Numerická matematika

Jiří Felcman

Univerzita Karlova v Praze

Matematicko-fyzikální fakulta

KNM PRESS • PRAHA
2010

PŘEDMLUVA

<u>1. přednáška</u>

1. `felcman@karlin.mff.cuni.cz`
   - Tel. 22191 3392
   - KNM č. dv. K458 (5042)
   -
2. Numerická matematika - anotace
   - Obecná informatika (Druh, program, ročník, obor: B.I.3.IOI)
3. Požadavky ke zkoušce
   - státnice (prospěl s vyznamenáním), souborná zkouška
   - sylabus
4. Tituly
   - Ph.D. (projekt + angličtina)
   - RNDr.
   - Mgr.
   - Bc
5. Studium v zahraničí - ERASMUS
6. Ceny udělované studentům
7. SVOČ
8. Hodnocení učitelů - srozumitelnost
9. Náhrada
   - 16.03.2010 (Zahraničí)
10. Zápočet: pátek 7. května 2010
11. Zkouška: úterý 11. května 2010 (14. května 2010 - ukončení výuky předmětů, které jsou uvedeny v doporučeném průběhu bakalářského studia pro 6. semestr.)
    část písemná
    část ústní

Praha, 2. března 2010 J. F.

Numerická matematika
MAI 042 p1a LS 2008/2009

1. přednáška Reálné situace, modely, diskretizace, počítačová realizace (fólie)

   Náměty do cvičení: Zdroje chyb v numerické matematice

2. přednáška Aproximace funkcí, interpolace, aproximace pomoci metody nejmenších čtverců, Lagrangeova báze, existence a jednoznačnost Lagrangeova interpolačního polynomu, chyba Lagrangeovy interpolace

   Náměty do cvičení: Newtonova báze, monomiální báze, Vandermondova matice a determinant, vyčíslení Lagrangeova interpolačního polynomu v jediném bodě (Aitkenovo-Nevilleovo schéma), Hornerovo schéma, vyčíslení hodnoty derivace polynomu pomocí Hornerova schématu, počet operací při vyčíslení hodnoty polynomu.

3. přednáška Přirozený kubický spline

   Náměty do cvičení: Konstrukce přirozeného kubického spline, řešitelnost soustavy rovnic pro momenty

4. přednáška Numerická integrace

   Náměty do cvičení: Symboly $o(h)$ a $O(h)$, substituce při výpočtu určitého integrálu (výpočet koeficientů N-C vzorců), odhad chyby N-C vzorců

5. přednáška Numerická integrace

   Náměty do cvičení: Rombergova kvadratura, ortogonální polynomy (důkaz věty 2.13), konstrukce ortogonálních polynomů p_0, p_1, p_2 na $[-1, 1]$, vyjádření polynomu jako lineární kombinace ortogonálních polynomů

6. přednáška Metody řešení nelineárních rovnic

   Náměty do cvičení: půlení intervalu, metoda sečen, metoda tečen, metoda regula falsi, důkaz lemmatu 3.3

7. přednáška Metody řešení nelineárních rovnic

   Náměty do cvičení: důkaz konvergence metody postupných aproximací (věta 3.7), Newtonova metoda jako speciální případ věty o pevném bodě, Hornerovo schéma pro výpočet derivace,zápis algoritmu Hornerova schématu

8. přednáška Soustavy lineárních rovnic

   Náměty do cvičení: zápis algoritmu Gaußovy eliminace, počet operací Gaußovy eliminace,

9. přednáška Soustavy lineárních rovnic

   Náměty do cvičení: Gaußova eliminace jako $\mathbb{LU}$ rozklad, $\mathbb{LU}$ rozklad obecně, $\mathbb{LU}$ rozklad matice 4×4, číslo podmíněnosti při $\mathbb{LU}$ rozkladu

10. přednáška Iterační metody řešení soustav lineárních rovnic

    Náměty do cvičení: výpočet jedné iterace Gaußovy–Seidelovy metody pro matici 4×4

11. přednáška Výpočet vlastních čísel matic

	Náměty do cvičení: motivace výpočtu vlastních čísel matic, aplikace, metody výpočtu vlastních čísel, vlastní čísla symetrických matic, diagonalizovatelnost matic, různé varianty mocninné metody
12. přednáška	Numerická integrace obyčejných diferenciálních rovnic
	Náměty do cvičení: Věta o existenci a jednoznačnosti řešení, soustavy obyčejných diferenciálních rovnic, Taylorův rozvoj funkce více proměnných, odvození Rungeovy–Kuttovy metody 2. řádu
13. přednáška	Gradientní metody
	Náměty do cvičení: pojem gradient funkce, motivace gradientních metod, odvození metody sdružených gradientů

OBSAH

ÚVOD

Numerická analýza: Studium algoritmů (jednoznačně definovaná konečná posloupnost aritmetických a logických operací) pro řešení problémů spojité matematiky. L.N. Trefethen, Bulletin IMA 1993

Numerická matematika: realizace matematických modelů na počítači

Fyzikální realita → matematický model → numerické řešení, t.j. realizace matematického modelu na počítači.

Validation (solving the right equations) – verification (solving the equations right)

Literatura k přednášce: (Quarteroni *et al.*, 2004), (Ueberhuber, 2000), (Segethová, 2000)

Předpokládané znalosti: Rolleova věta, definice normy funkce, definice seminormy, vlastní čísla, báze lineárního vektorového prostoru, Taylorova věta

1

APROXIMACE FUNKCÍ V $I\!R$

Jedna ze základních úloh numerické matematiky: aproximace dané funkce f jinou funkcí φ

Zadání aproximované funkce - analyticky, nebo je k dispozici

- tabulka hodnot $(x_i, f_i), x_i, f_i \in I\!R, i = 0, \ldots, n, n \in I\!N, f_i = f(x_i)$ (viz obr. 1.0.1)
- tabulka hodnot derivací do určitého řádu v uzlech x_i

Pro funkci f definovanou na uzavřeném intervalu $[a, b]$ uvažujeme *dělení intervalu* $[a, b]$, $a = x_0 < x_1, \ldots < x_n = b, n \in Z^+ = \{0, 1, \ldots\}$ a nazýváme ho *sítí*. x_i, $i = 0 \ldots, n$ nazýváme *uzly* (ekvidistantní, je-li $x_i = a + ih$, kde $h \in I\!R$ je krok sítě.)

Poznámka 1.1 Pojem síť se používá obecně v N-rozměrném prostoru, viz např. (Feistauer *et al.*, 2003, str. 185): Let $\Omega \subset I\!R^N$ be a domain. If $N = 2$, then by Ω_h we denote a polygonal approximation of Ω. This means that the boundary $\partial\Omega_h$ of Ω_h consists of a finite number of closed simple piecewise linear curves. For $N = 3$, Ω_h will denote a polyhedral approximation of Ω. For $N = 3$ we set $\Omega_h = \Omega$. The system $\mathcal{D}_h = \{D_i\}_{i \in J}$, where $J \subset Z^+ = \{0, 1, \ldots\}$ is an index set and $h > 0$, will be called a *finite volume mesh* in Ω_h, if D_i, $i \in J$, are *closed line segments* or *closed polygons* or *polyhedrons*, if $N = 1$ or $N = 2$ or 3, respectively, with mutually disjoint interiors such that

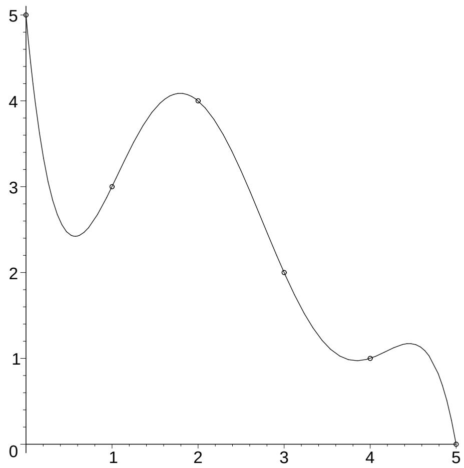
$$\overline{\Omega}_h = \bigcup_{i \in J} D_i.$$

The elements $D_i \in \mathcal{D}_h$ are called *finite volumes*. Two finite volumes D_i, $D_j \in \mathcal{D}_h$ are either disjoint or their intersection is formed by a common part of their boundaries ∂D_i and ∂D_j. If $\partial D_i \cap \partial D_j$ contains at least one straight segment or a plane manifold, if $N = 2$ or 3, respectively, then we call D_i and D_j *neighbouring finite volumes* (or simply neighbours).

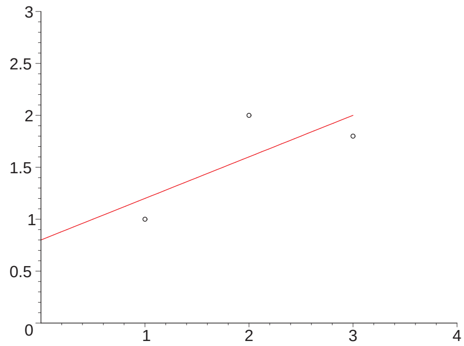
Požadavky na aproximující funkci φ

(A) jednoduchý tvar, snadno vyčíslitelná

- $*$ polynom $\{1, x, x^2, x^3, \ldots\}$
- $*$ trigonometrický polynom $\{1, \sin x, \cos x, \sin 2x, \cos 2x, \ldots\}$
- $*$ racionální funkce
- $*$ exponenciální funkce ae^{bx}



Obr. 1.0.1. Interpolační polynom nabývající v daných uzlech předepsaných hodnot



Obr. 1.0.2. Proložení přímky třemi body (ve smyslu nejmenších čtverců)

(B) $\varphi^{(j)}(x_i) = f^{(j)}(x_i), \quad \forall i = 0, \dots, n, j = 0, \dots, c_i$ (rovnost hodnot, event. derivací v uzlech)

(C) $\|\varphi - f\|$ 'malá', kde $\|\cdot\|$ značí normu

Poznámka 1.2 Od požadavku (B) někdy upouštíme (proložit třemi body přímku - viz obr. 1.0.2)

Nejčastější způsoby aproximace

1. Interpolace - k funkci f sestrojíme funkci φ z jisté třídy $\mathcal{M}$ splňující (B)
2. Aproximace metodou nejmenších čtverců - k funkci f sestrojíme funkci φ z jisté třídy $\mathcal{M}$ splňující (B) ve smyslu nejmenších čtverců
   - diskrétní případ

$$\sum_{i=0}^{n} w_i \big(f(x_i) - \varphi(x_i)\big)^2 = \min_{\psi \in \mathcal{M}} \sum_{i=0}^{n} w_i \big(f(x_i) - \psi(x_i)\big)^2$$

kde $w_i > 0, i = 0, \ldots, n$ jsou zadaná čísla, zvaná váhy. Název 'nejmenší čtverce' je patrný z následujícího příkladu:

Příklad 1.3 Pro dané dělení intervalu $[a, b]$ a dané kladné váhy w_i uvažujme normu funkce f danou vztahem

$$\|f\| := \sqrt{\sum_{i=0}^{n} w_i \big(f(x_i)\big)^2}$$

$\varphi \in \mathcal{M}$ se hledá tak, že

$$\|f - \varphi\|^2 = \min_{\psi \in \mathcal{M}} \|f - \psi\|^2$$

- spojitý případ

$$\int_a^b w(x)\big(f(x) - \varphi(x)\big)^2 \, dx = \min_{\psi \in \mathcal{M}} \int_a^b w(x)\big(f(x) - \psi(x)\big)^2 \, dx$$

w je váhová funkce (skoro všude kladná v $[a, b]$, $w \in L^2(a, b)$. Definice pojmu 'skoro všude' a prostoru $L^2(a, b)$ viz např. (Feistauer *et al.*, 2003, strana ...).)

3. Čebyševova (stejnoměrná) aproximace - k funkci f sestrojíme funkci φ z jisté třídy $\mathcal{M}$ splňující

$$\max_{[a,b]} |\varphi(x) - f(x)| \leq \max_{[a,b]} |\psi(x) - f(x)|$$

pro všechny funkce $\psi \in \mathcal{M}$, kde $\mathcal{M}$ je zvolená množina funkcí.

1.1 Lagrangeův interpolační polynom

Hledáme polynom L_n stupně nejvýše n (píšeme $L_n \in \Pi_n$ - prostor polynomů stupně nejvýše n) takový že

$$L_n(x_i) = f(x_i) \qquad i = 0, \ldots, n, \tag{1.1.1}$$

x_i - navzájem různé uzly, obecně neekvidistantní. Takový polynom nazveme *Lagranegeovým interpolačním polynomem.*

Věta 1.4 *Nechť $x_0, \ldots, x_n$ jsou navzájem různé uzly. Pak existuje právě jeden interpolační polynom $L_n \in \Pi_n$:*

$$L_n(x_i) = f(x_i) \qquad i = 0, \ldots, n.$$

Důkaz 1. Existence

Uvažujme polynomy

$$l_i(x) = \frac{(x-x_0)(x-x_1)\dots(x-x_{i-1})(x-x_{i+1})\dots(x-x_n)}{(x_i-x_0)(x_i-x_1)\dots(x_i-x_{i-1})(x_i-x_{i+1})\dots(x_i-x_n)}$$

(tzv. *Lagrangeovy polynomy*).

Platí

α) $\quad l_i(x) \in \Pi_n$,

β) $\quad l_i(x_j) = \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$ (Kroneckerovo delta).

Položme

$$L_n(x) = \sum_{i=0}^{n} f(x_i) l_i(x).$$

2. Jednoznačnost

Nechť $L_n^1, L_n^2 \in \Pi_n$ splňují (viz (1.1.1))

$$L_n^1(x_i) = L_n^2(x_i) = f(x_i) \qquad \forall i = 0, \dots, n.$$

Potom $L_n^1 - L_n^2 \in \Pi_n$ je polynom, který má $(n+1)$ různých kořenů. Podle základní věty algebry je $L_n^1 - L_n^2$ nulový polynom.

□

Poznámka 1.5 Položme

$$\omega_{n+1}(x) = (x-x_0)\dots(x-x_n).$$

Potom platí

$$\ell_i(x) = \frac{\omega_{n+1}(x)}{(x-x_i)\cdot\omega'_{n+1}(x_i)},$$

kde čárka označuje derivaci.

1.1.1 *Chyba Lagrangeovy interpolace*

Věta 1.6 *Nechť $f \in C^{n+1}(I)$, kde I je nejmenší interval obsahující $x_0, \dots, x_n, x^*$ a $x_0, \dots, x_n$ jsou navzájem různé uzly, . Nechť $L_n \in \Pi_n$ je Lagrangeův interpolační polynom pro funkci f. Pak $\exists\ \xi \in I$*

$$f(x^*) - L_n(x^*) = \frac{f^{(n+1)}(\xi)\cdot\omega_{n+1}(x^*)}{(n+1)!}$$

(chyba Lagrangeovy interpolace v bodě x^*).

