Přednáška 1, 20. února 2015

Nejprve motivace — souvislost primitivních funkcí s plochami rovinných útvarů. Funkce F je *primitivní* k funkci f, když na daném definičním oboru platí vztah $F' = f$. Pro nezápornou a spojitou funkci $f:\ [a,b] \to \mathbb{R}$ uvažme rovinný útvar

$$U(a,b,f) = \{(x,y) \in \mathbb{R}^2 \mid a \le x \le b \ \&\ 0 \le y \le f(x)\}\ .$$

Jeho plochu, ať je to cokoli, označíme jako

$$\int_a^b f := \mathrm{plocha}(U(a,b,f))\ .$$

Jde tedy o plochu části roviny vymezené osou x, grafem funkce f a svislými přímkami $y = a$ a $y = b$. Dva základní vztahy mezi plochou a derivací jsou následující. *První základní věta analýzy* říká, že pro každé $c \in [a,b]$ máme

$$\left(\int_a^x f\right)'(c) = f(c)$$

— derivace funkce, jejíž argument x je horní mez útvaru $U(a,x,f)$ a hodnota je jeho plocha, se rovná výchozí funkci f. Plocha $F(x) = \int_a^x f$ jako funkce je tedy primitivní funkcí k f. Podle *druhé základní věty analýzy* pro každou funkci g, která je na $[a,b]$ primitivní k f, je

$$\int_a^b f = g(b) - g(a)\ .$$

Známe-li nějakou funkci primitivní k f, a spoustu jich lze odvodit prostým obrácením pravidel pro derivování elementárních funkcí, můžeme ihned spočítat plochu útvaru $U(a,b,f)$. Obě věty přesně zformulujeme a dokážeme v části přednášky o Riemannově integrálu, kdy také přesně zavedeme pojem plochy $\int_a^b f$. Nejprve se ale musíme zabývat vlastnostmi primitivních funkcí — kdy má funkce primitivní funkci, zda je jednoznačně určena a podobně.

Funkce primitivní k dané funkci

Definice (primitivní funkce). *Pokud $I \subset \mathbb{R}$ je neprázdný interval a dvě funkce*

$$F, f : I \to \mathbb{R}$$

splňují na I vztah $F' = f$, pak F nazveme funkcí primitivní k funkci f (na intervalu I). *(V krajních bodech $x \in I$ se $F'(x)$ rozumí příslušná jednostranná derivace.)*

U limitění a derivování je výsledek operace jednoznačný, pokud existuje, ale primitivní funkce určená jednoznačně není. Hned uvidíme, že daná funkce buď nemá žádnou primitivní funkci nebo jich má nekonečně mnoho. Vzhledem k linearitě derivování je i operace nalezení primitivní funkce lineární: Je-li F na I primitivní k f, G ke g a $\alpha, \beta \in \mathbb{R}$, potom je funkce

$$\alpha F + \beta G$$

na I primitivní k funkci $\alpha f + \beta g$.

Tvrzení (o nejednoznačnosti primitivní funkce). *Je-li F na intervalu I primitivní k f, potom množina všech funkcí primitivních k f na I je*

$$\{F + c \mid c \in \mathbb{R}\} .$$

Všechny funkce primitivní k f se tedy dostanou posunem libovolné z nich o konstantu.

Důkaz. Derivace konstantní funkce je nulová, a tak $(F + c)' = F' + 0 = f$ pro každé $c \in \mathbb{R}$ a každou funkci F primitivní k f na I. Na druhou stranu, jsou-li F a G primitivní k f na I, pak jejich rozdíl $H = F - G$ má na I nulovou derivaci: pro každé $\gamma \in (a, b)$ je $H'(\gamma) = F'(\gamma) - G'(\gamma) = f(\gamma) - f(\gamma) = 0$. Pro libovolné dva body $\alpha < \beta$ z I tak podle Lagrangeovy věty o střední hodnotě máme rovnost

$$H(\beta) - H(\alpha) = (\beta - \alpha)H'(\gamma) = (\beta - \alpha)0 = 0 ,$$

pro nějaký bod $\gamma \in (\alpha, \beta)$, takže $H(\alpha) = H(\beta)$ a H je na I konstantní. Tedy existuje konstanta c, že $F(x) - G(x) = c$ pro každé $x \in I$ a $F = G + c$. $\square$

Tvrzení (spojitost primitivní funkce). *Je-li F na intervalu I primitivní k f, potom je F na I spojitá.*

Důkaz. Ze ZS víme, že existence vlastní (jednostranné, jde-li o krajní bod) derivace funkce v bodě implikuje její spojitost v daném bodě. Protože $F'(\alpha)$ existuje a rovná se $f(\alpha)$ pro každé $\alpha \in I$, je F na I spojitá. □

Věta (spojitá funkce má primitivní funkci). *Je-li f na intervalu I spojitá, pak má na I primitivní funkci F.*

Důkaz. To dokážeme podrobně a přesně později. Jak bylo naznačeno v úvodu, F se dá definovat jako plocha útvaru pod grafem funkce f. Jiný způsob, jak větu dokázat, je vyjádřit f jako $f = \lim f_n$ (stejnoměrná limita), kde f_n jsou lomené čáry s primitivními funkcemi F_n, a položit $F = \lim F_n$. □

Může mít nespojitá funkce primitivní funkci? Může.

Příklad (nespojitá funkce s primitivní funkcí). *Funkce $f : \mathbb{R} \to \mathbb{R}$ definovaná jako*

$$f(x) = 2x\sin(x^{-2}) - 2x^{-1}\cos(x^{-2}) \ \ \textit{pro } x \neq 0,\ f(0) = 0\ ,$$

má na reálné ose primitivní funkci, i když není spojitá v bodě 0.

Důkaz. Uvažme funkci $F : \mathbb{R} \to \mathbb{R}$ definovanou pro $x \neq 0$ jako $F(x) = x^2\sin(x^{-2})$ a pro $x = 0$ jako $F(0) = 0$. Derivování podle vzorců pro $x \neq 0$ dává $F' = f$. V nule podle definice derivace spočteme, že

$$F'(0) = \lim_{x\to 0} \frac{F(x) - F(0)}{x - 0} = \lim_{x\to 0} x\sin(x^{-2}) = 0\ ,$$

protože $|x\sin(x^{-2})| \leq |x|$ pro každé $x \neq 0$. Tedy $F'(0)$ existuje a opět $F'(0) = f(0)$. Tudíž $F' = f$ na $\mathbb{R}$ a F je na $\mathbb{R}$ primitivní k f. Funkce f není spojitá v 0, protože je v každém okolí nuly dokonce neomezená shora i zdola — pro $x \to 0$ její graf kmitá s neomezeně vzrůstající amplitudou i frekvencí. □

V MAI jsme si dokázali větu, že funkce spojitá na intervalu na něm nabývá všech mezihodnot, to jest zobrazuje ho zase na interval. Této vlastnosti funkcí se říká i *Darbouxova vlastnost*, podle francouzského matematika Jeana-Gastona Darbouxe (1842–1917). Darboux dokázal, že funkce s primitivní funkcí mají tuto vlastnost.

Věta (funkce s primitivní funkcí má Darbouxovu vlastnost). *Je-li F na intervalu I primitivní k f, potom f na I nabývá všech mezihodnot.*

Důkaz. Vezměme nějakou mezihodnotu c: $f(x_1) < c < f(x_2)$ pro nějaké dva body $x_1 < x_2$ z I. Nalezneme $x^* \in I$, dokonce $x^* \in (x_1, x_2)$, že $f(x^*) = c$. (Pokud $f(x_1) > c > f(x_2)$, následující argument se snadno upraví náhradou minima maximem.) Funkce

$$H(x) = F(x) - cx$$

je na I spojitá, dokonce tam má vlastní derivaci

$$H'(x) = (F(x) - cx)' = f(x) - c\ .$$

Podle věty ze ZS nabývá na kompaktním intervalu $[x_1, x_2]$ minimum v bodě $x^* \in [x_1, x_2]$. Protože $H'(x_1) = f(x_1) - c < 0$, je H klesající v bodě x_1, což dává, že pro nějaké $\delta > 0$ máme $x \in (x_1, x_1 + \delta) \Rightarrow H(x) < H(x_1)$. Tudíž $x^* \neq x_1$. Obdobně z $H'(x_2) > 0$ plyne, že $x^* \neq x_2$. Tedy $x^* \in (x_1, x_2)$ a podle kritéria extrému ze ZS musí být $H'(x^*) = f(x^*) - c = 0$. Tedy $f(x^*) = c$. □

Důsledek (příklad funkce bez primitivní funkce). *Funkce* $\mathrm{sgn} : \mathbb{R} \to \mathbb{R}$*, definovaná jako* $\mathrm{sgn}(x) = -1$ *pro* $x < 0$*,* $\mathrm{sgn}(0) = 0$ *a* $\mathrm{sgn}(x) = 1$ *pro* $x > 0$*, nemá na* $\mathbb{R}$ *(ani na žádném jiném intervalu obsahujícím* 0*) primitivní funkci.*

Důkaz. Podle Darbouxovy věty, protože sgn nenabývá všech mezihodnot: i když nabývá hodnotu -1 a 1, nenabývá nikde třeba hodnotu $\frac{1}{2} \in (-1, 1)$. □

Značení. Skutečnost, že funkce F je na intervalu I primitivní k funkci f, se značí jako

$$F = \int f, \quad \text{a píšeme též} \quad F = \int f + c\ ,$$

abychom zdůraznili, že každé posunutí F o konstantu je rovněž primitivní funkce k f. Symbolu $\int f$ je třeba rozumět tak, že označuje množinu všech funkcí primitivních k f na daném intervalu.

Pro derivaci součinu máme Leibnizův vzorec $(fg)' = f'g + fg'$. Invertováním obdržíme následující důležitý výsledek pro primitivní funkce.

Věta (integrace per partes). *Jsou-li* $f, g : I \to \mathbb{R}$ *spojité funkce (*I *je interval) a* F, G *jim odpovídající primitivní funkce, pak na* I *platí rovnost*

$$\int fG + \int Fg = FG + c\ .$$

Podrobněji řečeno, funkce fG a Fg mají na I primitivní funkce, jejichž součet se na I vždy rovná, až na aditivní konstantu c, funkci FG.

Důkaz. Podle předpokladu je f na I spojitá a podle hořejšího tvrzení tam je i primitivní funkce G spojitá. Takže součinová funkce fG je tam též spojitá a podle hořejší zatím nedokázané věty má na I primitivní funkci (funkce) $\int fG$. Podobně máme i primitivní funkci $\int Fg$. Podle hořejší poznámky o linearitě je součet $\int fG + \int Fg$ primitivní funkcí k funkci $fG + Fg$. K té je ale primitivní i funkce FG, protože Leibnizův vzorec dává $(FG)' = fG + Fg$. Podle tvrzení o nejednoznačnosti primitivní funkce tedy pro jistou konstantu c platí, že $\int fG + \int Fg = FG + c$. □

Vzorec pro integraci per partes se uvádí obvykle v ekvivalentním tvaru

$$\int F'G = FG - \int FG' \, .$$

Pokud tedy umíme spočítat pro dané dvě funkce F a G se spojitými derivacemi ($F' = f$ a $G' = g$) primitivní funkci k FG', dostáváme podle tohoto vzorce primitivní funkci k $F'G$.

Příklad. Díky $x' = 1$ a $(\log x)' = 1/x$ na intervalu $(0, +\infty)$ máme

$$\int \log x = \int x' \log x = x \log x - \int x(\log x)' = x \log x - \int 1 = x \log x - x \, .$$

Přesněji, $\int \log x = x \log x - x + c$ (na $(0, +\infty)$). Zderivováním snadno zkontrolujeme správnost odvozeného vzorce.

Přednáška 2, 27. února 2015

Obrácením vzorců pro derivace elementárních funkcí dostaneme tabulku základních primitivních funkcí (vynecháváme integrační konstantu c).

funkce	její primitivní funkce	na intervalu
$x^\alpha,\ \alpha \in \mathbb{R}\backslash\{-1\}$	$\dfrac{x^{\alpha+1}}{\alpha+1}$	$(0, +\infty)$
$x^\alpha,\ \alpha \in \mathbb{Z},\ \alpha < -1$	$\dfrac{x^{\alpha+1}}{\alpha+1}$	$(0, +\infty)$ i $(-\infty, 0)$
$x^\alpha,\ \alpha \in \mathbb{Z},\ \alpha > -1$	$\dfrac{x^{\alpha+1}}{\alpha+1}$	$\mathbb{R}$
$x^{-1} = \dfrac{1}{x}$	$\log\lvert x\rvert$	$(0, +\infty)$ i $(-\infty, 0)$
$\exp x = \mathrm{e}^x$	$\exp x = \mathrm{e}^x$	$\mathbb{R}$
$\sin x$	$-\cos x$	$\mathbb{R}$
$\cos x$	$\sin x$	$\mathbb{R}$
$\dfrac{1}{\cos^2 x}$	$\tan x = \dfrac{\sin x}{\cos x}$	$((k - \frac{1}{2})\pi, (k + \frac{1}{2})\pi),\ k \in \mathbb{Z}$
$\dfrac{1}{\sin^2 x}$	$-\cot x = -\dfrac{\cos x}{\sin x}$	$(k\pi, (k+1)\pi),\ k \in \mathbb{Z}$

funkce	její primitivní funkce	na intervalu
$\frac{1}{1+x^2}$	$\arctan x = (\tan x)^{\langle -1 \rangle}$ i $-\mathrm{arccot}\, x$	$\mathbb{R}$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x = (\sin x)^{\langle -1 \rangle}$ i $-\arccos x$	$(-1, 1)$

Tabulka nezahrnuje hyperbolické funkce (např. $\sinh x = \frac{\exp x - \exp(-x)}{2}$) ani další goniometrické funkce (např. sekans $\sec x = \frac{1}{\cos x}$, oblíbený v USA).

Úloha. *Formálním zderivováním máme* $(\log x)' = 1/x$*, ale i* $(\log(-x))' = (1/-x)(-1) = 1/x$*. Ale* $\log x$ *a* $\log(-x)$ *se neliší jen posunem o konstantu, takže funkce* $1/x$ *má dvě podstatně odlišné primitivní funkce, v rozporu s příslušným tvrzením. Jak je to možné?*

Invertování pravidla pro derivaci součinu dává vzorec pro integraci per partes a invertováním pravidla pro derivaci složené funkce dostaneme vzorec pro integraci substitucí. Má dva tvary, podle směru čtení rovnosti $f(\varphi)' = f'(\varphi)\varphi'$.

Věta (integrace substitucí). *Nechť* $\varphi : (\alpha, \beta) \to (a, b)$ *a* $f : (a, b) \to \mathbb{R}$ *jsou dvě funkce, přičemž existuje vlastní* φ' *na* (α, β)*.*

1. *Když* $F = \int f$ *na* (a, b)*, potom* $\int f(\varphi)\varphi' = F(\varphi) + c$ *na* (α, β)*.*
2. *Předpokládejme o* φ *navíc, že* $\varphi((\alpha, \beta)) = (a, b)$ *a* $\varphi' \neq 0$ *na* (α, β)*. Když* $G = \int f(\varphi)\varphi'$ *na* (α, β)*, potom* $\int f = G(\varphi^{\langle -1 \rangle}) + c$ *na* (a, b)*.*

Důkaz. První část plyne ihned zderivováním:

$$F(\varphi)' = F'(\varphi)\varphi' = f(\varphi)\varphi'$$

na (α, β), podle předpokladu o F a derivace složené funkce. Ve druhé části předpoklady o φ zaručují, že to je ostře rostoucí nebo ostře klesající bijekce z (α, β) na (a, b). Skutečně, musí být $\varphi' > 0$ nebo $\varphi' < 0$ na (α, β), jinak by podle věty z minulé přednášky musela funkce φ' nabýt mezihodnotu 0.

Podle výsledků ze ZS tedy φ na (α,β) ostře roste nebo ostře klesá. Takže to je prostá funkce a existuje k ní inverzní funkce

$$\varphi^{\langle -1\rangle}:\ (a,b)\to(\alpha,\beta)\ ,$$

již lze navíc derivovat podle pravidla pro derivaci inverzní funkce. Zderivování dává, podle předpokladu o G, derivace složené funce a derivace inverzní funkce, že $G(\varphi^{\langle -1\rangle})$ je na (a,b) primitivní k f:

$$G(\varphi^{\langle -1\rangle})' = G'(\varphi^{\langle -1\rangle})\cdot(\varphi^{\langle -1\rangle})' = f(\varphi(\varphi^{\langle -1\rangle}))\varphi'(\varphi^{\langle -1\rangle})\cdot\frac{1}{\varphi'(\varphi^{\langle -1\rangle})} = f\ .$$

□

Uvedeme si dva příklady, na oba tvary substitučního pravidla. **1.** Když $F(x)=\int f(x)\ dx$ na nějakém intervalu I a $a,b\in\mathbb{R}$, $a\neq 0$, pak podle prvního tvaru spočítáme, že

$$\int f(ax+b)\ dx = a^{-1}\int f(ax+b)\cdot(ax+b)'\ dx = a^{-1}F(ax+b)+c\ ,$$

na intervalu $J=a^{-1}(I-b)=\{a^{-1}(x-b)\mid x\in I\}$. Snadno se to zpětně ověří zderivováním. Vzali jsme $\varphi(x)=ax+b:\ J\to I$.

2. Chceme spočítat primitivní funkci k $\sqrt{1-t^2}$ na $(-1,1)$. Protože to připomíná poslední položku v hořejší tabulce, zkusíme substituci $t=\sin x$, tj. $t=\varphi(x)=\sin x:\ (-\frac{\pi}{2},\frac{\pi}{2})\to(-1,1)$. Předpoklady druhého tvaru substitučního pravidla jsou splněné a $\int\sqrt{1-t^2}\ dt$ nalezneme, když spočteme, na intervalu $(-\frac{\pi}{2},\frac{\pi}{2})$,

$$G(x)=\int\sqrt{1-\sin^2 x}\cdot(\sin x)'\ dx=\int\sqrt{\cos^2 x}\cdot\cos x\ dx=\int\cos^2 x\ dx\ .$$

Pomohli jsme si? Pomohli, protože poslední primitivní funkci snadno spočteme integrací per partes:

$$\begin{aligned}\int\cos^2 x &= \int\cos x(\sin x)' = \cos x\sin x+\int\sin^2 x\\ &= \cos x\sin x+\int(1-\cos^2 x)\\ &= \sin x\cos x+x-\int\cos^2 x\ ,\end{aligned}$$

takže

$$G(x) = \int \cos^2 x = \frac{\sin x \cos x + x}{2} + c = \frac{\sin x \sqrt{1 - \sin^2 x} + x}{2} + c\,.$$

Po dosazení $x = \varphi^{\langle -1 \rangle}(t) = \arcsin t$ dostáváme kýžený výsledek

$$\int \sqrt{1 - t^2} = G(\arcsin t) + c = \frac{t\sqrt{1 - t^2} + \arcsin t}{2} + c, \quad \text{na } (-1, 1)\,.$$

Derivační kontrolou se snadno ujistíme o jeho správnosti.

Obrat, že funkci f *lze vyjádřit pomocí elementárních funkcí* znamená, že f lze vyjádřit ze základních funkcí $\alpha \in \mathbb{R}$ (konstantní funkce), x (identická funkce), $\exp(x)$ (exponenciála), $\log x$, $\sin x$, $\arcsin x$, $\cos x$, $\arccos x$, $\tan x$ a $\arctan x$ opakovaným použitím aritmetických operací $+, -, \times, :$ a operace skládání funkcí. Mnohé primitivní funkce „jdou spočítat“, to jest lze je takto vyjádřit, ale stejně tak mnohé primitivní funkce spočítat nelze. Následující větu dokazovat nebudeme.

Věta (ne vše lze spočítat). *Například primitivní funkce*

$$F_1(x) = \int \exp(x^2),\ F_2(x) = \int \frac{\sin x}{x} \quad a \quad F_3(x) = \int \frac{1}{\log x}$$

(první dvě jsou na $\mathbb{R}$*, poslední na* $(0, +\infty)$*) nelze vyjádřit pomocí elementárních funkcí.*

Relativně širokou třídou funkcí, k nimž lze primitivní funkce spočítat, jsou *racionální funkce*, což jsou podíly polynomů. Uvedeme to jednoduchým příkladem. Nechť $I \subset \mathbb{R}$ je libovolný otevřený interval neobsahující ani -1 ani 1. Pak na I platí, že

$$\begin{aligned}
\int \frac{x^2}{x^2 - 1} &= \int \left(1 + \frac{1}{x^2 - 1}\right) = \int \left(1 + \frac{1/2}{x - 1} - \frac{1/2}{x + 1}\right) \\
&= \int 1 + \frac{1}{2} \int \frac{1}{x - 1} - \frac{1}{2} \int \frac{1}{x + 1} \\
&= x + \frac{\log|x - 1| - \log|x + 1|}{2} + c \\
&= x + \log(\sqrt{|(x - 1)/(x + 1)|}) + c\,.
\end{aligned}$$

Ukazuje se, že podobně lze spočítat primitivní funkci k libovolné racionální funkci. Klíčový je zřejmě rozklad na součet jednodušších racionálních funkcí na prvním řádku výpočtu, jemuž se říká *rozklad na parciální zlomky.* V důkazu následující věty použijeme některé výsledky z algebry, které zde nebudeme dokazovat. (A ani jsem tento důkaz na přednášce neuváděl pro časovou náročnost.)

