funkcí. Budeme pracovat v euklidovském prostoru $\mathbf{R}^m$ s obvyklým skalárním součinem $\langle x, y\rangle = \sum_{i=1}^m x_i y_i$, s odvozenou euklidovskou normou

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_m^2}$$

a euklidovskou metrikou

$$d(x,y) = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2 + \ldots + (x_m-y_m)^2},$$

a s funkcemi o m proměnných

$$f:\ D \to \mathbf{R}$$

definovanými na otevřených množinách D v $\mathbf{R}^m$.

Směrová derivace, parciální derivace, diferenciál. *Směrovou derivací* funkce $f:\ D \to \mathbf{R}$ v bodu a ve směru v, kde $D \subset \mathbf{R}^m$ je otevřená množina, bod a leží v D a v z $\mathbf{R}^m$ je nenulový vektor, rozumíme limitu

$$\mathrm{D}_v f(a) := \lim_{t \to 0} \frac{f(a+tv) - f(a)}{t},$$

pokud existuje. Představte si D jako oblast v třírozměrném euklidovském prostoru, kde funkce f měří teplotu a kterou prolétá po přímočaré dráze částice. Směrová derivace $\mathrm{D}_v f(a)$ pak udává okamžitou změnu teploty částice ve chvíli, kdy se nachází v bodu a a má vektor rychlosti v.

Parciální derivace funkce f v bodě a podle proměnné x_i je směrová derivace $\mathrm{D}_{e_i} f(a)$, kde e_i je i-tý vektor kanonické báze, tj. $e_i = (0, 0, \ldots, 0, 1, 0, 0, \ldots, 0)$ má na i-tém místě 1 a jinde nuly. Značíme ji $\frac{\partial f}{\partial x_i}(a)$. Explicitně,

$$\frac{\partial f}{\partial x_i}(a) = \lim_{h \to 0} \frac{f(a_1, \ldots, a_{i-1}, a_i + h, a_{i+1}, \ldots, a_m) - f(a_1, a_2, \ldots, a_m)}{h}.$$

Má-li f parciální derivaci podle x_i v každém bodě D, dostáváme funkci

$$\frac{\partial f}{\partial x_i}:\ D \to \mathbf{R},$$

která každému bodu a z D přiřazuje hodnotu $\frac{\partial f}{\partial x_i}(a)$. Vektor hodnot všech parciálních derivací funkce f v bodě a je *gradient* funkce f v a,

$$\nabla f(a) := (\tfrac{\partial f}{\partial x_1}(a), \tfrac{\partial f}{\partial x_2}(a), \ldots, \tfrac{\partial f}{\partial x_m}(a)).$$

Počítat parciální derivace už umíme, při výpočtu $\frac{\partial f}{\partial x_i}$ se proměnné různé od x_i berou jako konstanty a f tak derivujeme jako funkci jediné proměnné x_i. Například

$$\frac{\partial(x^3 y \sin(yz) + x \log z)}{\partial y} = x^3(\sin(yz) + zy\cos(yz)).$$

Funkce f má v bodě a *(totální) diferenciál*, jinými slovy f je v a *diferencovatelná*, když existuje takové lineární zobrazení $L:\ \mathbf{R}^m \to \mathbf{R}$, že

$$\lim_{\|h\| \to 0} \frac{f(a+h) - f(a) - L(h)}{\|h\|} = 0.$$

Toto lineární zobrazení L nazýváme *diferenciálem* a značíme $\mathrm{D}f(a)$, jeho hodnota $L(h)$ na vektoru h pak je $\mathrm{D}f(a)(h)$. Podstatný rozdíl ve srovnání se směrovou a parciální derivací je ten, že ty jsou pouhá čísla, kdežto diferenciál je složitější věc, lineární zobrazení.

Směrová derivace, parciální derivace a diferenciál funkce f v bodu a dávají lokální aproximace f poblíž a lineární funkcí:

$$\begin{aligned}
f(a+tv) &= f(a) + \mathrm{D}_v f(a) \cdot t + o(t),\ t \to 0, \\
f(a+te_i) &= f(a) + \frac{\partial f}{\partial x_i}(a) \cdot t + o(t),\ t \to 0, \\
f(a+h) &= f(a) + \mathrm{D}f(a)(h) + o(\|h\|),\ \|h\| \to 0.
\end{aligned}$$

V prvních dvou vztazích je t reálné číslo jdoucí k nule a aproximace platí pouze pro argumenty funkce na přímce jdoucí bodem a ve směru v, resp. ve směru i-té souřadnicové osy. Ve třetím vztahu h probíhá body $\mathbf{R}^m$ a aproximace platí pro všechny argumenty funkce v okolí bodu a. Diferencovatelnost je silnější vlastnost f než existence směrových nebo parciálních derivací, z nichž neplyne ani spojitost funkce v daném bodě.

Příklady. 1. Funkce $f = f(x,y):\ \mathbf{R}^2 \to \mathbf{R}$ definovaná jako 1 na množině $\{(x,y) \in \mathbf{R}^2:\ y = x^2, x \neq 0\}$ a jako 0 pro všechny zbylé body roviny má v počátku všechny směrové derivace (jsou rovné nule), ale není tam spojitá.

2. Podobně, definujeme-li f jako 1 na souřadnicových osách, tj. na množině $\{(x,y) \in \mathbf{R}^2:\ xy = 0\}$, a jako 0 pro všechny zbylé body roviny, má f v počátku obě parciální derivace (jsou rovné nule), ale kromě nich už žádnou další směrovou derivaci. Funkce f opět není v počátku spojitá.

Pojem diferenciálu rozšíříme na obecnější situaci, kdy $f:\ D \to \mathbf{R}^n$ ($D \subset \mathbf{R}^m$ je otevřená množina) je zobrazení dané n-ticí souřadnicových funkcí: $f =$

$(f_1, f_2, \ldots, f_n)$ a $f_i: D \to \mathbf{R}$. Řekneme, že zobrazení f má v bodě a z D *diferenciál* nebo že tam je *diferencovatelné*, existuje-li lineární zobrazení $L: \mathbf{R}^m \to \mathbf{R}^n$ takové, že

$$\lim_{\|h\| \to 0} \frac{\|f(a+h) - f(a) - L(h)\|}{\|h\|} = 0.$$

(Norma v čitateli je v $\mathbf{R}^n$, norma ve jmenovateli je v $\mathbf{R}^m$.) Lineární zobrazení L značíme $\mathrm{D}f(a)$. Z aproximačního pohledu to opět znamená, že

$$f(a+h) = f(a) + \mathrm{D}f(a)(h) + \alpha(h), \quad \text{kde } \|h\| \to 0 \Rightarrow \|\alpha(h)\|/\|h\| \to 0.$$

Tvrzení 2.1. *Buď dáno zobrazení $f = (f_1, f_2, \ldots, f_n): D \to \mathbf{R}^n$, kde $D \subset \mathbf{R}^m$ je otevřená množina, a bod a v D.*

1. *Diferenciál f v a je určený jednoznačně.*
2. *Zobrazení f je diferencovatelné v a, právě když je každá souřadnicová funkce f_i diferencovatelná v a.*
3. *Když je f diferencovatelné v bodu a, potom je v a spojité.*

Důkaz. 1. Úloha 6. 2. Úloha 7. 3. Zřejmé. □

Úlohy

1. Nechť $(X, \mathcal{T})$ je topologický prostor vzniklý z metriky, tj. tvořený otevřenými množinami nějakého metrického prostoru (X, d). Dokažte, že pro každé dva různé body a, b z X existují takové dvě otevřené množiny U, V z $\mathcal{T}$, že $a \in U$, $b \in V$ a $U \cap V = \emptyset$. Topologiím s touto vlastností se říká *Hausdorffovy.*
2. Uveďte příklad topologie, která není Hausdorffova, takže nevznikla z metriky.
3. Dokažte, že funkce $d(x, y) := \|x - y\|$ definovaná na normovaném prostoru je metrika.
4. Ověřte, že funkce $\|x\| := \sqrt{\langle x, x \rangle}$ na prostoru se skalárním součinem je norma.
5. Ukažte, že metrický prostor spojitých funkcí $\mathcal{C}[a, b]$ s metrikou danou skalárním součinem $\langle f, g \rangle = \int_a^b f(x)g(x)\,dx$ není úplný.
6. Dokažte část 1 Tvrzení 2.1.
7. Dokažte část 2 Tvrzení 2.1.

6. přednáška 5. listopadu 2007

Souvislost diferenciálu a parciálních derivací. Diferenciál implikuje parciální derivace a spojité parciální derivace implikují diferenciál.

Tvrzení 2.3. *Když je funkce*

$$f:\ U \to \mathbf{R},\quad U \subset \mathbf{R}^m \text{ je okolí bodu } a,$$

diferencovatelná v a, pak má v a všechny parciální derivace a jejich hodnoty určují diferenciál:

$$\begin{aligned} \mathrm{D}f(a)(h) &= \frac{\partial f}{\partial x_1}(a)\cdot h_1 + \frac{\partial f}{\partial x_2}(a)\cdot h_2 + \cdots + \frac{\partial f}{\partial x_m}(a)\cdot h_m \\ &= \langle \nabla f(a), h\rangle. \end{aligned}$$

Také má v a všechny směrové derivace a platí $\mathrm{D}_v f(a) = \mathrm{D}f(a)(v)$.

Důkaz. Z linearity diferenciálu $L = \mathrm{D}f(a)$ máme

$$L(h) = L(h_1e_1 + h_2e_2 + \cdots + h_me_m) = L(e_1)h_1 + \cdots + L(e_m)h_m,$$

kde e_i je i-tý vektor kanonické báze. Ovšem

$$\begin{aligned} \frac{\partial f}{\partial x_i}(a) &= \lim_{t\to 0}\frac{f(a+te_i)-f(a)}{t} \\ &= \lim_{t\to 0}\frac{L(te_i)+o(\|te_i\|)}{t} \\ &= \lim_{t\to 0}\frac{tL(e_i)+o(t)}{t} \\ &= L(e_i), \end{aligned}$$

a tak $L(e_i) = \frac{\partial f}{\partial x_i}(a)$. Tvrzení o směrové derivaci plyne z definice a z právě dokázané formule pro diferenciál. □

Obecně je pro zobrazení $f:\ D \to \mathbf{R}^n$ diferenciál $L = \mathrm{D}f(a):\ \mathbf{R}^m \to \mathbf{R}^n$ reprezentován maticí tvaru $n \times m$ a L se na vektor h aplikuje maticovým násobením:

$$L(h) = \begin{pmatrix} L(h)_1 \\ L(h)_2 \\ \vdots \\ L(h)_n \end{pmatrix} = \begin{pmatrix} l_{1,1} & l_{1,2} & \dots & l_{1,m} \\ l_{2,1} & l_{2,2} & \dots & l_{2,m} \\ \vdots & \vdots & \dots & \vdots \\ l_{n,1} & l_{n,2} & \dots & l_{n,m} \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_m \end{pmatrix}.$$

Podle předešlého tvrzení a bodu 2 Tvrzení 2.1 má tato matice v i-tém řádku gradient souřadnicové funkce f_i v bodě a, takže

$$l_{i,j} = \frac{\partial f_i}{\partial x_j}(a).$$

Důsledek. *Diferenciál zobrazení* $f: D \to \mathbf{R}^n$ *v bodě* a*, kde* $D \subset \mathbf{R}^m$ *je okolí* a *a* f *má souřadnicové funkce* $f = (f_1, f_2, \ldots, f_n)$*, je dán tzv.* Jacobiho maticí *zobrazení* f *v bodě* a*,*

$$\left(\frac{\partial f_i}{\partial x_j}(a)\right)_{i,j=1}^{n,m} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \ldots & \frac{\partial f_1}{\partial x_m}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \ldots & \frac{\partial f_2}{\partial x_m}(a) \\ \vdots & \vdots & \ldots & \vdots \\ \frac{\partial f_n}{\partial x_1}(a) & \frac{\partial f_n}{\partial x_2}(a) & \ldots & \frac{\partial f_n}{\partial x_m}(a) \end{pmatrix}.$$

Je-li tato matice čtvercová, nazývá se její determinant jacobiánem.

Věta 2.4. *Nechť* $U \subset \mathbf{R}^m$ *je okolí bodu* $a \in \mathbf{R}^m$*. Pokud má funkce* $f: U \to \mathbf{R}$ *na* U *všechny parciální derivace a ty jsou v bodě* a *spojité, pak je* f *v bodě* a *diferencovatelná.*

Důkaz. Pro jednoduchost nechť $m = 2$ a $a = \overline{0} = (0, 0)$. (Viz úlohu 1.) Označíme $h = (h_1, h_2)$ a $h' = (h_1, 0)$. Přírůstek $f(h) - f(\overline{0})$ napíšeme pomocí přírůstků ve směrech souřadnicových os:

$$f(h) - f(\overline{0}) = (f(h) - f(h')) + (f(h') - f(\overline{0})).$$

Na úsečkách $h'h$ a $\overline{0}h'$ funkce f závisí pouze na proměnné x_2, resp. na x_1. Použijeme Lagrangeovu větu o střední hodnotě:

$$f(h) - f(\overline{0}) = \frac{\partial f}{\partial x_2}(\zeta_2) \cdot h_2 + \frac{\partial f}{\partial x_1}(\zeta_1) \cdot h_1,$$

kde ζ_2 (resp. ζ_1) je jistý vnitřní bod úsečky $h'h$ (resp. $\overline{0}h'$). Oba body leží v otevřené kouli $B(\overline{0}, \|h\|)$. Díky spojitosti v počátku máme

$$\frac{\partial f}{\partial x_2}(\zeta_2) = \frac{\partial f}{\partial x_2}(\overline{0}) + \alpha(\zeta_2) \;\text{ a }\; \frac{\partial f}{\partial x_1}(\zeta_1) = \frac{\partial f}{\partial x_1}(\overline{0}) + \beta(\zeta_1),$$

kde $\alpha(h), \beta(h) = o(1)$ pro $h \to \overline{0}$ (tj. pro každé $\varepsilon > 0$ máme $\delta > 0$, že $\|h\| < \delta \Rightarrow |\alpha(h)| < \varepsilon$ a podobně pro $\beta(h)$). Tedy

$$f(h) - f(\overline{0}) = \frac{\partial f}{\partial x_2}(\overline{0}) \cdot h_2 + \frac{\partial f}{\partial x_1}(\overline{0}) \cdot h_1 + \alpha(\zeta_2)h_2 + \beta(h_1)h_1.$$

Díky nerovnostem $0 < \|\zeta_1\|, \|\zeta_2\| < \|h\|$ a $|h_1|, |h_2| \leq \|h\|$ je jasné, že $\alpha(\zeta_2)h_2 + \beta(\zeta_1)h_1 = o(h)$ pro $h \to \overline{0}$. Funkce f je diferencovatelná v počátku. □

Lagrangeova věta o střední hodnotě pro funkce více proměnných. Následující dvě tvrzení zobecňují Lagrangeovu větu o střední hodnotě a fakt, že nulovost derivace implikuje konstantnost funkce.

Tvrzení 2.5. *Nechť* $U \subset \mathbf{R}^m$ *je otevřená množina a* $u = ab$ *je úsečka s koncovými body* a *a* b *ležící v* U*. Nechť je funkce* $f: U \to \mathbf{R}$ *na* u *spojitá a má* *v každém vnitřním bodě u diferenciál. Pak existuje vnitřní bod ζ úsečky u s vlastností, že*

$$f(b) - f(a) = \mathrm{D}f(\zeta)(b-a).$$

Důkaz. Položíme $F(t) = f(a + th)$, kde $h = b - a$ a reálné číslo t probíhá interval $[0, 1]$. Funkce F je patrně spojitá na $[0, 1]$ a v $t \in (0, 1)$ má derivaci

$$\begin{aligned} F'(t) &= \lim_{\Delta \to 0} \frac{f(a+th+\Delta h) - f(a+th)}{\Delta} \\ &= \lim_{\Delta \to 0} \frac{\mathrm{D}f(a+th)(\Delta h) + o(\|\Delta h\|)}{\Delta} \\ &= \mathrm{D}f(a+th)(h). \end{aligned}$$

Podle Lagrangeovy věty o střední hodnotě existuje takové $t_0 \in (0, 1)$, že $F(1) - F(0) = F'(t_0)$. Odtud

$$f(b) - f(a) = F(1) - F(0) = F'(t_0) = \mathrm{D}f(a + t_0 h)(h) = \mathrm{D}f(\zeta)(h),$$

kde $\zeta = a + t_0 h$. □

Řekneme, že otevřená množina D v $\mathbf{R}^m$ je *souvislá*, když lze každé její dva body spojit lomenou čarou, která celá leží v D. Například koule s jednotkovým poloměrem v $\mathbf{R}^m$, celé $\mathbf{R}^m$ a $\mathbf{R}^3 \setminus L$, kde L je sjednocení konečně mnoha přímek, jsou souvislé otevřené množiny. Na druhou stranu množina $B \setminus R$, kde B je otevřená koule v $\mathbf{R}^3$ a R rovina protínající B, není souvislá.