Důkaz Pro $x^* = x_i$ je důkaz zřejmý. Pro $x^* \neq x_i$ uvažujme funkci :

$$F(x) = f(x) - L_n(x) - t \cdot \omega_{n+1}(x)$$

kde $t \in I\!R$. Platí:

$$F(x_i) = 0 \qquad \forall\, i = 0, \dots, n.$$

Pro vhodnou volbu

$$t := \frac{f(x^*) - L_n(x^*)}{\omega_{n+1}(x^*)} \tag{1.1.2}$$

platí, že $F(x^*) = 0$. F má tedy $n+2$ *nulových bodů* (uzly x_i a bod x^*). Podle Rolleovy věty:

$$\begin{array}{ll} F' & \text{má aspoň } n+1 \text{ nulových bodů,} \\ \vdots & \\ F^{(n+1)} & \text{má aspoň 1 nulový bod, označme ho } \xi. \end{array}$$

$$F^{(n+1)}(\xi) = 0 = f^{(n+1)}(\xi) - 0 - t \cdot (n+1)! \qquad \Big/ \frac{\omega_{n+1}(x^*)}{(n+1)!}$$

kde jsme využili toho, že $(n+1)$-ní derivace L_n je nulová a $(n+1)$-ní derivace ω_{n+1} je $(n+1)!$. Dosadíme-li za t ze vztahu (1.1.2), dostaneme

$$f(x^*) - L_n(x^*) = \frac{f^{(n+1)}(\xi) \cdot \omega_{n+1}(x^*)}{(n+1)!}.$$

□

Zkušební otázka 1.1! Chyba Lagrangeovy interpolace

2. přednáška

1.2 Kubický spline

Definice 1.7 *Nechť je dáno dělení intervalu $[a,b]$, $a = x_0 < x_1 < \ldots < x_n = b$ (x_i navzájem různé). Řekneme, že funkce $\varphi : [a,b] \to I\!R$ je kubický spline, jestliže*

1. *φ'' je spojitá ($\in C^2[a,b]$),*
2. *$\varphi|_{[x_i,x_{i+1}]}$ je kubický polynom, pro $i = 0, 1, \dots, n-1$.*

Poznámka 1.8 Spline - elastické pravítko používané při stavbě lodí

Poznámka 1.9 Kubický spline je speciálním případem *spline k-tého řádu* pro $k = 3$. Důvodem častého použití kubického spline je fakt, že lidské oko je schopné rozlišit ještě změny 2. derivace.

Poznámka 1.10 Kubický spline dobře aproximuje funkci, která popisuje tvar s minimální energií. Popíšeme-li tvar pružné laťky funkcí $y = f(x)$, potom

$$E(y) = \int_a^b \frac{y''(x)}{\left[1 + (y'(x))^2\right]^{3/2}}\, dx$$

měří její ohybovou energii. Lať se deformuje tak, že je tato energie minimální (Hamiltonův princip). Dá se ukázat, že mezi všemi funkcemi z $C^2[a, b]$ aproximuje kubický spline $\varphi :: \varphi(x_i) = f(x_i)$ velmi dobře funkci y^*, pro kterou se nabývá minima $E(y)$: $\min_y E(y) = E(y^*)$.

Věta 1.11 *Nechť $f \in C^2[a, b]$. Pak pro každý kubický spline φ splňující*

$$\varphi(x_i) = f(x_i), \qquad i = 0, \ldots, n,$$

platí

$$\|\varphi\| \leq \|f\|\,, \ \textit{kde}\ \|u\|^2 := \int_a^b |u''(x)|^2\, dx,$$

jestliže je splněna některá z následujících třech podmínek:

$$\begin{array}{lrcl} \text{(a)} & \varphi''(a) = & 0 & = \varphi''(b) \\ \text{(b)} & \varphi'(a) = f'(a) & a & \varphi'(b) = f'(b) \\ \text{(c)} & \varphi'(a) = \varphi'(b) & a & \varphi''(a) = \varphi''(b) \end{array} \tag{1.2.1}$$

Poznámka 1.12 (Pozor, $\|.\|$ ve větě 1.11 neznačí normu, ale pouze seminormu v Sobolevově prostoru $H^2(a, b)$, která se obvykle značí $|.|_{H^2(a,b)}$, detaily viz např. (Feistauer *et al.*, 2003, page ...))

Důkaz Viz cvičení k přednášce. □

Důsledek 1.13 *Ve všech třech případech (a), (b), (c) je kubický spline určen jednoznačně.*

1.2.1 *Konstrukce přirozeného kubického spline*

Značení:

$$\begin{array}{rcll} f_i & := & f(x_i) & \forall i = 0, \ldots, n, \\ \varphi_i & := & \varphi|_{[x_i, x_{i+1}]} & \forall i = 0, \ldots, n-1, \\ h_i & := & x_{i+1} - x_i & \forall i = 0, \ldots, n-1. \end{array}$$

Kubický polynom φ_i je na intervalu $[x_i, x_{i+1}]$ určen čtyřmi koeficienty. Počet intervalů je n, celkem máme tedy pro určení φ počet *stupňů volnosti* $4n$. Pro tyto stupně volnosti sestavíme příslušné rovnice.

Počet neznámých	$4 \times$ počet intervalů	$4n$
Počet rovnic	$\varphi(x_i) = f(x_i), i = 0, \ldots, n$	$n + 1$
	spojitost φ v $x_i, i = 1, \ldots, n-1$	$n - 1$
	spojitost φ' v $x_i, i = 1, \ldots, n-1$	$n - 1$
	spojitost φ'' v $x_i, i = 1, \ldots, n-1$	$n - 1$
		$4n - 2$

Počet rovnic je o dvě menší než počet neznámých. Doplníme je proto některou z podmínek (1.2.1), (a)–(b). Uvažujme např. podmínku (1.2.1), (a), tj. podmínku nulových druhých derivací v krajních bodech. Takový spline nazýváme *přirozeným kubickým splinem.* Pro určení přirozeného kubického splinu hledáme φ_i ve vhodném tvaru. Ukazuje se, že efektivní metoda není založena na vyjádření

$$\varphi_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i \qquad \text{(NEVHODNÉ viz cvičení)}$$

ani na vyjádření

$$\varphi_i(x) = a_i(x-x_i)^3 + b_i(x-x_i)^2 + c_i(x-x_i) + d_i \qquad \text{(MÉNĚ VHODNÉ viz cvičení)}$$

ale na vyjádření pomocí tzv. *momentů*, což jsou hodnoty druhé derivace φ v uzlech. Označme je M_i:

$$M_i := \varphi''(x_i),\ i = 0, \ldots, n$$

a předpokládejme, že tyto momenty známe. Později ukážeme, jak je určit. Platí

$$\begin{aligned} \varphi_i &- \text{kubický polynom} \\ \varphi_i' &- \text{parabola} \\ \varphi_i'' &- \text{přímka} \end{aligned}$$

Z předpokladu spojitosti druhé derivace φ v uzlech dostáváme

$$\begin{aligned} M_i &= \varphi_i''(x_i), \\ M_{i+1} &= \varphi_i''(x_{i+1}). \end{aligned}$$

Je tedy φ_i'' přímka, procházející body (x_i, M_i) a (x_{i+1}, M_{i+1}) (viz obr. 1.2.1).

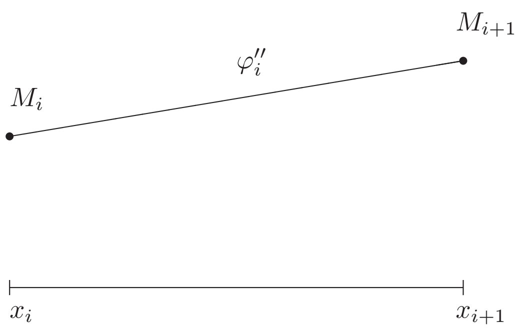
$$\begin{aligned} \varphi_i''(x) &= \frac{(x-x_i)\cdot M_{i+1}}{x_{i+1}-x_i} + \frac{M_i\cdot(x-x_{i+1})}{x_i - x_{i+1}}, \\ \varphi_i''(x) &= -\frac{M_i}{h_i}\cdot(x-x_{i+1}) + \frac{M_{i+1}}{h_i}\cdot(x-x_i). \end{aligned}$$

Integrací odvodíme

$$\varphi_i'(x) = -\frac{M_i}{2h_i}\cdot(x-x_{i+1})^2 + \frac{M_{i+1}}{2h_i}\cdot(x-x_i)^2 + A_i,$$

$$\varphi_i(x) = -\frac{M_i}{6h_i}\cdot(x-x_{i+1})^3 + \frac{M_{i+1}}{6h_i}\cdot(x-x_i)^3 + A_i(x-x_i) + B_i. \qquad (1.2.2)$$

vhodný rozpis integrační konstanty $\uparrow$



Obr. 1.2.1. Přímka φ_i''

Ve vyjádření φ_i ve tvaru (1.2.2) nejprve určíme koeficienty A_i, B_i, $i = 0, \dots, n-1$ pomocí momentů a potom sestavíme rovnice pro momenty. Využijeme k tomu podmínky

$$\varphi_i(x_i) = f_i,$$
$$\varphi_i(x_{i+1}) = f_{i+1},\ i = 0, \dots, n-1.$$

(Dvě rovnice pro dvě neznámé A_i, B_i, $i = 0, \dots, n-1$.) Dostaneme

$$\varphi_i(x_i) = \frac{M_i}{6} \cdot h_i^2 + B_i = f_i,$$

$$\rightarrow \quad B_i = f_i - \frac{M_i}{6} \cdot h_i^2,$$

$$\varphi_i(x_{i+1}) = \frac{M_{i+1}}{6} \cdot h_i^2 + A_i h_i + f_i - \frac{M_i}{6} \cdot h_i^2 = f_{i+1},$$

$$\rightarrow \quad A_i = \frac{f_{i+1} - f_i}{h_i} + \frac{M_i - M_{i+1}}{6} \cdot h_i.$$

Rovnice pro momenty sestavíme ekvivalentním vyjádřením podmínky spojitosti derivace kubického spline v uzlech:

$$\varphi_{i-1}'(x_i) = \varphi_i'(x_i), \quad i = 1, \dots, n.$$

Připomeňme si tvar φ_i'

$$\varphi_i'(x) = -\frac{M_i}{2h_i} \cdot (x - x_{i+1})^2 + \frac{M_{i+1}}{2h_i} \cdot (x - x_i)^2 + A_i$$

resp. φ_{i-1}'

$$\varphi_{i-1}'(x) = -\frac{M_{i-1}}{2h_{i-1}} \cdot (x - x_i)^2 + \frac{M_i}{2h_{i-1}} \cdot (x - x_{i-1})^2 + A_{i-1}.$$

S využitím vyjádření pro A_i, resp. A_{i-1} pomocí momentů dostaneme

$$\begin{aligned}\varphi'_{i-1}(x_i) &= 0 + \frac{M_i}{2h_{i-1}} \cdot h_{i-1}^2 + \frac{f_i - f_{i-1}}{h_{i-1}} + \frac{M_{i-1} - M_i}{6} \cdot h_{i-1} \\ &= -\frac{M_i}{2h_i} \cdot h_i^2 + 0 + \frac{f_{i+1} - f_i}{h_i} + \frac{M_i - M_{i+1}}{6} \cdot h_i = \varphi'_i(x_i).\end{aligned}$$

Protože konstruujeme přirozený kubický spline, je $M_0 = \varphi''(x_0) = 0 = \varphi''(x_n) = M_n$ a dostáváme tak $n-1$ rovnic $(i = 1, \dots, n-1)$ pro neznáme momenty $M_1, M_2, \dots, M_{n-1}$. Tyto rovnice lze přepsat ve tvaru

$$\frac{h_{i-1}}{6} M_{i-1} + \underbrace{\left(\frac{h_{i-1}}{2} - \frac{h_{i-1}}{6} + \frac{h_i}{2} - \frac{h_i}{6}\right)}_{\frac{h_{i-1}+h_i}{3}} \cdot M_i + \frac{h_i}{6} M_{i+1} = \underbrace{-\frac{f_i - f_{i-1}}{h_{i-1}} + \frac{f_{i+1} - f_i}{h_i}}_{g_i}.$$

Maticový zápis vede na soustavu s třídiagonální maticí.

$$\begin{pmatrix} \frac{h_0+h_1}{3} & \frac{h_1}{6} & & & & & & \\ \ddots & \ddots & \ddots & & & & & \\ & \ddots & \ddots & \ddots & & & & \\ & & \frac{h_{i-1}}{6} & \frac{h_{i-1}+h_i}{3} & \frac{h_i}{6} & & & \\ & & & \ddots & \ddots & \ddots & & \\ & & & & \ddots & \ddots & \ddots & \\ & & & & & \frac{h_{n-2}}{6} & \frac{h_{n-2}+h_{n-1}}{3} \end{pmatrix} \begin{pmatrix} M_1 \\ \vdots \\ M_{i-1} \\ M_i \\ M_{i+1} \\ \vdots \\ M_{n-1} \end{pmatrix} = \begin{pmatrix} g_1 \\ \vdots \\ g_{i-1} \\ g_i \\ g_{i+1} \\ \vdots \\ g_{n-1} \end{pmatrix}$$

Zkušební otázka 1.2! Konstrukce přirozeného kubického spline.

Příklad 1.14 Pro ekvidistantní dělení s krokem h má matice soustavy tvar

$$\frac{h}{6} \begin{pmatrix} 4 & 1 & & & \\ \ddots & \ddots & \ddots & & \\ & 1 & 4 & 1 & \\ & & \ddots & \ddots & \ddots \\ & & & 1 & 4 \end{pmatrix}$$

Při vyšetřování řešitelnosti této soustavy lze využít následující definici a větu z algebry:

Definice 1.15 *Řekněme, že matice $\mathbb{A}$ typu $n \times n, n \geq 2$ je ostře diagonálně dominantní (ODD), jestliže*

$$|a_{ii}| > \sum_{j=1, j\neq i}^{n} |a_{ij}| \qquad \forall i = 1, \dots, n.$$

Věta 1.16 *Nechť $\mathbb{A} \in \mathbb{R}^{n\times n}$ je ODD. Pak $\mathbb{A}$ je nesingulární.*

Důkaz pomocí Geršgorinových kruhů, viz (Quarteroni *et al.*, 2004, str. 184).

$\mathbb{A}$ je nesingulární $\Leftrightarrow$ $\det\mathbb{A} \neq 0 \Leftrightarrow$ rovnice $\det(\mathbb{A}-\lambda\mathbb{I}) = 0$ nemá kořen $\lambda = 0 \Leftrightarrow$ nula není vlastním číslem matice $\mathbb{A}$. Nechť λ je vlastní číslo matice $\mathbb{A}$

$$\begin{aligned}
&\mathbb{A}x = \lambda x \qquad y := \frac{x}{\|x\|}, \ \|x\| := \max_i |x_i| \\
&\mathbb{A}y = \lambda y \qquad |y_i| \leq 1, \ \exists i_0 :: |y_{i_0}| = 1 \\
&\qquad \sum_{j\neq i_0} a_{i_0 j} y_j + a_{i_0 i_0} y_{i_0} = \lambda y_{i_0} \\
&\qquad \Big|\sum_{j\neq i_0} a_{i_0 j} y_j\Big| = |\lambda - a_{i_0 i_0}||y_{i_0}| \\
&\qquad |\lambda - a_{i_0 i_0}| \leq \sum_{j\neq i_0} |a_{i_0 j}| \\
&\qquad \text{(Geršgorinův kruh o středu } a_{i_0 i_0} \text{ a poloměru } \sum_{j\neq i_0} |a_{i_0 j}|)
\end{aligned}$$

Kdyby $\lambda = 0$ bylo vlastním číslem

$$|a_{i_0 i_0}| \leq \sum_{j\neq i_0} |a_{i_0 j}| \qquad \text{Spor s ODD}$$

$\lambda = 0$ tedy není vlastní číslo a matice $\mathbb{A}$ je nesingulární. □

Matice soustavy rovnic pro momenty je ODD, soustava je tedy podle výše uvedené věty jednoznačně řešitelná a protože matice soustavy je třídiagonální, lze pro řešení použít např. Gaußovu eliminaci.

2

NUMERICKÁ INTEGRACE FUNKCÍ

3. přednáška

Nechť je dáno dělení intervalu $[a,b]$, $a \leq x_0 < x_1 < \ldots < x_n \leq b$ (x_i navzájem různé). Označme $h = \max_{i \in \{0,\ldots,n-1\}} |x_{i+1} - x_i|$.

$$\text{Cíl:} \quad I(f) = \int_a^b f(x)\,dx \approx I_h(f) = \sum_{i=0}^{n} \alpha_i\, f(x_i). \tag{2.0.1}$$

Vzorec

$$I_h(f) = \sum_{i=0}^{n} \alpha_i\, f(x_i)$$

se nazývá *kvadraturní formule*, α_i jsou *koeficienty* kvadraturní formule a x_i jsou *uzly* kvadraturní formule. Motivace hledání aproximace určitého integrálu ve tvaru lineární kominace hodnot funkce f v uzlech x_i je zřejmá z následujícího odstavce.