Věta (primitivní funkci k racionální funkci lze vždy spočítat). *Nechť $P(x)$ a $Q(x) \neq 0$ jsou polynomy s reálnými koeficienty a $I \subset \mathbb{R}$ je otevřený interval neobsahující žádný z kořenů polynomu $Q(x)$. Primitivní funkci*

$$F(x) = \int \frac{P(x)}{Q(x)} \quad \text{(na } I\text{)}$$

lze vyjádřit pomocí elementárních funkcí, konkrétně pomocí racionálních funkcí, logaritmů a arcustangent.

Důkaz. Búno je $Q(x)$ monický (má vedoucí koeficient rovný 1). Po vydělení $P(x)$ polynomem $Q(x)$ se zbytkem máme rozklad

$$\frac{P(x)}{Q(x)} = p(x) + \frac{R(x)}{Q(x)} \ ,$$

kde $p(x), R(x)$ jsou reálné polynomy a $R(x)$ má stupeň menší než $Q(x)$. Algebra nám dává jednoznačný rozklad $Q(x)$ na součin ireducibilních (tj. dále součinově nerozložitelných) reálných polynomů:

$$Q(x) = \prod_{i=1}^{k} (x - \alpha_i)^{m_i} \prod_{i=1}^{l} (x^2 + \beta_i x + \gamma_i)^{n_i} \ ,$$

kde $k, l \geq 0$ jsou celá čísla, $\alpha_i, \beta_i, \gamma_i \in \mathbb{R}$, $m_i, n_i \geq 1$ jsou celá čísla, čísla α_i jsou vzájemně různá, dvojice (β_i, γ_i) jsou vzájemně různé a vždy $\beta_i^2 - 4\gamma_i < 0$ (což znamená, že $x^2 + \beta_i x + \gamma_i$ nemá reálné kořeny a je opravdu ireducibilní). V algebře se dále dá dokázat, že $R(x)/Q(x)$ pak má jednoznačné vyjádření jako součet parciálních zlomků

$$\frac{R(x)}{Q(x)} = \sum_{i=1}^{k} \sum_{j=1}^{m_i} \frac{\delta_{i,j}}{(x - \alpha_i)^j} + \sum_{i=1}^{l} \sum_{j=1}^{n_i} \frac{\epsilon_{i,j} x + \theta_{i,j}}{(x^2 + \beta_i x + \gamma_i)^j} \ ,$$

kde $\delta_{i,j}, \epsilon_{i,j}, \theta_{i,j} \in \mathbb{R}$. V předešlém příkladu třeba máme $P(x) = x^2$, $Q(x) = x^2 - 1$, $p(x) = 1$, $R(x) = 1$, $k = 2$, $m_1 = m_2 = 1$, $l = 0$ (žádný kvadratický trojčlen se v rozkladu $Q(x)$ nevyskytuje), $\alpha_1 = 1$, $\alpha_2 = -1$, $\delta_{1,1} = \frac{1}{2}$ a $\delta_{2,1} = -\frac{1}{2}$. Takže primitivní funkce $\int \frac{P(x)}{Q(x)}$ se rovná součtu konečně mnoha primitivních funkcí tří typů:

$$\int p(x),\ \int \frac{\delta}{(x-\alpha)^j} \ \text{ a } \int \frac{\epsilon x + \theta}{(x^2+\beta x+\gamma)^j}\ ,$$

kde $p(x)$ je reálný polynom, $j \in \mathbb{N}$ a kromě x jsou ostatní symboly reálné konstanty, přičemž $\beta^2 - 4\gamma < 0$. Když takovéto primitivní funkce dokážeme vyjádřit elementárními funkcemi, budeme mít vyjádřenu i $\int \frac{P(x)}{Q(x)}$.

Je snadné spočítat primitivní funkce prvních dvou typů:

$$\int p(x) = \int (a_n x^n + \cdots + a_1 x + a_0) = \frac{a_n x^{n+1}}{n+1} + \cdots + \frac{a_1 x^2}{2} + a_0 x$$

na $\mathbb{R}$ a

$$\int \frac{\delta}{(x-\alpha)^j} = \frac{\delta}{(1-j)(x-\alpha)^{j-1}}\ (j \geq 2),\ \int \frac{\delta}{x-\alpha} = \delta \log|x-\alpha|$$

na $(-\infty, \alpha)$ i $(\alpha, +\infty)$ (pominuli jsme integrační konstanty). Poslední třetí typ je složitější. Máme

$$\int \frac{\epsilon x+\theta}{(x^2+\beta x+\gamma)^j} = \frac{\epsilon}{2} \int \frac{2x+\beta}{(x^2+\beta x+\gamma)^j} + (\theta - \epsilon\beta/2) \int \frac{1}{(x^2+\beta x+\gamma)^j}\ ,$$

a předposlední $\int$ je po substituci $y = x^2 + \beta x + \gamma$ druhého typu: na $\mathbb{R}$ máme ($x^2 + \beta x + \gamma$ nemá reálný kořen)

$$\int \frac{2x+\beta}{(x^2+\beta x+\gamma)^j} = \frac{1}{(1-j)(x^2+\beta x+\gamma)^{j-1}}\ (j \geq 2)$$

a

$$\int \frac{2x+\beta}{x^2+\beta x+\gamma} = \log|x^2+\beta x+\gamma| = \log(x^2+\beta x+\gamma)\ .$$

Zbývá tedy spočítat primitivní funkci $\int 1/(x^2+\beta x+\gamma)^j$. Označíme $\eta = \sqrt{\gamma - \beta^2/4}$ (nezapomeňme, že $\gamma - \beta^2/4 > 0$) a použijeme substituci $y = y(x) = x/\eta + \beta/2\eta$. Doplněním na čtverec dostaneme

$$\begin{aligned}
\int \frac{1}{(x^2+\beta x+\gamma)^j} &= \frac{1}{\eta^{2j-1}} \int \frac{1/\eta}{((x/\eta+\beta/2\eta)^2+1)^j} \\
&= \frac{1}{\eta^{2j-1}} \int \frac{y'}{((x/\eta+\beta/2\eta)^2+1)^j} \\
&= \frac{1}{\eta^{2j-1}} \int \frac{1}{(y^2+1)^j}\ .
\end{aligned}$$

Zbývá tak na $\mathbb{R}$ spočítat primitivní funkci

$$I_j = \int \frac{1}{(1+x^2)^j} \ .$$

Pro $j = 1$ je, jak podle úvodní tabulky víme, $I_1 = \arctan x$. Pro $j = 2, 3, \ldots$ vyjádříme I_j pomocí rekurence, kterou získáme integrací per partes:

$$\begin{aligned} I_j &= \int \frac{x'}{(1+x^2)^j} = \frac{x}{(1+x^2)^j} + \int \frac{2jx^2}{(1+x^2)^{j+1}} \\ &= \frac{x}{(1+x^2)^j} + 2j \int \frac{x^2+1}{(1+x^2)^{j+1}} - 2j \int \frac{1}{(1+x^2)^{j+1}} \\ &= \frac{x}{(1+x^2)^j} + 2jI_j - 2jI_{j+1} \ , \end{aligned}$$

čili

$$I_{j+1} = I_j(1 - 1/2j) + \frac{x}{2j(1+x^2)^j} \ .$$

Například

$$I_2 = \frac{\arctan x}{2} + \frac{x}{2(1+x^2)} \ \text{ a } \ I_3 = \frac{3 \arctan x}{8} + \frac{3x}{8(1+x^2)} + \frac{x}{4(1+x^2)^2} \ .$$

Celkem rekurence ukazuje, že pro každé $j = 1, 2, \ldots$ je I_j tvaru $I_j = \kappa \arctan x + r(x)$, kde κ je zlomek a $r(x)$ je racionální funkce. Tím jsme dokončili výpočet primitivní funkce třetího typu z vyjádření $R(x)/Q(x)$ součtem parciálních zlomků a vyjádření primitivní funkce $\int \frac{P(x)}{Q(x)}$ elementárními funkcemi je úplné. $\square$

Riemannův integrál

Nyní definujeme přesně pojem plochy, zejména pojem plochy útvaru $U(a, b, f)$ pod grafem funkce f. Nechť $-\infty < a < b < +\infty$ jsou dvě reálná čísla a $f : [a, b] \to \mathbb{R}$ je libovolná funkce, jež nemusí být ani spojitá ani omezená. Konečná $k+1$-tice bodů $D = (a_0, a_1, \ldots, a_k)$ z intervalu $[a, b]$ je jeho *dělením*, pokud

$$a = a_0 < a_1 < a_2 < \cdots < a_k = b \ .$$

Tyto body dělí interval $[a,b]$ na intervaly $I_i = [a_i, a_{i+1}]$. Délku intervalu označíme pomocí absolutní hodnoty: $|I_i| = a_{i+1} - a_i$ a $|[a,b]| = b - a$. Zřejmě

$$\sum_{i=0}^{k-1} |I_i| = (a_1 - a_0) + (a_2 - a_1) + \cdots + (a_k - a_{k-1}) = b - a = |[a,b]| .$$

Normou dělení λ rozumíme největší délku intervalů dělení:

$$\lambda = \lambda(D) = \max_{0 \le i \le k-1} |I_i| .$$

Dělením intervalu $[a,b]$ *s body* rozumíme dvojici (D,C), kde $D = (a_0, a_1, \ldots, a_k)$ je dělení tohoto intervalu a k-tice $C = (c_0, c_1, \ldots, c_{k-1})$ se skládá z nějakých bodů $c_i \in I_i$ (tj. $a_i \le c_i \le a_{i+1}$). *Riemannovu sumu* odpovídající funkci f a dělení s body (D,C) definujeme jako

$$R(f,D,C) = \sum_{i=0}^{k-1} |I_i| f(c_i) = \sum_{i=0}^{k-1} (a_{i+1} - a_i) f(c_i) .$$

Pokud $f \ge 0$ na $[a,b]$, je to součet ploch k obdélníků $I_i \times [0, f(c_i)]$, jejichž sjednocení aproximuje útvar $U(a,b,f)$. Tato suma je ovšem definovaná pro každou funkci f, bez ohledu na její znaménko na $[a,b]$. Následující definici zavedl Bernhard Riemann (1826–1866).

Definice (první definice Riemannova integrálu, Riemannova). *Řekneme, že funkce $f : [a,b] \to \mathbb{R}$ má na intervalu $[a,b]$ Riemannův integrál $I \in \mathbb{R}$, pokud pro každé $\varepsilon > 0$ existuje $\delta > 0$ takové, že pro každé dělení intervalu $[a,b]$ s body (D,C) platí, že*

$$\lambda(D) < \delta \Rightarrow |I - R(f,D,C)| < \varepsilon .$$

Požadujeme tedy $I \in \mathbb{R}$, nevlastní hodnoty $\pm\infty$ nejsou povoleny (lze samozřejmě zavést i nevlastní integrály, podobně jako nevlastní limity). Pokud takové číslo I existuje, píšeme

$$I = \int_a^b f(x)\, dx = \int_a^b f$$

a řekneme, že f je *riemannovsky integrovatelná* na intervalu $[a,b]$. Budeme pracovat s třídou všech riemannovsky integrovatelných funkcí

$$\mathcal{R}(a,b) := \{f \mid f \text{ je definovaná a riemannovsky integrovatelná na } [a,b]\} \ .$$

První definici Riemannova integrálu tedy můžeme shrnout vzorcem

$$\int_a^b f = \lim_{\lambda(D)\to 0} R(f,D,C) \in \mathbb{R} \ .$$

Limitu zde chápeme ve smyslu uvedeném v definici; jako symbol jsme definovali pouze limitu posloupnosti a limitu funkce v bodě.

Pro druhou, ekvivalentní, definici integrálu budeme potřebovat pár dalších pojmů. Pro funkci $f:\ [a,b]\to\mathbb{R}$ a dělení $D=(a_0,a_1,\ldots,a_k)$ intervalu $[a,b]$ definujeme *dolní*, respektive *horní, Riemannovu sumu* (budeme jim tak říkat, i když by se měly jmenovat po Darbouxovi) jako

$$s(f,D) = \sum_{i=0}^{k-1} |I_i| m_i, \ \text{ respektive } \ S(f,D) = \sum_{i=0}^{k-1} |I_i| M_i \ ,$$

kde

$$m_i = \inf_{x\in I_i} f(x) \ \text{ a } \ M_i = \sup_{x\in I_i} f(x) \ (I_i = [a_i, a_{i+1}]) \ .$$

Tyto součty jsou vždy definované, $s(f,D)\in\mathbb{R}\cup\{-\infty\}$ a $S(f,D)\in\mathbb{R}\cup\{+\infty\}$. *Dolní*, respektive *horní, Riemannův integrál* funkce f na intervalu $[a,b]$ definujeme jako

$$\underline{\int_a^b} f = \underline{\int_a^b} f(x)\, dx = \sup(\{s(f,D):\ D \text{ je dělení } [a,b]\}) \ ,$$

respektive

$$\overline{\int_a^b} f = \overline{\int_a^b} f(x)\, dx = \inf(\{S(f,D):\ D \text{ je dělení } [a,b]\}) \ .$$

Tyto výrazy jsou opět vždy definované a máme $\underline{\int_a^b} f, \overline{\int_a^b} f \in \mathbb{R}^* = \mathbb{R}\cup\{-\infty,+\infty\}$.

Definice (druhá definice Riemannova integrálu, Darbouxova). *Řekneme, že funkce $f: [a,b] \to \mathbb{R}$ má na intervalu $[a,b]$ Riemannův integrál, pokud*

$$\underline{\int_a^b} f(x)\, dx = \overline{\int_a^b} f(x)\, dx \in \mathbb{R}\,.$$

Tuto společnou konečnou hodnotu, když existuje, značíme

$$\int_a^b f(x)\, dx = \int_a^b f$$

a nazýváme Riemannovým integrálem funkce f na intervalu $[a,b]$.

Za chvíli dokážeme, že vždy

$$\underline{\int_a^b} f \le \overline{\int_a^b} f\,.$$

Dokážeme také, že obě definice jsou ekvivalentní: dávají stejné třídy riemannovsky integrovatelných funkcí a dávají stejnou hodnotu Riemannova integrálu, je-li definován.

Tvrzení (neomezené funkce nemají integrál). *Když funkce $f: [a,b] \to \mathbb{R}$ není omezená, potom nemá na $[a,b]$ Riemannův integrál, ani podle jedné z obou definic.*

Důkaz. Dokažte si to jako cvičení. $\square$

Přednáška 3, 6. března 2015

Joseph Liouville (1809–1882) jako první dokázal existenci transcendentních čísel a také našel kritéria pro neelementárnost primitivních funkcí.

Věta (Liouville, 1835). *Nechť $f, g \in \mathbb{R}(x)$ jsou racionální funkce (podíly polynomů). Pak se primitivní funkce*

$$\int fe^g$$

(na nějakém intervalu neobsahujícím zádný kořen jmenovatelů funcí f a g) dá vyjádřit elementárními funkcemi, právě když existuje taková racionální funkce $a \in \mathbb{R}(x)$, že $f = a' + ag'$.

Větu dokazovat nebudeme. Podstatou je implikace $\Rightarrow$, implikace $\Leftarrow$ je triviální, protože $(ae^g)' = (a' + ag')e^g$.

Příklad. Dokážeme podle L. věty neelementárnost primitivní funkce

$$\int e^{x^2},\ x \in \mathbb{R}\ .$$

Zde $f = 1$ a $g = x^2$. Primitivní funkce by byla elementární, pouze když $f = a' + ag'$ pro nějakou rac. funkci a, tedy

$$1 = a' + 2xa\ .$$

Ukažme, že se tato rovnice nedá splnit žádnou rac. funkcí $a = p/q$, kde $p, q \in \mathbb{R}[x]$, $q \neq 0$, jsou nesoudělné polynomy (nemají společný kořen). Jistě $p \neq 0$. Když je q konstantní, je $a = p$ polynom. Pak vlevo $\deg 1 = 0$, ale vpravo $\deg(a' + 2xa) = 1 + \deg a \geq 1$, spor. Když q není konstantní, vezmeme nějaký jeho kořen $\alpha \in \mathbb{C}$, s násobností $m \in \mathbb{N}$. Tedy $q(x) = (x - \alpha)^m r(x)$, kde $r(\alpha) \neq 0$, rovněž $p(\alpha) \neq 0$. Rovnici s $a = p/q$ přepíšeme ekvivalentně jako

$$0 = -q^2 + p'q - pq' + 2xpq\ .$$

V polynomech $-q^2$, $p'q$, $-pq'$ a $2xpq$ má α jako kořen po řadě násobnost $2m$, $\geq m$, $m - 1$ a $\geq m$ (pro $\alpha = 0$ je násobnost $m + 1$). Minimum násobností

$m-1$ se nabývá jednoznačně, pro jediný ze čtyř polynomů, takže se nemohou součtem zrušit a sečíst na 0. Rovnost je nemožná a máme opět spor. $\square$

Vrátíme se k Riemannově integrálu. Když $D = (a_0, a_1, \ldots, a_k)$ a $D' = (b_0, b_1, \ldots, b_l)$ jsou dělení intervalu $[a, b]$ a $D \subset D'$, to jest pro každé $i = 0, 1, \ldots, k$ existuje j, že $a_i = b_j$ (tudíž $k \leq l$), řekneme, že D' *je zjemnění* D nebo že D' *zjemňuje* D.

Lemma. *Když* $f : [a, b] \to \mathbb{R}$ *a* D, D' *jsou dvě dělení* $[a, b]$*, přičemž* D' *zjemňuje* D*, pak*

$$s(f, D') \geq s(f, D) \quad a \quad S(f, D') \leq S(f, D)\ .$$

Důkaz. Uvážíme-li definici sum $s(f, D)$ a $S(f, D)$ a to, že D' lze vytvořit z D postupným přidáváním bodů, vidíme, že obě nerovnosti stačí dokázat v situaci, kdy $D = (a_0 = a < a_1 = b)$ a $D' = (a'_0 = a < a'_1 < a'_2 = b)$. Podle definice infim hodnot funkce f máme

$$m_0 = \inf_{a_0 \leq x \leq a_1} f(x) \leq \inf_{a'_0 \leq x \leq a'_1} f(x) = m'_0 \ \text{ a } \ m_0 \leq \inf_{a'_1 \leq x \leq a'_2} f(x) = m'_1\ .$$

Tedy

$$\begin{aligned} s(f, D') &= (a'_1 - a'_0)m'_0 + (a'_2 - a'_1)m'_1 \\ &\geq (a'_1 - a'_0)m_0 + (a'_2 - a'_1)m_0 \\ &= (a'_2 - a'_0)m_0 = (b - a)m_0 \\ &= s(f, D)\ . \end{aligned}$$

Nerovnost $S(f, D') \leq S(f, D)$ se dokáže podobně. $\square$

Důsledek. *Když* $f : [a, b] \to \mathbb{R}$ *a* D, D' *jsou dvě dělení* $[a, b]$*, pak*

$$s(f, D) \leq S(f, D')\ .$$

Důkaz. Nechť $E = D \cup D'$ je společné zjemnění obou dělení. Podle předešlého lemmatu máme

$$s(f, D) \leq s(f, E) \leq S(f, E) \leq S(f, D')\ .$$

Přesněji, první a poslední nerovnost platí podle Lemmatu, a prostřední je triviální, z definice horní a dolní sumy. $\square$

Tvrzení (dolní integrál nepřesahuje horní). *Nechť* $f: [a,b] \to \mathbb{R}$, $m = \inf_{a \le x \le b} f(x)$, $M = \sup_{a \le x \le b} f(x)$ *a* D, D' *jsou dvě dělení intervalu* $[a,b]$*. Pak platí nerovnosti*

$$m(b-a) \le s(f,D) \le \underline{\int_a^b} f \le \overline{\int_a^b} f \le S(f,D') \le M(b-a) \,.$$

Důkaz. První a poslední nerovnost jsou speciální případy Lemmatu. Druhá a předposlední nerovnost plynou hned z definice dolního a horního integrálu jakožto suprema, respektive infima. Podle Důsledku je každý prvek množiny dolních sum, jejíž supremum je $\underline{\int_a^b} f$, menší či roven každému prvku množiny horních sum, jejíž infimum je $\overline{\int_a^b} f$. S použitím definice infima (největší dolní mez) a suprema (nejmenší horní mez) odtud vyplývá prostřední nerovnost: Pro každé dělení D je $s(f,D)$ dolní mezí druhé množiny, tedy $s(f,D) \le \overline{\int_a^b} f$, tedy $\overline{\int_a^b} f$ je horní mezí první množiny, tedy $\underline{\int_a^b} f \le \overline{\int_a^b} f$. $\square$

Tvrzení (kritérium integrovatelnosti). *Nechť* $f: [a,b] \to \mathbb{R}$*. Potom*

$$f \in \mathcal{R}(a,b) \iff \forall \varepsilon > 0 \; \exists D:\; 0 \le S(f,D) - s(f,D) < \varepsilon.$$

Jinými slovy, f *má Riemannův integrál, právě když pro každé* $\varepsilon > 0$ *je pro nějaké dělení* D *intervalu* $[a,b]$ *odpovídající horní suma o méně než* ε *větší než odpovídající dolní suma.*

Důkaz. Implikace $\Rightarrow$. Předpokládáme, že f má na $[a,b]$ R. integrál, tedy $\underline{\int_a^b} f = \overline{\int_a^b} f = \int_a^b f \in \mathbb{R}$. Nechť je dáno $\varepsilon > 0$. Podle definice dolního a horního integrálu existují dělení E_1 a E_2 tak, že

$$s(f,E_1) > \underline{\int_a^b} f - \frac{\varepsilon}{2} = \int_a^b f - \frac{\varepsilon}{2} \;\text{ a }\; S(f,E_2) < \overline{\int_a^b} f + \frac{\varepsilon}{2} = \int_a^b f + \frac{\varepsilon}{2} \,.$$

