

Úvod

Numerická matematika (v anglické terminologii “numerical/computational mathematics”) se zabývá zpracováním matematických modelů pomocí počítačů a realizuje tak přechod od čistě teoretické matematiky k prakticky použitelným výsledkům. Užívá se v celé řadě oborů, nejen v technických a přírodních vědách. Numerická matematika prakticky používá teoretické výsledky matematických věd a to nejen matematické analýzy a algebry, ale i mnoha dalších oborů jako je funkcionální analýza, atd. Samotná numerická matematika lze rozdělit na teoretickou část (numerická analýza) a praktickou část. Cílem teoretické numerické matematiky (numerické analýzy) je analýza vlastností numerických metod a algoritmů. Numerická matematika ve své praktické části se naopak zabývá použitím numerických metod pro řešení vybraných problémů.

Předmět Numerická matematika se zabývá vybranými partiemi z numerické matematiky, jedná se zejména o princip řešení (velkých) soustav lineárních a nelineárních rovnic, numerické řešení obyčejných diferenciálních rovnic a také numerická aproximace parciálních diferenciálních rovnic. Cílem těchto poznámek je poskytnout základní učební text studentům.

Šedou barvou pozadí (stejně jako tato část) jsou zvýrazněny teoreticky obtížnější části látky, které jsou vhodné pro porozumění ale nemusí být vždy plně obsahem přednášek a nebudou tedy ani obsahem zkoušky.

Opakování

Na přednáškách a v těchto materiálech budeme užívat následující označení: Symbolem \mathbb{R} značíme množinu reálných čísel a \mathbb{C} množinu komplexních čísel. Symbolem \mathbb{N} značíme množinu přirozených čísel a \mathbb{Z} množinu celých čísel. Symbolem $\mathbb{R}^{n \times m}$ značíme množinu všech reálných matic typu $n \times m$, tj. $\mathbf{A} \in \mathbb{R}^{n \times m}$ čteme “matice \mathbf{A} má n -řádků a m -sloupců”. Matice značíme symboly $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \dots$. Obdobně vektory značíme většinou tučně, tj. $\mathbf{u}, \mathbf{v}, \mathbf{x}$ nebo také i velkým písmenem $\mathbf{U}, \mathbf{V}, \mathbf{X}$. Tak jak je obvyklé vektorem rozumíme vždy sloupcový vektor (tj. matice typu $\mathbb{R}^{n \times 1}$). Symbolem \mathbf{E} nebo \mathbf{E}_n značíme jednotkovou matici $\mathbf{E} \in \mathbb{R}^{n \times n}$.

Symbolem $\mathcal{C}(D)$ značíme množinu všech spojitých funkcí na množině $D \subset \mathbb{R}^d$, $d \in \mathbb{N}$. Symbolem $\mathcal{C}^k(I)$, $I \subset \mathbb{R}$ značíme množinu všech funkcí, jejichž k -tá derivace je spojitá na I . Symbolem $\mathcal{C}^k(D)$, $D \subset \mathbb{R}^d$, $d > 1$ značíme množinu všech funkcí, které mají všechny parciální derivace až do řádu k spojitě na D .

Definice 0.1 Matice $\mathbf{B} = (b_{ij}) \in \mathbb{R}^{m \times n}$ je transponovaná matici $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{n \times m}$ pokud platí $b_{ij} = a_{ji}$ pro všechna i, j .

Transponovanou matici značíme symbolem \mathbf{A}^T .

Připomeňme, že platí následující pravidla pro transponování součinu matic.

Lemma 0.1 Nechť matice \mathbf{A}, \mathbf{B} jsou takového typu, že součin matic $\mathbf{A}\mathbf{B}$ je definován. Pak platí

$$(\mathbf{A}\mathbf{B})^T = \mathbf{B}^T \mathbf{A}^T.$$

Speciálně, je-li $\mathbf{A} \in \mathbb{R}^{n \times n}$ a $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ sloupcové vektory, pak platí

$$\mathbf{x}^T \mathbf{A} \mathbf{y} = (\mathbf{x}^T \mathbf{A} \mathbf{y})^T = \mathbf{y}^T \mathbf{A}^T \mathbf{x}$$

neboť výsledkem této operace je (symetrická) matice 1×1 .

Definice 0.2 Říkáme, že matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ je symetrická, pokud platí $a_{ij} = a_{ji}$ pro všechna i, j , tj. $\mathbf{A} = \mathbf{A}^T$.

Definice 0.3 Říkáme, že matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ má vlastní číslo $\lambda \in \mathbb{C}$ a vlastní vektor $\mathbf{u} \in \mathbb{C}^n$, $\mathbf{u} \neq \mathbf{0}$ příslušný tomuto číslu, pokud

$$\mathbf{A}\mathbf{u} = \lambda\mathbf{u}. \tag{1}$$

Lemma 0.2

Vlastní čísla matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ jsou kořeny charakteristického polynomu $\det(\mathbf{A} - \lambda\mathbf{E}) = 0$.

Vlastní vektor \mathbf{u}^* příslušný vlastnímu číslu λ^* je řešením homogenní rovnice $(\mathbf{A} - \lambda^* \mathbf{E})\mathbf{u} = \mathbf{0}$.

Věta 0.1 *Nechť matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ je symetrická. Pak její vlastní čísla jsou reálná. Jsou-li navíc \mathbf{u}_1 resp. \mathbf{u}_2 vlastní vektory příslušné vlastním číslům λ_1 resp. λ_2 , kde $\lambda_1 \neq \lambda_2$, pak $\mathbf{u}_1 \cdot \mathbf{u}_2 = 0$. Nebo-li: vlastní vektory symetrické matice příslušné různým vlastním číslům jsou navzájem kolmé.*

Tvrzení 0.1 *Nechť matice \mathbf{A} je symetrická $n \times n$. Pak existují vlastní čísla λ_i , $i = 1, \dots, n$, a také matice \mathbf{U} , $\mathbf{U}^T \mathbf{U} = \mathbf{E}$, taková, že*

$$\mathbf{U}^T \mathbf{A} \mathbf{U} = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Sloupce matice \mathbf{U} jsou tvořeny navzájem ortonormálními vlastními vektory.

1 Základní pojmy a prostá iterační metoda

1.1 Normy a vlastnosti matic

V předchozích kurzech (např. Matematika 1) jsme se seznámili s Euklidovskou normou vektoru definovanou jako

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

Nahrazením čísla 2 v této definici číslem $p \geq 1$ zobecníme Euklidovskou normu na normu $\|\cdot\|_p$ danou vztahem

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}.$$

V předmětu Numerická matematika budeme dále pracovat s touto normou pouze pro $p = 1$, $p = 2$ a pro limitní případ $p = \infty$. Norma vektoru lze ale zavést obecněji a zobecnit pro výpočet “velikosti” (tj. normy) matice.

Definice 1.1 (*Norma vektoru a matice*)

Normou vektoru v \mathbb{R}^n rozumíme takové zobrazení $\|\cdot\| : \mathbb{R}^n \rightarrow \langle 0, +\infty \rangle$, které pro libovolné vektory $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ a číslo $\alpha \in \mathbb{R}$ splňuje

$$(i) \|\alpha \mathbf{u}\| = |\alpha| \|\mathbf{u}\|, \quad (ii) \|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|, \quad (iii) \|\mathbf{u}\| = 0 \Leftrightarrow \mathbf{u} = \mathbf{0}. \quad (2)$$

Normou matice v $\mathbb{R}^{n \times n}$ rozumíme obecně takové zobrazení $\|\cdot\| : \mathbb{R}^{n \times n} \rightarrow \langle 0, +\infty \rangle$, které pro libovolné dvě matice $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$ a číslo $\alpha \in \mathbb{R}$ splňuje

$$\begin{aligned} (i) \quad & \|\mathbf{A}\| = 0 \text{ právě tehdy když } \mathbf{A} \text{ je nulová matice,} \\ (ii) \quad & \|\alpha \mathbf{A}\| = |\alpha| \|\mathbf{A}\|, \\ (iii) \quad & \|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|, \\ (iv) \quad & \|\mathbf{A} \mathbf{B}\| \leq \|\mathbf{A}\| \|\mathbf{B}\|, \end{aligned} \quad (3)$$

Maticovou normu lze definovat užitím normy vektorové. Čtvercové matice chápeme také jako zobrazení z \mathbb{R}^n do \mathbb{R}^n , které vektoru \mathbf{x} přiřadí vektor $\mathbf{A}\mathbf{x}$. Maticovou normu pak lze definovat jako poměr (vektorové) normy $\mathbf{A}\mathbf{x}$ vzhledem k vektorové normě \mathbf{x} , tedy

$$\|\mathbf{A}\| = \sup_{\mathbf{x}, \|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x}\| = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|}. \quad (4)$$

Pro takto definovanou normu matice je snadné ukázat, že splňuje vlastnosti (3) a je maticovou normou (příslušnou vektorové normě $\|\cdot\|$). Ve speciálních případech takovou normu lze explicitně vyjádřit výpočtovým vztahem, viz následující definice.

Definice 1.2 (Vektorové a maticové normy)

Označme symbolem \mathbf{x} vektor $\mathbf{x} \in \mathbb{R}^n$ a symbolem \mathbf{A} matici $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathcal{J} = \{1, \dots, n\}$. Definujeme normu řádkovou, sloupcovou a Euklidovskou.

$$\text{řádková norma} \quad \|\mathbf{x}\|_\infty = \max_{i \in \mathcal{J}} |x_i|, \quad \|\mathbf{A}\|_\infty = \max_{i \in \mathcal{J}} \sum_{j \in \mathcal{J}} |a_{i,j}|,$$

$$\text{sloupcová norma} \quad \|\mathbf{x}\|_1 = \sum_{i \in \mathcal{J}} |x_i| \quad \|\mathbf{A}\|_1 = \max_{j \in \mathcal{J}} \sum_{i \in \mathcal{J}} |a_{i,j}|,$$

$$\text{Frobeniova norma} \quad \|\mathbf{A}\|_E = \left(\sum_{i,j \in \mathcal{J}} |a_{i,j}|^2 \right)^{\frac{1}{2}},$$

Alternativní značení a spektrální norma. V rámci předmětu NMA se také můžete setkat se značením užívaným ve skriptech Numerická matematika, dřívějších písemkách nebo starších materiálech $\|\cdot\|_\infty = \|\cdot\|_m$, $\|\cdot\|_1 = \|\cdot\|_l$ a $\|\mathbf{x}\|_2 = \|\mathbf{x}\|_E$. Zatímco alternativní značení řádkové a sloupcové normy užívat nebudeme, značení Euklidovské/Frobeniovy normy symbolem $\|\cdot\|_E$ užívat budeme. S Euklidovskou normou je totiž často spojena také tzv. spektrální norma matice $\|\mathbf{A}\|_2$ definovaná jako

$$\|\mathbf{A}\|_2 = \sup_{\mathbf{x} \neq 0} \frac{\|\mathbf{A}\mathbf{x}\|_2}{\|\mathbf{x}\|_2}$$

Obdobně řádková a sloupcová norma matice \mathbf{A} je norma, která vznikne vyjádřením ze vztahu

$$\|\mathbf{A}\|_1 = \sup_{\mathbf{x} \neq 0} \frac{\|\mathbf{A}\mathbf{x}\|_1}{\|\mathbf{x}\|_1}, \quad \|\mathbf{A}\|_\infty = \sup_{\mathbf{x} \neq 0} \frac{\|\mathbf{A}\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty},$$

Pozn. Tato definice maticové normy je vlastně normou lineárního zobrazení $\mathbf{x} \rightarrow \mathbf{A}\mathbf{x}$ a lze rozšířit na případ libovolného lineárního zobrazení. Ačkoliv tato definice vypadá komplikovaně, její geometrická interpretace je poměrně názorná. Nakreslíme si množinu bodů \mathbf{x} takových, že $\|\mathbf{x}\| = 1$ a pak nakreslíme množinu obrazů $\mathbf{A}\mathbf{x}$. Norma nám pak udává největší možné “prodloužení” vektoru \mathbf{x} .

Příklad 1.1 Užijte vektorové normy $\|\cdot\|_1$ resp. $\|\cdot\|_\infty$ a odvodte vztah pro příslušnou maticovou normu.

Věta 1.1 Necht' $\|\cdot\|$ označuje příslušnou stejnou vektorovou i maticovou normu (tj. buď $\|\cdot\|_1$, $\|\cdot\|_\infty$ nebo $\|\cdot\|_E$). Pak pro libovolnou matici $\mathbf{A} \in \mathbb{R}^{n \times n}$ a vektor $\mathbf{x} \in \mathbb{R}^n$ platí

$$\|\mathbf{A}\mathbf{x}\| \leq \|\mathbf{A}\| \|\mathbf{x}\|. \quad (5)$$

Důkaz této věty je snadným podle definice maticové normy příslušné určité vektorové normě. Nicméně pro Frobeniovu normu je ještě nutné užít vztah mezi normou spektrální a Frobeniovou $\|\mathbf{A}\|_2 \leq \|\mathbf{A}\|_F$. Tento vztah nebudeme dokazovat, přestože tento důkaz není komplikovaný, užívá ale souvislost spektrální normy a spektrálního poloměru, který je definován v následující definici.

Definice 1.3 (Spektrální poloměr) *Spektrální poloměr matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ je nezáporné číslo $\rho(\mathbf{A})$ dané jako*

$$\rho(\mathbf{A}) = \max_{i=1, \dots, n} |\lambda_i|, \quad (6)$$

kde $\lambda_1, \dots, \lambda_n$ jsou vlastní čísla matice \mathbf{A} .

Pro spektrální normu matice \mathbf{A} platí $\|\mathbf{A}\|_2 = \sqrt{\rho(\mathbf{A}^T \mathbf{A})}$. Speciálně pro symetrickou matici tedy vidíme, že platí $\|\mathbf{A}\|_2 = \rho(\mathbf{A})$ (rozmyslete si proč?). Podívejme se, zda platí nějaký obecnější vztah mezi nějakou normou matice a spektrálním poloměrem. Vezměme λ vlastní číslo matice \mathbf{A} takové, že $|\lambda| = \rho(\mathbf{A})$. Označme \mathbf{x} příslušný vlastní vektor. Dle definice vlastního čísla a vektoru platí

$$|\lambda| \|\mathbf{x}\| = \|\lambda \mathbf{x}\| = \|\mathbf{A} \mathbf{x}\| \leq \|\mathbf{A}\| \|\mathbf{x}\|.$$

Z této nerovnosti již vyplývá platnost následující věty.

Věta 1.2 (Vztah mezi spektrálním poloměrem a normou) *Pro libovolnou normu matice a spektrální poloměr platí*

$$\rho(\mathbf{A}) \leq \|\mathbf{A}\|. \quad (7)$$

Pro matice je vypovídající také číslo podmíněnosti matice označované $\kappa(\mathbf{A}) = \|\mathbf{A}\| \|\mathbf{A}^{-1}\|$. Pro symetrické matice a spektrální normu matice platí $\kappa(\mathbf{A}) = \lambda_{\max} / \lambda_{\min}$, neboť $\|\mathbf{A}^{-1}\|_2 = \frac{1}{\lambda_{\min}}$. Číslo podmíněnosti matice je důležité např. při řešení soustav lineárních rovnic.

Číslo podmíněnosti, vlastní čísla, vlastní vektory a spektrální poloměr matice jsou důležité vlastnosti, ale jejich výpočet je zvláště v případě rozsáhlých soustav lineárních rovnic obtížný. V řadě případů lze ale u matic jednoduše ověřit jiné vlastnosti, které již o např. spektrálních vlastnostech vypovídají. Takovou vlastností je například diagonální dominance, symetrie nebo pozitivní definitnost.

Definice 1.4 (Ostře diagonálně dominantní matice) *Matici $\mathbf{A} \in \mathbb{R}^{n \times n}$ nazýváme ostře diagonálně dominantní (ODD), pokud platí*

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}| \quad \text{pro všechna } i \in \{1, \dots, n\}, \quad (8)$$

nebo

$$|a_{ii}| > \sum_{j \neq i} |a_{ji}| \quad \text{pro všechna } i \in \{1, \dots, n\}.$$

V odborné literatuře se také užívá pojem diagonálně dominantní (DD) matice v případě, kdy nerovnost v rovnici (8) je neostrá, tj.

$$|a_{ii}| \geq \sum_{j \neq i} |a_{ij}| \quad \text{pro všechna } i \in \{1, \dots, n\}.$$

Diagonálně dominantní maticí je ovšem např. i nulová matice, tedy podmínka DD matice není postačující pro regularitu matice.

Definice 1.5 (Pozitivně definitní matice) *Matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ nazýváme pozitivně definitní pokud pro libovolný vektor $\mathbf{x} \neq 0$ platí*

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_{i,j=1}^n x_i a_{ij} x_j > 0. \quad (9)$$

Obdobně lze definovat matice pozitivně semidefinitní v případě, že v nerovnici (9) platí neostrá nerovnost. Pozitivní semidefinitnost ani diagonální dominance ovšem nezaručují řešitelnost úlohy, na rozdíl od pozitivní definitnosti a ostré diagonální dominance.

Lemma 1.1 *Nechť matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ je ostře diagonálně dominantní nebo pozitivně definitní. Pak je regulární.*

Pro poz. def. matici je regularita matice jednoduchým důsledkem nerovnosti (9) z které vyplývá, že $\mathbf{A} \mathbf{x} \neq 0$ pro $\mathbf{x} \neq 0$ (Rozmyslete si!).

Pro ostře diagonálně dominantní matici \mathbf{A} provedeme důkaz sporem. Předpokládejme, že existuje vektor $\mathbf{x} \neq 0$ takový, že $\mathbf{A} \mathbf{x} = 0$. Vezměme takové i , že $|x_i| \geq |x_j|$ pro všechna $j \neq i$. Pak vidíme, že

$$0 = a_{ii}x_i + \sum_{j \neq i} a_{ij}x_j,$$

nebo-li

$$|a_{ii}||x_i| = \left| - \sum_{j \neq i} a_{ij}x_j \right|.$$

Dělením $|x_i|$ a opakovaným užitím trojúhelníkové nerovnosti $|\alpha + \beta| \leq |\alpha| + |\beta|$ dostaneme

$$|a_{ii}| = \left| - \sum_{j \neq i} a_{ij} \frac{x_j}{x_i} \right| \leq \sum_{j \neq i} |a_{ij}| \frac{|x_j|}{|x_i|} \leq \sum_{j \neq i} |a_{ij}|,$$

kde poslední nerovnost je ale sporem s předpokládanou ODD matice \mathbf{A} .

Pro ODD matici může být důkaz i jednodušší, neboť lze aplikovat Gerschgorinovu větu. Z té ihned vyplývá, že $\lambda = 0$ nemůže být vlastním číslem ODD matice \mathbf{A} a že matice je tedy regulární.

Lemma 1.2 (Gerschgorinova věta) *Nechť je dána matice $\mathbf{A} \in \mathbb{R}^{n \times n}$. Pak všechna její vlastní čísla $\lambda \in \mathbb{C}$ leží ve sjednocení kruhů K_i , jejichž střed je dán bodem $a_{ii} \in \mathbb{C}$ a poloměr součtem absolutních hodnot mimodiagonálních prvků v daném řádku i , tj. $\sum_{j \neq i} |a_{ij}|$. Tedy pro každé vlastní číslo $\lambda \in \mathbb{C}$ matice \mathbf{A} existuje kruh K_i , tak že $\lambda \in K_i$. Zde K_i jsou dány*

$$K_i = \left\{ \lambda \in \mathbb{C} : |\lambda - a_{ii}| \leq \sum_{j \neq i} |a_{ij}| \right\}. \quad (10)$$

Toto tvrzení je geometricky názorné a dává nám velmi jednoduše představu o poloze vlastních čísel. Jeho důkaz je přitom poměrně jednoduchý. Vezmeme vlastní číslo $\lambda \in \mathbb{C}$ a příslušný vlastní vektor \mathbf{x} . Volíme i tak, aby $|x_i| \geq |x_j|$ pro $j \neq i$. Z definice vlastního čísla dostáváme v řádku i

$$\left| \sum_{j \neq i} a_{ij} x_j \right| = |\lambda - a_{ii}| |x_i|.$$

Dělením $|x_i|$ a užitím $|x_i| \geq |x_j|$ již dostaneme požadovanou nerovnost.

Všimněme si, že zatímco ověření ostré diagonální dominance je relativně jednoduché, ověření pozitivní definitnosti již nikoliv. Připomeňme bez důkazu Sylvestrovu kritérium pro určování pozitivní definitnosti symetrické matice s využitím hlavních minorů, tj. determinantů

$$a_{11}, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \dots, \det \mathbf{A}$$

Věta 1.3 (Sylvestrovu kritérium) *Nechť matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ je symetrická. Pak platí: \mathbf{A} je pozitivně definitní tehdy a jen tehdy pokud má všechny hlavní minory kladné.*

Vidíme, že toto kritérium je v případě malých matic snadno ověřitelné. Na druhou stranu samotná definice poz. def. matice nám dává vlastnost, která matice charakterizuje lépe než Sylvestrovu kritérium. Např. užitím definice lze snadno ukázat, že symetrická matice je pozitivně definitní právě tehdy, když její všechna vlastní čísla jsou kladná. Pozitivní definitnost se někdy zavádí pouze pro symetrické matice. Zamyslete se sami, proč.

1.2 Prostá iterační metoda

Zabývejme se nyní řešením (velkých) soustav lineárních rovnic, kdy přímé řešení soustav lineárních rovnic je příliš náročné (na výpočetní čas i na operační paměť). V celé řadě případů jsou tyto velké soustavy tzv. řídké. To přibližně znamená, že počet jejich nenulových prvků je mnohem menší než počet jejich nulových prvků (např. matice třídiagonální). V těchto případech je snadné uchovat pouze nenulové prvky (např. cca n) oproti všem (tj. cca n^2). Také násobení takovou maticí vyžaduje pouze přibližně n

operací oproti n^2 operací v případě plné matice. Připomeňme, že Gaussova eliminace vyžaduje přibližně n^3 operací.

Uvažujme nyní soustavu rovnic zapsanou ve tvaru

$$\mathbf{Ax} = \mathbf{b}. \quad (11)$$

Vidíme, že tuto soustavu lze např. ekvivalentně přepsat ($\omega \neq 0$)

$$\mathbf{x} = \mathbf{x} + \omega(\mathbf{b} - \mathbf{Ax}) = (\mathbf{E} - \omega\mathbf{A})\mathbf{x} + \omega\mathbf{b}.$$

Pro tento zápis (tzv. Richardsonova iterační metoda) se nabízí užít iterační proces, tj. aproximaci řešení dosadit do pravé strany a vypočítat nové přiblížení. Pokud tento proces bude konvergovat, pak bude konvergovat k řešení původní soustavy. Všimněme si, že jednu iteraci této metody jsme pro řídké soustavy schopni provést “rychle”, tj. potřebujeme cca n operací. Původní rovnici (11) lze ale přepsat do tohoto tvaru i jinými způsoby (jak uvidíme dále).

