

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ
ÚSTAV MATEMATIKY

Moderní numerické metody

(Elektrotechnika, elektronika, komunikační a řídicí technika)

Jaromír Baštinec
Michal Novák



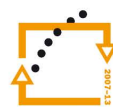
evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost



VYSOKÉ
UČENÍ
TECHNICKÉ
V BRNĚ

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Ústav matematiky FEKT VUT v Brně, 2014

<http://www.umat.feec.vutbr.cz>

Obrázky pomocí METAPOSTu a METAFONTu: Jaromír Kuben

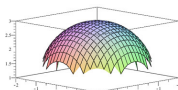
Animace a 3D objekty: Roman Plch



Součástí tohoto učebního textu jsou odkazy na tzv. maplety, tj. programy vytvořené v prostředí Maple. Tyto odkazy jsou v textu zvýrazněny barvou, příp. uvozeny slovem *maplet*. Maplety ke svému běhu nevyžadují software Maple – je však nutné mít na klientském počítači nainstalováno prostředí Java a nastavenou vhodnou úroveň zabezpečení prohlížeče i prostředí Java. Po kliknutí na odkaz mapletu se v závislosti na softwarovém prostředí klientského počítače zobrazí různá hlášení o zabezpečení – všechny dialogy je třeba povolit a spouštění požadovaných prvků neblokovat.



Součástí tohoto učebního textu jsou odkazy na tzv. webMathematica applety, tj. programy vytvořené v prostředí webMathematica. Tyto odkazy jsou v textu zvýrazněny barvou, příp. uvozeny slovy *matematický software*, *webMathematica applet* apod. Applety ke svému běhu nevyžadují software Mathematica – je však nutné mít na klientském počítači nainstalováno prostředí Java a nastavenou vhodnou úroveň zabezpečení prohlížeče i prostředí Java. Po kliknutí na odkaz appletu se v závislosti na softwarovém prostředí klientského počítače mohou zobrazit různá hlášení o zabezpečení – všechny dialogy je třeba povolit a spouštění požadovaných prvků neblokovat.



Součástí tohoto učebního textu jsou animace, resp. prvky 3D grafiky. Pro korektní zobrazení těchto multimediálních prvků a práci s nimi je nutné správně nastavit zabezpečení prohlížeče PDF souborů, a to zejména na záložkách typu *3D a multimédia*, *Důvěryhodnost multimedií (starší)*, *Multimédia (starší)*. Vlastnosti zobrazení těchto prvků lze ovlivnit pomocí položek jejich kontextového menu, příp. pomocí práce s myší nebo jiným polohovacím zařízením.



Doplňující součástí tohoto učebního textu jsou příklady zpracované v [elektronické bance příkladů](#).



Doplňující součástí tohoto učebního textu jsou příklady zpracované v [elektronické bance příkladů](#).

Obsah

Úvod	5
1 Princip numerických metod	6
1.1 Chyby při numerických výpočtech	6
1.2 Šíření chyb při aritmetických operacích	7
1.3 Podmíněnost úloh	9
1.4 Richardsonova extrapolace	11
1.5 Metriky, kontrakce, Banachova věta	13
2 Soustavy lineárních rovnic	20
2.1 Soustavy lineárních rovnic — Základní pojmy	20
2.1.1 Řešení soustav	22
2.2 Gaussova eliminační metoda	26
2.2.1 Úplný a částečný výběr hlavního prvku	28
2.2.2 Metoda LU-rozkladu	29
2.2.3 Řešení pomocí inverzní matice	32
2.3 Iterační metody řešení	32
2.3.1 Jacobiho iterační metoda	34
2.3.2 Gaussova – Seidelova iterační metoda	35
2.4 Relaxační metody	37
2.4.1 Superrelaxační metoda	39
2.4.2 Metoda největšího spádu	39
2.4.3 Metoda sdružených gradientů	39
2.5 Stabilita řešení numerické úlohy	41
2.6 Cvičení	43

3	Řešení rovnic.	44
3.1	Základní pojmy	44
3.2	Startovací metody – Separace kořenů	45
3.3	Iterační metody	46
3.3.1	Metoda prosté iterace	47
3.3.2	Metoda regula falsi	49
3.3.3	Metoda sečen	49
3.3.4	Metoda tečen	50
3.3.5	Modifikovaná Newtonova metoda	53
3.3.6	Newtonova metoda pro komplexní kořeny	54
3.3.7	Kombinovaná metoda sečen a tečen	55
3.4	Řád metody	56
3.5	Algebraické rovnice	58
3.5.1	Metoda Laguerrova	62
3.5.2	Metoda Graeffova – Lobačevského	62
3.5.3	Metoda Schurova	67
3.6	Cvičení	70
3.7	Animace	70
4	Vlastní čísla	74
4.1	Základní pojmy	74
4.2	Numerické metody pro hledání vlastních čísel	77
4.3	Klasické metody určení koeficientů charakteristického polynomu	78
4.3.1	Krylovova metoda	78
4.3.2	Faddějevova-Leverrierova metoda	79
4.4	Poloha a odhad vlastních čísel	80
4.4.1	Geršgorinovy věty	80
4.4.2	Metody výpočtu dominantního vlastního čísla	84
4.4.3	Metoda Rayleighova podílu	87
4.4.4	Výpočet dalších vlastních čísel mocninnou metodou	89
4.4.5	Výpočet vlastních čísel a vlastních vektorů symetrických matic	90
4.5	Householderova matice zrcadlení	99

4.5.1	Givensova-Householderova metoda	101
4.5.2	QR-rozklad	106
4.5.3	Srovnání algoritmů	115
4.5.4	QR-rozklad a vlastní čísla matice \mathbf{A} – QR -algoritmus	115
4.6	Podmíněnost problému vlastních čísel	117
4.6.1	Globální číslo podmíněnosti	117
4.6.2	Odhad chyby vypočítaného vlastního čísla	118
4.7	Cvičení	120
5	Soustavy nelineárních rovnic	121
5.1	Formulace úlohy	121
5.2	Metoda prosté iterace	122
5.3	Newtonova metoda	124
5.4	Gradientní metoda	128
5.5	Cvičení	129
6	Řešení obyčejných diferenciálních rovnic.	131
6.1	Cauchyova úloha.	132
6.2	Základní analytické metody.	134
6.2.1	Lineární rovnice	134
6.2.2	Bernoulliho rovnice	136
6.2.3	Exaktní rovnice	138
6.3	Jednokrokové metody	140
6.3.1	Diferenční metody řešení Cauchyovy úlohy založené na diskretizaci proměnné.	141
6.3.2	Eulerova metoda	141
6.3.3	Metody Rungeho – Kuttovy	144
6.4	Stabilita	146
6.5	Víceprokové metody.	148
6.5.1	Metoda prediktor – korektor	152
6.5.2	Metoda prediktor – modifikátor – korektor	153
6.6	Metody založené na užití derivací vyšších řádů	154
6.7	Metody Taylorovy řady	156

6.8	Diferenciální rovnice vyšších řádů	158
6.8.1	Numerické metody pro rovnice druhé řádu	159
6.8.2	Užití Taylorovy řady	160
6.9	Cvičení	162
6.10	Animace	163
7	Řešení soustav obyčejných diferenciálních rovnic	164
7.1	Základní pojmy	164
7.2	Eulerova metoda pro soustavy diferenciálních rovnic	165
7.3	Rungeho-Kuttova metoda pro soustavy diferenciálních rovnic	166
7.4	Metoda Taylorovy řady	169
7.5	Zaokrouhlovací chyby	170
7.5.1	Vhodnost počátečních podmínek	171
7.5.2	Řízení délky kroku	173
8	Řešení okrajových úloh pro obyčejné diferenciální rovnice.	175
8.1	Základní pojmy	175
8.1.1	Metoda střelby	177
8.1.2	Metoda konečných diferencí	179
8.2	Metoda konečných objemů	190
8.3	Cvičení	192
9	Metoda konečných prvků.	194
9.1	Základní pojmy	194
10	Parciální diferenciální rovnice	198
10.1	Základní pojmy	198
10.2	Parciální diferenciální rovnice prvního řádu	200
10.3	Formulace počáteční úlohy.	201
10.4	Nejjednodušší příklady parciálních rovnic prvního řádu	202
10.4.1	Rovnice typu $\frac{\partial z(x, y)}{\partial x} = 0$	202
10.4.2	Rovnice typu $\frac{\partial z(x, y)}{\partial x} = f(x, y)$	204

10.4.3	Rovnice typu $\frac{\partial z(x, y)}{\partial x} = f_1(x) \cdot f_2(z)$	206
10.4.4	Řešené příklady	208
10.4.5	Lineární homogenní parciální rovnice prvního řádu	210
10.5	Parciální diferenciální rovnice druhého řádu	211
10.5.1	Klasifikace rovnic na hyperbolické, parabolické a eliptické	211
10.6	Transformace proměnných.	213
10.6.1	Řešené příklady	220
10.7	Metoda konečných diferencí pro parciální diferenciální rovnice	222
10.8	Metoda konečných prvků pro parciální diferenciální rovnice	230
10.8.1	Základní pojmy	230
	Cvičení	236
	Výsledky	236
10.9	Cvičení	236

Literatura**237**

Označení

\mathbb{N}	množina přirozených čísel
\mathbb{Z}	množina celých čísel
\mathbb{R}	množina reálných čísel
\mathbb{Q}	množina racionálních čísel
\mathbb{I}	množina iracionálních čísel
\mathbb{C}	množina komplexních čísel
$P_n(x)$	polynom n -tého stupně proměnné x
$A_{m,n}$	matice typu m, n (s m řádky a n sloupci)
$A = (a_{ij})$	matice s prvky a_{ij}
I	jednotková matice
\mathcal{O}	nulová matice
$\det A = A $	determinant matice A
A^{-1}	matice inverzní k matici A
$\text{adj } A$	matice adjungovaná k matici A
A_{ks}	algebraický doplněk prvku a_{ks}
$\text{hod}(A)$	hodnota matice A
$(\mathbb{R}^n, +, \cdot)$	vektorový prostor všech uspořádaných n -tic
$\dim P$	dimenze prostoru P .
$a \cdot b$	skalární součin vektorů a, b
$\ x\ $	norma vektoru x
\square	konec důkazu, konec řešení příkladu
$\langle A \rangle$	lineární obal množiny A
$M_{\mathcal{A}'}^{\mathcal{A}}$	matice přechodu od báze \mathcal{A} k bázi \mathcal{A}'
$a \perp b$	vektor a je ortogonální na vektor b
$f _V = g$	zúžení funkce na podmnožinu
$A \times B$	kartézský součin množin A, B
$a \times b$	vektorový součin vektorů a, b
$[a, b, c]$	smíšený součin vektorů a, b, c
\sim	ekvivalence
\approx	aproximace, přibližné řešení
\doteq	přibližné řešení
$\text{diag}(\lambda_1, \dots, \lambda_n)$	diagonální matice s prvky $\lambda_1, \dots, \lambda_n$ na hlavní diagonále
$\forall x$	pro všechna x
\exists	existuje
$\exists!$	existuje právě jedno
$\varrho(C)$	spektrální poloměr matice C

1 Princip numerických metod

Průvodce studiem

Cílem této kapitoly je seznámit čtenáře se základy teorie chyb, rozdělením chyb a jejich šíření při provádění základních aritmetických operací.

Dále si připomeneme Banachovu větu o pevném bodě, která je teoretickým základem pro všechny iterační metody.

Seznámíme se s podmínkami, které nám zaručují konvergenci jednotlivých iteračních metod k přesnému řešení. Budeme se věnovat odhadům přesnosti výpočtu.

Cíle

Po prostudování této kapitoly budete schopni:

- Rozeznat rychlost šíření chyb během výpočtu.
- Chápat odvození dalších metod pomocí Richardsonovy extrapolace.
- Rozumět principům numerických metod založených na Banachově větě o pevném bodu.

1.1 Chyby při numerických výpočtech.

Omyly člověka, poruchy stroje či zařízení jsou také chybami, ale těmi se nebudeme zabývat. Slovo chyba budeme užívat ve smyslu odchylek, které jsou pravidelným a nerozlučným doprovodem každého numerického řešení.

Rozeznáváme chyby:

- vstupních dat
 - chyby měření
 - chyby způsobené zobrazením vstupních dat v počítači, (Např.: π , e , $\sqrt{2}$, ...).
- numerické metody

- limitu nahradíme členem posloupnosti s dosti vysokým indexem.
Např.: nahrazení funkce jejím rozvojem do Taylorovy řady.

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$$

Volme $x = 1$, potom při použití prvních tří členů řady budeme mít chybu nejvýše $\frac{1}{7!}$.

- úlohu nahradíme jednodušší.
Např.: Určit povrch Země. Použijeme-li vztah $S = 4\pi r^2$ pro povrch koule o poloměru r , potom se dopouštíme chyby, protože jsme Zemi - geoid nahradili koulí.

- zaokrouhlovací

- číslo s nekonečným dekadickým rozvojem nahradíme číslem s konečným počtem členů,
- velká čísla se v počítači zobrazují v semilogaritmickém tvaru,
- všechny nepřesnosti způsobené realizací algoritmu v počítači, včetně nepřesného provádění aritmetických operací.

Přitom je vhodné mít na paměti, že při řešení konkrétní úlohy se obvykle vyskytují všechny druhy chyb současně.

Při výpočtech jsme často nuceni nahradit číslo x jeho aproximací x_n .

$$\begin{aligned} x - x_n = \Delta x & \quad - \text{absolutní chyba aproximace } x_n. \\ |x - x_n| \leq \varepsilon(x_n) & \quad - \text{odhad absolutní chyby - s ní se pracuje} \end{aligned}$$

$$\frac{\Delta x}{x} = \frac{x - x_n}{x}, x \neq 0 \quad - \text{relativní chyba aproximace } x_n$$

$$\left| \frac{x - x_n}{x} \right| \leq \frac{\varepsilon(x_n)}{|x|} = \delta(x_n) \quad - \text{odhad relativní chyby.}$$

Absolutní hodnota relativní chyby se často uvádí v procentech.

1.2 Šíření chyb při aritmetických operacích

Věta 1.1. *Absolutní chyba součtu (rozdílu) je rovna součtu (rozdílu) absolutních chyb.*

$$\Delta(x \pm y) = \Delta x \pm \Delta y.$$

Důkaz. Máme $x \pm y = (x_n + \Delta x) \pm (y_n + \Delta y) = (x_n \pm y_n) + (\Delta x \pm \Delta y)$.
Odečtením získáme tvrzení věty. □

Věta 1.2. *Pro absolutní chybu součinu platí*

$$\Delta(x \cdot y) \approx x_n \Delta y + y_n \Delta x.$$

Důkaz.

$$x \cdot y = (x_n + \Delta x) \cdot (y_n + \Delta y) = x_n \cdot y_n + x_n \Delta y + y_n \Delta x + \Delta x \Delta y,$$

tvrzení věty dostaneme, jestliže zanedbáme člen $\Delta x \Delta y$, o kterém předpokládáme, že je dostatečně malý. \square

Věta 1.3. *Pro absolutní chybu podílu platí*

$$\Delta\left(\frac{x}{y}\right) \approx \frac{\Delta x}{y_n} - \frac{x_n \Delta y}{(y_n)^2}.$$

Důkaz.

Obdobně
$$\frac{x}{y} = \frac{x_n + \Delta x}{y_n + \Delta y} = \frac{x_n + \Delta x}{y_n} \left(1 + \frac{\Delta y}{y_n}\right)^{-1} =$$

poslední závorku vpravo chápeme jako součet geometrické řady a dostáváme

$$= \frac{x_n + \Delta x}{y_n} \left(1 - \frac{\Delta y}{y_n} + \dots\right) \approx$$

a zanedbáním vyšších řádů dostaneme

$$\approx \frac{x_n}{y_n} + \frac{\Delta x}{y_n} - \frac{x_n \Delta y}{(y_n)^2}$$

Odečtením dostaneme tvrzení věty. \square

Věta 1.4. *Relativní chyba součinu či podílu nepřevyší součet relativních chyb činitelů.*

$$\delta(xy) \leq \delta(x) + \delta(y),$$

$$\delta\left(\frac{x}{y}\right) \leq \delta(x) + \delta(y).$$

Důkaz. Analogicky jako u věty 1.3. Doporučuji provést samostatně jako cvičení. \square

Definice 1.5. Nechť A (množina vstupních dat), B (množina výstupních dat) jsou metrické normované prostory. Řekneme, že úloha

$$y = U(x), \quad x \in A, \quad y \in B$$

je *korektní* na (A, B) jestliže

- 1) $\forall x \in A \quad \exists! y \in B$ takové, že platí $y = U(x)$.
- 2) Toto řešení spojitě závisí na vstupních datech, t.j.

$$x_n \rightarrow x, \quad U(x_n) = y_n \Rightarrow U(x_n) \rightarrow U(x) = y.$$

Poznámka 1.6. Jestliže A, B jsou Banachovy prostory¹, pak ke spojitě závislosti řešení stačí

$$\|y_n - y\|_B \leq L\|x_n - x\|_A, \quad L = \text{Konst.}$$

Symbol $\|\cdot\|_B$ označuje normu v prostoru B .

Mnohdy dostačuje jen jiná formulace úlohy, aby se z nekorektní úlohy stala úloha korektní.

Příklad 1.7. „Určit všechny kořeny polynomu.“

Stupeň polynomu určuje počet řešení, která mohou být reálná či komplexní. Úloha proto není jednoznačná a proto se nejedná o korektní úlohu. Stačí ale doplnit podmínku jednoznačnosti: „Určit největší reálný kořen“ a máme korektní úlohu.

Běžně se za korektní úlohy považují i úlohy, které lze pouhou změnou formulace převést na korektní.

Příklad 1.8. Nekorektní úloha z časopisu²

„Kolik stojí kilo hrušek, když v poli stojí jedna?“

1.3 Podmíněnost úloh

Definice 1.9. Řekneme, že korektní úloha je *dobře podmíněna*, jestliže malá změna ve vstupních datech vyvolá malou změnu řešení. Číslo

$$C_p = \frac{\left\| \frac{\Delta y}{y} \right\|}{\left\| \frac{\Delta x}{x} \right\|} = \frac{\text{relativní chyba řešení}}{\text{relativní chyba vstupu}}$$

nazýváme *číslem podmíněnosti* úlohy $y = U(x)$.

Protože většinou umíme stanovit pouze odhady, tak

$$C_p \approx \frac{\delta(y)}{\delta(x)}.$$

Jestliže $C_p \approx 1$ je úloha velmi dobře podmíněná.

Pro C_p velké jde o špatně podmíněnou úlohu. Ale tento pojem je velmi relativní. V praxi se hovoří o špatně podmíněné úloze pro $C_p \geq 100$.

Někteří autoři místo o dobré či špatné podmíněnosti používají termín *malá či velká citlivost vzhledem ke vstupním datům*.

Příklad 1.10. Mějme dán polynom $p(x) = x^2 + x - 1150$. Určete hodnotu polynomu v bodech $x = 33$ a $x = \frac{100}{3}$ a odhadněte hodnotu čísla podmíněnosti.

¹Jsou definovány dále, viz poznámka (1.29)

²Publikováno v časopise „Mladý svět“ kolem roku 1970.

Řešení. Dosazením dostaneme

$$p(33) = -28, \quad p\left(\frac{100}{3}\right) \approx -5.6.$$

$$\Delta(x) = \frac{1}{3} \quad \Delta(y) = 22.4 \quad \Rightarrow C_p \approx 79.178.$$

Jedná se proto o nestabilní úlohu. Přitom jde o hodnoty blízké ke kořenu $x = 33.41533576$. \square

Podmíněnost matic.

Nechť máme soustavu lineárních algebraických rovnic $Ax = b$, kde A je matice koeficientů, b je sloupec pravých stran a x je sloupec neznámých. Nechť x_n je aproximace řešení této soustavy. Označme $r = b - Ax_n$, kde r je reziduum. Je-li složky vektoru r malé, ještě to nic neříká o přesnosti výpočtu, neboť

$$r = b - Ax_n = Ax - Ax_n = A(x - x_n),$$

odtud po úpravě dostáváme

$$A^{-1}r = x - x_n$$

a jsou-li prvky matice A^{-1} dostatečně velké, bude i rozdíl $x - x_n$ velký i pro velmi malé r .

Platí
$$C_p = \|A^{-1}\| \cdot \|A\|.$$

Symbolem $\|\cdot\|$ označujeme normu matice, viz definici (1.31).

Příklad 1.11. Mějme dánu soustavu $Ax = b$, kde

$$A = \begin{pmatrix} 4,1 & 2,8 \\ 9,7 & 6,6 \end{pmatrix}.$$

Jestliže zvolíme jako pravou stranu $b = \begin{pmatrix} 4,1 \\ 9,7 \end{pmatrix}$, potom dostaneme řešení $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Jestliže zvolíme jako pravou stranu $b = \begin{pmatrix} 4,11 \\ 9,7 \end{pmatrix}$, potom dostaneme řešení $x = \begin{pmatrix} 0,34 \\ 0,97 \end{pmatrix}$. Jde o špatně podmíněnou soustavu s $C_p > 2249,5$ (pro Eukleidovskou normu dostáváme $C_p > 1622$).

Řešení špatně podmíněné soustavy musíme interpretovat velmi opatrně. Je to vlastnost dané matice, nikoliv postupu řešení. Proto je vhodné se zabývat otázkou, zda je úloha dobře formulovaná, zda nejde úlohu nějak upravit a podobně, aby bylo možné se vyhnout použití špatně podmíněné matice při řešení.

Poznámka 1.12. Teoretické odhady chyb bývají značně pesimistické. Horní hranice chyb bývá většinou jen zřídka dosaženo, protože v průběhu výpočtu dochází k určité vzájemné kompenzaci chyb. Příkladem může být Eratosthenovo¹ měření obvodu Země.

1.4 Richardsonova extrapolace

Jde o univerzální postup, který nám umožňuje pomocí základní metody s nižší přesností vytvářet metody s přesností vyšší.

Nechť je základní metoda reprezentovaná funkcí $F(h)$ s parametrem h (může jím být například velikost kroku dané metody). Pomocí této metody umíme spočítat hodnotu $F(h)$ pro malé $h > 0$. Chceme co nejpřesněji aproximovat hodnotu $F(0)$, kterou ale neumíme určit přímo z funkce $F(h)$.

Nechť funkci $F(h)$ můžeme zapsat ve tvaru mocninné řady

$$F(h) = a_0 + a_1 h^2 + a_2 h^4 + a_3 h^6 + \dots \quad (1.1)$$

Pro malé h můžeme položit $h = 0$ a dostaneme aproximaci $F(0) \approx a_0$.

Hledejme lepší aproximaci. Podle (1.1) platí

$$F\left(\frac{h}{2}\right) = a_0 + a_1 \left(\frac{h}{2}\right)^2 + a_2 \left(\frac{h}{2}\right)^4 + a_3 \left(\frac{h}{2}\right)^6 + \dots \quad (1.2)$$

Odstraníme z rovnic (1.1) a (1.2) člen obsahující druhou mocninu. Ten totiž představuje největší chybu v rozdílu $a_0 - F(h)$ i $a_0 - F\left(\frac{h}{2}\right)$. Rovnici (1.2) vynásobíme čtyřmi a odečteme od ní (1.1), dostaneme

$$\begin{aligned} 4F\left(\frac{h}{2}\right) - F(h) &= 4a_0 + 4a_1 \left(\frac{h}{2}\right)^2 + 4a_2 \left(\frac{h}{2}\right)^4 + 4a_3 \left(\frac{h}{2}\right)^6 + \dots \\ &\quad - a_0 - a_1 h^2 - a_2 h^4 - a_3 h^6 - \dots \\ &= 3a_0 + 4a_2 \left(\frac{h}{2}\right)^4 - a_2 h^4 + 4a_3 \left(\frac{h}{2}\right)^6 - a_3 h^6 + \dots \end{aligned}$$

¹ **Eratosthenés z Kýreny** (275 — 195 př.n.l.) rodák z Kýreny v nynější Libyi. Základní vzdělání získal v Athénách. Později působil v Alexandrii, pracoval společně s Eukleidem, Apoloniem z Pergy (260 — 170 př.n.l.), byl přítelem Archimedovým (287 — 212 př.n.l.). Eratosthenés byl všestranným pracovníkem — věnoval se gramatice, literární historii, matematice, astronomii, chronologii, etice, geografii a kartografii. Pokusil se změřit a vypočítat obvod Země. Jeho výsledek je neuvěřitelně přesný (chyba je asi 0,8%). Jeho kolegové a spolupracovníci mu za jeho pracovitost a dosažené výsledky přezdívali “Pentathlos”, t.j. atlet–pětibojař, který dosahuje výborných výsledků v různých oblastech, ale ani v jedné z nich se nestane nejlepším. Tato přezdívka je spojena s jemnou výčitkou. Stoupenci specializace mu s jistou povýšeností přezdívali “Beta” — (název čísla 2) t.j. druhořadý. Někdy je také tato přezdívka vysvětlována tím, že Eratosthenés byl jako padesátiletý povolán ke dvoru Ptolemaia III. (vládl 247 — 221 př.n.l.), kde se stal vychovatelem následníka trůnu. A s tímto titulem, asi jako finanční zabezpečení, byl jmenován i *druhým* hlavním knihovníkem alexandrijské knihovny. Eratosthenés ke konci života oslepl a dobrovolně odešel ze světa. Z celého jeho rozsáhlého díla se dochovaly pouze zlomky jako citace pozdějších autorů. Eratosthenovo síto k určení prvočísel v množině $\{1, 2, \dots, n\}$ se stalo východiskem pro celou jednu část teorie čísel.

$$= 3a_0 + a_2 \left(\frac{1}{4} - 1\right) h^4 + a_3 \left(\frac{1}{8} - 1\right) h^6 + \dots$$

Rovnici vydělíme třemi a dostaneme novou funkci

$$F_2(h) = \frac{4F\left(\frac{h}{2}\right) - F(h)}{3} = a_0 + a_2^{(2)}h^4 + a_3^{(2)}h^6 + \dots \quad (1.3)$$

Přitom platí, že $|a_i^{(2)}| < |a_i|$, $i = 2, 3, \dots$. Proto je $F_2(h)$ lepší aproximací pro a_0 než $F(h)$.

Pro dosti malá h je také $F_2(h)$ lepší aproximací než $F\left(\frac{h}{2}\right)$, protože rozdíl $F_2(h) - a_0$ začíná až čtvrtou mocninou h .

Dostali jsme tak metodu F_2 , která je pro dostatečně malá h lepší než metoda F .

Analogickým způsobem můžeme odstranit z F_2 čtvrtou mocninu a získáme ještě lepší aproximaci $F(0)$.

Podle (1.3) platí

$$F_2\left(\frac{h}{2}\right) = a_0 + a_2^{(2)}h^4 + a_3^{(3)}h^6 + \dots \quad (1.4)$$

Rovnici (1.4) vynásobíme 16 a odečteme od ní rovnici (1.3), výsledek dělíme 15. Dostaneme

$$F_3(h) = \frac{16F_2\left(\frac{h}{2}\right) - F_2(h)}{15} = a_0 + a_3^{(3)}h^6 + \dots$$

Takto můžeme pokračovat dále a získávat stále lepší aproximace pro které platí

$$F_{i+1}(h) = \frac{4^i F_i\left(\frac{h}{2}\right) - F_i(h)}{4^i - 1} = a_0 + a_{i+1}^{(i+1)}h^{2i+2} + \dots$$

přičemž $F_1(h) = F(h)$. Výpočet si můžeme zapsat do tabulky (vyplňuje se po řádcích)

T_{00}				
T_{10}	T_{11}			
T_{20}	T_{21}	T_{22}		
T_{30}	T_{31}	T_{32}	T_{33}	
...

přičemž

$$T_{s0} = F\left(\frac{h}{2^s}\right), \quad s = 0, 1, \dots$$

$$T_{si} = \frac{4^i T_{s,i-1} - T_{s-1,i-1}}{4^i - 1}.$$

Výpočet ukončíme a hodnotu T_{ss} považujeme za dostatečně přesnou, jestliže $|T_{ss} - T_{s,s-1}| < \varepsilon$, kde ε je předem zadaná požadovaná přesnost.

Speciálním případem Richardsonovy extrapolace je vyjádření Simpsonova¹ pravidla pomocí lichoběžníkového pravidla při numerické integraci (viz Matematika 3). Platí

$$S_n = S_{2k} = \frac{1}{3}(4L_{2k} - L_k), \quad (1.5)$$

¹**T. Simpson** (1710 – 1761) anglický matematik, samouk. Pracoval jako tkadlec a od r.1743 vyučoval na vojenské akademii. Ve stejném roce odvodil svůj vzorec pro přibližnou integraci. Jeden ze zakladatelů teorie chyb.

kde L_{2k}, L_k jsou hodnoty získané lichoběžníkovým pravidlem a index udává počet dělení intervalu.

1.5 Metriky, kontrakce, Banachova věta

S pojmy metrika, norma, kontrakce jste se setkali už v předmětu Matematika 3. Zde si je znovu připomeneme, korektně zadefinujeme, odvodíme si jejich základní vlastnosti a ukážeme si jejich použití.

Definice 1.13. Nechť M je neprázdná množina. Zobrazení $d : M \times M \rightarrow \mathbb{R}$ splňující pro všechna $x, y, z \in M$ následující axiomy

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

nazveme *metrikou* a dvojici (M, d) *metrickým prostorem*.

Příklad 1.14.

- a) $M \equiv \mathbb{R}$, $d(x, y) = |x - y|$
- b) $M \equiv \mathbb{R}^n$, $d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$ (Eukleidovská metrika)
- c) $M \equiv \mathbb{R}^n$, $d(x, y) = \max_{i=1, \dots, n} |x_i - y_i|$ (krychlová metrika)
- d) $M \equiv \mathbb{R}^n$, $d(x, y) = \sum_{i=1}^n |x_i - y_i|$ (oktaetická metrika)

Definice 1.15. Bod $x \in M$ je *limitou* posloupnosti $\{x_n\}_{n=1}^{\infty}$ v metrickém prostoru (M, d) , jestliže $\forall \varepsilon > 0 \exists n_0 \in \mathbb{N}$, takové, že $\forall n > n_0 : d(x_n, x) < \varepsilon$. Posloupnost, která má limitu, se nazývá *konvergentní*. Označení $\lim_{n \rightarrow \infty} x_n = x$.

Věta 1.16. Každá posloupnost v metrickém prostoru (M, d) může mít nejvýše jednu limitu.

Důkaz byl proveden v 1. semestru v předmětu Matematika 1.

Definice 1.17. Posloupnost $\{x_n\}_{n=1}^{\infty}$ se nazývá *cauchyovská*¹, jestliže $\forall \varepsilon > 0 \exists n_0 \in \mathbb{N}$ takové, že $\forall n > n_0 \forall p \in \mathbb{N} : d(x_n, x_{n+p}) < \varepsilon$.

¹**L. A. Cauchy** (1789 – 1857) Francouzský matematik, jeden z tvůrců moderní matematiky. Je autorem více než 800 prací z teorie čísel, algebry, matematické analýzy, diferenciálních rovnic, mechaniky, aj. S využitím pojmu “limity” vybudoval systematicky základy analýzy. Po francouzské červencové revoluci v roce 1830 pobýval nějakou dobu i v Praze jako vychovatel vnuků sesazeného krále Karla X.

Věta 1.18. Každá konvergentní posloupnost je cauchyovská.

Důkaz. Protože je $\{x_n\}$ konvergentní, existuje limita $\lim_{n \rightarrow \infty} x_n = x$. To ale (podle předchozích definic) znamená, že $\forall \varepsilon/2 > 0 \exists n_0 \in \mathbb{N}$ takové, že $\forall n > n_0 : d(x_n, x) < \frac{1}{2}\varepsilon$. Potom, ale i $\forall p \in \mathbb{N}$ je i $d(x_{n+p}, x) < \frac{1}{2}\varepsilon$. Takže máme (s využitím trojúhelníkové nerovnosti)

$$d(x_n, x_{n+p}) \leq d(x_n, x) + d(x_{n+p}, x) < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon.$$

□

Z následujícího příkladu vyplývá, že větu nelze obrátit.

Příklad 1.19. $M = \mathbb{R}_+ = (0, +\infty)$, $x_n = \frac{1}{n}$, $\lim x_n = 0 \notin M$. Protože každá posloupnost v (M, d) může mít nejvýše jednu limitu, je tento příklad příkladem posloupnosti, která je cauchyovská, ale není v (M, d) konvergentní.

Definice 1.20. Metrický prostor (M, d) se nazývá *úplný*, jestliže v něm má každá cauchyovská posloupnost limitu.

Příklad 1.21. Každá uzavřená neprázdna podmnožina metrického prostoru je úplná.

Definice 1.22. Na množině M je definována *binární operace* $\circ : M \times M \rightarrow M$, jestliže $\forall x, y \in M$ platí $(x \circ y) \in M$.

Příklad 1.23. Na množině celých čísel tvoří sčítání binární operaci, protože součet libovolných dvou celých čísel je opět celé číslo.

Na množině celých čísel dělení netvoří binární operaci, protože dělení nulou není definováno a existuje nekonečně mnoho dvojic celých čísel, jejichž podíl není číslo celé, jako například čísla 2 a 3.

Definice 1.24. Nechť M je množina a (\circ) je binární operace na M . Uspořádaná dvojice (M, \circ) se nazývá *grupou*, jestliže platí

$$A) \quad (x \circ y) \circ z = x \circ (y \circ z), \quad \forall x, y, z \in M,$$

$$B) \quad \exists e \in M : e \circ x = x \circ e = x, \quad \forall x \in M,$$

$$C) \quad \forall x \in M \exists x^{-1} \in M : x \circ x^{-1} = x^{-1} \circ x = e.$$

Definice 1.25. Pokud v grupě (M, \circ) platí navíc komutativní zákon

$$D) \quad x \circ y = y \circ x, \quad \forall x, y \in M,$$

pak mluvíme o *komutativní grupě* (abelovské grupě).

Definice 1.26. Necht' $(L, +)$ je komutativní grupa, a necht' je dále definována binární operace $\cdot : \mathbb{R} \times L \rightarrow L$, $(\alpha, x) \mapsto \alpha \cdot x$, která splňuje následující podmínky $\forall x, y \in L, \forall \alpha, \beta \in \mathbb{R}$:

1. $\alpha \cdot (\beta \cdot x) = (\alpha\beta) \cdot x$ (asociativita pro násobení)
2. $\alpha \cdot (x + y) = (\alpha \cdot x) + (\alpha \cdot y)$ (distributivita I.)
 $(\alpha + \beta) \cdot x = (\alpha \cdot x) + (\beta \cdot x)$ (distributivita II.)
3. $1 \cdot x = x$

Potom uspořádaná trojice $(L, +, \cdot)$ tvoří *vektorový prostor* nad \mathbb{R} . Prvky z L budeme nazývat vektory, prvky z \mathbb{R} skaláry. Značit budeme vektory malými písmeny latinky a skaláry malými písmeny řecké abecedy.

Definice 1.27. Vektorový prostor \mathcal{V} nazveme *normovaným prostorem*, jestliže existuje zobrazení $\|\cdot\| : \mathcal{V} \rightarrow \mathbb{R}$ takové, že $\forall x, y \in \mathcal{V}, \forall \alpha \in \mathbb{R}$ platí:

1. $\|x\| \geq 0 \wedge \|x\| = 0 \Leftrightarrow x = \mathcal{O}$, kde \mathcal{O} je nulový vektor.
2. $\|\alpha x\| = |\alpha| \cdot \|x\|$.
3. $\|x + y\| \leq \|x\| + \|y\|$.

Toto zobrazení nazýváme *vektorovou normou*.

Věta 1.28. Necht' \mathcal{V} je normovaný prostor. Definujeme metriku d jako

$$\forall x, y \in \mathcal{V}, d(x, y) = \|x - y\|.$$

Potom (\mathcal{V}, d) je metrický prostor.

Důkaz. Provedeme prověrkou axiomů. Doporučuji provést jako cvičení. □

Poznámka 1.29. Metrický prostor, který je úplný v takto definované metrice, se nazývá *Banachův*¹.

Příklad 1.30.

a) $\mathcal{V} = \mathbb{R}^n$, $\|x\| = \max\{|x_1|, |x_2|, \dots, |x_n|\}$, potom lze $d(x, y)$ definovat takto

$$d(x, y) = \max_i |x_i - y_i|.$$

b) $\mathcal{V} = \mathbb{R}^n$, $\|x\| = \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}}$, $p \geq 1$, potom je metrikou funkce

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

¹**Stefan Banach** (1892 — 1945) vynikající polský matematik. Jeden ze zakladatelů moderní funkcionální analýzy.

c) $\mathcal{V} = C[a, b]$, tj. množina všech spojitých funkcí na $[a, b]$.¹ Skalární součin je $x \bullet y = \int_a^b x(t)y(t)dt$. Potom

$$\|x\| = \sqrt{x \bullet x}, \quad d(x, y) = \|x - y\| = \sqrt{\int_a^b (x(t) - y(t))^2 dt}.$$

Definice 1.31. Nechť M_n je množina všech čtvercových matic řádu n . Nechť $\forall A, B \in M_n, \forall \alpha \in \mathbb{R}$ zobrazení $\|\cdot\| : M_n \rightarrow \mathbb{R}$ splňuje axiomy

1. $\|A\| \geq 0 \wedge \|A\| = 0 \Leftrightarrow A = \mathcal{O}$, kde \mathcal{O} je nulová matice.
2. $\|\alpha A\| = |\alpha| \cdot \|A\|$.
3. $\|A + B\| \leq \|A\| + \|B\|$.
4. $\|AB\| \leq \|A\| \cdot \|B\|$.

Potom toto zobrazení nazveme *maticovou normou*.

Příklad 1.32. Mějme matici A s prvky $a_{ij}, i, j \in \{1, 2, \dots, n\}$, potom její eukleidovská norma má tvar

$$\|A\| = \left(\sum_{i,j=1}^n (a_{ij})^2 \right)^{\frac{1}{2}}.$$

Definice 1.33. Řekneme, že maticová norma $\|A\|_1$ je *souhlasná* s vektorovou normou $\|x\|_2$ v \mathbb{R}^n , jestliže $\|Ax\|_2 \leq \|A\|_1 \cdot \|x\|_2$.

Příklad 1.34. Souhlasné normy

$$\text{a) } \|x\|_2 = \max_i |x_i|, \quad \|A\|_1 = \max_i \sum_{j=1}^n |a_{ij}|, \quad x \in \mathbb{R}^n, A \in M_n,$$

$$\text{b) } \|x\|_2 = \sum_{i=1}^n |x_i|, \quad \|A\|_1 = \max_j \sum_{i=1}^n |a_{ij}|,$$

$$\text{c) } \|x\|_2 = \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}}, \quad \|A\|_1 \leq \left(\sum_{i,j=1}^n a_{ij}^2 \right)^{\frac{1}{2}}.$$

V posledním případě jde o odhad. Přesný tvar je $\|A\| = \sqrt{\varrho(A^T A)} = \sqrt{\varrho(AA^T)}$, kde $\varrho(B)$ je spektrální poloměr matice B .

¹ $[a, b]$ označuje uzavřený interval a (a, b) otevřený interval.

Definice 1.35. Spektrální poloměr matice B je $\rho(B) = \max_i \{|\lambda_i|\}$, kde λ_i jsou vlastní čísla matice B .

Definice 1.36. Nechť (M, d) je metrický prostor. Zobrazení $\varphi : M \rightarrow M$ nazveme *kontrakcí*, jestliže existuje konstanta k , $0 \leq k < 1$ taková, že platí

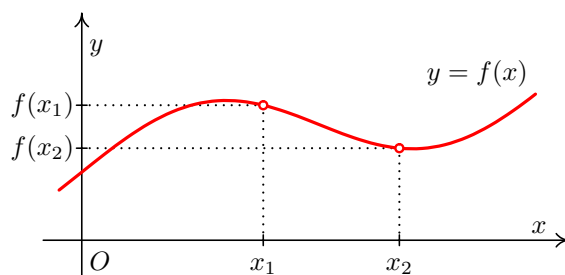
$$d(\varphi(x), \varphi(y)) \leq k \cdot d(x, y).$$

Číslo k se nazývá koeficient kontrakce.

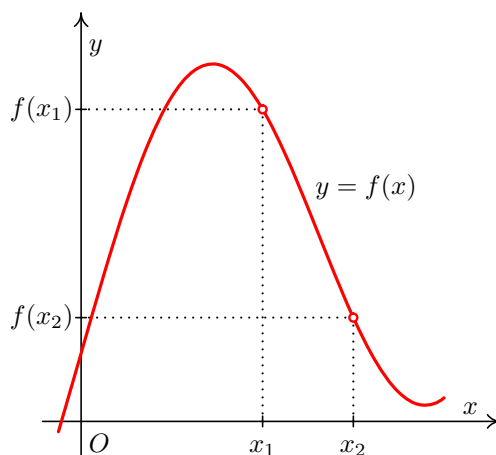
Definice 1.37. Nechť φ je zobrazení z M do M . Bod $\tilde{x} \in M$ nazveme *pevným bodem* zobrazení φ , jestliže je zobrazen sám na sebe, tj. jestliže platí

$$\tilde{x} = \varphi(\tilde{x}).$$

Jedna z možných interpretací kontraktivního zobrazení (na prostoru spojitých funkcí jedné proměnné) je znázorněna na Obr. 1.1 a 1.2.



Obr. 1.1: Funkce, která je kontraktivní



Obr. 1.2: Funkce, která není kontraktivní

Věta 1.38. Banachova věta o pevném bodu

Nechť (M, d) je úplný metrický prostor, φ je kontrakce na M s koeficientem k . Potom:

- 1) Existuje právě jeden bod $\tilde{x} \in M$ takový, že je splněna rovnice

$$\varphi(\tilde{x}) = \tilde{x}.$$

2) Zvolíme-li $x_1 \in M$ libovolně a sestrojíme-li posloupnost $\{x_i\}_{i=1}^{\infty}$ předpisem

$$x_{i+1} = \varphi(x_i), \text{ pak } \lim_{i \rightarrow \infty} x_i = \tilde{x}.$$

3) Platí odhady pro $n = 1, 2, \dots$

$$d(x_n, \tilde{x}) \leq \frac{k^{n-1}}{1-k} d(x_1, x_2),$$

$$d(x_n, \tilde{x}) \leq \frac{k}{1-k} d(x_{n-1}, x_n).$$

Důkaz. Posloupnost $\{x_i\}_{i=1}^{\infty}$, kde $x_{i+1} = \varphi(x_i)$, se nazývá *iterační proces*, x_1 je *počáteční aproximace*, x_n je *n-tá aproximace*.

a) Nejprve ukážeme, že pevný bod, pokud existuje, je určen jednoznačně. Předpokládejme, že existují dva různé body $a \in M, b \in M$ pro něž platí $a = \varphi(a), b = \varphi(b), d(a, b) \neq 0$. Potom

$$d(a, b) = d(\varphi(a), \varphi(b)) \leq k d(a, b) \quad (1.6)$$

a po zkrácení nenulovým $d(a, b)$ dostaneme

$$1 \leq k.$$

Protože φ je kontrakce, platí pro k nerovnost $0 \leq k < 1$. Dostali jsme spor. Odtud plyne, že nerovnost (1.6) je splněna pouze pro případ $d(a, b) = 0$, což je spor s předpokladem, že body a, b jsou různé a tedy pevně bod, pokud existuje, je určen jednoznačně.

b) Dále máme

$$\begin{aligned} d(x_j, x_{j+1}) &= d(\varphi(x_{j-1}), \varphi(x_j)) \leq k d(x_{j-1}, x_j) = \\ &= k d(\varphi(x_{j-2}), \varphi(x_{j-1})) \leq k^2 d(x_{j-2}, x_{j-1}) = \dots \Rightarrow \\ d(x_j, x_{j+1}) &\leq k^{j-1} d(x_1, x_2) \end{aligned}$$

a protože $0 \leq k < 1$ dostáváme, že $\lim_{j \rightarrow \infty} d(x_j, x_{j+1}) = 0$.

Pro $p \in \mathbb{N}$ pak máme $\forall n \in \mathbb{N}$

$$d(x_n, x_{n+p}) \leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \dots + d(x_{n+p-1}, x_{n+p}) \leq \quad (1.7)$$

$$\begin{aligned} &\leq (k^{n-1} + k^n + \dots + k^{n+p-2}) d(x_1, x_2) \leq \\ &\leq (k^{n-1} + k^n + \dots) d(x_1, x_2) = \frac{k^{n-1}}{1-k} d(x_1, x_2). \end{aligned} \quad (1.8)$$

Protože $0 \leq k < 1$, tak

$$\lim_{n \rightarrow \infty} d(x_n, x_{n+p}) = 0 \quad \text{pro } \forall p \in \mathbb{N},$$

což znamená, že naše posloupnost $\{x_j\}$ je cauchyovská a tedy existuje \tilde{x} tak, že $\tilde{x} = \lim_{n \rightarrow \infty} x_n$,

protože prostor (M, d) je úplný.

c) První odhad plyne přímo z (1.8) pro $p = \infty$. Druhý odhad dostaneme z (1.8), když v (1.7) položíme $p = \infty$. Potom

$$\begin{aligned} d(x_n, \tilde{x}) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \dots + d(x_{n+p-1}, x_{n+p}) + \dots \leq \\ &\leq d(x_{n-1}, x_n)(k + k^2 + k^3 + \dots) \leq \frac{k}{1-k} d(x_{n-1}, x_n). \end{aligned}$$

□

Pojmy k zapamatování

- Seznámili jsme se se základy teorie chyb a jejich šíření při provádění základních aritmetických operací.
- Vyslovili a dokázali jsme si Banachovu větu o pevném bodě a jako její důsledek jsme si odvodili vzorce pro odhad přesnosti numerického výpočtu.
- Ukázali jsme si použití Richardsonova extrapolace pro odvození přesnějších vzorců.
- Seznámili se s pojmy metrika a norma.

Kontrolní otázky

1. Co rozumíme pojmem kontrakce?
2. Může platit Banachova věta i pro záporné funkce?
3. Je možné ovlivnit rychlost šíření chyb, tj. narůstání chyby?

2 Soustavy lineárních rovnic

Průvodce studiem

Cílem této kapitoly je seznámit čtenáře s aplikacemi Banachovy věty při hledání řešení soustav lineárních algebraických rovnic. Nejdříve si však připomeneme finitní metody řešení těchto soustav, a to nejen Gaussovu a Jordanovu eliminační metodu, ale i řešení pomocí inverzní matice a pomocí LU rozkladu. Potom se seznámíme s aplikací Banachovy věty pro soustavy lineárních algebraických rovnic. Zde si zopakujeme Jacobiho a Gaussovu-Seidelovu iterační metodu a ukážeme si některé další metody, které se používají především pro řešení soustav vysokých řádů.

Seznámíme se s podmínkami, které nám zaručují konvergenci jednotlivých iteračních metod k přesnému řešení soustavy rovnic. Budeme se věnovat odhadům přesnosti výpočtu.

Cíle

Po prostudování této kapitoly budete schopni:

- Numericky řešit soustavy lineárních algebraických rovnic.
- Umět vybrat vhodnou metodu, která povede k cíli, tj. k řešení.
- Odhadnout nutný počet kroků pro dosažení požadované přesnosti.

2.1 Soustavy lineárních rovnic — Základní pojmy

Se soustavami lineárních algebraických rovnic a jejich řešením jste se už setkali v předmětu Matematika 1 a v předmětu Matematika 3 jste si uvedli i dvě numerické metody pro jejich řešení.

Ale až zde bude uveden solidní teoretický základ, všechny pojmy budou řádně definovány a budou odvozeny potřebné vztahy. Naším cílem je ukázat všechny souvislosti a ukázat rozsah jeho možného použití.

Definice 2.1. Maticová rovnice $Ax = b$, kde $A \in \mathbb{R}_{m,n}$, $b \in \mathbb{R}_{m,1}$, $x \in \mathbb{R}_{n,1}$ se nazývá *soustava lineárních algebraických rovnic*.

V rozepsaném tvaru máme

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ \cdot & \quad \cdot \quad \cdot \quad \cdot \quad \dots \quad \cdot \quad \cdot \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned} \tag{2.1}$$

A je matice koeficientů, b je sloupec pravých stran,

$$(A|b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \cdot & \cdot & \dots & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right) \text{ se nazývá matice rozšířená .}$$

Každý sloupec (sloupcová matice) α , pro který platí $A\alpha = b$, se nazývá *řešením* soustavy (2.1).

Definice 2.2.

Soustava (2.1) je *řešitelná*, má-li aspoň jedno řešení.

Soustava (2.1) je *jednoznačně řešitelná*, má-li právě jedno řešení.

Soustava (2.1) je *víceznačně řešitelná*, má-li více než jedno řešení.

Definice 2.3. Soustava lineárních algebraických rovnic se nazývá *homogenní*, jestliže je tvaru

$$Ax = \mathcal{O}, \tag{2.2}$$

kde \mathcal{O} je nulový sloupec. V opačném případě mluvíme o *nehomogenní* soustavě.

Definice 2.4. Je-li $Ax = b$ nehomogenní soustava, pak *přidruženou homogenní* soustavou rozumíme soustavu $Ax = \mathcal{O}$ (t.j. homogenní soustavu se stejnou maticí koeficientů jakou má nehomogenní soustava).

Příklad 2.5. Mějme danu nehomogenní soustavu

$$\begin{aligned} 3x_1 + x_2 - 4x_3 &= 1 \\ x_1 - 2x_2 + x_3 &= 5 \\ 2x_1 - x_2 - 3x_3 &= 4 \end{aligned}$$

Přidružená homogenní soustava má tvar

$$\begin{aligned} 3x_1 + x_2 - 4x_3 &= 0 \\ x_1 - 2x_2 + x_3 &= 0 \\ 2x_1 - x_2 - 3x_3 &= 0 \end{aligned}$$

2.1.1 Řešení soustav

Věta 2.6. *Nechť soustava $Ax = b$ má čtvercovou regulární matici koeficientů. Potom má tato soustava právě jedno řešení, které můžeme určit použitím **Cramerových**¹ vzorců: k -tý člen řešení je zlomek, v jehož jmenovateli je determinant matice koeficientů A a v čitateli determinant matice, kterou získáme z matice A nahrazením k -tého sloupce sloupcem pravých stran soustavy (2.1) a ostatní sloupce ponecháme.*

Důkaz. Máme $Ax = b$, $|A| \neq 0$, takže existuje A^{-1} . Potom

$$x = A^{-1}b = \frac{1}{|A|}(\text{adj } A)b = \frac{1}{|A|} \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \cdot & \cdot & \dots & \cdot \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix} b.$$

Teď si jen stačí uvědomit, že v prvním řádku matice $\text{adj } A$ jsou algebraické doplňky příslušné k prvnímu sloupci matice A . Potom součin prvního řádku matice $\text{adj } A$ se sloupcem b můžeme podle Laplaceovy věty o rozvoji determinantu chápat jako rozvoj determinantu matice podle prvního sloupce, kde matice má jako první sloupec sloupec b a zbývající sloupce jsou z matice A . Obdobně pro další prvky. \square

Příklad 2.7. Najděte řešení soustavy rovnic

$$\begin{aligned} 3x_1 + x_2 - 4x_3 &= 1 \\ x_1 - 2x_2 + x_3 &= 5 \\ 2x_1 - x_2 - 3x_3 &= 4 \end{aligned}$$

Řešení. Určíme si determinant matice koeficientů

$$|A| = \begin{vmatrix} 3 & 1 & -4 \\ 1 & -2 & 1 \\ 2 & -1 & -3 \end{vmatrix} = 14$$

Determinant matice A je nenulový, soustava je tedy jednoznačně řešitelná. Spočítáme si determinanty matic D_i , kde matice D_i vznikne z matice A nahrazením i -tého sloupce sloupcem pravých stran naší soustavy.

$$|D_1| = \begin{vmatrix} 1 & 1 & -4 \\ 5 & -2 & 1 \\ 4 & -1 & -3 \end{vmatrix} = 14, \quad |D_2| = \begin{vmatrix} 3 & 1 & -4 \\ 1 & 5 & 1 \\ 2 & 4 & -3 \end{vmatrix} = -28, \quad |D_3| = \begin{vmatrix} 3 & 1 & 1 \\ 1 & -2 & 5 \\ 2 & -1 & 4 \end{vmatrix} = 0.$$

Potom $x_i = \frac{|D_i|}{|A|}$, takže máme

$$x_1 = \frac{14}{14} = 1, \quad x_2 = \frac{-24}{14} = -2, \quad x_3 = \frac{0}{14} = 0.$$

\square

¹ **Gabriel Cramer** (31. 7. 1704 – 4. 1. 1752) švýcarský matematik, přírodovědec a technik. V matematice se věnoval hlavně geometrii a teorii pravděpodobnosti. V r. 1750 vydal knihu o algebraických křivkách, kde je v dodatku uveden způsob vyloučení $(n - 1)$ neznámých ze soustavy n rovnic o n neznámých. Přes nevhodnou symboliku tím položil základy teorie determinantů.

Poznámka 2.8. Cramerovy vzorce sice dávají přesné řešení, ale je zapotřebí pro ně vypočítat $(n + 1)$ determinantů n -tého řádu. Pro rozsáhlejší soustavy je jejich použití problematické, protože ani s pomocí výpočetní techniky nejsme schopni určit přesné hodnoty determinantů. Cramerovy vzorce nemají obecnou platnost, předpokládají regularitu matice koeficientů.

Věta 2.9. Frobeniova,¹ Kroneckerova² — Capelliho³, existenční.

Soustava (2.1) je řešitelná právě tehdy, když hodnota matice koeficientů se rovná hodnotě matice rozšířené.

Poznámka 2.10. Jinak řečeno:

$h(A) < h(A|b)$ — soustava (2.1) nemá řešení,

$h(A) = h(A|b)$ — soustava (2.1) je řešitelná,

$h(A) > h(A|b)$ — nemůže nikdy nastat. Přidáním dalšího sloupce můžeme hodnotu matice zvýšit, ale nikdy ne snížit.

Důkaz. věty (2.9): Necht' je soustava (2.1) řešitelná. Potom existuje sloupec

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)^T$$

takový, že platí $A\alpha = b$. Vezmeme si matici rozšířenou soustavy (2.1) a od posledního sloupce odečteme α_1 násobek prvního sloupce, α_2 násobek druhého sloupce, atd. až α_n násobek n -tého sloupce. Dostaneme

$$(A|b) \sim \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 - \alpha_1 a_{11} - \alpha_2 a_{12} - \dots - \alpha_n a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 - \alpha_1 a_{21} - \alpha_2 a_{22} - \dots - \alpha_n a_{2n} \\ \cdot & \cdot & \dots & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m - \alpha_1 a_{m1} - \alpha_2 a_{m2} - \dots - \alpha_n a_{mn} \end{array} \right).$$

Protože α je řešením soustavy (2.1), tak platí pro všechna $i = 1, 2, \dots, m$

$$b_i - \alpha_1 a_{i1} - \alpha_2 a_{i2} - \dots - \alpha_n a_{in} = 0.$$

V posledním sloupci budou stát samé nuly a my máme

$$(A|b) \sim \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 \\ \cdot & \cdot & \dots & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} & 0 \end{array} \right) \sim A$$

¹**Georg Ferdinand Frobenius** (26. 10. 1849 – 3. 8. 1917) německý matematik. Zabýval se hlavně algebrou. Vyslovil existenční větu pro řešení soustavy lineárních algebraických rovnic. Patří mu vynikající práce z oblasti kvadratických forem, maticového počtu a teorie konečných grup. Zavedl řadu pojmů moderní algebry.

²**Leopold Kronecker** (7. 12. 1823— 29. 12. 1891) německý matematik. Zabýval se teorií čísel, teorií kvadratických forem, teorií grup, teorií eliptických funkcí. Byl odpůrcem Cantorovy teorie množin. Odvodil metodu, kterou lze vždy nalézt všechny racionální kořeny polynomu s racionálními koeficienty (i když mnohdy obtížně a zdlouhavě). Dokázal existenční větu pro řešení soustavy lineárních algebraických rovnic.

³**Alfredo Capelli** (5. 6. 1855 – 28. 1. 1910) italský matematik, působil v Neapoli. Významně přispěl k rozvoji teorie algebraických rovnic. Dále se zabýval teorií funkcí komplexní proměnné a teorií diferenciálních rovnic.

Nulový sloupec jsme vynechali. Použili jsme pouze elementární úpravy, které nemění hodnotu matice. Proto platí: Jestliže je soustava (2.1) řešitelná, potom $h(A) = h(A|b)$.

Z druhé strany: Nechť je $h(A) = h(A|b) = k$. Potom existuje v matici A podmatice řádu k , jejíž determinat je nenulový. Nechť tato podmatice je vlevo nahoře. Prvních k sloupců matice A nazveme bázovými. Potom každý sloupec matice $(A|b)$ je lineární kombinací bázových sloupců.

$$\begin{aligned} \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix} &= \lambda_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{pmatrix} + \lambda_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \dots \\ a_{m2} \end{pmatrix} + \dots + \lambda_k \begin{pmatrix} a_{1k} \\ a_{2k} \\ \dots \\ a_{mk} \end{pmatrix} = \\ &= \lambda_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{pmatrix} + \lambda_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \dots \\ a_{m2} \end{pmatrix} + \dots + \lambda_k \begin{pmatrix} a_{1k} \\ a_{2k} \\ \dots \\ a_{mk} \end{pmatrix} + 0 \begin{pmatrix} a_{1,k+1} \\ a_{2,k+1} \\ \dots \\ a_{m,k+1} \end{pmatrix} + \dots + 0 \begin{pmatrix} a_{1n} \\ a_{2n} \\ \dots \\ a_{mn} \end{pmatrix}. \end{aligned}$$

To ovšem znamená, že sloupec $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k, 0, \dots, 0)^T$ je řešením soustavy (2.1). Neboli platí: Jestliže je $h(A) = h(A|b) = k$, potom je soustava (2.1) řešitelná. \square

Důsledek 2.11. *Je-li soustava (2.1) řešitelná, t.j. $h(A) = h(A|b) = h$, pak pro $h = n$ má soustava (2.1) právě jedno řešení a pro $h < n$ má soustava (2.1) nekonečně mnoho řešení, která závisí na $(n - h)$ parametrech.*

Důsledek 2.12. *Je-li soustava (2.1) řešitelná, t.j. $h(A) = h(A|b) = h$, potom nikdy nemůže nastat případ, že $h > n$.*

Příklad 2.13. Řešte soustavu

$$\begin{aligned} x + y + z &= 1 \\ 2x + y + 2z &= 1 \\ x + y + 3z &= 2 \end{aligned}$$

Řešení. Protože $|A| = -2 \neq 0$, jde o soustavu, jejíž řešení můžeme určit pomocí Cramerových vzorců. Dostaneme

$$x = -\frac{1}{2}, \quad y = 1, \quad z = \frac{1}{2}.$$

\square

Příklad 2.14. Řešte soustavu

$$\begin{aligned} x + y + z &= 1 \\ x + y + 2z &= 1 \\ x + y + 3z &= 2 \end{aligned}$$

Řešení. $|A| = 0$, proto nemůžeme použít Cramerových vzorců.

$$h(A) = 2, \quad h(A|b) = 3 \Rightarrow h(A) \neq h(A|b),$$

podle věty 2.9 nemá soustava řešení. \square

Příklad 2.15. Řešte soustavu

$$\begin{aligned}x + y + z &= 1 \\x + y + 2z &= 1 \\2x + 2z + 4z &= 2\end{aligned}$$

Řešení. $|A| = 0$, proto nemůžeme použít Cramerových vzorců.

$$h(A) = 2, h(A|b) = 2 \Rightarrow h(A) = h(A|b),$$

Řešení závisí na jednom parametru.

$$\begin{aligned}x &= 1 - t \\y &= t \\z &= 0.\end{aligned}$$

□

Věta 2.16. Homogenní soustava (2.2) je vždy řešitelná.

Důkaz. Nulový sloupec je vždy řešením.

□

Definice 2.17. Nulové řešení soustavy (2.2) nazveme *triviálním*.

Věta 2.18. Homogenní soustava má netriviální řešení právě tehdy, když hodnota matice koeficientů je menší jak počet neznámých.

Věta 2.19. Necht' u, v jsou řešením soustavy (2.2). Potom i jejich libovolná lineární kombinace $\alpha u + \beta v$ je řešením soustavy (2.2).

Důsledek 2.20. Každá lineární kombinace řešení soustavy (2.2) je opět řešením soustavy (2.2).

Důsledek 2.21. Věta 2.19 mluví jen o dvou řešeních, ale jejich počet není omezen.

Důkaz se provádí matematickou indukcí.

Definice 2.22. Maximální počet lineárně nezávislých řešení soustavy (2.2) nazveme *fundamentální soustavou řešení soustavy (2.2)*.

Věta 2.23. Každá víceznačně řešitelná soustava (2.2) má vždy fundamentální soustavu řešení.

Příklad 2.24. Řešte homogenní soustavu rovnic

$$\begin{aligned}3x_1 + 2x_2 + 5x_3 + 2x_4 + 7x_5 &= 0 \\6x_1 + 4x_2 + 7x_3 + 4x_4 + 5x_5 &= 0 \\3x_1 + 2x_2 - x_3 + 2x_4 - 11x_5 &= 0 \\6x_1 + 4x_2 + x_3 + 4x_4 - 13x_5 &= 0\end{aligned}$$

Řešení. Koeficienty soustavy si zapíšeme do matice a pomocí elementárních řádkových úprav si matici převedeme na stupňovitý tvar.

$$\begin{aligned} \begin{pmatrix} 3 & 2 & 5 & 2 & 7 \\ 6 & 4 & 7 & 4 & 5 \\ 3 & 2 & -1 & 2 & -11 \\ 6 & 4 & 1 & 4 & -13 \end{pmatrix} &\sim \begin{pmatrix} 3 & 2 & 5 & 2 & 7 \\ 0 & 0 & -3 & 0 & -9 \\ 0 & 0 & -6 & 0 & -18 \\ 0 & 0 & -9 & 0 & -27 \end{pmatrix} \sim \\ &\sim \begin{pmatrix} 3 & 2 & 0 & 2 & -8 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & \frac{2}{3} & 0 & \frac{2}{3} & \frac{-8}{3} \\ 0 & 0 & 1 & 0 & 3 \end{pmatrix}. \end{aligned}$$

Máme dvě rovnice o pěti neznámých. Volíme si proto tři parametry. Zvolme $x_2 = 3s$, $x_4 = 3t$, $x_5 = 3u$, kde $s, t, u \in \mathbb{R}$. Potom

$$x = \begin{pmatrix} -2s - 2t + 8u \\ 3s \\ -9u \\ 3t \\ 3u \end{pmatrix} = s \begin{pmatrix} -2 \\ 3 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} -2 \\ 0 \\ 0 \\ 3 \\ 0 \end{pmatrix} + u \begin{pmatrix} 8 \\ 0 \\ -9 \\ 0 \\ 3 \end{pmatrix}.$$

Trojice vektorů vpravo pak představuje fundamentální soustavu řešení. □

Věta 2.25. *Nechť p, q jsou řešení soustavy (2.1). Potom $(p - q)$ je řešením přidružené homogenní soustavy.*

Důsledek 2.26. *Součet parciálního řešení soustavy (2.1) a řešení přidružené homogenní soustavy je řešením soustavy (2.1).*

Důsledek 2.27. *Všetchna řešení soustavy (2.1) získáme jako součet jednoho (parciálního) řešení soustavy (2.1) a fundamentální soustavy řešení přidružené homogenní soustavy.*

2.2 Gaussova eliminační metoda

Definice 2.28. Dvě řešitelné soustavy lineárních rovnic se nazývají *ekvivalentní*, jestliže mají stejnou množinu řešení.

Dvě ekvivalentní soustavy mohou mít různý počet rovnic, ale musí mít stejný počet neznámých.

Mějme dvě takové soustavy

$$Ax = b, \quad A'x = b'.$$

Potom z podmínek řešitelnosti plyne, že

$$h(A) = h(A|b) = h(A') = h(A'|b').$$

Protože soustavy mají stejnou množinu řešení, tak platí:

$$A\alpha = b \Leftrightarrow A'\alpha = b'.$$

Potom konečným počtem řádkových elementárních úprav lze matici $(A|b)$ převést na matici $(A'|b')$.

Pozor: zde nelze zaměňovat řádkové a sloupcové úpravy.

Můžeme používat pouze řádkové úpravy a ze sloupcových pouze výměnu sloupců v matici A , což je vlastně přeznačení proměnných.

Pomocí elementárních úprav si upravíme soustavu $Ax = b$ na tvar

$$\begin{aligned} c_{1,1}y_1 + c_{1,2}y_2 + \cdots + c_{1,h}y_h + c_{1,h+1}y_{h+1} + \cdots + c_{1,n}y_n &= d_1 \\ c_{2,2}y_2 + \cdots + c_{2,h}y_h + c_{2,h+1}y_{h+1} + \cdots + c_{2,n}y_n &= d_2 \\ &\dots\dots\dots \dots \\ c_{h,h}y_h + c_{h,h+1}y_{h+1} + \cdots + c_{h,n}y_n &= d_h \end{aligned}$$

kde (y_1, y_2, \dots, y_n) je vhodná permutace proměnných (x_1, x_2, \dots, x_n) .

Je-li $h = n$, má soustava právě jedno řešení — jde o kramerovskou soustavu, tj. řešení můžeme získat pomocí Cramerových vzorců.

Je-li $h < n$, potom proměnné y_{h+1}, \dots, y_n prohlásíme za parametry a soustavu upravíme na tvar

$$\begin{aligned} c_{1,1}y_1 + c_{1,2}y_2 + \cdots + c_{1,h}y_h &= d_1 - c_{1,h+1}y_{h+1} - \cdots - c_{1,n}y_n \\ c_{2,2}y_2 + \cdots + c_{2,h}y_h &= d_2 - c_{2,h+1}y_{h+1} - \cdots - c_{2,n}y_n \\ &\dots\dots\dots \dots \\ c_{h,h}y_h &= d_h - c_{h,h+1}y_{h+1} - \cdots - c_{h,n}y_n \end{aligned}$$

Tato soustava je ekvivalentní s původní soustavou $Ax = b$ a každé volbě parametrů y_{h+1}, \dots, y_n odpovídá právě jedno řešení. Parametrů je celkem $(n - h)$. Jestliže za prvky y_{h+1}, \dots, y_n bereme sloupce regulární matice řádu $(n - h)$, potom bereme za parametry lineárně nezávislé prvky a obdržíme obecné řešení soustavy (2.1).

Tento postup se nazývá Gaussova ¹ eliminační metoda.

Jestliže budeme dále pokračovat v řádkových úpravách, můžeme soustavu (2.3) upravit na tvar

$$y_1 + 0y_2 + \cdots + 0y_h = g_1 - f_{1,h+1}y_{h+1} - \cdots - f_{1,n}y_n$$

¹ **K.F.Gauss** (1777 — 1855) německý matematik, fyzik, geofyzik, geodet, astronom. Jeden z nejvýznamnějších matematiků všech dob. Všestranný vědec, který pracoval ve všech oblastech matematiky. Všude dosáhl prvořadých výsledků a předznamenal mnohdy další rozvoj. Byl též velmi zručný numerický matematik, který objevil řadu numerických metod. Jako první dospěl k principům neeukleidovské geometrie, ale výsledky v této oblasti nechtěl pro jejich převratnost publikovat, proto patří priorita objevu N.I.Lobačevskému. V algebře jako první dokázal *Základní větu algebry*. Rozvinul teorii kvadratických forem, zavedl přesně komplexní čísla, rozvinul metody řešení soustav algebraických rovnic.

$$\begin{aligned}
y_2 + \cdots + 0y_h &= g_2 - f_{2,h+1}y_{h+1} - \cdots - f_{2,n}y_n \\
\cdots &\cdots \\
y_h &= g_h - f_{h,h+1}y_{h+1} - \cdots - f_{h,n}y_n
\end{aligned}$$

a nebo po vynechání nulových prvků

$$\begin{aligned}
y_1 &= g_1 - f_{1,h+1}y_{h+1} - \cdots - f_{1,n}y_n \\
y_2 &= g_2 - f_{2,h+1}y_{h+1} - \cdots - f_{2,n}y_n \\
\cdots &\cdots \\
y_h &= g_h - f_{h,h+1}y_{h+1} - \cdots - f_{h,n}y_n
\end{aligned}$$

zde máme na hlavní diagonále vlevo všechny koeficienty u neznámých rovny jedné a zbývající prvky nalevo jsou nulové. Tento postup se nazývá Jordanova ¹ eliminace.

2.2.1 Úplný a částečný výběr hlavního prvku

Při Gaussově eliminační metodě si postupně upravujeme soustavu $Ax = b$ na ekvivalentní soustavu $A^{(k)}x = b^{(k)}$, kde $A^{(k)}, b^{(k)}$ jsou matice koeficientů a sloupec pravých stran po k -tém kroku. Postup je následující:

První krok:

Nechť je prvek $a_{11} \neq 0$ (v opačném případě provedeme přehození řádků), prvky prvního řádku vynásobíme multiplikátorem $(-a_{k1}/a_{11})$ a přičteme k prvkům k -tého řádku, pro $k = 2, \dots, n$. Tím získáme v prvním sloupci na prvním místě nenulový prvek a zbývající prvky jsou nulové.

Druhý krok:

Nechť je prvek $a_{22} \neq 0$, pokud tomu tak není provedeme přehození řádků (vyjma prvního) a nebo sloupců (opět vyjma prvního). Prvky druhého řádku vynásobíme multiplikátorem $(-a_{k2}/a_{22})$ a přičteme k prvkům k -tého řádku, pro $k = 3, \dots, n$. Tím získáme v druhém sloupci na druhém místě nenulový prvek a zbývající prvky jsou nulové.

Pokračujeme dále stejným způsobem. Obecně

$$A = \left(a_{ij}^{(0)} \right), b = \left(a_{1,n+1}^{(0)}, a_{2,n+1}^{(0)}, \dots, a_{n,n+1}^{(0)} \right)^T$$

a pro $k = 1, 2, \dots, n - 1$

$$m_{ik} = -\frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}, \quad \text{pro } i = k + 1, k + 2, \dots, n,$$

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} + m_{ik}a_{kj}^{(k-1)}, \quad \text{pro } j = k + 1, k + 2, \dots, n, n + 1.$$

Pokud bude u multiplikátoru m_{ik} dělitel $a_{kk}^{(k-1)}$ příliš malý, budou nám narůstat zaokrouhlovací chyby, které velmi brzy znehodnotí celý výsledek. Proto se používá *eliminace s výběrem hlavního*

¹ **Camille Marie Edmond Jordan** (5. 1. 1838 – 21. 1. 1922) francouzský matematik. Do r. 1873 pracoval jako inženýr, pak vyučoval na polytechnice. Zabýval se algebrou, teorií čísel, teorií funkcí, geometrií, topologií, diferenciálními rovnicemi, teorií míry, a j.

prvku. Prvek $a_{kk}^{(k-1)}$ nazveme hlavním prvkem k -tého kroku eliminace. Abychom minimalizovali vliv zaokrouhlovacích chyb, je vhodné vybírat jako hlavní prvky takové prvky matice A , které mají největší absolutní hodnotu. Potom budeme vždy mít multiplikátor nejvýše roven jedné (v absolutní hodnotě) a každá dílčí zaokrouhlovací chyba se také násobí stejným číslem, t.j. nevětšuje se.

Pokud vybíráme hlavní prvek ze všech prvků, které v daném kroku přicházejí v úvahu, pak mluvíme o *úplném výběru hlavního prvku*. Tato metoda je sice přesnější, ale časově náročná. Proto se často používá *částečný výběr hlavního prvku*, kdy hlavní prvek vybíráme pouze z některých prvků, které v daném kroku přicházejí v úvahu. Nejčastěji se vybírají pouze z daného sloupce, t.j. z prvků $a_{k,k}^{(k-1)}, a_{k+1,k}^{(k-1)}, a_{k+2,k}^{(k-1)}, \dots, a_{n,k}^{(k-1)}$.

2.2.2 Metoda LU-rozkladu

Definice 2.29. Matici $A \in \mathbb{R}_{m,n}$ nazveme *horní trojúhelníkovou* maticí, když $a_{ij} = 0 \forall i > j$. Matici A nazveme *dolní trojúhelníkovou* maticí, když $a_{ij} = 0 \forall i < j$.

Věta 2.30. Mějme soustavu $Ax = b$ s regulární maticí A . Potom existují matice L, U takové, že L je dolní trojúhelníková matice a U je horní trojúhelníková matice a matice $A = LU$. Jestliže si zvolíme prvky na hlavní diagonále jedné z matic L, U , potom je rozklad matice A určen jednoznačně.

Důkaz. Metoda řešení soustavy lineárních rovnic LU-rozkadem spočívá v tom, že si nejdříve určíme matice L a U a potom řešíme dvě soustavy

$$Ly = b,$$

$$Ux = y.$$

V matici L zvolme za diagonální prvky jedničky, t.j. $l_{ii} = 1, i = 1, \dots, n$. Potom z rovnosti $A = LU$ plyne

$$u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik}u_{kj}, \quad i = 1, \dots, j,$$

$$l_{ij} = \frac{1}{u_{jj}} \left(a_{ij} - \sum_{p=1}^{j-1} l_{ip}u_{pj} \right), \quad i = j + 1, j + 2, \dots, n.$$

Postupně počítáme první řádek matice U , potom první sloupec matice L , druhý řádek matice U , druhý sloupec matice L , atd. \square

V případě, že matice koeficientů soustavy je speciálního tvaru, dostáváme varianty metody LU-rozkladu.

Definice 2.31. Matici A nazveme *třídiagonální* maticí, když obsahuje prvky na hlavní diagonále a prvky nad a pod hlavní diagonálou a všechny ostatní prvky jsou nulové.

Matici A nazveme *dvoudiagonální* maticí, když obsahuje prvky na hlavní diagonále a prvky buď nad a nebo pod hlavní diagonálou a všechny ostatní prvky jsou nulové.

Příklad 2.32. LU-rozkladem řešte soustavu $Ax = b$, kde matice A je třídiagonální

$$A = \begin{pmatrix} a_1 & c_2 & 0 & 0 & \dots & 0 & 0 \\ b_2 & a_2 & c_2 & 0 & \dots & 0 & 0 \\ 0 & b_3 & a_3 & c_3 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \dots & a_{n-1} & c_{n-1} \\ 0 & 0 & 0 & 0 & \dots & b_n & a_n \end{pmatrix}.$$

Řešení. LU rozkladem matice A získáme dvoudiagonální matice L a U , kde

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ \beta_2 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \beta_3 & 1 & 0 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & 0 & \dots & \beta_n & 1 \end{pmatrix}$$

a

$$U = \begin{pmatrix} \alpha_1 & c_2 & 0 & 0 & \dots & 0 & 0 \\ 0 & \alpha_2 & c_2 & 0 & \dots & 0 & 0 \\ 0 & 0 & \alpha_3 & c_3 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \dots & \alpha_{n-1} & c_{n-1} \\ 0 & 0 & 0 & 0 & \dots & 0 & \alpha_n \end{pmatrix}.$$

kde koeficienty $\alpha_i, \beta_j, i, j = 1, 2, \dots, n$, určíme podle vztahů

$$\alpha_1 = a_1,$$

$$\beta_i = \frac{b_i}{\alpha_{i-1}}, \quad \alpha_i = a_i - \beta_i c_{i-1}, \quad i = 2, 3, \dots, n.$$

Tento postup se označuje jako *Thomasův algoritmus*.

