

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

Fakulta stavební

Numerická analýza a programování II

doc. Ing. Zdeněk Vospěl, CSc.

1992

Ediční středisko ČVUT, Praha 6, Zikova 4

© Zdeněk Vospěl, 1992
ISBN 80-01-00807-X

K A P I T O L A 1.

Ú V O D

1.1 Co je numerická matematika

Řečeno zcela lapidárně, je to takové odvětví matematiky, které se zabývá řešením matematicky formulovaných úloh pomocí aritmetických a logických operací.

Mnohé matematické úlohy jsou formulovány přímo jako úlohy numerické a pak pro jejich řešení stačí pouze nalézt vhodný algoritmus, tj. deterministický popis konečné posloupnosti operací, která převede množinu vstupních dat na množinu dat výstupních. Tento typ úloh se vyskytuje např. v lineární algebře: výpočet determinantu, řešení soustav lineárních algebraických rovnic apod. Zcela jinak tomu bývá při numerickém řešení úloh matematické analýzy (např. při výpočtech derivací a integrálů), kdy matematickou úlohu musíme nahradit (aproximovat) nějakou úlohou numerickou a teprve její řešení nám slouží jako náhrada (aproximace) řešení původní úlohy. Kromě předcházejících dvou typů úloh jsou i takové numerické úlohy, které pro jejich složitost převádíme na úlohy jednodušší.

Můžeme tedy říci, že numerickou metodou rozumíme postup výpočtu numerické úlohy, nebo její převod na úlohu jednodušší či postup, který nahrazuje úlohu matematickou úlohou numerickou.

V numerické matematice hraje velkou roli i zkušenosť a intuice, které řešitelé pomohou při výběru "vhodné" metody. Tato "vhodnost" závisí kromě vlastní úlohy i na výpočetním prostředku, požadované přesnosti řešení, termínu ukončení úlohy apod.

Rozvoj numerické matematiky byl značně urychljen v posledních dvaceti letech, kdy rozvoj výpočetní techniky umožnil řešit i takové úlohy, které byly z ekonomických nebo časových důvodů dříve neřešitelné.

1.2 Chyby a jejich charakteristiky

Každá numerická úloha je při své realizaci zatížena třemi druhy chyb:

- chybami vstupních údajů (tzv. diskretizačními chybami), které mohou vznikat jako chyby z měření vstupních údajů, nebo jako chyby z generalizace fyzikální skutečnosti;
- chybou metody, tj. chybou z approximace matematické úlohy úlohou numerickou;
- zaokrouhlovacími chybami, které vznikají v průběhu výpočtů zaokrouhlováním mezi-výsledků.

Z technických důvodů a z důvodů existence chyb, které jsme právě klasifikovali, se výpočty provádějí s omezeným počtem cifer, a proto mnohdy ztrácejí platnost tvrzení, která mlžky předpokládají přesné údaje. Přesné číslo je tedy ve výpočtu nahrazeno přibližným číslem (approximací) a toto nahrazení je zatíženo určitou chybou. Proto při práci s přibližnými čísly zavádíme pojem absolutní a relativní chyby.

Definice 1.1 Je-li x přibližná hodnota čísla X , pak absolutní chyba $A(x)$, resp.

relativní chyba $R(x)$ přibližného čísla x jsou definovány těmito vztahy:

$$A(x) = |X - x|, \text{ resp. } R(x) = A(x)/|X|.$$

Ve většině případů věk číslo X zůstává neznámé a chyby $A(x)$ a $R(x)$ nemůžeme nikdy zjistit. Proto místo nich zavádíme jejich (horní) odhady.

Definice 1.2 Nechť x je přibližná hodnota čísla X a $A(x)$, $R(x)$ je absolutní a relativní chyba této aproximace. Pak odhadem absolutní, resp. relativní chyby jsou čísla $a(x)$, resp. $r(x)$, pro která platí

$$A(x) = |X - x| \leq a(x), \text{ resp. } R(x) = \frac{A(x)}{|X|} \leq r(x).$$

V praxi je $X \approx x$ a proto za odhad relativní chyby bereme $r(x) = a(x)/|x|$. Procentová chyba = $r(x) \cdot 100\%$.

Pomocí zavedených pojmu můžeme napsat velmi důležitou nerovnost

$$(1.1) \quad x - a(x) \leq X \leq x + a(x),$$

které se stručně zapisuje takto:

$$(1.2) \quad X = x \pm a(x).$$

Při hodnocení výsledků nás především zajímají cifry, které již nejsou zatíženy chybou, tzv. platné cifry.

Definice 1.3 Mějme přibližné číslo x zapsané pomocí cifer $x_1, x_2, x_3, \dots, x_n$, kde $x_1 \neq 0$, ve tvaru

$$x = \pm 0, x_1 x_2 \dots x_n \cdot 10^p.$$

Jestliže platí

$$A(x) \leq 0,5 \cdot 10^{p-n},$$

pak říkáme, že cifry $x_1 x_2 \dots x_n$ jsou platné (zaručené).

1.3 Šíření chyb ve výpočtech

K analýze chyb výsledku numerické úlohy je nezbytné znát odhad chyb výsledků základních početních operací, případně výsledku výpočtu funkční hodnoty.

Mějme dána čísla $X = x \pm a(x)$ a $Y = y \pm a(y)$, tedy podle (1.1)

$$x \in \langle x - a(x), x + a(x) \rangle, \quad y \in \langle y - a(y), y + a(y) \rangle.$$

Pak platí následující nerovnosti:

a) při sčítání

$$(1.3) \quad (x \pm y) - (a(x) + a(y)) \leq X \pm Y \leq (x \pm y) + (a(x) + a(y));$$

b) při násobení (předpokládejme, že čísla X, x, Y, y jsou kladná)

$$(1.4) \quad \begin{aligned} X \cdot Y &\leq (x + a(x))(y + a(y)) = xy + xa(y) + ya(x) + a(x) \cdot a(y), \\ X \cdot Y &\geq xy - xa(y) - ya(x) + a(x) \cdot a(y); \end{aligned}$$

c) při dělení (nechť opět X, x, Y, y jsou kladná čísla)

$$(1.5) \quad \frac{x}{y} \leq \frac{x + a(x)}{y - a(y)} = \frac{x}{y} + \frac{xa(y) + ya(x)}{y(y - a(y))},$$

$$\frac{x}{Y} \geq \frac{x - a(x)}{y + a(y)} = \frac{x}{y} + \frac{xa(y) + ya(x)}{y(y + a(y))}.$$

Předpoklad pozitivnosti čísel X, x, Y, y není na újmu obecnosti, protože šíření chyb na znaménku operandů nezávisí. Dále předpokládáme, že odhadы absolutních chyb $a(x), a(y)$ jsou mnohem menší než čísla x, y .

Z nerovností (1.3), (1.4) a (1.5) plyne tvrzení následující věty, jestliže v nerovnostech (1.4) zanedbáme členy $a(x).a(y)$, které jsou oproti ostatním členům velmi malé a v nerovnostech (1.5) hodnotu $a(y)$ ve jmenovateli, která je oproti y opět velmi malá.

Věta 1.1 Mějme dána dvě čísla X, Y svými přibližnými hodnotami x, y a odhadы absolutních chyb $a(x), a(y)$. Pak odhadы absolutních, resp. relativních chyb, aritmetických operací jsou dány vzorcemi:

$$\begin{aligned} a(x \pm y) &= a(x) + a(y), \\ a(x \cdot y) &= |x|a(y) + |y|a(x), \\ a(x / y) &= (|x|a(y) + |y|a(x))/y^2, \\ r(x \cdot y) &= r(x) + r(y), \\ r(x / y) &= r(x) + r(y). \end{aligned}$$

Nakonec ještě odvodíme odhad absolutní chyby funkční hodnoty funkce u v proměnných. Nechť tedy

$$u = f(t_1, t_2, \dots, t_n)$$

je diferencovatelná funkce, jejíž hodnotu chceme vypočítat pro argumenty

$X_1 = x_1 \pm a(x_1), X_2 = x_2 \pm a(x_2), \dots, X_n = x_n \pm a(x_n)$. Pro odhad absolutní chyby výsledku platí

$$(1.6) \quad a(u) = \sum_{i=1}^n \left| \frac{\partial f}{\partial t_i} \right| a(x_i),$$

kde derivace se počítají pro přibližné hodnoty $t_1 = x_1, t_2 = x_2, \dots, t_n = x_n$. Důkaz tohoto vzorce je např. v [9].

Speciálně pro funkci jedné proměnné $u = f(t)$ přejde vzorec (1.6) ve vzorec

$$(1.7) \quad a(u) = \left| \frac{df}{dt} \right| a(x).$$

Poznámka 1.1 Ve výpočtech se snažíme vyhnout odečítání blízkých čísel x, y , protože potom může být relativní chyba rozdílu značně velká, neboť

$$r(x - y) = \frac{a(x) + a(y)}{|x - y|}$$

a tím dochází ke ztrátě platných cifer. Je-li např. $X = 0,129 \pm 0,0003, Y = 0,124 \pm$

$\pm 0,0003$ pak $x - y = 0,005$, $a(x - y) = 0,0006$, $r(x - y) = 0,12$. Tedy podle definice 1.3 nemá rozdíl žádnou platnou cifru.

1.4 Numerická stabilita

Dříve, než začneme hovořit o stabilitě, zavedeme pojem korektní úlohy.

Definice 1.4 Nechť M , N jsou množiny vstupních a výstupních dat úlohy U , tedy

$$y = U(x), \quad x \in M, \quad y \in N.$$

Tuto úlohu nazýváme korektní na množinách M , N , jestliže platí:

- a) ke každému $x \in M$ existuje právě jedno řešení $y \in N$,
- b) řešení spojitě závisí na vstupních datech; tzn. když $x_n \xrightarrow{\varphi} x$, $x_n \in M$, $U(x_n) = y_n$, pak $U(x_n) \xrightarrow{G} U(x) = y$, kde φ a G jsou metriky na množinách M , N .

U korektních úloh nás bude zajímat jak sama úloha, případně její konkrétní algoritmus, jsou citlivé na změny vstupních údajů. Citlivost úlohy na tuto změnu nazýváme podmíněností.

Definice 1.5 Mějme dánu korektní úlohu z definice 1.4. Budeme říkat, že je dobré podmíněná, jestliže malá změna vstupních dat vyvolá malou změnu řešení. Podmíněnost hodnotíme číslem podmíněnosti úlohy

$$(1.8) \quad C_p = \frac{\|\Delta y\|}{\|y\|} / \frac{\|\Delta x\|}{\|x\|},$$

kde změně Δx odpovídá změna Δy .

V praxi vypočítáváme číslo podmíněnosti z relativních chyb, tedy

$$(1.9) \quad C_p \approx r(x)/r(y).$$

Je-li $C_p = 1$, je úloha (velmi) dobré podmíněná a při $C_p > 100$ hovoříme o špatné podmíněnosti. Podmíněnost se netýká jenom úlohy samotné, ale i způsobu řešení, tedy algoritmu úlohy. Proto bychom zcela obdobně mohli definovat citlivost algoritmu k zaokrouhlovacím chybám a chybám vstupních údajů.

Abychom se ve výpočtech vyvarovali nesmyslných výsledků, musíme si vybírat algoritmy, které jsou málo citlivé k zaokrouhlovacím chybám i chybám vstupních údajů. Takové algoritmy se nazývají stabilní.

Příklad 1.1 Mějme dánu soustavu

$$x + \alpha y = 1,$$

$$\alpha x + y = 0,$$

kde $\alpha \neq \pm 1$ je volitelný vstupní údaj. Máme zjistit podmíněnost úlohy při výpočtu hodnoty x pro různá α .

Řešení: Protože $x = 1/(1 - \alpha^2)$, pak číslo podmíněnosti podle (1.7) a (1.8) je

$$C_p \approx \left| \left(\alpha \frac{dx}{d\alpha} \right) / x \right| = \frac{2\alpha^2}{|1 - \alpha^2|},$$

z čehož plyne, že pro α^2 blízké 1 je úloha špatně podmíněná a pro $\alpha^2 \gg 1$ je úloha podmíněna dobrě.

KAPITOLA 2.

ÚLOHY LINEÁRNÍ ALGEBRY

2.1 Některé pojmy

Základní pojmy lineární algebry jsou matici a vektori, se kterými jste se již dostatečně obeznámili v matematice. Budeme tedy předpokládat znalost základních druhů matic jako jsou: jednotková, nulová, čtvercová, diagonální, trojúhelníková, transponovaná, symetrická a inverzní. Dále předpokládáme znalost pojmu determinant, jeho definici, způsob výpočtu, základní pravidla pro počítání s determinanty a konečně i pojem regulární a singulární matice.

2.1.1 Ortogonalita vektorů a matic

Pro potřeby tohoto učebního textu budeme matici typu $(n, 1)$ nazývat (n-členný sloupcový) vektor a označovat malým písmenem s podtržením, tedy

$$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (x_1, x_2, \dots, x_n)^T,$$

kde horní index T značí transpozici jako v maticovém počtu. Matici typu $(1, n)$ budeme nazývat (řádkový) vektor a označovat

$$\underline{y}^T = (y_1, y_2, \dots, y_n).$$

Skalární součin dvou aritmetických n-členných vektorů $\underline{u}, \underline{v}$ se členy $u_i, v_i, i = 1, 2, \dots, n$ budeme označovat $(\underline{u}, \underline{v})$ a počítat podle vzorce

$$(\underline{u}, \underline{v}) = u_1 v_1 + u_2 v_2 + \dots + u_n v_n.$$

Dále zavedeme méně obvyklý pojem ortogonality soustavy vektorů a matice.

Definice 2.1 Nechť $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n$ je soustava nenulových vektorů. Jestliže platí

$$(\underline{v}_i, \underline{v}_j) \begin{cases} = 0 & \text{pro } i \neq j \\ \neq 0 & (\text{resp. } = 1) \text{ pro } i = j \end{cases}$$

pak říkáme, že soustava vektorů je ortogonální (resp. ortonormovaná).

Definice 2.2 Čtvercová matice A se nazývá ortogonální, jestliže její řádky tvoří ortonormovanou soustavu vektorů.

2.1.2 Matice rozdělená na bloky

V obecné matici $A = (a_{ij})$ můžeme rozdělit řádky a sloupce vodorovnými a svislými čarami do skupin tak, že vzniknou bloky s tzv. submaticemi

$$A = \begin{pmatrix} a_{11} & \cdots & \underbrace{\quad}_{n_s} \\ \vdots & & \\ & \boxed{A_{rs}} & \\ & & \left. \right\} m_r \\ & & \vdots \\ \cdots & a_{mn} \end{pmatrix},$$

kdy např. r -tá skupina řádků a s -tá skupina sloupců definují submatice A_{rs} typu (m_r, n_s) . Matici rozdělenou na bloky budeme značit $A = (A_{rs})$, $r = 1, 2, \dots, t$, $s = 1, 2, 3, \dots, u$.

Pro matice rozdělené na bloky definujeme součet a součin.

Nechť $A = (A_{rs})$ a $B = (B_{rs})$ jsou dvě matice téhož typu se shodným dělením na bloky, pak jejich součet je definován vzorcem

$$(2.1) \quad A + B = C = (A_{rs} + B_{rs}) = (C_{rs}).$$

Tedy součet je matice stejného typu, opět se shodným dělením na bloky, kdy submatice součtu C_{rs} je součtem submatic A_{rs} a B_{rs} se stejnou indexací.

Součin definujeme pro matice A , B , kde $A = (A_{rs})$, $r = 1, 2, \dots, t$, $s = 1, 2, \dots, v$ je typu (m, n) a matice $B = (B_{rs})$, $r = 1, 2, \dots, t$, $s = 1, 2, \dots, u$ je typu (n, p) . Matice A má tedy rozdělené sloupce na stejný počet skupin jako matice B řádky. Dělení sloupců matice A a řádků matice B se musí shodovat nejen v počtu skupin, ale i v počtu řádků (sloupců) v odpovídajících si skupinách. Tedy submatice $A_{1k}, A_{2k}, \dots, A_{tk}$ mají stejný počet sloupců jako submatice $B_{k1}, B_{k2}, \dots, B_{ku}$ řádků. Pak součinem matic A , B je matice C , kde

$$(2.2) \quad C = A \cdot B = (C_{rs}),$$

$$C_{rs} = \sum_{k=1}^v A_{rk} B_{ks}, \quad r = 1, 2, \dots, t, \quad s = 1, 2, \dots, u.$$

Mějme např. matici $A = (A_{rs})$, $r = 1, 2, \dots, t$, $s = 1, 2, 3$, $B = (B_{rs})$, $r = 1, 2, \dots, t$, $s = 1, 2$, pak

$$\begin{array}{c} m_1 \left\{ \begin{array}{|c|c|c|} \hline A_{11} & A_{12} & A_{13} \\ \hline \hline \end{array} \right. \\ \hline \begin{array}{|c|c|c|} \hline A_{21} & A_{22} & A_{23} \\ \hline \hline \end{array} \right\} \begin{array}{c} P_1 \left\{ \begin{array}{|c|c|} \hline B_{11} & B_{12} \\ \hline \hline \end{array} \right. \\ \hline \begin{array}{|c|c|} \hline B_{21} & B_{22} \\ \hline \hline \end{array} \right\} \\ \hline \begin{array}{|c|c|} \hline B_{31} & B_{32} \\ \hline \hline \end{array} \right\} \end{array} = \begin{array}{c} m_1 \left\{ \begin{array}{|c|c|} \hline C_{11} & C_{12} \\ \hline \hline \end{array} \right. \\ \hline \begin{array}{|c|c|} \hline C_{21} & C_{22} \\ \hline \hline \end{array} \right\} \\ \hline \end{array},$$

$$\begin{array}{cc} P_1 & P_2 \\ \hline \hline \end{array} \quad \begin{array}{cc} n_1 & n_2 \\ \hline \hline \end{array}$$

kde

$$C_{11} = A_{11} \cdot B_{11} + A_{12} \cdot B_{21} + A_{13} \cdot B_{31} \quad C_{12} = A_{11} \cdot B_{12} + A_{12} \cdot B_{22} + A_{13} \cdot B_{32}$$

$$C_{21} = A_{21} \cdot B_{11} + A_{22} \cdot B_{21} + A_{23} \cdot B_{31} \quad C_{22} = A_{21} \cdot B_{12} + A_{22} \cdot B_{22} + A_{23} \cdot B_{32}$$

2.1.3 Normy matic a vektorů

Norma matice A je číslo, které v jistém smyslu charakterizuje míru její velikosti. Budeme ji označovat $\|A\|$ a žádat, aby měla tyto vlastnosti:

- 1) $\|A\| \geq 0$ a $\|A\| = 0$ právě tehdy, když $A = 0$,
- 2) $\|c \cdot A\| = |c| \cdot \|A\|$, pro $c \in \mathbb{R}$,
- 3) $\|A + B\| \leq \|A\| + \|B\|$ a $\|A \cdot B\| \leq \|A\| \cdot \|B\|$.

Tento požadavkům vyhovují následující definice, kde A je matice typu (m, n) s prvky a_{ij} :

$$\|A\|_R = \max_i \sum_{j=1}^n |a_{ij}|, \text{ tzv. řádková norma,}$$

$$\|A\|_S = \max_j \sum_{i=1}^m |a_{ij}|, \text{ tzv. sloupcová norma,}$$

$$\|A\|_E = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2}, \text{ tzv. eukleidovská norma.}$$

Pro vektor \underline{x} , který chápeme jako matici typu $(m, 1)$ dostáváme z předcházejícího tyto definice:

$$\|\underline{x}\|_R = \max_i |\underline{x}_i|, \quad \|\underline{x}\|_S = \sum_{i=1}^m |\underline{x}_i|, \quad \|\underline{x}\|_E = \sqrt{\sum_{i=1}^m |\underline{x}_i|^2}.$$

Máme-li tedy např. vektor $A \cdot \underline{x}$, pak pro odpovídající si normy platí podle třetího axiomu

$$\|A \cdot \underline{x}\| \leq \|A\| \cdot \|\underline{x}\|.$$

2.1.4 Speciální matice

Jestliže většina prvků matice je nulová, pak o ní říkáme, že je řídká. V opačném případě hovoříme o plné matici, případně pouze o matici. Termíny řídká a plná nejsou exaktně definovány a užívají se zcela volně. Toto dělení nabývá značného významu při zpracovávání velkých řídkých matic pomocí počítačů, protože můžeme vypustit operace s nulovými prvky včetně jejich uložení v paměti.

Definice 2.3 O matici A typu (m, n) říkáme, že je pásová, jestliže platí

$$a_{ij} = 0 \text{ pro } |i - j| > p,$$

kde $0 < p < \min\{m, n\}$. To znamená, že všechny prvky vně pásu kolem diagonály jsou nulové. Je-li $p = 1$, hovoříme o třídiagonální matici, při $p = 2$ o pětidiagonální apod.

Definice 2.4 O symetrické matici $A = (a_{ij})$ typu (n, n) říkáme, že je pozitivně definitní, jestliže pro libovolný nenulový vektor \underline{x} platí

$$(2.3) \quad (\underline{x}, A \cdot \underline{x}) > 0.$$

Definice 2.5 O matici $A = (a_{ij})$ typu (n, n) říkáme, že má dominantní diagonálu (je diagonálně dominantní), jestliže

$$(2.4) \quad |a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \quad \text{pro } i = 1, 2, \dots, n.$$

2.2 Základní úlohy lineární algebry

První základní úlohou lineární algebry je řešení nehomogenní soustavy lineárních algebraických rovnic, která je daná maticí A typu (n, n) a n -členným nenulovým vektorem \underline{b} . Máme za úkol nalézt takový n -členný vektor \underline{x} , pro který je splněna maticová rovnice

$$(2.5) \quad A \cdot \underline{x} = \underline{b}.$$

Soustava (2.5) je jednoznačně řešitelná, jestliže matice A je regulérní, tzn. když determinant matice A je nenulový.

V praxi se setkáváme s velmi rozličnými soustavami, od malých s několika rovniciemi až po velké se stovkami či tisíci rovnic, které jsou definovány plnými nebo řídícími maticemi, či maticemi pásovými apod. Tato rozmanitost úloh vedla k vytvoření mnoha metod řešení, které se rozdělují na přímé a iterační, z nichž každá má jisté výhody a také nevýhody, takže nějaká univerzální optimální metoda řešení neexistuje.

Druhá základní úloha (kterou se budeme zabývat jenom s ohledem na čtvercové matice), je řešení maticové rovnice

$$(2.6) \quad A \cdot X = B,$$

kde A, B jsou dané a X je hledaná matice typu (n, n) . Zapíšeme-li matice A, X, B pomocí složek, pak rovnice (2.6) je ekvivalentní s řešením n soustav rovnic

$$(2.7) \quad \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_{1j} \\ x_{2j} \\ \vdots \\ x_{nj} \end{pmatrix} = \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{pmatrix}, \quad j = 1, 2, \dots, n.$$

Speciálním případem této úlohy je nalezení inverzní matice ke čtvercové regulérní matici A , tj. matice označené A^{-1} , pro kterou platí

$$(2.8) \quad A \cdot A^{-1} = E,$$

kde E je jednotková matice.

Třetí základní úloha je úloha o nalezení vlastních čísel a vektorů čtvercové matice A . V této úloze hledáme všechna čísla λ , pro které existuje netriviální řešení homogenní soustavy

$$(2.9) \quad (A - \lambda \cdot E) \cdot \underline{x} = 0, \quad \text{resp.} \quad A \cdot \underline{x} = \lambda \cdot \underline{x}.$$

Taková λ nazýváme vlastní čísla a odpovídající řešení \underline{x} vlastní vektory matice A .

2.3 Přímé metody řešení soustav lineárních algebraických rovnic

Přímé metody řešení soustavy (2.5) jsou takové, kterými dostaneme přesný výsledek při realizaci konečného počtu aritmetických operací, když bychom je provedli přesně, tj. bez zaokrouhllování. Všechny přímé metody jsou jenom jistou obměnou základního Gaussova eliminacního postupu, jehož smyslem je převést zadanou soustavu s plnou maticí na soustavu s maticí trojúhelníkovou. Proto se budeme v tomto odstavci nej-

dříve zabývat řešením tzv. trojúhelníkových soustav, pak si probereme Gaussovu eliminacní metodu, její modifikace a některé metody navazující.

2.3.1 Řešení trojúhelníkových soustav

Mějme dány soustavy

$$(2.10) \quad L \cdot \underline{x} = \underline{b}, \quad \text{resp.} \quad U \cdot \underline{y} = \underline{d},$$

kde L je regulérní dolní a U je regulérní horní trojúhelníková matice. Označme $L = (\ell_{ij})$, $U = (u_{ij})$, $i, j = 1, 2, \dots, n$, pak soustavy (2.10) mají tvar

$$(2.11) \quad \begin{aligned} \ell_{11}x_1 &= b_1 & u_{11}y_1 + u_{12}y_2 + \dots + u_{1n}y_n &= d_1 \\ \ell_{21}x_1 + \ell_{22}x_2 &= b_2 & u_{22}y_2 + \dots + u_{2n}y_n &= d_2 \\ \dots & \dots & \dots & \dots \\ \ell_{n1}x_1 + \ell_{n2}x_2 + \dots + \ell_{nn}x_n &= b_n & u_{nn}y_n &= d_n \end{aligned}$$

tedy řešení je dáno vzorcí

$$(2.12) \quad x_i = \frac{1}{\ell_{ii}} \left(b_i - \sum_{j=1}^{i-1} \ell_{ij}x_j \right), \quad y_k = \frac{1}{u_{kk}} \left(d_k - \sum_{j=k+1}^n u_{kj}y_j \right)$$

pro $i = 1, 2, \dots, n$, resp. $k = n, n-1, \dots, 1$. Vzhledem k regularitě L a U je vždy $\ell_{ii} \neq 0$ a $u_{kk} \neq 0$. Pro $i = 1$, resp. $k = n$ se ve vzorcích (2.12) vynechávají příslušné sumy, protože by nedávaly smysl, tedy pokládáme $\sum_{r=s}^s \dots = 0$, při $r > s$.

2.3.2 Gaussova a Jordanova eliminacní metoda

Smyslem Gaussovy eliminacní metody je převod plné soustavy na soustavu s horní trojúhelníkovou maticí, a to postupnou eliminací neznámých, která se provádí vhodnými lineárními kombinacemi rovnic. Jordanova metoda postupuje obdobně, ale jejím cílem je získání ekvivalentní soustavy s diagonální maticí.

Mějme tedy soustavu $A \cdot \underline{x} = \underline{b}$, s regulérní maticí A typu (n,n) a nenulovým vektorom pravých stran \underline{b} . Při obvyklém značení prvků matice A a vektorů \underline{x} , \underline{b} má soustava tvar

$$(2.13) \quad \begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2, \\ \dots & \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= b_n. \end{aligned}$$

V prvním kroku Gaussovy eliminacní metody předpokládáme, že $a_{11} \neq 0$. Je-li $a_{11} = 0$, pak vzhledem k regularitě matice A je alespoň jeden z prvků prvního sloupce $a_{21}, a_{31}, \dots, a_{n1}$ nenulový. Stačí tedy příslušnou rovnici přeřadit na první místo a tím docílíme $a_{11} \neq 0$. Prvek a_{11} nazveme hlavní prvek. V dalším postupu přičítáme první, tzv. hlavní rovnici, násobenou postupně multiplikátory

$$m_i = -a_{i1}/a_{11}, \quad i = 2, 3, \dots, n$$

ke druhé až poslední rovnici. Tím dosáhneme redukci rovnic (2.13) na tvar s nulami pod hlavním prvkem

$$(2.14) \quad \begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1, \\ a_{22}^{(1)}x_2 + \dots + a_{2n}^{(1)}x_n &= b_2^{(1)}, \\ \dots \\ a_{n2}^{(1)}x_2 + \dots + a_{nn}^{(1)}x_n &= b_n^{(1)}, \end{aligned}$$

kde jsou redukované koeficienty 2. až n-té rovnice počítány podle vzorců

$$\left. \begin{aligned} a_{ij}^{(1)} &= a_{ij} + m_i \cdot a_{1j} \\ b_i^{(1)} &= b_i + m_i \cdot b_1 \end{aligned} \right\} \begin{aligned} i &= 2, 3, \dots, n \\ j &= 2, 3, \dots, n \end{aligned}$$

Ověřte si, že $a_{ii}^{(1)} = 0$ pro $i = 2, 3, \dots, n$.

Tím se podařilo eliminovat z druhé až n-té rovnice neznámou x_1 . Zcela analogicky provedeme druhý krok za předpokladu $a_{22}^{(1)} \neq 0$. Je-li $a_{22}^{(1)} = 0$, pak mezi prvky $a_{32}^{(1)}, a_{42}^{(1)}, \dots, a_{n2}^{(1)}$ vybereme nenulový (který vzhledem k regularitě existuje) a příslušnou rovnici přeřadíme na druhé místo, tím bude $a_{22}^{(1)} \neq 0$. Nyní má prvek $a_{22}^{(1)}$ funkci hlavního prvku a hlavní rovnici je rovnice druhá. Soustavu (2.14) upravíme tak, že 1. rovnici ponecháme a ke třetí až n-té rovnici přičteme hlavní rovnici (t.j. 2. rovnici z (2.14)) násobenou postupně multiplikátory

$$m_i^{(1)} = -a_{i2}^{(1)} / a_{22}^{(1)}, \quad i = 3, 4, \dots, n.$$

Soustava (2.14) bude mít po tomto kroku eliminován ze třetí až n-té rovnice neznámé x_1, x_2 .

Jestliže předcházející postup zobecníme, pak v každém k-tém kroku, $k = 1, 2, \dots, n-1$, dostáváme soustavu

$$(2.15) \quad \begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n &= b_1 \\ \dots \\ a_{kk}^{(k-1)}x_k + a_{k,k+1}^{(k-1)}x_{k+1} + \dots + a_{kn}^{(k-1)}x_n &= b_k^{(k-1)}, \\ a_{k+1,k+1}^{(k)}x_{k+1} + \dots + a_{k+1,n}^{(k)}x_n &= b_{k+1}^{(k)}, \\ \dots \\ a_{n,k+1}^{(k)}x_{k+1} + \dots + a_{nn}^{(k)}x_n &= b_n^{(k)} \end{aligned}$$

ve které se redukované prvky k-tého kroku počítají podle vzorců

$$(2.16) \quad \left. \begin{aligned} m_i^{(k-1)} &= -a_{ik}^{(k-1)} / a_{kk}^{(k-1)} \\ a_{ij}^{(k)} &= a_{ij}^{(k-1)} + m_i^{(k-1)} \cdot a_{kj}^{(k-1)} \\ b_i^{(k)} &= b_i^{(k-1)} + m_i^{(k-1)} \cdot b_k^{(k-1)} \end{aligned} \right\} \begin{aligned} i &= k+1, k+2, \dots, n \\ j &= k+1, k+2, \dots, n \end{aligned}$$

Jistě platí $a_{ik}^{(k)} = 0$ pro $i=k+1, k+2, \dots, n$. Při $k=1$, tj. na začátku, pokládáme $a_{ij}^{(0)} = a_{ij}$, $b_i^{(0)} = b_i$.

Po $n-1$ krocích bude mít soustava (2.15) trojúhelníkový tvar a její řešení je podle (2.12) dáno vzorcem

$$(2.17) \quad x_i = \frac{1}{a_{ii}^{(i-1)}} \left(b_i^{(i-1)} - \sum_{j=i+1}^n a_{ij}^{(i-1)} x_j \right), \quad i = n, n-1, \dots, 1.$$

Řešení trojúhelníkové soustavy se nazývá zpětný chod Gaussovy eliminacní metody.

Při programování metody podle vzorců (2.16) můžeme vynechat horní indexy, protože úpravy prvků matice A a vektoru b provádíme v témže paměťovém prostoru.

Příklad 2.1 Gaussovou eliminacní metodou řešme soustavu

$$(2.18) \quad \begin{aligned} x_1 + 2x_2 + x_3 &= 1, \\ -x_1 + x_2 &= -2, \\ 2x_1 - x_2 - x_3 &= 1. \end{aligned}$$

Řešení: Postupné kroky jsou přehledně sestaveny v následující tabulce 2.1.

k	i	m_i	a_{i1}	a_{i2}	a_{i3}	b_i
0	1		1	2	1	1
	2		-1	1	0	-2
	3		2	-1	-1	1
1	2	1		3	1	-1
	3	-2		-5	-3	-1
2	3	5/3			-4/3	-8/3

Tab. 2.1

Tím jsme získali trojúhelníkovou soustavu

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 1, \\ 3x_2 + x_3 &= -1, \\ -4/3x_3 &= -8/3, \end{aligned}$$

s řešením

$$x_3 = 2, \quad x_2 = (-1 - 2)/3 = -1, \quad x_1 = 1 - 2 + 2 = 1.$$

Obměnou uvedeného eliminacního postupu je tzv. Jordanova eliminacní metoda, která vytváří nulové prvky jak pod, tak také nad hlavním prvkem, a tím převádí danou soustavu na diagonální tvar. První krok Jordanovy eliminace je shodný s Gaussovou eliminací, tedy budeme vycházet ze soustavy (2.14), ve které pouze z důvodů označení položíme $a_{1j} = a_{1j}^{(1)}$, pro $j = 2, 3, \dots, n$ a $b_1 = b_1^{(1)}$. Za předpokladu $a_{22}^{(1)} \neq 0$ (není-li tomu tak, dovedeme si již poradit) přičteme ve druhém kroku k i -té rovnici,

$i = 1, 3, 4, \dots, n$, rovnici druhou, vynásobenou multiplikátorem

$$m_i^{(1)} = -a_{i2}^{(1)}/a_{22}^{(1)}.$$

Z (2.14) tak dostaneme soustavu, která má v 1. a 2. sloupci všechny nedigonální prvky nulové.

Za předpokladu nenulové hodnoty postupně vytvářených diagonálních prvků můžeme analogickými kroky pokračovat dále, takže v každém k -tém kroku, pro $k=1, 2, \dots, n$, nulujeme prvky k -tého sloupce s výjimkou prvku diagonálního. V k -tém kroku dostáváme soustavu

$$(2.19) \quad \begin{aligned} a_{11}x_1 + a_{1,k+1}^{(k)}x_{k+1} + \dots + a_{1n}^{(k)}x_n &= b_1^{(k)}, \\ \dots &\dots \\ a_{kk}^{(k-1)}x_k + a_{k,k+1}^{(k-1)}x_{k+1} + \dots + a_{kn}^{(k-1)}x_n &= b_k^{(k-1)}, \\ a_{k+1,k+1}^{(k)}x_{k+1} + \dots + a_{k+1,n}^{(k)}x_n &= b_{k+1}^{(k)}, \\ \dots &\dots \\ a_{n,k+1}^{(k)}x_{k+1} + \dots + a_{nn}^{(k)}x_n &= b_n^{(k)}, \end{aligned}$$

kde se redukované prvky počítají ze vzorců

$$(2.20) \quad \begin{aligned} m_i^{(k-1)} &= -a_{ik}^{(k-1)}/a_{kk}^{(k-1)}, \\ a_{ij}^{(k)} &= a_{ij}^{(k-1)} + m_i^{(k-1)} \cdot a_{kj}^{(k-1)}, \\ b_i^{(k)} &= b_i^{(k-1)} + m_i^{(k-1)} \cdot b_k^{(k-1)}, \end{aligned}$$

pro $i = 1, 2, \dots, k-1, k+1, \dots, n$ a $j = k+1, k+2, \dots, n$.

Při $k=1$ pokládáme $a_{ij}^{(0)} = a_{ij}$, $b_i^{(0)} = b_i$ pro $i, j = 1, 2, 3, \dots, n$.

Po n krocích bude mít soustava (2.19) diagonální tvar, tedy

$$(2.21) \quad a_{11}^{(0)}x_1 = b_1^{(n)}, \quad a_{22}^{(1)}x_2 = b_2^{(n)}, \quad \dots, \quad a_{nn}^{(n-1)}x_n = b_n^{(n)}$$

a potom

$$(2.22) \quad x_i = b_i^{(n)}/a_{ii}^{(i-1)}, \quad i = 1, 2, \dots, n.$$

Vzhledem k jednoduchosti vzorce (2.22) by se mohlo zdát, že Jordanův postup je jednodušší, avšak při bližší analýze počtu operací se ukáže, že Jordanova eliminace spotřebuje při "velkých" n asi o 50 % více operací, a tedy i strojového času. Proto se budeme orientovat většinou na Gaussův postup.

Při programování Jordanovy metody podle vzorců (2.20) jsou opět horní indexy nepodstatné, protože redukované prvky ukládáme v paměti na původní místa.

Příklad 2.2 Jordanovou eliminační metodou řešme soustavu (2.18).

Řešení: V následující tabulce jsou sestaveny výsledky jednotlivých kroků.

k	i	m_i	a_{il}	a_{i2}	a_{i3}	b
0	1		1	2	1	1
	2		-1	1	0	-2
	3		2	-1	-1	1
1	1		1	2	1	1
	2	1		3	1	-1
	3	-2		-5	-3	-1
2	1	-2/3	1		1/3	5/3
	2			3	1	-1
	3	5/3			-4/3	-8/3
3	1	1/4	1			1
	2	3/4		3		-3
	3				-4/3	-8/3

Tab. 2.2

Ze třetího kroku plyne $x_1 = 1$, $x_2 = -1$ a $x_3 = -2$.

Poznámka 2.1 Vzorce eliminačních metod jsou zapsány za předpokladu, že hlavní prvek je vždy prvkem diagonálním. Problémů, které jsou spojené s nediagonálními hlavními prvky, si všimneme až při výběru hlavních prvků v následujícím odstavci.

2.3.3 Výběr hlavních prvků

Mějme dánu soustavu (2.13). Gaussova eliminační metoda je teoreticky realizovatelná, jestliže je v každém kroku $a_{kk}^{(k-1)} \neq 0$, jak vidíme z (2.16). Je-li však v po- dílech $m_i^{(k-1)}$ absolutní hodnota jmenovatele vzhledem k čitateli malá, nastává pro- blém velkých zaokrouhlovacích chyb, jejichž hromadění může výpočet zcela znehodnotit. Proto provádíme tzv. výběr hlavních prvků, abychom omezili chyby podílů při výpočtech multiplikátorů $m_i^{(k-1)}$ v (2.16).

Nejdříve si popíšeme tzv. soupcový výběr hlavních prvků u Gaussovy eliminace. Pro $k = 1$ hledáme v 1. sloupci prvek s maximální absolutní hodnotou. Je-li

$$|a_{r_1 1}| = \max_{i=1,2,\dots,n} |a_{il}|,$$

pak první hlavní prvek je $a_{r_1 1}$ a první hlavní rovnice má index r_1 . V dalším po- stupu můžeme rovnici r_1 přemístit na první místo a pokračovat obvyklým způsobem, nebo ponechat rovnici r_1 na svém místě a neznámou x_1 eliminovat z ostatních rovnic tím, že k nim přičteme vhodné násobky hlavní rovnice. Pro $k = 2$ vybereme za druhý hlavní prvek $a_{r_2 2}^{(1)}$ takový, pro který platí

$$|a_{r_2^2}^{(1)}| = \max_{\substack{i=1,2,\dots,n \\ i \neq r_1}} |a_{i2}^{(1)}|,$$

tedy druhá hlavní rovnice má index r_2 . Postupujeme-li opět bez přemísťování rovnic, pak eliminujeme neznámou x_2 ze všech rovnic kromě první a druhé hlavní rovnice tím, že k nim přičteme vhodné násobky druhé hlavní rovnice.

Obecně tedy v k-tém kroku, pro $k = 1, 2, \dots, n-1$, hledáme k-tý hlavní prvek $a_{r_k^k}^{(k-1)}$, pro který platí

$$|a_{r_k^k}^{(k-1)}| = \max_{\substack{i=1,2,\dots,n \\ i \neq r_1, r_2, \dots, r_{k-1}}} |a_{ik}^{(k-1)}|.$$

Je-li takových prvků více, vezmeme např. ten, který má nejmenší řádkový index. Redučované prvky k-tého kroku dostaneme podle vzorců

$$\begin{aligned} a_{ij}^{(k)} &= a_{ij}^{(k-1)} - (a_{ik}^{(k-1)} / a_{r_k^k}^{(k-1)}) \cdot a_{r_k^k j}^{(k-1)} \\ b_i^{(k)} &= b_i^{(k-1)} - (a_{ik}^{(k-1)} / a_{r_k^k}^{(k-1)}) \cdot b_{r_k^k}^{(k-1)} \end{aligned}$$

pro $i = 1, 2, \dots, n$, ale $i \neq r_1, r_2, \dots, r_k$, $j = k+1, k+2, \dots, n$.

Po $n-1$ krocích je přímý chod Gaussovy eliminace ukončen, i když soustava nemá trojúhelníkový tvar, ale úpravou bychom jej mohli snadno získat. Zpětný chod bez přerovnávání soustavy lze zapsat takto:

$$(2.23) \quad x_i = \frac{1}{a_{r_i i}^{(i-1)}} \left(b_{r_i}^{(i-1)} - \sum_{j=i+1}^n a_{r_i j}^{(i-1)} x_j \right), \quad i = n, n-1, \dots, 1.$$

Při $i = n$ se suma ve vzorci (2.23) vypouští.

Obdobně provádime řádkový výběr hlavního prvku. V k-tém kroku vybereme za k-tý hlavní prvek takový prvek k-tého řádku $a_{ks_k}^{(k-1)}$, pro který platí

$$|a_{ks_k}^{(k-1)}| = \max_{\substack{j=1,2,\dots,n \\ j \neq s_1, s_2, \dots, s_{k-1}}} |a_{kj}^{(k-1)}|,$$

kde s_1, s_2, \dots, s_{k-1} jsou sloupcové indexy předcházejících hlavních prvků. Přičtení vhodného násobku k-té hlavní rovnice postupně ke zbývajícím rovnicím obstaráme nulové prvky v s_k -tém sloupci. Po $n-1$ krocích dostaneme "skrytě" trojúhelníkovou soustavu, která je opět snadno řešitelná za použití sloupcových indexů s_1, s_2, \dots, s_n , kde s_n je poslední zbývající sloupcový index.

Poslední a nejnáročnější je úplný výběr hlavních prvků, kdy se za hlavní prvek k-tého kroku vybere takový prvek $a_{rs}^{(k-1)}$ s minimálními indexy, pro který platí

$$|a_{rs}^{(k-1)}| = \max_{i,j} |a_{ij}^{(k-1)}|,$$

kde i, j probíhá indexy, které nebyly v předcházejících krocích rovny indexům hlavních prvků. Úplný výběr hlavních prvků je časově značně náročný, užívá se jenom zřídka, a proto se jím nebudeme podrobně zabývat.

Příklad 2.3 Řešme soustavu (2.18) Gaussovou eliminací se sloupcovým výběrem hlavních prvků.

Řešení: Výpočetní kroky jsou opět sestaveny v tabulce 2.3.

k	i	m _i	a _{il}	a _{i2}	a _{i3}	b _i
0	1		1	2	1	1
	2		-1	1	0	-2
	3		2	-1	-1	1
1	1	-1/2		5/2	3/2	1/2
	2	1/2		1/2	-1/2	-3/2
2	2	-1/5			-8/10	-16/10

Tab. 2.3

Dostali jsme tedy soustavu ve tvaru

$$\left. \begin{array}{l} \frac{5}{2}x_2 + \frac{3}{2}x_3 = \frac{1}{2} \\ -8/10x_3 = -\frac{16}{10} \\ 2x_1 - x_2 - x_3 = 1 \end{array} \right\} \Rightarrow \begin{array}{l} x_3 = 2 \\ x_2 = -1 \\ x_1 = 1 \end{array}$$

2.3.4 Metoda LU-rozkladu

Pod pojmem LU-rozklad (trojúhelníkový rozklad) regulární matice A budeme rozumět nalezení takové dolní L a horní U trojúhelníkové matice, aby platilo

$$(2.24) \quad A = L \cdot U .$$

Rozklad (2.24) je však víceznačný, a proto můžeme ještě volit hodnoty některých prvků matice L nebo U. Často volíme v matici L jednotkovou diagonálu. Existenci a jednoznačnost takového rozkladu zajišťuje následující věta.

Věta 2.1 Nechť A je čtvercová matice, jejíž hlavní subdeterminanty jsou nenulové. Pak existuje právě jedna dolní trojúhelníková matice L s jednotkovou diagonálou a právě jedna horní trojúhelníková matice U tak, že platí (2.24).

Důkaz je proveden např. v [18].

Máme tedy nalézt takové prvky ℓ_{ij} , u_{ij} , pro které platí:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} 1 & & & & \\ \ell_{21} & 1 & & & \\ \ell_{31} & \ell_{32} & 1 & & \\ \dots & \dots & \dots & \dots & \\ \ell_{n1} & \ell_{n2} & \ell_{n3} & \cdots & 1 \end{pmatrix} \cdot \begin{pmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ u_{22} & u_{23} & \cdots & u_{2n} \\ u_{33} & \cdots & u_{3n} \\ \dots & \dots & \dots & \dots & \\ u_{nn} & & & & \end{pmatrix}$$

Z definice násobení matic plyne, že pro:

$$\begin{aligned}
 a) \quad i \leq j \quad a_{ij} &= \sum_{k=1}^{i-1} \ell_{ik} u_{kj} + u_{ij}, \\
 \text{čili} \quad u_{ij} &= a_{ij} - \sum_{k=1}^{i-1} \ell_{ik} u_{kj}, \\
 b) \quad i > j \quad a_{ij} &= \sum_{k=1}^j \ell_{ik} u_{kj} = \sum_{k=1}^{j-1} \ell_{ik} u_{kj} + \ell_{ij} u_{jj}, \\
 \text{čili} \quad \ell_{ij} &= \left(a_{ij} - \sum_{k=1}^{j-1} \ell_{ik} u_{kj} \right) / u_{jj}.
 \end{aligned}$$

Výpočet můžeme vést např. po sloupcích, takže pro $j = 1, 2, 3, \dots, n$ vypočítáváme postupně

$$\begin{aligned}
 u_{ij} &= a_{ij} - \sum_{k=1}^{i-1} \ell_{ik} u_{kj}, \quad \text{pro } i = 1, 2, \dots, j, \\
 (2.25) \quad \ell_{ij} &= \left(a_{ij} - \sum_{k=1}^{j-1} \ell_{ik} u_{kj} \right) / u_{jj}, \quad \text{pro } i = j+1, j+2, \dots, n.
 \end{aligned}$$

(Je-li $r > s$, pak klademe $\sum_r^s \dots = 0.$)

Užijeme-li rozklad (2.24) při řešení soustavy $A \cdot \underline{x} = \underline{b}$, dostaneme soustavu $L \cdot U \cdot \underline{x} = \underline{b}$ a ta je ekvivalentní řešení dvou soustav

$$L \cdot \underline{y} = \underline{b} \quad \text{a} \quad U \cdot \underline{x} = \underline{y},$$

které řešíme podle vzorců (2.12). Tato metoda se nazývá Croutova, resp. Doolitlova.

Výhody LU-rozkladu jsou zřejmě teprve tehdy, když řešíme více soustav s odlišnými pravými stranami, protože metody rozkladu nemění pravou stranu soustavy.

Ještě si všimneme jedné modifikace LU-rozkladu. Je-li A symetrická pozitivně definitní matici, pak existuje právě jedna dolní trojúhelníková matici L s kladnými diagonálními prvky, pro kterou platí

$$(2.26) \quad A = L L^T.$$

Toto tvrzení je dokázáno např. ve [12]. Zmíněná modifikace se nazývá Choleského metoda.

Z definice součinu $L L^T$ odvodíme užitím rovnice (2.26) vzorce pro výpočet prvků matice L . Pro $j = 1, 2, \dots, n$ dostaneme

$$\begin{aligned}
 (2.27) \quad \ell_{jj} &= a_{jj} - \sum_{k=1}^{j-1} \ell_{jk}^2, \\
 \ell_{ij} &= \left(a_{ij} - \sum_{k=1}^{j-1} \ell_{ik} \ell_{kj} \right) / \ell_{jj}, \quad \text{pro } i = j+1, j+2, \dots, n.
 \end{aligned}$$

Poznámka 2.2 Ve [23] je ukázáno, že LU-rozklad není v podstatě nic jiného, než jiná formulace Gaussovy eliminaci metody bez výběru hlavních prvků. V této publikaci je uveden také algoritmus LU-rozkladu, který užívá výběr hlavních prvků a omezuje také zaokrouhlovací chyby.

Poznámka 2.3 Metoda Choleského je velmi užitečná při řešení velkých řídkých soustav se symetrickou pozitivně definitní maticí - viz [12].

Příklad 2.4 Řešme soustavu (2.18) pomocí LU-rozkladu.

Řešení: Ověřte si, že podle vzorců (2.25) dostaneme

$$\begin{pmatrix} 1 & 2 & 1 \\ -1 & 1 & 0 \\ 2 & -1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & & \\ -1 & 1 & \\ 2 & -5/3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 1 \\ & 3 & 1 \\ & & -4/3 \end{pmatrix} = L \cdot U .$$

Pak ze soustavy $L \cdot \underline{y} = \underline{b}$ dostaneme $y_1 = 1$, $y_2 = -1$, $y_3 = -8/3$ a ze soustavy $U \cdot \underline{x} = \underline{y}$ vypočítáme již hledané řešení $x_3 = 2$, $x_2 = -1$ a $x_1 = 1$.

2.4 I terační metody řešení soustav lineárních algebraických rovnic

Iterační metody jsou ty, jimiž hledáme řešení soustavy postupným přibližováním. Přesné řešení je pak limita posloupnosti postupných approximací. Velká skupina iteračních procesů pro soustavu $A \cdot \underline{x} = \underline{b}$ hledá postupné approximace $\underline{x}^{(k)}$ podle rekurentního vzorce

$$(2.28) \quad \underline{x}^{(k)} = \underline{x}^{(k-1)} + H_k(\underline{b} - A \cdot \underline{x}^{(k-1)}),$$

kde H_k jsou jisté matice a $\underline{x}^{(0)}$ je počáteční approximace řešení. Vzorec (2.28) je splněn pro libovolné H_k řešením rovnice $A \cdot \underline{x} = \underline{b}$. Volbou matic H_k docházíme k různým iteračním procesům. Je-li H_k konstantní nebo se cyklicky opakuje po jistém počtu kroků, dostáváme tzv. stacionární iterační metody. V opačném případě jde o metody nestacionární, v nichž iterační předpis na pravé straně (2.28) závisí na k . Nestacionární metody přesahují rámcem tohoto skripta a proto se jimi nebudeme zabývat. Čtenář se s nimi může seznámit např. v [23], [21], [11].

Při iterační metodě sestrojíme podle (2.28) z počáteční approximace $\underline{x}^{(0)}$ posloupnost $\underline{x}^{(1)}, \underline{x}^{(2)}, \dots$. Výpočet ukončíme buď po stanoveném počtu iterací nebo v okamžiku, kdy se po sobě jdoucí iterace "dostatečně málo" odlišují, tedy když platí

$$\| \underline{x}^{(k)} - \underline{x}^{(k-1)} \| < \varepsilon .$$

Pak říkáme, že approximace $\underline{x}^{(k)}$ je určena s přesností ε .

2.4.1 Prostá iterační metoda - Jacobiova metoda

V prosté iterační metodě upravíme danou regulární soustavu $A \cdot \underline{x} = \underline{b}$ na tvar $HA \cdot \underline{x} = H \cdot \underline{b}$, kde H je konstantní regulární matici "blízká" matici A^{-1} . Přičteme-li k oběma stranám vektor \underline{x} , pak dostaneme po úpravě rovnici

$$(2.29) \quad \underline{x} = \underline{x} + H(\underline{b} - A \cdot \underline{x}) \quad \text{čili} \quad \underline{x} = B \cdot \underline{x} + \underline{c},$$

kde $B = E - HA$ a $\underline{c} = H \cdot \underline{b}$,

která je základem pro iterační vzorec

$$(2.30) \quad \underline{x}^{(k)} = B \cdot \underline{x}^{(k-1)} + \underline{c} .$$

U soustav s nenulovou (raději dominantní) diagonálou je vhodné volit

$$H = (h_{ij}), \quad \text{kde} \quad h_{ij} = \begin{cases} 0 & \text{pro } i \neq j \\ \frac{1}{a_{ii}} & \text{pro } i = j \end{cases} ,$$

potom

$$(2.31) \quad B = - \begin{pmatrix} 0 & a_{12}/a_{11} & a_{13}/a_{11} & \cdots & a_{1n}/a_{11} \\ a_{21}/a_{22} & 0 & a_{23}/a_{22} & \cdots & a_{2n}/a_{22} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1}/a_{nn} & a_{n2}/a_{nn} & a_{n3}/a_{nn} & \cdots & 0 \end{pmatrix}, \quad c = \begin{pmatrix} b_1/a_{11} \\ b_2/a_{22} \\ \cdots \\ b_n/a_{nn} \end{pmatrix}.$$

Odpovídající iterační vzorec (2.30) můžeme přepsat do soustavy

$$(2.32) \quad x_i^{(k)} = \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k-1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k-1)} \right) / a_{ii},$$

pro $i = 1, 2, \dots, n$.

Iterační metoda řízená rovnicemi (2.32) se nazývá Jacobiova iterační metoda.

Poznámka 2.4 Všimněte si, že rovnice (2.32) můžeme získat ze soustavy $A \cdot \underline{x} = \underline{b}$ tím, že v každé i -té rovnici osamostatníme na levé straně neznámou x_i .

Při realizaci prosté iterační metody vzniká otázka volby počáteční approximace $\underline{x}^{(0)}$, otázka podmínek konvergence posloupnosti $\{\underline{x}^{(k)}\}$ k řešení rovnice $A \cdot \underline{x} = \underline{b}$ a konečně i otázka rychlosti konvergence.

Věta 2.2 Mějme dánu vzorcem (2.29), resp. (2.30) prostou iterační metodu pro regulární soustavu $A \cdot \underline{x} = \underline{b}$. Jestliže platí

$$(2.33) \quad \|B\| = \|E - HA\| < 1,$$

potom posloupnost approximací $\{\underline{x}^{(k)}\}$, počítaná podle vzorce (2.30), konverguje pro libovolnou počáteční approximaci $\underline{x}^{(0)}$ k řešení soustavy $A \cdot \underline{x} = \underline{b}$. Dále platí

$$(2.34) \quad \|\underline{x}^{(k)} - \underline{x}^*\| \leq \frac{\|B\|}{1 - \|B\|} \|\underline{x}^{(k)} - \underline{x}^{(k-1)}\|,$$

kde \underline{x}^* je řešení soustavy $A \cdot \underline{x} = \underline{b}$. Normy matic a vektorů jsou souhlasné.

Z této věty vidíme, že $\|B\| < 1$ je postačující podmínkou konvergence iteračního procesu, který vychází z libovolného vektoru $\underline{x}^{(0)}$. Rychlosť konvergence roste pro $\|B\| \ll 1$ a je malá, když $\|B\|$ je blízká 1.

Jestliže iterační proces ukončíme při splnění podmínky

$$\|\underline{x}^{(k)} - \underline{x}^{(k-1)}\| < \varepsilon,$$

pak podle (2.34) platí

$$\|\underline{x}^{(k)} - \underline{x}^*\| < \frac{\|B\|}{1 - \|B\|} \varepsilon.$$

Pro Jacobiovu iterační metodu platí ještě jedna snadno ověřitelná postačující podmínka konvergence daná následující větou.

Věta 2.3 Jestliže matice soustavy $A \cdot \underline{x} = \underline{b}$ je diagonálně dominantní, pak Jacobiova iterační metoda (2.32) konverguje nezávisle na volbě počáteční approximace k řešení zadáné soustavy, a to v řádkové nebo ve sloupcové normě.

Důkaz: Je-li A diagonálně dominantní, pak pro libovolné i nebo j platí

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|, \text{ nebo } |a_{jj}| > \sum_{\substack{i=1 \\ i \neq j}}^n |a_{ij}|.$$

Potom např. z první nerovnosti platí

$$1 > \sum_{\substack{j=1 \\ j \neq i}}^n \left| \frac{a_{ij}}{a_{ii}} \right| \quad \text{pro } i = 1, 2, \dots, n,$$

tedy

$$1 > \max_i \sum_{\substack{j=1 \\ j \neq i}}^n \left| \frac{a_{ij}}{a_{ii}} \right|.$$

Vzhledem k nulovým diagonálním prvkům matice B je na pravé straně poslední nerovnosti řádková norma matice B, tedy je splněn předpoklad věty 2.2 a Jacobiův iterativní proces konverguje.