Problém 1.1 (Prostá iterační metoda) *Je-li soustava lineárních rovnic zapsaná ve tvaru $\mathbf{x} = \mathbf{Ux} + \mathbf{v}$, pak prostá iterační metoda (PIM) je dána předpisem*

$$\mathbf{x}^{(k+1)} = \mathbf{Ux}^{(k)} + \mathbf{v}. \quad (12)$$

Problematiku konvergence si osvětlíme na případě jedné rovnice, $x = qx + v$, kdy přesné řešení je dáno dle $x^* = (1 - q)^{-1}v$. Toto řešení pro případ $|q| < 1$ lze zapsat jako součet geometrické řady

$$x^* = \sum_{k=0}^{\infty} q^k v.$$

Pro určení přibližného řešení je ovšem možné sečíst pouze n členů.

$$x^{(n)} = \sum_{k=0}^n q^k v.$$

Zde snadno vidíme, že platí

$$x^{(n+1)} = qx^{(n)} + v$$

v případě volby $x^{(0)} = v$. Tedy vidíme, že posloupnost daná prostou iterační metodou (pro jednu rovnici) konverguje pokud $|q| < 1$. V případě obecné volby počáteční podmínky bude posloupnost také konvergentní (rozmyslete si).

Definice 1.6 *Říkáme, že iterační metoda pro danou soustavu lineárních rovnic konverguje, pokud pro libovolné $\mathbf{x}^{(0)}$ posloupnost $\{\mathbf{x}^{(k)}\}$ získaná předpisem této iterační metody konverguje k přesnému řešení této soustavy rovnic.*

Zabývejme se nyní konvergencí prosté iterační metody. Pro zjednodušení uvažujme, že matice \mathbf{U} je symetrická. Označme \mathbf{x}^* přesné řešení soustavy.

$$\mathbf{x}^* = \mathbf{U}\mathbf{x}^* + \mathbf{v} \quad (13)$$

Označíme $\mathbf{e}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x}^*$ a odečtením rovnici (13) od rovnice (12) dostaneme vztah $\mathbf{e}^{(k+1)} = \mathbf{U}\mathbf{e}^{(k)}$. Užijeme Tvrzení 0.1 a chybu $\mathbf{e}^{(0)}$ zapíšeme jako lineární kombinaci vlastních vektorů \mathbf{w}_j , tedy

$$\mathbf{e}^{(0)} = \sum_{j=1}^n \alpha_j \mathbf{w}_j.$$

Pak vidíme, že

$$\mathbf{e}^{(k)} = \mathbf{U}^k \left(\sum_{j=1}^n \alpha_j \mathbf{w}_j \right) = \sum_{j=1}^n \alpha_j \lambda_j^k \mathbf{w}_j,$$

neboť násobení maticí je lineární operace a vektory \mathbf{w}_j jsou vlastní vektory matice \mathbf{U} . Tedy konvergence $\mathbf{e}^{(k)} \rightarrow 0$ je ekvivalentní konvergenci $|\lambda_j|^k \rightarrow 0$ pro $k \rightarrow \infty$. Vidíme, že pro konvergenci je nutné i postačující aby $|\lambda_j| < 1$ pro všechna j . Dostáváme nutnou a postačující podmínku konvergence PIM.

Věta 1.4 (Nutná a postačující podmínka PIM) *Prostá iterační metoda je konvergentní právě tehdy, když $\rho(\mathbf{U}) < 1$.*

Vzhledem ke vztahu mezi normou a spektrálním poloměrem (7) vidíme, že přímým důsledkem předchozí věty je postačující podmínce pro konvergenci PIM.

Věta 1.5 *Je-li $\|\mathbf{U}\| < 1$, pak je prostá iterační metoda konvergentní (k přesnému řešení \mathbf{x}^*) Navíc platí následující odhad chyby*

$$\|\mathbf{x}^{(k)} - \mathbf{x}^*\| \leq \frac{\|\mathbf{U}\|}{1 - \|\mathbf{U}\|} \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|, \quad \|\mathbf{x}^{(k)} - \mathbf{x}^*\| \leq \frac{\|\mathbf{U}\|^k}{1 - \|\mathbf{U}\|} \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|.$$

První tvrzení věty 1.5 je zřejmý důsledek věty 1.4 o nutné a postačující podmínce konvergence PIM, neboť pokud je norma matice \mathbf{U} menší než jedna, pak také pro spektrální poloměr (vzhledem ke vztahu spektrálního poloměru a normy) platí $\rho(\mathbf{U}) \leq \|\mathbf{U}\| < 1$.

Odhad ve větě 1.5 získáme následujícím postupem: Odečtením (12) od rovnice $\mathbf{x}^* = \mathbf{U}\mathbf{x}^* + \mathbf{v}$ dostáváme konvergenci díky odhadu chyby $\mathbf{e}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x}^*$

$$\|\mathbf{e}^{(k+1)}\| \leq \|\mathbf{U}\| \|\mathbf{e}^{(k)}\| \quad \text{a tedy} \quad \|\mathbf{e}^{(k+1)}\| \leq \|\mathbf{U}\|^{k+1} \|\mathbf{e}^{(0)}\|. \quad (14)$$

Odhad odvodíme pomocí vztahu ($\mathbf{e}^0 = \mathbf{x}^* - \mathbf{x}^{(0)}$, $\mathbf{x}^* = \lim_{k \rightarrow +\infty} \mathbf{x}^{(k)}$)

$$\|\mathbf{e}^{(0)}\| = \left\| \lim_{k \rightarrow +\infty} \mathbf{x}^{(k+1)} - \mathbf{x}^{(0)} \right\| = \left\| \sum_{k=0}^{+\infty} (\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) \right\| \leq \sum_{k=0}^{+\infty} \|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| \quad (15)$$

a

$$\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| = \|\mathbf{U}(\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)})\| = \|\mathbf{U}^k(\mathbf{x}^{(1)} - \mathbf{x}^{(0)})\| \leq \|\mathbf{U}\|^k \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|. \quad (16)$$

Dosazením rovnice (16) do (15), užitím vzorce pro součet geometrické řady a dosazením za $\mathbf{e}^{(0)}$ do druhého vztahu (14) dostáváme druhý odhad z tvrzení věty. První odhad je jen jeho speciálním případem (rozmyslete si).

2 Jacobiova a Gaussova-Seidelova iterační metoda

Jacobiova a Gaussova-Seidelova iterační metoda jsou speciálním případem převodu soustavy rovnic z tvaru

$$\mathbf{Ax} = \mathbf{b}$$

na tvar potřebný pro prostou iterační metodu. Tento převod se provádí v zápise po složkách spíše než v maticovém tvaru (ten ale použijeme později pro analýzu). Ukážeme si tento postup obecně pro soustavu rovnic o 3 neznámých (z 1. rovnice vyjádříme neznámou x_1 z 2. rovnice neznámou x_2 atd.)

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 & x_1 &= \frac{1}{a_{11}}(b_1 - a_{12}x_2 - a_{13}x_3) \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 & \rightarrow x_2 &= \frac{1}{a_{22}}(b_2 - a_{21}x_1 - a_{23}x_3) \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 & x_3 &= \frac{1}{a_{33}}(b_3 - a_{31}x_1 - a_{32}x_2) \end{aligned} \quad (17)$$

Soustava rovnic na pravé straně má již tvar soustavy zapsané ve tvaru pro prostou iterační metodu, stačí na levou resp. pravou stranu doplnit $k+1$ -ní nebo k -tou iteraci. Pro případ soustavy rovnic o n neznámých dostáváme následující vzorec.

Definice 2.1 (Složkový zápis Jacobiovy iterační metody) *Jacobiova iterační metoda pro řešení soustavy rovnic $\mathbf{Ax} = \mathbf{b}$ je dána předpisem pro výpočet složek $\mathbf{x}^{(k+1)}$ z předchozí iterace $\mathbf{x}^{(k)}$ dle*

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j<i} a_{ij}x_j^{(k)} - \sum_{j>i} a_{ij}x_j^{(k)} \right) \quad \text{pro všechna } i = 1, \dots, n. \quad (18)$$

Je zřejmé, že rovnice (18) lze ekvivalentně zapsat také jako

$$x_i^{(k+1)} = x_i^{(k)} + \frac{1}{a_{ii}} \left(b_i - \sum_j a_{ij}x_j^{(k)} \right).$$

Zápis (18) odpovídá prosté iterační metodě. Chceme dále vyšetřovat podmínky konvergence Jacobiovy iterační metody s využitím znalostí o konvergenci PIM. Vidíme, že budeme potřebovat zapsat iterační postup (18) v maticové podobě. Využijeme tzv. rozštěpení matice \mathbf{A} na diagonální matici \mathbf{D} a ostře trojúhelníkové matice \mathbf{L} (dolní) a \mathbf{P} (horní), tedy

$$\mathbf{A} = \mathbf{L} + \mathbf{D} + \mathbf{P}, \quad (19)$$

kde \mathbf{D} je diagonální matice s prvky $d_{ii} = a_{ii}$, \mathbf{L} je ostře dolní trojúhelníková matice s nenulovými prvky $l_{ij} = a_{ij}$ pro $i > j$, a matice \mathbf{P} je ostře horní trojúhelníková matice s nenulovými prvky $u_{ij} = a_{ij}$ pro $i < j$.

Lemma 2.1 (Maticový zápis Jacobiovy iterační metody) *Nechť $\mathbf{A} = \mathbf{L} + \mathbf{D} + \mathbf{P}$ je rozštěpení matice dané rovnicí (19). Pak rovnice (18) se dají zapsat jako*

$$\mathbf{x}^{(k+1)} = \mathbf{U}_J \mathbf{x}^{(k)} + \mathbf{v}_J, \quad \mathbf{U}_J = -\mathbf{D}^{-1}(\mathbf{L} + \mathbf{P}), \quad \mathbf{v}_J = \mathbf{D}^{-1}\mathbf{b}. \quad (20)$$

Důkaz 2.1 *Rovnice (18) se dají zapsat maticově (vynásobením i -té rovnice d_{ii}) jako*

$$\mathbf{D}\mathbf{x}^{k+1} = -(\mathbf{L} + \mathbf{P})\mathbf{x}^k + \mathbf{b}.$$

Lemma 2.2 *Číslo $\lambda \in \mathbb{C}$ je vlastní číslo matice \mathbf{U}_J právě když je kořenem rovnice*

$$\det(\mathbf{L} + \lambda\mathbf{D} + \mathbf{P}) = 0. \quad (21)$$

Důkaz 2.2 *Vlastní číslo λ matice \mathbf{U}_J splňuje*

$$\begin{aligned} 0 &= \det(\mathbf{U}_J - \lambda\mathbf{E}) = \det(-\mathbf{D}^{-1}(\mathbf{L} + \mathbf{P}) - \lambda\mathbf{E}) \\ &= \det(-\mathbf{D}^{-1}(\mathbf{L} + \mathbf{P}) - \lambda\mathbf{D}^{-1}\mathbf{D}) = \det(-\mathbf{D}^{-1}(\mathbf{L} + \mathbf{P} + \lambda\mathbf{D})). \end{aligned}$$

Vzhledem k tomu, že determinant součinu matic je součin determinantů jednotlivých matic, a determinant matice \mathbf{D}^{-1} je nenulový, je tvrzení lemmatu dokázáno.

Věta 2.1

Je-li matice \mathbf{A} ODD, pak je Jacobiova iterační metoda konvergentní.

Důkaz 2.3 *Důkaz uvedeme nejprve jen pro ODD v řádcích, tedy platí $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$ pro všechna i . Vzhledem k tomu jak je daná matice \mathbf{U}_J vidíme, že*

$$\|\mathbf{U}_J\|_\infty = \max_i \left(\frac{1}{|a_{ii}|} \sum_{j \neq i} |a_{ij}| \right) < 1.$$

Důkaz konvergence pro obecnou ODD matici pak provedeme zobecněním pojmu ODD i pro komplexní matice (tedy matice s komplexními prvky).

Dále: Je-li matice $\mathbf{A} = \mathbf{L} + \mathbf{D} + \mathbf{P}$ ODD, pak pro $|\lambda| \geq 1$ je také matice $\mathbf{L} + \lambda\mathbf{D} + \mathbf{P}$ ODD. Ověříme snadno, neboť

$$|\lambda||a_{ii}| \geq |a_{ii}|.$$

Je-li ovšem matice $\mathbf{L} + \lambda\mathbf{D} + \mathbf{P}$ ODD, je také regulární.

Regularita matice $\mathbf{L} + \lambda\mathbf{D} + \mathbf{P}$ pro $|\lambda| \geq 1$ ale znamená, že $|\lambda| \geq 1$ není kořenem charakteristické rovnice a není tedy vlastním číslem matice \mathbf{U}_J .

Všchna vlastní čísla matice \mathbf{U}_J jsou tedy menší než 1 (a tedy i spektrální poloměr $\rho(\mathbf{U}_J) < 1$).

Je splněna nutná a postačující podmínka pro konvergenci Jacobiovy iterační metody (viz následující věta), metoda je tedy konvergentní.

Věta 2.2

Jacobiova iterační metoda je konvergentní právě tehdy, když $\rho(\mathbf{U}_J) < 1$.

Věta je důsledkem nutné a postačující podmínky konvergence prosté iterační metody a maticového zápisu Jacobiovy metody.

2.1 Gaussova-Seidelova iterační metoda

Pro odvození Gaussovy-Seidelovy iterační metody se postupuje obdobně jako v případě (17). Soustavu rovnic přepíšeme stejným způsobem. Nyní si ale všimneme, že při výpočtu v pořadí odshora dolů můžeme do 2. a 3. rovnice dosadit již hodnoty z nové iterace. Dostáváme iterační proces

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 & x_1^{(k+1)} &= \frac{1}{a_{11}} \left(b_1 - a_{12}x_2^{(k)} - a_{13}x_3^{(k)} \right) \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 & \rightarrow x_2^{(k+1)} &= \frac{1}{a_{22}} \left(b_2 - a_{21}x_1^{(k+1)} - a_{23}x_3^{(k)} \right) \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 & x_3^{(k+1)} &= \frac{1}{a_{33}} \left(b_3 - a_{31}x_1^{(k+1)} - a_{32}x_2^{(k+1)} \right) \end{aligned} \quad (22)$$

Pro obecnou soustavu rovnic pak Gaussova-Seidelova metoda lze zapsat následujícím vzorcem.

Definice 2.2 (Složkový zápis Gaussovy-Seidelovy iterační metody)

Gaussova-Seidelova iterační metoda pro řešení soustavy rovnic $\mathbf{Ax} = \mathbf{b}$ je dána předpisem pro výpočet složek $\mathbf{x}^{(k+1)}$ z předchozí iterace $\mathbf{x}^{(k)}$ dle

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j<i} a_{ij}x_j^{(k+1)} - \sum_{j>i} a_{ij}x_j^{(k)} \right). \quad (23)$$

Zde je vhodné připomenout, že pro řídké matice je zápis (23) čistě formální, součet probíhá pouze přes nenulové prvky matice. Stejně jako pro Jacobiho iterační metodu zapíšeme i Gaussovu-Seidelovu iterační metodu v maticové podobě.

Lemma 2.3 (Maticový zápis Gaussovy-Seidelovy iterační metody) *Nechť $\mathbf{A} = \mathbf{L} + \mathbf{D} + \mathbf{P}$ je rozštěpení matice dané rovnicí (19). Pak se rovnice (23) dají zapsat v maticovém zápisu jako*

$$\mathbf{x}^{(k+1)} = \mathbf{U}_{GS} \mathbf{x}^{(k)} + \mathbf{v}_{GS}, \quad \mathbf{U}_{GS} = -(\mathbf{D} + \mathbf{L})^{-1} \mathbf{P}, \quad \mathbf{v}_{GS} = (\mathbf{D} + \mathbf{L})^{-1} \mathbf{b}.$$

Důkaz 2.4 *Rovnice (23) lze zapsat ve tvaru*

$$(\mathbf{D} + \mathbf{L})\mathbf{x}^{(k+1)} = -\mathbf{P}\mathbf{x}^k + \mathbf{b}.$$

Lemma 2.4 *Vlastní čísla matice U_{GS} jsou kořeny rovnice*

$$\det(\lambda \mathbf{L} + \lambda \mathbf{D} + \mathbf{P}) = 0$$

Důkaz 2.5 *Vlastní číslo λ matice U_{GS} splňuje*

$$\begin{aligned} 0 &= \det(\mathbf{U}_{GS} - \lambda \mathbf{E}) = \det(-(\mathbf{D} + \mathbf{L})^{-1} \mathbf{P} - \lambda \mathbf{E}) \\ &= \det(-(\mathbf{D} + \mathbf{L})^{-1} \mathbf{P} - \lambda (\mathbf{D} + \mathbf{L})^{-1} (\mathbf{D} + \mathbf{L})) = \det(-(\mathbf{D} + \mathbf{L})^{-1} (\lambda \mathbf{L} + \lambda \mathbf{D} + \mathbf{P})). \end{aligned}$$

Vzhledem k tomu, že determinant součinu matic je součin determinantů jednotlivých matic, a determinant matice $(\mathbf{D} + \mathbf{L})^{-1}$ je nenulový, je tvrzení lemmatu dokázáno.

Věta 2.3 (Postačující podmínky Gaussovy-Seidelovy iterační metody)

Je-li matice \mathbf{A} ODD, pak je Gaussova-Seidelova iterační metoda konvergentní.

Je-li matice \mathbf{A} SPD, pak je Gaussova-Seidelova iterační metoda konvergentní.

Důkaz 2.6 *Bez důkazu.*

Věta 2.4 (Nutná a postačující podmínka Gaussovy-Seidelovy iterační metody)

Gaussova-Seidelova iterační metoda je konvergentní právě tehdy, když $\rho(\mathbf{U}_{GS}) < 1$.

Důkaz 2.7 *Věta je snadným důsledkem nutné a postačující podmínky konvergence prosté iterační metody a maticového zápisu Gauss-Seidelovy iterační metody.*

3 Minimalizační problémy

3.1 Soustava lineárních rovnic a její řešení jako minimalizační problém

V této části se budeme zabývat řešením soustavy lineárních rovnic v případě, že matice $\mathbf{A} \in \mathbb{R}^{n \times n}$ je symetrická a pozitivně definitní. Namísto toho, abychom hledali řešení soustavy rovnic, budeme hledat minimum funkce F n -proměnných dané jako

$$F(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{x}^T \mathbf{b}.$$

Pro takovou funkci (někdy též funkcionál) se snadno - dosazením a užitím symetrie matice \mathbf{A} - ukáže, že platí vztah

$$F(\mathbf{x} + \alpha \mathbf{v}) - F(\mathbf{x}) = \alpha \mathbf{v}^T (\mathbf{A} \mathbf{x} - \mathbf{b}) + \frac{1}{2} \alpha^2 \mathbf{v}^T \mathbf{A} \mathbf{v} \quad (24)$$

pro libovolné vektory \mathbf{x} , \mathbf{v} a číslo $\alpha \in \mathbb{R}$. Důkaz: DCV.

Věta 3.1 (Řešení soustavy rovnic jako minimalizace funkce) *Nechť matice \mathbf{A} je symetrická a pozitivně definitní. Pak*

$$\mathbf{A} \mathbf{x}^* = \mathbf{b} \quad \Leftrightarrow \quad F(\mathbf{x}^*) \leq F(\mathbf{x}) \quad \forall \mathbf{x}$$

Důkaz 3.1 *Užitím vztahu (24) pro $\mathbf{x} = \mathbf{x}^*$ ihned vidíme, že pokud \mathbf{x}^* je řešením soustavy, pak $F(\mathbf{x}) \geq F(\mathbf{x}^*)$, neboť matice \mathbf{A} je pozitivně definitní, tedy $\mathbf{v}^T \mathbf{A} \mathbf{v} \geq 0$.*

Naopak, je-li \mathbf{x}^ minimem funkcionálu, pak funkce $g(\alpha) = F(\mathbf{x}^* + \alpha \mathbf{v}) - F(\mathbf{x}^*)$ má minimum pro $\alpha = 0$. Platí $g'(0) = 0$ a $\mathbf{v}^T (\mathbf{A} \mathbf{x}^* - \mathbf{b}) = 0$ pro libovolné \mathbf{v} . Vektor $\mathbf{A} \mathbf{x}^* - \mathbf{b}$ musí být nulový.*

Metoda největšího spádu. Iterační metoda největšího spádu pro řešení soustavy rovnic $\mathbf{A} \mathbf{x} = \mathbf{b}$ je dána předpisem

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{p}^{(k)},$$

kde směr $\mathbf{p}^{(k)}$ je volen jako směr největšího spádu funkce $F(\mathbf{x})$ a krok je volen jako optimální krok ve směru \mathbf{p} .

Ze vztahu (24) vidíme, že směr největšího spádu je dán jako směr

$$\mathbf{p}^{(k)} = \mathbf{b} - \mathbf{A} \mathbf{x}^{(k)}.$$

Užitím pravé strany rovnice (24) jako funkce α lze snadno určit minimum užitím derivace, viz následující lemma.

Lemma 3.1 (Optimální krok v metodě největšího spádu) *Optimální krok v metodě největšího spádu je dán vztahem*

$$\alpha_k = \frac{\mathbf{p}^T(\mathbf{b} - \mathbf{A}\mathbf{x})}{\mathbf{p}^T(\mathbf{A}\mathbf{p})}$$

Důkaz 3.2 *Vezmeme funkci $g(\alpha) = F(\mathbf{x}^{(k)} + \alpha\mathbf{p}) - F(\mathbf{x}^{(k)})$ viz rovnice (24). Spočteme její derivaci a položíme rovnou nule. Ze vztahu $g'(\alpha) = 0$ již vyjádříme optimální α .*

3.2 Metoda nejmenších čtverců

Kvadratická odchylka a metoda nejmenších čtverců Metoda nejmenších čtverců se zabývá problémem aproximace dané tabulky hodnot pomocí funkce zvoleného typu, např. polynomu. Nechť je dána tabulka dat x_i, y_i pro $i = 1, \dots, n$. Kvadratickou odchylkou polynomu $p(x)$ rozumíme

$$\delta^2(p(x)) = \sum_{i=1}^n (p(x_i) - y_i)^2.$$

Říkáme, že polynom $p^*(x)$ stupně nejvýše q aproximuje danou tabulku dat nejlépe ve smyslu nejmenších čtverců, pokud

$$\delta^2(p^*(x)) \leq \delta^2(p(x)) \quad \text{pro } p(x) \text{ libovolný polynom stupně nejvýše } q.$$

Odvození soustavy normálních rovnic pro kvadratický polynom V této části si odvodíme, jakým způsobem se získá optimální kvadratický polynom pro danou tabulku dat. V tomto případě minimalizujeme kvadratickou odchylku polynomu $p(x) = a_0 + a_1x + a_2x^2$. Vidíme, že kvadratická odchylka je funkce pouze 3 neznámých

$$\delta^2(p(x)) = \sum_{i=1}^n (a_0 + a_1x_i + a_2x_i^2 - y_i)^2 = F(a_0, a_1, a_2).$$

Všimněme si, že se jedná o kvadratickou funkci (tedy spojitou funkci) v proměnných a_i , a že tato funkce je zdola omezená. Z toho již vyplývá, že nabývá minima. Nutná podmínka pro minimum nám dá tři rovnice

$$\frac{\partial F}{\partial a_0} = 2 \sum_{i=1}^n (a_0 + a_1x_i + a_2x_i^2 - y_i) = 0,$$

$$\frac{\partial F}{\partial a_1} = 2 \sum_{i=1}^n (a_0 + a_1x_i + a_2x_i^2 - y_i) x_i = 0,$$

$$\frac{\partial F}{\partial a_2} = 2 \sum_{i=1}^n (a_0 + a_1 x_i + a_2 x_i^2 - y_i) x_i^2 = 0,$$

z kterých po úpravě dostáváme soustavu normálních rovnic pro tento případ

$$\begin{aligned} a_0 \left(\sum_{i=1}^n 1 \right) + a_1 \left(\sum_{i=1}^n x_i \right) + a_2 \left(\sum_{i=1}^n x_i^2 \right) &= \left(\sum_{i=1}^n y_i \right), \\ a_0 \left(\sum_{i=1}^n x_i \right) + a_1 \left(\sum_{i=1}^n x_i^2 \right) + a_2 \left(\sum_{i=1}^n x_i^3 \right) &= \left(\sum_{i=1}^n x_i y_i \right), \\ a_0 \left(\sum_{i=1}^n x_i^2 \right) + a_1 \left(\sum_{i=1}^n x_i^3 \right) + a_2 \left(\sum_{i=1}^n x_i^4 \right) &= \left(\sum_{i=1}^n x_i^2 y_i \right). \end{aligned}$$

Vidíme, že jsme tento problém převedli na problém řešení soustavy lineárních rovnic s maticí, která je symetrická. Lze ukázat, že - za předpokladu na dostatečný počet různých bodů x_i - je i pozitivně definitní.

