



NUMERICKÉ METODY

Radek Kučera

Vytvořeno v rámci projektu Operačního programu Rozvoje lidských zdrojů
CZ.04.1.03/3.2.15.1/0016

Studijní opory s převažujícími distančními prvky pro předměty teoretického
základu studia.

Tento projekt je spolufinancován Evropským sociálním fondem
a státním rozpočtem České republiky

ISBN 80-248-1198-7

STUDIJNÍ OPORY S PŘEVAŽUJÍCÍMI DISTANČNÍMI PRVKY PRO PŘEDMĚTY TEORETICKÉHO ZÁKLADU STUDIA

je název projektu, který uspěl v rámci první výzvy Operačního programu Rozvoj lidských zdrojů. Projekt je spolufinancován státním rozpočtem ČR a Evropským sociálním fondem. Partnery projektu jsou Regionální středisko výchovy a vzdělávání, s.r.o. v Mostě, Univerzita obrany v Brně a Technická univerzita v Liberci. Projekt byl zahájen 5.1.2006 a bude ukončen 4.1.2008.

Cílem projektu je zpracování studijních materiálů z matematiky, deskriptivní geometrie, fyziky a chemie tak, aby umožnily především samostatné studium a tím minimalizovaly počet kontaktních hodin s učitelem. Je zřejmé, že vytvořené texty jsou určeny studentům všech forem studia. Studenti kombinované a distanční formy studia je využijí k samostudiu, studenti v prezenční formě si mohou doplnit získané vědomosti. Všem studentům texty pomohou při procvičení a ověření získaných vědomostí. Nezanedbatelným cílem projektu je umožnit zvýšení kvalifikace širokému spektru osob, které nemohly ve studiu na vysoké škole z různých důvodů (sociálních, rodinných, politických) pokračovat bezprostředně po maturitě.

V rámci projektu jsou vytvořeny jednak standardní učební texty v tištěné podobě, koncipované pro samostatné studium, jednak e-learningové studijní materiály, přístupné prostřednictvím internetu. Součástí výstupů je rovněž banka testových úloh pro jednotlivé předměty, na níž si studenti ověří, do jaké míry zvládli prostudované učivo.

Bližší informace o projektu můžete najít na adrese <http://www.studopory.vsb.cz/>.

Přejeme vám mnoho úspěchů při studiu a budeme mít radost, pokud vám předložený text pomůže při studiu a bude se vám líbit. Protože nikdo není neomylný, mohou se i v tomto textu objevit nejasnosti a chyby. Předem se za ně omlouváme a budeme vám vděční, pokud nás na ně upozorníte.

ESF – ROVNÉ PŘÍLEŽITOSTI PRO VŠECHNY

POKYNY KE STUDIU

V úvodu si vysvětlíme jednotnou pevnou strukturu každé kapitoly textu, která by vám měla pomoci k rychlejší orientaci při studiu. Pro zvýraznění jednotlivých částí textu jsou používány ikony a barevné odlišení, jejichž význam nyní objasníme.



Průvodce studiem

vás stručně seznámí s obsahem dané kapitoly a s její motivací. Slouží také k instrukci, jak pokračovat dál po vyřešení kontrolních otázek nebo kontrolních textů.



Cíle

vás seznámí s učivem, které v dané kapitole poznáte a které byste po jejím prostudování měli umět.



Předpokládané znalosti

shrnují stručně učivo, které byste měli znát ještě dříve než kapitolu začnete studovat. Jsou nezbytným předpokladem pro úspěšné zvládnutí následující kapitoly.



Výklad

označuje samotný výklad učiva dané kapitoly, který je členěn způsobem obvyklým v matematice na definice, věty, případně důkazy.

Definice 1.1.1.

Zavádí základní pojmy v dané kapitole.

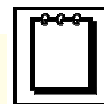
Věta 1.1.1.

Uvádí základní vlastnosti pojmů zavedených v dané kapitole.

Důkaz: Vychází z předpokladů věty a dokazuje tvrzení uvedené ve větě.

Poznámka

neformálně komentuje vykládanou látku..

**Řešené úlohy**

označují vzorové příklady, které ilustrují probrané učivo.

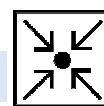


Příklad Uvádí zadání příkladu.

Řešení: Uvádí podrobné řešení zadaného příkladu.

Úlohy k samostatnému řešení

obsahují zadání příkladů k procvičení probraného učiva. Úlohy označené **✖** patří k obtížnějším a jsou určeny zájemcům o hlubší pochopení tématu.

**Výsledky úloh k samostatnému řešení**

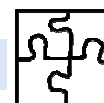
obsahují správné výsledky předchozích příkladů, slouží ke kontrole správnosti řešení.

**Kontrolní otázky**

obsahují soubor otázek k probranému učivu včetně několika odpovědí, z nichž je vždy alespoň jedna správná.

**Odpovědi na kontrolní otázky**

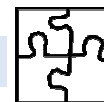
uvádějí správné odpovědi na kontrolní otázky.

**Kontrolní test**

obsahuje soubor příkladů k probranému učivu.

**Výsledky testu**

uvádějí správné odpovědi na příklady kontrolního testu.



**Shrnutí lekce**

obsahuje stručný přehled učiva, které by měl student po prostudování příslušné kapitoly zvládnout.

**Literatura**

obsahuje seznam knih, které byly použity při tvorbě příslušného textu a na které byly případně uvedeny odkazy k hlubšímu prostudování tématu.



Piktogram, který upozorňuje na důležité vztahy nebo vlastnosti, které je nezbytné si zapamatovat.

Obsah

1. Numerické metody a chyby	7
1.1. Obsah předmětu	7
1.2. Chyby v numerických výpočtech	13
2. Řešení nelineárních rovnic	21
2.1. Separace kořenů	21
2.2. Nejjednodušší metody	24
2.2.1. Metoda půlení intervalu	25
2.2.2. Metoda regula falsi	26
2.3. Newtonova metoda	29
2.4. Metoda prosté iterace	34
3. SLR – přímé metody	40
3.1. Formulace úlohy	40
3.2. Gaussova eliminační metoda (GEM)	43
3.2.1. Zpětný chod GEM	44
3.2.2. Dopředný chod GEM	45
3.2.3. Výběr hlavního prvku	46
3.3. LU-rozklad	48
3.4. Použití LU-rozkladu	54
3.4.1. Řešení soustav lineárních rovnic	54
3.4.2. Výpočet inverzní matice	56
3.4.3. Výpočet determinantu	58
3.5. Maticové normy a podmíněnost matic	61
4. SLR – iterační metody	66
4.1. Příklad iteračního výpočtu	67
4.2. Obecné iterační metody	71

4.2.1. Jacobiova metoda	72
4.2.2. Gauss-Seidelova metoda	74
4.3. Vlastní čísla a vlastní vektory matic	78
4.3.1. Výpočet vlastních čísel metodou LU-rozkladu	82
4.4. Konvergence iteračních metod	86
5. Interpolace a aproximace funkcí	92
5.1. Interpolační polynom	93
5.1.1. Lagrangeův tvar interpolačního polynomu	95
5.1.2. Newtonův tvar interpolačního polynomu	97
5.1.3. Interpolační chyba	99
5.2. Interpolační splajny	102
5.2.1. Lineární splajn	102
5.2.2. Kubický splajn	103
5.3. Aproximace metodou nejmenších čtverců	108
6. Numerické integrování a derivování	113
6.1. Newton-Cotesovy vzorce	114
6.2. Složené vzorce	118
6.3. Výpočet integrálu se zadanou přesností	122
6.4. Numerické derivování	127
7. ODR – počáteční úlohy	132
7.1. Formulace úlohy	132
7.2. Eulerova metoda	136
7.3. Jednokrokové metody vyššího řádu	139
7.4. Vícekrokové metody	143
Literatura	152

1. Numerické metody a chyby

1.1. Obsah předmětu

Průvodce studiem



Numerickou úlohou rozumíme jasný a jednoznačný popis vztahu mezi **konečným** počtem vstupních a výstupních dat (reálných čísel). Podstatná je přitom konečnost vstupního a výstupního souboru, která ve svém důsledku umožňuje při řešení použít počítač. Postupy řešení numerických úloh se pak nazývají *numerické* nebo *počítačové metody*.

Numerické úlohy patří do skupiny úloh *diskrétních*. Matematické modely se však často formulují jako *úlohy spojité*, u nichž se mezi vstupními nebo výstupními daty vyskytují spojité funkce. Pokud chceme takové úlohy řešit numerickými metodami, musíme je nejdříve na úlohy diskrétní převést, tj. *diskretizovat*.

Cíle



Na příkladech ukážeme diskrétní a spojité úlohy. Dále předvedeme diskretizaci a vysvětlíme pojmy diskretizační parametr a řád.

Předpokládané znalosti



Kvadratická rovnice, soustava lineárních rovnic, určitý integrál, počáteční úloha pro obyčejnou diferenciální rovnici.

Výklad



1) Úloha řešit kvadratickou rovnici $ax^2 + bx + c = 0$, $a \neq 0$, je úloha diskrétní. Vstupní data jsou koeficienty a, b, c , výstupní data jsou reálná čísla $\alpha_1, \beta_1, \alpha_2, \beta_2$, která určují dva komplexní kořeny $x_k = \alpha_k + i\beta_k$, $k = 1, 2$.

2) Diskrétní úlohou je také soustava lineárních rovnic

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

kde $\mathbf{A} = (a_{ij})$ je daná čtvercová matice řádu n , $\mathbf{b} = (b_i)$ je daný sloupcový vektor o n složkách a $\mathbf{x} = (x_i)$ je sloupcový vektor neznámých také o n složkách. Například pro $n = 3$ můžeme takovou soustavu psát v maticovém tvaru

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

nebo po jednotlivých rovnicích

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1,$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2,$$

$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3.$$

Vstupními daty jsou zde prvky matice a_{ij} a vektoru pravých stran b_i . Výstupními daty jsou složky x_i vektoru neznámých. Připomeňme ještě, že řešení může být jediné, nemusí existovat, nebo jich může být nekonečně mnoho.

3) Úloha vypočítat určitý integrál

$$I = \int_a^b f(x) dx$$

je spojitá úloha, protože jedním ze vstupů je spojitá funkce f . Na tomto příkladě si předvedeme diskretizaci.

Integrační interval $\langle a, b \rangle$ rozdělíme na n úseků o délce h pomocí bodů x_i , $i = 0, 1, \dots, n$, tak, že $x_i - x_{i-1} = h$, $x_0 = a$ a $x_n = b$. Pak můžeme psát

$$I = \int_{x_0}^{x_1} f(x) dx + \int_{x_1}^{x_2} f(x) dx + \dots + \int_{x_{n-1}}^{x_n} f(x) dx.$$

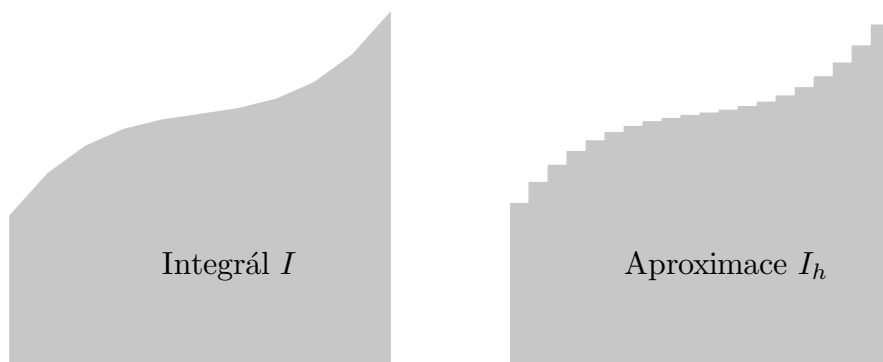
Každý dílčí integrál nahradíme jeho přibližnou hodnotou

$$\int_{x_{i-1}}^{x_i} f(x) dx \approx hf\left(\frac{x_{i-1} + x_i}{2}\right)$$

a místo hodnoty I budeme počítat její aproximaci

$$I_h = hf\left(\frac{x_0 + x_1}{2}\right) + hf\left(\frac{x_1 + x_2}{2}\right) + \dots + hf\left(\frac{x_{n-1} + x_n}{2}\right). \quad (1.1.1)$$

Výpočet podle posledního vzorce je již úloha diskrétní. Vstupními daty jsou funkční hodnoty $f\left(\frac{x_{i-1} + x_i}{2}\right)$, $i = 1, \dots, n$ a parametr h . Výstupní data představuje přibližná hodnota I_h .



Obrázek 1.1.1: Znázornění integrálu I a jeho aproximace I_h .

Smysl vzorce (1.1.1) ukazuje obrázek 1.1.1, kde jsou hodnoty I a I_h znázorněny jako velikosti plochy příslušného obrazce. Odtud můžeme usoudit, že při menším h bude I_h lépe aproximovat I , tj., že platí

$$\lim_{h \rightarrow 0^+} I_h = I. \quad (1.1.2)$$

Jinými slovy řešení diskretizované úlohy se může přiblížit libovolně přesně k řešení původní úlohy spojité, pokud zvolíme dostatečně malý *diskretizační parametr* h .

Kladné číslo p , pro něž platí

$$|I - I_h| \leq Ch^p, \quad (1.1.3)$$

kde $C > 0$ je konstanta nezávislá na h , se nazývá *řád diskretizace*. Výraz na levé straně nerovnosti (1.1.3) je velikost *diskretizační chyby*. Tato chyba bude při zmenšujícím se h klesat k nule tím rychleji, čím větší bude hodnota p . Diskretizace vysokého řádu je proto přesnější než diskretizace nízkého řádu; viz tabulka 1.1.1.

Tabulka 1.1.1: Odhady diskretizační chyby Ch^p pro $C = 1$.

h	$p = 1$	$p = 2$	$p = 3$
0.1	0.1	0.01	0.001
0.01	0.01	0.0001	0.000001
0.001	0.001	0.000001	0.000000001

Příklad 1.1.1. Pomocí vzorce (1.1.1) vypočítejte přibližnou hodnotu integrálu

$$I = \int_0^1 x^2 dx$$

pro $h = 0.5, 0.25$ a 0.125 . Z výsledků odhadněte, jaký je řád diskretizace.

Řešení: Přesná hodnota integrálu je $I = \frac{1}{3}$. Přibližné hodnoty vypočítáme takto:

$$I_{0.5} = 0.5(0.25^2 + 0.75^2) = 0.3125,$$

$$I_{0.25} = 0.25(0.125^2 + 0.375^2 + 0.625^2 + 0.875^2) = 0.328125,$$

$$I_{0.125} = 0.125(0.0625^2 + 0.1875^2 + \dots + 0.9375^2) = 0.33203125.$$

Diskretizační chyby mají hodnotu:

$$E_{0.5} = |I - I_{0.5}| = 0.0208333333333333,$$

$$E_{0.25} = |I - I_{0.25}| = 0.0052083333333333,$$

$$E_{0.125} = |I - I_{0.125}| = 0.0013020833333333.$$

Při odhadu řádu diskretizace budeme pro jednoduchost předpokládat, že v (1.1.3) nastane rovnost. Pro $h = 0.5$ pak dostáváme

$$\frac{E_{0.5}}{E_{0.25}} = \frac{Ch^p}{C(h/2)^p} = 2^p \implies p = \log_2 \frac{E_{0.5}}{E_{0.25}} \doteq 2.000000000000069.$$

Podobně pro $h = 0.25$ vypočteme $p \doteq 2.000000000000277$. Z těchto výsledků můžeme usoudit, že diskretizace podle vzorce (1.1.1) je druhého řádu. \square

4) Úloha najít funkci $y = y(x)$, která splňuje diferenciální rovnici

$$y' = x^2 - 0.2y \quad (1.1.4)$$

a vyhovuje počáteční podmínce $y(-2) = -1$, je spojitá úloha. Jak uvidíme později, diskretizace této úlohy bude v mnohém podobná postupu, který jsme použili při diskretizaci určitého integrálu.

Kontrolní otázky



Otázka 1. Jaký je rozdíl mezi diskrétní a spojitou úlohou?

Otázka 2. Co je to diskretizace? Jaký je význam diskretizačního parametru?

Otázka 3. Je přesnější diskretizace vysokého nebo nízkého řádu?

Úlohy k samostatnému řešení



1. Vyřešte rovnice: a) $x^2 + 3x + 1 = 0$; b) $x^2 + 2x + 1 = 0$; c) $x^2 + x + 1 = 0$.

2. Řešte následující soustavy lineárních rovnic:

$$a) \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 5 \end{pmatrix}, \quad b) \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ -3 \end{pmatrix},$$

$$c) \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Jak lze rozhodnout z hodnoty determinantu matice o existenci řešení?

3. Pomocí vzorce (1.1.1) vypočtete přibližnou hodnotu integrálu

$$I = \int_{-1}^1 x^2 dx$$

pro $h = 1, 0.5$ a 0.25 a určete diskretizační chyby.

4. Vyřešte diferenciální rovnici (1.1.4) pomocí známých analytických metod.

**Výsledky úloh k samostatnému řešení**

1. a) Dva kořeny $x_1 \doteq -2.6180$, $x_2 \doteq -0.3820$; b) jeden (dvojnásobný) kořen $x_1 = x_2 = -1$; c) dva komplexně sdružené kořeny $x_1 = x_2 \doteq -0.5 \pm i 0.8660$.
2. a) $\mathbf{x} = (3, 1)^\top$, $\det \mathbf{A} = -7$; b) nekonečně mnoho řešení $\mathbf{x} = (3 - 2t, 3 - 2t)^\top$, $t \in \mathbb{R}$, $\det \mathbf{A} = 0$; c) řešení neexistuje.
3. $I \doteq 0.6667$, $I_1 = 0.5$, $|I - I_1| \doteq 0.1667$, $I_{0.5} = 0.625$, $|I - I_{0.5}| \doteq 0.0417$, $I_{0.25} \doteq 0.6563$, $|I - I_{0.25}| \doteq 0.0104$.
4. Obecné řešení je $y(x) = 5x^2 - 50x + 250 + Ce^{-0.2x}$, řešení vyhovující počáteční podmínce je určeno konstantou $C = -248.688737$.

**Shrnutí lekce**

Ukázali jsem rozdělení matematických úloh na úlohy diskrétní a spojitě. Diskrétní úlohy lze zpravidla okamžitě řešit pomocí numerických metod. Spojité úlohy je potřeba nejdříve diskretizovat.

1.2. Chyby v numerických výpočtech

Průvodce studiem



Chyby, kterými jsou ovlivněny výsledky numerických výpočtů, mají různou podstatu. *Chyba matematického modelu* vzniká v důsledku toho, že místo skutečného technického nebo fyzikálního problému řešíme jeho matematický model. Je-li řešení tohoto modelu z nějakého důvodu náročné nebo nemožné, provedeme jeho aproximaci jednodušší úlohou, čímž vznikne *chyba aproximační*. Jejím speciálním případem je *diskretizační chyba*, kterou jsme zmínili v předchozím odstavci.

Dalším zdrojem chyb je počítání s „nepřesnými“ čísly. Sem patří *chyby vstupních dat* a *chyby zaokrouhlovací*. Vstupní data mohou být naměřené veličiny, jejichž nepřesnost je dána rozlišovací schopností měřících zařízení. K zaokrouhlování mezivýsledků dochází v průběhu celého výpočtu, protože pro ukládání čísel je k dispozici pouze omezený paměťový prostor.

Konečně jsou výsledky ovlivněny také *chybami lidského faktoru*. Jedná se o chyby v počítačových programech, špatná zadání vstupních dat, nevhodnou volbu matematického modelu nebo nesprávný výběr metody řešení.

Cíle



Budeme se zabývat chybami zaokrouhlovacími a ukážeme jejich vliv na stabilitu numerického výpočtu.

Předpokládané znalosti



Základní aritmetické operace, určitý integrál, rekurentní výpočty.



Výklad

Definice 1.2.1.

Nechť \bar{x} je přesná hodnota reálného čísla a x je jeho aproximace. Rozdíl

$$e(x) = \bar{x} - x$$

se nazývá absolutní chyba. Odhad absolutní chyby je číslo $\epsilon(x)$, pro které platí

$$|\bar{x} - x| \leq \epsilon(x). \quad (1.2.1)$$

Je-li $x \neq 0$, pak číslo

$$r(x) = \frac{\bar{x} - x}{x}$$

se nazývá relativní chyba. Odhad relativní chyby je číslo $\delta(x)$, pro které platí

$$\left| \frac{\bar{x} - x}{x} \right| \leq \delta(x).$$

Relativní chyba a její odhad se často udávají v procentech. Nerovnost (1.2.1) znamená $\bar{x} \in \langle x - \epsilon(x), x + \epsilon(x) \rangle$, což symbolicky zapisujeme $\bar{x} = x \pm \epsilon(x)$. Pokud nebude hrozit nedorozumnění, budeme psát e , r , ϵ a δ místo $e(x)$, $r(x)$, $\epsilon(x)$ a $\delta(x)$.

Příklad 1.2.1. Číslo $x = 2.72$ je aproximace Eulerova čísla $\bar{x} = 2.7182818\dots$. Absolutní chyba je $e = -0.001718\dots$ a její odhad je například číslo $\epsilon = 0.002$, protože $|e| \leq \epsilon$. Proto $\bar{x} = 2.72 \pm 0.002$. Relativní chyba je $r = -0.00063168\dots$ a za odhad relativní chyby můžeme vzít $\delta = 0.00064$, protože $|r| \leq \delta$.

Nyní ukážeme jak se šíří chyby při provádění aritmetických operací. Budeme přitom předpokládat, že vykonáváme přesné aritmetické operace s nepřesnými čísly, tj. s aproximacemi, a že známe chyby, respektive jejich odhady.

Nechť

$$\bar{x}_i = x_i + e(x_i), \quad |e(x_i)| \leq \epsilon(x_i), \quad |r(x_i)| \leq \delta(x_i), \quad i = 1, 2.$$

(a) Je-li $u = x_1 + x_2$ *aproximace součtu* $\bar{u} = \bar{x}_1 + \bar{x}_2$, potom

$$\bar{u} = x_1 + e(x_1) + x_2 + e(x_2) = u + e(u),$$

kde

$$e(u) = e(x_1) + e(x_2)$$

a platí

$$|e(u)| \leq |e(x_1)| + |e(x_2)| \leq \epsilon(x_1) + \epsilon(x_2).$$

(b) Je-li $v = x_1 - x_2$ *aproximace rozdílu* $\bar{v} = \bar{x}_1 - \bar{x}_2$, potom

$$e(v) = e(x_1) - e(x_2)$$

a platí

$$|e(v)| \leq |e(x_1)| + |e(x_2)| \leq \epsilon(x_1) + \epsilon(x_2).$$

(c) Je-li $w = x_1 x_2$ *aproximace součinu* $\bar{w} = \bar{x}_1 \bar{x}_2$, potom

$$\bar{w} = (x_1 + e(x_1))(x_2 + e(x_2)) = w + x_1 e(x_2) + x_2 e(x_1) + e(x_1) e(x_2)$$

a klademe

$$e(w) \approx x_1 e(x_2) + x_2 e(x_1).$$

Odtud

$$|e(w)| \lesssim |x_1| \epsilon(x_2) + |x_2| \epsilon(x_1).$$

(d) Je-li $z = x_1/x_2$ *aproximace podílu* $\bar{z} = \bar{x}_1/\bar{x}_2$, potom

$$\bar{z} = \frac{x_1 + e(x_1)}{x_2 + e(x_2)} = z + e(z),$$

kde

$$e(z) = \frac{x_2 e(x_1) - x_1 e(x_2)}{x_2(x_2 + e(x_2))} \approx \frac{x_2 e(x_1) - x_1 e(x_2)}{x_2^2}$$

a platí

$$|e(z)| \lesssim \frac{|x_2|\epsilon(x_1) + |x_1|\epsilon(x_2)}{|x_2|^2}.$$

Pro relativní chyby můžeme z pravidel (a), (b), (c) a (d) odvodit:

$$(A) \quad r(u) = \frac{1}{x_1 + x_2} \left(x_1 \frac{e(x_1)}{x_1} + x_2 \frac{e(x_2)}{x_2} \right) = \frac{1}{x_1 + x_2} (x_1 r(x_1) + x_2 r(x_2)),$$

$$|r(u)| \leq \frac{1}{|x_1 + x_2|} (|x_1|\delta(x_1) + |x_2|\delta(x_2)),$$

$$(B) \quad r(v) = \frac{1}{x_1 - x_2} \left(x_1 \frac{e(x_1)}{x_1} - x_2 \frac{e(x_2)}{x_2} \right) = \frac{1}{x_1 - x_2} (x_1 r(x_1) - x_2 r(x_2)),$$

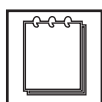
$$|r(v)| \leq \frac{1}{|x_1 - x_2|} (|x_1|\delta(x_1) + |x_2|\delta(x_2)),$$

$$(C) \quad r(w) \approx \frac{x_1 e(x_2) + x_2 e(x_1)}{x_1 x_2} = r(x_2) + r(x_1),$$

$$|r(w)| \lesssim \delta(x_2) + \delta(x_1),$$

$$(D) \quad r(z) \approx \frac{x_2 e(x_1) - x_1 e(x_2)}{x_2^2} \cdot \frac{x_2}{x_1} = r(x_1) - r(x_2),$$

$$|r(z)| \lesssim \delta(x_1) + \delta(x_2).$$



Poznámka

Při odčítání blízkých čísel (pravidlo (B)) má na velikost relativní chyby rozhodující vliv zlomek $1/|x_1 - x_2|$, který ukazuje, že dochází ke ztrátě relativní přesnosti.

Příklad 1.2.2. Nechť $\bar{x}_1 = 758\,320$, $x_1 = 758\,330$, $\bar{x}_2 = 757\,940$ a $x_2 = 757\,930$. Určete k jak velké ztrátě relativní přesnosti dojde při odčítání.

Řešení: Protože $e(x_1) = -10$, $e(x_2) = 10$, můžeme položit

$$\left| \frac{-10}{758330} \right| \doteq 1.32 \cdot 10^{-5} = \delta(x_1), \quad \left| \frac{10}{757930} \right| \doteq 1.32 \cdot 10^{-5} = \delta(x_2).$$

Dále je $\bar{v} = \bar{x}_1 - \bar{x}_2 = 380$ a $v = x_1 - x_2 = 400$, a proto

$$\left| \frac{\bar{v} - v}{v} \right| = \left| \frac{-20}{400} \right| \doteq 5 \cdot 10^{-2} = \delta(v).$$

Došlo ke ztrátě relativní přesnosti zhruba o tři řády. Podle pravidla (B) totiž platí

$$|r(v)| \leq \frac{758330}{400}(\delta(x_1) + \delta(x_2)) \approx 2000(\delta(x_1) + \delta(x_2)).$$

Výklad



Při provádění rozsáhlejších výpočtů může nastat situace, kdy se zaokrouhlovací chyby nekontrolovatelně hromadí a mohou znehodnotit výsledek. O takovém výpočtu říkáme, že je *numericky nestabilní*. Ukážeme to na příkladu.

Předpokládejme, že je naším úkolem vypočítat hodnoty integrálů

$$y_i = \int_0^1 \frac{x^i}{x+5} dx \quad \text{pro } i = 0, 1, \dots, 8. \quad (1.2.2)$$

Pomocí úpravy

$$y_i + 5y_{i-1} = \int_0^1 \frac{x^i + 5x^{i-1}}{x+5} dx = \int_0^1 x^{i-1} \frac{x+5}{x+5} dx = \int_0^1 x^{i-1} dx = \frac{1}{i}$$

odvodíme rekurentní vzorec

$$y_i = \frac{1}{i} - 5y_{i-1} \quad \text{pro } i = 1, 2, \dots, 8. \quad (1.2.3)$$

Při výpočtu budeme zaokrouhlovat na tři desetinná místa. Nejdříve určíme startovací hodnotu

$$y_0 = \int_0^1 \frac{1}{x+5} dx = [\ln|x+5|]_0^1 = 0.18232\dots \doteq 0.182$$

a potom pomocí (1.2.3) počítáme

$$\begin{aligned} y_1 &= 1 - 5y_0 = 1 - 5 \cdot 0.182 = 0.090, \\ y_2 &= \frac{1}{2} - 5y_1 = \frac{1}{2} - 5 \cdot 0.090 = 0.050, \\ y_3 &= \frac{1}{3} - 5y_2 = \frac{1}{3} - 5 \cdot 0.050 \doteq 0.083, \\ y_4 &= \frac{1}{4} - 5y_3 = \frac{1}{4} - 5 \cdot 0.083 = -0.165. \end{aligned}$$

Poslední hodnota y_4 je zjevně nesmyslná, protože všechny integrály musí být kladné. Správná hodnota je $y_4 = 0.03427\dots$. Výpočet podle vzorce (1.2.3) je tedy numericky nestabilní.

Nestability se zbavíme vhodnější organizací výpočtu. Rekurentní vzorec (1.2.3) přepíšeme pro výpočet v opačném směru, tj.

$$y_{i-1} = \frac{1}{5i} - \frac{1}{5}y_i \quad \text{pro } i = 9, 8, \dots, 1. \quad (1.2.4)$$

Startovací hodnotu y_9 určíme z přibližné rovnosti $y_{10} \doteq y_9$, odkud $y_9 \doteq \frac{1}{50} - \frac{1}{5}y_9$, a proto $y_9 \doteq 0.017$. Pomocí vzorce (1.2.4) dostaneme:

$$\begin{aligned} y_8 &= \frac{1}{45} - \frac{1}{5}y_9 = \frac{1}{45} - \frac{1}{5} \cdot 0.017 \doteq 0.019, \\ y_7 &= \frac{1}{40} - \frac{1}{5}y_8 = \frac{1}{40} - \frac{1}{5} \cdot 0.019 \doteq 0.021, \\ &\vdots \\ y_0 &= \frac{1}{5} - \frac{1}{5}y_1 = \frac{1}{5} - \frac{1}{5} \cdot 0.088 \doteq 0.182. \end{aligned}$$

Hodnota y_0 je přesná (na tři desetinná místa), takže výpočet podle vzorce (1.2.4) je numericky stabilní.

Definice 1.2.2.

Uvažujme úlohu $\bar{y} = U(\bar{x})$. Nechť x je porušená vstupní hodnota a y je odpovídající porušená hodnota výsledku. Číslem podmíněnosti úlohy U nazýváme číslo C_U , pro které platí

$$|r(y)| = C_U |r(x)|,$$

kde $r(x)$ a $r(y)$ jsou relativní chyby.

Číslo podmíněnosti vyjadřuje citlivost úlohy na poruchu ve vstupních datech. Je-li $C_U \approx 1$, říkáme, že úloha U je *dobře podmíněná*. Je-li C_U velké, říkáme, že úloha U je *špatně podmíněná*. Podle čísla podmíněnosti můžeme posoudit také citlivost úlohy na zaokrouhlovací chyby, protože je můžeme interpretovat jako

důsledek (teoretické) počáteční poruchy. Pokud umíme určit jenom odhady relativních chyb, stanovíme číslo podmíněnosti přibližně, tj.

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

Příklad 1.2.3. Určete číslo podmíněnosti úlohy U vypočítat hodnotu y_4 podle vzorců (1.2.3) při zaokrouhlování na tři desetinná místa.

Řešení: Dostáváme

$$\begin{aligned} r(y_0) &= \frac{0.18232\dots - 0.182}{0.182} \doteq 0.001758, \\ r(y_4) &= \frac{0.03427\dots + 0.165}{-0.165} \doteq -1.207, \\ C_U &= \frac{|r(y_4)|}{|r(y_0)|} \doteq 686.6. \end{aligned}$$

Všimněme si, že při výpočtu podle vzorců (1.2.3) se hodnota y_{i-1} násobí pěti, čímž dojde také k pětinasobnému zvětšení chyby. Vstupní porucha se v hodnotě y_4 promítne násobená číslem $5^4 = 625$, což je zhruba číslo podmíněnosti C_U .

Kontrolní otázky

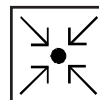


Otázka 1. Jak se definuje absolutní a relativní chyba a jejich odhady?

Otázka 2. Jak se chovají chyby při provádění aritmetických operací?

Otázka 3. Jak se definuje číslo podmíněnosti úlohy?

Úlohy k samostatnému řešení



1. Pro aproximaci $x = 3.14$ Ludolfova čísla $\bar{x} = 3.1415926\dots$ určete absolutní a relativní chybu a jejich odhady.

2. Pro data z příkladu 1.2.2. určete relativní chyby při sčítání, odčítání a dělení.

3. Určete číslo podmíněnosti úlohy vypočítat y_0 podle vzorců (1.2.4).

**Výsledky úloh k samostatnému řešení**

1. $e(x) = 0.0015926\dots$, $\epsilon(x) = 0.0016$, $r(x) = 0.000507197$, $\delta(x) = 0.00051$.
2. Pro $u = x_1 + x_2$ je $e(u) = 0$, $|e(u)| \leq 20$, $r(u) = 0$, $|r(u)| \leq 1.32 \cdot 10^{-5}$; pro $w = x_1 x_2$ je $e(w) = 3900$, $|e(w)| \leq 15162600$, $r(w) = 6.79 \cdot 10^{-9}$, $|r(w)| \leq 2.64 \cdot 10^{-5}$; pro $z = x_1/x_2$ je $e(z) = -2.64 \cdot 10^{-5}$, $|e(z)| \leq 2.64 \cdot 10^{-5}$, $r(z) = -2.64 \cdot 10^{-5}$, $|r(z)| \leq 2.64 \cdot 10^{-5}$.
3. $y_9 = 0.0169264\dots$, $r(y_9) = (y_9 - 0.017)/0.017 \doteq -0.004329$, $r(y_0) = (0.18232 - 0.1824)/0.1824 \doteq -4.01758$, $C_U = |r(y_0)|/|r(y_9)| \doteq 0.4061$.

**Shrnutí lekce**

Ukázali jsem, jak se šíří chyby při provádění aritmetických operací. Dále jsme ukázali jak posuzovat citlivost úloh na vstupní a zaokrouhlovací chyby.

2. Řešení nelineárních rovnic

Průvodce studiem



Budeme se zabývat výpočtem reálných kořenů nelineární rovnice

$$f(x) = 0, \quad (2.0.1)$$

kde f je v jistém smyslu „rozumná“ reálná funkce. Pro některé funkce (kvadratické, goniometrické atp.) umíme kořeny vypočítat pomocí (uzavřených) vzorců, pro drtivou většinu funkcí však žádné takové vzorce neexistují.

Metody, s nimiž se seznámíme v této kapitole, lze použít pro libovolnou funkci f . Patří do třídy metod *iteračních*, které počítají posloupnost $\{x^k\}$ konvergující pro $k \rightarrow \infty$ ke kořenu \bar{x} . Obecně platí, že konvergence nastane, pokud je počáteční aproximace x^0 zvolena dostatečně blízko u hledaného kořene. Jednotlivé iterační metody se pak liší rychlostí konvergence.

2.1. Separace kořenů

Cíle



Ukážeme několik možností, jak provést rozbor rovnice $f(x) = 0$, jehož výsledkem je separace kořenů v dostatečně krátkých intervalech.

Předpokládané znalosti



Spojitost funkce, grafy elementárních funkcí.

Výklad



- a) *Grafická separace 1.* Z grafu funkce f najdeme polohu průsečíků s x -ovou osou.
- b) *Grafická separace 2.* Rovnici $f(x) = 0$ převedeme na ekvivalentní rovnici $h(x) = g(x)$ a nakreslíme grafy funkcí g a h . Průsečíky těchto grafů promítneme

do x -ové osy, čímž zjistíme polohu kořenů.

c) Separace tabelací. Sestavíme tabulku funkčních hodnot funkce f a podle znaménkových změn určíme intervaly obsahující kořeny. Využíváme přitom následující větu.

Věta 2.1.1.

Nechť f je spojitá funkce na intervalu $\langle a, b \rangle$, pro niž platí

$$f(a)f(b) < 0. \quad (2.1.1)$$

Pak uvnitř intervalu (a, b) leží aspoň jeden kořen rovnice $f(x) = 0$.

Jinými slovy věta říká, že ze znaménkové změny u funkčních hodnot v krajních bodech intervalu $\langle a, b \rangle$ můžeme rozpoznat výskyt kořene uvnitř tohoto intervalu.

