NI-MPI handouts

Obsah

Část I
Vícerozměrné volné a vázané extrémy

1 Vícerozměrný prostor a funkce

1.1 Norma

Norma a vzdálenost – připomenutí

Definice 1.1 — Norma *norm*. **Norma** na vektorovém prostoru V (nad $\mathbb{R}$ nebo $\mathbb{C}$) je zobrazení $\|\cdot\| : V \to \mathbb{R}_0^+$ splňující:

1. $\|\mathbf{x}\| = 0 \Rightarrow \mathbf{x} = \mathbf{0}$,
2. $\|\alpha\mathbf{x}\| = |\alpha| \cdot \|\mathbf{x}\|$,
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (trojúhelníková nerovnost),

pro všechna $\mathbf{x}, \mathbf{y} \in V$ a všechny skaláry α.

Máme-li normu $\|\cdot\|$ na V, definujeme vzdálenost vektorů $\mathbf{x}, \mathbf{y} \in V$ jako $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$. Zřejmě platí

- $d(\mathbf{x}, \mathbf{y}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{y}$;
- $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ (symetrie);
- $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$ (trojúhelníková nerovnost).

Příklady norem

Pro $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$(nebo $\mathbb{C}^n$)

- $\|\mathbf{x}\|_2 := \sqrt[2]{\sum_{i=1}^n |x_i|^2}$ (eukleidovská norma)
- $\|\mathbf{x}\|_1 := \sum_{i=1}^n |x_i|$ (součtová norma)
- Obecně, pro libovolné $p \geq 1$

$$\|\mathbf{x}\|_p := \sqrt[p]{\sum_{i=1}^n |x_i|^p}$$

- $\|\mathbf{x}\|_\infty := \max\left\{|x_i| \,\middle|\, i \in \{1, \ldots, n\}\right\}$ (maximová norma)

(Ověření trojúhelníkové nerovnosti není vždy triviální!)

Okolí bodu

Zvolme nějakou normu $\|\cdot\|$ na $\mathbb{R}^n$.

Buď $\mathbf{x} \in \mathbb{R}^n$ a $\delta \in \mathbb{R}^+$, **δ-okolí bodu $\mathbf{x}$** je množina

$$H_\delta(\mathbf{x}) = \{\mathbf{b} \in \mathbb{R}^n \colon \|\mathbf{x} - \mathbf{b}\| < \delta\}.$$

(Taková množina se také nazývá otevřená koule o středu $\mathbf{x}$ a poloměru δ.)
Obecně pro okolí: pokud není třeba specifikovat parametr δ budeme psát jednoduše $H(\mathbf{x})$.

Hromadný bod

Mějme množinu $M \subset \mathbb{R}^n$.

Řekneme, že $\mathbf{x} \in \mathbb{R}^n$ je **hromadným bodem** M (*limit/cluster/accumulation point of* M), pokud pro všechna $r > 0$ platí $(H_r(\mathbf{x}) \setminus \{\mathbf{x}\}) \cap M \neq \emptyset$.

Bod $\mathbf{x} \in M$, který není hromadný, se nazývá **izolovaný**.

Limita posloupnosti

Mějme posloupnost bodů z $\mathbb{R}^n$, tedy

$$(\mathbf{x}_i)_{i=0}^{+\infty},$$

kde $\mathbf{x}_i \in \mathbb{R}^n$.

Řekneme, že posloupnost $(\mathbf{x}_i)_{i=0}^{+\infty}$ **má limitu** $L \in \mathbb{R}^n$, pokud

$$\forall \epsilon > 0 \quad \exists N \quad \forall n > N \quad \mathbf{x}_n \in H_\epsilon(L).$$

Značení:

$$\lim_{n\to\infty} \mathbf{x}_n = L$$

1.2 Funkce více proměnných

Funkce více proměnných

Reálnou funkcí více reálných proměnných rozumíme zobrazení

$$f : D_f \to \mathbb{R},$$

kde $D_f \subseteq \mathbb{R}^n$ (pro n kladné celé).

Tedy funkce, která má n reálných parametrů a vrací taktéž reálnou hodnotu.

D_f je **definiční obor** (*domain*).
$f(D_f)$ je **obor hodnot** (*range*).

Graf funkce

Graf funkce f je množina

$$\Gamma_f = \{(b_1, b_2, \ldots, b_n, f(b_1, b_2, \ldots, b_n)) \colon (b_1, b_2, \ldots, b_n) \in D_f\} \subset \mathbb{R}^{n+1}.$$

1.3 Limita funkce více proměnných

Limita funkce více proměnných

Řekneme, že funkce $f : D_f \to \mathbb{R}$, $D_f \subset \mathbb{R}^n$, **má limitu** $L \in \mathbb{R}$ **v hromadném bodě** $\mathbf{b}$ **množiny** D_f pokud

$$\forall H(L) \quad \exists H(\mathbf{b}) \quad \mathbf{x} \in (D_f \cap H(\mathbf{b})) \setminus \{\mathbf{b}\} \implies f(\mathbf{x}) \in H(L)$$

Značení:

$$\lim_{\mathbf{x}\to\mathbf{b}} f(\mathbf{x}) = L.$$

Limity pomocí posloupnosti

Věta 1.2 Mějme funkci $f : D_f \to \mathbb{R}$, $D_f \subset \mathbb{R}^n$. Funkce f má v hromadném bodě $\mathbf{b}$ množiny D_f limitu L, tedy !

$\lim_{\mathbf{x}\to\mathbf{b}} f(\mathbf{x}) = L$, právě tehdy, když pro všechny posloupnosti $(\mathbf{x}_i)_{i=0}^{+\infty} \subset D_f$, $\mathbf{x}_i \neq \mathbf{b}$, platí

$$\lim_{n\to+\infty} \mathbf{x}_n = \mathbf{b} \quad \Longrightarrow \quad \lim_{n\to+\infty} f(\mathbf{x}_n) = L$$

1.4 Spojitost

Spojitá funkce

Řekneme, že funkce $f : D_f \to \mathbb{R}$, $D_f \subset \mathbb{R}^n$, je **spojitá v bodě** $\mathbf{x}_0 \in D_f$ pokud

$$\forall \epsilon > 0 \quad \exists \delta > 0 : \quad x \in (D_f \cap H_\delta(\mathbf{x}_0)) \Longrightarrow f(x) \in H_\epsilon(f(\mathbf{x}_0)).$$

Funkce f je **spojitá** pokud je spojitá ve všech bodech z D_f.

Lze formulovat i pomocí limity: funkce f je spojitá, pokud pro všechny neizolované body $\mathbf{x}_0 \in D_f$ platí

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0).$$

1.5 Lokální extrémy

Lokální extrémy

Definice 1.3 Řekneme, že reálná funkce f má v bodě $\mathbf{b} \in D_f$

1. **lokální minimum**, pokud $\exists \delta > 0$, $\forall \mathbf{x} \in (D_f \cap H_\delta(\mathbf{b})), f(\mathbf{x}) \geq f(\mathbf{b})$;
2. **ostré lokální minimum**, pokud $\exists \delta > 0$, $\forall \mathbf{x} \in (D_f \cap H_\delta(\mathbf{b})) \setminus \{\mathbf{b}\}, f(\mathbf{x}) > f(\mathbf{b})$;
3. **globální minimum**, pokud $\forall \mathbf{x} \in D_f, f(\mathbf{x}) \geq f(\mathbf{b})$.

Bod $\mathbf{b}$ se nazývá postupně bodem lokálního minima, bodem ostrého lokálního minima a bodem globálního minima funkce f.

Pro globální minimum (na množině $D \subset D_f$) se též používá **argument minima** s tímto (zneužitým) značením:

$$\mathbf{b} = \operatorname*{argmin}_{\mathbf{x}\in D} f(\mathbf{x}).$$

(Obecně by argmin měl být vzor minima, tj. množina bodů.)
(Ostré) lokální maximum a globální maximum se definuje analogicky.

Extrémy na omezených a uzavřených množinách

Definice 1.4 Množina $D \subset \mathbb{R}^n$ se nazývá

- **omezená**, pokud je podmnožinou nějaké otevřené koule;
- **otevřená**, pokud s každým svým bodem obsahuje i nějaké jeho okolí;
- **uzavřená**, pokud $\mathbb{R}^n \setminus D$ je otevřená (obsahuje i svou hranici, tedy body, v jejichž každém okolí leží bod z D a bod mimo D).

(Uzavřenou množinu lze také charakterizovat tím, že obsahuje všechny své hromadné body.)

Věta 1.5 Je-li $D_f \subset \mathbb{R}^n$ **omezená** a **uzavřená**, pak má spojitá funkce $f : D_f \to \mathbb{R}$ v D_f globální minimum a globální maximum.

2 Parciální derivace a gradient

2.1 Parciální derivace

Definice parciální derivace

Označme jednotlivé proměnné jako $x_1, x_2, x_3, \ldots, x_n$.

Definice 2.1 Parciální derivace funkce f ve směru x_i (nebo podle x_i) v bodě $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in D_f$ takovém, že $\exists H(\mathbf{b}) \subset D_f$, je

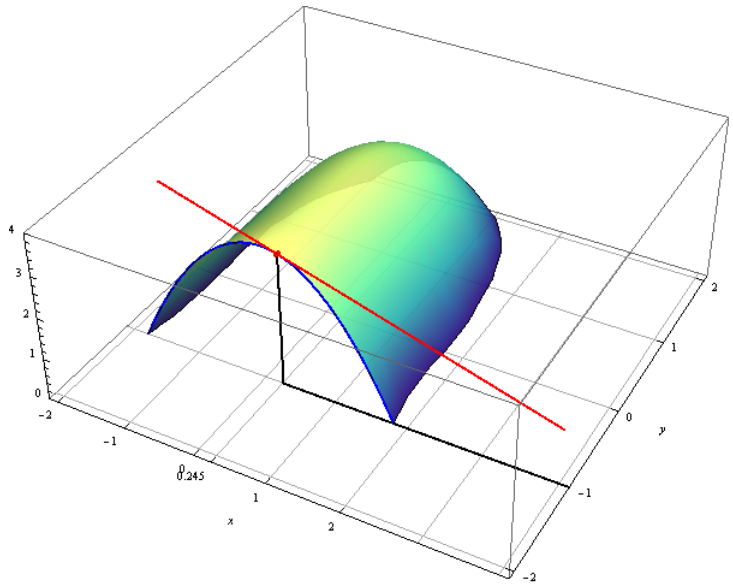
$$\lim_{h \to 0} \frac{f(b_1, b_2, \ldots, b_i + h, \ldots, b_n) - f(b_1, b_2, \ldots, b_i, \ldots, b_n)}{h} = L \in \mathbb{R},$$

pokud tato limita existuje.

Značení: $\frac{\partial f}{\partial x_i}(\mathbf{b}) = L$. (Jiná častá značení: $\partial_{x_i} f(\mathbf{b})$ a $f_{x_i}(\mathbf{b})$.)

Jedná se o směrnici tečny ke grafu funkce f ve směru osy x_i.

Parciální derivace - ilustrace



2.2 Gradient

Gradient

Definice 2.2 Gradient funkce f v bodě $\mathbf{b} \in D_f$ je (řádkový) vektor

$$\nabla f(\mathbf{b}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{b}), \frac{\partial f}{\partial x_2}(\mathbf{b}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{b}) \right).$$

Poznámka: jelikož se gradient opírá o pojem parc. derivace, bod $\mathbf{b}$ je nutně vnitřním bodem D_f (v D_f leží i nějaké jeho okolí).

Poznámka 2: uvedená definice je zjednodušená: lze ji použít pokud jsou všechny parciální derivace funkce f spojité na nějakém okolí bodu $\mathbf{b}$.

Nezjednodušená definice vypadá takto: Mějme vnitřní bod $\mathbf{b}$ def. oboru D_f. Gradient funkce f v bodě $\mathbf{b} \in D_f$ je řádkový vektor $\mathbf{v}$, pro který platí

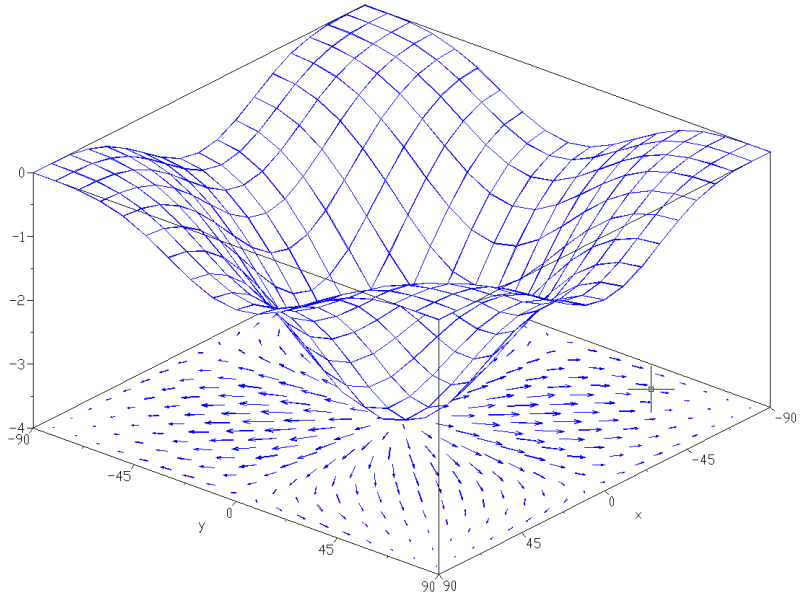
$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad : \quad \mathbf{x} \in (H_\delta(\mathbf{b}) \setminus \mathbf{b}) \Longrightarrow \|f(\mathbf{x}) - f(\mathbf{b}) - \mathbf{v} \cdot (\mathbf{x} - \mathbf{b})\| < \varepsilon \|\mathbf{x} - \mathbf{b}\|.$$

(Což lze i zapsat $\lim_{\mathbf{x} \to \mathbf{b}} \frac{\|f(\mathbf{x}) - f(\mathbf{b}) - \mathbf{v} \cdot (\mathbf{x} - \mathbf{b})\|}{\|\mathbf{x} - \mathbf{b}\|} = 0$.) Má-li f spojité parciální derivace na nějakém okolí $\mathbf{b}$, pak pro souřadnice vektoru $\mathbf{v} = \nabla f(\mathbf{b})$ platí rovnost v definici výše. Obráceně to neplatí.

Toto zjednodušení uvažujeme, protože s funkcemi, pro které to neplatí, vesměs pracovat nebudeme.

Geometrický význam: gradient ukazuje směr (v D_f) nejvyššího růstu funkce f.

Gradient - ilustrace



2.3 Derivace ve směru

Derivace ve směru

Parciální derivace je vždy ve směru některé ze souřadnicových os x_i. Co směrnice v jiných směrech (v D_f)?

Definice 2.3 Nechť $\mathbf{v} \in \mathbb{R}^{n,1} = \mathbb{R}^n$, $\|\mathbf{v}\| = 1$. **Derivace funkce f ve směru $\mathbf{v}$ v bodě $\mathbf{b} \in D_f$** takovém, že $\exists H(\mathbf{b}) \subset D_f$, je

$$\nabla_{\mathbf{v}} f(\mathbf{b}) = \lim_{h \to 0} \frac{f(\mathbf{b} + h\mathbf{v}) - f(\mathbf{b})}{h}.$$

Věta 2.4 Jsou-li všechny parciální derivace funkce f na nějakém okolí bodu $\mathbf{b}$ spojité, pak platí

$$\nabla_{\mathbf{v}} f(\mathbf{b}) = \nabla f(\mathbf{b}) \cdot \mathbf{v}.$$

Poznámka: někdy se nevyžaduje, aby směr $\mathbf{v}$ byl jednotkový, tj. $\|\mathbf{v}\| = 1$. (Potom se o $\mathbf{v}$ nemluví jako o směru, ale o obecném vektoru.)

2.4 Tečná rovina

Tečná nadrovina

Obdobně jako tečna pro $n = 1$, tak tečnou nadrovinou rozumíme objekt sídlící ve stejném prostoru jako graf funkce (tedy v $\mathbb{R}^{n+1}$), který je lineární varietou kodimenze 1 a který, v nějakém smyslu, nejlépe postihuje chování přírustku funkce f v daném bodě $\mathbf{b} \in D_f$.

Toto chování umíme popsat pro jednotlivé směry $\mathbf{v} \in \mathbb{R}^n$: pro každý směr umíme v tomto směru najít tečnu funkce f zúžené na tento směr, tj. zúžené na $D_f \cap \{\mathbf{x} \colon \mathbf{x} = \mathbf{b} + t\mathbf{v}, t \in \mathbb{R}\}$. Toto zúžení lze chápat jako funkci jedné proměnné, tedy $\mathbb{R} \to \mathbb{R}$, a u té umíme najít tečnu.

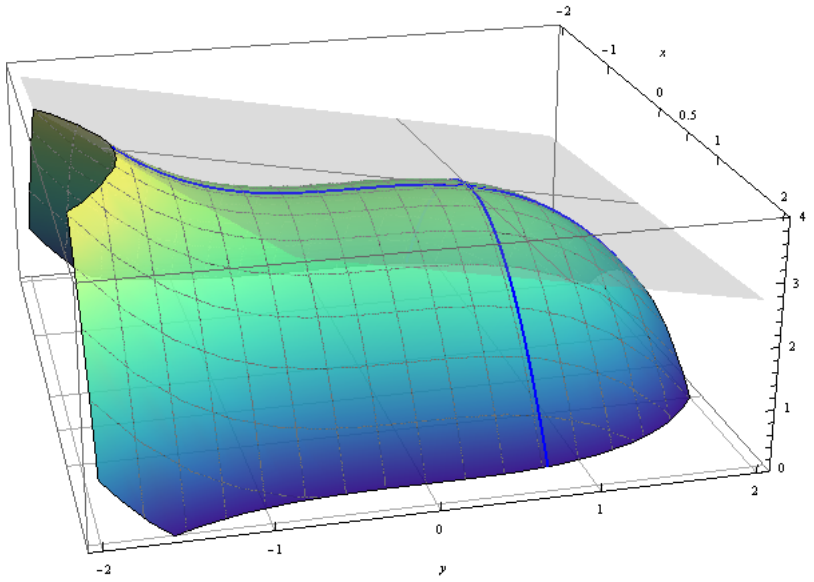
Sjednotíme-li tečny ve všech směrech (v bodě $\mathbf{b} \in D_f$), dostaneme **tečnou nadrovinu funkce f v bodě $\mathbf{b}$**. (Podmínkou pro existenci je existence gradientu f v bodě $\mathbf{b}$.)

Rovnice této nadroviny je

$$z = \frac{\partial f}{\partial x_1}(\mathbf{b})\,(x_1 - b_1) + \frac{\partial f}{\partial x_2}(\mathbf{b})\,(x_2 - b_2) + \cdots + \frac{\partial f}{\partial x_n}(\mathbf{b})\,(x_n - b_n) + f(\mathbf{b}).$$

Její normálový vektor je $(\frac{\partial f}{\partial x_1}(\mathbf{b}), \frac{\partial f}{\partial x_2}(\mathbf{b}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{b}), -1)$.

Tečná nadrovina - ukázka



3 Lokální extrémy a nutná podmínka existence

3.1 Nutná podmínka existence lok. extrému

Věta 3.1 — Nutná podmínka lokálního extrému. Nechť má funkce $f : D_f \to \mathbb{R}$, $D_f \subset \mathbb{R}^n$, v bodě $\mathbf{b}$ parciální derivaci podle i-té proměnné.

Pokud f má v bodě $\mathbf{b}$ lokální extrém, potom

$$\frac{\partial f}{\partial x_i}(\mathbf{b}) = 0.$$

!

Důkaz Věty 3.1. Zavedeme funkci jedné proměnné:

$$g(x_i) = f(b_1, b_2, \ldots, b_{i-1}, x_i, b_{i+1}, \ldots, b_n).$$

Funkce g je v bodě b_i diferencovatelná a pro její derivaci platí

$$g'(b_i) = \frac{\partial f}{\partial x_i}(\mathbf{b}).$$

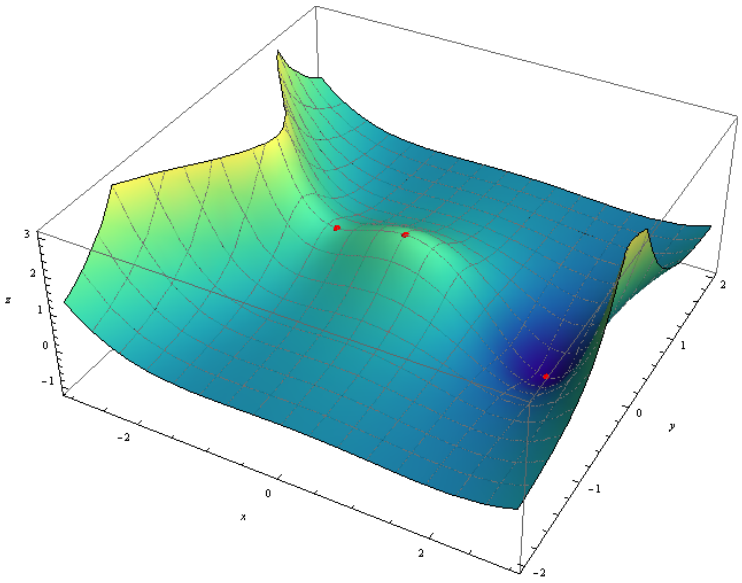
Protože f má v $\mathbf{b}$ lokální extrém, má g v b_i lokální extrém, a tedy $g'(b_i) = 0$. ■

Důsledek: pokud existuje gradient funkce f v bodě $\mathbf{b}$, pak existence lokálního extrému implikuje

$$\nabla f(\mathbf{b}) = \mathbf{0}.$$

Body $\mathbf{b} \in D_f$ splňující $\nabla f(\mathbf{b}) = 0$ se nazývají **stacionární**.

Stacionární body



Kritické body

Mějme funkci $f : D_f \to \mathbb{R}$, kde $D_f \subseteq \mathbb{R}^n$.

V úloze hledání (lokálních) extrémů funkce f jsou body podezřelé z extrému buď

- body stacionární

anebo

- body, ve kterých gradient f neexistuje.

Takové body se souhrně nazývají **kritické body**.

4 Lokální extrémy a postačující podmínka existence

4.1 Parciální derivace druhého řádu

Parciální derivace druhého řádu

První parciální derivace (nějaké funkce $f : D_f \to \mathbb{R}, D_f \subset \mathbb{R}^n$) je zobrazením z množiny všech bodů, kde parciální derivace existuje.

Pro jednoduchost předpokládejme, že tomu tak je pro všechny body z D_f (obecně je to pro nějakou podmnožinu D_f) a všechny proměnné x_i.

Taková zobrazení budou opět zobrazení typu $D_f \to \mathbb{R}$, tedy $\frac{\partial f}{\partial x_i} : D_f \to \mathbb{R}$

Můžeme je tedy opět zkusit (parciální derivace nemusí existovat) parciálně derivovat. Dostaneme parciální derivaci druhého řádu (parciálně derivujeme podle x_j parciální derivaci podle x_i) v bodě $\mathbf{b}$:

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

- Pokud $i \neq j$, hovoříme také o **smíšené** (druhé) parciální derivaci.
- Pokud $i = j$, zapisuje se též zkráceně: $\frac{\partial^2 f}{\partial x_i^2}(\mathbf{b})$.
- Opět lze chápat jako zobrazení z nějaké podmnožiny D_f.

Jiná značení (jako zobrazení): $\partial_{xy} f$, f_{xy}.

Hessova matice

Definice 4.1 Mějme funkci $f : D_f \to \mathbb{R}, D_f \subset \mathbb{R}^n$.
Existují-li všechny druhé parciální derivace funkce f v bodě $\mathbf{b}$, pak se zaznamenávají do matice takto:

$$\nabla^2 f(\mathbf{b}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{b}) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{b}) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{b}) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{b}) \end{pmatrix}.$$

Matici $\nabla^2 f(\mathbf{b})$ nazýváme **Hessovou maticí funkce f v bodě $\mathbf{b}$** (též Hessián).

Lze chápat jako zobrazení z podmnožiny D_f do $\mathbb{R}^{n,n}$.

Zaměnitelnost druhých parciálních derivací

Věta 4.2 Nechť $f : D_f \to \mathbb{R}, D_f \subset \mathbb{R}^2$ a $\mathbf{b} \in D_f$.
Pokud existuje $\frac{\partial^2 f}{\partial x \partial y}(\mathbf{b})$ a funkce $\frac{\partial^2 f}{\partial x \partial y}$ je v $\mathbf{b}$ spojitá, potom $\frac{\partial^2 f}{\partial y \partial x}(\mathbf{b})$ existuje a platí

$$\frac{\partial^2 f}{\partial x \partial y}(\mathbf{b}) = \frac{\partial^2 f}{\partial y \partial x}(\mathbf{b}).$$

Tedy pokud nějaká smíšená druhé parciální derivace existuje a je spojitá, pak nezávisí na pořadí parciálního derivování (vše v bodě $\mathbf{b}$).

Důsledky:

- Hessova matice je „často“ symetrická.
- Je-li nějaká n-tá parciální derivace dané funkce v daném bodě spojitá, pak záleží pouze na tom, kolikrát podle které proměnné derivujeme, nikoli na pořadí derivování.

Druhá derivace ve směru

Definice 4.3 Nechť $\mathbf{v} \in \mathbb{R}^{n,1} = \mathbb{R}^n$, $\|\mathbf{v}\| = 1$. **Druhá parciální derivace funkce f ve směru $\mathbf{v}$ v bodě $\mathbf{b} \in D_f$** takovém, že $\exists H(\mathbf{b}) \subset D_f$, je

$$\nabla_{\mathbf{v}} \left(\nabla_{\mathbf{v}} f\right)(\mathbf{b}).$$

Věta 4.4 Nechť $\mathbf{v} \in \mathbb{R}^{n,1}$, $\|\mathbf{v}\| = 1$. Mějme funkci $f : D_f \to \mathbb{R}, D_f \subset \mathbb{R}^n$ a bod $\mathbf{b} \in D_f$. Nechť existuje okolí $H(\mathbf{b}) \subset D_f$ takové, že f má na $H(\mathbf{b})$ spojité všechny druhé parciální derivace, potom

$$\nabla_{\mathbf{v}} \left(\nabla_{\mathbf{v}} f\right)(\mathbf{b}) = \mathbf{v}^T \cdot \nabla^2 f(\mathbf{b}) \cdot \mathbf{v}.$$

4.2 Definitnost matic

Definitnost matic

Definice 4.5 — Definitnost matic. Mějme $\mathbf{A} \in \mathbb{R}^{n,n}$. Řekneme, že matice $\mathbf{A}$ je

1. **pozitivně semidefinitní**, pokud $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$ pro $\forall \mathbf{x} \in \mathbb{R}^{n,1}$;
2. **pozitivně definitní**, pokud $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ pro $\forall \mathbf{x} \in \mathbb{R}^{n,1}, \mathbf{x} \neq 0$;
3. **negativně semidefinitní**, pokud $\mathbf{x}^T\mathbf{A}\mathbf{x} \leq 0$ pro $\forall \mathbf{x} \in \mathbb{R}^{n,1}$;
4. **negativně definitní**, pokud $\mathbf{x}^T\mathbf{A}\mathbf{x} < 0$ pro $\forall \mathbf{x} \in \mathbb{R}^{n,1}, \mathbf{x} \neq 0$;
5. **indefinitní**, pokud není pozitivně ani negativně semidefinitní.

Matice $\mathbf{A}$ je indefinitní právě tehdy, když $\exists \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ a $\mathbf{y}^T\mathbf{A}\mathbf{y} < 0$.

Charakterizace (semi)definitnosti

Věta 4.6 Buď $\mathbf{A} \in \mathbb{R}^{n,n}$ **symetrická** matice. Potom platí následující:

- Matice $\mathbf{A}$ je *pozitivně semidefinitní* právě tehdy, když všechna její vlastní čísla jsou *nezáporná*.
- Matice $\mathbf{A}$ je *pozitivně definitní* právě tehdy, když všechna její vlastní čísla jsou *kladná*.
- Matice $\mathbf{A}$ je *negativně semidefinitní* právě tehdy, když všechna její vlastní čísla jsou *nekladná*.
- Matice $\mathbf{A}$ je *negativně definitní* právě tehdy, když všechna její vlastní čísla jsou *záporná*.
- Matice $\mathbf{A}$ je *indefinitní* právě tehdy, když má alespoň jedno *kladné* a alespoň jedno *záporné* vlastní číslo.

Náznak důkazu: Jelikož je matice $\mathbf{A}$ symetrická, platí $\mathbf{A} = P^{-1}DP$, kde D je reálná diagonální a sloupce reálné regulární matice P sestávají z jednotkových vlastních vektorů. Jelikož platí $P^{-1} = P^T$, tak zbytek plyne přímo z definice (semi)definitnosti a indefinitnosti.

Sylvestrovo kritérium

Věta 4.7 — Sylvestrovo kritérium. Buď $\mathbf{A} \in \mathbb{R}^{n,n}$ **symetrická** matice. Pro matici $\mathbf{A} \in \mathbb{R}^{n,n}$ definujeme matice $A_1, A_2, \ldots, A_n$ takto: $A_k \in \mathbb{R}^{k,k}$ je čtvercová matice v levém horním rohu matice $\mathbf{A}$. Platí:

- Matice $\mathbf{A}$ je pozitivně definitní právě tehdy, když je determinant všech matic $A_1, A_2, \ldots, A_n$ kladný.
- Matice $\mathbf{A}$ je negativně definitní právě tehdy, když je determinant matic A_k záporný pro k liché a kladný pro k sudé.

Jak poznat indefinitnost

Věta 4.8 Pokud má matice $\mathbf{A} \in \mathbb{R}^{n,n}$ na diagonále dva prvky s různým znaménkem (jeden kladný a druhý záporný), pak je indefinitní.

Důkaz. Mějme $(\mathbf{A})_{i,i} > 0$ a $(\mathbf{A})_{j,j} < 0$. Pak jistě platí

$$\mathbf{e}_i^T \mathbf{A} \mathbf{e}_i > 0 \quad \text{a} \quad \mathbf{e}_j^T \mathbf{A} \mathbf{e}_j < 0$$

($\mathbf{e}_i$ je i-tý vektor standardní báze $\mathbb{R}^n$) a tedy matice $\mathbf{A}$ je indefinitní (přímo z definice). ■

4.3 Postačující podmínka existence lokálního extrému

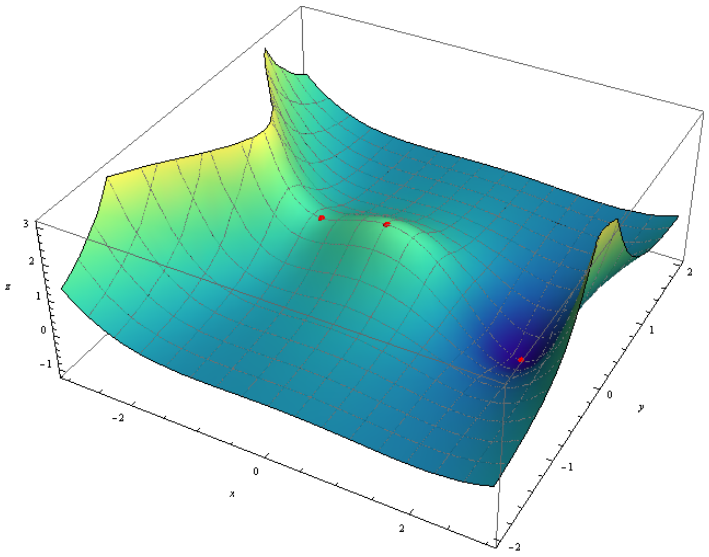
Postačující podmínka existence extrému a sedlového bodu

Stacionární bod, který není minimem ani maximem a na jehož nějakém okolí má funkce f spojité všechny parciální derivace, se nazývá **sedlovým** bodem.

Věta 4.9 — Postačující podmínka existence extrému a sedlového bodu. Nechť $\mathbf{b} \in D_f$ je stacionární bod funkce $f : D_f \to \mathbb{R}, D_f \subset \mathbb{R}^n$. Nechť existuje okolí $H(\mathbf{b}) \subset D_f$ takové, že f má na $H(\mathbf{b})$ spojité všechny druhé parciální derivace, potom

- je-li $\nabla^2 f(\mathbf{b})$ pozitivně definitní, pak $\mathbf{b}$ je ostré lokální minimum;
- je-li $\nabla^2 f(\mathbf{b})$ negativně definitní, pak $\mathbf{b}$ je ostré lokální maximum;
- je-li $\nabla^2 f(\mathbf{b})$ indefinitní, pak $\mathbf{b}$ je sedlový bod.

Minimum, maximum, sedlový bod



Co semidefinitnost?

Věta 4.10 — Nutná podmínka existence lokálního extrému. Nechť $\mathbf{b} \in D_f$ je stacionární bod funkce $f : D_f \to \mathbb{R}, D_f \subset \mathbb{R}^n$. Nechť existuje okolí $H(\mathbf{b}) \subset D_f$ takové, že f má na $H(\mathbf{b})$ spojité všechny druhé parciální derivace, potom

- je-li $\mathbf{b}$ lokální minimum, pak $\nabla^2 f(\mathbf{b})$ je pozitivně semidefinitní;
- je-li $\mathbf{b}$ lokální maximum, pak $\nabla^2 f(\mathbf{b})$ je negativně semidefinitní.

Toto tvrzení **nelze** obrátit.

5 Shrnutí: postup analytického hledání extrémů

Postup analytického hledání extrémů

1. Najít kritické body, tj. stacionární body a body, kde alespoň jedna parciální derivace neexistuje.
2. Pokud jsou všechny 2. parciální derivace v okolí stacionárního bodu $\mathbf{b}$ spojité, nalézt Hessovu matici. Pokud je tato matice
   a) pozitivně definitní, pak je bod $\mathbf{b}$ bodem ostrého lokálního minima;
   b) negativně definitní, pak je bod $\mathbf{b}$ bodem ostrého lokálního maxima;
   c) indefinitní, pak je bod $\mathbf{b}$ sedlovým bodem (tj. není extrémem).

V ostatních případech je třeba rozhodnout jiným způsobem (nelze pomocí Hessovy matice rozhodnout, např. protože neexistuje nebo je semidefinitní).

6 Konvexní funkce

6.1

Definice konvexní funkce

Definice 6.1 Funkce $f : D_f \to \mathbb{R}, D_f \subset \mathbb{R}^n$ je **konvexní** pokud je D_f konvexní množina a

$$\forall \mathbf{b}_1, \mathbf{b}_2 \in D_f \ \forall t \in [0,1] : \ f(t\mathbf{b}_1 + (1-t)\mathbf{b}_2) \leq tf(\mathbf{b}_1) + (1-t)f(\mathbf{b}_2).$$

Funkce f je **konkávní**, pokud $-f$ je konvexní.

Extrémy konvexní funkce

Funkce f, která má spojité všechny druhé parciální derivace, je konvexní právě tehdy, když je její Hessova matice pozitivně semidefinitní ve všech bodech vnitřku D_f (vnitřek množiny jsou všechny její body, které v ní leží s nějakým okolím, jinak řečeno je to množina bez své hranice).

Lokální minimum konvexní funkce je globálním minimem.

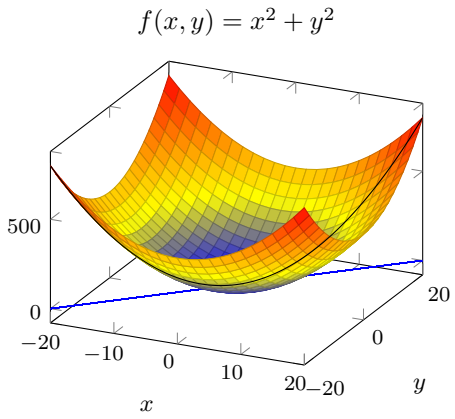
Řešené příklady: Extrémy a definitnost

Základní cvičení 17.1
Uvažme funkce

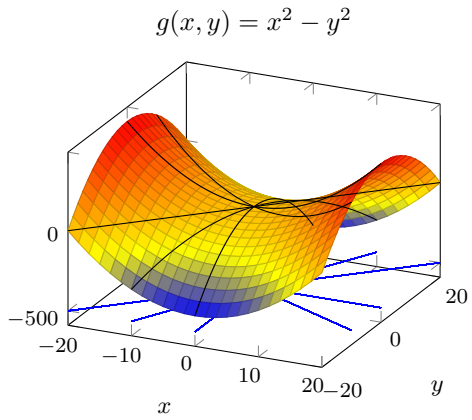
$$f(x,y) = x^2 + y^2$$
$$g(x,y) = x^2 - y^2$$
$$h(x,y) = x^2 + y^3$$
$$u(x,y) = xy$$
$$w(x,y) = (x+y)^2$$
$$z(x,y) = x^4 + y^4$$

Pro všechny funkce nalezněte všechny kritické body a zjistěte, zda se jedná o lokální minimum, lokální maximum nebo sedlový bod. Prof funkce f a g spočtěte první a druhou derivaci ve směru přímky $y = x$ v bodě $(1, 1)$.

$f(x,y) = x^2 + y^2$

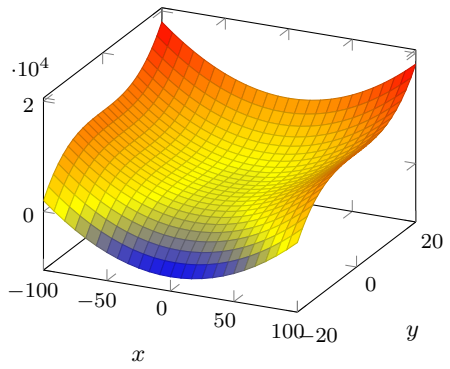


$g(x,y) = x^2 - y^2$



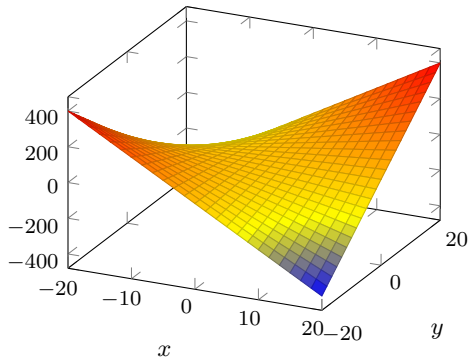
$h(x,y) = x^2 + y^3$

$h(x, y) = x^2 + y^3$



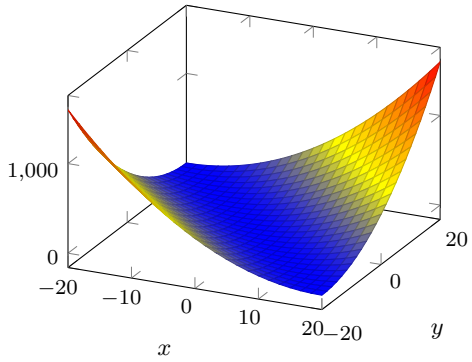
$u(x, y) = xy$

$u(x, y) = xy$



$w(x, y) = (x + y)^2$

$w(x, y) = (x + y)^2$



$f(x, y, z) = x^3 + y^2 + z^2 + 12xy + 2z$

Základní cvičení 17.2
Mějme

$$f(x, y, z) = x^3 + y^2 + z^2 + 12xy + 2z.$$

Nalezněte všechny kritické body a zjistěte, zda se jedná o lokální minimum, lokální maximum nebo sedlový bod.

$D_f =$
$\nabla f(x, y, z) =$

$\nabla f(x, y, z) = \left(3x^2 + 12y, 2y + 12x, 2z + 2\right)$

$\nabla f(x, y, z) = 0 \Leftrightarrow$

$\nabla f(x,y,z) = \left(3x^2 + 12y, 2y + 12x, 2z + 2\right)$
k. body $= \{(0,0,-1), (24,-144,-1)\}$

$\nabla^2 f(x,y,z) =$

k. body $= \{(0,0,-1), (24,-144,-1)\}$
$\nabla^2 f(x,y,z) = \begin{pmatrix} 6x & 12 & 0 \\ 12 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

$\nabla^2 f(0,0,-1) =$

k. body $= \{(0,0,-1), (24,-144,-1)\}$
$\nabla^2 f(x,y,z) = \begin{pmatrix} 6x & 12 & 0 \\ 12 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

$\nabla^2 f(24,-144,-1) =$

7 Vázané extrémy

7.1 Implicitní funkce

Věta 7.1 — Věta o implicitní funkci. Nechť $f\colon \mathbb{R}^2 \to \mathbb{R}$ a $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ jsou spojité v okolí bodu (a,b). Předpokládejme dále, že

$$f(a,b) = 0 \qquad \text{a} \qquad \frac{\partial f}{\partial y}(a,b) \neq 0.$$

(a) Pak existují $\varepsilon > 0$, $\delta > 0$ a obdélníkové okolí $R := \{(x,y) : |x-a| < \varepsilon, |y-b| < \delta\}$ bodu (a,b) takové, že pro každé $x \in (a-\varepsilon, a+\varepsilon)$ existuje právě jedno $y \in (b-\delta, b+\delta)$, které splňuje rovnici $f(x,y) = 0$. Tedy y je funkcí proměnné x a lze psát $y = \varphi(x)$. Definiční obor φ je interval $(a-\varepsilon, a+\varepsilon)$.