Odvození soustavy normálních rovnic pro obecný polynom Vzhledem k tomu, že polynom stupně nejvýše $q > 0$ má $q + 1$ neznámých, je třeba tyto určit z podmínky na nulové parciální derivace. Tím dostáváme soustavu tzv. normálních rovnic. Polynom $p^*(x)$ stupně nejvýše $q > 0$ aproximuje danou tabulku dat nejlépe ve smyslu nejmenších čtverců pokud $p^*(x) = \sum_{k=0}^q a_k x^k$, kde a_k jsou řešením soustavy lineárních rovnic

$$\begin{aligned} a_0 \left(\sum_i 1 \right) + a_1 \left(\sum_i x_i \right) + \dots + a_q \left(\sum_i x_i^q \right) &= \sum_i y_i, \\ a_0 \left(\sum_i x_i \right) + a_1 \left(\sum_i x_i^2 \right) + \dots + a_q \left(\sum_i x_i^{q+1} \right) &= \sum_i x_i y_i, \\ &\vdots \\ a_0 \left(\sum_i x_i^q \right) + a_1 \left(\sum_i x_i^{q+1} \right) + \dots + a_q \left(\sum_i x_i^{q+q} \right) &= \sum_i x_i^q y_i, \end{aligned}$$

K této soustavě lineárních rovnic lze dojít jednoduše pokud si kvadratickou odchylku zapíšeme ve vektorovém tvaru, tj.

$$\delta^2 = \|\mathbf{A}\alpha - \mathbf{y}\|_E^2,$$

kde matice \mathbf{A} je dána jako

$$\mathbf{A} = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^q \\ 1 & x_2 & x_2^2 & \dots & x_2^q \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^q \end{pmatrix}.$$

Máme tedy tvar

$$\delta^2(\alpha) = (\mathbf{A}\alpha - \mathbf{y})^T(\mathbf{A}\alpha - \mathbf{y}) = \mathbf{y}^T\mathbf{y} + \alpha^T\mathbf{A}^T\mathbf{A}\alpha - 2\mathbf{y}^T\mathbf{A}\alpha.$$

Víme, že matice $\mathbf{A}^T\mathbf{A}$ je symetrická a pozitivně definitní nebo pozitivně semidefinitní (toto závisí na počtu různých bodů x_i). Lze užít věty 3.1 a vidíme, že minimum této funkce je řešením soustavy lineárních rovnic

$$\mathbf{A}^T\mathbf{A}\alpha = \mathbf{A}^T\mathbf{y}.$$

O tomtéž se dá přesvědčit formálním derivováním podle složek α , kdy dostáváme

$$\frac{\partial\delta^2}{\partial\alpha_i} = 2e_i\mathbf{A}^T\mathbf{A}\alpha - 2e_i\mathbf{A}^T\mathbf{y} = 2e_i(\mathbf{A}^T\mathbf{A}\alpha - \mathbf{A}^T\mathbf{y}) = 0.$$

Vidíme, že

$$(\mathbf{A}^T\mathbf{A}\alpha - \mathbf{A}^T\mathbf{y}) = 0.$$

Poznámka 3.1 (Normální rovnice) *Soustava normálních rovnic se často zapisuje ve tvaru*

$$\mathbf{A}^T\mathbf{A}\alpha = \mathbf{A}^T\mathbf{b},$$

kde $\alpha = (a_0, \dots, a_q)^T$, $b_i = y_i$ a matice \mathbf{A} je dána jako

$$\mathbf{A} = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^q \\ 1 & x_2 & x_2^2 & \dots & x_2^q \\ \vdots & \vdots & & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^q \end{pmatrix}.$$

4 Nelineární rovnice a Newtonova metoda

Existence a jednoznačnost řešení soustav nelineárních rovnic. Víme, že i pro soustavu lineárních rovnic není vždy zaručena ani existence ani jednoznačnost řešení (regulární matici toto zaručuje). Existence a případně i jednoznačnost řešení soustav nelineárních rovnic je samozřejmě ještě komplikovanější (stačí si představit jednu nelineární rovnici $f(x) = 0$). Existence a jednoznačnost lze ukázat většinou jen ve speciálních případech. Často se ale používá speciální případ nelineární rovnice, tj. problém pevného bodu $\mathbf{x} = \mathbf{F}(\mathbf{x})$.

Definice 4.1 *Zobrazení (funkci) $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ nazýváme kontraktivní zobrazení na množině $D \subset \mathbb{R}^n$, pokud existuje $\kappa \in (0, 1)$ tak, že pro libovolné $\mathbf{x}, \mathbf{y} \in D$ platí*

$$\|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{y})\| < \kappa \|\mathbf{x} - \mathbf{y}\|.$$

Věta 4.1 (Banachova věta) *Nechť F je kontraktivní zobrazení na uzavřené neprázdné množině $D \subset \mathbb{R}^n$ a nechť $F(D) \subset D$. Pak existuje právě jeden pevný bod tohoto zobrazení.*

Ukažme si myšlenku důkazu pro $n = 1$. Nejprve ukážeme existenci pevného bodu. Sestrojíme posloupnost $\{x^{(k)}\} \subset \mathbb{R}$, která bude konvergovat k pevnému bodu. Volíme počáteční přiblížení $x^{(0)} \in D$ a sestrojíme posloupnost $x^{(k+1)} = F(x^{(k)})$. Vzhledem k tomu, že zobrazení je kontraktivní vidíme, že

$$|x^{(k+1)} - x^{(0)}| \leq \sum_{j=0}^k |x^{(j+1)} - x^{(j)}| \leq \sum_{j=0}^k \kappa^j |x^{(1)} - x^{(0)}|,$$

kde pravá strana je shora omezená součtem nekonečné geometrické řady, tedy

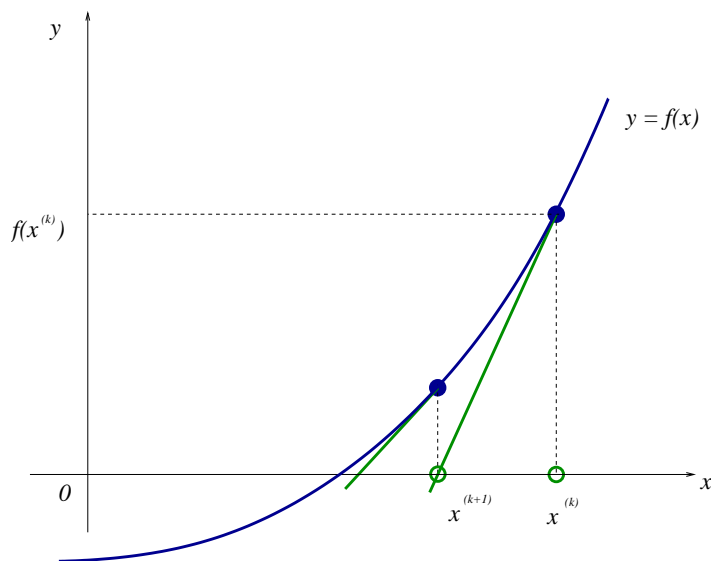
$$|x^{(k+1)} - x^{(0)}| \leq \frac{1}{1 - \kappa} |x^{(1)} - x^{(0)}|.$$

Posloupnost $\{x_k\}$ je omezená, tudíž z ní lze vybrat konvergentní podposloupnost. Pak lze snadno ukázat, že i celá posloupnost je konvergentní k téže limitě. Limitou této posloupnosti je hledaný pevný bod.

Jednoznačnost: vezmeme dva pevné body $x^* = F(x^*)$, $y^* = F(y^*)$. Spočtíme velikost rozdílu a uijme kontraktivitu F , tedy

$$|x^* - y^*| = |F(x^*) - F(y^*)| \leq \kappa |x^* - y^*|.$$

Z této nerovnosti ($\kappa < 1$) již vyplývá $x^* = y^*$.



Obrázek 1: Grafické znázornění Newtonovy metody

Newtonova metoda pro funkci jedné proměnné (1D) Zabýváme se dále řešením jedné nelineární rovnice $f(x) = 0$. Užijeme Newtonovu metodu, také zvanou metodu tečen, viz Obr. 1.

Hledáme x^* tak, že $f(x^*) = 0$. Užijeme Taylorův polynom v bodě x^k

$$0 = f(x^*) = f(x^{(k)}) + f'(x^{(k)})(x^* - x^{(k)}) + \frac{1}{2}f''(\xi)(x^* - x^{(k)})^2$$

Zanedbáním posledního členu (tj. rovnice tečny) dostáváme vzorec pro přiblížení $x^{(k+1)} \approx x^*$. Tedy

$$x^{(k+1)} = x^{(k)} - [f'(x^{(k)})]^{-1}f(x^{(k)}). \quad (25)$$

Newtonova metoda pro soustavu dvou rovnic o dvou neznámých (2D) Zabýváme se dále řešením soustavy nelineárních rovnic

$$f(x, y) = 0, \quad g(x, y) = 0.$$

Postupujeme obdobně jako pro 1D případ. Užijeme Taylorův polynom, kde zanedbáme členy vyššího řádu než jedna. V tomto případě je Taylorův polynom shodný s rovnicí tečné roviny. Tedy užijeme rovnici tečné roviny v bodě $A = [x^{(k)}, y^{(k)}]$ pro nahrazení funkce f a funkce g . Hledáme tedy řešení soustavy lineárních rovnic

$$0 = f(A) + f_x(A)(x - x^{(k)}) + f_y(A)(y - y^{(k)}),$$

$$0 = g(A) + g_x(A)(x - x^{(k)}) + g_y(A)(y - y^{(k)}),$$

kde f_x respektive f_y značíme parciální derivace funkce f podle x respektive y . Obdobně g_x, g_y . Řešení této soustavy rovnic označíme $x^{(k+1)}, y^{(k+1)}$. Dostaneme formální vzorec

$$\begin{pmatrix} x^{(k+1)} \\ y^{(k+1)} \end{pmatrix} = \begin{pmatrix} x^{(k)} \\ y^{(k)} \end{pmatrix} - \begin{pmatrix} f_x(A) & f_y(A) \\ g_x(A) & g_y(A) \end{pmatrix}^{-1} \begin{pmatrix} f(x^{(k)}, y^{(k)}) \\ g(x^{(k)}, y^{(k)}) \end{pmatrix}. \quad (26)$$

Všimněte si podobnosti se vzorcem (25). Zde je třeba zdůraznit, že inverzní matici nepočítáme, ale výpočet realizujeme pomocí řešení soustavy lineárních rovnic.

Newtonova metoda pro soustavu n -rovníc o n -neznámých Předchozí postup lze zobecnit pro řešení soustavy nelineárních rovnic

$$F(\mathbf{x}) = 0,$$

kde F je vektorová funkce $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$. I v tomto případě lze odvodit vzorec, který lze zapsat ve tvaru rovnice (25). Je ale nutné si uvědomit, co v tomto rozumíme výrazem F' . Jedná se o Jacobiho matici zobrazení F danou v následující definici.

Definice 4.2 (Jacobiho matice zobrazení F) *Jacobiho maticí zobrazení $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ rozumíme matici*

$$F'(\mathbf{x}) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(\mathbf{x}) & \frac{\partial F_1}{\partial x_2}(\mathbf{x}) & \dots & \frac{\partial F_1}{\partial x_n}(\mathbf{x}) \\ \frac{\partial F_2}{\partial x_1}(\mathbf{x}) & \frac{\partial F_2}{\partial x_2}(\mathbf{x}) & \dots & \frac{\partial F_2}{\partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1}(\mathbf{x}) & \frac{\partial F_n}{\partial x_2}(\mathbf{x}) & \dots & \frac{\partial F_n}{\partial x_n}(\mathbf{x}) \end{pmatrix}.$$

Definice 4.3 (Newtonova metoda) *Hledáme kořeny rovnice $F(\mathbf{x}) = \mathbf{0}$, $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ tak, že volíme počáteční přiblížení $\mathbf{x}^{(0)}$ a sestavíme posloupnost pomocí předpisu*

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - (F'(\mathbf{x}^{(k)}))^{-1} F(\mathbf{x}^{(k)}).$$

Věta 4.2 (Konvergence Newtonovy metody) *Nechť $F: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$, D - oblast, $F \in C^2(D)$. Nechť $F(x^*) = 0$ a $J = \det F'(x^*) \neq 0$. Pak existuje okolí bodu x^* takové, že Newtonova metoda je konvergentní k x^* pro libovolné $x^{(0)}$ z tohoto okolí.*

Důkaz 4.1 (Náznak důkazu pro $n = 1$) *Z předpisu pro Newtonovu metodu vidíme, že*

$$x^* - x^{(k+1)} = x^* - x^{(k)} + \frac{f(x^{(k)})}{f'(x^{(k)})}. \quad (27)$$

Užitím Taylorova polynomu s Lagrangeovým tvarem zbytku vidíme, že

$$\underbrace{f(x^*)}_0 = f(x^{(k)}) + f'(x^{(k)})(x^* - x^{(k)}) + \frac{f''(\xi)}{2}(x^* - x^{(k)})^2.$$

Dělením předchozí rovnice $f'(x^{(k)})$ dostaneme

$$0 = \frac{f(x^{(k)})}{f'(x^{(k)})} + (x^* - x^{(k)}) + \frac{f''(\xi)}{2f'(x^{(k)})}(x^* - x^{(k)})^2,$$

a tedy vzhledem ke vztahu (27) vidíme, že

$$|x^* - x^{(k+1)}| = \left| -\frac{f''(\xi)}{2f'(x^{(k)})}(x^* - x^{(k)})^2 \right| \leq M|x^* - x^{(k)}|^2,$$

kde $M = \max |f''(\xi)/(2f'(x^{(k)}))|$ - maximum je bráno na nějakém okolí bodu x^* (rozmyslete si proč je tento výraz konečný). Poslední nerovnost říká, že je-li $x^{(k)}$ dostatečně blízko přesnému řešení, tj. je-li $|x^* - x^{(k)}| \leq \frac{q}{M}$, pro nějaké $0 \leq q < 1$, pak také $|x^* - x^{(k+1)}| \leq \frac{q}{M}$ a navíc

$$|x^* - x^{(k+1)}| < q|x^* - x^{(k)}|.$$

Opakované použití poslední nerovnosti pak již dává konvergenci k řešení.

V předchozím důkazu byla použita věta o Taylorově polynomu z Matematiky I, připomeňme její znění. V této větě budeme symbolem $f^{(k)}(x)$ značit k -tou derivaci funkce f . Speciálně $f^{(1)}(x) = f'(x)$ a nultou derivací rozumíme přímo funkci $f^{(0)}(x) = f(x)$. Tato věta bude hrát velmi důležitou roli v další části přednášek!

Věta 4.3 (Taylorův polynom s Lagrangeovým tvarem zbytku, opakování) *Nechť f je funkce, která má na intervalu I spojité derivace až do řádu $n + 1$ včetně, tj. $f \in \mathcal{C}^{n+1}(I)$. Pak pro $x, x + h \in I$ platí*

$$f(x + h) = T_n(h) + R_{n+1}(h), \quad (28)$$

kde

$$T_n(h) = \sum_{k=0}^n \frac{f^{(k)}(x)}{k!} h^k, \quad R_{n+1}(h) = \frac{f^{(n+1)}(\xi)}{(n+1)!} h^{n+1}, \quad (29)$$

kde $\xi = \xi(h)$ je nějaký bod mezi body x a $x + h$ (zde h může být i záporné).

Tuto větu budeme speciálně používat pro případ $n = 3$, $f \in \mathcal{C}^4$, pro volbu $h := h$ a (formálně) $h := -h$, tedy

$$f(x + h) = f(x) + f'(x)h + \frac{1}{2}f''(x)h^2 + \frac{1}{3!}f'''(x)h^3 + \frac{1}{4!}f^{(IV)}(\xi)h^4$$

nebo

$$f(x - h) = f(x) - f'(x)h + \frac{1}{2}f''(x)h^2 - \frac{1}{3!}f'''(x)h^3 + \frac{1}{4!}f^{(IV)}(\xi)h^4.$$

Definice 4.4 (Symboly $\mathcal{O}(h^p)$, $\mathcal{O}(h^p + \tau^q)$)

Nechť pro funkci $g(h)$ definovanou pro všechna $h \in (0, h_{max})$ existuje konstanta $K > 0$ taková, že platí odhad

$$|g(h)| \leq Kh^p.$$

Pak užíváme zápis

$$g(h) = \mathcal{O}(h^p).$$

Nechť pro funkci $g(h, \tau)$ definovanou pro všechna $h \in (0, h_{max})$ a $\tau \in (0, \tau_{max})$ existují konstanty $K_1, K_2 > 0$ takové, že platí odhad

$$|g(h, \tau)| \leq K_1h^p + K_2\tau^q.$$

Pak užíváme zápis

$$g(h, \tau) = \mathcal{O}(h^p + \tau^q).$$

5 Numerické řešení Cauchyovy úlohy. Eulerova metoda.

V této části se budeme věnovat numerickým řešením obyčejných diferenciálních rovnic. Budeme se zabývat řešením úlohy nalezení funkce $y = y(x)$ splňující na intervalu $\langle a, b \rangle$ diferenciální rovnici

$$y' = f(x, y). \quad (30)$$

Aby byla zaručena jednoznačnost řešení doplníme rovnici (30) počáteční podmínkou $y(a) = \eta$, tj. budeme hledat řešení Cauchyovy úlohy

$$y' = f(x, y), \quad y(a) = \eta. \quad (31)$$

V případě úlohy (31) mluvíme o počáteční nebo také Cauchyově úloze. Rovnice (30) může reprezentovat buď jednu obyčejnou diferenciální rovnici nebo také soustavu diferenciálních rovnic, kdy $y = (y_1, \dots, y_d)^T$ je vektorová funkce $y : \mathbb{R} \rightarrow \mathbb{R}^d$ a pravá strana $f = (f_1, \dots, f_d)^T$ je d -rozměrná vektorová funkce.

Poznámka 5.1 (O existenci a jednoznačnosti řešení)

Postačující podmínkou existence řešení úlohy (31) je spojitost funkce f a spojitost parciální derivace $\frac{\partial f}{\partial y}$. V případě soustavy rovnic pak spojitost parciálních derivací $\frac{\partial f_i}{\partial y_j}$ pro všechna i, j . Viz podrobněji Matematika III.

Poznámka 5.2 (Převod rovnice vyššího řádu)

Každá obyčejná diferenciální rovnice k -tého řádu ($k > 1$) se dá formulovat ve tvaru (30). Toto odvození si zopakujeme pro případ rovnice druhého řádu

$$y'' = G(x, y, y'),$$

kdy zavedeme novou funkci $\mathbf{z} = (z_1, z_2)^T$, jejíž složky jsou dány pomocí vztahu

$$z_1 = y, \quad z_2 = y' \quad \text{a tedy} \quad z'_1 = z_2, \quad z'_2 = G(x, z_1, z_2).$$

Zcela obdobně postupujeme pro případ rovnice k -tého řádu ($k > 2$), tj. například pro rovnici

$$y''' = G(x, y, y', y'')$$

volíme $\mathbf{z} = (z_1, z_2, z_3)^T = (y, y', y'')^T$.

Věta 5.1 (Náhrady 1. derivace) Nechť funkce $f \in C^2(I)$, interval $I = \langle a, b \rangle$, body $x, x \pm h \in I$. Pak platí

$$f'(x) = \frac{f(x+h) - f(x)}{h} + \eta_1(h), \quad f'(x) = \frac{f(x) - f(x-h)}{h} + \eta_2(h),$$

kde $\eta_i(h) = \mathcal{O}(h)$.

Důkaz 5.1 Věta je okamžitým důsledkem věty o Taylorově polynomu s Lagrangeovým tvarem zbytku, viz věta 4.3 a poznámka pod ní.

Princip numerického řešení. Pro jednoduchost uvažujeme pouze skalární případ, tedy řešíme jednu diferenciální rovnici. Označme přesné řešení $y = y(x)$ pro $x \in \langle a, b \rangle$. Hodnoty přesného řešení hledáme v diskrétních bodech

$$a = x_0 < x_1 < \dots < x_N \leq b,$$

Obvykle řešení hledáme v bodech $x_n = a + nh$ daných pomocí kroku $h > 0$. Hodnoty přesného řešení pak aproximujeme hodnotami

$$y_n \approx y(x_n).$$

Přibližné hodnoty y_{n+1} můžeme získat například tak, že v rovnici (30) nahradíme derivaci y' v bodě $x = x_n$ nebo $x = x_{n+1}$ podle Věty 5.1. Tím dostaneme dvě varianty Eulerovy metody, které se často používají. První je tzv. Eulerova metoda daná předpisem

$$\frac{y_{n+1} - y_n}{h} = f(x_n, y_n),$$

z kterého je možné vyjádřit explicitně y_{n+1} (také nazývaná explicitní Eulerova metoda).

Definice 5.1 (Explicitní Eulerova metoda) Je dána Cauchyova úloha (31). Volíme $y_0 = \eta$, krok $h > 0$ a hodnoty přibližného řešení y_n vypočteme předpisem

$$y_{n+1} = y_n + hf(x_n, y_n).$$

Říkáme, že hodnoty y_n jsou vypočtené užitím (explicitní) Eulerovy metody.

Druhá metoda pak je tzv. implicitní Eulerova metoda daná předpisem

$$\frac{y_{n+1} - y_n}{h} = f(x_{n+1}, y_{n+1}). \quad (32)$$

Definice 5.2 (Implicitní Eulerova metoda) Je dána Cauchyova úloha (31). Volíme $y_0 = \eta$, krok $h > 0$ a vypočteme přibližné hodnoty y_n , pro které platí

$$y_{n+1} - hf(x_{n+1}, y_{n+1}) = y_n$$

Říkáme, že přibližné hodnoty y_n jsou vypočtené užitím implicitní Eulerovy metody.