Příklad 2.1.1. Pro rovnici

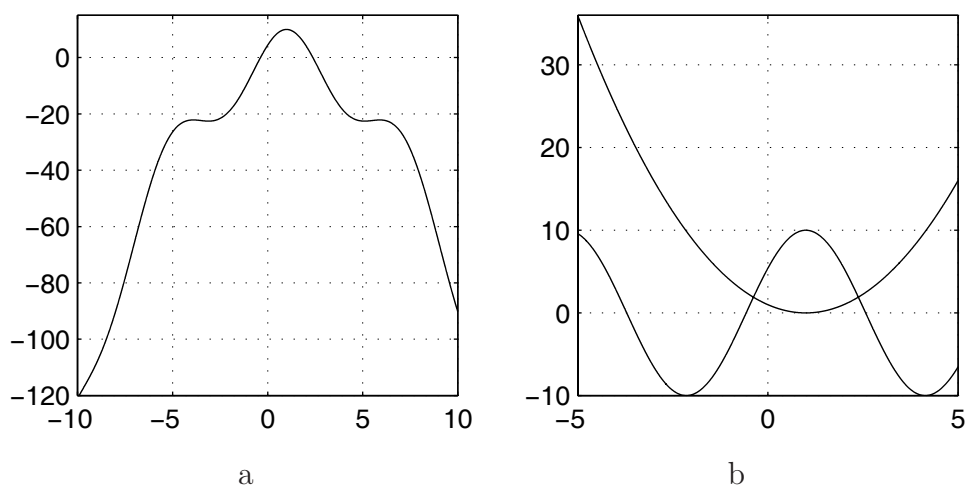
$$10 \cos(x - 1) - x^2 + 2x - 1 = 0$$

určete intervaly délky nejvýše 0.1 obsahující kořeny.

Řešení: Z grafu funkce $f(x) = 10 \cos(x - 1) - x^2 + 2x - 1$ na obrázku 2.1.1.a lze usoudit, že existují dva kořeny \bar{x}_1 a \bar{x}_2 , které leží v intervalu $\langle -5, 5 \rangle$. Zadanou rovnici přepíšeme do tvaru

$$10 \cos(x - 1) = x^2 - 2x + 1.$$

Grafy funkcí $g(x) = 10 \cos(x - 1)$ a $h(x) = x^2 - 2x + 1$ jsou znázorněny na obrázku 2.1.1.b. Odtud plyne, že $\bar{x}_1 \in \langle -1, 0 \rangle$ a $\bar{x}_2 \in \langle 2, 3 \rangle$. Další zpřesnění polohy kořenů provedeme pomocí tabelace. Z Tabulky 2.1.1 je patrné, že kořeny leží v intervalech $\bar{x}_1 \in \langle -0.4, -0.3 \rangle$ a $\bar{x}_2 \in \langle 2.3, 2.4 \rangle$.

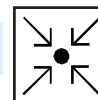
Obrázek 2.1.1: a) Graf funkce f ; b) Grafy funkcí g a h .Tabulka 2.1.1: Tabelace funkce f .

x	$f(x)$	x	$f(x)$
...
-0.5	-1.5426	2.2	2.1836
-0.4	-0.2603	2.3	0.9850
-0.3	0.9850	2.4	-0.2603
-0.2	2.1836	2.5	-1.5426
...

Kontrolní otázky

Otázka 1. Jak se provádí separace kořenů rovnic?

Otázka 2. Jaký je grafický smysl věty 2.1.1.?

Úlohy k samostatnému řešení1. Proveďte separaci kořenů rovnice $x^2 - x - \frac{6}{7} \ln x = 0$.**Výsledky úloh k samostatnému řešení**1. Dva kořeny: $\bar{x}_1 \in (0.9, 0.91)$, $\bar{x}_2 = 1$.

2.2. Nejjednodušší metody



Cíle

Seznámíme se s nejjednoduššími iteračními metodami pro výpočet kořenů rovnice $f(x) = 0$. Jsou založeny na postupném zkracování intervalu, který obsahuje kořen. Tato strategie zaručuje konvergenci pro každou spojitou funkci, výpočet je však pomalý.



Předpokládané znalosti

Určení polohy kořene pomocí znaménkových změn, věta 2.1.1. Rovnice přímky.



Výklad

Princip zkracování intervalu použijeme u dvou metod. Začneme proto nejdříve jeho obecným popisem. Budeme přitom předpokládat, že f je spojitá funkce na intervalu $\langle a^0, b^0 \rangle$, pro niž platí $f(a^0)f(b^0) < 0$. Zvolíme bod $x^1 \in (a^0, b^0)$, kterým rozdělíme původní interval na dvě části, a jako nový interval $\langle a^1, b^1 \rangle$ vezmeme tu část, v níž leží kořen \bar{x} . Rozhodujeme se takto:

- je-li $f(x^1) = 0$, potom x^1 je kořen, tj. $\bar{x} = x^1$;
- je-li $f(a^0)f(x^1) < 0$, položíme $\langle a^1, b^1 \rangle = \langle a^0, x^1 \rangle$;
- je-li $f(x^1)f(b^0) < 0$, položíme $\langle a^1, b^1 \rangle = \langle x^1, b^0 \rangle$.

Pokud nastane první případ, jsme hotovi. V opačném případě zopakujeme celý postup na intervalu $\langle a^1, b^1 \rangle$, tj. zvolíme bod $x^2 \in (a^1, b^1)$, který buďto je kořenem, nebo s jeho pomocí určíme další interval $\langle a^2, b^2 \rangle$ kořen obsahující atd..

Uvedeným postupem tedy vytvoříme posloupnosti $\{a^k\}$, $\{b^k\}$ a $\{x^k\}$ takové, že kořen \bar{x} leží uvnitř každého z intervalů $\langle a^k, b^k \rangle$. Abychom byli schopni určit číselnou hodnotu kořene \bar{x} , musí k němu konvergovat posloupnost $\{x^k\}$. To zajistíme vhodnou konkrétní volbou bodů x^k .

2.2.1. Metoda půlení intervalu

Bod x^{k+1} určíme jako střed intervalu $\langle a^k, b^k \rangle$ podle vzorce

$$x^{k+1} = \frac{a^k + b^k}{2}. \quad (2.2.1)$$

Intervaly tedy postupně půlíme a jejich středy tvořící posloupnost $\{x^k\}$ konvergují ke kořenu \bar{x} . Výpočet ukončíme při dosažení zadané přesnosti ϵ , tj. když platí

$$|\bar{x} - x^{k+1}| \leq \epsilon.$$

Otázkou je, jak takovou situaci rozpoznat, jelikož \bar{x} neznáme. Musí však platit

$$|\bar{x} - x^{k+1}| \leq \frac{b^k - a^k}{2},$$

protože kořen \bar{x} ležící v intervalu $\langle a^k, b^k \rangle$ se od středu x^{k+1} nemůže lišit víc než o polovinu délky intervalu. Pro ukončení výpočtu proto použijeme kritérium

$$\frac{b^k - a^k}{2} \leq \epsilon \quad (2.2.2)$$

a poslední střed x^{k+1} je pak aproximací kořene \bar{x} s přesností ϵ .

Algoritmus (Metoda půlení intervalu)

Vstup: f , a^0 , b^0 , ϵ .

Pro $k = 0, 1, \dots$ opakuj:

$$x^{k+1} := (a^k + b^k)/2;$$

je-li $f(x^{k+1}) = 0$, potom jdi na Výstup;

je-li $f(a^k)f(x^{k+1}) < 0$, potom $a^{k+1} := a^k$, $b^{k+1} := x^{k+1}$;

je-li $f(x^{k+1})f(b^k) < 0$, potom $a^{k+1} := x^{k+1}$, $b^{k+1} := b^k$;

dokud $b^{k+1} - a^{k+1} > \epsilon$.

Výstup: poslední hodnota x^{k+1} .

Příklad 2.2.1. Metodou půlení intervalu vypočtěte kořen rovnice

$$f(x) \equiv 10 \cos(x - 1) - x^2 + 2x - 1 = 0,$$

který leží v intervalu $\langle 2.3, 2.4 \rangle$ s přesností $\epsilon = 10^{-3}$.

Řešení: Na začátku je $a^0 = 2.3$, $b^0 = 2.4$ a první střed je $x^1 = 2.35$. Tabulka 2.2.1 ukazuje průběh výpočtu. Symbolem $+$ nebo $-$ za číslem označujeme znaménko funkční hodnoty funkce f v tomto bodě. Všimněme si, že x^{k+1} nahrazuje a^k nebo b^k tak, aby byla zachována znaménková změna. Aproximace kořene s přesností ϵ je poslední číslo ve sloupci x^{k+1} . Proto $\bar{x} = 2.378 \pm 10^{-3}$.

Tabulka 2.2.1: Metoda půlení intervalu.

k	a^k	b^k	x^{k+1}	$(b^k - a^k)/2$
0	2.3^+	2.4^-	2.35^+	0.05
1	2.35^+	2.4^-	2.375^+	0.025
2	2.375^+	2.4^-	2.3875^-	0.0125
3	2.375^+	2.3875^-	2.38125^-	0.00625
4	2.375^+	2.38125^-	2.378125^+	0.003125
5	2.378125^+	2.38125^-	2.3796875^-	0.0015625
6	2.378125^+	2.3796875^-	2.37890625^+	$0.00078125 < 10^{-3} = \epsilon$

2.2.2. Metoda regula falsi

V intervalu $\langle a^k, b^k \rangle$ zvolíme bod x^{k+1} jako kořen přímky p , která prochází krajními body grafu funkce f , viz obrázek 2.2.1.α. Uvažovaná přímka je dána předpisem

$$p(x) = f(a^k) + \frac{f(b^k) - f(a^k)}{b^k - a^k}(x - a^k)$$

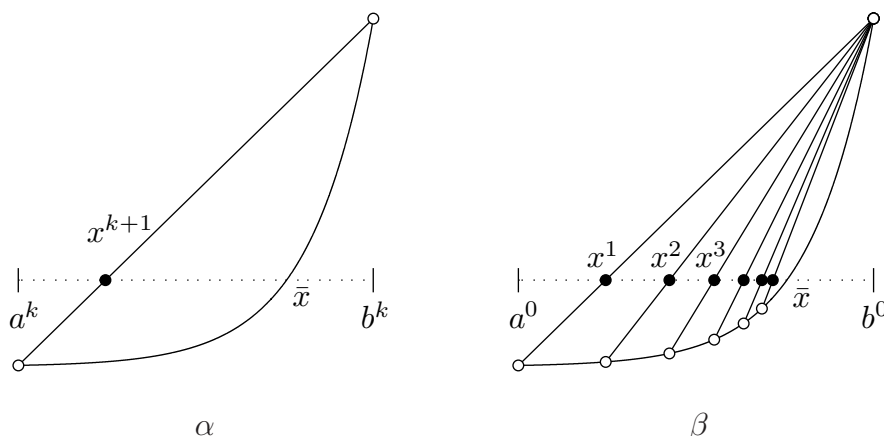
a její kořen je určen rovností $p(x^{k+1}) = 0$. Odtud lze snadno odvodit vzorec

$$x^{k+1} = a^k - \frac{b^k - a^k}{f(b^k) - f(a^k)} f(a^k), \quad (2.2.3)$$

který se používá při výpočtu.

Geometrický smysl metody regula falsi je patrný z obrázku 2.2.1.β. Ukončení iterací se provádí podle kritéria

$$|x^{k+1} - x^k| \leq \epsilon, \quad (2.2.4)$$



Obrázek 2.2.1: Metoda regula falsi.

kde $\epsilon > 0$ je dané malé číslo.

Algoritmus (Metoda regula falsi)

Vstup: f , a^0 , b^0 , ϵ , $x^0 := a^0$.

Pro $k = 0, 1, \dots$ opakuj:

$$x^{k+1} := a^k - (b^k - a^k)/(f(b^k) - f(a^k))f(a^k);$$

je-li $f(x^{k+1}) = 0$, potom jdi na Výstup;

je-li $f(a^k)f(x^{k+1}) < 0$, potom $a^{k+1} := a^k$, $b^{k+1} := x^{k+1}$;

je-li $f(x^{k+1})f(b^k) < 0$, potom $a^{k+1} := x^{k+1}$, $b^{k+1} := b^k$;

dokud $|x^{k+1} - x^k| > \epsilon$.

Výstup: poslední hodnota x^{k+1} .

Příklad 2.2.2. Metodou regula falsi vypočítejte kořen rovnice z příkladu 2.2.1.

Řešení: Na začátku je $a^0 = 2.3$, $b^0 = 2.4$ a položíme ještě $x^0 = a^0$. V první iteraci vypočítáme

$$x^1 := 2.3 - (2.4 - 2.3)f(2.3)/(f(2.4) - f(2.3)) \doteq 2.379095,$$

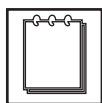
$$|x^1 - x^0| = |2.379095 - 2.3| = 0.079095.$$

Tabulka 2.2.2 zachycuje celý výpočet, který se řídí podobnými pravidly jako u me-

tody půlení intervalu. Výsledná aproximace kořene je $\bar{x} = 2.379 \pm 10^{-3}$.

Tabulka 2.2.2: Metoda regula falsi.

k	a^k	b^k	x^{k+1}	$ x^{k+1} - x^k $
0	2.3^+	2.4^-	2.379095^+	0.079095
1	2.379095^+	2.4^-	2.379363^+	$0.000268 < 10^{-3} = \epsilon$



Poznámka

Ukončovací kritérium (2.2.4) říká, že poslední dvě aproximace kořene se liší méně než ϵ . Může se ovšem stát, že obě jsou od skutečné hodnoty kořene vzdálené více než ϵ . Poznáme to tak, že u funkčních hodnot $f(x^k - \epsilon)$, $f(x^k)$, $f(x^k + \epsilon)$ nedojde ke znaménkové změně. V takovém případě je možno provést doplňující výpočet funkčních hodnot

$$\dots, f(x^k - 2\epsilon), f(x^k - \epsilon), f(x^k), f(x^k + \epsilon), f(x^k + 2\epsilon), \dots$$

který zastavíme, když dojde ke znaménkové změně.

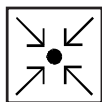


Kontrolní otázky

Otázka 1. V čem se shodují a v čem se liší metoda půlení intervalu a metoda regula falsi? Která z nich je rychlejší?

Otázka 2. Podrobně odvoďte vzorec (2.2.3).

Otázka 3. Proč nelze metodu regula falsi ukončovat podle kritéria (2.2.2)?



Úlohy k samostatnému řešení

1. Vypočtete kořeny rovnice $x^2 - x - \frac{6}{7} \ln x = 0$ metodou půlení intervalu, $\epsilon = 10^{-3}$.
2. Vypočtete kořeny rovnice z předchozí úlohy metodou regula falsi.



Výsledky úloh k samostatnému řešení

1. Začneme-li na intervalu $(0.9, 0.91)$, bude ve čtvrté iteraci $\bar{x} = 0.9019 \pm 10^{-3}$.
2. Začneme-li na stejném intervalu, bude ve druhé iteraci $\bar{x} = 0.9021 \pm 10^{-3}$.

2.3. Newtonova metoda

Cíle



Odvodíme Newtonovu metodu, která kromě funkčních hodnot používá také hodnoty první derivace. Důsledkem je vyšší řád metody a rychlejší konvergence. Nevýhodou je nutnost zvolit počáteční aproximaci „blízko“ kořene tak, aby byly splněny předpoklady zaručující konvergenci.

Předpokládané znalosti



Spojitosť funkce. Výpočty derivací. Taylorův rozvoj.

Výklad



Budeme předpokládat, že známe aproximaci x^k kořene \bar{x} rovnice $f(x) = 0$ a chceme určit další (přesnější) aproximaci x^{k+1} . Zapišeme-li danou rovnici pomocí Taylorova polynomu prvního stupně v okolí bodu x^k , dostaneme

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

kde ξ je blíže neurčené číslo mezi x a x^k . V tomto tvaru rovnice provedeme *linearizaci*, při níž vypustíme kvadratický člen na levé straně. Řešením linearizované rovnice určíme aproximaci x^{k+1} , tj.

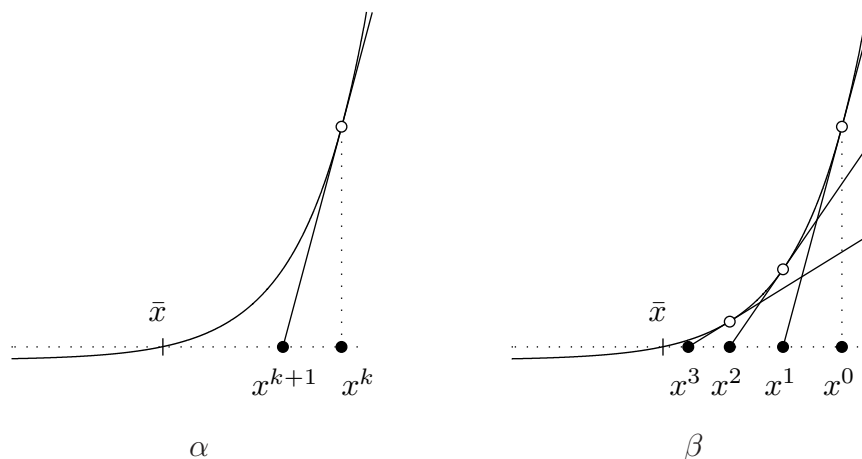
$$f(x^k) + (x^{k+1} - x^k)f'(x^k) = 0.$$

Odtud snadno vyjádříme vzorec

$$x^{k+1} = x^k - \frac{f(x^k)}{f'(x^k)}. \quad (2.3.1)$$

Všimněme si ještě, že $t(x) = f(x^k) + (x - x^k)f'(x^k)$ je rovnice tečny ke grafu funkce f v bodě x^k , jejíž kořen je aproximace x^{k+1} , viz obrázek 2.3.1.α. Geometrický smysl Newtonovy metody je znázorněn na obrázku 2.3.1.β.

Pro zahájení výpočtu podle vzorcem (2.3.1) musíme zadat počáteční aproximaci x^0 . O ukončení iterací rozhodujeme pomocí vhodného kritéria, budeme používat opět kritérium (2.2.4).



Obrázek 2.3.1: Newtonova metoda.

Algoritmus (Newtonova metoda)

Vstup: f , f' , x^0 , ϵ .

Pro $k = 0, 1, \dots$ opakuj:

$$x^{k+1} := x^k - f(x^k)/f'(x^k);$$

dokud $|x^{k+1} - x^k| > \epsilon$.

Výstup: poslední hodnota x^{k+1} .

Příklad 2.3.1. Newtonovou metodou vypočtete kořen rovnice z příkladu 2.2.1. s přesností $\epsilon = 10^{-6}$.

Řešení: Derivace funkce $f(x) = 10 \cos(x - 1) - x^2 + 2x - 1$ má tvar $f'(x) = -10 \sin(x - 1) - 2x + 2$. Newtonova metoda je proto dána vzorcem

$$x^{k+1} = x^k - \frac{10 \cos(x^k - 1) - (x^k)^2 + 2x^k - 1}{-10 \sin(x^k - 1) - 2x^k + 2}.$$

Počáteční aproximaci zvolíme například $x^0 = 2.4$ a dostaneme

$$x^1 = 2.4 - \frac{10 \cos(2.4 - 1) - 2.4^2 + 2 \cdot 2.4 - 1}{-10 \sin(2.4 - 1) - 2 \cdot 2.4 + 2} = 2.37942798004.$$

Celý průběh výpočtu je zachycen v tabulce 2.3.1. Výsledná aproximace kořene je $\bar{x} = 2.379364 \pm 10^{-6}$. \square

Tabulka 2.3.1: Newtonova metoda.

k	x^k	$ x^k - x^{k-1} $
1	2.37942798004	0.02057201996
2	2.37936459485	0.00006338519
3	2.37936459422	0.00000000062

V příkladu jsme měli možnost si všimnout velice rychlé konvergence Newtonovy metody. Následující analýza ukazuje, že se nejednalo o náhodu.

Věta 2.3.1.

Nechť f'' je spojitá, f' nenulová na $\langle a, b \rangle$ a nechť $\{x^k\}$ je posloupnost na $\langle a, b \rangle$ počítaná podle vzorce (2.3.1), která konverguje k číslu \bar{x} . Potom \bar{x} je kořenem rovnice $f(x) = 0$ a platí

$$|\bar{x} - x^{k+1}| \leq C |\bar{x} - x^k|^2 \quad \text{pro } k = 0, 1, 2, \dots, \quad (2.3.2)$$

kde konstanta $C \geq 0$ nezávisí na k .

Důkaz: Limitním přechodem ve vzorci (2.3.1) dostaneme $f(\bar{x}) = 0$, čili \bar{x} je kořen. Protože předpokládáme, že f'' je spojitá funkce, můžeme $f(\bar{x}) = 0$ zapsat pomocí Taylorova rozvoje ve tvaru

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

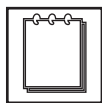
kde ξ je bod mezi x^k a \bar{x} . Vydělíme derivací $f'(x^k)$ a dosadíme ze vzorce (2.3.1), čímž dostaneme

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

Označíme-li $C = \max_{\xi, x \in \langle a, b \rangle} |f''(\xi)/2f'(x)|$, můžeme psát

$$|\bar{x} - x^{k+1}| = \left| \frac{f''(\xi)}{2f'(x^k)} \right| |\bar{x} - x^k|^2 \leq C |\bar{x} - x^k|^2.$$

\square

**Poznámka**

Nerovnost (2.3.2) dokazuje, že Newtonova metoda je **druhého řádu**. Důsledkem tohoto faktu je velmi rychlá konvergence, při níž se počet správných číslic za desetinnou tečkou v každé iteraci „přibližně“ zdvojnásobí.

Následující věta říká, za jakých předpokladů konvergence nastane.

Věta 2.3.2.

Nechť jsou splněny následující předpoklady:

- a) f' je nenulová na intervalu $\langle a, b \rangle$;
- b) f'' nemění znaménko v intervalu (a, b) ;
- c) platí $f(a)f(b) < 0$;
- d) platí $|f(a)/f'(a)| < b - a$ a $|f(b)/f'(b)| < b - a$.

Potom posloupnost $\{x^k\}$ počítaná podle vzorce (2.3.1) konverguje pro libovolnou počáteční aproximaci $x^0 \in \langle a, b \rangle$.

Příklad 2.3.2. Ukažte, že pro rovnici z příkladu 2.3.1. jsou na intervalu $\langle 2.3, 2.4 \rangle$ splněny předpoklady věty 2.3.2.

Řešení: Rovnice je zadána funkcí $f(x) = 10 \cos(x - 1) - x^2 + 2x - 1$, jejíž první a druhá derivace mají tvar $f'(x) = -10 \sin(x - 1) - 2x + 2$ a $f''(x) = -10 \cos(x - 1) - 2$. V tabulce 2.3.2 jsou vypočteny hodnoty f' , f'' , z nichž můžeme usoudit, že jsou splněny předpoklady a) a b).

Výpočtem dostáváme $f(2.3)f(2.4) \doteq -0.2564 < 0$, $|f(2.3)/f'(2.3)| \doteq 0.0805 < 0.1$ a $|f(2.4)/f'(2.4)| \doteq 0.0206 < 0.1$, což znamená, že jsou splněny také předpoklady c) a d). Počáteční aproximaci x^0 je proto možné zvolit libovolně na $\langle 2.3, 2.4 \rangle$.

Tabulka 2.3.2: Tabelace první a druhé derivace.

x	$f'(x)$	$f''(x)$
2.30	-12.2356	-4.6750
2.31	-12.2818	-4.5785
2.32	-12.3272	-4.4818
2.33	-12.3715	-4.3848
2.34	-12.4148	-4.2875
2.35	-12.4572	-4.1901
2.36	-12.4986	-4.0924
2.37	-12.5391	-3.9945
2.38	-12.5785	-3.8964
2.39	-12.6170	-3.7981
2.40	-12.6545	-3.6997

Kontrolní otázky

Otázka 1. Jak se odvozuje vzorec (2.3.1)? Jaký je jeho grafický smysl?

Otázka 2. Zkuste nakreslit situace, kdy Newtonova metoda diverguje.

Otázka 3. Jakého řádu je Newtonova metoda a jak se to projeví při výpočtu?

Úlohy k samostatnému řešení

1. Ověřte, že pro rovnici $x^2 - x - \frac{6}{7} \ln x = 0$ jsou na intervalu $\langle 0.9, 0.91 \rangle$ splněny předpoklady a)–d) z věty 2.3.2.

2. Vypočtete kořen z předchozí úlohy pomocí Newtonovy metody, $\epsilon = 10^{-6}$.

Výsledky úloh k samostatnému řešení

1. Z tabelace $f'(x) = 2x - 1 - \frac{6}{7x}$ a $f''(x) = 2 + \frac{6}{7x^2}$ zjistíme, že na uvedeném intervalu je první derivace záporná a druhá derivace kladná. Přímým výpočtem dostaneme $f(0.9)f(0.91) \doteq -3.28 \cdot 10^{-7} < 0$, $|f(0.9)/f'(0.9)| \doteq 2.03 \cdot 10^{-3} < 0.01$ a $|f(0.91)/f'(0.91)| \doteq 8.71 \cdot 10^{-3} < 0.01$.

2. Začneme-li z $x^0 = 0.9$, dostaneme ve třetí iteraci $\bar{x} = 0.9020709 \pm 10^{-6}$.

2.4. Metoda prosté iterace



Cíle

Seznámíme se s obecnou metodou pro výpočet pevného bodu funkce. Metody z předchozích odstavců můžeme chápat jako její speciální varianty. Pojem pevného bodu hraje důležitou roli v různých oblastech matematického modelování.



Předpokládané znalosti

Spojitosť funkce. Výpočty derivací. Věta o střední hodnotě diferenciálního počtu.



Výklad

Rovnici $f(x) = 0$ převedeme na ekvivalentní rovnici $x - g(x) = 0$, kde g je vhodná spojitá funkce. Místo původní rovnice budeme řešit rovnici v *iteračním tvaru*:

$$x = g(x). \quad (2.4.1)$$

Číslo \bar{x} , které je řešením rovnice (2.4.1), se nazývá *pevný bod funkce g* .

Věta 2.4.1.

(Brouwerova věta o pevném bodu) Nechť g je spojitá funkce na intervalu $\langle a, b \rangle$, pro níž platí

$$g(x) \in \langle a, b \rangle \quad \forall x \in \langle a, b \rangle. \quad (2.4.2)$$

Pak na intervalu $\langle a, b \rangle$ existuje pevný bod funkce g .

Důkaz: Položme $f(x) = x - g(x)$. Pokud $f(a) = 0$ resp. $f(b) = 0$, pak je pevným bodem a , resp. b . Nechť $f(a) \neq 0$ a $f(b) \neq 0$. Protože $g(a) \in \langle a, b \rangle$, platí $f(a) = a - g(a) < 0$. Podobně lze ukázat $f(b) > 0$. Dohromady dostáváme

$f(a)f(b) < 0$, takže existence pevného bodu plyne z věty 2.1.1. \square

O funkci g , která splňuje (2.4.2), říkáme, že zobrazuje interval $\langle a, b \rangle$ do sebe.

Nechť $x^0 \in \langle a, b \rangle$ je počáteční aproximace. *Metodou prostých iterací* nazýváme výpočet podle předpisu:

$$x^{k+1} = g(x^k), \quad k = 0, 1, 2, \dots \quad (2.4.3)$$

Jestliže posloupnost $\{x^k\}$ počítaná tímto postupem konverguje k číslu \bar{x} , pak limitním přechodem v (2.4.3) dostaneme, že \bar{x} je pevným bodem funkce g .

Výchozí rovnici $f(x) = 0$ můžeme převést na iterační tvar různými způsoby, ale jen některé vedou ke konvergentnímu výpočtu.

Příklad 2.4.1. Rovnici

$$f(x) \equiv e^x - x^2 + 1 = 0, \quad (2.4.4)$$

převeďte na iterační tvar a počítejte kořen, který leží v intervalu $\langle -1.2, -1.1 \rangle$.

Řešení: Navrhujeme tři iterační tvary:

$$x = -\sqrt{e^x + 1} \quad \equiv \quad g_a(x);$$

$$x = x + (e^x - x^2 + 1) \quad \equiv \quad g_b(x);$$

$$x = x - \frac{e^x - x^2 + 1}{e^x - 2x} \quad \equiv \quad g_c(x).$$

Průběh výpočtů pro $x^0 = -1.1$ ukazuje tabulka 2.4.1. Pro g_a výpočet konverguje pomalu, pro g_b výpočet diverguje a pro g_c výpočet konverguje rychle. \square

Tabulka 2.4.1: Metoda prosté iterace.

k	$x^{k+1} = g_a(x^k)$	$x^{k+1} = g_b(x^k)$	$x^{k+1} = g_c(x^k)$
0	-1.1	-1.1	-1.1
1	-1.1545003	-0.9771289	-1.1485105
2	-1.1468282	-0.5555196	-1.1477578
3	-1.1478861	+0.7096523	-1.1477576
4	-1.1477398	+3.2393301	-1.1477576
5	-1.1477600	+19.262693	
6	-1.1477572	∞	
7	-1.1477576		
8	-1.1477576		

Při studiu konvergence metody prostých iterací se používá pojem kontrakce.

Definice 2.4.1.

Funkce g se nazývá *kontrakce na intervalu* $\langle a, b \rangle$, jestliže existuje konstanta L , $0 < L < 1$, taková, že

$$|g(x) - g(y)| \leq L|x - y| \quad \forall x, y \in \langle a, b \rangle. \quad (2.4.5)$$

Věta 2.4.2.

Nechť g je spojitá kontrakce na $\langle a, b \rangle$, která zobrazuje tento interval do sebe. Pak na intervalu $\langle a, b \rangle$ existuje jediný pevný bod \bar{x} funkce g . Navíc posloupnost $\{x^k\}$ vypočítaná podle předpisu (2.4.3) konverguje k \bar{x} pro každou počáteční aproximaci $x^0 \in \langle a, b \rangle$.

Důkaz: Existence pevného bodu plyne z věty 2.4.1. Jednoznačnost dokážeme sporem. Nechť \tilde{x} je další pevný bod g . Pomocí (2.4.5) dostaneme

$$|\bar{x} - \tilde{x}| = |g(\bar{x}) - g(\tilde{x})| \leq L|\bar{x} - \tilde{x}|,$$

odkud $(1 - L)|\bar{x} - \tilde{x}| \leq 0$. Protože $1 - L > 0$, dostáváme $\bar{x} = \tilde{x}$.

Zbývá dokázat, že posloupnost $\{x^k\}$ vypočítaná podle předpisu (2.4.3) konverguje

k \bar{x} . Podle (2.4.5) je

$$|x^k - \bar{x}| = |g(x^{k-1}) - g(\bar{x})| \leq L|x^{k-1} - \bar{x}|,$$

odkud plyne

$$|x^k - \bar{x}| \leq L^k |x^0 - \bar{x}|. \quad (2.4.6)$$

Protože $L \in (0, 1)$, je $\lim_{k \rightarrow \infty} L^k = 0$, a proto $\lim_{k \rightarrow \infty} |x^k - \bar{x}| = 0$. \square

V konkrétních situacích je zpravidla nesnadné dokázat, že daná funkce g je kontrakce. Jednodušší je ověřovat následující silnější předpoklad:

$$\begin{aligned} &\text{nechť } g \text{ má v } (a, b) \text{ derivaci a} \\ &\exists L \in (0, 1) \text{ tak, že } |g'(\eta)| \leq L \quad \forall \eta \in (a, b). \end{aligned} \quad (2.4.7)$$

Věta 2.4.3.

Nechť g je spojitá funkce na $\langle a, b \rangle$, která zobrazuje tento interval do sebe a splňuje (2.4.7). Pak platí tvrzení věty 2.4.2..

Důkaz: Pomocí věty o střední hodnotě diferenciálního počtu dostáváme

$$|g(x) - g(y)| = |g'(\eta)| |x - y| \leq L|x - y|,$$

kde $\eta \in (x, y)$. Funkce g je proto kontrakce na intervalu $\langle a, b \rangle$ a věta 2.4.3. je tak důsledkem věty 2.4.2. \square

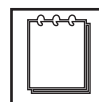
Následující poznámka dává návod, jak rozhodnout o konvergenci metody prostých iterací.

Poznámka

Je-li číslo

$$M_g = \max_{x \in (a, b)} |g'(x)|$$

menší než jedna, pak můžeme položit $L = M_g$ a funkce g bude kontrakce na intervalu $\langle a, b \rangle$. Rychlost konvergence lze posoudit podle velikosti L . Vztah (2.4.6) totiž ukazuje, že výpočet bude konvergovat rychleji pro menší hodnoty L .



Příklad 2.4.2. Rozhodněte o konvergenci metody prostých iterací u iteračních tvarů z příkladu 2.4.1.

Řešení: Derivováním g_a , g_b a g_c dostaneme

$$\begin{aligned} g'_a(x) &= -\frac{e^x}{2\sqrt{e^x+1}}, \\ g'_b(x) &= 1 + e^x - 2x, \\ g'_c(x) &= \frac{(e^x - x^2 + 1)(e^x - 2)}{(e^x - 2x)^2}. \end{aligned}$$

Tabulka 2.4.2 obsahuje absolutní hodnoty těchto derivací na intervalu $\langle -1.2, -1.1 \rangle$. Odtud $M_{g_a} \doteq 0.1442$, $M_{g_b} \doteq 3.7012$ a $M_{g_c} \doteq 0.0323$. Protože $M_{g_b} > 1$ iterační tvar b) diverguje. Pro iterační tvary a) resp. c) můžeme položit $L_{g_a} = M_{g_a}$ resp. $L_{g_c} = M_{g_c}$, takže funkce g_a a g_b jsou kontrakce a metoda prostých iterací konverguje. Protože $L_{g_c} < L_{g_a}$, je konvergence rychlejší u iteračního tvaru c).

Tabulka 2.4.2: Posouzení rychlosti konvergence metody prosté iterace.

x	$ g'_a(x) $	$ g'_b(x) $	$ g'_c(x) $
-1.2000	0.1320	3.7012	0.0323
-1.1875	0.1335	3.6800	0.0248
-1.1750	0.1350	3.6588	0.0172
-1.1625	0.1365	3.6377	0.0094
-1.1500	0.1380	3.6166	0.0014
-1.1375	0.1395	3.5956	0.0067
-1.1250	0.1410	3.5747	0.0149
-1.1125	0.1426	3.5537	0.0233
-1.1000	0.1442	3.5329	0.0319

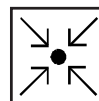


Kontrolní otázky

Otázka 1. Co nazýváme pevným bodem funkce? Jak se pevný bod počítá?

Otázka 2. Čím je zaručena konvergence metody prostých iterací?

Otázka 3. Jaký je vztah mezi metodou prosté iterace a Newtonovou metodou?

Úlohy k samostatnému řešení

1. Pro rovnici $x^2 - x - \frac{6}{7} \ln x = 0$ uvažujte iterační tvar $x = x^2 - \frac{6}{7} \log x \equiv g(x)$. Vypočítejte hodnotu konstanty M_g na intervalu $\langle 0.9, 0.91 \rangle$.
2. U iteračního tvaru z předchozí úlohy vypočítejte pevný bod s přesností $\epsilon = 10^{-6}$.

Výsledky úloh k samostatnému řešení

1. Z tabelace $g'(x) = 2x - \frac{6}{7x}$ zjistíme, že $M_g = L \doteq 0.8781$.
2. Začneme-li z $x^0 = 0.9$, dostaneme ve 38-mé iteraci $x^{38} \doteq 0.9020659$ a platí $|x^{38} - x^{37}| \doteq 0.00000086 < \epsilon$.

Shrnutí lekce

Probrali jsem základní metody pro řešení nelineárních rovnic a ukázali jsme jak tyto metody souvisí s vlastnostmi funkce, která rovnici popisuje. Metoda půlení intervalu a metoda regula falsi konvergují pro každou spojitou funkci. Rychlejší Newtonova metoda vyžaduje spojitost druhé derivace a splnění dalších předpokladů. Metoda prosté iterace zahrnuje ostatní iterační metody jako speciální případ.

3. SLR – přímé metody



Průvodce studiem

Přímé metody řešení soustav lineárních rovnic (SLR) jsou založeny na eliminaci neznámých. Výchozí myšlenka spočívá v tom, že z některé rovnice vyjádříme jednu neznámou a dosadíme ji do ostatních rovnic tak, aby soustava po eliminaci byla snáze řešitelná než soustava původní. Základní algoritmus tohoto typu je *Gaussova eliminační metoda*. V maticovém zápisu jí odpovídá *LU-rozklad matice*. Charakteristickým rysem přímých metod je výpočet (přesného) řešení po konečném počtu eliminací, tj. po konečném počtu aritmetických operací.

3.1. Formulace úlohy



Cíle

Připomeneme základní poznatky o soustavách lineárních rovnic.



Předpokládané znalosti

Operace s maticemi. Determinant. Matice jednotková a inverzní.