Tvrzení 2.6. *Má-li reálná funkce m proměnných v každém bodě otevřené a souvislé množiny nulový diferenciál, je na této množině konstantní.*

Důkaz. Nechť $U \subset \mathbf{R}^m$ je otevřená a souvislá množina a funkce $f : U \to \mathbf{R}$ má na U nulový diferenciál. Vezmeme dva libovolné body $a, b \in U$ a spojíme je lomenou čarou $s = s_1 s_2 \dots s_r$ ležící v U. Pro libovolnou úsečku $s_i = a_i b_i$ z s máme podle předchozího tvrzení a předpokladu o f, že

$$f(a_i) - f(b_i) = \mathrm{D}f(\zeta)(a_i - b_i) = 0$$

(ζ je nějaký vnitřní bod s_i), tedy $f(a_i) = f(b_i)$. Hodnoty funkce f na koncích všech úseček s_i se rovnají a tedy $f(a) = f(b)$. □

Tvrzení 2.3, 2.4 a 2.6 dávají následující důsledek.

Důsledek. *Má-li reálná funkce m proměnných v každém bodě otevřené a souvislé množiny každou parciální derivaci nulovou, je na této množině konstantní.*

Počítání s parciálními derivacemi a diferenciály. Pro dvě funkce $f, g : U \to \mathbf{R}$, které jsou definované na okolí $U \subset \mathbf{R}^m$ bodu $a \in U$ a mají v bodě a

i-tou parciální derivaci, máme pro i-tou parciální derivaci jejich lineární kombinace, součinu a podílu stejné vzorce jako v případě funkcí jedné proměnné (místo $\frac{\partial f}{\partial x_i}$ píšeme $\partial_i f$):

$$\begin{aligned} \partial_i(\kappa f + \lambda g)(a) &= \kappa \partial_i f(a) + \lambda \partial_i g(a) \\ \partial_i(fg)(a) &= g(a)\partial_i f(a) + f(a)\partial_i g(a) \\ \partial_i(f/g)(a) &= \frac{g(a)\partial_i f(a) - f(a)\partial_i g(a)}{g(a)^2} \quad \text{(pokud } g(a) \neq 0\text{).} \end{aligned}$$

Tyto vzorce fakticky jsou vzorce pro funkce jedné proměnné, protože ∂_i se počítá z funkce závisející jen na x_i. Podobně pro diferenciály.

Tvrzení 2.7. *Nechť $U \subset \mathbf{R}^m$ je otevřená množina, $a \in U$ a $f, g : U \to \mathbf{R}$ jsou dvě funkce, obě diferencovatelné v bodě a.*
1. *Pro všechny $\kappa, \lambda \in \mathbf{R}$ je i funkce $\kappa f + \lambda g$ v bodu a diferencovatelná a*

$$\mathrm{D}(\kappa f + \lambda g)(a) = \kappa \mathrm{D} f(a) + \lambda \mathrm{D} g(a).$$

2. *Součinová funkce fg je diferencovatelná v a a*

$$\mathrm{D}(fg)(a) = g(a)\mathrm{D} f(a) + f(a)\mathrm{D} g(a).$$

3. *Pokud $g(a) \neq 0$, je podílová funkce f/g diferencovatelná v a a*

$$\mathrm{D}(f/g)(a) = \frac{1}{g(a)^2}\Big(g(a)\mathrm{D} f(a) - f(a)\mathrm{D} g(a)\Big).$$

Důkaz. Tyto vzorce plynou z analogických vzorců pro parciální derivace a z Tvrzení 2.3. (Viz úlohu 2.) □

Vzorec pro diferenciál lineární kombinace v části 1 platí obecněji i pro zobrazení $f, g : U \to \mathbf{R}^n$.

Podíváme se na diferenciál složeného zobrazení. V následující větě budeme skládání funkcí a zobrazení zapisovat v pořadí zprava doleva podle pořadí aplikace: $(g \circ f)(x) = g(f(x))$.

Věta 2.8. *Nechť*

$$f : U \to V, \ g : V \to \mathbf{R}^k$$

jsou dvě zobrazení, kde $U \subset \mathbf{R}^m$ a $V \subset \mathbf{R}^n$ jsou otevřené množiny. Je-li zobrazení f diferencovatelné v bodě a z U a g je diferencovatelné v bodě $b = f(a)$ z V, je složené zobrazení

$$g \circ f = g(f) : U \to \mathbf{R}^k$$

diferencovatelné v bodě a a jeho diferenciál se rovná složenině diferenciálů zobrazení f a g:

$$\mathrm{D}(g \circ f)(a) = \mathrm{D} g(b) \circ \mathrm{D} f(a).$$

Než se pustíme do důkazu věty, připomeneme význam symbolů $o(h)$ a $O(h)$ a explicitně uvedeme jejich jednoduché vlastnosti, které v důkazu využijeme.

Pro zobrazení $z: U \to \mathbf{R}^n$ definované v okolí počátku $U \subset \mathbf{R}^m$ budeme psát stručně $z(x) = o(x)$ místo $\|z(x)\| = o(\|x\|)$ a $z(x) = O(x)$ místo $\|z(x)\| = O(\|x\|)$, bereme vždy $x \to \overline{0}$. Značení $z(x) = o(x)$ je zkratka pro

$$\forall \varepsilon > 0 \ \exists \delta > 0: \ \|x\| < \delta \Rightarrow \|z(x)\| < \varepsilon \|x\|$$

a $z(x) = O(x)$ pro

$$\exists c > 0 \ \exists \delta > 0: \ \|x\| < \delta \Rightarrow \|z(x)\| < c\|x\|.$$

Lemma. *Nechť $z_1, z_2: U \to \mathbf{R}^n$, kde $U \subset \mathbf{R}^m$ je okolí počátku, jsou dvě zobrazení. Nechť $u: U \to V$ a $v: V \to \mathbf{R}^k$ jsou dvě zobrazení, přičemž $U \subset \mathbf{R}^m$ a $V \subset \mathbf{R}^n$ jsou okolí počátku. V následujících tvrzeních $x \to \overline{0}$.*

1. *Když je z_1 lineární, potom $z_1(x) = O(x)$.*
2. *Když je $z_1(x) = o(x)$ a $z_2(x) = o(x)$, potom $z_1(x) + z_2(x) = o(x)$.*
3. *Když je $z_1(x) = o(x)$ a $z_2(x) = O(x)$, potom $z_1(x) + z_2(x) = O(x)$.*
4. *Pokud $u(x) = o(x)$ a $v = O(x)$, pak $v(u(x)) = o(x)$.*
5. *Pokud $u(x) = O(x)$ a $v(x) = o(x)$, pak $v(u(x)) = o(x)$.*

Důkaz. Úlohy 3 a 4. □

Důkaz věty 2.8. V okolí počátků souřadnic máme aproximace

$$g(b+h) = g(b) + \mathrm{D}g(b)(h) + \gamma(h) \quad \text{a} \quad f(a+h) = f(a) + \mathrm{D}f(a)(h) + \beta(h),$$

kde $\gamma(h)$ a $\beta(h)$ jsou $o(h)$. Rozdíl $f(a+h) - f(a)$ si označíme jako $\Delta(h)$. Pak $f(a+h) = f(a) + \Delta(h) = b + \Delta(h)$ a $\Delta(h) = \mathrm{D}f(a)(h) + \beta(h)$. Takže

$$\begin{aligned}
(g \circ f)(a+h) - (g \circ f)(a) &= g(f(a+h)) - g(f(a)) \\
&= g(b + \Delta(h)) - g(b) \\
&= \mathrm{D}g(b)(\Delta(h)) + \gamma(\Delta(h)) \\
&= \mathrm{D}g(b)(\mathrm{D}f(a)(h)) + \mathrm{D}g(b)(\beta(h)) + \gamma(\Delta(h)) \\
&= (\mathrm{D}g(b) \circ \mathrm{D}f(a))(h) + \alpha(h),
\end{aligned}$$

kde

$$\alpha(h) = \mathrm{D}g(b)(\beta(h)) + \gamma(\Delta(h)) = \mathrm{D}g(b)(\beta(h)) + \gamma(\mathrm{D}f(a)(h) + \beta(h)).$$

První sčítanec definující $\alpha(h)$ je $o(h)$ podle částí 1 a 4 lemmatu a druhý je také $o(h)$ podle částí 1, 3 a 5. Celkem $\alpha(h) = o(h)$ podle části 2. Vidíme, že $g \circ f$ má v a diferenciál rovný lineárnímu zobrazení $\mathrm{D}g(b) \circ \mathrm{D}f(a)$. $\square$

Z lineární algebry víme, že matice lineárního zobrazení $g \circ f$ složeného z lineárních zobrazení f a g se dostane jako součin matice zobrazení g a matice zobrazení f (v tomto pořadí). Jacobiho matice zobrazení f v bodě a je matice lineárního zobrazení $\mathrm{D}f(a)$ vzhledem ke kanonické bázi a její prvky jsou hodnoty parciálních derivací souřadnicových funkcí v bodě a. Pomocí matic tak větu 2.8 vyjádříme následovně.

Důsledek. *Za situace popsané v předchozí větě je Jacobiho matice složeného zobrazení $h = g \circ f$ v bodě a rovna součinu Jacobiho matice zobrazení g v bodě $b = f(a)$ a Jacobiho matice zobrazení f v bodě a:*

$$\begin{aligned}\left(\frac{\partial h_i}{\partial x_j}(a)\right)_{i,j=1}^{k,m} &= \left(\frac{\partial g_i}{\partial x_j}(b)\right)_{i,j=1}^{k,n} \left(\frac{\partial f_i}{\partial x_j}(a)\right)_{i,j=1}^{n,m} \\ &= \left(\sum_{r=1}^{n} \frac{\partial g_i}{\partial x_r}(b) \cdot \frac{\partial f_r}{\partial x_j}(a)\right)_{i,j=1}^{k,m}.\end{aligned}$$

Speciálně pro $k = 1$, kdy funkce h o m proměnných je složeninou

$$h = g(f_1, f_2, \ldots, f_n)$$

funkce g o n proměnných a n funkcí $f_i = f_i(x_1, x_2, \ldots, x_m)$, dostáváme *řetízkové pravidlo* pro parciální derivaci složené funkce:

$$\begin{aligned}\frac{\partial h}{\partial x_i}(a) &= \sum_{j=1}^{n} \frac{\partial g}{\partial x_j}(f(a)) \cdot \frac{\partial f_j}{\partial x_i}(a) \\ &= \langle \nabla g(f(a)), \partial_i f(a) \rangle,\end{aligned}$$

kde $i = 1, 2, \ldots, m$, $f = (f_1, f_2, \ldots, f_n)$ a $\partial_i f = (\partial_i f_1, \partial_i f_2, \ldots, \partial_i f_n)$.

Úlohy

1. Zobecněte důkaz Tvrzení 2.4 na více než dvě proměnné.
2. Rozmyslete si důkaz Tvrzení 2.7.
3. Dokažte části 1–3 lemmatu.
4. Dokažte části 4 a 5 lemmatu.
5. Lemma můžeme zobecnit na zobrazení mezi normovanými prostory (které jako vektorové prostory nemusejí už mít konečnou dimenzi). Pak ale jedna z částí 1–5 obecně přestane platit. Která?

7. přednáška 12. listopadu 2007

Geometrie diferenciálu a parciálních derivací. Nechť $U \subset \mathbf{R}^m$ je okolí bodu a a $f:\ U \to \mathbf{R}$ je funkce o m proměnných. Její okamžitý růst v bodu a ve směru v, kde v je jednotkový vektor (tj. $\|v\| = 1$) z $\mathbf{R}^m$, je dán směrovou derivací

$$\mathrm{D}_v f(a) = \lim_{t \to 0} \frac{f(a + tv) - f(a)}{t}.$$

V jakém směru roste funkce nejrychleji? Když je f v a diferencovatelná, pak podle tvrzení 2.3 a Cauchyovy–Schwarzovy nerovnosti (věta 2.1)

$$|\mathrm{D}_v f(a)| = |\mathrm{D}f(a)(v)| = |\langle \nabla f(a), v\rangle| \le \|\nabla f(a)\| \cdot \|v\| = \|\nabla f(a)\|$$

a rovnost se nabývá, právě když v je skalárním násobkem $\nabla f(a)$, to jest právě pro dva vektory

$$v^+ = \frac{\nabla f(a)}{\|\nabla f(a)\|} \ \text{ a } \ v^- = -\frac{\nabla f(a)}{\|\nabla f(a)\|}.$$

Ve směru v^+ svého normovaného gradientu tedy f roste nejrychleji a v opačném směru v^- stejnou měrou nejrychleji klesá:

$$\mathrm{D}_{v^+} f(a) = \langle \nabla f(a), v^+\rangle = \|\nabla f(a)\| \ \text{ a } \ \mathrm{D}_{v^-} f(a) = \langle \nabla f(a), v^-\rangle = -\|\nabla f(a)\|.$$

(Přesně řečeno, tohle je pravda, pokud je gradient $\nabla f(a)$ nenulový vektor. Je-li to nulový vektor, pak má f ve všech směrech nulový růst.)

Zobecníme pojem tečny ke grafu funkce jedné proměnné na (nad)rovinu tečnou ke grafu funkce více proměnných. Pro jednoduchost značení se omezíme na případ tečné roviny a dvou proměnných; obecná tečná nadrovina ke grafu funkce m proměnných se zavádí analogicky.

Nechť $(x_0, y_0) \in D \subset \mathbf{R}^2$, kde D je otevřená množina v rovině, a $f:\ D \to \mathbf{R}$ je funkce. Její graf

$$P = \{(x, y, z) \in \mathbf{R}^3 \mid (x, y) \in D, z = f(x, y)\}$$

je plocha v třírozměrném euklidovském prostoru. Na P leží bod (x_0, y_0, z_0), kde $z_0 = f(x_0, y_0)$. Předpokládejme, že funkce f je v bodě (x_0, y_0) diferencovatelná. Potom mezi všemi rovinami $z = L(x, y)$ (L je afinní funkce dvou proměnných), které obsahují bod (x_0, y_0, z_0), je pouze jediná splňující pro $(x, y) \to (x_0, y_0)$ aproximaci

$$f(x, y) = L(x, y) + o(\sqrt{(x - x_0)^2 + (y - y_0)^2}),$$

totiž rovina

$$T(x, y) = z_0 + \frac{\partial f}{\partial x}(x_0, y_0) \cdot (x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0) \cdot (y - y_0).$$

To plyne hned z existence diferenciálu a jeho jednoznačnosti, protože zřejmě $\mathrm{D}f(x_0, y_0)(x - x_0, y - y_0) = T(x, y) - z_0$. Graf funkce $T(x, y)$,

$$T = \{(x, y, z) \in \mathbf{R}^3 \mid (x, y) \in D, z = T(x, y)\}$$

se nazývá *tečnou rovinou ke grafu funkce* f *v bodě* (x_0, y_0, z_0).

Rovnici tečné roviny $z = T(x, y)$ přepíšeme ve tvaru

$$\begin{aligned} \frac{\partial f}{\partial x}(x_0, y_0) \cdot (x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0) \cdot (y - y_0) \ - (z - z_0) &= 0 \\ \langle V, (x - x_0, y - y_0, z - z_0) \rangle &= 0, \end{aligned}$$

kde V z $\mathbf{R}^3$ je vektor

$$V = \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0), -1 \right).$$

Označíme-li $X = (x, y, z)$ a $X_0 = (x_0, y_0, z_0)$, můžeme tečnou rovinu T definovat jako

$$T = \{X \in \mathbf{R}^3 \mid \langle V, X - X_0 \rangle = 0\}.$$

Je tedy tvořena právě těmi body, jejichž směrové vektory k bodu X_0 jsou kolmé na V. Vektor V se nazývá *normálovým vektorem ke grafu funkce* f *v bodě* X_0.