Řešení naší soustavy je potom určeno dvojicí dvoudiagonálních soustav

$$\begin{aligned} Ly &= b, \\ Ux &= y. \end{aligned}$$

Postupným dosazováním dostaneme

$$\begin{aligned} Ly = b &\Rightarrow y_1 = b_1, \\ y_i &= b_i - \beta_i y_{i-1}, \quad i = 2, 3, \dots, n \end{aligned}$$

a nakonec

$$\begin{aligned} Ux = y &\Rightarrow x_n = \frac{y_n}{\alpha_n} \\ x_i &= \frac{y_i - c_i x_{i-1}}{\alpha_i}, \quad i = n-1, n-2, \dots, 1. \end{aligned}$$

□

Definice 2.33. Matice A se nazývá *pásová*, jestliže existují taková přirozená čísla p, q , že

$$a_{ij} = 0 \quad \text{když} \quad j > i + p \quad \text{nebo} \quad i > j + q.$$

Číslo $p + q + 1$ se nazývá šířkou pásu.

V případě, že matice A je pásová, potom si můžeme výrazně zkrátit výpočet, pokud budeme počítat pouze s prvky uvnitř pásu.

Příklad 2.34. Matice $A = \begin{pmatrix} 3 & 1 & 0 & 0 \\ -5 & 1 & 3 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 3 & -3 \end{pmatrix}$ je pásovou maticí a současně jde i o

třídiagonální maticí.

Příklad 2.35. Matice $A = \begin{pmatrix} 3 & 1 & 2 & -1 & 0 & 0 & 0 \\ -5 & 1 & 3 & 1 & 4 & 0 & 0 \\ 0 & 1 & 1 & -1 & 5 & 1 & 0 \\ 0 & 0 & 3 & -3 & 9 & -7 & 0 \\ 0 & 0 & 0 & 1 & 6 & 3 & 1 \end{pmatrix}$ je pásovou maticí. Všimněte

si, že uvnitř „pásu“ se mohou vyskytovat nulové prvky. V našem příkladu je nula na pozici (4,7).

Definice 2.36. Matice $A = (a_{ij})$ se nazývá *symetrická*, jestliže platí

$$A^T = A, \quad \text{t.j.} \quad a_{ij} = a_{ji}.$$

Definice 2.37. Symetrická matice $A = (a_{ij})$ je *pozitivně definitní*, jestliže pro libovolný nenulový vektor $x = (x_1, x_2, \dots, x_n)^T$ platí

$$x^T A x = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j > 0.$$

Věta 2.38. Je-li matice A symetrická a pozitivně definitní, potom existuje právě jedna horní trojúhelníková matice U s kladnými diagonálními prvky, že platí

$$A = U^T U.$$

Důkaz. Postup odvození je analogický jako v předešlém případě. Dostaneme

$$u_{jj} = \sqrt{\left(a_{jj} - \sum_{k=1}^{j-1} u_{jk}^2 \right)}, \quad \text{pro } j = 1, 2, \dots, n,$$

$$u_{ij} = \frac{1}{u_{jj}} \left(a_{ij} - \sum_{p=1}^{j-1} u_{ip} u_{pj} \right) \quad \text{pro } i = j + 1, j + 2, \dots, n.$$

□

Příklad 2.39. Určete LU-rozklad matice $A = \begin{pmatrix} 3 & 1 & -1 & -2 \\ -5 & 1 & 3 & -4 \\ 2 & 0 & 1 & -1 \\ 1 & -5 & 3 & -3 \end{pmatrix}$.

Řešení.

$$A = \begin{pmatrix} 3 & 1 & -1 & -2 \\ -5 & 1 & 3 & -4 \\ 2 & 0 & 1 & -1 \\ 1 & -5 & 3 & -3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -5/3 & 1 & 0 & 0 \\ 2/3 & -1/4 & 1 & 0 \\ 1/3 & -2 & 3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 & 1 & -1 & -2 \\ 0 & 8/3 & 4/3 & -22/3 \\ 0 & 0 & 2 & -3/2 \\ 0 & 0 & 0 & -25/2 \end{pmatrix}$$

□

Příklad 2.40. Matici $A = \begin{pmatrix} 4 & 1 & -2 & 4 \\ 1 & 2.5 & 1 & 1.5 \\ -2 & 1 & 3 & 0 \\ 4 & 1.5 & 0 & 7 \end{pmatrix}$ vyjádřete ve tvaru $A = U^T U$.

Řešení.

$$A = \begin{pmatrix} 4 & 1 & -2 & 4 \\ 1 & 2.5 & 1 & 1.5 \\ -2 & 1 & 3 & 0 \\ 4 & 1.5 & 0 & 7 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 1/2 & 3/2 & 0 & 0 \\ -1 & 1 & 1 & 0 \\ 2 & 1/3 & 5/3 & 1/3 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1/2 & -1 & 2 \\ 0 & 3/2 & 1 & 1/3 \\ 0 & 0 & 1 & 5/3 \\ 0 & 0 & 0 & 1/3 \end{pmatrix}$$

□

2.2.3 Řešení pomocí inverzní matice

Jestliže je matice koeficientů soustavy $Ax = b$ regulární, potom existuje inverzní matice A^{-1} a můžeme použít postup

$$Ax = b,$$

$$A^{-1}Ax = A^{-1}b,$$

$$x = A^{-1}b.$$

Podmínkou je, že jsme schopni efektivně určit inverzní matici A^{-1} . Zvláště pro matice vyšších řádů jde o obtížnou úlohu.

2.3 Iterační metody řešení

Předpokládejme, že soustava (2.1) je řešitelná. Upravíme si ji na tvar

$$x = Cx + d. \tag{2.3}$$

Nechť $x^1 = (x_1^1, x_2^1, \dots, x_n^1)^T$ je libovolný vektor z \mathbb{R}^n . Definujme si posloupnost

$$x^{k+1} = Cx^k + d, \quad k = 1, 2, \dots,$$

neboli

$$x_i^{k+1} = \sum_{j=1}^n c_{ij}x_j^k + d_i, \quad i = 1, 2, \dots, n, k = 1, 2, \dots$$

Tyto vztahy nám definují prostý iterační proces. V případě, že řešení závisí na parametrech, tj. v případě víceznačně řešitelné soustavy, je třeba zvolit konkrétní hodnoty parametrů a pro ně potom hledat řešení. Je důležité určit podmínky konvergence, které nám zaručí existenci limity iteračního procesu a tím i existenci řešení.

Věta 2.41. Označme \mathcal{V}_n vektorový prostor dimenze n . Nechť je ve \mathcal{V}_n dána vektorová norma a s ní souhlasná maticová norma. Nechť matice C soustavy (2.3) splňuje podmínku $\|C\| < 1$. Potom

1. Soustava (2.3) (a tedy i soustava (2.1)) má právě jedno řešení \tilde{x} .
2. Iterační proces $x^{k+1} = Cx^k + d$, $k = 1, 2, \dots$ konverguje k \tilde{x} a navíc x^1 může být libovolný prvek z \mathcal{V}_n .
3. Platí odhady (pro $k = 1, 2, \dots$)

$$\|x^k - \tilde{x}\| \leq \frac{\|C\|}{1 - \|C\|} \|x^k - x^{k-1}\|,$$

$$\|x^k - \tilde{x}\| \leq \frac{\|C\|^{k-1}}{1 - \|C\|} \|x^2 - x^1\|.$$

Důkaz. Matice C je čtvercová. Norma určuje ve \mathcal{V}_n metriku $d(x, y) = \|x - y\|$. Vzhledem k této metrice je prostor \mathcal{V}_n úplný. (Plyne bezprostředně z definice.) Jestliže si definujeme zobrazení $\varphi: \mathcal{V}_n \rightarrow \mathcal{V}_n$, $\varphi(x) = Cx + d$, pak máme $\forall u, v \in \mathcal{V}_n$

$$\begin{aligned} d(\varphi(u), \varphi(v)) &= d(Cu + d, Cv + d) = \|Cu + d - Cv - d\| = \\ &= \|C(u - v)\| \leq \|C\| \cdot \|u - v\| = \|C\| \cdot d(u, v). \end{aligned}$$

Přitom jsme užili souhlasnosti obou norem a protože podle předpokladů věty je $\|C\| < 1$, máme, že zobrazení φ je kontrakce s koeficientem $\|C\|$. Použitím Banachovy věty o pevném bodě dostáváme zbytek důkazu. \square

Označme

$$\alpha_1 = \max_i \sum_{j=1}^n |c_{ij}|, \quad (\text{řádková norma})$$

$$\alpha_2 = \max_j \sum_{i=1}^n |c_{ij}|, \quad (\text{sloupcová norma})$$

$$\alpha_3 = \left(\sum_{i=1}^n \sum_{j=1}^n c_{ij}^2 \right)^{\frac{1}{2}}. \quad (\text{Eukleidovská norma})$$

Věta 2.42. Je-li některá z hodnot $\alpha_1, \alpha_2, \alpha_3$ menší než 1, pak

1. Soustava $x = Cx + d$ (a tedy i soustava $Ax = b$) má právě jedno řešení \tilde{x} .
2. Posloupnost iterací $x^{k+1} = Cx^k + d$ konverguje k tomuto řešení pro libovolnou počáteční aproximaci.

Důkaz. Zřejmý, jde o aplikaci předchozí věty pro konkrétní tvar maticové normy. □

Poznámka 2.43. Při konkrétním výpočtu je třeba vždy pracovat se souhlasnými normami.

Příklad 2.44. Mějme iterační proces s maticí $C = \begin{pmatrix} 0.3 & 0.6 \\ 0.8 & 0.1 \end{pmatrix}$. Rozhodněte o konvergenci procesu.

$$\begin{aligned} \alpha_1 &= 0.9 \\ \text{\textbf{Řešení.}} \text{ Máme } \alpha_2 &= 1.1 \\ \alpha_3 &= \sqrt{1.1} = 1.0488\dots \end{aligned}$$

Protože $\alpha_1 < 1$, bude iterační proces s touto maticí konvergovat. □

V praxi se nejčastěji používají následující metody s přesným algoritmem pro vytvoření iteračního procesu. Vždy přitom předpokládáme, že soustava (2.1) je jednoznačně řešitelná, neboli že matice A je regulární.

2.3.1 Jacobiho iterační metoda

U Jacobiho ¹ iterační metody se řechod od soustavy (2.1) k soustavě (2.3) provádí následovně: Předpokládáme, že $a_{ii} \neq 0$, potom prvky na hlavní diagonále soustavy $Ax = b$ ponecháme na místě a zbývající členy převedeme na pravou stranu, poté vydělíme koeficienty u neznámých na hlavní diagonále. Dostaneme

$$c_{ii} = 0, \quad c_{ij} = -\frac{a_{ij}}{a_{ii}}, \quad \text{pro } i \neq j, \quad d_i = \frac{b_i}{a_{ii}},$$

iterační vztahy mají tvar pro $k = 1, 2, \dots$

$$x_i^{k+1} = \sum_{j=1}^n c_{ij} x_j^k + d_i.$$

Jestliže $\exists i : a_{ii} = 0$, pak provedeme přehození pořadí rovnic tak, aby byla naše podmínka splněna. Vzhledem k regularitě matice A to lze vždy provést. (V opačném případě bychom totiž měli $a_{ij} = 0 \forall j \in \{1, \dots, n\}$, neboli získali bychom matici s nulovým sloupcem, tj. singulární matici, což je spor s předpokládanou regularitou matice A .)

¹**C. G. J. Jacobi** (1804 – 1851) německý matematik. Zabýval se teorií čísel, teorií funkcí, diferenciálními rovnicemi, algebrou.

Definice 2.45. Čtvercová matice A se nazývá *diagonálně dominantní*, jestliže platí

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

Věta 2.46. *Je-li matice A diagonálně dominantní, pak je regulární.*

Věta 2.47. *Je-li matice A diagonálně dominantní, pak Jacobiho iterační metoda konverguje pro libovolnou volbu počáteční aproximace.*

Důkaz.

$$\sum_{j=1}^n |c_{ij}| = c_{ii} + \sum_{j=1, j \neq i}^n c_{ij} = 0 + \sum_{j=1, j \neq i}^n \left| \frac{a_{ij}}{a_{ii}} \right| = \frac{1}{|a_{ii}|} \sum_{j=1, j \neq i}^n |a_{ij}| < 1$$

a tedy $\alpha_1 = \max_i \sum_{j=1}^n |c_{ij}| < 1$. Podle věty 4.49 dostáváme naše tvrzení. \square

2.3.2 Gaussova – Seidelova iterační metoda

U Gaussovy – Seidelovy ¹ metody se přechod od soustavy (2.1) k soustavě (2.3) provádí stejně jako u Jacobiho metody, tj. opět předpokládáme, že $a_{ii} \neq 0$, potom

$$c_{ii} = 0, \quad c_{ij} = -\frac{a_{ij}}{a_{ii}}, \quad i \neq j, \quad d_i = \frac{b_i}{a_{ii}},$$

ale iterační vztahy mají tvar

$$x_i^{k+1} = \sum_{j=1}^{i-1} c_{ij} x_j^{k+1} + \sum_{j=i+1}^n c_{ij} x_j^k + d_i.$$

Pro $a_{ii} = 0$ opět zaměníme pořadí rovnic.

Věta 2.48. *Platí:*

1. *Je-li matice A diagonálně dominantní, pak Gaussova – Seidelova metoda konverguje pro libovolný počátek.*
2. *Je-li některá z hodnot $\alpha_1, \alpha_2, \alpha_3$ menší než 1, pak Gaussova – Seidelova metoda konverguje pro libovolný počátek.*
3. *Je-li matice A symetrická a pozitivně definitní, pak Gaussova – Seidelova metoda konverguje pro libovolný počátek.*

¹**K. F. Gauss** (1777 – 1855) německý matematik. Poslední z matematiků, který pracoval prakticky ve všech částech matematiky.

P. L. Seidel – (1821 – 1896) německý matematik, zabýval se hlavně analýzou. V r.1874 navrhl iterační metodu řešení soustav algebraických rovnic

Poznámka 2.49. Všechny uvedené podmínky konvergence jsou postačující. Nutnou a postačující podmínkou konvergence iteračního procesu $x^{k+1} = Cx^k + d$ je $\rho(C) < 1$, kde ρ je spektrální poloměr matice C , viz definice 1.35.

Poznámka 2.50. Máme-li soustavu $Ax = b$, kde A je regulární matice, která však nesplňuje výše uvedené podmínky konvergence, pak vynásobením A^T zleva dostaneme soustavu

$$A^T Ax = A^T b,$$

kde matice koeficientů ($A^T A$) je symetrická a pozitivně definitní. Pro takovou soustavu Gaussova – Seidelova metoda již konverguje. Obecně však dosti pomalu. Dále může nastat situace, že zaokrouhlovací chyby zcela eliminují posun k přesnému řešení a tak i když teoreticky se máme přibližovat k přesnému řešení, prakticky k tomu nemusí dojít.

Obě metody (Jacobiho i Gaussovu-Seidelovu) si můžeme zapsat s využitím prvků původní matice A . Dostaneme tak předpis pro Jacobiho metodu ve tvaru

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

a pro Gaussovu-Seidelovu metodu ve tvaru

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

Gaussova – Seidelova iterační metoda konverguje většinou rychleji než Jacobiho metoda. Existují však výjimky.

Příklad 2.51. Jacobiho i Gaussovou-Seidelovou metodou řešte soustavu.

$$\begin{array}{rcl} 10x_1 & -2x_2 & -2x_3 = 6 \\ -x_1 & +10x_2 & -2x_3 = 7 \\ -x_1 & -x_2 & +10x_3 = 8 \end{array}$$

Porovnejte rychlost konvergence.

Řešení. Přesné řešení naší rovnice je $x_1 = x_2 = x_3 = 1$. Jestliže zvolíme nulovou počáteční aproximaci, dostaneme pro Jacobiho metodu posloupnost iterací

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \sim \begin{pmatrix} 0,6 \\ 0,7 \\ 0,8 \end{pmatrix} \sim \begin{pmatrix} 0,9 \\ 0,92 \\ 0,93 \end{pmatrix} \sim \begin{pmatrix} 0,970 \\ 0,976 \\ 0,986 \end{pmatrix} \sim \begin{pmatrix} 0,9918 \\ 0,9931 \\ 0,9958 \end{pmatrix} \sim \dots$$

Pro Gaussovu-Seidelovu metodu dostaneme

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \sim \begin{pmatrix} 0,6 \\ 0,76 \\ 0,936 \end{pmatrix} \sim \begin{pmatrix} 0,9392 \\ 0,98112 \\ 0,99203 \end{pmatrix} \sim \begin{pmatrix} 0,994630 \\ 0,997869 \\ 0,9992499 \end{pmatrix} \sim \dots$$

Větší rychlost konvergence u Gaussovy-Seidelovy metody je zřejmá. \square

Příklad 2.52. Jacobiho i Gaussovou-Seidelovou metodou řešte soustavu:

$$\begin{array}{rcl} x_1 & +x_2 & = 1 \\ 2(1-\varepsilon)x_1 & +x_2 & +x_3 = 2 \\ & & x_3 +x_4 = -1 \\ -(1-\varepsilon)^2x_1 & & +x_4 = 5 \end{array}$$

kde $0 < \varepsilon < 0,1$.

Řešení. Jacobiho metoda konverguje, Gaussova-Seidelova diverguje. □

Příklad 2.53. Jacobiho i Gaussovou-Seidelovou metodou řešte soustavu:

$$\begin{array}{rcl} 3x_1 & +2x_2 & +2x_3 = 1 \\ 2x_1 & +3x_2 & +2x_3 = 0 \\ 2x_1 & +2x_2 & +3x_3 = -1 \end{array}$$

Řešení. Jacobiho metoda diverguje, Gaussova-Seidelova konverguje. □

Příklad 2.54. Jacobiho i Gaussovou-Seidelovou metodou řešte soustavu:

$$\begin{array}{rcl} 2x_2 & +4x_3 & = 0 \\ x_1 & -x_2 & -x_3 = 0,375 \\ x_1 & -x_2 & +2x_3 = 0 \end{array}$$

Řešení. Obě metody divergují. □

2.4 Relaxační metody

Snaha o zrychlení konvergence iteračních metod vedla k vytvoření relaxačních metod. Do iteračního procesu je dodán další člen, který může, při správné volbě, výrazně zrychlit konvergenci procesu.

Relaxační metoda pro Jacobiho metodu má tvar

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

Relaxační metoda pro Gaussovou-Seidelovou metodu má tvar

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

Prvek $\omega > 0$ se nazývá relaxační parametr. Volíme jej tak, aby jsme urychlili konvergenci základní metody. Volbou $\omega = 1$ dostaneme původní metodu.

Efektivní volba relaxačního parametru ω závisí na zvolené základní metodě a na tvaru matice A .

Ve speciálních případech je možné relaxační parametr vypočítat, jinak záleží hlavně na zkušenosti a dobrém odhadu.

Věta 2.55. *Nechť A je třídiagonální, symetrická, pozitivně definitní matice. Potom pro spektrální poloměr platí $\varrho^2(C) < 1$ a optimální hodnotu relaxačního parametru můžeme určit podle vztahu*

$$\omega_{\text{optimum}} = \frac{2}{1 + \sqrt{1 - \varrho^2(C)}}$$

kde C je iterační matice pro Jacobiho metodu.

Příklad 2.56. Mějme danu soustavu

$$\begin{aligned} 2x_1 - x_2 &= 1, \\ -x_1 + 2x_2 - x_3 &= 0, \\ -x_2 + 2x_3 &= 1 \end{aligned}$$

Najděte pro ni optimální relaxační parametr.

Řešení. Jacobiho iterační matice má tvar

$$\begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{pmatrix}$$

Vlastní čísla této matice jsou $\lambda_1 = 0$, $\lambda_2 = \frac{\sqrt{2}}{2}$, $\lambda_3 = -\frac{\sqrt{2}}{2}$.

$$\varrho(C) = \frac{\sqrt{2}}{2}$$

a proto

$$\omega_{\text{optimum}} = \frac{2}{1 + \sqrt{1 - \frac{1}{2}}} \approx 1,172.$$

Pokud vezmeme $\omega = 1,17$, potom například pro dosažení přesnosti 10^{-3} musíme provést 5 iterací, zatímco při použití Gaussovy-Seidelovy iterační metody potřebujeme pro dosažení stejné přesnosti 10 iterací. \square

Poznámka 2.57. Jestliže v iteračním procesu

$$x^{k+1} = Cx^k + d$$

nastává cyklus, potom $\varrho(C) = 1$.

Opak však neplatí. Je-li $\varrho(C) = 1$ cyklus v iteračním procesu ještě nastat nemusí. Metoda může divergovat nebo konvergovat. Chování metody závisí na vlivu zaokrouhlovacích chyb.

Jestliže vznikne cyklus u soustavy dvou rovnic, potom se hodnoty řešení pohybují po elipse se středem v přesném řešení. Iterační proces je nekonečný a nedosáhne nikdy přesného řešení.

Podobně pro soustavy vyšších řádů se budou hodnoty řešení pohybovat po polochách druhého řádu.

2.4.1 Superrelaxační metoda

Pokud je v relaxační metodě

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

parametr $\omega > 1$, potom mluvíme o *superrelaxační metodě* a označujeme ji SOR.

Porovnáním SOR a Gaussovy-Seidelovy metody dostaneme tvar

$$x_i^{k+1} = x_i^k + \omega (GS(x_i^k) - x_i^k),$$

kde $GS(x_i^k)$ označuje iteraci získanou pomocí Gaussovy-Seidelovy metody. Opět pro $\omega = 1$ dostaneme původní Gaussovu-Seidelovu metodu.

Konvergence metody SOR je velmi citlivá na správnou volbu parametru ω .

SOR můžeme přepsat do tvaru

$$x^{k+1} = Bx^k + c,$$

kde B je iterační matice a $c = (I - B)A^{-1}b$. Nutnou a postačující podmínkou konvergence je opět $\rho(B) < 1$.

2.4.2 Metoda největšího spádu

Patří mezi gradientní metody. Hledáme při ní nejrychlejší změnu rezidua $r = Ax - b$. Její algoritmus je následující: pro $k = 0, 1, 2, \dots$

$$Ax^k - b = r^k,$$

$$x^{k+1} = x^k + \frac{(r^k)^T r^k}{(r^k)^T A r^k},$$

$$r^{k+1} = r^k - \alpha^k A r^k,$$

$$\alpha^k = \frac{(r^k)^T r^k}{(r^k)^T A r^k}.$$

Potom platí $(r^{k+1})^T r^k = 0$ pro všechna k .

Věta 2.58. *Nechť A je symetrická, pozitivně definitní matice. Potom metoda největšího spádu konverguje pro libovolný počátek.*

2.4.3 Metoda sdružených gradientů

Označujeme ji MSG (method of conjugate gradients). Mějme opět naši soustavu $Ax = b$, kde A je komplexní matice.

Definice 2.59. Řekneme, že matice A^H je *hermitovskky sdružená s maticí* A , jestliže $A^H = (\overline{A})^T$, kde \overline{A} je matice komplexně sdružená s maticí A .

Řekneme, že vektor v^H je *hermitovskky sdružený s vektorem* v , jestliže $v^H = (\overline{v})^T$, kde \overline{v} je vektor komplexně sdružený s vektorem v .

Algoritmus metody MSG je následující:

Mějme soustavu $Ax = b$, kde $A = A^H$. Nechť x^0 je vektor počáteční aproximace takové, že $Ax \neq b$, potom pro $k = 0, 1, 2, \dots$ máme

$$\begin{aligned} p^0 &= r^0 = b - Ax^0, \\ \alpha^k &= \frac{(r^k)^H r^k}{(p^k)^H A p^k}, \\ x^{k+1} &= x^k + \alpha^k p^k, r^{k+1} = r^k - \alpha^k A p^k, \\ \beta^k &= \frac{(r^{k+1})^H r^{k+1}}{(r^k)^H r^k}, \\ p^{k+1} &= r^{k+1} + \beta^k p^k. \end{aligned}$$

Koeficienty α^k, β^k jsou reálné i pro komplexní matice A .

Věta 2.60. *Jestliže matice A je hermitovská pozitivně definitní matice (tj. $A^H = A$ a $x^H A x > 0$ pro všechny nenulové vektory x), potom metoda MSG konverguje.*

Poznámka 2.61. Pro plné matice vyžaduje metoda MSG $2n^3 + O(n^2)$ kroků, kde n je řád matice soustavy. Proto při svém objevení v roce 1952 příliš nezaujala. Pro srovnání, klasická Gaussova eliminační metoda vyžaduje $\frac{2}{3}n^3 + O(n^2)$ kroků. V roce 1969 Srasen publikoval svoji variantu Gaussovy metody, která potřebuje už pouze $O(n^{2,81})$ kroků a tento výsledek byl později ještě zlepšen na hodnotu $O(n^{2,38})$. Další pokusy o zlepšení výsledku pokračují. Exponent zřejmě nebude menší jak 2, protože v plné matici je n^2 prvků a každého z nich se eliminační metoda týká.

Až později se u metody MSG zjistilo, že pokud budeme pracovat s řídkou maticí, potom se její konvergence výrazně zrychluje a počet kroků strmě klesá. (Matice se označuje za řídkou, jestliže obsahuje méně než 5% nenulových prvků)

Na příkladu si ukážeme, že i tato metoda může havarovat.

Příklad 2.62. Mějme homogenní soustavu s reálnou symetrickou maticí

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 4 & -2 \\ 2 & -2 & 0 \end{pmatrix}.$$

Řešení. Potom $b = \mathcal{O}$ a pro matici A platí

$$D_1 = 1, \quad D_2 = \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 0,$$

takže matice A není pozitivně definitní. Zvolme $x^0 = (4, 4, 3)^T$, potom dostaneme

$$p^0 = r^0 = (-18, -18, 0)^T,$$

$$x^1 = (0, 0, 3)^T,$$

$$r^1 = (-6, 6, 0)^T,$$

$$p^1 = (-8, 4, 0)^T,$$

$$(p^1)^T A p_1 = 0,$$

neboli algoritmus selhal. □

Každá hermitovská a pozitivně definitní matice má kladná vlastní čísla $\lambda_n \geq \lambda_{n-1} \geq \dots \geq \lambda_1 > 0$. A to i v případě, že matice A je komplexní. Definujme si číslo podmíněnosti matice A vztahem

$$K(A) = \frac{\max_i \lambda_i}{\min_i \lambda_i} = \frac{\lambda_n}{\lambda_1}.$$

Vždy platí $K(A) \geq 1$. Zavedeme si dále normu (tzv. energetickou normu) předpisem

$$\|x\|_A = \sqrt{x^H A x}.$$

Věta 2.63. *Nechť A je hermitovská pozitivně definitní matice a x^0 je počáteční aproximace přesného řešení x^* soustavy $Ax = b$. Pak pro algoritmus MSG platí odhad*

$$\|x^* - x^k\|_A \leq 2 \left(\frac{\sqrt{K(A)} - 1}{\sqrt{K(A)} + 1} \right) \|x^* - x^0\|_A$$

pro všechna k a kde $K(A)$ je číslo podmíněnosti matice A .

2.5 Stabilita řešení numerické úlohy

Při použití jakékoliv numerické metody vznikají zaokrouhlovací chyby – na počátku řešení při zadávání vstupních dat a během výpočtu při provádění početních operací, které se všechny promítnou do výsledku.

Pokud chceme získat smysluplné výsledky, musíme si vybírat stabilní algoritmy.

Algoritmus je přitom stabilní, jestliže je

1. dobře podmíněný, tj. málo citlivý na změny ve vstupních datech,

2. numericky stabilní, tj. málo citlivý na vliv zaokrouhlovacích chyb, které vznikají během výpočtu.

Jinak řečeno – úloha je stabilní, jestliže drobná změna vstupních hodnot vyvolá jen drobnou změnu ve výsledku.

Jakákoliv numerická úloha obsahující soustavy lineárních algebraických rovnic je obecně vzato nestabilní.

Příklad 2.64. Mějme danu soustavu

$$\begin{aligned} 2x + 6y &= 8, \\ 2x + 6.00001y &= 8.00001, \end{aligned}$$

kteřá má řešení $x = 1$, $y = 1$. Drobnou změnou zadání získáme soustavu

$$\begin{aligned} 2x + 6y &= 8, \\ 2x + 5.99999y &= 8.00002 \end{aligned}$$

kteřá ale má výrazně odlišné řešení $x = 10$, $y = -2$. Změna vstupních hodnot byla řádově 10^{-5} a u výstupních hodnot došlo ke skoku z jedné na deset.

U rozsáhlejších soustav mohou být změny ještě výraznější. Záleží na tvaru matice koeficientů soustavy.

Platí: jestliže mají matice A a matice A^{-1} srovnatelné prvky, potom je úloha $Ax = b$ stabilní. V opačném případě jde o nestabilní úlohu.

V předchozím příkladu máme

$$A = \begin{pmatrix} 2 & 6 \\ 2 & 6.00001 \end{pmatrix}$$

$$A^{-1} = \frac{1}{|A|} \text{adj } A = \frac{1}{2 \cdot 10^{-5}} \begin{pmatrix} 6.00001 & -6 \\ -2 & 2 \end{pmatrix} = \begin{pmatrix} 300000.5 & -300000 \\ -100000 & 100000 \end{pmatrix}.$$

Prvky matice A jsou řádově jednotky, prvky matice inverzní jsou řádově statisíce, máme tedy výrazně nestabilní matici. Pro druhou matici získáme analogický výsledek.

Pojmy k zapamatování

- Zopakovali jsme si metody řešení soustavy lineárních algebraických rovnic.
- Společným rysem všech uváděných metod, t.j. Gaussovy a Jordanovy eliminační metody, použití inverzní matice i LU rozkladu je, že po provedení konečného počtu kroků získáme výsledek, který je konečný, včetně případných chyb. Všechny uvedené metody jsou použitelné pro libovolnou řešitelnou soustavu.
- Naproti tomu iterační metody, se kterými jsme se seznámili později, nám zaručují dosažení výsledku pouze pro soustavy s maticí koeficientů, která splňuje podmínky konvergence = nedají se tedy použít vždy. Pokud je však můžeme použít, tj. pokud jsou splněny podmínky konvergence, potom se přiblížíme k přesnému řešení s libovolnou přesností.

Kontrolní otázky

1. Kdy je matice diagonálně dominantní?
2. Jak poznáme, že je matice pozitivně definitní?
3. Pro matice jakého řádu je vhodná metoda sdružených gradientů?

2.6 Cvičení

Maplety

Příklady pro samostatnou práci jsou uvedeny v samostatné sbírce příkladů. Pomocí následujících mapletů si můžete usnadnit některé dílčí výpočty, nebo zkontrolovat jejich správnost.

1. [Výpočet determinantu matice](#)
2. [Řešení soustav lineárních rovnic přímými metodami](#)
3. [Násobení matic](#)
4. [Vlastní čísla a vektory](#)

webMathematica

Příklady pro samostatnou práci jsou uvedeny v samostatné sbírce příkladů. Pomocí následujících aplikací si můžete usnadnit některé dílčí výpočty, nebo zkontrolovat jejich správnost.

1. [Definitnost matic](#)

3 Řešení rovnic.

Průvodce studiem

Cílem této kapitoly je seznámit čtenáře s numerickými metodami řešení rovnic typu $f(x) = 0$. S vybranými metodami jste se seznámili už v předmětu Matematika 3. Zde budou metody korektně zavedeny a odvozeny.

Začneme seznámením se startovacími metodami - grafickou, tabelovací a metodou půlení intervalu, které nám slouží k separaci kořenů, tj. k tomu, abychom dostatečně zúžili interval, na kterém hledáme řešení.

Potom se budeme věnovat iteračním metodám. Ty nám umožní se přiblížit k hledanému řešení s libovolnou přesností, ale pouze tehdy, když jsou splněny podmínky konvergence pro danou metodu.

Protože budeme požadovat splnění konvergenčních podmínek pouze na nějakém okolí hledaného řešení, bude pro nás vždy výhodné, pokud se dokážeme vhodnou startovací metodou přiblížit k řešení.

V aplikacích se často vyskytují algebraické rovnice, proto se jim budeme věnovat samostatně na konci kapitoly.

Cíle

Po prostudování této kapitoly budete schopni:

- Řešit nelineární a algebraické rovnice.
- Vybrat vhodnou metodu pro daný typ rovnice.
- Stanovit nutný počet kroků pro dosažení požadované přesnosti.

3.1 Základní pojmy

Definice 3.1. Kořenem rovnice $f(x) = 0$ nazveme každé číslo α takové, že $f(\alpha) = 0$.

Uvedeme si některé metody pro řešení obecné rovnice a potom speciální metody pro určení kořenů polynomu. Při řešení rovnic rozeznáváme

a) přímé metody, které využíváme např. pro hledání kořenů kvadratické rovnice,

b) iterační metody – těmi se budeme zabývat.

Vždy nás bude zajímat zda konvergují k správnému řešení, za jakých podmínek a jak rychle.

Věta 3.2. *Nechť $f(x)$ je spojitá na $[a, b]$, $f(a)f(b) < 0$. Potom existuje aspoň jedno $\alpha \in (a, b)$ takové, že $f(\alpha) = 0$.*

Důkaz. Plyne z věty o střední hodnotě pro konkrétní hodnotu $c = 0$. □

Věta 3.3. *Nechť $f(x)$ je spojitá na $[a, b]$, $f(a)f(b) < 0$, $f'(x)$ existuje a nemění znaménko $\forall x \in (a, b)$. Potom je v (a, b) právě jeden kořen rovnice $f(x) = 0$.*

Důkaz. Jde o důsledek předchozí věty pro případ ryze monotónní funkce. □

Příklad 3.4. Mějme danu rovnici $f(x) = x + e^x$. Rozhodněte o tom, zda má rovnice kořeny.

Řešení. Funkce $f(x)$ je spojitá pro všechna $x \in \mathbb{R}$. $f(+\infty) = +\infty$, $f(-\infty) = -\infty$, neboli funkce nabývá opačných hodnot, derivace $f'(x) = 1 + e^x$ je kladná pro všechna $x \in \mathbb{R}$. Máme splněny všechny podmínky věty 3.3 a proto existuje právě jeden reálný kořen naší rovnice. □

3.2 Startovací metody – Separace kořenů

Výpočet kořene rovnice můžeme podstatně urychlit, jestliže umíme odhadnout, kde leží kořen rovnice. Pro přibližné stanovení polohy kořene používáme startovací metody - grafickou, tabelovací a metodou půlení intervalu, které nám slouží k separaci kořenů, tj. k tomu, abychom dostatečně zúžili interval, na kterém hledáme řešení. Přitom vycházíme z toho, že na dostatečně úzkém intervalu se valná většina funkcí chová „rozumně“.

Grafická metoda

1. Nakreslíme si graf funkce $f(x)$. Z něj pak můžeme odečíst potřebné údaje.
2. Rovnici $f(x) = 0$ si upravíme na tvar $g(x) = h(x)$. Umíme-li nakreslit grafy obou těchto funkcí, potom jejich průsečíky určují kořeny.

Příklad 3.5. $x^2 - \cos x = 0$.

VYHODNOCENÍ

1. Velmi nepřesná, ale velmi rychlá metoda,
2. Pokud věnujeme grafu *potřebnou* pozornost, jsme schopni s jeho pomocí provést separaci kořenů a mnohdy i napomoci při určení, zda vůbec má rovnice nějaký reálný kořen.
3. V grafu nejdříve volíme měřítko a až po té do něj zakreslujeme hodnoty funkce.

4. Závěry vždy nutno ověřit výpočtem.

Tabelování funkce Na intervalu $[a, b]$ si libovolně volíme posloupnost $x_0 = a < x_1 < \dots < x_n = b$ a určíme příslušné funkční hodnoty. Jestliže $f(x_k)f(x_{k+1}) < 0$, potom podle věty 3.2 leží v intervalu $[x_k, x_{k+1}]$ aspoň jeden kořen.

Metoda bisekce – půlení intervalu

Nechť $f(x)$ je spojitá na $[a, b]$, $f(a)f(b) < 0$. Sestrojíme si posloupnost intervalů $[a, b] = I_0 \supset I_1 \supset I_2 \supset \dots \supset I_k = [a_k, b_k] \supset \dots$. Je-li $f(a_{k-1})f(b_{k-1}) < 0$, potom $[a_k, b_k]$ bude ten z intervalů $[a_{k-1}, s_k]$, $[s_k, b_{k-1}]$, $s_k = \frac{1}{2}(a_{k-1} + b_{k-1})$ v jehož koncových bodech má funkce f opačná znaménka. Pokud neuděláme chybu – t.j. pokud vyloučíme chybu lidského faktoru, tak bude kořen α ležet v každém z intervalů $[a_k, b_k]$.

Posloupnost $\{a_k\}$ je neklesající a shora omezená libovolným z čísel b_i . Posloupnost $\{b_k\}$ je nerostoucí a zdola omezená kterýmkoliv číslem a_j . To znamená, že každá z těchto posloupností má limitu. Dále platí

$$b_k - a_k = \frac{b - a}{2^k}$$

a tedy

$$\lim_{k \rightarrow \infty} (b_k - a_k) = 0 \Rightarrow \lim_{k \rightarrow \infty} a_k = \lim_{k \rightarrow \infty} b_k = \alpha.$$

Pro odhad chyby máme

$$|a_k - \alpha| \leq \frac{b - a}{2^k}, \quad |b_k - \alpha| \leq \frac{b - a}{2^k}.$$

VYHODNOCENÍ

1. Metoda konverguje pomalu.
2. Pokud v $[a, b]$ leží více kořenů, určíme tímto postupem jen jeden.
3. Je jednoduchá.
4. Prostý odhad chyby.
5. Nezáleží na vlastnostech funkce f .

3.3 Iterační metody

S iteračními metodami jste se setkali už v předmětu Matematika 3. Nyni si provedeme nejen opakování a doplnění, ale hlavně si zde uvedeme přesné znění vět a předpokladů.

Nechť f je spojitá na $[a, b]$, $f(a)f(b) < 0$. Potom podle věty 3.2 existuje aspoň jedno $\alpha \in [a, b]$ takové, že $f(\alpha) = 0$. Sestrojíme si posloupnost $x_1, x_2, \dots, x_k, \dots$ tak, že $\lim_{k \rightarrow \infty} x_k = \alpha$. Výpočet ukončíme buď v souladu s teoretickým odhadem chyby a nebo užitím empirického kritéria:

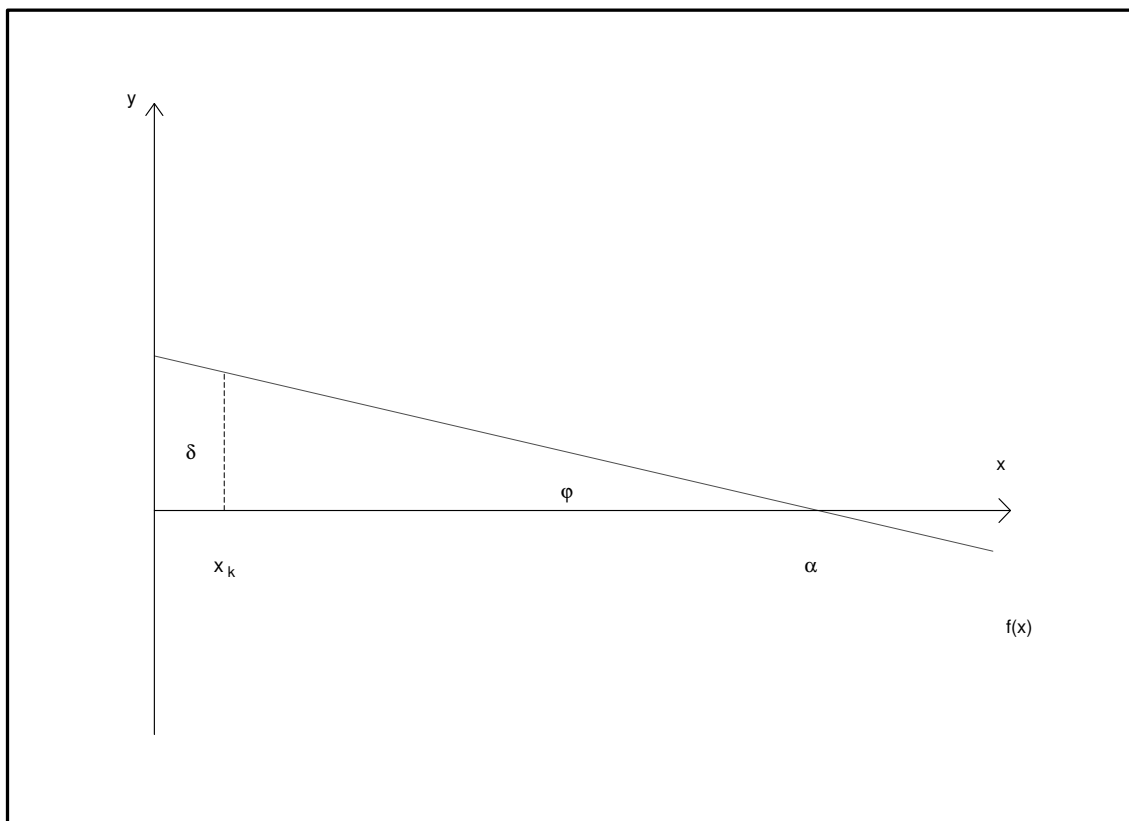
a) Je-li $|x_k - x_{k-1}| < \varepsilon$ pro zadané $\varepsilon > 0$.

Pokud máme pochybnosti, provedeme další šetření:

- Je-li posloupnost $\{x_k\}_{k=1}^{\infty}$ rostoucí a $f(x_k + \varepsilon)f(x_k) < 0$, pak platí $|x_k - \alpha| < \varepsilon$.
- Je-li posloupnost $\{x_k\}_{k=1}^{\infty}$ klesající a $f(x_k - \varepsilon)f(x_k) < 0$, pak platí $|x_k - \alpha| < \varepsilon$.

b) Je-li $|f(x_k)| < \delta$ pro zvolené $\delta > 0$.

Tato podmínka může být velmi ošidná. Viz obrázek 3.1 - jestliže bude úhel φ velmi malý,



Obr. 3.1: Problém malého úhlu φ

potom i pro malé δ je x_n výrazně odlišné od α .

Definice 3.6. Řekneme, že funkce $g(x)$ splňuje na $[a, b]$ Lipschitzovu¹ podmínku s konstantou $k \in [0, 1)$, jestliže

$$\forall x_1, x_2 \in [a, b] : |g(x_1) - g(x_2)| \leq k|x_1 - x_2|.$$

Důsledek 3.7. Jestliže $g(I) \subset I = [a, b]$, potom je g kontrakce.

3.3.1 Metoda prosté iterace

Nechť f je spojitá na $[a, b]$, $f(a)f(b) < 0$. Rovnici $f(x) = 0$ si upravíme na tvar $x = g(x)$. Úprava přitom nemusí být jednoznačná, většinou ji lze provést více způsoby, jak je vidět

¹R. O. S. Lipschitz (1832 – 1903) německý matematik. Zabýval se teorií čísel, diferenciálními rovnicemi, geometrií, teorií řad.

z následujícího příkladu.

Příklad 3.8. Ukázka převodu na tvar $x = g(x)$:

$$\begin{aligned} x^3 - 2x + 7 = 0 &\Rightarrow x = (x^3 + 7)/2, \\ &x = \sqrt[3]{2x - 7}, \\ &x = (2x - 7)/x^2, \\ &x = \sqrt{(2x - 7)/x}. \end{aligned}$$

Věta 3.9. *Nechť $\forall x \in I = [a, b]$ je $g(x) \subset I$, $g(x)$ splňuje na I Lipschitzovu podmínku s konstantou $k \in [0, 1)$. Potom má rovnice $x = g(x)$ v I právě jedno řešení α a posloupnost $\{x_n\}_{n=1}^{\infty}$ definovaná předpisem $x_{n+1} = g(x_n)$ k němu konverguje pro libovolný počátek $x_1 \in I$. Pro odhad chyby platí*

$$|x_n - \alpha| \leq \frac{k^{n-1}}{1 - k} |x_1 - x_2|.$$

Důkaz. Metrika je $d(x, y) = |x - y|$ a vzhledem k ní je (I, d) úplným metrickým prostorem. Jsou tedy splněny všechny podmínky Banachovy věty o pevném bodě. Jako její důsledek dostaneme zbytek důkazu. \square

Důsledek 3.10. *Nechť pro všechna $x \in I$ platí, že*

1. *existuje $g'(x)$,*
2. $\max_{x \in I} |g'(x)| \leq k < 1$.

Pak $g(x)$ splňuje na I Lipschitzovu podmínku.

Důkaz. Plyne z věty o střední hodnotě. \square

Lipschitzova podmínka se obtížně prověřuje. Důsledek 3.10 se prověřuje lépe a rychleji. Provedené zúžení třídy použitelných funkcí není podstatné.

Důsledek 3.11. *Mějme rovnici $f(x) = 0$ upravenou na tvar $x = g(x)$, potom:*

- a) *Jestliže $x = g(x)$, $g(I) \subset I$ a $\forall x \in I : |g'(x)| > 1$, potom přejdeme k inverzní funkci $g^{-1} = h$ a máme $x = h(x)$, $|h'(x)| < 1$.*
- b) *Vždy musí platit $|g'(x)| \leq \underline{\underline{k}} \leq 1$. Derivace $g'(x)$ se nemůže libovolně přibližovat k 1.*
- c) *Jestliže $-1 < g' \leq 0$, pak kořen leží mezi dvěma po sobě jdoucími aproximacemi.*
- d) *Pro odhad dosažené přesnosti platí pro konvergentní posloupnosti*

$$|x_n - \alpha| < \frac{k}{1 - k} |x_n - x_{n-1}|.$$

- e) *Nechť jsou splněny předpoklady věty 3.9. Jestliže pro některé k platí $x_k \notin I$, pak je ve výpočtu chyba a je nutné ho prověřit.*

3.3.2 Metoda regula falsi

Nechť $f(x)$ je spojitá na $[a, b]$, $f(a)f(b) < 0$. Spojíme body $(a, f(a))$, $(b, f(b))$ přímkou. Její průsečík s osou x je bod $(s, 0)$,

$$s = a - \frac{b-a}{f(b)-f(a)}f(a) = b - \frac{b-a}{f(b)-f(a)}f(b).$$

a) Jestliže $f(a)f(s) < 0$, potom položíme $a := a$, $b := s$.

b) Jestliže $f(b)f(s) < 0$, potom položíme $a := s$, $b := b$.

Výpočet opakujeme dokud nedosáhneme požadované přesnosti.

VYHODNOCENÍ:

1. Metoda regula falsi je vždy konvergentní.
2. Může být efektivnější než metoda půlení intervalu.
3. Je jednoduchá. Neklade přílišné nároky na separaci kořenů.
4. Určí vždy jen jeden kořen.
5. Konvergence je však obvykle dosti pomalá.
6. Neplatí odhad $|\alpha - x_n| \leq K |x_n - x_{n-1}|$.

3.3.3 Metoda sečen

Jde o variantu metody regula falsi. Nechť f je spojitá na $[a, b]$, $f(a)f(b) < 0$, $f''(x)$ nemění na $[a, b]$ znaménko, neboli $f(x)$ je na celém intervalu buď konvexní a nebo konkávní, a má tedy na intervalu (a, b) právě jeden kořen. Pak pro

a) $\text{sign } f(a) = \text{sign } f''(x)$ je

$$x_{n+1} = x_n - \frac{x_n - a}{f(x_n) - f(a)}f(x_n), \quad x_1 = b,$$

b) $\text{sign } f(a) \neq \text{sign } f''(x)$ je

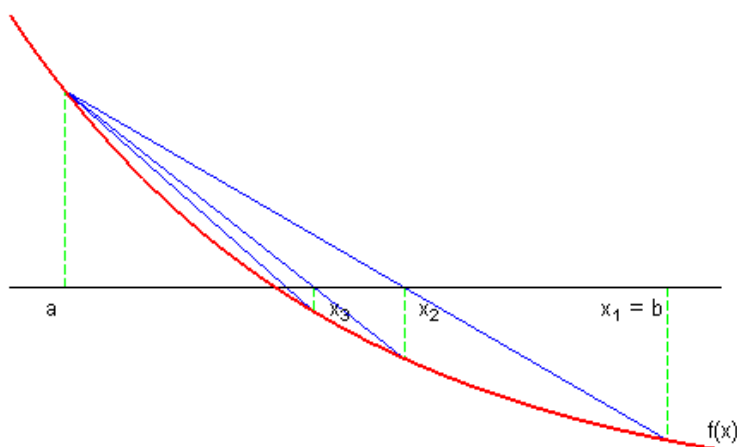
$$x_{n+1} = x_n - \frac{b - x_n}{f(b) - f(x_n)}f(x_n), \quad x_1 = a.$$

V obou případech jde o jednostrannou konvergenci.

Záleží na tom, z kterého bodu začnete provádět iterace.

VYHODNOCENÍ:

1. Metoda sečen je vždy konvergentní.
2. Je efektivnější než metoda půlení intervalu.
3. Konvergence je však obvykle dosti pomalá.
4. Předpokladem konvergence je pouze jeden prostý kořen na intervalu I .
5. Neplatí odhad $|\alpha - x_n| \leq K |x_n - x_{n-1}|$.



Obr. 3.2: Geometrický smysl metody sečen

3.3.4 Metoda tečen

Metoda tečen se často označuje jako Newtonova ¹ metoda.

Nechť rovnice $f(x) = 0$ má jednoduchý reálný kořen α na intervalu $I = [a, b]$. Předpokládejme, že existují na I nenulové derivace $f(x)$. Potom je možno rozvinout f do Taylorovy ² řady v okolí libovolného bodu $x_0 \in I$.

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + \dots$$

Nyní v rovnici $f(x) = 0$ nahradíme funkci $f(x)$ prvními dvěma členy Taylorova rozvoje (t.j. provedeme linearizaci)

$$f(x_0) + f'(x_0)(x - x_0) = 0$$

a určíme kořen x_1 této rovnice

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

V okolí bodu x_1 můžeme zase rovnici $f(x) = 0$ aproximovat lineární částí Taylorova rozvoje

$$f(x_1) + f'(x_1)(x - x_1) = 0$$

s kořenem

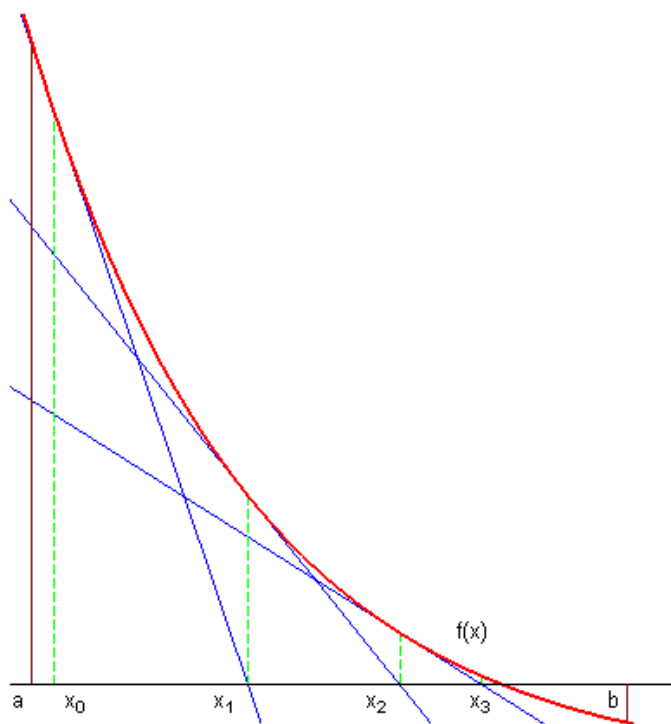
$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)},$$

¹**I. Newton** – (1642 – 1727) anglický matematik, fyzik, astronom, optik a filosof. Jeden z největších světových vědců všech dob. Prakticky současně s Leibnizem a nezávisle na něm vybudoval diferenciální a integrální počet. Významné jsou i jeho práce z algebry, teorie řad, numerické matematiky, a j.

²**B. Taylor** (1685 – 1737) anglický matematik, věnoval se matematické analýze, teorii řad, matematické fyzice – položil základy matematického popisu kmitající struny

stejným způsobem pak můžeme pokračovat dále. Dostaneme tak posloupnost

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, \dots$$



Obr. 3.3: Geometrický smysl metody tečen

Věta 3.12. Necht' $f(x)$ je definovaná a spojitá na intervalu I , $f(a)f(b) < 0$, $f'(x) \neq 0 \forall x \in I$, $f''(x)$ nemění znaménko na intervalu I , potom iterační proces

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, \dots$$

konverguje pro libovolné $x_0 \in I$ pro něž platí $f(x_0)f''(x_0) > 0$.

Věta 3.13. Necht' platí předpoklady věty 3.12 a dále

$$\left| \frac{f(a)}{f'(a)} \right| < b - a, \quad \left| \frac{f(b)}{f'(b)} \right| < b - a.$$

Potom Newtonova metoda konverguje pro libovolné $x_0 \in I$.

Pro odhad chyby platí:

Necht' $m_1 = \min |f'(x)|$, $M_2 = \max |f''(x)|$ pro $x \in I$. Potom

$$|\alpha - x_n| \leq \frac{M_2}{2m_1} (x_n - x_{n-1})^2.$$

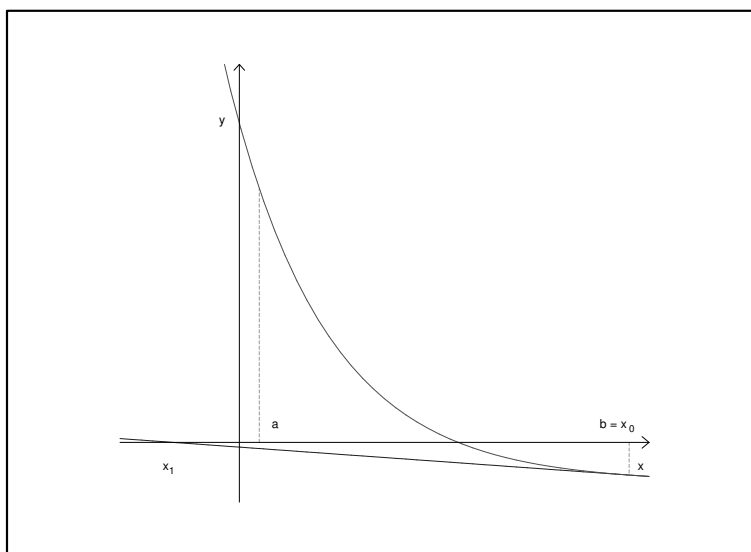
Tento odhad plyne z Taylorova rozvoje:

$$|\alpha - x_n| \leq \frac{M_2}{2m_1}(\alpha - x_0)^2.$$

Je-li tedy $\frac{M_2}{2m_1}(\alpha - x_0)^2 \leq k < 1$, potom Newtonova metoda konverguje, a to velmi rychle.

VYHODNOCENÍ:

1. Jestliže derivaci nahradíme diferencí, dostaneme metodu sečen.
2. Konverguje dostatečně rychle.
3. Podmínkou konvergence je u Newtonovy metody jednoduchý kořen α .
4. V případě vícenásobných kořenů je $f'(\alpha) = 0$ a tedy iterační vztah není v tomto bodě definován a nejsou splněny podmínky věty 3.12.
5. $f(x_0)f''(x_0) > 0$ je podmínka postačující, ne nutná. Ale při jejím nesplnění můžeme dojít ke sporu, kdy hodnota x_k bude ležet mimo interval I .



Obr. 3.4: Problém u metody tečen

Příklad 3.14. Najděte kladný kořen rovnice $\sin x - \frac{x}{2} = 0$.

Řešení. Grafickou metodou odhadneme, že kořen leží v intervalu $(\frac{\pi}{2}, \pi)$. Pro použití Newtonovy metody máme

$$f(x) = \sin x - \frac{x}{2},$$

$$f'(x) = \cos x - \frac{1}{2},$$

$$f''(x) = -\sin x.$$

V tomto případě můžeme použít obou krajních bodů a dostaneme se k cíli:

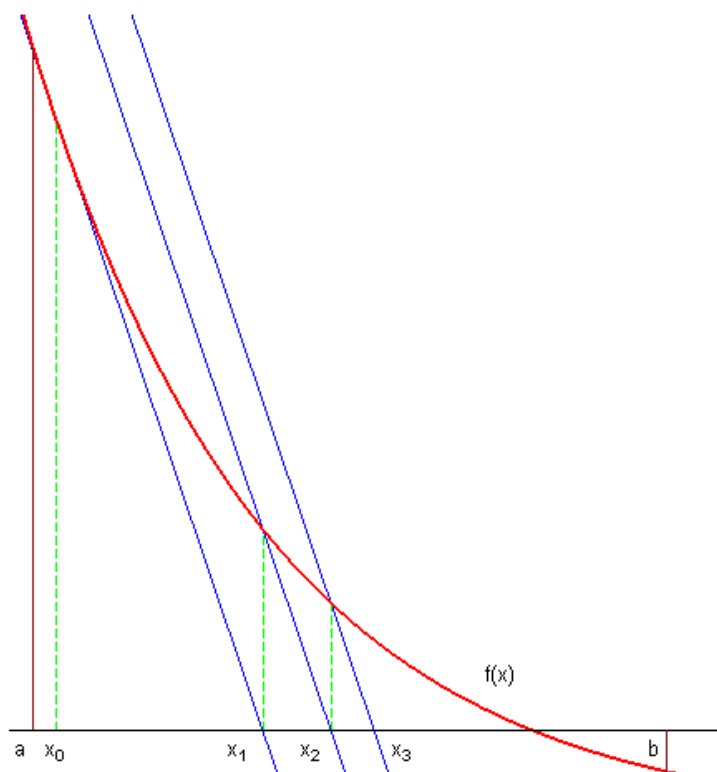
x_0	π	$\frac{\pi}{2}$
x_1	2.09440	2.0
x_2	1.91322	1.9010
x_3	1.89567	1.89551
x_4	1.89549	1,89549

□

3.3.5 Modifikovaná Newtonova metoda

Každý krok Newtonovy metody vyžaduje výpočet funkční hodnoty $f(x)$ a derivace $f'(x)$ v bodě x_n . Náročný a problematický může být zejména výpočet derivace. Jestliže se derivace podstatně nemění, je možno použít následující tvar iteračního vzorce:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_0)} = x_n - \frac{f(x_n)}{c}, \quad c = f'(x_0), \quad n = 0, 1, \dots$$



Obr. 3.5: Geometrický smysl modifikované metody tečen

Jestliže $f(x)$ je polynomem, při každém kroku ušetříme skoro polovinu operací. Dochází však ke zpomalení konvergence.

Existují i další modifikace, např. že se derivace počítá v každém druhém, každém desátém kroku, atd. Pokud při rozvoji do Taylorovy řady použijeme pro druhou derivaci její aproximaci interpolačním polynomem, dostaneme další modifikaci Newtonovy metody. Iterační proces má potom tvar

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)} - \frac{[f(x_i)]^2}{2[f'(x_i)]^3} \cdot \tilde{f}''(x_i),$$

kde

$$\tilde{f}''(x_i) = -\frac{6}{h_i^2} [f(x_i) - f(x_{i-1})] + \frac{2}{h_i} [2f'(x_i) + f'(x_{i-1})],$$

$$h_i = x_i - x_{i-1}.$$

Příklad 3.15. Řešme touto metodou stejnou úlohu jako v příkladu 3.14.

Řešení. Vezmeme si počáteční hodnoty $x_1 = \pi$, $x_2 = \frac{\pi}{2}$ a dostaneme

$$\begin{array}{ll} x_1 & \pi \\ x_2 & \frac{\pi}{2} \\ x_3 & 1.78659 \\ x_4 & 1.89414 \\ x_5 & 1.89549 \end{array}$$

Jde o výrazné zlepšení Newtonovy metody, které je však zapláceno tím, že potřebuje 2 výchozí hodnoty. \square

3.3.6 Newtonova metoda pro komplexní kořeny

Nechť $f(z)$ je komplexní funkce, která je analytická v okolí jejího izolovaného nulového bodu $w = \alpha + i\beta$, $f(w) = 0$. Potom stejně jako v případě reálných kořenů si odvodíme z Taylorovy řady

$$z_{n+1} = z_n - \frac{f(z_n)}{f'(z_n)}.$$

Věta 3.16. Je-li $f(z)$ analytická v uzavřeném okolí bodu z_0 o poloměru R a jestliže $\forall z : |z - z_0| \leq R$ platí

$$1. \quad \left| \frac{1}{f'(z_0)} \right| \leq A,$$

$$2. \quad \left| \frac{f(z_0)}{f'(z_0)} \right| \leq B \leq \frac{R}{2},$$

$$3. \quad |f''(z)| \leq C,$$

$$4. \quad ABC = \mu \leq 1,$$

má rovnice $f(z) = 0$ právě jeden kořen w v oblasti $|z - z_0| < R$ a iterační proces

$$z_{n+1} = z_n - \frac{f(z_n)}{f'(z_n)}, \quad n = 0, 1, \dots$$

konverguje k tomuto kořenu.

Jestliže chceme najít komplexní kořen, musíme volit z_0 jako komplexní číslo. Jinak se budeme pohybovat v reálném oboru.

3.3.7 Kombinovaná metoda sečen a tečen

Metoda sečen i tečen konvergují jednostranně. Konvergence se může podstatně urychlit, jestliže použijeme obě metody současně. Potom budou aproximace konvergovat k řešení z obou stran. Přitom se vyhneme problémům s nevhodnou konvergencí u jedné z metod pro speciální případy, kdy funkce protíná osu x pod nevhodným úhlem.

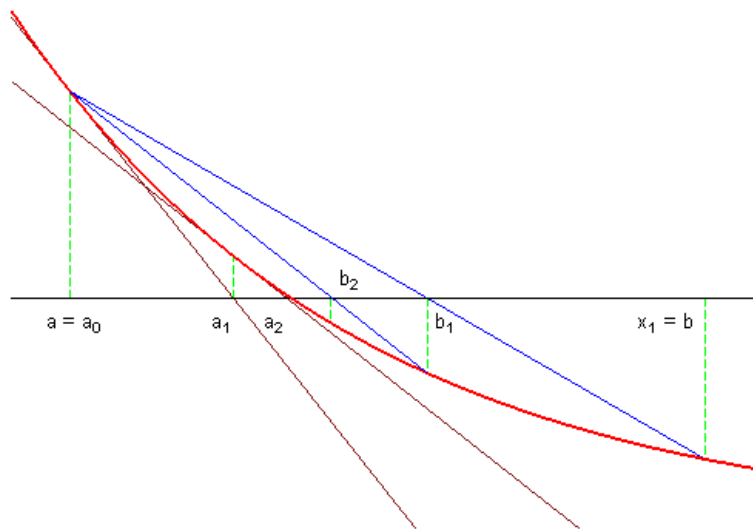
Věta 3.17. *Nechť jsou splněny předpoklady věty 3.12. Označme a_0 ten z bodů a, b pro který platí $f(a_0)f''(a_0) > 0$, druhý označme b_0 . Potom posloupnosti $\{a_k\}_{k=1}^{\infty}$, $\{b_k\}_{k=1}^{\infty}$ konvergují k řešení α rovnice $f(x) = 0$, t.j. $\lim_{k \rightarrow \infty} a_k = \lim_{k \rightarrow \infty} b_k = \alpha$, kde*

$$a_{k+1} = a_k - \frac{f(a_k)}{f'(a_k)},$$

$$b_{k+1} = \frac{b_k f(a_{k+1}) - a_{k+1} f(b_k)}{f(a_{k+1}) - f(b_k)}, \quad k = 0, 1, 2, \dots$$

VYHODNOCENÍ:

1. Kombinovaná metoda sečen a tečen má stejné výhody jako metody tečen a sečen.
2. Kořen vždy leží mezi prvky a_k, b_k pro libovolné k .
3. Lze snáze odhadnout chybu řešení.
4. Jde o oboustrannou konvergenci.



Obr. 3.6: Geometrický smysl kombinované metody sečen a tečen

3.4 Řád metody

Při numerickém řešení rovnice $f(x) = 0$ (a to i na počítači) je řada činností, které není účelné algoritmizovat, ale je účelnější je provádět mimo rámec výpočtu – separace kořenů, výběr metody, ověření splnitelnosti předpokladů, volba počáteční aproximace, odhad chyby, apod.

Definice 3.18. Iterační proces $x_{n+1} = \varphi(x_n)$ je r -tého řádu v bodě α , jestliže platí:

$$\begin{aligned}\varphi(\alpha) &= \alpha, \\ \varphi'(\alpha) &= \varphi''(\alpha) = \dots = \varphi^{(r-1)}(\alpha) = 0, \\ \varphi^{(r)}(\alpha) &\neq 0.\end{aligned}$$

Nechť jsou derivace φ do řádu r včetně spojitě v okolí bodu α . Potom podle Taylorova vzorce máme

$$\varphi(x) = \varphi(\alpha) + \frac{\varphi'(\alpha)}{1!}(x - \alpha) + \frac{\varphi''(\alpha)}{2!}(x - \alpha)^2 + \dots + \frac{\varphi^{(r-1)}(\alpha)}{(r-1)!}(x - \alpha)^{r-1} + \frac{\varphi^{(r)}(\xi)}{r!}(x - \alpha)^r,$$

$\xi \in (x, \alpha)$, což si můžeme podle definice 3.18 přepsat na tvar

$$\varphi(x) = \varphi(\alpha) + \frac{\varphi^{(r)}(\xi)}{r!}(x - \alpha)^r.$$

Protože $\varphi(\alpha) = \alpha$, $\varphi(x_n) = x_{n+1}$ máme

$$x_{n+1} = \alpha + \frac{\varphi^{(r)}(\xi_n)}{r!}(x_n - \alpha)^r,$$

kde ξ_n leží mezi x_n a α a závisí na x_n .

Absolutní chyba aproximace x_{n+1} je tedy

$$\alpha - x_{n+1} = -\frac{\varphi^{(r)}(\xi_n)}{r!}(x_n - \alpha)^r.$$

Máme tedy odhad

$$|\alpha - x_{n+1}| \leq K|(x_n - \alpha)^r|, \quad K = \sup \left| \frac{\varphi^{(r)}(\xi_n)}{r!} \right|.$$

Konstanta K se nazývá asymptotická konstanta chyby.

Čím je r větší, tím rychleji nám iterační proces konverguje.

- Metoda prosté iterace je řádu 1.
- Metoda sečen je řádu $\frac{1+\sqrt{5}}{2} \approx 1.618$.
- Metoda tečen je řádu 2.
- Dokonce i varianta Newtonovy metody pro násobné kořeny je v bodě α řádu 2.
- Varianta Newtonovy metody s 2. derivací je řádu 3.

Věta 3.19. *Nechť α je p -násobným kořenem rovnice $f(x) = 0$. Potom*

$$x_{n+1} = x_n - \frac{pf(x_n)}{f^{(p)}(x_n)}$$

je v okolí bodu α řádu 2.

Důkaz. Podle předpokladů je $f(\alpha) = f'(\alpha) = \dots = f^{(p-1)}(\alpha) = 0$, $f^{(p)}(\alpha) \neq 0$. Tím se v Taylorově rozvoji anulují první členy a my máme

$$f(x) = \frac{f^{(p)}(\alpha)}{p!}(x - \alpha)^p + \dots$$

Označme si $A = \frac{f^{(p)}(\alpha)}{p!}$ a získáme

$$\begin{aligned} x_{n+1} &= x_n - \frac{pA(x_n - \alpha)^p + \dots}{pA(x_n - \alpha)^{p-1} + \dots} = x_n - (x_n - \alpha) \frac{1 + \dots}{1 + \dots} = \\ &= x_n - (x_n - \alpha)R(x), \end{aligned}$$

kde $R(\alpha) = 1$. Označme $\varphi(x) = x - \frac{pf(x)}{f^{(p)}(x)}$. Potom

$$\varphi(\alpha) = \alpha, \quad \varphi'(x) = 1 - R(x) - (x - \alpha)R'(x),$$

takže $\varphi'(\alpha) = 0$, ale obecně je $\varphi''(\alpha) \neq 0$, a jedná se tedy o metodu 2. řádu. □

3.5 Algebraické rovnice

Definice 3.20. Algebraická rovnice je výraz tvaru $P_n(x) = 0$, kde

$$P_n(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0, \quad a_n \neq 0, \quad n \in \mathbb{N}.$$

Pro určení kořenů algebraické rovnice můžeme použít všechny předchozí metody pro určení kořene rovnice. Algebraické rovnice se však vyskytují velmi často v nejrůznějších aplikacích a mají své vlastní specifické vlastnosti. Proto se jimi budeme zabývat zvlášť.

Věta 3.21. O poloze kořenů.

Mějme rovnici $P_n(x) = 0$. Označme

$$A = \max\{|a_{n-1}|, |a_{n-2}|, \dots, |a_0|\}, \quad B = \max\{|a_n|, |a_{n-1}|, \dots, |a_1|\}.$$

Potom pro kořeny x_k polynomu $P_n(x)$ platí

$$\frac{|a_0|}{|a_0| + B} \leq |x_k| \leq \frac{|a_n| + A}{|a_n|}.$$

Příklad 3.22. Odhadněte polohu kořenů pro rovnici

$$P_6(x) = x^6 + 101x^5 + 425x^4 - 425x^2 - 101x - 1 = 0.$$

Řešení. Máme $A = 425$, $B = 425$. Dosazením dostaneme

$$\frac{1}{1 + 425} \leq |x_k| \leq \frac{1 + 425}{1},$$

$$\frac{1}{426} \leq |x_k| \leq 426.$$

Přesné řešení v tomto případě je

$$\begin{array}{ll} x_1 = 1 & x_4 = -0.241 \\ x_2 = -1 & x_5 = -4.143 \\ x_3 = -0.0104 & x_6 = -96.601 \end{array}$$

Vidíme, že všechny kořeny leží v námi určeném intervalu. □

Věta 3.23. Cauchyova o poloze kořenů.

Všechny kořeny polynomu $P_n(x)$ jsou obsaženy v kruhu Γ v komplexní rovině, kde

$$\Gamma = \{z \in \mathbb{C} : |z| \leq 1 + \eta_k\}$$

a

$$\eta_k = \max_{0 \leq k \leq n-1} \left| \frac{a_k}{a_n} \right|.$$

Věta 3.24. Descartesova¹

Počet kladných kořenů rovnice $P_n(x) = 0$ je roven počtu znaménkových změn v posloupnosti koeficientů $a_n, a_{n-1}, \dots, a_1, a_0$, přičemž nulové koeficienty jsou vynechány, a nebo je o sudý počet menší.

Důsledek 3.25. Polynom s kladnými koeficienty nemůže mít žádné kladné kořeny, protože v posloupnosti jeho koeficientů není žádná znaménková změna.

Důsledek 3.26. Počet záporných kořenů dostaneme z Descartesovy věty, když místo polynomu $P_n(x)$ vezmeme $P_n(-x)$.

Příklad 3.27. Odhadněte počet kladných a záporných kořenů pro rovnici

$$P_6(x) = x^6 + 101x^5 + 425x^4 - 425x^2 - 101x - 1 = 0.$$

Řešení. V posloupnosti $\{1, 101, 425, -425, -101, -1\}$ je pouze jedna znaménková změna, a proto bude mít tento polynom pouze jeden kladný kořen. Pro záporné kořeny máme

$$P_6(-x) = x^6 - 101x^5 + 425x^4 - 425x^2 + 101x - 1 = 0.$$

V posloupnosti $\{1, -101, 425, -425, 101, -1\}$ je pět znaménkových změn, a proto tento polynom může mít pět záporných kořenů (nebo tři nebo jeden).

Jde o stejný polynom jako v příkladu 3.22. Víme tedy, že náš polynom má pět záporných kořenů. \square

Definice 3.28. Sturmovou² posloupností pro polynom $P_n(x)$ nazveme posloupnost

$M(x), M_1(x), \dots, M_r(x)$, kde

$$M(x) = P_n(x),$$

$$M_1(x) = M'(x),$$

$M_2(x)$ je zbytek po dělení $M(x) : M_1(x)$ násobený číslem (-1) ,

$M_3(x)$ je zbytek po dělení $M_1(x) : M_2(x)$ násobený číslem (-1) ,

atd.

$M_r(x)$ je zbytek po dělení $M_{r-2}(x) : M_{r-1}(x)$ násobený číslem (-1) ,

přičemž zbytek $M_{r+1}(x)$ po dělení $M_{r-1}(x) : M_r(x)$ je nulový.

Věta 3.29. Sturmova.

Označme $N(c)$ počet znaménkových změn ve Sturmově posloupnosti pro $x = c$, přičemž nulové prvky vynecháme. Jestliže polynom $P_n(x)$ nemá násobné kořeny a jestliže je $P_n(a) \neq 0$, $P_n(b) \neq 0$, potom počet reálných kořenů tohoto polynomu ležících v intervalu (a, b) je roven $N(a) - N(b)$.

Příklad 3.30. Určete počet kladných a záporných kořenů rovnice $x^4 - 4x + 1 = 0$.

¹**R. Descartes** (1596 – 1650) francouzský matematik, filosof, fyzik, fyziolog. Jeden ze zakladatelů moderní vědy. Hlavní význam má jeho zavedení analytické geometrie. Pravděpodobně se zúčastnil bitvy na Bílé hoře na straně císařských.

²**J. Ch. F. Sturm** (1803 – 1855) francouzsko-švýcarský matematik. Zabýval se především diferenciálními rovnicemi.

Řešení. Máme rovnici $x^4 - 4x + 1 = 0$. Potom

$$M(x) = x^4 - 4x + 1,$$

$$M_1(x) = 4x^3 - 4 \Rightarrow M_1(x) = x^3 - 1.$$

Protože nás zajímá počet kladných a záporných kořenů a protože vždy dělíme polynom polynomem, můžeme proto polynom vynásobit kladným číslem, tím se ovlivní pouze velikost a ne znaménko. Obecně to ale neplatí. Pokud hledáme počet kořenů na intervalu $\langle a, b \rangle$, potom musíme pracovat s příslušnými tvary polynomů a nemůžeme si je dále upravovat.

$$M_2(x) = 3x - 1,$$

$$M_3(x) = 1.$$

Potom máme

	$-\infty$	0	$+\infty$
M	+	+	+
M_1	-	-	+
M_2	-	-	+
M_3	+	+	+
$N(c)$	2	2	0

Počet záporných kořenů je $N(-\infty) - N(0) = 2 - 2 = 0$.

Počet kladných kořenů je $N(0) - N(+\infty) = 2 - 0 = 2$.

Přesné řešení je

$$x_1 = 0.250992157490491,$$

$$x_2 = 1.4933585565601,$$

$$x_{3,4} = -0.872175357025343 \pm i1.38103159782422$$

Po dosazení do rovnice dostaneme chybu řádově 10^{-19} . □

Poznámka 3.31. Připomeňme si, jak se dělí polynomy. Dělíme mezi sebou první členy polynomů – v našem případě dělíme první člen, tj. x^4 členem x^3 a dostaneme

$$x^4 - 4x + 1 : x^3 - 1 = x,$$

(jde o první krok výpočtu, který ještě není ukončen, takže se nejedná o výsledek). Nyní prvkem x vynásobíme dělence a odečteme od dělitele

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

Protože stupeň výsledného polynomu je menší jak stupeň dělitele, ukončujeme výpočet. Pokud by byl stupeň výsledku větší nebo roven stupni dělitele, pokračujeme ve výpočtu stejným způsobem, tj. dělili bychom první členy polynomů mezi sebou a dále bychom odečítali násobky dělitele od dělence. Takže jsme dostali

$$x^4 - 4x + 1 : x^3 - 1 = x, \text{ zbytek } -3x + 1,$$

resp.

$$\frac{x^4 - 4x + 1}{x^3 - 1} = x + \frac{-3x + 1}{x^3 - 1}.$$

Příklad 3.32. Určete podíl polynomů P a Q , kde $P = x^3 - 9x^2 + 20x - 12$ a $Q = 3x^2 - 18x + 20$.