2.4.2 Gaussova - Seidelova iterativní metoda (G - S iterativní metoda)

G - S iterativní metoda vychází z jednoduchého principu Jacobiovy metody definované rovnicemi (2.32). V Jacobiově metodě se všechny složky nové (k-té) aproximace počítají jenom ze složek předcházející k - 1 aproximace, přestože při výpočtu složky $x_i^{(k)}$ máme již k dispozici opravené složky $x_1^{(k)}, \dots, x_{k-1}^{(k)}$. V Seidelově metodě se postupuje v zásadě podle vzorců (2.32), ale při výpočtu $x_i^{(k)}$ se užívá již opravených hodnot $x_1^{(k)}, \dots, x_{i-1}^{(k)}$, tedy se postupuje podle vzorců

$$(2.35) \quad x_i^{(k)} = \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^n a_{ij} x_j^{(k-1)} \right) / a_{ii}$$

pro $i = 1, 2, \dots, n$.

Tuto metodu můžeme formulovat také v maticovém vyjádření podle (2.28) tak, že zavedeme matice H_1, H_2, \dots, H_n , kde H_s ($s = 1, 2, \dots, n$) má všechny prvky nulové kromě $h_{ss} = 1/a_{ss}$. Dosadíme-li postupně matice H_s do (2.28), dostaneme rekurentní vztahy

$$(2.36) \quad \underline{x}^{(s)} = \underline{x}^{(s-1)} + H_s (b - A \cdot \underline{x}^{(s-1)}), \quad s = 1, 2, \dots, n,$$

které lze přepsat do iterativní soustavy

$$(2.37) \quad \begin{aligned} x_i^{(s)} &= x_i^{(s-1)}, \quad i = 1, 2, \dots, s-1, s+1, \dots, n \\ x_s^{(s)} &= \left(b_s - \sum_{\substack{j=1 \\ j \neq s}}^n a_{sj} x_j^{(s-1)} \right) / a_{ss} \end{aligned}$$

pro $i = 1, 2, \dots, n$.

Z předcházejícího je zřejmé, že když z počátečního vektoru $\underline{x}^{(0)}$ sestrojíme postupně vektory

$$\underline{x}^{(1)} = \varphi_1(\underline{x}^{(0)}), \underline{x}^{(2)} = \varphi_2(\underline{x}^{(1)}), \dots, \underline{x}^{(n)} = \varphi_n(\underline{x}^{(n-1)}),$$

kde $\varphi_1, \varphi_2, \dots, \varphi_n$ je obecné označení pro iterační vzorce (2.36), pak vektor $\underline{x}^{(n)}$ je shodný s vektorem $\underline{x}^{(1)}$ vypočteným podle (2.35).

Následující věta formuluje postačující podmínky konvergence G-S iterační metody přímo vzhledem k matici A.

Věta 2.4 Mějme dánu soustavu $A \cdot \underline{x} = \underline{b}$. Je-li matice A buď diagonálně dominantní nebo symetrická a pozitivně definitní, pak G - S iterační metoda definované vzorcem (2.35), resp. (2.37) konverguje pro libovolný počáteční vektor $\underline{x}^{(0)}$ k řešení soustavy $A \cdot \underline{x} = \underline{b}$.

Poznámka 2.5 Jacobiova metoda se v praxi užívá méně, protože G - S metoda konverguje skoro vždy, když konverguje Jacobiova metoda a většinou konverguje rychleji. G - S metoda může konvergovat i tehdy, když Jacobiova metoda nekonverguje.

Příklad 2.5 Řešme Jacobiovou a Gaussovou - Seidelovou iterační metodou soustavu

$$(2.38) \quad \begin{aligned} 5,63 x_1 - 1,53 x_2 + 1,06 x_3 &= 9,28 \\ - 1,53 x_1 + 4,36 x_2 + 2,01 x_3 &= - 1,87 \\ 1,06 x_1 + 2,01 x_2 - 5,03 x_3 &= - 11,01 \end{aligned}$$

Řešení: Daná soustava je diagonálně dominantní, proto obě metody konvergují. Převeďme tedy (2.38) na iterační tvar (při výpočtu budeme zapisovat čtyři cifry)

$$\begin{aligned} x_1 &= 1,648 + 0,2718 x_2 - 0,1883 x_3 \\ x_2 &= - 0,4289 + 0,3509 x_1 - 0,4610 x_3 \\ x_3 &= - 1,390 - 0,2107 x_1 + 0,3996 x_2 \end{aligned}$$

a vypočítejme pro každou metodu 20 iterací, které sestavíme do tabulky 2.4.

k	Jacobiova metoda			Seidelova metoda		
	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0	0,0	0,0	0,0	0,0	0,0	0,0
5	0,997 180	- 0,991 322	1,970 240	1,000 507	- 1,000 970	1,999 819
10	0,999 534	- 0,999 933	1,999 446	0,999 572	- 1,000 131	1,999 958
15	0,999 572	- 1,000 127	1,999 949	0,999 572	- 1,000 131	1,999 958
20	0,999 572	- 1,000 131	1,999 958	0,999 572	- 1,000 131	1,999 958

Tab. 2.4

2.4.3 Relaxační metoda

Seidelova metoda, tak jak jsme ji formulovali v maticovém tvaru pomocí matic H_1, \dots, H_n , patří do třídy tzv. relaxačních metod, které v jednom kroku modifikují jednu nebo více vybraných složek řešení tak, aby jedna nebo více rovnic byly v tomto kroku splněny přesně. Metody tohoto typu jsou vhodné pro "ruční" počítání, protože umožňují počtaři volit operativně postup tak, aby v k-tém kroku byly splněny přesně ty rovnice, které splňuje předcházející approximace s největší chybou.

Relaxační metoda probíhá takto:

a) Zvolme počáteční vektor $\underline{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$.

b) Vypočtěme tzv. reziduální vektor

$$\underline{r}^{(0)} = \underline{b} - A \cdot \underline{x}^{(0)}, \text{ kde } r_i^{(0)} = b_i - \sum_{j=1}^n a_{ij} x_j^{(0)}, \text{ pro } i = 1, 2, \dots, n.$$

c) Nalezněme tu složku reziduálního vektoru, která má maximální absolutní hodnotu.

Nechť tedy

$$|r_t^{(0)}| = \max_i |r_i^{(0)}|.$$

d) Vektor $\underline{x}^{(1)}$ bude mít oproti $\underline{x}^{(0)}$ změněnou pouze jednu složku, kterou vypočítáme tak, aby vektor

$$\underline{x}^{(1)} = \underline{b} - A \cdot \underline{x}^{(0)}$$

měl složku $r_t^{(1)}$ nulovou. Označíme-li $\underline{x}^{(1)} = (x_1^{(1)}, \dots, x_n^{(1)})$, pak platí $x_i^{(1)} = x_i^{(0)}$ pro $i \neq u$ a $x_u^{(1)} = x_u^{(0)} + \alpha$, kde zatím neznáme index u a přírůstek α . Pro α dostaneme podle předešlézejícího rovnici

$$0 = b_t - \sum_{j=1}^n a_{tj} x_j^{(1)} = b_t - \sum_{j=1}^n a_{tj} x_j^{(0)} - a_{tu} \alpha,$$

tedy

$$(2.39) \quad \alpha = r_t^{(0)} / a_{tu}.$$

S ohledem na přesnost podílu (2.39) najdeme index u tak, aby platilo

$$|a_{tu}| = \max_j |a_{tj}|.$$

e) Kroky b), c), d) opakujeme tak dlouho, pokud reziduální vektor nemá dostatečně malou normu.

Popsaný algoritmus je jenom základním principem relaxačních metod, které jsou velice mnohotvárné. Tento algoritmus není pro počítač příliš vhodný, protože neustále hledání maximálních hodnot je při větší soustavě příliš zdlouhavé (tedy i nákladné na počet operací i strojový čas).

2.5 Řešení maticové rovnice $AX = B$. Invertování matice

Mějme zadánu úlohu (2.6), která je ekvivalentní řešení systému soustav rovnic (2.7). Protože v systému (2.7) se jednotlivé soustavy liší jenom pravými stranami, pak se pro řešení nabízí eliminacní metody a metoda LU-rozkladu. Problémem řešitelnosti úlohy (2.6) se nebudeme zabývat, protože se dá převést na problém řešitelnosti soustav, který jsme diskutovali dříve.

2.5.1 Řešení rovnice $AX = B$ eliminací

Rovnice $AX = B$ lze zapsát ve tvaru

$$(2.40) \quad \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{nl} & \dots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_{11} & \dots & x_{1n} \\ \vdots & & \vdots \\ x_{nl} & \dots & x_{nn} \end{pmatrix} = \begin{pmatrix} b_{11} & \dots & b_{nl} \\ \vdots & & \vdots \\ b_{nl} & \dots & b_{nn} \end{pmatrix},$$

v němž je skryto soustavu (2.7). Při užití Gaussovy eliminacní metody vycházíme ze vzorců (2.16), kdy v každém kroku redukujeme pravé strany odpovídající rovnice ve všech soustavách. Z důvodů snadného označování budeme předpokládat $b_{il} = a_{i,n+1}$, $b_{i2} = a_{i,n+2}$, ..., $b_{in} = a_{i,2n}$. Tedy podle (2.16) počítáme pro $k = 1, 2, \dots, n-1$ redukce soustavu (2.40) podle vzorce

$$(2.41) \quad a_{ij}^{(k)} = a_{ij}^{(k-1)} - \left(a_{ik}^{(k-1)} / a_{kk}^{(k-1)} \right) \cdot a_{kj}^{(k-1)},$$

pro $i = k+1, k+2, \dots, n$, $j = k+1, k+2, \dots, 2n$, za předpokladu $a_{kk}^{(k-1)} \neq 0$. Míčky jsme předpokládali $a_{ij}^{(0)} \neq a_{ij}$. Po ukončení postupu podle (2.41) mají soustavy (2.40) příslušný trojúhelníkový tvar a jejich řešení je triviální, viz odstavec 2.3.1.

Analogickým postupem lze systém (2.40) řešit Jordanovou eliminacní metodou.

Příklad 2.6 Řešme rovnici $AX = B$, kde

$$(2.42) \quad A = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 1 & 0 \\ 2 & -1 & -1 \end{pmatrix} \quad a \quad B = \begin{pmatrix} 1 & 3 & 5 \\ -2 & 1 & 3 \\ 1 & -2 & -6 \end{pmatrix}$$

Řešení: Pro výpočet užijeme Gaussovy eliminace bez výběru hlavních prvků a výsledky sestavíme do tabulky 2.5.

k	i	m_i	a_{il}	a_{i2}	a_{i3}	b_{il}	b_{i2}	b_{i3}
0	1		1	2	1	1	3	5
	2		-1	1	0	-2	1	3
	3		2	-1	-1	1	-2	-6
1	2	1		3	1	-1	4	8
	3	-2		-5	-3	-1	-8	-16
2	3	5/3			-4/3	-8/3	-4/3	-8/3

Tab. 2.5

Dostali jsme tedy soustavy

$$\begin{pmatrix} 1 & 2 & 1 \\ 3 & 1 & 0 \\ -4/3 & & \end{pmatrix} \cdot \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} = \begin{pmatrix} 1 & 3 & 5 \\ -1 & 4 & 8 \\ -8/3 & -4/3 & -8/3 \end{pmatrix},$$

z nichž $x_{11} = 1$, $x_{21} = -1$, $x_{31} = 2$, $x_{12} = 0$, $x_{22} = 1$, $x_{32} = 1$, $x_{13} = -1$, $x_{23} = 2$, $x_{33} = 2$.

Poznámka 2.6 Při řešení rovnic $AX = B$ můžeme samozřejmě užívat výběr hlavních prvků ve sloupcích, řádcích, případně úplný.

2.5.2 Invertování maticy eliminací

Hledat inverzní matici k regulérní čtvercové matici A znamená podle (2.8) řešit maticovou rovnici $AX = E$, tedy řešit systém soustav

$$(2.43) \quad \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{nl} & \dots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_{11} & \dots & x_{ln} \\ \vdots & \ddots & \vdots \\ x_{nl} & \dots & x_{nn} \end{pmatrix} = \begin{pmatrix} 1 & & 0 \\ \ddots & \ddots & \ddots \\ 0 & & 1 \end{pmatrix}.$$

K řešení soustav (2.43) užijeme Jordanovy eliminační metody zcela analogicky jako v odstavci 2.5.1. Označme pro $i = 1, 2, \dots, n$ $a_{i,n+1} = e_{i1}$, $a_{i,n+2} = e_{i2}$, $\dots, a_{i,2n} = e_{in}$, kde e_{ij} jsou prvky jednotkové matice E. Při užití Jordanovy eliminace redukujeme postupně pro $k = 1, 2, \dots, n$ prvky a_{ij} podle vzorců

$$(2.44) \quad \begin{aligned} a_{ij}^{(k)} &= a_{ij}^{(k-1)} - (a_{ik}^{(k-1)} / a_{kk}^{(k-1)}) \cdot a_{kj}^{(k-1)}, \\ a_{kj}^{(k)} &= a_{kj}^{(k-1)}, \end{aligned}$$

kde $i = 1, 2, \dots, k-1, k+1, \dots, n$, $j = k+1, k+2, \dots, 2n$, za předpokladu $a_{kk}^{(k-1)} \neq 0$. Tím převedeme soustavy (2.43) na diagonální tvar, který je triviálně řešitelný. Ověřme tuto metodu na následujícím příkladě.

Příklad 2.7 Invertujme matici A z (2.42).

Řešení: Inverzní matici budeme hledat Jordanovou metodou bez výběru hlavních prvků.

k	i	A			E		
0	1	1	2	1	1	0	0
	2	-1	1	0	0	1	0
	3	2	-1	-1	0	0	1
1	1	1	2	1	1	0	0
	2	0	3	1	1	1	0
	3	0	-5	-3	-2	0	1
2	1	1	0	1/3	1/3	-2/3	0
	2	0	3	1	1	1	0
	3	0	0	-4/3	-1/3	5/3	1
3	1	1	0	0	1/4	-1/4	1/4
	2	0	3	0	3/4	9/4	3/4
	3	0	0	-4/3	-1/3	5/3	1
4	1	1	0	0	1/4	-1/4	1/4
	2	0	1	0	1/4	3/4	1/4
	3	0	0	1	1/4	-5/4	-3/4

Tab. 2.6

V silně orámované dolní části tabulky je pak matici A^{-1} .

2.5.3 Inverze LU-rozkladem

Předpokládejme, že jsme provedli rozklad matice A na součin LU , viz odstavec 2.3.4. Z maticového počtu víme, že když $A = L \cdot U$, potom $A^{-1} = U^{-1} \cdot L^{-1}$. Stačí tedy invertovat matice U , L např. řešením maticových rovnic $LX = E$, $UY = E$, potom $X = L^{-1}$ a $Y = U^{-1}$. O maticích X a Y ještě víme, že jsou dolní a horní trojúhelníkové. Jedná se tedy o řešení rovnic

$$\begin{pmatrix} 1 & & & \\ -\ell_{21} & 1 & & 0 \\ \dots & \dots & \dots & \dots \\ -\ell_{n1} & -\ell_{n2} & \dots & 1 \end{pmatrix} \begin{pmatrix} x_{11} & & & \\ x_{21} & x_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nn} \end{pmatrix} = E,$$

$$\begin{pmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ u_{22} & \dots & u_{2n} \\ \dots & \dots & \dots \\ u_{nn} \end{pmatrix} \begin{pmatrix} y_{11} & y_{12} & \dots & y_{1n} \\ y_{22} & \dots & y_{2n} \\ \dots & \dots & \dots \\ y_{nn} \end{pmatrix} = E.$$

Z definice součinu matic platí:

$$a) \quad \text{pro } i > j \quad \sum_{k=j}^{i-1} \ell_{ik} x_{kj} + x_{ij} = 0 \quad \text{a} \quad x_{jj} = 1,$$

$$b) \quad \text{pro } i < j \quad \sum_{k=i}^j u_{ik} y_{kj} = 0 \quad \text{a} \quad u_{jj} \cdot y_{jj} = 1.$$

Při výpočtu prvků x_{ij} a y_{ij} můžeme postupovat po sloupcích, tedy pro $j = 1, 2, \dots, n$ je

$$(2.45) \quad \begin{aligned} x_{jj} &= 1 & \text{a} \quad x_{ij} &= - \sum_{k=j}^{i-1} \ell_{ik} x_{kj} \quad \text{pro } i = j+1, j+2, \dots, n, \\ y_{jj} &= 1/u_{jj} & \text{a} \quad y_{ij} &= - \left(\sum_{k=i+1}^j u_{ik} y_{kj} \right) / u_{ii} \quad \text{pro } i = 1, 2, \dots, j-1. \end{aligned}$$

Příklad 2.8 Invertujme symetrickou matici ze soustavy (2.38) metodou LU-rozkladu.

Řešení: Podle vzorců (2.25) dostaneme (při výpočtu zapisujeme čtyři cifry):

$$L = \begin{pmatrix} 1.0 & & \\ -0.2718 & 1.0 & \\ 0.1883 & 0.5827 & 1.0 \end{pmatrix}, \quad U = \begin{pmatrix} 5.630 & -1.530 & 1.060 \\ & 3.944 & 2.298 \\ & & -6.569 \end{pmatrix}.$$

Potom podle (2.45) je

$$L^{-1} = \begin{pmatrix} 1.0 & & \\ 0.2718 & 1.0 & \\ -0.3467 & -0.5827 & 1.0 \end{pmatrix}, \quad U^{-1} = \begin{pmatrix} 0.1776 & 0.06889 & 0.05276 \\ & 0.2535 & 0.08868 \\ & & -0.1522 \end{pmatrix}$$

a konečně

$$A^{-1} = U^{-1} \cdot L^{-1} = \begin{pmatrix} 0.1780 & 0.03815 & 0.05276 \\ 0.03815 & 0.2018 & 0.08868 \\ 0.05276 & 0.8868 & -0.1522 \end{pmatrix}.$$

Ověřte si správnost inverze součinem $A \cdot A^{-1}$.

2.5.4 Inverze dělením na bloky - metoda vroubení

Rozdělme regulární matici A na bloky B, C, D, F

$$A = \left(\begin{array}{c|c} B & C \\ \hline D & F \end{array} \right) \left\{ \begin{array}{l} r \\ s \end{array} \right\},$$

kde B , resp. F je čtvercová matici řádu r , resp. s a C , resp. D je obdélníková matici typu (r, s) , resp. typu (s, r) . Stejné dělení provedeme u hledané inverzní matice, a to na bloky K, L, M, N . Podle definice inverzní matice platí

$$(2.46) \quad \left(\begin{array}{c|c} B & C \\ \hline D & F \end{array} \right) \cdot \left(\begin{array}{c|c} K & L \\ \hline M & N \end{array} \right) = \left(\begin{array}{c|c} E_{rr} & 0 \\ \hline 0 & E_{ss} \end{array} \right),$$

z čehož podle (2.2) dostaneme rovnice

$$\begin{aligned} BK + CM &= E_{rr} & BL + CN &= 0 \\ DK + FM &= 0 & DL + FN &= E_{ss} \end{aligned}$$

Z první dvojice rovnic vypočítáme matice K, M a ze druhé L, N .

$$\begin{aligned} M &= (DB^{-1}C - F)^{-1} \cdot DB^{-1}, & N &= (F - DB^{-1}C)^{-1}, \\ K &= B^{-1} - B^{-1} \cdot CM, & L &= -B^{-1}CN. \end{aligned}$$

Matice K, L, M, N budeme počítat v tomto pořadí:

$$(2.47) \quad N = (F - DB^{-1}C)^{-1}, \quad L = -B^{-1}CN, \quad M = -NDB^{-1}, \quad K = B^{-1} - B^{-1}CM.$$

Ve vzorcích se kromě maticových operací invertují pouze dvě matice a to B a $(F - DB^{-1}C)$, které jsou typu (r, r) a (s, s) .

Invertování matice dělené na bloky je podstatou tzv. metody vroubení, pro kterou označíme A_1, A_2, \dots, A_n hlavní submatice matice A . Invertování submatice A_1 je triviální, protože $A_1 = (a_{11})$ a $A_1^{-1} = (1/a_{11})$. V metodě vroubení postupně invertujeme matice A_k ($k = 2, 3, \dots, n$) ze znalosti matic A_{k-1}^{-1} . Rozdělme A_k a A_k^{-1} na bloky

$$A_k = \left(\begin{array}{c|c} A_{k-1} & \begin{matrix} a_{1k} \\ \vdots \\ a_{kk} \end{matrix} \\ \hline a_{k1} \dots a_{kk-1} & a_{kk} \end{array} \right), \quad A_k^{-1} = \left(\begin{array}{c|c} K & L \\ \hline M & N \end{array} \right).$$

Označíme-li tyto bloky podle (2.46), můžeme A_k^{-1} vypočítat ze vzorců (2.47). Výpočet je snadný, protože $B^{-1} = A_{k-1}^{-1}$, což známe z předcházejícího kroku a matice $(F - DB^{-1}C)$ je typu $(1, 1)$, tedy její inverze je triviální. Pak je výpočet K, L, M, N

realizován pouze základními operacemi s maticemi a vektory. Po nalezení A_n^{-1} je úloha ukončena, protože $A_n = A$, tedy $A^{-1} = A_n^{-1}$.

2.6 Špatně podmíněné soustavy rovnic

Vlivem chyb vstupních údajů neřešíme ve skutečnosti soustavu $A \underline{x} = \underline{b}$, ale soustavu

$$(A + \Delta A) \underline{x} = (\underline{b} + \Delta \underline{b}),$$

kde ΔA je matice chyb prvků a_{ij} a $\Delta \underline{b}$ je vektor chyb pravých stran b_i . Označíme-li $\Delta \underline{x}$ vektor chyb řešení a \underline{x}^* přesné řešení soustavy $A \underline{x} = \underline{b}$, pak platí

$$(2.48) \quad (A + \Delta A)(\underline{x}^* + \Delta \underline{x}) = (\underline{b} + \Delta \underline{b}).$$

Uvažujme na začátku jednodušší případ, kdy $\Delta A = 0$, tedy podle (2.48) je

$$A \Delta \underline{x} = \Delta \underline{b}, \text{ tj. } \Delta \underline{x} = A^{-1} \cdot \Delta \underline{b}, \text{ čili } \|\Delta \underline{x}\| \leq \|A^{-1}\| \cdot \|\Delta \underline{b}\|.$$

Protože $A \cdot \underline{x}^* = \underline{b}$, je tedy

$$\|\underline{b}\| \leq \|A\| \cdot \|\underline{x}^*\|, \text{ čili } \|\underline{x}^*\| \geq \|\underline{b}\| / \|A\|.$$

Pro relativní chybu řešení tedy platí

$$(2.49) \quad \frac{\|\Delta \underline{x}\|}{\|\underline{x}^*\|} \leq \|A\| \cdot \|A^{-1}\| \cdot \frac{\|\Delta \underline{b}\|}{\|\underline{b}\|},$$

kde $\|A\| \cdot \|A^{-1}\| = C_p$ se nazývá číslem podmíněnosti matice A (resp. soustavy $A \underline{x} = \underline{b}$) a je mírou citlivosti relativní chyby řešení k relativní chybě vstupních dat.

Je-li naopak $\Delta A \neq 0$ a $\Delta \underline{b} = 0$, dostaneme z (2.48) $\Delta \underline{x} = -A^{-1} \Delta A (\underline{x}^* + \Delta \underline{x})$ a odтud odhad relativní chyby řešení

$$\frac{\|\Delta \underline{x}\|}{\|\underline{x}^* + \Delta \underline{x}\|} \leq \|A^{-1}\| \cdot \|\Delta A\| = \|A^{-1}\| \cdot \|A\| \cdot \frac{\|\Delta A\|}{\|A\|}.$$

V obecném případě, tedy pro $\Delta A \neq 0$ i $\Delta \underline{b} \neq 0$, je v [1], resp. [6] odvozen vztah

$$(2.50) \quad \frac{\|\Delta \underline{x}\|}{\|\underline{x}^*\|} \leq C_p \frac{\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta \underline{b}\|}{\|\underline{b}\|}}{1 - C_p \frac{\|\Delta A\|}{\|A\|}}.$$

Pro velká C_p dostáváme špatně podmíněné soustavy, kde se malá změna vstupních dat projeví velkou chybou řešení. V takovém případě je vhodné přeformulovat zadанou úlohu tak, aby chom špatně podmíněné soustavy nemuseli řešit.

2.7 Výpočet vlastních čísel

V tomto odstavci se budeme věnovat třetí základní úloze, tj. úloze o vlastních číslech, která je naznačena v odstavci 2.2.

Definice 2.6 Nechť A je čtvercová matice řádu n. Číslo λ , pro které má homogenní soustava

$$(2.51) \quad A \cdot \underline{x} = \lambda \underline{x}, \quad \text{resp.} \quad (A - \lambda E) \cdot \underline{x} = 0$$

nenulové řešení, se nazývá vlastní číslo matice A a jemu odpovídající vektor vlastní vektor matice A.

Z lineární algebry víme, že homogenní soustava má netriviální řešení, když $\det(A - \lambda E) = 0$, tedy vlastní čísla matice A jsou kořeny tzv. charakteristického polynomu

$$(2.52) \quad p_A(\lambda) = \det(A - \lambda E) = (-1)^n \lambda^n + c_{n-1} \lambda^{n-1} + \dots + c_1 \lambda + c_0.$$

Z toho plyne, že matice A má právě n vlastních čísel $\lambda_1, \lambda_2, \dots, \lambda_n$, jestliže počítáme i s násobností. Tedy vlastní číslo λ je k-násobné, když je k-násobným kořenem polynomu $p_A(\lambda)$. Jistě může být λ i komplexní číslo.

Dosadíme-li vlastní číslo λ do homogenní soustavy (2.51), pak má nekonečně mnoho řešení, která vyplňují r-rozměrný vektorový prostor, kde $r = n - h$ a h je hodnota matice soustavy (2.51). Pak k vlastnímu číslu λ přísluší r lineárně nezávislých vektorů. Je-li vlastní číslo λ k-násobné, je $r \leq k$, a tedy k-násobnému vlastnímu číslu λ přísluší maximálně k lineárně nezávislých vlastních vektorů.

Příklad 2.9 Vypočítejme pro matice

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

vlastní čísla.

Řešení: Jistě platí $\det(A_1 - \lambda E) = \det(A_2 - \lambda E) = \det(A_3 - \lambda E) = (1 - \lambda)^3$, tedy $\lambda = 1$ je trojnásobné vlastní číslo každé z daných matic. Matice příslušných soustav tvaru (2.51) mají postupně hodnoty $h_1 = 0$, $h_2 = 1$ a $h_3 = 2$ (ověřte!). Pro matice A_1, A_2, A_3 přísluší trojnásobnému vlastnímu číslu $\lambda = 1$ postupně tři, dva a jeden lineárně nezávislé vlastní vektory. Vypočítejte je!

Připomeňme si ještě některé poznatky z lineární algebry:

1) Jsou-li A, B podobné matice (tzn. že existuje regulární matice podobnosti P tak, že $B = P^{-1}AP$, resp. $A = PBP^{-1}$), pak mají shodná vlastní čísla. To plyně z rovnosti

$$\det(B - \lambda E) = \det P^{-1}(A - \lambda E)P = \det(A - \lambda E).$$

2) Vlastní vektory příslušné k různým vlastním čislům jsou lineárně nezávislé.

3) Nechť čtvercová matice A řádu n má různá vlastní čísla $\lambda_1, \lambda_2, \dots, \lambda_s$ s násobnostmi k_1, k_2, \dots, k_s a nechť ke každému vlastnímu číslu přísluší k_1, k_2, \dots, k_s lineárně nezávislých vlastních vektorů. Pak k matici A lze vybrat n lineárně nezávislých vlastních vektorů, protože platí bod 2) a kromě toho

$$\sum_{i=1}^s k_i = n.$$

4) Je-li matice A symetrická, pak každému jejímu k-násobnému vlastnímu číslu přísluší k lineárně nezávislých vlastních vektorů a vlastní vektory odpovídající různým vlastním čislům jsou ortogonální. Z toho plyně, že symetrická matice řádu n má právě n lineárně nezávislých vektorů a lze je vybrat ortogonální.

5) Symetrická pozitivně definitní matice má všechna vlastní čísla kladná.

V řadě úloh technické praxe nemusíme hledat všechna vlastní čísla, ale stačí nalézt pouze čísla s extrémní absolutní hodnotou. Proto si nejdříve uvedeme jednoduchou iterační metodu pro stanovení maximálního vlastního čísla matice s n lineárně nezávislými vlastními vektory.

2.7.1 Mochinná metoda

Nechť matice A má n lineárně nezávislých vlastních vektorů $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$, které odpovídají vlastním číslům

$$(2.53) \quad |\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n| ,$$

kde λ_1 má maximální absolutní hodnotu a je jednonásobné. Zvolme počáteční vektor $\underline{v}^{(0)}$ a podle rekurentního vzorce

$$(2.54) \quad \underline{v}^{(k)} = A \cdot \underline{v}^{(k-1)} = A^k \cdot \underline{v}^{(0)}$$

počítejme členy posloupnosti $\underline{v}^{(0)}, \underline{v}^{(1)}, \dots$. Protože vektory $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ tvoří bázi v n-rozměrném vektorovém prostoru aritmetických n-členných vektorů, pak

$$\underline{v}^{(0)} = a_1 \cdot \underline{x}_1 + a_2 \cdot \underline{x}_2 + \dots + a_n \cdot \underline{x}_n$$

a vzhledem k (2.54) a rovnosti $A \cdot \underline{x}_i = \lambda_i \cdot \underline{x}_i$ také platí

$$\underline{v}^{(k)} = a_1 \lambda_1^k \cdot \underline{x}_1 + a_2 \lambda_2^k \cdot \underline{x}_2 + a_n \lambda_n^k \cdot \underline{x}_n .$$

Zvolme ještě vektor \underline{y} a vyšetřujme podíl skalárních součinů $(\underline{y}, \underline{v}^{(k+1)}) / (\underline{y}, \underline{v}^{(k)})$. Pak dostaneme

$$(2.55) \quad \frac{(\underline{y}, \underline{v}^{(k+1)})}{(\underline{y}, \underline{v}^{(k)})} = \frac{\lambda_1^{k+1} a_1 (\underline{y}, \underline{x}_1) + \lambda_2^{k+1} a_2 (\underline{y}, \underline{x}_2) + \dots + \lambda_n^{k+1} a_n (\underline{y}, \underline{x}_n)}{\lambda_1^k a_1 (\underline{y}, \underline{x}_1) + \lambda_2^k a_2 (\underline{y}, \underline{x}_2) + \dots + \lambda_n^k a_n (\underline{y}, \underline{x}_n)} = \\ = \frac{\lambda_1 + \lambda_1 \sum_{i=2}^n (\lambda_i / \lambda_1)^{k+1} (a_i / a_1) [(\underline{y}, \underline{x}_i) / (\underline{y}, \underline{x}_1)]}{1 + \sum_{i=2}^n (\lambda_i / \lambda_1)^k (a_i / a_1) [(\underline{y}, \underline{x}_i) / (\underline{y}, \underline{x}_1)]} ,$$

tedy

$$\lim_{k \rightarrow \infty} \frac{(\underline{y}, \underline{v}^{(k+1)})}{(\underline{y}, \underline{v}^{(k)})} = \lambda_1 ,$$

protože $\lim_{k \rightarrow \infty} (\lambda_i / \lambda_1)^k = 0$ pro $i = 2, 3, \dots, n$, neboť $|\lambda_i / \lambda_1| < 1$. Samozřejmě musíme předpokládat $a_1 \neq 0$ a $(\underline{y}, \underline{x}_1) \neq 0$. Pomocný vektor \underline{y} volíme v praxi tak, aby jeho složka odpovídající maximální složce vektoru $\underline{v}^{(k)}$ byla 1 a ostatní složky byly 0. Tato volba minimalizuje objem výpočtu i chybu podílu (2.55). Tím approximuje vlastní číslo λ_1 podílem odpovídajících si složek vektorů $\underline{v}^{(k+1)}$ a $\underline{v}^{(k)}$, přičemž pro podíl si vybíráme ty složky, kde má vektor $\underline{v}^{(k)}$ složku s maximální absolutní hodnotou.

Abychom se vyhnuli počítání s příliš velkými čísly, je vhodné nějakým způsobem vektory $\underline{v}^{(k)}$ normalizovat (např. udržovat maximální složku rovnou jedné).

Poznámka 2.7 Moeninnou metodu lze užít i v případě, kdy λ_1 je vícenásobné, nebo v případě, kdy $\lambda_1 = -\lambda_2$ a v dalších případech, viz [11].

Poznámka 2.8 Vektor $\underline{v}^{(k)}$ konverguje pro $k \rightarrow \infty$ k vlastnímu vektoru, který přísluší vlastnímu číslu λ_1 , protože

$$(2.57) \quad \underline{v}^{(k)} = a_1 \lambda_1^k \left(\underline{x}_1 + \sum_{i=2}^n a_i \left(\frac{\lambda_i}{\lambda_1} \right)^k \underline{x}_i \right) \rightarrow a_1 \lambda_1^k \underline{x}_1, \text{ pro } k \rightarrow \infty.$$

V praktických případech většinou iterační proces konverguje pomalu a je třeba nějak jej zrychlit.

2.7.2 Urychlení konvergence moeninné metody

Nechť A^T je transponovaná matici A . Obě mají samozřejmě stejná vlastní čísla (neboť $\det(A - \lambda E) = \det(A^T - \lambda E)$), pro která platí (2.53). Dále nechť $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$, resp. $\underline{z}_1, \underline{z}_2, \dots, \underline{z}_n$ jsou vlastní vektory matic A , resp. A^T , které tvoří lineárně nezávislé systémy.

Zvolme vektory $\underline{v}^{(0)}, \underline{u}^{(0)}$, které můžeme zapsat jako lineární kombinace

$$(2.58) \quad \underline{v}^{(0)} = \sum_{i=1}^n a_i \underline{x}_i, \quad \underline{u}^{(0)} = \sum_{i=1}^n b_i \underline{z}_i$$

a obdobně jako v předcházejícím počítejme vektory $\underline{v}^{(k)}, \underline{u}^{(k)}$ podle vzorců

$$(2.59) \quad \underline{v}^{(k)} = A \cdot \underline{v}^{(k-1)} = \sum_{i=1}^n a_i \lambda_i^k \underline{x}_i, \quad \underline{u}^{(k)} = A^T \cdot \underline{u}^{(k-1)} = \sum_{i=1}^n b_i \lambda_i^k \underline{z}_i.$$

Vyšetřujeme-li podíl $(\underline{v}^{(k+1)}, \underline{u}^{(k+1)}) / (\underline{v}^{(k)}, \underline{u}^{(k+1)})$, dostaneme

$$(2.60) \quad \frac{(\underline{v}^{(k+1)}, \underline{u}^{(k+1)})}{(\underline{v}^{(k)}, \underline{u}^{(k+1)})} = \frac{\lambda_1^{2k+2} a_1 b_1 + \lambda_2^{2k+2} a_2 b_2 + \dots + \lambda_n^{2k+2} a_n b_n}{\lambda_1^{2k+1} a_1 b_1 + \lambda_2^{2k+1} a_2 b_2 + \dots + \lambda_n^{2k+1} a_n b_n}$$

protože soustavy vlastních vektorů lze vybrat tak, aby byly vzájemně ortonormální, tedy aby platilo

$$(\underline{x}_i, \underline{z}_j) = 0 \text{ pro } i \neq j \text{ a } (\underline{x}_i, \underline{z}_i) = 1,$$

což je dokázáno např. v [11].

Protože platí $|\lambda_i/\lambda_1| < 1$ pro $i = 2, 3, \dots, n$ a předpokládáme, že $a_1 \neq 0$, $b_1 \neq 0$, pak

$$(2.61) \quad \lim_{k \rightarrow \infty} \frac{(\underline{v}^{(k+1)}, \underline{u}^{(k+1)})}{(\underline{v}^{(k)}, \underline{u}^{(k+1)})} = \lambda_1.$$

Je zřejmé, že konvergence je zrychlena, protože podíly λ_i / λ_1 se ve (2.60) vyskytují nejméně v mocnině $2k + 1$, zatímco ve (2.55) byly v mocnině k .

Tato metoda urychlení se také nazývá metoda skalárních součinů.

2.7.3 Mocninná metoda pro symetrickou matici

Předcházející metoda skalárních součinů je zvláště výhodná pro symetrickou matici ($A = A^T$), kde zvolíme $\underline{u}^{(0)} = \underline{v}^{(0)}$, tedy $\underline{u}^{(k)} = \underline{v}^{(k)}$. Potom podíl (2.60) obsahuje pouze vektory \underline{v} a (2.61) má tvar

$$(2.62) \quad \lim_{k \rightarrow \infty} \frac{(\underline{v}^{(k+1)}, \underline{v}^{(k+1)})}{(\underline{v}^{(k)}, \underline{v}^{(k+1)})} = \lambda_1.$$

Příklad 2.10 Vypočítejme přibližnou hodnotu maximálního vlastního čísla matice

$$A = \begin{pmatrix} 1 & 1 & 1/2 \\ 1 & 1 & 1/4 \\ 1/2 & 1/4 & 2 \end{pmatrix}$$

mocninnou metodou, a to jednak podle základního vztahu (2.56) s počátečním vektorem $\underline{v}^{(0)T} = (1, 1, 1)$ a potom s urychlením a s využitím symetrie, tedy podle (2.62).

Řešení: Postupné approximace jsou sestaveny v tabulce 2.7. Každý vektor $\underline{v}^{(k)}$ je před výpočtem další approximace "normalizován" tak, aby jeho maximální složka byla rovna 1. Veličiny $\lambda_1^{(k)}$, resp. $\tilde{\lambda}_1^{(k)}$ jsou počítány podle (2.55), resp. (2.60) z vektoru $\underline{v}^{(k)}$ a z vektoru $\underline{v}^{(k+1)}$ před "normalizací".

k	$v_1^{(k)}$	$v_2^{(k)}$	$v_3^{(k)}$	$\lambda_1^{(k)}$	$\tilde{\lambda}_1^{(k)}$
0	1	1	1		
1	0,9091	0,8182	1,000	2,75000	2,51667
5	0,7651	0,6674	1,000	2,55879	2,53647
10	0,7494	0,6508	1,000	2,53800	2,53653
15	0,7483	0,6497	1,000	2,53663	2,53653
20	0,7482	0,6497	1,000	2,53653	2,53653

Tab. 2.7

2.7.4 Výpočet dalších vlastních čísel

Jedna z možností, jak pokračovat ve výpočtu druhého a dalších vlastních čísel, je postavena na následující větě.

Věta 2.5 Nechť $\lambda_1, \underline{x}_1$ je vlastní číslo a jemu odpovídající vlastní vektor matice A. Zvolme vektor \underline{u} tak, aby $(\underline{u}, \underline{x}_1) = 1$. Pak matice

$$B = A - \lambda_1 \underline{x}_1 \underline{x}_1^T$$

(\underline{u}^T je řádkový vektor, viz označení v odstavci 2.1.1)
 má též vlastní čísla jako matice A, s výjimkou λ_1 , které je nahrazeno nulou.
Důkaz: viz např. [23].

Pak tedy maximální vlastní číslo matice B je λ_2 , které můžeme určit mocninnou metodou a obdobný proces opakovat pro hledání λ_3 atd. Problematická je však volba pomocného vektoru \underline{u} .

2.7.5 LU-rozklad pro úplný problém vlastních čísel

Uvedeme si bez důkazu jenom základní výpočetní princip. Nechť L.U je trojúhelníkový rozklad matice A jako v odstavci 2.3.4. Matice $B = U \cdot L$ je podobná matici A, protože $B = L^{-1}AL$, neboť $U = L^{-1}A$. Tedy A a B mají stejná vlastní čísla.

Z LU-rozkladu $A = A_0 = L_0 U_0$ sestrojíme podobnou matici $A_1 = U_0 L_0$, jejíž LU-rozklad je $A_1 = L_1 U_1$ atd. Tímto postupem dostaneme posloupnost podobných matic

$$(2.63) \quad A_k = U_{k-1} L_{k-1} = L_k U_k \quad \text{pro } k = 1, 2, \dots$$

Dá se dokázat, že když matice $B_k = L_1 \cdot L_2 \cdot \dots \cdot L_k$ konvergují k regulérní matici, pak matice A_k konvergují k horní trojúhelníkové matici s vlastními čísly matice A na diagonále.

Nevýhodou LU-rozkladu je pomalá konvergence a velký počet operací. Pro úplný problém vlastních čísel se užívá také metod založených na výpočtu charakteristického polynomu a mnoha dalších. Podrobněji se s těmito metodami můžete seznámit např. v publikacích [25], [11].

2.8 Cvičení

1. Gaussovou a Jordanovou eliminační metodou bez výběru hlavních prvků řešte následující soustavy lineárních algebraických rovnic. Výsledky porovnejte a zpětným dosazením do soustavy ověřte splnění zadávaných rovnic.

a) $0,197 x_1 + 0,305 x_2 - 0,206 x_3 - 0,084 x_4 = 0,137$

$0,468 x_1 + 0,713 x_2 - 0,474 x_3 + 0,052 x_4 = 0,119$

$0,886 x_1 + 0,764 x_2 - 0,108 x_3 + 0,802 x_4 = 0,253$

$0,145 x_1 + 0,590 x_2 + 0,613 x_3 + 0,365 x_4 = 0,661$

b) $6,036 x_1 + 0,732 x_2 - 2,134 x_3 + 1,324 x_4 = -0,930$

$-4,821 x_1 + 0,356 x_2 + 1,052 x_3 + 4,267 x_4 = 5,713$

$3,006 x_1 - 2,307 x_2 + 1,506 x_3 + 0,728 x_4 = 8,361$

$5,604 x_1 + 4,112 x_2 + 3,283 x_3 + 3,415 x_4 = 8,478$

c) $3,2 x_1 + 1,4 x_2 + 0,6 x_3 = 3,0 \quad \text{d)} \quad 9,8 x_1 + 2,3 x_2 - 3,7 x_3 = 5,8$

$1,4 x_1 - 2,7 x_2 + 1,1 x_3 = 6,3 \quad -3,1 x_1 + 8,4 x_2 + 1,6 x_3 = 23,4$

$-2,1 x_1 + 1,2 x_2 + 9,6 x_3 = 15,9 \quad -4,5 x_1 + 1,6 x_2 + 7,7 x_3 = -24,7$

2. Soustavy z předcházejícího cvičení řešte Gaussovou a Jordanovou eliminační metodou s řádkovým, sloupcovým, resp. úplným výběrem hlavních prvků. Přesnost výsledků ověřte zpětným dosazením do soustav a porovnejte s výsledky prvního cvičení.

3. Proveďte LU-rozklad matic

$$A = \begin{pmatrix} 2 & 5 & 0 \\ 6 & 12 & -2 \\ 10 & 4 & -13 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 5 & 6 \\ 4 & 13 & 19 \\ 6 & 27 & 50 \end{pmatrix}, \quad C = \begin{pmatrix} 4 & -1 & 1 \\ -4 & 6 & 1 \\ 8 & 13 & 10 \end{pmatrix}$$

a jeho pomocí vypočítejte jejich determinanty užitím vztahu ($A = L \cdot U$, ...)

$$\det A = \det L \cdot U = \det L \cdot \det U = \det U = u_{11} \cdot u_{22} \cdots u_{nn}.$$

4. Jacobiovou a Gaussovou-Seidelovou iterační metodou řešte soustavy lineárních algebraických rovnic

$$\begin{aligned} a) \quad 5,21 x_1 + 1,52 x_2 - 2,37 x_3 &= 1,32 \\ 1,72 x_1 - 2,97 x_2 + 0,21 x_3 &= 4,90 \\ 2,01 x_1 + 0,92 x_2 + 3,89 x_3 &= 4,98 \end{aligned}$$

$$\begin{aligned} b) \quad 12 x_1 - 2 x_2 + 3 x_3 &= 55 \\ -x_1 + 21 x_2 - 5 x_3 &= -70 \\ 2 x_1 - 5 x_2 - 13 x_3 &= -49 \end{aligned}$$

Za počáteční aproksimaci volte $x^{(0)} = (0, 0, 0)$ a výpočet ukončete podmínkou $\|x^{(k)} - x^{(k-1)}\| < 10^{-3}$.

5. Invertujte matice

$$A = \begin{pmatrix} 4,312 & 2,156 & 3,142 \\ 0,849 & -1,238 & 4,006 \\ -2,139 & 0,873 & 2,154 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & -2 & 1 \\ 2 & 0 & -1 \\ 1 & -3 & 3 \end{pmatrix}$$

a) eliminací, b) LU-rozkladem, c) metodou vroubení a porovnejte jednotlivé výsledky.

6. Vypočítejte vlastní číslo s maximální absolutní hodnotou pro matice

$$A = \begin{pmatrix} 4 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & 4 & 1 & 1 \\ 4 & 5 & 1 & 1 \\ 1 & 1 & 4 & 2 \\ 1 & 1 & 2 & 4 \end{pmatrix}.$$

Výsledky

1. a) $x_1 \doteq 0,911, \quad x_2 \doteq 0,335, \quad x_3 \doteq 1,058, \quad x_4 \doteq -0,868,$
- b) $x_1 \doteq 0,415, \quad x_2 \doteq -1,306, \quad x_3 \doteq 2,039, \quad x_4 \doteq 1,414,$
- c) $x_1 = 1, \quad x_2 = -1, \quad x_3 = 2,$
- d) $x_1 = -2, \quad x_2 = 3, \quad x_3 = -5.$

$$3. \quad A = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 5 & 7 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 5 & 0 \\ 0 & -3 & -2 \\ 0 & 0 & 1 \end{pmatrix}, \quad \det A = -6,$$

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 5 & 6 \\ 0 & 3 & 7 \\ 0 & 0 & 4 \end{pmatrix}, \quad \det B = 24,$$

$$C = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 4 & -1 & 1 \\ 0 & 5 & 2 \\ 0 & 0 & 2 \end{pmatrix}, \quad \det C = 40.$$

$$4. \quad a) \quad x_1 = 1, \quad x_2 = -1, \quad x_3 = 1, \quad b) \quad x_1 = 3, \quad x_2 = -2, \quad x_3 = 5$$

$$5. \quad A^{-1} = \begin{pmatrix} 0,112\ 096 & 0,034\ 573 & -0,227\ 811 \\ 0,189\ 090 & -0,291\ 134 & 0,265\ 630 \\ 0,034\ 670 & 0,152\ 327 & 0,130\ 370 \end{pmatrix}, \quad B^{-1} = \begin{pmatrix} 3 & -3 & -2 \\ 7 & -8 & -5 \\ 6 & -7 & -4 \end{pmatrix}$$

$$6. \quad \lambda_1(A) = 4,46, \quad \lambda_1(B) = 10.$$

K A P I T O L A 3.

A P R O X I M A C E F U N K C E

3.1 Úloha approximace

Aproximovat funkci znamená nalézt k ní nějakou kombinaci (obvykle lineární) speciálních funkcí, která bude v jistém smyslu blízká k funkci zadané. Důvodů approximace funkce je více a jsou různorodé. Může to být např. složitost funkčního předpisu, která znesnadňuje výpočet, nebo potřeba numerického výpočtu derivací a integrálů či jiných charakteristik. Funkci potřebujeme approximovat také tehdy, když neznáme její analytické vyjádření a chceme z tabulky (často jenom měřených) hodnot nalézt přibližné vyjádření funkční závislosti.

V celé této kapitole se budeme zabývat approximacemi spojité funkce $f(x)$ na intervalu $\langle a, b \rangle$, která je zadána buď analytickým vyjádřením nebo pro ni známe jenom tzv. tabulku funkce

$$(3.1) \quad Tf = \{(x_i, y_i) : y_i = f(x_i), x_i \in \langle a, b \rangle, i = 0, 1, \dots, m\}.$$

Diskrétní množina $X = \{x_0, x_1, \dots, x_m\} \subset \langle a, b \rangle$ obsahuje vzájemně různé body, které jsou obvykle seřazeny monotonně.

Za approximující funkci budeme brát tzv. zobecněný polynom n -tého stupně

$$(3.2) \quad \phi_n(x) = c_0 g_0(x) + c_1 g_1(x) + \dots + c_n g_n(x),$$

kde c_0, c_1, \dots, c_n jsou reálné koeficienty a $g_0(x), g_1(x), \dots, g_n(x)$ je lineárně nezávislý systém jednoduchých, dostatečně hladkých funkcí, definovaných na intervalu $\langle a, b \rangle$, které nazýváme základní funkce.

Jsou-li základní funkce (algebraické) polynomy i -tého stupně, $i = 0, 1, \dots, n$ (např. $1, x, \dots, x^n$ nebo $1, (x - x_0), \dots, (x - x_0)^n$, kde x_0 je pevné), pak (3.2) je (algebraický) polynom nejvýše n -tého stupně. Při základních funkcích $1, \cos x, \sin x, \cos 2x, \sin 2x, \dots, \cos Lx, \sin Lx$ ($n = 2L$) hovoříme o trigonometrickém polynomu n -tého stupně.

Úlohou approximace spojité funkce $f(x)$ na intervalu $\langle a, b \rangle$ rozumíme tedy najít takový zobecněný polynom (3.2), aby byl v jistém smyslu blízký funkci $f(x)$ na intervalu $\langle a, b \rangle$. Podle kritérií pro hledání konstant c_0, c_1, \dots, c_n si ukáže me tři typy approximací. Jsou to:

- 1) Interpolaci approximace, kdy konstanty c_0, c_1, \dots, c_n jsou voleny tak, aby v bodech $x_0, x_1, \dots, x_m \in \langle a, b \rangle$ byly hodnoty funkce $\phi_n(x)$ (případně některých derivací) stejné jako hodnoty (případně derivace) funkce $f(x)$.
- 2) Approximace metodou nejmenších čtverců (MNČ), kdy c_0, c_1, \dots, c_n hledáme z podmínky minimalizace buď

$$(3.3) \quad \sqrt{\int_a^b [f(x) - \phi_n(x)]^2 dx} \quad \text{nebo} \quad \sqrt{\sum_{i=0}^m [f(x_i) - \phi_n(x_i)]^2},$$

kde $m > n$. Pak hovoříme buď o spojitém nebo diskrétním případu approximace metodou nejmenších čtverců a funkci $\phi_n(x)$ nazýváme nejlepší approximace.

- 3) Čebyševova approximace, jejímž cílem je volit koeficienty c_0, c_1, \dots, c_n tak, aby se minimalizoval maximální rozdíl mezi $f(x)$ a $\phi_n(x)$ na intervalu $\langle a, b \rangle$, čili aby se minimalizovala hodnota

$$(3.4) \quad \max_{x \in \langle a, b \rangle} |f(x) - \phi_n(x)|.$$

Tento druh approximace se také nazývá nejlepší stejnoměrná approximace.

V tomto učebním textu se budeme téměř výhradně zabývat approximacemi pomocí polynomů. Tím rozumíme případ, kdy za základní funkce volíme polynomy nultého až n -tého stupně, např. 1, x, x^2, \dots, x^n . Approximace polynomy se v numerické matematice často užívají, protože jejich hodnota se dá vypočítat konečným počtem aritmetických operací, snadno se derivují a integraruji a i jinak se s nimi matematicky snadno pracuje. Všechny tyto a jiné výhody by však byly nepodstatné, kdyby neplatila následující věta, která zaručuje, že pomocí polynomů můžeme dosáhnout approximace libovolně vysoké přesnosti.

Věta 3.1 Je-li $f(x)$ spojitá na konečném intervalu $\langle a, b \rangle$, pak ke každému $\epsilon > 0$ existuje n (závislé na ϵ) a polynom n -tého stupně $p_n(x)$ tak, že

$$(3.5) \quad \max_{x \in \langle a, b \rangle} |f(x) - p_n(x)| < \epsilon$$

Důkaz viz např. [1] nebo [11].

Poznámka 3.1 Zcela analogická věta platí, jestliže místo tvrzení (3.5) zapíšeme nerovnost

$$\sqrt{\int_a^b [(f(x) - p_n(x))^2 dx]} < \epsilon, \quad \text{resp.} \quad \sqrt{\sum_{i=0}^m [f(x_i) - p_n(x_i)]^2} < \epsilon,$$

kde x_0, x_1, \dots, x_m ($m > n$) je množina diskrétních bodů z $\langle a, b \rangle$.

Poznámka 3.2 Analogické úvahy bychom mohli provést i pro approximaci pomocí trigonometrických polynomů, viz [1] a [11].

Poznámka 3.3 Tomuto pojednání approximace se poněkud vymyká odstavec 3.3, ve kterém se budeme zabývat tzv. approximací po částech. To je approximace, při níž rozdělíme interval $\langle a, b \rangle$ na subintervaly a v každém z nich zvolíme jinou approximační funkci. Ve "styčných" bodech musíme však zajistit nějaký stupeň hladkosti.

3.2 Interpolační approximace pomocí polynomů

Nechť pro spojitou funkci $f(x)$ na intervalu $\langle a, b \rangle$ jsou zadány funkční hodnoty, případně hodnoty některých derivací v bodech $x_0, x_1, \dots, x_m \in \langle a, b \rangle$, které nazýváme uzly interpolace. Předmětem interpolační approximace pomocí polynomů je nalezení takového polynomu nejvýše n -tého stupně

$$(3.5) \quad \phi_n(x) = c_0 + c_1 x + c_2 x^2 + \dots + c_n x^n,$$

který má v uzlech x_0, x_1, \dots, x_m s interpolovanou funkci shodné funkční hodnoty, případně některé derivace.

Požadujeme-li jenom shodu funkčních hodnot $\phi_n(x_i) = f(x_i)$ pro $i = 0, 1, \dots, m$, pak musí platit $m = n$.

3.2.1 Lagrangeův interpolační polynom

Mějme pro spojitou funkci $y = f(x)$ na intervalu $\langle a, b \rangle$ dánou tabulku Tf, viz (3.1). Pod pojmem Lagrangeův interpolační polynom budeme rozumět (algebraický) polynom (3.5) maximálně n -tého stupně ($n = m$), označený $L_n(x)$, pro který platí

$$(3.6) \quad L_n(x_0) = y_0, \quad L_n(x_1) = y_1, \dots, L_n(x_n) = y_n.$$

Podmínky (3.6) tvoří nehomogenní soustavu $n+1$ lineárních rovnic o $n+1$ neznámých c_0, c_1, \dots, c_n , jejíž determinant je tzv. Vandermondův determinant, který je nenulový, protože body x_0, x_1, \dots, x_n jsou vzájemně různé. Tím je zaručena jednoznačná existence polynomu $L_n(x)$, ale vyčíslení koeficientů c_0, c_1, \dots, c_n ze soustavy (3.6) je příliš těžkopádné, a proto budeme postupovat jinak.

Je zřejmé, že funkce

$$F_i(x) = \frac{(x - x_0)(x - x_1) \dots (x - x_{i-1})(x - x_{i+1}) \dots (x - x_n)}{(x_i - x_0)(x_i - x_1) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_n)}, \quad i = 0, 1, \dots, n,$$

jsou polynomy n -tého stupně s nulovými hodnotami v bodech $x_0, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$, pro které platí $F_i(x_i) = 1$. Pak lineární kombinace funkcí $F_0(x), F_1(x), \dots, F_n(x)$ s koeficienty y_0, y_1, \dots, y_n je hledaný Lagrangeův interpolační polynom, tedy

$$(3.7) \quad \begin{aligned} L_n(x) &= \sum_{i=0}^n y_i \cdot F_i(x) = \\ &= \sum_{i=0}^n y_i \frac{(x - x_0)(x - x_1) \dots (x - x_{i-1})(x - x_{i+1}) \dots (x - x_n)}{(x_i - x_0)(x_i - x_1) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_n)} = \\ &= \sum_{i=0}^n y_i \frac{\omega_n(x)}{(x - x_i) \omega'_n(x_i)}, \end{aligned}$$

kde $\omega_n(x) = (x - x_0)(x - x_1) \dots (x - x_n)$ a

$$\omega'_n(x_i) = (x_i - x_0)(x_i - x_1) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_n).$$

Jsou-li interpolační uzly uspořádány ekvidistantně, tzn. že

$$(3.8) \quad x_{i+1} - x_i = h \quad \text{pro} \quad i = 0, 1, 2, \dots, n-1,$$

kde konstanta h se nazývá krokem, pak můžeme zavést novou proměnnou t vztahem

$$(3.9) \quad t = (x - x_0)/h, \quad \text{čili} \quad x = x_0 + th,$$

tedy $x_i = x_0 + ih$, $x - x_i = h(t - i)$ a $x_i - x_j = h(i - j)$.