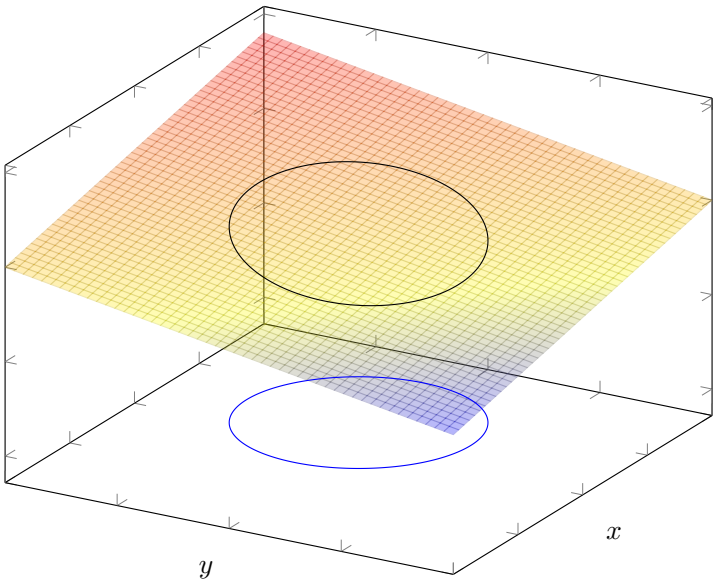
(b) Funkce φ určená v bodě (a) a její derivace φ' jsou spojité na intervalu $(a-\varepsilon, a+\varepsilon)$ a platí

$$\varphi'(x) = -\frac{\frac{\partial f}{\partial x}(x,\varphi(x))}{\frac{\partial f}{\partial y}(x,\varphi(x))} \text{ pro } x \in (a-\varepsilon, a+\varepsilon).$$

7.2 Definice problému

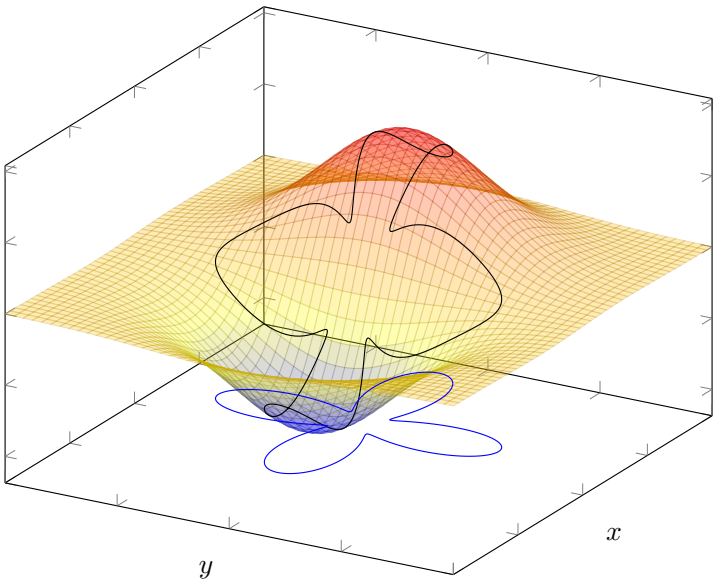
Ukázka - Obrázek 1

Nalezněte maximum a minimum, když se „pohybujeme“ na grafu pouze po černé křivce (= nad modrou křivkou):



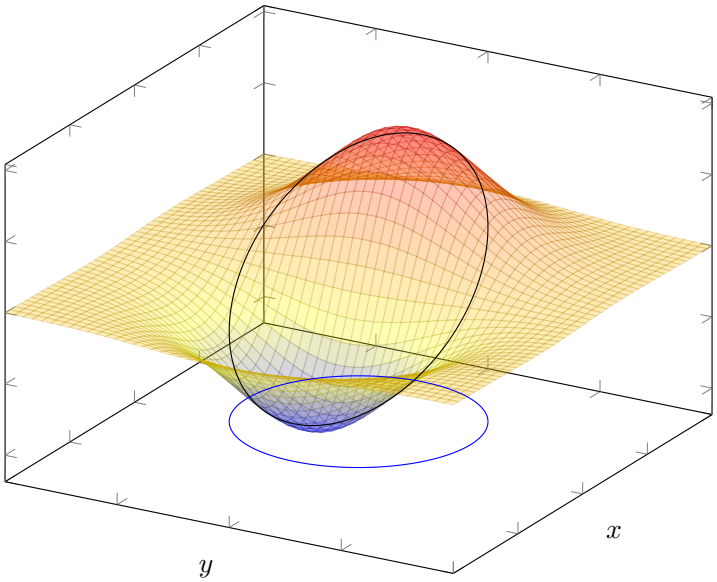
Ukázka - Obrázek 2

Nalezněte maximum a minimum, když se „pohybujeme“ na grafu pouze po černé křivce (= nad modrou křivkou):



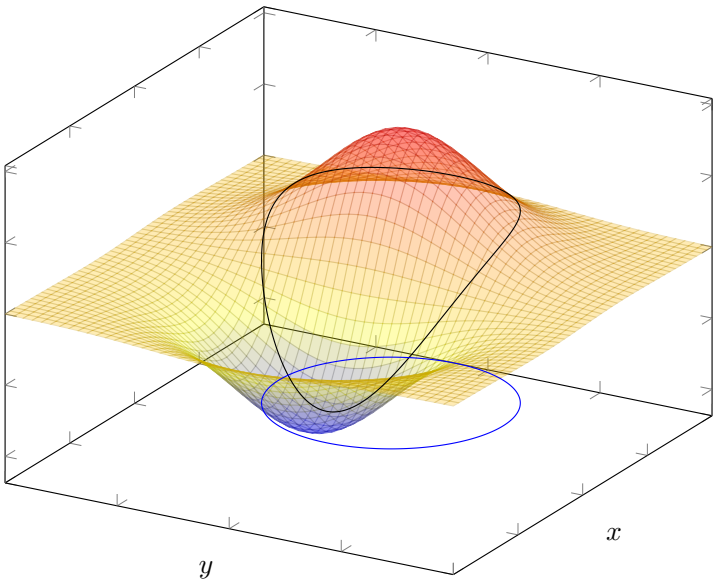
Ukázka - Obrázek 3

Nalezněte maximum a minimum, když se „pohybujeme“ na grafu pouze po černé křivce (= nad modrou křivkou):



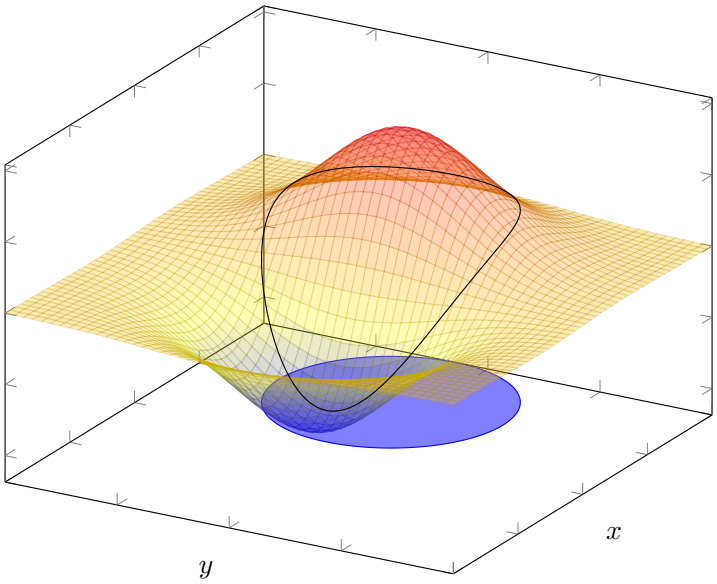
Ukázka - Obrázek 4

Nalezněte maximum a minimum, když se „pohybujeme“ na grafu pouze po černé křivce (= nad modrou křivkou):



Ukázka - Obrázek 5

Nalezněte maximum a minimum, když se „pohybujeme“ na grafu pouze **nad modrou množinou**:



Formulace problému

Označme $\hat{m} = \{1, \ldots, m\}$ a $\hat{p} = \{1, \ldots, p\}$ $(m, p \in \mathbb{N})$.

Úloha vázaného extrému (minima) (*constrained optimization/minimization*) je obecně následující úloha:

$$\begin{cases} \text{minimalizuj} & f(\mathbf{x}), \\ \text{za podmínek} & g_j(\mathbf{x}) = 0,\ j \in \hat{m}, \\ & h_k(\mathbf{x}) \leq 0,\ k \in \hat{p}, \end{cases} \tag{1}$$

kde f, g_j, h_k jsou funkce $D \to \mathbb{R}$, kde $D \subset \mathbb{R}^n$.

Terminologie

f: objektivní/účelová/minimalizovaná/optimalizovaná funkce

g_j: rovnostní podmínka/vazba | podmínky/vazby

h_k: nerovnostní podmínka/vazba | podmínky/vazby

Jsou-li všechny funkce lineární (a je-li alespoň jedna nerovnostní podmínka): **úloha lineárního programování**

Jsou-li všechny vazby lineární (a je-li alespoň jedna nerovnostní podmínka) a f kvadratická: **úloha kvadratického programování**

Jinak obecně: **úloha nelineárního programování**

Přípustná řešení a jiná formulace

Označme **množinu přípustných řešení:**

$$\mathcal{M} = \{\mathbf{x} \in D \colon (\forall j \in \hat{m})(g_j(\mathbf{x}) = 0) \ \wedge \ (\forall k \in \hat{p})(h_k(\mathbf{x}) \leq 0)\}$$

Definice 7.2 Mějme množiny $D \subset \mathbb{R}^n$ a $\mathcal{M} \subset D$. Řekneme, že funkce $f : D \to \mathbb{R}$ má v bodě $\mathbf{x}^* \in \mathcal{M}$ lokální minimum **vzhledem k množině** $\mathcal{M}$, pokud existuje okolí $H(\mathbf{x}^*)$ takové, že platí

$$\forall \mathbf{x} \in (H(\mathbf{x}^*) \cap \mathcal{M}) \quad f(\mathbf{x}^*) \leq f(\mathbf{x}).$$

Bod $\mathbf{x}^*$ se zove bodem lokálního minima funkce f vzhledem k množině $\mathcal{M}$.

Analogicky se definuje maximum a ostré extrémy vzhledem k množině, a body těchto extrémů.

7.3 Metody řešení

7.3.1 Obecně

Obecně

Obecná metoda závisí na funkcích f, g_j a h_k:

- substituce,
- lineární programování,
- kvadratické programování,
- konvexní optimalizace,
- ...

Máme-li k dispozici (druhé) parciální derivace všech funkcí, lze je zkusit (pro nelineární problémy) použít. Tyto analytické metody se opět opírají o geometrický význam první a druhé parciální derivace.

7.3.2 Metoda řešení při rovnostní vazbách

Metoda řešení při rovnostních vazbách

Lagrangeova funkce

Mějme úlohu nalézt lokální extrémy funkce $f : D \to \mathbb{R}$, kde $D \subset \mathbb{R}^n$, vzhledem k množině

$$\mathcal{M} = \{\mathbf{x} \in D \colon g_i(\mathbf{x}) = 0, i = 1, 2, \ldots, m\}, \text{kde } g_i : \mathbb{R}^n \to \mathbb{R}.$$

Definice 7.3 Funkci $L : \mathcal{M} \times \mathbb{R}^m \to \mathbb{R}$ definovanou jako

$$L(\mathbf{x}; \lambda) = f(\mathbf{x}) + \sum_{j=1}^{m} \lambda_j g_j(\mathbf{x})$$

nazýváme **Lagrangeovou funkcí** pro danou úlohu.

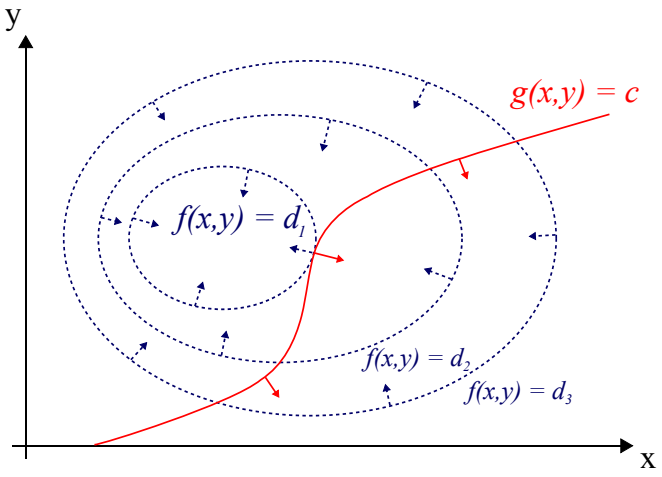
Koeficienty

$$\lambda = (\lambda_1, \ldots, \lambda_m)$$

nazýváme **Lagrangeovy multiplikátory** (*Lagrange multipliers*).

K čemu Lagrangeova funkce?

Uvažujme jednu rovnostní vazbu (tedy $m = 1$). Pokud v bodě x^* platí, že gradient vazby i naší funkce mají stejný směr, tedy $\nabla f(x^*) = -\lambda_1^* \nabla g_1(x^*)$, pak „vazba v daném bodě neprotíná vrstevnici“.



Zdroj: https://en.wikipedia.org/wiki/File:LagrangeMultipliers2D.svg

Věta 7.4 — Postačující podmínka existence ostrého lokálního minima pro rovnostní vazby. Nechť $f, g_j, j \in \{1, \dots, m\}$ mají spojité všechny druhé parciální derivace na nějaké otevřené nadmnožině $\tilde{\mathcal{M}} \supset \mathcal{M}$. Pokud dvojice $(x^*; \lambda^*) \in \mathbb{R}^n \times \mathbb{R}^m$ splňuje podmínky:

(0) (0. derivace) $x^* \in \mathcal{M}$;

(1) (1. derivace) $\forall i, \frac{\partial L}{\partial x_i}(x^*; \lambda^*) = 0$;

(2) (2. derivace) pro každý (sloupcový) vektor $0 \neq v \in \mathbb{R}^n$ splňující

$$\nabla g_j(x^*) \cdot v = 0, \quad \text{pro } \forall j \in \{1, \dots, m\},$$

platí

$$v^T \cdot \nabla_x^2 L(x^*; \lambda^*) \cdot v > 0;$$

kde $\nabla_x^2 L$ je Hessova matice funkce L vzhledem k proměnným $x = (x_1, x_2, \dots, x_n)$,

potom je x^* bodem ostrého lokálního minima.

Všimněme si, že body (0) a (1) jsou ekvivalentní rovnosti $\nabla L(x^*; \lambda^*) = 0$.

7.3.3 Metoda řešení při rovnostních i nerovnostních vazbách

Metoda řešení při rovnostních vazbách i nerovnostních vazbách

Lagrangeova funkce

Mějme úlohu nalézt lokální extrémy funkce $f : D \to \mathbb{R}$, kde $D \subset \mathbb{R}^n$, vzhledem k množině

$$\mathcal{M} = \{\mathbf{x} \in D \colon g_i(\mathbf{x}) = 0, h_j(\mathbf{x}) \leq 0, i = 1, 2, \dots, m, j = 1, \dots, p\}, \text{kde } g_i, h_j : \mathbb{R}^n \to \mathbb{R}.$$

Definice 7.5 — Lagrangeova funkce. Funkci $L : \mathcal{M} \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ definovanou

$$L(x; \lambda; \mu) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x) + \sum_{k=1}^{p} \mu_k h_k(x)$$

nazýváme **Lagrangeovou funkcí** pro danou úlohu (1). Koeficienty

$$\lambda = (\lambda_1, \ldots, \lambda_m) \quad \text{a} \quad \mu = (\mu_1, \ldots, \mu_p)$$

nazýváme **Lagrangeovy multiplikátory** (*Lagrange multipliers*).

Technické předpoklady diferencovatelnosti

Nechť $\widetilde{\mathcal{M}}$ je otevřená **nad**množina $\mathcal{M}$, na níž pro všechny funkce f, g_j pro $j \in \hat{m}$, h_k pro $k \in \hat{p}$ existují druhé parciální derivace.

Funkci L pak chápeme jako funkci $L : \widetilde{\mathcal{M}} \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$.

Pro jednoduchost lze uvažovat $\mathcal{M} \subset \widetilde{\mathcal{M}} = \mathbb{R}^n$.

(Množina $\mathcal{M}$ je uzavřená (díky spojitosti g_j a h_k).)

Aktivní omezení

Pokud máme nerovnostní vazby h_k, je třeba vědět, kdy je splněna rovnost:

Definice 7.6 — Množina aktivních omezení. Pro bod $x \in \mathcal{M}$ definujeme **množinu aktivních omezení**

$$\mathcal{B}(x) = \{k \in \hat{p} \colon h_k(x) = 0\}\,.$$

Tedy $\mathcal{B}(x)$ jsou indexy nerovnostních vazeb takových, že x je na hranici množiny $\{x \colon h_k(x) \leq 0\}$.

Věta 7.7 — Postačující podmínka existence ostrého lokálního minima. Nechť f, g_j, h_k pro $j \in \hat{m}, k \in \hat{p}$ mají spojité všechny druhé parciální derivace na nějaké otevřené nadmnožině $\tilde{\mathcal{M}} \supset \mathcal{M}$. Pokud trojice $(x^*; \lambda^*; \mu^*) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$ splňuje podmínky:

1. (0. derivace) $x^* \in \mathcal{M}$;
2. (1. derivace) $\forall i, \frac{\partial L}{\partial x_i}(x^*; \lambda^*; \mu^*) = 0$;
3. (aktivní a neaktivní vazby) Pro každé $k \in \hat{p}$, $\mu_k^* = 0$ nebo $h_k(x^*) = 0$;
4. (2. derivace) pro každý (sloupcový) vektor $0 \neq v \in \mathbb{R}^n$ splňující

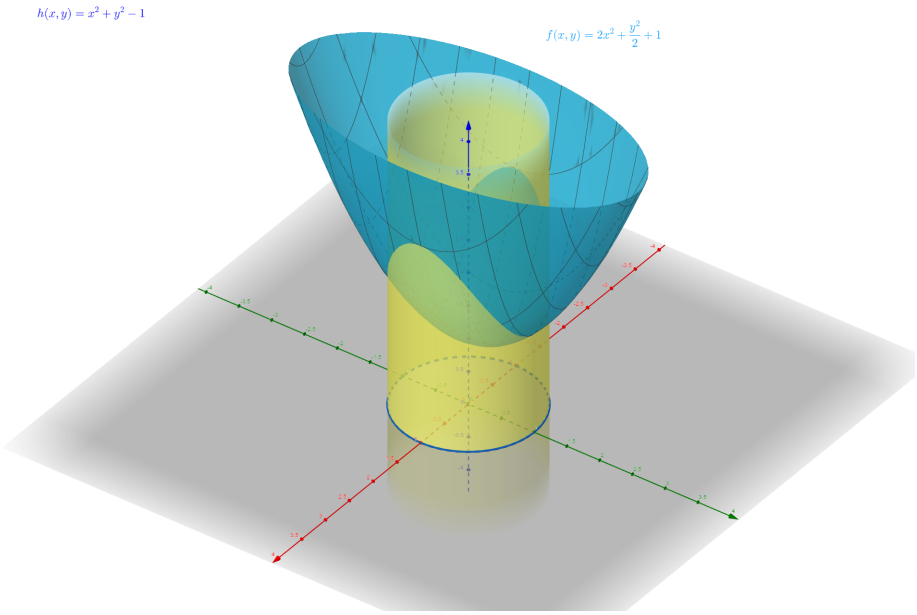
$$\begin{aligned} \nabla g_j(x^*) \cdot v &= 0, \quad \text{pro všechna } j \in \hat{m}, \\ \nabla h_k(x^*) \cdot v &= 0, \quad \text{pro všechna } k \in \hat{p}, \mu_k^* \neq 0, \text{ platí} \end{aligned}$$

$$v^T \cdot \nabla_x^2 L(x^*; \lambda^*; \mu^*) \cdot v > 0,$$

   kde $\nabla_x^2 L$ je Hessova matice funkce L vzhledem k proměnným $x = (x_1, x_2, \ldots, x_n)$;
5. (správný „směr“ od hranice $\mathcal{M}$) $\mu_k^* \geq 0$, pro každé $k \in \hat{p}$.

Potom je x^* bodem ostrého lokálního minima úlohy (1).

Znaménko multiplikátoru



Věta 7.8 — Postačující podmínka existence ostrého lokálního maxima. Nechť f, g_j, h_k pro $j \in \hat{m}, k \in \hat{p}$ mají spojité všechny druhé parciální derivace na nějaké otevřené nadmnožině $\tilde{\mathcal{M}} \supset \mathcal{M}$. Pokud trojice $(x^*; \lambda^*; \mu^*) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$ splňuje podmínky:

1. (0. derivace) $x^* \in \mathcal{M}$;
2. (1. derivace) $\forall i, \frac{\partial L}{\partial x_i}(x^*; \lambda^*; \mu^*) = 0$;
3. (aktivní a neaktivní vazby) Pro každé $k \in \hat{p}$, $\mu_k^* = 0$ nebo $h_k(x^*) = 0$;
4. (2. derivace) pro každý (sloupcový) vektor $0 \neq v \in \mathbb{R}^n$ splňující

$$\begin{aligned} \nabla g_j(x^*) \cdot v &= 0, \quad \text{pro všechna } j \in \hat{m}, \\ \nabla h_k(x^*) \cdot v &= 0, \quad \text{pro všechna } k \in \hat{p}, \mu_k^* \neq 0, \text{ platí} \end{aligned}$$

$$v^T \cdot \nabla_x^2 L(x^*; \lambda^*; \mu^*) \cdot v < 0,$$

   kde $\nabla_x^2 L$ je Hessova matice funkce L vzhledem k proměnným $x = (x_1, x_2, \ldots, x_n)$;
5. (správný „směr“ od hranice $\mathcal{M}$) $\mu_k^* \leq 0$, pro každé $k \in \hat{p}$.

Potom je x^* bodem ostrého lokálního maxima úlohy (1).

Řešené příklady: Vázané extrémy

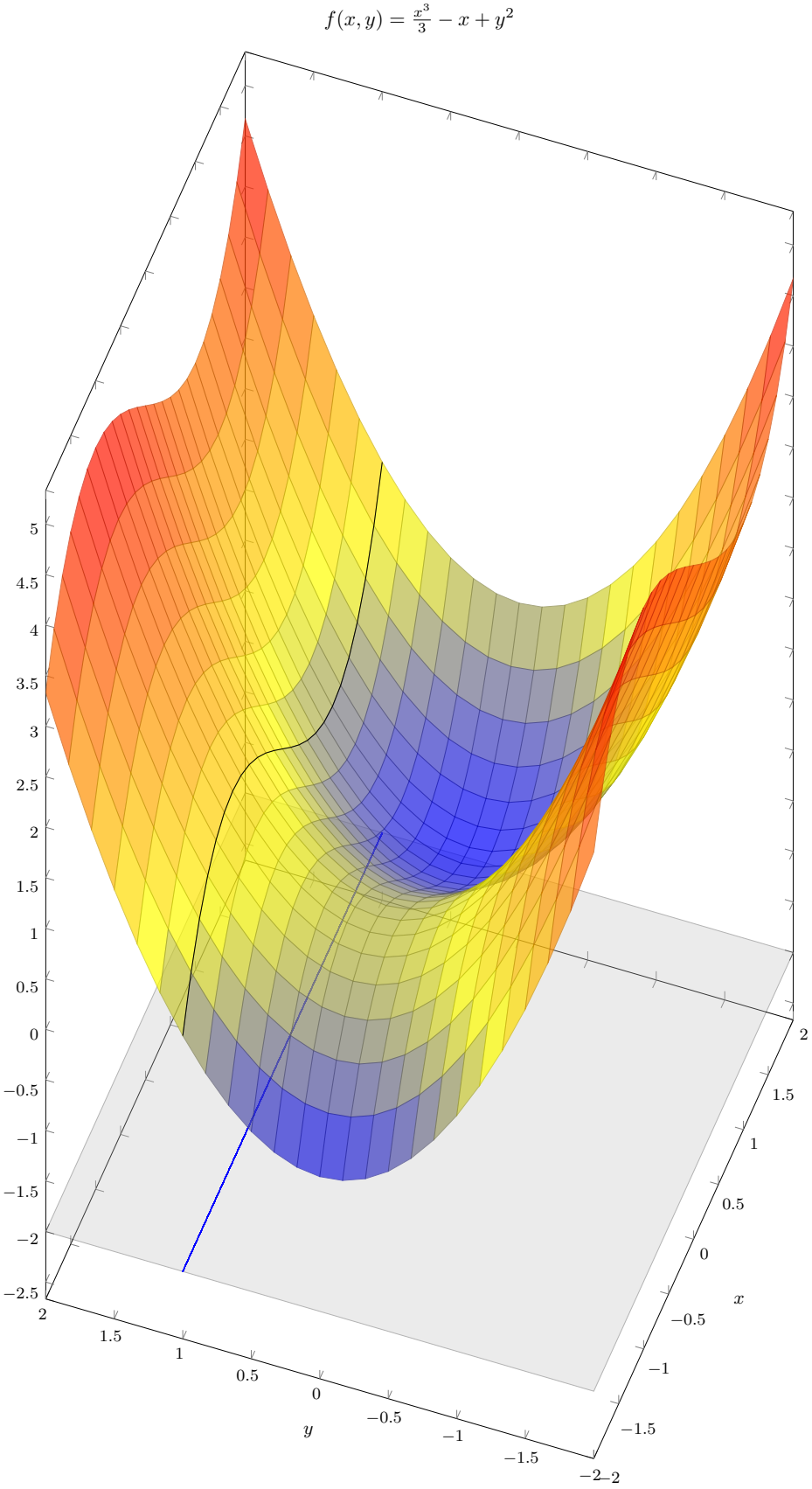
Základní cvičení 10.3
Najděte lokální extrémy funkce $f(x,y) = \frac{x^3}{3} - x + y^2$ za podmínky

a) $g(x,y) = y - 1 = 0$;

b) $g(x,y) = y = 0$;

c) $g(x,y) = x^2 + 2x + y^2 = 0$.

$f(x,y) = \frac{x^3}{3} - x + y^2$ **a** $g(x,y) = y - 1 = 0$



$f(x,y) = \frac{x^3}{3} - x + y^2$ **a** $g(x,y) = y - 1 = 0$

$f(x,y) = \frac{x^3}{3} - x + y^2$ za podmínky $g(x,y) = y - 1 = 0$
(Lze dosadit!)

Obecný postup pro rovnostní vazby:

$f(x,y) = \frac{x^3}{3} - x + y^2$ za podmínky $g(x,y) = y - 1 = 0$

$L(x,y,\lambda) =$

$\nabla L(x^*, y^*, \lambda^*) = 0 \Leftrightarrow$

$f(x,y) = \frac{x^3}{3} - x + y^2$ za podmínky $g(x,y) = y - 1 = 0$

$\nabla_{\mathbf{x}} L(x,y,\lambda) = \left(x^2 - 1, 2y + \lambda\right)$

$\nabla^2_{\mathbf{x}} L(x,y,\lambda) =$

$x^* = 1, y^* = 1, \lambda^* = -2$

$\nabla^2_{\mathbf{x}} L\left(1,1,-2\right) =$

$f(x,y) = \frac{x^3}{3} - x + y^2$ za podmínky $g(x,y) = y - 1 = 0$

$\nabla_{\mathbf{x}} L(x,y,\lambda) = \left(x^2 - 1, 2y + \lambda\right)$

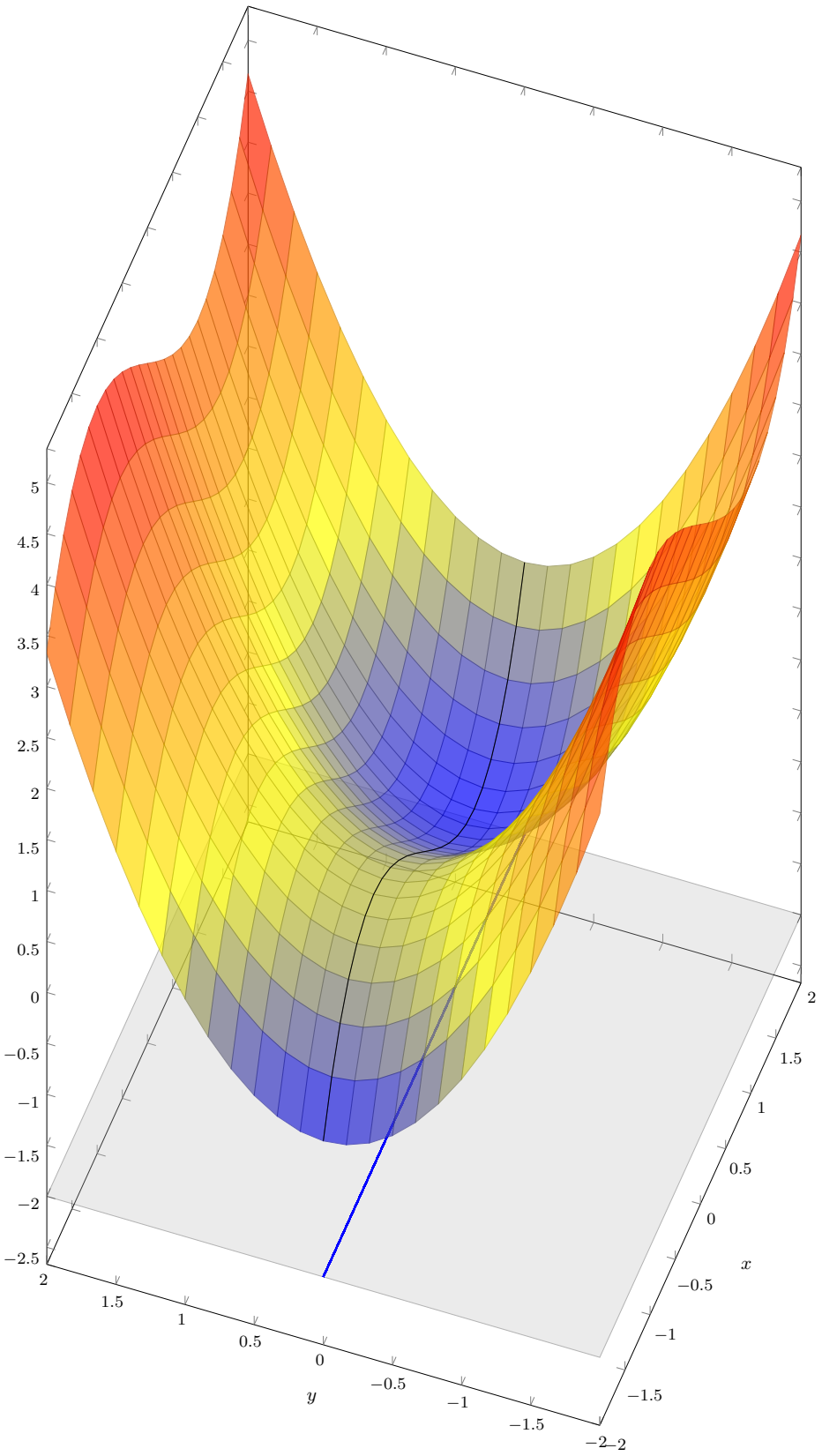
$\nabla^2_{\mathbf{x}} L(x,y,\lambda) = \begin{pmatrix} 2x & 0 \\ 0 & 2 \end{pmatrix}$

$x^* = -1, y^* = 1, \lambda^* = -2$

$\nabla^2_{\mathbf{x}} L\left(-1,1,-2\right) =$

$f(x,y) = \frac{x^3}{3} - x + y^2$ **a** $g(x,y) = y = 0$

$f(x,y) = \frac{x^3}{3} - x + y^2$



$f(x,y) = \frac{x^3}{3} - x + y^2$ **a** $g(x,y) = y = 0$

$f(x,y) = \frac{x^3}{3} - x + y^2$ za podmínky $g(x,y) = y = 0$

$L(x,y,\lambda) = \frac{x^3}{3} - x + y^2 + \lambda\,(y)$

$\nabla L(x^*, y^*, \lambda^*) = 0 \Leftrightarrow$

$f(x,y) = \frac{x^3}{3} - x + y^2$ za podmínky $g(x,y) = y = 0$

$\nabla_{\mathbf{x}} L(x,y,\lambda) = \left(x^2 - 1, 2y + \lambda\right)$

$\nabla^2_{\mathbf{x}} L(x,y,\lambda) = \begin{pmatrix} 2x & 0 \\ 0 & 2 \end{pmatrix}$

$x^* = 1, y^* = 0, \lambda^* = 0$

$\nabla^2_{\mathbf{x}} L\left(1,0,0\right) =$

$f(x,y) = \frac{x^3}{3} - x + y^2$ za podmínky $g(x,y) = y = 0$

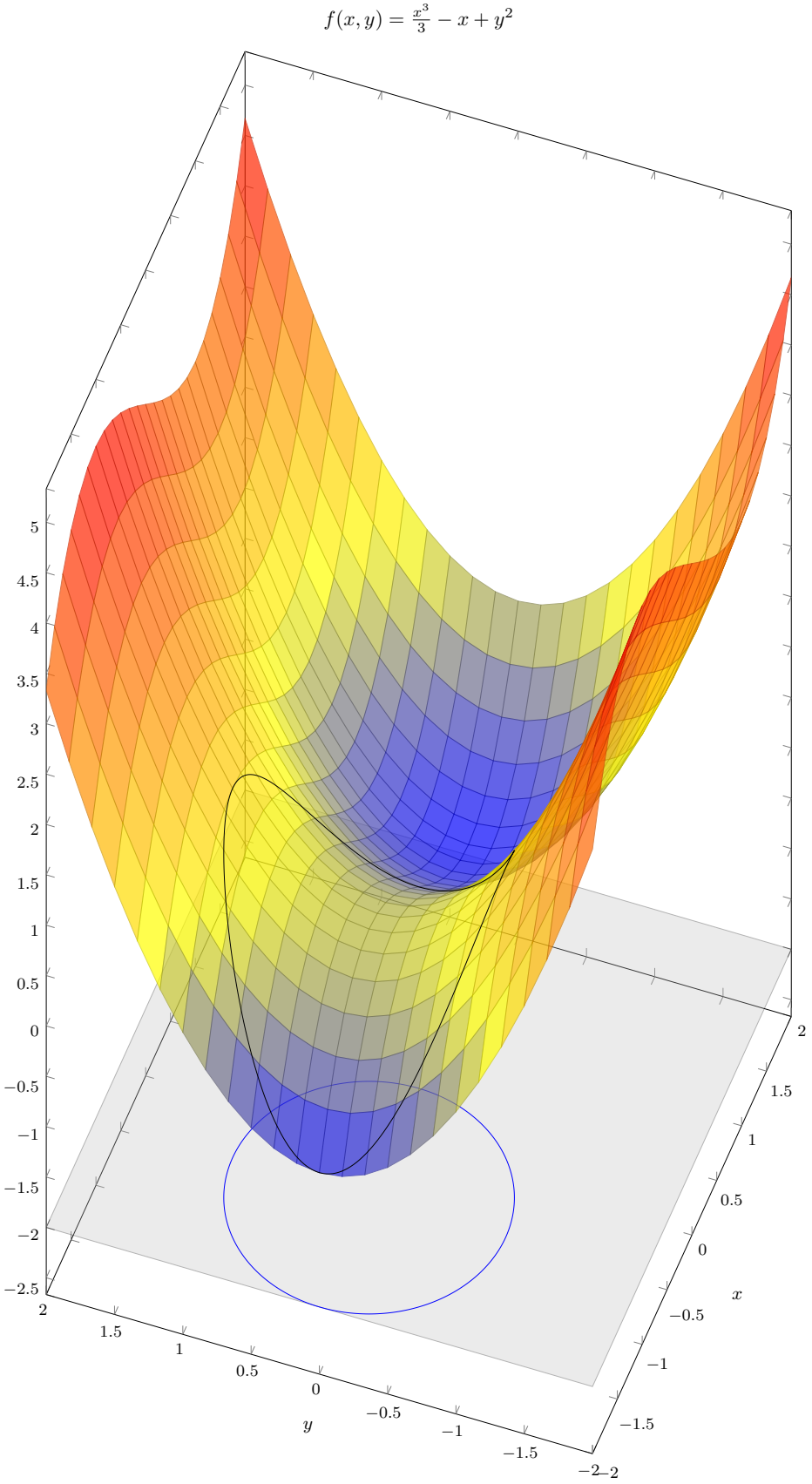
$\nabla_{\mathbf{x}} L(x,y,\lambda) = \left(x^2 - 1, 2y + \lambda\right)$

$\nabla^2_{\mathbf{x}} L(x,y,\lambda) = \begin{pmatrix} 2x & 0 \\ 0 & 2 \end{pmatrix}$

$x^* = -1, y^* = 0, \lambda^* = 0$

$\nabla^2_{\mathbf{x}} L\left(-1,0,0\right) =$

$f(x,y) = \frac{x^3}{3} - x + y^2$ **a** $g(x,y) = x^2 + 2x + y^2 = 0$



$f(x,y) = \frac{x^3}{3} - x + y^2$ **a** $g(x,y) = x^2 + 2x + y^2 = 0$

$f(x,y) = \frac{x^3}{3} - x + y^2$ za podmínky $g(x,y) = x^2 + 2x + y^2 = 0$

$L(x,y,\lambda) =$

$\nabla L(x^*, y^*, \lambda^*) = 0 \Leftrightarrow$

$f(x,y) = \frac{x^3}{3} - x + y^2$ za podmínky $g(x,y) = x^2 + 2x + y^2 = 0$

$\nabla_{\mathbf{x}} L(x,y,\lambda) = \left(x^2 - 1 + 2\lambda x + 2\lambda, 2y + 2\lambda y\right)$

Kritické body: $\left(0,0,\frac{1}{2}\right), \left(-2,0,\frac{3}{2}\right), (-1,1,-1), (-1,-1,-1)$

$\nabla^2_{\mathbf{x}} L(x,y,\lambda) =$

$\nabla^2_{\mathbf{x}} L\left(0,0,\frac{1}{2}\right) =$

$f(x,y) = \frac{x^3}{3} - x + y^2$ za podmínky $g(x,y) = x^2 + 2x + y^2 = 0$

Kritické body: $\left(0,0,\frac{1}{2}\right), \left(-2,0,\frac{3}{2}\right), (-1,1,-1), (-1,-1,-1)$

$\nabla^2_{\mathbf{x}} L(x,y,\lambda) = \begin{pmatrix} 2x + 2\lambda & 0 \\ 0 & 2 + 2\lambda \end{pmatrix}$

$\nabla^2_{\mathbf{x}} L\left(-2,0,\frac{3}{2}\right) =$

$f(x,y) = \frac{x^3}{3} - x + y^2$ za podmínky $g(x,y) = x^2 + 2x + y^2 = 0$

Kritické body: $\left(0,0,\frac{1}{2}\right), \left(-2,0,\frac{3}{2}\right), (-1,1,-1), (-1,-1,-1)$

$\nabla^2_{\mathbf{x}} L(x,y,\lambda) = \begin{pmatrix} 2x + 2\lambda & 0 \\ 0 & 2 + 2\lambda \end{pmatrix}$

$\nabla^2_{\mathbf{x}} L(-1,1,-1) =$

$f(x,y) = \frac{x^3}{3} - x + y^2$ za podmínky $g(x,y) = x^2 + 2x + y^2 = 0$

Kritické body: $\left(0,0,\frac{1}{2}\right), \left(-2,0,\frac{3}{2}\right), (-1,1,-1), (-1,-1,-1)$

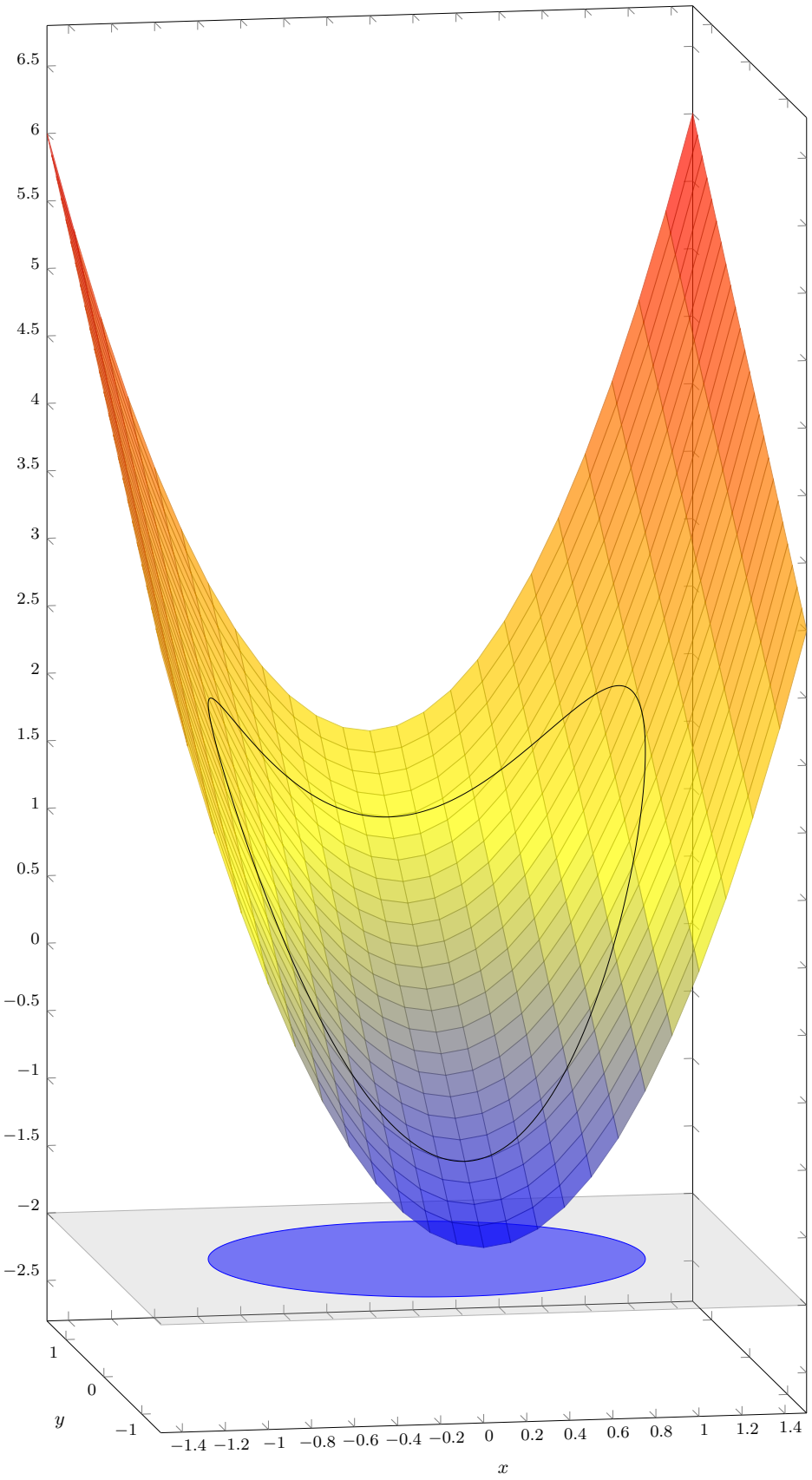
$\nabla^2_{\mathbf{x}} L(x,y,\lambda) = \begin{pmatrix} 2x + 2\lambda & 0 \\ 0 & 2 + 2\lambda \end{pmatrix}$

$\nabla^2_{\mathbf{x}} L(-1,-1,-1) =$

Základní cvičení 11.1
Najděte lokální extrémy funkce $f(x,y) = 2x^2 + y$ za podmínky

$$h(x,y) = x^2 + y^2 \leq 1.$$

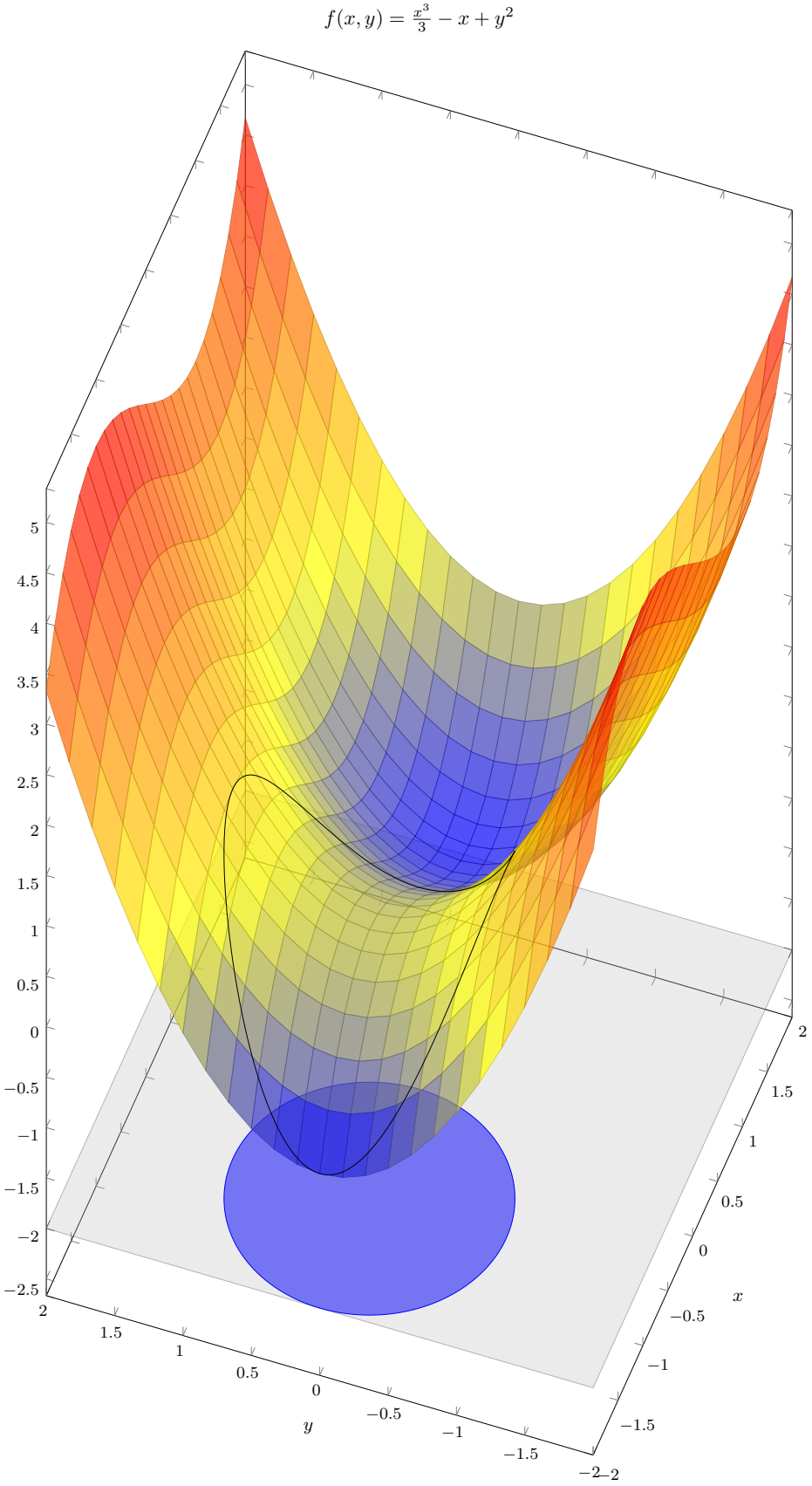
$f(x,y) = 2x^2 + y$ **a** $h(x,y) = x^2 + y^2 \leq 1$



Základní cvičení 11.2
Najděte lokální extrémy funkce $f(x,y) = \frac{x^3}{3} - x + y^2$ za podmínky

$$h(x,y) = x^2 + 2x + y^2 \leq 0.$$

$f(x,y) = \frac{x^3}{3} - x + y^2$ **a** $h(x,y) = x^2 + 2x + y^2 \leq 0$



$f(x,y) = \frac{x^3}{3} - x + y^2$ **a** $h(x,y) = x^2 + 2x + y^2 \leq 0$

$f(x,y) = \frac{x^3}{3} - x + y^2$ za podmínky $h(x,y) = x^2 + 2x + y^2 \leq 0$

Je třeba rozlišit, zda je vazba aktivní či ne.