V tomto případě je nutné hledat řešení této obecně nelineární rovnice. Implicitní metoda se nicméně často používá v případě lineárních ODR, tj. v případě kdy $f(x, Y) = \mathbf{A}Y + g(x)$, kde \mathbf{A} je matice. V takovém případě je (32) soustavou lineárních rovnic, kterou je možné (alespoň pro malá h) snadno vyřešit, tj.

$$(\mathbf{E} - h\mathbf{A})y_{n+1} = y_n + hg(x).$$

Příklad 5.1 Je dáno $h > 0$, $D > 0$, $A > 0$ a Cauchyova úloha $y' = -Ay$, $y(0) = D$. Užitím explicitní a implicitní Eulerovy metody a kroku h spočítáme aproximaci hodnot řešení v bodech $x_j = jh$, $j = 1, 2, 3$ a $j = n$.

Určíme přesné řešení $y(x)$ dané Cauchyovy úlohy a srovnáme s přibližným řešením.

Řešení: Nejprve řešme tento příklad pomocí explicitní Eulerovy metody. Hodnotu v bodě x_1 tedy spočteme dle vzorce

$$y_1 = y_0 + hf(x_0, y_0) = (1 - Ah)y_0 = (1 - Ah)D,$$

a dále obdobně

$$y_2 = y_1 + hf(x_1, y_1) = (1 - Ah)y_1 = (1 - Ah)^2 D,$$

a

$$y_3 = y_2 + hf(x_2, y_2) = (1 - Ah)y_2 = (1 - Ah)^3 D.$$

Obecně tedy máme

$$y_j = (1 - Ah)^j D.$$

Obdobně pro implicitní Eulerovu metodu (hodnoty aproximací označíme z_i , $z_0 = y_0 = D$)

$$z_1 - hf(x_1, z_1) = z_0, \quad \rightarrow (1 + Ah)z_1 = D,$$

a tedy $z_1 = (1 + Ah)^{-1}D$. Obdobně dostaneme

$$z_j = \frac{1}{(1 + Ah)^j} D.$$

Všimněme si, že přesné řešení dané úlohy je $y(x) = De^{-Ax}$ a že toto řešení je klesající pro $x \in \mathbb{R}$ ($A > 0$). Vidíme, že také hodnoty numerické aproximace z_j získané implicitní metodou jsou vždy klesající nezávisle na volbě $h > 0$. Aby hodnoty numerické aproximace y_i pro explicitní metodu splňovali totéž, je nutné aby $|1 - Ah| \leq 1$, tedy aby hodnoty Ah leželi v intervalu $\langle 0, 2 \rangle$. Uvažujeme-li hodnoty A komplexní, pak hodnoty Ah musí ležet v kruhu o poloměru 1 a středu $1 + 0i \in \mathbb{C}$. Dá se také snadno ukázat, že hodnoty získané ať už explicitní nebo implicitní metodou konvergují k hodnotám přesného řešení (zkuste si pro konkrétní volbu x a různá h), např. volba $x = 1 = x_n$ a $h = 1/n$ dává

$$\lim_{n \rightarrow \infty} y_n = \lim_{n \rightarrow \infty} \left(1 - \frac{A}{n}\right)^n = e^{-A}$$

a případně

$$\lim_{n \rightarrow \infty} z_n = \lim_{n \rightarrow \infty} \frac{1}{\left(1 + \frac{A}{n}\right)^n} = \frac{1}{e^A}.$$

V případě užití numerických metod pro vektorovou Cauchyovu úlohu, je třeba být opatrný vzhledem k tomu, že hodnoty y_n jsou vektory, které mají složky. Abychom odlišili hodnoty aproximace řešení v bodě x_n od hodnot jeho složek, značíme tyto aproximace někdy horním indexem namísto dolního, tedy $y^{(n)} \approx y(x_n)$ případně $Y^{(n)} \approx Y(x_n)$. Vzorec pro explicitní Eulerovu metodu pak zapisujeme

$$Y^{(n+1)} = Y^{(n)} + hf(x_n, Y^{(n)}).$$

Je zřejmé, že na postupu výpočtu toto značení nic nemění.

Příklad 5.2 Je dána Cauchyova úloha pro soustavu ODR

$$Y' = \begin{pmatrix} y_1 - y_2^2 \\ x - y_1 \end{pmatrix}, \quad Y(0) = (1, 1)^T, \quad Y = (y_1, y_2)^T$$

Užijte explicitní Eulerovy metody s krokem $h = 0.5$ a spočtěte hodnoty aproximace řešení $Y(0.5)$, $Y(1)$, $Y(2)$.

Řešení: V daném případě se jedná o soustavu nelineárních ODR prvního řádu, kde pravá strana je zadána jako vektorová funkce

$$f(x, Y) = \begin{pmatrix} y_1 - y_2^2 \\ x - y_1 \end{pmatrix}.$$

Všimněme si, že pravá strana f této soustavy ODR je spojitá a parciální derivace $\frac{\partial f_1}{\partial y_1}$, $\frac{\partial f_2}{\partial y_1}$, $\frac{\partial f_1}{\partial y_2}$ a také $\frac{\partial f_2}{\partial y_2}$. Počítáme tedy pro $Y^{(0)} = (1, 1)^T$ dle vzorce pro explicitní metodu

$$Y^{(1)} = Y^{(0)} + 0.5 f\left(0, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 0.5 \begin{pmatrix} 1 - 1^2 \\ 0 - 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0.5 \end{pmatrix}$$

a dále

$$Y^{(2)} = Y^{(1)} + 0.5 f\left(0.5, \begin{pmatrix} 1 \\ 0.5 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 0.5 \end{pmatrix} + 0.5 \begin{pmatrix} 1 - 0.5^2 \\ 0.5 - 1 \end{pmatrix} = \begin{pmatrix} 0.625 \\ 0.25 \end{pmatrix}.$$

$$Y^{(3)} = Y^{(2)} + 0.5 f\left(1, \begin{pmatrix} 0.625 \\ 0.25 \end{pmatrix}\right) = \begin{pmatrix} 0.625 \\ 0.25 \end{pmatrix} + 0.5 \begin{pmatrix} 0.625 - 0.25^2 \\ 1 - 0.625 \end{pmatrix} = \begin{pmatrix} 0.90625 \\ 0.4375 \end{pmatrix}.$$

Příklad 5.3 Nechť $\mathbf{A} \in \mathbb{R}^{d \times d}$, $\mathbf{u} \in \mathbb{R}^d$ je vlastní vektor matice \mathbf{A} příslušný reálnému vlastnímu číslu $\lambda < 0$. Je dána Cauchyova úloha

$$Y' = \mathbf{A}Y, \quad Y(0) = \mathbf{u},$$

Užitím explicitní a implicitní Eulerovy metody a kroku h spočítáme aproximaci řešení v bodech $x_j = jh$, $j = 1, 2, 3$ a $j = n$.

Určíme přesné řešení $y(x)$ dané Cauchyovy úlohy a srovnáme s přibližným řešením.

Uvážíme vlastní číslo λ a vlastní vektor \mathbf{u} komplexní, $\operatorname{Re} \lambda < 0$.

Řešení: V tomto případě postupujeme zcela obdobně jako v příkladě předchozím. Nejprve si všimněme, že funkce $Y(x) = e^{\lambda x} \mathbf{u}$ je přesným řešením dané Cauchyovy úlohy (ověřte si dosazením!). Dále aplikujeme explicitní Eulerovu metodu a užitím definice vlastního čísla a vektoru. Užitím zápisu s horními indexy, tedy $Y^{(0)} = \mathbf{u}$.

$$Y^{(1)} = Y^{(0)} + h\mathbf{A}Y^{(0)} = \mathbf{u} + h\mathbf{A}\mathbf{u} = (\mathbf{E} + h\mathbf{A})\mathbf{u} = (1 + \lambda h)\mathbf{u},$$

$$Y^{(2)} = Y^{(1)} + h\mathbf{A}Y^{(1)} = Y^{(1)} + h\mathbf{A}Y^{(1)} = (\mathbf{E} + h\mathbf{A})Y^{(1)} = (1 + \lambda h)^2\mathbf{u},$$

a tedy

$$Y^{(j+1)} = Y^{(j)} + h\mathbf{A}Y^{(j)} = (1 + \lambda h)^{j+1}\mathbf{u}.$$

Všimněme si podobnosti tohoto řešení a řešení z příkladu 5.1. Vidíme tedy také, že hodnoty aproximace řešení budou konvergovat k hodnotám přesného řešení. Hodnota kroku h by měla být volena tak, aby chování numerického řešení bylo obdobné jako chování přesného řešení, které v čase klesá ($\lambda < 0$). Tedy dostáváme podmínku

$$|1 + \lambda h| \leq 1,$$

a vzhledem ke znaménku potřebujeme splnit podmínku $h < -2/\lambda$.

Zde je ale nutné poznamenat, že v praktickém výpočtu pro tento konkrétní problém je třeba uvážit v určitých speciálních případech vliv zaokrouhlovacích chyb. Např. zaokrouhlením může z počáteční podmínky $Y(0) = \mathbf{u}$ vzniknout $Y(0) = \mathbf{u} + \epsilon \mathbf{v}$, kde ϵ je nějaké velmi malé číslo. Bude-li \mathbf{v} např. vlastní vektor příslušný jinému vlastnímu číslu λ_1 , kde $\lambda_1 > 0$ (víme dle MA III typ bodu rovnováhy pro danou ODR je sedlo). Přesné řešení této “zaokrouhlené” úlohy pak je $\tilde{Y}(x) = e^{\lambda x} \mathbf{u} + \epsilon e^{\lambda_1 x} \mathbf{v}$. Vidíme tedy, že pro velká $x > 0$ je první člen velmi malý ($\lambda x \rightarrow -\infty$ pro $x \rightarrow +\infty$), ale druhý člen naopak veliký a rostoucí ($\lambda_1 x \rightarrow +\infty$ pro $x \rightarrow +\infty$). Je zřejmé, že řešíme-li numericky danou úlohu, nemůžeme dobře aproximovat obě tyto řešení $Y(x)$ a $\tilde{Y}(x)$. Vidíme, že vliv zaokrouhlovací chyby v případě úlohy, kdy $\lambda < 0$ a $\lambda_1 > 0$, může být obrovský a nezávislý na volbě kroku h . Pokud ale všechna vlastní čísla jsou záporná, k takovéto situaci nedojde. Zdůrazněme tedy, že pro numerickou aproximaci řešení je nutné vědět něco nejen o vlastnostech numerických metod ale také i o řešení úlohy a vlastnostech jejího řešení (např. také i oblast, kde existuje jednoznačné řešení). Užití numerické metody v diskutovaném případě, kdy velmi malá změna počáteční podmínky vede k velké změně přesného řešení, je nebo může být nevhodné.

Věta 5.2 (Náhrada 1. derivace, centrální difference) *Nechť funkce $f \in C^3(I)$, interval $I = \langle a, b \rangle$, body $x, x \pm h \in I$. Pak platí*

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} + \eta(h),$$

kde $\eta(h) = \mathcal{O}(h^2)$.

Důkaz 5.2 *Věta je okamžitým důsledkem věty o Taylorově polynomu s Lagrangeovým tvarem zbytku, viz věta 4.3. Speciálně odečtením dvou rovnic za větou 4.3 dostaneme tvrzení věty 5.2.*

Užitím Věty 5.2 lze nyní zlepšit Eulerovu metodu: nahrazením derivace v bodě $x_{n+1/2} = x_n + h/2$ v ODR dostáváme

$$\frac{y_{n+1} - y_n}{h} \approx y'(x_{n+1/2}) = f(x_n + h/2, y(x_n + h/2)), \quad (33)$$

kde na pravé straně uijeme aproximaci pomocí Taylorova rozvoje

$$y(x_n + h/2) = y(x_n) + \frac{h}{2}y'(x_n) + \mathcal{O}(h^2) \approx y_n + \frac{h}{2}f(x_n, y_n). \quad (34)$$

Dosazením aproximace (34) do (33) a zanedbnáním členů $\mathcal{O}(h^2)$ dostáváme tzv. Collatzovu metodu (též metoda Rungeova-Kuttova).

Definice 5.3 (Collatzova metoda) *Je dána Cauchyova úloha (31). Volíme $y_0 = \eta$, krok $h > 0$ a vypočteme hodnoty aproximací Y^k předpisem*

$$y_{n+1} = y_n + hf \left(x_n + \frac{h}{2}, y_n + \frac{h}{2}f(x_n, y_n) \right) \quad (35)$$

Říkáme, že přibližné hodnoty y_n jsou vypočtené užitím Collatzovy metody.

Poznámka 5.3 *Collatzova metoda se také nazývá Rungeova-Kuttova metoda (2. řádu), případně 1. modifikací Eulerovy metody. Výpočet lze formálně v každém kroku n provádět postupným výpočtem*

$$\begin{aligned} k_1 &= f(x_n, y_n), \\ y_{pom} &= y_n + \frac{h}{2}k_1, \\ k_2 &= f \left(x_n + \frac{h}{2}, y_{pom} \right), \\ y_{n+1} &= y_n + hk_2. \end{aligned}$$

Hodnoty k_1 , k_2 a y_{pom} jsou tedy obecně různé v různých krocích n . Z tohoto důvodu se někdy užívá přesnější označení $k_1^{(n)}$, $k_2^{(n)}$ a $y_{pom}^{(n)}$. V tomto textu toto označení neužíváme a tyto vzorce bereme jen jako praktickou realizaci výpočtu vzorce (35) pro výpočet y_{n+1} z y_n .

Příklad 5.4 Řešte numericky Cauchyovu úlohu explicitní Eulerovou a Collatzovou metodou

$$y' = -x^2 + y^2, \quad y(0) = 0.1.$$

Řešení: Nejprve budeme numericky řešit Eulerovou metodou s krokem $h = 0.2$. Dle počáteční podmínky určíme $x_0 = 0$, $y_0 = 0.1$. Dále si všimněme, že pravé strana dané diferenciální rovnice je $f(x, y) = -x^2 + y^2$. Spočtěme hodnotu y_1 Eulerovou metodou, tedy

$$y_1 = y_0 + h f(x_0, y_0) = 0.1 + 0.2 \cdot f(0, 0.1) = 0.1 + 0.2(-0^2 + 0.1^2) = 0.102,$$

kde y_1 je hodnota aproximace řešení $y(0, 1)$ získaná Eulerovou metodou.

Ukažme si dále výpočet použitím Collatzovy metody, tj. $x_0 = 0$, $y_0 = 0.1$. Nejprve spočtěme

$$k_1 = f(0, 0.1) = 0.01,$$

pak

$$y_{pom} = y_0 + \frac{h}{2} k_1 = 0.1 + 0.05(0.01) = 0.1005,$$

a dále

$$k_2 = f(0.05, 0.1005) = 0.00760025.$$

Spočtěme hodnotu y_1 nyní Collatzovou metodou, tedy

$$y_1 = y_0 + h k_2 = 0.1 + 0.2 \cdot 0.00760025 = 0.10152005,$$

kde y_1 je hodnota aproximace řešení $y(0, 1)$ získaná Collatzovou metodou.

V případě použití pro soustavu rovnic je potřeba si uvědomit, že funkce \mathbf{f} je vektorová funkce stejně jako hodnoty aproximací \mathbf{y}_n , \mathbf{y}_{n+1} . Tedy i hodnoty \mathbf{k}_1 , \mathbf{k}_2 a \mathbf{y}_{pom} jsou vektory. Někdy se tedy používá značení \mathbf{Y}_n , \mathbf{Y}_{n+1} případně s horním indexem \mathbf{Y}^n , \mathbf{Y}^{n+1} . Viz též obdobná poznámka u Eulerovy explicitní metody.

6 Numerické řešení Cauchyovy úlohy, obecná jednokroková metoda.

(Explicitní) Eulerova i Collatzova metoda jsou jednokrokové metody, tj. metody které pro výpočet hodnoty aproximace $y_{n+1} \approx y(x_{n+1})$ užívají pouze hodnotu y_n . Pro vyšetření jejich vlastností, rychlosti konvergence atp. se budeme zabývat obecnými jednokrokovými metodami.

Definice 6.1 (Obecná jednokroková metoda pro řešení Cauchyovy úlohy) *Je dána Cauchyova úloha (31). Volíme $y_0 = \eta$, krok $h > 0$. Říkáme, že hodnoty aproximací y_n jsou dány obecnou jednokrokovou metodou, pokud platí*

$$y_{n+1} = y_n + h\Phi(x_n, y_n, h). \quad (36)$$

Funkci Φ nazýváme přírůstkovou funkcí $\Phi = \Phi(x, y, h)$.

Je zřejmé, že aby numerická metoda vůbec mohla být úspěšná, musí existovat jednoznačné řešení CÚ. Víme, že za určitých (jakých?) předpokladů na funkci f existuje právě jedno řešení Cauchyovy úlohy. Ukážeme, že tyto předpoklady nám budou stačit i ke konvergenci jednokrokové metody.

Definice 6.2 (Konsistentní metoda) *Říkáme, že jednokroková metoda je konsistentní s úlohou (31), pokud*

$$\Phi(x, y, 0) = f(x, y) \quad \text{pro všechna } x \in \langle a, b \rangle, y \in \mathbb{R},$$

a funkce $\Phi = \Phi(x, y, h)$ je spojitá.

Tato vlastnost sama o sobě nemusí být pro konvergenci dostačující (např. pro Eulerovu metodu je $\Phi(x, y, h) = f(x, y)$, ale pouhá spojitost funkce f nezaručuje jednoznačnost řešení. Konvergenci k nejednoznačnému řešení nelze splnit). Je tedy nutné přidat i další předpoklad na funkci Φ .

Poznámka 6.1 (Předpoklad na přírůstkovou funkci (P_Φ))

O přírůstkové funkci Φ budeme navíc požadovat, aby byla tzv. L -lipschitzovská v argumentu y , tedy existuje $C_L > 0$ takové, že pro všechna x, y, z, h platí

$$|\Phi(x, y, h) - \Phi(x, z, h)| \leq C_L |y - z|. \quad (37)$$

Tato podmínka je zaručena např. za předpokladu spojitosti parciální derivace $\frac{\partial \Phi}{\partial y}$. Vidíme, že např. v případě Eulerovy metody tento předpoklad většinou užíváme již pro ověření existence a jednoznačnosti řešení Cauchyovy úlohy. Rozmyslete si, jak je to v případě Collatzovy metody.

Pro vyšetření konvergence jednokrokové metody si nejprve definujeme, co pojem konvergence metody znamená.

Definice 6.3 (Konvergence jednokrokové metody) Označme pro dané $h > 0$ chybu $e(h)$ maximální chybu metody v intervalu $\langle a, b \rangle$, tj.

$$e(h) = \max_{x_n = a + nh \in \langle a, b \rangle} |y_n - y(x_n)|,$$

kde $y(x)$ je přesné řešení úlohy.

Řekneme, že obecná jednokroková metoda daná předpisem (36) je konvergentní, pokud platí

$$\lim_{h \rightarrow 0_+} e(h) = 0.$$

Poznámka 6.2 (Alternativní definice konvergence) Řekneme, že obecná jednokroková metoda daná předpisem (36) je konvergentní, pokud pro libovolné $x \in \langle a, b \rangle$

$$\lim_{h \rightarrow 0_+, x = x_n} y_n = y(x),$$

kde y_n je přibližné řešení v uzlu $x_n = x$ a $y(x)$ je přesné řešení.

Lze ukázat, že konsistentní jednokroková metoda je konvergentní (za předpokladu P_Φ). Budeme nyní odhadovat rychlost konvergence, tedy chybu jednokrokové metody. Dosadíme přesné řešení do (36) a získáme chybu v uzlu x_n , tj.

$$y(x_n + h) - y(x_n) - h\Phi(x_n, y(x_n), h) = h\delta_n,$$

tj. lokální relativní diskretizační chybu (krátce: lokální chybu). Nás ale zajímá globální chyba (také akumulovaná diskretizační chyba), tedy rozdíl $y_n - y(x_n) =: e_n$, kde y_n splňují vztah (36). Odečtením získáme

$$e_{n+1} = e_n + h(\Phi(x_n, y_n, h) - \Phi(x_n, y(x_n), h)) - h\delta_n,$$

kde pro třetí člen na pravé straně (lokální relativní aproximační chybu δ_n) máme odhad dle užití metody ale pro druhý člen musíme použít odhad z předpokladu P_Φ .

Definice 6.4 (Diskretizační chyba a chyba metody)

Přesný relativní přírůstek budeme označovat jako

$$\Delta(x, \tilde{y}(x), h) = \frac{\tilde{y}(x+h) - \tilde{y}(x)}{h}, \quad \text{kde } \tilde{y} \text{ je řešení ODR, které prochází bodem } [x, \tilde{y}(x)].$$

Lokální relativní diskretizační chybou (krátce diskretizační chybou) rozumíme rozdíl přírůstkové funkce a přesného relativního přírůstku

$$\delta(x, y, h) = \Phi(x, y, h) - \Delta(x, y, h)$$

Chybou metody pak rozumíme rozdíl přibližného a přesného řešení $y(x)$

$$e_n = e(x_n) = y_n - y(x_n)$$

Říkáme, že jednokroková metoda je řádu p , pokud existuje konstanta $C > 0$ tak, že pro libovolné $x \in \langle a, b \rangle$, $y \in \mathbb{R}$ a $h \in (0, h_0)$

$$|\delta(x, y, h)| = |\Phi(x, y, h) - \Delta(x, y, h)| \leq Ch^p.$$

Odvození 1 (Odhad pro chybu metody/akumulovanou diskretizační chyby)

Víme ze zadané počáteční podmínky, že platí

$$e_0 = y_0 - y(x_0) = 0.$$

Dle definice přesného relativního přírůstku a přírůstkové funkce víme

$$y_{n+1} = y_n + h\Phi(x_n, y_n, h), \quad y(x_{n+1}) = y(x_n) + h\Delta(x_n, y(x_n), h),$$

tedy pro chybu metody $e_{n+1} = y_{n+1} - y(x_{n+1})$ platí

$$\begin{aligned} e_{n+1} &= e_n + h(\Phi(x_n, y_n, h) - \Delta(x_n, y(x_n), h)) = \\ &= e_n + h[\Phi(x_n, y_n, h) - \Phi(x_n, y(x_n), h) + \Phi(x_n, y(x_n), h) - \Delta(x_n, y(x_n), h)]. \end{aligned}$$

Užitím předpokladu a řádu aproximace dostáváme

$$|e_{n+1}| \leq |e_n|(1 + C_L h) + Ch^{p+1},$$

tedy pro $A = 1 + C_L h$ máme

$$\begin{aligned} |e_1| &\leq Ch^{p+1}, \\ |e_2| &\leq A|e_1| + Ch^{p+1} \leq Ch^{p+1}(1 + A) \\ |e_3| &\leq A|e_2| + Ch^{p+1} \leq Ch^{p+1}(1 + A + A^2) \\ &\vdots \\ |e_n| &\leq A|e_{n-1}| + Ch^{p+1} \leq Ch^{p+1}(1 + A + \dots + A^{n-1}) = Ch^{p+1} \frac{A^n - 1}{A - 1} \end{aligned}$$

kde dosadíme-li za A do výrazu na pravé straně dostaneme

$$h \frac{A^n - 1}{A - 1} = \frac{(1 + C_L h)^n - 1}{C_L} \leq \frac{e^{C_L h n} - 1}{C_L} = \frac{e^{C_L(x_n - a)} - 1}{C_L}$$

a tedy máme odhad

$$|e_n| \leq \hat{C} h^p$$

Předchozí odvození si shrňme jako tvrzení.