Dosadíme-li nyní do (3.7), dostaneme Lagrangeův interpolační polynom pro ekvidistantní argumenty

$$(3.10) \quad L_n(x) = L_n(x_0 + th) = \sum_{i=0}^n y_i \frac{t(t-1)\dots(t-n)}{(t-i)i!(n-i)!(-1)^{n-i}} = \\ = \frac{t(t-1)\dots(t-n)}{n!} \sum_{i=0}^n (-1)^{n-i} \binom{n}{i} \frac{1}{t-i} y_i.$$

Příklad 3.1 Sestavme Lagrangeovy interpolační polynomy $L_1(x)$, $L_2(x)$, $L_3(x)$ pro hodnoty dané v tabulce 3.1, kdy $L_1(x)$ je definován na uzlech $x_0, x_1, L_2(x)$ na uzlech x_0, x_1, x_2 a $L_3(x)$ na uzlech x_0, x_1, x_2, x_3 . Dále vypočítejme přibližné hodnoty zadané funkce v bodě 0,5 užitím každého z polynomů $L_1(x)$, $L_2(x)$ a $L_3(x)$.

i	0	1	2	3
x_i	0	2	3	5
y_i	1	3	2	5

Tab. 3.1

Řešení:

$$L_1(x) = \frac{x-2}{-2} 1 + \frac{x-0}{2} 3 = x + 1$$

$$L_2(x) = \frac{(x-2)(x-3)}{(-2)(-3)} 1 + \frac{(x-0)(x-3)}{2(-1)} 3 + \frac{(x-0)(x-2)}{3 \cdot 1} 2 = \\ = -2/3x^2 + 7/3x + 1$$

$$L_3(x) = \frac{(x-2)(x-3)(x-5)}{(-2)(-3)(-5)} 1 + \frac{(x-0)(x-2)(x-5)}{2(-1)(-3)} 3 + \\ + \frac{(x-0)(x-2)(x-5)}{3 \cdot 1 \cdot (-2)} 2 + \frac{(x-0)(x-2)(x-3)}{5 \cdot 3 \cdot 2} 5 = 3/10x^3 - 13/6x^2 + 62/15x + 1$$

$$L_1(0,5) = 1,5 \quad L_2(0,5) = 2,0 \quad L_3(0,5) = 2,5625$$

3.2.2 Nevillův algoritmus

Při některých použitích interpolačního polynomu nepotřebujeme znát jeho tvar (jeho koeficienty), ale hledáme jenom jeho hodnotu $L_n(x)$ pro $x = \tilde{x}$. Pro tento účel si ukážeme postup, který se nazývá Nevillův algoritmus nebo také iterovaná interpolace – viz [3], resp. [20], kdy hodnotu $L_n(\tilde{x})$ můžeme vypočítat přímo ze vstupních dat pomocí jisté rekurence.

Opět předpokládejme, že pro spojitou funkci $f(x)$ na $\langle a, b \rangle$ známe tabulkou Tf – viz (3.1) – pro $m = n$. Ke každému uzlu x_i , $i = 0, 1, \dots, n$ můžeme vybrat skupinu $k+1$ uzlů $x_{i-k}, x_{i-k+1}, \dots, x_i$ pro $k = 0, 1, \dots, i$ a definovat nad

nimi Lagrangeův interpolační polynom k-tého stupně, který označme $L_{ik}(x)$. Pro $k = 0$ položíme $L_{i0}(x) = y_i$. Polynomy $L_{ik}(x)$ (zjednoúšeně L_{ik}), tzv. interpolanty, sestavíme spolu se zadanými hodnotami do tabulky 3.2, která se nazývá Nevilovo schéma.

x	y	$k = 0$	$k = 1$	$k = 2$...	$k = i-1$	$k = i$...	$k = n$
x_0	y_0	L_{00}							
x_1	y_1	L_{10}	L_{11}						
x_2	y_2	L_{20}	L_{21}	L_{22}					
....					
x_{i-1}	y_{i-1}	$L_{i-1,0}$	$L_{i-1,1}$	$L_{i-1,2}$...	$L_{i-1,i-1}$			
x_i	y_i	L_{i0}	L_{i1}	L_{i2}	...	$L_{i,i-1}$	L_{ii}		
....	
x_n	y_n	L_{n0}	L_{n1}	L_{n2}	...	$L_{n,i-1}$	L_{ni}	...	L_{nn}

Tab. 3.2

Snadno ověříme, že interpolant $L_{ik}(x)$ můžeme vypočítat z rekurentního vztahu

$$(3.11) \quad L_{ik}(x) = \frac{1}{x_i - x_{i-k}} \begin{vmatrix} L_{i-1,k-1}(x) & (x_{i-k} - x) \\ L_{i,k-1}(x) & (x_i - x) \end{vmatrix}.$$

Především je zřejmé, že na pravé straně je polynom k-tého stupně, protože $L_{i-1,k-1}$ a $L_{i,k-1}$ jsou polynomy stupně $k-1$. Dosadíme-li do pravé strany (3.11) uzly $x_{i-k}, x_{i-k+1}, \dots, x_i$ a provedeme naznačené operace, pak dostaneme hodnoty $y_{i-k}, y_{i-k+1}, \dots, y_i$. Tedy $L_{ik}(x)$ je interpolační polynom sestavený nad uzly $x_{i-k}, x_{i-k+1}, \dots, x_i$, což jsme chtěli dokázat. Ověřte si vzorec (3.11) na zadání příkladu 3.1.

Na základě vzorce (3.11) můžeme postupně vytvářet polynomy $L_{ik}(x)$ (v i-tém řádku tabulky 3.2) z předcházejících polynomů $L_{i-1,k-1}(x)$ a $L_{i,k-1}(x)$ podle vztahů

$$(3.12) \quad \begin{aligned} L_{i0}(x) &= y_i, \\ L_{ik}(x) &= L_{i,k-1}(x) + (x - x_i) \frac{L_{i,k-1}(x) - L_{i-1,k-1}(x)}{x_i - x_{i-k}}, \quad k = 1, 2, \dots, i. \end{aligned}$$

Polynom $L_{nn}(x)$, sestavený pro uzly x_0, x_1, \dots, x_n jsme dříve označovali s jedním indexem $L_n(x)$. Postup (3.12) se nazývá Nevillův algoritmus, který se však neužívá ke konstrukci polynomu $L_n(x)$, ale k výpočtu jeho hodnoty. Vztah (3.11), resp. (3.12) je samozřejmě splněn také pro hodnoty polynomů $L_{ik}(x)$, tedy pro případ, že $x = \tilde{x}$. Pak $L_{ik}(\tilde{x})$ jsou konstanty a $L_{nn}(\tilde{x})$ je hledaná hodnota získaná

bez výpočtu koeficientů polynomu $L_{nn}(x)$. Tento algoritmus je výhodný také proto, že pro případný výpočet hodnoty $L_{n+1,n+1}(\tilde{x})$, která má přesněji approximovat hodnotu $f(\tilde{x})$ než $L_{nn}(\tilde{x})$, stačí v tabulce 3.2 doložit Nevilllovým algoritmem (3.12) další řádek bez obtížného sestavení polynomu $L_{n+1,n+1}(x) = L_{n+1}(x)$.

Z odvození Nevilllova algoritmu vyplývá, že uzly nemusí být uspořádány monotóně. Je dokonce výhodné uspořádat je podle rostoucí vzdálenosti od \tilde{x} nebo tak, aby $x_0 < \tilde{x} < x_1$.

Příklad 3.2 Sestavme tabulku 3.2 pro zadání z příkladu 3.1.

Řešení: Nevillův algoritmus má podle (3.12) tvar:

$$\begin{aligned} \text{Vstupní údaje: } \tilde{x} &= 0,5, & x_0 &= 0, & x_1 &= 2, & x_2 &= 3, & x_3 &= 5 \\ y_0 &= 1, & y_1 &= 3, & y_2 &= 2, & y_3 &= 5 \end{aligned}$$

Pro $i = 0, 1, 2, 3$ postupně počítáme

$$\begin{aligned} L_{i0}(\tilde{x}) &= y_i, \\ L_{ik}(\tilde{x}) &= L_{i,k-1}(\tilde{x}) + (\tilde{x} - x_i) \frac{L_{i,k-1}(\tilde{x}) - L_{i-1,k-1}(\tilde{x})}{x_i - x_{i-k}}, \quad k = 1, 2, \dots, i. \end{aligned}$$

$$\text{Výstup: } L_3(0,5) = L_{33}(\tilde{x})$$

Numerické výsledky jsou sestaveny v následující tabulce 3.3.

i	x_i	y_i	$\tilde{x} - x_i$	L_{i0}	L_{i1}	L_{i2}	L_{i3}
0	0	1	0,5	1			
1	2	3	-1,5	3	1,5		
2	3	2	-2,5	2	4,5	2,0	
3	5	5	-4,5	5	-1,75	7,625	2,5625

3.2.3 Newtonův interpolační polynom

Dříve než začneme hovořit o Newtonově interpolačním polynomu, řekneme si pár slov o poměrných a obyčejných diferencích, které jsou spojeny s tabulkou funkčních hodnot.

Definice 3.1 Nechť pro spojitou funkci $f(x)$ na intervalu (a, b) máme zadanou tabulkou T_f . Potom podíly

$$(3.13) \quad f(x_i; x_{i+1}) = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}$$

nazýváme poměrné diferenze 1. řádu a podíly

$$(3.14) \quad f(x_i; x_{i+1}; \dots, x_{i+k}) =$$

$$= \frac{f(x_{i+1}; x_{i+2}; \dots; x_{i+k}) - f(x_i; x_{i+1}; \dots; x_{i+k-1})}{x_{i+k} - x_i}$$

pro $k = 2, 3, \dots$ nazýváme poměrné diference k-tého řádu.

Obdobně budeme rozdíly

$$(3.15) \quad \Delta^1 f_i = y_{i+1} - y_i$$

nazývat obyčejné diference 1. řádu a rozdíly

$$(3.16) \quad \Delta^k f_i = \Delta^{k-1} f_{i+1} - \Delta^{k-1} f_i$$

pro $k = 2, 3, \dots$ obyčejné diference k-tého řádu.

Poměrné i obyčejné diference budeme připisovat k tabulce funkce, viz tab. 3.4.

x	y	Poměrné diference			Obyčejné diference		
		k = 1	k = 2	k = 3	k = 1	k = 2	k = 3
x ₀	y ₀						
x ₁	y ₁	f(x ₀ ; x ₁)			$\Delta^1 f_0$		
x ₂	y ₂	f(x ₁ ; x ₂)	f(x ₀ ; x ₁ ; x ₂)		$\Delta^1 f_1$	$\Delta^2 f_0$	
x ₃	y ₃	f(x ₂ ; x ₃)	f(x ₁ ; x ₂ ; x ₃)	f(x ₀ ; x ₁ ; x ₂ ; x ₃)	$\Delta^1 f_2$	$\Delta^2 f_1$	$\Delta^3 f_0$

Tab. 3.4

Vzhledem k rekurentnímu vytváření diferencí můžeme každou diferenci vyjádřit pomocí hodnot $x_i, x_{i+1}, x_{i+2}, \dots$ a $y_i, y_{i+1}, y_{i+2}, \dots$

Věta 3.2 Pro poměrné a obyčejné diference platí

$$(3.17) \quad f(x_i; x_{i+1}; \dots; x_{i+k}) = \sum_{j=0}^k \frac{y_{i+j}}{(x_{i+j}-x_i)(x_{i+j}-x_{i+1}) \dots (x_{i+j}-x_{i+j-1})(x_{i+j}-x_{i+j+1}) \dots (x_{i+j}-x_{i+k})},$$

$$(3.18) \quad \Delta^k f_i = \sum_{j=0}^k (-1)^j \binom{k}{j} y_{i+k-j}.$$

Důkaz: Vztahy (3.17) a (3.18) lze dokázat úplnou matematickou indukcí. Podrobný zápis důkazu (3.17) je v [5] a důkaz (3.18) je snadný, a proto jej ponecháme na čtenáři.

Věta 3.3 Jestliže tabulka funkce z definice 3.2 je ekvidistantní a provedeme-li transformaci (3.9), pak platí

$$(3.19) \quad f(x_i; x_{i+1}; \dots; x_{i+k}) = \frac{\Delta^k f_i}{k! h^k}.$$

Důkaz úplnou matematickou indukcí je opět snadný, proto jej ponecháme na čtenáři.

Věnujme se již odvození Newtonova interpolačního polynomu $N_n(x)$, který zapíšeme jako speciální úpravu polynomu $L_n(x)$, a to

$$(3.20) \quad N_n(x) = L_0(x) + [L_1(x) - L_0(x)] + \dots + [L_n(x) - L_{n-1}(x)].$$

Polynom $L_k(x) - L_{k-1}(x)$ je polynom k -tého stupně s nulovými body v uzlech x_0, x_1, \dots, x_{k-1} , tedy má tvar

$$L_k(x) - L_{k-1}(x) = A(x - x_0)(x - x_1) \dots (x - x_{k-1}) = A \omega_{k-1}(x).$$

Konstantu A zjistíme dosazením $x = x_k$ a úpravou pomocí (3.7) a vzorců

$$\omega'_k(x_k) = \omega_{k-1}(x_k), \quad \omega'_k(x_j) = (x_j - x_k) \omega'_{k-1}(x_j).$$

Tedy

$$\begin{aligned} A &= \frac{L_k(x_k) - L_{k-1}(x_k)}{\omega'_{k-1}(x_k)} = \frac{y_k}{\omega'_k(x_k)} - \frac{1}{\omega'_{k-1}(x_k)} \sum_{j=0}^{k-1} y_j \frac{\omega'_{k-1}(x_k)}{(x_k - x_j) \omega'_{k-1}(x_j)} = \\ &= \frac{y_k}{\omega'_k(x_k)} + \sum_{j=0}^{k-1} \frac{y_j}{\omega'_k(x_j)} = \sum_{j=0}^k \frac{y_j}{\omega'_k(x_j)}. \end{aligned}$$

Vzhledem k větě 3.2 je zřejmé, že

$$A = f(x_0; x_1; \dots; x_k),$$

tedy po dosazení do (3.20) dostaneme Newtonův interpolační polynom pro neekvidistantní argumenty ve tvaru

$$(3.21) \quad \begin{aligned} N_n(x) &= f(x_0) + (x - x_0)f(x_0; x_1) + (x - x_0)(x - x_1)f(x_0; x_1; x_2) + \dots + \\ &+ (x - x_0)(x - x_1) \dots (x - x_{n-1})f(x_0; x_1; \dots; x_n). \end{aligned}$$

Jsou-li uzlové body ekvidistantní, pak užitím transformace (3.9) dostaneme z (3.21) tvar

$$(3.22) \quad \begin{aligned} N_n(x) &= N_n(x_0 + th) = f(x_0) + \frac{t}{1!} \Delta^1 f_0 + \frac{t(t-1)}{2!} \Delta^2 f_0 + \dots + \\ &+ \frac{t(t-1) \dots (t-n+1)}{n!} \Delta^n f_0, \end{aligned}$$

který se nazývá Newtonův interpolační polynom pro ekvidistantní argumenty.

Koeficienty Newtonových polynomů (3.21), (3.22) jsou čísla na diagonále tabulky 3.4. Proto je můžeme pro danou tabulku vypočítat jednou "provždy" a hodnoty polynomu $N_n(x)$ pak počítat postupným dosazováním. Další výhoda je v tom, že když přidáme

k tabulce další uzel x_{n+1} s hodnotou y_{n+1} , pak se polynomy $N_n(x)$ a $N_{n+1}(x)$ v prvních n členech shodují. Proto se u interpolačního polynomu v Newtonově tvaru pohodlně zvyšuje stupeň, zatímco v Lagrangeově tvaru bychom museli sestrojit vždy celý polynom $L_{n+1}(x)$ znova, nebo užít Nevillova algoritmu.

Příklad 3.3 Sestavme tabulku poměrných differencí pro tabulku funkce z příkladu 3.1. Dále nad uzly x_0, x_1, x_2, x_3 sestavme postupně interpolační polynomy $N_1(x)$, $N_2(x)$, $N_3(x)$ a vypočítejme jejich hodnoty pro $x = 0,5$.

Řešení:

i	x_i	y_i	$k = 1$	$k = 2$	$k = 3$
0	0	1			
1	2	3	1		
2	3	2	-1	-2/3	
3	5	5	3/2	5/6	3/10

$$\begin{aligned} N_1(x) &= 1 + x, \\ N_2(x) &= 1 + x - 2/3x(x-2), \\ N_3(x) &= 1 + x - 2/3x(x-2) + \\ &\quad + 3/10x(x-2)(x-3), \\ N_1(0,5) &= 1,5, \quad N_2(0,5) = 2,0, \\ N_3(0,5) &= 2,5625. \end{aligned}$$

Při některých úlohách potřebujeme sestrojit Newtonův interpolační polynom pro uzly, které v indexování předcházejí uzel x_0 . Označme je $x_0, x_{-1}, \dots, x_{-n}$. Pro takovou soustavu uzlů můžeme podle definice 3.1 zapsat poměrné i obyčejné diference takto:

x	y	Poměrné differenze			Obyčejné differenze		
		$k = 1$	$k = 2$	$k = 3$	$k = 1$	$k = 2$	$k = 3$
x_{-3}	y_{-3}						
x_{-2}	y_{-2}	$f(x_{-3}; x_{-2})$			$\Delta^1 f_{-3}$		
x_{-1}	y_{-1}	$f(x_{-2}; x_{-1})$	$f(x_{-3}; x_{-2}; x_{-1})$		$\Delta^1 f_{-2}$	$\Delta^2 f_{-3}$	
x_0	y_0	$f(x_{-1}; x_0)$	$f(x_{-2}; x_{-1}; x_0)$	$f(x_{-3}; x_{-2}; x_{-1}; x_0)$	$\Delta^1 f_{-1}$	$\Delta^2 f_{-2}$	$\Delta^3 f_{-3}$

Tab. 3.5

Předpokládejme, že pro uzlové body $x_{-k}, x_{-k+1}, \dots, x_0$, kde $k = 0, 1, 2, \dots, n$ máme sestrojeny Lagrangeovy interpolační polynomy $L_k(x)$. Analogickým postupem, kterým jsme z (3.20) dostali (3.21), můžeme odvodit Newtonův interpolační polynom pro uzly $x_0, x_{-1}, \dots, x_{-n}$ s výchozím bodem x_0 ve tvaru

$$\begin{aligned} N_n(x) &= f(x_0) + (x - x_0)f(x_{-1}; x_0) + (x - x_0)(x - x_{-1})f(x_{-2}; x_{-1}; x_0) + \\ (3.23) \quad &+ \dots + (x - x_0)(x - x_{-1}) \dots (x - x_{-n+1})f(x_{-n}; x_{-n+1}; \dots; x_0) \end{aligned}$$

Jestliže v ekvidistantním případě zavedeme transformaci (3.9), ve které bude $t \leq 0$, pak z (3.23) dostaneme podle věty 3.3 $N_n(x)$ ve tvaru

$$(3.24) \quad N_n(x) = N_n(x_0 + th) = f(x_0) + \frac{t}{1!} \Delta^1 f_{-1} + \frac{t(t+1)}{2!} \Delta^2 f_{-2} + \dots + \\ + \frac{t(t+1) \dots (t+n-1)}{n!} \Delta^n f_{-n}.$$

Poměrné i obyčejné diference pro polynomy (3.23) a (3.24) jsou seřazeny v tabulce diferencií, a to v řádku s hodnotami x_0, y_0 (viz tab. 3.5). Polynomy (3.23) a (3.24) nazýváme Newtonovy interpolační polynomy pro interpolaci vzad.

Příklad 3.4 Sestavme Newtonův interpolační polynom $N_3(x)$ pro interpolaci vzad k tabulce hodnot z příkladu 3.3. Označme $x_{-3} = 0, x_{-2} = 2, x_{-1} = 3, x_0 = 5$. Nejdříve vypočítejme $N_3(0,5)$.

Řešení:

i	x_i	y_i	$k = 1$	$k = 2$	$k = 3$
-3	0	1			
-2	2	3	1		
-1	3	2	-1	-2/3	
0	5	5	3/2	5/6	3/10

$$N_3(x) = 5 + (x - 5) \frac{3}{2} + \\ + (x - 5)(x - 3) \frac{5}{6} + \\ + (x - 5)(x - 3)(x - 2) \frac{3}{10} \\ N_3(0,5) = 2,5625$$

3.2.4 Hermitova interpolace

Mějme dánu tabulku T_f spojité funkce $f(x)$ na $a < x < b$. V každém uzlovém bodě x_i nechť jsou ještě dány hodnoty derivací

$$f'(x_i) = y'_i, f''(x_i) = y''_i, \dots, f^{(\alpha_i)}(x_i) = y^{(\alpha_i)}_i.$$

Hermitův interpolační polynom $H_n(x)$ je takový algebraický polynom n-tého stupně (3.5), pro který platí

$$(3.25) \quad H_n(x_i) = y_i, H'_n(x_i) = y'_i, \dots, H_n^{(\alpha_i)}(x_i) = y^{(\alpha_i)}_i, \quad i = 0, 1, \dots, m,$$

kde $n = m + \alpha_0 + \alpha_1 + \dots + \alpha_m$, tedy $m < n$.

Z rovnice (3.25) bychom mohli sice počítat koeficienty Hermitova polynomu, ale bylo by to příliš těžkopádné. Sestrojení $H_n(x)$ je snadnější, když užijeme Lagrangeova interpolačního polynomu $L_m(x)$, který je sestrojen nad uzly x_0, x_1, \dots, x_m . Rozdíl $H_n(x) - L_m(x)$ je polynom nejvyšše n-tého stupně s nulovými body v uzlech x_0, x_1, \dots, x_m , tedy

$$H_n(x) - L_m(x) = \omega_m(x) \cdot H_{n-m-1}(x),$$

kde $\omega_m(x) = (x - x_0)(x - x_1) \dots (x - x_m)$ a $H_{n-m-1}(x)$ je polynom maximálně $(n - m - 1)$ -ho stupně. Potom

$$H_n(x) = L_m(x) + \omega_m(x) \cdot H_{n-m-1}(x)$$

je polynom, který splňuje první z podmínek (3.25). Stačí tedy sestrojít polynom $H_{n-m-1}(x)$ tak, aby byly splněny i ostatní podmínky v (3.25), tj. podmínky

$$y'_i = H'_n(x_i) = L'_m(x_i) + \omega'_m(x_i) \cdot H_{n-m-1}(x_i),$$

$$y''_i = H''_n(x_i) = L''_m(x_i) + \omega''_m(x_i) \cdot H_{n-m-1}(x_i) + 2\omega'_m(x_i) \cdot H'_m(x_i),$$

$$\dots$$

$$y^{(\alpha_i)}_i = H^{(\alpha_i)}_n(x_i) = L^{(\alpha_i)}_m(x_i) + \omega^{(\alpha_i)}_m(x_i) \cdot H_{n-m-1}(x_i) + \dots,$$

pro $i = 0, 1, 2, \dots, m$. Úpravou dostaneme rovnice

$$H_{n-m-1}(x_i) = (y'_i - L'_m(x_i)) / \omega'_m(x_i) = z_i,$$

$$H''_{n-m-1}(x_i) = (y''_i - L''_m(x_i) - \omega''_m(x_i) \cdot z_i) / (2\omega'_m(x_i)) = z'_i,$$

$$\dots$$

$$H^{(\alpha_i-1)}_{n-m-1}(x_i) = (y^{(\alpha_i)}_i - L^{(\alpha_i)}_m(x_i) - \dots) / (\alpha_i \cdot \omega'_m(x_i)) = z^{(\alpha_i-1)}_i,$$

pro sestrojení $H_{n-m-1}(x)$, které jsou analogické k rovnicím (3.25) pro nalezení $H_n(x)$, tedy budeme postupovat zcela obdobně.

Po dostatečném počtu kroků zbude sestrojit Lagrangeův interpolační polynom na těch uzlových bodech, ve kterých byla podmínkami (3.25) předepsána nejvyšší derivace.

Podrobněji se Hermitovou interpolací zabývat nebudeme, protože přesahuje rámec tohoto skripta.

Příklad 3.5 Sestrojme Hermitův interpolační polynom pro tabulku 3.6 hodnot a derivací spojité funkce se dvěma spojitými derivacemi.

Rешení:

i	0	1	2	3
x	-1	0	1	2
y _i	1	0	1	16
y' _i	-4	0	4	-
y'' _i	-	0	12	-

$$n = m + \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 = \\ = 3 + 1 + 2 + 2 + 0 = 8.$$

Sestrojíme-li $L_3(x)$ nad uzly x_0, x_1, x_2, x_3 , pak dostaneme

Tab. 3.6

$$H_8(x) = L_3(x) + \omega_3(x) \cdot H_4(x) = \\ = 2x^3 + x^2 - 2x + (x^4 - 2x^3 - x^2 + 2x) \cdot H_4(x),$$

$$H'_8(x_i) = 6x_i^2 + 2x_i - 2 + (4x_i^3 - 6x_i^2 - 2x_i + 2) \cdot H_4(x_i) = y'_i$$

a

$$H''_8(x_i) = 12x_i^2 + 2 + (12x_i^2 - 12x_i - 2) \cdot H_4(x_i) + \\ + 2(4x_i^3 - 6x_i^2 - 2x_i + 2) \cdot H'_4(x_i) = y''_i.$$

Podmínky pro sestrojení polynomu $H_4(x)$ vypočítáme tedy z rovnic

$$H_4(x_i) = \frac{y'_i - (6x_i^2 + 2x_i - 2)}{4x_i^3 - 6x_i^2 - 2x_i + 2} = z_i,$$

$$H'_4(x_i) = \frac{y''_i - (12x_i + 2) - (12x_i^2 - 12x_i - 2)z_i}{2(4x_i^3 - 6x_i^2 - 2x_i + 2)} = z'_i,$$

a výsledky sestavíme do tabulky 3.7.

i	0	1	2
x	-1	0	1
z_i	1	1	1
z'_i	-	0	0

Tab. 3.7

V dalším kroku sestrojíme $L_2(x)$ nad uzly x_0, x_1, x_2 a dostaneme

$$H_4(x) = L_2(x) + \omega_2(x) \cdot H_1(x) =$$

$$= 1 + (x^3 - x) \cdot H_1(x)$$

a

$$H'_4(x_i) = (3x_i^2 - 1) \cdot H_1(x_i) = z'_i.$$

Pro sestrojení $H_1(x)$ jsme dostali dvě podmínky $H_1(0) = 0$ a $H_1(1) = 0$, tedy $H_1(x) \equiv 0$, potom $H_4(x) = 1$, čili hledaný Hermitův polynom je

$$H_8(x) = 2x^3 + x^2 - 2x + (x^4 - 2x^3 - x^2 + 2x) \cdot 1 = x^4.$$

3.2.5 Přesnost a konvergence interpolační approximace

Předpokládejme, že na intervalu $\langle a, b \rangle$ jsme sestrojili ke spojité funkci $f(x)$ interpolační polynom $L_n(x)$, tedy platí $L_n(x_i) = f(x_i)$ pro $x_0, x_1, \dots, x_n \in \langle a, b \rangle$. V tomto odstavci se budeme nejdříve zabývat vyjádřením (resp. odhadem) rozdílu $f(x) - L_n(x)$ pro $x \in \langle a, b \rangle$ a dále otázkou konvergence polynomu $L_n(x)$ k funkci $f(x)$, jestliže se na intervalu $\langle a, b \rangle$ zvyšuje počet uzlových bodů, tzn. když se zvyšuje stupeň polynomu $L_n(x)$. Abychom mohli velikost chyby, případně otázku konvergence zkoumat, musíme požadovat, aby funkce $f(x)$ byla dostatečně hladká, tj. měla dostatečný počet spojitých derivací.

Věta 3.4 Nechť $f(x)$ má na intervalu $\langle a, b \rangle$ spojité derivace až do n -tého řádu včetně a nechť $f^{(n+1)}(x)$ existuje ve všech bodech intervalu (a, b) . Dále nechť $L_n(x)$ je interpolační polynom pro funkci $f(x)$ a uzlové body $x_0, x_1, \dots, x_n \in \langle a, b \rangle$. Potom platí

$$(3.26) \quad R(x) = f(x) - L_n(x) = \omega_n(x) \frac{f^{(n+1)}(\xi)}{(n+1)!}, \quad x \in \langle a, b \rangle,$$

kde $\xi \in (\min(x, x_0, \dots, x_n), \max(x, x_0, \dots, x_n))$.

Důkaz založený na vícenásobném použití Rollovy věty najde čtenář např. v [5], [16], [19], [23].

Z (3.26) vidíme, že chyba interpolace závisí kromě x na $\omega_n(x)$, tj. na rozložení uzlů, na stupni interpolačního polynomu a na hladkosti interpolované funkce. Vzorec (3.26) není však pro výpočet chyby použitelný, protože ξ neznáme. Z tohoto

důvodu se musíme spokojit jenom odhadem chyby. Jestliže existuje konstanta M_{n+1} taková, že $|f^{(n+1)}(x)| \leq M_{n+1}$ pro $x \in (a, b)$, potom pro všechna $x \in [a, b]$ je

$$(3.26) \quad |R(x)| \leq \frac{M_{n+1}}{(n+1)!} \max_{x \in [a,b]} |\omega_n(x)|.$$

Chybu interpolace, jak je vidět z (3.25) a z (3.26), lze ovlivnit různým rozmištěním uzlových bodů na intervalu $[a, b]$. Jestliže ztotožníme uzlové body s kořeny tzv. Čebyševova polynomu $(n+1)$ -ního stupně (viz [9], str. 591), tj. s body

$$(3.27) \quad x_i = \frac{a+b}{2} - \frac{b-a}{2} \cos\left(\frac{2i+1}{2(n+1)}\pi\right), \quad i = 0, 1, \dots, n,$$

pak tím minimalizujeme $\max_{x \in [a,b]} |\omega_n(x)|$. V mnoha případech však nemůžeme toto optimální rozložení provést, protože dvojice x_i, y_i pro $i = 0, 1, 2, \dots, n$ jsou dány např. z měření nebo předcházejících výpočtů a při interpolaci je nemůžeme ovlivnit.

Nyní ještě několik slov o konvergenci interpolačního polynomu $L_n(x)$ při $n \rightarrow \infty$ k interpolované funkci $f(x)$ na intervalu $[a, b]$. Předpokládejme, že na intervalu $[a, b]$ postupně volíme množiny uzlových bodů $U_n = \{x_0^{(n)}, x_1^{(n)}, \dots, x_n^{(n)}\} \subset [a, b]$ pro $n = 1, 2, \dots$, a to tak, že pro $n \rightarrow \infty$

$$\max_{i=1,2,\dots,n} |x_i^{(n)} - x_{i-1}^{(n)}| \rightarrow 0.$$

Ke každé množině uzlů U_n a k hodnotám interpolované funkce y_0, y_1, \dots, y_n se strojíme interpolační polynom $L_n(x)$. Zajímá nás, zdali bude pro $n \rightarrow \infty$ nějakým způsobem $L_n(x)$ konvergovat k $f(x)$ na intervalu $[a, b]$. Na tuto otázku nelze dát bohužel jednoznačně kladnou odpověď, jak ihned objasníme.

Pro libovolné zhušťování uzlů existuje vždy spojitá funkce tak, že její interpolační polynomy $L_n(x)$ k ní nekonvergují stejnou rychlostí na $[a, b]$. Volíme-li uzly ekvidistantní, pak např. pro \sqrt{x} na intervalu $[0, 1]$ nekonverguje interpolační proces ani bodově na celém intervalu $[0, 1]$, nebo pro $|x|$ na $[-1, 1]$ nekonverguje $L_n(x)$ v žádném vnitřním bodě s výjimkou 0. Navíc interpolace s ekvidistantními uzly je pro velká n špatně podmíněná úloha, tj. malé chyby ve vstupních datech se projeví velkou chybou v hodnotě interpolačního polynomu.

Podrobněji se o konvergenci interpolačních polynomů přeše např. v [5].

3.3 Interpolace spline-funkcemi

Zvyšování přesnosti interpolační approximace funkce na intervalu $[a, b]$ zvyšováním stupně interpolačního polynomu má značné nevýhody a je i problematické, jak jsme naznačili v odstavci 3.2.5. Především se komplikuje výpočet hodnot interpolačního polynomu a pak se polynomy vysokých stupňů značně "vlní".

Tyto nevýhody lze odstranit tím, že rozdělíme interval $\langle a, b \rangle$ body $a = x_0 < x_1 < \dots < x_n = b$ na subintervaly a na každém z nich budeme funkci $f(x)$ interpolovat jiným polynomem. Budou-li subintervaly "dostatečně" malé, můžeme užít polynomů nízkých stupňů.

Interpolační funkce k $f(x)$ na $\langle a, b \rangle$ je pak po částech polynomiální funkce. Tento druh interpolace se s oblibou realizuje pomocí tzv. kubických spline-funkcí.

Definice 3.2 Nechť interval $\langle a, b \rangle$ je rozdělen body x_0, x_1, \dots, x_n tak, že $a = x_0 < x_1 < \dots < x_n = b$. Kubickou spline-funkcí $S(x)$ (jednoduše S-funkcí) nazveme každou funkci, pro kterou platí:

- $S(x) \in C^2 \langle a, b \rangle$, tzn. že $S(x)$ je spojitá na intervalu $\langle a, b \rangle$ spolu se svojí první a druhou derivací,
- $S(x)$ je na každém subintervalu $\langle x_{i-1}, x_i \rangle$, $i = 1, 2, \dots, n$, vyjádřena polynomem 3. stupně.

Kubická spline-funkce se tedy skládá z kubických polynomů definovaných nad jednotlivými subintervaly, které jsou spojené v bodech x_1, x_2, \dots, x_{n-1} tak, aby $S(x) \in C^2 \langle a, b \rangle$.

Definice 3.3 Nechť v bodech x_0, x_1, \dots, x_n z definice 3.2 jsou zadány funkční hodnoty y_0, y_1, \dots, y_n spojité funkce $f(x)$ na intervalu $\langle a, b \rangle$.

Potom každá kubická spline-funkce, pro kterou platí

$$(3.28) \quad S(x_i) = y_i, \quad i = 0, 1, \dots, n,$$

se nazývá interpolační kubickou spline-funkcí (jednoduše interpolační S-funkcí) k funkci $f(x)$ na intervalu $\langle a, b \rangle$.

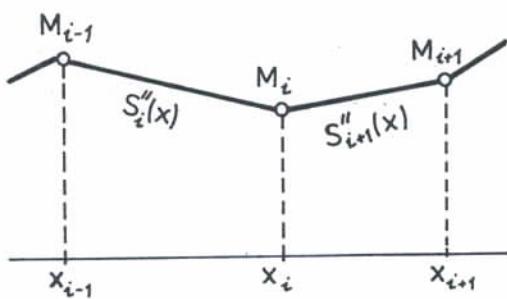
3.3.1 Konstrukce interpolační spline-funkce

Interpolační S-funkce není však podmínkami (3.28) určena jednoznačně. K jednoznačnému určení S-funkce potřebujeme celkem $4n$ konstant, neboť na každém subintervalu $\langle x_{i-1}, x_i \rangle$ je polynom 3. stupně definován čtyřmi konstantami. K určení těchto konstant máme $n + 1$ podmínek (3.28) a dále $3(n - 1)$ podmínek, které plynou ze spojitosti S-funkce a její 1. a 2. derivace v bodech x_1, x_2, \dots, x_{n-1} . Celkově máme $4n - 2$ podmínek k určení $4n$ konstant. Zbyvající dvě podmínky volíme obvykle tak, že v krajních bodech $a = x_0, b = x_n$ předepisujeme hodnoty první nebo druhé derivace interpolační S-funkce. Těmto dvěma podmínkám říkáme okrajové podmínky.

Nyní přistoupíme ke konstrukci interpolační S-funkce, která splňuje jisté okrajové podmínky. Z definice 3.2 plyne, že na každém subintervalu $\langle x_{i-1}, x_i \rangle$, $i = 1, 2, \dots, n$, má spline-funkce $S(x)$ tvar polynomu 3. stupně

$$(3.29) \quad S(x) = S_i(x) = a_i + b_i x + c_i x^2 + d_i x^3,$$

a tedy $S_i'(x)$ je polynom 1. stupně. Protože pro hledanou S-funkci požadujeme spojitost $S''(x)$ na $\langle a, b \rangle$, musí platit



Obr. 3.1

K této soustavě momentů přidáme hodnoty momentů v bodech x_0, x_n , tedy $M_0 = S''(x_0)$, $M_n = S''(x_n)$. V dalším ukážeme, že hledaná interpolační S-funkce je svými momenty

$$(3.30) \quad M_i, \text{ kde } M_i = S''(x_i) \text{ pro } i = 0, 1, 2, \dots, n,$$

plně určena. Podle obr. 3.1 snadno odvodíme, že pro $i = 1, 2, \dots, n$ je

$$(3.31) \quad S''_i(x) = M_{i-1} \frac{(x_i - x)}{h_i} + M_i \frac{(x - x_{i-1})}{h_i}, \quad x \in \langle x_{i-1}, x_i \rangle,$$

kde $h_i = x_i - x_{i-1}$. Po dvojí integraci dostaneme

$$S_i(x) = M_{i-1} \frac{(x_i - x)^3}{6h_i} + M_i \frac{(x - x_{i-1})^3}{6h_i} + A_i(x_i - x) + B_i(x - x_{i-1}),$$

kde $A_i(x_i - x) + B_i(x - x_{i-1})$ je vhodné vyjádření integračního polynomu 1. stupně. Konstanty A_i a B_i se snadno vypočítají na základě podmínek

$$y_{i-1} = S_i(x_{i-1}) = M_{i-1}h_i^2/6 + A_i h_i,$$

$$y_i = S_i(x_i) = M_i h_i^2/6 + B_i h_i.$$

Potom hledaná interpolační S-funkce $S(x)$ má tvar

$$(3.32) \quad S(x) = S_i(x) = M_{i-1} \frac{(x_i - x)^3}{6h_i} + M_i \frac{(x - x_{i-1})^3}{6h_i} + \\ + \left(y_{i-1} - \frac{M_{i-1}h_i^2}{6} \right) \frac{x_i - x}{h_i} + \left(y_i - \frac{M_i h_i^2}{6} \right) \frac{x - x_{i-1}}{h_i},$$

pro $x \in \langle x_{i-1}, x_i \rangle$, $i = 1, 2, \dots, n$ a je jednoznačně definované momenty M_0, M_1, \dots, M_n (jejich hodnoty ale ještě neznáme).

Funkce $S(x)$ z (3.32) nesplňuje doposud podmínsku spojitosti $S'(x)$. Abychom toto splnili, musí pro $i = 1, 2, \dots, n-1$ platit

$$\lim_{x \rightarrow x_{i-1}^-} S'_i(x) = \lim_{x \rightarrow x_i^+} S'_{i+1}(x),$$

čili

$$\frac{y_i - y_{i-1}}{h_i} + M_{i-1} \frac{h_i}{6} + M_i \frac{h_i}{3} = \frac{y_{i+1} - y_i}{h_{i+1}} - M_i \frac{h_{i+1}}{3} - M_{i+1} \frac{h_{i+1}}{6},$$

a tím dostaneme soustavu n rovnic

$$(3.33) \quad \frac{h_i}{h_i + h_{i+1}} M_{i-1} + 2M_i + \frac{h_{i+1}}{h_i + h_{i+1}} = \frac{6}{h_i + h_{i+1}} \left(\frac{y_{i+1} - y_i}{h_{i+1}} - \frac{y_i - y_{i-1}}{h_i} \right)$$

o $n + 1$ neznámých M_0, M_1, \dots, M_n . Aby soustava (3.33) byla čtvercová, přidáme k ní ještě dvě lineární rovnice tak, abychom neporušili její pásový charakter, a ty nám umožní splnit okrajové podmínky. Tuto soustavu $n + 1$ rovnic o $n + 1$ neznámých zapíšeme ve tvaru

$$(3.34) \quad \begin{aligned} 2M_0 + b_0 M_1 &= c_0 \\ a_1 M_0 + 2M_1 + b_1 M_2 &= c_1 \\ a_2 M_1 + 2M_2 + b_2 M_3 &= c_2 \\ \cdot &\quad \cdot & \cdot \\ \cdot &\quad \cdot & \cdot \\ a_{n-1} M_{n-2} + 2M_{n-1} + b_{n-1} M_n &= c_{n-1} \\ a_n M_{n-1} + 2M_n &= c_n , \end{aligned}$$

kde koeficienty první a poslední rovnice, tj. b_0, a_n, c_0, c_n , slouží k vyjádření okrajových podmínek a další koeficienty vycházejí z (3.33), tj.

$$(3.35) \quad a_i = \frac{h_i}{h_i + h_{i+1}}, \quad b_i = \frac{h_{i+1}}{h_i + h_{i+1}}, \quad c_i = \frac{6}{h_i + h_{i+1}} \left(\frac{y_{i+1} - y_i}{h_{i+1}} - \frac{y_i - y_{i-1}}{h_i} \right),$$

pro $i = 1, 2, \dots, n - 1$.

Ukažme volby b_0, c_0, a_n, c_n , které zajistí splnění některých typických okrajových podmínek:

I. Požadujeme $S'(a) = y'_0, S'(b) = y'_n$, potom

$$\lim_{x \rightarrow a_+} S'(x) = y'_0 \quad \text{a} \quad \lim_{x \rightarrow b_-} S'(x) = y'_n ,$$

z čehož dostaneme tvar první a poslední rovnice z (3.34)

$$(3.36) \quad \begin{aligned} 2M_0 + M_1 &= \frac{6}{h_1} \left(\frac{y_1 - y_0}{h_1} - y'_0 \right), \\ M_{n-1} + 2M_n &= \frac{6}{h_n} \left(y'_n - \frac{y_n - y_{n-1}}{h_n} \right), \end{aligned}$$

tedy $b_0 = a_n = 1$ a c_0, c_n jsou pravé strany v (3.36).

II. Požadujeme $S''(a) = y''_0, S''(b) = y''_n$, pak podle (3.30) je $b_0 = a_n = 0$ a $c_0 = 2y''_0, c_n = 2y''_n$, tedy první a poslední rovnice v (3.34) mají tvar

$$(3.37) \quad 2M_0 = 2y''_0, \quad 2M_n = 2y''_n .$$

Soustava (3.34) je třídiagonální a pro $|b_0| < 2$, $|a_n| < 2$ má dominantní diagonálu, protože $a_i + b_i = 1$. Taková soustava je jednoznačně řešitelná a její řešení najdeme např. Gaussovým eliminacním postupem.

Jak tedy postupujeme při výpočtu interpolační S-funkce pro dané hodnoty y_0, y_1, \dots, y_n v bodech x_0, x_1, \dots, x_n :

1. Podle (3.35) vypočítáme koeficienty a_i, b_i, c_i pro $i = 1, 2, \dots, n - 1$.
2. Podle okrajových podmínek vypočítáme b_0, a_n, c_0, c_n .
3. Ze soustavy (3.34) nalezneme M_0, M_1, \dots, M_n a dosadíme je do (3.32).

Příklad 3.6 Mějme dánu tabulku funkce $f(x)$ z příkladu 3.1 a předpokládejme, že $\langle a, b \rangle = \langle 0, 5 \rangle$. Sestrojme k ní interpolační kubickou spline-funkci $S(x)$, pro kterou bude platit $S''(0) = S''(5) = 0$.

Rешení: Podle (3.35) dostaneme $a_1 = 2/3, a_2 = 1/3, b_1 = 1/3, b_2 = 2/3, c_1 = -4, c_2 = 5$. Vzhledem k okrajovým podmínkám dostaneme soustavu

$$\begin{aligned} 2M_0 &= 0 \\ 2/3M_0 + 2M_1 + 1/3M_2 &= -4 \\ 1/3M_1 + 2M_2 + 2/3M_3 &= 5 \\ 2M_3 &= 0, \end{aligned}$$

ze které vypočítáme $M_0 = 0, M_1 = -87/35, M_2 = 102/35, M_3 = 0$. Hledaná S-funkce má podle (3.32) tvar

$$S(x) = \begin{cases} -\frac{87}{420}x^3 + \frac{768}{420}x + 1, & x \in \langle 0, 2 \rangle, \\ -\frac{87}{210}(3-x)^3 + \frac{102}{210}(x-2)^3 + \frac{717}{210}(3-x) + \frac{318}{210}(x-2), & x \in \langle 2, 3 \rangle, \\ \frac{102}{420}(5-x)^3 + \frac{12}{420}(5-x) + \frac{5}{2}(x-3), & x \in \langle 3, 5 \rangle. \end{cases}$$

Ověřte si, že $S(x), S'(x)$ i $S''(x)$ jsou na $\langle 0, 5 \rangle$ spojité.

3.3.2 Konvergence spline-interpolace

Ještě několik slov o konvergenci interpolační S-funkce. V odstavci 3.2.5 jsme upozornili na některé nepříjemnosti při interpolaci pomocí polynomů. Naproti tomu se dá o interpolačních S-funkcích dokázat, že za jistých podmínek hladkosti interpolované funkce $f(x)$ konverguje interpolační S-funkce sestrojené pro zjemňující se dělení intervalu $\langle a, b \rangle$ k funkci $f(x)$ stejnomořně. Přesněji tuto vlastnost popisuje následující věta.

Věta 3.5 Nechť $f(x) \in C^q \langle a, b \rangle$, $q = 0, 1, 2, 3, 4$. Dále nechť interval $\langle a, b \rangle$ je rozdělen na subintervaly délky h_i , $i = 1, 2, \dots, n$, kde $\max_i h_i = h$. Zavedme ještě konstantu K , pro kterou $K \geq \max_i h_i / \min_i h_i$. Je-li $S(x)$ interpolač-

mí spline-funkce splňující okrajové podmínky I, resp. II, pak pro všechna $x \in \langle a, b \rangle$ a $p=0, 1, 2, 3$ platí

$$(3.38) \quad |f^{(p)}(x) - S^{(p)}(x)| \leq CK h^{q-p},$$

kde C je konstanta, která nezávisí na x ani na způsobu dělení intervalu $\langle a, b \rangle$.

Důkaz věty je uveden např. v [25], [27] a [2].

Z této věty plyne, že rychlosť konvergencie interpolačnej S-funkcie nebo jejich derivácií je tím väčší, čím má funkcia $f(x)$ vyšší stupeň hladkosti.

Podrobnejši sa může čtenář dočíti o spline-funkciach ve [2].

3.4 Aproximace metodou nejmenších čtverců (MNČ) - diskrétní případ

Mějme dánu tabulku T_f spojité funkce $f(x)$ na intervalu $\langle a, b \rangle$ a o funkčních hodnotách předpokládejme, že jsou zatíženy nějakou chybou (jedná se např. o měřené veličiny). Nebylo by tedy "rozumné" approximovat tuto funkci interpolačním polynomem nebo interpolační S-funkcií, protože při interpolaci se požaduje plná shoda interpolační i interpolované funkce v bodech tabulky.

Aproximační funkci budeme nadále hledat ve tvaru zobecněného polynomu (3.2), tedy $\phi_n(x) = c_0 g_0 + c_1 g_1 + \dots + c_n g_n$, avšak koeficienty c_0, c_1, \dots, c_n chceme nalézt tak, aby rozdíl dvou srovnatelných vektorů (y_0, y_1, \dots, y_m) a $(\phi_n(x_0), \phi_n(x_1), \dots, \phi_n(x_m))$ měl minimální eukleidovskou normu, čili aby tzv. střední kvadratická odchylka

$$(3.39) \quad \sigma_m = \sqrt{\sum_{i=0}^m [f(x_i) - \phi_n(x_i)]^2}$$

byla minimální na množině všech možných koeficientů c_0, c_1, \dots, c_n , při pevně zvolené soustavě základních funkcí g_0, g_1, \dots, g_n . To je základní charakteristika approximace MNČ v diskrétním případě. Takto získaný zobecněný polynom n -tého stupně se nazývá polynomem nejlepší diskrétní approximace. Approximace MNČ vyrovnává vliv náhodných chyb v daných (měřených) veličinách a praxe ukazuje, že polynom nejlepší approximace "mnohem lépe" nahrazuje reálnou funkci $f(x)$ než interpolační polynom.

3.4.1 Normální rovnice

Koeficienty zobecněného polynomu tedy hledáme tak, aby odchylka (3.39) byla minimální, nebo aby její kvadrát

$$(3.40) \quad \sigma_m^2 = \sum_{i=0}^m [f(x_i) - \sum_{k=0}^n c_k g_k(x_i)]^2$$

byl minimální, což je jistě ekvivalentní úloha. Protože σ_m^2 je funkcií $n+1$ proměnných c_0, c_1, \dots, c_n , pak σ_m^2 může dosáhnout svého minima jenom pro takové hodnoty proměnných, pro které platí

$$(3.41) \quad \frac{\partial \varphi_m^2(c_0, c_1, \dots, c_n)}{\partial c_k} = 0, \quad \text{pro } k = 0, 1, \dots, n.$$

Provedeme-li naznačené derivace, pak z (3.41) dostaneme soustavu

$$(3.42) \quad \sum_{j=0}^n \left[\sum_{i=0}^m g_k(x_i) g_j(x_i) \right] c_j = \sum_{i=0}^m f(x_i) g_k(x_i), \quad k = 0, 1, \dots, n,$$

která se nazývá soustavou normálních rovnic, a je jednoznačně řešitelná, když systém základních funkcí g_0, g_1, \dots, g_n je diskrétně lineárně nezávislý, tzn. když aritmetické vektory

$$(g_k(x_0), g_k(x_1), \dots, g_k(x_m)), \quad k = 0, 1, \dots, n,$$

jsou lineárně nezávislé. Vyšetřením výrazu $\varphi_m^2(\tilde{c}_0 + \Delta c_0, \tilde{c}_1 + \Delta c_1, \dots, \tilde{c}_n + \Delta c_n)$, kde $\tilde{c}_0, \tilde{c}_1, \dots, \tilde{c}_n$ je řešení soustavy (3.42), zjistíme, že toto řešení realizuje minimum výrazu (3.40).

Soustavu (3.42) zapíšeme přehledněji pomocí tzv. diskrétního skalárního součinu, který se pro funkce $\varphi(x), \Psi(x)$ označuje $(\varphi, \Psi)_m$ a definuje vztahem

$$(3.43) \quad (\varphi, \Psi)_m = \sum_{i=0}^m \varphi(x_i) \cdot \Psi(x_i),$$

kde $\varphi(x_i), \Psi(x_i)$ pro $i = 0, 1, \dots, m$ jsou hodnoty funkcí φ, Ψ v bodech x_0, x_1, \dots, x_m . Soustava (3.42) má tedy tvar

$$(3.44) \quad \sum_{j=0}^n (g_k, g_j)_m c_j = (f, g_k)_m, \quad k = 0, 1, \dots, n.$$

Tato soustava je symetrická a dokonce pozitivně definitní (viz např. [5]). Její řešení můžeme vypočítat např. Gaussovou eliminační metodou nebo LU-rozkladem apod.

Nalezené koeficienty c_0, c_1, \dots, c_n dosadíme do zobecněného polynomu (3.2) a tím získáme hledaný polynom nejlepší (diskrétní) approximace.

3.4.2 Volba základních funkcí

Klasická volba základních funkcí je $g_k(x) = x^k$ pro $k = 0, 1, \dots, n$. Soustava (3.44) pak má tvar

$$(3.45) \quad \sum_{j=0}^n \left[\sum_{i=0}^m x_i^{k+j} \right] c_j = \sum_{i=0}^m x_i^k f(x_i), \quad k = 0, 1, \dots, n.$$

Systém $1, x, x^2, \dots, x^n$ je pro množinu $X = \{x_0, x_1, \dots, x_m\}$ diskrétně lineárně nezávislý, tedy soustava (3.45) je jednoznačně řešitelná.

x_i	-1	-0,5	0	0,5	1
y_i	0	-1	0	1	0

Tab. 3.8

Příklad 3.7 K tabulce 3.8 funkce $f(x)$ na intervalu $(-1, 1)$ sestrojme MNČ polynom nejlepší approximace 3. stupně, jestliže za základní funkce zvolíme $1, x, x^2, x^3$.

Řešení: Pro základní funkce $1, x, x^2, x^3$ má aproximační polynom $\phi_3(x)$ tvar

$$\phi_3(x) = c_0 + c_1x + c_2x^2 + c_3x^3,$$

kde c_0, c_1, c_2, c_3 určíme podle (3.45) ze soustavy

$$5c_0 + 2,5c_2 = 0, \quad 2,5c_1 + 2,125c_3 = 1,$$

$$2,5c_0 + 2,125c_2 = 0, \quad 2,125c_1 + 2,03125c_3 = 0,25,$$

tedy $c_0 = c_2 = 0$ a $c_1 = -c_3 = 8/3$. Hledaný polynom nejlepší approximace má tvar

$$\phi_3(x) = 8/3x - 8/3x^3 = 8/3(x - x^3).$$

Soustava (3.45) je však pro větší n špatně podmíněná, tedy je řešení citlivé na zaokrouhlovací chyby a chyby vstupních dat. Zvláště při ekvidistantních tabulkách již při $n \approx 9$ je podmíněnost matice (3.45) tak špatná, že získané výsledky jsou velice problematické. Tato skutečnost svědčí proti užívání základního systému $g_i(x) = x^i$ s výjimkou velmi malých hodnot n (viz [23]).

Nastíněné nepříjemnosti odpadají, jestliže za základní funkce vezmeme tzv. ortogonální polynomy na diskrétní množině $X = \{x_0, x_1, \dots, x_n\}$.