Podle lemmatu tyto nerovnosti platí i po náhradě E_1 a E_2 jejich společným zjemněním $D = E_1 \cup E_2$. Sečtením obou nerovností dostaneme

$$S(f,D) - s(f,D) < \int_a^b f + \frac{\varepsilon}{2} + \left(- \int_a^b f + \frac{\varepsilon}{2} \right) = \varepsilon \,.$$

Implikace $\Leftarrow$. Předpokládáme platnost uvedené podmínky s ε a D. Pro dané $\varepsilon > 0$ vezmeme odpovídající dělení D a podle definice dolního a horního integrálu dostaneme

$$\overline{\int_a^b} f \le S(f, D) < s(f, D) + \varepsilon \le \underline{\int_a^b} f + \varepsilon, \text{ tedy } \overline{\int_a^b} f - \underline{\int_a^b} f < \varepsilon \; .$$

Tato nerovnost platí pro každé $\varepsilon > 0$, a tak podle předchozího tvrzení máme $\overline{\int_a^b} f = \underline{\int_a^b} f \in \mathbb{R}$. Tedy f má na $[a, b]$ R. integrál. $\square$

Příklad (omezená funkce bez integrálu). *Funkce* $f : [0, 1] \to \{0, 1\}$ *definovaná jako* $f(\alpha) = 1$*, když* α *je racionální číslo, a* $f(\alpha) = 0$*, když* α *je iracionální, jíž se říká* Dirichletova funkce*, nemá na* $[0, 1]$ *R. integrál, i když je omezená.*

To je jasné, každý interval (s kladnou délkou) obsahuje body, kde má f hodnotu 0 a rovněž i body, kde má hodnotu 1. Tedy $s(f, D) = 0$ a $S(f, D) = 1$ pro každé dělení D a

$$\underline{\int_0^1} f = 0 < \overline{\int_0^1} f = 1 \; .$$

Jako druhý příklad spočteme podle definice, že

$$\int_0^1 x^2 \; dx = 1/3 \; .$$

Pro $n = 1, 2, \ldots$ vezmeme dělení $D_n = (0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1)$. Pak

$$s(f, D_n) = \sum_{i=1}^{n} \frac{1}{n} \left(\frac{i-1}{n} \right)^2 = n^{-3}(1^2 + 2^2 + \cdots + (n-1)^2)$$

a podobně

$$S(f, D_n) = \sum_{i=1}^{n} \frac{1}{n} \left(\frac{i}{n} \right)^2 = n^{-3}(1^2 + 2^2 + \cdots + n^2) \; .$$

Tedy $S(f, D_n) - s(f, D_n) = \frac{1}{n} \to 0$ pro $n \to \infty$ a $f(x) = x^2$ má na $[0, 1]$ R. integrál podle předchozího kritéria. Stačí ukázat, že pro $n \to \infty$ je

$$S_n := 1^2 + 2^2 + \cdots + n^2 = n^3/3 + O(n^2) \; .$$

Protože $1 + 2 + \cdots + n \leq n^2$, z

$$(n+1)^3 - 1^3 = \sum_{i=1}^{n}((i+1)^3 - i^3) = \sum_{i=1}^{n}(3i^2 + 3i + 1) = 3S_n + 3\sum_{i=1}^{n} i + n$$

opravdu máme, že

$$S_n = \frac{n^3 + 3n^2 + 2n}{3} - \sum_{i=1}^{n} i = n^3/3 + O(n^2) \; .$$

Henri Lebesgue (1877–1941) dokázal větu charakterizující funkce s Riemannovým integrálem, kterou si uvedeme bez důkazu. Nejprve ale zavedeme množiny míry nula. Množina $M \subset \mathbb{R}$ má *nulovou (Lebesgueovu) míru*, když pro každé $\varepsilon > 0$ existuje posloupnost intervalů $I_1, I_2, \ldots$ taková, že

$$\sum_{i=1}^{\infty} |I_i| < \varepsilon \;\text{ a }\; M \subset \bigcup_{i=1}^{\infty} I_i \; .$$

Definice říká, že se M dá pokrýt intervaly libovolně malé celkové délky. Uveďme základní vlastnosti množin reálných čísel s nulovou mírou. Důkazy si rozmyslete jako cvičení.

- Každá konečná nebo spočetná množina má nulovou míru.
- Podmnožina množiny s nulovou mírou má také nulovou míru.
- Má-li každá z množin $A_1, A_2, \ldots$ nulovou míru, má i jejich sjednocení

  $$\bigcup_{n=1}^{\infty} A_n$$

  nulovou míru.
- Interval s kladnou délkou nemá nulovou míru.

Například celá množina racionálních čísel $\mathbb{Q}$ má nulovou míru.

Věta (Lebesgueova, charakterizující integrovatelné funkce) *Funkce $f : [a, b] \to \mathbb{R}$ má Riemannův integrál, právě když je omezená a množina jejích bodů nespojitosti má nulovou míru.*

Otázka posluchače. *Existuje nespočetná množina reálných čísel s nulovou mírou?*

Existuje. Klasickým příkladem je tzv. *Cantorovo diskontinuum.* Je to množina

$$X_C := \bigcap_{n=1}^{\infty} [0, \tfrac{1}{3^n}] \cup [\tfrac{2}{3^n}, \tfrac{3}{3^n}] \cup [\tfrac{4}{3^n}, \tfrac{5}{3^n}] \cup [\tfrac{6}{3^n}, \tfrac{7}{3^n}] \cup \cdots \cup [\tfrac{3^n-1}{3^n}, 1]$$

— to, co zbude z intervalu $[0,1]$ vyhodíme-li prostřední třetinu $(\frac{1}{3}, \frac{2}{3})$, pak prostřední třetinu z každé z obou zbylých krajních třetin, tj. pak vyhodíme $(\frac{1}{9}, \frac{2}{9}) \cup (\frac{7}{9}, \frac{8}{9})$, pak vyhodíme z každé ze zbylých čtyřech devítin její prostřední třetinu a tak dále do nekonečna. Jinými slovy

$$X_C = \{\alpha = \textstyle\sum_{i=1}^{\infty} a_i/3^i \mid a_i \in \{0,2\}\}$$

— X_C jsou právě ta čísla z $[0,1]$, v jejichž rozvoji při základu 3 se vyskytují pouze cifry 0 a 2 (a nikde cifra 1). Například

$$1/3 = (0.02222\ldots)_3 \in X_C \quad \text{nebo} \quad 1/4 = (0.020202\ldots)_3 \in X_C\ ,$$

ale $5/8 = (0.121212\ldots)_3 \notin X_C$. Z této korespondence s nekonečnými posloupnostmi 0 a 2 je vidět, že X_C je nespočetná. Není ani těžké ukázat, že má nulovou míru (cvičení).

Přednáška 4, 13. března 2015

Tvrzení (monotonie ⇒ integrovatelnost). *Je-li funkce $f:\ [a,b] \to \mathbb{R}$ na intervalu $[a,b]$ nerostoucí nebo neklesající, potom má Riemannův integrál.*

Důkaz. Nechť f neklesá (pro nerostoucí f se argumentuje podobně). Pro každý podinterval $[\alpha,\beta] \subset [a,b]$ pak máme $\inf_{[\alpha,\beta]} f = f(\alpha)$ a $\sup_{[\alpha,\beta]} f = f(\beta)$. Buď dáno $\varepsilon > 0$. Vezmeme libovolné dělení $D = (a_0, a_1, \ldots, a_{k-1})$ intervalu $[a,b]$ s $\lambda(D) < \varepsilon$ a máme

$$
\begin{aligned}
S(f,D) - s(f,D) &= \sum_{i=0}^{k-1}(a_{i+1}-a_i)(\sup_{I_i} f - \inf_{I_i} f) \\
&= \sum_{i=0}^{k-1}(a_{i+1}-a_i)(f(a_{i+1}) - f(a_i)) \\
&\leq \varepsilon \sum_{i=0}^{k-1}(f(a_{i+1}) - f(a_i)) \\
&= \varepsilon(f(a_k) - f(a_0)) = \varepsilon(f(b) - f(a))\ .
\end{aligned}
$$

Tuto mez lze zmenšováním ε učinit libovolně malou. Podle kritéria integrovatelnosti tedy $f \in \mathcal{R}(a,b)$. (Kontrolní otázka: proč v předešlém výpočtu nelze místo $\leq$ psát $<$?) □

I spojitost postačuje pro integrovatelnost. Musíme se však seznámit s její silnější podobou. Řekneme, že funkce $f:\ I \to \mathbb{R}$, kde I je interval, je *stejnoměrně spojitá* (na I), pokud

$$\forall \varepsilon > 0\ \exists \delta > 0:\ x, x' \in I,\ |x - x'| < \delta \Rightarrow |f(x) - f(x')| < \varepsilon\ .$$

Požaduje se tedy silněji, aby jediná mez δ fungovala pro všechny dvojice bodů x, x' z I. V obyčejné spojitosti může δ záviset na poloze x a x'. Stejnoměrná spojitost implikuje triviálně spojitost, ale naopak to obecně neplatí. Například funkce

$$f(x) = 1/x:\ I = (0,1) \to \mathbb{R}$$

je na I spojitá, ale ne stejnoměrně spojitá: $f(1/(n+1)) - f(1/n) = 1$, i když $1/(n+1) - 1/n \to 0$ pro $n \to \infty$. Na *kompaktním intervalu I*, což je interval typu $[a,b]$ s $-\infty < a \leq b < +\infty$, však naštěstí oba typy spojitosti splývají.

Tvrzení (na kompaktu: spojitost $\Rightarrow$ stejnoměrná spojitost). *Je-li funkce $f: [a,b] \to \mathbb{R}$ na intervalu $[a,b]$ spojitá, je na něm stejnoměrně spojitá.*

Důkaz. Pro spor předpokládáme, že $f: [a,b] \to \mathbb{R}$ je spojitá v každém bodě intervalu $[a,b]$ (tedy jednostraně v krajních bodech a a b), ale že není na $[a,b]$ stejnoměrně spojitá. Odvodíme spor. Negace stejnoměrné spojitosti znamená, že

$$\exists \varepsilon > 0\ \forall \delta > 0\ \exists x, x' \in I:\ |x - x'| < \delta\ \&\ |f(x) - f(x')| \ge \varepsilon\ .$$

Což znamená, že pro $\delta = 1/n$ a $n = 1, 2, \ldots$ existují body $x_n, x'_n \in [a,b]$, že $|x_n - x'_n| < 1/n$, ale $|f(x_n) - f(x'_n)| \ge \varepsilon$. Díky Bolzanově–Weierstrassově větě ze ZS můžeme bez újmy na obecnosti předpokládat, že posloupnosti (x_n) a (x'_n) obě konvergují a (nevyhnutelně) k témuž bodu α z $[a,b]$. (Podle této věty existuje posloupnost přir. čísel $k_1 < k_2 < \ldots$, že (x_{k_n}) konverguje. Opět podle této věty existuje posloupnost přir. čísel $l_1 < l_2 < \ldots$, že $(x'_{k_{l_n}})$ konverguje. Posloupnost $(x_{k_{l_n}})$ zůstává konvergentní, protože je podposloupností posloupnosti (x_{k_n}). Protože $|x_{k_{l_n}} - x'_{k_{l_n}}| < 1/k_{l_n} \le 1/n \to 0$,

$$\lim_{n\to\infty} x_{k_{l_n}} = \lim_{n\to\infty} x'_{k_{l_n}} = \alpha\ .$$

Abychom se vyhnuli vícenásobným indexům, přeznačíme $x_{k_{l_n}}$ jako x_n a $x'_{k_{l_n}}$ jako x'_n.) Podle Heineho definice limity, spojitosti f v bodě α a aritmetiky limit máme

$$0 = f(\alpha) - f(\alpha) = \lim f(x_n) - \lim f(x'_n) = \lim(f(x_n) - f(x'_n))\ .$$

Jsme ve sporu s tím, že $|f(x_n) - f(x'_n)| \ge \varepsilon$ pro každé n. $\square$

Tvrzení (spojitost $\Rightarrow$ integrovatelnost). *Je-li funkce $f: [a,b] \to \mathbb{R}$ na intervalu $[a,b]$ spojitá, potom má Riemannův integrál.*

Důkaz. Nechť f je na $[a,b]$ spojitá. Buď dáno $\varepsilon > 0$. Podle předchozího tvrzení vezmeme $\delta > 0$, že $|f(x) - f(x)'| < \varepsilon$ platí, jakmile $x, x' \in [a,b]$ jsou blíže než δ. Tedy

$$\sup_{[\alpha,\beta]} f - \inf_{[\alpha,\beta]} f \le \varepsilon$$

platí pro každý podinterval $[\alpha, \beta] \subset [a, b]$ délky menší než δ (proč?). Vezmeme jakékoli dělení $D = (a_0, a_1, \ldots, a_{k-1})$ intervalu $[a, b]$ s $\lambda(D) < \delta$ a máme

$$\begin{aligned} S(f, D) - s(f, D) &= \sum_{i=0}^{k-1} (a_{i+1} - a_i)(\sup_{I_i} f - \inf_{I_i} f) \\ &\leq \sum_{i=0}^{k-1} (a_{i+1} - a_i)\varepsilon \\ &= \varepsilon(a_k - a_0) = \varepsilon(b - a) \ . \end{aligned}$$

Tuto mez lze zmenšováním ε učinit libovolně malou. Podle kritéria integrovatelnosti tedy $f \in \mathcal{R}(a, b)$. $\square$

Že monotonie i spojitost postačují k integrovatelnosti jsme dokázali přímo, i když obojí vyplývá hned jako důsledek z Lebesgueovy věty, což si rozmyslete jako cvičení. Zmíníme její další důsledky.

Tvrzení (spojitost(integrovatelnost)=integrovatelnost). *Má-li funkce $f : [a, b] \to [c, d]$ Riemannův integrál a $g : [c, d] \to \mathbb{R}$ je na $[c, d]$ spojitá, potom má složená funkce $g(f) : [a, b] \to \mathbb{R}$ Riemannův integrál.*

Důkaz. Protože vnější funkce g je omezená, jako spojitá funkce na kompaktním intervalu, je i složená funkce $g(f)$ omezená. Je-li f spojitá v bodě α z $[a, b]$, je i složená funkce $g(f)$ spojitá v α, protože g je spojitá v $f(\alpha)$ a spojitost se skládáním zachovává, jak jsme si dokázali v ZS. Množina M bodů nespojitosti funkce $g(f)$ je tedy obsažena v množině N bodů nespojitosti funkce f. Podle předpokladu a L. věty má N nulovou míru. Takže i M má nulovou míru a podle L. věty má $g(f)$ Riemannův integrál. $\square$

Proto z $f \in \mathcal{R}(a, b)$ plyne například $f^2 \in \mathcal{R}(a, b)$ nebo $|f| \in \mathcal{R}(a, b)$. Jako cvičení si rozmyslete, proč a jak z $f, g \in \mathcal{R}(a, b)$ plyne, že i

$$fg \in \mathcal{R}(a, b) \ \text{ a } \ \max(f, g) \in \mathcal{R}(a, b) \ .$$

Nyní se podíváme na linearitu R. integrálu. Nejprve ukážeme linearitu $\int_a^b f$ jako funkce integrandu f, a pak jako funkce integračních mezí a a b.

Tvrzení (linearita $\int$ v integrandu). *Nechť $f, g \in \mathcal{R}(a, b)$ jsou dvě funkce mající R. integrál a $\alpha, \beta \in \mathbb{R}$. Potom i*

$$\alpha f + \beta g \in \mathcal{R}(a, b) \ \ a \ \ \int_a^b (\alpha f + \beta g) = \alpha \int_a^b f + \beta \int_a^b g \ .$$

Důkaz. Stačí prověřit tři speciální případy lineárních kombinací, totiž $-f$, αf s $\alpha \geq 0$ a $f+g$, ostatní se z těchto již odvodí. Buď dáno $\varepsilon > 0$. Podle kritéria integrovatelnosti existuje dělení D intervalu $[a,b]$, že

$$S(f,D) - s(f,D) < \varepsilon \;\&\; S(g,D) - s(g,D) < \varepsilon\;.$$

(Jistě máme dvě taková dělení, D_1 pro f a D_2 pro g. Přechodem ke společnému zjemnění dosáhneme, že $D_1 = D_2$.) Podle definice infima a suprema množiny reálných čísel, pro libovolný podinterval $I \subset [a,b]$ platí, že (pro $\alpha \geq 0$)

$$\inf_I(-f) = -\sup_I f,\;\; \inf_I \alpha f = \alpha \inf_I f,\;\; \inf_I(f+g) \geq \inf_I f + \inf_I g$$

a analogicky pro suprema (prohodíme inf a sup a poslední nerovnost otočíme). Podle definice dolní, popř. horní, sumy jako lineární kombinace (s >0 koeficienty) infim, popř. suprem,

$$S(-f,D) - s(-f,D) = -s(f,D) - (-S(f,D)) = S(f,D) - s(f,D) < \varepsilon\;,$$

$$S(\alpha f,D) - s(\alpha f,D) = \alpha S(f,D) - \alpha s(f,D) \leq \alpha\varepsilon\;(\alpha \geq 0)$$

a

$$\begin{aligned} S(f+g,D) - s(f+g,D) &\leq (S(f,D)+S(g,D)) - (s(f,D)+s(g,D)) \\ &= S(f,D) - s(f,D) + S(g,D) - s(g,D) \\ &< 2\varepsilon\;. \end{aligned}$$

Takže, podle kritéria integrovatelnosti, i $-f, \alpha f, f+g \in \mathcal{R}(a,b)$. Navíc, podle nerovností mezi dolními a horními sumami a integrálem, $\int_a^b f \in [s(f,D), S(f,D)]$ a totéž platí pro funkci g. Tedy $\int_a^b(-f)$ leží v intervalu

$$[s(-f,D), S(-f,D)] = [-S(f,D), -s(f,D)] \ni -\int_a^b f$$

a čísla $\int_a^b(-f)$ a $-\int_a^b f$ se tak liší o méně než ε. Tedy $\int_a^b(-f) = -\int_a^b f$. Podobně $\int_a^b \alpha f$ leží v intervalu

$$[s(\alpha f,D), S(\alpha f,D)] = [\alpha s(f,D), \alpha S(f,D)] \ni \alpha\int_a^b f$$

o délce nejvýše $\alpha\varepsilon$, a tak $\int_a^b \alpha f = \alpha \int_a^b f$. Konečně $\int_a^b (f+g)$ leží v intervalu

$$[s(f+g,D), S(f+g,D)] \subset [s(f,D)+s(g,D), S(f,D)+S(g,D)] \ni \int_a^b f + \int_a^b g$$

o délce méně než 2ε, a tak $\int_a^b (f+g) = \int_a^b f + \int_a^b g$. □

Podle předchozího tvrzení množina riemannovsky integrovatelných funkcí $\mathcal{R}(a,b)$ tvoří vektorový prostor nad tělesem $\mathbb{R}$ a $f \mapsto \int_a^b f$ je lineární zobrazení z $\mathcal{R}(a,b)$ do $\mathbb{R}$. Říkáme, že R. integrál je *lineární funkcionál.* Předchozí důkaz linearity pomocí dolních a horních sum jsem na přednášce z časových důvodů neuváděl a odvolal jsem se na ekvivalenci obou definic R. integrálu, kterou nebudeme dokazovat.

Věta (ekvivalence obou definic R. $\int$). *Obě definice Riemannova integrálu jsou ekvivalentní: pro každou funkci $f:\ [a,b] \to \mathbb{R}$*

$$\lim_{\lambda(D)\to 0} R(f,D,C) \ \textit{existuje} \quad \Longleftrightarrow \quad \underline{\int_a^b} f = \overline{\int_a^b} f \in \mathbb{R}$$

a obě hodnoty, pokud existují, se rovnají.

Jako cvičení si rozmyslete důkaz tvrzení o linearitě $\int$ v integrandu pomocí první definice R. integrálu.

Přejdeme k linearitě $\int$ jako funkce integračních mezí. Nejprve mírně rozšíříme definici $\int_a^b f$:

$$\int_a^a f := 0 \ \text{ a } \ \int_a^b f := -\int_b^a f \ \text{ pro } \ a > b\ .$$

Pro $f:\ [a,b] \to \mathbb{R}$ a podinterval $I \subset [a,b]$ označíme zúžení funkce f na I v následujícím tvrzení pro jednoduchost opět jako f.

Tvrzení (linearita $\int$ v integračních mezích). *Nechť $f:\ [a,b] \to \mathbb{R}$ je funkce a $c \in (a,b)$. Potom*

$$f \in \mathcal{R}(a,b) \iff f \in \mathcal{R}(a,c)\ \&\ f \in \mathcal{R}(c,b)$$

a, jsou-li tyto integrály definované,

$$\int_a^b f = \int_a^c f + \int_c^b f\ .$$

Důkaz. Jako $f_1 : [a,c] \to \mathbb{R}$ a $f_2 : [c,b] \to \mathbb{R}$ označíme zúžení funce f na uvedený podinterval. Nechť $f \in \mathcal{R}(a,b)$. Pro dané $\varepsilon > 0$ podle kritéria integrovatelnosti máme dělení D intervalu $[a,b]$, že $S(f,D) - s(f,D) < \varepsilon$. Můžeme předpokládat (přechodem ke zjemnění), že $c \in D$. Bod c dělí D na dělení D' intervalu $[a,c]$ a dělení D'' intervalu $[c,b]$. Protože $S(f,D) = S(f_1,D') + S(f_2,D'')$ a $s(f,D) = s(f_1,D') + s(f_2,D'')$, z

$$\varepsilon > S(f,D) - s(f,D) = (S(f_1,D') - s(f_1,D')) + (S(f_2,D'') - s(f_2,D''))$$

plyne, díky nezápornosti rozdílu horní a dolní sumy, že i $\varepsilon > S(f_1,D') - s(f_1,D')$ a $\varepsilon > S(f_2,D'') - s(f_2,D'')$. (Obecně samozřejmě z $\gamma > \alpha + \beta$ neplyne, že $\gamma > \alpha$ a $\gamma > \beta$.) Tedy, podle kritéria integrovatelnosti, $f_1 \in \mathcal{R}(a,c)$ a $f_2 \in \mathcal{R}(c,b)$. Máme $\int_a^c f_1 \in [s(f_1,D'), S(f_1,D')]$, $\int_c^b f_2 \in [s(f_2,D''), S(f_2,D'')]$ a $\int_a^b f \in [s(f,D), S(f,D)]$, z čehož plyne, že $\int_a^c f_1 + \int_c^b f_2$ a $\int_a^b f$ se liší o méně než ε. Tedy $\int_a^c f_1 + \int_c^b f_2 = \int_a^b f$.