Tvrzení 6.1 (Odhad akumulované diskretizační chyby) *Pokud je jednokroková metoda konsistentní, řádu p a přírůstková funkce Φ je lipschitzovsky spojitá v proměnné y (viz nerovnost (37)), pak akumulovaná diskretizační chyba je také řádu p , tedy*

$$|e_n| = |y(x_n) - y_n| \leq \hat{C}h^p.$$

Zde je třeba upozornit, že tento odhad může být poměrně nepřesný a obsahuje (neznámou) konstantu C . O této konstantě víme pouze, že existuje, ale neznáme její hodnotu. Z praktického hlediska je tedy tato rovnice jako odhad nepoužitelná, dává pouze představu o (limitním) chování pro h blížící se k nule.

7 Numerické řešení Cauchyovy úlohy

7.1 Příklady jednokrokových metod

Předchozí tvrzení dává návod, jakým způsobem postupovat při odvození jednokrokové metody, která je vyššího než prvního řádu. Pokud jsou splněny ostatní předpoklady tvrzení 6.1, stačí ukázat, že lokální chyba je řádu p . Vidíme, že získáváme odhad pro globální chybu Eulerovy metody (lokální chyba je totiž prvního řádu).

Příklad 7.1 *Rozhodněte, zda explicitní Eulerova metoda splňuje předpoklady z Pozn. 6.1 a zda je konsistentní. Zdůvodněte proč konverguje a jakého řádu je globální chyba.*

Příklad 7.2 *Metoda založená na přímém užití Taylorova polynomu, řádu 2.*

$$y(x+h) = y(x) + y'(x)h + \frac{h^2}{2}y''(x) + \mathcal{O}(h^3),$$

kde pro řešení $y(x)$ úlohy (31) platí

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

a tedy

$$y''(x) = \frac{d}{dx}(f(x, y(x))).$$

Definice 7.1 (Metody typu Runge-Kutta) *Metody Runge-Kutta jsou jednokrokové metody, kde přírůstkovou funkci hledáme ve tvaru*

$$\Phi(x, y, h) = \sum_{i=1}^M \omega_i \mathbf{k}_i, \quad \text{kde } \mathbf{k}_i = f\left(x + \alpha_i h, y + h \sum_{j=1}^{i-1} \beta_{ij} \mathbf{k}_j\right).$$

Koeficienty ω_i , α_i , β_{ij} hledáme tak, aby metoda byla co nejpřesnější.

Pro $M = 2$ vidíme, že

$$\Phi(x, y, h) = \omega_1 \mathbf{k}_1 + \omega_2 \mathbf{k}_2$$

kde

$$\mathbf{k}_1 = f(x, y), \quad \mathbf{k}_2 = f(x + \alpha h, y + h\beta \mathbf{k}_1).$$

Užijeme Taylorova rozvoje funkce $\Phi(x, y, h)$ jako funkce jedné proměnné h a dostaneme

$$\Phi(x, y, h) = (\omega_1 + \omega_2)f(x, y) + \omega_2 \alpha h \frac{\partial f}{\partial x}(x, y) + \omega_2 \beta h \frac{\partial f}{\partial y}(x, y) f(x, y) + \mathcal{O}(h^2).$$

Srovnáním s přesným relativním přírůstkem v bodě $y = y(x)$

$$\Delta(x, y(x), h) = \frac{y(x+h) - y(x)}{h} = y'(x) + \frac{1}{2}hy''(x) + \mathcal{O}(h^2),$$

kde y'' vyjádříme derivováním (derivace složené funkce více proměnných) rovnice

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

tj.

$$y''(x) = \frac{\partial f}{\partial x}(x, y(x)) + \frac{\partial f}{\partial y}(x, y(x))y'(x).$$

Dosazením do vyjádření pro přesný relativní přírůstek Δ dostáváme

$$\Delta(x, y(x), h) = f(x, y(x)) + \frac{1}{2}h \frac{\partial f}{\partial x}(x, y(x)) + \frac{1}{2}h \frac{\partial f}{\partial y}(x, y(x))f(x, y(x)) + \mathcal{O}(h^2).$$

Srovnáním přesného relativního přírůstku Δ a přírůstkové funkce Φ vidíme, že pro volbu $\omega_1 + \omega_2 = 1$, $\omega_2\alpha = \frac{1}{2}$ a $\omega_2\beta = \frac{1}{2}$ platí

$$\Delta(x, y(x), h) - \Phi(x, y(x), h) = \mathcal{O}(h^2),$$

tedy, že daná jednokroková metoda je druhého řádu. Tyto rovnosti jsou speciálně splněny pro $\omega_1 = 0$, $\omega_2 = 1$, $\alpha = \beta = \frac{1}{2}$. Tato volba dává vzorec pro Collatzovu metodu, Collatzova metoda je tedy druhého řádu.

Poznámka 7.1 (Collatzova metoda a Eulerova metoda) *Eulerova a Collatzova metoda jsou metody typu Runge-Kutta.*

Příklad 7.3 *Rozhodněte, zda Collatzova metoda splňuje předpoklady z Pozn. 6.1 a zda je konsistentní. Zdůvodněte proč konverguje a jakého řádu je globální chyba.*

Definice 7.2 (RK3) *Je dána Cauchyova úloha (31). Volíme $y_0 = \eta$, krok $h > 0$ a vypočteme hodnoty aproximací y_n předpisem*

$$\begin{aligned} y_{n+1} &= y_n + \frac{h}{6}(k_1 + 4k_2 + k_3), \text{ kde} \\ k_1 &= f(x_n, y_n), \\ k_2 &= f(x_n + h/2, y_n + k_1 h/2), \\ k_3 &= f(x_n + h, y_n + 2h k_2 - h k_1), \end{aligned}$$

Pak říkáme, že hodnoty aproximace y_n jsou vypočteny pomocí Rungeovy-Kuttovy metody 3. řádu. Lokální relativní aproximační chyba této metody 3. řádu, tj. $\mathcal{O}(h^3)$.

Definice 7.3 (RK4) Je dána Cauchyova úloha (31). Volíme $y_0 = \eta$, krok $h > 0$ a vypočteme posloupnost aproximací $\{y_n\}$ předpisem

$$\begin{aligned} y_{n+1} &= y_n + \frac{h}{6} (k_1 + 2k_2 + 2k_3 + k_4), \text{ kde} \\ k_1 &= f(x_n, y_n), \\ k_2 &= f(x_n + h/2, y_n + k_1 h/2), \\ k_3 &= f(x_n + h/2, y_n + k_2 h/2), \\ k_4 &= f(x_n + h, y_n + k_3 h). \end{aligned}$$

Pak říkáme, že hodnoty aproximací y_n je vypočtena pomocí Rungeovy-Kuttovy metody 4. řádu. Lokální relativní aproximační chyba je 4. řádu, tj. $O(h^4)$.

Pro RK3 vidíme, že podle Tvzení 6.1 je také globální chyba řádu 3 za předpokladu Lipschitzovské spojitosti přírůstkové funkce Φ . Tento předpoklad ale bude splněn, pokud funkce f bude mít např. spojitou parciální derivaci vzhledem k proměnné y . Metoda je tedy třetího řádu přesnosti. Obdobně pro RK4.

7.2 Metoda polovičního kroku

Vezměme nyní přesné řešení $y(x)$ Cauchyovy úlohy a hodnoty přibližného řešení y_n získané jednokrokovou metodou řádu p s krokem h . Předpokládejme dále, že akumulovaná diskretizační chyba v bodě $x \in \langle a, b \rangle$ lze zapsat ve tvaru

$$e(x; h) = e(x)h^p + \mathcal{O}(h^{p+1}),$$

tedy, že pro aproximace hodnot řešení v bodě $x = x_n = a + nh$ platí

$$y_n - y(x_n) = e(x)h^p + \mathcal{O}(h^{p+1}).$$

Spočteme nyní řešení s polovičním krokem $\tilde{h} = h/2$, označíme ho jako \tilde{y}_n . V bodě $x = \tilde{x}_{2n} = a + 2n\tilde{h} = x_n$ tedy platí

$$\tilde{y}_{2n} - y(x_n) = e(x) \frac{h^p}{2^p} + \mathcal{O}(h^{p+1}).$$

Odečtením dostaneme

$$y_n - \tilde{y}_{2n} = e(x)h^p \frac{2^p - 1}{2^p} + \mathcal{O}(h^{p+1}). \quad (38)$$

Vidíme, že rozdíl hodnot aproximací spočtených s krokem h a s krokem $h/2$ lze využít pro přibližné určení chyby.

7.3 Náhrady 1. a 2. derivace pomocí diferencí

Připomeňme nyní, že jsme v předchozích částech ukázali, že pro dostatečně hladkou funkci $f \in C^2(I)$, je možné aproximovat 1. derivaci funkce f podle Věty 5.1 např.

$$f'(x) = \frac{f(x+h) - f(x)}{h} + \eta_1(h),$$

nebo

$$f'(x) = \frac{f(x) - f(x-h)}{2h} + \eta_2(h),$$

kde $\eta_i(h) = \mathcal{O}(h)$.

První derivace funkce lze ale nahradit také dle Věty 5.2, tedy pro funkci $f \in C^3(I)$ platí

$$f'(x) = \frac{f(x+h) - f(x-h)}{h} + \eta(h), \quad (39)$$

kde $\eta(h) = \mathcal{O}(h^2)$.

Obdobně můžeme odvodit náhradu 2. derivace pomocí vztahu platného pro funkci $f \in C^4(I)$ zapsaného

$$f''(x) = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} + \eta(h), \quad (40)$$

kde $\eta(h) = \mathcal{O}(h^2)$.

8 Okrajová úloha pro ODR.

V technických problémech se často řeší namísto počátečních úloh úlohy okrajové. Víme, že pro rovnici 2. řádu se předepisují dvě počáteční podmínky. V okrajové úloze pro tutéž rovnici jednu z těchto podmínek “přesuneme” z bodu a do bodu b . Řešení hledáme na intervalu $I = (a, b)$, v krajních bodech máme předepsány tzv. okrajové podmínky. Zatímco pro počáteční úlohu byla existence jednoznačnosti zaručena spojitostí pravé strany $f(x, y)$ a její parciální derivace podle y , pro úlohy okrajové žádná takto jednoduchá podmínka neexistuje. A to dokonce ani pro lineární rovnici, rozmyslete se jak je to s řešitelností rovnice $y'' + y = 0$ s okrajovými podmínkami

$$y(0) = 0, y(\pi/2) = A$$

případně

$$y(0) = 0, y(\pi) = A.$$

Všimněte si, že daná úloha obsahuje pouze nekonečně hladké funkce (konstanty) jako svoje parametry.

Existenci a jednoznačnost úloh, které jsou zapsány ve speciálním tvaru - tzv. samo-adjungovaném tvaru - je ale možné dokázat. Tento tvar je tvar, v kterém jsou fyzikální úlohy často zapsány.

Problém 8.1 (Okrajová úloha pro lineární ODR 2. řádu) Hledáme funkci $y = y(x)$ takovou, že platí

$$\begin{aligned} a) \quad & -(p(x)y')' + q(x)y = f(x), \quad \text{pro } x \in (a, b), \\ b) \quad & \alpha_1 y'(a) + \alpha_2 y(a) = \alpha_3, \quad \beta_1 y'(b) + \beta_2 y(b) = \beta_3, \end{aligned} \tag{41}$$

kde $|\alpha_1| + |\alpha_2| \neq 0$, $|\beta_1| + |\beta_2| \neq 0$.

Definice 8.1 (Řešení úlohy) (Klasickým) řešením úlohy (41) rozumíme takovou funkci $y \in C^2(\langle a, b \rangle)$ která splňuje rovnici (41a) a okrajové podmínky (41b).

Věta 8.1 (Existence a jednoznačnost řešení) Nechť dále $p(x)$, $p'(x)$, $q(x)$, $f(x)$ jsou spojité funkce na $\langle a, b \rangle$ a navíc pro všechna $x \in \langle a, b \rangle$ platí $p(x) > 0$, $q(x) \geq 0$.

Pak existuje právě jedno řešení $y(x)$ dané okrajové úlohy (s výjimkou případu $\alpha_2 = \beta_2 = 0, q(x) \equiv 0$).

Tuto větu uvádíme bez důkazu, viz také skripta Numerická matematika, doplňkové skriptum.

Poznámka 8.1 (Převod na samoadjungovaný tvar) Pokud řešíme úlohu, která není zadaná v samoadjungovaném tvaru, můžeme ji vždy do tohoto tvaru převést a užít předchozí věty pro zdůvodnění existence a jednoznačnosti řešení (samozřejmě ale získaný tvar nemusí vždy splňovat postačující předpoklady pro existenci a jednoznačnost řešení). Ukažme si nyní postup pro převod: Vezmeme lineární ODR 2.řádu ve tvaru

$$y'' + f_1(x)y' + f_2(x)y = g(x),$$

který přenásobíme (zatím neznámou) funkcí $-p(x)$.

$$-p(x)y'' - p(x)f_1(x)y' - p(x)f_2(x)y = -p(x)g(x).$$

Srovnáme s rovnicí (41a), ve které zderivujeme člen v závorce podle pravidla pro derivaci součinu, tedy dostaneme

$$-p(x)y'' - p'(x)y' + q(x)y = f(x).$$

Srovnáním koeficientů u y a na pravé straně dostaneme vzorec pro $q(x)$ a $f(x)$. Srovnáním koeficientu u y' pak dostaneme obyčejnou diferenciální rovnici se separovanými koeficienty, kterou umíme snadno vyřešit

$$p' = pf_1(x).$$

Je zřejmé, že naopak i rovnice v samoadjungovaném tvaru lze zapsat ve tvaru

$$y'' + f_1(x)y' + f_2(x)y = g(x),$$

kde první a druhou derivaci můžeme nahradit dle (39) a (40). Tato diskretizace se někdy také používá, vede ale na soustavu rovnic s maticí, u které není možné zaručit ani diagonální dominanci ani pozitivní definitnost. Získaná matice navíc je vždy nesymetrická. Budeme používat postup, který vede k soustavě rovnic, která je symetrická, pozitivně definitní, diagonálně dominantní (někdy také ostře diagonálně dominantní).

8.1 Okrajová úloha pro lineární ODR - princip metody sítí

Řešení hledáme na $I = \langle a, b \rangle$. Volíme krok $h = (b - a)/n$ a síť $x_i = a + ih$. V bodech x_i budeme hledat hodnoty aproximace řešení $y_i \approx y(x)$. Tyto aproximace získáme tak, že rovnici (41a) nahradíme v síťových bodech x_i pomocí metody sítí. Tedy postupně aproximujeme v bodě $x = x_i$

$$qy|_{x=x_i} \approx q_i y_i, \quad f|_{x=x_i} = f_i,$$

kde $q_i = q(x_i)$, $f_i = f(x_i)$. Derivace v rovnici (41) pak nahradíme postupně pomocí vztahu (39) s krokem $h/2$, tj. v bodě $x = x_i$ dostaneme

$$-(py')'|_{x_i} \approx \frac{p_{i+1/2}y'(x_i + h/2) - p_{i-1/2}y'(x_i - h/2)}{h}.$$

Dalším užitím vztahu (39) s krokem $h/2$ dává aproximaci

$$y'(x_i - h/2) \approx \frac{y_i - y_{i-1}}{h}, \quad y'(x_i + h/2) \approx \frac{y_{i+1} - y_i}{h}$$

Dostaneme soustavu lineárních rovnic $A_h \mathbf{Y} = F_h$, kde i -tá rovnice je dána jako

$$-p_{i-1/2}y_{i-1} + (p_{i+1/2} + p_{i-1/2} + h^2q_i)y_i - p_{i+1/2}y_{i+1} = h^2f_i \quad (42)$$

V obou náhradách jsme se sice formálně dopustili chyby $O(h^2)$, nicméně v prvním případě jsme ještě h^2 dělili. Dá se ale ukázat, že náhradou (42) se dopouštíme chyby $O(h^2)$.

Lemma 8.1

Nechť funkce $p(x)$ a $q(x)$ jsou spojité a splňují $p(x) > 0$, $q(x) \geq 0$ pro libovolné $x \in \langle a, b \rangle$. Pak matice soustavy získaná aproximací okrajové úlohy pro ODR v samoadjungovaném tvaru je diagonálně dominantní, symetrická, třídiagonální a pozitivně definitní. Je-li navíc $q(x) > 0$ pak matice soustavy je ostře diagonálně dominantní.

Příklad 8.1 Je dána Dirichletova okrajová úloha pro ODR v samoadjungovaném tvaru

$$-(xy')' + x^2y = 2 - x, y(1) = 0, y(3) = 0.$$

Zapište síťové rovnice pro danou úlohu a krok $h = 0.5$.

Řešení: Vidíme, že $x_0 = 1$, $x_1 = 1.5$, $x_2 = 2$, $x_3 = 2.5$, $x_4 = 3$. Hodnoty $y_0 = y(1) = 0$ a $y_4 = y(3) = 0$ jsou dány okrajovou podmínkou, pro hodnoty y_1, y_2, y_3 sestavíme soustavu lineárních rovnic. Vyčíslíme nejprve hodnoty funkce $f(x) = 2 - x$ a funkce $q(x) = x^2$ v uzlech x_i , tedy

x	1.5	2	2.5
$f(x)$	0.5	0	-0.5
$q(x)$	2.25	4	6.25

a dále hodnoty funkce $p(x) = x$ v uzlech $x_{i\pm 1/2}$, tedy

x	1.25	1.75	2.25	2.75
$p(x)$	1.25	1.75	2.25	2.75

Sestavíme síťové rovnice dle (42) postupně pro $i = 1, 2, 3$, tedy

$$\begin{aligned} -1.25 \cdot \underbrace{0}_{y_0} + (1.25 + 1.75 + 0.5^2 \cdot 2.25)y_1 - 1.75y_2 &= 0.5^2 \cdot 0.5, \\ -1.75 \cdot y_1 + (1.75 + 2.25 + 0.5^2 \cdot 4)y_2 - 2.25y_3 &= 0.5^2 \cdot 0, \\ -2.25 \cdot y_2 + (2.25 + 2.75 + 0.5^2 \cdot 6.25)y_3 - 2.75 \cdot \underbrace{0}_{y_4} &= -0.5^2 \cdot 0.5. \end{aligned}$$

a tedy

$$\begin{aligned} 3.5625 y_1 - 1.75 y_2 &= 0.125, \\ -1.75 y_1 + 5 y_2 - 2.25 y_3 &= 0, \\ -2.25 y_2 + 6.5625 y_3 &= -0.125. \end{aligned}$$

Vidíme, že tato soustava je ostře diagonálně dominantní, symetrická. Lze snadno ověřit, že je pozitivně definitní (např. Sylvestrovým kritériem). Poznamenejme, že stejné vlastnosti by splňovala i (větší) soustava rovnic získaná pro menší krok h .

Odhad chyby a konvergence. Při odvození náhrady síťové rovnice (42) jsme zanedbali členy o velikosti $\mathcal{O}(h^2)$. Uvažujme speciálně úlohu

$$-y'' = f(x), \quad y(0) = 0, y(1) = 0$$

Diskrétní problém pak je zapsán ve tvaru ($h = 1/(n+1)$, $\mathbf{Y} = (y_1, \dots, y_n)$)

$$\mathbf{AY} = \mathbf{b},$$

kde

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & 0 & \dots & 0 \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ 0 & \dots & 0 & -1 & 2 & -1 \\ 0 & \dots & \dots & 0 & -1 & 2 \end{pmatrix}, \quad \mathbf{b} = h^2(f(x_1), \dots, f(x_n))^T.$$

Pro hodnoty přesného řešení $\mathbf{y} = (y(x_1), \dots, y(x_{n-1}))$, platí

$$\mathbf{Ay} = \mathbf{b} + \boldsymbol{\eta}_h,$$

kde $\boldsymbol{\eta}_h = \mathcal{O}(h^4)$.

Pro chybu $\mathbf{e} = \mathbf{Y} - \mathbf{y}$ tedy platí

$$\mathbf{Ae} = \boldsymbol{\eta}_h, \quad \Rightarrow \mathbf{e} = \mathbf{A}^{-1}\boldsymbol{\eta}_h.$$

Užijeme odhad v Euklidovské normě

$$\|\mathbf{e}\|_2 \leq \|\mathbf{A}^{-1}\|_2 \|\boldsymbol{\eta}_h\|_2, \quad (43)$$

kde spektrální norma matice $\|\mathbf{A}^{-1}\|_2$ je v tomto případě rovna spektrálnímu poloměru (\mathbf{A} a tedy i \mathbf{A}^{-1} jsou symetrická matice). Pro spektrální poloměr matice \mathbf{A} platí

$$\|\mathbf{A}^{-1}\|_2 = \rho(\mathbf{A}^{-1}) = 1/\lambda_{\min},$$

kde $\lambda_{\min} > 0$ je nejmenší vlastní číslo matice \mathbf{A} (SPD). Vlastní vektor a vlastní číslo je možné určit např. v MATLABu. Ukážeme, že $\lambda_{\min} = (2 - 2 \cos(\pi/(n+1)))$ a tedy $\lambda_{\min} \approx h^2$ dle Taylorova rozvoje funkce \cos . Nakonec tedy dostáváme z nerovnice (43) odhad pro chybu

$$\|\mathbf{e}\|_2 \leq O(h^2).$$

Vlastní vektory a vlastní čísla matice \mathbf{A} . Složky vlastního vektoru \mathbf{u} příslušné nejmenšímu vlastnímu číslu dle 1. cvičení jsou dány jako

$$u_j = \sin(\pi j/(n+1)) = \text{Im} e^{i\pi j/(n+1)}.$$

Ověřte užitím MATLABu. Vlastní číslo určíme ze vztahu

$$\begin{aligned} -u_{j-1} + 2u_j - u_{j+1} &= \text{Im}((2 - e^{-i\pi/(n+1)} - e^{+i\pi/(n+1)}) e^{i\pi j/(n+1)}) \\ &= \text{Im}((2 - 2 \cos(\pi/(n+1))) e^{i\pi j/(n+1)}) \\ &= (2 - 2 \cos(\pi/(n+1))) u_j, \end{aligned}$$

nebo opět MATLABem $\lambda_{\min} = (2 - 2 \cos(\pi/(n+1)))$.

Tím získáváme odhad pro normu matice

$$\|\mathbf{A}^{-1}\|_2 = 1/\lambda_{\min} = \frac{1}{2} \frac{1 + \cos(\pi/(n+1))}{\sin^2(\pi/(n+1))} \approx \frac{1}{\pi^2 h^2}.$$

9 Parciální diferenciální rovnice, okrajová úloha pro Poissonovu rovnici

V další části se budeme zabývat parciálními diferenciálními rovnicemi, speciálně se budeme zabývat lineárními rovnicemi 2. řádu (tj. obsahující parciální derivace 2. a nižšího řádu) pro funkci dvou nezávislých proměnných. Takovéto rovnice se dělí na tři základní typy, eliptické, parabolické a hyperbolické. Příkladem eliptické rovnice je tzv. Poissonovu rovnice

$$-\Delta u = f$$

v Ω , kde $\Omega \subset \mathbb{R}^2$ je omezená oblast, $f = f(x, y)$ je funkce dvou proměnných a Δ je tzv. Laplaceův operátor. Laplaceův operátor je diferenciální operátor definovaný ve dvourozměrném prostoru

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}, \quad \text{nebo také} \quad \Delta = \nabla \cdot \nabla,$$

kde operátor “nabla” je definován $\nabla = (\partial_x, \partial_y)$. Aby úloha byla jednoznačně řešitelná je třeba jí doplnit okrajovou podmínkou, zabýváme se tedy okrajovou úlohou pro Poissonovu rovnici. Okrajové podmínky bývají zadány ve formulaci problému, fyzikální význam mají např. tzv. Dirichletova okrajová podmínka (předepisujeme hodnotu funkce u) a Neumannova okrajová podmínka (předepisujeme její tok přes hranici). Pro jednoduchost se v tomto textu budeme zabývat numerickým řešením problému pouze s Dirichletovou okrajovou podmínkou, tento problém budeme nazývat Dirichletovou okrajovou úlohou pro Poissonovu rovnici.