Výklad

V předchozích odstavcích jsme řešili rovnici $f(x) = 0$. Nejjednodušším příkladem je lineární rovnice $ax = b$. Je-li $a \neq 0$, můžeme řešení zapsat ve tvaru $x = a^{-1}b$. Nyní se budeme zabývat zobecněním této úlohy:

Nechť $\mathbf{A} = (a_{ij})$ je daná čtvercová matice řádu n s prvky a_{ij} a nechť

$\mathbf{b} = (b_i)$ je daný sloupcový vektor s n prvky b_i , $i = 1, \dots, n$, $j =$

$1, \dots, n$; hledáme sloupcový vektor \mathbf{x} takový, že

$$\mathbf{Ax} = \mathbf{b}. \quad (3.1.1)$$

[illegible]

Je-li matice \mathbf{A} regulární, pak můžeme obě strany rovnice (3.1.1) vynásobit inverzní maticí \mathbf{A}^{-1} , tj. $\mathbf{A}^{-1}(\mathbf{A}\mathbf{x}) = \mathbf{A}^{-1}\mathbf{b}$. Protože $\mathbf{A}^{-1}(\mathbf{A}\mathbf{x}) = (\mathbf{A}^{-1}\mathbf{A})\mathbf{x} = \mathbf{I}\mathbf{x} = \mathbf{x}$, dostáváme

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}. \quad (3.1.2)$$

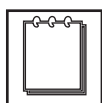
Příklad 3.1.1. Pomocí vzorce (3.1.2) vyřešte soustavu lineárních rovnic:

$$\begin{aligned} x_1 + x_2 + x_3 &= 6, \\ 2x_1 + 4x_2 + 2x_3 &= 16, \\ -x_1 + 5x_2 - 4x_3 &= -3. \end{aligned} \tag{3.1.3}$$

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 2 \\ -1 & 5 & -4 \end{pmatrix}, \quad \mathbf{A}^{-1} = \begin{pmatrix} \frac{13}{3} & -\frac{3}{2} & \frac{1}{3} \\ -1 & \frac{1}{2} & 0 \\ -\frac{7}{3} & 1 & -\frac{1}{3} \end{pmatrix}$$

(stačí ověřit $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$), dostáváme řešení

$$\mathbf{x} = \begin{pmatrix} \frac{13}{3} & -\frac{3}{2} & \frac{1}{3} \\ -1 & \frac{1}{2} & 0 \\ -\frac{7}{3} & 1 & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} 6 \\ 16 \\ -3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$



Poznámka

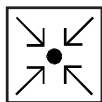
U rozsáhlejších soustav lineárních rovnic je použití vzorce (3.1.2) neekonomické, protože nalezení inverzní matice vyžaduje velké množství výpočtů.



Kontrolní otázky

Otázka 1. Kdy je matice regulární a kdy je singulární?

Otázka 2. Připomeňte si metody výpočtu determinantu a inverzní matice.



Úlohy k samostatnému řešení

1. Uvažujme soustavu lineárních rovnic

$$-x_1 - 3x_2 + 2x_3 = -9,$$

$$-6x_1 - 19x_2 + 10x_3 = -59,$$

$$3x_1 + 9x_2 - 5x_3 = 28.$$

Vypočtěte inverzní matici a soustavu vyřešte pomocí vzorce (3.1.2).



Výsledky úloh k samostatnému řešení

1.

$$\mathbf{A} = \begin{pmatrix} -1 & -3 & 2 \\ -6 & -19 & 10 \\ 3 & 9 & -5 \end{pmatrix}, \quad \mathbf{A}^{-1} = \begin{pmatrix} 5 & 3 & 8 \\ 0 & -1 & -2 \\ 3 & 0 & 1 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}.$$

3.2. Gaussova eliminační metoda (GEM)

Cíle

Připomeneme GEM a posoudíme výpočetní náročnost jejích částí. Pro jednoduchost budeme uvažovat soustavy s regulární maticí.



Předpokládané znalosti

Provádění eliminací u soustav lineárních rovnic.



Výklad

Nejdříve ukážeme GEM na příkladu. Soustavu lineárních rovnic (3.1.3) můžeme zapsat ve tvaru:



$$\mathbf{Ax} = \mathbf{b}, \text{ tj. } \begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 2 \\ -1 & 5 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 16 \\ -3 \end{pmatrix}. \quad (3.2.1)$$

V první fázi eliminujeme v prvním sloupci. První rovnici vynásobíme číslem $m_{21} = -2$ resp. $m_{31} = 1$ a přičteme ke druhé resp. třetí rovnici. Dostaneme:

$$\mathbf{A}_1\mathbf{x} = \mathbf{b}_1, \text{ tj. } \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 6 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \\ 3 \end{pmatrix}.$$

Ve druhé fázi eliminujeme ve druhém sloupci. Druhou rovnici vynásobíme číslem $m_{32} = -3$ a přičteme ke třetí rovnici. Dostaneme:

$$\mathbf{A}_2\mathbf{x} = \mathbf{b}_2, \text{ tj. } \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \\ -9 \end{pmatrix}$$

Poslední soustavu s *horní trojúhelníkovou maticí* budeme zapisovat jako $\mathbf{Ux} = \mathbf{y}$, tj. $\mathbf{U} = \mathbf{A}_2$, $\mathbf{y} = \mathbf{b}_2$. Pro lepší názornost použijeme zápis po rovnicích:

$$\begin{aligned} x_1 + x_2 + x_3 &= 6, \\ 2x_2 &= 4, \\ -3x_3 &= -9. \end{aligned} \tag{3.2.2}$$

Odtud $x_3 = 3$, $x_2 = 2$ a $x_1 = 1$.

Na příkladu jsme viděli, že GEM má dvě části:

- *dopředný chod* je úprava výchozí soustavy $\mathbf{Ax} = \mathbf{b}$ na soustavu $\mathbf{Ux} = \mathbf{y}$ s horní trojúhelníkovou maticí \mathbf{U} ;
- *zpětný chod* je výpočet řešení ze soustavy $\mathbf{Ux} = \mathbf{y}$.

3.2.1. Zpětný chod GEM

Uvažujme soustavu lineárních rovnic $\mathbf{U}\mathbf{x} = \mathbf{y}$ s horní trojúhelníkovou maticí $\mathbf{U} = (u_{ij})$, $u_{ij} = 0$, $i > j$, a vektorem pravé strany $\mathbf{y} = (y_i)$. Zapišeme-li tuto soustavu po jednotlivých rovnicích, dostaneme

$$\begin{aligned} u_{11}x_1 + u_{12}x_2 + \dots + u_{1n}x_n &= y_1, \\ u_{22}x_2 + \dots + u_{2n}x_n &= y_2, \\ &\vdots \\ u_{nn}x_n &= y_n. \end{aligned}$$

Výpočet řešení $\mathbf{x} = (x_i)$ se provádí postupným dosazováním „od konce”.

Algoritmus (Zpětný chod GEM)

Vstup: $\mathbf{U} = (u_{ij}), \mathbf{y} = (y_i)$.

$$x_n := y_n / u_{nn}.$$

Pro $i = n - 1, \dots, 1$ počítej i -tou neznámou:

$$x_i := (y_i - u_{in}x_n - \dots - u_{ii+1}x_{i+1})/u_{ii}.$$

Výstup: $\mathbf{x} = (x_i)$.

Všimněme si počtu operací. Při výpočtu i -té neznámé potřebujeme provést

jedno dělení a $n - i$ odčítání a násobení. Celkový počet operací je

$$\sum_{i=n}^1 [1 + 2(n - i)] = n^2 = \mathcal{O}(n^2).$$

3.2.2. Dopředný chod GEM

Dopředný chod GEM se dělí na *fáze*. V k -té fázi, $1 \leq k \leq n - 1$, se provádí eliminace v k -tém sloupci matice. Pro jednoduchost budeme předpokládat, že není potřeba měnit pořadí řádků tak, jak tomu bylo v našem úvodním příkladu. Na tomto příkladu by si měl čtenář také ilustrovat všechny níže uvedené pojmy.

Prvky matice na začátku k -té fáze označíme $a_{ij}^{(k)}$ a prvky vektoru pravé strany $a_{in+1}^{(k)}$ (na začátku 1. fáze je $a_{ij}^{(1)} = a_{ij}$ a $a_{in+1}^{(1)} = b_i$). Eliminace provádíme v k -tém sloupci matice pod jejím diagonálním prvkem $a_{kk}^{(k)}$, kterému říkáme *hlavní prvek k -té fáze*. Nejprve počítáme *multiplikátory k -té fáze*

$$m_{ik} = -\frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}, \quad i = k + 1, \dots, n \quad (3.2.3)$$

a pak přičteme m_{ik} -násobek k -tého řádku k řádku i -tému, tj.

$$a_{ij}^{(k+1)} := a_{ij}^{(k)} + m_{ik} a_{kj}^{(k)}, \quad j = k + 1, \dots, n + 1,$$

pro $i = k + 1, \dots, n + 1$.

Algoritmus (Dopředný chod GEM)

Vstup: $\mathbf{A} = (a_{ij})$, $\mathbf{b} = (b_i)$.

$a_{ij}^{(1)} := a_{ij}$, $a_{in+1}^{(1)} := b_i$, $i, j = 1, \dots, n$.

Pro $k = 1, \dots, n - 1$ proved' k -tou fázi:

Pro $i = k + 1, \dots, n$ přičti m_{ik} -násobek k -tého řádku k i -tému řádku:

$$m_{ik} := -a_{ik}^{(k)} / a_{kk}^{(k)};$$

Pro $j = k + 1, \dots, n + 1$ proved' přičítání v j -tém sloupci:

$$a_{ij}^{(k+1)} := a_{ij}^{(k)} + m_{ik} a_{kj}^{(k)}.$$

Polož $u_{ij} := a_{ij}^{(n)}$ pro $i \leq j$, $u_{ij} := 0$ pro $i > j$, $y_i := a_{in+1}^{(n)}$, $i, j = 1, \dots, n$.

Výstup: $\mathbf{U} = (u_{ij})$, $\mathbf{y} = (y_i)$.

Ze vzorce (3.2.3) je vidět, že hlavní prvek musí být nenulový. Pokud není, provedeme na začátku k -té fáze *výběr hlavního prvku*, viz příští odstavec.

Všimněme si ještě počtu operací. V k -té fázi musíme provést $n - k$ dělení při výpočtu multiplikátorů a kromě toho počítáme $(n - k)(n - k + 1)$ prvků $a_{ij}^{(k+1)}$ pomocí dvou operací (násobení a sčítání). Celkový počet operací je proto

$$\sum_{k=1}^{n-1} [(n - k) + 2(n - k)(n - k + 1)] = \frac{2}{3}n^3 - \frac{1}{2}n^2 + \frac{1}{3}n.$$

Pro velká n převažuje v posledním výrazu člen $\frac{2}{3}n^3$. Říkáme, že dopředný chod GEM vyžaduje $\mathcal{O}(\frac{2}{3}n^3)$ operací.

3.2.3. Výběr hlavního prvku

Cílem je vybrat hlavní prvek tak, aby byl co největší v absolutní hodnotě.

V první fázi dopředného chodu GEM se za hlavní prvek vybere v absolutní hodnotě největší číslo z prvního sloupce matice a eliminace se provedou ve všech řádcích neobsahujících hlavní prvek. V k -té fázi se celý výpočet omezí na řádky, v nichž dosud hlavní prvek nebyl vybrán. Nejprve se jako hlavní prvek vybere v absolutní hodnotě největší z čísel ležících v k -tém sloupci a příslušných řádcích a pak se provedou eliminace ve zbývajících řádcích.

Tento postup lze snadno realizovat přehazováním řádků. Do algoritmu dopředného chodu GEM stačí na začátek k -té fáze vsunout následující doplněk:

$$\begin{cases} \text{Najdi } p, p \geq k, \text{ takové, že } |a_{pk}^{(k)}| = \max\{|a_{ik}^{(k)}|, i \geq k\}; \\ \text{Prohod } p\text{-tý a } k\text{-tý řádek matice v } k\text{-té fázi.} \end{cases}$$

Algoritmus dopředného chodu GEM s výběrem hlavního prvku je použitelný pro každou regulární matici.

Příklad 3.2.1. Soustavu lineárních rovnic (3.2.1) řešte pomocí GEM s výběrem hlavního prvku.

Řešení: Výběr hlavního prvku a eliminace v první fázi:

$$\begin{pmatrix} 2 & 4 & 2 \\ 1 & 1 & 1 \\ -1 & 5 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 16 \\ 6 \\ -3 \end{pmatrix}, \begin{pmatrix} 2 & 4 & 2 \\ 0 & -1 & 0 \\ 0 & 7 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 16 \\ -2 \\ 5 \end{pmatrix}.$$

Výběr hlavního prvku a eliminace ve druhé fázi:

$$\begin{pmatrix} 2 & 4 & 2 \\ 0 & 7 & -3 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 16 \\ 5 \\ -2 \end{pmatrix}, \begin{pmatrix} 2 & 4 & 2 \\ 0 & 7 & -3 \\ 0 & 0 & -\frac{3}{7} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 16 \\ 5 \\ -\frac{9}{7} \end{pmatrix}$$

Zpětným chodem GEM vypočítáme $x_3 = 3$, $x_2 = 2$ a $x_1 = 1$.

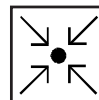
Kontrolní otázky



Otázka 1. Z jakých částí se skládá algoritmus GEM a jak jsou výpočetně náročné?

Otázka 2. Proč se provádí výběr hlavního prvku?

Úlohy k samostatnému řešení



1. Soustavu lineárních rovnic

$$-x_1 - 3x_2 + 2x_3 = -9,$$

$$-6x_1 - 19x_2 + 10x_3 = -59,$$

$$3x_1 + 9x_2 - 5x_3 = 28$$

řešte pomocí GEM bez výběru a s výběrem hlavního prvku.

Výsledky úloh k samostatnému řešení



1. Řešením je vektor $\mathbf{x} = (2, 3, 1)^\top$.

3.3. LU–rozklad



Cíle

Z Gaussovy eliminační metody odvodíme LU-rozklad matice.



Předpokládané znalosti

Dopředný a zpětný chod GEM. Výběr hlavního prvku. Permutační matice.



Výklad

K regulární čtvercové matici \mathbf{A} budeme hledat dolní trojúhelníkovou matici \mathbf{L} a horní trojúhelníkovou matici \mathbf{U} tak, aby platilo

$$\mathbf{A} = \mathbf{L}\mathbf{U}.$$

Začneme příkladem s maticí soustavy (3.2.1). Navážeme přitom na příklad ze začátku odstavce 3.2, kdy jsme pomocí dopředného chodu GEM vytvořili z \mathbf{A} v první fázi \mathbf{A}_1 a ve druhé fázi \mathbf{A}_2 :

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 2 \\ -1 & 5 & -4 \end{pmatrix}, \quad \mathbf{A}_1 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 6 & -3 \end{pmatrix}, \quad \mathbf{A}_2 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{pmatrix}.$$

Protože \mathbf{A}_2 je horní trojúhelníková matice, položíme $\mathbf{U} = \mathbf{A}_2$. Zbývá ukázat, jak vypadá dolní trojúhelníková matice \mathbf{L} .

První fázi zapíšeme jako násobení maticí \mathbf{M}_1 , kterou sestavíme z multiplikátorů $m_{21} = -2$ a $m_{31} = 1$:

$$\mathbf{A}_1 = \mathbf{M}_1\mathbf{A}, \quad \text{kde } \mathbf{M}_1 = \begin{pmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Podobně druhou fázi zapíšeme jako násobení maticí \mathbf{M}_2 , která je určena multiplikátorem $m_{32} = -3$:

$$\mathbf{A}_2 = \mathbf{M}_2 \mathbf{A}_1, \text{ kde } \mathbf{M}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & m_{32} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{pmatrix}.$$

Dosazením dostaneme

$$\mathbf{U} = \mathbf{A}_2 = \mathbf{M}_2 \mathbf{M}_1 \mathbf{A}$$

a odtud

$$\mathbf{A} = \mathbf{M}_1^{-1} \mathbf{M}_2^{-1} \mathbf{U}.$$

Zdá se, že matice \mathbf{L} by mohl být součin $\mathbf{M}_1^{-1} \mathbf{M}_2^{-1}$. Musíme však ještě prověřit, zda je to opravdu dolní trojúhelníková matice. Nejdříve si všimněme, že inverzní matice \mathbf{M}_1^{-1} a \mathbf{M}_2^{-1} mají tvar

$$\mathbf{M}_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -m_{21} & 1 & 0 \\ -m_{31} & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix},$$

$$\mathbf{M}_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -m_{32} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix}$$

(stačí ověřit, že platí $\mathbf{M}_1^{-1} \mathbf{M}_1 = \mathbf{I}$ a $\mathbf{M}_2^{-1} \mathbf{M}_2 = \mathbf{I}$). Vynásobením dostaneme

$$\mathbf{M}_1^{-1} \mathbf{M}_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -m_{21} & 1 & 0 \\ -m_{31} & -m_{32} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 3 & 1 \end{pmatrix}.$$

Proto můžeme položit $\mathbf{L} = \mathbf{M}_1^{-1} \mathbf{M}_2^{-1}$ a platí

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{pmatrix}.$$

Uvedený postup lze zobecnit pro matici libovolného řádu n .

Věta 3.3.1.

Nechť \mathbf{A} je matice řádu n , kterou lze dopředným chodem GEM bez výběru hlavního prvku upravit na horní trojúhelníkovou matici \mathbf{U} . Necht' m_{ik} , $k = 1, \dots, n-1$, $i = k+1, \dots, n$ jsou multiplikátory k -té fáze, z nichž vytvoříme dolní trojúhelníkovou matici $\mathbf{L} = (l_{ik})$ tak, že $l_{ik} = -m_{ik}$, $i > k$, $l_{ii} = 1$ a $l_{ik} = 0$, $i < k$. Potom platí

$$\mathbf{A} = \mathbf{LU}.$$

Z předchozího odstavce víme, že dopředný chod GEM *bez* výběru hlavního prvku nelze provést pro každou matici \mathbf{A} . Obecný tvar LU-rozkladu proto obsahuje ještě permutační matici, která popisuje přehazování řádků, k nimž dochází při výběru hlavního prvku.

Věta 3.3.2.

Nechť \mathbf{A} je regulární matice řádu n . Pak existují dolní trojúhelníková matice \mathbf{L} , horní trojúhelníková matice \mathbf{U} a permutační matice \mathbf{P} řádu n takové, že

$$\mathbf{PA} = \mathbf{LU}. \quad (3.3.1)$$

Důkaz: Princip důkazu je následující. Přehazování řádků v průběhu dopředného chodu GEM se zaznamená v permutační matici \mathbf{P} . Jestliže se vrátíme na začátek a vytvoříme \mathbf{PA} (tj. přehodíme řádky matice \mathbf{A} tak, jak to vyžaduje průběh výpočtu), pak můžeme pro tuto matici najít její LU-rozklad podle věty 3.3.1. \square

Při praktickém výpočtu LU-rozkladu (3.3.1) postupujeme například takto:

- vytvoříme pomocné matice $\tilde{\mathbf{U}} = \mathbf{A}$, $\tilde{\mathbf{P}} = \mathbf{I}$ a $\tilde{\mathbf{L}} = \mathbf{I}$;
- v matici $\tilde{\mathbf{U}}$ provádíme dopředný chod GEM s výběrem hlavního prvku;
- v matici $\tilde{\mathbf{P}}$ přehazujeme řádky stejně jako v matici $\tilde{\mathbf{U}}$;
- do matice $\tilde{\mathbf{L}}$ zapíšeme v každé fázi multiplikátory (s opačnými znaménky) a při přehození řádků v $\tilde{\mathbf{U}}$ přehodíme v $\tilde{\mathbf{L}}$ řádky i sloupce;
- nakonec dostáváme $\mathbf{P} = \tilde{\mathbf{P}}$, $\mathbf{L} = \tilde{\mathbf{L}}$ a $\mathbf{U} = \tilde{\mathbf{U}}$.

Příklad 3.3.1. Vypočtete LU-rozklad (3.3.1) pro matici

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 2 \\ -1 & 5 & -4 \end{pmatrix}.$$

Řešení:

$$\tilde{\mathbf{U}} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 2 \\ -1 & 5 & -4 \end{pmatrix}, \quad \tilde{\mathbf{P}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \tilde{\mathbf{L}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Výběr hlavního prvku v první fázi:

$$\tilde{\mathbf{U}} = \begin{pmatrix} 2 & 4 & 2 \\ 1 & 1 & 1 \\ -1 & 5 & -4 \end{pmatrix}, \quad \tilde{\mathbf{P}} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \tilde{\mathbf{L}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Eliminace v první fázi s multiplikátory $m_{21} = -\frac{1}{2}$ a $m_{31} = \frac{1}{2}$:

$$\tilde{\mathbf{U}} = \begin{pmatrix} 2 & 4 & 2 \\ 0 & -1 & 0 \\ 0 & 7 & -3 \end{pmatrix}, \quad \tilde{\mathbf{P}} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \tilde{\mathbf{L}} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{pmatrix}.$$

Výběr hlavního prvku ve druhé fázi:

$$\tilde{\mathbf{U}} = \begin{pmatrix} 2 & 4 & 2 \\ 0 & 7 & -3 \\ 0 & -1 & 0 \end{pmatrix}, \quad \tilde{\mathbf{P}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \tilde{\mathbf{L}} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{pmatrix}.$$

Eliminace ve druhé fázi s multiplikátorem $m_{32} = \frac{1}{7}$:

$$\tilde{\mathbf{U}} = \begin{pmatrix} 2 & 4 & 2 \\ 0 & 7 & -3 \\ 0 & 0 & -\frac{3}{7} \end{pmatrix}, \quad \tilde{\mathbf{P}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \tilde{\mathbf{L}} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{2} & -\frac{1}{7} & 1 \end{pmatrix}.$$

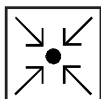
Výsledek je $\mathbf{P} = \tilde{\mathbf{P}}$, $\mathbf{L} = \tilde{\mathbf{L}}$ a $\mathbf{U} = \tilde{\mathbf{U}}$.



Kontrolní otázky

Otázka 1. Jak souvisí LU-rozklad s GEM?

Otázka 2. Jak se provádí výpočet LU-rozkladu?



Úlohy k samostatnému řešení

1. Pro matici

$$\mathbf{A} = \begin{pmatrix} -1 & -3 & 2 \\ -6 & -19 & 10 \\ 3 & 9 & -5 \end{pmatrix}.$$

vypočtete LU-rozklad $\mathbf{A} = \mathbf{LU}$.

2. Pro matici z předchozí úlohy vypočtete LU-rozklad $\mathbf{PA} = \mathbf{LU}$.

3. Jaká je výpočetní náročnost LU-rozkladu?



Výsledky úloh k samostatnému řešení

1.

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 \\ 6 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix}, \quad \mathbf{U} = \begin{pmatrix} -1 & -3 & 2 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

2.

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \mathbf{L} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{6} & -\frac{1}{3} & 1 \end{pmatrix}, \quad \mathbf{U} = \begin{pmatrix} -6 & -19 & 10 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}.$$

3. Výpočetní náročnost je zhruba stejná jako u dopředného chodu GEM. Objem výpočtu se zmenšil o úpravu vektoru pravé strany, což představuje $\mathcal{O}(n^2)$ operací. Platí $\mathcal{O}(\frac{2}{3}n^3) - \mathcal{O}(n^2) = \mathcal{O}(\frac{2}{3}n^3)$.

3.4. Použití LU-rozkladu



Cíle

LU-rozklad použijeme pro řešení základních úloh lineární algebry.



Předpokládané znalosti

Výpočet LU-rozkladu.



Výklad

3.4.1. Řešení soustav lineárních rovnic

Uvažujme soustavu lineárních rovnic

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

s regulární čtvercovou maticí řádu n a předpokládejme, že \mathbf{P} , \mathbf{L} a \mathbf{U} jsou matice, které tvoří LU-rozklad $\mathbf{PA} = \mathbf{LU}$. Platí následující ekvivalence:

$$\mathbf{Ax} = \mathbf{b} \quad \Leftrightarrow \quad \mathbf{PAx} = \mathbf{Pb} \quad \Leftrightarrow \quad \mathbf{LUx} = \mathbf{Pb}.$$

Poslední rovnici rozložíme s využitím pomocné proměnné \mathbf{y} na dvě rovnice

$$\mathbf{Ly} = \mathbf{Pb}, \quad \mathbf{Ux} = \mathbf{y}$$

a dostáváme následující algoritmus.

Algoritmus

Vstup: \mathbf{A} , \mathbf{b} .

Krok 1: Vypočti matice \mathbf{P} , \mathbf{L} a \mathbf{U} , které tvoří LU-rozklad $\mathbf{PA} = \mathbf{LU}$.

Krok 2: Vyřeš soustavu lineárních rovnic $\mathbf{Ly} = \mathbf{Pb}$.

Krok 3: Vyřeš soustavu lineárních rovnic $\mathbf{Ux} = \mathbf{y}$.

Výstup: \mathbf{x} .

Protože matice \mathbf{L} a \mathbf{U} jsou trojúhelníkové, stačí k provedení kroků 2 a 3 zhruba $2\mathcal{O}(n^2)$ operací. Krok 1 je podstatně pracnější, vyžaduje totiž $\mathcal{O}(\frac{2}{3}n^3)$ operací. Podrobným rozбором se dá ukázat, že pracnost celého algoritmu je naprosto stejná jako pracnost Gaussovy eliminační metody.

Příklad 3.4.1. Pomocí LU-rozkladu $\mathbf{PA} = \mathbf{LU}$ řešte soustavu

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 2 \\ -1 & 5 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 16 \\ -3 \end{pmatrix}.$$

Řešení: LU-rozklad pro matici této soustavy jsme vypočítali v Příkladu 3.3.1.:

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \mathbf{L} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{2} & -\frac{1}{7} & 1 \end{pmatrix}, \quad \mathbf{U} = \begin{pmatrix} 2 & 4 & 2 \\ 0 & 7 & -3 \\ 0 & 0 & -\frac{3}{7} \end{pmatrix}. \quad (3.4.1)$$

Pro druhý krok algoritmu potřebujeme nachystat pravou stranu

$$\mathbf{Pb} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 6 \\ 16 \\ -3 \end{pmatrix} = \begin{pmatrix} 16 \\ -3 \\ 6 \end{pmatrix}.$$

Máme tedy řešit soustavu

$$\begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{2} & -\frac{1}{7} & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 16 \\ -3 \\ 6 \end{pmatrix}.$$

Odtud postupně vypočítáme $y_1 = 16$, $y_2 = 5$ a $y_3 = -\frac{9}{7}$. Soustava ve třetím

kroku algoritmu má tvar

$$\begin{pmatrix} 2 & 4 & 2 \\ 0 & 7 & -3 \\ 0 & 0 & -\frac{3}{7} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 16 \\ 5 \\ -\frac{9}{7} \end{pmatrix}.$$

Z ní postupně vypočítáme řešení $x_3 = 3$, $x_2 = 2$ a $x_1 = 1$.

3.4.2. Výpočet inverzní matice

Připomeňme, že pro inverzní matici platí $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$. Označíme-li $\mathbf{a}^{(i)}$ i -tý sloupec matice inverzní \mathbf{A}^{-1} a $\mathbf{e}^{(i)}$ i -tý sloupec matice jednotkové \mathbf{I} , pak můžeme uvedenou rovnost zapsat jako $\mathbf{A}(\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(n)}) = (\mathbf{e}^{(1)}, \dots, \mathbf{e}^{(n)})$ a po roznásobení jako

$$(\mathbf{A}\mathbf{a}^{(1)}, \dots, \mathbf{A}\mathbf{a}^{(n)}) = (\mathbf{e}^{(1)}, \dots, \mathbf{e}^{(n)}).$$

Odtud je zřejmé, že musí být splněny soustavy lineárních rovnic

$$\mathbf{A}\mathbf{a}^{(i)} = \mathbf{e}^{(i)}, \quad i = 1, \dots, n.$$

Protože matice je u všech soustav stejná, stačí při jejich řešení vypočítat LU-rozklad jenom jednou.

Algoritmus

Vstup: \mathbf{A} .

Krok 1: Vypočti matice \mathbf{P} , \mathbf{L} a \mathbf{U} , které tvoří LU-rozklad $\mathbf{PA} = \mathbf{LU}$.

Pro $i = 1, \dots, n$ vypočti i -tý sloupec inverzní matice:

Krok 2: Vyřeš soustavu lineárních rovnic $\mathbf{L}\mathbf{y} = \mathbf{P}\mathbf{b}$, kde $\mathbf{b} = \mathbf{e}^{(i)}$;

Krok 3: Vyřeš soustavu lineárních rovnic $\mathbf{U}\mathbf{x} = \mathbf{y}$ a polož $\mathbf{a}^{(i)} = \mathbf{x}$.

Výstup: $\mathbf{A}^{-1} = (\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(n)})$.

Výpočetní náročnost je $\mathcal{O}(\frac{2}{3}n^3)$ v kroku 1 a n -krát $2\mathcal{O}(n^2)$ v krocích 2 a 3. Celkem tedy vyžaduje algoritmus $\mathcal{O}(\frac{8}{3}n^3)$ operací, což je zhruba čtyřikrát víc než při řešení soustavy lineárních rovnic.

Příklad 3.4.2. Vypočítejte inverzní matici \mathbf{A}^{-1} k matici

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 2 \\ -1 & 5 & -4 \end{pmatrix}.$$

Řešení: LU-rozklad tvoří matice (3.4.1). Podle algoritmu dále počítáme postupně jednotlivé sloupce inverzní matice.

Pro $i = 1$ v kroku druhém řešíme $\mathbf{Ly} = \mathbf{P} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$:

$$\begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{2} & -\frac{1}{7} & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \implies \mathbf{y} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Pro $i = 1$ v kroku třetím řešíme $\mathbf{Ux} = \mathbf{y}$:

$$\begin{pmatrix} 2 & 4 & 2 \\ 0 & 7 & -3 \\ 0 & 0 & -\frac{3}{7} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \implies \mathbf{x} = \begin{pmatrix} \frac{13}{3} \\ -1 \\ -\frac{7}{3} \end{pmatrix} = \mathbf{a}^{(1)}.$$

Pro $i = 2$ v kroku druhém řešíme $\mathbf{Ly} = \mathbf{P} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$:

$$\begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{2} & -\frac{1}{7} & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \implies \mathbf{y} = \begin{pmatrix} 1 \\ \frac{1}{2} \\ -\frac{3}{7} \end{pmatrix}.$$

Pro $i = 2$ v kroku třetím řešíme $\mathbf{Ux} = \mathbf{y}$:

$$\begin{pmatrix} 2 & 4 & 2 \\ 0 & 7 & -3 \\ 0 & 0 & -\frac{3}{7} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{1}{2} \\ -\frac{3}{7} \end{pmatrix} \implies \mathbf{x} = \begin{pmatrix} -\frac{3}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix} = \mathbf{a}^{(2)}.$$

Pro $i = 3$ v kroku druhém řešíme $\mathbf{Ly} = \mathbf{P} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$:

$$\begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{2} & -\frac{1}{7} & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \implies \mathbf{y} = \begin{pmatrix} 0 \\ 1 \\ \frac{1}{7} \end{pmatrix}.$$

Pro $i = 3$ v kroku třetím řešíme $\mathbf{Ux} = \mathbf{y}$:

$$\begin{pmatrix} 2 & 4 & 2 \\ 0 & 7 & -3 \\ 0 & 0 & -\frac{3}{7} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ \frac{1}{7} \end{pmatrix} \implies \mathbf{x} = \begin{pmatrix} \frac{1}{3} \\ 0 \\ -\frac{1}{3} \end{pmatrix} = \mathbf{a}^{(3)}.$$

Vypočítali jsme inverzní matici:

$$\mathbf{A}^{-1} = (\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \mathbf{a}^{(3)}) = \begin{pmatrix} \frac{13}{3} & -\frac{3}{2} & \frac{1}{3} \\ -1 & \frac{1}{2} & 0 \\ -\frac{7}{3} & 1 & -\frac{1}{3} \end{pmatrix}.$$

3.4.3. Výpočet determinantu

Použijeme jedno ze základních pravidel pro počítání s determinanty, které říká, že determinant ze součinu (čtvercových) matic se rovná součinu jejich determinantů.

Jestliže matice \mathbf{P} , \mathbf{L} a \mathbf{U} tvoří LU-rozklad $\mathbf{PA} = \mathbf{LU}$, pak můžeme psát

$$\det \mathbf{A} = (\det \mathbf{P})^{-1} \cdot \det \mathbf{L} \cdot \det \mathbf{U}.$$

Determinanty trojúhelníkových matic vypočítáme snadno jako součiny jejich diagonálních prvků. Determinant permutační matice je $+1$, resp. -1 podle toho, jestli vznikla z jednotkové matice sudým, resp. lichým počtem přehození řádků.

Příklad 3.4.3. Vypočítejte determinat matice \mathbf{A} z příkladu 3.3.1.

Řešení: Pomocí výsledku příkladu 3.3.1. dostáváme

$$\begin{aligned}\det \mathbf{L} &= 1 \cdot 1 \cdot 1 = 1, \\ \det \mathbf{U} &= 2 \cdot 7 \cdot \frac{-3}{7} = -6.\end{aligned}$$

Protože při výpočtu LU-rozkladu došlo ke dvěma záměnám řádků, bude $\det \mathbf{P} =$

1. Celkem je $\det \mathbf{A} = 1 \cdot 1 \cdot (-6) = -6$.

Kontrolní otázky

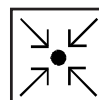


Otázka 1. Jak se pomocí LU-rozkladu řeší soustava lineárních rovnic?

Otázka 2. Jak se pomocí LU-rozkladu počítá inverzní matice?

Otázka 3. Jak se pomocí LU-rozkladu počítá determinant?

Úlohy k samostatnému řešení



1. Soustavu lineárních rovnic

$$\begin{aligned}-x_1 - 3x_2 + 2x_3 &= -9, \\ -6x_1 - 19x_2 + 10x_3 &= -59, \\ 3x_1 + 9x_2 - 5x_3 &= 28\end{aligned}$$

řešte pomocí LU-rozkladu $\mathbf{A} = \mathbf{LU}$.

2. Soustavu lineárních rovnic z první úlohy řešte pomocí LU-rozkladu $\mathbf{PA} = \mathbf{LU}$.

3. Pro matici

$$\mathbf{A} = \begin{pmatrix} -1 & -3 & 2 \\ -6 & -19 & 10 \\ 3 & 9 & -5 \end{pmatrix}$$

vypočtete inverzní matici pomocí LU-rozkladu $\mathbf{A} = \mathbf{LU}$.

4. K předchozí matici vypočtete inverzní matici pomocí LU-rozkladu $\mathbf{PA} = \mathbf{LU}$.

5. Vypočtete determinant matice z 3. úlohy pomocí LU-rozkladu $\mathbf{A} = \mathbf{LU}$.
6. Vypočtete determinant matice z 3. úlohy pomocí LU-rozkladu $\mathbf{PA} = \mathbf{LU}$.
7. Kolik operací je potřeba při výpočtu determinantu matice pomocí LU-rozkladu?

**Výsledky úloh k samostatnému řešení**

1. Dostaneme $\mathbf{y} = (-9, -5, 1)^\top$ a $\mathbf{x} = (2, 3, 1)^\top$.
2. Dostaneme $\mathbf{y} = (-59, -\frac{3}{2}, \frac{1}{3})^\top$ a $\mathbf{x} = (2, 3, 1)^\top$.
3. Pro $i = 1$ je $\mathbf{y} = (1, 6, -3)^\top$, pro $i = 2$ je $\mathbf{y} = (0, 1, 0)^\top$ a pro $i = 3$ je $\mathbf{y} = (0, 0, 1)^\top$. Inverzní matice má tvar

$$\mathbf{A}^{-1} = \begin{pmatrix} 5 & 3 & 8 \\ 0 & -1 & -2 \\ 3 & 0 & 1 \end{pmatrix}.$$

4. Pro $i = 1$ je $\mathbf{y} = (0, 0, 1)^\top$, pro $i = 2$ je $\mathbf{y} = (1, \frac{1}{2}, 0)^\top$ a pro $i = 3$ je $\mathbf{y} = (0, 1, \frac{1}{3})^\top$. Inverzní matice je stejná jako v předchozí úloze.
5. $\det \mathbf{L} = 1$, $\det \mathbf{U} = 1$ a $\det \mathbf{A} = 1$.
6. $\det \mathbf{P} = 1$, $\det \mathbf{L} = 1$, $\det \mathbf{U} = 1$ a $\det \mathbf{A} = 1$.
7. Zhruba $\mathcal{O}(\frac{2}{3}n^3)$ operací.

3.5. Maticové normy a podmíněnost matic

Cíle



V první kapitole jsme v definici 1.2.2. zavedli číslo podmíněnosti úlohy, které vyjadřuje citlivost na různé typy poruch. Stejný smysl má *číslo podmíněnosti matice*.

Předpokládané znalosti



Maticové operace. Absolutní hodnota.

Výklad



Při posuzování poruch u maticových výpočtů je potřeba nějakým způsobem měřit „velikost“ matic a vektorů. Pro tyto účely se zavádí *norma matice*, která je zobecněním funkce absolutní hodnota pro reálná čísla.

Definice 3.5.1.

Norma matice je zobrazení, které každé matici $\mathbf{A} = (a_{ij})$ typu $m \times n$ přiřadí číslo $\|\mathbf{A}\|$ tak, že platí:

- (i) $\|\mathbf{A}\| \geq 0$ a přitom $\|\mathbf{A}\| = 0$, právě když \mathbf{A} je matice nulová;
- (ii) $\|\alpha \mathbf{A}\| = |\alpha| \cdot \|\mathbf{A}\|$ pro každé reálné číslo α ;
- (iii) $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$ pro každou matici \mathbf{B} stejného typu jako je matice \mathbf{A} .