Parciální derivace vyšších řádů. Pokud má funkce $f : U \to \mathbf{R}$ definovaná na otevřené množině $U \subset \mathbf{R}^m$ v každém jejím bodě parciální derivaci $F = \partial_i f$ a ta má má v bodě $a \in U$ parciální derivaci $\partial_j F(a) = \partial_j \partial_i f(a)$, řekneme, že f má v bodě a *parciální derivaci druhého řádu podle proměnných* x_i *a* x_j a její hodnotu značíme

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

Induktivně definujeme parciální derivace vyšších řádů: má-li f v každém bodě $x \in U$ parciální derivaci

$$F = \frac{\partial^{k-1} f}{\partial x_{i_{k-1}} \partial x_{i_{k-2}} \dots \partial x_{i_1}}(x)$$

a ta má v bodě $a \in U$ parciální derivaci $\partial_j F(a)$, řekneme, že f má v bodě a *parciální derivaci* k*-tého řádu podle proměnných* $x_{i_1}, \dots, x_{i_{k-1}}, x_j$ a její hodnotu značíme

$$\frac{\partial^k f}{\partial x_j \partial x_{i_{k-1}} \dots \partial x_{i_1}}(a).$$

Na pořadí proměnných pří parciálním derivování obecně záleží, jak ukazuje příklad funkce

$$f(x, y) = \begin{cases} \frac{xy(x^2 - y^2)}{x^2 + y^2} & \text{pro } x^2 + y^2 \neq 0 \\ 0 & \text{pro } x^2 + y^2 = 0, \end{cases}$$

která má v počátku obě smíšené parciální derivace druhého řádu, ale s různými hodnotami:

$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) = 1 \neq -1 = \frac{\partial^2 f}{\partial y \partial x}(0, 0)$$

(úloha 1). Nicméně při spojitých parciálních derivacích na pořadí proměnných nezáleží.

Tvrzení 2.9. *Nechť funkce $f : U \to \mathbf{R}$ má na okolí $U \subset \mathbf{R}^m$ bodu a parciální derivace druhého řádu $\partial_j\partial_i f$ a $\partial_i\partial_j f$ a ty jsou v a spojité. Potom*

$$\partial_j\partial_i f(a) = \partial_i\partial_j f(a).$$

Důkaz. Pro jednoduchost buď $m = 2$ a $a = (0,0)$. Díky spojitosti obou parciálních derivací v počátku stačí nalézt pro každé $h > 0$ ve čtverci $[0,h]^2$ dva body σ a τ, v nichž $\partial_x\partial_y f(\sigma) = \partial_y\partial_x f(\tau)$.

Vrcholy čtverce označíme $a = (0,0)$, $b = (0,h)$, $c = (h,0)$, $d = (h,h)$ a uvážíme číslo $f(d) - f(b) - f(c) + f(a)$. Lze ho dvěma způsoby napsat jako rozdíl rozdílů:

$$\begin{aligned} f(d) - f(b) - f(c) + f(a) &= (f(d) - f(b)) - (f(c) - f(a)) = \psi(h) - \psi(0) \\ &= (f(d) - f(c)) - (f(b) - f(a)) = \phi(h) - \phi(0), \end{aligned}$$

kde

$$\psi(t) = f(h,t) - f(0,t) \ \text{ a } \ \phi(t) = f(t,h) - f(t,0).$$

Máme $\psi'(t) = \partial_y f(h,t) - \partial_y f(0,t)$ a $\phi'(t) = \partial_x f(t,h) - \partial_x f(t,0)$. Lagrangeova věta o střední hodnotě dává vyjádření

$$\begin{aligned} f(d) - f(b) - f(c) + f(a) &= \psi'(t_0)h = (\partial_y f(h,t_0) - \partial_y f(0,t_0))h \\ &= \phi'(s_0)h = (\partial_x f(s_0,h) - \partial_x f(s_0,0))h, \end{aligned}$$

kde $0 < s_0, t_0 < h$. Použijeme ji ještě jednou na rozdíly parciálních derivací f a máme

$$f(d) - f(b) - f(c) + f(a) = \partial_x\partial_y f(s_1,t_0)h^2 = \partial_y\partial_x f(s_0,t_1)h^2,\ s_1,t_1 \in (0,h).$$

Body $\sigma = (s_1,t_0)$ a $\tau = (s_0,t_1)$ leží ve čtverci $[0,h]^2$ a $\partial_x\partial_y f(\sigma) = \partial_y\partial_x f(\tau)$. $\square$

Rovnost hodnot obou derivací lze dokázat i za slabšího předpokladu: existuje-li $\partial_x\partial_y f$ v okolí bodu a a je v něm spojitá, potom existuje i $\partial_y\partial_x f(a)$ a $\partial_y\partial_x f(a) = \partial_x\partial_y f(a)$.

Pro otevřenou množinu $U \subset \mathbf{R}^m$ označíme symbolem $\mathcal{C}^k(U)$ množinu funkcí $f : U \to \mathbf{R}$, jejichž parciální derivace až do řádu k včetně jsou na U definované a spojité.

Důsledek. *Pro každou funkci $f = f(x_1, x_2, \ldots, x_m)$ z $\mathcal{C}^k(U)$ hodnoty jejích parciálních derivací až do řádu k nezávisí na pořadí proměnných—pro $l \le k$ a $a \in U$ platí*

$$\frac{\partial^l f}{\partial x_{i_l}\partial x_{i_{l-1}}\ldots\partial x_{i_1}}(a) = \frac{\partial^l f}{\partial x_{j_l}\partial x_{j_{l-1}}\ldots\partial x_{j_1}}(a),$$

jakmile je posloupnost $(i_1, \ldots, i_l)$ permutací posloupnosti $(j_1, \ldots, j_l)$.

Důkaz. Když je posloupnost $v = (j_1, \ldots, j_l)$ pouze permutací posloupnosti $u = (i_1, \ldots, i_l)$, dokážeme u transformovat ve v prohazováním dvojic členů v u, dokonce vystačíme s prohazováním sousedních členů: v u nalezneme člen j_1 a necháme ho „propadnout“ až dolů, pak necháme propadnout na správné místo j_2 atd. Rovnost hodnot parciálních derivací tak plyne z tvrzení 2.9. □

V případě spojitých parciálních derivací tak záleží jen na multimnožině proměnných, podle kterých se derivuje, ale ne na jejich pořadí. Místo $\partial_x \partial_x$ píšeme stručně ∂x^2 apod. Například, pro f z $C^5(U)$ máme

$$\frac{\partial^5 f}{\partial y\, \partial x\, \partial y\, \partial y\, \partial z} = \frac{\partial^5 f}{\partial y^2\, \partial x\, \partial z\, \partial y} = \frac{\partial^5 f}{\partial x\, \partial z\, \partial y^3} = \cdots$$

Důležitým nástrojem při studiu vlastností funkcí je Taylorův rozvoj, jehož verzi pro více proměnných nyní odvodíme. Jak rozumět použitému symbolickému zápisu diferenciálního operátoru vysvětlíme na příkladu, v němž $f = f(x, y, z)$ je funkce a $a \in \mathbf{R}^3$, $\alpha, \beta \in \mathbf{R}$ jsou konstanty. Zápisem

$$(\alpha\partial_y + \beta\partial_z)^3 f(a)$$

se rozumí

$$\begin{aligned} & (\alpha^3(\partial_y)^3 + 3\alpha^2\beta(\partial_y)^2\partial_z + 3\alpha\beta^2\partial_y(\partial_z)^2 + \beta^3(\partial_z)^3)f(a) \\ = \; & \alpha^3\frac{\partial^3 f}{\partial y^3}(a) + 3\alpha^2\beta\frac{\partial^3 f}{\partial y^2\partial z}(a) + 3\alpha\beta^2\frac{\partial^3 f}{\partial y\partial z^2}(a) + \beta^3\frac{\partial^3 f}{\partial z^3}(a). \end{aligned}$$

Tvrzení 2.10. *Nechť $U \subset \mathbf{R}^m$ je otevřená množina, $a \in U$ je bod a $f : U \to \mathbf{R}$ je funkce z $C^n(U)$. Potom v okolí bodu a máme Taylorův rozvoj*

$$\begin{aligned} f(a+h) &= \sum_{i=0}^{n} \frac{1}{i!}(h_1\partial_1 + h_2\partial_2 + \cdots + h_m\partial_m)^i f(a) + o(\|h\|^n) \\ &= \sum \frac{1}{i_1! \dots i_m!} \cdot \frac{\partial^{i_1+\cdots+i_m} f}{\partial x_1^{i_1} \dots \partial x_m^{i_m}}(a) \cdot h_1^{i_1} \dots h_m^{i_m} + o(\|h\|^n) \\ &= f(a) + \sum_{i=1}^{m} \partial_{x_i} f(a) h_i + \sum_{i<j} \partial_{x_i}\partial_{x_j} f(a) h_i h_j + \frac{1}{2}\sum_{i=1}^{m} \partial_{x_i}^2 f(a) h_i^2 + \\ &\quad + \cdots + o(\|h\|^n). \end{aligned}$$

V první sumě mocninu chápeme symbolicky (ve smyslu operátorového počtu) a ve druhé sumě sčítáme přes všechny m-tice nezáporných celých čísel $i_1, i_2, \dots, i_m$, jejichž součet je nanejvýš n.

Důkaz. Vezmeme Taylorův rozvoj až do řádu n pomocné funkce jedné proměnné $F(t) = f(a + th)$, kde $t \in [0, 1]$. Opakovaným použitím řetízkového pravidla ($F = f \circ l$, kde l je lineární zobrazení, přímka $l(t) = a + th$) pro $k \le n$ dostáváme

$$F^{(k)}(t) = \sum_{i_1, i_2, \dots, i_k} \frac{\partial^k f}{\partial x_{i_1}\, \partial x_{i_2} \dots \partial x_{i_k}}(a + th) \cdot h_{i_1} h_{i_2} \dots h_{i_k},$$

kde $i_1, \ldots, i_k$ probíhají nezávisle na sobě všechny indexy $1, 2, \ldots, m$. Dosazením do Taylorova rozvoje funkce F se zbytkem v Lagrangeově tvaru

$$f(a+h) = F(1) = \sum_{i=0}^{n-1} \frac{1}{i!} F^{(i)}(0) + \frac{F^{(n)}(\theta)}{n!},\ 0 < \theta < 1,$$

dostáváme, s využitím kompaktního symbolického zápisu parciálních derivací, první formuli pro $f(a+h)$. Druhá formule vyplývá z první pomocí multinomické věty:

$$(h_1\partial_1 + h_2\partial_2 + \cdots + h_m\partial_m)^i = \sum_{i_1, i_2, \ldots, i_m} \binom{i}{i_1, i_2, \ldots, i_m} \prod_{j=1}^{m} (h_j\partial_j)^{i_j},$$

kde $i_1, i_2, \ldots, i_m$ probíhají nezáporná celá čísla se součtem i a

$$\binom{i}{i_1, i_2, \ldots, i_m} = \frac{i!}{i_1! \cdot i_2! \cdot \ldots \cdot i_m!}$$

je multinomický koeficient. $\square$

Sčítance odpovídající $i = 0$ a 1 jsou $f(a)$ a $\mathrm{D}f(a)(h)$. Taylorova formule zobecňuje lokální aproximaci pomocí diferenciálu, kterou dostáváme pro $n = 1$.

Extrémy funkcí více proměnných. Symetrická (tj. $a_{i,j} = a_{j,i}$) reálná $n \times n$ matice $A \in \mathbf{R}^{n \times n}$ definuje kvadratickou formu

$$P(x_1, x_2, \ldots, x_n) = xAx^T = \sum_{i,j=1}^{n} a_{i,j} x_i x_j : \mathbf{R}^n \to \mathbf{R}$$

(x je řádkový vektor $(x_1, x_2, \ldots, x_n)$). Připomeňme si, že A se nazývá

- *pozitivně (negativně) definitní*, když $P(x) > 0$ ($P(x) < 0$) pro všechny $x \in \mathbf{R}^n \backslash \{\overline{0}\}$;
- *pozitivně (negativně) semidefinitní*, když $P(x) \geq 0$ ($P(x) \leq 0$) pro všechny $x \in \mathbf{R}^n$;
- *indefinitní*, není-li ani pozitivně ani negativně semidefinitní, tj. $P(x) > 0$ a $P(y) < 0$ pro nějaké dva vektory $x, y \in \mathbf{R}^n$.

Hessova matice funkce f v bodě a, kde $f : U \to \mathbf{R}$ je definovaná na okolí $U \subset \mathbf{R}^m$ bodu a a má na U všechny derivace druhého řádu, je matice zaznamenávající hodnoty těchto derivací:

$$H_f(a) := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)_{i,j=1}^{m}.$$

Podle tvrzení 2.9 je Hessova matice funkcí z $\mathcal{C}^2(U)$ symetrická.

Odvodíme kritérium existence lokálních extrémů funkcí m proměnných, které zobecňuje výsledek pro funkce jedné proměnné. Roli hodnoty druhé derivace přebírá Hessova matice. Připomeňme si, že funkce $f : U \to \mathbf{R}$, kde $U \subset \mathbf{R}^m$ je otevřená množina, má v bodě $a \in U$ ostré lokální minimum, pokud existuje $\delta > 0$ takové, že $0 < \|x - a\| < \delta$ implikuje $f(x) > f(a)$. (Neostré) lokální minimum znamená, že $\|x - a\| < \delta$ implikuje $f(x) \geq f(a)$. Podobně pro ostré a neostré lokální maximum. Funkce f nemá v a ani neostrý lokální extrém, nemá-li v tomto bodě ani lokální neostré minimum ani lokální neostré maximum, to jest pro každé $\delta > 0$ existují body x, y takové, že $\|x - a\|, \|y - a\| < \delta$ a $f(x) > f(a)$, $f(y) < f(a)$.

Věta 2.11. *Nechť $f \in C^2(U)$, kde $U \subset \mathbf{R}^m$ je otevřená množina, a $a \in U$ je bod.*

- *Pokud $\nabla f(a) \neq \overline{0}$, nemá f v a ani neostrý lokální extrém.*
- *Pokud $\nabla f(a) = \overline{0}$ a $H_f(a)$ je pozitivně (negativně) definitní, potom má f v a ostré lokální minimum (maximum).*
- *Pokud $\nabla f(a) = \overline{0}$ a $H_f(a)$ je indefinitní, nemá f v a ani neostrý lokální extrém.*

Úlohy

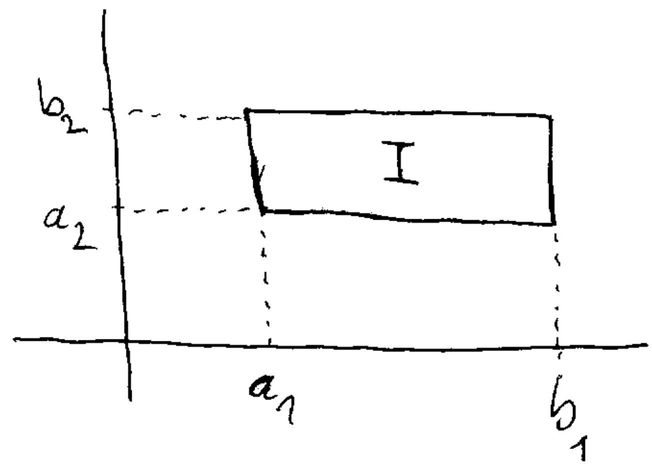
1. Ověřte, že uvedená funkce $f(x, y)$ má v počátku různé smíšené parciální derivace druhého řádu.