Definice 3.4 Systém funkcí g_0, g_1, \dots, g_n nazýváme ortogonální (resp. ortonormální) na diskrétní množině $X = \{x_0, x_1, \dots, x_m\}$, jestliže

$$(3.46) \quad (g_k, g_j)_m \begin{cases} = 0 & \text{pro } k \neq j, \\ \neq 0 \text{ (resp. } = 1) & \text{pro } k = j. \end{cases}$$

Je-li základní systém g_0, g_1, \dots, g_n ortogonální na diskrétní množině X , pak je na ní diskrétně lineárně nezávislý (důkaz viz [10]), tedy příslušná soustava normálních rovnic (3.44) je jednoznačně řešitelná a má v důsledku (3.46) diagonální tvar

$$(3.47) \quad (g_k, g_k)_m c_k = (f, g_k)_m, \quad k = 0, 1, 2, \dots, n.$$

Hledané koeficienty se vypočítají podle vzorce

$$(3.48) \quad c_k = (f, g_k)_m / (g_k, g_k)_m, \quad k = 0, 1, \dots, n.$$

Je-li systém g_0, g_1, \dots, g_n ortonormální, pak (3.44) má dokonce jednotkovou diagonálu, tedy

$$(3.49) \quad c_k = (f, g_k)_m.$$

Zobecněný polynom nejlepší approximace pro ortogonální (resp. ortonormální) základní systém má pro zadanou tabulku funkce tvar

$$(3.50) \quad \phi_n(x) = \sum_{k=0}^n [(f, g_k)_m / (g_k, g_k)_m] \cdot g_k(x),$$

$$\text{resp. } \phi_n(x) = \sum_{k=0}^n (f, g_k)_m \cdot g_k(x),$$

kde

$$(f, g_k)_m = \sum_{i=0}^m f(x_i) \cdot g_k(x_i), \quad (g_k, g_k)_m = \sum_{i=0}^m g_k^2(x_i).$$

3.4.3 Čebyševovy polynomy

Mějme diskrétní množinu $T = \{0, 1, \dots, m\}$. Na ní tvoří tzv. Čebyševovy polynomy $P_{km}(t)$, $k = 0, 1, 2, \dots, n$ ($n \leq m$), kde k je stupeň polynomu, ortogonální systém. Polynomy $P_{km}(t)$ jsou definovány vzorcem

$$(3.51) \quad P_{km}(t) = \sum_{\ell=0}^k (-1)^\ell \left(\frac{k}{\ell}\right) \binom{k+\ell}{\ell} \frac{\omega(t, \ell)}{\omega(m, \ell)}, \quad k = 0, 1, \dots, n.$$

kde $\omega(t, \ell) = t(t-1) \dots (t-\ell+1)$,

$$\omega(m, \ell) = m(m-1) \dots (m-\ell+1)$$

a

$$\omega(t, 0) = \omega(m, 0) = 1.$$

Vzhledem k ortogonalitě polynomů (3.51) na množině T a vzhledem k definici diskrétního skalárního součinu, viz (3.43), platí pro $j, k = 0, 1, \dots, n = m$

$$(3.52) \quad (P_{jm}, P_{km}) = \sum_{t=0}^m P_{jm}(t) \cdot P_{km}(t) \begin{cases} = 0, & \text{při } j \neq k, \\ = \frac{\omega(m+k+1, k+1)}{(2k+1)\omega(m, k)}, & \text{při } j = k. \end{cases}$$

Ověřte si, že podle (3.51) a (3.52) platí pro $m = 4$:

$$(3.53) \quad \begin{aligned} P_{04}(t) &= 1, & (P_{04}, P_{04}) &= 5, \\ P_{14}(t) &= 1 - t/2, & (P_{14}, P_{14}) &= 2,5, \\ P_{24}(t) &= 1 - 3t/2 + t(t-1)/2, & (P_{24}, P_{24}) &= 3,5, \\ P_{34}(t) &= 1 - 3t + 5t(t-1)/2 - 5t(t-1)(t-2)/6, & (P_{34}, P_{34}) &= 10. \end{aligned}$$

Zbývá ukázat, jak se Čebyševovy polynomy při MNČ používají. Máme nalézt polynom nejlepší approximace $\phi_n(x)$ pro spojitou funkci $f(x)$ na intervalu $\langle a, b \rangle$, je-li pro ni dána tabulka T_f s ekvidistantními argumenty. Lineární transformací (3.9) převedeme množinu

$$X = \{x_0, x_1, \dots, x_m\} \text{ na } T = \{0, 1, \dots, m\}$$

a tabulkou T_f , viz (3.1), na tabulku

$$(3.54) \quad \{(t, y(t)) : y(t) = f(x_0 + th), \text{ pro } t = 0, 1, \dots, m\}.$$

Tím jsme zadanou úlohu převedli na nalezení polynomu nejlepší approximace $\phi_n(t)$ ($n \leq m$) pro funkci $f(x_0 + th)$ s hodnotami danými tabulkou (3.54). Zvolíme-li za základní funkce MNČ Čebyševovy polynomy (3.51), potom

$$(3.55) \quad \phi_n(t) = c_0 P_{0m}(t) + c_1 P_{1m}(t) + \dots + c_n P_{nm}(t).$$

Protože polynomy $P_{km}(t)$, $k = 0, 1, \dots, n$ jsou na množině T ortogonální, pak soustava normálních rovnic má tvar (3.47) a řešení je dáno vzorcem (3.48).

Vezmeme-li v úvahu (3.52), pak bude

$$(3.56) \quad c_k = \frac{(f, P_{km})_m}{(P_{km}, P_{km})_m} = \frac{(2k+1)\omega(m, k)}{\omega(m+k+1, k+1)} \cdot \sum_{t=0}^m y_t \cdot P_{km}(t),$$

pro $k = 0, 1, \dots, n$.

Polynom nejlepší approximace v proměnné x , $\phi_n(x)$, dostaneme opět užitím transformace (3.9), tedy

$$\phi_n(x) = \sum_{k=0}^n c_k P_{km}((x - x_0)/h).$$

Příklad 3.8 Pro funkci z příkladu 3.7 s hodnotami v tabulce 3.8 hledejme polynom nejlepší approximace třetího stupně užitím Čebyševových polynomů (3.51).

Rešení: Transformací $t = 2(x + 1)$ převedeme tabulku 3.8 na tabulku 3.9. Approximační polynom $\phi_3(t)$ k tabulce 3.9 je

t	0	1	2	3	4
y(t)	0	-1	0	1	0

Tab. 3.9

$$\phi_3(t) = c_0 P_{04}(t) + c_1 P_{14}(t) + c_2 P_{24}(t) + c_3 P_{34}(t).$$

Protože polynomy P_{04} až P_{34} máme vypočítány v (3.53), pak podle (3.56) snadno dostaneme

$$c_0 = 0 \quad c_1 = -2/5 \quad c_2 = 0 \quad c_3 = 2/5,$$

tedy

$$\phi_3(t) = -\frac{8}{3}t + 2t^2 - \frac{1}{3}t^3$$

a po dosazení $2(x + 1)$ za proměnnou t dostaneme hledaný polynom ve tvaru shodném s výsledkem příkladu 3.7 (ověřte!), kde $\phi_3(x) = 8/3(x - x^3)$.

3.4.4 Trigonometrické polynomy

Při úlohách v reálném čase se velmi často užívají za základní funkce při approximaci MNC v diskrétním případě trigonometrické polynomy

$$(3.57) \quad 1, \cos x, \sin x, \cos 2x, \sin 2x, \dots,$$

které jsou ortogonální na diskrétní množině

$$x = \{x_0, x_1, x_2, \dots, x_{2N}\},$$

kde $x_i = 2\pi i/(2N+1)$, $i = 0, 1, 2, \dots, 2N$.

Množina X tedy obsahuje lichý počet bodů, které jsou ekvidistantním dělením intervalu $(0, 2\pi)$. Funkce (3.57) splňují tyto vztahy

$$\sum_{i=0}^{2N} \sin^2 nx_i = \sum_{i=0}^{2N} \cos^2 nx_i = \frac{2N+1}{2}, \quad n = 1, 2, \dots$$

Při užívání trigonometrických funkcí (3.57) pro approximaci MNČ se postupuje zcela obdobně jako při užívání Čebyševových polynomů v odstavci 3.4.3. Podrobněji viz [23].

3.5 Approximace metodou nejménších čtverců (MNČ) - spojitý případ

Nechť $f(x)$ je spojitá funkce na intervalu $\langle a, b \rangle$. Při spojité approximaci funkce $f(x)$ MNČ hledáme takový zábečněný polynom

$$(3.58) \quad \phi_n(x) = c_0 g_0(x) + c_1 g_1(x) + \dots + c_n g_n(x),$$

viz (3.2), jehož koeficienty minimalizují tzv. střední kvadratickou odchylku

$$(3.59) \quad \varphi = \sqrt{\int_a^b [f(x) - \phi_n(x)]^2 dx}.$$

Zábečněný polynom (3.58) se nyní nazývá polynom nejlepší (spojité) approximace.

3.5.1 Normální rovnice

Obdobně jako v 3.4.1 hledáme koeficienty c_0, c_1, \dots, c_n tak, aby minimalizovaly funkci φ^2 , což je jistě ekvivalentní úloha. Funkce φ^2 může dosahovat svého minima jenom pro takové hodnoty c_0, c_1, \dots, c_n , pro které platí

$$(3.60) \quad \frac{\partial \varphi^2}{\partial c_k} = \int_a^b \left[f(x) - \sum_{j=0}^n c_j g_j(x) \right] g_k(x) dx = 0, \quad k = 0, 1, \dots, n.$$

Tato soustava se opět nazývá soustava normálních rovnic. Protože systém základních funkcí je na intervalu $\langle a, b \rangle$ lineárně nezávislý, je soustava (3.60) jednoznačně řešitelná a její řešení realizuje minimum funkce φ^2 , tedy i φ .

Soustavu (3.60) zapíšeme pomocí tzv. skalárního součinu spojitych funkcí φ, ψ na intervalu $\langle a, b \rangle$, který označíme (φ, ψ) a definujeme vzorcem

$$(3.61) \quad (\varphi, \psi) = \int_a^b \varphi(x) \cdot \psi(x) dx.$$

Pak normální rovnice mají tvar

$$(3.62) \quad \sum_{j=0}^n (g_j, g_k) c_j = (f, g_k), \quad k = 0, 1, \dots, n.$$

Řešením (3.62) jsou koeficienty c_0, c_1, \dots, c_n , které definují zábečněný polynom (3.58) jako polynom nejlepší (spojité) approximace funkce $f(x)$ na intervalu $\langle a, b \rangle$.

3.5.2 Volba základních funkcí

Klasická volba základních funkcí $g_k(x) = x^k$, $k = 0, 1, \dots, n$ je sice možná, ale přináší obdobné potíže, o nichž jsme psali v 3.4.1, proto se budeme i zde orientovat na tzv. ortogonální funkce na intervalu $\langle a, b \rangle$.

Definice 3.5 Systém spojitých funkcí g_0, g_1, \dots, g_n budeme nazývat ortogonální (resp. ortonormální) na intervalu $\langle a, b \rangle$, jestliže

$$(3.63) \quad (g_j, g_k) \begin{cases} = 0 & \text{pro } j \neq k, \\ \neq 0 \text{ (resp. } = 1) & \text{pro } j = k. \end{cases}$$

Je-li systém základních funkcí ortogonální, resp. ortonormální, pak soustava (3.62) má diagonální tvar

$$(3.64) \quad \begin{aligned} (g_k, g_k) c_k &= (f, g_k), \quad k = 0, 1, \dots, n, \\ \text{resp.} \quad c_k &= (f, g_k), \quad k = 0, 1, \dots, n, \end{aligned}$$

s řešením

$$(3.65) \quad c_k = \frac{(f, g_k)}{(g_k, g_k)}, \text{ resp. } c_k = (f, g_k), \quad k = 0, 1, \dots, n.$$

Polynom nejlepší (spojitý) approximace je pak zapsán vzorcem

$$(3.66) \quad \begin{aligned} \phi_n(x) &= \sum_{k=0}^n \frac{(f, g_k)}{(g_k, g_k)} \cdot g_k(x), \\ \text{resp.} \quad \phi_n(x) &= \sum_{k=0}^n (f, g_k) \cdot g_k(x). \end{aligned}$$

Uvedeme dvě volby základních funkcí:

a) Legendreovy polynomy, které jsou ortogonální na intervalu $\langle -1, 1 \rangle$, se definují vzorcem

$$(3.67) \quad P_k(x) = \frac{1}{2^k k!} \frac{d^k}{dx^k} [x^2 - 1]^k, \quad k = 0, 1, \dots, n.$$

Platí $(P_j, P_k) = 0$ pro $j \neq k$ a $(P_k, P_k) = 2/(2k + 1)$. Tedy

$$\begin{array}{ll} P_0(x) = 1 & (P_0, P_0) = 2 \\ P_1(x) = x & (P_1, P_1) = 2/3 \\ P_2(x) = (3x^2 - 1)/2 & (P_2, P_2) = 2/5 \\ P_3(x) = (5x^3 - 3x)/2 & (P_3, P_3) = 2/7 \end{array}$$

b) Trigonometrické funkce

$$(3.68) \quad 1, \sin x, \cos x, \sin 2x, \cos 2x, \sin 3x, \cos 3x, \dots$$

tvoří ortogonální systém na intervalu $\langle -\pi, \pi \rangle$ (dokonce na každém intervalu $\langle a, a + 2\pi \rangle$). Pro tento systém platí $(1, 1) = 2\pi$, $(\sin nx, \sin nx) =$

$= (\cos nx, \cos nx) = \pi$, pro $n = 1, 2, 3, \dots$, tedy z něho můžeme odvodit ortonormovaný systém

$$(3.69) \quad \frac{1}{\sqrt{2\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \frac{\sin 3x}{\sqrt{\pi}}, \frac{\cos 3x}{\sqrt{\pi}}, \dots$$

Na intervalu $\langle -\pi, \pi \rangle$ jsou také ortogonální systémy

$$(3.70) \quad \begin{aligned} & \sin x, \sin 2x, \sin 3x, \sin 4x, \dots, \\ & 1, \cos x, \cos 2x, \cos 3x, \cos 4x, \dots, \end{aligned}$$

z nichž lze odvodit ortonormované systémy

$$(3.71) \quad \begin{aligned} & \frac{\sin x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \frac{\sin 3x}{\sqrt{\pi}}, \frac{\sin 4x}{\sqrt{\pi}}, \dots, \\ & \frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \frac{\cos 3x}{\sqrt{\pi}}, \frac{\cos 4x}{\sqrt{\pi}}, \dots \end{aligned}$$

Systémy (3.70) jsou dokonce ortogonální na intervalu $\langle 0, \pi \rangle$, neboť pro $n \neq m$ je

$$\int_0^\pi \sin nx \sin mx dx = \int_0^\pi \cos nx \cos mx dx = 0$$

$$\int_0^\pi \cos nx dx = 0 \quad (n, m = 1, 2, 3, \dots).$$

Protože

$$\int_0^\pi dx = \pi \quad \text{a} \quad \int_0^\pi \sin^2 nx dx = \int_0^\pi \cos^2 nx dx = \frac{\pi}{2},$$

pak jsou systémy

$$(3.72) \quad \begin{aligned} & \sqrt{\frac{2}{\pi}} \sin x, \sqrt{\frac{2}{\pi}} \sin 2x, \sqrt{\frac{2}{\pi}} \sin 3x, \dots, \\ & \sqrt{\frac{1}{\pi}}, \sqrt{\frac{2}{\pi}} \cos x, \sqrt{\frac{2}{\pi}} \cos 2x, \sqrt{\frac{2}{\pi}} \cos 3x, \dots \end{aligned}$$

na intervalu $\langle 0, \pi \rangle$ ortonormované.

Užijeme-li pro approximaci základní systémy (3.67) až (3.72), pak musíme daný problém převést lineární transformací

$$(3.73) \quad t = \left(x - \frac{b+a}{2} \right) \frac{2}{b-a}, \quad \text{resp. } t = \left(x - \frac{b+a}{2} \right) \frac{2\pi}{b-a}, \quad \text{resp. } t = (x - a) \frac{\pi}{b-a}$$

na interval $\langle -1, 1 \rangle$, resp. $\langle -\pi, \pi \rangle$, resp. $\langle 0, \pi \rangle$. Po vyřešení této nové úlohy se lineární transformací vrátíme k proměnné x . Sledujte a ověřte si postupy na následujících příkladech.

Příklad 3.9 Hledejme na intervalu $\langle 0, 1 \rangle$ k funkci $y = \sqrt{x}$ polynom nejlepší (spojité) approximace druhého stupně při základním systému $g_0 = 1, g_1 = x, g_2 = x^2$.

Řešení: Systém $1, x, x^2$ je sice lineárně nezávislý na intervalu $\langle 0, 1 \rangle$, ale ne-ní na něm ortogonální, proto nejdříve vypočítáme všechny skalérní součiny soustavy (3.62)

$$(g_0, g_0) = 1, \quad (g_0, g_1) = 1/2, \quad (g_0, g_2) = 1/3, \quad (g_1, g_1) = 1/3, \\ (g_1, g_2) = 1/4, \quad (g_2, g_2) = 1/5, \quad (g_0, \sqrt{x}) = 2/3, \quad (g_1, \sqrt{x}) = 2/5, \\ (g_2, \sqrt{x}) = 2/7$$

Řešením soustavy (3.62) s předcházejícími hodnotami dostaneme $c_0 = 6/35$, $c_1 = 48/35$, $c_2 = -4/7$. Pak hledaný polynom nejlepší approximace je

$$\phi_2(x) = 6/35 + 48/35x - 4/7x^2.$$

Příklad 3.10 Řešme ještě jednou úlohu z předcházejícího příkladu, avšak se základním systémem $g_0 = P_0$, $g_1 = P_1$, $g_2 = P_2$, viz (3.67).

Řešení: Nejdříve provedeme transformaci problému podle prvního vzorce v (3.73), tedy

$$t = 2x - 1, \quad \text{resp.} \quad x = (t + 1)/2,$$

takže máme approximovat funkci $\sqrt{(t + 1)/2}$ na intervalu $(-1, 1)$ funkcí

$$(3.74) \quad \phi_2(t) = \sum_{k=0}^2 c_k P_k(t) = c_0 + c_1 t + c_2 (3t^2 - 1)/2.$$

Vzhledem k ortogonalitě funkcí P_0 , P_1 , P_2 na intervalu $(-1, 1)$ má soustava normálních rovnic tvar (3.64). Snadno vypočítáme

$$(f, g_0) = \int_{-1}^1 \sqrt{(t + 1)/2} dt = \frac{4}{3}, \quad (f, g_1) = \int_{-1}^1 t \sqrt{(t + 1)/2} dt = \frac{4}{15}, \\ (f, g_2) = \int_{-1}^1 \sqrt{(t + 1)/2} [3t^2 - 1]/2 dt = -\frac{4}{105}.$$

Podle (3.65) je $c_0 = 2/3$, $c_1 = 2/5$ a $c_2 = -2/21$. Dosazením do (3.74) dostaneme

$$\phi_2(t) = 2/3 + 2/5t - 1/21(3t^2 - 1)$$

a vrátíme-li se transformací $t = 2x - 1$ opět k proměnné x a původní úloze, dostaneme

$$\phi_2(x) = 6/35 + 48/35x - 4/7x^2,$$

což je samozřejmě stejný výsledek jako v příkladě 3.9, protože polynom nejlepší approximace je jediný.

3.6 Čebyševova approximace

Třetí skupina approximací, jak jsme je kategorizovali v odstavci 3.1, jsou Čebyševovy approximace, neboli approximace, které minimalizují maximální chybu. Tyto approximace mají velký význam při sestavování podprogramů pro výpočet funkcí na počítači. Při výpočtech jsou standardní funkce mnohokrát vypočítávány, a to s rozmanitými argumenty z nějakého intervalu (a, b) , proto je musíme approximovat na celém intervalu. Nejdůležitější charakteristikou této approximace je maximální chyba. Snažíme se tedy najít takovou approximaci, která maximální chybu minimalizuje. Nazveme ji nejlepší stejnoměrná approximace.

Až doposud jsme za třídu approximujících funkcí brali polynomy. U Čebyševovy approximace tuto třídu rozšíříme na racionální funkce, tj. na funkce ve tvaru

$$(3.75) \quad R_{pq}(x) = \frac{P_p(x)}{Q_q(x)} = \frac{\sum_{i=0}^p a_i x^i}{\sum_{i=0}^q b_i x^i},$$

kde $P_p(x)$, $Q_q(x)$ jsou nesoudělné algebraické polynomy p -tého, q -tého stupně. Racionální funkce je nejobecnější funkci, kterou počítac může vypočítat přímo. Do této třídy patří jistě i polynomy, stačí položit $q = 0$ a $b_0 = 1$.

Úlohu Čebyševovy approximace formulujeme takto: Nechť $f(x)$ je spojitá funkce na intervalu $\langle a, b \rangle$. Sestrojme k ní takovou racionální funkci (3.75) s indexem $n = p + q$, která minimalizuje hodnotu

$$(3.76) \quad \Delta_{pq} = \max_{x \in \langle a, b \rangle} |f(x) - R_{pq}(x)|.$$

Po teoretické stránce řeší úlohu následující věta.

Věta 3.6 Je-li dána úloha Čebyševovy approximace, pak existuje právě jedna racionální funkce $R_{pq}(x)$ s $n + 1$ koeficienty (jeden z koeficientů $a_0, a_1, \dots, a_p, b_0, b_1, \dots, b_q$ v (3.75) můžeme zvolit libovolně), která minimalizuje Δ_{pq} , viz (3.76). Označíme-li ji $R_{pq}^*(x)$, pak platí:

a) V $\langle a, b \rangle$ existuje alespoň $n + 2$ bodů x_0, x_1, \dots, x_{n+1} tak, že v nich absolutní hodnota chyby approximace nabývá maxima, tzn. že

$$(3.77) \quad |f(x_i) - R_{pq}^*(x_i)| = \max_{x \in \langle a, b \rangle} |f(x) - R_{pq}^*(x)|, \quad i = 0, 1, \dots, n + 1.$$

b) Jestliže $a \leq x_0 < x_1 < \dots < x_{n+1} \leq b$, pak extrémy chyby approximace střídají v těchto bodech znaménka, tj. platí

$$(3.78) \quad f(x_i) - R_{pq}^*(x_i) = - (f(x_{i+1}) - R_{pq}^*(x_{i+1})) \quad \text{pro } i = 0, 1, \dots, n.$$

3.6.1 Konstrukce Čebyševovy approximace

Sestrojit Čebyševovu approximaci $R_{pq}^*(x)$ je v praxi obtížné, protože předcházející věta neříká nic o tom, jak body x_i najít. Kdybychom naopak body x_0, x_1, \dots, x_{n+1} již znali, pak hodnoty $n + 1$ koeficientů vypočítáme z rovnice (3.78) celkem snadno. Pro konstrukci $R_{pq}^*(x)$ ke spojité funkci $f(x)$ na $\langle a, b \rangle$ byla vymyšlena řada iteracních postupů, z nichž si uvedeme tzv. Remesův algoritmus, který vychází z předpokladu znalosti nějaké počáteční approximace $x_0^{(0)}, x_1^{(0)}, \dots, x_{n+1}^{(0)}$ bodů x_0, x_1, \dots, x_{n+1} .

Remesův algoritmus:

- 1) K $n + 2$ bodům $x_0^{(0)}, x_1^{(0)}, \dots, x_{n+1}^{(0)}$ sestrojíme podle rovnice (3.78) nulté přiblížení funkce $R_{pq}^*(x)$, které označíme $R_{pq}^{(0)}(x)$. Pro $q = 0$ je (3.78) lineární soustava a pro $q \neq 0$ nelineární.

2) Najdeme extrémy funkce $f(x) - R_{pq}^{(0)}(x)$. Body, v nichž extrémy nastávají, označíme $x_0^{(1)}, x_1^{(1)}, \dots, x_{n+1}^{(1)}$ a vezmeme je za další aproksimaci bodů x_0, x_1, \dots, x_{n+1} .

Opakováním těchto dvou bodů vytváříme postupně

$$(3.79) \quad \{x_0^{(k)}, x_1^{(k)}, \dots, x_{n+1}^{(k)}\}, R_{pq}^{(k)}(x) \quad \text{pro } k = 0, 1, 2, \dots.$$

V [22] je uveden důkaz konvergence $R_{pq}^{(k)}(x) \rightarrow R_{pq}^*(x)$. Zde jenom poznamenáme, že pro $q = 0$ (pak je $R_{pq}(x)$ polynom) Remesův algoritmus konverguje pro libovolnou volbu počátečních hodnot $x_1^{(0)}, x_2^{(0)}, \dots, x_{n+1}^{(0)}$. Je však vhodné volit tyto počáteční hodnoty v bodech, v nichž Čebyševův polynom $T_{n+1}(x)$ (viz např. [10], str. 65) definovaný na intervalu $\langle a, b \rangle$ nabývá extrémních hodnot, tj. v bodech

$$(3.80) \quad x_i = \frac{a+b}{2} + \frac{b-a}{2} \cos\left(\frac{n+1-i}{n+1}\pi\right), \quad i = 0, 1, \dots, n+1.$$

Příklad 3.11 Remesovým algoritmem sestrojme k \sqrt{x} na intervalu $\langle 0, 1 \rangle$ polynom 1. stupně nejlepší stejnomořná approximace.

Řešení: Zvolme $x_0^{(0)} = 0, x_1^{(0)} = 1/2, x_2^{(0)} = 1$. Pro $f(x) = \sqrt{x}$ a $R_{10}(x) = a_0 + a_1 x$ má soustava (3.78) tvar

$$\begin{aligned} \sqrt{0} - (a_0 + a_1 \cdot 0) &= -\sqrt{1/2} - (a_0 + a_1 \cdot 1/2) \\ \sqrt{1/2} - (a_0 + a_1 \cdot 1/2) &= -\sqrt{1} - (a_0 + a_1 \cdot 1) \end{aligned}$$

a řešení je $a_1 = 1, a_0 = (\sqrt{2} - 1)/4$. Tedy

$$R_{10}^{(0)}(x) = x + (\sqrt{2} - 1)/4 \quad \text{a} \quad \max_{x \in \langle 0, 1 \rangle} |\sqrt{x} - R_{10}^{(0)}(x)| = 0,146.$$

Funkce $\sqrt{x} - R_{10}^{(0)}(x)$ nabývá maxima v bodech $x_0^{(1)} = 0, x_1^{(1)} = 1/4, x_2^{(1)} = 1$ a opakováním postupu dostaneme

$$R_{10}^{(1)}(x) = 1/8 + x \quad \text{a} \quad \max_{x \in \langle 0, 1 \rangle} |\sqrt{x} - R_{10}^{(1)}(x)| = 0,125.$$

Protože funkce $\sqrt{x} - R_{10}^{(1)}(x)$ má extrémy opět v bodech 0, 1/4, 1, je tedy $R_{10}^{(1)}(x)$ hledanou nejlepší stejnomořnou approximací.

3.7 Numerické derivování

Vzorce pro numerický výpočet derivací získáme buď derivováním interpolačního polynomu nebo interpolační spline-funkce nebo jiné approximační funkce.

3.7.1 Derivování interpolačního polynomu

Mějme tedy k tabulce $\{(x_i, y_i) : y_i = f(x_i), i = 1, 2, \dots, n\}$ funkce $y = f(x)$ sestavenou Lagrangeuvou (3.7) nebo Newtonovou (3.21) interpolační polynomem, pak

přibližnou hodnotu k-té derivace ($k \leq n$) funkce $f(x)$ vyjádříme vzorcem

$$(3.81) \quad f^{(k)}(x) \approx L_n^{(k)}(x) = \sum_{i=0}^n \frac{y_i}{\omega_n'(x_i)} \cdot \frac{d^k}{dx^k} \left[\frac{\omega_n(x)}{x - x_i} \right] = \\ = N_n^{(k)}(x) = \sum_{i=1}^n f(x_0; x_1; \dots; x_i) \frac{d^k}{dx^k} \left[\omega_{i-1}(x) \right].$$

Obdobně bychom mohli při ekvidistantní interpolaci derivovat polynomy (3.10) a (3.22) při interpolaci vpřed, resp. polynomy (3.23) a (3.24) při interpolaci vzad.

Chyba při výpočtu derivace pomocí interpolačních polynomů se určí podle (3.26) vztahem (označme tuto chybu $R_n^{(k)}(x)$)

$$(3.82) \quad R_n^{(k)}(x) = \frac{1}{(n+1)!} \frac{d^k}{dx^k} \left[\omega_n(x) f^{(n+1)}(\xi) \right].$$

V tomto vztahu je ξ neznámá funkce proměnné x a proto je vyšetřování chyby velmi komplikované a nepřináší příliš velký užitek. Obecným vyšetřováním chyby se tedy nebudeme zabývat, ale ukážeme její tvar alespoň v případě první derivace, kterou navíc vyjádříme jenom v uzlových bodech interpolace.

Odvodíme nyní některé vzorce numerického derivování při užití Lagrangeova interpolačního polynomu pro ekvidistantní argument. Podle (3.10) je

$$f^{(k)}(x) \approx \frac{d^k}{dx^k} L_n(x) = \frac{1}{h^k} \sum_{i=0}^n \frac{y_i(-1)^{n-i}}{i!(n-i)!} \cdot \frac{d^k}{dt^k} \left[\frac{t(t-1)\dots(t-n)}{t-i} \right].$$

Pro první derivaci (tzn. pro $k = 1$) dostaneme při $n = 1$ vzorec (místo \approx budeme psát =)

$$(3.83) \quad f'(x) = \frac{1}{h} [-y_0 + y_1] + \frac{1}{2} hf''(\xi).$$

Pro $n = 2$ můžeme kromě vzorce pro první derivaci najít i vzorec pro přibližný výpočet druhé derivace

$$(3.84) \quad f'(x) = \frac{1}{2h} [y_0(2t-3) - 2y_1(2t-2) + y_2(2t-1)] + R_2^{(1)}(x), \\ f''(x) = \frac{1}{h^2} [y_0 - 2y_1 + y_2] + R_2^{(2)}(x).$$

Z prvního vzorce v (3.84) můžeme ještě odvodit vzorec pro první derivaci v uzlových bodech x_0, x_1, x_2 , ve kterých je $t = 0, 1, 2$, proto

$$(3.85) \quad f'(x_0) = \frac{1}{2h} [-3y_0 + 2y_1 - y_2] + \frac{1}{3} h^2 f'''(\xi), \\ f'(x_1) = \frac{1}{2h} [-y_0 + y_2] - \frac{1}{6} h^2 f'''(\xi), \\ f'(x_2) = \frac{1}{2h} [y_0 - 4y_1 + 3y_2] + \frac{1}{3} h^2 f'''(\xi).$$

Pro $n = 3$ dostaneme tzv. čtyřbodové vzorce

$$(3.86) \quad \begin{aligned} f'(x) &= \frac{1}{h} \left[-\frac{1}{6} y_0 (3t^2 - 12t + 11) + \frac{1}{2} y_1 (3t^2 - 10t + 6) - \right. \\ &\quad \left. - \frac{1}{2} y_2 (3t^2 - 8t + 3) + \frac{1}{6} y_3 (3t^2 - 6t + 2) \right] + R'_3(x), \\ f''(x) &= \frac{1}{h^2} \left[-\frac{1}{6} y_0 (6t - 12) + \frac{1}{2} y_1 (6t - 10) - \right. \\ &\quad \left. - \frac{1}{2} y_2 (6t - 8) + \frac{1}{6} y_3 (6t - 6) \right] + R''_3(x), \end{aligned}$$

ze kterých snadno odvodíme $f'(x_0)$, $f'(x_1)$, $f'(x_2)$, $f'(x_3)$ a $f''(x_0)$, $f''(x_1)$, $f''(x_2)$, $f''(x_3)$, když do vzorců (3.86) postupně dosadíme $t = 0, 1, 2, 3$, tedy

$$\begin{aligned} f'(x_0) &= \frac{1}{6h} [-11y_0 + 18y_1 - 9y_2 + 2y_3] - \frac{h^3}{4} f^{(4)}(\xi), \\ f''(x_0) &= \frac{1}{6h^2} [12y_0 - 30y_1 + 24y_2 - 6y_3] + \frac{11}{12} h^2 f^{(4)}(\xi) - \frac{h^3}{10} f^{(5)}(\xi_1), \end{aligned}$$

atd.

Při počítání chyb podle (3.82) platí $\omega_n(x_i) = 0$, pro $i = 0, 1, \dots, n$.

Obdobně bychom mohli odvozovat i vzorce, které vyplývají z Newtonova interpolačního polynomu nebo interpolačních polynomů jiných. Prostudujte si následující příklad, v němž je užit Newtonův interpolační polynom.

Příklad 3.12 Vypočítejme přibližnou hodnotu derivace funkce $\log \Gamma(x)$ v bodě $\tilde{x} = 10,0154$, je-li dána následující tabulka funkce

x	$\log \Gamma(x)$	$\Delta^1 f_i$	$\Delta^2 f_i$
10,00	5,559 7630		
10,02	5,579 3306	19 5676	
10,04	5,598 9165	19 5859	183

Rешení: Krok $h = 0,02$, $x_0 = 10$, $\tilde{x} = x_0 + 0,77 \cdot h$. Pro danou tabulku sestavme Newtonův interpolační polynom (3.22) pro $n = 2$. Potom

$$f'(x) = N'_2(x) + R'_2(x) = \frac{1}{2!} \left[\Delta^1 f_0 + \frac{2t - 1}{2!} \Delta^1 f_0 \right] + R'_2(x)$$

$$f'(10,0154) \approx (0,019 5676 + 0,27 \cdot 0,000 0183) \cdot 50 = 0,978 6270$$

3.7.2 Derivování interpolační spline-funkce

Předpokládejme, že pro tabulku $\{(x_i, y_i) : y_i = f(x_i), x_0 < x_1 < \dots < x_n\}$ funkce $f(x)$ jsme sestrojili pro nějaké okrajové podmínky S-funkci (3.32), potom

$$S'(x) = S'_i(x) = -\frac{(x_i - x)^2}{2h_i} M_{i-1} + \frac{(x - x_{i-1})^2}{2h_i} M_i -$$

$$-\left(y_{i-1} - \frac{M_{i-1}h_i^2}{6}\right) \frac{1}{h_i} + \left(y_i - \frac{M_i h_i^2}{6}\right) \frac{1}{h_i}$$

pro $x \in \langle x_{i-1}, x_i \rangle$, $i = 1, 2, \dots, n$. Tedy

$$S'(x_i) = \begin{cases} S'_i(x_i) = \frac{h_i}{6} M_{i-1} + \frac{h_i}{3} M_i + \frac{1}{h_i} (y_i - y_{i-1}) & \text{pro } i = 1, 2, \dots, n \\ S'_{i+1}(x_i) = -\frac{h_{i+1}}{3} M_i - \frac{h_{i+1}}{6} M_{i+1} + \frac{1}{h_{i+1}} (y_{i+1} - y_i) & \text{pro } i = 0, 1, \dots, n-1 \end{cases}$$

Pro druhou derivaci dostaneme

$$S''(x) = S''_i(x) = \frac{x_i - x}{h_i} M_{i-1} + \frac{(x - x_{i-1})}{h_i} M_i ,$$

tedy

$$S''(x_i) = \begin{cases} S''_i(x_i) = M_i \\ S''_{i+1}(x_i) = M_i \end{cases}$$

(srovnejte s (3.30) a (3.31)).

3.8 Numerické integraci (kvadratura)

Nechť je dána spojitá funkce $f(x)$ na intervalu $\langle a, b \rangle$. Mějme za úkol nalézt její vlastní určitý integrál

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

V mnoha případech jde tato úloha řešit analytickou cestou pomocí tzv. primitivní funkce, jak jste se s tím seznámili v základním kursu matematiky. Často však je funkce $f(x)$ zadána složitým předpisem nebo pro ni známe jenom tabulku hodnot a pak nám nezbývá než approximovat danou funkci buď interpolačním polynomem, nebo některými jinými funkcemi a s jejich pomocí hledat vzorce pro přibližný výpočet integrálu. Takové vzorce se nazývají kvadraturní vzorce a postup numerického výpočtu integrálu kvadraturou.

Budeme hledat takové kvadraturní vzorce, které převádějí výpočet integrálu na výpočet lineární kombinace

$$(3.87) \quad \sum_{i=0}^n c_i^{(n)} \cdot f(x_i^{(n)}) \approx I ,$$

kde $c_i^{(n)}$ jsou reálné koeficienty a $x_i^{(n)}$ jsou tzv. uzly integrace. Kvadraturní vzorce, pro které platí

$$\lim_{n \rightarrow \infty} \sum_{i=0}^n c_i^{(n)} \cdot f(x_i^{(n)}) = \int_a^b f(x) dx ,$$

nazýváme konvergentní.

Často užíváme vzorce založené na approximaci integrované funkce interpolačním polynomem n -tého stupně, kdy pro uzly interpolace platí $a \leq x_0 < x_1 < \dots < x_n \leq b$. Užijeme-li např. Lagrangeův interpolační polynom (3.7), pak má kvadraturní vzorec (označený $K(f)$) tvar

$$(3.88) \quad K(f) = \int_a^b L_n(x) dx = \sum_{i=0}^n \int_a^b F_i(x) dx \cdot f(x_i).$$

V takovém případě (srovnejme-li (3.87) a (3.88)) se shodují uzly interpolace s uzly integrace. Tyto tzv. interpolační kvadraturní vzorce nazýváme otevřené, jestliže $a \neq x_0$, $b \neq x_n$, a uzavřené, když $a = x_0$, $b = x_n$. Vzorec (3.88) nemusí být ale konvergentní, protože jak jsme uvedli v odstavci 3.2.5, nemusí, zejména při ekvidistantní interpolaci, konvergovat interpolační polynom pro $n \rightarrow \infty$ k approximované spojité funkci.

Zvyšování přesnosti numerického výpočtu integrálu se dosáhne při interpolační kvadratuře zvyšováním stupně interpolačního polynomu nebo užíváním tzv. složených interpolačních kvadraturních vzorců. Složené vzorce dostaneme tím, že interval $\langle a, b \rangle$ rozdělíme na konečný počet subintervalů a na každém z nich užijeme interpolační kvadraturní vzorec (tzv. základní). "Sečtemi těchto vzorců dostaneme složený vzorec. Přesněji to objasníme později.

Kvadraturní vzorce (3.88) mohou být pro pevné n hledány také z podmínky, aby byly přesné pro polynomy co nejvyššího stupně. Toho lze dosáhnout vhodným rozložením uzelů integrace. Získané vzorce se nazývají Gaussovy kvadraturní vzorce.

V současné době je však velmi často užívána k výpočtu určitých integrálů tzv. Rombergova integrační metoda, která zpřesňuje výsledky lichoběžníkové metody, viz 3.8.2. Její algoritmus uvedeme v odstavci 3.8.5.

3.8.1 Newtonovy - Cotesovy kvadraturní vzorce

Základní tvar těchto vzorců dostaneme tak, že integrovanou funkci nahradíme Lagrangeovým interpolačním polynomem pro ekvidistantní argumenty (3.10). Rozdělme tedy interval $\langle a, b \rangle$ ekvidistantními uzly x_0, x_1, \dots, x_n tak, že $x_i = a + ih$, kde $h = (b - a)/n$ a na těchto uzlech nahradíme funkci $f(x)$ interpolačním polynomem (3.10), potom

$$(3.89) \quad \int_a^b f(x) dx = \int_a^b L_n(x) dx + \int_a^b R_n(x) dx = h \sum_{i=0}^n c_i^{(n)} \cdot f(x_i) + \int_a^b R_n(x) dx,$$

kde

$$c_i^{(n)} = \frac{(-1)^{n-i}}{i!(n-i)!} \int_0^n \frac{t(t-1)\dots(t-n)}{t-i} dt.$$

$\int_a^b R_n(x) dx$ je chyba vzorce, která vyplývá z chyby interpolace $R_n(x)$, viz (3.26). Konstanty $c_i^{(n)}$ nejsou závislé na $f(x)$ ani na a, b , a proto se mohou pro různá n, i tabelovat. Vzorce

$$(3.90) \quad K_{NC}(f) = h \sum_{i=0}^n c_i^{(n)} \cdot f(x_i)$$

s koeficienty $c_i^{(n)}$ počítanými podle (3.89), se nazývají Newtonovy - Cotesovy kvadraturní vzorce uzavřeného typu.

x	f(x)
0,0	0,398 94
0,1	0,396 95
0,2	0,391 04
0,3	0,381 37
0,4	0,368 25
0,5	0,352 06
0,6	0,333 33
0,7	0,312 24
0,8	0,289 68
0,9	0,266 09
1,0	0,242 06

Tab. 3.10

Příklad 3.13 Vypočítejme přibližnou hodnotu integrálu $I = \int_0^1 e^{-(x^2/2)} / \sqrt{2\pi} dx$, jestliže máme pro integrovanou funkci danou tabulkou 3.10.

Řešení: Přibližný výpočet integrálu provedeme podle vzorců (3.90) a (3.89) pro $h = 0,2$, $x_0 = 0,0$, $x_1 = 0,2$, ..., $x_5 = 1,0$. Podle (3.89) dostaneme

$$c_0^{(5)} = c_5^{(5)} = 95/288 = 0,329 86,$$

$$c_1^{(5)} = c_4^{(5)} = 250/288 = 0,868 06,$$

$$c_2^{(5)} = c_3^{(5)} = 375/288 = 1,302 08,$$

tedy podle (3.90) je

$$I \approx 0,2 [c_0^{(5)} f(0) + c_1^{(5)} f(0,2) + c_2^{(5)} f(0,4) + c_3^{(5)} f(0,6) + \\ + c_4^{(5)} f(0,8) + c_5^{(5)} f(1)] = 0,341 34.$$

3.8.2 Lichoběžníková metoda

Zvolme ve vzorcích (3.90) $n = 1$, potom $c_0^{(1)} = c_1^{(1)} = 1/2$ a tím dostaneme základní vzorec lichoběžníkové metody

$$\int_a^b f(x) dx \approx (b-a) \left[\frac{1}{2} f(a) + \frac{1}{2} f(b) \right] = L_z(f).$$

Rozdělíme-li interval a, b rovnoměrně na m subintervalů body $x_i = a + ih$, kde $h = (b-a)/m$, pak ze základního vzorce dostaneme

$$(3.91) \quad \int_a^b f(x) dx = \sum_{i=0}^{m-1} \int_{x_i}^{x_{i+1}} f(x) dx \approx \sum_{i=0}^{m-1} h \left[\frac{1}{2} f(x_i) + \frac{1}{2} f(x_{i+1}) \right] = \\ = h \left[\frac{1}{2} f(x_0) + f(x_1) + f(x_2) + \dots + f(x_{m-1}) + \frac{1}{2} f(x_m) \right] = L(f, h),$$

což je tzv. složený vzorec lichoběžníkové metody.

Příklad 3.14 Vypočítejme přibližnou hodnotu integrálu z příkladu 3.13, kde je pro integrovanou funkci zadáná tabulka 3.10. K výpočtu užijme složeného vzorce lichoběžníkové metody (3.91) pro $h = 0,1$ a $m = 10$.

Řešení:

$$I \approx 0,1 \left[\frac{1}{2} 0,398 94 + 0,396 95 + 0,391 04 + \dots + \right. \\ \left. + 0,266 09 + \frac{1}{2} 0,242 06 \right] = 0,341 14.$$

3.8.3 Simpsonova metoda

Zvolme ve vzorci (3.90) $n = 2$, potom $c_0^{(2)} = 1/3$, $c_1^{(2)} = 4/3$ a $c_2^{(2)} = 1/3$. Tím dostaneme základní vzorec Simpsonovy metody

$$\int_a^b f(x) dx \approx \frac{b-a}{2} \left[\frac{1}{2} f(a) + \frac{4}{3} f\left(\frac{a+b}{2}\right) + \frac{1}{3} f(b) \right] = S_z(f).$$

Rozdělme nyní interval $\langle a, b \rangle$ rovnoměrně na $2m$ subintervalů body $x_i = a + ih$, kde $h = (b - a)/2m$, pak ze základního vzorce dostaneme

$$\begin{aligned} \int_a^b f(x) dx &= \sum_{i=0}^{m-1} \int_{x_{2i}}^{x_{2i+2}} f(x) dx \approx \sum_{i=0}^{m-1} \frac{h}{3} [f(x_{2i}) + 4f(x_{2i+1}) + f(x_{2i+2})] = \\ (3.92) \quad &= \frac{h}{3} [f(x_0) + 4f(x_1) + 2f(x_2) + \dots + 4f(x_{2m-2}) + f(x_{2m})] = S(f, h), \end{aligned}$$

což je složený vzorec Simpsonovy metody.

Příklad 3.15 Vypočítejme přibližnou hodnotu integrálu z příkladu 3.13, kde je pro integrovanou funkci zadána tabulka 3.10. K výpočtu užijme složeného vzorce Simpsonovy metody (3.92) pro $h = 0,1$ a $m = 5$.

Řešení:

$$\begin{aligned} I &\approx \frac{0,1}{3} [0,39894 + 4 \cdot 0,39695 + 2 \cdot 0,39104 + 4 \cdot 0,38137 + \\ &\quad + 4 \cdot 0,38137 + 2 \cdot 0,36825 + \dots + 4 \cdot 0,26609 + 0,24206] = \\ &= 0,34134. \end{aligned}$$

3.8.4 Gaussovy kvadraturní vzorce

Hledejme kvadraturní vzorec (3.87), tedy rozložení uzlů x_0, x_1, \dots, x_n na intervalu $\langle a, b \rangle$ a hodnoty konstant $c_0^{(n)}, c_1^{(n)}, \dots, c_n^{(n)}$ tak, aby tento vzorec byl přesný pro polynom co nejvyššího stupně. Je-li $p(x) = a_0 + a_1 x + \dots + a_m x^m$, pak požadujeme, aby platilo

$$(3.93) \quad \int_a^b p(x) dx = \sum_{i=0}^n c_i^{(n)} \cdot p(x_i)$$

pro maximální možné m , tedy aby platila rovnost

$$\sum_{k=0}^m a_k \int_a^b x^k dx = \sum_{k=0}^m a_k \sum_{i=0}^n c_i^{(n)} x_i^k.$$

Označíme-li integrál v posledním vzorci I_k , pak z něho odvodíme soustavu $m+1$ rovnic pro $2(n+1)$ neznámých $c_i^{(n)}$ a x_i ($i=0,1,2,\dots,n$), které mají tvar

$$\begin{aligned}
 I_0 &= c_0^{(n)} x_0^0 + c_1^{(n)} x_1^0 + \dots + c_n^{(n)} x_n^0, \\
 I_1 &= c_0^{(n)} x_0^1 + c_1^{(n)} x_1^1 + \dots + c_n^{(n)} x_n^1, \\
 &\dots \dots \dots \dots \dots \dots \dots \dots \\
 I_m &= c_0^{(n)} x_0^m + c_1^{(n)} x_1^m + \dots + c_n^{(n)} x_n^m,
 \end{aligned} \tag{3.94}$$

kde k-tou rovnici soustavy dostaneme v důsledku libovolné volby konstant a_0, a_1, \dots, a_n tak, že zvolíme $a_k \neq 0$, $a_j = 0$ pro $j \neq k$. Protože soustava má $m+1$ rovnic o $2(n+1)$ neznámých, pak musí platit $m = 2n+1$. Tedy vzorec (3.93) je přesný maximálně pro polynomy $(2n+1)$ -ního stupně. Řešit soustavu (3.94) je obtížné, a proto bylo hledáno řešení této úlohy pomocí ortogonálních polynomů. Dá se ukázat, že když zvolíme uzly x_0, x_1, \dots, x_n v kořenech ortogonálních (např. Legendreových) polynomů na intervalu $\langle a, b \rangle$ a sestrojíme na těchto uzlech interpolační kvadraturní vzorec (3.88), pak je tento vzorec přesný pro polynomy nejvýše $(2n+1)$ -ního stupně.

Nechť Legendreův polynom $P_n(x)$, viz (3.67), má na intervalu $\langle -1, 1 \rangle$ kořeny t_0, t_1, \dots, t_n (seřazené podle velikosti), které lineární transformací $x_i = 1/2(b-a)t_i + 1/2(b+a)$ převedeme na body x_0, x_1, \dots, x_n intervalu $\langle a, b \rangle$. Na soustavě těchto uzlů sestrojíme pro funkci $f(x)$ Lagrangeův interpolační polynom (3.7) a podle (3.88) otevřený interpolační kvadraturní vzorec

$$\sum_{i=0}^n \int_a^b \frac{\omega_n(x)}{(x-x_i)\omega'_n(x_i)} dx \cdot f(x_i). \tag{3.95}$$

Pro výpočet integrálů ve vzorci (3.95) zvolíme substituci $x = 1/2(b-a)t + 1/2(b+a)$, tedy

$$\int_a^b \frac{\omega_n(x)}{(x-x_i)\omega'_n(x_i)} dx = \frac{1/2(b-a)}{\omega'_n(t_i)} \int_{-1}^1 \frac{\omega_n(t)}{t-t_i} dt. \tag{3.96}$$

Pak dostaneme hledaný Gaussův kvadraturní vzorec

$$G(f) = \frac{b-a}{2} \sum_{i=0}^n c_i^{(n)} \cdot f(x_i), \tag{3.97}$$

kde koeficienty $c_i^{(n)}$ jsou dány vzorcem

$$c_i^{(n)} = \frac{1}{\omega'_n(t_i)} \int_{-1}^1 \frac{\omega_n(t)}{t-t_i} dt \tag{3.98}$$

a mohou se pro různá i, n tabelovat, viz tab. 3.11. Vzhledem k symetrii kořenů t_0, t_1, \dots, t_n podle počátku platí i symetrie bodů x_0, x_1, \dots, x_n podle středu intervalu $\langle a, b \rangle$ a také symetrie koeficientů $c_i^{(n)}$, takže platí $c_i^{(n)} = c_{n-i}^{(n)}$.

n	i	t_i	$c_i^{(n)}$
2	0,2	$\bar{t} 0,774\ 596\ 67$	0,555 555 56
	1	0	0,888 888 89
3	0,3	$\bar{t} 0,861\ 136\ 31$	0,347 854 84
	1,2	$\bar{t} 0,339\ 981\ 04$	0,652 145 16
4	0,4	$\bar{t} 0,906\ 175\ 85$	0,236 926 88
	1,3	$\bar{t} 0,538\ 469\ 31$	0,478 628 68
	2	0	0,568 888 89
5	0,5	$\bar{t} 0,932\ 469\ 51$	0,171 324 50
	1,4	$\bar{t} 0,661\ 209\ 39$	0,360 761 58
	2,3	$\bar{t} 0,238\ 619\ 19$	0,467 913 94

Tab. 3.11

Příklad 3.15 Vypočítejme přibližnou hodnotu integrálu $I = \int_0^1 \sqrt{1+2x} dx$ pomocí Gaussova kvadraturního vzorce (3.97) pro $n = 2$.

Řešení: Z tabulky 3.11 vybereme pro $n = 2$ hodnoty kořenů t_0 , t_1 a t_2 a po substituci dostaneme uzlové body integrace $x_0 = 0,5 + 0,5t_0 = 0,112\ 702$, $x_1 = 0,5 + 0,5t_1 = 0,5$ a $x_2 = 0,5 + 0,5t_2 = 0,887\ 298$. Potom

$$I \approx 0,5 \sum_{i=0}^2 c_i^{(2)} \sqrt{1+2x_i} = 1,3987,$$

kde $c_0^{(2)}$, $c_1^{(2)}$ a $c_2^{(2)}$ jsou z tabulky 3.11. Přesná hodnota integrálu je $I = \sqrt{3} - 1/3 = 1,39872$.

3.8.5 Rombergova metoda

Ve [23], resp. [15] je dokázáno, že má-li funkce $f(x)$ na intervalu $\langle a, b \rangle$ spojitou $(2n+2)$ -hou derivaci, pak existují čísla a_1, a_2, \dots, a_n tak, že

$$(3.99) \quad L_{k,0}(f, h) - \int_a^b f(x) dx = \sum_{j=1}^n a_j h^{2j} + O(h^{2n+2}),$$

kde $L_{k,0}$ je vzorec lichoběžníkové metody (3.91) pro $m = 2^k$, $h = (b-a)/2^k$, čísla a_1, a_2, \dots závisejí jenom na $a, b, f(x)$. $O(h^{2n+2})$ je chyba řádu $2n+2$.

Z hodnot $L_{k,0}$ pro $k = 0, 1, 2, \dots$ definujme čísla

$$(3.100) \quad L_{k,1} = \frac{1}{3} (4L_{k,0} - L_{k-1,0}), \quad k = 1, 2, \dots$$

Označíme-li integrál v (3.99) I a dosadíme-li ze (3.99) do (3.100), pak pro $k = 1, 2, \dots$ dostaneme

$$(3.101) \quad L_{k,1} - I = \frac{1}{3} \sum_{j=1}^n (2^{2j-4}) a_j \left(\frac{b-a}{2^k}\right)^{2j} + O_1 \left(\left(\frac{b-a}{2^{k-1}}\right)^{2n+2}\right).$$

Hodnota $L_{k,1}$ approximuje přesněji integrál I oproti $L_{k,0}$, protože první člen chyby ve (3.99) obsahuje h^2 , kdežto první člen chyby ve (3.101) je nula a druhý člen obsahuje h^4 . Tím se řád chyby zdvojnásobil.

S tímto postupem zpřesňování můžeme analogicky pokračovat podle vzorce

$$(3.102) \quad L_{k,s} = \frac{1}{4^s - 1} (4^s L_{k,s-1} - L_{k-1,s-1}), \quad k = 0, 1, \dots, s = 1, 2, \dots$$

Obdobně jako ve (3.101) můžeme ukázat, že řád chyby approximace $L_{k,s}$ je $2(s+1)$.

Výpočet podle vzorce (3.102) uspořádáme do schématu

$$(3.103) \quad \begin{array}{ccccccccc} & & L_{0,0} & & & & & & \\ & L_{1,0} & & L_{1,1} & & & & & \\ L_{2,0} & & L_{2,1} & & L_{2,2} & & & & \\ \cdots & & \cdots & & \cdots & & & & \\ L_{k,0} & & L_{k,1} & & L_{k,2} & \cdots & & L_{k,k} \end{array}$$

V prvním sloupci je $k+1$ hodnot počítaných podle lichoběžníkové metody pro $m = 1, 2, 4, 8, \dots, 2^k$. V dalších sloupcích počítáme zpřesňování podle vzorce (3.102), přitom posloupnost $L_{0,0}, L_{1,1}, \dots, L_{k,k}$ konverguje k I mnohem rychleji než posloupnost $L_{0,0}, L_{1,0}, L_{2,0}, \dots, L_{k,0}$, protože hlavní člen chyby (tj. člen s minimální mocninou kroku) approximace $L_{k,k}$ obsahuje $[(b-a)/2^k]^{2(k+1)}$, zatímco hlavní člen chyby approximace $L_{k,0}$ obsahuje $[(b-a)/2^k]^2$.

Postup výpočtu approximace $L_{k,k}$ integrálu I podle vzorce (3.103) na základě lichoběžníkové metody se nazývá Rombergova metoda.

Algoritmus Rombergovy metody je založen na výpočtu lichoběžníkového pravidla $L(f, h)$ - viz (3.91) - a můžeme jej přehledně zapsat takto:

1. $L_{0,0} = L(f, h = b - a)$

2. $k = 1$

3. Výpočet $L_{k,0} = L(f, h = (b - a)/2^k)$

4. Pro $s = 1, 2, \dots, k$ vypočítáváme postupně $L_{k,1}, L_{k,2}, \dots, L_{k,k}$ podle vzorce (3.103) a sledujeme rozdíly $|L_{k,s} - L_{k-1,s}|$. Je-li tento rozdíl menší než předešaná chyba, pak $L_{k,s}$ považujeme za přibližnou hodnotu počítaného integrálu.

5. Když nebyl výpočet ukončen z hlediska přesnosti v předcházejícím kroku, pak zvětšíme k o jedničku a opakujeme třetí a další kroky.

Lichoběžníková metoda má při Rombergově postupu tu výhodu, že k výpočtu $L_{k,0}$ můžeme využít $L_{k-1,0}$. Stačí si uvědomit, že při výpočtu $L_{k,0}$ užíváme polovičního kroku než při výpočtu $L_{k-1,0}$ a tedy užijeme všech funkčních hodnot, z nichž je stanovena hodnota $L_{k-1,0}$.

Příklad 3.16 Vypočítejme $I = \int_0^{0,8} (\sin x)/x \, dx$ Rombergovou metodou s přesností 10^{-6} po sobě následujících approximací.

Řešení: Numerický výpočet úlohy přesně podle algoritmu Rombergovy metody je zapsán v tabulce 3.12.

k	h	$L_{k,0}$	$L_{k,1}$	$L_{k,2}$
0	0,8	0,758 680		
1	0,4	0,768 760	0,772 120	
2	0,2	0,771 262	0,772 096	0,772 094
3	0,1	0,771 887	0,772 095	0,772 095

Tab. 3.12

Protože $|L_{3,1} - L_{2,1}| = 10^{-6}$, položíme $I \approx L_{3,1}$. Přesná hodnota na šest desetinných míst je $I \doteq 0,772\ 095$.

3.8.6 Chyby a konvergance metod

Chyby základních vzorců lichoběžníkové a Simpsonovy metody můžeme odvodit z (3.89) na základě vyjádření $R_n(x)$ vzorcem (3.26). Celý postup je však poměrně složitý, protože (3.26) obsahuje ξ , které je neznámou funkcí x , proto uvedeme jenom výsledky; celý postup je např. ve [23]. Pro dostatečně hladkou funkci $f(x)$ na intervalu (a, b) tedy pro lichoběžníkovou a Simpsonovu metodu platí

$$\int_a^b f(x) dx = L_z(f) - (b-a)^3 \cdot \frac{1}{12} \cdot f''(\xi_L), \quad (3.104)$$

$$\int_a^b f(x) dx = S_z(f) - \left(\frac{b-a}{2}\right)^5 \cdot \frac{1}{90} \cdot f^{IV}(\xi_S),$$

kde ξ_L, ξ_S jsou blíže neurčené body z intervalu (a, b) .