Řešení. Máme

$$\begin{aligned} x^3 - 9x^2 + 20x - 12 & : 3x^2 - 18x + 20 = \frac{1}{3}x - 1 \\ -(x^3 - 6x^2 + \frac{20}{3}x) & \\ -3x^2 + \frac{40}{3}x - 12 & \\ -(-3x^2 + 18x - 20) & \\ -\frac{14}{3}x + 8 & \end{aligned}$$

Dostali jsme

$$\frac{P}{Q} = \frac{x^3 - 9x^2 + 20x - 12}{3x^2 - 18x + 20} = \frac{1}{3}x - 1 + \frac{-\frac{14}{3}x + 8}{3x^2 - 18x + 20}.$$

□

Ve Sturmově větě požadujeme, aby $P_n(x)$ měl pouze prosté kořeny. Musíme proto umět určit, kdy je polynom má, resp. zajistit, aby je měl.

Věta 3.33. *Jestliže pro poslední člen Sturmovy posloupnosti platí $M_r = \text{Konst.}$, pak má $P_n(x)$ pouze prosté kořeny.*

Důkaz. Nechť $P_n(x)$ má i násobné kořeny, zapíšeme si polynom $P_n(x)$ ve tvaru součinu kořenových činitelů

$$P_n(x) = a_n(x - x_1)^{k_1}(x - x_2)^{k_2} \dots (x - x_m)^{k_m},$$

kde x_1, x_2, \dots, x_m jsou navzájem různé kořeny $P_n(x)$ a k_1, k_2, \dots, k_m jsou jejich násobnosti. Potom

$$\begin{aligned} P'_n(x) &= a_n[k_1(x - x_1)^{k_1-1}(x - x_2)^{k_2} \dots (x - x_m)^{k_m} + \\ &+ (x - x_1)^{k_1}k_2(x - x_2)^{k_2-1} \dots (x - x_m)^{k_m} + \dots + \\ &+ (x - x_1)^{k_1}(x - x_2)^{k_2} \dots k_m(x - x_m)^{k_m-1}] = \\ &= a_n(x - x_1)^{k_1-1}(x - x_2)^{k_2-2} \dots (x - x_m)^{k_m-1} \cdot \\ &\cdot [k_1(x - x_2) \dots (x - x_m) + (x - x_1)k_2(x - x_3) \dots (x - x_m) + \\ &+ \dots + (x - x_1)(x - x_2) \dots (x - x_{m-1})k_m] = \\ &= R(x)Q(x), \end{aligned}$$

kde $Q(x_k) \neq 0 \forall k = 1, 2, \dots, m$. Polynom $R(x)$ je největším společným dělitelem $P_n(x)$ a $P'_n(x)$. Určíme jej Eukleidovým¹ algoritmem. Potom

$$\frac{P_n(x)}{R(x)} = (x - x_1)(x - x_2) \dots (x - x_m)$$

¹**Eukleides z Alexandrie** (asi 340 př.n.l. – asi 278 př.n.l.) Starořecký matematik, autor nejvýznamnější matematické knihy celé dosavadní historie. Zabýval se geometrií, optikou, teorií hudby. Jako jeden z prvních se začal zabývat logickými základy matematiky. Jeho hlavním dílem je kniha “Základy” (řecky “Stoicheia”), která byla skoro 2000 let učebnicí matematiky a dodnes neztratila svoji důležitost. Obsahuje planimetrii, stereometrii, geometrickou algebru, řešení kvadratických rovnic, teorii čísel aj. Je to první pokus o axiomatickou výstavbu matematické teorie.

má pouze prosté kořeny. Když si uvědomíme, že Sturmova posloupnost je shodná s posloupností získanou Eukleidovým algoritmem (pro polynomy $P_n(x)$ a $P'_n(x)$) s přesností do znaménka, máme důkaz věty. \square

Důsledek 3.34. *Je-li α kořenem $P_n(x)$ a $P'_n(\alpha) = 0$, pak je α násobným kořenem. Je-li α kořenem $P_n(x)$ a je-li $P'_n(\alpha) \neq 0$, pak je α prostým kořenem.*

3.5.1 Metoda Laguerrova

Předpokládejme, že koeficienty polynomu $P_n(x)$ jsou reálné a všechny jeho kořeny jsou reálné a navíc prosté. Seřadíme-li si kořeny podle velikosti tak, aby platilo $\alpha_1 < \alpha_2 < \dots < \alpha_n$. Definujme si dále $\alpha_0 = -\infty$, $\alpha_{n+1} = +\infty$, a označme $I_i = \langle \alpha_i, \alpha_{i+1} \rangle$, $i = 0, 1, \dots, n$. Potom pro každé reálné $\alpha \neq \alpha_j$, $j = 0, 1, \dots, n + 1$ existuje právě jedno i takové, že platí $\alpha \in I_i$.

Sestrojíme si nyní kvadratickou funkci, která bude mít oba kořeny v I_i a v bodě α bude mít zápornou hodnotu. Takových funkcí můžeme sestavit nekonečně mnoho.

Hlavní myšlenkou Laguerrovy metody je sestavit takovou parabolu, tj. kvadratickou funkci, která bude protínat osu x co nejbližše koncům intervalu I_i . Tím dostáváme iterační posloupnost

$$\alpha^{k+1} = \alpha^k - \frac{nP_n(\alpha^k)}{(P_n(\alpha^k))' \pm \sqrt{H(\alpha^k)}}, \quad (3.1)$$

kde

$$H(\alpha) = (n-1) \left((n-1)(P'_n(\alpha))^2 - nP_n(\alpha)P''_n(\alpha) \right).$$

Věta 3.35. *Nechť $P_n(x)$ je reálný polynom stupně $n \geq 1$, jehož všechny kořeny jsou reálné a navzájem různé. Nechť α^0 je počáteční aproximace a $P(\alpha^0) \neq 0$. Potom iterační proces (3.1), kde znaménko před odmocninou volíme rovné $\text{sgn } P'(\alpha^k)$, vytváří posloupnost $\{\alpha^k\}_{k=0}^\infty$, která konverguje monotonně a kubicky k některému z kořenů polynomu $P_n(x)$.*

Je-li $P'(\alpha^k) = 0$, potom volíme znaménko před odmocninou v první iteraci libovolně. Přitom $P'(\alpha^k) \neq 0$ pro $k = 1, 2, \dots$.

Příklad 3.36. Mějme polynom $P(x) = x \left(\frac{1}{3}x^2 + a^2 \right)$, $a > 0$. Potom $P(a) \neq 0$. Jestliže zvolíme $x^0 = a$, potom budeme dostávat

$$x^{\text{sudé číslo}} = a, \quad x^{\text{liché číslo}} = -a.$$

Ke konvergenci nám stačí zvolit jinou počáteční aproximaci.

3.5.2 Metoda Graeffova – Lobačevského

Metoda Graeffova – Lobačevského ¹ slouží k přibližnému stanovení polohy všech kořenů polynomu. Nulové kořeny přitom předem vyloučíme, takže všude dále předpokládáme, že

¹ V anglicky a nebo německy psané literatuře se metoda označuje pouze prvním jménem jako metoda Graeffova.

$a_0 \neq 0$.

Zapišme si rozklad $P_n(x)$ na kořenové činitele

$$a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 = a_n (x - x_1)(x - x_2) \dots (x - x_n).$$

Jestliže nyní pravou stranu roznásobíme, pak srovnáním koeficientů dostaneme tzv. Vietovy ¹ vzorce.

$$\begin{aligned} \frac{a_{n-1}}{a_n} &= -(x_1 + x_2 + \dots + x_n), \\ \frac{a_{n-2}}{a_n} &= (-1)^2 (x_1 x_2 + x_1 x_3 + \dots + x_{n-1} x_n), \\ &\dots \\ \frac{a_0}{a_n} &= (-1)^n (x_1 x_2 x_3 \dots x_n). \end{aligned}$$

Princip Graeffovy-Lobačevského metody si ukážeme na příkladu. Mějme kvadratickou rovnici $a_2 x^2 + a_1 x + a_0 = 0$, potom podle Vietových vzorců platí

$$x_1 + x_2 = -\frac{a_1}{a_2}, \quad x_1 x_2 = \frac{a_0}{a_2}. \quad (3.2)$$

Jestliže se kořeny (v absolutní hodnotě) od sebe dostatečně liší, t.j. $x_1 \gg x_2$, potom

$$\left| \frac{x_2}{x_1} \right| \ll 1,$$

a

$$\begin{aligned} x_1 + x_2 &= -\frac{a_1}{a_2}, \\ x_1 \left(1 + \frac{x_2}{x_1}\right) &= -\frac{a_1}{a_2}. \end{aligned}$$

Zanedbáním druhého sčítance dostaneme

$$x_1 \approx -\frac{a_1}{a_2}.$$

Dosazením této hodnoty do rovnice

$$x_1 x_2 = \frac{a_0}{a_2}$$

K. L. Graeffe (1799 – 1873) německý matematik, žák K.F.Gausse. Věnoval se algebře, teoretické mechanice, historii matematiky. Metodu řešení algebraické rovnice navrhnul v r.1837.

N. I. Lobačevskij (1792 – 1856) ruský matematik. Vystudoval universitu v Kazani a celý život tam působil. Do historie matematiky vešel jako tvůrce neeukleidovské geometrie. První práce jsou z r.1826. Přestože byl vystaven nevybíravým útokům a zůstal prakticky nepochopen a osamocen, pokračoval v rozvíjení své geometrie. Uznání se mu dostalo až posmrtně.

¹**F. Viete** (1540 – 1603) francouzský matematik a právník. Jeho matematické práce byly psány velmi těžkým jazykem a proto dlouho nevěšly v obecnou známost. Jako první zavedl označování nejen neznámých ale i koeficientů písmeny, přitom počáteční písmena abecedy vyhradil pro koeficienty a koncová písmena pro neznámé. Rozluštil kód, který používali Španělé ve válce proti Francii.

dostaneme
$$x_2 \approx -\frac{a_0}{a_1}.$$

Tyto hodnoty můžeme brát jako první iteraci.

Umocníme (3.2) a po úpravě máme

$$x_1^2 + x_2^2 = \frac{a_1^2 - 2a_0a_2}{a_2^2}, \quad x_1^2x_2^2 = \frac{a_0^2}{a_2^2}.$$

Protože známe součet a součin druhých mocnin kořenů, můžeme si sestavit rovnici $b_2y^2 + b_1y + b_0 = 0$ s kořeny $y_1 = x_1^2$, $y_2 = x_2^2$, kde

$$b_2 = a_2^2, \quad b_1 = -(a_1^2 - 2a_0a_2), \quad b_0 = a_0^2.$$

Zcela analogicky můžeme určit i jejich aproximace

$$y_1 \approx -\frac{b_1}{b_2}, \quad y_2 \approx -\frac{b_0}{b_1},$$

a odtud plyne, že

$$|x_1| \approx \sqrt{\left|\frac{b_1}{b_2}\right|}, \quad |x_2| \approx \sqrt{\left|\frac{b_0}{b_1}\right|}.$$

Přičemž jestliže je $\left|\frac{x_2}{x_1}\right| \ll 1$, pak člen $\left|\frac{x_2^2}{x_1^2}\right|$ bude ještě menší, tedy aproximace bude lepší. O znaménku rozhodneme na základě dosazení do původní rovnice. Takto můžeme pokračovat dále.

Obecně chceme najít aproximaci kořenů x_1, \dots, x_n rovnice $P_n(x) = a_nx^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0 = 0$, $a_n \neq 0$, přičemž předpokládáme, že kořeny tohoto polynomu jsou dobře rozlišitelné, tj. že $P_n(x)$ má pouze reálné prosté kořeny, $|x_1| > |x_2| > \dots > |x_n|$. Vytvoříme si posloupnost polynomů $P^0, P^1, \dots, P^k, \dots$, kde

$$\begin{aligned} P^0 &= P_n(x), \\ P^k(x) &= a_n^k x^n + a_{n-1}^k x^{n-1} + \dots + a_1^k x + a_0^k, \\ a_n^k &= (a_n^{k-1})^2, \\ -a_{n-1}^k &= (a_{n-1}^{k-1})^2 - 2a_n^{k-1}a_{n-2}^{k-1}, \\ a_{n-2}^k &= (a_{n-2}^{k-1})^2 - 2a_{n-1}^{k-1}a_{n-3}^{k-1} + 2a_n^{k-1}a_{n-4}^{k-1}, \\ &\dots \\ (-1)^{n-1}a_1^k &= (a_1^{k-1})^2 - 2a_2^{k-1}a_0^{k-1}, \\ (-1)^na_0^k &= (a_0^{k-1})^2. \end{aligned}$$

Pro kořeny pak platí

$$|x_j| = \sqrt[2^k]{\left|\frac{a_{n-j}^k}{a_{n-j+1}^k}\right|}, \quad j = 1, 2, \dots, n.$$

Výpočet zastavujeme, jestliže platí $|a_i^k| \approx (a_i^{k-1})^2 \forall i = 0, 1, 2, \dots, n$.

VYHODNOCENÍ:

1. Podmínka dobré *rozlišitelnosti kořenů* znamená, že předpokládáme, že $P_n(x)$ má pouze reálné prosté kořeny.
2. Jestliže se absolutní hodnoty kořenů liší od sebe velmi málo, potom G – L metoda konverguje velmi pomalu – nevýhoda této metody.
3. Nepožaduje se separace kořenů. Stačí odstranit násobné, jako u Sturmovy věty.
4. Problematická je i otázka přesnosti.
5. Metoda nám určí *slušné* aproximace všech kořenů, které je potom nutno zpřesnit jinou metodou.

Jestliže v posloupnosti $\{a_i^k\}, k = 1, 2, \dots, i = 0, 1, \dots, n$ dochází ke znaménkovým změnám, znamená to (ve většině případů), že polynom má i komplexní kořeny.

Předpokládejme, že

$$|\alpha_1| > |\alpha_2| > \dots > |\alpha_i| = |\alpha_{i+1}| > |\alpha_{i+2}| > \dots > |\alpha_n|,$$

$$\alpha_i = r(\cos \varphi + i \sin \varphi),$$

$$\alpha_{i+1} = r(\cos \varphi - i \sin \varphi),$$

$$|\alpha_i| = |\alpha_{i+1}| = r.$$

Potom

$$r^2 = |\alpha_i| \cdot |\alpha_{i+1}| \approx \sqrt[2^k]{\left| \frac{a_{n-i}^k}{a_{n-i+1}^k} \right|} \sqrt[2^k]{\left| \frac{a_{n-i-1}^k}{a_{n-i}^k} \right|} = \sqrt[2^k]{\left| \frac{a_{n-i-1}^k}{a_{n-i+1}^k} \right|}$$

a potom

$$\alpha_1 + \alpha_2 + \dots + \alpha_n = -\frac{a_{n-1}}{a_n},$$

$$\alpha_i + \alpha_{i+1} = 2r \cos \varphi,$$

$$2r \cos \varphi = -\frac{a_{n-1}}{a_n} - \sum_{k=1, k \neq i, i+1}^n \alpha_k.$$

Známe součet a součin kořenů, aproximaci kořenů pak získáme řešením kvadratické rovnice

$$\xi^2 - (2r \cos \varphi)\xi + r^2 = 0,$$

$$\xi^2 - \left(-\frac{a_{n-1}}{a_n} - \sum_{k=1, k \neq i, i+1}^n \alpha_k \right) \xi + \sqrt[2^k]{\left| \frac{a_{n-i-1}^k}{a_{n-i+1}^k} \right|} = 0.$$

Opět je nutné počítat s chybou a výsledek zpřesnit jinou metodou.

Příklad 3.37. Graeffovou-Lobačevského metodou určete kořeny rovnice

$$z^4 - 4z^3 + 3z^2 + 2z - 6 = 0$$

Řešení. Máme polynom 4. stupně. Sestavíme si odpovídající posloupnost polynomů s koeficienty a_i , $i = 0, 1, 2, 3, 4$:

k	a_4	a_3	a_2	a_1	a_0
0	1	-4	3	2	-6
1	1	-10	13	-40	36
2	1	-74	-559	-664	1296
3	1	-6594	218801	-1889824	1679616

Potom

$$|\alpha_1| \approx \sqrt[8]{\frac{6594}{1}} \doteq 3.001882007.$$

Dosazením do polynomu dostaneme

$$P_4(3.001882007) \approx 0.0377,$$

$$P_4(-3.001882007) \approx 204.44.$$

Proto za aproximaci prvního kořene přijímáme hodnotu $+3.001882007$. Dále

$$|\alpha_2| \approx \sqrt[8]{\frac{1679616}{1889824}} \doteq 0.9853682858.$$

Dosazením do polynomu dostaneme

$$P_4(0.9853682858) \approx -4.006,$$

$$P_4(-0.9853682858) \approx -0.288.$$

Proto za aproximaci prvního kořene přijímáme hodnotu -0.9853682858 . V dalších sloupcích dochází ke střídání znamének, budeme proto hledat komplexní kořeny:

$$r^2 = \sqrt[8]{\frac{1889824}{6594}} \approx 2.028425455 \approx |\alpha_3\alpha_4|,$$

$$2r \cos \varphi = 4 - \alpha_1 - \alpha_2 \doteq 1.983486276.$$

Určíme si řešení kvadratické rovnice

$$\zeta^2 - 1.983486276\zeta + 2.028425455 = 0,$$

$$\alpha_{3,4} = 0.991743139 \pm j1.022289317.$$

□

3.5.3 Metoda Schurova

Mějme polynom

$$f(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0, \quad (3.3)$$

který je obecně vzato komplexní. Označme

$$f^*(z) = z^n \bar{f}(\bar{z}^{-1}) = \bar{a}_n + \bar{a}_{n-1} z + \cdots + \bar{a}_0 z^n,$$

kde \bar{u} je číslo komplexně sdružené k číslu u . Dále

$$T(f(z)) = \bar{a}_0 f(z) - a_n f^*(z).$$

Speciálně platí

$$T(f(0)) = \bar{a}_0 a_0 - a_n \bar{a}_n = |a_0|^2 - |a_n|^2,$$

neboli $T(f(0))$ je reálné číslo.

Protože $T(f(z))$ je polynomem stupně nejvýše $n - 1$, můžeme si definovat posloupnost polynomů

$$T^j(f(z)) = T(T^{j-1}(f(z))).$$

Protože stupeň polynomů klesá, je jejich posloupnost konečná.

Věta 3.38. Označme k nejmenší celé číslo, pro které platí $T^k(f(0)) = 0$. Necht' $f(0) \neq 0$. Jestliže pro $h : 0 < h < k$ platí, $T^h(f(0)) < 0$, potom má polynom f nejméně jeden kořen uvnitř jednotkového kruhu se středem v počátku.

Je-li naopak $T^i(f(0)) > 0$ pro $1 \leq i \leq k$ a $T^{k-1}(f(z))$ je konstanta, neleží uvnitř jednotkového kruhu se středem v počátku žádný kořen.

Na základě této věty platí:

1. Je-li $f(0) = 0$, potom existuje kořen $z = 0$.
Je-li $f(0) \neq 0$, potom pokračujeme bodem 2.
2. Vypočteme $T(f(z))$. Je-li $T(f(0)) < 0$, existuje uvnitř jednotkového kruhu kořen.
Je-li $T(f(z)) \geq 0$ pokračujeme bodem 3.
3. Vypočítáme postupně $T^j(f(z))$ pro $j = 1, 2, \dots$, až buď $T^j(f(z)) < 0$ pro nějaké $j < k$, nebo $T^k(f(0)) = 0$.
V prvním případě leží uvnitř jednotkového kruhu aspoň jeden kořen.
Nastane-li druhý případ a je-li polynom $T^{k-1}(f(z))$ konstantní, neleží uvnitř jednotkového kruhu žádný kořen.

Poznámka 3.39. Všimněte si, že je-li $T^k(f(0)) = 0$ a $T^{k-1}(f(z))$ není konstanta, potom větu (3.38) nelze využít.

Má-li polynom $f(z)$ kořen uvnitř kruhu $|z| = \varrho$, potom má polynom

$$g(z) = f(\varrho z)$$

kořen uvnitř jednotkového kruhu.

A obecně: má-li $f(z)$ kořen uvnitř kruhu $|z - c| = \varrho$, má polynom

$$g(z) = f(\varrho z + c)$$

kořen uvnitř jednotkového kruhu.

Při hledání kořene proto postupujeme následovně:

1. Nemá-li polynom $f(z)$ kořen uvnitř jednotkového kruhu, budeme uvažovat polynom $g(z) = f(2z)$ a zkoumat, má-li $g(z)$ kořen uvnitř jednotkového kruhu. Pokud nemá, budeme zkoumat polynom $f(2^2z)$, atd. Budeme pokračovat až nalezneme mezikruží

$$R = z^j \leq |z| < z^{j+1} = 2R, \quad (3.4)$$

takové, že polynom f má kořen v tomto mezikruží a současně nemá žádný kořen uvnitř kruhu o poloměru R .

Má-li polynom $f(z)$ kořen uvnitř jednotkového kruhu, potom budeme púlit poloměr tak dlouho, až dostaneme nerovnost typu (3.4). Opět tedy budeme mít mezikruží, ve kterém leží kořen.

2. Mezikruží (3.4) lze úplně pokrýt 8 překrývajícími se kruhy o poloměru $\frac{4}{5}R$ a se středy v bodech

$$\frac{3R}{2 \cos \frac{1}{8}\pi} e^{\frac{2\pi jk}{8}}, \quad k = 0, 1, \dots, 7.$$

Postupně prověřujeme jednotlivé kruhy, až nalezneme kruh, ve kterém leží aspoň jeden kořen.

3. Označme střed tohoto kruhu C_1 a pokračujeme jako v bodě 1. Přitom v každém kruhu púlíme poloměr. Začínáme od $\frac{4}{5}R$ až dostaneme mezikruží

$$R_1 = \frac{4}{5}R \cdot 2^{-j_1} \leq |z - C_1| < \frac{4}{5}R \cdot 2^{-(j_1-1)} = 2R_1,$$

které obsahuje kořen polynomu f . Stejně jako v bodě 2 pokryjeme mezikruží 8 kruhy.

Postup opakujeme tak dlouho, dokud je třeba.

Příklad 3.40. Určete kořeny polynomu $f(z) = 4z^3 - 8z^2 + 9z - 18$.

Řešení. Sestrojíme si posloupnost polynomů:

$$f^*(z) = -18z^3 + 9z^2 - 8z + 4.$$

$$f_1(z) = T(f(z)) = -18f(z) - 4f^*(z) = 108z^2 - 130z + 308,$$

$$T(f(0)) = 308,$$

$$f_1^*(z) = 308z^2 - 130z + 108,$$

$$f_2(z) = T(f_1(z)) = T^2(f(z)) = 308f_1(z) - 108f_1^*(z) = 400(-65z + 208),$$

$$T^2(f(0)) = 83200,$$

$$f_2^*(z) = 400(208z - 65),$$

$$T(f_2(z)) = T^3(f(z)) = (208^2 - 65^2) \cdot 400^2,$$

$$T^3(z) = \textit{konstanta},$$

$$T^4(f(0)) = 0.$$

Tím jsme prokázali, že uvnitř jednotkového kruhu neleží žádný kořen. Pro $g(z) = f(2z)$ dostaneme

$$T(g(z)) = 700z - 700,$$

neboli

$$T(g(0)) = -700 < 0,$$

a proto má $f(z)$ kořen v mezikruží $1 \leq |z| < 2$. Stejným způsobem můžeme pokračovat dále, tj. mezikruží pokryjeme 8 kruhy a určíme v kterém z nich leží kořen. Přesné řešení je

$$z_1 = +2, \quad z_{2,3} = \pm \frac{3}{2}j.$$

□

Pojmy k zapamatování

- Seznámili jsme se s numerickými metodami řešení lineárních i nelineárních rovnic.
- Po úvodním seznámením se startovacími metodami - grafickou, tabelovací a metodou půlení intervalu, jsme přešli k iteračním metodám různých typů.
- Iterační metody vyžadují pro svoji konvergenci splnění dalších doplňujících podmínek, proto bude pro nás vždy výhodné, pokud dokážeme vhodnou startovací metodou zúžit interval, na kterém hledáme kořen naší rovnice.
- Seznámili jsme se s metodou prosté iterace, s metodami tečen a sečen a s jejich modifikacemi.
- V závěru jsme se věnovali algebraickým rovnicím. Při řešení algebraických rovnic můžeme použít všechny dříve uvedené metody, ale protože algebraické rovnice mají i své specifické vlastnosti, které můžeme využít při hledání kořenů, věnovali jsme se jim samostatně.

Kontrolní otázky

1. Co rozumíme pojmem kořen rovnice?
2. Je možné použít Newtonovu metodu pro libovolnou rovnici?
3. Je možné u každé metody provést odhad chyby?

3.6 Cvičení

Maplety

Příklady pro samostatnou práci jsou uvedeny v samostatné sbírce příkladů. Pomocí následujících mapletů si můžete usnadnit některé dílčí výpočty, nebo zkontrolovat jejich správnost.

1. [Derivování](#)
2. [Výpočet funkčních hodnot funkcí jedné proměnné](#)
3. [Kreslení grafů elementárních funkcí](#)
4. [Kreslení grafů funkcí jedné proměnné](#)
5. [Dělení polynomů](#)

webMathematica

Příklady pro samostatnou práci jsou uvedeny v samostatné sbírce příkladů. Pomocí následujících aplikací si můžete usnadnit některé dílčí výpočty, nebo zkontrolovat jejich správnost.

1. [Násobení matic](#)

3.7 Animace

Na následujících animacích jsou znázorněny postupy některých numerických metod studovaných v této kapitole.

Animace ovládáte tlačítky pod obrázkem. Kvalita animací je dána skutečností, že se jedná o bitmapové obrázky exportované ze softwaru Maple a následně vkládané do textu v jiné velikosti. Animace se nacházejí vždy na samostatné stránce.

Obr. 3.7: Regula falsi (animace)

Obr. 3.8: Newtonova metoda (animace)

Obr. 3.9: Modifikovaná Newtonova metoda (animace)

4 Vlastní čísla

Průvodce studiem

Budeme se věnovat vlastním číslům a vlastním vektorům matic. Definujeme si oba pojmy a ukážeme si jejich základní vlastnosti. Dále se seznámíme s vybranými metodami pro jejich přímý výpočet a také s některými numerickými metodami pro jejich určení.

Vlastní čísla matic hrají důležitou úlohu v mnoha oborech, například v teorii systémů diferenciálních rovnic.

Cíle

Po prostudování této kapitoly budete schopni:

- Definovat vlastní číslo a vlastní vektor matice.
- Vypočítat vlastní čísla a vlastní vektory a to jak přímo, tak i pomocí numerických metod.
- Odhadnout polohu vlastních čísel matice.

4.1 Základní pojmy

Definice 4.1. Nechť \mathbf{A} je čtvercová matice řádu n . Její *vlastní čísla* $\lambda_1, \dots, \lambda_n$ jsou kořeny rovnice

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0$$

zvané *charakteristická rovnice*.

Po roznásobení dostaneme charakteristickou rovnici ve tvaru

$$f(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0 = 0.$$

Definice 4.2. Ke každému vlastnímu číslu λ_i existuje aspoň jedno nenulové řešení soustavy rovnic $\mathbf{A}\mathbf{x} = \lambda_i\mathbf{x}$. Toto řešení \mathbf{x}_i , kde $\mathbf{x}_i^T = (x_i^{(1)}, x_i^{(2)}, \dots, x_i^{(n)})$, nazveme *pravým vlastním vektorem* matice \mathbf{A} .

Všude v dalším bude pojem vlastní vektor značit výhradně pravý vlastní vektor.

Definice 4.3. Levý vlastní vektor \mathbf{y}_i odpovídající vlastnímu číslu λ_i je řešením rovnice $\mathbf{y}^T \mathbf{A} = \lambda_i \mathbf{y}^T$.

Levý vlastní vektor matice \mathbf{A} je tedy vlastním vektorem transponované matice \mathbf{A}^T a snadno lze ukázat, že odpovídá-li levý vlastní vektor \mathbf{y}_k vlastnímu číslu λ_k a pravý vlastní vektor \mathbf{x}_i vlastnímu číslu λ_i a platí $\lambda_k \neq \lambda_i$ jsou vektory \mathbf{y}_k a \mathbf{x}_i ortogonální.

Definice 4.4. Vektory u a v se nazývají *ortogonálními*, jestliže jejich skalární součin je roven nule.

Připomeňme si, že skalární součin vektorů $\vec{u} = (u_1, u_2, \dots, u_n)^T$ a $\vec{v} = (v_1, v_2, \dots, v_n)^T$ je definován jako $u_1v_1 + u_2v_2 + \dots + u_nv_n = \sum_{i=1}^n u_iv_i$.

Definice 4.5. Vektory u a v se nazývají *ortonormálními*, jestliže jsou ortogonální, tj. jejich skalární součin je roven nule, a jestliže navíc je jejich norma rovna jedné.

Ve většině dále uvedených příkladů se budou vyskytovat reálné matice, budeme předpokládat, pokud nebude řečeno jinak, že matice \mathbf{A} je reálná. Mnohé věty budou však platit i pro komplexní matice nebo budeme-li předpokládat symetrii, pro hermitovské matice.

Příklad 4.6. Určete vlastní čísla a vlastní vektory matice

$$A = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Řešení. Pomocí Sarrusova pravidla určíme charakteristický polynom

$$|A - \lambda I| = \begin{vmatrix} (-1 - \lambda) & 1 & 1 \\ 1 & (-1 - \lambda) & 1 \\ 1 & 1 & (1 - \lambda) \end{vmatrix} = -\lambda^3 - \lambda^2 + 4\lambda + 4.$$

Charakteristická rovnice $-\lambda^3 - \lambda^2 + 4\lambda + 4 = 0$ je celočíselná. Podle Vietových vzorců musí každý kořen dělit absolutní člen. V našem případě je to číslo 4. Pomocí Hornerova schématu budeme nejdříve hledat celočíselné kořeny, pokud existují. Dostaneme řešení

$$\lambda_1 = -1, \quad \lambda_2 = 2, \quad \lambda_3 = -2.$$

Dosazením do rovnice $(A - \lambda I)x = \mathcal{O}$ dostaneme homogenní soustavu a jejím řešením jsou vlastní vektory:

$$\lambda_1 : \begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{vmatrix} \sim \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{vmatrix} \sim \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{vmatrix}$$

$$v_1 = (-t, -t, t), \quad t \in \mathbb{R}.$$

$$\lambda_2 : \begin{vmatrix} -3 & 1 & 1 \\ 1 & -3 & 1 \\ 1 & 1 & -1 \end{vmatrix} \sim \begin{vmatrix} 1 & 1 & -2 \\ 0 & -4 & 2 \\ 0 & 4 & -2 \end{vmatrix} \sim \begin{vmatrix} 2 & 0 & 1 \\ 0 & -2 & 1 \\ 0 & 0 & 0 \end{vmatrix}$$

$$v_2 = (s, s, 2s), \quad s \in \mathbb{R}.$$

$$\lambda_3 : \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{vmatrix} \sim \begin{vmatrix} 1 & 1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{vmatrix} \sim \begin{vmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{vmatrix}$$

$$v_3 = (-u, u, 0), \quad u \in \mathbb{R}.$$

Volíme $t = s = u = 1$ a dostaneme vlastní vektory

$$v_1 = (-1, -1, 1)^T, \quad v_2 = ((1, 1, 2)^T, \quad v_3 = (-1, 1, 0)^T.$$

□

Prímý výpočet podle definice nemusí být vždy nejlepší a nejrychlejší.

Příklad 4.7. Určete vlastní čísla matice

$$A = \begin{pmatrix} 3 & -32 & -24 \\ 2 & -35 & -27 \\ -4 & 44 & 33 \end{pmatrix}.$$

Řešení. Budeme postupně upravovat determinant

$$\begin{aligned} |A - \lambda I| &= \begin{vmatrix} 3 - \lambda & -32 & -24 \\ 2 & -35 - \lambda & -27 \\ -4 & 44 & 33 - \lambda \end{vmatrix} = \begin{vmatrix} 1 - \lambda & 3 + \lambda & 3 \\ 2 & -35 - \lambda & -27 \\ -2 & 9 - \lambda & 6 - \lambda \end{vmatrix} = \\ & \begin{vmatrix} 1 - \lambda & \lambda & 3 \\ 2 & -8 - \lambda & -27 \\ -2 & 3 & 6 - \lambda \end{vmatrix} = \begin{vmatrix} 1 & \lambda & 3 \\ -6 - \lambda & -8 - \lambda & -27 \\ 1 & 3 & 6 - \lambda \end{vmatrix} = \begin{vmatrix} 1 & \lambda - 1 & -1 \\ -6 - \lambda & -2 & -3 + 4\lambda \\ 1 & 2 & 2 - \lambda \end{vmatrix} = \\ & \begin{vmatrix} 0 & \lambda - 3 & -3 + \lambda \\ -5 - \lambda & 0 & -1 + 3\lambda \\ 1 & 2 & 2 - \lambda \end{vmatrix} = \begin{vmatrix} 0 & \lambda - 3 & 0 \\ -5 - \lambda & 0 & -1 - 3\lambda \\ 1 & 2 & -\lambda \end{vmatrix} = (*). \end{aligned}$$

Poslední determinant rozvineme podle prvního řádku a dostaneme

$$\begin{aligned} (*) &= -(\lambda - 3) \cdot \begin{vmatrix} -5 - \lambda & -1 - 3\lambda \\ 1 & -\lambda \end{vmatrix} = -(\lambda - 3) \cdot ((-5 - \lambda)(-\lambda) - (-1 - 3\lambda)) = \\ &= -(\lambda - 3)(5\lambda + \lambda^2 + 1 - 3\lambda) = -(\lambda - 3)(\lambda^2 + 2\lambda + 1) = -(\lambda - 3)(\lambda + 1)^2. \end{aligned}$$

Takže jsme dostali

$$\lambda_1 = 3, \quad \lambda_{2,3} = -1.$$

□

Věta 4.8. Jsou-li $\lambda_1, \dots, \lambda_n$ vlastní čísla matice \mathbf{A} , má matice \mathbf{A}^k vlastní čísla $\lambda_1^k, \dots, \lambda_n^k$. Obecněji, je-li $p(\mathbf{x})$ libovolný polynom, má matice $p(\mathbf{A})$ vlastní čísla

$$p(\lambda_1), \dots, p(\lambda_n).$$

Věta 4.9. Je-li matice \mathbf{A} reálná a symetrická, jsou všechna její vlastní čísla a všechny příslušné vlastní vektory reálné. Kromě toho vlastní vektory příslušné různým vlastním číslům jsou ortogonální a levý vlastní vektor a pravý vlastní vektor příslušné témuž vlastnímu číslu jsou si rovny.

Věta 4.10. Podobnostní transformace $\mathbf{P} \cdot \mathbf{A} \cdot \mathbf{P}^{-1}$ nemění vlastní čísla matice \mathbf{A} .

Věta 4.11. (Cayley-Hamiltonova) Nechť je $f(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = 0$ charakteristická rovnice matice \mathbf{A} . Pak platí $f(\mathbf{A}) = \mathcal{O}$, kde \mathcal{O} je nulová matice. Tj. jestliže

$$f(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0 = 0,$$

potom

$$\mathbf{A}^n + a_{n-1}\mathbf{A}^{n-1} + \dots + a_1\mathbf{A} + a_0\mathbf{I} = \mathcal{O}.$$

Neboli matice \mathbf{A} je kořenem charakteristické rovnice.

Věta 4.12. Vlastní čísla horní (resp. dolní) trojúhelníkové matice jsou prvky na její diagonále.

Věta 4.13. Má-li libovolná matice \mathbf{A} kompletní soubor n lineárně nezávislých vlastních vektorů, potom matice \mathbf{A} je podobná diagonální matici \mathbf{D} , kde na hlavní diagonále stojí vlastní čísla matice \mathbf{A} .

4.2 Numerické metody pro hledání vlastních čísel

Podle základní definice víme, že vlastní čísla dané matice jsou kořeny jejího charakteristického polynomu. Z algebraické teorie víme, že kořeny polynomu stupně $n > 4$ nemůžeme algebraicky (tj. pomocí operací $\pm, \times, \div, \sqrt{\quad}$ s koeficienty rovnice) vyjádřit ve tvaru vzorce. Proto se obecně nedají získat vlastní čísla přesně (až na zaokrouhlovací chyby) po konečném počtu operací. K řešení našeho problému můžeme přistupovat více způsoby.

1. Použijeme libovolnou metodu na hledání kořenů charakteristického polynomu $p(\lambda)$. Pro jednoduchý kořen můžeme použít *Newtonovu metodu*

$$c_{i+1} = c_i - \frac{p(c_i)}{p'(c_i)}, \quad i = 1, 2, \dots$$

při vhodné volbě počáteční aproximace c_0 , *metodu sečen*, *metodu půlení intervalu* atd. Modifikovaná Newtonova metoda se dá použít i na hledání násobných kořenů. V případě komplexně sdružené dvojice kořenů můžeme použít např. *Bairstowovu metodu*. Hledání velkého počtu kořenů tímto způsobem je však dost náročné a problém bývá obvykle nestabilní vlivem nakupení zaokrouhlovacích chyb.

2. Získání vlastních čísel bez znalostí charakteristického polynomu, jen využíváním vlastností podobných matic. Cílem je najít podobnou matici v jednodušším tvaru, ze kterého se dá vlastní číslo určit (například z diagonální nebo trojúhelníkové matice). Takovou matici (někdy jen některé její vlastní číslo) můžeme získat jako limitu posloupnosti podobnostních transformací. Výběr těchto transformací bývá založen na speciálních vlastnostech matic a jejich vlastních vektorech
3. Nelineární přístup, kdy vlastní problém

$$(\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0}$$

uvažujeme jako soustavu n rovnic pro $n + 1$ neznámých $\mathbf{x}_1, \dots, \mathbf{x}_n, \lambda$, kterou doplníme normovanou podmínkou například $\sum_i \mathbf{x}_i^2 = 1$ na soustavu $n + 1$ nelineárních rovnic. Tato soustava se dá řešit například Newtonovou metodou. Přitom se však nevyužívají algebraické vlastnosti soustavy, které mohou výpočet značně ulehčit. Proto je tento postup značně neefektivní.

Poznámka 4.14. Pod pojmem *úplný problém vlastních čísel* se rozumí úloha najít všechna vlastní čísla a případně i příslušné vlastní vektory. Pojem *částečný problém vlastních čísel* znamená najít jedno nebo více vlastních čísel spolu s příslušnými vlastními vektory.

Úplný a částečný problém vyžadují odlišné metody řešení.

Řešení úplného problému je náročnější. Navíc neexistuje (nebo přesněji není dosud znám) univerzální algoritmus, který by byl stejně efektivní pro všechny typy matic.

4.3 Klasické metody určení koeficientů charakteristického polynomu

Zpočátku se většina metod na výpočet vlastních čísel zakládala právě na výpočtu koeficientů charakteristického polynomu. Jejich výpočet pomocí součtu hlavních minorů je však neefektivní. Existují mnohem jednodušší metody na určení koeficientů, které mají stejný charakter (tj. při výpočtu bez zaokrouhlování získáme po konečném počtu kroků přesné koeficienty). Zaokrouhlovací chyby však mohou vypočítané koeficienty hodně oddálit od jejich přesných hodnot.

4.3.1 Krylovova metoda

Charakteristickou rovnici můžeme zapsat ve tvaru

$$p(\lambda) = \lambda^n + \sum_{i=0}^{n-1} b_i \lambda^i = 0.$$

Z Cayleyovy-Hamiltonovy věty plyne

$$\mathbf{A}^n + \sum_{i=0}^{n-1} b_i \mathbf{A}^i = \mathbf{0}.$$

Tedy pro každý vektor y platí

$$\mathbf{A}^n y + \sum_{i=0}^{n-1} b_i \mathbf{A}^i y = \mathbf{0}. \quad (4.1)$$

Rovnice (4.1) je soustava n lineárních rovnic pro n neznámých b_0, \dots, b_{n-1} .

Poznámka 4.15. K výpočtu vektoru $\mathbf{A}^i y$ podle rovnice $\mathbf{A}^i y = \mathbf{A}(\mathbf{A}^{i-1} y)$ je třeba provést n^2 násobení, takže k sestavení soustavy (4.1) je třeba řádově n^3 operací.

4.3.2 Faddějevova-Leverrierova metoda

Definice 4.16. Stopou matice \mathbf{A} nazveme součet prvků na hlavní diagonále. Označujeme ji $tr \mathbf{A}$.

Metoda Faddějevova-Leverrierova se opírá o fakt, že součet vlastních čísel libovolné matice je roven její stopě. Algoritmus Faddějevovy-Leverrierovy metody počítá jednoduchým způsobem kořeny charakteristické rovnice.

Algoritmus 1. Je dána matice \mathbf{A} řádu n .

Krok 1: Položme $\mathbf{B}_1 = \mathbf{A}$ pak $p_1 = tr(\mathbf{B}_1)$

Krok 2: $\mathbf{B}_2 = \mathbf{A}(\mathbf{B}_1 - p_1 \mathbf{I})$ a $p_2 = \frac{1}{2} tr(\mathbf{B}_2)$

⋮

Krok n: $\mathbf{B}_n = \mathbf{A}(\mathbf{B}_{n-1} - p_{n-1} \mathbf{I})$ a $p_n = \frac{1}{n} tr(\mathbf{B}_n)$

Krok n+1: Charakteristický polynom je ve tvaru

$$p(\lambda) = \lambda^n - p_1 \lambda^{n-1} - \dots - p_{n-1} \lambda - p_n.$$

Poznámka 4.17. Pro inverzní matici \mathbf{A}^{-1} platí

$$\mathbf{A}^{-1} = \frac{1}{p_n} (\mathbf{B}_{n-1} - p_{n-1} \mathbf{I}).$$

Příklad 4.18. Najděte koeficienty charakteristického polynomu užitím Faddějevovy-Leverrierovy metody pro matici

$$\mathbf{A} = \begin{pmatrix} 8 & -1 & 3 & -1 \\ -1 & 6 & 2 & 0 \\ 3 & 2 & 9 & 1 \\ -1 & 0 & 1 & 7 \end{pmatrix}.$$

Řešení.

$$\mathbf{B}_1 = \mathbf{A} \Rightarrow \operatorname{tr}(\mathbf{B}_1) = 30 \Rightarrow p_1 = 30$$

$$\mathbf{B}_2 = \mathbf{A}(\mathbf{B}_1 - 30\mathbf{I}) = \begin{pmatrix} -165 & 22 & -42 & 18 \\ 22 & -139 & -33 & 3 \\ -42 & -33 & -175 & -17 \\ 18 & 3 & -17 & -159 \end{pmatrix}$$

$$\Rightarrow p_2 = \frac{1}{2}\operatorname{tr}(\mathbf{B}_2) = \frac{1}{2}(-638) = -319$$

$$\mathbf{B}_3 = \mathbf{A}(\mathbf{B}_2 + 319\mathbf{I}) = \begin{pmatrix} 1066 & -106 & 146 & -70 \\ -106 & 992 & 132 & -34 \\ 146 & 132 & 1087 & -67 \\ 70 & -34 & 67 & 1085 \end{pmatrix}$$

$$\Rightarrow p_3 = \frac{1}{3}\operatorname{tr}(\mathbf{B}_3) = \frac{1}{3}4230 = 1410$$

$$\mathbf{B}_4 = \mathbf{A}(\mathbf{B}_3 - 1410\mathbf{I}) = \begin{pmatrix} -2138 & 0 & 0 & 0 \\ 0 & -2138 & 0 & 0 \\ 0 & 0 & -2138 & 0 \\ 0 & 0 & 0 & -2138 \end{pmatrix}$$

$$\Rightarrow p_4 = \frac{1}{4}\operatorname{tr}(\mathbf{B}_4) = \frac{1}{4}(-8552) = -2138$$

Tím jsme získali všechny koeficienty charakteristického polynomu.

$$p(\lambda) = \lambda^4 - 30\lambda^3 + 319\lambda^2 - 1410\lambda + 2138.$$

□

Poznámka 4.19. F.-L. metoda je i přes jednoduchý algoritmus méně výhodná než Krylovova metoda, protože vyžaduje skutečně počítat matice \mathbf{A}_k pro $k = 1, \dots, n$.

4.4 Poloha a odhad vlastních čísel

4.4.1 Geršgorinovy věty

Přesná znalost vlastních čísel dané matice nás v některých praktických aplikacích nemusí zajímat. Často stačí jen znát polohu vlastních čísel v určitých oblastech komplexní roviny. Tyto informace můžeme získat i bez přímých výpočtů vlastních čísel dané matice. K nalezení polohy vlastních čísel lze použít následující větu.

Věta 4.20. Geršgorinova věta

Nechť $\mathbf{A} = \{a_{ij}\}$ je čtvercová matice řádu n . Definujme

$$r_i := \sum_{j=1, j \neq i}^n |a_{ij}| \quad i = 1, \dots, n. \quad (4.2)$$

Potom každé vlastní číslo λ matice \mathbf{A} splňuje aspoň jednu z následujících nerovností

$$|\lambda - a_{ii}| \leq r_i \quad i = 1, \dots, n. \quad (4.3)$$

Jinými slovy, všechna vlastní čísla matice \mathbf{A} leží v oblasti

$$K = \bigcup_{i=1}^n R_i, \quad (4.4)$$

kde R_i jsou kruhy o poloměru r_i a středu a_{ii} .

Důkaz. Nechť λ je vlastní číslo matice \mathbf{A} a \mathbf{x} je vlastní vektor odpovídající vlastnímu číslu λ . Potom ze vztahu $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ nebo ze vztahu $(\mathbf{A} - \lambda\mathbf{I}) = 0$ dostaneme

$$(\lambda - a_{ii})x_i = \sum_{j=1, j \neq i}^n a_{ij}x_j \quad i = 1, \dots, n$$

kde x_i je i -tý prvek vektoru \mathbf{x} .

Nechť x_k je (v absolutní hodnotě) největší prvek vektoru \mathbf{x} . Protože $\frac{|x_j|}{|x_k|} \leq 1$ pro $j \neq k$, je

$$|\lambda - a_{kk}| \leq \sum_{j=1}^n |a_{kj}| \frac{|x_j|}{|x_k|} \leq \sum_{j=1, j \neq k}^n |a_{kj}|. \quad (4.5)$$

Tedy λ leží v kruhu $|\lambda - a_{kk}| \leq r_k$. □

Definice 4.21. Kruhy $R_i := \{z : |z - a_{ii}| \leq r_i\}$ $i = 1, \dots, n$, se nazývají Geršgorinovy kruhy v komplexní rovině.

Poznámka 4.22. Věta 4.20 nám nezaručuje, že v každém kruhu bude nějaké vlastní číslo. Pouze říká, že vlastní čísla matice \mathbf{A} leží ve sjednocení Geršgorinových kruhů.

Následující věta polohu vlastních čísel upřesňuje.

Věta 4.23. Geršgorinova zobecněná věta

Nechť r Geršgorinových kruhů je disjunktních. Pak právě r vlastních čísel matice \mathbf{A} leží ve sjednocení těchto kruhů.

Poznámka 4.24. Určení polohy vlastního čísla dané matice pomocí Geršgorinových vět je poměrně jednoduché. Pro zajímavost uvedeme ještě jednu větu, která sice také určuje polohu vlastních čísel, ale její použití je už složitější a v určitých příkladech nepraktické.

Věta 4.25. *Nechť \mathbf{A} je čtvercová (obecně komplexní) matice n -tého řádu, nechť α je (komplexní) číslo, pro které stopa matice*

$$\operatorname{tr}((\alpha\mathbf{I} - \mathbf{A})^{-1}) \neq 0.$$

Pak v každém uzavřeném kruhu obsahujícím čísla α a $\tilde{\alpha}$, kde

$$\tilde{\alpha} = \alpha - \frac{n}{\operatorname{tr}((\alpha\mathbf{I} - \mathbf{A})^{-1})}$$

leží alespoň jedno vlastní číslo matice \mathbf{A} .

Definujme $r = \frac{n}{2\operatorname{tr}((\alpha\mathbf{I} - \mathbf{A})^{-1})}$, pak v kruhu o středu $\frac{(\alpha - \tilde{\alpha})}{2}$ a poloměru r leží alespoň jedno vlastní číslo matice \mathbf{A} .

Poznámka 4.26. Tato věta není obecně známa a vyplývá z vět o kořenech algebraické rovnice.

Příklad 4.27. Užitím Geršgorinových vět určete přibližnou polohu vlastních čísel komplexní matice

$$\begin{pmatrix} 1 & -\frac{1}{2} & \frac{1}{4} & -\frac{1}{4} \\ \frac{1}{4} & 1 + 2i & 0 & \frac{1}{4} \\ -\frac{1}{2} & \frac{1}{4} & -1 & \frac{1}{2} \\ \frac{1}{4} & -\frac{1}{2} & \frac{1}{2} & -2 - 2i \end{pmatrix}$$

Řešení.

$$r_1 = \sum_{i=1, i \neq 1}^4 |a_{1i}| = \frac{1}{2} + \frac{1}{4} + \frac{1}{2} = 1,$$

$$r_2 = \sum_{i=1, i \neq 2}^4 |a_{2i}| = \frac{1}{4} + 0 + \frac{1}{4} = \frac{1}{2},$$

$$r_3 = \sum_{i=1, i \neq 3}^4 |a_{3i}| = \frac{1}{2} + \frac{1}{4} + \frac{1}{2} = \frac{5}{4},$$

$$r_4 = \sum_{i=1, i \neq 4}^4 |a_{4i}| = \frac{1}{4} + \frac{1}{2} + \frac{1}{2} = \frac{5}{4}.$$

Potom

$$R_1 = \{z : |z - 1| \leq 1\},$$

$$R_2 = \left\{ z : |z - 1 - 2i| \leq \frac{1}{2} \right\},$$

$$R_3 = \left\{ z : |z + 1| \leq \frac{5}{4} \right\},$$

$$R_4 = \left\{ z : |z + 2 + 2i| \leq \frac{5}{4} \right\}.$$

Podle Geršgorinových vět tedy leží jedno vlastní číslo v kruhu K_1 , jedno v kruhu K_2 a zbylá dvě ve sjednocení kruhů $K_3 \cup K_4$. viz obr. DOPLNIT

Uveďme přesnou hodnotu vlastních čísel:

$$\begin{aligned}\lambda_1 &= 0.9824 + i2.0775, \\ \lambda_2 &= 0.9317 - i0.0535, \\ \lambda_3 &= -0.8879 - i0.08061, \\ \lambda_4 &= -2.0261 - i1.9433,\end{aligned}$$

což přesně odpovídá poloze určené pomocí Geršgorinových kruhů. □

Důsledek 4.28. Pro spektrální poloměr $\rho(\mathbf{A})$ matice \mathbf{A} platí nerovnost

$$\rho(\mathbf{A}) \leq \max_i \sum_{k=1}^n |a_{ik}|.$$

Důsledek 4.29. Pro spektrální poloměr matice \mathbf{A}^{-1} platí

$$\frac{1}{\rho(\mathbf{A}^{-1})} \geq \min_i (a_{ii} - \sum_{k=1, k \neq i}^n |a_{ik}|).$$

Poznámky ke Geršgorinově větě

1. Ze vztahu (4.5) pro maximální souřadnici $|x_i|$ můžeme získat odhad

$$|\lambda_i| \geq |a_{ii}| - \sum_{j \neq i} |a_{ij}| \geq \min_k (|a_{kk}| - \sum_{j \neq i} |a_{kj}|)$$

a

$$\min_i |\lambda_i| \geq (|a_{ii}| - \sum_{j \neq i} |a_{ij}|).$$

Pro diagonálně dominantní matici platí

$$0 < \min_i (|a_{ii}| - \sum_{j \neq i} |a_{ij}|) \leq |\lambda_i| \leq \max_i \sum_{j \neq i} |a_{ij}| = \|\mathbf{A}\|_\infty$$

2. K matici \mathbf{A} můžeme pomocí jednoduché podobnostní transformace $\mathbf{D}^{-1}\mathbf{A}\mathbf{D} = \mathbf{B}$, kde (\mathbf{D} je diagonální matice) získat podobnou matici \mathbf{B} , která má jiné Geršgorinovy kruhy. Potom všechna vlastní čísla leží v oblasti $K_{\mathbf{A}} \cap K_{\mathbf{B}}$. Cílem těchto transformací je rozklad oblasti K na souvislé komponenty, případně izolace jednoho kruhu, ve kterém pak můžeme zaručit existenci právě jednoho vlastního čísla.

3. Pokud $\det(\lambda\mathbf{I} - \mathbf{A}) = \det(\lambda\mathbf{I} - \mathbf{A}^T)$ můžeme vytvořit Geršgorinovy kruhy i pro matici \mathbf{A}^T a získat oblast $K_{\mathbf{A}} \cap K_{\mathbf{A}^T}$, ve které vlastní čísla leží.

Příklad 4.30. Matice $\mathbf{A}_1 = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}$ a $\mathbf{A}_2 = \begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix}$ mají stejné oblasti $K_{\mathbf{A}_1} = K_{\mathbf{A}_2} := K_{\mathbf{A}}$.

Obě matice mají stejná vlastní čísla, které jsou $\lambda_1 = 3.7321$, $\lambda_2 = 0.2679$.

4.4.2 Metody výpočtu dominantního vlastního čísla

Definice 4.31. Očíslujeme-li vlastní čísla dané matice \mathbf{A} tak, aby platilo

$$|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$$

(každé vlastní číslo píšeme tolikrát, kolik činí jeho násobnost), pak budeme vlastní číslo λ_1 nazývat *dominantní* vlastní číslo.

Mocninná metoda

Mocninná metoda je nejčastěji používanou metodou pro nalezení dominantního vlastního čísla a příslušného vlastního vektoru dané matice. Metoda je obzvláště vhodná pro řídké matice, protože spočívá pouze v násobení maticových vektorů. Základní předpoklad k užití této metody je, že daná matice má dominantní vlastní číslo λ_1 a že nemá nelineární elementární dělitele, tj. že existuje n lineárně nezávislých vlastních vektorů této matice, kde n je řád matice.

Konstrukce:

Nechť \mathbf{x} je libovolný vektor, $\mathbf{x} \in \mathbb{R}^n$, za předpokladu, že $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ je množina lineárně nezávislých vlastních vektorů, můžeme vektor \mathbf{x} vyjádřit jako lineární kombinaci vektorů $\mathbf{v}_i, i = 1, \dots, n$

$$\mathbf{x} = \sum_{i=1}^n \alpha_i \mathbf{v}_i. \quad (4.6)$$

Násobením obou stran rovnice (4.6) maticemi $\mathbf{A}, \mathbf{A}^2, \dots, \mathbf{A}^k$ dostaneme systém rovnic

$$\begin{aligned} \mathbf{A}\mathbf{x} &= \sum_{i=1}^n \alpha_i \mathbf{A}\mathbf{v}_i = \sum_{i=1}^n \alpha_i \lambda_i \mathbf{v}_i \\ \mathbf{A}^2\mathbf{x} &= \sum_{i=1}^n \alpha_i \mathbf{A}^2\mathbf{v}_i = \sum_{i=1}^n \alpha_i \lambda_i^2 \mathbf{v}_i \\ &\vdots \\ \mathbf{A}^k\mathbf{x} &= \sum_{i=1}^n \alpha_i \mathbf{A}^k\mathbf{v}_i = \sum_{i=1}^n \alpha_i \lambda_i^k \mathbf{v}_i \end{aligned} \quad (4.7)$$

Pro λ_1^k , které jsme vypočítali ze systému (4.7), dostáváme

$$\mathbf{A}^k \mathbf{x} = \lambda_1^k \sum_{i=1}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k \mathbf{v}_i.$$

Z předpokladu, že λ_1 je dominantní vlastní číslo a tedy $|\lambda_1| > |\lambda_j| \quad j = 2, \dots, n$, plyne, že pro $j > 1$ platí

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

a tedy

$$\lim_{k \rightarrow \infty} \mathbf{A}^k \mathbf{x} = \lim_{k \rightarrow \infty} \lambda_1^k \alpha_1 \mathbf{v}_1. \quad (4.8)$$

Tento proces bude konvergovat k nule, jestliže $|\lambda_1| < 1$, a divergovat, jestliže $|\lambda_1| \geq 1$, ovšem za předpokladu, že $\alpha_1 \neq 0$.

Poznámka 4.32. Popsaná konstrukce je i důkazem následující věty.

Věta 4.33. Von Misesova

Jestliže matice \mathbf{A} má n lineárně nezávislých vlastních vektorů $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ příslušných vlastním číslům $\lambda_1, \lambda_2, \dots, \lambda_n$ a je-li vlastní číslo λ_1 dominantní a pro vektor $\mathbf{x}_0 \in \mathbb{R}^n$ platí, že skalární součin $\langle \mathbf{x}_0, \mathbf{v}_1 \rangle \neq 0$. Pak

$$\lim_{k \rightarrow \infty} \left(\frac{\mathbf{A}^k \mathbf{x}_0}{\lambda_1^k} \right) = \alpha_1 \mathbf{v}_1. \quad (4.9)$$

Důsledek 4.34. Je-li \mathbf{y} libovolný vektor, který není ortogonální k vlastnímu vektoru \mathbf{v}_1 , plyne z věty 4.33, že

$$\lambda_1 = \lim_{k \rightarrow \infty} \left(\frac{\mathbf{y}^T \mathbf{x}_{k+1}}{\mathbf{y}^T \mathbf{x}_k} \right),$$

kde

$$\mathbf{x}_{k+1} = \mathbf{A} \mathbf{x}_k = \mathbf{A}^k \mathbf{x}_0.$$

Definice 4.35. Čísla $\mathbf{y}^T \mathbf{x}_{k+1} = \mathbf{y}^T \mathbf{A} \mathbf{x}_k$ se nazývají Schwarzovými konstantami.

Uvedeme si nyní postup při výpočtu dominantního vlastního čísla:

Algoritmus 2. Je zadána matice \mathbf{A} .

Krok 1: Zvolíme \mathbf{x}_0 .

Krok 2: Použijeme iterační formuli $\widehat{\mathbf{x}}_{k+1} = \mathbf{A} \mathbf{x}_k$, kde $\widehat{\mathbf{x}}_{k+1}$ je odhad, který dále zpřesníme.

Krok 3: $\mathbf{x}_{k+1} = \frac{\mathbf{x}_k}{\max_j \{|\mathbf{x}_k^{(j)}|\}} \Rightarrow \lambda_1^{(i)} = \max_{j=1, \dots, n} \{|\mathbf{x}_n^j|\}$

⋮

Krok n: Zastavení výpočtu po n krocích $\Rightarrow \lambda_1^{(n)} = \max_{j=1, \dots, n} \{|\mathbf{x}_n^{(j)}|\}$

nebo zastavení výpočtu pro $|\lambda_1^{(k+1)} - \lambda_1^{(k)}| < \delta$.

Nejčastější volbou počátečního vektoru \mathbf{x}_0 je vektor $\mathbf{x}_0 = (1, \dots, 1)^T$.

Příklad 4.36. Najděte dominantní vlastní číslo matice

$$\mathbf{A} = \begin{pmatrix} 4 & 2 & 3 & 2 & 2 \\ 3 & 0 & 4 & 1 & 3 \\ 1 & 2 & 5 & 0 & 4 \\ 2 & 6 & 0 & 2 & 1 \\ 3 & 6 & 5 & 1 & 2 \end{pmatrix}.$$

Řešení. Zvolíme $\mathbf{x}_0 = (1, 1, 1, 1, 1)^T$

$$\hat{\mathbf{x}}_1 = \mathbf{A}\mathbf{x}_0 = \begin{pmatrix} 13 \\ 11 \\ 12 \\ 11 \\ 17 \end{pmatrix}, \quad \lambda_1^{(1)} = 17, \quad \mathbf{x}_1 = \begin{pmatrix} 0.7647 \\ 0.6471 \\ 0.7059 \\ 0.6471 \\ 1 \end{pmatrix},$$

$$\hat{\mathbf{x}}_2 = \mathbf{A}\mathbf{x}_1 = \begin{pmatrix} 9.7647 \\ 8.7647 \\ 9.5882 \\ 7.7059 \\ 12.3529 \end{pmatrix}, \quad \lambda_1^{(2)} = 12.3529, \quad \mathbf{x}_2 = \begin{pmatrix} 0.7905 \\ 0.7095 \\ 0.7762 \\ 0.6238 \\ 1 \end{pmatrix}.$$

Po dalších krocích získáme

$$\mathbf{x}_{10} = \begin{pmatrix} 0.7731 \\ 0.6957 \\ 0.7735 \\ 0.6125 \\ 1 \end{pmatrix},$$

$$\hat{\mathbf{x}}_{11} = \mathbf{A}\mathbf{x}_{10} = \begin{pmatrix} 10.0285 \\ 9.0247 \\ 10.0307 \\ 7.9454 \\ 12.9722 \end{pmatrix}, \quad \lambda_1^{(11)} = 12.9722.$$

Vlastní čísla matice \mathbf{A} jsou

$$\lambda_1 = 12.9722, \lambda_2 = 3.8755, \lambda_3 = -3.0794, \lambda_{4,5} = -0.3851 \pm i0.2130.$$

Je tedy vidět, že po jedenácti krocích jsme dostali přesné řešení zadaného příkladu. \square

Příklad 4.37. Pro matici $\mathbf{A} = \begin{pmatrix} 1.5 & -2 & 0.4 \\ 3 & 0.86 & -0.5 \\ 2 & 1.5 & 1.5 \end{pmatrix}$ však metoda nebude konvergovat, protože číselné hodnoty budou oscilovat. V tomto případě máme

$$\lambda_1 = 2.13746 \quad \lambda_{2,3} = 0.86127 \pm i2.52118 \quad \Rightarrow |\lambda_{2,3}| = 2.66423.$$

Absolutní hodnoty dvou vlastních čísel jsou si rovny. Mocnná metoda nedokáže určit dominantní vlastní číslo.

Poznámka 4.38. Nevýhody mocnné metody:

- Problematický odhad chyby.
- Není zaručena konvergence metody, protože obvykle v praxi nevíme, zda jsou splněny předpoklady mocnné metody.
- Otázka volby počáteční hodnoty \mathbf{x}_0 . Bude-li vektor \mathbf{x}_0 takovou lineární kombinací vlastních vektorů, že koeficient u vlastního vektoru odpovídajícího dominantnímu vlastnímu číslu bude roven 0, potom mocnná metoda nevypočte dominantní vlastní číslo.

Rychlost konvergence mocnné metody závisí hlavně na vhodné volbě vektoru \mathbf{x}_0 a na velikosti podílu $\frac{|\lambda_2|}{|\lambda_1|}$.

4.4.3 Metoda Rayleighova podílu

Metoda Rayleighova podílu je modifikovanou mocnnou metodou a zaměřuje se na výpočet dominantního vlastního čísla symetrické matice.

Dále tedy budeme vždy předpokládat, že matice \mathbf{A} je symetrická. Potom jsou všechna vlastní čísla reálná a vlastní vektory příslušné různým vlastním číslům jsou ortogonální (tj. $\mathbf{v}_i^T \mathbf{v}_j = 0$ pro $i \neq j$). Dále budeme předpokládat, že vlastní vektory jsou ortonormální (tj. že platí $\mathbf{v}_i^T \mathbf{v}_i = 1$).

Metodu lze odvodit takto:

1. Zvolíme \mathbf{x}_0 jako lineární kombinaci vlastních vektorů

$$\mathbf{x}_0 = \sum_{i=1}^n \alpha_i \mathbf{v}_i.$$

2. Sestrojíme posloupnost

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1}, \text{ tj. } \mathbf{x}_k = \mathbf{A}^k \mathbf{x}_0,$$

$$\mathbf{x}_k = \alpha_1 \mathbf{A}^k \mathbf{v}_1 + \dots + \alpha_n \mathbf{A}^k \mathbf{v}_n.$$

3. Platí $\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_i$, potom

$$\mathbf{x}_k = \alpha_1\lambda_1^k\mathbf{v}_1 + \alpha_2\lambda_2^k\mathbf{v}_2 + \dots + \alpha_n\lambda_n^k\mathbf{v}_n,$$

kde λ_1 je dominantní vlastní číslo.

4. Dostaneme

$$\mathbf{x}_k = \lambda_1^k \left[\alpha_1\mathbf{v}_1 + \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k \mathbf{v}_i \right].$$

Součet v hranaté závorce označíme jako \mathbf{w}_k , $\mathbf{w}_k \rightarrow \mathcal{O}$.

5. Vyjádříme součin $\mathbf{x}_k^T\mathbf{x}_k$

$$\begin{aligned} \mathbf{x}_k^T\mathbf{x}_k &= \lambda_1^k \left[\alpha_1\mathbf{v}_1^T + \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k \mathbf{v}_i^T \right] \lambda_1^k \left[\alpha_1\mathbf{v}_1 + \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k \mathbf{v}_i \right] = \\ &= \lambda_1^{2k} \left[\alpha_1^2 + \sum_{i=2}^n \alpha_i^2 \left(\frac{\lambda_i}{\lambda_1} \right)^{2k} \right] = \lambda_1^{2k} [\alpha_1^2 + \mathbf{w}_k^T\mathbf{w}_k] \end{aligned}$$

a součin $\mathbf{x}_k^T\mathbf{x}_{k+1}$

$$\begin{aligned} \mathbf{x}_k^T\mathbf{x}_{k+1} &= \lambda_1^k \left[\alpha_1\mathbf{v}_1^T + \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k \mathbf{v}_i^T \right] \lambda_1^{k+1} \left[\alpha_1\mathbf{v}_1 + \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^{k+1} \mathbf{v}_i \right] = \\ &= \lambda_1^{2k+1} \left[\alpha_1^2 + \sum_{i=2}^n \alpha_i^2 \left(\frac{\lambda_i}{\lambda_1} \right)^{2k+1} \right] = \lambda_1^{2k+1} [\alpha_1^2 + \mathbf{w}_k^T\mathbf{w}_{k+1}]. \end{aligned}$$

Dostáváme

$$\lim_{k \rightarrow \infty} \frac{\mathbf{x}_k^T\mathbf{A}\mathbf{x}_k}{\mathbf{x}_k^T\mathbf{x}_k} = \lim_{k \rightarrow \infty} \frac{\mathbf{x}_k^T\mathbf{A}\mathbf{x}_{k+1}}{\mathbf{x}_k^T\mathbf{x}_{k+1}} = \frac{\lambda_1^{2k+1}(\alpha_1^2 + \overbrace{\mathbf{w}_k^T\mathbf{w}_{k+1}}^{\rightarrow 0})}{\lambda_1^{2k}(\alpha_1^2 + \underbrace{\mathbf{w}_k^T\mathbf{w}_{k+1}}_{\rightarrow 0})} = \lambda_1$$

Poznámka 4.39. Součin $\mathbf{w}_k^T\mathbf{w}_k$ konverguje k nule pro $k \rightarrow \infty$ dvakrát rychleji než \mathbf{w}_k k nulovému vektoru, z čehož vyplývá, že metoda Raleighova podílu bude rychlejší než mocinná metoda.

Příklad 4.40. Metodou Rayleighova podílu určete dominantní vlastní číslo matice

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Řešení. Volíme $\mathbf{x}_0 = (1, \quad 1, \quad 1)^T$

$$\mathbf{x}_1 = \mathbf{A}\mathbf{x}_0 = \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix}, \quad \lambda_1^{(1)} = \frac{\mathbf{x}_0^T \mathbf{x}_1}{\mathbf{x}_0^T \mathbf{x}_0} = 2.3333$$

$$\mathbf{x}_2 = \mathbf{A}\mathbf{x}_1 = \begin{pmatrix} 5 \\ 7 \\ 5 \end{pmatrix}, \quad \lambda_1^{(2)} = \frac{\mathbf{x}_1^T \mathbf{x}_2}{\mathbf{x}_1^T \mathbf{x}_1} = 2.4118$$

$$\mathbf{x}_3 = \mathbf{A}\mathbf{x}_2 = \begin{pmatrix} 12 \\ 17 \\ 12 \end{pmatrix}, \quad \lambda_1^{(3)} = \frac{\mathbf{x}_2^T \mathbf{x}_3}{\mathbf{x}_2^T \mathbf{x}_2} = 2.4142$$

Vlastní čísla matice \mathbf{A} jsou

$$\lambda_1 = 2.4142, \quad \lambda_2 = 1, \quad \lambda_3 = -0.4142.$$

Tedy už po třech krocích jsme dostali přesné řešení. Všimněte si, že všechna vlastní čísla jsou reálná. Jde o důsledek toho, že jsme pracovali s reálnou symetrickou maticí. \square

4.4.4 Výpočet dalších vlastních čísel mocninnou metodou

Pokud již známe vlastní číslo λ_1 matice \mathbf{A} a k němu příslušný vlastní vektor \mathbf{v}_1 , můžeme vypočítat následující vlastní číslo λ_2 a vlastní vektor \mathbf{v}_2 opět mocninnou metodou, pro kterou použijeme na redukovanou matici.

Věta 4.41. O redukci

Nechť $\lambda_1 \neq 0$ je vlastní číslo matice \mathbf{A} s vlastním vektorem \mathbf{v}_1 a vektor \mathbf{x} je libovolný vektor s vlastností $\mathbf{x}^T \mathbf{v}_1 = 1$. Potom vlastní čísla matice

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

jsou $0, \lambda_2, \dots, \lambda_n$, kde $\lambda_1, \lambda_2, \dots, \lambda_n$ jsou vlastní čísla matice \mathbf{A} .

Důkaz. Nechť $\mathbf{e}^1 = (1, 0, \dots, 0)$ je jednotkový vektor s 1 na prvním místě a

$$\mathbf{J} = \mathbf{V}^{-1} \mathbf{A} \mathbf{V} = \begin{pmatrix} \lambda_1 & \delta_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \delta_2 & \cdots & \\ \vdots & \vdots & \vdots & \ddots & \\ \vdots & \vdots & \vdots & \ddots & \delta_{n-1} \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

je Jordanův tvar matice, kde $\delta_i \in \{0, 1\}, i = 1, \dots, n - 1$. Jsou-li $\mathbf{v}_1, \dots, \mathbf{v}_n$ sloupce matice \mathbf{V} , potom matice

$$\mathbf{C} = \mathbf{V}^{-1}\mathbf{B}\mathbf{V}$$

má tvar

$$\begin{aligned} \mathbf{C} &= \mathbf{J} - \lambda_1 \mathbf{V}^{-1} \mathbf{x}^T \mathbf{V} = \mathbf{J} - \lambda_1 \mathbf{e}^1 (\mathbf{x}^T \mathbf{v}_1, \dots, \mathbf{x}^T \mathbf{v}_n) = \\ &= \mathbf{J} - \lambda_1 \begin{pmatrix} 1 & \mathbf{x}^T \mathbf{v}_2 \dots \mathbf{x}^T \mathbf{v}_n \\ \mathbf{0}_{1,n-1} & \mathbf{0}_{n-1,n-1} \end{pmatrix} = \\ &= \begin{pmatrix} 0 & \delta_1 - \lambda_1 \mathbf{x}^T \mathbf{v}_2 & -\lambda_1 \mathbf{x}^T \mathbf{v}_3 & \dots & -\lambda_1 \mathbf{x}^T \mathbf{v}_n \\ 0 & \lambda_2 & \delta_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \delta_{n-1} \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix}, \end{aligned}$$

což větu dokazuje (vlastní čísla jsou na diagonále). \square

Výběr vektoru \mathbf{x} :

Věta o redukci zaručuje široké možnosti výběru vektoru \mathbf{x} . Např.

1. Wielandtova redukce

Výhoda této metody je v tom, že v každé další fázi pracujeme s menší maticí a provádíme méně výpočtů. Položíme

$$\mathbf{x} = \frac{1}{\lambda_1} v_1^j r_j^T$$

kde r_j je j -tý řádek matice \mathbf{A} a $v_1^j \neq 0$. Index j vybereme tak, aby odpovídal největší složce vektoru \mathbf{x} .

2. Hotellingova redukce

Zde položíme $\mathbf{x} = \mathbf{y}_1$, kde \mathbf{y}_1 je levý vlastní vektor k λ_1 a je normalizován tak, že platí $\mathbf{y}_1^T \mathbf{x} = 1$. Protože \mathbf{y}_1 obvykle neznáme, používá se tato metoda nejnázve u symetrických matic, v tomto případě je $\mathbf{x}_i = \mathbf{v}_i$.

4.4.5 Výpočet vlastních čísel a vlastních vektorů symetrických matic

Jacobiho metoda

Pomocí *Jacobiho metody* můžeme najít všechna vlastní čísla a jim odpovídající vlastní vektory symetrické matice \mathbf{A} . Metoda je vhodná hlavně pro *plné matice*, tj. matice bez nulových prvků.

Definice 4.42. Báze vektorového prostoru se nazývá *ortogonální*, jestliže je složena z navzájem ortogonálních vektorů. Báze se nazývá *ortonormální*, jestliže je ortogonální a norma každého z vektorů je rovna 1.

Definice 4.43. Matice se nazývá *ortogonální*, jestliže je složena z lineárně nezávislých řádků, které jsou navzájem ortogonální a norma každého řádku je rovna 1.

Poznámka 4.44. Ortonormální báze zapsaná do matice tvoří ortogonální matici.

Věta 4.45. *Nechť A je ortogonální matice. Potom $A^{-1} \equiv A^T$.*

Součin dvou ortogonálních matic téhož řádu je opět ortogonální maticí.

Nechť \mathbf{A} je symetrická matice, potom existuje ortonormální báze složená z vlastních vektorů taková, že

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

kde λ_i jsou reálná vlastní čísla matice \mathbf{A} a \mathbf{B} je matice složená z báze ortonormálních vlastních vektorů.

Na začátku *Jacobiho metody* položíme $\mathbf{A} = \mathbf{A}_1$ a sestrojíme posloupnost elementárních ortogonálních matic $\{\mathbf{S}_k\}, k = 1, 2, \dots$, takovou, aby

$$\mathbf{A}_{k+1} = \mathbf{S}_k^T \mathbf{A}_k \mathbf{S}_k = (\mathbf{S}_1 \dots \mathbf{S}_k)^T \mathbf{A} (\mathbf{S}_1 \dots \mathbf{S}_k), \quad k = 1, 2, \dots$$

konvergující k diagonální matici $\text{diag}(\lambda_1, \dots, \lambda_n)$. Protože všechny matice \mathbf{A}_{k+1} jsou podobné matici \mathbf{A} , potom mají stejná vlastní čísla jako matice \mathbf{A} .