1) nechť $h(x,y) = 0$ — vazba je aktivní — již máme kritické body pro úlohu s rovnostní vazbou
$(x^*, y^*, \mu^*) \in \left\{ \left(0, 0, \frac{1}{2}\right), \left(-2, 0, \frac{3}{2}\right), (-1, 1, -1), (-1, -1, -1) \right\}$

$f(x,y) = \frac{x^3}{3} - x + y^2$ za podmínky $h(x,y) = x^2 + 2x + y^2 \leq 0$

2) vazba není aktivní — řešíme bez ní a kontrolujeme náležitost do zbytku množiny přípustných řešení $\mathcal{M}$, tj. zde $h(x,y) < 0$, a nastavíme $\mu = 0$, abychom splnili podmínku 3. Věty **??**:

$\nabla_{\mathbf{x}} L(x,y,0) = \left(\frac{\partial L}{\partial x}(x,y,0), \frac{\partial L}{\partial y}(x,y,0) \right)^T = \nabla f(x,y) =$

$\nabla_{\mathbf{x}} L(x,y,0) = 0 \Leftrightarrow$

Část II
Vícerozměrný integrál

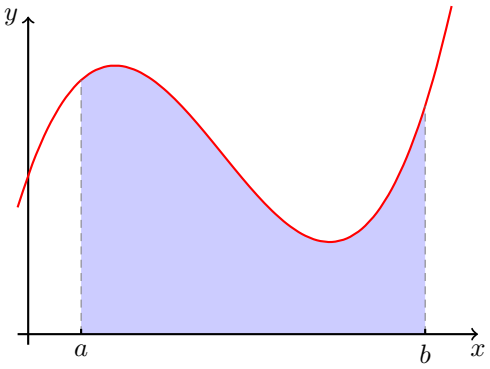
8 Připomenutí: integrace funkce 1 proměnné

8.1

(Určitý) integrál

Integrál je nástroj pro výpočet obsahu „pod grafem“[1] nějaké funkce. Tuto úlohu lze najít v mnoha dalších úlohách:

- objem těles, hledání těžiště, hledání průměrné hodnoty, hledání střední hodnoty náhodné veličiny, hledání pravděpodobnosti (integrace hustoty pravděpodobnosti), …



8.2 Darbouxův/Riemannův integrál funkce jedné proměnné

Konstrukce integrálu – postup

Úkol: spočítejte obsah funkce pod grafem funkce $f(x)$ na uzavřeném intervalu $[a, b]$

- Hlavní myšlenka konstrukce je aproximace plochy pod křivkou pomocí obdélníků:
  - interval rozdělíme na malé kousky (tzv. rozdělení intervalu),
  - na těchto kouscích aproximujeme funkci $f(x)$ vhodně zvolenými konstantními funkcemi (dostaneme takzvané *stupňovité* funkce),
  - obsah pod grafem stupňovité funkce je součet obsahu obdélníků, a tedy snadno spočítatelná veličina.
- Zjemňujeme rozdělení a tím získáváme přesnější a přesnější aproximace hledaného obsahu.
- Přesnou hodnotu získáme tak, že v limitě „pošleme“ šířku výše uvedených malých kousků k nule.

Rozdělení intervalu $[a, b]$

Definice 8.1 Buď dán interval $[a, b]$. Konečnou množinu

$$\sigma = \{x_0, x_1, \dots, x_n\}$$

takovou, že

$$a = x_0 < x_1 < \cdots < x_n = b$$

nazýváme **rozdělením intervalu** $[a, b]$. Bodům x_k, $k = 1, 2, \dots, n-1$, říkáme **dělící body intervalu** $[a, b]$. Číslo

$$\nu(\sigma) = \max\{\Delta_k \colon k = 1, 2, \dots, n\}, \quad \text{kde } \Delta_k = x_k - x_{k-1},\ k = 1, 2, \dots, n,$$

nazýváme **normou rozdělení** σ.

[1] Přesněji mezi grafem, osou parametru a kolmicemi na osu parametrů, které procházejí okraji intervalu, přes který integrál počítáme.

Ekvidistantní rozdělení

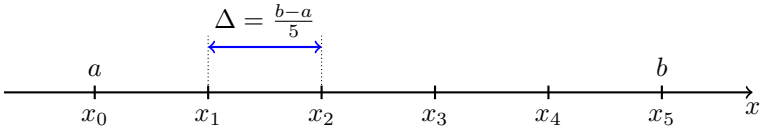
■ **Příklad 8.2 — Ekvidistantní rozdělení.** Pro interval $[a, b]$ a kladné celé n položme $\Delta = \frac{b-a}{n}$ a

$$x_i = a + i \cdot \Delta, \quad i = 0, 1, \ldots, n.$$

Tedy

$$\sigma = \big\{a,\ a+\Delta,\ a+2\Delta, \ldots,\ a+(n-1)\Delta,\ a+n\Delta = b\big\}.$$

■



Definice 8.3 Nechť funkce f je definovaná na intervalu $[a, b]$ a $\sigma = \{x_0, x_1, \ldots, x_n\}$ je rozdělení tohoto intervalu. Označme

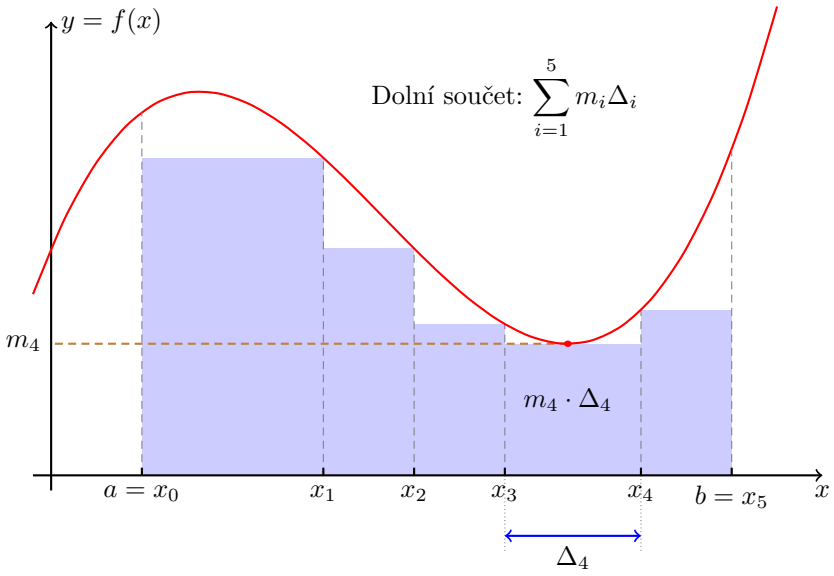
$$M_i = \sup_{x \in [x_{i-1}, x_i]} f(x) \quad \text{a} \quad m_i = \inf_{x \in [x_{i-1}, x_i]} f(x).$$

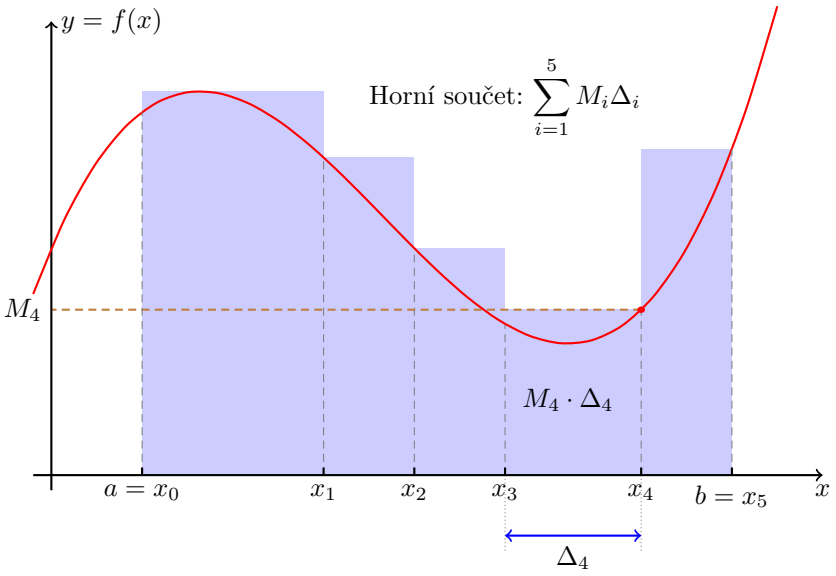
pro každé $i = 1, 2, \ldots, n$. Potom

$$S_f(\sigma) = \sum_{i=1}^{n} M_i \Delta_i \quad \text{a} \quad s_f(\sigma) = \sum_{i=1}^{n} m_i \Delta_i$$

nazýváme **horním**, resp. **dolním**, **(Darbouxovým) součtem funkce** f při rozdělení σ.

Dolní, resp. horní, součty představují obsah plochy tvořené obdélníky pod, resp. nad, grafem funkce. Následující obrázky jsou ilustrativní.





Definice Darbouxova intergrálu

Horní Darbouxův integrál (funkce f na $[a,b]$) je

$$D_f = \inf\left\{S_f(\sigma)\colon \sigma \text{ je rozdělení } [a,b]\right\}$$

a **dolní Darbouxův integrál (funkce f na $[a,b]$)** je

$$d_f = \sup\left\{s_f(\sigma)\colon \sigma \text{ je rozdělení } [a,b]\right\}.$$

Pokud $D_f = d_f$, nazveme tuto hodnotu **Darbouxovým integrálem** funkce f na intervalu $[a,b]$ a značíme ji

$$\int_a^b f(x)\mathrm{d}x = D_f = d_f.$$

Říkáme, že f je **(Darbouxovsky) integrabilní** na $[a,b]$.

Jiné značení: $\int_a^b f$.
Posloupnost rozdělení σ_n nazveme **normální**, pokud pro její normy platí

$$\lim_{n\to\infty} \nu(\sigma_n) = 0.$$

Věta 8.4 Buď f spojitá na $[a,b]$. Potom existuje $\int_a^b f(x)\mathrm{d}x$. Je-li σ_n normální posloupnost rozdělení, potom

$$\lim_{n\to\infty} s_f(\sigma_n) \quad \text{a} \quad \lim_{n\to\infty} S_f(\sigma_n)$$

existují a jsou rovny $\int_a^b f(x)\mathrm{d}x$.

Riemannův integrál je definovaný velice podobně, jen se použije jiná stupňovitá funkce. Ve výsledku je to ale jedno, Riemannova a Darbouxova definice je ekvivalentní. Darbouxova je o něco málo názornější, proto ji používáme.

■ **Příklad 8.5** Vypočtěte integrál funkce $f(x) = x$ na intervalu $J = [0,1]$.

Zvolme normální posloupnost (σ_n) ekvidistantních rozdělení intervalu J.

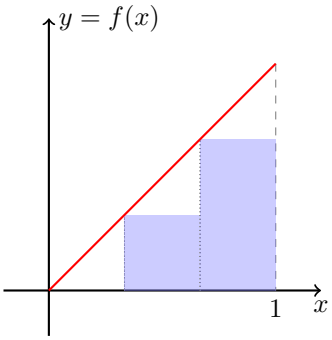
$$\sigma_n = \left\{0 = x_0^{(n)}, x_1^{(n)}, \ldots, x_n^{(n)} = 1\right\}, \quad x_i^{(n)} = i \cdot \frac{1}{n}, \quad i = 0,1,\ldots,n.$$

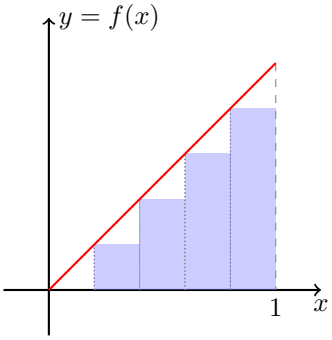
Pro dolní součet při rozdělení σ_n dostáváme

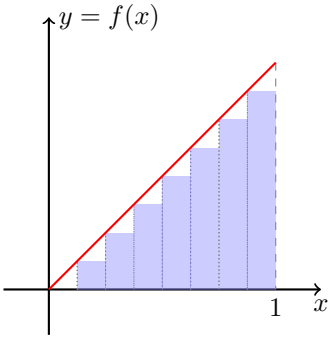
$$s_f(\sigma_n) = \sum_{i=1}^{n} x_{i-1}^{(n)} \cdot \frac{1}{n} = \frac{1}{n^2} \cdot \frac{n(n-1)}{2}.$$

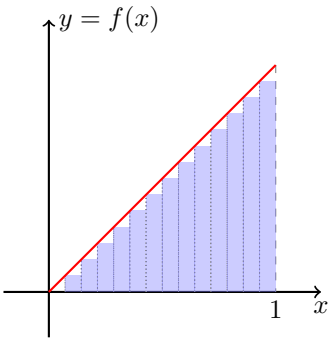
Protože f je spojitá, platí

$$\int_0^1 x \,\mathrm{d}x = \lim_{n \to +\infty} s_f(\sigma_n) = \frac{1}{2}.$$









Vlastnosti Darbouxova/Riemannova integrálu

Věta 8.6 — Aditivita integrálu. Nechť f a g jsou spojité funkce na intervalu $[a, b]$. Potom pro integrál funkce $f + g$!

(která je také automaticky spojitá na $[a,b]$) platí

$$\int_a^b (f+g)(x)\mathrm{d}x = \int_a^b f(x)\mathrm{d}x + \int_a^b g(x)\mathrm{d}x.$$

Věta 8.7 — Multiplikativita integrálu. Nechť f je spojitá na intervalu $[a,b]$ a $c \in \mathbb{R}$ je konstanta. Potom pro integrál funkce cf platí

$$\int_a^b (cf)(x)\mathrm{d}x = c\int_a^b f(x)\mathrm{d}x.$$

!

Primitivní funkce

Pokud si nepamatujete, co je primitivní funkce, projděte si relevantní přednášku předmětu BI-MA2. Zde si zopakujeme pouze definici.

Primitivní funkce (resp. neurčitý integrál) k funkci f je taková funkce F, pro kterou platí že $f = F'$. Hledání primitivní funkce je tedy něco jako inverzní proces k derivování.

Definice 8.8 Nechť funkce f je definována v intervalu (a,b), kde $-\infty \le a < b \le +\infty$. Funkci F splňující podmínku

$$F'(x) = f(x) \text{ pro každé } x \in (a,b)$$

nazýváme **primitivní funkcí** k funkci f v intervalu (a,b).

Primitivní funkce elementárních funkcí

Ze znalosti derivací můžeme ihned sestavit tabulku primitivních funkcí:

vzorec	interval, parametry
$\int x^n\mathrm{d}x = \frac{x^{n+1}}{n+1} + C$	$x \in \mathbb{R}$, $n \in \mathbb{N}$
$\int x^n\mathrm{d}x = \frac{x^{n+1}}{n+1} + C$	$x \in \mathbb{R} \setminus \{0\}$, $n \in \mathbb{Z}$, $n \le -2$
$\int x^\alpha\mathrm{d}x = \frac{x^{\alpha+1}}{\alpha+1} + C$	$x \in (0,+\infty)$, $\alpha \notin \mathbb{Z}$
$\int \frac{1}{x}\mathrm{d}x = \ln\lvert x\rvert + C$	$x \in \mathbb{R} \setminus \{0\}$
$\int a^x\mathrm{d}x = \frac{a^x}{\ln a} + C$	$x \in \mathbb{R}$, $a > 0$ a $a \ne 1$
$\int \sin(x)\mathrm{d}x = -\cos(x) + C$	$x \in \mathbb{R}$

vzorec	interval, parametry
$\int \cos(x)\mathrm{d}x = \sin(x) + C$	$x \in \mathbb{R}$
$\int \frac{1}{\cos^2(x)}\mathrm{d}x = \mathrm{tg}(x) + C$	$x \in \left(-\frac{\pi}{2} + k\pi,\ \frac{\pi}{2} + k\pi\right)$, $k \in \mathbb{Z}$
$\int \frac{1}{\sin^2(x)}\mathrm{d}x = -\mathrm{cotg}(x) + C$	$x \in \left(k\pi,\ \pi + k\pi\right)$, $k \in \mathbb{Z}$
$\int \frac{1}{\sqrt{1-x^2}}\mathrm{d}x = \arcsin(x) + C$	$x \in (-1,1)$
$\int \frac{1}{1+x^2}\mathrm{d}x = \mathrm{arctg}(x) + C$	$x \in \mathbb{R}$

Newtonova formule

Následující věta odhaluje vztah mezi určitým a neurčitým integrálem. Umožňuje nám počítat integrál bez explicitního použití definice s limitou.

Věta 8.9 — Newtonova formule. Nechť f je funkce spojitá na intervalu $[a,b]$ s primitivní funkcí F na (a,b). Pak platí rovnost

$$\int_a^b f(x)\mathrm{d}x = \lim_{x\to b_-} F(x) - \lim_{x\to a_+} F(x).$$

Rozdíl na pravé straně má zaběhlé značení: $\Big[F(x)\Big]_a^b$.

Per partes pro určitý integrál

Věta 8.10 Nechť f a g jsou funkce spojité na $[a,b]$, f má spojitou derivaci na intervalu $[a,b]$ a nechť G je primitivní funkce k funkci g na intervalu $[a,b]$. Potom

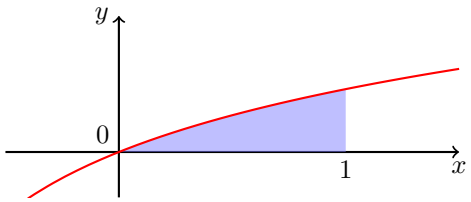
$$\int_a^b f(x)g(x)\mathrm{d}x = \big[f(x)G(x)\big]_a^b - \int_a^b f'(x)G(x)\mathrm{d}x.$$

■ **Příklad 8.11** Vypočtěte

$$\int_0^1 \ln(1+x)\mathrm{d}x.$$

Derivujeme $\ln(1+x)$ a integrujeme 1,

$$\int_0^1 \ln(1+x)\mathrm{d}x = \Big[x\ln(1+x)\Big]_0^1 - \int_0^1 \frac{x}{1+x}\mathrm{d}x =$$
$$= \ln(2) - \Big[x - \ln|1+x|\Big]_0^1 = 2\ln(2) - 1.$$



■

Substituce v určitém integrálu

Zavádíme následující značení

- $\int_a^a f = 0$,
- pro $a > b$ klademe $\int_a^b f = -\int_b^a f$.

Věta 8.12 — O substituci. Nechť pro funkce f a φ platí

1. φ a její derivace φ' jsou spojité na $[\alpha,\beta]$,
2. f je spojitá na $\varphi\big([\alpha,\beta]\big)$.

Potom

$$\int_\alpha^\beta f\big(\varphi(t)\big)\cdot\varphi'(t)\mathrm{d}t = \int_{\varphi(\alpha)}^{\varphi(\beta)} f(x)\mathrm{d}x.$$

■ **Příklad 8.13** Vypočtěte integrál

$$\int_0^{\ln(2)} \frac{e^{-x}}{\frac{1}{2}+e^{-x}}\mathrm{d}x.$$

Použijeme substituci $y = \varphi(x) = \frac{1}{2} + e^{-x}$. Potom

$$\int_0^{\ln(2)} \frac{e^{-x}}{\frac{1}{2} + e^{-x}} \mathrm{d}x = -\int_{\frac{3}{2}}^{1} \frac{1}{y} \mathrm{d}y = \Big[\ln|y|\Big]_1^{\frac{3}{2}} = \ln\frac{3}{2}.$$

■

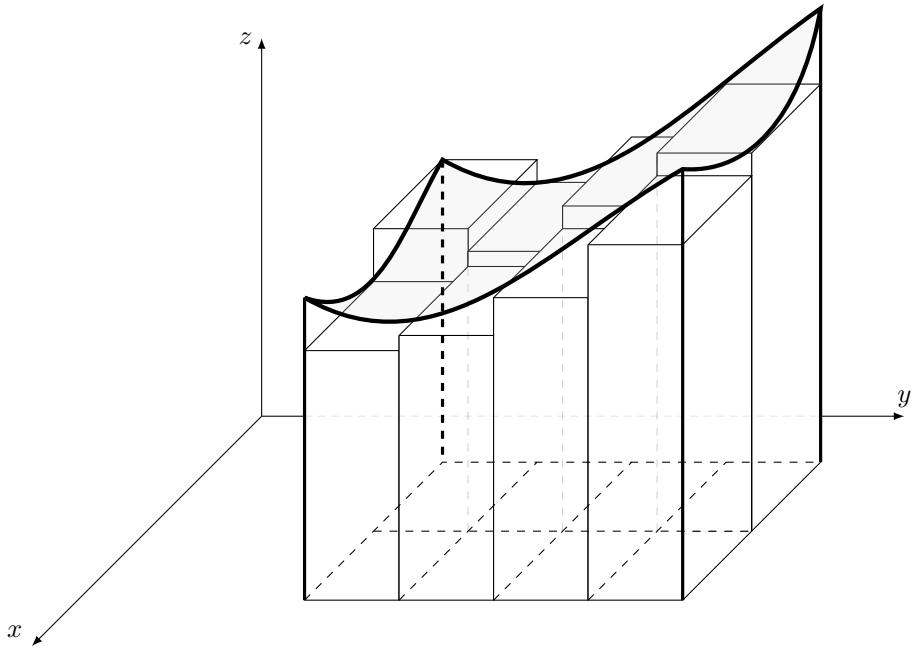
9 Vícerozměrný integrál

9.1 Funkce 2 proměnných

Funkce 2 proměnných

Mějme $f : D \to \mathbb{R}$, kde $D = [a, b] \times [c, d]$.

Graf této funkce si lze představit jako část povrchu nějakého předmětu.
Integrálem této funkce budeme počítat objem pod tímto grafem.



9.1.1 Obdélníková oblast

Definice

Nechť $\sigma_x = (x_i)_{i=0}^n$ je rozdělení $[a, b]$ a $\sigma_y = (y_j)_{j=0}^m$ rozdělení $[c, d]$.
$\sigma = \sigma_x \times \sigma_y$ je rozdělením $D = [a, b] \times [c, d]$.

Označme $M_{i,j} = \sup\Big\{f(x, y) \colon (x, y) \in [x_{i-1}, x_i] \times [y_{j-1}, y_j]\Big\}$ a $m_{i,j} = \inf\Big\{f(x, y) \colon (x, y) \in [x_{i-1}, x_i] \times [y_{j-1}, y_j]\Big\}$.

Horní Darbouxova suma f vzhledem k rozdělení σ je

$$S_f(\sigma) = \sum_{i=1}^{n}\sum_{j=1}^{m} M_{i,j}(x_i - x_{i-1})(y_j - y_{j-1})$$

a **dolní Darbouxova suma f vzhledem k rozdělení σ** je

$$s_f(\sigma) = \sum_{i=1}^{n}\sum_{j=1}^{m} m_{i,j}(x_i - x_{i-1})(y_j - y_{j-1})$$

Definice ...

Horní Darbouxův integrál (funkce f na D) je

$$D_f = \inf\Big\{S_f(\sigma)\colon \sigma \text{ je rozdělení } D\Big\}$$

a **dolní Darbouxův integrál (funkce f na D)** je

$$d_f = \sup\Big\{s_f(\sigma)\colon \sigma \text{ je rozdělení } D\Big\}.$$

Pokud $D_f = d_f$, tak tuto hodnotu nazýváme **(*dvojitým*) Darbouxovým integrálem** funkce f na D a značíme ji

$$\iint_D f(x,y)\mathrm{d}x\mathrm{d}y = D_f = d_f.$$

Zkrácené značení: $\iint_D f$, $\int_D f$, $\int_D f(x,y)\mathrm{d}x\mathrm{d}y$.
Řekneme, že f je **(Darbouxovsky) integrabilní** na D.

Definice je ekvivalentní s Riemannovou definicí.

Spojité funkce

Řekneme, že posloupnost rozdělení $\sigma_n = \sigma_{x,n} \times \sigma_{y,n}$ množiny D je **normální**, jsou-li $\sigma_{x,n}$ i $\sigma_{y,n}$ normální.

Analogicky k jednorozměrnému případu:
je-li f spojitá na D, pak integrál $\iint_D f$ existuje a je roven $\lim_{n\to\infty} s_f(\sigma_n)$ a $\lim_{n\to\infty} S_f(\sigma_n)$ pro libovolnou normální posloupnost rozdělení σ_n množiny D.

Výpočet dvojného integrálu nad obdélníkovou oblastí

Následující věta nám říká, jak převést problém výpočtu dvojného integrálu na dva jednodimenzionální podproblémy.

Věta 9.1 Buď $f(x,y)$ integrabilní funkce na $D = [a,b] \times [c,d]$. Pokud existuje jeden z integrálů

$$\int_a^b \left(\int_c^d f(x,y)\mathrm{d}y\right)\mathrm{d}x \qquad \text{nebo} \qquad \int_c^d \left(\int_a^b f(x,y)\mathrm{d}x\right)\mathrm{d}y$$

potom je roven dvojnému integrálu

$$\iint_D f(x,y)\mathrm{d}x\mathrm{d}y.$$

Výpočet dvojného integrálu tedy můžeme provést tak, že funkci nejdříve zintegrujeme vzhledem k jedné proměnné a druhou považujeme za konstantu. Výsledek této integrace (získaný pomocí Newtonovy formule) potom již závisí pouze na jedné proměnné, vzhledem ke které provedeme druhou integraci.

9.1.2 Obecná oblast

Obecná oblast

Je-li D omezená podmnožina $\mathbb{R}^2$, pak definujeme Darbouxův integrál na D následovně:

Definice 9.2 Mějme $f : D \to \mathbb{R}$, kde $D \subset \widetilde{D} = [a,b] \times [c,d]$.
Definujeme *dvojitý Darbouxův integrál* funkce f na D jako hodnotu

$$\iint_D f := \iint_{\widetilde{D}} \tilde{f},$$

kde

$$\tilde{f}(x) = \begin{cases} f(x) & \text{pro } x \in D \\ 0 & \text{pro } x \in \widetilde{D} \setminus D, \end{cases}$$

pokud existuje.

Uvedená definice nezávisí na volbě obdélníka $\widetilde{D}$.

Množina míry nula

Definice 9.3 Řekneme, že množina $Z \subset \mathbb{R}^2$ má *míru nula* pokud pro každé $\varepsilon > 0$ existují obdélníky $R_i = [a_i, b_i] \times [c_i, d_i]$ pro $i = 1, \ldots, n$ tak, že

$$Z \subset \bigcup_{i=1}^{n} R_i \qquad \text{a} \qquad \sum_{i=1}^{n} |b_i - a_i||d_i - c_i| < \varepsilon.$$

- Množiny míry nula mají tu vlastnost, že jsou pro hodnotu integrálu „zanedbatelné“.
- Graf spojité funkce $\varphi : [a, b] \to \mathbb{R}$ má míru nula.
- Mluvíme-li o nějaké vlastnosti bodů množiny $M \subset \mathbb{R}^2$, řekneme, že platí *skoro všude* (*almost everywhere*), pokud množina, kde neplatí, má míru nula. Říkáme tak např., že funkce f a g jsou rovny skoro všude, pokud množina $\{x \in \mathbb{R}^2 : f(x) \neq g(x)\}$ má míru nula.
- V pravděpodobnosti (např. v předmětu NI-VSM) se v obdobném kontextu používá termín *skoro jistě*.
- Množinu míry nula lze zavést analogicky i v obecném $\mathbb{R}^n$. Jak?

Věta 9.4 Omezená funkce $f : D \to \mathbb{R}$, kde $D = [a, b] \times [c, d]$, je integrabilní, pokud množina $\{x \in D : f$ není spojitá v $x\}$ má míru nula. (Jinými slovy, pokud f je spojitá skoro všude na D.)

Připomeňme, že *hranice* množiny $D \subset \mathbb{R}^n$ je množina všech bodů $x \in \mathbb{R}^n$ takových, že každé okolí $H(x)$ má neprázdný průnik jak s D tak s $\mathbb{R}^n \setminus D$.

Důsledek 9.5 Omezená spojitá funkce $f : D \to \mathbb{R}$ na omezené množině D mající hranici míry nula, je integrabilní.

Vlastnosti dvojného integrálu

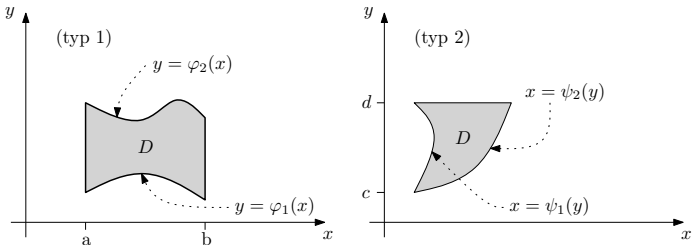
- Pokud $D = D_1 \cup D_2$, kde D_1 i D_2 jsou uzavřené omezené množiny, $D_1 \cap D_2$ má míru nula a f je integrabilní na D, pak platí
$$\iint_D f = \iint_{D_1} f + \iint_{D_2} f.$$
- Platí-li pro (skoro) všechna $(x, y) \in D$ a pro integrabilní funkce f_1 a f_2, že $f_1(x, y) \leq f_2(x, y)$, potom
$$\iint_D f_1 \leq \iint_D f_2.$$
- Pro $c \in \mathbb{R}$ a integrabilní funkci f platí
$$\iint_D c \cdot f(x, y)\mathrm{d}x\mathrm{d}y = c \cdot \iint_D f(x, y)\mathrm{d}x\mathrm{d}y.$$

Oblasti s hranicí danou spojitými funkcemi

Jak integrál nad obecnou oblastí spočítat? (Jinak než z definice...)

Zatím jsme si ukázali, jak spočítat integrál funkce přes obdélníkovou oblast. Teď si ukážeme, jak integrovat i přes oblasti, které jsou vymezené spojitými funkcemi. Budeme uvažovat dva typy oblastí D:

- (typ 1) x je z intervalu $[a,b]$ a y je omezené spoj. funkcemi $\varphi_1(x)$ a $\varphi_2(x)$ splňujícími $\varphi_1(x) \leq \varphi_2(x)$,
- (typ 2) y je z intervalu $[c,d]$ a x je omezené spoj. funkcemi $\psi_1(y)$ a $\psi_2(y)$ splňujícími $\psi_1(y) \leq \psi_2(y)$.



Výpočet dvojného integrálu nad obecnou oblastí – myšlenka

Myšlenka:

- Pro oblast typu 1 zafixujme hodnotu x na x_0, nad vzniklým řezem oblasti D nám vznikne funkce $f(x_0, y)$ jedné proměnné y.
- Plocha nad tímto řezem závisí na x_0 a je rovna $p(x_0) = \int_{\varphi_1(x_0)}^{\varphi_2(x_0)} f(x_0, y)\mathrm{d}y$.
- Nyní „posčítáme“ takto získané jednorozměrné plochy přes všechna x od a do b a dostaneme

$$\iint_D f(x,y)\mathrm{d}x\mathrm{d}y = \int_a^b \left(\underbrace{\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\mathrm{d}y}_{=p(x)} \right) \mathrm{d}x.$$

Výpočet dvojného integrálu nad obecnou oblastí

Věta 9.6 Pokud integrály (ve vzorcích níže) existují, platí pro oblast D, že

- je-li D typu 1, máme
$$\iint_D f(x,y)\mathrm{d}x\mathrm{d}y = \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\mathrm{d}y \right) \mathrm{d}x.$$
- je-li D typu 2, máme
$$\iint_D f(x,y)\mathrm{d}x\mathrm{d}y = \int_c^d \left(\int_{\psi_1(y)}^{\psi_2(y)} f(x,y)\mathrm{d}x \right) \mathrm{d}y.$$

Výpočet dvojného integrálu nad obecnou oblastí příklad

■ **Příklad 9.7** Vypočítejte integrál $f(x,y) = xy$ nad oblastí D typu 1, kde $x \in [0,1]$ a y je sevřené funkcemi x^2 a x^3:

$$\iint_D xy\ \mathrm{d}x\mathrm{d}y = \int_0^1 \left(\int_{x^3}^{x^2} xy\ \mathrm{d}y \right) \mathrm{d}x = \int_0^1 \left(\left[\frac{xy^2}{2} \right]_{y=x^3}^{y=x^2} \right) \mathrm{d}x =$$
$$\int_0^1 \frac{xx^4 - xx^6}{2} \mathrm{d}x =$$
$$= \frac{1}{2}\left[\frac{x^6}{6} - \frac{x^8}{8} \right]_0^1 = \frac{1}{12} - \frac{1}{16} = \frac{1}{48}.$$

■

9.1.3 Aplikace

Aplikace (dvojného) integrálu

Pomocí dvojného integrálu můžeme spočítat několik užitečných čísel charakterizujících daný objem pod grafem funkce f nad oblastí D:

- průměr (jako objem lomeno povrch oblasti D):
$$\left(\iint_D f(x,y)\mathrm{d}x\mathrm{d}y\right) / \left(\iint_D 1\ \mathrm{d}x\mathrm{d}y\right).$$
- těžiště desky D s proměnnou hustotou $\rho(x,y)$ má souřadnice $(\bar{x},\bar{y})$:
$$\bar{x} = \left(\iint_D x\rho(x,y)\mathrm{d}x\mathrm{d}y\right) / \left(\iint_D \rho(x,y)\ \mathrm{d}x\mathrm{d}y\right),$$
$$\bar{y} = \left(\iint_D y\rho(x,y)\mathrm{d}x\mathrm{d}y\right) / \left(\iint_D \rho(x,y)\ \mathrm{d}x\mathrm{d}y\right).$$
- Povrch grafu $f(x,y)$ nad D je
$$\iint_D \sqrt{1+\left(\frac{\partial f}{\partial x}(x,y)\right)^2+\left(\frac{\partial f}{\partial y}(x,y)\right)^2}\mathrm{d}x\mathrm{d}y.$$

9.2 Funkce více proměnných

Trojný integrál

Konstrukce trojného integrálu je naprosto analogická konstrukci integrálu dvojného, pouze obdélníčky Δ_{ij} jsou nahrazeny „kvádříčky“ Δ_{ijk} a integrujeme funkci tří proměnných $f(x,y,z)$:

$$\iiint_D f(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z$$

Výpočet lze opět převést na tři výpočty jednorozměrného integrálu, např.

$$\int_e^f\int_c^d\int_a^b f(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z = \int_e^f\left(\int_c^d\left(\int_a^b f(x,y,z)\mathrm{d}x\right)\mathrm{d}y\right)\mathrm{d}z\,.$$

Existuje ovšem 3! možných pořadí integrování.

Konstrukce pro obecné funkce nad $\mathbb{R}^n$ probíhá zcela analogicky a platí analogická tvrzení. (Nicméně se pro definici v naprosté většině případů nepoužívá Darbouxova/Riemannova definice, ale jiná, např. Lebesgueova.)

Definice 9.8 Mějme $\Psi : \mathbb{R}^n \to \mathbb{R}^n$, $\Psi(\mathbf{v}) = (\Psi_1(\mathbf{v}),\dots,\Psi_n(\mathbf{v}))$. **Jacobiho matice** zobrazení Ψ je následující zobrazení $\mathbb{R}^n \to \mathbb{R}^{n,n}$ (pro $\mathbf{v} = (v_1, v_2, \dots, v_n)$)

$$J_\Psi = \begin{pmatrix} \frac{\partial\Psi_1}{\partial v_1} & \cdots & \frac{\partial\Psi_1}{\partial v_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial\Psi_n}{\partial v_1} & \cdots & \frac{\partial\Psi_n}{\partial v_n} \end{pmatrix},$$

pokud všechny parciální derivace existují.

Jacobiho matice zobrzení Ψ má na řádcích složky gradientů jednotlivých složek Ψ. Toto se (s drobným zneužitím značením) zapíše takto:

$$J_\Psi = \begin{pmatrix} \nabla\Psi_1 \\ \vdots \\ \nabla\Psi_n \end{pmatrix}.$$

Věta 9.9 — Věta o substituci. Nechť D je omezená uzavřená množina na $\mathbb{R}^n$. Nechť $\Psi : \mathbb{R}^n \to \mathbb{R}^n$ má spojité všechny parciální derivace (všech složek) na nějaké otevřené nadmnožině množiny D a *skoro všude* na D platí, že

1. Ψ je bijekce a
2. $\det J_\Psi$ je nenulový.

Potom pro každou spojitou funkci $f : D \to \mathbb{R}$ platí

$$\int_{\psi(D)} f(\mathbf{x})\mathrm{d}\mathbf{x} = \int_D f(\Psi(\mathbf{v}))\,|\det J_\Psi(\mathbf{v})|\,\mathrm{d}\mathbf{v}$$

kde $\mathbf{x} = (x_1, x_2, \ldots, x_n)$.

Řešené příklady: Integrály

Příklad 1: obdélníková oblast

Základní cvičení 12.1
Buď $f(x,y) = \frac{x^2}{1+y^2}$ a $D = [2,3] \times [0,3]$. Spočítejte

$$\int_D f(x,y)\mathrm{d}x\mathrm{d}y.$$

Příklad 2: obecnější oblast

Základní cvičení 13.1
Buď $f(x,y) = xy$. Spočítejte

$$\int_D f(x,y)\mathrm{d}x\mathrm{d}y,$$

kde D je omezená množina ohraničená křivkami $y^2 = x$ a $y = x - 2$.

$\int_D xy = \int_{-1}^{2} \left(\int_{y^2}^{y+2} xy\mathrm{d}x \right) \mathrm{d}y$

Příklad 3: obecnější oblast

Základní cvičení 14.1
Buď $f(x,y) = 3x + 2y - 1$. Spočítejte

$$\int_{\widetilde{D}} f(x,y)\mathrm{d}x\mathrm{d}y,$$

kde $\widetilde{D} = \big\{(x,y)\colon 1 \le x^2 + y^2 \le 4 \text{ a } x \le y\big\}$.