Vícerozměrný Riemannův integrál, Fubiniova věta

(n-rozměrný) **box** $I \subset \mathbb{R}^n$ je kart. součin

$$I = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n],$$ kde

$-\infty < a_i < b_i < +\infty$. Např. v $\mathbb{R}^2$:

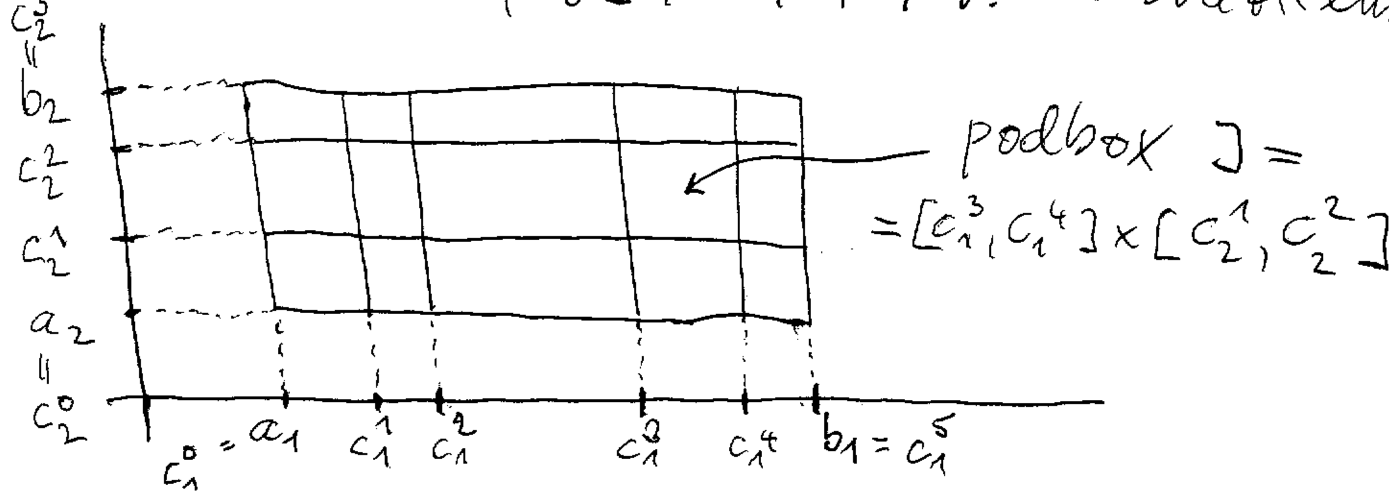


Objem boxu je $|I| = \prod_{i=1}^{n} (b_i - a_i)$

Dělení boxu na podboxy: D je dělení I, když

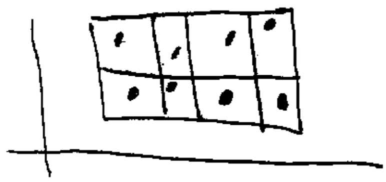
$$D = \{ \underbrace{[c_1^{a_1}, c_1^{a_1+1}] \times [c_2^{a_2}, c_2^{a_2+1}] \times \cdots \times [c_n^{a_n}, c_n^{a_n+1}]}_{\text{podbox}} \mid 0 \le a_i < k_i,\ 1 \le i \le n \},$$ kde

$a_i = c_i^0 < c_i^1 < \cdots < c_i^{k_i-1} < c_i^{k_i} = b_i$ jsou dělení intervalů $[a_i, b_i]$, $i = 1, 2, \ldots, n$. Obrázkem:



<u>norma dělení D</u> je $\lambda(D) := \max\limits_{\substack{1 \le i \le n \\ 0 \le j < k_i}} (c_i^{j+1} - c_i^j)$

<u>Dělení D boxu I s body ξ</u> je dvojice (D, ξ), kde $\xi : D \to \mathbb{R}^n$, $\xi(J) \in J$ $\forall$ podbox J:



Riemannova definice vícerozměrného integrálu

$I \subset \mathbb{R}^n$ box, (D, ξ) jeho dělení s body, $f : I \to \mathbb{R}$ <u>omezená</u> funkce, R.-ova suma je

$$S(f, D, \xi) := \sum_{J \in D} |J| \cdot f(\xi(J)).$$

<u>Integrál f přes I je limita</u>

$$\int_I f := \lim_{\substack{(D,\xi) \\ \lambda(D) \to 0}} S(f, D, \xi),$$ když existuje,

$$\Big(\int_I f \in \mathbb{R}, \text{ ne } \pm\infty\Big)$$

to jest

$\forall \varepsilon > 0\ \exists \delta > 0\ \forall (D, \xi)$ dělení I s body, $\lambda(D) < \delta$:

$$: \Big| \int_I f - S(f, D, \xi) \Big| < \varepsilon.$$

Darbouxova definice: D dělení I, $f: I \to \mathbb{R}$ omezená, $m(J) := \inf_{x \in J} f(x)$, $M(J) = \sup_{x \in J} f(x)$

$$s(f,D) := \sum_{J \in D} |J| \cdot m(J) \text{ – dolní součet},$$

$$S(f,D) := \sum_{J \in D} |J| \cdot M(J) \text{ – horní součet}.$$

$$\underline{\int_I} f = \sup\{s(f,D) \mid D \text{ dělení } I\} \text{ – dolní } \textstyle\int,$$

$$\overline{\int_I} f = \inf\{S(f,D) \mid \text{–"–}\} \text{ – horní } \textstyle\int.$$

Opět platí: $-\infty < s(f,D) \leq \underline{\int_I} f \leq \overline{\int_I} f \leq S(f,D) < +\infty$, pro $\forall D$.

$$\boxed{\int_I f := \underline{\int_I} f = \overline{\int_I} f}\text{, když se dolní a horní } \textstyle\int \text{ rovnají.}$$

Platí: f má $\int$ podle R.-ovy def. $\Longleftrightarrow$ f má $\int$ podle D.-ovy def. a obě hodnoty se rovnají.

$\mathcal{R}(I) = \{f \mid f \text{ má R.-ův } \int \text{ přes } I\}$ – množina funkcí R.-ovsky integrovatelných přes box I.

$\forall \varepsilon > 0$

$E \subset \mathbb{R}^n$ má (Lebesgueovu) míru 0: ex. taková posl. boxů $I_1, I_2, \dots$ v $\mathbb{R}^n$, že $\sum_{i=1}^{\infty} |I_i| < \varepsilon$ a $E \subset \bigcup_{i=1}^{\infty} I_i$.

Věta (Lebesgue) $I \subset \mathbb{R}^n$ box, $f: I \to \mathbb{R}$, pak $f \in \mathcal{R}(I) \iff f$ je na I omezená a mn. bodů nespojitosti f má míru 0.

Např. každá omezená f nespojitá jen v nejvýše spočetně mnoha bodech má R.-ův $\int$ přes I.

Důsledek $f: I \to \mathbb{R}$, $f \in \mathcal{R}(I)$, $f \geq 0$ na I, $\int_I f = 0 \Rightarrow f = 0$ na I až na mn. bodů míry 0.

Fubiniova věta, přesněji věta Fubiniova typu, umožňuje převést výpočet vícerozměrného $\int$ na posloupnost obyčejných (jednorozměrných) $\int$-ů.

$X \subset \mathbb{R}^m$, $Y \subset \mathbb{R}^n$, $Z = X \times Y \subset \mathbb{R}^{m+n}$ buďte m-rozm., n-rozm. a $(m+n)$-rozm. boxy.

Věta (Fubini) Nechť $f: \underset{X\times Y}{\underset{\parallel}{Z}} \to \mathbb{R}$, $f \in \mathcal{R}(Z)$.
Všechny tři integrály

$$\int_Z f, \quad \int_X dx \int_Y f(x,y)\,dy \quad a \quad \int_Y dy \int_X f(x,y)\,dx$$

existují a rovnají se. Vysvětlení značení!

Integrál $\int_Z f$ existuje podle předpokladu o f.
Definujeme funkci $F: X \to \mathbb{R}$, $F(x) := \int_Y f(x,y)\,dy$;
pokud pro nějaké $x_0 \in X$ integrál neexistuje, definujeme $F(x_0)$ jako libovolnou hodnotu z intervalu $\left[\underline{\int_Y} f(x_0,y)\,dy, \overline{\int_Y} f(x_0,y)\,dy\right]$. Podobně $G: Y \to \mathbb{R}$, $G(y) := \int_X f(x,y)\,dx$. Věta říká, že

$F \in \mathcal{R}(X)$, $G \in \mathcal{R}(Y)$ a $\int_Z f = \int_X F = \int_Y G$.

V důkazu též ukážeme fakticky, že mn. bodů $x_0 \in X$, pro něž $f(x_0,y) \notin \mathcal{R}(Y)$, má míru 0 a podobně v y-ové souřadnici.

<u>Důkaz.</u> Dokážeme, že $F \in \mathcal{R}(X)$ a $\int_Z f = \int_X F$, pro G je důkaz podobný. Každé dělení D boxu Z je „součinem" $D_1 \times D_2$ dělení D_1 boxu X a dělení D_2 boxu Y, tj. $\forall J \in D$ je součinem $J = J_1 \times J_2$ pro $J_1 \in D_1$, $J_2 \in D_2$. Buď dáno $\varepsilon > 0$. Pak ex. dělení D boxu Z, že $s(f, D) > \int_Z f - \varepsilon$. Vezmeme dělení D_1 a D_2, že D „=" $D_1 \times D_2$. Pak, podle vlastnosti infima,

$$s(f, D) = \sum_{J \in D} |J| \cdot \inf_J (f) = \sum_{J \in D} \underbrace{|J_1 \times J_2|}_{=|J_1|\cdot|J_2|} \cdot \inf_{\substack{x \in J_1 \\ y \in J_2}} f(x,y)$$

proč platí? – zamyslete se!

$$\leq \sum_{J_1 \in D_1} |J_1| \cdot \inf_{x \in J_1} \Bigg(\underbrace{\sum_{J_2 \in D_2} |J_2| \inf_{y \in J_2} f(x,y)}_{= s(f(x,\cdot), D_2) \leq \underline{\int_Y} f(x,y)\,dy \leq F(x)} \Bigg) \leq$$

$$\leq \sum_{J_1 \in D_1} |J_1| \cdot \inf_{x \in J_1} F(x) = s(F, D_1).$$ Takže

$s(F, D_1) > \int_Z f - \varepsilon$. Analogicky se dokáže, že pro dané $\varepsilon > 0$ ex. dělení D_1' boxu X_1 že $S(F, D_1') < \int_Z f + \varepsilon$. Pro $\varepsilon \to 0$ to podle D.-ovy def. $\int$ znamená, že $F \in \mathcal{R}(X)$ a $\int_X F = \int_Z f$. $\boxtimes$

$\boxed{\text{Příklad}}$ $f(x,y,z) = z \cdot \sin(x+y)$, $I \subset \mathbb{R}^3$ dán ~~jako~~ vztahy $0 \le x \le \pi$, $|y| \le \frac{\pi}{2}$, $0 \le z \le 1$. Pak

$$\iiint_I f(x,y,z)\,dx\,dy\,dz = \int_0^1 dz \int_{-\pi/2}^{\pi/2} dy \int_0^{\pi} z \cdot \sin(x+y)\,dx =$$

$$= \int_0^1 dz \int_{-\pi/2}^{\pi/2} \left(-z \cdot \cos(x+y) \Big|_{x=0}^{\pi} \right) dy = \int_0^1 dz \int_{-\pi/2}^{\pi/2} 2z \cdot \cos y\,dy =$$

$$= \int_0^1 \left(2z \sin y \Big|_{y=-\pi/2}^{\pi/2} \right) dz = \int_0^1 4z\,dz = 2.$$

<u>Integrál přes množinu $E \subset \mathbb{R}^n$</u>. Pojem integrálu rozšíříme z boxů $I \subset \mathbb{R}^n$ na obecnější množiny. Zavedeme i objem množiny.

<u>Definice</u> Množina $E \subset \mathbb{R}^n$ je <u>přípustná</u>, když je <u>omezená</u> a její <u>hranice ∂E má míru nula.</u>

$(\partial E = \{x \in \mathbb{R}^n \mid \forall r>0 : B(x,r)\cap E \neq \emptyset,\ B(x,r)\cap(\mathbb{R}^n\setminus E) \neq \emptyset\})$

<u>Příklady</u> ,  jsou přípustné, $\mathbb{Q}\cap([0,1]^2)$

není přípustná.

<u>Definice</u> <u>Objem</u> omezené množiny $E \subset \mathbb{R}^n$ je integrál $\mathrm{vol}(E) := \int_I \chi_E$, kde I je box obsahující E

a $\chi_E(x) = \begin{cases} 1 \ldots & x\in E \\ 0 \ldots & x \in I\setminus E \end{cases}$, když existuje.

Dá se dokázat:

$\boxed{\text{Tvrzení}}$ $E\subset\mathbb{R}^n$ má objem (podle definice)

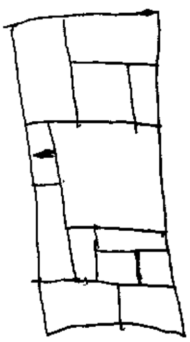
$\iff$ E je přípustná.

Nechť $E\subset\mathbb{R}^n$ je omezená. <u>Integrál funkce $f: E\to\mathbb{R}$ přes E</u> definujeme jako

$\int_E f := \int_I \bar{f}$, kde I je box obsahující E a $\bar{f}$ je rozšíření f: $\bar{f}(x) = \begin{cases} f(x) \ldots & x\in E \\ 0 \ldots & x\in I\setminus E. \end{cases}$

Úlohy: **1.** Nechť $I \subset \mathbb{R}^n$ je box a D je jeho dělení. Dokažte, že $|I| = \sum_{J \in D} |J|$.

2*. Nechť $I \subset \mathbb{R}^n$ je box a $I = \bigcup_{i=1}^{k} J_i$ je rozklad I na boxy J_i, které mají disjunktní vnitřky, např. v $\mathbb{R}^2$:



Dokažte, že $|I| = \sum_{i=1}^{k} |J_i|$.

3. Nechť $f: I \to \mathbb{R}$ je neomezená ($I \subset \mathbb{R}^n$ je box). Co se stane v R-ově def. $\int$? Jak vypadají $\underline{\int_I} f$ a $\overline{\int_I} f$?

4. Zdůvodněte, proč ve Fub. větě má množina $\{x_0 \in X \mid \int_Y f(x_0, y)\,dy \text{ neexistuje}\}$ míru 0.

5. Zdůvodněte, proč objem $\mathrm{vol}(E)$ a integrál $\int_E f$ pro množinu $E \subset \mathbb{R}^n$ nezávisejí na volbě boxu $I \subset \mathbb{R}^n$ obsahujícího E (když existují).

* * * * * * * * * * * * *

9. přednáška 26. listopadu 2007

Věta 2.11. *Nechť $f \in C^2(U)$, kde $U \subset \mathbf{R}^m$ je otevřená množina, a $a \in U$ je bod.*

- *Pokud $\nabla f(a) \neq \overline{0}$, nemá f v a ani neostrý lokální extrém.*
- *Pokud $\nabla f(a) = \overline{0}$ a $H_f(a)$ je pozitivně (negativně) definitní, potom má f v a ostré lokální minimum (maximum).*
- *Pokud $\nabla f(a) = \overline{0}$ a $H_f(a)$ je indefinitní, nemá f v a ani neostrý lokální extrém.*

Důkaz. 1. Pokud $\nabla f(a) \neq \overline{0}$, pak např. $\partial_{x_1} f(a) > 0$ (pro $\partial_{x_1} f(a) < 0$ postupujeme obdobně), a $f(a_1+h, a_2, \ldots, a_m) = f(a)+\partial_{x_1} f(a)h+o(h)$. Existuje tedy takové $\delta > 0$, že pro $h \in (-\delta, 0)$ máme $f(a_1+h, a_2, \ldots, a_m)-f(a) < \frac{1}{2}\partial_{x_1} f(a)h < 0$ a pro $h \in (0, \delta)$ máme $f(a_1 + h, a_2, \ldots, a_m) - f(a) > \frac{1}{2}\partial_{x_1} f(a)h > 0$. Funkce f nemá v a ani neostrý lokální extém.

2 a 3. Nyní $\nabla f(a) = \overline{0}$. Kvadratickou formu $xH_f(a)x^T$ označíme jako $P(x)$ a f rozvineme v okolí a do Taylorova rozvoje řádu $n = 2$ (tvrzení 2.10). Sčítanec $f(a)$ odpovídající $i = 0$ převedeme vlevo, sčítanec s $i = 1$ zmizí, protože $\nabla f(a) = \overline{0}$. $P(x)$ je homogenní polynom stupně 2, takže

$$\begin{aligned}
f(a+h) - f(a) &= \sum_{i=1}^{2} \frac{1}{i!}(h_1\partial_1 + h_2\partial_2 + \cdots + h_m\partial_m)^i f(a) + o(\|h\|^2) \\
&= \frac{1}{2}\sum_{i,j=1}^{m} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j + o(\|h\|^2) \\
&= \frac{1}{2}P(h_1, h_2, \ldots, h_m) + o(\|h\|^2) \\
&= \frac{1}{2}\|h\|^2\Big(P(h_1/\|h\|, h_2/\|h\|, \ldots, h_m/\|h\|) + o(1)\Big) \\
&= \frac{1}{2}\|h\|^2(P(e) + o(1)),
\end{aligned}$$

kde vektor $e = e(h) = (h_1/\|h\|, h_2/\|h\|, \ldots, h_m/\|h\|)$ leží na jednotkové sféře $S = \{x \in \mathbf{R}^m \mid \|x\| = 1\}$. S je kompaktní podmnožina $\mathbf{R}^m$ (je uzavřená a omezená) a spojitá funkce $P(x)$ na ní proto nabývá minima a maxima:

$$\mu = P(\alpha) = \min_{\|x\|=1} P(x) \quad \text{a} \quad M = P(\beta) = \max_{\|x\|=1} P(x)$$

pro nějaké vektory α a β z S. Pozitivní (negativní) definitnost $H_f(a)$ je ekvivalentní nerovnostem $0 < \mu \leq M$ ($\mu \leq M < 0$) a indefinitnost je ekvivalentní $\mu < 0 < M$.