Základní maticové normy jsou:

- *řádková norma*: $\|\mathbf{A}\|_R = \max_{i=1,\dots,m} \sum_{j=1}^n |a_{ij}|$;
- *sloupcová norma*: $\|\mathbf{A}\|_S = \max_{j=1,\dots,n} \sum_{i=1}^m |a_{ij}|$;

- *Frobeniova norma*: $\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2}$.

Příklad 3.5.1. Vypočítejte řádkovou, sloupcovou a Frobeniovu normu pro matici

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 2 \\ -1 & 5 & -4 \end{pmatrix}$$

a pro matici inverzní \mathbf{A}^{-1} .

Řešení: Dostáváme

$$\|\mathbf{A}\|_R = \max\{1 + 1 + 1, 2 + 4 + 2, 1 + 5 + 4\} = 10,$$

$$\|\mathbf{A}\|_S = \max\{1 + 2 + 1, 1 + 4 + 5, 1 + 2 + 4\} = 10,$$

$$\|\mathbf{A}\|_F = \sqrt{1 + 1 + 1 + 4 + 16 + 4 + 1 + 25 + 16} = \sqrt{69} \doteq 8.3066.$$

Matici inverzní \mathbf{A}^{-1} jsem vypočítali v příkladu 3.4.2. Pomocí tohoto výsledku dostaneme

$$\|\mathbf{A}^{-1}\|_R = \frac{37}{6}, \quad \|\mathbf{A}^{-1}\|_S = \frac{23}{3}, \quad \|\mathbf{A}^{-1}\|_F \doteq 5.38.$$

Věta 3.5.1.

Nechť \mathbf{A} je matice typu $m \times n$ a \mathbf{B} je matice typu $n \times p$. Pro řádkovou, sloupcovou a Frobeniovu normu platí:

$$\|\mathbf{AB}\| \leq \|\mathbf{A}\| \|\mathbf{B}\|. \quad (3.5.1)$$

Důkaz: Platnost tvrzení ukážeme pouze pro řádkovou normu, ostatní případy ponecháme jako cvičení. Prvky matice součinu $\mathbf{C} = \mathbf{AB}$ jsou určeny předpisem $c_{ij} = \sum_{k=1}^n a_{ik}b_{kj}$. Proto

$$\|\mathbf{AB}\|_R = \max_{i=1,\dots,m} \sum_{j=1}^p |c_{ij}| = \max_{i=1,\dots,m} \sum_{j=1}^p \left| \sum_{k=1}^n a_{ik}b_{kj} \right| \leq$$

$$\begin{aligned}
&\leq \max_{i=1,\dots,m} \sum_{j=1}^p \sum_{k=1}^n |a_{ik}| |b_{kj}| = \max_{i=1,\dots,m} \sum_{k=1}^n |a_{ik}| \sum_{j=1}^p |b_{kj}| \\
&\leq \max_{i=1,\dots,m} \sum_{k=1}^n |a_{ik}| \max_{l=1,\dots,n} \sum_{j=1}^p |b_{lj}| = \|A\|_R \|B\|_R.
\end{aligned}$$

□

Definice 3.5.2.

Číslo podmíněnosti regulární čtvercové matice \mathbf{A} je definováno předpisem

$$\kappa(\mathbf{A}) = \|\mathbf{A}\| \|\mathbf{A}^{-1}\|.$$

Příklad 3.5.2. Vypočítejte číslo podmíněnosti matice z příkladu 3.5.1. pomocí řádkové, sloupcové a Frobeniovy normy.

Řešení: S využitím výsledků příkladu 3.5.1. dostaneme:

$$\kappa_R(\mathbf{A}) \doteq 61.67, \quad \kappa_S(\mathbf{A}) \doteq 76.67, \quad \kappa_F(\mathbf{A}) \doteq 44.69.$$

Věta 3.5.2.

Nechť \mathbf{A} je regulární čtvercová matice řádu n a nechť \mathbf{b} a \mathbf{x} jsou nenulové n -složkové vektory takové, že platí:

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

Dále nechť $\tilde{\mathbf{b}}$ a $\tilde{\mathbf{x}}$ jsou n -složkové vektory takové, že platí

$$\mathbf{A}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}.$$

Potom

$$\frac{\|\mathbf{x} - \tilde{\mathbf{x}}\|}{\|\mathbf{x}\|} \leq \kappa(\mathbf{A}) \frac{\|\mathbf{b} - \tilde{\mathbf{b}}\|}{\|\mathbf{b}\|}. \quad (3.5.2)$$

Důkaz: Zřejmě platí

$$\mathbf{b} = \mathbf{Ax} \quad \text{resp.} \quad \mathbf{x} - \tilde{\mathbf{x}} = \mathbf{A}^{-1}(\mathbf{b} - \tilde{\mathbf{b}})$$

a pomocí (3.5.1) odtud dostaneme

$$\|\mathbf{x}\|^{-1} \leq \|\mathbf{A}\| \|\mathbf{b}\|^{-1} \quad \text{resp.} \quad \|\mathbf{x} - \tilde{\mathbf{x}}\| \leq \|\mathbf{A}^{-1}\| \|\mathbf{b} - \tilde{\mathbf{b}}\|.$$

Vynásobením těchto nerovností vznikne tvrzení (3.5.2). \square

Nerovnost (3.5.2) říká, že při velké hodnotě $\kappa(\mathbf{A})$ může malá porucha ve vektoru \mathbf{b} vyvolat velkou změnu v řešení. Výpočty s maticí, která má velké číslo podmíněnosti, jsou zpravidla znehodnoceny kumulací zaokrouhlovacích chyb, jak ukazuje následující příklad.

Příklad 3.5.3. Vypočteme čísla podmíněnosti *Hilbertovy* matice

$$\mathbf{A} = \begin{pmatrix} 1 & 1/2 & 1/3 & \dots \\ 1/2 & 1/3 & 1/4 & \dots \\ 1/3 & 1/4 & 1/5 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

řádů $n = 5, 10, 15, 20, 25$ a pokusíme se vypočítat inverzní matici.

Řešení: Čísla podmíněnosti jsou zaznamenána v tabulce 3.5.1. Poslední sloupec

Tabulka 3.5.1: Podmíněnost Hilbertovy matice \mathbf{A} .

n	$\kappa(\mathbf{A})$	$\ \mathbf{A}\mathbf{A}^{-1} - \mathbf{I}\ $
5	4.8×10^5	1.4×10^{-11}
10	1.6×10^{13}	3.3×10^{-3}
15	1.1×10^{18}	2.8×10^3
20	2.5×10^{28}	2.6×10^{11}
25	1.0×10^{36}	1.3×10^{19}

tabulky ukazuje, jak se (na počítači) podařilo vypočítat inverzní matice. Je vidět, že pro řád $n = 15$ a vyšší jsou výsledky naprosto nesmyslné.

Kontrolní otázky

Otázka 1. Jak se definuje norma matice?

Otázka 2. Co vyjadřuje číslo podmíněnosti matice?

Úlohy k samostatnému řešení

1. Pro matici

$$\mathbf{A} = \begin{pmatrix} -1 & -3 & 2 \\ -6 & -19 & 10 \\ 3 & 9 & -5 \end{pmatrix}.$$

vypočtete číslo podmíněnosti pomocí řádkové, sloupcové a Frobeniovy normy.

Výsledky úloh k samostatnému řešení

1. $\|\mathbf{A}\|_R = 35$, $\|\mathbf{A}^{-1}\|_R = 16$, $\kappa_R(\mathbf{A}) = 560$; $\|\mathbf{A}\|_S = 31$, $\|\mathbf{A}^{-1}\|_S = 11$, $\kappa_S(\mathbf{A}) = 341$; $\|\mathbf{A}\|_F = 25.02$, $\|\mathbf{A}^{-1}\|_F = 10.63$, $\kappa_F(\mathbf{A}) = 265.96$.

Shrnutí lekce

V kapitole jsme se zabývali Gaussovou eliminační metodou vyjádřenou na maticové úrovni jako LU-rozklad. Dále jsme ukázali, že výpočty mohou být podstatně ovlivněny nahromaděním zaokrouhlovacích chyb, což lze rozpoznat pomocí čísla podmíněnosti.

4. SLR – iterační metody



Průvodce studiem

Iterační metody umožňují řešit soustavy lineárních rovnic pomocí postupného přibližování k přesnému řešení. Počítá se posloupnost vektorů aproximací $\{\mathbf{x}^{(k)}\}$ taková, že

$$\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}, \quad \text{kde } \mathbf{x} \text{ je řešením } \mathbf{Ax} = \mathbf{b}.$$

Přesné řešení tedy dostaneme v limitě, tj. formálně po nekonečném počtu kroků. Výhody iteračních metod jsou tyto:

- V každé iteraci známe aproximaci řešení $\mathbf{x}^{(k)}$. Pokud je tato aproximace dostatečně přesná, pak výpočet ukončíme.
- V každé iteraci je nejpracnější operací násobení matice a vektoru. Jedná se o operaci, která je algoritmicky podstatně jednodušší než GEM a lze ji snadno provést i pro rozsáhlé řídké matice, tj. pro matice s velkým počtem (neuložených) nulových prvků.
- Iterační metody jsou méně citlivé na zaokrouhlovací chyby než metody přímé. Na každou iteraci můžeme nahlížet jako na počáteční. Zaokrouhlovací chyby z předchozích iterací proto vymizí, pokud v dalším výpočtu dojde ke konvergenci. Některé speciální iterační metody byly navrženy pro zpřesnění výsledků vypočítaných pomocí přímých metod.

Zhruba platí následující dělení: přímé metody se používají, je-li matice soustavy malá ($1 \leq n \leq 10000$), plná a dobře podmíněná; iterační metody se používají pro velké soustavy ($n > 10000$) s řídkou maticí.

4.1. Příklad iteračního výpočtu

Cíle



Uvedeme příklady iteračního řešení soustav lineárních rovnic a ukážeme na nich, že výpočet může konvergovat, ale i divergovat.

Předpokládané znalosti



Řešení soustav lineárních rovnic. Provádění rekurentních výpočtů.

Výklad



Uvažujme soustavu lineárních rovnic

$$\begin{aligned} 11x_1 + 2x_2 + x_3 &= 15, \\ x_1 + 10x_2 + 2x_3 &= 16, \quad \text{resp.} \\ 2x_1 + 3x_2 - 8x_3 &= 1, \end{aligned} \quad \begin{pmatrix} 11 & 2 & 1 \\ 1 & 10 & 2 \\ 2 & 3 & -8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 15 \\ 16 \\ 1 \end{pmatrix}. \quad (4.1.1)$$

Soustavu nejdříve převedeme na tvar vhodný pro výpočet iterací, tzv. *iterační tvar*. Provádí se to tak, že z každé rovnice vyjádříme jednu neznámou. Například:

$$\begin{aligned} x_1 &= \frac{1}{11}(15 - 2x_2 - x_3), \\ x_2 &= \frac{1}{10}(16 - x_1 - 2x_3), \\ x_3 &= \frac{1}{8}(-1 + 2x_1 + 3x_2), \end{aligned} \quad (4.1.2)$$

tj.

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & -\frac{2}{11} & -\frac{1}{11} \\ -\frac{1}{10} & 0 & -\frac{1}{5} \\ \frac{1}{4} & \frac{3}{8} & 0 \end{pmatrix}}_{\mathbf{C}_J} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \underbrace{\begin{pmatrix} \frac{15}{11} \\ \frac{8}{5} \\ -\frac{1}{8} \end{pmatrix}}_{\mathbf{d}_J}.$$

Jiná možnost:

$$\begin{aligned}x_1 &= 15 - 10x_1 - 2x_2 - x_3, \\x_2 &= 16 - x_1 - 9x_2 - 2x_3, \\x_3 &= -1 + 2x_1 + 3x_2 - 7x_3,\end{aligned}\tag{4.1.3}$$

tj.

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \underbrace{\begin{pmatrix} -10 & -2 & -1 \\ -1 & -9 & -2 \\ 2 & 3 & -7 \end{pmatrix}}_{\mathbf{C}_B} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \underbrace{\begin{pmatrix} 15 \\ 16 \\ -1 \end{pmatrix}}_{\mathbf{d}_B}.$$

Takových převodů existuje zřejmě nekonečně mnoho, ale jenom některé povedou ke konvergentnímu výpočtu.

Z rovnic (4.1.2) a (4.1.3) dostaneme rekurentní vzorce připsáním iteračního indexu $k + 1$ k neznámým na levé straně a k k neznámým na pravé straně.

Dostáváme

$$\left. \begin{aligned}x_1^{(k+1)} &= \frac{1}{11}(15 - 2x_2^{(k)} - x_3^{(k)}), \\x_2^{(k+1)} &= \frac{1}{10}(16 - x_1^{(k)} - 2x_3^{(k)}), \\x_3^{(k+1)} &= \frac{1}{8}(-1 + 2x_1^{(k)} + 3x_2^{(k)}),\end{aligned} \right\} \quad \text{tj. } \mathbf{x}^{(k+1)} = \mathbf{C}_J \mathbf{x}^{(k)} + \mathbf{d}_J, \tag{4.1.4}$$

a

$$\left. \begin{aligned}x_1^{(k+1)} &= 15 - 10x_1^{(k)} - 2x_2^{(k)} - x_3^{(k)}, \\x_2^{(k+1)} &= 16 - x_1^{(k)} - 9x_2^{(k)} - 2x_3^{(k)}, \\x_3^{(k+1)} &= -1 + 2x_1^{(k)} + 3x_2^{(k)} - 7x_3^{(k)},\end{aligned} \right\} \quad \text{tj. } \mathbf{x}^{(k+1)} = \mathbf{C}_B \mathbf{x}^{(k)} + \mathbf{d}_B. \tag{4.1.5}$$

Nyní zvolíme počáteční aproximaci $\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, x_3^{(0)})^\top$, např. $\mathbf{x}^{(0)} = (0, 0, 0)^\top$.

Tyto hodnoty dosadíme do pravé strany rekurentních vzorců (4.1.4) a dostaneme

$$\mathbf{x}^{(1)} = \left(\frac{15}{11}, \frac{16}{10}, -\frac{1}{8} \right)^\top.$$

Jestliže takto pokračujeme dále, dostáváme

$$\mathbf{x}^{(2)} = (1.0841, 1.4886, 0.81591)^\top,$$

$$\mathbf{x}^{(3)} = (1.0188, 1.3284, 0.70426)^\top,$$

.....

Provedeme-li několik dalších iterací, zjistíme, že se číslice na prvních desetinných místech začnou po chvíli opakovat. Dostaneme přitom vektor

$$\mathbf{x} = (1.0564, 1.3642, 0.65069)^\top,$$

o němž se můžeme domnívat, že je aproximací přesného řešení soustavy (4.1.1).

Jestliže analogicky počítáme podle rekurentních vzorců (4.1.5), dostaneme

$$\mathbf{x}^{(1)} = (15, 16, -1)^\top,$$

$$\mathbf{x}^{(2)} = (-166, -141, 84)^\top,$$

$$\mathbf{x}^{(3)} = (1873, 1283, -1344)^\top,$$

.....

Zde žádnou tendenci ke konvergenci nevidíme a je proto pravděpodobné, že posloupnost iterací diverguje.

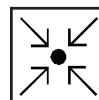
Kontrolní otázky



Otázka 1. V čem spočívá základní rozdíl mezi přímými a iteračními metodami?

Otázka 2. Jaké jsou výhody a nevýhody přímých a iteračních metod?

Úlohy k samostatnému řešení



1. Pro soustavu lineárních rovnic

$$-x_1 - 3x_2 + 2x_3 = -9,$$

$$-6x_1 - 19x_2 + 10x_3 = -59,$$

$$3x_1 + 9x_2 - 5x_3 = 28$$

navrhněte dva iterační tvary pomocí postupů z odstavce 4.1.

2. U kterého z navržených iteračních tvarů výpočet konverguje?



Výsledky úloh k samostatnému řešení

1. Rekurentní vzorce pro první iterační tvar:

$$\begin{aligned}x_1^{(k+1)} &= 9 - 3x_2^{(k)} + 2x_3^{(k)}, \\x_2^{(k+1)} &= \frac{1}{19}(59 - 6x_1^{(k)} + 10x_3^{(k)}), \\x_3^{(k+1)} &= \frac{1}{5}(-28 + 3x_1^{(k)} + 9x_2^{(k)}); \end{aligned}$$

rekurentní vzorce pro druhý iterační tvar:

$$\begin{aligned}x_1^{(k+1)} &= 9 - 3x_2^{(k)} + 2x_3^{(k)}, \\x_2^{(k+1)} &= 59 - 6x_1^{(k)} - 18x_2^{(k)} + 10x_3^{(k)}, \\x_3^{(k+1)} &= -28 + 3x_1^{(k)} + 9x_2^{(k)} - 4x_3^{(k)}.\end{aligned}$$

2. Jestliže zkusíme výpočet provést, zjistíme, že dochází k divergenci v obou případech. Rozpoznáním konvergentního výpočtu z vlastností matice soustavy se budeme zabývat v dalších odstavcích.

4.2. Obecné iterační metody

Cíle

Ukážeme obecný postup pro iterační řešení soustav lineárních rovnic a uvedeme jeho dvě základní varianty.



Předpokládané znalosti

Příklady iteračního řešení soustav lineárních rovnic.



Výklad

Uvažujme soustavu lineárních rovnic



$$\mathbf{Ax} = \mathbf{b} \quad (4.2.1)$$

s regulární čtvercovou maticí $\mathbf{A} = (a_{ij})$ řádu n , vektorem pravé strany $\mathbf{b} = (b_i)$ a vektorem neznámých $\mathbf{x} = (x_i)$. Soustavu (4.2.1) přepíšeme na ekvivalentní soustavu v *iteračním tvaru*

$$\mathbf{x} = \mathbf{Cx} + \mathbf{d}, \quad (4.2.2)$$

kde \mathbf{C} je *iterační matice* řádu n a \mathbf{d} je sloupcový vektor. Musí přitom platit, že rovnice (4.2.1) a (4.2.2) mají stejné řešení.

Nechť $\mathbf{x}^{(0)}$ je daná počáteční aproximace. Iterační výpočet provádíme podle rekurentního vzorce

$$\mathbf{x}^{(k+1)} = \mathbf{Cx}^{(k)} + \mathbf{d}, \quad k = 0, 1, 2, \dots \quad (4.2.3)$$

Jestliže posloupnost vektorů $\{\mathbf{x}^k\}$ konverguje k vektoru \mathbf{x} , pak limitním přechodem v (4.2.3) dostaneme, že \mathbf{x} je řešením rovnice (4.2.2) a také (4.2.1).

Jak uvidíme později, volba počáteční aproximace neovlivní konvergenci, takže vektor $\mathbf{x}^{(0)}$ můžeme zvolit libovolně. Výpočet ukončíme, jestliže dvě poslední

aproximace se od sebe liší ne více, než kolik je požadovaná přesnost, tj. jestliže je splněno ukončovací kritérium

$$\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| \leq \epsilon, \quad (4.2.4)$$

kde $\epsilon > 0$ je dané malé číslo a $\|\cdot\|$ je vhodná norma. V našich příkladech použijeme ukončovací kritérium s řádkovou normou.

Algoritmus (Iterační řešení SLR)

Vstup: \mathbf{C} , \mathbf{d} , $\mathbf{x}^{(0)}$, ϵ .

Opakuj

$$\mathbf{x}^{(k+1)} := \mathbf{C}\mathbf{x}^{(k)} + \mathbf{d};$$

dokud $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| > \epsilon$.

Výstup: $\mathbf{x}^{(k)}$.

4.2.1. Jacobiova metoda

Jacobiovu metodu jsme si již ukázali při řešení soustavy (4.1.1) v odstavci 4.1. Jsou to rekurentní vzorce (4.1.4). Nyní si ji probereme obecně.

Budeme předpokládat, že diagonální prvky matice soustavy (4.2.1) jsou nenulové, tj. $a_{ii} \neq 0$. Z i -té rovnice

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i, \quad i = 1, \dots, n,$$

vyjádříme i -tou neznámou

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

Jacobiova metoda je určena rekurentními vzorci

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)} \right), \quad i = 1, \dots, n, \quad (4.2.5)$$

pro $k = 0, 1, 2, \dots$

Všimněme si ještě, že vzorce (4.2.5) můžeme zapsat v obecném maticovém tvaru (4.2.3), jestliže položíme $\mathbf{C} = \mathbf{C}_J$ a $\mathbf{d} = \mathbf{d}_J$, kde

$$\mathbf{C}_J = \begin{pmatrix} 0 & -\frac{a_{12}}{a_{11}} & -\frac{a_{13}}{a_{11}} & \cdots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & -\frac{a_{23}}{a_{22}} & \cdots & -\frac{a_{2n}}{a_{22}} \\ -\frac{a_{31}}{a_{33}} & -\frac{a_{32}}{a_{33}} & 0 & \cdots & -\frac{a_{3n}}{a_{33}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & -\frac{a_{n3}}{a_{nn}} & \cdots & 0 \end{pmatrix}, \quad \mathbf{d}_J = \begin{pmatrix} \frac{b_1}{a_{11}} \\ \frac{b_2}{a_{22}} \\ \frac{b_3}{a_{33}} \\ \vdots \\ \frac{b_n}{a_{nn}} \end{pmatrix}. \quad (4.2.6)$$

Pomocí aditivního rozkladu matice $\mathbf{A} = (a_{ij})$,

$$\mathbf{A} = \mathbf{L} + \mathbf{D} + \mathbf{U}, \quad (4.2.7)$$

kde $\mathbf{L} = (l_{ij})$, $l_{ij} = a_{ij}$, $i > j$, $l_{ij} = 0$, $i \leq j$, je dolní trojúhelníková část, $\mathbf{D} = (d_{ij})$, $d_{ii} = a_{ii}$, $d_{ij} = 0$, $i \neq j$, je diagonální část a $\mathbf{U} = (u_{ij})$, $u_{ij} = 0$, $i \geq j$, $u_{ij} = a_{ij}$, $i < j$, je horní trojúhelníková část, můžeme stručně psát

$$\mathbf{C}_J = -\mathbf{D}^{-1}(\mathbf{L} + \mathbf{U}), \quad \mathbf{d}_J = \mathbf{D}^{-1}\mathbf{b}.$$

Příklad 4.2.1. Soustavu lineárních rovnic (4.1.1) řešte pomocí Jacobiovy metody s přesností $\epsilon = 10^{-4}$.

Řešení: Výpočet se provádí podle rekurentních vzorců (4.1.4). Začátek výpočtu jsme naznačili v odstavci 4.1. Nyní vše shrneme v tabulce 4.2.1, kde kromě aproximací $\mathbf{x}^{(k)}$ uvádíme řádkové normy $\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|_R$. Naznačme ještě výpočet prvních dvou norem:

$$\|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|_R = \max\{|\frac{15}{11} - 0|, |\frac{16}{10} - 0|, |-\frac{1}{8} - 0|\} = 1.6,$$

$$\|\mathbf{x}^{(2)} - \mathbf{x}^{(1)}\|_R = \max\{|1.0841 - \frac{15}{11}|, |1.4886 - \frac{16}{10}|, |0.8159 + \frac{1}{8}|\} = 0.9409,$$

atd.

Výpočet jsme ukončili po desáté iteraci, protože $\|\mathbf{x}^{(10)} - \mathbf{x}^{(9)}\|_R = 0.00005 \leq 10^{-4}$, a výsledek je $x_1 = 1.0564 \pm 10^{-4}$, $x_2 = 1.3642 \pm 10^{-4}$, $x_3 = 0.6507 \pm 10^{-4}$.

Tabulka 4.2.1: Iterace Jacobiovy metody.

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\ \mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\ _R$
0	0	0	0	—
1	1.3636	1.6000	-0.1250	1.60000
2	1.0841	1.4886	0.8159	0.94091
3	1.0188	1.3284	0.7043	0.16023
4	1.0581	1.3573	0.6279	0.07641
5	1.0598	1.3686	0.6485	0.02064
6	1.0558	1.3643	0.6532	0.00468
7	1.0562	1.3638	0.6506	0.00260
8	1.0565	1.3643	0.6505	0.00048
9	1.0565	1.3643	0.6507	0.00027
10	1.0564	1.3642	0.6507	0.00005

4.2.2. Gauss-Seidelova metoda

Začneme příkladem. Pro řešení soustavy lineárních rovnic (4.1.1) jsme použili Jacobiovu metodu, která je určena rekurentními vzorci (4.1.4). Podle těchto vzorců se počítají v k -té iteraci složky nové aproximace $\mathbf{x}^{(k+1)}$ postupně, tj. nejdříve $x_1^{(k+1)}$ pak $x_2^{(k+1)}$ a nakonec $x_3^{(k+1)}$. Přitom se stále používají složky z předchozí aproximace, tj. $x_1^{(k)}$, $x_2^{(k)}$ a $x_3^{(k)}$. Tento postup můžeme snadno vylepšit. Stačí si uvědomit, že při výpočtu $x_2^{(k+1)}$ můžeme použít přesnější aproximaci $x_1^{(k+1)}$ namísto méně přesné $x_1^{(k)}$. Podobně můžeme při výpočtu $x_3^{(k+1)}$ použít přesnější aproximace $x_1^{(k+1)}$ a $x_2^{(k+1)}$ namísto méně přesných $x_1^{(k)}$ a $x_2^{(k)}$. Původní rekurentní vzorce (4.1.1) se tak změní na tvar

$$\begin{aligned}
 x_1^{(k+1)} &= \frac{1}{11}(15 - 2x_2^{(k)} - x_3^{(k)}), \\
 x_2^{(k+1)} &= \frac{1}{10}(16 - x_1^{(k+1)} - 2x_3^{(k)}), \\
 x_3^{(k+1)} &= \frac{1}{8}(-1 + 2x_1^{(k+1)} + 3x_2^{(k+1)}),
 \end{aligned} \tag{4.2.8}$$

což je Gauss-Seidelova metoda.

Příklad 4.2.2. Soustavu lineárních rovnic (4.1.1) řešte pomocí Gauss-Seidelovy metody s přesností $\epsilon = 10^{-4}$.

Řešení: Výpočet podle vzorců (4.2.8) je zaznamenán v tabulce 4.2.2.

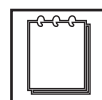
Tabulka 4.2.2: Iterace Gauss-Seidelovy metody.

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\ \mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\ _R$
0	0	0	0	—
1	1.3636	1.4636	0.7648	1.46364
2	1.0280	1.3442	0.6361	0.33564
3	1.0614	1.3666	0.6528	0.03341
4	1.0558	1.3639	0.6504	0.00559
5	1.0565	1.3643	0.6507	0.00073
6	1.0564	1.3642	0.6507	0.00011
7	1.0564	1.3642	0.6507	0.00001

Výpočet jsme ukončili už po sedmé iteraci, protože $\|\mathbf{x}^{(7)} - \mathbf{x}^{(6)}\|_R = 0.00001 \leq 10^{-4}$, a výsledek je $x_1 = 1.0564 \pm 10^{-4}$, $x_2 = 1.3642 \pm 10^{-4}$, $x_3 = 0.6507 \pm 10^{-4}$.

Poznámka

Z příkladů 4.2.1. a 4.2.2. je vidět, že Gauss-Seidelova metoda je rychlejší než metoda Jacobiova. Existují ale příklady, kdy Jacobiova metoda konverguje, zatímco Gauss-Seidelova metoda diverguje.



Pro obecnou soustavu (4.2.1), kde $a_{ii} \neq 0$, je Gauss-Seidelova metoda určena rekurentními vzorci

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right), \quad i = 1, \dots, n \quad (4.2.9)$$

pro $k = 0, 1, 2, \dots$

Nyní není na první pohled jasné, jak tyto vzorce zapsat v maticovém tvaru (4.2.3). Podívejme se proto ještě na náš příklad. Jestliže v (4.2.8) převedeme na

levou stranu všechny členy obsahující složky nové aproximace, dostaneme

$$\begin{aligned} 11x_1^{(k+1)} &= 15 - 2x_2^{(k)} - x_3^{(k)}, \\ x_1^{(k+1)} + 10x_2^{(k+1)} &= 16 - 2x_3^{(k)}, \\ -2x_1^{(k+1)} - 3x_2^{(k+1)} + 8x_3^{(k+1)} &= -1. \end{aligned}$$

Odtud je vidět, že výpočet nové aproximace $\mathbf{x}^{(k+1)}$ z aproximace předchozí $\mathbf{x}^{(k)}$ můžeme interpretovat také jako řešení soustavy lineárních rovnic s dolní trojúhelníkovou maticí. Pomocí aditivního rozkladu (4.2.7) matice \mathbf{A} to zapíšeme jako

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

a po úpravě

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

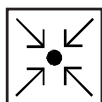
Obecné rekurentní vzorce (4.2.3) proto popisují Gauss-Seidelovu metodu, když v nich položíme $\mathbf{C} = \mathbf{C}_{GS}$ a $\mathbf{d} = \mathbf{d}_{GS}$, kde

$$\mathbf{C}_{GS} = -(\mathbf{L} + \mathbf{D})^{-1}\mathbf{U}, \quad \mathbf{d}_{GS} = (\mathbf{L} + \mathbf{D})^{-1}\mathbf{b}.$$



Kontrolní otázky

- Otázka 1. Jak vypadá obecné schéma iteračního řešení soustav lineárních rovnic?
Otázka 2. Jak se provádí výpočet u Jacobiovy a Gauss-Seidelovy metody? Která z nich je rychlejší?



Úlohy k samostatnému řešení

1. Soustavu lineárních rovnic

$$\begin{aligned} 4x_1 - x_2 + 2x_3 &= -12, \\ 2x_1 + 5x_2 + x_3 &= 5, \\ x_1 + x_2 - 3x_3 &= -4 \end{aligned}$$

řešte pomocí Jacobiovy metody s přesností $\epsilon = 10^{-2}$.

2. V předchozí úloze použijte při řešení Gauss-Seidelovu metodu.

3. Upravte obecný algoritmus pro iterační řešení soustav lineárních rovnic tak, aby vyjadřoval Jacobiovu resp. Gauss-Seidelovu metodu.

Výsledky úloh k samostatnému řešení



1. Rekurentní vzorce pro Jacobiovu metodu mají tvar

$$\begin{aligned}x_1^{(k+1)} &= \frac{1}{4}(-12 + x_2^{(k)} - 2x_3^{(k)}), \\x_2^{(k+1)} &= \frac{1}{5}(5 - 2x_1^{(k)} - x_3^{(k)}), \\x_3^{(k+1)} &= \frac{1}{3}(4 + x_1^{(k)} + x_2^{(k)}).\end{aligned}$$

Při nulové počáteční aproximaci dojdeme na požadovanou přesnost v jedenácté iteraci; $x_1 = -2.9955 \pm 10^{-2}$, $x_2 = 2.0019 \pm 10^{-2}$, $x_3 = 1.0011 \pm 10^{-2}$.

2. Rekurentní vzorce pro Gauss-Seidelovu metodu mají tvar

$$\begin{aligned}x_1^{(k+1)} &= \frac{1}{4}(-12 + x_2^{(k)} - 2x_3^{(k)}), \\x_2^{(k+1)} &= \frac{1}{5}(5 - 2x_1^{(k+1)} - x_3^{(k)}), \\x_3^{(k+1)} &= \frac{1}{3}(4 + x_1^{(k+1)} + x_2^{(k+1)}).\end{aligned}$$

Při nulové počáteční aproximaci dojdeme na požadovanou přesnost ve čtvrté iteraci; $x_1 = -2.9991 \pm 10^{-2}$, $x_2 = 1.9998 \pm 10^{-2}$, $x_3 = 1.0002 \pm 10^{-2}$.

3. Vstupní parametry \mathbf{C} a \mathbf{d} u původního algoritmu nahradíme za $\mathbf{A} = (a_{ij})$ a $\mathbf{b} = (b_i)$. Maticový výpočet nové iterace $\mathbf{x}^{(k+1)}$ zapíšeme rekurentními vzorci (4.2.5) resp. (4.2.9).

4.3. Vlastní čísla a vlastní vektory matic



Cíle

Připomeneme, jak se definují a počítají vlastní čísla a vlastní vektory matic. Poznatky z tohoto odstavce jsou základem při posuzování konvergence iteračních metod.



Předpokládané znalosti

Výpočty determinantů, řešení algebraických rovnic, řešení soustav lineárních rovnic se singulární maticí, lineární závislost a nezávislost vektorů, ortogonalita vektorů, báze.



Výklad

Úlohu na vlastní čísla připomeneme na příkladu.

Příklad 4.3.1. Uvažujme matici

$$\mathbf{A} = \begin{pmatrix} 2 & 0 & 0 \\ 2 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}. \quad (4.3.1)$$

Určete taková čísla λ , pro která má soustava lineárních rovnic $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ nenulové řešení a toto řešení vypočtěte.

Řešení: Soustavu přepíšeme do tvaru:

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \begin{pmatrix} 2 - \lambda & 0 & 0 \\ 2 & 2 - \lambda & 1 \\ 1 & 1 & 2 - \lambda \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Odtud je vidět, že jedním řešením je vždy nulový vektor. Nás však zajímají

situace, kdy existuje ještě další řešení nenulové. V takovém případě ale musí být nulový determinant matice soustavy, takže pro číslo λ platí

$$\det(\mathbf{A} - \lambda \mathbf{I}) = -\lambda^3 + 6\lambda^2 - 11\lambda + 6 = 0.$$

Dostali jsme algebraickou rovnici třetího stupně a pouze pro její tři kořeny $\lambda_1 = 3$, $\lambda_2 = 2$ a $\lambda_3 = 1$ bude mít uvažovaná soustava nenulová řešení. Najdeme je ze tří soustav lineárních rovnic

$$(\mathbf{A} - 3\mathbf{I})\mathbf{v} = \mathbf{0}, \quad (\mathbf{A} - 2\mathbf{I})\mathbf{v} = \mathbf{0}, \quad (\mathbf{A} - \mathbf{I})\mathbf{v} = \mathbf{0},$$

tj.

$$\begin{pmatrix} -1 & 0 & 0 \\ 2 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Vyřešením těchto soustav dostaneme

$$\mathbf{v}_1 = (0, r, r)^\top, \quad \mathbf{v}_2 = (s, -s, -2s)^\top, \quad \mathbf{v}_3 = (0, t, -t)^\top,$$

kde r, s a t jsou libovolná nenulová čísla. Vidíme, že každá soustava má nekonečně mnoho řešení. Konkrétní volbou r, s a t dostaneme například

$$\mathbf{v}_1 = (0, 1, 1)^\top, \quad \mathbf{v}_2 = (1, -1, -2)^\top, \quad \mathbf{v}_3 = (0, 1, -1)^\top.$$

Všimněme si ještě, že čísla λ_1 , λ_2 a λ_3 jsou vzájemně různá a že vektory \mathbf{v}_1 , \mathbf{v}_2 a \mathbf{v}_3 jsou lineárně nezávislé.

Definice 4.3.1.

Nechť \mathbf{A} je čtvercová matice řádu n . Číslo λ (obecně komplexní), pro které má soustava

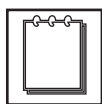
$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}, \quad \text{resp.} \quad (\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$$

nenulové řešení, se nazývá *vlastní číslo* matice \mathbf{A} a jemu odpovídající nenulové řešení $\mathbf{v} = (v_1, v_2, \dots, v_n)^\top$ se nazývá *vlastní vektor* matice \mathbf{A} .

Je zřejmé, že číslo λ je vlastním číslem matice \mathbf{A} právě tehdy, když je kořenem *charakteristického polynomu*

$$p_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = (-1)^n \lambda^n + c_1 \lambda^{n-1} + \dots + c_{n-1} \lambda + c_n.$$

Odtud plyne, že každá čtvercová matice řádu n má právě n vlastních čísel, pokud každé vlastní číslo počítáme tolikrát, kolik je násobnost kořene.

**Poznámka**

Vlastní čísla můžeme hledat jako řešení rovnice $p_{\mathbf{A}}(\lambda) = 0$ metodami z kapitoly 2. Tento postup je však obtížný pro větší hodnoty n , protože je pracné vypočítat koeficienty charakteristického polynomu c_i .

Snadno lze určit vlastní čísla u horní trojúhelníkové matice $\mathbf{U} = (u_{ij})$, $u_{ij} = 0$ pro $i > j$. Jsou to všechny diagonální prvky u_{ii} , protože z definice determinantu plyne, že charakteristický polynom má tvar

$$p_{\mathbf{U}}(\lambda) = (u_{11} - \lambda)(u_{22} - \lambda) \dots (u_{nn} - \lambda).$$

Stejně tvrzení platí samozřejmě i pro dolní trojúhelníkovou matici.

Při vyšetřování konvergence iteračních metod budeme využívat následující větu.

Věta 4.3.1.

Nechť λ je vlastní číslo matice \mathbf{A} , které odpovídá vlastnímu vektoru \mathbf{v} , c je dané reálné číslo a k je číslo přirozené. Potom $c\lambda^k$ je vlastní číslo matice $c\mathbf{A}^k$, které odpovídá vlastnímu vektoru \mathbf{v} .

