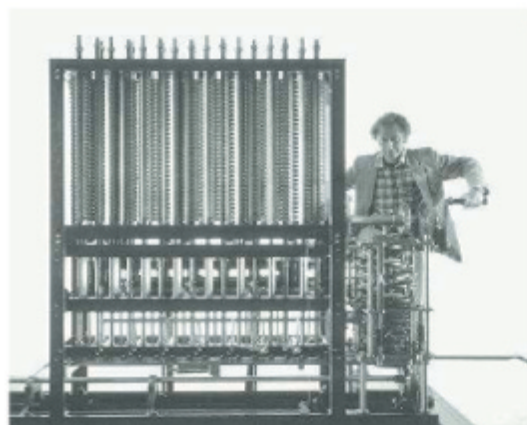


ZÁKLADY NUMERICKEJ MATEMATIKY

Teória a riešené príklady

Jana Gabková, Daniela Richtáriková



ZÁKLADY NUMERICKEJ MATEMATIKY

Teória a riešené príklady

Jana Gabková, Daniela Richtáriková

Všetky práva vyhradené. Nijaká časť textu nesmie byť použitá na ďalšie šírenie akoukoľvek formou bez predchádzajúceho súhlasu autorov alebo nakladateľstva.

© RNDr. Jana Gabková, PhD., RNDr. Daniela Richtáriková, PhD.

Recenzenti: doc. RNDr. Ladislav Halada, CSc.
doc. RNDr. Jana Dobráková, CSc.

Schválila Vedecká rada Strojníckej fakulty STU v Bratislave.

ISBN 978-80-227-3658-9

Obsah

Predhovor	5
1 Numerická matematika	7
1.1 Numerické metódy	8
1.1.1 Priame metódy	9
1.1.2 Iteračné metódy	10
1.2 Chyby	14
1.2.1 Typy chýb	14
1.2.2 Definície chýb	16
1.2.3 Chyby aritmetických operácií	17
1.2.4 Reprezentácia čísel	21
1.2.5 Podmienenosť numerických úloh a numerická stabilita algoritmov	28
1.3 Riešené príklady	29
2 Numerické riešenie nelineárnych rovníc	39
2.1 Základné pojmy	39
2.2 Numerické riešenie rovníc typu $f(x) = 0$	40
2.2.1 Separácia koreňov rovnice	41
2.2.2 Výpočet koreňa rovnice	45
2.3 Riešené príklady	55
3 Vlastné čísla, vlastné vektory matice	59
3.1 Základné pojmy lineárnej algebry	59
3.2 Základné pojmy pre vlastné čísla, vektory matice	66
3.3 Výpočet vlastných čísel matice z charakteristického polynómu	67
3.4 Odhad polohy vlastných čísel matice	69
3.5 Čiastočný problém vlastných čísel	70
3.5.1 Mocninová metóda	71
3.6 Riešené príklady	73
4 Metódy riešenia systémov lineárnych rovníc	79
4.1 Priame metódy riešenia SLR	79
4.1.1 Trojuholníkové matice	81
4.1.2 Riešenie SLR s trojuholníkovými maticami	82
4.1.3 LU rozklad matice systému	83
4.1.4 Riešené príklady	87
4.2 Iteračné metódy riešenia systémov lineárnych rovníc	90
4.2.1 Jacobiho metóda	95
4.2.2 Gauss-Seidlova metóda	97

4.2.3	Riešené príklady	101
5	Aproximácia funkcií	111
5.1	Interpolácia	115
5.1.1	Hľadanie interpolačného polynómu pomocou tabuľky diferencií	117
5.1.2	Lagrangeov polynóm	119
5.1.3	Odhad absolútnej chyby interpolačného polynómu	120
5.1.4	Riešené príklady	122
5.2	Aproximácia v zmysle MNŠ	128
5.2.1	MNŠ – aproximácia polynómom	129
5.2.2	MNŠ – aproximácia exponenciálnou funkciou	132
5.2.3	aproximácia ďalšími funkciami pomocou linearizácie	135
5.2.4	Riešené príklady	140
6	Numerické derivovanie	146
6.1	Riešené príklady	151
7	Numerické integrovanie	155
7.1	Numerické metódy na výpočet určitého integrálu	155
7.2	Newton – Cotesove kvadráturálne vzorce stupňa n	157
7.2.1	Obdĺžniková metóda	157
7.2.2	Lichobežníková metóda	161
7.2.3	Simpsonova metóda	163
7.2.4	Presnosť zložených metód	168
7.3	Riešené príklady	170
8	Numerické riešenie diferenciálnych rovníc	178
8.1	Základné pojmy	178
8.2	Numerické metódy riešenia Cauchyho úlohy	182
8.2.1	Princíp numerických metód riešenia CU	183
8.2.2	Chyby numerických metód a ich rád	185
8.2.3	Rozdelenie numerických metód na riešenie CU	186
8.3	Jednokrokové metódy riešenia Cauchyho úlohy	187
8.3.1	Eulerova metóda	190
8.3.2	1. modifikovaná Eulerova metóda	192
8.3.3	2. modifikovaná Eulerova metóda (Heunova metóda)	193
8.3.4	Metóda Runge – Kutta 4. rádu	195
8.3.5	Metóda Runge – Kutta 3. rádu	197
8.3.6	Presnosť jednokrokových metód	199
8.4	Riešené príklady	201
	Literatúra	212

Pripraviť sa! Pozor! Začíname! Klapka č.1, prosím!

Milé študentky, milí študenti!

Držite v rukách skriptá určené na štúdium predmetu *Numerická matematika* v druhom ročníku bakalárskeho štúdia na SjF STU v Bratislave. Pri ich písaní sme sledovali dva základné ciele. Prvý z nich je zoznámiť Vás so základmi princípov numerickej matematiky a ukázať Vám, ako ich použiť pri riešení technických problémov, pri ktorých bežné analytické metódy matematiky zlyhávajú, alebo sa dajú riešiť len s veľkou námahou. Druhý cieľ, priamo nadväzujúci na prvý, je uľahčiť Vám štúdium tohto predmetu odstránením roky pretrvávajúceho problému nedostatku vhodnej literatúry.

Do skriptu sme zahrnuli výklad tých základných metód Numerickej matematiky, ktoré sú potrebné na štúdium odborných predmetov a zároveň zodpovedajú rozsahu učebných plánov na bakalárskom stupni štúdia. Na zvýšenie úrovne pochopenia a schopnosti riešiť úlohy z oblasti špecifického a nešpecifického transferu sme na koniec každej kapitoly pridali aj riešené príklady. Prečo? Pretože práve riešené príklady sú často najlepší zdroj na pochopenie a osvojenie si preberaného učiva. I keď sú metódy numerickej matematiky predurčené na počítačové spracovanie, snažili sme sa odpútať od nárokov na vstupné vedomosti z programovania a sústredili sme sa viac na samotné princípy. V kľúčových momentoch sme však zaradili aj algoritmy príslušných metód. Ich podrobnejšie rozpracovanie nájdete v učebných materiáloch určených na cvičenia z tohto predmetu. Obrázky a niektoré čiastkové riešenia úloh boli generované v algebraickom systéme *Mathematica*.

Zvládnutie vysvetlenej problematiky a riešenia úloh sú stavebný základ na riešenie matematických modelov technických problémov vo Vašich bakalárskych či diplomových prácach. Skriptá sú vhodné aj pre študentov externého štúdia a aj pre ďalších záujemcov o numerickejšiu matematiku.

Podnetné pripomienky recenzentov doc. RNDr. Ladislava Haladu, PhD. a doc. RNDr. Jany Dobrakovovej, PhD. prispeli k zlepšeniu odbornej a metodologickej úrovne skriptu. Ďakujeme Vám! Zároveň ďakujeme za pomoc aj Mgr. Márii Makovej.

Prajeme Vám úspešné zvládnutie „numeriky“ a celého ďalšieho štúdia!

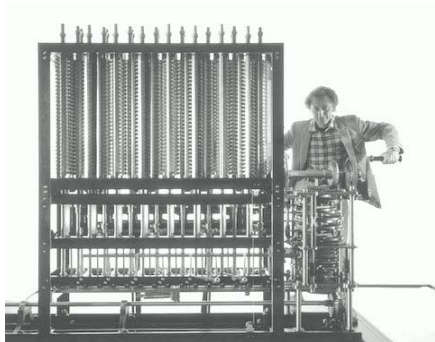
Bratislava 2011

Autorky

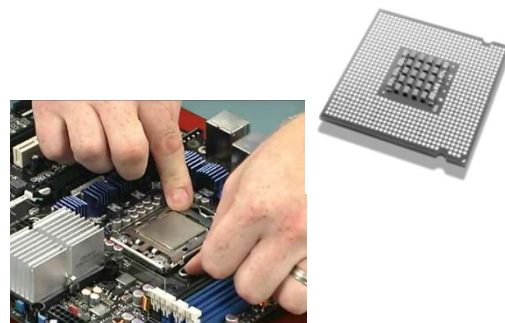
1 NUMERICKÁ MATEMATIKA

Numerická matematika (numerická analýza) patrí do skupiny matematických vied, ktorá na riešenie matematických úloh používa **iba aritmetické operácie** (sčítanie, odčítanie, násobenie, delenie) a **porovnávanie dvoch reálnych čísel**. Na rozdiel od ostatných disciplín, výsledok vo všeobecnosti nie je vyjadrený presne, ale iba **približne**, prostredníctvom konkrétnych čísel s určitou konečnou presnosťou. Dôvodov je niekoľko. V praktickom živote obyčajne potrebujeme hodnoty čísel vyjadriť numericky, t. j. pomocou číslíc. Je známe, že niektoré hodnoty (napríklad iracionálne čísla π , e , $\sqrt{5}$, $\sin 7$, $\log 2$ a pod.) sa nedajú vyjadriť pomocou konečného počtu číslíc. Ďalší dôvod je fakt, že presné riešenia matematických úloh nie vždy sú dostupné, niektoré neexistujú alebo sú také zložité, že sa v podstate nedajú použiť.

História numerickej matematiky siaha do čias Babylonskej ríše, odkiaľ pochádza prvá numerická aproximácia, približné vyjadrenie dĺžky uhlopriečky štvorca so stranou 1 – hodnota čísla $\sqrt{2}$. Vyvíjala sa v súčinnosti s ostatnými matematickými disciplínami, najmä s algebrou a analýzou. Bola a je úzko spätá so súdobou výpočtovou technikou, pracujúcou aj teraz na báze aritmetických a relačných operácií, pre ktorú hľadá, vyvíja a analyzuje metódy umožňujúce ostatným vedným a technickým disciplínam získať relevantné, v praxi použiteľné výsledky. Až do nástupu samostatných počítačích strojov jej práca bola založená na rozsiahlych, ručne počítaných interpolačných tabuľkách a vzorcoch. Dnes sú tabuľky, ako aj algoritmické postupy numerických výpočtov, súčasťou matematických častí softvérov a numerická matematika je samostatný vedný odbor.



Obr. 1.1. *Analytický počítačový stroj
Ch. Babbaga, z roku 1833¹*



Obr. 1.2. *Procesor v súčasnosti*

Numerická matematika je vedná disciplína, ktorá skúma, vyvíja, analyzuje a aplikuje metódy výpočtov s číslami s konečnou presnosťou, pričom sa opiera o modernú výpočtovú techniku. Stručne povedané – numerická matematika:

- a) transformuje matematickú úlohu na úlohu numerickú,
- b) vytvára algoritmy a metódy numerických úloh, pričom na ich riešenie používa výpočtovú techniku.

Riešenie reálnych problémov praxe obyčajne začína formuláciou odborného problému, pokračuje rozdelením na podproblémy, ktoré sa ďalej formulujú ako matematické úlohy. Po ich uspokojivom vyriešení sa výsledky spätne interpretujú v danom odbore a začlenia sa do postupu celkového riešenia. Numerická matematika v prípade potreby vstupuje do procesu vo fáze riešenia matematického problému a jej úlohou je matematickými prostriedkami dosiahnuť výsledky, ktoré sú pre prax použiteľné. V danom procese dochádza v záujme riešiteľnosti a efektivity riešenia k zámerným zjednodušeniam a zanedbaniu rôznych minoritných vplyvov, čo prirodzene zaťažuje výsledok chybou. Aby bol výsledok použiteľný, je nutné vedieť chybu vyčíslieť, t. j. vypočítať alebo kvalifikovane odhadnúť.

1.1 NUMERICKÉ METÓDY

Spôsoby hľadania numerických riešení nazývame **numerické metódy**. Realizujú sa pomocou počítačových programov. Na začiatku, do procesu numerického riešenia vstupujú vstupné údaje, výsledkom sú výstupné údaje a proces sa dá algoritmicke opísať. Napríklad. pri riešení kvadratickej rovnice $ax^2 + bx + c = 0$ vstupnými údajmi sú koeficienty $a, b, c \in R, a \neq 0$ a výstupnými údajmi sú 2 riešenia x_1, x_2 . Algoritmom je presný postup jednotlivých krokov, ako z koeficientov a, b, c vypočítame riešenia rovnice. V numerickej úlohe

- **vstupné údaje** reprezentuje konečná množina čísel, ktorá je potrebná na jednoznačné riešenie danej numerickej úlohy,
- **výstupné údaje** algoritmu reprezentuje konečná množina čísel, ktorá predstavuje riešenie danej úlohy,
- **algoritmus** je konečná postupnosť aritmetických a logických operácií, ktoré vstupným údajom jednoznačne priradia údaje výstupné. Dá sa opísať slovne, graficky (vývojový diagram) alebo priamo programom prostredníctvom niektorého programovacieho jazyka. Zložitejšia úloha sa rieši ako postupnosť jednoduchších úloh.

Mnohé matematické úlohy sa z rôznych dôvodov nedajú presne riešiť aritmetickými prostriedkami. Spojité úlohy, ktoré sa riešia na spojitých množinách, napríklad výpočet derivácií, integrálov, riešenie diferenciálnych rovníc, a pod. nahrádzame úlohami diskrétno charakteru, t. j.

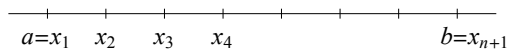
a) spojité intervaly nahrádzame postupnosťou izolovaných bodov:

$$\langle a, b \rangle \sim \{x_1, x_2, x_3, \dots, x_{n+1}\},$$

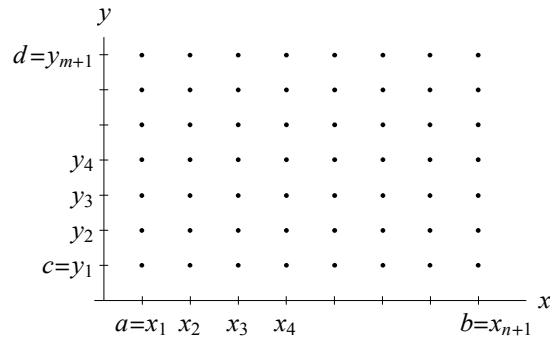
b) spojité dvojrozmerné oblasti typu $\langle a, b \rangle \times \langle c, d \rangle$ nahrádzame sieťou izolovaných bodov $[x_i, y_j]$, kde

$$\langle a, b \rangle \sim \{x_1, x_2, x_3, \dots, x_{n+1}\} \text{ a } \langle c, d \rangle \sim \{y_1, y_2, y_3, \dots, y_{m+1}\}.$$

Riešenie úlohy diskrétno charakteru potom s istou malou chybou aproximuje riešenie pôvodnej úlohy a dá sa vypočítať pomocou aritmetických operácií.



Obr. 1.3. Diskretizácia intervalu $\langle a, b \rangle$



Obr. 1.4. Diskretizácia oblasti $\langle a, b \rangle \times \langle c, d \rangle$

Pri numerických riešeníach používame dva základné druhy výpočtov, podľa ktorých potom metódy delíme na priame a iteračné.

1.1.1 PRIAME METÓDY

Pri priamych metódach dostaneme **výsledok po konečnom počte krokov** a ak v procese nepoužijeme zaokrúhlenie, tak výsledok je **presné číslo**. Takmer všetky výpočty, s ktorými sme sa stretli v základnom kurze matematiky, sú priame. Napríklad riešenie nasledujúcej rovnice pozostáva z krokov:

1. $4x^2 = 7$
2. $x^2 = \frac{7}{4}$
3. $|x| = \frac{\sqrt{7}}{2}$
4. $x_1 = \frac{\sqrt{7}}{2}, x_2 = -\frac{\sqrt{7}}{2}$

Úloha má dve riešenia, ktoré dostaneme vo 4. kroku.

Iným príkladom môže byť riešenie systému lineárnych rovníc dosadzovacou metódou. V prvom kroku z prvej rovnice vypočítame x_1 . V druhom kroku dosadíme x_1 do druhej rovnice a vypočítame x_2 , v treťom kroku dosadíme x_1 a x_2 do tretej rovnice a vypočítame x_3 atď.

$$\begin{aligned}
 1. \quad a_{11}x_1 &= b_1 \quad \Rightarrow \quad x_1 = \frac{1}{a_{11}}b_1 \\
 2. \quad a_{21}x_1 + a_{22}x_2 &= b_2 \quad \Rightarrow \quad x_2 = \frac{1}{a_{22}}(b_2 - a_{21}x_1) \\
 3. \quad a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \quad \Rightarrow \quad x_3 = \frac{1}{a_{33}}(b_3 - a_{31}x_1 - a_{32}x_2) \\
 &\dots\dots\dots \\
 n. \quad a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= b_n \quad \Rightarrow \quad x_n = \frac{1}{a_{nn}}(b_n - a_{n1}x_1 - a_{n2}x_2 - \dots - a_{nn-1}x_{n-1})
 \end{aligned}$$

Na začiatku sú známe hodnoty koeficientov a_{ij} a b_i , $a_{ii} \neq 0$, $i = 1, 2, \dots, n$ a postupne

vypočítame n hodnôt x_i pomocou všeobecného vzorca

$$x_i = \frac{1}{a_{ii}}(b_i - a_{i1}x_1 - a_{i2}x_2 - \dots - a_{ii-1}x_{i-1}) = \frac{1}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j\right), i = 1, 2, \dots, n \quad (1.1)$$

alebo ešte všeobecnejšie:

$$x_i = q_i(x_{i-1}, x_{i-2}, \dots, x_{i-n}), i = 1, 2, \dots, n \quad (1.2)$$

Riešením je n hodnôt x_i , postupnosť $\{x_1, x_2, \dots, x_n\}$. Vzťah (1.1) resp. (1.2) nazývame **rekurentný** vzorec. Všetky vypočítané hodnoty x_i sú súčasťou výsledku. V tomto prípade je to riešenie systému rovníc, súradnice vektora $\bar{x} = (x_1, x_2, \dots, x_n)^T$

1.1.2 ITERAČNÉ METÓDY

Iteračné metódy pracujú na podobnej báze ako rekurentné procesy, avšak s tým rozdielom, že proces výpočtu

$$x_{i+1} = f(x_i) \quad (1.3)$$

kde $i = 0, 1, \dots, k, \dots$ a $f(x)$ je reálna funkcia, vo všeobecnosti **nie je konečný**. Výsledkom nie je postupnosť $\{x_1, x_2, \dots, x_k, \dots\}$ ale iba **jedna hodnota x_k** , ktorá s istou presnosťou **aproximuje** riešenie. Proces sa začína v štartovacom bode x_0 , ktorý musí byť na začiatku procesu známy. Postupne počítame

$$\begin{aligned} x_1 &= f(x_0) \\ x_2 &= f(x_1) \\ x_3 &= f(x_2), \text{ atď.} \end{aligned}$$

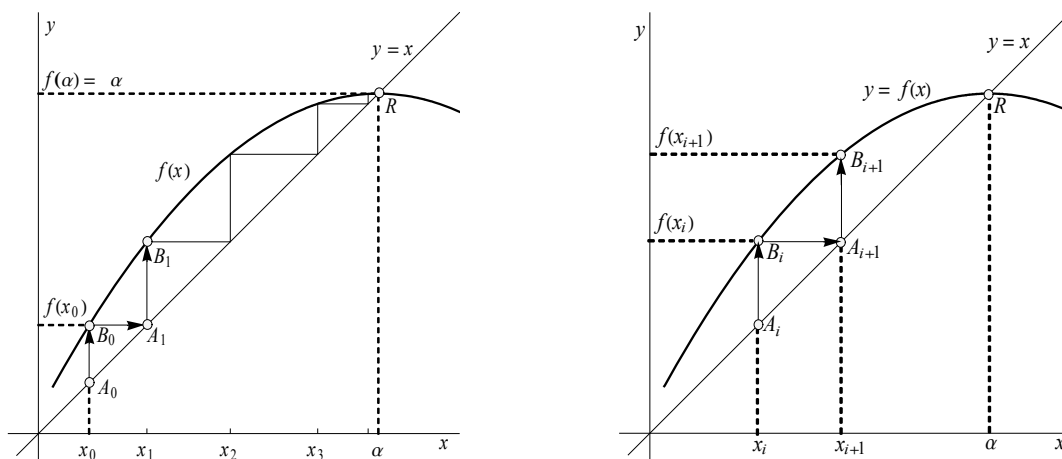
Výpočet jednej hodnoty x_{i+1} nazývame jeden krok iteračného procesu, resp. jedna iterácia. Výsledok, ktorý je výstupom v jednom kroku, v ďalšom kroku vstupuje do iteračného procesu ako vstupná hodnota, premenná x_i . Napríklad hodnotu x_2 v rekurzívnom (spätnom) zápise dostaneme ako

$$x_2 = f(x_1) = f(f(x_0)),$$

všeobecne:

$$x_{i+1} = f(x_i) = f(f(x_{i-1})) = f(f(f(x_{i-2}))) = \dots = f(f(f(\dots f(x_0)\dots))).$$

Graficky môžeme tento proces zaznamenať takto. Zobrazíme graf funkcie $f(x)$ a priamku $y = x$. Budeme kresliť lomenú čiaru (Obr. 1.5). Začneme v bode $[x_0, 0]$ na osi x alebo v bode $A_0[x_0, x_0]$ na priamke $y = x$ a kreslíme priamku rovnobežnú s osou y , nájdeme priesečník s grafom $f(x)$, bod $B_0[x_0, f(x_0)]$. Bodom B_0 vedieme priamku rovnobežnú s osou x a nájdeme priesečník s priamkou $y = x$, bod $A_1[x_1, x_1]$. Bodom A_1 vedieme priamku rovnobežnú s osou y a nájdeme priesečník s grafom $f(x)$, bod $B_1[x_1, f(x_1)]$ atď.



Obr. 1.5. Konvergentný iteračný proces

Ak je proces konvergentný, tak postupnosť $\{x_0, x_1, x_2, \dots, x_k, \dots\}$ konverguje k číslu α , t. j.

$$\lim_{i \rightarrow \infty} x_i = \alpha$$

a postupnosť bodov grafu $\{A_0, \dots, A_k, \dots\}$ ako aj $\{B_0, \dots, B_k, \dots\}$ (Obr. 1.5) konverguje k bodu $R[\alpha, \alpha]$ – priesečníku grafu funkcie $f(x)$ a priamky $y = x$:

$$\lim_{i \rightarrow \infty} A_i = \lim_{i \rightarrow \infty} B_i = R.$$

Číslo α je presným riešením $x = f(x)$, t. j. $\alpha = f(\alpha)$. Čísla x_i , $i = 0, 1, \dots, k, \dots$ nazývame postupné aproximácie (približenia) alebo postupné iterácie čísla α (nultá, prvá, druhá, i -ta aproximácia).

Poznámka. Termín aproximácia, resp. iterácia používame tak pre jeden krok iteračného procesu ako aj pre jeho výsledok, hodnotu x_i .

Je prirodzené, že na počítačoch zariadeniach vykonávame len konečný počet iterácií. Cyklus výpočtov končí buď vykonaním pevného, dopredu stanoveného počtu iterácií alebo ukončovacou (výstupnou) podmienkou. Ostatnú vypočítanú iteráciu x_k tak obvykle považujeme za numerické riešenie $x = f(x)$, t. j. $\alpha \doteq x_k$.

ÚLOHA

Vypočítajte hodnotu $\sqrt{5}$ pomocou iteračného predpisu $x_{i+1} = \frac{1}{2} \left(x_i + \frac{a}{x_i} \right)$, kde a je základ druhej odmocniny a platí $\sqrt{a} = \lim_{i \rightarrow \infty} x_i$ pre ľubovoľnú štartovaciu hodnotu $x_0 > 0$. Počítajte

hej odmocniny a platí $\sqrt{a} = \lim_{i \rightarrow \infty} x_i$ pre ľubovoľnú štartovaciu hodnotu $x_0 > 0$. Počítajte

a) 4 iterácie,

b) s toleranciou $|x_{i+1} - x_i| \leq \varepsilon$, $\varepsilon = 10^{-3}$.

RIEŠENIE

a) Výpočet 4 iterácií

Zvolíme štartovací bod x_0 , $x_0 > 0$ a vypočítame postupnosť hodnôt x_1, x_2, x_3 a x_4 .

$$x_0 = 1$$

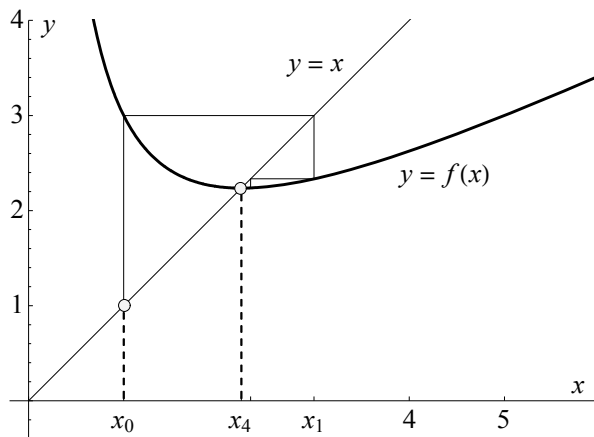
$$x_1 = \frac{1}{2} \left(x_0 + \frac{5}{x_0} \right) = 3$$

$$x_2 = \frac{1}{2} \left(x_1 + \frac{5}{x_1} \right) = 2,333333333$$

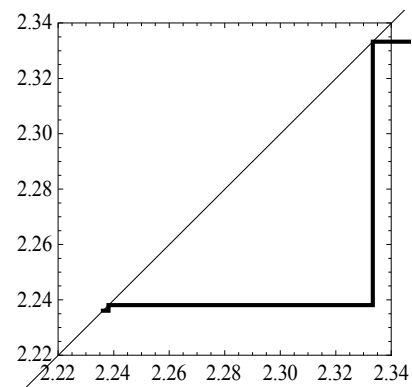
$$x_3 = \frac{1}{2} \left(x_2 + \frac{5}{x_2} \right) = 2,238095238$$

$$x_4 = \frac{1}{2} \left(x_3 + \frac{5}{x_3} \right) = 2,236068896$$

Výpočet končí štvrtou iteráciou. Hodnota čísla $\sqrt{5} \doteq x_4 = 2,236068896$.



a)



b) detail

Obr. 1.6. Grafické zobrazenie iteračného procesu

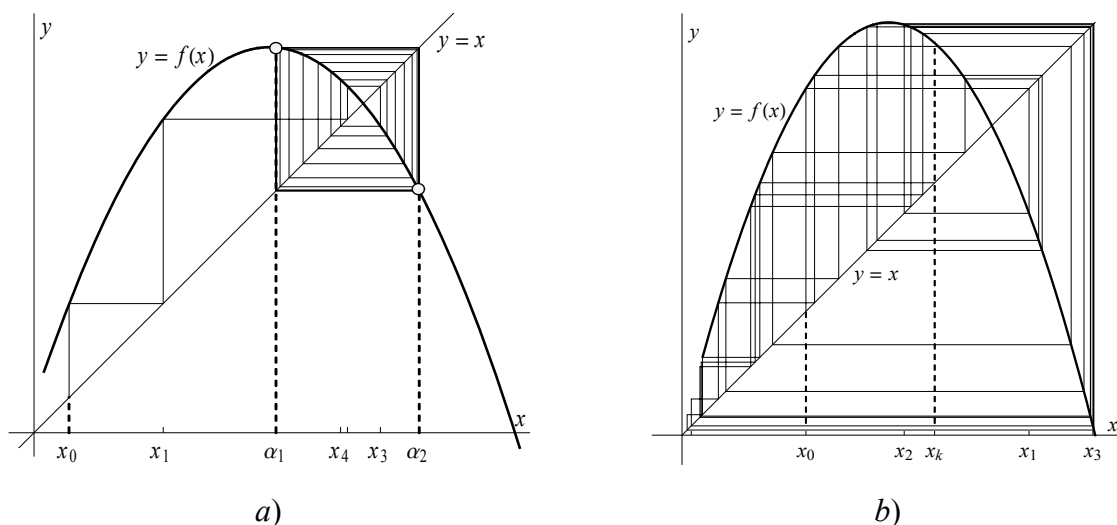
b) Výpočet s toleranciou $|x_{i+1} - x_i| \leq \varepsilon$, $\varepsilon = 10^{-3}$.

Počítame podobne ako v a) len s tým rozdielom, že výpočet ukončíme, keď ostatné dve iterácie budú dostatočne blízko, t. j. budú sa líšiť o menej ako predpísané ε .

$x_0 = 1$	
$x_1 = 3$	$\Delta x = x_1 - x_0 = 2,000000000 \dots > \varepsilon$
$x_2 = 2,333333333$	$\Delta x = x_2 - x_1 = 0,666666667 \dots > \varepsilon$
$x_3 = 2,238095238$	$\Delta x = x_3 - x_2 = 0,095238095 \dots > \varepsilon$
$x_4 = 2,236068896$	$\Delta x = x_4 - x_3 = 0,002026342 \dots > \varepsilon$
$x_5 = 2,236067977$	$\Delta x = x_5 - x_4 = 9,181433855 \times 10^{-7} < \varepsilon$

Pre piatu iteráciu je rozdiel $|x_5 - x_4| = 9,181433855 \times 10^{-7} < 10^{-3}$, podmienka je splnená, a tým sa iteračný proces končí. $\sqrt{5} \doteq x_5 = 2,236067977$

Nie všetky iteračné procesy konvergujú. Ukážky dvoch divergentných procesov vidíme na Obr. 1.7 a) a b). Na Obr. 1.7 a) čísla x_i nadobúdajú od istého i počnúc striedavo dve hodnoty α_1 a α_2 . Na Obr. 1.7 b) čísla x_i nadobúdajú hodnoty neusporiadane, chaoticky.



Obr. 1.7. Divergentný iteračný proces

Pred použitím iteračnej metódy budeme vždy zisťovať jej konvergenciu. Základnou vetou, o ktorú sa opierame pri zisťovaní konvergencie iteračného procesu (1.3) je Banachova veta o pevnom bode.

Veta 1.1 (Banachova veta o pevnom bode)

Nech funkcia $f(x)$ je

- reálna funkcia zobrazujúca množinu M do seba, t. j. funkcia, ktorá každému vzoru $x \in M \subseteq \mathbb{R}$ priradí práve jedno $y \in M$,
- **kontraktívna** funkcia s koeficientom s , t. j. existuje také reálne číslo $s \in (0, 1)$, že pre ľubovoľnú dvojicu čísel x_i, x_j z M platí

$$|f(x_i) - f(x_j)| \leq s |x_i - x_j|$$

(vzdialenosť obrazov je menšia nanajvýš rovná s násobku vzdialenosti vzorov, t. j. vzdialenosti bodov sa po zobrazení funkciou f zmenšia (Obr. 1.8)).

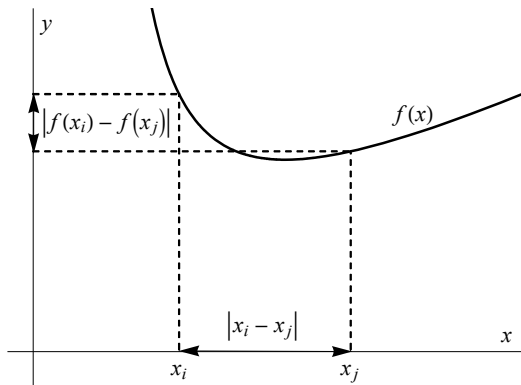
Potom existuje práve jeden **pevný bod** $\alpha \in M$ funkcie $f(x)$, t. j. také číslo α , pre ktoré

$$f(\alpha) = \alpha.$$

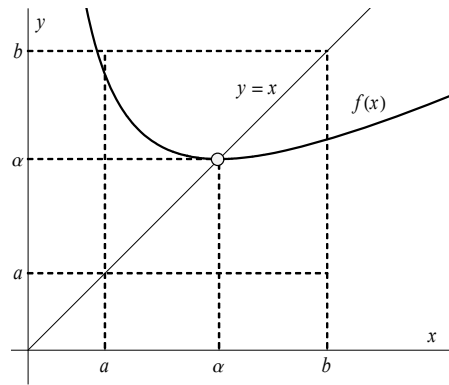
Zároveň platí: ak $x_{i+1} = f(x_i)$, potom $\lim_{i \rightarrow \infty} x_i = \alpha$ pre všetky $x_i \in M$.

Poznámka. Bod α nazývame tiež **atraktorom** funkcie $f(x)$.

Veta 1.2 Nech funkcia $f(x)$ je diferencovateľná (existuje derivácia) na množine M a nech existuje $s < 1$ také, že pre všetky $x \in M$ platí $|f'(x_i)| \leq s$. Potom funkcia $f(x)$ je na množine M **kontraktívna** a koeficient kontrakcie je s .



Obr. 1.8. Kontraktívna funkcia $f(x)$.
 $|f(x_i) - f(x_j)| < |x_i - x_j|$



Obr. 1.9. Pevný bod α funkcie $f(x)$ na intervale $\langle a, b \rangle$

1.2 CHYBY

1.2.1 TYPY CHÝB

Podľa toho, kde a ako chyby vznikajú, môžeme ich rozdeliť do skupín.

1. Chyby matematického modelu

Pri vytváraní matematického modelu, napr. fyzikálnych, chemických či mechanických procesov často zanedbávame vplyvy, ktoré majú nepodstatný význam alebo zjednodušujú formuláciu úloh. Vznikne matematický model, ktorý idealizuje reálny problém. Chyby, ktoré pramenia z rozdielov medzi idealizujúcim modelom a reálnym problémom, sú chybami matematického modelu. Napríklad:

- ideálny plyn je matematickým modelom reálneho plynu (v ideálnom plyne predpokladáme, že molekuly nezaberajú žiadny objem, nepôsobia na seba silami, a pod.),
- hmotný bod nahrádza teleso. Zanedbávame rozmery, tvar, objem telesa atď. Hmotnosť telesa je sústredená do jediného bodu.
- pri výpočte veľkosti povrchu Zeme uvažujeme povrch gule ($S=4\pi \cdot r^2$), i keď vieme, že Zem má komplikovanejší tvar aj povrch,
- pri opise pohybu fyzikálneho kyvadla neuvažujeme odpor prostredia atď.

2. Chyby vstupných údajov

Väčšinou sú spôsobené nepresnosťou pri meraní fyzikálnych veličín (chyba meracieho prístroja, chyby merania, náhodné chyby, ...). Meraním získavame len približnú hodnotu meranej veličiny. Novým meraním vo všeobecnosti dostávame nový výsledok.

3. Chyby numerickej metódy

Pri riešení konkrétnych matematických úloh je nutné robiť určité zjednodušenia a isté aproximácie. Napríklad

- derivácie a integrály funkcie nahrádzame výrazom obsahujúcim súčet konečného počtu funkčných hodnôt,
- v okolí skúmaných bodov elementárne funkcie (goniometrické, logaritmické, ...) nahrádzame konečným počtom členov Taylorovho radu,
- limitu postupnosti nahrádzame niektorým (dostatočne veľkým) členom postupnosti:

$$\lim_{i \rightarrow \infty} a_i \doteq a_k .$$

Chyby metódy vznikajú pri riešení matematickej úlohy, keď zvolená metóda neposkytuje presné teoretické riešenie, presný výsledok.

4. Chyby zaokrúhľovania

Pri výpočtoch reálne čísla nahrádzame ich aproximáciami, t. j. číslami s konečným počtom platných číslic. Súvisí to s reprezentáciou čísla v počítačoch zariadeniach. Tieto chyby sa pri výpočtoch môžu kumulovať ale aj navzájom eliminovať. Ich veľkosť závisí od počtu číslic použitých na ich vyjadrenie, napr. $\sqrt{655} = 25,59296\dots \doteq 25,6$, $\pi = 3,14159\dots \doteq 3,14$.

5. Šírenie chýb

Pod šírením chýb rozumieme chybu, ktorá vzniká v priebehu výpočtov v dôsledku predchádzajúcich chýb. Je to chyba, ktorá aditívnym spôsobom zväčšuje lokálnu chybu, t. j. k ostatným chybám sa pripočítava. Nebezpečné sú situácie, keď chyba výpočtami nekontrolovane rastie a jej veľkosť úplne zatieni samotné riešenie a kompletne znehodnotí výsledok.

6. Hrubé chyby a ľudské chyby

Počítacie stroje robia chyby veľmi zriedkavo. Hrubé chyby ako aj ľudské chyby vznikajú pôsobením človeka. Pramenia z omylu, nepozornosti, neprecíznosti, a pod. Odhaľujú sa testovaním programov a rôznymi kontrolnými mechanizmami. Kontrola vkladaných údajov sa často robí ešte aj dnes pomocou tzv. kontrolných súčtov (nájdeme ich v tabuľkách pre štatistické, daňové a iné úrady).

Každá z uvedených chýb sa môže pri riešení reálneho problému vyskytovať samostatne bez vplyvu ostatných chýb, no zvyčajne sa vyskytujú všetky chyby súčasne.

1.2.2 DEFINÍCIE CHÝB

Označme x – presnú hodnotu čísla a x^* – jej približnú hodnotu (aproximáciu). Na vyjadrenie ich odlišnosti zavádzame tieto pojmy:

1. Absolútna chyba

Absolútnou chybou aproximácie čísla x^* nazývame rozdiel

$$\Delta x = |x - x^*| \quad (1.4)$$

Pre presnú hodnotu x potom platí $x = x^* + \Delta x$ pre $x > x^*$ alebo $x = x^* - \Delta x$ pre $x < x^*$.

Absolútna chyba aproximácie dáva informáciu o absolútnej veľkosti rozdielu presnej a približnej hodnoty skúmaného čísla. Ak napríklad meriame fyzikálnu veličinu v nejakých jednotkách, potom aj absolútna chyba sa bude udávať v týchto jednotkách.

2. Odhad absolútnej chyby

Odhadom absolútnej chyby nazývame ľubovoľné nezáporné číslo, pre ktoré platí:

$$\Delta x \leq \varepsilon(x), \quad \varepsilon(x) \geq 0. \quad (1.5)$$

Odhad absolútnej chyby zavádzame, pretože zvyčajne presnú hodnotu čísla x nepoznáme a teda nevieme vypočítať ani presnú hodnotu absolútnej chyby Δx . Číslo $\varepsilon(x)$ nemôže byť menšie ako skutočná chyba, ale nemalo by byť ani oveľa väčšie. Z $|x - x^*| \leq \varepsilon(x)$ dostávame $x^* - \varepsilon(x) \leq x \leq x^* + \varepsilon(x)$, v skrátenej zápise $x = x^* \pm \varepsilon(x)$. Odhad chyby $\varepsilon(x)$ určuje dĺžku intervalu, v ktorom sa nachádza presná hodnota x : $x \in \langle x^* - \varepsilon(x), x^* + \varepsilon(x) \rangle$. Čím je odhad chyby menší, tým je interval v ktorom garantujeme presnú hodnotu x menší, a teda tým je aj odhad lepší. Očakáva sa, že odhadom chyby vyjadríme najmenší možný odhad.

Nech je chyba vážiaceho zariadenia 2 kg. Znamená to, že presná hodnota m (hmotnosť) bude z intervalu $m \in \langle m^* - 2\text{kg}, m^* + 2\text{kg} \rangle$. Je to veľa alebo málo? Z hodnoty absolútnej chyby a ani z jej odhadu nevieme usúdiť, či je chyba malá alebo veľká. Ak pôjde o vážiace zariadenie nákladných áut, potom je toto vážiace zariadenie veľmi presné. Ak však budeme hovoriť o osobnej váhe, potom táto váha je značne nepresná.

3. Relatívna chyba:

$$\frac{|x - x^*|}{|x|} = \frac{\Delta x}{|x|} \quad (1.6)$$

Relatívna chyba je pomerové číslo, vyjadruje pomer absolútnej chyby vo vzťahu ku skutočnej hodnote daného čísla a teda udáva koľkokrát je absolútna chyba aproximácie Δx väčšia ako skutočná hodnota x . Na jej základe už vieme posúdiť, či ide o chybu veľkú alebo malú, zanedbateľnú alebo nie.

Keďže je to pomerové číslo, často sa vyjadruje v percentách:

$$\frac{\Delta x}{|x|} \cdot 100\%$$

Vzhľadom na to, že presnú hodnotu čísla x zvyčajne nepoznáme, nahrádzame vo vzorci absolútnu chybu jej odhadom a presnú hodnotu x jej aproximáciou x^* , t. j.

$$\frac{\Delta x}{|x|} \approx \frac{\Delta x}{|x^*|} \leq \frac{\varepsilon(x)}{|x^*|}$$

4. Odhad relatívnej chyby

Odhadom relatívnej chyby nazývame ľubovoľné nezáporné číslo $\delta(x)$, pre ktoré platí:

$$\frac{\Delta x}{|x|} \leq \delta(x), \quad \delta(x) \geq 0, \quad x \neq 0 \quad (1.7)$$

Analogicky ako pri odhade absolútnej chyby, $\delta(x)$ je číslo, ktoré nemôže byť menšie ako

skutočná relatívna chyba, ale nemalo by byť ani o veľa väčšie. Z $\frac{|x-x^*|}{|x^*|} \leq \delta(x)$ dostávame

$$x^* - |x^*| \cdot \delta(x) \leq x \leq x^* + |x^*| \cdot \delta(x), \text{ v skrátrenom zápise } x = x^* (1 \pm \delta(x)).$$

Z dvoch aproximácií x_1^* a x_2^* považujeme za presnejšiu tú, ktorá má menšiu relatívnu chybu.

typ	vzorec	poznámka
absolútna chyba	$\Delta x = x - x^* $	
odhad absolútnej chyby	$ x - x^* \leq \varepsilon(x)$	$\varepsilon(x) \geq 0$, veľmi malé
relatívna chyba	$\frac{ x - x^* }{ x } = \frac{\Delta x}{ x }$	$\frac{ x - x^* }{ x } = \frac{\Delta x}{ x } \leq \frac{\varepsilon(x)}{ x }$,
relatívna chyba v %	$\frac{\Delta x}{ x } \cdot 100\%$	$x \neq 0$
odhad relatívnej chyby	$\delta(x) \geq \frac{\Delta x}{ x }$	$\delta(x) \geq 0$, veľmi malé, $x \neq 0$

Tabuľka 1.1. *Prehľad typov chýb, ich odhadov a vzorcov*

1.2.3 CHYBY ARITMETICKÝCH OPERÁCIÍ

Na aritmetické operácie budeme nahliadať ako na funkcie viac premenných, napr. pre súčet n sčítancov budeme uvažovať funkciu: $f(x_1, x_2, \dots, x_n) = x_1 + x_2 + \dots + x_n$.

1.2.3.1 Odhad chyby funkcie viac premenných

Nech x_1, x_2, \dots, x_n sú presné čísla a $x_1^*, x_2^*, \dots, x_n^*$ sú ich aproximácie. Nech na oblasti $G \subseteq R_n$ má funkcia $f(\bar{x}) = f(x_1, x_2, \dots, x_n)$ deriváciu, pričom $\bar{x} = (x_1, x_2, \dots, x_n)$ a $\bar{x}^* = (x_1^*, x_2^*, \dots, x_n^*)$ ležia v oblasti G . V prípade, že sa funkcia $f(\bar{x})$ v okolí aproximácie \bar{x}^* málo mení a absolútne chyby súradníc Δx_i , resp. ich odhady $\varepsilon(x_i)$ sú dostatočne malé, potom absolútnu chybu môžeme približne vyjadriť pomocou diferenciálu funkcie $f(\bar{x})$ a platí:

$$\Delta f(\bar{x}) = |f(\bar{x}^*) - f(\bar{x})| \leq \sum_{i=1}^n \left| \frac{\partial f(\bar{x})}{\partial x_i} \right| \cdot \Delta x_i \quad (1.8)$$

t. j.

$$\Delta f(\bar{x}) = |f(\bar{x}^*) - f(\bar{x})| \leq \left| \frac{\partial f(\bar{x})}{\partial x_1} \right| \cdot \Delta x_1 + \left| \frac{\partial f(\bar{x})}{\partial x_2} \right| \cdot \Delta x_2 + \dots + \left| \frac{\partial f(\bar{x})}{\partial x_n} \right| \cdot \Delta x_n$$

Pri praktických výpočtoch, keď presnú hodnotu \bar{x} nepoznáme, odhad absolútnej chyby budeme počítať pomocou diferenciálu funkcie $df(\bar{x}, \bar{x}^*)$ a odhadu absolútnych chýb premenných $\varepsilon(x_i^*)$. Parciálne derivácie ako aj odhady chýb jednotlivých premenných počítame v približnej hodnote \bar{x}^* :

$$\varepsilon(\bar{x}^*) \approx \sum_{i=1}^n \left| \frac{\partial f(\bar{x}^*)}{\partial x_i} \right| \cdot \varepsilon(x_i^*) \quad (1.9)$$

t. j.

$$\varepsilon(\bar{x}^*) \approx \left| \frac{\partial f(\bar{x}^*)}{\partial x_1} \right| \cdot \varepsilon(x_1^*) + \left| \frac{\partial f(\bar{x}^*)}{\partial x_2} \right| \cdot \varepsilon(x_2^*) + \dots + \left| \frac{\partial f(\bar{x}^*)}{\partial x_n} \right| \cdot \varepsilon(x_n^*)$$

V praxi sa často vyskytuje úloha – s akými absolútnymi chybami treba vziať približné hodnoty premenných funkcie, aby absolútna chyba približnej hodnoty funkcie so rovnala vopred predpísanej hodnote. Túto úlohu nazývame „**obrátená úloha**“ teórie chýb. Pri riešení obrátenej úlohy teórie chýb sa používa „princíp rovnakého vplyvu“. Podľa tohto princípu sa predpokladá, že všetky členy $\left| \frac{\partial f(\bar{x})}{\partial x_i} \right| \cdot \Delta x_i$, $i = 1, 2, \dots, n$ majú rovnaký vplyv na $\Delta f(\bar{x}^*)$.

1. Súčet

Pre súčet dvoch čísel má funkcia tvar $f(x_1, x_2) = x_1 + x_2$. Parciálne derivácie sú $\frac{\partial f(\bar{x})}{\partial x_1} = 1$, $\frac{\partial f(\bar{x})}{\partial x_2} = 1$. Odhad absolútnej chyby vyjadríme podľa (1.8) a (1.9) ako

$$\varepsilon(x_1 + x_2) = \left| \frac{\partial f(\bar{x})}{\partial x_1} \right| \cdot \Delta x_1 + \left| \frac{\partial f(\bar{x})}{\partial x_2} \right| \cdot \Delta x_2 = \Delta x_1 + \Delta x_2 \approx \varepsilon(x_1) + \varepsilon(x_2). \quad (1.10)$$

Odhad relatívnej chyby bude

$$\delta(x_1 + x_2) = \frac{\varepsilon(x_1 + x_2)}{|x_1 + x_2|} = \frac{\Delta x_1 + \Delta x_2}{|x_1 + x_2|} \approx \frac{\varepsilon(x_1) + \varepsilon(x_2)}{|x_1 + x_2|}. \quad (1.11)$$

Keď vo vzťahu (1.11) chceme použiť relatívne chyby sčítancov, potom $\Delta x = \frac{\Delta x}{|x|} \cdot |x|$ a

$$\delta(x_1 + x_2) = \frac{\Delta x_1 + \Delta x_2}{|x_1 + x_2|} = \left| \frac{x_1}{x_1 + x_2} \right| \cdot \frac{\Delta x_1}{|x_1|} + \left| \frac{x_2}{x_1 + x_2} \right| \cdot \frac{\Delta x_2}{|x_2|} \approx \frac{|x_1| \delta(x_1) + |x_2| \delta(x_2)}{|x_1 + x_2|}$$

2. Násobok

Nech m je prirodzené číslo. Potom

$$\varepsilon(m x_1) = \varepsilon(x_1) \quad (1.12)$$

$$\delta(m x_1) = \delta(x_1). \quad (1.13)$$

3. Rozdiel

Rozdiel $x_1 - x_2$ môžeme považovať za súčet $x_1 + (-x_2)$. Potom odhad absolútnej chyby vypočítame ako

$$\varepsilon(x_1 - x_2) = \Delta x_1 + \Delta x_2 \approx \varepsilon(x_1) + \varepsilon(x_2). \quad (1.14)$$

Odhad relatívnej chyby bude

$$\delta(x_1 - x_2) = \frac{\varepsilon(x_1 - x_2)}{|x_1 - x_2|} = \frac{\Delta x_1 + \Delta x_2}{|x_1 - x_2|} \approx \frac{\varepsilon(x_1) + \varepsilon(x_2)}{|x_1 - x_2|}. \quad (1.15)$$

Keď vo vzťahu (1.15) použijeme relatívne chyby argumentov, potom

$$\delta(x_1 - x_2) = \frac{\Delta x_1 + \Delta x_2}{|x_1 - x_2|} = \left| \frac{x_1}{x_1 - x_2} \right| \cdot \frac{\Delta x_1}{|x_1|} + \left| \frac{x_2}{x_1 - x_2} \right| \cdot \frac{\Delta x_2}{|x_2|} \approx \frac{|x_1| \delta(x_1) + |x_2| \delta(x_2)}{|x_1 - x_2|}.$$

Veľkosť relatívnej chyby zásadným spôsobom ovplyvňuje veľkosť, resp. malosť menovateľa. Nebezpečnou je situácia, keď menovateľ $x_1 - x_2$ je veľmi malé číslo, t. j. hodnoty $x_1 \doteq x_2$. Vtedy odhad relatívnej chyby rozdielu značne prevýši odhady relatívnych chýb argumentov.

4. Súčin

Pre súčin dvoch čísel bude mať funkcia tvar $f(x_1, x_2) = x_1 \cdot x_2$. Parciálne derivácie sú

$\frac{\partial f(\bar{x})}{\partial x_1} = x_2$, $\frac{\partial f(\bar{x})}{\partial x_2} = x_1$. Odhad absolútnej chyby vyjadríme podľa (1.8) a (1.9) ako

$$\varepsilon(x_1 \cdot x_2) = |x_2| \cdot \Delta x_1 + |x_1| \cdot \Delta x_2 \approx |x_2| \cdot \varepsilon(x_1) + |x_1| \cdot \varepsilon(x_2) \quad (1.16)$$

a odhad relatívnej chyby bude

$$\delta(x_1 \cdot x_2) = \frac{\Delta x_1}{|x_1|} + \frac{\Delta x_2}{|x_2|} \approx \frac{\varepsilon(x_1)}{|x_1|} + \frac{\varepsilon(x_2)}{|x_2|}, \quad (1.17)$$

pri vyjadrení relatívnych chýb argumentov

$$\delta(x_1 \cdot x_2) \approx \delta(x_1) + \delta(x_2).$$

5. Podiel

Teraz funkcia f bude mať tvar $f(x_1, x_2) = \frac{x_1}{x_2}$. Parciálne derivácie sú $\frac{\partial f(\bar{x})}{\partial x_1} = \frac{1}{x_2}$,

$\frac{\partial f(\bar{x})}{\partial x_2} = \frac{-x_1}{x_2^2}$, $x_2 \neq 0$. Približný odhad absolútnej chyby vyjadríme podľa (1.8) a (1.9) ako

$$\varepsilon\left(\frac{x_1}{x_2}\right) = \left|\frac{1}{x_2}\right| \cdot \Delta x_1 + \left|\frac{x_1}{x_2^2}\right| \cdot \Delta x_2 \approx \left|\frac{1}{x_2}\right| \cdot \varepsilon(x_1) + \left|\frac{x_1}{x_2^2}\right| \cdot \varepsilon(x_2), \quad (1.18)$$

odhad relatívnej chyby bude

$$\delta\left(\frac{x_1}{x_2}\right) = \left|\frac{x_2}{x_1}\right| \cdot \left(\left|\frac{1}{x_2}\right| \cdot \Delta x_1 + \left|\frac{x_1}{x_2^2}\right| \cdot \Delta x_2\right) \approx \left|\frac{1}{x_2}\right| \cdot \varepsilon(x_1) + \left|\frac{x_1}{x_2^2}\right| \cdot \varepsilon(x_2) \quad (1.19)$$

a pri použití relatívnych chýb argumentov

$$\delta\left(\frac{x_1}{x_2}\right) = \delta(x_1) + \delta(x_2).$$

operácia	absolútna chyba a jej odhad	relatívna chyba a jej odhad
súčet, rozdiel	$\Delta(x_1 \pm x_2) \leq \Delta x_1 + \Delta x_2$ $\varepsilon(x_1 \pm x_2) \approx \varepsilon(x_1) + \varepsilon(x_2)$	$\left \frac{\Delta(x_1 \pm x_2)}{x_1 \pm x_2}\right \leq \left \frac{x_1}{x_1 \pm x_2}\right \cdot \left \frac{\Delta x_1}{x_1}\right + \left \frac{x_2}{x_1 \pm x_2}\right \cdot \left \frac{\Delta x_2}{x_2}\right $ $\delta(x_1 \pm x_2) \approx \frac{\varepsilon(x_1) + \varepsilon(x_2)}{ x_1 \pm x_2 } = \frac{ x_1 \delta(x_1) + x_2 \delta(x_2)}{ x_1 \pm x_2 }$
násobok	$\Delta(m \cdot x_1) = \Delta(x_1)$ $\varepsilon(m \cdot x_1) = \varepsilon(x_1)$	$\frac{\Delta(m \cdot x_1)}{ m \cdot x_1 } = \frac{\Delta(x_1)}{ x_1 }$ $\delta(m \cdot x_1) = \delta(x_1)$
súčin	$\Delta(x_1 \cdot x_2) \leq x_2 \cdot \Delta x_1 + x_1 \cdot \Delta x_2$ $\varepsilon(x_1 \cdot x_2) \approx x_2 \cdot \varepsilon(x_1) + x_1 \cdot \varepsilon(x_2)$	$\left \frac{\Delta(x_1 \cdot x_2)}{x_1 \cdot x_2}\right \leq \frac{\Delta x_1}{ x_1 } + \frac{\Delta x_2}{ x_2 }$ $\delta(x_1 \cdot x_2) \approx \frac{\varepsilon(x_1)}{ x_1 } + \frac{\varepsilon(x_2)}{ x_2 } = \delta(x_1) + \delta(x_2)$
podiel	$\Delta\left(\frac{x_1}{x_2}\right) \leq \left \frac{1}{x_2}\right \cdot \Delta x_1 + \left \frac{x_1}{x_2^2}\right \cdot \Delta x_2$ $\varepsilon\left(\frac{x_1}{x_2}\right) \approx \left \frac{1}{x_2}\right \cdot \varepsilon(x_1) + \left \frac{x_1}{x_2^2}\right \cdot \varepsilon(x_2)$	$\left \frac{\Delta\left(\frac{x_1}{x_2}\right)}{\frac{x_1}{x_2}}\right \leq \frac{\Delta x_1}{ x_1 } + \left \frac{x_1}{x_2}\right \cdot \frac{\Delta x_2}{ x_2 }$ $\delta\left(\frac{x_1}{x_2}\right) \approx \left \frac{1}{x_2}\right \cdot \varepsilon(x_1) + \left \frac{x_1}{x_2^2}\right \cdot \varepsilon(x_2) = \delta(x_1) + \delta(x_2)$

Tabuľka 1.2. Prehľad vzorcov na výpočet odhadov chýb aritmetických operácií

1.2.4 REPREZENTÁCIA ČÍSEL

Pri numerických výpočtoch nepracujeme na počítačoch zariadeniach (kalkulačky, počítače) s presnými číslami, ale s ich reprezentáciami. Základné východiská práce s číslami si ukážeme na dekadickom systéme, ktorý je najpoužívanejší tak vo vede, technike ako aj v občianskom živote. Reálne čísla v desiatkovej sústave sú reprezentované v nasledujúcom tvare:

$$x = \pm c_n \cdot 10^n + c_{n-1} \cdot 10^{n-1} + \dots + c_1 \cdot 10^1 + c_0 \cdot 10^0 + \frac{d_1}{10^1} + \frac{d_2}{10^2} + \frac{d_3}{10^3} + \dots,$$

kde c_i a d_i sú z množiny číslic (cifrier) $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, n je nezáporné celé číslo a pre $x \neq 0$ je prvá z uvedených číslic nenulová. Na desatinných miestach obyčajne uvádzame číslice d_i až po poslednú nenulovú číslicu. Mocniny desiatok jednoznačne určujú pozície číslic a pri zápise sa nevypisujú:

$$x = \pm c_n c_{n-1} \dots c_1 c_0, d_1 d_2 d_3 \dots \quad (1.20)$$

Čiarka medzi číslicami sa nazýva desatinná čiarka a oddeľuje celú časť čísla od desatinnej. V anglicky hovoriacich krajinách ako aj vo veľkej časti Ázie sa ako oddeľovač používa desatinná bodka.

Pri určovaní presnosti čísla má veľký význam **počet platných číslic**. Za platné číslice považujeme všetky číslice posudzovaného čísla okrem núl nachádzajúcich sa pred prvou nenulovou číslicou s výnimkou celých čísel, keď nuly za poslednou nenulovou číslicou môžu byť platné ale aj neplatné. Napríklad 4,320 má 4 platné číslice, 0,00365 má 3 platné číslice a číslo 4300 môže mať 2, 3 ale aj všetky 4 platné číslice. Pri celých číslach sa v prípade potreby počet platných číslic uvádza v komentári alebo sa označuje dohovorenou značkou, napríklad vodorovnou čiarkou pod, alebo nad poslednou platnou číslicou: 4300 označuje 3 platné číslice. Počet platných číslic sa používa najmä v spojení so zaokrúhľovaním.

1.2.4.1 Zakrúhľovanie

Pri zaokrúhľovaní čísla (1.20) na k platných číslic postupujeme takto:

1. Ak číslica nachádzajúca sa na $(k+1)$ -vom platnom mieste je číslica v celej časti čísla
 - a) $c_i < 5$ potom zaokrúhľujeme smerom nadol, t. j. číslica na k -tom mieste sa nemení a zvyšné nasledujúce číslice c_i položíme rovné 0. Napríklad pri zaokrúhľovaní na 2 platné číslice číslo $7842,6 \doteq 7800$.
 - b) $c_i \geq 5$ potom zaokrúhľujeme smerom nahor, t. j. číslica na k -tom mieste sa zväčší o 1 a zvyšné nasledujúce číslice c_i položíme rovné 0. Napríklad pri zaokrúhľovaní na 2 platné číslice číslo $7862,6 \doteq 7900$ alebo číslo $797 \doteq 800$.
2. Ak číslica nachádzajúca sa na $(k+1)$ -vom platnom mieste je číslica v desatinnej časti čísla
 - a) $d_i < 5$ potom zaokrúhľujeme smerom nadol, t. j. číslica na k -tom mieste sa nemení a zvyšné nasledujúce číslice d_i neuvádzame. Napríklad pri zaokrúhľovaní na 2 platné číslice číslo $19,3 \doteq 19$ alebo $0,734 \doteq 0,73$.

- b) $d_i \geq 5$ potom zaokrúľujeme smerom nahor, t. j. číslica na k -tom mieste sa zväčší o 1 a zvyšné nasledujúce číslice d_i neuvádzame. Napríklad pri zaokrúľovaní na 2 platné číslice číslo $18,5 \doteq 19$ alebo číslo $0,798 \doteq 0,80$.

Zaokrúľovanie na istý počet platných číslic má pri práci s presnosťou oveľa väčší význam ako zaokrúľovanie na istý počet desatinných miest. Napríklad ak zaokrúhlime čísla $x = 0,2365$ a $y = 23650$ na 3 platné číslice, potom $x^* = 0,237$ a $y^* = 23700$ majú takú istú relatívnu chybu. Presvedčte sa o tom.

1.2.4.2 Semilogaritmický tvar čísla

Existuje veľa spôsobov zápisu reálneho čísla, napr.

$$\begin{aligned} -273,16 &= -2,7316 \cdot 10^2 \\ &= -0,27316 \cdot 10^3 \\ &= -27316 \cdot 10^{-2} \\ &= -2,7316e2, \end{aligned}$$

pričom $-2,7316e2$ je analogickým zápisom $-2,7316 \cdot 10^2$, kde namiesto „ 10^2 “ uvádzame iba exponent 2, „ $e2$ “.

Jednotlivé zápisy sa od seba líšia umiestnením desatinnej čiarky. V takomto prípade hovoríme o tvare čísla s pohyblivou alebo plávajúcou desatinnou čiarkou. V snahe zjednotiť zápis reálnych čísel bol zavedený (najmä na prácu s počítačami zariadeniami) semilogaritmický tvar čísla s normalizovanou mantisou.

Označme $M(q)$ systém normalizovaných čísel s pohyblivou (plávajúcou) rádovou čiarkou, kde číslo $q \geq 2$ je základ číselnej sústavy ($q \in N$). Podľa použitého základu (rádu) číselnej sústavy q používame príslušný názov číselnej sústavy:

- napr. pre $q = 2$ – binárna (dvojková) sústava,
 $q = 10$ – dekadická (desiatková) sústava,
 $q = 16$ – hexadecimálna (šest'nástková) sústava,

ako aj názov rádovej čiarky:

- $q = 2$ – binárna čiarka,
 $q = 10$ – desatinná čiarka,
 $q = 16$ – hexadecimálna čiarka.

Každé číslo $x \in M(q)$ má tvar

$$x = \pm a \cdot q^b = \pm \left(\sum_{i=1}^{\infty} d_i \cdot q^{-i} \right) \cdot q^b = \pm \left(\frac{d_1}{q^1} + \frac{d_2}{q^2} + \frac{d_3}{q^3} + \dots + \frac{d_p}{q^p} + \dots \right) \cdot q^b, \quad d_1 \neq 0$$

kde

- a je normalizovaná mantisa ($d_1 \neq 0$), obsahujúca číslice $d_i \in \{0, 1, 2, \dots, b-1\}$, $i=1, 2, \dots, p$, $p \in N$ je počet číslic (miest pre číslice) v mantise,
- $q \geq 2$ je základ číselnej sústavy,
- $b \in Z$ je exponent.

Napríklad v čísle $x = -273,16 = -0,27316 \cdot 10^3$ normalizovaná mantisa je $a = 0,27316$, základ číselnej sústavy je $q = 10$ a exponent $b = 3$.

Definícia 1.1 Nech $x^* = a^* \cdot 10^b$ je aproximáciou čísla $x = a \cdot 10^b$ (a je normalizovaná mantisa). Hovoríme, že j -ta dekadická číslica aproximácie x^* je platná, ak platí

$$|x - x^*| \leq 0,5 \cdot 10^{b-j} \quad \text{alebo} \quad |a - a^*| \leq 0,5 \cdot 10^{-j}.$$

Ak uvedené nerovnosti platia pre $j = 1, 2, \dots, t$, ale neplatia už pre $j = t+1$, potom hovoríme, že x^* má t platných číslic.

Poznámka. Počet platných číslic približnej hodnoty x^* vieme určiť z absolútnej chyby $|x - x^*|$ jednoduchým odčítaním počtu núl t za desatinnou čiarkou (počítaných zľava), za ktorými nasleduje jedna z číslic 0 až 5 (napríklad ak $x = 0,320 \cdot 10^3$ a $x^* = 0,31842 \cdot 10^3$, potom $|0,320 - 0,31842| = 0,0168$, za desatinnou čiarkou sa nachádza jedna 0, za ktorou je číslica 1 (jedna z číslic 0 až 5), t. j. x^* má jednu platnú číslicu: $0,0168 \leq 0,5 \cdot 10^{-1}$.

Poznámka

- V správne zaokrúhlenom čísle je každá číslica platná.
- U čísel $x^* = a^* \cdot 10^b$, kde a^* nie je normalizovaná mantisa, sa často používa termín platné desatinné miesto.

Nasledujúca veta pojednáva o tom, ako súvisí veľkosť relatívnej chyby aproximácie x^* čísla x s počtom jeho platných číslic.

Veta 1.3 Nech aproximácia x^* čísla x je kladná a má t platných číslic. Potom pre jej relatívnu chybu platí

$$\frac{|x - x^*|}{|x|} \leq \frac{1}{d_1} \cdot 10^{1-t},$$

kde d_1 je prvá platná číslica aproximácie x^* .

Dôsledok. Ak aproximácia x^* čísla x má aspoň 2 platné číslice, t. j. $t \geq 2$, potom

$$\frac{|x - x^*|}{|x|} \leq \frac{1}{2d_1} \cdot 10^{1-t}$$

1.2.4.3 Reprezentácia čísla v počítači

Reálne čísla sú v počítači reprezentované v systéme čísel s plávajúcou rádovou čiarkou (počet číslic mantisy je konštantný). Niekedy, ak je to výhodnejšie, používa sa na reprezentáciu určitej skupiny čísel systém s pevnou rádovou čiarkou (počet miest za čiarkou je konštantný). Sem patria napr. čísla typu „INTEGER“ v programovacom jazyku Fortran.

Rozšírime systém $M(q)$ normalizovaných čísel s plávajúcou rádovou čiarkou na systém $M(q, p, L, U)$, skrátene $M(q, p)$, kde

- $p \geq 1$ je počet číslic mantisy ($p \in N$),
- L, U sú hranice rozsahu exponentu ($L < 0 < U$).

Parametre p, L a U sú dané konštrukciou počítača.

Každé číslo $x \in M(q, p)$ má potom tvar

$$x = \pm a \cdot q^b = \pm \left(\sum_{i=1}^p d_i \cdot q^{-i} \right) \cdot q^b = \pm \left(\frac{d_1}{q^1} + \frac{d_2}{q^2} + \frac{d_3}{q^3} + \dots + \frac{d_p}{q^p} \right) \cdot q^b, \quad d_1 \neq 0.$$

Počítač, ktorý pracuje s číslami z $M(q, p)$, voláme q -adický p -miestny počítač.

Pri vkladaní vstupných údajov do počítača počítač každému číslu $x \in R$ priradí číslo $x^* \in M(q, p)$. Počet čísel, ktoré tvoria systém $M(q, p)$ je konečný (čísla majú konečný rozvoj daný počtom miest mantisy). Ak číslo x medzi nimi nie je, potom x sa aproximuje číslom x^* . Priradenie sa realizuje orezaním alebo zaokrúhlením čísla x , čím dochádza k chybe zobrazenia.

1. Orezanie čísla x

Ak číslo

$$x = \pm a \cdot q^b = \pm \left(\sum_{i=1}^{\infty} d_i \cdot q^{-i} \right) \cdot q^b = \pm \left(\frac{d_1}{q^1} + \frac{d_2}{q^2} + \frac{d_3}{q^3} + \dots + \frac{d_p}{q^p} + \dots \right) \cdot q^b$$

orežeme na konečný počet p číslic, dostaneme číslo

$$x^* = \pm a^* \cdot q^b = \pm \left(\sum_{i=1}^p d_i \cdot q^{-i} \right) \cdot q^b = \pm \left(\frac{d_1}{q^1} + \frac{d_2}{q^2} + \frac{d_3}{q^3} + \dots + \frac{d_p}{q^p} \right) \cdot q^b,$$

napr. číslu $x = 14,5714732 = 0,145714732 \cdot 10^2$ bude v $M(10,6)$ priradené orezaním číslo $x^* = 0,145714 \cdot 10^2 = 14,5714$.

Úprava čísel orezaním je typická napríklad pre procesory AMD.

2. Zaokrúhlenie čísla x

Ak číslo

$$x = \pm a \cdot q^b = \pm \left(\sum_{i=1}^{\infty} d_i \cdot q^{-i} \right) \cdot q^b = \pm \left(\frac{d_1}{q^1} + \frac{d_2}{q^2} + \frac{d_3}{q^3} + \dots + \frac{d_p}{q^p} + \frac{d_{p+1}}{q^{p+1}} \dots \right) \cdot q^b$$

zaokrúhlime na konečný počet p číslic, potom

ak číslica $d_{p+1} < \frac{q}{2}$ dostaneme číslo

$$x^* = \pm a^* \cdot q^b = \pm \left(\sum_{i=1}^p d_i \cdot q^{-i} \right) \cdot q^b = \pm \left(\frac{d_1}{q^1} + \frac{d_2}{q^2} + \frac{d_3}{q^3} + \dots + \frac{d_p}{q^p} \right) \cdot q^b.$$

ak číslica $d_{p+1} \geq \frac{q}{2}$ dostaneme číslo

$$x^* = \pm a^* \cdot q^b = \pm \left(\sum_{i=1}^{p-1} d_i \cdot q^{-i} + (d_p + 1) \cdot q^{-p} \right) \cdot q^b = \pm \left(\frac{d_1}{q^1} + \frac{d_2}{q^2} + \frac{d_3}{q^3} + \dots + \frac{d_{p-1}}{q^{p-1}} + \frac{d_p + 1}{q^p} \right) \cdot q^b,$$

t. j. číslica d_p sa zaokrúhli číslicou d_{p+1} podľa pravidiel zaokrúhľovania.

napr. číslu $x=14,5714732=0,145714732 \cdot 10^2$ bude v $M(10,6)$ priradené zaokrúhlením číslo $x^*=0,145715 \cdot 10^2=14,5715$.

Úprava čísel zaokrúhlením je typická napríklad pre procesory Intel.

Množina čísel $M(q, p, L, U)$, s ktorými pracuje dnešný počítač je síce veľká, ale vždy je konečná. Jej veľkosť je určená architektúrou procesnej jednotky počítača.

Počet čísel v $M(q, p, L, U)$ sa dá vyjadriť ako

$$2 \cdot (q-1) \cdot q^{p-1} \cdot (U-L+1) + 1$$

lebo

- môžeme mať 2 znamienka (+, -),
- $(q-1)$ možností pre prvú číslicu mantisy,
- q možností pre ostatné číslice mantisy,
- $(U-L+1)$ možností hodnôt exponentu b ,
- (+1) je jedna nula.

Hardvér počítača pracuje v binárnej (dvojkovej) sústave. V tejto sústave sú číselné údaje reprezentované najjednoduchším spôsobom – pomocou 2 symbolov sústavy, číslic 0 a 1, napr.

$$729_{10} = 1011011001_2$$

$$729_{10} = 7 \cdot 10^2 + 2 \cdot 10^1 + 9 \cdot 10^0$$

$$1011011001_2 = 1 \cdot 2^9 + 0 \cdot 2^8 + 1 \cdot 2^7 + 1 \cdot 2^6 + 0 \cdot 2^5 + 1 \cdot 2^4 + 1 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0$$

Reprezentácia čísel v počítači je ovplyvnená viacerými vlastnosťami, napríklad

- celkovým disponibilným počtom bitov,
- spôsobom priradenia mantisy a exponentu,
- rozdelením bitov pre mantisu a exponent,
- poradím mantisy a exponentu atď.

Pri voľbe počtu bitov pre zobrazenie mantisy a exponentu rozhoduje požadovaný:

1. **rozsah** (angl. range) systému $M(q, p)$, t. j. najmenšie a najväčšie číslo v absolútnej hodnote, rôzne od nuly, ktoré sa dá zobraziť.

a) najmenšie kladné číslo v $M(q, p)$ sa zvyčajne označuje ako UFL (z angl. underflow level), hranica podtečenia. Prvá číslica mantisy sa rovná jednej, ostatné sú nulové, pričom exponent je najmenší možný.

$$\text{UFL} = q^L$$

b) najväčšie kladné číslo v $M(q, p)$ sa zvyčajne označuje ako OFL (z angl. overflow level), hranica pretečenia. Všetky číslice mantisy má rovné $q-1$, pričom exponent je najväčší možný.

$$\text{OFL} = (q - q^{1-p}) \cdot q^U$$

2. **presnosť** (angl. accuracy), ktorá hovorí o blízkosti aproximácie čísla x^* k presnej hodnote x , napr. číslo π môže byť zobrazené ako 3,142 alebo 3,14159. Ich absolútne chyby potom sú $|\pi - 3,142| = 0,0004073\dots$ a $|\pi - 3,14159| = 0,000002654\dots$
3. **počet platných číslíc** (angl. precision), tzv. rozlišovacia schopnosť. Je to presnosť zobrazenia čísla, napr. $\pi = 3,142$ (počet platných číslíc 4) alebo $\pi = 3,14159$ (počet platných číslíc 6).