2.1 Newtonovy-Cotesovy vzorce

Pro dané $n \in I\!N$ uvažujme ekvidistantní dělení intervalu $[a,b]$ s krokem $h = \frac{b-a}{n}$, $x_i = a + ih$, $i = 0,\ldots,n$. Aproximujeme-li funkci f Lagrangeovým interpolačním polynomem L_n pro uzly $x_0,\ldots,x_n$, lze určitý integrál z funkce f aproximovat následujícím způsobem:

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

$$\int_a^b \sum_{i=0}^{n} f(x_i) \overbrace{\frac{(x-x_0)\ldots(x-x_{i-1})(x-x_{i+1})\ldots(x-x_n)}{(x_i-x_0)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots(x_i-x_n)}}^{\ell_i(x)}\,dx =$$

$$\sum_{i=0}^{n} \underbrace{\int_a^b \ell_i(x)\,dx}_{\alpha_i}\; f(x_i) \tag{2.1.1}$$

Tento vzorec nazýváme pro ekvidistantní uzly Newtonův-Cotesův. Pro výpočet koeficientů α_i použijeme následující substituci

$$\begin{aligned} subst. \quad & x = a + th \\ & x_i = a + ih, \qquad h = \frac{b-a}{n} \end{aligned}$$

$$\alpha_i := \int_a^b \prod_{j=0, j\neq i}^{n} \frac{(x-x_j)}{(x_i-x_j)}\, dx = \frac{b-a}{n} \int_0^n \prod_{j=0, j\neq i}^{n} \frac{(t-j)}{(i-j)}\, dt \tag{2.1.2}$$

Z konstrukce Lagrangeovy interpolace L_n funkce $f \in \Pi_n$ plyne, že $L_n(x) = f(x)$, a tedy N-C vzorec je přesný pro polynomy stupně nejvýše n. To nás vede k následující definici.

Definice 2.1 *Řekneme, že kvadraturní formule $\sum_{i=0}^n \alpha_i\, f(x_i)$ má řád přesnosti m, jestliže $m \in \mathbb{N} \cup \{0\}$ je maximální číslo takové, že*

$$\int_a^b p(x)\, dx = \sum_{i=0}^{n} \alpha_i\, p(x_i) \qquad \forall p \in \Pi_m. \tag{2.1.3}$$

Zkušební otázka 2.1 Řád kvadraturní formule.

Lemma 2.2 *Je-li kvadraturní formule $\sum_{i=0}^n \alpha_i\, f(x_i)$ symetrická, t.j. pro $i = 0, \dots, n$ platí*

$$\begin{aligned} b - x_{n-i} &= x_i - a, \\ \alpha_i &= \alpha_{n-i}, \end{aligned}$$

a je-li její řád $\geq n$, n sudé, pak je její řád $\geq n+1$.

Lemma 2.3 *Newtonův-Cotesův vzorec je symetrická kvadraturní formule.*

Důsledek 2.4 *Pro n sudé je řád N-C vzorce $\geq n+1$.*

Zkušební otázka 2.2 Odvoďte Newtonův-Cotesův vzorec.

Lemma 2.5 *(Odhad chyby lichoběžníkového pravidla)*
Nechť $f \in C^2[a,b]$. Označme $T_h(f)$ N-C vzorec pro $n = 1$ (lichoběžníkové pravidlo). Pak $\exists \xi \in [a,b]$, :: (při značení $I(f) = \int_a^b f(x)\, dx$)

$$I(f) - T_h(f) = -\frac{f''(\xi)}{2} \cdot \frac{h^3}{6}, \qquad h = (b-a). \tag{2.1.4}$$

Lemma 2.6 *(Odhad chyby Simpsonova pravidla)*
Nechť $f \in C^3[a,b]$. Označme $S_h(f)$ N-C vzorec pro $n = 2$ (Simpsonovo pravidlo). Pak $\exists \xi \in [a,b]$, ::

$$I(f) - S_h(f) = -\frac{h^5}{90} \cdot f'''(\xi), \qquad h = \frac{(b-a)}{2}. \tag{2.1.5}$$

Definice 2.7 *(zbytek kvadraturního vzorce) Rozdíl*

$$E_h(f) = I(f) - I_h(f),$$

kde

$$I(f) = \int_a^b f(x)\,dx, \;\; I_h(f) = \sum_{i=0}^{n} \alpha_i f(x_i),$$

nazýváme zbytek kvadraturního vzorce.

2.1.1 *Složené Newtonovy–Cotesovy vzorce*

Newtonovy–Cotesovy vzorce lze také aplikovat tak, že interval $[a, b]$ rozdělíme na m ekvidistantních subintervalů $[x_i, x_{i+1}]$ velikosti H a na každém z těchto subintervalů použijeme Newtonův–Cotesův vzorec pro n ekvidistantních uzlů $x_i = x_{i_0} < \cdots < x_{i_n} = x_{i+1}$ s krokem h

$$x_i = a + iH, \;\; H = \frac{b-a}{m}, \;\; i = 0, \ldots m, \;\; x_{i_j} = x_i + jh, \;\; h = \frac{H}{n} \quad m, n \in \mathbb{N}$$

$$I(f) := \sum_{i=0}^{m-1} \int_{x_i}^{x_{i+1}} f(x)\,dx \approx \sum_{i=0}^{m-1} I_h^i(f) = \sum_{i=0}^{m-1} \sum_{j=0}^{n} \alpha_{i_j} f(x_{i_j}) =: I_h(f)$$

Věta 2.8 *(složené N-C vzorce) Nechť $f \in C^{n+1}[a, b]$. Pak pro složené N-C vzorce platí*

$$|I(f) - I_h(f)| \le ch^{n+1}, \tag{2.1.6}$$

kde $c > 0$ je konstanta nezávislá na h.

Důkaz plyne z odhadu chyby Lagrangeova interpolačního polynomu □

2.2 Rombergova kvadratura

Výpočet $\int_a^b f(x)\,dx$ pomocí složeného lichoběžníkového pravidla pro $n+1$ uzlů.

$$\begin{aligned} m &= n, \\ H &= \frac{b-a}{m}, \\ h &= H. \end{aligned}$$

Věta 2.9 *(Eulerova-MacLaurinova) Nechť $f \in C^{2N+2}[a, b]$, $h = \frac{b-a}{n}$, $n \in \mathbb{N}$. Potom pro složené lichoběžníkové pravidlo (označme ho $CT_h(f)$) platí:*

$$\begin{aligned} CT_h(f) &= p(h^2) + O(h^{2N+2}) && (2.2.1) \\ &= I(f) + a_1h^2 + a_2h^4 + \cdots + a_Nh^{2N} + O(h^{2N+2}), && (2.2.2) \end{aligned}$$

$$\textit{kde} \quad p \in \Pi_N, \quad p = p(t) = a_0 + a_1t + \cdots + a_Nt,$$

$$a_0 = p(0) = \int_a^b f(x)\,dx = I(f).$$

Důkaz viz Stör Numerische Mathematik I □

Rombergova kvadratura: konstruujeme lineární kombinaci vzorců $CT_h(f)$ pro vhodné h tak, abychom získali vzorec, který je přesnější:

$$\begin{aligned} CT_h(f) &= I(f) + a_1 h^2 + O(h^4) \quad / -1 \\ CT_{\frac{h}{2}}(f) &= I(f) + a_1 \frac{h^2}{4} + O(h^4) \quad /4 \\ \frac{4CT_{\frac{h}{2}}(f) - CT_h(f)}{3} &= I(f) + O(h^4) \\ \text{lin. k.} &= I(f) + \text{chyba} \qquad (N = 1) \end{aligned}$$

Vhodnou lineární kombinací vzorců, z nichž každý aproximuje integrál $I(f)$ s chybou $O(h^2)$, jsme tak odvodili vzorec, který aproximuje integrál $I(f)$ s chybou $O(h^4)$. Za předpokladu dostatečné hladkosti funkce f (viz Eulerova–MacLaurinova věta) můžeme tímto způsobem odvodit vzorec, který aproximuje integrál $I(f)$ s chybou $O(h^{2N+2})$. Všimněme si např., jakou roli hraje v tomto postupu vyčíslení Lagrangeova interpolačního polynomu L_2 pro uzly $\frac{h^2}{16}$, $\frac{h^2}{4}$, h^2 a hodnoty $CT_{\frac{h}{4}}$, $CT_{\frac{h}{2}}$, CT_h zapsaného ve tvaru $L_2(0) + b_1 t + b_2 t^2$ (předpokládáme $h << 1$).

$$\begin{array}{rcll} & \text{Euler–MacLaurin} \downarrow & \text{Lagrange} \downarrow & \\ CT_h(f) = & I(f) + a_1 h^2 + a_2 h^4 + O(h^6) & = L_2(0) + b_1 h^2 + b_2 h^4 & (2.2.3) \\ CT_{\frac{h}{2}}(f) = & I(f) + a_1 \frac{h^2}{4} + a_2 \frac{h^4}{16} + O(h^6) & = L_2(0) + b_1 \frac{h^2}{4} + b_2 \frac{h^4}{16} & (2.2.4) \\ CT_{\frac{h}{4}}(f) = & I(f) + a_1 \frac{h^2}{16} + a_2 \frac{h^4}{256} + O(h^6) & = L_2(0) + b_1 \frac{h^2}{16} + b_2 \frac{h^4}{256} & (2.2.5) \\ \hline \text{lin. k.} \uparrow = & I(f) + 0 + 0 + O(h^6) & = L_2(0) + 0 + 0 & (N = 2) \end{array}$$

kde $L_2(t) = \underbrace{b_0}_{L_2(0)} + b_1 t + b_2 t^2$ je Lagrangeův interpolační polynom pro tabulku

0	$\frac{h^2}{16}$	$\frac{h^2}{4}$	h^2
lin. k.	$CT_{\frac{h}{4}}(f)$	$CT_{\frac{h}{2}}(f)$	$CT_h(f)$

Závěr: $L_2(0)$ aproximuje $\int_a^b f(x)\,dx$ s chybou $O(h^6)$. Při konstrukci $L_2(0)$ se jedná o tzv. Richardsonovu extrapolaci. Uvedený postup lze provést až do řádu $2N+2$ pro uzly $(\frac{h}{2^i})^2$ a hodnoty $CT_{\frac{h}{2^i}}$, $i = 0, \ldots, N$, pomocí nichž konstruujeme L_N.

Problém: Vyčíslení Lagrangeova interpolačního polynomu v jediném bodě (zde konkrétně v 0) aniž bychom Lagrangeův interpolační polynom sestavovali. Zde nepotřebujeme tvar Lagrangeova interpolačního polynomu, ale pouze jeho hodnotu v jediném bodě. K tomu se používá Aitkenovo–Nevilleovo schéma.

Zkušební otázka 2.3! Odvoďte Rombergův kvadraturní vzorec, který aproximuje hodnotu $\int_a^b f(x)\,dx$ s chybou $O(h^4)$. Vysvětlete význam Lagrangeova interpolačního polynomu při konstrukci Rombergova kvadraturního vzorce.

2.3 Gaußova kvadratura

4. přednáška

Víme, že N-C vzorce mají řád aspoň n (pro n sudé dokonce aspoň $n+1$). Jakého řádu může být formule typu $\sum_{i=0}^n \alpha_i f(x_i)$? Uvažujme pro dané dělení intervalu $[a,b]$, $a \le x_0 < \ldots < x_n \le b$ kvadraturní formuli

$$I_h(f) = \sum_{i=0}^{n} \alpha_i\, f(x_i). \tag{2.3.1}$$

Lemma 2.10 *(Řád kvadraturní formule) Řád kvadraturní formule* (2.3.1) *je nejvýše* $2n+1$.

Důkaz Uvažujme polynom $\tilde{p}(x) = \prod_{i=0}^n (x-x_i)^2 \in \Pi_{2n+2}$. Tento polynom je nezáporná funkce na intervalu $[a,b]$ a platí pro něho

$$\int_a^b \tilde{p}(x) > 0.$$

Kvadraturní formule typu (2.3.1) dává pro tento polynom

$$\sum_{i=0}^{n} \alpha_i \tilde{p}(x_i) = 0.$$

Pro polynom $\tilde{p}$ není tedy kvadraturní formule (2.3.1) přesná a její řád je tedy nejvýše $2n+1$. $\square$

Gaußova kvadratura je způsob konstrukce vzorce $\sum_{i=0}^n \alpha_i f(x_i)$, který je přesný pro všechny polynomy stupně nejvýše $2n+1$.

Definice 2.11 *(skalární součin polynomů) Skalární součin v* $C[a,b]$ *je definován*

$$(u,v) = \int_a^b u(x)v(x)\,dx. \tag{2.3.2}$$

Definice 2.12 *Množina normovaných polynomů*

$$\tilde{\Pi}_n = \{p \in \Pi_n; \quad p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0\}. \tag{2.3.3}$$

Myšlenka konstrukce Gaußovy kvadratury:

x_i (uzly):	kořeny polynomu p_{n+1} z množiny ortogonálních polynomů $\{p_0, p_1, \ldots, p_{n+1}\}$,
α_i (koeficienty):	určíme tak, aby $\int_a^b q(x)\,dx = \sum_{i=0}^n \alpha_i q(x_i) \quad \forall q \in \Pi_{2n+1}$.

Věta 2.13 *(Ortogonální polynomy) Existují jednoznačně určené polynomy p_i, pro které platí*

1. $p_i \in \tilde{\Pi}_i, \quad i \in \mathbb{N} \cup \{0\},$

$$(p_i, p_j) = 0, \quad i \neq j, \qquad (\textit{pozn.} \quad p_0(x) = 1)$$

2. *Kořeny $x_0, \ldots, x_n$ polynomu p_{n+1}, $n \in \mathbb{N} \cup \{0\}$, jsou reálné, jednoduché a leží v (a, b)*
3. $$A = \begin{pmatrix} p_0(x_0) & p_0(x_1) & \cdots & p_0(x_n) \\ p_1(x_0) & p_1(x_1) & \cdots & p_1(x_n) \\ & & \ddots & \\ p_n(x_0) & p_n(x_1) & \cdots & p_n(x_n) \end{pmatrix} \qquad \textit{je nesingulární.}$$

Důkaz viz cvičení k přednášce. □

S využitím ortogonálních polynomů $p_0, \ldots, p_{n+1}$ a kořenů x_i polynomu p_{n+1} určíme koeficienty α_i Gaußovy kvadraturní formule tak, aby platilo:

$$\int_a^b q(x)\,dx = \sum_{i=0}^{n} \alpha_i q(x_i), \quad \forall q \in \Pi_{2n+1}. \tag{2.3.4}$$

K tomu vyjádříme polynom q ve tvaru

$$q(x) = r(x)p_{n+1}(x) + s(x), \quad r, s \in \Pi_n,$$

(dělení polynomu q polynomem p_{n+1}) a polynomy $r(x), s(x) \in \Pi_n$ vyjádříme jako lineární kombinaci ortogonálních polynomů (existence takového vyjádření viz cvičení k přednášce), specielně nechť

$$s(x) = \sum_{j=0}^{n} \gamma_j p_j(x).$$

Na základě tohoto vyjádření má výraz na levé straně v (2.3.4) tvar

$$\int_a^b q(x)\,dx = \underbrace{\int_a^b r(x)p_{n+1}(x)\,dx}_{=0} + \int_a^b s(x)\,dx$$

$$= \int_a^b \sum_{j=0}^{n} \gamma_j p_j(x)\,dx = \int_a^b \sum_{j=0}^{n} \gamma_j \overbrace{p_0(x)}^{=1} p_j(x)\,dx$$

$$= \sum_{j=0}^{n} \gamma_j \int_a^b p_0(x) p_j(x)\, dx = \gamma_0 \int_a^b p_0(x) p_0(x)\, dx = \gamma_0 \int_a^b dx.$$

Levá strana v (2.3.4) je tedy rovna

$$\gamma_0(b-a).$$

Pravá strana v (2.3.4) má na základě výše uvedených vyjádření tvar

$$\sum_{i=0}^{n} \alpha_i [\underbrace{r(x_i) p_{n+1}(x_i)}_{=0} + s(x_i)] = \sum_{i=0}^{n} \alpha_i \sum_{j=0}^{n} \gamma_j p_j(x_i).$$

Vidíme, že levou a pravou stranu v (2.3.4) lze tedy vyjádřit jako lineární kombinací jistých výrazů s koeficienty γ_j

$$\begin{aligned} &\gamma_0(b-a) + \quad \gamma_1 \cdot 0 \quad + \dots + \gamma_n \cdot 0 \\ = &\gamma_0 \sum_{i=0}^{n} p_0(x_i)\alpha_i + \gamma_1 \sum_{i=0}^{n} p_1(x_i)\alpha_i + \dots + \gamma_n \sum_{i=0}^{n} p_n(x_i)\alpha_i \end{aligned}$$

Porovnáním výrazů u koeficientů γ_j na levé a pravé straně dostaneme rovnice pro určení hledaných koeficientů α_i:

$$\left.\begin{array}{rcl} \sum_{i=0}^{n} p_0(x_i)\alpha_i &=& (b-a) \\ \sum_{i=0}^{n} p_1(x_i)\alpha_i &=& 0 \\ \vdots & & \vdots \\ \sum_{i=0}^{n} p_n(x_i)\alpha_i &=& 0 \end{array}\right\} \Leftrightarrow \begin{pmatrix} p_0(x_0) & \cdots & p_0(x_n) \\ p_1(x_0) & \cdots & p_1(x_n) \\ & \ddots & \\ p_n(x_0) & \cdots & p_n(x_n) \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} b-a \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Z hlediska stability je výhodné, že koeficienty α_i Gaußova kvadraturního vzorce $\sum_{i=0}^{n} \alpha_i f(x_i)$ jsou kladné.