Nechť \mathbf{S} je matice tvaru

$$\mathbf{S} = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & \cos \alpha & \cdots & \sin \alpha & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & -\sin \alpha & \cdots & \cos \alpha & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix},$$

je to tzn. rovinná rotace nebo Givensova transformace, kde prvky $\cos \alpha$ jsou na pozicích (p,p) a (q,q) , $\sin \alpha$ na pozici (p,q) a $-\sin \alpha$ na pozici (q,p) . Pak platí věta

Věta 4.46. *Nechť p, q jsou přirozená čísla, $1 \leq p < q \leq n$, α je reálné číslo, nechť \mathbf{S} je ortogonální matice. Potom platí*

1. Je-li $\mathbf{A} = (a_{ij})$ symetrická, potom je $\mathbf{B} = \mathbf{S}^T \mathbf{A} \mathbf{S} = (b_{ij})$ symetrická a

$$\sum_{i,j=1}^n b_{ij}^2 = \sum_{i,j=1}^n a_{ij}^2$$

2. Je-li $a_{pq} \neq 0$, existuje jediné $\alpha \in \langle -\pi/4, 0 \rangle \cup (0, \pi/4)$ tak, že

$$b_{pq} = 0.$$

Jedná se o jediné řešení rovnice

$$\cot g 2\alpha = \frac{a_{qq} - a_{pp}}{2a_{pq}}$$

ležící v této množině. Potom

$$\sum_{i=1}^n b_{ii}^2 = \sum_{i=1}^n a_{ii}^2 + 2a_{pq}^2.$$

Důkaz. 1. Protože $\mathbf{A} = \mathbf{S}^T \cdot \mathbf{B} \cdot \mathbf{S}$ a víme, že pro dvě matice platí, že

$$\text{tr}(\mathbf{K} \cdot \mathbf{L}) = \text{tr}(\mathbf{L} \cdot \mathbf{K}),$$

máme

$$\begin{aligned} \sum_{i,j=1}^n a_{ij}^2 &= \text{tr}(\mathbf{A}^T \cdot \mathbf{A}) = \text{tr}(\mathbf{S} \cdot \mathbf{B}^T \cdot \mathbf{S}^T \cdot \mathbf{S} \cdot \mathbf{B} \cdot \mathbf{S}^T) = \\ &= \text{tr}(\mathbf{S} \cdot \mathbf{B}^T \cdot \mathbf{B} \cdot \mathbf{S}^T) = \text{tr}(\mathbf{B} \cdot \mathbf{S}^T \cdot \mathbf{S} \cdot \mathbf{B}^T) = \text{tr}(\mathbf{B} \cdot \mathbf{B}^T) = \\ &= \sum_{i,j=1}^n b_{ij}^2. \end{aligned}$$

2. Transformace na pozicích (p,q);(q,q);(p,p);(q,p) má tvar

$$\begin{aligned} \begin{bmatrix} b_{pp} & b_{pq} \\ b_{qp} & b_{qq} \end{bmatrix} &= \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \cdot \begin{bmatrix} a_{pp} & a_{pq} \\ a_{qp} & a_{qq} \end{bmatrix} \cdot \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix} \\ &= \begin{bmatrix} a_{pp} \cos \alpha - a_{pq} \sin \alpha & a_{pq} \cos \alpha - a_{qq} \sin \alpha \\ a_{pp} \sin \alpha + a_{pq} \cos \alpha & a_{pq} \sin \alpha + a_{qq} \cos \alpha \end{bmatrix} \cdot \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix} \end{aligned}$$

a tedy

•

$$\begin{aligned} b_{pp} &= a_{pp} \cos^2 \alpha - 2a_{pq} \sin \alpha \cos \alpha + a_{qq} \sin^2 \alpha = \\ &= a_{pp} \cos^2 \alpha + a_{qq} \sin^2 \alpha - a_{pq} \sin 2\alpha. \end{aligned}$$

•

$$\begin{aligned} b_{pq} &= b_{qp} = \\ &= a_{pp} \cos \alpha \sin \alpha + a_{pq} \sin^2 \alpha + a_{pq} \cos^2 \alpha - a_{qq} \sin \alpha \cos \alpha = \\ &= a_{pq} \cos 2\alpha + 1/2(a_{pq} - a_{qq}) \sin 2\alpha. \end{aligned}$$

•

$$\begin{aligned} b_{qq} &= a_{pp} \sin^2 \alpha + 2a_{pq} \sin \alpha \cos \alpha + a_{qq} \cos^2 \alpha = \\ &= a_{pp} \sin^2 \alpha + a_{qq} \cos^2 \alpha + a_{pq} \sin 2\alpha. \end{aligned}$$

Stejně jako v (1)

$$a_{pp}^2 + a_{qq}^2 + 2a_{pq}^2 = b_{pp}^2 + b_{qq}^2 + 2b_{pq}^2$$

pro libovolné α .

Zvolíme-li α tak, aby platilo

$$\cotg 2\alpha = -\frac{a_{pp} - a_{qq}}{2a_{pq}}$$

je $b_{pq} = b_{qp} = 0$ a tedy

$$b_{pp}^2 + b_{qq}^2 = a_{pp}^2 + a_{qq}^2 + 2a_{pq}^2$$

ostatní $a_{ii} = b_{ii}$ pro $i \neq p, q$. □

Poznámka 4.47. Při transformaci

$$\mathbf{A} \rightarrow \mathbf{B} = \mathbf{S}^T \cdot \mathbf{A} \cdot \mathbf{S}$$

se mění pouze p -té a q -té řádky a sloupce, přesněji pro libovolné α :

•

$$b_{ij} = a_{ij} \text{ pro } i \neq p, q \quad \text{a} \quad j \neq p, q$$

•

$$b_{pi} = b_{ip} = a_{pi} \cos \alpha - a_{qi} \sin \alpha \text{ pro } i \neq p, q$$

•

$$b_{qi} = b_{iq} = a_{pi} \sin \alpha - a_{qi} \cos \alpha \text{ pro } i \neq p, q$$

•

$$b_{pp} = a_{pp} \cos^2 \alpha + a_{qq} \sin^2 \alpha - a_{pq} \sin 2\alpha$$

•

$$b_{qq} = a_{pp} \sin^2 \alpha + a_{qq} \cos^2 \alpha + a_{pq} \sin 2\alpha$$

•

$$b_{pq} = b_{qp} = a_{pq} \cos 2\alpha + \frac{1}{2}(a_{pp} - a_{qq}) \sin 2\alpha$$

Použijeme-li vztahy mezi goniometrickými funkcemi, lze prvky matice \mathbf{B} vyjádřit pomocí prvků matice \mathbf{A} .

Postup výpočtu:

1. Nejprve položíme

$$K = \frac{a_{qq} - a_{pp}}{2a_{pq}} \quad (= \cotg 2\alpha)$$

2. Označíme-li $t = \operatorname{tg} \alpha$, je

$$t = \begin{cases} \text{kořen} & t^2 + 2Kt - 1 \quad \text{pro } K \neq 0 \\ 1 & \text{pro } K = 0 \end{cases}$$

3. Dále

$$c = \frac{1}{\sqrt{1+t^2}} \quad (= \cos \alpha),$$

$$s = \frac{t}{\sqrt{1+t^2}} \quad (= \sin \alpha).$$

4. Pro prvky matice \mathbf{B} platí vztahy:

$$b_{pi} = b_{iq} = c \cdot a_{pi} - s \cdot a_{qi}, \quad i \neq p, q,$$

$$b_{qi} = b_{ip} = c \cdot a_{qi} + s \cdot a_{pi}, \quad i \neq p, q,$$

$$b_{pi} = b_{ip} = a_{pp} - t \cdot a_{pq},$$

$$b_{pi} = b_{ip} = a_{qq} + t \cdot a_{pq}.$$

Uved' me odvození na příklad pro b_{qq}

$$b_{qq} = a_{pp} \sin^2 \alpha + a_{qq}(1 - \sin^2 \alpha) + a_{pq} \sin 2\alpha =$$

$$a_{qq} - (a_{qq} + a_{pp}) \sin^2 \alpha + a_{pq} \sin 2\alpha =$$

$$a_{qq} + a_{pq}(\sin 2\alpha - 2 \cotg 2\alpha \sin^2 \alpha).$$

Protože

$$-2 \cotg 2\alpha \sin^2 \alpha + \sin 2\alpha = \frac{\sin^2 2\alpha - 2 \cos^2 2\alpha \sin^2 \alpha}{2 \sin \alpha \cos \alpha}$$

a dále čitatel

$$4 \sin^2 \alpha \cos^2 \alpha - 2 \sin^2 \alpha \cos^2 \alpha + 2 \sin^4 \alpha =$$

$$2 \sin^2 \alpha (\sin^2 \alpha + \cos^2 \alpha) = 2 \sin^2 \alpha,$$

je

$$b_{qq} = a_{qq} + \frac{\sin \alpha}{\cos \alpha} a_{pq} = a_{qq} + t \cdot a_{pq}.$$

Krok Jacobiho metody:

Máme-li sestrojenou matici $\mathbf{A}_k = [a_{ij}^{(k)}]$, vybereme (p, q) tak, aby

$$a_{p,q}^{(k)} \neq 0.$$

Sestrojíme \mathbf{S}_k jako ve větě 4.46, určíme $\alpha \in (-\pi/4, 0) \cup (0, \pi/4)$ tak, aby

$$\cotg 2\alpha_k = \frac{a_{qq}^{(k)} - a_{pp}^{(k)}}{2a_{pq}^{(k)}},$$

položíme

$$\mathbf{A}_{k+1} = \mathbf{S}_k^T \cdot \mathbf{A} \cdot \mathbf{S}_k = [a_{ij}^{(k+1)}].$$

Strategie pro volbu (p,q):

1. Klasická Jacobiho metoda:

Zvolíme (p,q) taková, aby platilo

$$|a_{pq}^{(k)}| = \max_{i \neq j} |a_{ij}^{(k)}|$$

a (p,q) se mění pro různá k .

2. Cyklická Jacobiho metoda:

Nulují se všechny nediagonální prvky cyklickou smyčkou, např. (p,q) volíme

$$(1, 2) \quad (1, 3) \quad \dots \quad (1, n); (2, 3) \quad \dots \quad (2, n); \quad \dots \quad ; (n-1, n).$$

Zřejmě, je-li některý prvek nulový, postupujeme dále (tj. volíme $\alpha_k = 0$ nebo $\mathbf{S}_k = \mathbf{I}$).

3. Prahová Jacobiho metoda:

Postupujeme jako u cyklické Jacobiho metody, ale nediagonální prvky, které jsou v absolutní hodnotě menší než „jistá“ mez, která se zmenšuje s každou smyčkou, se nenulují, tj. přeskočíme je.

Poznámka 4.48. Co se týče konvergence, ukážeme myšlenku důkazu pro nejjednodušší případ. Označíme P_n množinu všech permutací čísel $1, 2, \dots, n$.

Věta 4.49. *Posloupnost matic $\{\mathbf{A}_k\}_{k=1}^{\infty}$ získaných klasickou Jacobiho metodou je konvergentní,*

$$\lim_{k \rightarrow \infty} \mathbf{A}_k = \text{diag}(\lambda_{s(i)})$$

pro jistou permutaci $s \in P_n$.

K důkazu potřebujeme následující lemma.

Lemma 4.50. Bud' \mathbf{X} konečnědimenzionální normovaný vektorový prostor, $\{\mathbf{x}_k\}$ ohraničená posloupnost v \mathbf{X} , která má pouze konečný počet hromadných bodů, nechť

$$\lim_{k \rightarrow \infty} \|\mathbf{x}_{k+1} - \mathbf{x}_k\| = 0.$$

Potom je posloupnost $\{\mathbf{x}_k\}$ konvergentní.

Důkaz. Věty 4.49

Označme $\mathbf{A}_k = [a_{ij}^{(k)}] = \mathbf{D}_k + \mathbf{B}_k$, $\mathbf{D}_k = \text{diag}(a_{ii}^{(k)})$.

Nejprve dokážeme, že $\lim_{k \rightarrow \infty} \mathbf{B}_k = \mathbf{0}$. Označme

$$\Omega_k = \sum_{i \neq j} |a_{ij}^{(k)}|^2,$$

potom zřejmě

$$\Omega_k \leq n(n-1)|a_{pq}^{(k)}|^2$$

protože máme $n(n-1)$ nediagonálních prvků a prvek $|a_{pq}^{(k)}|$ z nich byl maximální. Dále podle věty 4.46

$$\Omega_{k+1} = \Omega_k - 2|a_{ij}^{(k)}|^2,$$

tedy

$$\Omega_{k+1} \leq \left(1 - \frac{2}{n(n-1)}\right)\Omega_k$$

tj.

$$\lim_{k \rightarrow \infty} \Omega_k = 0.$$

Nyní dokážeme, že $\lim_{k \rightarrow \infty} (\mathbf{D}_{k+1} - \mathbf{D}_k) = \mathcal{O}$. Pro diagonální prvky matice \mathbf{A}_{k+1} platí

$$a_{ii}^{(k+1)} - a_{ii}^{(k)} = \begin{cases} 0 & i \neq p, q \\ -\tan \alpha_k a_{pq}^{(k)} & i = p \\ \tan \alpha_k a_{pq}^{(k)} & i = q. \end{cases}$$

Protože $|\alpha_k| \leq \pi/4$ a $\lim_{k \rightarrow \infty} a_{pq}^{(k)} = 0$, je důkaz proveden.

Nechť $\{\mathbf{D}_{k'}\}$ je posloupnost, která konverguje k matici \mathbf{D} , potom také $\lim_{k' \rightarrow \infty} \mathbf{A}_{k'} = \mathbf{D}$, protože

$$\mathbf{A}_{k'} = \mathbf{D}_{k'} + \mathbf{B}_{k'} \quad \text{a} \quad \lim_{k' \rightarrow \infty} \mathbf{B}_{k'} = 0.$$

Tedy

$$\det(\lambda \mathbf{I} - \mathbf{D}) = \lim_{k' \rightarrow \infty} \det(\lambda \mathbf{I} - \mathbf{A}_{k'}) = \det(\lambda \mathbf{I} - \mathbf{A}).$$

Matice $\mathbf{A}_{k'}$ a \mathbf{A} jsou podobné, tedy

$$\det(\lambda \mathbf{I} - \mathbf{A}_{k'}) = \det(\lambda \mathbf{I} - \mathbf{a})$$

pro všechna k' .

Tedy \mathbf{D} a \mathbf{A} mají stejné charakteristické polynomy, tedy i stejná vlastní čísla. \mathbf{D} proto musí být diagonální, $\mathbf{D} = \text{diag}(\lambda_{s(i)})$.

Posloupnost $\{\mathbf{D}_k\}$ je ohraničená, neboť

$$\|\mathbf{D}_k\|_2 = \left(\sum_{i,j=1}^n |d_{ij}^{(k)}|^2 \right)^{1/2} \leq \left(\sum_{i,j=1}^n |a_{ij}^{(k)}|^2 \right)^{1/2} = \|\mathbf{A}_k\|_2 = \|\mathbf{A}\|_2.$$

Jsou tedy splněny předpoklady lemmatu a posloupnost $\{\mathbf{A}_k\}$ konverguje. \square

Příklad 4.51. Klasickou Jacobiho metodou určete všechna vlastní čísla matice

$$\mathbf{A} = \begin{pmatrix} 8 & -1 & 3 & -1 \\ 1 & 6 & 2 & 0 \\ 3 & 2 & 9 & 1 \\ -1 & 0 & 1 & 7 \end{pmatrix}$$

Řešení. Maximální prvek (v absolutní hodnotě) je 3 na pozici (1,3) $\Rightarrow p = 1 \quad q = 3$

$$K = \frac{a_{33} - a_{11}}{2a_{13}} = \frac{1}{6} \neq 0 \Rightarrow$$

t je kořen (s menší absolutní hodnotou) polynomu

$$t^2 + \frac{1}{3}t - 1 = 0,$$

$$t = 0.84712708838304,$$

$$c = \frac{1}{\sqrt{1+t^2}} = 0.76301998247272, \quad s = \frac{t}{\sqrt{1+t^2}} = 0.64637489613020,$$

$$b_{13} = b_{31} = 0,$$

$$b_{11} = a_{11} - t \cdot a_{13} = 5.45861873485088,$$

$$b_{33} = a_{33} + t \cdot a_{13} = 11.54138126514912,$$

$$b_{12} = c \cdot a_{12} - s \cdot a_{32} = -2.05576977473312 = b_{21},$$

$$b_{14} = c \cdot a_{14} - s \cdot a_{34} = -1.40939487860292 = b_{41},$$

$$b_{32} = c \cdot a_{32} + s \cdot a_{12} = 0.87966506881525 = b_{23},$$

$$b_{34} = c \cdot a_{34} + s \cdot a_{14} = 0.11664508634253 = b_{43},$$

$$b_{22} = a_{22} \quad b_{44} = a_{44} \quad b_{42} = b_{24} = a_{24}.$$

Pak dostaneme matici

$$\begin{pmatrix} 5.45861873485088 & -2.05576977473312 & 0 & -1.40939487860292 \\ -2.05576977473312 & 6 & 0.87966506881525 & 0 \\ 0 & 0.87966506881525 & 11.54138126514912 & 0.11664508634253 \\ -1.40939487860292 & 0 & 0.11664508634253 & 7 \end{pmatrix}.$$

Nyní opět vybereme maximální prvek a stejným způsobem postupujeme dále. Po 7 krocích se dostaneme k matici

$$B_7 = \begin{pmatrix} 3.79407218081762 & 0.07086171427580 & -0.00393661412823 & 0.00516622055919 \\ 0.07086171427580 & 6.40219536739289 & -0.08436498867668 & -0.06428537120075 \\ -0.00393661412823 & -0.08436498867668 & 11.76776520507119 & 0 \\ 0.00516622055919 & -0.06428537120075 & 0 & 8.03596724671830 \end{pmatrix}.$$

Zde už je vidět, že nediagonální prvky konvergují k nule. Po dalších sedmi krocích už dostaneme diagonální matici

$$B = \begin{pmatrix} 3.2957 & 0 & 0 & 0 \\ 0 & 6.5923 & 0 & 0 \\ 0 & 0 & 11.7043 & 0 \\ 0 & 0 & 0 & 8.4077 \end{pmatrix},$$

kde diagonální prvky jsou vlastní čísla zadané matice \mathbf{A} . □

Nyní se budeme zabývat konvergencí vlastních vektorů klasické Jacobiho metody, kterou dokážeme pomocí následující věty. Připomeňme, že

$$\mathbf{A}_{k+1} = \mathbf{S}_k^T \cdot \mathbf{A}_k \cdot \mathbf{S}_k = \mathbf{Q}_k^T \cdot \mathbf{A} \cdot \mathbf{Q}_k$$

kde $\mathbf{Q}_k = \mathbf{S}_1 \dots \mathbf{S}_k$.

Věta 4.52. *Předpokládejme, že všechna vlastní čísla matice \mathbf{A} jsou vzájemně různá. Potom posloupnost matic \mathbf{Q}_k , $k = 1, 2, \dots$, konstruovaných klasickou Jacobiho metodou konverguje k ortogonální matici, jejíž sloupce tvoří ortonormální množinu vlastních vektorů matice \mathbf{A} .*

Důkaz. Opět použijeme lemma 4.50. Nejprve ověříme jeho předpoklady. $\{\mathbf{Q}_k\}$ má pouze konečný počet hromadných bodů, které jsou nutně ve tvaru

$$[\pm p_{s(1)} \pm p_{s(2)} \pm \dots \pm p_{s(n)}], \quad s \in P_n,$$

kde p_1, \dots, p_n jsou sloupce ortonormální matice \mathbf{Q} , pro niž $\mathbf{Q}^T \cdot \mathbf{A} \cdot \mathbf{Q} = \text{diag}(\lambda_i)$.

Nechť $\{\mathbf{Q}_{k'}\}$ je podposloupnost posloupnosti $\{\mathbf{Q}_k\}$, $\mathbf{Q}_{k'} \rightarrow \mathbf{Q}_k$. Podle věty 4.49 existují $s \in P_n$ tak, že

$$\text{diag}(\lambda_{s(i)}) = \lim_{k' \rightarrow \infty} \mathbf{A}_{k'} = \lim_{k' \rightarrow \infty} (\mathbf{Q}_{k'}^T \cdot \mathbf{A}_{k'} \cdot \mathbf{Q}_{k'}) = \mathbf{Q}_{k'}^T \cdot \mathbf{A}_{k'} \cdot \mathbf{Q}_{k'}$$

což bylo dokázáno. Všechna vlastní čísla jsou různá, tedy existuje pouze konečně mnoho hromadných bodů.

Dále pro úhly určující \mathbf{S}_k máme

$$\tan 2\alpha_k = \frac{2a_{pq}^{(k)}}{a_{qq}^{(k)} - a_{pp}^{(k)}}, \quad |\alpha_k| \leq \pi/4.$$

Podle věty 4.49 odtud plyne, že existuje l tak, že pro $k \geq l$ je

$$|a_{qq}^{(k)} - a_{pp}^{(k)}| \geq \frac{1}{2} \min_{i \neq j} |\lambda_i - \lambda_j| > 0.$$

Protože se dvojice (p, q) mění s k , nemůžeme dokázat, že posloupnosti $a_{qq}^{(k)}$ a $a_{pp}^{(k)}$ konvergují.

$$\lim_{k \rightarrow \infty} a_{pq}^{(k)} = 0,$$

tedy

$$\lim_{k \rightarrow \infty} \alpha_k = 0 \quad a \quad \lim_{k \rightarrow \infty} \mathbf{S}_k = \mathbf{I}$$

$$\mathbf{Q}_{k+1} - \mathbf{Q}_k = \mathbf{Q}_k(\mathbf{S}_k - \mathbf{I}) \rightarrow 0.$$

A konečně posloupnost $\{\mathbf{Q}_k\}$ je ohraničená, protože $\|\mathbf{Q}_k\| = 1$. □

Při výpočtu můžeme průběžně kontrolovat výsledky tím, že po každém kroku zjišťujeme, zda

$$a_{pp}^{(k+0)} + a_{qq}^{(k+1)} = a_{pp}^{(k)} + a_{qq}^{(k)}.$$

Nebo vypočítáme matici $\mathbf{S} \cdot \mathbf{D} \cdot \mathbf{S}^T$, která by se měla rovnat matici \mathbf{A} .

Přesnost Jacobiho metody závisí na tom, jak přesně se vypočítají odmocniny pro určení $\sin \alpha_k$ a $\cos \alpha_k$.

Ačkoliv se Jacobiho metoda používá převážně pro symetrické matice, pracuje často dobře i v případě nesymetrických matic. V tomto případě ovšem konverguje k trojúhelníkové matici a má-li výchozí matice komplexní vlastní čísla, je nutné použít místo matic \mathbf{S}_k vhodné unitární matice.

4.5 Householderova matice zrcadlení

Definice 4.53. Matice tvaru

$$\mathbf{H}(\mathbf{u}) := \mathbf{I} - \frac{2\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T\mathbf{u}} = \mathbf{I} - \frac{2\mathbf{u}\mathbf{u}^T}{\|\mathbf{u}\|^2}$$

se nazývá *Householderova matice* (někdy též elementární zrcadlení nebo *Householderova transformace*).

Vlastnosti Householderovy matice:

- označení *matice zrcadlení* se používá proto, že aplikujeme-li matici $\mathbf{H}(\mathbf{u})$ pro nějaké \mathbf{u} na vektor $\mathbf{x} \in \mathbb{R}^n$, pak je vektor $\mathbf{H}(\mathbf{u})\mathbf{x}$ souměrný s vektorem \mathbf{x} podle nadroviny ortogonální k vektoru \mathbf{v} ;
- matice \mathbf{I} bývá považována za speciální případ Householderovy transformace – pro $\mathbf{u} = \mathbf{o}$ je $\mathbf{H}(\mathbf{o}) = \mathbf{I}$;
- $\|\mathbf{H}\mathbf{x}\|_2 = \|\mathbf{x}\|_2$ pro každé $\mathbf{x} \in \mathbb{R}^n$, tj. zrcadlení tedy nemění délku vektoru;
- $\mathbf{H}\mathbf{y} = \mathbf{y}$ pro každé $\mathbf{y} \in P = \{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{v}^T\mathbf{u} = 0\}$;
- \mathbf{H} má jednoduchou vlastní hodnotu -1 a $(n - 1)$ násobnou vlastní hodnotu 1 .

Důkaz. Neboť $\mathbf{y} \in P = \{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{v}^T\mathbf{u} = 0\}$ má $n - 1$ lineárně nezávislých vektorů $\mathbf{y}_1, \dots, \mathbf{y}_{n-1}$ a $\mathbf{H}\mathbf{y}_i = \mathbf{y}_i$ pro $i = 1, 2, \dots, n - 1$. Takže 1 je $(n - 1)$ -násobná vlastní hodnota. \mathbf{H} také zrcadlí \mathbf{u} na $-\mathbf{u}$, tj. $\mathbf{H}\mathbf{u} = -\mathbf{u}$. Takže -1 je vlastní hodnota matice \mathbf{H} , která musí být jednoduchá, neboť \mathbf{H} má pouze n vlastních hodnot. \square

- z věty o spektrálním rozkladu plyne

$$\det(\mathbf{H}) = (-1)1 \cdots 1 = -1;$$

- Matice \mathbf{H} je ortogonální a symetrická

Důkaz. Symetrie plyne z

$$\mathbf{H}^T(\mathbf{u}) = \mathbf{I}^T - 2\left(\frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T\mathbf{u}}\right)^T = \mathbf{I} - \frac{2\mathbf{u}\mathbf{u}^T}{\|\mathbf{u}\|^2} = \mathbf{H}(\mathbf{u}).$$

Protože platí

$$\mathbf{H}^2(\mathbf{u}) = \left(\mathbf{I} - \frac{2\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T\mathbf{u}}\right) \left(\mathbf{I} - \frac{2\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T\mathbf{u}}\right) = \mathbf{I}^2 - 4\frac{\mathbf{u}\mathbf{u}^T}{\|\mathbf{u}\|^2} + 4\frac{\mathbf{u}\mathbf{u}^T\mathbf{u}\mathbf{u}^T}{\|\mathbf{u}\|^4} = \mathbf{I},$$

je matice $\mathbf{H}(\mathbf{u})$ ortogonální. □

Věta 4.54. Pro každé dva vektory $\mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ takové, že $\mathbf{y} \neq \mathbf{z}$ a $\|\mathbf{y}\|_2 = \|\mathbf{z}\|_2$, platí

$$\mathbf{y} = \mathbf{H}(\mathbf{y} - \mathbf{z})\mathbf{z}.$$

Jinými slovy, každé dva různé vektory o stejné normě lze převést jeden na druhý Householderovou transformací.

Důkaz. Platí

$$\begin{aligned} \mathbf{H}(\mathbf{y} - \mathbf{z})\mathbf{z} &= \left(\mathbf{I} - \frac{2(\mathbf{y} - \mathbf{z})(\mathbf{y} - \mathbf{z})^T}{\|\mathbf{y} - \mathbf{z}\|_2^2}\right)\mathbf{z} = \mathbf{z} - 2\frac{\mathbf{y}^T\mathbf{z} - \|\mathbf{z}\|_2^2}{\|\mathbf{y} - \mathbf{z}\|_2^2}(\mathbf{y} - \mathbf{z}) = \\ &= \mathbf{z} + \frac{\|\mathbf{y}\|_2^2 + \|\mathbf{z}\|_2^2 - 2\mathbf{y}^T\mathbf{z}}{\|\mathbf{y} - \mathbf{z}\|_2^2}(\mathbf{y} - \mathbf{z}) = \mathbf{z} + \frac{\|\mathbf{y} - \mathbf{z}\|_2^2}{\|\mathbf{y} - \mathbf{z}\|_2^2}(\mathbf{y} - \mathbf{z}) = \mathbf{y}. \end{aligned}$$

□

Důsledek 4.55. Jsou-li \mathbf{y}, \mathbf{z} dva vektory o stejné normě, potom existuje ortogonální matice \mathbf{Q} taková, že $\mathbf{y} = \mathbf{Q}\mathbf{z}$.

Důkaz. Pro $\mathbf{y} \neq \mathbf{z}$ stačí vzít $\mathbf{Q} = \mathbf{H}(\mathbf{y} - \mathbf{z})$, jinak $\mathbf{Q} = \mathbf{I}$. □

Věta 4.56. Pro každé $x \in \mathbb{R}^n$ je

$$\mathbf{H} = \begin{cases} \mathbf{H}(x + \operatorname{sgn}(x_1)\|x\|_2\mathbf{e}_1) & \text{pro } x_1 \neq \|x\|_2, \\ \mathbf{I} & \text{pro } x_1 = \|x\|_2 \end{cases}$$

ortogonální matice s vlastností

$$\mathbf{H}\mathbf{x} = \|x\|_2\mathbf{e}_1.$$

Jinými slovy, aplikujeme-li vhodnou matici \mathbf{H} na vektor \mathbf{x} , dostaneme vektor, který má všechny složky až na první nulové.

Důkaz. Je-li $x_1 = \|x\|_2$, potom z $x_1^2 = x_1^2 + \dots + x_n^2$ plyne, že $x_2 = \dots = x_n = 0$. Tedy $\mathbf{x} = x_1\mathbf{e}_1 = \|x\|_2\mathbf{e}_1 = \mathbf{I}\mathbf{x} = \mathbf{H}\mathbf{x}$.

Je-li $x_1 \neq \|x\|_2$, potom $\mathbf{x} + \operatorname{sgn}(x_1)\|x\|_2\mathbf{e}_1 \neq 0$, takže vektory $\mathbf{y} = \operatorname{sgn}(x_1)\|x\|_2\mathbf{e}_1$ a $\mathbf{z} = \mathbf{x}$ jsou různé a platí pro ně $\|\mathbf{y}\|_2 = \|x\|_2 = \|\mathbf{z}\|_2$, a odtud je

$$\mathbf{y} = \operatorname{sgn}(x_1)\|x\|_2\mathbf{e}_1 = \mathbf{H}(\mathbf{y} - \mathbf{z})\mathbf{z} = \mathbf{H}(-\mathbf{x} - \operatorname{sgn}(x_1)\|x\|_2\mathbf{e}_1)\mathbf{x}.$$

□

Poznámka 4.57. Pro vektor určující Householderovu matici lze volit buď $+\|\mathbf{x}\|_2\mathbf{e}_1$ nebo $-\|\mathbf{x}\|_2\mathbf{e}_1$. Z důvodu minimalizace numerických chyb volíme stejné znaménko jako u první složky vektoru \mathbf{x} .

Věta 4.58. Pro každé \mathbf{x} takové, že $\|\mathbf{x}\|_2 = 1$, je

$$\mathbf{H} = \begin{cases} \mathbf{H}(\mathbf{x} + \operatorname{sgn}(x_1)\mathbf{w}_1) & \text{pro } \mathbf{x} \neq \mathbf{e}_1, \\ \mathbf{I} & \text{pro } \mathbf{x} = \mathbf{e}_1 \end{cases}$$

ortogonální matice, jejímž prvním sloupcem je vektor \mathbf{x} .

Důkaz. Pro $\mathbf{x} = \mathbf{e}_1$ je zřejmý.

Nechť tedy $\mathbf{x} \neq \mathbf{e}_1$. Protože $\|\mathbf{x}\|_2 = 1 = \|\mathbf{e}_1\|_2$, je podle Věty 4.58

$$\mathbf{x} = \mathbf{H}(\mathbf{x} + \operatorname{sgn}(x_1)\mathbf{e}_1) = \mathbf{H}\mathbf{e}_1 = \mathbf{H}\cdot\mathbf{1},$$

což je tvrzením věty. □

Díky těmto větám tedy umíme najít vektor \mathbf{u} tak, že daný nenulový vektor \mathbf{x} se transformuje na vektor, který má nenulovou pouze první složku.

Příklad 4.59. Lze

$$\mathbf{x} = (-1, -2, 7)^T \xrightarrow{\mathbf{H}(\mathbf{u})} (\alpha, 0, 0)^T?$$

Řešení. Protože $\|\mathbf{x}\|_2 = 3\sqrt{6}$, položíme $\mathbf{u} = \mathbf{x} - \|\mathbf{x}\|_2\mathbf{e}_1 = (-1 - 3\sqrt{6}, -2, 7)^T$ a $\|\mathbf{u}\|_2 = 6(18 + \sqrt{6})$. Dále

$$\mathbf{u}\mathbf{u}^T = \begin{pmatrix} -1 - 3\sqrt{6} \\ -2 \\ 7 \end{pmatrix} (-1 - 3\sqrt{6}, -2, 7) = \begin{pmatrix} 55 + 6\sqrt{6} & 2 + 6\sqrt{6} & -7 - 21\sqrt{6} \\ 2 + 6\sqrt{6} & 4 & -14 \\ -7 - 21\sqrt{6} & -14 & 49 \end{pmatrix},$$

takže

$$\mathbf{H}(\mathbf{u}) = \frac{1}{3(18 + \sqrt{6})} \begin{pmatrix} -1 - 3\sqrt{6} & -2 - 6\sqrt{6} & 7 + 21\sqrt{6} \\ -2 - 6\sqrt{6} & 50 + 3\sqrt{6} & 14 \\ 7 + 21\sqrt{6} & 14 & 5 + 3\sqrt{6} \end{pmatrix}.$$

Snadno lze ověřit, že

$$\mathbf{H}(\mathbf{u})\mathbf{x} = (3\sqrt{6}, 0, 0)^T.$$

Odpověď na položenou otázku je proto kladná. □

4.5.1 Givensova-Householderova metoda

Jedná se o metodu speciálně vhodnou k hledání některých vlastních čísel symetrických matic, např. všech vlastních čísel obsažených v předem zadaném intervalu. Metoda umožňuje počítat vlastní čísla s různou přesností. Na druhé straně nám neposkytuje informace o vlastních vektorech. Tato metoda má dvě etapy:

tj. po vhodné volbě $\widetilde{\mathbf{v}}_k$ máme další část třídiagonální matice.

Matici \mathbf{H}_k můžeme popsat také jako Householderovu matici příslušnou vektoru $\mathbf{v}_k = [0, \dots, 0, \widetilde{\mathbf{v}}_k]^T$.

Máme dvě možné volby vektoru \mathbf{v}_k :

$$\mathbf{v}_k = \left[0, \dots, 0, a_{k+1,k}^{(k)} \pm \left(\sum_{i=k+1}^n |a_{ik}^{(k)}|^2 \right)^{1/2}, a_{k+2,k}^{(k)}, \dots, a_{n,k}^{(k)} \right]^T,$$

znaménko se volí stejné jako je znaménko u $a_{k+1,k}^{(k)}$. Máme-li určen vektor \mathbf{v}_k , potom prvky $a_{ij}^{(k+1)}$, $k+1 \leq i, j \leq n$ matice $\mathbf{A}_{k+1} = \left(a_{ij}^{(k+1)} \right)$ určíme následovně:

Postupně určíme vektory

$$\mathbf{w}_k = (\mathbf{v}_k^T \mathbf{v}_k)^{-1/2} \mathbf{v}_k,$$

$$\mathbf{q}_k = 2(\mathbf{I} - \mathbf{w}_k \mathbf{w}_k^T) \mathbf{A}_k \mathbf{w}_k,$$

jejichž složky označíme $w_i^{(k)}$, $q_i^{(k)}$. Potom matice \mathbf{A}_{k+1} má tvar

$$\mathbf{A}_{k+1} = \mathbf{A}_k - \mathbf{w}_k \mathbf{q}_k^T - \mathbf{q}_k \mathbf{w}_k^T$$

tj.

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - w_i^{(k)} q_j^{(k)} - q_i^{(k)} w_j^{(k)}$$

$k+1 \leq i, j \leq n$.

Příklad 4.60. Householderovou transformací převed'te matici

$$\mathbf{A} = \begin{pmatrix} 4 & 2 & 2 & 1 \\ 2 & -3 & 1 & 1 \\ 2 & 1 & 3 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}$$

na třídiagonální tvar.

Řešení.

$$\mathbf{v}_0 = (0 \quad 2 + (\sqrt{2^2 + 2^2 + 1^2}) \quad 2 \quad 1)^T,$$

$$\mathbf{w}_0 = (\mathbf{v}_0^T \mathbf{v}_0)^{-1/2} \mathbf{v}_0 = (0 \quad 0.912871 \quad 0.365148 \quad 0.182574)^T,$$

$$\mathbf{q}_0 = 2(\mathbf{I} - \mathbf{w}_0 \mathbf{w}_0^T) \mathbf{A} \mathbf{w}_0 = (5.477224 \quad -2.7386095 \quad 5.03904626 \quad 3.61496813)^T,$$

$$\mathbf{A}_1 = \mathbf{A} - \mathbf{w}_0 \mathbf{q}_0^T - \mathbf{q}_0 \mathbf{w}_0^T = \begin{pmatrix} 4 & -3 & 0 & 0 \\ -3 & 2 & -2.6 & -1.8 \\ 0 & -2.6 & -0.68 & -1.24 \\ 0 & -1.8 & -1.24 & 0.68 \end{pmatrix},$$

$$\begin{aligned} \mathbf{v}_1 &= (0 \quad 0 \quad -2.6 - \sqrt{-2.6^2 + (-1.8)^2} \quad -1.8), \\ \mathbf{w}_1 &= (0 \quad 0 \quad -0.954514 \quad -0.298168)^T, \\ \mathbf{q}_1 &= (0 \quad 6.0365793 \quad -0.3770516 \quad 1.207794)^T, \\ \mathbf{A}_2 &= \mathbf{A}_1 - \mathbf{w}_1 \mathbf{q}_1^T - \mathbf{q}_1 \mathbf{w}_1^T = \begin{pmatrix} 4 & -3 & 0 & 0 \\ -3 & 2 & 3.162278 & 0 \\ 0 & 3.162278 & -1.4 & -0.2 \\ 0 & 0 & -0.2 & 1.4 \end{pmatrix}. \end{aligned}$$

□

Givensova metoda

Metoda slouží k určení vlastních čísel symetrické třídiagonální matice

$$\mathbf{B} = \begin{pmatrix} b_1 & c_1 & & & \\ c_1 & b_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & c_{n-1} & b_{n-1} & c_{n-1} \\ & & & c_{n-2} & b_n \end{pmatrix}.$$

Pokud je některé z c_i nula, rozpadá se matice \mathbf{B} na dvě třídiagonální matice stejného typu. Tedy bez újmy na obecnosti můžeme předpokládat, že $c_i \neq 0$, ($i = 1, \dots, n-1$). Označme

$$\mathbf{B}_i = \begin{pmatrix} b_1 & c_1 & & & \\ c_1 & b_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & c_{i-1} & b_{i-1} & c_{i-1} \\ & & & c_{i-2} & b_i \end{pmatrix},$$

$i = 1, \dots, n$.

Věta 4.61. Polynomy $p_i(\lambda)$, $\lambda \in \mathbb{R}$, definované pro $i = 1, \dots, n$ rekurentně

$$p_0(\lambda) = 1,$$

$$p_1(\lambda) = b_1 - \lambda,$$

$$p_i(\lambda) = (b_i - \lambda)p_{i-1}(\lambda) - c_{i-1}^2 p_{i-2}(\lambda), \quad 2 \leq i \leq n,$$

mají následující vlastnosti:

1. Polynom p_i je charakteristický polynom matice \mathbf{B}_i .

$$p_i(\lambda) = \det(\mathbf{B}_i - \lambda \mathbf{I}).$$

2.

$$\lim_{\lambda \rightarrow \infty} p_i(\lambda) = +\infty, \quad i = 1, \dots, n.$$

3. Jestliže $p_i(\lambda_0) = 0$, potom $p_{i-1}(\lambda_0) \cdot p_{i+1}(\lambda_0) < 0$, $i = 1, \dots, n-1$.4. Polynom p_i má i vzájemně různých kořenů, které oddělují $i+1$ kořenů polynomu p_{i+1} , $i = 1, \dots, n$.**Důkaz.** ad 1 Plyne z rozvoje $\det(\mathbf{B}_i - \lambda \mathbf{I})$.ad 2 $p_i(\lambda) = (-1)^i \lambda^i - \dots \rightarrow \infty$ pro $\lambda \rightarrow \infty$.ad 3 Necht' $p_i(\lambda_0) = 0$ pro nějaké i , $i = 1, \dots, n-1$. Z definice p_i plyne

$$p_{i+1}(\lambda_0) = -c_i^2 \cdot p_{i-1}(\lambda_0).$$

Protože $c_i \neq 0$, dostaneme buď

$$p_{i-1}(\lambda_0) \cdot p_{i+1}(\lambda_0) < 0$$

nebo

$$p_{i-1}(\lambda_0) = p_{i-1}(\lambda_0) = p_{i+1}(\lambda_0),$$

což by indukcí vedlo k tomu, že

$$p_i(\lambda_0) = p_{i-1}(\lambda_0) = \dots = p_1(\lambda_0) = p_0(\lambda_0),$$

což je spor, protože $p_0(\lambda_0) = 1$.ad 4 Plyne z 2 a 3. □**Poznámka 4.62.** Posloupnost polynomů splňující 2 – 4 se nazývá **Sturmova posloupnost**, která se používá při výpočtu kořenu polynomů. Už jsme o ní mluvili v předchozí kapitole, viz definici 3.28.**Příklad 4.63.** Pomocí charakteristického polynomu určete vlastní čísla třídiagonální matice \mathbf{A}_2 z příkladu 4.60.**Řešení.**

$$\mathbf{A}_2 = \begin{pmatrix} 4 & -3 & 0 & 0 \\ -3 & 2 & 3.162278 & 0 \\ 0 & 3.162278 & -1.4 & -0.2 \\ 0 & 0 & -0.2 & 1.4 \end{pmatrix}.$$

$$p_0(\lambda) = 1,$$

$$p_1(\lambda) = 4 - \lambda,$$

$$p_2(\lambda) = (-2 - \lambda)(4 - \lambda) - 9,$$

$$p_3(\lambda) = (-1.4 - \lambda)[(-2 - \lambda)(4 - \lambda) - 9] - 10(4 - \lambda),$$

$$p_4(\lambda) = (1.4 - \lambda)[(-1.4 - \lambda)[(-2 - \lambda)(4 - \lambda) - 9] - 10(4 - \lambda)] - 0.04[(-2 - \lambda)(4 - \lambda) - 9] = \lambda^4 - 2\lambda^3 - 29\lambda^2 + 58\lambda - 22.$$

Kořeny polynomu $p_4(\lambda)$ jsou

$$\lambda_1 = -5.4355 \quad \lambda_2 = 5.4907 \quad \lambda_3 = 1.4289 \quad \lambda_4 = 0.5159$$

□

Věta 4.64. *Bud' i přirozené číslo, $1 \leq i \leq n$. Pro dané $\mu \in \mathbb{R}$, položme*

$$\operatorname{sgn} p_i(\mu) = \begin{cases} \operatorname{sgn} p_i(\mu) & \text{je-li } p_i(\mu) \neq 0, \\ \operatorname{sgn} p_{i-1}(\mu) & \text{je-li } p_i(\mu) = 0. \end{cases}$$

Potom $N(i, \mu)$, což je počet znaménkových změn v posloupnosti po sobě jdoucích prvků uspořádané množiny $N(i, \mu) = \{+, \operatorname{sgn} p_1(\mu), \dots, \operatorname{sgn} p_i(\mu)\}$, se rovná počtu kořenů polynomu p_i , které jsou menší než μ .

Tato věta umožňuje aproximaci (s libovolnou přesností) vlastních čísel matice $\mathbf{B} = \mathbf{B}_n$ a dokonce přímý výpočet vlastního čísla na dané pozici. Předpokládejme např., že chceme aproximaci i -tého vlastního čísla $\lambda_i^{(n)} = \lambda_i$ matice \mathbf{B} (jako předtím předpokládáme, že $\lambda_1, \dots, \lambda_n$ jsou vzájemně různá a uspořádaná sestupně).

Krok 1: Určíme interval $\langle a_0, b_0 \rangle$, v němž leží žádané vlastní číslo, např. $-a_0 = b_0 = \|\mathbf{B}\|_\infty$.

Krok 2: Vypočteme $c_0 = \frac{a_0 + b_0}{2}$, spočteme $N(n, c_0)$. Potom buď

$$N(n, c_0) \geq i \quad a \quad \lambda_i \in \langle a_0, c_0 \rangle$$

nebo

$$N(n, c_0) < i \quad a \quad \lambda_i \in \langle c_0, b_0 \rangle.$$

Tím získáme interval $\langle a_1, b_1 \rangle$, v němž leží kořen λ_i .

Postupně získáme posloupnost intervalů $\langle a_k, b_k \rangle$, $k \geq 0$ takových, že $\lambda_i \in \langle a_k, b_k \rangle$ a

$$b_k - a_k = 2^{-k}(b_0 - a_0), \quad k \geq 0.$$

4.5.2 QR-rozklad

Definice 4.65. Dvojici matic \mathbf{Q} a \mathbf{R} nazveme **QR-rozkladem** matice \mathbf{A} , pokud platí, že

$$\mathbf{A} = \mathbf{Q}\mathbf{R},$$

přičemž \mathbf{Q} je ortogonální matice a \mathbf{R} je horní trojúhelníková matice.

Nyní uvedeme věty o existenci QR-rozkladu a jeho jednoznačnosti.

Věta 4.66. *K libovolné reálné matici $\mathbf{A} \in \mathbb{R}^{m \times n}$, kde $m \geq n$, existuje ortogonální matice $\mathbf{Q} \in \mathbb{R}^{m \times m}$ a horní trojúhelníková matice $\mathbf{R} \in \mathbb{R}^{m \times n}$ tak, že platí $\mathbf{A} = \mathbf{Q}\mathbf{R}$.*

Věta 4.67. *Jsou-li sloupce matice $\mathbf{A} \in \mathbb{R}^{m \times n}$, $m \geq n$, lineárně nezávislé, potom v QR-rozkladu jsou matice \mathbf{R} a prvních n sloupců matice \mathbf{Q} určeny až na znaménko jednoznačně.*

Důkazy obou vět viz [49].

Konstrukce QR-rozkladu

QR-rozklad pomocí Gram-Schmidtova algoritmu

Věta 4.68 (Gram-Schmidtův QR-rozklad). *K libovolné reálné matici $\mathbf{A} \in \mathbb{R}^{m \times n}$, kde $m \geq n$, existuje ortogonální matice $\mathbf{Q} \in \mathbb{R}^{m \times m}$ a horní trojúhelníková matice $\mathbf{R} \in \mathbb{R}^{m \times n}$ s nezápornými prvky na diagonále tak, že platí $\mathbf{A} = \mathbf{QR}$. V případě lineárně nezávislých sloupců matice \mathbf{A} jsou prvky na diagonále kladné.*

Základní myšlenka důkazu: Máme-li matici $\mathbf{A} \in \mathbb{R}^{m \times n}$, pak aplikací zobecněného Gram-Schmidtova ortogonalizačního procesu na sloupce matice \mathbf{A} (ty mohou být lineárně závislé i nezávislé) a doplněním těchto vektorů na bázi v \mathbb{R}^m získáme sloupce matice \mathbf{Q} . Uvažujme matici $\mathbf{A} = (\mathbf{a}_1 | \dots | \mathbf{a}_n)$ složenou ze sloupcových vektorů. Pak

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{a}_1, & \mathbf{e}_1 &= \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \\ \mathbf{u}_2 &= \mathbf{a}_2 - p_{\mathbf{e}_1} \mathbf{a}_2, & \mathbf{e}_2 &= \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}, \\ \mathbf{u}_3 &= \mathbf{a}_3 - p_{\mathbf{e}_1} \mathbf{a}_3 - p_{\mathbf{e}_2} \mathbf{a}_3, & \mathbf{e}_3 &= \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|}, \\ & \vdots & & \\ \mathbf{u}_k &= \mathbf{a}_k - \sum_{j=1}^{k-1} p_{\mathbf{e}_j} \mathbf{a}_k, & \mathbf{e}_k &= \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}, \end{aligned}$$

kde $p_{\mathbf{u}} \mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{u}$. Po úpravě obdržíme vzorce pro vektory \mathbf{a}_i

$$\begin{aligned} \mathbf{a}_1 &= \mathbf{e}_1 \|\mathbf{u}_1\|, \\ \mathbf{a}_2 &= p_{\mathbf{e}_1} \mathbf{a}_2 + \mathbf{e}_2 \|\mathbf{u}_2\|, \\ \mathbf{a}_3 &= p_{\mathbf{e}_1} \mathbf{a}_3 + p_{\mathbf{e}_2} \mathbf{a}_3 + \mathbf{e}_3 \|\mathbf{u}_3\|, \\ & \vdots \\ \mathbf{a}_k &= \sum_{j=1}^{k-1} p_{\mathbf{e}_j} \mathbf{a}_k + \mathbf{e}_k \|\mathbf{u}_k\|. \end{aligned}$$

Označme $\mathbf{Q} = (\mathbf{e}_1 | \dots | \mathbf{e}_n)$. Nyní máme

$$\mathbf{R} = \mathbf{Q}^T \mathbf{A} = \begin{pmatrix} \langle \mathbf{e}_1, \mathbf{a}_1 \rangle & \langle \mathbf{e}_1, \mathbf{a}_2 \rangle & \langle \mathbf{e}_1, \mathbf{a}_3 \rangle & \cdots & \langle \mathbf{e}_1, \mathbf{a}_n \rangle \\ 0 & \langle \mathbf{e}_2, \mathbf{a}_2 \rangle & \langle \mathbf{e}_2, \mathbf{a}_3 \rangle & \cdots & \langle \mathbf{e}_2, \mathbf{a}_n \rangle \\ 0 & 0 & \langle \mathbf{e}_3, \mathbf{a}_3 \rangle & \cdots & \langle \mathbf{e}_3, \mathbf{a}_n \rangle \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \langle \mathbf{e}_n, \mathbf{a}_n \rangle \end{pmatrix},$$

neboť $\mathbf{Q}\mathbf{Q}^T = \mathbf{E}$ a $\langle \mathbf{e}_j, \mathbf{a}_j \rangle = \|\mathbf{u}_j\|$, $\langle \mathbf{e}_j, \mathbf{a}_k \rangle = 0$ pro $j > k$.

Příklad 4.69. Provedme QR -rozklad matice $\mathbf{A} = \begin{pmatrix} 12 & -51 & 4 \\ 6 & 167 & -68 \\ -4 & 24 & -41 \end{pmatrix}$.

Řešení. Gram-Schmidtovým algoritmem dostaneme

$$\mathbf{U} = (\mathbf{u}_1 \mid \mathbf{u}_2 \mid \mathbf{u}_3) = \begin{pmatrix} 12 & -69 & -58 \\ 6 & 158 & 6 \\ -4 & 30 & -165 \end{pmatrix}.$$

Matici \mathbf{Q} potom získáme jako

$$\mathbf{Q} = \left(\frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} \mid \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} \mid \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} \right) = \begin{pmatrix} 6/7 & -69/175 & -58/175 \\ 3/7 & 158/175 & 6/175 \\ -2/7 & 6/35 & -33/35 \end{pmatrix}.$$

$\mathbf{A} = \mathbf{Q}\mathbf{Q}^T\mathbf{A} = \mathbf{Q}\mathbf{R}$, takže

$$\mathbf{R} = \mathbf{Q}^T\mathbf{A} = \begin{pmatrix} 14 & 21 & -14 \\ 0 & 175 & -70 \\ 0 & 0 & 35 \end{pmatrix}.$$

□

Algoritmus 3. Mějme matici \mathbf{A} . Položme

$$r_{11} = \|\mathbf{a}_1\|, \quad \mathbf{q}_1 = \frac{\mathbf{a}_1}{r_{11}},$$

pro $k = 2, \dots, n$ spočítejme:

$$r_{jk} = \langle \mathbf{q}_j, \mathbf{a}_k \rangle \quad \text{pro } j = 1, \dots, k-1,$$

$$z_k = \mathbf{a}_k - \sum_{j=1}^{k-1} r_{jk}\mathbf{q}_j,$$

$$r_{kn}^2 = \langle z_k, z_n \rangle$$

$$\mathbf{q}_k = \frac{z_k}{r_{kk}}.$$

Metodu lze také upravit tak, že zaměníme pořadí operací. Tedy položme

$$\mathbf{A}^0 \equiv \mathbf{A}.$$

Pak pro $k = 2, \dots, n$ spočtěme

$$r_{kk} = \|\mathbf{a}_k^{(k-1)}\|_2, \quad \mathbf{q}_k = \frac{\mathbf{a}_k^{(k-1)}}{r_{kk}},$$

$$r_{ki} = \mathbf{q}_k^T \mathbf{a}_i^{(k-1)} \quad \text{pro } i = k+1, \dots, n,$$

$$\mathbf{A}^{(k)} = \mathbf{A}^{(k-1)} - \mathbf{q}_k \mathbf{r}_k^T.$$

Z formálního hlediska jde o změnu pořadí operací, ovšem z numerického hlediska obdržíme různé výsledky.

QR-rozklad pomocí Householderovy matice

Věta 4.70 (Householderův QR-rozklad). Každou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ lze pomocí $s = \min\{n, m - 1\}$ Householderových matic rozložit na součin \mathbf{QR} , a to tak, že platí

$$\mathbf{H}_s \cdots \mathbf{H}_2 \mathbf{H}_1 \mathbf{A} = \mathbf{Q}^T \mathbf{A} = \begin{cases} \begin{pmatrix} \mathbf{R}_1 \\ 0 \end{pmatrix} & m > n, \\ \begin{pmatrix} \mathbf{R}_1, 0 \end{pmatrix} & m < n, \\ \mathbf{R} & m = n. \end{cases}$$

Důkaz. Konstrukce QR-rozkladu

Mějme reálnou matici \mathbf{A}

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

Krok 1.: Zkonstruujeme Householderovu matici \mathbf{H}_1 tak, aby $\mathbf{H}_1 \mathbf{A}$ měla v prvním sloupci pouze nuly s výjimkou pozice (1, 1), tj. aby

$$\mathbf{H}_1 \mathbf{A} = \begin{pmatrix} \boxplus & \cdots \\ 0 & \cdots \\ \vdots & \vdots \\ 0 & \cdots \end{pmatrix}.$$

K tomu stačí získat vektor \mathbf{u}_n (dle předchozího) tak, že pro

$$\mathbf{H}_1 = \mathbf{E} - 2 \frac{\mathbf{u}_n \mathbf{u}_n^T}{\mathbf{u}_n^T \mathbf{u}_n}$$

platí

$$\mathbf{H}_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} = \begin{pmatrix} \boxplus \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Označme $\mathbf{A}^{(1)} := \mathbf{H}_1 \mathbf{A}$. $\mathbf{A}^{(1)}$ je tvaru

$$\mathbf{A}^{(1)} = \begin{pmatrix} a_{11} & \cdots \\ 0 & \cdots \\ \vdots & \vdots \\ 0 & \cdots \end{pmatrix}.$$

Krok 2.: Zkonstruuje Householderovu matici \mathbf{H}_2 tak, že $\mathbf{H}_2\mathbf{A}^{(1)}$ má ve druhém sloupci 0 pod pozicí (2, 2) při zachování požadavku prvního kroku, tj.

$$\mathbf{A}^{(2)} := \mathbf{H}_2\mathbf{A}^{(1)} = \begin{pmatrix} \boxplus & \square & \cdots \\ 0 & \boxplus & \cdots \\ 0 & 0 & \\ \vdots & \vdots & \vdots \\ 0 & 0 & \cdots \end{pmatrix}.$$

Matici \mathbf{H}_2 získáme tak, že nejdříve zkonstruuje Householderovu matici o rozměru $(m-1) \times (n-1)$

$$\widehat{\mathbf{H}}_2 := \mathbf{E}_{n-1} - 2 \frac{\mathbf{u}_{n-1}\mathbf{u}_{n-1}^T}{\mathbf{u}_{n-1}^T\mathbf{u}_{n-1}}$$

takovou, že

$$\widehat{\mathbf{H}}_2 \begin{pmatrix} a_{22} \\ a_{32} \\ \vdots \\ a_{m2} \end{pmatrix} = \begin{pmatrix} \boxplus \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

a definujeme

$$\mathbf{H}_2 := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & \widehat{\mathbf{H}}_2 & & \\ 0 & & & \end{pmatrix}.$$

Tím získáme matici $\mathbf{A}^{(2)} = \mathbf{H}_2\mathbf{A}^{(1)}$. Analogicky pokračujeme dále.

Pro $k \leq s$.

Krok k-tý: Obecně vytváříme Householderovu matici

$$\widehat{\mathbf{H}}_k := \mathbf{E}_{n-k+1} - 2 \frac{\mathbf{u}_{n-k+1}\mathbf{u}_{n-k+1}^T}{\mathbf{u}_{n-k+1}^T\mathbf{u}_{n-k+1}}$$

o rozměru $(m-k+1) \times (n-k+1)$ takovou, že

$$\widehat{\mathbf{H}}_k \begin{pmatrix} a_{kk} \\ \vdots \\ a_{mk} \end{pmatrix} = \begin{pmatrix} \boxplus \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Definujeme

$$\mathbf{H}_k := \begin{pmatrix} \mathbf{E}_{k-1} & 0 \\ 0 & \widehat{\mathbf{H}}_k \end{pmatrix},$$

čili můžeme spočítat $\mathbf{A}^{(k)} = \mathbf{H}_k\mathbf{A}^{(k-1)}$.

Tímto způsobem po s krocích obdržíme matici $\mathbf{A}^{(s)}$, která bude v horním trojúhelníkovém

tvaru a bude právě maticí \mathbf{R} .

Protože

$$\mathbf{A}^{(k)} = \mathbf{H}_k \mathbf{A}^{(k-1)} \quad k = 2, \dots, s,$$

máme

$$\mathbf{R} = \mathbf{A}^{(s)} = \mathbf{H}_s \mathbf{A}^{(s-1)} = \mathbf{H}_s \mathbf{H}_{s-1} \mathbf{A}^{(s-2)} = \dots = \mathbf{H}_s \mathbf{H}_{s-1} \dots \mathbf{H}_2 \mathbf{H}_1 \mathbf{A}.$$

Položme

$$\mathbf{Q}^T = \mathbf{H}_s \mathbf{H}_{s-1} \dots \mathbf{H}_2 \mathbf{H}_1.$$

Máme hledanou ortogonální matici (neboť každá z \mathbf{H}_i je ortogonální). Celkem

$$\mathbf{R} = \mathbf{Q}^T \mathbf{A},$$

tj.

$$\mathbf{A} = \mathbf{Q} \mathbf{R}.$$

(Zopakujme si, že $\mathbf{Q} = \mathbf{H}_1^T \mathbf{H}_2^T \dots \mathbf{H}_s^T = \mathbf{H}_1 \mathbf{H}_2 \dots \mathbf{H}_s$.)

□

Příklad 4.71. Proved'te QR-rozklad matice

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}.$$

Řešení. Krok 1.: Konstrukce \mathbf{H}_1 .

$$\mathbf{H}_1 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \boxplus \\ 0 \\ 0 \end{pmatrix}.$$

Potom tedy dle Příkladu 4.59 spočteme

$$\mathbf{u}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \sqrt{2} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \sqrt{2} \\ 1 \\ 1 \end{pmatrix},$$

takže

$$\mathbf{H}_1 = \mathbf{I}_3 - 2 \frac{\mathbf{u}_3 \mathbf{u}_3^T}{\mathbf{u}_3^T \mathbf{u}_3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 0 & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{\sqrt{2}} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Určeme

$$\mathbf{A}^{(1)} = \mathbf{H}_1 \mathbf{A} = \begin{pmatrix} -\sqrt{2} & -\frac{3\sqrt{2}}{2} & 2\sqrt{2} \\ 0 & -\frac{1-\sqrt{2}}{2} & -\frac{2-\sqrt{2}}{2} \\ 0 & -\frac{1+\sqrt{2}}{2} & -\frac{2+\sqrt{2}}{2} \end{pmatrix}.$$

Krok 2.: Zkonstruuje

$$\widehat{\mathbf{H}}_2 = \begin{pmatrix} -0,2071 \\ -1,2071 \end{pmatrix} = \begin{pmatrix} \boxplus \\ 0 \end{pmatrix},$$

$$\mathbf{u}_2 = \begin{pmatrix} -0,2071 \\ -1,2071 \end{pmatrix} - 1,2247 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1,4318 \\ -1,2071 \end{pmatrix},$$

$$\widehat{\mathbf{H}}_2 = \begin{pmatrix} -0,1691 & -0,9856 \\ -0,9856 & 0,1691 \end{pmatrix},$$

tzn.

$$\mathbf{H}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -0,1691 & -0,9856 \\ 0 & -0,9856 & 0,1691 \end{pmatrix},$$

a spočítáme

$$\mathbf{A}^{(2)} = \mathbf{H}_2 \mathbf{A}^{(1)} = \mathbf{H}_2 \mathbf{H}_1 \mathbf{A} = \begin{pmatrix} -1,4142 & -2,1213 & -2,8284 \\ 0 & 1,2247 & 1,6330 \\ 0 & 0 & -0,5774 \end{pmatrix} = \mathbf{R}$$

Pro \mathbf{Q} nyní platí

$$\mathbf{Q} = \mathbf{H}_2 \mathbf{H}_1 = \begin{pmatrix} 0 & 0,8165 & 0,5774 \\ -0,7071 & 0,4082 & -0,5774 \\ -0,7071 & -0,4082 & 0,5774 \end{pmatrix}.$$

Celkem tedy

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0,8165 & 0,5774 \\ -0,7071 & 0,4082 & -0,5774 \\ -0,7071 & -0,4082 & 0,5774 \end{pmatrix} \begin{pmatrix} -1,4142 & -2,1213 & -2,8284 \\ 0 & 1,2247 & 1,6330 \\ 0 & 0 & -0,5774 \end{pmatrix} = \mathbf{QR}.$$

□

QR-rozklad pomocí Givensovy matice

Definice 4.72. Matice tvaru

$$\mathbf{G}(i, j, c, s) := \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & c & \cdots & s & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & -s & \cdots & c & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix} = \mathbf{I} + (c-1)(\mathbf{e}_i \mathbf{e}_i^T + \mathbf{e}_j \mathbf{e}_j^T) + s(\mathbf{e}_i \mathbf{e}_i^T - \mathbf{e}_j \mathbf{e}_j^T),$$

kde $c^2 + s^2 = 1$, se nazývá *Givensova matice*.

Můžeme volit $c = \cos \alpha$ a $s = \sin \alpha$ pro nějaké α . Pak značíme Givensovu matici jako $\mathbf{G}(i, j, \alpha)$.

Opět chceme sestavit matice $\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_s$ tentokrát však pomocí Givensových matic tak, aby $\mathbf{A}^{(1)} = \mathbf{Q}_1 \mathbf{A}$ měla nuly pod prvkem (1, 1) v prvním sloupci, matice $\mathbf{A}^{(2)} = \mathbf{Q}_2 \mathbf{A}^{(1)}$ měla nuly pod (2, 2) ve druhém sloupci, atd. Každou z matic \mathbf{Q}_i lze sestavit jako součin Givensových matic – ten je přitom možné sestavit takto:

$$\begin{aligned}\mathbf{Q}_1 &:= \mathbf{G}(1, m, \alpha) \mathbf{G}(1, m-1, \alpha) \cdots \mathbf{G}(1, 3, \alpha) \mathbf{G}(1, 2, \alpha) \\ \mathbf{Q}_2 &:= \mathbf{G}(2, m, \alpha) \mathbf{G}(2, m-1, \alpha) \cdots \mathbf{G}(2, 3, \alpha) \\ &\vdots\end{aligned}$$

Bud' $s = \min\{m-1, n\}$. Pak

$$\mathbf{R} = \mathbf{A}^{(s)} = \mathbf{Q}_s \mathbf{A}^{(s-1)} = \cdots = \mathbf{Q}_s \mathbf{Q}_{s-1} \cdots \mathbf{Q}_2 \mathbf{Q}_1 \mathbf{A} = \mathbf{Q}^T \mathbf{A}.$$

Nyní máme $\mathbf{A} = \mathbf{Q}\mathbf{R}$, kde $\mathbf{Q}^T = \mathbf{Q}_s \cdots \mathbf{Q}_2 \mathbf{Q}_1$. Tento fakt lze zformulovat do následující věty.

Věta 4.73 (Givensův QR-rozklad). *Bud' \mathbf{A} matice typu $m \times n$ a necht' $s = \min\{m-1, n\}$. Existuje s ortogonálních matic $\mathbf{Q}_1, \dots, \mathbf{Q}_s$ definovaných jako*

$$\mathbf{Q}_i := \mathbf{G}(i, m, \alpha) \mathbf{G}(i, m-1, \alpha) \cdots \mathbf{G}(i, i+1, \alpha),$$

takových, že pro

$$\mathbf{Q} = \mathbf{Q}_1^T \mathbf{Q}_2^T \cdots \mathbf{Q}_s^T$$

platí

$$\mathbf{A} = \mathbf{Q}\mathbf{R},$$

kde \mathbf{R} je matice $m \times n$ s nulami pod hlavní diagonálou.

Znáznorněme si schématicky Givensovu metodu redukce matice $\mathbf{A} \in \mathbb{R}^{3 \times 3}$ na horní trojúhelníkový tvar (symbol \bullet značí prvky, které se transformací nezměnily, a \pm značí prvky, které se změnily):

$$\begin{aligned}\mathbf{A} &= \begin{pmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{pmatrix} \xrightarrow{\mathbf{G}(1, 2, \alpha)} \begin{pmatrix} \pm & \pm & \pm \\ 0 & \pm & \pm \\ \bullet & \bullet & \bullet \end{pmatrix} \xrightarrow{\mathbf{G}(1, 3, \alpha)} \\ &\xrightarrow{\mathbf{G}(1, 3, \alpha)} \begin{pmatrix} \pm & \pm & \pm \\ 0 & \bullet & \bullet \\ 0 & \pm & \pm \end{pmatrix} \xrightarrow{\mathbf{G}(2, 3, \alpha)} \begin{pmatrix} \bullet & \bullet & \bullet \\ 0 & \pm & \pm \\ 0 & 0 & \pm \end{pmatrix} = \mathbf{R}.\end{aligned}$$

Příklad 4.74. Necht'

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}.$$

Givensovou metodou určete QR-rozklad.

Řešení. Krok 1.: Najděme c a s tak, aby

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} = \begin{pmatrix} \boxplus \\ 0 \end{pmatrix}.$$

Neboť $a_{11} = 0$ a $a_{21} = 1$, musí být $c = 0$ a $s = 1$, tedy

$$\mathbf{G}(1, 2, \alpha) = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Pak dostaneme

$$\tilde{\mathbf{A}} = \mathbf{G}(1, 2, \alpha)\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Nyní najděme c a s tak, aby

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} \tilde{a}_{11} \\ \tilde{a}_{31} \end{pmatrix} = \begin{pmatrix} \boxplus \\ 0 \end{pmatrix}.$$

Neboť $\tilde{a}_{11} = 1$ a $\tilde{a}_{31} = 1$, bude $c = \frac{1}{\sqrt{2}}$ a $s = \frac{1}{\sqrt{2}}$, tedy

$$\mathbf{G}(1, 3, \alpha) = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Celkem

$$\mathbf{A}^{(1)} = \mathbf{G}(1, 3, \alpha)\tilde{\mathbf{A}} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} \sqrt{2} & \frac{3}{\sqrt{2}} & 2\sqrt{2} \\ 0 & -1 & -1 \\ 0 & -\frac{1}{\sqrt{2}} & -\sqrt{2} \end{pmatrix}.$$

Krok 2.: Určeme c a s tak, aby

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} a_{22}^{(1)} \\ a_{32}^{(1)} \end{pmatrix} = \begin{pmatrix} \boxplus \\ 0 \end{pmatrix}.$$

Neboť $a_{22}^{(1)} = -1$ a $a_{32}^{(1)} = -\frac{1}{\sqrt{2}}$, bude $c = -\sqrt{\frac{2}{3}}$ a $s = -\frac{1}{\sqrt{3}}$, tedy

$$\mathbf{G}(1, 3, \Theta)\mathbf{A}^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\sqrt{\frac{2}{3}} & -\frac{1}{\sqrt{3}} \\ 0 & -\frac{1}{\sqrt{3}} & -\sqrt{\frac{2}{3}} \end{pmatrix} \begin{pmatrix} \sqrt{2} & \frac{3}{\sqrt{2}} & 2\sqrt{2} \\ 0 & -1 & -1 \\ 0 & -\frac{1}{\sqrt{2}} & -\sqrt{2} \end{pmatrix} = \begin{pmatrix} \sqrt{2} & \frac{3}{\sqrt{2}} & 2\sqrt{2} \\ 0 & \sqrt{\frac{3}{2}} & 2\sqrt{\frac{2}{3}} \\ 0 & 0 & \frac{1}{\sqrt{3}} \end{pmatrix} = R.$$

□

4.5.3 Srovnání algoritmů

Při výpočtu QR -rozkladu pomocí Householderovy matice je počet provedených operací roven číslu

$$n^2\left(3 - \frac{n}{3}\right).$$

K explicitnímu vyjádření matice \mathbf{Q} je navíc potřeba

$$2(m^2n - mn^2 + \frac{1}{3}n^3)$$

operací, tedy celkem

$$2m^2n - mn^2 + \frac{1}{3}n^3.$$

Pro QR -rozklad pomocí Givensovy matice je tento počet dvojnásobný, tj.

$$2n^2\left(3 - \frac{n}{3}\right).$$

Ovšem pokud v metodě s Givensovou maticí nahradíme matici rotace $\begin{pmatrix} c & s \\ -s & c \end{pmatrix}$ a matice odrazu $\begin{pmatrix} c & s \\ s & -c \end{pmatrix}$ maticemi $\begin{pmatrix} 1 & a \\ -a & 1 \end{pmatrix}$ a $\begin{pmatrix} a & 1 \\ 1 & -a \end{pmatrix}$ s ortogonálními sloupci, pak se nám podaří snížit počet operací na úroveň metody využívající Householderovy matice – jedná se o tzv. **matice rychlé Givensovy transformace**.

Householderova matice má však tu nevýhodu, že v matici, kterou ji násobíme, nám změní všechny prvky (zatímco Givensova matice jen i -tý a k -tý řádek), takže může například z řídké matice vytvořit matici plnou.

V modifikované metodě s Gram-Schmidtovým algoritmem je počet operací mn^2 .

4.5.4 QR-rozklad a vlastní čísla matice \mathbf{A} – QR -algoritmus

Základní **QR-algoritmus**: Mějme matici \mathbf{A} . Sestrojme její QR -rozklad, tj.

$$\mathbf{A} = \mathbf{A}_0 = \mathbf{Q}_0\mathbf{R}_0,$$

určíme

$$\mathbf{A}_1 := \mathbf{R}_0\mathbf{Q}_0.$$

Nyní sestrojme QR -rozklad matice \mathbf{A}_1 , tj.

$$\mathbf{A}_1 = \mathbf{Q}_1\mathbf{R}_1,$$

a spočtíme

$$\mathbf{A}_2 = \mathbf{R}_1\mathbf{Q}_1.$$

Takto pokračujeme analogicky dále.

Jistě platí

$$\begin{aligned} \mathbf{A}_{k+1} &= \mathbf{R}_k\mathbf{Q}_k = \mathbf{Q}_k^T\mathbf{A}_k\mathbf{Q}_k = \mathbf{Q}_k^T\mathbf{R}_{k-1}\mathbf{Q}_{k-1}\mathbf{Q}_k = \\ &= \mathbf{Q}_k^T\mathbf{Q}_{k-1}^T\mathbf{A}_{k-1}\mathbf{Q}_{k-1}\mathbf{Q}_k = \cdots = (\mathbf{Q}_0\mathbf{Q}_1\cdots\mathbf{Q}_k)^T\mathbf{A}(\mathbf{Q}_0\mathbf{Q}_1\cdots\mathbf{Q}_k). \end{aligned}$$

Tedy matice $\mathbf{A}_0, \mathbf{A}_1, \dots$ jsou *kongruentní* (tj. $\mathbf{A} \equiv \mathbf{B} \iff \mathbf{A} = \mathbf{P}^T \mathbf{B} \mathbf{P}$). Navíc díky ortogonálnosti matic $\mathbf{Q}_0, \mathbf{Q}_1, \dots$ jsou matice $\mathbf{A}_0, \mathbf{A}_1, \dots$ také *podobné* (tj. $\mathbf{A} \sim \mathbf{B}$ (též $\mathbf{A} \approx \mathbf{B}$) $\iff \mathbf{A} = \mathbf{P}^{-1} \mathbf{B} \mathbf{P}$). Tyto matice mají díky podobnosti stejná vlastní čísla jako matice \mathbf{A} . Posloupnost těchto matic konverguje za určitých předpokladů k horní trojúhelníkové (resp. horní blokově trojúhelníkové) matici, která má vlastní čísla na diagonále (resp. diagonální bloky mají vlastní čísla se stejnou absolutní hodnotou) seřazena podle velikosti počínaje největším vlastním číslem. Poddiagonální prvky (resp. poddiagonální bloky) konvergují k nule. Ovšem důkazy konvergence existují jen pro některé speciální typy matic. Například má-li matice \mathbf{A} kladná vlastní čísla, pak \mathbf{Q}_k konverguje k jednotkové matici a posloupnost matic \mathbf{A}_k k horní trojúhelníkové matici, přičemž diagonální prvky této matice jsou vlastní čísla matice \mathbf{A} .

Příklad 4.75. Určete vlastní čísla matice

$$\mathbf{A} = \begin{pmatrix} 2 & \frac{1}{2} & 1 \\ 3 & -\frac{3}{2} & 1 \\ 0 & \frac{11}{9} & \frac{5}{3} \end{pmatrix}.$$

Řešení. Lze snadno ověřit, že vlastní čísla matice \mathbf{A} jsou $\lambda_1 = 1$, $\lambda_2 = -2$ a $\lambda_3 = 3$. Výsledky získané *QR*-algoritmem:

k	$a_{11}^{(k)}$	$a_{22}^{(k)}$	$a_{33}^{(k)}$
1	2,0	-1,6666667	1,6666667
5	3,1781374	-2,2260322	1,0478949
10	2,9486278	-1,9471270	0,9984996
15	3,0003596	-2,0064061	1,0000468
20	2,9991547	-1,9991527	0,9999984
25	3,0001104	-2,0001098	0,9999999

□

Ke zrychlení konvergence lze využít tzv. *posunutí* a počítat nikoli rozklad matice $\mathbf{A}_k = \mathbf{Q}_k \mathbf{R}_k$, nýbrž matice

$$\mathbf{A}_k - \sigma_k \mathbf{I} = \tilde{\mathbf{Q}}_k \tilde{\mathbf{R}}_k.$$

Původní spektrum matice \mathbf{A} se tímto posune o σ_k (je výhodné volit jej jako nějakou aproximaci vlastního čísla; matice $\mathbf{A}_k - \sigma_k \mathbf{E}$ má vlastní čísla $\lambda_j - \sigma_k$, jsou-li λ_j vlastní čísla matice \mathbf{A}_k). Zaznamenáváme-li velikosti posunutí σ_k , snadno ze znalosti spektra matice \mathbf{A}_k najdeme spektrum matice \mathbf{A} . Protože ještě (v metodě bez posunutí)

$$\mathbf{A}_k = (\mathbf{Q}_0 \mathbf{Q}_1 \cdots \mathbf{Q}_{k-1})^T \mathbf{A} (\mathbf{Q}_0 \mathbf{Q}_1 \cdots \mathbf{Q}_{k-1}),$$

je

$$\mathbf{A} = (\mathbf{Q}_0 \mathbf{Q}_1 \cdots \mathbf{Q}_{k-1})^T \mathbf{A}_k (\mathbf{Q}_0 \mathbf{Q}_1 \cdots \mathbf{Q}_{k-1}).$$

Vlastní vektory \mathbf{y} matice \mathbf{A} dostaneme z vlastních vektorů matice \mathbf{A}_k podle vzorce

$$\mathbf{y} = \mathbf{Q}_0 \mathbf{Q}_1 \cdots \mathbf{Q}_{k-1} \mathbf{z}.$$

Týž vzorec (s maticemi $\tilde{\mathbf{Q}}_j$) zůstává v platnosti i pro metodu s posunutími, protože při posunutí se vlastní vektory nemění.

Volí-li se posunutí speciálně, dostáváme v některých důležitých případech i kubickou konvergenci (tzn. zhruba řečeno, počet platných míst se v každém kroku přibližně ztrojnásobí).