Reálne čísla, ktoré sa dajú v systéme $M(q, p)$ zobraziť presne sa nazývajú **strojové čísla**. Ostatné čísla sa musia aproximovať „blízkymi“ strojovými číslami. **Strojová presnosť** – tzv. **strojové epsilon** udáva číslo $\varepsilon_m = q^{1-p}$.

Pri numerických výpočtoch vždy pracujeme s číslami vyjadrenými na príslušný počet desatinných (rádových) miest. Prvé verzie počítačov prebytočné miesta odstraňovali, čo viedlo k nepresnostiam. Snahou bolo presadiť lepšiu manipuláciu s reálnymi číslami. V počítačoch vyvinutých od roku 1985 sa reálne čísla zobrazujú podľa štandardu IEEE 754.

Štandard IEEE 754 (IEEE Standard for Floating Point Arithmetic) je najrozšírenejší štandard na výpočty v plávajúcej rádovej čiarky, ktorý požíva veľa mikroprocesorov a jednotiek FPU. Aktuálna verzia štandardu je IEEE 754 - 2008, ktorá bola zverejnená v auguste 2008. Zahŕňa takmer všetky pôvodné verzie IEEE 754 do r. 1985 a IEEE štandard pre RADIX - independent floating point.

IEEE 754 definuje 5 formátov čísla, 3 binárne (používajúce 32, 64 a 128 bitov) pre dvojkovú aritmetiku v plávajúcej binárnej čiarky a 2 dekadické (používajúce 64 a 128 bitov) pre desiatkovú aritmetiku v plávajúcej desatinnej čiarky. Najčastejšie sú používané nasledovné dva binárne formáty:

1. Jednoduchá presnosť (single precision, 32 bitov)

Používa 4 bajty, t. j. 32 bitov, z toho je 23 bitov rezervovaných pre mantisu, 8 bitov pre exponent, 1 bit pre znamienko. Pretože mantisa je normalizovaná, pre $x \neq 0$ je $d_1 = 1$. Táto číslica sa neukladá. Mantisa má približne 7 dekadických číslíc presnosti.

Jednoduchá presnosť (Single Precision)	
Počet bitov mantisy	$p = 24$ (7 miest v dek. sústave)
Rozsah exponentu	$-126 \leq b \leq 127$
Strojová presnosť	$\varepsilon_m = 2^{-23} \doteq 1,2 \cdot 10^{-7}$
Hranica podtečenia	$UFL = 2^{-126} \doteq 1,2 \cdot 10^{-38}$
Hranica pretečenia	$OFL = (2 - 2^{-23}) \cdot 2^{127} \doteq 3,4 \cdot 10^{38}$

Tabuľka 1.3. Parametre pre jednoduchú presnosť

2. Dvojnásobná presnosť (double precision, 64 bitov)

Používa 8 bajtov, t. j. 64 bitov, z toho je 52 bitov rezervovaných pre mantisu, 11 bitov pre exponent, 1 bit pre znamienko. Mantisa je normalizovaná, pre $x \neq 0$ je $d_1 = 1$. Táto číslica sa neukladá. Mantisa má približne 17 dekadických číslic presnosti.

Dvojnásobná presnosť (Double Precision)	
Počet bitov mantisy	$p = 53$ (17 miest v dek. sústave)
Rozsah exponentu	$-1022 \leq b \leq 1023$
Strojová presnosť	$\varepsilon_m = 2^{-52} \doteq 2,2 \cdot 10^{-16}$
Hranica podtečenia	$UFL = 2^{-1022} \doteq 2,2 \cdot 10^{-308}$
Hranica pretečenia	$OFL = (2 - 2^{-52}) \cdot 2^{1023} \doteq 1,8 \cdot 10^{308}$

Tabuľka 1.4. Parametre pre dvojnásobnú presnosť

Splnenie štandardu vyžaduje implementáciu najmenej jedného zo základných piatich formátov, ostatné sú voliteľné.

Štandard definuje aj ďalšie kódovania pre reprezentáciu čísel v plávajúcej rádovej čiarko:

- INF pre výraz typu $+\infty$, výsledok operácie $1/0$,
- -INF pre výraz typu $-\infty$, výsledok operácie $-1/0$,
- NaN (not a number) pre „nie je číslo“, výsledok operácie $0/0$,
- UFLs pre denormalizované čísla (subnormálne, nenulové nenormalizované) s možným exponentom $b = L$, $UFLs = \varepsilon_m \cdot UFL$.

Počítačová aritmetika podľa štandardu IEEE:

- Výsledok aritmetickej operácie vykonanej v počítači je rovnaký ako keď operáciu vykonáme presne a potom získaný výsledok vložíme do počítača.
- Ak je absolútna hodnota výsledku aritmetickej operácie je väčšia ako OFL, dochádza k tzv. pretečeniu (overflow, výsledok je INF).
- Ak je absolútna hodnota výsledku aritmetickej operácie je menšia ako UFL, dochádza k tzv. podtečeniu (underflow, výsledok je -INF).

Dôsledky používania štandardu pre výpočty v plávajúcej rádovej čiarko:

1. Pripočítanie malého nenulové čísla k číslu nemá vplyv na výsledok

$$5,18 \cdot 10^2 + 4,37 \cdot 10^{-1} = 5,18 \cdot 10^2 + 0,00437 \cdot 10^2 = 5,18437 \cdot 10^2 \doteq 5,18 \cdot 10^2.$$

2. Strojové epsilon (najväčšie epsilon, pre ktoré platí $1 + \varepsilon = 1$) udáva, aký je maximálny počet platných číslic pre číslo, ktoré je vyjadrené pohyblivou rádovou čiarko.

3. Vo všeobecnosti neplatí:

a) **asociatívny zákon** $(a+b)+c \neq a+(b+c)$.

b) **rovnosti** typu: $x \frac{1}{x} = 1$, $(1+x)-1 = x$, a pod.

4. Pri aritmetických operáciách dochádza **ku strate platných čísiel**.

1.2.5 PODMIENENOSŤ NUMERICKÝCH ÚLOH A NUMERICKÁ STABILITA ALGORITMOV

Medzi numerickými úlohami sa vyskytujú úlohy, ktoré sú citlivé na zmeny vstupných údajov. Prejavuje sa to veľkým rozptylom riešení (výstupných údajov), t. j. malá zmena na vstupe spôsobuje veľkú zmenu na výstupe. Preto v numerických postupoch je nutná analýza závislosti vstupných a výstupných údajov.

Definícia 1.2 Nech A je množina vstupných údajov a B je množina výstupných údajov. Hovoríme, že numerická úloha $y = f(x)$, kde $x \in A$ a $y \in B$ je korektná na dvojici množín (A, B) , ak

a) pre každé $x \in A$ existuje jediné riešenie $y \in B$,

b) toto riešenie spojitostou závisí na vstupných údajoch, t. j. keď $x_n \rightarrow x$, $U(x_n) = y_n$, potom $U(x_n) \rightarrow U(x)$, t. j. $y_n \rightarrow y$.

V opačnom prípade hovoríme o nekorektných úlohách. Veľkú skupinu nekorektných úloh tvoria nejednoznačne riešiteľné úlohy. My budeme skúmať len korektné úlohy.

Definícia 1.3 Hovoríme, že korektná úloha je dobre podmienená, ak malá zmena vo vstupných údajoch vyvolá malú zmenu riešenia (výstupných údajov).

Definícia 1.4 Číslo c_p , ktoré určuje veľkosť týchto zmien, nazývame číslom podmienenosti úlohy a definujeme ho ako

$$c_p = \frac{\text{relatívna chyba na výstupe}}{\text{relatívna chyba na vstupe}} = \frac{\delta(\text{výstup})}{\delta(\text{vstup})}.$$

Presná hodnota relatívnej chyby zvyčajne nie je známa. Na jej určenie používame odhad. Ak $c_p \approx 1$, je úloha dobre podmienená, pre veľké c_p (>100) je úloha zle podmienená.

Podmienenosť úlohy vyjadruje závislosť nepresnosti výsledku úlohy od nepresnosti vstupných údajov:

$$c_p = \frac{\delta(\text{výstup})}{\delta(\text{vstup})} \Rightarrow \delta(\text{výstup}) = c_p \cdot \delta(\text{vstup}).$$

Pri realizácii algoritmu na počítači pri vykonávaní aritmetických operácií vždy dochádza k zaokrúhľovacím chybám, pretože počítač pracuje s číslami zo systému $M(q, p)$. Tieto chyby nemožno vylúčiť aj keby vstupné údaje boli presné. Otázkou je ako posúdiť citlivosť každého konkrétneho algoritmu na zaokrúhľovacie chyby a chyby vstupných údajov. Keď sa

chceme vyvarovať nezmyselným výsledkom, musíme vyberať algoritmy málo citlivé na uvedené vplyvy, tzv. stabilné algoritmy.

Aby bol **algoritmus stabilný**, musí byť:

- a) **dobře podmienený**, t. j. málo citlivý na poruchy vo vstupných údajoch,
- b) **numericky stabilný**, t. j. jeho numerická realizácia musí byť málo citlivá na vplyv zaokrúhľovacích chýb počas výpočtu.

Predstavu o stabilite algoritmu a presnosti získaných výsledkov môžeme okrem teoretického spôsobu dostať napríklad **metódou počítačových experimentov**, keď riešime niekoľko rovnakých úloh so zmenenými vstupnými údajmi (malá zmena) a sledujeme, ako sa prejavia poruchy vo vstupných údajoch na výsledku (výstupných údajoch). Takýto postup nazývame **metóda experimentálnych perturbácií**.

Sledovanie stability algoritmov je v numerickej matematike jedna z kľúčových úloh.

1.3 RIEŠENÉ PRÍKLADY

PRÍKLAD č. 1

Nech $x^* = 2,71$ je približnou hodnotou (aproximáciou) čísla $x = e = 2,7182818\dots$ (Vzniklo orezaním.)

1. Vypočítajte
 - a) absolútnu chybu aproximácie x^* ,
 - b) relatívnu chybu aproximácie x^* ,
2. Navrhните dva odhady absolútnej chyby.
3. Na základe odhadov absolútnej chyby stanovte interval a jeho dĺžku, v ktorom musí ležať presná hodnota e . Overte. Ktorý z odhadov je lepší?
4. Overte vetu 1.3 a jej dôsledok.

RIEŠENIE

1. Výpočet chýb

- a) absolútna chyba aproximácie x^* :

$$\Delta x = |x - x^*| = |e - 2,71| = 0,0082818\dots$$

- b) relatívna chyba aproximácie x^* :

$$\frac{\Delta x}{|x|} = \frac{0,0082818\dots}{|e|} = \frac{0,0082818\dots}{2,7182818\dots} = 0,00304671 = 0,304671\%$$

2. Návrh dvoch odhadov absolútnej chyby.

Pre odhad absolútnej chyby platí: $\Delta x \leq \varepsilon(x)$. Navrhujeme:

prvý odhad: $\varepsilon_1(x) = 0,01$. Odhad je správny, $0,0082\dots < 0,01$.

druhý odhad: $\varepsilon_2(x) = 0,009$. Odhad je správny, $0,0082\dots < 0,009$.

3. Stanovenie intervalov v ktorom leží presná hodnota e .

Pre presnú hodnotu x platí: $x \in \langle x^* - \varepsilon(x), x^* + \varepsilon(x) \rangle$.

Prvý odhad: $\varepsilon_1(x) = 0,01$; $e \in \langle x^* - 0,01, x^* + 0,01 \rangle = \langle 2,70; 2,72 \rangle$.

Dĺžku intervalu vypočítame ako: $d = 2 \cdot \varepsilon_1(x) = 0,02$ alebo ako rozdiel hornej a dolnej hranice intervalu: $d = 2,72 - 2,70 = 0,02$.

Presná hodnota $x = e = 2,7182818... \in \langle 2,70; 2,72 \rangle$

Druhý odhad: $\varepsilon_2(x) = 0,009$; $e \in \langle x^* - 0,009, x^* + 0,009 \rangle = \langle 2,701; 2,719 \rangle$.

Dĺžka intervalu je $d = 2 \cdot \varepsilon_2(x) = 0,018 = 2,719 - 2,701$.

Presná hodnota $x = e = 2,7182818... \in \langle 2,701; 2,719 \rangle$

Druhý odhad je lepší, pretože je menší ako prvý odhad. Presnejšie určuje presnú hodnotu e . Dĺžka intervalu (0,018), v ktorom leží presná hodnota e je menšia, ako dĺžka intervalu určitého 1. odhadom (0,02).

4. Overenie vety Veta 1.3.

Podľa vety Veta 1.3 pre odhad relatívnej chyby platí $\frac{|x - x^*|}{|x|} \leq \frac{1}{d_1} \cdot 10^{1-t}$, kde x^* je kladná aproximácia čísla x a má t platných číslíc.

Najskôr zistíme počet platných číslíc aproximácie $x^* = 2,71$.

Na zistenie počtu platných číslíc môžeme použiť definíciu Definícia 1.1 alebo poznámku k tejto definícii. Číslo x a x^* vyjadríme s normalizovanou mantisou.

$$x = e = 0,27182818... \cdot 10^{-1}, x^* = 0,271 \cdot 10^{-1}.$$

Podľa poznámky: $\Delta x = |x - x^*| = 0,00082818... \cdot 10^{-1}$. Za desatinnou čiarkou sa nachádzajú dve nuly, za nimi je číslica 0 (jedna z číslíc 0 až 5), t. j. x^* má dve platné číslice.

Podľa definície: $0,00082818... \leq 0,005$, t. j. $0,00082818... \leq 0,5 \cdot 10^{-2}$. x^* má dve platné číslice.

Overenie vety:

$$x^* = 2,71 \text{ má dve platné číslice, t. j. vo vzorci } \frac{|x - x^*|}{|x|} \leq \frac{1}{d_1} \cdot 10^{1-t} \text{ } t = 2 \text{ a } d_1 = 2.$$

$$\frac{|x - x^*|}{|x|} \leq \frac{1}{2} \cdot 10^{1-2}, \text{ t. j. } 0,00304671 \leq 0,05, \text{ čo je pravdivý výrok a veta Veta 1.3 platí.}$$

Overenie dôsledku vety Veta 1.3.

Číslo $x^* = 2,71$ má aspoň dve platné číslice a podľa dôsledku vety Veta 1.3 by malo platiť:

$$\frac{|x - x^*|}{|x|} \leq \frac{1}{2d_1} \cdot 10^{1-t}, \text{ t. j. } \frac{|x - x^*|}{|x|} \leq \frac{1}{2 \cdot 2} \cdot 10^{1-2}, 0,00304671 \leq 0,025, \text{ čo takisto platí.}$$

PRÍKLAD č. 2

Na koľko desatinných miest treba vypočítať približnú hodnotu $x^* \doteq \sqrt{30}$, aby jej relatívna chyba bola menšia alebo rovná 0,001?

RIEŠENIE

Aby sme mohli použiť vzorec $\frac{|x-x^*|}{|x|} \leq \frac{1}{d_1} \cdot 10^{1-t}$, musíme zistiť prvú platnú číslicu d_1

v aproximácii $x^* \doteq \sqrt{30} = 5,477\dots$, t. j. $d_1 = 5$. Potom podľa uvedeného vzorca dostávame

$$\frac{|x-x^*|}{|x|} \leq \frac{1}{5} \cdot 10^{1-t} \leq 0,001$$

$$\frac{1}{5} \cdot 10^{1-t} \leq 10^{-3}$$

$$1-t \leq -3$$

$$t \geq 4$$

Ak približnú hodnotu $x^* \doteq \sqrt{30}$ vypočítame aspoň na 4 platné číslice, t. j. $x^* = 5,477$ bude mať aspoň 3 desatinné miesta, potom jej relatívna chyba bude menšia alebo rovná 0,001. Overíme:

$$\frac{|x-x^*|}{|x|} = \frac{|\sqrt{30}-5,477|}{|\sqrt{30}|} = 0,0000411842 < 0,001.$$

V prípade, že číslo $x^* \doteq \sqrt{30}$ bude mať viac ako 2 platné číslice, môžeme odhad relatívnej chyby zmenšiť na základe vzťahu

$$\frac{|x-x^*|}{|x|} \leq \frac{1}{2d_1} \cdot 10^{1-t}.$$

Počítame

$$\frac{|x-x^*|}{|x|} \leq \frac{1}{2 \cdot 5} \cdot 10^{1-t} \leq 0,001$$

$$10^{-t} \leq 10^{-3}$$

$$-t \leq -3$$

$$t \geq 3$$

Ak vezmeme približnú hodnotu $x^* \doteq \sqrt{30}$ aspoň na 3 platné číslice, t. j. $x^* = 5,48$ bude mať aspoň 2 desatinné miesta, potom jej relatívna chyba by mala byť tiež menšia alebo sa rovnáť 0,001. Overíme:

$$\frac{|x-x^*|}{|x|} = \frac{|\sqrt{30}-5,48|}{|\sqrt{30}|} = 0,000506538 < 0,001.$$

PRÍKLAD č. 3

Čísla $x = 2,372$ a $y = 2,370$ sú správne zaokrúhlené. Nájdite odhad absolútnej chyby, odhad relatívnej chyby rozdielu $(x - y)$ a dĺžku najmenšieho z intervalov, o ktorých viete, že v ňom leží presná hodnota rozdielu $(x - y)$.

RIEŠENIE

Obidve čísla sú správne zaokrúhlené na tri desatinné miesta, preto ich odhad absolútnej chyby je $\varepsilon(x) = \varepsilon(y) = 0,0005$.

Rozdiel $(x - y)$ vyjadríme ako funkciu dvoch premenných: $f(x, y) = x - y$. Hodnoty jej partiálnych derivácií sú $\frac{\partial f(x, y)}{\partial x} = 1$, $\frac{\partial f(x, y)}{\partial y} = -1$ pre všetky x, y . Odhad absolútnej chyby rozdielu potom podľa (1.8) a (1.9) bude

$$\varepsilon(x - y) \approx \left| \frac{\partial f(x, y)}{\partial x} \right| \cdot \varepsilon(x) + \left| \frac{\partial f(x, y)}{\partial y} \right| \cdot \varepsilon(y) = \varepsilon(x) + \varepsilon(y) = 0,0005 + 0,0005 = 0,001$$

Pretože hodnota rozdielu $x - y = 2,372 - 2,370 = 0,002$, bude odhad relatívnej chyby rozdielu

$$\delta(x - y) \approx \frac{\varepsilon(x) + \varepsilon(y)}{|x^* - y^*|} = \frac{0,001}{|0,002|} = 0,5 = 50\%$$

Presná hodnota skúmaného rozdielu bude ležať v intervale:

$$(x - y) \in \langle 0,002 - 0,001; 0,002 + 0,001 \rangle = \langle 0,001; 0,003 \rangle,$$

ktorého dĺžka je 0,002.

PRÍKLAD č. 4

Vyšetrte podmienenosť úlohy, v ktorej treba vypočítať hodnotu funkcie $f(x) = \frac{1}{10 - 7x}$ pre x^* z blízkeho okolia bodu x :

- $x = \sqrt{2}$, ak x^* je hodnota $x = \sqrt{2}$ zaokrúhlená na dve a tri platné číslice.
- $x = 10$, ak $x^* = 10,1$ a $x^* \doteq 10,01$.

Zdôvodnite.

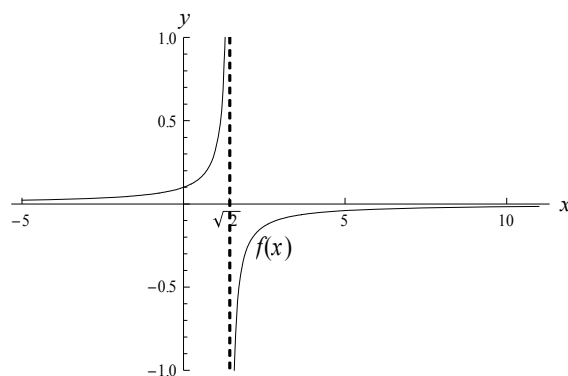
RIEŠENIE

Podmienenosť úlohy zistíme pomocou čísla podmienenosti

$$c_p = \frac{\text{relatívna chyba na výstupe}}{\text{relatívna chyba na vstupe}} = \frac{\delta(\text{výstup})}{\delta(\text{vstup})}.$$

- Podmienenosť úlohy v okolí bodu $x = \sqrt{2}$

V okolí bodu $x = \sqrt{2}$ vyšetríme podmienenosť úlohy pre $x^* \doteq \sqrt{2}$ zaokrúhlenú na dve a tri platné číslice.

Obr. 1.10. Graf funkcie $f(x)$ v okolí skúmaných bodov $x = \sqrt{2}$ a $x = 10$

$$x^* = 1,4$$

Relatívna chyba na vstupe:

$$\delta(\text{vstup}) = \frac{|x - x^*|}{|x|} = \frac{|\sqrt{2} - 1,4|}{|\sqrt{2}|} = 0,0100505\dots \doteq 1\%$$

Relatívna chyba na výstupe:

$$\delta(\text{výstup}) = \frac{|f(x) - f(x^*)|}{|f(x)|} = \frac{|f(\sqrt{2}) - f(1,4)|}{|f(\sqrt{2})|} = \frac{|9,94975\dots - 5|}{9,94975\dots} = 0,497475\dots \doteq 49,75\%$$

Číslo podmienenosti:

$$c_p = \frac{\text{relatívna chyba na výstupe}}{\text{relatívna chyba na vstupe}} = \frac{0,497475\dots}{0,010050\dots} = 49,49747\dots \doteq 49,50,$$

a teda $\delta(\text{výstup}) = 49,50 \cdot \delta(\text{vstup})$.

$$x^* = 1,41$$

Relatívna chyba na vstupe:

$$\frac{|x - x^*|}{|x|} = \frac{|\sqrt{2} - 1,41|}{|\sqrt{2}|} = 0,002979\dots \doteq 0,30\%$$

Relatívna chyba na výstupe:

$$\frac{|f(x) - f(x^*)|}{|f(x)|} = \frac{|f(\sqrt{2}) - f(1,41)|}{|f(\sqrt{2})|} = \frac{|9,94975\dots - 7,69230\dots|}{9,94975\dots} = 0,226884\dots \doteq 22,69\%$$

Číslo podmienenosti:

$$c_p = \frac{\text{relatívna chyba na výstupe}}{\text{relatívna chyba na vstupe}} = \frac{0,226884\dots}{0,002979\dots} = 76,14996\dots \doteq 76,15,$$

t. j. $\delta(\text{výstup}) = 76,15 \cdot \delta(\text{vstup})$

V okolí bodu $x = \sqrt{2}$ je úloha **zle podmienená**. Pre $x^* = 1,4$ číslo podmienenosti $c_p \doteq 49,50$ dosahuje hodnotu takmer 50 a pre $x^* = 1,41$ je číslo podmienenosti ešte väčšie $c_p \doteq 76,15$. Malá chyba na vstupe spôsobuje veľkú chybu na výstupe. Graf funkcie je v okolí bodu $x = \sqrt{2}$ veľmi strmý (Obr. 1.10). Príčinou je veľmi malá hodnota menovateľa funkcie $f(x)$ v danej oblasti, keď sa v menovateli nachádza rozdiel dvoch blízkych čísel: pre $x = \sqrt{2}$:

$$10 - 7x = 10 - 7\sqrt{2} = 10 - 9,89949 = 0,100505\dots$$

b) *Podmienenosť úlohy v okolí bodu $x = 10$*

V okolí bodu $x = 10$ vyšetříme podmienenosť úlohy pre $x^* \doteq 10,1$ a pre $x^* \doteq 10,01$

$$\boxed{x^* \doteq 10,1}$$

Relatívna chyba na vstupe:

$$\delta(\text{vstup}) = \frac{|x - x^*|}{|x|} = \frac{|10 - 10,1|}{|10|} = 0,01 = 1\%$$

Relatívna chyba na výstupe:

$$\delta(\text{výstup}) = \frac{|f(x) - f(x^*)|}{|f(x)|} = \frac{|f(10) - f(10,1)|}{|f(10)|} = \frac{|-0,01666\dots + 0,01647\dots|}{|-0,01666\dots|} = 0,01153\dots \doteq 1,15\%$$

Číslo podmienenosti:

$$c_p = \frac{\text{relatívna chyba na výstupe}}{\text{relatívna chyba na vstupe}} = \frac{0,01153\dots}{0,01} = 1,153\dots \doteq 1,15$$

t. j. $\delta(\text{výstup}) \doteq 1,15 \cdot \delta(\text{vstup})$.

$$x^* = 10,01$$

Relatívna chyba na vstupe:

$$\delta(\text{vstup}) = \frac{|x - x^*|}{|x|} = \frac{|10 - 10,01|}{|10|} = 0,001 = 0,10\%$$

Relatívna chyba na výstupe:

$$\delta(\text{výstup}) = \frac{|f(x) - f(x^*)|}{|f(x)|} = \frac{|f(10) - f(10,01)|}{|f(10)|} = \frac{|-0,01666\dots + 0,0166472|}{|-0,01666\dots|} = 0,001165\dots \doteq 0,12\%$$

Číslo podmienenosti:

$$c_p = \frac{\text{relatívna chyba na výstupe}}{\text{relatívna chyba na vstupe}} = \frac{0,001165\dots}{0,001} = 1,165 \doteq 1,17$$

t. j. $\delta(\text{výstup}) \doteq 1,17 \cdot \delta(\text{vstup})$.

V okolí bodu $x = 10$ je úloha **dobře podmienená**. Pre $x^* = 10,1$ číslo podmienenosti dosahuje hodnotu $c_p \doteq 1,15$ a pre $x^* = 10,01$ je číslo podmienenosti $c_p \doteq 1,17$. Malá chyba na vstupe

spôsobuje malú chybu na výstupe. Graf funkcie je v okolí bodu $x = 10$ primerane strmý (Obr. 1.10).

PRÍKLAD č. 5

Daná je polynomickeá funkcia reálnej premennej x vyjadrená:

$$f(x) = x^3 + 0,345x^2 - 5,67x - 1,46 \text{ a jej vnorený (Hornerov) tvar}$$

$$g(x) = -1,46 + x(-5,67 + x(0,345 + x)).$$

1. Ukážte, že $f(x) = g(x)$.
2. Vypočítajte $f(4,56)$ a $g(4,56)$.
 - a) presne,
 - b) približne, s presnosťou na 3 platné číslice **zaokrúhlením**, t. j. všetky výsledky aj medzivýsledky zaokrúhlite na 3 platné číslice,
 - c) približne, s presnosťou na 3 platné číslice **orezaním**, t. j. všetky výsledky aj medzivýsledky orežete na 3 platné číslice.

Pre približné hodnoty $f^*(4,56)$ a $g^*(4,56)$ vypočítajte absolútnu chybu, relatívnu chybu a počet platných číslic.

Ktorý tvar polynomickej funkcie je vhodnejší na výpočet hodnoty funkcie v istom bode? Zdôvodnite (Všimajte si počet a druh aritmetických operácií).

RIEŠENIE

1. *Rovnosť funkcií* $f(x) = g(x)$: Dve funkcie sa rovnajú, ak sa rovnajú ich definičné obory a ak pre každé x z definičného oboru je $f(x) = g(x)$.

$$D_f = R = D_g. \text{ Definičné obory sa rovnajú.}$$

Upravme:

$$\begin{aligned} g(x) &= -1,46 + x(-5,67 + x(0,345 + x)) = 1,46 + x(-5,67 + 0,345x + x^2) \\ &= -1,46 - 5,67x + 0,345x^2 + x^3 = f(x) \end{aligned}$$

Boli použité ekvivalentné úpravy, $f(x) = g(x)$ pre všetky $x \in R$.

2. Na výpočet hodnôt $f(4,56)$ použijeme nasledujúcu tabuľku (Tabuľka 1.5):

hodnota	x	x^2	$x \cdot x^2 = x^3$	$0,345x^2$	$-5,67x$
presná	4,56	20,7936	94,818816	7,173792	-25,8552
zaokrúhlená na 3 platné číslice	4,56	20,8	$94,848 \doteq 94,8$	$7,176 \doteq 7,18$	$-25,8552 \doteq 25,9$
orezaná na 3 platné číslice	4,56	20,7	$94,392 \doteq 94,3$	$7,1415 \doteq 7,14$	$-25,8552 \doteq -25,8$

Tabuľka 1.5

a) *presný výpočet:*

$$f(4,56) = (4,56)^3 + 0,345 \cdot (4,56)^2 - 5,67 \cdot 4,56 - 1,46 = 94,818816 + 7,173792 - 25,8552 = 74,677408$$

b) *približný výpočet so zaokrúhlením:*

$$f_z^*(4,56) = (4,56)^3 + 0,345 \cdot (4,56)^2 - 5,67 \cdot 4,56 - 1,46 = 94,8 + 7,18 - 25,9 - 1,46 = 102_{(101,98)} - 25,9 - 1,46 = 76,1 - 1,46 = 74,6_{(74,64)}$$

– absolútna chyba: $\varepsilon_z(4,56) = |f(4,56) - f_z^*(4,56)| = |74,677408 - 74,6| = 0,077408$

– relatívna chyba: $\frac{\varepsilon_z(4,56)}{f(4,56)} = \frac{0,077408}{74,677408} = 0,00103656... \doteq 0,1\%$

– počet platných čísiel:

$f(4,56)$ vyjadríme ako číslo s normalizovanou mantisou: $f(4,56) = 0,74677408 \cdot 10^{-2}$

$f_z^*(4,56)$ vyjadríme ako číslo s rovnakým exponentom: $f_z^*(4,56) = 0,746 \cdot 10^{-2}$

počet platných čísiel: (podľa definície Definícia 1.1):

$$|0,74677408 - 0,746| = 0,00077408 \leq 0,005 \leq 0,5 \cdot 10^{-2} \Rightarrow f_z^*(4,56) \text{ má 2 platné čísiel}$$

(podľa poznámky: v $|0,74677408 - 0,746| = 0,00077408$ sa za desatinnou čiarkou nachádzajú 2 nuly, za nimi je číslica 0 (jedna z čísiel 0 až 5), t. j. $f_z^*(4,56)$ má 2 platné čísiel).

c) *približný výpočet s orezaním:*

$$f_o^*(4,56) = (4,56)^3 + 0,345 \cdot (4,56)^2 - 5,67 \cdot 4,56 - 1,46 = 94,3 + 7,14 - 25,8 - 1,46 = 101_{(101,44)} - 25,8 - 1,46 = 75,2 - 1,46 = 73,7_{(74,74)}$$

absolútna chyba: $\varepsilon_o(4,56) = |f(4,56) - f_o^*(4,56)| = |74,677408 - 73,7| = 0,977408$

relatívna chyba: $\frac{\varepsilon_o(4,56)}{f(4,56)} = \frac{0,977408}{74,677408} = 0,0130884... \doteq 1,3\%$

počet platných čísiel:

- $f(4,56)$ vyjadríme ako číslo s normalizovanou mantisou: $f(4,56) = 0,74677408 \cdot 10^{-2}$

- $f_o^*(4,56)$ vyjadríme ako číslo s rovnakým exponentom: $f_o^*(4,56) = 0,737 \cdot 10^{-2}$

- počet platných čísiel: (podľa definície Definícia 1.1):

$$|0,74677408 - 0,737| = 0,00977408 \leq 0,05 \leq 0,5 \cdot 10^{-1} \Rightarrow f_o^*(4,56) \text{ má 1 platnú číslicu}$$

(podľa poznámky: v $|0,74677408 - 0,737| = 0,00977408$ sa za desatinnou čiarkou nachádza 1 nula, za ňou je číslica 0 (jedna z čísiel 0 až 5), t. j. $f_o^*(4,56)$ má 1 platnú číslicu).

2. Na výpočet hodnôt $g(4,56)$ použijeme nasledujúcu tabuľku (Tabuľka 1.6):

$$g(x) = -1,46 + x(-5,67 + x(0,345 + x)).$$

hodnota	x	$b_1 = 0,345 + x$	$b_2 = x \cdot b_1$	$b_3 = -5,67 + b_2$	$b_4 = x \cdot b_3$
presná	4,56	4,905	22,3668	16,6868	76,137408
zaokrúhlená na 3 platné číslice	4,56	4,91	$22,3896 \doteq 22,4$	$16,73 \doteq 16,7$	$76,152 \doteq 76,2$
orezaná na 3 platné číslice	4,56	4,90	$22,344 \doteq 22,3$	$16,63 \doteq 16,6$	$75,696 \doteq 75,6$

Tabuľka 1.6

$$g(x) = -1,46 + x(-5,67 + x(0,345 + x)) = -1,46 + b_4$$

a) *presný výpočet:*

$$g(4,56) = -1,46 + 76,137408 = 74,677408 = f(4,56)$$

b) *približný výpočet so zaokrúhlením:*

$$g_z^*(4,56) = -1,46 + 76,2 = 74,7_{(74,74)}$$

$$\text{absolútna chyba: } \varepsilon_{g_z}(4,56) = |g(4,56) - g_z^*(4,56)| = |74,677408 - 74,7| = 0,022592$$

$$\text{relatívna chyba: } \frac{\varepsilon_{g_z}(4,56)}{g(4,56)} = \frac{0,022592}{74,677408} = 0,000302527\dots \doteq 0,03\%$$

počet platných číslic:

- $g(4,56)$ vyjadríme ako číslo s normalizovanou mantisou: $g(4,56) = 0,74677408 \cdot 10^{-2}$
- $g_z^*(4,56)$ vyjadríme ako číslo s rovnakým exponentom: $g_z^*(4,56) = 0,747 \cdot 10^{-2}$
- počet platných číslic: (podľa definície Definícia 1.1):
 $|0,74677408 - 0,747| = 0,00022592 \leq 0,0005 \leq 0,5 \cdot 10^{-3} \Rightarrow g_z^*(4,56)$ má 3 platné číslice
 (podľa poznámky: v $|0,74677408 - 0,747| = 0,00022592$ sa za desatinnou čiarkou nachádzajú 3 nuly, za nimi je číslica 2 (jedna z číslic 0 až 5), t. j. $g_z^*(4,56)$ má 3 platné číslice).

c) *približný výpočet s orezaním:*

$$g(4,56) = -1,46 + 75,6 = 74,1_{(74,14)}$$

$$\text{absolútna chyba: } \varepsilon_{g_o}(4,56) = |g(4,56) - g_o^*(4,56)| = |74,677408 - 74,1| = 0,577408$$

$$\text{relatívna chyba: } \frac{\varepsilon_{g_o}(4,56)}{g(4,56)} = \frac{0,577408}{74,677408} = 0,00773203\dots \doteq 0,7\%$$

počet platných číslic:

- $g(4,56)$ vyjadríme ako číslo s normalizovanou mantisou: $g(4,56) = 0,74677408 \cdot 10^{-2}$
- $g_o^*(4,56)$ vyjadríme ako číslo s rovnakým exponentom: $g_o^*(4,56) = 0,741 \cdot 10^{-2}$

- počet platných číslíc: (podľa definície Definícia 1.1):

$$|0,74677408 - 0,741| = 0,00577408 \leq 0,05 \leq 0,5 \cdot 10^{-1} \Rightarrow g_o^*(4,56) \text{ má 1 platnú číslicu}$$

(podľa poznámky: v $|0,74677408 - 0,741| = 0,00577408$ sa za desatinnou čiarkou nachádzajú 1 nula, za ňou je číslica 0 (jedna z číslíc 0 až 5), t. j. $g_o^*(4,56)$ má 1 platnú číslicu).

Výpočet	$f(4,56)$ $g(4,56)$		absolútna chyba		relatívna chyba		počet plat. číslíc	
	$f^*(4,56)$	$g^*(4,56)$	$f^*(4,56)$	$g^*(4,56)$	$f^*(4,56)$	$g^*(4,56)$	$f^*(4,56)$	$g^*(4,56)$
presne	74,677408		$f^*(4,56)$	$g^*(4,56)$	$f^*(4,56)$	$g^*(4,56)$	$f^*(4,56)$	$g^*(4,56)$
zaokr. na 3 platné číslice	74,6	74,7	0,077408	0,022592	0,1%	0,03%	2	3
orezaný na 3 platné číslice	73,7	74,1	0,977408	0,577408	1,3%	0,7%	1	1

Tabuľka 1.7. Prehľad výsledkov – príklad č. 5

Z porovnania výsledkov (Tabuľka 1.7) vyplýva, že na výpočet hodnoty polynomickej funkcie v istom bode je z numerického hľadiska **výhodnejší vnorený tvar polynómu**. Pri počítaní s platnosťou na 3 platné číslice vnorený tvar polynómu dosiahol tak pri zaokrúhlení ako aj orezaní menšiu absolútnu chybu, menšiu relatívnu chybu a menší najvyšší počet platných číslíc.

Zdôvodnenie. Základný tvar $f(x) = x^3 + 0,345x^2 - 5,67x - 1,46$ vyžaduje 3 sčítania a 5 násobení, zatiaľ čo vnorený tvar polynómu $g(x) = -1,46 + x(-5,67 + x(0,345 + x))$ vyžaduje iba 3 sčítania a 2 násobenia, šírenie zaokrúhľovacích chýb je tak oveľa menšie.

ⁱ <http://www.sciencemuseum.org.uk/>

2 NUMERICKÉ RIEŠENIE NELINEÁRNYCH ROVNÍC

V strojárskkej praxi sa často stretávame s úlohami, riešením ktorých je nájsť maximálny priehyb prizmatického nosníka, kritickú vzpernú silu, maximálnu rýchlosť na hranici prevrátenia sa automobilu atď.

2.1 ZÁKLADNÉ POJMY

Matematická formulácia vyššie uvedených problémov znie:

Nájdite riešenie rovnice

$$f(x) = 0, \quad (2.1)$$

kde f je reálna funkcia reálnej premennej x .

Riešiť rovnicu (2.1) znamená nájsť všetky také čísla α , pre ktoré platí $f(\alpha) = 0$.

Číslo α nazývame **koreňom** alebo **riešením** rovnice (2.1).

Ak v rovnici (2.1)

1. $f(x) = P_n(x)$, potom rovnicu (2.1) nazývame **algebraickou rovnicou stupňa n** ,
2. $f(x) \neq P_n(x)$, potom hovoríme o **transcendentnej** rovnici,

kde $P_n(x) = a_n x^n + a_{n-1} x^{n-1} + a_{n-2} x^{n-2} + \dots + a_1 x^1 + a_0$ je polynóm n -tého stupňa,

$n \in \mathbb{N}$ je stupeň polynómu,

$a_i \in \mathbb{R}$ sú koeficienty polynómu, pričom $a_0 \neq 0$, $i = 0, 1, 2, \dots, n$.

Pre algebraické rovnice stupňa $n < 5$ a niektoré transcendentné rovnice poznáme algoritmy presného riešenia, t. j. vzorce a postupy vedúce po konečnom počte krokov k určeniu presnej hodnoty koreňa α (Tabuľka 2.1).

stupeň	typ	tvar	riešenie	poznámka
$n = 1$	lineárna	$ax + b = 0$	$x = -\frac{b}{a}$	$a, b \in \mathbb{R}$, $a \neq 0$
$n = 2$	kvadratická	$ax^2 + bx + c = 0$	$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$	$a, b, c \in \mathbb{R}$, $a \neq 0$
-	transcendentná	$a^x = b$	$x = \log_a b$	$a, b \in \mathbb{R}^+$, $a \neq 1$
		$\log_a(x - c) = b$	$x = a^b + c$	$a, b, c \in \mathbb{R}$, $x > c$, $a > 0$, $a \neq 1$
		$\cos x = a$	$x = \pm \arccos a$	$a \in \langle -1, 1 \rangle$

Tabuľka 2.1. Riešenia niektorých typov rovníc

Vzorce pre presné riešenie algebrických rovníc 3. a 4. stupňa sa pre ich zložitost' používajú zriedkavo, majú skôr teoretický význam. Korene (presné riešenie) algebrických rovníc stupňa $n > 4$ alebo zložitejších transcendentných rovníc nemožno vo všeobecnosti získať konečným počtom aritmetických operácií a odmocnín, na ich určenie sa používajú **numerické metódy**.

Bližšie sa budeme venovať numerickým metódam riešenia rovníc typu $f(x) = 0$ na množine reálnych čísel.

2.2 NUMERICKÉ RIEŠENIE ROVNÍC TYPU $f(x) = 0$

Rovnica (2.1) môže mať v oblasti reálnych čísel žiadne, jedno, viac (konečný počet) alebo nekonečne veľa riešení. Numerickými metódami hľadáme vždy konečný počet riešení, t. j. rovnicu riešime na uzavretom intervale a každé riešenie rovnice hľadáme osobitne.

ÚLOHA

Nájdite riešenie rovnice $g(x) = h(x)$ na intervale $\langle a, b \rangle$, kde g a h sú reálne funkcie reálnej premennej x , ktorá je definovaná na množine M .

RIEŠENIE

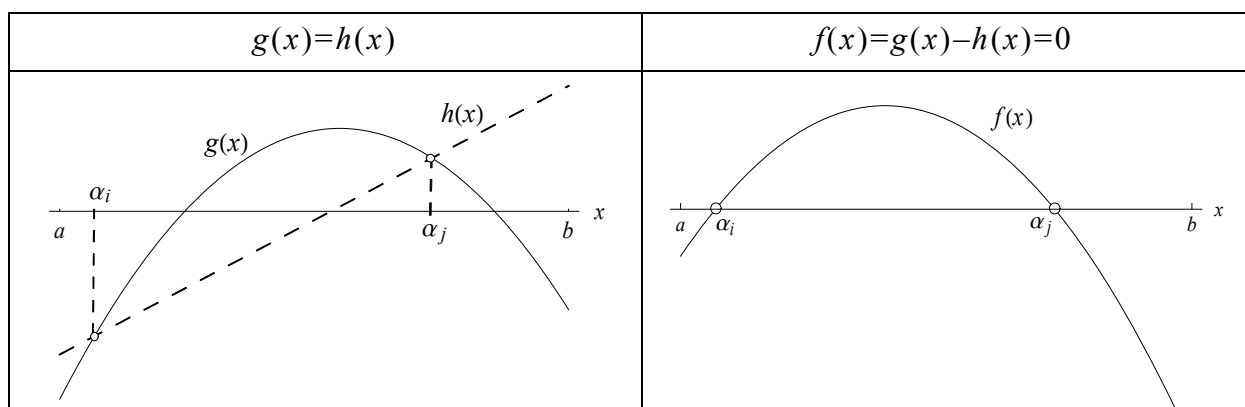
Rovnicu

$$g(x) = h(x) \tag{2.2}$$

upravíme na typový tvar (2.1):

$$f(x) = 0,$$

kde $f(x) = g(x) - h(x)$, alebo $f(x) = h(x) - g(x)$.



Obr. 2.1. Grafické riešenie rovnice $g(x) = h(x)$ a $f(x) = g(x) - h(x) = 0$

Ekvivalentnou úpravou na typový tvar $f(x) = 0$ získame možnosť využiť pri riešení vlastnosti funkcie $f(x)$. Koreňmi rovnice (2.1) a teda aj zadanej rovnice (2.2) sú potom nulové body $f(x)$, t. j. body, v ktorých funkcia $f(x)$ nadobúda nulové hodnoty. K určaniu numerického

riešenia rovnice $f(x)=0$ budeme potrebovať informácie o počte koreňov rovnice na danom intervale, o ich polohe a o priebehu funkcie $f(x)$ v okolí hľadaného koreňa. V prípade, že proces nie je automatizovaný a riešenie hľadá človek, žiadané informácie najľahšie získa z grafu funkcie. Na Obr. 2.1 je zobrazené grafické riešenie rovnice. Pri hľadaní koreňov rovnice $g(x)=h(x)$ určíme x -ové súradnice priesečníkov grafov funkcií, pri hľadaní koreňov rovnice $f(x)=0$ určíme x -ové súradnice priesečníkov grafu funkcie $f(x)$ s osou x . Na danom obrázku rovnica nadobúda na intervale $\langle a, b \rangle$ dva korene α_i a α_j . Pri počítačovom spracovaní na lokalizáciu reálnych koreňov (2.1) spravidla využívame tabeláciu $f(x)$ (pozri odsek 2.2.1 b)) a počet koreňov vyhodnotíme ako počet zmien znamienka hodnôt $f(x_i)$.

Numerické riešenie rovnice (2.1) pozostáva:

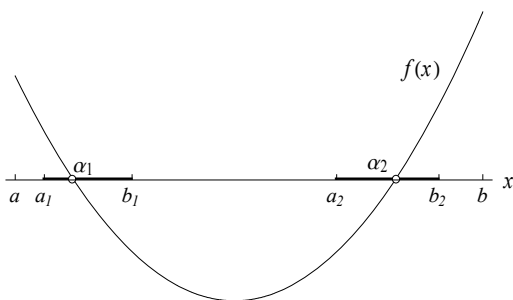
1. zo separácie koreňov rovnice,
2. z výpočtu koreňov rovnice numerickou metódou,
3. z odhadu absolútnej chyby jednotlivých koreňov.

2.2.1 SEPARÁCIA KOREŇOV ROVNICE

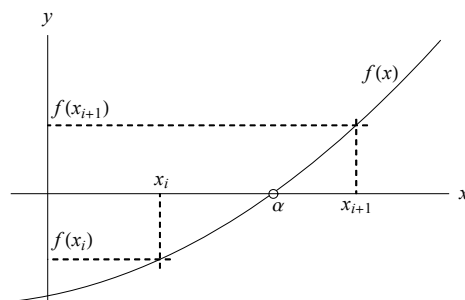
Pod separáciou koreňov rovnice (2.1) rozumieme určenie intervalov $\langle a_i, b_i \rangle$, v ktorých leží práve jeden jej koreň $\alpha_i, i = 1, 2, \dots$ (Obr. 2.2).

Separáciu koreňov danej rovnice vykonávame tromi spôsobmi:

- a) analýzou priebehu $f(x)$, čo je zvyčajne zdĺhavé a náročné na výpočet,
- b) tabeláciou $f(x)$. Interval $\langle a, b \rangle$ rozdelíme $(n+1)$ deliacimi bodmi $a = x_0 < x_1 < \dots < x_n = b$ na n čiastkových intervalov $\langle x_i, x_{i+1} \rangle, i = 0, 1, 2, \dots, n$, s dĺžkou $h = \frac{b-a}{n} = \text{konšt.}$, pričom skúmame hodnoty $f(x_i), i = 0, 1, 2, \dots, n$. Koreň $\alpha \in \langle x_i, x_{i+1} \rangle$ ak $f(x_i) \cdot f(x_{i+1}) < 0$ (Obr. 2.3).

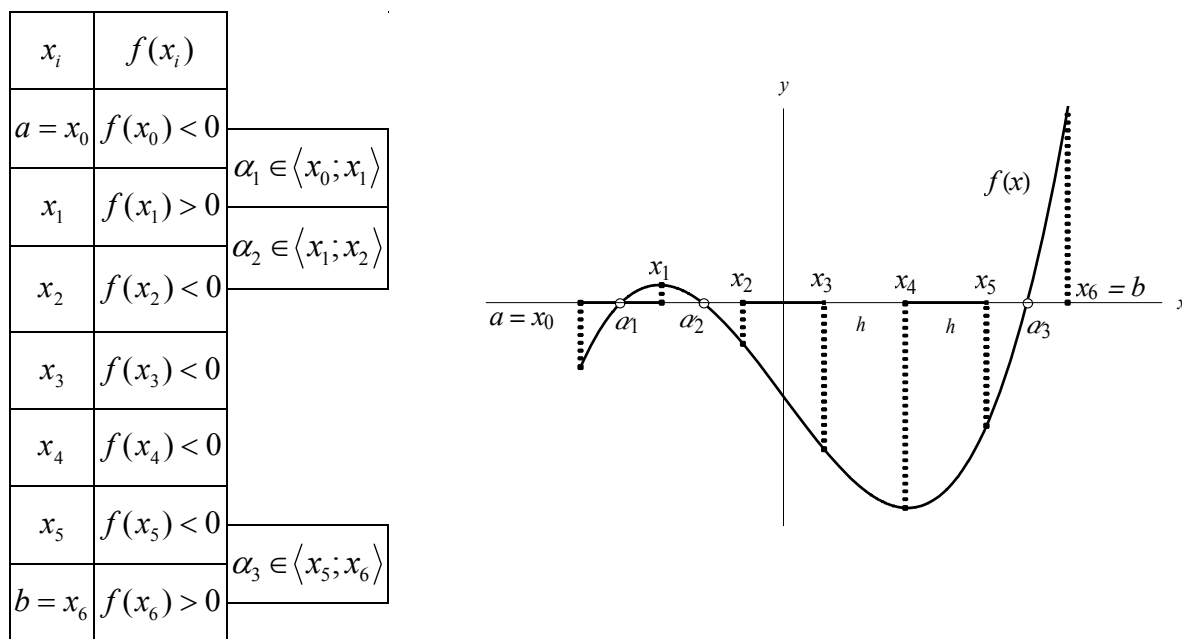


Obr. 2.2



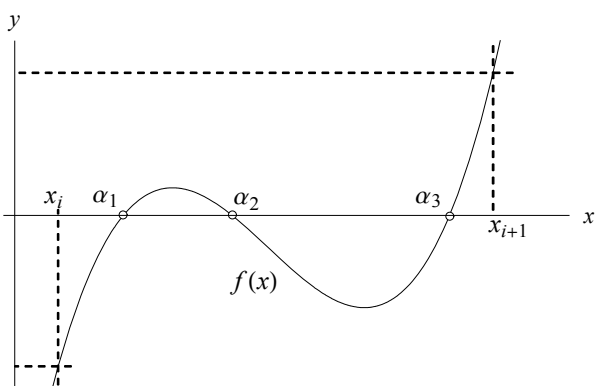
Obr. 2.3

Napríklad rozdelíme interval $\langle a, b \rangle$ na 6 čiastkových intervalov s konštantnou dĺžkou h a vykonajme tabeláciu funkcie $f(x)$ (Obr. 2.4). Zmena znamienka $f(x)$ nastáva medzi deliacimi bodmi x_0 a x_1 , x_1 a x_2 a medzi x_5 a x_6 . Znamená to, že v každom z intervalov $\langle x_0, x_1 \rangle$, $\langle x_1, x_2 \rangle$ a $\langle x_5, x_6 \rangle$ existuje minimálne jeden koreň α_1 , α_2 a α_3 .

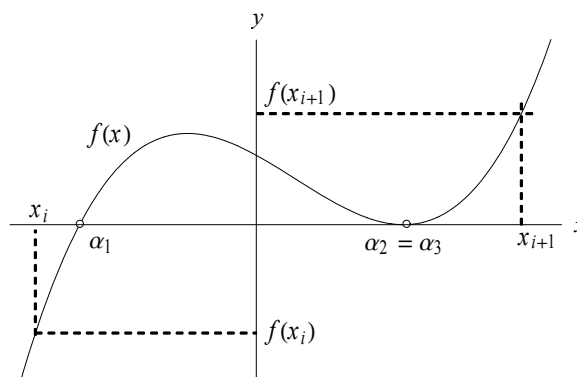


Obr. 2.4. Tabelácia funkcie $f(x)$ a jej graf

Nevýhodou tabelácie $f(x)$ je možná strata jednoduchých (Obr. 2.5) alebo viacnásobných (Obr. 2.6) koreňov, ku ktorej dochádza vtedy, keď je splnená podmienka $f(x_i) \cdot f(x_{i+1}) < 0$, avšak v intervale $\langle x_i, x_{i+1} \rangle$ leží viacero koreňov α_i .



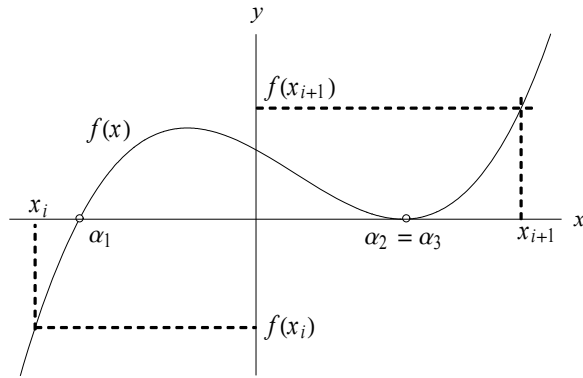
Obr. 2.5. Strata jednoduchých koreňov



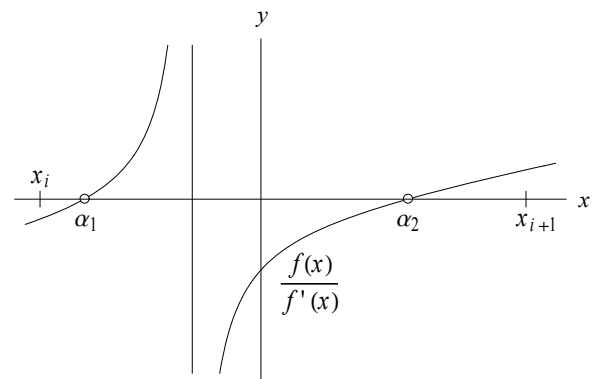
Obr. 2.6. Strata viacnásobného koreňa

V danej situácii nám pomôže znalosť intervalov monotónnosti prípadne ak máme podozrenie, že v intervale $\langle x_i, x_{i+1} \rangle$ existuje viacnásobný koreň, môžeme využiť nasledujúcu vetu:

Veta 2.1 Ak funkcia $f(x)$ má viacnásobné korene, potom funkcia $u(x) = \frac{f(x)}{f'(x)}$ má tie isté korene ako funkcia $f(x)$, ale každý koreň je jednoduchý (Obr. 2.7, Obr. 2.8).



Obr. 2.7. Jednoduchý koreň α_1 a dvojnásobný koreň $\alpha_2 = \alpha_3$ rovnice $f(x) = 0$



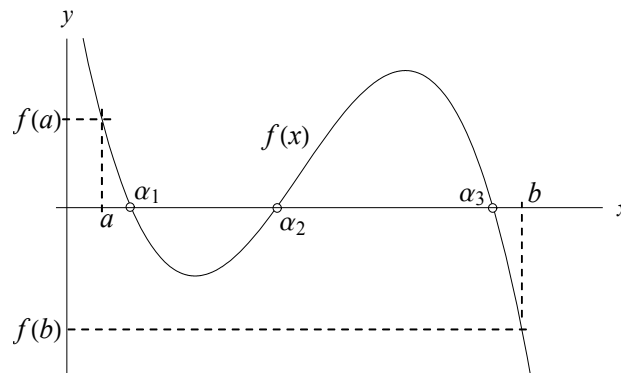
Obr. 2.8. Jednoduché korene α_1 a α_2 rovnice $\frac{f(x)}{f'(x)} = 0$

c) Graficky. Intervaly, v ktorých leží práve jeden koreň rovnice, určíme z grafu.

Pri separácii koreňov rovnice sú užitočné nasledujúce vety:

Veta 2.2 (Veta o existencii koreňa) Nech funkcia $f(x)$ je spojitá na $\langle a, b \rangle$ a platí $f(a) \cdot f(b) < 0$, potom existuje aspoň jedno $\alpha \in (a, b)$ také, že $f(\alpha) = 0$.

Veta nehovorí nič o počte koreňov v intervale $\langle a, b \rangle$. Jej platnosť dokazuje len existenciu koreňa v tomto intervale. Môže nastať prípad, že $f(a) \cdot f(b) < 0$, a v intervale $\langle a, b \rangle$ ležia napr. 3 korene rovnice (Obr. 2.9). Z grafu je zrejmé, že koreň rovnice bude jediný, ak na (a, b) bude funkcia rýdzomonotónna.



Obr. 2.9

Veta 2.3 (Veta o jednoznačnosti koreňa) Nech na $\langle a, b \rangle$ existuje koreň rovnice $f(x)=0$. Nech funkcia $f(x)$ je na $\langle a, b \rangle$ rýdzomonotónna, t. j. nech existuje derivácia $f'(x)$, ktorá na tomto intervale nemení znamienko, čiže:

- a) $\forall x \in \langle a, b \rangle: f'(x) > 0; f(x)$ je rastúca na $\langle a, b \rangle$ alebo
- b) $\forall x \in \langle a, b \rangle: f'(x) < 0; f(x)$ je klesajúca na $\langle a, b \rangle$.

Potom **existuje** práve jedno $\alpha \in (a, b)$ také, že $f(\alpha) = 0$.

ZHRNUTIE

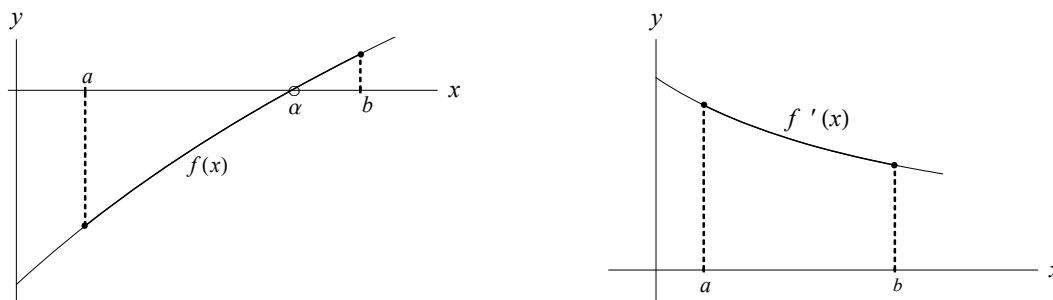
Nech:

- a) $f(x)$ je spojitá na $\langle a, b \rangle$,
- b) $f(a) \cdot f(b) < 0$,
- c) $\forall x \in \langle a, b \rangle: f'(x) > 0$ alebo $f'(x) < 0$,

potom na intervale $\langle a, b \rangle$ leží **práve jeden** koreň rovnice $f(x)=0$.

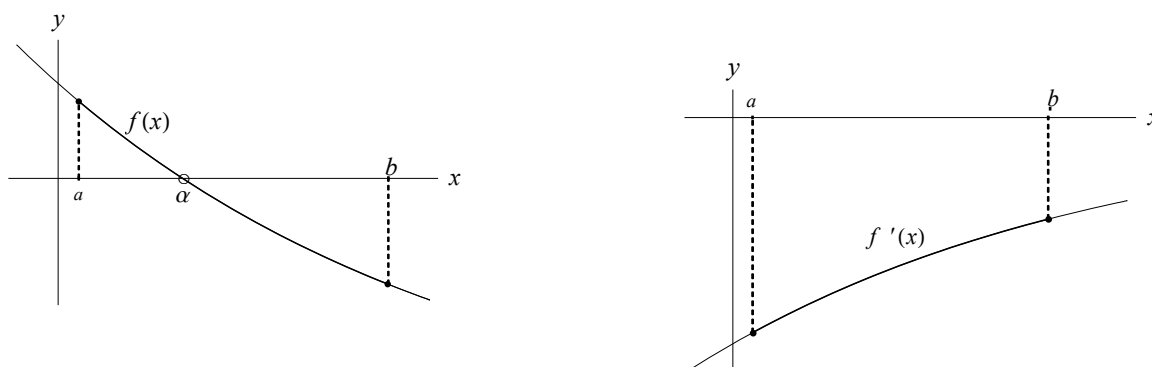
Všimnite si!

Ak funkcia $f(x)$ je na $\langle a, b \rangle$ rastúca, potom $f'(x) > 0$ a graf funkcie $f'(x)$ leží nad osou x .



Obr. 2.10. Graf rastúcej funkcie a jej derivácie

Ak funkcia $f(x)$ je na $\langle a, b \rangle$ klesajúca, potom $f'(x) < 0$ a graf funkcie $f'(x)$ leží pod osou x .



Obr. 2.11. Graf klesajúcej funkcie a jej derivácie

2.2.2 VÝPOČET KOREŇA ROVNICE

Na výpočet koreňa rovnice $f(x) = 0$ použijeme numerické metódy. Je známych viacero numerických metód, my sa bližšie budeme venovať dvom, ktoré spočívajú v jednoduchých, počítačom realizovateľných postupoch.

2.2.2.1 Numerické metódy – základné pojmy

Predpokladáme, že na intervale $\langle a, b \rangle$ leží práve jeden koreň rovnice (2.1). Pod numerickou metódou riešenia rovnice $f(x)=0$ rozumieme algoritmus postupného výpočtu postupnosti aproximácií $x_0, x_1, \dots, x_i, \dots$ koreňa rovnice α . Základom je iteračný, rekurzívny predpis, ktorý pri výpočte nasledovníka x_{i+1} využíva n predchodcov x_i, x_{i-1}, \dots, x_0 . Postupnosť aproximácií koreňa zostrojíme podľa iteračnej formuly:

$$x_{i+1} = q(x_{i-(n-1)}, x_{i-(n-2)}, \dots, x_i), \quad n = 1, 2, \dots, \quad i = n-1, n, n+1, \dots \quad (2.3)$$

$$i = n-1 \Rightarrow x_{i+1} = x_n = q(x_0, x_1, x_2, \dots, x_{n-1})$$

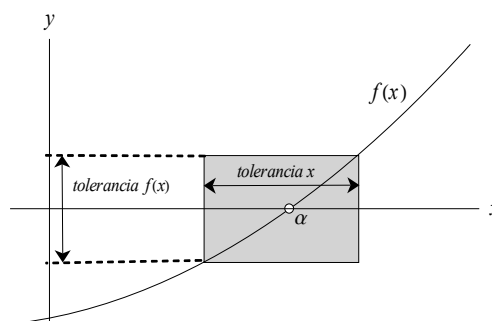
$$i = n \Rightarrow x_{i+1} = x_{n+1} = q(x_1, x_2, x_3, \dots, x_n)$$

$$i = n+1 \Rightarrow x_{i+1} = x_{n+2} = q(x_2, x_3, x_4, \dots, x_{n+1}) \text{ atď.},$$

kde

- q je iteračná funkcia,
- členy postupnosti $\{x_0, x_1, x_2, \dots, x_k, \dots\}$ nazývame postupné aproximácie alebo iterácie koreňa α ,
- výpočet postupnosti aproximácií nazývame iteračný proces,
- iteračný proces zapísaný iteračnou formulou (2.3) nazývame n -krokový iteračný proces (vopred treba poznať n hodnôt x_0, x_1, \dots, x_{n-1}) a x_0, x_1, \dots, x_{n-1} sú začiatkové (štartovacie) hodnoty iteračného procesu.
- numerická metóda je stacionárna, ak na výpočet všetkých iterácií x_{i+1} používame tú istú iteračnú funkciu q .

Ak iteračný proces (2.3) konverguje, tak konverguje k presnému riešeniu $f(x)=0$. Numerickou metódou nehľadáme presné riešenie. Jej cieľom je nájsť „dostatočne presné“ riešenie, ktoré sa od presného riešenia líši najviac o vopred stanovenú hodnotu – tolerancia, presnosť (Obr. 2.12).



Obr. 2.12

Postup výpočtu približnej hodnoty koreňa α spočíva v 3 základných krokoch:

1. definovanie štartovacích hodnôt x_0, x_1, \dots, x_{n-1} ,

2. postupný výpočet jednotlivých aproximácií $x_n, x_{n+1}, x_{n+2}, \dots$, koreňa α rovnice (2.1) podľa iteračnej formuly (2.3),
3. ukončenie iteračného procesu.

V prípade, že iteračný proces konverguje, výpočet sa ukončuje:

a) tzv. **zastavovacou podmienkou**. Výpočet aproximácií x_n, x_{n+1}, \dots, x_k , sa ukončí, keď je splnená jedna z nasledovných testovacích podmienok:

i) $|f(x_k)| < \varepsilon$, kde $\varepsilon > 0$ je vopred stanovená dostatočne malá hodnota, (napr. 10^{-4} , 10^{-10}). Nazývame ju aj **požadovaná chyba (tolerancia)** koreňa α vzhľadom na $|f(x_k)|$. Tiež hovoríme, že presné riešenie α rovnice (2.1) je určené aproximáciou x_k s toleranciou ε , $x_k \doteq \alpha$. Toleranciu vzhľadom na $|f(x_k)|$ volíme vtedy, ak funkcia je v okolí hľadaného koreňa dostatočne strmá, t. j. interval hodnôt x_k , pre ktoré $|f(x_k)| < \varepsilon$ nie je príliš dlhý.

ii) $|x_k - x_{k-1}| < \varepsilon$, kde $\varepsilon > 0$ je vopred stanovená tolerancia vzhľadom na vzdialenosť dvoch po sebe idúcich iterácií; $x_k \doteq \alpha$. Túto toleranciu volíme vtedy, keď vzdialenosť dvoch po sebe idúcich iterácií sa dostatočne rýchlo skracuje.

iii) $\varepsilon(x_k) < \varepsilon$, kde $\varepsilon(x_k)$ je odhad absolútnej chyby približnej hodnoty koreňa x_k (pozri Veta 2.4). Toleranciu vzhľadom na odhad absolútnej chyby volíme vtedy, ak požadujeme výpočet koreňa s vopred určenou maximálnou chybou, s danou **presnosťou** ε .

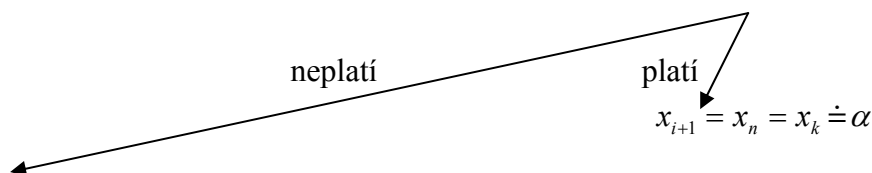
b) vopred stanoveným, dostatočne veľkým, počtom iterácií $pit = k - (n - 1)$, pričom x_{n-1} je posledný štartovací bod.

Ostatnú vypočítanú iteráciu x_k potom prehlásime za numerické riešenie koreňa rovnice (2.1), napr. $k = 50$, $x_{50} \doteq \alpha$.

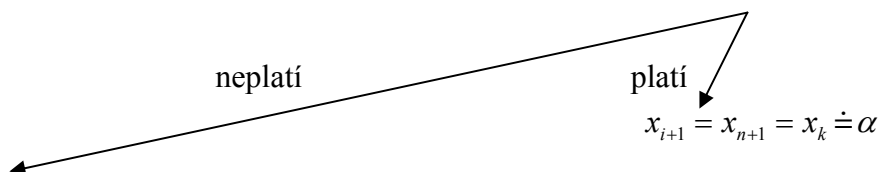
V prípade, že iteračný proces diverguje, výpočet sa ukončuje bez určenia α .

Postup výpočtu jednotlivých aproximácií x_n, x_{n+1}, x_{n+2} atď., koreňa α :

$i = (n-1)$, $x_{i+1} = x_n = q(x_0, x_1, x_2, \dots, x_{n-1})$ a zisťujeme, či platí zastavovacia podmienka



$i = n$, $x_{i+1} = x_{n+1} = q(x_1, x_2, x_3, \dots, x_n)$ a zisťujeme, či platí zastavovacia podmienka



$i = (n+1)$, $x_{i+1} = x_{n+2} = q(x_2, x_3, x_4, \dots, x_{n+1})$ a zisťujeme, či platí zastavovacia podmienka atď.

Poznámka. Pri počítačovom spracovaní z bezpečnostných dôvodov odporúčame pracovať s konečným, vopred určeným počtom iterácií. Pri korektnom zadaní sa tak výpočet vždy ukončí a to buď zastavovacou podmienkou alebo počtom iterácií. Pri stanovovaní zastavovacej podmienky je vhodné si uvedomiť, že ak by sme ako zastavovaciu podmienku zvolili $f(x_k)=0$ alebo $x_k = x_{k+1}$, dostali by sme presné riešenie $x_k = \alpha$. Avšak, vzhľadom na fakt, že výpočty sa vždy realizujú s konečnou presnosťou a sú zaťažované chybou (pozri odsek 1.2.4.3), testovacia podmienka daná rovnosťou je splnená iba ojedinele. Preto pri stanovovaní zastavovacej podmienky používame niektorú z vyššie uvedených možností i) až iii).

Algoritmus výpočtu x_k bez zastavovacej podmienky, s vopred stanoveným počtom iterácií pit :

1. určíme vhodné štartovacie body x_0, x_1, \dots, x_{n-1} ,
2. pre $i = (n-1)$ až $pit+(n-2)$ s krokom 1 počítame

$$x_{i+1} = q(x_{i-(n-1)}, x_{i-(n-2)}, \dots, x_i),$$
3. spolu so štartovacími hodnotami dostaneme postupnosť $\{x_0, x_1, \dots, x_n, \dots, x_k\}$,
 $k = pit + (n-1)$. Výsledkom je $x_k \doteq \alpha$.

Algoritmus výpočtu x_k so zastavovacou podmienkou:

1. určíme vhodné štartovacie body x_0, x_1, \dots, x_{n-1} , a zvolíme dostatočne veľký počet iterácií pit ,
2. pre $i = (n-1)$ až $pit+(n-2)$, s krokom 1 postupne počítame
 - a) $x_{i+1} = q(x_{i-(n-1)}, x_{i-(n-2)}, \dots, x_i)$,
 - b) testujeme platnosť danej zastavovacej podmienky
 $(|f(x_{i+1})| < \varepsilon \text{ alebo } |x_{i+1} - x_i| < \varepsilon \text{ alebo } \varepsilon(x_i) < \varepsilon)$:
 - Ak podmienka neplatí, pokračujeme výpočtom ďalšej iterácie.
 - Ak podmienka platí, výpočet je ukončený. Spolu so štartovacími hodnotami dostaneme postupnosť $\{x_0, x_1, \dots, x_n, \dots, x_k\}$, $k = i+1$.

Výsledkom, numerickým riešením je $x_k \doteq \alpha$.

V prípade, že iteračný výpočet nie je ukončený podmienkou, ale pevným počtom zvolených iterácií, počet iterácií zvýšime a pokračujeme vo výpočtoch, až kým nie je splnená zastavovacia podmienka.

Pri iteračných metódach, ktoré konvergujú k presnému riešeniu $f(x)=0$, nahradzujeme presné riešenie α niektorou aproximáciou x_k . V súvislosti s hľadaním numerického riešenia $x_k \doteq \alpha$ nás zaujímajú dve základné otázky:

1. Kedy je numerická metóda konvergentná?

Numerická metóda je konvergentná, ak konverguje postupnosť aproximácií koreňa $x_0, x_1, \dots, x_k, \dots$ vytvorená iteračnou formulou (2.3) k presnému riešeniu α rovnice (2.1), t. j.

$$\lim_{k \rightarrow \infty} x_k = \alpha. \quad (2.4)$$

Každú z aproximácií x_k , ktorá spĺňa (2.4), môžeme považovať za približnú hodnotu koreňa α .

Za **numerické riešenie rovnice** (2.1) budeme považovať takú aproximáciu x_k , zo vzťahu (2.4), ktorá sa od presného riešenia α líši o dostatočne malú hodnotu ε , $\varepsilon > 0$, $|x_k - \alpha| < \varepsilon$, resp. je splnená jedna z vyššie uvedených zastavovacích podmienok.

Rýchlosť konvergence numerickej metódy (tzv. rád metódy) vyjadruje $r \in \mathbb{R} \wedge r \geq 1$, pre ktoré platí

$$\lim_{k \rightarrow \infty} \frac{|x_{k+1} - \alpha|}{|x_k - \alpha|^r} = \lim_{k \rightarrow \infty} \frac{|\varepsilon_{k+1}|}{|\varepsilon_k|^r} = C, \quad C \neq 0,$$

kde C je chybová konštanta, ε_k je chyba k -tej iterácie, α je presná hodnota koreňa. Ak $r = 1$, potom $C < 1$. Rád konvergence metódy takto definujeme len pre konvergentné numerické metódy.

2. Akej chyby sa dopustíme pri nájdení numerického riešenia rovnice (2.1)?

Pri numerickom riešení rovníc hlavným zdrojom chýb sú:

- *chyba numerickej metódy*, kedy presné riešenie nahrádzame riešením približným (limitu nekonečnej postupnosti nahrádzame niektorým jej členom x_k , napr. členom s dostatočne veľkým indexom k),
- *zaokrúhľovacie chyby*, nepresnosti spôsobené výpočtom postupnosti aproximácií majú spravidla menší význam, preto ich zvyčajne možno zanedbať.

Odhad chyby k -tej aproximácie koreňa

Veta 2.4 Nech α je presná hodnota koreňa a x_k je približná hodnota koreňa rovnice

$f(x) = 0$. Nech pre $\forall x \in \langle a, b \rangle$ je $|f'(x)| \geq m > 0$, kde $m = \min_{x \in \langle a, b \rangle} |f'(x)|$, potom pre **odhad**

absolútnej chyby $\varepsilon(x_k)$ približnej hodnoty x_k koreňa α rovnice (2.1) platí:

$$|x_k - \alpha| \leq \varepsilon(x_k) \quad \wedge \quad \varepsilon(x_k) = \frac{|f(x_k)|}{m}. \quad (2.5)$$

Najčastejšie používané metódy v praxi:

1. metóda bisekcie – polenia intervalu,
2. Newtonova metóda – metóda dotyčníc,
3. metódy lineárnej interpolácie – metódy sečníc. Sem patrí napr. metóda Regula Falsi, Secant metóda.