Dirichletova úloha pro Poissonovu rovnici. Poissonovu rovnici budeme řešit na omezené oblasti $\Omega \subset \mathbb{R}^2$ s hranicí $\partial\Omega$. Na hranici oblasti budeme klást ještě další požadavky, v literatuře se používá hladká ($\partial\Omega \in \mathcal{C}^2$) nebo lipschitzovsky spojitá hranice. V případě Dirichletovy úlohy pro Poissonovu rovnici pak hledáme funkci $u = u(x, y) : \overline{\Omega} \rightarrow \mathbb{R}$ takovou, že splňuje

$$-\Delta u = f \quad \text{v oblasti } \Omega, \quad (44)$$

a navíc platí

$$u = \varphi \quad \text{na hranici } \partial\Omega. \quad (45)$$

Zde f a φ jsou funkce, tj. $f = f(x, y)$ a $\varphi = \varphi(x, y)$. Řešením (nebo také klasickým řešením) Dirichletovy úlohy pro Poissonovu rovnici nazveme funkci $u \in \mathcal{C}^2(\overline{\Omega})$ takovou, že splňuje rovnici (44) a okrajovou podmínku (45).

Zatímco v případě okrajové úlohy pro ODR bylo možné ověřit existenci a jednoznačnost řešení, v případě Poissonovy rovnice je situace komplikovanější. Řešení je ovlivněno volbou oblasti Ω , zejména jeho hranicí $\partial\Omega$. Existence řešení závisí také na funkci f a předepsané okrajové podmínce. V některých jednoduchých případech je možné existenci a jednoznačnost řešení ukázat, zde uvádíme dvě věty přejaté ze skript Numerická matematika, doplňkové skriptum.

Např. je-li Ω oblast s hladkou hranicí (\mathcal{C}^2), $\varphi \in \mathcal{C}(\overline{\Omega})$, $f \in \mathcal{C}^1(\overline{\Omega})$, pak existuje právě jedno řešení $u \in \mathcal{C}^2(\overline{\Omega})$.

Nebo je-li Ω konvexní s Lipschitzovskou spojitou hranicí, $\varphi \in \mathcal{C}(\overline{\Omega})$, $f \equiv 0$, pak existuje také právě jedno řešení $u \in \mathcal{C}^2(\overline{\Omega})$.

Existence a jednoznačnost řešení V obecném případě je ale otázka existence a jednoznačnosti klasického řešení složitá. Uvažujme zde ale speciální případ, kdy oblast $\Omega = (0, 1)^2$ (čtverec o straně jedna). Úloha je doplněna nulovými Dirichletovými okrajovými podmínkami. Snadno ověříme (ověřte si!), že funkce

$$u(x, y) = \sin(k\pi x) \sin(m\pi y),$$

je řešením rovnice

$$-\Delta u = (k^2 + m^2) \sin(k\pi x) \sin(m\pi y).$$

Vidíme, že je-li pravá strana zapsaná ve tvaru

$$f(x, y) = \sum_{k,m=1}^N \alpha_{k,m} \sin(k\pi x) \sin(m\pi y),$$

jsme schopni napsat přesné řešení úlohy (rozmyslete si!) vzhledem k linearitě Poissonovy rovnice. Pokud bychom provedli limitní přechod $N \rightarrow \infty$, vidíme, že pokud se dá pravá strana f rozvinout do nekonečné Fourierovy sinové řady, můžeme také řešení zapsat ve tvaru součtu Fourierovy sinové řady. Zde je ale třeba uvážit, zda součet takovéto nekonečné řady je dostatečně hladká funkce. V numerické matematice ale budeme postupovat jinak, užijeme pro aproximaci řešení metodu sítí.

Princip metody sítí pro Dirichletovy úlohy pro Poissonovu rovnici. Metoda sítí (nebo také metoda konečných diferencí) spočívá v aproximaci rovnice na zvolené

množině bodů (tzv. síťových uzlech). Budeme tedy aproximovat řešení Dirichletovy úlohy pro Poissonovu rovnici tak, že rovnici

$$-\Delta u = f$$

nahradíme pomocí diferenčního schématu v uzlech sítě, která je daná pomocí prostoro-
vého kroku $h > 0$ (stejný v obou směrech). Síťové čáry jsou pak dány rovnicemi
 $x = x_i$ resp. $y = y_j$, kde $x_i = x_0 + ih$, $y_j = y_0 + jh$. Bod $[x_0, y_0]$ je nějaký vhodně
zvolený bod. Řešení u pak aproximujeme v síťových uzlech $P_{i,j}$ pomocí $U_{i,j}$, tj.

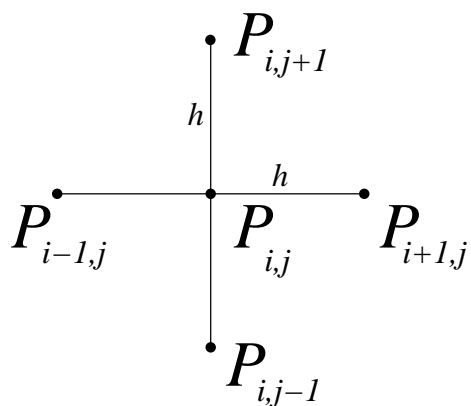
$$P_{i,j} = [x_i, y_j], \quad U_{i,j} \approx u(P_{i,j}) = u(x_i, y_j) = u_{ij}.$$

Na hranici oblasti jsou v případě Dirichletovy okrajové podmínky hodnoty dány okrajovou podmínkou. Parciální derivace v bodě $P_{i,j}$ nahradíme diferencemi. Vzhledem k tomu, že parciální derivace jsou derivace vzhledem k jedné proměnné, uijeme následující větu (viz její odvození v minulých přednáškách).

Věta 9.1 (Diferenční náhrada druhé derivace) *Nechť funkce $f \in C^4(I)$ a interval I obsahuje body $x, x \pm h$. Pak platí*

$$\frac{f(x+h) - 2f(x) + f(x-h)}{h^2} = f''(x) + \eta(h). \quad (46)$$

kde $\eta(h) = \mathcal{O}(h^2)$



Obrázek 2: Schématické znázornění regulárního uzlu

Předpokládáme, že řešení úlohy splňuje $u \in C^4(\bar{\Omega})$. Užitím náhrady (46) pro funkci $u(x, y_j)$ a $u(x_i, y)$ vidíme, že

$$\frac{\partial^2 u}{\partial x^2}(P_{i,j}) = \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{h^2} + \mathcal{O}(h^2),$$

a dále

$$\frac{\partial^2 u}{\partial y^2}(P_{ij}) = \frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{h^2} + \mathcal{O}(h^2).$$

Užitím těchto náhrad pro Δu dostáváme

$$\Delta u|_{P_{ij}} = \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{h^2} + \frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{h^2} + \mathcal{O}(h^2).$$

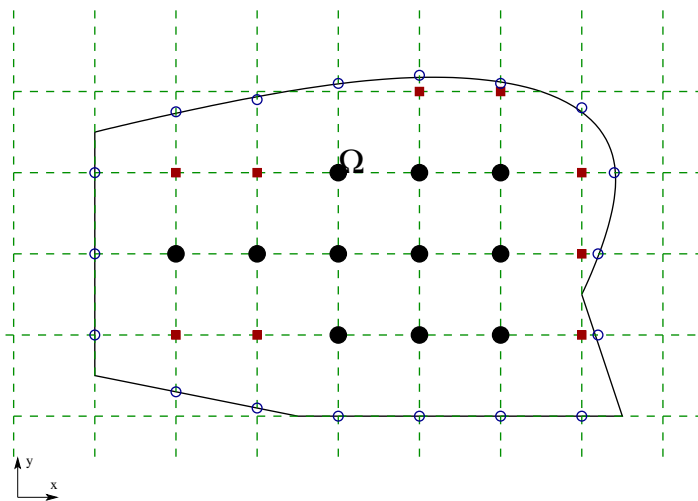
Dosadíme do rovnice (44), vynásobíme h^2 a dostáváme

$$-u_{i,j-1} - u_{i-1,j} + 4u_{i,j} - u_{i,j+1} - u_{i+1,j} = h^2 f_{i,j} + \mathcal{O}(h^4).$$

Zanedbáním členů $\mathcal{O}(h^4)$ dostáváme rovnici pro aproximaci hodnot řešení, tj.

$$-U_{i,j-1} - U_{i-1,j} + 4U_{i,j} - U_{i,j+1} - U_{i+1,j} = h^2 f_{i,j}.$$

Tuto náhradu můžeme použít jen v takovém bodě $P_{ij} \in \Omega$ (tzv. regulární uzel), jehož všichni jeho sousedi (ale i spojnice) leží v oblasti Ω nebo na její hranici, viz Obr. 2.



Obrázek 3: Příklad sítě s regulárními (černá plná kolečka) a neregulárními (červené čtverečky) uzly. Hraníční uzly jsou vyznačeny modře.

Realizace okrajové podmínky v neregulárních uzlech Hodnotu v neregulárním uzlu můžeme brát jako hodnotu okrajové podmínky z nejbližšího uzlu na hranici (tzv. *přímý přenos* okrajové podmínky) nebo užít tzv. *lineární interpolace*.

V případě *přímého přenosu okrajové podmínky* tedy pro neregulární uzel N předepíšeme okrajovou podmínku

$$U_N = u(Q),$$

kde $u(Q) = \varphi(Q)$ je hodnota řešení daná okrajovou podmínkou v nejbližším hraničním uzlu Q . V tomto případě se ale dopouštíme chyby, která je pouze prvního řádu přesnosti (tedy $\mathcal{O}(h)$).

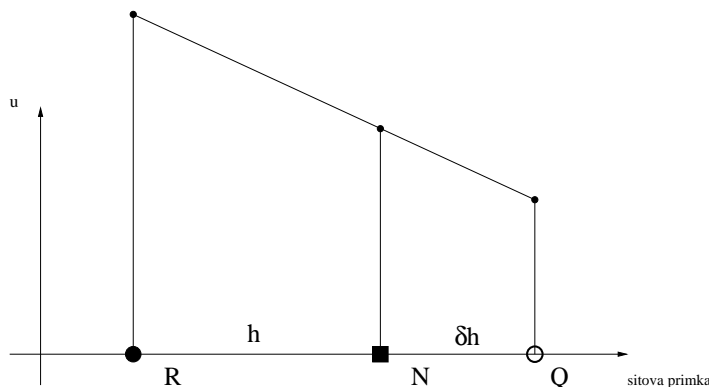
Toto lze napravit užitím **lineární interpolace**, kdy hodnotu v regulárním uzlu určíme jako lineární interpolaci z hodnoty dané na hranici a hodnoty v uzlu regulárním, viz obrázek 4. Z podobnosti trojúhelníků dostáváme

$$\frac{U_N - U_Q}{\delta \cdot h} = \frac{U_R - U_Q}{(1 + \delta) \cdot h},$$

a tedy po vynásobení δh dostáváme náhradu v neregulárním uzlu

$$(1 + \delta)U_N - \delta U_R = \varphi(Q).$$

Při užití této náhrady se dopouštíme chyby druhého řádu přesnosti, tedy stejné chyby jako při náhradě v regulárním uzlu.



Obrázek 4: Lineární interpolace v neregulárním uzlu.

Příklad 9.1 (Ukázka aproximace metodou sítí) Je dána okrajová úloha $-\Delta u = 12xy$ v oblasti tvořené čtyřúhelníkem s vrcholy $[0, 0]$, $[1.8, 0]$, $[0, 1.5]$, $[1.5, 1.5]$, na hranici je předepsána okrajová podmínka $u(x, y) = x + y$.

- Uveďte o jaký typ rovnice se jedná a pomocí parciálních derivací rozepište symbol Δu . Ověřte, zda pro funkci $u(x, y) = xy(97 - x^2 - y^2)$ platí $-\Delta u = 12xy$.
- Zapište, jak se v regulárním uzlu $P_{i,j} = [x_i, y_j]$ nahradí parciální derivace uvedené v části a). Odvoďte rovnici pro náhradu dané rovnice metodou sítí v uzlu $P_{i,j}$.
- Volte krok $h = 0.5$ a síť tak, aby obsahovala bod $[0, 0]$. Sestavte síťové rovnice v uzlech sítě ležících na přímce $y = 1$. V neregulárních uzlech užitě lineární interpolaci.

Řešení: Zabývejme se pouze řešením části c), řešení zbývajících částí čtenář snadno nalezne ve zbývajících částech textu nebo jednoduše ověří dosazením. Danou oblast si znázorníme na obrázku do kterého také zakreslíme síť, regulární a neregulární uzly, viz Obr. 5. Regulární uzly jsou uzly $A = [0.5, 0.5]$, $B = [1, 0.5]$, $D = [0.5, 1]$, $E = [1, 1]$. Neregulární uzly jsou uzly $C = [1.5, 0.5]$ a $F = [1.5, 1]$. Nejprve pro daný krok h sestavíme rovnice v regulárních uzlech

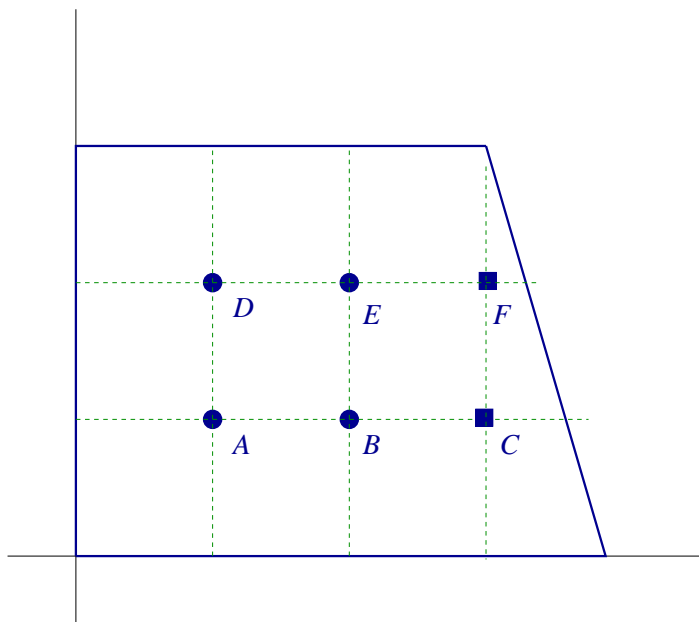
$$\begin{aligned} 4U_A - U_B - U_D - (0.5 + 0) - (0 + 0.5) &= 0.5^2 \cdot 12 \cdot 0.5 \cdot 0.5, \\ 4U_B - U_C - U_A - U_E - (1 + 0) &= 0.5^2 \cdot 12 \cdot 1 \cdot 0.5, \\ 4U_D - U_A - U_E - (0 + 1) - (0.5 + 1.5) &= 0.5^2 \cdot 12 \cdot 0.5 \cdot 1, \\ 4U_E - U_D - U_F - U_B - (1 + 1.5) &= 0.5^2 \cdot 12 \cdot 1 \cdot 1. \end{aligned} \tag{47}$$

Uzly C a F jsou neregulární. Dle obrázku je možné zjistit souřadnice hraničních uzlů $Q_C = [1.7, 0.5] = [1.5 + \frac{2}{5}h, 0.5]$ a $Q_F = [1.6, 1] = [1.5 + \frac{1}{5}h, 1]$, leží totiž na přímkách spojující body $[1.5, 1.5]$ a $[1.8, 0]$. Vidíme tedy, že náhrada v neregulárním uzlu užitím lineární interpolace dává rovnice

$$\begin{aligned} (1 + \frac{2}{5})U_C - \frac{2}{5}U_B &= 1.7 + 0.5, \\ (1 + \frac{1}{5})U_F - \frac{1}{5}U_E &= 1.6 + 1. \end{aligned} \tag{48}$$

Získali jsme soustavu 6 rovnic (47-48) o šesti neznámých ($U_A, U_B, U_C, U_D, U_E, U_F$). Matice této soustavy (jak se snadno přesvědčíte) není symetrická, ale je ostře diagonálně dominantní. Lze ukázat, že je také pozitivně definitní (jak?). Symetrickou matici je možné získat v případě, kdy se pro hodnoty v neregulárních uzlech U_C a U_F uzlech použije přímý přenos okrajové podmínky. Případně lze vyjádřit U_C a U_F z rovnic (48) a dosadit do (47), čímž získáme symetrickou a pozitivně definitní matici (eliminace neznámých v neregulárních uzlech). Rozmyslete si, že pro menší hodnoty kroku h by získaná matice již nebyla ostře diagonálně dominantní (proč?), ale výše zmíněný postup (eliminace neznámých v neregulárních uzlech) by vedl na soustavu rovnic se symetrickou pozitivně definitní maticí.

Soustavu rovnic (47-48) je tedy možné přímo řešit užitím Jacobiovy nebo Gaussovy-Seidelovy iterační metody. Všimněme si, že tuto soustavu je možné takto řešit dokonce i v případě, že celou tuto soustavu vůbec nebudeme sestavovat, ale vždy užitíme jen jednu její konkrétní rovnici. Rozmyslete si, že v důsledku Gerschgorinovy věty platí: symetrická matice diagonálně dominantní regulární matice je pozitivně definitní.



Obrázek 5: Oblast pro Příklad 9.1 a vyznačení uzlů.

Odhad chyby a konvergence. Uvažujme speciálně úlohu

$$-\Delta u = f(x, y), \quad \text{v } \Omega = (0, 1)^2, u = 0 \text{ na } \partial\Omega.$$

Diskrétní problém pak je zapsán ve tvaru ($h = 1/(n + 1)$, $\mathbf{U} = (U_{ij})$)

$$\mathbf{A}\mathbf{U} = \mathbf{b},$$

kde matice \mathbf{A} je blokově třídiagonální matice, viz také příklady ze cvičení (MATLAB).

$$\mathbf{A} = \begin{pmatrix} \mathbf{B} & -\mathbf{E} & 0 & \dots & 0 \\ -\mathbf{E} & \mathbf{B} & -\mathbf{E} & 0 & \dots & 0 \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ 0 & \dots & 0 & -\mathbf{E} & \mathbf{B} & -\mathbf{E} \\ 0 & \dots & \dots & 0 & -\mathbf{E} & \mathbf{B} \end{pmatrix}, \text{ kde } \mathbf{B} = \begin{pmatrix} 4 & -1 & 0 & \dots & 0 \\ -1 & 4 & -1 & 0 & \dots & 0 \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ 0 & \dots & 0 & -1 & 4 & -1 \\ 0 & \dots & \dots & 0 & -1 & 4 \end{pmatrix}.$$

Pro hodnoty přesného řešení $\mathbf{z} = (z_{ij})$, $z_{ij} = u(P_{ij})$ platí

$$\mathbf{A}\mathbf{z} = \mathbf{b} + \boldsymbol{\eta}_h,$$

kde $\|\boldsymbol{\eta}_h\|_2 = O(h^4)$.

Pro chybu $\mathbf{e} = \mathbf{z} - \mathbf{U}$ tedy platí

$$\mathbf{A}\mathbf{e} = \boldsymbol{\eta}_h, \quad \Rightarrow \mathbf{e} = \mathbf{A}^{-1}\boldsymbol{\eta}_h.$$

Užijeme odhad v Euklidovské normě

$$\|\mathbf{e}\|_2 \leq \|\mathbf{A}^{-1}\|_2 \|\boldsymbol{\eta}_h\|_2,$$

kde $\|\mathbf{A}^{-1}\|_2$ je v tomto případě rovna spektrálnímu poloměru (\mathbf{A} a tedy i \mathbf{A}^{-1} jsou symetrická matice).

Pro spektrální poloměr matice \mathbf{A} platí

$$\|\mathbf{A}^{-1}\|_2 = \rho(\mathbf{A}^{-1}) = 1/\lambda_{\min},$$

kde $\lambda_{\min} > 0$ je nejmenší vlastní číslo matice \mathbf{A} (matice je SPD). Lze ověřit, že nejmenší vlastní číslo matice \mathbf{A} splňuje $\lambda_{\min} \approx \frac{1}{h^2}$ (řádově). Tedy nakonec dostáváme odhad pro chybu

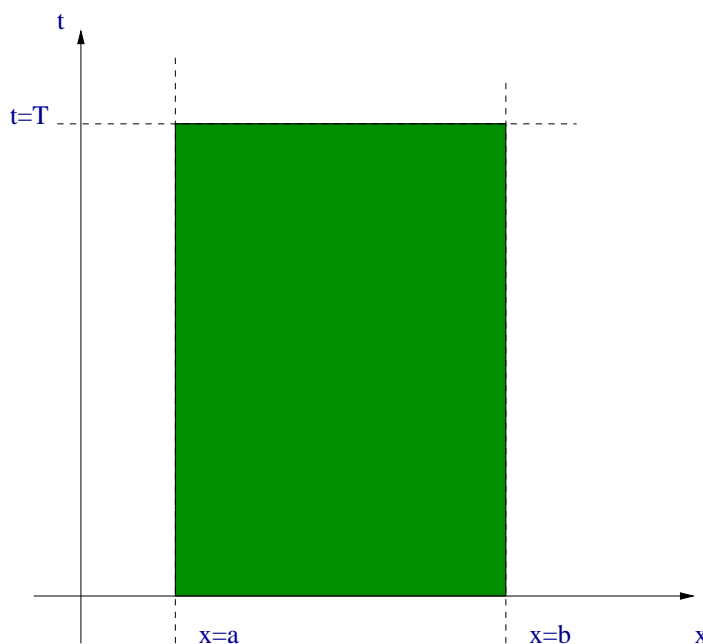
$$\|\mathbf{e}\|_2 \leq O(h^2).$$

10 Rovnice vedení tepla

Smíšená úloha pro rovnici vedení tepla Příkladem parabolické diferenciální rovnice je rovnice vedení tepla, tedy rovnice zapsaná ve tvaru

$$\frac{\partial u}{\partial t} = p \frac{\partial^2 u}{\partial x^2} + f(x, t).$$

V této rovnici má funkce $u = u(x, t)$ význam teploty, proměnná t označuje čas, p je kladná konstanta a $f(x, t)$ tepelné zdroje. Rovnici budeme řešit na oblasti $Q_T = (a, b) \times (0, T)$. Formulaci úlohy je třeba doplnit počáteční podmínkou a okrajovými podmínkami, tím dostáváme tzv. smíšenou úlohu.