Je-li $H_f(a)$ pozitivně definitní, máme $P(e) \geq \mu > 0$ pro každé $e \in S$, a tak existuje $\delta > 0$ takové, že pro každé h splňující $0 < \|h\| < \delta$ platí

$$f(a+h) - f(a) = \frac{1}{2}\|h\|^2(P(e) + o(1)) > \frac{\|h\|^2}{2} \cdot \frac{\mu}{2} > 0$$

—f má v a ostré lokální minimum. Analogicky pro negativně definitní $H_f(a)$ dostáváme ostré lokální maximum. Když je $H_f(a)$ indefinitní, pak existuje $\delta > 0$ takové, že pro každé $t \in (0, \delta)$ máme

$$\begin{aligned} f(a+t\alpha) - f(a) &= \frac{t^2}{2}(P(\alpha) + o(1)) < \frac{t^2}{2} \cdot \frac{\mu}{2} < 0 \\ f(a+t\beta) - f(a) &= \frac{t^2}{2}(P(\beta) + o(1)) > \frac{t^2}{2} \cdot \frac{M}{2} > 0 \end{aligned}$$

—f nemá v a ani neostrý lokální extrém. □

Důležité poznámky. Podle této věty funkce, která má v každém bodu otevřené množiny U gradient, může mít lokální extrém pouze v bodech, v nichž je gradient nulový. Těmto bodům se říká *stacionární body*. Dostaneme je jako řešení rovnice $\nabla f(a) = 0$. Když je matice $H_f(a)$ semidefinitní, neříká věta nic, funkce může mít v a extrém nebo nemusí. Konečně zdůrazněme, že se věta týká otevřených množin U, respektive vnitřních bodů a množiny U. Pokud je bod a v U ale není jejím vnitřním bodem, pak může f mít v a lokální extrém vzhledem k U, i když je $\nabla f(a)$ nenulový. Lokálními extrémy v hraničních bodech množin se budeme zabývat později (v partii o Lagrangeových multiplikátorech).

Příklad. Nalezněte lokální a globální extrémy funkce

$$f: \mathbf{R}^2 \to \mathbf{R},\ f(x,y) = y^2 + y\cos x - \sin x - 2.$$

Definiční obor $\mathbf{R}^2$ je otevřená množina, pro hledání lokálních extrémů můžeme bez problémů použít větu 2.11. Máme

$$\nabla f(x,y) = (\partial_x f, \partial_y f) = (-y\sin x - \cos x, 2y + \cos x)$$

a

$$H_f(x,y) = \begin{pmatrix} \partial^2_{xx} f & \partial^2_{xy} f \\ \partial^2_{yx} f & \partial^2_{yy} f \end{pmatrix} = \begin{pmatrix} -y\cos x + \sin x & -\sin x \\ -\sin x & 2 \end{pmatrix}.$$

Soustava rovnic $\nabla f(x,y) = (0,0)$ se snadno vyřeší a dává stacionární body

$$s_k = (\pi/2 + k\pi, 0),\ k \in \mathbf{Z}.$$

Tedy

$$H_f(s_k) = \begin{pmatrix} (-1)^k & (-1)^{k+1} \\ (-1)^{k+1} & 2 \end{pmatrix}$$

a

$$H_f(s_k) = \begin{pmatrix} -1 & 1 \\ 1 & 2 \end{pmatrix} \text{ pro liché } k \text{ a } H_f(s_k) = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \text{ pro sudé } k.$$

První matice je indefinitní,

$$P(x,y) = -x^2 + 2xy + 2y^2 = -(x-y)^2 + 3y^2,$$

a druhá je pozitivně definitní,

$$P(x,y) = x^2 - 2xy + 2y^2 = (x-y)^2 + y^2.$$

Pro liché k v s_k není lokální extrém a pro sudé k je v s_k ostré lokální minimum, vždy s hodnotou

$$f(s_{2k}) = -3.$$

Jediné lokální extrémy funkce f tedy jsou tato ostrá lokální minima.

Globální maximum neexistuje, protože f je shora neomezená: $f(\pi/2, y) = y^2 - 3$. Jiný důvod je ten, že f nemá žádné lokální maximum (a globální maximum by muselo být i lokálním maximem). Nalezneme globální minimum. Definiční obor $\mathbf{R}^2$ není kompaktní, nelze hned použít větu o extrémech spojitých funkcí na kompaktech. Funkce f je však 2π-periodická v x a pro vyšetření globálních minim stačí uvážit její hodnoty v pásu

$$P = \{(x,y) \mid 0 \le x \le 2\pi, y \in \mathbf{R}\}.$$

Na jeho hranici máme

$$f(0,y) = f(2\pi, y) = y^2 + y - 2 = \left(y + \frac{1}{2}\right)^2 - \frac{9}{4} \ge -\frac{9}{4} > -3.$$

Ještě ale nejsme hotovi. I když hodnoty f na hranici pásu nejsou menší než -3, pás sám je nekompaktní a pro $y \to \pm\infty$ by někde uprostřed něj mohla f klesat k hodnotám menším než -3, třeba do $-\infty$, a globální minimum by nemuselo existovat. Jednoduchý odhad však ukazuje, že se f tak nechová. Pro $|y| \ge 2$ a libovolné $x \in \mathbf{R}$ máme

$$f(x,y) \ge y^2 - |y| - 3 = \left(y \pm \frac{1}{2}\right)^2 - \frac{13}{4} \ge -1 > -3.$$

Když tedy pás P rozložíme na disjunktní sjednocení

$$P = P_1 \cup P_2,$$

kde $P_1 = [0, 2\pi] \times [-2, 2]$ je kompaktní obdélník a P_2 je nekompaktní zbytek, pro každé $a \in P_2$ platí $f(a) \ge -1 > f(s_0) = -3$ a $s_0 \in P_1$. Na hranici obdélníka P_1 má f vždy hodnotu alespoň $-9/4 > -3$ a na jeho vnitřku má f jediné lokální minimum $f(s_0) = -3$. Proto má f na obdélníku P_1 a na celém pásu P jediné ostré globální minimum $f(s_0) = -3$. Z 2π-periodičnosti v proměnné x plyne, že hodnoty $f(s_{2k}) = -3$, $k \in \mathbf{Z}$, jsou právě všechna neostrá globální minima funkce f na $\mathbf{R}^2$.

Věta o implicitních funkcích. Uvažujme soustavu n rovnic o $m+n$ neznámých

$$\begin{aligned} F_1(x_1,\dots,x_m,y_1,\dots,y_n) &= 0 \\ F_2(x_1,\dots,x_m,y_1,\dots,y_n) &= 0 \\ &\vdots \\ F_n(x_1,\dots,x_m,y_1,\dots,y_n) &= 0. \end{aligned}$$

F_i jsou reálné funkce definované na okolí bodu (x_0,y_0) v $\mathbf{R}^{m+n}$, kde x_0 je v $\mathbf{R}^m$ a y_0 v $\mathbf{R}^n$, který je řešením této soustavy, to jest $F_1(x_0,y_0) = F_2(x_0,y_0) = \ldots = F_n(x_0,y_0) = 0$. Nedaly by se neznámé $y_1,\dots,y_n$ ze soustavy eliminovat a nedaly by se vyjádřit, alespoň lokálně v okolí x_0, jako funkce $y_i = f_i(x_1,\dots,x_m)$ neznámých $x_1,\dots,x_m$? Následující věta ukazuje, že jistých předpokladů to možné je.

Zavedeme značení. Pro zobrazení $F = (F_1,F_2,\dots,F_n)$ a $f = (f_1,f_2,\dots,f_n)$, přičemž $F_i = F_i(x_1,\dots,x_m,y_1,\dots,y_n)$ a $f_j = f_j(x_1,\dots,x_m)$, označíme $x = (x_1,x_2,\dots,x_m)$, $y = (y_1,y_2,\dots,y_n)$ a

$$\begin{aligned}
F'_x(x,y) = \left(\frac{\partial F_i}{\partial x_j}\right)_{i,j=1}^{n,m}(x,y) &= \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots & \frac{\partial F_1}{\partial x_m} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial F_n}{\partial x_1} & \frac{\partial F_n}{\partial x_2} & \cdots & \frac{\partial F_n}{\partial x_m} \end{pmatrix}(x,y) \\
F'_y(x,y) = \left(\frac{\partial F_i}{\partial y_j}\right)_{i,j=1}^{n}(x,y) &= \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \frac{\partial F_1}{\partial y_2} & \cdots & \frac{\partial F_1}{\partial y_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial F_n}{\partial y_1} & \frac{\partial F_n}{\partial y_2} & \cdots & \frac{\partial F_n}{\partial y_n} \end{pmatrix}(x,y) \\
f'(x) = \left(\frac{\partial f_i}{\partial x_j}\right)_{i,j=1}^{n,m}(x) &= \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_m} \end{pmatrix}(x).
\end{aligned}$$

První a třetí matice mají rozměr $n \times m$, druhá matice je čtvercová s rozměrem $n \times n$.

Věta 2.12. *Nechť*

$$F = (F_1,F_2,\dots,F_n) :\ W \to \mathbf{R}^n$$

je zobrazení definované na okolí $W \subset \mathbf{R}^{m+n}$ bodu (x_0,y_0), kde $x_0 \in \mathbf{R}^m$ a $y_0 \in \mathbf{R}^n$, které splňuje následující podmínky.

1. *$F_i = F_i(x,y) \in C^1(W)$ pro $1 \le i \le n$.*

2. *$F_i(x_0,y_0) = 0$ pro $1 \le i \le n$.*

3. $\det(F'_y(x_0, y_0)) \neq 0$.

Potom existují okolí $U \subset \mathbf{R}^m$ a $V \subset \mathbf{R}^n$ bodů x_0 a y_0 taková, že $U \times V \subset W$ a pro každý bod $x \in U$ existuje právě jeden bod $y \in V$ splňující $F_i(x, y) = 0$ pro $1 \leq i \leq n$. Jinak řečeno, existuje zobrazení $f = (f_1, f_2, \ldots, f_n) : U \to V$ takové, že

$$\forall (x, y) \in U \times V : \ F(x, y) = \overline{0} \iff y = f(x).$$

Navíc každá funkce f_i je v $C^1(U)$, takže zobrazení f je diferencovatelné na U a jeho Jacobiho matice $f'(x)$ v bodě $x \in U$ splňuje

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

Důkaz této věty dělat nebudeme. Naznačíme ale, jak ze vztahů

$$F_k(x, f_1(x), \ldots, f_n(x)) = 0, \ 1 \leq k \leq n \ \text{ a } \ x \in U,$$

a z $f_i \in C^1(U)$ plyne hořejší formule pro $f'(x)$ a také praktičtější explicitní formule pro $\partial_i f_j(x)$. Parciálním derivováním těchto n rovnic podle proměnné x_i dostáváme n vztahů

$$\frac{\partial F_k}{\partial x_i}(x, f(x)) + \sum_{j=1}^{n} \frac{\partial F_k}{\partial y_j}(x, f(x)) \cdot \frac{\partial f_j}{\partial x_i}(x) = 0, \ 1 \leq k \leq n.$$

To je soustava n rovnic s n neznámými $\partial_i f_j(x)$, $1 \leq j \leq n$, kterou zapíšeme maticově jako

$$F'_y \cdot \partial_i f = -\partial_i F,$$

kde $F'_y = F'_y(x, f(x))$, $\partial_i F$ je sloupcový vektor $(\partial_{x_i} F_1, \partial_{x_i} F_2, \ldots, \partial_{x_i} F_n)^T$, $\partial_i f$ je analogický sloupcový vektor pro f a argumenty parciálních derivací $x, f(x)$ a x pro stručnost vynecháváme. Odtud už pomocí lineární algebry plynou vztahy

$$f'(x) = -(F'_y(x, f(x)))^{-1} \cdot F'_x(x, f(x))$$

a

$$\frac{\partial f_j}{\partial x_i} = -\frac{\det(\partial_{y_1} F, \ldots, \partial_{y_{j-1}} F, \partial_{x_i} F, \partial_{y_{j+1}} F, \ldots, \partial_{y_n} F)}{\det(\partial_{y_1} F, \partial_{y_2} F, \ldots, \partial_{y_n} F)}$$

(v bodech $x \in U$ a $(x, f(x)) \in U \times V$). Podrobnosti viz úloha 1.

Úlohy

1. Rozmyslete si odvození vzorce pro Jacobiho matici implicitních funkcí ve tvaru součinu dvou matic a pro jejich parciální derivace ve tvaru podílu determinantů.

10. přednáška 3. prosince 2007

Věta 2.12 tedy říká, že soustavu

$$F_1(x_1,\ldots,x_m,y_1,\ldots,y_n)=\ldots=F_n(x_1,\ldots,x_m,y_1,\ldots,y_n)=0$$

lze lokálně vyřešit pomocí implicitních funkcí $y_1 = f_1(x_1,\ldots,x_m),\ldots,y_n = f_n(x_1,\ldots,x_m)$, pokud F_i mají lokálně spojité první parciální derivace a v daném bodě je determinant matice derivací funkcí F_i podle y_j nenulový. Navíc pak f_i také mají lokálně spojité první parciální derivace a ty se spočtou jako minus podíl dvou determinantů.

Příklad. Ukažte, že soustava rovnic

$$x+y-\sin z=0 \ \text{ a } \ -x^3-y^3+\mathrm{e}^z-1=0$$

definuje v okolí bodu $x=0$ funkce $y=y(x)$ a $z=z(x)$ třídy C^1 splňující $y(0)=z(0)=0$ a spočítejte hodnoty derivací $y'(0)$ a $z'(0)$.

Pro $F_1(x,y,z)=x+y-\sin z$, $F_2(x,y,z)=-x^3-y^3+\mathrm{e}^z-1$ a $F=(F_1,F_2)$ máme skutečně $F(0,0,0)=(0,0)$ a

$$\begin{aligned}\det(\partial_y F(0,0,0),\partial_z F(0,0,0)) &= \det\begin{pmatrix}1 & -\cos z\\ -3y^2 & \mathrm{e}^z\end{pmatrix}(0,0,0)\\ &= \det\begin{pmatrix}1 & -1\\ 0 & 1\end{pmatrix}\\ &= 1\\ &\neq 0.\end{aligned}$$

Předpoklady věty o implicitních funkcích jsou splněny a uvedené funkce $y(x)$ a $z(x)$ jsou na okolí nuly definovány. Protože

$$\partial_x F(0,0,0)=\begin{pmatrix}1\\ -3x^2\end{pmatrix}(0,0,0)=\begin{pmatrix}1\\ 0\end{pmatrix},$$

podle vztahů uvedených na konci předešlé přednášky máme

$$y'(0)=-\frac{\det\begin{pmatrix}1 & -1\\ 0 & 1\end{pmatrix}}{1}=-1 \ \text{ a } \ z'(0)=-\frac{\det\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}}{1}=0.$$

Diferenciál inverzního zobrazení. Důsledkem věty o implicitních funkcích je zobecnění formule pro derivaci inverzní funkce pro více proměnných.