Ze vztahů (3.104) snadno odvodíme vyjádření chyb pro složené vzorce $L(f, h)$, $S(f, h)$ lichoběžníkové a Simpsonovy metody, neboť podle (3.91), (3.92) a (3.104) platí

$$\int_a^b f(x) dx = \sum_{i=0}^{m-1} \left[L_z(f) - \frac{h^3}{12} f''(\xi_{L_i}) \right] = L(f, h) - (b-a) \frac{h^2}{12} f''(\xi_L), \quad (3.105)$$

$$\int_a^b f(x) dx = \sum_{i=0}^{m-1} \left[S_z(f) - \frac{h^5}{90} f^{IV}(\xi_{S_i}) \right] = S(f, h) - (b-a) \frac{h^4}{180} f^{IV}(\xi_S),$$

kde $\xi_{L_i} \in (x_i, x_{i+1})$, $\xi_{S_i} \in (x_i, x_{i+2})$ a $\xi_L, \xi_S \in (a, b)$. Poslední úpravy chyb by ve vzorcích (3.105) plynou z předpokladu spojitosti f'' na (a, b) , protože pak ze základních vlastností spojitých funkcí plyne existence bodů ξ_L, ξ_S tak, že

$$- \frac{h^3}{12} \sum_{i=0}^{m-1} f''(\xi_{L_i}) = - \frac{h^3}{12} m \cdot f''(\xi_L)$$

a

$$- \frac{h^5}{90} \sum_{i=0}^{m-1} f^{IV}(\xi_{S_i}) = - \frac{h^5}{90} m \cdot f^{IV}(\xi_S).$$

Pro konečný tvar chyby u lichoběžníkové metody lze užít $h \cdot m = b - a$ a u Simpsonovy metody $h \cdot 2m = b - a$.

Ještě poznamenejme, že oba složené vzorce reprezentují konvergentní kvadratury, takže pro ně platí

$$\lim_{m \rightarrow \infty} L(f, h = (b - a)/m) = \int_a^b f(x) dx$$

$$\lim_{m \rightarrow \infty} S(f, h = (b - a)/(2m)) = \int_a^b f(x) dx .$$

3.9 Cvičení

- Sestavte Lagrangeovy a Newtonovy interpolační polynomy funkcií, pro které jsou zadány hodnoty
 - $f(1) = 2, f(5) = 30, f(3) = -4, f(0) = 5,$
 - $f(-1) = -5, f(0) = -7, f(1) = -3, f(5) = 193,$
 - $f(-2) = -10, f(-1) = -12, f(0) = -10, f(0,5) = -10,3125,$
 $f(1) = -10.$
- Jsou dány hodnoty funkce $f(x) : f(0,2) = 0,962, f(0,4) = 0,862, f(0,6) = 0,735,$
 $f(0,8) = 0,610$. Vypočítejte Nevilllovým algoritmem přibližnou hodnotu $f(0,5)$ při interpolaci kubickým polynomem.
- Jsou dány hodnoty funkce $f(x) : f(-1) = 2, f(0) = 1, f(1) = 2, f(3) = 0$. Sestrojte interpolační spline-funkci $S(x)$, pro kterou bude platit $S''(-1) = S''(3) = 0$. Porovnejte graficky sestrojenou spline-funkci s interpolačním polynomem 3. stupně.
- Jsou dány hodnoty funkce $y = f(x) : f(-2) = 7, f(-1) = 2, f(0) = -0,5,$
 $f(1) = -2, f(2) = -0,5$. Metodou nejmenších čtverců sestrojte polynom nejvýše 2. stupně jako polynom nejlepší approximace zadaných hodnot funkce, jestliže za základní funkce zvolíte:
 - $1, x, x^2,$
 - Čebyševovy polynomy P_{04}, P_{14}, P_{24} .
- Metodou nejmenších čtverců sestrojte k funkci $y = \sin x$ na intervalu $\langle -\pi/2, \pi/2 \rangle$ polynom nejlepší approximace 3. stupně, jestliže za základní funkce zvolíte:
 - polynomy $1, x, x^2, x^3,$
 - Legendreovy polynomy P_0, P_1, P_2, P_3 .
 Výsledky porovnejte graficky.
- Vypočítejte přibližné hodnoty integrálů

$$I_1 = \int_0^1 \sqrt{x} \sin x dx, \quad I_2 = \int_0^1 \frac{dx}{2+x}, \quad I_3 = \int_0^1 e^{x^2} dx ,$$

složenými vzorci lichoběžníkové a Simpsonovy metody. Při výpočtu budete interval $\langle 0, 1 \rangle$ postupně dělit na 2, 4, 8, 16, ... dílů a počítat approximace integrálů oběma metodami tak dlouho, pokud se nebudou po sobě jdoucí approximace lišit o méně než 10^{-3} .

7. Přibližné hodnoty integrálů z příkladu 6. vypočítejte ještě Rombergovou metodou a srovnajte s předcházejícím výpočtem rychlosť konvergencie.

Výsledky

1. a) $x^3 - 4x^2 + 5$, b) $x^3 + 3x^2 - 7$, c) $x^4 + x^3 - 2x^2 - 10$

2. $f(0,5) \approx 0,8001$

3. $M_0 = 0$, $M_1 = 84/23$, $M_2 = -60/23$, $M_3 = 0$

4. $\phi_2(x) = x^2 - 1,9x - 0,8$

5. $\phi_3(x) = 0,988\ 792\ x - 0,145\ 062\ x^3$

6. $I_1 \doteq 0,3647$, $I_2 \doteq 0,4055$, $I_3 \doteq 1,4806$

K A P I T O L A 4 .

Ř E Š E N Í N E L I N E Á R N Í C H R O V N I C A J E J I C H S O U S T A V

V této kapitole se budeme zabývat dvěma úlohami, jednak řešením rovnice

$$(4.1) \quad f(x) = 0,$$

kde $f(x)$ je nelineární funkce jedné reálné proměnné, která je definovaná a spojita na konečném nebo nekonečném intervalu I , a jednak řešením soustavy

$$\begin{aligned} f_1(x_1, x_2, x_3, \dots, x_n) &= 0, \\ f_2(x_1, x_2, x_3, \dots, x_n) &= 0, \\ &\dots \\ f_n(x_1, x_2, x_3, \dots, x_n) &= 0, \end{aligned}$$

kde $f_1, f_2, f_3, \dots, f_n$ jsou nelineární funkce n proměnných, definované a spojité v jisté oblasti G .

Podstatná část této kapitoly bude věnována výpočtu přibližných hodnot tzv. izolovaných kořenů rovnice (4.1), tj. takových hodnot, které rovnici (4.1) splňují a existuje k nim okolí, v němž žádný druhý kořen neexistuje. Rovnice (4.1) nemusí mít samozřejmě jenom reálné kořeny, ale i kořeny komplexní. Ve většině případů se však budeme zabývat jenom výpočtem reálných kořenů, protože komplexní kořeny se kromě algebraických rovnic vyskytují velmi zřídka.

Přibližný výpočet izolovaných reálných kořenů rovnice (4.1) se skládá ze dvou fází:

- 1) Separace kořenů, tj. určení intervalů, které obsahují právě jeden izolovaný kořen.
- 2) Zpřesnění přibližných hodnot kořenů.

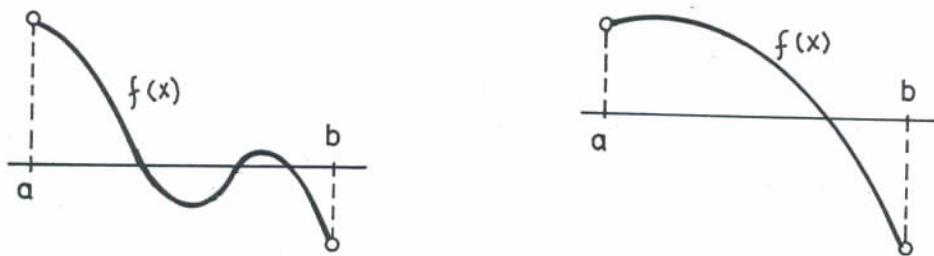
Řešení soustav n nelineárních rovnic o n neznámých může být mnohdy velice komplikované, proto se jím nebudeme příliš široce zabývat. Uvedeme pouze dvě metody (v odstavci 4.4), které jsou při numerickém řešení nejužívanější.

4.1 Separace kořenů nelineárních rovnic

Pro separaci reálných kořenů rovnice (4.1) lze užít následující větu, která je dobře známá ze základního kursu matematiky.

Věta 4.1 Nechť $f(x)$ je spojitá funkce na intervalu I . Jestliže $f(a) \cdot f(b) < 0$ pro $\langle a, b \rangle \subset I$, pak uvnitř intervalu $\langle a, b \rangle$ leží alespoň jeden reálný kořen rovnice (4.1). Jestliže navíc $f'(x)$ na intervalu (a, b) nemění znaménko, pak je tento kořen právě jeden (geometrická demonstrace je na obr. 4.1).

Důkaz těchto tvrzení viz [Budinský: Matematika, skripta ČVUT FSv].



Obr. 4.1

Při separaci reálných kořenů podle věty 4.1 určujeme znaménka funkčních hodnot v krajních a v některých vnitřních bodech intervalu I . Tím je interval I rozdělen na vhodné subintervaly, ve kterých zkoumáme splnění předpokladů věty 4.1. Při tomto postupu se snažíme vyhledat takové intervaly $\langle a, b \rangle \subset I$, ve kterých leží právě jeden kořen.

Příklad 4.1 Separujme kořeny rovnice $x^4 - 4x - 1 = 0$ pro $I = (-\infty, \infty)$.

Řešení: Jistě platí $f(-\infty) > 0$, $f(0) < 0$, $f(\infty) > 0$, $f'(x) = 4(x^3 - 1) = 0$ pro $x = 1$, $f'(x) < 0$ pro $x \in (-\infty, 1)$ a $f'(x) > 0$ pro $x \in (1, \infty)$. Z těchto údajů tedy plyne, že interval $(-\infty, 0)$ obsahuje právě jeden kořen, interval $(0, 1)$ žádný a interval $(1, \infty)$ také právě jeden kořen. Tedy zmíněná rovnice má pouze dva reálné kořeny, a to po jednom v intervalech $(-\infty, 0)$, $(1, \infty)$. Protože $f(-1) > 0$ a $f(2) > 0$, pak jsou tyto kořeny dokonce v intervalech $(-1, 0)$ a $(1, 2)$.

Je-li rovnice (4.1) algebraická, tedy když

$$(4.2) \quad f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 = 0, \quad a_n \neq 0,$$

pak dovedeme podle věty 4.2, resp. 4.3, určit hranice všech kořenů (i komplexních), resp. hranice kladných a záporných kořenů.

Věta 4.2 Všechny kořeny algebraické rovnice (4.2) jsou rozloženy v mezikruží

$$\frac{|a_0|}{B + |a_0|} \leq |x| \leq 1 + \frac{A}{|a_n|},$$

kde $A = \max \{|a_{n-1}|, |a_{n-2}|, \dots, |a_0|\}$ a $B = \max \{|a_n|, |a_{n-1}|, \dots, |a_1|\}$.

Důkaz je uveden v [5] a platí samozřejmě i pro komplexní kořeny.

Věta 4.3 Mějme dánu rovnici (4.2), kde $a_n > 0$, a_{n-k} je první záporný koeficient a A je maximum z absolutních hodnot všech záporných koeficientů. Potom všechny kladné kořeny rovnice (4.2) jsou menší než $R = 1 + \sqrt[k]{A/a_n}$.

Důkaz viz [5].

Dolní hranici kladných kořenů rovnice (4.2) určíme pomocí horní hranice kladných kořenů rovnice

$$(4.3) \quad a_n + a_{n-1} y + a_{n-2} y^2 + \dots + a_0 y^n = 0,$$

kterou dostaneme ze (4.2) po substituci $x = 1/y$. Je-li R' horní hranice kladných kořenů rovnice (4.3), pak $r = 1/R'$ je dolní hranice kladných kořenů rovnice (4.2).

Hranice záporných kořenů rovnice (4.2) určíme podle věty 4.3 pomocí rovnice, kterou ze (4.2) získáme substitucí $x = -y$.

Příklad 4.2 Určíme hranice kladných a záporných kořenů rovnice $x^4 + 4x^3 + 3x^2 - 2x - 27 = 0$.

Řešení: Podle věty 4.3 je $a_4 = 1 > 0$, $k = 3$, $A = 27$, tedy $R = 1 + \sqrt[3]{27} = 4$. Provedeme-li v zadané rovnici substituci $x = 1/y$, pak dostaneme rovnici $27y^4 + 2y^3 - 3y^2 - 4y - 1 = 0$, jejíž horní hranice kladných kořenů je $1 + \sqrt{4/27}$, tedy dolní hranice kladných kořenů zadané rovnice je $r = 1/(1 + \sqrt{4/27}) = 0,722$. Užitím substituce $x = -y$ můžeme zcela analogicky zjistit, že záporné kořeny patří do intervalu $(-28, -0,9)$ (ověřte!).

Při separaci reálných kořenů algebraické rovnice (4.2) je velice užitečná tzv. Sturmova věta, podle které dovedeme určit přesný počet reálných kořenů na nějakém intervalu.

Věta 4.4 (Sturmova). Nechť algebraická rovnice (4.2) má pouze jednoduché kořeny. Dále nechť $f_1(x) = f'(x)$, $f_2(x)$ je zbytek při dělení $f(x)/f_1(x)$ násobený číslem -1 , $f_3(x)$ je zbytek při dělení $f_1(x)/f_2(x)$ násobený číslem -1 atd. Postup ukončíme až dostaneme konstantu. Tím získáme Sturmovo posloupnost

$$f(x), f_1(x), f_2(x), \dots$$

Potom počet reálných kořenů rovnice (4.2) na intervalu $\langle a, b \rangle$ je roven rozdílu mezi počtem znaménkových změn ve Sturmově posloupnosti pro $x = a$ a $x = b$.

Důkaz viz např. [Kuroš: Kurs vyšší algebry, Moskva 1946].

Při výpočtu Sturmovy posloupnosti můžeme každou funkci f, f_1, f_2, \dots násobit libovolnou kladnou konstantou.

Má-li rovnice (2.4) násobné kořeny, pak je možné získat posloupnost $f, f_1, \dots, f_k, 0$, ale f_k nebude konstanta. Dělíme-li všechny členy posloupnosti funkcí f_k , dostaneme novou posloupnost, na kterou můžeme užít Sturmovu větu a dostaneme počet kořenů rovnice (4.2) na intervalu $\langle a, b \rangle$ bez uvažování jejich násobnosti.

Příklad 4.3 Určíme počet reálných kořenů rovnice $4x^3 - 2x^2 - 4x - 3 = 0$.

Řešení: Sturmova posloupnost je: $f(x) = 4x^3 - 2x^2 - 4x - 3$, $f_1(x) = 3x^2 - x - 1$, $f_2(x) = 26x + 29$, $f_3(x) = -1$.

	$-\infty$	0	2	∞
$\operatorname{sgn} f(x)$	-	-	+	+
$\operatorname{sgn} f_1(x)$	+	-	+	+
$\operatorname{sgn} f_2(x)$	-	+	+	+
$\operatorname{sgn} f_3(x)$	-	-	-	-
Počet znaménkových změn	2	2	1	1

V uvedené tabulce 4.1 jsou sestavena znaménka hodnot posloupnosti v bodech $-\infty, 0, 2, \infty$, z čehož vidíme, že daná rovnice má jeden reálný kořen v intervalu $\langle 0, 2 \rangle$.

Tab. 4.1

4.2 Výpočet kořenů nelineární rovnice

Předpokládejme, že pro rovnici (4.1) jsme již provedli separaci reálných kořenů a že stojíme před druhou fází přibližného výpočtu izolovaného kořenu rovnice (4.1). Známe tedy konečný interval $\langle a, b \rangle$, ve kterém leží právě jeden reálný kořen ξ rovnice $f(x) = 0$, přičemž $f(x)$ je na $\langle a, b \rangle$ alespoň spojitá. V dalším vyložíme některé iterační metody pro výpočet přibližné hodnoty kořenu a přitom řekneme, za jakých podmínek a jak rychle konvergují. Rychlosť i sama konvergence závisí u některých metod na kvalitě počáteční approximace, ale existují i tzv. vždy konvergentní metody, jejichž konvergence na počáteční approximaci nezávisí. Takové metody konvergují pomalu a hodí se především pro nalezení vhodné počáteční approximace pro metodu rychlejší.

Někdy je vhodné hledat počáteční approximaci grafickou cestou, kdy rovnici (4.1) nahradíme ekvivalentní rovnici

$$(4.4) \quad g(x) = h(x),$$

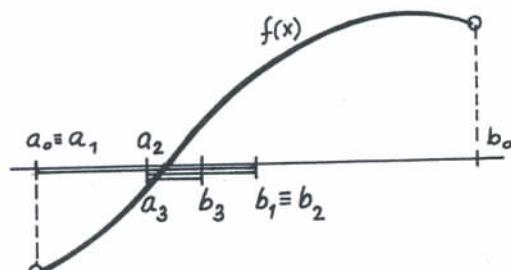
kde $g(x)$ a $h(x)$ jsou jednodušší funkce, jejichž grafy snadno sestrojíme. Pak approximace hledaného kořenu je x -ová souřadnice průsečíku těchto grafů. Grafická metoda je velmi závislá na vhodnosti volby funkcí ve (4.4) a na zkušenostech řešitele.

4.2.1 Metoda půlení intervalu (bisekce)

Je to první ze vždy konvergentních metod pro hledání reálného kořene ξ rovnice (4.1), který je separován na intervalu $\langle a, b \rangle$. Přitom předpokládáme, že

$$(4.5) \quad f(a) \cdot f(b) < 0,$$

tzn., že funkční hodnoty v krajních bodech intervalu $\langle a, b \rangle$ mají opačné znaménka.



Obr. 4.2

Metodou půlení intervalu vytváříme posloupnost uzavřených intervalů (viz obrázek 4.2)

$$\langle a_0, b_0 \rangle \supset \langle a_1, b_1 \rangle \supset \langle a_2, b_2 \rangle \supset \dots \supset \langle a_3, b_3 \rangle \supset \dots,$$

kde $a_0 = a$, $b_0 = b$, takto: Postupně pro $k = 1, 2, \dots$ vypočítáme střed intervalu $\langle a_{k-1}, b_{k-1} \rangle$

$$(4.6) \quad x_k = (b_{k-1} + a_{k-1})/2$$

a je-li $f(x_k) = 0$, je x_k hledaným kořenem ξ . Je-li $f(x_k) \neq 0$, potom za interval $\langle a_k, b_k \rangle$ vybereme ten z intervalů $\langle a_{k-1}, x_k \rangle$, $\langle x_k, b_{k-1} \rangle$, v jehož krajních bodech má funkce $f(x)$ opačná znaménka, tedy platí $f(a_k) \cdot f(b_k) < 0$.

Proces půlení končí, je-li $f(x_k) = 0$, nebo když délka intervalu $\langle a_k, b_k \rangle$ je menší než předem dané $\varepsilon > 0$. Pak approximace x_k určuje hledaný kořen ξ s přesností ε , protože

$$(4.7) \quad |x_k - \xi| < (b_k - a_k)^k = (b - a)/2^k < \varepsilon.$$

Ze (4.7) plyne $\lim_{k \rightarrow \infty} (b_k - a_k) = 0$, tedy $\lim_{k \rightarrow \infty} a_k = \lim_{k \rightarrow \infty} b_k = L$. Snadno zjistíme, že $f(L) = 0$, neboť limitním přechodem v nerovnosti $f(a_k) \cdot f(b_k) < 0$ dostaneme v důsledku spojitosti funkce $f(x)$ nerovnost $f^2(L) \leq 0$, tedy $f(L) = 0$, čili $L = \xi$. Protože při každém kroku platí $a_k \leq x_k \leq b_k$, pak $\lim_{k \rightarrow \infty} x_k = \xi$. Čili jedinou podmínkou konvergence metody půlení intervalu je spojitost funkce $f(x)$.

Příklad 4.4 Metodou půlení intervalu vypočítejme přibližnou hodnotu kořene rovnice $1 - 4x - \operatorname{tg} x = 0$ v intervalu $\langle 0, 1 \rangle$ s přesností $\varepsilon = 10^{-3}$.

Řešení: Předpoklady metody jsou splněny, protože $f(x) = 1 - 4x - \operatorname{tg} x$ je na $\langle 0, 1 \rangle$ spojitá a $f(0) \cdot f(1) < 0$. Numerické hodnoty výpočtu jsou uvedeny v tabulce 4.2.

k	a_k	b_k	x_k	$b_k - a_k$
0	0	1		1
1	0	0,500 00	0,500 00	0,5
2	0	0,250 00	0,250 00	0,25
3	0,125 00	0,250 00	0,125 00	0,125
4	0,187 50	0,250 00	0,187 50	0,0625
5	0,187 50	0,218 75	0,218 75	0,031 25
6	0,187 50	0,203 12	0,203 12	0,015 62
7	0,195 31	0,203 12	0,195 31	0,007 81
8	0,199 22	0,203 12	0,199 22	0,003 90
9	0,199 22	0,201 17	0,201 17	0,001 95
10	0,199 22	0,200 19	0,200 19	0,000 97

Tab. 4.2

Hodnota kořene na sedm desetinných míst je 0,199 4624.

Z příkladu vidíme, že rychlosť konvergencie je velmi malá, protože každý krok zpřesní výsledek pouze o jednu dvojkovou cifru.

Kromě jednoduchosti a snadného splnění podmínek konvergencie je výhoda této metody i v tom, že ji můžeme užít i tehdy, když v $\langle a, b \rangle$ je více kořenů. Počítané approximace pak konvergují k jednomu z nich.

4.2.2 Metoda prosté iterace

Je to jedna z nejdůležitějších metod numerického řešení rovnice (4.1), při které nahradíme rovnici (4.1) ekvivalentní rovnici ve tvaru

$$(4.8) \quad x = \varphi(x),$$

tj. rovnici, ve které je $\varphi(x)$ spojitou funkcí v $\langle a, b \rangle$ a která má v intervalu $\langle a, b \rangle$ tentýž izolovaný kořen ξ jako rovnice (4.1). Rovnici (4.8) lze sestrojit více způsoby, např. algebraickými úpravami rovnice (4.1) nebo pouhým přičtením x k oběma stranám rovnice (4.1) nebo nějakým mnohem promyšlenějším postupem.

Zvolme dále nějakým způsobem $x_0 \in \langle a, b \rangle$ za počáteční approximaci kořenu a podle iterační rovnice

$$(4.9) \quad x_k = \varphi(x_{k-1})$$

počítejme postupně posloupnost approximací x_1, x_2, x_3, \dots . Jestliže posloupnost $\{x_k\}$ je konvergentní, tedy je-li $\lim_{k \rightarrow \infty} x_k = L$, pak v důsledku spojitosti funkce $\varphi(x)$ můžeme přejít ve (4.9) k limitě, a potom $\lim_{k \rightarrow \infty} x_k = \varphi(\lim_{k \rightarrow \infty} x_k)$, čili L je hledaný kořen ξ , který lze podle rovnice (4.9) určit s libovolnou přesností.

Zbývá tedy zodpovědět otázku, za jakých předpokladů a s jakou rychlostí konverguje posloupnost $\{x_k\}$. Na tuto otázku odpovídá následující věta.

Věta 4.5 Nechť funkce $\varphi(x)$ je spojitá a diferencovatelná v $\langle a, b \rangle$ a nechť $\varphi(x)$ zobrazuje interval do sebe, tzn. že pro $x \in \langle a, b \rangle$ je $\varphi(x) \in \langle a, b \rangle$. Existuje-li číslo q tak, že

$$(4.10) \quad |\varphi'(x)| \leq q < 1 \quad \text{pro } x \in \langle a, b \rangle,$$

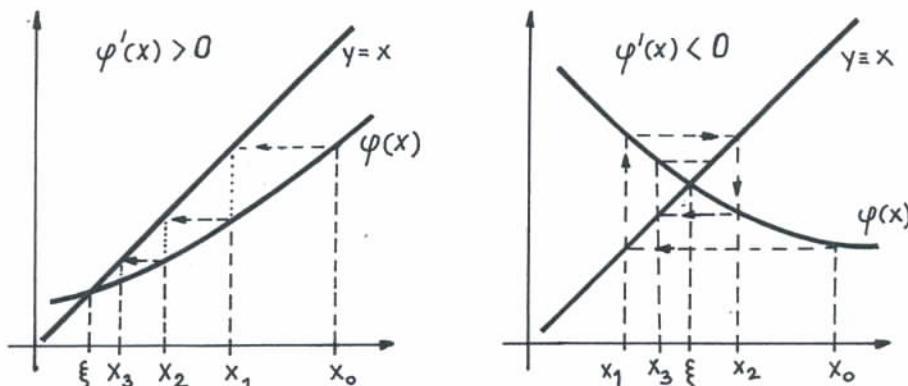
potom iterační proces (4.9), tedy posloupnost $\{x_n\}$, konverguje nezávisle na volbě počáteční approximace $x_0 \in \langle a, b \rangle$, a to k jedinému kořenu ξ rovnice (4.8) na intervalu $\langle a, b \rangle$.

Přitom platí

$$(4.11) \quad |x_k - \xi| \leq \frac{q}{1-q} |x_k - x_{k-1}| \leq \frac{q^k}{1-q} |x_1 - x_0|.$$

Důkaz této věty je uveden v [5], [9], [16], [19], [22] a dalších.

Na obrázcích 4.3 je geometrická interpretace iterační metody (4.9) pro případ $\varphi'(x) > 0$ a $\varphi'(x) < 0$.



Obr. 4.3

Příklad 4.5 Metodou prosté iterace řešme úlohu z příkladu 4.4.

Řešení: Danou rovnici převedeme na tvar $x = (1 - \operatorname{tg} x)/4$, tedy $\varphi(x) = (1 - \operatorname{tg} x)/4$. Funkce $\varphi(x)$ je v intervalu $\langle 0, 1 \rangle$ diferencovatelná a platí $|\varphi'(x)| < 1$ pro $x \in \langle 0, 1 \rangle$, ale nezobrazuje interval $\langle 0, 1 \rangle$ do sebe. Proto se omezíme např. na interval $\langle 0, 0.5 \rangle$, kde jsou jistě splněny všechny předpoklady věty 4.5. Zvolme tedy $x_0 = 0.5$ a počítejme approximace x_1, x_2, \dots podle vzorce (4.9); numerické výsledky jsou v tabulce 4.3.

k	x_k
0	0,5
1	0,1134 24
2	0,2215 22
3	0,1936 96
4	0,2009 61
5	0,1990 72
6	0,1995 64

Protože jistě platí

$$|\varphi'(x)| < 0,33 < 1 ,$$

pak podle vzorce (4.11) dostaneme

$$|x_6 - \xi| \leq \frac{0,33^6}{1 - 0,33} \cdot 0,4 = 0,000\ 77 = \\ = 0,77 \cdot 10^{-3} < 10^{-3} ,$$

tedy podmínka přesnosti approximace je splněna.

Tab. 4.3

4.2.3 Newtonova metoda (metoda tečen)

Řešme opět rovnici (4.1), o které víme, že uvnitř intervalu $\langle a, b \rangle$ leží právě jeden izolovaný reálný kořen ξ . Dále předpokládejme, že $f'(x)$ a $f''(x)$ jsou spojité a zachovávají na intervalu $\langle a, b \rangle$ znaménko a že $x_k \in \langle a, b \rangle$ je k-tá approximace kořene ξ . Když položíme $\xi = x_k + h_k$, pak na základě Taylorovy věty zapíšeme

$$(4.12) \quad 0 = f(\xi) = f(x_k + h_k) = f(x_k) + h_k f'(x_k) + \frac{h_k^2}{2} f''(x_k + \Theta h_k) ,$$

kde Θ je nějaké číslo z intervalu $(0,1)$.

Zanedbáme-li ve (4.12) poslední člen, který představuje zbytek Taylorova rozvoje, pak

$$h_k \approx -f(x_k)/f'(x_k) ,$$

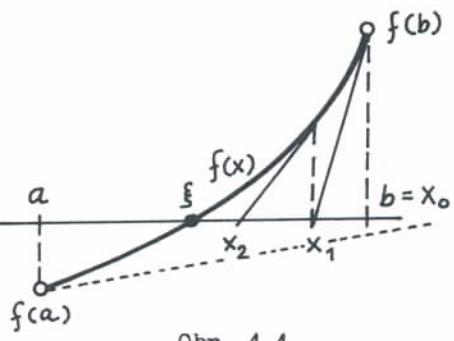
tedy číslo $x_k + h_k = x_{k+1}$, resp.

$$(4.13) \quad x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

bude další approximací kořene ξ . Iterační rovnici (4.13) nazýváme Newtonovou iteracní metodou.

Také Newtonovu metodu můžeme interpretovat geometricky, viz obr. 4.4.

Předpokládejme, že $f(a) < 0$, $f(b) > 0$, $f'(x) > 0$ a $f''(x) > 0$ pro $x \in \langle a, b \rangle$. Položíme-li $x_0 = b$, pak číslo x_1 vypočítané podle (4.13) dostaneme geometricky jako průsečík tečny ke grafu funkce $f(x)$ v bodě $[x_0, f(x_0)]$ s osou x . Další approximaci x_2 můžeme interpretovat jako průsečík tečny v bodě $[x_1, f(x_1)]$ s osou x atd.



Obr. 4.4

Všimněte si ještě, že Newtonova metoda je v podstatě metodou prosté iterace, při které zvolíme $\varphi(x) = x - f(x)/f'(x)$, tedy podmínky konvergence metody jsou v zásadě dány větou 4.5. Vraťme se ještě k obr. 4.4. Postup, který jsme popsali, začal od počáteční approximace $x_0 = b$, takže byla splněna nerovnost $f(x_0) \cdot f''(x_0) > 0$. Zvolíme-li naopak $x_0 = a$, čili platí nerovnost $f(x_0) \cdot f''(x_0) < 0$, pak se může stát (viz obr. 4.4), že $x_1 \notin \langle a, b \rangle$ a v metodě nelze pokračovat, protože jsme porušili základní předpoklad věty 4.5. Takže volba $x_0 \in \langle a, b \rangle$ je podstatnou podmínkou konvergence Newtonovy metody.

Věta 4.6 Nechť v intervalu $\langle a, b \rangle$ existuje právě jeden reálný kořen ξ rovnice (4.1). Dále nechť $f(a) \cdot f(b) < 0$, $f'(x)$ a $f''(x)$ jsou spojité a zachovávají v $\langle a, b \rangle$ znaménko a nechť pro počáteční approximaci $x_0 \in \langle a, b \rangle$ platí

$$f(x_0) \cdot f''(x_0) > 0.$$

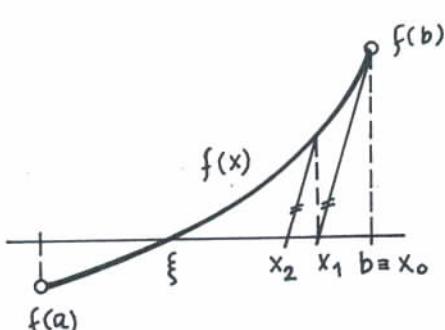
Pak posloupnost $\{x_k\}$ počítaná podle vzorce (4.13) konverguje ke kořenu ξ , a přitom platí

$$(4.14) \quad |x_k - \xi| < \frac{M_2}{2m_1} (x_k - x_{k-1})^2,$$

$$\text{kde } M_2 = \max_{x \in \langle a, b \rangle} |f''(x)| \text{ a } m_1 = \min_{x \in \langle a, b \rangle} |f'(x)|.$$

Důkaz viz [23], [9], [25] a další.

Newtonova metoda konverguje velmi rychle, ale závisí na vhodné volbě počáteční approximace, při nevhodné volbě nemusí konvergovat vůbec.



Obr. 4.5

Jestliže $f'(x)$ se na intervalu $\langle a, b \rangle$ mění málo, pak můžeme položit $f'(x) \approx f'(x_0)$ a tím dostaneme tzv. modifikovanou Newtonovu metodu

$$(4.15) \quad x_{k+1} = x_k - \frac{f(x_k)}{f'(x_0)},$$

která je výpočetně jednodušší, ale má sníženou rychlosť konvergence. Geometrická interpretace viz obr. 4.5.

Příklad 4.6 Newtonovou metodou řešme úlohu z příkladu 4.4.

Řešení: Protože $f(x) = 1 - 4x - \operatorname{tg} x$, $f'(x) = -4 - 1/\cos^2 x$, $f''(x) = -2 \sin x \cos^{-3} x$, pak $f(0) > 0$, $f(1) < 0$, $f'(x) < 0$ a $f''(x) < 0$ v intervalu $\langle 0, 1 \rangle$, tedy zvolíme $x_0 = 1$ a approximace kořene ξ počítáme podle vzorce

$$x_{k+1} = x_k + (1 - 4x - \operatorname{tg} x)/(4 + 1/\cos^2 x),$$

pak $x_1 = 0,386251$, $x_2 = 0,202009$, $x_3 = 0,199463$.

Odhadněme ještě přesnost approximace x_3 podle vzorce (4.14). Vyšetřováním $f'(x)$ a $f''(x)$ na intervalu $\langle 0, 1 \rangle$ dostaneme $M_2 = 10,7$ a $m_1 = 5$, tedy

$$|x_3 - \xi| < \frac{10,7}{10} (x_3 - x_2)^2 = 0,69 \cdot 10^{-5} < 10^{-3}.$$

Porovnáme-li výsledky s řešením příkladů 4.4 a 4.5, vidíme, že konvergencie Newtonovy metody je mnohem rychlejší.

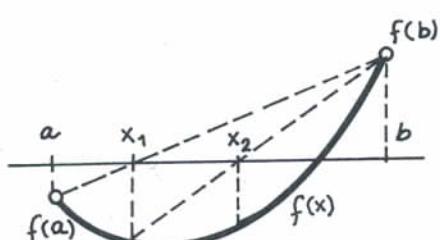
4.2.4 Metoda regula falsi (tětiv)

Je jednou z nejstarších metod řešení rovnice (4.1) na intervalu $\langle a, b \rangle$, pro který platí $f(a) \cdot f(b) < 0$. Při metodě regula falsi nahradíme funkci $f(x)$ na intervalu $\langle a, b \rangle$ interpolačním polynomem 1. stupně $L_1(x)$ s uzlovými body a, b a najdeme první approximaci x_1 kořene ξ řešením rovnice $L_1(x) = 0$, tedy

$$L_1(x) = \frac{x - b}{a - b} f(a) + \frac{x - a}{b - a} f(b) = 0,$$

čili

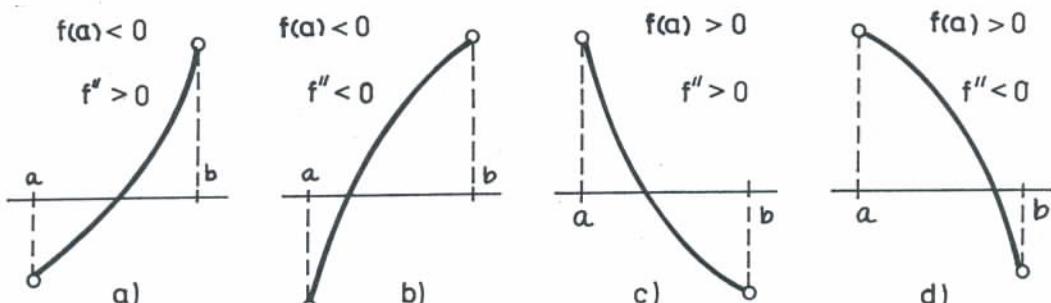
$$(4.16) \quad x_1 = x = a - \frac{b - a}{f(b) - f(a)} f(a).$$



Obr. 4.6

Protože $f(a) \cdot f(b) < 0$, pak bod x_1 dělí interval $\langle a, b \rangle$ na dva subintervaly $\langle a, x_1 \rangle$, $\langle x_1, b \rangle$. Uvedený postup opakujeme na tom z nich, v jehož koncových bodech má $f(x)$ opačná znaménka a tím dostaneme x_2 atd. Geometricky je tento postup demonstrován na obr. 4.6.

Aby uvedený postup byl konvergentní, je po- stávající, aby v $\langle a, b \rangle$ ležel právě jeden ko- řen a aby $f''(x)$ zachovala v $\langle a, b \rangle$ znamén- ko. Vzhledem ke znaménku čísla $f(a)$ a znaménku $f''(x)$ mohou nastat čtyři varian- ty, viz obr. 4.7.



Obr. 4.7

Pro případy a), d) se základní vzorec (4.16) užívá postupně na intervalech $\langle x_k, b \rangle$, tedy položíme-li $x_0 = a$, potom má metoda regula falsi vzorec

$$(4.17) \quad x_{k+1} = x_k - \frac{b - x_k}{f(b) - f(x_k)} \cdot f(x_k), \quad k = 0, 1, 2, \dots$$

Obdobně pro b), c), kdy se základní vzorec užívá na intervalech $\langle a, x_k \rangle$, dostaneme ze (4.16) pro $x_0 = b$ po vhodné úpravě vzorec

$$(4.18) \quad x_{k+1} = x_k - \frac{x_k - a}{f(x_k) - f(a)} \cdot f(x_k), \quad k = 0, 1, 2, \dots$$

Věta 4.7 Nechť v $\langle a, b \rangle$ je právě jeden reálný kořen ξ rovnice (4.1), $f(a) \cdot f(b) < 0$ a nechť $f''(x)$ zachovává na $\langle a, b \rangle$ znaménko. Platí-li $f(b) \cdot f''(b) > 0$, resp.

$f(a) \cdot f''(a) > 0$, pak posloupnost $\{x_k\}$, sestrojená podle vzorce (4.17), resp. (4.18), konverguje k jedinému kořenu ξ rovnice (4.1) na intervalu $\langle a, b \rangle$, a přitom platí

$$(4.19) \quad |\xi - x_k| \leq \frac{M_1 - m_1}{m_1} |x_k - x_{k-1}|,$$

kde M_1 a m_1 jsou maximální a minimální hodnoty funkce $|f'(x)|$ v intervalu $\langle a, b \rangle$.

Důkaz viz např. [10].

Příklad 4.7 Metodou regula falsi řešte úlohu z příkladu 4.4.

k	x_k	$f(x_k)$
0	0.	1.000 000
1	0.179 940	0.098 332
2	0.197 260	0.011 010
3	0.199 211	0.001 267
4	0.199 434	0.000 143

Tab. 4.4

Řešení: Z příkladu 4.6 plyne, že $f(1) < 0$ a $f''(x) < 0$ pro $x \in \langle 0, 1 \rangle$, tedy budeme vést výpočet podle vzorce (4.17), pro který $x_0 = 0$, $b = 1$, $f(b) = -4,557 408$. Výsledky viz tab. 4.4.

Odhad chyby provedeme podle vzorce (4.19) pro $m_1 = 5$ a $M_1 = 7,5$, tedy

$$|\xi - x_4| \leq \frac{7,5 - 5}{5} |x_4 - x_3| = 0,11 \cdot 10^{-3} < 10^{-3}.$$

Metoda regula falsi je velmi užitečná, protože se dá užít i v případech, kdy kromě nerovnosti $f(a) \cdot f(b) < 0$ nemáme o funkci $f(x)$ žádné další informace. Konvergence může pak být ale velmi pomalá.

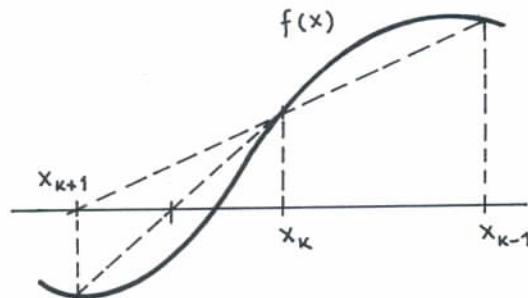
4.2.5 Metoda sečen

Jestliže v Newtonově metodě nahradíme ve vzorci (4.13) derivaci $f'(x_k)$ podílem diferencí

$$f'(x_k) \approx \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}},$$

pak dostaneme vzorec metody sečen

$$(4.20) \quad x_{k+1} = x_k - f(x_k) \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}.$$



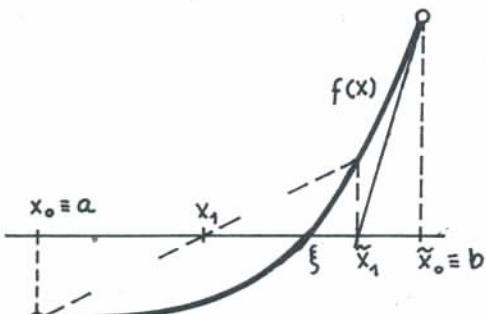
Obr. 4.8

Je to dvoubodová iterační metoda, protože k určení x_{k+1} užíváme dvou předcházejících approximací, viz obr. 4.8.

K témuž výsledku bychom dospěli úpravou metody regula falsi, když upustíme od požadavku, aby funkce $f(x)$ měla v krajních bodech intervalu opačná znaménka. Ze cvičených důvodů doporučujeme čtenáři, aby si tuto modifikaci prověřil.

4.2.6 Kombinovaná metoda

Řešme rovnici (4.1) na intervalu $\langle a, b \rangle$ za předpokladů: $f(a) \cdot f(b) < 0$, $f'(x)$ a $f''(x)$ jsou spojité a zachovávají na $\langle a, b \rangle$ znaménko. Spojením metody Newtonovy a metody regula falsi dostaneme kombinovanou metodu, při které je kořen ξ approximován zdola i shora čísly x_k a \tilde{x}_k . Vzorce metody regula falsi i výběr počátečních approximací x_0 , \tilde{x}_0 pro obě metody závisejí jistým způsobem na znaménkách $f(x)$, $f'(x)$ a $f''(x)$ v některých bodech nebo na intervalu $\langle a, b \rangle$, proto se omezíme jenom na případ, kdy $f(a) < 0$, $f(b) > 0$, $f'(x) > 0$ a $f''(x) > 0$ pro $x \in \langle a, b \rangle$, viz obr. 4.9. Vzorce kombinované metody uvedeme jenom pro tento případ, protože v ostatních případech jsou vzorce zcela analogické a čtenář si je snadno sestaví sám.



Obr. 4.9

Položme tedy $x_0 = a$, $\tilde{x}_0 = b$ a pak již zcela logicky ze vzorců (4.17) a (4.13) plynou vzorce kombinované metody

$$\tilde{x}_{k+1} = \tilde{x}_k - \frac{f(\tilde{x}_k)}{f'(\tilde{x}_k)},$$

(4.21)

$$x_{k+1} = x_k - \frac{x_{k+1} - x_k}{f(x_{k+1}) - f(x_k)} f(x_k),$$

pro $k = 0, 1, 2, \dots$. V těchto vzorcích se metoda tětiv užívá vždy na intervalu $\langle x_k, \tilde{x}_{k+1} \rangle$.

Velkou výhodou kombinované metody je ta skutečnost, že hledané řešení je approximováno současně zdola i shora, protože metoda tětiv i metoda tečen approximují kořen ξ z jedné strany. Tedy pro $k=1, 2, \dots$ platí $\xi \in \langle x_k, \tilde{x}_k \rangle$.

Kombinovaná metoda konverguje vždy, když jsou splněny postačující podmínky konvergence metody tětiv a metody tečen.

4.3 Výpočet kořenů algebraické rovnice

Věnujme se nyní hledání kořenů algebraických rovnic (4.2), pro které byla vymyšlena řada speciálních metod, uvedených v [5], [8], [9], [14], [19], [22] a dalších.

4.3.1 Metoda Laguerrova

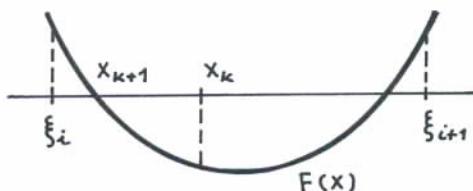
Předpokládejme, že kořeny polynomu $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$ jsou všechny různé reálné, tedy je můžeme seřadit tak, aby platilo

$$\xi_1 < \xi_2 < \xi_3 < \dots < \xi_{n-1} < \xi_n.$$

Tím jsme definovali posloupnost intervalů $I_0 = (-\infty, \xi_0)$, $I_1 = (\xi_1, \xi_2)$, ..., $\dots, I_{n-1} = (\xi_{n-1}, \xi_n)$, $I_n = (\xi_n, \infty)$. Nechť x_k je approximace kořenu ξ_i (resp. ξ_{i+1}) polynomu $f(x)$ a platí $x_k \in I_i$. Podstata Laguerrovy metody záleží v tom, že sestrojíme polynom druhého stupně se dvěma reálnými kořeny na intervalu I_i , z nichž jeden leží blíže hledanému kořenu než approximace x_k . Takový polynom druhého stupně může mít tvar

$$(4.22) \quad F(x) = \left\{ \sum_{j=1}^n \left(\frac{\lambda - \xi_j}{x_k - \xi_j} \right)^2 - \left(\frac{\lambda - x}{x_k - x} \right)^2 \right\} (x_k - x)^2 = \\ = (x_k - x) \sum_{j=1}^n \left(\frac{\lambda - \xi_j}{x_k - \xi_j} \right)^2 - (\lambda - x)^2,$$

kde λ je volitelný parametr. Tvar (4.22) zdánlivě závisí na kořenech polynomu $f(x)$, ale tuto závislost lze vyloučit. Polynom (4.22) má skutečně předepsané vlastnosti, protože pro $\lambda \neq x_k$ plyne z druhého (upraveného) vyjádření $F(x)$ ve vzorci (4.22), že $F(x_k) < 0$; dále platí $F(\xi_i) > 0$, $F(\xi_{i+1}) > 0$, což lze snadno ověřit z prvního vyjádření $F(x)$ ve vzorci (4.22). Průběh polynomu $F(x)$ je naznačen na obrázku 4.10.



Obr. 4.10

Naším úkolem je vyloučit ze (4.22) kořeny polynomu $f(x)$ a najít takovou hodnotu parametru λ , aby jeden z kořenů polynomu $F(x)$ byl co nejbliže kořenu ξ_i nebo ξ_{i+1} . Dále by následovaly rozsáhlé a únavné výpočty, které přeskočíme a řekneme pouze jejich efekt. Takže po vyloučení kořenů ξ_1 až ξ_n z $F(x)$ a určení parametru λ lze zjistit, že hledaný kořen polynomu $F(x)$, který je

lepší approximací kořenu ξ_i (resp. ξ_{i+1}) než x_k , je vyjádřen vzorcem

$$(4.23) \quad x_{k+1} = x_k - \frac{n \cdot f(x_k)}{f'(x_k) \pm \sqrt{H(x_k)}},$$

kde $H(x) = (n-1) [(n-1)(f(x))^2 - n \cdot f(x) \cdot f''(x)]$

je při reálných kořenech polynomu $f(x)$ nezáporná funkce, tedy $\sqrt{H(x)}$ je reálné číslo. Znaménko +, - ve (4.23) vybíráme tak, aby souhlasilo se znaménkem čísla $f'(x_k)$.

Vzorec (4.23) definuje pro $k = 0, 1, 2, \dots$ Laguerrovu iteraci, pro kterou je x_0 počáteční approximací některého kořenu polynomu $f(x)$.

Posloupnost $\{x_k\}$ počítaná podle (4.23) je z libovolného x_0 konvergentní k některému z kořenů $\xi_1, \xi_2, \dots, \xi_n$. Je-li $x_0 < \xi_1$, pak $\{x_k\} \rightarrow \xi_1$, je-li $x_0 \in (\xi_i, \xi_{i+1})$, pak $\{x_k\} \rightarrow \xi_i$ nebo $\{x_k\} \rightarrow \xi_{i+1}$ a je-li $x_0 > \xi_n$, pak $\{x_k\} \rightarrow \xi_n$.

Laguerrova metoda je velmi účinná a rychle konverguje, jsou-li všechny kořeny jednoduché a reálné. Je-li některý kořen vícenásobný, pak metoda sice konverguje, ale v okolí násobného kořenu je velmi pomalá.

Metoda Laguerrova se dá použít i pro komplexní kořeny, ale touto modifikací se nebude vůbec zabývat, protože pro tyto případy je o konvergenci Laguerrovy metody

známo velice málo. Praktické zkušenosti však ukazují, že případy, kdy je konvergence porušena, jsou velmi řídké, viz [23].

Příklad 4.8 Vypočítejme Laguerrovou metodou kořeny rovnice $x^3 - 5x^2 - 17x + 21 = 0$.

k	x_k
0	1000.0
1	7.478 5207
2	7.000 1011
3	7.000 0000

Tab. 4.5

Řešení: Přesné kořeny této rovnice jsou 1, -3, 7. Zvolíme-li za počáteční approximaci $x_0 = 10^6$ (což je velmi špatná approximace), pak podle (4.23) dostaneme již po třech krocích největší kořen s přesností sedmi desetinných míst, viz tab. 4.5.

4.3.2 Graeffova metoda

Základní myšlenkou Graeffovy metody je nahradit rovnici (4.2) takovou rovnicí, která je opět algebraickou rovnicí n-tého stupně, ale její kořeny jsou kvadráty kořenů rovnice (4.2). Jestliže kořeny původní rovnice jsou v absolutní hodnotě různé, pak opakováním tohoto postupu docílíme dostatečně velkou separaci kořenů a můžeme je pak vypočítat z koeficientů, jak uvidíme později.

Nechť tedy pro kořeny rovnice (4.2) platí

$$|\xi_1| > |\xi_2| > |\xi_3| > \dots > |\xi_{n-1}| > |\xi_n|,$$

pak (4.2) má tvar

$$a_n(x - \xi_1)(x - \xi_2) \dots (x - \xi_{n-1})(x - \xi_n) = 0.$$

Provedeme-li v této rovnici substituci $x = -x$, dostaneme po vynásobení číslem $(-1)^n$ rovnici

$$a_n(x + \xi_1)(x + \xi_2) \dots (x + \xi_{n-1})(x + \xi_n) = 0.$$

Vynásobením těchto dvou rovnic dostaneme rovnici

$$a_n^2(x^2 - \xi_1^2)(x^2 - \xi_2^2) \dots (x^2 - \xi_{n-1}^2)(x^2 - \xi_n^2) = 0,$$

ve které ještě položíme $-x^2 = y$ a po vynásobení $(-1)^n$ dostaneme rovnici

$$(4.24) \quad a_n^2(y + \xi_1^2)(y + \xi_2^2) \dots (y + \xi_{n-1}^2)(y + \xi_n^2) = 0,$$

jejíž kořeny $-\xi_1^2, -\xi_2^2, \dots, -\xi_{n-1}^2, -\xi_n^2$ jsou od sebe vzdálenější než kořeny původní rovnice (4.2).

Proveďme nyní tyto operace s rovinicí (4.2), takže vynásobíme polynomy $f(x) \cdot f(-x)$ a dostaneme

$$\left[a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \right] \left[a_n (-x)^n + a_{n-1} (-x)^{n-1} + \dots + a_1 (-x) + a_0 \right].$$

Při násobení dostáváme postupně členy, které obsahují sudé mocniny proměnné x a členy s lichými mocninami x .

Členy se sudými mocninami:

$$\begin{aligned}
 & a_i^2 x^{2i} (-1)^i + \left[a_{i-1} a_{i+1} x^{2i} (-1)^{i+1} + a_{i+1} a_{i-1} x^{2i} (-1)^{i-1} \right] + \\
 & + \left[a_{i-2} a_{i+2} x^{2i} (-1)^{i+2} + a_{i+2} a_{i-2} x^{2i} (-1)^{i-2} \right] + \dots + \\
 & + \left[a_{i-1} a_{i+k} x^{2i} (-1)^{i+k} + a_{i+k} a_{i-k} x^{2i} (-1)^{i-k} \right] + \dots = \\
 & = a_i^2 (-x^2)^i + (-1)^1 2 \cdot a_{i-1} a_{i+1} (-x^2)^i + (-1)^2 2 \cdot a_{i-2} a_{i+2} (-x^2)^i + \dots + \\
 & + (-1)^k 2 \cdot a_{i-k} a_{i+k} (-x^2)^i + \dots ,
 \end{aligned}$$

kde k se zvětšuje tak dlouho, pokud není splněno $a_{i+k} = a_n$ nebo $a_{i-k} = a_0$.

Členy s lichými mocninami:

$$\begin{aligned}
 & \left[a_i a_{i-1} x^{2i-1} (-1)^{i-1} + a_{i-1} a_i x^{2i-1} (-1)^i \right] + \\
 & + \left[a_{i+1} a_{i-2} x^{2i-1} (-1)^{i-2} + a_{i-2} a_{i+1} x^{2i-1} (-1)^{i+1} \right] + \dots + \\
 & + \left[a_{i+k} a_{i-(k+1)} x^{2i-1} (-1)^{i-(k+1)} + a_{i-(k+1)} a_{i+k} x^{2i-1} (-1)^{i+k} \right] + \dots = 0 ,
 \end{aligned}$$

protože obsah každé lomené závorky je nula. Tedy

$$f(x) \cdot f(-x) = f_n y^n + b_{n-1} y^{n-1} + \dots + b_1 y + b_0 ,$$

kde $y = -x^2$ a pro $i = n, n-1, \dots, 1, 0$ je

$$(4.25) \quad b_i = a_i^2 - 2a_{i-1} a_{i+1} + 2a_{i-2} a_{i+2} - 2a_{i-3} a_{i+3} + \dots ,$$

přitom ve (4.25) jsou pouze takové členy, jejichž indexy mají v naší úloze smysl.

Tento proces můžeme pochopitelně opakovat vícekrát. Proto v počáteční rovnici (4.2) zavedeme ještě horní indexy

$$a_n^{(0)} x^n + a_{n-1}^{(0)} x^{n-1} + \dots + a_1^{(0)} x + a_0^{(0)} = 0$$

a po k výše uvedených krocích, tzv. kvadratizacích, dostaneme

$$(4.26) \quad a_n^{(k)} x^n + a_{n-1}^{(k)} x^{n-1} + \dots + a_1^{(k)} x + a_0^{(k)} + a_o^{(k)} = 0 ,$$

kde pro $i = n, n-1, \dots, 1, 0$

$$(4.27) \quad a_i^{(k)} = a_i^{(k-1)} - 2 \cdot a_{i-1}^{(k-1)} a_{i+1}^{(k-1)} + 2 \cdot a_{i-2}^{(k-1)} a_{i+2}^{(k-1)} - 2 \cdot a_{i-3}^{(k-1)} a_{i+3}^{(k-1)} + \dots .$$

Kvadratizační proces ukončíme, jestliže pro některé koeficienty $a_i^{(k)}$ platí

$$a_i^{(k)} \approx \left(a_i^{(k-1)} \right)^2 .$$

Potom předpokládáme, že kořeny jsou dostatečně oddělené, a začneme je vypočítávat.

Označme kořeny poslední rovnice (4.26) x_1, x_2, \dots, x_n , pak

$$(4.28) \quad |\xi_i| \approx \sqrt[2k]{|x_i|} \approx \sqrt[2k]{|a_i^{(k)} / a_{i+1}^{(k)}|}, \quad i = 1, 2, \dots, n.$$

Odvození viz [5], resp. [23]. Nakonec je třeba ještě zjistit, která z hodnot $\pm |\xi_i|$ je kořenem zadáné rovnice, a to buď odhadem, dosazením nebo z jiných informací o rovnici (4.2).

Při užívání této metody na počítači se musíme ještě vypořádat s problémem růstu absolutních hodnot koeficientů nějakým normalizačním opatřením.

Příklad 4.9 Vypočítejme Graeffovou metodou kořeny rovnice z příkladu 4.8.

Řešení: Výpočet koeficientů $a_3^{(k)}, a_2^{(k)}, a_1^{(k)}, a_0^{(k)}$ a kořenů je seřazen v tabulce 4.6.

k	$a_3^{(k)}$	$a_2^{(k)}$	$a_1^{(k)}$	$a_0^{(k)}$	ξ_1	ξ_2	ξ_3
0	1	- 5	- 17	21	5,0	- 3,4	1,2
1	1	59	499	441	7,68	- 2,91	0,94
2	1	2483	196 963	194 481	7,06	- 2,98	0,997
3	1	$5771 \cdot 10^3$	$3783 \cdot 10^7$	$2782 \cdot 10^7$	7,001	- 2,999	0,9997

Tab. 4.6

Graeffova metoda se může užívat i pro případ komplexních kořenů. Podrobněji se s ní může čtenář seznámit v [5].