Věta 2.14 *(pozitivita α_i) Koeficienty α_i Gaußova kvadraturního vzorce jsou kladné.*

Důkaz Položme:

$$\tilde{p}_k(x) = \prod_{i=0, i\neq k}^{n} (x - x_i)^2 \quad \in \Pi_{2n}$$

$$0 < \int_a^b \tilde{p}_k(x)\, dx = \sum_{i=0}^{n} \alpha_i \tilde{p}_k(x_i) = \alpha_k \underbrace{\tilde{p}_k(x_k)}_{>0}$$

$$\Rightarrow \alpha_k \text{ kladné } \forall k = 0, 1, \dots, n.$$

□

Zkušební otázka 2.4! Odvoďte Gaußův kvadraturní vzorec řádu $2n+1$ na intervalu $[a, b]$. Odvoďte Gaußův kvadraturní vzorec řádu 3 na intervalu $[-1, 1]$ (uvažujte ortogonální polynomy $\{1, x, x^2 - \frac{1}{3}\}$.

3

METODY ŘEŠENÍ NELINEÁRNÍCH ROVNIC

5. přednáška

Nechť je dáno nelineární zobrazení

$$F : \mathbb{R}^N \to \mathbb{R}^N.$$

$$\text{Hledáme } \alpha :: \ F(\alpha) = 0.$$

Metody pro řešení výše uvedené úlohy jsou většinou iterační. Cíl je generovat posloupnost $\{x^{(k)}\}$ takovou, že $\lim x^{(k)} = \alpha$, kde $F(\alpha) = 0$.

3.1 Newtonova metoda

Problém $F(x) = 0$ nahradíme posloupností lineárních problémů $L_k(x) = 0$, $L_k : \mathbb{R}^N \to \mathbb{R}^N$, $k = 0, 1, \ldots$, takových, že jejich řešení tvoří posloupnost konvergující k řešení problému $F(x) = 0$.

$$\alpha \approx x^{k+1}, \text{ kde } L_k(x^{k+1}) = 0.$$

Nechť $x^{(0)}$ je dáno (později ukážeme, jak ho volit). Pro danou aprixamaci $x^{(k)}$ uvažujeme $L_k(x)$ jako lineární část Taylorova rozvoje zobrazení F v bodě $x^{(k)} \in \mathbb{R}^N$ ($J(x)$ značí Jakobiho matici zobrazení F v bodě x):

$$F(x) = \underbrace{F(x^{(k)}) + J(x^{(k)})(x - x^{(k)})}_{L_k(x)} + O(|x - x^{(k)}|^2).$$

(za předpokladu dostatečné hladkosti zobrazení F). Nelineární problém nahradíme problémem lineárním

$$\{F(x) = 0\} \qquad \approx \qquad \{\underbrace{F(x^{(k)}) + J(x^{(k)})(x - x^{(k)})}_{L_k(x)} = 0 \quad (3.1.1)$$

řešení α nelin. pb. aproximujeme řešením $x^{(k+1)}$ lin. pb. (3.1.2)

$$\alpha \qquad \approx \qquad x^{(k+1)} := x^{(k)} - J^{-1}(x^{(k)})F(x^{(k)}) \quad (3.1.3)$$

Vzorec v (3.1.3), kterým je definována $(k+1)$-ní aproximace $x^{(k+1)}$ řešení nelineárního problému je formální, ve skutečnosti se inverzní matice nepočítá a algoritmus má následující dva kroky:

Algoritmus:

1. $J(x^{(k)})\underbrace{(x - x^k)}_{\delta x^{(k)}} = -F(x^{(k)})$ - řešíme lineární úlohu pro $\delta x^{(k)}$
2. $x^{(k+1)} := x^{(k)} + \delta x^{(k)}$ - provedeme update předchozí aproximace.

Pro $N = 1$ (nelineární skalární rovnice pro jednu neznámou) má Newtonova metoda názorný geometrický význam. Nelineární funkci $f(x)$ nahradíme lineární funkcí (přímkou), která je tečnou ke grafu funkce f v bodě $(x^{(k)}, f(x^{(k)})$ (má tedy směrnici $f'(x^{(k)})$ a prochází bodem $(x^{(k)}, f(x^{(k)}))$. V tomto případě se Newtonova metoda nazývá metodou tečen.

$N = 1:$

$$x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(k)})}, \quad x^{(0)} \text{ dáno, } f'(x^{(k)}) \neq 0.$$

Zkušební otázka 3.1! Odvoďte Newtonovu metodu pro soustavy nelineárních rovnic a její algoritmizaci. Popište algoritmus v případě jedné skalární rovnice.

Způsob, jak nahradit funkci f přímkou procházející bodem $(x^{(k)}, f(x^{(k)}))$ není jediný. Další možnosti jsou metoda sečen, jednobodová metoda sečen nebo metoda regula falsi (viz cvičení k přednášce).

Poznámka 3.1 Newtonova metoda je speciálním případem náhrady funkce f lineární funkcí

$$l_k(x) := f(x^{(k)}) + (x - x^{(k)})q_k,$$

kde směrnice q_k se volí

$$q_k := f'(x^{(k)}).$$

3.1.1 *Důkaz konvergence Newtonovy metody*

Věta 3.2 *(Konvergence Newtonovy metody pro soustavy) Nechť $F \in C(D), D \subset \mathbb{R}^N$ konvexní, otevřená množina, která obsahuje α :: $F(\alpha) = 0$. Nechť $\exists\, J^{-1}(\alpha)$, nechť $\exists\, R > 0, c > 0, L > 0$:*

$$\begin{aligned} \left\|J^{-1}(\alpha)\right\| &\leq c, \\ \underbrace{\|J(x) - J(y)\|}_{\textit{maticová norma}} &\leq L \underbrace{\|x - y\|}_{\textit{vekt. norma}} \qquad \forall x, y \in B(\alpha, R), \end{aligned}$$

kde $B(\alpha, R)$ je koule o středu α a poloměru R. Potom $\exists\, r$, $\forall x^{(0)} \in B(\alpha, r)$, posloupnost 3.1.3 je jednoznačně definována a konverguje k α a platí

$$\left\|\alpha - x^{(k+1)}\right\| \leq cL \left\|\alpha - x^{(k)}\right\|^2 \tag{3.1.4}$$

<u>Motivace:</u> $N = 1$

$$x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(k)})} \qquad \text{Newtonova metoda}$$

Z Taylorova rozvoje dostáváme:

$$f(\alpha) = f(x^{(k)}) + f'(x^{(k)})(\alpha - x^{(k)}) + \frac{f''(\boldsymbol{\xi})(\alpha - x^{(k)})^2}{2}.$$

Zajímá nás chyba $(\alpha - x^{(k+1)})$. Chceme ukázat, že $\left|\alpha - x^{(k+1)}\right| \to 0$. Odečtením α od obou stran vzorce pro Newtonovu metodu získáme vyjádření

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

Úpravou Taylorova rozvoje (uvědomíme-li si, že $f(\alpha) = 0$) dostaneme

$$0 = \frac{f(x^{(k)})}{f'(x^{(k)})} + (\alpha - x^{(k)}) + \frac{f''(\boldsymbol{\xi})(\alpha - x^{(k)})^2}{2 \cdot f'(x^{(k)})}.$$

Z předchozích dvou rovnic snadno nahlédneme, že platí

$$\alpha - x^{(k+1)} = -\frac{f''(\boldsymbol{\xi})(\alpha - x^{(k)})^2}{2 \cdot f'(x^{(k)})}.$$

Předpokládejme nyní, že existuje konstanta $\tilde{c}$ taková, že podíl derivací na pravé straně předchozího výrazu lze odhadnout

$$\left|\frac{f''(\boldsymbol{\xi})}{2 \cdot f'(x^{(k)})}\right| < \tilde{c} \quad \forall \boldsymbol{\xi}\ \forall x^{(k)}.$$

Potom dostaneme

$$\left|\alpha - x^{(k+1)}\right| \le \tilde{c}\left|\alpha - x^{(k)}\right|^2 \le \tilde{c}\left(\tilde{c}\left|\alpha - x^{(k-1)}\right|^2\right)^2 = \frac{1}{\tilde{c}}\left(\tilde{c}\left|\alpha - x^{(k-1)}\right|\right)^4$$

$$\le \ldots \le \frac{1}{\tilde{c}}\left(\tilde{c}\left|\alpha - x^{(0)}\right|\right)^{2^{k+1}}.$$

Pravá strana konverguje k nule pro $k \to +\infty$ za předpokladu

$$\tilde{c}\left|\alpha - x^{(0)}\right| < 1, \text{tj. jestliže je } x^{(0)} \text{ dostatečně blízko k } \alpha.$$

Pro důkaz konvergence Newtonovy metody pro jednu skalární rovnici jsme tedy využili tyto předpoklady

1. $x^{(0)}$ dostatečně blízko α,
2. f'' omezená shora
3. $\frac{1}{f'}$ omezená shora (tj. předpokládáme $f'(\alpha) \neq 0$),

které korespondují s předpoklady Věty 3.2. Jak, to je patrné z následujícího důkazu.

Důkaz věty 3.2. $x^{(0)}$ zvolíme v $B(\alpha, r)$, kde r učíme tak, aby $J^{-1}(x^{(0)})$ existovala. (Jinými slovy, $x^{(0)}$ volíme dostatečně blízko α.) K tomu využijeme následující tvrzení z algebry:

Lemma 3.3

$$\|\mathbb{A}\| < 1 \Rightarrow (\mathbb{I} - \mathbb{A})^{-1} \qquad \textit{existuje a platí}$$

$$\left\|(\mathbb{I} - \mathbb{A})^{-1}\right\| \leq \frac{1}{1 - \|\mathbb{A}\|}.$$

Důkaz viz cvičení k přednášce. □

Definujme matici

$$\mathbb{A} := \mathbb{I} - J^{-1}(\alpha)J(x^{(0)})$$

kde $x^{(0)}$ zvolíme tak (blízko α), aby $\|\mathbb{A}\| < 1$, konkrétně zvolíme $x^{(0)}$ tak, aby

$$\|\mathbb{A}\| \leq \frac{1}{2}.$$

K tomu využijeme předpoklady Věty 3.2 týkající se odhadu inverze Jacobiho matice v bodě α a lipschitzovskosti Jacobiho matice:

$$\left\|\overbrace{\mathbb{I} - J^{-1}(\alpha)J(x^{(0)})}^{\mathbb{A}}\right\| = \left\|J^{-1}(\alpha)(J(\alpha) - J(x^{(0)}))\right\| \leq cL\left\|\alpha - x^{(0)}\right\|. \tag{3.1.5}$$

$x^{(0)}$ zvolíme tak, aby poslední výraz v (3.1.5) $\leq \frac{1}{2}$. Tím dostáváme podmínku na $x^{(0)}$:

$$\left\|\alpha - x^{(0)}\right\| \leq \frac{1}{cL} \text{ a zároveň } \left\|\alpha - x^{(0)}\right\| \leq R \text{ (platnost podmínky lipschitzovskosti).}$$

Pro r dostáváme

$$r := \min\left(\frac{1}{2cL}, R\right).$$

V množine $B(\alpha, r)$ existuje podle výše uvedeného tvrzení z algebry $J^{-1}(x^{(0)})$. To plyne ze vztahů

$$(\mathbb{I} - \mathbb{A}) = J^{-1}(\alpha)J(x^0), \qquad (\mathbb{I} - \mathbb{A})^{-1} = J^{-1}(x^{(0)})J(\alpha).$$

Lze tedy spočítat první iteraci Newtonovy metody a odhadnout její chybu:

$$\alpha - x^{(1)} = \alpha - x^{(0)} + J^{(-1)}(x^{(0)})F(x^{(0)}).$$

Úpravou Taylorova rozvoje (uvědomíme-li si, že $F(\alpha) = 0$) dostaneme

$$0 = F(\alpha) = F(x^{(0)}) + J(x^{(0)})(\alpha - x^{(0)}) + \text{ zbytek},$$

$$0 = J^{-1}(x^{(0)})F(x^{(0)}) + (\alpha - x^{(0)}) + J^{-1}(x^{(0)}) \text{ zbytek}.$$

S využitím odhadu zbytku Taylorova rozvoje dostaneme

$$\left\|\alpha - x^{(1)}\right\| \leq \left\|J^{-1}(x^{(0)})\right\| \overbrace{\frac{1}{2}L\left\|\alpha - x^{(0)}\right\|^2}^{\text{odhad zbytku}}$$

a důkaz dokončíme pomocí odhadu normy inverzní matice

$$\left\|J^{-1}(x^{(0)})\right\| = \left\|\underbrace{J^{(-1)}(x^{(0)})J(\alpha)}_{(\mathbb{I}-\mathbb{A})^{-1}} J^{(-1)}(\alpha)\right\| \leq \frac{1}{1-\underbrace{\|\mathbb{A}\|}_{\leq\frac{1}{2}}}c \leq 2c.$$

Pro odhad chyby máme tedy vztah

$$\left\|\alpha - x^{(1)}\right\| \leq cL\left\|\alpha - x^{(0)}\right\|^2 = \underbrace{\left(cL\left\|\alpha - x^{(0)}\right\|\right)}_{\leq\frac{1}{2}}\left\|\alpha - x^{(0)}\right\|,$$

z něhož plyne dále indukcí konvergence Newtonovy metody. □

Zkušební otázka 3.2 Dokažte větu o konvergenci Newtonovy metody.

6. přednáška

Poznámka 3.4 Modifikace Newtonovy metody:

- Jacobiho matice se nemění pro $p \geq 2$ kroků
- nepřesné řešení soustavy lin. rovnic
- vyčíslení Jacobiho matice pomocí diferencí $f'(x) \approx \frac{f(x+h)-f(x)}{h}$

3.1.2 *Řád konvergence*

Definice 3.5 *(řád konvergence iterační metody pro řešení $F(x) = 0$) Řekneme, že posloupnost $\{x^{(k)}\}$ generovaná numerickou metodou konverguje k α s řádem $p \geq 1$, pokud $\exists\, c > 0$*

$$\frac{\left\|\alpha - x^{(k+1)}\right\|}{\left\|\alpha - x^{(k)}\right\|^p} \leq c \qquad \forall\, k \geq k_0.$$

V takovém případě se numerická metoda nazývá řádu p.