Důsledek 2.13. *Nechť $U\subset\mathbf{R}^m$ je okolí bodu $a\in\mathbf{R}^m$ a*

$$f:\ U\to\mathbf{R}^m$$

je zobrazení z $C^1(U)$ *splňující* $\det(f'(a)) \neq 0$*. Potom existují okolí* $U_1 \subset U$*, respektive* $V \subset \mathbf{R}^m$ *bodů* a*, respektive* $b = f(a)$ *taková, že* $f:\ U_1 \to V$ *je bijekce, inverzní zobrazení*

$$f^{-1}:\ V \to U_1$$

je z $C^1(V)$ *a pro každé* $x \in U_1$ *v bodě* $y = f(x) \in V$ *máme*

$$\mathrm{D}f^{-1}(y) = (\mathrm{D}f(x))^{-1}.$$

Jacobiho matice inverzního zobrazení f^{-1} *v bodě* y *je tedy inverzní k Jacobiho matici zobrazení* f *v bodě* x*.*

Důkaz. Uvažme zobrazení o $2m$ proměnných

$$F(x,y) = f(x) - y:\ U \times \mathbf{R}^m \to \mathbf{R}^m,$$

kde $F = (F_1, \ldots, F_m)$ a $F_i(x,y) = f_i(x) - y_i$. Pak $F_i(a,b) = 0$ pro $1 \le i \le n$. Na soustavu $F(x,y) = \overline{0}$ aplikujeme větu 2.12. Doplňte další podrobnosti důkazu jako cvičení (úloha 1). □

Vázané extrémy. Dalším důsledkem věty o implicitních funkcích je zobecnění první části věty 2.11 (nutnou podmínkou lokálního extrému funkce v bodě otevřené množiny je nulovost všech parciálních derivací) na extrémy na množině zadané soustavou rovnic. Nechť $U \subset \mathbf{R}^m$ je otevřená množina a

$$f, F_1, \ldots, F_n:\ U \to \mathbf{R}$$

jsou funkce z $C^1(U)$, přičemž $n < m$. Hledáme lokální extrémy funkce f na množině

$$H = \{x \in U \mid F_1(x) = F_2(x) = \cdots = F_n(x) = 0\}.$$

Typicky tato množina nemá žádný vnitřní bod a nelze použít větu 2.11. Příkladem je jednotková sféra v $\mathbf{R}^m$:

$$H = \{x \in \mathbf{R}^m \mid x_1^2 + x_2^2 + \ldots + x_m^2 - 1 = 0\}.$$

Důsledek 2.14 (Lagrangeovy multiplikátory). *Nechť* a *je bod z* H*. Když jsou vektory* $\nabla F_1(a), \ldots, \nabla F_n(a)$ *z* $\mathbf{R}^m$ *lineárně nezávislé a vektor* $\nabla f(a)$ *není jejich lineární kombinací, pak* f *nemá v bodu* a *na množině* H *ani neostrý lokální extrém.*

Ekvivalentně řečeno, když jsou $\nabla F_1(a), \ldots, \nabla F_n(a)$ *lineárně nezávislé a funkce* f *má v bodě* a *na množině* H *(ostrý či neostrý) lokální extrém, potom existují taková reálná čísla* $\lambda_1, \ldots, \lambda_n$ *(tzv.* Lagrangeovy multiplikátory*), že*

$$\nabla f(a) - \sum_{i=1}^{n} \lambda_i \nabla F_i(a) = (0, 0, \ldots, 0),$$

neboli $\partial_{x_j} f(a) - \lambda_1 \partial_{x_j} F_1(a) - \cdots - \lambda_n \partial_{x_j} F_n(a) = 0$ *pro* $1 \le j \le m$*.*

Důkaz. V důkazu použijeme druhou formulaci a předpokládáme, že f má v a na H lokální extrém. Lineární nezávislost uvedených vektorů znamená, že když je složíme jako řádky do matice (Jacobiho matice zobrazení $F = (F_1, \ldots, F_n)$), má tato $n \times m$ matice takových n sloupců, že determinant odpovídající čtvercové podmatice je nenulový. Pro jednoduchost značení předpokládáme, že to je posledních n sloupců. Přeznačíme-li tedy proměnné jako

$$\begin{aligned} x_1, x_2, \ldots, x_m &= y_1, y_2, \ldots, y_{m-n}, z_1, z_2, \ldots, z_n \\ &= y, z, \end{aligned}$$

máme $\det(\partial_{z_1} F(a), \ldots, \partial_{z_n} F(a)) \neq 0$. Podle věty 2.12 existují taková okolí U_1 a V_1 bodů

$$y_0 = (a_1, \ldots, a_{m-n}) \ \text{ a } \ z_0 = (a_{m-n+1}, \ldots, a_m)$$

a takové zobrazení $g = (g_1, \ldots, g_n) : U_1 \to V_1$, že pro (y, z) probíhající $U_1 \times V_1$ máme

$$F_i(y, z) = 0 \ \text{ pro } 1 \leq i \leq n \iff z = g(y)$$

a speciálně $g(y_0) = z_0$. Uvažme nyní funkci

$$h(y) = f(y, g_1(y), \ldots, g_n(y)) : U_1 \to \mathbf{R}.$$

Protože má v y_0 lokální extrém (nyní už bez vazby), první část věty 2.11 dává $\nabla h(y_0) = \overline{0}$. Parciálním derivováním složené funkce máme

$$\frac{\partial f}{\partial y_i}(y_0, g(y_0)) + \sum_{j=1}^{n} \frac{\partial f}{\partial z_j}(y_0, g(y_0)) \cdot \frac{\partial g_j}{\partial y_i}(y_0) = 0, \ 1 \leq i \leq m - n.$$

V řeči Jacobiho matic,

$$f_y'(y_0, z_0) + f_z'(y_0, z_0) \cdot g'(y_0) = \overline{0}.$$

Za $g'(y_0)$ dosadíme vyjádření podle vzorce ve větě 2.12:

$$f_y'(y_0, z_0) - f_z'(y_0, z_0) \cdot F_z'(y_0, z_0)^{-1} \cdot F_y'(y_0, z_0) = \overline{0}.$$

Označíme-li

$$\lambda = (\lambda_1, \ldots, \lambda_n) = f_z'(y_0, z_0) \cdot F_z'(y_0, z_0)^{-1},$$

dostáváme

$$f_y'(y_0, z_0) - \lambda F_y'(y_0, z_0) = \overline{0}.$$

Ale z $\lambda = f_z'(y_0, z_0) \cdot (F_z')^{-1}(y_0, z_0)$ úpravou plyne, že stejný vztah platí i v z-ových proměnných: $f_z'(y_0, z_0) - \lambda F_z'(y_0, z_0) = \overline{0}$. Celkem v y-ových i z-ových proměnných máme

$$f'(y_0, z_0) - \lambda F'(y_0, z_0) = \overline{0},$$

takže $\nabla f(a) = \lambda_1 \nabla F_1(a) + \ldots + \lambda_n \nabla F_n(a)$. □

Všimněte si, že $\nabla f(a) = \overline{0}$ je lineární kombinací gradientů $\nabla F_i(a)$ vždy a pro nulový gradient f v a tedy důsledek 2.14 (stejně jako první část věty 2.11) nic neříká. Pomocí *Lagrangeovy funkce*

$$L(x, \lambda) = L(x_1, \dots, x_m, \lambda_1, \dots, \lambda_n) = f(x) - \sum_{i=1}^{n} \lambda_i F_i(x)$$

se důsledek 2.14 dá hezky přeformulovat. Protože

$$\nabla L = (\partial_{x_1} f - \textstyle\sum_1^n \lambda_i \partial_{x_1} F_i,\ \dots,\ \partial_{x_m} f - \sum_1^n \lambda_i \partial_{x_m} F_i,\ -F_1,\ \dots,\ -F_n)$$

(v bodech (x, λ) a x), je $\nabla L(a, \lambda) = \overline{0}$ přesně ekvivalentní tomu, že bod a leží na ploše H (posledních n souřadnic gradientu) a že koeficienty $\lambda = (\lambda_1, \dots, \lambda_n)$ jsou Lagrangeovy multiplikátory (prvních m souřadnic gradientu). Nutnou podmínku lokálního extrému funkce f v bodě a vzhledem k H tedy můžeme zformulovat i takto:

> Když má funkce f v bodě $a \in H$ lokální vázaný extrém, existuje takový bod $\lambda \in \mathbf{R}^n$, že $\nabla L(a, \lambda) = \overline{0}$.

Zde se o náležení a do H nemusíme starat, protože je v podmínce $\nabla L(a, \lambda) = \overline{0}$ automaticky zahrnuto.

Uvedeme ještě jednu ekvivalentní formulaci důsledku 2.14. Nechť $\nabla f(a) \neq \overline{0}$ a vektory $\nabla F_1(a), \dots, \nabla F_n(a)$ jsou lineárně nezávislé. Uvažme dva vektorové podprostory $\mathbf{R}^m$ složené z vektorů kolmých na $\nabla f(a)$, respektive z vektorů kolmých na každý z vektorů $\nabla F_1(a), \dots, \nabla F_n(a)$:

$$\begin{aligned} TN_a &= \{x \in \mathbf{R}^m \mid \langle \nabla f(a), x \rangle = 0\} \\ TH_a &= \{x \in \mathbf{R}^m \mid \langle \nabla F_1(a), x \rangle = \cdots = \langle \nabla F_n(a), x \rangle = 0\}. \end{aligned}$$

Podprostor TN_a má dimenzi $m-1$ a TH_a má dimenzi $m-n$. Pomocí implicitních funkcí se dá ukázat, že $a + TH_a$ je tečným afinním podprostorem k ploše $H = \{x \in \mathbf{R}^m \mid F_1(x) = \cdots = F_n(x) = 0\}$ v bodě a a podobně je $a + TN_a$ tečnou afinní nadrovinou v bodě a k „vrstevnicové“ ploše

$$N = \{x \in \mathbf{R}^m \mid f(x) = f(a)\}.$$

Podprostorům TN_a a TH_a se říká *tečné prostory* (k odpovídajícím plochám v bodě a). Z lineární algebry (teorie ortogonálních doplňků) víme, že $\nabla f(a)$ je lineární kombinací vektorů $\nabla F_1(a), \dots, \nabla F_n(a)$, právě když $TH_a \subset TN_a$. Nutná podmínka lokálního vázaného extrému má tedy i tuto geometrickou formulaci.

> Když má funkce f v bodě $a \in H$ lokální vázaný extrém, je tečný prostor TH_a k ploše H v bodě a obsažený v tečném prostoru TN_a k vrstevnicové ploše N funkce f v bodě a,
>
> $$TH_a \subset TN_a.$$

Příklad: auto na horské silnici. Podíváme se na situaci $m = 2$ a $n = 1$. Funkce $f : \mathbf{R}^2 \to \mathbf{R}$ například udává nadmořskou výšku $f(x)$ bodu v terénu se zeměpisnými souřadnicemi x a křivka $H = \{x \in \mathbf{R}^2 \mid F(x) = 0\}$ je silnice. Vrstevnice $N = \{x \in \mathbf{R}^2 \mid f(x) = f(a) = b\}$, kde $a \in H$, je též rovinná křivka. Nechť $\nabla f(a)$ a $\nabla F(a)$ jsou nenulové vektory. Tečné prostory TH_a a TN_a pak mají dimenzi 1 a přímky $p = a + TH_a$ a $q = a + TN_a$ jsou tečny ke křivkám H a N v jejich průsečíku a. Předpokládejme, že $TH_a \not\subset TN_a$. Pak p a q jsou dvě různé přímky procházející společným bodem a. Vrstevnice N, která poblíž a úzce sleduje q, musí v průsečíku a přecházet z jedné strany silnice H na její druhou stranu, protože H zase úzce sleduje p. Pokud nadmořskou výšku vrstevnice b málo změníme, zmenšíme nebo zvětšíme, vrstevnice N se též změní jen málo a stále musí přecházet z jedné strany H na druhou a musí tak H protínat. Na silnici se tedy v okolí a nacházejí body jak s menší tak s větší nadmořskou výškou, než má a, a proto f nemá v a na H lokální extrém. Pokud na silnici H zastaví v bodě a auto, v neutrálu se bez ruční brzdy určitě rozjede!

Úlohy

1. (budou doplněny)

11. přednáška 10. prosince 2007

Kapitola 3. Úvod do teorie diferenciálních rovnic.

Obyčejná diferenciální rovnice řádu n (ODR řádu n) je vztah

$$F(x, y, y', y'', \ldots, y^{(n)}) = 0$$

mezi argumentem x funkce jedné proměnné y, její hodnotou $y = y(x)$ v x a hodnotami jejích prvních n derivací $y'(x), \ldots, y^{(n)}(x)$ v x. Je zadaná funkcí s $n + 2$ proměnnými

$$F(x_1, x_2, \ldots, x_{n+2}).$$

Dvojice (I, y), kde $I \subset \mathbf{R}$ je otevřený interval a $y: I \to \mathbf{R}$ na něm definovaná funkce, je jejím *řešením*, když má y na I derivace až do řádu n a pro každé x z I leží $(n + 2)$-tice čísel $(x, y(x), y'(x), y''(x), \ldots, y^{(n)}(x))$ v definičním oboru funkce F a ta na ní nabývá nulovou hodnotu.

Přívlastek „obyčejný" vymezuje, že neznámou v rovnici je funkce jedné proměnné. *Parciálními diferenciálními rovnicemi* (PDR), které mají jako neznámé funkce více proměnných a svazují hodnoty jejich parciálních derivací, se na této přednášce nebudeme zabývat. ODR, a ještě více PDR, jsou důležité jako matematické modely problémů z fyziky, techniky, biologie, ekonomie, ... Uvedeme dva příklady. V obou jako definiční interval I bereme celé $\mathbf{R}$.

Volný pád a radioaktivní rozpad. Newtonův zákon síly $ma = F$ (m je hmotnost, a zrychlení a F síla) se vyjadřuje diferenciální rovnicí

$$mx'' = F,$$

kde $x = x(t) \in \mathbf{R}$ je poloha částice o hmotnosti m v čase t (uvažujeme jen jednoduchý jednorozměrný případ) vystavené působení síly F. Hmotnost i sílu bereme jako konstanty, i když v mnoha situacích také závisejí na čase t (a/nebo poloze částice x a dalších parametrech). Nechť F představuje třeba působení tíhového pole Země, které je v jednoduchém modelu konstantní (nemění se s časem, nezávisí na poloze částice atd.). Pak máme *rovnici volného pádu*

$$mx'' = -mg$$

($g \approx 9.81$ je konstanta tíhového zrychlení). Záporné znaménko znamená, že tíhová síla je orientována směrem do $-\infty$ reálné osy. Jejím řešením je každá funkce

$$x(t) = -\frac{1}{2}gt^2 + c_1 t + c_2,$$

kde c_1 a c_2 jsou libovolné konstanty. Ty vyjadřují skutečnost, že pohyb částice v tíhovém poli je úplně určen teprve zadáním její polohy $x(t_0)$ a rychlosti $x'(t_0)$ v nějakém okamžiku t_0. Například pro $t_0 = 0$, $x(0) = 0$ a $x'(0) = v > 0$ máme $c_1 = v$ a $c_2 = 0$. Funkce

$$x(t) = -\frac{1}{2}gt^2 + vt$$

pak popisuje pohyb hmotného bodu vrženého v čase $t_0 = 0$ z bodu 0 rychlostí v směrem vzhůru. (V čase $t_1 = 2v/g$, jenž je druhým řešení rovnice $x(t) = 0$ vedle $t_0 = 0$, částice znovu proletí bodem $x = 0$.)

Rovnice radioaktivního rozpadu

$$R' = -kR$$

popisuje vývoj množství $R = R(t)$ rozpadajícího se radioaktivního materiálu v čase t a $k > 0$ je materiálová konstanta. Tuto rovnici snadno odvodíme z fyzikálního předpokladu, že pro malé $\Delta > 0$ má každý atom v daném množství radioaktivního materiálu v každém okamžiku t_0 pravděpodobnost $k\Delta$, že se v náledující časovém intervalu $[t_0, t_0 + \Delta]$ rozpadne. Je jasné, že každá funkce

$$R(t) = c\exp(-kt),$$

kde c je konstanta, je řešením této rovnice.

Dvě věty o existenci a jednoznačnosti řešení ODR prvního řádu. V předchozích dvou příkladech jsme uhádli řešení diferenciální rovnice, ale nebylo jasné, zda neexistují ještě jiná řešení. Uvedeme dvě obecné věty zaručující existenci a za silnějšího předpokladu i jednoznačnost řešení. Uvažme ODR 1. řádu s počáteční podmínkou, která je navíc vyřešená vhledem k derivaci:

$$(*)\left\{\begin{array}{rcl} y(a) & = & b \\ y'(x) & = & f(x, y(x)). \end{array}\right.$$

Předpokládáme, že rovnicová funkce f je spojitá na nějaké otevřené množině $\Omega \subset \mathbf{R}^2$. Následující větu nebudeme dokazovat.