Důkaz: Jestliže $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$, potom $c\mathbf{A}^k\mathbf{v} = c\lambda\mathbf{A}^{k-1}\mathbf{v} = \dots = c\lambda^k\mathbf{v}$. \square

Nyní si všimneme vlastních vektorů. V úvodním příkladu jsme viděli, že vlastní vektory odpovídající různým vlastním číslům jsou lineárně nezávislé. Toto tvrzení platí obecně, takže matice řádu n , může mít (nejvýše) n lineárně nezávislých vektorů. Tyto vektory pak tvoří bázi v prostoru n -složkových aritmetických vektorů. Následující příklad ukazuje, že matice nemusí mít vždy plný počet lineárně nezávislých vlastních vektorů.

Příklad 4.3.2. Určete vlastní čísla a vlastní vektory pro matice

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

$$\mathbf{C} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Řešení: Všechny matice jsou trojúhelníkové a mají stejný charakteristický polynom

$$p_{\mathbf{A}}(\lambda) = p_{\mathbf{B}}(\lambda) = p_{\mathbf{C}}(\lambda) = p_{\mathbf{D}}(\lambda) = (2 - \lambda)^3.$$

Číslo $\lambda = 2$ je tedy (trojnásobným) vlastním číslem všech čtyř matic. Postupem z příkladu 4.3.1. zjistíme, že matice \mathbf{A} má tři lineárně nezávislé vlastní vektory

$$\mathbf{v}_1 = (1, 0, 0)^\top, \quad \mathbf{v}_2 = (0, 1, 0)^\top, \quad \mathbf{v}_3 = (0, 0, 1)^\top$$

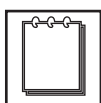
(každá nenulová lineární kombinace těchto vektorů je také vlastním vektorem matice \mathbf{A}). Pro matici \mathbf{B} se podaří najít pouze dva lineárně nezávislé vlastní vektory \mathbf{v}_1 a \mathbf{v}_3 . Matice \mathbf{C} má opět dva vlastní vektory, nyní to jsou vektory \mathbf{v}_1 a \mathbf{v}_2 . Konečně matice \mathbf{D} má jediný vlastní vektor \mathbf{v}_1 . \square

V aplikacích se často vyskytují symetrické matice, pro něž platí následující tvrzení.

Věta 4.3.2.

Nechť \mathbf{A} je symetrická čtvercová matice, tj. $\mathbf{A} = \mathbf{A}^\top$. Potom platí:

- (i) všechna vlastní čísla jsou reálná;
- (ii) vlastní vektory odpovídající různým vlastním číslům jsou ortogonální;
- (iii) k -násobnému vlastnímu číslu odpovídá k lineárně nezávislých vlastních vektorů, které lze zvolit tak, aby byly ortogonální.



Poznámka

Z věty plyne, že pro symetrickou matici řádu n můžeme vždy najít n ortogonálních vlastních vektorů. Protože ortogonální vektory jsou lineárně nezávislé, budou tvořit bázi v prostoru n -složkových aritmetických vektorů.

4.3.1. Výpočet vlastních čísel metodou LU-rozkladu

Ukážeme iterační metodu výpočtu vlastních čísel, která se v literatuře nazývá LR-algoritmus. Jejím základem jsou vlastnosti podobných matic.

Definice 4.3.2.

Dvě čtvercové matice \mathbf{A} a \mathbf{B} řádu n se nazývají *podobné*, jestliže existuje regulární čtvercová matice \mathbf{C} taková, že platí $\mathbf{A} = \mathbf{C}^{-1}\mathbf{B}\mathbf{C}$.

Věta 4.3.3.

Podobné matice mají stejná vlastní čísla.

Důkaz: Nechť λ je vlastní číslo matice \mathbf{A} odpovídající vlastnímu vektoru \mathbf{v} , tj. platí $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$. Potom

$$\underbrace{\mathbf{C}^{-1}\mathbf{A}\mathbf{C}}_{\mathbf{B}} \underbrace{\mathbf{C}^{-1}\mathbf{v}}_{\tilde{\mathbf{v}}} = \mathbf{C}^{-1}\mathbf{A}\mathbf{v} = \lambda \underbrace{\mathbf{C}^{-1}\mathbf{v}}_{\tilde{\mathbf{v}}}.$$

Odtud plyne, že λ je vlastní číslo matice \mathbf{B} odpovídající vlastnímu vektoru $\tilde{\mathbf{v}}$. \square

Metoda LU-rozkladu je založena na následujícím pozorování. Nechť \mathbf{L} a \mathbf{U} tvoří LU-rozklad matice \mathbf{A} podle věty 3.3.1., tj. platí $\mathbf{A} = \mathbf{L}\mathbf{U}$. Definujme matici $\mathbf{A}_1 = \mathbf{U}\mathbf{L}$. Protože

$$\mathbf{A}_1 = \mathbf{U}\mathbf{L} = \mathbf{L}^{-1}\mathbf{L}\mathbf{U}\mathbf{L} = \mathbf{L}^{-1}\mathbf{A}\mathbf{L},$$

vidíme, že matice \mathbf{A} a \mathbf{A}_1 jsou podobné a mají proto stejná vlastní čísla. Analogicky můžeme k matici \mathbf{A}_1 vytvořit podobnou matici \mathbf{A}_2 atd. Dostaneme posloupnost podobných matic a budeme se zajímat o vlastní čísla limitní matice.

Algoritmus (Metoda LU-rozkladu)

Položíme $\mathbf{A}_0 = \mathbf{A}$ a pro $k = 1, 2, \dots$ provedeme:

- 1) LU-rozklad matice \mathbf{A}_{k-1} , tj. určíme \mathbf{L}_k a \mathbf{U}_k tak, že $\mathbf{A}_{k-1} = \mathbf{L}_k\mathbf{U}_k$;
- 2) vypočítáme součin $\mathbf{A}_k := \mathbf{U}_k\mathbf{L}_k$.

Všimněme si posloupností $\{\mathbf{A}_k\}$ a $\{\mathbf{U}_k\}$. Lze dokázat, že za jistých předpokladů konvergují obě tyto posloupnosti ke stejné limitní matici \mathbf{M} ; viz [1]. Tato matice je nutně horní trojúhelníková a má stejná vlastní čísla jako \mathbf{A} . Hledaná vlastní čísla proto určíme jako diagonální prvky matice \mathbf{M} .

Příklad 4.3.3. Pomocí metody LU-rozkladu vypočtěte vlastní čísla matice

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \quad (4.3.2)$$

s přesností na dvě desetinná místa.

Řešení: Pro $\mathbf{A}_0 = \mathbf{A}$ určíme LU-rozklad

$$\mathbf{L}_0 = \begin{pmatrix} 1 & 0 & 0 \\ -0.50 & 1 & 0 \\ 0 & -0.67 & 1 \end{pmatrix}, \quad \mathbf{U}_0 = \begin{pmatrix} 2.00 & -1.00 & 0 \\ 0 & 1.50 & -1.00 \\ 0 & 0 & 1.33 \end{pmatrix}$$

a vynásobením dostaneme

$$\mathbf{A}_1 = \mathbf{U}_0 \mathbf{L}_0 = \begin{pmatrix} 2.50 & -1.00 & 0 \\ -0.75 & 2.17 & -1.00 \\ 0 & -0.89 & 1.33 \end{pmatrix}.$$

Podobně pro \mathbf{A}_1 vypočítáme LU-rozklad

$$\mathbf{L}_1 = \begin{pmatrix} 1 & 0 & 0 \\ -0.30 & 1 & 0 \\ 0 & -0.48 & 1 \end{pmatrix}, \quad \mathbf{U}_1 = \begin{pmatrix} 2.50 & -1.00 & 0 \\ 0 & 1.87 & -1.00 \\ 0 & 0 & 0.86 \end{pmatrix}$$

a opět vynásobením dostaneme

$$\mathbf{A}_2 = \mathbf{U}_1 \mathbf{L}_1 = \begin{pmatrix} 2.80 & -1.00 & 0 \\ -0.56 & 2.34 & -1.00 \\ 0 & -0.41 & 0.86 \end{pmatrix}.$$

Čísla pod diagonálou u matic \mathbf{A}_k se začínají přibližovat k nule, což ukazuje na tendenci ke konvergenci. Jestliže takto pokračujeme dále, dostaneme

$$\mathbf{A}_{13} = \begin{pmatrix} 3.41 & -1.00 & 0 \\ 0.00 & 2.00 & -1.00 \\ 0 & 0.00 & 0.58 \end{pmatrix} \approx \mathbf{M}$$

a v dalších iteracích se již čísla na diagonále (na prvních dvou desetinných místech) nemění. Přibližné hodnoty vlastních čísel jsou $\lambda_1 \doteq 3.41$, $\lambda_2 \doteq 2.00$ a $\lambda_3 \doteq 0.58$.

**Kontrolní otázky**

Otázka 1. Jak se definují vlastní čísla a vlastní vektory matic?

Otázka 2. Kolik vlastních čísel a vlastních vektorů má matice řádu n ?

Otázka 3. Na jaké vlastnosti je založena metoda LU-rozkladu?

Úlohy k samostatnému řešení

1. Pomocí charakteristického polynomu vypočtete vlastní čísla matice (4.3.2).

Vypočtete také vlastní vektory.

2. Metodou LU-rozkladu vypočtete vlastní čísla matice

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

s přesností na čtyři desetinná místa.

Výsledky úloh k samostatnému řešení

1. $p_{\mathbf{A}}(\lambda) = \lambda^3 - 6\lambda^2 + 10\lambda - 4$, $\lambda_1 = 3.414214$, $\lambda_2 = 2$, $\lambda_3 = 0.585786$. Vlastní vektory jsou například $\mathbf{v}_1 = (-1, \sqrt{2}, -1)^\top$, $\mathbf{v}_2 = (-\sqrt{2}, 0, \sqrt{2})^\top$, $\mathbf{v}_3 = (1, \sqrt{2}, 1)^\top$.

2. Vlastní čísla s požadovanou přesností jsou na diagonále matice \mathbf{A}_{20} ; $\lambda_1 \doteq 3.7320$, $\lambda_2 \doteq 2.0000$, $\lambda_3 \doteq 0.2679$.

4.4. Konvergence iteračních metod



Cíle

Odvodíme podmínky, které zaručují konvergenci iteračních metod.



Předpokládané znalosti

Iterační řešení soustav lineárních rovnic. Vlastní čísla a vlastní vektory.



Výklad

Připomeňme, že při iteračním řešení převádíme soustavu lineárních rovnic $\mathbf{Ax} = \mathbf{b}$ na (ekvivalentní) soustavu v iteračním tvaru

$$\mathbf{x} = \mathbf{Cx} + \mathbf{d}. \quad (4.4.1)$$

K řešení se potom přibližujeme pomocí posloupnosti $\{\mathbf{x}^{(k)}\}$, kterou počítáme podle rekurentního vzorce

$$\mathbf{x}^{(k+1)} = \mathbf{Cx}^{(k)} + \mathbf{d}. \quad (4.4.2)$$

Jestliže odečteme (4.4.1) a (4.4.2) a označíme přitom $\mathbf{e}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x}$, dostaneme

$$\mathbf{e}^{(k+1)} = \mathbf{Ce}^{(k)}.$$

Tímto vzorcem se řídí *iterační chyba* $\mathbf{e}^{(k)}$. Jeho opakovaným použitím dostaneme $\mathbf{e}^{(k+1)} = \mathbf{Ce}^{(k)} = \mathbf{C}^2\mathbf{e}^{(k-1)} = \dots = \mathbf{C}^{k+1}\mathbf{e}^{(0)}$. Proto

$$\mathbf{e}^{(k)} = \mathbf{C}^k\mathbf{e}^{(0)}, \quad (4.4.3)$$

kde $\mathbf{e}^{(0)}$ je *počáteční chyba*, která je určena volbou počáteční aproximace $\mathbf{x}^{(0)}$.

Je zřejmé, že iterační výpočet bude konvergovat, jestliže $\lim_{k \rightarrow \infty} \mathbf{e}^{(k)} = \mathbf{0}$, tj. když se iterační chyba blíží k nulovému vektoru. Tuto limitu budeme vyšetřovat pomocí vzorce (4.4.3). Budeme přitom předpokládat, že iterační matice \mathbf{C} má

vlastní čísla $\lambda_1, \dots, \lambda_n$, kterým odpovídají vlastní vektory $\mathbf{v}_1, \dots, \mathbf{v}_n$, které tvoří bázi. Připomeňme, že taková situace nastane podle věty 4.3.2. a následně poznámky například v případě, kdy je \mathbf{C} symetrická matice. Vektor $\mathbf{e}^{(0)}$ pak můžeme zapsat jako lineární kombinaci vlastních vektorů, tj. existují konstanty c_1, \dots, c_n , pro něž platí

$$\mathbf{e}^{(0)} = c_1 \mathbf{v}_1 + \dots + c_n \mathbf{v}_n. \quad (4.4.4)$$

Jestliže dosadíme (4.4.4) do (4.4.3) dostaneme s pomocí věty 4.3.1. vztah

$$\begin{aligned} \mathbf{e}^{(k)} &= c_1 \mathbf{C}^k \mathbf{v}_1 + \dots + c_n \mathbf{C}^k \mathbf{v}_n \\ &= c_1 \lambda_1^k \mathbf{v}_1 + \dots + c_n \lambda_n^k \mathbf{v}_n. \end{aligned} \quad (4.4.5)$$

Odtud je vidět, že

$$\lim_{k \rightarrow \infty} \mathbf{e}^{(k)} = \mathbf{0} \iff \lim_{k \rightarrow \infty} \lambda_i^k = 0 \quad \text{pro } i = 1, \dots, n.$$

Uvedené limity budou nulové, právě když $|\lambda_i| < 1$ pro $i = 1, \dots, n$. Dokázali jsme následující tvrzení:

Věta 4.4.1.

Nechť \mathbf{C} je iterační matice, která má n lineárně nezávislých vlastních vektorů. Iterační metoda daná vzorcem (4.4.2) konverguje pro každou počáteční aproximaci $\mathbf{x}^{(0)}$, právě když všechna vlastní čísla matice \mathbf{C} jsou v absolutní hodnotě menší než jedna.

Příklad 4.4.1. Určete vlastní čísla iteračních matic \mathbf{C}_J a \mathbf{C}_B z odstavce 4.1. a porovnejte průběhy iteračních výpočtů s tvrzením poslední věty.

Řešení: Z charakteristického polynomu $p_{\mathbf{C}_J}(\lambda) = \lambda^3 + \frac{7}{88}\lambda - \frac{1}{80}$ určíme vlastní čísla matice \mathbf{C}_J : $\lambda_1 \doteq 0.1297$, $\lambda_2 \doteq -0.0649 + i0.3036$, $\lambda_3 \doteq -0.0649 - i0.3036$. Z absolutních hodnot $|\lambda_1| = \lambda_1$, $|\lambda_2| = |\lambda_3| \doteq 0.3104$ vidíme, že iterační výpočet musí být konvergentní, což je v souladu s naším pozorováním z odstavce 4.1.

Podobně z charakteristického polynomu $p_{\mathbf{C}_B}(\lambda) = \lambda^3 + 26\lambda^2 + 229\lambda + 683$ určíme vlastní čísla matice \mathbf{C}_B . Stačí si povšimnout, že jedno z vlastních čísel je $\lambda_1 \doteq -8.7373$. Protože $|\lambda_1| > 1$, nemůže iterační výpočet (obecně) konvergovat, což je rovněž v souladu s pozorováním z odstavce 4.1.

Viděli jsme, že o konvergenci iterační metody lze rozhodnout na základě vlastních čísel. Výpočet vlastních čísel je však zpravidla složitější než řešení soustavy lineárních rovnic. V další větě proto ukážeme jednodušší, i když slabší konvergenční podmínku.

Věta 4.4.2.

Nechť \mathbf{C} je iterační matice, která má n lineárně nezávislých vlastních vektorů. Iterační metoda daná vzorcem (4.4.2) konverguje pro každou počáteční aproximaci $\mathbf{x}^{(0)}$, jestliže pro některou normu platí $\|\mathbf{C}\| < 1$.

Důkaz: Nechť $\|\mathbf{C}\| < 1$ a nechť λ je libovolné vlastní číslo matice \mathbf{C} odpovídající vlastnímu vektoru \mathbf{v} , tj. $\mathbf{C}\mathbf{v} = \lambda\mathbf{v}$. Protože $|\lambda|\|\mathbf{v}\| = \|\lambda\mathbf{v}\| = \|\mathbf{C}\mathbf{v}\| \leq \|\mathbf{C}\|\|\mathbf{v}\|$, platí $|\lambda| \leq \|\mathbf{C}\| < 1$, takže všechna vlastní čísla matice \mathbf{C} jsou v absolutní hodnotě menší než jedna. Iterační metoda proto konverguje podle věty 4.4.1. \square

U Jacobiovy a Gauss-Seidelovy metody lze konvergenční podmínku z poslední věty formulovat pomocí matice soustavy \mathbf{A} . Používá se přitom terminologie z následující definice.

Definice 4.4.1.

Řekneme, že čtvercová matice $\mathbf{A} = (a_{ij})$ řádu n je *ostře diagonálně dominantní*, jestliže platí

$$|a_{i1}| + \dots + |a_{ii-1}| + |a_{ii+1}| + \dots + |a_{in}| < |a_{ii}| \quad \text{pro } i = 1, \dots, n. \quad (4.4.6)$$

Příklad 4.4.2. Rozhodněte, která z následujících matic je ostře diagonálně dominantní:

$$\mathbf{A} = \begin{pmatrix} 11 & 2 & 1 \\ 1 & 10 & 2 \\ 2 & 3 & -8 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} -8 & 2 & 1 \\ -2 & 1 & -3 \\ 1 & 1 & 3 \end{pmatrix}.$$

Řešení: Matice \mathbf{A} je ostře diagonálně dominantní, protože platí $2 + 1 < 11$, $1 + 2 < 10$ a $2 + 3 < 8$. Matice \mathbf{B} není ostře diagonálně dominantní, protože ve druhém řádku je $2 + 3 > 1$.

Věta 4.4.3.

Nechť $\mathbf{Ax} = \mathbf{b}$ je daná soustava lineárních rovnic. Jacobiova iterační metoda konverguje pro každou počáteční aproximaci $\mathbf{x}^{(0)}$, jestliže matice \mathbf{A} je ostře diagonálně dominantní.

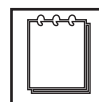
Důkaz: Nechť \mathbf{A} je ostře diagonálně dominantní. Podmínku (4.4.6) může přepsat do tvaru

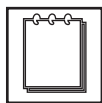
$$\frac{|a_{i1}|}{|a_{ii}|} + \dots + \frac{|a_{ii-1}|}{|a_{ii}|} + \frac{|a_{ii+1}|}{|a_{ii}|} + \dots + \frac{|a_{in}|}{|a_{ii}|} < 1, \quad \text{pro } i = 1, \dots, n.$$

Pro iterační matici Jacobiovy metody \mathbf{C}_J (viz (4.2.6)) to znamená, že součet absolutních hodnot prvků v každém řádku je menší než jedna. V řádkové normě proto platí $\|\mathbf{C}_J\|_R < 1$ a konvergence Jacobiovy metody plyne z věty 4.4.2. \square

Poznámka

Také Gauss-Seidelova metoda konverguje, je-li matice soustavy ostře diagonálně dominantní. Důkaz je však o něco složitější; viz [1].



**Poznámka**

Konvergenci Jacobiový i Gauss-Seidelovy metody zajistíme tak, že řešenou soustavu předem upravíme na ekvivalentní soustavu s ostře diagonálně dominantní maticí.

Příklad 4.4.3. Soustavu lineárních rovnic

$$-8x_1 + 2x_2 + x_3 = -1,$$

$$-2x_1 + x_2 - 3x_3 = -9,$$

$$x_1 + x_2 + 3x_3 = 12$$

upravte na tvar s ostře diagonálně dominantní maticí.

Řešení: Úpravy, které nemění řešení, jsou tři: záměna pořadí rovnic, vynásobení rovnice nenulovým číslem a přičtení nenulového násobku rovnice k jiné rovnici. V našem případě stačí přičíst třetí rovnici k rovnici druhé:

$$-8x_1 + 2x_2 + x_3 = -1,$$

$$-x_1 + 2x_2 = -3,$$

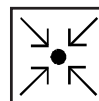
$$x_1 + x_2 + 3x_3 = 12.$$

Provedeme-li výpočet Jacobiový nebo Gauss-Seidelovy metody pro tuto soustavu, budeme mít zaručenu konvergenci.

**Kontrolní otázky**

Otázka 1. Jaké podmínky zaručují konvergenci obecné iterační metody?

Otázka 2. Jak lze zajistit konvergenci u Jacobiový a Gauss-Seidelovy metody?

Úlohy k samostatnému řešení

1. Soustavu lineárních rovnic

$$4x_1 + x_2 + x_3 = 6,$$

$$x_1 + 4x_2 + x_3 = 6,$$

$$6x_1 + 6x_2 + 6x_3 = 18$$

upravte na tvar s ostře diagonálně dominantní maticí.

2. Napište iterační matici pro Jacobiovu metodu a vypočtěte její vlastní čísla.

Výsledky úloh k samostatnému řešení

1. Od třetí rovnice odečteme rovnici první i druhou. Dostaneme:

$$4x_1 + x_2 + x_3 = 6,$$

$$x_1 + 4x_2 + x_3 = 6,$$

$$x_1 + x_2 + 4x_3 = 6.$$

2. Iterační matice má tvar:

$$\mathbf{C}_J = \begin{pmatrix} 0 & -1/4 & -1/4 \\ -1/4 & 0 & -1/4 \\ -1/4 & -1/4 & 0 \end{pmatrix}.$$

Z charakteristického polynomu $p_{\mathbf{C}_J}(\lambda) = \lambda^3 - \frac{3}{16}\lambda + \frac{1}{32}$ určíme kořeny $\lambda_1 = -\frac{1}{2}$, $\lambda_2 = \lambda_3 = \frac{1}{4}$.

Shrnutí lekce

Ukázali jsme iterační způsob řešení soustav lineárních rovnic a odvodili jsem podmínky, které zaručují konvergenci.

5. Interpolace a aproximace funkcí



Průvodce studiem

Často je potřeba „složitou“ funkci f nahradit funkcí „jednodušší“. V této kapitole budeme předpokládat, že u funkce f známe její funkční hodnoty $f_i = f(x_i)$ v *uzlech* x_i pro $i = 0, \dots, n$. Budeme rozlišovat dvě úlohy.

Interpolační úloha: Hledáme funkci φ , pro niž je

$$\varphi(x_i) = f_i, \quad i = 0, \dots, n. \quad (5.0.1)$$

Aproximace metodou nejmenších čtverců: Hledáme funkci φ , pro niž je

$$\varphi(x_i) \approx f_i, \quad i = 0, \dots, n, \quad (5.0.2)$$

kde přibližná rovnost „ \approx “ je určena tak, aby součet druhých mocnin odchylek mezi předepsanými hodnotami f_i a předpokládanými hodnotami $\varphi(x_i)$ byl minimální.

Jestliže tyto úlohy znázorníme graficky, bude řešení interpolační úlohy procházet přes body (x_i, f_i) , $i = 0, \dots, n$, kdežto řešení aproximační úlohy bude (obecně) procházet jejich blízkým okolím.

Formulace obou úloh je zatím příliš obecná, protože jsme neřekli jakého typu má být funkce φ . Ukážeme tři volby: polynom, splajn (spline-funkce) a lineární kombinace obecných funkcí. Polynom je jednoduchý z hlediska matematických operací (snadno se derivuje, integruje atp.), jeho graf však často osciluje. Lepší tvary grafu mají splajny. Kombinace obecných funkcí se používá zpravidla v situacích, kdy je známo, jakou závislost daná data popisují (pro periodickou závislost je dobré použít funkce goniometrické, pro strmě rostoucí data se hodí funkce exponenciální atp.).

5.1. Interpolační polynom

Cíle



Ukážeme metody pro sestavení interpolačního polynomu a odvodíme vzorec pro interpolační chybu.

Předpokládané znalosti



Polynomy. Řešení soustav lineárních rovnic. Věta o střední hodnotě diferenciálního počtu.

Výklad



Funkci φ v úloze (5.0.1) budeme hledat jako *interpolační polynom* stupně nejvýše n , tj. položíme $\varphi = p_n$, kde

$$p_n(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n. \quad (5.1.1)$$

Začneme příkladem.

Příklad 5.1.1. Jsou dány uzly $x_0 = -2$, $x_1 = -1$, $x_2 = 1$, $x_3 = 2$ a funkční hodnoty $f_0 = 10$, $f_1 = 4$, $f_2 = 6$, $f_3 = 3$. Určete interpolační polynom p_3 .

Řešení: Hledaný polynom má obecný tvar

$$p_3(x) = a_0 + a_1x + a_2x^2 + a_3x^3.$$

Koeficienty a_0 , a_1 , a_2 , a_3 určíme tak, aby platilo (5.0.1). Každá interpolační rovnost určuje jednu rovnici:

$$p_3(-2) = 10 \Rightarrow a_0 - 2a_1 + 4a_2 - 8a_3 = 10,$$

$$p_3(-1) = 4 \Rightarrow a_0 - a_1 + a_2 - a_3 = 4,$$

$$p_3(1) = 6 \Rightarrow a_0 + a_1 + a_2 + a_3 = 6,$$

$$p_3(2) = 3 \Rightarrow a_0 + 2a_1 + 4a_2 + 8a_3 = 3.$$

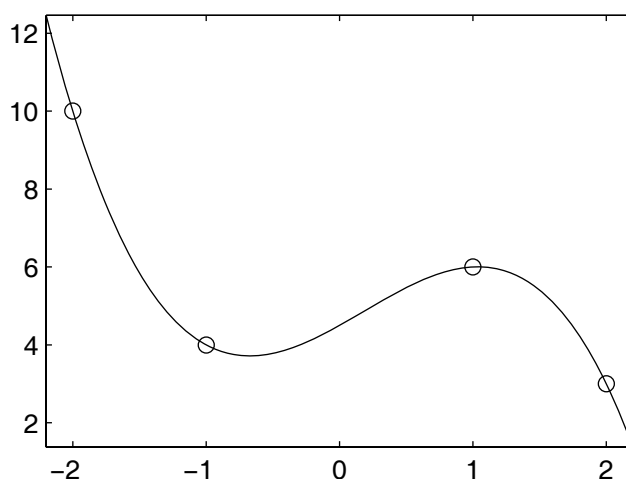
Dostali jsme soustavu lineárních rovnic

$$\begin{pmatrix} 1 & -2 & 4 & -8 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 10 \\ 4 \\ 6 \\ 3 \end{pmatrix},$$

jejímž řešením (na tři desetinná místa) jsou koeficienty $a_0 = 4.500$, $a_1 = 1.917$, $a_2 = 0.500$ a $a_3 = -0.917$. Interpolační polynom má tvar

$$p_3(x) = 4.500 + 1.917x + 0.500x^2 - 0.917x^3.$$

Jeho graf je na obrázku 5.1.1.



Obrázek 5.1.1: Graf interpolačního polynomu p_3 .

Rozborem postupu z příkladu dokážeme následující větu.

Věta 5.1.1.

Nechť jsou dány vzájemně různé uzly x_i a funkční hodnoty f_i , $i = 0, \dots, n$. Existuje právě jeden interpolační polynom stupně nejvýše n .

Důkaz: Dosazením obecného tvaru polynomu (5.1.1) do interpolačních rovností

(5.0.1) dostaneme soustavu lineárních rovnic

$$p_n(x_i) = a_0 + a_1x_i + a_2x_i^2 + \dots + a_nx_i^n = f_i, \quad i = 0, \dots, n,$$

kterou lze zapsat maticově jako

$$\begin{pmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ 1 & x_2 & x_2^2 & \dots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} f_0 \\ f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}.$$

Matice této soustavy má nenulový determinant (Vandermodův determinant). Odtud plyne existence jediného řešení soustavy lineárních rovnic a také interpolačního polynomu. \square

5.1.1. Lagrangeův tvar interpolačního polynomu

Ukážeme postup, při němž se obejdeme bez řešení soustavy lineárních rovnic. Interpolační polynom budeme hledat ve tvaru

$$p_n(x) = f_0\varphi_0(x) + f_1\varphi_1(x) + \dots + f_n\varphi_n(x). \quad (5.1.2)$$

Rovnosti $p_n(x_i) = f_i$, $i = 0, 1, \dots, n$ budou splněny, jestliže bude platit

$$\varphi_i(x_j) = \begin{cases} 1 & \text{pro } i = j, \\ 0 & \text{pro } i \neq j. \end{cases}$$

Z věty 5.1.1. víme, že interpolační polynom je stupně nejvýše n , takže také všechny funkce φ_i musí být polynomy stupně nejvýše n . Uvedeným požadavkům vyhovuje následující definice:

$$\varphi_i(x) = \frac{(x - x_0) \dots (x - x_{i-1})(x - x_{i+1}) \dots (x - x_n)}{(x_i - x_0) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_n)} \quad (5.1.3)$$

pro $i = 0, 1, \dots, n$. Čitatel je totiž polynom, který nabývá nulových hodnot ve všech uzlech kromě x_i . V uzlu x_i pak nabývá nenulové hodnoty, která je obsažena ve jmenovateli zlomku, takže platí $\varphi_i(x_i) = 1$.

Polynomům φ_i , $i = 0, 1, \dots, n$ se říká *Lagrangeova báze* interpolační úlohy a vzorec (5.1.2) se nazývá *Lagrangeův tvar* interpolačního polynomu.

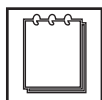
Příklad 5.1.2. Mějme dány uzly $x_0 = -2$, $x_1 = -1$, $x_2 = 1$, $x_3 = 2$ a funkční hodnoty $f_0 = 10$, $f_1 = 4$, $f_2 = 6$, $f_3 = 3$. Napište Lagrangeův tvar interpolačního polynomu.

Řešení: Nejdříve sestavíme Lagrangeovu bázi. Podle (5.1.3) je

$$\begin{aligned}\varphi_0(x) &= \frac{(x+1)(x-1)(x-2)}{(-2+1)(-2-1)(-2-2)} = -\frac{1}{12}(x+1)(x-1)(x-2), \\ \varphi_1(x) &= \frac{(x+2)(x-1)(x-2)}{(-1+2)(-1-1)(-1-2)} = \frac{1}{6}(x+2)(x-1)(x-2), \\ \varphi_2(x) &= \frac{(x+2)(x+1)(x-2)}{(1+2)(1+1)(1-2)} = -\frac{1}{6}(x+2)(x+1)(x-2), \\ \varphi_3(x) &= \frac{(x+2)(x+1)(x-1)}{(2+2)(2+1)(2-1)} = \frac{1}{12}(x+2)(x+1)(x-1).\end{aligned}$$

Dosazením do (5.1.2) dostaneme výsledek

$$\begin{aligned}p_3(x) &= -\frac{5}{6}(x+1)(x-1)(x-2) + \frac{2}{3}(x+2)(x-1)(x-2) - \\ &\quad -(x+2)(x+1)(x-2) + \frac{1}{4}(x+2)(x+1)(x-1).\end{aligned}$$



Poznámka

Interpolační polynom je podle věty 5.1.1. určen jednoznačně. Úpravou Lagrangeova tvaru proto musíme nutně dojít k polynomu, který jsem vypočítali v příkladu 5.1.1. (ověřte).

5.1.2. Newtonův tvar interpolačního polynomu

Uvažujme zápis polynomu ve tvaru:

$$p_n(x) = a_0 + a_1(x-x_0) + a_2(x-x_0)(x-x_1) + \dots + a_n(x-x_0)\dots(x-x_{n-1}). \quad (5.1.4)$$

Jestliže dosadíme do interpolačních rovností $p_n(x_i) = f_i$, $i = 0, 1, \dots, n$, dostaneme soustavu lineárních rovnic s dolní trojúhelníkovou maticí:

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & (x_1 - x_0) & 0 & \dots & 0 \\ 1 & (x_2 - x_0) & (x_2 - x_0)(x_2 - x_1) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} f_0 \\ f_1 \\ f_2 \\ \vdots \end{pmatrix}, \quad (5.1.5)$$

Odud můžeme postupně vyjádřit koeficienty a_k :

$$\begin{aligned} a_0 &= f_0, & a_1 &= \frac{f_1 - a_0}{x_1 - x_0} = \frac{f_1 - f_0}{x_1 - x_0}, \\ a_2 &= \frac{f_2 - a_1(x_2 - x_0) - a_0}{(x_2 - x_0)(x_2 - x_1)} = \frac{\frac{f_2 - f_1}{x_2 - x_1} - \frac{f_1 - f_0}{x_1 - x_0}}{x_2 - x_0}, \end{aligned}$$

atd..

Výrazy na pravých stranách jsou poměrné difference, jejichž označení zavádíme v následující definici.

Definice 5.1.1.

Nechť jsou dány vzájemně různé uzly x_i a funkční hodnoty f_i , $i = 0, \dots, n$. Poměrné difference k -tého řádu $f[x_{i+k}, \dots, x_i]$, $i = 0, 1, \dots, n - k$ definujeme rekurentně:

- pro $k = 0$: $f[x_i] = f_i$;
- pro $k = 1$: $f[x_{i+1}, x_i] = \frac{f_{i+1} - f_i}{x_{i+1} - x_i}$;
- pro $k \leq n$: $f[x_{i+k}, \dots, x_i] = \frac{f[x_{i+k}, \dots, x_{i+1}] - f[x_{i+k-1}, \dots, x_i]}{x_{i+k} - x_i}$.

Porovnáním poměrných diferencí s koeficienty a_k vidíme, že

$$a_k = f[x_k, \dots, x_0], \quad k = 0, 1, \dots, n.$$

Dosazením do (5.1.4) dostaneme *Newtonův tvar* interpolačního polynomu:

$$p_n(x) = f_0 + f[x_1, x_0](x - x_0) + \dots + f[x_n, \dots, x_0](x - x_0) \dots (x - x_{n-1}). \quad (5.1.6)$$

Při jeho sestavování potřebujeme vypočítat poměrné difference. Vše ukážeme v následujícím příkladu.

Příklad 5.1.3. Mějme dány uzly $x_0 = -2$, $x_1 = -1$, $x_2 = 1$, $x_3 = 2$ a funkční hodnoty $f_0 = 10$, $f_1 = 4$, $f_2 = 6$, $f_3 = 3$. Napište Newtonův tvar interpolačního polynomu.

Řešení: Potřebujeme vypočítat poměrné difference:

$$f[x_1, x_0], \quad f[x_2, x_1, x_0], \quad f[x_3, x_2, x_1, x_0].$$

Podle definice je

$$\begin{aligned} f[x_1, x_0] &= \frac{4 - 10}{-1 + 2} = -6, \\ f[x_2, x_1, x_0] &= \frac{f[x_2, x_1] - f[x_1, x_0]}{x_2 - x_0} = \frac{\frac{6-4}{1+1} + 6}{1 + 2} = \frac{7}{3}, \\ f[x_3, x_2, x_1, x_0] &= \frac{f[x_3, x_2, x_1] - f[x_2, x_1, x_0]}{x_3 - x_0} = \frac{\frac{\frac{3-6}{2-1} - \frac{6-4}{1+1}}{2+1} - \frac{7}{3}}{2 + 2} = -\frac{11}{12}. \end{aligned}$$

Dosazením do (5.1.6) dostaneme výsledek

$$p_3(x) = 10 - 6(x + 2) + \frac{7}{3}(x + 2)(x + 1) - \frac{11}{12}(x + 2)(x + 1)(x - 1).$$

Přehledně můžeme výpočet poměrných diferencí provést v tabulce (tabulka 5.1.1), kde do prvních dvou sloupců zapíšeme zadané uzly a funkční hodnoty a v každém dalším sloupci pak vypočítáme všechny (!) poměrné difference postupně se zvyšujících řádů. Pro napsání interpolačního polynomu potřebujeme z této tabulky hodnoty diferencí z prvního řádku.

Tabulka 5.1.1: Výpočet poměrných diferencí.

i	x_i	f_i	$f[x_{i+1}, x_i]$	$f[x_{i+2}, x_{i+1}, x_i]$	$f[x_3, x_2, x_1, x_0]$
0	-2	10	-6	$\frac{7}{3}$	$-\frac{11}{12}$
1	-1	4	1	$-\frac{4}{3}$	
2	1	6	-3		
3	2	3			

5.1.3. Interpolační chyba

Předpokládejme, že hodnoty f_i jsou funkčními hodnotami funkce f v uzlech x_i , tj. $f_i = f(x_i)$. Bude nás zajímat *interpolační chyba*

$$f(x) - p_n(x).$$

V uzlech x_i je interpolační chyba nulová, ale mimo uzly může být velká.

Věta 5.1.2.

Nechť uzly x_i , $i = 0, 1, \dots, n$, jsou vzájemně různé a leží na intervalu $\langle a, b \rangle$. Nechť funkce f má na tomto intervalu $n+1$ spojitých derivací. Pak pro každé $x \in \langle a, b \rangle$ existuje $\xi = \xi(x)$ v (a, b) tak, že platí

$$f(x) - p_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \pi_{n+1}(x), \quad (5.1.7)$$

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

Důkaz: Pro $x = x_i$ je rovnost (5.1.7) splněna, protože obě její strany jsou nulové.