Bližšie sa budeme venovať Newtonovej metóde a Secant metóde.

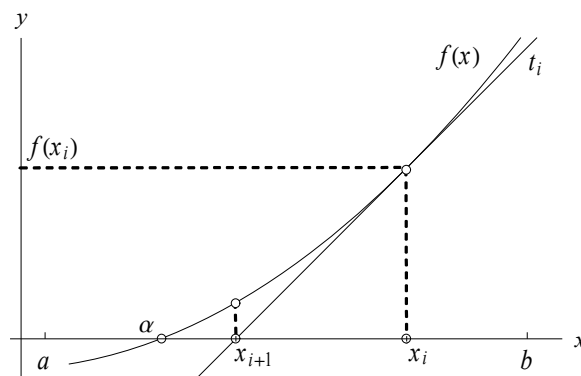
2.2.2.2 Newtonova metóda

Nazýva sa aj metóda dotyčníc. Jej podstata spočíva v tom, že v okolí hľadaného koreňa α nahradzujeme graf funkcie $y = f(x)$ jeho dotyčnicou. Predpokladá sa, že funkcia je na $\langle a, b \rangle$ spojitá a diferencovateľná, $f'(x) \neq 0$.

Iteračná formula má tvar

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}, \quad i=0,1,2,\dots, \quad (2.6)$$

čo je vlastne výpočet priesečníka x_{i+1} dotyčnice grafu $f(x)$ s osou x , vedenej bodom grafu funkcie $[x_i, f(x_i)]$ (overtu). (Obr. 2.13).



Obr. 2.13. Metóda dotyčníc (Newtonova metóda)

Metóda je jednokroková, v iteračnej formule vystupuje jeden predchodca x_i . Výpočet sa začína voľbou vhodného štartovacieho bodu x_0 a pokračuje výpočtom nasledovníkov x_{i+1} :

$$i = 0 \Rightarrow x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}; \quad \text{na výpočet } x_1 \text{ (prvá vypočítaná aproximácia) je potrebné poznať jednu (predchádzajúcu) hodnotu } x_0$$

$$i = 1 \Rightarrow x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}; \quad \text{na výpočet } x_2 \text{ je potrebné poznať jednu hodnotu } x_1$$

$$i = 2 \Rightarrow x_3 = x_2 - \frac{f(x_2)}{f'(x_2)}; \quad \text{na výpočet } x_3 \text{ je potrebné poznať jednu hodnotu } x_2$$

.....

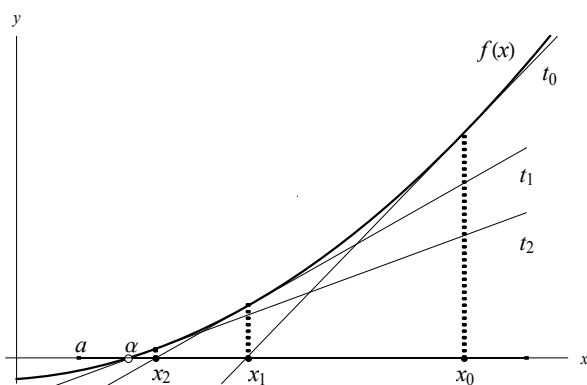
$$i = k-1 \Rightarrow x_k = x_{k-1} - \frac{f(x_{k-1})}{f'(x_{k-1})}; \quad \text{na výpočet } x_k \text{ je potrebné poznať jednu hodnotu } x_{k-1} \text{ atď}$$

Dostávame tak postupnosť iterácií $\{x_0, x_1, \dots, x_k, \dots\}$, ktorá pri zabezpečení podmienok konvergence metódy, konverguje k hľadanému koreňu α .

Ďalšie vlastnosti

- Ak metóda na danom intervale konverguje, štartovacím bodom môže byť ľubovoľný bod intervalu $\langle a, b \rangle$. Zvyčajne volíme krajný bod intervalu a alebo b , resp. bod ktorý leží čo možno najbližšie k hľadanému koreňu α .
- Metóda je stacionárna – každú iteráciu počítame podľa tej istej iteračnej funkcie (2.6).
- Rád konvergence Newtonovej metódy je $r = 2$. Pre jednoduchý koreň Newtonova metóda konverguje kvadraticky.
- Z hľadiska rýchlosti konvergence metódy je metódu vhodné použiť vtedy, keď je graf funkcie $f(x)$ v okolí koreňa strmý.
- Na výpočet odhadu chyby k -tej aproximácie koreňa existuje viacero vzorcov, my budeme používať vzťah (2.5).

Geometrická interpretácia metódy



Obr. 2.14

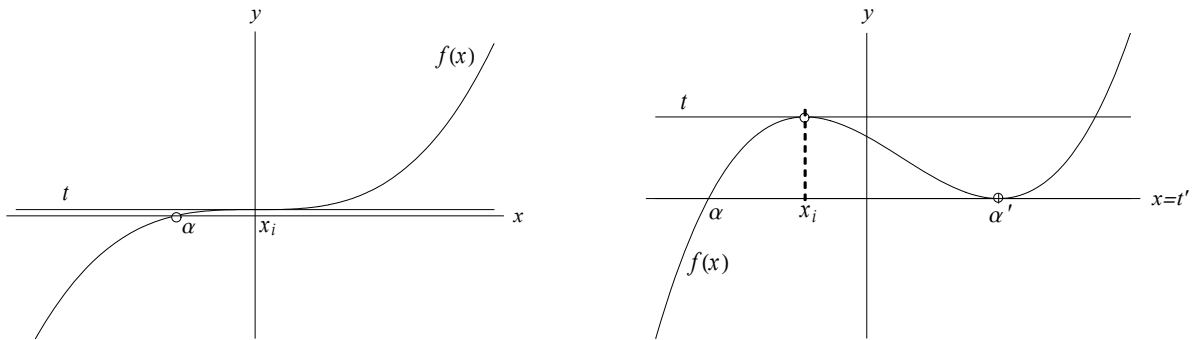
Zvolíme začiatočnú aproximáciu x_0 koreňa α . Bodom $[x_0, f(x_0)]$ vedieme dotyčnicu ku grafu funkcie $f(x)$. Jej priesečník s osou x označíme x_1 . Potom vedieme dotyčnicu bodom $[x_1, f(x_1)]$ a jej priesečník s osou x označíme x_2 atď. Dostávame postupnosť aproximácií koreňa $\{x_0, x_1, x_2, \dots\}$, ktorá, ak metóda konverguje, konverguje k hľadanému koreňu α .

„Úskalia“ metódy

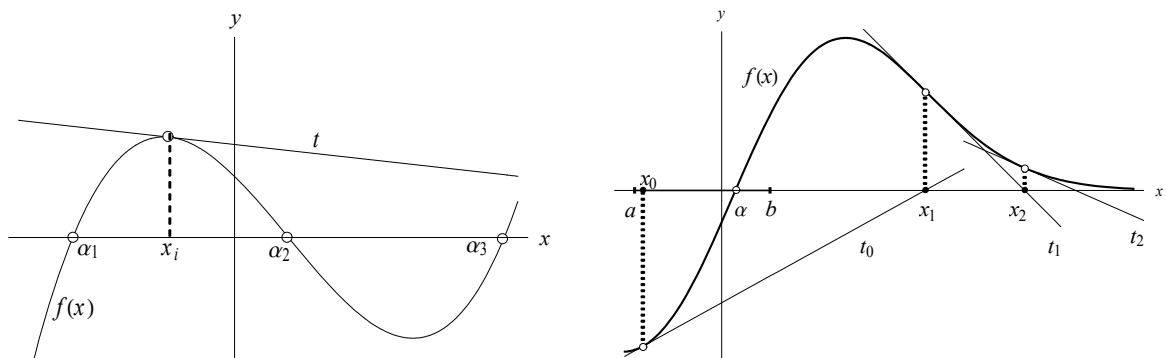
Newtonova metóda si vyžaduje v každom kroku okrem výpočtu $f(x_i)$ aj výpočet derivácie $f'(x_i)$ (pozri vzťah (2.6)). Problémy spôsobujú:

- tzv. kritické body x_i , kde $f'(x_i) = 0$. Sú to stacionárne body funkcie $f(x)$, pre ktoré vzťah (2.6) nie je definovaný. Dotyčnica v týchto bodoch je rovnobežná s osou x . Môžu nastať dva prípady: stacionárny bod nie je koreňom (2.1) a dotyčnica nemá s osou x žiadny vlastný priesečník (Obr. 2.15), alebo stacionárny bod je koreňom a dotyčnica má s osou x nekonečne veľa vlastných spoločných bodov, dotyčnica a os x splývajú (Obr. 2.17, koreň α' na Obr. 2.15).
- Nebezpečné sú nielen samotné stacionárne body $f(x)$, ale aj body v ich blízkom okolí. V takých prípadoch dotyčnica ku grafu funkcie $f(x)$ zvierá s osou x veľmi malý uhol. Jej priesečník s osou x leží ďaleko od hľadaného koreňa α_i , mimo intervalu $\langle a_i, b_i \rangle$ (pozri kap. 2.2.1). Ak rovnica má viac koreňov, metóda môže skonvergovať k inému koreňu ale-

bo môže divergovať. Nie nepodstatnou je tiež skutočnosť, že hodnota $|f'(x_i)|$ je v takýchto bodoch veľmi malá a bude spôsobovať veľkú zaokrúhľovaciú chybu vo výpočte (2.6). Z tohto dôvodu sa pokiaľ možno vyhýbame všetkým stacionárnym bodom funkcie a intervaly $\langle a_i, b_i \rangle$ volíme bez lokálnych extrémov (minimum, maximum), inflexných bodov a bodov, ktoré ležia v ich blízkom okolí.



Obr. 2.15. Dotyčnice v stacionárnym bodoch



Obr. 2.16. Dotyčnica v bode blízkom stacionárnemu bodu

Typický príklad, kde Newtonova metóda zlyhá, je riešenie kubických rovníc typu $x^3 = 0$ (Obr. 2.17).

Riešenie:

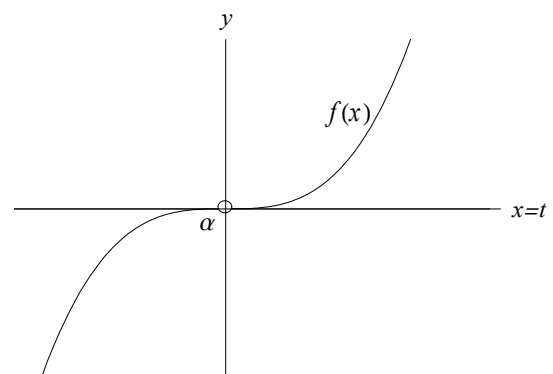
Funkcia $f(x) = x^3$

Stacionárne body funkcie:

$$f'(x) = 3x^2; f'(x) = 0 \Rightarrow x = 0$$

Inflexné body funkcie:

$$f''(x) = 6x; f''(x) = 0 \Rightarrow x = 0$$



Obr. 2.17

Dotyčnica t ku grafu funkcie $f(x)$ vedená bodom $[0, 0]$ je totožná s osou x a má s ňou nekonečne veľa priesečníkov.

Voľba štarovacej hodnoty x_0 kľúčovým spôsobom ovplyvňuje konvergenciu Newtonovej metódy a to tak, že metóda môže:

- konvergovať k hľadanému koreňu rovnice,
- konvergovať k inému koreňu rovnice,
- divergovať (nekonverguje k žiadnemu koreňu rovnice).

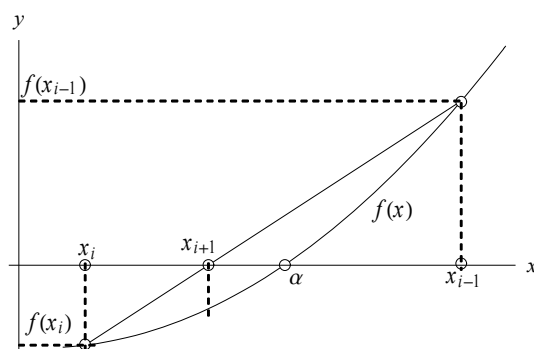
2.2.2.3 Secant metóda

Nazýva sa aj metóda sečníc (tetív). Patrí do skupiny metód lineárnej interpolácie (Secant, Regula Falsi), ktoré využívajú v okolí hľadaného koreňa α aproximáciu grafu funkcie $y = f(x)$ priamkou, jej sečnicou. Predpokladá sa, že funkcia $f(x)$ je na $\langle a, b \rangle$ spojitá a neobsahuje body s rovnakou funkčnou hodnotou.

Iteračná formula má tvar

$$x_{i+1} = x_i - f(x_i) \frac{x_{i-1} - x_i}{f(x_{i-1}) - f(x_i)}, \quad i = 1, 2, \dots, \quad (2.7)$$

čo je vlastne výpočet priesečníka x_{i+1} sečnice grafu funkcie $f(x)$ s osou x , prechádzajúcej bodmi grafu funkcie $[x_i, f(x_i)]$ a $[x_{i-1}, f(x_{i-1})]$ (overte). (Obr. 2.18).



Obr. 2.18. Metóda sečníc (Secant metóda)

Metóda je dvojkroková. V iteračnej formule vystupujú dva predchádzajúce body x_i a x_{i-1} , preto výpočet začína voľbou dvoch vhodných štarovacích bodov x_0 a x_1 a pokračuje výpočtom nasledovníkov x_{i+1} :

$$i = 0 \Rightarrow x_2 = x_1 - f(x_1) \frac{x_0 - x_1}{f(x_0) - f(x_1)} ; \text{ na výpočet } x_2 \text{ (prvá vypočítaná aproximácia) je potrebné poznať 2 (predchádzajúce) hodnoty } x_1, x_0,$$

$$i = 1 \Rightarrow x_3 = x_2 - f(x_2) \frac{x_1 - x_2}{f(x_1) - f(x_2)} ; \text{ na výpočet } x_3 \text{ je potrebné poznať 2 hodnoty } x_2, x_1,$$

$$i = 2 \Rightarrow x_4 = x_3 - f(x_3) \frac{x_2 - x_3}{f(x_2) - f(x_3)} ; \text{ na výpočet } x_4 \text{ je potrebné poznať 2 hodnoty } x_3, x_2,$$

...

$i = k-1 \Rightarrow x_k = x_{k-1} - f(x_{k-1}) \frac{x_{k-2} - x_{k-1}}{f(x_{k-2}) - f(x_{k-1})}$; na výpočet x_k je potrebné poznať 2 predchádzajúce hodnoty x_{k-1}, x_{k-2} ,

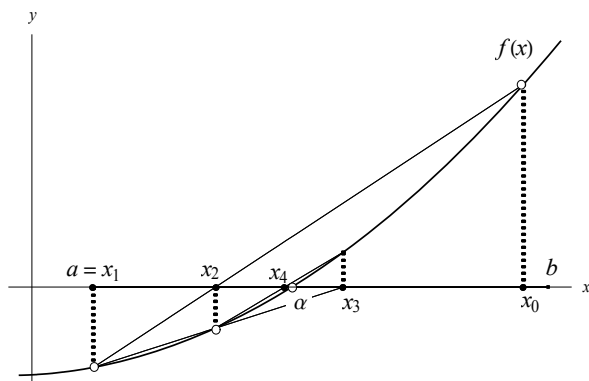
..... atď.

Dostávame tak postupnosť iterácií $\{x_0, x_1, \dots, x_k, \dots\}$, ktorá (pri zabezpečení podmienok konvergenzie metódy) konverguje k hľadanému koreňu α .

Ďalšie vlastnosti

- Ak metóda na danom intervale konverguje, štartovacími bodmi môžu byť ľubovoľné dva body intervalu $\langle a, b \rangle$. Zvyčajne volíme krajné body intervalu a alebo b , resp. body, ktoré ležia čo možno najbližšie k hľadanému koreňu α .
- Metóda je stacionárna – každú iteráciu počítame podľa tej istej iteračnej funkcie q .
- Dá sa odvodiť, že rýchlosť konvergenzie metódy je rádu $r = \frac{1}{2}(1 + \sqrt{5}) \doteq 1,618$.
- Metóda konverguje pomalšie ako Newtonova metóda.
- Metóda si nevyžaduje v každom kroku výpočet $f'(x_i)$, stačí poznať len $f(x_{i-1})$ a $f(x_i)$.
- Ak v Newtonovej metóde nahradíme $f'(x_i)$ pomernou diferenciou $\frac{f(x_{i-1}) - f(x_i)}{x_{i-1} - x_i}$, dostaneme Secant metódu.

Geometrická interpretácia metódy



Obr. 2.19. Metóda sečníc

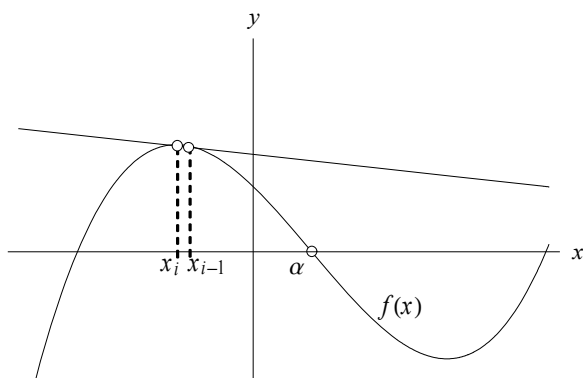
Zvolíme začiatočné aproximácie x_0, x_1 koreňa α . Bodmi $[x_0, f(x_0)], [x_1, f(x_1)]$ vedieme sečnicu. Jej priesečník s osou x označíme x_2 . Potom vedieme sečnicu bodmi $[x_1, f(x_1)], [x_2, f(x_2)]$ a jej priesečník s osou x označíme x_3 atď. Dostávame postupnosť aproximácií koreňa $\{x_0, x_1, x_2, x_3, \dots\}$, ktorá, ak metóda konverguje, konverguje k hľadanému koreňu α (Obr. 2.19).

„Úskalia ” metódy

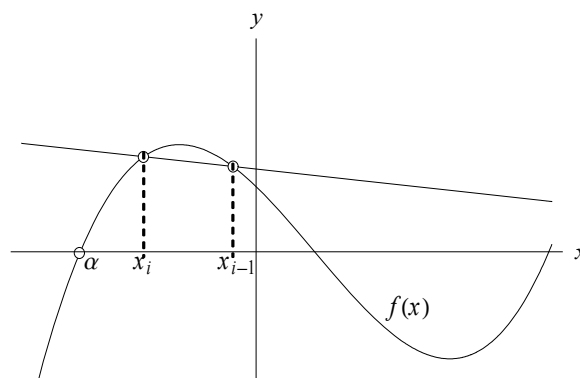
Metóda je zradná, ak:

- $|f(x_{i-1}) - f(x_i)| = 0$, t. j. $f(x_{i-1}) = f(x_i)$, ale $x_i \neq x_{i-1}$, sečnica je rovnobežná s osou x , os x nepretína vo vlastnom bode.

- $|f(x_{i-1}) - f(x_i)| \rightarrow 0$, t. j. $f(x_i) \approx f(x_{i-1})$, ale $x_i \neq x_{i-1}$ (pozri (2.7)). Potom sečnica grafu funkcie $f(x)$ zvierá s osou x veľmi malý uhol (Obr. 2.20). Jej priesečník s osou x leží ďaleko od hľadaného koreňa, mimo intervalu $\langle a_i, b_i \rangle$ (pozri kap.2.2.1). Ak rovnica má viac koreňov, metóda skonverguje k inému koreňu alebo bude divergovať. Podobne ako v prípade Newtonovej metódy malá hodnota menovateľa v iteračnej formule bude spôsobovať veľkú chybu vo výpočte. Z tohto dôvodu sa pokiaľ možno vyhýbame všetkým stacionárnym bodom funkcie a intervaly $\langle a_i, b_i \rangle$ volíme bez lokálnych extrémov (minimum, maximum), inflexných bodov a bodov, ktoré ležia v ich blízkom okolí.
- Interval $\langle a_i, b_i \rangle$ obsahuje lokálne extrém, minimum alebo maximum. Extrém ležiaci medzi x_i a x_{i-1} môže spôsobiť, že sečnica nepretne os x v blízkom okolí koreňa a iteračný proces skonverguje buď k inému koreňu alebo bude divergovať (Obr. 2.21).



Obr. 2.20



Obr. 2.21

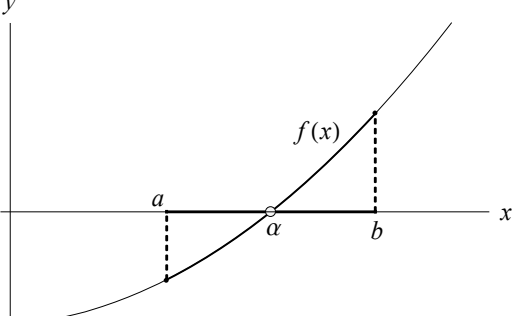
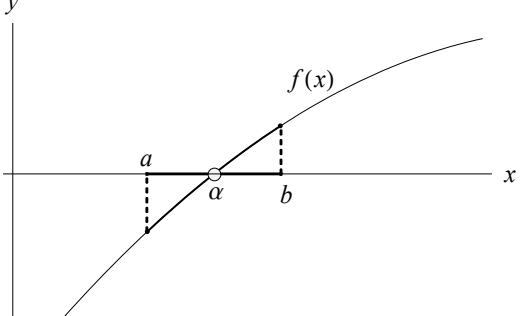
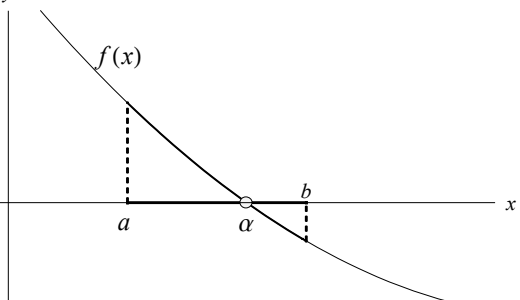
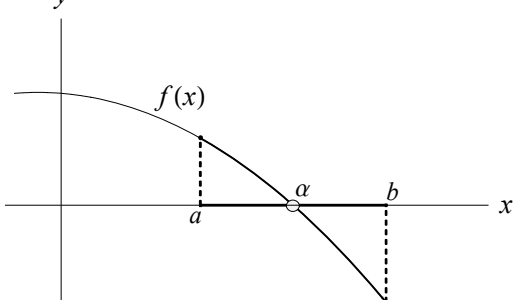
2.2.2.4 Konvergencia Newtonovej a Secant metódy

Konvergenciu oboch metód na $\langle a, b \rangle$ zaručuje táto veta:

Veta 2.5 Nech:

1. $f(x)$ je spojitá na $\langle a, b \rangle$,
2. $f(a) \cdot f(b) < 0$,
3. na $\langle a, b \rangle$ je $f'(x)$ spojitá a nemení znamienko, t. j. pre $\forall x \in \langle a, b \rangle$ je $f'(x) > 0$ alebo $f'(x) < 0$ ($f(x)$ je rýdzomonotónna – rastúca alebo klesajúca),
4. na $\langle a, b \rangle$ je $f''(x)$ spojitá a nemení znamienko, t. j. pre $\forall x \in \langle a, b \rangle$ je $f''(x) > 0$ alebo $f''(x) < 0$ ($f(x)$ je rýdzonkonvexná alebo rýdzokonkávna),

potom Newtonova aj Secant metóda konverguje ku koreňu α rovnice $f(x)=0$ pre ľubovoľné štartovacie body z intervalu $\langle a, b \rangle$. (pozri literatúru [3])

<p><i>Rastúca konvexná funkcia</i> $f'(x) > 0 \wedge f''(x) > 0$</p>	<p><i>Rastúca konkávna funkcia</i> $f'(x) > 0 \wedge f''(x) < 0$</p>
	
<p><i>Klesajúca konvexná funkcia</i> $f'(x) < 0 \wedge f''(x) > 0$</p>	<p><i>Klesajúca konkávna funkcia</i> $f'(x) < 0 \wedge f''(x) < 0$</p>
	

Obr. 2.22. Konvergencia Newtonovej a Secant metódy – priebeh funkcie

2.3 RIEŠENÉ PRÍKLADY

PRÍKLAD č. 1

Jeden z koreňov rovnice $x \cdot 2^x = 1$ leží v intervale $\langle 0, 1 \rangle$.

1. Overte podmienky existencie a jednoznačnosti tohto koreňa na danom intervale.
2. Newtonovou metódou vypočítajte prvé dve iterácie x_1, x_2 koreňa rovnice pre štartovací bod $x_0 = 0,8$.

Môžeme iteráciu x_2 považovať za hľadaný koreň rovnice pri voľbe tolerancie $\varepsilon(x_k) < 10^{-3}$ (t.j. za numerické riešenie s prenosťou 10^{-3})? Svoje tvrdenie zdôvodnite.

3. Metódou lineárnej interpolácie – Secant metódou vypočítajte iteráciu x_2 koreňa rovnice pre štartovacie body $x_0 = 0,7, x_1 = 0,75$.

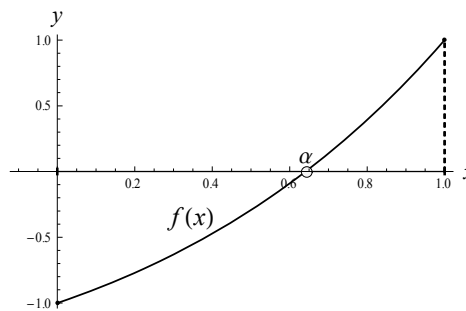
Môžeme iteráciu x_2 považovať za hľadaný koreň rovnice pri voľbe tolerancie $\varepsilon(x_k) < 10^{-3}$? Svoje tvrdenie zdôvodnite.

RIEŠENIE

Rovnicu najskôr upravíme na tvar $f(x) = 0$:

$$x \cdot 2^x = 1$$

$$x \cdot 2^x - 1 = 0 \Rightarrow f(x) = x \cdot 2^x - 1$$



Obr. 2.23. Graf $f(x)$

1. Podmienky existencie a jednoznačnosti koreňa na danom intervale

Z grafu funkcie $f(x)$ (Obr. 2.23) vidíme, že $f(x)$ je na danom intervale spojitá, rastúca a konvexná, graf $f(x)$ pretína os x , čo zaručuje existenciu a jednoznačnosť koreňa na danom intervale. Overíme analyticky:

Spojitosť $f(x)$ na $\langle 0,1 \rangle$: $D(f) = \mathbb{R} \Rightarrow \langle 0,1 \rangle \subset D(f) \Rightarrow f(x)$ je spojitá aj na $\langle 0,1 \rangle$.

Existencia koreňa na $\langle 0,1 \rangle$:

$$f(0) \cdot f(1) < 0$$

$$(0 \cdot 2^0 - 1) \cdot (1 \cdot 2^1 - 1) = (-1) \cdot 1 = -1 < 0 \Rightarrow \text{na } \langle 0,1 \rangle \text{ existuje koreň rovnice}$$

Jednoznačnosť koreňa rovnice:

Stacionárne body:

$$f'(x) = 0$$

$$1 \cdot 2^x + x \cdot 2^x \cdot \ln(2) - 0 = 0$$

$$2^x(1 + x \cdot \ln(2)) = 0 \Rightarrow 1 + x \cdot \ln(2) = 0 \text{ (pre } \forall x \in \mathbb{R} \text{ je } 2^x > 0)$$

$$x = -\frac{1}{\ln(2)}$$

$$x = -1,4427$$

$x = -1,4427 \notin \langle 0,1 \rangle$ a ani neleží v blízkosti tohto intervalu $\Rightarrow f'(x)$ nemení na $\langle 0,1 \rangle$ znamienko, t. j. v intervale $\langle 0,1 \rangle$ neleží žiadny stacionárny bod, $f(x)$ je na $\langle 0,1 \rangle$ rýdzomonotónna (funkcia je rastúca (pozri graf $f(x)$, Obr. 2.23).

Inflexné body funkcie:

$$f''(x) = 0$$

$$2^x \cdot \ln(2) + 2^x \cdot \ln(2) + x \cdot \ln(2) \cdot 2^x \cdot \ln(2) = 0$$

$$2^x \cdot \ln(2) \cdot (2 + x \cdot \ln(2)) = 0 \Rightarrow 2 + x \cdot \ln(2) = 0 \text{ (pre } \forall x \in \mathbb{R} \text{ je } 2^x \cdot \ln(2) > 0)$$

$$x = -\frac{2}{\ln(2)}$$

$$x = -2,88539$$

$x = -2,88539 \notin \langle 0,1 \rangle$ a ani neleží v blízkosti intervalu $\Rightarrow f''(x)$ nemení na $\langle 0,1 \rangle$ znamienko, $f''(0) = 2 \ln 2 > 0$, t. j. $f(x)$ je na $\langle 0,1 \rangle$ rýdzokonvexná (pozri graf $f(x)$, Obr. 2.23).

2. Newtonova metóda

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}, \quad i = 0, 1, 2, \dots; \quad f(x) = x \cdot 2^x - 1; \quad f'(x) = 2^x \cdot (1 + x \cdot \ln(2))$$

$$x_0 = 0,8 \quad (\text{štartovací bod})$$

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 0,8 - \frac{0,8 \cdot 2^{0,8} - 1}{2^{0,8} \cdot (1 + 0,8 \cdot \ln(2))} = 0,8 - \frac{0,39288}{2,70657} = 0,65484$$

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = 0,65484 - \frac{0,65484 \cdot 2^{0,65484} - 1}{2^{0,65484} \cdot (1 + 0,65484 \cdot \ln(2))} = 0,65484 - \frac{0,31011}{2,28909} = 0,64129$$

Daná je tolerancia: $\varepsilon(x_k) < 10^{-3}$.

Odhad chyby pre iteráciu koreňa x_2 :

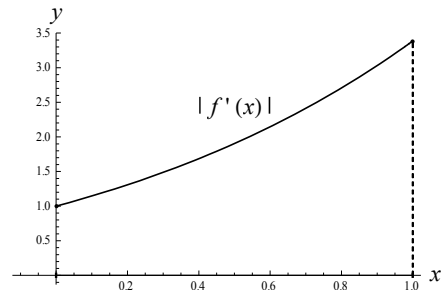
$$|x_2 - \alpha| \leq \varepsilon(x_2) \wedge \varepsilon(x_2) = \frac{|f(x_2)|}{m}, \quad m = \min |f'(x)|, \quad x \in \langle 0, 1 \rangle$$

Z grafu (Obr. 2.24):

$$m = |f'(0)| = |2^0 \cdot (1 + 0 \cdot \ln(2))| = 1$$

$$\varepsilon(x_2) = \frac{|f(x_2)|}{m} = \frac{|0,64129 \cdot 2^{0,64129} - 1|}{1}$$

$$= 0,00024 = 0,24 \cdot 10^{-3} < 10^{-3}$$



Obr. 2.24. Graf $|f'(x)|$ pre $x \in \langle 0, 1 \rangle$

Záver. Pri voľbe štartovacieho bodu $x_0 = 0,8$ a tolerancie vzhľadom na odhad chyby $\varepsilon(x_k) < 10^{-3}$ môžeme iteráciu $x_2 = 0,64129$ považovať za hľadaný koreň α s presnosťou 10^{-3} , t. j. $\alpha \doteq 0,64129$, $\alpha \in \langle x_2 - \varepsilon, x_2 + \varepsilon \rangle = \langle 0,65384; 0,65584 \rangle$ alebo zapíšeme $\alpha = 0,65484 \pm 10^{-3}$.

3. Secant metóda

$$x_{i+1} = x_i - f(x_i) \frac{x_{i-1} - x_i}{f(x_{i-1}) - f(x_i)}, \quad i = 1, 2, \dots, \quad f(x) = x \cdot 2^x - 1$$

$$x_0 = 0,7, \quad x_1 = 0,5 \quad (\text{štartovacie body})$$

$$x_2 = x_1 - f(x_1) \frac{x_0 - x_1}{f(x_0) - f(x_1)}$$

$$= 0,5 - (0,5 \cdot 2^{0,5} - 1) \cdot \frac{0,7 - 0,5}{(0,7 \cdot 2^{0,7} - 1) - (0,5 \cdot 2^{0,5} - 1)}$$

$$= 0,5 - (-0,29289) \cdot \frac{0,2}{0,13715 - (-0,29289)}$$

$$= 0,63621$$

Daná je tolerancia: $\varepsilon(x_k) < 10^{-3}$.

Odhad chyby pre iteráciu koreňa x_2 :

$$|x_2 - \alpha| \leq \varepsilon(x_2) \wedge \varepsilon(x_2) = \frac{|f(x_2)|}{m}, \quad m = \min |f'(x)|, \quad x \in \langle 0, 1 \rangle$$

Z grafu (Obr. 2.24):

$$m = |f'(0)| = |2^0 \cdot (1 + 0 \cdot \ln(2))| = 1$$

$$\varepsilon(x_2) = \frac{|f(x_2)|}{m} = \frac{|0,63621 \cdot 2^{0,63621} - 1|}{1} = 0,01116 = 11,16 \cdot 10^{-3} > 10^{-3}$$

Záver. Pri voľbe štartovacích bodov $x_0 = 0,7$ a $x_1 = 0,5$ a tolerancie $\varepsilon(x_k) < 10^{-3}$ iteráciu $x_2 = 0,63621$ **nemôžeme** považovať za hľadaný koreň, numerické riešenie rovnice $x \cdot 2^x = 1$.

3 VLASTNÉ ČÍSLA, VLASTNÉ VEKTORY MATICE

S vlastnými číslami a vlastnými vektormi matice sa stretávame vo viacerých oblastiach technickej praxe. Napríklad v mechanike motorov sa každé zo zalomení kľukového hriadeľa piestového spaľovacieho motora nahrádza torznou sústavou. Riešením tejto sústavy (budenie neuvažujeme) sú vlastné frekvencie (vlastné čísla) a zároveň aj vlastné tvary kmitania (vlastné vektory). Ďalej sa s nimi môžeme stretnúť pri chvení fyzikálnych sústav, v pružnosti pri vzperoch či v úlohách z kvantovej fyziky.

3.1 ZÁKLADNÉ POJMY LINEÁRNEJ ALGEBRY

V tejto časti stručne uvádzame tie základné pojmy lineárnej algebry, ktoré sú potrebné na osvojenie si problematiky uvedenej v tejto kapitole a v kapitole č.4.

Matica

Definícia 3.1 Matica A typu $m \times n$, $m, n \in N$ je množina $m \cdot n$ vo všeobecnosti komplexných čísel a_{ij} usporiadaných do m riadkov a n stĺpcov

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix}.$$

Stručne zapisujeme $A = A_{m \times n} = (a_{ij}) = (a_{i,j})$, $a_{ij} \in R \vee a_{ij} \in C$, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$.

Prvky a_{ii} , $i = 1, 2, \dots, \min(m, n)$ nazývame *diagonálnymi prvkami* matice A .

Tieto prvky ležia na *hlavnej diagonále* (hlavnej uhlopriečke) matice A .

Prvky $a_{1,n}, a_{2,n-1}, a_{3,n-2}, \dots, a_{m,1}$ ležia na *vedľajšej diagonále* (vedľajšej uhlopriečke) matice A .

Poznámka. Budeme pracovať len s maticami, ktorých prvky sú reálne čísla.

Definícia 3.2 Maticu \bar{x} typu $(1, n)$, $n \in N$, nazývame riadkovým vektorom a označujeme

$$\bar{x} = (x_1, x_2, \dots, x_n).$$

Definícia 3.3 Maticu \bar{x} typu $(n, 1)$, $n \in N$, nazývame stĺpcovým vektorom a označujeme

$$\bar{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}.$$

Operácie s maticami:

- súčet 2 matic: $A_{m \times n} + B_{m \times n} = (a_{ij} + b_{ij})_{m \times n}$, $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$,
- rozdiel 2 matic: $A_{m \times n} - B_{m \times n} = (a_{ij} - b_{ij})_{m \times n}$, $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$,
- súčin 2 matic: $A_{m \times n} \cdot B_{n \times p} = C_{m \times p}$; $c_{ij} = \sum_{k=1}^n a_{ik} \cdot b_{kj}$,
resp. $c_{ij} = iR_A \cdot jS_B$, kde iR_A je i -ty riadok matice A , jS_B je j -ty stĺpec matice B ,
- násobenie matice konštantou: $c \cdot A = (c \cdot a_{ij})$, $c \in R$, $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$.

Poznámka. Na označenie súčinu matic ako aj násobenia čísel (skalárov) je použitý ten istý symbol „ \cdot “. Upozorňujeme však, že operácia násobenia matic je iná ako násobenie čísel!

Typy matic a jej prvkov

Definícia 3.4 Matica $A = A_{m \times n}$ sa nazýva **obdĺžniková**, ak má rôzny počet stĺpcov a riadkov, t. j. $m \neq n$, napr.

$$A_{2 \times 3} = \begin{pmatrix} 1 & -2 & -3 \\ 0 & 4 & -7 \end{pmatrix}.$$

Definícia 3.5 Matica $A = A_{m \times n}$ sa nazýva **štvorcová** matica stupňa alebo rádu n , ak má taký istý počet riadkov ako stĺpcov, t. j. keď platí $m = n$; označujeme $A_{n \times n}$ alebo iba stručne A_n , napr.

$$A_2 = \begin{pmatrix} 1 & 2 \\ -3 & 0 \end{pmatrix}.$$

Definícia 3.6 Štvorcová matica $D = D_n$ sa nazýva **diagonálna** matica, ak má všetky prvky okrem prvkov hlavnej diagonály rovné 0, t. j. keď pre jej prvky platí $d_{i,j} = 0$ pre všetky $i \neq j$, $i, j = 1, 2, \dots, n$, napr.

$$D_3 = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Definícia 3.7 Matica $A = A_{m \times n}$ sa nazýva **nulová** matica, keď všetky jej prvky sa rovnajú nule, t. j. $a_{ij} = 0$, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$, napr.

$$A_{2 \times 3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Definícia 3.8 Štvorcová matica $E = E_n$ sa nazýva **jednotková** matica, ak matica E_n je diagonálna a všetky prvky na hlavnej diagonále sú rovné 1, t. j. keď pre jej prvky platí $e_{ii} = 1$, $e_{i,j} = 0$, $i \neq j$, $i, j = 1, 2, \dots, n$; napr.

$$E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Poznámka. Platí $E_n \cdot A_n = A_n \cdot E_n = A_n$.

Definícia 3.9 Matica $A^T = A_{n \times m}^T$ sa nazýva **transponovaná** matica k matici $A_{m \times n}$, keď platí $a_{ij}^T = a_{ji}$, napr.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Determinant matice

Definícia 3.10 Nech $A = A_n$ je štvorcová matica stupňa n . **Determinantom** matice A nazývame číslo $\det(A) = |A|$, ktoré je matici A priradené jednoznačne a vypočíta sa takto:

- pre maticu stupňa $n = 1$: $|A| = |a_{11}| = a_{11}$,
- pre maticu stupňa $n = 2$ podľa tzv. krížového pravidla: $|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}$,
- pre maticu stupňa $n = 3$ podľa tzv. Sarrusovho pravidla:

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = (a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13}) - (a_{31}a_{22}a_{13} + a_{12}a_{21}a_{33} + a_{23}a_{32}a_{11}),$$

Na výpočet môžeme použiť schému súčinov. Pod determinant podpíšeme prvé 2 riadky alebo k determinantu pripíšeme sprava prvé 2 stĺpce, pomocnými priamkami spojíme činitele v jednotlivých súčinoch. Súčiny orientované zľava doprava smerom dolu majú znamienko + a súčiny orientované zľava doprava smerom hore majú znamienko –.

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \quad \text{alebo} \quad |A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{vmatrix}$$

- pre maticu stupňa $n \geq 4$ použijeme
 - rozvoj podľa i -teho riadku: $|A| = \sum_{j=1}^n (-1)^{i+j} \cdot a_{ij} \cdot |A_{ij}|$,
 - rozvoj podľa j -teho stĺpca: $|A| = \sum_{i=1}^n (-1)^{i+j} \cdot a_{ij} \cdot |A_{ij}|$.

Matica A_{ij} sa nazýva **submatica** matice A a $|A_{ij}|$ je subdeterminantom matice A .

Definícia 3.11 **Subdeterminantom** (minorom) matice $A = A_n$, $n \geq 2$ k prvku a_{ij} , nazývame determinant submatice $|A_{ij}|$ stupňa $(n-1)$, ktorú dostaneme vynechaním i -teho riadku a j -teho stĺpca matice A .

Dalšie typy matíc

Definícia 3.12 Štvorcová matica $A = A_n$ sa nazýva **regulárna** matica, keď

$$|A| \neq 0 \quad (\det(A) \neq 0).$$

Definícia 3.13 Štvorcová matica A^{-1} sa nazýva **inverzná** matica k (regulárnej) matici A , keď platí

$$A^{-1} = \frac{1}{|A|} \cdot \begin{pmatrix} \alpha_{11} & \alpha_{21} & \dots & \alpha_{n1} \\ \alpha_{12} & \alpha_{22} & \dots & \alpha_{n2} \\ \dots & \dots & \dots & \dots \\ \alpha_{1n} & \alpha_{2n} & \dots & \alpha_{nn} \end{pmatrix},$$

kde $\alpha_{ij} = (-1)^{i+j} |A_{ij}|$, je algebraický doplnok k prvku a_{ij} matice A .

Poznámka.

• Pre regulárne matice A , B platí:

- $(A^{-1})^{-1} = A$,
- $A \cdot A^{-1} = A^{-1} \cdot A = E$,
- $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$,
- $|A^{-1}| = |A|^{-1}$.

Na nájdenie inverznej matice existuje viac metód, napr.

- Gauss-Jordanova metóda: $(A | E) \sim \text{ekviv. úpravy} \sim (E | A^{-1})$ alebo,
- pre maticu druhého stupňa platí: $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Rightarrow A^{-1} = \frac{1}{|A|} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$,

- pre diagonálnu maticu stupňa n platí: $D^{-1} = \begin{pmatrix} \frac{1}{d_{11}} & 0 & \dots & 0 \\ 0 & \frac{1}{d_{22}} & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & \frac{1}{d_{nn}} \end{pmatrix}, d_{ii} \neq 0.$

Definícia 3.14 Štvorcovú maticu $A = A_n$ považujeme za **riedku**, keď má rádovo n nenulových prvkov.

Poznámka. Ak maticu $A = A_{m+n}$ rozčleníme na 4 submatice $M_{m \times m}, N_{m \times n}, P_{n \times m}$ a $R_{n \times n}$,

$$A_{(m+n) \times (m+n)} = \begin{pmatrix} M_{m \times m} & N_{m \times n} \\ P_{n \times m} & R_{n \times n} \end{pmatrix}, \text{ potom aspoň jedna z matic } N_{m \times n}, P_{n \times m} \text{ je nulová } \Rightarrow |A| = |M| \cdot |R|,$$

napr.

$$A_5 = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 1 & 4 & 0 & 0 & 0 \\ 1 & 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 4 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}, M_2 = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}, N_{2 \times 3} = 0, P_{3 \times 2} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, R_3 = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 2 \end{pmatrix} \Rightarrow |A_5| = |M_2| \cdot |R_3| = 7 \cdot 12 = 84.$$

Definícia 3.15 Opakom riedkej matice je **plná** (všeobecná) matica A_n .

Definícia 3.16 Matica A sa nazýva **špeciálna** matica, ak má aspoň jednu z nasledujúcich vlastností:

Definícia 3.17 Štvorcová matica $A = A_n$ je **pásová** matica rádu n , keď existujú také čísla $p, q \in \mathbb{N}$, že $a_{ij} = 0$ pre $j > i + p$ alebo $i > j + q$, t. j. má:

- p pásov nenulových prvkov nad hlavnou diagonálou,
- q pásov nenulových prvkov pod hlavnou diagonálou,

Číslo $w = p+q+1$ sa nazýva šírka pásu.

Pásová matica je špeciálnym prípadom riedkej matice.

$$\text{Např. pásová matica } A_6 = \begin{pmatrix} a_{11} & * & * & * & 0 & 0 \\ \square & a_{22} & * & * & * & 0 \\ \square & \square & a_{33} & * & * & * \\ 0 & \square & \square & a_{44} & * & * \\ 0 & 0 & \square & \square & a_{55} & * \\ 0 & 0 & 0 & \square & \square & a_{66} \end{pmatrix}, w = p+w+1 = 3+2+1 = 6.$$

Definícia 3.18 Štvorcová matica $A = A_n$ je **ostro diagonálne dominantná** matica (ODD),

keď pre jej diagonálne prvky a_{ii} platí $|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|$, $i = 1, 2, \dots, n$, napr.

$$A = \begin{pmatrix} -5 & 2 & 1 \\ 2 & 4 & -1 \\ 1 & 0 & 3 \end{pmatrix}.$$

Definícia 3.19 Štvorcová matica $A = A_n$ sa nazýva **symetrická** práve vtedy, keď $A^T = A$, t. j. keď $a_{ij} = a_{ji}$ pre všetky dvojice indexov i, j , napr.

$$A = \begin{pmatrix} 3 & 1 & -2 \\ 1 & 0 & 4 \\ -2 & 4 & -7 \end{pmatrix}.$$

Definícia 3.20 Symetrická matica $A = A_n$ je **pozitívne definitná**, keď pre každý nenulový stĺpcový vektor $\bar{x} = (x_1, x_2, \dots, x_n)^T$ platí:

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

Je zrejmé, že výsledkom je číslo. Priamo overiť túto podmienku nie je ľahké. Inú, ľahšiu možnosť overenia, či je symetrická matica A pozitívne definitná, ponúka často uvádzaná (nutná a postačujúca) podmienka pozitívnej definitnosti známa ako Sylvesterovo kritérium:

Veta 3.1 (Sylvesterovo kritérium) Štvorcovú maticu $A = A_n$ nazveme pozitívne definitnou práve vtedy, keď všetky jej hlavné subdeterminanty A_i sú kladné, t. j. keď platí

$$|A_1| = |a_{11}| > 0, |A_2| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, |A_3| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0, \dots, |A_n| = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix} > 0.$$

Napr.

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 5 & 2 \\ 1 & 2 & 3 \end{pmatrix}, |A_1| = |1| = 1, |A_2| = \begin{vmatrix} 1 & 2 \\ 2 & 5 \end{vmatrix} = 1, |A_3| = \begin{vmatrix} 1 & 2 & 1 \\ 2 & 5 & 2 \\ 1 & 2 & 3 \end{vmatrix} = 2$$

Norma matice

Reálne číslo $\|A\|$ priradené matici A , ktoré je v určitom zmysle mierou jej veľkosti, sa nazýva **norma matice** A . Existuje veľa možností ako definovať normu matice. Obmedzíme sa na nižšie uvedené definície:

Definícia 3.21 Nech A je reálna matica. Potom pre túto maticu sú definované normy:

a) **riadková:** $\|A\|_R = \max_{i=1,2,\dots,n} \sum_{j=1}^n |a_{ij}|$ – (maximálny súčet v riadkoch),

b) **stĺpcová:** $\|A\|_S = \max_{j=1,2,\dots,n} \sum_{i=1}^n |a_{ij}|$ – (maximálny súčet v stĺpcoch),

c) **euklidovská:** $\|A\|_E = \sqrt{\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2}$.

Definícia 3.22 Pre normy reálneho stĺpcového vektora $\bar{x} = (x_1, x_2, \dots, x_n)^T$ platí:

a) **riadková:** $\|\bar{x}\|_R = \max_{i=1,2,\dots,n} |x_i|$,

b) **stĺpcová:** $\|\bar{x}\|_S = \sum_{i=1}^n |x_i|$,

c) **euklidovská:** $\|\bar{x}\|_E = \sqrt{\sum_{i=1}^n x_i^2}$.

Vlastnosti normy matice:

- $\|A\| \geq 0$, $\|A\| = 0$, ak A je nulová matica,
- $\|\alpha \cdot A\| = |\alpha| \cdot \|A\|$, kde $\alpha \in R$,
- $\|A + B\| \leq \|A\| + \|B\|$, trojuholníková nerovnosť,
- $\|A \cdot B\| \leq \|A\| \cdot \|B\|$,
- $\|A \cdot \bar{x}\| \leq \|A\| \cdot \|\bar{x}\|$, podmienka konzistencie.

3.2 ZÁKLADNÉ POJMY PRE VLASTNÉ ČÍSLA, VEKTORY MATICE

Definícia 3.23 Nech $A = A_{n \times n}$ je štvorcová matica. Číslo $\lambda \in C$ (vo všeobecnosti), pre ktoré má systém

$$A \cdot \bar{v} = \lambda \cdot \bar{v} \tag{3.1}$$

nenulové riešenie $\bar{v} \neq 0$, sa nazýva **vlastné číslo** matice A .

Definícia 3.24 Nenulový vektor

$$\bar{v} \neq \bar{0}, \bar{v} = (v_1, v_2, \dots, v_n)^T = \begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{pmatrix}$$

prislúchajúci vlastnému číslu λ , ktorý je riešením (3.1), sa nazýva **vlastný vektor** matice A .

Veta 3.2 Vlastné vektory prislúchajúce rôznym vlastným číslam sú lineárne nezávislé.

Definícia 3.25 Množinu vlastných čísel matice A voláme **spektrum** matice A .

Definícia 3.26 Maticu λ_A nazývame **spektrálnou maticou** prislúchajúcou matici A , pričom $\lambda_i, i = 1, 2, \dots, n$ sú všetky vlastné čísla matice A (vrátane ich násobnosti):

$$\lambda_A = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Definícia 3.27 Nech $\lambda_1, \lambda_2, \dots, \lambda_n$ sú všetky vlastné čísla matice A . Číslo $\rho(A)$ nazývame **spektrálny polomerom** matice A , kde $\rho(A) = \max\{|\lambda_1|, |\lambda_2|, \dots, |\lambda_n|\}$.

Veta 3.3 Nech $\rho(A)$ je spektrálny polomer matice A a $\|A\|$ je ľubovoľná norma matice A .

Potom platí $\rho(A) \leq \|A\|$.

Hľadanie vlastných čísel a vlastných vektorov sa často nazýva ako úplný alebo čiastočný problém vlastných čísel.

1. Úplný problém vlastných čísel spočíva v určení všetkých vlastných čísel, a to:
 - a) metódami založenými na výpočte vlastných čísel z charakteristického polynómu alebo,
 - b) metódami využívajúcimi vzťah podobnosti matíc.
2. Čiastočný problém vlastných čísel spočíva v určení len niektorých vlastných čísel, napr. dominantného (najväčšieho) resp. najmenšieho vlastného čísla, alebo vlastného čísla ležiaceho v určitom intervale.

3.3 VÝPOČET VLASTNÝCH ČÍSEL MATICE Z CHARAKTERISTICKÉHO POLYNÓMU

Pri výpočte vlastných čísel matice vychádzajme z rovnosti (3.1). Pripomíname, že matica A je stupňa n , $\lambda \in C$. Úpravami dostaneme:

$$\begin{aligned} A \cdot \bar{v} &= \lambda \cdot \bar{v} \\ A \cdot \bar{v} - \lambda \cdot \bar{v} &= \bar{0} \\ (A - \lambda \cdot E) \cdot \bar{v} &= \bar{0}, \end{aligned}$$

kde E je jednotková matica rádu n .

Systém $(A - \lambda \cdot E) \cdot \bar{x} = \bar{0}$ je homogénny systém, ktorý má nenulové riešenie, keď determinant matice tohto systému sa rovná nule, t. j.

$$|A - \lambda \cdot E| = 0. \quad (3.2)$$

Rovnicu (3.2) nazývame **charakteristická rovnica**.

Všimnime si ľavú stranu charakteristickej rovnice

$$\begin{aligned} |A - \lambda \cdot E| &= \left| \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix} - \lambda \cdot \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix} \right| = \\ &= \left| \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix} - \begin{pmatrix} \lambda & 0 & 0 & 0 & \dots & 0 \\ 0 & \lambda & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & \lambda \end{pmatrix} \right| = \left| \begin{matrix} a_{11} - \lambda & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} - \lambda \end{matrix} \right| = P_n(\lambda). \end{aligned}$$

Polynóm $P_n(\lambda) = |A_n - \lambda \cdot E_n|$ sa nazýva **charakteristický polynóm**. Je stupňa n , $n \in N$.

Z definície Definícia 3.23 vyplýva, že vlastné čísla matice A sú korene charakteristického polynómu $P_n(\lambda)$.

Poznámka. Korene polynómu $P_n(\lambda)$ sú vo všeobecnosti komplexné čísla. Komplexné čísla majú tvar $z = a + bi$, kde a, b sú reálne čísla a i je imaginárna (komplexná) jednotka, $i^2 = -1$. Zobrazujú sa v komplexnej rovine ako body z so súradnicami $[a, b]$. V množine komplexných čísel rozlišujeme:

- a) reálne čísla $z = a + bi$, $b = 0$,
- b) imaginárne čísla $z = a + bi$, $b \neq 0$.

Čísla $z_1 = a + bi$ a $z_2 = a - bi$ nazývame komplexne združené čísla.

Ak matica A má n vlastných čísel $(\lambda_1, \lambda_2, \dots, \lambda_n)$, potom polynóm $P_n(\lambda)$ môžeme rozložiť na súčin koreňových činiteľov takto:

$$P_n(\lambda) = (\lambda - \lambda_1) \cdot (\lambda - \lambda_2) \cdot (\lambda - \lambda_3) \cdot (\lambda - \lambda_4) \cdot \dots \cdot (\lambda - \lambda_{n-1}) \cdot (\lambda - \lambda_n).$$

Zo stupňa polynómu vieme potom určiť koľko reálnych a koľko imaginárnych (komplexne združených) vlastných čísel môže mať matica A . Pomôže nám v tom nasledujúca úvaha:

- Ak n je **párne** číslo, potom $P_n(\lambda) = M_2(\lambda) \cdot N_2(\lambda) \cdot \dots \cdot R_2(\lambda)$, kde
 $M_2(\lambda) = (\lambda - \lambda_1) \cdot (\lambda - \lambda_2)$, $N_2(\lambda) = (\lambda - \lambda_3) \cdot (\lambda - \lambda_4)$, ..., $R_2(\lambda) = (\lambda - \lambda_{n-1}) \cdot (\lambda - \lambda_n)$.
- Ak n je **nepárne** číslo, potom $P_n(\lambda) = M_2(\lambda) \cdot N_2(\lambda) \cdot \dots \cdot R_2(\lambda) \cdot G_1(\lambda)$, kde
 $M_2(\lambda) = (\lambda - \lambda_1) \cdot (\lambda - \lambda_2)$, ..., $R_2(\lambda) = (\lambda - \lambda_{n-2}) \cdot (\lambda - \lambda_{n-1})$, $G_1(\lambda) = (\lambda - \lambda_n)$.

Dá sa ukázať, že absolútny člen charakteristického polynómu $P_n(\lambda)$ sa rovná determinantu matice A a zároveň súčinu vlastných čísel. Pripomíname, že korene polynómu 2. stupňa môžu byť 2 reálne alebo 2 komplexne združené čísla. Polynóm 1. stupňa ($G_1(\lambda)$) má len 1 reálny koreň.

Príklad 1.

Matici $A_{7 \times 7}$ bude prislúchať 7 vlastných čísel, ktoré sú riešením charakteristickej rovnice (3.2). Charakteristický polynóm bude 7. stupňa, pričom

$$P_7(\lambda) = M_2(\lambda) \cdot N_2(\lambda) \cdot R_2(\lambda) \cdot G_1(\lambda).$$

Vlastné čísla matice $A_{7 \times 7}$ preto môžu byť:

typ \ možnosti	1.	2.	3.	4.
reálne	7	5	3	1
imaginárne (komplexne združené)	0	2	4	6
Σ	7	7	7	7

Príklad 2.

Matici $A_{6 \times 6}$ bude prislúchať 6 vlastných čísel, ktoré sú riešením (3.2). Charakteristický polynóm bude 6. stupňa, pričom $P_6(\lambda) = M_2(\lambda) \cdot N_2(\lambda) \cdot R_2(\lambda)$.

Vlastné čísla matice $A_{6 \times 6}$ preto môžu byť:

typ \ možnosti	1.	2.	3.	4.
reálne	6	4	2	0
imaginárne (komplexne združené)	0	2	4	6
Σ	6	6	6	6

Vlastnosti vlastných čísel

- Násobnosťou vlastného čísla nazývame násobnosť koreňa rovnice (3.2).
- Ak matica A je stupňa n , potom má n vlastných čísel ($\lambda_1, \lambda_2, \dots, \lambda_n$) vrátane ich násobnosti.
- Vlastné čísla reálnej a symetrickej matice A sú reálne čísla.

- Vlastné čísla reálnej (nie symetrickej) matice A môžu byť aj reálne aj komplexne združené čísla.
- Ak matica A je symetrická a pozitívne definitná, potom všetky jej vlastné čísla sú kladné čísla.
- Ak matica A je trojuholníková (pozri definície 4.2 a 4.3 v kapitole 4), potom jej vlastné čísla sú prvky hlavnej diagonály.

3.4 ODHAD POLOHY VLASTNÝCH ČÍSEL MATICE

Výpočet vlastných čísel matice A z charakteristickej rovnice (3.2) nie je jednoduchý, najmä ak ide o rozsiahlejšie matice vyššieho rádu. Preto vlastné čísla matice, ako aj hodnotu jej spektrálneho polomeru $\rho(A)$ väčšinou odhadujeme. Na výpočet odhadu existujú rôzne numerické metódy.

Na odhad polohy vlastných čísel v komplexnej rovine môžeme použiť niektorú z nasledovných viet:

Veta 3.4 (GERŠGORINOVA VETA 1)

Nech je daná štvorcová matica $A = A_n$. Potom všetky jej vlastné čísla ležia v oblasti $D = \bigcup_{i=1}^n C_i$

, kde $C_i = \{z, z \in C, |z - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|\}, i = 1, 2, \dots, n\}$, t. j. C_i je kruh so stredom $S_i = (a_{ii}, 0)$

a polomerom $r_i = \sum_{\substack{i=1 \\ i \neq j}}^n |a_{ij}|$.

Polomer r_i sa rovná súčtu absolútnych hodnôt prvkov a_{ij} okrem prvku a_{ii} v i -tom riadku.

Veta 3.5 (GERŠGORINOVA VETA 2)

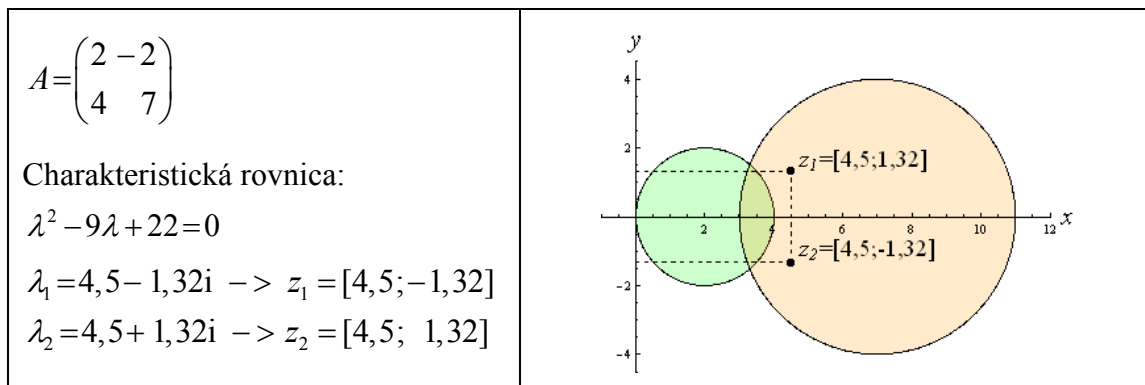
Keď zjednotenie kruhov z vety Veta 3.4 tvorí niekoľko súvislých množín, potom v každej z nich leží toľko vlastných čísel, koľko kruhov ju tvorí.

Poznámka. Veta nehovorí nič o tom, v ktorej časti súvislej množiny vlastné čísla ležia. Môže to byť tak prienik kruhov ako aj oblasť mimo neho.

Napríklad matici $A = \begin{pmatrix} 2 & -2 \\ 4 & 7 \end{pmatrix}$ prislúchajú dve vlastné čísla. Ich polohu v komplexnej

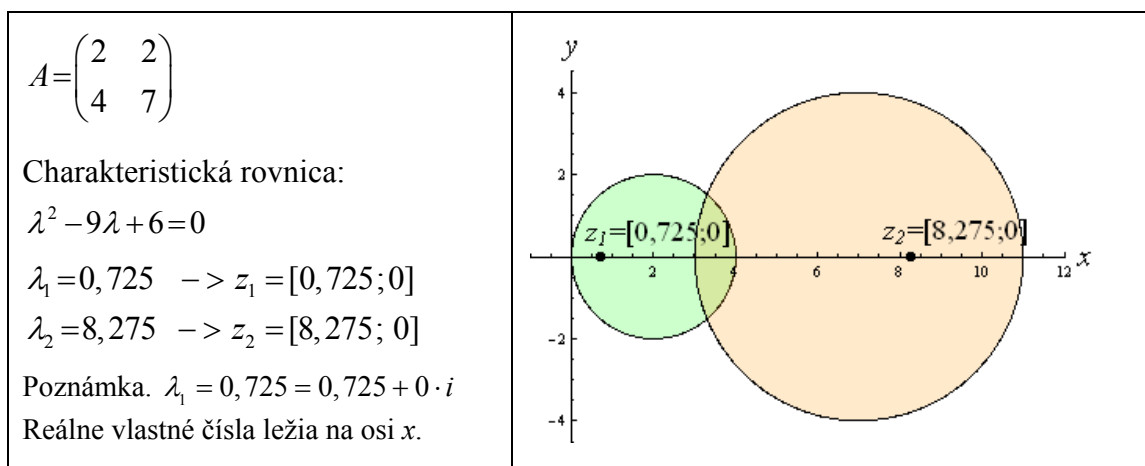
rovine vieme odhadnúť pomocou oblasti $D = \bigcup_{i=1}^2 C_i$ a ich presné hodnoty získame riešením

charakteristickej rovnice $\lambda^2 - 9\lambda + 22 = 0$, (Obr. 3.1).



Obr. 3.1. Presné riešenie a graf polohy vlastných čísel matice A

Matici $A = \begin{pmatrix} 2 & 2 \\ 4 & 7 \end{pmatrix}$ prislúchajú tiež dve vlastné čísla, tentoraz reálne (Obr. 3.2).



Obr. 3.2. Presné riešenie a graf polohy vlastných čísel matice A

3.5 ČIASŤOČNÝ PROBLÉM VLASTNÝCH ČÍSEL

Čiastočný problém vlastných čísel spočíva v určení len niektorých vlastných čísel.

Definícia 3.28 Dominantným (najväčším v absolútnej hodnote) vlastným číslom matice A nazývame vlastné číslo λ_d s najväčšou absolútnou hodnotou:

$$|\lambda_d| = \max \{ |\lambda_1|, |\lambda_2|, |\lambda_3|, \dots, |\lambda_n| \}, \text{ kde } \lambda_i \text{ sú vlastné čísla matice } A.$$

Definícia 3.29 Najmenším vlastným číslom matice A nazývame vlastné číslo λ_m s najmenšou absolútnou hodnotou:

$$|\lambda_m| = \min \{ |\lambda_1|, |\lambda_2|, |\lambda_3|, \dots, |\lambda_n| \}.$$

Polohu najmenšieho vlastného čísla vieme odhadnúť na základe tejto vety:

Veta 3.6 Pre najmenšie vlastné číslo matice A platí odhad:

$$|\lambda_m| \leq \sqrt[n]{k_1 \cdot k_2 \cdot k_3 \cdot \dots \cdot k_n}, \text{ kde } k_i = \sqrt{\sum_{j=1}^n a_{ij}^2}, i = 1, 2, \dots, n.$$

V praxi väčšinou hľadáme dominantné vlastné číslo matice. Na jeho výpočet sa používa viacero metód. Najznámejšia z nich je mocninová metóda a jej modifikácie (napr. Rayleighho podiel).

3.5.1 MOCNINOVÁ METÓDA

Veta 3.7 Predpoklady:

- a) Nech matica $A = A_{n \times n}$ má n lineárne nezávislých vlastných vektorov $\bar{v}_i, i = 1, 2, \dots, n$.
- b) Nech pre vlastné čísla matice A platí:

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|, \tag{3.3}$$

a teda existuje jediné dominantné vlastné číslo $\lambda_d = \lambda_1$ matice A .

Zostrojme postupnosť vektorov $\{\bar{y}^{(k)}\}_{k=0}^n$ takú, že

$$\bar{y}^{(k+1)} = A \cdot \bar{y}^{(k)}, \quad k = 0, 1, 2, \dots (\text{iterácie}),$$

t. j.

$$\begin{aligned} \bar{y}^{(1)} &= A \cdot \bar{y}^{(0)} \\ \bar{y}^{(2)} &= A \cdot \bar{y}^{(1)} \\ &\text{atď,} \end{aligned}$$

kde štartovací vektor $\bar{y}^{(0)} = \sum_{i=1}^n \alpha_i \cdot \bar{v}_i = \alpha_1 \cdot \bar{v}_1 + \alpha_2 \cdot \bar{v}_2 + \dots + \alpha_n \cdot \bar{v}_n$, pričom $\alpha_i \in R$ a $\alpha_1 \neq 0$. Na základe definície (3.1):

$$\begin{aligned} \bar{y}^{(1)} &= A \cdot \bar{y}^{(0)} = \sum_{i=1}^n \alpha_i \cdot A \cdot \bar{v}_i = \sum_{i=1}^n \alpha_i \cdot \lambda_i \cdot \bar{v}_i, \\ \bar{y}^{(2)} &= A \cdot \bar{y}^{(1)} = A^2 \cdot \bar{y}^{(0)} = \sum_{i=1}^n \alpha_i \cdot A^2 \cdot \bar{v}_i = \sum_{i=1}^n \alpha_i \cdot \lambda_i^2 \cdot \bar{v}_i \\ &\dots \\ \bar{y}^{(k)} &= A \cdot \bar{y}^{(k-1)} = A^k \cdot \bar{y}^{(0)} = \sum_{i=1}^n \alpha_i \cdot A^k \cdot \bar{v}_i = \sum_{i=1}^n \alpha_i \cdot \lambda_i^k \cdot \bar{v}_i \\ &= \lambda_1^k \left(\alpha_1 \cdot \bar{v}_1 + \alpha_2 \cdot \frac{\lambda_2^k}{\lambda_1^k} \cdot \bar{v}_2 + \dots + \alpha_n \cdot \frac{\lambda_n^k}{\lambda_1^k} \cdot \bar{v}_n \right), \end{aligned}$$

ďalej $\lim_{k \rightarrow \infty} \frac{\lambda_i^k}{\lambda_1^k} = 0$ a $\lim_{k \rightarrow \infty} \frac{1}{\lambda_1^k} \bar{y}^{(k)} = \alpha_1 \bar{v}_1$ pre všetky $i = 2, 3, \dots, n$.

Potom pre ľubovoľný vektor \bar{z} , pre ktorý $\bar{z}^T \cdot \bar{y}^{(k)} \neq 0$ a pre najväčšie vlastné číslo matice A platí

$$\lambda_d = \lim_{k \rightarrow \infty} \frac{\bar{z}^T \cdot \bar{y}^{(k+1)}}{\bar{z}^T \cdot \bar{y}^{(k)}}. \quad (3.4)$$

Poznámka. Dominantnému vlastnému číslu $\lambda_d = \lambda_1$ zodpovedá vlastný vektor \bar{v}_1 .

Algoritmus výpočtu dominantného čísla matice A mocninovou metódou

1. *Bez zastavovacej podmienky, s vopred stanoveným počtom iterácií pit:*
 - a) Určíme vhodný štartovací vektor $\bar{y}^{(0)}$ a dostatočný počet iterácií *pit*. (Najčastejšie sa používa $\bar{y}^{(0)} = (1, 1, \dots, 1)^T$.)
 - b) Pre $k = 0$ až *pit* počítame:
 - i) $\bar{y}^{(k+1)} = A \cdot \bar{y}^{(k)}$,
 - ii) $\lambda_d^{(k+1)}$.
 - c) Výsledkom je $\lambda_d \doteq \lambda_d^{(k+1)}$.
2. *So zastavovacou podmienkou $|\lambda_d^{(k+1)} - \lambda_d^{(k)}| < \varepsilon$, kde $\varepsilon > 0$ (veľmi malé) je požadovaná presnosť (tolerancia) výpočtu:*
 - a) Určíme vhodný štartovací vektor $\bar{y}^{(0)}$ a dostatočný počet iterácií *pit*.
 - b) Pre $k=0$ až *pit* počítame:
 - i) $\bar{y}^{(k+1)} = A \cdot \bar{y}^{(k)}$,
 - ii) $\lambda_d^{(k+1)}$,
 - iii) zisťujeme platnosť danej zastavovacej podmienky:
 - ak podmienka neplatí, pokračujeme výpočtom ďalšej iterácie,
 - ak podmienka platí, výpočet je ukončený.
 - c) Výsledkom je $\lambda_d \doteq \lambda_d^{(k+1)}$, posledná vypočítaná aproximácia.

Vzhľadom na to, že na výpočet dominantného vlastného čísla (3.4) môžeme použiť ľubovoľný vektor \bar{z} , existuje veľa spôsobov, ako určiť $(k+1)$ -vú iteráciu dominantného vlastného čísla $\lambda_d^{(k+1)}$. Najčastejšie:

- vo vektore $\bar{y}^{(k)}$ nájdeme súradnicu s najväčšou absolútnou hodnotou a vo vektore \bar{z} zvolíme odpovedajúcu súradnicu rovnú 1, ostatné súradnice sa rovnajú 0:

$$\lambda_d^{(k+1)} = \frac{y_j^{(k+1)}}{y_j^{(k)}}, \text{ kde } |y_j^{(k)}| = \|\bar{y}^{(k)}\|_R \text{ (} y_j^{(k)} \text{ je v absolútnej hodnote najväčšia súradnica vektora}$$

$\bar{y}^{(k)}$). Takýmto spôsobom sa minimalizuje zaokrúhľovacia chyba výpočtu (pozri kap. 1) a zároveň sa dosiahne rýchla konvergencia metódy.

- vektor $\bar{z} = \bar{y}^{(k)}$, $\lambda_d^{(k+1)} = \frac{\bar{y}^{(k)T} \cdot \bar{y}^{(k+1)}}{\bar{y}^{(k)T} \cdot \bar{y}^{(k)}} = \frac{\bar{y}^{(k)T} \cdot \bar{y}^{(k+1)}}{\|\bar{y}^{(k)}\|_E}$.

Vlastný vektor prislúchajúci dominantnému vlastnému číslu $\lambda_d \doteq \lambda_d^{(k+1)}$ vypočítame:

$$\bar{v}_1^{(k+1)} = \frac{\bar{y}^{(k+1)}}{\|\bar{y}^{(k)}\|_E} = \left(\frac{y_1^{(k+1)}}{\|\bar{y}^{(k)}\|_E}, \frac{y_2^{(k+1)}}{\|\bar{y}^{(k)}\|_E}, \dots, \frac{y_n^{(k+1)}}{\|\bar{y}^{(k)}\|_E} \right)^T,$$

kde $\|\bar{y}^{(k)}\|_E$ je kladné číslo, euklidovská norma vektora $\bar{y}^{(k)}$.

„Úskalia“ metódy

Hlavný problém metódy je, že nevieme vopred zaručiť jej konvergenciu, rozhodnúť, či sú splnené požadované predpoklady vety Veta 3.7. V prípade, že:

- vlastné vektory nie sú lineárne nezávislé, matica A je singulárna, metóda zvyčajne konverguje veľmi pomaly, najmä ak závislé vektory odpovedajú dvom najväčším vlastným číslam λ_1, λ_2 v zmysle zoradenia (3.3).
- $|\lambda_1| = |\lambda_2|$, vtedy výsledkom iteračného procesu je trieda vektorov, ktoré sú lineárnou kombináciou λ_1 a λ_2 ,
- vektor $\bar{y}^{(0)}$ je takou lineárnou kombináciou vlastných vektorov \bar{v}_i ,

$$\bar{y}^{(0)} = \sum_{i=1}^n \alpha_i \cdot \bar{v}_i = \alpha_1 \cdot \bar{v}_1 + \alpha_2 \cdot \bar{v}_2 + \dots + \alpha_n \cdot \bar{v}_n,$$

že koeficient α_1 pri vlastnom vektore \bar{v}_1 , ktorému prislúcha dominantné vlastné číslo λ_1 , sa rovná 0, potom metóda neskonverguje ku správne vlastnému číslu. Táto situácia nastáva najmä v prípade, ak súradnice vektora $\bar{y}^{(0)}$ volíme celé čísla. Najjednoduchšou cestou, ako sa tomuto problému vyhnúť, je voliť súradnice náhodné desatinné čísla, ktoré, ak by aj spôsobili, že $\alpha_1 = 0$ (čo je veľmi málo pravdepodobné), obyčajne v dôsledku zaokrúhľovacích chýb v iteračnom procese skonvergujú tak, že $\alpha_1 \neq 0$.