Nechť $f_1 \in \mathcal{R}(a,c)$ a $f_2 \in \mathcal{R}(c,b)$. Pro dané $\varepsilon > 0$ podle kritéria integrovatelnosti máme dělení D' intervalu $[a,c]$ a dělení D'' intervalu $[c,b]$, že $S(f_1,D') - s(f_1,D') < \varepsilon$ a $S(f_2,D'') - s(f_2,D'') < \varepsilon$. Sečtením dostaneme

$$2\varepsilon > S(f_1,D') - s(f_1,D') + S(f_2,D'') - s(f_2,D'') = S(f,D) - s(f,D) ,$$

kde D je dělení intervalu $[a,b]$ vzniklé spojením D' a D''. Tedy, podle kritéria integrovatelnosti, i $f \in \mathcal{R}(a,b)$. □

Důsledek ($\int$ přes cyklus je 0). *Nechť $a,b,c \in \mathbb{R}$, $d = \min(a,b,c)$, $e = \max(a,b,c)$ a $f \in \mathcal{R}(d,e)$. Potom následující tři integrály existují a*

$$\int_a^b f + \int_b^c f + \int_c^a f = 0 .$$

Důkaz. Nechť například $d = a < e = c$. Podle předchozího tvrzení máme

$$\int_a^b f + \int_b^c f = \int_a^c f .$$

Celý součet pak je $\int_a^c f + \int_c^a f = \int_a^c f - \int_a^c f = 0$. Ostatní možnosti jsou podobné. □

Přednáška 5, 20. března 2015

Věta (1. základní věta analýzy). *Nechť $f \in \mathcal{R}(a,b)$ a funkce $F : [a,b] \to \mathbb{R}$ je definována předpisem*

$$F(x) = \int_a^x f \ .$$

Potom (i) F je na $[a,b]$ spojitá a (ii) v každém bodě spojitosti $x_0 \in [a,b]$ funkce f existuje vlastní $F'(x_0)$ a $F'(x_0) = f(x_0)$ (platí to jednostranně, pokud $x_0 = a$ nebo $x_0 = b$).

Důkaz. Nechť $c > 0$ je horní mez pro hodnoty $|f(x)|$, $a \le x \le b$ (f je integrovatelná a tedy omezená). Pro každé dva body $x, x_0 \in [a,b]$ máme

$$|F(x) - F(x_0)| = \left| \int_a^x f - \int_a^{x_0} f \right| = \left| \int_{x_0}^x f \right| \le |x - x_0| c \ ,$$

podle definice F, linearity $\int$ v integračních mezích a odhadu $\int$ horní sumou pro dělení (x_0, x) či (x, x_0) interválku s krajními body x a x_0. Pro $x \to x_0$ máme $F(x) \to F(x_0)$ a F je v x_0 spojitá.

Nechť $x_0 \in [a,b]$ je bod spojitosti f. Pro dané $\varepsilon > 0$ máme $\delta > 0$, že $f(x_0) - \varepsilon < f(x) < f(x_0) + \varepsilon$, jakmile $|x - x_0| < \delta$. Pro $0 < x - x_0 < \delta$ tedy

$$f(x_0) - \varepsilon \le \frac{\int_{x_0}^x f}{x - x_0} = \frac{F(x) - F(x_0)}{x - x_0} \le f(x_0) + \varepsilon \ ,$$

podle triviálního odhadu $\int_{x_0}^x f$ dolní a horní sumou pro dělení (x_0, x). Pro $-\delta < x - x_0 < 0$ platí tytéž nerovnosti (v čitateli i jmenovateli zlomku se změní znaménko). Pro $x \to x_0$, $x \ne x_0$, tak máme $\frac{F(x)-F(x_0)}{x-x_0} \to f(x_0)$, čili $F'(x_0) = f(x_0)$. $\square$

Vyřídíme rest z 1. přednášky.

Důsledek (spojitá funkce má primitivní funkci). *Je-li $f : [a,b] \to \mathbb{R}$ na $[a,b]$ spojitá, potom má na $[a,b]$ primitivní funkci F.*

Důkaz. Stačí použít předchozí větu a položit $F(x) = \int_a^x f$. $\square$

Věta (2. základní věta analýzy). *Pokud* $f \in \mathcal{R}(a,b)$ *a* $F : [a,b] \to \mathbb{R}$ *je na* $[a,b]$ *primitivní k* f, *pak*

$$\int_a^b f = F(b) - F(a) \ .$$

Důkaz. Nechť $D = (a_0, a_1, \ldots, a_k)$ je libovolné dělení intervalu $[a,b]$. Použijeme-li na každý interval $I_i = [a_i, a_{i+1}]$ a funkci F Lagrangeovu větu o střední hodnotě, dostaneme vztah

$$F(b) - F(a) = \sum_{i=0}^{k-1} (F(a_{i+1}) - F(a_i)) = \sum_{i=0}^{k-1} f(b_i)(a_{i+1} - a_i) \ ,$$

pro nějaké mezibody $a_i < b_i < a_{i+1}$ (neboť $F'(b_i) = f(b_i)$). Tedy (neboť $\inf_{I_i} f \le f(b_i) \le \sup_{I_i} f$)

$$s(f,D) \le F(b) - F(a) \le S(f,D) \ .$$

Odtud a z integrovatelnosti f ihned plyne, že $F(b) - F(a) = \int_a^b f$. □

Pro funkci $F : [a,b] \to \mathbb{R}$ se rozdíl jejích hodnot v krajních bodech často značí jako

$$F|_a^b := F(b) - F(a) \ .$$

Předchozí výsledky shrneme.

Důsledek ($\int$ pomocí primitivní funkce). *Je-li* $f : [a,b] \to \mathbb{R}$ *na* $[a,b]$ *spojitá, potom* $f \in \mathcal{R}(a,b)$, f *má na* $[a,b]$ *primitivní funkci* F *a pro tuto i všechny ostatní primitivní funkce je*

$$\int_a^b f = F|_a^b = F(b) - F(a) \ .$$

Newtonův integrál. Staletí předtím, než přišel Riemann (po jistých pokusech Cauchyho) s přesnou definicí integrálu, počítali matematici integrály bez zábran přímo z primitivních funkcí tzv. *Newtonovým integrálem.* Připomeneme ho a porovnáme s integrálem Riemannovým.[1]

[1] Zájemcům o historii a další druhy integrálů doporučujeme knihu Š. Schwabik a P. Šarmanová, *Malý průvodce historií integrálu*, Prometheus, 1996.

Nechť $f: (a,b) \to \mathbb{R}$, f má na (a,b) primitivní funkci F a ta má v krajních bodech vlastní jednostranné limity $F(a^+) = \lim_{x\to a^+} F(x)$ a $F(b^-) = \lim_{x\to b^-} F(x)$. Newtonův integrál funkce f na (a,b) pak definujeme jako

$$(N)\int_a^b f = F(b^-) - F(a^+)\ .$$

Protože různé primitivní funkce k f se liší jen posunem o konstantu, nezávisí tento rozdíl na volbě F a definice je korektní. Množinu funkcí newtonovsky integrovatelných na (a,b) označíme jako $\mathcal{N}(a,b)$. Jako $C(a,b)$ označíme množinu funkcí spojitých na $[a,b]$.

Tvrzení (porovnání Newtonova a Riemannova $\int$). *Máme*

$$C(a,b) \subset \mathcal{N}(a,b) \cap \mathcal{R}(a,b)\ .$$

Pokud $f \in \mathcal{N}(a,b) \cap \mathcal{R}(a,b)$, pak

$$(N)\int_a^b f = (R)\int_a^b f\ .$$

Množina $\mathcal{N}(a,b)\backslash\mathcal{R}(a,b)$ i $\mathcal{R}(a,b)\backslash\mathcal{N}(a,b)$ je neprázdná: existují funkce, které mají Newtonův integrál, ale ne Riemannův, i naopak.

Důkaz. Je-li f na $[a,b]$ spojitá, je (jak jsme pomocí stejnoměrné spojitosti dokázali) $f \in \mathcal{R}(a,b)$ a (podle 1. ZVA) $F(x) = \int_a^x f$ je na $[a,b]$ primitivní k f. Máme $F(a^+) = F(a) = 0$ a $F(b^-) = F(b) = \int_a^b f$, takže i $f \in \mathcal{N}(a,b)$.

Nechť $f \in \mathcal{N}(a,b) \cap \mathcal{R}(a,b)$. Protože $f \in \mathcal{N}(a,b)$, má na (a,b) primitivní funkci F a existují jednostranné vlastní limity $F(a^+)$ a $F(b^-)$. Protože $f \in \mathcal{R}(a,b)$, je pro každé $\delta > 0$ i $f \in \mathcal{R}(a+\delta, b-\delta)$ a podle 2. ZVA máme

$$(R)\int_{a+\delta}^{b-\delta} f = F(b-\delta) - F(a+\delta)\ .$$

Pro $\delta \to 0^+$ jde levá strana k $(R)\int_a^b f$ (f je na $[a,b]$ omezená, takže integrály f přes $[a, a+\delta]$ a $[b-\delta, b]$ jdou k 0) a pravá strana jde k $F(b^-) - F(a^+) = (N)\int_a^b f$.

Funkce $f(x) = x^{-1/2}: (0,1] \to \mathbb{R}$, $f(0) = 2013$, má na $(0,1)$ Newtonův integrál: $F(x) = 2x^{1/2}$ je tam k ní primitivní, $F(0^+) = 0$ a $F(1^-) = 2$, takže $(N)\int_0^1 f = 2$. Tato funkce ale není na $[0,1]$ omezená, a proto $f \notin \mathcal{R}(0,1)$.

Funkce znaménka $\mathrm{sgn}(x)$ je na $[-1,1]$ neklesající a tedy má na $[-1,1]$ Riemanův integrál. Na $(-1,1)$ ale nemá Newtonův integrál — jak jsme ukázali v 1. přednášce, nemá na $(-1,1)$ primitivní funkci. $\square$

K závěrečným příkladům poznamenejme, že nicméně

$$\lim_{\delta\to 0^+} (R)\int_\delta^1 \frac{dx}{\sqrt{x}} = 2\ ,$$

čili útvar vymezený grafem $y = 1/\sqrt{x}$ a intervalem $(0,1]$ má plochu 2, a že i když nemůžeme spočítat

$$(R)\int_{-1}^1 \mathrm{sgn}(x)\ dx$$

okamžitě přímo pomocí 2. ZVA (protože primitivní funkce na daném intervalu neexistuje), po rozkladu $[-1,1] = [-1,0] \cup [0,1]$ už můžeme počítat pomocí primitivních funkcí:

$$\begin{aligned}
(R)\int_{-1}^1 \mathrm{sgn}(x)\ dx &= (R)\int_{-1}^0 \mathrm{sgn}(x)\ dx + (R)\int_0^1 \mathrm{sgn}(x)\ dx \\
&= (R)\int_{-1}^0 (-1)\ dx + (R)\int_0^1 1\ dx \\
&= (-x)|_{-1}^0 + x|_0^1 = (-1) + 1 \\
&= 0
\end{aligned}$$

(ve výpočtu jsme změnili hodnotu funkce $\mathrm{sgn}(x)$ v $x = 0$, ale to nemá na integrovatelnost a hodnotu integrálu žádný vliv).

V dalším už budeme integrálem opět rozumět výhradně Riemannův integrál a místo $(R)\int$ psát pouze $\int$. Dvě metody výpočtu primitivní funkce, per partes a substituční, se díky 2. ZVA odrážejí i ve výpočtech R. integrálů.

Tvrzení (integrace per partes). *Nechť $f, g : [a,b] \to \mathbb{R}$ mají na $[a,b]$ spojité derivace f' a g'. Potom*

$$\int_a^b fg' = fg|_a^b - \int_a^b f'g\ .$$

Důkaz. Cvičení. $\square$

Tvrzení (integrace substitucí). *Nechť* $\varphi: [\alpha,\beta] \to [a,b]$ *a* $f: [a,b] \to \mathbb{R}$ *jsou dvě funkce, přičemž* φ *má na* $[\alpha,\beta]$ *spojitou* φ' *a* $\varphi(\alpha) = a$*,* $\varphi(\beta) = b$ *nebo* $\varphi(\alpha) = b$*,* $\varphi(\beta) = a$*.*

1. *Je-li* f *spojitá na* $[a,b]$*, pak*
$$\int_\alpha^\beta f(\varphi)\varphi' = \int_{\varphi(\alpha)}^{\varphi(\beta)} f = \begin{cases} \int_a^b f \quad \textit{nebo} \\ \\ \int_b^a f = -\int_a^b f \ . \end{cases}$$

2. *Je-li* φ *na* $[\alpha,\beta]$ *rostoucí nebo klesající a* $f \in \mathcal{R}(a,b)$*, platí opět předchozí rovnost integrálů.*

Důkaz. 1. Funkce f je na $[a,b]$ spojitá a má tam tedy primitivní funkci F. Derivace složené funkce dává na $[\alpha,\beta]$ rovnost $F(\varphi)' = f(\varphi)\varphi'$. Takže $F(\varphi)$ je na $[\alpha,\beta]$ primitivní k $f(\varphi)\varphi'$. Funkce $f(\varphi)\varphi'$ je na $[\alpha,\beta]$ spojitá (je součinem spojitých funkcí), takže $f(\varphi)\varphi' \in \mathcal{R}(\alpha,\beta)$. Dvojím užitím 2. ZVA (v první a třetí rovnosti) máme
$$\int_\alpha^\beta f(\varphi)\varphi' = F(\varphi)|_\alpha^\beta = F|_{\varphi(\alpha)}^{\varphi(\beta)} = \int_{\varphi(\alpha)}^{\varphi(\beta)} f \ .$$

2. Ponecháme z časových důvodů bez důkazu. □

Aplikace integrálů. Odhadneme tzv. *harmonická čísla* H_n,
$$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \ .$$

Pro funkci $f(x) = 1/x : (0,+\infty) \to (0,+\infty)$ a dělení $D = (1,2,\ldots,n+1)$ intervalu $[1,n+1]$ zřejmě máme
$$s(f,D) = \sum_{i=1}^{n} 1 \cdot \frac{1}{i+1} = H_{n+1} - 1 \ \text{ a } \ S(f,D) = \sum_{i=1}^{n} 1 \cdot \frac{1}{i} = H_n \ .$$

Protože $s(f,D) < \int_1^{n+1} 1/x = \log(n+1) < S(f,D)$, pro $n \geq 2$ dostáváme odhad
$$\log(n+1) < H_n < 1 + \log n \ .$$

Cvičení: dokažte, že pro $n \geq 2$ nikdy H_n není celé číslo.

Podobně *faktoriál* $n! = 1 \cdot 2 \cdot \ldots \cdot n$: pro funkci $f(x) = \log x : [1, +\infty) \to [0, +\infty)$ a dělení $D = (1, 2, \ldots, n+1)$ intervalu $[1, n+1]$ zřejmě máme

$$s(f, D) = \sum_{i=1}^{n} 1 \cdot \log i = \log(n!) \;\; \text{a} \;\; S(f, D) = \sum_{i=1}^{n} 1 \cdot \log(i+1) = \log((n+1)!) \ .$$

Protože $s(f, D) < \int_1^{n+1} \log x = (n+1)\log(n+1) - (n+1) + 1 < S(f, D)$, pro $n \geq 2$ dostáváme odhad

$$n \log n - n + 1 < \log(n!) < (n+1)\log(n+1) - n$$

a tedy

$$e\left(\frac{n}{e}\right)^n < n! < e\left(\frac{n+1}{e}\right)^{n+1} \ .$$

Přednáška 6, 27. března 2015

Rafinovanějším počítáním s integrály lze odvodit[1] přesnou asymptotiku faktoriálu, tzv. *Stirlingovu formuli*

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n,\ n \to \infty\ .$$

Podobně lze odhadovat i nekonečné řady a jejich součty, ale k tomu jsou třeba integrály přes nekonečné intervaly. Pro $a \in \mathbb{R}$ a $f:\ [a, +\infty) \to \mathbb{R}$, kde $f \in \mathcal{R}(a, b)$ pro každé $b > a$, položíme

$$\int_a^{+\infty} f := \lim_{b \to +\infty} \int_a^b f\ ,$$

pokud tato limita existuje (povolíme i nevlastní hodnotu limity).

Tvrzení (integrální kritérium konvergence). *Nechť a je celé číslo a funkce $f:\ [a, +\infty) \to \mathbb{R}$ je na $[a, +\infty)$ nezáporná a nerostoucí. Pak*

$$\sum_{n=a}^{\infty} f(n) = f(a) + f(a+1) + f(a+2) + \cdots < +\infty \iff \int_a^{+\infty} f < +\infty\ .$$

Řada tedy konverguje, právě když konverguje odpovídající integrál.

Důkaz. Posloupnost částečných součtů řady je neklesající a její limita tedy existuje a je vlastní nebo $+\infty$. Díky monotonii f je $f \in \mathcal{R}(a, b)$ pro každé reálné $b > a$. Díky nezápornosti f je $\int_a^{b'} f \ge \int_a^b f$, jakmile $b' \ge b$. Podobně tedy $\lim_{b\to+\infty} \int_a^b f$ existuje a je vlastní nebo $+\infty$. Pro celé číslo $b > a$ vezmeme dělení $D = (a, a+1, a+2, \ldots, b)$ intervalu $[a, b]$. Máme nerovnosti

$$\sum_{i=a+1}^{b} f(i) = s(f, D) \le \int_a^b f \le S(f, D) = \sum_{i=a}^{b-1} f(i)\ .$$

Z nich je jasné, že omezenost posloupnosti částečných součtů řady implikuje omezenost hodnot integrálů $\int_a^b f$ pro každé reálné $b > a$ a naopak. Obě limity jsou tedy současně vlastní nebo současně $+\infty$. □

[1] Odvození se najde třeba v mém textu *Kombinatorické počítání* na `http://kam.mff.cuni.cz/~klazar/kpoc99.ps`

V prvním příkladu na aplikaci tohoto kritéria rozhodneme o konvergenci řady

$$\sum_{n=1}^{\infty}\frac{1}{n^s},\ \ s>0\ .$$

Pro $s\neq 1$ je

$$\int_1^{+\infty}\frac{dx}{x^s}=\left.\frac{x^{1-s}}{1-s}\right|_1^{+\infty}=(1-s)^{-1}(\lim_{x\to+\infty}x^{1-s}-1)\ ,$$

což se rovná $+\infty$ pro $0<s<1$ a $(s-1)^{-1}$ pro $s>1$ (proč?). Pro $s=1$ je

$$\int_1^{+\infty}\frac{dx}{x}=\log x|_1^{+\infty}=\lim_{x\to+\infty}\log x=+\infty\ .$$

Podle integrálního kritéria tedy řada konverguje, právě když $s>1$. V druhém příkladu rozhodneme o konvergenci řady

$$\sum_{n=2}^{\infty}\frac{1}{n\log n}\ .$$

Zde

$$\int_2^{+\infty}\frac{dx}{x\log x}=\log\log x|_2^{+\infty}=\lim_{x\to+\infty}\log\log x-\log\log 2=+\infty\ .$$

Podle integrálního kritéria tedy řada diverguje. Cvičení: rozhodněte o konvergenci řady $\sum_{n\geq 2}1/n(\log n)^s$, $s>1$.

Integrálem jsme faktoriál $n!$ odhadli a nyní ukážeme, opět s pomocí integrálu, jak lze $n!$ rozšířit při zachování rekurence $n!=n\cdot(n-1)!$ na funkci definovanou na celém intervalu $[1,+\infty)$.[2]

Tvrzení (funkce Gamma). *Funkce*

$$\Gamma(x):=\int_0^{+\infty}t^{x-1}e^{-t}\,dt:\ [1,+\infty)\to(0,+\infty)$$

[2]Následující klasická integrální definice funguje na větším intervalu $x\in(0,+\infty)$, ale pro jednoduchost (pro $x<1$ je nutné integrál i u dolní integrační meze 0 definovat limitou) se omezujeme na menší interval $x\in[1,+\infty)$.

splňuje na intervalu $[1, +\infty)$ *funkcionální rovnici*

$$\Gamma(x+1) = x\Gamma(x) \ .$$

Máme $\Gamma(1) = 1$ *a* $\Gamma(n) = (n-1)!$ *pro celá čísla* $n \geq 2$.