$\widetilde{D}$
$\Psi(D) = \widetilde{D}$

$\Psi : \begin{cases} x = \\ y = \end{cases}$

$\Psi\left(\quad\right) = \left(\qquad\qquad\qquad\right)$

$J_\Psi\left(\quad\right) = \begin{pmatrix} & \\ & \end{pmatrix}$
$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \mathbf{v} = \begin{pmatrix} r \\ \varphi \end{pmatrix}, \mathbf{x} = \Psi(\mathbf{v})$
$\Psi^{-1}(\widetilde{D}) = D = [1,2] \times \left[\frac{\pi}{4}, \frac{5\pi}{4}\right]$
$J_\Psi \begin{pmatrix} r \\ \varphi \end{pmatrix} = \begin{pmatrix} \cos(\varphi) & -r\sin(\varphi) \\ \sin(\varphi) & r\cos(\varphi) \end{pmatrix}$

$\det J_\Psi \begin{pmatrix} r \\ \varphi \end{pmatrix} =$

$\int_D (3x + 2y - 1)\mathrm{d}x\mathrm{d}y =$
$\int_{\widetilde{D}} (3x + 2y - 1)\mathrm{d}x\mathrm{d}y = \int_{\frac{\pi}{4}}^{\frac{5\pi}{4}} \left(\int_1^2 (3r\cos\varphi + 2r\sin\varphi - 1)|r|\mathrm{d}r \right) \mathrm{d}\varphi =$

Doplnění k větě o postupném integrování

Mějme $D = [0,1] \times [0,1]$.
Uvažme funkci

$$f(x,y) = \begin{cases} 1 & \text{pokud } x = \frac{1}{2} \text{ a } y = \frac{\ell}{2^i} \text{ pro všechna kladná celá } \ell, i; \\ 0 & \text{jinak.} \end{cases}$$

$\iint_D f(x,y)\mathrm{d}x\mathrm{d}y$ existuje (a je roven 0), ale $\int_0^1 f\left(\frac{1}{2}, y\right) \mathrm{d}y$ neexistuje.
Mějme $D = [0,1] \times [0,1]$.
Mějme $\tau(t) = \begin{cases} 1 & \text{pro } t \in \left[0, \frac{1}{2}\right]; \\ 0 & \text{pro } t \in \left(\frac{1}{2}, 1\right], \end{cases}$
a

$$g(x,y) = \begin{cases} \tau(x) & \text{pokud } y = \frac{\ell}{2^i} \text{ pro všechna kladná celá } \ell, i; \\ 1 - \tau(x) & \text{jinak.} \end{cases}$$

$\iint_D g(x,y)\mathrm{d}x\mathrm{d}y$ neexistuje, ale $\int_0^1 g(x,y)\mathrm{d}x$ existuje a navíc je roven $\frac{1}{2}$, a tedy $\int_0^1 \left(\int_0^1 g(x,y)\mathrm{d}x\right) \mathrm{d}y$ existuje.

Část III
Strojová čísla a numerická matematika

10 Numerická matematika

10.1 Co to je?

Numerická matematika

Numerická matematika se věnuje matematickým metodám hledajícím přibližná řešení matematických úloh a jejich spolehlivosti.

Zahrnuje například metody pro…

1. řešení soustav lineárních rovnic,
2. řešení (obyčejných i parciálních) diferenciálních rovnic,
3. výpočet integrálů,
4. vyhodnocování funkčních hodnot,
5. odhadování chyb při výpočtech,
6. hledání lokálních a globálních extrémů (optimalizační úlohy),
7. výpočet vlastních čísel a vlastních vektorů,
8. faktorizace matic,
9. …

Typicky k řešení úloh využívá počítačů.

Z dějin neúspěchu…

- Chyba v raketě Patriot (28 mrtvých)

$$(0.1)_{10} = (0.0001100110011001100110011001100110011\ldots)_2$$

- Exploze rakety Ariane 5 konverze z 64-bitového čísla s plovoucí čárkou na 16-bitové celé se znaménkem. ($7 miliard na vývoj, raketa a náklad za půl miliardy.)

[http://ta.twi.tudelft.nl/users/vuik/wi211/disasters.html]

Neznamená to, že by metody nefungovaly: naopak v naprosté většině případů fungují dobře.

10.2 Chyby podle jejich původu

Kategorizace chyb

Při návrhu algoritmu pro hledání chyb budeme používat různé **aproximace**. Budeme se tedy dopouštět různých chyb, které lze rozdělit podle jejich původu:

1. chyba **modelu**: matematický model řešené úlohy je nějakým způsobem zjednodušený, např. je zanedbáno tření, nebo se používají průměrné místo aktuálních hodnot.
2. chyba **dat**: data často pocházejí z měření, která nemají absolutní přesnost (vesměs všechna měření fyzikálních veličin).
3. chyba **algoritmu**: nemusíme mít k dispozici algoritmus, který v konečném počtu kroků najde přesné řešení.
4. **zaokrouhlovací** chyba: při samotném výpočtu (v rámci algoritmu) dochází k chybám (např. při aritmetických operacích).

Začneme od zaokrouhlovacích chyb.

11 Počítačová aritmetika

11.1 Reprezentace s pohyblivou řádovou čárkou

Reprezentace s pohyblivou řádovou čárkou

K ukládání čísel v počítači se většinou používá binární báze, například

$$(6)_{10} = (110)_2, \quad (0.1)_{10} = (0.0001100110011001100110011001100110011...)_2.$$

Pro necelá čísla x se používá tzv. **vědecký zápis čísel**, v binární bázi vypadá takto

$$x = \pm q \cdot 2^e,$$

kde

- q je **signifikant**[2] (*significand*) mající pevný počet cifer / pevnou délku, těmto cifrám se také říká **platné cifry**.
- e je **exponent** mající pevný počet cifer / pevnou délku.

Standard IEEE-754 (1985)

Strojové číslo lze reprezentovat znaménkem s a celými kladnými čísly e a m. Označme $(m_2)_2 = m$. Standard IEEE-754 klade na e a m následující podmínky a omezuje tak množinu reprezentovatelných reálných čísel:

přesnost	délka m	d = délka e	parametr b
poloviční (binary16, *half precision*)	10	5	15
jednoduchá (binary32, *single precision*)	23	8	127
dvojitá (binary64, *double precision*)	52	11	1023
čtyřnásobná (binary128, *quadruple prec.*)	112	15	16383

Reprezentovaná hodnota x se pak určí následujícím způsobem:

- pokud $e = 2^d - 1$ a $m \neq 0$, tak $x =$ `NaN` (*Not a Number*)
- pokud $e = 2^d - 1$ a $m = 0$, tak $x = (-1)^s \cdot$ `Inf`
- pokud $0 < e < 2^d - 1$, pak $x = (-1)^s \cdot (1.m_2)_2 \cdot 2^{e-b}$ (tzv. **normalizovaná čísla**)
- pokud $e = 0$ a $m \neq 0$, pak $x = (-1)^s \cdot (0.m_2)_2 \cdot 2^{1-b}$ (tzv. **subnormální čísla**)
- pokud $e = 0$ a $m = 0$, pak $x = (-1)^s \cdot 0$

[2] Často též „mantisa“. Další čtení k této terminologii na Wikipedii.

Skrytá jednička

Všimněme si, že pro normalizované číslo $x = (-1)^s \cdot (1.m_2)_2 \cdot 2^{e-b}$ jsme vlastně uložili o 1 platnou cifru více, než kolik je délka m. (Tedy např. v jednoduché přesnosti uložíme 24 platných cifer.)

Této konvenci se říká různě: skrytá jednička, skrytý bit, vedoucí bit, implicitní bit.

Strojová čísla (1/3)

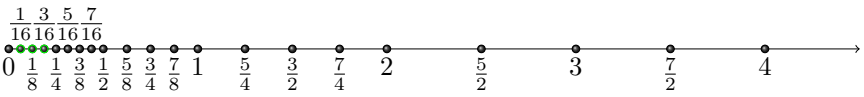
Reálná čísla, která lze reprezentovat popsaným způsobem, se nazývají **strojová čísla**.

Příklad: Vezměme dvoubitové m, exponent e s třemi bity (tj. $d = 3$) a $b = 3$.

Dostaneme následující množinu strojových čísel (vypisujeme jen nezáporná):

$$\left\{0, \frac{1}{16}, \frac{1}{8}, \frac{3}{16}, \frac{1}{4}, \frac{5}{16}, \frac{3}{8}, \frac{7}{16}, \frac{1}{2}, \frac{5}{8}, \frac{3}{4}, \frac{7}{8}, 1, \frac{5}{4}, \frac{3}{2}, \frac{7}{4}, 2, \frac{5}{2}, 3, \frac{7}{2}, 4, 5, 6, 7, 8, 10, 12, 14\right\}$$

Subnormální strojová čísla jsou zvýrazněna zeleně.



Množina všech strojových čísel s danou přesností (tj. specifikací m, e a b) nemá nic moc společného s reálnými čísly. Jde o konečnou podmnožinu racionálních čísel.

Strojová čísla (2/3)

Označme množinu strojových čísel symbolem $F \equiv F(|m|, |e|, b)$. Tímto zápisem zdůrazňujeme, že F závisí na zvoleném počtu bitů pro m, e a parametru b.

Množina F, jakožto konečná podmnožina $\mathbb{R}$ má samozřejmě největší a nejmenší prvek ($\min F$ a $\max F$). Dále jsou zajímavé následující hodnoty:

přesnost	max. č.	min. kladné norm. č.	min. kladné subnorm. č.
single	$(2 - 2^{-23})2^{127}$ $\approx 3.4 \cdot 10^{38}$	2^{-126} $\approx 1.2 \cdot 10^{-38}$	$2^{-126-23} = 2^{-149}$ $\approx 1.4 \cdot 10^{-45}$
double	$(2 - 2^{-52})2^{1023}$ $\approx 1.8 \cdot 10^{308}$	2^{-1022} $\approx 2.2 \cdot 10^{-308}$	$2^{-1022-52} = 2^{-1074}$ $\approx 4.9 \cdot 10^{-324}$

Skutečně, pro **jednoduchou** přesnost máme pro m 23 bitů, e 8 bitů a $b = 127$. Proto

$$\max F = (1.1\ldots1)_2 \cdot 2^{254-127} = \frac{1 - (1/2)^{24}}{1 - 1/2} \cdot 2^{127} = (2 - 2^{-23}) \cdot 2^{127}$$

$$\text{minimální kladné normalizované} = (1.0\ldots0)_2 \cdot 2^{1-127} = 2^{-126},$$

$$\text{minimální kladné subnormální} = (0.0\ldots1)_2 \cdot 2^{1-127} = 2^{-23-126} = 2^{-149}.$$

Strojová čísla (3/3)

F je charakterizováno pomocí **strojové přesnosti** ε_F (*machine epsilon*), což je vzdálenost čísla $1 = +1 \cdot 2^0$ od nejbližšího většího čísla v F, tj.

$$\varepsilon_F = (1.0\ldots01)_2 \cdot 2^0 - (1.0\ldots00)_2 \cdot 2^0.$$

Pro jednoduchou přesnost proto platí $\varepsilon_F = 2^{-23}$ a pro dvojitou $\varepsilon_F = 2^{-52}$.

Tvrzení 11.1. *Vzdálenost libovolného normalizovaného čísla $x \in F$ od jeho nejbližších sousedů z F je nejméně $\varepsilon_F \frac{|x|}{2}$ a nejvíce $\varepsilon_F |x|$.*

Reprezentace reálných čísel (1/3)

Nechť $\mathrm{fl} : \mathbb{R} \to F$ je zobrazení, které přiřadí každému $x \in \mathbb{R}$ „nejbližší" strojové číslo. (Zkratka fl je od slova „float".)

„Nejbližší" je určeno podle vybrané strategie zaokrouhlování (k nejbližšímu, k $\pm$ nekonečnu, náhodně)[3] či usekávání (zaokrouhlování směrem k nule)[4].

Při pokusu o reprezentaci čísel mimo rozsah dochází k **přetečení** (*overflow*) respektive **podtečení** (*underflow*).

Definice 11.2 Nechť číslo $\alpha \in F$ je přibližnou hodnotou čísla $a \in \mathbb{R}$.

- **Absolutní chybou** reprezentace a pomocí α rozumíme hodnotu $|\alpha - a|$.
- Pro $a \neq 0$ je **relativní chyba** reprezentace a pomocí α rovna

$$\frac{|\alpha - a|}{|a|}.$$

Reprezentace reálných čísel (2/3)

Uvažujme reálné číslo x takové, že lze psát

$$x = q \cdot 2^{\ell}, \qquad \text{kde } 1 \leq q < 2 \text{ a } -126 \leq \ell \leq 126.$$

Hrubě řečeno, x je v rozsahu normalizovaných čísel v jednoduché přesnosti.

Jaká je **chyba** vzniknuvší při zaokrouhlení na nejbližší strojové číslo?
Pro jednoduchost budeme *zaokrouhlovat směrem k nule*, tedy usekneme bity přesahující délku signifikandu (x je kladné). Nechť

$$x = (1.m_1 m_2 m_3 m_4 \ldots)_2 \cdot 2^{\ell},$$

pak

$$\mathrm{fl}(x) = (1.m_1 m_2 \ldots m_{23})_2 \cdot 2^{\ell},$$

Pro absolutní chybu platí $|x - \mathrm{fl}(x)| \leq 2^{-23+\ell}$ a pro relativní chybu

$$\frac{|x - \mathrm{fl}(x)|}{|x|} \leq \frac{2^{-23+\ell}}{q \cdot 2^{\ell}} \leq 2^{-23}.$$

Reprezentace reálných čísel (3/3)

Této mezi pro relativní chybu se říká **zaokrouhlovací jednotka** (*unit roundoff error*) a značí se $\mathbf{u} = 2^{-23}$.
Pozor, tato definice není ustálená a je zaměňována s výše uvedenou strojovou přesností (s různými variantami detailů).

Pokud bychom použili *zaokrouhlování směrem k nejbližšímu*[5], dostaneme $\mathbf{u} = 2^{-24}$.

Tvrzení 11.3. *Nechť $x \in \mathbb{R}$ leží mezi největším a nejmenším normalizovaným kladným číslem množiny F. Pak platí*

$$\mathrm{fl}(x) = x(1 + \delta), \qquad \textit{kde } |\delta| \leq \mathbf{u}.$$

Poznámka: δ zde závisí na x, tj. $\delta = \delta(x)$. Například pro strojové číslo x je $\delta = 0$.

[3] *Round to nearest, ties to even*; *Round to nearest, ties away from zero*; *Round towards +infinity*; *Round towards -infinity*; *Stochastic rounding*
[4] *Round towards zero*
[5] nezáleží, jak si poradíme s nerozhodnutelnými právě mezi 2 nejbližšími

11.2 Aritmetické operace

Aritmetické operace a chyba při jejich provádění

Se strojovými čísly můžeme provádět obvyklé základní číselné operace, např.: $+ : (x, y) \mapsto \mathrm{fl}(x + y)$.

Z předchozího tvrzení o zaokrouhlování ihned plyne:

Tvrzení 11.4. *Nechť $x, y \in F$ a $\odot$ značí operaci sčítání, odečítání, násobení nebo dělení. Pokud nedojde k přetečení nebo podtečení (zůstali jsme v intervalu normalizovaných čísel), tak platí*

$$\mathrm{fl}(x \odot y) = (x \odot y)(1 + \delta), \qquad \textit{kde } |\delta| \leq \mathbf{u}.$$

- Uvědomme si, že sčítáním/odečítáním/násobením/dělením dvou strojových čísel nemusíme nutně dostat opět strojové číslo! Obecně jde o reálné číslo, které je potřeba zaokrouhlit. (Tedy nemusí platit $\delta = 0$.)

- Model je příliš jednoduchý pro dnešní procesory, některé např. umí FMA: Fused Multiply Add, které počítá $(x, y, z) \mapsto x \pm yz$ s jedním zaokrouhlením.

Aritmetické operace – katastrofická ukázka 1/2

Mějme funkci $f : \mathbb{R}^2 \to \mathbb{R}$ definovanou takto:

$$f(x, y) = 333.75y^6 + x^2 \left(11x^2y^2 - y^6 - 121y^4 - 2\right) + 5.5y^8 + \frac{x}{2y}.$$

Vyhodnotíme $f(77617, 33096)$ v různých přesnostech

Aritmetické operace – katastrofická ukázka 2/2

přesnost	$f(77617, 33096)$
SageMath (přesnost 23 bitů)	1.17260
SageMath (přesnost 24 bitů)	$-6.33825 \cdot 10^{-29}$
SageMath (přesnost 53 bitů)	$-1.18059162071741 \cdot 10^{21}$
SageMath (přesnost 54 bitů)	$1.18059162071741 \cdot 10^{21}$
SageMath (přesnost 100 bitů)	$1.1726039400531786318588349045$
SageMath (přesnost 121 bitů)	$1.172603940053178631858834904520183 71$
SageMath (přesnost 122 bitů)	$-0.82739605994682136814116509547981 6292$

Přesný výsledek je $f(77617, 33096) = -\frac{54767}{66192} \approx -0.827396$.

[S. M. Rump: *Algorithms for verified inclusions – theory and practice*, 1988]

Ztráta platných cifer (1/3)

Došlo ke kumulaci chyb při provádění aritmetických operací.

Velké problémy může způsobit tzv. **krácení** (*cancellation*), které ovšem na první pohled nemusí být patrné.

Předvedeme si jej na ilustrativním příkladě. Představme si, že počítač počítá v desítkové soustavě a zaokrouhluje na 10 platných cifer.

Chceme vyhodnotit výraz $x - \sin(x)$ pro $x = \frac{1}{15}$.

$$\begin{aligned} x &\leftarrow 6.6666\ \ 66667 \ \cdot\ 10^{-2} \\ \sin(x) &\leftarrow 6.6617\ \ 29492 \ \cdot\ 10^{-2} \\ x - \sin(x) &\leftarrow 0.0049\ \ 37175 \ \cdot\ 10^{-2} \\ x - \sin(x) &\leftarrow 4.9371\ \ 75000 \ \cdot\ 10^{-5} \end{aligned}$$

Poslední 3 *nuly* nejsou *správné* platné cifry. Během výpočtu jsme o ně přišli.

Spočítejme relativní chybu našeho výpočtu...

Ztráta platných cifer (2/3)

Relativní chyba výpočtu je

$$\frac{\left|\frac{1}{15} - \sin\left(\frac{1}{15}\right) - \mathrm{fl}\left(\mathrm{fl}\left(\frac{1}{15}\right) - \sin\left(\mathrm{fl}\left(\frac{1}{15}\right)\right)\right)\right|}{\left|\frac{1}{15} - \sin\left(\frac{1}{15}\right)\right|} \approx 1.4 \cdot 10^{-7}.$$

To je hodně v porovnání se zaokrouhlovací jednotkou v této aritmetice

$$\frac{|x - \mathrm{fl}(x)|}{|x|} \leq 5 \cdot 10^{-10} = \mathbf{u},$$

kde x je v rozsahu normalizovaných čísel.

Tvrzení 11.5. *Nechť x a y jsou normalizovaná strojová čísla a platí $x > y > 0$. Pokud $2^{-p} \leq 1 - \frac{y}{x} \leq 2^{-q}$ pro nějaká kladná celá p a q, tak platí, že* nejvíce p *a* nejméně q *platných binárních bitů je ztraceno při provedení odečítání $x - y$.*

Ztráta platných cifer (3/3)

Krácení se lze vyhnout několika technikami:

- přeformulováním problému tak, aby nedocházelo k odečítání,
- použitím rozvojů funkcí do řad (např. do Taylorovy řady),
- použitím jiných rovností …
- (použitím přesné aritmetiky)

Další čtení

Základní potíže při práci s čísly s plovoucí čárkou I

Základní potíže při práci s čísly s plovoucí čárkou II

Implementace funkce sinus v libm

11.3 Závěr

Zaokrouhlovací chyby – shrnutí

Původ zaokrouhlovacích chyb:

- zaokrouhlovací chyby jednotlivých operací a jejich kumulace,
- krácení.

Několik poznámek k zaokrouhlovacím chybám:

- zvýšení přesnosti nemusí dát přesnější výsledek,
- krácení může být někdy výhodné – lze tak vyrušit zaokrouhlovací či jiné chyby,
- málo operací s malými čísly neznamená, že chyba bude malá.

Nebereme v úvahu hardware (např x87 vs SSE2, FMA...).

Zaokrouhlovací chyby – alternativní přístupy

Jeden z **problémů** strojových čísel (IEEE-754 apod.) spočívá v ignoraci vzniknuvší chyby při výpočtu.

Možné alternativy (stručně):

- Použít exaktní aritmetiku $\mathbb{Z}$, $\mathbb{Q}$ či $GF(p)$ (není vždy možné a vhodné).
- Intervalová aritmetika (místo jednoho strojového čísla pracujeme s dvěma reprezentujícími krajní body intervalu, jehož délka představuje neurčitost ve znalosti příslušného „čísla“). (IEEE 1788–2015)
- Unum.
- …

Řešené příklady: Integrály

Příklad 1

Základní cvičení 24.2
Jak přesně bude vypadat 32 bitů reprezentujících následující čísla (uvažujeme jednoduchou přesnost, pouze normalizovaná čísla a zaokrouhlování k nejbližšímu, nerozhodné směrem od nuly – první bit je znaménko, pak exponent a pak signifikant):

a) $-1/5$,

b) $2/3$,

c) součet těchto (reprezentovaných) čísel.

$$
\begin{aligned}
-\frac{1}{5} &= -1.1001\ 1001\ 1001\ 1001\ 1001\ 1001\ldots \cdot 2^{-3} \\
\mathrm{fl}\left(-\frac{1}{5}\right) &= -1.1001\ 1001\ 1001\ 1001\ 1001\ 101 \cdot 2^{-3} \\
\frac{2}{3} &= 1.01\ 01\ 01\ 01\ 01\ 01\ 01\ 01\ 01\ 01\ 01\ 01\ \ldots \cdot 2^{-1} \\
\mathrm{fl}\left(\frac{2}{3}\right) &= 1.01\ 01\ 01\ 01\ 01\ 01\ 01\ 01\ 01\ 01\ 01\ 1 \cdot 2^{-1}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{fl}\left(\frac{2}{3}\right) &= \quad 1.01\ 01\ 01\ 01\ 01\ 01\ 01\ 01\ 01\ 01\ 01\ 100 \cdot 2^{-1} \\
\mathrm{fl}\left(-\frac{1}{5}\right) &= -0.01\ 10\ 01\ 10\ 01\ 10\ 01\ 10\ 01\ 10\ 01\ 101 \cdot 2^{-1}
\end{aligned}
$$

12 Přímé a iterační metody obecně

Přímé metody

Přímá metoda počítá řešení nějakého problému v konečném počtu kroků tak, že v teoretické absolutní přesnosti dává (přesné) řešení.

Příklady z lineární algebry:

- Gaussova eliminační metoda (GEM),
- metoda hledání inverze matice pomocí GEM,
- …pomocí rozkladů matic (Choleského, LU, QR, …),
- …

Obecná myšlenka iteračních metod

Iterační metody hledají přibližná řešení matematických problémů tak, že konstruují posloupnost přibližných „řešení“:

$$x_0,\ x_1,\ x_2, \ldots$$

Každé další přibližné „řešení“ je odvozeno z předchozího:

$$x_k = T(x_{k-1}),$$

pro $k > 0$ a zatím blíže neurčené zobrazení T.

Zobrazení T je voleno tak, aby posloupnost $(x_k)_{k=0}^{\infty}$ **měla limitu**[6], **která je skutečným řešením** dané úlohy.

Poznámka: Pokud je T neměnné pro všechny iterace k, metoda se nazývá **stacionární**.

13 Vlastní čísla a vektory: připomenutí

Vlastní čísla a vektory

- Komplexní číslo λ nazýváme vlastním číslem matice $M \in \mathbb{C}^{n,n}$, právě když existuje nenulový vektor $\mathbf{u} \in \mathbb{C}^n$ splňující rovnici
$$M\mathbf{u} = \lambda\mathbf{u}.$$
Takovýto vektor $\mathbf{u}$ pak nazýváme vlastním vektorem matice M příslušejícím k vlastnímu číslu λ.

- Všechny vlastní vektory matice M příslušející vlastnímu číslu λ spolu s nulovým vektorem tvoří podprostor $\ker(M - \lambda E)$.

- Vlastní čísla matice M jsou právě kořeny charakteristického polynomu matice M, tj. polynomu
$$p_M(\lambda) := \det(M - \lambda E).$$
Každá takováto matice M má proto nejvýše n různých komplexních vlastních čísel.

[6] v nějaké normě

Vlastní čísla a vektory

- Hledat kořeny charakteristického polynomu „velkého“ stupně není snadné. Postup pro hledání vlastních čísel známý z lineární algebry není použitelný.

- Dokonce platí následující pozorování: umíme-li hledat vlastní čísla matice, pak umíme hledat kořeny polynomů. Skutečně, pro monický polynom

$$q(x) = x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0$$

jsou vlastní čísla matice (tzv. **matice společnice** (*companion matrix*) monického polynomu q)

$$C(q) := \begin{pmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{n-1} \end{pmatrix}$$

přesně kořeny polynomu q. Polynom q je totiž (případně až na multiplikativní znaménko) charakteristickým polynomem matice $C(q)$.

Diagonalizovatelnost matice

- Matice $M \in \mathbb{C}^{n,n}$ je diagonalizovatelná, právě když existují diagonální matice $D \in \mathbb{C}^{n,n}$ a regulární matice $P \in \mathbb{C}^{n,n}$ splňující

$$M = PDP^{-1}.$$

- **Připomenutí**: V předchozí přednášce jsme zkoumali mocniny matice M, platí $M^k = PD^kP^{-1}$. Vlastní čísla kontrolují asymptotické chování těchto mocnin pro $k \to +\infty$.

- **Poznámka**: Matice P obsahuje ve sloupcích vlastní vektory matice M (Lze ověřit. Tyto vlastní vektory tvoří bázi $\mathbb{C}^n$. Prvky na diagonále matice D jsou právě vlastní čísla matice M (včetně násobností).

Využití vlastních čísel

Vlastní čísla hrají důležitou roli v řadě aplikací, namátkou

- Klasifikace kuželoseček a kvadrik (geometrie).

- Kvantové počítání, kvantová mechanika, asymptotické chování různých systémů (fyzika).

- Analýza hlavních komponent: PCA neboli *principal component analysis* (big data).

- Rozpoznávání 2D i 3D objektů pomocí spektrálních metod (AI).

- Konkrétnější příklad: **PageRank** měří relativní důležitost WWW stránek zkoumáním odkazů mezi nimi.
  - Jeho výpočet spočívá ve výpočtu vlastního vektoru k dominantnímu číslu modifikované matice sousednosti grafu těchto odkazů.
  - K výpočtu **PageRank**u se používá se tzv. **mocninná metoda** (*power method*), kterou si probereme dále.

14 Norma – připomenutí

Norma – dpřipomenutí

Definice 14.1 — Norma (*norm*). **Norma** na vektorovém prostoru V (nad $\mathbb{R}$ nebo $\mathbb{C}$) je zobrazení $\|\cdot\| : V \to \mathbb{R}_0^+$ splňující:

1. $\|x\| = 0 \Rightarrow x = 0$,
2. $\|\alpha x\| = |\alpha| \cdot \|x\|$,
3. $\|x + y\| \leq \|x\| + \|y\|$ (trojúhelníková nerovnost),

pro všechna $x, y \in V$ a všechny skaláry α.

Na $\mathbb{R}^n$ ($\mathbb{C}^n$) je nejznámější pravděpodobně **eukleidovská** norma:

$$\|x\|_2 := \left(\sum_{i=1}^{n} |x_i|^2 \right)^{\frac{1}{2}},$$

kde $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ ($\mathbb{C}^n$). (Ověření trojúhelníkové nerovnosti není triviální!)

Norma – další příklady znovu

Dalšími často užívanými normami jsou:

$$\|x\|_\infty := \max\left\{ |x_i| \,\middle|\, i \in \{1, \ldots, n\} \right\} \qquad \text{(tzv. maximová norma)}$$

$$\|x\|_1 := \sum_{i=1}^{n} |x_i| \qquad \text{(tzv. součtová norma)}$$

Obecně, pro libovolné $p \geq 1$ je

$$\|x\|_p := \sqrt[p]{\sum_{i=1}^{n} |x_i|^p}$$

normou na $\mathbb{R}^n$ (resp. $\mathbb{C}^n$). (Ověření trojúhelníkové nerovnosti není triviální!)

Ekvivalence norem

Řekneme, že dvě normy $\|\cdot\|_a$ a $\|\cdot\|_b$ na prostoru V jsou *ekvivalentní*, pokud existují konstanty C_1 a C_2 takové, že

$$\forall x \in V, \quad C_1\|x\|_b \leq \|x\|_a \leq C_2\|x\|_b.$$

Věta 14.2 Nechť V je vektorový prostor konečné dimenze nad $\mathbb{R}$ nebo $\mathbb{C}$ a nechť $\|\cdot\|_a$ a $\|\cdot\|_b$ jsou dvě normy na V. Platí, že $\|\cdot\|_a$ a $\|\cdot\|_b$ jsou ekvivalentní.

15 Mocninná metoda

Mocninná metoda: úvod a předpoklady (1/3)

- Ve své základní variantě slouží mocninná metoda k nalezení v absolutní hodnotě největšího vlastního čísla (takovému se říká dominantní vlastní číslo) a příslušného vlastního vektoru.

- Mějme matici $M \in \mathbb{C}^{n,n}$. **Předpokládejme**, že je diagonalizovatelná, tj. existuje regulární matice $P \in \mathbb{C}^{n,n}$ splňující
$$M = PDP^{-1},$$
kde $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ a dále navíc **předpokládejme**, že můžeme její vlastní čísla označit takto:
$$|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|.$$

Poznámky:

- Z výše uvedeného plyne, že dominantní vlastní číslo λ_1 není degenerované (příslušný vlastní podprostor matice M má dimenzi 1).

- Předpoklad diagonalizovatelnosti lze odstranit, v zájmu jednoduchosti výkladu ho ale přijmeme.

Mocninná metoda: úvod a předpoklady (2/3)

Jak moc je předpoklad o vlastních číslech restriktivní?

Splňují jej velké třídy matic: kladné matice a obecněji primitivní matice (tzv. Perronova-Frobeniova věta).

Definice 15.1 Čtvercová matice M je **primitivní**, pokud má nezáporné (reálné) prvky a existuje kladné k takové, že M^k má kladné prvky.

Jak moc je předpoklad o diagonalizovatelnosti restriktivní?

Není, ten máme pouze pro zjednodušení. V obecnější variantě lze opět použít Jordanův normální tvar (bez změny principu).

Mocninná metoda: úvod a předpoklady (3/3)

- Hledáme vlastní vektor přidružený k vlastnímu číslu λ_1, tedy nenulový vektor $\mathbf{u}_1$ splňující
$$M\mathbf{u}_1 = \lambda_1 \mathbf{u}_1.$$

- Mocninná metoda je **iterativní metoda**. Zvolme $\mathbf{x}_0 \in \mathbb{C}^n$ a sestrojme posloupnost $(\mathbf{x}_k)_{k=0}^{\infty}$ zadanou rekurentně vztahem
$$\mathbf{x}_{k+1} = M\mathbf{x}_k, \quad k \in \mathbb{N}.$$
Ekvivalentně máme explicitní vyjádření
$$\mathbf{x}_k = M^k \mathbf{x}_0, \quad k \in \mathbb{N}.$$

- Tento vzorec je původem vžitého názvu *mocninná* metoda (též *power iteration*).

Mocninná metoda: vlastní vektor (1/3)

- Vektor $\mathbf{x}_0$ lze napsat jako lineární kombinaci vlastních vektorů matice M, tj.
$$\mathbf{x}_0 = P\boldsymbol{\alpha},$$
kde $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)^T \in \mathbb{C}^n$ je vektor obsahující koeficienty příslušné linární kombinace. **Předpokládejme**, že $\alpha_1 \neq 0$. (Při náhodné volbě vektoru $\mathbf{x}_0$ je rovnost $\alpha_1 = 0$ nepravděpodobná.)

- Pro naší posloupnost $(\mathbf{x}_k)_{k=0}^{\infty}$ potom platí

$$\mathbf{x}_k = M^k\mathbf{x}_0 = (PDP^{-1})^k\mathbf{x}_0 = PD^k\boldsymbol{\alpha} = P\,\mathrm{diag}(\lambda_1^k,\ldots,\lambda_n^k)\,\boldsymbol{\alpha} = \\ = \lambda_1^k P\,\mathrm{diag}\left(1,\left(\frac{\lambda_2}{\lambda_1}\right)^k,\ldots,\left(\frac{\lambda_n}{\lambda_1}\right)^k\right)\boldsymbol{\alpha},$$

  kde $D = \mathrm{diag}(\lambda_1,\ldots,\lambda_n)$.

Mocninná metoda: vlastní vektor (2/3)

- Nyní vektor $\mathbf{x}_k$ normalizujme tak, aby jeho první složka byla rovna 1 (vizte poznámku dále), tedy

$$\mathbf{y}_k := \frac{\mathbf{x}_k}{(\mathbf{x}_k)_1} = \frac{P\,\mathrm{diag}\left(1,\left(\frac{\lambda_2}{\lambda_1}\right)^k,\ldots,\left(\frac{\lambda_n}{\lambda_1}\right)^k\right)\boldsymbol{\alpha}}{\left(P\,\mathrm{diag}\left(1,\left(\frac{\lambda_2}{\lambda_1}\right)^k,\ldots,\left(\frac{\lambda_n}{\lambda_1}\right)^k\right)\boldsymbol{\alpha}\right)_1} \longrightarrow \\ \longrightarrow \frac{P\,\mathrm{diag}(1,0,\ldots,0)\boldsymbol{\alpha}}{(P\,\mathrm{diag}(1,0,\ldots,0)\boldsymbol{\alpha})_1} =: \mathbf{y}_\infty \in \ker(M-\lambda_1 E),$$

  když $k \to +\infty$. Skutečně, dle našeho předpokladu totiž

$$\lim_{k\to\infty}\left(\frac{\lambda_i}{\lambda_1}\right)^k = 0$$

  pro každé $i = 2,3,\ldots,n$ a

$$P\,\mathrm{diag}(1,0,\ldots,0)\boldsymbol{\alpha} = \alpha_1\mathbf{u}_1 \in \ker(M-\lambda_1 E).$$

Mocninná metoda: vlastní vektor (3/3)

Hned na tomto místě je vhodné učinit několik poznámek:

- Při normalizaci můžeme/musíme zvolit i **jinou složku než první**, která je nenulová. Vlastní vektor je nenulový a jedna z jeho složek proto nutně musí být nenulová.

- Kdybychom věděli, že $\lambda_1 > 0$, pak lze vektory $\mathbf{x}_k$ skutečně normalizovat a tím se také zbavit mocnin λ_1^k. Ve výpočtu výše bychom kladli (stručný zápis)

$$\mathbf{y}_k := \frac{\mathbf{x}_k}{\|\mathbf{x}_k\|} = \frac{\lambda_1^k(P\ldots)}{\|\lambda_1^k(P\ldots)\|} = \frac{\lambda_1^k(P\ldots)}{|\lambda_1^k|\,\|(P\ldots)\|} = \frac{\lambda_1^k(P\ldots)}{\lambda_1^k\,\|(P\ldots)\|} = \frac{(P\ldots)}{\|(P\ldots)\|}$$

  a posloupnost $(\mathbf{y}_k)_{k=1}^{+\infty}$ by opět byla konvergentní.

- Pro kladné a nezáporné ireducibilní matice platí $\lambda_1 > 0$ a složky $\mathbf{u}_1$ jsou také kladné (opět Perronova-Frobeniova věta).

Mocninná metoda: vlastní číslo

- Vlastní číslo, resp. jeho aproximaci, nyní snadno zjistíme následovně

$$\frac{\langle\mathbf{y}_k, M\mathbf{y}_k\rangle}{\|\mathbf{y}_k\|^2} \longrightarrow \frac{\langle\mathbf{y}_\infty, M\mathbf{y}_\infty\rangle}{\|\mathbf{y}_\infty\|^2} = \frac{\langle\mathbf{y}_\infty, \lambda_1\mathbf{y}_\infty\rangle}{\|\mathbf{y}_\infty\|^2} = \lambda_1,$$

  zde $\langle x, y\rangle$ značí standardní skalární součin[7] na $\mathbb{C}^n$ a $\|x\|$ je euklidovská norma.

- Alternativně bychom mohli vzít libovolné (nenulové na $\ker(M-\lambda_1 E)$) lineární zobrazení $\varphi : \mathbb{C}^n \to \mathbb{C}$ a počítat podíly

$$\frac{\varphi(M\mathbf{y}_k)}{\varphi(\mathbf{y}_k)} \longrightarrow \frac{\varphi(M\mathbf{y}_\infty)}{\varphi(\mathbf{y}_\infty)} = \frac{\varphi(\lambda_1\mathbf{y}_\infty)}{\varphi(\mathbf{y}_\infty)} = \lambda_1.$$

  Například $\varphi(\mathbf{v}) = (\mathbf{v})_i$ pro zvolené $i \in \{1,2,\ldots,n\}$.

[7] Tedy $\langle x, y\rangle = \sum \overline{x_i} y_i$.

Mocninná metoda: konvergence

- **Kritérium zastavení** ($\|\cdot\|$ je jistá norma na $\mathbb{C}^n$): máme-li poslední iteraci aproximace vlastního vektoru $\mathbf{y}_k$ a vlastního čísla $\lambda_1^{(k)}$, pak otestujeme reziduum

$$\|M\mathbf{y}_k - \lambda_1^{(k)}\mathbf{y}_k\| < \varepsilon.$$

- Pro rychlost konvergence posloupnosti platí (odhad explicitně neprovádíme)

$$\|\mathbf{y}_k - \mathbf{y}_\infty\| = \mathcal{O}\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right)$$

  a vlastních čísel platí

$$\left\|\lambda_1^{(k)} - \lambda_1\right\| = \mathcal{O}\left(\left|\frac{\lambda_2}{\lambda_1}\right|^{2k}\right).$$

- Obecné **pozorování**: Jsou-li vlastní čísla „blízko u sebe“, pak může být výpočet velmi pomalý.

Mocninná metoda: poznámky

- Dejme tomu, že jsme mocninnou metodou nalezli dominantní vlastní číslo λ_1 a jemu příslušející normalizovaný (Euklidovou normou) vlastní vektor $\mathbf{u}_1$. Jak hledat **další** vlastní čísla?

- Předpokládejme, že matice M je normální ($MM^\dagger = M^\dagger M$, dýka označuje transpozici a komplexní sdružení), pak má ortogonální vlastní vektory.

- Můžeme matici upravit následovně (tečka označuje maticové násobení[8]; pruh komplexní sdružení složek)

$$M' := M - \lambda_1 \mathbf{u}_1 \cdot \mathbf{u}_1^\dagger.$$

- Matice M' má nyní vektor $\mathbf{u}_1$ také jako vlastní vektor, ale s vlastním číslem 0! Skutečně,

$$M'\mathbf{u}_1 = M\mathbf{u}_1 - \lambda_1 \mathbf{u}_1 \cdot \|\mathbf{u}_1\|_2^2 = \lambda_1\mathbf{u}_1 - \lambda_1\mathbf{u}_1 = 0.$$

- Nyní aplikujeme mocninnou metodu (jsou-li splněny její předpoklady) na M' a získáme **druhé** v absolutní hodnotě největší vlastní číslo matice M.

Mocninná metoda: ukázka (1/2)

Uvažme matici

$$M = \begin{pmatrix} 36408 + 16769i & -5412 - 2481i & 107256 + 49397i & -492 - 214i \\ -10656 - 5164i & 1584 + 762i & -31392 - 15210i & 144 + 66i \\ -12876 - 5954i & 1914 + 881i & -37932 - 17539i & 174 + 76i \\ 4329 - 262i & -643 + 39i & 12753 - 771i & -58 + 6i \end{pmatrix}$$

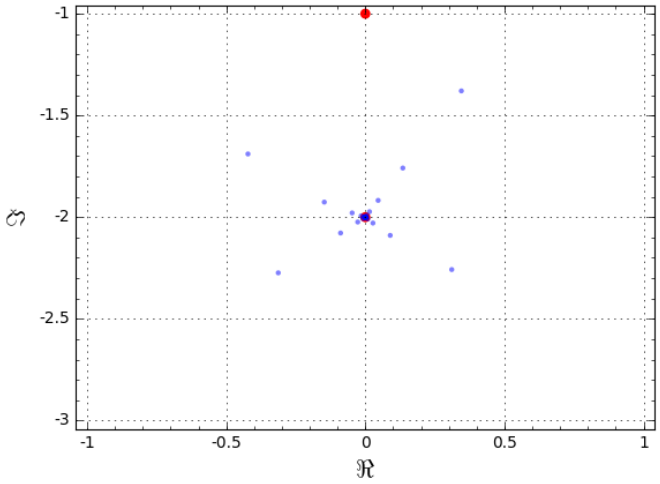
Je zkonstruována tak, že její vlastní čísla jsou $-2i$, $-i$, $3i/2$, $3/2$. Požadujeme přesnost (mez pro reziduum)

$\varepsilon = 10^{-6}$. Aproximace λ_1 z posledních 7 iterací:

[8] a vektor je stále sloupec

$$0.000047758815096087 2 - 1.99991424541241i$$
$$-0.0000479821875446196 - 1.99998019901599i$$
$$-0.0000272650944159076 - 2.00002375338328i$$
$$0.0000271520045767515 - 2.00002973125038i$$
$$0.0000154506695115737 - 1.99997272532314i$$
$$-0.0000152424622193764 - 1.99999349337182i$$

Mocninná metoda: ukázka (2/2)



16 QR algoritmus

QR faktorizace a QR algoritmus (1/2)

- Mocninná metoda se nehodí na hledání všech vlastních čísel matice M.
- Navíc maticové násobení je obecně náročně a vyplatí se až pro velké řídké matice.

- Existují další algoritmy založené na sérii podobnostních transformací: cílem je sestrojit posloupnost matic $(M_k)_{k=0}^{\infty}$, $M_0 = M$ tak, že
$$M_k = P_k M_{k-1} P_k^{-1}, \quad k \in \mathbb{N},$$
kde každá P_k je regulární, $M_k \to M_\infty$ a pro limitní matici umíme snadno spočítat vlastní čísla (například je horní trojúhelníková).

QR faktorizace a QR algoritmus (2/2)

- **QR faktorizace** matice spočívá ve vyjádření reálné (resp. komplexní) matice M ve tvaru součinu
$$M = Q \cdot R,$$
kde Q je ortogonální (resp. unitární) a R je horní trojúhelníková.

- Existuje několik algoritmů počítajících tuto faktorizaci (Gram-Schmidt, Hauseholderova transformace, Givensova rotace).

- **QR algoritmus** tuto faktorizaci provádí v každém kroku. Zhruba řečeno, pro M_k spočteme její QR faktorizaci
$$M_k = Q_k R_k$$
a položíme
$$M_{k+1} := R_k Q_k = Q_k^{-1} Q_k R_k Q_k = Q_k^{-1} M_k Q_k$$
Iterace začíná s $M_0 = M$. Všechny matice M_k jsou podobné naší M a mají tedy stejná vlastní čísla. Za jistých předpokladů M_k konverguje k trojúhelníkové matici.