Poznámka 4.76. Nejvýhodnější se jeví upravit nejdříve matici \mathbf{A} do tzv. *Hessenbergova tvaru* (tj. $a_{ij} = 0$ pro $j < i - 1$, $i, j = 1, \dots, n$) pomocí Gaussovy eliminace a pak na tuto upravenou matici použít *QR-rozklad*. Konvergence je potom rychlejší (obzvláště použijeme-li metodu posunu, kde za σ_k volíme tzv. *Rayleighův podíl*

$$\frac{(\mathbf{e}^k)^T \mathbf{H} \mathbf{e}^k}{(\mathbf{e}^k)^T \mathbf{e}^k},$$

kde \mathbf{H} je právě matice \mathbf{A} v Hessenbergově tvaru). Navíc platí, že je-li matice \mathbf{A} v Hessenbergově tvaru, pak každá z matic \mathbf{H}_k je také v Hessenbergově tvaru, a to i při metodě posunutí.

4.6 Podmíněnost problému vlastních čísel

Důležitou charakteristikou libovolného problému je jeho *podmíněnost*, která udává, jak významně se změní řešení problému, pokud změňme vstupní hodnoty. Podmíněnost problému vlastních čísel můžeme popsat pomocí tzv. *globálního čísla podmíněnosti*.

4.6.1 Globální číslo podmíněnosti

Vzhledem k zaokrouhlování řešíme ve skutečnosti problém

$$(\mathbf{A} + \mathbf{E} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0}$$

namísto

$$(\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0}.$$

Následkem zaokrouhlování při výpočtu dostáváme řešení problému $\tilde{\lambda}$, $\tilde{\mathbf{x}}$, který je přesným řešením problému s poruchou

$$(\mathbf{A} - \mathbf{E}_M - \tilde{\lambda} \mathbf{I})\tilde{\mathbf{x}} = \mathbf{0},$$

kde \mathbf{E}_M zahrnuje zaokrouhlovací chyby během výpočtu.

Poznámka 4.77. Analýza stability vlastního problému je velmi složitá a dá se uapokojivě provést jen pro jednoduché vlastní číslo anebo pro matici, která je diagonalizovatelná.

Definice 4.78. Matice \mathbf{A} je *diagonalizovatelná*, když existuje regulární matice \mathbf{X} taková, že

$$\mathbf{X}^{-1} \cdot \mathbf{A} \cdot \mathbf{X} = \mathbf{D},$$

kde \mathbf{D} je diagonální matice.

Věta 4.79. Pokud je \mathbf{A} diagonalizovatelná matice s vlastními čísly $\lambda_1, \dots, \lambda_n$, potom vlastní čísla matice $\mathbf{A} + \mathbf{E}$ leží v jednotkovém kruhu

$$K_i = \{z; |z - \lambda_i| \leq c(\mathbf{X}) \cdot \|\mathbf{E}\|\},$$

kde $c(\mathbf{X}) = \|\mathbf{X}\| \|\mathbf{X}^{-1}\|$ je číslo podmíněnosti matice vlastních vektorů v maticové normě $\|\cdot\|$.

Důkaz. Necht' $(\mathbf{A} + \mathbf{E})\mathbf{x} = \lambda\mathbf{x}$. Potom buď $\lambda = \lambda_i$ pro nějaký index i a potom $\lambda \in K_i$, a nebo $\lambda \neq \lambda_i$ pro $i = 1, \dots, n$. Potom $\lambda\mathbf{I} - \mathbf{A}$ je regulární matice. Matice

$$(\lambda\mathbf{I} - \mathbf{A})^{-1}(\lambda\mathbf{I} - \mathbf{A} - \mathbf{E}) = \mathbf{I} - (\lambda\mathbf{I} - \mathbf{A})^{-1}\mathbf{E}$$

je singulární. Proto podle vztahu

$$\rho[(\lambda\mathbf{I} - \mathbf{A})^{-1}\mathbf{E}] \geq 1$$

platí

$$\begin{aligned} 1 &\leq \|(\lambda\mathbf{I} - \mathbf{A})^{-1}\mathbf{E}\| = \|(\lambda\mathbf{I} - \mathbf{XDX}^{-1})^{-1}\mathbf{E}\| = \|\mathbf{X}(\lambda\mathbf{I} - \mathbf{D})^{-1}\mathbf{X}^{-1}\mathbf{E}\| \leq \\ &\leq \|\mathbf{X}\|\|\mathbf{X}^{-1}\|\|\mathbf{E}\|\|(\lambda\mathbf{I} - \mathbf{D})^{-1}\| = c(\mathbf{X})\|\mathbf{E}\|\|(\lambda\mathbf{I} - \mathbf{D})^{-1}\| = \\ &= c(\mathbf{X})\|\mathbf{E}\| \max_i \left[\frac{1}{|\lambda - \lambda_i|} \right]. \end{aligned}$$

Přitom jsme využili skutečnost, že maticová norma diagonální matice je daná jejím maximálním diagonálním prvkem v absolutní hodnotě. Odtud

$$\min_i |\lambda - \lambda_i| \leq c(\mathbf{X})\|\mathbf{E}\|.$$

□

4.6.2 Odhad chyby vypočítaného vlastního čísla

Přesnost vypočítaného vlastního čísla a vlastního vektoru ověřujeme pomocí reziduí (zbytků).

Věta 4.80 (Odhad chyby vypočítaného vlastního čísla). *Necht' určené vlastní číslo $\tilde{\lambda}$ a k němu příslušný vypočítaný vlastní vektor $\tilde{\mathbf{x}}$ dávají (přesný) reziduální vektor*

$$\mathbf{r} = \mathbf{A}\tilde{\mathbf{x}} - \tilde{\lambda}\tilde{\mathbf{x}}.$$

Potom $\tilde{\lambda}$, $\tilde{\mathbf{x}}$ jsou přesné hodnoty vlastního čísla a vlastního vektoru matice s poruchou $\mathbf{A} + \mathbf{E}$, kde

$$\mathbf{E} = -\frac{\mathbf{r}\tilde{\mathbf{x}}^*}{\|\tilde{\mathbf{x}}\|_2^2}$$

a platí odhad

$$|\lambda - \tilde{\lambda}| \leq \frac{\|\mathbf{y}\|_2\|\mathbf{x}\|_2\|\mathbf{r}\|_2}{|\mathbf{y}^T\mathbf{x}|\|\tilde{\mathbf{x}}\|_2}, \quad (4.10)$$

kde \mathbf{y} je levý vlastní vektor matice \mathbf{A} příslušný vlastnímu číslu λ .

Důkaz.

$$\begin{aligned} \left(\mathbf{A} - \frac{\mathbf{r}\tilde{\mathbf{x}}^*}{\|\tilde{\mathbf{x}}\|_2^2} \right) \tilde{\mathbf{x}} &= \mathbf{A}\tilde{\mathbf{x}} - \frac{\|\tilde{\mathbf{x}}\|_2^2}{\|\tilde{\mathbf{x}}\|_2^2} \mathbf{r} = \mathbf{A}\tilde{\mathbf{x}} - \mathbf{r} = \tilde{\lambda}\tilde{\mathbf{x}} \\ |\lambda - \tilde{\lambda}| &\leq \frac{\|\mathbf{y}\|_2\|\mathbf{x}\|_2\|\mathbf{E}\|_2}{|\mathbf{y}^T\mathbf{x}|} + O\left(\frac{\|\mathbf{E}\|_2}{\|\mathbf{A}\|_2} \right)^2. \end{aligned}$$

Když

$$\|\mathbf{E}\|_2 = \frac{\|\mathbf{r}\|_2\|\tilde{\mathbf{x}}^*\|_2}{\|\tilde{\mathbf{x}}\|_2} = \frac{\|\mathbf{r}\|_2}{\|\tilde{\mathbf{x}}\|_2}$$

dosadíme do nerovnosti, dostáváme

$$|\lambda - \tilde{\lambda}| \leq \frac{\|\mathbf{y}\|_2 \|\mathbf{x}\|_2 \|\mathbf{r}\|_2}{|\mathbf{y}^T \mathbf{x}| \|\tilde{\mathbf{x}}\|_2} + O\left(\frac{\|\mathbf{E}\|_2}{\|\mathbf{A}\|_2}\right)^2 = c(\lambda) \frac{\|\mathbf{r}\|_2}{\|\tilde{\mathbf{x}}\|_2} + O\left(\frac{\|\mathbf{E}\|_2}{\|\mathbf{A}\|_2}\right)^2.$$

□

Poznámka 4.81. 1. Pro symetrické matice

$$c(\lambda) = \frac{\|\mathbf{y}\|_2 \|\mathbf{x}\|_2}{|\mathbf{y}^T \mathbf{x}|} = 1 \quad a \quad |\lambda - \tilde{\lambda}| \leq \frac{\|\mathbf{r}\|_2}{\|\tilde{\mathbf{x}}\|_2}$$

2. V praxi přesné hodnoty \mathbf{x} , \mathbf{y} neznáme, proto se ve vztahu (7.1) nahrazují hodnotami $\tilde{\mathbf{x}}$, $\tilde{\mathbf{y}}$, tj.

$$|\lambda - \tilde{\lambda}| \approx \frac{\|\tilde{\mathbf{y}}\|_2 \|\tilde{\mathbf{x}}\|_2 \|\mathbf{r}\|_2}{|\tilde{\mathbf{y}}^T \tilde{\mathbf{x}}| \|\tilde{\mathbf{x}}\|_2} = \frac{\|\tilde{\mathbf{y}}\|_2 \|\mathbf{r}\|_2}{|\tilde{\mathbf{y}}^T \tilde{\mathbf{x}}|} \quad (4.11)$$

Relativní chyba vypočítaného vlastního čísla

Pro jednoduché vlastní číslo $\lambda \neq 0$ můžeme pomocí (7.1) vyjádřit relativní chybu takto

$$\begin{aligned} \frac{|\Delta\lambda|}{|\lambda|} &\approx \left| \varepsilon \frac{\mathbf{y}^T \mathbf{B} \mathbf{x}}{\mathbf{y}^T \mathbf{x}} \frac{1}{\lambda} \right| \leq \varepsilon \frac{\|\mathbf{y}\|_2 \|\mathbf{x}\|_2 \|\mathbf{B}\|_2}{|\mathbf{y}^T \mathbf{x}| |\lambda|} = \varepsilon c(\lambda) \frac{\|\mathbf{B}\|_2}{|\lambda|} \approx \\ &\approx \varepsilon c(\lambda) \frac{\|\mathbf{A}\|_2}{|\lambda|} = \varepsilon c(\lambda) \frac{\|\mathbf{A}\|_2}{\rho(\mathbf{A})} \frac{\rho(\mathbf{A})}{|\lambda|}. \end{aligned}$$

Pojmy k zapamatování

- Zopakovali jsme si definici vlastních čísel matice a jejich základní vlastnosti.
- Ukázali jsme si, jak můžeme odhadnout polohu vlastních čísel.
- Seznámili jsme se s vybranými metodami určení vlastních čísel, a to jak pomocí finitních metod, tak i pomocí numerických metod.
- Řešili jsme jak úplný problém vlastních čísel, tj. nalezení všech vlastních čísel matice, tak i částečný problém vlastních čísel, tj. nalezení dominantního vlastního čísla.

Kontrolní otázky

1. Co rozumíme pojmem dominantní vlastní číslo.?
2. Jsou vlastní čísla definována pro libovolnou matici?
3. Kde se využívá vlastních čísel matice a jejich vlastností?

4.7 Cvičení

Maplety

Příklady pro samostatnou práci jsou uvedeny v samostatné sbírce příkladů. Pomocí následujících mapletů si můžete usnadnit některé dílčí výpočty, nebo zkontrolovat jejich správnost.

1. [Nalezení charakteristického polynomu, vlastních čísel a vektorů](#)
2. [Výpočet determinantu matice](#)
3. [Násobení matic](#)

webMathematica

Příklady pro samostatnou práci jsou uvedeny v samostatné sbírce příkladů. Pomocí následujících aplikací si můžete usnadnit některé dílčí výpočty, nebo zkontrolovat jejich správnost.

1. [Výpočet vlastních vektorů](#)

5 Soustavy nelineárních rovnic

Průvodce studiem

V druhé kapitole jsme se věnovali řešení soustav lineárních algebraických rovnic. V nejrůznějších aplikacích se ale vyskytují i nelineární rovnice (těm jsme se věnovali ve třetí kapitole) a soustavy nelineárních rovnic.

Cílem této kapitoly je seznámit čtenáře s numerickými metodami řešení soustav nelineárních rovnic. S těmito pojmy jste se setkali už v předmětu Matematika 3. Zde uvádíme opět všechny souvislosti a přesné znění vět a podmínek konvergence.

Nebudeme schopni řešit každou soustavu nelineárních rovnic. Pouze pokud budou splněny konvergenční podmínky, jsme schopni najít řešení. Proto je důležité se umět vhodně přiblížit k řešení a nebo umět odhadnout polohu řešení.

Z celé řady používaných metod si ukážeme použití prosté iterační metody a Newtonovy metody pro soustavy nelineárních rovnic a stanovíme si podmínky pro konvergenci těchto metod.

Opět nás bude zajímat i rychlost konvergence a možnost odhadu chyby po k -tém kroku iterace.

Cíle

Po prostudování této kapitoly budete schopni:

- Rozhodnout, zda je možné danou soustavu rovnic řešit.
- Vybrat vhodnou metodu pro řešení dané soustavy.
- Provést odhad počtu kroků, které budou garantovat dosažení požadované přesnosti.

5.1 Formulace úlohy

Soustavu n nelineárních rovnic si můžeme zapsat ve tvaru

$$F_1(x_1, x_2, \dots, x_n) = 0,$$

$$F_2(x_1, x_2, \dots, x_n) = 0,$$

$$\vdots \quad \quad \quad \vdots$$

$$F_n(x_1, x_2, \dots, x_n) = 0.$$

Nebo ve vektorovém tvaru

$$\mathcal{F}(X) = \mathcal{O},$$

kde $\mathcal{F}(X) = (F_1(X), F_2(X), \dots, F_n(X))^T$, $X = (x_1, x_2, \dots, x_n)$, \mathcal{O} je nulový sloupec.

5.2 Metoda prosté iterace

Nechť \mathcal{F} je spojitě zobrazení na oblasti D . Rovnici $\mathcal{F}(X) = \mathcal{O}$ si přepíšeme na tvar $X^T = \Phi(X)$, zvolíme si $X_0 \in D$ a iterační vztah

$$(X_{k+1})^T = \Phi(X_k), \quad \Phi(X) = (\varphi_1(X), \varphi_2(X), \dots, \varphi_n(X))^T.$$

Věta 5.1. *Nechť v uzavřené oblasti D platí*

$$1) \quad \Phi(X) \in D \quad \forall X \in D,$$

$$2) \quad \exists q \in [0, 1) \text{ tak, že } \|\Phi(X) - \Phi(Y)\| \leq q\|X - Y\| \quad \forall X, Y \in D,$$

potom posloupnost $\{X_k\}$ konverguje k jedinému řešení \mathcal{R} a platí

$$\|X_k - \mathcal{R}\| \leq \frac{q}{1 - q} \|X_k - X_{k-1}\|.$$

Důkaz. Plyne z Banachovy věty o pevném bodě 1.38. □

Poznámka 5.2. Jestliže Φ je diferencovatelná, pak podmínku 2 můžeme nahradit podmínkou

$$2') \quad \exists q : \|\Phi'(X)\| \leq q < 1 \quad \forall X \in D,$$

kde $\Phi'(X)$ je matice s prvky

$$a_{ij} = \frac{\partial \varphi_i(x_1, \dots, x_n)}{\partial x_j}.$$

Maticová norma v 2') přitom musí být souhlasná s vektorovou normou.

Většinou bývá obtížné najít Φ a D tak, abychom měli zajištěnou konvergenci, zvláště u rozsáhlejších soustav. Někdy se užívá metoda pokusů. Pozor, v takovém případě je nutno zvláště pečlivě prověřit výsledek.

Příklad 5.3. Metodou prosté iterace najděte kladné řešení (pokud existuje) pro soustavu

$$\begin{aligned} \sin(x + 1) - y &= 0.5, \\ x^2 + y^2 &= 1. \end{aligned}$$

Řešení. Jde o průsečík kružnice se středem v počátku a o jednotkovém poloměru s posunutou goniometrickou funkcí.

Grafickou metodou můžeme odhadnout polohu kořene

$$0.8 \leq x \leq 1, \quad 0.4 \leq y \leq 0.5.$$

Upravíme si rovnice do iteračního tvaru

$$\begin{aligned} x &= \sqrt{1 - y^2}, \\ y &= \sin(x + 1) - 0.5. \end{aligned}$$

Potom

$$\Phi'(X) = \begin{pmatrix} 0 & -\frac{y}{\sqrt{1 - y^2}} \\ \cos(x + 1) & 0 \end{pmatrix},$$

$$\|\Phi'(X)\| \doteq 0.5,$$

při použití řádkové normy. Zvolíme si počáteční aproximaci $x_0 = 0.85$, $y_0 = 0.46$ a dosazením do iteračních vztahů dostaneme

$$x_1 = \sqrt{1 - (y_0)^2} \doteq 0.882,$$

$$y_1 = \sin(x_0 + 1) - 0.5 \doteq 0.45.$$

Výpočet opakujeme. Výsledky si zapíšeme do tabulky

k	x	y
0	0.85	0.46
1	0.882	0.45
2	0.893	1.448
3	0.894	0.448

Po dosazení posledních hodnot do jedné z rovnic dostaneme

$$(x_3)^2 + (y_3)^2 = 0.9999.$$

Chyba je řádově 10^{-4} . Pokud je tato přesnost postačující, ukončujeme výpočet, pokud ne, pokračuje dále ve výpočtech. □

Příklad 5.4. Metodou prosté iterace najděte řešení soustavy

$$\begin{aligned} x &= 0.2 + 0.1(-xy^2 + 3x), \\ y &= 0.6 + 0.1(x^2y^3 - 2y), \end{aligned}$$

kteří leží v oblasti $\Omega = \{0 \leq x \leq 1, 0 \leq y \leq 1\}$.

Řešení. Nejdříve si ověříme první podmínku věty 5.1, tj. že $\Phi(\Omega) \subset \Omega$.

$$\begin{aligned} 0 &\leq 0.2 + 0.1(-xy^2 + 3x) \leq 1, \\ 0 &\leq 0.6 + 0.1(x^2y^3 - 2y) \leq 1. \end{aligned}$$

Potom

$$\Phi'(X) = \begin{pmatrix} -0.1y^2 + 0.3 & -0.2xy \\ -0.2xy^3 & -0.3x^2y^2 - 0.2 \end{pmatrix}.$$

Určíme si odhad $\|\Phi'(X)\|$ v řádkové normě a dostaneme

$$\begin{aligned} \|\Phi'(X)\| &= \max_{x,y \in \langle 0;1 \rangle} \{|-0.1y^2 + 0.3| + |0.2xy|, |0.2xy^3| + |-0.3x^2y^2 - 0.2|\} \\ &\leq \max\{0.3 + 0.2, 0.2 + 0.5\} = 0.7. \end{aligned}$$

Máme splněnu druhou podmínku věty 5.1. Protože máme zaručenou konvergenci metody, můžeme iterovat. Volíme $x_0 = y_0 = 0$ a postupně dostáváme výsledky, které si opět zapíšeme do tabulky

k	x	y
0	0	0
1	0,2	0,6
2	0,2528	0,479136
3	0,270036	0,503402
...
8	0,275882	0,499209
9	0,27589	0,499211

Rozdíl posledních dvou iterací je řádově 10^{-5} . Pokud je tato přesnost je postačující, ukončujeme výpočet, pokud ne, pokračuje dále v iterování. \square

5.3 Newtonova metoda

Mějme soustavu $\mathcal{F}(X) = \mathcal{O}$, kde vektorová funkce \mathcal{F} je diferencovatelná v oblasti D . Potom rozvojem do Taylorovy řady v okolí bodu $X_0 = (x_1^0, x_2^0, \dots, x_n^0)$ dostaneme

$$\mathcal{F}(X) = \mathcal{F}(X_0) + \mathcal{F}'(X_0)(X - X_0) + \dots,$$

kde

$$\mathcal{F}' = \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1} & \frac{\partial F_n}{\partial x_2} & \cdots & \frac{\partial F_n}{\partial x_n} \end{pmatrix},$$

je Jakobián funkce \mathcal{F} , resp. v souřadnicích pro $i = 1, 2, \dots, n$

$$F_i(X) = F_i(X_0) + \sum_{j=1}^n \frac{\partial F_i(X_0)}{\partial x_j} (x_j - x_j^0) + \dots$$

Zanedbáním vyšších řádů pak dostaneme soustavu

$$\mathcal{F}(X) \approx \mathcal{F}(X_0) + \mathcal{F}'(X_0)(X - X_0).$$

Předpokládáme, že $\mathcal{F}(X) = \mathcal{O}$. Je-li \mathcal{F}' regulární, potom má soustava jediné řešení

$$X_1 = X_0 - (\mathcal{F}'(X_0))^{-1}\mathcal{F}(X_0).$$

Jestliže je \mathcal{F}' regulární v bodě X_1 , postup opakujeme. Máme tak definovaný iterační proces

$$X_{n+1} = X_n - (\mathcal{F}'(X_n))^{-1}\mathcal{F}(X_n).$$

Problémy nastávají vždy s určením inverzní matice. Vezmeme si nejjednodušší případ: $\mathcal{F}(X) = \mathcal{O}$ je soustava dvou rovnic o dvou neznámých. Mějme

$$F_1(x, y) = 0,$$

$$F_2(x, y) = 0.$$

Nechť F_1, F_2 jsou diferencovatelné v oblasti D . Potom rozvoj do Taylorovy řady nám v okolí bodu (x_k, y_k) dává pro $i = 1, 2$

$$F_i(x, y) = F_i(x_k, y_k) + \frac{\partial F_i(x_k, y_k)}{\partial x}(x - x_k) + \frac{\partial F_i(x_k, y_k)}{\partial y}(y - y_k) + \dots$$

a opět zanedbáme členy vyšších řádů a dostaneme aproximaci lineární funkcí.

$$\mathcal{F}(X) \approx \mathcal{F}(X_k) + \mathcal{F}'(X_k)(X - X_k),$$

kde \mathcal{F}' je Jacobián funkce \mathcal{F} ,

$$\mathcal{F}'(X) = \begin{pmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} \end{pmatrix}, \quad (X - X_k) = \begin{pmatrix} x - x_k \\ y - y_k \end{pmatrix}.$$

Je-li \mathcal{F}' regulární v okolí bodu X_k , má naše soustava jediné řešení X_{k+1} .

$$X_{k+1} = X_k - (\mathcal{F}'(X_k))^{-1}(\mathcal{F}(X_k)),$$

resp. v souřadnicích

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ y_k \end{pmatrix} - \frac{1}{\det \mathcal{F}'(X_k)} \begin{pmatrix} \frac{\partial F_2}{\partial y} & -\frac{\partial F_1}{\partial y} \\ -\frac{\partial F_2}{\partial x} & \frac{\partial F_1}{\partial x} \end{pmatrix} \begin{pmatrix} F_1(x_k, y_k) \\ F_2(x_k, y_k) \end{pmatrix}.$$

Zde jsme využili vzorce z algebry: Pro $ad - bc \neq 0$ platí

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Výpočet ukončíme při splnění podmínky $\|X_{k+1} - X_k\| < \varepsilon$.

Příklad 5.5. Určete řešení soustavy rovnic:

$$F_1(x, y) = x^2 - 0.4x + y^2 - 4 = 0,$$

$$F_2(x, y) = x^2y + y - 1 = 0.$$

Řešení.

$$\frac{\partial F_1}{\partial x} = 2x - 0.4, \quad \frac{\partial F_1}{\partial y} = 2y,$$

$$\frac{\partial F_2}{\partial x} = 2xy, \quad \frac{\partial F_2}{\partial y} = x^2 + 1,$$

$$\begin{aligned} |\mathcal{F}'| &= \begin{vmatrix} 2x - 0.4 & 2y \\ 2xy & x^2 + 1 \end{vmatrix} = (x^2 + 1)(2x - 0.4) - 4xy^2 = \\ &= 2x^3 + 2x - 0.4x^2 - 0.4 - 4xy^2. \end{aligned}$$

Provedeme separaci kořenů a dostaneme $x \in (2; 3)$, $y \in (0; 1)$. Pro určení oblasti D si stačí uvědomit, že $F_1(x, y) = 0$ je rovnicí kružnice a y nemůže nabývat záporných hodnot. Rovnice $F_2(x, y) = 0$ si můžeme upravit na tvar $y(x^2 + 1) = 1$, odkud už plyne, že $y \geq 0$. $|\mathcal{F}'(2.2; 0.2)| = 23.008$. V okolí tohoto bodu je derivace nenulová, tj. můžeme začít iterovat podle vztahů:

$$\begin{aligned} x_{k+1} &= x_k - \frac{(x_k^2 + 1)(x_k^2 - 0.4x_k + y_k^2 - 4) - 2y_k(x_k^2y_k + y_k - 1)}{2x_k^3 + 2x_k - 0.4x_k^2 - 0.4 - 4x_ky_k^2}, \\ y_{k+1} &= y_k - \frac{(-2x_ky_k)(x_k^2 - 0.4x_k + y_k^2 - 4) + (2x_k - 0.4)(x_k^2y_k + y_k - 1)}{2x_k^3 + 2x_k - 0.4x_k^2 - 0.4 - 4x_ky_k^2}. \end{aligned}$$

□

Poznámka 5.6. Newtonovou metodou můžeme určit i kořen komplexní rovnice $f(z) = 0$. Necht $z = x + jy$, $Re f = u(x, y)$, $Im f = v(x, y)$, potom máme soustavu dvou rovnic

$$u(x, y) = 0,$$

$$v(x, y) = 0,$$

kteřou řešíme podle výše uvedených vztahů.

Příklad 5.7. Newtonovou metodou řešte soustavu

$$f_1 : x^2 + y^2 + 2x - 3 = 0,$$

$$f_2 : x^2 + y^2 - 4y - 5 = 0.$$

Řešení. Nejdříve ovedeme separaci kořenů. Upravíme si první rovnici, tj. doplníme ji na úplný kvadrát.

$$\begin{aligned} x^2 + 2x + y^2 - 3 &= 0, \\ (x + 1)^2 - 1 + y^2 - 3 &= 0, \end{aligned}$$

$$(x + 1)^2 + y^2 = 4.$$

Dostali jsme rovnici kružnice se středem $S = (-1; 0)$ a poloměrem $r = 2$. Analogicky pro druhou rovnici dostaneme

$$\begin{aligned} x^2 + y^2 - 4y - 5 &= 0, \\ x^2 + (y - 2)^2 - 4 - 5 &= 0, \\ x^2 + (y - 2)^2 &= 9. \end{aligned}$$

Dostali jsme rovnici kružnice se středem $S = (0; 2)$ a poloměrem $r = 3$. Grafickou metodou odhadneme polohu kořenů

$$\Omega = \{0.8 < x < 0.9, -0.9 < y < -1\}.$$

Určíme si matici inverzní

$$\mathcal{F}'(X) = \begin{pmatrix} 2(x+1) & 2y \\ 2x & 2(y-2) \end{pmatrix},$$

Sestavíme si iterační rovnice

$$\begin{aligned} x^{k+1} &= x^k + \frac{1}{2(y^k - 2x^k) - 4} [y^k f_2 - (y^k - 2)f_1], \\ y^{k+1} &= y^k + \frac{1}{2(y^k - 2x^k) - 4} [x^k f_1 - (x^k + 1)f_2]. \end{aligned}$$

Protože máme zaručenou konvergenci metody, můžeme iterovat, přičemž výsledky si opět zapíšeme do tabulky

k	x	y
0	0.8	-0.9
1	0.8111	-0.9056
2	0.7889	-0.8845
...

Výpočet ukončíme v závislosti na požadované přesnosti. □

Příklad 5.8. Newtonovou metodou řešte soustavu

$$\begin{aligned} f(x, y) : x^3 - xy^2 - 1 &= 0, \\ g(x, y) : y^3 - 2x^2y + 2 &= 0. \end{aligned}$$

Řešení. Odhadneme polohu kořenů. Ze tří kořenů si vybereme levý horní, který leží v oblasti

$$\Omega = \{-2 \leq x \leq -1, 1 \leq y \leq 2\}.$$

Výpočet ukončíme, když

$$\max\{|f(x^{k+1}, y^{k+1})|, |g(x^{k+1}, y^{k+1})|\} < 10^{-5}.$$

Volíme $x^0 = -1$, $y^0 = 1$ a dostaneme

k	x	y
0	-1	1
1	-1.5	2
2	-1.379562	1.673966
3	-1.392137	1.629879
4	-1.394072	1.631182
5	-1.394069	1.631182

Všimněte si, že v okolí kořene $x = -1.39407, y = 1.63118$ je konvergence velmi rychlá. \square

Příklad 5.9. Newtonovou metodou řešte soustavu

$$\begin{aligned} e^{x^2+y^2} - 1 &= 0, \\ e^{x^2-y^2} - 1 &= 0. \end{aligned}$$

Řešení. Je zřejmé, že řešení je $x = 0, y = 0$.

Pokud zvolíme počáteční bod jako $x^0 = \begin{pmatrix} 0.1 \\ 0.1 \end{pmatrix}$, potom po 15 iteracích dostaneme řešení $x = 0.61 \cdot 10^{-5}, y^{(15)} = 0.61 \cdot 10^{-5}$.

Pokud zvolíme počáteční bod jako $x^0 = \begin{pmatrix} 10 \\ 10 \end{pmatrix}$, potom dostaneme řešení po 220 iteracích.

Pokud zvolíme počáteční bod jako $x^0 = \begin{pmatrix} 20 \\ 20 \end{pmatrix}$, potom nám bude Newtonova metoda havarovat. \square

5.4 Gradientní metoda

Nechť $\bar{f}(\bar{x}) = (f_1(\bar{x}), f_2(\bar{x}), \dots, f_n(\bar{x}))$, kde $\bar{x} = (x_1, x_2, \dots, x_n), n \geq 2$. Soustava rovnic

$$\bar{f}(\bar{x}) = \mathcal{O}, \quad (5.1)$$

je ekvivalentní jedné rovnici

$$\Psi(\bar{x}) = \mathcal{O}, \quad (5.2)$$

kde $\Psi(\bar{x}) = f_1^2(\bar{x}) + f_2^2(\bar{x}) + \dots + f_n^2(\bar{x})$.

Je zřejmé, že řešením rovnice (5.2) budou body nulových minim funkce

$$\Psi(\bar{x}) = \Psi(x_1, x_2, \dots, x_n).$$

Předpokládejme, že funkce $\Phi(\bar{x})$ je dvakrát spojitě diferencovatelná v oblasti D , která obsahuje izolované řešení $\bar{x}^* = (x_1^*, x_2^*, \dots, x_n^*)$. Necht' v okolí \bar{x}^* jsou vrstevnice funkce Ψ ve tvaru (viz obr. DOPLNIT)

Volíme počáteční aproximaci \bar{x}^0 a hledáme minimum funkce

$$\Psi(\bar{x}^0 - \lambda \text{grad}\Psi(\bar{x}^0)),$$

kteřá závisí pouze na jedné proměnné - parametru λ .

Prakticky hledáme minimální kladný kořen $\lambda = \lambda_0$ rovnice

$$\frac{d}{d\lambda} \Psi(\bar{x}^0 - \lambda \operatorname{grad} \Psi(\bar{x}^0)) = 0.$$

Položíme

$$\bar{x}^1 = \bar{x}^0 - \lambda_0 \operatorname{grad} \Psi(\bar{x}^0)$$

a obecně

$$\bar{x}^k = \bar{x}^{k-1} - \lambda_{k-1} \operatorname{grad} \Psi(\bar{x}^{k-1}), \quad (5.3)$$

$k = 1, 2, \dots$, kde λ_{k-1} je minimální kladný kořen rovnice

$$\frac{d}{d\lambda} \Psi(\bar{x}^{k-1} - \lambda \operatorname{grad} \Psi(\bar{x}^{k-1})) = 0.$$

Konvergenci postupných aproximací (5.3) k řešení x^* rovnice (5.2) ale nemáme obecně zajištěnu, protože můžeme dokonvergovat k bodu lokálního minima a ne globálního, které hledáme. Proto je nutné prověřit výsledek, zda se jedná o lokální nebo globální minimum.

Pojmy k zapamatování

- V aplikacích se často vyskytují soustavy rovnic a to lineárních i nelineárních.
- Zabývali jsme se hledáním řešení soustav nelineárních rovnic. Ukázali jsme si podmínky, které nám zaručují konvergenci prosté iterační metody a Newtonovy metody pro soustavy nelineárních rovnic.
- Zabývali jsme se rychlostí konvergence těchto metod a možnostmi odhadu chyby po k -tém kroku iterace.
- Všechny tyto podmínky a odhady vycházejí z Banachovy věty o pevném bodu.

Kontrolní otázky

1. Co rozumíme pojmem soustava nelineárních rovnic?
2. Odkud vyplývají podmínky řešitelnosti soustavy nelineárních rovnic?
3. Které soustavy nelineárních rovnic neumíme řešit?

5.5 Cvičení

Maplety

Příklady pro samostatnou práci jsou uvedeny v samostatné sbírce příkladů. Pomocí následujících mapletů si můžete usnadnit některé dílčí výpočty, nebo zkontrolovat jejich správnost.

1. Parciální derivace
2. Řešení soustav lineárních rovnic

6 Řešení obyčejných diferenciálních rovnic.

Průvodce studiem

Matematicky popisujeme většinu fyzikálních a technických dějů pomocí diferenciálních rovnic. Pokud je popisovaný děj funkcí jedné proměnné, jde při popisu o obyčejné diferenciální rovnice. Pouze malou část z nich jsme schopni řešit analyticky, tj. vyjádřit řešení v explicitním tvaru $y = f(x)$ nebo v implicitním tvaru $F(x, y) = 0$.

Základní pojmy o obyčejných diferenciálních rovnicích jste měli v předmětu Matematika 2. Podrobněji jste se mohli seznámit s teorií obyčejných diferenciálních rovnic ve volitelném předmětu Diferenciální rovnice v elektrotechnice.

Cílem této kapitoly je seznámit čtenáře s numerickými metodami řešení obyčejných diferenciálních rovnic. Numerické řešení hledáme tehdy, když nejsme schopni nalézt analytické řešení nebo pokud by jeho nalezení zabralo příliš mnoho času, případně bylo pro nás příliš obtížné.

Nejdříve si připomeneme základní pojmy z teorie diferenciálních rovnic. Zformulujeme si znovu Cauchyovu úlohu. Zopakujeme si analytické řešení vybraných typů obyčejných diferenciálních rovnic.

Poté přejdeme k numerickým metodám. Začneme u jednokrokových metod. Budeme se zabývat Eulerovou metodou a jejími modifikacemi a metodou Rungeho-Kutty. Zmíníme se i o stabilitě řešení. V této části půjde o opakování z předmětu Matematika 3.

Pak se zaměříme na vícekové metody. Probereme si Adamsovy metody, metody prediktor-korektor a prediktor-modifikátor-korektor.

Všechny metody budou založeny na diskretizaci proměnné. Nebudeme hledat spojitou funkci, která vyhovuje dané rovnici, ale diskrétní funkci, která bude definována pouze na konečném počtu bodů a bude se v limitě blížit k přesnému řešení.

Cíle

Po prostudování této kapitoly budete schopni:

- Zformulovat základní úlohu teorie diferenciálních rovnic.
- Přímým výpočtem řešit vybrané typy rovnic.
- Zvládat numerické řešení diferenciálních rovnic a to jak pomocí jednokrokových metod, tak

i pomocí více krokových metod.

- Umět rozhodnout o ukončení výpočtu po dosažení zadané přesnosti.
- Rozhodnout o stabilitě rovnice.

6.1 Cauchyova úloha.

Řád diferenciální rovnice je definován jako řád nejvyšší derivace neznámé funkce, která se v rovnici vyskytuje. My se nejdříve budeme věnovat rovnicím prvního řádu.

Na intervalu $[a, b]$ máme řešit diferenciální rovnici prvního řádu

$$y'(x) = f(x, y) \quad (6.1)$$

s počáteční podmínkou

$$y(x_0) = y_0, \quad x_0 \in [a, b]. \quad (6.2)$$

Nejčastěji se bere $x_0 = a$.

Definice 6.1. Úloha (6.1), (6.2) se nazývá *počáteční úloha* (nebo *Cauchyova úloha*).

Věta 6.2. Picardova věta - o existenci a jednoznačnosti řešení.

Mějme oblast $D = \{(x, y) : |x - x_0| \leq p, |y - y_0| \leq q\}$. Nechť

1. $f(x, y)$ je spojitá v D pro obě proměnné a tedy je v D ohraničená, t.j. $\forall (x, y) \in D : |f(x, y)| \leq M$, M je kladná konstanta.
2. $f(x, y)$ splňuje v D Lipschitzovu podmínku $|f(x, y) - f(x, u)| \leq L|y - u|$, kde L je kladná konstanta a $(x, y), (x, u)$ jsou libovolné body z D .

Potom má úloha (6.1), (6.2) právě jedno řešení, které je spojitě a spojitě diferencovatelné na intervalu $|x - x_0| < h$, $h = \min\{p, \frac{q}{M}\}$ a při těchto hodnotách je $|y - y_0| \leq q$, t.j. zůstává v D .

Poznámka 6.3. Podmínka 2 věty 6.2 bude splněna automaticky, jestliže budou v D spojitě a tedy i ohraničené parciální derivace funkce $f(x, y)$ podle y pro libovolný bod $(x, y) \in D$, t.j.

$$\left| \frac{\partial f(x, y)}{\partial y} \right| \leq K, \quad K > 0.$$

Velmi často se vyskytují úlohy, při nichž je funkce $f(x, y)$ spojitá a diferencovatelná, ale parciální derivace podle y jsou neohraničené. Potom je existence a hlavně jednoznačnost řešení garantována pouze v nějakém okolí bodu x_0 , které může být velmi malé.

Příklad 6.4. Mějme danu rovnici

$$y' = 2\sqrt{y}, \quad (y \geq 0).$$

Potom pro $y \neq 0$ máme

$$\frac{y'}{2\sqrt{y}} = 1 \quad \Rightarrow \quad (\sqrt{y})' = 1.$$

Integrací dostaneme

$$|\sqrt{y}| = x + C,$$

a po odstranění absolutní hodnoty máme

$$\sqrt{y} = x + C, \quad \text{kde } x + C > 0 \quad \text{a tedy } x > -C,$$

$$y = \begin{cases} (x + C)^2, & x > -C \\ 0 & \text{jinak} \end{cases}.$$

Jde tedy o pravé větve parabol.

Při řešení jsme vyloučili $y = 0$. Je třeba prověřit, zda tím neztrácíme jedno řešení. V našem případě je funkce $y = 0$ také řešením. Je to parciální řešení.

Pro parciální derivace platí

$$f(x, y) = 2\sqrt{y} \quad \Rightarrow \quad \frac{\partial f}{\partial y} = \frac{1}{\sqrt{y}},$$

parciální derivace není definovaná pro $y = 0$ a pro $y > 0$ je derivace neohrazenou funkcí v okolí $y = 0$. Nemůžeme proto garantovat jednoznačnost pro všechny body osy x . Naopak, každým bodem $(x, 0)$ prochází nekonečně mnoho řešení.

Příklad 6.5. Rovnice

$$x(y')^2 - 2yy' + 4x = 0,$$

je kvadratická rovnice vzhledem k y' . (Stále se jedná o diferenciální rovnici prvního řádu. Řád rovnice určuje nejvyšší řád derivace, ne její mocnina.) Po jejím vyřešení dostaneme

$$y' = \frac{y \pm \sqrt{y^2 - 4x^2}}{x}, \tag{6.3}$$

kde $y^2 - 4x^2 > 0$, $x^2 + y^2 \neq 0$.

Označme

$$y = zx,$$

kde $z = z(x)$ je nová neznámá funkce. Dosazením do (6.3) dostaneme

$$z'x + z = \frac{zx \pm \sqrt{z^2x^2 - 4x^2}}{x},$$

a po úpravě

$$z'x = \pm \sqrt{z^2 - 4}.$$

Pro $x \neq 0$, $\sqrt{z^2 - 4} \neq 0$ dostaneme

$$\frac{dz}{\pm \sqrt{z^2 - 4}} = \frac{dx}{x},$$

máme rovnici se separovanými proměnnými a její integrací dostaneme

$$\begin{aligned}z \pm \sqrt{z^2 - 4} &= Cx, \\ \pm \sqrt{z^2 - 4} &= Cx - z, \\ z^2 - 4 &= C^2x^2 - 2Cxz + z^2.\end{aligned}$$

Zpětným dosazením $z = \frac{y}{x}$ dostaneme

$$C^2x^2 - 2Cy + 4 = 0,$$

což je řešení pro

$$\begin{aligned}y^2 - 4x^2 &> 0, y > 0, \\ y^2 - 4x^2 &> 0, y < 0.\end{aligned}\tag{6.4}$$

V průběhu výpočtu jsme předpokládali, že $x \neq 0$, $z^2 - 4 \neq 0$. Je nutno prověřit, zda nejde o parciální řešení. V tomto případě ano. A pro $y = \pm 2x$ neplatí jednoznačnost.

Každým bodem v okolí bodu (x_0, y_0) , náležejícím do oblasti (6.4), procházejí právě dvě integrální křivky a to $y = 2x$ a $y = \frac{x^2}{x_0} + x_0$. Např. pro $y = \frac{1}{2}x^2 + 2$ a $y = x^2 + 1$ máme společný bod $(\sqrt{2}, 3)$.

6.2 Základní analytické metody.

6.2.1 Lineární rovnice

Lineární rovnice jsou rovnice ve tvaru

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

Příklad 6.6. Najděte řešení rovnice

$$y' + 2y = 4x.$$

Řešení. Rovnici si upravíme na kanonický tvar

$$y' = -2y + 4x.$$

Jedná se o lineární nehomogenní rovnici. Najdeme řešení přidružené homogení rovnice

$$y' = -2y.$$

$$\frac{dy}{y} = -2dx,$$

$$\ln |y| = -2x + \ln C,$$

$$y = C \cdot e^{-2x}.$$

Metodou variace konstanty najdeme řešení nehomogenní rovnice.

$$C \equiv K(x) \Rightarrow y = K(x)e^{-2x}.$$

Dosazením do původní rovnice dostaneme

$$K'(x)e^{-2x} + K(x)e^{-2x}(-2) = -2K(x)e^{-2x} + 4x,$$

$$K'(x)e^{-2x} = 4x,$$

$$K'(x) = 4xe^{2x},$$

$$K(x) = 4 \int xe^{2x} dx = \left| \begin{array}{l} u = x, \quad u' = 1 \\ v' = e^{2x}, \quad v = \frac{1}{2}e^{2x} \end{array} \right| =$$

$$4 \left[\frac{1}{2}xe^{2x} - \frac{1}{2} \int e^{2x} dx \right] = 2xe^{2x} - 2 \left[\frac{1}{2}e^{2x} \right] + L,$$

$$K(x) = 2xe^{2x} - e^{2x} + L.$$

Dosazením do řešení homogenní rovnice dostaneme výsledek

$$y(x) = K(x)e^{-2x},$$

$$y(x) = Le^{-2x} + 2x - 1.$$

□

Příklad 6.7. Najděte řešení rovnice

$$y' + 2xy = xe^{-x^2}.$$

Řešení. Rovnici si upravíme na kanonický tvar

$$y' = -2xy + xe^{-x^2}.$$

Jedná se o lineární nehomogenní rovnici. Najdeme řešení homogenní rovnice

$$y' = -2xy.$$

$$\frac{dy}{y} = -2xdx,$$

$$\ln |y| = -x^2 + \ln C,$$

$$y = C \cdot e^{-x^2}.$$

Metodou variace konstanty najdeme řešení nehomogenní rovnice.

$$C \equiv K(x) \Rightarrow y = K(x)e^{-x^2}.$$

Dosazením do původní rovnice dostaneme

$$K'(x)e^{-x^2} + K(x)e^{-x^2}(-2x) = -2xK(x)e^{-x^2} + xe^{-x^2},$$

$$K'(x) = x,$$

$$K(x) = \frac{x^2}{2} + L.$$

Dosazením do řešení homogenní rovnice dostaneme výsledek

$$y(x) = K(x)e^{-x^2},$$

$$y(x) = e^{-x^2} \left(\frac{x^2}{2} + L \right).$$

□

6.2.2 Bernoulliho rovnice

Řešení Bernoulliho rovnice můžeme hledat buď vhodnou substitucí nebo metodou variace konstanty. Jedná se o rovnici

$$y' = a(x)y + b(x)y^r, \quad r \in \mathbb{R}, r \neq 1.$$

Volbou $u = y^{1-r}$ převedeme Bernoulliho rovnici na rovnici lineární. Proved'te.

Ukážeme si jiný postup a to použití metody *variace konstanty*.

Příklad 6.8. Najděte řešení rovnice

$$y' + 2xy = 2x^3y^3.$$

Řešení. Rovnici si upravíme na kanonický tvar

$$y' = -2xy + 2x^3y^3,$$

což je Bernoulliho rovnice. Najdeme řešení homogenní rovnice

$$y' = -2xy.$$

$$\frac{dy}{y} = -2x dx,$$

$$\ln |y| = -x^2 + \ln C,$$

$$y = C \cdot e^{-x^2}.$$

Metodou variace konstanty najdeme řešení Bernoulliho rovnice.

$$C \equiv K(x) \Rightarrow y = K(x)e^{-x^2}.$$

Dosazením do původní rovnice dostaneme

$$K'(x)e^{-x^2} + K(x)e^{-x^2}(-2x) = -2xK(x)e^{-x^2} + 2x^3K^3(x)e^{-3x^2},$$

$$K'(x)e^{-x^2} = 2x^3K^3(x)e^{-3x^2},$$

$$K'(x) = 2x^3K^3(x)e^{-2x^2},$$

$$\frac{dK}{K^3(x)} = 2x^3e^{-2x^2} dx.$$

Budeme integrovat každou stranu rovnice samostatně:

$$\int \frac{dK}{K^3(x)} = \int K^{-3}(x) dK = \frac{K^{-2}(x)}{-2}.$$

$$\begin{aligned} \int 2x^3e^{-2x^2} dx &= \left| \begin{array}{l} x^2 = t, \\ 2x dx = dt \end{array} \right| = \int te^{-2t} dt = \left| \begin{array}{l} u = t, \quad u' = 1 \\ v' = e^{-2t}, \quad v = \frac{1}{-2}e^{-2t} \end{array} \right| = \\ &= -\frac{1}{2}te^{-2t} + \frac{1}{2} \int e^{-2t} dx = -\frac{1}{2}te^{-2t} - \frac{1}{4}e^{-2t} + L = -\frac{1}{2}x^2e^{-2x^2} - \frac{1}{4}e^{-2x^2} + L. \end{aligned}$$

Kde L je integrační konstanta. Dostali jsme

$$\frac{K^{-2}(x)}{-2} = -\frac{1}{2}x^2 e^{-2x^2} - \frac{1}{4}e^{-2x^2} + L,$$

$$\frac{1}{K^2(x)} = x^2 e^{-2x^2} + \frac{1}{2}e^{-2x^2} + M,$$

kde $M = -2L$ je konstanta. Dosazením do řešení homogenní rovnice dostaneme výsledek

$$y(x) = K(x)e^{-x^2},$$

$$y^2(x) = K^2(x)e^{-2x^2},$$

$$\frac{1}{y^2} = \frac{1}{K^2(x)e^{-2x^2}},$$

$$\frac{1}{y^2} = \frac{e^{2x^2}}{K^2(x)},$$

$$\frac{1}{y^2} = x^2 + \frac{1}{2} + Me^{2x^2}.$$

□

Příklad 6.9. Najděte řešení rovnice

$$xy' + y = y^2 \cdot \ln x.$$

Řešení. Rovnici si upravíme na kanonický tvar

$$y' = -\frac{1}{x}y + \frac{\ln x}{x}y^2.$$

Máme Bernoulliho rovnici. Najdeme řešení homogenní rovnice

$$y' = -\frac{1}{x}y.$$

$$\frac{dy}{y} = -\frac{1}{x}dx,$$

$$\ln |y| = -\ln x + \ln C,$$

$$y = \frac{C}{x}.$$

Metodou variace konstanty najdeme řešení Bernoulliho rovnice.

$$C \equiv K(x) \Rightarrow y = \frac{K(x)}{x}.$$

Dosazením do původní rovnice dostaneme

$$K'(x)\frac{1}{x} + K(x)\frac{-1}{x^2} = -\frac{1}{x}\frac{K(x)}{x} + \frac{\ln x}{x}\frac{K^2(x)}{x^2},$$

$$K'(x) = \frac{\ln x}{x^2}K^2(x),$$

$$\frac{dK}{K^2(x)} = \frac{\ln x}{x^2} dx.$$

Budeme integrovat každou stranu rovnice samostatně:

$$\begin{aligned} \int \frac{dK}{K^2(x)} &= \int K^{-2}(x) dK = \frac{K^{-1}(x)}{-1} = -\frac{1}{K(x)}. \\ \int \frac{\ln x}{x^2} dx &= \left| \begin{array}{l} u = \ln x, \quad u' = \frac{1}{x} \\ v' = x^{-2}, \quad v = -x^{-1} \end{array} \right| = \\ &= -\frac{\ln x}{x} + \int \frac{1}{x^2} dx = -\frac{\ln x}{x} - \frac{1}{x} + L \end{aligned}$$

Dostali jsme

$$-\frac{1}{K(x)} = -\frac{\ln x}{x} - \frac{1}{x} + L.$$

Dosazením do řešení homogenní rovnice dostaneme výsledek

$$\begin{aligned} \frac{1}{y(x)} &= \frac{x}{K(x)}, \\ \frac{1}{y(x)} &= x \left(\frac{\ln x}{x} + \frac{1}{x} + L \right), \\ y(x)(\ln x + 1 + Lx) &= 1. \end{aligned}$$

□

6.2.3 Exaktní rovnice

Rovnice $P(x, y) dx + Q(x, y) dy = 0$ se nazývá exaktní diferenciální rovnice, jestliže P a Q jsou funkce dvou proměnných, které jsou spojité a mají spojité parciální derivace prvního řádu v nějaké obdélníkové oblasti $R = \{(x, y) \in \mathbb{R}^2 : a < x < b, c < y < d\}$, kde a, b, c, d jsou konstanty, a platí $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$. Potom exaktní rovnice je totálním diferenciálem nějaké funkce F , kterou nazýváme kmenovou funkcí.

Příklad 6.10. Najděte řešení rovnice

$$(2x^3 - xy^2)dx + (2y^3 - x^2y)dy = 0.$$

Řešení. Máme

$$P(x, y) = 2x^3 - xy^2, \quad Q(x, y) = 2y^3 - x^2y.$$

Potom

$$P_y = -2xy, \quad Q_x = -2xy.$$

Protože $P_y = Q_x$, jde o exaktní rovnici. Budeme hledat její kmenovou funkci.

$$\begin{aligned} dF(x, y) &= P(x, y)dx + Q(x, y)dy, \\ F(x, y) &= \int P(x, y)dx = \int (2x^3 - xy^2)dx = \frac{1}{2}x^4 - \frac{1}{2}x^2y^2 + C(y). \end{aligned}$$

Potom

$$F_y = -x^2y + C'(y).$$

A současně $F_y = Q$.

$$\begin{aligned} -x^2y + C'(y) &= 2y^3 - x^2y, \\ C'(y) &= 2y^3, \\ C(y) &= \frac{1}{2}y^4. \end{aligned}$$

Kmenová funkce má tedy tvar

$$F(x, y) = \frac{1}{2}x^4 - \frac{1}{2}x^2y^2 + \frac{1}{2}y^4$$

a obecné řešení je tvaru

$$x^4 - x^2y^2 + y^4 = L,$$

kde L je konstanta. □

Příklad 6.11. Najděte řešení rovnice

$$\frac{xdy}{x^2 + y^2} - \left(\frac{y}{x^2 + y^2} - 1 \right) dx = 0$$

Řešení. Máme

$$P(x, y) = - \left(\frac{y}{x^2 + y^2} - 1 \right), \quad Q(x, y) = \frac{x}{x^2 + y^2}.$$

Potom

$$P_y = - \frac{x^2 + y^2 - 2y^2}{(x^2 + y^2)^2} = - \frac{x^2 - y^2}{(x^2 + y^2)^2}, \quad Q_x = \frac{x^2 + y^2 - 2x^2}{(x^2 + y^2)^2} = \frac{y^2 - x^2}{(x^2 + y^2)^2}.$$

Protože $P_y = Q_x$, jde o exaktní rovnici. Budeme hledat její kmenovou funkci.

$$dF(x, y) = P(x, y)dx + Q(x, y)dy,$$

$$\begin{aligned} F(x, y) &= \int P(x, y)dx = - \int \left(\frac{y}{x^2 + y^2} + 1 \right) dx = \\ &= - \int \left(\frac{y}{y^2 \left(\frac{x^2}{y^2} + 1 \right)} + 1 \right) dx = - \frac{1}{y} \int \frac{1}{\frac{x^2}{y^2} + 1} dx - \int dx = - \arctan x - x + C(y). \end{aligned}$$

Potom

$$\begin{aligned} F_y &= - \frac{1}{1 + \frac{x^2}{y^2}} \frac{-x}{y^2} + C'(y). \\ F_y &= \frac{x}{x^2 + y^2} + C'(y). \end{aligned}$$

A současně $F_y = Q$.

$$\frac{x}{x^2 + y^2} + C'(y) = \frac{x}{x^2 + y^2},$$

$$C'(y) = 0,$$

$$C(y) = K.$$

Kmenová funkce má tedy tvar

$$F(x, y) = x + \arctan \frac{x}{y}$$

a obecné řešení je tvaru

$$x + \arctan \frac{x}{y} = L,$$

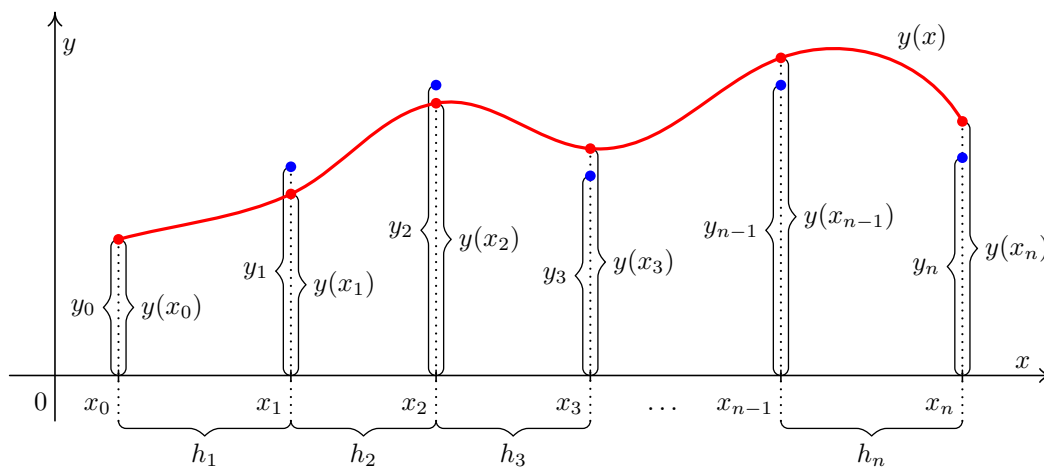
kde L je konstanta. □

6.3 Jednokrokové metody

Numerické metody, kterými se budeme zabývat, nebudou rozlišovat, zda řešení existuje či neexistuje, ani zda je jednoznačné či nikoliv. Proto je třeba předem provést prověrku podmínek věty 6.2.

My budeme vždy předpokládat, že jsou splněny podmínky věty 6.2 a že hledáme řešení diferenciální rovnice, které existuje a je jednoznačné.

Princip hledání numerického řešení počáteční úlohy $y' = f(x, y(x))$, $y(x_0) = y_0$ znáte z předmětu *Matematika 3*. Pro připomenutí, na obrázku 6.1 vidíme přesné řešení diferenciální rovnice, které je vykresleno plnou černou čarou a přibližné hodnoty řešení v uzlových bodech, vyznačené kroužky nad / pod přesným řešením.



Obr. 6.1: Princip numerického řešení počáteční úlohy $y' = f(x, y)$, $y(x_0) = y_0$; $y(x)$ je přesné řešení

6.3.1 Diferenční metody řešení Cauchyovy úlohy založené na diskretizaci proměnné.

Definice 6.12. Body $x_i : a = x_0 < x_1 < \dots < x_n = b$ nazveme *uzly sítě*, $h_i = x_{i+1} - x_i$ je *krok* sítě. Je-li $h_i = h = \text{Konst.}$, pak mluvíme o pravidelné síti, tj. síti s *ekvidistantními uzly*.

Definice 6.13. *Numerickým řešením* úlohy (6.1), (6.2) rozumíme určení hodnot y_0, y_1, \dots, y_n , které aproximují hodnoty $y(x)$ v uzlových bodech, t.j. $y_i \approx y(x_i), i = 1, 2, \dots, n$.

Takto získanou diskrétní funkci pak můžeme převést na spojitou funkci některou ze známých metod, jako je třeba interpolační polynom, splajn, metoda nejmenších čtverců atd.

Nechť $y_{i+1} = y_i + \Delta y_i$.

Definice 6.14. Jestliže $\Delta y_i = \varphi(x_i, y_i, h_i)$, pak jde o jedнокrokovou metodu.

Jestliže $\Delta y_i = \varphi(x_i, y_i, h_i, x_{i-1}, y_{i-1}, h_{i-1}, \dots, x_{i-k}, y_{i-k}, h_{i-k})$, $k > 1$ pak jde o vícekrokovou metodu.

Konvergenci metody chápeme takto:

$$\lim_{h_i \rightarrow 0} y_i = y(x_i), \text{ pro } i = 0, 1, \dots, n - 1,$$

pak říkáme, že numerické řešení konverguje k teoretickému.

Určit chybu metody a řád konvergence vyžaduje dodatečný matematický aparát, a proto tuto problematiku vynecháme. Kontrola se provádí nejčastěji metodou polovičního kroku tak, že srovnáváme hodnoty

$$y_h(x_i) \text{ a } y_{\frac{h}{2}}(x_{2i}),$$

kde $y_h(x_i)$ je hodnota funkce y v bodě $x_i = x_0 + ih$ pro krok $h, i = 1, \dots, n$ a $y_{\frac{h}{2}}(x_{2i})$ je hodnota funkce y ve stejném bodě $x_i = x_0 + 2i \frac{h}{2}$ pro krok $\frac{h}{2}, i = 1, \dots, n$.

6.3.2 Eulerova metoda

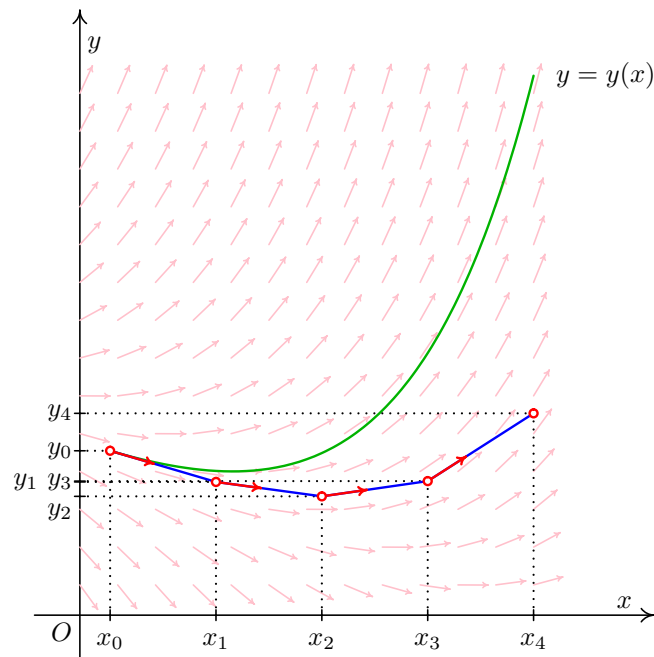
Eulerova ¹ metoda pro slouží k řešení úlohy $y' = f(x, y), y(a) = y_0$ na intervalu $[a, b]$. Označíme si $y_i = y(x_i), h_i = x_{i+1} - x_i, h = \frac{b-a}{n}$. Potom z Taylorova rozvoje, po

¹ **L. Euler** (1707 – 1783) švýcarské matematik, fyzik, astronom, mechanik. Jeden z nejvýznamnějších matematiků všech dob. V letech 1741 – 1766 působil v Berlíně, v letech 1727 – 1741 a 1766 – 1783 působil v Petrohradu v Rusku. Zabýval se prakticky všemi přírodními vědami, ve kterých bylo v jeho době možné aplikovat matematiku. Je autorem přes 860 prací. Přes polovinu jich diktoval už jako slepý. Věnoval se výrazně analýze, vybudoval trigonometrii prakticky do dnešní podoby, vytvořil analytickou teorii čísel, zavedl komplexní funkce, stojí u zrodu teorie grafů, ... Byl vynikajícím počtářem a zpracoval řadu numerických metod.

zanedbání vyšších řádů, dostaneme

$$y_{i+1} = y_i + h_i f(x_i, y_i), \quad i = 0, 1, \dots, n-1. \quad (6.5)$$

Pro $h_i \rightarrow 0$ posloupnost $\{y_i\}$ konverguje k přesnému řešení. Geometrický smysl této metody - viz obrázek 6.2.



Obr. 6.2: Přibližné řešení diferenciální rovnice Eulerovou metodou

Příklad 6.15. Eulerovou metodou s krokem $h = 0.1$ řešte na intervalu $[1, 1.5]$ úlohu

$$y' = y + (1 + x)y^2, \quad y(1) = -1.$$

Řešení. Je uvedeno v tabulce:

i	x_i	y_i	$f(x_i, y_i)$
0	1	-1	1
1	1.1	-0.9	0.801
2	1.2	-0.8199	0.659019
3	1.3	-0.753998	0.553582
4	1.4	-0.698640	0.47294
5	1.5	-0.651361	

□

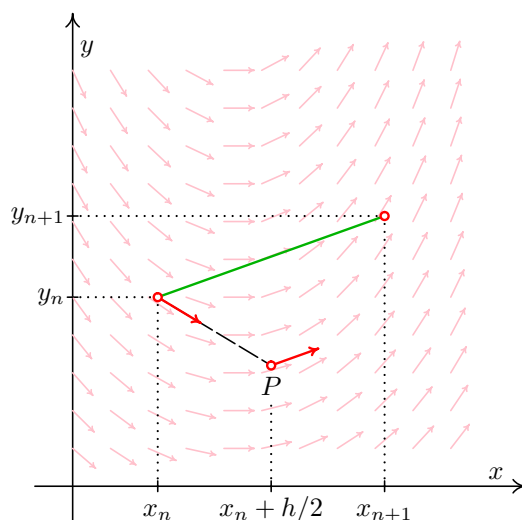
Eulerova metoda je nejjednodušší numerickou metodou pro řešení diferenciálních rovnic. I když se již v praxi příliš nepoužívá, hlavně pro její malou přesnost, její význam proto není menší. Je především pro svou jednoduchost východiskem mnoha teoretických úvah a dále jejím zobecněním můžeme dostat nové přesnější metody.

První modifikace Eulerovy metody

Místo vzorce (6.5) použijeme

$$y_{i+1} = y_i + hf\left(x_i + \frac{h}{2}, y_i + \frac{h}{2}f(x_i, y_i)\right)$$

Geometrická interpretace je zřejmá z obrázku 6.3



Obr. 6.3: První modifikace Eulerovy metody

Druhá modifikace Eulerovy metody

Místo vzorce (6.5) použijeme

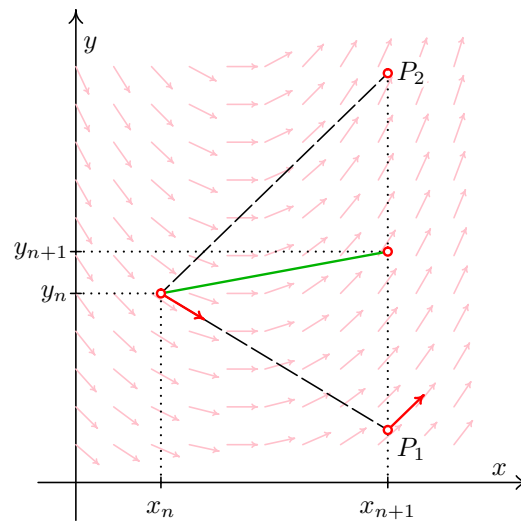
$$y_{i+1} = y_i + \frac{1}{2}h(f(x_i, y_i) + f(x_i + h, y_i + hf(x_i, y_i)))$$

Geometrická interpretace plyne z obrázku 6.4.

Existují i další varianty metody, které si však už nebudeme odvozovat, jen si u každé z nich uvedeme vzorec pro výpočet:

1. Eulerova metoda vpřed. Je to metoda se kterou jsme začínali.

$$y_{n+1} = y_n + hf_n.$$



Obr. 6.4: Druhá modifikace Eulerovy metody

2. Eulerova metoda vzad

$$y_{n+1} = y_n + h f_{n+1}.$$

3. Crank-Nicolsonova metoda

$$y_{n+1} = y_n + \frac{h}{2} [f_n + f_{n+1}].$$

4. Heumova metoda

$$y_{n+1} = y_n + \frac{h}{2} [f_n + f(x_{n+1}, y_n + h f_n)].$$

6.3.3 Metody Rungeho – Kuttovy

Metody Rungeho – Kuttovy ¹ jsou dány rekurentními vztahy

$$y_{j+1} = y_j + h_j \sum_{i=1}^r \alpha_i k_i, \quad j = 0, 1, \dots, n-1.$$

kde $k_1 = f(x_j, y_j)$, $k_i = f(x_j + \lambda_i h_j, y_j + \mu_i h_j k_{i-1})$, $i > 1$. Na každém kroku se vypočítávají hodnoty k_i a pomocí nich se určuje hodnota přírůstku. Poté se výpočet opakuje. Tyto vztahy se odvozují z Taylorova rozvoje pro přesný relativní přírůstek a pro jeho přibližnou hodnotu. To vede na soustavu rovnic, která má nekonečně mnoho řešení. Vhodnou volbou parametrů pak získáme jednotlivé typy metody. Obě modifikace Eulerovy metody je možné získat i tímto postupem. Odvození viz např. [17], str.107, [3], str. 353 — 358.

¹ **K. D. T. Runge** (1856 – 1927) německý matematik a fyzik. První profesor aplikované matematiky v Německu.

Nejčastěji se používá Rungeho – Kuttova metoda 4.řádu. Její tvar pro ekvidistantní uzly je:

$$y_{j+1} = y_j + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4),$$

$$k_1 = hf(x_j, y_j),$$

$$k_2 = hf\left(x_j + \frac{h}{2}, y_j + \frac{k_1}{2}\right),$$

$$k_3 = hf\left(x_j + \frac{h}{2}, y_j + \frac{k_2}{2}\right),$$

$$k_4 = hf(x_j + h, y_j + k_3).$$

Příklad 6.16. Rungeho – Kuttovou metodou řešte

$$y' = y - \frac{2x}{y}, \quad y(0) = 1, \quad h = 0.2, \quad a = 0, \quad b = 1.$$

Řešení.

Analytické řešení je

$$y = \sqrt{2x + 1}.$$

Numerické řešení je uvedeno v tabulce.

j	x	y	k_i	přírůstek
0	0	1	$k_1 = 0.2$	
	0.1	1.1	$k_2 = 0.1838$	
	0.1	1.0918	$k_3 = 0.1817$	
	0.2	1.1817	$k_4 = 0.1686$	
				$\Delta y = 0.1832$
1	0.2	1.1832	$k_1 = 0.1690$	
	0.3	1.2677	$k_2 = 0.1589$	
	0.3	1.2626	$k_3 = 0.1575$	
	0.4	1.3407	$k_4 = 0.1488$	
				$\Delta y = 0.1584$
2	0.4	1.3416	...	

Sami si dopočítejte a srovnajte numerické řešení s analytickým. □

Příklad 6.17. Cauchyovu úlohu

$$y' = x^2 + \frac{1}{10}y^2, \quad y(0) = 0, \quad a = 0, \quad b = 1,$$

řešte postupně Eulerovou metodou, první a druhou modifikací Eulerovy metody a metodou Rungeho-Kuttovou, výsledky srovnajte.

Přesné řešení je $y(1) = 0.334930314854697$.

Řešení. Numerické řešení je uvedeno v tabulce. V prvním sloupci je uvedeno na kolik dílů dělíme interval $[0; 1]$. Jednotlivými metodami jsme se už zabývali, proto je v dalších sloupcích vždy uvedena až konečná hodnota $y_n(1)$, získaná danou metodou.

n	EM	1.mEM	2.mEM	R-K
5	0.240284311	0.331333136	0.341556796	0.334932164
10	0.285738187	0.334026264	0.336596144	0.334930446
20	0.309855335	0.334703844	0.335348142	0.334930324
50	0.324783702	0.334894044	0.334997310	0.334930315
100	0.329837423	0.334921244	0.334947076	0.334930315
200	0.332378958	0.334928047	0.334934507	0.334930315
500	0.333908592	0.334929952	0.334930986	0.334930315

Povšimněte si rychlosti konvergence u jednotlivých metod. □

Další často užívanou numerickou metodou Rungeho-Kuttova typu je tzv. *tříosminové pravidlo*. Jeho tvar pro ekvidistantní uzly je:

$$\begin{aligned}
 y_{n+1} &= y_n + \frac{h}{8}(k_1 + 3k_2 + 3k_3 + k_4), \\
 k_1 &= f(x_n, y_n), \\
 k_2 &= f\left(x_n + \frac{h}{3}, y_n + \frac{h}{3}k_1\right), \\
 k_3 &= f\left(x_n + \frac{2h}{3}, y_n - \frac{h}{3}k_1 + hk_2\right), \\
 k_4 &= f(x_n + h, y_n + h(k_1 - k_2 + k_3)).
 \end{aligned}$$

Jinou variantou metod Rungeho-Kuttova typu je *Ralstonova metoda*

$$\begin{aligned}
 y_{i+1} &= y_i + \frac{1}{9}h(2k_1 + 3k_2 + 4k_3), \\
 k_1 &= f(x_i, y_i), \\
 k_2 &= f\left(x_i + \frac{1}{2}h, y_i + \frac{1}{2}hk_1\right), \\
 k_3 &= f\left(x_i + \frac{3}{4}h, y_i - \frac{3}{4}hk_2\right).
 \end{aligned}$$

6.4 Stabilita

Metody typu Rungeho-Kuttovy jsou relativně přesné a mají vyhovující rychlost konvergence. Ale i zde se mohou vyskytnout problémy:

Příklad 6.18. Na intervalu $[a, b]$ hledáme řešení úlohy

$$y' = 5y - x^2 + 0.4x, \quad y(a = 0) = 0.$$

Řešení. Analytickým řešením této úlohy je funkce $y(x) = 0.2x^2$. Při numerickém řešení Rungeho-Kuttovou metodou dostaneme postupně pro krok $h = 0.2$ a $h = 0.1$:

x	$y(x)$
0.0	0.0
0.2	0.01
0.4	0.03
0.6	0.07
0.8	0.12
1.0	0.19
1.2	0.25
1.4	0.29
1.6	0.23
1.8	-0.12
2.0	-1.27
2.2	-4.64
2.4	-14.04

x	$y(x)$
0.0	0
0.1	0
0.2	0.01
0.3	0.02
0.4	0.03
0.5	0.05
0.6	0.07
0.7	0.10
0.8	0.13
0.9	0.16
1.0	0.2
1.1	0.24
1.2	0.28
1.3	0.33
1.4	0.38
1.5	0.44
1.6	0.49
1.7	0.54
1.8	0.58
1.9	0.62
2.0	0.62
2.1	0.59
2.2	0.49
2.3	0.27
2.4	-0.15
2.5	-0.90

Vidíme, že velmi brzy budeme v obou případech dostávat záporné hodnoty y . Nejde tedy o chybu metody. Důvodem je nestabilita této rovnice.

Pokud budeme mít nastavenou automatickou volbu kroku, potom až budeme pracovat s velmi malým krokem, tak se teprve můžeme vyhnout havarování výpočtu. \square

Mějme lineární diferenciální rovnici

$$y' = p(x)y + q(x),$$

jejím řešením je

$$y(x) = e^{\int_a^x p(t)dt} \left(\int_a^x q(t)e^{-\int_a^t p(\tau)d\tau} dt + K \right).$$

(Všimněte si, že v příkladu 6.18 se v řešení nevyskytují exponenciály.)

Předpokládejme, že na naši rovnici působí nějaká drobná porucha, popsaná funkcí $\delta(x)$, $|\delta(x)| \ll |f(x, y)| \equiv |p(x)y + q(x)|$. Potom

$$y' = f(x, y) + \delta(x) = p(x)y + q(x) + \delta(x),$$

a řešení dostaneme ve tvaru

$$\tilde{y} = \tilde{K} e^{\int_a^x p(t) dt} + \int_a^x (q(t) + \delta(t)) e^{\left(\int_t^x p(\tau) d\tau\right)} dt.$$

Je zřejmé, že vliv poruch bude minimální, jestliže

$$\frac{\partial f(x, y)}{\partial y} = p(x) < 0.$$

Takové lineární rovnice se pak nazývají *stabilními*. V případě rovnic vyšších řádů nebo rovnic nelineárních je také možné odvodit podmínky stability. Ty však jsou podstatně složitější.

VYHODNOCENÍ:

1. Vždy prověřovat podmínky konvergence řešení.
2. Vždy prověřovat podmínky existence a jednoznačnosti řešení.
3. Není vhodné hledat řešení na příliš dlouhém intervalu.

6.5 Vícekrokové metody.

Opět hledáme diskrétní funkci, která nám bude aproximovat přesné řešení diferenciální rovnice $y' = f(x, y)$. Nechť $y_{i+1} = y_i + \Delta y_i$. Jestliže

$$\Delta y_i = \varphi(x_i, y_i, h_i, x_{i-1}, y_{i-1}, h_{i-1}, \dots, x_{i-k}, y_{i-k}, h_{i-k}), \quad k > 1$$

pak jde o vícekrokovou metodu.

Adamsovy metody

Mějme dánu úlohu (6.1), (6.2). Integrací (6.1) dostaneme

$$y(x) - y(x_0) = \int_{x_0}^x f(t, y(t)) dt.$$

Integrovanou funkci $f(t, y(t))$ nahradíme interpolačním polynomem s uzly x_0, x_1, \dots, x_s . Přitom předpokládáme, že známe v těchto bodech hodnoty funkce y . Potom můžeme integrál numericky určit a dostáváme

$$y(x) - y(x_0) = \sum_{k=0}^s c_k f(x_k, y(x_k)) + R,$$

kde R je chyba příslušného kvadraturního vzorce. Předpokládejme, že máme ekvidistantní uzly $x_i = x_0 + ih, i = 0, 1, \dots, s$. Nechť dále integrační meze splynou s některými uzly, t.j. $x_0 \equiv x_p, x \equiv x_q, p < q$. Použitím Lagrangeova interpolačního polynomu dostaneme

$$y(x_q) - y(x_p) = h \sum_{k=0}^s A_k f(x_k, y(x_k)) + R, \quad (6.6)$$

$$A_k = \frac{(-1)^{s-k}}{k!(s-k)!} \int_p^q \prod_{i=0, i \neq k}^s (t-i) dt,$$

$$R = h \int_p^q F[t, 0, 1, \dots, s] \prod_{i=0}^s (t-i) dt.$$

Zde vystupuje poměrná diference funkce F , pro kterou platí

$$F(t) = y'(x_0 + ht).$$

Vztah (6.6) můžeme vzít za základ pro numerickou metodu.

$$y_q = y_p + h \sum_{k=0}^s A_k f(x_k, y_k) + R.$$

Volbou $p = s$, $q = s + 1$ získáme *Adamsovy extrapolační metody*. Říká se jim tak proto, že integrační interval se nachází vně intervalu, na němž je určen interpolační polynom.