3.6 RIEŠENÉ PRÍKLADY

PRÍKLAD č. 1

Vypočítajte riadkovú, stĺpcovú a euklidovskú normu

a) matice $A = \begin{pmatrix} 1 & 2 & 3 \\ -4 & 5 & -6 \\ 7 & 8 & 9 \end{pmatrix}$,

b) vektorov $\bar{x} = (1, -2, 3)^T$, $\bar{y} = (0, 2, 3)^T$.

RIEŠENIE

$$a) \|A\|_R = \max \{|1|+|2|+|3|; |-4|+|5|+|-6|; |7|+|8|+|9|\} = \max \{6; 15; 24\} = 24$$

$$\|A\|_S = \max \{|1|+|-4|+|7|; |2|+|5|+|8|; |3|+|-6|+|9|\} = \max \{12; 15; 18\} = 18$$

$$\|A\|_E = \sqrt{1^2 + 2^2 + 3^2 + (-4)^2 + 5^2 + (-6)^2 + 7^2 + 8^2 + 9^2} = \sqrt{285} \doteq 16,882$$

$$b) \|\bar{x}\|_R = \max \{1, 2, 3\} = 3 \qquad \|\bar{y}\|_R = \max \{0, 2, 3\} = 3$$

$$\|\bar{x}\|_S = |1|+|2|+|3| = 6 \qquad \|\bar{y}\|_S = |0|+|2|+|3| = 5$$

$$\|\bar{x}\|_E = \sqrt{1^2 + (-2)^2 + 3^2} = \sqrt{14} \doteq 3,742 \qquad \|\bar{y}\|_E = \sqrt{0^2 + 2^2 + 3^2} = \sqrt{13} \doteq 3,606$$

Poznámka. Všimnite si, že $\|\bar{x}\|_R = \|\bar{y}\|_R$. Z tejto skutočnosti ale nevyplýva, že $\bar{x} = \bar{y}$ (pozri zadanie).

PRÍKLAD č. 2

Daná je matica A .

1. Koľko vlastných čísel bude prislúchať matici A ?
2. Koľko reálnych a koľko imaginárnych vlastných čísel môže mať matica A ?

$$A = \begin{pmatrix} 1 & 1 & 3 \\ 0 & -2 & 3 \\ -2 & 3 & -16 \end{pmatrix}$$

3. Použitím Geršgorinových viet:
 - a) zobrazte oblasť $D = \bigcup D_i$, v ktorej ležia všetky vlastné čísla matice A ,
 - b) určte tie oblasti D_i , v ktorých určite ležia reálne vlastné čísla matice A , stanovte ich počet a bližšie odhadnite ich polohu,
 - c) určte tie oblasti D_i , v ktorých môžu ležať reálne vlastné čísla matice A , bližšie odhadnite ich polohu a stanovte ich počet,
 - d) určte tie oblasti D_i , v ktorých môžu ležať vlastné imaginárne čísla matice A a určte ich počet v príslušnej oblasti D_i ,
 - e) pomocou charakteristickej rovnice vypočítajte všetky vlastné čísla matice A ,
 - f) pre jedno reálne vlastné číslo matice A nájdite odpovedajúci vlastný vektor,
 - g) nájdite spektrálny polomer matice A ,
 - h) určte dominantné vlastné číslo matice A a overte ho Mocninovou metódou.

Počítajte s toleranciou $|\lambda_d^{(k+1)} - \lambda_d^{(k)}| < \varepsilon$, $\varepsilon = 0,1$ a $\bar{y}^{(0)} = (1, 1, -3)^T$.

RIEŠENIE

1. Počet vlastných čísel matice A

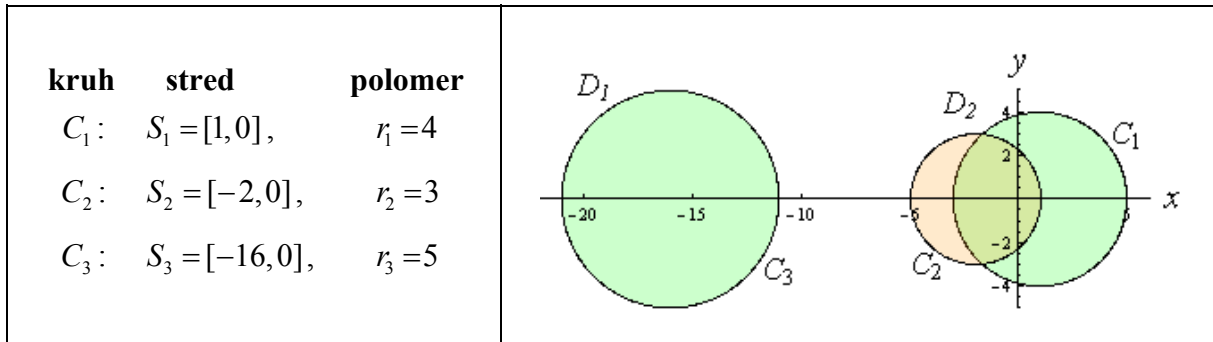
Matica A je typu 3×3 , budú jej prislúchať 3 vlastné čísla.

2. Typ vlastných čísel matice A

Matica A má všetky prvky reálne čísla, nie je symetrická \Rightarrow vlastné čísla matice A môžu byť aj reálne aj imaginárne, a to buď všetky reálne (0 komplexne združených) alebo 1 reálne vlastné číslo a 2 komplexne združené vlastné čísla

3. Geršgorinove vety:

a) oblasť $D = \bigcup D_i$:



Obr. 3.3. Oblasť vlastných čísel $D = D_1 \cup D_2$

Všetky vlastné čísla matice A ležia v oblasti $D = D_1 \cup D_2$ (Obr. 3.3).

Reálne vlastné číslo určite leží v oblasti D_1 . Tvorí ju jeden kruh – C_3 , to znamená, že v nej leží jedno vlastné číslo, ktoré môže byť iba reálne. Reálne vlastné číslo bude ležať v intervale $\langle -21, -11 \rangle$.

Reálne vlastné čísla môžu ležať ešte v oblasti $D_2 = C_1 \cup C_2$. Oblasť D_2 tvoria 2 kruhy, t. j. ležia v nej dve vlastné čísla. Ak budú reálne, potom budú ležať v intervale $\langle -5, 5 \rangle$.

Imaginárne vlastné čísla môžu ležať v oblasti $D_2 = C_1 \cup C_2$. Imaginárne vlastné čísla môžu byť iba komplexne združené. Oblasť D_2 tvoria 2 kruhy, môže v nej ležať dvojica komplexne združených vlastných čísel (mimo osi x).

Vlastné čísla matice A vypočítame z charakteristickej rovnice:

$$|A - \lambda E| = 0 \Leftrightarrow \begin{vmatrix} 1 - \lambda & 1 & 3 \\ 0 & -2 - \lambda & 3 \\ -2 & 3 & -16 - \lambda \end{vmatrix} = 0$$

$$(1 - \lambda) \cdot (-2 - \lambda) \cdot (-16 - \lambda) + 1 \cdot 3 \cdot (-2) + 0 \cdot 3 \cdot 3 - (-2) \cdot (-2) \cdot 3 - 3 \cdot 3 \cdot (1 - \lambda) - 1 \cdot 0 \cdot (-16 - \lambda) = 0$$

$$-\lambda^3 - 17\lambda^2 - 11\lambda + 5 = 0$$

$$-\lambda^3 - 17\lambda^2 - 16\lambda + 5\lambda + 5 = 0$$

$$-\lambda(\lambda^2 + 17\lambda + 16) + 5(\lambda + 1) = 0$$

$$-\lambda(\lambda + 1)(\lambda + 16) + 5(\lambda + 1) = 0$$

$$-(\lambda + 1)(\lambda(\lambda + 16) - 5) = 0$$

$$\begin{aligned}
 & -(\lambda+1)(\lambda^2+16\lambda-5)=0 \\
 & -(\lambda+1)(\lambda-(-8+\sqrt{69}))(\lambda-(-8-\sqrt{69}))=0 \\
 & \lambda_1=-1, \lambda_2=-8+\sqrt{69}, \lambda_3=-8-\sqrt{69}
 \end{aligned}$$

Vlastný vektor \vec{v}_1 prislúchajúci vlastnému číslu $\lambda_1=-1$ vypočítame nasledovne:

$$\begin{aligned}
 & (A-\lambda_1 E) \cdot \vec{v}_1 = \vec{0} \\
 & \begin{pmatrix} 1-(-1) & 1 & 3 \\ 0 & -2-(-1) & 3 \\ -2 & 3 & -16-(-1) \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 2 & 1 & 3 \\ 0 & -1 & 3 \\ -2 & 3 & -15 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
 & 2v_1 + v_2 + 3v_3 = 0 \\
 & -v_2 + 3v_3 = 0 \\
 & -2v_2 + 3v_3 - 15v_3 = 0
 \end{aligned}$$

Sčítaním 1. a 3. rovnice dostávame:

$$\begin{aligned}
 & 4v_2 - 12v_3 = 0 \\
 & -v_2 + 3v_3 = 0
 \end{aligned}$$

Vidíme, že rovnice sú lineárne závislé (1. rovnica je (-4) -násobok 2. rovnice. Jednu z nich preto vynecháme. Volíme jednoduchšiu z rovníc:

$$\begin{aligned}
 & -v_2 + 3v_3 = 0 \Rightarrow v_2 = 3v_3 \\
 & \text{vyjadríme} \quad v_3 = t, \quad t \in \mathbb{R} - \{0\} \quad - \text{parameter} \\
 & v_2 = 3t
 \end{aligned}$$

$$\text{Z rovnice } 2v_1 + v_2 + 3v_3 = 0 \Rightarrow v_1 = \frac{-v_2 - 3v_3}{2} = \frac{-3t - 3t}{2} = -3t$$

Záver. Vlastnému číslu $\lambda_1=-1$ prislúcha (nenulový) vlastný vektor:

$$\vec{v}_1 = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} -3t \\ 3t \\ t \end{pmatrix} = \begin{pmatrix} -3 \\ 3 \\ 1 \end{pmatrix} \cdot t = t \cdot (-3, 3, 1)^T, \quad t \in \mathbb{R} - \{0\}.$$

Spektrálny polomer matice A :

$$\begin{aligned}
 \rho(A) &= \max\{|\lambda_1|, |\lambda_2|, |\lambda_3|\} = \max\{|-1|, |-8+\sqrt{69}|, |-8-\sqrt{69}|\} = \\
 &= \max\{1; 0,3066; 16,3066\} = 16,3066 = |\lambda_3|.
 \end{aligned}$$

Dominantným vlastným číslom matice A je vlastné číslo matice A s najväčšou absolútnou hodnotou, $|\lambda_d| = \max\{|\lambda_1|, |\lambda_2|, |\lambda_3|\}$. V našej úlohe je to $\lambda_d = \lambda_3 = -8 - \sqrt{69} = -16,3066$.

Overíme ho mocninovou metódou (Tabuľka 3.1).

Vo výpočte použijeme $\lambda_d^{(k+1)} = \frac{y_j^{(k+1)}}{y_j^{(k)}}$, kde $|y_j^{(k)}| = \|\bar{y}^{(k)}\|_R$, zastavovacia podmienka má tvar $|\lambda_d^{(k+1)} - \lambda_d^{(k)}| < \varepsilon$, kde $\varepsilon = 0,1$ a $\bar{y}^{(0)} = (1, 1, -3)^T$, $\lambda_d^{(0)} = 0$.

postupnosť vektorov	dominantné vl. číslo	zastav. podmienka
$\bar{y}^{(k+1)} = A \cdot \bar{y}^{(k)}$, $\bar{y}^{(0)} = (1, 1, -3)^T$	$\lambda_d^{(k+1)}$	$ \lambda_d^{(k+1)} - \lambda_d^{(k)} < 0,1$
$\bar{y}^{(1)} = A \cdot \bar{y}^{(0)}$ $\bar{y}^{(1)} = \begin{pmatrix} 1 & 1 & 3 \\ 0 & -2 & 3 \\ -2 & 3 & -16 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ -3 \end{pmatrix} = \begin{pmatrix} -7 \\ -11 \\ 49 \end{pmatrix}$	$\lambda_d^{(1)} = \frac{y_3^{(1)}}{y_3^{(0)}} = \frac{49}{-3}$ $\lambda_d^{(1)} = -16,333$	$ \lambda_d^{(1)} - \lambda_d^{(0)} = 16,333$ $16,333 > 0,1$
$\bar{y}^{(2)} = A \cdot \bar{y}^{(1)}$ $\bar{y}^{(2)} = \begin{pmatrix} 1 & 1 & 3 \\ 0 & -2 & 3 \\ -2 & 3 & -16 \end{pmatrix} \cdot \begin{pmatrix} -7 \\ -11 \\ 49 \end{pmatrix} = \begin{pmatrix} 129 \\ 169 \\ -803 \end{pmatrix}$	$\lambda_d^{(2)} = \frac{y_3^{(2)}}{y_3^{(1)}} = \frac{-803}{49}$ $\lambda_d^{(2)} = -16,3878$	$ \lambda_d^{(2)} - \lambda_d^{(1)} = 0,054422$ $0,054422 < 0,1$

Tabuľka 3.1. Určenie dominantného čísla mocninovou metódou

Záver. Pri voľbe štartovacieho vektora $\bar{y}^{(0)} = (1, 1, -3)^T$ a tolerancie $\varepsilon = 0,1$ dominantným vlastným číslom matice A je 2. iterácia $\lambda_d^{(2)} = -16,3878$.

PRÍKLAD č. 2

Ukážte, že matica A je symetrická a pozitívne definitná. Využitím týchto vlastností matice A , Geršgorinových viet a odhadu pre λ_{\min} určte interval, v ktorom leží najmenšie vlastné číslo matice A .

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 5 \end{pmatrix}$$

RIEŠENIE

Matica $A = (a_{ij})$ je reálna ($a_{ij} \in R$) a symetrická ($A = A^T$):

1. riadok = 1. stĺpec
2. riadok = 2. stĺpec
3. riadok = 3. stĺpec

\Rightarrow všetky jej vlastné čísla budú reálne čísla (pozri vlastnosti vlastných čísel).

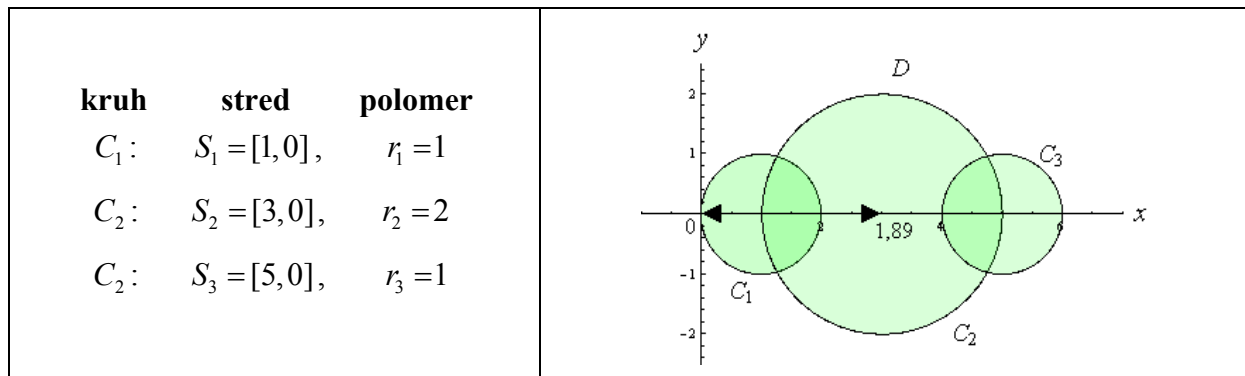
Zistíme, či matica A je aj pozitívne definitná, t. j. či všetky jej hlavné subdeterminanty $|A_i|$, $i = 1, 2, 3$ sú kladné čísla:

$$|A_1| = |1| = 1 > 0, \quad |A_2| = \begin{vmatrix} 1 & 1 \\ 0 & 3 \end{vmatrix} = 1 \cdot 3 - 0 \cdot 1 = 3 > 0, \quad |A_3| = |A| = \begin{vmatrix} 1 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 5 \end{vmatrix} = 9 > 0.$$

Všetky hlavné subdeterminanty $|A_i| > 0$, matica A je symetrická a pozitívne definitná \Rightarrow všetky jej vlastné čísla sú *kladné reálne čísla*.

Zobrazme oblasť $D = \bigcup C_i$, v ktorej ležia všetky vlastné čísla matice A (Obr. 3.4).

Všetky vlastné čísla matice A ležia v oblasti $D = C_1 \cup C_2 \cup C_3$. Vzhľadom na to, že sú to kladné reálne čísla, sú z intervalu $(0, 6)$.



Obr. 3.4. Oblasť vlastných čísel $D = C_1 \cup C_2 \cup C_3$

Pre veľkosť najmenšieho vlastného čísla λ_m matice A platí odhad:

$$|\lambda_m| \leq \sqrt[n]{k_1 \cdot k_2 \cdot \dots \cdot k_n}, \quad k_i = \sqrt{\sum_{j=1}^n a_{ij}^2}, \quad i=1, 2, \dots, n$$

$$k_1 = \sqrt{1^2 + 1^2 + 0^2} = \sqrt{2}, \quad k_2 = \sqrt{1^2 + 3^2 + 1^2} = \sqrt{11}, \quad k_3 = \sqrt{0^2 + 1^2 + 5^2} = \sqrt{26}$$

$$|\lambda_m| \leq \sqrt[3]{k_1 \cdot k_2 \cdot k_3}$$

$$|\lambda_m| \leq \sqrt[3]{\sqrt{2} \cdot \sqrt{11} \cdot \sqrt{26}}$$

$$|\lambda_m| \leq \sqrt[3]{\sqrt{572}}$$

$$|\lambda_m| \leq \sqrt[6]{572}, \quad \sqrt[6]{572} = 2,88115\dots$$

$$|\lambda_m| < 2,89 \Leftrightarrow -2,89 < \lambda_m < 2,89 \Leftrightarrow \lambda_m \in (-2,89; 2,89)$$

súčasne $\lambda_m > 0 \Rightarrow \lambda_m \in (0; 2,89)$.

Záver. Vzhľadom na oblasť D , v ktorej ležia všetky vlastné čísla matice A a vlastnosti matice A , môžeme tvrdiť, že $\lambda_m \in (0; 2,89)$.

4 METÓDY RIEŠENIA SYSTÉMOV LINEÁRNYCH ROVNÍC

Veľká časť problémov v praxi sa transformuje na systém n lineárnych rovníc o n neznámych, napr. prúťové sústavy v statike, prúdenie vzduchu okolo karosérie automobilu, šírenie, prestupy a prenosy tepla; v pružnosti a technickej mechanike – Metóda konečných prvkov a Metóda konečných diferencií atď.

Definícia 4.1 Systém rovníc

$$\begin{array}{l}
 a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n = b_1 \\
 a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n = b_2 \\
 a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \dots + a_{3n}x_n = b_3 \\
 \dots\dots\dots \\
 a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \dots + a_{nn}x_n = b_n
 \end{array}
 \xrightarrow[\text{tvar}]{\text{maticový}}
 \begin{array}{c}
 A \cdot \bar{x} = \bar{b} \\
 \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \dots \\ b_n \end{pmatrix}
 \end{array}$$

nazývame systémom n lineárnych rovníc o n neznámych (SLR), kde:

$$A_{n \times n} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix} \text{ je regulárna matica systému } (|A| \neq 0), \text{ , } a_{ij} \in R,$$

$$\bar{x}_{n \times 1} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = (x_1, x_2, \dots, x_n)^T \text{ je vektor riešení,}$$

$$\bar{b}_{n \times 1} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{pmatrix} = (b_1, b_2, \dots, b_n)^T \text{ je vektor pravých strán, } b_i \in R.$$

Na riešenie systémov lineárnych rovníc sa používajú metódy priame a iteračné.

4.1 PRIAME METÓDY RIEŠENIA SLR

Dávajú riešenie po konečnom počte elementárnych aritmetických operácií. Založené sú na eliminácii neznámych. Prvé tri z uvádzaných metód sú známe zo základného kurzu matematiky, v tejto kapitole budeme venovať pozornosť najmä 4. metóde.

1. Riešenie SLR pomocou inverznej matice A^{-1} k matici systému A

Metódu môžeme použiť len vtedy, keď A^{-1} k matici systému A existuje, t. j. keď matica systému A je regulárna ($|A| \neq 0$).

Hľadané riešenie:

$$\begin{aligned} A \cdot \bar{x} &= \bar{b} & / \cdot A^{-1} \\ A^{-1} \cdot A \cdot \bar{x} &= A^{-1} \cdot \bar{b} \\ E \cdot \bar{x} &= A^{-1} \cdot \bar{b} \\ \bar{x} &= A^{-1} \cdot \bar{b}. \end{aligned}$$

2. Riešenie SLR pomocou Cramerovho pravidla

Hľadané riešenie $\bar{x} = (x_1, x_2, \dots, x_n)^T = \left(\frac{|A_1|}{|A|}, \frac{|A_2|}{|A|}, \dots, \frac{|A_n|}{|A|} \right)^T$, kde $|A| \neq 0$, $|A_i|$, $i = 1, 2, \dots, n$, sú

subdeterminanty, ktoré dostaneme z $|A|$ tak, že v $|A|$ i -ty stĺpec nahradíme b , napríklad:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \bar{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, |A_1| = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}, |A_2| = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}, x_1 = \frac{|A_1|}{|A|}, x_2 = \frac{|A_2|}{|A|}.$$

Riešenie SLR pomocou Cramerovho pravidla sa vzhľadom na náročnosť výpočtov používa ojedinele, má skôr teoretický význam.

Osobitnú skupinu metód tvoria metódy, ktoré sú založené na úprave daného systému ekvivalentnými úpravami na ľahko riešiteľný systém s trojuholníkovou maticou.

3. Gaussova eliminačná metóda (GEM)

Jej základom je úprava rozšírenej matice systému na horný trojuholníkový tvar. Ten dostaneme pomocou ekvivalentných úprav, pričítaním vhodných násobkov vybraných riadkov matice k ostatným riadkom matice. Napr.

$$\begin{array}{c} A \quad | \bar{b} \\ \left(\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array} \right) \end{array} \approx \text{ekvivalentné úpravy} \approx \begin{array}{c} U \quad | \bar{y} \\ \left(\begin{array}{ccc|c} u_{11} & u_{12} & u_{13} & y_1 \\ 0 & u_{22} & u_{23} & y_2 \\ 0 & 0 & u_{33} & y_3 \end{array} \right) \end{array}$$

Hľadané riešenie: $u_{33} x_3 = y_3 \rightarrow u_{22} x_2 + u_{23} x_3 = y_2 \rightarrow u_{11} x_1 + u_{12} x_2 + u_{13} x_3 = y_1$.

Riešenie SLR pomocou inverznej matice, Gaussova eliminačná metóda ako aj Cramerovo pravidlo vedú priamo k riešeniu SLR. Keby sme sa nedopúšťali chýb zaokrúhľovania, našli by sme presné riešenie SLR. Avšak pre prácu s väčšími maticami (v praxi sa často počíta s maticami stupňa 100 a viac) priame metódy 1. – 3. nie sú pre ich náročnosť veľmi použiteľné. Napríklad GEM je náročná tak z časového, ako aj pamäťového hľadiska. Ak matica systému je stupňa n , potom metóda „potrebuje“ na nájdenie riešenia rádovo n^3 aritmetických operácií. V dôsledku veľkého počtu aritmetických operácií dochádza k mohutnému šíreniu chýb, a tak v snahe eliminovať tento nepriaznivý jav, sa výpočty realizujú až na dvojnásobný

počet desatinných miest ako je počet aritmetických operácií potrebných na nájdenie riešenia SLR, napr. ak matica A je stupňa 100, potom „potrebujeme“ vykonať rádovo $100^3 = 10^6$ aritmetických operácií a výpočty realizovať na 6-12 desatinných miest. GEM je vhodná na hľadanie riešenia nie príliš rozsiahlych systémov s plnou maticou. Významne menší počet aritmetických operácií ako GEM „potrebuje“ metóda LU rozkladu matice systému a iteračné metódy. Tieto metódy sú ľahko naprogramovateľné a použiteľné pri ďalšom spracovaní pomocou počítača, avšak ich použitie je limitované kapacitou pamäte počítača.

4. LU rozklad matice systému A

Metóda LU rozkladu je založená na rozklade matice systému na trojuholníkové matice.

4.1.1 TROJUHLNÍKOVÉ MATICE

Definícia 4.2 Matica $U = U_{n \times n}$, ktorá má všetky prvky pod hlavnou diagonálou rovné nule, sa nazýva horná trojuholníková matica, čo zapíšeme: $U = (u_{ij})$, $u_{ij} \neq 0$ pre $i \leq j$ a $u_{ij} = 0$ pre $i > j$, napr.

$$U = \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix}.$$

Definícia 4.3 Matica $L = L_{n \times n}$, ktorá má všetky prvky nad hlavnou diagonálou rovné nule, sa nazýva dolná trojuholníková matica, čo zapíšeme: $L = (l_{ij})$, $l_{ij} \neq 0$ pre $i \geq j$ a $l_{ij} = 0$ pre $i < j$, napr.

$$L = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix}.$$

Poznámka. Diagonálna matica je súčasne horná aj dolná trojuholníková matica.

Napr.

$$D = \begin{pmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{pmatrix}.$$

Dôležité vlastnosti trojuholníkových matíc rádu n :

- Ak T_1 a T_2 sú horné (dolné) trojuholníkové matice, potom sú opäť horné (dolné) trojuholníkové matice aj matice $T_1 + T_2$, $T_1 - T_2$, $T_1 \cdot T_2$, $T_2 \cdot T_1$.
- Ak T je trojuholníková matica, potom determinant matice T je rovný súčinu prvkov na hlavnej diagonále, t. j. $|T| = t_{11} \cdot t_{22} \cdot t_{33} \cdot \dots \cdot t_{nn}$.

- Ak $|T| \neq 0$, potom T^{-1} existuje a je to opäť horná (dolná) trojuholníková matica, ak T bola horná (dolná) trojuholníková matica.

4.1.2 RIEŠENIE SLR S TROJUHLNÍKOVÝMI MATICAMI

Princíp riešenia SLR s trojuholníkovými maticami ukážeme pre systém 3 lineárnych rovníc o 3 neznámych. Pre väčšie systémy bude postup analogický, len rozsah výpočtov bude väčší.

4.1.2.1 Riešenie SLR s dolnou trojuholníkovou maticou

$$L \cdot \bar{x} = \bar{b}$$

$$\begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \Leftrightarrow \begin{aligned} l_{11}x_1 + 0 + 0 &= b_1 & (1) \\ l_{21}x_1 + l_{22}x_2 + 0 &= b_2 & (2) \\ l_{31}x_1 + l_{32}x_2 + l_{33}x_3 &= b_3 & (3) \end{aligned}$$

Hľadané riešenie $\bar{x} = (x_1, x_2, x_3)^T$ nájdeme za predpokladu, že $l_{ii} \neq 0$ pre $i = 1, 2, 3$. Z rovnice

(1) dostávame $x_1 = \frac{b_1}{l_{11}}$, z rovnice (2) vypočítame $x_2 = \frac{b_2 - l_{21}x_1}{l_{22}}$ a z rovnice (3) vypočítame

$x_3 = \frac{b_3 - (l_{31}x_1 + l_{32}x_2)}{l_{33}}$. Všeobecne pre SLR typu $n \times n$ za predpokladu $l_{ii} \neq 0$ pre $i = 1, 2, \dots, n$

definujeme:

$$x_k = \frac{1}{l_{kk}} \left(b_k - \sum_{j=1}^{k-1} l_{kj} \cdot x_j \right), \quad k=1, 2, \dots, n.$$

4.1.2.2 Riešenie SLR s hornou trojuholníkovou maticou

$$U \cdot \bar{x} = \bar{b}$$

$$\begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \Leftrightarrow \begin{aligned} u_{11}x_1 + u_{12}x_2 + u_{13}x_3 &= b_1 & (1) \\ 0 + u_{22}x_2 + u_{23}x_3 &= b_2 & (2) \\ 0 + 0 + u_{33}x_3 &= b_3 & (3) \end{aligned}$$

Hľadané riešenie $\bar{x} = (x_1, x_2, x_3)^T$ nájdeme za predpokladu, že $u_{ii} \neq 0$ pre $i = 1, 2, 3$. Z rovnice

(3) dostávame $x_3 = \frac{b_3}{u_{33}}$, z rovnice (2) vypočítame $x_2 = \frac{b_2 - u_{23}x_3}{u_{22}}$ a z rovnice (1) vypočítame

$x_1 = \frac{b_1 - (u_{12}x_2 + u_{13}x_3)}{u_{11}}$. Všeobecne pre SLR typu $n \times n$ za predpokladu $u_{ii} \neq 0$ pre $i = 1, 2, \dots, n$

definujeme:

$$x_k = \frac{1}{u_{kk}} \left(b_k - \sum_{s=k+1}^n u_{ks} \cdot x_s \right), \quad k = n, n-1, \dots, 1.$$

Poznámka. Inverznú maticu k trojuholníkovej matici možno nájsť pomocou riešenia SLR s trojuholníkovými maticami. Princíp výpočtu naznačíme v nasledujúcej úlohe pre maticu stupňa $n = 3$. Pre maticu stupňa $n > 3$ by postup výpočtu bol analogický, len rozsah výpočtov by bol väčší.

ÚLOHA

Nájdite inverznú maticu L^{-1} k matici $L = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix}$, $l_{ii} \neq 0$, $i = 1, 2, 3$.

RIEŠENIE

Vieme, že L^{-1} k matici L existuje vtedy a len vtedy, keď matica L je regulárna, t. j. $|L| \neq 0$.

Pretože matica L má na hlavnej diagonále nenulové prvky, jej determinant $|L| = l_{11} \cdot l_{22} \cdot l_{33} \neq 0$, t. j. k matici L existuje inverzná matica L^{-1} .

Nech $L^{-1} = \begin{pmatrix} \tilde{l}_{11} & 0 & 0 \\ \tilde{l}_{21} & \tilde{l}_{22} & 0 \\ \tilde{l}_{31} & \tilde{l}_{32} & \tilde{l}_{33} \end{pmatrix}$ je hľadaná matica. Vzhľadom k tomu, že L je dolná trojuholníková

matica, tak aj L^{-1} bude dolná trojuholníková matica. Pretože $L \cdot L^{-1} = E$, dostávame:

$$\begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix} \cdot \begin{pmatrix} \tilde{l}_{11} & 0 & 0 \\ \tilde{l}_{21} & \tilde{l}_{22} & 0 \\ \tilde{l}_{31} & \tilde{l}_{32} & \tilde{l}_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

kde:

- prvky $\tilde{l}_{11}, \tilde{l}_{21}, \tilde{l}_{31}$ vypočítame pomocou súčinov všetkých riadkov matice L a 1. stĺpca matice L^{-1} ,
- prvky $\tilde{l}_{22}, \tilde{l}_{32}$ vypočítame pomocou súčinov všetkých riadkov matice L s 2. stĺpcom matice L^{-1} ,
- posledný prvok \tilde{l}_{33} vypočítame pomocou súčinov všetkých riadkov matice L s 3. stĺpcom matice L^{-1} .

Napríklad ak $L = \begin{pmatrix} 1 & 0 & 0 \\ 6 & 2 & 0 \\ 4 & 4 & 0 \end{pmatrix}$, potom k nej inverzná matica je $L^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 0,5 & 0 \\ 4 & -1 & 0,5 \end{pmatrix}$. Overte.

4.1.3 LU ROZKLAD MATICE SYSTÉMU

Riešenie systémov s hornou, resp. dolnou trojuholníkovou maticou je veľmi výhodné a zložitou nenáročné. Výhodou je najmä efektívnosť použitia tohto rozkladu. Pre $A_{3 \times 3}$ je efektívnosť LU rozkladu porovnateľná s Gaussovou eliminačnou metódou, pri maticiach vyššieho rádu ($n > 3$) efektívnosť narastá.

LU rozklad matice systému je výhodné použiť pri riešení systémov, ktorých matica je symetrická a pozitívne definitná alebo pásová ako aj pre systémy, ktoré majú rovnakú maticu systému A , ale rôzne matice pravých strán b . Takéto systémy nájdeme pri riešení parciálnych diferenciálnych rovníc v Metóde konečných prvkov alebo v Metóde konečných diferencií.

Uvažujme systém lineárnych rovníc, ktorý má práve jedno riešenie. Matica systému A je regulárna, t. j. $|A| \neq 0$.

Otázka: Môžeme maticu systému A rozložiť na hornú U a dolnú L trojuholníkovú maticu, a tak zjednodušiť náročnosť riešenia systému?

Odpoveď: Áno, za istých podmienok existuje rozklad matice systému A na hornú U a dolnú L trojuholníkovú maticu, t. j. $A = L \cdot U$.

V nasledujúcej vete uvádzame podmienky pre LU rozklad matice systému A .

Veta 4.1 Ak A je regulárna štvorcová matica, ktorá má všetky hlavné subdeterminanty nenulové, potom existuje taká dolná L a taká horná U trojuholníková matica, že platí $A = L \cdot U$.

Poznámka. Prv než pristúpime k LU rozkladu matice systému A , upravíme maticu A (súčasne aj vektor pravých strán \bar{b}) tak, aby na hlavnej diagonále nemala nulové prvky.

Pre jednoduchosť uvažujme o systéme 3 lineárnych rovníc o 3 neznámych.

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{array} \xrightarrow[\text{tvar}]{\text{maticový}} \begin{array}{l} A \cdot \bar{x} = \bar{b} \\ \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \end{array}$$

Rozložme maticu systému A na dolnú L a hornú U trojuholníkovú maticu, t. j.

$$A = L \cdot U$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix} \cdot \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix}$$

Tu je potrebné si uvedomiť, že prvky matice A poznáme, prvky matic L , U nepoznáme, chceme ich vypočítať.

Pre prvky matice A platí:

- prvok a_{11} vypočítame, keď vynásobíme 1. riadok matice L s 1. stĺpcom matice U , t. j.
 $a_{11} = 1R_L \cdot 1S_U$,
- prvok a_{12} vypočítame, keď vynásobíme 1. riadok matice L s 2. stĺpcom matice U , t. j.
 $a_{12} = 1R_L \cdot 2S_U$,
- ...
- prvok a_{33} vypočítame, keď vynásobíme 3. riadok matice L s 3. stĺpcom matice U , t. j.
 $a_{33} = 3R_L \cdot 3S_U$,

Dostaneme systém 9 lineárnych rovníc o 12 neznámých – 6 neznámých v matici L ($l_{11}, l_{21}, l_{22}, l_{31}, l_{32}, l_{33}$) a 6 neznámých v matici U ($u_{11}, u_{12}, u_{13}, u_{22}, u_{23}, u_{33}$):

$$\begin{aligned}
 a_{11} &= l_{11} \cdot u_{11} \\
 a_{12} &= l_{11} \cdot u_{12} \\
 a_{13} &= l_{11} \cdot u_{13} \\
 &\dots \\
 a_{33} &= l_{31} \cdot u_{13} + l_{32} \cdot u_{23} + l_{33} \cdot u_{33}.
 \end{aligned}
 \tag{4.1}$$

Aby mal systém (4.1) jednoznačné riešenie, treba v ňom počet neznámých zmenšiť na deväť. Tri neznáme zvolíme. Možností voľby 3 neznámých je veľa, pre jednoduchosť výpočtu položíme prvky ležiace na hlavnej diagonále buď v matici L alebo v matici U rovné 1. Tak bude **LU rozklad matice A daný jednoznačne**:

- Ak v matici L volíme diagonálne prvky $l_{ii}=1$, potom hovoríme o **DOOLITTLOVOM** rozklade:

$$A = L \cdot U$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{pmatrix} \cdot \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix}.$$

Postup výpočtu prvkov matice L a U :

$$\begin{array}{lll}
 a_{11} = 1R_L \times 1S_U : & a_{11} = 1 \cdot u_{11} + 0 \cdot 0 + 0 \cdot 0 = u_{11} & \Rightarrow u_{11} \\
 a_{21} = 2R_L \times 1S_U : & a_{21} = l_{21} \cdot u_{11} + 1 \cdot 0 + 0 \cdot 0 = l_{21} \cdot u_{11} & \Rightarrow l_{21} \\
 a_{31} = 3R_L \times 1S_U : & a_{31} = l_{31} \cdot 1 + l_{32} \cdot 0 + 1 \cdot 0 = l_{31} & \Rightarrow l_{31} \\
 \\ \\
 a_{12} = 1R_L \times 2S_U : & a_{12} = 1 \cdot u_{12} + 0 \cdot u_{22} + 0 \cdot 0 = u_{12} & \Rightarrow u_{12} \\
 a_{22} = 2R_L \times 2S_U : & a_{22} = l_{21} \cdot u_{12} + 1 \cdot u_{22} + 0 \cdot 0 = l_{21} \cdot u_{12} + u_{22} & \Rightarrow u_{22} \\
 a_{32} = 3R_L \times 2S_U : & a_{32} = l_{31} \cdot u_{12} + l_{32} \cdot u_{22} + 1 \cdot 0 = l_{31} \cdot u_{12} + l_{32} \cdot u_{22} & \Rightarrow l_{32}
 \end{array}$$

$$\begin{aligned}
 a_{13} = 1R_L \times 3S_U : \quad a_{13} &= 1 \cdot u_{13} + 0 \cdot u_{23} + 0 \cdot u_{33} = u_{13} & \Rightarrow & u_{13} \\
 a_{23} = 2R_L \times 3S_U : \quad a_{23} &= l_{21} \cdot u_{13} + 1 \cdot u_{23} + 0 \cdot u_{33} = l_{21} \cdot u_{13} + u_{23} & \Rightarrow & u_{23} \\
 a_{33} = 3R_L \times 3S_U : \quad a_{33} &= l_{31} \cdot u_{13} + l_{32} \cdot u_{23} + 1 \cdot u_{33} = l_{31} \cdot u_{13} + l_{32} \cdot u_{23} + u_{33} & \Rightarrow & u_{33}
 \end{aligned}$$

- Ak v matici U volíme diagonálne prvky $u_{ii}=1$, potom hovoríme o **CROUTOVOM** rozklade:

$$\begin{matrix} A & = & L & \cdot & U \\ \left(\begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{matrix} \right) & = & \left(\begin{matrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{matrix} \right) & \cdot & \left(\begin{matrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{matrix} \right)
 \end{matrix}$$

Postup výpočtu prvkov matice L a U :

$$\begin{aligned}
 a_{11} = 1R_L \times 1S_U : \quad a_{11} &= l_{11} \cdot 1 + 0 \cdot 0 + 0 \cdot 0 = l_{11} & \Rightarrow & l_{11} \\
 a_{12} = 1R_L \times 2S_U : \quad a_{12} &= l_{11}u_{12} + 0 \cdot 1 + 0 \cdot 1 = l_{11}u_{12} & \Rightarrow & u_{12} \\
 a_{13} = 1R_L \times 3S_U : \quad a_{13} &= l_{11} \cdot u_{13} + 0 \cdot u_{23} + 0 \cdot 1 = l_{11} \cdot u_{13} & \Rightarrow & u_{13}
 \end{aligned}$$

$$\begin{aligned}
 a_{21} = 2R_L \times 1S_U : \quad a_{21} &= l_{21} \cdot 1 + l_{22} \cdot 0 + 0 \cdot 0 = l_{21} & \Rightarrow & l_{21} \\
 a_{22} = 2R_L \times 2S_U : \quad a_{22} &= l_{21} \cdot u_{12} + l_{22} \cdot 1 + 0 \cdot 0 = l_{21} \cdot u_{12} & \Rightarrow & l_{22} \\
 a_{23} = 2R_L \times 3S_U : \quad a_{23} &= l_{21} \cdot u_{13} + l_{22} \cdot u_{23} + 0 \cdot 1 = l_{21} \cdot u_{13} + l_{22} \cdot u_{23} & \Rightarrow & u_{23}
 \end{aligned}$$

$$\begin{aligned}
 a_{31} = 3R_L \times 1S_U : \quad a_{31} &= l_{31} \cdot 1 + l_{32} \cdot 0 + l_{33} \cdot 0 = l_{31} & \Rightarrow & l_{31} \\
 a_{32} = 3R_L \times 2S_U : \quad a_{32} &= l_{31} \cdot u_{12} + l_{32} \cdot 1 + l_{33} \cdot 0 = l_{31} \cdot u_{12} + l_{32} & \Rightarrow & l_{32} \\
 a_{33} = 3R_L \times 3S_U : \quad a_{33} &= l_{31} \cdot u_{13} + l_{32} \cdot u_{23} + l_{33} \cdot 1 = l_{31} \cdot u_{13} + l_{32} \cdot u_{23} + l_{33} & \Rightarrow & l_{33}
 \end{aligned}$$

Riešenie systému $A \cdot \bar{x} = \bar{b}$ nájdeme postupom:

$$\begin{aligned}
 A \cdot \bar{x} &= \bar{b} \\
 L \cdot U \cdot \bar{x} &= \bar{b} \Rightarrow 1. \quad L \cdot \bar{y} = \bar{b}, \quad \bar{y} = ? \\
 & \quad \quad \quad 2. \quad U \cdot \bar{x} = \bar{y}, \quad \bar{x} = ? \quad - \text{hľadané riešenie systému,}
 \end{aligned}$$

kde:

$$L \cdot \bar{y} = \bar{b} \qquad U \cdot \bar{x} = \bar{y}$$

je systém s dolnou trojuholníkovou maticou je systém s hornou trojuholníkovou maticou

$$L \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \qquad U \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

Poznámka

- Všimnime si, že v Doolittlovom rozklade sú prvky matic L a U počítané po stĺpcoch – najskôr vypočítame všetky prvky, ktoré ležia v 1. stĺpci v matici L a U , potom všetky prvky, ktoré ležia v 2. stĺpci v matici L a U a nakoniec všetky prvky ležiace v 3. stĺpci v matici L a U . V Croutovom rozklade sú zasa prvky matic L a U počítané po riadkoch. Vo všeobecnosti nie je nutné dodržať toto poradie. Poradie počítania jednotlivých prvkov neovplyvní výsledok, avšak treba dbať na to, aby pri výpočte jednotlivých prvkov boli známe všetky potrebné hodnoty.
- Keď porovnáme prvky matic A , L a U , zistíme, že v Doolittlovom rozklade matica A a matica U majú rovnaký prvý riadok ($a_{11}=u_{11}$, $a_{12}=u_{12}$, $a_{13}=u_{13}$), v Croutovom rozklade matica A a matica L majú rovnaký prvý stĺpec ($a_{11}=l_{11}$, $a_{21}=l_{21}$, $a_{31}=l_{31}$). Z výpočtu je zrejmé, že je tak tomu vždy a výpočet môžeme obmedziť iba na výpočet zostávajúcich, v našom prípade, šiestich prvkov.

Okrem riešenia systémov lineárnych rovníc LU rozklad efektívne využívame aj pri výpočte determinantu matice A a determinantu inverznej matice A^{-1} :

- $|A|=|L|\cdot|U|$. Determinant matice A je rovný súčinu determinantov matice L a matice U .
- $|A^{-1}|=|L^{-1}|\cdot|U^{-1}|$. Za predpokladu, že matice L a U sú regulárne matice, determinant inverznej matice A^{-1} je rovný súčinu determinantov inverzných matic L^{-1} a U^{-1} .

4.1.4 RIEŠENÉ PRÍKLADY

PRÍKLAD č. 1

Je daný systém lineárnych rovníc:

$$A \cdot \bar{x} = \bar{b} \quad \Leftrightarrow \quad \begin{pmatrix} 2 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & -2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \\ 1 \end{pmatrix}.$$

- Je tento systém riešiteľný metódou LU rozkladu matice systému? Svoje tvrdenie zdôvodnite. Ak systém nie je riešiteľný metódou LU rozkladu matice, upravte ho ekvivalentnými úpravami tak, aby bol riešiteľný touto metódou.
- Nájdite riešenie systému Doolittlovým rozkladom.
- Nájdite riešenie systému Croutovým rozkladom.

RIEŠENIE

a) *Overme predpoklady existencie LU rozkladu matice systému A .*

- *Jednoznačnosť riešenia SLR* – overujeme regulárnosť matice A :

$$|A| = \begin{vmatrix} 2 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & -2 \end{vmatrix} = -1 \neq 0 \Rightarrow \text{systém rovníc má páve jedno riešenie.}$$

- Existencia LU rozkladu matice A – overujeme, či hlavné subdeterminanty matice A sú rôzne od nuly:

$$|A_1|=2 \neq 0, \quad |A_2|=\begin{vmatrix} 2 & 2 \\ 1 & 1 \end{vmatrix}=2 \cdot 1 - 1 \cdot 2 = 0, \quad |A_3|=|A|=1 \neq 0$$

Subdeterminant $|A_2|=0$, matica A sa v tomto tvare nedá rozložiť na matice L a U . Systém sa pokúsime upraviť ekvivalentnými úpravami tak, aby všetky hlavné subdeterminanty matice systému boli nenulové. Vymeňme v systéme 1. a 3. rovnicu, t. j. majme zmenený systém:

$$A \cdot \bar{x} = \bar{b} \quad \Leftrightarrow \quad \begin{pmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}.$$

Opäť overme predpoklad existencie LU rozkladu matice systému A : Sú všetky hlavné subdeterminanty v A rôzne od nuly?

$$|A_1|=1 \neq 0, \quad |A_2|=\begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix}=1 \cdot 1 - 1 \cdot 2 = -1 \neq 0, \quad |A_3|=\begin{vmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{vmatrix}=1 \neq 0$$

Predpoklady sú splnené, vykonáme LU rozklad matice systému A .

b) Doolittlov rozklad:

$$A = L \cdot U$$

$$\begin{pmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{pmatrix} \cdot \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix}$$

Postup výpočtu prvkov matice L a U :

$$\begin{aligned} a_{11} = 1R_L \times 1S_U : & \quad 1 = 1 \cdot u_{11} + 0 \cdot 0 + 0 \cdot 0 & \quad 1 = u_{11} & \quad \Rightarrow u_{11} = 1 \\ a_{21} = 2R_L \times 1S_U : & \quad 1 = l_{21} \cdot u_{11} + 1 \cdot 0 + 0 \cdot 0 & \quad 1 = l_{21} \cdot 1 & \quad \Rightarrow l_{21} = 1 \\ a_{31} = 3R_L \times 1S_U : & \quad 2 = l_{31} \cdot u_{11} + l_{32} \cdot 0 + 1 \cdot 0 & \quad 2 = l_{31} \cdot 1 & \quad \Rightarrow l_{31} = 2 \end{aligned}$$

$$\begin{aligned} a_{12} = 1R_L \times 2S_U : & \quad 2 = 1 \cdot u_{12} + 0 \cdot u_{22} + 0 \cdot 0 & \quad 2 = u_{12} & \quad \Rightarrow u_{12} = 2 \\ a_{22} = 2R_L \times 2S_U : & \quad 1 = l_{21} \cdot u_{12} + 1 \cdot u_{22} + 0 \cdot 0 & \quad 1 = 1 \cdot 2 + 1 \cdot u_{22} & \quad \Rightarrow u_{22} = -1 \\ a_{32} = 3R_L \times 2S_U : & \quad 2 = l_{31} \cdot u_{12} + l_{32} \cdot u_{22} + 1 \cdot 0 & \quad 2 = 2 \cdot 2 + l_{32} \cdot (-1) & \quad \Rightarrow l_{32} = 2 \end{aligned}$$

$$\begin{aligned} a_{13} = 1R_L \times 3S_U : & \quad -2 = 1 \cdot u_{13} + 0 \cdot u_{23} + 0 \cdot u_{33} & \quad -2 = u_{13} & \quad \Rightarrow u_{13} = -2 \\ a_{23} = 2R_L \times 3S_U : & \quad 1 = l_{21} \cdot u_{13} + 1 \cdot u_{23} + 0 \cdot u_{33} & \quad 1 = 1 \cdot (-2) + 1 \cdot u_{23} & \quad \Rightarrow u_{23} = 3 \\ a_{33} = 3R_L \times 3S_U : & \quad 1 = l_{31} \cdot u_{13} + l_{32} \cdot u_{23} + 1 \cdot u_{33} & \quad 1 = 2 \cdot (-2) + 2 \cdot 3 + u_{33} & \quad \Rightarrow u_{33} = -1 \end{aligned}$$

$$A = L \cdot U$$

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

iešenie systému

$$A \cdot \bar{x} = \bar{b}$$

$$L \cdot U \cdot \bar{x} = \bar{b} \quad \Rightarrow 1. \quad L \cdot \bar{y} = \bar{b}, \quad \bar{y} = ?$$

$$2. \quad U \cdot \bar{x} = \bar{y}, \quad \bar{x} = ?$$

1. $L \cdot \bar{y} = \bar{b}$ – systém s dolnou trojuholníkovou maticou

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} \Rightarrow \begin{array}{l} y_1 = 1 \\ y_1 + y_2 = 3 \\ 2y_1 + 2y_2 + y_3 = 5 \end{array} \Rightarrow \begin{array}{l} y_1 = 1 \\ y_2 = 2 \\ y_3 = -1 \end{array} \Rightarrow \bar{y} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$$

2. $U \cdot \bar{x} = \bar{y}$ – systém s hornou trojuholníkovou maticou

$$\begin{pmatrix} 1 & 2 & -2 \\ 0 & -1 & 3 \\ 0 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \Rightarrow \begin{array}{l} x_1 + 2x_2 - 2x_3 = 1 \\ -x_2 + 3x_3 = 2 \\ -x_3 = -1 \end{array} \Rightarrow \begin{array}{l} x_1 = 1 \\ x_2 = 1 \\ x_3 = 1 \end{array} \Rightarrow \bar{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = (1, 1, 1)^T$$

hľadané riešenie

c) Croutov rozklad:

$$A = L \cdot U$$

$$\begin{pmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{pmatrix} = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix} \cdot \begin{pmatrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{pmatrix}$$

Postup výpočtu prvkov matice L a U :

$$a_{11} = 1R_L \times 1S_U: \quad 1 = l_{11} \cdot 1 + 0 \cdot 0 + 0 \cdot 0 \quad 1 = l_{11} \quad \Rightarrow l_{11} = 1$$

$$a_{12} = 1R_L \times 2S_U: \quad 2 = l_{11} \cdot u_{12} + 0 \cdot 1 + 0 \cdot 1 \quad 2 = 1 \cdot u_{12} \quad \Rightarrow u_{12} = 2$$

$$a_{13} = 1R_L \times 3S_U: \quad -2 = l_{11} \cdot u_{13} + 0 \cdot u_{23} + 0 \cdot 1 \quad -2 = 1 \cdot u_{13} \quad \Rightarrow u_{13} = -2$$

$$a_{21} = 2R_L \times 1S_U: \quad 1 = l_{21} \cdot 1 + l_{22} \cdot 0 + 0 \cdot 0 \quad 1 = l_{21} \quad \Rightarrow l_{21} = 1$$

$$a_{22} = 2R_L \times 2S_U: \quad 1 = l_{21} \cdot u_{12} + l_{22} \cdot 1 + 0 \cdot 0 \quad 1 = 1 \cdot 2 + l_{22} \quad \Rightarrow l_{22} = -1$$

$$a_{23} = 2R_L \times 3S_U: \quad 1 = l_{21} \cdot u_{13} + l_{22} \cdot u_{23} + 0 \cdot 1 \quad 1 = 1 \cdot (-2) + (-1)u_{23} \quad \Rightarrow u_{23} = -3$$

$$a_{31} = 3R_L \times 1S_U: \quad 2 = l_{31} \cdot 1 + l_{32} \cdot 0 + l_{33} \cdot 0 \quad 2 = l_{31} \quad \Rightarrow l_{31} = 2$$

$$a_{32} = 3R_L \times 2S_U: \quad 2 = l_{31} \cdot u_{12} + l_{32} \cdot 1 + l_{33} \cdot 0 \quad 2 = 2 \cdot 2 + l_{32} \quad \Rightarrow l_{32} = -2$$

$$a_{33} = 3R_L \times 3S_U: \quad 1 = l_{31} \cdot u_{13} + l_{32} \cdot u_{23} + l_{33} \cdot 1 \quad 1 = 2 \cdot (-2) + (-2) \cdot (-3) + l_{33} \quad \Rightarrow l_{33} = -1$$

$$A = L \cdot U$$

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

Riešenie systému:

$$A \cdot \bar{x} = \bar{b}$$

$$L \cdot U \cdot \bar{x} = \bar{b} \Rightarrow 1. \ L \cdot \bar{y} = \bar{b}, \ \bar{y} = ?$$

$$2. \ U \cdot \bar{x} = \bar{y}, \ \bar{x} = ?$$

$L \cdot \bar{y} = \bar{b}$ – systém s dolnou trojuholníkovou maticou

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ 2 & -2 & -1 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} \Rightarrow \begin{matrix} y_1 = 1 \\ y_1 - y_2 = 3 \\ 2y_1 - 2y_2 - y_3 = 5 \end{matrix} \Rightarrow \begin{matrix} y_1 = 1 \\ y_2 = -2 \\ y_3 = 1 \end{matrix} \Rightarrow \bar{y} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

$U \cdot \bar{x} = \bar{y}$ – systém s hornou trojuholníkovou maticou

$$\begin{pmatrix} 1 & 2 & -2 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \Rightarrow \begin{matrix} x_1 + 2x_2 - 2x_3 = 1 \\ x_2 - 3x_3 = -2 \\ x_3 = 1 \end{matrix} \Rightarrow \begin{matrix} x_1 = 1 \\ x_2 = 1 \\ x_3 = 1 \end{matrix} \Rightarrow \bar{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = (1, 1, 1)^T$$

hladané riešenie

4.2 ITERAČNÉ METÓDY RIEŠENIA SYSTÉMOV LINEÁRNYCH ROVNÍC

Princíp iteračných metód ukážeme pre systém 3 lineárnych rovníc o 3 neznámých. Pre väčšie systémy bude postup analogický, len rozsah výpočtov bude väčší.

Majme teda systém 3 lineárnych rovníc o 3 neznámých:

$$\begin{matrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{matrix} \xrightarrow[\text{tvar}]{\text{maticový}} \begin{matrix} A \cdot \bar{x} = \bar{b} \\ \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \end{matrix}, \quad (4.2)$$

kde $A = A_{3 \times 3} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ je regulárna matica systému, ($|A| \neq 0$), $a_{ij} \in R$, $a_{ii} \neq 0$,

$$\bar{x}_{3 \times 1} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = (x_1, x_2, x_3)^T \text{ je vektor riešeni,}$$

$$\bar{b}_{3 \times 1} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = (b_1, b_2, b_3)^T \text{ je vektor pravých strán, } b_i \in \mathbb{R}.$$

Na základe predpokladu má systém $A \cdot \bar{x} = \bar{b}$ jediné presné riešenie $\bar{x}^p = A^{-1} \cdot \bar{b} = (x_1, x_2, x_3)^T$.

Preved'me systém $A \cdot \bar{x} = \bar{b}$ ekvivalentnými úpravami na systém $\bar{x} = H \cdot \bar{x} + \bar{g}$ tak, aby jeho riešením bolo riešenie \bar{x}^p . Dostávame:

$$\begin{array}{l} x_1 = h_{11}x_1 + h_{12}x_2 + h_{13}x_3 + g_1 \\ x_2 = h_{21}x_1 + h_{22}x_2 + h_{23}x_3 + g_2 \\ x_3 = h_{31}x_1 + h_{32}x_2 + h_{33}x_3 + g_3 \end{array} \xrightarrow[\text{tvar}]{\text{maticový}} \begin{array}{l} \bar{x} = H \cdot \bar{x} + \bar{g} \\ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix} \end{array}$$

Zostrojme postupnosť aproximácií vektora \bar{x} : $\{\bar{x}^{(k)}\}_{k=0}^{\infty} = \{\bar{x}^{(0)}, \bar{x}^{(1)}, \bar{x}^{(2)}, \dots, \bar{x}^{(n)}, \dots\}$ podľa iteračnej formuly:

$$\begin{aligned} \bar{x}^{(k+1)} &= H \cdot \bar{x}^{(k)} + \bar{g}, \quad k = 0, 1, 2, \dots & (4.3) \\ \bar{x}^{(1)} &= H \cdot \bar{x}^{(0)} + \bar{g} \\ \bar{x}^{(2)} &= H \cdot \bar{x}^{(1)} + \bar{g} \\ \bar{x}^{(3)} &= H \cdot \bar{x}^{(2)} + \bar{g}, \quad \dots \text{ atď.}, \end{aligned}$$

kde: $H = H_{3 \times 3}$ je iteračná matica,

$\bar{g} = \bar{g}_{3 \times 1}$ je stĺpcový vektor,

$x_i^{(k)}$ je k -ta iterácia i -tej zložky (súradnice) vektora \bar{x} , napr.

$$\bar{x}^{(3)} = H \cdot \bar{x}^{(2)} + \bar{g} \quad \Leftrightarrow \quad \begin{pmatrix} x_1^{(3)} \\ x_2^{(3)} \\ x_3^{(3)} \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \cdot \begin{pmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{pmatrix} + \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix}.$$

Ak iteračný proces (4.3) konverguje, tak konverguje k riešeniu systému $\bar{x} = H \cdot \bar{x} + \bar{g}$ a tiež aj k riešeniu systému $A \cdot \bar{x} = \bar{b}$.

Ak existuje $\lim_{k \rightarrow \infty} \bar{x}^{(k)} = \bar{x}^*$, potom zrejme aj $\lim_{k \rightarrow \infty} \bar{x}^{(k+1)} = \bar{x}^*$, a teda

$$\bar{x}^* = \lim_{k \rightarrow \infty} \bar{x}^{(k+1)} = \lim_{k \rightarrow \infty} (H \cdot \bar{x}^{(k)} + \bar{g}) = H \cdot \lim_{k \rightarrow \infty} \bar{x}^{(k)} + \bar{g} = H \cdot \bar{x}^* + \bar{g} \Rightarrow \bar{x}^* = H \cdot \bar{x}^* + \bar{g}$$

Záver:

$$\begin{array}{ccc} A \cdot \bar{x} = \bar{b} & \xrightarrow{\quad \quad \quad} & \bar{x}^p = A^{-1} \cdot \bar{b} \\ \text{ekviv. } \downarrow \text{ úpravy} & \quad \quad \quad |A| \neq 0 & \\ \bar{x} = H \cdot \bar{x} + \bar{g} & \xrightarrow{\quad \quad \quad} & \bar{x}^* = H \cdot \bar{x}^* + \bar{g} \end{array} \left. \vphantom{\begin{array}{ccc} A \cdot \bar{x} = \bar{b} \\ \bar{x} = H \cdot \bar{x} + \bar{g} \end{array}} \right\} \bar{x}^p = \bar{x}^*$$

$\bar{x}^{(k+1)} = H \cdot \bar{x}^{(k)} + \bar{g}, \quad \lim_{k \rightarrow \infty} \bar{x}^{(k+1)}$

Ak sa matica H a vektor \bar{g} počas iteračného procesu nemenia, jedná sa o **stacionárnu iteračnú metódu**.

K spusteniu iteračného procesu je nutné poznať:

- iteračnú formulu, v našom prípade (4.3),
- štartovací vektor $\bar{x}^{(0)}$ (volíme ľubovoľne, zvyčajne $\bar{x}^{(0)} = (0, 0, \dots, 0)^T$, resp. $\bar{x}^{(0)} = \bar{g}$),
- spôsob ukončenia:
 - a) vopred stanoveným (dostatočne veľkým) počtom iterácií *pit*, ktorý chceme vypočítať,
 - b) pomocou tzv. zastavovacej podmienky.

Numerickým riešením $\bar{x}^{(n)}$ systému (4.2) je pri konvergentných procesoch ostatná vypočítaná iterácia $\bar{x}^{(k+1)} = \bar{x}^{(n)}$, napr. pre $n = 50$, $\bar{x}^{(50)} \doteq \bar{x}^p$.

Pri numerických metódach nás vždy zaujíma veľkosť chyby, ktorej sa dopustíme pri výpočte. Pre presné riešenie \bar{x}^p platí: $\bar{b} - A \cdot \bar{x}^p = \bar{0} = (0, 0, 0)^T$, pre aproximáciu $\bar{x}^{(k)}$ zrejme $\bar{b} - A \cdot \bar{x}^{(k)} \neq \bar{0}$. Označme $\bar{b} - A \cdot \bar{x}^{(k)} = \bar{r}^{(k)}$.

Definícia 4.4 Vektor $\bar{r}^{(k)} = \bar{b} - A \cdot \bar{x}^{(k)}$, resp. $\bar{r}^{(k)} = A \cdot \bar{x}^{(k)} - \bar{b}$, $k = 0, 1, 2, \dots$ sa nazýva **rezíduum (reziduálny vektor)** k -tej iterácie riešenia SLR.

Je to polohový vektor odchýlky od žiadaného nulového vektora. Jeho norma (euklidovská) udáva jeho veľkosť, vzdialenosť, o ktorú je jeho koncový bod vzdialený od 0 (začiatok súradnicovej sústavy).

Vektor chyby numerického riešenia označme $\bar{e}^{(n)} = \bar{x}^p - \bar{x}^{(n)}$. Potom $\bar{x}^p = \bar{x}^{(n)} + \bar{e}^{(n)}$.

Počítajme:

$$\begin{aligned} A \cdot \bar{x}^p &= \bar{b} \\ A \cdot (\bar{x}^{(n)} + \bar{e}^{(n)}) &= \bar{b} \\ A \cdot \bar{x}^{(n)} + A \cdot \bar{e}^{(n)} &= \bar{b} \\ A \cdot \bar{e}^{(n)} &= \bar{b} - A \cdot \bar{x}^{(n)}, \text{ kde } \bar{r}^{(n)} = \bar{b} - A \cdot \bar{x}^{(n)} \\ A \cdot \bar{e}^{(n)} &= \bar{r}^{(n)} \end{aligned} \tag{4.4}$$

$$\bar{e}^{(n)} = A^{-1} \cdot \bar{r}^{(n)} \tag{4.5}$$

V prípade dobre podmienených SLR (pozri kapitolu 1) je rezíduum dobrým reprezentantom chyby numerického riešenia. Ak sa z rovnice (4.4) dá vypočítať vektor chyby $\bar{e}^{(n)}$, potom ho vieme efektívne použiť na spresnenie každého numerického riešenia, vektora $\bar{x}^{(n)}$: $\bar{x}^p \doteq \bar{x}^{(n)} + \bar{e}^{(n)}$.

Porovnaním noriem matíc v rovniciach (4.4) a (4.5) dostaneme dolné a horné ohraničenie pre číselné vyjadrenie chyby numerického riešenia – normy vektora $\bar{e}^{(n)}$:

z rovnice (4.4) dostaneme: $\|A\| \cdot \|\bar{e}^{(n)}\| \geq \|\bar{r}^{(n)}\| \Rightarrow \|\bar{e}^{(n)}\| \geq \frac{\|\bar{r}^{(n)}\|}{\|A\|}$,

z rovnice (4.5) dostaneme $\|\bar{e}^{(n)}\| \leq \|A^{-1}\| \cdot \|\bar{r}^{(n)}\|$,

a teda $\frac{\|\bar{r}^{(n)}\|}{\|A\|} \leq \|\bar{e}^{(n)}\| \leq \|A^{-1}\| \cdot \|\bar{r}^{(n)}\|$.

Zastavovacia podmienka

Zastavovacia podmienka v iteračnom procese je kľúčovým prostriedkom na ovplyvnenie presnosti výsledku. Určujeme ňou, akým kritériom sa má iteračný proces ukončiť, a tým aj akú presnosť dosiahne vypočítaný výsledok. Priamy výpočet vektora chyby jednotlivých aproximácií riešenia SLR nie je jednoduchý – vyžaduje inverznú maticu. Na určenie výstupnej podmienky obyčajne používame tieto zjednodušenia:

- a) rozdiel presného riešenia a ostatnej iterácie $\bar{e}^{(k+1)} = \bar{x}^p - \bar{x}^{(k+1)}$ nahradíme rozdielom ostatných dvoch iterácií $\bar{x}^{(k+1)} - \bar{x}^{(k)}$, absolútnu chybu (normu vektora $\|\bar{e}^{(k+1)}\|$) normou vektora

$$\|\bar{x}^{(k+1)} - \bar{x}^{(k)}\| \text{ a relatívnu chybu podielom } \frac{\|\bar{x}^{(k+1)} - \bar{x}^{(k)}\|}{\|\bar{x}^{(k+1)}\|}.$$

Výstupnú podmienku formulujeme ako $\|\bar{x}^{(k+1)} - \bar{x}^{(k)}\| < \varepsilon$ alebo $\frac{\|\bar{x}^{(k+1)} - \bar{x}^{(k)}\|}{\|\bar{x}^{(k+1)}\|} < \varepsilon$, $\varepsilon > 0$.

Hovoríme, že presné riešenie \bar{x}^p je určené aproximáciou (iteráciou, numerickým riešením) $\bar{x}^{(n)}$, $\bar{x}^{(n)} = \bar{x}^{(k+1)}$, s požadovanou presnosťou (chybou, toleranciou) ε vzhľadom na rozdiel dvoch ostatných iterácií, resp. vzhľadom na podiel uvedený vyššie; $\bar{x}^{(n)} \doteq \bar{x}^p$.

- b) rozdiel presného riešenia a ostatnej iterácie $\bar{e}^{(k+1)} = \bar{x}^p - \bar{x}^{(k+1)}$ nahradíme rezíduom ostatnej iterácie $\bar{r}^{(k+1)} = \bar{b} - A \cdot \bar{x}^{(k+1)}$, resp. $\bar{r}^{(k+1)} = A \cdot \bar{x}^{(k+1)} - \bar{b}$. Absolútnu chybu nahradíme normou rezídua. Výstupná podmienka je tvaru $\|\bar{r}^{(k+1)}\| < \varepsilon$.

Hovoríme, že presné riešenie \bar{x}^p je určené aproximáciou (iteráciou, numerickým riešením) $\bar{x}^{(n)}$, ($\bar{x}^{(n)} = \bar{x}^{(k+1)}$) s požadovanou presnosťou (chybou, toleranciou) ε vzhľadom na rezíduum ostatnej iterácie; $\bar{x}^{(n)} \doteq \bar{x}^p$.

Poznámka. Vo výpočtoch zvyčajne používame euklidovskú normu vektora. Pre $\bar{x} = (x_1, x_2, x_3)^T$ počítame $\|\bar{x}\|_E = \sqrt{x_1^2 + x_2^2 + x_3^2} = \sqrt{\bar{x} \cdot \bar{x}^T}$, pričom platí $\|\bar{x}\|_E = |\bar{x}|$.

Algoritmus výpočtu riešenia SLR iteračnými metódami:

1. bez zastavovacej podmienky s vopred stanoveným počtom iterácií *pit*:

a) určíme vhodný štartovací vektor $\bar{x}^{(0)}$ a vhodný počet iterácií *pit*,

b) pre $k = 0$ až *pit*, s krokom 1 počítame:

$$\bar{x}^{(k+1)} = H \cdot \bar{x}^{(k)} + \bar{g},$$

c) výsledkom je $\bar{x}^{(n)} = \bar{x}^{(k+1)}$, $\bar{x}^{(n)} \doteq \bar{x}^p$.

2. so zastavovacou podmienkou:

a) určíme vhodný štartovací vektor $x^{(0)}$ a vhodný počet iterácií pit ,

b) pre $k = 0$ až pit , s krokom 1 počítame:

i) $\bar{x}^{(k+1)} = H \cdot \bar{x}^{(k)} + \bar{g}$,

ii) pomocné výpočty k zastavovacej podmienke,

iii) zisťujeme platnosť danej zastavovacej podmienky:

$$\left(\|\bar{x}^{(k+1)} - \bar{x}^{(k)}\| < \varepsilon \text{ alebo } \frac{\|\bar{x}^{(k+1)} - \bar{x}^{(k)}\|}{\|\bar{x}^{(k+1)}\|} < \varepsilon \text{ alebo } \|\bar{r}^{(k+1)}\| < \varepsilon \right), \text{ ak:}$$

- podmienka neplatí, pokračujeme výpočtom ďalšej iterácie,
- podmienka platí, výpočet je ukončený;

c) výsledkom je $\bar{x}^{(n)} = \bar{x}^{(k+1)}$, $\bar{x}^{(n)} \doteq \bar{x}^p$.

V prípade, že iteračný výpočet nie je ukončený podmienkou, ale počtom zvolených iterácií, počet iterácií navýšime a pokračujeme vo výpočtoch, až pokiaľ nie je zastavovacia podmienka splnená.

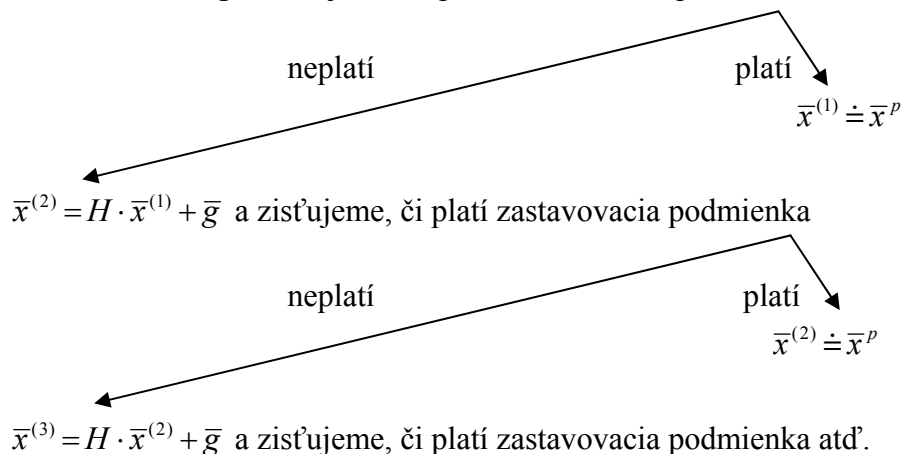
Poznámka. Iteračný proces môžeme vykonávať aj spôsobom: Opakuj, kým platí

Schéma podmieneného iteračného procesu:

a) zvolíme (ľubovoľne) štartovací vektor $\bar{x}^{(0)}$,

b) počítame:

$\bar{x}^{(1)} = H \cdot \bar{x}^{(0)} + \bar{g}$ a zisťujeme, či platí zastavovacia podmienka



Existuje viacero spôsobov prevedenia systému $A \cdot \bar{x} = \bar{b}$ na systém $\bar{x} = H \cdot \bar{x} + \bar{g}$, ale nie každý tvar $\bar{x} = H \cdot \bar{x} + \bar{g}$ konverguje k hľadanému riešeniu \bar{x}^p . Odlišnosti sú v iteračnej matici H a vektore \bar{g} . Napríklad systém:

$$\begin{aligned} 2x + y &= 2 \\ x + 4y + z &= 0 \\ y + 2z &= -2 \end{aligned}$$

môžeme previesť na tvar $\bar{x} = H \cdot \bar{x} + \bar{g}$ takto:

$$\begin{aligned} x &= 0x - \frac{1}{2}y + 0z + 1 & x &= -x - y + 0y + 2 \\ y &= -\frac{1}{4}x + 0y - \frac{1}{4}z + 0 & \text{ale aj} & y &= -x - 3y - z + 0 \\ z &= 0x - \frac{1}{2}y + 0z - 1 & z &= 0x - y - z - 2 \end{aligned}$$

My sa budeme zaoberať 2 numerickými metódami na riešenie SLR, a to Jacobiho metódou a Gauss-Seidlovou metódou.

4.2.1 JACOBIHO METÓDA

Opäť uvažujme o systéme 3 lineárnych rovníc o 3 neznámych:

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

kde $a_{ii} \neq 0$.

Upravme systém $A \cdot \bar{x} = \bar{b}$ ekvivalentnými úpravami na systém $\bar{x} = H \cdot \bar{x} + \bar{g}$ takto:

- z 1. rovnice vyjadríme x_1 za predpokladu, že $a_{11} \neq 0$,
- z 2. rovnice vyjadríme x_2 za predpokladu, že $a_{22} \neq 0$,
- z 3. rovnice vyjadríme x_3 za predpokladu, že $a_{33} \neq 0$.