Věta 3.2 říká, že Newtonova metoda je kvadraticky konvergentní,

$$\left\|\alpha - x^{(k+1)}\right\| \leq cL\left\|\alpha - x^{(k)}\right\|^2,$$

pokud je $x^{(0)}$ dostatečně blízko α a pokud je $J(\alpha)$ nesingulární.

3.2 Metoda postupných aproximací pro nelineární rovnice

Metoda postupných aproximací je založena na faktu, že pro dané zobrazení $F : M \subset \mathbb{R}^N \to \mathbb{R}^N$ je vždy možné transformovat problém $F(x) = 0$ na ekvivalentní problém $x - \phi(x) = 0$, kde pomocná funkce ϕ je volena tak, aby $\phi(\alpha) = \alpha$ právě když $F(\alpha) = 0$. Nalezení nulových bodů zobrazení F se tak převede na nalezení *pevného bodu* zobrazení ϕ, které se realizuje pomocí následujícího algoritmu:

Dáno $x^{(0)}$, $x^{(k+1)} := \phi(x^{(k)}), \quad k \geq 0.$

Definice 3.6 *(kontrahující zobrazení) Řekneme, že zobrazení $G : D \subset \mathbb{R}^N \to \mathbb{R}^n$ je kontrahující na $D_0 \subset D$, jestliže $\exists L < 1 ::$*

$$\|G(x) - G(y)\| \leq L \|x - y\| \qquad \forall x, y \in D_0.$$

Věta 3.7 *(věta o pevném bodě) Nechť $G : D \subset \mathbb{R}^N \to \mathbb{R}^N$ kontrahující na uzavřené množině $D_0 \subset D$, $G(x) \in D_0 \quad \forall x \in D_0$. Pak G má právě jeden pevný bod. Tento bod je limitou posloupnosti $x^{(k+1)} = \phi(x^{(k)})$, $x^{(0)} \in D_0$ libovolné.*

Důkaz jednoznačnost, existence (Cauchyovská posloupnost, spojitost G), viz cvičení k přednášce. □

Poznámka 3.8 Newtonova metoda jako speciální případ věty o pevném bodě. (Viz cvičení k přednášce.)

3.3 Kořeny polynomu

Nalezení

- lokalizace kořenů v $\mathbb{C}$
- aproximace kořenů

Věta 3.9. (Descartes) *Počet kladných kořenů (včetně násobnosti) polynomu $p_n(\alpha) = a_0 + a_1 x + \cdots + a_n x^n$ je roven počtu znaménkových změn v posloupnosti $a_0, a_1, \ldots, a_n$, nebo je o sudé číslo menší.*

Věta 3.10. (Cauchy) *Kořeny polynomu leží v kruhu*

$$\Gamma = \left\{ z \in \mathbb{C}; |z| \leq 1 + \eta, \eta = \max_{0 \leq k \leq n-1} \left| \frac{a_k}{a_n} \right| \right\}$$

Poznámka 3.11 $1 \ll \eta$: translace a změna souřadnic

3.3.1 *Hornerovo schema*

V dalším budeme potřebovat vyčíslení hodnoty polynomu

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

v daném bodě x. Vyčíslení polynomu:

1. neefektivní

   $r = 1; \; s = a_0;$

for $i = 1$ to n **do**
$r = r \cdot x;$
$s = s + a_i \cdot r;$
end for
$p_n(x) = s$, počet násobení $2n$.

2. Hornerovo schéma

$s = a_n;$
for $i = n-1$ downto 0 **do**
$s = s \cdot x + a_i;$
end for
$p_n(x) = s$, počet násobení n.

Poznámka 3.12 Zapišme Hornerovo schéma pro vyčíslení $p_n(z)$ takto:

$b_n = a_n;$
for $i = n-1$ downto 0 **do**
$b_i = b_{i+1} \cdot z + a_i;$
end for
$p_n(z) = b_0$.

Ukážeme, že tento zápis je vhodný pro vyčíslení derivace p_n' (a následně použijeme Newtonovu metodu pro určení kořene $p_n(x)$). Pro dělení polynomu polynomem platí

$$(a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0) : (x - z) = \underbrace{a_n}_{b_n} x^{n-1} + (\underbrace{a_{n-1} + a_n z}_{b_{n-1}}) x^{n-2} + \cdots + b_1 + \text{zbytek}$$

$$p_n(x) = q_{n-1}(x; z)(x - z) + b_0$$
$$\text{kde } q_{n-1}(x; z) = b_n x^{n-1} + b_{n-1}x^{n-2} + \cdots + b_1$$

Je-li z kořen, pak $b_0 = 0$.

Nyní apliklujeme Newtonovu metodu pro nalezení kořene polynomu p_n.

$$\text{Newtonova metoda: } x^{(k+1)} = x^{(k)} - \frac{\overbrace{p_n(x^{(k)})}^{\text{Hornerovo sch.}}}{\underbrace{p_n'(x^{(k)})}_{\text{Hornerovo sch.}}}, \quad x^{(0)} \text{ dáno}$$

Vzorec, který dostaneme s využitím Hornerova schématu, se nazývá Newtonova-Hornerova metoda:

$$x^{(k+1)} = x^{(k)} - \frac{p_n(x^{(k)})}{q_{n-1}(x^{(k)}; x^{(k)})}$$

Výraz ve jmenovateli dostaneme z následujících vztahů

$$p_n'(x) \;=\; q_{n-1}'(x; z)(x - z) + q_{n-1}(x; z),$$

$$p_n'(z) = q_{n-1}(z; z),$$
$$z := x^{(k)}.$$

Algoritmus pro nalezení kořenů polynomu p_n:
for $m = n$ downto 1 **do**
  Najdi kořen r polynomu p_m (Newtonova metoda)
  Vyčísli koeficienty $q_{m-1}(x; r)$ (pomocí Hornerova schematu)
  $p_{m-1} := q_{m-1}$
end for

Zkušební otázka 3.3 Odvoďte Newtonovu-Hornerovu metodu nalezení kořene polynomu.

Poznámka 3.13 Začít od kořene nejmenšího v absolutní hodnotě (kvůli zaokrouhlovacím chybám).

Poznámka 3.14 Restartovat algoritmus, t.j. použít původní polynom (je-li $\tilde{r}_j$ aproximace kořene r_j, jít zpět k $p_n(x)$ a hledat novou aproximaci s $r_j^{(0)} = \tilde{r}_j$).

4

SOUSTAVY LINEÁRNÍCH ROVNIC

Hledáme $x \in I\!R^N$ takové, že

$$\mathbb{A}x = b, \quad \mathbb{A} \in I\!R^{N\times N}, \mathbb{A}\text{-nesingulární}.$$

Metody:

- **přímé** - konečný předem známý počet kroků pro nalezení řešení
- **iterační** - konstruujeme (nekonečnou) posloupnost vektorů konvergujících k řešení

4.1 Podmíněnost matic

Matice se nazývá **dobře podmíněná**, jestliže relativně malé změny v koeficientech způsobí relativně malé změny v řešení. Matice se nazývá **špatně podmíněná**, jestliže relativně malé změny v koeficientech způsobí relativně velké změny v řešení.

Analýza zaokrouhlovacích chyb - chyby ve výpočtu se obvykle reprezentují chybami ve vstupních datech. Vzhledem k zaokrouhlovacím chybám poskytuje numerická metoda přibližné řešení, které splňuje perturbovaný systém. Numerická metoda poskytuje (přesné) řešení $x + \delta x$ perturbovaného systému

$$(A + \delta A)(x + \delta x) = b + \delta b.$$

δx lze ("zhruba") odhadnout následujícím způsobem

$$\begin{aligned} x + \delta x &= (A + \delta A)^{-1}(b + \delta b) = [A(I + A^{-1}\delta A)]^{-1}(b + \delta b) \\ &= (\underbrace{I + A^{-1}\delta A}_{\text{nahradíme}})^{-1}A^{-1}(b + \delta b) \\ &\approx \underbrace{(I - A^{-1}\delta A)}_{\text{motivace}\ \downarrow}(x + A^{-1}\delta b) = x + A^{-1}\delta b - A^{-1}\delta A x - A^{-1}\delta A A^{-1}\delta b. \end{aligned}$$

$$f(x) = \frac{1}{1+x} = 1 + xf'(0) + chyba = 1 + x(-1) + chyba \quad \left(\frac{1}{1+x} \approx 1 - x\right).$$

$$\begin{gathered} \delta x \doteq A^{-1}\delta b - A^{-1}\delta A x, \\ \|\delta x\| \leq \left\|A^{-1}\right\| \|\delta b\| + \left\|A^{-1}\right\| \|\delta A\| \|x\|, \end{gathered}$$

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{\left\|A^{-1}\right\| \, \|\delta b\|}{\|x\|} + \frac{\left\|A^{-1}\right\| \, \|\delta A\| \, \|A\|}{\|A\|}$$
$$\leq \frac{\left\|A^{-1}\right\| \, \|A\| \, \|x\| \, \|\delta b\|}{\|x\| \|b\|} + \frac{\left\|A^{-1}\right\| \, \|\delta A\| \, \|A\|}{\|A\|}$$

Závěr:

$$\frac{\|\delta x\|}{\|x\|} \leq \underbrace{\|A\| \left\|A^{-1}\right\|}_{\text{číslo podmíněnosti } K(A).} \left(\frac{\|\delta b\|}{\|b\|} + \frac{\|\delta A\|}{\|A\|} \right)$$

Poznámka 4.1 Nejčastěji používané normy v $\mathbb{C}^N$, $x \in \mathbb{C}^N$, $A \in \mathbb{C}^{N \times N}$

$$\|x\|_1 = \sum_i |x_i| \,,$$
$$\|x\|_2 = \sqrt{\left(\sum_i |x_i|^2 \right)} \qquad \text{Euklidova,}$$
$$\|x\|_p = \left(\sum_i |x_i|^p \right)^{\frac{1}{p}} \qquad 1 \leq p < \infty,$$
$$\|x\|_\infty = \max_i |x_i| \,,$$
$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|},$$
$$\|A\|_1 = \max_j \sum_i \underbrace{|a_{ij}|}_{\text{sloupcový součet}} \,,$$
$$\|A\|_2 = \sqrt{\rho(A^H A)} = \sqrt{\rho(A A^H)},$$

A^H – transponovaná a kompl. združená (hermitovská),

$\rho(B)$ – největší v abs. hodnotě vlastní číslo B (spektrální poloměr),

$$\|A\|_F = \sqrt{\sum_{i,j} |a_{ij}|^2} \qquad \text{Frobeniova,}$$
$$\|A\|_\infty = \max_i \sum_j |a_{ij}| \qquad \text{řádkový součet,}$$

- $\|I\|_F = \sqrt{N}$
- $\|I\| = 1$, $\|Ax\| \leq \|A\| \cdot \|x\|$
- $\|AB\| \leq \|A\| \, \|B\|$ sub-multiplikativita

4.2 Gaußova eliminace

Cíl:

$$Ax = b \Leftrightarrow Ux = \hat{b}, \qquad \text{kde } U \text{ je horní trojúhelníková}$$

Algoritmus 4.2

for sloupec $j = 1$ to $n - 1$ **do**
  najdi $a_{pj} \neq 0, \qquad p \in \{j, \dots, n\}$
  if $a_{pj} = 0\ \forall p$ **then**
    STOP (singularita)
  else
    záměna p a j-tého řádku
  end if
  for řádek $i = j + 1$ to n **do**
    $l_{ij} = \frac{a_{ij}}{a_{jj}}$;
    for $k = j + 1$ to n **do**
      $a_{ik} = a_{ik} - l_{ij} a_{jk}$;
    end for
    $b_i = b_i - l_{ij} b_j$;
  end for
end for

$$\begin{array}{ll} u_{ij},\ i \leq j & \text{jsou pak poslední hodnoty } a_{ij} \\ \hat{b}_i & \text{jsou pak poslední hodnoty } b_i \end{array}$$

Počet operací	v j-tém kroku	celkem
Hledání $a_{pj} \neq 0$	$n - j + 1$	$\sum_{j=2}^{n} j = \frac{(2+n)(n-1)}{2}$
Výpočet l_{ij}	$n - j$	$\sum_{j=1}^{n-1} j = \frac{n(n-1)}{2}$
Výpočet a_{ik}	$2(n-j)^2$	$2\sum_{j=1}^{n-1} j^2 = 2\frac{2n^3-3n^2+n}{6}$
Výpočet b_i	$2(n-j)$	$2\sum_{j=1}^{n-1} j = 2\frac{n(n-1)}{2}$

$$\text{Celkový počet operací: } \frac{2}{3}n^3 + O(n^2)$$

Počet operací pro řešení $Ux = \hat{b}$:

násobení	sčítání
$\frac{(n+1)n}{2}$	$\frac{n(n-1)}{2}$

Zkušební otázka 4.1 Zdůvodněte odhad počtu operací v Gaußově eliminaci.

4.2.1 *Pivotace*

Výpočet $l_{ij} = \frac{a_{ij}}{a_{jj}}$ v Algoritmu 4.2, $a_{jj} \neq 0$.

$$\text{Částečná pivotace} \quad |a_{pj}| = \max_{l=j,\dots,n} |a_{lj}| \tag{4.2.1}$$

$$\text{Úplná pivotace} \quad |a_{pj}| = \max_{l,m=j,\dots,n} |a_{lm}| \tag{4.2.2}$$

Důvod: I když Guaßova eliminace je proveditelná bez záměny řádků a sloupců, mohou malé hodnoty a_{jj} způsobit velké chyby v řešení.

Příklad 4.2

$$\begin{pmatrix} 6 & 1 & & & \\ 8 & 6 & 1 & & \\ & 8 & 6 & 1 & \\ & & \ddots & \ddots & \ddots \\ & & & 8 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{50} \end{pmatrix} = \begin{pmatrix} 7 \\ 15 \\ 15 \\ \vdots \\ 14 \end{pmatrix}, \qquad x_{GE} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ -3 \times 10^7 \end{pmatrix}$$

Gaußova eliminace je numericky nestabilní. Pivotace je podstatná pro stabilitu elim. procesu. Ani velké hodnoty pivotů však nejsou zárukou dostatečně přesného řešení.

Důvod: velké změny v koeficientech

Náprava: škálování, dělení i-tého řádku $d_i = \sum_{j=1}^{n} |a_{ij}|$, ale toto dělení opět vnáší zaokrouhlovací chyby.

7. přednáška

4.3 Gaußova eliminace jako faktorizační metoda

$$Ax = b \Leftrightarrow LUx = b \quad \begin{cases} Ux = \hat{b}, \\ L\hat{b} = b. \end{cases} \tag{4.3.1}$$

Nechť P_j je matice, která v j-tém kroku Gaußovy eliminace realizuje záměnu p-tého a j-tého řádku matice A v Algoritmu 4.2

$$\begin{array}{c c} & \begin{array}{cccccc} & j & & & p & \end{array} \\ \begin{array}{c} \\ j \\ \\ \\ p \\ \\ \end{array} & \begin{pmatrix} 1 & \cdot & & & \cdot & \\ \cdot & 0 & \cdot & \cdot & 1 & \cdot \\ & \cdot & 1 & & \cdot & \\ & \cdot & & 1 & \cdot & \\ \cdot & 1 & \cdot & \cdot & 0 & \cdot \\ & \cdot & & & \cdot & 1 \end{pmatrix} \end{array} \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ \circ & \circ & \circ & \circ & \circ & \circ \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \diamond & \diamond & \diamond & \diamond & \diamond & \diamond \\ \times & \times & \times & \times & \times & \times \end{pmatrix} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ \diamond & \diamond & \diamond & \diamond & \diamond & \diamond \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \circ & \circ & \circ & \circ & \circ & \circ \\ \times & \times & \times & \times & \times & \times \end{pmatrix}$$

a nechť L_j je matice, pomocí níž se provádí nulování prvků j-tého sloupce pod diagonálou.

$$\begin{pmatrix} 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & -\ell_{43} & 1 & \cdot & \cdot \\ \cdot & \cdot & -\ell_{53} & \cdot & 1 & \cdot \\ \cdot & \cdot & -\ell_{63} & \cdot & \cdot & 1 \end{pmatrix} \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \end{pmatrix} = \begin{pmatrix} \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \end{pmatrix}$$

Algoritmus Gaußovy eliminace lze maticově zapsat (GE s částeční pivotací):

$$\underbrace{L_{n-1}P_{n-1}\cdots L_1P_1}_{M} A = U.$$

Označme

$$P = P_{n-1}\cdots P_1,$$

$$M = L_{n-1}P_{n-1}\cdots L_1P_1.$$

Potom

$$\begin{aligned} MA &= U, \\ MP^{-1}PA &= U, \\ PA &= \underbrace{PM^{-1}}_{L}U, \\ PA &= LU. \end{aligned}$$

Lze-li provést Gaußovou eliminaci bez záměny řádků a sloupců, dostáváme

$$A = LU. \tag{4.3.2}$$

Věta 4.3 *Nechť $A \in \mathbb{R}^{n\times n}, A$ regulární. Pak existuje permutační matice $P \in \mathbb{R}^{n\times n}$, nesingulární U a L s jedničkami na diagonále ::*

$$PA = LU \tag{4.3.3}$$

Algoritmus

Matici L výše uvedenou dostaneme pomocí Algoritmu 4.2 tak, že l_{ij} uložíme do a_{ij}, jejichž hodnoty nejsou v Gaußově eliminaci potřeba a při pivotaci je zaměníme.