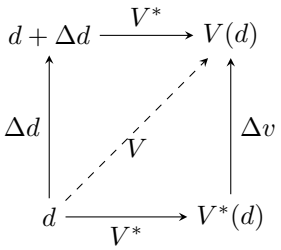
17 Podmíněnost úlohy a stabilita algoritmů

Dopředná a zpětná chyba

Nechť V je nějaký numerický algoritmus, jehož teoretický (přesný) výstup označíme $V^*(d)$, kde d jsou vstupní data.

Výsledek výpočtu v konečné (strojové) aritmetice označíme $V(d)$.
Označme $\Delta v = V^*(d) - V(d)$. Tato hodnota je tzv. **dopředná/přímá chyba** (*forward error*). Je to odchylka spočítaného řešení od přesného řešení.

Nejmenší (v normě) číslo Δd takové, že $V^*(d+\Delta d) = V(d)$ je **zpětná chyba** (*backward error*). Jedná se promítnutí chyby algoritmu V do jeho vstupu – jaký problém byl ve skutečnosti vyřešen.



Pokud je pro všechny vstupy d zpětná chyba relativně malá, řekneme, že algoritmus je **zpětně stabilní** (*backward stable*). „Malá“ závisí na kontextu.

Podmíněnost úlohy

Podmíněnost úlohy vyjadřuje závislost změny výstupu na změně vstupních dat - jejich malé perturbaci δd.

Relativní číslo podmíněnosti (*relative condition number*) úlohy je

$$C_r = \lim_{\epsilon \to 0^+} \sup_{\substack{d+\delta d \in D \\ \|\delta d\| \leq \epsilon}} \frac{\frac{\|V^*(d+\delta d) - V^*(d)\|}{\|V^*(d)\|}}{\frac{\|\delta d\|}{\|d\|}},$$

kde D je zkoumaný definiční obor V potažmo V^*.

Je-li $C_r \approx 1$, řekneme, že úloha je **dobře podmíněná** (*well-conditioned*).
Je-li velké, řekneme, že úloha je **špatně podmíněná** (*ill-conditioned*).

18 Soustavy lineárních rovnic

18.1 Značení

Soustavy lineárních rovnic

V této přednášce se budeme soustředit na známý problém z lineární algebry:

Chceme řešit soustavu $n \in \mathbb{N}$ lineárních rovnic pro n neznámých. Zapíšeme ji v maticovém tvaru

$$Ax = b,$$

kde $A \in \mathbb{R}^{n,n}$ je regulární **matice soustavy**, $b \in \mathbb{R}^{n,1}$ je **vektor pravých stran** a $x \in \mathbb{R}^{n,1}$ je hledané řešení.

Řešení takovéhoto problému je velmi často dílčím úkolem v nějaké větší úloze. V lineární algebře tuto úlohu řešíme pomocí Gaussovy eliminace (GEM).

GEM a numerické chyby

- Typickým problémem přímých metod je to, že vznikne-li v jednom kroku numerická chyba, tak se projevuje i při dalších výpočtech: přímé metody obecně nejsou „samoopravující se“, ale spíše nahrávají ke kumulování chyb.

- Pro některé úlohy se drobné chyby během výpočtů projeví pouze jako drobná odchylka ve výsledku, ovšem pro některé mohou znamenat řádovou změnu.

- Při řešení soustavy lineárních rovnic může nastat obojí. Čím to je?

Soustavy lineárních rovnic: příklad (1/2)

Uvažujme dvě soustavy dvou rovnic o dvou neznámých:

$$\begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{pmatrix} x = \begin{pmatrix} 3/2 \\ 1 \end{pmatrix} \quad \text{a} \quad \begin{pmatrix} 1 & 1/5 \\ 1/5 & -1 \end{pmatrix} x = \begin{pmatrix} 3/2 \\ 1 \end{pmatrix}.$$

Řešením těchto rovnic je

$$x = (0, 3)^T \text{ resp. } x = (85/52, -35/52)^T \approx (1.6346, -0.67308)^T.$$

Zkusme simulovat chybu na vstupu či chybu vzniklou během výpočtu záměnou 1 na pravé straně rovnice za 5/6:

$$\begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{pmatrix} x = \begin{pmatrix} 3/2 \\ 5/6 \end{pmatrix} \quad \text{a} \quad \begin{pmatrix} 1 & 1/5 \\ 1/5 & -1 \end{pmatrix} x = \begin{pmatrix} 3/2 \\ 5/6 \end{pmatrix}.$$

Řešení se změní na

$$x = (1, 1)^T \text{resp. } x = (125/78, -20/39)^T \approx (1.6026, -0.51282)^T.$$

Soustavy lineárních rovnic: příklad (1/2)

Pravou stranu jsme změnili o

$$\begin{pmatrix} 3/2 \\ 1 \end{pmatrix} - \begin{pmatrix} 3/2 \\ 5/6 \end{pmatrix} = \begin{pmatrix} 0 \\ 1/6 \end{pmatrix},$$

vektor euklidovské délky $\frac{1}{6}$ (norma relativní chyby je 0.09).

Změna v řešení **první rovnice** byla

$$\begin{pmatrix} 0 \\ 3 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

(norma relativní chyby je 0.75) a **druhé**

$$\begin{pmatrix} 85/52 \\ -35/52 \end{pmatrix} - \begin{pmatrix} 125/78 \\ -20/39 \end{pmatrix} = \begin{pmatrix} 5/156 \\ -25/156 \end{pmatrix}$$

(norma relativní chyby je 0.09).

V prvním případě je relativní chyba řádově větší než relativní chyba pravé strany. U druhé rovnice jsou relativní chyby řádově stejné.

18.2 Maticová norma

Maticová norma

Pro nějakou vektorovou normu $\|\cdot\|$ na $\mathbb{R}^n \equiv \mathbb{R}^{n,1}$ definujeme **přidruženou maticovou normu** matice $A \in \mathbb{R}^{n,n}$ následujícím způsobem

$$\|A\| := \sup\big\{\|Ax\| \colon x \in \mathbb{R}^{n,1} \text{ a } \|x\| = 1\big\}.$$

Připomenutí: **supremum** neprázdné omezené množiny $M \subset \mathbb{R}$ je číslo

$$\sup M = \min\{y \in \mathbb{R} \colon \forall x \in M : \, x \leq y\}.$$

Takto definované zobrazení je skutečně normou (rozmyslete!) a platí pro ni $\forall A, B \in \mathbb{R}^{n,n}, \forall x \in \mathbb{R}^n$:

- $\|E\| = 1$ (zde E je jednotková matice),
- $\|Ax\| \leq \|A\| \cdot \|x\|$ (konzistence normy),
- $\|AB\| \leq \|A\| \cdot \|B\|$ (submultiplikativita).

(Obecná maticová norma je někdy definová tak, aby splňovala submultiplikativitu.)

Maticová norma – příklady

Jaké maticové normy jsou přidružené dříve uvedeným normám $\|\cdot\|_1, \|\cdot\|_2$ a $\|\cdot\|_\infty$?

- Pro normu $\|\cdot\|_1$ dostáváme:

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^{n} |a_{i,j}|, \quad \text{tedy maximum součtu absolutních hodnot ve sloupci.}$$

- Pro normu $\|\cdot\|_\infty$ dostáváme:

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{i,j}|, \quad \text{tedy maximum součtu absolutních hodnot v řádku.}$$

- Pro normu $\|\cdot\|_2$ dostáváme:

$$\|A\|_2 = \text{ odmocnina z největšího vlastního čísla matice } A^T A,$$

kde $A = (a_{i,j})_{i,j=1}^{n}$ je matice z $\mathbb{R}^{n,n}$.

18.3 Podmíněnost úlohy

Podmíněnost úlohy (1/3)

Uvažujme soustavu rovnic $Ax = b \neq 0$ s regulární maticí A. Budeme zkoumat, co se stane, pokud pravou stranu b lehce změníme o *perturbaci* δb. Změnu v řešení $x = A^{-1}b$ pak označíme δx, platí tedy

$$Ax = b \quad \text{a} \quad A(x + \delta x) = Ax + A\delta x = b + \delta b.$$

Tudíž $A\delta x = \delta b$.

Platí $\|b\| = \|Ax\| \leq \|A\| \cdot \|x\|$, z čehož plyne $\frac{1}{\|x\|} \leq \frac{\|A\|}{\|b\|}$.

Dále $\|\delta x\| = \|A^{-1}\delta b\| \leq \|A^{-1}\| \cdot \|\delta b\|$.

Nakonec dostaneme

$$\frac{\|\delta x\|}{\|x\|} \leq \|A\| \cdot \|A^{-1}\| \cdot \frac{\|\delta b\|}{\|b\|}.$$

Podmíněnost úlohy (2/3)

Odvodili jsme nerovnost

$$\frac{\|\delta x\|}{\|x\|} \leq \|A\| \cdot \|A^{-1}\| \cdot \frac{\|\delta b\|}{\|b\|}.$$

Číslo $\kappa(A) := \|A\| \cdot \|A^{-1}\|$ se nazývá **číslo podmíněnosti** matice A.

Nerovnost výše můžeme číst takto: relativní chyba v řešení x soustavy $Ax = b$ je menší než relativní chyba pravé strany b vynásobená číslem $\kappa(A)$.

Čím je $\kappa(A)$ větší, tím je úloha hůře podmíněná a při jejím numerickém řešení musíme být obezřetní, zejména je-li velká chyba při napočítávání Ax nebo b (které může být řešením nějaké jiné úlohy či výsledek měření).

Hodnota čísla podmíněnosti závisí na zvolené normě: závěry jsou tedy vždy vzhledem k této použité normě.

Podmíněnost úlohy (3/3)

Vraťme se ke dvou maticím z ukázkové úlohy uvedené dříve:

$$A = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{pmatrix} \quad \text{a} \quad B = \begin{pmatrix} 1 & 1/5 \\ 1/5 & -1 \end{pmatrix}.$$

Jejich inverze jsou

$$A^{-1} = \begin{pmatrix} 4 & -6 \\ -6 & 12 \end{pmatrix} \quad \text{resp.} \quad B^{-1} \approx \begin{pmatrix} 0.961538 & 0.192308 \\ 0.192308 & -0.961538 \end{pmatrix}.$$

Pro výpočet podmíněnosti $\kappa(A) = \|A\|\|A^{-1}\|$ použijeme např. normu $\|A\|_\infty$:

$$\kappa(A) = \|A\|_\infty \|A^{-1}\|_\infty = \frac{3}{2} 18 = 27 \quad \text{a} \quad \kappa(B) = \frac{18}{13} \approx 1.3846056.$$

Úloha s maticí A je tedy výrazně **hůře podmíněná**, než ta s maticí B, což koresponduje s našimi dřívějšími výpočty.

18.4 Popis iterační metody

Schéma základní iterační metody pro řešení $Ax = b$

Nejdříve si uvedeme obecný popis metody, která „spadne z nebe“, a následně si vysvětlíme, kdy a jak funguje:

- Naším cílem je algoritmus, který konstruuje **posloupnost vektorů** $x_0, x_1, x_2, \ldots$, která se „blíží“ k přesnému řešení rovnice $Ax = b$.
- Startovací vektor x_0 zvolíme náhodně, o řešení nemáme žádnou informaci, takže chceme, aby algoritmus fungoval pro libovolnou volbu x_0.
- Zvolíme si **regulární matici** Q (různé volby této matice pak povedou na různé metody).
- Členy posloupnosti $x_0, x_1, x_2, \ldots$ budeme napočítávat podle následujícího předpisu:

  $$Qx_k = (Q - A)x_{k-1} + b, \quad \text{pro všechna } k > 0.$$

  resp.

  $$x_k := Q^{-1}\big((Q - A)x_{k-1} + b\big), \quad \text{pro všechna } k > 0.$$

Základní iterační metody pro řešení $Ax = b$: myšlenka

Kdyby byla posloupnost $(x_k)_{k=0}^{\infty}$ konvergentní s limitou x^*, potom je toto x^* hledané řešení! Pošleme-li totiž v rovnici

$$Qx_k = (Q - A)x_{k-1} + b,$$

k do nekonečna, dostaneme

$$Qx^* = (Q - A)x^* + b,$$

a tedy $Ax^* = b$. **Poznámka:** Skutečně. Ze základních vlastností normy plyne implikace: pokud $\lim_{k\to\infty} x_k = x^*$, potom $\lim_{k\to\infty} Mx_k = Mx^*$. **Myšlenka**: budeme volit Q tak, aby výše definovaná posloupnost $(x_k)_{k=0}^{\infty}$ konvergovala k nějakému x^*.

18.5 Konvergence

Konvergence – volba Q

Rovnost $x_k = Q^{-1}\big((Q - A)x_{k-1} + b\big)$ dosadíme do

$$\begin{aligned} x_k - x &= Q^{-1}\big((Q - A)x_{k-1} + b\big) - x \\ &= (E - Q^{-1}A)x_{k-1} - x + Q^{-1}b \\ &= (E - Q^{-1}A)x_{k-1} - (E - Q^{-1}A)x \\ &= (E - Q^{-1}A)(x_{k-1} - x), \end{aligned}$$

kde x je vektor splňující $Ax = b$ a E jednotková matice.

Označme $W := E - Q^{-1}A$. Dále označme **vektor chyby** $e_k := x_k - x$, pak platí

$$e_k = We_{k-1} = W^2e_{k-2} = \cdots = W^ke_0.$$

Naším cílem je, aby se e_k pro rostoucí k zmenšovalo a blížilo se k nule, vágně řečeno platí: e_k bude „menší“ než e_{k-1} pokud bude W „malé“.

Tj. potřebujeme W pro které $\lim_{k\to\infty} W^k = 0$.

Konvergence – Spektrální poloměr

Spektrální poloměr matice M je číslo $\rho(M) \geq 0$ definované jako absolutní hodnota největšího (v absolutní hodnotě) vlastního čísla:

$$\rho(M) := \max\{|\lambda| \colon \lambda \text{ je vlastním číslem } M\}.$$

Věta 18.1 Nechť $M \in \mathbb{C}^{n,n}$. Potom platí

$$\lim_{k\to+\infty} M^k = 0 \quad \Leftrightarrow \quad \rho(M) < 1.$$

!

Tedy v našem případě máme zajištěnou konvergenci iterační metody **právě tehdy, když**

$$\rho(W) < 1,$$

neboli všechna vlastní čísla matice $W = E - Q^{-1}A$ jsou v absolutní hodnotě menší než 1.

Důkaz věty (1/3)

Ukážeme si implikaci

$$\rho(M) < 1 \;\Rightarrow\; \lim_{k\to\infty} M^k = 0$$

pro speciální případ.

Předpokládejme, že matice M je diagonalizovatelná, neboli že existuje regulární matice P taková, že $M = PDP^{-1}$, kde

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

a $\lambda_1, \ldots, \lambda_n$ jsou vlastní čísla M.

Platí $M^k = PDP^{-1}PD^{k-1}P^{-1} = \cdots = PD^kP^{-1}$.

Důkaz věty (2/3)

Zřejmě platí

$$D^k = \begin{pmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{pmatrix}.$$

Protože $\rho(M) < 1$ jistě pro všechna $i = 1, \ldots, n$ platí $|\lambda_i| < 1$. Tudíž $\lim_{k\to\infty} D^k = 0$.

Celkem tedy

$$\lim_{k\to\infty} M^k = P\Big(\lim_{k\to+\infty} D^k\Big)P^{-1} = 0.$$

Pro nediagonalizovatelnou matici M se postupuje velice podobně, jen se použije Jordanův normální tvar matice.

Důkaz věty (3/3)

Dokažme druhou implikaci

$$\lim_{k\to\infty} M^k = 0 \;\Rightarrow\; \rho(M) < 1$$

sporem.

Mějme vlastní číslo λ matice M, $|\lambda| \geq 1$, s vlastním vektorem $v \neq 0$.

Potom

$$0 = \lim_{k\to\infty} \|M^k v\| = \lim_{k\to\infty} |\lambda|^k \cdot \|v\| \neq 0.$$

Což je spor.

Rychlost konvergence

Jak rychle se vektor chyby e_k blíží k nule?

Máme

$$e_k = W^k e_0.$$

Odhadneme normu (vizte vlastnosti maticové normy!)

$$\|e_k\| = \|W^k e_0\| \leq \|W^k\| \cdot \|e_0\| \leq \|W\|^k \cdot \|e_0\|.$$

Odhad vpravo bude ostře klesající pokud $\|W\| < 1$.

Pro tento odhad tedy záleží na volbě normy. Jak jsme viděli pro symetrické matice je chování W^k ovlivněno hodnotou $\rho(W)$. Obecně platí jen $\rho(M) \leq \|M\|$ (pro všechny přidružené maticové normy). Jelikož pro symetrickou W platí $\|W\|_2 = \rho(W)$, je volba normy $\|\cdot\|_2$ pro tento účel ta nejlepší, ačkoliv v praxi se jedná o výpočetně náročnou normu a spíše se nepoužívá. Pro nesymetrickou W je situace obdobná, jen je nutné se opřít o pojem singulární hodnoty. Obecně lze učinit závěr, že čím menší hodnota $\|W\|$ (a menší než 1), tím rychlejší konvergenci můžeme očekávat.

Kdy iterování ukončit? (1/2)

Iterační metodu ukončíme v kroku k, dosáhne-li x_k požadované přesnosti.
Tady vzniká **chyba algoritmu**.

Požadovaná přesnost a podmínka na ni většinou plyne z úlohy.

Pro případ $\|W\| < 1$ máme zajištěno, že posloupnost $(\|e_k\|)$ je ostře klesající a iterace lze zastavit, když nastane

$$\|e_k\| = \|x_k - x\| < \epsilon,$$

kde konstanta ϵ je uživatelem zadaný parametr. To je samozřejmě nepraktické, protože nemáme přesné řešení x.

Napočítáme v kroku k tzv. **reziduum** $Ax_k - b$ a tzv. **kritérium konvergence** bude

$$\|Ax_k - b\| < \epsilon.$$

Kdy iterování ukončit? (2/2)

Někdy se místo počítání rezidua volí výpočetně méně náročné kritérium

$$\|x_{k+1} - x_k\| < \epsilon.$$

Platí totiž

$$\begin{aligned}\|e_k\| = \|x_k - x\| &= \|x_k - x_{k+1} + x_{k+1} - x\| \\ &\leq \|x_k - x_{k+1}\| + \|\underbrace{x_{k+1} - x}_{=e_{k+1}}\| \\ &< \epsilon + \|W\| \cdot \|e_k\|,\end{aligned}$$

kde za předpokladu $\|W\| < 1$ z poslední nerovnosti plyne

$$\|e_k\| < \frac{\epsilon}{1 - \|W\|}.$$

Tedy kritérium lze výhodně použít je-li $\|W\|$ menší než 1, ale nepříliš blízko 1.

Poznámka: Výpočty v konečné přesnosti

Všechny dosud uvedené úvahy byly v teoretické (absolutní) přesnosti.
Při počítání v nepřesné aritmetice samozřejmě nemusí metoda konvergovat k přesnému řešení, i když máme $\|W\| < 1$.

Pokud ovšem v nepřesné aritmetice posloupnost konverguje, pak mají iterační metody tu výhodu, že si „nepamatují“ chyby z předchozích iterací – v každém kroku se z aproximace vyrobí „lepší“ řešení úlohy a začíná se znovu.

V nepřesné aritmetice metoda nemusí konvergovat i v případě, kdy úloha není špatně podmíněná.

V praxi je tedy ještě dalším parametrem tohoto algoritmu **maximální počet iterací**. Při jeho překročení metoda selhala – po daném maximálním počtu iterací nebylo nalezeno přibližné řešení, které by splnilo podmínku konvergence dané konstantou ϵ.

Poznámka 2: iterační zpřesnění

Základní metodu lze dále různě vylepšovat: například použitím tzv. iteračního zpřesnění (*iterative refinement*):

1. V každém kroku napočteme reziduum: $r_k = Ax_k - b$.
2. Vyřešíme systém $Ay_k = r_k$ (přímou metodou).
3. Vypočtené řešení y_k použijeme k vylepšení x_k:
$$x'_k = x_k + y_k.$$

18.6 Konkrétní algoritmy

Konkrétní volby Q

Vraťme se zpět k základnímu algoritmu $x_k := Q^{-1}\big((Q-A)x_{k-1}+b\big)$.

Označme $a_{i,j}$ prvky matice A a položme

$$L := \begin{pmatrix} 0 & 0 & \cdots & 0 \\ a_{2,1} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n-1} & 0 \end{pmatrix} \quad \text{a} \quad D := \begin{pmatrix} a_{1,1} & 0 & \cdots & 0 \\ 0 & a_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{n,n} \end{pmatrix}.$$

a $U := A - L - D$, tj. platí

$$A = L + D + U.$$

Zmíníme následující volby Q:

- Richardsonova metoda $Q = E$ (E je jednotková matice),
- Jacobiho metoda $Q = D$,
- Gaussova-Seidlova metoda $Q = D + L$ (obecně superrelaxační (*successive overrelaxation*) metoda / SOR metoda $Q = \frac{1}{\omega}D + L$).

Richardsonova metoda

Richardsonova metoda odpovídá volbě $Q = E$.

Tedy iterace jsou jednoduše dány

$$x_k = Q^{-1}\big((Q-A)x_{k-1}+b\big) = (E-A)x_{k-1} + b.$$

Konvergenci kontroluje matice $W = E - Q^{-1}A = E - A$.

Konvergenci proto máme zajištěnou pro velmi úzkou třídu matic: A musí být blízko E tak, aby například platilo

$$\|E - A\| < 1.$$

Jacobiho metoda

Jacobiho metoda odpovídá volbě $Q = D$.

Iterace se napočítávají takto

$$x_k = Q^{-1}\big((Q-A)x_{k-1}+b\big) = D^{-1}(-L-U)x_{k-1} + D^{-1}b.$$

Konvergence je kontrolována maticí $W = E - Q^{-1}A = E - D^{-1}A$. Dále máme následující postačující podmínku.

Tvrzení 18.2. *Pokud je matice A ostře diagonálně dominantní, pak Jacobiho metoda konverguje pro všechny volby x_0.*

Poznámka: Matice je **ostře diagonálně dominantní** tehdy, pokud pro každý její řádek platí, že součet absolutních hodnot prvků vyjma diagonálního je menší než absolutní hodnota diagonálního prvku. Důkaz vynecháváme.

SOR metoda

SOR metoda odpovídá volbě $Q = \frac{1}{\omega}D + L$, kde $\omega \in \mathbb{R} \setminus \{0\}$.

Iterace se napočítávají takto

$$\left(\frac{1}{\omega}D + L\right) x_k = \left(\frac{1}{\omega}D + L - A\right) x_{k-1} + b = \left(\left(\frac{1}{\omega} - 1\right) D - U\right) x_{k-1} + b.$$

Tvrzení 18.3. *SOR metoda konverguje pokud $0 < \omega < 2$ a A je symetrická a pozitivně definitní s kladnými prvky na diagonále.*

Parametr ω se používá k urychlení konvergence.

Závěr: Algoritmus

Vstup: matice A, Q, vektor b, požadovaná přesnost ϵ, maximální počet iterací K

1. zvol náhodně x_0
2. pro k od 1 do K prováděj
   (a) $x_k = Q^{-1}(Q - A)x_{k-1} + Q^{-1}b$
   (b) pokud $\|Ax_k - b\| < \epsilon$, vrať x_k (nebo obecně je-li splněno jiné kritérium konvergence)
3. vrať „řešení nebylo nalezeno po K iteracích“

18.7 Ukázka

Ukázka – Jacobiho algoritmus (1/2)

Máme matici $A = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}$.

$$\|E - D^{-1}A\| = \frac{1}{2}.$$

Použijeme Jacobiho algoritmu pro výpočet řešení pro pravou stranu rovnou $b = (3, 5)^T$. Přesné řešení je $(1, 1)^T$.

Kritérium konvergence použijeme $\varepsilon := \|Ax_k - b\| < 10^{-2}$.

k	x_k	$\|Ax_k - b\|$
0	$(0.5, 1.5)$	1.58113883008
1	$(0.75, 1.125)$	0.450693909433
2	$(0.9375, 1.0625)$	0.197642353761
3	$(0.96875, 1.015625)$	0.0563367386791
4	$(0.9921875, 1.0078125)$	0.0247052942201
5	$(0.99609375, 1.001953125)$	0.00704209233489

Ukázka – Jacobiho algoritmus (2/2)

…stejná úloha, jen začneme vektorem x_0, který je dále od přesného řešení.

k	x_k	$\|Ax_k - b\|$
0	$(-10, 10)$	28.1780056072
1	$(-3.5, 3.75)$	9.01734439844
2	$(-0.375, 2.125)$	3.5222507009
3	$(0.4375, 1.34375)$	1.1271680498
4	$(0.828125, 1.140625)$	0.440281337613
5	$(0.9296875, 1.04296875)$	0.140896006226
6	$(0.978515625, 1.017578125)$	0.0550351672016
7	$(0.9912109375, 1.00537109375)$	0.0176120007782
8	$(0.997314453125, 1.002197265625)$	0.0068793959002

Odhady zaokrouhlovacích chyb - ukázky

18.8 Pomocná tvrzení pro vyhodnocení zaokrouhlovacích chyb

Připomenutí

Nechť $x, y \in F$ a $\odot$ značí operaci sčítání, odečítání, násobení nebo dělení. Pokud nedojde k přetečení nebo podtečení (zůstali jsme v intervalu normalizovaných čísel), tak platí

$$\mathrm{fl}(x \odot y) = (x \odot y)(1+\delta), \qquad \text{kde } |\delta| \le \mathbf{u}.$$

Mez $\mathbf{u}$ je **zaokrouhlovací jednotka**.

Načítání chyb

Lemma 18.4. *Pokud $|\delta_i| \le \mathbf{u}$ a $|\rho_i| = 1$ pro všechna $i \in \{1, \dots, n\}$, $n\mathbf{u} < 1$, tak platí*

$$\prod_{i=1}^{n} (1+\delta_i)^{\rho_i} = 1 + \Theta_n,$$

kde $|\Theta_n| \le \dfrac{n\mathbf{u}}{1-n\mathbf{u}}$.

Důkaz. Dokážeme indukcí.
$n = 1$ a $\rho_1 = 1$, pak $|\Theta_1| = |\delta_1| \le \mathbf{u} < \dfrac{\mathbf{u}}{1-\mathbf{u}}$.
$n = 1$ a $\rho_1 = -1$, pak $\frac{1}{1+\delta_1} = 1 - \frac{\delta_1}{1+\delta_1}$ a tedy $|\Theta_1| = \frac{|\delta_1|}{|1+\delta_1|} \le \dfrac{\mathbf{u}}{1-\mathbf{u}}$. ■

Načítání chyb - pokračování

pokračování. Předpokládejme, že tvrzení platí pro $n = j-1$.
Nechť $\rho_j = 1$, pak $\prod_{i=1}^{j}(1+\delta_i)^{\rho_i} = (1+\Theta_{j-1})(1+\delta_j) = 1 + \underbrace{\delta_j + \Theta_{j-1} + \delta_j\Theta_{j-1}}_{\Theta_j}$.

$$|\Theta_j| \le |\delta_j| + |\Theta_{j-1}| + |\delta_j\Theta_{j-1}| \le \mathbf{u} + \frac{(j-1)\mathbf{u}}{1-(j-1)\mathbf{u}} + \frac{(j-1)\mathbf{u}^2}{1-(j-1)\mathbf{u}} = \frac{j\mathbf{u}}{1-(j-1)\mathbf{u}} \le \frac{j\mathbf{u}}{1-j\mathbf{u}}.$$

Nechť $\rho_j = -1$, pak $\prod_{i=1}^{j}(1+\delta_i)^{\rho_i} = \dfrac{1+\Theta_{j-1}}{1+\delta_j} = 1 + \underbrace{\dfrac{-\delta_j + \Theta_{j-1}}{1+\delta_j}}_{\Theta_j}$.

$$|\Theta_j| \le \frac{|\delta_j|}{|1+\delta_j|} + \frac{|\Theta_{j-1}|}{|1+\delta_j|} \le \frac{\mathbf{u}}{1-\mathbf{u}} + \frac{1}{1-\mathbf{u}} \cdot \frac{(j-1)\mathbf{u}}{1-(j-1)\mathbf{u}} = \frac{j\mathbf{u} - (j-1)\mathbf{u}^2}{1-j\mathbf{u}+(j-1)\mathbf{u}^2} \le \frac{j\mathbf{u}}{1-j\mathbf{u}}.$$ ■

Značení

Budeme používat výhodné značení $\langle n \rangle = \prod_{i=1}^{n}(1+\delta_i)^{\rho_i}$ pro počítání počtu nakumulovaných relativních chyb.

Platí

$$\langle j \rangle \cdot \langle k \rangle = \langle j+k \rangle \qquad \text{a} \qquad \frac{\langle j \rangle}{\langle k \rangle} = \langle j+k \rangle.$$

Pozor: toto značení nám velmi zjednoduší zápis (pomůže jednoduchému počítání nakumulovaných relativních chyb), ale formálně není korektní, protože oba výrazy $\prod_{i=1}^{n}(1+\delta_i)^{\rho_i}$ a $\prod_{i=1}^{n}(1+\delta_i')^{\rho_i'}$ označíme $\langle n \rangle$ a přitom se obecně nerovnají. Při jejich používání je tedy třeba dávat pozor.

Skalární součin

Základní cvičení 24.2

Mějme pevně danou množinu strojových čísel F (např. v jednoduché přesnosti) a uvažujme standardní model aritmetických operací.

Uvažujme algoritmus $V : F^n \to F$, který počítá skalární součin, tedy

$$V(x_1, x_2, \ldots, x_n) = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n,$$

kde $\alpha_i \in F$ jsou pevně zvolené parametry.

1. Odhadněte dopřednou chybu.
2. Odhadněte zpětnou chybu.

Předpokládáme, že nedojde k podtečení, přetečení apod.

Označme $s_i = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_i x_i$.

$\widehat{s_1} = \mathrm{fl}(\alpha_1 x_1) = \alpha_1 x_1(1 + \delta_1)$, kde $|\delta_1| \leq \mathbf{u}$. Tedy $\widehat{s_1} = \alpha_1 x_1 \langle 1 \rangle$.

$\widehat{s_2} =$

$\widehat{s_3} =$
$\widehat{s_n} = \widehat{V}(x_1, x_2, \ldots, x_n) = \alpha_1 x_1 \langle n \rangle + \alpha_2 x_2 \langle n \rangle + \alpha_3 x_3 \langle n-1 \rangle + \cdots + \alpha_n x_n \langle 2 \rangle$ pro $n > 1$.

$\left| \widehat{V}(x_1, x_2, \ldots, x_n) - V(x_1, x_2, \ldots, x_n) \right| \leq$

$\widehat{V}(\mathbf{x}) = V(x_1 + x_1\Theta_n, x_2 + x_2\Theta_n, x_3 + x_3\Theta_{n-1}, \ldots, x_n + x_n\Theta_2) = V(\mathbf{x} + \Delta\mathbf{x})$.
$\|\Delta\mathbf{x}\|_1 \leq$

Funkční hodnota polynomu

Základní cvičení 24.3
Mějme pevně danou množinu strojových čísel F (např. v jednoduché přesnosti) a uvažujme standardní model aritmetických operací.

Uvažujme zobrazení $p : x \mapsto (x-2)^9$ a 3 způsoby jeho výpočtu:

a) $p_a(x) = (x-2)^9$;

b) $p_b(x) = x^9 - 18x^8 + 144x^7 - 672x^6 + 2016x^5 - 4032x^4 + 5376x^3 - 4608x^2 + 2304x - 512$;

c) $p_c(x) = -512 + x\,(2304 + x\,(-4608 + x\,(\ldots)))$ (Hornerova metoda/pravidlo).

Uvažujme algoritmus $V_z : F \to F : x \mapsto p_z(x)$ pro $z \in \{a, b, c\}$ (tedy počítá fukční hodnotu $p(x)$ 3 výše uvedenými způsoby). Pro všechny 3 varianty

1. odhadněte dopřednou chybu, a
2. odhadněte zpětnou chybu.

Předpokládáme, že nedojde k podtečení, přetečení apod.

Metoda a

$p_a(x) = (x-2)^9$

$\widehat{p_a}(x) =$

$|\widehat{p_a}(x) - p_a(x)| \leq$

$\widehat{p_a}(x) = p_a(x + \Delta x)$, kde $|\Delta x| \leq$

Metoda b

$p_b(x) = x^9 - 18x^8 + 144x^7 - 672x^6 + 2016x^5 - 4032x^4 + 5376x^3 - 4608x^2 + 2304x - 512$

$\widehat{p_b}(x) =$

$|\widehat{p_b}(x) - p_b(x)| \leq$

$\widehat{p_b}(x) = \widehat{p_b}(x + \Delta x)$, kde $|\Delta x| \leq$

18.9 Hornerova metoda - obecně

Hornernova metoda

Chceme vyhodnotit funkční hodnotu polynomu

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

Hornerova metoda spočívá ve vyhodnocení

$$p(x) = \Big(\big(\cdots ((a_n)x + a_{n-1})x + a_{n-2}\big)x + \cdots + a_2\Big)x + a_1\Big)x + a_0,$$

kde je potřeba $2n$ operací s plovoucí čárkou[9] pro $n > 0$.

Hornerova metoda - chyba

Spočítáme chybu vzniklou použitím operací se strojovými čisly.

$$(a_n x\langle 1\rangle + a_{n-1})\langle 1\rangle = a_n x\langle 2\rangle + a_{n-1}\langle 1\rangle.$$

$$\begin{aligned}\big((a_n x\langle 2\rangle + a_{n-1}\langle 1\rangle)x\langle 1\rangle + a_{n-2}\big)\langle 1\rangle &= \\ (a_n x^2\langle 3\rangle + a_{n-1}x\langle 2\rangle + a_{n-2})\langle 1\rangle &= \\ a_n x^2\langle 4\rangle + a_{n-1}x\langle 3\rangle + a_{n-2}\langle 1\rangle.&\end{aligned}$$

Indukcí dostaneme celkovou chybu napočítané hodnoty polynomu p v bodě x označené $\widehat{p}(x)$:

$$\widehat{p}(x) = a_0\langle 1\rangle + a_1 x\langle 3\rangle + \cdots a_{n-1}x^{n-1}\langle 2n-1\rangle + a_n x^n\langle 2n\rangle.$$

[9] *flops*

Hornerova metoda - dopředná chyba

Platí $\langle i \rangle = 1 + \Theta_i$, kde $|\Theta_i| \leq \dfrac{i\mathbf{u}}{1 - i\mathbf{u}} = \gamma_i$.
Odhadneme dopřednou chybu následovně

$$|p(x) - \widehat{p}(x)| \leq \gamma_{2n} \sum_{i=0}^{n} |a_i||x|^i.$$

Pro relativní chybu pak platí

$$\frac{|p(x) - \widehat{p}(x)|}{|p(x)|} \leq \gamma_{2n} \frac{\sum_{i=0}^{n} |a_i||x|^i}{|p(x)|}$$

Toto je *apriorní teoretický odhad* dopředné chyby, a v některých případech je odhad velice nadsazený. Lehce nahlédneme, že pravá strana může být jakkoliv velká.

Místo tohoto horního odhadu si spočítáme k jaké zaokrouhlovací chybě došlo přesněji (tzv. *Running error analysis*).

Hornernova metoda - aposteriorní odhad dopředné chyby (1/2)

Pro dané x definujme posloupnost (q_i) takto: $q_n = a_n$ a $q_i = q_{i+1}x + a_i$ pro všechna $i \in \{0, \ldots, n-1\}$.

Tedy $q_0 = p(x)$.

V i-tém kroku Hornerovy metody platí

$$(1 + \epsilon_i)\widehat{q}_i = \widehat{q}_{i+1}x(1 + \delta_i) + a_i, \quad \text{kde } |\delta_i|, |\epsilon_i| \leq \mathbf{u}.$$

Označme $\widehat{q}_i = q_i + f_i$, dostaneme

$$(1 + \epsilon_i)\widehat{q}_i = (q_{i+1} + f_{i+1})x + x\widehat{q}_{i+1}\delta_i + a_i.$$

Vyjádříme f_i:

$$f_i = \underbrace{q_{i+1}x - q_i + a_i}_{=0} + f_{i+1}x + x\widehat{q}_{i+1}\delta_i + \epsilon_i\widehat{q}_i.$$

Platí $f_n = 0$.

Hornernova metoda - aposteriorní odhad dopředné chyby (2/2)

Odhadneme f_i:

$$|f_i| \leq |f_{i+1}||x| + \mathbf{u}(|x||\widehat{q}_{i+1}| + |\widehat{q}_i|).$$

Označme posloupnost (π_i) tak, aby $|f_i| \leq \mathbf{u}\pi_i$, tedy

$$\pi_n = 0 \quad \text{a} \quad \pi_i = |x|\pi_{i+1} + |x||\widehat{q}_{i+1}| + |\widehat{q}_i|.$$

Aposteriorní dopřednou chybu π_0 tedy můžeme výhodně napočítat během výpočtu $\widehat{q}_0 = \widehat{p}(x)$ a bude platit

$$|\widehat{p}(x) - p(x)| \leq \pi_0\mathbf{u}.$$

Hornerova metoda - zpětná chyba

$$\widehat{p}(x) = a_0\langle 1 \rangle + a_1 x\langle 3 \rangle + \cdots a_{n-1}x^{n-1}\langle 2n-1 \rangle + a_n x^n \langle 2n \rangle.$$

Tedy zpětná chyba je:
$\widehat{p}(x) = p(x + \Delta(x))$, kde $|\Delta(x)| \leq \gamma_{2n}|x|$

Výpočet funkční hodnoty polynomu - relativní podmíněnost

Předpokládejme $p(x) \neq 0$.

Podle věty o střední hodnotě platí

$$|p(x+\delta x) - p(x)| = |p'(\epsilon)\delta x|$$

pro nějaké $\epsilon \in (x, x+\delta x)$ nebo $\epsilon \in (x+\delta x, x)$.

Platí

$$\frac{|p(x+\delta x) - p(x)|}{|p(x)|} = \frac{|p'(\epsilon)\delta x|}{|p(x)|}$$

a tedy

$$C_r = \frac{|p(x+\delta x) - p(x)|/|p(x)|}{|\delta x|/|x|} = \frac{|xp'(\epsilon)|}{|p(x)|}$$

Část IV
Obecná algebra

V matematice a v oblastech, kde se matematika používá, se opakovaně objevují objekty, které mají podobnou strukturu, přestože jsou na první pohled zcela odlišné. Matematici si těchto podobností všímali a postupem času vyvinuli hierarchii takových obecných *nadřazených* struktur, která pokrývá širokou škálu (matematických) objektů. Ta část matematiky, která se těmito strukturami zabývá, se nazývá (obecná[10]) algebra.

19 Úvod do teorie grup

Hledání skrytých podobností ...

Uvažujme následující objekty:

- množina $\mathbb{Z}$ celých čísel a jejich sčítání,
- množina matic $\mathbb{R}^{n,n}$ s operací násobení matic,
- množinu relací na množině A s operací skládání relací,
- množina zobrazení $f : A \to A$ z množiny A do množiny A a operaci skládání zobrazení,
- množinu manipulací s šestiúhelníkem, které jej nemění (zrcadlení dle os, dle středu, otočení, ...) a jejich skládání,
- množinu $\{0, 1, 2, 3\}$ s operací násobení modulo 4,
- množinu konečných automatů a jejich skládání,
- množinu všech konečných řetězců nad zadanou abecedou a jejich spojování,
- množinu všech barev a operaci „míchání“,
- ...

Co mají společného?

Společná struktura!

Všechny uvedené objekty mají stejnou strukturu. Skládají se ze dvou ingrediencí:

- Neprázdné (konečné či nekonečné) **množiny objektů**.
- **Binární operace**, která každé dvojici objektů z této množiny jednoznačně přiřadí objekt z uvažované množiny.

Binární operací na množině M máme na mysli zobrazení $M \times M \to M$.

Obecně se tedy jedná o dvojici množina a binární operace na ní a proto budeme (většinou) používat značení $(M, \cdot)$ (multiplikativní zápis) resp. $(M, +)$ (aditivní zápis), resp. $(M, \circ)$ (obecný zápis), kde

- M je **neprázdná** množina,
- a pro binární operaci platí $\cdot : M \times M \to M$, resp. $+ : M \times M \to M$, resp. $\circ : M \times M \to M$.

[10] Slovo „obecná“ se používá zejména v případech, kdy by někdo mohl nabýt dojmu, že se jedná o lineární algebru.

O co jde v teorii grup?

- Dvojici „množina a binární operace na ní“ mohou, jak bylo dříve na příkladech ukázáno, tvořit velice odlišné struktury. My je budeme klasifikovat podle vlastností, které mají.

- O této dvojici nás budou zajímat například následující otázky: Je operace asociativní? Je komutativní? Existují v množině prvky s vlastnostmi jako má jednička a nula v číselných množinách? …

- Obecná algebra se dále zabývá bohatšími strukturami jako jsou okruhy, tělesa a další. S těmi se setkáme později během semestru.

A proč to děláme?

Pokud dokážeme nějaké tvrzení pro obecnou strukturu $(M, \circ)$, kde $\circ$ je asociativní operace, bude (automaticky) toto tvrzení platit pro všechny konkrétní struktury s asociativní binární operací.

Důkaz tohoto tvrzení se zredukuje na (triviální) důkaz asociativity operace!

Obecnou strukturu můžeme chápat jako **nadřazený objekt**, od kterého konkrétní struktury **dědí** všechny jeho vlastnosti.

Příklad „dědičnosti“ (1/4)

Na množině nenulových reálných čísel dokážeme následující větu:

Věta 19.1 Pro všechna $b, c \in \mathbb{R} \setminus \{0\}$ má rovnice $bx = c$ jediné řešení $x = \frac{c}{b}$. **!**

Důkaz. Následující rovnosti jsou za předpokladu $b, c \in \mathbb{R} \setminus \{0\}$ ekvivalentní.

$$\begin{aligned} bx &= c && \{\text{vyděl } b, \text{ lze pro } \forall b \neq 0\} \\ \frac{bx}{b} &= \frac{c}{b} && \{\text{použijeme asociativitu násobení}\} \\ \frac{b}{b}x &= \frac{c}{b} && \{\text{víme, že pro lib. } b \neq 0 \text{ je } \frac{b}{b} = 1\} \\ 1x &= \frac{c}{b} && \{\text{pro lib. nenulové } d \text{ je } 1d = d\} \end{aligned}$$ ■

Co jsme potřebovali: asociativitu, umět dělit nenulovým reálným číslem, existenci jedničky.