Věta 3.1 (Peanova). *Nechť $(a, b) \in \Omega$ a $f \in \mathcal{C}(\Omega)$. Potom existuje takové $\delta > 0$, že na intervalu $(a - \delta, a + \delta)$ má rovnice $(*)$ řešení $y(x)$.*

Pouhá spojitost rovnicové funkce však nezaručuje jednoznačnost řešení. Nechť $\Omega = \mathbf{R}^2$ a uvažme rovnici $y(0) = 0, y' = xy^{2/3}$ (zde $y^{2/3}$ bereme jako $(y^2)^{1/3}$, takže mocnina je definovaná pro každé $y \in \mathbf{R}$). V okolí 0, a vlastně na celém $\mathbf{R}$, má dvě řešení: $y_1(x) \equiv 0$ a $y_2(x) = x^6/6^3$. Obecněji, zvolíme-li $c > 0$, potom funkce $y(x)$ definovaná jako $(x^2 - c)^3/6^3$ pro $x \in \mathbf{R}\backslash(-\sqrt{c}, \sqrt{c})$ a jako 0 pro $x \in [-\sqrt{c}, \sqrt{c}]$ je řešením. Máme dokonce nekonečně mnoho řešení.

Řekneme, že funkce $f(x, y)$ je *lokálně lipschitzovská na množině Ω vzhledem k proměnné y*, když pro každý bod $a \in \Omega$ existují konstanty $\varepsilon > 0$ a $K > 0$ takové, že pro každé dva body (x_0, y_1) a (x_0, y_2) z ε-ového okolí bodu a platí $|f(x_0, y_1) - f(x_0, y_2)| < K|y_1 - y_2|$. Lokální lipschitzovskost vyplývá například ze spojitosti parciální derivace $\partial_y f$ na Ω.

Věta 3.2 (Picardova). *Nechť $(a, b) \in \Omega$, $f \in \mathcal{C}(\Omega)$ a f je na Ω lokálně lipschitzovská vzhledem k proměnné y. Potom existuje $\delta > 0$ takové, že na intervalu $(a - \delta, a + \delta)$ má rovnice $(*)$ právě jedno řešení $y(x)$.*

Tuto větu jsme dokázali jako větu 1.11 na 4. přednášce. Pro rovnicovou funkci $f(x,y) = xy^{2/3}$ z předchozího příkladu tuto větu nelze pro bod $(0,0)$ použít, f není v jeho okolí lipschitzovská vzhledem k y.

Lineární ODR prvního řádu. Vyřešíme lineární diferenciální rovnici prvního řádu

$$y' + a(x)y = b(x).$$

Zde $y = y(x)$ je neznámá funkce a funkce $a, b:\ I \to \mathbf{R}$ jsou spojité na nějakém otevřeném intervalu I.

Řešení metodou integračního faktoru. Nejprve nalezneme takovou funkci $c = c(x)$, tzv. integrační faktor, že $c(y' + ay) = (cy)'$. Pak $cy' + acy = cy' + c'y$ a c musí splňovat rovnici $ac = c'$, čili $(\log c)' = a$. Funkce $c = \mathrm{e}^A$, kde $A = A(x)$ je nějaká primitivní funkce k $a(x)$, má tedy požadovanou vlastnost. Výchozí lineární rovnici vynásobíme integračním faktorem a dostaneme

$$(cy)' = c(y' + ay) = cb.$$

Takže $(cy)' = cb$ a $cy = D + c_0$, kde D je primitivní funkce k cb a c_0 je integrační konstanta. Máme řešení $y = c^{-1}(D + c_0)$. Shrnuto,

$$y(x) = \mathrm{e}^{-A(x)} \left(\int \mathrm{e}^{A(x)} b(x)\, dx + c_0 \right), \quad \text{kde } A(x) = \int a(x)\, dx.$$

Všimněte si, že $y(x)$ je definovaná na celém I (definičním oboru funkcí a a b) a že každé počáteční podmínce $y(x_0) = y_0$ odpovídá přesně jedna hodnota integrační konstanty c_0, pro níž je splněna. Zavedení integrační konstanty pro A, tj. nahrazení $A(x)$ obecnějším výrazem $A(x) + c_1$, už nedává obecnější řešení, které by se nedalo dostat jen s pomocí konstanty c_0.

Řešení metodou variace konstant. Nejprve vyřešíme homogenní rovnici $y' + ay = 0$. Odtud $y'/y = -a$ a $(\log y)' = -a$. Dostáváme $\log y = -A + c$ a $y = \mathrm{e}^c \mathrm{e}^{-A} = K\mathrm{e}^{-A}$, kde A je primitivní funkce k a a c a K jsou konstanty. Konstantu K v řešení $y(x) = K\mathrm{e}^{-A(x)}$ homogenní rovnice nahradíme funkcí $K = K(x)$ a obecnou funkci $K(x)\mathrm{e}^{-A(x)}$ dosadíme do původní rovnice, čímž dostaneme podmínku na $K(x)$:

$$\begin{aligned} (K\mathrm{e}^{-A})' + a \cdot K\mathrm{e}^{-A} &= b \\ K'\mathrm{e}^{-A} - Ka\mathrm{e}^{-A} + Ka\mathrm{e}^{-A} &= b \\ K' &= b\mathrm{e}^{A}. \end{aligned}$$

Takže $K(x) = \int b(x)\mathrm{e}^{A(x)}\, dx + c$ a po dosazení do $y(x) = K(x)\mathrm{e}^{-A(x)}$ dostáváme opět shora uvedený vzorec.

Příklad. Volný pád s odporem prostředí. Uvažujme částici o hmotnosti m, která z klidu padá vlivem konstantní tíže a na kterou kromě tíže působí i odpor prostředí. Předpokládejme, že síla odporu závisí lineárně na rychlosti částice—to je samozřejmě zjednodušení, ve skutečnosti je závislost složitější. Newtonův zákon síly dává pohybovou rovnici

$$m\frac{\mathrm{d}v}{\mathrm{d}t} = \text{tíže} - \text{odpor} = mg - kv,$$

kde $v = v(t)$ je rychlost částice v čase t, g je konstanta tíhového zrychlení a $k > 0$ je konstanta odporu prostředí. Máme lineární diferenciální rovnici

$$v' + av = b,$$

kde $a = k/m$ a $b = g$ jsou konstanty. Integrační faktor tedy je $c = \mathrm{e}^{kt/m}$ a podle hořejšího vzorce máme řešení

$$v(t) = \frac{mg}{k} + c_1\mathrm{e}^{-kt/m}.$$

Z počáteční podmínky $v(0) = 0$ vypočteme hodnotu integrační konstanty $c_1 = -mg/k$. Takže

$$v(t) = \frac{mg}{k}\left(1 - \mathrm{e}^{-kt/m}\right).$$

Pro $t \to \infty$ se tedy rychlost částice blíží k limitní rychlosti

$$v_{lim} = \frac{mg}{k}.$$

Tento vzorec plyne také uvážením rovnovážného stavu, kdy se tíže rovná síle odporu.

ODR prvního řádu se separovanými proměnnými. Je to diferenciální rovnice tvaru

$$y' = f(x)g(y),$$

kde $f(x)$ a $g(y)$ jsou funkce definované a spojité na nějakém otevřeném intervalu I a $g \neq 0$ na I. Jedná se obecně o nelineární diferenciální rovnici, v níž na pravé straně můžeme od sebe oddělit—separovat—proměnné x a y.

Rovnici upravíme do tvaru

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

a ten přepíšeme pomocí funkce $G(t)$, jež je primitivní k funkci $1/g(t)$ na intervalu I, jako $G(y(x))' = f(x)$. Odtud dostáváme vztah $G(y(x)) = F(x) + c$, kde $F(x)$ je primitivní funkce k $f(x)$ na I a c je integrační konstanta. Řešení původní diferenciální rovnice je tedy dáno jako implicitní funkce vztahem

$$G(y(x)) = F(x) + c, \quad \text{kde} \quad G(t) = \int \frac{dt}{g(t)} \quad \text{a} \quad F(x) = \int f(x)\,dx.$$

Postup při řešení rovnice se separovanými proměnnými se obvykle zapisuje takto:

$$\begin{aligned}\frac{\mathrm{d}y}{\mathrm{d}x} &= f(x)g(y)\\ g(y)^{-1}\mathrm{d}y &= f(x)\mathrm{d}x\\ \int g(y)^{-1}\,\mathrm{d}y &= \int f(x)\,\mathrm{d}x\\ G(y) &= F(x)+c.\end{aligned}$$

Dva důležité speciální případy jsou rovnice $y' = f(x)$ a $y' = g(y)$. Řešení první z nich jsou právě funkce primitivní k $f(x)$ na I. Řešení rovnice $y' = g(y)$ je dáno implicitně jako $G(y(x)) = x + c$ a je to tedy funkce inverzní ke $G(x)+c$:

$$y(x) = \left(\int \frac{dx}{g(x)} + c\right)^{\langle -1\rangle}.$$

Úlohy

1. (budou doplněny)

12. přednáška 17. prosince 2007

Soustavy lineárních diferenciálních rovnic prvního řádu. Jedna diferenciální rovnice n-tého řádu se takto převede na ekvivalentní soustavu diferenciálních rovnic prvého řádu: funkce $y = y(x)$ je na intervalu I řešením rovnice

$$F(x, y, y', y'', \dots, y^{(n)}) = 0,$$

právě když n-tice funkcí $y, y_1, \dots, y_{n-1}$ je na intervalu I řešením soustavy rovnic prvního řádu

$$y' = y_1,\ y_1' = y_2,\ \dots, y_{n-2}' = y_{n-1},\ F(x, y, y_1, y_2, \dots, y_{n-1}, y_{n-1}') = 0.$$

Za snížení řádu rovnice jsme ovšem zaplatili zavedením dalších $n - 1$ funkcí.

Lineární diferenciální rovnice n-tého řádu

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_0 y = b,$$

kde a_i a b jsou zadané funkce, je tedy ekvivalentní speciální soustavě lineárních rovnic prvního řádu

$$y' = y_1,\ y_1' = y_2,\ \dots, y_{n-2}' = y_{n-1},\ y_{n-1}' = -a_{n-1}y_{n-1} - \dots - a_1 y_1 - a_0 y + b.$$

Budeme se proto zabývat teorií obecných soustav lineárních diferenciálních rovnic prvního řádu

$$y_i' = a_{i,1}y_1 + a_{i,2}y_2 + \dots + a_{i,n}y_n + b_i,\ 1 \le i \le n,$$

kde $a_{i,j} = a_{i,j}(x)$ a $b_i = b_i(x)$, $1 \le i, j \le n$, je $n^2 + n$ zadaných funkcí, definovaných na nějakém otevřeném intervalu I, a $y_i = y_i(x)$, $1 \le i \le n$, jsou neznámé funkce. V maticovém zápisu,

$$y' = Ay + b,$$

kde $A:\ I \to \mathbf{R}^{n\times n}$ a $b:\ I \to \mathbf{R}^n$ je daná maticová a daná vektorová funkce a $y:\ I \to \mathbf{R}^n$ je neznámá vektorová funkce. V dalším budeme vždy předpokládat, že funkce $a_{i,j}$ a b_i jsou na intervalu I spojité.

Věta 3.3. *Nechť $I \subset \mathbf{R}$ je otevřený interval, $\alpha \in I$ a $\beta \in \mathbf{R}^n$ jsou počáteční podmínky a $a_{i,j},\ b_i:\ I \to \mathbf{R}$, $1 \le i, j \le n$, jsou spojité funkce. Soustava lineárních diferenciálních rovnic s počátečními podmínkami*

$$\begin{aligned} y(\alpha) &= \beta \\ y'(x) &= A(x) \cdot y(x) + b(x) \end{aligned}$$

má na intervalu I jediné řešení—existuje jediná n-tice funkcí $y_1, \dots, y_n$ z $C^1(I)$, která pro každé $i \in \{1, 2, \dots, n\}$ a $x \in I$ splňuje rovnosti

$$y_i(\alpha) = \beta_i \quad a \quad y_i'(x) = \sum_{j=1}^{n} a_{i,j}(x) y_j(x) + b_i(x).$$

Důkaz věty 3.3, který je opět založený na větě o kontrahujícím zobrazení, nebudeme na přednášce dělat. Na rozdíl od vět 3.1 a 3.2 dostáváme globální existenci a jednoznačnost řešení na celém intervalu I. Z věty 3.3 plyne, že pokud se dvě řešení z a u soustavy $y' = Ay + b$ shodují v jednom bodě $x_0 \in I$ (tj. $z(x_0)$ a $u(x_0)$ je tatáž n-tice z $\mathbf{R}^n$), potom se shodují na celém I, $z(x) = u(x)$ pro $\forall x \in I$.

Uvažme množinu řešení homogenní soustavy $y' = Ay$ a množinu řešení nehomogenní soustavy $y' = Ay + b$:

$$H = \{y \in C^1(I)^n \mid y' = Ay \text{ na } I\} \text{ a } N = \{y \in C^1(I)^n \mid y' = Ay + b \text{ na } I\}.$$

Obě jsou obsažené v množině n-tic funkcí $C^1(I)^n$, což je vektorový prostor nad $\mathbf{R}$ nekonečné dimenze.

Tvrzení 3.4. *H je vektorový podprostor $C^1(I)^n$ s dimenzí n. N je afinní podprostor $C^1(I)^n$ s dimenzí n. Pro každé řešení $y \in N$ platí, že $N = y + H = \{y + z \mid z \in H\}$.*

Důkaz. Díky linearitě derivování a maticového násobení je zřejmé, že H je vektorový podprostor: Pokud $y, z \in H$ a $\alpha, \beta \in \mathbf{R}$, pak $(\alpha y + \beta z)' = \alpha y' + \beta z' = \alpha Ay + \beta Az = A(\alpha y + \beta z)$ a $\alpha y + \beta z \in H$. Stejně se dokážou i implikace $y, z \in N \Rightarrow y - z \in H$ a $y \in N, z \in H \Rightarrow y + z \in N$, které dávají, že $N = y + H$. Existence alespoň jednoho řešení $y \in N$ plyne z věty 3.3.

Dokážeme, že $\dim H = n$. Odtud plyne, že $\dim N = n$. Nechť $x_0 \in I$ je libovolné číslo, $\{e^i \in \mathbf{R}^n \mid 1 \le i \le n\}$ je kanonická báze $\mathbf{R}^n$ (i-tá složka e^i je 1 a ostatní jsou 0) a $\{y^i \in H \mid 1 \le i \le n\}$ jsou řešení homogenní soustavy splňující počáteční podmínky $y^i(x_0) = e^i$, $1 \le i \le n$—tato řešení existují podle věty 3.3. Je jasné (podle hodnot v x_0), že $\{y^1, \dots, y^n\}$ je lineárně nezávislá množina v $C^1(I)^n$. Je-li $y \in H$ libovolné řešení, které má v x_0 hodnoty

$$y(x_0) = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbf{R}^n,$$

potom funkce $z(x) = \sum_{i=1}^n \alpha_i y^i(x)$ patří do H a $z(x_0) = y(x_0)$. Podle věty 3.3 máme $z(x) = y(x)$ pro každé $x \in I$ a tedy $y = \sum_{i=1}^n \alpha_i y^i$. Takže $H = Lin(\{y^1, \dots, y^n\})$ a $\dim H = n$. □

Každá báze prostoru H se nazývá *fundamentální systém řešení (FSŘ)* homogenní soustavy $y' = Ay$.

Wronskián. *Wronského determinant* neboli *wronskián* n-tice vektorových funkcí

$f^1, \dots, f^n : I \to \mathbf{R}^n$ je determinant matice $n \times n$ s $f^1, \dots, f^n$ ve sloupcích:

$$W(x) = W_{f^1,\dots,f^n}(x) = \det \begin{pmatrix} f_1^1(x) & f_1^2(x) & \dots & f_1^n(x) \\ f_2^1(x) & f_2^2(x) & \dots & f_2^n(x) \\ \vdots & \vdots & \vdots & \vdots \\ f_n^1(x) & f_n^2(x) & \dots & f_n^n(x) \end{pmatrix}.$$

Připomeňme si, že $f^1, \dots, f^n$ jsou *lineárně závislé (LZ)*, existují-li konstanty $\alpha_1, \dots, \alpha_n \in \mathbf{R}$, ne všechny nulové, že

$$\sum_{i=1}^{n} \alpha_i f^i$$

je identicky nulová funkce, to jest pro každé $x \in I$ máme $\sum_{i=1}^n \alpha_i f^i(x) = \overline{0}$. Zřejmě

$$f^1, \dots, f^n \text{ jsou LZ} \implies W_{f^1,\dots,f^n}(x) = 0 \text{ pro } \forall x \in I$$

(matice definující $W(x)$ má pro každé $x \in I$ lineárně závislé sloupce). Opačná implikace obecně neplatí (úloha 1). Platí však v případě, že $f^1, \dots, f^n$ jsou řešení homogenní soustavy $y' = Ay$.