Řešení úlohy $Ax = b$ ve třech krocích

1. $PA = LU$
2. $PAx = L\underbrace{Ux}_{\hat{b}} = Pb$

   $L\hat{b} = Pb$
3. $Ux = \hat{b}$

Měření kvality řešení: $r = b - A\tilde{x}$ - reziduum

Věta 4.4 *(Odhad rezidua) [Prager/Oettli]*

Nechť $\tilde{x}$ je přibližné řešení $Ax = b$, $r = b - A\tilde{x}$ reziduum. Nechť je dáno $0 \le \delta A \in \mathbb{R}^{n\times n}, 0 \le \delta b \in \mathbb{R}^n$. Pak $\tilde{x}$ je přesné řešení

$$\tilde{A}\tilde{x} = \tilde{b}, \qquad \textit{kde} \tag{4.3.4}$$

$$\left|\tilde{A} - A\right| \le \delta A, \ \left|\tilde{b} - b\right| \le \delta b \qquad \textit{(po složkách)} \tag{4.3.5}$$

právě když

$$|r| \le \delta A\,|\tilde{x}| + \delta b. \tag{4.3.6}$$

Důkaz (pouze $\Rightarrow$)

Nechť $\tilde{A}\tilde{x} = \tilde{b}$ ($\tilde{x}$ je přesné řešení perturbovaného systému) a pro perturbace platí odhad

$$\left|\tilde{A} - A\right| \leq \delta A$$
$$\left|\tilde{b} - b\right| \leq \delta b$$

$$\begin{aligned} \tilde{A} &= A + \Delta A \qquad & |\Delta A| \leq \delta A \\ \tilde{b} &= b + \Delta b \qquad & |\Delta b| \leq \delta b \end{aligned}$$

$$\begin{aligned} |r| = |b - A\tilde{x}| = &\left|\tilde{b} - \Delta b - \underbrace{\tilde{A}\tilde{x}}_{\tilde{b}} + \Delta A\tilde{x}\right| \leq \\ &\leq |\Delta A\tilde{x} - \Delta b| \leq \delta A\,|\tilde{x}| + \delta b. \end{aligned}$$

□

4.4 *LU* rozklad v obecném případě

$$\begin{aligned} A &= \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \qquad \exists!\ LU \text{ rozklad} \\ A &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{neexistuje } LU \text{ rozklad} \\ A &= \begin{pmatrix} 0 & 1 \\ 0 & 2 \end{pmatrix} \quad LU \text{ není jednoznačný} \end{aligned}$$

Při konstrukci LU rozkladu matice $A \in \mathbb{R}^{n\times n}$ postupujeme tak, že postupně počítáme m-tý řádek matice U a m-tý sloupec matice L, $m = 1, \dots n$. Příslušné vzorce odvodíme pomocí vzorce pro násobení matic.

$$A = LU,$$

$$a_{ij} = \sum_{k=1}^{n} l_{ik}u_{kj}.$$

Máme n^2 rovnic pro určení neznámých l_{ij}, $i \leq j$ a u_{ij}, $i \geq j$ (prvků dolní trojúhelníkové matice L a horní trojúhelníkové matice U). Počet neznámých je $2\,(1 + n)n/2 = n^2 + n$. Předepíšeme tedy hodnoty některých prvků, například položíme diagonální prvky matice L rovny jedné. Dostáváme následující vzorce pro $m = 1, \dots, n$:

m-tý řádek matice U, u_{mj}, $j \geq m$ (křížky označují již spočtené hodnoty, počítáme prvek $\diamond$:

$$
\underbrace{\begin{array}{c c}
 & \begin{array}{cccccc} & & j & & & \end{array} \\
m & \left(\begin{array}{cccccc}
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \boxed{\times} & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
\end{array}}_{A}
=
\underbrace{\begin{array}{c c}
m & \left(\begin{array}{cccccc}
1 & & & & & \\
\times & 1 & & & & \\
\times & \times & 1 & & & \\
\times & \times & \cdot & 1 & & \\
\times & \times & \cdot & \cdot & 1 & \\
\times & \times & \cdot & \cdot & \cdot & 1
\end{array}\right)
\end{array}}_{L}
\underbrace{\begin{array}{c}
\begin{array}{cccccc} & & j & & & \end{array} \\
\left(\begin{array}{cccccc}
\times & \times & \times & \times & \times & \times \\
 & \times & \times & \times & \times & \times \\
 & & \diamond & \rightarrow & \cdot & \cdot \\
 & & & \cdot & \cdot & \cdot \\
 & & & & \cdot & \cdot \\
 & & & & & \cdot
\end{array}\right)
\end{array}}_{U}
$$

$$a_{mj} = \sum_{k=1}^{n} l_{mk} u_{kj} = \sum_{k=1}^{m-1} l_{mk} u_{kj} + 1 \cdot u_{mj}, \quad u_{mj} = \diamond,$$

m-tý sloupec matice L, l_{im}, $i > m$:

$$
\underbrace{\begin{array}{c c}
 & \begin{array}{cccccc} & & m & & & \end{array} \\
i & \left(\begin{array}{cccccc}
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \boxed{\times} & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
\end{array}}_{A}
=
\underbrace{\begin{array}{c c}
i & \left(\begin{array}{cccccc}
1 & \cdot & \cdot & \cdot & \cdot & \cdot \\
\times & 1 & \cdot & \cdot & \cdot & \cdot \\
\times & \times & 1 & \cdot & \cdot & \cdot \\
\times & \times & \circ & 1 & \cdot & \cdot \\
\times & \times & \downarrow & \cdot & 1 & \cdot \\
\times & \times & \cdot & \cdot & \cdot & 1
\end{array}\right)
\end{array}}_{L}
\underbrace{\begin{array}{c}
\begin{array}{cccccc} & & m & & & \end{array} \\
\left(\begin{array}{cccccc}
\times & \times & \times & \times & \times & \times \\
\cdot & \times & \times & \times & \times & \times \\
\cdot & \cdot & \diamond & \diamond & \diamond & \diamond \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
\end{array}}_{U}
$$

$$a_{im} = \sum_{k=1}^{n} l_{ik} u_{km} = \sum_{k=1}^{m-1} l_{ik} u_{km} + l_{im} \cdot u_{mm}, \quad l_{im} = \circ.$$

Zkušební otázka 4.2! Odvoďte vzorce pro konstrukci LU rozkladu matice A.

Věta 4.5 *Nechť $A \in \mathbb{R}^{n \times n}$ je obecná matice. Faktorizace $A = LU$ existuje a je jednoznačná právě když všechny hlavní minory A, t.j.* $\det \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ & \ddots & \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$, $k = 1, \ldots, n-1$ *jsou nenulové.*

Věta 4.6 *Je-li matice řádkově nebo sloupcově diagonálně dominantní, t.j.*

$$|a_{ii}| \geq \sum_{j=1, j \neq i}^{n} |a_{ij}|, \qquad \textit{(řádkově)} \tag{4.4.1}$$

nebo

$$|a_{jj}| \geq \sum_{i=1, i \neq j}^{n} |a_{ij}|, \qquad \textit{(sloupcově)} \tag{4.4.2}$$

pak LU rozklad existuje. Speciálně, je-li matice sloupcově diagonálně dominantní, je $|l_{ij}| \leq 1 \; \forall i, j = 1, \ldots, n$.

4.4.1 *Vliv zaokrouhlovacích chyb*

Uvažujeme-li zaokrouhlovací chyby, faktorizační proces produkuje matice $\hat{L}$, $\hat{U}$ takové, že

$$\hat{L}\hat{U} = A + \delta A. \tag{4.4.3}$$

Lze odhadnout (viz (Higham, 1989))

$$|\delta A| \leq \frac{nu}{1-nu}\left|\hat{L}\right|\left|\hat{U}\right|, \qquad u = \frac{1}{2}\varepsilon_M, \tag{4.4.4}$$

kde $B = |A|$ znamená matici $n \times n$ s prvky $b_{ij} = |a_{ij}|$, $C \leq D$ má význam $c_{ij} \leq d_{ij}$ (po prvcích), $i, j = 1, \ldots n$ a ε_M je nejmenší číslo číslo takové, že $1 + \varepsilon_M > 1$ (strojové epsilon, roundoff unit). Z (4.4.4) je vidět ($l_{ij} = \frac{a_{ij}}{a_{jj}}$, viz Gaußova eliminace), že přítomnost malých pivotů může způsobit neomezenost pravé strany a v důsledku toho ztrátu kontroly kontroly δA. Je tedy vhodné najít odhad

$$|\delta A| \leq \underbrace{g(u)}_{\text{vhodná funkce}} |A|$$

Příklad 4.7 Nechť $\hat{L} \geq 0$, $\hat{U} \geq 0$, pak $\left|\hat{L}\right|\left|\hat{U}\right| = \left|\hat{L}\hat{U}\right|$

$$\left|\hat{L}\right|\left|\hat{U}\right| = \left|\hat{L}\hat{U}\right| = |A + \delta A| \leq |A| + |\delta A| \leq |A| + \frac{nu}{1-nu}\left|\hat{L}\right|\left|\hat{U}\right|$$

Odtud

$$\left|\hat{L}\right|\left|\hat{U}\right|\left(1 - \frac{nu}{1-nu}\right) \leq |A|$$

$$\left|\hat{L}\right|\left|\hat{U}\right| \leq \left(\frac{1-2nu}{1-nu}\right)^{-1} |A|$$

a z 4.4.4 dostáváme

$$|\delta A| \leq \underbrace{\frac{nu}{1-2nu}}_{g(u)} |A| \tag{4.4.5}$$

Pivotace umožňuje obdržet odhad obdobný (4.4.5) pro libovolnou matici.

4.5 Choleského rozklad

Věta 4.8 *Pro každou symetrickou, pozitivně definitní ($x^TAx > 0$, $\forall x \neq 0$, $x \in \mathbb{R}^n$) matici $A \in \mathbb{R}^{n\times n}$ existuje právě jedna dolní trojúhelníková matice L s kladnými prvky na diagonále tak, že platí*

$$A = L \cdot L^T \tag{4.5.1}$$

Důkaz indukcí □

Věta 4.9 *Nechť $A \in \mathbb{R}^{n\times n}$ je symetrická, ostře diag. dominantní ($|a_{ii}| > \sum_{j\neq i}|a_{ij}|$), $a_{ii} > 0$, pak A je pozitivně definitní.*

4.6 QR rozklad

Věta 4.10 *Ke každé nesingulární matici $A \in \mathbb{R}^{n\times n}$ existuje ortogonální matice $Q \in \mathbb{R}^{n\times n}$ ($QQ^T = Q^TQ = I$) a nesingulární horní trojúhelníková R taková, že*

$$A = Q \cdot R. \tag{4.6.1}$$

Poznámka 4.11 Transformace, která (na rozdíl od LU) nezvyšuje číslo podmíněnosti ($K(U) \leq 4^{n-1}K(PA)$).

5

ITERAČNÍ METODY ŘEŠENÍ SOUSTAV LINEÁRNÍCH ROVNIC

8. přednáška

$$A \in I\!R^{n\times n},\ b \in I\!R^n,\ \det A \neq 0.$$

Hledáme $x \in I\!R^n$::

$$Ax = b. \tag{5.0.1}$$

Přímé metody (např. Gaußova eliminace) :

- pro libovolné plné matice
- počet operací $O(\frac{2}{3}n^3)$

Nevýhoda:

a) nevyužívají informaci o struktuře matice (řídkost, blokově diagonální)
b) nákladné, je-li n velké
c) pro řídké matice mohou být nevhodné (zaplnění)

Iterační metody

- formálně poskytují řešení po nekonečném počtu kroků
- v každém kroku požadují výpočet rezidua, výpočetní náročnost $O(n^2)$
- mohou soupeřit s přímými metodami, je-li počet iterací k získání řešení s danou tolerancí nezávislý na n nebo menší než n
- používají se, stačí-li získat řešení pouze s určitou přesností (Fyzika → model → matematický model)

Idea iteračních metod: konstrukce $\{x^{(k)}\}$

$$x = \lim_{k\to\infty} x^{(k)}, \quad \text{kde } x \text{ je řešení Ax=b.} \tag{5.0.2}$$

Poznámka 5.1 Posloupnost $x^{(0)}, x^{(1)}, \dots, x^{(k)}, \dots,$ je nekonečná, cílem je nalezení řešení x^* s předepsanou přesností, t.j. $\left\|x^* - x^{(k)}\right\| \le \varepsilon$. Otázkou je určení vhodného stopping kriteria (např. omezenost rezidua $\left\|b - Ax^{(k)}\right\| \le \varepsilon$).

Princip iteračních metod je na základě předchozích aproximací konstrukce nové aproximace

$$x^{(k+1)} = \varphi(x^{(k)}), \quad \text{resp. } x^{(k+1)} = \varphi(x^{(k)}, x^{(k-1)}, \dots, x^{(0)})$$

takové, že

$$x^{(k+1)} \to x \text{ pro } k \to \infty,$$

kde x je hledané řešení. Požadavky:

- rychlá konvergence
- snadné vyčíslení φ (méně operací než matice $\times$ vektor, řádově $O(n)$)
- řešení s předepsanou přesností

5.1 Klasické iterační metody

Idea: věta o pevném bodě

$$Ax = b \Leftrightarrow x = G(x). \tag{5.1.1}$$

Pro dané $x^{(0)}$ se řešení hledá jako limita posloupnosti $x^{(k+1)} = G(x^{(k)})$.