Obrázek 6: Oblast řešení smíšené úlohy pro rovnici vedení tepla

V rovnici vedení tepla je dán koeficient $p > 0$ a zdrojové členy $f = f(x, t)$. Úlohu řešíme na oblasti $Q_T = (a, b) \times (0, T)$. Řešení u definované v bodech $[x, t] \in Q_T$ má v Q_T splňovat rovnici

$$\frac{\partial u}{\partial t} = p \frac{\partial^2 u}{\partial x^2} + f, \quad (49)$$

dále počáteční podmínku

$$u(x, 0) = \varphi(x) \text{ pro } x \in \langle a, b \rangle \quad (50)$$

a okrajové podmínky pro $t \geq 0$

$$u(a, t) = \alpha(t), u(b, t) = \beta(t). \quad (51)$$

Všimněme si, že v každém z bodů $[a, 0]$ a $[b, 0]$ jsou pro hodnoty u předepsány dvě podmínky, viz (50) a (51). Je tedy zřejmé, že aby bylo zadání úlohy korektní je třeba, aby obě podmínky předepisovaly stejné hodnoty u , tj. aby byly splněny tzv. **podmínky souhlasu**

$$\varphi(a) = \alpha(0), \quad \varphi(b) = \beta(0). \quad (52)$$

Definice 10.1 (Řešení smíšené úlohy pro rovnici vedení tepla) Funkci $u = u(x, t)$, $u \in \mathcal{C}^2(Q_T) \cup \mathcal{C}(\overline{Q_T})$ nazveme řešením úlohy (také klasickým řešením), pokud u splňuje rovnici (49) v oblasti Q_T , počáteční podmínku (50) a okrajové podmínky (51).

Pro tuto rovnici se nebudeme zabývat otázkou existence a jednoznačnosti řešení, ale podíváme jaké mají vlastnosti řešení ve speciálním případě. Uvažujme nyní úlohu pro případ intervalu $(a, b) = (0, \pi)$, $f \equiv 0$ a nulové okrajové podmínky. Počáteční podmínka nechť je dána jako

$$\varphi(x) = \sin(kx),$$

kde $k \in \mathbb{N}$. Řešení hledejme ve tvaru $u(x, t) = e^{-At} \sin(kx)$, kde konstantu A určíme tak, aby platila rovnice (49). Snadno ověříme (ověřte), že pak

$$u(x, t) = e^{-pk^2 t} \sin(kx) \quad (53)$$

je řešením dané úlohy.

V případě, že počáteční podmínka je dána jako lineární kombinace

$$\varphi(x) = \sum_{k=1}^n b_k \sin(kx),$$

užijeme linearitu rovnice a snadno ukážeme (rozmyslete si!), že řešení má tvar

$$u(x, t) = \sum_{k=1}^n \beta_k e^{-pk^2 t} \sin(kx).$$

Všimněte si, jakým způsobem se chovají jednotlivé členy pro různá k v čase (proměnná t) a prostoru (proměnná x). Obecně je-li φ rozvinuto do Fourierovy sinové řady (limitní případ $n \rightarrow +\infty$), pak také řešení hledáme ve tvaru součtu Fourierovy sinové řady. Opět ale upozorníme, že součet nekonečné řady hladkých funkcí obecně nemusí být hladká funkce, více viz např. Matematika II. V numerické matematice budeme ale řešení aproximovat pomocí metody sítí.

10.1 Metoda sítí pro rovnici vedení tepla.

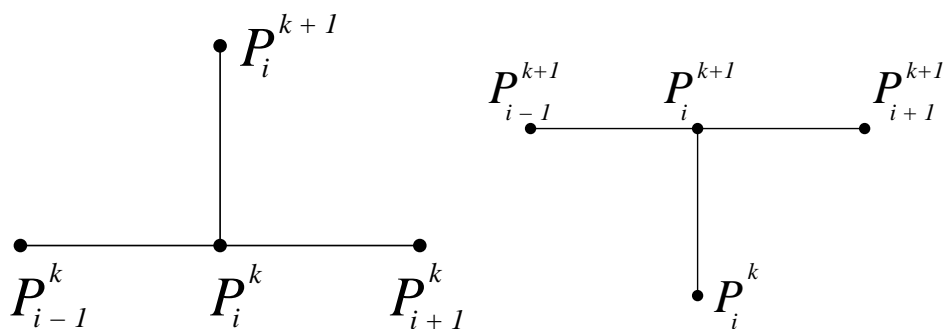
Princip metody sítí. Budeme aproximovat řešení smíšené úlohy pro rovnici

$$\frac{\partial u}{\partial t} = p \frac{\partial^2 u}{\partial x^2} + f,$$

v uzlech sítě, která je daná pomocí prostorového kroku $h = (b - a)/n$ a kroku $\tau > 0$. Síťové čáry jsou pak dány rovnicemi $x_i = a + ih$, $t_k = k\tau$ a hodnoty řešení $u(x_i, t_k)$ pak aproximujeme v síťových uzlech P_i^k jako přibližné řešení U_i^k , tj.

$$P_i^k = [x_i, t_k], \quad U_i^k \approx u_i^k \stackrel{\text{def}}{=} u(x_i, t_k),$$

kde u_i^k budeme značit hodnotu přesného řešení v bodě P_i^k . Z počáteční podmínky určíme hodnoty $\vec{U}^0 = (U_i^0)_i$ a sestavíme schéma pro výpočet hodnot $\vec{U}^{k+1} = (U_i^{k+1})_{i=1}^{n-1}$ z \vec{U}^k , které použijeme pro $k = 0, 1, 2, \dots$



Obrázek 7: Znázornění explicitního (nalevo) a implicitního (napravo) schématu pro rovnici vedení tepla

10.2 Explicitní schéma

V případě explicitního schématu nahradíme časovou derivaci diferencí hodnot v bodech P_i^k, P_i^{k+1} a prostorovou derivaci nahradíme v k -té časové vrstvě jak ukazuje následující postup. Předpokládejme, že řešení u je dostatečně hladké ($u \in C^4(\overline{Q_T})$). Pro hodnoty přesného řešení $u_i^k = u(x_i, t_k)$ v uzlech P_i^k platí

$$\frac{\partial^2 u}{\partial x^2}(P_i^k) = \frac{u_{i-1}^k - 2u_i^k + u_{i+1}^k}{h^2} + \eta_1(h),$$

kde $\eta_1(h) = \mathcal{O}(h^2)$. Dále pro derivaci vzhledem k času dostáváme

$$\frac{\partial u}{\partial t}(P_i^k) = \frac{u_i^{k+1} - u_i^k}{\tau} + \eta_2(\tau),$$

kde $\eta_2(\tau) = \mathcal{O}(\tau)$. Hodnoty přesného řešení tedy splňují rovnici

$$\frac{u_i^{k+1} - u_i^k}{\tau} = p \frac{u_{i-1}^k - 2u_i^k + u_{i+1}^k}{h^2} + f_i^k + \mathcal{O}(\tau + h^2),$$

kde $f_i^k = f(P_i^k)$. Zanedbáme tedy člen $\mathcal{O}(\tau + h^2)$ a dostaneme vztah pro přibližné řešení

$$\frac{U_i^{k+1} - U_i^k}{\tau} = p \frac{U_{i-1}^k - 2U_i^k + U_{i+1}^k}{h^2} + f_i^k \quad (54)$$

Rovnici (54) násobíme τ , označíme $\sigma = \frac{p\tau}{h^2}$ a vyjádříme

$$U_i^{k+1} = \sigma U_{i-1}^k + (1 - 2\sigma)U_i^k + \sigma U_{i+1}^k + \tau f_i^k. \quad (55)$$

Lze snadno ověřit, že výpočet explicitním schématem není možné provádět pro libovolnou volbu h a τ , a to dokonce ani v případě, že tyto hodnoty jsou hodně malé. Aby bylo možné dosáhnout toho, že řešení je “blízko” přesnému řešení, je třeba splnit ještě dodatečnou podmínku, tzv. stability schématu. V našem případě je stabilita schématu zaručena pro hodnotu $\sigma = \frac{p\tau}{h^2} \leq \frac{1}{2}$. Naopak pro velké hodnoty sigma je schéma tzv. nestabilní, hodnoty které takto vypočteme jsou nesmyslné. Všimněme si, že schéma je nestabilní např. pro případ $p = 1$, $h = 10^{-10}$ a $\tau = 10^{-10}$. Hodnoty h a τ jsou v tomto případě velmi malé, ale získaná aproximace řešení je přesto velmi nepřesná.

Podmínka stability explicitního schématu Podmínka stability explicitního schématu je dána nerovnicí

$$\sigma = \frac{p\tau}{h^2} \leq \frac{1}{2}.$$

Ukážeme si, jak lze tuto podmínku stability odvodit. Užijeme zjednodušený postup pro případ nulové pravé strany $f \equiv 0$ a nulových okrajových podmínek. Jak jsme viděli ve zjednodušeném případě tak např. amplituda řešení by měla být v čase klesající nebo alespoň nerostoucí, viz (53). V maticovém zápisu vidíme, že rovnice (55) lze zapsat

$$\vec{U}^{k+1} = \mathbf{A}\vec{U}^k,$$

kde matice \mathbf{A} je třídiagonální matice s prvky $(1 - 2\sigma)$ na hlavní diagonále a s prvky σ na diagonálách vedlejších. Opakovaným užitím dostáváme

$$\vec{U}^k = \mathbf{A}^k \vec{U}^0.$$

Tedy aby \vec{U}^k pro $k \rightarrow \infty$ bylo nerostoucí, je nutné (i postačující) aby spektrální poloměr matice \mathbf{A} byl menší nebo roven jedné. Všimněme si v první řadě, že vlastní čísla

matice \mathbf{A} jsou reálná (rozmyslete si, proč !?). Užitím Gerschgorinovy věty dostaneme podmínku stability: dle (10) vidíme, že vlastní čísla $\lambda \in \mathbb{R}$ matice \mathbf{A} splňují

$$|\lambda - a_{ii}| = |\lambda - 1 + 2\sigma| \leq \sum_{j \neq i} |a_{ij}| = 2\sigma$$

nebo-li

$$-2\sigma \leq \lambda - 1 + 2\sigma \leq 2\sigma.$$

Přičtením $1 - 2\sigma$ dostáváme odhad pro vlastní čísla

$$1 - 4\sigma \leq \lambda \leq 1.$$

Vidíme tedy, že pro $0 \leq \sigma \leq 1/2$ dostáváme $|\lambda| \leq 1$ a tedy $\rho(\mathbf{A}) \leq 1$.

Použití explicitního schématu Pro lepší pochopení ukážeme na jakým způsobem se používá explicitní schéma ukážeme jeho použití pro jednoduchý příklad a výpočet několika málo hodnot aproximací ve vybraných bodech.

Příklad 10.1 *Je dána smíšená úloha pro rovnici vedení tepla*

$$\frac{\partial u}{\partial t} = \frac{1}{5} \frac{\partial^2 u}{\partial x^2} + xt,$$

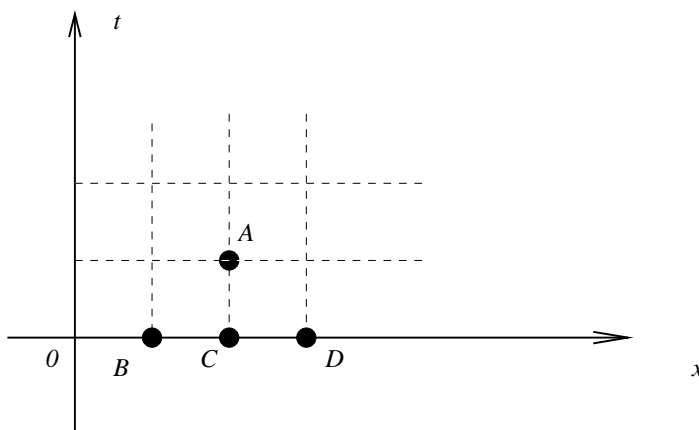
s počáteční podmínkou $u(x, 0) = x - x^2$ pro $x \in \langle 0; 1 \rangle$, a okrajovými podmínkami

$$u(0, t) = 0 \quad u(1, t) = 0 \quad \text{pro } t \geq 0$$

- Ověřte, zda jsou splněny podmínky souhlasu.
- Určete maximální krok τ tak, aby byla splněna podmínka stability pro explicitní metodu s prostorovým krokem $h = 0.25$. Odvoďte schéma pro explicitní metodu.
- Volte prostorový krok 0.25 a maximální časový krok tak aby bod $A = [0.5; 0.1]$ byl uzlem sítě. Určete pro dané h a τ přibližně hodnotu řešení v bodě A explicitním schématem.

Řešení: Podmínky souhlasu jsou pro zadanou úlohu splněny jak čtenář snadno ověří. Podmínka stability pro explicitní schéma a danou rovnici vedení tepla je $\sigma = \frac{\frac{1}{5}\tau}{h^2} \leq \frac{1}{2}$. Volba $h = 0.25$ dává omezení

$$\tau \leq \frac{5}{2} \frac{1}{4^2} = \frac{5}{32} \approx 0.156.$$



Obrázek 8: Oblast řešení, síť a uzly sítě

Maximální časový krok je tedy $\tau_{max} = \frac{5}{32}$.

Pro část d) tedy volíme krok $h = 0.25$ a $\tau = 0.1 \leq \tau_{max}$. Pro tuto volbu vyjde $\sigma = \frac{\frac{1}{2}\tau}{h^2} = 0.32$, budeme tedy počítat podle vzorce

$$U_i^{k+1} = 0.32U_{i+1}^k + 0.36U_i^k + 0.32U_{i-1}^k + 0.1f_i^k.$$

Hodnotu v bodě A určíme z hodnot v bodech v nulové časové vrstvě, které si označíme $B = [0; 0]$, $C = [0.25; 0]$ a $D = [0.5; 0]$, viz Obrázek 8. Hodnoty U_B , U_C , U_D určíme z počáteční podmínky

$$\begin{aligned} U_B &= u(0.25, 0) = 0.1875, \\ U_C &= u(0.5, 0) = 0.25, \\ U_D &= u(0.75, 0) = 0.1875. \end{aligned}$$

Užitím vzorce pro explicitní schéma dostáváme hodnotu aproximace v bodě A

$$u(0.5, 0.1) \approx U_A = 0.32U_B + 0.36U_C + 0.32U_D + 0.1 \cdot f(C) = 0.21$$

kde hodnota $f(C) = 0$ (rozmyslete si proč je funkce f vyčíslena právě v bodě C).

Výpočet dalších aproximací v ostatních bodech sítě na 1. časové vrstvě by probíhal obdobně. Stejný postup bychom použili pro výpočet hodnot v dalších časových vrstvách.

10.3 Implicitní schéma

Pro odvození implicitního schématu postupujeme obdobně, ale derivace nahradíme v bodě P_i^{k+1} . Pro hodnoty přesného řešení $u_i^{k+1} = u(x_i, t_k)$ v uzlech P_i^{k+1} platí

$$\frac{\partial^2 u}{\partial x^2}(P_i^{k+1}) = \frac{u_{i-1}^{k+1} - 2u_i^{k+1} + u_{i+1}^{k+1}}{h^2} + \eta_1(h),$$

kde $\eta_1(h) = \mathcal{O}(h^2)$. Dále pro derivaci vzhledem k času uijeme

$$\frac{\partial u}{\partial t}(P_i^{k+1}) = \frac{u_i^{k+1} - u_i^k}{\tau} + \eta_2(\tau),$$

kde $\eta_2(\tau) = \mathcal{O}(\tau)$. Hodnoty přesného řešení tedy splňují rovnici

$$\frac{u_i^{k+1} - u_i^k}{\tau} = p \frac{u_{i-1}^{k+1} - 2u_i^{k+1} + u_{i+1}^{k+1}}{h^2} + f_i^{k+1} + \mathcal{O}(\tau + h^2).$$

Zanedbáme-li člen $\mathcal{O}(\tau + h^2)$ dostaneme vztah pro hodnoty přibližného řešení

$$\frac{U_i^{k+1} - U_i^k}{\tau} = p \frac{U_{i-1}^{k+1} - 2U_i^{k+1} + U_{i+1}^{k+1}}{h^2} + f_i^{k+1} \quad (56)$$

Rovnici (56) násobíme τ , označíme $\sigma = \frac{p\tau}{h^2}$ a vyjádříme

$$-\sigma U_{i-1}^{k+1} + (1 + 2\sigma)U_i^{k+1} - \sigma U_{i+1}^{k+1} = U_i^k + \tau f_i^{k+1}. \quad (57)$$

Dostáváme soustavu rovnic, se symetrickou pozitivně definitní maticí, která je také ostře diagonálně dominantní.

Podmínka stability implicitního schématu Při výpočtu implicitním schématem se ukazuje, že výpočet lze provádět pro libovolné hodnoty h a τ . Není tedy potřeba splnit další podmínku jak tomu bylo u explicitní metody, implicitní schéma je stabilní vždy. Nicméně i v případě implicitní metody je chyba řádu $\mathcal{O}(h^2 + \tau)$, krok τ by tedy měl být volen kvůli přesnosti zhruba jako h^2 .

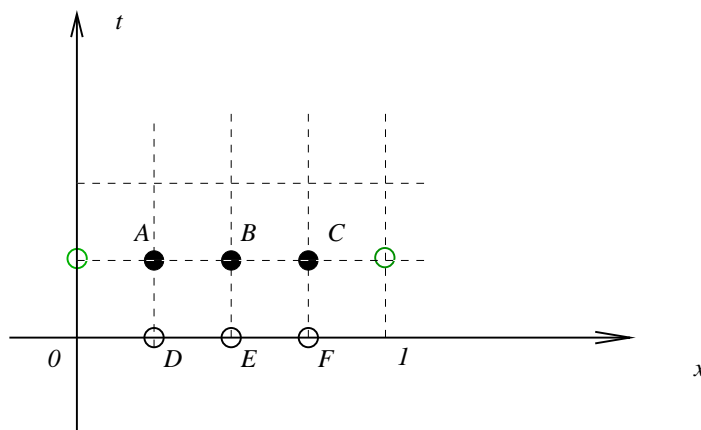
Splnění podmínky stability si ukážeme také zjednodušeným postupem pro případ nulové pravé strany $f \equiv 0$ a nulových okrajových podmínek, tedy řešení (nebo jeho absolutní hodnota) by mělo být v čase klesající nebo alespoň nerostoucí. V maticovém zápisu vidíme, že rovnice (57) lze zapsat

$$\mathbf{A}_I \vec{U}^{k+1} = \vec{U}^k, \quad \vec{U}^{k+1} = \mathbf{A}_I^{-1} \vec{U}^k$$

kde matice \mathbf{A}_I je třídiagonální matice s prvky $(1 + 2\sigma)$ na hlavní diagonále a s prvky $-\sigma$ na diagonálách vedlejších. Opakovaným užitím dostáváme

$$\vec{U}^k = (\mathbf{A}_I^{-1})^k \vec{U}^0.$$

Tedy aby \vec{U}^k pro $k \rightarrow \infty$ bylo nerostoucí, je nutné (i postačující) aby spektrální poloměr matice \mathbf{A}_I^{-1} byl menší nebo roven jedné. Všimněme si v první řadě, že vlastní čísla matice \mathbf{A}_I jsou reálná (rozmyslete si, proč !?) a dle Gerschgorinovy věty v intervalu



Obrázek 9: Oblast řešení, síť a uzly sítě

$\langle 1, 1 + 4\sigma \rangle$. Vlastní čísla matice \mathbf{A}_I^{-1} jsou tedy v intervalu $\langle (1 + 4\sigma)^{-1}, 1 \rangle$ a vidíme tedy, že spektrální poloměr matice \mathbf{A}_I^{-1} je menší roven jedné pro libovolné $\sigma \geq 0$. Implicitní schéma je tedy stabilní pro libovolnou volbu h a τ ! Nicméně i v případě implicitní metody je chyba řádu $\mathcal{O}(h^2 + \tau)$, krok τ by tedy měl být volen kvůli přesnosti zhruba jako h^2 .

Použití implicitního schématu Stejně jako jsme si ukázali jakým způsobem se používá explicitní schéma pro jednoduchý problém, ukážeme si také použití schématu implicitního pro jednoduchost pro tentýž problém, viz zadání Příkladu 10.1. Pro implicitní schéma nejsme pro volbu h a τ nijak omezení podmínkou stability (implicitní schéma je stabilní vždy), volme tedy např. $h = 0.25$ a $\tau = 0.1$. Spočteme

$$\sigma = \frac{1}{5} \frac{\tau}{h^2} = \frac{0.1}{\frac{5}{16}} = 0.32,$$

vidíme, že implicitní schéma budeme používat ve tvaru

$$-0.32U_{i-1}^{k+1} + 1.64U_i^{k+1} - 0.32U_{i+1}^{k+1} = U_i^k + \tau f_i^{k+1}.$$

Rovnice implicitního schématu sestavíme v bodech $A = [0.25, 0.1]$, $B = [0.5, 0.1]$, $C = [0.75, 0.1]$ viz obrázek 9.

Hodnoty U_D , U_E a U_F v bodech $D = [0.25, 0]$, $E = [0.5, 0]$ a $F = [0.75, 0]$ jsou dány počáteční podmínkou. Dále z okrajových podmínek určíme hodnoty $u(0, 0.1) = 0$ a $u(1, 0.1) = 0$. Nyní sestavíme soustavu rovnic pro neznámé U_A , U_B , U_C . Tedy v našem případě

$$-0.32 \cdot 0 + 1.64U_A - 0.32U_B = U_D + 0.1 \cdot (0.25) \cdot (0.1),$$

$$\begin{aligned}-0.32U_A + 1.64U_B - 0.32U_C &= U_E + 0.1 \cdot (0.5) \cdot (0.1), \\ -0.32U_B + 1.64U_C - 0.32 \cdot 0 &= U_F + 0.1 \cdot (0.75) \cdot (0.1),\end{aligned}$$

nebo-li

$$\begin{aligned}1.64U_A \quad -0.32U_B &= 0.19, \\ -0.32U_A \quad 1.64U_B \quad -0.32U_C &= 0.255, \\ -0.32U_B \quad +1.64U_C &= 0.195.\end{aligned}$$

11 Vlnová rovnice

Příkladem hyperbolické diferenciální rovnice je vlnová rovnice, tedy rovnice zapsaná ve tvaru

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} + f(x, t).$$

Fyzikální interpretace této úlohy jsou např. kmity struny, kdy hodnoty u v tomto případě mají význam výchylky (polohy) a hodnoty $\frac{\partial u}{\partial t}$ pak hodnoty rychlosti. Na levé straně rovnice pak druhá derivace má význam zrychlení. Parameter $c > 0$ pak můžeme chápat jako rychlost, jakou se tyto kmity v tělese šíří. Funkce f pak udává působící sílu v místě x a v čase t .

Vlnovou rovnici budeme řešit na oblasti $Q_T = (a, b) \times (0, T)$. Formulaci úlohy je třeba doplnit počáteční podmínkou a okrajovými podmínkami, tím dostáváme tzv. smíšenou úlohu. Na rozdíl od rovnice vedení tepla je v tomto případě potřeba předepsat počáteční podmínku jednak pro u (polohu) ale také i pro $\frac{\partial u}{\partial t}$ (rychlost).