Důkaz. Ukažme, že $\Gamma(x)$ je korektně definovaná. Pro každé pevné $x \geq 1$ je integrand nezáporná spojitá funkce (pro $x = 1$ a $t = 0$ klademe $0^0 = 1$). Protože $\lim_{t \to +\infty} t^{x-1} e^{-t/2} = 0$ (exponenciála porazí polynom), pro každé $t \in [0, +\infty)$ máme nerovnost

$$t^{x-1} e^{-t} = t^{x-1} e^{-t/2} \cdot e^{-t/2} \leq c e^{-t/2} \ ,$$

kde $c > 0$ je konstanta závisející jen na x. Integrály přes konečné intervaly $[0, b]$ jsou tedy definované, pro $b \to +\infty$ neklesají a mají vlastní limitu:

$$\int_0^b t^{x-1} e^{-t} \ dt \leq \int_0^b c e^{-t/2} = c(1 - e^{-b/2}/2) \ dt < c \ .$$

Hodnota $\Gamma(x)$ je tak korektně definovaná pro každé $x \geq 1$. Pro $x = 1$ máme

$$\Gamma(1) = \int_0^{+\infty} e^{-t} \ dt = (-e^{-t})|_0^{+\infty} = 0 - (-1) = 1 \ .$$

Funkcionální rovnici odvodíme integrací per partes:

$$\begin{aligned} \Gamma(x+1) &= \int_0^{+\infty} t^x e^{-t} \ dt = t^x(-e^{-t})|_0^{+\infty} - \int_0^{+\infty} x t^{x-1}(-e^{-t}) \ dt \\ &= 0 - 0 + x \int_0^{+\infty} t^{x-1} e^{-t} \ dt \\ &= x\Gamma(x) \ . \end{aligned}$$

Hodnoty $\Gamma(n)$ z ní plynou indukcí. □

Ještě tři poznámky k funkci $\Gamma(x)$. Zaprvé, samotné rozšíření faktoriálu vyřešením rekurence $f(x+1) = xf(x)$ nějakou funkcí $f: \ [1, +\infty) \to \mathbb{R}$ není nic zázračného a dá se udělat jednoduše i bez integrálů. Vezmeme si funkci f definovanou na $[1, 2)$ jako $f(1) = 1$ a jinak úplně libovolně, například jako $f(x) = 1$, a na celý interval $[1, +\infty)$ ji protáhneme právě pomocí vztahu $f(x+1) = xf(x)$. Pro $f = 1$ na $[1, 2)$ tak dostaneme $f(x) = x - 1$ na $[2, 3)$, $f(x) = (x-1)(x-2)$ na $[3, 4)$ a tak dál. Tato funkce splňuje na $[1, +\infty)$

rovnici $f(x+1) = xf(x)$ a $f(1) = 1$, takže i $f(n) = (n-1)!$. Dokonce je na definičním intervalu spojitá. Její graf však nevypadá moc hezky, protože v bodech $2, 3, 4, \ldots$ má „špičky“ — f v nich nemá derivaci (má různé jednostranné derivace). Pokud navíc po $f(x)$ chceme, aby na $[1, +\infty)$ měla první derivaci, popřípadě i další derivace, tak tato jednoduchá konstrukce nestačí. Pak se musíme obrátit ke $\Gamma(x)$, o níž se dá ukázat, že na $[1, +\infty)$ má derivace všech řádů. Zadruhé, bylo by hezčí mít rozšíření faktoriálu splňující přímo vztah $f(n) = n!$. Z tohoto hlediska je hořejší definice funkce $\Gamma(x)$ trochu nešikovná. Vznikla však historicky, stejně jako název funkce, a již je zavedená (a nelze s tím nic udělat). Zatřetí, jaké jsou nějaké další hodnoty $\Gamma(x)$, kromě bodů $x = 1, 2, 3, \ldots$? Dá se třeba spočítat, že

$$\Gamma(3/2) = \frac{\sqrt{\pi}}{2} = 0.8862\ldots\ .$$

Funkce $\Gamma(x)$ splňuje řadu dalších zajímavých vztahů a identit.

Jako závěrečnou aplikaci integrálu připomeneme vzorce pro plochu, délku křivky a objem rotačního tělesa. Plochu rovinného útvaru $U(a, b, f)$ (jsou to body (x, y) v rovině splňující $a \le x \le b$ a $0 \le y \le f(x)$) pod grafem funkce f jsme víceméně definovali jako $\int_a^b f$.

Pro funkci $f : [a, b] \to \mathbb{R}$ můžeme délku jejího grafu $G = \{(x, f(x)) \in \mathbb{R}^2 \mid a \le x \le b\}$ jakožto oblouku křivky definovat jako limitu délek lomených čar L spojujících konce G a s body zlomu na G, když délka nejdelší úsečky v L jde k 0. Když je f pěkná funkce, například f' je spojitá, tato limita existuje a můžeme ji spočítat R. integrálem. Úsečka v L spojující body $(x, f(x))$ a $(x + \Delta, f(x + \Delta))$ má podle Pythagorovy věty délku

$$\sqrt{\Delta^2 + (f(x+\Delta) - f(x))^2} = \Delta \sqrt{1 + \left(\frac{f(x+\Delta) - f(x)}{\Delta}\right)^2}\ ,$$

a hodnota zlomku je podle Lagrangeovy věty o střední hodnotě rovna $f'(\alpha)$ v nějakém mezibodě α (ležícím mezi x a $x + \Delta$). Délka lomené čáry L tak je vlastně přímo Riemannova suma pro jisté dělení intervalu $[a, b]$ s body a funkci $\sqrt{1 + (f'(t))^2}$. Dostáváme následující vzorec.

Tvrzení (délka oblouku křivky). *Nechť $f : [a, b] \to \mathbb{R}$ má na $[a, b]$ spojitou derivaci f' (takže $\sqrt{1 + (f')^2} \in \mathcal{R}(a, b)$). Pak*

$$\text{délka}(\{(x, f(x)) \in \mathbb{R}^2 \mid a \le x \le b\}) = \int_a^b \sqrt{1 + (f'(t))^2}\, dt\ .$$

Pro podmnožinu $M \subset \mathbb{R}^3$ trojrozměrného prostoru můžeme její objem definovat jako limitu, pro $n \to \infty$, součtu objemů $1/n^3$ krychliček K v množině

$$\{K = [\tfrac{a}{n}, \tfrac{a+1}{n}] \times [\tfrac{b}{n}, \tfrac{b+1}{n}] \times [\tfrac{c}{n}, \tfrac{c+1}{n}] \mid a, b, c \in \mathbb{Z} \;\&\; K \subset M\} \ .$$

Je-li M hezká, tato limita existuje a můžeme ji spočítat integrálem. Uvedeme si jeden speciální případ.

Tvrzení (objem rotačního tělesa). *Nechť $f \in \mathcal{R}(a, b)$ a $f \geq 0$ na $[a, b]$. Pro objem rotačního tělesa*

$$V = \{(x, y, z) \in \mathbb{R}^3 \mid a \leq x \leq b \;\&\; \sqrt{y^2 + z^2} \leq f(x)\}$$

vzniklého rotací (v $\mathbb{R}^3$) rovinného útvaru $U(a, b, f)$ pod grafem funkce f kolem osy x platí vztah

$$\text{objem}(V) = \pi \int_a^b f(t)^2 \, dt \ .$$

Vzorec se dostane rozřezáním V rovinami kolmými na osu x na plátky tloušťky $\Delta > 0$ a sečtením jejich objemů. Objem plátku mezi rovinami kolmými v bodech $(x, 0, 0)$ a $(x + \Delta, 0, 0)$ je přibližně $\pi f(x)^2 \Delta$, neboť to je zhruba válec s (kruhovou) podstavou o poloměru $f(x)$ a výškou Δ.

Cvičení: pomocí prvního vzorce spočtěte délku obvodu kružnice a pomocí druhého objem koule.

Diferenciální počet funkcí několika proměnných

Budeme pracovat v m-rozměrném *euklidovském prostoru* $\mathbb{R}^m$, $m \in \mathbb{N}$, což je množina všech uspořádaných m-tic reálných čísel $x = (x_1, x_2, \ldots, x_m)$ s $x_i \in \mathbb{R}$. Je to m-rozměrný vektorový prostor nad tělesem $\mathbb{R}$ — jeho prvky můžeme mezi sebou sčítat a odečítat a skalárně je násobit reálnými čísly. Vzdálenosti v $\mathbb{R}^m$ měříme pomocí *(euklidovské) normy*, což je zobrazení $\|\cdot\| : \mathbb{R}^m \to [0, +\infty)$ dané formulí

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

Vlastnosti normy ($a \in \mathbb{R}$, $x, y \in \mathbb{R}^m$):

1. (nezápornost a nenulovost) $\|x\| \geq 0$ a $\|x\| = 0 \iff x = \overline{0} = (0, 0, \ldots, 0)$,
2. (homogenita) $\|ax\| = |a| \cdot \|x\|$ a
3. (trojúhelníková nerovnost) $\|x + y\| \leq \|x\| + \|y\|$.

Z normy dostáváme *(euklidovskou) vzdálenost* $d(x, y) : \mathbb{R}^m \times \mathbb{R}^m \to [0, +\infty)$ mezi dvěma body x a y v $\mathbb{R}^m$:

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

Vlastnosti vzdálenosti ($x, y, z \in \mathbb{R}^m$):

1. (nezápornost a nenulovost) $d(x, y) \geq 0$ a $d(x, y) = 0 \iff x = y$,
2. (symetrie) $d(x, y) = d(y, x)$ a
3. (trojúhelníková nerovnost) $d(x, y) \leq d(x, z) + d(z, y)$.

Až na trojúhelníkovou nerovnost, jejíž důkaz dá trochu práci, jsou všechny vlastnosti normy i vzdálenosti zřejmé z definice.

(Otevřená) koule $B(s, r)$ s poloměrem $r > 0$ a středem $s \in \mathbb{R}^m$ je množina bodů v $\mathbb{R}^m$ se vzdáleností od s menší než r:

$$B(s, r) = \{x \in \mathbb{R}^m \mid \|x - s\| < r\}\ .$$

Otevřená množina v $\mathbb{R}^m$ je podmnožina $M \subset \mathbb{R}^m$ s tou vlastností, že s každým svým bodem x obsahuje i nějakou kouli se středem v x:

$$M \text{ je otevřená} \iff \forall x \in M\ \exists r > 0 :\ B(x, r) \subset M\ .$$

Přednáška 7, 3. dubna 2015

Jako cvičení zdůvodněte tyto vlastnosti otevřených množin v $\mathbb{R}^m$: (i) množiny $\emptyset$ a $\mathbb{R}^m$ jsou otevřené, (ii) sjednocení $\bigcup_{i\in I} A_i$ libovolného systému $\{A_i \mid i \in I\}$ otevřených množin A_i je otevřená množina a (iii) průnik dvou (a tedy i konečně mnoha) otevřených množin je otevřená množina. Průnik nekonečně mnoha otevřených množin už ale nemusí být otevřená množina. *Okolí bodu* $a \in \mathbb{R}^m$ je libovolná otevřená množina v $\mathbb{R}^m$ obsahující a.

Budeme pracovat s funkcemi $f:\ M \to \mathbb{R}$, $f = f(x_1, x_2, \ldots, x_m)$, definovanými na podmnožinách $M \subset \mathbb{R}^m$ (které budou většinou otevřené), nebo i obecněji se zobrazeními

$$f:\ M \to \mathbb{R}^n,\ M \subset \mathbb{R}^m,\ f = (f_1, f_2, \ldots, f_n)\ ,$$

kde $f_i = f_i(x_1, x_2, \ldots, x_m)$ jsou *souřadnicové funkce*. Naším cílem bude jednak zobecnit derivaci jako lineární aproximaci z funkcí jedné proměnné na funkce několika proměnných, a pak stejně zobecnit kritéria existence extrému.

Než se pustíme do derivací, zobecníme spojitost: je-li $U \subset \mathbb{R}^m$ okolí bodu a, funkci $f:\ U \to \mathbb{R}$ nazveme *spojitou v bodě* a, když

$$\forall \varepsilon > 0\ \exists \delta > 0:\ \|x - a\| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon\ .$$

Stejně se definuje spojitost pro zobrazení $f:\ U \to \mathbb{R}^n$, pouze absolutní hodnotu $|f(x) - f(a)|$ (což je norma v $\mathbb{R}^1$) nahradíme normou $\|f(x) - f(a)\|$ v $\mathbb{R}^n$.

Směrová derivace, parciální derivace, diferenciál. Nechť $U \subset \mathbb{R}^m$ je okolí bodu a a $f:\ U \to \mathbb{R}$ je funkce. Její *směrovou derivací* v bodu a ve směru vektoru $v \in \mathbb{R}^m \backslash \{\overline{0}\}$ rozumíme limitu

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

existuje-li. Představme si U jako oblast v třírozměrném euklidovském prostoru, kde funkce f měří teplotu a kterou prolétává nějaká částice. Směrová derivace $\mathrm{D}_v f(a)$ udává okamžitou změnu teploty částice ve chvíli, kdy se nalézá 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,

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

Vektor hodnot všech parciálních derivací funkce f v bodě a je *gradient funkce f v a*,

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

Funkce f má v bodě a *(totální) diferenciál*, jinými slovy f je v a *diferencovatelná*, když existuje takové lineární zobrazení $L:\ \mathbb{R}^m \to \mathbb{R}$, 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 $\mathrm{D}f(a)(h)$. Obecněji, pokud $f:\ U \to \mathbb{R}^n$ je zobrazení, řekneme, že je v a *diferencovatelné*, když existuje takové lineární zobrazení $L:\ \mathbb{R}^m \to \mathbb{R}^n$, že

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

(norma ve jmenovateli se bere v $\mathbb{R}^m$ a norma v čitateli v $\mathbb{R}^n$). L opět nazýváme *diferenciálem* a značíme $\mathrm{D}f(a)$. 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 $\mathbb{R}^m$ 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. Například funkce

$f = f(x, y) : \mathbb{R}^2 \to \mathbb{R}$ definovaná jako $f(x, y) = 1$ pokud $xy = 0$ a $f(x, y) = 0$ jinde (tj. f je 1 na souřadnicových osách a jinak 0) má v počátku obě parciální derivace a jejich hodnota je 0, ale není v počátku spojitá. Jako cvičení vymyslete funkci z $\mathbb{R}^2$ do $\mathbb{R}$, která má v počátku všechny směrové derivace, ale přesto tam není spojitá.

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 pro funkci $f = f(x, y, z) = x^3 y \sin(yz) + x \log z$ máme

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

Jednoduchý důkaz následujícího tvrzení ponecháváme jako cvičení.

Tvrzení (vlastnosti diferenciálu). *Nechť $f = (f_1, f_2, \ldots, f_n) : U \to \mathbb{R}^n$ je zobrazení, kde $U \subset \mathbb{R}^m$ je okolí bodu a.*

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

Tvrzení (diferenciál $\Rightarrow \partial$). *Když je funkce $f : U \to \mathbb{R}$, $U \subset \mathbb{R}^m$ je okolí bodu a, diferencovatelná v a, pak má v a všechny parciální derivace a jejich hodnoty diferenciál určují:*

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

(tj. hodnota diferenciálu v h se dostane jako obvyklý skalární součin h a gradientu f v a). Funkce f pak má také 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

$$L(h) = L(h_1 e_1 + h_2 e_2 + \cdots + h_m e_m) = L(e_1)h_1 + \cdots + L(e_m)h_m \ ,$$

kde e_i je i-tý vektor kanonické báze. Ovšem ($f(a+te_i) = f(a) + L(te_i) + o(\|te_i\|)$ pro $t \to 0$ vzhledem k diferencovatelnosti f v a)

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

a tak $L(e_i) = \frac{\partial f}{\partial x_i}(a)$.

Nechť $v \in \mathbb{R}^m$ je nenulový vektor. Protože $v = v_1e_1 + \cdots + v_me_m$, kde e_i jsou vektory kanonické báze $\mathbb{R}^m$, a f má v a všechny parciální derivace, z definice směrové derivace plyne, že

$$\mathrm{D}_v f(a) = v_1 \cdot (\partial f/\partial x_1)(a) + v_2 \cdot (\partial f/\partial x_2)(a) + \cdots + v_m \cdot (\partial f/\partial x_m)(a)\ ,$$

což je právě hodnota diferenciálu $\mathrm{D}f(a)$ ve v. $\square$

Obecné zobrazení $f : U \to \mathbb{R}^n$ má diferenciál $L = \mathrm{D}f(a) : \mathbb{R}^m \to \mathbb{R}^n$ popsaný maticí typu $n \times m$ 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 předešlých tvrzení má tato matice v i-tém řádku gradient souřadnicové funkce f_i v bodě a:

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

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

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

Je-li Jacobiho matice čtvercová, její determinant se nazývá *jacobián.*

Věta ($\partial \Rightarrow$ diferenciál). *Nechť $U \subset \mathbb{R}^m$ je okolí bodu $a \in \mathbb{R}^m$. Má-li funkce $f : U \to \mathbb{R}$ na U všechny parciální derivace a ty jsou v bodě a spojité, pak je f v bodě a diferencovatelná.*

Důkaz si řekneme příště.

Přednáška 8, 10. dubna 2015

Důkaz. Omezíme se na případ dvou proměnných x a y ($m = 2$) a bod $a = \overline{0} = (0, 0)$, pro více proměnných se postupuje podobně. Rovněž můžeme předpokládat, že $U \subset \mathbb{R}^2$ je koule (tedy otevřený kruh) se středem v počátku. Nechť $h = (h_1, h_2) \in U$ a $h' = (h_1, 0)$. Přírůstek $f(h) - f(\overline{0})$ napíšeme pomocí přírůstků ve směrech obou souřadnicových os:

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

Úsečky $h'h$ a $\overline{0}h'$ obě leží v U, funkce f je na nich definovaná a na první úsečce závisí pouze na proměnné y a na druhé jen na x. Pro obě úsečky a funkci f použijeme Lagrangeovu větu o střední hodnotě (pro funkce jedné proměnné):

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

kde ζ_2 (resp. ζ_1) je nějaký vnitřní bod úsečky $h'h$ (resp. $\overline{0}h'$). Body ζ_1 a ζ_2 leží v otevřené kouli $B(\overline{0}, \|h\|)$. Díky spojitosti obou parciálních derivací v počátku máme

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

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

$$f(h) - f(\overline{0}) = \frac{\partial f}{\partial y}(\overline{0}) \cdot h_2 + \frac{\partial f}{\partial x}(\overline{0}) \cdot h_1 + \alpha(\zeta_2)h_2 + \beta(\zeta_1)h_1\ .$$

Z trojúhelníkové nerovnosti a nerovností $0 < \|\zeta_1\|, \|\zeta_2\| < \|h\|$ a $|h_1|, |h_2| \leq \|h\|$ plyne, že když $\|h\| < \delta$, tak

$$|\alpha(\zeta_2)h_2 + \beta(\zeta_1)h_1| \leq |\alpha(\zeta_2)| \cdot \|h\| + |\beta(\zeta_1)| \cdot \|h\| \leq 2\varepsilon\|h\|\ .$$

Tedy $\alpha(\zeta_2)h_2 + \beta(\zeta_1)h_1 = o(\|h\|)$ pro $h \to \overline{0}$. Podle definice diferenciálu je funkce f diferencovatelná v počátku. $\square$

Zobecníme Lagrangeovu větu o střední hodnotě pro funkce více proměnných.

Tvrzení (Lagrange pro funkce několika proměnných). *Nechť $U \subset \mathbb{R}^m$ je otevřená množina obsahující úsečku $u = ab$ s koncovými body a a b a $f : U \to \mathbb{R}$ je funkce, jež je spojitá v každém bodě u a má v každém vnitřním bodě u diferenciál. Pak pro nějaký vnitřní bod ζ úsečky u platí, že*

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

Rozdíl hodnot funkce v krajních bodech úsečky se tedy rovná hodnotě jejího diferenciálu v nějakém vnitřním bodě úsečky na směrovém vektoru úsečky.

Důkaz. Udělejte si jako cvičení, s pomocí funkce $F(t) = f(a + t(b - a))$, kde reálné číslo t probíhá interval $[0, 1]$. □

Otevřená množina $D \subset \mathbb{R}^m$ je *souvislá*, když lze každé dva její body spojit lomenou čarou, jež celá leží v D. Příklady souvislých otevřených množin: koule v $\mathbb{R}^m$, celé $\mathbb{R}^m$ a $\mathbb{R}^3 \backslash L$, kde L je sjednocení konečně mnoha přímek. Na druhou stranu ale $B \backslash R$, kde B je otevřená koule v $\mathbb{R}^3$ a R rovina protínající B, je otevřená avšak nikoli souvislá množina.

Důsledek ($\partial = 0 \Rightarrow f \equiv$ const.). *Má-li funkce m proměnných f v každém bodě otevřené a souvislé množiny U nulový diferenciál, je na U konstantní. Totéž platí, má-li f v každém bodě U nulovou každou parciální derivaci.*

Důkaz. Nechť $U \subset \mathbb{R}^m$ je otevřená a souvislá množina a funkce $f : U \to \mathbb{R}$ má na U nulový diferenciál. Vezmeme dva libovolné body $a, b \in U$ a spojíme je lomenou čarou $s = s_1 s_2 \ldots s_r$ ležící v U. Pro libovolnou úsečku $s_i = a_i b_i$ z s máme díky předchozímu 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 $f(a) = f(b)$.

Má-li f v každém bodě U nulovou každou parciální derivaci, je podle tvrzení z minulé přednášky v každém bodě U diferencovatelná a (podle vyjádření diferenciálu pomocí parciálních derivací) její diferenciál je vždy nulový, čímž jsme v předchozí situaci. □

Počítání s parciálními derivacemi a diferenciály. Pro dvě funkce $f, g : U \to \mathbb{R}$, které jsou definované na okolí $U \subset \mathbb{R}^m$ bodu $a \in U$ a mají v bodě a parciální derivaci podle proměnné x_i, máme pro parciální derivaci podle

x_i jejich lineární kombinace, součinu a podílu stejné vzorce jako v případě funkcí jedné proměnné (místo $\frac{\partial f}{\partial x_i}$ píšeme stručně $\partial_i f$):

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

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

Tvrzení (počítání s diferenciály). *Nechť $U \subset \mathbb{R}^m$ je okolí bodu a a $f, g : U \to \mathbb{R}$ jsou dvě funkce, obě diferencovatelné v a.*
1. *Když $\alpha, \beta \in \mathbb{R}$, pak i funkce $\alpha f + \beta g$ je v a diferencovatelná a*

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

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

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

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

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

Všimněme si, že výsledný diferenciál je vždy lineární kombinace diferenciálů funkcí f a g.