Máme tedy vztah

$$y_{s+1} = y_s + h \sum_{k=0}^s A_k^s f_k + R.$$

Horní index u A_k^s přitom zdůrazňuje závislost na s , tj. na počtu uzlů užitých při interpolaci. Zdůrazňujeme na POČTU, nikoliv na samotných uzlech. Koeficienty A_k^s nezávisí na uzlech, ani na jejich hodnotách. Proto je možné vzorec “posunovat” po síti o celočíselné násobky h , takže jej můžeme psát ve tvaru

$$y_{n+1} = y_n + h \sum_{k=0}^s A_k^s f_{n-s+k} + R.$$

Budeme-li psát k místo $s - k$, pak přerovnááním sčítanců do pořadí od nejbližšího uzlu k nejbvzdálenějšímu dostaneme

$$y_{n+1} = y_n + h \sum_{k=0}^s B_k^s f_{n-k} + R,$$

$$B_k^s = \frac{(-1)^k}{k!(s-k)!} \int_s^{s+1} \prod_{i=0, i \neq s-k}^s (t-i) dt, k = 0, 1, \dots, s.$$

Tyto vztahy nám definují *Adamsovu extrapolační metodu řádu $s + 1$* . (Někdy se označuje jako *Adams – Bashforthova* metoda). Chybu této metody můžeme popsat

$$R(x) = h^{s+2} \frac{y^{(s+2)}(\xi)}{(s+1)!} \int_s^{s+1} \prod_{i=0}^s (t-i) dt,$$

$\xi \in [x_0, x_{s+1}]$ a nebo po posunutí je $\xi \in [x_{n-s}, x_{n+1}]$.

Pro volbu $s = 1, 2, \dots$, $f_n = f(x_n, y_n)$ dostaneme:

Adamsovy extrapolační vzorce

$$\begin{aligned}
y_{n+1} &= y_n + hf_n + O\left(\frac{1}{2}y''(\xi)h^2\right) \\
y_{n+1} &= y_n + \frac{1}{2}h(3f_n - f_{n-1}) + O\left(\frac{5}{12}y'''(\xi)h^3\right) \\
y_{n+1} &= y_n + \frac{1}{12}h(23f_n - 16f_{n-1} + 5f_{n-2}) + O\left(\frac{3}{8}y^{(4)}(\xi)h^4\right) \\
y_{n+1} &= y_n + \frac{1}{24}h(55f_n - 59f_{n-1} + 37f_{n-2} - 9f_{n-3}) + O\left(\frac{251}{720}y^{(5)}(\xi)h^5\right) \\
y_{n+1} &= y_n + \frac{1}{720}h(1901f_n - 2774f_{n-1} + 2616f_{n-2} - 1274f_{n-3} + 251f_{n-4}) + \\
&\quad + O\left(\frac{95}{288}y^{(6)}(\xi)h^6\right) \\
y_{n+1} &= y_n + \frac{1}{1440}h(4277f_n - 7923f_{n-1} + 9982f_{n-2} - 7298f_{n-3} + 2877f_{n-4} - \\
&\quad - 475f_{n-5}) + O(y^{(7)}h^7) \\
&\dots \quad \dots
\end{aligned}$$

Zde symbolem $O(a)$ rozumíme, že chyba je řádově rovna a .

Řešíme-li rovnici Adamsovou metodou řádu $s + 1$, musíme znát hodnoty funkce v $(s + 1)$ předchozích bodech. Protože však bývá zadána pouze jedna počáteční podmínka $y(x_0) = y_0$, musíme si dopočítat zbývající hodnoty y_1, \dots, y_s některou jednokrokovou metodou. Nejčastěji se používá Runge-Kuttova metoda čtvrtého řádu.

Příklad 6.19. Adamsovou extrapolací metodou řešte na intervalu $[0; 1]$ úlohu

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

Řešení. Použijeme Adamsovu metodu čtvrtého řádu popsanou rovnicí

$$y_{n+1} = y_n + \frac{1}{24}h(55f_n - 59f_{n-1} + 37f_{n-2} - 9f_{n-3}).$$

Vedle počáteční podmínky, která je součástí zadání úlohy, potřebuje znát ještě hodnoty funkce $y(x)$ v dalších třech bodech. Výsledky jsou uvedeny v tabulce:

x	$y(x)$	
0.0	0.0	Počáteční podmínka
0.1	0.0003333	Vypočteno metodou Rungeho-Kutty
0.2	0.0026667	Vypočteno metodou Rungeho-Kutty
0.3	0.0090003	Vypočteno metodou Rungeho-Kutty
0.4	0.0213354	
0.5	0.0416772	
0.6	0.0720404	
0.7	0.1144567	
0.8	0.1709877	
0.9	0.2437425	
1.0	0.3349011	

Kontrola se provádí metodou polovičního kroku. Zde je nutné znovu vypočítat potřebné první hodnoty pomocí Rungeho-Kuttovy metody. Doporučuji počítat všechny potřebné hodnoty znovu, protože obecně

$$y_{h,i} \neq y_{\frac{h}{2},2i}.$$

□

Jestliže v rovnici (6.6) zvolíme $p = s - 1$, $q = s$, pak integrační interval přiléhá k pravému okraji interpolačního intervalu. Poté zcela stejným způsobem si odvodíme Adamsovu interpolační metodu (někdy se označuje jako *Adamsova-Moultonova* metoda).

$$y_{n+1} = y_n + h \sum_{k=-1}^s C_k^s f_{n-k},$$

$$C_k^s = \frac{(-1)^{k+1}}{(k+1)!(s-k)!} \int_s^{s+1} \prod_{i=0, i \neq s-k}^{s+1} (t-i) dt, \quad k = -1, 0, 1, \dots, s.$$

Chyba této metody je dána výrazem

$$R = h^{s+3} \frac{y^{(s+3)}(\xi)}{(s+2)!} \int_s^{s+1} \prod_{i=0}^{s+1} (t-i) dt,$$

kde $\xi \in [x_{n-s}, x_{n+1}]$.

Adamsovy interpolační vzorce

$$\begin{aligned} y_{n+1} &= y_n + hf_{n+1} + O\left(-\frac{1}{2}y''(\xi)h^2\right) \\ y_{n+1} &= y_n + \frac{1}{2}h(f_{n+1} + f_n) + O\left(-\frac{1}{12}y'''(\xi)h^3\right) \\ y_{n+1} &= y_n + \frac{1}{12}h(5f_{n+1} - 8f_n - f_{n-1}) + O\left(-\frac{1}{24}y^{(4)}(\xi)h^4\right) \\ y_{n+1} &= y_n + \frac{1}{24}h(9f_{n+1} + 19f_n - 5f_{n-1} + f_{n-2}) + O\left(-\frac{19}{720}y^{(5)}(\xi)h^5\right) \\ y_{n+1} &= y_n + \frac{1}{720}h(251f_{n+1} + 64f_n - 264f_{n-1} + 106f_{n-2} - 19f_{n-3}) + O(y^{(6)}(\xi)h^6) \\ y_{n+1} &= y_n + \frac{1}{1440}h(475f_{n+1} + 1427f_n - 798f_{n-1} + 486f_{n-2} - 173f_{n-3} + 27f_{n-4}) + \\ &\quad + O(y^{(7)}(\xi)h^7) \\ \dots &\quad \dots \end{aligned}$$

Také při použití Adamsových interpolačních metod musíme k počáteční podmínce $y(x_0) = y_0$ dopočítat dalších s hodnot jinou vhodnou metodou. Další problém je ten, že hodnota y_{n+1} se vyskytuje na obou stranách každé rovnice, neboť $f_{n+1} = f(x_{n+1}, y_{n+1})$.

Možnost explicitního vyjádření y_{n+1} závisí na konkrétním tvaru funkce $f(x, y)$. Nejčastěji se používá pro výpočet iterační metoda. Zvolíme si vhodnou nultou iteraci y_{n+1}^0 a sestavíme si posloupnost $\{y_{n+1}^i\}$:

$$y_{n+1}^i = y_n + hC_{-1}^s f(x_{n+1}, y_{n+1}^{i-1}) + h \sum_{k=0}^s C_k^s f_{n-k}.$$

Hledaná hodnota y_{n+1} je pak limitou této posloupnosti, pokud jsme si zvolili vhodnou první iteraci. Je možné stanovit podmínky konvergence.

6.5.1 Metoda prediktor – korektor

Adamsovy metody se už většinou nepoužívají samostatně. Jejich kombinace, zvaná *metody typu prediktor – korektor* využívají na každém kroku obou Adamsových metod:

1. **Prediktor** – explicitní metodou určíme y_{n+1}^k , $k = 0$.
2. Vypočteme si hodnotu $f_{n+1}^k = f(x_{n+1}, y_{n+1}^k)$.
3. **Korektor** – implicitní metodou vypočítáme lepší aproximaci y_{n+1}^{k+1} .
4. Zpřesníme výpočet pravé strany $f_{n+1}^{k+1} = f(x_{n+1}, y_{n+1}^{k+1})$.

Přitom body 3, 4 je možno opakovat – jde zde zase o iterační proces a prediktor nám pro něj dává dobré přiblížení. Obě metody bereme stejného řádu.

$$\text{Prediktor} \quad y_{n+1}^0 = y_n + \frac{h}{24}(55f_n - 59f_{n-1} + 37f_{n-2} - 9f_{n-3}),$$

$$\text{Korektor} \quad y_{n+1}^{k+1} = y_n + \frac{h}{24}(9f_{n+1}^k + 19f_n - 5f_{n-1} + f_{n-2}).$$

Další často používaná metoda prediktor – korektor má tvar

$$y_{n+1}^0 = y_{n-3} + \frac{4}{3}h(2f_n - f_{n-1} + 2f_{n-2}),$$

$$y_{n+1}^{k+1} = \frac{9}{8}y_n - \frac{1}{8}y_{n-2} + \frac{3}{8}h(f(x_{n+1}, y_{n+1}^k) + 2f_n - f_{n-1}).$$

Zde je prediktorem otevřený Simpsonův vzorec pro integraci a korektorem stabilní diferenční vzorec.

6.5.2 Metoda prediktor – modifikátor – korektor

Mějme metodu prediktor – korektor. Prediktorem určíme y_{n+1}^{0p} . Tuto hodnotu si zpřesníme modifikátorem

$$y_{n+1}^0 = y_{n+1}^{0p} + \frac{251}{270}(y_n - y_n^{0p}),$$

kde v závorce je rozdíl hodnoty funkce y v předchozím bodě v konečném tvaru a její predikce. Při prvním kroku metody prediktor – korektor nemáme žádnou hodnotu y_n^{0p} a proto se modifikátor počítá až od druhého kroku. Modifikovanou hodnotu pak použijeme v korektoru. Další postup je shodný s metodou prediktor – korektor.

Existují i další varianty metod tohoto typu. Vždy závisí na tom, jakou integrační metodu či tvar interpolačního polynomu použijeme.

Jednou z nejstarších je metoda *Milnova*, kde je prediktorem otevřený Simpsonův vzorec a korektorem uzavřený Simpsonův vzorec:

$$y_{n+1}^0 = y_{n-3} + \frac{4}{3}h(2f_n - f_{n-1} + 2f_{n-2}) + \left(\frac{14}{45}y^{(5)}(\xi)h^5\right),$$

$$y_{n+1}^{k+1} = y_{n-1} + \frac{h}{3}(f(x_{n+1}, y_{n+1}^k) + 4f_n + f_{n-1}) - \frac{1}{90}(y^{(5)}(\xi)h^5), k = 0, 1, \dots$$

Pro použití je třeba znát hodnoty ve čtyřech předcházejících uzlech. Odhad chyby je na každém kroku určen

$$R \leq \frac{1}{29}|y_i^p - y_i^k|,$$

kde je y_i^p prediktor a y_i^k jeho korekce. Často je nutno v průběhu výpočtu měnit krok h tak, aby chyba zůstala přiměřená – t.j. pokud je R velké, tak zmenšit krok, je-li R podstatně malé, tak je možno krok zvětšit.

Mnohdy bývá jednodušší provést celý výpočet na intervalu $[a, b]$ s krokem h a potom zopakovat výpočet s krokem $h/2$. Jestliže je i nyní chyba velká, provedeme výpočet s krokem $h/4$, atd. V případě zmenšení kroku je opět nutno dopočítat potřebné hodnoty funkce y vhodnou jednokrokovou metodou.

Modifikátor má u této metody tvar

$$y_{n+1}^0 = y_{n+1}^{0p} + \frac{28}{29}(y_n - y_n^{0p}).$$

Věta 6.20. *Nechť jsou splněny podmínky věty 6.2. Potom pro úlohu (6.1), (6.2) má metoda prediktor – korektor, kde je prediktor nejméně stejného řádu jako korektor, celkovou chybu řádově stejnou jako metoda, která používá korektor s nekonečným opakováním.*

Podmínky této věty jsou splněny pro všechny výše uvedené konkrétní typy metod prediktor – korektor a prediktor – modifikátor – korektor. Je tedy vhodnější je používat přímo v uvedeném tvaru, bez iterace, a vyšší přesnost zajistit zmenšením kroku.

6.6 Metody založené na užití derivací vyšších řádů

Jestliže použijeme ve vzorcích pro řešení diferenciální rovnice prvního řádu derivace vyšších řádů, můžeme dosáhnout snížení chyby. Takový vzorec můžeme odvodit z Eulerova-McLaurinova sumačního vzorce, kdy místo $f(x)$ dosadíme derivaci $y'(x)$ a po úpravě dostaneme

$$y_{n+1} = y_1 + h \sum_{i=-1}^{n-1} y'_{n-i} - \frac{h}{2}(y'_{n+1} + y'_1) - \sum_{k=1}^m \frac{B_{2k}}{(2k)!} h^{2k-1} (y_{n+1}^{(2k)} - y_1^{(2k)}), \quad (6.7)$$

kde máme lokální chybu

$$R_m = \frac{nh^{2m+2} B_{2m+2}}{(2m+2)!} y^{(2m+3)}(\xi),$$

kde B_{2k} jsou Bernoulliova čísla. Platí

$$\frac{t}{e^t - 1} = \sum_{k=0}^{\infty} B_k \frac{t^k}{k!},$$

$$\forall k > 0 : B_{2k+1} = 0,$$

$$B_0 = 1, \quad B_1 = -\frac{1}{2}, \quad B_2 = \frac{1}{6}, \quad B_4 = -\frac{1}{30}, \dots$$

Budeme se nyní věnovat vzorcům typu prediktor - korektor, které používají první i druhou derivaci. Vzorce pro korektor dostaneme tak, že Hermitův interpolační vzorec $h_j(x)$ pro body $x_{n+1}, x_n, \dots, x_{n-p}$ integrujeme od x_n do x_{n+1} . Jestliže \bar{f} označuje funkci komplexně sdruženou s funkcí f , dostaneme

$$y_{n+1} = y_n + \sum_{i=-1}^p \left[\int_{x_n}^{x_{n+1}} h_{n-i}(x) dx \right] y'_{n-i} + \sum_{i=-1}^p \left[\int_{x_n}^{x_{n+1}} \bar{h}_{n-i}(x) dx \right] y''_{n-i}.$$

Lokální chyba je dána výrazem

$$R_n = \frac{y^{(2p+5)}}{(2p+4)!} \int_{x_n}^{x_{n+1}} [(x - x_{n+1}) \dots (x - x_{n-p})] dx.$$

Pro $p = 0$ dostaneme korektor čtvrtého řádu ve tvaru

$$y_{n+1} = y_n + \frac{1}{2}h (y'_{n+1} + y'_n) + \frac{1}{12}h^2 (-y''_{n+1} + y''_n),$$

přítom

$$y'' = \frac{\partial f}{\partial y} \cdot y' + \frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} f + \frac{\partial f}{\partial x}.$$

Pro získání prediktoru použijeme Hermitův vzorec s uzly x_n, \dots, x_{n-p} a analogickými úpravami dostaneme výsledek. Takže celkově máme:

Prediktor

$$y_{n+1} = y_n + \frac{1}{2} (-y'_n + 3y'_{n-1}) + \frac{1}{12}h^2 (17y''_n + 7y''_{n-1}),$$

Modifikátor

$$\begin{aligned}\tilde{y}_{n+1} &= y_{n+1}^0 + \frac{31}{30}(y_n - y_n^0), \\ (\tilde{y}_{n+1})' &= f(x_{n+1}, \tilde{y}_{n+1}), \\ (\tilde{y}_{n+1})'' &= f'_y(x_{n+1}, \tilde{y}_{n+1})(\tilde{y}_{n+1})' + f'_x(x_{n+1}, \tilde{y}_{n+1}).\end{aligned}$$

Korektor

$$y_{n+1} = y_n + \frac{1}{2}h(y'_{n+1} + y'_n) + \frac{1}{12}h^2(-y''_{n+1} + y''_n).$$

Příklad 6.21. Numericky řešte úlohu $y' = -y$, $y(0) = 1$. Výsledek srovnejte s analytickým řešením $Y(x) = e^{-x}$.

Řešení. Výsledky jsou zapsány v tabulce

x	y	chyba
0	1	0
0.1	0.90483753	$-1.1 \cdot 10^{-7}$
0.2	0.81873093	$-1.8 \cdot 10^{-7}$
\ddots	\ddots	\ddots
3	$4.9787110 \cdot 10^{-2}$	$-4.2 \cdot 10^{-8}$
\ddots	\ddots	\ddots
15	$3.0590305 \cdot 10^{-7}$	$-7.3 \cdot 10^{-3}$

Povšimněte si, že na konci výpočtu je už chyba v absolutní hodnotě větší než funkční hodnota. Současně je ale třeba vzít v úvahu i počet kroků, které jsme vykonali, tj. 150. \square

Příklad 6.22. Numericky řešte úlohu $y' = y$, $y(0) = 1$. Výsledek srovnejte s analytickým řešením $Y(x) = e^x$.

Řešení. Výsledky jsou zapsány v tabulce

x	y	chyba
0	1	0
0.1	1.1051708	$1 \cdot 10^{-7}$
0.2	1.2214024	$4 \cdot 10^{-7}$
0.3	1.3498582	$6 \cdot 10^{-7}$
\ddots	\ddots	\ddots
3	20.085520	$1.7 \cdot 10^{-5}$
\ddots	\ddots	\ddots
15	3269011.1	6.3

\square

Příklad 6.23. Numericky řešte úlohu $y' = \frac{1}{1 + \tan^2 y}$, $y(0) = 0$. Výsledek srovnejte s analytickým řešením $Y(x) = \arctan x$.

Řešení. Výsledky jsou zapsány v tabulce

x	y	chyba
0	0	0
0.1	$9.9668686 \cdot 10^{-2}$	$-3.4 \cdot 10^{-8}$
0.2	0.19739560	$-4 \cdot 10^{-8}$
0.3	0.29145683	$-4 \cdot 10^{-8}$
\ddots	\ddots	\ddots
3	1.2490458	0
\ddots	\ddots	\ddots
15	1.5042272	$1 \cdot 10^{-6}$

□

6.7 Metody Taylorovy řady

Eulerovu metodu jsme si odvodili z Taylorovy řady

$$y(x+h) = y(x) + y'(x)h + \frac{y''(x)h^2}{2!} + \frac{y'''(x)h^3}{3!} + \dots + \frac{y^{(k)}(x)h^k}{k!} + O(h^{(k+1)}),$$

když jsme položili $k = 1$ a zanedbali jsme členy vyšších řádů. Přitom jsme se dopustili chyby řádově rovné $O(h^2)$. Dosazením za první derivaci $y'(x) = f(x, y)$ jsme dostali

$$y_{i+1} = y_i + h_i f(x_i, y_i), \quad i = 0, 1, \dots, n-1.$$

Podobně můžeme dostat pro $k = 2$ metodu druhého řádu. Stačí jen zanedbat členy řádů vyšších než dvě a vyjádřit si druhou derivaci $y''(x)$. Dostáváme

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

Dosazením tohoto vztahu do Taylorovy řady dostaneme vztah pro metodu Taylorovy řady druhého řádu:

$$y_{n+1} = y_n + hf(x_n, y_n) + \frac{1}{2}h^2 \left[\frac{\partial f(x_n, y_n)}{\partial x} + \frac{\partial f(x_n, y_n)}{\partial y} \cdot f(x_n, y_n) \right].$$

Obdobně můžeme pokračovat dále a dostaneme

$$y_{n+1} = y_n + hf_n + \frac{1}{2}h^2 f_n^{(1)} + \dots + \frac{1}{k!}f_n^{(k)}.$$

Výrazy pro vyšší derivace funkce $f(x, y)$ mohou být obecně velmi složité. Pro některé speciální případy však tyto výrazy mohou nabývat jednoduchých tvarů. Potom je možné metodu Taylorovy řady vyššího řádu použít s velmi dobrými výsledky.

Příklad 6.24. Hledejme řešení úlohy $y' = qy$, $y(0) = y_0$ metodou Taylorovy řady.

Řešení. Rozvojem do Taylorovy řady dostaneme, pokud použijeme prvních l členů:

$$y_{i+1} = \sum_{j=0}^l \frac{h^j}{j!} y_i^{(j)} = \sum_{j=0}^l \xi_{i,j},$$

kde $\xi_{i,j} = \frac{h^j}{j!} y_i^{(j)}$. Současně ale platí

$$y^{(j)} = qy^{(j-1)}, \Rightarrow y_i^{(j)} = qy_i^{(j-1)}.$$

Potom

$$\frac{\xi_{i,j}}{\xi_{i,j-1}} = \frac{\frac{h^j}{j!} y_i^{(j)}}{\frac{h^{j-1}}{(j-1)!} y_i^{(j-1)}} = \frac{h y_i^{(j)}}{j y_i^{(j-1)}} = \frac{h}{j} \cdot q,$$

neboli

$$\xi_{i,j} = \frac{h}{j} q \xi_{i,j-1}, \quad j = 1, 2, \dots, l.$$

Když si navíc uvědomíme, že $\xi_{i,0} = y_i$, tak jsme získali rekurentní posloupnost, která nám umožňuje spočítat řešení s libovolnou přesností. Všimněte si přitom, že všechny posloupnosti jsou stejné a nezávisí na i . Navíc je možné podle potřeby průběžně měnit krok h . \square

V obecném případě má metoda Taylorovy řady tvar:

Na intervalu $I = [a, b]$ hledáme řešení Cauchyovy úlohy $y' = f(x, y)$, $y(x_0) = y_0$, $x_0 \in I$, $a, b, y_0 \in \mathbb{R}$. Funkci y si rozvineme do Taylorovy řady

$$y(x+h) = y(x) + hy'(x) + \frac{1}{2!} h^2 y''(x) + \frac{1}{3!} h^3 y'''(x) + \dots + \frac{1}{k!} h^k y^{(k)}(x) + O(h^{k+1}).$$

Pro $k = 1$ dostaneme, při zanedbání vyšších řádů, Eulerovu metodu s chybou $O(h^2)$.

Pokud si zvolíme $k = 2$, dostaneme chybu $O(h^3)$. Stačí si jen vyjádřit druhou derivaci

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

Neboli

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

Pro $k = 3$ potřebujeme ještě třetí derivaci.

$$y'''(x) = \frac{d}{dx} (y''(x)) = \frac{d}{dx} (f'_x + f'_y \cdot f) = f''_{xx} + f'_{xy} \cdot y' + f'_{yx} \cdot f + f'_y \cdot f'_x + f''_{yy} \cdot y' \cdot f + f'_y \cdot f'_y \cdot y'.$$

6.8 Diferenciální rovnice vyšších řádů

Definice 6.25. Cauchyova úloha pro diferenciální rovnici n -tého řádu má tvar

$$y^{(n)}(x) = f(x, y(x), y'(x), \dots, y^{(n-1)}(x)), \quad (6.8)$$

$$y(x_0) = k_0, y'(x_0) = k_1, \dots, y^{(n-1)}(x_0) = k_{n-1}, \quad (6.9)$$

kde $x \in I = [a, b]$, $k_i, i = 0, 1, \dots, n-1$ jsou konstanty a f je funkce $n+1$ proměnných definovaná na otevřené množině $\Omega \subset \mathbb{R}^{n+1}$.

Řešením pak rozumíme funkci $y(x)$ definovanou a spojitou na I , která je na I n krát spojitě diferencovatelná, pro každé $x \in I$ splňuje rovnici (6.8) a vyhovuje počátečním podmínkám (6.9).

Definice 6.26. Funkce $f(x, z_0, z_1, \dots, z_{n-1})$ splňuje v bodě $(x_0, k_0, k_1, \dots, k_{n-1}) \in \Omega \subset \mathbb{R}^{n+1}$ lokálně Lipschitzovu podmínku, jestliže existuje konstanta $K > 0$ a okolí U bodu $(x_0, k_0, k_1, \dots, k_{n-1})$, $U \subset \Omega$, takové, že pro každé dva body $(x_0, a_0, a_1, \dots, a_{n-1}) \in U$, $(x_0, b_0, b_1, \dots, b_{n-1}) \in U$, platí

$$|f(x_0, a_0, a_1, \dots, a_{n-1}) - f(x_0, b_0, b_1, \dots, b_{n-1})| \leq K \sum_{i=0}^{n-1} |a_i - b_i|.$$

Přitom okolím bodu $(x_0, k_0, k_1, \dots, k_{n-1})$ rozumíme libovolnou otevřenou kouli v \mathbb{R}^{n+1} se středem v tomto bodě.

Věta 6.27. Lokální Lipschitzova podmínka v Ω pro funkci $f(x, z_0, z_1, \dots, z_{n-1})$ bude splněna, jestliže v Ω existují lokálně ohraničené parciální derivace $\frac{\partial f}{\partial z_0}, \frac{\partial f}{\partial z_1}, \dots, \frac{\partial f}{\partial z_{n-1}}$.

Lokální Lipschitzova podmínka je splněna automaticky, jsou-li tyto derivace spojité v Ω .

Věta 6.28. (O existenci a jednoznačnosti řešení.)

Nechť funkce $f(x, z_0, z_1, \dots, z_{n-1})$ je spojitá na otevřené množině $\Omega \in \mathbb{R}^{n+1}$. Pak pro každé $(x_0, k_0, k_1, \dots, k_{n-1}) \in \Omega$ má úloha (6.8), (6.9) aspoň jedno řešení.

Je-li navíc v každém bodě Ω splněna lokálně Lipschitzova podmínka, je toto řešení právě jedno.

Opět je vždy nutno prověřit splnitelnost podmínek existence a jednoznačnosti řešení, protože jednotlivé numerické metody řešení se tím nezabývají a vždy předpokládají jejich platnost.

Mějme úlohu (6.8), (6.9). Zavedeme si nové neznámé funkce $y_i(x)$ předpisem: $y = y_1, y' = y_2, \dots, y^{(n-1)} = y_n$, potom máme soustavu

$$\begin{aligned} y_1' &= y_2, \\ y_2' &= y_3, \\ \cdot &\quad \cdot \\ y_{n-1}' &= y_n, \\ y_n' &= f(x, y_1, y_2, \dots, y_n). \end{aligned} \quad (6.10)$$

Stejnou transformaci provedeme i s počátečními podmínkami (6.9).

$$\begin{aligned} y(x_0) &= y_1(x_0) = k_0, \\ y'(x_0) &= y_2(x_0) = k_1, \\ &\vdots \\ y^{(n-1)}(x_0) &= y_n(x_0) = k_n. \end{aligned} \tag{6.11}$$

Tímto způsobem převedeme numerické řešení diferenciálních rovnic vyšších řádů na řešení soustavy diferenciálních rovnic prvního řádu.

Věta 6.29. *Funkce $y(x)$, $x \in I$ je řešením rovnice (6.8) právě tehdy, když jsou funkce $(y(x), y'(x), \dots, y^{(n-1)}(x))$ řešením soustavy (6.10).*

Diferenciální rovnice vyšších řádů mají rozsáhlé aplikace. Například v mechanice, teorii pružnosti, teorii elektrických obvodů, abychom se zmínili alespoň o těch oborech, se kterými se budete nejčastěji setkávat. Přitom jedna rovnice se může vyskytovat v několika oborech.

Příklad 6.30. Homogenní lineární diferenciální rovnice druhého řádu s konstantními koeficienty

$$y'' + 2ay' + b^2y = 0, \quad a \geq 0, b > 0,$$

popisuje

1. Kmity struny
2. Matematické kyvadlo
3. Elektrický obvod RLC

Jde o tzv. rovnici lineárního oscilátoru.

Příklad 6.31. Nehomogenní rovnice tvaru

$$y'' + 2ay' + b^2y = f(x), \quad a \geq 0, b > 0,$$

popisuje tzv. nucené kmity. Výsledek potom podstatně závisí na konkrétním tvaru budícího „členu“ $f(x)$.

6.8.1 Numerické metody pro rovnice druhé řádu

Protože se rovnice druhé řádu často vyskytují v různých aplikacích, ukážeme si speciální metody pro jejich řešení. Obecný tvar rovnice druhé řádu je

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

s počátečními podmínkami

$$y(x_0) = k_0, \quad y'(x_0) = k_1,$$

kde $k_0, k_1 \in \mathbb{R}$ jsou libovolné konstanty. Tuto rovnici lze přepsat na soustavu tvaru

$$\begin{aligned} z' &= f(x, y, z), \\ y' &= z. \end{aligned}$$

s počátečními podmínkami

$$\begin{aligned} y(x_0) &= k_0, \\ z(x_0) &= k_1. \end{aligned}$$

Pro její řešení můžeme použít libovolnou z obecných metod.

Pokud však pravá strana nezávisí na y' , což se dosti často objevuje v aplikacích, můžeme dosáhnout podstatného zlepšení. Rovnice je potom ve tvaru

$$y''(x) = f(x, y), \quad y(x_0) = k_0, \quad y'(x_0) = k_1.$$

Pro přímé odvození metody pro řešení rovnice je přirozené vzít numerickou metodu ve tvaru (označení je stejné jako v předchozím výkladu)

$$y_{n+1} = \sum_{i=0}^p a_i y_{n-i} + h^2 \sum_{i=-1}^p b_i y''_{n-i},$$

kde $p \in \mathbb{N}$ nám určuje řád metody. Potom pro $b_{-1} = 0$ máme prediktor a pro $b_{-1} \neq 0$ máme korektor. Metodou neurčitých koeficientů si odvodíme parametry a_i, b_i a dostaneme např.:

Prediktor

$$y_{n+1} = 2y_{n-1} + y_{n-3} + \frac{4}{3}h^2 (y''_n + y''_{n-1} + y''_{n-2}),$$

Korektor

$$y_{n+1} = 2y_n - y_{n-1} + \frac{1}{12}h^2 (y''_{n+1} + 10y''_n + y''_{n-1}).$$

A když nyní využijeme toho, že $y'' = f(x, y)$ dostaneme

Prediktor

$$y_{n+1} = 2y_{n-1} + y_{n-3} + \frac{4}{3}h^2 (f_n + f_{n-1} + f_{n-2}),$$

Korektor

$$y_{n+1} = 2y_n - y_{n-1} + \frac{1}{12}h^2 (f_{n+1} + 10f_n + f_{n-1}).$$

Pomocí metody polovičního kroku dostaneme řešení s potřebnou přesností.

6.8.2 Užití Taylorovy řady

Pro hledání řešení diferenciálních rovnic vyšších řádů můžeme opět použít Taylorovu řadu

$$y(x) = y(x_0) + y'(x_0)(x - x_0) + \frac{y''(x_0)}{2!}(x - x_0)^2 + \frac{y'''(x_0)}{3!}(x - x_0)^3 + \dots$$

Pokud dovedeme určit hodnoty derivací v bodě x_0 , jde o velmi efektivní metodu. Postup si ukážeme na příkladě.

Příklad 6.32. Najděte řešení počáteční úlohy

$$y'' + xy' + y = 0, \quad y(0) = 0, \quad y'(0) = 1.$$

Řešení: Rovnici si upravíme na tvar

$$y'' = -xy' - y. \quad (6.12)$$

Dosazením počátečních podmínek dostaneme

$$y''(0) = -0 \cdot 1 - 0 = 0.$$

Dále derivací (6.12) postupně dostáváme

$$\begin{aligned} y''' &= -xy'' - 2y', \\ y^{(IV)} &= -xy''' - 3y'', \\ y^{(V)} &= -xy^{(IV)} - 4y''', \dots \end{aligned}$$

Postupným dosazováním už známých počátečních podmínek dostáváme

$$\begin{aligned} y'''(0) &= -0 \cdot 0 - 2 \cdot 1 = -2, \\ y^{(IV)}(0) &= 0, \quad y^{(V)}(0) = 8, \dots \end{aligned}$$

Dosazením vypočítaných hodnot do Taylorovy řady

$$y(x) = y(0) + y'(0)x + \frac{y''(0)}{2!}x^2 + \frac{y'''(0)}{3!}x^3 + \dots$$

dostaneme

$$y(x) = x - \frac{x^3}{3} + \frac{x^5}{15} + \dots,$$

což je námi hledané řešení rovnice v okolí bodu 0, které máme vyjádřeno pomocí řady.

Pojmy k zapamatování

- Zabývali jsme se řešením obyčejných diferenciálních rovnic. Připomenuli jsme si základní pojmy z teorie obyčejných diferenciálních rovnic a analytické metody řešení některých typů těchto rovnic. Analytické řešení jsme získali buď v explicitním tvaru $y = f(x)$ a nebo v implicitním tvaru $F(x, y) = 0$. Takto řešitelných rovnic je ale menšina.
- Proto jsme se dále zabývali numerickými metodami řešení obyčejných diferenciálních rovnic. Numerické jsme hledali na základě diskretizace proměnné. Místo spojité funkce, která je řešením u analytických metod, jsme hledali diskrétní funkci, která se až v limitě bude blížit přesnému řešení.
- Z numerických metod jsme probrali jednokrokové metody - Eulerovu a její modifikace, metodou Rungeho-Kutty. Z vícekových metod jsme probrali Adamsovy metody, metody prediktor-korektor a prediktor-modifikátor-korektor.

- Krátce jsme se zmínili i o stabilitě řešení. Obsah kurzu nepředpokládá u čtenáře znalosti o stabilitě. Jde pouze o informaci pro lepší pochopení problematiky numerických výpočtů.
- Připomněli jsme si definici diferenciální rovnice vyššího řádu a jejího řešení.
- Ukázali jsme si, jak lze diferenciální rovnici řádu n převést na soustavu n diferenciálních rovnic prvního řádu. A odkázali jsme se na následující kapitolu, kde se budeme této problematice věnovat.
- Ukázali jsme si speciální metody pro rovnice druhého řádu, včetně použití Taylorovy řady.

Kontrolní otázky

1. Co rozumíme pojmem stabilita řešení?
2. Kolik kroků je vhodné provádět u metody Rungeho-Kutty?
3. Existuje nějaké omezení na přesnost řešení?
4. Co rozumíme řešením diferenciální rovnice vyššího řádu ?
5. Proč se speciálně zabýváme rovnicemi druhého řádu?
6. Je metoda Taylorovy řady univerzální?

6.9 Cvičení

Maplety

Příklady pro samostatnou práci jsou uvedeny v samostatné sbírce příkladů. Pomocí následujících mapletů si můžete usnadnit některé dílčí výpočty, nebo zkontrolovat jejich správnost.

1. [Derivování](#)
2. [Integrovaní](#)
3. [Výpočet funkčních hodnot](#)
4. [Úprava algebraických výrazů](#)
5. [Analytické řešení soustavy lineárních rovnic](#)
6. [Určení typu diferenciální rovnice](#)
7. [Separovatelné diferenciální rovnice](#)
8. [Bernoulliova diferenciální rovnice](#)
9. [Lineární diferenciální rovnice prvního řádu](#)
10. [Rozhodnutí, zda funkce je nebo není řešením dané diferenciální rovnice](#)

6.10 Animace

Na následujících animacích jsou znázorněny některé aspekty numerických metod studovaných v této kapitole.

Animace ovládáte tlačítky pod obrázkem. Kvalita animací je dána skutečností, že se jedná o bitmapové obrázky exportované ze softwaru Maple a následně vkládané do textu v jiné velikosti. Animace se nacházejí vždy na samostatné stránce.

Obr. 6.5: Vliv délky kroku na přesnost numerického řešení počáteční úlohy; analytické řešení znázorněno zeleně, použita Eulerova metoda (animace)

7 Řešení soustav obyčejných diferenciálních rovnic

Průvodce studiem

V předchozí kapitole jsme si ukázali, jak můžeme řešení diferenciální rovnice vyššího řádu převést na řešení soustavy diferenciálních rovnic prvního řádu.

Cílem této kapitoly je seznámit čtenáře s numerickými metodami řešení soustav obyčejných diferenciálních rovnic.

Budeme se zabývat Eulerovou metodou a Rungeho-Kuttovou metou pro soustavy diferenciálních rovnic. Půjde o zobecnění už známých metod pro jednorozměrné řešení na vícedimenzionální případ.

Řešení budeme zase hledat ve tvaru diskrétní funkce, která za určitých podmínek v limitě konverguje k přesnému řešení.

Cíle

Po prostudování této kapitoly budete schopni:

- Řešit soustavu diferenciálních rovnic prvního řádu.
- Soustavy diferenciálních rovnic vyšších řádů převést na soustavu diferenciálních rovnic prvního řádu.
- Zvolit vhodnou metodu pro řešení dané soustavy.

7.1 Základní pojmy

Definice 7.1. *Cauchyovou úlohou pro soustavu diferenciálních rovnic prvního řádu rozumíme soustavu rovnic (ve vektorovém tvaru)*

$$Y' = F(x, Y), \quad (7.1)$$

s počáteční podmínkou

$$Y(x_0) = Y_0, \quad (7.2)$$

kde $x \in I = [a, b]$, $Y = (y_1(x), y_2(x), \dots, y_n(x))$, $Y' = (y_1'(x), y_2'(x), \dots, y_n'(x))^T$, $F(x, Y)$ je vektorová funkce $F(x, Y) = (f_1(x, Y), f_2(x, Y), \dots, f_n(x, Y))^T$ definovaná na otevřené množině $\Omega \subset \mathbb{R}^{n+1}$ a $Y_0 = (y_1^0, y_2^0, \dots, y_n^0)$ je vektor z \mathbb{R}^n .

Soustavu 7.1 si můžeme zapsat i ve skalárním tvaru

$$y_i' = f_i(x, y_1, y_2, \dots, y_n), i = 1, 2, \dots, n.$$

Věta 7.2. *Nechť $F(x, Y)$ je spojitá na Ω . Pak pro každé $(x_0, Y_0) \in \Omega$ má úloha (7.1), (7.2) alespoň jedno řešení.*

Splňuje-li navíc $F(x, Y)$ v každém bodě Ω lokálně Lipschitzovu podmínku, je řešení právě jedno.

Věta 7.3. *Nechť je dána soustava (7.1) a nechť funkce $f_i, \partial f_i / \partial y_j, i, j = 1, 2, \dots, n$, jsou spojité v oblasti $\Omega \in \mathbb{R}^{n+1}$. Potom pro libovolně bod $[x_0, Y_0] \in \Omega$ existuje právě jedno řešení Y soustavy (7.1), splňující podmínky (7.2), definované v takovém intervalu I , že $x_0 \in I$ a pro každé $x \in I$ je bod $[x, Y(x)] \in \Omega$.*

Jde tady vlastně o zeslabení podmínek předchozí věty, protože ze spojitosti parciálních derivací v oblasti Ω plyne platnost lokální Lipschitzovy podmínky.

Numerické metody řešení Cauchyovy úlohy pro soustavy diferenciálních rovnic jsou podobné metodám řešení Cauchyovy úlohy pro rovnici $y' = f(x, y)$.

7.2 Eulerova metoda pro soustavy diferenciálních rovnic

Mějme dánu úlohu (7.1), (7.2). Předpokládáme, že F je spojitá na Ω a splňuje lokálně Lipschitzovu podmínku, takže máme podle věty 7.2 zajištěnou existenci a jednoznačnost řešení. Potom můžeme postupovat analogicky jako u rovnice (6.1). Z Taylorova rozvoje dostaneme, po zanedbání derivací vyšších řádů, ve vektorovém tvaru pro krok $h = \frac{b-a}{n}$

$$Y_{i+1} = Y_i + hF(x_i, Y_i), i = 0, 1, \dots, n-1. \quad (7.3)$$

V případě soustavy dvou diferenciálních rovnic máme

$$u' = f(x, u(x), v(x)),$$

$$v' = \varphi(x, u(x), v(x)),$$

$$u(x_0) = u_0, v(x_0) = v_0.$$

Jestliže si pro tuto soustavu rozepíšeme vztahy (7.3), dostaneme

$$u_{i+1} = u_i + hf(x_i, u_i, v_i),$$

$$v_{i+1} = v_i + h\varphi(x_i, u_i, v_i),$$

$$i = 0, 1, \dots, n, x_0 = a, x_n = b, h = (b-a)/n.$$

Příklad 7.4. Řešte soustavu

$$\begin{aligned}u' &= v, \\v' &= 1 + e^u, \\u(0) &= 0, \quad v(0) = 0, \quad a = 0, \quad b = 0.4, \quad h = 0.1.\end{aligned}$$

Řešení. Platí, že

$$\begin{aligned}u_1 &= u_0 + 0.1v_0 = 0, \\v_1 &= v_0 + 0.1(1 + e^{u_0}) = 0 + 0.1(1 + e^0) = 0.2, \\u_2 &= u_1 + 0.1v_1, \\v_2 &= v_1 + 0.1(1 + e^{u_1}).\end{aligned}$$

Další výsledky jsou uvedeny v tabulce

i	x	u	v
0	0	0	0
1	0.1	0	0.2
2	0.2	0.02	0.4
3	0.3	0.06	0.602
4	0.4	0.1202	0.8082

□

7.3 Rungeho-Kuttova metoda pro soustavy diferenciálních rovnic

Postup je opět analogický s řešením rovnice (6.1). Odlišnost je pouze v tom, že y , f a k_j , $j = 1, 2, 3, 4$, chápeme jako vektory. Jinak zůstávají všechny vztahy beze změny. Konkrétní postup si ukážeme na příkladě:

Příklad 7.5. Mějme danu Cauchyovu úlohu pro soustavu dvou diferenciálních rovnic

$$\begin{aligned}y'(x) &= f(x, y(x), z(x)), \\z'(x) &= \varphi(x, y(x), z(x)), \\y(x_0) &= y_0, \quad z(x_0) = z_0.\end{aligned}$$

Nechť $y_i = y_0 + ih$, $z_i = z_0 + ih$, $h = (b - a)/n$ je krok. Potom R-K. metoda čtvrtého řádu je popsána vztahy $i = 1, 2, \dots, n - 1$,

$$y_{i+1} = y_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4), \quad z_{i+1} = z_i + \frac{1}{6}(l_1 + 2l_2 + 2l_3 + l_4),$$

$$\begin{aligned}
k_1 &= hf(x_i, y_i, z_i), & l_1 &= h\varphi(x_i, y_i, z_i), \\
k_2 &= hf\left(x_i + \frac{h}{2}, y_i + \frac{k_1}{2}, z_i + \frac{l_1}{2}\right), & l_2 &= h\varphi\left(x_i + \frac{h}{2}, y_i + \frac{k_1}{2}, z_i + \frac{l_1}{2}\right), \\
k_3 &= hf\left(x_i + \frac{h}{2}, y_i + \frac{k_2}{2}, z_i + \frac{l_2}{2}\right), & l_3 &= h\varphi\left(x_i + \frac{h}{2}, y_i + \frac{k_2}{2}, z_i + \frac{l_2}{2}\right), \\
k_4 &= hf(x_i + h, y_i + k_3, z_i + l_3), & l_4 &= h\varphi(x_i + h, y_i + k_3, z_i + l_3).
\end{aligned}$$

Příklad 7.6. Pro soustavu tří rovnic máme:

$$\begin{aligned}
y'(x) &= f(x, y(x), z(x), u(x)), \\
z'(x) &= \varphi(x, y(x), z(x), u(x)), \\
u'(x) &= g(x, y(x), z(x), u(x)), \\
y(x_0) &= y_0, \quad z(x_0) = z_0, \quad u(x_0) = u_0.
\end{aligned}$$

Nechť $y_i = y_0 + ih$, $z_i = z_0 + ih$, $u_i = u_0 + ih$, $h = (b - a)/n$ je krok. Potom Rungeho-Kuttova metoda čtvrtého řádu je popsána vztahy

$$\begin{aligned}
y_{i+1} &= y_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4), & z_{i+1} &= z_i + \frac{1}{6}(l_1 + 2l_2 + 2l_3 + l_4), \\
u_{i+1} &= u_i + \frac{1}{6}(m_1 + 2m_2 + 2m_3 + m_4), \\
k_1 &= hf(x_i, y_i, z_i, u_i), & l_1 &= h\varphi(x_i, y_i, z_i, u_i), \\
m_1 &= hg(x_i, y_i, z_i, u_i), \\
k_2 &= hf\left(x_i + \frac{h}{2}, y_i + \frac{k_1}{2}, z_i + \frac{l_1}{2}, u_i + \frac{m_1}{2}\right), & l_2 &= h\varphi\left(x_i + \frac{h}{2}, y_i + \frac{k_1}{2}, z_i + \frac{l_1}{2}, u_i + \frac{m_1}{2}\right), \\
m_2 &= hg\left(x_i + \frac{h}{2}, y_i + \frac{k_1}{2}, z_i + \frac{l_1}{2}, u_i + \frac{m_1}{2}\right), \\
k_3 &= hf\left(x_i + \frac{h}{2}, y_i + \frac{k_2}{2}, z_i + \frac{l_2}{2}, u_i + \frac{m_2}{2}\right), & l_3 &= h\varphi\left(x_i + \frac{h}{2}, y_i + \frac{k_2}{2}, z_i + \frac{l_2}{2}, u_i + \frac{m_2}{2}\right), \\
m_3 &= hg\left(x_i + \frac{h}{2}, y_i + \frac{k_2}{2}, z_i + \frac{l_2}{2}, u_i + \frac{m_2}{2}\right), \\
k_4 &= hf(x_i + h, y_i + k_3, z_i + l_3, u_i + m_3), & l_4 &= h\varphi(x_i + h, y_i + k_3, z_i + l_3, u_i + m_3), \\
m_4 &= hg(x_i + h, y_i + k_3, z_i + l_3, u_i + m_3).
\end{aligned}$$

Pro rozsáhlejší soustavy je postup analogický.

Příklad 7.7. Řešme stejnou úlohu jako u Eulerovy metody, tj. příklad 7.4.

$$\begin{aligned}
u' &= v, \\
v' &= 1 + e^u, \\
u(0) &= 0, \quad v(0) = 0, \quad a = 0, \quad b = 0.4, \quad h = 0.1.
\end{aligned}$$

Řešení. Potom dostaneme

n	x	u	v	k	l
0	0	0	0	$k_1 = 0$	$l_1 = 0.2$
	0.05	0	0.1	$k_2 = 0.01$	$l_2 = 0.2$
	0.05	0.005	0.1	$k_3 = 0.01$	$l_3 = 0.2005013$
	0.1	0.01	0.2005013	$k_4 = 0.0200501$	$l_4 = 0.2003346$
				$\Delta u = 0.0100084$	$\Delta v = 0.2003346$
1	0.1	0.0100084	0.2003346		
2	0.2	0.040135	0.402705		
3	0.3	0.090689	0.609292		
4	0.4	0.162216	0.822595		

V tabulce je rozepsán pouze první krok. U dalších kroků jsou uvedeny pouze výsledky. \square

Pokud máme soustavu diferenciálních rovnic, která obsahuje i rovnice vyšších řádů, tak si každou takovou rovnici převedeme na soustavu rovnic prvního řádu stejným postupem, který byl uveden v předchozí části.

Příklad 7.8. Mějme dānu soustavu

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

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

a počáteční podmínky $y(x_0) = y_0$, $y'(x_0) = y_1$, $z(x_0) = z_0$. Transformací $y(x) = u(x)$, $y'(x) = v(x)$ dostaneme úlohu

$$u' = v(x),$$

$$v' = f(x, u(x), v(x), z(x)),$$

$$z' = g(x, u(x), v(x), z(x)),$$

$$u(x_0) = y_0, v(x_0) = y_1, z(x_0) = z_0.$$

Tím máme soustavu tří diferenciálních rovnic prvního řādu. O tē uŹ vīme, jak ji řešit.

I u tēchto úloh hraje dūležitou roli stabilita vypočtu. Například mějme soustavu

$$y_1' = y_2,$$

$$y_2' = -1001y_2 - 1000y_1.$$

s počátečními podmínkami

$$y_1(0) = 1, \quad y_2(0) = -1.$$

Přesné řešení tēto úlohy je

$$y_1(x) = e^{-x}, \quad y_2(x) = -e^{-x}$$

a obecné řešení soustavy je

$$y_i(x) = A_i e^{-x} + B_i e^{-1000x},$$

$i = 1, 2, A, B \in \mathbb{R}$. Všimněte si, že přesné řešení neobsahuje členy $B_i e^{-1000x}$. Vlastní čísla matice soustavy jsou $\lambda_1 = -1, \lambda_2 = -1000$. Odtud plyne nestabilita výpočtu.

V tomto případě pro krok $h \geq 0.003$ při klasické Rungeho-Kuttově metodě čtvrtého řádu vůbec nedostaneme použitelné výsledky.

Opět je nutné hlídat podmínky existence a jednoznačnosti řešení.

7.4 Metoda Taylorovy řady

Mějme soustavu diferenciálních rovnic (ve vektorovém tvaru)

$$\frac{d\mathbb{Y}}{dx} = \mathbb{F}(x, \mathbb{Y}^T),$$

kde

$$\mathbb{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad \mathbb{F} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}, \quad f_i = f(x, y_1, y_2, \dots, y_n).$$

Rozvinutím do Taylorovy řady v okolí bodu x_0 dostaneme

$$\mathbb{Y}(x) = \sum_{k=0}^{\infty} \frac{\mathbb{Y}^{(k)}(x_0)}{k!} (x - x_0)^k,$$

což znamená, že každá složka \mathbb{Y} se rozkládá do Taylorovy řady, kde

$$\mathbb{Y}^{(2)} = \frac{\partial^2 \mathbb{Y}}{\partial x^2} = \frac{\partial \mathbb{F}}{\partial x} + \frac{\partial \mathbb{F}}{\partial \mathbb{Y}} \cdot \frac{\partial \mathbb{Y}}{\partial x} = \frac{\partial \mathbb{F}}{\partial x} + \frac{\partial \mathbb{F}}{\partial \mathbb{Y}} \cdot \mathbb{F}$$

a

$$\frac{\partial \mathbb{F}}{\partial \mathbb{Y}} = \begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial y_2} & \cdots & \frac{\partial f_1}{\partial y_n} \\ \frac{\partial f_2}{\partial y_1} & \frac{\partial f_2}{\partial y_2} & \cdots & \frac{\partial f_2}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial y_1} & \frac{\partial f_n}{\partial y_2} & \cdots & \frac{\partial f_n}{\partial y_n} \end{pmatrix}.$$

Pro derivace vyšších řádů postupujeme analogicky. Postup si opět ukážeme na příkladě:

Příklad 7.9. Určete řešení soustavy rovnic

$$\frac{dx}{dt} = x \cos t - y \sin t, \tag{7.4}$$

$$\frac{dy}{dt} = x \sin t + y \cos t, \tag{7.5}$$

s počátečními podmínkami

$$y(0) = 0, \quad x(0) = 1.$$

Řešení. Rozvoje do Taylorovy řady v okolí bodu $t = 0$ mají tvar

$$x(t) = x(0) + x'(0)t + \frac{x''(0)}{2!}t^2 + \frac{x'''(0)}{3!}t^3 + \dots \quad (7.6)$$

$$y(t) = y(0) + y'(0)t + \frac{y''(0)}{2!}t^2 + \frac{y'''(0)}{3!}t^3 + \dots \quad (7.7)$$

Pro jejich použití si musíme umět vyjádřit hodnoty derivací. Z počátečních podmínek pro $t = 0$ dostaneme

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

Derivujeme (7.4),(7.5) a dostaneme

$$\frac{d^2x}{dt^2} = -x \sin t - y \cos t + \frac{dx}{dt} \cos t - \frac{dy}{dt} \sin t, \quad (7.8)$$

$$\frac{d^2y}{dt^2} = x \cos t - y \sin t + \frac{dx}{dt} \sin t + \frac{dy}{dt} \cos t. \quad (7.9)$$

Odtud po dosazení dostaneme

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

Derivujeme (7.8),(7.9) a dostaneme

$$\frac{d^3x}{dt^3} = -x \cos t + y \sin t + 2 \left(-\frac{dx}{dt} \sin t - \frac{dy}{dt} \cos t \right) + \frac{d^2x}{dt^2} \cos t - \frac{d^2y}{dt^2} \sin t,$$

$$\frac{d^3y}{dt^3} = -x \sin t - y \cos t + 2 \left(\frac{dx}{dt} \cos t - \frac{dy}{dt} \sin t \right) + \frac{d^2x}{dt^2} \sin t + \frac{d^2y}{dt^2} \cos t.$$

Odtud po dosazení dostaneme

$$x'''(0) = 0, \quad y'''(0) = 3.$$

A můžeme pokračovat dále. Pokud nám dostačují získané hodnoty, potom jejich dosazením do (7.6) a (7.7) dostaneme aproximaci řešení v okolí bodu $t = 0$:

$$x(t) = 1 + t + \frac{1}{2}t^2 + \dots$$

$$y(t) = \frac{1}{2}t^2 + \frac{1}{2}t^3 + \dots$$

□

7.5 Zaokrouhlovací chyby

Při řešení každé numerické úlohy se vyskytují zaokrouhlovací chyby. Ukážeme si, že jejich vliv nemůžeme zanedbat. Pro jednoduchost si vezmeme Eulerovu metodu pro řešení rovnice (soustavy rovnic)

$$y' = f(x, y), \quad y(x_0) = y_0.$$

Numerickým postupem s krokem $h = \frac{b-a}{n}$ získáme aproximace přesného řešení $y = y(x)$, které budeme označovat jako \tilde{y}_i . Předpokládejme, že

$$\tilde{y}_{i+1} = \tilde{y}_i + h \cdot f(x_i, \tilde{y}_i) + \varepsilon_i, \quad i = 0, 1, \dots, n-1$$

a $|\varepsilon_i| < \varepsilon$, kde ε je zadaná přesnost. Potom pro chybu r_i na $i+1$ -tém kroku $r_i = y_i - \tilde{y}_i$ platí

$$r_{i+1} = r_i + h(f(x_i, y_i) - f(x_i, \tilde{y}_i)) - \varepsilon_i.$$

Odtud s využitím Lipschitzovy podmínky (s konstantou L) dostaneme

$$|r_{i+1}| \leq |r_i| + h \cdot L|y_i - \tilde{y}_i| + |\varepsilon_i| \leq (1 + hL)|r_i| + \varepsilon.$$

Protože $r_0 = 0$, dostáváme

$$\begin{aligned} |r_1| &\leq \varepsilon, \\ |r_2| &\leq (1 + hL)\varepsilon + \varepsilon, \\ &\dots \\ |r_i| &\leq \varepsilon(1 + (1 + hL) + (1 + hL)^2 + \dots + (1 + hL)^{i-1}) = \varepsilon \frac{(1 + hL)^i - 1}{hL} \leq \\ &\leq \varepsilon \frac{e^{ihL} - 1}{hL} \leq K \cdot \varepsilon \cdot h^{-1}, \end{aligned}$$

kde $K = \frac{e^{(b-a)L} - 1}{L}$. Pro celkovou chybu potom máme

$$\max_i |y_i - \tilde{y}_i| \leq K_1 h + \varepsilon K h^{-1},$$

kde h je dostatečně malý krok a K_1, K jsou konstanty na h nezávislé.

Protože předpokládáme, že konstanta ε je malá (jde přece o velikost přípustné chyby), tak se nám vliv zaokrouhlovacích chyb projeví až po provedení velkého počtu kroků, neboli pro dostatečně malé h . Potom může dokonce dojít k převážení účinků zaokrouhlovacích chyb a vypočtené hodnoty \tilde{y}_i tak mohou být zcela bezcenné, i přes vynaloženou výpočetní snahu.

7.5.1 Vhodnost počátečních podmínek

Mějme diferenciální rovnici

$$y' = \frac{2y}{x}$$

a počáteční podmínku $y(0) = 0$.

Řešením naší rovnice bude funkce $y(x) = x^2$, protože potom

$$y' = 2x,$$

a současně

$$\frac{2y}{x} = \frac{2x^2}{x} = 2x.$$

Pokud použijeme pro řešení Eulerovu metodu s krokem h , dostaneme

$$\begin{aligned}y_0 &= 0, \\y_1 &= y_0 + hf(x, y) = y_0 + \frac{2y_0}{x_0}.\end{aligned}$$

Dostali jsme neurčitý výraz typu „ $\frac{0}{0}$ “, neboli metoda nemůže pokračovat.

Jiný problém. Hledáme řešení rovnice

$$y' = xy.$$

Jestliže $y \neq 0$, potom máme

$$\begin{aligned}\frac{dy}{y} &= x dx, \\ \ln y &= \frac{x^2}{2} + \ln K,\end{aligned}$$

kde K je konstanta, potom

$$y = K \cdot e^{\frac{x^2}{2}}.$$

O správnosti výpočtu se můžeme přesvědčit zkouškou:

$$y' = K \cdot e^{\frac{x^2}{2}} \cdot \frac{2x}{2} = x \cdot K e^{\frac{x^2}{2}} = xy.$$

Pokud budeme mít počáteční podmínku ve tvaru $y(0) = 0$, potom z ní dostáváme, že $K = 0$. Proto jediným řešením, které vyhovuje této počáteční podmínce je nulová funkce $y(x) \equiv 0$, kterou jsme ale hned na začátku výpočtu vyloučili..

V obecném případě: Mějme rovnici $y' = f(x, y)$ s počáteční podmínkou $y(x_0) = 0$ takovou, že $f(x, 0) = 0$.

Potom nám použití Eulerovy metody dává

$$\begin{aligned}y_1 &= y_0 + hf(x_0, y_0) = 0 + hf(x_0, 0) = 0, \\y_2 &= y_1 + hf(x_1, y_1) = 0 + hf(x_1, 0) = 0, \\&\dots\end{aligned}$$

Pokud přejdeme k Rungeho-Kuttově metodě, dostaneme

$$\begin{aligned}y_1 &= y_0 + \frac{1}{6}(k_1 + 2k_2 + 2Kk_3 + k_4), \\k_1 &= h \cdot f(x_0, y_0) = 0, \\k_2 &= h \cdot f\left(x_0 + \frac{h}{2}, y_0 + \frac{k_1}{2}\right) = h \cdot f\left(x_0 + \frac{h}{2}, 0\right) = 0, \\k_3 &= h \cdot f\left(x_0 + \frac{h}{2}, y_0 + \frac{k_2}{2}\right) = h \cdot f\left(x_0 + \frac{h}{2}, 0\right) = 0, \\k_4 &= h \cdot f(x_0 + h, y_0 + k_3) = h \cdot f(x_0 + h, 0) = 0, \\y_1 &= 0.\end{aligned}$$

A stejné výsledky budeme dostávat i ve všech dalších krocích.

Pokud přejdeme k více krokové metodě, budeme opět dostávat stejné výsledky. A to přesto, že se nemusí jednat o rovnici, která má pouze nulové řešení.

Ve většině případů pomůže, když místo $y_0 = 0$ vezmeme nějaké dostatečně malé číslo ($\varepsilon = 10^{-6}$).

Pokud se při řešení diferenciální rovnice objeví opakovaně stejná hodnota y_i , potom je vhodné si znovu prověřit podmínky existence a jednoznačnosti řešení.

7.5.2 Řízení délky kroku

Zatím jsme vždy předpokládali, že délka kroku h je u dané metody konstantní. Ale ani z odvození metody ani z jejího tvaru nevyplývá nutnost tohoto požadavku. V každém kroku můžeme měnit délku kroku h . Dále si uvedeme jak.

V ideálním případě budeme mít, že numerický výsledek y_i se bude lišit od skutečného řešení $y(x_i)$ nejvýše o povolenou toleranci, tj. o požadovanou přesnost $\varepsilon > 0$, neboli bude platit

$$\|e_i\| = \|y_i - y(x_i)\| \leq \varepsilon.$$

Realita je však jiná. Současné programy pro řešení obyčejných diferenciálních rovnic dokáží zajistit pouze to, že pro dostatečně malé $\varepsilon > 0$ bude globální chyba $\|e_i\|$ dosti malá. Jednotlivé metody tohoto cíle dosahují tím, že délku kroku h vybírají tak, aby velikost lokální chyby $\|le_i\|$ nabývala stále zhruba stejné hodnoty ε . Velikost globální chyby $\|e_i\|$ dokážeme pouze odhadnout. Řídit ji zatím neumíme.

Dále je nutné mít na paměti, že sledování velikosti globální chyby je časově *velmi náročné*, proto většina programů ani sledování globální chyby neobsahuje.

Pojmy k zapamatování

- Seznámili jsme se s numerickými metodami řešení soustav obyčejných diferenciálních rovnic prvního řádu.
- Zabývali jsme se Eulerovou metodou a Rungeho-Kuttovou metou pro soustavy diferenciálních rovnic prvního řádu. Jde o analogický postup, jako v případě jedné diferenciální rovnice. Pouze je třeba změnit zápis na vektorový a pak budeme mít i stejné vzorce jako dříve.
- Dále jsme si ukázali možnosti použití Taylorovy řady.
- Opět získáme jako řešení diskrétní funkci, která při splnění konvergenčních podmínek nám v limitě konverguje k přesnému řešení.
- Všechna řešení se týkala explicitních soustav.
- Existují sice i numerické metody pro řešení implicitních rovnic $F(x, y, y') = 0$ a jejich soustav a také poloimplicitních rovnic $y' = \varphi(x, y, y')$ a jejich soustav. Tato problematika ale přesahuje možnosti našeho kurzu.

Kontrolní otázky

1. Co rozumíme pojmem řešení soustavy diferenciálních rovnic?
2. Jaký je rozdíl mezi explicitní a implicitní diferenciální rovnicí a nebo soustavou diferenciálních rovnic?
3. Jak stanovíme konec výpočtu řešení soustavy diferenciálních rovnic?

8 Řešení okrajových úloh pro obyčejné diferenciální rovnice.

Průvodce studiem

V této kapitole se budeme zabývat diferenciálními rovnicemi alespoň druhého řádu. Pouze pro takové rovnice má smysl hledat řešení okrajové úlohy. Nejdříve si zavedeme pojem okrajové úlohy a jejího řešení. Ukážeme si rozdíl mezi okrajovou a počáteční úlohou.

Cílem této kapitoly je seznámit čtenáře s numerickými metodami řešení okrajových úloh pro diferenciální rovnice druhého řádu.

Okrajovou úlohu můžeme převést na počáteční a hledat řešení počáteční úlohy některou z dříve probraných metod - to je princip metody střelby.

Jinou možností je diskretizace proměnných a převedení okrajové úlohy na řešení soustavy lineárních algebraických rovnic - to je princip metody konečných diferencí a metody konečných objemů.

Cíle

Po prostudování této kapitoly budete schopni:

- Řešit okrajovou úlohu metodou střelby.
- Řešit okrajovou úlohu metodou konečných diferencí,
- Převést úlohu na samoadjungovaný tvar.
- Řešit okrajovou úlohu metodou konečných objemů.

8.1 Základní pojmy

Definice 8.1. Okrajová úloha pro diferenciální rovnici druhého řádu je tvořena rovnicí

$$F(x, y, y', y'') = 0 \quad (8.1)$$

a okrajovými podmínkami

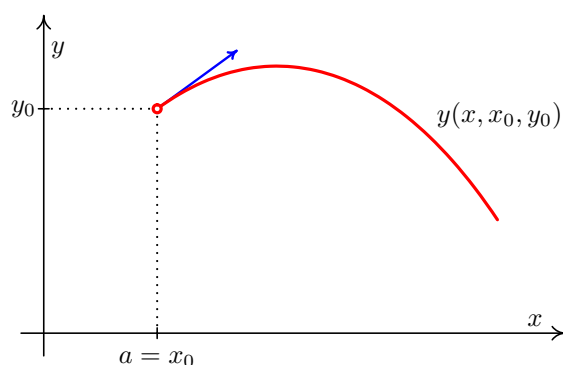
$$\alpha_1 y(a) + \alpha_2 y'(a) = A, \quad \beta_1 y(b) + \beta_2 y'(b) = B, \quad (8.2)$$

kde $A, B, \alpha_1, \alpha_2, \beta_1, \beta_2$ jsou konstanty, pro něž platí $|\alpha_1| + |\alpha_2| \neq 0$, $|\beta_1| + |\beta_2| \neq 0$.

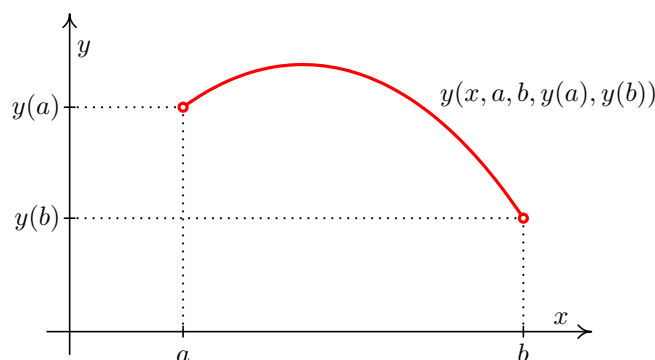
Řešení hledáme na intervalu $I = [a, b]$, $a < b$, jako spojitou a dvakrát spojitě diferencovatelnou funkci, která splňuje (8.1), (8.2).

Budeme se tedy zabývat pouze tzv. “klasickým řešením”.

Nebudeme studovat obecný tvar (8.1), ale budeme se zabývat pouze rovnicemi v jednodušším tvaru $y'' = f(x, y, y')$.



Obr. 8.1: Počáteční úloha – zadané je $y(x_0)$ a $y'(x_0)$ (tj. směrnice tečny)



Obr. 8.2: Okrajová úloha pro $\alpha_1 = 1$, $\alpha_2 = 0$, $\beta_1 = 1$, $\beta_2 = 0$

Základní rozdíl mezi počáteční (Cauchyovou) a okrajovou úlohou je v tom, že řešení počáteční úlohy existuje a je dokonce i jednoznačně určitelné pro velmi širokou třídu rovnic, zatímco u okrajové úlohy se může stát, že i pro jednoduchou lineární rovnici řešení neexistuje nebo je jich nekonečně mnoho. U rovnic vyšších řádů se přirozeně taková situace může vyskytovat s mnohem větší pravděpodobností.

Okrajová úloha pro rovnici prvního stupně nemá smysl. Řešení diferenciální rovnice prvního stupně, pokud existuje, tj. pokud jsou splněny podmínky Picardovy věty, je jednoznačně určeno už počátečním bodem. Proto se budeme zabývat rovnicemi druhého řádu.

Příklad 8.2. Mějme rovnici

$$y'' + y = 0.$$

Řešení. Její obecné řešení má tvar

$$\tilde{y}(x) = A \sin x + B \cos x$$

kde $A, B \in \mathbb{R}$ jsou libovolné konstanty.

Zvolme si okrajové podmínky

$$y(0) = 0, \quad y(\pi) = 1.$$

Potom z první podmínky dostaneme

$$\tilde{y}(0) = A \sin 0 + B \cos 0 = 0,$$

$$B \cdot 1 = 0, \quad \Rightarrow \quad B = 0.$$

Druhá podmínka nám dává

$$\tilde{y}(\pi) = 1.$$

A současně

$$\tilde{y}(\pi) = A \sin(\pi) = A \cdot 0 = 0.$$

Dostali jsme spor. Neboli neexistuje žádné řešení naší rovnice, které vyhovuje zvoleným počátečním podmínkám.

Naopak pro volbu počátečních podmínek

$$y(0) = 0, \quad y(\pi) = 0$$

existuje nekonečně mnoho řešení $\tilde{y}_1(x) = K \sin x$, $K \in \mathbb{R}$.

Okrajovým podmínkám

$$y(0) = 0, \quad y\left(\frac{\pi}{2}\right) = 1$$

vyhovuje pouze jediné řešení $\tilde{y}_2(x) = \sin x$. □

Z příkladu je zřejmé, že možností, kdy existuje právě jedno řešení či nekonečně mnoho řešení či neexistuje řešení, lze najít také nekonečně mnoho.

8.1.1 Metoda střelby

Někdy se označuje jako *balistická metoda*.

Základem metody střelby je převedení okrajové úlohy na počáteční úlohu. Mějme danu okrajovou úlohu v nejjednodušším možném tvaru

$$y'' = f(x, y, y'), \quad y(a) = A, \quad y(b) = B,$$

tj. v nejjednodušším tvaru okrajových podmínek (8.2).

Volíme si libovolně $y'(a) = k_0$ a řešíme počáteční úlohu

$$y'' = f(x, y, y'), \quad y(a) = A, \quad y'(a) = k_0$$

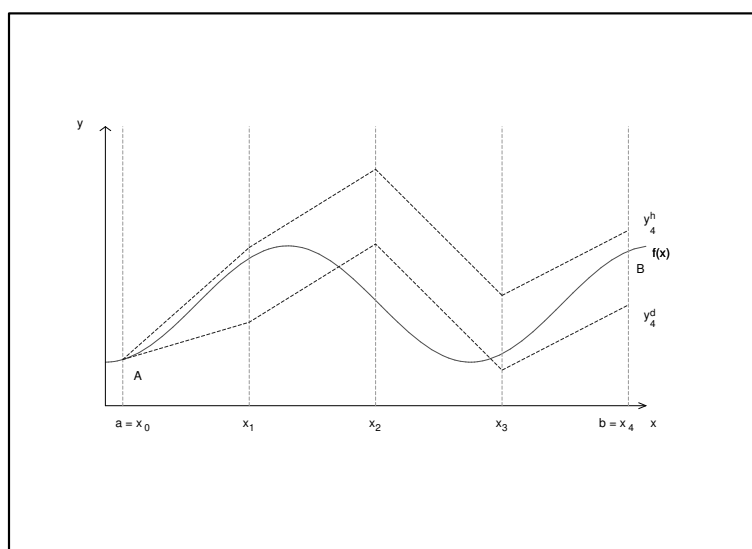
na intervalu $[a, b]$ s krokem $h = \frac{b-a}{n}$. Substitucí $y = u, y' = v$ získáme úlohu

$$u' = v(x),$$

$$v' = f(x, u(x), v(x)),$$

$$u(a) = A, v(a) = k_0.$$

Jde o soustavu diferenciálních rovnic prvního řádu, kterou už umíme řešit. Výsledně tak získáme diskrétní funkci $(x_i, y_i^0), i = 0, 1, \dots, n$. Jestliže $y_n^0 = y^0(b) = B$, zastavujeme výpočet. V opačném případě si volíme novou hodnotu $y'(a) = k_1$ a znovu provádíme výpočet. Výpočet a tedy i volbu $y'(a) = k_i$ opakujeme tak dlouho, až se hodnota y_n^i přiblíží k B s požadovanou přesností. Viz obr. 8.3



Obr. 8.3: Geometrický smysl metody střelby

Pro vlastní výpočet používáme některou z numerických metod pro řešení počátečních úloh. Často se používá zejména Rungeho-Kuttova metoda. Pro větší interval I a nebo pro větší počet uzlů je vhodnější použít Rungeho-Kuttovu metodu na počáteční přiblížení a pak použít přesnější metody typu prediktor-korektor.

Jakmile dosáhneme stavu, že pro volbu počátečních podmínek k_i, k_j platí

$$y_n^i < B < y_n^j,$$

potom můžeme použít pro další výběr hodnot k_r vhodnou numerickou metodu, např. bisekci. Výstupem metody střelby je diskrétní funkce $\{y_i \approx y(x_i)\}, i = 0, 1, \dots, n$. Kontrola přesnosti se provádí metodou polovičního kroku.

Touto metodou lze řešit lineární i nelineární diferenciální rovnice a taktéž soustavy rovnic.

8.1.2 Metoda konečných diferencí

S metodou konečných diferencí jste se setkali už v předmětu Matematika 3. Zde si ji zasadíme do širších souvislostí a odvodíme si potřebná tvrzení. Mějme danu úlohu

$$-y'' + \sigma(x)y = f(x), \quad (8.3)$$

$$y(a) = \alpha, \quad y(b) = \beta, \quad (8.4)$$

kde $a < b$, α, β jsou konstanty. Hledáme spojitou, dvakrát spojitě diferencovatelnou funkci $y(x)$, která splňuje (8.3), (8.4).

Věta 8.3. *Nechť $\sigma(x), f(x)$ jsou spojité na $[a, b]$, $\sigma(x) \geq 0$. Potom existuje právě jedno řešení úlohy (8.3), (8.4).*

Rovnici (8.3) můžeme řešit přímo analyticky, jestliže $\sigma(x)$ je konstanta. Předpokládejme, že platí $\sigma(x) \equiv 0$, potom dvojnásobnou integrací dostaneme řešení:

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

$$y' = \int -f(x)dx + C \quad \Rightarrow \quad y(x) = \int \left(\int -f(x)dx \right) dx + Cx + D.$$

Parametry C, D si určíme tak, aby byly splněny okrajové podmínky (8.4).

Nechť nyní je $\sigma(x) \in \mathbb{R}$, $\sigma > 0$, tj. σ je kladná konstanta. Potom charakteristická rovnice má tvar

$$-\lambda^2 + \sigma = 0,$$

$$\lambda_{1,2} = \pm\sqrt{\sigma}$$

a obecné řešení rovnice (8.3) je tvaru

$$y(x) = Ae^{\sqrt{\sigma}x} + Be^{-\sqrt{\sigma}x},$$

$A, B \in \mathbb{R}$. Parametry A, B si dále určíme tak, aby byly splněny okrajové podmínky (8.4). Ověříme si dále správnost našeho výpočtu:

$$y'(x) = Ae^{\sqrt{\sigma}x} \sqrt{\sigma} + Be^{-\sqrt{\sigma}x} (-\sqrt{\sigma}),$$

$$y''(x) = Ae^{\sqrt{\sigma}x} \sigma + Be^{-\sqrt{\sigma}x} \sqrt{\sigma} = \sigma \left(Ae^{\sqrt{\sigma}x} + Be^{-\sqrt{\sigma}x} \right) = \sigma y.$$

Pro $\sigma < 0$ dostaneme analogicky řešení

$$y(x) = Ce^{j\sqrt{\sigma}x} + De^{-j\sqrt{\sigma}x},$$

$j = \sqrt{-1}, C, D \in \mathbb{R}$. Parametry C, D si opět určíme tak, aby byly splněny okrajové podmínky (8.4).

Nechť nyní je $\sigma(x)$ funkcí. Budeme nyní hledat numerické řešení rovnice (8.3) metodou konečných diferencí. Vytvoříme si síť:

$$x_0 = a, x_i = x_0 + ih, i = 0, 1, \dots, n+1, h = \frac{b-a}{n+1}.$$

Body x_0, x_{n+1} jsou hraniční, zbývající body jsou vnitřní. Hodnoty funkce v hraničních bodech známe – jsou to hodnoty $y_0 = \alpha, y_{n+1} = \beta$. Zbývající hodnoty musíme dopočítat. **Pozor:** Zde máme jiné dělení intervalu (a, b) , než bylo používáno dříve. To proto, abyste předchozí tvar nepokládali za jediný možný.