Dostávame:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \quad / a_{11} \neq 0 & x_1 &= \frac{b_1 - a_{12}x_2 - a_{13}x_3}{a_{11}} \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \quad / a_{22} \neq 0 & \Rightarrow x_2 &= \frac{b_2 - a_{21}x_1 - a_{23}x_3}{a_{22}} \quad \Rightarrow \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \quad / a_{33} \neq 0 & x_3 &= \frac{b_3 - a_{31}x_1 - a_{32}x_2}{a_{33}} \end{aligned}$$

$$\begin{aligned} x_1 &= 0x_1 - \frac{a_{12}}{a_{11}}x_2 - \frac{a_{13}}{a_{11}}x_3 + \frac{b_1}{a_{11}} \\ x_2 &= -\frac{a_{21}}{a_{22}}x_1 + 0x_2 - \frac{a_{23}}{a_{22}}x_3 + \frac{b_2}{a_{22}} \\ x_3 &= -\frac{a_{31}}{a_{33}}x_1 - \frac{a_{32}}{a_{33}}x_2 + 0x_3 + \frac{b_3}{a_{33}} \end{aligned} \quad \Rightarrow \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 & -\frac{a_{12}}{a_{11}} & -\frac{a_{13}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & -\frac{a_{23}}{a_{22}} \\ -\frac{a_{31}}{a_{33}} & -\frac{a_{32}}{a_{33}} & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} \frac{b_1}{a_{11}} \\ \frac{b_2}{a_{22}} \\ \frac{b_3}{a_{33}} \end{pmatrix}$$

Iteračnú formulu môžeme jednoznačne zapísať v maticovom tvare:

$$\bar{x}^{(k+1)} = H_J \cdot \bar{x}^{(k)} + \bar{g}_J, \quad k = 0, 1, 2, \dots$$

$$\begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ x_3^{(k+1)} \end{pmatrix} = \begin{pmatrix} 0 & -\frac{a_{12}}{a_{11}} & -\frac{a_{13}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & -\frac{a_{23}}{a_{22}} \\ -\frac{a_{31}}{a_{33}} & -\frac{a_{32}}{a_{33}} & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ x_3^{(k)} \end{pmatrix} + \begin{pmatrix} \frac{b_1}{a_{11}} \\ \frac{b_2}{a_{22}} \\ \frac{b_3}{a_{33}} \end{pmatrix},$$

resp. v tvare rovníc:

$$\begin{aligned} x_1^{(k+1)} &= 0 x_1 - \frac{a_{12}}{a_{11}} x_2^{(k)} - \frac{a_{13}}{a_{11}} x_3^{(k)} + \frac{b_1}{a_{11}} \\ x_2^{(k+1)} &= -\frac{a_{21}}{a_{22}} x_1^{(k)} + 0 x_2 - \frac{a_{23}}{a_{22}} x_3^{(k)} + \frac{b_2}{a_{22}} \\ x_3^{(k+1)} &= -\frac{a_{31}}{a_{33}} x_1^{(k)} - \frac{a_{32}}{a_{33}} x_2^{(k)} + 0 x_3^{(k)} + \frac{b_3}{a_{33}}, \end{aligned}$$

kde iteračná matica $H_J = \begin{pmatrix} 0 & -\frac{a_{12}}{a_{11}} & -\frac{a_{13}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & -\frac{a_{23}}{a_{22}} \\ -\frac{a_{31}}{a_{33}} & -\frac{a_{32}}{a_{33}} & 0 \end{pmatrix} = \begin{pmatrix} 0 & h_{12} & h_{13} \\ h_{21} & 0 & h_{23} \\ h_{31} & h_{32} & 0 \end{pmatrix}$ a $\bar{g}_J = \begin{pmatrix} \frac{b_1}{a_{11}} \\ \frac{b_2}{a_{22}} \\ \frac{b_3}{a_{33}} \end{pmatrix}$.

Všimnite si! Iteračná matica H_J pre Jacobiho metódu je špecifická tým, že má na hlavnej diagonále samé nuly – prvky h_{ii} sa musia rovnať nule. Prvky matice H_J , ktoré neležia na hlavnej diagonále, môžu aj nemusia sa rovnať nule.

Uvažovaný systém zovšeobecňujeme na systém n lineárnych rovníc o n neznámych. Z i -tej rovnice vyjadríme x_i ($a_{ii} \neq 0$), dostaneme

$$\begin{aligned} x_i &= \frac{1}{a_{ii}} (b_i - (a_{i1}x_1 + a_{i2}x_2 + \dots + a_{i(i-1)}x_{i-1} + a_{i(i+1)}x_{i+1} + \dots + a_{in}x_n)), \\ x_i^{(k+1)} &= \frac{1}{a_{ii}} (b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)}), \quad i = 1, 2, \dots, n, \quad k = 0, 1, \dots \end{aligned}$$

Výpočet práve počítanej súradnice $x_i^{(k+1)}$ prostredníctvom ostatných súradníc iterácie $x^{(k)}$ symbolicky vyjadríme takto:

$$x_i^{(k+1)} = q(x_1^{(k)}, x_2^{(k)}, \dots, x_{i-1}^{(k)}, x_{i+1}^{(k)}, \dots, x_n^{(k)}).$$

Ak matica systému $A = L_o + D + U_o$, kde L_o , D a U_o sú matice:

$$L_o = \begin{pmatrix} 0 & 0 & 0 \\ a_{21} & 0 & 0 \\ a_{31} & a_{32} & 0 \end{pmatrix} \quad D = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix} \quad U_o = \begin{pmatrix} 0 & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & 0 \end{pmatrix}, \quad (4.6)$$

ostro dolná trojuholníková diagonálna ostro horná trojuholníková

potom po prevedení systému $A \cdot \bar{x} = \bar{b}$ (ekvivalentnými úpravami) na systém $\bar{x} = H_J \cdot \bar{x} + \bar{g}_J$, dostávame:

$$\begin{aligned} A \cdot \bar{x} &= \bar{b} & \Rightarrow & \bar{x} = H_J \cdot \bar{x} + \bar{g}_J \\ (L_o + D + U_o) \cdot \bar{x} &= \bar{b} \\ D \cdot \bar{x} + (L_o + U_o) \cdot \bar{x} &= \bar{b} \\ D \cdot \bar{x} &= -(L_o + U_o) \cdot \bar{x} + \bar{b} \\ \bar{x} &= -D^{-1} \cdot (L_o + U_o) \cdot \bar{x} + D^{-1} \cdot \bar{b} \Rightarrow H_J = -D^{-1} \cdot (L_o + U_o) \\ & & \bar{g}_J &= D^{-1} \cdot \bar{b} \end{aligned}$$

Poznámka. Ak $D = D_{3 \times 3}$ je diagonálna matica, ktorej prvky hlavnej diagonály $d_{ii} \neq 0$, potom inverzná

$$\text{matica } D^{-1} = \begin{pmatrix} \frac{1}{d_{11}} & 0 & 0 \\ 0 & \frac{1}{d_{22}} & 0 \\ 0 & 0 & \frac{1}{d_{33}} \end{pmatrix}.$$

4.2.2 GAUSS-SEIDLOVA METÓDA

Podobne ako u Jacobiho metóde uvažujme o systéme 3 lineárnych rovníc o 3 neznámych:

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

Upravme systém $A \cdot \bar{x} = \bar{b}$ ekvivalentnými úpravami na systém $\bar{x} = H \cdot \bar{x} + \bar{g}$.

Za predpokladu, že $a_{ii} \neq 0$, vyjadrime x_i z i -tej rovnice, $i = 1, 2, 3$, dostávame:

$$\begin{aligned} x_1 &= 0 x_1 - \frac{a_{12}}{a_{11}} x_2 - \frac{a_{13}}{a_{11}} x_3 + \frac{b_1}{a_{11}} \\ x_2 &= -\frac{a_{21}}{a_{22}} x_1 + 0 x_2 - \frac{a_{23}}{a_{22}} x_3 + \frac{b_2}{a_{22}} \\ x_3 &= -\frac{a_{31}}{a_{33}} x_1 - \frac{a_{32}}{a_{33}} x_2 + 0 x_3 + \frac{b_3}{a_{33}} \end{aligned}$$

Gauss-Seidlova metóda je podobná Jacobiho metóde, avšak na výpočet iterácií jednotlivých súradníc vektora riešení používa predchádzajúce súradnice ihneď po ich vypočítaní, t. j. pri výpočte $(k+1)$ -vej iterácie i -tej súradnice $x_i^{(k+1)}$ používa všetky predchádzajúce súradnice z práve počítanej iterácie $(k+1)$: $x_1^{(k+1)}, x_2^{(k+1)}, \dots, x_{i-1}^{(k+1)}$. Ostatné súradnice, ktoré vystupujú vo výpočte, sú z predchádzajúcej iterácie (k) : $x_{i+1}^{(k)}, x_{i+2}^{(k)}, \dots, x_n^{(k)}$.

$$\begin{aligned} x_1^{(k+1)} &= 0 x_1 - \frac{a_{12}}{a_{11}} x_2^{(k)} - \frac{a_{13}}{a_{11}} x_3^{(k)} + \frac{b_1}{a_{11}} \\ x_2^{(k+1)} &= -\frac{a_{21}}{a_{22}} x_1^{(k+1)} + 0 x_2 - \frac{a_{23}}{a_{22}} x_3^{(k)} + \frac{b_2}{a_{22}} \\ x_3^{(k+1)} &= -\frac{a_{31}}{a_{33}} x_1^{(k+1)} - \frac{a_{32}}{a_{33}} x_2^{(k+1)} + 0 x_3 + \frac{b_3}{a_{33}} \end{aligned} \quad (4.7)$$

V (4.7) vidíme, že napr. $x_2^{(k+1)}$ počítame pomocou $x_1^{(k+1)}$ a $x_3^{(k)}$.

Uvažovaný systém opäť zovšeobecňujeme na systém n lineárnych rovníc o n neznámych. Z i -tej rovnice vyjadríme x_i ($a_{ii} \neq 0$), dostaneme:

$$\begin{aligned} x_i &= \frac{1}{a_{ii}} \cdot (b_i - (a_{i1}x_1 + a_{i2}x_2 + \dots + a_{i(i-1)}x_{i-1} + a_{i(i+1)}x_{i+1} + \dots + a_{in}x_n)) \\ x_i^{(k+1)} &= \frac{1}{a_{ii}} \cdot (b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)}), \quad i = 1, 2, \dots, n, \quad k = 0, 1, \dots \end{aligned}$$

Symbolicky zápis výpočtu práve počítanej súradnice $x_i^{(k+1)}$ pomocou ostatných súradníc iterácie $\bar{x}^{(k+1)}$ vyzerá takto:

$$x_i^{(k+1)} = q(x_1^{(k+1)}, x_2^{(k+1)}, \dots, x_{i-1}^{(k+1)}, x_{i+1}^{(k)}, \dots, x_n^{(k)}).$$

V tejto súvislosti je potrebné si uvedomiť, že $x_1^{(k+1)}, x_2^{(k+1)}, \dots, x_{i-1}^{(k+1)}$ sú súradnice iterácie $\bar{x}^{(k+1)}$ nachádzajúce sa pred súradnicou $x_i^{(k+1)}$ a $x_{i+1}^{(k)}, x_{i+2}^{(k)}, \dots, x_n^{(k)}$ sú súradnice iterácie $\bar{x}^{(k)}$ nachádzajúce sa za súradnicou $x_i^{(k)}$.

Ak chceme iteračnú formulu pre Gauss-Seidlovu metódu tiež zapísať v tvare $\bar{x}^{(k+1)} = H \cdot \bar{x}^{(k)} + \bar{g}$, $k = 0, 1, 2, \dots$, potom postupujeme takto.

Vychádzajme zo zjednodušeného tvaru systému (4.7):

$$x_1^{(k+1)} = 0x_1 + B_1x_2^{(k)} + C_1x_3^{(k)} + D_1 \quad (1)$$

$$x_2^{(k+1)} = A_2x_1^{(k+1)} + 0x_2 + C_2x_3^{(k)} + D_2 \quad (2)$$

$$x_3^{(k+1)} = A_3x_1^{(k+1)} + B_3x_2^{(k+1)} + 0x_3 + D_3 \quad (3)$$

Dosaďme $x_1^{(k+1)}$ z rovnice (1) do rovnice (2) a po úprave dostávame:

$$x_1^{(k+1)} = 0x_1 + B_1x_2^{(k)} + C_1x_3^{(k)} + D_1 \quad (1)$$

$$x_2^{(k+1)} = 0x_1 + \tilde{B}_2x_2^{(k)} + \tilde{C}_2x_3^{(k)} + \tilde{D}_2 \quad (2)$$

$$x_3^{(k+1)} = A_3x_1^{(k+1)} + B_3x_2^{(k+1)} + 0x_3 + D_3 \quad (3)$$

Do rovnice (3) dosadíme $x_1^{(k+1)}$ z rovnice (1) a $x_2^{(k+1)}$ z rovnice (2). Po úprave dostávame:

$$x_1^{(k+1)} = 0x_1 + B_1x_2^{(k)} + C_1x_3^{(k)} + D_1 \quad (1)$$

$$x_2^{(k+1)} = 0x_1 + \tilde{B}_2x_2^{(k)} + \tilde{C}_2x_3^{(k)} + \tilde{D}_2 \quad (2)$$

$$x_3^{(k+1)} = 0x_1 + \tilde{B}_3x_2^{(k)} + \tilde{C}_3x_3^{(k)} + \tilde{D}_3 \quad (3)$$

v maticovom tvare:

$$\bar{x}^{(k+1)} = H_S \cdot \bar{x}^{(k)} + \bar{g}_S, \quad k = 0, 1, 2, \dots,$$

kde iteračná matica $H_S = \begin{pmatrix} 0 & h_{12} & h_{13} \\ 0 & h_{22} & h_{23} \\ 0 & h_{32} & h_{33} \end{pmatrix}$ a $\bar{g}_S = \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix}$.

Všimnite si! Iteračná matica H_S pre Gauss-Seidlovu metódu je špecifická tým, že má v 1. stĺpci nuly, prvky a_{i1} , $i = 1, 2, 3$ sa musia rovnať nule. Prvky matice H_S , ktoré neležia v 1. stĺpci, môžu aj nemusia sa rovnať nule.

Nech matica systému $A_{3 \times 3} = L_0 + D + U_0$, kde L_0 , D a U_0 , sú matice (4.6), potom po prevedení systému $A \cdot \bar{x} = \bar{b}$ ekvivalentnými úpravami na systém $\bar{x} = H_S \cdot \bar{x} + \bar{g}_S$, dostávame:

$$\begin{aligned} A \cdot \bar{x} &= \bar{b} & \Rightarrow \bar{x} &= H_S \cdot \bar{x} + \bar{g}_S \\ (L_0 + D + U_0) \cdot \bar{x} &= \bar{b} \\ (L_0 + D) \cdot \bar{x} + U_0 \cdot \bar{x} &= \bar{b} \\ (L_0 + D) \cdot \bar{x} &= -U_0 \cdot \bar{x} + \bar{b} \\ \bar{x} &= -(L_0 + D)^{-1} \cdot U_0 \cdot \bar{x} + (L_0 + D)^{-1} \cdot \bar{b} \Rightarrow H_S = -(L_0 + D)^{-1} \cdot U_0 \\ & & \bar{g}_S &= (L_0 + D)^{-1} \cdot \bar{b} \end{aligned}$$

Podmienky konvergencie iteračných metód

Veta 4.2 Iteračný proces $\bar{x}^{(k+1)} = H \cdot \bar{x}^{(k)} + \bar{g}$ konverguje k riešeniu systému $A \cdot \bar{x} = \bar{b}$ pre ľubovoľný štartovací vektor $\bar{x}^{(0)}$, keď platí:

podmienka	metóda		poznámka
	Jacobiho	Gauss-Seidlova	
nutná a postačujúca	$\rho(H) < 1$	$\rho(H) < 1$	$\rho(H) = \max\{ \lambda_1 , \lambda_2 , \dots, \lambda_n \}$ $\rho(H)$ je spektrálny polomer H λ_i sú vlastné čísla H
postačujúce	$\ H\ < 1$	$\ H\ < 1$	$\ H\ $ je ľubovoľná norma H
	A je ODD	A je ODD	ODD – ostro diagonálne dominantná
		A je SPD	SPD – symetrická a pozitívne definitná

Poznámka

- Podmienky kladené na maticu A sú postačujúce, t. j. ak Jacobiho metóda konverguje, matica A nemusí byť ODD. Podobne, ak Gauss-Seidlova metóda konverguje, matica A nemusí byť ani ODD, ani SPD.
- Ak matica A je SPD a Jacobiho metóda konverguje, potom Gauss-Seidlova metóda konverguje 2-krát rýchlejšie ako Jacobiho metóda.
- Existujú systémy, pre ktoré Jacobiho metóda konverguje a Gauss-Seidlova metóda diverguje a naopak, napr.

$$\begin{array}{rcl} x - 0,464y & & = 0,536 \\ 2,047x + & y - 0,464z & = 2,583 \\ -0,464x & & + z = 0,536 \end{array}$$

Nutná a postačujúca podmienka konvergenencie sa vo všeobecnosti overuje veľmi ťažko, najmä ak matica H je veľkých rozmerov. Narážame tu na problém výpočtu vlastných čísel matice H (kap.3). Preto v nešpecifických prípadoch využívame možnosť rozhodnúť o konvergencii metód na základe vlastností matice A (ODD, SPD). Pozitívnu definitnosť neoverujeme z definície, ale využívame Sylvestrovo kritérium (pozri kap.3, veta 3.1) alebo môžeme použiť tiež nasledujúcu vetu, ktorá nám síce pomôže zaručiť konvergenciu Gauss-Seidlovej metódy, avšak v takomto prípade môže metóda konvergovať pomaly.

Veta 4.3 Nech $A = A_{n \times n}$ je regulárna matica ($|A| \neq 0$) a nech A^T je k nej transponovaná matica, potom $A^T \cdot A$ je symetrická a pozitívne definitná matica.

Znamená to, že miesto systému $A \cdot \bar{x} = \bar{b}$ budeme riešiť systém $A^T \cdot A \cdot \bar{x} = A^T \cdot \bar{b}$, ktorý ale konverguje k riešeniu pomalšie ako pôvodný systém $A \cdot \bar{x} = \bar{b}$.

V prípade konvergenencie oboch metód Gauss-Seidlova metóda konverguje spravidla rýchlejšie ako Jacobiho metóda, avšak výhodou Jacobiho metódy je, že môže výpočty realizovať paralelne, pričom Gauss-Seidlova metóda je vďaka okamžitému použitiu vypočítaných súradníc odkázaná na simultánny výpočet.

Niekoľko poznámok na záver

- Iteračné metódy patria medzi tie metódy riešenia systémov lineárnych rovníc (SLR), ktorými nehľadáme presné riešenie, ale riešenie s istou, vopred danou presnosťou.
- Na začiatku zvolíme začiatočnú aproximáciu riešenia $\bar{x}^{(0)}$, ktorú iteračným procesom spresňujeme. Zaokrúhľovacie chyby sa iteračným procesom samé opravujú, t. j. chyba vo výpočte jednotlivých iterácií sa na konečnom výsledku neprejaví. Na konečnom výsledku sa prejaví len zaokrúhľovacia chyba poslednej iterácie. Je to spôsobené tým, že chybou zaťažená aproximácia vystupuje v ďalšej iterácii ako nová začiatočná iterácia. Na rozdiel od priamych metód riešenia SLR sú iteračné metódy numericky stabilné.

- Štartovací vektor $\bar{x}^{(0)}$ môžeme voliť ľubovoľne. Je však vhodné voliť ho ako nejaké približné riešenie SLR, napr. veľmi hrubý odhad riešenia. Často sa volí $\bar{x}^{(0)} = \bar{g}$.
- Výpočet iterácií zvyčajne ukončujeme zastavovacou podmienkou, napr. keď je norma rozdielu 2 po sebe idúcich iterácií vektora riešenia dostatočne malá (napr. $\varepsilon < 10^{-6}$).
- Iteračné metódy môžu aj divergovať. Konvergencia, resp. divergencia iteračnej metódy závisí od vlastností matice systému A .
- Iteračné metódy sú vhodné na riešenie veľkých sústav s riedkou maticou systému. Pri veľkých sústavách Gaussova eliminačná metóda (GEM) zlyháva – presné riešenie značne ovplyvňujú zaokrúhľovacie chyby. (Ak matica systému je rádu n , pri GEM je potrebné vykonať rádovo n^3 aritmetických operácií.)
- Výhodné je, keď diagonálne prvky matice systému A sú veľké v porovnaní s ostatnými ne-diagonálnymi prvkami (v absolútnej hodnote). Na prvkoch pravých strán nezáleží. Čím je $|a_{ii}|$ väčšia ako $\sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|$, $i = 1, 2, \dots$, tým metóda rýchlejšie konverguje.

4.2.3 RIEŠENÉ PRÍKLADY

PRÍKLAD č. 1

Riešte systém lineárnych rovníc:

$$\begin{aligned} x + y - 3z &= -4 \\ 2x + 5y + z &= 5 \\ 4x - y + 2z &= -12 \end{aligned} \quad (4.8)$$

1. Jacobiho metódou.

- a) Ak je potrebné, upravte systém (ekvivalentnými úpravami) tak, aby bola zaručená konvergencia metódy. Svoju úpravu zdôvodnite.
- b) Nájdite riešenie systému touto metódou pre štartovací vektor $\bar{x}^{(0)} = (0, 0, 0)^T$. Počítajte s toleranciou $\|\bar{x}^{(k+1)} - \bar{x}^{(k)}\| < \varepsilon$, $\varepsilon = 0,5$. V zastavovacej podmienke voľte euklidovskú normu vektora.
- c) Pre každú iteráciu vypočítajte rezíduum $\bar{r}^{(i)}$ a jeho normu.
- d) Vypočítajte presnosť poslednej iterácie riešenia. (Nájdite hodnotu (euklidovskú normu), ktorou sa líši posledná iterácia riešenia od predposlednej iterácie riešenia.)

2. Gauss-Seidlovou metódou.

- a) Ak je potrebné, upravte systém (ekvivalentnými úpravami) tak, aby bola zaručená konvergencia metódy. Svoju úpravu zdôvodnite.

- b) Nájdite riešenie systému touto metódou pre štartovací vektor $\bar{x}^{(0)} = (0, 0, 0)^T$. Počítajte s toleranciou $\|\bar{x}^{(k+1)} - \bar{x}^{(k)}\| < \varepsilon, \varepsilon = 0,5$. V zastavovacej podmienke voľte euklidovskú normu vektora.
- c) Pre každú iteráciu vypočítajte rezíduum $\bar{r}^{(i)}$ a jeho normu.
- d) Vyjadrite presnosť poslednej iterácie riešenia (vyjadrite hodnotu euklidovskej normy vektora, ktorým sa líši posledná iterácia riešenia od predposlednej iterácie riešenia.)

RIEŠENIE

SLR prepíšeme na tvar $A \cdot \bar{x} = \bar{b}$, kde $A = \begin{pmatrix} 1 & 1 & -3 \\ 2 & 5 & 1 \\ 4 & -1 & 2 \end{pmatrix}$, $\bar{b} = \begin{pmatrix} -4 \\ 5 \\ -12 \end{pmatrix}$.

Overíme, či má systém práve jedno riešenie: $|A| \neq 0$.

$|A|=77$, systém má práve jedno riešenie, na výpočet môžeme použiť numerické metódy.

1. Jacobiho metóda

- a) *Overme postačujúce podmienky konverencie Jacobiho metódy:*

Je A ostro diagonálne dominantná matica?

$$A = \begin{pmatrix} 1 & 1 & -3 \\ 2 & 5 & 1 \\ 4 & -1 & 2 \end{pmatrix} \quad \begin{array}{l} |1| > |1| + |-3| \quad \dots\dots \text{neplatí} \\ |5| > |1| + |1| \quad \dots\dots \text{platí} \\ |2| > |4| + |-1| \quad \dots\dots \text{neplatí} \end{array}$$

Matica systému A nie je ostro diagonálne dominantná.

Aby sme zaručili konverenciu metódy, skúsme upraviť maticu A ekvivalentnými úpravami na ostro diagonálne dominantnú maticu. Výmenou 1. a 3. rovnice v systéme dostávame:

$$A \cdot \bar{x} = \bar{b} \Rightarrow \begin{pmatrix} 4 & -1 & 2 \\ 2 & 5 & 1 \\ 1 & 1 & -3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -12 \\ 5 \\ -4 \end{pmatrix}$$

Pre maticu systému teraz platí:

$$\begin{array}{l} |4| > |-1| + |2| \\ |5| > |2| + |1| \\ |-3| > |1| + |1| \end{array}$$

Po úprave matica A už ostro diagonálne dominantná je a iteračný proces Jacobiho (tiež aj Gauss-Seidlovej) metódy bude konvergovať.

- b) *Nájdeme riešenie.*

Systém upravíme na iteračný tvar Jacobiho metódy $\bar{x}^{(k+1)} = H_J \cdot \bar{x}^{(k)} + \bar{g}_J$.

Iteračnú maticu H_J a vektor \bar{g}_J môžeme dostať buď úpravou rovníc:

$$4x - y + 2z = -12 \Rightarrow x = \frac{1}{4}y - \frac{1}{2}z - 3$$

$$2x + 5y + z = 5 \Rightarrow y = -\frac{2}{5}x - \frac{1}{5}z + 1$$

$$x + y - 3z = -4 \Rightarrow z = \frac{1}{3}x + \frac{1}{3}y + \frac{4}{3}$$

$$H_J = \begin{pmatrix} 0 & \frac{1}{4} & -\frac{1}{2} \\ -\frac{2}{5} & 0 & -\frac{1}{5} \\ \frac{1}{3} & \frac{1}{3} & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0,25 & -0,5 \\ -0,4 & 0 & -0,2 \\ 0,333 & 0,333 & 0 \end{pmatrix}, \quad \bar{g}_J = \begin{pmatrix} -3 \\ 1 \\ \frac{4}{3} \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \\ 1,333 \end{pmatrix},$$

alebo výpočtom pomocou trojuholníkových matic L_o , U_o a D :

$$L_o = \begin{pmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \quad D = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -3 \end{pmatrix} \quad U_o = \begin{pmatrix} 0 & -1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$H_J = -D^{-1} \cdot (L_o + U_o) = - \begin{pmatrix} \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{5} & 0 \\ 0 & 0 & -\frac{1}{3} \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 & 2 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{4} & -\frac{1}{2} \\ -\frac{2}{5} & 0 & -\frac{1}{5} \\ \frac{1}{3} & \frac{1}{3} & 0 \end{pmatrix},$$

$$\bar{g}_J = D^{-1} \cdot \bar{b} = \begin{pmatrix} \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{5} & 0 \\ 0 & 0 & -\frac{1}{3} \end{pmatrix} \cdot \begin{pmatrix} -12 \\ 5 \\ -4 \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \\ \frac{4}{3} \end{pmatrix}.$$

Iteračný proces:

Štartovacím vektorom je $\bar{x}^{(0)} = (0, 0, 0)^T$.

V každom kroku iteračného procesu:

- vykonáme výpočet vektora $\bar{x}^{(k+1)} = H_J \cdot \bar{x}^{(k)} + \bar{g}_J$,
- vypočítame rezíduum $r^{(k+1)}$,
- zistíme, či platí $\|\bar{x}^{(k+1)} - \bar{x}^{(k)}\|_E < 0,5$.

Iteračný proces ukončíme, keď bude splnená zastavovacia podmienka $\|\bar{x}^{(k+1)} - \bar{x}^{(k)}\| < \varepsilon$, $\varepsilon = 0,5$.

Výpočet 1. iterácie riešenia:

$$\bar{x}^{(1)} = H_J \cdot \bar{x}^{(0)} + \bar{g}_J$$

$$\begin{pmatrix} x^{(1)} \\ y^{(1)} \\ z^{(1)} \end{pmatrix} = \begin{pmatrix} 0 & 0,25 & -0,5 \\ -0,4 & 0 & -0,2 \\ 0,333 & 0,333 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -3 \\ 1 \\ 1,333 \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \\ 1,333 \end{pmatrix}$$

Rezíduum: $\bar{r}^{(1)} = \bar{b} - A \cdot \bar{x}^{(1)} = \begin{pmatrix} -12 \\ 5 \\ -4 \end{pmatrix} - \begin{pmatrix} 4 & -1 & 2 \\ 2 & 5 & 1 \\ 1 & 1 & -3 \end{pmatrix} \cdot \begin{pmatrix} -3 \\ 1 \\ 1,333 \end{pmatrix} = \begin{pmatrix} -1,667 \\ 4,667 \\ 2 \end{pmatrix}$

Norma rezídua: $\|r^{(1)}\|_E = \sqrt{(-1,667)^2 + (4,667)^2 + (2)^2} = 5,34413$

Zistíme platnosť zastavovacej podmienky: $\|\bar{x}^{(1)} - \bar{x}^{(0)}\|_E < 0,5$:

$$\|\bar{x}^{(1)} - \bar{x}^{(0)}\|_E = \sqrt{(-3-0)^2 + (1-0)^2 + (1,333-0)^2} = \sqrt{3^2 + 1^2 + 1,333^2} = 3,432 > 0,5 \Rightarrow \text{neplatí}$$

Výpočet 2. iterácie riešenia:

$$\bar{x}^{(2)} = H_J \cdot \bar{x}^{(1)} + \bar{g}_J$$

$$\begin{pmatrix} x^{(2)} \\ y^{(2)} \\ z^{(2)} \end{pmatrix} = \begin{pmatrix} 0 & 0,25 & -0,5 \\ -0,4 & 0 & -0,2 \\ 0,333 & 0,333 & 0 \end{pmatrix} \cdot \begin{pmatrix} -3 \\ 1 \\ 1,333 \end{pmatrix} + \begin{pmatrix} -3 \\ 1 \\ 1,333 \end{pmatrix} = \begin{pmatrix} -3,414 \\ 1,933 \\ 0,667 \end{pmatrix}$$

Rezíduum: $\bar{r}^{(2)} = \bar{b} - A \cdot \bar{x}^{(2)} = \begin{pmatrix} -12 \\ 5 \\ -4 \end{pmatrix} - \begin{pmatrix} 4 & -1 & 2 \\ 2 & 5 & 1 \\ 1 & 1 & -3 \end{pmatrix} \cdot \begin{pmatrix} -3,414 \\ 1,933 \\ 0,667 \end{pmatrix} = \begin{pmatrix} 2,555 \\ 1,499 \\ -0,518 \end{pmatrix}$

Norma rezídua: $\|r^{(2)}\|_E = \sqrt{(2,555)^2 + (1,499)^2 + (-0,518)^2} = 6,00147$

Zistíme platnosť zastavovacej podmienky: $\|\bar{x}^{(2)} - \bar{x}^{(1)}\|_E < 0,5$:

$$\|\bar{x}^{(2)} - \bar{x}^{(1)}\|_E = \sqrt{(-3,414 - (-3))^2 + (1,933 - 1)^2 + (0,667 - 1,333)^2}$$

$$= \sqrt{0,413^2 + 0,933^2 + 0,666^2} = 1,219 > 0,5 \Rightarrow \text{neplatí}$$

Výpočet 3. iterácie riešenia:

$$\bar{x}^{(3)} = H_J \cdot \bar{x}^{(2)} + \bar{g}_J$$

$$\begin{pmatrix} x^{(3)} \\ y^{(3)} \\ z^{(3)} \end{pmatrix} = \begin{pmatrix} 0 & 0,25 & -0,5 \\ -0,4 & 0 & -0,2 \\ 0,333 & 0,333 & 0 \end{pmatrix} \cdot \begin{pmatrix} -3,414 \\ 1,933 \\ 0,667 \end{pmatrix} + \begin{pmatrix} -3 \\ 1 \\ 1,333 \end{pmatrix} = \begin{pmatrix} -2,85 \\ 2,233 \\ 0,838 \end{pmatrix}$$

Rezíduum: $\bar{r}^{(3)} = \bar{b} - A \cdot \bar{x}^{(3)} = \begin{pmatrix} -12 \\ 5 \\ -4 \end{pmatrix} - \begin{pmatrix} 4 & -1 & 2 \\ 2 & 5 & 1 \\ 1 & 1 & -3 \end{pmatrix} \cdot \begin{pmatrix} -2,85 \\ 2,233 \\ 0,838 \end{pmatrix} = \begin{pmatrix} -0,043 \\ -1,303 \\ -0,869 \end{pmatrix}$

Norma rezídua: $\|r^{(3)}\|_E = \sqrt{(-0,043)^2 + (-1,303)^2 + (-0,869)^2} = 1,56679$

Zistíme platnosť zastavovacej podmienky: $\|\bar{x}^{(3)} - \bar{x}^{(2)}\|_E < 0,5$:

$$\begin{aligned}\|\bar{x}^{(3)} - \bar{x}^{(2)}\|_E &= \sqrt{(-2,85 - (-3,416))^2 + (2,233 - 1,933)^2 + (0,838 - 0,667)^2} \\ &= \sqrt{0,565^2 + 0,3^2 + 0,171^2} = 0,663 > 0,5 \Rightarrow \text{neplatí}\end{aligned}$$

Výpočet 4. iterácie riešenia:

$$\bar{x}^{(4)} = H_J \cdot \bar{x}^{(3)} + \bar{g}_J$$

$$\begin{pmatrix} x^{(4)} \\ y^{(4)} \\ z^{(4)} \end{pmatrix} = \begin{pmatrix} 0 & 0,25 & -0,5 \\ -0,4 & 0 & -0,2 \\ 0,333 & 0,333 & 0 \end{pmatrix} \cdot \begin{pmatrix} -2,85 \\ 2,233 \\ 0,838 \end{pmatrix} + \begin{pmatrix} -3 \\ 1 \\ 1,333 \end{pmatrix} = \begin{pmatrix} -2,861 \\ 1,972 \\ 1,129 \end{pmatrix}$$

$$\text{Rezíduum: } \bar{r}^{(4)} = \bar{b} - A \cdot \bar{x}^{(4)} = \begin{pmatrix} -12 \\ 5 \\ -4 \end{pmatrix} - \begin{pmatrix} 4 & -1 & 2 \\ 2 & 5 & 1 \\ 1 & 1 & -3 \end{pmatrix} \cdot \begin{pmatrix} -2,861 \\ 1,972 \\ 1,129 \end{pmatrix} = \begin{pmatrix} -0,842 \\ -0,267 \\ 0,276 \end{pmatrix}$$

$$\text{Norma rezídua: } \|\bar{r}^{(4)}\|_E = \sqrt{(-0,842)^2 + (-0,267)^2 + (0,276)^2} = 1,56679$$

Zistíme platnosť zastavovacej podmienky: $\|\bar{x}^{(4)} - \bar{x}^{(3)}\|_E < 0,5$:

$$\begin{aligned}\|\bar{x}^{(4)} - \bar{x}^{(3)}\|_E &= \sqrt{(2,861 - (-2,85))^2 + (1,972 - 2,233)^2 + (1,129 - 0,838)^2} \\ &= \sqrt{(-0,011)^2 + 0,261^2 + 0,291^2} = 0,3911 \Rightarrow \text{platí}\end{aligned}$$

c) Presnosť poslednej iterácie riešenia je euklidovská norma vektora:

$$\|\bar{x}^{(4)} - \bar{x}^{(3)}\|_E = 0,3911$$

Záver

Výsledky zapíšeme do tabuľky (Tabuľka 4.1):

k	$\bar{x}^{(k)}$	$\ \bar{x}^{(k+1)} - \bar{x}^{(k)}\ _E$	$\bar{r}^{(k)} = \bar{b} - A \cdot \bar{x}^{(k)}$	$\ \bar{r}^{(k)}\ _E$
0	$\bar{x}^{(0)} = (0, 0, 0)^T$	-	$\bar{r}^{(0)} = (-12, 5, -4)^T$	13,6015
1	$\bar{x}^{(1)} = (-3; 1; 1,33)^T$	3,432	$\bar{r}^{(1)} = (-1,667; 4,667; 2)^T$	5,34413
2	$\bar{x}^{(2)} = (-3,414; 1,933; 0,667)^T$	1,219	$\bar{r}^{(2)} = (-0,01; 5,495; 2,413)^T$	6,00147
3	$\bar{x}^{(3)} = (-2,85; 2,233; 0,838)^T$	0,663	$\bar{r}^{(3)} = (-0,043; -1,303; -0,869)^T$	1,56679
4	$\bar{x}^{(4)} = (-2,861; 1,972; 1,129)^T$	0,3911	$\bar{r}^{(4)} = (-0,842; -0,267; 0,276)^T$	0,83548

Poznámka. Presné riešenie systému je $\bar{x}^p = (-3; 2; 1)^T$.

Tabuľka 4.1. Riešenie systému (4.8) Jacobiho metódou

Výpočet sme ukončili po 4. iterácii, pretože $\|\bar{x}^{(4)} - \bar{x}^{(3)}\|_E = 0,3911 < 0,5$. Hľadaným numerickým riešením s danou toleranciou je 4. iterácia riešenia $\bar{x}^{(4)} = (-2,861; 1,972; 1,129)^T$.

2. Gauss-Seidlova metóda

a) Overme postačujúce podmienky konverencie Gauss-Seidlovej metódy:

Matica A nie je symetrická \Rightarrow nemusíme overovať, či je aj pozitívne definitná. Aby sme zaručili konvergenciu metódy, upravíme maticu A ekvivalentnými úpravami na ostro diagonálne dominantnú maticu (pozri riešenie Jacobiho metóda). Po úprave dostaneme systém:

$$A \cdot \bar{x} = \bar{b} \quad \Rightarrow \quad \begin{pmatrix} 4 & -1 & 2 \\ 2 & 5 & 1 \\ 1 & 1 & -3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -12 \\ 5 \\ -4 \end{pmatrix}$$

b) Nájdeime riešenie.

Systém upravíme na iteračný tvar Gauss-Seidlovej metódy $\bar{x}^{(k+1)} = H_S \cdot \bar{x}^{(k)} + \bar{g}_S$:

$$4x - y + 2z = -12 \quad \Rightarrow \quad x^{(k+1)} = \frac{1}{4}y^{(k)} - \frac{1}{2}z^{(k)} - 3 \quad (R1)$$

$$2x + 5y + z = 5 \quad \Rightarrow \quad y^{(k+1)} = -\frac{2}{5}x^{(k+1)} - \frac{1}{5}z^{(k)} + 1 \quad (R2)$$

$$x + y - 3z = -4 \quad \Rightarrow \quad z^{(k+1)} = \frac{1}{3}x^{(k+1)} + \frac{1}{3}y^{(k+1)} + \frac{4}{3} \quad (R3)$$

Dosaďme do rovnice (R2) za $\bar{x}^{(k+1)}$ z rovnice (R1) a do rovnice (R3) za $\bar{x}^{(k+1)}$ a $\bar{y}^{(k+1)}$

z rovníc (R1) a (R2), po úprave dostávame hľadanú iteračnú schému $\bar{x}^{(k+1)} = H_S \cdot \bar{x}^{(k)} + \bar{g}_S$

$$x^{(k+1)} = \frac{1}{4}y^{(k+1)} - \frac{1}{2}z^{(k+1)} - 3$$

$$y^{(k+1)} = -\frac{1}{10}y^{(k+1)} + 0z^{(k+1)} + \frac{11}{5}$$

$$z^{(k+1)} = \frac{1}{20}y^{(k+1)} - \frac{1}{6}z^{(k+1)} + \frac{16}{15}$$

$$H_S = \begin{pmatrix} 0 & \frac{1}{4} & -\frac{1}{2} \\ 0 & -\frac{1}{10} & 0 \\ 0 & \frac{1}{20} & -\frac{1}{6} \end{pmatrix} = \begin{pmatrix} 0 & 0,25 & -0,5 \\ 0 & -0,1 & 0 \\ 0 & 0,05 & -0,167 \end{pmatrix}, \quad \bar{g}_S = \begin{pmatrix} -3 \\ \frac{11}{5} \\ \frac{16}{15} \end{pmatrix} = \begin{pmatrix} -3 \\ 2,2 \\ 1,067 \end{pmatrix}$$

Iteračný proces:

Štartovacím vektorom je $\bar{x}^{(0)} = (0, 0, 0)^T$.

V každom kroku iteračného procesu:

- vykonáme výpočet vektora $\bar{x}^{(k+1)} = H_S \cdot \bar{x}^{(k)} + \bar{g}_S$,
- vypočítame rezíduum $r^{(k+1)}$,
- zistíme, či platí $\|\bar{x}^{(k+1)} - \bar{x}^{(k)}\|_E < 0,5$

Iteračný proces ukončíme, keď bude splnená zastavovacia podmienka $\|\bar{x}^{(k+1)} - \bar{x}^{(k)}\| < \varepsilon$, $\varepsilon=0,5$.

Výpočet 1. iterácie riešenia:

$$\bar{x}^{(1)} = H_S \cdot \bar{x}^{(0)} + \bar{g}_S$$

$$\begin{pmatrix} x^{(1)} \\ y^{(1)} \\ z^{(1)} \end{pmatrix} = \begin{pmatrix} 0 & 0,25 & -0,5 \\ 0 & -0,1 & 0 \\ 0 & 0,05 & -0,167 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -3 \\ 2,2 \\ 1,067 \end{pmatrix} = \begin{pmatrix} -3 \\ 2,2 \\ 1,067 \end{pmatrix}$$

Rezíduum: $\bar{r}^{(1)} = \bar{b} - A \cdot \bar{x}^{(1)} = \begin{pmatrix} -12 \\ 5 \\ -4 \end{pmatrix} - \begin{pmatrix} 4 & -1 & 2 \\ 2 & 5 & 1 \\ 1 & 1 & -3 \end{pmatrix} \cdot \begin{pmatrix} -3 \\ 2,2 \\ 1,067 \end{pmatrix} = \begin{pmatrix} 0,067 \\ -1,067 \\ 0 \end{pmatrix}$

Norma rezídua: $\|r^{(1)}\|_E = \sqrt{(0,067)^2 + (-1,067)^2 + (0)^2} = 1,0691$

Zistíme platnosť zastavovacej podmienky: $\|\bar{x}^{(1)} - \bar{x}^{(0)}\|_E < 0,5$:

$$\|\bar{x}^{(1)} - \bar{x}^{(0)}\|_E = \sqrt{(-3-0)^2 + (2,2-0)^2 + (1,067-0)^2} = \sqrt{3^2 + 2,2^2 + 1,067^2} = 3,870 > 0,5$$

Výpočet 2. iterácie riešenia:

$$\bar{x}^{(2)} = H_S \cdot \bar{x}^{(1)} + \bar{g}_S$$

$$\begin{pmatrix} x^{(2)} \\ y^{(2)} \\ z^{(2)} \end{pmatrix} = \begin{pmatrix} 0 & 0,25 & -0,5 \\ 0 & -0,1 & 0 \\ 0 & 0,05 & -0,167 \end{pmatrix} \cdot \begin{pmatrix} -3 \\ 2,2 \\ 1,067 \end{pmatrix} + \begin{pmatrix} -3 \\ 2,2 \\ 1,067 \end{pmatrix} = \begin{pmatrix} -2,983 \\ 1,98 \\ 0,999 \end{pmatrix}$$

Rezíduum: $\bar{r}^{(2)} = \bar{b} - A \cdot \bar{x}^{(2)} = \begin{pmatrix} -12 \\ 5 \\ -4 \end{pmatrix} - \begin{pmatrix} 4 & -1 & 2 \\ 2 & 5 & 1 \\ 1 & 1 & -3 \end{pmatrix} \cdot \begin{pmatrix} -2,983 \\ 1,98 \\ 0,999 \end{pmatrix} = \begin{pmatrix} -0,086 \\ 0,067 \\ 0 \end{pmatrix}$

Norma rezídua: $\|r^{(1)}\|_E = \sqrt{(-0,086)^2 + (0,067)^2 + (0)^2} = 0,10902$

Zistíme platnosť zastavovacej podmienky: $\|\bar{x}^{(2)} - \bar{x}^{(1)}\|_E < 0,5$:

$$\begin{aligned} \|\bar{x}^{(2)} - \bar{x}^{(1)}\|_E &= \sqrt{(-2,983 - (-3))^2 + (1,98 - 2,2)^2 + (0,999 - 1,067)^2} \\ &= \sqrt{0,017^2 + (-0,22)^2 + (-0,068)^2} = 0,231 < 0,5 \quad \Rightarrow \text{platí.} \end{aligned}$$

d) *Presnosť poslednej iterácie riešenia je euklidovská norma vektora:*

$$\|\bar{x}^{(2)} - \bar{x}^{(1)}\|_E = \sqrt{0,017^2 + 0,22^2 + 0,068^2} = 0,231$$

Záver

Výpočty zapíšeme do tabuľky:

k	$\bar{x}^{(k)}$	$\ \bar{x}^{(k+1)} - \bar{x}^{(k)}\ _E$	$\bar{r}^{(k)} = \bar{b} - A \cdot \bar{x}^{(k)}$	$\ \bar{r}^{(k)}\ _E$
0	$\bar{x}^{(0)} = (0, 0, 0)^T$	-	$\bar{r}^{(0)} = (-12, 5, -4)^T$	13,6015
1	$\bar{x}^{(1)} = (-3; 2, 2; 1, 067)^T$	3,870	$\bar{r}^{(1)} = (-0, 067; -1, 067; 0)^T$	1,0691
2	$\bar{x}^{(2)} = (-2, 983; 1, 98; 0, 99)^T$	0,231	$\bar{r}^{(2)} = (-0, 086; 0, 067; 0)^T$	0,10902
Poznámka. Presné riešenie systému je $\bar{x}^p = (-3; 2; 1)^T$				

Tabuľka 4.2. Riešenie systému (4.8) Gauss-Seidlovou metódou

Výpočet sme ukončili po 2. iterácii, pretože $\|\bar{x}^{(2)} - \bar{x}^{(1)}\|_E = 0,231 < 0,5$. Numerickým riešením daného SLR so zadanou toleranciou je 2. iterácia riešenia $\bar{x}^{(2)} = (-2, 983; 1, 98; 0, 99)^T$.

PRÍKLAD č. 2

Daný je systém lineárnych rovníc:

$$\begin{aligned} 5x + 7y + 2z &= -26 \\ x + 5y + 3z &= -4 \\ -2x + y + 4z &= 24. \end{aligned}$$

- a) Upravte systém (ekvivalentnými úpravami) tak, aby bola zaručená konvergencia Jacobiho metódy. Svoju úpravu zdôvodnite.
- b) Nájdite iteračnú maticu H_J a vektor \bar{g}_J .
- c) Nájdite jednu iteračnú schému $\bar{x}^{(k+1)} = H \cdot \bar{x}^{(k)} + \bar{g}$, ktorá nekonverguje k riešeniu systému. Zdôvodnite, prečo nájdená iteračná schéma nekonverguje k riešeniu systému.

RIEŠENIE

SLR prepíšeme na tvar $A \cdot \bar{x} = \bar{b}$, kde $A = \begin{pmatrix} 5 & 7 & 2 \\ 1 & 5 & 3 \\ -2 & 1 & 4 \end{pmatrix}$, $\bar{b} = \begin{pmatrix} -26 \\ -4 \\ 24 \end{pmatrix}$.

- a) *Podmienky konvergence metódy*

Overíme, či matica systému A je ostro diagonálne dominantná:

$$\begin{aligned} |5| > |7| + |2| &\dots\dots\dots \text{neplatí} \\ |5| > |1| + |3| &\dots\dots\dots \text{platí} \\ |4| > |-2| + |1| &\dots\dots\dots \text{platí} \end{aligned}$$

Aby sme zaručili konvergenciu metódy, upravíme maticu A ekvivalentnými úpravami na ostro diagonálne dominantnú maticu. Výmena riadkov matice A nám, žiaľ, nepomôže, urobíme inú ekvivalentnú úpravu. Vynásobíme 2. riadok matice A číslom (-1) a pripočítame ho k 1. riadku

$$\begin{pmatrix} 5 & 7 & 2 \\ 1 & 5 & 3 \\ -2 & 1 & 4 \end{pmatrix} \cdot (-1) \approx \begin{pmatrix} 4 & 2 & -1 \\ 1 & 5 & 3 \\ -2 & 1 & 4 \end{pmatrix}$$

Matica A je už ostro diagonálne dominantná: $|4| > |2| + |1|$
 $|5| > |1| + |3|$
 $|4| > |-2| + |1|$

Tou istou ekvivalentnou úpravou upravíme aj vektor pravých strán \bar{b} :

$$\begin{pmatrix} -26 \\ -4 \\ 24 \end{pmatrix} \cdot (-1) \approx \begin{pmatrix} -22 \\ -4 \\ 24 \end{pmatrix}$$

Riešime upravený systém $A \cdot \bar{x} = \bar{b} \Leftrightarrow \begin{pmatrix} 4 & 2 & -1 \\ 1 & 5 & 3 \\ -2 & 1 & 4 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -22 \\ -4 \\ 24 \end{pmatrix}$.

b) *Iteračná matica H_J a vektor g_J Jacobiho metódy*

Systém prevedieme na iteračný tvar Jacobiho metódy: $\bar{x}^{(k+1)} = H_J \cdot \bar{x}^{(k)} + \bar{g}_J$

$$4x + 2y - z = -22 \Rightarrow x = 0x - \frac{1}{2}y + \frac{1}{4}z - \frac{11}{2}$$

$$x + 5y + 3z = -4 \Rightarrow y = -\frac{1}{5}x + 0y - \frac{3}{5}z - \frac{4}{5}$$

$$-2x + y + 4z = -24 \Rightarrow z = \frac{1}{2}x - \frac{1}{4}y + 0z - 6,$$

pričom $H_J = \begin{pmatrix} 0 & -\frac{1}{2} & \frac{1}{4} \\ -\frac{1}{5} & 0 & -\frac{3}{5} \\ \frac{1}{2} & -\frac{1}{4} & 0 \end{pmatrix}$, $\bar{g}_J = \begin{pmatrix} -\frac{11}{2} \\ -\frac{4}{5} \\ -6 \end{pmatrix}$.

Overíme postačujúcu podmienku konvergencie iteračnej schémy $\bar{x}^{(k+1)} = H \cdot \bar{x}^{(k)} + \bar{g}$ k riešeniu systému $A \cdot \bar{x} = \bar{b}$, $\|H_J\| < 1$, kde $\|H_J\|$ je ľubovoľná norma matice H_J . Vypočítame napr. riadkovú normu H_J :

$$\begin{aligned} \|H_J\|_R &= \max \left\{ \left| 0 \right| + \left| -\frac{1}{2} \right| + \left| \frac{1}{4} \right|; \left| -\frac{1}{5} \right| + \left| 0 \right| + \left| -\frac{3}{5} \right|; \left| \frac{1}{2} \right| + \left| -\frac{1}{4} \right| + \left| 0 \right| \right\} \\ &= \max \left\{ \frac{3}{4}, \frac{4}{5}, \frac{3}{4} \right\} = \max \{ 0,75; 0,8; 0,75 \} = 0,75 < 1 \end{aligned}$$

c) Iteračná schéma $x^{(k+1)} = H \cdot x^{(k)} + g$, ktorá nekonverguje k riešeniu systému $A \cdot x = b$:

$$4x + 2y - z = -22 \Rightarrow (x + 3x) + 2y - z = -22 \Rightarrow x = -3x - 2y + z - 22$$

$$x + 5y + 3z = -4 \Rightarrow x + (y + 4y) + 3z = -4 \Rightarrow y = -x - 4y - 3z - 4$$

$$-2x + y + 4z = -24 \Rightarrow -2x + y + (x + 3z) = -24 \Rightarrow z = 2x - y - 3z - 24,$$

pričom
$$H = \begin{pmatrix} -3 & -2 & 1 \\ -1 & -4 & -3 \\ 2 & -1 & -3 \end{pmatrix}, \quad \bar{g} = \begin{pmatrix} -22 \\ -4 \\ -24 \end{pmatrix}.$$

Vypočítame $\|H\|_R$:

$$\|H\|_R = \max\{|-3| + |-2| + |1|; |-1| + |-4| + |-3|; |2| + |-1| + |-3|\} = \max\{6, 8, 6\} = 8 > 1$$

Pretože $\|H\| > 1$, iteračná schéma $\bar{x}^{(k+1)} = H \cdot \bar{x}^{(k)} + \bar{g}$ nekonverguje k riešeniu systému $A \cdot \bar{x} = \bar{b}$.

5 APROXIMÁCIA FUNKCIÍ

Pri sledovaní závislosti dvoch parametrov, napríklad x a y , často potrebujeme určiť hodnotu závislej premennej y v nejakom bode a . Ak poznáme funkčnú závislosť medzi premennými x a y , predpis funkcie $y = f(x)$, potom hodnotu funkcie $f(x)$ v bode $x = a$ vypočítame veľmi jednoducho, dosadením $x = a$ do predpisu funkcie, napr. ak $y = 3x$ a $x = 3$, potom $f(3) = 9$.

V technickej praxi často nepoznáme predpis funkcie $y = f(x)$, ktorá popisuje pozorovaný jav. Vieme len, že medzi x a y existuje určitá funkčná závislosť a z meraní máme k dispozícii údaje (dáta) len v izolovaných bodoch – hodnoty x_i a y_i , ktoré sú usporiadané v tabuľke, $x_1 < x_2 < \dots < x_{n+1}$ (Tabuľka 5.1).

x_i	x_1	x_2	\dots	x_{n+1}
$y_i = f(x_i)$	y_1	y_2	\dots	y_{n+1}

Tabuľka 5.1. Funkcia daná tabuľkou

Problém nastáva, ak potrebujeme poznať funkčnú hodnotu v bode, ktorá sa v tabuľke nenachádza, informáciu o priebehu funkcie v okolí nejakého bodu, resp. na danom intervale. Podobne je tomu aj v situácii, keď síce pracujeme so známou funkciou, ale jej predpis je príliš náročný na počítanie a relevantnú hodnotu funkcie, derivácie, či integrálu nevieme obvyklým matematickým výpočtom získať. V takýchto prípadoch funkciu $f(x)$ nahrádzame, **aproximujeme** inou, podobnou funkciou $g(x)$, ktorá je funkcii $f(x)$ v istom zmysle blízka.

Poznámka. V matematike pod označením $y = f(x)$ (príp. iba $y(x)$) rozumieme funkčnú závislosť dvoch abstraktných premenných: závisle premennej y od nezávisle premennej x . V jednotlivých vedných ako aj technických odboroch však toto všeobecné označenie zvyčajne nadobúda konkrétny význam so svojím typickým pomenovaním ako napríklad:

- $s(t)$ – závislosť dráhy („polohy telesa“) s od času t ,
- $v(T)$ – závislosť rýchlosti šírenia zvuku vo vzduchu v od teploty vzduchu T ,
- $v(S)$ – závislosť rýchlosti prúdiacej tekutiny v rúre v od prierezu rúry S ,
- $T(x)$ – závislosť teploty T od miesta prestupu tepla x .

Možností, ako určovať a posudzovať **mieru podobnosti funkcií** je niekoľko. Ich cieľom je zistiť a vhodným spôsobom vyčíslieť odchýlku, ktorou sa funkcia $f(x)$ líši od $g(x)$. Väčšinou vychádzame z porovnania funkčných hodnôt, rozdielu $|f(x) - g(x)|$, ktorý na grafe funkcií reprezentuje vzdialenosť odpovedajúcich si bodov (Obr. 5.1, Obr. 5.2):

$$|[x, f(x)], [x, g(x)]| = \sqrt{(x - x)^2 + (f(x) - g(x))^2} = |f(x) - g(x)|.$$

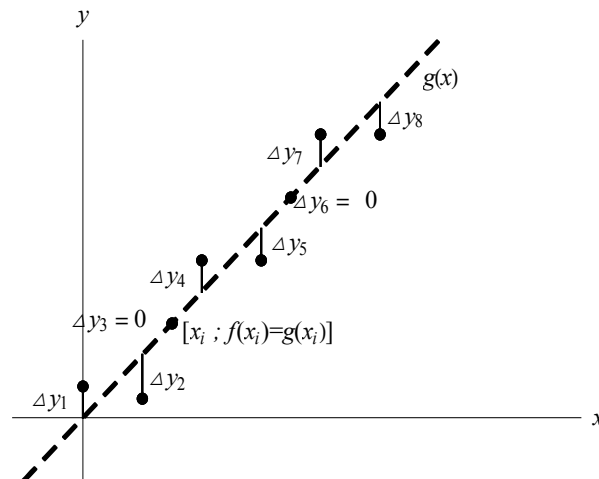
Analogicky tak mieru podobnosti funkcií niekedy zjednodušene nazývame aj **vzdialenosť funkcií**. Otázka je, čo zo sledovania vzdialeností zodpovedajúcich si bodov zobrať za základ

dostatočne reprezentujúci mieru podobnosti – vzdialenosť funkcií. Uvedieme dva najznámejšie prístupy jej definovania.

Definícia 5.1 Nech funkcie $f(x)$ a $g(x)$ sú spojité na intervale $\langle a, b \rangle$. Mierou podobnosti (vzdialenosťou) funkcií $f(x)$ a $g(x)$ nazveme číslo:

$$d(f, g) = \max |f(x) - g(x)|, x \in \langle a, b \rangle.$$

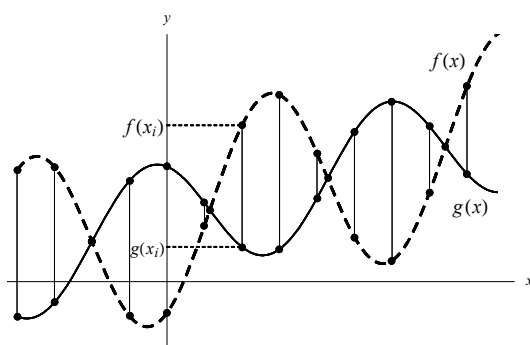
Podľa tejto definície za vzdialenosť dvoch funkcií považujeme maximálnu vzdialenosť zo vzdialeností zodpovedajúcich si bodov na grafoch $f(x)$ a $g(x)$.



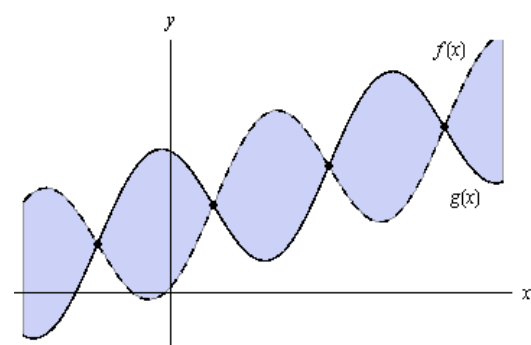
Obr. 5.1. Vzdialenosti $|f(x_i) - g(x_i)| = \Delta y_i$. Funkcia $f(x)$ je daná tabuľkou

V diskkrétnej forme (pre funkcie dané izolovanými bodmi) používame vzťah:

$$d(f, g) = \max |f(x_i) - g(x_i)|, x_i \in \langle a, b \rangle.$$



Obr. 5.2. Vzdialenosti zodpovedajúcich si bodov $|f(x_i) - g(x_i)|$



Obr. 5.3. Plocha ohraničená grafmi funkcií $f(x)$ a $g(x)$

Iný prístup je sledovanie veľkosti plochy ohraničenej grafmi funkcií (Obr. 5.3):

Definícia 5.2 Nech funkcie $f(x)$ a $g(x)$ sú spojité na intervale $\langle a, b \rangle$. Mierou podobnosti (vzdialenosťou) funkcií $f(x)$ a $g(x)$ nazveme číslo:

$$d(f, g) = \int_a^b |f(x) - g(x)| dx,$$

pričom pre všetky $x \in \langle a, b \rangle$ je buď $f(x) \geq g(x)$, alebo $f(x) \leq g(x)$.

V diskkrétnej forme sa bude jednať o súčet vzdialeností jednotlivých zodpovedajúcich si dvojíc bodov (Obr. 5.2):

$$d(f, g) = \sum_{i=1}^{n+1} |f(x_i) - g(x_i)|, x_i \in \langle a, b \rangle.$$

Keďže práca s absolútnou hodnotou je komplikovaná (navyše v nulových bodoch nie je funkcia s absolútnou hodnotou diferencovateľná), najčastejšie na vyjadrenie vzdialenosti sa používa jej druhá mocnina, štvorec vzdialenosti.

Definícia 5.3 Nech funkcie $f(x)$ a $g(x)$ sú spojité na intervale $\langle a, b \rangle$. Mierou podobnosti (vzdialenosťou) funkcií $f(x)$ a $g(x)$ nazveme číslo:

$$d(f, g) = \int_a^b (f(x) - g(x))^2 dx,$$

v diskkrétnej forme:

$$d(f, g) = \sum_{i=1}^{n+1} (f(x_i) - g(x_i))^2, x_i \in \langle a, b \rangle.$$

V štatistike ju poznáme pod názvom kvadratická odchýlka.

Pri aproximácii chceme funkciu $f(x)$ čo možno najvernejšie nahradiť funkciou $g(x)$. Hľadáme takú funkciu $g(x)$, ktorá sa v čo najmenej možnej miere líši od funkcie $f(x)$, vzdialenosť funkcií je minimálna. Metódy aproximácie založené na definícii Definícia 5.1 sa nazývajú **metódy minimalizácie maximálnej chyby** (minimax). Metóda aproximácie založená na definícii Definícia 5.3 sa nazýva **metóda najmenších štvorcov** (MNS).

Od funkcie $g(x)$, ktorou aproximujeme funkciu $f(x)$ budeme požadovať:

- aby bola ľahko vyjadriteľná pomocou množiny dát $\{[x_i, f(x_i)]\}_{i=1}^{n+1}$,
- jej hodnoty by sa mali dať ľahko vypočítať,
- možno ju ľahko derivovať a integrovať.

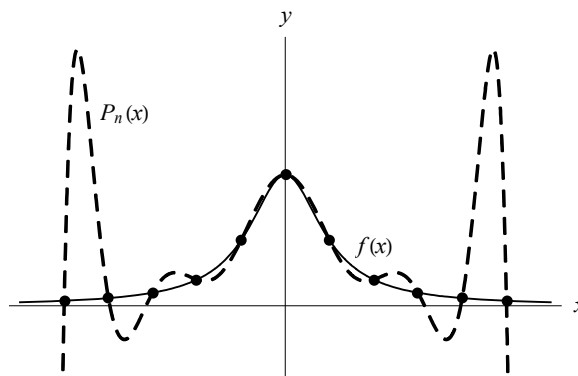
Z tohto pohľadu je veľmi vhodná aproximačná funkcia $g(x)$ polynóm (polynomická funkcia). Získať ho môžeme z konečného počtu vstupných údajov pomocou 3 základných aritmetických operácií (súčet, súčin, rozdiel), ľahko sa derivuje a integruje, výsledkom operácií je opäť polynóm a možno ním aproximovať spojitú funkciu s ľubovoľnou presnosťou. Dá sa dokázať nasledujúca veta:

Veta 5.1 (Weierstrassova veta o aproximácii) Nech $f(x)$ je spojitá funkcia reálnej premennej x na uzavretom intervale $\langle a, b \rangle$ a $P_n(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$, $a_i \in R, a_n \neq 0$ je polynóm (polynomická funkcia) stupňa n . Potom pre každé $\varepsilon > 0$ existuje taký polynóm $P_n(x)$, že pre všetky $x \in \langle a, b \rangle$ platí:

$$|f(x) - P_n(x)| < \varepsilon.$$

Stupeň polynómu n závisí od hodnoty ε .

Nevýhodou polynómu je, že grafy polynómov vyšších stupňov sa v krajných častiach intervalu nadmerne vlnia, kmitajú. Medzi jednotlivými bodmi $[x_i, y_i]$ sa tak môže nachádzať aj viacero extrémov, dochádza k prekmitom, a preto sú na aproximáciu v týchto oblastiach nevhodné. Počet extrémov polynómu úzko súvisí s jeho stupňom, je len o jeden menší. To znamená, ak stupeň polynómu je n , potom počet extrémov je $n - 1$.



Obr. 5.4. Kmitanie polynómov vyšších stupňov

Okrem polynómu sa na aproximáciu môžu vzhľadom na priebeh funkcie $f(x)$ s istými úpravami použiť aj ďalšie elementárne funkcie, ako racionálne lomené funkcie, trigonometrické funkcie, exponenciálna funkcia, ap.

Body z množiny $\{x_1, x_2, \dots, x_{n+1}\}$ sa nazývajú **uzly** (sú v nich dané alebo v nich počítame hodnoty funkcie). Usporiadané dvojice $[x_i, f(x_i)]$, $i = 1, 2, \dots, n+1$ nazývame **póly** (body grafu funkcie).

Podľa charakteru množiny vstupných bodov $\{[x_i, f(x_i)]\}_{i=1}^{n+1}$, môžeme voliť medzi dvoma základnými spôsobmi aproximácie:

1. od grafu aproximačnej funkcie požadujeme, aby prechádzal všetkými bodmi

$\{[x_i, f(x_i)]\}_{i=1}^{n+1}$. Množina bodov zvyčajne nie je veľká, tvoria ju relatívne presne namerané údaje. Hovoríme o **interpolácii**. Používame Newtonov interpolačný polynóm, Lagrangeov interpolačný polynóm, Čebyševov interpolačný polynóm, Hermitove polynómy, interpolačné splajny atď.

2. graf aproximačnej funkcie nemusí prechádzať každým bodom množiny $\{[x_i, f(x_i)]\}_{i=1}^{n+1}$, avšak dostatočne presne kopíruje priebeh funkcie $f(x)$. Počet dát (bodov) môže značne prekračovať počet bodov nutných na jednoznačné určenie vhodnej aproximačnej funkcie alebo vstupné údaje môžu byť zaťažené chybou merania. Hovoríme o **aproximácii** (v užšom zmysle slova). Patrí sem napríklad Metóda najmenších štvorcov, aproximačné splajny.

V prípade, že chceme určiť hodnotu funkcie v bode x , ktorý sa medzi vstupnými uzlami nenachádza, $x \notin \{x_1, x_2, \dots, x_{n+1}\}$, rozlišujeme medzi dvomi polohami bodu x :

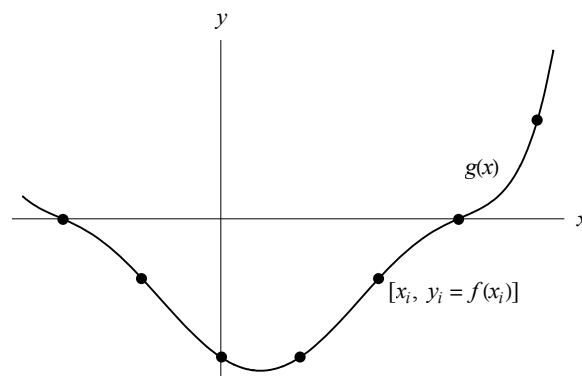
- a) $x \in \langle x_1, x_{n+1} \rangle$. Na určenie funkčnej hodnoty použijeme **interpoláciu** – medzi dva známe póly vkladáme ďalší pól.
- b) $x \notin \langle x_1, x_{n+1} \rangle$. Hovoríme o **extrapolácii**. Interpoláčnne techniky môžeme použiť iba v prípade, že bod x leží veľmi blízko krajných bodov intervalu $\langle x_1, x_{n+1} \rangle$.

5.1 INTERPOLÁCIA

Daných je $n+1$ hodnôt reálnej funkcie $f(x)$ reálnej premennej x v bodoch $x_1 < x_2 < \dots < x_{n+1}$, funkcia daná tabuľkou (Tabuľka 5.1).

Definícia 5.4 Pod **interpoláciou** rozumieme aproximáciu funkcie $f(x)$ takou funkciou $g(x)$, že v daných bodoch x_i nadobúda funkcia $g(x)$ hodnoty $y_i = f(x_i)$, tj.

$$f(x_i) = g(x_i), \quad i = 1, 2, \dots, n+1.$$



Obr. 5.5. Interpolácia funkcie $f(x)$ funkciou $g(x)$

Ide o špeciálny prípad aproximácie, kedy sa funkcie $f(x)$ a $g(x)$ zhodujú vo všetkých tabuľkových bodoch x_i , $|f(x_i) - g(x_i)| = 0$, $i = 1, 2, \dots, n+1$. Body x_1, x_2, \dots, x_{n+1} nazývame **uzly interpolácie** a $X_i[x_i, y_i]$ **póly interpolácie**. Funkcia $g(x)$ je **interpoláčnna funkcia**, ak je ňou polynóm $g(x) = P(x)$, tak je to **interpoláčnny polynóm** $P(x)$.

Niekedy, okrem zhody funkčných hodnôt vo všetkých tabuľkových bodoch x_i , sa navyše požadujú aj ďalšie vlastnosti, ako rovnosť prvých derivácií $f'(x_i) = P'(x_i)$ alebo

$$\int_{x_1}^{x_{n+1}} f(x) dx = \int_{x_1}^{x_{n+1}} g(x) dx.$$

Ďalej sa budeme zaoberať interpoláciou polynómom na uzavretom intervale $\langle x_1, x_{n+1} \rangle$.

Veta 5.2 Daná je množina usporiadaných dvojíc $\{[x_i; y_i]\}_{i=1}^{n+1}$, $x_1 < x_2 < \dots < x_{n+1}$. Existuje práve jeden polynóm $P(x)$ stupňa najviac n taký, že $P_n(x_i) = y_i$.

Opäť predpokladáme, že máme nameraných $(n+1)$ hodnôt neznámej spojitej funkcie $f(x)$ (Tabuľka 5.2).

x_i	x_1	x_2	\dots	x_{n+1}
$y_i = f(x_i)$	y_1	y_2	\dots	y_{n+1}

Tabuľka 5.2

Hľadaný interpolačný polynóm $P_n(x)$, pre ktorý $P_n(x_i) = y_i$ bude mať tvar:

$$P_n(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n, \quad a_i \in \mathbb{R}, a_n \neq 0, i = 0, 1, \dots, n \quad (5.1)$$

Dosadením $n+1$ dvojíc $[x_i, y_i]$ do (5.1) dostávame systém $n+1$ lineárnych rovníc o $n+1$ neznámych a_0, a_1, \dots, a_n :

$$\begin{aligned}
 P_n(x_1) = y_1 & \Leftrightarrow a_0 + a_1x_1 + a_2x_1^2 + a_3x_1^3 + \dots + a_nx_1^n = y_1 \\
 P_n(x_2) = y_2 & \Leftrightarrow a_0 + a_1x_2 + a_2x_2^2 + a_3x_2^3 + \dots + a_nx_2^n = y_2 \\
 P_n(x_3) = y_3 & \Leftrightarrow a_0 + a_1x_3 + a_2x_3^2 + a_3x_3^3 + \dots + a_nx_3^n = y_3 \\
 & \dots \dots \dots \\
 P_n(x_{n+1}) = y_{n+1} & \Leftrightarrow a_0 + a_1x_{n+1} + a_2x_{n+1}^2 + a_3x_{n+1}^3 + \dots + a_nx_{n+1}^n = y_{n+1}
 \end{aligned} \quad (5.2)$$

v maticovom tvare:

$$\begin{pmatrix} 1 & x_1 & x_1^2 & x_1^3 & \dots & x_1^n \\ 1 & x_2 & x_2^2 & x_2^3 & \dots & x_2^n \\ 1 & x_3 & x_3^2 & x_3^3 & \dots & x_3^n \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n+1} & x_{n+1}^2 & x_{n+1}^3 & \dots & x_{n+1}^n \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \dots \\ a_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \dots \\ y_{n+1} \end{pmatrix} \quad (5.3)$$

Systém (5.2), resp. (5.3) má práve jedno riešenie, ak determinant matice systému je rôzny od 0. Na jednoznačné určenie polynómu 1. stupňa (grafom je priamka) potrebujeme dva rôzne body $\{[x_1, y_1], [x_2, y_2]\}$, na jednoznačné určenie polynómu 2. stupňa (grafom je parabola) potrebujeme tri rôzne body atď.; $n+1$ bodov jednoznačne určuje polynóm n -tého stupňa. Matica systému sa nazýva **Vandermontova matica**, determinant matice je Vandermontov determinant.

Tento postup hľadania interpolačného polynómu je veľmi náročný (pozri kapitolu 4), v praxi sa používa len ojedinele. Numerické metódy používajú výhodnejšie spôsoby výpočtu koeficientov polynómu. Pri ich výbere a použití si treba uvedomiť, že pri rovnakých vstupných údajoch pôjde vždy iba o rôzne tvary toho istého polynómu. Podrobnejšie sa budeme venovať dvom numerickým metódam, konštrukcii Newtonovho a Lagrangeovho polynómu.