1. Richardsonova metoda

$$\begin{aligned} x &= x + b - Ax, \\ x &= \underbrace{(I - A)}_{B_R} x + b, \\ x^{(k+1)} &= B_R x^{(k)} + b. \end{aligned}$$

2. Jacobiho metoda

$$A = E + D + F, \quad \text{kde}$$

$$\begin{array}{ll} E & \text{je ostře dolní trojúhelníková,} \\ D & \text{je diagonální,} \\ F & \text{je ostře horní trojúhelníková.} \end{array}$$

$$\begin{aligned} Ax = b &\iff (E + D + F)x = b, \\ Dx &= -(E + F)x + b, \\ x &= \underbrace{-D^{-1}(E + F)}_{B_J} x + \underbrace{D^{-1}b}_{f_J}, \\ x^{(k+1)} &= B_J x^{(k)} + f_J. \end{aligned}$$

3. Gaußova–Seidelova metoda

$$\begin{aligned} Ax = b &\iff (D + E)x + Fx = b, \\ (D + E)x &= -Fx + b, \\ x &= \underbrace{-(D + E)^{-1}F}_{B_{GS}} x + \underbrace{(D + E)^{-1}b}_{f_{GS}} \\ x^{(k+1)} &= B_{GS} x^{(k)} + f_{GS}. \end{aligned}$$

Poznámka 5.2 Porovnejme způsob algoritmizace Jacobiho a Gaußovy-Seidelovy metody. K tomu je třeba si nejprve uvědomit, že

$$D^{-1} = \begin{pmatrix} \frac{1}{a_{11}} & & & & & \\ & \frac{1}{a_{22}} & & & & \\ & & \frac{1}{a_{33}} & & & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & \frac{1}{a_{nn}} \end{pmatrix},$$

$$D^{-1}X = \begin{pmatrix} \frac{1}{a_{11}}\times & \frac{1}{a_{11}}\times & \frac{1}{a_{11}}\times & \frac{1}{a_{11}}\times & \frac{1}{a_{11}}\times & \frac{1}{a_{11}}\times \\ \frac{1}{a_{22}}\times & \frac{1}{a_{22}}\times & \frac{1}{a_{22}}\times & \frac{1}{a_{22}}\times & \frac{1}{a_{22}}\times & \frac{1}{a_{22}}\times \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{1}{a_{nn}}\times & \frac{1}{a_{nn}}\times & \frac{1}{a_{nn}}\times & \frac{1}{a_{nn}}\times & \frac{1}{a_{nn}}\times & \frac{1}{a_{nn}}\times \end{pmatrix}.$$

Vyčíslení inverzní matice v Gaußově-Seidelově metodě se vyhneme následujícím způsobem. Na základě vyjádření

$$x^{(k+1)} = \underbrace{-(D+E)^{-1}F}_{B_{GS}}\, x^{(k)} + \underbrace{(D+E)^{-1}b}_{f_{GS}}$$

přepíšeme Gaußovu–Seidelovu metodu ve tvaru

$$\begin{aligned} (D+E)x^{(k+1)} &= -Fx^{(k)} + b, \\ Dx^{(k+1)} &= -Ex^{(k+1)} - Fx^{(k)} + b, \\ x^{(k+1)} &= -D^{-1}Ex^{(k+1)} - D^{-1}Fx^{(k)} + D^{-1}b, \end{aligned}$$

$$x^{(k+1)} = -D^{-1}E \quad x^{(k+1)} \;-\; D^{-1}F \quad x^{(k)} \;+\; \underbrace{D^{-1}b}_{f_J}, \tag{5.1.2}$$

t.j.

$$\overset{x^{(k+1)}}{\begin{pmatrix} \times \\ \times \\ \times \\ \times \end{pmatrix}} = - \begin{pmatrix} \cdot & & & \\ \times & \cdot & & \\ \times & \times & \cdot & \\ \times & \times & \times & \cdot \end{pmatrix} \overset{x^{(k+1)}}{\begin{pmatrix} \times \\ \times \\ \times \\ \times \end{pmatrix}} - \begin{pmatrix} \cdot & \times & \times & \times \\ & \cdot & \times & \times \\ & & \cdot & \times \\ & & & \cdot \end{pmatrix} \overset{x^{(k)}}{\begin{pmatrix} \times \\ \times \\ \times \\ \times \end{pmatrix}}$$

$$
+\begin{array}{c} {\color{red} f_J} \\ \begin{pmatrix} \times \\ \times \\ \times \\ \times \end{pmatrix} \end{array}.
$$

Při použití Jacobiho metody

$$
x^{(k+1)} = B_J x^{(k)} + f_J,
$$

t.j.

$$
\begin{array}{c} x^{k+1} \\ \begin{pmatrix} \times \\ \times \\ \times \\ \times \end{pmatrix} \end{array} = - \begin{array}{c} \\ \begin{pmatrix} \cdot & \times & \times & \times \\ \times & \cdot & \times & \times \\ \times & \times & \cdot & \times \\ \times & \times & \times & \cdot \end{pmatrix} \end{array} \begin{array}{c} x^{k} \\ \begin{pmatrix} \times \\ \times \\ \times \\ \times \end{pmatrix} \end{array} + \begin{array}{c} f_J \\ \begin{pmatrix} \times \\ \times \\ \times \\ \times \end{pmatrix} \end{array},
$$

je třeba si pamatovat celý vektor $x^{(k)}$ pro výpočet nové iterace $x^{(k+1)}$. U metody Gaußovy-Seidelovy se v paměti počítače rezervuje místo pro jediný vektor $x^{(k)}$, na jehož místo se postupně ukládají složky vektoru $x^{(k+1)}$ jak vyplývá z rozepsání po složkách vztahu (5.1.2):

$$
\begin{aligned}
\textstyle\sum_{j=1}^{n} a_{ij} x_j &= b_i, \\
\textstyle\sum_{j=1}^{i-1} a_{ij} x_j + a_{ii} x_i + \sum_{j=i+1}^{n} a_{ij} x_j &= b_i, \\
a_{ii} x_i &= -\textstyle\sum_{j=1}^{i-1} a_{ij} x_j - \sum_{j=i+1}^{n} a_{ij} x_j + b_i, \\
x_i^{(k+1)} &= \tfrac{1}{a_{ii}} \left(-\textstyle\sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)} \right) + \tfrac{b_i}{a_{ii}}.
\end{aligned}
$$

Dostáváme tak algoritmus Gaußovy–Seidelovy metody, který lze vyjádřit následujícím způsobem. Vyjdeme z Jacobiho metody a spočtenou složku $x_i^{(k+1)}$ uložíme do $x_i^{(k)}$ a následně počítáme $x_{i+1}^{(k+1)}, i = 1, \ldots, n$.

$$
\begin{array}{cc} & x^{k+1} \\ \begin{array}{c} \\ i+1 \\ \\ \\ \end{array} & \begin{pmatrix} \boxed{\times} \\ \boxed{\times} \\ \times \\ \times \end{pmatrix} \end{array} = - \begin{array}{c} \\ \begin{pmatrix} \cdot & \times & \times & \times \\ \times & \cdot & \times & \times \\ \times & \times & \cdot & \times \\ \times & \times & \times & \cdot \end{pmatrix} \end{array} \begin{array}{c} x^{k} \\ \begin{pmatrix} \boxed{\times} \\ \times \\ \times \\ \times \end{pmatrix} \end{array} + \begin{array}{c} f_J \\ \begin{pmatrix} \times \\ \times \\ \times \\ \times \end{pmatrix} \end{array}.
$$

4. Metoda SOR (superrelaxační)

$$
\begin{aligned}
Ax &= (E + D + F)x = b, \\
\tilde{x}^{(k+1)} &= -D^{-1} E x^{(k+1)} - D^{-1} F x^{(k)} + D^{-1} b, \\
x^{(k+1)} &= x^{(k)} + \omega(\tilde{x}^{(k+1)} - x^{(k)}),
\end{aligned}
$$

$$
\begin{aligned}
x^{(k+1)} &= (1-\omega)x^{(k)} + \omega\tilde{x}^{(k+1)},\\
x^{(k+1)} &= -\omega D^{-1}Ex^{(k+1)} + \left[(1-\omega)I - \omega D^{-1}F\right]x^{(k)} + \omega D^{-1}b,\\
\\
x^{(k+1)} &= \left(D^{-1}ID + \omega D^{-1}E\right)^{-1}\left[(1-\omega)D^{-1}ID - \omega D^{-1}F\right]x^{(k)}\\
&\quad + (D^{-1}ID + \omega D^{-1}E)^{-1}\omega D^{-1}b,\\
\\
x^{(k+1)} &= \underbrace{(D+\omega E)^{-1}[(1-\omega)D - \omega F]}_{B_{SOR}}x^{(k)}\\
&\quad + \underbrace{(D+\omega E)^{-1}\omega b}_{f_{SOR}}.
\end{aligned}
$$

$$x^{(k+1)} = B_{SOR}x^{(k)} + f_{SOR}.$$

Pro výpočty pomocí výše uvedených metod se používá jejich zápis do složek:

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

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

$$\tilde{x}_i^{(k+1)} = \frac{1}{a_{ii}}\left(-\sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}\right) + \frac{b_i}{a_{ii}}, \qquad \text{(SOR)}$$

$$x_i^{(k+1)} = x_i^{(k)} + \omega(\tilde{x}_i^{(k+1)} - x_i^{(k)}).$$

Zkušební otázka 5.1! Odvoďte Jacobiho, Gaußovu–Seidelovu a SOR metodu pro řešení úlohy $Ax = b$. Zapište je maticově a rozepsané do složek bez použití inverze matic.

Uvažujme iterační metodu

$$x^{(k+1)} = Bx^{(k)} + f. \tag{5.1.3}$$

Definice 5.3 *Řekneme, že iterační metoda $x^{(k+1)} = Bx^{(k)} + f$ je konzistentní s $Ax = b$, jestliže*

$$x = Bx + f, \text{ kde } x \text{ je řešení úlohy } Ax = b.$$

Ekvivalentně

$$f = (I - B)x = (I - B)A^{-1}b.$$

Věta 5.4 *Nechť* $x^{(k+1)} = Bx^{(k)} + f$ *je konzistentní metoda. Pak posloupnost* $\{x^{(k)}\}$ *konverguje k* x^**, kde* x^* *splňuje* $Ax^* = b$*, pro libovolné* $x^{(0)}$*, právě když* $\rho(B)$ *(spektrální poloměr matice* B*,* $\rho(B) = \max_{\{\lambda \text{ vl. č. } B\}} |\lambda|$*) je menší než* 1.

Důkaz

$$\begin{aligned} x^* &= Bx^* + f \quad \text{(podmínka konzistence)},\\ x^{(k+1)} &= Bx^{(k)} + f \quad \text{(iterační metoda)},\\ e^{(k+1)} &:= x^* - x^{(k+1)} \quad \text{(chyba v)}(k+1)\text{-ní iteraci)}. \end{aligned}$$

Chybu v k-té iteraci lze vyjádřit jako součin k-té mocniny matice B a chyby počáteční aproximace

$$e^{(k)} = Be^{(k-1)} = B^2 e^{(k-2)} = \cdots = B^k e^{(0)}.$$

Podle definice limity

$$x^{(k)} \to x^* \Leftrightarrow \left\| e^{(k)} \right\| \to 0.$$

Platí

$$\left\| e^{(k)} \right\| \to 0 \Leftrightarrow \left\| B^k e^{(0)} \right\| \to 0 \Leftrightarrow \rho(B) < 1.$$

Poslední ekvivalenci dokážeme na základě následující věty z algebry. Vyhneme se tak klasickému důkazu pomocí převedení matice B na **Jordanův** kanonický tvar.

Lemma 5.5 *Nechť* $A \in \mathbb{C}^{n\times n}$*,* $\varepsilon > 0$*. Pak existuje konzistentní* ($\|Ax\| \leq \|A\| \, \|x\|$) *maticová norma* $\|.\|_{A,\varepsilon}$ *taková, že*

$$\|A\|_{A,\varepsilon} \leq \rho(A) + \varepsilon.$$

Pokračování v důkazu předchozí věty

$\boxed{\Leftarrow}$ Nechť $\rho(B) < 1$. Potom $\exists \varepsilon :: \rho(B) < 1 - \varepsilon$ a dále existuje $\|.\|_{B,\varepsilon} ::$

$$\|B\|_{B,\varepsilon} \leq \rho(B) + \varepsilon < 1$$

a tedy

$$\left\|B^k\right\|_{B,\varepsilon} \leq \|B\|_{B,\varepsilon}^k \to 0.$$

Platí tedy

$$\left\| e^{(k)} \right\| = \left\| B^k e^{(0)} \right\| \leq \left\| B^k \right\| \left\| e^{(0)} \right\| \leq \|B\|^k \left\| e^{(0)} \right\| \to 0.$$

$\boxed{\Rightarrow}$ Předpokládejme sporem, že $\rho(B) > 1$. Existuje tedy vlastní číslo λ matice B takové, že $|\lambda| > 1$. Zvolme počáteční aproximaci $x^{(0)}$ tak, že $e^{(0)}$ je vlastní vektor odpovídající vlastnímu číslu λ. Potom

$$e^{(k)} = B^k e^{(0)} = \lambda^k e^{(0)}.$$

V důsledku tohoto vztahu $e^{(k)}$ nekonverguje k nule (pro danou volbu $x^{(0)}$), protože $|\lambda| > 1$.

$\square$

Poznámka 5.6 Lemma 5.5 jsme využili pro důkaz vztahu $\rho(B) < 1 \Rightarrow B^k \to 0$. Obrácená implikace se dokáže snadno.

Pro důkaz konvergence výše uvedených klasických iteračních metod se využívá řada kritérií, která vycházejí z přímo z vlastností matice A. Detaily viz cvičení k přednášce.

6

VÝPOČET VLASTNÍCH ČÍSEL MATIC

<u>9. přednáška</u>

$$A \in \mathbb{C}^{n\times n}.$$

Hledáme $\lambda \in \mathbb{C} :: \exists x \in \mathbb{C}^n,\ x \neq 0,$

$$Ax = \lambda x. \tag{6.0.1}$$

- aplikace: kvantová mechanika, strukturální vibrace, analýza elektrických sítí, analýza numerických metod (výpočet optimálních parametrů relaxačních metod), analýza stability numerických metod pro řešení soustav obyčejných diferenciálních rovnic
- omezíme se na výpočet dominantního vlastního čísla

6.1 Mocninná metoda

$$A \in \mathbb{C}^{n\times n},\ A \text{ diagonalizovatelná}$$

$$A = X\Lambda X^{-1},\ X = \begin{pmatrix} \vdots & & \vdots \\ x_1 & \cdots & x_n \\ \vdots & & \vdots \end{pmatrix} \in \mathbb{C}^{n\times n},$$

x_i vlastní vektory ($Ax_i = \lambda_i x_i$), $\|x_i\| = 1$. Nechť $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots |\lambda_n|$, λ_1 má násobnost 1. Pak λ_1 nazveme dominantním vlastním číslem.

Nechť je dáno $q^{(0)} \in \mathbb{C}^n$, $\left\|q^{(0)}\right\| = 1$ ($\|\cdot\| := \|\cdot\|_2$ - Euklidovská). Konstruujeme posloupnost vektorů

$$q^{(k)} = \frac{Aq^{(k-1)}}{\left\|Aq^{(k-1)}\right\|} = \cdots = \frac{A^k q^{(0)}}{\left\|A^k q^{(0)}\right\|} \text{ (odtud název mocninná metoda).}$$

Je-li A diagonalizovatelná, má matice X za sloupce vlastní vektory matice A. Tyto vlastní vektory jsou lineárně nezávislé a tvoří bázi $\mathbb{C}^n$. Lze tedy psát:

$$q^{(0)} = \sum_{i=1}^{n} \alpha_i x_i,\ \alpha_i \in \mathbb{C},\ i = 1, \ldots, n.$$

Budeme uvažovat takové $q^{(0)}$, pro které $\alpha_1 \neq 0$. Jinak bychom v dalším postupu narazili na problém dělení nulou. Dále $Ax_i = \lambda_i x_i$ pro $i = 1, \dots, n$ a vytkneme-li $\alpha_1 \lambda_1^k$, dostaneme

$$q^{(k)} = \frac{\sum_{i=1}^n \alpha_i \lambda_i^k x_i}{\left\| \sum_{i=1}^n \alpha_i \lambda_i^k x_i \right\|} = \frac{\alpha_1 \lambda_1^k (x_1 + y^{(k)})}{|\alpha_1 \lambda_1^k| \left\| x_1 + y^{(k)} \right\|}$$

kde $(\sum_{i=1}^n \alpha_i \lambda_i^k x_i = \alpha_1 \lambda_1^k (x_1 + \frac{\alpha_2 \lambda_2}{\alpha_1 \lambda_1^k} x_2 + \ldots + \frac{\alpha_n \lambda_n}{\alpha_1 \lambda_1^k} x_n)$

$$y^{(k)} = \sum_{i=2}^n \frac{\alpha_i}{\alpha_1} \left(\frac{\lambda_i}{\lambda_1} \right)^k x_i$$

a v důsledku přepokladu, že λ_1 je dominantní vlastní číslo matice A a $\|x_i\| = 1$,

$$\left\| y^{(k)} \right\| = \left\| \sum_{i=2}^n \frac{\alpha_i}{\alpha_1} \left(\frac{\lambda_i}{\lambda_1} \right)^k x_i \right\| \leq \underbrace{\sum_{i=2}^n \left| \frac{\alpha_i}{\alpha_1} \right|}_{C} \left| \frac{\lambda_2}{\lambda_1} \right|^k \leq C \left| \frac{\lambda_2}{\lambda_1} \right|^k .$$

Odtud dostáváme

$$y^{(k)} \to 0 \qquad \text{pro } k \to \infty.$$

Pro $k \to \infty$ se tedy směr $q^{(k)}$ bude blížit směru x_1. O rychlosti konvergence rozhoduje podíl $|\lambda_2/\lambda_1|$.