Druhou derivaci funkce $y(x)$ nahradíme diferencí

$$-\frac{d^2y(x_i)}{dx^2} = -\frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} + \frac{h^2}{12}y^{(4)}(\xi),$$

kde $y_i = y(x_i), \xi \in [x_{i-1}, x_{i+1}]$. Je-li h dostatečně malé a $\max_{x \in I} |y^{(4)}(x)| < K < \infty$, můžeme zbytkový člen zanedbat. Protože jsme předpokládali jen spojitost funkce $y(x)$ včetně první a druhé derivace, plyne odtud i další požadavek na ohraničenost čtvrté derivace. Ovšem funkci $y(x)$ předem neznáme. Proto se o tomto požadavku předpokládá, že je splněný. Označme $\sigma(x_i) = \sigma_i, f(x_i) = f_i$. Pak dosazením do rovnice (8.3) dostaneme, při zanedbání zbytkového členu,

$$-\frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} + \sigma_i y_i = f_i, i = 1, 2, \dots, n$$

a po úpravě

$$-y_{i-1} + (2 + h^2\sigma_i)y_i - y_{i+1} = h^2f_i, i = 1, 2, \dots, n.$$

Dosazením hodnot $y_0 = \alpha, y_{n+1} = \beta$ dostaneme soustavu rovnic

$$\begin{array}{rccccccc} (2 + h^2\sigma_1)y_1 & & -y_2 & & & & = h^2f_1 + \alpha \\ -y_1 & + (2 + h^2\sigma_2)y_2 & & & -y_3 & & = h^2f_2 \\ & & -y_2 & & + (2 + h^2\sigma_3)y_3 & & -y_4 = h^2f_3 \\ & \ddots & & & & & \ddots \\ & & -y_{n-2} & + (2 + h^2\sigma_{n-1})y_{n-1} & & & -y_n = h^2f_{n-1} \\ & & & -y_{n-1} & + (2 + h^2\sigma_n)y_n & & = h^2f_n + \beta \end{array}$$

To je soustava lineárních algebraických rovnic s třídiagonální maticí, která je pro $\sigma > 0$ diagonálně dominantní a tedy regulární a navíc je i symetrická a pozitivně definitní. Existuje tedy jediné řešení této soustavy a iterační metody (Jacobiho i Gauss-Seidelova) konvergují k tomuto řešení.

Pro $\sigma \geq 0$ jde o symetrickou matici, která má minimálně první a poslední rovnici diagonálně ostře dominantní, u zbývajících rovnic platí neostrá diagonální dominantnost. I pro takovéto soustavy bude Gauss-Seidelova metoda konvergovat.

V případě $\sigma(x) \equiv 0$ máme rovnici $-y'' = f(x)$ a dvojitou integrací dostaneme hledané řešení přímo a podstatně rychleji.

Pro odhad přesnosti platí:

Věta 8.4. Necht' $\tilde{y}(x)$ je přesné řešení úlohy (8.3), (8.4) na intervalu $[a, b]$, kde $\sigma(x) \geq 0$. Necht' y_1, y_2, \dots, y_n jsou diskrétní aproximace řešení získané metodou konečných diferencí. Potom

$$|\tilde{y}(x_i) - y_i| \leq \frac{M_4 h^2}{24} (x_i - a)(b - x_i),$$

kde $x_i = a + ih = a + i \frac{b-a}{n+1}$, $0 \leq i \leq n+1$, $M_4 = \max_{x \in [a, b]} |f^{(4)}(x)|$. Dále platí

$$\max_{1 \leq i \leq n} |\tilde{y}(x_i) - y_i| \leq \frac{M_4 (b-a)^2}{96} h^2.$$

Pozor: Zde zase máme celkem $n+2$ bodů, přičemž n_0 a n_{n+1} jsou okrajové podmínky. Z věty 8.4 plyne, že při splnění předpokladů věty 8.3, pro dostatečně malý krok h , lze dosáhnout libovolné předepsané přesnosti. Neboli pro $h \rightarrow 0$ diskrétní aproximace konvergují k přesnému řešení.

Poznámka 8.5. Rovnici $y'' - \sigma(x)y = f(x)$, $\sigma \geq 0$ si substitucí $z = -y$ upravíme na tvar požadovaný větou 8.3 (a nebo vynásobením rovnice číslem (-1) obdržíme požadovaný tvar). Neboli — pokud $\sigma(x)$ nemění znaménko na intervalu I a znaménka u y'' a $\sigma(x)$ jsou opačná, je úloha (8.3), (8.4) jednoznačně řešitelná metodou konečných diferencí.

Příklad 8.6. Řešte metodou konečných diferencí s přesností ε okrajovou úlohu

$$-y'' + (1 + x^2)y = -1,$$

$$y(-1) = y(1) = 0.$$

Řešení. Máme $a = -1$, $b = 1$. Zvolme krok $h = \frac{1}{2}$. Sestavíme si soustavu rovnic

$$\begin{aligned} \left(2 + \frac{1}{4} \left(1 + \frac{1}{4}\right)\right) y_1 - y_2 &= -\frac{1}{4}, \\ -y_1 + \left(2 + \frac{1}{4}\right) y_2 - y_3 &= -\frac{1}{4}, \\ -y_2 + \left(2 + \frac{1}{4} \left(1 + \frac{1}{4}\right)\right) y_3 &= -\frac{1}{4}. \end{aligned}$$

A po úpravě

$$\begin{aligned} 37y_1 - 16y_2 &= -4 \\ -4y_1 + 9y_2 - 4y_3 &= -1 \\ -16y_2 + 37y_3 &= -4 \end{aligned}$$

Její řešení jsou hledané hodnoty:

$$y_1 = -0.25365, \quad y_2 = -0.33658, \quad y_3 = -0.25365.$$

Máme první přiblížení. Zmenšíme krok na polovinu a opakujeme výpočet. Jestliže je $\forall i \quad |y_{i,h} - y_{2i, \frac{h}{2}}| < \varepsilon$, pak zastavujeme výpočet. Neplatí-li tato podmínka, zmenšíme opět

krok na polovinu a opakujeme postup. Pokud použijeme Eukleidovskou normu, bude mít

podmínka tvar $\sqrt{\sum_{i=1}^n (y_{i,h} - y_{2i,\frac{h}{2}})^2} < \varepsilon$. □

V obecnějším případě budeme mít složitější tvar.

Definice 8.7. *Okrajová úloha* pro obyčejnou lineární diferenciální rovnici 2.řádu je: Určit takové řešení $y(x)$ rovnice

$$y'' + f_1(x)y' + f_2(x)y = f_3(x), \quad (8.5)$$

které na intervalu $I = [a, b]$ splňuje okrajové podmínky

$$\alpha_1 y(a) + \beta_1 y'(a) = \gamma_1, \quad \alpha_2 y(b) + \beta_2 y'(b) = \gamma_2, \quad |\alpha_i| + |\beta_i| > 0, \quad i = 1, 2. \quad (8.6)$$

Řešení hledáme jako spojitou funkci, která je na intervalu $[a, b]$ spojitě dvakrát diferencovatelná.

Podmínky (8.6) se většinou označují jako *Sturmovy* podmínky. Jejich speciálním případem jsou *Dirichletovy*¹ podmínky

$$y(a) = \alpha, \quad y(b) = \beta,$$

a nebo *Neumannovy*² podmínky

$$y'(a) = \alpha, \quad y'(b) = \beta.$$

Věta 8.8. *Nechť je dána rovnice (8.5) a nechť $f_1 \in C(I)$. Potom lze tuto rovnici převést na samoadjungovaný tvar*

$$-(p(x)y')' + q(x)y = f(x) \quad (8.7)$$

kde

$$p(x) = e^{\int f_1(x)dx}, \quad q(x) = -f_2(x)p(x), \quad f(x) = -f_3(x)p(x).$$

Důkaz. Podle předpokladu je $f_1 \in C(I)$, potom tedy existuje $(\int f_1(x)dx)$ pro $\forall x \in I$. Označme $p(x) = \exp(\int f_1(x)dx)$. Rovnici (8.5) vynásobíme $(-p(x))$. Dostaneme

$$-y''p(x) - f_1(x)p(x)y' - f_2(x)p(x)y = -f_3(x)p(x). \quad (8.8)$$

Dále je

$$y''p(x) + f_1(x)p(x)y' = (p(x)y')',$$

neboť

$$(p(x)y')' = p'(x)y' + p(x)y'',$$

$$p'(x) = \left(e^{\int f_1(x)dx}\right)' = \left(e^{\int f_1(x)dx}\right) f_1(x) = p(x)f_1(x).$$

Dosazením tohoto výsledku do (8.8) a použitím označení p, q, f definovaných ve větě 8.8 dostaneme rovnici (8.7). □

¹**P. G. L. Dirichlet** (1805 – 1859) německý matematik. Pracoval v teorii čísel, matematické analýze, matematické fyzice. Ve všech těchto oborech dosáhl významných výsledků.

²**K. G. Neumann** (1832 – 1925) německý matematik. Věnoval se hlavně teorii logaritmického potenciálu, diferenciálními rovnicemi, matematické fyzice.

Věta 8.9. *Nechť máme lineární diferenciální rovnici v samoadjungovaném tvaru (8.7). Nechť dále platí*

$$q, f \in C(I), \quad p \in C^1(I), \\ p(x) > 0, \quad q(x) \geq 0 \quad \forall x \in I.$$

Potom existuje právě jedno řešení $u(x)$ diferenciální rovnice (8.7), které splňuje okrajové podmínky $u(a) = \alpha$, $u(b) = \beta$, kde α, β jsou libovolná reálná čísla.

Na základě této věty platí i postup řešení, který jsme si předvedli jako první, neboť rovnice (8.3) je speciálním případem rovnice (8.5) a tedy i speciálním případem rovnice (8.7).¹

Numerické řešení opět budeme hledat metodou konečných diferencí. Předpokládáme přitom, že jsou splněny předpoklady věty 8.9, které nám zajišťují existenci a jednoznačnost řešení. Označme

$$p_{i \pm \frac{1}{2}} = p\left(x_i \pm \frac{h}{2}\right), \\ q_i = q(x_i), \quad f_i = f(x_i), \quad h = \frac{b-a}{n}, \quad i = 0, 1, \dots, n.$$

Výraz $(p(x)y)'$ nahradíme výrazem

$$(p(x)y)' \approx \frac{1}{h} \left(p_{i+\frac{1}{2}} \frac{y_{i+1} - y_i}{h} - p_{i-\frac{1}{2}} \frac{y_i - y_{i-1}}{h} \right).$$

Po dosazení do (8.7) dostaneme

$$-\frac{1}{h^2} \left(p_{i+\frac{1}{2}}(y_{i+1} - y_i) - p_{i-\frac{1}{2}}(y_i - y_{i-1}) \right) + q_i y_i = f_i$$

a po úpravě máme soustavu lineárních algebraických rovnic

$$-p_{i-\frac{1}{2}} y_{i-1} + (p_{i-\frac{1}{2}} + p_{i+\frac{1}{2}} + h^2 q_i) y_i - p_{i+\frac{1}{2}} y_{i+1} = h^2 f_i, \quad i = 1, 2, \dots, n-1.$$

¹Mějme rovnici

$$(-py)' + qy = f$$

a okrajové podmínky

$$y(x_0) = y_0, \quad y(x_n) = y_n.$$

V případě, že p je konstanta, máme rovnici

$$-py'' + qy = f.$$

Vydělením celé rovnice p dostaneme stejný tvar jako má rovnice (8.3). Jestliže budeme počítat přímo a nahradíme derivaci diferencí, potom dostaneme, při stejném označení jako dříve,

$$-p \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} + q_i y_i = f_i$$

a po úpravě

$$-py_{i+1} + (2p + h^2 q_i) y_i - py_{i-1} = h^2 f_i.$$

Máme soustavu s reálnou symetrickou maticí. Jestliže je $p > 0$ a $q \geq 0$, potom jde o matici diagonálně dominantní a proto je soustava s ní jednoznačně řešitelná.

Je to soustava lineárních algebraických rovnic s třídiagonální symetrickou maticí koeficientů. Dá se ukázat, že tato matice je regulární: Podle podmínek věty 8.9 je $p > 0$, $q \geq 0$. Matice je diagonálně dominantní pro $q > 0$, a proto podle věty 4.54 je regulární a existuje jediné řešení této soustavy. Bude k němu konvergovat Jacobiho iterační metoda i Gauss-Seidelova iterační metoda.

Pro $q \geq 0$ jde o symetrickou matici, která má minimálně první a poslední rovnici diagonálně ostře dominantní, u zbývajících platí neostrá diagonální dominantnost. I pro takovéto soustavy bude Gauss-Seidelova metoda konvergovat.

V případě $q \equiv 0$ máme rovnici $-(p(x)y')' = f(x)$ a dvojitou integrací dostaneme hledané řešení.

$$\begin{aligned} -(p(x)y')' &= f(x), \\ -p(x)y' &= \int f(x)dx + K, \end{aligned}$$

protože $p(x) > 0$ pro všechna x z I ,

$$\begin{aligned} y' &= \frac{-1}{p(x)} \int f(x)dx + K, \\ y &= \int \left(\frac{-1}{p(x)} \int f(x)dx + K \right) dx + C. \end{aligned}$$

Příklad 8.10. Metodou konečných diferencí s krokem $h = 0.25$ řešte na intervalu $I = [1; 2]$ úlohu

$$\begin{aligned} y'' - \frac{2}{x}y' - \frac{4}{x^2}y &= \frac{6}{x}, \\ y(1) = 0, \quad y(2) &= -1.5. \end{aligned}$$

Řešení. Funkce $f_1(x) = -2/x$ je na I spojitá. Proto podle věty 8.8 si rovnici převedeme na samoadjungovaný tvar, kde

$$p(x) = \exp\left(\int -\frac{2}{x}dx\right) = \exp\left(-2 \int \frac{dx}{x}\right) = \exp(-2 \ln x) = \exp(\ln x^{-2}) = x^{-2} = \frac{1}{x^2}.$$

Dostáváme

$$-\left(\frac{y'}{x^2}\right)' + \frac{4}{x^4}y = -\frac{6}{x^3}. \quad (8.9)$$

Funkce $p = x^{-2}$, $p' = -2x^{-3}$, $q = 4x^{-4}$, $f = -6x^{-3}$ jsou spojitě na I , $p > 0$, $q > 0$, jsou tedy splněny všechny podmínky věty 8.9 a řešení této úlohy existuje a je právě jedno.

Po dosazení do (8.9) dostaneme soustavu lineárních algebraických rovnic

$$\begin{aligned} 1.4214y_1 - 0.5289y_2 &= -0.192, \\ -0.5289y_1 + 0.987y_2 - 0.37869y_3 &= -0.11111, \\ -0.37869y_2 + 0.68979y_3 &= -0.496. \end{aligned}$$

Jejím řešením je $y_1 = -0.441$, $y_2 = -0.82349$, $y_3 = -1.1712$.

Tento numerický výsledek si můžete srovnat s analytickým řešením

$$y^*(x) = \frac{1}{x} - x,$$

potom je $y_1^* = -0.45$, $y_2^* = -0.83\bar{3}$, $y_3^* = -1.178$.

Dosáhli jsme přesnosti 10^{-1} . Provedeme zmenšení kroku a výpočet opakujeme do té doby, než dosáhneme požadované přesnosti. T.j. pokud se hodnoty v uzlových bodech neustálí. \square

Mějme obecný případ *okrajové úlohy* pro obyčejnou lineární diferenciální rovnici 2.řádu:

$$a_0(x)y'' + a_1(x)y' + a_2(x)y = f(x), \quad (8.10)$$

kde $x \in [a, b]$, a_0, a_1, a_2, f jsou spojité funkce na $[a, b]$ a $a_0(x) \neq 0$ pro všechna x z intervalu $[a, b]$. Hledáme takové řešení rovnice (8.10), které na intervalu $I = [a, b]$ splňuje okrajové podmínky

$$\alpha_1 y(a) + \alpha_2 y'(a) = d_0, \quad \beta_1 y(b) + \beta_2 y'(b) = d_1, \quad (8.11)$$

kde $\alpha_1, \alpha_2, \beta_1, \beta_2$ jsou daná reálná čísla, $|\alpha_1| + |\alpha_2| \neq 0$, $|\beta_1| + |\beta_2| \neq 0$. Řešení hledáme jako spojitou funkci, která je na intervalu $[a, b]$ spojitě dvakrát diferencovatelná.

Postupujeme analogicky jako v předchozích případech. Označme $a_i(x_j) = a_{ij}$, $y(x_j) = y_j$, $h = \frac{b-a}{n}$. Derivace nahradíme diferencemi

$$y''(x_j) \approx \frac{y_{j+1} - 2y_j + y_{j-1}}{h^2},$$

$$y'(x_j) \approx \frac{y_{j+1} - y_{j-1}}{2h}.$$

Dosazením do rovnice (8.10) dostaneme soustavu rovnic

$$\frac{a_{0j}}{h^2} (y_{j+1} - 2y_j + y_{j-1}) + \frac{a_{1j}}{2h} (y_{j+1} - y_{j-1}) + a_{2j}y_j = f_j$$

a po úpravě

$$y_{j-1} \left(\frac{a_{0j}}{h^2} - \frac{a_{1j}}{2h} \right) + y_j \left(-2\frac{a_{0j}}{h^2} + a_{2j} \right) + y_{j+1} \left(\frac{a_{0j}}{h^2} + \frac{a_{1j}}{2h} \right) = f_j. \quad (8.12)$$

Pokud odstraníme zlomky, dostaneme

$$y_{j-1} (2a_{0j} - ha_{1j}) + y_j (-4a_{0j} + 2h^2 a_{2j}) + y_{j+1} (2a_{0j} + ha_{1j}) = 2h^2 f_j. \quad (8.13)$$

V okrajových podmínkách (8.11) nahradíme derivaci v krajních bodech intervalu diferencí

$$y'(x_0) \approx \frac{1}{2h} (-3y_0 + 4y_1 - y_2),$$

$$y'(x_n) \approx \frac{1}{2h} (3y_n - 4y_{n-1} + y_{n-2}).$$

Dosazením do počátečních podmínek dostaneme

$$\alpha_1 y_0 + \frac{\alpha_2}{2h} (-3y_0 + 4y_1 - y_2) = d_0, \quad (8.14)$$

$$\beta_1 y_n + \frac{\beta_2}{2h} (3y_n - 4y_{n-1} + y_{n-2}) = d_1. \quad (8.15)$$

Rovnice (8.12), (8.14), (8.15) a nebo (8.13), (8.14), (8.15) tvoří soustavu $n + 1$ rovnic o $n + 1$ neznámých y_0, y_1, \dots, y_n . Jejím vyřešením získáme diskrétní aproximaci úlohy (8.10), (8.11).

Věta 8.11. *Nechť a_0, a_1, a_2, f jsou spojité funkce na intervalu $[a, b]$ takové, že platí*

$$a_0(x) \geq c > 0, a_2(x) \leq 0$$

na $[a, b]$ a dále $|\alpha_1| + |\alpha_2| > 0, |\beta_1| + |\beta_2| > 0$. Potom má úloha (8.10), (8.11) právě jedno řešení a diskrétní aproximace získaná pomocí rovnic (8.12), (8.14), (8.15), respektive (8.13), (8.14), (8.15) konverguje k tomuto řešení.

Příklad 8.12. Metodou konečných diferencí řešte úlohu

$$y'' - y = x,$$

$$y(0) = 1,$$

$$2y(1) + 3y'(1) = -1.$$

Řešení. Volme $n = 4$, potom $h = 0.25$ a máme uzlové body $x_0 = 0, x_1 = 0.25, x_2 = 0.5, x_3 = 0.75, x_4 = 1$. Máme $a_0 = 1, a_1 = 0, a_2 = -1, f = x$, jsou tedy splněny požadavky věty 8.11 a po dosazení do (8.12) dostaneme soustavu

$$\begin{array}{rcl} -y_0 & +2.0625y_1 & -y_2 & = -0.015625, \\ & -y_1 & +2.0625y_2 & -y_3 & = -0.031250, \\ & & -y_2 & +2.0625y_3 & -y_4 & = -0.046875. \end{array}$$

Z první okrajové podmínky vyplývá, že

$$y_0 = 1$$

a z druhé okrajové podmínky vyplývá, po dosazení do (8.15), rovnice

$$2y_4 + 6(3y_4 - 4y_3 + y_2) = -1.$$

Po úpravě dostaneme soustavu

$$\left(\begin{array}{cccc|c} 2.065 & -1 & 0 & 0 & 0.984375 \\ -1 & 2.0625 & -1 & 0 & -0.031250 \\ 0 & -1 & 2.0625 & -1 & -0.046875 \\ 0 & 0 & -11.6250 & 14 & -1.281250 \end{array} \right),$$

s třídiagonální maticí koeficientů, která je diagonálně dominantní. Jejím řešením je

$$\begin{aligned}y_0 &= 1, \\y_1 &= 0.0688479, \\y_2 &= 0.435612, \\y_3 &= 0.241121, \\y_4 &= 0.108782.\end{aligned}$$

Pokud porovnáme tuto aproximaci s analytickým řešením

$$y = \frac{4e + 1}{5e^2 + 1}e^x + \frac{5e^2 - 4e}{5e^2 + 1}e^{-x} - x, \quad x \in [0, 1],$$

zjistíme, že chyba je $|\varepsilon| \leq 0.00546$. Vzhledem k velikosti kroku jde o velmi dobrou aproximaci. \square

Příklad 8.13. Řešte rovnici

$$-((1+x)y')' + (2-x)y = x^4 - 2x^3 + 8x^2 + 8x + 1$$

s okrajovými podmínkami

$$\begin{aligned}y'(0) &= 2y(0) + 1, \\y(1) &= 0.\end{aligned}$$

Řešení. Naše úloha má analytické řešení $\tilde{y} = x - x^3$, které můžeme použít pro porovnání s numerickým řešením.

Numerické řešení budeme hledat pomocí metody konečných diferencí. Volme $n = 2^k$, $k = 1, 2, \dots$, potom pro $k = 1$ máme $n = 2$ a dostaneme soustavu

$$\begin{pmatrix} 10 & -5 \\ -5 & 13.5 \end{pmatrix} \cdot \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} = \begin{pmatrix} -2.5 \\ 4.8125 \end{pmatrix}.$$

Hodnotu $y_2 = y(1)$ známe. Řešením naší soustavy je

$$\begin{aligned}y_0 &= -0,088068, \\y_1 &= 0,323864.\end{aligned}$$

Pro $k = 2$ máme $n = 4$ a dostaneme soustavu

$$\begin{pmatrix} 27 & -18 & 0 & 0 \\ -18 & 41.75 & -22 & 0 \\ 0 & -22 & 49.5 & -26 \\ 0 & 0 & -26 & 57,25 \end{pmatrix} \cdot \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -4.5 \\ 1.472656 \\ 4.8125 \\ 8.972656 \end{pmatrix}.$$

Řešením naší soustavy je

$$y_0 = -0,022, \quad y_1 = 0,216453, \quad y_2 = 0.362128, \quad y_3 = 0.321188.$$

Hodnotu $y_4 = y(1)$ známe. A můžeme stejným způsobem pokračovat dále. \square

V aplikacích se objevují i obecnější rovnice, např.

$$-(p(x)y'(x))' + r(x)y'(x) + g(x)y(x) = f(x). \quad (8.16)$$

Tato rovnice popisuje mezi jiným i transport chemické příměsi v kapalině a proto se někdy označuje jako *difuzní rovnice*.

Člen ry' , který je zde navíc, ve srovnání s předchozím typem rovnic, aproximujeme výrazem

$$r(x_i)y'(x_i) = r_i \frac{y_{i+1} - y_{i-1}}{2h} + O(h^2).$$

Po dosazení do (8.16) rovnice dostaneme soustavu

$$-\left(p_{i-\frac{1}{2}} + \frac{1}{2}hr_i\right)y_{i-1} + \left(p_{i-\frac{1}{2}} + p_{i+\frac{1}{2}} + h^2g_i\right)y_i - \left(p_{i+\frac{1}{2}} - \frac{1}{2}hr_i\right)y_{i+1} = h^2f_i.$$

Podmínky řešitelnosti rovnice (8.16) se opět převádí na podmínku řešitelnosti soustavy lineárních algebraických rovnic.

Způsob provedení diskretizace přímo ovlivňuje řešení, viz následující příklad.

Příklad 8.14. Pro $x \in [0, l]$ řešme rovnici

$$-pu'' + ru' = 0,$$

s okrajovými podmínkami

$$u(0) = \alpha, \quad u(l) = \beta,$$

kde $p > 0, r \neq 0, \alpha \neq \beta$ jsou konstanty.

Řešení. Naše úloha má přesné řešení

$$u(x) = \alpha + (\beta - \alpha) \frac{1 - e^{\frac{rx}{p}}}{1 - e^{\frac{rl}{p}}}.$$

Jde to monotónní funkci, která je pro $\alpha < \beta$ rostoucí a pro $\beta < \alpha$ klesající.

Při hledání numerického řešení nahradíme derivace centrálními diferencemi a dostaneme

$$-p \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + r \frac{u_{i+1} - u_{i-1}}{2h} = 0.$$

Označme $\lambda = \frac{rh}{p}$. Potom po úpravě dostaneme

$$-\left(1 + \frac{1}{2}\lambda\right)u_{i-1} + 2u_i - \left(1 - \frac{1}{2}\lambda\right)u_{i+1} = 0.$$

Pro $\lambda = 2$ dostaneme řešení

$$u_i = u_{i-1},$$

neboli

$$\begin{aligned} u_i &= \alpha \quad \text{pro } i < n, \\ u_n &= \beta. \end{aligned}$$

Pro $\lambda \neq 2$ dostaneme řešení

$$u_i = A + B \left(\frac{1 + \frac{1}{2}\lambda}{1 - \frac{1}{2}\lambda} \right)^i,$$

kde konstanty A, B určíme z okrajových podmínek.

Pokud použijeme pro náhradu první derivace diferenci vzad, neboli

$$u' = \frac{u_i - u_{i-1}}{h},$$

potom dostaneme pro $r > 0$ (a tedy i $\lambda > 0$, protože podle předpokladů je $p > 0$ a h je velikost kroku, která je také kladná)

$$-p \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + r \frac{u_i - u_{i-1}}{h} = 0$$

a po úpravě dostaneme

$$-(1 + \lambda)u_{i-1} + (2 + \lambda)u_i - u_{i+1} = 0.$$

Tato diferenční rovnice má řešení

$$u_i = C + D(1 + \lambda)^i,$$

ktelé je ryze monotonní. Konstanty C, D určíme z okrajových podmínek.

V případě, že máme $r < 0$, postupujeme obdobně. První derivaci nahradíme diferencí vpřed, tj.

$$u' = \frac{u_{i+1} - u_i}{h}$$

a dostaneme

$$-p \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + r \frac{u_{i+1} - u_i}{h} = 0.$$

V tomto případě je $\lambda < 0$, (protože podle předpokladů je $p > 0$ a h je velikost kroku, která je také kladná). Po úpravě potom dostaneme

$$-u_{i-1} + (2 - \lambda)u_i - (1 - \lambda)u_{i+1} = 0.$$

Tato diferenční rovnice má řešení

$$u_i = E + F(1 - \lambda)^i,$$

ktelé je opět ryze monotonní. Konstanty E, F určíme z okrajových podmínek. □

Poznámka 8.15. První postup je použitelný jen pro krok, jehož délka splňuje podmínku $|\lambda| < 2 \Rightarrow h < \frac{2p}{|r|}$.

8.2 Metoda konečných objemů

Jinou metodou pro řešení okrajových úloh je metoda konečných objemů.

Hledáme řešení rovnice

$$(-pu')' + qu = f, \quad (8.17)$$

kde $p > 0, q \neq 0$, na intervalu $[0, l]$ s krokem h . Opět budeme hledat řešení v jednotlivých uzlových bodech, kde máme $x_i = x_0 + ih = ih$, protože $x_0 = 0$ a jako obvykle $h = \frac{l}{n}$.

Ke každému i -tému kroku přiřadíme konečný objem B_i takto

a) pro vnitřní uzly $B_i = \langle x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}} \rangle, \quad i = 1, 2, \dots, n-1.$

b) pro hraniční uzly $B_0 = \langle 0, x_{\frac{1}{2}} \rangle, B_n = \langle x_{n-\frac{1}{2}}, x_n \rangle$. Potom

$$\langle 0, l \rangle = \bigcup_{i=0}^n B_i.$$

Vezmeme si naši rovnici (8.17) a budeme ji integrovat přes B_i .

$$\int_{B_i} [(-pu')' + qu] dx = \int_{B_i} f dx.$$

Předpokládejme nejdříve, že B_i přísluší vnitřnímu uzlu. Potom dostaneme

$$-pu' \Big|_{x=x_{i-\frac{1}{2}}}^{x=x_{i+\frac{1}{2}}} + \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} qu dx = \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} f dx,$$

$$-pu'(x_{i+\frac{1}{2}}) - u'(x_{i-\frac{1}{2}}) + \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} qu dx = \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} f dx.$$

Derivace nahradíme diferencemi

$$u'(x_i - \frac{1}{2}) = \frac{u_i - u_{i-1}}{h} + O(h^2),$$

$$u'(x_i + \frac{1}{2}) = \frac{u_{i+1} - u_i}{h} + O(h^2).$$

Hodnoty obou integrálů odhadneme pomocí obdélníkového pravidla, přitom budeme vždy vycházet ze středu intervalu. Neboli

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} qu dx = hq_i u_i + O(h^3),$$

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} f dx = hf_i + O(h^3).$$

Dosadíme a po zanedbání chybových funkcí dostaneme po úpravě rovnici

$$-\frac{p_{i-\frac{1}{2}}}{h}u_{i-1} + \left(\frac{p_{i-\frac{1}{2}}}{h} + \frac{p_{i+\frac{1}{2}}}{h} + hq_i\right)u_i - \frac{p_{i-\frac{1}{2}}}{h}u_{i+1} = hf_i.$$

Nechť máme zadané Neumannovy okrajové podmínky v bodech $x = 0$ a $x = l$. Potom pro B_0 máme

$$-pu'|_{x=x_0} + \int_{x_0}^{x_{\frac{1}{2}}} qu dx = \int_{x_0}^{x_{\frac{1}{2}}} f dx.$$

Derivaci v bodě $x_{\frac{1}{2}}$ nahradíme diferencí

$$u'(x_{\frac{1}{2}}) = \frac{u_1 - u_0}{h} + O(h^2)$$

dosadíme za výraz $p(x_0)u'(x_0) \equiv p(0)u(0)$ podle okrajové podmínky a integrály počítáme podle obdelníkového pravidla přičemž budeme vycházet z levého konce intervalu, tj.

$$\int_{x_0}^{x_{\frac{1}{2}}} qu dx = \frac{1}{2}hq_0u_0 + O(h^2),$$

$$\int_{x_0}^{x_{\frac{1}{2}}} f dx = \frac{1}{2}hf_0 + O(h^2).$$

Analogicky pro bod B_n , kde použijeme obdelníkové pravidlo a vycházíme z pravého konce intervalu.

Nakonec dostáváme soustavu lineárních algebraických rovnic, která je (vzhledem k podmínkám) jednoznačně řešitelná.

Můžeme použít i jiný postup, jehož základem je integrace podle lichoběžníkového pravidla. Vezměme si opět naši rovnici (8.17), budeme ji integrovat na intervalu $\langle x_i, x_{i+1} \rangle$ a dostaneme

$$\begin{aligned} -pu'|_{x_i} + \int_{x_i}^{x_{i+1}} qu dx &= \int_{x_i}^{x_{i+1}} f dx, \\ -p_{i+1}u'_{i+1} + p_iu'_i + \int_{x_i}^{x_{i+1}} qu dx &= \int_{x_i}^{x_{i+1}} f dx. \end{aligned}$$

Derivace nahradíme diferencemi, tj.

$$u'_{i+1} = \frac{u_{i+1} - u_i}{h} + O(h^2),$$

$$u'_i = \frac{u_i - u_{i-1}}{h} + O(h^2)$$

a integrály určíme podle lichoběžníkového pravidla

$$\int_{x_i}^{x_{i+1}} qu dx = \frac{q_{i+1}u_{i+1} + q_iu_i}{2} \cdot h + O(h^2),$$

$$\int_{x_i}^{x_{i+1}} f dx = \frac{f_{i+1} + f_i}{2} \cdot h + O(h^2).$$

Po dosazení dostaneme

$$-p_{i+1} \left(\frac{u_{i+1} - u_i}{h} \right) + p_i \left(\frac{u_i - u_{i-1}}{h} \right) + \frac{q_{i+1}u_{i+1} + q_i u_i}{2} \cdot h = \frac{h}{2} (f_{i+1} + f_i).$$

Po úpravě dostáváme

$$\left(-\frac{p_{i+1}}{h} + \frac{hq_{i+1}}{2} \right) u_{i+1} + \left(\frac{p_{i+1}}{h} + \frac{p_i}{h} + \frac{hq_i}{2} \right) u_i + \left(-\frac{p_i}{h} \right) u_{i-1} = \frac{h}{2} (f_{i+1} + f_i).$$

Opět máme soustavu lineárních algebraických rovnic. Protože požaduje jednoznačnou řešitelnost, dostáváme podmínku $|q_i| \geq |q_{i+1}| \forall i$. Pouze v tomto případě můžeme garantovat jednoznačnou řešitelnost.

Pojmy k zapamatování

- Zabývali jsme se hledáním řešení okrajových úloh pro diferenciální rovnice řádu aspoň dvě. Vysvětlili jsme si, že pouze pro ně má smysl hledat řešení okrajové úlohy.
- Ukázali jsme si dva z možných přístupů.
- Okrajovou úlohu lze převést na počáteční a hledat řešení počáteční úlohy některou z dříve probraných metod - to je princip metody střelby.
- Nebo lze na základě diskretizace proměnných převést okrajovou úlohu na řešení soustavy lineárních algebraických rovnic - to je princip metody konečných diferencí a metody konečných objemů.
- Formulace úlohy a použití hraničních bodů nám zaručuje jednoznačnou řešitelnost soustavy. Nalezené diskrétní řešení v limitě konverguje k přesnému řešení.

Kontrolní otázky

1. Co rozumíme pojmem okrajová úloha?
2. Proč se nezabýváme řešením okrajových úloh pro rovnice prvního řádu?
3. Je řešitelná každá okrajová úloha? Nutno zdůvodnit.

8.3 Cvičení

Maplety

Příklady pro samostatnou práci jsou uvedeny v samostatné sbírce příkladů. Pomocí následujících mapletů si můžete usnadnit některé dílčí výpočty, nebo zkontrolovat jejich správnost.

1. Výpočet determinantu matice
2. Analytické řešení soustavy lineárních rovnic
3. Výpočet funkčních hodnot funkcí jedné proměnné
4. Parciální derivace

9 Metoda konečných prvků.

Průvodce studiem

Cílem této kapitoly je seznámit čtenáře s další numerickou metodou pro nalezení řešení okrajové úlohy pro obyčejnou diferenciální rovnici.

Metoda konečných prvků je v současnosti velmi často používána při řešení celé řady nejrůznějších úloh v řadě aplikačních oblastí.

*Na rozvoji metody konečných prvků se podílela celá řada vynikajících matematiků. Pro nás je důležité, že jako první dokázal konvergenci metody konečných prvků v roce 1968 prof. RNDr. Miloš Zlámal, DrSc. pracovník VUT v Brně, který svůj výsledek publikoval v *Math. Nachrichten*, a současně s ním prof. RNDr. Alexander Ženíšek, DrSc. (emeritní profesor na FSI VUT), který však svůj výsledek publikoval v češtině a proto je celosvětově přiznáváno prvenství prof. Zlámalovi.*

Metoda konečných prvků se používá takřka výhradně pro řešení úloh s parciálními diferenciálními rovnicemi. Při použití na řešení okrajové úlohy pro obyčejné diferenciální rovnice dává, při splnění podmínek konvergence, stejný výsledek jako metoda konečných diferencí. My se jí ale budeme přesto zabývat, protože v jednodimenzionálním případě je základní princip metody snáze pochopitelný.

Cíle

Po prostudování této kapitoly budete schopni:

- Chápat podstatu metody konečných prvků.
- Sestrojit vhodné bázové funkce.

9.1 Základní pojmy

Metoda vznikla při řešení okrajových úloh rovinné pružnosti. Její princip je, že se rovinná oblast rozdělí na vhodné části, nazývané *konečné prvky* – obvykle to bývají nepřekrývající se trojúhelníky, a celá oblast se pak chápe jako konečný systém prvků, které na sebe vzájemně působí. V této podobě se metoda konečných prvků objevila v inženýrské praxi. Brzy se ukázalo, že jde o variantu *Rietzovy metody*, ale název jí už zůstal původní.

Např. celková energie soustavy, která je popsána parciální diferenciální okrajovou úlohou eliptického typu, je dána integrálem, který se proto nazývá energetický. Řešení úlohy pak má

tu vlastnost, že tomuto integrálu dává nejmenší hodnotu. Vidíme tedy, že původní okrajová úloha je ekvivalentní s určením takové funkce, která by minimalizovala zmíněný funkcionál. Jde o *variační formulaci úlohy*.

V případě funkcí jedné proměnné se jedná o nalezení funkce y splňující určité okrajové podmínky a pro níž je funkcionál

$$F(y) = \int_a^b g(x, y, y') dx \quad (9.1)$$

minimální. K přibližnému řešení se užívá Rietzova metoda. Řešení hledáme ve tvaru

$$v(x) = \varphi(x) + \sum_{j=1}^n \alpha_j \cdot b_j(x),$$

kde φ splňuje okrajové podmínky, $\varphi(a) = \alpha$, $\varphi(b) = \beta$ a kde b_j jsou lineárně nezávislé (bázové) funkce splňující homogenní okrajové podmínky (tj. jsou nulové na dané hranici). Protože φ splňuje okrajové podmínky a funkce b_j jsou na hranicích nulové, tak i funkce v splňuje okrajové podmínky. Po dosazení do funkcionálu (9.1) dostaneme

$$F(v) = \Phi(\alpha_1, \dots, \alpha_n),$$

tedy funkci n proměnných. Hledáme nyní její minimum metodami matematické analýzy.

Je jasné, že obecně neplatí $v(x) = y(x)$, protože $v(x)$ má speciální tvar určený jejími bázovými funkcemi. Můžeme tedy pouze očekávat $v(x) \approx y(x)$. Přesnost je přitom přímo závislá na volbě bázových funkcí a na jejich počtu. Metodu konečných prvků dostaneme speciální volbou bázových funkcí.

Mějme úlohu na intervalu $[a, b]$, $a < b$,

$$-y'' + \sigma(x)y = f(x), \quad \sigma(x) \geq 0, \quad (9.2)$$

$$y(a) = \alpha, \quad y(b) = \beta. \quad (9.3)$$

Existenci a jednoznačnost řešení jsme si ukázali už dříve.

Nejdříve převedeme tuto okrajovou úlohu na variační úlohu:

Nechť $w(x)$ je libovolná funkce, která má po částech spojitou první derivaci a splňuje okrajové podmínky, t.j. $w(a) = \alpha$, $w(b) = \beta$. Sestrojíme si funkcionál

$$F(w) \equiv \frac{1}{2} \int_a^b ([w'(x)]^2 + \sigma(x)w^2(x) - 2w(x)f(x)) dx \quad (9.4)$$

a ukážeme, že funkce, která jej minimalizuje, je řešením naší úlohy (9.2), (9.3).

Nechť $w(x) = y(x) + \varepsilon(x)$, kde $y(x)$ je řešením okrajové úlohy (9.2) a $\forall x \in I : \varepsilon(x) \geq 0$. Potom po dosazení máme po úpravě

$$F(w) = F(y) + \int_a^b (\varepsilon'y' + \varepsilon\sigma y - \varepsilon f) dx + \frac{1}{2} \int_a^b ([\varepsilon']^2 + \sigma\varepsilon^2) dx. \quad (9.5)$$

První člen prvního integrálu integrujeme "per partes"

$$\int_a^b \varepsilon'y' dx = \left| \begin{array}{ll} u = y' & u' = y'' \\ v' = \varepsilon' & v = \varepsilon \end{array} \right| = \varepsilon y' \Big|_a^b - \int_a^b \varepsilon y'' dx =$$

$$= \varepsilon(b)y'(b) - \varepsilon(a)y'(a) - \int_a^b \varepsilon y'' dx.$$

Takže celý první integrál z (9.5) si přepíšeme na tvar

$$\varepsilon(b)y'(b) - \varepsilon(a)y'(a) + \int_a^b \varepsilon(-y'' + \sigma y - f) dx$$

a celý tento výraz je roven nule, neboť y je řešením úlohy (9.2), (9.3) a $\varepsilon(a) = \varepsilon(b) = 0$, protože $\varepsilon = w - y$ a funkce w a y splňují tytéž okrajové podmínky. Takže z (9.5) máme

$$F(w) = F(y) + \frac{1}{2} \int_a^b ([\varepsilon'(x)]^2 + \sigma \varepsilon^2) dx \geq F(y),$$

protože pod integrálem je nezáporná funkce. Takže y skutečně minimalizuje funkcionál F v množině dostatečně hladkých funkcí $w(x)$ splňujících tytéž okrajové podmínky.

Omezíme se při hledání minima v funkce F na množinu spojitých, po částech lineárních funkcí a na případ ekvidistantních uzlů

$$x_j = x_0 + jh, \quad h = \frac{b-a}{n+1}, \quad v_0 = \alpha, \quad v_{n+1} = \beta, \quad v_j = v(x_j).$$

Protože lomená čára je jednoznačně určena svými vrcholy, budeme hledat funkci v ve tvaru

$$v(x) = \sum_{j=0}^n v_j \cdot b_j(x),$$

kde v_j jsou konstanty a báze funkce $b_j(x)$ jsou dány předpisem

$$b_j(x) = \begin{cases} \frac{1}{h}(x - x_{j-1}), & x \in [x_{j-1}, x_j], \\ -\frac{1}{h}(x - x_{j+1}), & x \in [x_j, x_{j+1}], \\ 0 & \text{jinak.} \end{cases}$$

Z této definice plyne, že $b_j(x_j) = 1$, $b_j(x_k) = 0 \forall k \neq j$. Zvláště je důležité, že funkce b_j má velmi malý nosič – tj. je nenulová jen na úzkém intervalu (x_{j-1}, x_{j+1}) .

V případě $b_0(x), b_{n+1}(x)$ bereme do úvahy pouze tu část, která leží uvnitř intervalu $[a, b]$. Protože derivace $v'(x)$ je po částech konstantní $v'(x) = \frac{1}{h}(v_{j+1} - v_j)$, $x \in [x_j, x_{j+1}]$, tak

$$\begin{aligned} \int_a^b (v'(x))^2 dx &= \sum_{j=0}^n \int_{x_j}^{x_{j+1}} (v'(x))^2 dx = \frac{1}{h^2} \sum_{j=0}^n \underbrace{(x_{j+1} - x_j)}_{=h} (v_{j+1} - v_j)^2 = \\ &= \frac{1}{h} \sum_{j=0}^n (v_{j+1} - v_j)^2. \end{aligned}$$

Dosažením $v(x)$ a $v'(x)$ do rovnice (9.4) dostaneme

$$F(v) = \frac{1}{2h} \sum_{j=0}^n (v_{j+1} - v_j)^2 + \frac{1}{2} \int_a^b \sigma(x) v^2(x) dx - \int_a^b v(x) f(x) dx.$$

F je nyní kvadratickou funkcí parametrů v_1, v_2, \dots, v_n . Hledáme nyní minimum této funkce

$$\frac{\partial F}{\partial v_i} = 0, \quad i = 0, 1, \dots, n.$$

Dostáváme soustavu lineárních rovnic.

Přitom se v_j objevuje pouze ve sčítancích s indexem $j - 1, j$. Takže máme

$$\frac{\partial v}{\partial v_j} = b_j(x),$$

a tedy

$$\frac{\partial(v^2)}{\partial v_j} = 2v \frac{\partial v}{\partial v_j} = 2v \cdot b_j(x).$$

Odtud ovšem plyne, že se stačí omezit na interval $[x_{i-1}, x_{i+1}]$, protože mimo něj je funkce b_j nulová.

Dostáváme tedy soustavu pro $i = 1, 2, \dots, n$,

$$\frac{1}{h}(2v_i - v_{i-1} - v_{i+1}) + \int_{x_{i-1}}^{x_{i+1}} \sigma(x)v(x)b_i(x)dx = \int_{x_{i-1}}^{x_{i+1}} f(x)b_i(x)dx.$$

Toto je opět soustava s třídiagonální maticí.

Když si ještě uvědomíme, že

$$v(x)b_i(x) = \left(\sum_{j=0}^n v_j b_j(x) \right) b_i(x) = v_i b_i^2(x) + v_{i-1} b_{i-1}(x)b_i(x) + v_{i+1} b_{i+1}(x)b_i(x).$$

A protože báze funkce známe, tak známe i tyto koeficienty.

V případě, že $\sigma(x) \equiv 0$ dostaneme tentýž výsledek jako metodou konečných diferencí.

Pojmy k zapamatování

- Seznámili jsme se s další numerickou metodou pro nalezení řešení okrajové úlohy pro obyčejnou diferenciální rovnici - s metodou konečných prvků, která je v současnosti velmi často používána při řešení celé řady nejrůznějších úloh v řadě aplakačních oblastí.
- Na rozvoji metody konečných prvků se podílela celá řada vynikajících matematiků. Pro nás je důležité, že první dokázal konvergenci metody v roce 1968 profesor Miloš Zlámal, pracovník VUT Brno. Na dalším rozvoji metody se podíleli například prof. A. Ženíšek z FSI VUT v Brně, prof F. Melkes z FEKT VUT v Brně, prof. M. Křížek z MU AV ČR, Praha, a mnoho dalších.
- Ačkoliv se metoda konečných prvků používá především pro řešení úloh z parciálními diferenciálními rovnicemi, zabývali jsme se jí, protože v jednodimenzionálním případě je základní princip metody snáze pochopitelný.

Kontrolní otázky

1. Co rozumíme pojmem báze funkce?
2. Jakou hlavní vlastnost musí mít báze funkce?
3. Co nám bude zaručovat řešitelnost soustavy rovnic?
4. Co rozumíme pod pojmem funkcionál?
5. Jaký byl přínos brněnských matematiků pro rozvoj teoretických základů metody konečných prvků?

10 Parciální diferenciální rovnice

Průvodce studiem

V předchozích kapitolách jsme se zabývali obyčejnými diferenciálními rovnicemi. S pojmem parciální derivace jste se setkali v předmětu Matematika 2. S parciálními diferenciálními rovnicemi jste se mohli setkat v předmětu Diferenciální rovnice v elektrotechnice. Zde se jimi budeme zabývat širěji a s hlavním důrazem na hledání numerického řešení.

Cílem této kapitoly je seznámit čtenáře se základy teorie parciálních diferenciálních rovnic. Stanovíme si co budeme rozumět řešením parciální diferenciální rovnice a jaké úlohy budeme řešit.

Potom se zaměříme na parciální diferenciální rovnice prvního řádu. Zformulujeme si požadavky na počáteční úlohu pro parciální diferenciální rovnici prvního řádu.

Ukážeme si způsoby řešení nejjednodušších parciálních diferenciálních rovnic prvního řádu.

Cíle

Po prostudování této kapitoly budete schopni:

- Odlišit obyčejné a parciální diferenciální rovnice a jejich soustavy.
- Umět řešit základní typy parciálních diferenciálních rovnic prvního řádu.

10.1 Základní pojmy

Definice 10.1. Parciální diferenciální rovnicí rozumíme rovnici, která obsahuje neznámou funkci více proměnných a její parciální derivace. Řád nejvyšší derivace, která se v rovnici vyskytuje, se nazývá řádem dané rovnice. Řešením parciální diferenciální rovnice rozumíme každou funkci, která je definovaná v zadané oblasti, včetně svých parciálních derivací, až do řádu rovnice včetně, a vyhovuje dané rovnici v zadané oblasti.

Obecný tvar parciální diferenciální rovnice je

$$F \left(x_1, \dots, x_n, u(x_1, \dots, x_n), \frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \dots, \frac{\partial u}{\partial x_n}, \frac{\partial^2 u}{\partial x_1^2}, \dots, \frac{\partial^k u}{\partial x_1^k}, \dots \right) = 0. \quad (10.1)$$

Příklad 10.2. Hledejme funkci u , která vyhovuje rovnici

$$\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} = 0. \quad (10.2)$$

Řešení. Máme parciální diferenciální rovnici prvního řádu. Jejím řešením je funkce

$$u(x, y) = x - y + \pi,$$

protože

$$\frac{\partial u}{\partial x} = 1, \quad \frac{\partial u}{\partial y} = -1,$$

v každém bodě roviny Oxy . Číslo π můžeme nahradit jiným libovolným reálným číslem.

Řešením bude i funkce

$$u(x, y) = \alpha x - \alpha y + \beta,$$

pro libovolná $\alpha, \beta \in \mathbb{R}$, protože

$$\frac{\partial u}{\partial x} = \alpha, \quad \frac{\partial u}{\partial y} = -\alpha,$$

v každém bodě roviny Oxy .

Obdobně se můžeme přesvědčit, že řešením bude i funkce

$$u(x, y, z) = x - y + z,$$

protože

$$\frac{\partial u}{\partial x} = 1, \quad \frac{\partial u}{\partial y} = -1, \quad \frac{\partial u}{\partial z} = 0, \quad \frac{\partial u}{\partial y} = 0.$$

Lehce si ověříme, že řešením bude i každá funkce

$$u(x, y, z) = x - y + f(z),$$

kde f je libovolná funkce proměnné z . □

Z uvedeného příkladu plyne jeden podstatný rozdíl mezi parciálními rovnicemi a obyčejnými diferenciálními rovnicemi. Ze zápisu obyčejné diferenciální rovnice $y' = x + y$ poznáme okamžitě, že hledaná funkce y závisí pouze na x . Ze zápisu parciální rovnice (10.2) nepoznáme na kolika proměnných závisí řešení. Víme, že hledaná funkce u závisí na proměnných x, y , ale nevíme, zda se jedná o funkci dvou, tří či více proměnných.

Přijmeme proto hned na místě úmluvu, že řešení dané parciální rovnice budeme hledat pouze mezi funkcemi těch proměnných, které se přímo v rovnici vyskytují. Bude-li hledaná neznámá funkce záviset i na proměnných, které se v rovnici nevyskytují, bude to při zápisu rovnice výslovně zdůrazněno.

Příklad 10.3. Najděte funkci dvou proměnných $u(x, y)$, která vyhovuje rovnici

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0. \quad (10.3)$$

Řešení. Máme parciální diferenciální rovnici druhého řádu. Jejím řešením je funkce

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

protože

$$\frac{\partial u}{\partial x} = 2x, \quad \frac{\partial^2 u}{\partial x^2} = 2, \quad \frac{\partial u}{\partial y} = -2y, \quad \frac{\partial^2 u}{\partial y^2} = -2,$$

v každém bodě roviny Oxy .

Obdobně se můžeme přesvědčit, že řešením budou i funkce

$$u(x, y) = x^2 - y^2 + ax + by, \quad a, b \in R,$$

nebo

$$u(x, y) = e^x \sin y.$$

□

Podobně jako u obyčejných diferenciálních rovnic máme i u parciálních rovnic dva základní problémy:

1. Najít obecné řešení dané parciální rovnice = najít všechna řešení.
2. Najít takové řešení dané parciální rovnice, které vyhovuje některým doplňujícím podmínkám (které obvykle plynou z daného technického problému, který řešíme a nebo jsou součástí zadání matematického úkolu).

Ukážeme si později, že najít obecné řešení parciální rovnice je mnohem těžší než u obyčejných diferenciálních rovnic. Proto se u parciálních rovnic studují častěji problémy s dodatečnými podmínkami.

10.2 Parciální diferenciální rovnice prvního řádu

Definice 10.4. Rovnici

$$f\left(x, y, u(x, y), \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right) = 0. \quad (10.4)$$

nazýváme *parciální diferenciální rovnici prvního řádu*, kde funkce $f(x, y, u, p, q)$ je definovaná na otevřené množině D proměnných x, y, u, p, q .

Definice 10.5. Řešením rovnice (10.4) v oblasti G proměnných x, y nazveme každou takovou funkci u , definovanou a spojitou v G , pro kterou platí:

1. Funkce u má v oblasti G spojitě parciální derivace prvního řádu.
2. Pro každé $(x, y) \in G$ platí, že $\left(x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right) \in D$.
3. Funkce $u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}$ splňují rovnici (10.4).

Příklad 10.6. Řešením rovnice

$$x \frac{\partial u}{\partial x} + y \frac{\partial u}{\partial y} + 2u = 0$$

je funkce

$$u = \frac{1}{x^2 + y^2},$$

kteřá je definována v oblasti $G = \{(x, y) : x^2 + y^2 \neq 0\}$ (neboli funkce u je definována pro všechny body roviny Oxy kromě počátku, t.j. bodu $(0, 0)$).

Příklad 10.7. Jedním z řešení rovnice

$$u \left(\frac{\partial u}{\partial x}\right)^2 + \frac{\partial u}{\partial y} \ln(xy) = \frac{x^2 + y}{x^2 y} \ln(xy)$$

je funkce $u = \ln(xy)$, která je definována v oblasti $G = \{(x, y) : xy > 0\}$, neboli u je definována pro všechny vnitřní body prvního a třetího kvadrantu roviny Oxy .

Poznámka 10.8. Pro určení řádu rovnice je rozhodující, jakého řádu jsou parciální derivace, které se v ní vyskytují, ne mocniny těchto derivací.

10.3 Formulace počáteční úlohy.

U obyčejných diferenciálních rovnic jsme vždy hledali obecné řešení Cauchyovy úlohy $y' = f(x, y), y(x_0) = y_0$. U parciálních diferenciálních rovnic bude situace složitější.

Definice 10.9. *Cauchyovou úlohou* pro rovnici (10.4) rozumíme dvojici: rovnici

$$f\left(x, y, z, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right) = 0. \quad (10.5)$$

a počáteční křivku Θ zadanou parametricky

$$x = \varphi(t), \quad y = \psi(t), \quad z = \chi(t), \quad t \in (a, b). \quad (10.6)$$

Funkci $z = h(x, y)$, která má spojitě parciální derivace v G , nazveme *řešením Cauchyovy úlohy* (10.5), (10.6), jestliže funkce h splňuje v G rovnici (10.5) a pro všechna $t \in (a, b)$ křivka $(x = \varphi(t), y = \psi(t))$ leží v G a navíc platí $\chi(t) = h(\varphi(t), \psi(t))$.

O křivce Θ budeme všude dále předpokládat, že je hladká a jednoduchá.

10.4 Nejjednodušší příklady parciálních rovnic prvního řádu

10.4.1 Rovnice typu $\frac{\partial z(x, y)}{\partial x} = 0$.

Řešením rovnice

$$\frac{\partial z(x, y)}{\partial x} = 0 \quad (10.7)$$

je buď libovolná konstanta nebo libovolná funkce závisající pouze na proměnné y , která bude mít spojitě parciální derivace,

$$z(x, y) = h(y). \quad (10.8)$$

Podívejme se, jaký je geometrický význam rovnice (10.7).

V prostoru $Oxyz$ jde o rovnici válcové plochy, jejíž přímky jsou kolmé na rovinu Oyz a jsou tedy rovnoběžné s x -ovou souřadnicovou osou.

Potom lze lehce řešit Cauchyovu úlohu pro rovnici (10.7). Máme-li danou křivku Θ , potom každým bodem křivky vedeme přímku rovnoběžnou s osou x . Dostaneme tak válcovou plochu, která je zřejmě řešením rovnice (10.7) a prochází křivkou Θ .

Křivka Θ přitom může být i prostorová, tj. nepožadujeme, aby byla závislá pouze na proměnné y .

Vraťme se nyní zpět k analytickému řešení Cauchyovy úlohy pro rovnici (10.7).

Nechť je křivka Θ zadána parametricky

$$x = \varphi(t), \quad y = \psi(t), \quad z = \chi(t), \quad a < t < b.$$

Předpokládejme, že pro funkce $\psi(t), \chi(t)$ platí, že mají spojitě derivace v intervalu (a, b) a že funkce $\psi(t)$ je ryze monotónní, tj. že pro ni existuje funkce inverzní $\psi^{-1}(t)$ taková, že $\psi(\psi^{-1}(y)) = y$.

Předpokládejme, že existuje řešení $z = H(y)$ Cauchyovy úlohy pro rovnici (10.7) a že známe křivku Θ . Potom dosazením zjistíme, že platí

$$\chi(t) = H(\psi(t)). \quad (10.9)$$

Dosadíme do této rovnice $t = \psi^{-1}(y)$, dostaneme

$$\chi(\psi^{-1}(y)) = H(y), \quad (10.10)$$

a protože $H(y) = z$, máme

$$\chi(\psi^{-1}(y)) = z.$$

Tím máme dokázanou jednoznačnost Cauchyovy úlohy.

Z druhé strany, definujeme funkci H pomocí rovnosti (10.10). Potom ale platí

$$\frac{\partial H}{\partial x} = 0,$$

neboli funkce (10.10) je řešením rovnice (10.7) a současně platí (10.9), což znamená, že řešení prochází křivkou Θ .

Příklad 10.10. Najděte řešení rovnice

$$\frac{\partial z(x, y)}{\partial x} = 0,$$

které prochází křivkou

$$x = t, \quad y = t^3, \quad z = t^2, \quad t \in (-\infty, +\infty).$$

Řešení. Podle (10.10) platí, že jediným řešením této úlohy je

$$z = \chi(\psi^{-1}(y)).$$

V našem případě máme

$$\chi(t) = t^2, \quad \psi^{-1}(y) = \sqrt[3]{y}.$$

Řešením je proto

$$z = \sqrt[3]{y^2}.$$

□

Příklad 10.11. Najděte řešení rovnice

$$\frac{\partial z(x, y)}{\partial x} = 0,$$

které prochází křivkou

$$x = 0, \quad z = y^2, \quad y \in (-1, 1).$$

Řešení. Křivka Θ je v našem případě parabola, která leží v rovině Oyz . Podle předchozího platí, že řešením je válcová plocha, která prochází křivkou Θ a je rovnoběžná s x -ovou osou, takže

$$z = y^2$$

je řešením.

□

Poznámka 10.12.

1. Zcela analogicky jako rovnice $\frac{\partial z(x, y)}{\partial x} = 0$ se řeší rovnice

$$\frac{\partial z(x, y)}{\partial y} = 0.$$

2. Předpoklad existence inverzní funkce ψ^{-1} je nezbytný pro existenci a jednoznačnost řešení. Pokud není splněn, Cauchyova úloha nemusí být řešitelná nebo může mít nekonečně mnoho řešení.
3. Postup použitý při hledání řešení rovnice (10.7) můžeme zobecnit i pro parciální diferenciální rovnici více než dvou proměnných

$$\frac{\partial z(x_1, x_2, \dots, x_n)}{\partial x_i} = 0 \quad i = \{1, 2, \dots, n\}.$$

Řešením této rovnice bude funkce $z = f(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$, která nezávisí na x_i .

10.4.2 Rovnice typu $\frac{\partial z(x, y)}{\partial x} = f(x, y)$.

Řešení rovnice

$$\frac{\partial z(x, y)}{\partial x} = f(x, y), \quad (10.11)$$

za předpokladu, že $f(x, y)$ je spojitá funkce v oblasti G , je dáno vztahem

$$z = \int f(x, y) dx + H(y). \quad (10.12)$$

Důkaz existence a jednoznačnosti řešení se provádí stejně jako v předchozím případě.

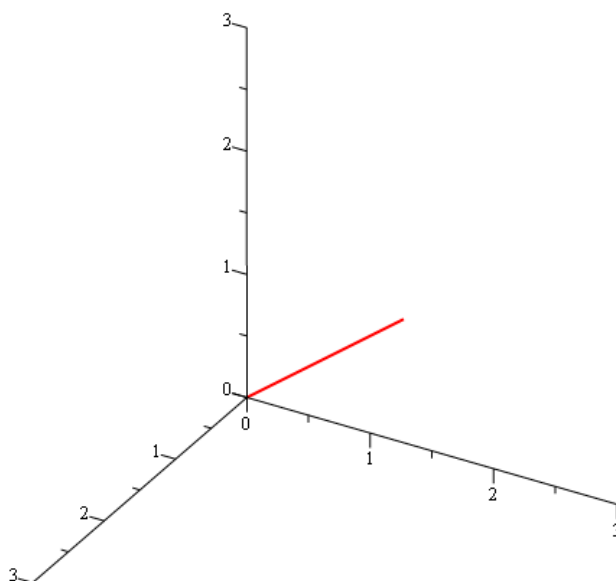
Příklad 10.13. Najděte řešení rovnice

$$\frac{\partial z}{\partial x} = x^2 + xy - y^2,$$

které prochází křivkou

$$x = t, \quad y = t, \quad z = t, \quad t > 0.$$

Řešení. Křivka Θ je v našem případě osou prvního oktantu – viz obrázek (10.1).



Obr. 10.1: Počáteční křivka z příkladu 10.13

Řešení získáme integrací podle proměnné x :

$$z = \frac{x^3}{3} + \frac{x^2 y}{2} - xy^2 + H(y).$$

Řešení musí procházet počáteční křivkou Θ . Dosadíme-li za x, y, z parametrické vyjádření křivky Θ , dostaneme, že musí platit

$$t = \frac{t^3}{3} + \frac{t^3}{2} - t^3 + H(t).$$

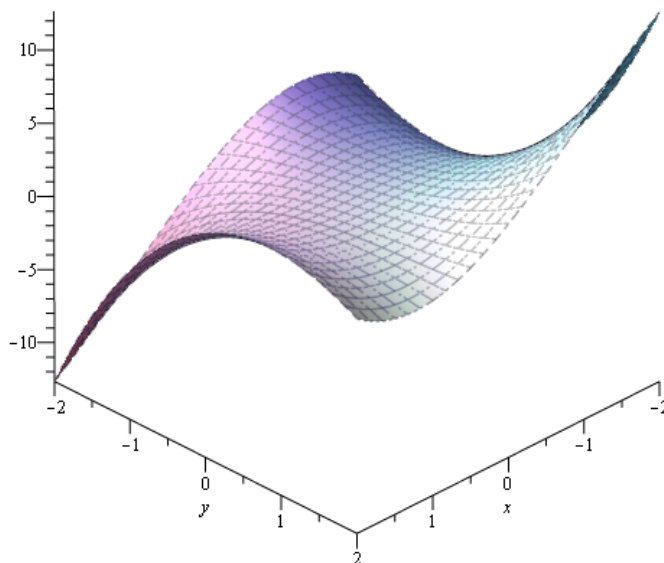
Odtud dostáváme

$$H(t) = t + \frac{t^3}{6}.$$

Řešením naší úlohy je funkce

$$z = \frac{x^3}{3} + \frac{x^2y}{2} - xy^2 + y + \frac{y^3}{6}.$$

Povšimněte si, že řešením je plocha – viz obrázek 10.2.



Obr. 10.2: Zobrazení řešení příkladu 10.13

□

Hledejme nyní řešení rovnice (10.11), které vyhovuje podmínce

$$z(0, y) = \omega(y). \quad (10.13)$$

To značí, že řešení $z = z(x, y)$ má procházet křivkou

$$x = 0, \quad z = \omega(y),$$

kteřá leží v rovině Oyz . Podle vzorce (10.12) má řešení tvar

$$z(x, y) = \int_0^x f(\xi, y) d\xi + H(y). \quad (10.14)$$

Dosadíme do této rovnice podle podmínky (10.13) hodnotu $x = 0$, dostaneme

$$\omega(y) = H(y).$$

Hledané řešení má proto tvar

$$z(x, y) = \int_0^x f(\xi, y) d\xi + \omega(y).$$

Příklad 10.14. Najděte řešení rovnice

$$\frac{\partial z(x, y)}{\partial x} = 3x^2 + y \sin x,$$

kteřé prochází křivkou $z(0, y) = y^3$.

Řešení. Podle předchozího můžeme hned psát

$$z(x, y) = \int_0^x (3\xi^2 + y \sin \xi) d\xi + y^3.$$

Po integraci dostaneme řešení ve tvaru

$$z(x, y) = x^3 + y^3 + y(1 - \cos x).$$

□

Poznámka 10.15.

1. Analogicky jako rovnice $\frac{\partial z}{\partial x} = f(x, y)$ se řeší i rovnice $\frac{\partial z}{\partial y} = f(x, y)$.
2. Obdobně jako u rovnice $\frac{\partial z}{\partial x} = 0$ i u rovnice $\frac{\partial z}{\partial x} = f(x, y)$ se může stát, že rovnice nemá řešení nebo je řešení nekonečně mnoho.

10.4.3 Rovnice typu $\frac{\partial z(x, y)}{\partial x} = f_1(x) \cdot f_2(z)$.

Mějme rovnici

$$\frac{\partial z(x, y)}{\partial x} = f_1(x) \cdot f_2(z), \quad (10.15)$$

kde funkce f_1 je spojitá na intervalu (a, b) a f_2 je spojitá a nenulová na intervalu (c, d) .

Řešení parciální rovnice (10.15) je analogické řešení obyčejné diferenciální rovnice se separovanými proměnnými.

$$\frac{\partial z(x, y)}{\partial x} = f_1(x) \cdot f_2(z),$$

$$\int \frac{1}{f_2(z)} dz = \int f_1(x) dx + H(y).$$

Označme $F_1(x) = \int f_1(x) dx$, $F_2(z) = \int \frac{1}{f_2(z)} dz$, potom můžeme předchozí rovnici přepsat na tvar

$$F_2(z) = F_1(x) + H(y).$$

Pokud je funkce F_2 ryze monotonní, potom můžeme z rovnice vypočítat funkci z a dostaneme

$$z(x, y) = F_2^{-1}[F_1(x) + H(y)]. \quad (10.16)$$

Přítom musíme požadovat, aby funkce $F_1(x) + H(y)$ ležela v definičním oboru funkce F_2^{-1} . Pouze tehdy má řešení smysl.

Příklad 10.16. Najděte řešení rovnice

$$\frac{\partial z(x, y)}{\partial x} = x \cdot e^z \quad (10.17)$$

Řešení. Rovnici si upravíme podle předchozího.

$$F_1(x) = \int x dx = \frac{x^2}{2},$$

$$F_2(z) = \int e^{-z} dz = -e^{-z}$$

Po dosazení máme

$$-e^{-z} = \frac{x^2}{2} + H(y).$$

Protože na pravé straně máme ryze monotonní funkci, můžeme rovnici vyřešit vzhledem k z a dostaneme

$$z = -\ln\left(-\frac{x^2}{2} - H(y)\right), \quad (10.18)$$

což je řešení rovnice (10.17). Výraz bude mít smysl pouze pro

$$\left(-\frac{x^2}{2} - H(y)\right) > 0 \quad \Rightarrow \quad \left(\frac{x^2}{2} + H(y)\right) < 0.$$

□

Pokud v rovnici (10.15) neplatí podmínka $f_2(z) \neq 0$, potom mohou existovat i řešení, která nezískáme pomocí vztahu (10.16).

Příklad 10.17. Najděte řešení rovnice

$$\frac{\partial z}{\partial x} = z. \quad (10.19)$$

Řešení. Rovnici (10.19) si upravíme podle předchozího postupu.

$$\begin{aligned}\frac{\partial z}{\partial x} &= z, \\ \int \frac{dz}{z} &= \int dx, \\ \ln z &= x + H(y), \\ z &= e^{x+H(y)}, \\ z &= K(y)e^x,\end{aligned}$$

kde $K(y) = e^{H(y)}$.

Rovnice (10.19) má ale ještě řešení $z = 0$, které nezískáme žádnou volbou funkce $H(y)$, respektive $K(y)$. \square

Právě uvedený postup můžeme použít i pro ty rovnice, které obsahují pouze jednu parciální derivaci, tedy pro rovnice tvaru

$$\frac{\partial z(x, y)}{\partial x} = \varphi(x, y, z)$$

nebo

$$\frac{\partial z(x, y)}{\partial y} = \psi(x, y, z).$$

Vztah (10.16) můžeme použít i pro řešení Cauchyovy úlohy pro odpovídající typy rovnic. Protože obecný postup by byl příliš nepřehledný, ukážeme si použití na příkladech.

10.4.4 Řešené příklady

Příklad 10.18. Určete řešení Cauchyovy úlohy

$$\begin{aligned}\frac{\partial z(x, y)}{\partial x} &= z(x, y), \\ z(0, y) &= y^2.\end{aligned}$$

Řešení. Máme stejnou rovnici jako v předchozím příkladu. Řešení proto bude mít tvar

$$z(x, y) = K(y)e^x.$$

Dosadíme do něj $x = 0$ a dostaneme

$$z(0, y) = y^2 = K(y).$$

Řešením Cauchyovy úlohy je proto funkce

$$z(x, y) = y^2 e^x.$$

\square

Příklad 10.19. Určete řešení Cauchyovy úlohy

$$\frac{\partial z(x, y)}{\partial x} = ax + by + c \cdot z(x, y) + d,$$

$$z(0, y) = y^3 + 2y,$$

kde a, b, c, d jsou konstanty, přičemž $c \neq 0$.

Řešení. Prohlásíme y za konstantu a řešíme parciální rovnici jako obyčejnou diferenciální rovnici.

$$\frac{dz}{dx} - c \cdot z = ax + by + d.$$

Použijeme metodu variace konstanty. Nejprve řešíme homogenní rovnici (tj. o pravé straně předpokládáme že je rovna nule).

$$\frac{dz}{dx} - c \cdot z = 0,$$

$$\frac{dz}{dx} = c \cdot z,$$

$$\frac{dz}{z} = c dx.$$

$$\ln z = cx + K,$$

$$z = Le^{cx},$$

kde $L = e^K$. Nyní předpokládáme, že $L = L(x)$. Potom

$$z' = L'e^{cx} + Lce^{cx},$$

$$L'e^{cx} + Lce^{cx} = ax + by + cLe^{cx} + d,$$

$$L'e^{cx} = ax + by + d,$$

$$L' = (ax + by + d)e^{-cx}.$$

Integrací *per partes* dostaneme

$$\left| \begin{array}{ll} u = ax + by + d & u' = a \\ v' = e^{-cx} & v = e^{-cx} \cdot \left(\frac{-1}{c}\right) \end{array} \right|$$

$$L = (ax + by + d)e^{-cx} \cdot \left(\frac{-1}{c}\right) + \frac{a}{c} \int e^{-cx} dx,$$

$$L = (ax + by + d)e^{-cx} \cdot \left(\frac{-1}{c}\right) - \frac{a}{c^2} e^{-cx} + M.$$

Po dosazení dostaneme řešení ve tvaru

$$z = -\frac{1}{c}(ax + by + d) - \frac{a}{c^2} + Me^{cx}.$$

Tím jsme získali řešení diferenciální rovnice. Pro řešení parciální rovnice je nutné místo konstanty M brát funkci $H(y)$.

$$z = -\frac{1}{c}(ax + by + d) - \frac{a}{c^2} + H(y)e^{cx}.$$

Nyní najdeme řešení, které vyhovuje naší počáteční podmínce. Dosazením do předchozí rovnice hodnoty $x = 0$ dostaneme

$$y^3 + 2y = -\frac{b}{c}y - \frac{d}{c} - \frac{a}{c^2} + H(y),$$

a odtud plyne

$$H(y) = y^3 + 2y + \frac{b}{c}y + \frac{d}{c} + \frac{a}{c^2}.$$

Řešením Cauchyovy úlohy je tedy funkce

$$z = -\frac{1}{c}(ax + by + d) - \frac{a}{c^2} + \left(y^3 + \left(2 + \frac{b}{c} \right) y + \frac{d}{c} + \frac{a}{c^2} \right) e^{cx}.$$

□

10.4.5 Lineární homogenní parciální rovnice prvního řádu

Definice 10.20. *Lineární parciální diferenciální rovnici nazveme výraz*

$$a(x, y) \frac{\partial z(x, y)}{\partial x} + b(x, y) \frac{\partial z(x, y)}{\partial y} = c(x, y) \cdot z(x, y) + d(x, y), \quad (10.20)$$

kde a, b, c, d jsou funkce proměnných x, y definované na otevřené množině G .

Jestliže $c(x, y) \equiv 0, d(x, y) \equiv 0$, potom se rovnice (10.20) nazývá homogenní.

Pokud je $a \equiv 0$ a nebo $b \equiv 0$ můžeme použít pro řešení dříve uvedené metody. Pokud jsou obě funkce nenulové, musíme hledat řešení jinými metodami, které se probírají například v předmětu Diferenciální rovnice v elektrotechnice.

10.5 Parciální diferenciální rovnice druhého řádu

10.5.1 Klasifikace rovnic na hyperbolické, parabolické a eliptické

Definice 10.21. Parciální diferenciální rovnicí druhého řádu dvou proměnných rozumíme rovnici

$$F\left(x, y, u(x, y), \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial^2 u}{\partial x^2}, \frac{\partial^2 u}{\partial x \partial y}, \frac{\partial^2 u}{\partial y^2}\right) = 0.$$

Definice 10.22. Parciální diferenciální rovnicí druhého řádu tří proměnných rozumíme rovnici

$$\Phi\left(x, y, z, u(x, y, z), \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z}, \frac{\partial^2 u}{\partial x^2}, \frac{\partial^2 u}{\partial x \partial y}, \frac{\partial^2 u}{\partial x \partial z}, \frac{\partial^2 u}{\partial y^2}, \frac{\partial^2 u}{\partial y \partial z}, \frac{\partial^2 u}{\partial z^2}\right) = 0.$$

Příklad 10.23. Parciální diferenciální rovnice druhého řádu se často používají při popisu fyzikálních a technických dějů a procesů. Následují některé parciální rovnice, které se používají v elektrotechnice a příbuzných oborech.

$$\Delta u \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0 \quad \text{Laplaceova rovnice (elektrostatické pole),}$$

$$\Delta u = f \quad \text{Poissonova rovnice (elektrostatické pole s volnými náboji),}$$

$$\Delta u = a^2 u \quad \text{Helmholzova rovnice (stacionární vlnová rovnice),}$$

$$\Delta u = a \frac{\partial u}{\partial t} \quad \text{difuzní rovnice,}$$

$$\Delta u = a^2 \frac{\partial^2 u}{\partial t^2} \quad \text{vlnová rovnice (šíření elektromagnetických vln),}$$

$$\frac{\partial}{\partial x} \left(v \frac{\partial u}{\partial x} \right) + \frac{\partial}{\partial y} \left(v \frac{\partial u}{\partial y} \right) = -J, v = v(|\text{grad } u|) \quad \text{rovinné stacionární magnetické pole,}$$

$$\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial t^2} = 0 \quad \text{rovnice pro kmity struny,}$$

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0 \quad \text{rovnice vedení tepla.}$$

Definice 10.24. Kvazilineární parciální diferenciální rovnicí druhého řádu dvou proměnných rozumíme rovnici

$$A \frac{\partial^2 u}{\partial x^2} + B \frac{\partial^2 u}{\partial x \partial y} + C \frac{\partial^2 u}{\partial y^2} + K = 0, \quad (10.21)$$

kde funkce A, B, C, K jsou funkcemi $x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}$.