Pro pevně zvolené $x \neq x_i$ definujme funkci

$$g(t) = f(t) - p_n(t) - \frac{\pi_{n+1}(t)}{\pi_{n+1}(x)} (f(x) - p_n(x)), \quad (5.1.8)$$

kde t je proměnná a x je parametr. Funkce g má zřejmě $n+2$ kořenů, kterými jsou body x_0, \dots, x_n a x . Každá derivace funkce g má o jeden kořen méně, takže $(n+1)$ -ní derivace má jediný kořen v nějakém bodě $\xi \in (a, b)$. Derivujeme-li $(n+1)$ -krát

výraz (5.1.8) (podle t) a použijeme přitom $p_n^{(n+1)}(t) = 0$ a $\pi_{n+1}^{(n+1)}(t) = (n+1)!$, dostaneme

$$0 = g^{(n+1)}(\xi) = f^{(n+1)}(\xi) - \frac{(n+1)!}{\pi_{n+1}(x)} (f(x) - p_n(x)).$$

Jestliže odtud vyjádříme interpolační chybu, vznikne rovnost (5.1.7). \square

Na průběh interpolační chyby v intervalu $\langle a, b \rangle$ má podstatný vliv tvar polynomu π_{n+1} , jak ukazuje následující příklad.

Příklad 5.1.4. (Rungeho příklad) Nakreslíme graf funkce

$$f(x) = \frac{1}{1+x^2}$$

a graf interpolačního polynomu odpovídajícího uzlům $x_i = -5+i$, $i = 0, 1, \dots, 10$.

Výsledek porovnáme s grafem polynomu

$$\pi_{11}(x) = (x+5)(x+4) \dots (x-5).$$

Řešení: Obrázek 5.1.2.a ukazuje graf polynomu π_{11} . Z jeho průběhu lze usoudit, že největší interpolační chyby budou poblíž krajních uzlů $x_0 = -5$ a $x_{10} = 5$. Na obrázku 5.1.2.b vidíme, že graf interpolačního polynomu osciluje kolem grafu funkce f a že oscilace jsou největší právě na krajích intervalu $\langle -5, 5 \rangle$. Poznamenejme ještě, že při zvětšení počtu interpolačních uzlů nedojde ke zmenšení interpolační chyby, ale naopak k jejímu zvětšení.

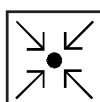


Kontrolní otázky

Otázka 1. Jaké znáte metody pro sestavení interpolačního polynomu?

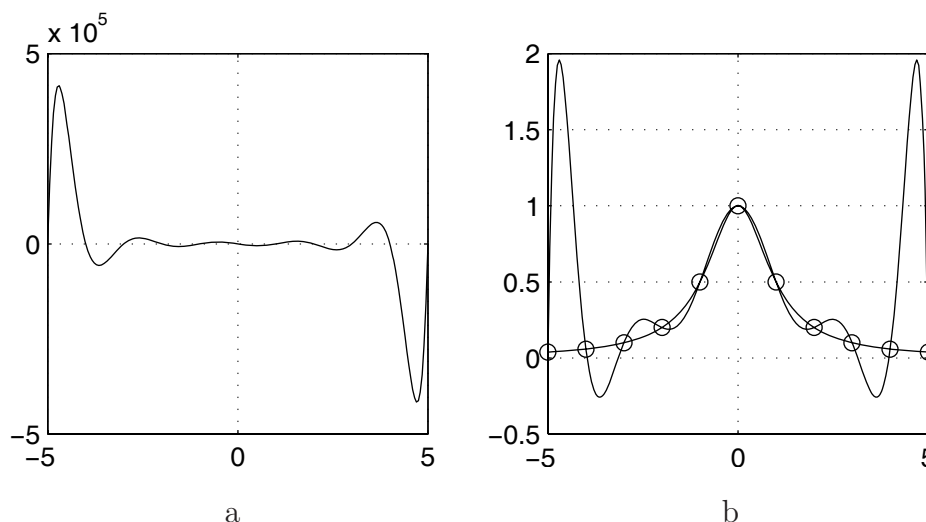
Otázka 2. Jakého stupně je interpolační polynom?

Otázka 3. Jak se chová interpolační chyba?



Úlohy k samostatnému řešení

1. Pro uzly $x_0 = -1$, $x_1 = 0$, $x_2 = 2$, $x_3 = 3$, $x_4 = 5$ a funkční hodnoty $f_0 = -2$,



Obrázek 5.1.2: a) Graf π_{11} ; b) Grafy f (neoscilující) a p_{10} (oscilující).

$f_1 = 1, f_2 = 0, f_3 = 2, f_4 = -1$ vypočtete interpolační polynom ve tvaru (5.1.1).

2. Pro předchozí data vypočtete Lagrangeův a Newtonův tvar interpolačního polynomu.

Výsledky úloh k samostatnému řešení

1. $p_4(x) = -\frac{3}{20}x^4 + \frac{11}{10}x^3 - \frac{109}{60}x^2 - \frac{1}{15}x + 1.$

2. Lagrangeův tvar: $p_4(x) = -\frac{1}{36}x(x-2)(x-3)(x-5) - \frac{1}{30}(x+1)(x-2)(x-3)(x-5) - \frac{1}{12}(x+1)x(x-2)(x-5) - \frac{1}{180}(x+1)x(x-2)(x-3);$

Newtonův tvar: $p_4(x) = -2 + 3(x+1) - \frac{35}{30}(x+1)x + \frac{1}{2}(x+1)x(x-2) - \frac{3}{20}(x+1)x(x-2)(x-3).$



5.2. Interpolační splajny



Cíle

Viděli jsme, že graf interpolačního polynomu může nepříjemně oscilovat. Tato situace nastává při předsání většího počtu dat, protože interpolační polynom je pak vysokého stupně. Zdá se proto rozumné při řešení interpolační úlohy použít funkci, která bude *po částech* polynomem nízkého stupně, jejíž jednotlivé části budou na sebe navazovat dostatečně hladce. Takovým funkcím se říká *splajn* (z angl. „spline”). Ukážeme dva nejčastěji používané splajny: lineární a kubický.



Předpokládané znalosti

Interpolační polynom. Spojitost derivace. Řešení soustav lineárních rovnic.



Výklad

Abychom se vyhnuli komplikacím při popisu, budeme předpokládat, že uzly interpolace tvoří rostoucí posloupnost, tzn. $x_0 < x_1 < \dots < x_n$. Vzdálenost dvou sousedních uzlů označíme h_i , tj. $h_i = x_i - x_{i-1}$, $i = 1, \dots, n$.

5.2.1. Lineární splajn

Definice 5.2.1.

Lineárním splajnem nazýváme funkci s_1 , která je spojitá na intervalu $\langle x_0, x_n \rangle$ a na každém podintervalu $\langle x_{i-1}, x_i \rangle$, $i = 1, \dots, n$, je polynomem prvního stupně.

Lineární *interpolační* splajn je řešením úlohy (5.0.1), tzn. že pro něj platí $s_1(x_i) = f_i$, $i = 0, \dots, n$. Můžeme ho zapsat po částech pro $i = 1, \dots, n$:

$$s_1(x) = f_{i-1}(1-t) + f_i t, \quad t = (x - x_{i-1})/h_i, \quad x \in \langle x_{i-1}, x_i \rangle. \quad (5.2.1)$$

Grafem lineárního splajnu je *lomená čára*.

Příklad 5.2.1. Mějme dány uzly $x_0 = -2$, $x_1 = -1$, $x_2 = 1$, $x_3 = 2$ a funkční hodnoty $f_0 = 10$, $f_1 = 4$, $f_2 = 6$, $f_3 = 3$. Napište lineární interpolační splajn.

Řešení: Zapišeme jej pomocí předpisu (5.2.1):

$$s_1(x) = \begin{cases} 10\varphi_1(t) + 4\varphi_2(t), & t = x + 2 & \text{pro } x \in \langle -2, -1 \rangle, \\ 4\varphi_1(t) + 6\varphi_2(t), & t = (x + 1)/2 & \text{pro } x \in \langle -1, 1 \rangle, \\ 6\varphi_1(t) + 3\varphi_2(t), & t = x - 1 & \text{pro } x \in \langle 1, 2 \rangle, \end{cases}$$

kde $\varphi_1(t) = 1 - t$, $\varphi_2(t) = t$. Graf je znázorněn na obrázku 5.2.1.

5.2.2. Kubický splajn

Definice 5.2.2.

Kubickým splajnem nazýváme funkci s_3 , která má na intervalu $\langle x_0, x_n \rangle$ dvě spojitě derivace a na každém podintervalu $\langle x_{i-1}, x_i \rangle$, $i = 1, \dots, n$, je polynomem třetího stupně.

Kubický *interpolační* splajn, je řešení interpolační úlohy (5.0.1). Jeho konstrukce je složitější než u lineárního splajnu. Vyjdeme opět z vyjádření po částech pro $i = 1, \dots, n$:

$$\begin{aligned} s_3(x) = & f_{i-1}(1 - 3t^2 + 2t^3) + f_i(3t^2 - 2t^3) \\ & + m_{i-1}h_i(t - 2t^2 + t^3) + m_i h_i(-t^2 + t^3), \end{aligned} \quad (5.2.2)$$

kde $t = (x - x_{i-1})/h_i$ a $x \in \langle x_{i-1}, x_i \rangle$. Tento předpis je navržen tak, aby parametry f_{i-1} , f_i a m_{i-1} , m_i měly význam funkčních hodnot a hodnot první derivace v krajních bodech intervalu $\langle x_{i-1}, x_i \rangle$, tj. platí

$$s_3(x_{i-1}) = f_{i-1}, \quad s_3(x_i) = f_i, \quad (5.2.3)$$

$$s'_3(x_{i-1}) = m_{i-1}, \quad s'_3(x_i) = m_i. \quad (5.2.4)$$

O splnění rovností (5.2.3) a (5.2.4) se můžeme přesvědčit dosazením x_{i-1} a x_i do (5.2.2) a do první derivace s'_3 , kterou vyjádříme z (5.2.2) podle pravidla o derivování složené funkce:

$$\begin{aligned} s'_3(x) &= f_{i-1}(-6t + 6t^2)/h_i + f_i(6t - 6t^2)/h_i \\ &\quad + m_{i-1}(1 - 4t + 3t^2) + m_i(-2t + 3t^2). \end{aligned} \quad (5.2.5)$$

Předpis (5.2.2) zaručuje spojitost první derivace s'_3 na celém intervalu $\langle x_0, x_n \rangle$ pro libovolné hodnoty m_i . Spojitost druhé derivace vynutíme speciální volbou m_i . Budeme požadovat

$$\lim_{x \rightarrow x_i^-} s''_3(x) = \lim_{x \rightarrow x_i^+} s''_3(x) \quad (5.2.6)$$

ve vnitřních uzlech x_i , $i = 1, \dots, n-1$. Potřebný výraz pro druhou derivaci vypočteme z (5.2.5) opět podle pravidla o derivování složené funkce:

$$\begin{aligned} s''_3(x) &= f_{i-1}(-6 + 12t)/h_i^2 + f_i(6 - 12t)/h_i^2 \\ &\quad + m_{i-1}(-4 + 6t)/h_i + m_i(-2 + 6t)/h_i. \end{aligned} \quad (5.2.7)$$

Levou stranu v (5.2.6) vyjádříme z (5.2.7) pro $t = 1$:

$$\lim_{x \rightarrow x_i^-} s''_3(x) = 6f_{i-1}/h_i^2 - 6f_i/h_i^2 + 2m_{i-1}/h_i + 4m_i/h_i. \quad (5.2.8)$$

Pravou stranu v (5.2.6) vyjádříme z (5.2.7) pro $t = 0$, když současně posuneme indexování:

$$\lim_{x \rightarrow x_i^+} s''(x) = -6f_i/h_{i+1}^2 + 6f_{i+1}/h_{i+1}^2 - 4m_i/h_{i+1} - 2m_{i+1}/h_{i+1}. \quad (5.2.9)$$

Dosadíme-li (5.2.8) a (5.2.9) do (5.2.6), dostaneme po jednoduché úpravě

$$\begin{aligned} h_{i+1}m_{i-1} + 2(h_{i+1} + h_i)m_i + h_im_{i+1} = \\ 3 \left[-\frac{h_{i+1}}{h_i}f_{i-1} + \left(\frac{h_{i+1}}{h_i} - \frac{h_i}{h_{i+1}} \right) f_i + \frac{h_i}{h_{i+1}}f_{i+1} \right], \quad i = 1, \dots, n-1. \end{aligned} \quad (5.2.10)$$

Tyto rovnosti tvoří soustavu $n - 1$ rovnic pro $n + 1$ neznámých m_i , $i = 0, 1, \dots, n$.

Abychom dostali jediné řešení, určíme m_0 a m_n například jako přibližné derivace:

$$m_0 = \frac{f_1 - f_0}{h_1}, \quad m_n = \frac{f_n - f_{n-1}}{h_n}. \quad (5.2.11)$$

Příklad 5.2.2. Mějme dány uzly $x_0 = -2$, $x_1 = -1$, $x_2 = 1$, $x_3 = 2$ a funkční hodnoty $f_0 = 10$, $f_1 = 4$, $f_2 = 6$, $f_3 = 3$. Napište kubický interpolační splajn.

Řešení: Nejdříve vypočítáme parametry m_i , $i = 0, 1, 2, 3$. Podle (5.2.11) je

$$m_0 = \frac{4 - 10}{-1 + 2} = -6, \quad m_3 = \frac{3 - 6}{2 - 1} = -3.$$

Soustava (5.2.10) má dvě rovnice:

$$\begin{aligned} 2(h_2 + h_1)m_1 + h_1m_2 &= 3 \left[-\frac{h_2}{h_1}f_0 + \left(\frac{h_2}{h_1} - \frac{h_1}{h_2} \right) f_1 + \frac{h_1}{h_2}f_2 \right] - h_2m_0, \\ h_3m_1 + 2(h_3 + h_2)m_2 &= 3 \left[-\frac{h_3}{h_2}f_1 + \left(\frac{h_3}{h_2} - \frac{h_2}{h_3} \right) f_2 + \frac{h_2}{h_3}f_3 \right] - h_2m_3, \end{aligned}$$

které můžeme psát jako

$$\begin{pmatrix} 6 & 1 \\ 1 & 6 \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \begin{pmatrix} -21 \\ -9 \end{pmatrix}.$$

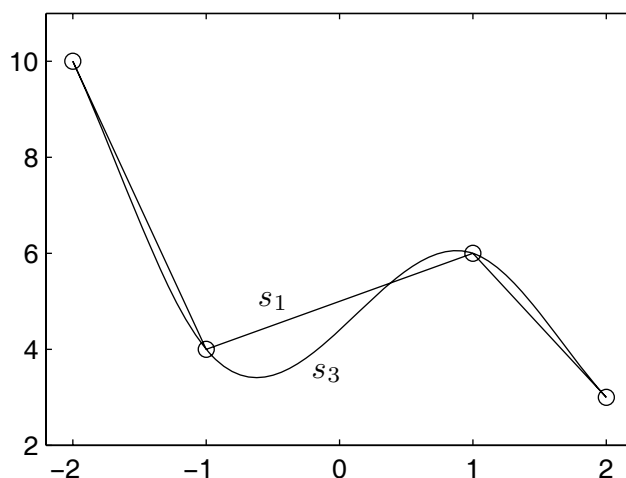
Vyřešením dostaneme $m_1 = -\frac{234}{70}$, $m_2 = -\frac{66}{70}$. Výsledný splajn zapíšeme podle (5.2.2) po částech:

$$s_3(x) = \begin{cases} 10\varphi_1(t) + 4\varphi_2(t) - 6\varphi_3(t) - \frac{117}{35}\varphi_4(t), & t = x + 2 \text{ pro } x \in \langle -2, -1 \rangle, \\ 4\varphi_1(t) + 6\varphi_2(t) - \frac{234}{35}\varphi_3(t) - \frac{66}{35}\varphi_4(t), & t = (x + 1)/2 \text{ pro } x \in \langle -1, 1 \rangle, \\ 6\varphi_1(t) + 3\varphi_2(t) - \frac{33}{35}\varphi_3(t) - 3\varphi_4(t), & t = x - 1 \text{ pro } x \in \langle 1, 2 \rangle, \end{cases}$$

kde $\varphi_1(t) = 1 - 3t^2 + 2t^3$, $\varphi_2(t) = 3t^2 - 2t^3$, $\varphi_3(t) = t - 2t^2 + t^3$, $\varphi_4(t) = -t^2 + t^3$.

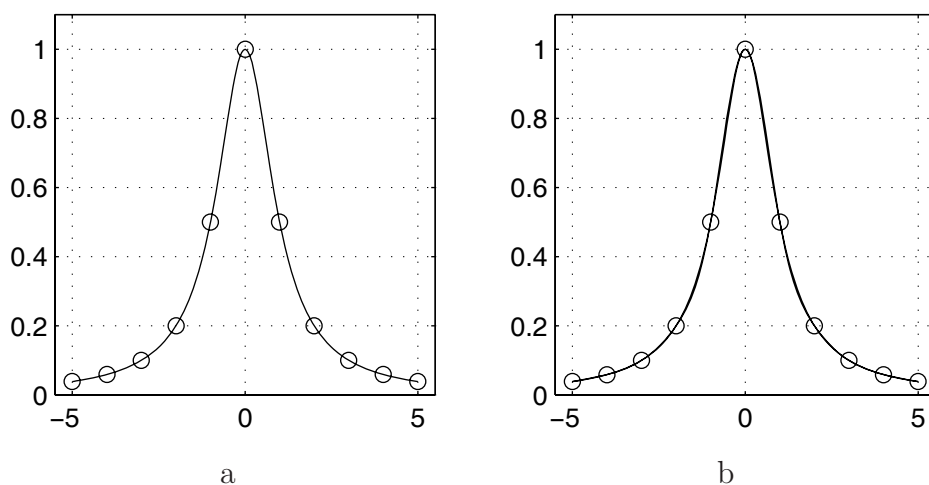
Graf je znázorněn na obrázku 5.2.1.

Příklad 5.2.3. (Rungeho příklad, pokračování) Nakreslíme graf interpolačního kubického splajnu pro funkci f a uzly x_i z příkladu 5.1.4. a porovnáme ho s grafem interpolačního polynomu.



Obrázek 5.2.1: Graf lineárního (s_1) a kubického interpolačního (s_3) splajnu.

Řešení: Na obrázku 5.2.2 vidíme, že splajn s_3 neosciluje a je proto mnohem lepší aproximací interpolované funkce f než interpolační polynom p_{10} , viz obrázek 5.1.2.b.



Obrázek 5.2.2: a) Funkce f ; b) Kubický interpolační splajn s_3 .

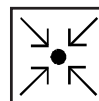


Kontrolní otázky

Otázka 1. Co je to splajn? Jak se definuje a počítá splajn lineární a kubický?

Otázka 2. Jak se chovají při interpolaci splajny v porovnání s polynomy?

Úlohy k samostatnému řešení



1. Pro uzly $x_0 = -1, x_1 = 0, x_2 = 2, x_3 = 3, x_4 = 5$ a funkční hodnoty $f_0 = -2, f_1 = 1, f_2 = 0, f_3 = 2, f_4 = -1$ sestavte lineární interpolační splajn.
2. Pro stejná data sestavte kubický interpolační splajn.

Výsledky úloh k samostatnému řešení



1.

$$s_1(x) = \begin{cases} -2\varphi_1(t) + \varphi_2(t), & t = x + 1 & \text{pro } x \in \langle -1, 0 \rangle, \\ \varphi_1(t), & t = x/2 & \text{pro } x \in \langle 0, 2 \rangle, \\ 2\varphi_2(t), & t = x - 2 & \text{pro } x \in \langle 2, 3 \rangle, \\ 2\varphi_1(t) - 1\varphi_2(t), & t = (x - 3)/2 & \text{pro } x \in \langle 3, 5 \rangle, \end{cases}$$

kde $\varphi_1(t) = 1 - t, \varphi_2(t) = t$.

2. Krajní parametry jsou $m_0 = 3, m_4 = -\frac{3}{2}$, ostatní dostaneme ze soustavy

$$\begin{pmatrix} 6 & 1 & 0 \\ 1 & 6 & 2 \\ 0 & 2 & 6 \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix} = \begin{pmatrix} \frac{21}{2} \\ \frac{21}{2} \\ 9 \end{pmatrix},$$

takže $m_1 = \frac{97}{62}, m_2 = \frac{69}{62}, m_3 = \frac{35}{31}$ a konečně

$$s_3(x) = \begin{cases} -2\varphi_1(t) + \varphi_2(t) + 3\varphi_3(t) + \frac{97}{62}\varphi_4(t), & t = x + 1 \text{ pro } x \in \langle -1, 0 \rangle, \\ \varphi_1(t) + \frac{97}{31}\varphi_3(t) + \frac{69}{31}\varphi_4(t), & t = x/2 \text{ pro } x \in \langle 0, 2 \rangle, \\ 2\varphi_2(t) + \frac{69}{62}\varphi_3(t) + \frac{35}{31}\varphi_4(t), & t = x - 2 \text{ pro } x \in \langle 2, 3 \rangle, \\ 2\varphi_1(t) - \varphi_2(t) + \frac{70}{31}\varphi_3(t) - 3\varphi_4(t), & t = (x - 3)/2 \text{ pro } x \in \langle 3, 5 \rangle, \end{cases}$$

kde $\varphi_1(t) = 1 - 3t^2 + 2t^3, \varphi_2(t) = 3t^2 - 2t^3, \varphi_3(t) = t - 2t^2 + t^3, \varphi_4(t) = -t^2 + t^3$.

5.3. Aproximace metodou nejmenších čtverců



Cíle

V mnoha situacích, v nichž je potřeba danou funkci f nahradit funkcí „jednodušší“, je nevhodné nebo vůbec nelze použít interpolaci. Jsou-li například v uzlech zadány nepřesné hodnoty, přenáší se tato nepřesnost i na interpolant. Interpolace je nepoužitelná, jestliže je požadován jistý charakter aproximující funkce a přitom žádná funkce tohoto charakteru není interpolantem. V těchto případech je rozumné použít metodu nejmenších čtverců.



Předpokládané znalosti

Lineární závislost a nezávislost. Určení minima funkce pomocí derivace. Řešení soustav lineárních rovnic.



Výklad

Začneme příkladem.

Příklad 5.3.1. Mějme dány uzly $x_0 = -2$, $x_1 = -1$, $x_2 = 1$, $x_3 = 2$ a funkční hodnoty $f_0 = 10$, $f_1 = 4$, $f_2 = 6$, $f_3 = 3$. Najděte přímku

$$\varphi(x) = c_1 + c_2x, \quad (5.3.1)$$

která je „blízko“ předepsaným hodnotám.

Řešení: Nejdříve se musíme rozhodnout jak chápat slovo „blízko“. Už jsme to vlastně řekli, když jsme popisovali smysl přibližných rovností v aproximační úloze (5.0.2). Přímku φ určíme tak, aby minimalizovala součet druhých mocnin odchylek $\sum_{i=0}^3 (\varphi(x_i) - f_i)^2$. Jestliže sem dosadíme (5.3.1), dostaneme úlohu na minimalizaci funkce dvou proměnných $\Psi(c_1, c_2) = \sum_{i=0}^3 (c_1 + c_2x_i - f_i)^2$. Minimum

c_1^* , c_2^* vyhovuje rovnicím

$$\frac{\partial \Psi}{\partial c_1}(c_1^*, c_2^*) = 0, \quad \frac{\partial \Psi}{\partial c_2}(c_1^*, c_2^*) = 0.$$

Po vyjádření parciálních derivací dostáváme

$$2 \sum_{i=0}^3 (c_1^* + c_2^* x_i - f_i) = 0, \quad 2 \sum_{i=0}^3 (c_1^* + c_2^* x_i - f_i) x_i = 0,$$

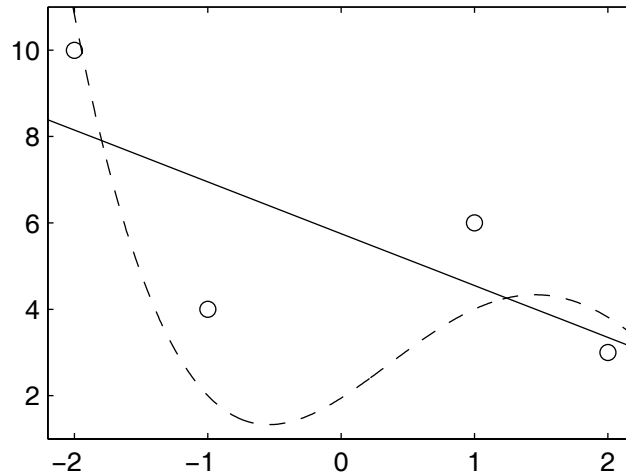
což je soustava lineárních rovnic

$$\begin{pmatrix} \sum_{i=0}^3 1 & \sum_{i=0}^3 x_i \\ \sum_{i=0}^3 x_i & \sum_{i=0}^3 x_i^2 \end{pmatrix} \begin{pmatrix} c_1^* \\ c_2^* \end{pmatrix} = \begin{pmatrix} \sum_{i=0}^3 f_i \\ \sum_{i=0}^3 f_i x_i \end{pmatrix}, \text{ tj. } \begin{pmatrix} 4 & 0 \\ 0 & 10 \end{pmatrix} \begin{pmatrix} c_1^* \\ c_2^* \end{pmatrix} = \begin{pmatrix} 23 \\ -12 \end{pmatrix}.$$

Tato soustava má jediné řešení $c_1^* = \frac{23}{4}$, $c_2^* = -\frac{6}{5}$, takže hledaná přímka $\varphi^* = \varphi$ je určena předpisem

$$\varphi^*(x) = \frac{23}{4} - \frac{6}{5}x. \quad (5.3.2)$$

Její graf je znázorněn na obrázku 5.3.1. □



Obrázek 5.3.1: Aproximace metodou nejmenších čtverců; přímka (5.3.2) plně; funkce (5.3.8) čárkovaně.

Postup z příkladu nyní zobecníme. Budeme předpokládat, že je dán systém funkcí $\varphi_j = \varphi_j(x)$, $j = 1, \dots, m$ a budeme uvažovat všechny funkce ve tvaru

$$\varphi(x) = c_1\varphi_1(x) + \dots + c_m\varphi_m(x) = \sum_{j=1}^m c_j\varphi_j(x), \quad (5.3.3)$$

kde koeficienty c_1, \dots, c_m jsou libovolná čísla. Funkci φ^* , pro niž platí

$$\sum_{i=1}^n (\varphi^*(x_i) - f_i)^2 \leq \sum_{i=1}^n (\varphi(x_i) - f_i)^2 \quad \forall \varphi \quad (5.3.4)$$

nazýváme *aproximací podle metody nejmenších čtverců*. Její koeficienty c_1^*, \dots, c_m^* určíme jako minimum funkce

$$\Psi(c_1, \dots, c_m) = \sum_{i=1}^n \left(\sum_{j=1}^m c_j \varphi_j(x_i) - f_i \right)^2, \quad (5.3.5)$$

které vyhovuje rovnicím

$$\frac{\partial \Psi}{\partial c_k}(c_1^*, \dots, c_m^*) = 0, \quad k = 1, \dots, m. \quad (5.3.6)$$

Vyjádříme-li parciální derivace

$$\frac{\partial \Psi}{\partial c_k} = 2 \sum_{i=1}^n \left(\sum_{j=1}^m c_j \varphi_j(x_i) - f_i \right) \varphi_k(x_i)$$

a dosadíme je do (5.3.6), dostaneme po jednoduché úpravě soustavu lineárních rovnic

$$\sum_{j=1}^m \left(\sum_{i=1}^n \varphi_j(x_i) \varphi_k(x_i) \right) c_j^* = \sum_{i=1}^n f_i \varphi_k(x_i), \quad k = 1, \dots, m. \quad (5.3.7)$$

Soustava (5.3.7) se nazývá *soustava normálních rovnic*.

Věta 5.3.1.

Nechť jsou dány vzájemně různé uzly x_i a funkční hodnoty f_i , $i = 0, \dots, n$. Nechť je dán systém funkcí φ_j , $j = 1, \dots, m$, které jsou lineárně nezávislé. Potom existuje jediná funkce φ^* , která splňuje (5.3.4) a její koeficienty c_1^*, \dots, c_m^* jsou řešením soustavy normálních rovnic (5.3.7).

Důkaz: V bodě c_1^*, \dots, c_m^* , který vyhovuje rovnicím (5.3.6), nabývá funkce Ψ minima, jestliže matice druhých derivací je symmetrická a pozitivně definitní (kladná). Druhé derivace jsou určeny vzorcem

$$\frac{\partial^2 \Psi}{\partial c_k \partial c_l} = 2 \sum_{i=1}^n \varphi_l(x_i) \varphi_k(x_i),$$

odkud je symetrie zřejmá na první pohled (prohozením indexů k a l se nic nezmění). Necht' d_1, \dots, d_m jsou čísla ne všechny současně nulová. Potom

$$\sum_{k=1}^m \sum_{l=1}^m d_k d_l \frac{\partial^2 \Psi}{\partial c_k \partial c_l} = 2 \sum_{i=1}^n \left(\sum_{l=1}^m d_l \varphi_l(x_i) \right) \left(\sum_{k=1}^m d_k \varphi_k(x_i) \right) = 2 \sum_{i=1}^n \tilde{\varphi}(x_i)^2 > 0,$$

kde $\tilde{\varphi}(x) = \sum_{k=1}^m d_k \varphi_k(x)$, takže matice druhých derivací je pozitivně definitní. Odtud také plyne, že matice soustavy normálních rovnic je regulární, což znamená, že existuje její jediné řešení c_1^*, \dots, c_m^* , které určuje jedinou funkci φ^* . \square

Při aproximaci metodou nejmenších čtverců se musíme nejdříve rozhodnout pro nějaký lineárně nezávislý systém funkcí φ_j , $j = 1, \dots, m$. Poté stačí sestavit a vyřešit soustavu normálních rovnic (5.3.7).

Příklad 5.3.2. Napište normální soustavu lineárních rovnic odpovídající systému funkcí

$$\varphi_1(x) = e^{-x}, \quad \varphi_2(x) = \sin x.$$

Aproximujte data z příkladu 5.3.1.

Řešení: Obecně má soustava normálních rovnic tvar

$$\begin{pmatrix} \sum_{i=0}^n e^{-2x_i} & \sum_{i=0}^n e^{-x_i} \sin x_i \\ \sum_{i=0}^n e^{-x_i} \sin x_i & \sum_{i=0}^n \sin^2 x_i \end{pmatrix} \begin{pmatrix} c_1^* \\ c_2^* \end{pmatrix} = \begin{pmatrix} \sum_{i=0}^n f_i e^{-x_i} \\ \sum_{i=0}^n f_i \sin x_i \end{pmatrix}.$$

Po dosazení dostaneme

$$\begin{pmatrix} 62.1409 & -8.5736 \\ -8.5736 & 3.0698 \end{pmatrix} \begin{pmatrix} c_1^* \\ c_2^* \end{pmatrix} = \begin{pmatrix} 87.3770 \\ -4.6821 \end{pmatrix}$$

a odtud vypočítáme $c_1^* = 1.9452$, $c_2^* = 3.9076$, tj.

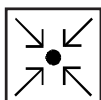
$$\varphi^*(x) = 1.9452e^{-x} + 3.9076 \sin x. \quad (5.3.8)$$

Graf je znázorněn na obrázku 5.3.1.



Kontrolní otázky

- Otázka 1. Kdy je vhodné použít metodu nejmenších čtverců?
- Otázka 2. Graficky znázorníte smysl výrazu pro součet druhých mocnin odchylek?
- Otázka 3. Co je to normální soustava lineárních rovnic a jak vznikne?
- Otázka 4. Co se stane, když v (5.3.3) a (5.3.4) bude $m = n + 1$?



Úlohy k samostatnému řešení

- Napište soustavu normálních lineárních rovnic pro systém funkcí $\varphi_1(x) = 1$, $\varphi_2(x) = x$, $\varphi_3(x) = x^2$.
- Data $x_0 = -1$, $x_1 = 0$, $x_2 = 2$, $x_3 = 3$, $x_4 = 5$ a $f_0 = -2$, $f_1 = 1$, $f_2 = 0$, $f_3 = 2$, $f_4 = -1$ aproximujte metodou nejmenších čtverců pomocí systému funkcí z předchozí úlohy.



Výsledky úloh k samostatnému řešení

- Normální soustava má tvar:

$$\begin{pmatrix} \sum_{i=0}^n 1 & \sum_{i=0}^n x_i & \sum_{i=0}^n x_i^2 \\ \sum_{i=0}^n x_i & \sum_{i=0}^n x_i^2 & \sum_{i=0}^n x_i^3 \\ \sum_{i=0}^n x_i^2 & \sum_{i=0}^n x_i^3 & \sum_{i=0}^n x_i^4 \end{pmatrix} \begin{pmatrix} c_1^* \\ c_2^* \\ c_3^* \end{pmatrix} = \begin{pmatrix} \sum_{i=0}^n f_i \\ \sum_{i=0}^n f_i x_i \\ \sum_{i=0}^n f_i x_i^2 \end{pmatrix}.$$

- $\varphi^*(x) = -0.2835x^2 + 1.2359x - 0.0130$.



Shrnutí lekce

Ukázali jsme základní postupy pro aproximaci funkcí (dat) pomocí interpolace a metody nejmenších čtverců.

6. Numerické integrování a derivování

Průvodce studiem



Při řešení různých úloh je potřeba nalézt hodnotu určitého integrálu. Některé integrály lze vypočítat snadno pomocí tabulek a klasických integračních metod jako je per partes nebo substituce. Tak například integrály

$$\int_0^1 e^x dx \quad \text{a} \quad \int_0^\pi \cos x dx$$

mají hodnotu $e - 1$ a 0. Často však stačí malá změna v integrované funkci a klasický výpočet se stane složitým nebo dokonce nemožným. Pokuste se vypočítat integrály

$$\int_0^1 e^{x^2} dx \quad \text{a} \quad \int_0^\pi \cos x^2 dx$$

a uvidíte jak daleko se vám podaří dojít! Klasické integrační metody nelze použít vůbec, jestliže integrovaná funkce je dána tabulkou (např. to může být výsledek měření nebo předchozího výpočtu).

V této kapitole se budeme zabývat metodami numerickými pro obecně zadanou úlohu: předpokládáme, že je dána spojitá funkce f na intervalu $\langle a, b \rangle$ a máme vypočítat hodnotu určitého integrálu

$$I = \int_a^b f(x) dx. \quad (6.0.1)$$

Z geometrického pohledu představuje číslo I velikost plochy obrazce, který je vymezen grafem funkce f . Numerické metody jsou navrhovány tak, že počítají velikost plochy přibližného obrazce pomocí několika funkčních hodnot. Lze je tedy použít pro jakoukoliv funkci f , dokonce i pro funkci danou tabulkou, a jejich pracnost je relativně malá.

Při numerickém výpočtu derivace je situace podobná.

6.1. Newton-Cotesovy vzorce



Cíle

Odvodíme Newton-Cotesovy vzorce a ukážeme jejich použití.



Předpokládané znalosti

Interpolační polynom. Interpolační chyba.