Uvažujme $Aq^{(k)}$ a ${q^{(k)}}^{\mathrm{H}} A q^{(k)}$ ($x^{\mathrm{H}} = \overline{x}^{\mathrm{T}}$). Ukážeme, že

$${q^{(k)}}^{\mathrm{H}} A q^{(k)} \to \lambda_1 \qquad \text{pro } k \to \infty.$$

$${q^{(k)}}^{\mathrm{H}} = \frac{\alpha_1 \lambda_1^k {(x_1 + y^{(k)})}^{\mathrm{H}}}{|\alpha_1 \lambda_1^k| \left\| x_1 + y^{(k)} \right\|}, \quad A q^{(k)} = \frac{\alpha_1 \lambda_1^k (\lambda_1 x_1 + A y^{(k)})}{|\alpha_1 \lambda_1^k| \left\| x_1 + y^{(k)} \right\|},$$

$${q^{(k)}}^{\mathrm{H}} A q^{(k)} = \frac{(\alpha_1 \lambda_1^k)^2 (\lambda_1 + x_1^{\mathrm{H}} A y^{(k)} + \lambda_1 {y^{(k)}}^{\mathrm{H}} x_1 + {y^{(k)}}^{\mathrm{H}} A y^{(k)})}{\left| \alpha_1 \lambda_1^k \right|^2 \left\| x_1 + y^{(k)} \right\|^2} \to \lambda_1,$$

kde jsme využili toho, že $x_1^{\mathrm{H}} x_1 = 1$.

Zkušební otázka 6.1! Dokažte konvergenci mocninné metody pro výpočet dominantního vlastního čísla matice A.

Dále ukážeme, že $q^{(k)} \to x_1$. K tomu uvažujme

$$\begin{aligned} A q^{(k)} - \lambda_1 q^{(k)} &= \frac{\alpha_1 \lambda_1^k (\lambda_1 x_1 + A y^{(k)})}{|\alpha_1 \lambda_1^k| \left\| x_1 + y^{(k)} \right\|} - \frac{\alpha_1 \lambda_1^k (\lambda_1 x_1 + \lambda_1 y^{(k)})}{|\alpha_1 \lambda_1^k| \left\| x_1 + y^{(k)} \right\|} \\ &= \frac{\alpha_1 \lambda_1^k (A y^{(k)} - \lambda_1 y^{(k)})}{|\alpha_1 \lambda_1^k| \left\| x_1 + y^{(k)} \right\|} \to 0 \end{aligned}$$

v důsledku toho, že $y^{(k)} \to 0$ pro $k \to \infty$.

7

NUMERICKÁ INTEGRACE OBYČEJNÝCH DIFERENCIÁLNÍCH ROVNIC

7.1 Formulace problému

10. přednáška

Dáno $f: [a,b] \times I\!R \to I\!R$, $f = f(x,y)$, $x \in [a,b]$, $y \in I\!R$. Dána tzv. počáteční podmínka $\eta \in I\!R$. Hledáme zobrazení $y: [a,b] \to I\!R$ splňující

$$\begin{aligned} y'(x) &= f(x, y(x)), \quad x \in [a,b], \\ y(a) &= \eta. \end{aligned}$$

Vyšetřování:

- (lokální) existence a jednoznačnost - matematická analýza: o funkci f předpokládáme, že je spojitá a dále předpokládáme, že f je (lokálně) lipschitzovská v druhé proměnné
- nalezení řešení
  - $*$ analyticky
  - $*$ numericky

7.2 Jednokrokové metody

Uvažujme dělení intervalu $[a,b]$ s uzly $x_i = a + ih,\ i = 0, \ldots, n$ (s konstantním krokem, obecně lze uvažovat nekonstantní).

Hodnotu řešení $y(x_i)$ aproximujeme pomocí hodnoty y_i:

$$y(x_i) \approx y_i, \qquad i = 0, \ldots, n.$$

V uzlu x_i platí

$$y'(x_i) = f(x_i, y(x_i)).$$

Předpokládejme, že funkce y je dostatečně hladká. Z Taylorova rozvoje dostaneme

$$y'(x_i) = \frac{y(x_{i+1}) - y(x_i)}{h} + O(h).$$

Dosadíme-li tento vztah do diferenciální rovnice, dostaneme

$$\frac{y(x_{i+1}) - y(x_i)}{h} + O(h) = f(x_i, y(x_i)).$$

To nás vede k myšlence, zanedbat chybu řádu $O(h)$ a počítat přibližné hodnoty $y_i, i = 1, \ldots, n$ ze vztahu

$$\frac{y_{i+1} - y_i}{h} = f(x_i, y_i)$$

tak, že přibližnou hodnotu funkce y v uzlu x_{i+1} vyjádříme pomocí přibližné hodnoty v uzlu x_i:

$$y_{i+1} = y_i + h\ f(x_i, y_i), \qquad y_0 = \eta \text{ (dáno)}.$$

Tato metoda se nazývá Eulerova metoda pro řešení úlohy $y' = f(x, y)$ s danou počáteční podmínkou.

Zkušební otázka 7.1 Odvoďte Eulerovu metodu pro pro řešení úlohy $y' = f(x, y)$.

Obecně uvažujme metody typu:

$$y_{i+1} = y_i + h\ \underbrace{\phi(x_i, y_i, h).}_{\text{přírustkové zobrazení}} \tag{7.2.1}$$

Tato metoda se nazývá jednokroková, protože hodnotu aproximace y_{i+1} počítáme pomocí hodnoty y_i. Pro Eulerovu metodu máme

$$\phi(x_i, y_i, h) := f(x_i, y_i).$$

Při použití Eulerovy metody dále platí pro hodnoty přesného řešení

$$\underbrace{\frac{y(x+h) - y(x)}{h}}_{\text{přesný relativní přírustek}} = y'(x) + O(h) = f(x, y(x)) + O(h) = \phi(x, y(x), h) + O(h).$$

V Eulerově metodě se tedy liší přesný relativní přírustek a přírustkové zobrazeni o veličinu řádu $O(h)$:

$$\frac{y(x+h) - y(x)}{h} = \phi(x, y(x), h) + O(h).$$

To nás vede k definici řádu metody:

Definice 7.1 *Řekneme, že metoda* (7.2.1) *je řádu p, jestliže*

$$\frac{y(x+h) - y(x)}{h} = \phi(x, y(x), h) + O(h^p). \tag{7.2.2}$$

Jinými slovy definice říká, že obecná jednokroková metoda (7.2.1) je řádu p, jestliže přesné řešení splňuje vztah (7.2.1) s chybou $hO(h^p)$.

Definice 7.2 *Řekneme, že obecná jednokroková metoda je konvergentní, jestliže*

$$\forall i = 0, \ldots, n, \qquad |y(x_i) - y_i| \le \varphi(h),$$

kde $\varphi(h)$ je infinitesimální vzhledem k h. V takovém případě řekneme, že metoda je konvergentní s řádem p, jestliže $\varphi(h) = O(h^p)$.

Věta 7.3 *Metoda (7.2.1) je konvergentní, právě když $f(x,y) = \phi(x,y,0)$, za předpokladu spojitosti f, ϕ a* **lipschitzovskosti** *f a ϕ v druhé proměnné.*

Věta 7.4 *(Odhad chyby) Je-li metoda řádu p, potom ∃ konstanta $C \geq 0$ taková, že*

$$|y(x_i) - y_i| \leq C \cdot h^p \cdot \frac{e^{L(x_i - x_0)} - 1}{L},$$

za předpokladu spojitosti f, ϕ a **lipschitzovskosti** *f a ϕ v druhé proměnné. Zde L je konstanta lipschitzovskosti přírustkového zobrazení ϕ.*

Poznámka 7.5 Obdobně se odvodí jednokrokové metody pro řešení soustav obyčejných diferenciálních rovnic

$$\begin{aligned} y'(x) &= f(x, y(x)), \quad x \in [a,b], \\ y(a) &= \eta, \end{aligned}$$

$f : [a,b] \times I\!R^m \to I\!R^m, \eta \in I\!R^m$.

7.2.1 *Metody typu Runge–Kutta*

Podobně jako při konstrukci Eulerovy metody, která je metodou prvního řádu, můžeme postupovat při odvození metody vyššího řádu. Z Taylorova rozvoje funkce y dostaneme

$$\begin{aligned} y(x_{i+1}) &= y(x_i) + h\, y'(x_i) + \tfrac{1}{2}h^2\, y''(x_i) + O(h^3), \\ y(x_{i+1}) &= y(x_i) + h\, f(x_i, y(x_i)) + \tfrac{1}{2}h^2\, \tfrac{df}{dx}(x_i, y(x_i)) + O(h^3). \end{aligned}$$

Podle věty o derivaci složené funkce máme

$$\frac{df}{dx} = f_x + f_y\, f.$$

Můžeme tak zkonstruovat metodu druhého řádu s přírustkovým zobrazením

$$\phi(x,y,h) = f(x,y) + \frac{1}{2}h\left(f_x(x,y) + f_y(x,y) f(x,y)\right),$$

které ale závisí na derivacích zobrazení f. Proto se používají metody typu Runge-Kutta: konstruuje se ϕ, splňující (7.2.2), bez použití derivací f. Základní myšlenka spočívá v tom, že přírustkové zobrazení se hledá ve speciálním tvaru tak, aby se lišilo od přesného relativního přírustku o veličinu $O(h^p)$. Tvar, ve kterém se hledá přírustkové zobrazení, je následující:

$$\phi(x,y,h) = \sum_{i=1}^{s} \omega_i k_i = \omega_1 k_1 + \omega_2 k_2 + \cdots + \omega_s k_s,$$

kde ω_i jsou konstanty. Veličiny k_i jsou vyjádřeny pomocí hodnot zobrazení f bez použití jeho derivací.

$$k_1 = f(x,y),$$

$$
\begin{aligned}
k_2 &= f(x+\alpha_2 h, y+\beta_{21} h k_1),\\
&\vdots\\
k_i &= f(x+\alpha_i h, y+h\sum_{j=1}^{i-1}\beta_{ij}k_j),\\
&\vdots\\
k_s &= f(x+\alpha_s h, y+h\sum_{j=1}^{s-1}\beta_{sj}k_j),
\end{aligned}
$$

kde α_i, β_{ij} jsou konstanty. Ve výše uvedených vzorcích je obecně

$$s \neq p.$$

Pro požadovaný řád metody $p \leq 4$ lze volit $s := p$. Pro $p > 4$ musí být $s > p$.

7.2.1.1 *Rungeova–Kuttova metoda 2. řádu* Ukážeme, jak se určí konstanty $\omega_i, \alpha_i, \beta_{ij}$ na příkladu odvození Rungeovy-Kuttovy metody 2. řádu, tj. pro $p = 2$. Přírustkové zobrazení hledáme ve tvaru

$$\phi(x, y, h) = \omega_1 f(x, y) + \omega_2 f(x + \alpha h, y + \beta h f(x, y)).$$

Cílem je určit konstanty $\omega_1, \omega_2, \alpha, \beta$ tak, aby metoda byla 2. řádu, tj. aby

$$\frac{y(x+h) - y(x)}{h} = \phi(x, y, h) + O(h^2).$$

Myšlenka je založena na vyjádření přesného relativního přírustku pomocí Taylorova rozvoje ve tvaru

$$\frac{y(x+h) + y(x)}{h} = \text{výraz } 1 + O(h^2)$$

a vyjádření přírustkového zobrazení ϕ ve tvaru

$$\phi(x, y, h) = \text{výraz } 2 + O(h^2).$$

Konstanty $\omega_1, \omega_2, \alpha, \beta$ ve 'výraz 2' nastavíme tak, aby 'výraz 1 = 'výraz 2'. Z Taylorova rozvoje funkce y dostaneme

$$y(x+h) = y(x) + hf + \frac{1}{2}h^2\Big(\underbrace{f_x + f_y f}_{y''(x)=\frac{d}{dx}f(x,y(x))=f_x+f_y f}\Big) + O(h^3),$$

odkud

$$\frac{y(x+h) - y(x)}{h} = \underbrace{f + \frac{1}{2}hf_x + \frac{1}{2}hf_y f}_{\text{výraz } 1} + O(h^2).$$

Na základě Taylorova rozvoje funkce dvou proměnných f

$$f(x+h_1, y+h_2) = f(x,y) + h_1 \frac{\partial f}{\partial x}(x,y) + h_2 \frac{\partial f}{\partial y}(x,y) + O(|h|^2)$$

pro $h_1 = \alpha h, h_2 = h\beta f$ a definice přírustkového zobrazení v metodě typu Runge–Kutta

$$\phi(x,y,h) = \omega_1 f(x,y) + \omega_2 f(x+\alpha h, y + \beta h f(x,y))$$

vyjádříme přírustkové zobrazení ve tvaru

$$\phi(x,y,h) = \underbrace{\omega_1 f + \omega_2[f + \alpha h f_x + \beta h f f_y]}_{\text{výraz 2}} + O(h^2).$$

'výraz 2' ještě upravíme, abychom ho mohli porovnat s 'výrazem 1':

$$\phi(x,y,h) = \underbrace{(\omega_1+\omega_2) f + \omega_2 \alpha h f_x + \omega_2 \beta h f f_y}_{\text{výraz 2}} + O(h^2).$$

Porovnáním koeficientů u f, hf_x a $hf_y f$ ve 'výraz 1' a 'výraz 2' získáme rovnice pro $\omega_1, \omega_2, \alpha$ a β

$$1 = \omega_1 + \omega_2, \qquad \frac{1}{2} = \alpha\omega_2, \qquad \frac{1}{2} = \beta\omega_2.$$

Odvodili jsme tak 3 rovnice pro 4 neznámé. Zvolíme např. $\omega_1 = 0$ a určíme zbývající konstanty:

$$\omega_2 = 1,$$
$$\alpha = \frac{1}{2}, \;\; \beta = \frac{1}{2}.$$

Runge–Kuttovu metodu 2. řádu lze tedy zapsat ve tvaru

$$y_{i+1} = y_i + h\phi(x_i, y_i, h),$$

$$\phi(x_i, y_i, h) = f\Big(x_i + \frac{1}{2}h,\; y_i + \frac{1}{2}h f(x_i, y_i)\Big).$$

Zkušební otázka 7.2! Odvoďte Rungeovu–Kuttovu metodu 2. řádu.

8

GRADIENTNÍ METODY

8.1 Formulace problému

11. přednáška

↓↓↓↓↓

Pokračování příště.

BIBLIOGRAFIE

Feistauer, M., Felcman, J., and Straškraba, I. (2003). *Mathematical and Computational Methods for Compressible Flow.* Oxford University Press, Oxford.

Higham, N. (1989). The accuracy of solutions to triangular systems. *SIAM J. Appl. Math.*, **26(5)**, 1252–1265.

Quarteroni, A., Sacco, R., and Saleri, F. (2004). *Numerical Mathematics* (2nd edn), Volume 37 of *Texts in Applied Mathematics.* Springer, Berlin. ISBN 0-387-98959-5.

Segethová, J. (2000). *Základy numerické matematiky.* Karolinum, Praha.

Ueberhuber, W. (2000). *Numerical Computation 1, 2: Methods, Software, and Analysis.* Springer, Berlin.

INDEX