Důkaz. Tyto vzorce plynou z analogických vzorců pro parciální derivace a z vyjádření diferenciálu pomocí parciálních derivací. □

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

Zobecníme vzorec pro derivaci složené funkce na obecný případ skládání 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 (diferenciál složeného zobrazení). *Nechť*

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

jsou dvě zobrazení, kde $U \subset \mathbb{R}^m$ je okolí bodu a a $V \subset \mathbb{R}^n$ je okolí bodu $b = f(a)$. Je-li zobrazení f diferencovatelné v a a g je diferencovatelné v b, je složené zobrazení

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

diferencovatelné v 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, připomeneme význam symbolů $o(h)$ a $O(h)$ a explicitně uvedeme jejich jednoduché vlastnosti, které v důkazu využijeme.

Pro zobrazení $z : \; U \to \mathbb{R}^n$, kde $U \subset \mathbb{R}^m$ je okolí počátku, budeme psát stručně $z(x) = o(x)$ místo $\|z(x)\| = o(\|x\|)$ a $z(x) = O(x)$ místo $\|z(x)\| = O(\|x\|)$, vždy $x \to \overline{0}$. Připomeňme si, že značení $z(x) = o(x)$ je zkratka pro

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

(to jest $\|z(x)\|/\|x\| \to 0$ pro $x \to \overline{0}$) a $z(x) = O(x)$ je zkratka pro

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

(to jest podíl $\|z(x)\|/\|x\|$ je v prstencovém okolí $\overline{0}$ omezený).

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

1. *Když je z_1 lineární zobrazení, potom $z_1(x) = O(x)$.*
2. *Když $z_1(x) = o(x)$ a $z_2(x) = o(x)$, potom $z_1(x) + z_2(x) = o(x)$.*
3. *Když $z_1(x) = o(x)$ 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. Cvičení. $\square$

Důkaz věty. 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 $\gamma(h) = o(h)$ a $\beta(h) = o(h)$. 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

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

kde

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

Zbývá ukázat, že pro $h \to \overline{0}$ je $\alpha(h) = o(h)$. První sčítanec ve vyjádření $\alpha(h)$ je $o(h)$ podle částí 1 a 4 lemmatu (lineární, tedy O, zobrazení složené s o dává o) a druhý je rovněž $o(h)$ podle částí 1, 3 a 5 (o zobrazení složené se součtem O a o je o složené s O a tedy o). Celkem $\alpha(h) = o(h)$ podle části 2. Takže $g \circ f$ má v a diferenciál rovný lineárnímu zobrazení $\mathrm{D}g(b) \circ \mathrm{D}f(a)$. $\square$

Přednáška 9, 17. dubna 2015

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:

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

Speciálně pro $k = 1$, kdy funkce $h = h(x_1, x_2, \ldots, x_m)$ o m proměnných je složeninou

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

funkce $g = g(x_1, x_2, \ldots, x_n)$ o n proměnných s n funkcemi $f_i = f_i(x_1, x_2, \ldots, x_m)$, 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 $\partial_i f = (\partial_i f_1, \partial_i f_2, \ldots, \partial_i f_n)$.

Geometrie parciálních derivací. 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 $(x_0, y_0) \in U \subset \mathbb{R}^2$, kde U je otevřená množina v rovině, a $f: U \to \mathbb{R}$ je funkce. Její graf

$$G_f = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in U, z = f(x, y)\}$$

je plocha v třírozměrném euklidovském prostoru. Na G_f leží bod (x_0, y_0, z_0), kde $z_0 = f(x_0, y_0)$. Nechť je funkce f v bodě (x_0, y_0) diferencovatelná. Potom mezi všemi afinními funkcemi dvou proměnných $L(x, y)$ (tj. $L(x, y) = \alpha +$

$\beta x + \gamma y)$, jejichž graf obsahuje 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 funkce

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

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

$$G_T = \{(x,y,z) \in \mathbb{R}^3 \mid (x,y) \in \mathbb{R}^2, z = T(x,y)\}$$

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

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

kde $V \in \mathbb{R}^3$ 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 G_T zapsat i jako

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

Tvoří ji tedy právě ty 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 $f : U \to \mathbb{R}$ definovaná na okolí $U \subset \mathbb{R}^m$ bodu a v každém bodě U parciální derivaci $F = \partial_i f$ a tato funkce $F : U \to \mathbb{R}$ má v bodě a 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

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

Podobně definujeme parciální derivace vyšších řádů: má-li $f = f(x_1, x_2, \ldots, x_m)$ v každém bodě $x \in U$ parciální derivaci ($i_1, i_2, \ldots, i_{k-1}, j \in \{1, 2, \ldots, m\}$)

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

a F 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 $x_{i_1}, \ldots, x_{i_{k-1}}, x_j$* a její hodnotu značíme

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

Na pořadí proměnných při parciálním derivování obecně záleží: jako cvičení dokažte, že funkce $f : \mathbb{R}^2 \to \mathbb{R}$,

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

má v počátku obě smíšené parciální derivace druhého řádu s různými hodnotami

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

Při spojitých parciálních derivacích však na pořadí proměnných nezáleží.

Tvrzení (obvykle $\partial_x \partial_y f = \partial_y \partial_x f$). *Nechť funkce $f : U \to \mathbb{R}$ má na okolí $U \subset \mathbb{R}^m$ bodu a parciální derivace druhého řádu $\partial_j \partial_i f$ a $\partial_i \partial_j f$, $i \neq j$, a ty jsou v a spojité. Pak*

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

Důkaz. Nechť $m = 2$ a $a = \overline{0} = (0, 0)$, obecný případ je velmi podobný. Díky spojitosti obou parciálních derivací v počátku stačí nalézt pro každé (dosti malé) $h > 0$ ve čtverci $[0, h]^2$ dva body σ a τ, v nichž $\partial_x \partial_y f(\sigma) = \partial_y \partial_x f(\tau)$. Pro $h \to 0^+$ pak totiž $\sigma, \tau \to \overline{0}$ a limitní přechod a spojitost obou parciálních derivací v $\overline{0}$ dávají, že $\partial_x \partial_y f(\overline{0}) = \partial_y \partial_x f(\overline{0})$.

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 dvě 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 $0 < s_0, t_0 < h$ jsou mezibody. 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 máme $\partial_x\partial_y f(\sigma) = \partial_y\partial_x f(\tau)$ (protože obě hodnoty se rovnají témuž číslu $(f(d) - f(b) - f(c) + f(a))/h^2$). □

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

Pro otevřenou množinu $U \subset \mathbb{R}^m$ označíme symbolem $\mathcal{C}^k(U)$ množinu funkcí $f : U \to \mathbb{R}$, jejichž všechny parciální derivace 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 se posloupnosti $(i_1, \ldots, i_l)$ a $(j_1, \ldots, j_l)$ liší jen pořadím členů.

Důkaz. Když je posloupnost $v = (j_1, \ldots, j_l)$ pouze permutací posloupnosti $u = (i_1, \ldots, i_l)$, dokážeme u proměnit ve v prohazováním dvojic členů v u, dokonce stačí prohazovat sousední členy: v u nalezneme člen j_1 a necháme ho „propadnout“ až dolů na první místo, pak necháme propadnout na druhé místo j_2 atd. Rovnost hodnot parciálních derivací tak plyne z předchozího tvrzení. □

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 stručněji ∂x^2 apod. Například, pro f z $\mathcal{C}^5(U)$ na U máme

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

Důležitým nástrojem při studiu funkcí je Taylorův polynom, jenž nyní zobecníme pro více proměnných. Na příkladu vysvětlíme, jak rozumět použitému symbolickému zápisu mocniny diferenciálního operátoru. Nechť $f = f(x, y, z)$ je funkce z $\mathcal{C}^3(U)$ a $a \in \mathbb{R}^3$, $\alpha, \beta \in \mathbb{R}$ jsou konstanty. Například zápisem

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

se rozumí

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

Podobně pro jiné mocniny.

Věta (zobecnění Taylorova polynomu). *Nechť $U \subset \mathbb{R}^m$ je okolí bodu a a $f : U \to \mathbb{R}$ je funkce z $\mathcal{C}^n(U)$. Potom pro každý bod $h = (h_1, h_2, \ldots, h_m)$, že $a + h \in U$, máme Taylorův rozvoj*

$$\begin{aligned}
f(a+h) &= \sum_{i=0}^{n} \frac{1}{i!}(h_1\partial_1 + h_2\partial_2 + \cdots + h_m\partial_m)^i f(a) + e(h) \\
&= \sum \frac{1}{i_1! i_2! \ldots i_m!} \cdot \frac{\partial^{i_1+i_2+\cdots+i_m} f}{\partial x_1^{i_1} \partial x_2^{i_2} \ldots \partial x_m^{i_m}}(a) \cdot h_1^{i_1} h_2^{i_2} \ldots h_m^{i_m} + e(h) \\
&= f(a) + \sum_{i=1}^{m} \partial_{x_i} f(a) h_i + \sum_{1 \le i < j \le m} \partial_{x_i}\partial_{x_j} f(a) h_i h_j + \\
&\quad + \frac{1}{2}\sum_{i=1}^{m} \partial_{x_i}^2 f(a) h_i^2 + \cdots + e(h)\ ,
\end{aligned}$$

kde $e(h)$ je chybová funkce splňující pro $h \to \overline{0}$ odhad $e(h) = o(\|h\|^n)$, tj. $\lim_{h\to\overline{0}} e(h)/\|h\|^n = 0$. V prvním výrazu mocninu chápeme symbolicky (ve *výše popsaném smyslu) a ve druhém, kde jsme ji rozvinuli podle multinomické věty, v sumě sčítáme přes všechny* m*-tice nezáporných celých čísel* $i_1, i_2, \ldots, i_m$ *se součtem nejvýše* n. *Ve třetím výrazu jsme uvedli začátek rozvoje pro hodnoty* $i = 0, 1$ *a* 2.

Přednáška 10, 24. dubna 2015

Vzpomeňme si na klasickou postačující podmínku pro existenci lokálního extrému funkce $f(x)$ jedné proměnné v bodě $a \in \mathbb{R}$: z jejího Taylorova polynomu stupně 2 (se středem v a)

$$f(a+h) = f(a) + f'(a)h + \tfrac{1}{2}f''(a)h^2 + o(h^2),\ h \to 0\ ,$$

ihned vidíme, že

1. pokud $f'(a) \neq 0$, funkce f nemá v a lokální extrém;
2. pokud $f'(a) = 0$ a $f''(a) > 0$, funkce f má v a ostré lokální minimum a
3. pokud $f'(a) = 0$ a $f''(a) < 0$, funkce f má v a ostré lokální maximum.

Pokud $f'(a) = f''(a) = 0$, nelze bez další analýzy o existenci extrému v a říci nic. Pokud $f'(a) = 0$ (a je „podezřelý“ bod), nemůžeme tedy jen ze samotné hodnoty druhé derivace $f''(a)$ nikdy vydedukovat *neexistenci* lokálního extrému. Jak uvidíme, pro funkce více proměnných je situace jiná. Uvedenou postačující podmínku nyní zobecníme na tyto funkce.

Nejprve ale zavedeme značení a oživíme si pár věcí z lineární algebry. Nechť $A = (a_{i,j}) \in \mathbb{R}^{n\times n}$ je symetrická ($a_{i,j} = a_{j,i}$) reálná $n \times n$ matice. Přiřadíme jí kvadratickou formu (= homogenní polynom stupně 2) o n proměnných

$$P_A(x_1, x_2, \ldots, x_n) = xAx^T = \sum_{i,j=1}^{n} a_{i,j}x_ix_j :\ \mathbb{R}^n \to \mathbb{R}\ ,$$

kde x označuje řádkový vektor $(x_1, x_2, \ldots, x_n)$ a x^T je týž vektor psaný transponovaně ve sloupci. Zřejmě $P_A(\overline{0}) = 0$ a $P_A(tx) = t^2P_A(x)$ (díky homogenitě P) pro každou matici A, vektor $x \in \mathbb{R}^n$ a skalár $t \in \mathbb{R}$. Matice A se nazývá

- *pozitivně (resp. negativně) definitní*, když $P_A(x) > 0$ (resp. $P_A(x) < 0$) pro každý bod $x \in \mathbb{R}^n\backslash\{\overline{0}\}$;
- *pozitivně (resp. negativně) semidefinitní*, když $P_A(x) \geq 0$ (resp. $P_A(x) \leq 0$) pro každý bod $x \in \mathbb{R}^n$ a
- *indefinitní*, není-li ani pozitivně ani negativně semidefinitní, to jest $P_A(x) > 0$ a $P_A(y) < 0$ pro nějaké dva body $x, y \in \mathbb{R}^n$.

Vzhledem ke zmíněné rovnosti $P_A(tx) = t^2 P_A(x)$ určují definitnost A už hodnoty P_A na *jednotkové sféře*

$$S = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$$

(proč?). Jak poznat definitnost A se o něco podrobněji zmíníme později.

Připomeneme nomenklaturu extrémů. Funkce $f: U \to \mathbb{R}$, kde $U \subset \mathbb{R}^m$ je okolí bodu a, má v a *ostré lokální minimum*, existuje-li takové $\delta > 0$, že $0 < \|x - a\| < \delta \Rightarrow f(x) > f(a)$. *(Neostré) lokální minimum* v a znamená, že $\|x - a\| < \delta \Rightarrow f(x) \geq f(a)$. Podobně pro ostré lokální maximum a (neostré) lokální maximum. Funkce f *nemá v a lokální extrém*, nemá-li v a ani lokální minimum ani lokální maximum, to jest pro každé $\delta > 0$ existují takové dva body x, y, že $\|x - a\|, \|y - a\| < \delta$ a $f(y) < f(a) < f(x)$. Funkce $f: M \to \mathbb{R}$, kde $M \subset \mathbb{R}^m$, *nabývá na množině M maximum v bodě $a \in M$*, když $f(a) \geq f(b)$ pro každý bod $b \in M$. Podobně pro nabývání minima.

V ZS jsme dokázali větu pro funkce jedné proměnné: každá funkce $f: I \to \mathbb{R}$, spojitá na kompaktním intervalu I (tj. $I = [a, b]$, $-\infty < a \leq b < +\infty$), nabývá na I maximum i minimum. Budeme potřebovat ji zobecnit na více proměnných. Množina $M \subset \mathbb{R}^m$ je *omezená*, když existuje takový poloměr $R > 0$, že $M \subset B(\overline{0}, R)$, a je *uzavřená*, když její doplněk $\mathbb{R}^m \backslash M$ je otevřená množina (to jest každý bod b mimo M leží mimo M i s nějakou celou koulí se středem v b). Množina $M \subset \mathbb{R}^m$ je *kompaktní*, když je současně omezená a uzavřená.

Věta (nabývání extrémů na kompaktu). *Nechť je funkce $f: M \to \mathbb{R}$, kde $M \subset \mathbb{R}^m$ je neprázdná kompaktní množina, na M spojitá. Pak f nabývá na M minimumm i maximum.*

Důkaz. Zatím bez důkazu. Podáme ho pravděpodobně později v partii o metrických prostorech. □

Například výše definovaná jednotková sféra S je kompaktní podmnožina $\mathbb{R}^n$, a proto každá spojitá funkce $f: S \to \mathbb{R}$ nabývá na S minimum i maximum.

Poslední definice před větou o lokálních extrémech: *Hessova matice $H_f(a)$ funkce f v bodě a*, kde $f: U \to \mathbb{R}$ a $U \subset \mathbb{R}^m$ je okolí bodu a a f 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 o $\partial_x \partial_y = \partial_y \partial_x$ je pro funkci z $\mathcal{C}^2(U)$ (tj. se spojitými druhými derivacemi na U) její Hessova matice symetrická.

Věta (o lokálních extrémech). *Nechť $f \in \mathcal{C}^2(U)$, kde $U \subset \mathbb{R}^m$ je okolí bodu a. Připomeňme, že (gradient) $\nabla f(a)$ je vektor hodnot prvních derivací a (Hessova matice) $H_f(a)$ je matice hodnot druhých derivací.*

1. *Pokud $\nabla f(a) \neq \overline{0}$, nemá f v a ani neostrý lokální extrém. (Zde stačí předpokládat pouze existenci gradientu $\nabla f(a)$.)*
2. *Pokud $\nabla f(a) = \overline{0}$ a $H_f(a)$ je pozitivně (resp. negativně) definitní, potom má f v a ostré lokální minimum (resp. maximum).*
3. *Pokud $\nabla f(a) = \overline{0}$ a $H_f(a)$ je indefinitní, nemá f v a 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 $f(a_1 + h, a_2, \ldots, a_m) = f(a) + \partial_{x_1} f(a)h + o(h)$ pro $h \to 0$. 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 $f(a_1 + h, a_2, \ldots, a_m) - f(a) > \frac{1}{2}\partial_{x_1} f(a)h > 0$. Proto funkce f nemá v a ani neostrý lokální extrém.

2. Nyní $\nabla f(a) = \overline{0}$. Kvadratickou formu $xH_f(a)x^T$ označíme jako $P(x)$ a f aproximujeme v okolí a Taylorovým polynomem stupně $n = 2$ (podle věty o Taylorově polynomu). Sčítanec $f(a)$ odpovídající $i = 0$ převedeme vlevo, sčítanec s $i = 1$ zmizí, protože $\nabla f(a) = \overline{0}$. Dále je $P(x)$ homogenní polynom stupně 2. Pro $\|h\| \to 0$ tak máme vyjádření

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

kde vektor $e(h) = (h_1/\|h\|, h_2/\|h\|, \ldots, h_m/\|h\|)$ leží na jednotkové sféře S. Jak jsme se již zmínili, S je kompaktní podmnožina $\mathbb{R}^m$ a funkce $P(x) : \mathbb{R}^m \to \mathbb{R}$, která je jistě spojitá na celém $\mathbb{R}^m$, na S nabývá minimum a maximum:

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

pro nějaké dva vektory $\alpha, \beta \in S$. Pozitivní (resp. negativní) definitnost $H_f(a)$ je ekvivalentní nerovnostem $0 < \mu \leq M$ (resp. $\mu \leq M < 0$) a indefinitnost je ekvivalentní nerovnostem $\mu < 0 < M$.

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

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

Takže f má v a ostré lokální minimum. Podobně pro negativně definitní $H_f(a)$ dostáváme v a ostré lokální maximum.

3. Když je $H_f(a)$ indefinitní, pak existuje takové $\delta > 0$, že pro každé $t \in (0, \delta)$ máme

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

Takže f nemá v a lokální extrém. □

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 a nemusí. Konečně zdůrazněme, že se věta týká otevřených množin U, respektive vnitřních bodů množin. Pokud $f:\ M \to \mathbb{R}$, kde $M \subset \mathbb{R}^m$, a bod $a \in M$ není vnitřním bodem M, to jest neleží v M spolu s nějakým svým okolím, pak stále může f mít v a lokální extrém vzhledem k M, i když je gradient $\nabla f(a)$ nenulový vektor. Lokálními extrémy v hraničních bodech množin se budeme zabývat později, v partii o Lagrangeových multiplikátorech.

Poznámky o definitnosti matic. Nechť $A = (a_{i,j}) \in \mathbb{R}^{n\times n}$ je symetrická matice a $P = P(x_1, x_2, \ldots, x_n) = \sum_{i,j=1}^n a_{i,j}x_ix_j$ je jí odpovídající kvadratická forma. Z lineární algebry víme, že existuje taková regulární matice $B = (b_{i,j}) \in \mathbb{R}^{n\times n}$, že po změně proměnných $x^T = By^T$ přejde P do tvaru

$$P(x_1, x_2, \ldots, x_n) = P(b_{1,1}y_1 + \cdots + b_{1,n}y_n, \ldots, b_{n,1}y_1 + \cdots + b_{n,n}y_n) = \sum_{i=1}^{n} d_i y_i^2 \ ,$$

kde $d_i \in \{-1, 0, 1\}$. Ekvivalentně,

$$P(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} d_i (c_{i,1}x_1 + c_{i,2}x_2 + \cdots + c_{i,n}x_n)^2 \ ,$$

kde $C = (c_{i,j}) \in \mathbb{R}^{n\times n}$ je inverzní k B. Počty koeficientů d_i rovných $-1, 0$ a 1 v tomto vyjádření jsou určené jednoznačně maticí A (nezávisejí na změně proměnných B) a udávají tzv. signaturu kvadratické formy P. Jsou-li všechny d_i rovny 1 (resp. -1), je A pozitivně (resp. negativně) definitní. Je-li některé d_i rovno 1 a jiné -1, je A indefinitní. Jsou-li všechny d_i rovny 1 nebo 0, je A pozitivně semidefinitní, a ve zbývajícím případě -1 a 0 je A negativně semidefinitní. Do tohoto tvaru $P =$ součet $\pm$ čtverců lze P při malém počtu proměnných $n = 2$ či $n = 3$ transformovat snadno ručně, viz počítání v následujícím příkladu.

Připomeňme ještě *Sylvestrovo kritérium definitnosti* z lineární algebry: pokud jsou všechny subdeterminanty $d_m = \det(a_{i,j})_{i,j=1}^m$, $1 \le m \le n$, nenulové, pak, jsou-li všechny kladné, je matice A pozitivně definitní, nastává-li $(-1)^m d_m > 0$, $1 \le m \le n$, je A negativně definitní, a jinak je indefinitní; o případu, kdy $d_m = 0$ pro alespoň jedno m, Sylvestrovo kritérium neříká nic.