4.3.3 Metody užívající syntetické dělení

Mějme opět zadánu rovnici $f(x) = 0$ s polynomem n-tého stupně

$$(4.29) \quad f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 .$$

Dělíme-li tento polynom dvojčlenem $(x - x_k)$, pak podílem je polynom stupně $(n - 1)$ a zbytek je reálné číslo, které označíme R_k , tedy

$$(4.30) \quad f(x) = (x - x_k) (b_{n-1} x^{n-1} + b_{n-2} x^{n-2} + \dots + b_1 x + b_0) + R_k .$$

Přitom si všimněte, že ze (4.30) plyne $R_k = f(x_k)$.

Toto dělení se nazývá syntetické. Abychom odvodili rekurence pro výpočet koeficientů $b_{n-1}, b_{n-2}, \dots, b_1, b_0$ a zbytku R_k , provedeme v (4.30) naznačené násobení a porovnáme koeficienty stejných mocnin (4.30) a (4.29). Dostaneme

$$a_n = b_{n-1}$$

$$a_i = b_{i-1} - x_k b_i, \quad \text{pro } i = n-1, n-2, \dots, 2, 1,$$

$$a_0 = R_k - x_k b_0$$

a z toho plyne

$$(4.31) \quad \begin{aligned} b_{n-1} &= a_n, \\ b_i &= a_{i+1} + x_k b_{i+1}, \quad \text{pro } i = n-2, n-3, \dots, 1, 0, \\ R_k &= a_0 + x_k b_0. \end{aligned}$$

Jestliže podíl ve (4.30) ještě jednou vydělíme dvojčlenem $(x - x_k)$, potom bude podílem polynom stupně $(n-2)$ a zbytek je číslo R'_k , tedy

$$\begin{aligned} (b_{n-1}x^{n-1} + b_{n-2}x^{n-2} + \dots + b_0) &= \\ = (x - x_k)(c_{n-2}x^{n-2} + c_{n-3}x^{n-3} + \dots + c_0) + R'_k. \end{aligned}$$

Dosadíme-li do (4.30), dostaneme

$$(4.32) \quad f(x) = (x - x_k)^2(c_{n-2}x^{n-2} + c_{n-3}x^{n-3} + \dots + c_1x + c_0) + (x - x_k)R'_k + R_k,$$

z čehož plyne, že $R'_k = f'(x_k)$. Je zřejmé, že koeficienty c_{n-2} až c_0 vypočítáme z rovnice

$$(4.33) \quad \begin{aligned} c_{n-2} &= b_{n-1}, \\ c_i &= b_{i+1} + x_k c_{i+1}, \quad \text{pro } i = n-3, n-4, \dots, 1, 0, \\ R'_k &= b_0 + x_k c_0, \end{aligned}$$

které se odvodí obdobně jako rovnice (4.31).

Nyní ukážeme, jakým způsobem opakování syntetického dělení užit při řešení algebraické rovnice (4.2). Předpokládejme, že $x_0, x_1, x_2, \dots, x_k$ je posloupnost approximací kořenu ξ , pak čísla R_{k-1}, R_k můžeme dosadit do vzorce metody sečen (4.20) a vypočítat následující approximaci

$$(4.34) \quad x_{k+1} = x_k - \frac{x_k - x_{k-1}}{R_k - R_{k-1}} R_k,$$

protože $R_k = f(x_k)$ a $R_{k-1} = f(x_{k-1})$.

Také lze dosazovat čísla R'_k a R_k do vzorce Newtonovy metody (4.13) a počítat následující approximaci kořenu vzorcem

$$(4.35) \quad x_{k+1} = x_k - \frac{R_k}{R'_k}.$$

Obdobně můžeme užívat dělení polynomu (4.29) kvadratickým trojčlenem pro hledání komplexních kořenů rovnice (4.2), viz [23].

Příklad 4.10 Vypočítejme Newtonovou metodou pomocí opakování syntetického dělení jeden z kořenů rovnice z příkladu 4.8.

Řešení: Vyšetřováním první a druhé derivace zjistíme, že na intervalu $<0, 10/6>$ platí $f(0) > 0$, $f(10/6) < 0$, $f'(x) < 0$ a $f''(x) > 0$, čili v tomto intervalu existuje jeden izolovaný kořen. Za počáteční approximaci pro (4.35) zvolíme $x_0 = 0$, přestože bychom z takového bodu neměli startovat. Podle výsledků v tabulce 4.7 ale vidíme, že $x_1 = 1.2$ a pro tuto hodnotu již platí $f(1.2) \cdot f''(1.2) < 0$, takže od tohoto kroku je už konvergence podle věty 4.6 zajištěna. Pro lepší pochopení si celou situaci nakreslete.

Newtonova metoda: $a_3 = 1$, $a_2 = -5$, $a_1 = -17$, $a_0 = 21$.

k	b_2	b_1	b_0	R_k	c_1	c_0	R'_k	x_k
1	1	-5	-17	21	1	-5	-17	1.235 294
2	1	-3.764 706	-21.650 519	-5.744 756	1	-2.529 412	-24.775 086	1.003 418
3	1	-3.996 582	-21.010 242	-0.082 055	1	-2.993 164	-24.013 637	1.000 001

Tab. 4.7

4.4 Řešení nelineárních soustav

Nechť je dána soustava nelineárních rovnic

$$(4.36) \quad \begin{aligned} f_1(x_1, x_2, \dots, x_n) &= 0, \\ f_2(x_1, x_2, \dots, x_n) &= 0, \\ &\dots \\ f_n(x_1, x_2, \dots, x_n) &= 0. \end{aligned}$$

Přitom předpokládejme, že funkce f_1, f_2, \dots, f_n jsou nelineární, spojité a mají v nějaké oblasti G , v níž soustavu řešíme, první parciální derivace a že tzv. Jacobiho matice

$$(4.37) \quad Jf(x_1, x_2, \dots, x_n) = \left(\begin{array}{cccc} \frac{\partial f_1}{\partial x_1}, & \frac{\partial f_1}{\partial x_2}, & \dots, & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1}, & \frac{\partial f_2}{\partial x_2}, & \dots, & \frac{\partial f_2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_n}{\partial x_1}, & \frac{\partial f_n}{\partial x_2}, & \dots, & \frac{\partial f_n}{\partial x_n} \end{array} \right)$$

nemá v této oblasti identicky nulový determinant.

Jestliže proměnné x_1, x_2, \dots, x_n budeme považovat za n -členný vektor \underline{x} a analogicky i soustavu funkcí $f_1(\underline{x}), f_2(\underline{x}), \dots, f_n(\underline{x})$ za n -členný vektor \underline{f} (vektorovou funkci), pak soustavu (4.36) lze zkráceně zapsat ve tvaru

$$(4.38) \quad \underline{f}(\underline{x}) = \underline{0},$$

kde $\underline{0}$ je nulový vektor.

K řešení soustavy nelineárních rovnic (4.36) užíváme nejčastěji metodu prosté iterace a (zobecněnou) Newtonovu metodu, které dále vysvětlíme.

4.4.1 Metoda prosté iterace

Soustavu (4.36) přepíšeme vhodným způsobem na tvar

$$(4.39) \quad \begin{aligned} x_1 &= \varphi_1(x_1, x_2, \dots, x_n), \\ x_2 &= \varphi_2(x_1, x_2, \dots, x_n), \\ &\dots \\ x_n &= \varphi_n(x_1, x_2, \dots, x_n) \end{aligned}$$

tak, aby (4.39) a (4.36) měly stejné řešení. Přitom předpokládejme, že $\varphi_1, \varphi_2, \dots, \varphi_n$ jsou reálné funkce definované a spojité v jistém okolí izolovaného řešení soustavy (4.36).

Pro snadnost zápisů budeme opět užívat vektorového označení, tedy

$$(4.40) \quad \underline{\varphi}(\underline{x}) = (\varphi_1(\underline{x}), \varphi_2(\underline{x}), \dots, \varphi_n(\underline{x}))^T,$$

pak soustava (4.39) je zapsána tvarem

$$(4.41) \quad \underline{x} = \underline{\varphi}(\underline{x}).$$

Zvolme $\underline{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})^T$ jako počáteční approximaci řešení $\underline{\xi} = (\xi_1, \xi_2, \dots, \xi_n)^T$ soustavy (4.39) a podle iteračního vzorce

$$(4.42) \quad \underline{x}^{(k+1)} = \underline{\varphi}(\underline{x}^{(k)}),$$

počítejme postupně $\underline{x}^{(1)}, \underline{x}^{(2)}, \dots$. Postačující podmínky konvergence $\{\underline{x}^{(k)}\} \rightarrow \underline{\xi}$ určuje následující věta.

Věta 4.8 Mějme dánu soustavu (4.39), resp. (4.41). Dále předpokládejme, že v jistém okolí G řešení $\underline{\xi}$ soustavy (4.39) platí:

$$(4.43) \quad \begin{aligned} a) \quad &\underline{\varphi}(\underline{x}) \in G \quad \text{pro } \underline{x} \in G, \\ b) \quad &\text{existuje } q \in (0, 1) \text{ tak, že pro } \underline{x}, \underline{y} \in G \text{ je} \\ &\|\underline{\varphi}(\underline{x}) - \underline{\varphi}(\underline{y})\| \leq q \|\underline{x} - \underline{y}\|. \end{aligned}$$

Potom posloupnost $\{\underline{x}^{(k)}\}_{k=0}^\infty$, definovaná rekurentním vztahem (4.42), konverguje k řešení $\underline{\xi}$ pro libovolné $\underline{x}^{(0)} \in G$ a přitom platí:

$$(4.44) \quad \|\underline{x}^{(k)} - \underline{\xi}\| \leq \frac{q}{1-q} \|\underline{x}^{(k)} - \underline{x}^{(k-1)}\|.$$

Důkaz viz např. [5].

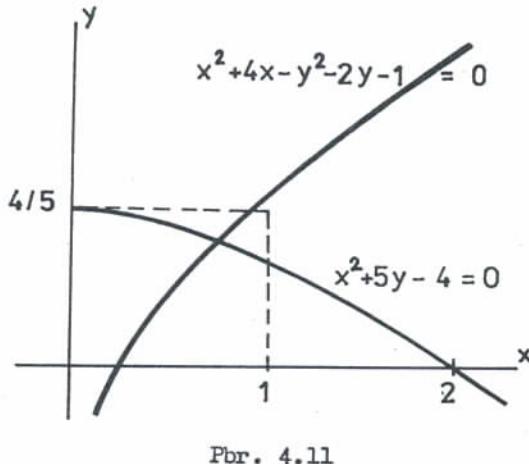
Jestliže je funkce $\underline{\varphi}(\underline{x})$ diferencovatelná, pak podmínce (4.43) b) můžeme nahradit podmínkou

$$(4.45) \quad \|\underline{\varphi}'(\underline{x})\| \leq q < 1,$$

kde $\underline{\varphi}'(\underline{x})$ označuje Jacobiovu matici $J\varphi(x_1, x_2, \dots, x_n)$, složenou z parciálních derivací $\partial \varphi_i / \partial x_j$, viz (4.37).

I když výpočet approximací podle vzorce (4.42) je jednoduchý, není vždycky snadné najít k soustavě (4.36) vhodnou soustavu (4.39) a ověřit podmínky konvergence iteračního procesu.

Příklad 4.11 Najděme prostou iterační metodou řešení soustavy



$$x^2 + 4x - y^2 - 2y - 1 = 0,$$

$$x^2 + 5y - 4 = 0,$$

které leží v 1. kvadrantu.

Řešení: Z grafického znázornění, viz obr. 4.11, vidíme, že řešení leží v oblasti

$$\begin{aligned} G = \{ (x, y) : & x \in \langle 0, 1 \rangle, \\ & y \in \langle 0, \frac{4}{5} \rangle \} . \end{aligned}$$

Soustavu upravíme na tvar

$$\begin{aligned} x &= ((y+1)^2 - x^2)/4 \equiv \varphi_1(x, y), \\ (4.46) \quad y &= (4 - x^2)/5 \equiv \varphi_2(x, y). \end{aligned}$$

Protože $\underline{\varphi}$ je diferencovatelná, ověřme splnění podmínky (4.45) pro

$$\underline{\varphi}'(\underline{x}) = J\varphi(x, y) = \begin{pmatrix} \partial \varphi_1 / \partial x, & \partial \varphi_1 / \partial y \\ \partial \varphi_2 / \partial x, & \partial \varphi_2 / \partial y \end{pmatrix} = \begin{pmatrix} -x/2 & (y+1)/2 \\ -2x/5 & 0 \end{pmatrix}.$$

Jestliže budeme počítat s maximalistickou sloupcovou normou, potom

$$\begin{aligned} \|\underline{\varphi}'(\underline{x})\| &= \|J\varphi(x, y)\| = \max_{(x,y) \in G} (|x|9/10, |y+1|/2) = \max(9/10, 9/10) = \\ &= 9/10 < 1. \end{aligned}$$

k	x ^(k)	y ^(k)
0	0	0
2	0.794 38	0.787 50
4	0.597 67	0.717 82
6	0.641 86	0.715 91
8	0.637 34	0.719 84
10	0.636 80	0.718 61
12	0.637 29	0.718 84
14	0.637 08	0.718 83
15	0.637 13	0.718 82
16	0.637 10	0.718 81

Tab. 4.7

Podmínu (č.43) a) ověřovat nebudeme, ale sledujme, zdali postupné approximace budou ležet v oblasti G. Zvolme $x^{(0)} = y^{(0)} = 0$ a počítejme podle (4.46) další approximace, které jsou sestaveny v tabulce 4.7.

Pro odhad chyby platí (jestliže užíváme u vektorů opět sloupcovou normu) podle (4.44)

$$\|(x^{(16)}, y^{(16)})^T - (\xi_1, \xi_2)^T\| \leq$$

$$\leq \frac{0.9}{0,1} \|(x^{(16)}, y^{(16)})^T - (x^{(15)}, y^{(15)})^T\| =$$

$$= 9 \cdot 4 \cdot 10^{-5} = 3,6 \cdot 10^{-4}.$$

4.4.2 Newtonova (zobecněná) metoda

Mějme dánu soustavu (4.36) a jistou oblast G , ve které má soustava (4.36) izolované řešení $\underline{f} = (\xi_1, \xi_2, \dots, \xi_n)$. Nechť jsou v této oblasti funkce f_1, f_2, \dots, f_n spojité spolu se svými l. parciálními derivacemi a nechť Jacobiova matice (4.37) je v G regulární. Dále předpokládejme, že již známe k-tou approximaci řešení

$$\underline{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})^T,$$

tedy přesné řešení lze vyjádřit ve tvaru

$$(4.47) \quad \underline{\xi} = \underline{x}^{(k)} + \underline{h}^{(k)} = \left(x_1^{(k)} + h_1^{(k)}, x_2^{(k)} + h_2^{(k)}, \dots, x_n^{(k)} + h_n^{(k)} \right)^T,$$

kde $\underline{h}^{(k)} = (h_1^{(k)}, h_2^{(k)}, \dots, h_n^{(k)})^T$ je chyba k-té approximace, kterou přibližně vyjádříme pomocí Taylorova polynomu l. stupně z rovnic (4.36) takto:

$$(4.48) \quad 0 = f_i(\underline{x}^{(k)} + \underline{h}^{(k)}) \approx f_i(\underline{x}^{(k)}) + \sum_{j=1}^n \frac{\partial f_i}{\partial x_j} \Big|_{\underline{x}=\underline{x}^{(k)}} \cdot h_j^{(k)}, \quad i = 1, 2, \dots, n.$$

Rovnice (4.48) představují soustavu n lineárních rovnic o n neznámých

$$(4.49) \quad \begin{pmatrix} \frac{\partial f_1}{\partial x_1}, \frac{\partial f_1}{\partial x_2}, \dots, \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1}, \frac{\partial f_2}{\partial x_2}, \dots, \frac{\partial f_2}{\partial x_n} \\ \vdots \\ \frac{\partial f_n}{\partial x_1}, \frac{\partial f_n}{\partial x_2}, \dots, \frac{\partial f_n}{\partial x_n} \end{pmatrix} \cdot \begin{pmatrix} h_1^{(k)} \\ h_2^{(k)} \\ \vdots \\ h_n^{(k)} \end{pmatrix} = \begin{pmatrix} -f_1(\underline{x}^{(k)}) \\ -f_2(\underline{x}^{(k)}) \\ \vdots \\ -f_n(\underline{x}^{(k)}) \end{pmatrix},$$

jejíž matice je Jacobiova matice s parciálními derivacemi počítanými pro $\underline{x} = \underline{x}^{(k)}$, vektorem neznámých $\underline{h}^{(k)}$ a vektorem pravých stran $-f(\underline{x}^{(k)})$. Označíme-li Jacobiovu matici podle (4.37) $Jf(\underline{x}^{(k)})$, pak (4.49) zapíšeme symbolicky jako

$$Jf(\underline{x}^{(k)}) \cdot \underline{h}^{(k)} = -f(\underline{x}^{(k)}),$$

tedy řešení je

$$\underline{h}^{(k)} = -[Jf(\underline{x}^{(k)})]^{-1} \cdot f(\underline{x}^{(k)}).$$

V důsledku přibližné rovnosti (4.48) platí

$$\underline{\xi} \approx \underline{x}^{(k)} + \underline{h}^{(k)} = \underline{x}^{(k+1)},$$

tedy jsme dostali další approximaci řešení podle vzorce

$$(4.50) \quad \underline{x}^{(k+1)} = \underline{x}^{(k)} - [Jf(\underline{x}^{(k)})]^{-1} \cdot f(\underline{x}^{(k)}).$$

Vzorec (4.50) se nazývá Newtonova metoda pro řešení soustavy n nelineárních rovnic o n neznámých.

Podmínkami konvergence posloupnosti $\{\underline{x}^{(k)}\}$, jejíž prvky jsou počítané podle vzorce (4.50), se bude zabývat věta 4.9, i když ve většině případů se velmi obtížně ověřují. Při výpočtech sledujeme, jestli v jistém okolí \underline{x} kořene \underline{x} je Jacobiova matice regulární a zdali jednotlivé aproximace jsou prvky tohoto okolí.

Věta 4.9 Nechť $\underline{f}(\underline{x}) = \underline{0}$ je soustava (4.36) a G je oblast obsahující izolované řešení \underline{x} . Dále nechť jsou splněny tyto podmínky:

- 1) V G mají funkce f_1, f_2, \dots, f_n druhé parciální derivace a jejich absolutní hodnoty jsou menší než D_2 .
- 2) V bodě $\underline{x}^{(0)} \in G$ je Jacobiova matice $J\underline{f}(\underline{x}^{(0)})$ regulární.
- 3) Platí nerovnost

$$h = M^2 D_2 \delta n^2 \leq \frac{1}{2},$$

kde

$$(4.51) \quad \begin{aligned} |f_i(\underline{x}^{(0)})| &\leq \delta \quad \text{pro } i = 1, 2, \dots, n, \\ \|(\underline{J}\underline{f}(\underline{x}^{(0)}))^{-1}\| &\leq M. \end{aligned}$$

Potom posloupnost $\{\underline{x}^{(k)}\}$, počítaná podle vzorce (4.50), konverguje k řešení \underline{x} a rychlosť této konvergencie je dáná vzorcem

$$(4.52) \quad \|\underline{x}^{(k)} - \underline{x}\| \leq \frac{1}{2^{k-1}} (2h)^{2^{k-1}} \cdot \delta.$$

Důkaz viz [5].

Příklad 4.12 Newtonovou metodou řešme úlohu z příkladu 4.11.

Řešení: Podle (4.37) je Jacobiova matice

$$\underline{J}\underline{f}(x, y) = \begin{pmatrix} 2x + 4, & -2y - 2 \\ 2x, & 5 \end{pmatrix},$$

a její determinant

$$\det \underline{J}\underline{f}(x, y) = 5(2x + 4) + 2x(-2y - 2) \neq 0$$

pro všechna $x > 0, y > 0$. Potom

$$[\underline{J}\underline{f}(x, y)]^{-1} = \frac{1}{14x + 4xy + 20} \begin{pmatrix} 5, & 2y + 2 \\ -2x, & 2x + 4 \end{pmatrix},$$

tedy iterační rovnice (4.50) má tvar

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} - \frac{1}{14x + 4xy + 20} \begin{pmatrix} 5, & 2y + 2 \\ -2x, & 2x + 4 \end{pmatrix} \begin{pmatrix} x^2 + 4x - y^2 - 2y - 1 \\ x^2 + 5y - 4 \end{pmatrix},$$

jestliže na levé straně položíme $x = x^{(k+1)}$, $y = y^{(k+1)}$ a na straně pravé položíme $x = x^{(k)}$, $y = y^{(k)}$. Podle obrázku 4.8 zvolíme $x^{(0)} = 0,7$, $y^{(0)} = 0,7$.

k	$x^{(k)}$	$y^{(k)}$
0	0.700 000	0.700 000
1	0.638 098	0.719 332
2	0.637 614	0.718 690
3	0.637 108	0.718 819

Numerické výsledky jsou sestaveny v tabulce 4.8. Porovnáním výsledků s výsledky v tabulce 4.7 vidíme, že konvergance je mnohem rychlejší.

Tab. 4.8

4.5 Cvičení

1. Všemožnými iteračními metodami z odstavce 4.2 vypočítejte reálný kořen rovnic

- | | |
|--------------------------|------------------------------|
| a) $4x - \cos x = 0,$ | g) $\ln(5x/2) = \sin \pi x,$ |
| b) $x^3 + x - 1000 = 0,$ | h) $x^2 = \cos x,$ |
| c) $x - \sin x = 1/4,$ | i) $-3x^2 - 2x + 4 = 0,$ |
| d) $2 - x - \ln x = 0,$ | j) $x + e^x = 0,$ |
| e) $x^5 - x - 1 = 0,$ | k) $x - \cos x = 0,$ |
| f) $x \ln x = 100,$ | l) $x \ln x - 14 = 0,$ |
- s přesností 10^{-3} . Přibližné polohy kořenů určete graficky.

2. Metodami z odstavce 4.3 najděte všechny reálné kořeny rovnic

- | | |
|---------------------------------|--------------------------------|
| a) $x^3 - 5x + 3 = 0,$ | d) $x^4 - 10x^3 + 1 = 0,$ |
| b) $x^4 + 10x - 100 = 0,$ | e) $x^3 + 5x^2 - 15x - 7 = 0,$ |
| c) $x^3 + 4x^2 - 14x - 12 = 0,$ | f) $2x^3 - x^2 - 7x + 5 = 0,$ |
- s přesností 10^{-3} .

3. Metodou prosté iterace a Newtonovou metodou najděte jedno řešení soustav

- | | |
|-------------------------|------------------------------|
| a) $x^2 + y^2 - 1 = 0,$ | b) $2x^2 - xy - 5x + 1 = 0,$ |
| $x^3 - y = 0,$ | $x + 3 \ln x - y^2 = 0,$ |

s přesností 10^{-3} . Přibližné hodnoty řešení najděte graficky.

Výsledky

1. a) $x \doteq 0,243,$ e) $x \doteq 1,167,$ i) $x \doteq 0,901$
 b) $x \doteq 9,967,$ f) $x \doteq 29,537,$ j) $x \doteq -0,568$
 c) $x \doteq 1,171,$ g) $x \doteq 0,772$ k) $x \doteq 0,739$
 d) $x \doteq 1,557$ h) $x \doteq 0,824$ l) $x \doteq 7,128$
2. a) $x_1 \doteq 0,657,$ $x_2 \doteq 1,834,$ $x_3 \doteq -2,491$
 b) $x_1 \doteq -3,402,$ $x_2 \doteq 2,903,$
 c) $x_1 \doteq 6,$ $x_2 \doteq -0,732,$ $x_3 \doteq 2,732$

$$d) \quad x_1 \doteq 0,472, \quad x_2 \doteq 9,999$$

$$e) \quad x_1 \doteq -7, \quad x_2 \doteq -0,414, \quad x_3 = 2,414$$

$$f) \quad x_1 \doteq -1,951, \quad x_2 \doteq 0,756$$

3. a) $x_1 \doteq 0,826, \quad y_1 \doteq 0,564$

b) $x_1 \doteq 1,373, \quad y_1 \doteq -1,525.$

KAPITOLA 5.

NUMERICKÉ ŘEŠENÍ OBYČEJNÝCH DIFERENCIÁLNÍCH ROVNIC

Numerické řešení obyčejných diferenciálních rovnic patří k základním partiím numerické matematiky, protože pomocí nich je formulováno mnoho fyzikálních úloh z rozličných technických i jiných oblastí.

Podstatnou část této kapitoly budeme věnovat počáteční úloze pro obyčejnou diferenciální rovnici 1. řádu, rozřešenou vzhledem k derivaci a počáteční úloze pro soustavy takových rovnic. Protože taková soustava může vzniknout také transformací diferenciální rovnice m -tého řádu, pak řešení soustav diferenciálních rovnic 1. řádu zahrnuje vlastně i řešení obyčejné diferenciální rovnice m -tého řádu.

Nakonec se budeme ještě krátce zabývat řešením okrajových úloh pro vybrané typy obyčejných diferenciálních rovnic.

5.1 Počáteční úloha pro rovnice 1. řádu

V tomto odstavci se budeme zabývat řešením diferenciální rovnice 1. řádu

$$(5.1) \quad y'(x) = f(x, y(x)), \quad x \in \langle a, b \rangle,$$

s počáteční podmínkou

$$(5.2) \quad y(a) = y_0.$$

Přitom vždy předpokládejme, že:

a) $f(x, y)$ je definovaná a spojitá jako funkce dvou proměnných v oblasti

$$G = \{(x, y) : a \leq x \leq b, -\infty < y < +\infty\}.$$

b) $f(x, y)$ splňuje v G tzv. Lipschitzovu podmíinku, tj. existuje konstanta L taková, že pro každé $x \in \langle a, b \rangle$ a libovolné y, \tilde{y} platí

$$(5.3) \quad |f(x, y) - f(x, \tilde{y})| \leq L|y - \tilde{y}|.$$

Tyto podmínky postačují k tomu, aby daná úloha byla jednoznačně řešitelná, tzn. že existuje právě jedna funkce $y(x)$, která je v intervalu $\langle a, b \rangle$ spojitá se svou první derivací a splňuje rovnici (5.1) i okrajovou podmíinku (5.2).

Úlohu, kterou jsme právě definovali, lze snadno rozšířit na počáteční úlohu pro soustavu diferenciálních rovnic 1. řádu, neboť $y(x)$ v (5.1) můžeme považovat za vektor $(y_1(x), y_2(x), \dots, y_m(x))$ a tím $f(x, y)$ za vektorovou funkci $m+1$ proměnných. Problém (5.1), (5.2) by byl potom velice obecný, protože každá rovnice nebo soustava rovnic vyššího řádu se dá za jistých podmínek převést na soustavu rovnic prvního řádu rozřešených vzhledem k derivaci. Z tohoto důvodu se budeme zabývat úlohou (5.1) a (5.2) jako základní úlohou při řešení obyčejných diferenciálních rovnic.

Základem numerických metod řešení počátečních úloh je tzv. diskretizace proměnných, tzn. že řešení hledáme ve tvaru tabulky hodnot $y_0, y_1, y_2, \dots, y_n$, které approximují hodnoty přesného řešení $y(x_0), y(x_1), \dots, y(x_n)$. Body x_0, x_1, \dots, x_n budeme volit vždy jako ekvidistantní dělení intervalu $\langle a, b \rangle$, tedy $x_i = a + ih$, $i = 0, 1, \dots, n$ a krok dělení $h = (b - a)/n$.

Dále probírané numerické metody dělíme na jednokrokové, které přibližnou hodnotu řešení y_{i+1} počítají pouze pomocí hodnot $y_i, f(x_i, y_i)$, a vícekrokové, které hodnotu y_{i+1} počítají z hodnot $y_i, y_{i-1}, \dots, y_{i-k+1}, f(x_i, y_i), f(x_{i-1}, y_{i-1}), \dots, f(x_{i-k+1}, y_{i-k+1})$ pro $k > 1$.

5.2 Jednokrokové metody řešení počáteční úlohy

Jednokrokovou metodou pro řešení úlohy (5.1), (5.2) nazveme každý algoritmus, který počítá přibližnou hodnotu řešení y_{i+1} pouze z hodnot x_i, y_i a h . Obecně budeme psát

$$(5.4) \quad y_{i+1} = y_i + h \phi(x_i, y_i, h), \quad i = 0, 1, \dots, n-1,$$

kde ϕ , tzv. přírůstková funkce, je spojitá v oblasti $\Omega = \{(x, y, h) : a \leq x \leq b, -\infty < y < \infty, 0 \leq h \leq H\}$ (H je pevné kladné číslo) a závisí na $f(x, y)$. V dalším si ukážeme některé způsoby konstrukce přírůstkové funkce.

5.2.1 Eulerova metoda

Tato metoda je jednou z nejstarších a spočívá na jednoduchém požadavku nahradit řešení úlohy (5.1), (5.2) na intervalu $\langle x_i, x_{i+1} \rangle$ přímou se směrnicí $y'(x_i) = f(x_i, y_i)$, která je vedena bodem $[x_i, y_i]$.

Z toho plynne, že hodnoty y_1, y_2, \dots, y_n se Eulerovou metodou počítají podle vzorce

$$(5.5) \quad y_{i+1} = y_i + h \cdot f(x_i, y_i), \quad i = 0, 1, 2, \dots, n-1.$$

Při srovnání s (5.4) můžeme zapsat Eulerovu metodu vzorcem

$$(5.6) \quad y_{i+1} = y_i + h \cdot \phi_E(x_i, y_i, h), \quad i = 0, 1, 2, \dots, n-1,$$

kde $\phi_E = f(x_i, y_i)$.

Příklad 5.1 Eulerovou metodou řešme rovnici $y' = y + x^2$ s počáteční podmínkou $y(0) = 1$, pro $x \in \langle 0; 0,5 \rangle$ a $h = 0,1$.

Řešení: Řešení podle vzorce (5.5) je přehledně uspořádáno v tabulce 5.1.

i	x_i	y_i	$f(x_i, y_i)$	$hf(x_i, y_i)$	Přesné řešení $y(x_i)$	$ y_i - y(x_i) $
0	0,0	1,000	1,000	0,100	1,000	0,000
1	0,1	1,100	1,110	0,111	1,106	0,006
2	0,2	1,211	1,251	0,125	1,224	0,013
3	0,3	1,336	1,426	0,143	1,360	0,024
4	0,4	1,479	1,639	0,164	1,515	0,036
5	0,5	1,643			1,696	0,053

Tab. 5.1

5.2.2 Metody Taylorova typu

Předpokládejme, že řešení úlohy (5.1), (5.2) je dostatečně hladké, tzn. že má dostatečný počet spojitých derivací v $\langle a, b \rangle$. Potom pro libovolné dva body dělení $x_i, x_{i+1} \in \langle a, b \rangle$ platí pro řešení $y(x)$ podle Taylorovy věty

$$(5.7) \quad \begin{aligned} y(x_{i+1}) &= y(x_i) + h \cdot y'(x_i) + \frac{h^2}{2!} y''(x_i) + \dots + \frac{h^p}{p!} y^{(p)}(x_i) + \frac{h^{p+1}}{(p+1)!} y^{(p+1)}(\xi) = \\ &= y(x_i) + h \cdot \sum_{s=1}^p \frac{h^{s-1}}{s!} f^{(s-1)}(x_i, y(x_i)) + \frac{h^{p+1}}{(p+1)!} f^{(p)}(\xi, y(\xi)), \end{aligned}$$

kde $\xi \in (x_i, x_{i+1})$ a $f, f', f'', \dots, f^{(p)}$ jsou derivace pravé strany rovnice (5.1) podle x a nahrazují příslušné derivace řešení $y', y'', \dots, y^{(p)}, y^{(p+1)}$. Derivace funkce $f(x, y)$ podle x počítáme podle věty o derivování implicitní funkce, tedy

$$(5.8) \quad \begin{aligned} y'(x) &= f(x, y), \\ y''(x) &= f'(x, y) = f_x + f_y \cdot f, \\ y'''(x) &= f''(x, y) = f_{xx} + 2f_{xy} \cdot f + f_{yy} \cdot f^2 + (f_x + f_y \cdot f) \cdot f_y, \\ &\text{atd.} \end{aligned}$$

kde $f_x, f_y, f_{xx}, f_{xy}, f_{yy}, \dots$ jsou parciální derivace pravé strany $f(x, y)$.

V jednokrokových metodách Taylorova typu volíme vzhledem k (5.7) přírůstkovou funkci ϕ , kterou nyní označíme ϕ_T , ve tvaru

$$(5.9) \quad \begin{aligned} \phi_T &= \sum_{s=1}^p \frac{h^{s-1}}{s!} f^{(s-1)}(x_i, y_i) = f(x_i, y_i) + \\ &+ \frac{h}{2!} f'(x_i, y_i) + \dots + \frac{h^{p-1}}{p!} f^{(p-1)}(x_i, y_i), \end{aligned}$$

kde derivace (5.8) jsou počítány pro $x = x_i$, $y = y_i$ (y_i je přibližná hodnota přesného řešení $y(x_i)$).

Metody Taylorova typu jsou tedy dány vzorcem

$$(5.10) \quad y_{i+1} = y_i + h \cdot \phi_T(x_i, y_i, h), \quad i = 0, 1, \dots, n-1,$$

kde přírůstková funkce ϕ_T je definována vztahem (5.9). Je zřejmé, že pro $p = 1$ je $\phi_T = \phi_E$.

Velkou nevýhodou těchto metod je pracný výpočet derivací (5.8), i když pro malá p jsou vzorce (5.10) užitečné, avšak při řešení na počítačích se málo užívají.

Příklad 5.2 Taylorovým vzorcem (5.10), (5.9) pro $p = 2$ řešme úlohu zadanou v příkladu 5.1.

i	x _i	y _i
0	0	1,000
1	0,1	1,105
2	0,2	1,223
3	0,3	1,358
4	0,4	1,513
5	0,5	1,693

Tab. 5.2

Řešení: Pro $p = 1$ dostaneme z (5.10), (5.9) a (5.8) vzorec

$$\begin{aligned} y_{i+1} &= y_i + h \left[f(x_i, y_i) + \frac{h}{2} f'(x_i, y_i) \right] = \\ &= y_i + h \left[x_i^2 + y_i + \frac{h}{2} (2x_i + x_i^2 + y_i) \right]. \end{aligned}$$

Numerické hodnoty řešení jsou v následující tabulce 5.2. Výsledky porovnejte s řešením příkladu 5.1.

5.2.3 Metody Rungovy-Kuttovy

Vzhledem k nevýhodám metod (5.10) byly hledány takové přírůstkové funkce, které by neobsahovaly derivace funkce $f(x, y)$ a jejichž rozvoj podle proměnné h by se shodoval s (5.7) do co nejvyšší mocniny h . Z těchto požadavků byly Rungovy-Kuttovy metody odvozeny. Jejich základem je vyjádření přírůstkové funkce (označme ji nyní ϕ_{RK}) lineární kombinací

$$(5.11) \quad \phi_{RK}(x_i, y_i, h) = p_{rl} k_1(h) + p_{r2} k_2(h) + \dots + p_{rr} k_r(h),$$

kde $r \geq 1$, p_{rl} až p_{rr} jsou reálné koeficienty a

$$\begin{aligned} (5.12) \quad k_1(h) &= f(x_i, y_i), \\ k_2(h) &= f(x_i + \alpha_2 h, y_i + h \beta_{21} k_1), \\ &\dots \\ k_r(h) &= f(x_i + \alpha_r h, y_i + h (\beta_{rl} k_1 + \beta_{r2} k_2 + \dots + \beta_{r,r-1} k_{r-1})), \end{aligned}$$

jsou funkce proměnné h , v nichž α , β jsou reálná čísla. Tedy vzorce jednobodové Rungovy-Kuttovy metody mají tvar

$$(5.13) \quad y_{i+1} = y_i + h \sum_{s=1}^r p_{rs} k_s(h),$$

ve kterém konstanty p, α, β určíme při pevném r tak, aby se rozvoj přírůstkové funkce ϕ_{RK} podle h shodoval s přírůstkovou funkcí (5.9) metod Taylorova typu do co nejvyšší možné mocniny proměnné h . Zvolíme-li $r = 1$, pak funkce

$$\phi_{RK} = p_{11} \cdot k_1(h) = p_{11} \cdot f(x_i, y_i)$$

není závislá na h , tedy její rozvoj podle h je pouze $p_{11} \cdot f(x_i, y_i)$ a ten se může shodovat jenom s prvním členem v (5.9), tedy $p_{11} = 1$. Pak (5.13) pro $r = 1$ definuje Eulerovu metodu stejně jako (5.10) pro $p = 1$.

Ještě odvodíme vzorec (5.13) pro $r = 2$, kdy přírůstková funkce (5.11) a její rozvoj podle h mají tvar

$$\begin{aligned}
 \phi_{RK}(x_i, y_i, h) &= p_{21}f(x_i, y_i) + p_{22}f(x_i + \alpha_2 h, y_i + \beta_{21} h f(x_i, y_i)) = \\
 &= \phi_{RK}(x_i, y_i, 0) + h \phi'_{RK}(x_i, y_i, 0) + O(h^2) = \\
 (5.14) \quad &= p_{21}f(x_i, y_i) + p_{22}f(x_i, y_i) + \\
 &+ h [p_{22}(f_x(x_i, y_i)\alpha_2 + f_y(x_i, y_i)\beta_{21}f(x_i, y_i))] + O(h^2),
 \end{aligned}$$

kde $O(h^2)$ je zbytek rozvoje, který obsahuje h ve druhé mocnině.

Porovnáme-li výrazy umulté a první mocniny h v (5.14) a (5.9), dostaneme rovnice

$$\begin{aligned}
 (p_{21} + p_{22})f(x_i, y_i) &= f(x_i, y_i), \\
 p_{22}\alpha_2 \cdot f_x(x_i, y_i) + p_{22}\beta_{21} \cdot f_y(x_i, y_i) \cdot f(x_i, y_i) &= \\
 &= \frac{1}{2} [f_x(x_i, y_i) + f_y(x_i, y_i) \cdot f(x_i, y_i)],
 \end{aligned}$$

které budou splněny, jestliže

$$p_{21} + p_{22} = 1, \quad p_{22}\alpha_2 = 1/2, \quad p_{22}\beta_{21} = 1/2.$$

Tato soustava má nekonečně mnoho řešení, která budou záviset na volbě některé neznámé. Zvolíme-li např. $\alpha_2 = 1$, potom $p_{22} = 1/2$, $\beta_{21} = 1$ a $p_{21} = 1/2$, tedy podle (5.13) má vzorec Rungovy-Kuttovy metody tvar

$$(5.15) \quad y_{i+1} = y_i + h/2 [k_1 + k_2],$$

$$\text{kde } k_1 = f(x_i, y_i),$$

$$k_2 = f(x_i + h, y_i + h \cdot k_1).$$

Ověřte si, že koeficienty u h^2 v (5.14) a (5.9) se nemohou pro obecnou funkci $f(x, y)$ již shodovat.

Obdobným postupem můžeme pro $r = 3$ dostat Rungův-Kuttův vzorec (viz [16])

$$(5.16) \quad y_{i+1} = y_i + \frac{h}{6} [k_1 + 4k_2 + k_3],$$

$$\text{kde } k_1 = f(x_i, y_i),$$

$$k_2 = f(x_i + h/2, y_i + h \cdot k_1/2),$$

$$k_3 = f(x_i + h/2, y_i + h \cdot k_2).$$

Nejužívanější metodou Rungova-Kuttova typu je tzv. klasická Rungova-Kuttova metoda, kdy pro $r = 4$ docílíme shody rozvoje funkce ϕ_{RK} s (5.9) až do třetí mocniny

ny h , tedy shody prvních čtyř členů rozvoje funkce ϕ_{RK} s prvními čtyřmi členy v (5.9). Tím získáme nelineární soustavu (viz [23], str. 223), která má nekonečně mnoho řešení. Jedno z nich je

$$\alpha_2 = \alpha_3 = 1/2, \alpha_4 = 1, \beta_{21} = \beta_{32} = 1/2, \beta_{43} = 1, \beta_{31} = \beta_{41} = \beta_{42} = 0,$$

$$p_{41} = p_{44} = 1/6 \text{ a } p_{42} = p_{43} = 1/3.$$

Dosazením do (5.13) a (5.12) dostaneme vzorce klasické Rungovy-Kuttovy metody

$$(5.17) \quad y_{i+1} = y_i + \frac{h}{6} [k_1 + 2k_2 + 2k_3 + k_4],$$

$$\begin{aligned} k_1 &= f(x_i, y_i), \\ k_2 &= f(x_i + h/2, y_i + h \cdot k_1/2), \\ k_3 &= f(x_i + h/2, y_i + h \cdot k_2/2), \\ k_4 &= f(x_i + h, y_i + h \cdot k_3). \end{aligned}$$

Obecně neplatí, že pro libovolné r se rozvoj funkce ϕ_{RK} z (5.11) shoduje s r členy vzorce (5.9), i když to pro $r = 1, 2, 3$ a 4 platí. Označíme-li počet shodných členů p , pak platí $p(r)$ a pro $r \rightarrow \infty$ také $p(r) \rightarrow \infty$.

Podrobněji se těmito otázkami nebudeme zabývat.

Příklad 5.3 Klasickou Rungovou-Kuttovou metodou 4. řádu řešme rovnici $y' = y + x^2$ s počáteční podmínkou $y(0) = 1$ pro $x \in \langle 0; 0,5 \rangle$ a $h = 0,1$.

Řešení: Řešení podle vzorců (5.17) je uspořádáno v tabulce 5.3.

i	x_i	y_i	k_1	k_2	k_3	k_4
0	0,0	1,000	1,0000	1,0525	1,0551	1,1155
1	0,1	1,106	1,1160	1,1843	1,1877	1,2054
2	0,2	1,224	1,2640	1,3497	1,3540	1,4494
3	0,3	1,359	1,4490	1,5539	1,5592	1,6049
4	0,4	1,514	1,6740	1,8002	1,8065	1,9447
5	0,5	1,693				

Tab. 5.3

5.2.4 Odhad chyby a konvergence

Přírůstkové funkce jednokrokových metod pro řešení úlohy (5.1), (5.2), které jsme až doposud probrali, splňují kromě požadavku spojitosti ještě Lipschitzovu podmínu vzhledem k y . Existuje tedy konstanta L taková, že pro každé dva body $(x, y), (x, \tilde{y}) \in G$ a libovolné $h \in \langle 0, H \rangle$ platí

$$(5.18) \quad |\phi(x, y, h) - \phi(x, \tilde{y}, h)| \leq L|y - \tilde{y}|.$$

Splnění (5.18) plyne z tvaru funkcí $\phi_E, \phi_T, \phi_{RK}$ a ze vztahu (5.3). Jednokrokové metody, které splňují (5.18), se nazývají regulární. Probrané metody jsou navíc konzistentní, což znamená, že pro libovolné $(x, y) \in G$ je

$$(5.19) \quad \phi(x, y, 0) = f(x, y).$$

Ověření (5.19) pro ϕ_E , ϕ_T a ϕ_{RK} je triviální.

Chybou jednokrokové metody jsou dvojího druhu, jednak jsou to lokální a jednak globální diskretizační chyby. Lokální diskretizační chyba d_i je nepřesnost, s níž vyjadřuje příslušný vzorec rozdíl $y(x_{i+1}) - y(x_i)$, tedy je to nepřesnost vzorce metody. Můžeme ji vyjádřit pomocí (5.4) vztahem

$$d_i = y(x_{i+1}) - y(x_i) - h\phi(x_i, y(x_i), h).$$

Pro metody Taylorova typu vidíme z (5.8), že lokální chyba

$$(5.20) \quad d_i = \frac{h^{p+1}}{(p+1)!} y^{(p+1)}(\xi) = \frac{h^{p+1}}{(p+1)!} f^{(p)}(\xi, y(\xi)).$$

S pojmem lokální diskretizační chyby je úzce spojen pojem řád metody, který definuje rychlosť, s níž při pevném n a $h \rightarrow 0$ konverguje chyba d_n k nule. Řád metody je největší přirozené číslo p takové, že pro danou metodu aplikovanou na úlohu s dostatečně hladkým řešením platí odhad

$$d_n = O(h^{p+1}).$$

Tedy metody Taylorova typu jsou p -tého řádu. Rungovy-Kuttovy metody jsou p -tého řádu, když se v rozvoji příruškové funkce (5.11) a v (5.7) shoduje prvních p členů, které obsahují h v multě až $(p-1)$ -ní mocnině.

Podstatnější než lokální je globální (také akumulovaná) diskretizační chyba e_i , která vznikne nahromaděním lokálních chyb v průběhu užívání vzorce metody, tj. při postupu od y_1 až k y_{i+1} . Globální chyba je měřítkem toho, jak přesně approximují hodnoty y_1, y_2, \dots, y_n hodnoty přesného řešení $y(x_1), y(x_2), \dots, y(x_n)$ počáteční úlohy.

Tedy

$$e_i = y(x_i) - y_i, \quad i = 1, 2, \dots, n.$$

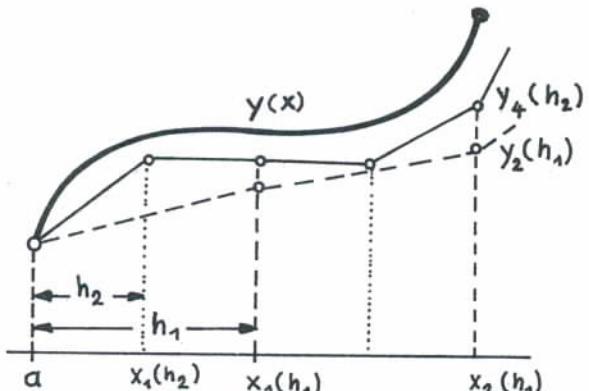
Globální chyba závisí na zvolené metodě, velikosti kroku h , na indexu i , ale také na zaokrouhlovacích chybách, které vznikají při výpočtu s neúplnými čísly.

V dalším se vyhneme složitým úvahám o chybách, ale ukážeme si praktický odhad globální diskretizační chyby metodou polovičního kroku. Sestrojme dvě posloupnosti přibližných hodnot $\{y_i(h)\}, \{y_i(h/2)\}$, jednokrokovou metodu p -tého řádu, kde $y_i(h)$ a $y_i(h/2)$ jsou přibližné metody přesného řešení $y(x_i)$, získané při kroku h a $h/2$. Potom chyba ℓ_i numerického řešení $y_i(h/2)$ je dána odhadem

$$(5.21) \quad e_i \approx \frac{y_i(h/2) - y_i(h)}{2^p - 1}.$$

Důkaz vzorce (5.21) je uveden např. ve [13].

Základní vlastnosti, kterou od každé metody požadujeme, je konvergence, tzn. že chceme, aby pro $h \rightarrow 0$ konvergovaly přibližné hodnoty k přesnému řešení. Tento požadavek můžeme přesněji definovat pomocí globální diskretizační chyby e_i . O jednokrokové metodě budeme říkat, že je konvergentní, když



Obr. 5.1

$$(5.22) \quad \lim_{h \rightarrow 0} \max_i |e_i| = 0.$$

Tuto definici můžeme snadno geometricky interpretovat, viz obr. 5.1. Proložme pro zvolené h_1, h_2, \dots body $(x_i, y_i(h_1)), (x_i, y_i(h_2)), \dots$ (kde $y_i(h_1), y_i(h_2), \dots$ jsou přibližná řešení při kroku h_1, h_2, \dots) lomené čáry, pak (5.22) znamená, že posloupnost lomených čar konverguje stejnoměrně k přesnému řešení, jestliže posloupnost h_1, h_2, \dots konverguje k nule.

Bez důkazu, který je uveden např. ve [14], uvedeme větu o konvergenci jednokrokových metod.

Věta 5.1 Každá jednokroková regulérní a konzistentní metoda pro řešení úlohy (5.1), (5.2) je konvergentní.

U jednokrokových metod se výpočet zahajuje z počáteční podmínky $y(a) = y_0$, avšak výpočet k-krokovou metodou můžeme zahájit nejdříve pro y_k a přitom musíme již znát hodnoty y_0, y_1, \dots, y_{k-1} (resp. ještě y_k při implicitní metodě). Tyto tzv. "startovací" hodnoty určíme (s dostatečnou přesností) pomocí vhodné jednokrokové metody.

Příklad 5.4 Řešme úlohu z příkladu 5.1 s krokem $h = 0,05$ a odhadněme chybu e_i podle vzorce (5.21).

Řešení: Výsledky pro krok 0,05 budeme zapisovat jenom pro $x = 0,0; 0,1; \dots, 0,5$, protože jenom v takových bodech můžeme odhad chyby podle (5.21) vypočítat. Řešení je v tabulce 5.4.

i	x_i	$y_i(h)$	$y_i(h/2)$	$ y(x_i) - y_i(h/2) $	e_i
0	0,0	1,000	1,000	0,000	0,000
1	0,1	1,100	1,103	0,003	0,003
2	0,2	1,211	1,217	0,007	0,004
3	0,3	1,336	1,347	0,013	0,011
4	0,4	1,479	1,496	0,019	0,017
5	0,5	1,643	1,668	0,028	0,025

Tab. 5.4

5.3 Vícekrokové metody řešení počáteční úlohy

Jednokrokové metody užívaly k výpočtu y_{i+1} jenom informace vztažené k bodu x_i , a to přibližnou hodnotu řešení a jeho derivaci. Zvyšování přesnosti se dosahovalo užíváním vyšších derivací, což se ve svém důsledku přenášelo na vícenásobný výpočet pravé strany rovnice (5.1). To je obtížné zvláště tehdy, když funkce $f(x, y)$ je složitá. Zdá se tedy být rozumné hledat takové algoritmy pro výpočet y_{i+1} , které využijí hodnoty řešení získané v předcházejících bodech sítě (tedy hodnoty y_i, y_{i-1}, \dots a f_i, f_{i-1}, \dots), aby bylo dosaženo vyšší přesnosti výpočtu. Na této myšlence jsou založeny vícekrokové (k-krokové) metody, které se vesměs dají vyjádřit pomocí vzorce

$$(5.23) \quad y_{i+1} = \sum_{j=1}^k a_j y_{i+1-j} + h \sum_{j=0}^k b_j f_{i+1-j},$$

ve kterém a, b jsou reálné konstanty a $y_{i+1}, y_i, y_{i-1}, \dots, y_{i+1-k}$ (resp. $f_{i+1}, f_i, f_{i-1}, \dots, f_{i+1-k}$) jsou přibližné hodnoty řešení (resp. pravé strany) v bodech $x_{i+1}, x_i, x_{i-1}, \dots, x_{i+1-k}$. Přesněji řečeno, definuje vzorec (5.23) lineární k-krokové metody, protože pravá strana (5.23) je lineární kombinací hodnot f_{i+1-j}, y_{i+1-j} s reálnými koeficienty a_j, b_j .

Je-li ve vzorci (5.23) $b_0 = 0$, pak se nazývá explicitní a při $b_0 \neq 0$ implicitní. Implicitní vzorce přinášejí v algoritmech značné komplikace, které jsou však vyváženy mnoha dobrými vlastnostmi. Je-li $f(x, y)$ lineární vzhledem k y , pak lze implicitní vzorce (5.23) rozřešit vzhledem k y_{i+1} , tedy najít explicitní vztah.

Uvedeme dva jednoduché příklady. Zvolme $k = 1$, $a_1 = 1$, $b_0 = 1$, $b_1 = 0$, pak dostaneme z (5.23) vzorec

$$(5.24) \quad y_{i+1} = y_i + h \cdot f(x_{i+1}, y_{i+1}),$$

který se nazývá implicitní Eulerova metoda. Pro $k = 2$, $a_1 = 0$, $a_2 = 1$, $b_0 = 0$, $b_1 = 2$, $b_2 = 0$ dostaneme tzv. obdélníkové pravidlo

$$(5.25) \quad y_{i+1} = y_{i-1} + 2 \cdot h \cdot f(x_i, y_i).$$

V dalším ukážeme některé způsoby konstrukce vzorců vícekrokových metod.

5.3.1 Adamsovy extrapolační metody

Vícekrokové metody jsou často založeny na numerickém výpočtu integrálů. Z matematické analýzy víme, že řešení úlohy (5.1), (5.2) můžeme pomocí integrálu zapsat takto:

$$y(x) = y(a) + \int_a^x f(x, y(x)) dx,$$

z čehož plyne

$$(5.26) \quad y(x_{i+1}) = y(x_{i+1-m}) + \int_{x_{i+1-m}}^{x_{i+1}} f(x, y(x)) dx.$$

Sestrojme interpolační polynom $(k - 1)$ -ního stupně pro hodnoty

$$\{(x_{i+1-j}, f_{i+1-j}) : j = 1, 2, \dots, k\}, \text{ kde } f_{i+1-j} = f(x_{i+1-j}, y_{i+1-j})$$

a použijeme jej za approximaci funkce $f(x, y)$ na intervalu $\langle x_{i+1-m}, x_{i+1} \rangle$.

Předpokládejme, že $k \geq m$. Funkci $f(x, y)$ jsme vlastně na intervalu $\langle x_i, x_{i+1} \rangle$ approximovali extrapolacním způsobem, tj. prodloužením interpolačního polynomu za poslední uzel, tedy za x_i . Proto se dále odvozené metody budou nazývat extrapolačními. Zapíšeme-li interpolační polynom např. v Lagrangeově tvaru (3.7), potom

$$f(x, y(x)) \approx L_{k-1}(x) = \sum_{j=k}^1 F_{i+1-j}(x) f_{i+1-j}.$$

Dosadíme-li $L_{k-1}(x)$ do (5.26) místo $f(x, y(x))$, dostaneme vzorec

$$(5.27) \quad y_{i+1} = y_{i+1-m} + \int_{x_{i+1-m}}^{x_{i+1}} \sum_{j=k}^1 F_{i+1-j}(x) f_{i+1-j} dx.$$

Provedeme-li ještě substituci $x = x_i + th$, pak (5.27) přejde na tvar

$$(5.28) \quad y_{i+1} = y_{i+1-m} + h \cdot \sum_{j=1}^k b_j f_{i+1-j}, \text{ kde } b_j = \int_{1-m}^1 F_{i+1-j}(x_i + th) dt.$$

Vzorec (5.28) definuje Adamsovy extrapolacní metody (AE - metody) jako explicitní k -krokové metody, protože (5.28) neobsahuje hodnotu f_{i+1} . Ve většině případů se ve vzorci (5.28) volí $m = 1$, tedy

$$(5.29) \quad y_{i+1} = y_i + h \cdot \sum_{j=1}^k b_j f_{i+1-j}, \text{ kde } b_j = \int_0^1 F_{i+1-j}(x_i + th) dt.$$

Při odvozování konkrétních vzorců této metody můžeme samozřejmě užívat různé tvarы interpolačních polynomů. Jako příklad sestrojme explicitní vzorec AE-metody pro $m = 1$ a $k = 3$. Funkci $f(x, y)$ nahradíme Newtonovým interpolačním polynomem (3.24), sestrojeným pro hodnoty $(x_{i-2}, f_{i-2}), (x_{i-1}, f_{i-1}), (x_i, f_i)$. Analogicky k (5.27) bude mít hledaný vzorec tvar

$$y_{i+1} = y_i + \int_{x_i}^{x_{i+1}} N_2(x) dx = y_i + h \int_0^1 N_2(x_i + th) dt,$$

kde podle (3.24) platí

$$\begin{aligned} y_{i+1} &= y_i + h \int_0^1 \left(f_i + t \Delta^1 f_{i-1} + \frac{t(t+1)}{2!} \Delta^2 f_{i-2} \right) dt = \\ &= y_i + \frac{h}{12} (23f_i - 16f_{i-1} + 5f_{i-2}). \end{aligned}$$

Velké množství vzorců AE-metody je uvedeno např. v [16], ale také si je může čtenář snadno sestavit samostatně.