Výklad

Naším cílem je navrhnout vzorce pro přibližný výpočet integrálu (6.0.1). Ze základního kurzu integrálního počtu víme, že snadno lze integrovat polynomy. Budeme proto postupovat tak, že k funkci f sestavíme interpolační polynom p_n a ten integrujeme místo f , tj.

$$I = \int_a^b f(x) dx \approx \int_a^b p_n(x) dx. \quad (6.1.1)$$

Pro jednoduchost uvažujme na intervalu $\langle a, b \rangle$ ekvidistantní uzly

$$x_i = a + i\tau, \quad i = 0, 1, \dots, n,$$

kde $\tau = (b-a)/n$. Interpolační polynom k funkci f můžeme zapsat v Lagrangeově tvaru (viz odstavec 5.1.1), tj.

$$p_n(x) = \sum_{i=0}^n f(x_i) \varphi_i(x), \quad \text{kde} \quad \varphi_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j}.$$

Dosazením výrazu pro p_n do pravé strany přibližné rovnosti (6.1.1) dostaneme *Newton-Cotesovy vzorce*:

$$\int_a^b f(x) dx \approx \sum_{i=0}^n w_i f(x_i), \quad (6.1.2)$$

kde

$$w_i = \int_a^b \varphi_i(x) dx, \quad i = 0, 1, \dots, n \quad (6.1.3)$$

jsou *integrační váhy*. Abychom lépe ilustrovali tento postup, odvodíme základní integrační pravidla.

a) Obdélníkové pravidlo. V tomto případě položíme $n = 0$, $x_0 = \frac{a+b}{2}$ a za interpolační polynom vezmeme konstantní funkci $p_0(x) = f(\frac{a+b}{2})$. Proto

$$\int_a^b f(x) dx \approx (b-a)f\left(\frac{a+b}{2}\right) = I_{Obd}. \quad (6.1.4)$$

b) Lichoběžníkové pravidlo dostaneme pro $n = 1$ a $x_0 = a$, $x_1 = b$. Interpolační polynom je přímka

$$p_1(x) = \frac{x-b}{a-b}f(a) + \frac{x-a}{b-a}f(b).$$

Po její integraci je

$$\int_a^b f(x) dx \approx \frac{b-a}{2}[f(a) + f(b)] = I_{Lich}. \quad (6.1.5)$$

c) Simpsonovo pravidlo vznikne pro $n = 2$ a $x_0 = a$, $x_1 = \frac{a+b}{2}$, $x_2 = b$, kdy je interpolační polynom kvadratická funkce. Vypočtěme integrační váhy w_0 , w_1 a w_2 . Dostáváme

$$\begin{aligned} w_0 &= \int_a^b \varphi_0(x) dx = \int_a^b \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} dx \\ &= \int_{-1}^1 \frac{t(t-1)}{2} \frac{b-a}{2} dt = \frac{b-a}{6} \end{aligned}$$

s použitím substituce $x = \frac{b-a}{2}t + \frac{a+b}{2}$. Podobně vypočítáme $w_1 = \frac{4}{6}(b-a)$ a $w_2 = \frac{1}{6}(b-a)$, takže

$$\int_a^b f(x) dx \approx \frac{b-a}{6}[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)] = I_{Simps}. \quad (6.1.6)$$

Příklad 6.1.1. Pomocí Newton-Cotesových vzorců (6.1.4), (6.1.5) a (6.1.6) vypočtete přibližnou hodnotu integrálu

$$I = \int_{-1}^1 e^x dx.$$

Řešení: Máme $a = -1$, $b = 1$ a $f(x) = e^x$. Podle obdélníkového pravidla (6.1.4) je

$$I_{Obd} = 2e^0 = 2.$$

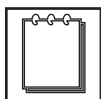
Lichoběžníkové pravidlo (6.1.5) dává

$$I_{Lich} = \frac{2}{2}[e^{-1} + e^1] = 3.086161$$

a konečně pomocí Simpsonova pravidla (6.1.6) vypočteme

$$I_{Simps} = \frac{2}{6}[e^{-1} + 4e^0 + e^1] = 2.362054.$$

Poznamenejme ještě, že přesná hodnota integrálu je $I = e^1 - e^{-1} = 2.350402$.



Poznámka

Nejpřesnější hodnoty jsme dosáhli pomocí Simpsonova pravidla. To nás může vést k domněnce, že ještě přesnějších hodnot dosáhneme, použijeme-li Newton-Cotesovy vzorce odvozené z interpolačních polynomů vyšších stupňů. Obecně to však není pravda. V Rungeho příkladu (příklad 5.1.4.) jsme viděli, že interpolační polynom vysokého stupně může být velice špatnou aproximací. Newton-Cotesovy vzorce dávají v takovém případě nesmyslné výsledky.

Následující věta ukazuje vyjádření integrační chyby.

Věta 6.1.1.

(i) Nechť funkce f má na intervalu $\langle a, b \rangle$ spojitou druhou derivaci. Pak platí:

$$\begin{aligned} I - I_{Obd} &= \frac{f''(\xi)}{24}(b-a)^3, \\ I - I_{Lich} &= -\frac{f''(\xi)}{12}(b-a)^3. \end{aligned} \quad (6.1.7)$$

(ii) Nechť funkce f má na intervalu $\langle a, b \rangle$ spojitou čtvrtou derivaci. Pak platí:

$$I - I_{Simps} = -\frac{f^{(4)}(\xi)}{90}(b-a)^3.$$

Ve všech případech je ξ blíže neurčený bod z intervalu (a, b) .

Důkaz: Důkaz provedeme pouze pro lichoběžníkové pravidlo. Vyjádření interpolační chyby z věty 5.1.2. má pro interpolační polynom p_1 s uzly $x_0 = a$ a $x_1 = b$ tvar:

$$f(x) - p_1(x) = \frac{f''(\xi)}{2}(x-a)(x-b).$$

Integrací (a užitím věty o střední hodnotě integrálního počtu) dostaneme

$$\begin{aligned} I - I_{Lich} &= \frac{f''(\xi)}{2} \int_a^b (x-a)(x-b) dx \\ &= \frac{f''(\xi)}{2} \int_0^{b-a} t(t+a-b) dx = -\frac{f''(\xi)}{12}(b-a)^3. \end{aligned}$$

Použili jsme substituci $t = x - a$. □

Kontrolní otázky



Otázka 1. Jak vzniknou Newton-Cotesovy vzorce?

Otázka 2. Jaké jsou základní pravidla pro numerický výpočet integrálu?

Otázka 3. Znázorněte graficky smysl obdélníkového a lichoběžníkového pravidla.

Otázka 4. Proč není rozumné používat Newton-Cotesovy vzorce odvozené z interpolačních polynomů vysokého stupně?

Úlohy k samostatnému řešení



1. Pomocí obdélníkového, lichoběžníkového a Simpsonova pravidla vypočítejte přibližné hodnoty integrálů

$$\text{a) } \int_0^1 e^{x^2} dx; \quad \text{b) } \int_0^\pi \cos x^2 dx.$$

Výsledky úloh k samostatnému řešení



1. a) $I_{Obd} = 1.284025$, $I_{Lich} = 1.859141$, $I_{Simps} = 1.475731$; b) $I_{Obd} = 1.961189$, $I_{Lich} = -0.675916$, $I_{Simps} = 1.082154$.

6.2. Složené vzorce



Cíle

V předchozím odstavci jsme odvodili vzorce pro přibližný výpočet interálu integrací interpolačního polynomu. Zároveň jsme upozornili na to, že vzorce odvozené z polynomů vysokého stupně mohou dávat nesmyslné výsledky. Pro přesnější výpočty integrálů se proto používají složené vzorce, které dostaneme „slepením“ základních integračních pravidel.



Předpokládané znalosti

Obdélníkové, lichoběžníkové a Simpsonovo pravidlo. Chyby těchto pravidel.



Výklad

Interval $\langle a, b \rangle$ rozdělíme na m stejně dlouhých dílků s krokem $h = (b - a)/m$, $m \geq 2$, tj. použijeme uzly

$$x_i = a + ih = a + \frac{i}{m}(b - a), \quad i = 0, 1, \dots, m.$$

Integrál rozložíme následovně:

$$I = \int_a^b f(x) dx = \sum_{i=1}^m \int_{x_{i-1}}^{x_i} f(x) dx. \quad (6.2.1)$$

Jestliže každý dílčí integrál nahradíme pomocí obdélníkového nebo lichoběžníkového pravidla, dostaneme odpovídající pravidlo složené.

a) Složené obdélníkové pravidlo jsme již odvodili v odstavci 1.1. Má tvar

$$\begin{aligned} I_{SO} &= h \left[f\left(\frac{x_0 + x_1}{2}\right) + f\left(\frac{x_1 + x_2}{2}\right) + \dots + f\left(\frac{x_{m-1} + x_m}{2}\right) \right] \\ &= h \sum_{i=1}^m f\left(\frac{x_{i-1} + x_i}{2}\right). \end{aligned} \quad (6.2.2)$$

b) Složené lichoběžníkové pravidlo. V (6.2.1) provedeme náhradu podle vzorce (6.1.5), tj.

$$\int_{x_{i-1}}^{x_i} f(x) dx \approx \frac{h}{2}[f(x_{i-1}) + f(x_i)],$$

a dostaneme

$$\begin{aligned} I_{SL} &= h \left[\frac{1}{2}f(x_0) + f(x_1) + \dots + f(x_{m-1}) + \frac{1}{2}f(x_m) \right] \\ &= \frac{1}{2}h \left[f(x_0) + 2 \sum_{i=1}^{m-1} f(x_i) + f(x_m) \right]. \end{aligned} \quad (6.2.3)$$

c) Složené Simpsonovo pravidlo. V tomto případě rozdělíme interval $\langle a, b \rangle$ na $2m$ stejně dlouhých dílků, $m \geq 2$, s krokem $h = (b - a)/(2m)$, takže budeme mít lichý počet uzlů $x_i = a + ih$, $i = 0, 1, \dots, 2m$. Simpsonovo pravidlo (6.1.6) použijeme na každém intervalu $\langle x_{2i-2}, x_{2i} \rangle$, $i = 1, \dots, m$, což dává

$$I = \int_a^b f(x) dx = \sum_{i=1}^m \int_{x_{2i-2}}^{x_{2i}} f(x) dx \approx \sum_{i=1}^m \frac{2h}{6} [f(x_{2i-2}) + 4f(x_{2i-1}) + f(x_{2i})].$$

Výraz na pravé straně přepíšeme do přehlednějšího tvaru:

$$\begin{aligned} I_{SS} &= \frac{h}{3} [f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + \dots + f(x_{2m})] \\ &= \frac{h}{3} \left[f(x_0) + 4 \sum_{i=1}^m f(x_{2i-1}) + 2 \sum_{i=1}^{m-1} f(x_{2i}) + f(x_{2m}) \right]. \end{aligned} \quad (6.2.4)$$

Příklad 6.2.1. Pomocí pravidel (6.2.2), (6.2.3) a (6.2.4) vypočtete přibližnou hodnotu integrálu

$$I = \int_{-1}^1 e^x dx.$$

Interval $\langle -1, 1 \rangle$ přitom rozdělte na čtyři části.

Řešení: (a) Položíme $m = 4$, takže krok bude $h = 0.5$ a dostaneme uzly $x_0 = -1$, $x_1 = -0.5$, $x_2 = 0$, $x_3 = 0.5$ a $x_4 = 1$. Složené obdélníkové pravidlo (6.2.2) má tvar

$$I_{SO} = 0.5[e^{-0.75} + e^{-0.25} + e^{0.25} + e^{0.75}] \doteq 2.326096.$$

(b) Složené lichoběžníkové pravidlo (6.2.3) vypočítáme pro stejné uzly, tj.

$$I_{SL} = \frac{1}{2} \cdot 0.5[e^{-1} + 2e^{-0.5} + 2e^0 + 2e^{0.5} + e^1] \doteq 2.399166.$$

(c) U složeného Simpsonova pravidla (6.2.4) vezmeme $m = 2$, což znamená, že použijeme stejné dělení intervalu jako v předchozích případech (protože $h = 2/(2m) = 0.5$). Dostaneme

$$I_{SS} = \frac{0.5}{3}[e^{-1} + 4e^{-0.5} + 2e^0 + 4e^{0.5} + e^1] \doteq 2.351195.$$

V následující větě popisujeme řád jednotlivých integračních pravidel.

Věta 6.2.1.

(i) Nechť má funkce f na intervalu $\langle a, b \rangle$ spojitou druhou derivaci. Složené obdélníkové a složené lichoběžníkové pravidlo jsou druhého řádu, tj. platí:

$$|I - I_{SO}| \leq C_1 h^2,$$

$$|I - I_{SL}| \leq C_2 h^2.$$

(ii) Nechť funkce f má na intervalu $\langle a, b \rangle$ spojitou čtvrtou derivaci. Složené Simpsonovo pravidlo je čtvrtého řádu, tj. platí:

$$|I - I_{SS}| \leq C_3 h^4.$$

Konstanty C_1 , C_2 a C_3 jsou blíže neurčená nezáporná čísla, která nezávisí na kroku h .

Důkaz: Omezíme se na lichoběžníkové pravidlo. Stačí si uvědomit, že jednoduché lichoběžníkové pravidlo je ve složeném lichoběžníkovém pravidle obsaženo m -krát. Jestliže tedy chceme vyjádřit chybu, sečteme m -krát výraz pro chybu jednoduchého pravidla (6.1.7). Dostaneme

$$|I - I_{SL}| = \left| \sum_{i=1}^m -\frac{f''(\xi_i)}{12} h^3 \right| \leq \sum_{i=1}^m \left| \frac{f''(\xi_i)}{12} h^3 \right| \leq \frac{M}{12} h^2 \sum_{i=1}^m h = \frac{M(b-a)}{12} h^2,$$

kde jsme použili $M = \max_{\zeta \in \langle a, b \rangle} |f''(\zeta)|$. Vidíme, že $C_2 = M(b-a)/12$. \square

Příklad 6.2.2. Pomocí složených integračních pravidel (6.2.2), (6.2.3) a (6.2.4) vypočtete přibližné hodnoty integrálu

$$I = \int_{-1}^1 e^x dx$$

s krokem $h = 0.5$, $h = 0.25$, $h = 0.125$ a $h = 0.0625$ a porovnejte je s přesnou hodnotou.

Řešení: Přesná hodnota je $I = 2.350402387$. Přibližné hodnoty a jejich porovnání s přesnou hodnotou uvádíme v tabulce 6.2.1. Odtud je vidět, že složené obdélníkové a složené lichoběžníkové pravidlo se chovají podobně, zatímco složené Simpsonovo pravidlo dává výsledky, které jsou přesnější o několik řádů.

Tabulka 6.2.1: Složené integrační vzorce.

h	I_{SO}	$ I - I_{SO} $	I_{SL}	$ I - I_{SL} $	I_{SS}	$ I - I_{SS} $
0.5	2.326096	0.024306	2.399166	0.048764	2.351195	0.000792
0.25	2.344293	0.006110	2.362631	0.012229	2.350453	0.000051
0.125	2.348873	0.001530	2.353462	0.003060	2.350406	0.000003
0.0625	2.350020	0.000383	2.351167	0.000765	2.350402	0.000000

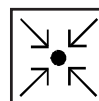
Kontrolní otázky



Otázka 1. Jak se odvozují složené integrační vzorce?

Otázka 2. Jakého řádu jsou základní složená integrační pravidla?

Úlohy k samostatnému řešení



1. Pomocí složeného obdélníkového, lichoběžníkového a Simpsonova pravidla vypočtete přibližné hodnotu integrálů

$$\text{a) } \int_0^1 e^{x^2} dx \quad \text{pro } h = 0.1; \quad \text{b) } \int_0^\pi \cos x^2 dx \quad \text{pro } h = \pi/10.$$

Výsledky úloh k samostatnému řešení



1. a) $I_{SO} = 1.460393$, $I_{SL} = 1.467175$, $I_{SS} = 1.462681$; b) $I_{SO} = 0.553752$, $I_{SL} = 0.588876$, $I_{SS} = 0.566030$.

6.3. Výpočet integrálu se zadanou přesností



Cíle

U integračních pravidel z předchozího odstavce není jasné jak dosáhnout zadané přesnosti. Pro tyto účely použijeme dvojný přepočet, který lze kombinovat s extrapolacími vzorci.



Předpokládané znalosti

Složená integrační pravidla a jejich řád. Taylorův rozvoj.



Výklad

Ve větě 6.2.1. jsme uvedli vzorce pro odhad chyby složených integračních pravidel. Jejich použitím lze určit krok h , který zajistí, aby chyba vypočtené hodnoty integrálu byla menší než je zadaná tolerance. Tento postup má však dva háčky. Především (jak je patrné z důkazu zmíněné věty) výraz pro odhad chyby závisí na derivacích integrované funkce, které se odhadují pracně. Druhý problém spočívá v tom, že výsledné odhady jsou pesimistické, takže integrály jsou počítány mnohem přesněji než je požadováno, což vyžaduje zbytečně velký objem výpočtů. Obvykle se proto postupuje tak, že hledaný integrál vyčísľujeme opakovaně, stále přesněji a ze shody výsledků se usuzuje, zda již byla dosažena požadovaná přesnost. Pro tyto účely se osvědčilo postupné zdvojnásobování počtu dílků, na něž rozdělujeme integrační interval $\langle a, b \rangle$. Hovoříme pak o *dvojném přepočtu*.

Celý postup zapíšeme v podobě algoritmu, kde $I(h)$ označuje přibližnou hodnotu integrálu vypočtenou s krokem h pomocí některého složeného integračního pravidla. Na začátku předpokládáme, že integrační interval $\langle a, b \rangle$ se rozdělí na m úseků. Symbolem ϵ označujeme zadanou přesnost.

Algoritmus (Dvojný přepočet)

Vstup: f, a, b, m, ϵ .

Vypočti $h := (b - a)/m$ a $I(h)$.

Opakuj:

 polož $h := h/2$,

 vypočti $I(h)$,

dokud $|I(2h) - I(h)| > \epsilon$.

Výstup: $I(h)$.

Příklad 6.3.1. Pomocí dvojného přepočtu vypočtete přibližnou hodnotu integrálu

$$I = \int_{-1}^1 e^x dx$$

s přesností $\epsilon = 0.0001$. Použijte složené lichoběžníkové pravidlo a začněte s rozdělením integračního intervalu na $m = 4$ dílků.

Řešení: Výpočet je zaznamenán v tabulce 6.3.1. Jako výsledek dostáváme $I = 2.3504 \pm 0.0001$. □

Tabulka 6.3.1: Dvojný přepočet.

m	h	$I(h)$	$ I(2h) - I(h) $
4	0.5000000	2.3991662	—
8	0.2500000	2.3626313	0.0365349
16	0.1250000	2.3534620	0.0091693
32	0.0625000	2.3511674	0.0022945
64	0.0312500	2.3505936	0.0005737
128	0.0156250	2.3504502	0.0001434
256	0.0078125	2.3504143	0.0000358

V dalším ukážeme efektivnější variantu dvojného přepočtu s extrapolacími vzorci, které nejdříve odvodíme. Na hodnotu $I(h)$ můžeme pohlížet jako na funkci proměnné h , takže pro ni můžeme psát Taylorův rozvoj. Odhady pro chybu podle věty 6.2.1. ukazují, že se v tomto rozvoji budou vyskytovat mocniny h

odpovídající řádu příslušného integračního pravidla a vyšší, tj.

$$I(h) = I + Ch^p + \mathcal{O}(h^r), \quad (6.3.1)$$

kde p je řád a $r > p$. Jestliže vzorec (6.3.1) napíšeme pro $2h$ dostaneme

$$I(2h) = I + 2^p Ch^p + \mathcal{O}(h^r), \quad (6.3.2)$$

protože $\mathcal{O}((2h)^r) = \mathcal{O}(h^r)$. Vyloučíme-li z rovností (6.3.1) a (6.3.2) výraz Ch^p dostáváme

$$I = I(h) + \frac{I(h) - I(2h)}{2^p - 1} + \mathcal{O}(h^r). \quad (6.3.3)$$

Odtud vidíme následující:

1) Veličina

$$I_1(h) = I(h) + \frac{I(h) - I(2h)}{2^p - 1} \quad (6.3.4)$$

je lepší aproximací I než $I(h)$ (je vyššího řádu r).

2) Výraz

$$E(h) = \frac{I(h) - I(2h)}{2^p - 1} \quad (6.3.5)$$

je aproximace chyby přibližné hodnoty integrálu $I(h)$, kterou můžeme vzhledem k bodu 1) použít také pro odhad chyby aproximace $I_1(h)$.

Postup pro zvyšování přesnosti výsledků získaných numerickou metodou podle vzorce (6.3.3) se nazývá *Richardsonova extrapolace*. Pomocí extrapolacních vzorců (6.3.3) a (6.3.4) upravíme algoritmus dvojného přepočtu.

Algoritmus (Dvojný přepočet s extrapolací)

Vstup: f , a , b , m , ϵ .

Vypočti $h := (b - a)/m$ a $I(h)$.

Opakuj:

polož $h := h/2$,

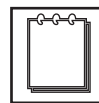
vypočti $I(h)$,

dokud $|E(h)| > \epsilon$.

Výstup: $I_1(h)$.

Poznámka

V algoritmu potřebujeme znát řád p používaného integračního vzorce. Podle věty 6.2.1. je $p = 2$ pro složené obdélníkové a lichoběžníkové pravidlo a $p = 4$ pro složené Simpsonovo pravidlo.



Příklad 6.3.2. Pomocí dvojného přepočtu s extrapolací vypočtete přibližnou hodnotu integrálu

$$I = \int_{-1}^1 e^x dx$$

s přesností $\epsilon = 10^{-4}$. Použijete složené lichoběžníkové pravidlo a začnete pro $m = 4$.

Řešení: Protože $p = 2$, budou mít extrapolační vzorce tvar

$$I_1(h) = I(h) + \frac{I(h) - I(2h)}{3} \quad \text{a} \quad E(h) = \frac{I(h) - I(2h)}{3}.$$

Průběh výpočtu je zaznamenán v tabulce 6.3.2.

Tabulka 6.3.2: Dvojný přepočet s extrapolací, lichoběžníkové pravidlo.

m	h	$I(h)$	$ E(h) $
4	0.5000000	2.3991662	—
8	0.2500000	2.3626313	0.0121783
16	0.1250000	2.3534620	0.0030564
32	0.0625000	2.3511674	0.0007649
64	0.0312500	2.3505936	0.0001913
128	0.0156250	2.3504502	0.0000478

Z posledních dvou hodnot ve sloupci $I(h)$ vypočítáme zpřesněnou aproximaci

$$I_1(h) = 2.3504502 + \frac{2.3504502 - 2.3505936}{3} = 2.3504024.$$

Jako výsledek dostáváme $I = 2.3504024 \pm 0.0001$. Porovnáním s přesnou hodnotou integrálu $I = 2.350402387\dots$ vidíme, že dosažená přesnost je velmi vysoká.

Příklad 6.3.3. V předchozím příkladu použijte složené Simpsonovo pravidlo.

Řešení: Použijeme extrapolační vzorce

$$I_1(h) = I(h) + \frac{I(h) - I(2h)}{15} \quad \text{a} \quad E(h) = \frac{I(h) - I(2h)}{15}.$$

Průběh výpočtu je zaznamenán v tabulce 6.3.3.

Tabulka 6.3.3: Dvojný přepočít s extrapolací, Simpsonovo pravidlo.

m	h	$I(h)$	$ E(h) $
4	0.5000000	2.3511948	—
8	0.2500000	2.3504530	0.000049

Pomocí extrapolace vypočítáme zpřesněnou aproximaci

$$I_1(h) = 2.3504530 + \frac{2.3504530 - 2.3511948}{15} = 2.350403.$$

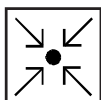
Porovnáním s výpočtem v předchozím příkladu vidíme, že požadované přesnosti jsme dosáhli mnohem dříve.



Kontrolní otázky

Otázka 1. Vysvětlete princip dvojného přepočtu.

Otázka 2. Jak vzniknou extrapolační vzorce a jakou informaci poskytují?



Úlohy k samostatnému řešení

1. Pomocí dvojného přepočtu vypočtete hodnoty integrálů

$$\text{a) } \int_0^1 e^{x^2} dx; \quad \text{b) } \int_0^\pi \cos x^2 dx$$

s přesností $\epsilon = 0.0001$. Použijte složené lichoběžníkové pravidlo.

2. V předchozí úloze použijte složené Simpsonovo pravidlo a extrapolaci.



Výsledky úloh k samostatnému řešení

- Požadovanou přesnost dosáhneme: a) pro $m = 128$, kdy $I = 1.462679 \pm 10^{-4}$;
b) pro $m = 512$, kdy $I = 0.565702 \pm 10^{-4}$.
- Požadovanou přesnost dosáhneme: a) pro $m = 8$, kdy $I = 1.462658 \pm 10^{-4}$;
b) pro $m = 32$, kdy $I = 0.565689 \pm 10^{-4}$.

6.4. Numerické derivování

Cíle

Ukážeme jak přibližně vypočítat hodnoty derivací $f'(x)$ a $f''(x)$ ze známých funkčních hodnot $f(x-h)$, $f(x)$ a $f(x+h)$.



Předpokládané znalosti

Newtonův tvar interpolačního polynomu. Taylorův rozvoj.



Výklad

Budeme postupovat podobně jako při odvození Newton-Cotesových vzorců v odstavci 6.1. K funkci f sestavíme interpolační polynom p_n a ten pak derivujeme místo f . Použijeme přitom Newtonův tvar interpolačního polynomu (5.1.6).



1) Pro uzly $x_0 = x$ a $x_1 = x + h$ sestavíme lineární interpolační polynom a vyjádříme jeho první derivaci:

$$p_1(t) = f(x) + f[x + h, x](t - x) \quad \Rightarrow \quad p'_1(t) = f[x + h, x].$$

Z definice poměrných diferencí dostaneme

$$p'_1(x) = \frac{f(x+h) - f(x)}{h} \approx f'(x). \quad (6.4.1)$$

2) Pro uzly $x_0 = x-h$, $x_1 = x$ a $x_2 = x+h$ sestavíme kvadratický interpolační polynom:

$$p_2(t) = f(x-h) + f[x, x-h](t-x+h) + f[x+h, x, x-h](t-x+h)(t-x).$$

První a druhá derivace mají tvar

$$p'_2(t) = f[x, x-h] + f[x+h, x, x-h](2t-2x+h),$$

$$p''_2(t) = 2f[x+h, x, x-h].$$

Dosazením $t = x$ a úpravou dostáváme

$$\begin{aligned} p'_2(x) &= \frac{f(x) - f(x-h)}{h} + \frac{f(x+h) - 2f(x) + f(x-h)}{2h^2}h \\ &= \frac{f(x+h) - f(x-h)}{2h} \approx f'(x) \end{aligned} \quad (6.4.2)$$

a

$$p''_2(x) = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} \approx f''(x). \quad (6.4.3)$$

Příklad 6.4.1. Vypočtěte přibližné hodnoty první a druhé derivace pro funkci $f(x) = \sin x$ v bodě $x = 1$ s krokem $h = 0.01$ pomocí vzorců (6.4.1), (6.4.2) a (6.4.3). Porovnejte je s přesnými hodnotami.

Řešení: Budeme potřebovat funkční hodnoty $\sin 1 = 0.841471$, $\sin 1.01 = 0.846832$ a $\sin 0.99 = 0.836026$. Podle vzorec (6.4.1), resp. (6.4.2) dostaneme

$$f'(1) \approx p'_1(1) = \frac{0.846832 - 0.841471}{0.01} = 0.536100,$$

resp.

$$f'(1) \approx p'_2(1) = \frac{0.846832 - 0.836026}{2 \cdot 0.01} = 0.540300.$$

Přesná hodnota je $f'(1) = \cos 1 = 0.540302$, takže druhá přibližná hodnota je podstatně přesnější. Podle vzorce (6.4.3) vypočítáme druhou derivaci

$$f''(1) \approx p''_2(1) = \frac{0.846832 - 2 \cdot 0.841471 + 0.836026}{0.01^2} = -0.840000.$$

Nyní je přesná hodnota $f''(1) = -\sin 1 = -0.841471$.

Následující věta ukazuje jaký je řád jednotlivých vzorců.

Věta 6.4.1.

Nechť je funkce f dostatečně hladká (má v okolí bodu x spojitou druhou, třetí, resp. čtvrtou derivaci). Pak platí:

$$\begin{aligned} p_1'(x) - f'(x) &= h \frac{f''(\xi)}{2}, \\ p_2'(x) - f'(x) &= h^2 \frac{f^{(3)}(\xi)}{3}, \\ p_2''(x) - f''(x) &= h^2 \frac{f^{(4)}(\xi)}{12}. \end{aligned} \tag{6.4.4}$$

Ve všech případech je ξ blíže neurčený bod z okolí x .

Důkaz: Nejjednodušší způsob jak získat tato tvrzení je použít Taylorův rozvoj.

Například pro odvození (6.4.4) stačí vyjádřit první derivaci z

$$f(x+h) = f(x) + hf'(x) + h^2 \frac{f''(\xi)}{2}.$$

□

Výpočet přibližných hodnot derivací mohou podstatně ovlivnit zaokrouhlovací chyby. Jmenovatele vzorců totiž obsahují parametr h , který musí být malý, abychom dostali dostatečně přesnou aproximaci derivace. Současně ovšem malá hodnota jmenovatele zlomku zvětšuje zaokrouhlovací chyby v čitateli. Největší dosažitelná přesnost proto musí být jistým kompromisem mezi chybou aproximace a chybou zaokrouhlení. Na příkladu vzorce (6.4.1) si ukážeme analýzu tohoto problému.

Označme $E_{celk}(h)$ horní odhad celkové chyby, který vznikne jako součet odhadu chyby aproximace $e(h)$ a chyby zaokrouhlení $z(h)$, tj.

$$E_{celk}(h) = e(h) + z(h).$$

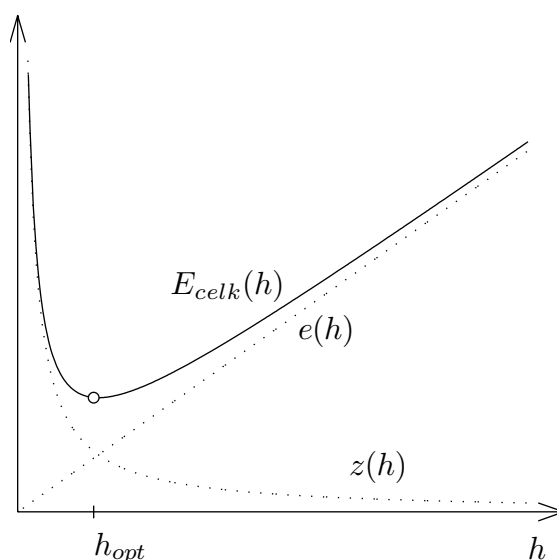
Podle (6.4.4) je $e(h) = Ch$. Dále budeme předpokládat, že při výpočtu $f(x)$ dostaneme vlivem zaokrouhlení porušenou hodnotu $f^*(x)$, přičemž velikost poruchy je nejvýše κ , tj. $|f^*(x) - f(x)| \leq \kappa$. Pro vzorec (6.4.1) dostáváme následující odhad:

$$\left| \frac{f(x+h) - f(x)}{h} - \frac{f^*(x+h) - f^*(x)}{h} \right| \leq \frac{1}{h} (|f^*(x) - f(x)| + |f^*(x+h) - f(x+h)|) \leq \frac{2\kappa}{h} = z(h).$$

Odhad celkové chyby má proto tvar $E_{celk}(h) = Ch + \frac{2\kappa}{h}$, viz obrázek 6.4.1. Funkce $E_{celk}(h)$ má jediné minimum v bodě

$$h_{opt} = \sqrt{\frac{2\kappa}{C}} \quad (6.4.5)$$

a odpovídající hodnota $E_{celk}(h_{opt})$ představuje nejmenší horní odhad celkové chyby. Při výpočtu derivace podle (6.4.1) tedy dojde pro h menší než h_{opt} paradoxně ke ztrátě přesnosti.



Obrázek 6.4.1: Odhad celkové chyby $E_{celk}(h)$.

Příklad 6.4.2. Vypočtěte přibližnou hodnotu první derivace funkce $f(x) = \sin x$ v bodě $x = 1$ s krokem $h = 0.01$ a $h = 0.001$ pomocí vzorce (6.4.1). Zaokrouhľujte přitom na čtyři desetinná místa. Výsledky porovnejte s optimálním krokem h_{opt} .

Řešení: V našem případě je $\kappa = 5 \cdot 10^{-5}$ a

$$\left| \frac{f''(\xi)}{2} \right| = \left| \frac{-\sin \xi}{2} \right| \leq 0.5 = C.$$

Dosazením do vzorce (6.4.5) vypočítáme $h_{opt} = 0.014142$, takže výsledek pro $h = 0.01$ by měl být přesnější. Přibližné derivace mají hodnotu:

$$f'(1) \approx \frac{\sin(1.01) - \sin(1)}{0.01} \doteq \frac{0.8468 - 0.8415}{0.01} = 0.53,$$

$$f'(1) \approx \frac{\sin(1.001) - \sin(1)}{0.001} \doteq \frac{0.8420 - 0.8415}{0.001} = 0.50.$$

Jejich porovnání s hodnotou $f'(1) = \cos 1 = 0.540302$ potvrzuje předchozí závěr.

Kontrolní otázky



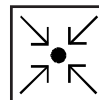
Otázka 1. Jak se odvozují vzorce pro přibližný výpočet derivace?

Otázka 2. Znázorněte graficky smysl vzorců pro výpočet první derivace.

Otázka 3. Jak ovlivňují výpočet derivací zaokrouhlovací chyby?

Otázka 4. Odvoďte podrobně vztah (6.4.5).

Úlohy k samostatnému řešení



1. Vypočítejte přibližně první a druhou derivaci funkce $f(x) = \cos x$ v bodě $x = 1.5$ s krokem $h = 0.01$.

Výsledky úloh k samostatnému řešení



1. $f'(1.5) \approx -0.997832$ podle vzorce (6.4.1); $f'(1.5) \approx -0.997478$ podle vzorce (6.4.2); $f''(1.5) \approx -0.070737$ podle vzorce (6.4.3).

Shrnutí lekce



Ukázali jsme použití základních integračních pravidel a vzorců numerického derivování. Viděli jsme, že numerické výpočty integrálů jsou stabilní vzhledem k zaokrouhlovacím chybám. Zatímco přibližné výpočty derivací jsou na zaokrouhlovací chyby poměrně citlivé.

7. ODR – počáteční úlohy



Průvodce studiem

Jen velmi málo diferenciálních rovnic, které se vyskytují při popisu praktických úloh, se dá řešit exaktně, a i když dokážeme najít vzorce popisující analytické řešení, bývají natolik složité, že je obtížné s nimi dále pracovat. Numerické metody jsou proto často jedinou možností jak danou diferenciální rovnici vyřešit.

V této kapitole představíme základní numerické metody pro řešení počátečních úloh pro obyčejné diferenciální rovnice (ODR). Nejdříve si na příkladu připomeneme některé analytické postupy a nachystáme si modelovou úlohu. Pak stručně odvodíme nejjednodušší Eulerovu metodu, na níž ukážeme podstatné rysy společné všem numerickým metodám. U ostatních metod pak uvedeme pouze vzorce a na příkladech předvedeme jejich použití a vlastnosti.

7.1. Formulace úlohy



Cíle

Zavedeme obecnou formulaci počáteční úlohy pro ODR prvního řádu a zmíníme předpoklady, které zajišťují existenci a jednoznačnost řešení. Na příkladu připomeneme exaktní metody pro výpočet analytického řešení.



Předpokládané znalosti

Exaktní metody řešení počátečních úloh pro ODR.

Výklad

Budeme se zabývat numerickým výpočtem funkce $y = y(x)$, která na intervalu $\langle a, b \rangle$ vyhovuje rovnici

$$y'(x) = f(x, y(x)), \quad (7.1.1)$$

kde $f = f(x, y)$ je daná pravá strana. Rovnice (7.1.1) je *obyčejná diferenciální rovnice prvního řádu* a z teorie je o ní známo, že řešení (pokud existuje) je určeno až na jeden volitelný parametr. Aby bylo řešení určeno jednoznačně, budeme požadovat splnění *počáteční podmínky* ve tvaru

$$y(a) = c. \quad (7.1.2)$$

Úloha (7.1.1), (7.1.2) je *počáteční úloha pro obyčejnou diferenciální rovnici*. Ovšem ani tato úloha nemusí mít jediné řešení. Vše závisí na vlastnostech pravé strany f . O této funkci budeme v celé kapitole předpokládat, že je spojitá v první proměnné a splňuje *Lipschitzovu podmínku* ve druhé proměnné, tj. existuje konstanta L (nezávislá na x a y) taková, že

$$|f(x, y) - f(x, z)| \leq L|y - z| \quad \forall x \in \langle a, b \rangle \quad \forall y, z \in R. \quad (7.1.3)$$

Za těchto předpokladů má počáteční úloha jediné řešení.

Příklad 7.1.1. Počáteční úloha

$$y' = y^2, \quad y(0) = 1$$

nemá jediné řešení na intervalu $\langle 0, 2 \rangle$, protože pravá strana na uvedeném intervalu nesplňuje Lipschitzovu podmínku.

V následujícím příkladu si připomeneme klasické postupy pro řešení počátečních úloh. Bude se jednat o úlohu, kterou v dalším textu použijeme jako modelový příklad při testování numerických metod.

Příklad 7.1.2. Ověřte, že počáteční úloha

$$y' = x^2 - 0.2y, \quad y(-2) = -1 \quad (7.1.4)$$

má jediné řešení. Toto řešení vypočtete exaktně pomocí známých analytických metod.