Tvrzení 3.5. *Nechť vektorové funkce $f^1, \dots, f^n : I \to \mathbf{R}^n$ na I splňují $(f^i)' = Af^i$, pro danou maticovou funkci $A : I \to \mathbf{R}^{n\times n}$ se spojitými položkami, a W je jejich wronskián. Pak*

$$\exists x \in I : W(x) = 0 \implies f^1, \dots, f^n \textit{ jsou LZ.}$$

Máme tedy ekvivalenci

$$\exists x \in I : W(x) = 0 \iff \forall x \in I : W(x) = 0.$$

Důkaz. Pokud $W(x_0) = 0$ pro $x_0 \in I$, matice hodnot vektorových funkcí $f^i(x_0)$ má lineárně závislé sloupce: $\sum_{i=1}^n \alpha_i f^i(x_0) = \overline{0}$ pro nějaká $\alpha_i \in \mathbf{R}$, ne všechny nulové. Lineární kombinace

$$f(x) = \sum_{i=1}^{n} \alpha_i f^i(x)$$

je pak na I řešením soustavy $f' = Af$ a splňuje počáteční podmínku $f(x_0) = \overline{0}$. Jiným řešením $y' = Ay$ splňujícím $y(x_0) = \overline{0}$ je ovšem identicky nulová vektorová funkce. Podle věty 3.3 se obě řešení na I rovnají a funkce f je tedy identicky nulová. Takže $\sum_{i=1}^n \alpha_i f^i(x) = \overline{0}$ pro každé $x \in I$ a $f^1, \dots, f^n$ jsou LZ. V ekvivalenci je implikace $\Leftarrow$ triviální a $\Rightarrow$ plyne spojením právě dokázané implikace a implikace uvedené před tvrzením. $\square$

Wronskián n-tice řešení $f^1, \dots, f^n$ homogenní soustavy $y' = Ay$ je tedy na I buď vždy nenulový a $f^1, \dots, f^n$ tvoří FSŘ, nebo je na I vždy nulový a $f^1, \dots, f^n$ jsou LZ a netvoří FSŘ.

Variace konstant pro soustavy. Následující vzorec ukazuje, jak pomocí FSŘ homogenní soustavy $y' = Ay$ dostat jedno (tzv. *partikulární*) řešení nehomogenní soustavy $y' = Ay + b$.

Věta 3.6. *Nechť $I \subset \mathbf{R}$ je otevřený interval, $x_0 \in I$ a $y^0 \in \mathbf{R}^n$ jsou počáteční podmínky, $A: I \to \mathbf{R}^{n \times n}$ a $b: I \to \mathbf{R}^n$ je maticová a vektorová funkce se spojitými položkami, $y^1, \dots, y^n : I \to \mathbf{R}^n$ je FSŘ homogenní soustavy $y' = Ay$ a*

$$Y = Y(x) = \begin{pmatrix} y_1^1(x) & y_1^2(x) & \dots & y_1^n(x) \\ y_2^1(x) & y_2^2(x) & \dots & y_2^n(x) \\ \vdots & \vdots & \vdots & \vdots \\ y_n^1(x) & y_n^2(x) & \dots & y_n^n(x) \end{pmatrix}$$

je matice hodnot vektorových funkcí y^i. Potom je vektorová funkce $z: I \to \mathbf{R}^n$ definovaná formulí

$$z(x) = Y(x) \left(\int_{x_0}^{x} Y(t)^{-1} \cdot b(t)\, dt + Y(x_0)^{-1} \cdot y^0 \right)$$

(vektorovou funkci v integrandu integrujeme po složkách) řešením nehomogenní soustavy $z' = Az + b$ a splňuje počáteční podmínku $z(x_0) = y^0$.

Důkaz. Řešení soustavy $z' = Az + b$ budeme hledat ve tvaru kombinace

$$z = \sum_{i=1}^{n} c_i y^i = Yc,$$

kde $c_1(x), \dots, c_n(x)$ jsou neznámé funkce a c je jejich sloupcový vektor. Máme

$$\begin{aligned} z' &= \left(\sum_{i=1}^{n} c_i y^i \right)' = \sum_{i=1}^{n} (c_i y^i)' \\ &= \sum_{i=1}^{n} c_i (y^i)' + \sum_{i=1}^{n} c_i' y^i = \sum_{i=1}^{n} c_i A y^i + \sum_{i=1}^{n} c_i' y^i \\ &= A \sum_{i=1}^{n} c_i y^i + \sum_{i=1}^{n} c_i' y^i \\ &= Az + \sum_{i=1}^{n} c_i' y^i. \end{aligned}$$

Takže bude platit $z' = Az + b$, pokud

$$Yc' = \sum_{i=1}^{n} c_i' y^i = b.$$

Funkce $y^1, \dots, y^n$ tvoří FSŘ soustavy $y' = Ay$, jejich wronskián $W = \det Y$ je podle Tvrzení 3.5 v každém bodě x intervalu I nenulový a matice $Y(x)$ je invertibilní. Tudíž

$$c' = Y^{-1}b \quad \text{a} \quad c(x) = \int_{x_0}^{x} Y(t)^{-1} \cdot b(t)\, dt + d,$$

kde d je sloupcový vektor integračních konstant. Celkem

$$z(x) = Y(x) \left(\int_{x_0}^{x} Y(t)^{-1} \cdot b(t)\, dt + d \right).$$

Zvolíme-li $d = Y(x_0)^{-1} y^0$, je splněna počáteční podmínka $z(x_0) = y^0$. $\square$

Úlohy

1. Uveďte příklad lineárně nezávislé n-tice vektorových funkcí, jejichž wronskián je identicky nulový.

13. přednáška 7. ledna 2008

FSŘ lineární diferenciální rovnice s konstantními koeficienty. Je to rovnice

$$R(y) = a_n y^{(n)} + \cdots + a_1 y' + a_0 y = 0,$$

kde $a_i \in \mathbf{R}$ jsou konstanty, $a_n \neq 0$ a $y = y(x)$ je neznámá funkce s definičním intervalem $I = \mathbf{R}$. *Charakteristický polynom* této rovnice je

$$p(x) = a_n x^n + \cdots + a_1 x + a_0.$$

Jako $K(p) = \{\lambda \in \mathbf{C} \mid p(\lambda) = 0\}$ označíme množinu jeho kořenů a pro kořen λ symbolem $n(\lambda) \in \mathbf{N}$ označíme jeho násobnost. Definujeme dvě množiny funkcí:

$$\mathcal{F}(R, \mathbf{C}) = \{x^k \mathrm{e}^{\lambda x} \mid \lambda \in K(p), 0 \leq k < n(\lambda)\}$$

a

$$\begin{aligned}
&\mathcal{F}(R, \mathbf{R}) = \{x^k \mathrm{e}^{\lambda x} \mid \lambda \in K(p) \cap \mathbf{R}, 0 \leq k < n(\lambda)\} \\
&\cup \{x^k \mathrm{e}^{\lambda x} \sin(\mu x) \mid \lambda + \mu i \in K(p), \lambda, \mu \in \mathbf{R}, \mu > 0, 0 \leq k < n(\lambda + \mu i)\} \\
&\cup \{x^k \mathrm{e}^{\lambda x} \cos(\mu x) \mid \mathrm{dtto}\}.
\end{aligned}$$

Funkce v $\mathcal{F}(R, \mathbf{C})$ obsahují komplexní exponenciálu a jsou to obecně komplexní funkce reálné proměnné. Funkce v $\mathcal{F}(R, \mathbf{R})$ jsou reálné. Seznam $\mathcal{F}(R, \mathbf{R})$ vznikl z $\mathcal{F}(R, \mathbf{C})$ náhradou dvojic komplexních funkcí

$$x^k \mathrm{e}^{(\lambda + \mu i)x},\ x^k \mathrm{e}^{(\lambda - \mu i)x}$$

(nereálné kořeny p se vyskytují ve dvojicích $\lambda + \mu i, \lambda - \mu i$ komplexně sdružených kořenů se stejnými násobnostmi—úloha 1) dvojicemi reálných funkcí

$$x^k \mathrm{e}^{\lambda x} \sin(\mu x),\ x^k \mathrm{e}^{\lambda x} \cos(\mu x).$$

V obou množinách je n různých funkcí, $|\mathcal{F}(R, \mathbf{C})| = |\mathcal{F}(R, \mathbf{R})| = n$ (úloha 2). Dokážeme, že funkce v obou množinách jsou řešení rovnice $R(y) = 0$ a tvoří lineárně nezávislé n-tice, $\mathcal{F}(R, \mathbf{C})$ a $\mathcal{F}(R, \mathbf{R})$ jsou tedy její FSŘ.

Tvrzení 3.7. *Funkce v $\mathcal{F}(R, \mathbf{C})$ a v $\mathcal{F}(R, \mathbf{R})$ jsou řešení rovnice $R(y) = 0$.*

Důkaz. Protože $(\mathrm{e}^{\lambda x})^{(m)} = \lambda^m \mathrm{e}^{\lambda x}$, pro každý kořen $\lambda \in K(p)$ (p je charakteristický polynom rovnice $R(y) = 0$) máme $R(\mathrm{e}^{\lambda x}) = \mathrm{e}^{\lambda x} p(\lambda) = 0$ a $\mathrm{e}^{\lambda x}$ je řešením. Abychom vyrobili další řešení, uvažme "derivovanou" rovnici řádu $n - 1$

$$R'(y) = n a_n y^{(n-1)} + \cdots + 2 a_2 y' + a_1 y = 0.$$

Její charakteristický polynom je $p'(x)$, derivace charakteristického polynomu původní rovnice. Podobně definujeme rovnici $R''(y) = 0$ atd. Nechť $f = f(x)$ je funkce a $R(f) = R'(f) = 0$. Díky $(xf)^{(m)} = m f^{(m-1)} + x f^{(m)}$ máme

$$R(xf) = R'(f) + x R(f) = 0.$$

Takže

$$R(f) = R'(f) = 0 \Rightarrow R(xf) = 0.$$

Má-li kořen $\lambda \in K(p)$ násobnost $m = n(\lambda)$, je $e^{\lambda x}$ řešením všech rovnic $R(y) = 0, R'(y) = 0, \ldots, R^{(m-1)}(y) = 0$, protože λ je kořenem všech jejich charakteristických polynomů $p, p', \ldots, p^{(m-1)}$. Opakovaným užitím právě dokázané implikace dostáváme, že $R(e^{\lambda x}) = R(xe^{\lambda x}) = \cdots = R(x^{m-1}e^{\lambda x}) = 0$. Tím jsme dokázali, že každá funkce v $\mathcal{F}(R, \mathbf{C})$ je řešením rovnice $R(y) = 0$.

Pro dvojici komplexně sdružených kořenů $\lambda + \mu i, \lambda - \mu i$ v $K(p)$, $\mu > 0$, si označíme funkce v odpovídajících dvojicích v $\mathcal{F}(R, \mathbf{C})$ a v $\mathcal{F}(R, \mathbf{R})$ jako

$$f_1 = x^k e^{(\lambda+\mu i)x}, f_2 = x^k e^{(\lambda-\mu i)x} \quad \text{a} \quad g_1 = x^k e^{\lambda x}\sin(\mu x), g_2 = x^k e^{\lambda x}\cos(\mu x).$$

Díky $e^{\varphi i} = \cos\varphi + i\sin\varphi$ máme

$$g_2 = \frac{f_1 + f_2}{2} \quad \text{a} \quad g_1 = \frac{f_1 - f_2}{2i}.$$

Protože je množina řešení rovnice $R(y) = 0$ uzavřená na lineární kombinace, z $R(f_1) = R(f_2) = 0$ plyne i $R(g_1) = R(g_2) = 0$. Pro reálný kořen $\lambda \in K(p)$ je odpovídající funkce v $\mathcal{F}(R, \mathbf{C})$ a v $\mathcal{F}(R, \mathbf{R})$ stejná. Dokázali jsme, že i každá funkce v $\mathcal{F}(R, \mathbf{R})$ je řešení rovnice $R(y) = 0$. □

Věta 3.8. *Množiny funkcí $\mathcal{F}(R, \mathbf{C})$ a $\mathcal{F}(R, \mathbf{R})$ jsou FSŘ rovnice $R(y) = 0$.*

Důkaz. Víme, že to jsou řešení—zbývá ukázat, že obě n-tice funkcí jsou lineárně nezávislé. Dokážeme to jen pro $\mathcal{F}(R, \mathbf{C})$. Lineární nezávislost $\mathcal{F}(R, \mathbf{R})$ pak plyne z toho, že $\mathcal{F}(R, \mathbf{R})$ vznikla z $\mathcal{F}(R, \mathbf{C})$ lineárními úpravami zachovávajícími lineární nezávislost (úloha 3). Ukážeme obecněji, že každá r-tice různých funkcí $f_1, \ldots, f_r$ z množiny

$$\mathcal{F} = \{x^k e^{\lambda x} \mid k \in \mathbf{N}_0, \lambda \in \mathbf{C}\}$$

je lineárně nezávislá nad $\mathbf{R}$.

V lineární kombinaci

$$a_1 f_1 + \ldots + a_r f_r,$$

kde $f_i \in \mathcal{F}$ jsou vzájemně různé funkce (tj. odpovídající různým dvojicím parametrů k, λ) a a_i jsou nenulové reálné (či komplexní, to je jedno) koeficienty, dáme k sobě stejné exponenciály a upravíme ji tak na tvar

$$T(x) = p_1(x)e^{\lambda_1 x} + \ldots + p_s(x)e^{\lambda_s x},$$

kde $s \geq 1$ (zřejmě $s \leq r$), λ_i jsou vzájemně různá komplexní čísla a $p_i(x)$ jsou nenulové polynomy. Dokážeme, že žádná funkce $T(x)$ tohoto typu není identicky nulová. Je to jasné pro $s = 1$, protože $p_1(x)e^{\lambda_1 x} = 0$ jen když $x \in \mathbf{C}$ je kořen polynomu $p_1(x)$. Předpokládejme pro spor, že $T(x) = 0$ pro každé $x \in \mathbf{C}$ pro nějakou funkci $T(x)$. Vezmeme takovou funkci $T_0(x)$ s nejmenší délkou s, nutně $s \geq 2$, a jako d označíme stupeň polynomu $p_1(x)$ v $T_0(x)$. Pak

$$\begin{aligned} T_1(x) &:= \left(\frac{T_0(x)}{\mathrm{e}^{\lambda_1 x}}\right)^{(d+1)} \\ &= \left(p_1(x) + p_2(x)e^{(\lambda_2-\lambda_1)x} + \ldots + p_s(x)e^{(\lambda_s-\lambda_1)x}\right)^{(d+1)} \\ &= q_2(x)e^{(\lambda_2-\lambda_1)x} + \ldots + q_s(x)e^{(\lambda_s-\lambda_1)x} \end{aligned}$$

pro nějaké nenulové polynomy $q_i(x)$. To vyplývá z rovnosti

$$\left(p(x)\mathrm{e}^{\lambda x}\right)' = (p'(x) + \lambda p(x))\mathrm{e}^{\lambda x} = q(x)\mathrm{e}^{\lambda x},$$

kde pro $\lambda \neq 0$ polynomy $p(x)$ a $q(x)$ mají stejný stupeň. Takže $T_1(x)$ je kratší funkce daného typu. Ale také $T_1(x) = 0$ pro každé $x \in \mathbf{C}$. Máme spor s minimalitou s. $\square$

Úlohy

1. Proč se nereálné kořeny charakteristického polynomu vyskytují v komplexně sdružených dvojicích se stejnými násobnostmi?
2. Ukažte, že opravdu $|\mathcal{F}(R,\mathbf{C})| = |\mathcal{F}(R,\mathbf{R})| = n$, to jest nestane se, aby dvě různé dvojice resp. trojice parametrů dávaly stejnou funkci.
3. Ukažte podrobně, jak z lineární nezávislosti $\mathcal{F}(R,\mathbf{C})$ plyne lineární nezávislost $\mathcal{F}(R,\mathbf{R})$.
4. Popište komplexní a reálný FSŘ rovnice $y^{(4)} + 2y''' - y'' + 2y' + y = 0$.