5.3.2 Adamsovy interpolační metody

Ve vzorcích AE-metody je funkce $f(x, y)$, jak jsme již řekli, approximována extrapolačním způsobem. Této situaci se Adamsovy interpolační metody (AI-metody) vyhýbají tím, že interpolační polynom funkce $f(x, y)$ sestrojí pro data

$$\{(x_{i+1-j}, f_{i+1-j}) : j = 0, 1, \dots, k\}.$$

Při použití Lagrangeova interpolačního polynomu bude platit

$$f(x, y(x)) \approx L_k(x) = \sum_{j=0}^k F_{i+1-j}^{(j)} \cdot f_{i+1-j}.$$

Analogickým postupem, kterým jsme v odstavci 5.3.1 odvodili vzorec (5.28), dostaneme nyní vzorec

$$(5.31) \quad y_{i+1} = y_{i+1-m} + h \cdot \sum_{j=0}^k b_j f_{i+1-j}, \quad \text{kde } b_j = \int_{1-m}^1 F_{i+1-j}(x_i + th) dt,$$

který definuje Adamsovy interpolační metody jako implicitní k-krokové metody. I v těchto vzorcích volíme obvykle $m = 1$.

Vzhledem k tomu, že (5.31) je implicitní metoda, představuje tento vzorec (obecně) nelineární rovnici

$$(5.32) \quad y_{i+1} = h \cdot b_0 \cdot f(x_{i+1}, y_{i+1}) + c, \\ \text{kde } c = y_{i+1-m} + h \cdot \sum_{j=1}^k b_j f_{i+1-j},$$

kterou musíme řešit vzhledem k y_{i+1} . Pro řešení si vybereme nějakou metodu ze čtvrté kapitoly.

AI-metody jsou sice složitější, ale mají větší přesnost než AE-metody (při též k).

Najděme např. interpolační vzorec pro $m = 1$ a $k = 3$, když užijeme opět interpolační polynom (3.24) pro tabulku $\{(x_{i+1-j}, f_{i+1-j}) : j = 0, 1, 2, 3\}$, který má nyní tvar

$$N_3(x) = N_3(x_{i+1} + th) = f_{i+1} + t \Delta^1 f_i + \frac{t(t+1)}{2!} \Delta^2 f_{i-1} + \\ + \frac{t(t+1)(t+2)}{3!} \Delta^3 f_{i-2}.$$

Podle (5.26) dostaneme

$$y_{i+1} = y_i + \int_{x_i}^{x_{i+1}} N_3(x) dx = y_i + h \int_0^1 [f_{i+1} + t \Delta^1 f_i + \\ + \frac{t(t+1)}{2!} \Delta^2 f_{i-1} + \frac{t(t+1)(t+2)}{3!} \Delta^3 f_{i-2}] dt,$$

tedy

$$(5.33) \quad y_{i+1} = y_i + \frac{h}{24} (9f_{i+1} + 19f_i - 5f_{i-1} + f_{i-2}) .$$

5.3.3 Konstrukce vícekrokových metod

Vráťme se k obecnému vzorci k-krokových metod (5.22) a ukážeme postup určování koeficientů a_j a b_j metodou, kterou nazýváme metoda neurčitých koeficientů. Pro tento účel zapišeme k-krokovou metodu vztahem

$$(5.34) \quad \sum_{j=0}^k a_j y_{i+j} = h \sum_{j=0}^k b_j y'_{i+j} ,$$

který je v matematické literatuře obvyklý. Řešení $y(x)$ (úlohy (5.1), (5.2)) a jeho derivaci $y'(x)$ vyjádříme za předpokladu dostatečné hladkosti pomocí Taylorovy věty v okolí bodu x_i ve tvaru

$$y(x_{i+j}) = y(x_i) + jhy'(x_i) + \frac{(jh)^2}{2!} y''(x_i) + \dots + \frac{(jh)^p}{p!} y^{(p)}(x_i) + R ,$$

$$y'(x_{i+j}) = y'(x_i) + jhy''(x_i) + \dots + \frac{(jh)^{p-1}}{(p-1)!} y^{(p)}(x_i) + S .$$

Vyjádříme-li zbytky R , S v Lagrangeově tvaru a užijeme-li sumičních znaků, pak

$$y(x_{i+j}) = \sum_{s=0}^p \left[\frac{(jh)^s}{s!} y^{(s)}(x_i) \right] + \frac{(jh)^{p+1}}{(p+1)!} y^{(p+1)}(\xi) ,$$

(5.35)

$$y'(x_{i+j}) = \sum_{s=1}^p \left[\frac{(jh)^{s-1}}{(s-1)!} y^{(s)}(x_i) \right] + \frac{(jh)^p}{p!} y^{(p+1)}(\xi) ,$$

kde $\xi \in (x_i, x_{i+j})$. Sumy ve vztazích (5.35) můžeme považovat za přibližné hodnoty $y(x_{i+j})$, $y'(x_{i+j})$, tedy je dosadíme do (5.34) a tím dostaneme rovnost dvou polynomů v proměnné h

$$a_0 y_i + \sum_{j=1}^k a_j \sum_{s=0}^p \frac{(jh)^s}{s!} y^{(s)}(x_i) = h b_0 y'_i + h \sum_{j=1}^k b_j \sum_{s=1}^p \frac{(jh)^{s-1}}{(s-1)!} y^{(s)}(x_i) ,$$

která je splněna právě tehdy, když se rovnají koeficienty u h^0, h^1, \dots, h^p . Zaměníme-li pořadí sumování, pak z porovnání koeficientů dostaneme rovnice

$$(5.37) \quad \begin{aligned} \sum_{j=0}^k a_j &= 0 , \\ \sum_{j=1}^k j a_j - \sum_{j=0}^k b_j &= 0 , \\ \sum_{j=1}^k j^s a_j - \sum_{j=1}^k s j^{s-1} b_j &= 0 , \quad \text{kde } s = 2, 3, \dots, p , \end{aligned}$$

ze kterých určíme a_j , b_j . Protože konstant je $2(k+1)$, ale rovnic $p+1$, pak p může být maximálně rovno $2k+1$.

O vzorcích (5.34) říkáme, že jsou p-tého řádu, protože po dosazení (5.32) do (5.34) jsme zanedbali členy s $(p+1)$ -ní mocninou h .

Příklad 5.5 Sestavme explicitní 2-krokovou metodu třetího řádu.

Řešení: Hledaný vzorec má tvar

$$a_0 y_i + a_1 y_{i+1} + a_2 y_{i+2} = h(b_0 y'_i + b_1 y'_{i+1} + b_2 y'_{i+2}).$$

Podle (5.37) dostaneme soustavu ($k=2$, $p=3$) rovnic

$$a_0 + a_1 + a_2 = 0,$$

$$a_1 + 2a_2 - b_0 - b_1 - b_2 = 0,$$

$$a_1 + 4a_2 - 2b_1 - 4b_2 = 0,$$

$$a_1 + 8a_2 - 3b_1 - 12b_2 = 0,$$

ve které můžeme dvě neznámé zvolit, protože $k=2$ a $p=3$. Zvolme $a_2=1$ a $b_2=0$, protože chceme dostat explicitní vzorec. Řešením soustavy dostaneme $a_0=-5$, $b_0=2$, $a_1=4$ a $b_1=4$, tedy hledaný vzorec má tvar

$$y_{i+2} = -4y_{i+1} + 5y_i + h(4y'_{i+1} + 2y'_i).$$

Přejdeme-li ke značení zavedenému dříve, pak dostaneme hledaný vzorec ve tvaru

$$y_{i+1} = -4y_i + 5y_{i-1} + h(4f_i + 2f_{i-1}).$$

5.3.4 Chyby vícekrokových metod

Při odvozování k-krokové metody v odstavci 5.3.3 můžeme do vzorce (5.34) dosadit z (5.35) celé pravé strany a pak se lokální diskretizační chyba odhadne rozdílem

$$(5.38) \quad \sum_{j=1}^k \left(a_j \frac{j^{p+1}}{(p+1)!} - b_j \frac{j^p}{p!} \right) h^{p+1} y^{(p+1)}(\xi) = c_{p+1} h^{p+1} y^{(p+1)}(\xi).$$

Lokální přesnost k-krokové metody p-tého řádu charakterizujeme číslem c_{p+1} , které je u běžně užívaných metod vypočítáno předem podle (5.38).

Znovu připomeňme, že při odvozování k-krokových metod jsme předpokládali, že řešení je dostatečně hladké, tedy že má alespoň $p+1$ spojitých derivací. Pokud užijeme metodu p-tého řádu pro počáteční úlohu, jejíž řešení má nižší hladkost, pak lokální chyba bude obecně větší. Z toho plyne, že předpokladem efektivního užívání metod vyšších řádů je podmínka dostatečné hladkosti řešení, čili funkce $f(x, y)$. Stejně tak je tomu u metod jednokrokových.

Studium globálních diskretizačních chyb je mnohem komplikovanější a přesahuje rozsah těchto skript, proto se jím nebudeme zabývat a odkážeme čtenáře na [26] nebo [4].

Při řešení počáteční úlohy k-krokovou metodou nás ještě zajímá konvergence přibližných hodnot k přesnému řešení. Pro zahájení výpočtu k-krokovou metodou musíme již znát přibližné hodnoty y_0, y_1, \dots, y_{k-1} (kde $y_0 = y(x_0)$, viz (5.2)), které vypočí-

táme např. nějakou jednokrokovou metodou. Tyto hodnoty závisejí na h , tedy $y_s = y_s(h)$ pro $s = 1, 2, \dots, k-1$. Jistě požadujeme, aby

$$(5.39) \quad \lim_{h \rightarrow 0} y_s(h) = y_0.$$

O k -krokové metodě budeme říkat, že je konvergentní, když platí

$$(5.40) \quad \lim_{\substack{h \rightarrow 0 \\ x_{i+1} \rightarrow x}} y_{i+1} = y(x)$$

pro každé $x \in (a, b)$ a každé y_{i+1} počítané podle (5.23), resp. (5.34). Přitom předpokládáme, že pro startovní hodnoty platí (5.39).

Aby k -kroková metoda byla konvergentní, musí být konzistentní a stabilní. Metoda, která je alespoň 1. řádu, se nazývá konzistentní.

Pojem stability je poněkud komplikovanější a souvisí s vlivem tzv. parazitních složek řešení, který u stabilních metod pro $h \rightarrow 0$ mizí. Podle [1] můžeme o stabilitě rozhodnout na základě koeficientů a_0, a_1, \dots, a_k . Vícekroková metoda (5.34) je stabilní, jestliže polynom

$$p(x) = \sum_{j=0}^k a_j x^j$$

má kořeny z intervalu $(-1, 1)$ a je-li kořen 1 nebo -1, pak musí být jednoduchý. Všechny k -krokové metody, zapsané v tabulce 5.5, jsou stabilní.

	k	p	Vzorec	$ c_{p+1} \cdot h^{p+1} $
1.	2	2	$y_{i+1} = y_i + h/2(3y'_i - y'_{i-1})$	$5/12 \cdot h^3$
2.	1	2	$y_{i+1} = y_i + h/2(y'_{i+1} + y'_i)$	$1/12 \cdot h^3$
3.	3	2	$y_{i+1} = y_{i-2} + 3h/2(y'_i + y'_{i-1})$	$3/4 \cdot h^3$
4.	2	3	$y_{i+1} = y_i + h/12(5y'_{i+1} + 8y'_i - y'_{i-1})$	$1/24 \cdot h^4$
5.	3	4	$y_{i+1} = y_i + (y_{i-1} - y_{i-2})/9 + h/27(10y'_{i+1} + 22y'_i - 8y'_{i-1})$	$19/810 \cdot h^5$
6.	3	4	$y_{i+1} = y_{i-3} + 4h/3(2y'_i - y'_{i-1} + 2y'_{i-2})$	$14/15 \cdot h^5$
7.	2	4	$y_{i+1} = y_{i-1} + h/3(y'_{i+1} + 4y'_i + y'_{i-1})$	$1/90 \cdot h^5$

Tab. 5.5

5.3.5 Algoritmus prediktor-korektor

Všimněte si 2. a 3. vzorce v tabulce 5.5. Vzorec 2. je implicitní a 3. je explicitní. Porovnáme-li čísla c_{p+1} , vidíme, že vzorec 2. je asi 9krát přesnější než vzorec 3. I obecně platí, že implicitní vzorce jsou při též řádu podstatně přesnější. Proto si nyní ukážeme vhodný algoritmus pro užívání implicitních metod.

Především předpokládejme, že známe dostatečně přesně počáteční hodnoty y_0, y_1, \dots, y_{k-1} , spočítané např. některou jednokrokovou metodou.

Pro výpočet y_{i+1} užijeme dvou k-krokových metod (nejvhodněji stejného řádu). Nejprve to bude explicitní metoda, kterou budeme nazývat prediktor a její approximaci hodnoty $y(x_{i+1})$ označíme $y_{i+1}^{(0)}$. Dále bude následovat užití implicitní metody, které říkáme korektor. Iteračním řešením rovnice, kterou definuje korektor, počítáme z $y_{i+1}^{(0)}$ postupně $y_{i+1}^{(1)}, y_{i+1}^{(2)}, \dots$. Takový algoritmus výpočtu y_{i+1} se nazývá algoritmus prediktor-korektor. Výpočet hodnoty y_{i+1} tedy probíhá takto:

- 1) Položíme $s = 0$ a pomocí nějakého explicitního vzorce - prediktoru - vypočítáme přibližnou hodnotu řešení v bodě x_{i+1} a označíme ji $y_{i+1}^{(s)}$, tedy

$$y_{i+1}^{(s)} = \sum_{j=1}^k a_j y_{i+1-j} + h \sum_{j=1}^k b_j f_{i+1-j} \quad \text{- prediktor.}$$

- 2) Iteračním krokem zpřesníme $y_{i+1}^{(s)}$ pomocí některého implicitního vzorce a zpřesněnou hodnotu označíme $y_{i+1}^{(s+1)}$, tedy

$$y_{i+1}^{(s+1)} = \sum_{j=1}^k a_j y_{i+1-j} + h \left[b_0 f(x_{i+1}, y_{i+1}^{(s)}) + \sum_{j=1}^k b_j f_{i+1-j} \right] \quad \text{- korektor.}$$

- 3) Je-li $|y_{i+1}^{(s+1)} - y_{i+1}^{(s)}| < \varepsilon$ (ε - zadaná nepřesnost po sobě následujících iteracích), pak $y_{i+1} = y_{i+1}^{(s+1)}$ a výpočet přechází na další uzlový bod. Není-li nerovnost splněna, zvětšíme s o jedničku a opakujeme algoritmus od bodu 2).

Jedním z nejstarších algoritmů prediktor-korektor je Milnova metoda, která užívá těchto vzorců:

$$(5.41) \quad \begin{aligned} \text{prediktor: } y_{i+1}^{(0)} &= y_{i-3} + \frac{4}{3} h(2y'_i - y'_{i-1} + 2y'_{i-2}), \\ \text{korektor: } y_{i+1}^{(s+1)} &= y_{i-1} + \frac{1}{3} h(f(x_{i+1}, y_{i+1}^{(s)}) + 4y'_i + y'_{i-1}), \end{aligned}$$

viz tab. 5.5, vzorce 6. a 7.

Pro metodu prediktor-korektor můžeme užít také 3. a 2. vzorce z tabulky 5.5, na které jsme upozornili v začátku tohoto odstavce. Tedy

$$\begin{aligned} \text{prediktor: } y_{i+1}^{(0)} &= y_{i-2} + \frac{3}{2} h(y'_i + y'_{i-1}), \\ \text{korektor: } y_{i+1}^{(s+1)} &= y_i + \frac{1}{2} h(f(x_{i+1}, y_{i+1}^{(s)}) + y'_i). \end{aligned}$$

Na samočinných počítačích se nejvíce užívá metod prediktor-korektor 4. řádu, protože tyto metody zaručují dostatečnou přesnost při většině problémů a jsou i dostačně jednoduché pro odvozování. Dále je z mnoha důvodů (které nebudeme upřesňovat) vhodné, volit vzorce prediktoru a korektoru stejného řádu přesnosti. Často se také kritérium přesnosti výpočtu hodnoty y_{i+1} (které je dané číslem ε) nahrazuje předepsaným počtem iterací korektorem. Tím se sníží pracnost výpočtu a zabrání se možnosti vzniku nekonečného cyklu.

Metody prediktor-korektor mají ještě mnoho zajímavých vlastností a různých modifikací, proto odkazujeme čtenáře na rozsáhlější literaturu [23], [13] apod.

Příklad 5.6 Řešme algoritmem prediktor-korektor počáteční úlohu $y' = x - y$, $y(0) = 1$ na intervalu $\langle 0, 1 \rangle$. Přitom za dvojici prediktor-korektor zvolme 1. a 2. vzorec z tabulky 5.5, tedy

$$y_{i+1}^{(0)} = y_i + \frac{1}{2} h (3y_i' - y_{i-1}') \quad - \text{prediktor},$$

$$y_{i+1}^{(s+1)} = y_i + \frac{1}{2} h (f(x_{i+1}), y_{i+1}^{(s)}) + y_i' \quad - \text{korektor}.$$

Korektorem provedeme vždy pouze jeden krok, čili $y_{i+1} = y_{i+1}^{(1)}$.

i	x_i	$y(x_i)$	$y_i^{(0)}$	$y_i = y_i^{(1)}$
0	0	1,000	-	-
1	0,2	0,837	-	-
2	0,4	0,741	0,746	0,739
3	0,6	0,698	0,701	0,695
4	0,8	0,699	0,700	0,696
5	1,0	0,736	0,737	0,733

Tab. 5.6

Řešení: Zvolíme $h = 0,2$ a za počáteční hodnoty y_0, y_1 , potřebné ke startu algoritmu vezmeme pro jednoduchost hodnoty přesného řešení $y(x) = 2e^{-x} + x - 1$. Výsledky jsou seštaveny v tabulce 5.6.

Všimněte si, že $y_i^{(0)}$ dávají přesnější approximační řešení než $y_i^{(1)}$. Je to způsobeno nevhodnou volbou ukončení algoritmu.

5.4 Okrajové úlohy

Okrajovou úlohou rozumíme úlohu o nalezení takové funkce $y(x)$, která na intervalu $\langle a, b \rangle$ splňuje diferenciální rovnici a jisté podmínky v krajních bodech a, b , které jsou definovány spojitými funkcemi. Proto se tato úloha nazývá také dvoubodová. Metody řešení budeme demonstrovat na okrajové úloze pro samoadjungovanou lineární diferenciální rovnici 2. rádu

$$(5.42) \quad - [p(x)y'(x)]' + q(x)y(x) = r(x), \quad x \in \langle a, b \rangle,$$

s okrajovými podmínkami tří typů

- (5.43)
- 1. typ $y(a) = \alpha, y(b) = \beta$ - Dirichletova úloha,
 - 2. typ $y'(a) = \alpha, y'(b) = \beta$ - Neumanova úloha,
 - 3. typ $\begin{aligned} -\alpha_1 y(a) + \alpha_2 y'(a) &= \alpha \\ \beta_1 y(b) + \beta_2 y'(b) &= \beta \end{aligned}$ - smíšená úloha,

kde funkce $p(x), p'(x), q(x), r(x)$ jsou spojité pro $x \in \langle a, b \rangle$ a $q(x) \geq 0$ pro $x \in \langle a, b \rangle$, $q(x) \neq 0$. Dále předpokládejme, že α, β jsou reálná čísla a $\alpha_1, \alpha_2, \beta_1, \beta_2$ jsou nezáporná čísla, která splňují podmínky $\alpha_1 + \alpha_2 > 0, \beta_1 + \beta_2 > 0$.

Takto definovaná úloha má právě jedno řešení pro libovolná α, β .

5.4.1 Metoda sítí pro rovnice 2. řádu

Mějme dánu okrajovou úlohu pro rovnici (5.42) s podmínkami (5.43). Metoda sítí pro řešení této úlohy spočívá v tom, že přibližné hodnoty řešení hledáme v konečné síti bodů intervalu $\langle a, b \rangle$, a to jako řešení soustavy lineárních algebraických rovnic, na kterou zadanou úlohu převedeme. Rovnice této soustavy dostaneme tak, že pro každý vnitřní bod dělení intervalu $\langle a, b \rangle$ zapíšeme diferenciální rovnici (5.42), ve které diskretizujeme zúčastněné funkce a derivace nahradíme nějakou lineární kombinací přibližných hodnot řešení v sousedních bodech. Stejným způsobem přepíšeme i okrajové podmínky v bodech a, b .

Nyní tedy přesněji. Rozdělíme interval $\langle a, b \rangle$ ekvidistantní sítí bodů x_0, x_1, \dots, x_n , tedy $a = x_0 < x_1 < \dots < x_n = b$, kde $x_i = x_0 + ih$ a $h = (b - a)/n$. Dále označme y_0, y_1, \dots, y_n přibližné hodnoty řešení $y(x)$ v bodech x_0, x_1, \dots, x_n . Protože v každém vnitřním bodě dělení má být splněna rovnice (5.42) a v krajních bodech okrajová podmínka (5.43) (např. třetího typu), můžeme napsat soustavu rovnic

$$(5.44) \quad \begin{aligned} -\alpha_1 \cdot y_0 + \alpha_2 \cdot y'(x)|_{x=a} &= \alpha, \\ -[p(x) \cdot y'(x)]'|_{x=x_i} + q(x_i) \cdot y_i &= r(x_i), \quad i = 1, 2, \dots, n-1, \\ \beta_1 \cdot y_n + \beta_2 \cdot y'(x)|_{x=b} &= \beta, \end{aligned}$$

ve které je ještě třeba nahradit derivace pomocí nějakého jednoduchého vzorce numerického derivování. Podle odstavce 3.7 provedeme např. tyto náhrady

$$(5.45) \quad \begin{aligned} y'(x)|_{x=x_i} &\approx \frac{y_{i+1} - y_i}{h}, \quad \text{pro } i = 0, 1, 2, \dots, n-1, \\ y'(x)|_{x=b} &\approx \frac{y_n - y_{n-1}}{h}, \\ [p(x) \cdot y'(x)]'|_{x=x_i} &\approx \frac{1}{h} \left[p_{i+1/2} \frac{y_{i+1} - y_i}{h} - p_{i-1/2} \cdot \frac{y_i - y_{i-1}}{h} \right], \end{aligned}$$

pro $i = 1, 2, \dots, n-1$, kde $p_{i+1/2} = p(a + (i + 1/2)h)$ a $p_{i-1/2} = p(a + (i - 1/2)h)$. Koeficienty p bereme ve středu intervalů $\langle x_{i-1}, x_i \rangle$, $\langle x_i, x_{i+1} \rangle$, abychom zvýšili přesnost aproximace. Dosadme nyní (5.45) do (5.44) a označme $q_i = q(x_i)$, $r_i = r(x_i)$, tedy

$$(5.46) \quad \begin{aligned} -\alpha_1 \cdot y_0 + \alpha_2 \cdot (y_1 - y_0)/h &= \alpha, \\ -\frac{1}{h} \left[p_{i+1/2} \frac{y_{i+1} - y_i}{h} - p_{i-1/2} \cdot \frac{y_i - y_{i-1}}{h} \right] + q_i \cdot y_i &= r_i, \quad i = 1, 2, \dots, n-1, \\ \beta_1 \cdot y_n + \beta_2 \cdot (y_n - y_{n-1})/h &= \beta. \end{aligned}$$

To je hledaná soustava metody sítí s $n+1$ lineárními algebraickými rovnicemi o $n+1$ neznámých y_0, y_1, \dots, y_n .

Věta 5.2 Jsou-li splněny předpoklady okrajové úlohy (5.42), (5.43), pak soustava lineárních algebraických rovnic (5.46) má pro libovolnou funkci $r(x)$ a libovolné α, β právě jedno řešení.

Důkaz je uveden např. ve [4], [26].

Soustavu (5.46) můžeme řešit např. Gaussovou eliminační metodou, ve které využijeme tu skutečnost, že soustava je třídiagonální. Soustavu lze řešit i Gaussovou - Seidelovou iterační metodou, protože její matice je dokonce pozitivně definitní, důkaz viz [4].

Dále nás bude zajímat otázka konvergance metody sítí, tzn. zdali volbou dostatečně malého h můžeme udělat rozdíl $y_i - y(x_i)$ libovolně malý pro všechna i . Tuto skutečnost vyjadřuje následující věta, ze které můžeme odvodit i rychlosť konvergencie.

Věta 5.3 Nechť jsou splněny předpoklady okrajové úlohy (5.42), (5.43) a nechť má navíc funkce $p(x)$ tři spojité derivace a $q(x)$, $r(x)$ dvě spojité derivace na intervalu $\langle a, b \rangle$. Je-li y_0, y_1, \dots, y_n řešení soustavy (5.46), pak existuje taková konstanta M (nezávislá na h), že platí

$$(5.47) \quad |y_i - y(x_i)| \leq M \cdot h \quad \text{pro } i = 0, 1, \dots, n.$$

Jestliže bude ještě platit $\alpha_k > 0$, $\beta_k > 0$ pro $k = 1, 2$, pak

$$(5.48) \quad |y_i - y(x_i)| \leq M \cdot h^2 \quad \text{pro } i = 0, 1, \dots, n.$$

Příklad 5.7 Řešme metodou sítí okrajovou úlohu pro rovnici $-y'' + y = 0$ na intervalu $\langle 0, 1 \rangle$, s okrajovými podmínkami $y'(0) - y(0) = 1$, $y'(1) + y(1) = 1$.

Řešení: Rozdělme interval $\langle 0, 1 \rangle$ body x_0, x_1, \dots, x_{10} , které představují ekvidistantní dělení s krokem $h = 0,1$. První a druhou derivaci budeme podle (3.91) a (3.92) nahrazovat vzorcí

$$y'(0) \approx \frac{y_1 - y_0}{h}, \quad y''(x_i) \approx \frac{1}{h^2} (y_{i-1} - 2y_i + y_{i+1}), \quad y'(1) \approx \frac{y_{10} - y_9}{h},$$

tedy soustava (5.46) bude mít tvar

$$\begin{aligned} 1, ly_0 - y_1 &= -0,1, \\ y_{i-1} - 2,0ly_i + y_{i+1} &= 0, \quad i = 1, 2, \dots, 9, \\ y_9 - 1, ly_{10} &= -0,1. \end{aligned}$$

Výsledky řešení jsou v následující tabulce a pro srovnání jsou uvedeny hodnoty přesného řešení $y(x) = -e^{-x}/2$.

i	0	1	2	3	4	5	6	7	8	9	10
x_i	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0
y_i	-0,512	-0,464	-0,420	-0,380	-0,344	-0,312	-0,283	-0,256	-0,231	-0,209	-0,190
$y(x_i)$	-0,500	-0,452	-0,409	-0,370	-0,335	-0,303	-0,274	-0,248	-0,225	-0,203	-0,184

Tab. 5.7

5.4.2 Variační metody

V tomto odstavci se budeme zabývat řešením lineární diferenciální rovnice s homogenními okrajovými podmínkami a nehomogenní pravou stranou. Pro takové úlohy bude zapisovat diferenciální rovnici symbolicky ve tvaru

$$(5.49) \quad A u(x) = f(x), \quad x \in \langle a, b \rangle$$

a okrajové podmínky ve tvaru

$$(5.50) \quad L_m^{(a)} u = 0, \quad L_m^{(b)} u = 0, \quad m = 1, 2, \dots, k,$$

kde A je lineární diferenciální operátor $2k$ -tého řádu (tzn. že nejvyšší derivace, kterou obsahuje, je spojitá a je řádu $2k$) se spojitými koeficienty $a_i(x)$, tedy

$$(5.51) \quad A u = \sum_{i=2k}^0 a_i(x) \cdot u^{(i)}(x),$$

$f(x)$ je spojitá funkce pravé strany a $L_m^{(a)}$, $L_m^{(b)}$ jsou lineární diferenciální operátory nejvýše $(k-1)$ -ního řádu.

Tuto symbolikou přepíšeme rovnici (5.42) s okrajovými podmínkami třetího typu. Potom ($k=1$)

$$A u(x) = - (p(x) \cdot u'(x))' + q(x), \quad f(x) = r(x),$$

$$L_1^{(a)} = \alpha_2 \cdot u'(a) - \alpha_1 \cdot u(a),$$

$$L_1^{(b)} = \beta_2 \cdot u'(b) + \beta_1 \cdot u(b).$$

Řešení úloh s nehomogenními okrajovými podmínkami se dá převést na řešení homogenní úlohy tím, že se hledané řešení vyjádří ve tvaru součtu dvou funkcí, z nich jedna splňuje nehomogenní okrajové podmínky a nesplňuje diferenciální rovnici a druhou najdeme jako řešení úlohy s homogenními podmínkami, která se zadánou rovnici jistým způsobem souvisí. Podrobně viz [24].

Abychom mohli vyložit Ritzovu variační metodu pro řešení těchto úloh, uvedeme nejdříve některé vlastnosti lineárních operátorů v prostoru $L^2 \langle a, b \rangle$.

Mějme dán lineární diferenciální operátor A v prostoru $L^2 \langle a, b \rangle$, jehož definiční obor D_A je množina hustá v $L^2 \langle a, b \rangle$. Takový operátor se nazývá symmetrický, jestliže

$$(5.52) \quad (Au, v) = (u, Av) \quad \text{pro libovolné } u, v \in D_A.$$

Symetrický operátor se nazývá pozitivní, resp. pozitivně definitní, jestliže pro $u \in D_A$

$$(5.53) \quad (Au, u) \geq 0 \quad \text{a} \quad (Au, u) = 0 \quad \text{pouze pro } u = 0,$$

$$(5.54) \quad \text{resp.} \quad (Au, u) \geq \gamma^2 \|u\|^2, \quad \text{kde } \gamma \text{ je konstanta.}$$

Příklad 5.8 Dokažme, že lineární diferenciální operátor $Au = -u''$ s definičním oborem

$$D_A \equiv \left\{ u \in L^2 \langle 0, 1 \rangle : u, u', u'' \text{ jsou spojité v } \langle 0, 1 \rangle \right\}$$

a vyhovují okrajovým podmínkám $u(0) = u(1) = 0 \}$

je pozitivně definitní.

Řešení: Pro $u, v \in L^2[0, 1]$ je $(u, v) = \int_0^1 u' v' dx$, $\|u\| = \sqrt{(u, u)}$, $\varphi = \|u - v\|$.

Dále pro $u, v \in D_A$ platí na základě integrace per partes

$$(Au, v) = - \int_0^1 u'' v' dx = \int_0^1 u' v' dx = - \int_0^1 u' v'' dx = (u, Av),$$

tzn., že operátor A je symetrický. Je však také pozitivní, protože $(Au, u) = \int_0^1 (u')^2 dt \geq 0$ a $(Au, u) = 0$ jedině pro $u' \equiv 0$, tedy pro $u = \text{konstanta}$, která musí být v důsledku okrajových podmínek nulová, čili pro $u \equiv 0$.

Zbývá ukázat, že je operátor A také pozitivně definitní. Podle Schwartzovy nerovnosti, kterou známe z integrálního počtu, platí

$$(5.55) \quad \left(\int_0^1 u' v' dx \right)^2 \leq \int_0^1 u'^2 dx \cdot \int_0^1 v'^2 dx.$$

Z okrajové podmínky $u(0) = 0$ a předcházející nerovnosti plyne

$$u(x) = \int_0^x u' dt, \text{ pak } u^2(x) = \left(\int_0^x u' dt \right)^2 \leq \int_0^x 1^2 dt \int_0^x (u')^2 dt \leq x \int_0^x (u')^2 dt.$$

Protože $x \in [0, 1]$ a $(u')^2 \geq 0$ v $[0, 1]$, pak z poslední nerovnosti vyplývá

$$u^2(x) \leq x \int_0^1 (u')^2 dt, \text{ tedy také } \int_0^1 u'^2 dx \leq \int_0^1 x dx \int_0^1 (u')^2 dt,$$

neboli

$$\int_0^1 (u')^2 dt \geq 2 \int_0^1 u'^2 dx, \text{ pak } (Au, u) \geq 2 \|u\|^2.$$

Tedy operátor A je pozitivně definitní.

Dále se věnujme řešení rovnice

$$(5.56) \quad Au = f \text{ pro } f \in L^2[a, b],$$

kde A je pozitivně definitní lineární diferenciální operátor. Protože operátor A je pozitivní, pak rovnice (5.56) má maximálně jedno řešení, což lze snadno dokázat sporem:

Předpokládejme, že řešení jsou dvě, a to u_1, u_2 . Potom $Au_1 = f$ a $Au_2 = f$, čili $Au_1 - Au_2 = 0$. V důsledku linearity operátoru A také $A(u_1 - u_2) = 0$. Vynásobíme-li poslední rovnici skalárně prvkem $u_1 - u_2$, dostaneme

$$(A(u_1 - u_2), u_1 - u_2) = 0,$$

což je však u pozitivního operátoru splněno pouze tehdy, když $u_1 = u_2 = 0$, čili pro $u_1 = u_2$. Tedy řešení je jediné.

K rovnici (5.56) definujme ještě kvadratický funkcionál

$$(5.57) \quad F(u) = (Au, u) - 2(f, u), \quad \text{pro } u \in D_A,$$

jehož minimální hodnota se významně váže k řešení rovnice (5.56). Tato vazba je dána následující větou, která se nazývá věta o minimu kvadratického funkcionálu.

Věta 5.4 Mějme dánu rovnici (5.56) s pozitivně definitním operátorem A. Má-li rovnice (5.56) řešení $u_0 \in D_A$, tj. platí-li

$$Au_0 = f, \quad u_0 \in D_A,$$

pak funkcionál (5.57) nabývá v D_A minimální hodnoty právě pro prvek u_0 .

Naopak, nechť funkcionál (5.57) nabývá pro prvek $u_0 \in D_A$ minimální hodnoty, pak u_0 je řešením rovnice (5.56).

Důkaz věty najdete např. v [17], [24].

Věta 5.4 má podmíněný charakter. Vyjadřuje ekvivalence řešení rovnice (5.56) s hledáním minima funkcionálu (5.57) za předpokladu, že je známa existence řešení nebo existence minima.

Předpokládejme, že řešení rovnice (5.56) existuje, a hledejme je jako minimum funkcionálu (5.57) pomocí tzv. minimalizující posloupnosti. Jestliže označíme d minimum funkcionálu (5.57), pak libovolná posloupnost $\{u_n\} \in D_A$ se nazývá minimalizující, jestliže

$$d = \lim_{n \rightarrow \infty} \{F(u_n)\}.$$

Protože operátor A je pozitivně definitní, konverguje minimalizující posloupnost k hledanému řešení v tom smyslu, že

$$\lim_{n \rightarrow \infty} (A(u - u_n), u - u_n) = 0, \quad \text{tedy podle (5.54) také}$$

$$\lim_{n \rightarrow \infty} (u - u_n, u - u_n) = 0.$$

Základní metoda pro hledání minimalizující posloupnosti je Ritzova metoda. Zvolme v D_A posloupnost tzv. souřadnicových funkcí $\{\varphi_n\}$, které splňují tyto požadavky:

(5.58) $\{\varphi_n\}$ je lineárně nezávislý systém,

(5.59) $\{\varphi_n\}$ je úplný systém podle energie operátoru A, tzn. že pro $u \in D_A$ a $\varepsilon > 0$ existuje n a čísla a_1, a_2, \dots, a_n tak, že

$$\left(A \left(u - \sum_{i=1}^n a_i \varphi_i \right), u - \sum_{i=1}^n a_i \varphi_i \right) < \varepsilon.$$

Proto budeme prvky minimalizující posloupnosti hledat ve tvaru

$$(5.60) \quad u_n = \sum_{i=1}^n a_i \varphi_i.$$

Dosadíme-li (5.60) do $F(u)$, dostaneme

$$F(u_n) = \left(A \sum_{i=1}^n a_i \varphi_i, \sum_{i=1}^n a_i \varphi_i \right) - 2 \left(f, \sum_{i=1}^n a_i \varphi_i \right),$$

což je funkce n argumentů a_1, a_2, \dots, a_n , která dosahuje své minimální hodnoty pro taková a_1, a_2, \dots, a_n , pro která platí

$$(5.61) \quad \frac{\partial F(u_n)}{\partial a_i} = 0 \quad \text{pro } i = 1, 2, \dots, n.$$

Jestliže v soustavě (5.61) vypočítáme naznačené derivace a příslušně je upravíme, pak dostaneme tvar

$$(5.62) \quad \begin{aligned} (A \varphi_1, \varphi_1) a_1 + (A \varphi_1, \varphi_2) a_2 + \dots + (A \varphi_1, \varphi_n) a_n &= (f, \varphi_1), \\ (A \varphi_2, \varphi_1) a_1 + (A \varphi_2, \varphi_2) a_2 + \dots + (A \varphi_2, \varphi_n) a_n &= (f, \varphi_2), \\ \dots &\dots \\ (A \varphi_n, \varphi_1) a_1 + (A \varphi_n, \varphi_2) a_2 + \dots + (A \varphi_n, \varphi_n) a_n &= (f, \varphi_n). \end{aligned}$$

Tato soustava je vždy jednoznačně řešitelná, protože systém $\{\varphi_n\}$ je lineárně nezávislý. Řešením soustavy (5.62) najdeme prvek minimalizující posloupnosti ve tvaru (5.60).

Vraťme se nyní k rovnici (5.42) s okrajovými podmínkami (5.43) např. 3. typu, které však bude homogenní, tj. $\alpha = \beta = 0$ a alespoň jedno z čísel α_2, β_2 je nějake. Takto definovaná úloha má pozitivně definitní operátor, tedy můžeme hledat approximaci řešení Ritzovou metodou (důkaz např. v [4]).

Příklad 5.9 Řešme rovnici $-y'' - (1 + x^2)y = 1$ na intervalu $(-1, 1)$ s okrajovými podmínkami $y(-1) = y(1) = 0$.

Řešení: Z předcházejících úvah plyne, že úloha má pozitivně definitní operátor $Ay = -y'' - (1 + x^2)y$. $D_A = \{y \in L^2(-1, 1) : y, y', y'' \text{ jsou spojité a } y(-1) = y(1) = 0\}$. Systém souřadnicových funkcí musí být z D_A , tedy musí splňovat okrajové podmínky, proto volíme

$$\varphi_i = (1 - x^i)(1 + x^i) = 1 - x^{2i}, \quad i = 1, 2, \dots, n.$$

Pro výpočet koeficientů soustavy (5.57) počítejme

$$\begin{aligned} (A \varphi_i, \varphi_j) &= \int_{-1}^1 (-\varphi_i' - (1 + x^2)\varphi_i) \varphi_j \, dx = \int_{-1}^1 \varphi_i' \varphi_j' - (1 + x^2) \varphi_i \varphi_j \, dx, \\ (f, \varphi_i) &= \int_{-1}^1 f \varphi_i \, dx, \end{aligned}$$

tedy pro $i = 1, 2$ dostaneme

$$\begin{aligned} (A \varphi_1, \varphi_1) &= 152/105, \quad (A \varphi_1, \varphi_2) = (A \varphi_2, \varphi_1) = 16/9, \\ (A \varphi_2, \varphi_2) &= 42488/3645, \quad (f, \varphi_1) = 4/3, \quad (f, \varphi_2) = 8/5, \end{aligned}$$

a řešením soustavy

$$\begin{aligned} 152/105 \cdot a_1 + 16/9 \cdot a_2 &= 4/3 \\ 16/9 \cdot a_1 + 42488/3645 \cdot a_2 &= 8/5 \end{aligned}$$

získáme přibližné řešení ve tvaru

$$y_2 = a_1 \varphi_1 + a_2 \varphi_2 = 0,934 - 0,988x^2 + 0,054x^4 .$$

Další variační metoda, jejíž princip vyložíme, se nazývá Galerkinova metoda. Mějme řešit diferenciální rovnici $Au = f$, viz (5.49), s homogenními okrajovými podmínkami (5.50), kde A je lineární diferenciální operátor (5.51), definovaný na husté podmnožině D_A Hilbertova prostoru $L^2(a, b)$.

Obdobně jako v Ritzově metodě zvolíme v D_A posloupnost souřadnicových funkcí $\{\varphi_n\}$, která splňuje požadavky (5.58) a (5.59). Požadavek úplnosti posloupnosti $\{\varphi_n\}$, viz (5.59), je ekvivalentní s tímto požadavkem: Je-li pro nějakou funkci u splněno

$$(5.63) \quad (Au, \varphi_n) = 0 \quad \text{pro všechna } n ,$$

pak je u nutně nulová funkce.

V Galerkinově metodě hledáme přibližné řešení rovnice $Au = f$ ve tvaru

$$(5.64) \quad u_n = \sum_{i=1}^n a_i \varphi_i$$

a koeficienty (vzhledem k (5.63)) určíme z podmínky, aby rozdíl $Au_n - f$ byl ortogonální k souřadnicovým funkcím $\varphi_1, \varphi_2, \dots, \varphi_n$, tedy aby platilo

$$(5.65) \quad (Au_n - f, \varphi_i) = \int_a^b (Au_n - f) \varphi_i dx = 0 , \quad i = 1, 2, \dots, n .$$

Z (5.65) dostaneme soustavu lineárních rovnic

$$(5.66) \quad \begin{aligned} (A \varphi_1, \varphi_1) a_1 + (A \varphi_1, \varphi_2) a_2 + \dots + (A \varphi_1, \varphi_n) a_n &= (f, \varphi_1) , \\ (A \varphi_2, \varphi_1) a_1 + (A \varphi_2, \varphi_2) a_2 + \dots + (A \varphi_2, \varphi_n) a_n &= (f, \varphi_2) , \\ \dots &\dots \\ (A \varphi_n, \varphi_1) a_1 + (A \varphi_n, \varphi_2) a_2 + \dots + (A \varphi_n, \varphi_n) a_n &= (f, \varphi_n) , \end{aligned}$$

která je formálně shodná se soustavou (5.62). Z toho plyne, že metoda Galerkinova je shodná s Ritzovou metodou, když A je pozitivně definitní operátor.

Vzhledem k lineární nezávislosti souřadnicových funkcí je soustava (5.66) vždy jednoznačně řešitelná a přibližné řešení u_n rovnice $Au = f$ konverguje k přesnému řešení u v tom smyslu, že

$$(5.67) \quad \lim_{n \rightarrow \infty} (u_n - u, u_n - u) = \lim_{n \rightarrow \infty} \|u_n - u\|^2 = 0 ,$$

jestliže je rovnice $Au = f$ jednoznačně řešitelná.

Galerkinova metoda má mnohem obecnější použití, protože s ní můžeme řešit i rovnice $Au = f$ s nelineárním operátorem. Pak ovšem soustava (5.65) představuje nelineární soustavu, kterou nemůžeme převést na tvar (5.66), ale musíme ji řešit nějakými jinými postupy, viz 4. kapitola.

5.5 Počáteční úloha pro soustavy rovnic
1. řádu

Nechť je dána soustava lineárních diferenciálních rovnic prvního řádu, rozřešených vzhledem k první derivaci hledaných funkcí

$$(5.68) \quad \begin{aligned} y'_1(x) &= f_1(x, y_1(x), y_2(x), \dots, y_m(x)), \\ y'_2(x) &= f_2(x, y_1(x), y_2(x), \dots, y_m(x)), \quad x \in \langle a, b \rangle, \\ &\dots \dots \dots \dots \dots \dots \dots \dots \\ y'_m(x) &= f_m(x, y_1(x), y_2(x), \dots, y_m(x)), \end{aligned}$$

a počáteční podmínky

$$(5.69) \quad y_1(a) = y_1^{(0)}, \quad y_2(a) = y_2^{(0)}, \quad \dots, \quad y_m(a) = y_m^{(0)}.$$

Předpoklady, které zajišťují existenci a jednoznačnost spojitého řešení na intervalu $\langle a, b \rangle$, jsou uvedeny v následující větě.

Věta 5.2 Předpokládejme, že funkce f_1, f_2, \dots, f_m jsou:

a) definované a spojité jako funkce $m+1$ proměnných na oblasti

$$G = \{(x, y_1, y_2, \dots, y_m) : x \in \langle a, b \rangle, y_t \in (-\infty, \infty)$$

$$\text{pro } t = 1, 2, \dots, m\},$$

b) splňují vzhledem k y_1, y_2, \dots, y_m Lipschitzovu podmíinku, tzn. že existuje konstanta K tak, že

$$(5.70) \quad \begin{aligned} &|f_t(x, y_1, y_2, \dots, y_m) - f_t(x, \tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_m)| \leq \\ &\leq K \{ |y_1 - \tilde{y}_1| + |y_2 - \tilde{y}_2| + \dots + |y_m - \tilde{y}_m| \}, \end{aligned}$$

pro $t = 1, 2, \dots, m$,

platí pro libovolnou dvojici

$$(x, y_1, y_2, \dots, y_m), (x, \tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_m) \in G.$$

Potom existuje jediné řešení soustavy (5.68) $y_1(x), y_2(x), \dots, y_m(x)$, které je v $\langle a, b \rangle$ spojité se svojí první derivací a které splňuje i počáteční podmínky (5.69).

Důkaz najdete např. v [13].

Při práci s těmito soustavami se často přechází na vektorové značení. Pak zadání soustava se zapíše ve tvaru

$$(5.71) \quad \underline{y}'(x) = \underline{f}(x, \underline{y}(x)), \quad x \in \langle a, b \rangle$$

a počáteční podmínka (5.69) jako

$$(5.72) \quad \underline{y}(a) = \underline{y}^{(0)},$$

$$\text{кде } \underline{y}(x) = (y_1(x), y_2(x), \dots, y_m(x)), \quad \underline{f}(x, \underline{y}) = (f_1(x, \underline{y}), f_2(x, \underline{y}), \dots, f_m(x, \underline{y}))^T, \quad \underline{y}(a) = (y_1(a), y_2(a), \dots, y_m(a)) \quad \text{а } \underline{y}^{(0)} = (y_1^{(0)}, y_2^{(0)}, \dots, y_m^{(0)}).$$

Numerické metody řešení úlohy (5.71), (5.72) budou pouze rozšířením metod, které jsme uvedli v předcházejících odstavcích pro řešení rovnice (5.1). Nebudeme se tedy v následujícím příliš zdržovat podrobným odvozováním, ale většinou uvedeme pouze výsledky.

V numerických metodách zavedeme ještě vektor

$$(5.73) \quad \underline{y}^{(i)} = (y_1^{(i)}, y_2^{(i)}, \dots, y_m^{(i)}),$$

kde $y_1^{(i)}, y_2^{(i)}, \dots, y_m^{(i)}$ jsou přibližné hodnoty hledaných funkcí v bodě x_i , tedy $y_1^{(i)} \approx y_1(x_i)$, $y_2^{(i)} \approx y_2(x_i)$, ..., $y_m^{(i)} \approx y_m(x_i)$. Bod $x_i = a + i \cdot h$, $i = 0, 1, 2, \dots, n$, kde h je krok rovnoramenného dělení intervalu $[a, b]$.

5.5.1 Eulerova metoda

Vzorce Eulerovy metody pro výpočet přibližných hodnot řešení soustavy (5.71) za podmínek (5.72) dostaneme ze vzorce (5.5) tím, že funkci $f(x, y)$ nahradíme vektorovou funkcí $\underline{f}(x, \underline{y})$ a y_{i+1} , resp. y_i , vektory $\underline{y}^{(i+1)}$, resp. $\underline{y}^{(i)}$ podle (5.73), tedy

$$(5.74) \quad \underline{y}^{(i+1)} = \underline{y}^{(i)} + h \cdot \underline{f}(x_i, \underline{y}^{(i)}), \quad i = 0, 1, 2, \dots, n-1.$$

Pro výpočet rozepíšeme symbolickou rovnici (5.74) podrobněji

Příklad 5.10 Eulerovou metodou vypočítejme přibližné hodnoty řešení soustavy

$$(5.76) \quad \begin{aligned} y'(x) &= y(x) + z(x), & x \in \langle 0, 1/2 \rangle, \\ z'(x) &= y(x) - z(x), \end{aligned}$$

s počáteční podmínkou

$$(5.77) \quad y(0) = 1, \quad z(0) = 1,$$

v bodech $x_i = i.h$, $i = 0, 1, 2, 3, 4, 5$, $h = 0, 1.$

Řešení: Vzhledem k rovnicím (5.75) hledáme přibližné hodnoty řešení podle vzorců

$$y^{(i+1)} = y^{(i)} + h \cdot (y^{(i)} + z^{(i)}) ,$$

$$z^{(i+1)} = z^{(i)} + h \cdot (y^{(i)} - z^{(i)}) ,$$

kde $i = 0, 1, 2, 3, 4$. Numerické hodnoty řešení jsou sestaveny v tabulce 5.8.

i	x_i	$y^{(i)}$	$z^{(i)}$	Přesné řešení	
				$y(x_i)$	$z(x_i)$
0	0,0	1,000	1,000	1,000	1,000
1	0,1	1,200	1,000	1,211	1,010
2	0,2	1,420	1,020	1,446	1,040
3	0,3	1,664	1,060	1,710	1,091
4	0,4	1,936	1,120	2,008	1,164
5	0,5	2,242	1,202	2,346	1,261

Tab. 5.8

V pátém a šestém sloupci tabulky 5.8 jsou přesná řešení, která snadno vypočítáme, protože (5.76) je lineární homogenní soustava s konstantními koeficienty. Tedy řešení hledáme ve tvaru

$$(5.78) \quad y(x) = c_1 e^{\lambda x}, \quad z(x) = c_2 e^{\lambda x},$$

kde c_1, c_2 a λ vypočítáme dosazením (5.78) do (5.76) a užitím počátečních podmínek (5.77). Hledané řešení má tvar

$$\begin{aligned} y(x) &= \frac{1}{2} \left[(1 + \sqrt{2}) e^{x\sqrt{2}} + (1 - \sqrt{2}) e^{-x\sqrt{2}} \right], \\ z(x) &= \frac{1}{2} \left[1 \cdot e^{x\sqrt{2}} + 1 \cdot e^{-x\sqrt{2}} \right]. \end{aligned}$$

5.5.2 Metody Taylorova typu

Předpokládáme, že funkce f_1, f_2, \dots, f_m z (5.68), jako funkce proměnné x , mají v (a, b) spojité derivace až do p -tého řádu včetně a v (a, b) spojitou derivaci $(p+1)$ -ho řádu. Potom můžeme každou z funkcí y_1 až y_n vyjádřit v bodě x_{i+1} podle Taylorovy věty vzorcem

$$(5.79) \quad y_t(x_{i+1}) = y_t(x_i) + h \sum_{s=1}^p \frac{h^{s-1}}{s!} y_t^{(s)}(x_i) + \frac{h^{p+1}}{(p+1)!} y_t^{(p+1)}(\xi_t),$$

pro $t = 1, 2, \dots, m$, kde $\xi_t \in (a, b)$.

Zanedbáme-li zbytek ve vzorci (5.79) a nahradíme-li přesné hodnoty přibližnými, pak dostaneme vzorce metod Taylorova typu

$$(5.80) \quad y_t^{(i+1)} = y_t^{(i)} + h \sum_{s=1}^p \frac{h^{s-1}}{s!} y_t^{(s)}(x_i) =$$

$$= y_t^{(i)} + h \sum_{s=1}^p \frac{h^{s-1}}{s!} f_t^{(s-1)}(x_i, y_1^{(i)}, y_2^{(i)}, \dots, y_m^{(i)})$$

pro t = 1, 2, ..., m a i = 0, 1, 2, ..., n - 1.

Vzorce (5.80) obsahují derivace funkcí f_1, f_2, \dots, f_m a proto jsou pro praktické užívání velice těžkopádné, jak jsme o tom hovořili již v odstavci 5.2.2.

5.5.3 Metody Rungovy-Kuttovy

V metodách Rungeho-Kutty se rozdíly $y_t(x_{i+1}) - y_t(x_i)$ approximují lineárními kombinacemi

$$(5.81) \quad h [p_{rlt} \cdot k_{1t}(h) + p_{r2t} \cdot k_{2t}(h) + \dots + p_{rrt} \cdot k_{rt}(h)], \quad t = 1, 2, \dots, m,$$

kde $r \geq 1$ je zvolené pevně, p_{r11} až p_{rrm} jsou reálné koeficienty a k_{11} až k_{rm} jsou funkce proměnné h , srov. s (5.11). Potom, analogicky k (5.13), mají vzorce pro výpočet přibližných hodnot řešení soustavy (5.71) metodou Rungovou-Kuttovou tvar

$$(5.82) \quad y_t^{(i+1)} = h \sum_{s=1}^r p_{rst} \cdot k_{st}(h), \quad t = 1, 2, \dots, m.$$

kde

$$k_{1t}(h) = f_t(x_i, y_1^{(i)}, y_2^{(i)}, \dots, y_m^{(i)}),$$

$$k_{2t}(h) = f_t(x_i + \alpha_{2t} \cdot h, y_1^{(i)} + h \cdot \beta_{211} \cdot k_{11}, y_2^{(i)} + h \cdot \beta_{212} \cdot k_{12}, \\ \dots, y_m^{(i)} + h \cdot \beta_{21m} \cdot k_{1m}) .$$

$$k_{rt}(h) = f_t \left(x_i + \alpha_{rt} \cdot h, y_1^{(i)} + h \sum_{s=1}^{r-1} \beta_{rs1} \cdot k_{s1}, \right.$$

$$\left. y_2^{(i)} + h \sum_{s=1}^{r-1} \beta_{rs2} \cdot k_{s2}, \dots, y_m^{(i)} + h \sum_{s=1}^{r-1} \beta_{rsm} \cdot k_{sm} \right),$$

pro $t = 1, 2, \dots, m$.

Koeficienty p, α, β se pro pevné r určují podle stejných zásad jako v odstavci 5.2.3. Požadujeme tedy, aby se rozvoj sumy v (5.82) podle h shodoval se sumou v (5.80) do co nejvyšší možné mocniny proměnné h . Tyto výpočty nebudeme provádět, ale pouze uvedeme výsledek pro $r = 4$, kdy dostaneme vzorce tzv. klasické Rungovy-Kuttovy metody, a to

$$(5.83) \quad y_t^{(i+1)} = y_t^{(i)} + \frac{h}{6} [k_{1t} + 2k_{2t} + 2k_{3t} + k_{4t}], \quad t = 1, 2, \dots, m.$$

kde

$$k_{1t} = f_t(x_i, y_1^{(i)}, y_2^{(i)}, \dots, y_m^{(i)}),$$

$$k_{2t} = f_t(x_i + h/2, y_1^{(i)} + h \cdot k_{11}/2, y_2^{(i)} + h \cdot k_{12}/2, \dots, y_m^{(i)} + h \cdot k_{1m}/2),$$

$$k_{3t} = f_t(x_i + h/2, y_1^{(i)} + h \cdot k_{21}/2, y_2^{(i)} + h \cdot k_{22}/2, \dots,$$

$$y_m^{(i)} + h \cdot k_{2m}/2),$$

$$k_{4t} = f_t(x_i + h, y_1^{(i)} + h \cdot k_{31}, y_2^{(i)} + h \cdot k_{32}, \dots, y_m^{(i)} + h \cdot k_{3m}).$$

Příklad 5.11 Úlohu z příkladu 5.10 řešme pomocí Rungovy-Kuttovy metody 4. řádu, tj. podle vzorců (5.83).

Řešení: Numerické hodnoty jsou počítány podle vzorců

$$y^{(i+1)} = y^{(i)} + \frac{h}{6} [k_{11} + 2k_{21} + 2k_{31} + k_{41}],$$

$$z^{(i+1)} = z^{(i)} + \frac{h}{6} [k_{12} + 2k_{22} + 2k_{32} + k_{42}],$$

kde

$$k_{11} = y^{(i)} + z^{(i)}, \quad k_{12} = y^{(i)} + z^{(i)},$$

$$k_{21} = y^{(i)} + z^{(i)} + \frac{h}{2} (k_{11} + k_{12}), \quad k_{22} = y^{(i)} - z^{(i)} + \frac{h}{2} (k_{11} - k_{12}),$$

$$k_{31} = y^{(i)} + z^{(i)} + \frac{h}{2} (k_{21} + k_{22}), \quad k_{32} = y^{(i)} - z^{(i)} + \frac{h}{2} (k_{21} - k_{22}),$$

$$k_{41} = y^{(i)} + z^{(i)} + h(k_{31} + k_{32}), \quad k_{42} = y^{(i)} - z^{(i)} + h(k_{31} - k_{32})$$

i	x _i	y ⁽ⁱ⁾	z ⁽ⁱ⁾
0	0,0	1,000	1,000
1	0,1	1,211	1,010
2	0,2	1,443	1,039
3	0,3	1,704	1,088
4	0,4	1,998	1,159
5	0,5	2,332	1,253

Tab. 5.9

a sestaveny v tabulce 5.9. Srovnejte je s výsledky Eulerovy metody a s přesným řešením, viz tabulka 5.8.