Příklad. Nalezněte lokální a globální extrémy funkce

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

Řešení (nebylo uvedeno na přednášce). Definiční obor $\mathbb{R}^2$ je otevřená množina, a pro hledání lokálních extrémů tak můžeme bez problémů použít větu o lokálních extrémech. Máme

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

a

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

Vyřešíme soustavu rovnic $\nabla f(x,y) = (0,0)$:

$$-y\sin x - \cos x = 0,\ 2y + \cos x = 0 \ .$$

Sečtením rovnic obdržíme $y(2 - \sin x) = 0$. Nutně (neboť $|\sin x| \le 1$) $y = 0$, a tedy $\cos x = 0$. Dostáváme stacionární body

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

Podle věty o lokálních extrémech má f lokální extrémy pouze v těchto bodech.

Máme

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

to jest

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

První matice je indefinitní, protože

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

a druhá je pozitivně definitní, protože

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

Pro liché k není v s_k 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 minima. Definiční obor $\mathbb{R}^2$ není kompaktní (není omezený), nelze hned použít větu o extrémech spojitých funkcí na kompaktech. Funkce f je však 2π-periodická v x (tj. $f(x \pm 2\pi, y) = f(x,y)$ pro každé $x, y \in \mathbb{R}$) a pro vyšetření globálních minim stačí uvážit její hodnoty ve svislém nekonečném pásu

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

Na jeho hranici máme

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

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

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

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 $f(a) \geq -1 > f(s_0) = -3$, přičemž $s_0 \in P_1$. Všechny hodnoty f na P_2 jsou větší než hodnota f v bodu s_0. Na hranici obdélníka P_1 má f vždy hodnotu alespoň $\min(-9/4,-1) = -9/4 > -3$ a na jeho vnitřku, což je otevřená množina, má f jediné lokální minimum $f(s_0) = -3$. Proto má f na obdélníku P_1 i na celém pásu P jediné ostré globální minimum $f(s_0) = -3$. Z 2π-periodičnosti v proměnné x plyne, že hodnoty $f(s_{2k}) = -3$, $k \in \mathbb{Z}$, jsou právě všechna neostrá globální minima funkce f na $\mathbb{R}^2$.

Závěr. Jediné lokální extrémy f jsou ostrá lokální minima $f(s_{2k}) = -3$ v (nekonečně mnoha) bodech $s_{2k} = (\frac{\pi}{2} + 2k\pi, 0), k \in \mathbb{Z}$. Tyto body jsou i body neostrého globálního minima funkce f. Globální maximum f nemá. □

Implicitní funkce. Jak víme z lineární algebry, soustava n lineárních rovnic o n neznámých $a_{i,1}y_1 + a_{i,2}y_2 + \cdots + a_{i,n}y_n + b_i = 0$, $i = 1,2,\ldots,n$, kde $a_{i,j}, b_i \in \mathbb{R}$ jsou dané a $\det(a_{i,j})_{i,j=1}^n \neq 0$, má pro každou volbu n konstant b_i jednoznačné řešení $y_1, y_2, \ldots, y_n$. Navíc toto řešení y_j je jakožto funkce volených konstant b_i homogenní lineární funkce: $y_j(b_1, b_2, \ldots, b_n) = c_{j,1}b_1 + c_{j,2}b_2 + \cdots + c_{j,n}b_n$, $j = 1,2,\ldots,n$, pro jisté konstanty $c_{j,i} \in \mathbb{R}$ (to plyne z Cramerova vzorce vyjadřujícího řešení nehomogenní lineární soustavy ve tvaru podílu dvou determinantů).

Tento výsledek nyní zobecníme na situaci, kdy jsou lineární funkce nahrazeny obecnými funkcemi a kdy v každé rovnici lze předem zvolit více než jeden parametr. Vezmeme soustavu n rovnic o $m+n$ neznámých

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

kde F_i jsou reálné funkce definované na okolí bodu (x_0,y_0) v $\mathbb{R}^{m+n}$, kde $x_0 \in \mathbb{R}^m$ a $y_0 \in \mathbb{R}^n$, který je řešením této soustavy, to jest $F_1(x_0,y_0) = F_2(x_0,y_0) = \cdots = F_n(x_0,y_0) = 0$. Jak uvidíme, za jistých předpokladů lze neznámé $y_1,y_2,\ldots,y_n$ ze soustavy eliminovat a vyjádřit je, lokálně v okolí x_0, jako funkce $y_i = f_i(x_1,x_2,\ldots,x_m)$ neznámých $x_1,x_2,\ldots,x_m$. Nejprve ale zavedeme značení. Pro zobrazení $F = (F_1,F_2,\ldots,F_n)$ a $f = (f_1,f_2,\ldots,f_n)$, přičemž $F_i = F_i(x_1,\ldots,x_m,y_1,\ldots,y_n)$ a $f_j = f_j(x_1,\ldots,x_m)$, označíme $x = (x_1,x_2,\ldots,x_m)$, $y = (y_1,y_2,\ldots,y_n)$ a

$$F'_x(x,y) = \left(\frac{\partial F_i}{\partial x_j}\right)_{i,j=1}^{n,m}(x,y) \;=\; \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots & \frac{\partial F_1}{\partial x_m} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial F_n}{\partial x_1} & \frac{\partial F_n}{\partial x_2} & \cdots & \frac{\partial F_n}{\partial x_m} \end{pmatrix}(x,y)$$

$$F'_y(x,y) = \left(\frac{\partial F_i}{\partial y_j}\right)_{i,j=1}^{n}(x,y) \;=\; \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \frac{\partial F_1}{\partial y_2} & \cdots & \frac{\partial F_1}{\partial y_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial F_n}{\partial y_1} & \frac{\partial F_n}{\partial y_2} & \cdots & \frac{\partial F_n}{\partial y_n} \end{pmatrix}(x,y)$$

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

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

Věta (o implicitních funkcích). *Nechť*

$$F = (F_1,F_2,\ldots,F_n) :\; W \to \mathbb{R}^n$$

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

1. $F_i = F_i(x, y) \in \mathcal{C}^1(W)$ *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 $U \subset \mathbb{R}^m$ *a* $V \subset \mathbb{R}^n$ *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* $1 \le i \le n$*. 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* $\mathcal{C}^1(U)$*, 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

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

a z $f_i \in \mathcal{C}^1(U)$ 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

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

To je soustava n rovnic s n neznámými $\partial_i f_j(x)$, $1 \le j \le n$, 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 $(x, f(x)) \in U \times V$).

Vícerozměrný Riemannův integrál a Fubiniova věta

(sepsáno podle 11. kapitoly knihy V. A. Zoricha, *Mathematical Analysis II*, Springer, 2004)

Boxy. Box, přesněji n-rozměrný *box* $I \subset \mathbb{R}^n$, je kartézský součin intervalů

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

kde $-\infty < a_i < b_i < +\infty$. Např. v euklidovské rovině $\mathbb{R}^2$ to je uzavřený obdélník se stranami rovnoběžnými s osami souřadnic a s kladnými délkami. *Objem boxu* je

$$|I| := \prod_{i=1}^{n} (b_i - a_i) \ .$$

Dělení D boxu I na podboxy je množina boxů

$$D = \{[c_1^{j_1}, c_1^{j_1+1}] \times [c_2^{j_2}, c_2^{j_2+1}] \times \cdots \times [c_n^{j_n}, c_n^{j_n+1}] \mid 0 \le j_i < k_i, 1 \le i \le n\} \ ,$$

kde $a_i = c_i^0 < c_i^1 < \cdots < c_i^{k_i-1} < c_i^{k_i} = b_i$ jsou nějaká dělení intervalů $[a_i, b_i]$, $i = 1, 2, \ldots, n$. *Norma dělení je*

$$\lambda(D) := \max_{1 \le i \le n,\, 0 \le j < k_i} (c_i^{j+1} - c_i^j)$$

— maximální délka hrany podboxu. *Dělení D boxu I s body ζ* je dvojice (D, ζ), kde D je dělení boxu I a $\zeta :\ D \to \mathbb{R}^n$ je zobrazení splňující $\zeta(J) \in J$ pro každý podbox J. Prostě řečeno, v každém podboxu je zvolený nějaký bod.

Riemannova definice vícerozměrného integrálu. Nechť $I \subset \mathbb{R}^n$ je box, (D, ζ) je jeho dělení s body a $f :\ I \to \mathbb{R}$ je funkce. *Riemannova suma* je definována jako

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

Integrál funkce f přes box I je vlastní limita

$$\int_I f := \lim_{(D,\zeta),\, \lambda(D) \to 0} S(f, D, \zeta) \ ,$$

existuje-li, takže $\int_I f \in \mathbb{R}$ je takové číslo, že

$$\forall \varepsilon\, \exists \delta > 0\, \forall (D, \zeta) :\ \lambda(D) < \delta \Rightarrow \left| \int_I f - S(f, D, \zeta) \right| < \varepsilon \ .$$

Darbouxova definice vícerozměrného integrálu. Nechť D je dělení boxu I a $f : I \to \mathbb{R}$ je funkce. Pak definujeme pro každý podbox J tohoto dělení: $m(J) = \inf_{x \in J} f(x)$, $M(J) = \sup_{x \in J} f(x)$ a *dolní, resp. horní, součet*

$$s(f, D) := \sum_{J \in D} |J| \cdot m(J), \text{ resp. } S(f, D) := \sum_{J \in D} |J| \cdot M(J) .$$

Dolní, resp. horní, integrál je

$$\underline{\int_I} f = \sup(\{s(f, D) \mid D \text{ je dělení } I\}) ,$$

resp.

$$\overline{\int_I} f = \inf(\{S(f, D) \mid D \text{ je dělení } I\}) .$$

Opět (jako v jedné dimenzi) platí, že pro každé dělení D boxu I je

$$-\infty \leq s(f, D) \leq \underline{\int_I} f \leq \overline{\int_I} f \leq S(f, D) \leq +\infty .$$

Integrál funkce f přes box I pak definujeme jako reálné číslo

$$\int f = \underline{\int_I} f = \overline{\int_I} f \in \mathbb{R} ,$$

když se dolní a horní integrál rovnají společné vlastní hodnotě.

Platí: f má $\int$ podle Riemannovy definice $\iff$ f má $\int$ podle Darbouxovy definice, a v případě existence integrálu se obě hodnoty rovnají. Množinu funkcí riemannovsky integrovatelných přes box I označíme jako

$$\mathcal{R}(I) = \{f \mid f \text{ má Riemannův integrál přes } I\} .$$

Lebesgueova věta. Řekneme, že množina $E \subset \mathbb{R}^n$ má (n-rozměrnou Lebesgueovu) míru 0, když pro každé $\varepsilon > 0$ existuje taková posloupnost boxů $I_1, I_2, \ldots$ v $\mathbb{R}^n$, že

$$\sum_{n=1}^{\infty} |I_n| < \varepsilon \text{ a } E \subset \bigcup_{n=1}^{\infty} I_n .$$

Věta (Lebesgueova). *Nechť $I \subset \mathbb{R}^n$ je box a $f : I \to \mathbb{R}$ je na něm definovaná fukce. Pak $f \in \mathcal{R}(I) \iff f$ je na I omezená a množina jejích bodů nespojitosti má míru 0.*

Například každá omezená funkce $f:\ I \to \mathbb{R}$ nespojitá pouze ve spočetně mnoha bodech má Riemannův $\int_I f$. Z Lebesgueovy věty a definice integrálu dostáváme (dokažte si to jako úlohu 1):

Důsledek. *Když $f:\ I \to \mathbb{R}$ je funkce, jež je na boxu $I \subset \mathbb{R}^n$ nezáporná a $\int_I f = 0$, potom $f = 0$ na I až na množinu bodů s mírou 0.*

Fubiniova věta. (Přesněji, věta Fubiniova typu.) Umožňuje převést výpočet vícerozměrného integrálu na posloupnost obyčejných jednorozměrných integrálů. Nechť $X \subset \mathbb{R}^m$, $Y \subset \mathbb{R}^n$ a $Z = X \times Y \subset \mathbb{R}^{m+n}$ je m-rozměrný, n-rozměrný a $(m+n)$-rozměrný box.

Věta (Fubiniova). *Nechť $f:\ Z = X \times Y \to \mathbb{R}$, $f \in \mathcal{R}(Z)$. Pak všechny tři integrály*

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

existují a rovnají se.

Vysvětlíme značení a smysl věty. 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 $x = x_0 \in X$ tento integrál neexistuje, definujeme $F(x_0)$ jako libovolnou hodnotu z intervalu $[\underline{\int_Y} f(x_0,y)\ dy, \overline{\int_Y} f(x_0,y)\ dy]$. Podobným způsobem definujeme funkci

$$G:\ Y \to \mathbb{R},\ G(y) := \int_X f(x,y)\,dx\ .$$

Fubiniova věta říká, že $F \in \mathcal{R}(X)$, $G \in \mathcal{R}(Y)$ a $\int_Z f = \int_X F = \int_Y G$. Z důkazu též vyplyne (úloha 6), že množina bodů $x_0 \in X$ s $f(x_0,y) \notin \mathcal{R}(Y)$ má míru 0, a podobně v y-ové souřadnici.

Důkaz. Dokážeme, že $F \in \mathcal{R}(X)$ a $\int_Z f = \int_X F$, pro funkci 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, to jest každý box $J \in D$ je součinem $J = J_1 \times J_2$

pro $J_1 \in D_1$, $J_2 \in D_2$. Buď dáno $\varepsilon > 0$. Pak existuje 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 vlastností infima platí, že

$$\begin{aligned}
s(f, D) &= \sum_{J \in D} |J| \cdot \inf_{z \in J} f(z) = \sum_{J \in D} \overbrace{|J_1 \times J_2|}^{=|J_1| \cdot |J_2|} \cdot \inf_{x \in J_1, y \in J_2} f(x, y) \\
\text{(proč? — úloha 2)} &\le \sum_{J_1 \in D_1} |J_1| \cdot \inf_{x \in J_1} \Bigg(\underbrace{\sum_{J_2 \in D_2} |J_2| \inf_{y \in J_2} f(x, y)}_{=s(f(x,\cdot), D_2) \le \underline{\int}_Y f(x,y)\,dy \le F(x)} \Bigg) \\
&\le \sum_{J_1 \in D_1} |J_1| \cdot \inf_{x \in J_1} F(x) = s(F, D_1)\ .
\end{aligned}$$

Takže $s(F, D_1) > \int_Z f - \varepsilon$. Analogicky se dokáže, že pro dané $\varepsilon > 0$ existuje dělení D_1' boxu X, že $S(F, D_1') < \int_Z f + \varepsilon$. Pro $\varepsilon \to 0$ to podle Darbouxovy definice integrálu znamená, že $F \in \mathcal{R}(X)$ a $\int_X F = \int_Z f$. □

Příklad. Nechť $f(x, y, z) = z \sin(x + y)$ a box $I \subset \mathbb{R}^3$ je dán intervaly $0 \le x \le \pi$, $|y| \le \pi/2$ a $0 \le z \le 1$. Pak, podle Fubiniovy věty,

$$\begin{aligned}
\int\!\!\int\!\!\int_I f(x, y, z)\, dx\, dy\, dz &= \int_0^1 \left(\int_{-\pi/2}^{\pi/2} \left(\int_0^\pi z \sin(x + y)\, dx \right) dy \right) dz \\
&= \int_0^1 \left(\int_{-\pi/2}^{\pi/2} (-z \cos(x + y)|_{x=0}^{\pi})\, dy \right) dz \\
&= \int_0^1 \left(\int_{-\pi/2}^{\pi/2} 2z \cos y\, dy \right) dz \\
&= \int_0^1 (2z \sin y|_{y=-\pi/2}^{\pi})\, dz = \int_0^1 4z\, dz = 2\ .
\end{aligned}$$

Integrál přes množinu $E \subset \mathbb{R}^n$. Integrál rozšíříme z boxů na obecnější množiny. Zavedeme i objem množiny. Množina $E \subset \mathbb{R}^n$ je *přípustná*, když je omezená a její hranice ∂E (což jsou ty body $x \in \mathbb{R}^n$, že každé okolí x protíná jak E tak $\mathbb{R}^n \backslash E$) má míru 0. Například krychle, otevřená či uzavřená koule jsou přípustné množiny, kdežto $\mathbb{Q} \cap [0, 1]^n$ není přípustná množina. *Objem* omezené množiny $E \subset \mathbb{R}^n$ je integrál (když existuje)

$$\mathrm{vol}(E) := \int_I \chi_E\ ,$$

kde $I \subset \mathbb{R}^n$ je box obsahující E a χ_E je charakteristická funkce množiny E (takže $\chi_E(x) = 1$ pro $x \in E$ a $\chi_E(x) = 0$ pro $x \in I \backslash E$). Dá se dokázat:

Tvrzení. *$E \subset \mathbb{R}^n$ má objem $\iff$ E je přípustná.*

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

$$\int_E f := \int_I \overline{f} \;,$$

kde $I \subset \mathbb{R}^n$ je box obsahující E a $\overline{f}$ je rozšíření f:

$$\overline{f}(x) = \begin{cases} f(x) & \dots \quad x \in E \\ 0 & \dots \quad x \in I \backslash E \;. \end{cases}$$

Tato definice (jakož i definice objemu) je korektní: úloha 7.

Úlohy

1. Dokažte důsledek Lebesgueovy věty.

2. Proč platí ta nerovnost v důkazu Fubiniovy věty?

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

4. Nechť $I \subset \mathbb{R}^n$ je box a $I = \bigcup_{i=1}^k J_i$ je sjednocení boxů, které mají disjunktní vnitřky. Dokažte, že $|I| = \sum_{i=1}^k |J_i|$.

5. Nechť $f : I \to \mathbb{R}$ je neomezená funkce definovaná na boxu v $\mathbb{R}^n$. Co se stane v Riemannově definici integrálu? Jak vypadají $\underline{\int_I} f$ a $\overline{\int_I} f$?

6. Zdůvodněte, proč ve Fubiniově větě množina

   $$\{x_0 \in X \mid \textstyle\int_Y f(x_0, y)\, dy \text{ neexistuje}\}$$

   má míru nula.

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

8. Dokažte, že pro každý box $I \subset \mathbb{R}^n$ je $|I| = \mathrm{vol}(I)$.

Přednáška 11, 15. května 2015

Jako příklad na větu o implicitních funkcích ukážeme, že soustava rovnic

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

definuje v okolí bodu $x = 0$ dvě funkce $y = y(x)$ a $z = z(x)$ třídy $\mathcal{C}^1$ s hodnotami $y(0) = z(0) = 0$ a spočteme hodnoty derivací $y'(0)$ a $z'(0)$.

Položíme $F_1(x, y, z) = x + y - \sin z$, $F_2(x, y, z) = -x^3 - y^3 + \mathrm{e}^z - 1$ a $F = (F_1, F_2)$. Skutečně $F(0, 0, 0) = (0, 0)$ a jacobián soustavy J je nenulový:

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

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

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

podle vztahů uvedených na konci předešlé přednášky máme

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

Vázané extrémy. Z věty o implicitních funkcích lze odvodit (pro důkaz však nemáme čas) zobecnění první části věty o lokálních extrémech — nutná podmínka pro lokální extrém funkce v bodě otevřené množiny je nulovost všech parciálních derivací — na extrémy na množině zadané soustavou rovnic. Nechť $U \subset \mathbb{R}^m$ je otevřená množina a

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

jsou funkce z $\mathcal{C}^1(U)$, 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 větu o lokální extrémech nelze použít. Příkladem je jednotková sféra v $\mathbb{R}^m$:

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

Následující tvrzení udává nutnou podmínku pro existenci lokálního extrému funkce f v bodě množiny H.

Důsledek (Lagrangeovy multiplikátory). *Nechť $a \in H$. Jsou-li vektory $\nabla F_1(a), \ldots, \nabla F_n(a)$ z $\mathbb{R}^m$ lineárně nezávislé a vektor $\nabla f(a)$ není jejich lineární kombinací, pak f nemá v bodu a vzhledem k množině H ani neostrý lokální extrém.*

Ekvivalentně: jsou-li $\nabla F_1(a), \ldots, \nabla F_n(a)$ lineárně nezávislé a funkce f má v bodě a vzhledem k množině H (ostrý či neostrý) lokální extrém, potom existují čísla $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, tzv. Lagrangeovy multiplikátory*, že*

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

to jest $\partial_{x_j} f(a) - \lambda_1 \partial_{x_j} F_1(a) - \cdots - \lambda_n \partial_{x_j} F_n(a) = 0$ pro $1 \leq j \leq m$.

Pro ilustraci metody si spočteme dva jednoduché příklady. V **prvním příkladu** nalezneme extrémy funkce $f(x, y) = x + y$ vzhledem k množině $H \subset \mathbb{R}^2$ dané rovnicí

$$H: \ F(x, y) = x^2 + y^2 - 1 = 0 ,$$

což je jednotková kružnice se středem v počátku. Máme $\nabla F = (2x, 2y)$ a $\nabla f = (1, 1)$. Patrně $\nabla F = \overline{0}$ pouze v $\overline{0} \not\in H$, tedy $\nabla F \neq \overline{0}$ na H a předpoklad lineární nezávislosti gradientů rovnicových funkcí je splněn. Podle Lagrangeových multiplikátorů jsou body, v nichž má f lokální extrém vzhledem k H, obsaženy v řešeních soustavy

$$x^2 + y^2 - 1 = 0, \ 1 = 2\lambda x \ \text{ a } \ 1 = 2\lambda y .$$

Odečtením posledních dvou rovnic dostáváme $\lambda(x - y) = 0$. Protože λ nemůže být 0, je $x = y$. Dosazením do první rovnice dostaneme, že $x = y = \lambda = \pm 1/\sqrt{2}$, což dává přesně dvě řešení dané soustavy. Podezřelé body tak jsou

$$a = (-1/\sqrt{2}, -1/\sqrt{2}) \ \text{ a } \ b = (1/\sqrt{2}, 1/\sqrt{2}) .$$

Množina H je kompaktní a f je na ní spojitá, a proto f nabývá na H minimum i maximum. To jsou i lokální extrémy f. Protože $f(a) = -\sqrt{2}$ a $f(b) = \sqrt{2}$, má f na H v a globální minimum a v b globální maximum.