Řešení: Ověření Lipschitzovy podmínky pro $f(x, y) = x^2 - 0.2y$ provedeme dosazením do

$$\frac{f(x, y) - f(x, z)}{y - z} = \frac{x^2 - 0.2y - x^2 + 0.2z}{y - z} = -0.2,$$

odkud vidíme, že (7.1.3) je splněno například pro konstantu $L = 0.2$. Řešení zadané počáteční úlohy je proto jediné. Při jeho výpočtu nejdříve vyřešíme homogenní diferenciální rovnici $u' = -0.2u$ pomocí separace proměnných. Postup je následující:

$$\begin{aligned} \frac{du}{dx} = -0.2u &\Rightarrow \frac{du}{u} = -0.2 dx \Rightarrow \int \frac{du}{u} = \int -0.2 dx \Rightarrow \\ \ln |u| = -0.2x + c &\Rightarrow |u| = e^{-0.2x+c} \Rightarrow \\ u(x) = Ce^{-0.2x}, \quad C > 0. \end{aligned}$$

Pro výpočet (partikulárního) řešení y_p rovnice (7.1.4) použijeme metodu variace konstant, kdy předem předpokládáme, že řešení bude ve tvaru $y_p(x) = C(x)e^{-0.2x}$. Dosazením do diferenciální rovnice (7.1.4) dostaneme

$$C'(x)e^{-0.2x} + C(x)(-0.2)e^{-0.2x} = x^2 - 0.2C(x)e^{-0.2x} \Rightarrow C'(x) = x^2e^{0.2x}$$

a odtud integrací per partes vypočteme

$$C(x) = \int x^2 e^{0.2x} dx = e^{0.2x}(5x^2 - 50x + 250),$$

takže $y_p(x) = 5x^2 - 50x + 250$. Obecné řešení dané diferenciální rovnice lze zapsat ve tvaru $y(x) = y_p(x) + u(x)$, tj. $y(x) = 5x^2 - 50x + 250 + Ce^{-0.2x}$. Konečně

z počáteční podmínky určíme konstantu C

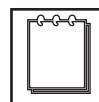
$$-1 = y(-2) = 20 + 100 + 250 + Ce^{0.4} \Rightarrow C = -248.688737$$

a tím i řešení naší výchozí úlohy:

$$y(x) = 5x^2 - 50x + 250 - 248.018417e^{-0.2x}. \quad (7.1.5)$$

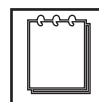
Poznámka

Úlohu (7.1.1), (7.1.2) můžeme chápat vektorově: hledáme vektorovou funkci $y(x) = (y_1(x), \dots, y_n(x))^T$ vyhovující rovnici (7.1.1), kde pravá strana je reprezentována vektorovou funkcí $f(x, y) = (f_1(x, y), \dots, f_n(x, y))^T$, a splňující počáteční podmínku (7.1.2) zadanou vektorem $c = (c_1, \dots, c_n)^T$. Všechny numerické metody, s nimiž se v dalším seznámíme, lze proto použít i pro soustavy obyčejných diferenciálních rovnic prvního řádu.

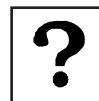


Poznámka

Skutečnost, že se omezujeme pouze na rovnice prvního řádu, není žádné podstatné omezení. Z teorie diferenciálních rovnic je totiž známo, že počáteční úlohu pro rovnici vyššího řádu než 1 lze jednoduchou substitucí převést na soustavu rovnic prvního řádu.



Kontrolní otázky



Otázka 1. Jak vypadá obecná formulace počáteční úlohy pro obyčejnou diferenciální rovnici prvního řádu?

Otázka 2. Zopakujte si exaktní metody pro výpočet přesného (analytického) řešení.

7.2. Eulerova metoda



Cíle

Odvodíme rekurentní vzorce pro Eulerovu metodu a ukážeme jejich použití.



Předpokládané znalosti

Provádění rekurentních výpočtů, analytické řešení počátečních úloh pro ODR.



Výklad

Základem, z něhož vycházejí numerické metody je diskretizace proměnných. Přibližné řešení se nekonstruuje jako spojitá funkce, ale postupně se generuje posloupnost uzlů $x_0 = a, x_1, x_2, \dots$ a pro ně se stanoví čísla $y_0 = c, y_1, y_2, \dots$, která aproximují hodnoty přesného řešení $y(x_0), y(x_1), y(x_2), \dots$. Pro jednoduchost budeme předpokládat, že uzly jsou *ekvidistantní* s krokem h a platí $x_n = b$, tj. $h = (b - a)/n$ a $x_i = a + ih, i = 0, 1, \dots, n$.

Při odvozování vzorců pro přibližné řešení počátečních úloh můžeme postupovat například tak, že v diferenciální rovnici (7.1.1) zapsané v uzlu x_i nahradíme přesnou hodnotu $y(x_i)$ její aproximací y_i a derivaci na levé straně vyjádříme numerickým vzorcem (6.4.1), tj.

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

Odtud jsme schopni při známých hodnotách x_i a y_i vypočítat y_{i+1} . Po přepisu do explicitního tvaru dostáváme rekurentní vzorce *Eulerovy metody*:

$$\left. \begin{aligned} y_0 &= c, \\ y_{i+1} &= y_i + hf(x_i, y_i), \quad i = 0, 1, \dots, n-1. \end{aligned} \right\} \quad (7.2.1)$$

Příklad 7.2.1. Počáteční úlohu (7.1.4) řešte na intervalu $\langle -2, 3 \rangle$ pomocí Eulerovy metody s krokem $h = 1$ a 0.5 . Výsledky porovnejte s přesným řešením.

Řešení: V našem případě je $x_0 = -2$, $y(-2) = y_0 = -1$ a $f(x, y) = x^2 - 0.2y$.

Pro $h = 1$ dostáváme:

$$\begin{aligned}y_1 &= -1 + 1 \cdot ((-2)^2 - 0.2 \cdot (-1)) = 3.2, \\y_2 &= 3.2 + 1 \cdot ((-1)^2 - 0.2 \cdot 3.2) = 3.56, \\y_3 &= 2.848, \quad y_4 = 3.2784, \quad y_5 = 6.6227.\end{aligned}$$

Podobně pro $h = 0.5$ vypočteme:

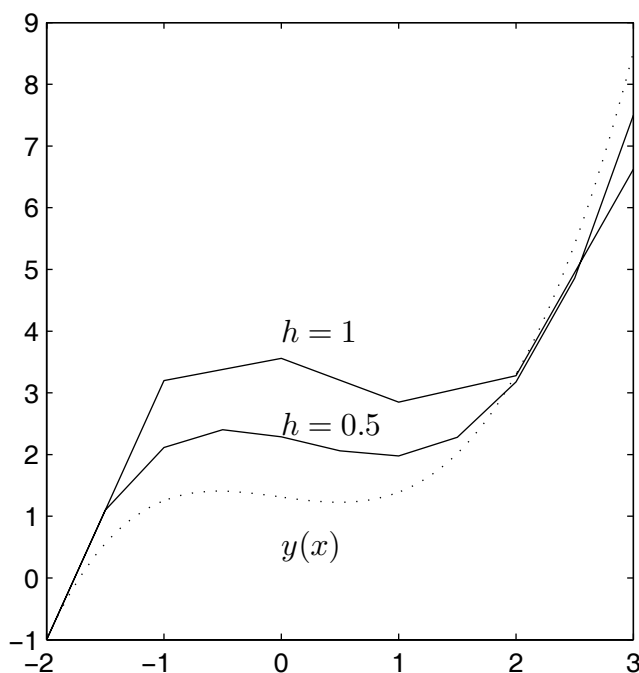
$$\begin{aligned}y_1 &= -1 + 0.5 \cdot ((-2)^2 - 0.2 \cdot (-1)) = 1.1, \\y_2 &= 1.1 + 0.5 \cdot ((-1.5)^2 - 0.2 \cdot 1.1) = 2.115, \\y_3 &= 2.4035, \quad \dots, \quad y_{10} = 7.4988.\end{aligned}$$

Pro lepší názornost zapišme obě přibližná řešení do tabulky 7.2.1, v níž provedeme také porovnání s přesným řešením $y(x)$ určeným vzorcem (7.1.5). Z tabulky je

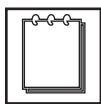
Tabulka 7.2.1: Eulerova metoda.

$h = 1$				$h = 0.5$			
i	x_i	y_i	$ y_i - y(x_i) $	i	x_i	y_i	$ y_i - y(x_i) $
0	-2	-1	0	0	-2	-1	0
1	-1	3.2	1.9491	1	-1.5	1.1	0.5447
2	0	3.56	2.2487	2	-1	2.1150	0.8641
3	1	2.848	1.4571	3	-0.5	2.4035	0.9971
4	2	3.2784	0.0206	4	0	2.2882	0.9769
5	3	6.6227	1.8940	5	0.5	2.0593	0.8322
				6	1	1.9784	0.5875
				7	1.5	2.2806	0.2637
				8	2	3.1775	0.1214
				9	2.5	4.8598	0.5529
				10	3	7.4988	1.0179

vidět, že řešení vypočtené s menším krokem h je přesnější. Tento závěr potvrzuje také obrázek 7.2.1, kde jsou přibližná řešení znázorněna jako lomené čáry.



Obrázek 7.2.1: Přibližná řešení z Eulerovy metody a přesné řešení (tečkovaně).



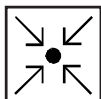
Poznámka

Eulerova metoda je metodou prvního řádu, přestože jsme ji odvodili ze vzorec (6.4.1), který je řádu druhého. Při výpočtu podle Eulerovy metody dochází ke kumulaci (diskretizační) chyby, což má za důsledek snížení řádu.



Kontrolní otázky

Otázka 1. Odvoďte Eulerovu metodu. Jakého je řádu a proč?



Úlohy k samostatnému řešení

1. Eulerovou metodou řešte diferenciální rovnici $y' = 1/(x^2+1) - 0.1y$ s počáteční podmínkou $y(-2) = -1$ na intervalu $\langle -2, 3 \rangle$ s krokem $h = 0.5$.



Výsledky úloh k samostatnému řešení

1. $y_0 = -1$, $y_1 = -0.85$, $y_2 = -0.6537$, $y_3 = -0.3710$, $y_4 = 0.0476$, $y_5 = 0.5452$, $y_6 = 0.9179$, $y_7 = 1.1220$, $y_8 = 1.2198$, $y_9 = 1.2588$, $y_{10} = 1.2648$.

7.3. Jednokrokové metody vyššího řádu

Cíle

Ukážeme použití rekurentních vzorců Heunovy metody a klasické Rungeovy-Kuttovy metody.



Předpokládané znalosti

Eulerova metoda, analytické řešení počátečních úloh pro ODR.



Výklad

Následující vzorce popisují metody vyššího řádu. Při jejich použití dosáhneme srovnatelné přesnosti jako u Eulerovy metody při podstatně větším kroku h , tedy při výrazně menším celkovém objemu výpočtů.

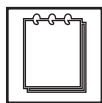


Heunova metoda, která je druhého řádu, je určena rekurentními vzorci:

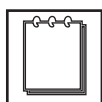
$$\left. \begin{aligned} y_0 &= c, \\ k_1 &= hf(x_i, y_i), \\ k_2 &= hf(x_i + h, y_i + k_1), \\ y_{i+1} &= y_i + \frac{1}{2}(k_1 + k_2), \quad i = 0, \dots, n-1. \end{aligned} \right\} \quad (7.3.1)$$

Rungeova-Kuttova metoda čtvrtého řádu (RK4) je určena rekurentními vzorci:

$$\left. \begin{aligned} y_0 &= c, \\ k_1 &= hf(x_i, y_i), \\ k_2 &= hf(x_i + \frac{1}{2}h, y_i + \frac{1}{2}k_1), \\ k_3 &= hf(x_i + \frac{1}{2}h, y_i + \frac{1}{2}k_2), \\ k_4 &= hf(x_i + h, y_i + k_3), \\ y_{i+1} &= y_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4), \quad i = 0, \dots, n-1. \end{aligned} \right\} \quad (7.3.2)$$

**Poznámka**

Obě uvedené metody jsou jednokrokové, čímž máme namysli, že pro výpočet y_{i+1} stačí hodnota z jediného předchozího kroku y_i .

**Poznámka**

Ze vzorců (7.3.1), (7.3.2) je dále vidět, že zvýšení řádu u jednokrokové metody je spojeno s větším objemem výpočtů, zejména je v každém kroku potřeba počítat hodnoty pravé strany $f(x, y)$ ve více bodech.

Příklad 7.3.1. Počáteční úlohu (7.1.4) řešte na intervalu $\langle -2, 3 \rangle$ pomocí Heunovy metody a metody RK4 s krokem $h = 1$. Výsledky porovnejte s přesným řešením a s řešením získaným Eulerovou metodou.

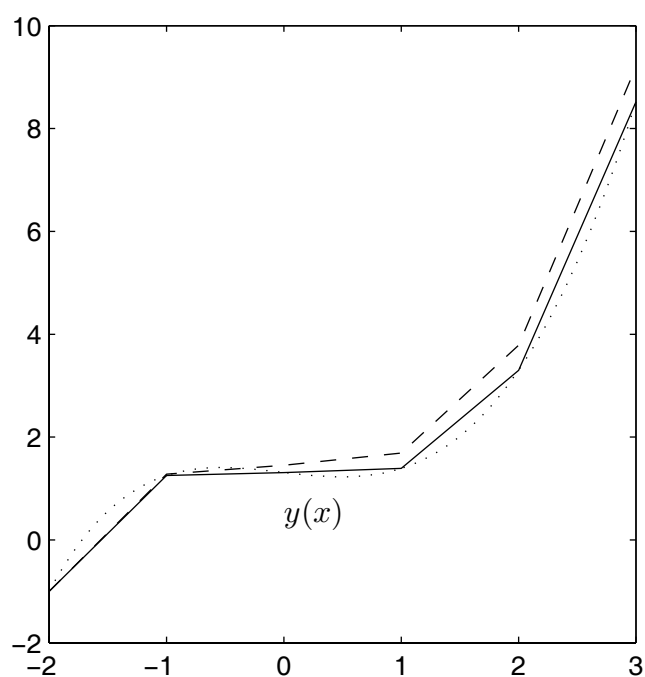
Řešení: Opět vyjdeme z $x_0 = -2$, $y(-2) = y_0 = -1$ a $f(x, y) = x^2 - 0.2y$. Naznačíme první dva kroky Heunovy metody:

$$\begin{aligned} k_1 &= 1 \cdot ((-2)^2 - 0.2 \cdot (-1)) = 4.2, \\ k_2 &= 1 \cdot ((-2 + 1)^2 - 0.2 \cdot (-1 + 4.2)) = 0.36, \\ y_1 &= -1 + \frac{1}{2}(4.2 + 0.36) = 1.28, \\ k_1 &= 1 \cdot ((-1)^2 - 0.2 \cdot 1.28) = 0.744, \\ k_2 &= 1 \cdot ((-1 + 1)^2 - 0.2 \cdot (1.28 + 0.744)) = -0.4048, \\ y_2 &= 1.28 + \frac{1}{2}(0.744 + (-0.4048)) = 1.4496, \\ &\text{atd.} \end{aligned}$$

Obě přibližná řešení zapíšeme do tabulky 7.3.1 a porovnáme je s přesným řešením $y(x)$ určeným vzorcem (7.1.5). Porovnáním se sloupcem $h = 1$ v tabulce 7.2.1 vidíme, že dosažené výsledky jsou přesnější než u Eulerovy metody, přičemž nejpresnější výsledky dává metoda RK4, viz také obrázek 7.3.1.

Tabulka 7.3.1: Metody vyššího řádu.

Heunova metoda				Metoda RK4			
i	x_i	y_i	$ y_i - y(x_i) $	i	x_i	y_i	$ y_i - y(x_i) $
0	-2	-1	0	0	-2	-1	0
1	-1	1.2800	0.0291	1	-1	1.2508	0.0001
2	0	1.4496	0.1383	2	0	1.3112	0.0001
3	1	1.6887	0.2978	3	1	1.3910	0.0001
4	2	3.7847	0.4858	4	2	3.2994	0.0004
5	3	9.2035	0.6867	5	3	8.5175	0.0008

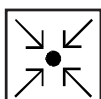


Obrázek 7.3.1: Přibližné řešení vypočítané metodou RK4 (plně), Heunovou metodou (čárkovaně) a přesné řešení (tečkovaně).

**Kontrolní otázky**

Otázka 1. Jaké znáte metody vyššího řádu? Co je jejich výhodou oproti metodám prvního řádu?

Otázka 2. Jaký je vztah mezi řádem metody a početní náročností jednoho kroku?

**Úlohy k samostatnému řešení**

1. Diferenciální rovnici $y' = y(1 + \sin x) - x^3$, $y(1) = 0$ řešte na intervalu $\langle 1, 2 \rangle$ Heunovou metodou a metodou RK4 s krokem $h = 0.2$.

**Výsledky úloh k samostatnému řešení**

1. Heunova metoda: $y_0 = 0$, $y_1 = -0.3114$, $y_2 = -0.9732$, $y_3 = -2.2320$, $y_4 = -4.4495$, $y_5 = -8.1186$; metoda RK4: $y_0 = 0$, $y_1 = -0.3210$, $y_2 = -1.0087$, $y_3 = -2.3257$, $y_4 = -4.6586$, $y_5 = -8.5351$.

7.4. Vícekrokové metody

Cíle



Ukážeme použití vícekrokových metod pro řešení počátečních úloh pro ODR. Pomocí základních vícekrokových vzorců odvodíme algoritmus prediktor-korektor.

Předpokládané znalosti



Jednokrokové a analytické metody řešení počátečních úloh pro ODR.

Výklad



Definice 7.4.1.

Daná metoda se nazývá k -kroková, jestliže při výpočtu y_{i+1} je potřeba dosadit do vzorce metody hodnoty řešení z k předchozích kroků, tj. hodnoty $y_i, y_{i-1}, \dots, y_{i-k+1}$.

Všechny dosud uvedené metody byly jednokrokové. Příkladem dvoukrokové metody je vzorec

$$y_{i+1} = y_{i-1} + 2hf(x_i, y_i), \quad (7.4.1)$$

který lze odvodit podobně jako Eulerovu metodu (7.2.1) pomocí numerického derivování. Vzorec (7.4.1) se říká *metoda skákající žáby*, protože hodnoty přibližného řešení většinou oscilují kolem řešení přesného.

Definice 7.4.2.

Daná metoda se nazývá l -bodová, jestliže při výpočtu y_{i+1} vyžaduje vzorec metody vypočítat hodnoty pravé strany f v l bodech.

Eulerova metoda (7.2.1) je jednobodová, Heunova metoda (7.3.1) je dvoubodová a Runge-Kuttova metoda (7.3.2) je čtyřbodová. Vidíme, že u jednokro-

kových metod je pro dosažení vyššího řádu počítat hodnoty pravé strany ve více bodech, čímž se ovšem zvětšuje pracnost výpočtu. Tuto nepříjemnou vlastnost odstraňují vícekrokové metody. Hodnoty přibližného řešení z předchozích kroků totiž představují v daný okamžik již napočtenou informaci, kterou lze využít pro dosažení vyššího řádu. Všimněme si, že metoda (7.4.1) je jednobodová, dvukroková a druhého řádu (její odvození je založeno na použití vzorce numerického derivování (6.4.2) druhého řádu).

Jistou komplikací u k -krokové metody je skutečnost, že při zahájení výpočtu potřebujeme znát hodnoty y_0, y_1, \dots, y_{k-1} , tzv. *počáteční úsek*. Jeho výpočet se zpravidla provádí vhodnou jednokrokovou metodou (stejného nebo vyššího řádu).

Všechny dosud uvedené metody byly *explicitní*, tzn. že v jejich vzorci se y_{i+1} vyskytovalo pouze na levé straně. U vícekrokových metod se často používají vzorce *implicitní*, u nichž se y_{i+1} vyskytuje i na straně pravé. Použití takového vzorce představuje v každém kroku rovnici (zpravidla nelineární), jejímž řešením je y_{i+1} . Při řešení této rovnice lze použít několik kroků vhodné iterační metody, což je podstata algoritmů typu *prediktor-korektor*.

Adamsovy-Bashforthovy vzorce

Pod názvem *Adamsovy-Bashforthovy vzorce* rozumíme explicitní vícekrokové metody, které lze odvodit integrací interpolačního polynomu sestaveného pro derivaci řešení diferenciální rovnice. Ukažme odvození tříkrokového vzorce.

Nechť $y(x)$ je řešení úlohy (7.1.1) a označme $y'_i = y'(x_i)$. Derivaci $y'(x)$ můžeme vyjádřit ve tvaru

$$y'(x) = p_2(x) + e(x), \quad (7.4.2)$$

kde $p_2(x)$ je Newtonův tvar interpolačního polynomu v uzlech x_i, x_{i-1}, x_{i-2} a $e(x)$ je interpolační chyba:

$$p_2(x) = y'_i + y'[x_{i-1}, x_i](x - x_i) + y'[x_{i-2}, x_{i-1}, x_i](x - x_i)(x - x_{i-1}),$$

$$e(x) = \frac{y^{(4)}(\xi)}{3!}(x - x_i)(x - x_{i-1})(x - x_{i-2}).$$

V (7.4.2) provedeme integraci na intervalu $\langle x_i, x_{i+1} \rangle$:

$$\begin{aligned} \int_{x_i}^{x_{i+1}} y'(x) dx &= y(x_{i+1}) - y(x_i), \\ \int_{x_i}^{x_{i+1}} p_2(x) dx &= y'_i \cdot h + y'[x_{i-1}, x_i] \cdot \frac{h^2}{2} + y'[x_{i-2}, x_{i-1}, x_i] \cdot \frac{5h^3}{6} \\ &= y'_i \cdot h + \frac{y'_i - y'_{i-1}}{h} \cdot \frac{h^2}{2} + \frac{y'_i - 2y'_{i-1} + y'_{i-2}}{2h^2} \cdot \frac{5h^3}{6} \\ &= \frac{h}{12}(23y'_i - 16y'_{i-1} + 5y'_{i-2}), \\ \int_{x_i}^{x_{i+1}} e(x) dx &= \frac{y^{(4)}(\xi)}{3!} \cdot \frac{9}{4} \cdot h^4 = \frac{3}{8} \cdot h^4 \cdot y^{(4)}(\xi), \end{aligned}$$

takže dohromady dostáváme

$$y(x_{i+1}) - y(x_i) = \frac{h}{12}(23y'_i - 16y'_{i-1} + 5y'_{i-2}) + \frac{3}{8}h^4y^{(4)}(\xi).$$

Vynecháním posledního členu vznikne rekurentní vzorec, který lze použít pro přibližný výpočet $y(x_{i+1})$. Stačí si uvědomit, že (přibližné) hodnoty derivací y'_i, y'_{i-1}, y'_{i-2} lze snadno vypočítat dosazením do pravé strany f naší diferenciální rovnice (7.1.1). Označíme-li $y_i \approx y(x_i)$ a $f_i = f(x_i, y_i) \approx y'_i$, obdržíme následující tvar vzorce:

$$y_{i+1} = y_i + \frac{h}{12}(23f_i - 16f_{i-1} + 5f_{i-2}).$$

Jedná se zřejmě o *tříkrokovou* metodu *třetího* řádu (srovnej s poznámkou u Eulerovy metody). Současně je to metoda *jednobodová*! Stačí si totiž uvědomit, že při výpočtu y_{i+1} není potřeba vyčíslit hodnoty f_{i-1} a f_{i-2} , protože k tomu již došlo v předchozích krocích.

Analogickým postupem lze odvodit další Adamsovy-Bashforthovy vzorce lišící se počtem kroků. Obvykle používané vzorce jsou nejvýše čtyřkrokové. Zde je jejich přehled:

$$y_{i+1} = y_i + hf_i, \quad (7.4.3)$$

$$y_{i+1} = y_i + \frac{h}{2}(3f_i - f_{i-1}), \quad (7.4.4)$$

$$y_{i+1} = y_i + \frac{h}{12}(23f_i - 16f_{i-1} + 5f_{i-2}), \quad (7.4.5)$$

$$y_{i+1} = y_i + \frac{h}{24}(55f_i - 59f_{i-1} + 37f_{i-2} - 9f_{i-3}). \quad (7.4.6)$$

Příklad 7.4.1. Počáteční úlohu (7.1.4) řešte na intervalu $\langle -2, 3 \rangle$ pomocí tříkrokové Adamsovy-Bashforthovy metody s krokem $h = 1$. Výsledky porovnejte s přesným řešením.

Řešení: Počáteční úsek vypočítáme metodou RK4. Jsou to první tři hodnoty x_i a y_i z druhé části tabulky 7.3.1, tj. $x_0 = -2$, $x_1 = -1$, $x_2 = 0$ a $y_0 = -1$, $y_1 = 1.2508$, $y_2 = 1.3112$. Protože $f(x, y) = x^2 - 0.2y$, vypočítáme $f_0 = (-2)^2 - 0.2 \cdot (-1) = 4.2000$, $f_1 = (-1)^2 - 0.2 \cdot 1.2508 = 0.7498$ a $f_2 = (0)^2 - 0.2 \cdot 1.3112 = -0.2622$. Pomocí vzorce (7.4.5) pak provedeme všechny další výpočty:

$$y_3 = -1.3112 + \frac{h}{12}(23f_2 - 16f_1 + 5f_0) = 1.5588,$$

$$f_3 = 1^2 - 0.2 \cdot 1.5588 = 0.6882,$$

$$y_4 = 1.5588 + \frac{h}{12}(23f_3 - 16f_2 + 5f_1) = 3.5400,$$

$$f_4 = 2^2 - 0.2 \cdot 3.5400 = 3.2920,$$

$$y_5 = 3.5400 + \frac{h}{12}(23f_4 - 16f_3 + 5f_2) = 8.8227.$$

Výsledky jsou shrnuty v tabulce 7.4.1.

Adamsovy-Moultonovy vzorce

Adamsovy-Moultonovy vzorce jsou implicitní vícekrokové metody, které lze odvodit podobným postupem jako Adamsovy-Bashforthovy vzorce, tj. provede se integrace interpolačního polynomu sestaveného pro derivaci řešení diferenciální rovnice. Nyní se integruje na intervalu $\langle x_{i-1}, x_i \rangle$ a pak se posune indexování. Uvedme pouze přehled vzorců:

$$y_{i+1} = y_i + hf_{i+1}, \quad (7.4.7)$$

$$y_{i+1} = y_i + \frac{h}{2}(f_{i+1} + f_i), \quad (7.4.8)$$

$$y_{i+1} = y_i + \frac{h}{12}(5f_{i+1} + 8f_i - f_{i-1}), \quad (7.4.9)$$

$$y_{i+1} = y_i + \frac{h}{24}(9f_{i+1} + 19f_i - 5f_{i-1} + f_{i-2}). \quad (7.4.10)$$

Například (7.4.9) předstává *dvoukrokovou, jednobodovou* metodu *třetího* řádu. Použití uvedených vzorců vyžaduje v každém kroku řešit rovnici pro neznámou y_{i+1} . Iterační způsob řešení, který jsme zmínili v úvodu, rozebereme až v dalším odstavci. Zde si ukažme jednodušší postup, kdy se do zvoleného vzorce dosadí pravá strana diferenciální rovnice f a pak se vyjádřením y_{i+1} vzorec převede na explicitní tvar. Takto můžeme ovšem postupovat pouze v případech, kdy pravá strana f není „příliš složitá funkce“.

Tabulka 7.4.1: Adamsova-Bashforthova a Adamsova-Moultonova metoda.

Ad.-Bash. 3. řádu				Ad.-Moul. 3. řádu			
i	x_i	y_i	$ y_i - y(x_i) $	i	x_i	y_i	$ y_i - y(x_i) $
0	-2	-1	0	0	-2	-1	0
1	-1	1.2508	0.0001	1	-1	1.2508	0.0001
2	0	1.3112	0.0001	2	0	1.2929	0.0183
3	1	1.5588	0.1679	3	1	1.3613	0.0296
4	2	3.5400	0.2411	4	2	3.2628	0.0362
5	3	8.8227	0.3060	5	3	8.4773	0.0394

Příklad 7.4.2. Počáteční úlohu (7.1.4) řešte na intervalu $\langle -2, 3 \rangle$ pomocí dvoukrokové Adamsova-Moultonovy metody pro $h = 1$. Výsledky porovnejte s přesným řešením.

Řešení: Do pravé strany vzorce (7.4.9) dosadíme $f_{i+1} = f(x_{i+1}, y_{i+1}) = x_{i+1}^2 - 0.2y_{i+1}$:

$$y_{i+1} = y_i + \frac{h}{12}(5(x_{i+1}^2 - 0.2y_{i+1}) + 8f_i - f_{i-1}).$$

Vyjadříme y_{i+1} :

$$y_{i+1} = 12(y_i + \frac{h}{12}(5x_{i+1}^2 + 8f_i - f_{i-1}))/ (12 + h).$$

Tento přepis Adamsova-Moultonova vzorce použijeme ve výpočtech. Jako počáteční úsek vezmeme hodnotu vypočítanou metodou RK4, tj. použijeme x_i a y_i z prvních dvou řádků tabulky 7.3.1. Výsledky jsou shrnuty v tabulce 7.4.1.

Algoritmus prediktor-korektor

Princip odvození algoritmu ukážeme na dvoukrokovém Adamsově-Moultonově vzorci:

$$y_{i+1} = y_i + \frac{h}{12}(5f(x_{i+1}, y_{i+1}) + 8f_i - f_{i-1}).$$

Předpokládejme, že známe hodnoty y_i , y_{i-1} , y_{i-2} a chceme vypočítat y_{i+1} . Na uvedený vzorec můžeme pohlížet jako na rovnici v iteračním tvaru (2.4.1), pro jejíž vyřešení je přirozené použít metodu prostých iterací (2.4.3). Tato úvaha vede na následující výpočetní postup:

Pro $i = 1, \dots, n-1$ vypočti y_{i+1} takto:

- (a) urči počáteční odhad y_{i+1}^0 ,*
- (b) počáteční odhad zpřesňuj pomocí iterací:*

$$y_{i+1}^{k+1} = y_i + \frac{h}{12}(5f(x_{i+1}, y_{i+1}^k) + 8f_i - f_{i-1}), \quad k = 0, 1, 2, \dots$$

Krok (a) se nazývá *prediktor*, krok (b) je *korektor*. Varianta *algoritmu prediktor-korektor* uvedená níže používá explicitní Adamsův-Bashforthův vzorec třetího řádu (7.4.5) jako prediktor. V korektoru se pak vykonává jediné iterační zpřesnění. S ohledem na stručný zápisu algoritmu použijeme symbol přiřazení „:=“ pro změnu hodnoty proměnné, což umožňuje vynechat iterační index k . Dostáváme následující algoritmus:

$$\left. \begin{aligned}
y_0 &= c; \\
y_1, y_2 &\text{ vypočti pomocí jednokrokové metody;} \\
\text{Pro } i = 2, \dots, n-1 &\text{ vypočti } y_{i+1} \text{ takto:} \\
(P) \quad y_{i+1} &:= y_i + \frac{h}{12}(23f_i - 16f_{i-1} + 5f_{i-2}), \\
(E) \quad f_{i+1} &:= f(x_{i+1}, y_{i+1}), \\
(C) \quad y_{i+1} &:= y_i + \frac{h}{12}(5f_{i+1} + 8f_i - f_{i-1}), \\
(E) \quad f_{i+1} &:= f(x_{i+1}, y_{i+1}).
\end{aligned} \right\} \quad (7.4.11)$$

Uvedený algoritmus je metoda *tříkroková*, *dvoubodová* a *třetího* řádu, pro pro niž se používá označení PECE. Podobně se používá označení $\text{PE}(\text{CE})^k$, $\text{P}(\text{EC})^k$ nebo $\text{P}(\text{EC})^k\text{E}$, pro varianty algoritmu prediktor-korektor s k vnitřními kroky a s různou organizací dalších výpočtů.

Příklad 7.4.3. Počáteční úlohu (7.1.4) řešte na intervalu $\langle -2, 3 \rangle$ pomocí PECE metody (7.4.11) s krokem $h = 1$. Výsledky porovnejte s přesným řešením.

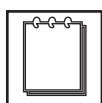
Řešení: Počáteční úsek vypočítáme metodou RK4. Jsou to první tři hodnoty x_i a y_i z druhé části tabulky 7.3.1, tj. $x_0 = -2$, $x_1 = -1$, $x_2 = 0$ a $y_0 = -1$, $y_1 = 1.2508$, $y_2 = 1.3112$. Pomocí těchto hodnot vypočítáme $f_0 = 4.2000$, $f_1 = 0.7498$ a $f_2 = -0.2622$. U dalších výpočtů podle (7.4.11) uvádíme pouze pořadí v jakém vznikají jednotlivé hodnoty:

$$\begin{aligned}
y_3 &:= 1.5588, \quad f_3 := 0.6882, \quad y_3 := 1.3607, \quad f_3 := 0.7279, \\
y_4 &:= 3.4178, \quad f_4 := 3.3164, \quad y_4 := 3.2496, \quad f_4 := 3.3501, \\
y_5 &:= 8.5908, \quad f_5 := 7.2818, \quad y_5 := 8.4564, \quad f_5 := 7.3087.
\end{aligned}$$

Výsledky jsou v tabulce 7.4.2.

Tabulka 7.4.2: Algoritmus prediktor-korektor.

i	x_i	y_i	$ y_i - y(x_i) $
0	-2	-1	0.0000
1	-1	1.2508	0.0001
2	0	1.3112	0.0001
3	1	1.3607	0.0302
4	2	3.2496	0.0493
5	3	8.4564	0.0603

**Poznámka**

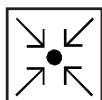
Z porovnáním výsledků v tabulkách 7.4.1 a 7.4.2 je vidět, že implicitní metoda je podstatně přesnější než explicitní metoda stejného řádu. Algoritmus prediktor-korektor představuje „nepřesnou“ implicitní metodu, takže také dosažená přesnost je o něco menší.

**Kontrolní otázky**

Otázka 1. V čem spočívá hlavní přínos vícekových metod oproti metodám jednokrokovým?

Otázka 2. Jak se odvozují Adamsovy-Bashforthovy a Adamsovy-Moultonovy vzorce?

Otázka 3. Na čem je založena konstrukce algoritmu prediktor-korektor?

**Úlohy k samostatnému řešení**

1. Diferenciální rovnici $y' = y(1 + \sin x) - x^3$, $y(1) = 0$ řešte na intervalu $\langle 1, 2 \rangle$ s krokem $h = 0.2$ pomocí Adamsovy-Bashforthovy metody čtvrtého řádu.
2. Vzorec Adamsovy-Moultonovy metody čtvrtého řádu přepište na explicitní tvar pro diferenciální rovnici z předchozí úlohy a proveďte výpočet.
3. Sestavte algoritmy prediktor korektor PECE a PECEC založené na Adamsových-Bashforthových a Adamsových-Moultonových vzorcích čtvrtého řádu a vyřešte diferenciální rovnici z první úlohy.

Výsledky úloh k samostatnému řešení

- Počáteční úsek vypočítáme metodou RK4, pak $y_4 = -4.6497$, $y_5 = -8.5164$.
- Odvodíme $y_{i+1} = 8(y_i + \frac{h}{24}(-9x_{i+1}^3 + 19f_i - 5f_{i-1} + f_{i-2}))/ (8 - 3h(1 + \sin x_{i+1}))$.
Pomocí y_0, y_1, y_2 z RK4 vypočítáme: $y_3 = -2.3270$, $y_4 = -4.6615$, $y_5 = -8.5396$.
- Algoritmus PECEC má následující podobu:

$$\left. \begin{aligned}
 &y_0 = c; \\
 &y_1, y_2, y_3 \text{ vypočti pomocí jednokrokové metody (RK4);} \\
 &\text{Pro } i = 3, \dots, n-1 \text{ vypočti } y_{i+1} \text{ takto:} \\
 &\quad (P) \quad y_{i+1} := y_i + \frac{h}{24}(55f_i - 59f_{i-1} + 37f_{i-2} - 9f_{i-3}), \\
 &\quad (E) \quad f_{i+1} := f(x_{i+1}, y_{i+1}), \\
 &\quad (C) \quad y_{i+1} := y_i + \frac{h}{24}(9f_{i+1} + 19f_i - 5f_{i-1} + f_{i-2}), \\
 &\quad (E) \quad f_{i+1} := f(x_{i+1}, y_{i+1}), \\
 &\quad (C) \quad y_{i+1} := y_i + \frac{h}{24}(9f_{i+1} + 19f_i - 5f_{i-1} + f_{i-2}).
 \end{aligned} \right\}$$

Algoritmus PECE dostaneme vynecháním posledního kroku (C). Počáteční úsek vypočítáme opět metodou RK4. Výsledky podle PECE: $y_4 = -4.6581$, $y_5 = -8.5342$; výsledky podle PECEC: $y_4 = -4.6594$, $y_5 = -8.5360$.

Shrnutí lekce

Ukázali jsme princip numerického řešení počátečních úloh pro obyčejné diferenciální rovnice (prvního řádu). Metody tohoto typu lze nalézt ve standardních knihovnách počítačových programů. Popisovaná problematika je poměrně obsáhlá a zahrnuje například řízení dosažené přesnosti nebo speciální metody pro diferenciální rovnice, jejichž řešení je citlivé na zaokrouhlovací chyby.

Úlohy zapsané pomocí diferenciálních rovnic (různých typů nejen ODR) se vyskytují velice často při modelování fyzikálních nebo technických úloh. Jejich řešení na počítači je atraktivní oblastí výzkumu, který v posledních desetiletích prochází bouřlivým vývojem.

**Literatura**

- [1] Fiedler M. Speciální matice a jejich použití v numerické matematice. SNTL, Praha, 1981.
- [2] Míka S.: Numerické metody algebry. SNTL, Praha, 1985.
- [3] Přikryl P. Numerické metody matematické analýzy. SNTL, Praha, 1988.
- [4] Ralston A. Základy numerické matematiky. Academia, Praha, 1978.
- [5] Sülli E., Mayers D. An introduction to numerical analysis. Cambridge University Press, Cambridge, 2003. (anglicky)
- [6] Van Loan Ch. F. Introduction to scientific computing. Prentice Hall, NJ, 1999. (anglicky)
- [7] Vitásek E. Numerické metody. SNTL, 1987.