5.1.1 HĽADANIE INTERPOLAČNÉHO POLYNÓMU POMOCOU TABUĽKY DIFERENCIÍ

Tabuľku diferencií používame pri hľadaní interpolačného polynómu v tvare:

$$N_n(x) = a_0 + a_1(x - x_1) + a_2(x - x_1)(x - x_2) + \dots + a_n(x - x_1)(x - x_2)\dots(x - x_n). \quad (5.4)$$

Uvedený tvar je Newtonov tvar a polynóm sa nazýva *Newtonov interpolačný polynóm* (NIP). Podľa definície interpolačného polynómu: $g(x_i) = f(x_i) \wedge g(x) = N_n(x)$, $i = 1, 2, \dots, n+1$. Po dosadení bodov $[x_i, y_i]$ do $N_n(x_i) = y_i$, dostávame systém lineárnych rovníc s dolnou trojuhlovou maticou:

$$\begin{aligned} N_n(x_1) = y_1 & \Leftrightarrow a_0 & = y_1 \\ N_n(x_2) = y_2 & \Leftrightarrow a_0 + a_1(x_2 - x_1) & = y_2 \\ N_n(x_3) = y_3 & \Leftrightarrow a_0 + a_1(x_3 - x_1) + a_2(x_3 - x_1)(x_3 - x_2) & = y_3 \\ N_n(x_{n+1}) = y_{n+1} & \Leftrightarrow a_0 + a_1(x_{n+1} - x_1) + a_2(x_{n+1} - x_1)(x_{n+1} - x_2) + \dots + a_n(x_{n+1} - x_1)\dots(x_{n+1} - x_n) & = y_{n+1} \end{aligned}$$

v maticovom tvare:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & (x_2 - x_1) & 0 & 0 & \dots & 0 \\ 1 & (x_3 - x_1) & (x_3 - x_1)(x_3 - x_2) & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & (x_{n+1} - x_1) & (x_{n+1} - x_1)(x_{n+1} - x_2) & \dots & (x_{n+1} - x_1)(x_{n+1} - x_2)\dots(x_{n+1} - x_n) & \dots \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \dots \\ a_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \dots \\ y_{n+1} \end{pmatrix}$$

Koeficienty a_i , $i = 0, 1, 2, \dots, n$, vypočítame takto:

$$\begin{aligned} a_0 = y_1 & \Leftrightarrow a_0 = \Delta^0 f(x_1) \\ a_0 + a_1(x_2 - x_1) = y_2 & \Leftrightarrow a_1 = \frac{y_2 - a_0}{x_2 - x_1} = \frac{y_2 - y_1}{x_2 - x_1} = \frac{\Delta^0 f(x_2) - \Delta^0 f(x_1)}{x_2 - x_1} = \Delta^1 f(x_1) \\ a_0 + a_1(x_3 - x_1) + a_2(x_3 - x_1)(x_3 - x_2) = y_3 & \Leftrightarrow a_2 = \frac{\frac{y_3 - y_2}{x_3 - x_2} - \frac{y_2 - y_1}{x_2 - x_1}}{x_3 - x_1} = \frac{\Delta^1 f(x_2) - \Delta^1 f(x_1)}{x_3 - x_1} = \Delta^2 f(x_1) \\ \dots, \text{ atď.} & \text{ Výrazy } \Delta^0 f(x_1), \Delta^1 f(x_1), \Delta^2 f(x_1) \text{ atď, sú tzv. pomerné diferencie.} \end{aligned}$$

Označme $y_i = f(x_i) = f_i$.

Definícia 5.5 Pre uzly x_i a funkčné hodnoty $y_i = f(x_i) = f_i$, $i = 1, 2, \dots, n+1$, $x_1 < x_2 < \dots < x_{n+1}$ **pomerné diferencie** k -tého rádu $\Delta^k f_i$, $k = 0, 1, \dots, n$, definujeme rekurentným spôsobom:

$$\begin{aligned} k = 0 & : \Delta^0 f_i = f_i \\ k = 1 & : \Delta^1 f_i = \frac{\Delta^0 f_{i+1} - \Delta^0 f_i}{x_{i+1} - x_i} \\ k = 2 & : \Delta^2 f_i = \frac{\Delta^1 f_{i+1} - \Delta^1 f_i}{x_{i+2} - x_i} \\ k < n+1 & : \Delta^k f_i = \frac{\Delta^{k-1} f_{i+1} - \Delta^{k-1} f_i}{x_{i+k} - x_i} \end{aligned}$$

Pre Newtonov interpolačný polynóm vzhľadom na pomerné diferencie potom platí:

$$\begin{aligned} N_n(x) &= a_0 + a_1(x - x_1) + a_2(x - x_1)(x - x_2) + \dots + a_n(x - x_1)(x - x_2)\dots(x - x_n) \\ &= \Delta^0 f_1 + \Delta^1 f_1(x - x_1) + \Delta^2 f_1(x - x_1)(x - x_2) + \dots + \Delta^n f_1(x - x_1)(x - x_2)\dots(x - x_n) \end{aligned} \quad (5.5)$$

Interpolačný polynóm vyjadrený vzťahom (5.5) sa nazýva **Newtonov interpolačný polynóm** stupňa n **vpred**.

Newtonov interpolačný polynóm stupňa n môžeme hľadať aj v tvare:

$$\begin{aligned} N_n(x) &= \hat{a}_0 + \hat{a}_1(x - x_{n+1}) + \hat{a}_2(x - x_{n+1})(x - x_n) + \dots + \hat{a}_n(x - x_{n+1})(x - x_n)\dots(x - x_2) = \\ &= \Delta^0 f_{n+1} + \Delta^1 f_n(x - x_{n+1}) + \Delta^2 f_{n-1}(x - x_{n+1})(x - x_n) + \dots + \Delta^n f_1(x - x_{n+1})(x - x_n)\dots(x - x_2), \end{aligned} \quad (5.6)$$

ktorý sa nazýva **Newtonov interpolačný polynóm** stupňa n **vzad**.

Koeficienty Newtonovho interpolačného polynómu stupňa n (vpred aj vzad) určíme zo **všeobecnej tabuľky diferencií**, tzv. **Fraserovho diagramu** (Tabuľka 5.3).

Nespornou výhodou Newtonovho interpolačného polynómu je fakt, že pridaním ďalšieho uzla stačí dopočítať len 1 riadok tabuľky diferencií. Ktorú podobu polynómu (vpred či vzad) použijeme, záleží od umiestnenia bodu a , v ktorom chceme čo možno najpresnejšie určiť chýbajúcu hodnotu funkcie $f(a)$, $a \in \langle x_1, x_{n+1} \rangle$, $a \neq x_i$. Ak a sa nachádza na začiatku intervalu $\langle x_1, x_{n+1} \rangle$, použijeme NIP vpred, ak a sa nachádza na konci intervalu $\langle x_1, x_{n+1} \rangle$, použijeme NIP vzad. Vhodným výberom NIP eliminujeme šírenie chýb vo výpočte. Nežiaduce kmitanie interpolačných polynómov vyšších stupňov riešime tak, že miesto všetkých bodov x_i , ktoré máme k dispozícii, berieme do úvahy iba uzly z najbližšieho okolia bodu $x = a$ odhadovanej hodnoty $f(a)$. Fraserov diagram potom zostrojíme iba nad vybranými uzlami. Počet použitých uzlov závisí od stupňa interpolačného polynómu.

x_i	$y_i = \Delta^0 f_i$	$\Delta^1 f_i$	$\Delta^2 f_i$	$\Delta^3 f_i$	$\Delta^4 f_i$	$\Delta^5 f_i$...	$\Delta^n f_i$
x_1	$y_1 = a_0$							
x_2	y_2	$\Delta^1 f_1 = a_1$						
x_3	y_3	$\Delta^1 f_2$	$\Delta^2 f_1 = a_2$					
x_4	y_4	$\Delta^1 f_3$	$\Delta^2 f_2$	$\Delta^3 f_1 = a_3$				
...	$\Delta^4 f_1 = a_4$			
x_{n-2}	y_{n-2}	$\Delta^1 f_{n-2}$	$\Delta^2 f_{n-3}$	$\Delta^3 f_{n-3}$	$\Delta^4 f_{n-4}$	$\Delta^5 f_1 = a_5$		
x_{n-1}	y_{n-1}	$\Delta^1 f_{n-1}$	$\Delta^2 f_{n-2}$	$\Delta^3 f_{n-2} = \bar{a}_3$	$\Delta^4 f_{n-3} = \bar{a}_4$	$\Delta^5 f_2$		
x_n	y_n	$\Delta^1 f_n = \bar{a}_1$	$\Delta^2 f_{n-1} = \bar{a}_2$					
x_{n+1}	$y_{n+1} = \bar{a}_0$							$\Delta^n f_1 = a_n = \bar{a}_n$

Tabuľka 5.3. Fraserov diagram

5.1.2 LAGRANGEOV POLYNÓM

Francúzsky matematik J. L. Lagrange použil na konštrukciu interpolačného polynómu tvar:

$$P_n(x) = \sum_{k=0}^n f(x_k) \cdot L_k(x) = f(x_0) \cdot L_0(x) + f(x_1) \cdot L_1(x) + \dots + f(x_n) \cdot L_n(x), \quad (5.7)$$

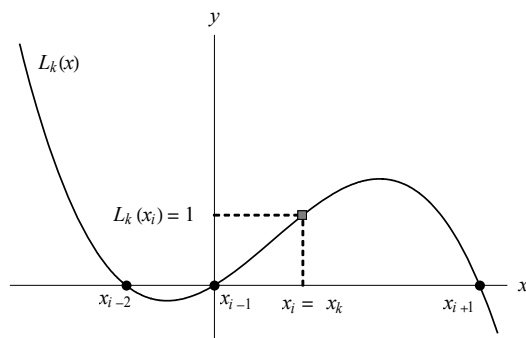
kde definoval tzv. fundamentálne polynómy:

$$L_k(x) = \prod_{\substack{i=0 \\ i \neq k}}^n \frac{x - x_i}{x_k - x_i} = \frac{(x - x_0) \cdot (x - x_1) \cdot \dots \cdot (x - x_{k-1}) \cdot (x - x_{k+1}) \cdot \dots \cdot (x - x_n)}{(x_k - x_0) \cdot (x_k - x_1) \cdot \dots \cdot (x_k - x_{k-1}) \cdot (x_k - x_{k+1}) \cdot \dots \cdot (x_k - x_n)}, \quad (5.8)$$

pričom

$$\begin{aligned} L_k(x_i) &= 1 \quad \text{pre } k = i, \\ L_k(x_i) &= 0 \quad \text{pre } k \neq i, \end{aligned}$$

čím sú splnené podmienky interpolácie $P_n(x_i) = \sum_{k=0}^n f(x_k) \cdot L_k(x_i) = y_i, i = 0, 1, \dots, n$. (Obr. 5.6)



Obr. 5.6

Lagrangeov interpolačný polynóm má potom tvar:

$$\begin{aligned} P_n(x) &= f(x_0) \cdot L_0(x) + f(x_1) \cdot L_1(x) + \dots + f(x_n) \cdot L_n(x) = \\ &= f(x_0) \cdot \frac{(x-x_1)(x-x_2)\dots(x-x_n)}{(x_0-x_1)(x_0-x_2)\dots(x_0-x_n)} + f(x_1) \cdot \frac{(x-x_0)(x-x_2)\dots(x-x_n)}{(x_1-x_0)(x_1-x_2)\dots(x_1-x_n)} + \\ &\quad + \dots + f(x_n) \cdot \frac{(x-x_0)(x-x_1)\dots(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)\dots(x_n-x_{n-1})} \end{aligned}$$

Poznámka. V Lagrangeovom type polynómu sú uzly číslované od 0: $i = 0, 1, \dots, n$.

Na rozdiel od Newtonovho interpolačného polynómu pridaním každého ďalšieho uzla sa všetky polynómy $L_k(x)$ musia znovu prepočítať.

Forma Lagrangeovho interpolačného polynómu je vhodná pri teoretickom skúmaní vlastností interpolačných polynómov ([14]), my ju použijeme pri odvodzovaní vzorcov pre výpočet určitého integrálu (kap. 7).

Pripomíname, že Newtonov aj Lagrangeov polynóm sú len rozdielne formy toho istého polynómu. V Newtonovom interpolačnom polynóme sú hodnoty funkcie $f(x)$ použité sprostredkovane – prostredníctvom pomerných diferencií, zatiaľ čo v Lagrangeovom polynóme sú použité priamo.

5.1.3 ODHAD ABSOLÚTNEJ CHYBY INTERPOLAČNÉHO POLYNÓMU

Veta 5.3 Nech funkcia $f(x)$ je na intervale $\langle a, b \rangle$ spojitá a nech tam má spojité derivácie $f'(x)$, $f''(x)$, ..., $f^{(n+1)}(x)$. Nech $P_n(x)$ je interpolačný polynóm stupňa n určený hodnotami $f(x_i)$ v uzloch x_1, \dots, x_{n+1} , potom pre ľubovoľné $x \in \langle a, b \rangle$ platí:

$$|f(x) - P_n(x)| \leq \frac{M_{n+1}}{(n+1)!} |(x-x_1)(x-x_2)\dots(x-x_{n+1})|,$$

kde $M_{n+1} = \max |f^{(n+1)}(x)|$, $x \in \langle a, b \rangle$.

Veta hovorí o odhade chyby pri určovaní hodnoty $f(x)$ pomocou interpolačného polynómu $P_n(x)$. Z definície Definícia 5.4 je zrejmé, že v uzloch je chyba nulová,

$|f(x_i) - g(x_i)| = 0$, $i = 1, 2, \dots, n+1$. Výraz $|(x - x_1)(x - x_2) \dots (x - x_{n+1})|$ nadobúda pre určité rozloženie uzlov x_1, \dots, x_{n+1} v intervale $\langle a, b \rangle$ svoje maximum a minimum. Najmenšia chyba vzniká v blízkosti uzlov, naopak, čím je bod x vzdialenejší od uzlových bodov (mimo interval $\langle x_1, x_{n+1} \rangle$), chyba prudko narastá. Chyba interpolácie takisto narastá pridávaním ďalších uzlov. To je ďalší z dôvodov, prečo sa vo všeobecnosti interpolácia polynómami vyšších stupňov neodporúča.

Zhrnutie

Interpoláčny polynóm určený hodnotami $f(x_i)$ v uzlových bodoch x_1, \dots, x_{n+1} je na intervale $\langle x_1, x_{n+1} \rangle$ istá „náhrada“ za funkciu $f(x)$. Cieľom aproximácie je, aby sa prostredníctvom polynómu dali na funkcii $f(x)$ vykonať operácie, ktorých realizácia by inak bola komplikovaná alebo nemožná.

Od grafu interpolačnej funkcie sa požaduje, aby prechádzal zadanými bodmi $\{[x_i, f(x_i)]\}_{i=1}^{n+1}$. Množina bodov zvyčajne nie je veľká, tvoria ju relatívne presne namerané údaje. Mohlo by sa zdať, že zvýšením počtu uzlov sa zväčší presnosť aproximácie $f(x)$ interpolačným polynómom. Nie je tomu tak. Zvýšením počtu uzlov sa zvyšuje počet extrémov a tým aj riziko nežiaducich kmitov v okrajových častiach intervalu, rastie chyba metódy (odsek 5.1.3) a vplyv šírenia zaokrúhľovacích chýb. Aproximáciu $f(x)$ môžeme zlepšiť interpolačným **splajnom** (angl. spline) – funkciou, ktorá je po častiach polynómom nízkeho stupňa, pričom jednotlivé časti polynomickej funkcie nadväzujú na seba dostatočne hladko (v nadväzujúcich bodoch sa zvyčajne požaduje rovnosť prvých a druhých derivácií). Najčastejšie používané splajny sú lineárny a kubický. Označenie „spline“ prvýkrát použil Shoenberg v roku 1946.

Najvernejšie informácie o funkcii $f(x)$ sú na intervale $\langle x_1, x_{n+1} \rangle$, ktorý obsahuje uzly interpolácie x_1, x_2, \dots, x_{n+1} . Ak počítame hodnoty $f(x)$ mimo intervalu $\langle x_1, x_{n+1} \rangle$, môžeme hovoriť len o nejakej prognóze – **extrapolácii**, pretože tu už žiadne informácie o funkcii $f(x)$ nemáme. Uspokojivé výsledky dostávame len v prípade, ak počítame hodnoty $f(a)$ v blízkosti intervalu $\langle x_1, x_{n+1} \rangle$.

Pri určovaní hodnoty $f(a)$, $a \in \langle x_1, x_{n+1} \rangle$, $a \neq x_i$, nie je nutné interpolačný polynóm prekladať cez všetky uzly x_1, \dots, x_{n+1} . Stačí zobrať príslušný počet uzlov z najbližšieho okolia $x = a$, pričom ich počet uzlov závisí od stupňa zvoleného interpolačného polynómu.

V minulosti sa interpolačné polynómy využívali na výpočet hodnôt funkcie, ktoré neboli v tabuľkách. Dnešné počítače a kalkulačky to zvládnu jednoducho a rýchlo. V súčasnosti sa interpolačné polynómy využívajú hlavne pri vyhodnocovaní experimentálneho merania.

5.1.4 RIEŠENÉ PRÍKLADY

PRÍKLAD č. 1

Meraním bolo zistených 6 hodnôt spojitej funkcie $f(x)$, ktoré sú uvedené v nasledujúcej tabuľke (Tabuľka 5.4):

i	1	2	3	4	5	6
x_i	-3	-1	0	1	3	6
$y_i = f(x_i)$	3	2	1	2	0	3

Tabuľka 5.4

- Pomocou tabuľky diferencií nájdite Newtonov interpolačný polynóm 3. stupňa $N_3(x)$ (vpred aj vzad), pomocou ktorého čo najpresnejšie vypočítate $f(2)$ a $f'(2)$.
- Nájdite Newtonove interpolačné polynómy (vpred) $N_2(x)$ a $N_4(x)$, pomocou ktorých čo najpresnejšie vypočítate $f(2)$ a $f'(2)$.
- Presnosť výpočtu $f(2)$ interpolačnými polynómami porovnajte na základe absolútnej chyby interpolácie v čísle $x = 2$ za predpokladu, že $f(2) = 1,5$.
- Nakreslite graf funkcie $f(x)$ a interpolačných polynómov do spoločného obrázku.

RIEŠENIE

a) *Newtonov interpolačný polynóm 3. stupňa $N_3(x)$*

Výber bodov: Hľadáme polynóm 3. stupňa \Rightarrow potrebujeme vybrať 4 hodnoty zo 6 nameraných hodnôt $f(x)$. Vzhľadom na to, že chceme vypočítať $f(2)$ a $f'(2)$, budeme vyberať 4 namerané hodnoty z najbližšieho okolia bodu $x = 2$.

Výber realizujeme na základe najmenšej vzdialenosti bodu $x = 2$ od jednotlivých uzlov x_i (Tabuľka 5.5).

		$x = 2$				
x_i	-3	-1	0	1	3	6
$y_i = f(x_i)$	3	2	1	2	0	3
$ 2 - x_i $	$ 2 - (-3) = 5$	$ 2 - (-1) = 3$	$ 2 - 0 = 2$	$ 2 - 1 = 1$	$ 2 - 3 = 1$	$ 2 - 6 = 4$
Najmenšiu vzdialenosť od $x = 2$ majú tieto 4 uzly.						

Tabuľka 5.5. *Výber bodov*

Vybrané namerané hodnoty $[x_i, y_i = f(x_i)]$ sú: $[-1, 2]$, $[0, 1]$, $[1, 2]$, $[3, 0]$.

Zostrojíme tabuľku diferencií (Tabuľka 5.6).

i	x_i	y_i	$\Delta^1 f_i$	$\Delta^2 f_i$	$\Delta^3 f_i$
1	-1	$2 = a_0$	$\Delta^1 f_1 = -1 = a_1$	$\Delta^2 f_1 = 1 = a_2$	$\Delta^3 f_1 = -\frac{5}{12} = a_3 = \hat{a}_3$
2	0	1			
3	1	2	$\Delta^1 f_2 = 1$	$\Delta^2 f_2 = -\frac{2}{3} = \hat{a}_2$	
4	3	$0 = \hat{a}_0$	$\Delta^1 f_3 = -1 = \hat{a}_1$		

Tabuľka 5.6. Tabuľka diferencií pre $N_3(x)$

Výpočet diferencií:

$$\begin{aligned}\Delta^1 f_1 &= \frac{y_2 - y_1}{x_2 - x_1} = \frac{1 - 2}{0 - (-1)} = -1, & \Delta^1 f_2 &= \frac{y_3 - y_2}{x_3 - x_2} = \frac{2 - 1}{1 - 0} = 1, & \Delta^1 f_3 &= \frac{y_4 - y_3}{x_4 - x_3} = \frac{0 - 2}{3 - 1} = -1 \\ \Delta^2 f_1 &= \frac{\Delta^1 f_2 - \Delta^1 f_1}{x_3 - x_1} = \frac{1 - (-1)}{1 - (-1)} = 1, & \Delta^2 f_2 &= \frac{\Delta^1 f_3 - \Delta^1 f_2}{x_4 - x_2} = \frac{-1 - 1}{3 - 0} = -\frac{2}{3}, \\ \Delta^3 f_1 &= \frac{\Delta^2 f_2 - \Delta^2 f_1}{x_4 - x_1} = \frac{-\frac{2}{3} - 1}{3 - (-1)} = -\frac{5}{12}.\end{aligned}$$

Newtonov interpolačný polynóm 3. stupňa vpred:

$$\begin{aligned}N_3(x) &= a_0 + a_1(x - x_1) + a_2(x - x_1)(x - x_2) + a_3(x - x_1)(x - x_2)(x - x_3) \\ &= 2 - 1(x - (-1)) + 1(x - (-1))(x - 0) - \frac{5}{12}(x - (-1))(x - 0)(x - 1) \\ &= 2 - (x + 1) + (x + 1)x - \frac{5}{12}(x + 1)x(x - 1) \\ &= -\frac{5}{12}x^3 + x^2 + \frac{5}{12}x + 1\end{aligned}$$

Newtonov interpolačný polynóm 3. stupňa – vzad:

$$\begin{aligned}N_3(x) &= a_0 + a_1(x - x_4) + a_2(x - x_4)(x - x_3) + a_3(x - x_4)(x - x_3)(x - x_2) \\ &= 0 + (-1)(x - 3) - \frac{2}{3}(x - 3)(x - 1) - \frac{5}{12}(x - 3)(x - 1)(x - 0) \\ &= -(x - 3) - \frac{2}{3}(x - 3)(x - 1) - \frac{5}{12}(x - 3)(x - 1)x \\ &= -\frac{5}{12}x^3 + x^2 + \frac{5}{12}x + 1\end{aligned}$$

Výpočet hľadaných hodnôt $f(2)$ a $f'(2)$:

$$f(2) \approx N_3(2) \Leftrightarrow N_3(2) = -\frac{5}{12}2^3 + 2^2 + \frac{5}{12}2 + 1 = \frac{5}{2} = 2,5 \quad \Rightarrow \quad f(2) \approx 2,5$$

$$f'(2) \approx N_3'(2) \Leftrightarrow N_3'(2) = \left(-\frac{5}{12}3x^2 + 2x + \frac{5}{12}\right)_{x=2} = -\frac{7}{12} = -0,58333 \Rightarrow f'(2) \approx -0,58333$$

b) Newtonove interpolačné polynómy (vpred) $N_2(x)$ a $N_4(x)$, pomocou ktorých čo najpresnejšie vypočítame $f(2)$ a $f'(2)$

$N_2(x)$:

Na jednoznačné určenie polynómu druhého stupňa potrebujeme hodnoty $f(x)$ v troch uzloch. Opäť vyberáme uzly ležiace najbližšie ku $x = 2$. Koeficienty $N_2(x)$ určíme pomocou tabuľky diferencií (Tabuľka 5.7).

i	x_i	y_i	$\Delta^1 f_i$	$\Delta^2 f_i$
1	0	$1 = a_0$	$\Delta^1 f_1 = 1 = a_1$	$\Delta^2 f_1 = -\frac{2}{3} = a_2$
2	1	2		
3	3	$0 = \hat{a}_0$	$\Delta^1 f_2 = -1 = \hat{a}_1$	

Tabuľka 5.7. Tabuľka diferencií pre $N_2(x)$

Hľadaný Newtonov interpolačný polynóm 2. stupňa bude mať tvar:

$$N_2(x) = a_0 + a_1(x - x_1) + a_2(x - x_1)(x - x_2) = 1 + 1(x - 0) - \frac{2}{3}(x - 0)(x - (-1)) = -\frac{2}{3}x^2 + \frac{5}{3}x + 1$$

Výpočet hľadaných hodnôt $f(2)$ a $f'(2)$:

$$f(2) \approx N_2(2) \Rightarrow f(2) \approx \frac{5}{3} \doteq 1,66667 \qquad f'(2) \approx N_2'(2) \Rightarrow f'(2) \approx -1$$

$N_4(x)$:

Na jednoznačné určenie polynómu 4. stupňa potrebujeme hodnoty $f(x)$ v najbližších piatich uzloch ku $x = 2$. V tabuľke diferencií (Tabuľka 5.6) dopočítame 1 riadok:

i	x_i	y_i	$\Delta^1 f_i$	$\Delta^2 f_i$	$\Delta^3 f_i$	$\Delta^4 f_i$
1	-1	$2 = a_0$	$-1 = a_1$	$1 = a_2$	$-\frac{5}{12} = a_3$	$\frac{107}{1260} = a_4$
2	0	1				
3	1	2	-1			
4	3	0	1	$\frac{2}{5}$		
5	6	3			$\frac{8}{45}$	

Tabuľka 5.8. Tabuľka diferencií pre $N_4(x)$

Newtonov interpolačný polynóm 4. stupňa:

$$\begin{aligned} N_4(x) &= a_0 + a_1(x-x_1) + a_2(x-x_1)(x-x_2) + a_3(x-x_1)(x-x_2)(x-x_3) + a_4(x-x_1)(x-x_2)(x-x_3)(x-x_4) \\ &= 2 - 1(x-(-1)) + 1(x-(-1))(x-0) - \frac{5}{12}(x-(-1))(x-0)(x-1) + \frac{107}{1260}(x-(-1))(x-0)(x-1)(x-3) \\ &= \frac{107}{1260}x^4 - \frac{47}{70}x^3 + \frac{1153}{1260}x^2 + \frac{47}{70}x + 1 \end{aligned}$$

Výpočet hľadaných hodnôt $f(2)$ a $f'(2)$:

$$f(2) \approx N_4(2) \Rightarrow f(2) \approx \frac{209}{105} \doteq 1,99048 \quad f'(2) \approx N_4'(2) \Rightarrow f'(2) \approx -\frac{127}{126} \doteq -1,00794$$

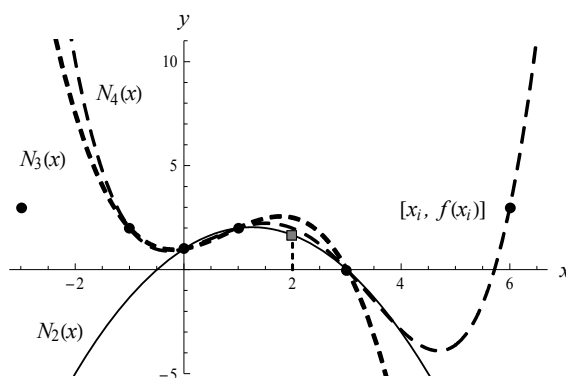
c) Absolútna chyba interpolácie v čísle $x = 2$: $|f(2) - N_2(2)| = |1,5 - 1,66667| = 0,166667$

$$|f(2) - N_3(2)| = |1,5 - 2,5| = 1$$

$$|f(2) - N_4(2)| = |1,5 - 1,99048| = 0,490476$$

Najmenšia absolútna chyba interpolácie v čísle $x = 2$ je pri aproximácii $f(x)$ interpolačným polynómom $N_2(x)$.

d) Graf funkcie $f(x)$ a interpolačných polynómov



Obr. 5.7. Graf funkcie $f(x)$ a interpolačných polynómov $N_2(x)$, $N_3(x)$ a $N_4(x)$

PRÍKLAD č. 2

S akou minimálnou presnosťou ε môžeme pomocou Newtonovho interpolačného polynómu určiť hodnotu $\sqrt{115}$ pri voľbe uzlov $x_1 = 100$, $x_2 = 121$, $x_3 = 144$? Výsledok vyjadrite na 4 desatinné miesta.

RIEŠENIE

Na určenie hodnoty $\sqrt{115}$ použijeme predpis funkcie $f(x)=\sqrt{x}$. Cez tri uzly $x_1=100, x_2=121, x_3=144$ môžeme preložiť Newtonov interpolačný polynóm maximálne 2. stupňa.

Použijeme odhad pre absolútnu chybu interpolácie, dostávame:

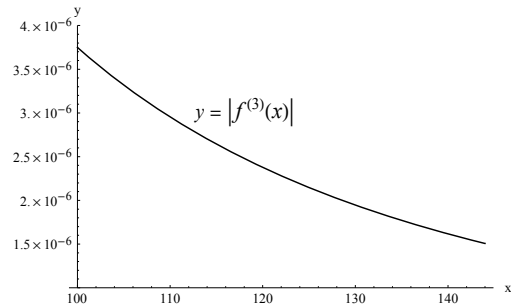
$$|f(x) - N_2(x)| \leq \frac{M_3}{3!} |(x-x_1)(x-x_2)(x-x_3)| < \varepsilon, \text{ kde } M_3 = \max |f^{(3)}(x)|, x \in <100, 144>$$

$$f^{(3)}(x) = \frac{3}{8\sqrt{x^5}}$$

$$M_3 = f^{(3)}(100) = \left| \frac{3}{8\sqrt{100^5}} \right| = 3,75 \cdot 10^{-6}$$

$$|f(x) - N_2(x)| \leq \frac{3,75 \cdot 10^{-6}}{3!} |(115-100)(115-121)(115-144)|$$

$$|f(x) - N_2(x)| \leq 0,00163125 < 0,0017$$



Obr. 5.8

Pri danej voľbe uzlov môžeme určiť hodnotu $\sqrt{115}$ pomocou Newtonovho interpolačného polynómu s minimálnou presnosťou $\varepsilon=1,7 \cdot 10^{-3}$.

PRÍKLAD č. 3

Funkcia $f(x)$ je daná tabuľkou Tabuľka 5.9. Nájdite Lagrangeov interpolačný polynóm pre uzlové body funkcie $f(x)$ a zobrazte ho. Zobrazte tiež fundamentálne polynómy $L_k(x)$.

i	0	1	2	3	4
x_i	-2	0	2	4	6
$y_i = f(x_i)$	2	1	-1	2	4

Tabuľka 5.9

RIEŠENIE

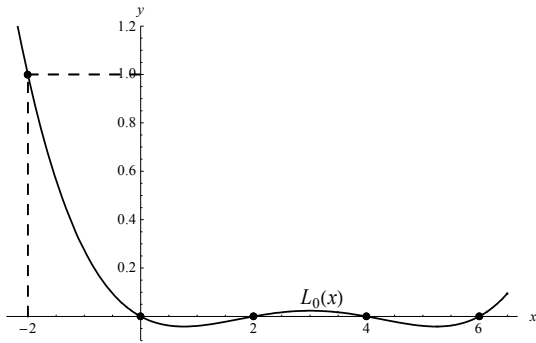
Funkcia $f(x)$ je daná 5 bodmi, hľadaný Lagrangeov interpolačný polynóm bude 4. stupňa:

$$P_4(x) = \sum_{k=0}^4 f(x_k) \cdot L_k(x) = f(x_0) \cdot L_0(x) + f(x_1) \cdot L_1(x) + f(x_2) \cdot L_2(x) + f(x_3) \cdot L_3(x) + f(x_4) \cdot L_4(x).$$

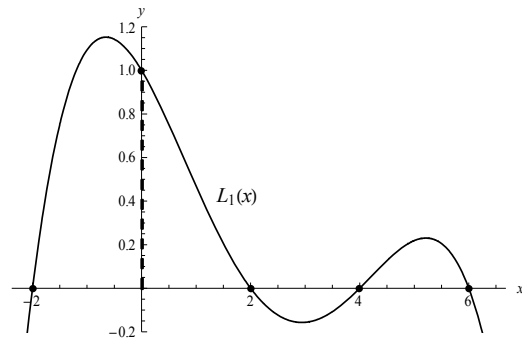
Nájdime fundamentálne polynómy $L_k(x)$, $k = 0, 1, \dots, 4$ (pozri (5.8)) pre zadané hodnoty a nakreslime ich grafy.

$$x_0 = -2: L_0(x) = \frac{(x-0) \cdot (x-2) \cdot (x-4) \cdot (x-6)}{(-2-0) \cdot (-2-2) \cdot (-2-4) \cdot (-2-6)} = \frac{x^4 - 12x^3 + 44x^2 - 48x}{384}$$

$$x_1 = 0: L_1(x) = \frac{(x-(-2)) \cdot (x-2) \cdot (x-4) \cdot (x-6)}{(0-(-2)) \cdot (0-2) \cdot (0-4) \cdot (0-6)} = \frac{-x^4 + 10x^3 - 20x^2 - 40x + 96}{96}$$



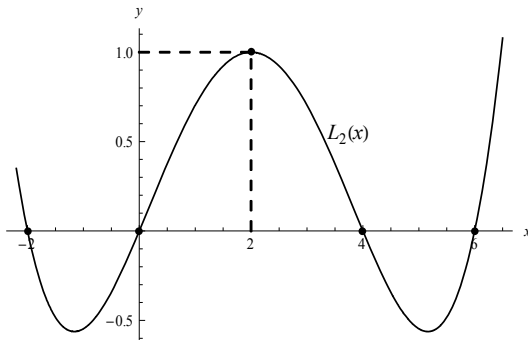
Obr. 5.9. Graf polynómu $L_0(x)$



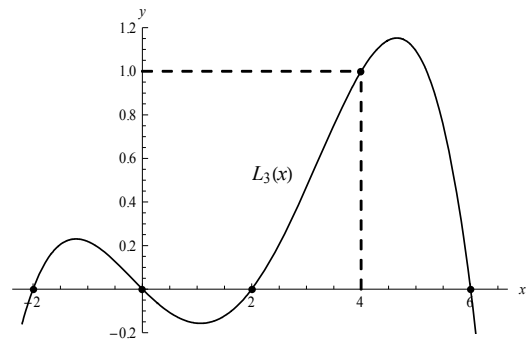
Obr. 5.10. Graf polynómu $L_1(x)$

$$x_2 = 2: L_2(x) = \frac{(x - (-2)) \cdot (x - 0) \cdot (x - 4) \cdot (x - 6)}{(2 - (-2)) \cdot (2 - 0) \cdot (2 - 4) \cdot (2 - 6)} = \frac{x^4 - 8x^3 + 4x^2 + 48x}{64}$$

$$x_3 = 4: L_3(x) = \frac{(x - (-2)) \cdot (x - 0) \cdot (x - 2) \cdot (x - 6)}{(4 - (-2)) \cdot (4 - 0) \cdot (4 - 2) \cdot (4 - 6)} = \frac{-x^4 + 6x^3 + 4x^2 - 24x}{96}$$



Obr. 5.11. Graf polynómu $L_2(x)$

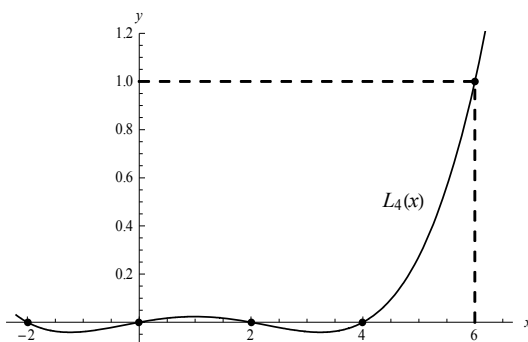


Obr. 5.12. Graf polynómu $L_3(x)$

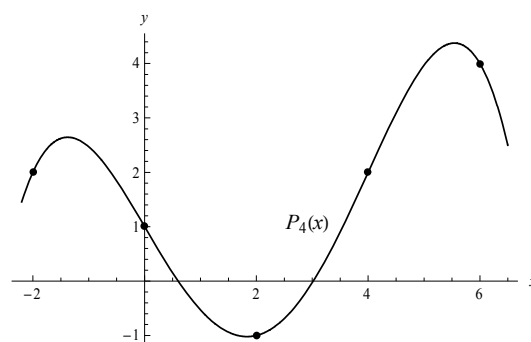
$$x_4 = 6: L_4(x) = \frac{(x - (-2)) \cdot (x - 0) \cdot (x - 2) \cdot (x - 4)}{(6 - (-2)) \cdot (6 - 0) \cdot (6 - 2) \cdot (6 - 4)} = \frac{x^4 - 4x^3 - 4x^2 + 16x}{384}$$

Potom hľadaný Lagrangeov interpolačný polynóm má tvar:

$$P_4(x) = 2 \cdot L_0(x) + 1 \cdot L_1(x) + (-1) \cdot L_2(x) + 2 \cdot L_3(x) + 4 \cdot L_4(x) = -\frac{x^4}{32} + \frac{x^3}{4} - \frac{7x}{4} + 1$$



Obr. 5.13. Graf polynómu $L_4(x)$



Obr. 5.14. Graf Lagrangeovho interpolačného polynómu $P_4(x)$

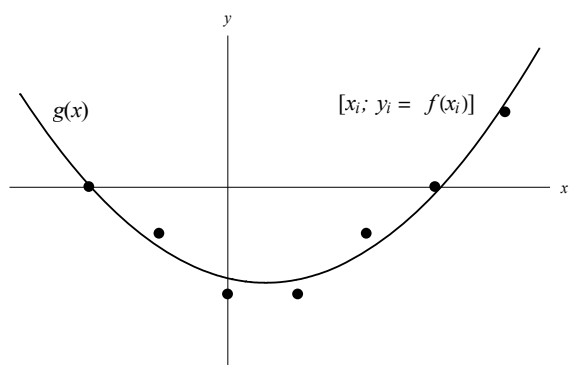
5.2 APROXIMÁCIA V ZMYSLE MNŠ

Majme nameraných $n+1$ navzájom rôznych hodnôt neznámej spojitej funkcie $f(x)$ (Tabuľka 5.1). Funkciu $f(x)$ chceme nahradiť aproximačnou funkciou $g(x)$, ktorá je v istom zmysle blízka funkcii $f(x)$. Pri interpolácii – špeciálnom prípade aproximácie, sme požadovali, aby graf funkcie $g(x)$ prechádzal všetkými známymi bodmi $[x_i, y_i]$, $i = 1, 2, \dots, n+1$ funkcie $f(x)$. Stotožnenie hodnôt funkcií v uzlových bodoch však nie vždy je vyhovujúce, najmä ak:

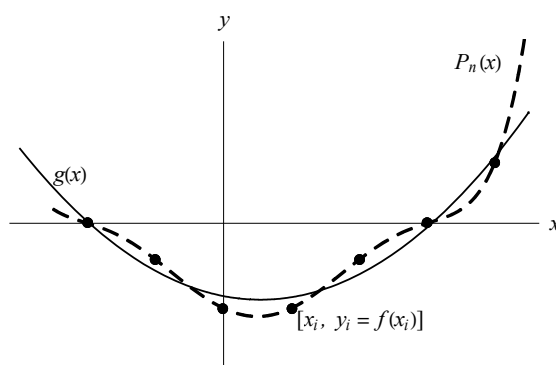
- údaje boli namerané experimentálne a sú zaťažené rôznymi chybami (chyba meracieho prístroja, presnosť merania, ...), tieto chyby sa potom premietajú aj do aproximačnej funkcie (Obr. 5.16);
- počet nameraných hodnôt je veľmi veľký, potom aj stupeň interpolačného polynómu je veľmi vysoký. Nájdenie interpolačného polynómu vysokého stupňa je náročné, podlieha šíreniu zaokrúhľovacích chýb a navyše takéto interpolačné polynómy kmitajú, čo v daných oblastiach značne znehodnocuje aproximáciu (Obr. 5.4);
- graf funkcie $f(x)$ obsahuje aj body, ktoré sú veľmi vzdialené od ostatných. Medzi nameranými dátami sú body, ktoré pravdepodobne obsahujú náhodnú chybu, v ich okolí je aproximačná funkcia znehodnotená;
- z praxe dopredu vieme, že skúmaná závislosť, funkcia $f(x)$, má istý charakter, napríklad lineárny, kvadratický, kubický, exponenciálny, a pod. Počet dát vtedy značne prekračuje počet bodov nutných na jednoznačné určenie danej aproximačnej funkcie a vhodnú interpolačnú funkciu nie je možné pomocou všetkých dát určiť.

V takýchto a podobných prípadoch od aproximačnej funkcie $g(x)$ požadujeme, aby dostatočne presne kopírovala priebeh funkcie $f(x)$, graf $g(x)$ nemusí prechádzať jednotlivými bodmi $[x_i, y_i]$, $i = 1, 2, \dots, n+1$, stačí, keď je dostatočne blízko. Miera podobnosti, vzdialenosť funkcií $f(x)$ a $g(x)$ je v zmysle definície Definícia 5.3 minimálna (Obr. 5.16). Pri konštrukcii aproximačnej funkcie budeme prihliadať:

- na druh aproximačnej funkcie (polynomická, racionálna, exponenciálna, a pod.),
- na minimálnu vzdialenosť medzi funkciami $f(x)$ a $g(x)$.



Obr. 5.15. Aproximácia polynómom $g(x)$



Obr. 5.16. Aproximácia interpolačným polynómom $P_n(x)$ a aproximačným polynómom $g(x)$

Aproximačnú funkciu $g(x)$ hľadáme v tvare:

$$g(x) = \sum_{i=0}^m a_i \varphi_i(x) = a_0 \varphi_0(x) + a_1 \varphi_1(x) + a_2 \varphi_2(x) + \dots + a_m \varphi_m(x),$$

kde $a_i \in R$ a $\varphi_i(x)$ sú funkcie.

Medzi najčastejšie používané aproximácie patrí aproximácia polynómom alebo aproximácia trigonometrickou funkciou, trigonometrickým polynómom.

Pri aproximácii polynómom systém funkcií $\{\varphi_0, \varphi_1, \varphi_2, \dots, \varphi_m\}$ nahradíme systémom funkcií $\{x^0, x^1, x^2, \dots, x^m\}$, t.j.

$$g(x) = \sum_{i=0}^m a_i x^i = a_0 x^0 + a_1 x^1 + a_2 x^2 + \dots + a_m x^m.$$

Aproximačný polynóm je stupňa m a pri zostupnom usporiadaní mocnín má vyjadrenie

$$P_m(x) = a_m x^m + a_{m-1} x^{m-1} + \dots + a_2 x^2 + a_1 x + a_0.$$

Aproximáciu trigonometrickým polynómom používame, keď funkcia $f(x)$ má periodický charakter. Vtedy systém funkcií $\{\varphi_0, \varphi_1, \varphi_2, \dots, \varphi_m\}$ nahradíme systémom funkcií $\{1, \cos(x), \sin(x), \cos(2x), \sin(2x), \dots, \cos(mx), \sin(mx)\}$:

$$g(x) = \sum_{i=0}^m a_i \cos(ix) + \sum_{i=0}^m b_i \sin(ix).$$

5.2.1 MNŠ – APROXIMÁCIA POLYNÓMOM

Majme nameraných $n+1$ navzájom rôznych hodnôt neznámej spojitej funkcie $f(x)$ (Tabuľka 5.1), ktoré chceme aproximovať funkciou $g(x)$, pričom požadujeme:

- $g(x) = P(x)$, $P(x)$ – polynóm,
- vzdialenosť $d(f, g)$ je minimálna.

Otázne je, ako zvoliť stupeň aproximačného polynómu $P(x)$, pretože:

- musí byť dostatočne vysoký, aby bol dobrou aproximáciou funkcie $f(x)$,
- zároveň by nemal byť príliš vysoký, aby nezachovával chyby merania.

Zvyčajne vychádzame z tvaru grafu $f(x)$, pričom zvolený aproximačný polynóm je najviac 5. stupňa. Polynómy 6. stupňa a viac používame iba vtedy, ak predpokladáme, že dáta funkcie $f(x)$ majú polynomiálny charakter.

Definícia 5.6 Hovoríme, že polynóm $P_m(x) = a_m x^m + a_{m-1} x^{m-1} + \dots + a_2 x^2 + a_1 x + a_0$, $a_i \in R$, $a_m \neq 0$, $m \leq n$ je aproximačným polynómom množiny bodov $[x_i; y_i = f(x_i)]$, $i = 1, 2, \dots, n+1$ v zmysle MNŠ, keď je súčet odchýlok

$$\sum_{i=1}^{n+1} (y_i - P_m(x_i))^2 = \sum_{i=1}^{n+1} (P_m(x_i) - y_i)^2 \quad (5.9)$$

minimálny.

Hľadáme, kedy (5.9) nadobúda minimum. Vyjadrime:

$$\sum_{i=1}^{n+1} (P_m(x_i) - y_i)^2 = (P_m(x_1) - y_1)^2 + (P_m(x_2) - y_2)^2 + \dots + (P_m(x_{n+1}) - y_{n+1})^2$$

$$\sum_{i=1}^{n+1} (P_m(x_i) - y_i)^2 = \sum_{i=1}^{n+1} (a_m x_i^m + \dots + a_2 x_i^2 + a_1 x_i + a_0 - y_i)^2 \quad (5.10)$$

Tu je potrebné si uvedomiť, že hodnoty x_i, y_i poznáme (namerané hodnoty), ale nepoznáme koeficienty $a_m, \dots, a_2, a_1, a_0$ polynómu $P_m(x)$, ktorými je tento polynóm určený a ktoré chceme vypočítať. Výraz (5.9) resp. (5.10) je funkciou koeficientov $a_m, \dots, a_2, a_1, a_0$ polynómu $P_m(x)$:

$$S(a_m, a_{m-1}, \dots, a_1, a_0) = \sum_{i=1}^{n+1} (a_m x_i^m + \dots + a_2 x_i^2 + a_1 x_i + a_0 - y_i)^2. \quad (5.11)$$

Koeficienty $a_m, \dots, a_2, a_1, a_0$ polynómu $P_m(x)$ určíme tak, že minimalizujeme funkciu (5.11). Vychádzajúc z diferenciálneho počtu funkcie viac premenných, hľadáme lokálne minimum funkcie $S = S(a_m, \dots, a_2, a_1, a_0)$. Funkcia S nadobudne minimum v stacionárnom bode $[a_m, \dots, a_2, a_1, a_0]$, ktorý je riešením systému **normálnych (normálových) rovníc**:

$$\frac{\delta S(a_m, \dots, a_0)}{\delta a_i} = 0 \Leftrightarrow \frac{\delta S(a_m, \dots, a_0)}{\delta a_m} = 0 \wedge \frac{\delta S(a_m, \dots, a_0)}{\delta a_{m-1}} = 0 \wedge \dots \wedge \frac{\delta S(a_m, \dots, a_0)}{\delta a_0} = 0 \quad (5.12)$$

t.j.

$$\frac{\delta S(a_m, \dots, a_0)}{\delta a_m} = 0 \Leftrightarrow \sum_{i=1}^{n+1} 2(a_m x_i^m + \dots + a_2 x_i^2 + a_1 x_i + a_0 - y_i) \cdot x_i^m = 0$$

$$\frac{\delta S(a_m, \dots, a_0)}{\delta a_{m-1}} = 0 \Leftrightarrow \sum_{i=1}^{n+1} 2(a_m x_i^m + a_{m-1} x_i^{m-1} + \dots + a_1 x_i + a_0 - y_i) \cdot x_i^{m-1} = 0$$

.....

$$\frac{\delta S(a_m, \dots, a_0)}{\delta a_2} = 0 \Leftrightarrow \sum_{i=1}^{n+1} 2(a_m x_i^m + \dots + a_2 x_i^2 + a_1 x_i + a_0 - y_i) \cdot x_i^2 = 0$$

$$\frac{\delta S(a_m, \dots, a_0)}{\delta a_1} = 0 \Leftrightarrow \sum_{i=1}^{n+1} 2(a_m x_i^m + \dots + a_2 x_i^2 + a_1 x_i + a_0 - y_i) \cdot x_i = 0$$

$$\frac{\delta S(a_m, \dots, a_0)}{\delta a_0} = 0 \Leftrightarrow \sum_{i=1}^{n+1} 2(a_m x_i^m + \dots + a_2 x_i^2 + a_1 x_i + a_0 - y_i) \cdot 1 = 0$$

Po úprave:

$$a_m \sum_{i=1}^{n+1} x_i^{2m} + a_{m-1} \sum_{i=1}^{n+1} x_i^{2m-1} + \dots + a_1 \sum_{i=1}^{n+1} x_i^{m-1} + a_0 \sum_{i=1}^{n+1} x_i^m = \sum_{i=1}^{n+1} x_i^m y_i$$

$$a_m \sum_{i=1}^{n+1} x_i^{2m-1} + a_{m-1} \sum_{i=1}^{n+1} x_i^{2m-2} + \dots + a_1 \sum_{i=1}^{n+1} x_i^{m-2} + a_0 \sum_{i=1}^{n+1} x_i^{m-1} = \sum_{i=1}^{n+1} x_i^{m-1} y_i$$

.....

$$\begin{aligned}
 a_m \sum_{i=1}^{n+1} x_i^{m+1} + a_{m-1} \sum_{i=1}^{n+1} x_i^m + \dots + a_1 \sum_{i=1}^{n+1} x_i^3 + a_0 \sum_{i=1}^{n+1} x_i^2 &= \sum_{i=1}^{n+1} x_i^2 y_i & (5.13) \\
 a_m \sum_{i=1}^{n+1} x_i^{m+1} + a_{m-1} \sum_{i=1}^{n+1} x_i^m + \dots + a_1 \sum_{i=1}^{n+1} x_i^2 + a_0 \sum_{i=1}^{n+1} x_i^1 &= \sum_{i=1}^{n+1} x_i^1 y_i \\
 a_m \sum_{i=1}^{n+1} x_i^m + a_{m-1} \sum_{i=1}^{n+1} x_i^{m-1} + \dots + a_1 \sum_{i=1}^{n+1} x_i^1 + a_0 \sum_{i=1}^{n+1} x_i^0 &= \sum_{i=1}^{n+1} x_i^0 y_i
 \end{aligned}$$

a zavedení označenia:

$$\sum_{i=1}^{n+1} 1 = \sum_{i=1}^{n+1} x_i^0 = [x^0], \quad \sum_{i=1}^{n+1} x_i^k = [x^k], \quad \sum_{i=1}^{n+1} x_i^k y_i = [x^k y] \quad (5.14)$$

dostávame systém (5.13) v maticovom tvare:

$$\begin{pmatrix} [x^{2m}] & [x^{2m-1}] & \dots & [x^{m+1}] & [x^m] \\ [x^{2m-1}] & [x^{2m-2}] & \dots & [x^m] & [x^{m-1}] \\ \dots & \dots & \dots & \dots & \dots \\ [x^{m+2}] & [x^{m+1}] & \dots & [x^3] & [x^2] \\ [x^{m+1}] & [x^m] & \dots & [x^2] & [x^1] \\ [x^m] & [x^{m-1}] & \dots & [x^1] & [x^0] \end{pmatrix} \cdot \begin{pmatrix} a_m \\ a_{m-1} \\ \dots \\ a_2 \\ a_1 \\ a_0 \end{pmatrix} = \begin{pmatrix} [x^m y] \\ [x^{m-1} y] \\ \dots \\ [x^2 y] \\ [x^1 y] \\ [x^0 y] \end{pmatrix} \quad (5.15)$$

Systém (5.15), resp.(5.13), či (5.12) je systémom $m+1$ lineárnych rovníc o $m+1$ neznámych. Jeho riešením sú súradnice lokálneho minima $[a_m, \dots, a_2, a_1, a_0]$ funkcie S , kde $a_m, \dots, a_2, a_1, a_0$ sú koeficienty polynómu $P_m(x)$. Funkcia S je nezáporná, dá sa ukázať, že vo vypočítanom stacionárnom bode $[a_m, \dots, a_2, a_1, a_0]$ funkcia S nadobúda lokálne (globálne) minimum.

i	x_i	y_i	x_i^2	x_i^3	...	x_i^{2m}	$x_i y_i$	$x_i^2 y_i$...	$x_i^m y_i$
1	x_1	y_1	x_1^2	x_1^3	...	x_1^{2m}	$x_1 y_1$	$x_1^2 y_1$...	$x_1^m y_1$
2	x_2	y_2	x_2^2	x_2^3	...	x_2^{2m}	$x_2 y_2$	$x_2^2 y_2$...	$x_2^m y_2$
3	x_3	y_3	x_3^2	x_3^3	...	x_3^{2m}	$x_3 y_3$	$x_3^2 y_3$...	$x_3^m y_3$
...
$n+1 = [x^0]$	x_{n+1}	y_{n+1}	x_{n+1}^2	x_{n+1}^3	...	x_{n+1}^{2m}	$x_{n+1} y_{n+1}$	$x_{n+1}^2 y_{n+1}$...	$x_{n+1}^m y_{n+1}$
$\sum_{i=1}^{n+1}$	$[x^1]$	$[x^0 y]$	$[x^2]$	$[x^3]$...	$[x^{2m}]$	$[x^1 y]$	$[x^2 y]$...	$[x^m y]$

Tabuľka 5.10. Tabuľka na výpočet koeficientov systému (5.15)

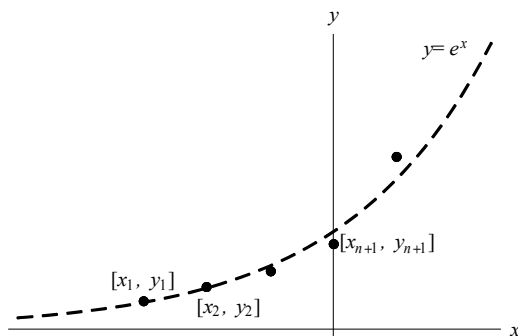
V prípade, keď počet nameraných hodnôt $n+1$ funkcie $f(x)$ je o 1 väčší ako stupeň aproximačného polynómu m , t.j. $n+1 = m+1$, dostávame **špeciálny prípad aproximácie – interpoláciu** polynómom $P_m(x)$.

5.2.2 MNŠ – APROXIMÁCIA EXPONENCIÁLNOU FUNKCIOU

Majme nameraných $n+1$ navzájom rôznych hodnôt neznámej spojitej funkcie $f(x)$

x_i	x_1	x_2	\dots	x_{n+1}
$y_i = f(x_i)$	y_1	y_2	\dots	y_{n+1}

Zobrazme tieto namerané hodnoty.



Obr. 5.17

Vidíme, že graf nameraných hodnôt $[x_i; y_i = f(x_i)]$ sa podobá na graf exponenciálnej (nelineárnej) funkcie $y = e^x$. Aproximačnú funkciu $g(x)$ budeme preto hľadať v tvare $g(x) = C e^{Ax}$, $C, A \in \mathbb{R} - \{0\}$.

Aproximácia funkcie $f(x)$ funkciou $g(x)$ metódou MNŠ spočíva v určení koeficientov C, A funkcie $g(x) = C e^{Ax}$, ktoré minimalizujú funkciu

$$S(C, A) = \sum_{i=1}^{n+1} (C e^{Ax_i} - y_i)^2$$

Opäť hľadáme lokálne minimum (stacionárny bod) funkcie $S = S(C, A)$, ktorý je riešením systému normálnych rovníc:

$$\begin{aligned} \frac{\delta S(C, A)}{\delta C} = 0 \quad \wedge \quad \frac{\delta S(C, A)}{\delta A} = 0 \\ \frac{\delta S(C, A)}{\delta C} = 0 \Leftrightarrow \sum_{i=1}^{n+1} 2 \cdot (C e^{Ax_i} - y_i) \cdot e^{Ax_i} = 0 \\ \frac{\delta S(C, A)}{\delta A} = 0 \Leftrightarrow \sum_{i=1}^{n+1} 2 \cdot (C e^{Ax_i} - y_i) \cdot C e^{Ax_i} \cdot x_i = 0 \end{aligned}$$

Po úprave dostávame

$$\begin{aligned} C \sum_{i=1}^{n+1} e^{2Ax_i} &= \sum_{i=1}^{n+1} e^{Ax_i} y_i \\ C^2 \sum_{i=1}^{n+1} e^{2Ax_i} x_i &= C \sum_{i=1}^{n+1} e^{Ax_i} x_i y_i \end{aligned} \tag{5.16}$$

Vidíme, že systém rovníc (5.16) je nelineárny. Jedna z možností, ako vypočítať koeficienty C a A vo funkcii $g(x) = C e^{Ax}$, sú numerické metódy na riešenie systémov nelineárnych rovníc (iteračné metódy, nie sú obsahom predmetu) alebo na intervaloch, kde funkcia má približne lineárny charakter, môžeme použiť *linearizáciu* funkcií $f(x)$ a $g(x)$. Funkciu $g(x)$ transformujeme vhodným spôsobom na lineárnu funkciu typu „ $y = ax + b$ “. Spoločne s transformáciou funkcie $g(x)$ vykonáme aj transformáciu nameraných dát funkcie $f(x)$.

Najsôr vykonáme linearizáciu exponenciálnej funkcie $g(x) = C e^{Ax}$.

$$y = C e^{Ax} \quad / \ln() \text{ (rovniciu logaritmujeme pri základe } e)$$

$$\ln(y) = \ln(C e^{Ax})$$

$$\ln(y) = \ln(C) + \ln(e^{Ax})$$

$$\ln(y) = \ln(C) + Ax \ln(e)$$

$$\ln(y) = \ln(C) + Ax$$

$$Y = B + Ax$$

Dostávame už lineárnu funkciu, polynóm $P_1(x) = Y = B + Ax$.

Následne linearizujeme aj namerané hodnoty $[x_i; y_i = f(x_i)] \rightarrow [x_i; Y_i = \ln(y_i)]$.

	Aproximačná $g(x)$	Množina dát $f(x)$	Transformačné vzťahy
Pred linearizáciou	$y = C e^{Ax}$	$[x_i; y_i = f(x_i)]$	$x \rightarrow x$
Linearizácia	$\ln(y) = \ln(C e^{Ax})$ $\ln(y) = \ln(C) + Ax$	$Y_i = \ln(y_i)$	$y \rightarrow Y \wedge Y = \ln(y)$ $A \rightarrow A$
Po linearizácii	$Y = B + Ax$	$[x_i; Y_i]$	$C \rightarrow e^B \wedge B = \ln(C)$

Tabuľka 5.11. Postup pri linearizácii $f(x)$ a $g(x)$

Ďalej pokračujeme aproximáciou transformovanej množiny dát $[x_i; Y_i], i = 1, 2, \dots, n+1$ polynómom prvého stupňa (pozri odsek 5.2.1). Hľadáme lokálne minimum funkcie:

$$S(A, B) = \sum_{i=1}^{n+1} (Ax_i + B - Y_i)^2$$

riešením systému normálnych rovníc:

$$\frac{\delta S(A, B)}{\delta A} = 0 \Leftrightarrow \sum_{i=1}^{n+1} 2 \cdot (Ax_i + B - Y_i) \cdot x_i = 0$$

$$\frac{\delta S(A, B)}{\delta B} = 0 \Leftrightarrow \sum_{i=1}^{n+1} 2 \cdot (Ax_i + B - Y_i) \cdot 1 = 0$$

Po úprave dostávame:

$$A \sum_{i=1}^{n+1} x_i^2 + B \sum_{i=1}^{n+1} x_i = \sum_{i=1}^{n+1} x_i^1 Y_i \tag{5.17}$$

$$A \sum_{i=1}^{n+1} x_i^1 + B \sum_{i=1}^{n+1} x_i^0 = \sum_{i=1}^{n+1} x_i^0 Y_i$$

a po zavedení označenia (5.14) zapíšeme systém (5.17) v maticovom tvare:

$$\begin{pmatrix} [x^2] & [x^1] \\ [x^1] & [x^0] \end{pmatrix} \cdot \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} [x^1 Y] \\ [x^0 Y] \end{pmatrix} \quad (5.18)$$

Systém (5.18) je už lineárnym systémom 2 rovníc o 2 neznámych. Jeho riešením sú koeficienty A a B polynómu $P_1(x)$.

Po aproximácii vykonáme spätnú transformáciu. Lineárnu funkciu $P_1(x) = Y = Ax + B$ spätne „odlinearizujeme“ – transformujeme na pôvodnú (nelineárnu) exponenciálnu funkciu $g(x) = y = C e^{Ax}$ a vypočítame koeficient C .

$$\begin{aligned} Y &= Ax + B & / \text{ dosadíme } Y = \ln(y) \\ \ln(y) &= Ax + B \\ y &= e^{Ax+B} \Rightarrow y = e^{Ax} \cdot e^B \\ & & y = e^{Ax} \cdot C = C \cdot e^{Ax} \end{aligned}$$

Je dobré si uvedomiť, že linearizáciou síce dostávame rozumnú aproximáciu, avšak nemusí to byť optimálna aproximácia v zmysle pôvodnej nelineárnej aproximácie. Na intervaloch, kde má funkcia $f(x)$ veľmi nelineárny priebeh, linearizácia spôsobuje veľké odchýlky a výsledná aproximácia potom vykazuje veľkú chybu.

Chyby pri aproximácii funkcie $f(x)$ funkciou $g(x)$

1. **absolútna chyba** aproximácie v ľubovoľnom uzle x_i : $\Delta y_i = |f(x_i) - g(x_i)|$,

2. **celková chyba** aproximácie:

a) **maximová**: $E_\infty = \max_{i=1,2,\dots,n+1} |f(x_i) - g(x_i)|$,

veľmi jednoduchá na výpočet, stačí však jedna zle nameraná hodnota, ktorá znehodnotí jej výsledok,

b) **priemerová**: $E = \frac{1}{n+1} \sum_{i=1}^{n+1} |f(x_i) - g(x_i)|$,

veľmi jednoduchá na výpočet, všetky namerané hodnoty majú rovnaký vplyv na jej výsledok,

c) **RMS** (z angl. Root Mean Square, štandardná odchýlka): $\varepsilon = \sqrt{\frac{1}{n+1} \sum_{i=1}^{n+1} (f(x_i) - g(x_i))^2}$

náročná na výpočet, potláča vplyv malých odchýlok a zdôrazňuje vplyv veľkých odchýlok nameraných hodnôt od aproximačnej funkcie.

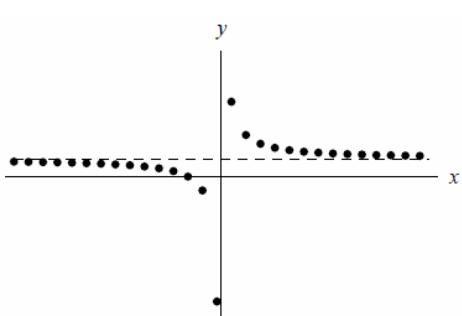
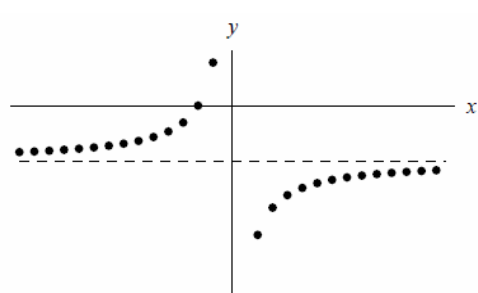
Postup pri aproximácii $f(x)$ funkciou $g(x)$:

1. Zobrazíme graf $f(x)$: $\{[x_i; y_i = f(x_i)], i = 1, 2, \dots, n + 1\}$.
2. Na základe grafu $f(x)$ vyberieme vhodný typ aproximačnej funkcie $g(x)$, ktorá v zmysle MNŠ čo možno najlepšie aproximuje namerané údaje.

3. Rozhodneme, či $g(x)$ má polynomiálny charakter:
 - a) Ak $g(x) = P(x)$, aproximujeme polynómom (odsek 5.2.1).
 - b) Ak $g(x) \neq P(x)$, aproximujeme vhodnou nelineárnou funkciou (odsek 5.2.2). Rozhodneme, či môžeme použiť linearizáciu.
 - i) Ak môžeme použiť linearizáciu:
 - linearizujeme $g(x)$ a $f(x)$
 - aproximujeme linearizovanú funkciu $f(x)$ polynómom 1. stupňa (priamkou). (Vypočítame koeficienty A a B linearizovanej funkcie $g(x)$).
 - vykonáme spätnú transformáciu linearizovanej funkcie $g(x)$ na pôvodnú (nelineárnou) funkciu $g(x)$.
 - ii) Ak linearizácia nie je vhodná, aproximujeme nelineárnou MNS (nie je obsahom predmetu).
4. Vypočítame celkovú chybu aproximácie.

5.2.3 APROXIMÁCIA ĎALŠÍMI FUNKCIAMI POMOCOU LINEARIZÁCIE

Postup, ktorým sme nelineárnou funkciu $g(x) = y = C e^{-Ax}$ transformovali na lineárnu funkciu $P_1(x) = Y = Ax + B$ môžeme aplikovať aj na ďalšie nelineárne funkcie. Aproximáciu vždy robíme na intervale, kde $g(x)$ je spojitá.

1.	Aproximačná $g(x)$	Množina dát $f(x)$	Transformačné vzťahy
Pred linearizáciou	$y = \frac{A}{x} + B$	$[x_i; y_i = f(x_i)]$	$x \rightarrow X \wedge X = \frac{1}{x}$
Linearizácia	$y = A \frac{1}{x} + B$	$X_i = \frac{1}{x_i}$	$y \rightarrow Y$ $A \rightarrow A$
Po linearizácii	$y = AX + B$	$[X_i; Y_i]$	$B \rightarrow B$
Graf nameraných hodnôt $[x_i; y_i = f(x_i)]$			
$A > 0 \wedge B > 0$		$A < 0 \wedge B < 0$	
			

Tabuľka 5.12

2.	Aproximačná $g(x)$	Množina dát $f(x)$	Transformačné vzťahy
Pred linearizáciou	$y = \frac{A}{x+B}$	$[x_i; y_i = f(x_i)]$	$x \rightarrow X \wedge X = x \cdot y$ $y \rightarrow y$ $A \rightarrow -\frac{D}{C} \wedge D = \frac{A}{\left(-\frac{1}{C}\right)}$ $B \rightarrow -\frac{1}{C} \wedge C = -\frac{1}{B}$
Linearizácia	$y = \frac{\frac{A}{B}}{\left(\frac{x}{B} + 1\right)}$ $y \left(\frac{x}{B} + 1\right) = \frac{A}{B}$ $y = -\frac{1}{B}(xy) + \frac{A}{B}$	$X_i = x_i y_i$	
Po linearizácii	$y = CX + D$	$[X_i; y_i]$	
Graf nameraných hodnôt $[x_i; y_i = f(x_i)]$			
$A > 0 \wedge B > 0$		$A > 0 \wedge B < 0$	

Tabuľka 5.13

3.	Aproximačná $g(x)$	Množina dát $f(x)$	Transformačné vzťahy
Pred linearizáciou	$y = \frac{1}{Ax+B}$	$[x_i; y_i = f(x_i)]$	$x \rightarrow x$ $y \rightarrow Y \wedge Y = \frac{1}{y}$ $A \rightarrow A$ $B \rightarrow B$
Linearizácia	$\frac{1}{y} = Ax + B$	$Y_i = \frac{1}{y_i}$	
Po linearizácii	$Y = Ax + B$	$[x_i; Y_i]$	
Graf nameraných hodnôt $[x_i; y_i = f(x_i)]$			
$A > 0 \wedge B > 0$		$A < 0 \wedge B > 0$	

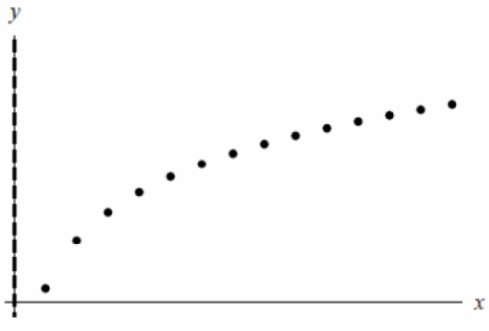
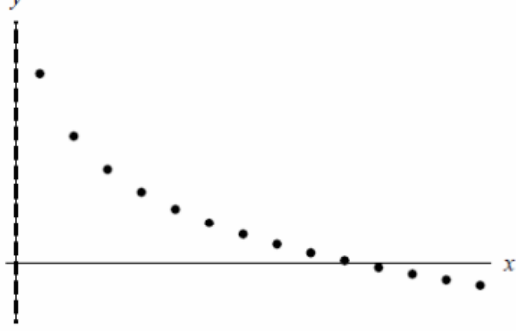
Tabuľka 5.14

4.	Aproximačná $g(x)$	Množina dát $f(x)$	Transformačné vzťahy
Pred linearizáciou	$\frac{1}{y} = \frac{A}{B+x} + C$	$[x_i; y_i = f(x_i)]$	$x \rightarrow x$ $y \rightarrow Y \wedge Y = \frac{1}{\left(\frac{1}{y} - C\right)}$ $A \rightarrow \frac{1}{E}$ $B \rightarrow \frac{D}{E}$
Linearizácia	$\frac{1}{y} - C = \frac{A}{B+x}$ $\frac{1}{\left(\frac{1}{y} - C\right)} = \frac{B+x}{A}$ $\frac{1}{\left(\frac{1}{y} - C\right)} = \frac{B}{A} + \frac{1}{A} \cdot x$	$Y_i = \frac{1}{\left(\frac{1}{y_i} - C\right)}$	
Po linearizácii	$Y = D + Ex$	$[X_i; Y_i]$	
Graf nameraných hodnôt $[x_i; y_i = f(x_i)]$			
$A > 0 \wedge B > 0 \wedge C > 0$		$A < 0 \wedge B < 0 \wedge C < 0$	

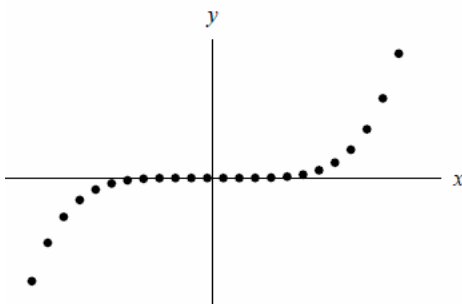

Tabuľka 5.15

5.	Aproximačná $g(x)$	Množina dát $f(x)$	Transformačné vzťahy
Pred linearizáciou	$y = \frac{1}{(Ax+B)^2}$	$[x_i; y_i = f(x_i)]$	$x \rightarrow x$ $y \rightarrow Y \wedge Y = \frac{1}{\sqrt{y}}$ $A \rightarrow A$ $B \rightarrow B$
Linearizácia	$\frac{1}{y} = (Ax+B)^2$ $\frac{1}{\sqrt{y}} = Ax+B $	$Y_i = \frac{1}{\sqrt{y_i}}$	
Po linearizácii	$Y = Ax+B $	$[x_i; Y_i]$	
Graf nameraných hodnôt $[x_i; y_i = f(x_i)]$			
$A > 0 \wedge B > 0$		$A < 0 \wedge B > 0$	

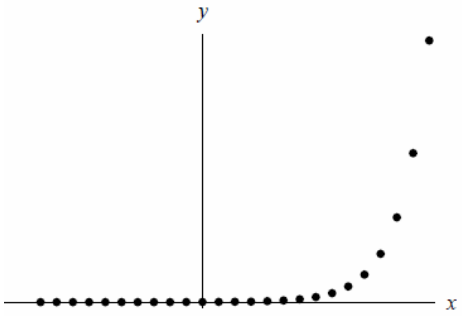
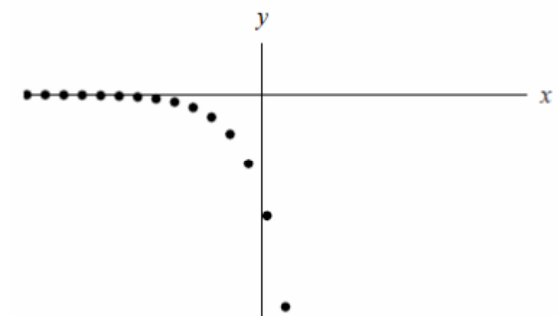
Tabuľka 5.16

6.	Aproximačná $g(x)$	Množina dát $f(x)$	Transformačné vzťahy
Pred linearizáciou	$y = A \ln(x) + B$	$[x_i; y_i = f(x_i)]$	$x \rightarrow X \wedge X = \ln(x)$
Linearizácia	$y = A \ln(x) + B$	$X_i = \ln(x_i)$	$y \rightarrow y$
Po linearizácii	$y = AX + B$	$[X_i; y_i]$	$A \rightarrow A$ $B \rightarrow B$
Graf nameraných hodnôt $[x_i; y_i = f(x_i)]$			
$A > 0 \wedge B > 0$		$A < 0 \wedge B < 0$	
			

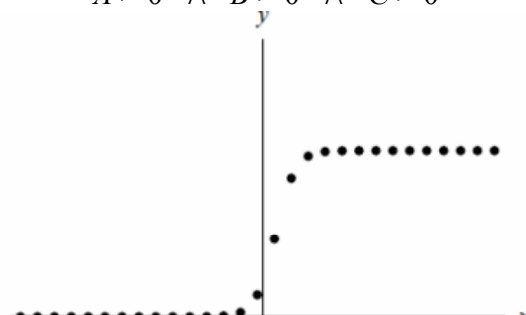
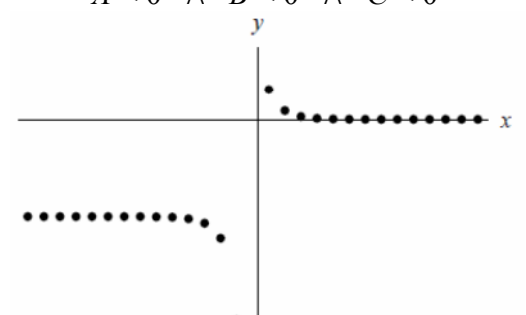
Tabuľka 5.17

7.	Aproximačná $g(x)$	Množina dát $f(x)$	Transformačné vzťahy
Pred linearizáciou	$y = B x^A$	$[x_i; y_i = f(x_i)]$	$x \rightarrow X \wedge X = \ln(x)$
Linearizácia	$\ln(y) = \ln(B x^A)$ $\ln(y) = \ln(B) + A \ln(x)$	$X_i = \ln(x_i)$ $Y_i = \ln(y_i)$	$y \rightarrow Y \wedge Y = \ln(y)$ $A \rightarrow A$
Po linearizácii	$Y = C + AX$	$[X_i; Y_i]$	$B \rightarrow e^C \wedge C = \ln(B)$
Graf nameraných hodnôt $[x_i; y_i = f(x_i)]$			
$A > 0 \wedge B > 0$		$A < 0 \wedge B < 0$	
			

Tabuľka 5.18

8.	Aproximačná $g(x)$	Množina dát $f(x)$	Transformačné vzťahy
Pred linearizáciou	$y = A \cdot B^x$ ($B > 0, B \neq 1$)	$[x_i; y_i = f(x_i)]$	$x \rightarrow x$ $y \rightarrow Y \wedge Y = \ln(y)$
Linearizácia	$\ln(y) = \ln(A \cdot B^x)$ $\ln(y) = \ln(A) + x \cdot \ln(B)$	$Y_i = \ln(y_i)$	$A \rightarrow e^C \wedge C = \ln(A)$ $B \rightarrow e^D \wedge D = \ln(B)$
Po linearizácii	$Y = C + Dx$	$[X_i; Y_i]$	
Graf nameraných hodnôt $[x_i; y_i = f(x_i)]$			
$A > 0 \wedge B > 0$		$A < 0 \wedge B > 0$	
			

Tabuľka 5.19

9.	Aproximačná $g(x)$	Množina dát $f(x)$	Transformačné vzťahy
Pred linearizáciou	$y = \frac{C}{1 + B e^{-Ax}}$	$[x_i; y_i = f(x_i)]$	$x \rightarrow x$ $y \rightarrow Y \wedge Y = \ln\left(\frac{C}{y} - 1\right)$
Linearizácia	$\frac{C}{y} - 1 = B e^{-Ax}$ $\ln\left(\frac{C}{y} - 1\right) = \ln(B e^{-Ax})$ $\ln\left(\frac{C}{y} - 1\right) = \ln(B) + (-A)x$	$Y_i = \ln\left(\frac{C}{y_i} - 1\right)$	$A \rightarrow -D$ $B \rightarrow e^E \wedge E = \ln(B)$ $C \rightarrow C$
Po linearizácii	$Y = E + Dx$	$[X_i; Y_i]$	
Graf nameraných hodnôt $[x_i; y_i = f(x_i)]$			
$A > 0 \wedge B > 0 \wedge C > 0$		$A < 0 \wedge B < 0 \wedge C < 0$	
			

Tabuľka 5.20

5.2.4 RIEŠENÉ PRÍKLADY

PRÍKLAD č. 1

Funkcia $f(x)$ je daná tabuľkou:

i	1	2	3	4
x_i	0	1	2	3
$y_i = f(x_i)$	1,05	2,99	4,52	6,98

Tabuľka 5.21

- Funkciu $f(x)$ aproximujte v zmysle metódy najmenších štvorcov polynómom vhodne zvoleného stupňa.
- Vypočítajte absolútnu chybu aproximácie v uzle x_2 .
- Vypočítajte celkovú chybu aproximácie.