Příklad „dědičnosti“ (2/4)

Uvažujme nyní množinu M všech matic z $\mathbb{R}^{n,n}$ s operací násobení matic.

- Je operace násobení matic asociativní?
  Ano. Pro $\forall A, B, C \in M$ platí $A(BC) = (AB)C$.

- Existuje jednička (neutrální prvek) vůči maticovému násobení?
  Ano. Jednotková matice E má vlastnost $EA = AE = A$ platící pro $\forall A \in M$.

- Existuje ke každé matici $A \in M$ matice inverzní?

  Neexistuje! Musíme se omezit na množinu všech **regulárních** matic z M, označovanou též $M_{\text{reg}} := \{A \in M \mid A \text{ je regulární}\}$.

Poznámka: Všechna tato tvrzení již známe z Lineární algebry!

Příklad „dědičnosti“ (3/4)

Máme vše co potřebujeme, větu můžeme přeformulovat pro matice:

Věta 19.2 Pro všechna $B, C \in M_{\text{reg}}$ má rovnice $BX = C$ jediné řešení $X = B^{-1}C$. !

Důkaz. Následující rovnosti jsou za předpokladu $B, C \in M_{\text{reg}}$ ekvivalentní.

$$\begin{aligned} BX &= C && \{\text{vynás. inverz. prvkem } B^{-1} \text{ zleva, existuje pro } \forall B\} \\ B^{-1}(BX) &= B^{-1}C && \{\text{přesuneme závorky díky asociativitě}\} \\ (B^{-1}B)X &= B^{-1}C && \{\text{víme, že pro lib. } B \text{ je } B^{-1}B = E\} \\ EX &= B^{-1}C && \{\text{pro lib. matici } D \text{ je } ED = D\} \\ X &= B^{-1}C \end{aligned}$$

■

Co jsme potřebovali: asociativitu, existenci inverzní matice, existenci jednotkové matice.

Příklad „dědičnosti“ (4/4)

Dvojici $(M, \circ)$, kde $\circ : M \times M \to M$, platí asociativní zákon, existuje neutrální prvek a ke každému prvku b existuje inverzní prvek (značený b^{-1}) budeme říkat grupa. Obecná věta pak bude znít:

Věta 19.3 Pro libovolné prvky b, c z grupy $(M, \circ)$ má rovnice $b \circ x = c$ jediné řešení $x = b^{-1} \circ c$. !

Důkaz. Následující rovnosti jsou ekvivalentní.

$$\begin{aligned} b \circ x &= c \\ b^{-1} \circ (b \circ x) &= b^{-1} \circ c \\ (b^{-1} \circ b) \circ x &= b^{-1} \circ c \\ x &= b^{-1} \circ c \end{aligned}$$

■

V předchozích příkladech pro neutrální prvky a inverzní prvky platilo:

- $(\mathbb{R} \setminus \{0\}, \cdot)$: neutrálním prvkem je číslo 1, inverzním prvkem k $b \neq 0$ je $\frac{1}{b} \neq 0$.
- $(M_{\text{reg}}, \cdot)$: neutrálním prvkem je jednotková matice příslušného rozměru, inverzním prvkem k $A \in M_{\text{reg}}$ je matice k ní inverzní.

?

Kontrolní otázka 19.1. *Platí Věta 19.3 i pro dvojici $(\mathbb{Z}, +)$? Jak vypadá inverzní prvek a neutrální prvek pro tuto dvojici?*

?

Kontrolní otázka 19.2. *V grupě $(\mathbb{R} \setminus \{0\}, \cdot)$ můžeme zavést operaci „dělení“ jako násobení inverzním prvkem:*

$$\frac{a}{b} := a \cdot b^{-1}.$$

*Je možné takto zavést dělení i v množině regulárních matic M_{reg}? Proč **ne**?*

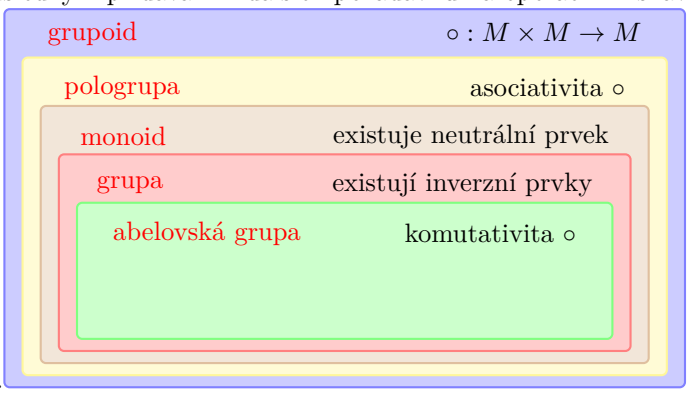
?

Kontrolní otázka 19.3. *Zkuste vymyslet, jaké barvy bychom museli přidat do množiny barev, aby byla vzniklá množina s operací „míchání barev“ grupou.*

20 Hierarchie množin s jednou binární operací

Množiny s jednou binární operací

Libovolnou dvojici tvořenou neprázdnou množinou a binární operací na ní, $(M, \circ : M \times M \to M)$, budeme nazývat grupoid. Následným přidáváním dalších požadavků na operaci $\circ$ získáváme další struktury...



Grupoid, pologrupa, monoid, grupa

Definice 20.1 Grupoid (*magma*) je uspořádaná dvojice $(M, \circ)$, kde M je libovolná neprázdná množina a $\circ$ je binární operace na M.

- Pologrupa (*semigroup*) je grupoid $(M, \circ)$, pro který je $\circ$ asociativní operace.
- Monoid je pologrupa $(M, \circ)$, ve které existuje neutrální prvek $e \in M$ takový, že

$$\text{pro všechna } a \in M \quad \text{platí} \quad e \circ a = a \circ e = a.$$

- Grupa (*group*) je monoid $(M, \circ)$, ve kterém ke každému $a \in M$ existuje inverzní prvek $b \in M$ takový, že

$$b \circ a = a \circ b = e.$$

- Komutativní (abelovská) grupa je grupa $(M, \circ)$, kde $\circ$ je komutativní operace.

První příklady

- Pro dvojici $(\mathbb{Z}, +)$ platí asociativní i komutativní zákon, neutrálním prvkem je 0 a inverzní prvek k prvku a je prvek $-a$, součet dvou celých čísel je celé číslo, jedná se tedy o abelovskou grupu.
- Pro dvojici $(\mathbb{R} \setminus \{0\}, \cdot)$ platí asociativní i komutativní zákon, neutrálním prvkem je 1 a inverzní prvek k (nenulovému) prvku a je $\frac{1}{a}$, součin dvou nenulových reálných čísel je nenulové reálné číslo, jedná se tedy o abelovskou grupu.
- Pro dvojici $(M_{\text{reg}}, \cdot)$ platí asociativní zákon, neutrální prvek i inverní prvky existují (k $A \in M_{\text{reg}}$ je inverzním prvkem inverzní matice A^{-1}), ale neplatí komutativní zákon! Jedná se tedy o grupu, nikoli ovšem vždy abelovskou. Příklad nekomutujících matic:

$$\begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}.$$

Otázka: Je součin dvou regulárních matic regulární matice?

?

Kontrolní otázka 20.1. *Co se změní, pokud namísto* $(\mathbb{R} \setminus \{0\}, \cdot)$ *uvažujeme* $(\mathbb{R}, \cdot)$*?*

?

Kontrolní otázka 20.2. *Co se změní, pokud namísto* $(M_{reg}, \cdot)$ *uvažujeme* $(\mathbb{R}^{n,n}, \cdot)$*?*

?

Kontrolní otázka 20.3. *Jaký je rozdíl mezi dvojicí* $\{1,2,3,4\}$ *s násobením* $\pmod 5$ *a dvojicí* $\{1,2,3,4,5\}$ *s násobením* $\pmod 6$*?*

Matematická analogie k OOP

- Na grupoid, monoid, atd. se můžeme dívat jako na matematický (abstraktní) objekt či **třídu**, pro který je definována nějaká neprázdná množina a binární operace s danými vlastnostmi.
- Pro tyto abstraktní třídy můžeme dokázat různá tvrzení (jako např. Větu o řešení „lineární“ rovnice pro grupy).
- Pokud potom pro nějakou instanci $(M, \circ)$ ukážeme, že je grupoid, monoid, atp., znamená to, že všechny tato tvrzení „zdědí“ a nemusíme je tedy dokazovat zvlášť.

Poznámky ke značení a terminologii

- O množině M také mluvíme jako o nosiči (*carrier*) grupy $G = (M, \circ)$.
- Zápisem $a \in G$, kde $G = (M, \circ)$, máme na mysli $a \in M$.

Značení 20.2. ***obecný zápis***

$(M, \circ)$

neutrální prvek: e

inverzní prvek k $a \in M$*:* a^{-1}

$\underbrace{a \circ a \circ \cdots \circ a}_{n-1 \text{ operací}} = a^n$

$a^0 = e$
$a^n \circ a^{-n} = e$
multiplikativní zápis

$(M, \cdot)$

neutrální prvek: 1

inverzní prvek k $a \in M$*:* a^{-1}

$\underbrace{a \cdot a \cdots a}_{n-1 \text{ operací}} = a^n$

$a^0 = 1$
$a^n \cdot a^{-n} = 1$
aditivní zápis

$(M, +)$

neutrální prvek: 0

inverzní prvek k a*:* $-a$

$\underbrace{a + a + \cdots + a}_{n-1 \text{ operací}} = n \times a$

$0 \times a = 0$
$\big(n \times a\big) + \big((-n) \times a\big) = 0$

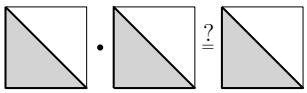
Uzavřenost množiny vůči binární operaci

V definici klademe na binární operaci $\circ$ podmínku, aby byla „binární operací na M“, což znamená, že výsledek aplikování binární operace na dva prvky z M opět patří do M, stručněji $\circ : M \times M \to M$. Též říkáme, že **množina M je uzavřená vůči $\circ$.**

■ **Příklad 20.3** Dvojice $(\mathbb{Z}^-, \cdot)$ záporných celých čísel s klasickým násobením není ani grupoid, neboť $\mathbb{Z}^-$ není uzavřená vůči násobení: například $(-1) \cdot (-1) = 1 \notin \mathbb{Z}^-$. ■

Uzavřenost či neuzavřenost vůči binární operaci nemusí být vždy očividná:

■ **Příklad 20.4** Uvažujme dvojici $(M_{\text{troj}}, \cdot)$ dolních trojúhelníkových matic s klasickým maticovým násobením. Je M_{troj} vůči operaci $\cdot$ uzavřená?



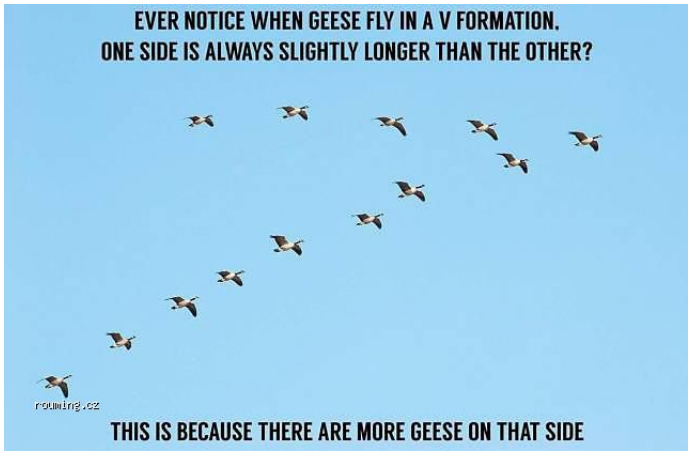
■

Ověřování vlastností $(M, \circ)$

Máme-li zadanou dvojici „množina a operace“ a chceme-li zjistit, jestli se jedná o grupoid, pologrupu, monoid, (abelovskou) grupu, můžeme systematicky postupovat (nepřekvapivě) v následujícím pořadí:

1. Je množina uzavřená vůči operaci? Pokud ano, je to grupoid, pokud ne, konec.
2. Je operace asociativní? Pokud ano, je to pologrupa, pokud ne, konec.
3. Existuje neutrální prvek? Pokud ano, je to monoid, pokud ne, konec.
4. Existuje ke každému prvku inverzní prvek? Pokud ano, je to grupa, pokud ne, konec.
5. Je operace komutativní? Pokud ano, je to abelovská grupa, pokud ne, konec.

Ověřování vlastností $(M, \circ)$



Uvedená „kuchařka“ je úplně zbytečná pokud znáte definice pojmů a chápete jejich hierarchii.

21 Příklady

Příklad č. 1

■ **Příklad 21.1** Uvažujme grupoid $(\mathbb{Q}, \circ)$, kde binární operace $\circ$ je definována jako aritmetický průměr:

$$a \circ b := \frac{a+b}{2}\,.$$

Jedná se o grupoid / pologrupu / monoid / grupu? ■

Pro racionální čísla a a b je výraz $\frac{a+b}{2}$ racionální číslo. Jde tedy o grupoid. V pologrupě musí platit asociativní zákon. Tvrdíme, že pro takto definovanou operaci $\circ$ *neplatí* a dokážeme to *protipříkladem*:

$$(2 \circ -2) \circ 4 = 0 \circ 4 = 2 \quad \text{ale} \quad 2 \circ (-2 \circ 4) = 2 \circ 1 = \frac{3}{2}\,.$$

Skutečně, pro takto zvolené prvky $\mathbb{Q}$ asociativní zákon neplatí a nejedná se proto o pologrupu. Z toho je již jasné, že $\mathbb{Q}$ s takto definovanou operací není ani monoid a ani grupa.

Příklad č. 2

■ **Příklad 21.2** Uvažujme grupoid $(\mathbb{R}^+, \circ)$, kde binární operace $\circ$ je definována takto:

$$a \circ b := \frac{a \cdot b}{a+b}\,.$$

Jedná se o pologrupu / monoid / grupu? ■

Poznámka: $\mathbb{R}^+ = (0, +\infty)$. V pologrupě musí platit asociativní zákon. Tvrdíme, že pro takto definovanou operaci $\circ$ asociativní zákon platí. Tzn., že chceme ukázat, že platí rovnost $(a \circ b) \circ c = a \circ (b \circ c)$. Postupujeme tak, že si napíšeme levou stranu rovnice a *pomocí definice operace a vlastností klasického násobení a dělení* levou stranu upravíme na pravou:

$$\begin{aligned}(a \circ b) \circ c = \{\text{ definice } \circ\} = \frac{a \cdot b}{a+b} \circ c = \{\text{ definice } \circ\} = \frac{\frac{a\cdot b}{a+b} \cdot c}{\frac{a\cdot b}{a+b} + c} = \\ = \frac{(a \cdot b) \cdot c}{a \cdot b + c \cdot (a+b)} = \frac{a \cdot (b \cdot c)}{a \cdot (b+c) + b \cdot c} = \frac{a \cdot \frac{b\cdot c}{b+c}}{a + \frac{b\cdot c}{b+c}} = \\ = \{\text{ definice } \circ\} = a \circ \frac{c \cdot b}{c+b} = \{\text{ definice } \circ\} = a \circ (b \circ c)\,.\end{aligned}$$

Asociativní zákon platí a jedná se o pologrupu.

Příklad č. 2

(…pokračování…)

Dokázali jsme, že $(\mathbb{R}^+, \circ)$ je pologrupa, je to též monoid? Neboli existuje neutrální prvek $e \in \mathbb{R}^+$ tak, že

$$(\forall a \in \mathbb{R}^+)(e \circ a = a \circ e = a)\ ?$$

Neutrální prvek e musí pro všechna $a \in \mathbb{R}^+$ splňovat rovnici

$$(a \circ e = a) \Rightarrow \left(\frac{a \cdot e}{a+e} = a\right) \Rightarrow (a \cdot e = a \cdot (a+e)) \Rightarrow (e = a+e) \Rightarrow (0 = a)\,,$$

která ale platí pouze pro $a = 0$. Neutrální prvek tedy neexistuje a o monoid se nejedná.

Příklad č. 3

■ **Příklad 21.3** Uvažujme grupoid $(\mathbb{R}, \cdot)$, kde za binární operaci $\cdot$ bereme klasické násobení čísel.

Jedná se o pologrupu / monoid / grupu? ■

Asociativita $\cdot$ je známá vlastnost násobení reálných čísel. Jedná se tedy o pologrupu.

Aby se jednalo o monoid, musí existovat neutrální prvek. Existuje?

Pro neutrální prvek e musí platit: $(\forall a \in \mathbb{R})(e \cdot a = a \cdot e = a)$ a takový prvek v $\mathbb{R}$ existuje a je to číslo 1. Grupoid $(\mathbb{R}, \cdot)$ je tedy dokonce monoid.

A je to grupa? Neboli, existuje ke každému reálnému číslu a číslo a^{-1}, tak že $a \cdot a^{-1} = a^{-1} \cdot a = 1$? Existuje, ovšem pouze pokud a není nula. O grupu se tedy nejedná.

Důsledek

Z definice plyne, že každá grupa je monoid, každý monoid je pologrupa a každá pologrupa je grupoid. To můžeme symbolicky zapsat takto:

$$\text{grupoid} \supset \text{pologrupa} \supset \text{monoid} \supset \text{grupa}\,.$$

Díky předchozím třem příkladům můžeme tuto hierarchii o něco zpřesnit:

$$\text{grupoid} \supsetneq \text{pologrupa} \supsetneq \text{monoid} \supsetneq \text{grupa}\,,$$

neboť jsme našli grupoid, který není pologrupou, pologrupu, která není monoidem a monoid, který není grupou.

Příklad: grupa $\mathbb{Z}_n^+$

■ **Příklad 21.4 — $\mathbb{Z}_n^+$.** Pro kladné $n \in \mathbb{N}$ je

$$\mathbb{Z}_n^+ := \big(\{0, 1, \ldots, n-1\}, +_n\big),$$

kde $+_n$ je sčítání modulo n, abelovská grupa. ■

- Množina $\mathbb{Z}_n := \{0, 1, \ldots, n-1\}$ je množina zbytkových tříd po dělení n. Množina M je uzavřená vůči operaci sčítání modulo n.

- Uvažme $a, b, c \in \mathbb{Z}_n$ libovolné, potom existují $j_i, k_i \in \mathbb{Z}$, $i = 1, 2$, tak, že

$$\begin{aligned}(a +_n b) +_n c &= (a +_n b) + c + j_1 n = (a + b) + c + (k_1 + j_1)n,\\ a +_n (b +_n c) &= a + (b +_n c) + j_2 n = a + (b + c) + (k_2 + j_2)n.\end{aligned}$$

  Zde symbol $+$ představuje standardní sčítání celých čísel.

  Díky asociativitě $+$ tedy platí

$$(a +_n b) +_n c \equiv a +_n (b +_n c) \pmod{n}$$

  a dle definice $+_n$ pak i $(a +_n b) +_n c = a +_n (b +_n c)$.

- Neutrálním prvkem vůči $+_n$ je 0.

- Inverzním prvkem k $a \in \mathbb{Z}_n$, $a \neq 0$, je $n - a$. Inverzním prvkem k 0 je 0.

- Operace sčítání modulo n je komutativní díky komutativitě $+$.

Příklad: grupa $\mathbb{Z}_n^\times$ (1 ze 2)

■ **Příklad 21.5** — $\mathbb{Z}_n^\times$. Pro kladné $n \in \mathbb{N}$, $n \geq 2$, je

$$\mathbb{Z}_n^\times := \big(\{k \in \{1, \ldots, n-1\} : k \text{ a } n \text{ jsou nesoudělná}\}, \times_n\big),$$

kde $\times_n$ je násobení modulo n, abelovská grupa. ■

Tvrdíme: jsou-li k a ℓ z $\{1, \ldots, n-1\}$ nesoudělná s n, pak i $k \times_n \ell$ je nesoudělné s n.

Důkaz sporem: buďte k a ℓ nesoudělná s n a $d \in \{0, \ldots, n-1\}$ takové, že $k\ell = jn + d$ pro nějaké $j \in \mathbb{Z}$. Předpokládejme, že d je soudělné s n. Jelikož d je soudělné s n, existuje prvočíslo dělící čísla d i n, které musí dělit i k nebo ℓ, což je spor.

Máme tedy uzavřenost množiny vzhledem k dané operaci.

Asociativita operace $\times_n$ plyne z asociativity násobení reálných čísel podobně jako v předcházejícím příkladě.

Neutrální prvek je 1.

Příklad: grupa $\mathbb{Z}_n^\times$ (2 ze 2)

Inverzní prvky: je-li k nesoudělné s n, existují $\alpha, \beta \in \mathbb{Z}$ tak, že

$$\alpha k + \beta n = 1 \quad \Rightarrow \quad \alpha k \equiv 1 \pmod{n}.$$

Jelikož je takové číslo α nesoudělné s n, tak $\alpha \bmod n$ je hledaný inverzní prvek.

Existence čísel α a β plyne z tvrzení, které je známě jako Bézoutova rovnost. Též se jim říká Bézoutovy koeficienty.

$\times_n$ je komutativní. $\mathbb{Z}_n^\times$ je tedy skutečně abelovská grupa.

Nemohlo by v nosiči být více zbytků po dělení n (vzhledem k $\times_n$)?

Ne, čísla **soudělná s** n nemají inverzní prvek: uvažujme jakékoli číslo $1 < k < n$ soudělné s n. Předpokládejme, že ℓ je inverzní prvek k prvku k, tedy že platí $k\ell \equiv 1 \pmod{n}$. Tedy pro nějaké j platí $k\ell = 1 + jn$, a tedy $1 = k\ell - jn = d(\frac{k}{d}\ell - j\frac{n}{d})$, kde d, $d > 1$, je společný faktor k a n. Tím dostaneme spor, tedy k prvku k bychom nenašli inverzní prvek.

Terminologie: modulární grupy celých čísel

$\mathbb{Z}_n^+$ je **aditivní modulární grupa (celých čísel) modulo** n.

$\mathbb{Z}_n^\times$ je **multiplikativní modulární grupa (celých čísel) modulo** n.

Jiná běžná značení: $\mathbb{Z}_n^*$, $(\mathbb{Z}/n\mathbb{Z})^\times$.

22 Vlastnosti neutrálních a inverzních prvků

Jednoznačnost neutrálního prvku

Věta 22.1 V monoidu existuje právě jeden neutrální prvek. !

Důkaz. Buď $(G, \circ)$ monoid a e nějaký neutrální prvek (z definice víme, že tam alespoň jeden je!). Dokážeme *sporem*, že e je jediný neutrální prvek.

Pro spor předpokládejme, že v monoidu existuje další neutrální prvek $\overline{e}$ různý od e, tj. $e \neq \overline{e}$. Potom platí

$$\overline{e} = \overline{e} \circ e = e,$$

kde jsme použili vlastnost neutrálního prvku danou definicí. Tím dostáváme spor s tím, že $\overline{e}$ a e jsou různé. ■

Jednoznačnost inverzního prvku

Věta 22.2 V grupě má každý prvek právě jeden inverzní prvek. !

Důkaz. Buď $(G, \circ)$ grupa, a libovolný prvek této grupy a nějaký k němu inverzní prvek b (z definice grupy víme, že tam alespoň jeden je). Dokážeme *sporem*, že b je jediný inverzní prvek.

Pro spor předpokládejme, že v grupě existuje jiný inverzní prvek c různý od b. Potom platí

$$c = c \circ e = c \circ (a \circ b) = (c \circ a) \circ b = e \circ b = b,$$

kde e je neutrální prvek. Tím dostáváme spor s tím, že c a b jsou různé. ■

Inverzní prvek - značení

Značení 22.3. *V souladu se zavedeným značením 20.2 značíme inverzní prvek k prvku a grupy $(G, \circ)$ pro obecné (multiplikativní) značení takto:*

$$a^{-1},$$

a pro aditivní značení takto:

$$-a.$$

23 Znázornění grup

23.1 Cayleyho tabulka grupy

Cayleyho tabulka pro konečné grupy

Pokud má množina M z dvojice $(M, \circ)$ konečný počet prvků, lze její strukturu (danou operací $\circ$) kompletně zachytit v tzv. Cayleyho tabulce, jejíž konstrukce je zřejmá z následujícího příkladu.

■ **Příklad 23.1** Uvažujme grupu $\mathbb{Z}_4^+$. Jelikož má její nosič 4 prvky, bude mít Cayleyho tabulka 4 řádky a 4 sloupce označené těmito čtyřmi prvky (takže bude typicky znázorněná jako tabulka 5x5).

$+_4$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Nyní do pole na řádku m a v sloupci n vyplníme výsledek $m +_4 n = m + n \pmod 4$. Například do řádku 2 a sloupce 3 vyplníme $2 + 3 \pmod 4 = 1$. ■

Jak poznat různé věci z Cayleyovy tabulky

Cayleho tabulka skýtá o dané množině a operaci veškeré informace. Některé vlastnosti lze z tabulky vyčíst velmi snadno, jiné už hůře:

- Uzavřenost množiny M vůči operaci $\circ$ poznáme tak, že všechny pole tabulky obsahují prvky z množiny M.
- Asociativitu operace z tabulky poznáme těžko (projděte všechny trojice prvků a ověřte splnění asociativity…).
- Neutrální prvek e v tabulce poznáme tak, že v „jeho“ řádku a sloupci se přesně opakují označení řádku a sloupce tabulky.
- Inverzní prvek k prvku a najdeme tak, že najdeme v příslušně označeném řádku a sloupci neutrální prvek e …

- Operace je komutativní, pokud je tabulka symetrická vůči hlavní diagonále.

Poznámka: Cayleyho tabulka není příliš praktická v případě, kdy množina M má velký počet prvků.

Cayleyho tabulka a latinský čtverec (1/4)

Otázka: Lze z Cayleyho tabulky snadno poznat, jestli se jedná o tabulku grupy?

Odpověď: Skoro.

Věta 23.2 Cayleyho tabulka každé grupy tvoří latinský čtverec. !

- Latinský čtverec pro n prvkovou množinu M je matice $n \times n$ taková, že v každém řádku i sloupci jsou vždy všechny prvky množiny M.

- Větu dokážeme tak, že dokážeme jinou větu, ze které už bude důkaz této věty triviálně vyplývat.

- Bohužel ne každá Cayleho tabulka tvořící latinský čtverec je tabulkou grupy. Později si ukážeme protipříklad.

Cayleyho tabulka a latinský čtverec (2/4)

Věta 23.3 V každé grupě lze jednoznačně dělit. !
Tzn.: V každé grupě $(G, \circ)$ mají pro libovolné $a, b \in G$ rovnice

$$a \circ x = b \quad \text{a} \quad y \circ a = b \quad \text{jediné řešení.}$$

Tuto větu jsme dokázali dříve v této přednášce ještě před zavedením pojmu grupa, přesto zde uvedeme další důkaz, lehce „algebraičtější“.

Důkaz. Jelikož se jedná o grupu, každý prvek má (jediný) inverzní prvek, a snadno tedy zjistíme, že řešením rovnic jsou prvky $a^{-1} \circ b$ resp. $b \circ a^{-1}$.

Jednoznačnost se dokáže sporem: nechť existuje řešení $x_1 \neq a^{-1} \circ b$, potom

$$x_1 = (a^{-1} \circ a) \circ x_1 = a^{-1} \circ (a \circ x_1) = a^{-1} \circ b\,. \qquad \blacksquare$$

Lze ukázat, že grupa je pologrupa, ve které lze jednoznačně dělit, tzn. že jednoznačnost dělení zajišťuje existenci neutrálního prvku a inverzních prvků, ovšem za předpokladu, že je operace asociativní!

Cayleyho tabulka a latinský čtverec (3/4)

Nyní dokážeme předchozí větu, která říká že Cayleyho tabulka grupy je latinský čtverec:

Důkaz. Dokážeme to sporem:

- Předpokládejme, že tabulka nějaké grupy $(G, \circ)$ není latinský čtverec.
- To znamená, že v nějakém řádku nebo sloupci se jeden prvek, označme jej b, opakuje dvakrát. Bez újmy na obecnosti předpokládejme, že je to v řádku n ve sloupcích m_1 a m_2.

$\circ$	$\cdots$	m_1	$\cdots$	m_2	$\cdots$
$\vdots$		$\vdots$		$\vdots$	
n	$\cdots$	b	$\cdots$	b	$\cdots$
$\vdots$		$\vdots$		$\vdots$	

- Z toho ale okamžitě vyplývá, že rovnice $n \circ x = b$ má dvě různá řešení m_1 a m_2, což je **spor s předchozí větou**! $\blacksquare$

Cayleyho tabulka a latinský čtverec (4/4)

- Ukázali jsme, že to, že Cayleyho tabulka je latinský čtverec, je *nutnou* podmínkou pro to, aby daná množina a operace byla grupou.
- Jak ukazuje následující protipříklad, nejedná se o podmínku *postačující*.

■ **Příklad 23.4** Uvažujme množinu $M = \{a, b, c\}$ s operací zadanou touto Cayleyho tabulkou:

$\circ$	a	b	c
a	b	a	c
b	c	b	a
c	a	c	b

Tato tabulka tvoří latinský čtverec, ale přesto není tabulkou grupy. (Proč?) ■

23.2 Cayleyho graf grupy

Cayleyho graf grupy

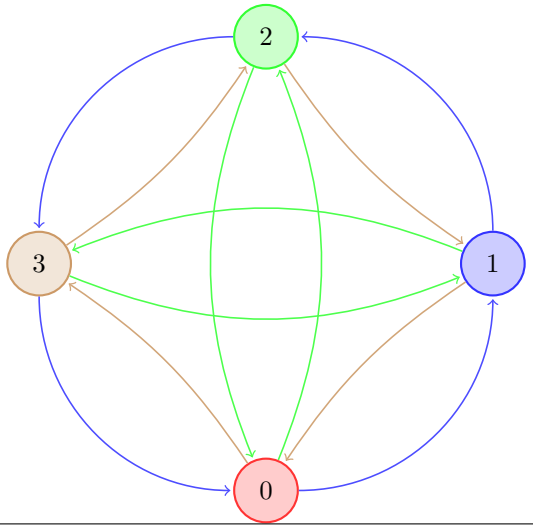
Vizualizovat konečnou grupu $G = (M, \circ)$ lze dále pomocí Cayleyho orientovaného grafu (V, E), kde

- vrcholy grafu V jsou prvky grupy, tj. $V = M$,
- orientovaná hrana (a, b) patří do E, právě když $b = a \circ c$ pro jisté $c \in N$.

Množinu N vhodně zvolíme předem (většinou neobsahuje neutrální prvek grupy G; typicky jde o generující množinu, viz příští přednáška). Hrany lze pro přehlednost obarvit podle příslušnosti k c. Podobně jako u Cayleyho tabulek platí,

že se Cayleyho grafy hodí pro grupy s menším počtem prvků. Lépe v nich navíc vynikne struktura grupy (viz dále). Pokud by grupa nebyla abelovská, museli bychom kreslit i hrany (a, b) pro $a = b \circ c$ pro $c \in N$.

Cayleyho graf grupy: příklad $\mathbb{Z}_4^+$



Odpovědi na některé kontrolní otázky

Odpověď na kontrolní otázku 19.1. Pro dvojici $(\mathbb{Z}, +)$ věta platí v tomto tvaru: pro všechna b, c má rovnice $b + x = c$ právě jedno řešení $x = -b + c$, neb inverzní prvek k b (vzhledem k operaci $+$) je prvek $-b$.

Odpověď na kontrolní otázku 19.2. V M_{reg} dělení zavést nelze, neboť operace násobení matic není komutativní, a tedy $AB^{-1} \neq B^{-1}A$. To znamená, že by nebylo jasné, který z těchto dvou výrazů by se rovnal $\frac{A}{B}$.

Odpověď na kontrolní otázku 20.3. Při násobení $\pmod 6$ nemají všechny prvky inverzní prvek, tedy neexistuje x tak, aby $ax \equiv 1 \pmod 6$, kde 1 je neutrální prvek a $a \in \{2, 3, 4\}$. Navíc se nejedná ani o grupoid, neboť $2 \cdot 3 \pmod 6 \notin \{1, 2, 3, 4, 5\}$.

24 Podgrupy

Příklad $\mathbb{Z}_{12}^+$ (1 ze 3)

■ **Příklad 24.1** Uvažme množinu $\mathbb{Z}_{12} := \{0, 1, 2, \ldots, 11\}$ se sčítáním modulo 12. Podle minulé přednášky tato dvojice tvoří aditivní grupu modulo 12 značenou $\mathbb{Z}_{12}^+$. ■

Otázka: Jaká jiná množina M tvoří s operací sčítání $\bmod 12$ grupu?

Aby tato binární operace měla smysl, musí být $M \subseteq \mathbb{Z}_{12}$:

Otázka (upřesnění): Jaké podmnožiny $\mathbb{Z}_{12}$ tvoří s operací sčítání $\bmod 12$ grupu?

Odpověď: Je jich poměrně hodně a abychom přišli na to, jak je získat, položme si podotázku:

Podotázka: Jaká nejmenší podmnožina $\mathbb{Z}_{12}$ tvoří s operací $+_{12}$ (sčítání modulo 12) grupu a obsahuje číslo $2 \in \mathbb{Z}_{12}$?

Příklad $\mathbb{Z}_{12}^+$ (2 ze 3)

Hledáme množinu $M \subseteq \mathbb{Z}_{12}$ tak, aby $2 \in M$ a $(M, +_{12})$ byla grupa:

- M musí být vůči sčítání modulo 12 uzavřená:
  - musí proto obsahovat $2 +_{12} 2 = \mathbf{4}$, $2 +_{12} 4 = \mathbf{6}$, $4 +_{12} 6 = \mathbf{10}, \ldots$
  - množina $\{0, 2, 4, 6, 8, 10\}$ už je uzavřená a tedy máme grupoid
- operace sčítání modulo 12 i na této podmnožině zůstává asociativní, je to pologrupa,
- 0 zůstává neutrálním prvkem, je to monoid,
- a každý prvek má inverzní prvek patřící do této množiny ($-0 = 0$, $-2 = 10$, $-4 = 8$, $-6 = 6$, $-8 = 4$, $-10 = 2$), je to grupa.

Hledaná množina je tedy $M = \{0, 2, 4, 6, 8, 10\}$: říkáme, že M je podgrupa generovaná množinou $\{2\}$.

Příklad $\mathbb{Z}_{12}^+$ (3 ze 3)

Podobně jako jsme vygenerovali množinu pro prvek 2, můžeme postupovat i pro jiné prvky $\mathbb{Z}_{12}$:

$$
\begin{array}{lcl}
\{0\} \to & \{0\} & \\
\{1\} \to & \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\} & \leftarrow \{11\} \\
\{2\} \to & \{0, 2, 4, 6, 8, 10\} & \leftarrow \{10\} \\
\{3\} \to & \{0, 3, 6, 9\} & \leftarrow \{9\} \\
\{4\} \to & \{0, 4, 8\} & \leftarrow \{8\} \\
\{5\} \to & \{0, 5, 10, 3, 8, 1, 6, 11, 4, 9, 2, 7\} & \leftarrow \{7\} \\
\{6\} \to & \{0, 6\} &
\end{array}
$$

Závěr: nalezli jsme 6 různých množin M takových, že $(M, +_{12})$ tvoří grupu.

Fakt, že těchto množin je právě 6 a my jsme žádnou neopomněli, není zcela zřejmý. My jsme našli všechny takové množiny, které lze vygenerovat jedním prvkem (takovým budeme říkat cyklické): např. $M = \{0, 2, 4, 6, 8, 10\}$ byla vygenerovaná prvkem 2. Později si ukážeme, že cyklická grupa (a Z_{12}^+ je cyklická) má všechny podgrupy (zde množiny M) taktéž cyklické a tedy jsme skutečně našli všechny! To si prosím rozmyslete!!

Definice podgrupy

Definice 24.2 — Podgrupa (*subgroup*). Buď $G = (M, \circ)$ grupa. Podgrupou grupy G nazveme libovolnou dvojici $H = (N, \circ)$ takovou, že

- $N \subseteq M$,
- $(N, \circ)$ je grupa.

- Tato konstrukce, kde se z dané struktury vezme podstruktura, která má stejné vlastnosti, je v matematice častá: vzpomeňte lineární prostor a podprostor.
- Podobně bychom mohli definovat podgrupoid, podpologrupu a podmonoid, ale nebudeme.
- Binární operaci v grupě $G = (M, \circ)$ chápeme jako zobrazení z $M \times M$ do M, operace v podgrupě $H = (N, \circ)$ je tedy exaktně řečeno zúžení původní operace na množinu $N \times N$.

Poznámka k předchozí poznámce 246: příkladem grupy, kde bychom nebyli s to najít všechny podgrupy tak, že bychom je vygenerovali z jednotlivých prvků, jako jsme učinili v případě $\mathbb{Z}_{12}^{+}$, je tzv. Dihedrální grupa D_4, více informací o této grupě na http://mathworld.wolfram.com/DihedralGroupD4.html.

(Ne)triviální podgrupy

V každé grupě $G = (M, \circ)$ s alespoň dvěma prvky existují vždy alespoň dvě (různé) podgrupy:

- grupa obsahující pouze neutrální prvek: $(\{e\}, \circ)$
- a grupa samotná: $G = (M, \circ)$.

Těmto dvěma grupám se říká triviální podgrupy, ostatním podgrupám se říká netriviální nebo vlastní podgrupy.

Kontrolní otázka 24.1. *Je-li H podgrupa grupy G, musí být vždy neutrální prvek v H shodný s neutrálním prvkem G?*

Průnik podgrup je opět podgrupa

Věta 24.3 Pro každé i z indexové množiny $\mathcal{I}$ buď H_i podgrupa grupy $G = (M, \circ)$, potom platí, že

$$H' = \bigcap_{i \in \mathcal{I}} H_i \quad \text{je také podgrupa grupy } G.$$

Důkaz. Jistě je H' podmnožinou M a je neprázdná.

- H' je uzavřená vůči operaci $\circ$: pro všechna $a, b \in H'$ platí: pro všechna $i \in \mathcal{I}$ máme $a, b \in H_i$, a jelikož H_i jsou uzavřené vůči $\circ$, tak platí $a \circ b \in H_i$; celkem tedy pro všechna $a, b \in H'$ máme $a \circ b \in H'$ a uzavřenost je dokázána.
- Operace jistě zůstává asociativní, neutrální prvek zůstává stejný jako v G.
- Uzavřenost vůči inverzi prvku se ukáže stejně jako uzavřenost vůči operaci $\circ$. ■

Indexová množina může být konečná (např. $\mathcal{I} = \{1, 2, \ldots, n\}$) i nekonečná.

Kritérium pro „být podgrupou“

Při ověřování, zda-li jistá podmnožina již známe grupy vytváří podgrupu, nemusíme ověřovat všechny podmínky v definici grupy. Máme k dispozici následující užitečnou větu:

Věta 24.4 Buď $G = (M, \circ)$ grupa a $N \subseteq M$ neprázdná množina. Dvojice $H = (N, \circ)$ je podgrupa grupy G, právě když pro každé $a, b \in N$ platí $a \circ b^{-1} \in N$. !

Důkaz. Implikace zleva doprava je zřejmá. Ověřme implikaci zprava doleva.

Postupně zkontrolujeme, že $(N, \circ)$ je grupa.

- Operaci $\circ$ jsme nijak nezměnili a její asociativita je tedy zachována.
- Vezměme $a \in N$, potom $e = a \circ a^{-1} \in N$.
- Uvažme $a \in N$, potom $a^{-1} = e \circ a^{-1} \in N$.
- Pro $a, b \in N$ platí $a \circ b = a \circ (b^{-1})^{-1} \in N$. ■

?

Kontrolní otázka 24.2. *Kolik existuje různých latinských čtverců pro tříprvkovou množinu?*

?

Kontrolní otázka 24.3. *Kolik existuje různých tříprvkových grup?*

Odpovědi na některé kontrolní otázky

Odpověď na kontrolní otázku 24.1. Předpokládejme, že e_H je neutrální prvek podgrupy H. Platí tedy $e_H \circ e_H = e_H$. Nyní přenásobme tuto rovnost zleva prvkem $e_H^{-1} \in G$, dostaneme

$$e_H^{-1} \circ e_H \circ e_H = e_H^{-1} \circ e_H.$$

Označíme-li e_G neutrální prvek grupy G, dostaneme

$$e_G \circ e_H = e_H = e_G.$$

Z této vlastnosti pak rovnou plyne i podobná vlasnost týkající se inverzních prvků: inverzní prvek k prvku a v podgrupě H grupy G je inverzním prvkem k prvku a v grupe G.

Odpověď na kontrolní otázku 24.2. Je jich $12 = 2 \cdot 3!$, neboť nastavení prvního řádku už nám dá jenom dvě možnosti pro druhý …

Odpověď na kontrolní otázku 24.3. Odpověď je 3, vybereme neutrální prvek, a pak už žádná volba nezbývá.

25 Řád grupy a Lagrangeova věta

Řád grupy

Definice 25.1 — Řád (*order*). Řád grupy $G = (M, \circ)$ nazýváme počet prvků množiny M. Je-li M nekonečná množina, řekneme, že její řád je nekonečno. Řád grupy G značíme $\#G$. Má-li G konečný řád, řekneme, že G je konečná (grupa), jinak nekonečná (grupa).

■ **Příklad 25.2 — pokračování.** Grupa $\mathbb{Z}_{12}^+$ je řádu 12. Existuje v ní 6 podgrup:

dvě triviální

$$\{0\} \quad \text{a} \quad \{0,1,2,3,4,5,6,7,8,9,10,11\}$$

řádů 1 a 12 a čtyři vlastní

$$\{0,6\},\ \{0,4,8\},\ \{0,3,6,9\} \quad \text{a} \quad \{0,2,4,6,8,10\}$$

řádů $2, 3, 4$ a 6. ■

Lagrangeova věta

Věta 25.3 — Lagrangeova. Buď H podgrupa konečné grupy G. Potom řád H je dělitelem řádu G.

Důkaz. Myšlenka důkazu:

1. V grupě G definujeme relaci: $x \sim y$ pokud existuje $h \in H$ tak, že $x = yh$.
2. Ukážeme, že se jedná o ekvivalenci, tedy o reflexivní, symetrickou a tranzitivní relaci.
3. Jelikož se jedná o ekvivalenci, tvoří prvky G, které jsou v relaci $\sim$, třídy ekvivalence (viz např. předmět BI-DML).
4. Jednou z těchto tříd ekvivalence je podgrupa H (plyne z uzavřenosti H vůči binární operaci).
5. Ukážeme-li, že všechny třídy ekvivalencí mají stejný počet prvků, je důkaz hotov, neb třídy ekvivalencí (jak známo) tvoří disjunktní rozklad množiny prvků grupy G.
6. Buďte $[a]_\sim$ a $[b]_\sim$ dvě různé třídy ekvivalence s reprezentanty a a b. Definujeme zobrazení z $[a]_\sim$ do $[b]_\sim$ takto:
$$f(x) = ba^{-1}x\,.$$
7. Jelikož toto zobrazení má inverzi (viz vlastnost „v grupě lze jednoznačně dělit“)
$$f^{-1}(y) = ab^{-1}y\,,$$
jedná se o bijekci a tedy $[a]_\sim$ a $[b]_\sim$ jsou skutečně stejně mohutné (mají stejně prvků).