Druhý příklad ukazuje, že předpoklad lineární nezávislosti gradientů rovnicových funkcí nelze pominout. Vezměme si množinu $H \subset \mathbb{R}^2$ danou rovnicí

$$H: \ F(x,y) = y^2 - x^3 = 0 \ ,$$

což je sjednocení grafů fukcí $y = x^{3/2}$ a $y = -x^{3/2}$ pro $x \geq 0$. Hledejme na H extrémy funkce $f(x,y) = x$. Máme $\nabla F = (-3x^2, 2y)$ a $\nabla f = (1,0)$. Pro podezřelé body dávají Lagrangeovy multiplikátory soustavu

$$y^2 - x^3 = 0, \ 1 = -3\lambda x^2 \ \text{ a } \ 0 = 2\lambda y \ .$$

Ta nemá řešení (z třetí rovnice je $\lambda = 0$ nebo $y = 0$ a obojí vede ke sporu s druhou rovnicí). Takže f nemá na H lokální extrém. To je však **chybný závěr**, protože f má zjevně v bodě $(0,0) \in H$ na H ostré globální minimum s hodnotou $f = 0$. Udělali jsme tu chybu, že jsme neověřili nenulovost vektoru ∇F v bodě $(0,0)$. A právě v něm se gradient ∇F anuluje, tj. předpoklad lineární nezávislosti v něm není splněn.

Trochu o metrických prostorech

Metrický prostor je struktura formalizující jev vzdálenosti. Je to dvojice (M,d) složená z množiny $M \neq \emptyset$ a funkce dvou proměnných

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

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). Uvedeme si pár příkladů metrických prostorů. Axiomy a) a b) se ověří obvykle snadno. Dokázat trojúhelníkovou nerovnost bývá často obtížnější.

Příklad 1. $M = \mathbb{R}^n$ 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 $\mathbb{R}$ a pro $p = 2, n \geq 2$ *euklidovskou metriku*

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

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 $f :\ X \to \mathbb{R}$ 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. Pro souvislý graf $G = (M, E)$ s množinou vrcholů M máme metriku

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

Příklad 4. Je-li A množina (abeceda), máme na množině $M = A^m$ slov délky m nad abecedou A tak zvanou *Hammingovu metriku* ($u = a_1 a_2 \ldots a_m, v = b_1 b_2 \ldots b_m$)

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

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

V rychlosti zavedeme pár základních pojmů; s mnohými jsme se již setkali u eukleidovských prostorů. Nechť (M, d) je metrický prostor. Pak

- *(otevřená) koule* v M se středem v bodu $a \in M$ a poloměrem $\mathbb{R} \ni r > 0$ je množina $B(a, r) = \{x \in M \mid d(a, x) < r\}$;
- $A \subset M$ je *otevřená množina*, pokud $\forall a \in A \; \exists r > 0 : \; B(a, r) \subset A$;
- $A \subset M$ je *uzavřená množina*, je-li $M \backslash A$ otevřená množina;
- $A \subset M$ je *omezená množina*, pokud existuje bod $a \in M$ a poloměr $r > 0$, že $A \subset B(a, r)$;
- $A \subset M$ je *kompaktní množina*, pokud každá posloupnost bodů $(a_n) \subset A$ má konvergentní podposloupnost, jejíž limita leží v A.

Konvergence a limita se zobecňují z reálné osy na obecný metrický prostor zřejmým způsobem: posloupnost $(a_n) \subset M$ je *konvergentní a za limitu má bod* $a \in M$, psáno $\lim_{n\to\infty} a_n = a$, když

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

Jinak řečeno, $\lim_{n\to\infty} d(a_n, a) = 0$ (převedli jsme to na limitu reálné posloupnosti).

Již jsme dříve zmínili vlastnosti otevřených množin: $\emptyset$ i M jsou otevřené, sjednocení libovolného systému otevřených množin je otevřená množina a průnik libovolného konečného systému otevřených množin je otevřená množina (důkazy si rozmyslete jako cvičení). Přechodem k doplňkům máme duální vlastnosti uzavřených množin: $\emptyset$ i M jsou uzavřené, sjednocení libovolného konečného systému uzavřených množin je uzavřená množina a průnik libovolného systému uzavřených množin je uzavřená množina. Následující tvrzení ukazuje, že uzavřené množiny jsou uzavřené na limity.

Tvrzení (charakterizace uzavřených množin). *$A \subset M$ je uzavřená množina v metrickém prostoru M, právě když limita každé její konvergentní podposloupnosti $(a_n) \subset A$ leží v A.*

Důkaz. Nechť $A \subset M$ je uzavřená množina a $(a_n) \subset A$ je konvergentní posloupnost. Kdyby $\lim_{n\to\infty} a_n = a \notin A$, existoval by poloměr $r > 0$, že

$B(a,r) \subset M \backslash A$. Pak ale $d(a_n, a) \geq r$ pro každé n, ve sporu s $\lim_{n\to\infty} a_n = a$. Tedy $a \in A$. Naopak, není-li $A \subset M$ uzavřená množina, podle definice existuje takový bod $a \in M \backslash A$, že pro každý poloměr $r > 0$ je $B(a,r) \cap A \neq \emptyset$. Položíme $r = 1/n$, $n = 1, 2, \ldots$, a pro každé n zvolíme libovolně bod $a_n \in B(a, 1/n) \cap A$. Pak $(a_n) \subset A$ a je to konvergentní posloupnost s $\lim_{n\to\infty} a_n = a$, ale $a \notin A$. $\square$

Přednáška 12, 22. května 2015

Topologické prostory. Podle sylabu se zmíníme o další abstraktní struktuře. Dvojice $T = (X, \mathcal{T})$, kde X je množina a $\mathcal{T}$ systém jejích podmnožin, je *topologický prostor*, má-li $\mathcal{T}$ tyto vlastnosti: (i) $\emptyset, X \in \mathcal{T}$, (ii) pro každý podsystém $\mathcal{U} \subset \mathcal{T}$ je i $\bigcup \mathcal{U} \in \mathcal{T}$ a (iii) pro každý konečný podsystém $\mathcal{U} \subset \mathcal{T}$ je i $\bigcap \mathcal{U} \in \mathcal{T}$. Množinám v systému $\mathcal{T}$ se říká *otevřené množiny* topologického prostoru T (jejich doplňky do X pak jsou *uzavřené množiny* prostoru T). Příkladem topologického prostoru jsou otevřené množiny každého metrického prostoru. Ovšem je spousta topologických prostorů, které nejsou metrizovatelné (tj. nepocházejí z metrického prostoru). Cvičení: vymyslete příklady takových topologických prostorů. Návod: metrizovatelná topologie má vždy tu vlastnost, že pro každé dva různé body a, b máme takové dvě otevřené množiny A, B, že $a \in A, b \in B$ a $A \cap B = \emptyset$.

Spojitá zobrazení. Jsou-li (M, d) a (N, e) dva metrické prostory, je zobrazení

$$f:\ M \to N$$

spojité, pokud

$$\forall a \in M, \varepsilon > 0\ \exists \delta > 0:\ b \in M, d(a, b) < \delta \Rightarrow e(f(a), f(b)) < \varepsilon\ .$$

Ekvivalentně můžeme spojitost definovat v Heineho stylu: f je spojité, právě když pro každou konvergentní posloupnost $(a_n) \subset M$ s limitou $a \in M$ je i $(f(a_n)) \subset N$ konvergentní a má limitu $f(a) \in N$. Jiná ekvivalentní definice spojitosti je následující.

Tvrzení (topologická definice spojitosti). *Zobrazení $f:\ M \to N$ mezi metrickými prostory je spojité, právě když pro každou otevřenou množinu $B \subset N$ je i její vzor $f^{-1}(B) = \{x \in M \mid f(x) \in B\}$ otevřená množina v M.*

Kompaktní podmnožinu A v metrickém prostoru jsme si definovali požadavkem, aby každá posloupnost bodů v A měla podposloupnost konvergentní v A (tj. konvergentní podposloupnost s limitou ležící v A). Jiná ekvivalentní definice kompaktnosti je následující.

Tvrzení (topologická definice kompaktnosti). *Množina $A \subset M$ v metrickém prostoru je kompaktní, právě když v každém systému $\mathcal{T}$ otevřených* *množin v M splňujícím $\bigcup \mathcal{T} \supset A$ existuje konečný podsystém $\mathcal{U} \subset \mathcal{T}$, který stále pokrývá A: $\bigcup \mathcal{U} \supset A$.*

Předchozí definice kompaktnosti se vyslovuje takto: každé otevřené pokrytí má konečné podpokrytí.

Tvrzení (spojité zobrazení zachovává kompaktnost). *Je-li zobrazení $f : M \to N$ mezi metrickými prostory spojité a M je kompaktní, je obraz $f(M) = \{f(x) \mid x \in M\}$ kompaktní podmnožina N.*

Důkaz. Je jednoduchý. Je-li $(b_n) \subset f(M)$ libovolná posloupnost, máme posloupnost $(a_n) \subset M$, že $f(a_n) = b_n$. Protože M je kompaktní, (a_n) má konvergentní podposloupnost (a_{k_n}) s limitou $a \in M$. Protože f je spojité zobrazení,

$$\lim_{n\to\infty} b_{k_n} = \lim_{n\to\infty} f(a_{k_n}) = f\big(\lim_{n\to\infty} a_{k_n}\big) = f(a)\ .$$

Takže (b_{k_n}) je konvergentní podposloupnost posloupnosti (b_n) s limitou $f(a) \in f(M)$. Tedy $f(M)$ je kompaktní. □

Tvrzení (kompaktnost ⇒ uzavřenost a omezenost). *Každá kompaktní množina v metrickém prostoru je uzavřená a omezená.*

Důkaz. Nechť $A \subset M$ je podmnožina v metrickém prostoru (M, d). Když A není uzavřená, existuje konvergentní posloupnost $(a_n) \subset A$, jejíž limita a leží mimo A. Každá podposloupnost (a_n) je zřejmě též konvergentní a má tutéž limitu a. To ale znamená, že žádná podposloupnost (a_n) není konvergentní v rámci A (limita je určena jednoznačně) a A není kompaktní. Když A není omezená, není obsažena v žádné kouli $B(a, r)$ a snadno sestrojíme posloupnost $(a_n) \subset A$ s vlastností, že $d(a_m, a_n) \geq 1$ pro každé dva indexy $1 \leq m < n$. Tato vlastnost popírá konvergenci posloupnosti (proč?) a má ji i každá podposloupnost (a_n), takže (a_n) nemá žádnou konvergentní podposloupnost. A opět není kompaktní.

Posloupnost $(a_n) \subset A$ s uvedenou vlastností setrojíme indukcí. První bod $a_1 \in A$ vezmeme libovolně. Nechť už máme body $a_1, a_2, \ldots, a_r$ z A, z nichž každé dva mají vzdálenost alespoň 1. Pak vezmeme libovolnou kouli $B(a, r)$, která obsahuje všechny tyto body (každá konečná množina je omezená) a uvážíme kouli $B(a, r + 1)$. Protože A není omezená, existuje bod $a_{r+1} \in A$, který není v $B(a, r+1)$. Podle trojúhelníkové nerovnosti je $d(a_{r+1}, x) \geq 1$ pro každý bod $x \in B(a, r)$ (proč?). Tedy a_{r+1} má od každého bodu $a_1, a_2, \ldots, a_r$ vzdálenost alespoň 1 a $a_1, a_2, \ldots, a_r$ můžeme prodloužit na $a_1, a_2, \ldots, a_r, a_{r+1}$.

Takto definovaná posloupnost a_r, $r = 1, 2, \dots$, má tedy požadovanou vlastnost. $\square$

Asi nejjednodušší příklad ukazující, ze opačná implikace obecně neplatí je tento. Nechť (M, d) je triviální metrický prostor, kde $d(x, y) = 1$ pro $x \neq y$ a $d(x, x) = 0$ (ověřte, že jde o metrický prostor), a množina M je nekonečná. Pak každá posloupnost $(a_n) \subset M$, kde a_n jsou vzájemně různé body (pro existenci takové posloupnosti potřebujeme nekonečnost M), splňuje, že $d(a_m, a_n) \geq 1$ pro každé dva indexy $1 \leq m < n$. Jak víme, taková posloupnost nemá žádnou konvergentní podposloupnost a proto M není kompaktní množina. Ovšem M je uzavřená množina (je to celý prostor) a je i omezená, protože patrně $M \subset B(a, 2)$ pro každý bod $a \in M$.

Jak už jsme se na dřívější přednášce zmínili, opačná implikace platí pro eukleidovské prostory.

Věta (uzavřenost a omezenost $\Rightarrow$ kompaktnost v $\mathbb{R}^k$). *Každá uzavřená a omezená množina v eukleidovském prostoru $\mathbb{R}^k$ je kompaktní.*

Důkaz. Nechť $A \subset \mathbb{R}^k$ je uzavřená a omezená podmnožina eukleidovského prostoru. Díky omezenosti A pro dostatečně velké $a > 0$ máme $A \subset [-a, a]^k$ (A je obsažena v dostatečně velké k-rozměrné krychli se středem v počátku). Souřadnice bodů $x \in \mathbb{R}^k$ označíme jako $x = (x(1), x(2), \dots, x(k))$. Nechť $(a_n) \subset A$ je libovolná posloupnost. Podle věty ze ZS má (a_n) podposloupnost (b_n), která konverguje v prvních souřadnicích (tyto souřadnice leží v kompaktním intervalu $[-a, a]$). Z posloupnosti (b_n) vybereme podposloupnost (c_n), která konverguje v druhých souřadnicích, a tak dál. Po k takových výběrech dostaneme posloupnost řekněme (d_n), jež je podposloupností posloupnosti (a_n) a konverguje v každé z k souřadnic:

$$\lim_{n\to\infty} d_n(j) = e(j) \in \mathbb{R},\ j = 1, 2, \dots, k\ .$$

Lehce se vidí, že pak i $\lim_{n\to\infty} d_n = e = (e(1), e(2), \dots, e(k)) \in \mathbb{R}^k$. (To plyne třeba z nerovnosti $\|e - d_n\| \leq \sum_{j=1}^{k} |e(j) - d_n(j)|$.) Protože A je uzavřená množina, $e \in A$ a (d_n) je podposloupnost (b_n), která konverguje v A. Takže A je kompaktní. $\square$

Věta (spojitá funkce nabývá na kompaktu extrém). *Nechť $f: M \to \mathbb{R}$ je spojitá funkce z metrického prostoru (M, d) do eukleidovského prostoru $\mathbb{R}^1$ a M je kompaktní. Pak f nabývá na M minimum i maximum.*

Důkaz. Podle jednoho z předchozích tvrzení je obraz $f(M)$ kompaktní podmnožina v $\mathbb{R}$. Tedy (podle dalšího z předchozích tvrzení) je $f(M) \subset \mathbb{R}$ uzavřená a omezená množina. Protože $f(M)$ je neprázdná a shora omezená, existuje supremum $h = \sup(f(M)) \in \mathbb{R}$. Podle aproximační vlastnosti suprema je h limitou bodů z množiny $f(M)$. Díky uzavřenosti $f(M)$ je $h \in f(M)$ a vlastně $h = \max(f(M))$. Takže f nabývá na M maximum. Pro mimimum argumentujeme stejně pomocí infima. □

Přednášku zakončíme aplikací tohoto výsledku v důkazu tzv. Základní věty algebry, že každý nekonstantní komplexní polynom má kořen. Důkaz sám jsem z časových důvodů na přednášce neuvedl. Pro jeho zajímavost a kvůli důležitost celého výsledku ho uvádím zde. Komplexní rovina $\mathbb{C}$ se v něm bere jako eukleidovský metrický prostor $\mathbb{R}^2$ s obvyklou vzdáleností.

Věta (Základní věta algebry). *Pro každý komplexní polynom* $p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$ *stupně alespoň* 1 *(takže* $a_j \in \mathbb{C}$*,* $a_n \neq 0$ *a* $n \geq 1$*) existuje komplexní číslo* $\alpha \in \mathbb{C}$*, že* $p(\alpha) = 0$*.*

Důkaz. Nechť $a \in \mathbb{C}$ s $a \neq 0$ a $k \in \mathbb{N}$. Připomeneme si chování komplexní funkce $z \mapsto az^k$.

- Když z probíhá v komplexní rovině $\mathbb{C}$ kružnici $K = \{z \mid |z| = r > 0\}$, pak az^k probíhá kružnici $L = \{z \mid |z| = |a|r^k > 0\}$, přičemž jednomu proběhnutí K odpovídá k proběhnutí L; speciálně má každý bod v L v této funkci přesně k vzorů v K.

To plyne z goniometrického tvaru nenulového komplexního čísla: $K \ni z = r\exp(\varphi i) = r(\cos\varphi + i\sin\varphi)$, kde $r = |z| > 0$ je modul čísla z a úhel $\varphi \in [0, 2\pi)$ je jeho argument. Odtud máme následující fakt, klíčový pro důkaz:

- Když $\alpha \in \mathbb{C}$ a $|p(\alpha)| > 0$, pak existuje číslo $\beta \in \mathbb{C}$, že $|p(\beta)| < |p(\alpha)|$.

Dokažme to. Nechť $|p(\alpha)| > 0$. Můžeme předpokládat, že $\alpha = 0$. (Jinak substitucí $w = z - \alpha$ přejdeme k polynomu $q(w) = p(z) = p(w + \alpha)$, který má stejný stupeň jako $p(z)$ a splňuje $q(0) = p(\alpha)$.) Napíšeme $p(z)$ od nejnižších mocnin a rozdělíme ho na tři sčítance:

$$p(z) = a_0 + p_1(z) + p_2(z) := a_0 + az^k + \sum_{j=k+1}^{n} a_j z^j \ ,$$

kde obě čísla $a_0, a = a_k \in \mathbb{C}$ jsou nenulová, $k \in \mathbb{N}$ a $a_j \in \mathbb{C}$ jsou zbylé koeficienty $p(z)$. Lehce se vidí, že

$$|z| = r \to 0^+ \Rightarrow \frac{|p_1(z)|}{|a_0|} \to 0 \ \text{ i } \ \frac{|p_2(z)|}{|p_1(z)|} \to 0\,.$$

Pro dostatečně malé $r > 0$ tedy každé číslo $z \in \mathbb{C}$ se $|z| = r$ splňuje, že $0 < |p_1(z)| < |a_0|$ a $|p_2(z)| < \frac{1}{2}|p_1(z)|$. Podle hořejšího připomenutí lze navíc mezi takovými z zvolit $z = \beta$ tak, že argumenty čísel $p_1(\beta) = a\beta^k$ a a_0 se liší přesně o π (tj. $p_1(\beta)$ jako vektor směřuje na opačnou stranu než vektor a_0). Pak ale $|a_0 + p_1(\beta)| = |a_0| - |p_1(\beta)|$. Takže

$$\begin{aligned} |p(\beta)| &= |a_0 + p_1(\beta) + p_2(\beta)| \\ &\leq |a_0 + p_1(\beta)| + |p_2(\beta)| \\ &= |a_0| - |p_1(\beta)| + |p_2(\beta)| \\ &< |a_0| - |p_1(\beta)|/2 < |a_0| \\ &= |p(0)| = |p(\alpha)| \end{aligned}$$

a $|p(\beta)| < |p(\alpha)|$.

Teď už jen stačí dokázat, že funkce

$$f:\ \mathbb{C} \to \mathbb{R},\ f(z) = |p(z)|$$

nabývá v nějakém bodě $\alpha \in \mathbb{C}$ na $\mathbb{C}$ minimum. Pak $|p(\alpha)| > 0$ nemůže nastat kvůli právě dokázanému faktu, tedy $|p(\alpha)| = 0$, $p(\alpha) = 0$ a jsme hotovi. Funkce $f(z)$ je na $\mathbb{C}$ spojitá, ale $\mathbb{C}$ není kompaktní množina a musíme jít oklikou. Každý uzavřený kruh $K_R = \{z \in \mathbb{C} \mid |z| \leq R\}$ kompaktní je. Napíšeme $p(z)$ od nejvyšších mocnin a rozdělíme ho na dva sčítance:

$$p(z) = z^n(a_n + q(z)) := z^n\left(a_n + \textstyle\sum_{j=0}^{n-1} a_j z^{j-n}\right)\,,$$

kde, jak víme, číslo $a_n \in \mathbb{C}$ je nenulové. Lehce se vidí, že

$$|z| = R \to +\infty \Rightarrow |q(z)| \to 0\,.$$

Existuje tedy $R > 0$, že pro každé $z \in \mathbb{C}$ se $|z| > R$ je

$$f(z) = |p(z)| = |z|^n|a_0 + q(z)| > R^n(|a_0| - |q(z)|) > R^n(|a_0|/2) > |p(0)|\,.$$

Na kompaktní množině K_R nabývá $f(z)$ v nějakém bodě $\alpha \in K_R$ minimum s hodnotou $f(\alpha) = |p(\alpha)| \leq |p(0)|$, neboť $0 \in K_R$. Protože $f(z) = |p(z)| > |p(0)| \geq |p(\alpha)| = f(\alpha)$ pro každý bod z mimo K_R, nabývá $f(z)$ v α minimum na celém $\mathbb{C}$. Tím je důkaz dokončen. $\square$