RIEŠENIE

a) *Aproximácia v zmysle MNS polynómom vhodného stupňa*

Funkcia $f(x)$ je daná 4 bodmi:

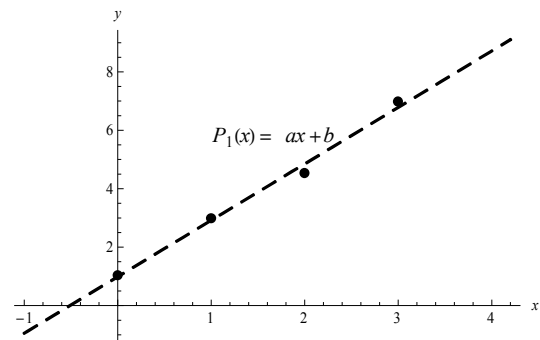
$$X_1 = [0; 1,05]$$

$$X_2 = [1; 2,99]$$

$$X_3 = [2; 4,52]$$

$$X_4 = [3; 6,98]$$

Z grafu vidíme, že hodnoty $f(x)$ ležia takmer na priamke, preto $f(x)$ budeme aproximovať polynómom 1. stupňa $P_1(x) = ax + b$.



Obr. 5.18. Graf $f(x)$ a $P_1(x)$

Chceme vypočítať koeficienty a , b polynómu $P_1(x)$.

Minimalizujeme funkciu:

$$S(a, b) = \sum_{i=1}^4 (P_1(x_i) - y_i)^2 = \sum_{i=1}^4 (ax_i + b - y_i)^2$$

Riešime systém normálnych rovníc:

$$\frac{\partial S(a, b)}{\partial a} = 0 \quad \wedge \quad \frac{\partial S(a, b)}{\partial b} = 0$$

$$\frac{\partial S(a, b)}{\partial a} = 0 \quad \Leftrightarrow \quad \sum_{i=1}^4 2 \cdot (ax_i + b - y_i) \cdot x_i = 0 \quad \Leftrightarrow \quad a[x^2] + b[x^1] = [x^1 y]$$

$$\frac{\partial S(a, b)}{\partial b} = 0 \quad \Leftrightarrow \quad \sum_{i=1}^4 2 \cdot (ax_i + b - y_i) \cdot 1 = 0 \quad \Leftrightarrow \quad a[x^1] + b[x^0] = [x^0 y]$$

i	x_i	y_i	x_i^2	$x_i y_i$
1	0	1,05	0	0
2	1	2,99	1	2,99
3	2	4,52	4	9,04
$[x^0] = 4$	3	6,98	9	20,94
$\sum_{i=1}^4$	$6=[x^1]$	$15,54 = [x^0 y]$	$14 = [x^2]$	$32,97 = [x^1 y]$

Tabuľka 5.22

Do normálnych rovníc dosadíme hodnoty príslušných súčtov z tabuľky (Tabuľka 5.22). Na výpočet môžeme použiť aj zápis rovníc v maticovom tvare:

$$\begin{pmatrix} [x^2] & [x^1] \\ [x^1] & [x^0] \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} [x^1 y] \\ [x^0 y] \end{pmatrix} \Leftrightarrow \begin{pmatrix} 14 & 6 \\ 6 & 4 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 32,97 \\ 15,54 \end{pmatrix}$$

$$14a + 6b = 32,97 \quad / \cdot (2)$$

$$6a + 4b = 15,54 \quad / \cdot (-3)$$

$$28a + 12b = 65,94$$

$$-18a - 12b = -46,62$$

Po sčítaní rovníc dostávame:

$$10a = 19,32 \Rightarrow a = 1,932.$$

Z rovnice:

$$6a + 4b = 15,54 \Rightarrow b = \frac{15,54 - 6a}{4} = \frac{15,54 - 6 \cdot 1,932}{4} = 0,987$$

Záver. Hľadaný aproximačný polynóm: $P_1(x) = ax + b = 1,932x + 0,987$.

b) *Absolútna chyba aproximácie v uzle x_2 :*

$$\Delta y_2 = |y_2 - P_1(x_2)| = |2,99 - P_1(1)| = |2,99 - 2,919| = 0,071.$$

c) *Celková chyba aproximácie:* $\varepsilon = \sqrt{\frac{1}{4} \sum_{i=1}^4 (y_i - P_1(x_i))^2} = \sqrt{\frac{0,15738}{4}} = 0,198356.$

x_i	y_i	$P_1(x_i)$	$P_1(x_i) - y_i$	$(P_1(x_i) - y_i)^2$
0	1,05	0,987	-0,063	0,003969
1	2,99	2,919	-0,071	0,005041
2	4,52	4,851	0,331	0,109561
3	6,98	6,783	-0,197	0,038809
Σ				0,15738

Tabuľka 5.23. Tabuľka výpočtu celkovej chyby aproximácie

PRÍKLAD č. 2

Funkciu $f(x)$ danú tabuľkou aproximujte v zmysle Metódy najmenších štvorcov polynómom vhodne zvoleného stupňa.

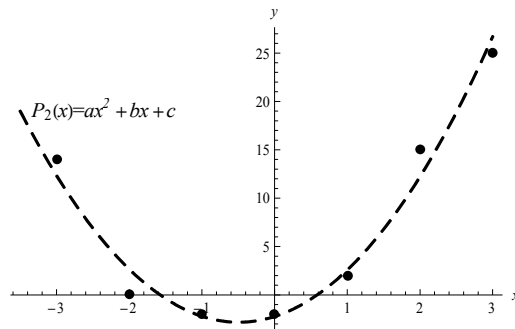
i	1	2	3	4	5	6	7
x_i	-3	-2	-1	0	1	2	3
$y_i = f(x_i)$	14	0	-2	-2	2	15	25

Tabuľka 5.24

RIEŠENIE

7 bodov grafu $f(x)$:

- $X_1 = [-3; 14]$
- $X_2 = [-2; 0]$
- $X_3 = [-1; -2]$
- $X_4 = [0; -2]$
- $X_5 = [1; 2]$
- $X_6 = [2; 15]$
- $X_7 = [3; 25]$



Obr. 5.19. Graf $f(x)$ a $P_2(x)$

Body grafu $f(x)$ ležia približne na parabole, preto $f(x)$ budeme aproximovať polynómom 2. stupňa: $P_2(x) = ax^2 + bx + c$.

Chceme vypočítať koeficienty a, b, c polynómu $P_2(x)$.

Minimalizujeme funkciu:

$$S(a, b, c) = \sum_{i=1}^7 (P_2(x_i) - y_i)^2 = \sum_{i=1}^7 (ax_i^2 + bx_i + c - y_i)^2$$

Riešime systém normálnych rovníc:

$$\frac{\partial S(a, b, c)}{\partial a} = 0 \quad \wedge \quad \frac{\partial S(a, b, c)}{\partial b} = 0 \quad \wedge \quad \frac{\partial S(a, b, c)}{\partial c} = 0$$

$$\frac{\partial S(a, b, c)}{\partial a} = 0 \quad \Leftrightarrow \quad \sum_{i=1}^7 2(ax_i^2 + bx_i + c - y_i) \cdot x_i^2 = 0 \quad \Leftrightarrow \quad a[x^4] + b[x^3] + c[x^2] = [x^2 y]$$

$$\frac{\partial S(a, b, c)}{\partial b} = 0 \quad \Leftrightarrow \quad \sum_{i=1}^7 2(ax_i^2 + bx_i + c - y_i) \cdot x_i = 0 \quad \Leftrightarrow \quad a[x^3] + b[x^2] + c[x^1] = [x^1 y]$$

$$\frac{\partial S(a, b, c)}{\partial c} = 0 \quad \Leftrightarrow \quad \sum_{i=1}^7 2(ax_i^2 + bx_i + c - y_i) \cdot 1 = 0 \quad \Leftrightarrow \quad a[x^2] + b[x^1] + c[x^0] = [x^0 y]$$

i	x_i	y_i	x_i^2	x_i^3	x_i^4	$x_i y_i$	$x_i^2 y_i$
1	-3	14	9	-27	81	-42	126
2	-2	0	4	-8	16	0	0
3	-1	-2	1	-1	1	2	-2
4	0	-2	0	0	0	0	0
5	1	2	1	1	1	2	2
6	2	15	4	8	16	30	60
$[x^0] = 7$	3	25	9	27	81	75	225
$\sum_{i=1}^7$	$0=[x^1]$	$52=[x^0 y]$	$28=[x^2]$	$0=[x^3]$	$196=[x^4]$	$67=[x^1 y]$	$411=[x^2 y]$

Tabuľka 5.25

Na výpočet môžeme opäť využiť zápis normálnych rovníc v maticovom tvare. Z tabuľky (Tabuľka 5.26) dosadíme hodnoty príslušných súčtov:

$$\begin{pmatrix} [x^4] & [x^3] & [x^2] \\ [x^3] & [x^2] & [x^1] \\ [x^2] & [x^1] & [x^0] \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} [x^2 y] \\ [x^1 y] \\ [x^0 y] \end{pmatrix} \Leftrightarrow \begin{pmatrix} 196 & 0 & 28 \\ 0 & 28 & 0 \\ 28 & 0 & 7 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 411 \\ 67 \\ 52 \end{pmatrix}$$

$$196a + 28c = 411$$

$$28b = 67$$

$$28a + 7c = 52$$

Riešenie systému je $a=2,41667$, $b=2,39286$, $c=-2,2381$.

Záver. Hľadaný aproximačný polynóm je

$$P_2(x) = ax^2 + bx + c = 2,41667x^2 + 2,39286x - 2,2381.$$

PRÍKLAD č. 3

Funkciu $f(x)$ danú tabuľkou Tabuľka 5.26 aproximujte v zmysle Metódy najmenších štvorcov exponenciálnou funkciou $y = Ce^{Ax}$.

i	1	2	3	4
x_i	-0,5	0	0,5	1
$y_i = f(x_i)$	0,05	0,14	0,37	1

Tabuľka 5.26

RIEŠENIE

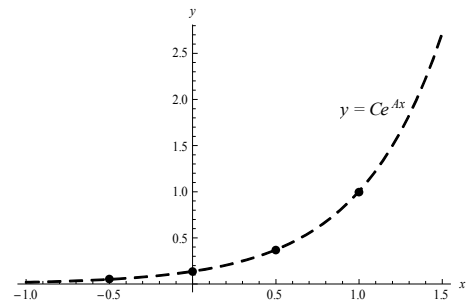
Funkcia $f(x)$ je daná 4 bodmi:

$$X_1 = [-0,05; 0,05]$$

$$X_2 = [0; 0,14]$$

$$X_3 = [0,5; 0,37]$$

$$X_4 = [1; 1]$$



Body grafu $f(x)$ ležia približne na grafe exponenciálnej funkcie $y = Ce^{Ax}$.

Obr. 5.20. Graf $f(x)$ a $y = Ce^{Ax}$

Exponenciálnu funkciu $y = Ce^{Ax}$ budeme linearizovať, t.j. transformovať na lineárnu funkciu – polynóm 1. stupňa: $P_1(x) = Ax + B$ (Tabuľka 5.27).

	Aproximačná $g(x)$	Množina dát $f(x)$	Transformačné vzťahy
Pred linearizáciou	$y = Ce^{Ax}$	$[x_i; y_i = f(x_i)]$	
Linearizácia	$y = Ce^{Ax} / \ln_e$ $\ln(y) = \ln(Ce^{Ax})$ $\ln(y) = \ln(C) + Ax \ln(e)$ $Y = B + Ax$	$Y_i = \ln y_i$	$x \rightarrow x$ $y \rightarrow Y \quad \wedge \quad Y = \ln y$ $A \rightarrow A$ $C \rightarrow e^B \quad \wedge \quad B = \ln(C)$
Po linearizácii	$Y = B + Ax$ $(P_1(x) = B + Ax)$	$[X_i; Y_i]$	

Tabuľka 5.27

Zároveň linearizujeme aj funkciu $f(x)$ (pozri stĺpec $Y_i = \ln(y_i)$ v tabuľke Tabuľka 5.28 a vypočítame koeficienty A a B polynómu $P_1(x)$.

Minimalizujeme chybu:

$$S(A, B) = \sum_{i=1}^4 (P_1(x_i) - Y_i)^2 = \sum_{i=1}^4 (Ax_i + B - Y_i)^2$$

$$\frac{\partial S(A, B)}{\partial A} = 0 \quad \wedge \quad \frac{\partial S(A, B)}{\partial B} = 0$$

a riešime systém normálnych rovníc:

$$\frac{\partial S(A, B)}{\partial A} = 0 \quad \Leftrightarrow \quad \sum_{i=1}^4 2(Ax_i + B - Y_i) \cdot x_i = 0 \quad \Leftrightarrow \quad A[x^2] + B[x^1] = [x^1 Y]$$

$$\frac{\partial S(A, B)}{\partial B} = 0 \quad \Leftrightarrow \quad \sum_{i=1}^4 2(Ax_i + B - Y_i) \cdot 1 = 0 \quad \Leftrightarrow \quad A[x^1] + B[x^0] = [x^0 Y]$$

i	x_i	y_i	$Y_i = \ln(y_i)$	x_i^2	$x_i Y_i$
1	-0,5	0,05	-2,9957	0,25	1,4979
2	0	0,14	-1,9661	0	0
3	0,5	0,37	-0,9943	0,25	-0,4972
$[x^0] = 4$	1	1	0	1	0
$\sum_{i=1}^4$	$1 = [x^1]$		$-5,9561 = [x^0 Y]$	$1,5 = [x^2]$	$1,0007 = [x^1 Y]$

Tabuľka 5.28

Hodnoty príslušných súčtov dosadíme z tabuľky (Tabuľka 5.28). V maticovom zápise: dostaneme:

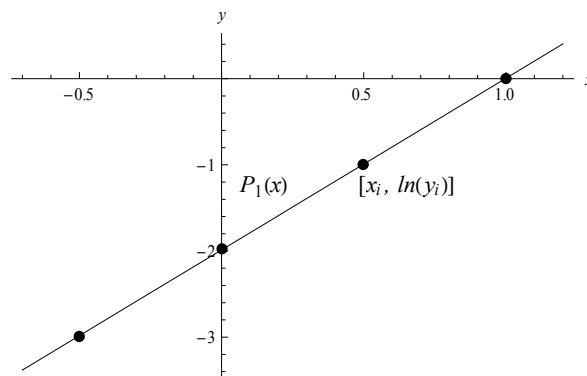
$$\begin{pmatrix} [x^2] & [x^1] \\ [x^1] & [x^0] \end{pmatrix} \cdot \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} [x^1 Y] \\ [x^0 Y] \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1,5 & 1 \\ 1 & 4 \end{pmatrix} \cdot \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 1,0007 \\ -5,9561 \end{pmatrix}$$

$$1,5A + B = 1,0007$$

$$A + 4B = -5,956$$

Riešenie systému je $A = 1,992$, $B = -1,9866$.

Hľadaný aproximačný polynóm $P_1(x) = Ax + B = 1,992x - 1,9866$.

Obr. 5.21. Graf linearizovanej $f(x)$ a $P_1(x)$

Teraz ešte musíme lineárnu funkciu $Y = 1,9866 + 1,992x$ „odlinearizovať“, t.j. spätne transformovať z lineárnej funkcie na funkciu exponenciálnu $y = Ce^{Ax}$.

Dostávame:

$$Y = 1,992x - 1,9866$$

$$\ln(y) = 1,992x - 1,9866 \Rightarrow y = e^{1,992x - 1,9866}$$

$$y = e^{1,992x} \cdot e^{-1,9866}$$

$$y = 0,1371e^{1,992x}$$

Záver. Hľadaná aproximačná exponenciálna funkcia $y = Ce^{Ax}$ je $y = 0,1371e^{1,992x}$.

6 NUMERICKÉ DERIVOVANIE

Medzi meranými veličinami x a y často hľadáme okrem predpisu funkčnej závislosti $y = f(x)$ aj hodnoty jej derivácií v žiadaných bodoch x .

Pre spojité funkcie na výpočet derivácie v danom bode x_0 používame exaktný vzorec matematickej analýzy:

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}, \quad (6.1)$$

kde funkcia $f(x)$ je spojitá v okolí bodu x_0 a limita skúmaného podielu v bode x_0 sa rovná limite sprava aj zľava:

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \rightarrow x_0^+} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \rightarrow x_0^-} \frac{f(x) - f(x_0)}{x - x_0}.$$

Ak limita v danom bode existuje iba jednostranná, t. j. iba zľava alebo iba sprava, derivácia v danom bode je tiež iba jednostranná:

$$f'_+(x) = \lim_{x \rightarrow x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \qquad f'_-(x) = \lim_{x \rightarrow x_0^-} \frac{f(x) - f(x_0)}{x - x_0}$$

Pri hľadaní derivácie funkcie danej izolovanými bodmi vychádzame z jej nahradenia interpolačným polynómom – Newtonovým alebo Lagrangeovým (pozri kap.5), pričom polynóm nederivujeme pomocou definície derivácie spojitaj funkcie (6.1), ale na jej určenie využívame pomerné diferencie.

ÚLOHA

Funkcia $y = f(x)$ je daná tabuľkou (Tabuľka 6.1), pričom uzly

- sú usporiadané vzostupne: $x_1 < x_2 < \dots < x_{n+1}$,
- majú ekvidistantné delenie (vzdialenosť medzi uzlami je rovnaká):

$$x_{i+1} = x_i + h, \text{ čiže } |x_{i+1} - x_i| = h = \text{konšt.}, \quad i = 1, 2, \dots, n.$$

Vypočítajte $f'(x_i)$ a $f''(x_i)$.

x_i	x_1	x_2	\dots	x_{n+1}
$y_i = f(x_i)$	y_1	y_2	\dots	y_{n+1}

Tabuľka 6.1

RIEŠENIE:

Funkciu $y = f(x)$ nahradíme interpolačným polynómom $P_n(x)$ a vieme, že pre ľubovoľné $x \in \langle x_1, x_{n+1} \rangle$ platí: $f(x) \approx P_n(x)$, $f'(x) \approx P_n'(x)$, resp. $f''(x) \approx P_n''(x)$. Platí to aj pre derivácie vyšších rádov.

Výpočet $f'(x_i)$ a $f''(x_i)$ ilustrujme na Newtonovom interpolačnom polynóme 2. stupňa, ktorého graf prechádza bodmi $X_1[x_1, f(x_1)]$, $X_2[x_2, f(x_2)]$, $X_3[x_3, f(x_3)]$. Označme $y_i = f(x_i) = f_i$.

Zostavíme tabuľku diferencií (Tabuľka 6.2).

x_i	f_i	$\Delta^1 f_i$	$\Delta^2 f_i$
x_1	f_1	$\Delta^1 f_1 = \frac{f_2 - f_1}{x_2 - x_1} = \frac{f_2 - f_1}{h}$	$\Delta^2 f_i = \frac{\Delta^1 f_2 - \Delta^1 f_1}{x_3 - x_1} = \frac{\frac{f_3 - f_2}{x_3 - x_2} - \frac{f_2 - f_1}{x_2 - x_1}}{2h} = \frac{f_3 - 2f_2 + f_1}{2h^2}$
x_2	f_2		
x_3	f_3		

Tabuľka 6.2. Tabuľka diferencií – Fraserov diagram

Potom

$$N_2(x) = f_1 + \frac{f_2 - f_1}{h} \cdot (x - x_1) + \frac{f_3 - 2f_2 + f_1}{2h^2} \cdot (x - x_1) \cdot (x - x_2)$$

$$N_2'(x) = \frac{f_2 - f_1}{h} + \frac{f_3 - 2f_2 + f_1}{2h^2} \cdot (2x - x_1 - x_2)$$

$$N_2''(x) = \frac{f_3 - 2f_2 + f_1}{2h^2} \cdot 2 = \frac{f_3 - 2f_2 + f_1}{h^2} = \text{konšt.}$$

Vypočítajme 1. deriváciu funkcie v jednotlivých uzloch x_1, x_2, x_3 :

$$N_2'(x_1) = \frac{f_2 - f_1}{h} + \frac{f_3 - 2f_2 + f_1}{2h^2} (x_1 - x_2) = \frac{f_2 - f_1}{h} + \frac{f_3 - 2f_2 + f_1}{2h^2} (-h) = \frac{-3f_1 + 4f_2 - f_3}{2h} \approx f'(x_1)$$

$$N_2'(x_2) = \frac{f_2 - f_1}{h} + \frac{f_3 - 2f_2 + f_1}{2h^2} (x_2 - x_1) = \frac{f_2 - f_1}{h} + \frac{f_3 - 2f_2 + f_1}{2h^2} (h) = \frac{f_3 - f_1}{2h} \approx f'(x_2)$$

$$\begin{aligned} N_2'(x_3) &= \frac{f_2 - f_1}{h} + \frac{f_3 - 2f_2 + f_1}{2h^2} ((x_3 - x_1) + (x_3 - x_2)) = \frac{f_2 - f_1}{h} + \frac{f_3 - 2f_2 + f_1}{2h^2} (2h) = \\ &= \frac{3f_3 - 4f_2 + f_1}{2h} \approx f'(x_3) \end{aligned}$$

Vidíme, že:

- vo výpočte derivácií sa používajú pomerné diferencié,
- výpočet $N_2'(x_2) \approx f'(x_2)$ používa na výpočet uzly vľavo i vpravo od uzla x_2 (obe hodnoty f_1 a f_3), a preto bude najpresnejší,

- výpočet $N_2'(x_1) \approx f'(x_1)$ používa na výpočet aktuálny uzol x_1 a nasledujúce uzly x_2 a x_3 ,
- výpočet $N_2'(x_3) \approx f'(x_3)$ používa na výpočet aktuálny uzol x_3 a predchodcov x_2 a x_1 ,
- druhá derivácia polynómu druhého stupňa je konštanta, ktorej hodnota závisí od všetkých troch hodnôt $y_i = f(x_i) = f_i$:

$$N_2''(x_i) = \frac{f_3 - 2f_2 + f_1}{h^2} \approx f''(x_i), i = 1, 2, 3.$$

Podobne by sme mohli pokračovať vo výpočte derivácií pre interpolačné polynómy vyšších stupňov určených izolovanými bodmi.

Je zrejmé, že hodnota prvej aj druhej derivácie bude závisieť:

- v 1. uzle od hodnôt funkcie v nasledujúcich uzloch,
- v stredných uzloch od hodnôt funkcie v predchádzajúcich aj v nasledujúcich uzloch,
- v poslednom uzle od hodnôt funkcie v predchádzajúcich uzloch.

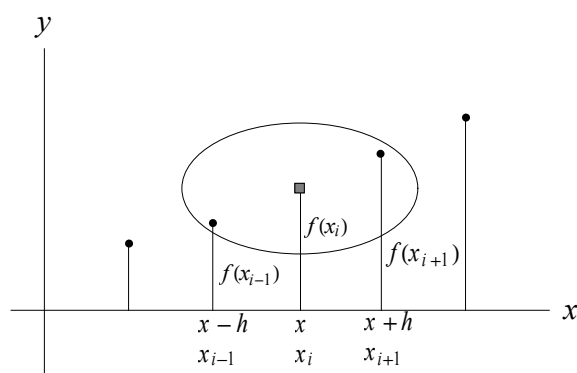
Podľa použitých hodnôt funkcie vzorce rozdelíme na:

- formuly vpred,
- formuly centrálné (symetrické),
- formuly vzad.

V ďalšom uvádzame prehľad najpoužívanejších vzorcov na výpočet 1. a 2. derivácie funkcie $f(x)$ danej tabuľkou (Tabuľka 6.1):

Centrálna diferenčná formula (CDF) rádu $O(h^2)$ – symetrická	
1. derivácia	2. derivácia
$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h}$ <p style="text-align: center;">resp.</p> $f'(x_i) \approx \frac{f(x_{i+1}) - f(x_{i-1}))}{2h}$	$f''(x) \approx \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}$ <p style="text-align: center;">resp.</p> $f''(x_i) \approx \frac{f(x_{i+1}) - 2f(x_i) + f(x_{i-1}))}{h^2}$

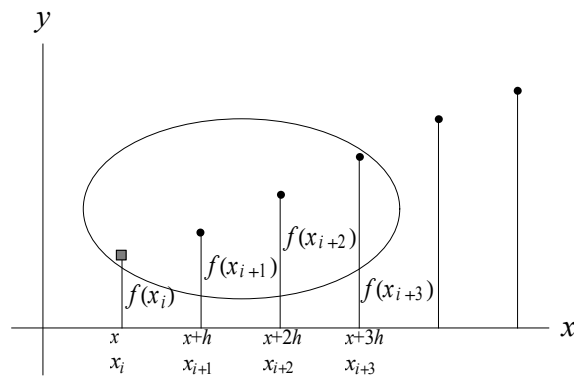
Tabuľka 6.3



Obr. 6.1

Nesymetrická formula vpred (NFvp) rádu $O(h^2)$	
1. derivácia	2. derivácia
$f'(x) \approx \frac{-3f(x) + 4f(x+h) - f(x+2h)}{2h}$	$f''(x) \approx \frac{2f(x) - 5f(x+h) + 4f(x+2h) - f(x+3h)}{h^2}$
resp.	resp.
$f'(x_i) \approx \frac{-3f(x_i) + 4f(x_{i+1}) - f(x_{i+2})}{2h}$	$f''(x_i) \approx \frac{2f(x_i) - 5f(x_{i+1}) + 4f(x_{i+2}) - f(x_{i+3})}{h^2}$

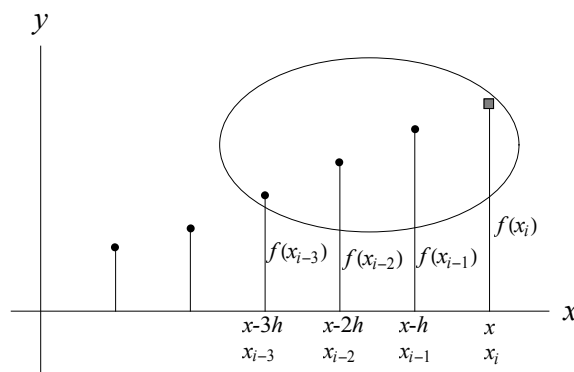
Tabuľka 6.4



Obr. 6.2

Nesymetrická formula vzad (NFvz) rádu $O(h^2)$	
1. derivácia	2. derivácia
$f'(x) \approx \frac{3f(x) - 4f(x-h) + f(x-2h)}{2h}$	$f''(x) \approx \frac{2f(x) - 5f(x-h) + 4f(x-2h) - f(x-3h)}{h^2}$
resp.	resp.
$f'(x_i) \approx \frac{-3f(x_i) + 4f(x_{i-1}) - f(x_{i-2})}{2h}$	$f''(x_i) \approx \frac{2f(x_i) - 5f(x_{i-1}) + 4f(x_{i-2}) - f(x_{i-3})}{h^2}$

Tabuľka 6.5



Obr. 6.3

Na základe uvedených vzťahov môžeme konštatovať, že pri výpočte derivácie v ľubovoľnom uzle x_i a použití formúl:

- CDF – musíme poznať hodnotu funkcie v predchádzajúcom uzle x_{i-1} (umiestnený vľavo od uzla x_i) a hodnotu funkcie v nasledujúcom uzle x_{i+1} (umiestnený vpravo od uzla x_i). Znamená to, že táto formula neumožňuje výpočet 1. a 2. derivácie funkcie $f(x)$ v prvom x_1 a ani poslednom uzle x_{n+1} .
- NFvp – musíme poznať minimálne 2, resp. 3 nasledujúce uzly $x_{i+1}, x_{i+2}, x_{i+3}$. Znamená to, že táto formula umožňuje výpočet 1. a 2. derivácie funkcie $f(x)$ už v prvom uzle x_1 , ale len po uzol x_{n-1} (pre 1. deriváciu), resp. x_{n-2} (pre 2. deriváciu).
- NFvz – si vyžaduje poznať minimálne 2, resp. 3 uzly $x_{i-1}, x_{i-2}, x_{i-3}$ vľavo od neho. Znamená to, že táto formula umožňuje výpočet 1. a 2. derivácie funkcie $f(x)$ aj v poslednom uzle x_{n+1} , ale len od uzla x_4 .

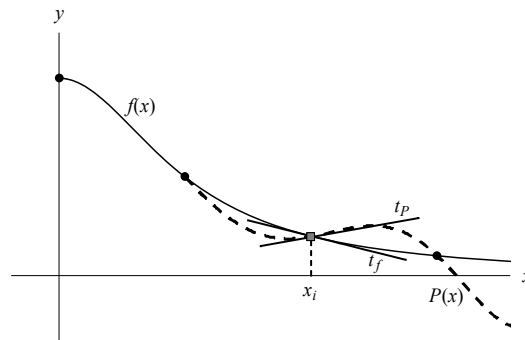
Výhodou použitia vzorcov je, že nemusíme určovať predpis pre príslušný interpolačný polynóm, ale derivácie počítame priamo z hodnôt funkcie $f(x_i)$. Zdôrazňujeme, že delenie intervalu $\langle x_1, x_{n+1} \rangle$ musí byť **ekvidištantné**, t. j. krok $h = |x_{i+1} - x_i|$ je **rovnaký** pre všetky uzly x_i .

Ak by sme chceli použiť interpolačný polynóm vyššieho stupňa (s väčším počtom uzlov), po jeho zderivovaní by sme získali ďalšie formuly na výpočet približných hodnôt derivácie v jednotlivých uzloch. Vo všeobecnosti presnosť výpočtu približných hodnôt derivácie $f(x)$ podľa formúl narastá zvyšovaním stupňa interpolačného polynómu.

Na prvý pohľad by sa zdalo, že pri výpočte derivácií z interpolačného polynómu nevznikajú žiadne ťažkosti. Žiaľ, nie je to tak. Napríklad:

- veľký vplyv má zaokrúhľovacia chyba, ktorá pri samotnej interpolácii nemá význam, a však pri numerickom výpočte derivácie je nielen významná, ale často až zničujúca. Dá sa totiž ukázať, že malé vzdialenosti (krok h) medzi uzlami spôsobujú veľké zväčšenie zaokrúhľovacej chyby na funkčnej hodnote derivácie (zdôvodnite, kap.1) a väčšie vzdialenosti (krok h) medzi uzlami spôsobujú zase veľkú chybu metódy – zaokrúhľovacia chyba je nepriamo úmerná kroku h a chyba metódy je úmerná kroku h . Dá sa vypočítať (Ralston – Základy NM) optimálny krok h tak, aby chyba metódy a zaokrúhľovacia chyba boli čo najmenšie:
- pre rôzne veľkosti kroku h dostávame rozdielne hodnoty derivácie $f'(x)$. Bod x má iných predchodcov ($x - h$) a nasledovníkov ($x + h$),
- dá sa očakávať, že aj keď sa interpolačná funkcia $P(x)$ len málo odlišuje od zadanej funkcie $f(x)$, ich derivácie (smernice dotyčníc) sa budú líšiť oveľa viac (Obr. 6.4).

Poznámka. Ak hodnoty funkcie sú už primárne zaťažené chybou merania, resp. je ich príliš veľký počet, potom na výpočet derivácie je lepšie použiť aproximačný polynóm získaný metódou najmenších štvorcov.



Obr. 6.4. Dotyčnice v bode x_i majú rôzne smernice

6.1 RIEŠENÉ PRÍKLADY

PRÍKLAD č. 1

Daná je funkcia $f(x)=0,5 \cdot e^{2x}$ na intervale $\langle 2;2,5 \rangle$.

- Určte derivácie $f'(x)$ a $f''(x)$ v uzloch $x = 2$, $x = 2,1$ a $x = 2,5$
 - analyticky (presne),
 - numericky. Na výpočet použite vhodné formuly s krokom $h = 0,1$ a $h = 0,05$.
- Vypočítané hodnoty derivácií porovnajte.

RIEŠENIE

Použijeme označenie:

- $f'(x)$ a $f''(x)$ pre derivácie danej funkcie počítané analyticky (presne),
- $\tilde{f}'(x)$ a $\tilde{f}''(x)$ pre derivácie danej funkcie počítané numericky.

Hodnoty funkcie a jej derivácií vyjadríme na 6 platných číslic.

- Výpočet derivácií $f'(x)$ a $f''(x)$ v uzloch $x = 2$, $x = 2,1$ a $x = 2,5$

a) analyticky

$$f(x)=0,5 \cdot e^{2x} \qquad f'(x)=0,5 \cdot 2 \cdot e^{2x} = e^{2x} \qquad f''(x)=(e^{2x})' = 2e^{2x}$$

Funkcie $f(x)$, $f'(x)$, $f''(x)$ sú spojité v každom bode $x \in \langle 2;2,5 \rangle$. Vypočítame príslušné derivácie v daných bodoch. Dostávame

$f'(x)=e^{2x}$	$f''(x)=2e^{2x}$
$f'(2)=54,5982$	$f''(2)=109,196$
$f'(2,1)=66,6863$	$f''(2,1)=133,373$
$f'(2,5)=148,413$	$f''(2,5)=296,826$

b) numericky

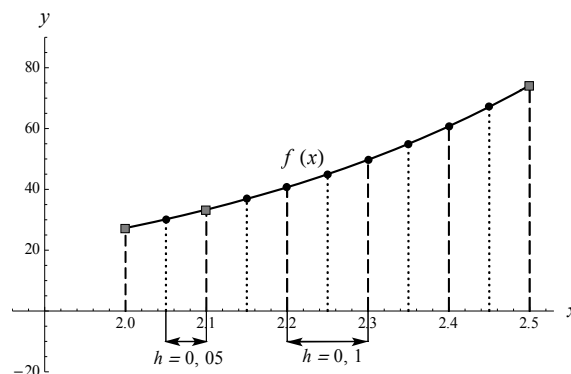
Výpočet približných hodnôt $\tilde{f}'(x)$ a $\tilde{f}''(x)$ s krokom $h = 0,1$

Rozdelíme interval $\langle 2; 2,5 \rangle$ s krokom $h = 0,1$ na $\frac{2,5-2}{0,1} = 5$ čiastkových intervalov deliacimi

bodmi (uzlami) $x_{i+1} = x_i + h$ (alebo $x_{i+1} = x_1 + i \cdot h = 2 + i \cdot h$) pre $i = 1, 2, \dots, 5$. Vypočítajme hodnoty funkcie $f(x)$ v jednotlivých uzloch a zapíšme ich do tabuľky (Tabuľka 6.6).

i	1	2	3	4	5	6
x_i	2	2,1	2,2	2,3	2,4	2,5
$y_i = 0,5e^{2x_i}$	27,2991	33,3432	40,7254	49,7422	60,7552	74,2066

Tabuľka 6.6



Obr. 6.5. Graf funkcie $f(x)$ s vyznačenými bodmi pre krok $h = 0,1$ a $h = 0,05$

Pretože $x_1 = 2$ leží na začiatku $\langle 2; 2,5 \rangle$, príslušné derivácie vypočítame nesymetrickou formulou vpred

$$\tilde{f}'(x_1) = \frac{-3f(x_1) + 4f(x_2) - f(x_3)}{2h} = \frac{-3 \cdot 27,2991 + 4 \cdot 33,3432 - 40,7254}{2 \cdot 0,1} = 53,7505$$

$$\tilde{f}''(x_1) = \frac{2f(x_1) - 5f(x_2) + 4f(x_3) - f(x_4)}{h^2} = \frac{2 \cdot 27,2991 - 5 \cdot 33,3432 + 4 \cdot 40,7254 - 49,7422}{0,1^2} = 104,160.$$

Na výpočet derivácií v $x_2 = 2,1$ použijeme napr. centrálnu diferenčnú formulu, ale môžeme použiť opäť nesymetrickú formulu vpred:

$$\tilde{f}'(x_2) = \frac{f(x_3) - f(x_1)}{2h} = \frac{40,7254 - 27,2991}{2 \cdot 0,1} = 67,1315$$

$$\tilde{f}''(x_2) = \frac{f(x_3) - 2f(x_2) + f(x_1)}{h^2} = \frac{40,7254 - 2 \cdot 33,3432 + 27,2991}{0,1^2} = 133,810$$

Keďže $x_6 = 2,5$ leží na konci $\langle 2; 2,5 \rangle$, príslušné derivácie vypočítame nesymetrickou formulou vzad:

$$\tilde{f}'(x_6) = \frac{3f(x_6) - 4f(x_5) + f(x_4)}{2h} = \frac{3 \cdot 74,2066 - 4 \cdot 60,7552 + 49,7422}{2 \cdot 0,1} = 146,706$$

$$\tilde{f}''(x_6) = \frac{2f(x_6) - 5f(x_5) + 4f(x_4) - f(x_3)}{h^2} = \frac{2 \cdot 74,2066 - 5 \cdot 60,7552 + 4 \cdot 49,7422 - 40,7254}{0,1^2} = 288,060.$$

Výpočet hodnôt $f'(x)$ a $f''(x)$ s krokom $h = 0,05$

Interval $\langle 2; 2,5 \rangle$ opäť rozdelíme, ale s krokom $h = 0,05$, na $\frac{2,5 - 2}{0,05} = 10$ čiastkových intervalov deliacimi bodmi $x_{i+1} = x_i + h$ pre $i = 1, 2, \dots, 10$. Vypočítajme hodnoty funkcie $f(x)$ v týchto uzloch a zapíšme ich do tabuľky (Tabuľka 6.7).

i	1	2	3	4	5	6
x_i	2	2,05	2,1	2,15	2,2	2,25
$y_i = 0,5 e^{2x_i}$	27,2991	30,1701	33,3432	36,8499	40,7254	45,0086
i	7	8	9	10	11	/
x_i	2,3	2,35	2,4	2,45	2,5	
$y_i = 0,5 e^{2x_i}$	49,7422	54,9736	60,7552	67,1449	74,2066	

Tabuľka 6.7

$\tilde{f}'(x)$ a $\tilde{f}''(x)$ v príslušných uzloch vypočítame tými istými formulami ako pri kroku $h = 0,1$.

$x_1 = 2$:

$$\tilde{f}'(x_1) = \frac{-3f(x_1) + 4f(x_2) - f(x_3)}{2h} = \frac{-3 \cdot 27,2991 + 4 \cdot 30,1701 - 33,3432}{2 \cdot 0,05} = 54,399$$

$$\tilde{f}''(x_1) = \frac{2f(x_1) - 5f(x_2) + 4f(x_3) - f(x_4)}{h^2} = \frac{2 \cdot 27,2991 - 5 \cdot 30,1701 + 4 \cdot 33,3432 - 36,8499}{0,05^2} = 108,24$$

$x_3 = 2,1$:

$$\tilde{f}'(x_3) = \frac{f(x_4) - f(x_2)}{2h} = \frac{36,8499 - 30,1701}{2 \cdot 0,05} = 66,798$$

$$\tilde{f}''(x_3) = \frac{f(x_4) - 2f(x_3) + f(x_2)}{h^2} = \frac{36,8499 - 2 \cdot 33,3432 + 30,1701}{0,05^2} = 133,44$$

$x_{11} = 2,5$:

$$\tilde{f}'(x_{11}) = \frac{+3f(x_{11}) - 4f(x_{10}) + f(x_9)}{2h} = \frac{3 \cdot 74,2066 - 4 \cdot 67,1449 + 60,7552}{2 \cdot 0,05} = 147,954$$

$$\tilde{f}''(x_{11}) = \frac{2f(x_{11}) - 5f(x_{10}) + 4f(x_9) - f(x_8)}{h^2} = \frac{2 \cdot 74,2066 - 5 \cdot 67,1449 + 4 \cdot 60,7552 - 54,9736}{0,05^2} = 294,36$$

2. Porovnanie

Porovnanie presných a približných hodnôt derivácií vykonáme na základe absolútnych chýb:

$$\varepsilon(\tilde{f}'(x)) = |f'(x) - \tilde{f}'(x)|, \quad \varepsilon(\tilde{f}''(x)) = |f''(x) - \tilde{f}''(x)|$$

uzol	$h = 0,1$		$h = 0,05$	
	$\varepsilon(\tilde{f}'(x))$	$\varepsilon(\tilde{f}''(x))$	$\varepsilon(\tilde{f}'(x))$	$\varepsilon(\tilde{f}''(x))$
$x = 2$	0,8477	5,036	0,1992	0,956
$x = 2,1$	0,4452	0,437	0,1117	0,067
$x = 2,5$	1,707	8,766	0,459	2,466

Tabuľka 6.8

Z porovnania presných a približných hodnôt 1. a 2. derivácie funkcie $f(x)$ vidíme, že ak krok $h = 0,1$ zmenšíme na polovicu ($h = 0,05$), absolútne chyby $\varepsilon(\tilde{f}'(x))$ a $\varepsilon(\tilde{f}''(x))$ tiež poklesnú, t. j. zmenšením kroku h vzrastie presnosť výpočtu približných hodnôt derivácií $\tilde{f}'(x)$ a $\tilde{f}''(x)$. Výsledky derivácií v krajných uzloch sú menej presné ako v strednom uzle $x = 2,1$, pre ktorý boli použité centrálné formuly.

7 NUMERICKÉ INTEGROVANIE

Súčasťou riešenia mnohých technických problémov je výpočet určitého integrálu. Bežným matematickým spôsobom riešenia je nájdenie primitívnej funkcie a ďalšie uplatnenie Newton-Leibnitzovho vzorca:

$$\int_a^b f(x) dx = [F(x)]_a^b = F(b) - F(a),$$

kde $f(x)$ je spojitá na $\langle a, b \rangle$ a $F(x)$ je primitívna funkcia k funkcii $f(x)$.

Geometrická interpretácia určitého integrálu

Ak $f(x) \geq 0$ na $\langle a, b \rangle$, potom

$$\int_a^b f(x) dx = P(L),$$

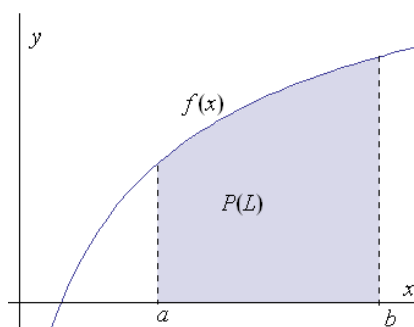
kde $P(L)$ je plocha krivočiareho lichobežníka ohraničeného grafom $f(x)$, osou x a priamkami $x = a$, $x = b$.

Hoci vieme, že každá spojitá funkcia na intervale $\langle a, b \rangle$ je integrovateľná a teda k nej existuje primitívna funkcia, nepoznáme žiadny univerzálny návod, ako primitívnu funkciu $F(x)$ nájsť. Jedným z dôvodov je fakt, že na

rozdiel od derivácie, primitívnu funkciu k elementárnej funkcii nemusí byť opäť elementárna funkcia, t. j. funkcia, ktorá sa dá vyjadriť vzorcom. Typickým príkladom integrálu, ku

ktorému nevieme určiť primitívnu funkciu, je Gaussov integrál $\int_a^b e^{-x^2} dx$.

Výpočet určitého integrálu patrí k štandardným úlohám numerickej matematiky. Na jeho riešenie existuje viacero numerických metód, my sa budeme venovať iba ich úzkemu výberu.



Obr. 7.1

7.1 NUMERICKÉ METÓDY NA VÝPOČET URČITÉHO INTEGRÁLU

Numerické metódy na výpočet určitého integrálu používame vtedy, keď:

- primitívnu funkciu $F(x)$ nevieme nájsť,
- primitívnu funkciu $F(x)$ poznáme, ale použitie Newton – Leibnitzovho vzorca je veľmi náročné,
- funkcia $f(x)$ je daná tabuľkou.

Základná myšlienka

1. Funkciu $f(x)$ nahradíme na $\langle a, b \rangle$ nejakou inou (jednoduchšou) funkciou $\varphi(x)$ (interpoláčny, aproximačný polynóm), ktorú potom analyticky integrujeme. Ak $\varphi(x)$ je dobrou aproximáciou funkcie $f(x)$, potom

$\int_a^b \varphi(x) dx$ je dobrou aproximáciou $\int_a^b f(x) dx$, t. j.

$$\int_a^b f(x) dx \approx \int_a^b \varphi(x) dx.$$

2. Vychádzame z definície určitého integrálu prostredníctvom limity integrálnych súčtov:

- a) interval $\langle a, b \rangle$ rozdelíme na n čiastkových intervalov (podintervalov) **rovnakej dĺžky**

h . Deliacimi bodmi je množina usporiadaných bodov $\{x_i\}_{i=1}^{n+1}$ s ekvidistantným delením, kde:

- i) $x_1 < x_2 < x_3 < \dots < x_n < x_{n+1}$, pričom $a = x_1$ a $b = x_{n+1}$,

$$\langle a, b \rangle = \langle x_1, x_2 \rangle \cup \langle x_2, x_3 \rangle \cup \dots \cup \langle x_n, x_{n+1} \rangle,$$

- ii) dĺžka každého čiastkového intervalu $\langle x_i, x_{i+1} \rangle$ je $h = \frac{b-a}{n}$,

$$|x_2 - x_1| = |x_3 - x_2| = \dots = |x_{n+1} - x_n| = h \text{ (konšt.)},$$

- b) na každom čiastkovom intervale nahradíme funkciu $f(x)$ vhodnou aproximačnou funkciou $\varphi(x)$:

$$\int_{x_i}^{x_{i+1}} f(x) dx \approx \int_{x_i}^{x_{i+1}} \varphi(x) dx,$$

- c) výpočet integrálu vyjadríme pomocou základného vzorca univerzálne pre každý podinterval $\langle x_i, x_{i+1} \rangle$ zvlášť (pozri ďalej jednoduchá metóda),

- d) sčítaním vzorcov jednoduchej metódy cez všetky podintervaly získame výsledný zložený vzorec (pozri ďalej zložená metóda), t. j. z vlastností určitého integrálu:

$$\int_a^b f(x) dx = \int_{a=x_1}^{b=x_{n+1}} f(x) dx = \int_{x_1}^{x_2} f(x) dx + \int_{x_2}^{x_3} f(x) dx + \dots + \int_{x_n}^{x_{n+1}} f(x) dx,$$

$$\text{Ak } f(x) \approx \varphi(x), \text{ potom } \int_a^b f(x) dx \approx \int_{x_1}^{x_2} \varphi(x) dx + \int_{x_2}^{x_3} \varphi(x) dx + \dots + \int_{x_n}^{x_{n+1}} \varphi(x) dx.$$

Geometrická interpretácia. Plochu rovinného útvaru, ktorú predstavuje $\int_a^b f(x) dx$, nahradíme súčtom plošných obsahov jednoduchších rovinných útvarov.

7.2 NEWTON – COTESOVE KVADRATÚRNE VZORCE STUPŇA n

Už z predchádzajúcich kapitol vieme, že vhodné aproximačné funkcie sú polynómy, ktoré sa ľahko derivujú aj integrujú.

Ak $f(x)$ nahradíme Lagrangeovým interpolačným polynóm n -tého stupňa (pozri kap.5),

$$f(x) \approx P_n(x),$$

potom dostávame tzv. Newton – Cotesove kvadrátúrne vzorce stupňa n :

- základný vzorec (jednoduchá metóda) pre interval $\langle x_i, x_{i+1} \rangle$,
- zložený vzorec (zložená metóda) pre interval

$$\langle a, b \rangle = \langle x_1, x_2 \rangle \cup \langle x_2, x_3 \rangle \cup \dots \cup \langle x_n, x_{n+1} \rangle = \bigcup_{i=1}^n \langle x_i, x_{i+1} \rangle.$$

7.2.1 OBDĽŽNIKOVÁ METÓDA

V tejto metóde funkciu $f(x)$ nahradíme Lagrangeovým interpolačným polynómom stupňa $n = 0$, čiže konštantnou funkciou, ktorej grafom je priamka rovnobežná s osou x , t. j.

$$f(x) \approx P_0(x), \quad P_0(x) = k, \quad k \in \mathbb{R}.$$

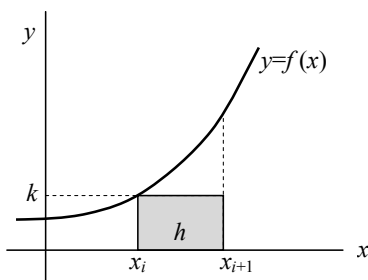
7.2.1.1 Jednoduchá obdĺžniková metóda

Funkciu $f(x)$ nahradíme na intervale $\langle x_i, x_{i+1} \rangle$ s dĺžkou intervalu $|x_{i+1} - x_i| = h$.

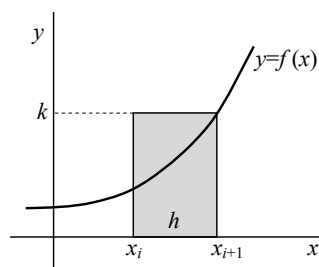
Geometrická interpretácia

Obsah rovinného útvaru, ktorý predstavuje určitý integrál, nahradíme obsahom obdĺžnika:

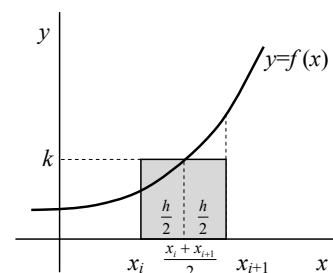
$$\int_{x_i}^{x_{i+1}} f(x) dx \stackrel{obd}{\approx} h \cdot k.$$



Obr. 7.2



Obr. 7.3



Obr. 7.4

Ak za uzlové body interpolácie zoberieme krajné body intervalu x_i, x_{i+1} , Lagrangeov interpolačný polynóm bude mať tvar $P_0(x) = f(x_i)$, $f(x_i) = k$ (Obr. 7.2). Analogickú aproximáciu môžeme urobiť s hodnotou funkcie v ktoromkoľvek známom bode $x \in \langle x_i, x_{i+1} \rangle$, čo obyčajne

ešte môže byť krajný bod intervalu x_{i+1} (Obr. 7.3), prípadne bod v strede intervalu $\frac{x_i + x_{i+1}}{2} = x_i + \frac{h}{2}$ (Obr. 7.4). K dispozícii tak máme obdĺžnikovú metódu:

Typ 1 – uzlový bod interpolácie je krajný bod intervalu:

a) $k = f(x_i)$ (Obr. 7.2)

$$\int_{x_i}^{x_{i+1}} f(x) dx \approx \int_{x_i}^{x_{i+1}} P_0(x) dx = \int_{x_i}^{x_{i+1}} f(x_i) dx = f(x_i)(x_{i+1} - x_i) = hf(x_i),$$

b) $k = f(x_{i+1})$ (Obr. 7.3)

$$\int_{x_i}^{x_{i+1}} f(x) dx \approx \int_{x_i}^{x_{i+1}} P_0(x) dx = \int_{x_i}^{x_{i+1}} f(x_{i+1}) dx = f(x_{i+1})(x_{i+1} - x_i) = hf(x_{i+1}).$$

Ak na intervale má $f(x)$ spojitú $f'(x)$, potom pre odhad absolútnej chyby metódy platí:

$$\varepsilon \leq \frac{M_1}{2}(x_{i+1} - x_i)^2 = \frac{M_1}{2}h^2, \quad M_1 = \max|f'(x)|, \quad x \in \langle x_i, x_{i+1} \rangle.$$

Typ 2 – uzlový bod interpolácie je bod v strede intervalu (Obr. 7.4):

$$k = f\left(\frac{x_i + x_{i+1}}{2}\right) = f\left(x_i + \frac{h}{2}\right)$$

$$\begin{aligned} \int_{x_i}^{x_{i+1}} f(x) dx &\approx \int_{x_i}^{x_{i+1}} P_0(x) dx = \int_{x_i}^{x_{i+1}} f\left(\frac{x_i + x_{i+1}}{2}\right) dx = f\left(\frac{x_i + x_{i+1}}{2}\right)(x_{i+1} - x_i) = hf\left(\frac{x_i + x_{i+1}}{2}\right) = \\ &= hf\left(x_i + \frac{h}{2}\right). \end{aligned}$$

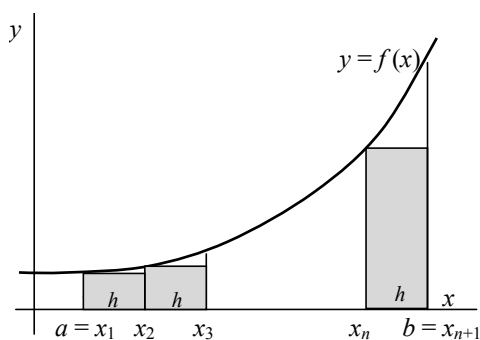
Ak na intervale má $f(x)$ spojitú $f''(x)$, potom pre odhad absolútnej chyby metódy platí:

$$\varepsilon = \frac{M_2}{24}(x_{i+1} - x_i)^3 = \frac{M_2}{24}h^3, \quad M_2 = \max|f''(x)|, \quad x \in \langle x_i, x_{i+1} \rangle.$$

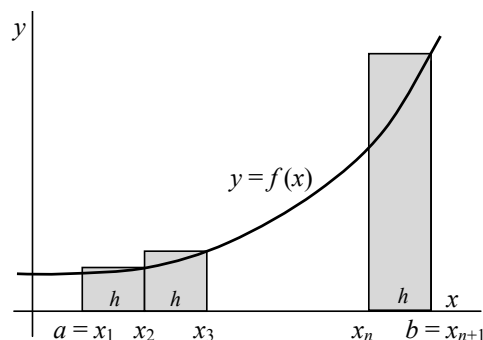
7.2.1.2 Zložená obdĺžniková metóda (ZOM)

Pracujeme na intervale $\langle a, b \rangle$, ktorý je zložený z n čiastkových intervalov $\langle x_i, x_{i+1} \rangle$ rovnakej dĺžky $h = \frac{b-a}{n}$. V každom podintervale na výpočet integrálu použijeme jednoduchú obdĺžnikovú metódu.

Typ 1



Obr. 7.5



Obr. 7.6

a) $k = f(x_i)$ (Obr. 7.5)

$$\begin{aligned}
 \int_a^b f(x) dx &= \int_{x_1}^{x_2} f(x) dx + \int_{x_2}^{x_3} f(x) dx + \dots + \int_{x_n}^{x_{n+1}} f(x) dx \approx \\
 &\approx \int_{x_1}^{x_2} P_0(x) dx + \int_{x_2}^{x_3} P_0(x) dx + \dots + \int_{x_n}^{x_{n+1}} P_0(x) dx = \\
 &= \int_{x_1}^{x_2} f(x_1) dx + \int_{x_2}^{x_3} f(x_2) dx + \dots + \int_{x_n}^{x_{n+1}} f(x_n) dx = \\
 &= hf(x_1) + hf(x_2) + \dots + hf(x_n) = \\
 &= h \sum_{i=1}^n f(x_i)
 \end{aligned} \tag{7.1}$$

b) $k = f(x_{i+1})$ (Obr. 7.6)

$$\begin{aligned}
 \int_a^b f(x) dx &= \int_{x_1}^{x_2} f(x) dx + \int_{x_2}^{x_3} f(x) dx + \dots + \int_{x_n}^{x_{n+1}} f(x) dx = \\
 &\approx \int_{x_1}^{x_2} f(x_2) dx + \int_{x_2}^{x_3} f(x_3) dx + \dots + \int_{x_n}^{x_{n+1}} f(x_{n+1}) dx = hf(x_2) + hf(x_3) + \dots + hf(x_{n+1}) = \\
 &= h \sum_{i=1}^n f(x_{i+1}) = h \sum_{i=2}^{n+1} f(x_i)
 \end{aligned} \tag{7.2}$$

Odhad absolútnej chyby metódy a rád metódy

Pri výpočte vychádzame z jednoduchšej metódy. Predpokladáme, že $f(x)$ má na celom intervale spojitú $f'(x)$:

- pre jeden čiastkový interval: $\varepsilon = \frac{M_1}{2}(x_{i+1} - x_i)^2 = \frac{M_1}{2}h^2$,
- pre n čiastkových intervalov: $\varepsilon = n \frac{M_1}{2} \left(\frac{b-a}{n} \right)^2 = \frac{M_1(b-a)^2}{2n} = \frac{M_1(b-a)}{2}h$.

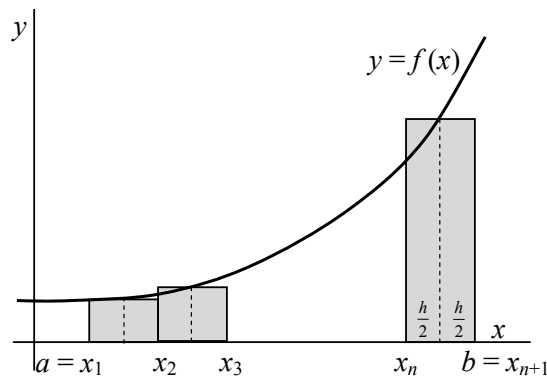
Veta 7.1 Nech funkcia $f(x)$ má na celom intervale spojitú prvú deriváciu $f'(x)$. Potom odhad absolútnej chyby zloženej obdĺžnikovej metódy (Typ 1) vypočítame podľa vzťahu:

$$\varepsilon = \frac{M_1(b-a)^2}{2n} = \frac{M_1(b-a)}{2}h, \quad M_1 = \max|f'(x)|, \quad x \in \langle a, b \rangle. \quad (7.3)$$

Rád metódy je $O(h)$ (pozri odsek 7.2.4).

Typ 2

$$k = f\left(\frac{x_i + x_{i+1}}{2}\right) = f\left(x_i + \frac{h}{2}\right) \text{ (Obr. 7.7)}$$



Obr. 7.7

$$\begin{aligned} \int_a^b f(x) dx &= \int_{x_1}^{x_2} f(x) dx + \int_{x_2}^{x_3} f(x) dx + \dots + \int_{x_n}^{x_{n+1}} f(x) dx \approx \\ &\approx \int_{x_1}^{x_2} f\left(x_1 + \frac{h}{2}\right) dx + \int_{x_2}^{x_3} f\left(x_2 + \frac{h}{2}\right) dx + \dots + \int_{x_n}^{x_{n+1}} f\left(x_n + \frac{h}{2}\right) dx \\ &= h \sum_{i=1}^n f\left(x_i + \frac{h}{2}\right) \end{aligned} \quad (7.4)$$

resp.

$$= h \sum_{i=1}^n f\left(\frac{x_i + x_{i+1}}{2}\right) \quad (7.5)$$

Odhad absolútnej chyby metódy a rád metódy

Tak ako pre jednoduchú metódu predpokladáme, že $f(x)$ má na intervale spojitú $f''(x)$:

- pre jeden čiastkový interval: $\varepsilon = \frac{M_2}{24}(x_{i+1} - x_i)^3 = \frac{M_2}{24}h^3$,
- pre n čiastkových intervalov: $\varepsilon = n \frac{M_2}{24} \left(\frac{b-a}{n}\right)^3 = \frac{M_2(b-a)}{24}h^2$.

Veta 7.2 Nech funkcia $f(x)$ má na celom intervale $\langle a, b \rangle$ spojitú druhú deriváciu $f''(x)$, potom odhad absolútnej chyby zloženej obdĺžnikovej metódy (Typ 2) stanovíme podľa vzťahu:

$$\varepsilon = \frac{M_2(b-a)^3}{24n^2} = \frac{M_2(b-a)}{24}h^2, \quad M_2 = \max|f''(x)|, \quad x \in \langle a, b \rangle. \quad (7.6)$$

Rád metódy je $O(h^2)$.

7.2.2 LICHOBĚŽNÍKOVÁ METÓDA

V lichobežníkovej metóde funkciu $f(x)$ nahradzame Lagrangeovým interpolačným polynómom stupňa $n = 1$, lineárnou funkciou (grafom je priamka), t. j. $f(x) \approx P_1(x)$.

7.2.2.1 Jednoduchá lichobežníková metóda

Funkciu $f(x)$ nahradzame na intervale $\langle x_i, x_{i+1} \rangle$ Lagrangeovým polynómom:

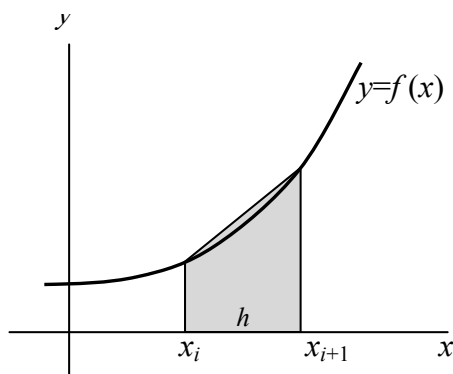
$$P_1(x) = f(x_i) \frac{x - x_{i+1}}{x_i - x_{i+1}} + f(x_{i+1}) \frac{x - x_i}{x_{i+1} - x_i},$$

dĺžka intervalu je $|x_{i+1} - x_i| = h$. Grafom $P_1(x)$ je priamka prechádzajúca bodmi $[x_i, f(x_i)]$, $[x_{i+1}, f(x_{i+1})]$.

Geometrická interpretácia

Obsah rovinného útvaru, ktorý predstavuje určitý integrál, nahradíme obsahom lichobežníka:

$$\int_{x_i}^{x_{i+1}} f(x) dx \stackrel{\text{Lich}}{\approx} h \frac{f(x_i) + f(x_{i+1})}{2}.$$



Obr. 7.8. Jednoduchá lichobežníková metóda

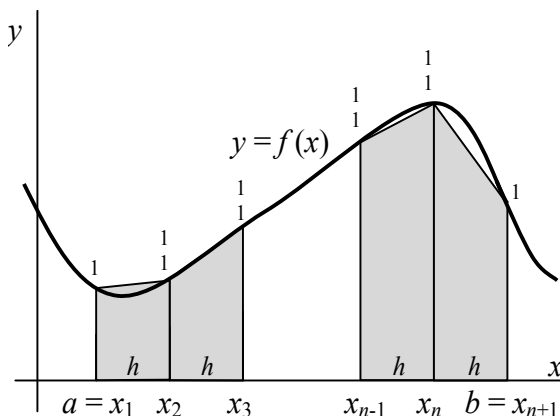
$$\begin{aligned} \int_{x_i}^{x_{i+1}} f(x) dx &\approx \int_{x_i}^{x_{i+1}} P_1(x) dx = \int_{x_i}^{x_{i+1}} \left(f(x_i) \frac{x - x_{i+1}}{x_i - x_{i+1}} + f(x_{i+1}) \frac{x - x_i}{x_{i+1} - x_i} \right) dx \\ &= \frac{x_{i+1} - x_i}{2} (f(x_i) + f(x_{i+1})) = \frac{h}{2} (f(x_i) + f(x_{i+1})) \end{aligned}$$

Ak $f(x)$ má na $\langle a, b \rangle$ spojitú $f''(x)$, potom pre odhad absolútnej chyby metódy platí:

$$\varepsilon = \frac{M_2}{12}(x_{i+1} - x_i)^3 = \frac{M_2}{12}h^3, \quad M_2 = \max|f''(x)|, \quad x \in \langle x_i, x_{i+1} \rangle.$$

7.2.2.2 Zložená lichobežníková metóda (ZLM)

Pracujeme na intervale $\langle a, b \rangle$, ktorý je zložený z n čiastkových intervalov $\langle x_i, x_{i+1} \rangle$ rovnakej dĺžky $h = \frac{b-a}{n}$.



Obr. 7.9. Zložená lichobežníková metóda

Na každom podintervale $\langle x_i, x_{i+1} \rangle$, použijeme jednoduchú lichobežníkovú metódu. Sčítaním vzorcov jednoduchej metódy cez všetky podintervaly získame výsledný vzorec ZLM:

$$\begin{aligned} \int_a^b f(x)dx &= \int_{x_1}^{x_2} f(x)dx + \int_{x_2}^{x_3} f(x)dx + \dots + \int_{x_n}^{x_{n+1}} f(x)dx \approx \\ &\approx \int_{x_1}^{x_2} P_1(x)dx + \int_{x_2}^{x_3} P_1(x)dx + \dots + \int_{x_n}^{x_{n+1}} P_1(x)dx = \\ &= \frac{h}{2}(f(x_1) + f(x_2)) + \frac{h}{2}(f(x_2) + f(x_3)) + \dots + \frac{h}{2}(f(x_n) + f(x_{n+1})) = \\ &= \frac{h}{2}(f(x_1) + f(x_{n+1}) + 2(f(x_2) + f(x_3) + \dots + f(x_n))) = \\ &= \frac{h}{2} \left(f(x_1) + f(x_{n+1}) + 2 \sum_{i=2}^n f(x_i) \right). \end{aligned}$$

Pre zloženú lichobežníkovú metódu teda dostávame:

$$\int_a^b f(x)dx \stackrel{\text{Lich}}{\approx} \frac{h}{2} \left(f(x_1) + f(x_{n+1}) + 2 \sum_{i=2}^n f(x_i) \right). \quad (7.7)$$

Keďže $x_1 = a$ a $x_{n+1} = b$, vzorec ZLM môžeme použiť aj v tvare:

$$\int_a^b f(x) dx \stackrel{\text{Lich}}{\approx} \frac{h}{2} \left(f(a) + f(b) + 2 \sum_{i=2}^n f(x_i) \right) \quad (7.8)$$

alebo pri použití označenia $y_i = f(x_i)$:

$$\int_a^b f(x) dx \stackrel{\text{Lich}}{\approx} \frac{h}{2} \left(y_1 + y_{n+1} + 2 \sum_{i=2}^n y_i \right). \quad (7.9)$$

Odhad absolútnej chyby metódy a rád metódy

Predpokladáme, že $f(x)$ má na $\langle a, b \rangle$ spojitú $f''(x)$. Odhad chyby ZLM vyjadríme opäť pomocou odhadu chyby pre jednoduchú metódu:

- pre jeden čiastkový interval: $\varepsilon = \frac{M_2}{12} (x_{i+1} - x_i)^3 = \frac{M_2}{12} h^3$,
- pre n čiastkových intervalov: $\varepsilon = n \frac{M_2}{12} \left(\frac{b-a}{n} \right)^3 = \frac{M_2 (b-a)^3}{12n^2} = \frac{M_2 (b-a)}{12} h^2$.

Veta 7.3 Nech funkcia $f(x)$ má na celom intervale $\langle a, b \rangle$ spojitú druhú deriváciu $f''(x)$. Potom odhad absolútnej chyby zloženej lichobežníkovej metódy vypočítame podľa vzťahu:

$$\varepsilon = \frac{M_2 (b-a)^3}{12n^2} = \frac{M_2 (b-a)}{12} h^2, \quad M_2 = \max |f''(x)|, \quad x \in \langle a, b \rangle. \quad (7.10)$$

Rád metódy je $O(h^2)$.

7.2.3 SIMPSONOVA METÓDA

V tejto metóde funkciu $f(x)$ nahradíme Lagrangeovým interpolačným polynómom stupňa $n = 2$, kvadratickou funkciou (grafom je parabola), t. j. $f(x) \approx P_2(x)$.

7.2.3.1 Jednoduchá Simpsonova metóda

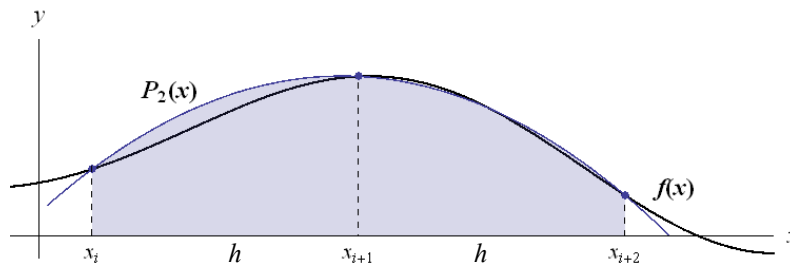
Funkciu $f(x)$ nahradzame Lagrangeovým polynómom 2. stupňa. Keďže na jednoznačné určenie polynómu 2. stupňa potrebujeme 3 body, základný interval bude obsahovať tri uzlové body x_i , x_{i+1} a x_{i+2} a budú ho tvoriť dva čiastkové intervaly $\langle x_i, x_{i+1} \rangle \cup \langle x_{i+1}, x_{i+2} \rangle = \langle x_i, x_{i+2} \rangle$.

$$P_2(x) = f(x_i) \frac{(x-x_{i+1})(x-x_{i+2})}{(x_i-x_{i+1})(x_i-x_{i+2})} + f(x_{i+1}) \frac{(x-x_i)(x-x_{i+2})}{(x_{i+1}-x_i)(x_{i+1}-x_{i+2})} + f(x_{i+2}) \frac{(x-x_i)(x-x_{i+1})}{(x_{i+2}-x_i)(x_{i+2}-x_{i+1})}$$

Dĺžka základného intervalu je $|x_{i+2} - x_i| = 2h$. Grafom $P_2(x)$ je parabola prechádzajúca bodmi $[x_i, f(x_i)]$, $[x_{i+1}, f(x_{i+1})]$, $[x_{i+2}, f(x_{i+2})]$.