5.5.4 Vícekrokové metody

Obecný vzorec k-krokové metody (5.34) můžeme rozšířit tak, aby platil i pro soustavu (5.68) s počátečními podmínkami (5.69). Stačí nahradit y_{i+j} , y'_{i+j} , a_j , b_j postupně hodnotami $y_t^{(i+j)}$, $y_t^{(i+j)}$, a_{jt} , b_{jt} pro $t = 1, 2, \dots, m$,

..., m. Potom k-kroková metoda pro soustavu lineárních diferenciálních rovnic 1. řádu je definován vzorcí

$$(5.84) \quad \sum_{j=0}^k a_{jt} y_t^{(i+j)} = h \cdot \sum_{j=0}^k b_{jt} y_t^{(i+j)}, \quad t = 1, 2, \dots, m,$$

kde a_{jt} , b_{jt} jsou reálné konstanty a $y_t^{(i+j)}$, $y_t^{(i+j)}$ jsou přibližné hodnoty řešení a jeho derivace v bodě x_{i+j} , tedy pro $t = 1, 2, \dots, m$ je

$$y_t^{(i+j)} \approx y_t(x_{i+j}), \quad y_t^{(i+j)} \approx y'_t(x_{i+j}) = f_t(x_{i+j}, y_1(x_{i+j}), \dots, y_n(x_{i+j})).$$

Jestliže funkce y_1, y_2, \dots, y_m mají p spojitých derivací na intervalu $\langle a, b \rangle$, pak můžeme hodnoty $y_t^{(i+j)}$, $y_t^{(i+j)}$ vyjádřit Taylorovými polynomy, tedy

$$(5.85) \quad y_t^{(i+j)} = \sum_{s=0}^p \frac{(jh)^s}{s!} y_t^{(s)}(x_i),$$

$$y_t^{(i+j)} = \sum_{s=1}^p \frac{(jh)^{s-1}}{(s-1)!} y_t^{(s)}(x_i).$$

Dosadíme-li (5.85) do (5.84) a porovnáme-li koeficienty u h^0, h^1, \dots, h^p , dostaneme rovnice

$$(5.86) \quad \begin{aligned} \sum_{j=0}^k a_{jt} &= 0, \\ \sum_{j=1}^k ja_{jt} - \sum_{j=0}^k b_{jt} &= 0, \\ \sum_{j=1}^k j^s a_{jt} - \sum_{j=1}^k s j^{s-1} b_{jt} &= 0, \quad s = 2, 3, \dots, p, \end{aligned}$$

pro $t = 1, 2, \dots, m$. Z rovnic (5.86) vypočítáme koeficienty a_{jt}, b_{jt} . Protože v soustavě je $2(k+1)$ neznámých a $p+1$ rovnic, pak p je maximálně rovno $2k+1$.

Zvolíme-li např. $k=1$ a $p=2$, dostaneme z (5.86) pro $t=1, 2, \dots, m$ rovnice

$$\begin{aligned} a_{0t} + a_{1t} &= 0 \\ a_{1t} - b_{0t} - b_{1t} &= 0 \\ a_{1t} - 2b_{1t} &= 0, \end{aligned}$$

ve kterých zvolíme $a_{1t} = 1$, pak $a_{0t} = -1$, $b_{1t} = 1/2$, $b_{0t} = 1/2$, tedy vzorec (5.84) má tvar

$$(5.87) \quad y_t^{(i+1)} = y_t^{(i)} + \frac{h}{2} \left[y_t^{(i)} + y_t^{(i+1)} \right], \quad t = 1, 2, \dots, m.$$

Srovnejte tento výsledek s druhým vzorcem v tabulce 5.5.

Příklad 5.12 Pomocí metody (5.87) řešme úlohu z příkladu 5.10.

Řešení: Vzorce (5.87) budou mít pro soustavu (5.76) tvar

$$(5.88) \quad \begin{aligned} y^{(i+1)} &= y^{(i)} + \frac{h}{2} \left[y^{(i)} + z^{(i)} + y^{(i+1)} + z^{(i+1)} \right], \\ z^{(i+1)} &= z^{(i)} + \frac{h}{2} \left[y^{(i)} - z^{(i)} + y^{(i+1)} - z^{(i+1)} \right]. \end{aligned}$$

Vzhledem k tomu, že (5.88) jsou implicitní vzorce a pravé strany (5.76) jsou lineární funkce, rozřešíme soustavu (5.88) vzhledem k $y^{(i+1)}, z^{(i+1)}$ a dostaneme vzorce

$$y^{(i+1)} = \frac{2}{2-h^2} \left[(1+h+\frac{h^2}{2})y^{(i)} + hz^{(i)} \right],$$

$$z^{(i+1)} = \frac{2}{2-h^2} \left[hy^{(i)} + (1-h+\frac{h^2}{2})z^{(i)} \right],$$

tedy

$$y^{(i+1)} = 1,00502 [1,105y^{(i)} + 0,1z^{(i)}],$$

$$z^{(i+1)} = 1,00502 [0,1y^{(i)} + 0,905z^{(i)}].$$

i	x_i	$y^{(i)}$	$z^{(i)}$
0	0,0	1,000	1,000
1	0,1	1,211	1,010
2	0,2	1,446	1,040
3	0,3	1,711	1,092
4	0,4	2,010	1,165
5	0,5	2,349	1,261

Numerické hodnoty tohoto výpočtu jsou v tabulce 5.10. Srovnajte je s přesným řešením v tabulce 5.8.

Tab. 5.10

5.6 Počáteční úloha pro rovnice m-tého řádu

Omezíme se na řešení diferenciální rovnice m-tého řádu rozřešené vzhledem k nejvyšší derivaci, tedy na řešení rovnice

$$(5.89) \quad y^{(m)}(x) = f(x, y(x), y'(x), \dots, y^{(m-1)}(x)), \quad x \in \langle a, b \rangle$$

a počátečními podmínkami

$$(5.90) \quad y(a) = y_0, \quad y'(a) = y'_0, \quad \dots, \quad y^{(m-1)}(a) = y_0^{(m-1)}.$$

Každou rovnici (5.89) můžeme převést na soustavu (5.68), zavedeme-li funkce $y_1(x), y_2(x), \dots, y_m(x)$, definované pomocí řešení $y(x)$ vztahy

$$(5.91) \quad y_1(x) = y(x), \quad y_2(x) = y'(x), \quad y_3(x) = y''(x), \quad \dots, \quad y_m(x) = y^{(m-1)}(x).$$

Z těchto rovnic a z rovnice (5.89) s podmínkami (5.90) dostaneme soustavu

$$(5.92) \quad \begin{aligned} y'_1(x) &= y_2(x), \\ y'_2(x) &= y_3(x), \\ &\dots \\ y'_{m-1}(x) &= y_m(x), \\ y'_m(x) &= f(x, y_1(x), y_2(x), \dots, y_m(x)), \end{aligned}$$

s počátečními podmínkami

$$(5.93) \quad y_1(a) = y_0, \quad y_2(a) = y'_0, \quad \dots, \quad y_m(a) = y_0^{(m-1)}.$$

Z tohoto převodu plyne, že pro řešení počáteční úlohy pro rovnice m -tého řádu můžeme užít všechny metody odvozené pro řešení počáteční úlohy pro soustavy rovnic prvního řádu v předcházejícím odstavci.

Diferenciální rovnice vyšších řádů můžeme však řešit i bez převodu na soustavu rovnic 1. řádu. Takový postup budeme demonstrovat na řešení rovnice

$$(5.94) \quad y''(x) = f(x, y, y'), \quad x \in \langle a, b \rangle$$

s okrajovými podmínkami

$$(5.95) \quad y(a) = y_0, \quad y'(a) = y'_0.$$

Integrujme rovnici (5.94) podle x v mezích od x_i do $x_i + th$, pak

$$(5.96) \quad y'(x_i + th) = y'(x_i) + \int_{x_i}^{x_i + th} f(x, y, y') dx$$

a integrací podle t v mezích od 0 do 1 dostaneme

$$(5.97) \quad y(x_{i+1}) = y(x_i) + hy'(x_i) + h \int_0^1 \int_{x_i}^{x_i + th} f(x, y, y') dx dt.$$

Předpokládejme, že známe hodnoty $y(x)$ a $y'(x)$ v bodech $x_i, x_{i-1}, \dots, x_{i-k}$, tedy lze spočítat $f_i, f_{i-1}, \dots, f_{i-k}$, kde $f_{i-j} = f(x_{i-j}, y(x_{i-j}), y'(x_{i-j}))$, pro $j = 0, 1, 2, \dots, k$. Sestrojme z těchto hodnot interpolační polynom $L_k(x)$ k funkci $f(x, y, y')$ a dosadme jej do (5.96) a (5.97), dále položme v (5.96) $t = 1$ a nahradíme přesné hodnoty $y(x_{i+1}), y(x_i), y'(x_{i+1}), y'(x_i)$ přibližnými $y_{i+1}, y_i, y'_{i+1}, y'_i$. Pak dostaneme vzorce

$$(5.98) \quad \begin{aligned} y'_{i+1} &= y'_i + \int_{x_i}^{x_{i+1}} L_k(x) dx, \\ y_{i+1} &= y_i + hy'_i + h \int_0^1 \int_{x_i}^{x_i + th} L_k(x) dx dt, \end{aligned}$$

které představují vzorce vícekrokové metody pro řešení rovnice (5.94).

V praxi se často setkáváme s rovnicemi, ve kterých pravá strana f nezávisí na y' , tedy s rovnicemi

$$(5.99) \quad y''(x) = f(x, y).$$

Potom bychom při užívání vzorců (5.98) zbytečně počítali y'_{i+1} . Ukážeme postup, jak se tomu můžeme vyhnout. Integraci rovnice (5.99) podle x v mezích od x_i do $x_i + th$ dostaneme analogii vzorce (5.96). Budeme-li tento analogický vzorec integrovat podle z , jednou v intervalu $\langle 0, 1 \rangle$ a podruhé v intervalu $\langle -1, 0 \rangle$, dostaneme vztahy

$$y(x_{i+1}) = y(x_i) + hy'(x_i) + h \int_0^{x_i+th} f(x, y) dx dt ,$$

$$y(x_i) = y(x_{i-1}) + hy'(x_i) + h \int_{-1}^0 \int_{x_i}^{x_i+th} f(x, y) dx dt$$

a z nich

$$(5.100) \quad y(x_{i+1}) = 2y(x_i) - y(x_{i-1}) + h \left[\int_0^1 \int_{x_i}^{x_i+th} f(x, y) dx dt - \int_{-1}^0 \int_{x_i}^{x_i+th} f(x, y) dx dt \right].$$

Dosadíme-li do (5.100) místo $f(x, y)$ interpolační polynom a místo přesných hodnot řešení hodnoty přibližné, dostaneme vzorec

$$(5.101) \quad y_{i+1} = 2y_i - y_{i-1} + h \left[\int_0^1 \int_{x_i}^{x_i+th} L_k(x) dx dt - \int_{-1}^0 \int_{x_i}^{x_i+th} L_k(x) dx dt \right],$$

který je vzorcem vícekrokové metody pro řešení rovnice (5.99) za podmínek (5.95).

Zvolíme-li za interpolační polynom např. Newtonův interpolační polynom (3.24) pro interpolaci vzad, pak

$$\int_{x_i}^{x_i+th} L_k(x) dx = h \int_0^t \left(f(x_i) + t \Delta^1 f_{i-1} + \dots + \frac{t(t+1)\dots(t+k-1)}{k!} \Delta^k f_{i-k} \right) dt ,$$

tedy vzorce (5.98) budou mít tvar

$$y'_{i+1} = y'_i + h \left[f_i + \frac{1}{2} \Delta^1 f_{i-1} + \frac{5}{12} \Delta^2 f_{i-2} + \frac{3}{8} \Delta^3 f_{i-3} + \dots \right] ,$$

$$(5.102) \quad y_{i+1} = y_i + hy'_i + h^2 \left[\frac{1}{2} f_i + \frac{1}{6} \Delta^1 f_{i-1} + \frac{1}{8} \Delta^2 f_{i-2} + \frac{19}{180} \Delta^3 f_{i-3} + \dots \right].$$

Zcela analogickým výpočtem příslušných integrálů v (5.101) dostaneme pro výpočet rovnice (5.99) vzorec

$$(5.103) \quad y_{i+1} = 2y_i - y_{i-1} + h^2 \left[f_i + \frac{1}{12} \Delta^2 f_{i-2} + \frac{1}{12} \Delta^3 f_{i-3} + \dots \right].$$

Zvolíme-li nakonec např. $k = 2$, dostaneme z (5.102), resp. (5.103) konečné extrapolační vzorce

$$y'_{i+1} = y'_i + h \left[\frac{23}{12} f_i - \frac{16}{12} f_{i-1} + \frac{5}{12} f_{i-2} \right] ,$$

$$(5.104) \quad y_{i+1} = y_i + hy'_i + h^2 \left[\frac{19}{24} f_i - \frac{5}{12} f_{i-1} + \frac{1}{12} f_{i-2} \right] ,$$

resp.

$$(5.105) \quad y_{i+1} = 2y_i - y_{i-1} + h^2 \left[\frac{13}{12} f_i - \frac{1}{12} f_{i-1} + \frac{1}{12} f_{i-2} \right].$$

Jistě bychom mohli dále postupovat v odvozování interpolacních vzorců, případně vzorců pro obecnější rovnici $y^{(m)}(x) = f(x, y)$ atd., ale tyto případy již rozebírat nebudeme, protože přesahují rozsah těchto skript.

Příklad 5.13 Řešme diferenciální rovnici $y'' = y$ v intervalu $\langle 0, 0,5 \rangle$ s počátečními podmínkami $y(0) = 1$, $y'(0) = 1$ při kroku $h = 0,1$. Přibližné řešení počítejme podle vzorce (5.105).

i	x_i	y_i	Přesné řešení $y(x_i)$
0	0,0	1,000	1,000
1	0,1	1,105	1,105
2	0,2	1,221	1,221
3	0,3	1,350	1,350
4	0,4	1,493	1,492
5	0,5	1,651	1,649

Řešení: Pro výpočet podle vzorce (5.105) potřebujeme startovací hodnoty y_0, y_1, y_2 , které vypočítáme pro jednoduchost z přesného řešení $y(x) = e^x$. Numerické hodnoty přibližného řešení jsou sestaveny v tabulce 5.11.

Tab. 5.11

5.7 Cvičení

1. Užitím jednokrokových metod z odstavce 5.2 vypočítejte přibližné hodnoty, řešení počátečních úloh:

- a) $y' = x \cdot y / 2$, $y(0) = 1$, $x \in \langle 0, 1 \rangle$, $h = 0,1$,
- b) $y' = y - 2x/y$, $y(0) = 1$, $x \in \langle 0, 1 \rangle$, $h = 0,1$,
- c) $y' = x - y$, $y(0) = 1$, $x \in \langle 0, 1 \rangle$, $h = 0,1$,
- d) $y' = (x + y)^2$, $y(0) = 0$, $x \in \langle 0, 1 \rangle$, $h = 0,1$,
- e) $y' = \frac{y^2}{x} - x$, $y(1) = 0$, $x \in \langle 1, 2 \rangle$, $h = 0,1$,
- f) $y' = -y^2 + \frac{y}{x}$, $y(1) = 1$, $x \in \langle 1, 2 \rangle$, $h = 0,1$.

Výsledky úloh porovnejte s analytickým řešením, případně odhadněte chybu metodou polovičního kroku.

2. Užitím vícekrokových vzorců z tabulky 5.5 řešte úlohy z předcházejícího příkladu a porovnejte přibližné hodnoty řešení s řešením přesným. Odhadněte lokální diskretizační chybu i globální chybu metodou polovičního kroku.

3. Úlohy z prvního příkladu řešte algoritmem prediktor - korektor.

4. Síťovou metodou řešte okrajové úlohy:

- a) $y'' + (1 + x^2)y = -1$, $x \in \langle -1, 1 \rangle$,
 $y(-1) = y(1) = 0$, $h = 0,5$,
- b) $y'' - 6y/x^2 = 0$, $x \in \langle 1, 2 \rangle$, $y(1) = 0$,
 $y(2) = 1$, $h = 0,2$,

c) $y'' + 4xy' + (4x^2 + 2)y = 0, \quad x \in \langle 0, 1 \rangle,$
 $y(0) = 0, \quad y(1) = 1, \quad h = 0,2.$

5. Ritzovou metodou řešte rovnici $y'' + y - x = 0$ na intervalu $\langle 0, 2 \rangle$ za okrajových podmínek $y(0) = y(2) = 0$.

Návod: Za souřadnicové funkce zvolte $\varphi_i = (2 - x)x^i$, $i = 1, 2, \dots, n$ a n volte pouze 3 nebo 4.

6. Řešte počáteční úlohy

a) $8y'' + 2y' - 3y = 0, \quad x \in \langle 0, 1 \rangle, \quad y(0) = -6, \quad y'(0) = 7,$

b) $y'' + 4y' + 4y = 0, \quad x \in \langle 0, 5, 1 \rangle, \quad y(0,5) = 0,5, \quad y'(0,5) = -4,$

převodem na soustavu rovnic 1. řádu nebo speciálními metodami, viz odstavec 5.6.

Výsledky

1. Přesná řešení jsou

a) $y = e^{(x^2/4)},$

b) $y = \sqrt{2x - 1},$

c) $y = 2e^{-x} - (1 - x),$

d) $y = -x + \operatorname{tg} x,$

e) $y = (1 - x^2)/(1 + x^2),$

f) $y = 2x/(x^2 + 1).$

4. a) $y(-1) = 0, \quad y(-0,5) = 0,721, \quad y(0) = 0,967, \quad y(0,5) = 0,721, \quad y(1) = 0,$

b) $y = c_1 x^3 + c_2/x^2, \quad c_1 = 4/31, \quad c_2 = -4/31,$

c) $y = (c_1 + c_2 x) e^{-x^2}, \quad c_1 = 0, \quad c_2 = e.$

5. $y = -2 \sinh x / \cosh 2 + x.$

6. a) $y = 2e^{(x/2)} - 8e^{(3x/4)},$

b) $y = (2e - 3ex) e^{-2x}.$

K A P I T O L A 6.

M E T O D A S Í T Í P R O P A R C I Á L N Í D I F E R E N C I Á L N Í R O V N I C E

V celé této kapitole se budeme zabývat metodou sítí pro řešení počátečních a okrajových problémů parciálních diferenciálních rovnic 2. řádu se dvěma proměnnými, i když lze tuto metodu užít i pro řešení jiných parciálních diferenciálních rovnic. Taková diferenciální rovnice má obecně tvar

$$(6.1) \quad F(x, y, u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}) = 0,$$

kde x, y jsou nezávisle proměnné, u je hledaná funkce, u_x je parciální derivace hledané funkce podle x atd.

Řešením rovnice (6.1) nazveme každou funkci $u(x, y)$, která má spojité parciální derivace 2. řádu a po jejímž dosazení do (6.1) přejde tato rovnice v identitu (samořejmě v jisté rovinné oblasti G).

Rovnice (6.1) se nazývá lineární, dá-li se zapsat ve tvaru

$$(6.2) \quad Au_{xx} + 2Bu_{xy} + Cu_{yy} + Du_x + Eu_y + Fu = H(x, y),$$

přičemž A až H jsou spojité funkce dvou proměnných x, y v uvažované oblasti G . Jsou-li navíc koeficienty A až F konstantní, pak mluvíme o lineární parciální diferenciální rovnici 2. řádu s konstantními koeficienty.

Platí-li ve všech bodech oblasti G

$$AC - B^2 > 0, \text{ resp. } AC - B^2 = 0, \text{ resp. } AC - B^2 < 0,$$

nazývá se taková lineární rovnice eliptická, resp. parabolická, resp. hyperbolická v oblasti G .

Například rovnice

$$(6.3) \quad u_{xx} + u_{yy} = 0$$

je tzv. Laplaceova rovnice a je eliptická v celé rovině xy , protože $A = C = 1$, $B = D = E = F = H = 0$ a $AC - B^2 = 1$ pro libovolné x, y .

Příkladem parabolické rovnice je rovnice vedení tepla v tenkém homogenním prutu

$$(6.4) \quad u_t - a^2 u_{xx} = H(x, t),$$

kde koeficient a závisí na fyzikálních vlastnostech prutu, $H(x, t)$ je funkce závislá na zdrojích tepla a $u(x, t)$ je hledaná teplota v okamžiku t v bodě x daného prutu. Protože $C = B = 0$, pak je tato rovnice parabolická.

Příkladem hyperbolické rovnice je tzv. jednorozměrná vlnová rovnice, čili rovnice kmitání homogenní struny

$$(6.5) \quad u_{tt} - a^2 u_{xx} = H(x, t),$$

kde koeficient a je opět závislý na fyzikálních vlastnostech struny, $H(x, t)$ závisí na síle, která kmitání působí a $u(x, t)$ je výchylka bodu x v okamžiku t .

Parciální diferenciální rovnici vyhovuje obecně nekonečně mnoho funkcí, proto úloha řešit parciální diferenciální rovnici není jednoznačná. Na řešení musíme proto klást ještě doplňující podmínky a teprve potom lze jednoznačnosti dosáhnout. Doplňující podmínky se zpravidla týkají hodnot hledané funkce nebo jejích derivací na hranici oblasti G , proto je nazýváme okrajové podmínky. Vyjadřuje-li některá z nezávisle proměnných čas, pak je "krajní" hodnotou této proměnné počátek děje a příslušná podmínka se nazývá počáteční podmínka. Proto se v matematice mluví o počátečních nebo okrajových problémech (či úlohách) při řešení parciálních diferenciálních rovnic.

Pro rovnici (6.3) se definuje tzv. Dirichletova (resp. Neumannova) okrajová úloha, při které hledáme takové řešení rovnice (6.3), které nabývá na hranici Γ oblasti G předepsaných funkčních hodnot (resp. hodnot derivací podle vnější normály hranice Γ). Hodnoty řešení (resp. derivací podle vnější normály) jsou zadány spojitou funkcí na Γ .

Okrajová podmínka může být na Γ formulována také jako vztah mezi hledanou funkcí u a její derivací ve směru vnější normály, tzv. Newtonova podmínka. Vyskytuje se i smíšené problémy, např. na jedné části hranice Γ je předepsána Dirichletova a na druhé Neumannova podmínka.

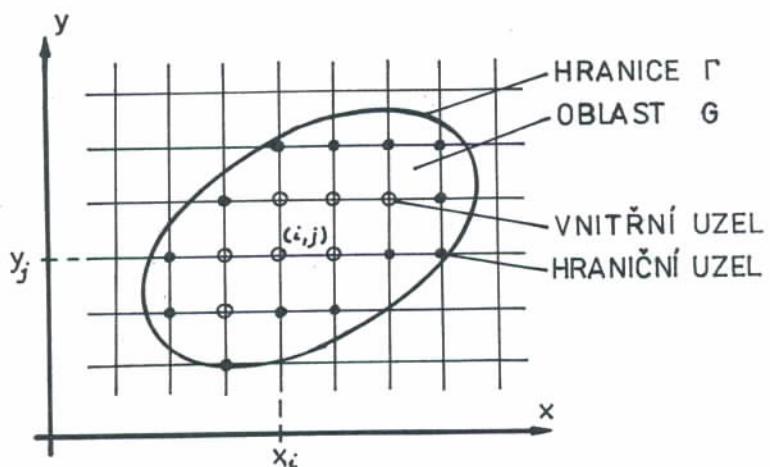
Pro rovnice (6.4), resp. (6.5), které obsahují čas, definujeme tzv. Cauchyovu úlohu (úlohu s počátečními podmínkami), tj. úlohu o nalezení takového řešení rovnice (6.4), resp. (6.5), které splňuje v počátečním čase jistou podmíanku.

Často se však pro rovnice (6.4), resp. (6.5) vyskytuje smíšené úlohy, které obsahují jak počáteční tak okrajové podmínky.

6.1 Základní princip metody sítí

Metoda sítí, nazývaná též metodou diferenční, je nejčastěji užívanou metodou při řešení parciálních diferenciálních rovnic. Je velmi oblíbená zřejmě proto, že její základní myšlenka je velice prostá. Spočívá v tom, že oblast G , ve které hledáme řešení dané diferenciální rovnice, pokryjeme konečnou sítí uzlových bodů a nahradíme derivace hledané funkce diferenčními. Tím dostaneme místo původní úlohy soustavu diferenčních rovnic pro hodnoty hledané funkce v uzlových bodech.

Mějme tedy řešit diferenciální rovnici (6.2) v omezené oblasti G roviny (x, y) , ohrazené po částech hladkou neprotínající se křivkou Γ . Pokryjeme tuto oblast prav-



Obr. 6.1

vouhlou (nejlépe čtvercovou) sítí složenou z přímek $x = x_0 + ih$, $y = y_0 + jk$, kde $i, j = 0, \pm 1, \pm 2, \dots$, (x_0, y_0) je pevně zvolený bod v rovině (x, y) a h, k udávají rozměr sítě ve směru x , resp. y . Průsečíky těchto přímek nazveme uzlové body, jednoduše uzly (viz obr. 6.1), které budeme označovat (i, j) a hodnotu hledané funkce v uzlu (i, j) stručně u_{ij} . Každý uzel $(i, j) \in \bar{G}$ má tedy čtveřici sousedních uzlů $(i-1, j)$, $(i+1, j)$, $(i, j-1)$, $(i, j+1)$. Leží-li všechny sousední uzly v \bar{G} , pak se uzel nazývá vnitřním, jinak hraničním. V každém vnitřním uzlu (i, j) nahradíme parciální derivace diferenčními vztahy:

$$(6.6) \quad \begin{aligned} \left(\frac{\partial u}{\partial x} \right)_{ij} &\approx \frac{u_{i+1,j} - u_{i,j}}{h}, \text{ nebo } \left(\frac{\partial u}{\partial x} \right)_{ij} \approx \frac{u_{i,j} - u_{i-1,j}}{h}, \\ \text{nebo } \left(\frac{\partial u}{\partial x} \right)_{ij} &\approx \frac{u_{i+1,j} - u_{i-1,j}}{2h}, \end{aligned}$$

obdobně

$$(6.7) \quad \begin{aligned} \left(\frac{\partial u}{\partial y} \right)_{ij} &\approx \frac{u_{i,j+1} - u_{i,j}}{k}, \text{ nebo } \left(\frac{\partial u}{\partial y} \right)_{ij} \approx \frac{u_{i,j} - u_{i,j-1}}{k}, \\ \text{nebo } \left(\frac{\partial u}{\partial y} \right)_{ij} &\approx \frac{u_{i,j+1} - u_{i,j-1}}{2k}, \end{aligned}$$

analogicky pro druhé derivace dostaneme

$$(6.8) \quad \begin{aligned} \left(\frac{\partial^2 u}{\partial x^2} \right)_{ij} &\approx \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2}, \\ \left(\frac{\partial^2 u}{\partial y^2} \right)_{ij} &\approx \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{k^2}, \text{ atd.} \end{aligned}$$

Jestliže ve vnitřních uzlech sítě označíme hodnoty funkcí A až H jako A_{ij} až H_{ij} , pak má rovnice (6.2) v každém vnitřním uzlu tvar lineární algebraické rovnice s neznámými u_{ij} , $u_{i+1,j}$, $u_{i-1,j}$, $u_{i,j+1}$, $u_{i,j-1}$.

Tím jsme řešení rovnice (6.2), až na splnění okrajových podmínek, přivedli na řešení soustavy lineárních algebraických rovnic. Přepis okrajových podmínek diferenčními rovnicemi si probereme až při řešení konkrétních okrajových problémů.

6.2 Metoda sítí pro eliptické rovnice

Mějme řešit Dirichletovu úlohu pro samoadjungovanou diferenciální rovnici druhého řádu, tzn. že máme řešit rovnici

$$(6.9) \quad L(u) \equiv \frac{\partial}{\partial x} (a(x, y) \frac{\partial u}{\partial x}) + \frac{\partial}{\partial y} (b(x, y) \frac{\partial u}{\partial y}) - q(x, y) u = f(x, y),$$

$$(x, y) \in G,$$

s okrajovou podmínkou

$$(6.10) \quad u(x, y) = \varphi(x, y) \quad \text{pro} \quad (x, y) \in \Gamma,$$

kde G je omezená rovinná oblast ohrazená po částech hladkou neprotínající se křivkou Γ . Dále předpokládáme, že $a(x, y)$, $b(x, y)$ jsou spojité, kladné a spojitě differencovatelné, $q(x, y)$ je spojitá a nezáporná a $f(x, y)$ je spojitá v G . Dále $\varphi(x, y)$ je spojitá na Γ .

Užitím diferenčních vztažů (6.6) a (6.7) vypočítáme přibližné hodnoty derivací $\partial u / \partial x$, $\partial u / \partial y$ v bodech $(x_0 + (i \pm 1/2)h, y_0 + jk)$, $(x_0 + ih, y_0 + (j \pm 1/2)k)$, které symbolicky označíme $(i \pm 1/2, j)$, $(i, j \pm 1/2)$, takto:

$$\frac{\partial u}{\partial x} \Big|_{(i-1/2,j)} \approx \frac{u_{i,j} - u_{i-1,j}}{h}, \quad \frac{\partial u}{\partial x} \Big|_{(i+1/2,j)} \approx \frac{u_{i+1,j} - u_{i,j}}{h},$$

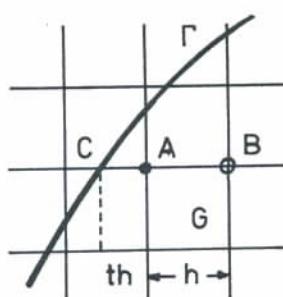
$$\left. \frac{\partial u}{\partial y} \right|_{(i,j-1/2)} \approx \frac{u_{i,j} - u_{i,j-1}}{k}, \quad \left. \frac{\partial u}{\partial y} \right|_{(i,j+1/2)} \approx \frac{u_{i,j+1} - u_{i,j}}{k}.$$

Dosadíme-li tyto vztahy do (6.9) a nahradíme-li analogicky i "vnější" derivace, pak bude mít rovnice (6.9) v každém vnitřním uzlu tvar

$$(6.11) \quad + \frac{1}{k} \left(b_{i,j+1/2} \frac{u_{i,j+1} - u_{i,j}}{k} - b_{i,j-1/2} \frac{u_{i,j} - u_{i,j-1}}{k} \right) - q_{ij} u_{ij} = f_{ij},$$

kde $a_{i+1/2,j}$, $b_{i,j+1/2}$, q_{ij} a f_{ij} jsou hodnoty funkcí a , b , q , f v bodech s příslušnou indexací.

Soustava rovnic (6.11) má menší počet rovnic než neznámých, proto je třeba k ní připojit další rovnice, které získáme z okrajové podmínky (6.10). Jestliže je oblast G čtverec nebo obdélník, pak lze síť uzlových bodů vytvořit tak, aby všechny hraniční body ležely na Γ . Potom hodnoty řešení v hraničních bodech jsou definovány hodnotami funkce $\psi(x, y)$ z podmínky (6.10). Jestliže je hranice Γ obecnější, pak je třeba podmítku (6.10) nějakým způsobem přenést do hraničních uzlů. Nejjednodušší postup je ten, že funkci $\psi(x, y)$ spojitě rozšíříme na celou oblast G a do hraničních uzlů klademe hodnoty tohoto spojitého prodloužení. Připojíme-li takto získané rovnice k rovnicím (6.11), dostaneme soustavu N lineárních rovnic o N neznámých hodnotách u_{ij} v bodech $(i, j) \in \bar{G}$, která je jednoznačně řešitelná (viz [4]). Takto získaná approximace bude sice pro $h, k \rightarrow 0$ konvergovat k přesnému řešení, ale rychlosť té konvergence bude malá. Proto se, podle Collatzze, častěji užívá přepisu okrajové podmínky pomocí lineární interpolace. Buď tedy A hraniční uzel, viz obr. 6.2, B jeho vnitřní soused a C sousedící přímky AB sítě s hranicí Γ .

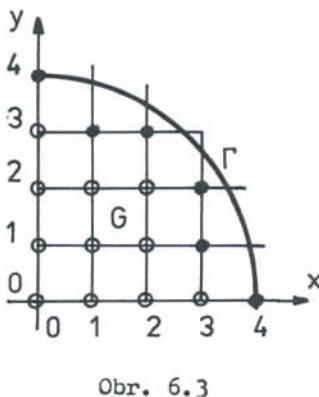


Obr. 6.2

Rovnice (6.11) a (6.12) vytvoří opět jednoznačně řešitelnou soustavu.

Příklad 6.1 Metodou sítí hledejme přibližné řešení rovnice $u_{xx} + u_{yy} = 0$ na oblasti $G = \{(x, y) : x^2 + y^2 < 16\}$ s hranicí $\Gamma = \{(x, y) : x^2 + y^2 = 16\}$, které splňuje podmítku $u(x, y) = x^2y^2$ pro $(x, y) \in \Gamma$.

Řešení: Vzhledem k symetričnosti problému se můžeme omezit jenom na čtvrtinu kruhové



Obr. 6.3

oblasti, viz obr. 6.3. Zvolme $h = k = 1$, pak dostaneme ve vnitřních uzlech, tj. v uzlech $(0, 1)$ až $(0, 3)$, $(1, 0)$ až $(1, 2)$, $(2, 0)$ až $(2, 2)$ a $(3, 0)$ podle (6.8) diferenční rovnice

$$(6.13) \quad u_{ij} = \frac{1}{4} (u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1}),$$

ve kterých vzhledem k symetrii pokládáme $u_{0,-j} = u_{0,j}$ a $u_{-i,0} = u_{i,0}$.

V hraničních uzlech dostaneme podle (6.12) a podle okrajové podmínky následující rovnice

$$(6.14) \quad \begin{aligned} u_{04} &= 0 \\ u_{13} &= \frac{\sqrt{15}-3}{\sqrt{15}-2} u_{12} + \frac{1}{\sqrt{15}-2} \cdot 15 \\ u_{23} &= \frac{\sqrt{12}-3}{\sqrt{12}-2} u_{22} + \frac{1}{\sqrt{12}-2} \cdot 48 = u_{32} \\ u_{31} &= \frac{\sqrt{15}-3}{\sqrt{15}-2} u_{21} + \frac{1}{\sqrt{15}-2} \cdot 15 \\ u_{04} &= 0 \end{aligned}$$

Spojením rovnic (6.13) a (6.14) dostaneme soustavu 17 rovnic o 17 neznámých u_{00} až u_{04} , u_{10} až u_{13} , u_{20} až u_{23} , u_{30} až u_{32} , u_{40} , ve které můžeme provést ještě mnohá zjednodušení, vzhledem k symetrii podle osy prvního kvadrantu.

Po výpočtu dostaneme

$$\begin{aligned} u_{00} &= u_{10} = u_{01} \doteq 18, \quad u_{11} \doteq 20, \quad u_{22} \doteq 32, \\ u_{02} &= u_{20} \doteq 15, \quad u_{12} = u_{21} \doteq 21, \quad u_{03} = u_{30} \doteq 13, \\ u_{13} &= u_{31} \doteq 18, \quad u_{40} = u_{04} = 0, \quad u_{23} = u_{32} \doteq 43. \end{aligned}$$

Srovnáme-li tyto výsledky s přesným řešením

$$u(x, y) = x^2 y^2 + \frac{1}{8} [256 - (x^2 + y^2)],$$

vidíme, že vykazují značné rozdíly. Přesnější výsledky docílíme zjemněním sítě nebo užitím tzv. Liebmanna procesu, viz např. [16].

6.3 Metoda sítí pro parabolické rovnice

Pro jednoduchost se omezíme na řešení smíšeného problému pro rovnici vedení tepla (6.4) s jednou prostorovou proměnnou. Postup při řešení obecné parabolické rovnice o n proměnných je zcela analogický.

Mějme na oblasti $G = \{(x, t) : 0 < x < L, 0 < t < T\}$ dánu diferenciální rovnici

$$(6.15) \quad \frac{\partial u}{\partial t} = \frac{\partial}{\partial x} \left(a(x, t) \frac{\partial u}{\partial x} \right) + H(x, t)$$

s počáteční podmínkou

$$(6.16) \quad u(x, 0) = g(x), \quad 0 < x < L$$

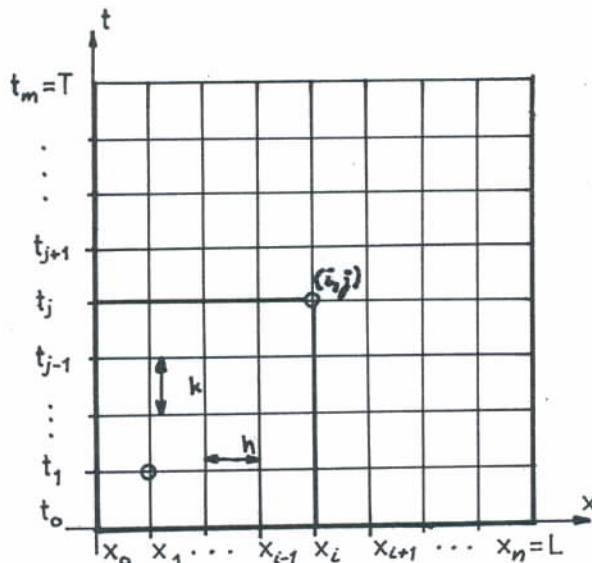
a okrajovou podmínkou

$$(6.17) \quad u(0, t) = f_1(t), \quad u(L, t) = f_2(t), \quad 0 \leq t \leq T.$$

Přitom funkce a , H jsou spojité a funkce $a(x, t)$ je v G spojitě diferencovatelná a platí

$$0 < a_0 \leq a(x, t) \leq a_1, \quad (x, t) \in G,$$

kde a_0, a_1 jsou konstanty. Funkce g, f_1, f_2 jsou spojité.



Obr. 6.4

Pro řešení metodou sítí pokryjeme G obdélníkovou sítí bodů $(i, j) = (x_i, t_j) = (ih, jk)$, kde $i = 0, 1, 2, \dots, n, j = 0, 1, 2, \dots, m$, tedy $h = L/n, k = T/m$, viz obr. 6.4.

V uzlových bodech (i, j) pro $i = 1, 2, \dots, n-1, j = 0, 1, 2, \dots, m-1$ sestavíme diferenční rovnice tak, že časovou derivaci nahradíme dopřednou differencí

$$\left. \frac{\partial u}{\partial t} \right|_{(i,j)} \approx \frac{u_{i,j+1} - u_{i,j}}{k}$$

a derivaci podle x nahradíme diferenčními centrálními. Jestliže $(i \pm 1/2, j)$ je uzlový bod $((i \pm 1/2)h, jk)$, potom

$$\left. \frac{\partial u}{\partial x} \right|_{(i+1/2,j)} \approx \frac{u_{i+1,j} - u_{i,j}}{h},$$

$$\left. \frac{\partial u}{\partial x} \right|_{(i-1/2,j)} \approx \frac{u_{i,j} - u_{i-1,j}}{h},$$

$$\left. \frac{\partial}{\partial x} \left(a(x, t) \frac{\partial u}{\partial x} \right) \right|_{(i,j)} \approx \frac{1}{h} \left[a_{i+1/2,j} \frac{u_{i+1,j} - u_{i,j}}{h} - a_{i-1/2,j} \frac{u_{i,j} - u_{i-1,j}}{h} \right].$$

Nakonec dostaneme podle (6.15) diferenční rovnice ve tvaru

$$(6.18) \quad u_{i,j+1} = u_{i,j} + \frac{k}{h^2} \left[a_{i-1/2,j} u_{i-1,j} - (a_{i-1/2,j} + a_{i+1/2,j}) u_{i,j} + a_{i+1/2,j} u_{i+1,j} \right] + H_{ij}$$

Z podmínek (6.16) a (6.17) dostaneme

$$(6.19) \quad u_{i,0} = g(ih), \quad i = 1, 2, \dots, n-1,$$

$$(6.20) \quad u_{0,j} = f_1(jk), \quad u_{n,j} = f_2(jk), \quad j = 0, 1, \dots, m.$$

Tím jsme převedli úlohu (6.15), (6.16), (6.17) na výpočet algebraických rovnic (6.18), (6.19), (6.20). Vzorec (6.18) dovoluje postupně explicitně vypočítávat přibližné hodnoty $u_{i,j+1}$ v následujícím časovém řádku z řádku předcházejícího. Proto hovoříme o tzv. explicitním schématu. Aby byla zaručena konvergencie a stabilita explicitního schématu musí být časový a prostorový krok svázán touto podmínkou

$$(6.21) \quad a_1 \frac{k}{h^2} \leq \frac{1}{2}.$$

Rovnice (6.18) nejsou však jedinou možností diskretizace zadání úlohy. Kdybychom např. časovou derivaci nahradili zpětnou diferencí

$$\frac{\partial u}{\partial t} \Big|_{(i,j)} \approx \frac{u_{ij} - u_{i,j-1}}{k},$$

pak budou mít diferenční rovnice tvar

$$(6.22) \quad \frac{u_{ij} - u_{i,j-1}}{k} - \frac{1}{h^2} \left[a_{i-1/2,j} u_{i-1,j} - (a_{i-1/2,j} + a_{i+1/2,j}) u_{ij} + a_{i+1/2,j} u_{i+1,j} \right] = H_{ij},$$

pro $i = 1, 2, \dots, n-1$, $j = 1, 2, \dots, m$ a okrajové podmínky budou přepsány opět do rovnic (6.19) a (6.20).

Soustava rovnic (6.19), (6.20), (6.22) tvoří tzv. implicitní schéma a musíme ji řešit některými metodami lineární algebry.

Explicitní schéma má sice praktické početní výhody ve své jednoduchosti, ale podmínka (6.21) někdy citelně omezuje velikost časového kroku. Naproti tomu implicitní schéma nemá tyto požadavky a z toho důvodu bývá výhodné užít implicitní schéma i za cenu řešení soustavy lineárních algebraických rovnic (6.19), (6.20), (6.22).

Příklad 6.2 Je dána rovnice vedení tepla v homogenní tyči

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

s podmínkami

$$u(x, 0) = \sin \pi x, \quad x \in \langle 0, 1 \rangle, \\ u(0, t) = u(1, t) = 0, \quad t \in \langle 0, 0.025 \rangle.$$

Sestavme pro tuto úlohu explicitní schéma metody sítí a vypočítejme přibližné hodnoty řešení.

Řešení: Zvolíme-li $h = 0.1$, potom podle (6.21) musí být $k \leq h^2/2 = 0.005$. Položme tedy $k = 0.005$. Vzorce (6.18), (6.19) a (6.20) budou mít tvar

$$u_{i,j+1} = \frac{1}{2} (u_{i-1,j} + u_{i+1,j}), \quad i = 1, 2, \dots, 9, \quad j = 0, 1, 2, 3, 4, \\ u_{i,0} = \sin(0, 1 \pi i), \quad i = 1, 2, \dots, 9, \\ u_{0,j} = u_{10,j} = 0, \quad j = 0, 1, 2, 3, 4, 5.$$

Pоловина hodnot vypočítaných podle těchto rovnic je v tabulce 6.1.

<i>t</i>	<i>x</i>	0	0,1	0,2	0,3	0,4	0,5
	<i>i</i>	0	1	2	3	4	5
	<i>j</i>	0					
0,000	0	0	0,309	0,588	0,809	0,951	1,000
0,005	1	0	0,394	0,559	0,770	0,904	0,951
0,010	2	0	0,379	0,532	0,732	0,860	0,904
0,015	3	0	0,266	0,506	0,696	0,818	0,860
0,020	4	0	0,253	0,481	0,662	0,778	0,818
0,025	5	0	0,240	0,457	0,629	0,740	0,778

Tab. 6.1

6.4 Metoda sítí pro hyperbolické rovnice

Pro demonstraci metody budeme vyšetřovat smíšenou úlohu pro rovnici kmitání struny, která je dána rovnicí

$$(6.23) \quad \frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2},$$

s počátečními podmínkami

$$(6.24) \quad u(x, 0) = g_1(x), \quad \left. \frac{\partial u}{\partial t} \right|_{t=0} = g_2(x), \quad 0 < x < L,$$

a okrajovými podmínkami

$$(6.25) \quad u(0, t) = f_1(t), \quad u(L, t) = f_2(t), \quad 0 \leq t \leq T.$$

Pro jednoduchost ještě položíme $a = 1$. Oblast $G = \{(x, t) : x \in (0, L), t \in (0, T)\}$ pokryjeme sítí jako v předcházejícím odstavci, viz obr. 6.4.

Ve vnitřních uzlech sítě přepíšeme diferenciální rovnici (6.23) užitím vzorce (6.8) na diferenční rovnici

$$(6.26) \quad \frac{u_{i,j+1} + 2u_{i,j} + u_{i,j-1}}{k^2} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2},$$

pro $i = 1, 2, \dots, n-1$, $j = L, 2, \dots, m-1$. Po úpravě pak dostaneme

$$(6.27) \quad u_{i,j+1} = 2u_{i,j} - u_{i,j-1} + \frac{k^2}{h^2} [u_{i+1,j} - 2u_{i,j} + u_{i-1,j}].$$

Abychom podle explicitních rovnic (6.27) mohli počítat přibližné hodnoty řešení v uzlových bodech, potřebujeme znát u_{ij} v prvních dvou časových řadách, tj. pro $j = 0, 1$ a $i = 0, 1, 2, \dots, n$. K tomu nám poslouží podmínky (6.24) a (6.25).

Ze (6.25) a z první rovnice v (6.24) plyne

$$(6.28) \quad \begin{aligned} u_{0,j} &= f_1(jk), & u_{n,j} &= f_2(jk), & j &= 0, 1, \dots, m, \\ u_{i,0} &= g_1(ik), & i &= 1, 2, \dots, n-1. \end{aligned}$$

Jestliže ve druhé rovnici v podmínce (6.24) nahradíme derivaci vztahem

$$\frac{u_{il} - u_{i0}}{k} \approx \left. \frac{\partial u}{\partial t} \right|_{\substack{x=ih \\ t=0}} = g_2(ih),$$

potom

$$(6.29) \quad u_{il} = u_{i0} + kg_2(ih) = g_1(ih) + kg_2(ih)$$

pro $i = 1, 2, \dots, n-1$.

Explicitní schéma metody sítí pro rovnici (6.23) a podmínky (6.24), (6.25) je definováno diferenčními rovnicemi (6.27), (6.28), (6.29). Toto schéma je konvergentní a stabilní, jestliže zvolíme k, h tak, aby platilo

$$(6.30) \quad k \leq \frac{h}{a}.$$

LITERATURA

- [1] A ch i e s e r , N.I.: Vorlesungen über Approximationstheorie. Berlin 1953.
- [2] A h l b e r g , J.H., N i l s o n , E.N., W a l s h , J.L.: The Theory of Spline and Their Applications. New York 1967 (ruský překlad Moskva 1972).
- [3] A i t k e n , A.C.: On Interpolation by Iteration of Proportional Parts, without the Use of Differences. Proc. Edinburg Math. Soc. 3, series 2, s. 56-76, (1932).
- [4] B a b u š k a , I., P r á g e r , M., V i t á s e k , E.: Numerické řešení diferenciálních rovnic. Praha, SNTL 1964.
- [5] B e r e z i n , I.S., Ž i d k o v , N.P.: Metody vyčislenij I, II. Moskva, Fizmatgiz 1966.
- [6] C o h e n , A.M.: Numerical Analysis. London, McGraw-Hill Book Company 1973.
- [7] C o l l a t z , L.: Funkcionální analýza a numerická matematika. Praha, SNTL 1970.
- [8] D a h l q u i s t , G., B j ö r c k , A.: Numerical methods. New York, Prentice-Hall 1974.
- [9] D ě m i d o v i č , B.P., M a r o n , I.A.: Základy numerické matematiky. Praha, SNTL 1966.
- [10] D ě m i d o v i č , B.P., M a r o n , I.A., Š u v a l o v a , E.Z.: Číslenyje metody analiza. Moskva, Nauka 1967.
- [11] F a d d ě j e v , D.K., F a d d ě j e v o v á , V.N.: Numerické metody lineární algebry. Praha, SNTL 1964.
- [12] G e o r g e , A., L i n , J.: Computer Solution of Large Sparse Positive Definite Systems. New York, Prentice-Hall 1981
- [13] H e n r i c i , P.: Discrete Variable Methods in Ordinary Differential Equations. New York, Wiley 1968.
- [14] H e n r i c i , P.: Elements of Numerical Analysis. New York, Wiley 1964.
- [15] J a r n í k , V.: Integrální počet II. Praha, Academia 1976.
- [16] J o k l , E., V o s p ě l , Z., Z a h r a d n í k , V., Z e l e n k a , M.: Základy numerické matematiky. Praha, ČVUT 1975.
- [17] M i c h l i n , S.G.: Variacionnyje metody v matematičeskoj fizike. Moskva, Gostechizdat 1957.
- [18] M í k a , S.: Numerické metody algebry. MVŠT. Praha, SNTL 1982.
- [19] N e k v i n d a , M., Š r u b a ř , J., V i l d , J.: Úvod do numerické matematiky. Praha, SNTL 1976.
- [20] N e v i l l e , E.H.: Iterative interpolation. J. Indian Math. Soc., series 20, s. 87-120 (1934).
- [21] P r á g e r , M.: Numerická matematika I. Praha, UK 1981.
- [22] R a l s t o n , A.: Rational Chebyshev Approximation by Remes' Algorithms. Numer. Math. 7, s. 322-330 (1965).

- [23] R a l s t o n , A.: Základy numerické matematiky. Praha, Academia 1973.
- [24] R e k t o r y s , K.: Variační metody. Praha, SNTL 1974.
- [25] S t o e r , J., B u l i r s c h , R.: Introduction to Numerical Analysis. New York - Heidelberg - Berlin, Springer-Verlag 1980.
- [26] V i t á s e k , E.: Numerická matematika II. Praha, UK 1981.
- [27] Z a v j a l o v , J.S., K v a s o v , B.I., M i r o š n i č e n k o , V.L.: Metody splajn-funkcij. Moskva, Nauka 1980.

O B S A H

Kap. 1. <u>Úvod</u>	3
1.1 Co je numerická matematika	3
1.2 Chyby a jejich charakteristiky	3
1.3 Šíření chyb ve výpočtech	4
1.4 Numerická stabilita	6
Kap. 2. <u>Úlohy lineární algebry</u>	7
2.1 Některé pojmy	7
2.1.1 Ortogonalita vektorů a matic	7
2.1.2 Matice rozdělená na bloky	7
2.1.3 Normy matic a vektorů	8
2.1.4 Speciální matice	9
2.2 Základní úlohy lineární algebry	10
2.3 Přímé metody řešení soustav lineárních algebraických rovnic	10
2.3.1 Řešení trojúhelníkových soustav	11
2.3.2 Gaussova a Jordanova eliminační metoda	11
2.3.3 Výběr hlavních prvků	15
2.3.4 Metoda LU-rozkladu	17
2.4 Iterační metody řešení soustav lineárních algebraických rovnic ..	19
2.4.1 Prostá iteriční metoda - Jacobiova metoda	19
2.4.2 Gaussova - Seidelova iteriční metoda	21
2.4.3 Relaxační metoda	22
2.5 Řešení maticové rovnice $AX = B$. Invertování matic	23
2.5.1 Řešení rovnice $AX = B$ eliminací	23
2.5.2 Invertování matic eliminací	25
2.5.3 Inverze LU-rozkladem	26
2.5.4 Inverze dělením na bloky - metoda vroubení	27
2.6 Špatně podmíněné soustavy rovnic	28
2.7 Výpočet vlastních čísel	28
2.7.1 Mocninná metoda	30
2.7.2 Urychlení konvergence mocninné metody	31
2.7.3 Mocninná metoda pro symetrickou matici	32
2.7.4 Výpočet dalších vlastních čísel	32
2.7.5 LU-rozklad pro úplný problém vlastních čísel	33
2.8 Cvičení	33
Kap. 3. <u>Aproximace funkce</u>	36
3.1 Úloha approximace	36
3.2 Interpolaciální approximace pomocí polynomů	37
3.2.1 Lagrangeův interpolaciální polynom	38
3.2.2 Nevillův algoritmus	39
3.2.3 Newtonův interpolaciální polynom	41
3.2.4 Hermitova interpolace	45
3.2.5 Přesnost a konvergence interpolaciální approximace	47
3.3 Interpolace spline-funkcemi	48

3.3.1	Konstrukce interpolační spline-funkce	49
3.3.2	Konvergance spline-interpolace	52
3.4	Aproximace metodou nejmenších čtverců (MNČ) - diskrétní případ	53
3.4.1	Normální rovnice	53
3.4.2	Volba základních funkcí	54
3.4.3	Čebyševovy polynomy	56
3.4.4	Trigonometrické polynomy	57
3.5	Aproximace metodou nejmenších čtverců (MNČ) - spojitý případ	58
3.5.1	Normální rovnice	58
3.5.2	Volba základních funkcí	59
3.6	Čebyševova aproximační	61
3.6.1	Konstrukce Čebyševovy aproximační	62
3.7	Numerické derivování	63
3.7.1	Derivování interpolačního polynomu	63
3.7.2	Derivování interpolační spline-funkce	65
3.8	Numerické integrování (kvadratura)	66
3.8.1	Newtonovy - Cotesovy kvadraturní vzorce	67
3.8.2	Lichoběžníková metoda	68
3.8.3	Simpsonova metoda	69
3.8.4	Gaussovy kvadraturní vzorce	69
3.8.5	Rombergova metoda	71
3.8.6	Chyby a konvergance metod	73
3.9	Cvičení	74
 Kap. 4.	<u>Řešení nelineárních rovnic a jejich soustav</u>	76
4.1	Separace kořenů nelineárních rovnic	76
4.2	Výpočet kořenů nelineární rovnice	79
4.2.1	Metoda půlení intervalu (bisekce)	79
4.2.2	Metoda prosté iterace	80
4.2.3	Newtonova metoda (metoda tečen)	82
4.2.4	Metoda regula falsi (tětiv)	84
4.2.5	Metoda sečen	85
4.2.6	Kombinovaná metoda	86
4.3	Výpočet kořenů algebraické rovnice	86
4.3.1	Metoda Laguerrova	86
4.3.2	Graeffova metoda	88
4.3.3	Metody užívající syntetické dělení	90
4.4	Řešení nelineárních soustav	92
4.4.1	Metoda prosté iterace	93
4.4.2	Newtonova (zobecněná) metoda	95
4.5	Cvičení	97
 Kap. 5.	<u>Numerické řešení obyčejných diferenciálních rovnic</u>	99
5.1	Počáteční úloha pro rovnice 1. řádu	99
5.2	Jednokrokové metody řešení počáteční úlohy	100
5.2.1	Eulerova metoda	100
5.2.2	Metody Taylorova typu	101
5.2.3	Metody Rungovy - Kuttovy	102
5.2.4	Odhad chyby a konvergence	104
5.3	Vícekrokové metody řešení počáteční úlohy	107

5.3.1	Adamsovy extrapoláční metody	107
5.3.2	Adamsovy interpolační metody	109
5.3.3	Konstrukce vícekrokových metod	110
5.3.4	Chyby vícekrokových metod	111
5.3.5	Algoritmus prediktor-korektor	112
5.4	Okrajové úlohy	114
5.4.1	Metoda sítí pro rovnice 2. řádu	115
5.4.2	Variační metody	117
5.5	Počáteční úloha pro soustavy rovnic 1. řádu	122
5.5.1	Eulerova metoda	123
5.5.2	Metody Taylorova typu	124
5.5.3	Metody Rungovy - Kuttovy	125
5.5.4	Vícekrokové metody	126
5.6	Počáteční úloha pro rovnice m-tého řádu	128
5.7	Cvičení	131
Kap. 6.	Metoda sítí pro parciální diferenciální rovnice	133
6.1	Základní princip metody sítí	134
6.2	Metoda sítí pro eliptické rovnice	135
6.3	Metoda sítí pro parabolické rovnice	137
6.4	Metoda sítí pro hyperbolické rovnice	140
Literatura	142