■

- Tato věta spojuje abstraktní strukturu grupy s pojmem dělitelnosti a tedy i s pojmem prvočísla.
- **Důsledek**: Grupa s prvočíselným řádem má pouze triviální podgrupy.
- Například grupa $\mathbb{Z}_{11}^+$ má pouze dvě podgrupy $\{0\}$ a sebe samu.

Kontrolní otázka 25.1. *Buď G grupa řádu n a $k \in \mathbb{N}$ takové, že k dělí n. Může nastat situace, že v G neexistuje podgrupa řádu k?*

Pro zvídavé: jak je to tedy s dalšími podgrupami?

Věta 25.4 — Sylowova věta. Buď G grupa konečného řádu n a číslo p prvočíselný dělitel čísla n. Pokud p^k dělí n (pro k kladné celé), pak grupa G obsahuje podgrupu řádu p^k.

(Pro $k = 1$ též Cauchyho věta.)

26 Generující množiny a generátory grup

Grupa generovaná množinou (1 ze 3)

Připomenutí: V lineární algebře hraje důležitou roli báze vektorového prostoru. V teorii grup má podobný význam generující množina, resp. generátor. Při zafixované grupě G a neprázdné množině N, $N \subseteq G$, zavádíme toto značení:

$$\langle N \rangle := \bigcap \{H : H \text{ je podgrupa grupy } G \text{ obsahující } N\}.$$

Věta 26.1 Buď $G = (M, \circ)$ grupa a $N \subseteq M$ neprázdná množina. Množina $\langle N \rangle$ je podgrupou grupy G obsahující množinu N.

Důkaz. $\langle N \rangle$ je grupa dle předchozí věty o průniku grup.

Každá z H obsahuje množinu N a proto i průnik všech těchto H, tj. $\langle N \rangle$, obsahuje N. ■

Grupa generovaná množinou (2 ze 3)

Definice 26.2 Podgrupu $\langle N \rangle$ grupy G pro neprázdnou N, $N \subseteq M$, nazýváme podgrupou generovanou množinou N. Množinu N pak nazýváme generující množinou grupy $\langle N \rangle$.

Speciálně pro jednoprvkovou množinu $N = \{a\}$ zavádíme značení $\langle a \rangle := \langle \{a\} \rangle$. V tomto případě o a mluvíme jako o generátoru grupy $\langle a \rangle$.

Poznámka: $\langle N \rangle$ je nejmenší podgrupa grupy G obsahující množinu N. Z kontextu musí být čtenáři jasné, o jaké grupě se bavíme (v notaci $\langle N \rangle$ je to potlačeno).

■ **Příklad 26.3** V grupě $\mathbb{Z}_{12}^+$ jsme ukázali, že $\langle 2 \rangle = \big(\{0, 2, 4, 6, 8, 10\}, +_{12}\big)$. ■

■ **Příklad 26.4** Grupa $\mathbb{Z}_{12}^+$ je generována např. množinami $\{1\}$ a $\{5\}$, tzn.

$$\langle 1 \rangle = \langle 5 \rangle = \mathbb{Z}_{12}^+.$$

Ekvivalentně řečeno, prvky 1 i 5 jsou generátory $\mathbb{Z}_{12}^+$. ■

Mocnina prvku (připomenutí)

V grupě $G = (M, \circ)$ s neutrálním prvkem e definujeme pro každý prvek $g \in M$ a kladné $n \in \mathbb{N}$ n-tou mocninu a $(-n)$-tou mocninu prvku g takto:

$$\begin{aligned} g^0 &= e \\ g^n &= \underbrace{g \circ g \circ \cdots \circ g}_{n \text{ krát}} \\ g^n \circ g^{-n} &= e \end{aligned}$$

- Výše uvedené značení používáme i v grupě $(M, \cdot)$ s multiplikativní notací.
- Pro aditivní zápis grupy $G = (M, +)$ se používá n-tý násobek prvku g a značí se $n \times g$ resp. $-n \times g = n \times (-g)$.
- Uvědomte si, že $g \circ g \circ \cdots \circ g$ můžeme psát bez závorek díky asociativitě.
- Pro všechna $n, m \in \mathbb{Z}$ platí zažité $g^{n+m} = g^n \circ g^m$ a $g^{nm} = (g^n)^m$.

Grupa generovaná množinou (3 ze 3)

Věta 26.5 Buď $G = (M, \circ)$ grupa a $N \subseteq M$ neprázdná množina. Potom všechny prvky patřící do $\langle N \rangle$ lze získat pomocí „grupového obalu“

$$\langle N \rangle = \left\{ a_1^{k_1} \circ a_2^{k_2} \circ \cdots \circ a_n^{k_n} \ : \ n \in \mathbb{N},\, k_i \in \mathbb{Z},\, a_i \in N \right\}.$$

Důkaz. Označme $K = \left\{ a_1^{k_1} \circ a_2^{k_2} \circ \cdots \circ a_n^{k_n} \ : \ n \in \mathbb{N},\, k_i \in \mathbb{Z},\, a_i \in N \right\}$. Je-li H podgrupa obsahující množinu N, pak nutně $K \subseteq H$. Platí tedy $K \subseteq \bigcap \{H : H \text{ je podgrupa grupy } G \text{ obsahující } N\} = \langle N \rangle$.

Dále stačí dokázat, že K je také podgrupa (např. pomocí věty 24.4 nebo z definice). Jelikož $N \subseteq K$, pak $\langle N \rangle \subseteq K$. Celkem tedy $K = \langle N \rangle$. ■

Důsledek: $\langle a \rangle = \{a^k \ : \ k \in \mathbb{Z}\}$.

27 Cyklické grupy

Příklad: Jak nagenerovat $\mathbb{Z}_n^+$?

Poznámka: Viděli jsme, že grupa $\mathbb{Z}_{12}^+$ je rovna $\langle 1 \rangle$, $\langle 5 \rangle$, $\langle 7 \rangle$ a $\langle 11 \rangle$. **Snadné pozorování**: platí $\langle 1 \rangle = \mathbb{Z}_n^+$. Ovšem v $\mathbb{Z}_n^\times$ platí $\langle 1 \rangle = \{1\}$.

Věta 27.1 Grupa $\mathbb{Z}_n^+$ je rovna $\langle k \rangle$, $k \in \mathbb{Z}_n^+$, tehdy, a jen tehdy, když k a n jsou nesoudělná čísla.

Tato věta bude důsledkem obecné věty, kterou si dokážeme později, a faktu, že $\langle 1 \rangle = \mathbb{Z}_n^+$ pro všechna $n \geq 2$.

Příklad: Jak nagenerovat $\mathbb{Z}_n^\times$?

■ **Příklad 27.2** V grupě $\mathbb{Z}_{11}^\times$ platí

$2^1 = 2$	$2^3 = 8$	$2^5 = 10$	$2^7 = 7$	$2^9 = 6$
$2^2 = 4$	$2^4 = 5$	$2^6 = 9$	$2^8 = 3$	$2^{10} = 1$

a proto $\langle 2 \rangle = \mathbb{Z}_{11}^\times$, tj. 2 je její generátor. ■

■ **Příklad 27.3** V grupě $\mathbb{Z}_8^\times$ (její nosič je $\{1, 3, 5, 7\}$) platí

$$3^2 = 1 \qquad 5^2 = 1 \qquad 7^2 = 1$$

a proto tato grupa nemá jednoprvkovou generující množinu, nemá generátor.

Na druhou stranu ale platí $\langle \{3, 5\} \rangle = \mathbb{Z}_8^\times$. Což je pořád lepší než takřka triviální $\langle \{3, 5, 7\} \rangle = \mathbb{Z}_8^\times$. ■

K problému existence generátorů grup $\mathbb{Z}_n^\times$ se vrátíme později v přednášce.

Definice cyklické grupy

Ne každá grupa má jednoprvkovou generující množinu (generátor). Dává proto smysl zavést následující pojem, jehož název, jak uvidíme později, odkazuje na strukturu takovýchto grup.

Definice 27.4 — Cyklická grupa (*cyclic group*). Grupa $G = (M, \circ)$ se nazývá cyklická, pokud existuje prvek $a \in M$ takový, že $\langle a \rangle = G$. Tomuto prvku se říká generátor cyklické grupy G.

Na předchozích příkladech jsme si ukázali:

- $\mathbb{Z}_n^+$ jsou cyklické grupy pro všechna n a generátorem jsou všechna kladná $k \leq n$ nesoudělná s n. Speciálně, číslo 1 je generátor každé $\mathbb{Z}_n^+$.
- Dokonce i grupa $(\mathbb{Z}, +)$ je cyklická: má právě dva generátory 1 a -1. Tj. $(\mathbb{Z}, +) = \langle 1 \rangle = \langle -1 \rangle$. Rozmyslete! Toto je častá problematická státnicová otázka!
- $\mathbb{Z}_{11}^\times$ je také cyklická, s generátorem 2. $\mathbb{Z}_8^\times$ není cyklická.

Řád prvku

Definice 27.5 Buď g prvek grupy G. Pokud existuje kladné celé číslo m splňující $g^m = e$, pak nejmenší m s touto vlastností nazýváme řádem prvku g. Pokud takové m neexistuje, pak řekneme, že řád prvku g je nekonečno. Řád prvku g značíme $\mathrm{ord}(g)$.

Poznámka: Řád prvku g je roven řádu grupy $\langle g \rangle$, platí tedy rovnost

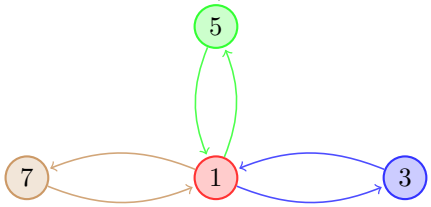
$$\mathrm{ord}(g) = \#\langle g \rangle.$$

Dále platí: nechť $k \in \mathbb{Z}$, pak $g^k = e \iff k = \ell \cdot \mathrm{ord}(g)$ pro nějaké celé ℓ.

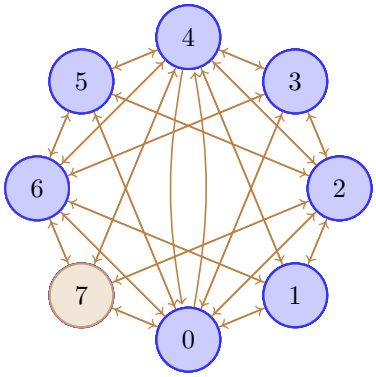
Kdy je $\mathbb{Z}_n^\times$ cyklická?

Věta 27.6 $\mathbb{Z}_n^\times$ je cyklická, právě když n je 2, 4, p^k, nebo $2p^k$, kde p je liché prvočíslo a k je kladné celé číslo.

Například $\mathbb{Z}_8^\times$ není cyklická, jak je pěkně vidět z jejího Cayleyho grafu (barva hrany udává jakým prvkem se násobí a odpovídá barvě vrcholů).



Proč „cyklická“?



Uvažujme aditivní grupu $\mathbb{Z}_8^+$.

$\langle 1 \rangle = \mathbb{Z}_8^+$, 1 je generátor $\mathbb{Z}_8^+$.

$\langle 2 \rangle = \{0, 2, 4, 6\}$, 2 není generátor $\mathbb{Z}_8^+$.

$\langle 3 \rangle = \mathbb{Z}_8^+$, 3 je generátor $\mathbb{Z}_8^+$.

$\langle 4 \rangle = \{0, 4\}$, 4 není generátor $\mathbb{Z}_8^+$.

$\langle 5 \rangle = \mathbb{Z}_8^+$, 5 je generátor $\mathbb{Z}_8^+$.

$\langle 6 \rangle = \{0, 2, 4, 6\}$, 6 není generátor $\mathbb{Z}_8^+$.

$\langle 7 \rangle = \mathbb{Z}_8^+$, 7 je generátor $\mathbb{Z}_8^+$.

Jak najít všechny generátory (1 ze 3)

Obecně najít generátory není úplně jednoduchý úkol (např. v grupách $\mathbb{Z}_p^\times$ to moc neumíme), ale pokud už jeden najdeme, je snadné najít i ty ostatní.

Věta 27.7 Je-li $(G, \circ)$ cyklická grupa řádu n a a nějaký její generátor, potom a^k je také generátor tehdy, a jen tehdy, když k a n jsou nesoudělná (tj. $\gcd(k, n) = 1$). !

Důkaz. ($\Rightarrow$) Jelikož $\langle a^k \rangle = G$, pak existuje $u \in Z$ tak, že $(a^k)^u = a$. Tedy $a^{uk-1} = e$. Prvek a je generátor G, tedy existuje $v \in \mathbb{Z}$ tak, že $uk - 1 = v \cdot \mathrm{ord}(g)$. Protože $\mathrm{ord}(g) = n$, dostaneme $1 = uk - vn$ a to již implikuje nesoudělnost, neboť:

Pomocné lemma: Buď $D = \{mk + \ell n \mid m, l \in \mathbb{Z}\}$ a $d = \min\{|x| \mid x \in D \setminus \{0\}\}$, potom $d = \gcd(k, n)$.
Důkaz tohoto lemmatu:

...pokračuje ... ■

Jak najít všechny generátory (2 ze 3)

Důkaz: pokračování.

- d dělí n: kdyby ne, je $n = jd + r$ pro $0 < r < d$, ale pak $r = n - jd \in D$ je menší než d, spor.
- d dělí k: stejně
- d' dělí k a n, pak ale dělí i d, neb d je celočíselná lin. kombinace těchto dvou čísel. A tedy $d' \leq d$.

(konec důkazu lemmatu)

($\Leftarrow$): $\gcd(k, n) = 1$ a tedy existují u a v tak, že $un + vk = 1$ a tedy

$$a = a^{un+vk} = a^{un} a^{vk} = a^{u\,\mathrm{ord}(g)} a^{vk} = a^{vk} = (a^k)^v.$$

Z toho už plyne, že a^k je generátor, neb jím jde vygenerovat jiný generátor.

Vskutku: mějme $b \in G$, hledejme $\ell \in \mathbb{Z}$ takové, že $(a^k)^\ell = b$. Protože $\langle a \rangle = G$, existuje $t \in \mathbb{Z}$ takové, že $b = a^t$. Tedy $b = (a^t)^{vk} = (a^k)^{vt}$ a tedy $\ell = vt$. ■

Jak najít všechny generátory (3 ze 3)

Důsledek 27.8 V cyklické grupě řádu n je počet generátorů roven $\varphi(n)$.

- φ je Eulerova funkce, která každému kladnému celému číslu n přiřazuje počet kladných celých čísel menších než n, která jsou s ním nesoudělná,
- Pro prvočíslo p je $\mathbb{Z}_p^\times$ cyklická grupa řádu $p - 1$ a má tedy $\varphi(p - 1)$ generátorů.
- Není znám efektivní algoritmus pro výpočet $\varphi(n)$.

Podgrupy cyklické grupy jsou cyklické

Věta 27.9 Libovolná podgrupa cyklické grupy je opět cyklická grupa. !

Důkaz.
- Buď $G = \langle a \rangle$ cyklická grupa a H její vlastní podgrupa. Buď q nejmenší kladné celé číslo takové, že $a^q \in H$. Ukážeme, že $H = \langle a^q \rangle$ a tím bude důkaz hotov.
- Jistě platí, že $\langle a^q \rangle \subseteq H$. Dokážeme-li, že platí $H \subseteq \langle a^q \rangle$, pak se musejí tyto množiny rovnat.
- Buď x nějaký prvek H a p takové, že $x = a^p$. Existují celá čísla u, v tak, že $d = \gcd(q, p) = uq + vp$. Potom $a^d = (a^q)^u \circ (a^p)^v \in H$ a tedy $d \geq q$, Současně $\gcd(q, p) = d \leq q$, proto $d = q$ a existuje k tak, že $p = kq$. Odtud konečně dostáváme $x = a^p = (a^q)^k \in \langle a^q \rangle$ a důkaz je hotov. ■

28 (Malá) Fermatova věta

Malá Fermatova věta (1 ze 2)

Důsledkem Lagrangeovy věty je:

Věta 28.1 V grupě $G = (M, \circ)$ řádu n platí pro všechny prvky $a \in M$, že !

$$a^n = e, \quad \text{kde } e \text{ je neutrální prvek.}$$

Důkaz.
- Cyklická grupa $\langle a \rangle$ je podgrupou grupy G.
- Dle Lagrangeovy věty tedy platí, že $\#\langle a \rangle$ dělí n, tzn. že existuje $k \in \mathbb{N}$ takové, že $n = k \cdot \#\langle a \rangle$.
- Máme tedy $a^n = a^{k \cdot \#\langle a \rangle} = (a^{\#\langle a \rangle})^k = (a^{\operatorname{ord} a})^k = e^k = e$. ■

Malá Fermatova věta (2 ze 2)

- Grupa $\mathbb{Z}_p^\times$ má řád $p - 1$.
- Aplikováním předchozí věty na tuto grupu získáme malou Fermatovu větu.

Důsledek 28.2 — Malá Fermatova věta. Pro libovolné prvočíslo p a libovolné $1 \leq a < p$ platí

$$a^{p-1} \equiv 1 \pmod{p}.$$

Odpovědi na některé kontrolní otázky

Odpověď na kontrolní otázku 25.1. Odpověď je ano, může, ale není úplně triviální takovou grupu najít. Poznamenejme, že pokud k je prvočíslo, pak zmíněna podgrupa existuje.

29 Homomorfismy a izomorfismy

29.1 Motivační příklad

Různé grupy – stejná struktura (1 z 6)

$\mathbb{Z}_5^\times$	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

řád: 4

podgrupy: $\{1\}$, $\{1,4\}$, $\{1,2,3,4\}$

neutrální prvek: 1

inverze: $1^{-1}=1$, $2^{-1}=3$, $3^{-1}=2$, $4^{-1}=4$

$\mathbb{Z}_4^+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

řád: 4

podgrupy: $\{0\}$, $\{0,2\}$, $\{0,1,2,3\}$

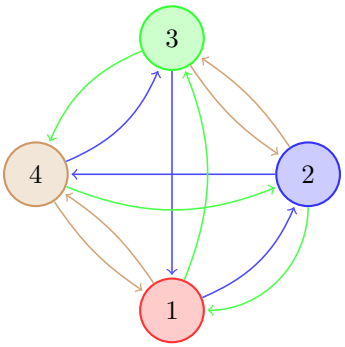
neutrální prvek: 0

inverze: $0^{-1}=0$, $1^{-1}=3$, $2^{-1}=2$, $3^{-1}=1$

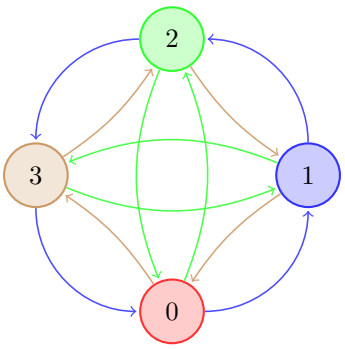
Nejsou $\mathbb{Z}_4^+$ a $\mathbb{Z}_5^\times$ vlastně stejné grupy lišící se pouze ve „jménech“ svých prvků a označení operace?

Různé grupy – stejná struktura (2 z 6)

$\mathbb{Z}_5^\times$	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1



$\mathbb{Z}_4^+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2



Různé grupy – stejná struktura (3 z 6)

$\mathbb{Z}_5^\times$	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

$\mathbb{Z}_4^+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Zkusme přejmenovat prvky grupy $\mathbb{Z}_5^\times$ tak, abychom dostali $\mathbb{Z}_4^+$:

- neutrální prvek má velmi speciální a jedinečné vlastnosti, proto přejmenujme 1 na 0,
- pokud se má zachovat kompletně struktura, musí jediné dvouprvkové podgrupě $\{1,4\}$ (v $\mathbb{Z}_5^\times$) odpovídat podgrupa $\{0,2\}$ (v $\mathbb{Z}_4^+$), proto $4 \leftrightarrow 2$,
- nyní už stačí přejmenovat 2 a 3: zjistíme, že obě zbývající možnosti fungují, zvolme tedy např. $3 \leftrightarrow 1$ a $2 \leftrightarrow 3$,
- a nyní už stačí přeházet řádky …a máme Cayleyho tabulku $\mathbb{Z}_4^+$.

Různé grupy – stejná struktura (4 z 6)

- Našli jsme způsob, jak přeznačit prvky v jedné tabulce, abychom dostali přesně tabulku druhou (po přeházení řádků a sloupců).

- Toto přejmenování je vlastně **prosté** zobrazení množiny $\{1,2,3,4\}$ **na** množinu $\{0,1,2,3\}$, označme jej h_1:

$$h_1(1) = 0, \quad h_1(2) = 3, \quad h_1(3) = 1, \quad h_1(4) = 2\,.$$

- Jak jsme naznačili, fungovalo by i h_2 (jen bychom museli na závěr jinak zpřeházet řádky a sloupce):

$$h_2(1) = 0, \quad h_2(2) = 1, \quad h_2(3) = 3, \quad h_2(4) = 2\,.$$

Nefungovaly by tedy všechny bijekce? A jestli ne, tak čím jsou tyto dvě výjimečné?

Různé grupy – stejná struktura (5 z 6)

Přejmenujme prvky grupy $\mathbb{Z}_5^\times$ podle bijekce h_3:

$$h_3(1) = 0, \quad h_3(2) = 3, \quad h_3(3) = 2, \quad h_3(4) = 1\,.$$

$\mathbb{Z}_5^\times$	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

$\mathbb{Z}_4^+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

- Výsledná tabulka ale neodpovídá grupě $\mathbb{Z}_4^+$ s operací sčítání modulo 4, neboť například $3+3 \pmod 4 \neq 1$.

- Bijekce h_3 tedy nevytvoří stejnou strukturu jakou má grupa $\mathbb{Z}_4^+$, takovou vlastnost mají pouze h_1 a h_2.

Různé grupy – stejná struktura (6 z 6)

Hledaná vlastnost bijekce h, kterou mají pouze bijekce h_1 a h_2, je tato:

$$\text{pro všechna } n, m \in \{1,2,3,4\} \text{ platí} \quad h(n \times_5 m) = h(n) +_4 h(m),$$

kde $\times_5$ značí operaci v grupě $\mathbb{Z}_5^\times$ a $+_4$ v grupě $\mathbb{Z}_4^+$.

Slovy: Jestliže na libovolné prvky v grupě $\mathbb{Z}_5^\times$ aplikujeme operaci $\times_5$ a pak je zobrazíme do $\mathbb{Z}_4^+$ pomocí h, dostaneme vždy stejný výsledek, jako kdybychom je nejdříve pomocí h zobrazili do $\mathbb{Z}_4^+$ a **potom** aplikovali operaci $+_4$.

$$\begin{array}{ccc} n, m \in \mathbb{Z}_5^\times & \xrightarrow{\times_5} & n \times_5 m \in \mathbb{Z}_5^\times \\ \Big\downarrow h & & \Big\downarrow h \\ h(n), h(m) \in \mathbb{Z}_4^+ & \xrightarrow{+_4} & h(n) +_4 h(m) = h(n \times_5 m) \end{array}$$

Bijekce navíc musí tzv. zachovávat operaci. Vzpomeňte na lineární zobrazení…

29.2 Definice a vlastnosti

Homomorfismus a izomorfismus

Definice 29.1 Buďte $G = (M, \circ_G)$ a $H = (N, \circ_H)$ dva grupoidy. Zobrazení $h : M \to N$ nazveme homomorfismem G do H jestliže

$$\text{pro všechna } x, y \in M \text{ platí} \quad h(x \circ_G y) = h(x) \circ_H h(y).$$

Je-li navíc h injektivní, resp. surjektivní, resp. bijektivní, říkáme že h je monomorfismus, resp. epimorfismus, resp. izomorfismus.

- Homomorfismus tedy zachovává strukturu danou binární operací: je jedno jestli nejdříve aplikuji operaci a pak homomorfismus, nebo naopak.

- Jediné, co potřebujeme pro definování této vlastnosti, je uzavřenost množiny vůči binární operaci, proto jsme homomorfismus definovali pro nejobecnější grupoidy.

- Definice se přímo přenáší na grupy, a používá se termín (homo|mono|epi|izo)morfismus grup.

Řecké jazykové okénko

- morfismus: z řecké slova *morfé*, znamenajícího forma, tvar
- homo: *homós*, stejný,
- izo: *ísos*, sobě rovný,
- epi: *epí*, na,
- mono: *monós*, samotný, jediný.

Izomorfní grupy

Definice 29.2 Grupy G a H nazýváme izomorfní, právě když existuje izomorfismus $G \to H$. O grupě G také říkáme, že je izomorfní s grupou H.

- Vlastnost dvou grup „být izomorfní“ je relace ekvivalence na třídě všech grup.
- Příkladem izomorfních grup jsou $\mathbb{Z}_5^\times$ a $\mathbb{Z}_4^+$: našli jsme dokonce dva různé izomorfismy h_1 a h_2.
- Je jasné, že izomorfní grupy musí mít stejný řád!

Základní vlastnosti homomorfismu (1 ze 2)

Věta 29.3 Buď h homomorfismus grupy $G = (M, \circ_G)$ do grupoidu $H = (N, \circ_H)$. Potom $h(G) := (h(M), \circ_H)$ je grupa.

Důkaz. Ukážeme postupně že v $h(G)$ platí asoc. zákon, existuje neutrální prvek a každý prvek má inverzi.

- Každý prvek $h(G)$ lze napsat jako $h(x)$ pro nějaké vhodné x.
- Pro všechna $x, y, z \in M$ platí
$$(h(x) \circ_H h(y)) \circ_H h(z) = h(x \circ_G y) \circ_H h(z) = h((x \circ_G y) \circ_G z) = \\ = h(x \circ_G (y \circ_G z)) = h(x) \circ_H (h(y) \circ_H h(z))$$
- Označme e_G neutr. prvek v G, potom $h(e_G)$ je neutrální prvek v $h(G)$, neboť pro všechna x platí $h(e_G) \circ_H h(x) = h(e_G \circ_G x) = h(x)$.
- Podobně se ukáže, že inverzí k $h(x)$ je $h(x^{-1})$. ∎

Základní vlastnosti homomorfismu (2 ze 2)

Je-li H grupa, tak předchozí věta a její důkaz mají následující důsledky:

- Neutrální prvek jedné grupy se homomorfismem zobrazí vždy na neutrální prvek té druhé grupy.
- Také inverze se zachovávají v následujícím smyslu: $h(x^{-1}) = h(x)^{-1}$.
- Je-li h homomorfismus grupy G do H, pak $h(G)$ je podgrupa v H.
- Např. $h(n) = 2n$ je homomorfismus grupy $\mathbb{Z}_4^+$ do $\mathbb{Z}_8^+$ a $h(\mathbb{Z}_4^+)$ je podgrupa $\{0, 2, 4, 6\}$.

...až na izomorfismus (1 ze 4)

Izomorfní grupy jsou vlastně totožné, liší se pouze pojmenováním prvků, jak jsme viděli v případě grup $\mathbb{Z}_4^+$ a $\mathbb{Z}_5^\times$. Řekneme-li, že existuje pouze jedna grupa s jistou vlastností až na izomorfismus, znamená to, že všechny grupy s touto jistou vlastností jsou navzájem izomorfní. Ukážeme si tři známá tvrzení tohoto typu.

Věta 29.4 Libovolné dvě cyklické grupy mající stejný řád jsou izomorfní. !

Důkaz: náznak – doladit za domácí úkol. Buď $G = \langle a \rangle$ cyklická grupa s generátorem a. Ukážeme, že libovolná nekonečná cyklická grupa je izomorfní s grupou $(\mathbb{Z}, +)$ a že libovolná cyklická grupa řádu n je izomorfní s $\mathbb{Z}_n^+$. Zbytek už plyne z tranzitivity relace „být izomorfní“. Hledaný izomorfismus je bijekce (ze $\mathbb{Z}$ či $\mathbb{Z}_n^+$ na G) definovaná pro všechna k jako $h(k) = a^k$. ■

$(\mathbb{Z}, +)$ a $\mathbb{Z}_n^+$ jsou tedy jedinými cyklickými grupami až na izomorfismus.

...až na izomorfismus (2 ze 4)

■ **Příklad 29.5 — Kleinova grupa.** Kleinova grupa je grupa $(\mathbb{Z}_2 \times \mathbb{Z}_2, \circ)$, kde

$$\mathbb{Z}_2 \times \mathbb{Z}_2 = \{(0,0), (0,1), (1,0), (1,1)\}$$

a $\circ$ je sčítání modulo 2 po složkách: např. $(1,0) \circ (1,1) = (0,1)$. ■

Kleinova grupa není cyklická a tedy nemůže být izomorfní se $\mathbb{Z}_4^+$! Kleinova grupa je izomorfní $\mathbb{Z}_8^\times$. Lze dokonce ukázat toto (zkuste si to, je to jednoduché):

Věta 29.6 Existují pouze dvě neizomorfní grupy řádu 4. !

$\mathbb{Z}_4^+$ a Kleinova grupa jsou tedy až na izomorfismus jediné grupy řádu 4.

...až na izomorfismus (3 ze 4)

■ **Příklad 29.7 — Symetrická grupa.** Symetrickou grupou množiny $\{1, 2, 3, \ldots, n\}$ nazveme množinu všech permutací této množiny s operací skládání zobrazení a značíme ji S_n. ■

- Permutace je bijekce z množiny do stejné množiny, tedy v našem případě z $\{1, 2, 3, \ldots, n\}$ do $\{1, 2, 3, \ldots, n\}$.
- Každou permutaci $\pi \in S_n$ můžeme zadat výčtem hodnot:

$$\begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \pi(1) & \pi(2) & \pi(3) & \cdots & \pi(n) \end{pmatrix},$$

  první řádek navíc můžeme vynechat, takže např. $(1\ 2\ 4\ 3\ 5) \in S_5$ je permutace prohazující 3. a 4. prvek.
- Složením permutací $(1\ 2\ 4\ 3\ 5) \circ (2\ 1\ 3\ 5\ 4)$ je $(2\ 1\ 4\ 5\ 3)$.
- Skládání zobrazení je asociativní, permutace $(1\ 2\ 3\ \cdots n)$ je neutrální prvek a inverzním prvkem je inverzní zobrazení – jedná se tedy skutečně o grupu. Řád S_n je $n!$.

...až na izomorfismus (4 ze 4)

Podgrupy symetrické grupy S_n nazýváme grupami permutací.

- Například permutace $(1\ 2\ 4\ 3\ 5) \in S_5$ prohazující 3. a 4.prvek generuje podgrupu grupy S_5 obsahující dva prvky

$$(1\ 2\ 4\ 3\ 5) \quad \text{a} \quad (1\ 2\ 3\ 4\ 5)\,.$$

- Struktura podgrup S_n je velice (v jistém slova smyslu maximálně) bohatá, o čemž svědčí následující věta.

Věta 29.8 — Cayleyova. Libovolná konečná grupa je izomorfní s nějakou grupou permutací. !

Důkaz: náznak pro zájemce. Buď a prvek grupy G řádu n s binární operací $\circ$. Definujeme $\pi_a(x) = a \circ x$. Jelikož lze v grupě jednoznačně dělit, je π_a bijekce a tedy permutace! Hledaný monomorfismus (tj. izomorfismus s podgrupou S_n) je zobrazení definované pro každý prvek a takto: $h(a) = \pi_a$ ... ■

30 Aplikace teorie grup v kryptografii

30.1 Problém diskrétního logaritmu

Diskrétní logaritmus obecně

Problém diskrétního logaritmu můžeme definovat v libovolné cyklické grupě:

Definice 30.1 — problém diskrétního logaritmu v grupě $G = (M, \cdot)$. Buď $G = (M, \cdot)$ cyklická grupa řádu n, α nějaký její generátor a β její prvek. Řešit problém diskrétního logaritmu znamená najít celé číslo $1 \leq k \leq n$ takové, že

$$\alpha^k = \beta.$$

Použijeme-li aditivní značení:

Definice 30.2 — problém diskrétního logaritmu v grupě $G = (M, +)$. Buď $G = (M, +)$ cyklická grupa řádu n, α nějaký její generátor a β její prvek. Řešit problém diskrétního logaritmu znamená najít celé číslo $1 \leq k \leq n$ takové, že

$$k \times \alpha = \beta.$$

Poznámka: Zahodíme-li požadavek na cykličnost grupy G, pak má úloha diskrétního logaritmu řešení pouze pokud β patří do $\langle\alpha\rangle$, cyklické podgrupy generované α.

Ne ve všech grupách je to těžké

Uvažujme grupu $\mathbb{Z}_p^+$. To je cyklická grupa prvočíselného řádu p, a každé kladné $\alpha \leq p-1$ je její generátor. Problém diskrétního logaritmu v této grupě má formu rovnice

$$k\alpha \equiv \beta \pmod{p}.$$

Tu umíme snadno vyřešit. Najdeme inverzi α^{-1} k α v grupě $\mathbb{Z}_p^\times$ (pomocí polynomiálního `EEA` (v délce vstupu), viz další přednášky a dřívější studium) a řešením je

$$k = \beta\alpha^{-1} \bmod p.$$

Alice	**Bob**
zvolí soukromý klíč $a \in \{2, \ldots, p-2\}$	zvolí soukromý klíč $b \in \{2, \ldots, p-2\}$
spočte veřejný klíč $A \equiv \alpha^a \bmod p$	spočte veřejný klíč $B \equiv \alpha^b \bmod p$
	výměna veřejných klíčů A a B $\longleftrightarrow$
spočítá $k_{AB} \equiv B^a \bmod p$	spočítá $k_{AB} \equiv A^b \bmod p$

■ **Příklad 30.3** Uvažujme $p = 11, \alpha = 3, \beta = 5$. Hledáme k tak, aby

$$k \cdot 3 \equiv 5 \pmod{11}.$$

Snadno ověříme, že v $\mathbb{Z}_{11}^\times$ je $3^{-1} = 4$ a tedy $k = (5 \cdot 4) \bmod 11 = 9$. ■

Výpočet diskrétního logaritmu obecně?

Obecně není známý žádný rozumně rychlý algoritmus řešící problém diskrétního logaritmu.

V případě grupy $\mathbb{Z}_p^\times$ je počet kroků známých algoritmů úměrný $\sqrt{p}$, což pro p délky 1024 bitů dává cca 2^{512} operací. (Obecně se je počet kroků úměrný $\sqrt{n}$, kde n je řád základu logaritmu.)

Inverzní operaci k logaritmu, tedy mocnění, umíme v $\mathbb{Z}_p^\times$ rychle.

Získáváme tedy jednosměrnou (*one-way*) funkci, kterou lze použít pro **asymetrickou šifru**: najít $\beta \equiv \alpha^x \pmod p$ je lehké, známe-li x, α a p, najít x, známe-li β, α a p je **obecně** velmi obtížné

Poznámka: Pro konstrukci RSA byla použita jednosměrná funkce „násobení prvočísel“: násobit prvočísla je lehké a rychlé, hledat prvočíselný rozklad výsledku je **obecně** složité.

30.2 Diffie-Hellman Key Exchange

Diffie-Hellman Key Exchange

Inicializace: Alice si najde velké prvočíslo p a nějaký generátor α grupy $\mathbb{Z}_p^\times$. **Zveřejní p a α.** (najít velké prvočíslo a generátor nejsou lehké úkoly!)

Princip

Diffie-Hellman Key Exchange stojí na následujících faktech:

- Mocnění v $\mathbb{Z}_p^\times$ je komutativní a tedy vypočtené k_{AB} je pro Alici i Boba stejné:

$$\begin{aligned} k_{AB} &\equiv (\alpha^b)^a \equiv \alpha^{ab} \mod p \\ k_{AB} &\equiv (\alpha^a)^b \equiv \alpha^{ab} \mod p, \end{aligned}$$

- mocnění není výpočetně náročné (square & multiply),
- inverzní operace k mocnění, tedy diskrétní logaritmus, je výpočetně velmi náročná.

?

Kontrolní otázka 30.1. *Víme, že grupy $\mathbb{Z}_p^\times$ a $\mathbb{Z}_{p-1}^+$ jsou izomorfní a tedy vlastně totožné. Není to problém pro Diffie-Hellmana, nedělá to z řešení diskrétního logaritmu v $\mathbb{Z}_p^\times$ lehký problém?*

Doteď jsem se zabývali strukturami, které vzniknou přidáním jedné binární operace k neprázdné množině. Jako grupu jsme definovali takovou strukturu, kde má daná operace něco jako svou inverzi, což je analogie k tomu co známe z klasických množin čísel: odečítání je přičítání inverze podobně jako dělení je násobení inverzí. Ovšem abychom měli aritmetiku kompletní, potřebujeme stejně jako na (reálných) číslech jak sčítání s odečítáním, tak také násobení s dělením, abychom mohli definovat například už tak základní pojem jako je polynom. Proto se v této přednášce, po krátkém výletě po eliptických křivkách, budeme věnovat právě strukturám se dvěma binárními operacemi, zejména okruhům a tělesům, které jsou právě zobecněním reálných čísel s dobře definovaným sčítáním a násobením i operacemi k nim obrácenými.

31 Množiny se dvěma binárními operacemi

Množiny se dvěma binárními operacemi

- Doposud jsme se zabývali množinami vybavenými jednou binární operací:
  - grupoidy,
  - pologrupy,
  - monoidy,
  - (abelovské) grupy.
- Například čísla, či matice, umíme vzájemně jak „sčítat“ tak „násobit“. Budeme proto dále uvažovat dvě binární operace. V následujících přednáškách se seznámíme s
  - **okruhy** a
  - **tělesy**.

31.1 Okruh

Definice okruhu

Definice 31.1 — Okruh (*Ring*). Buďte M neprázdná množina a $+$ a $\cdot$ binární operace na této množině. Řekneme, že trojice $R = (M, +, \cdot)$ je okruh, pokud platí:

- $(M, +)$ je **abelovská grupa**,
- $(M, \cdot)$ je **monoid**,
- platí (levý a pravý) **distributivní zákon**:

$$(\forall a, b, c \in M)\big(a \cdot (b + c) = a \cdot b + a \cdot c \quad \wedge \quad (b + c) \cdot a = b \cdot a + c \cdot a\big).$$

- Dodržujeme standardní konvenci, že násobení má vyšší prioritu než sčítání. Tím si ušetříme práci se psaním některých závorek, $a + b \cdot c$ chápeme jako $a + (b \cdot c)$. Tečku pro násobení navíc většinou ani nepíšeme, tj. $a + (b \cdot c) = a + bc$.
- Někteří autoři nevyžadují existenci neutrálního prvku v $(M, \cdot)$.

Názvosloví

Buď $R = (M, +, \cdot)$ okruh.

- Je-li $\cdot$ komutativní, je R komutativní okruh,
- $(M, +)$ se nazývá aditivní grupa okruhu R,
- $(M, \cdot)$ se nazývá multiplikativní monoid okruhu R,
- neutrální prvek grupy $(M, +)$ se nazývá nulový prvek a značí se 0, inverzní prvek vůči $+$ k $a \in M$ pak značíme $-a$,
- v okruhu **můžeme definovat odečítání** předpisem

$$a - b := a + (-b),$$

- neutrálnímu prvku multiplikativního monoidu budeme zpravidla říkat jednička a značit jej 1.

Jednoduché příklady okruhů

- triviální okruh je $(\{0\}, +, \cdot)$,
- $(\mathbb{Z}, +, \cdot)$ je okruh, (ale $(\mathbb{N}, +, \cdot)$ není okruh, neb $(\mathbb{N}, +)$ není grupa),
- množina $(\mathbb{R}^{n,n}, +, \cdot)$ čtvercových reálných matic se sčítáním po prvcích a maticovým násobením je okruh, nulový prvek je nulová matice (podobně pro komplexní matice),
- množina všech polynomů (s komplexními / reálnými / celočíselnými koeficienty) je okruh, nulový prvek je nulový polynom, tj. polynom splňující $p(x) = 0$ pro každé x,
- množina všech zbytkových tříd $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ po dělení $n \in \mathbb{N}$ se sčítáním a násobením modulo n je okruh.

Základní vlastnosti okruhu

V libovolném okruhu $(M, +, \cdot)$ platí:

- **Násobení nulovým prvkem dává opět nulový prvek**, tj.

$$(\forall a \in M)(a \cdot 0 = 0 \wedge 0 \cdot a = 0).$$

  Vskutku: $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$.

- Z toho plyne, že $(\forall a, b \in M)((-a) \cdot b = -a \cdot b)$.
- **Levý i pravý distributivní zákon pro odečítání**, tj. $\forall a, b, c \in M$

$$c(b - a) = cb - ca.$$

  Vskutku: $ca + c(b - a) = c(a + b - a) = cb \implies c(b - a) = cb - ca$.

Definice 31.2 — Obor integrity (*Integral domain*). *Nenulové* prvky $a, b \in M$ z okruhu $(M, +, \cdot)$ nazýváme dělitelé nuly, právě když $a \cdot b = b \cdot a = 0$. Obor integrity je komutativní okruh, ve kterém neexistují dělitelé nuly.

31.2 Těleso

Definice tělesa

Definice 31.3 — těleso (*Field*). Okruh $T = (M, +, \cdot)$ se nazývá těleso, jestliže $(M \setminus \{0\}, \cdot)$ je abelovská grupa. Tuto grupu nazýváme multiplikativní grupou tělesa T.

Jazyková vsuvka:

- V anglosaské literatuře narazíte na tento objekt pod názvem *field*, což by v češtině odpovídalo slovu „pole“. Pole (vektorové) má ale v české matematické terminologii již jiný význam.
- Česká terminologie se v tomto směru drží francouzké (*corps*) a německé (*Körper*). V obou případech je překladem „tělo“, či „těleso“.

Upozornění: pod pojmem těleso se někdy rozumími struktura, kde $\cdot$ není komutativní.

Proč musíme vyjmout nulový prvek v $(M \setminus \{0\}, \cdot)$?

1 je neutrální prvek $M \setminus \{0\}$, tedy v tělese vždy platí $1 \neq 0$.

Protože $a \cdot 0 = 0 \cdot a = 0$, nemůže k nule existovat inverzní prvek (vzhledem k násobení), tj. nelze dělit nulou.