Definice 10.25. Lineární parciální diferenciální rovnici druhého řádu dvou proměnných rozumíme rovnicí

$$a(x, y) \frac{\partial^2 u}{\partial x^2} + b(x, y) \frac{\partial^2 u}{\partial x \partial y} + c(x, y) \frac{\partial^2 u}{\partial y^2} + d(x, y) \frac{\partial u}{\partial x} + e(x, y) \frac{\partial u}{\partial y} + g(x, y)u = f(x, y). \quad (10.22)$$

Funkce a, b, c, d, e, g, f jsou spojité v oblasti $\Omega \subset \mathbb{R}^2$, přičemž aspoň jedna z funkcí a, b, c je nenulová v každém bodě $(x, y) \in \Omega$.

Řešením rovnice (10.21) v oblasti Ω rozumíme každou funkci $u(x, y) \in C^2(\Omega)$, která v Ω identicky splňuje (10.21).

Označme $D = b^2(x, y) - 4a(x, y)c(x, y)$, potom jestliže

$D < 0 \quad \forall (x, y) \in \Omega$ pak se jedná o eliptický typ,

$D = 0 \quad \forall (x, y) \in \Omega$ pak se jedná o parabolický typ,

$D > 0 \quad \forall (x, y) \in \Omega$ pak se jedná o hyperbolický typ.

Ve zbývajících případech mluvíme o smíšeném typu.

Příklad 10.26. Rovnice

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

je eliptického typu,

$$\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = 0$$

je hyperbolického typu,

$$\frac{\partial u}{\partial x} - \frac{\partial^2 u}{\partial y^2} = 0$$

je parabolického typu.

Určení typu někdy záleží na oblasti, ve které je funkce definovaná.

Příklad 10.27. Rovnice

$$\frac{\partial^2 u}{\partial x^2} + y \frac{\partial^2 u}{\partial y^2} = 0$$

má $D = -4y$, a proto je eliptického typu pro všechny body horní poloroviny, tj. $y > 0$, je hyperbolického typu v dolní polorovině, tj. $y < 0$, a je parabolického typu v bodech osy x , t.j. $y = 0$.

Příklad 10.28. Mějme rovnici kvazistacionárního elektromagnetického pole

$$\frac{\partial}{\partial x} \left(v \frac{\partial u}{\partial x} \right) + \frac{\partial}{\partial y} \left(v \frac{\partial u}{\partial y} \right) = \gamma(x, y) - J, \quad v = v(|\text{grad } u|).$$

Tato rovnice je v elektricky vodivých podoblastech parabolického typu, protože v těchto podoblastech je elektrická vodivost $\gamma(x, y)$ kladná, zatímco v nevodivých podoblastech je elektrická vodivost nulová, a proto se v nich jedná o eliptický typ.

10.6 Transformace proměnných.

Věta 10.29. Každou lineární parciální diferenciální rovnici druhého řádu dvou proměnných, eliptickou, nebo hyperbolickou, nebo parabolickou, lze vhodnou lokální transformací souřadnic převést v okolí každého bodu $(x_0, y_0) \in \Omega$ na kanonický tvar.

U rovnice eliptického typu na tvar

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + a_1(x, y) \frac{\partial u}{\partial x} + b_1(x, y) \frac{\partial u}{\partial y} + c_1(x, y)u = f_1(x, y),$$

u rovnice hyperbolického typu na tvar

$$\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} + a_2(x, y) \frac{\partial u}{\partial x} + b_2(x, y) \frac{\partial u}{\partial y} + c_2(x, y)u = f_2(x, y)$$

a u rovnice parabolického typu na tvar

$$\frac{\partial^2 u}{\partial x^2} + a_3(x, y) \frac{\partial u}{\partial x} + b_3(x, y) \frac{\partial u}{\partial y} + c_3(x, y)u = f_3(x, y), \quad a_3(x, y) \neq 0.$$

Mějme nyní lineární rovnici druhého řádu s konstantními koeficienty

$$A \frac{\partial^2 u}{\partial x^2} + B \frac{\partial^2 u}{\partial x \partial y} + C \frac{\partial^2 u}{\partial y^2} + D \frac{\partial u}{\partial x} + E \frac{\partial u}{\partial y} + Gu = F, \quad (10.23)$$

kde $|A| + |B| + |C| \neq 0$. Dané rovnici přiřadíme charakteristickou rovnici

$$A(dy)^2 - Bdx dy + C(dx)^2 = 0. \quad (10.24)$$

Řešením charakteristické rovnice bude každá dvojice funkcí

$$x = \varphi(t), \quad y = \psi(t), \quad t \in (\alpha, \beta), \quad (10.25)$$

která vyhovuje dané rovnici. Přitom za dx dosadíme výraz $\frac{dx}{dt}$ a za dy dosadíme výraz $\frac{dy}{dt}$ a nakonec celou rovnici vydělíme výrazem $(dt)^2$.

Definice 10.30. Jestliže jsou funkce $x(t) = \varphi(t)$, $y(t) = \psi(t)$ řešením rovnice (10.24) a jestliže jsou rovnice (10.25) parametrickými rovnicemi hladké křivky, potom tuto křivku nazveme *charakteristikou*.

Příklad 10.31. Najděte charakteristiky rovnice

$$\frac{\partial^2 z}{\partial x^2} - \frac{\partial^2 z}{\partial y^2} = 0.$$

Řešení. Naší rovnici přiřadíme charakteristickou rovnici:

$$(dy)^2 - (dx)^2 = 0,$$

neboli

$$(dy)^2 = (dx)^2 \implies \frac{(dy)^2}{(dx)^2} = 1,$$

$$\frac{dy}{dx} = \pm 1.$$

Charakteristikami naší rovnice jsou potom přímky $y = x + P$ a $y = -x + P$, kde $P \in \mathbb{R}$. Máme tak dvě třídy přímek a každým bodem roviny Oxy prochází právě jedna přímka z každé třídy. \square

Příklad 10.32. Najděte charakteristiky pro rovnici vedení tepla

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0.$$

Řešení. Rovnici přiřadíme charakteristickou rovnici:

$$(dy)^2 = 0,$$

neboli

$$(dy) = 0, \implies y = P, \quad P \in \mathbb{R}.$$

Rovnice vedení tepla má charakteristiky množinu přímek $y = \text{konst.}$ \square

Příklad 10.33. Najděte charakteristiky Laplaceovy rovnice

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

Řešení. Laplaceově rovnici přiřadíme charakteristickou rovnici:

$$(dy)^2 + (dx)^2 = 0.$$

Součet dvou nezáporných hodnot je roven nule tehdy a jen tehdy, pokud obě hodnoty jsou současně nulové, takže máme

$$(dy) = 0, \implies y = P, \quad (dx) = 0, \implies x = R, \quad P, R \in \mathbb{R}.$$

Vztahy $x = R, y = P$, ale nepopisují žádnou křivku, proto Laplaceova rovnice nemá žádnou charakteristiku. \square

Z výše uvedených příkladů plyne, že každým bodem (x, y) může procházet právě jedna charakteristika, nebo charakteristik může být více a nebo nemusí existovat žádná.

Ukážeme si použití charakteristik při převodu lineární rovnice (10.23) na tzv. kanonický tvar. Nechť máme rovnici hyperbolického typu, tj. $B^2 - 4AC > 0$. Bez omezení obecnosti můžeme předpokládat, že $A \neq 0$. Charakteristickou rovnicí

$$A(dy)^2 - Bdx dy + C(dx)^2 = 0.$$

si upravíme na tvar

$$A \left(\frac{dy}{dx} \right)^2 - B \frac{dy}{dx} + C = 0.$$

Označme $\lambda = \frac{dy}{dx}$. Potom hledáme řešení rovnice

$$A\lambda^2 - B\lambda + C = 0.$$

Vzhledem k podmínce pro rovnici hyperbolického typu $B^2 - 4AC > 0$ máme, že naše kvadratická rovnice bude mít za řešení dvě různá reálná čísla. Označme je $\lambda_1, \lambda_2 \in \mathbb{R}$. Potom

$$\frac{dy}{dx} = \lambda_1,$$

$$dy = \lambda_1 dx,$$

$$y = \lambda_1 x + \xi, \quad \xi \in \mathbb{R}.$$

A analogicky

$$y = \lambda_2 x + \eta, \quad \eta \in \mathbb{R}.$$

Tím jsme dostali, že charakteristikami jsou dvě třídy přímek

$$y = \lambda_1 x + \xi, \quad y = \lambda_2 x + \eta.$$

Každým bodem prochází právě jedna přímka z každé třídy. Jinak řečeno, známe-li bod (x, y) , potom známe i čísla ξ, η a naopak známe-li čísla ξ, η , známe i souřadnice průsečíku charakteristik

$$x = \frac{\xi - \eta}{\lambda_2 - \lambda_1}, \quad y = \frac{\lambda_2 \xi - \lambda_1 \eta}{\lambda_2 - \lambda_1}$$

a naopak

$$\xi = y - \lambda_1 x, \quad \eta = y - \lambda_2 x.$$

Můžeme tak každou funkci proměnných x, y považovat za funkci proměnných ξ, η . Přitom platí

$$u(x, y) = u \left(\frac{\xi - \eta}{\lambda_2 - \lambda_1}, \frac{\lambda_2 \xi - \lambda_1 \eta}{\lambda_2 - \lambda_1} \right) = U(\xi, \eta).$$

A naopak

$$U(\xi, \eta) = U(y - \lambda_1 x, y - \lambda_2 x) = u(x, y).$$

Budeme hledat rovnici, kterou splňuje funkce U , jestliže funkce u je řešením rovnice (10.23). Určíme si derivace

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial U}{\partial \xi} \cdot \frac{\partial \xi}{\partial x} + \frac{\partial U}{\partial \eta} \cdot \frac{\partial \eta}{\partial x} = -\lambda_1 \frac{\partial U}{\partial \xi} - \lambda_2 \frac{\partial U}{\partial \eta}, \\ \frac{\partial u}{\partial y} &= \frac{\partial U}{\partial \xi} \cdot \frac{\partial \xi}{\partial y} + \frac{\partial U}{\partial \eta} \cdot \frac{\partial \eta}{\partial y} = \frac{\partial U}{\partial \xi} + \frac{\partial U}{\partial \eta}, \end{aligned}$$

$$\begin{aligned}\frac{\partial^2 u}{\partial x^2} &= \lambda_1^2 \frac{\partial^2 U}{\partial \xi^2} + 2\lambda_1 \lambda_2 \frac{\partial^2 U}{\partial x \partial y} + \lambda_2^2 \frac{\partial^2 U}{\partial \eta^2}, \\ \frac{\partial^2 u}{\partial x \partial y} &= -\lambda_1 \frac{\partial^2 U}{\partial \xi^2} - (\lambda_1 + \lambda_2) \frac{\partial^2 U}{\partial \xi \partial \eta} - \lambda_2 \frac{\partial^2 U}{\partial \eta^2}, \\ \frac{\partial^2 u}{\partial y^2} &= \frac{\partial^2 U}{\partial \xi^2} + 2 \frac{\partial^2 U}{\partial \xi \partial \eta} + \frac{\partial^2 U}{\partial \eta^2}.\end{aligned}$$

Dosadíme tyto hodnoty do rovnice (10.23) a dostaneme

$$\begin{aligned}A \frac{\partial^2 u}{\partial x^2} + B \frac{\partial^2 u}{\partial x \partial y} + C \frac{\partial^2 u}{\partial y^2} + D \frac{\partial u}{\partial x} + E \frac{\partial u}{\partial y} + Gu - F &= 0, \\ A \left(\lambda_1^2 \frac{\partial^2 U}{\partial \xi^2} + 2\lambda_1 \lambda_2 \frac{\partial^2 U}{\partial x \partial y} + \lambda_2^2 \frac{\partial^2 U}{\partial \eta^2} \right) + B \left(-\lambda_1 \frac{\partial^2 U}{\partial \xi^2} - (\lambda_1 + \lambda_2) \frac{\partial^2 U}{\partial \xi \partial \eta} - \lambda_2 \frac{\partial^2 U}{\partial \eta^2} \right) + \\ C \left(\frac{\partial^2 U}{\partial \xi^2} + 2 \frac{\partial^2 U}{\partial \xi \partial \eta} + \frac{\partial^2 U}{\partial \eta^2} \right) + \dots &= 0,\end{aligned}$$

pro přehlednost zde uvádíme pouze parciální derivace druhého řádu, které jsou pro nás rozhodující, totéž i dále,

$$(A\lambda_1^2 - B\lambda_1 + C) \frac{\partial^2 U}{\partial \xi^2} + (2A\lambda_1\lambda_2 - B(\lambda_1 + \lambda_2) + 2C) \frac{\partial^2 U}{\partial \xi \partial \eta} + (A\lambda_2^2 - B\lambda_2 + C) \frac{\partial^2 U}{\partial \eta^2} + \dots = 0.$$

Čísla λ_1, λ_2 jsme získali jako kořeny kvadratické rovnice. Proto platí

$$A\lambda_i^2 - B\lambda_i + C = 0, \quad i = 1, 2,$$

proto jsou koeficienty u druhých derivací podle ξ, η nulové.

Navíc u kvadratické rovnice platí, podle Vietových vzorců, že součin jejích dvou kořenů je roven absolutnímu členu, dělenému koeficientem u nejvyšší mocniny a jejich součet, záporně vzatý, je roven koeficientu u první mocniny, dělenému koeficientem u nejvyšší mocniny, takže

$$2A\lambda_1\lambda_2 - B(\lambda_1 + \lambda_2) + 2C = 2A \frac{C}{A} - B \frac{B}{A} + 2C = \frac{1}{A}(4AC - B^2) \neq 0.$$

Proto bude koeficient u smíšené derivace nenulový a můžeme jím dělit celou rovnici. Tedy jsme dostali

$$\frac{\partial^2 U}{\partial \xi \partial \eta} + \dots = 0.$$

Máme nyní naši rovnici v *kanonickém tvaru*, který se používá pro určení řešení. Skutečně, jestliže nyní provedeme substituci

$$\xi = r + s, \quad \eta = r - s,$$

spočítáme si příslušné parciální derivace, $U(\xi, \eta) = V(r, s)$,

$$\begin{aligned}\frac{\partial U}{\partial \xi} &= \frac{\partial V}{\partial r} \cdot \frac{1}{2} + \frac{\partial V}{\partial s} \cdot \frac{1}{2}, \\ \frac{\partial^2 U}{\partial \xi \partial \eta} &= \frac{1}{4} \frac{\partial^2 V}{\partial r^2} - \frac{1}{4} \frac{\partial^2 V}{\partial r \partial s} + \frac{1}{4} \frac{\partial^2 V}{\partial s \partial r} - \frac{1}{4} \frac{\partial^2 V}{\partial s^2}.\end{aligned}$$

Dosažením dostaneme

$$\frac{\partial^2 V}{\partial r^2} - \frac{\partial^2 V}{\partial s^2} + \dots = 0.$$

Tím jsme opět získali kanonický tvar pro rovnici hyperbolického typu.

Mějme nyní rovnici (10.23) jako rovnici parabolického typu, tj. platí $4AC - B^2 = 0$. Opět budeme předpokládat, že platí $A \neq 0$. V opačném případě provedeme přeznačení proměnných. Charakteristická rovnice

$$A\lambda^2 - B\lambda + C = 0$$

má potom jeden dvojnásobný kořen $\lambda = \frac{B}{2A}$. Budeme dále předpokládat, že $B \neq 0$, protože pokud je $B = 0$, potom z charakteristické rovnice plyne, že i $C = 0$. O A jsme už dříve předpokládali, že je nenulové. Rovnice (10.23) by pak byla tvaru

$$A \frac{\partial^2 u}{\partial x^2} + D \frac{\partial u}{\partial x} + E \frac{\partial u}{\partial y} + Gu = F,$$

odkud po úpravě dostaneme

$$\frac{\partial^2 u}{\partial x^2} + \dots = 0,$$

a tedy už je přímo v kanonickém tvaru. Opět jsme uvedli pouze parciální derivace druhého řádu.

Jestliže je $B \neq 0$, potom je také $\lambda \neq 0$. Obdobně jako v předchozím případě položíme

$$\xi = y - \lambda x.$$

Protože máme pouze jeden kořen charakteristické rovnice, položíme dále $\eta = x$. Máme tedy $U(\xi, \eta) = u(\eta, \xi + \lambda x)$. Spočítáme si parciální derivace.

$$\frac{\partial u}{\partial x} = \frac{\partial U}{\partial \xi} \cdot \frac{\partial \xi}{\partial x} + \frac{\partial U}{\partial \eta} \cdot \frac{\partial \eta}{\partial x} = -\lambda \frac{\partial U}{\partial \xi} + \frac{\partial U}{\partial \eta},$$

$$\frac{\partial u}{\partial y} = \frac{\partial U}{\partial \xi} \cdot \frac{\partial \xi}{\partial y} + \frac{\partial U}{\partial \eta} \cdot \frac{\partial \eta}{\partial y} = \frac{\partial U}{\partial \xi},$$

$$\frac{\partial^2 u}{\partial x^2} = \lambda^2 \frac{\partial^2 U}{\partial \xi^2} - 2\lambda \frac{\partial^2 U}{\partial x \partial y} + \frac{\partial^2 U}{\partial \eta^2},$$

$$\frac{\partial^2 u}{\partial x \partial y} = -\lambda \frac{\partial^2 U}{\partial \xi^2} + \frac{\partial^2 U}{\partial \xi \partial \eta},$$

$$\frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 U}{\partial \xi^2}.$$

Dosadíme tyto hodnoty do naší rovnice (10.23) a dostaneme

$$A \frac{\partial^2 u}{\partial x^2} + B \frac{\partial^2 u}{\partial x \partial y} + C \frac{\partial^2 u}{\partial y^2} + D \frac{\partial u}{\partial x} + E \frac{\partial u}{\partial y} + Gu - F = 0,$$

$$A \left(\lambda^2 \frac{\partial^2 U}{\partial \xi^2} - 2\lambda \frac{\partial^2 U}{\partial \xi \partial \eta} + \frac{\partial^2 U}{\partial \eta^2} \right) + B \left(-\lambda \frac{\partial^2 U}{\partial \xi^2} + \frac{\partial^2 U}{\partial \xi \partial \eta} \right) + C \left(\frac{\partial^2 U}{\partial \xi^2} \right) + \dots = 0,$$

(pro přehlednost opět uvádíme pouze parciální derivace druhého řádu, které jsou rozhodující). Po úpravě dostaneme

$$(A\lambda^2 - B\lambda + C)\frac{\partial^2 U}{\partial \xi^2} + (-2\lambda A + B)\frac{\partial^2 U}{\partial \xi \partial \eta} + A\frac{\partial^2 U}{\partial \eta^2} + \dots = 0.$$

Z charakteristické rovnice nám plyne $A\lambda^2 - B\lambda + C = 0$ a z jejího řešení $\lambda = \frac{B}{2A}$ plyne platnost rovnice $B - 2A\lambda = 0 \Rightarrow -2A\lambda + B = 0$. První dva koeficienty v předešlé rovnici jsou proto nulové. Protože jsme navíc předpokládali, že $A \neq 0$, můžeme rovnici vydělit koeficientem A a dostaneme

$$\frac{\partial^2 U}{\partial \eta^2} + \dots = 0.$$

Tím jsme získali kanonický tvar pro rovnici parabolického typu.

Postup si opět ukážeme na příkladu.

Příklad 10.34. Převeďte na kanonický tvar rovnici

$$4\frac{\partial^2 u}{\partial x^2} + 4\frac{\partial^2 u}{\partial x \partial y} + \frac{\partial^2 u}{\partial y^2} = 0$$

a najděte její řešení.

Řešení. Máme $A = 4, B = 4, C = 1$. Takže platí

$$B^2 - 4AC = 4^2 - 4 \cdot 4 \cdot 1 = 16 - 16 = 0.$$

Rovnice je parabolického typu. Rovnici přiřadíme její charakteristickou rovnici:

$$4\lambda^2 - 4\lambda + 1 = 0$$

a určíme si její kořeny. Dostaneme

$$\lambda_{1,2} = \frac{1}{2}.$$

Podle předchozího volíme

$$\xi = y - \frac{1}{2}x, \quad \eta = x.$$

Určíme si parciální derivace

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} &= \frac{1}{4} \frac{\partial^2 U}{\partial \xi^2} - \frac{\partial^2 U}{\partial \xi \partial \eta} + \frac{\partial^2 U}{\partial \eta^2}, \\ \frac{\partial^2 u}{\partial x \partial y} &= -\frac{1}{2} \frac{\partial^2 U}{\partial \xi^2} + \frac{\partial^2 U}{\partial \xi \partial \eta}, \\ \frac{\partial^2 u}{\partial y^2} &= \frac{\partial^2 U}{\partial \xi^2}. \end{aligned}$$

Po dosazení dostaneme

$$\begin{aligned} 4\frac{\partial^2 u}{\partial x^2} + 4\frac{\partial^2 u}{\partial x \partial y} + \frac{\partial^2 u}{\partial y^2} &= 0, \\ 4\left(\frac{1}{4} \frac{\partial^2 U}{\partial \xi^2} - \frac{\partial^2 U}{\partial \xi \partial \eta} + \frac{\partial^2 U}{\partial \eta^2}\right) + 4\left(-\frac{1}{2} \frac{\partial^2 U}{\partial \xi^2} + \frac{\partial^2 U}{\partial \xi \partial \eta}\right) + \left(\frac{\partial^2 U}{\partial \xi^2}\right) &= 0, \end{aligned}$$

$$\begin{aligned}\frac{\partial^2 U}{\partial \xi^2} - 4 \frac{\partial^2 U}{\partial \xi \partial \eta} + 4 \frac{\partial^2 U}{\partial \eta^2} - 2 \frac{\partial^2 U}{\partial \xi^2} + 4 \frac{\partial^2 U}{\partial \xi \partial \eta} + \frac{\partial^2 U}{\partial \xi^2} &= 0, \\ 4 \frac{\partial^2 U}{\partial \eta^2} &= 0, \\ \frac{\partial^2 U}{\partial \eta^2} &= 0.\end{aligned}$$

Budeme nyní hledat řešení této rovnice. Protože druhá derivace U podle η je nulová, integrací dostaneme

$$\begin{aligned}\frac{\partial U}{\partial \eta} &= \int \frac{\partial^2 U}{\partial \eta^2} d\eta = f(\xi), \\ U &= \int f(\xi) d\eta = \eta f(\xi) + g(\xi),\end{aligned}$$

kde f, g jsou funkce se spojitými parciálními derivacemi. Dosazením za ξ, η dostaneme řešení naší rovnice ve tvaru

$$u = x f\left(y - \frac{1}{2}x\right) + g\left(y - \frac{1}{2}x\right).$$

□

Mějme nyní rovnici (10.23) eliptického typu, tj. platí $B^2 - 4AC < 0$. Přímo z této podmínky nám vyplývá, že $A \neq 0, C \neq 0$. Charakteristická rovnice

$$A\lambda^2 - B\lambda + C = 0$$

má dva komplexní kořeny $\lambda_1 = \alpha + i\beta, \lambda_2 = \alpha - i\beta, \beta \neq 0$.

Položíme

$$\xi = y - \alpha x, \quad \eta = -\beta x.$$

Uurčíme si parciální derivace pro funkci $u(x, y) = U(\xi, \eta) = u\left(\frac{\eta}{-\beta}, \xi - \frac{\alpha}{\beta}\eta\right)$.

$$\begin{aligned}\frac{\partial u}{\partial x} &= -\alpha \frac{\partial U}{\partial \xi} - \beta \frac{\partial U}{\partial \eta}, \\ \frac{\partial u}{\partial y} &= \frac{\partial U}{\partial \xi}, \\ \frac{\partial^2 u}{\partial x^2} &= \alpha^2 \frac{\partial^2 U}{\partial \xi^2} + 2\alpha\beta \frac{\partial^2 U}{\partial \xi \partial \eta} + \beta^2 \frac{\partial^2 U}{\partial \eta^2}, \\ \frac{\partial^2 u}{\partial y^2} &= \frac{\partial^2 U}{\partial \xi^2}.\end{aligned}$$

Dosadíme tyto hodnoty do naší rovnice (10.23) a dostaneme

$$\begin{aligned}A \frac{\partial^2 u}{\partial x^2} + B \frac{\partial^2 u}{\partial x \partial y} + C \frac{\partial^2 u}{\partial y^2} + D \frac{\partial u}{\partial x} + E \frac{\partial u}{\partial y} + Gu - F &= 0, \\ A \left(\alpha^2 \frac{\partial^2 U}{\partial \xi^2} + 2\alpha\beta \frac{\partial^2 U}{\partial \xi \partial \eta} + \beta^2 \frac{\partial^2 U}{\partial \eta^2} \right) + B \left(\alpha^2 \frac{\partial^2 U}{\partial \xi^2} + 2\alpha\beta \frac{\partial^2 U}{\partial \xi \partial \eta} + \beta^2 \frac{\partial^2 U}{\partial \eta^2} \right) +\end{aligned}$$

$$+C \left(\frac{\partial^2 U}{\partial \xi^2} \right) + \dots = 0,$$

(pro přehlednost opět uvádíme pouze parciální derivace druhého řádu, které jsou rozhodující). Po úpravě dostaneme

$$(\alpha^2 A - \alpha B + C) \frac{\partial^2 U}{\partial \xi^2} + \beta(2\alpha A - B) \frac{\partial^2 U}{\partial \xi \partial \eta} + A\beta^2 \frac{\partial^2 U}{\partial \eta^2} + \dots = 0.$$

$\alpha + i\beta$ je kořen charakteristické rovnice a proto platí

$$\alpha = \frac{B}{2A}, \quad \beta^2 = \frac{4AC - B^2}{4A^2}.$$

Potom ale

$$2\alpha A - B = 0$$

a

$$\begin{aligned} \alpha^2 A - \alpha B + C &= \frac{B^2}{4A^2} A - \frac{B}{2A} B + C = \frac{B^2}{4A} - \frac{B^2}{2A} + C = \\ &= \frac{-B^2 + 4AC}{4A} = A\beta^2. \end{aligned}$$

Neboli koeficient u smíšené derivace je nulový a koeficienty u zbývajících se sobě rovnají. Tím jsme získali rovnici

$$\frac{\partial^2 U}{\partial \xi^2} + \frac{\partial^2 U}{\partial \eta^2} + \dots = 0,$$

která je kanonickým tvarem pro rovnici eliptického typu. Pokud jsou koeficienty rovnice (10.23) funkcemi, potom postupujeme analogicky. Musíme pouze rozlišit oblasti, ve kterých je rovnice stejného typu.

Definice 10.35. *Okrajová úloha* pro lineární parciální diferenciální rovnici 2.řádu je: Najít řešení $u = u(x, y)$ rovnice (10.21) v oblasti $\Omega \subset \mathbb{R}^2$, jestliže známe hodnoty řešení na hranici $\Gamma(\Omega)$.

$$u(x, y)|_{(x,y) \in \Gamma} = \varphi(x, y).$$

10.6.1 Řešené příklady

Příklad 10.36. Najděte kanonický tvar rovnice

$$\frac{\partial^2 u}{\partial x^2} + 8 \frac{\partial^2 u}{\partial x \partial y} + 15 \frac{\partial^2 u}{\partial y^2} = 0$$

a její řešení.

Řešení. Máme $A = 1, B = 8, C = 15$. Takže platí

$$B^2 - 4AC = 8^2 - 4 \cdot 1 \cdot 15 = 64 - 60 = 4 > 0.$$

Rovnice je hyperbolického typu. Rovnici přiřadíme charakteristickou rovnici:

$$\lambda^2 - 8\lambda + 15 = 0$$

a určíme si její kořeny. Dostaneme

$$\lambda_1 = 3, \quad \lambda_2 = 5.$$

Podle předchozího volíme

$$\xi = y - 3x, \quad \eta = y - 5x.$$

Určíme si parciální derivace

$$\frac{\partial^2 u}{\partial x^2} = 9 \frac{\partial^2 U}{\partial \xi^2} + 30 \frac{\partial^2 U}{\partial \xi \partial \eta} + 25 \frac{\partial^2 U}{\partial \eta^2},$$

$$\frac{\partial^2 u}{\partial x \partial y} = -3 \frac{\partial^2 U}{\partial \xi^2} - 8 \frac{\partial^2 U}{\partial \xi \partial \eta} - 5 \frac{\partial^2 U}{\partial \eta^2},$$

$$\frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 U}{\partial \xi^2} + 2 \frac{\partial^2 U}{\partial \xi \partial \eta} + \frac{\partial^2 U}{\partial \eta^2}.$$

Po dosazení dostaneme

$$\frac{\partial^2 U}{\partial \xi \partial \eta} = 0.$$

Budeme nyní hledat řešení této rovnice. Zvolme

$$\frac{\partial U}{\partial \eta} = v.$$

Potom

$$\frac{\partial v}{\partial \xi} = 0,$$

a proto je $v = f(\eta)$, kde f je libovolná funkce. Takže máme

$$\frac{\partial U}{\partial \eta} = f(\eta),$$

$$U = \int f(\eta) d\eta = F(\eta) + G(\xi),$$

kde F, G jsou funkce se spojitými parciálními derivacemi. Dosazením za ξ, η dostaneme řešení naší rovnice ve tvaru

$$u = F(y - 5x) + G(y - 3x).$$

□

Příklad 10.37. Převeďte na kanonický tvar rovnici

$$4 \frac{\partial^2 u}{\partial x^2} + 12 \frac{\partial^2 u}{\partial x \partial y} + 13 \frac{\partial^2 u}{\partial y^2} = 0.$$

Řešení. Máme $A = 4, B = 12, C = 13$. Takže platí

$$B^2 - 4AC = 12^2 - 4 \cdot 4 \cdot 13 = 144 - 208 = -64.$$

Rovnice je eliptického typu. Rovnici přiřadíme její charakteristickou rovnici:

$$4\lambda^2 - 12\lambda + 13 = 0$$

a určíme si její kořeny. Dostaneme

$$\lambda_{1,2} = \frac{12 \pm \sqrt{144 - 4 \cdot 4 \cdot 13}}{8} = \frac{12 \pm \sqrt{-64}}{8} = \frac{3}{2} \pm i.$$

Máme $\alpha = \frac{3}{2}, \beta = 1$. Podle předchozího volíme

$$\xi = y - \frac{3}{2}x, \quad \eta = -x.$$

Určíme si parciální derivace

$$\frac{\partial^2 u}{\partial x^2} = \frac{9}{4} \frac{\partial^2 U}{\partial \xi^2} + 3 \frac{\partial^2 U}{\partial \xi \partial \eta} + \frac{\partial^2 U}{\partial \eta^2},$$

$$\frac{\partial^2 u}{\partial x \partial y} = -\frac{3}{2} \frac{\partial^2 U}{\partial \xi^2} - \frac{\partial^2 U}{\partial \xi \partial \eta},$$

$$\frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 U}{\partial \xi^2}.$$

Po dosazení dostaneme

$$4 \frac{\partial^2 u}{\partial x^2} + 12 \frac{\partial^2 u}{\partial x \partial y} + 13 \frac{\partial^2 u}{\partial y^2} = 0,$$

$$4 \left(\frac{9}{4} \frac{\partial^2 U}{\partial \xi^2} + 3 \frac{\partial^2 U}{\partial \xi \partial \eta} + \frac{\partial^2 U}{\partial \eta^2} \right) + 12 \left(-\frac{3}{2} \frac{\partial^2 U}{\partial \xi^2} - \frac{\partial^2 U}{\partial \xi \partial \eta} \right) + 13 \left(\frac{\partial^2 U}{\partial \xi^2} \right) = 0,$$

$$9 \frac{\partial^2 U}{\partial \xi^2} + 12 \frac{\partial^2 U}{\partial \xi \partial \eta} + 4 \frac{\partial^2 U}{\partial \eta^2} - 18 \frac{\partial^2 U}{\partial \xi^2} - 12 \frac{\partial^2 U}{\partial \xi \partial \eta} + 13 \left(\frac{\partial^2 U}{\partial \xi^2} \right) = 0,$$

$$4 \frac{\partial^2 U}{\partial \xi^2} + 4 \frac{\partial^2 U}{\partial \eta^2} = 0,$$

$$\frac{\partial^2 U}{\partial \xi^2} + \frac{\partial^2 U}{\partial \eta^2} = 0.$$

□

10.7 Metoda konečných diferencí pro parciální diferenciální rovnice

Nyní se budeme věnovat některým numerickým metodám pro určení řešení parciálních diferenciálních rovnic.

Definice 10.38. *Okrajová úloha* pro lineární parciální diferenciální rovnici 2.řádu je: Najít řešení $u = u(x, y)$ rovnice (10.22) v oblast $\Omega \subset \mathbb{R}^2$, jestliže známe hodnoty řešení na hranici $\Gamma(\Omega)$.

$$u(x, y)|_{(x,y) \in \Gamma} = \varphi(x, y).$$

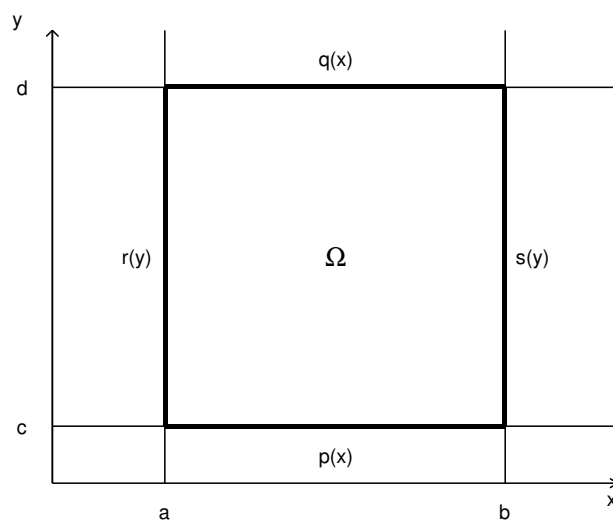
Vezměme si nejjednodušší případ: Mějme parciální lineární diferenciální rovnici eliptického typu

$$-\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} + \sigma(x, y)u = f(x, y), \quad (10.26)$$

kde $u = u(x, y)$, $\Omega = \{(x, y) : a \leq x \leq b, c \leq y \leq d\}$, $\sigma(x, y) \geq 0 \forall x, y \in \Omega$, σ, f jsou spojité na Ω .

Nechť je splněna tzv. Dirichletova okrajová podmínka na hranicích oblasti Ω :

$$\begin{aligned} u(x, c) = p(x), \quad u(x, d) = q(x), \quad a \leq x \leq b, \\ u(a, y) = r(y), \quad u(b, y) = s(y), \quad c \leq y \leq d, \\ p(a) = r(c), \quad p(b) = s(c), \quad q(a) = r(d), \quad q(b) = s(d). \end{aligned}$$



Obr. 10.3: Hranice oblasti

Poslední řádek nám zaručuje spojitost okrajových podmínek v „rozích“ oblasti Ω . Viz obrázek 10.3

Opět si vytvoříme síť na oblasti Ω (nejčastěji se používají čtvercové a nebo obdélníkové sítě).

$$\begin{aligned} x_i = a + ih, \quad i = 0, 1, \dots, n, n+1, \quad h = \frac{b-a}{n+1}, \\ y_j = c + jk, \quad j = 0, 1, \dots, m, m+1, \quad k = \frac{d-c}{m+1}. \end{aligned}$$

Uzly jsou pak body (x_i, y_j) . Označme $u_{ij} = u(x_i, y_j)$. Za předpokladu, že platí

$$\left| \frac{\partial^4 u}{\partial x^4} \right| \leq M_4, \quad \left| \frac{\partial^4 u}{\partial y^4} \right| \leq M_4, \quad M_4 \in \mathbb{R}, \quad M_4 < \infty,$$

t.j. $u(x, y)$ je spojitá a má ohraničené parciální derivace do čtvrtého řádu včetně. Pak si můžeme vyjádřit derivace pomocí diferencí a dostaneme

$$-\frac{\partial^2 u(x_i, y_j)}{\partial x^2} = - \left(\frac{u_{i+1,j} - 2u_{ij} + u_{i-1,j}}{h^2} - \frac{h^2}{12} u^{(4)}(\xi_i, y_j) \right),$$

$$-\frac{\partial^2 u(x_i, y_j)}{\partial y^2} = - \left(\frac{u_{i,j+1} - 2u_{ij} + u_{i,j-1}}{k^2} - \frac{k^2}{12} u^{(4)}(x_i, \eta_j) \right),$$

kde $x_{i-1} < \xi_i < x_{i+1}$, $y_{j-1} < \eta_j < y_{j+1}$. Jestliže předpokládáme spojitost $u(x, y)$, pak pro dostatečně malé h, k můžeme zanedbat chybové funkce. Potom dosazením do (10.26) dostáváme pro $i = 1, 2, \dots, n$, $j = 1, 2, \dots, m$:

$$\frac{2u_{ij} - u_{i+1,j} - u_{i-1,j}}{h^2} + \frac{2u_{ij} - u_{i,j+1} - u_{i,j-1}}{k^2} + \sigma_{ij} u_{ij} = f_{ij}.$$

Vynásobením této rovnice koeficientem hk dostaneme

$$\left(2 \left(\frac{h}{k} + \frac{k}{h} \right) + hk\sigma_{ij} \right) u_{ij} - \frac{k}{h}(u_{i+1,j} + u_{i-1,j}) - \frac{h}{k}(u_{i,j+1} + u_{i,j-1}) = hkf_{ij}.$$

Dosadíme podle počátečních podmínek

$$u_{i,0} = p_i, \quad u_{i,m+1} = q_i, \quad u_{0,j} = r_j, \quad u_{n+1,j} = s_j,$$

kde $p_i = p(x_i)$, $q_i = q(x_i)$, $r_j = r(y_j)$, $s_j = s(y_j)$. Dostaneme tak soustavu lineárních algebraických rovnic a po jejím vyřešení získáme hodnoty u_{ij} . V případě pravidelné čtvercové sítě, t.j. $h = k$, se tato soustava dále zjednoduší na tvar

$$(4 + h^2\sigma_{ij}) u_{ij} - u_{i+1,j} - u_{i-1,j} - u_{i,j+1} - u_{i,j-1} = h^2 f_{ij}. \quad (10.27)$$

Všimněte si, že matice soustavy je v obou případech pro $\sigma_{ij} \neq 0$ diagonálně dominantní, a proto můžeme použít i iterační metody řešení. V případě $\sigma_{ij} \equiv 0$ jde o soustavu, kde je diagonální dominance neostrá, ale je ostrá pro všechny rovnice, v nichž je alespoň jeden hraniční bod. Pro takovéto matice nám bude opět konvergovat Gauss-Seidelova metoda, ale obecně dosti pomalu. Při větším počtu rovnic se vyplatí používat relaxační nebo superrelaxační metodu nebo metodu sdružených gradientů, které nám podstatně zlepšují konvergenci a hlavně rychlost výpočtu. Zvláště v tomto případě, kdy matice koeficientů soustavy je řídká – tj. obsahuje velké množství nulových prvků, tady máme v každé rovnici nejvýše pět nenulových koeficientů. Pro takovéto případy je zvlášť vhodná metoda sdružených gradientů.

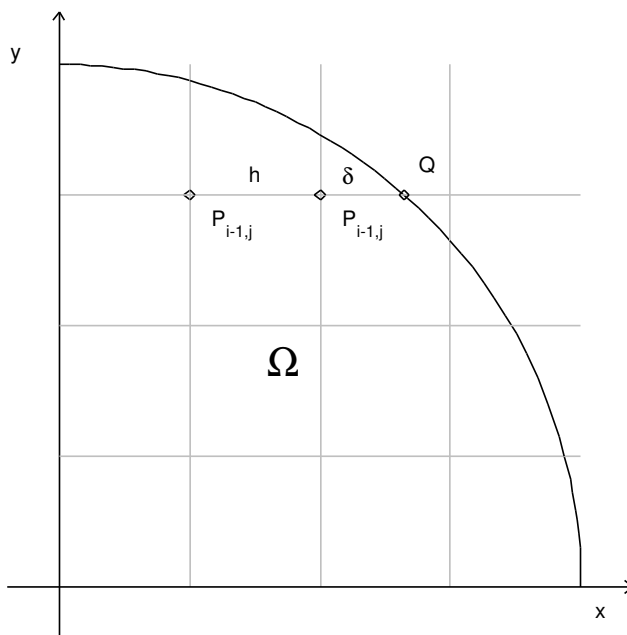
Analogický postup můžeme použít pro numerické řešení všech lineárních parciálních diferenciálních rovnic druhého řádu. Postup řešení je vždy stejný: derivace nahradíme diferencemi a hledáme řešení soustavy lineárních algebraických rovnic.

Pro další parciální derivace se používají aproximace:

$$\begin{aligned}\frac{\partial u(x_i, y_j)}{\partial x} &= \frac{u_{i+1,j} - u_{i,j}}{h}, & \frac{\partial u(x_i, y_j)}{\partial y} &= \frac{u_{i,j+1} - u_{i,j}}{k}, \\ \frac{\partial u(x_i, y_j)}{\partial x} &= \frac{u_{i,j} - u_{i-1,j}}{h}, & \frac{\partial u(x_i, y_j)}{\partial y} &= \frac{u_{i,j} - u_{i,j-1}}{k}, \\ \frac{\partial u(x_i, y_j)}{\partial x} &= \frac{u_{i+1,j} - u_{i-1,j}}{2h}, & \frac{\partial u(x_i, y_j)}{\partial y} &= \frac{u_{i,j+1} - u_{i,j-1}}{2k}. \\ \frac{\partial^2 u(x_i, y_j)}{\partial x \partial y} &= \frac{u_{i+1,j+1} - u_{i+1,j-1} - u_{i-1,j+1} + u_{i-1,j-1}}{2h2k}.\end{aligned}$$

Problémy při řešení mohou vzniknout, pokud oblast Ω není obdélníková.

Definice 10.39. Bod $P_{ij} = (x_i, y_j)$ sítě na oblasti Ω nazveme *vnitřním*, jestliže všechny body úseček spojujících jej se sousedními body $P_{i\pm 1,j}$, $P_{i,j\pm 1}$ leží v Ω a nazveme jej *hraničním* v opačném případě.



Obr. 10.4: Problém u metody konečných diferencí

Pro určení hodnoty funkce u v hraničních bodech se nejčastěji používá lineární interpolace či extrapolace.

Nechť hraniční bod Q leží na spojnici uzlů $P_{ij}, P_{i+1,j}$, ve vzdálenosti δ od bodu P_{ij} . Nechť se hodnoty funkce u mění lineárně podél této spojnice. Potom

$$\frac{u(Q) - u(P_{ij})}{\delta} = \frac{u(P_{ij}) - u(P_{i-1,j})}{h}.$$

Protože podle definice 10.38 můžeme psát $u(Q) = \varphi(Q)$ a dále $u(P_{ij}) = u_{ij}$, tak po úpravě dostaneme

$$\varphi(Q) = \left(1 + \frac{\delta}{h}\right) u_{ij} - \frac{\delta}{h} u_{i-1,j}$$

a nebo

$$u_{ij} = \frac{h\varphi(Q) + \delta u_{i-1,j}}{h + \delta} = \frac{h}{h + \delta} \varphi(Q) + \frac{\delta}{h + \delta} u_{i-1,j},$$

protože známe hodnotu v bodě Q a potřebujeme dopočítat hodnotu v bodě P_{ij} .

Příklad 10.40. Metodou konečných diferencí řešte okrajovou úlohu

$$u_{xx} + u_{yy} = 8x$$

$$u(x, 0) = x^3, \quad u(0, y) = 0, \quad u(x, y)|_{x^2+y^2=10} = 10x(y+1),$$

kde oblast Ω je vnitřní část čtvrtkruhu

$$x \geq 0, \quad y \geq 0, \quad x^2 + y^2 \leq 10.$$

Řešení. Rovnici si upravíme na stejný tvar jako (10.26)

$$-\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = -8x.$$

Zvolme čtvercovou síť s krokem $h = 1$. Potom máme hraniční body $u(0, 0) = 0, u(1, 0) = 1, u(2, 0) = 8, u(3, 0) = 27, u(0, 1) = u(0, 2) = u(0, 3) = 0, u(1, 3) = 40, u(3, 1) = 60$. Ještě potřebujeme znát hodnotu v bodě $P_{2,2}$. Protože $Q = (2.449; 2)$, $\varphi(Q) = 73.485$, $\delta = 0.449$, tak lineární interpolací dostaneme

$$u_{2,2} = \frac{1 \cdot 73.485 + 0.449 \cdot u_{1,2}}{1 + 0.449} = 50.697 + 0.310u_{1,2}.$$

Nyní pro 3 vnitřní uzly sestavíme síťové rovnice podle (10.27), přitom hraniční uzly jsou podtrženy.

$$4u_{1,1} - \underline{u_{0,1}} - \underline{u_{2,1}} - \underline{u_{1,0}} - \underline{u_{1,2}} = -8,$$

$$4u_{1,2} - u_{1,1} - \underline{u_{1,3}} - \underline{u_{0,2}} - \underline{u_{2,2}} = -8,$$

$$4u_{2,1} - u_{1,1} - \underline{u_{3,1}} - \underline{u_{2,0}} - \underline{u_{2,2}} = -16.$$

Pak přidáním odvozeného vztahu pro $u_{2,2}$ dostaneme soustavu 4 rovnic o čtyřech neznámých. Po úpravě

$$u_{1,1} = \frac{1}{4}(0 + u_{2,1} + 1 + u_{1,2}) - \frac{1}{4}8,$$

$$\begin{aligned}u_{2,1} &= \frac{1}{4}(u_{1,1} + u_{2,2} + 60 + 8) - \frac{1}{4}16, \\u_{1,2} &= \frac{1}{4}(0 + 40 + u_{1,1} + u_{2,2}) - \frac{1}{4}8, \\u_{2,2} &= 50.697 + 0.310u_{1,2}.\end{aligned}$$

Jejím řešením je pak

$$\begin{aligned}u_{1,1} &= 12.384, \\u_{2,1} &= 30.768, \\u_{1,2} &= 25.768, \\u_{2,2} &= 58.688.\end{aligned}$$

Další postup je pak obvyklý, tj. zmenšíme krok a opakujeme výpočet, až se nám odchylky v uzlových bodech ustálí.

Nechť tedy je $h = \frac{1}{2}$. Potom dostaneme síťové rovnice pro 19 vnitřních bodů a ještě musíme dopočítat hodnoty 5 hraničních uzlů lineární interpolací. Dostaneme tedy soustavu 24 rovnic o 24 neznámých. Přitom každá rovnice obsahuje nejvýše 5 nenulových hodnot proměnných.

Pro hraniční uzly máme

$$\begin{aligned}u_{6,1} &= 37.651 + 0.196u_{5,1}, \\u_{5,3} &= 44.388 + 0.223u_{4,3}, \\u_{4,4} &= 38.717 + 0.473u_{3,4}, \\u_{3,5} &= 36.196 + 0.446u_{2,5},\end{aligned}$$

a pro poslední hodnotu bereme buď

$$u_{1,6} = \frac{1}{2}(u_{0,6} + u_{2,6}) = 20,$$

nebo si tuto hodnotu vypočítáme, přitom bereme sousední uzly po vertikále (doposud jsme je brali po horizontále), pak dostaneme rovnici

$$u_{1,6} = 16.567 + 0.196u_{1,5}.$$

Pro vnitřní uzly pak budeme mít soustavu

$$\begin{aligned}-4u_{11} + u_{01} + u_{10} + u_{21} + u_{12} &= 1, \\-4u_{12} + u_{02} + u_{11} + u_{22} + u_{13} &= 1, \\-4u_{13} + u_{03} + u_{12} + u_{23} + u_{14} &= 1, \\-4u_{14} + u_{04} + u_{13} + u_{24} + u_{15} &= 1, \\-4u_{15} + u_{05} + u_{14} + u_{25} + u_{16} &= 1, \\-4u_{21} + u_{11} + u_{20} + u_{31} + u_{22} &= 2,\end{aligned}$$

$$\begin{aligned}
-4u_{22} + u_{12} + u_{21} + u_{32} + u_{23} &= 2, \\
-4u_{23} + u_{13} + u_{22} + u_{33} + u_{24} &= 2, \\
-4u_{24} + u_{14} + u_{23} + u_{34} + u_{25} &= 2, \\
-4u_{25} + u_{15} + u_{24} + u_{35} + u_{26} &= 2, \\
-4u_{31} + u_{21} + u_{30} + u_{41} + u_{32} &= 3, \\
-4u_{32} + u_{22} + u_{31} + u_{42} + u_{33} &= 3, \\
-4u_{33} + u_{23} + u_{32} + u_{43} + u_{34} &= 3, \\
-4u_{34} + u_{24} + u_{33} + u_{44} + u_{35} &= 3, \\
-4u_{41} + u_{31} + u_{40} + u_{51} + u_{42} &= 4, \\
-4u_{42} + u_{32} + u_{41} + u_{52} + u_{43} &= 4, \\
-4u_{43} + u_{33} + u_{42} + u_{53} + u_{44} &= 4, \\
-4u_{51} + u_{41} + u_{50} + u_{61} + u_{52} &= 5, \\
-4u_{52} + u_{42} + u_{51} + u_{62} + u_{53} &= 5.
\end{aligned}$$

Tím máme celou soustavu hotovou a zbývá jí „jen“ vyřešit. \square

Pro rovnice jiných typů je postup analogický. Vždy požadujeme, aby výsledná soustava lineárních algebraických rovnic byla jednoznačně řešitelná. Z této podmínky plynou i požadavky na tvar rovnice, které nám potom zaručí jednoznačnost a konvergenci řešení.

Postup si ukážeme na rovnici parabolického typu, půjde o zobecnění rovnice vedení tepla, kdy přidáme ještě další člen,

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(x, t), \quad (10.28)$$

počáteční podmínku

$$u(x, 0) = \varphi(x), \quad (10.29)$$

a okrajové podmínky

$$u(0, t) = \mu_1(t), \quad (10.30)$$

$$u(l, t) = \mu_2(t), \quad (10.31)$$

kde f, φ, μ_1, μ_2 jsou zadané spojité funkce, $x \in \langle 0, l \rangle$, $t \in [0, T]$.

Rozdělíme si interval $\langle 0, l \rangle$ na n dílů délky $\Delta x = \frac{l}{n}$. Na ose t si zvolíme velikost kroku Δt tak, aby platilo $\Delta t = \rho \Delta x^2$. Důvod bude zřejmý z dalšího výkladu. Derivace v rovnici (10.28) nahradíme diferencemi a dostaneme

$$\frac{u_{i,k+1} - u_{i,k}}{\Delta t} = \frac{u_{i+1,k} - 2u_{i,k} + u_{i-1,k}}{\Delta x^2} + f(x_i, t_k).$$

Nyní vynásobíme celou rovnici výrazem $\Delta t = \rho \Delta x^2$ a dostaneme po úpravě

$$u_{i,k+1} = (1 - 2\rho)u_{i,k} + \rho(u_{i-1,k} + u_{i+1,k}) + \Delta t f(x_i, t_k). \quad (10.32)$$

Dosazením do počátečních a okrajových podmínek dostaneme

$$u_{i,0} = \varphi(x_i), u_{0,k} = \mu_1(t_k), u_{l,k} = \mu_2(t_k). \quad (10.33)$$

Tím jsme získali soustavu lineárních algebraických rovnic. Pro její řešitelnost je třeba volit $\varrho \leq \frac{1}{2}$. V opačném případě nastává numerická nestabilita a výpočet se bude výrazně lišit od skutečnosti.

10.8 Metoda konečných prvků pro parciální diferenciální rovnice

10.8.1 Základní pojmy

Základ metody je stejný jako v části 9.

Hledáme řešení okrajové úlohy pro eliptickou rovnici metodou konečných prvků.

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

kde $(x, y) \in \Omega$. Jde o tzv. Poissonovu rovnici. Mějme homogenní okrajové podmínky, tj. $u(x, y) = 0 \quad \forall (x, y) \in \Gamma(\Omega)$. Užitím Greenovy věty se dá dokázat, že funkcionál

$$F(w) = \frac{1}{2} \int \int_{\Omega} \left[\left(\frac{\partial w}{\partial x} \right)^2 + \left(\frac{\partial w}{\partial y} \right)^2 - 2wf \right] dx dy$$

je minimální, když w je řešením naší úlohy.

Protože $\text{grad } w = (w'_x, w'_y)$, můžeme psát

$$F(w) = \frac{1}{2} \int \int_{\Omega} [(\text{grad } w)^2 - 2wf] dx dy.$$

Přitom opět předpokládáme, že $w = w(x, y)$ probíhá všechny funkce splňující danou okrajovou podmínku a mající derivace integrovatelné v kvadrátě.

Předpokládáme dále, že hranice Γ je tvořena uzavřeným polygonem složeným s konečného počtu částí. Provedeme triangulaci oblasti Ω – rozdělíme ji trojúhelníky, které mají nejvýše jeden společný vrchol nebo jednu společnou hranu a tvoří pokrytí Ω . Vrcholy si očíslováme $P_j, j \in \mathcal{J}$.

Při hledání minima se opět omezíme na množinu funkcí spojitých na Ω , rovných nule na Γ a lineárních na každém trojúhelníku. (V obecném případě můžeme použít i jinou množinu funkcí než lineárních, například polynomy stupně $k, k > 1, \dots$. Budeme pouze požadovat spojitost na oblasti Ω a rovnost nule na Γ . Na každém z trojúhelníků pak budeme pracovat s funkcí z vybrané množiny.)

Analogicky jako v předchozím případě můžeme hledanou funkci vyjádřit ve tvaru

$$v(P) = \sum_{j \in \mathcal{J}} v_j \cdot p_j(P),$$

kde bázové funkce $p_j(P)$ mají tvar pláště pyramidy o výšce jedna s podstavou složenou ze všech trojúhelníků obsahujících vrchol P . Platí

$$\text{grad } v = \sum_{j \in \mathcal{J}} v_j \cdot \text{grad } p_j$$

a dostáváme

$$F(v) = \frac{1}{2} \int \int_{\Omega} \left[\left(\sum_{j \in \mathcal{J}} v_j \cdot \text{grad } p_j \right)^2 - 2f \sum_{j \in \mathcal{J}} v_j \cdot p_j \right] dx dy$$

a opět obdržíme kvadratickou funkci parametrů v_j a hledáme její minimum. Soustavu

$$\frac{\partial F}{\partial v_i} = 0, \quad i = 1, 2, \dots, n,$$

si přepíšeme do tvaru

$$\sum_{j \in \mathcal{J}} v_j \int \int_{\Omega} (\text{grad } p_j \cdot \text{grad } p_i) dx dy = \int \int_{\Omega} (f p_i) dx dy, \quad i = 1, 2, \dots, n$$

a opět jsme získali soustavu lineárních algebraických rovnic. A zase zde bude hrát důležitou roli, že funkce p_i mají malý nosič.

Metodu konečných prvků charakterizují tři základní vlastnosti:

1. Diskretizace oblasti Ω , kdy se Ω vyjádří jako sjednocení konečného počtu zvolených podmnožin – podoblastí. Tyto podmnožiny se nazývají *konečné prvky*. Obvykle to jsou trojúhelníky. Jen vyjíměčně se používají čtyřúhelníky či jiné útvary. V případě, že konečné prvky jsou trojúhelníky, pak se místo o diskretizace mluví o *triangulaci* oblasti Ω .
2. Volba prostoru funkcí, které jsou na každém prvku polynomem zvoleného stupně. Tento prostor budeme značit \mathcal{V} .
3. Existence takové báze prostoru \mathcal{V} , že bázové funkce mají malý nosič – tj. jsou nenulové pouze na několika sousedních prvcích.

Řešením dané okrajové úlohy pak rozumíme funkci z \mathcal{V} , která je určena jako lineární kombinace zvolených bázových funkcí.

Ukážeme si nyní použití stejné metody pro obecnější typ rovnice.

Budeme se zabývat metodou konečných prvků pro okrajovou úlohu pro eliptickou rovnici druhého řádu dvou proměnných na oblasti $\Omega \in \mathbb{R}^2$

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + a(x, y)u = f(x, y), \quad (10.34)$$

s okrajovou podmínkou

$$u(x, y)|_{(x, y) \in \Gamma(\Omega)} = \varphi(x, y). \quad (10.35)$$

Nebudeme se zabývat obecnou diskretizací, ale omezíme se na rozklad oblasti $\Omega \in \mathbb{R}^2$ na systém disjunktních trojúhelníků.

Triangulace:

Oblast Ω aproximujeme sjednocením konečného počtu disjunktních trojúhelníků. Hranice $\Gamma(\Omega)$ je potom aproximována polygonem – lomenou čarou.

Množinu trojúhelníků $\tau = \{T_1, T_2, \dots, T_k\}$ budeme nazývat *přípustnou triangulací*, jestliže platí:

1. $\Omega \cap \left(\bigcup_{i=1}^k T_i \right) \subset \Omega$.
2. Jsou-li T_a, T_b dva různé trojúhelníky triangulace τ , pak jejich vnitřky mají prázdný průnik.
3. $\forall i = 1, 2, \dots, k$ je každá strana T_i buď částí hranice Γ a nebo stranou jiného trojúhelníka z τ .

Trojúhelníky, které mají společnou stranu, se nazývají sousední. Bez omezení obecnosti budeme dále předpokládat, že nejvýše jedna strana trojúhelníka T_i je částí hranice Γ – tento předpoklad dodáváme pro jednodušší konstrukci algoritmu řešení. Při dostatečně jemném dělení oblasti Ω je tato podmínka splněna automaticky.

Konstrukce vhodné přípustné triangulace oblast Ω není vůbec snadnou záležitostí. Je přitom třeba dodržovat tyto zásady:

1. Nepoužívat trojúhelníky s velmi malými nebo velmi velkými vnitřními úhly.
2. V těch částech oblasti Ω , kde se očekávají velké změny v chování hledaného řešení dané úlohy, je nutno volit jemnější triangulaci.

Kromě vrcholů trojúhelníků T_i , $i = 1, \dots, k$, se někdy používají při konstrukci řešení další body trojúhelníku T_i , jako jsou středy stran, těžiště. Všem takovým bodům se souhrně říká uzly triangulace. Ty uzly, které leží na hranici se nazývají *hraniční*, zbývající jsou *vnitřní*. Strana trojúhelníka, která patří hranici $\Gamma(\Omega)$ se nazývá *hraniční*.

V uzlech triangulace zadáváme hodnoty koeficientů rovnice či okrajových podmínek a hodnoty pravých stran rovnice 10.34. Současně v nich hledáme hodnoty přibližného řešení. Všem těmto hodnotám říkáme *uzlové parametry*.

K dosažení vyšší přesnosti je třeba provést jemnější triangulaci, tj. zvolit menší trojúhelníky. Samozřejmě toto má smysl pouze tehdy, když je řešení naší okrajové úlohy dostatečně hladké.

Bázové funkce

Za uzly triangulace budeme volit pouze vrcholy trojúhelníků. Vrcholy trojúhelníka T_s označme M_1^s, M_2^s, M_3^s . Číslování provádíme vždy v *kladném smyslu*, tj. proti směru pohybu

hodinových ručiček, a začínáme vždy zprava, respektive vpravo dole, pokud máme vpravo dva vrcholy. Každý uzel má tedy lokální index, vázaný na příslušný trojúhelník, a současně i globální index, vázaný na místo daného vrcholu v pořadí všech vrcholů. Hovoříme pak o lokálním a globálním číslování.

Jedna z možností globálního číslování je $M_i^s \equiv M_{si}$, $i = 1, 2, 3$; $s = 1, 2, \dots, N$, kde N je počet všech uzlů příslušné triangulace.

Nechť M_n je uzel triangulace s globálním indexem n . Definujeme si po částech lineární báze funkce $v_n = v_n(x, y)$ následovně:

1. Na každém trojúhelníku T_s s vrcholem M_n je v_n lineárním polynomem tvaru

$$N^s(x, y) = a^s + b^s x + c^s y, \quad a^s, b^s, c^s \in \mathbb{R}.$$

2. Funkce v_n splňují vrcholové podmínky

$$v_n(M_n) = 1, \quad v_n(M_m) = 0 \quad \forall m \neq n.$$

3. v_n je nenulová pouze na těch trojúhelnících, jejichž společným vrcholem je uzel M_n . Tyto trojúhelníky tvoří *nosič* funkce v_n . Všude jinde je funkce v_n identicky rovna nule.

Takto definované báze funkce v_1, v_2, \dots, v_N jsou spojité na $\bar{\Omega} = \Omega \cup \Gamma(\Omega)$ a jsou lineárně nezávislé. Lineární prostor všech lineárních kombinací báze funkcí označíme \mathcal{V}_∞ . Funkcím z \mathcal{V}_∞ říkáme lineární *splajny*. Libovolnou funkci z \mathcal{V}_∞ budeme označovat $v = v(x, y)$. Tato funkce je vždy spojitá a lze ji vyjádřit ve tvaru lineární kombinace báze funkcí

$$v(x, y) = \sum_{n=1}^N \alpha_n v_n(x, y), \quad \alpha_n \in \mathbb{R}.$$

Konstrukce báze funkce

Souřadnice vrcholů trojúhelníka T_s označíme $M_1^s = (x_1^s, y_1^s)$, $M_2^s = (x_2^s, y_2^s)$, $M_3^s = (x_3^s, y_3^s)$. Restrikce 3 báze funkcí příslušných těmto vrcholům na T_s označíme

$$N_1^s = a_1^s + b_1^s x + c_1^s y,$$

$$N_2^s = a_2^s + b_2^s x + c_2^s y,$$

$$N_3^s = a_3^s + b_3^s x + c_3^s y.$$

Tyto funkce jsou jednoznačně určeny vrcholovými podmínkami:

$$N_1^s : \quad N_1^s(M_1^s) = 1, \quad N_1^s(M_2^s) = 0, \quad N_1^s(M_3^s) = 0,$$

$$N_2^s : \quad N_2^s(M_1^s) = 0, \quad N_2^s(M_2^s) = 1, \quad N_2^s(M_3^s) = 0,$$

$$N_3^s : \quad N_3^s(M_1^s) = 0, \quad N_3^s(M_2^s) = 0, \quad N_3^s(M_3^s) = 1.$$

Koeficienty lineárních funkcí jsou pak určeny vztahy

$$\begin{aligned} a_1^s &= \frac{1}{D^s}(x_2^s y_3^s - x_3^s y_2^s), \quad b_1^s = \frac{1}{D^s}(y_2^s - y_3^s), \quad c_1^s = \frac{1}{D^s}(x_3^s - x_2^s), \\ a_2^s &= \frac{1}{D^s}(x_3^s y_1^s - x_1^s y_3^s), \quad b_2^s = \frac{1}{D^s}(y_3^s - y_1^s), \quad c_2^s = \frac{1}{D^s}(x_1^s - x_3^s), \\ a_3^s &= \frac{1}{D^s}(x_1^s y_2^s - x_2^s y_1^s), \quad b_3^s = \frac{1}{D^s}(y_1^s - y_2^s), \quad c_3^s = \frac{1}{D^s}(x_2^s - x_1^s), \end{aligned}$$

kde

$$D^s = \det \begin{pmatrix} 1 & x_1^s & y_1^s \\ 1 & x_2^s & y_2^s \\ 1 & x_3^s & y_3^s \end{pmatrix}.$$

D^s je dvojnásobek obsahu trojúhelníka T_s .

Diskretizace úlohy

Na oblasti $\Omega \subset \mathbb{R}^2$ máme okrajovou úlohu

$$-\operatorname{div}(p(x, y)\operatorname{grad} u) + q(x, y)u = f(x, y),$$

$$\sigma u + p \frac{\partial u}{\partial n} \Big|_{\Gamma} = g.$$

Zvolíme si triangulaci oblasti Ω a sestrojíme prostor $\mathcal{V}_{\infty}^{\langle}$ funkcí spojitých na $\bar{\Omega}$ a lineárních na každém trojúhelníku T_s zvolené triangulace. To znamená, že sestrojíme systém bázových funkcí v_1, v_2, \dots, v_N . Přibližné řešení úlohy v prostoru \mathcal{V}_1^h je taková po částech lineární funkce $u_h \in \mathcal{V}_1^h$, pro kterou platí

$$a(u_h, v_h) = F(v_h)$$

pro každou funkci $v_h \in \mathcal{V}_{\infty}^{\langle}$. Funkce u_h lze vyjádřit jako lineární kombinace po částech lineárních bázových funkcí

$$u_h = \sum_{n=1}^N U_n v_n,$$

kde $U_n = u_h(M_n) \approx u(M_n)$ jsou hledané parametry, které dostaneme řešením soustavy lineárních algebraických rovnic

$$A_h U_h = F_h,$$

kde $U_h = (U_1, U_2, \dots, U_N)^T$, $F_h = (F_1, F_2, \dots, F_N)^T$.

$$F_n = F(v_n) = \int \int_{\Omega} f(x, y)v_n(x, y) dx dy + \int_{\delta\Omega} g(x, y)v_n(x, y) ds, \quad n = 1, 2, \dots, N,$$

a A_h je symetrická pozitivně definitní matice s prvky

$$a_{nk} = a(v_n, v_k) = \int \int_{\Omega} [p(x, y)\operatorname{grad} v_n(x, y) \cdot \operatorname{grad} v_k(x, y) + q(x, y)v_n(x, y)v_k(x, y)] dx dy +$$

$$+ \int_{\delta\Omega} \sigma(x, y) v_n(x, y) v_k(x, y) ds, \quad n, k = 1, 2, \dots, N.$$

Je-li S počet trojúhelníků zvolené triangulace, potom je

$$\bar{\Omega} = \bigcup_{s=1}^S T_s$$

a pro hranice oblasti platí

$$\delta\Omega = \bigcup_{r=1}^R L_r,$$

neboli hranici aproximujeme sjednocením hraničních stran L_r hraničních trojúhelníků triangulace. Přitom přirozeně platí, že $S \geq R$.

Potom vzorce pro výpočet prvků matice A_h a prvků vektoru F_h můžeme psát ve tvaru

$$\begin{aligned} a_{nk} &= \sum_{s=1}^S \int \int_{T_s} [p(x, y) \operatorname{grad} v_n(x, y) \cdot \operatorname{grad} v_k(x, y) + q(x, y) v_n(x, y) v_k(x, y)] dx dy + \\ &\quad + \sum_{r=1}^R \int_{L_r} \sigma(x, y) v_n(x, y) v_k(x, y) ds, \quad n, k = 1, 2, \dots, N, \\ F_n &= \sum_{s=1}^S \int_{T_s} f(x, y) v_n(x, y) dx dy + \sum_{r=1}^R \int_{L_r} g(x, y) v_n(x, y) ds. \end{aligned}$$

Tím máme určeny koeficienty matice a můžeme hledat řešení celé soustavy.

Je vhodné se vyhýbat při triangulaci oblasti Ω obecným trojúhelníkům. Při použití tupouhlých trojúhelníků totiž hrozí (při limitním přechodu) ztráta linearit a s tím spojené havarování výpočtu. Tomuto nebezpečí se můžeme vyhnout, pokud budeme používat pouze ostroúhlé trojúhelníky, například při použití pravidelné sítě. (Nejjednodušší variantou potom je používat při triangulaci pravidelné šestiúhelníky.) Přitom na různých částech oblasti Ω může být velikost trojúhelníků různá. Podstatné zde je, že stále pracujeme s ostroúhlými trojúhelníky.

Pro řešení je vždy vhodné použít matematický software. Při řešení aplikačních úloh budete dostávat soustavy lineárních algebraických hodnot vysokých řádů. Použití MATLABu je popsáno například v práci [26], str 877.

Pojmy k zapamatování

- Uvedli jsme si základní pojmy z teorie parciálních diferenciálních rovnic, co je řešení a jaké okrajové podmínky se používají.
- Ukázali jsme si nejjednodušší metody pro řešení parciálních diferenciálních rovnic prvního řádu.

Kontrolní otázky

1. Metoda konečných prvků vede na řešení soustavy algebraických rovnic. Jaký bude její řád?

Cvičení

1. Najděte funkci $z = z(x, y)$, vyhovující diferenciální rovnici

$$\frac{\partial z(x, y)}{\partial x} = 1.$$

2. Najděte funkci $z = z(x, y)$, vyhovující diferenciální rovnici

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

Výsledky

1. Řešením úlohy je funkce

$$z(x, y) = x + \varphi(y),$$

kde $\varphi(y)$ je libovolná funkce.

2. Řešením úlohy je funkce

$$z(x, y) = \frac{x^2}{2} + xy + \varphi(y),$$

kde $\varphi(y)$ je libovolná funkce.

10.9 Cvičení

Maplety

Příklady pro samostatnou práci jsou uvedeny v samostatné sbírce příkladů. Pomocí následujících mapletů si můžete usnadnit některé dílčí výpočty, nebo zkontrolovat jejich správnost.

1. [Kanonické tvary parciálních diferenciálních rovnic 2. řádu](#)

Literatura

- [1] L. Bican: *Lineární algebra*, SNTL 1979, rozšířené vydání 2001.
- [2] G. Birkhoff, T. C. Barteo: *Aplikovaná algebra*, Alfa, Bratislava 1981.
- [3] G. Birkhoff, S. MacLane: *Algebra*, Alfa, Bratislava 1973.
- [4] R. Černá, M. Machlický, J. Vogel, Č. Zlatník: *Základy numerické matematiky a programování*. SNTL 1987.
- [5] Biswa Nath Datta: *Numerical linear algebra and applications*. Brooks and Cole Publishing Company, California, 1995.
- [6] M. Demlová, J. Nagy: *Algebra*, MVŠT —III, SNTL 1982.
- [7] J. Diblík, A. Haluzíková, J. Baštinec: *Numerické metody a matematická statistika*. VUT Brno, 1987 (skriptum).
- [8] J. Diblík, J. Baštinec: *Matematika IV*. Nakladatelství VUT v Brně, 1991 (skriptum).
- [9] F. Fabian, Z. Kluiber: *Metoda Monte Carlo a možnosti jejího uplatnění*, Prospektrum, Praha, 1998.
- [10] D. K. Faddejev, V. N. Faddejevova: *Computation Methods of Linear Algebra*. Moskva, Fizmatgiz, 1963.
- [11] M. Fiedler: *Speciální matice a jejich použití v numerické matematice*. Praha : SNTL, 1981.
- [12] L. E. Garner: *Calculus and analytic geometry*, London, 1988
- [13] A. Granas, J. Dugundji: *Fixed Point Theory*, Springer, Berlin, Heidelberg, New York, 2003, ISBN 0-387-00173-5.
- [14] E. Hairer, C. Lubich, G. Wanner: *Geometric Numerical Integration, Structure-Preserving Algorithms for Ordinary Differential Equations, second edition*. Springer, Berlin, Heidelberg, New York, 2006, ISBN 3-540-30663-3.

- [15] E. Hairer, S. P. Norsett, G. Wanner: *Solving Ordinary Differential Equations I, Non-stiff Problems, Second Revised Edition*. Springer, Berlin Heidelberg, New York, 2000, ISBN 3-540-56670-8.
- [16] E. Hairer, S.P. Nørsett, G. Wanner: *Solving Ordinary Differential Equations II, Stiff and Differential-Algebraic Problems, Second Revised Edition*. Springer, Berlin Heidelberg, New York, 2002, ISBN 3-540-60452-9.
- [17] A. Haluzíková: *Numerické metody*. Redakce VN MON VUT Brno, 1989 (skriptum).
- [18] V. Havel, J. Holenda: *Lineární algebra*, SNTL 1984.
- [19] N. J. Higham: *Accuracy and stability of numerical algorithms*. Philadelphia: Society for Industrial and Applied Mathematics.
- [20] R. A. Horn, Ch. R. Johnson: *Matrix Analysis*. Cambridge : Cambridge University Press, 1986. (ruský překlad, Moskva : Mir, 1989).
- [21] I. Horová: *Numerické metody*. Brno, Masarykova univerzita, 1999.
- [22] Z. Horský: *Množiny a matematické struktury*, MVŠT — I, SNTL 1980.
- [23] Z. Horský: *Vektorové prostory*, MVŠT — II, SNTL 1980.
- [24] Z. Horský: *Diferenciální počet*, MVŠT - V., Praha 1982.
- [25] B. Hruža, H. Mrhačová: *Cvičení z algebry a geometrie*, VUT, 1990.
- [26] S. C. Chapra, R. P. Canale: *Numerical methods for Engineers, fifth edition*. McGraw-Hill, New York, 2006, ISBN 007-124429-8.
- [27] V. Jarník: *Diferenciální počet I, II.*, Nakladatelství ČSAV, Praha 1963.
- [28] V. Jarník: *Integrální počet I, II.*, Nakladatelství ČSAV, Praha 1963.
- [29] P. Kaprálik, J. Tvarožek: *Zbierka riešených príkladov a úloh z lineárnej algebry a analytickej geometrie*, Alfa, Bratislava, 1987.
- [30] Kolektiv: *OBOROVÉ ENCYKLOPEDIÉ: Aplikovaná matematika A až Ž*. Praha, SNTL, 1978.
- [31] P. E. Kloeden, E. Platen: *Numerical solution of stochastic differential equations*, Springer, Berlin, Heidelberg, New York, 2008, ISBN 978-3-540-54062-5.
- [32] J. Kuben: *Diferenciální rovnice*. VA Brno 2000.
- [33] R. J. Leveque: *Finite Volume Methods for Hyperbolic Problems*. Cambridge University press, 2006, ISBN 0-521-00924-3.
- [34] J. D. Logan: *Applied partial differential equations*, Springer, Berlin Heidelberg, New York, 2004, ISBN 0-387-20953-0.

- [35] G. I. Marčuk: *Metody numerické matematiky*. Academia Praha 1987.
- [36] S. Míka: *Numerické metody algebry*, MVŠT — IV, SNTL 1982.
- [37] J. Nagy, E. Nováková, M. Vacek: *Integrální počet*, MVŠT — VI, SNTL Praha 1984.
- [38] J. Nagy, E. Nováková, M. Vacek: *Vektorová analýza*, MVŠT - VIII., Praha 1984.
- [39] M. Nekvinda, J. Šrubař, J. Vild: *Úvod do numerické matematiky*. SNTL 1976.
- [40] P. Pták: *Diferenciální rovnice, Laplaceova transformace*. ČVUT Praha 1999.
- [41] P. Příkryl: *Numerické metody matematické analýzy*. MVŠT — XXIV, SNTL 1985.
- [42] A. Quarteroni, R. Sacco, F. Saleri: *Numerical mathematics (Text in Applied Mathematics)*, Springer, Berlin Heidelberg, New York, 2006.
- [43] A. Quarteroni, F. Saleri: *Scientific Computing with MATLAB and Octave*, Second Edition, Springer, Berlin, Heidelberg, New York, 2006. ISBN 3-540-32612-X.
- [44] A. Quarteroni, A. Valli: *Numerical approximation of partial differential equations* Springer, Berlin, Heidelberg, New York, 2008, ISBN 978-3-540-85267-4.
- [45] A. Ralston: *A First Course in Numerical Analysis*. N. Y. : Mc Graw-Hill Book Company, 1965 (český překlad Praha : Academia, 1973).
- [46] K. Rektorys a kol.: *Přehled užití matematiky*. SNTL Praha 2009.
- [47] Z. Riečanová a kol.: *Numerické metody a matematická štatistika*. Alfa Bratislava 1987.
- [48] T. Šalát: *Metrické priestory*, Alfa, Bratislava 1981.
- [49] F. Šik: *Lineární algebra zaměřená na numerickou analýzu*. Brno, Masarykova univerzita, 1998.
- [50] M. Šikulová, Z. Karpíšek: *Matematika IV – Pravděpodobnost a matematická statistika*. VUT Brno, 1987 (skriptum).
- [51] E. Vitásek: *Numerické metody*. Praha: SNTL, 1987.
- [52] J. H. Wilkinson: *The Algebraic Eigenvalue Problem*. Oxford : Clarendon Press, 1965. (ruský překlad Moskva : Nauka, 1970).