Smíšená úloha pro vlnovou rovnici. Hledáme funkci $u \in \mathcal{C}^2(Q_T) \cup \mathcal{C}(\overline{Q_T})$, $Q_T = (a, b) \times (0, T)$, takovou, že je splněna (vlnová) rovnice ($c > 0$) v oblasti Q_T

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} + f(x, t), \quad (58)$$

a jsou splněny počáteční a okrajové podmínky, tj. pro všechna $x \in \langle a, b \rangle$

$$u(x, 0) = \varphi(x), \quad \frac{\partial u}{\partial t}(x, 0) = \psi(x), \quad (59)$$

a pro $t \geq 0$

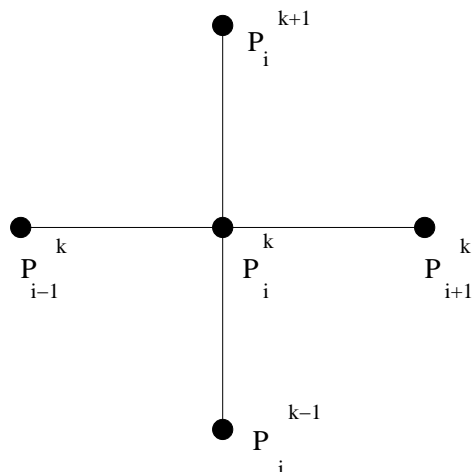
$$u(a, t) = \alpha(t), \quad u(b, t) = \beta(t), \quad (60)$$

Poznámka 11.1 (Podmínky souhlasu) Všimněme si, že v každém z bodů $[a, 0]$ a $[b, 0]$ jsou pro hodnoty u a $\frac{\partial u}{\partial t}$ předepsány dvě podmínky, viz (59) a (60). Je zřejmé, že pro řešitelnost je třeba, aby obě podmínky předepisovaly stejné hodnoty u a $\frac{\partial u}{\partial t}$, tj. aby byly splněny podmínky souhlasu.

Obdobně jako pro rovnici vedení tepla uvažujme nyní speciální případ, kdy interval $(a, b) = (0, \pi)$, $f \equiv 0$ a okrajové podmínky jsou nulové. Počáteční podmínka necht' je dána jako

$$u(x, 0) = \sin(kx), \quad \frac{\partial u}{\partial t}(x, 0) = 0$$

kde $k \in \mathbb{N}$. Řešení hledáme ve tvaru $u(x, t) = \cos(At) \sin(kx)$, kde konstantu A určíme. Vidíme, že pro $k \in \mathbb{N}$ funkce u splňuje okrajové podmínky. Konstanta A pak je dána jako $A = ck$. Rozmyslete si, jaké typické vlastnosti toto řešení má.



Obrázek 10: Znázornění explicitního schématu pro vlnovou rovnici

11.1 Metoda sítí pro vlnovou rovnici

Budeme aproximovat řešení smíšené úlohy pro rovnici

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} + f,$$

v uzlech sítě, která je daná pomocí prostorového kroku $h = (b - a)/n$ a kroku $\tau > 0$. Síťové čáry jsou pak dány rovnicemi $x_i = a + ih$, $t_k = k\tau$ a řešení u pak aproximujeme v síťových uzlech P_i^k jako přibližné řešení U_i^k , tj.

$$P_i^k = [x_i, t_k], \quad U_i^k \approx u_i^k \stackrel{\text{def}}{=} u(x_i, t_k),$$

kde u_i^k budeme značit hodnotu přesného řešení v bodě P_i^k . Z počátečních podmínek určíme hodnoty na nulté a první časové vrstvě, tj. $\vec{U}^0 = (U_i^0)_i$ a $\vec{U}^1 = (U_i^1)_i$. Dále sestavíme schéma pro výpočet hodnot \vec{U}^{k+1} z \vec{U}^k a \vec{U}^{k-1} , které použijeme pro $k = 1, 2, \dots$

Náhrada na první časové vrstvě Na první časové vrstvě uijeme Taylorova rozvoje

$$U_i^1 \approx u(x_i, \tau) = u(x_i, 0) + \frac{\partial u}{\partial t}(x_i, 0)\tau + \mathcal{O}(\tau^2),$$

kde na pravé straně zanedbáme člen $\mathcal{O}(\tau^2)$, a hodnoty u a $\frac{\partial u}{\partial t}$ v čase $t = 0$ dosadíme dle počátečních podmínek. Dostaneme vzorec

$$U_i^1 = \phi(x_i) + \psi(x_i)\tau.$$

11.2 Explicitní schéma

V případě explicitního schématu nahrazujem derivace v bodě P_i^k . Předpokládejme, že řešení u je dostatečně hladké ($u \in C^4(\overline{Q_T})$). Pro hodnoty přesného řešení $u_i^k = u(x_i, t_k)$ v uzlech P_i^k platí

$$\frac{\partial^2 u}{\partial x^2}(P_i^k) = \frac{u_{i-1}^k - 2u_i^k + u_{i+1}^k}{h^2} + \eta_1(h),$$

kde $\eta_1(h) = \mathcal{O}(h^2)$. Dále pro derivaci vzhledem k času dostáváme

$$\frac{\partial^2 u}{\partial t^2}(P_i^k) = \frac{u_i^{k+1} - 2u_i^k + u_i^{k-1}}{\tau^2} + \eta_2(\tau),$$

kde $\eta_2(\tau) = \mathcal{O}(\tau^2)$. Hodnoty přesného řešení tedy splňují rovnici

$$\frac{u_i^{k+1} - 2u_i^k + u_i^{k-1}}{\tau^2} = c^2 \frac{u_{i-1}^k - 2u_i^k + u_{i+1}^k}{h^2} + f_i^k + \mathcal{O}(\tau^2 + h^2),$$

kde $f_i^k = f(P_i^k)$. Zanedbáme tedy člen $\mathcal{O}(\tau^2 + h^2)$ a dostaneme vztah pro přibližné řešení

$$\frac{U_i^{k+1} - 2U_i^k + U_i^{k-1}}{\tau^2} = c^2 \frac{U_{i+1}^k - 2U_i^k + U_{i-1}^k}{h^2} + f_i^k$$

Násobíme τ^2 , označíme $\sigma^2 = \frac{c^2 \tau^2}{h^2}$ a explicitně vyjádříme U_i^{k+1} . Dostaneme síťovou rovnici v uzlu P_i^k pro explicitní schéma pro vlnovou rovnici

$$U_i^{k+1} = \sigma^2 U_{i+1}^k + 2(1 - \sigma^2)U_i^k + \sigma^2 U_{i-1}^k - U_i^{k-1} + \tau^2 f_i^k.$$

Podmínka stability explicitního schématu Obdobně jako tomu bylo u rovnice vedení tepla lze snadno ověřit, že výpočet explicitním schématem není možné provádět pro libovolnou volbu h a τ , ale je třeba splnit ještě dodatečnou podmínku, tzv. stability schématu. V našem případě je stabilita schématu zaručena pro hodnotu $\sigma \leq 1$. Naopak pro velké hodnoty σ je schéma tzv. nestabilní, hodnoty které takto vypočteme jsou pak obvykle nesmyslné.

Stabilita tohoto schématu je zaručena pro

$$\sigma = \frac{c\tau}{h} \leq 1.$$

Použití explicitního schématu Pro lepší pochopení ukážeme na jakým způsobem se používá explicitní schéma ukážeme jeho použití pro jednoduchý příklad a výpočet několika málo hodnot aproximací ve vybraných bodech.

Příklad 11.1 Je dána smíšená úloha pro vlnovou rovnici

$$\frac{\partial^2 u}{\partial t^2} = \left(\frac{9}{5}\right)^2 \frac{\partial^2 u}{\partial x^2} + xt$$

$$\begin{aligned} u(x, 0) &= x^2, & \frac{\partial u}{\partial t}(x, 0) &= 1 - x^2 & \text{pro } x \in \langle -1; 1 \rangle \\ u(-1, t) &= 1, & u(1, t) &= \frac{1}{1+t^2} & \text{pro } t \in \langle 0; \infty \rangle \end{aligned}$$

- Ověřte splnění podmínek souhlasu (pro polohu a rychlost)
- Určete maximální krok τ tak, aby byla splněna podmínka stability pro explicitní metodu s prostorovým krokem $h = 0.2$. Odvoďte schéma pro explicitní metodu.
- Odvoďte soustavu sít'ových rovnic pro určení přibližných hodnot řešení první časové vrstvě s chybou $\mathcal{O}(\tau^2)$.
- Stanovte přibližnou hodnotu řešení v bodě $A = [0.2; 0.2]$ explicitní metodou, volta krok $h = 0.2$ a τ (maximální) tak, aby explicitní schéma bylo stabilní a bod A byl uzlem sítě.

Řešení: Podmínky souhlasu pro polohu (u) i rychlost $\frac{\partial u}{\partial t}$ jsou splněny jak čtenář snadno ověří. Podmínka stability pro explicitní schéma a danou vlnovou rovnici je $\sigma^2 = \frac{81}{25} \frac{\tau^2}{h^2} \leq 1$ (případně $\sigma = \frac{9}{5} \frac{\tau}{h} \leq 1$ - zde ale pozor na možnou záměnu c^2 a c). Volba $h = 0.2$ dává

$$\frac{\frac{81}{25} \tau^2}{0.04} \leq 1 \quad \text{tedy } \tau^2 \leq 0.04 \cdot 25/81 = 1/81 = 1/9^2.$$

Maximální časový krok je tedy $\tau_{max} = 1/9$

Pro část d) volíme krok $h = 0.2$ a $\tau = 0.1 \leq \tau_{max}$. Pro tuto volbu vyjde $\sigma^2 = \frac{81}{25} \frac{0.01}{0.04} = \frac{81}{100} = 0.81$, budeme tedy počítat podle vzorce

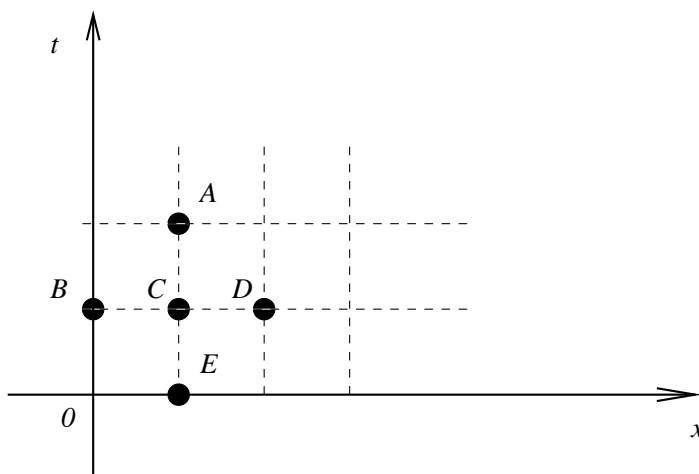
$$U_i^{k+1} = 0.81U_{i+1}^k + 0.38U_i^k + 0.81U_{i-1}^k - U_i^{k-1} + 0.01 f_i^k.$$

Hodnotu v bodě A určíme z hodnot v bodech, které si označíme $B = [0; 0.1]$, $C = [0.2; 0.1]$ a $D = [0.4; 0.1]$. Body B, C, D leží v první časové vrstvě, tedy přibližné hodnoty U_B, U_C, U_D určíme užitím náhrady na první časové vrstvě.

$$U_B = u(0, 0) + 0.1 \frac{\partial u}{\partial t}(0, 0) = 0^2 + 0.1(1 - 0^2) = 0.1,$$

$$U_C = u(0.2, 0) + 0.1 \frac{\partial u}{\partial t}(0.2, 0) = 0.2^2 + 0.1(1 - 0.2^2) = 0.136,$$

$$U_D = u(0.4, 0) + 0.1 \frac{\partial u}{\partial t}(0.4, 0) = 0.4^2 + 0.1(1 - 0.4^2) = 0.244.$$



Obrázek 11: Oblast řešení, síť a uzly sítě

Budeme dále ještě potřebovat hodnotu v bodě $E = [0.2, 0]$, který leží v nulté časové vrstvě a hodnota aproximace je dána počáteční podmínkou, tj. $U_E = u(E) = 0.2^2 = 0.04$. Užitím vzorce pro explicitní schéma dostáváme hodnotu aproximace v bodě A

$$u(0.2, 0.2) \approx U_A = 0.81 U_B + 0.38 U_C + 0.81 U_D - U_E + 0.01 \cdot f(C) = 0.29052,$$

kde hodnota $f(C) = 0.2 \cdot 0.1$ (rozmyslete si proč je funkce f vyčíslena právě v bodě C).

Výpočet dalších aproximací v bodech sítě na 2. časové vrstvě by probíhal obdobně, například hodnotu v bodě $[0.8, 0.2]$ bychom vypočítaly z hodnot aproximací v bodech $[0.6, 0.1]$, $[0.8, 0.1]$ (první časová vrstva), aproximace v bodě $[1, 0.1]$ (okrajová podmínka) a hodnoty $[0.8, 0]$ (nultá časová vrstva). Hodnoty v dalších vrstvách bychom získali stejným postupem.

11.3 Implicitní schéma

Předpokládejme, že řešení u je dostatečně hladké ($u \in C^4(\overline{Q_T})$). V případě implicitního schématu užitíme vztah

$$\frac{\partial^2 u}{\partial x^2}(P_i^k) = \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(P_i^{k+1}) + \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(P_i^{k-1}) + \mathcal{O}(h^2).$$

Ověřte jeho platnost! Dále pak už postupujeme analogicky jako u explicitního schématu, tj. pro hodnoty přesného řešení $u_i^k = u(x_i, t_k)$ platí

$$\frac{\partial^2 u}{\partial x^2}(P_i^k) = \frac{1}{2} \frac{u_{i-1}^{k+1} - 2u_i^{k+1} + u_{i+1}^{k+1}}{h^2} + \frac{1}{2} \frac{u_{i-1}^{k-1} - 2u_i^{k-1} + u_{i+1}^{k-1}}{h^2} + \eta_1(h),$$

kde $\eta_1(h) = \mathcal{O}(h^2)$. Dále pro derivaci vzhledem k času

$$\frac{\partial^2 u}{\partial t^2}(P_i^k) = \frac{u_i^{k+1} - 2u_i^k + u_i^{k-1}}{\tau^2} + \eta_2(\tau),$$

kde $\eta_2(\tau) = \mathcal{O}(\tau^2)$. Hodnoty přesného řešení tedy splňují rovnici

$$\frac{u_i^{k+1} - 2u_i^k + u_i^{k-1}}{\tau^2} = \frac{c^2}{2} \frac{u_{i-1}^{k+1} - 2u_i^{k+1} + u_{i+1}^{k+1}}{h^2} + \frac{c^2}{2} \frac{u_{i-1}^{k-1} - 2u_i^{k-1} + u_{i+1}^{k-1}}{h^2} + f_i^k + \mathcal{O}(\tau^2 + h^2),$$

kde $f_i^k = f(P_i^k)$. Zanedbáme tedy člen $\mathcal{O}(\tau^2 + h^2)$ a dostaneme vztah pro přibližné řešení

$$\frac{U_i^{k+1} - 2U_i^k + U_i^{k-1}}{\tau^2} = \frac{c^2}{2} \frac{U_{i+1}^{k+1} - 2U_i^{k+1} + U_{i-1}^{k+1}}{h^2} + \frac{c^2}{2} \frac{U_{i+1}^{k-1} - 2U_i^{k-1} + U_{i-1}^{k-1}}{h^2} + f_i^k$$

Vynásobíme rovnici τ^2 , označíme

$$\sigma^2 = \frac{c^2 \tau^2}{h^2}$$

a převedeme členy U_i^{k+1} na levou stranu. Pro pevné k dostáváme soustavu rovnic

$$-\frac{1}{2}\sigma^2 U_{i-1}^{k+1} + (1 + \sigma^2)U_i^{k+1} - \frac{1}{2}\sigma^2 U_{i+1}^{k+1} = \frac{1}{2}\sigma^2 U_{i-1}^{k-1} - (1 + \sigma^2)U_i^{k-1} + \frac{1}{2}\sigma^2 U_{i+1}^{k-1} + 2U_i^k + \tau^2 f_i^k,$$

pro $i = 1, \dots, n-1$, kde matice této soustavy je symetrická, pozitivně definitní a ostře diagonálně dominantní.

Podmínka stability implicitního schématu Stejně jako tomu bylo u rovnice vedení tepla, také zde je implicitní schéma vždy stabilní. Také diskretizační chyba, které se dopouštíme je v případě explicitní i implicitní metody stejná, $\mathcal{O}(h^2 + \tau^2)$.

Použití implicitního schématu Použijme nyní implicitní schéma pro řešení Příkladu 11.1. Jak víme, tak implicitní schéma je vždy stabilní, takže nejsme nijak omezeni pro volbu h a τ . V praktických výpočtech volíme tak malé, aby získaná aproximace byla co možná nejpřesnější. Zde ale zvolíme h a τ tak, aby se nám snadno prováděl výpočet. Volme $h = 0.5$ a $\tau = 0.5$. Hodnota $\sigma^2 = \frac{81}{25} = 3.24$ (implicitní metoda je stabilní). Hodnoty na nulté a první časové vrstvě získáme z počátečních podmínek stejně jako u explicitního schématu. Tedy dostaneme následující tabulku hodnot

i	0	1	2	3	4
x_i	-1	-0.5	0	0.5	1
$1 - x_i^2$	0	0.75	1	0.75	0
$U_i^0 = x_i^2$	1	0.25	0	0.25	1
U_i^1	1	0.625	0.5	0.625	0.8

Pro hodnoty U_i^2 na 2. časové vrstvě sestavíme soustavu rovnic dle implicitního schématu pro dané σ

$$-1.62U_{i-1}^{k+1} + 4.24U_i^{k+1} - 1.62U_{i+1}^{k+1} = 1.62U_{i-1}^{k-1} - 4.24U_i^{k-1} + 1.62U_{i+1}^{k-1} + 2U_i^k + \tau^2 f_i^k.$$

Nejprve určíme z okrajových podmínek hodnoty $U_0^2 = 1$ a $U_4^2 = 0.5$. Dále sestavíme soustavu rovnic pro neznámé U_1^2, U_2^2, U_3^2 . Tedy v našem případě

$$\begin{aligned} -1.62 \cdot 1 + 4.24U_1^2 - 1.62U_2^2 &= 1.62 \cdot 1 - 4.24 \cdot 0.25 + \\ &+ 1.62 \cdot 0 + 2 \cdot 0.625 + 0.25 \cdot (-0.5) \cdot (0.5). \\ -1.62 \cdot U_1^2 + 4.24U_2^2 - 1.62U_3^2 &= 1.62 \cdot 0.25 - 4.24 \cdot 0 + \\ &+ 1.62 \cdot 0.25 + 2 \cdot 0.5 + 0.25 \cdot (0) \cdot (0.5). \\ -1.62U_2^2 + 4.24U_3^2 - 1.62 \cdot 0.5 &= 1.62 \cdot 0 - 4.24 \cdot 0.25 + \\ &+ 1.62 \cdot 1 + 2 \cdot 0.625 + 0.25 \cdot (0.5) \cdot (0.5). \end{aligned}$$

nebo-li

$$\begin{aligned} 4.24U_1^2 - 1.62U_2^2 &= 1.8375, \\ -1.62 \cdot U_1^2 + 4.24U_2^2 - 1.62U_3^2 &= -0.185, \\ -1.62U_2^2 + 4.24U_3^2 &= 0.8285. \end{aligned}$$

Vidíme, že matice soustavy je třídiagonální, symetrická, ostře diagonálně dominantní a pozitivně definitní (a tedy regulární). Existuje jediné řešení, které lze nalézt např. Jacobiovou nebo Gaussovou-Seidelovou iterační metodou. Vzhledem k tomu, že matice je ostře diagonálně dominantní v řádcích, lze získat odhad pro řádkovou normu matice a tedy i odhad pro počet iterací potřebných k tomu, aby bylo dosaženo potřebné přesnosti. Vlastnosti soustavy pro menší krok h a tedy větší počet neznámých jsou stejné. Získané řešení dává hodnoty aproximací přesného řešení na druhé časové vrstvě. Pro výpočet hodnot na třetí časové vrstvě sestavíme novou soustavu a provedeme její vyřešení. Takto získáváme postupně hodnoty na 2., 3., 4., ... časové vrstvě.

12 Klasifikace PDR, opakování

V předchozích přednáškách jsme se seznámily s parciálními diferenciálními rovnicemi 2. řádu. Viděli jsme, že se podle typu dělí na eliptické, parabolické a hyperbolické. Viděli jsme, že charakteristické vlastnosti diferenciálních rovnic a vlastnosti jejich řešení se dle typu liší. Podle tohoto typu je také nutné volit způsob numerického řešení a to ať již pro metodu konečných diferencí nebo i pro jiné numerické metody (metoda konečných prvků nebo metoda konečných objemů).

Příkladem eliptické rovnice je Poissonova rovnice, parabolické pak rovnice vedení tepla a hyperbolické rovnice vlnová.

$$\text{Poissonova rovnice} \quad -\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = f,$$

$$\text{Rovnice vedení tepla } (p > 0) \quad \frac{\partial u}{\partial t} = p \frac{\partial^2 u}{\partial x^2} + f,$$

$$\text{Vlnová rovnice } (c > 0) \quad \frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} + f.$$

PDR jako formulace fyzikálních zákonů Uvedme další příklady parciálních diferenciálních rovnic, které se vyskytují v technických problémech. Rovnice vedení tepla v tělese reprezentovaném oblastí Ω je popsána

$$\rho \frac{\partial \theta}{\partial t} - \nabla \cdot (\lambda \nabla \theta) = f.$$

Jedná se o parabolickou rovnici. V případě hledání stacionárního stavu (uvažujeme řešení, která jsou nezávislá na proměnné t) se jedná o stacionární problém popsany eliptickou rovnicí

$$-\nabla \cdot (\lambda \nabla \theta) = f.$$

Dalším případem eliptické rovnice jsou Lamého rovnice popisující malé deformace isotropního elastického tělesa (lineární elasticita) pro stacionární případ

$$-\lambda \nabla (\nabla \cdot \mathbf{u}) + \mu \nabla \cdot (\nabla \mathbf{u} + (\nabla \mathbf{u})^T) = \mathbf{f}.$$

Pro nestacionární případ je pohyb tělesa popsán rovnicemi ve tvaru

$$\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} - \lambda \nabla (\nabla \cdot \mathbf{u}) + \mu \nabla \cdot (\nabla \mathbf{u} + (\nabla \mathbf{u})^T) = \mathbf{f}.$$

Vidíme že se jedná o rovnici hyperbolickou (vlnová rovnice). Ve všech předchozích případech se jednalo o rovnice lineární.

Příkladem nelineární rovnice (soustavy rovnic) jsou Navierovy-Stokesovy rovnice popisující proudění nestlačitelné vazké tekutiny, které jsou diferenciálním vyjádřením zákona zachování hybnosti a zákona zachování hmoty, tj.

$$\rho \frac{\partial \mathbf{u}}{\partial t} + \rho (\mathbf{u} \cdot \nabla) \mathbf{u} - \mu \Delta \mathbf{u} + \nabla p = \mathbf{f}.$$

V těchto rovnicích jsou neznámé složky vektoru rychlosti \mathbf{u} a tlak p , vidíme, že se typově jedná o parabolické rovnice. Vzhledem k počtu neznámých a počtu rovnic, je nutné doplnit systém Navierových-Stokesových rovnic další rovnicí, rovnicí kontinuity, která je diferenciálním vyjádřením zákona zachování hmoty. Rovnice kontinuity má tvar

$$\nabla \cdot \mathbf{u} = 0,$$

zde ale vidíme, že se jedná o rovnici 1. řádu (nemá tedy smysl hovořit o typu rovnice).

V případě stlačitelné tekutiny je její proudění vyjádřeno pomocí zákona zachování hmoty, hybnosti a energie. Pro jednoduchost formulujme jenom zákon zachování hmoty, který v tomto případě má tvar

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0.$$