Všemi jinými prvky tělesa dělit umíme!

$$\text{dělení} = \text{násobení inverzním prvkem}$$

$$\frac{a}{b} := a \cdot b^{-1} \quad \text{pro } b \neq 0.$$

Tato notace má dobrý smysl díky komutativitě operace $\cdot$.

Příklady těles

- Okruh celých čísel $(\mathbb{Z}, +, \cdot)$ není těleso, neb $(\mathbb{Z} \setminus \{0\}, \cdot)$ není grupa (chybí inverzní prvky).
- Okruh racionálních čísel $(\mathbb{Q}, +, \cdot)$ je těleso. Dokonce nejmenší číselné těleso (s obvyklými aritmetickými operacemi).
- Nejmenší těleso je tzv. triviální těleso $(\{0, 1\}, +, \cdot)$ s operacemi danými násl. tabulkami:

$+$	0	1
0	0	1
1	1	0

a

$\cdot$	0	1
0	0	0
1	0	1

První tabulka odpovídá bitové operaci `XOR` a druhá `AND`, nebo také sčítání a násobení modulo 2.

Některé vlastnosti těles

V každém tělese máme definované aritmetické operace:

sčítání, odčítání, násobení, dělení a všechny z nich odvozené, jako mocnění, odmocňování, logaritmování, ...

Triviální těleso nám tyto všechny operace definuje nad jedním bitem. Později si ukážeme, jak je rozšířit nad libovolný počet bitů.

Věta 31.4 Pokud pro a, b z tělesa T platí $ab = 0$ potom $a = 0$ nebo $b = 0$.

Důkaz. Sporem: kdyby $a \neq 0$ a $b \neq 0$ potom $ab \neq 0$, protože multiplikativní grupa $(T \setminus \{0\}, \cdot)$ je uzavřená na násobení. ■

Každé těleso je tedy oborem integrity.

Homomorfismus a izomorfismus okruhů a těles

Podobně jako u grup zavádíme homomorfismus a izomorfismus okruhů a těles.

Definice 31.5 Zobrazení h z okruhu R do okruhu S je homomorfismus těchto okruhů, jestliže je h homomorfismem z aditivní grupy R do aditivní grupy S, homomorfismem[a] z multiplikativního monoidu R do multiplikativního monoidu S a platí $h(1_R) = 1_S$.

Zobrazení h z tělesa R do tělesa S je homomorfismus těchto těles, jestliže je h homomorfismem z aditivní grupy R do aditivní grupy S a homomorfismem z multiplikativní grupy R do multiplikativní grupy S.

Je-li navíc h bijekce (prosté a „na"), jedná se o izomorfismus těchto okruhů (resp. těles).

[a] použijeme definici pro grupoid

Izomorfní tělesa

Definice 31.6 Tělesa T a K nazýváme izomorfní, právě když existuje izomorfismus z T na K. V tomto případě také říkáme, že těleso T je izomorfní s tělesem K.

Relace „být izomorfní" na třídě všech těles je relace ekvivalence.

32 Okruhy polynomů

32.1 Definice a základní vlastnosti

Polynom nad okruhem

Definice 32.1 — Polynom nad okruhem. Mějme okruh R a $a_i \in R$, $i = 0, 1, \ldots, n$. Formální výraz tvaru

$$P(x) = \sum_{i=0}^{n} a_i x^i$$

nazýváme polynomem nad okruhem R (s formální proměnnou x).

Používáme standardní názvosloví:

- a_i, $i = 0, 1, \ldots, n$, nazýváme **koeficienty** polynomu $P(x)$.

- x nazýváme **formální proměnnou** polynomu $P(x)$.
- Dva polynomy se rovnají, pokud se rovnají jejich příslušné koeficienty.
- Členy s nulovým koeficientem se často vynechávají, tedy např. $1 + 0x = 1$.
- Pokud pro polynom $P(x)$ existuje $k \in \{0, 1, \ldots, n\}$ takové, že $a_k \neq 0$, pak největší z těchto k nazýváme **stupněm polynomu** $P(x)$, značený $\deg(P(x))$.
- Polynom $P(x) = 0$ nazýváme **nulový polynom** a jeho stupeň nedefinujeme.

Okruh polynomů nad okruhem

Abychom mohli sčítat, odčítat a násobit polynomy, potřebujeme pouze vědět jak sčítat, odčítat a násobit jejich koeficienty. Obecně tedy můžeme vybudovat okruh polynomů podobný tomu, který známe z reálných resp. komplexních čísel, nad libovolným okruhem (a tedy i tělesem).

Věta 32.2 — Okruh polynomů (*polynomial ring*). Buď R okruh. Potom množina všech polynomů nad okruhem R spolu s operacemi sčítání a násobení definovanými předpisy

$$\sum_{i=0}^{n} a_i x^i + \sum_{i=0}^{n} b_i x^i := \sum_{i=0}^{n} (a_i + b_i) x^i$$

$$\left(\sum_{j=0}^{n} a_j x^j \right) \cdot \left(\sum_{k=0}^{m} b_k x^k \right) := \sum_{i=0}^{n+m} \left(\sum_{j+k=i} a_j b_k \right) x^i,$$

kde $a_i, b_i \in R$ pro všechny hodnoty i, tvoří okruh polynomů nad okruhem R. Tento okruh značíme $R[x]$.

!

Základní vlastnosti (1 ze 4)

Lemma 32.3 (o násobení polynomů)**.** *Buď T těleso a $f(x), g(x) \in T[x]$ nenulové polynomy. Platí*

$$\deg(f(x) \cdot g(x)) = \deg(f(x)) + \deg(g(x)).$$

Základní vlastnosti (2 ze 4)

Lemma 32.4 (o dělení polynomů)**.** *Buď T těleso a $f(x), g(x) \in T[x]$ nenulové polynomy. Pak existují jednoznačně určené polynomy $q(x), r(x) \in T[x]$ takové, že*

$$f(x) = q(x)g(x) + r(x),$$

kde $r(x)$ je buď nulový nebo má stupeň ostře menší než stupeň $g(x)$.

Základní vlastnosti (3 ze 4)

Buďte $f(x), g(x) \in T[x]$. Potom polynom $h(x) \in T[x]$ nazveme největší společný dělitel polynomů $f(x)$ a $g(x)$, jestliže

- $h(x)$ dělí $f(x)$ (tj. existuje $q(x) \in T[x]$, tak. že $f(x) = q(x)h(x)$),
- $h(x)$ dělí $g(x)$,
- každý polynom, který dělí $f(x)$ i $g(x)$, dělí také $h(x)$.

Tento polynom značíme $\gcd(f(x), g(x))$ (jedná se o drobné zneužití značení, protože polynom je jednoznačný až na multiplikativní konstantu).

Věta 32.5 — Bézoutova rovnost pro polynomy. Buďte $f(x)$ a $g(x)$ nenulové polynomy nad tělesem T. Pak existují polynomy $u(x), v(x) \in T[x]$ tak, že $\gcd(f(x), g(x)) = u(x)f(x) + v(x)g(x)$.

Důkaz se provádí indukcí na součet stupňů polynomů $f(x)$ a $g(x)$ (lze aplikovat i na čísla, kde jsme to dokazovali jinak) a využívá faktu (rozmyslet), že pro libovolný polynom $q(x) \in T[x]$ platí

$$\gcd(f(x), g(x)) = \gcd(f(x) - q(x)g(x), g(x)).$$

Označme stupeň $f(x)$ jako s a stupeň $g(x)$ jako t a předpokládejme bez újmy na obecnosti, že $s \geq t$ a že oba polynomy jsou monické[11]. Je-li $s + t = 0$, je $\gcd(f(x), g(x)) = 1$, a tedy můžeme zvolit $u(x) = 0$ a $v(x) = 1$.

Buď nyní $s + t = n$. Předpokládejme, že pro polynomy, jejichž součet stupňů je menší než n, tvrzení platí (= indukční předpoklad). Pokud $g(x)$ dělí $f(x)$, je $\gcd(f(x), g(x)) = g(x)$, a tedy můžeme opět zvolit $u(x) = 0$ a $v(x) = 1$. Předpokládejme tedy, že $g(x)$ nedělí $f(x)$. Potom existují polynomy $q(x), r(x) \in T[x]$ takové, že

$$f(x) = q(x)g(x) + r(x),$$

kde $r(x)$ je ostře menšího stupně než je stupeň $g(x)$. Dle indukčního předpokladu (součet stupňů $r(x)$ a $g(x)$ je menší než n) existují $\tilde{u}(x)$ a $\tilde{v}(x)$ tak, že

$$\gcd(r(x), g(x)) = \gcd(f(x) - q(x)g(x), g(x)) = \tilde{u}(x)r(x) + \tilde{v}(x)g(x) = \tilde{u}(x)(f(x) - q(x)g(x)) + \tilde{v}(x)g(x),$$

a proto můžeme položit $u(x) = \tilde{u}(x)$ a $v(x) = \tilde{v}(x) - \tilde{u}(x)q(x)$. Tím je věta dokázána.

Základní vlastnosti (4 ze 4)

Věta 32.6 — *Polynomial factor theorem*. Buď T těleso a $p(x) \in T[x]$ polynom stupně n. Prvek $\xi \in T$ je kořen polynomu p právě tehdy, když $p(x) = (x - \xi)g(x)$, kde $g(x) \in T[x]$ je stupně $n - 1$.

Důkaz. Mějme $\xi \in T$ takové, že $p(\xi) = 0$. Z věty o dělení polynomů (lemma 32.4) plyne existence polynomů $g(x), r(x) \in T[x]$ takových, že

$$p(x) = (x - \xi)g(x) + r(x)$$

a stupeň polynomu r je ostře menší než $\deg(x - \xi) = 1$ nebo r je nulový polynom. Protože platí $p(\xi) = r(\xi)$, tak r je nulový polynom, a tedy $p(x) = (x - \xi)g(x)$.

Pro dokázání druhého směru stačí ověřit, že ξ je kořenem. ■

32.2 Ireducibilní polynom

Ireducibilní polynom

Definice 32.7 Buď $P(x) \in K[x]$ stupně alespoň 1. Řekneme, že $P(x)$ je ireducibilní nad okruhem K, jestliže pro každé dva polynomy $A(x)$ a $B(x)$ z $K[x]$ platí

$$A(x) \cdot B(x) = P(x) \quad \Longrightarrow \quad (\deg(A(x)) = 0 \text{ NEBO } \deg(B(x)) = 0).$$

Ireducibilní polynomy jsou definovány analogicky jako jsou prvočísla v množině přirozených celých čísel.

Poznámka: „analogicky“ v předchozí větě je použito volně. Prvočísla jsou obecně tzv. prvoelementy (*prime elements*), kdežto my potřebujeme ireducibilní prvky. Oba tyto pojmy splývají, pokud lze ve zkoumaném komutativním okruhu (tedy $K[x]$) jednoznačně faktorizovat (na ireducibilní prvky).

■ **Příklad 32.8** $x^2 + 1$ je ireducibilní nad $\mathbb{Q}$, $x^2 - 1 = (x + 1)(x - 1)$ není.

$x^2 + 1$ není ireducibilní nad $(\mathbb{Z}_2, +_2, \times_2)$, zde totiž platí

$$x^2 + 1 = (x + 1)(x + 1) = x^2 + 2x + 1.$$

■

[11] U nejvyšší mocniny mají jedničku tělesa T. Takový polynom získáme vynásobením původního polynomu inverzí (pro násobení v tělese T) k číslu u nejvyšší mocniny x.

Ireducibilní polynomy: co o nich víme (1 z 2)

Věta 32.9 Mějme celé $n > 1$ a prvočíslo p. Označme $N(p, n)$ počet monických polynomů stupně n ireducibilních nad $\mathbb{Z}_p$. Potom

$$N(p,n) = \frac{1}{n}\sum_{d|n}\mu(d)p^{\frac{n}{d}} \geq \frac{1}{n}\left(p^n - \sum_{q|n, q \text{ prvoč.}} p^{\frac{n}{q}}\right).$$

Kde μ je Möbiova funkce definovaná pro celé $n > 0$ takto:

$$\mu(n) = \begin{cases} 1 & \text{pokud } n \text{ neobsahuje čtverec prvočísla a má sudý počet prvoč. faktorů,} \\ -1 & \text{pokud } n \text{ neobsahuje čtverec prvočísla a má lichý počet prvoč. faktorů,} \\ 0 & \text{pokud } n \text{ obsahuje čtverec prvočísla.} \end{cases}$$

a monický polynom je takový polynom, který má za koeficient u nejvyšší mocniny jedničku.

Ireducibilní polynomy: co o nich víme (2 z 2)

Otázka: Jak najít ireducibilní polynom? Poznat, jestli je daný polynom ireducibilní je snazší, než poznat jestli dané číslo je prvočíslo:

Dokonce existují polynomiální algoritmy, které nejen rozhodnou o ireducibilitě, ale dokonce najdou rozklad na ireducibilní polynomy (obdoba prvočíselného rozkladu).

Jedná se o Berlekampův a Cantor–Zassenhausův algoritmus: detaily např. v D. Knuth, *The Art of Computer Programming*, Vol. 2, sekce 4.6.

33 Konečná tělesa

33.1

Konečná tělesa

Definice 33.1 — konečné těleso (*finite field*). Těleso, které má konečný počet prvků, se nazývá konečné.

Řádem tělesa se, podobně jako u grup, označuje počet prvků tělesa. Tedy konečná tělesa jsou tělesa konečného řádu.

Základní příklad konečného tělesa je množina (zbytkových tříd modulo p) $\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$ s operacemi modulo **prvočíslo** p (viz minulé přednášky).

Např. pro $p = 5$ dostáváme těleso s násl. operacemi

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

a

$\cdot$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

O tělese $(\mathbb{Z}_p, +, \cdot)$

Aditivní grupa je $\mathbb{Z}_p^+$:

- Má řád p.

- Každý nenulový prvek je její generátor a má tedy řád p (to platí pro všechny grupy s prvočíselným řádem).
- $(\mathbb{Z}_p, +)$ je grupou i pro p, které není prvočíslo.

Multiplikativní grupa je $\mathbb{Z}_p^\times$:

- Má řád $p-1$ a to není jakožto sudé číslo prvočíslo (s výjimkou $p=3$)!
- $\mathbb{Z}_p^\times$ je cyklická (tj. existuje v ní generátor).
- Počet generátorů závisí na jejím řádu $p-1$, a je roven počtu čísel nesoudělných s $p-1$, tedy $\varphi(p-1)$.
- ($(\mathbb{Z}_p \setminus \{0\}, \times_p)$ je grupou pouze pro prvočíselné p, jinak obsahuje dělitele nuly.)
- Nechť $k < p$ dělí $p-1$, pak v $\mathbb{Z}_p^\times$ existuje podgrupa řádu k a obsahuje právě ty prvky $a \in \mathbb{Z}_p$, pro které $a^k = 1$.

Existují další konečná tělesa?

- V předchozí části přednášky jsme pro každé prvočíslo p nalezli konečné těleso mající p prvků, $(\mathbb{Z}_p, +_p, \times_p)$.
- Přirozeně se nabízí otázka: existují i tělesa s jiným počtem prvků než prvočíselným?
- Například na těleso $(\mathbb{Z}_2, +_2, \times_2)$ se můžeme dívat jako na model jednoho bitu. Co kdybychom chtěli pracovat se několikabitovým slovem jako s prvkem tělesa (toho se opět využívá v kryptologii, viz další část přednášky).
- Přirozeně by člověka napadlo, například, na $\mathbb{Z}_2^8$ zavést operace po složkách modulo 2. Tj.

$$(a_1, \ldots, a_8) + (b_1, \ldots, b_8) := (a_1 +_2 b_1, \ldots, a_8 +_2 b_8),$$
$$(a_1, \ldots, a_8) \cdot (b_1, \ldots, b_8) := (a_1 \times_2 b_1, \ldots, a_8 \times_2 b_8).$$

  Tato struktura ovšem tvoří **pouze** komutativní okruh. Není to ani obor integrity. (Rozmyslete!)

Existují další konečná tělesa? Ano!

Následující konstrukce nám již dá těleso neprvočíselného řádu. Mějme zadáno prvočíslo p a celé $n \geq 2$.

1. Uvažme těleso T mající p prvků (ta již umíme konstruovat).
2. Sestrojme okruh $T[x]$ všech polynomů nad tělesem T (ten má nekonečně mnoho prvků).
3. Pro zadané kladné celé n nalezněme polynom $P(x) \in T[x]$ ireducibilní nad T mající stupeň n.
4. Uvažme množinu F všech polynomů z $T[x]$ stupně menšího nebo rovno $n-1$, včetně nulového polynomu, (těch je p^n) a zaveďme na této množině operace:

sčítání: stejně jako v $T[x]$;

násobení: $R(x) \cdot S(x) := \big(R(x) \cdot_{T[x]} S(x)\big) \bmod P(x)$.

F s takto zavedenými operacemi tvoří těleso mající p^n prvků.

Příklad (1 z 2)

■ **Příklad 33.2** Výše uvedenou konstrukci demonstrujeme na $p = 2$, $T = (\mathbb{Z}_2, +_2, \times_2)$, $n = 4$ a

$$P(x) = x^4 + x + 1.$$

Pro zjednodušení zápisu ztotožněme polynomy s řetězci, tj.

$$\sum_{i=0}^{3} a_i x^i \longleftrightarrow a_3 a_2 a_1 a_0$$

pro $a_i \in T = \mathbb{Z}_2$, $i = 0, 1, 2, 3$. ■

Sečtěte 1011 a 0111:

$$1011 + 0111 = 1100.$$

Skutečně pouze sečteme koeficienty polynomů u stejných mocnin modulo 2, protože tyto koeficienty žijí v T.

Příklad (2 z 2)

Vynásobte 1101 a 0110: Součinem polynomů $R(x) = x^3 + x^2 + 1$ a $S(x) = x^2 + x$ v $T[x]$ je polynom

$$Q(x) := R(x) \cdot S(x) = x^5 + x^3 + x^2 + x.$$

Pomocí známého algoritmu dělení polynomu $Q(x)$ polynomem $P(x)$ získáme vztah

$$Q(x) = x \cdot P(x) + x^3.$$

Tudíž zbytkem po dělení polynomu $Q(x)$ polynomem $P(x)$ je polynom x^3. To je výsledek operace násobení, uzavíráme

$$1101 \cdot 0110 = 1000.$$

Znovu zdůrazňujeme: nejde o násobení modulo 2 (`AND`) po složkách, to by nám dalo $1101 \cdot 0110 = 0100$. S touto operací ale **nedostaneme** těleso, viz předchozí poznámky.

Tělesa kterých řádů existují?

Zatím jsme si ukázali konstrukci konečných těles řádu $p = p^1$ a p^n, kde p je prvočíslo a n je kladné celé číslo. Existují tělesa libovolného řádu?

Věta 33.3 Řádem konečného tělesa musí být mocnina prvočísla, tedy číslo zapsatelné jako p^n, kde p je prvočíslo a n je kladné celé číslo.
Navíc platí, že všechna tělesa řádu p^n jsou **navzájem izomorfní**.

Důsledek: neexistuje těleso s 6, 10, 12, 14, …prvky.

Definice 33.4 Těleso s p^n prvky nazýváme *konečné těleso* nebo též *Galoisovo těleso* (*Galois field*) a značíme ho $GF(p^n)$. Prvočíslo p se nazývá charakteristikou tělesa $GF(p^n)$.

$GF(p^n)$: aditivní grupa

Co víme o aditivní grupě tělesa $GF(p^n)$:

- Má řád p^n.

- Neutrální prvek je $0 = 00 \cdots 0 = 0^n$.
- Inverze k prvku $b_1 b_2 \cdots b_n$ je $(p - b_1)(p - b_2) \cdots (p - b_n)$.
- Pro $n > 1$ není cyklická, dokonce pro každý prvek v platí, že $(p + 1) \times v = v$, resp. $p \times v = 0$.

$GF(p^n)$: multiplikativní grupa

Co víme o multiplikativní grupě tělesa $GF(p^n)$:

- Má řád $p^n - 1$.
- Neutrální prvek je $00 \cdots 1 = 0^{n-1}1$.
- Inverzi ke každému prvku umíme nalézt pomocí EEA v polynomiálním čase.
- Je vždy cyklická (důkaz není moc složitý, ale zabral by nám moc času).

34 Aplikace konečných těles v kryptografii

34.1 AES

Symetrické šifrování (více v předmětu MI-BHW)

- Při šifrované výměně delšího textu, jsou asymetrické šifry (RSA, Diffie-Hellman a spol.) neefektivní.
- Proto se používá **symetrické šifrování**, kde se předpokládá, že Alice a Bob znají nějaký společný soukromý klíč, který nikdo jiný nezná a který šifrování výrazně usnadní. Asymetrické šifry se použijí pouze k výměně (resp. vytvoření) tohoto společného soukromého klíče.
- Velmi používaná metoda je bloková šifra (*block cipher*) zvaná *Advanced Encryption Standard* (AES). Zde se seznámíme s matematickým podhoubím této metody.

Bloková šifra AES

- Kódovaný text si rozdělíme na bloky o (např.) 8 bitech. Ty zašifrujeme pomocí klíče tak, že dešifrování lze snadno provést pouze se znalostí toho samého klíče.
- Toto šifrování v AES je založeno na tom, že operace s $n = 8$ bity lze chápat jako aritmetické operace v konečném tělese s 2^n prvky pro $n = 8$. Tělesa s 2^n prvky zveme binární tělesa a značíme $GF(2^n)$.
- Dle specifikace AES se násobení počítá modulo

  $$x^8 + x^4 + x^3 + x + 1.$$

  K reprezentaci prvků se tedy používá 8 (= stupeň polynomu výše) bitů.
- (Aritmetické operace v $GF(2^n)$ nejsou jedinými operacemi s bloky v AES.)

34.2 Eliptické křivky

Eliptická křivka

- Pod **eliptickou křivkou** nad tělesem T rozumíme množinu všech bodů $(x, y) \in T^2$ splňující zjednodušenou Weierstrassovu rovnici
$$y^2 = x^3 + ax + b,$$
případně Weierstrassovu rovnici
$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6,$$
kde $a, b \in T$ a $a_i \in T$ jsou zadány.

- Za jistých předpokladů na koeficienty lze na množině všech bodů na eliptické křivce obohacené o jeden nový prvek zavést binární operaci vytvářející strukturu abelovské grupy.

Počítání na eliptických křivkách

S body $(x, y) \in T^2$ na eliptické křivce dané zjednodušenou Weierstrassovou rovnicí se počítá následovně (operace se tradičně, ovšem bez zvláštního důvodu, značí $+$):

Definice 34.1 Pro dva body $P = (x_1, y_1)$ a $Q = (x_2, y_2)$, $x_1 \neq x_2$, definujeme $P + Q = (x_3, y_3)$ takto:

$$x_3 = s^2 - x_1 - x_2$$
$$y_3 = s(x_1 - x_3) - y_1$$
kde
$$s = \begin{cases} \frac{y_2 - y_1}{x_2 - x_1} & \text{pokud } P \neq Q \\ \frac{3x_1^2 + a}{2y_1} & \text{pokud } P = Q\,. \end{cases}$$

Parametr a je převzat z rovnice dané eliptické křivky. Neutrální prvek $\mathcal{O}$ je „uměle“ přidán tak, aby měl vlastnosti neutrálního prvku.

Dále pro dva body $P = (x_1, y_1)$ a $Q = (x_1, y_2)$ klademe $P + Q = \mathcal{O}$.

Eliptické křivky nad tělesem reálných čísel (1 ze 3)

Názorněji se dá vysvětlit počítání nad eliptickými křivkami, pokud bereme body (x, y) ze *souvislé* roviny $\mathbb{R}^2$.

Definice 34.2 Eliptickou křivkou rozumíme množinu bodů splňující rovnici
$$y^2 = x^3 + ax + b\,,$$
kde pro reálné koeficienty a, b platí, že $-16(4a^3 + 27b^2) \neq 0$.

Grupovou operaci sčítání definovanou dříve pak lze interpretovat geometricky (pro názornou grafickou ukázku viz Wolfram Demonstrations Project).

Eliptické křivky nad tělesem reálných čísel (2 ze 3)

Mějme dva různé body P a Q na eliptické křivce E nad $\mathbb{R}$, potom pro jejich součet platí:

- Sestroj přímku p procházející body P a Q.

- Pokud přímka p má s E ještě jeden průsečík $R = (x, y)$ různý od P a Q, pak $P + Q = (x, -y)$.

- Pokud průnik přímky p a eliptické křivky E je právě množina $\{P, Q\}$, pak $P + Q = \mathcal{O}$.

Mějme jeden bod P na eliptické křivce E, potom pro součet $P + P$ platí

- Sestroj tečnu p křivky E v bodě P.

- Pokud tato tečna prochází ještě jedním bodem $R = (x, y)$ křivky E, pak $P + P = (x, -y)$

- Pokud tato tečna protíná E právě v P, pak $P + P = \mathcal{O}$.

Eliptické křivky nad tělesem reálných čísel (3 ze 3)

- **Poznámka**: Při sčítání bodů tedy hledáme průsečíky nějaké přímky $y = sx + d$, kde s je směrnice a $d \in \mathbb{R}$, a eliptické křivky $y^2 = x^3 + ax + b$. To vede na řešení rovnice

$$(sx + d)^2 = x^3 + ax + b\,,$$

  což je polynomiální rovnice 3. stupně a ta, jak víme, může mít 1 až 3 různé reálné kořeny (jedno řešení máme vždy ze zadání).

- Např. situace, kdy pro různé P a Q dostaneme pouze dva kořeny, odpovídá tomu, že $Q = -P$ a výsledkem součtu je neutrální prvek $\mathcal{O}$.

- Místo nad $\mathbb{R}$ můžeme pracovat nad tělesem $GF(p^n)$. I v tomto případě dále mluvíme o „eliptické křivce“, ačkoliv příslušnou diskrétní množinu bodů bychom na obrázku za křivku jistě neoznačili.

Řešené příklady: Konečná tělesa

Připomenutí: tělesa kterých řádů existují?

Věta 34.3 Řádem konečného tělesa musí být mocnina prvočísla, tedy číslo zapsatelné jako p^n, kde p je prvočíslo a n je kladné celé číslo.
Navíc platí, že všechna tělesa řádu p^n jsou **navzájem izomorfní**.

Definice 34.4 Těleso s p^n prvky nazýváme *Galois field* a značíme ho $GF(p^n)$. Prvočíslo p se nazývá charakteristikou tělesa $GF(p^n)$.

Konstrukce konečného tělesa řádu p^n

Buď $p(x) \in \mathbb{Z}_p[x]$ ireducibilní polynom nad $\mathbb{Z}_p$ stupně n.

Trojice

$$\Big(\{q(x) \in \mathbb{Z}_p[x] \colon \deg(q(x)) \text{ je menší než } n \text{ nebo není definován } \}, +, \cdot \bmod p(x) \Big),$$

kde $+$ a $\cdot$ jsou operace sčítání a násobení polynomů (z okruhu polynomů $\mathbb{Z}_p[x]$), je konečným tělesem řádu p^n (tedy $GF(p^n)$).

$GF(p^n)$: aditivní grupa

Co víme o aditivní grupě tělesa $GF(p^n)$:

- Má řád p^n.
- Neutrální prvek je $0 = 00\cdots0 = 0^n$.
- Inverze k prvku $b_1b_2\cdots b_n$ je $(p-b_1)(p-b_2)\cdots(p-b_n)$.
- Pro $n > 1$ není cyklická, dokonce pro každý prvek v platí, že $(p+1) \times v = v$, resp. $p \times v = 0$.

$GF(p^n)$: multiplikativní grupa

Co víme o multiplikativní grupě tělesa $GF(p^n)$:

- Má řád $p^n - 1$.
- Neutrální prvek je $00\cdots1 = 0^{n-1}1$.
- Je vždy cyklická (důkaz není moc složitý, ale zabral by nám moc času).

Konečná tělesa – příklady

Základní cvičení 23.1

Základní cvičení 23.1
V tělese $\mathbb{Z}_{263}$ najděte multiplikativní inverzi k prvku 112.

Základní cvičení 23.2
Rozhodněte, zda je polynom $P(x)$ ireducibilní nad tělesem $\mathbb{Z}_5$, kde

(a) $P_a(x) = x^3 + 2x + 1$;

(b) $P_b(x) = x^2 + 2x + 2$;

Základní cvičení 23.3
Rozhodněte, zda je polynom $P(x)$ ireducibilní nad tělesem $\mathbb{Z}_3$, kde

(a) $P_a(x) = 2x^4 + x^3 + 2x + 1$;

(b) $P_b(x) = x^4 + x^3 + x + 2$;

(c) $P_c(x) = x^4 + x + 2$.

Základní cvičení 23.7
V tělese $GF(3^2)$, kde se násobí modulo ireducibilní polynom $x^2 + 2x + 2$, najděte

(a) všechna y taková, aby $21(y + 11) = 01 + y$,

(b) najděte všechny generátory multiplikativní grupy tohoto tělesa.

Základní cvičení 23.16
Mějme těleso $GF(5^3)$, kde se násobí modulo ireducibilní polynom $x^3 + 3x + 3$.

(a) Nalezněte inverzní prvek vzhledem k násobení k prvku $x^2 + 2x$.

(b) Nalezněte všechna $y \in GF(5^3)$, která splňují $120 \cdot y^2 = 111$.

35 Aritmetické operace v konečném tělese $GF(p^n)$

Aritmetické operace

Uvedeme si základní přístupy pro provádění aritmetických operací v tělese $GF(p^n)$.

Uvažujeme, že konečné těleso máme vybudováno kanonicky tedy pomocí polynomů nad $\mathbb{Z}_p$ a ireducibilního polynomu (nad $\mathbb{Z}_p$) stupně n.

Budeme uvažovat obecné p a n a elementární operace v $\mathbb{Z}_p$. Např. pro $n = 1$ a $p = 2$ jsou známy lepší algoritmy (co do komplexity).

35.1 Sčítání

Sčítání polynomů

Po koeficientech, tedy $\mathcal{O}(n)$.

Násobení

Násobení polynomů: $\mathcal{O}(n^2)$.

Hledání zbytku po dělení: opět $\mathcal{O}(n^2)$.
(Pro polynomy velkých stupňů existují algoritmy, které mohou být efektivnější.)

Celkem: $\mathcal{O}(n^2)$.

35.2 Mocnění

Mocnění

Mocněním myslíme výpočet g^k, kde $g \in GF(p^n)$ a $k < p^n$.

Klasický přístup: **metoda opakovaných čtverců** *Square & multiply*, v aditivních grupách *Double & add*

Počet přenásobení dle exponentu: $\mathcal{O}(n \log p)$
V každém kroce je násobení
Celkem tedy $\mathcal{O}(n^3 \log p)$.

35.3 Multiplikativní inverze

Strategie pro hledání multiplikativních inverzí

Hledáme multiplikativní inverzi k prvku g.

1. hrubou silou: $\mathcal{O}(p^n)$
2. výpočtem g^{p^n-2}: $\mathcal{O}(n^3 \log p)$.
3. EEA: $\mathcal{O}(n^2)$ (toto je výsledek jemnější analýzy, nahrubo vyjde $\mathcal{O}(n^3)$)
4. ...

Např. Itoh-Tsujiho inverze, využívající následující větu

Věta 35.1 Nechť $g \in GF(q^m)^*$ a $r = \frac{q^m-1}{q-1}$. Platí

$$g^{-1} = (g^r)^{-1} g^{r-1}$$

!

Jelikož g^r je prvkem podstruktury se stejnými vlastnostmi vzhledem k daným operaci (konkrétně podtělesa řádu q), lze dosáhnout lepších výsledků.

EEA: ukázka v $GF(3^3)$

V $GF(3^3)$ počítejme modulo ireducibilní polynom $P(x) = x^3 + 2x + 1$. Hledejme inverzi k $Q(x) = x^2 + 2x + 2$, tj. prvku 122.

1. Počítáme $\gcd(P(x), Q(x))$, vyjde $P(x) = (x+1)Q(x) + x + 2$.

2. Počítáme $\gcd(Q(x), x+2)$, vyjde $Q(x) = x \cdot (x+2) + 2$ a proto $2 = Q(x) - x \cdot (x+2) = (x^2+x+1) \cdot Q(x) + 2x \cdot P(x)$.

Tudíž

$$2 \equiv (x^2 + x + 1) \cdot Q(x) \pmod{P(x)},$$

a jelikož $2^{-1} = 2$ v $GF(3)$, dostaneme

$$1 \equiv (2x^2 + 2x + 2) \cdot Q(x) \pmod{P(x)}.$$

Hledanou inverzí k 122 v tomto tělese proto je 222.

36 Hledání izomorfismů mezi dvěma konečnými tělesy

36.1 Konstrukce izomorfismu

Hledání izomorfismu

Mějme prvočíslo p a kladné celé číslo n. Zkonstruujme dvě tělesa řádu p^n pomocí dvou ireducibilních polynomů $f_1 \in \mathbb{Z}_p[x], f_2 \in \mathbb{Z}_p[y]$ stupně n.

Označme tato tělesa F_1 (tady násobíme modulo f_1) a F_2 (tady násobíme modulo f_2).

Víme, že F_1 je izomorfní s F_2. Jak lze najít takový izomorfismus?

(Pro $n = 1$ již umíme...)

Pro lepší čitelnost budeme prvky F_1 psát s formální proměnnou x, prvky F_2 s proměnnou y.

Konstrukce izomorfismu

Mějme $t \in F_1$. Do polynomu f_1 můžeme dosadit t a operace chápat jako operace v F_1, a dostat tedy $f_1(t) \in F_1$.

Prvek $x \in F_1$ splňuje $f_1(x) = 0$ (neboli je kořenem polynomu f_1 nad F_1).

Protože F_1 je izomorfní s F_2, tak nějaký zvolený izomorfismus Ψ zobrazuje $x \in F_1$ na $\Psi(x) \in F_2$. Protože x je kořenem polynomu f_1 a nutně $\Psi(0) = 0$, tak $\Psi(x)$ je kořenem polynomu f_1 **nad** F_2. (Tím myslíme polynom z $\mathbb{Z}_p[y]$ se stejnými koeficienty jako f_1.)

Tedy polynom f_1 nad F_2 má v F_2 kořen. Označme nějaký takový kořen symbolem θ ($\theta \in F_2$).

Označíme-li obecný prvek tělesa F_1 jako $g(x)$ (polynom nad $\mathbb{Z}_p$ stupně nejvýše $n-1$), pak hledané zobrazení definujeme:

$$\Psi : \underbrace{g(x)}_{\in F_1} \mapsto \underbrace{g(\theta)}_{\in F_2}.$$

Náznak důkazu (1 ze 2)

I. Zobrazení Ψ je homomorfismus z F_1 do F_2:

Důkaz: nechť $g_1(x), g_2(x) \in F_1$, pak

$$\Psi(g_1(x) + g_2(x)) = \Psi((g_1 + g_2)(x)) = (g_1 + g_2)(\theta)$$

$$\Psi(g_1(x)) + \Psi(g_2(x)) = g_1(\theta) + g_2(\theta) = (g_1 + g_2)(\theta)$$

$$\Psi(g_1(x) \cdot g_2(x)) = \Psi((g_1(x) \cdot g_2(x) \bmod f_1)(x)) = (g_1(x) \cdot g_2(x) \bmod f_1)(\theta) \bmod f_2$$

$$\begin{aligned}\Psi(g_1(x)) \cdot \Psi(g_2(x)) &= g_1(\theta) \cdot g_2(\theta) \bmod f_2 \\ &= (g_1(x) \bmod f_1)(\theta) \cdot (g_2(x) \bmod f_1)(\theta) \bmod f_2 \\ &= (g_1(x) \cdot g_2(x) \bmod f_1)(\theta) \bmod f_2\end{aligned}$$

Náznak důkazu (2 ze 2)

II. Zobrazení Ψ je bijekce F_1 a F_2.

Náznak (opravdu!) důkazu: předpokládejme, že $g_1(x), g_2(x) \in F_1$ a $\Psi(g_1(x)) = \Psi(g_2(x))$.

(…)
Protože θ není kořenem nad F_2 žádného polynomu nad $\mathbb{Z}_p$ stupně menšího než n, tak $g_1(x) = g_2(x)$.

36.2 Ukázka

Ukázka hledání izomorfismu

Mějme $p = 3, n = 3$ a $f_1(x) = x^3 + 2x + 1$ a $f_2(y) = y^3 + 2y + 2$.

Hledáme θ: kořen f_1 nad F_2.
Označme $\theta = a + by + cy^2 \in F_2$ a hledejme

$$\begin{aligned}0 = f_1(\theta) &= f_1(a + by + cy^2) \\ &= (a + by + cy^2)^3 + 2(a + by + cy^2) + 1 \\ &= (a^3 + b^3y^3 + c^3y^6) + 2(a + by + cy^2) + 1 \\ &= (a + by^3 + cy^6) + 2(a + by + cy^2) + 1 \\ &= (a + b(y + 1) + c(y^2 + 2y + 1)) + 2(a + by + cy^2) + 1 \\ &= 2cy + b + c + 1\end{aligned}$$

A tedy $c = 0$, $b = 2$ a a lze zvolit (z $\mathbb{Z}_3$).
Volme $a = 0$ a tedy hledané $\theta = 2y$, a $\Psi(g(x)) = g(2y)$.
Obecně tedy $\Psi(a + bx + cx^2) = a + b2y + c4y^2 = a + 2by + cy^2$.

37 Soustavy lineárních kongruencí

Soustavy lineárních kongruencí (1 ze 2)

Problém: řešíme soustavu rovnic

$$\begin{aligned}a &\equiv a_1 \pmod{m_1} \\ a &\equiv a_2 \pmod{m_2} \\ &\vdots \\ a &\equiv a_N \pmod{m_N}\end{aligned} \tag{2}$$

kde $m_1, \ldots, m_N$ jsou navzájem nesoudělná $a_1 \in \mathbb{Z}_{m_1}$, …, $a_N \in \mathbb{Z}_{m_N}$, a $a \in \mathbb{Z}_n$ je neznámá.

- Pokud nejsou m_i navzájem nesoudělná, nemusí řešení existovat a situace se celkově komplikuje.

- **Myšlenka**: zkusíme zkonstruovat čísla $x_1, \ldots, x_N$ tak, že x_i bude řešit i-tou rovnici, tj. $x_i \equiv a_i \pmod{m_i}$, a pro ostatní rovnice bude $x_i \equiv 0 \pmod{m_k}$, kde $k \neq i$.

- Potom bude jistě $a = x_1 + x_2 + \cdots + x_N$ **řešením soustavy** (2).

Soustavy lineárních kongruencí (2 ze 2)

Jak taková x_i najdeme?

- Položme $M = \prod_{i=1}^{N} m_i$ a $M_i = \frac{M}{m_i}$ a pomocí postupu uvedeného dříve vyřešme rovnici

$$y_i M_i \equiv 1 \pmod{m_i}.$$

s neznámou y_i.

- Položíme-li

$$x_i = y_i M_i a_i,$$

dostáváme čísla s požadovanými vlastnostmi, a tedy

$$a = y_1 M_1 a_1 + y_2 M_2 a_2 + \cdots + y_N M_N a_N.$$

- Tím jsme (skoro) dokázali slavnou čínskou větu o zbytcích.

Čínská věta o zbytcích

Věta 37.1 — Čínská věta o zbytcích (*Chinese remainder theorem – CRT*). Nechť $m_1, \ldots, m_N$ jsou navzájem nesoudělná čísla a nechť $M = \prod_{i=1}^{N} m_i$. Pro libovolnou N-tici $a_1 \in \mathbb{Z}_{m_1}$, …, $a_N \in \mathbb{Z}_{m_N}$ existuje jednoznačně určené $a \in \mathbb{Z}_M$ tak, že

$$a \equiv a_i \pmod{m_i} \quad \text{pro všechna } i = 1, \ldots, N.$$

Platí

$$a \equiv \sum_{i=1}^{N} a_i y_i M_i \pmod{M},$$

kde $M_i = \frac{M}{m_i}$ a pro všechna i a $j \neq i$ platí

$$y_i M_i \equiv 1 \pmod{m_i} \quad \text{a} \quad y_i M_i \equiv 0 \pmod{m_j}.$$

Existenci řešení jsme ukázali, dokonce i postup jak jej nalézt. Zbývá dokázat jednoznačnost.

CRT: důkaz jednoznačnosti řešení

Při zachování značení z předchozí věty označme

$$\Gamma : \mathbb{Z}_M^+ \mapsto \mathbb{Z}_{m_1}^+ \times \cdots \times \mathbb{Z}_{m_N}^+$$

zobrazení, které číslu $a \in \mathbb{Z}_M^+$ přiřadí N-tici $(a_1, \ldots, a_N)$, kde platí $a \equiv a_i \pmod{m_i}$ pro všechna i.

- CRT nám říká, jak najít vzor N-tice $(a_1, \ldots, a_N)$ při zobrazení Γ (rozmyslet!).

- Zatím jsme si ukázali, že zobrazení Γ je surjektivní, neb pro každou N-tici $(a_1, \ldots, a_N)$ umíme najít $a \in \mathbb{Z}_M$ tak, že $\Gamma(a) = (a_1, \ldots, a_N)$.

- Jelikož jsou ale množiny $\mathbb{Z}_M^+$ a $\mathbb{Z}_{m_1}^+ \times \cdots \times \mathbb{Z}_{m_N}^+$ stejně velké, musí toto zobrazení být i injektivní, a tedy se jedná o bijekci.

- Je tedy nemožné, aby dvě různé N-tice měly dva různé vzory a jednoznačnost řešení a z CRT je dokázána!

Residue number system

Zobrazení Γ je (dokažte si!):

- izomorfismus grupy $\mathbb{Z}_M^+$ a grupy $\mathbb{Z}_{m_1}^+ \times \cdots \times \mathbb{Z}_{m_N}^+$, kde bereme operaci sčítání po složkách: i-tou složku sčítáme modulo m_i;

- izomorfismus grupy $\mathbb{Z}_M^\times$ a grupy $\mathbb{Z}_{m_1}^\times \times \cdots \times \mathbb{Z}_{m_N}^\times$, kde bereme operaci násobení po složkách: i-tou složku násobíme modulo m_i.

Zobrazení Γ určuje tzv. Residue number system. Místo modulo M počítáme v systému modulo $(m_1, \ldots, m_N)$. Jelikož zobrazení lze chápat též jako izomorfismus mezi **okruhy** $\mathbb{Z}_M$ a

$$\mathbb{Z}_{m_1} \times \cdots \times \mathbb{Z}_{m_N},$$

jediná problematická operace v tomto číselném systému zůstane dělení.