Geometrická interpretácia

Obsah rovinného útvaru, ktorý predstavuje hľadaný určitý integrál, nahradíme obsahom krivočiareho lichobežníka ohraničeného osou x , priamkami $x = x_i$, $x = x_{i+2}$ a parabolou s analytickým vyjadrením $y = P_2(x)$ (Obr. 7.10).



Obr. 7.10. Jednoduchá Simpsonova metóda

$$\begin{aligned} \int_{x_i}^{x_{i+2}} f(x) dx &\approx \int_{x_i}^{x_{i+2}} P_2(x) dx = \\ &= \int_{x_i}^{x_{i+2}} \left(f(x_i) \frac{(x-x_{i+1})(x-x_{i+2})}{(x_i-x_{i+1})(x_i-x_{i+2})} + f(x_{i+1}) \frac{(x-x_i)(x-x_{i+2})}{(x_{i+1}-x_i)(x_{i+1}-x_{i+2})} + f(x_{i+2}) \frac{(x-x_i)(x-x_{i+1})}{(x_{i+2}-x_i)(x_{i+2}-x_{i+1})} \right) dx = \\ &= \frac{x_{i+2}-x_i}{6} (f(x_i) + 4f(x_{i+1}) + f(x_{i+2})) = \frac{2h}{6} (f(x_i) + 4f(x_{i+1}) + f(x_{i+2})). \end{aligned}$$

Pre jednoduchú Simpsonovu metódu dostaneme vzorec:

$$\int_{x_i}^{x_{i+2}} f(x) dx \stackrel{\text{Simps}}{\approx} \frac{h}{3} (f(x_i) + 4f(x_{i+1}) + f(x_{i+2})). \quad (7.11)$$

Ak $f(x)$ má na $\langle x_i, x_{i+2} \rangle$ spojitú štvrtú deriváciu $f^{(4)}(x)$, potom pre odhad absolútnej chyby metódy platí:

$$\varepsilon = \frac{M_4}{2880} (x_{i+2} - x_i)^5 = \frac{M_4}{2880} (2h)^5, \quad M_4 = \max |f^{(4)}(x)| \quad x \in \langle x_i, x_{i+2} \rangle.$$

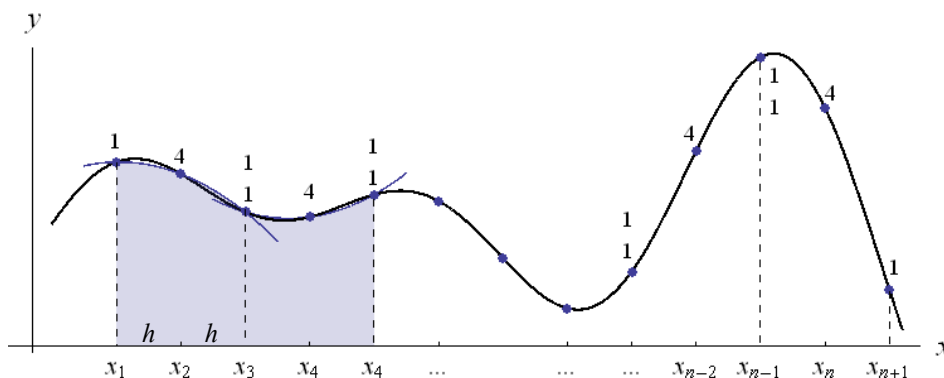
7.2.3.2 Zložená Simpsonova metóda (ZSM)

Pri konštruovaní zloženej Simpsonovej metódy si je treba uvedomiť, že Simpsonova metóda pracuje s trojicami uzlových bodov, vždy nad základným intervalom, ktorý tvoria dva čiastkové intervaly. Dĺžka čiastkového intervalu je h , dĺžka základného intervalu je $2h$. Ak dodržíme doteraz použité značenie, tak počet čiastkových intervalov je n , n musí byť párne číslo. Na označenie počtu základných intervalov zavedme novú premennú k , $2k = n$. Proces skladania môžeme riadiť buď z pohľadu čiastkového intervalu, keď budeme skladat' n čiastkových intervalov, alebo z pohľadu základného intervalu, keď skladáme k základných intervalov.

Prvý spôsob výpočtu – skladanie cez čiastkové intervaly

Pracujeme na intervale $\langle a, b \rangle$, ktorý je zložený z n čiastkových intervalov $\langle x_i, x_{i+1} \rangle$ rovnakej

dĺžky $h = \frac{b-a}{n}$.



Obr. 7.11. Zložená Simpsonova metóda (cez n)

$$\begin{aligned} \int_a^b f(x) dx &= \int_{x_1}^{x_3} f(x) dx + \int_{x_3}^{x_5} f(x) dx + \dots + \int_{x_{n-1}}^{x_{n+1}} f(x) dx \approx \\ &\approx \int_{x_1}^{x_3} P_2(x) dx + \int_{x_3}^{x_5} P_2(x) dx + \dots + \int_{x_{n-1}}^{x_{n+1}} P_2(x) dx = \\ &= \frac{h}{3} (f(x_1) + 4f(x_2) + f(x_3)) + \frac{h}{3} (f(x_3) + 4f(x_4) + f(x_5)) + \dots + \frac{h}{3} (f(x_{n-1}) + 4f(x_n) + f(x_{n+1})) \\ &= \frac{h}{3} \left(f(x_1) + f(x_{n+1}) + 4 \sum_{\substack{i=2 \\ \text{párne}}}^n f(x_i) + 2 \sum_{\substack{i=3 \\ \text{nepárne}}}^{n-1} f(x_i) \right). \end{aligned}$$

Pre zloženú Simpsonovu metódu teda dostávame:

$$\int_a^b f(x) dx \stackrel{\text{Simps}}{\approx} \frac{h}{3} \left(f(x_1) + f(x_{n+1}) + 4 \sum_{\substack{i=2 \\ \text{párne}}}^n f(x_i) + 2 \sum_{\substack{i=3 \\ \text{nepárne}}}^{n-1} f(x_i) \right). \quad (7.12)$$

Po dosadení $x_1 = a$ a $x_{n+1} = b$, vzorec ZSM môžeme použiť aj v tvare:

$$\int_a^b f(x) dx \stackrel{\text{Simps}}{\approx} \frac{h}{3} \left(f(a) + f(b) + 4 \sum_{\substack{i=2 \\ \text{párne}}}^n f(x_i) + 2 \sum_{\substack{i=3 \\ \text{nepárne}}}^{n-1} f(x_i) \right) \quad (7.13)$$

alebo pri použití označenia $y_i = f(x_i)$:

$$\int_a^b f(x) dx \stackrel{\text{Simps}}{\approx} \frac{h}{3} \left(y_1 + y_{n+1} + 4 \sum_{\substack{i=2 \\ \text{párne}}}^n y_i + 2 \sum_{\substack{i=3 \\ \text{nepárne}}}^{n-1} y_i \right). \quad (7.14)$$

Poznámka. $\sum_{i=2}^n y_i = y_2 + y_4 + \dots + y_n$, $\sum_{i=3}^{n-1} y_i = y_3 + y_5 + \dots + y_{n-1}$. V počítačovom spracovaní programujeme sumu cez $i =$ od 2 po n s krokom 2, resp. cez $i =$ od 3 po $n-1$ s krokom 2.

Odhad absolútnej chyby metódy a rád metódy

Ak $f(x)$ má na $\langle a, b \rangle$ spojitú štvrtú deriváciu $f^{(4)}(x)$, potom odhad zloženej metódy vyjadríme pomocou odhadu jednoduchej metódy:

- pre dva čiastkové intervaly: $\varepsilon = \frac{M_4}{2880} (x_{i+2} - x_i)^5 = \frac{M_4}{2880} (2h)^5$,
- pre $\frac{n}{2}$ dvojíc čiastkových intervalov: $\varepsilon = \frac{n}{2} \cdot \frac{M_4}{2880} \cdot (2h)^5 = \frac{n}{2} \cdot \frac{M_4}{2880} \cdot \left(2 \cdot \frac{b-a}{n}\right)^5 = \frac{M_4 \cdot 2^4 \cdot (b-a)^5}{2880n^4} = \frac{M_4 \cdot (b-a)^5}{180n^4} = \frac{M_4 \cdot (b-a)}{180} \cdot h^4$.

Veta 7.4 Nech funkcia $f(x)$ má na celom intervale $\langle a, b \rangle$ spojitú štvrtú deriváciu $f^{(4)}(x)$, potom pre odhad absolútnej chyby zloženej Simpsonovej metódy platí:

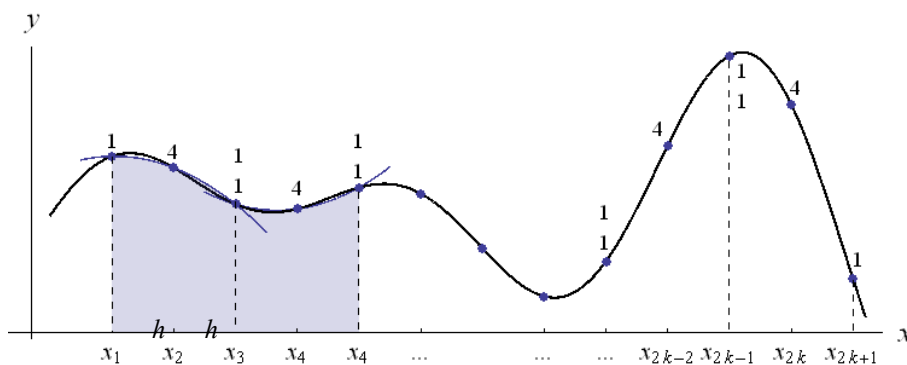
$$\varepsilon = \frac{M_4 \cdot (b-a)^5}{180 \cdot n^4} = \frac{M_4 \cdot (b-a)}{180} \cdot h^4, \quad M_4 = \max |f^{(4)}(x)| \quad x \in \langle a, b \rangle, \quad (7.15)$$

pričom h je dĺžka čiastkového intervalu, $h = \frac{b-a}{n}$, n počet čiastkových intervalov $\langle x_i, x_{i+1} \rangle$.

Rád metódy je $O(h^4)$.

Druhý spôsob výpočtu – skladanie cez základné intervaly

Pracujeme na intervale $\langle a, b \rangle$, ktorý je zložený z k základných intervalov $\langle x_i, x_{i+2} \rangle$, rovnakej dĺžky $2h = 2 \cdot \frac{b-a}{n}$. Základný interval tvoria dva čiastkové intervaly, počet základných intervalov je $k = \frac{n}{2}$, ($n = 2k$).



Obr. 7.12. Zložená Simpsonova metóda (cez $k = \frac{n}{2}$)

$$\begin{aligned}
 \int_a^b f(x) dx &= \int_{x_1}^{x_3} f(x) dx + \int_{x_3}^{x_5} f(x) dx + \dots + \int_{x_{2k-1}}^{x_{2k+1}} f(x) dx \approx \\
 &\approx \int_{x_1}^{x_3} P_2(x) dx + \int_{x_3}^{x_5} P_2(x) dx + \dots + \int_{x_{2k-1}}^{x_{2k+1}} P_2(x) dx = \\
 &= \frac{h}{3} (f(x_1) + 4f(x_2) + f(x_3)) + \frac{h}{3} (f(x_3) + 4f(x_4) + f(x_5)) + \dots + \frac{h}{3} (f(x_{2k-1}) + 4f(x_{2k}) + f(x_{2k+1})) \\
 &= \frac{h}{3} \left(f(x_1) + f(x_{2k+1}) + 4 \sum_{i=1}^k f(x_{2i}) + 2 \sum_{i=1}^{k-1} f(x_{2i+1}) \right) \tag{7.16}
 \end{aligned}$$

$$= \frac{h}{3} \left(f(a) + f(b) + 4 \sum_{i=1}^k f(x_{2i}) + 2 \sum_{i=1}^{k-1} f(x_{2i+1}) \right). \tag{7.17}$$

Odhad absolútnej chyby metódy a rád metódy

Postup odhadu absolútnej chyby pre zloženú metódu zopakujeme.

Ak $f(x)$ má na $\langle a, b \rangle$ spojitú štvrtú deriváciu $f^{(4)}(x)$, potom:

- pre jeden základný interval: $\varepsilon = \frac{M_4}{2880} (x_{i+2} - x_i)^5 = \frac{M_4}{2880} (2h)^5$,
- pre k základných intervalov: $\varepsilon = k \cdot \frac{M_4}{2880} \cdot (2h)^5 = k \cdot \frac{M_4}{2880} \cdot \left(\frac{b-a}{k}\right)^5 = \frac{M_4(b-a)^5}{2880k^4} =$
 $= \frac{M_4(b-a)}{2880} \cdot \left(\frac{b-a}{k}\right)^4 = \frac{M_4(b-a)}{2880} \cdot (2h)^4 = \frac{M_4(b-a)}{180} \cdot h^4$.

Veta 7.5 Nech funkcia $f(x)$ má na celom intervale $\langle a, b \rangle$ spojitú štvrtú deriváciu $f^{(4)}(x)$, potom odhad absolútnej chyby pre zloženú Simpsonovu metódu je

$$\varepsilon = \frac{M_4(b-a)^5}{2880k^4} = \frac{M_4 \cdot (b-a)}{180} \cdot h^4, \quad M_4 = \max |f^{(4)}(x)| \quad x \in \langle a, b \rangle, \tag{7.18}$$

pričom k je počet základných intervalov $\langle x_i, x_{i+2} \rangle$, h je dĺžka čiastkového intervalu, $h = \frac{b-a}{2k}$.

Rád metódy je $O(h^4)$.

Poznámka. Obidva spôsoby ZSM realizujú ten istý výpočet, líšia sa iba spôsobom zápisu, premennou, s ktorou pracujú. V 1. spôsobe je to počet čiastkových intervalov n , v 2. spôsobe je to počet základných intervalov k . Keďže $n = 2k$, výsledkom je tá istá približná hodnota integrálu ako aj odhadu chyby.

7.2.4 PRESNOSŤ ZLOŽENÝCH METÓD

Ak chceme vypočítať integrál $\int_a^b f(x) dx$ s požadovanou maximálnou chybou ε , potom pre každú z metód vieme z odhadu absolútnej chyby (pozri (7.3), (7.6), (7.10) a (7.18)) stanoviť minimálny počet čiastkových intervalov n , $n \in \mathbb{N}$, na ktorý treba rozdeliť interval $\langle a, b \rangle$.

Z odhadu absolútnej chyby metódy ďalej vyplýva, že presnosť každej z metód závisí od dĺžky čiastkového intervalu – kroku h , ktorým vykonáme alebo je dané delenie intervalu $\langle a, b \rangle$ na n čiastkových intervalov. Je zrejmé, že metódy sú presnejšie pri malých krokoch, **čím menší je krok, tým je chyba menšia**.

Závislosť chyby metódy od veľkosti kroku vyjadrujeme prostredníctvom mocniny r kroku h , – *rádom numerickej metódy* $O(h^r)$.

- Zložená obdĺžniková metóda využívajúca interpolačné uzly krajné body intervalu (Typ 1) je rádu $O(h)$. Odhad chyby (7.3) závisí od kroku lineárne, t. j. napríklad krok $h = 0,1$ v odhade chyby ε spôsobí zväčšenie 0,1-krát. Inými slovami, ak pri zjemňovaní kroku zmenšíme krok rádovo 10-krát (10^{-1}), odhad chyby sa rádovo zmenší 10-krát (10^{-1}).
- Zložená lichobežníková metóda ako aj zložená obdĺžniková metóda využívajúca interpolačný uzol stredný bod intervalu je rádu $O(h^2)$. Odhad chyby metódy (7.6) a (7.10) závisí od kroku kvadraticky, t. j. napríklad krok $h = 0,1$ spôsobí v odhade chyby ε zväčšenie 0,01-krát; inými slovami, ak zmenšíme krok rádovo 10-krát (10^{-1}), odhad chyby sa rádovo zmenší 100-krát (10^{-2}).
- Zložená Simpsonova metóda je rádu $O(h^4)$, odhad chyby metódy (7.18) závisí od 4. mocniny kroku, t. j. napríklad krok $h = 0,1$ v odhade chyby ε spôsobí zväčšenie 0,0001-krát alebo inak, ak zmenšíme krok rádovo 10-krát (10^{-1}), odhad chyby sa rádovo zmenší 10 000-krát (10^{-4}).

Z porovnania uvedených numerických metód vyplýva, že pri malých krokoch je vo výpočte integrálu $\int_a^b f(x) dx$ najpresnejšia Simpsonova metóda a najmenej presná je obdĺžniková metóda (Typ 1), ktorá sa v praxi používa len zriedka.

Označme $I_p = \int_a^b f(x) dx$ – presná hodnota integrálu, $I_{NM} = \int_a^b f(x) dx$ – hodnota integrálu vypočítaná príslušnou numericou metódou, $\varepsilon(I)$ – odhad absolútnej chyby pre príslušnú numericú metódu, potom platí:

$$\left| I_p - I_{NM} \right| \leq \varepsilon(I) \iff I_p = I_{NM} \pm \varepsilon(I) \iff I_p \in \langle I_{NM} - \varepsilon(I), I_{NM} + \varepsilon(I) \rangle.$$

Nech požadovaná presnosť na výpočet $\int_a^b f(x) dx$ je rádovo 10^{-3} . Znamená to, že potrebujeme zaručiť, aby $|I_p - I_{NM}| < \varepsilon$, $\varepsilon = O(10^{-3})$.

Realizujme výpočet integrálu zloženou lichobežníkovou metódou, ktorá je rádu $O(h^2)$. Ak zvolíme krok $h = 0,1 = 10^{-1}$, potom pre odhad absolútnej chyby platí:

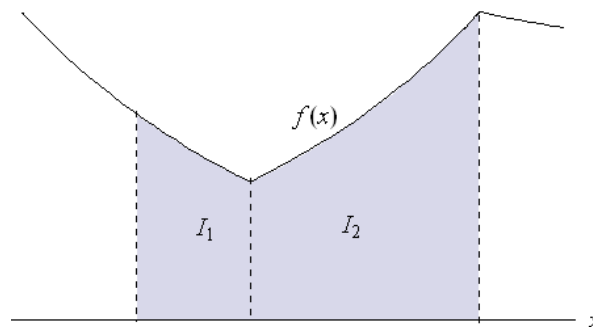
$$|I_p - I_{ZLM}| \leq O(h^2) \Leftrightarrow |I_p - I_{ZLM}| \leq O((10^{-1})^2) \Leftrightarrow |I_p - I_{ZLM}| \leq O(10^{-2}),$$

čo znamená, že pri výpočte integrálu sa dopúšťame absolútnej chyby rádovo 10^{-2} a menšej (na 2. desatinnom mieste a menej). Ak požadujeme presnosť výpočtu rádovo 10^{-3} , mali by sme sa dopustiť absolútnej chyby rádovo menšej ako na 3. desatinnom mieste, čo zapíšeme $|I_p - I_{ZLM}| \leq O(10^{-2}) < O(10^{-3})$. Vidíme, že neplatí $10^{-2} < 10^{-3}$, preto krok $h = 0,1$ nie je vhodný na výpočet integrálu uvedenou numerickou metódou pri požadovanej presnosti. Pre požadovanú presnosť a uvedenú numerickú metódu vyhovuje napr. krok $h = 0,01 = 10^{-2}$, pretože je splnená podmienka $|I_p - I_{ZLM}| \leq O((10^{-2})^2) < O(10^{-3})$.

Ak pri voľbe kroku $h = 0,1$ a požadovanej presnosti rádovo 10^{-3} použijeme na výpočet integrálu zloženú Simpsonovu metódu, ktorá je rádu $O(h^4)$, potom sme krok $h = 0,1$ zvolili správne, pretože je splnená podmienka $|I_p - I_{ZLM}| \leq O((10^{-1})^4) < O(10^{-3})$.

Potom v každej metóde z odhadu absolútnej chyby vieme stanoviť minimálny počet čiastkových intervalov n , $n \in \mathbb{N}$, na ktorý treba rozdeliť interval $\langle a, b \rangle$ (pozri riešené príklady).

Skutočnosť, že Simpsonova metóda je z uvádzaných metód najpresnejšia, môže viesť k domnienke, že väčšia presnosť výpočtu integrálu sa môže dosiahnuť použitím Newton-Cotesových vzorcov, v ktorých $f(x)$ nahradíme polynómom vyššieho stupňa. Všeobecne to ale neplatí. Interpoláčne polynómy vysokého stupňa sú veľmi zlou aproximáciou $f(x)$ a Newton-Cotesove vzorce by v takomto prípade dávali nezmyselné výsledky. Pri komplikovanejších priebehoch funkcií uprednostňujeme výpočet integrálu po častiach. Interval integrácie rozdelíme na vhodné časti (Obr. 7.13) a funkciu nahradzame po častiach polynómami nižších stupňov.



Obr. 7.13

7.3 RIEŠENÉ PRÍKLADY

PRÍKLAD č. 1

1. Vypočítajte určitý integrál $\int_{-1}^{0,2} \sqrt{1+x^2} dx$:

- zloženou obdĺžnikovou metódou (ľavý krajný bod intervalu) s krokom $h = 0,1$;
- zloženou lichobežníkovou metódou s krokom $h = 0,1$;
- zloženou Simpsonovou metódou s krokom $h = 0,1$.

Pre každú metódu spolu s približnou hodnotou integrálu vypočítajte aj odhad chyby metódy.

2. Pre každú numerickú metódu vypočítajte minimálny počet čiastkových intervalov, na ktoré treba rozdeliť interval $\langle -1; 0,2 \rangle$, ak chceme daný integrál vypočítať s presnosťou $\varepsilon = 10^{-4}$.

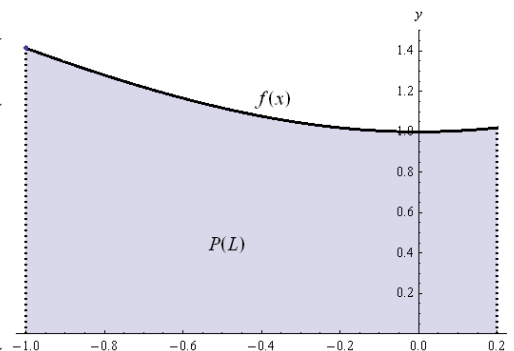
3. Použité numerické metódy porovnajte na základe získaných výsledkov.

Poznámky k riešeniu:

Funkcia $f(x) = \sqrt{1+x^2}$ je definovaná pre všetky $x \in \mathbb{R}$, pre ktoré platí $f(x) \geq 1$. Integrál riešime na intervale $\langle -1; 0,2 \rangle \subset D(f) = \mathbb{R}$, na ktorom je $f(x)$ spojitá a aj ohraničená.

Geometrická interpretácia

Chceme vypočítať plochu krivočiareho lichobežníka $P(L)$ ohraničeného grafom funkcie $f(x) = \sqrt{1+x^2}$, osou x a priamkami $x = -1$ a $x = 0,2$.



Obr. 7.14

Prv než pristúpime k riešeniu úlohy, uvedieme niekoľko dôležitých poznámok:

- Uvedený integrál sa dá vypočítať aj bežným matematickým spôsobom. Jeho výpočet je však náročný a komplikovaný, dávame do pozornosti iba výsledok:

$$\int_{-1}^{0,2} \sqrt{1+x^2} dx = \left[\frac{1}{2} \left(x \cdot \sqrt{1+x^2} + \arcsin h(x) \right) \right]_{-1}^{0,2} =$$

$$= \frac{1}{50} \left(25\sqrt{2} + \sqrt{26} + 25 \arcsin h\left(\frac{1}{5}\right) + 25 \arcsin h(1) \right) \doteq 1,34911902$$

- Pri výpočte odhadu chyby uvedených numerických metód integrovania bude potrebné určiť hodnotu $M_n = \max |f^{(n)}(x)|$ pre $x \in \langle -1; 0,2 \rangle$, $n = 1, 2, 4$. Je na mieste uvedomiť si základný rozdiel medzi funkciami $f^{(n)}(x)$ a $|f^{(n)}(x)|$:

- pre definičné obory týchto funkcií platí: $D(f^{(n)}(x)) = D(|f^{(n)}(x)|)$,
- pre obory hodnôt týchto funkcií rovnosť vo všeobecnosti neplatí:

$$H(f^{(n)}(x)) \neq H(|f^{(n)}(x)|), \text{ pretože } H(f^{(n)}(x)) \subset \mathbb{R}, H(|f^{(n)}(x)|) \subset \mathbb{R}_0^+.$$

Ak sú obe funkcie definované a spojité na uzavretom intervale $\langle a, b \rangle$, potom sú na tomto intervale aj ohraničené a pre určenie ich maxima M_n na intervale platí:

- $f^{(n)}(x)$: $M_n = \max\{f(a), f(b), f(x_f)\}$, kde $x_f \in \langle a, b \rangle$ sú stacionárne body $f^{(n)}(x)$,
- $|f^{(n)}(x)|$: $M_n = \max\{|f(a)|, |f(b)|, |f(x_{|f|})|\}$, kde $x_{|f|} \in \langle a, b \rangle$ sú stacionárne body $|f^{(n)}(x)|$ pričom $x_f = x_{|f|}$.

RIEŠENIE

1. Vypočítajte $\int_{-1}^{0,2} \sqrt{1+x^2} dx$.

Pretože na riešenie daného integrálu použijeme numerické metódy, je potrebné previesť funkciu $f(x) = \sqrt{1+x^2}$ danú predpisom na funkciu $f(x)$ danú tabuľkou s požadovaným krokom $h = 0,1$ na danom intervale (Tabuľka 7.1).

i	x	$f(x) = \sqrt{1+x^2}$
1	-1.	1.41421356
2	-0.9	1.3453624
3	-0.8	1.28062485
4	-0.7	1.22065556
5	-0.6	1.16619038
6	-0.5	1.11803399
7	-0.4	1.07703296
8	-0.3	1.04403065
9	-0.2	1.0198039
10	-0.1	1.00498756
11	0.	1.
12	0.1	1.00498756
13	0.2	1.0198039

Tabuľka 7.1

Stanovme:

- počet čiastkových intervalov $n = \frac{b-a}{h} = \frac{0,2 - (-1)}{0,1} = 12$,
- počet deliacich bodov, uzlov $n + 1 = 13$.

a) Výpočet integrálu zloženou obdĺžnikovou metódou (ľavý krajný bod intervalu) s krokom $h = 0,1$

Výpočet integrálu:

$$\int_{-1}^{0,2} \sqrt{1+x^2} dx \approx h \sum_{i=1}^n f(x_i) = h \sum_{i=1}^{12} f(x_i) = 0,1(f(-1) + f(-0,9) + \dots + f(0,1) + f(0,2)) = 0,1(1,41421356 + 1,3453624 + \dots + 1,00498756 + 1,0198039) = 1,36959234$$

Pre odhad chyby platí: $\varepsilon = \frac{M_1(b-a)^2}{2n}$, kde $M_1 = \max |f'(x)|$, $x \in \langle -1; 0,2 \rangle$.

Určenie M_1 : Vypočítajme najskôr 1. deriváciu funkcie $f(x)$: $f'(x) = \frac{x}{\sqrt{1+x^2}}$. Stacionárne

body funkcie $|f'(x)|$ nájdeme z podmienky: $(f'(x))' = 0 \Leftrightarrow f''(x) = 0$, pretože $x_f = x_{|f'|}$.

Potom $f''(x) = 0 \Leftrightarrow \frac{1}{\sqrt{(1+x^2)^3}} = 0$, odkiaľ vyplýva, že stacionárne body funkcie $f'(x)$ a

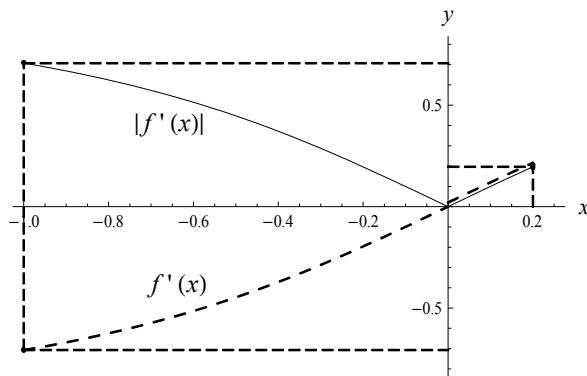
teda ani funkcie $|f'(x)|$ neexistujú. Preto

$$M_1 = \max\{|f'(-1)|, |f'(0,2)|\} = \max\left\{\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{26}}\right\} = \frac{1}{\sqrt{2}} \doteq 0,7071068.$$

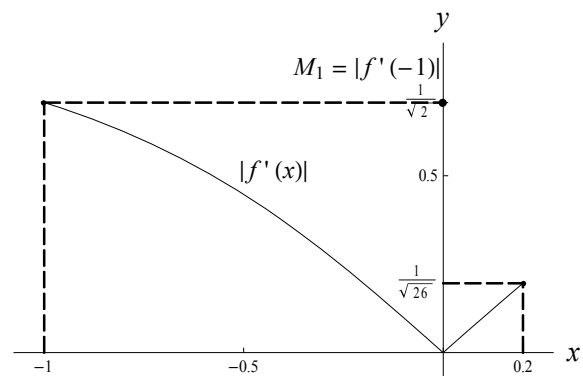
V prípade, že poznáme graf funkcie $|f'(x)|$ na danom intervale (Obr. 7.16), môžeme ho pri určovaní M_1 použiť. Z grafu ľahko určíme, či funkcia $|f'(x)|$ dosahuje maximum v jednom z koncových bodov intervalu. Ak je tomu tak, stacionárne body nepočítame a priamo vypočítame hodnotu maxima v príslušnom koncovom bode.

Odhad chyby pre ZOM:

$$\varepsilon = \frac{M_1(b-a)^2}{2n} = \frac{M_1(b-a)}{2} h = \frac{1}{\sqrt{2}} \frac{(-1-(0,2))}{2 \cdot 20} \cdot 0,1 = \frac{3}{50\sqrt{2}} \doteq 0,0424264.$$



Obr. 7.15. Graf $f'(x)$ a $|f'(x)|$



Obr. 7.16. Určovanie M_1 pomocou grafu

Záver. Zloženou obdĺžnikovou metódou (ľavý krajný bod intervalu) sme vypočítali:

$$\int_{-1}^{0,2} \sqrt{1+x^2} dx = 1,36959234 \pm 0.0424264$$

$$\int_{-1}^{0,2} \sqrt{1+x^2} dx \in \langle 1,36959234 - 0.0424264; 1,36959234 + 0.0424264 \rangle$$

$$\int_{-1}^{0,2} \sqrt{1+x^2} dx \in \langle 1,32716593; 1,41201875 \rangle.$$

b) Výpočet integrálu zloženou lichobežníkovou metódou s krokom $h = 0,1$

Výpočet integrálu:

$$\int_{-1}^{0,2} \sqrt{1+x^2} dx \approx \frac{Lich}{2} h \cdot \left(f(a) + f(b) + 2 \sum_{i=2}^n f(x_i) \right) = \frac{0,1}{2} \cdot \left(f(-1) + f(0,2) + 2 \sum_{i=2}^{12} f(x_i) \right) =$$

$$= \frac{0,1}{2} (f(-1) + f(0,2) + 2(f(-0,9) + \dots + f(0,1))) =$$

$$= 0,05 \cdot (1,4142136 + 1,0198039 + 2(1,3453624 + \dots + 1,00498756)) \doteq 1,34987186$$

Pre odhad chyby platí: $\varepsilon = \frac{M_2(b-a)^3}{12n^2}$, kde $M_2 = \max |f''(x)|$, $x \in \langle -1; 0,2 \rangle$.

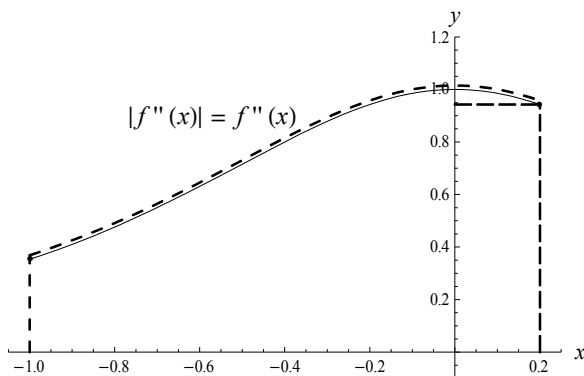
Určenie M_2 : Nájdime stacionárne body funkcie $|f''(x)|$ z podmienky:

$$(f''(x))' = 0 \Leftrightarrow f'''(x) = 0 \Leftrightarrow -\frac{3x}{\sqrt{(1+x^2)^5}} = 0 \Leftrightarrow x = 0.$$

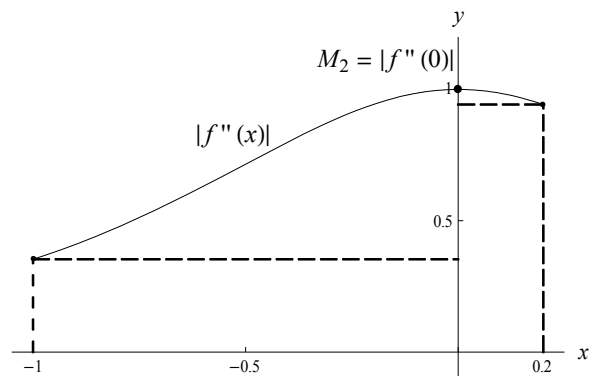
Stacionárny bod funkcie $f''(x)$, a teda aj funkcie $|f''(x)|$ je $x = 0$ a leží v intervale $\langle -1; 0,2 \rangle$.

Preto

$$M_2 = \max \{|f''(-1)|, |f''(0,2)|, |f''(0)|\} = \max \left\{ \frac{1}{2\sqrt{2}}, \frac{125}{26\sqrt{26}}, 1 \right\} \doteq \max \{0,35355; 0,94287; 1\} = 1.$$



Obr. 7.17. Graf $f''(x)$ a $|f''(x)|$



Obr. 7.18. $M_2 = \max |f''(x)|$, $x \in \langle -1; 0,2 \rangle$

Odhad chyby ZLM: $\varepsilon = \frac{M_2(b-a)^3}{12n^2} = \frac{M_2(b-a)}{12} h^2 = \frac{1 \cdot (0,2 - (-1))}{12} \cdot 0,1^2 = 0,001.$

Záver. Zloženou lichobežníkovou metódou sme vypočítali:

$$\int_{-1}^{0,2} \sqrt{1+x^2} dx = 1,34987186 \pm 0,001$$

$$\int_{-1}^{0,2} \sqrt{1+x^2} dx \in \langle 1,34987186 - 0,001; 1,34987186 + 0,001 \rangle$$

$$\int_{-1}^{0,2} \sqrt{1+x^2} dx \in \langle 1,34887186; 1,35087186 \rangle$$

c) Výpočet integrálu zloženou Simpsonovou metódou s krokom $h = 0,1$

Pri výpočte integrálu Simpsonovou metódou sa vyžaduje páry počet podintervalov. V našom prípade $n = 12$, čo je párne číslo, a teda Simpsonovou metódou môžeme vypočítať približnú hodnotu integrálu na celom intervale $\langle -1, 0,2 \rangle$.

Výpočet ukážeme pre obidva spôsoby skladania podintervalov:

Výpočet integrálu Simpsonovou metódou cez n :

$$\begin{aligned} \int_{-1}^{0,2} \sqrt{1+x^2} dx &\stackrel{\text{Simps}}{\approx} \frac{h}{3} \left(f(a) + f(b) + 4 \sum_{\substack{i=2 \\ \text{párne}}}^n f(x_i) + 2 \sum_{\substack{i=3 \\ \text{nepárne}}}^{n-1} f(x_i) \right) = \\ &= \frac{0,1}{3} \left(f(-1) + f(0,2) + 4 \sum_{\substack{i=2 \\ \text{párne}}}^{12} f(x_i) + 2 \sum_{\substack{i=3 \\ \text{nepárne}}}^{11} f(x_i) \right) = \\ &= \frac{0,1}{3} (f(-1) + f(0,2) + 4(f(x_2) + f(x_4) + \dots + f(x_{12})) + 2(f(x_3) + f(x_5) + \dots + f(x_{11}))) = \\ &= \frac{0,1}{3} \cdot (1,4142123 + 1,0198039 + 4 \cdot 26,95223092 + 2 \cdot 11,08730418) \doteq 1,34911842. \end{aligned}$$

Výpočet integrálu Simpsonovou metódou cez $k = \frac{n}{2}$:

$$\begin{aligned} \int_{-1}^1 \sqrt{1+x^2} dx &\stackrel{\text{Simps}}{\approx} \frac{h}{3} \left(f(a) + f(b) + 4 \sum_{i=1}^k f(x_{2i}) + 2 \sum_{i=1}^{k-1} f(x_{2i+1}) \right) = \\ &= \frac{0,1}{3} \left(f(-1) + f(0,2) + 4 \sum_{i=1}^6 f(x_{2i}) + 2 \sum_{i=1}^5 f(x_{2i+1}) \right) = \\ &= \frac{0,1}{3} (f(-1) + f(0,2) + 4(f(x_2) + f(x_4) + \dots + f(x_{12})) + 2(f(x_3) + f(x_5) + \dots + f(x_{11}))) = \\ &= \frac{0,1}{3} \cdot (1,4142123 + 1,0198039 + 4 \cdot 26,95223092 + 2 \cdot 11,08730418) \doteq 1,34911842. \end{aligned}$$

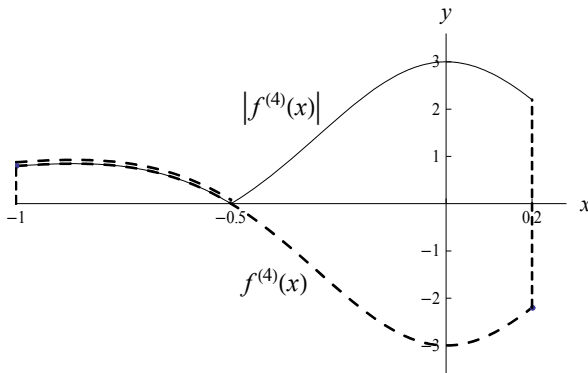
Pre odhad chyby platí : $\varepsilon = \frac{M_4(b-a)^5}{180n^4}$, kde $M_4 = \max |f^{(4)}(x)|$, $x \in \langle -1; 0,2 \rangle$.

Určenie M_4 pomocou grafu (Obr. 7.20):

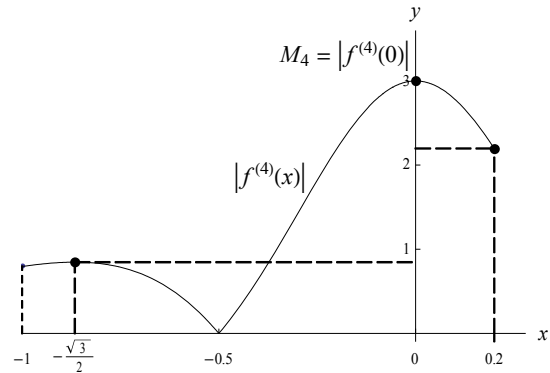
Na grafe $|f^{(4)}(x)|$ vidíme, že funkcia $|f^{(4)}(x)|$ nadobúda na intervale $\langle -1, 0,2 \rangle$ maximum pravdepodobne pre $x = 0$. Stacionárny bod $x = 0$ overíme výpočtom:

$$(f^{(4)}(x))' = 0 \Leftrightarrow f^{(5)}(x) = 0 \Leftrightarrow \frac{45x - 60x^3}{\sqrt{(1+x^2)^9}} = 0 \Leftrightarrow 45x - 60x^3 = 0 \Leftrightarrow x(45 - 60x^2) = 0.$$

Stacionárne body funkcie $f^{(4)}(x)$, a teda aj funkcie $|f^{(4)}(x)|$ na danom intervale sú $x = 0$, $x = -\frac{\sqrt{3}}{2} \doteq -0,86602540$, $x = \frac{\sqrt{3}}{2} \doteq 0,86602540$. Bod $x = 0$ je stacionárnym bodom, a teda $M_4 = |f^{(4)}(0)| = 3$.



Obr. 7.19. Graf $f^{(4)}(x)$ a $|f^{(4)}(x)|$



Obr. 7.20. $M_4 = \max |f^{(4)}(x)|$, $x \in \langle -1, 0, 2 \rangle$

Odhad chyby:

$$\varepsilon = \frac{M_4(b-a)^5}{2880n^4} = \frac{M_4 \cdot (b-a)}{180} \cdot h^4 = \frac{3 \cdot (0,2 - (-1))}{180} \cdot 0,1^4 = 0,000002.$$

Záver. Zloženou Simpsonovou metódou sme vypočítali:

$$\int_{-1}^{0,2} \sqrt{1+x^2} dx = 1,34911842 \pm 0,000002$$

$$\int_{-1}^{0,2} \sqrt{1+x^2} dx \in \langle 1,34911842 - 0,000002; 1,34911842 + 0,000002 \rangle$$

$$\int_{-1}^{0,2} \sqrt{1+x^2} dx \in \langle 1,34911642; 1,34912042 \rangle.$$

2. Pre každú numerickú metódu vypočítajte minimálny počet čiastkových intervalov, na ktoré treba rozdeliť interval $\langle -1, 1 \rangle$, ak chceme daný integrál vypočítať s presnosťou $\varepsilon = 10^{-4}$.

Pri stanovení minimálneho počtu čiastkových intervalov n (pozor, je to prirodzené číslo) vychádzame z odhadu chyby pre danú metódu. Na danom intervale môže daná zložená metóda nadobudnúť chybu maximálne:

• Zložená obdĺžniková metóda:

$$\varepsilon = \frac{M_1(b-a)^2}{2n} \stackrel{n \in \mathbb{N}}{\Rightarrow} n \geq \frac{M_1(b-a)^2}{2\varepsilon} \Rightarrow n \geq \frac{1}{\sqrt{2}} \cdot (0,2 - (-1))^2 \stackrel{n \in \mathbb{N}}{\Leftrightarrow} n \geq 5091,7 \Rightarrow n = 5092$$

Ak chceme metódou ZOM vypočítať $\int_{-1}^{0,2} \sqrt{1+x^2} dx$ s toleranciou $\varepsilon = 10^{-4}$, musíme interval $\langle -1; 0,2 \rangle$ rozdeliť na minimálne 5092 čiastkových intervalov (musíme sčítať plochu minimálne 5092 obdĺžnikov).

- Zložená lichobežníková metóda:

$$\varepsilon = \frac{M_2(b-a)^3}{12n^2} \stackrel{n \in \mathbb{N}}{\Rightarrow} n \geq \sqrt{\frac{M_2(b-a)^3}{12\varepsilon}} \Rightarrow n \geq \sqrt{\frac{1 \cdot (0,2 - (-1))^3}{12 \cdot 0,0001}} \Leftrightarrow n \geq 37,9473 \stackrel{n \in \mathbb{N}}{\Rightarrow} n = 38$$

Ak chceme metódou ZLM vypočítať $\int_{-1}^{0,2} \sqrt{1+x^2} dx$ s toleranciou $\varepsilon = 10^{-4}$, musíme interval $\langle -1; 0,2 \rangle$ rozdeliť na minimálne 38 čiastkových intervalov (musíme sčítať plochu minimálne 38 lichobežníkov).

- Zložená Simpsonova metóda:

Výpočet n (musí byť párne) ukážeme pre obidva spôsoby skladania podintervalov. Z odhadu chyby Simpsonovej metódy opäť predpokladáme, že na danom intervale môže ZSM nadobudnúť chybu maximálne:

pre Simpsonovu metódu cez n :

$$\varepsilon = \frac{M_4(b-a)^5}{180n^4} \stackrel{n \in \mathbb{N}}{\Rightarrow} n \geq \sqrt[4]{\frac{M_4(b-a)^5}{180\varepsilon}} \Rightarrow n \geq \sqrt[4]{\frac{3 \cdot (0,2 - (-1))^5}{180 \cdot 0,0001}} \Leftrightarrow n \geq 4,51272 \stackrel{n \text{ je párne}}{\Rightarrow} n = 6$$

pre Simpsonovu metódu cez k :

$$\varepsilon = \frac{M_4(b-a)^5}{2880k^4} \stackrel{k \in \mathbb{N}}{\Rightarrow} k \geq \sqrt[4]{\frac{M_4(b-a)^5}{2880\varepsilon}} \Rightarrow k \geq \sqrt[4]{\frac{3 \cdot (0,2 - (-1))^5}{2880 \cdot 0,0001}} \Leftrightarrow k \geq 2,25636 \stackrel{k \in \mathbb{N}}{\Rightarrow} k = 3,$$

pričom $n = 2k = 6$.

Ak chceme zloženou Simpsonovou metódou vypočítať $\int_{-1}^{0,2} \sqrt{1+x^2} dx$ s toleranciou $\varepsilon = 10^{-4}$, musíme interval $\langle -1; 0,2 \rangle$ rozdeliť na minimálne 6 čiastkových intervalov (musíme sčítať plochu minimálne $k = \frac{n}{2} = 3$ krivočiarych lichobežníkov).

3. Na základe získaných výsledkov porovnajte jednotlivé numerické metódy

Prehľad výsledkov je uvedený v tabuľke (Poznámka. $\langle c, d \rangle = \langle I_{NM} - \varepsilon_{NM}; I_{NM} + \varepsilon_{NM} \rangle$)

Tabuľka 7.2).

Výsledky výpočtov v súlade s rádom metódy ukázali, že daný integrál bol najpresnejšie vypočítaný Simpsonovou metódou (rád metódy je $O(h^4)$) a najmenej presne obdĺžnikovou metódou (rád $O(h)$).

NM	$\int_{-1}^{0,2} \sqrt{1+x^2} dx \in \langle c, d \rangle$	dĺžka $\langle c, d \rangle$	ε_{NM}	$n_{MIN} (\varepsilon = 10^{-4})$
obdĺžniková	$\langle 1,32716593; 1,41201875 \rangle$	$8,48528 \cdot 10^{-2}$	$4,24264 \cdot 10^{-2}$	5092
lichobežníková	$\langle 1,34887186; 1,35087186 \rangle$	$2 \cdot 10^{-3}$	$1 \cdot 10^{-3}$	38
Simpsonova	$\langle 1,34911642; 1,34912042 \rangle$	$4 \cdot 10^{-6}$	$2 \cdot 10^{-6}$	6
presne:	$\int_{-1}^{0,2} \sqrt{1+x^2} = \frac{1}{50} \left(25\sqrt{2} + \sqrt{26} + 25 \arcsin h\left(\frac{1}{5}\right) + 25 \arcsin h(1) \right) \doteq 1,34911902$			

Poznámka. $\langle c, d \rangle = \langle I_{NM} - \varepsilon_{NM}; I_{NM} + \varepsilon_{NM} \rangle$

Tabuľka 7.2. Prehľad výsledkov

Keďže integrál je možné vypočítať aj presne, odhad chyby máme možnosť porovnať so skutočnou absolútnou chybou (Tabuľka 7.3).

NM	$ I_p - I_{NM} $	ε_{NM}
obdĺžniková	$2,04733 \cdot 10^{-2}$	$4,24264 \cdot 10^{-2}$
lichobežníková	$7,52835 \cdot 10^{-4}$	$1 \cdot 10^{-3}$
Simpsonova	$6,01247 \cdot 10^{-7}$	$2 \cdot 10^{-6}$

Tabuľka 7.3

Odhad chyby ε_{NM} je pre všetky metódy väčší ako skutočná absolútna chyba $|I_p - I_{NM}|$.

8 NUMERICKÉ RIEŠENIE DIFERENCIÁLNYCH ROVNÍC

História diferenciálnych rovníc siaha až do 17. storočia a súvisí so vznikom infinitezimálneho počtu. Názov diferenciálna rovnica pochádza od nemeckého matematika G. W. Leibniza. V 18. storočí vznikla teória diferenciálnych rovníc, ktorej zakladateľom bol Euler. Známa je jeho metóda variácie konštánt, či riešenie diferenciálnej rovnice n -tého rádu s konštantnými koeficientami. V 19. storočí Cauchy a Lipschitz „priniesli“ do teórie diferenciálnych rovníc podmienky existencie a jednoznačnosti riešenia diferenciálnej rovnice. Ďalší významní matematici, ktorí sa zaoberali problémom diferenciálnych rovníc, boli bratia Bernoulli, d'Alembert, Lagrange či Fourier. Rozmach v teórii parciálnych diferenciálnych rovníc úzko súvisí s rozvojom fyziky.

8.1 ZÁKLADNÉ POJMY

Pri numerickom riešení diferenciálnej rovnice, vychádzame z poznatkov, ktoré sú známe zo základného kurzu matematiky.

Diferenciálna rovnica je rovnica, v ktorej sa vyskytuje neznáma funkcia a jej derivácie. Ak derivácie tejto funkcie sú obyčajné, hovoríme o obyčajnej diferenciálnej rovnici, ak sú parciálne, hovoríme o parciálnej diferenciálnej rovnici.

Budeme sa zaoberať obyčajnou diferenciálnou rovnicou 1. rádu.

Definícia 8.1 Obyčajná diferenciálna rovnica 1. rádu (ďalej len DR) je rovnica tvaru

$$F(x, y(x), y'(x)) = 0, \text{ stručne } F(x, y, y') = 0,$$

kde $y = y(x)$ je neznáma funkcia reálnej premennej x a $y'(x)$ je derivácia $y(x)$.

Definícia 8.2 Rád DR určuje najvyššia derivácia, ktorá vystupuje v DR. (DR nemusí obsahovať všetky uvedené členy.)

Pri riešení DR zapísanej v implicitnom tvare $F(x, y, y') = 0$ zvyčajne postupujeme tak, že ju najskôr, ak je to možné, upravíme na DR v explicitnom tvare $y'(x) = f(x, y(x))$, stručne $y' = f(x, y)$.

Funkcia $f(x, y)$ je reálnou funkciou 2 premenných x, y a je definovaná na oblasti $\Omega \subset E_2$:

$$D(f(x, y)) = \{\forall [x, y] \in \Omega \subset E_2; \Omega = D(x) \times D(y)\},$$

kde $D(x)$ je množina, na ktorej je definovaná premenná x a $D(y)$ je množina, na ktorej je definovaná premenná y .

Definícia 8.3 Riešením DR $y' = f(x, y)$ na intervale $I \subset R$ je každá funkcia $y = \varphi(x)$, pre ktorú platí:

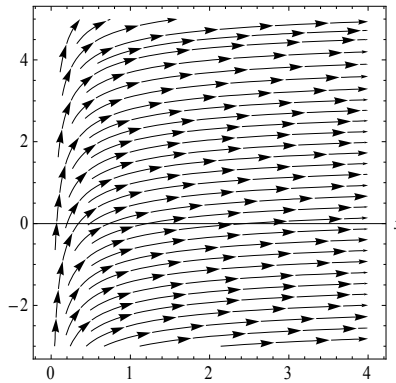
- $\varphi(x)$ je definovaná pre $\forall x \in I$,

- pre $\forall x \in I$ existuje $\varphi'(x)$,
- pre $\forall x \in I$ je $\varphi'(x) = f(x, \varphi(x))$.

Pretože DR sa často riešia integrovaním, riešenie DR sa nazýva aj **integrál** a **grafom** riešenia je **integrálna krivka**. Riešiť DR znamená nájsť všetky jej riešenia.

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

Nech $f(x, y)$ je spojitá na oblasti $\Omega \subset E_2$. Každému bodu $[\tilde{x}, \tilde{y}] \in \Omega$ je jednoznačne priradené číslo $f(\tilde{x}, \tilde{y})$. Predpokladajme, že bodom $[\tilde{x}, \tilde{y}]$ prechádza integrálna krivka $y = \varphi(x)$ diferenciálnej rovnice $y' = f(x, y)$, t. j. platí $\tilde{y} = \varphi(\tilde{x})$ a $y'(\tilde{x}) = f(\tilde{x}, y(\tilde{x})) = f(\tilde{x}, \tilde{y})$. Číslo $f(\tilde{x}, \tilde{y})$ je teda smernica dotyčnice k integrálnej krivke $y = \varphi(x)$ (riešeniu DR) prechádzajúcej príslušným bodom $[\tilde{x}, \tilde{y}]$. Všetky smernice vytvárajú na danej oblasti Ω pre danú diferenciálnu rovnicu **smerné pole** (Obr. 8.1).



Obr. 8.1. *Smerné pole diferenciálnej rovnice $2\sqrt{x}y' - 1 = 0$ na oblasti $\langle 0, 4 \rangle \times \langle -3, 5 \rangle$*

Definícia 8.4 *Všeobecným riešením DR $y' = f(x, y)$ je množina všetkých funkcií $y = \varphi(x, c)$, ktoré sa navzájom líšia o konštantu $c \in R$.*

Definícia 8.5 *Singulárnym riešením DR $y' = f(x, y)$ nazývame také riešenie DR, ktoré sa nedá dostať zo všeobecného riešenia DR žiadnou voľbou konštanty c . Každým bodom grafu singulárneho riešenia prechádza aspoň 1 riešenie DR. Toto riešenie môže byť obálkou alebo asymptotou všeobecných riešení.*

ÚLOHA 1

Ukážte výpočtom aj graficky, že funkcia $y = \frac{x^2}{4}$ je singulárnym riešením DR $y'^2 - xy' + y = 0$.

RIEŠENIE

$y'^2 - xy' + y = 0$ je DR, ktorú nevieme previesť na tvar $y' = f(x, y)$. Je to separovateľná DR, ktorá má na množine R ($x \in R$) všeobecné riešenie $y = cx - c^2$, $c \in R$. Presvedčme sa o tom.

Vypočítajme najskôr $y' = (cx - c^2)' = c$ a dosadíme za y' do DR. Dostaneme:

$$y'^2 - xy' + y = 0 \iff c^2 - xc + cx - c^2 = 0 \iff 0 = 0, \text{ platí pre } \forall x \in \mathbb{R}.$$

Overme singulárne riešenie DR.

Z predpisu funkcie $y = cx - c^2$, $c \in \mathbb{R}$ všeobecného riešenia DR je zrejmé, že sa jedná o lineárne funkcie – priamky, ktoré sa navzájom líšia konštantou $c \in \mathbb{R}$.

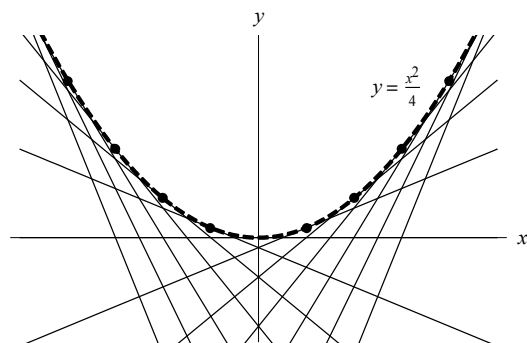
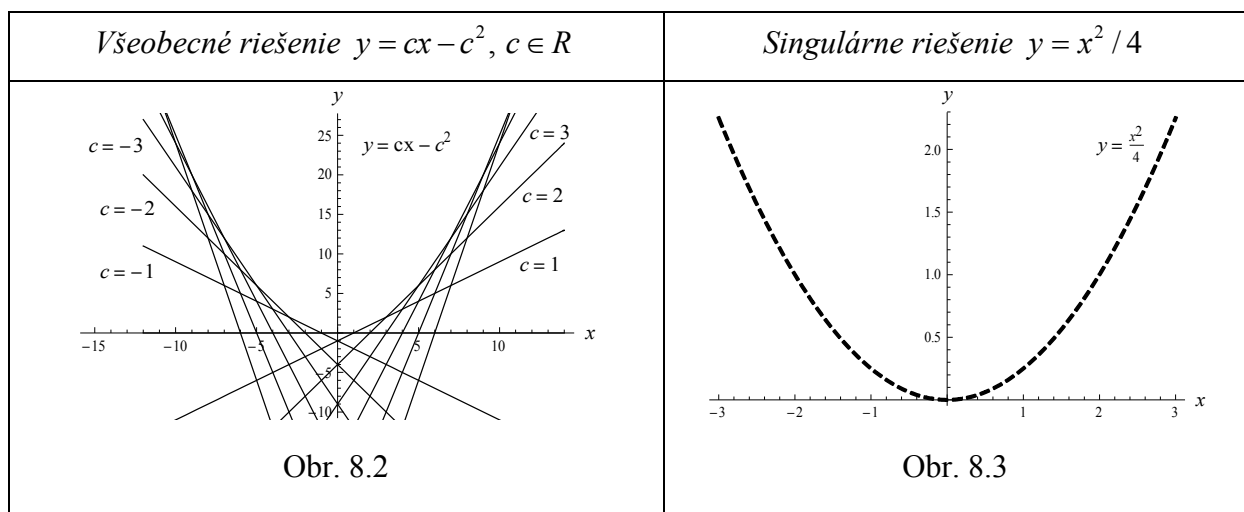
Z predpisu funkcie $y = \frac{x^2}{4}$ singulárneho riešenia DR vyplýva, že sa jedná o kvadratickú funkciu – parabolu, ktorú nie je možné dostať z lineárnej funkcie (zo všeobecného riešenia DR) žiadnou voľbou konštanty c .

- Overme výpočtom, či funkcia $y = \frac{x^2}{4}$ je singulárnym riešením danej DR:

Vypočítame $y' = \left(\frac{x^2}{4}\right)' = \frac{2x}{4} = \frac{x}{2}$ a dosadíme do DR, dostávame:

$$y'^2 - xy' + y = 0 \iff \left(\frac{x}{2}\right)^2 - x \frac{x}{2} + \frac{x^2}{4} = 0 \iff 0 = 0, \text{ platí pre } \forall x \in \mathbb{R}.$$

- Overme graficky, či funkcia $y = \frac{x^2}{4}$ je singulárnym riešením danej DR:



Obr. 8.4

Záver. Funkcia $y = \frac{x^2}{4}$ je singulárnym riešením danej DR. Každým bodom jej grafu prechádza graf aspoň jedného riešenia DR:(Obr. 8.4)

Definícia 8.6 Ak hľadáme riešenie DR $y' = f(x, y)$, ktoré vyhovuje podmienke $y(x_0) = y_0$, kde x_0, y_0 sú nejaké čísla, hovoríme o tzv. *Cauchyho úlohe (CU)* a podmienku $y(x_0) = y_0$ nazývame *Cauchyho začiatočná podmienka*.

Definícia 8.7 *Partikulárne riešenie DR* $y' = f(x, y)$ je výsledkom riešenia Cauchyho úlohy. Dostaneme ho tak, že do všeobecného riešenia DR $y = \varphi(x, c)$ dosadíme začiatočnú podmienku $y(x_0) = y_0$ a vypočítame konštantu c . Znamená to, že z množiny všetkých funkcií všeobecného riešenia vyberieme práve tú jednu funkciu $y = \varphi(x)$, ktorej graf prechádza bodom $[x_0, y_0]$.

ÚLOHA č. 2

Nájdite riešenie diferenciálnej rovnice $2\sqrt{x}y' - 1 = 0$, $y(4) = 2$.

RIEŠENIE

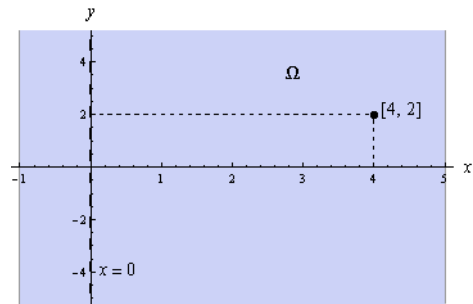
DR vieme upraviť na tvar $y' = f(x, y)$:

$$2\sqrt{x}y' - 1 = 0 \Rightarrow y' = \frac{1}{2\sqrt{x}} \Rightarrow f(x, y) = \frac{1}{2\sqrt{x}}$$

Určme definičný obor $f(x, y)$:

$$D(f(x, y)) = \{ \forall [x, y_i] \in \Omega \subset E_2; x > 0; y \in R \}$$

$$D(f(x, y)) = (0, \infty) \times R.$$



Obr. 8.5

Funkcia $f(x, y)$ je definovaná a spojitá na oblasti $\Omega = (0, \infty) \times R$, $[4, 2] \in \Omega$.

Všeobecné riešenie	Partikulárne riešenie
$2\sqrt{x}y' - 1 = 0$ $y' = \frac{1}{2\sqrt{x}} \quad / \int dx$ $y = \sqrt{x} + c, \quad c \in R$	$2\sqrt{x}y' - 1 = 0, \quad y(4) = 2$ $y = \sqrt{x} + c$ $2 = \sqrt{4} + c \Rightarrow c = 0$ $y = \sqrt{x}$
Obr. 8.6	Obr. 8.7

8.2 NUMERICKÉ METÓDY RIEŠENIA CAUCHYHO ÚLOHY

V technickej praxi sa často stretávame s problémami, ktorých jadrom je riešenie diferenciálnej rovnice. No iba pre relatívne málo DR vieme nájsť presné, analytické riešenie.

Približné, numerické metódy riešenia DR uprednostňujeme v prípade, že

- analytické riešenie neexistuje,
- analytické riešenie je zložité a náročné na výpočet.

Delíme ich na metódy:

- *analytické* – hľadáme presné riešenie v tvare funkcie na nejakom intervale. Patrí sem napr. metóda mocninových radov, metóda postupných aproximácií,
- *numerické* – s istou presnosťou hľadáme približné riešenie na zvolenej množine bodov (uzlov) z daného intervalu,
- *grafické* – metódy založené na jednoduchých geometrických konštrukciách. Používali sa skôr v minulosti.

Poznámka. Pod analytickým riešením DR rozumieme presné riešenie DR, funkciu danú predpisom.

Napr. analytickým riešením DR $1 + e^y \cdot y' = 0$ je funkcia $y = \ln(c - x)$, $c \in \mathbb{R}$.

Podrobnejšie sa budeme zaoberať numerickými metódami na riešenie obyčajnej diferenciálnej rovnice 1. rádu so začiatočnou podmienkou (Cauchyho úloha).

ÚLOHA č. 3

Riešte Cauchyho úlohu: $y' = f(x, y)$, $y(x_0) = y_0$ na intervale $\langle a, b \rangle$.

RIEŠENIE

Hľadáme partikulárne riešenie definované na intervale $\langle a, b \rangle$, čiže takú funkciu $y = \varphi(x)$, že $\varphi(x_0) = y(x_0) = y_0$, $x_0 = a$, $y_0 \in \mathbb{R}$. Funkcia $f(x, y)$ je funkcia spojitá na oblasti

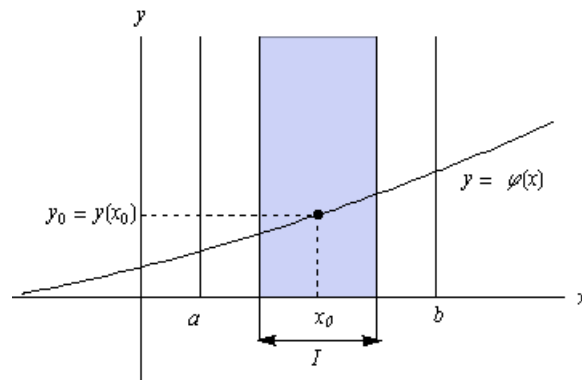
$$\Omega = \langle a, b \rangle \times M \subset E_2, M \subset \mathbb{R}.$$

Pri numerickom riešení CU je potrebné zaručiť existenciu a jednoznačnosť jej riešenia.

Postačujúce podmienky existencie a jednoznačnosti riešenia CU

Veta 8.1 Nech je daná DR $y' = f(x, y)$. Nech funkcie $f(x, y)$ a $\frac{\partial f(x, y)}{\partial y}$ sú spojité na ob-

lasti $\Omega = \langle a, b \rangle \times M \subset E_2$, kde $M \subset \mathbb{R}$. Potom pre ľubovoľný pevný bod $[x_0, y_0] \in \Omega$ existuje práve jedno riešenie (funkcia) $y = \varphi(x)$ danej DR, ktoré spĺňa začiatočnú podmienku $\varphi(x_0) = y_0$. Toto riešenie je definované na takom intervale $I \subseteq \langle a, b \rangle$, že $x_0 \in I$ a pre každé $x \in I$ je $[x, \varphi(x)] \in \Omega$.



Obr. 8.8

Spojitosť $f(x, y)$ na oblasti Ω zabezpečuje existenciu riešenia CU.

Spojitosť $\frac{\partial f(x, y)}{\partial y}$ na oblasti Ω zabezpečuje jednoznačnosť riešenia CU.

Splnenie postačujúcich podmienok existencie a jednoznačnosti riešenia $y' = f(x, y)$ na oblasti Ω však nezaručuje existenciu riešenia na celej oblasti Ω , ale len v istom okolí bodu x_0 . Napríklad pre CU $y' = y^2 + 1$, $y(0) = 0$ sú postačujúce podmienky existencie a jednoznačnosti riešenia splnené na oblasti $\Omega = R \times R = E_2$, ale jej riešenie $y = \operatorname{tg}(x)$ je definované len na intervale $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \subset R$.

Ak daná CU nespĺňa podmienky vyslovenej vety, potom používame na dôkaz existencie a jednoznačnosti jej riešenia iné (zoslabené) podmienky, napr. Lipschitzovu alebo Osgoodovu podmienku.

Veta 8.2 Ak funkcia $f(x, y)$ je spojitá na oblasti $\Omega = \langle a, b \rangle \times R \subset E_2$ a na intervale $\langle a, b \rangle$ spĺňa Lipschitzovu podmienku vzhľadom na premennú y , t. j. existuje také číslo L , že pre $\forall x \in \langle a, b \rangle$ a $\forall y_1, y_2 \in R$ platí $|f(x, y_1) - f(x, y_2)| \leq L|y_1 - y_2|$. Potom existuje práve jedno riešenie rovnice $y' = f(x, y)$ spĺňajúce začiatočnú podmienku $y(x_0) = y_0$, kde $x_0 \in \langle a, b \rangle$ a $y_0 \in R$.

Pre praktické aplikácie je jednoduchšie použiť podmienky uvedené vo vete Veta 8.1.

8.2.1 PRINCÍP NUMERICKÝCH METÓD RIEŠENIA CU

Riešime Cauchyho úlohu: $y' = f(x, y)$, $y(x_0) = y_0$ na intervale $\langle a, b \rangle$, $x_0 = a$.

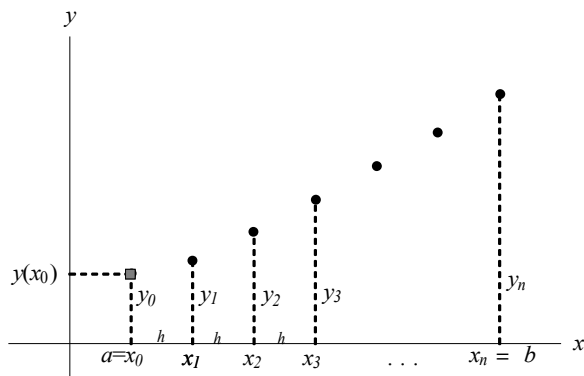
Nech $y(x)$ je presné partikulárne riešenie CU, $x \in \langle a, b \rangle$. Pod numerickým riešením CU budeme rozumieť funkciu danú tabuľkou (Tabuľka 8.1) – množinu usporiadaných dvojíc $\{[x_i, y_i], i = 0, 1, 2, \dots, n\}$, kde $y_i \approx y(x_i)$. Súradnice x_i a y_i získame nasledovne:

1. súradnice x_i , sú deliace body (uzly) x_0, x_1, \dots, x_n intervalu $\langle a, b \rangle$, ktorými je interval $\langle a, b \rangle$ rozdelený na n čiastočných intervalov $\langle x_i, x_{i+1} \rangle$, $i = 0, 1, \dots, n-1$ s dĺžkou $h = \frac{b-a}{n}$ (krok metódy). Počet uzlov je $n+1$, majú ekvidištantné delenie a sú usporiadané vzostupne: $x_0 < x_1 < \dots < x_n$. Začiatkový uzol x_0 je známy zo začiatkovej podmienky $y(x_0) = y_0$, $x_0 = a$. Každý ďalší uzol vypočítame podľa iteračného vzťahu $x_{i+1} = x_i + h$, čiže $|x_{i+1} - x_i| = h = \text{konšt.}$, $i = 0, 1, \dots, n-1$; pre posledný uzol x_n platí $x_n = b$.

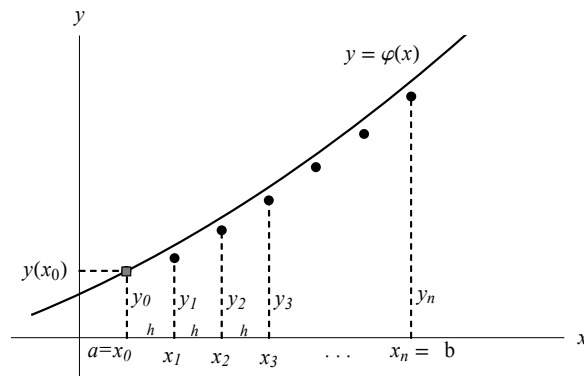
x_i	y_i
x_0	y_0
x_1	y_1
x_2	y_2
x_3	y_3
x_4	y_4
...	...
x_n	y_n

Tabuľka 8.1

2. y -ová súradnica v začiatkovom bode intervalu $\langle a, b \rangle$ je známa zo začiatkovej podmienky $y(x_0) = y_0$, $x_0 = a$. Ďalšie súradnice y_i , $i = 1, \dots, n$ vypočítame pomocou iteračných vzťahov zvolenej numerickej metódy.



Obr. 8.9. Graf numerického riešenia DR



Obr. 8.10. Graf numerického a presného riešenia DR

Základná vlastnosť, ktorú požadujeme od každej numerickej metódy riešenia CU je, aby pri znižovaní kroku h , $h \rightarrow 0^+$, množina hodnôt $\{y_0, y_1, y_2, \dots, y_n\}$ približného riešenia konvergovala k množine hodnôt $\{y(x_0), y(x_1), y(x_2), \dots, y(x_n)\}$ presného partikulárneho riešenia, čo vyjadríme zápisom

$$\lim_{\substack{h \rightarrow 0^+ \\ (i \rightarrow \infty)}} y_i^{(h)} = y(x_i).$$

Pre $h \rightarrow 0^+$ počet uzlov rastie, $i \rightarrow \infty$. Rýchlosť konvergencie numerickej metódy závisí od kvality použitej numerickej metódy, jej rádu.

Definícia 8.8 Rád p numerickej metódy na riešenie CU je také najväčšie $p \in \mathbb{N}$, že pre pevné $h \rightarrow 0^+$ platí, že lokálna chyba numerickej metódy je rádu $O(h^{p+1})$ a globálna chyba numerickej metódy je rádu $O(h^p)$.

8.2.2 CHYBY NUMERICKÝCH METÓD A ICH RÁD

Lokálna chyba d_i

Tejto chyby sa dopustíme v uzle x_i pri jednom použití vzorca numerickej metódy. Je to rozdiel približného riešenia y_i a presného riešenia $\tilde{y} = \varphi(x)$, ktoré v predchádzajúcom uzle x_{i-1} spĺňa začiatočnú podmienku, t. j. nadobúda hodnotu $\tilde{y}(x_{i-1}) = y_{i-1}$,

$$d_i = |y_i - \tilde{y}(x_i)|.$$

Graf presného riešenia $\tilde{y} = \varphi(x)$ prechádza bodom $[x_{i-1}, y_{i-1}]$ (Obr. 8.11).

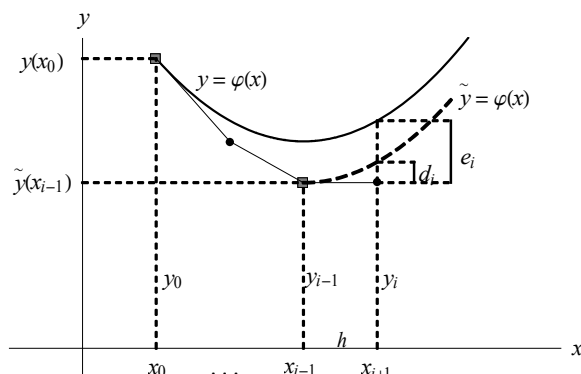
Globálna (akumulovaná) chyba e_i

Je chyba, ktorej sa dopustíme v uzle x_i po viacnásobnom použití vzorca numerickej metódy. Je výsledkom akumulácie (súčtu) lokálnych chýb pri postupe výpočtu hodnôt približného riešenia y_1, y_2, \dots, y_n DR, $y(x_0) = y_0$.

Globálna chyba je odchýlka presného partikulárneho a približného partikulárneho riešenia v uzle x_i :

$$e_i = |y(x_i) - y_i|.$$

Graf presného riešenia $y = \varphi(x)$ prechádza bodom $[x_0, y_0]$ (Obr. 8.11).



Obr. 8.11

Chyba numerickej metódy.

Rozumieme pod ňou maximálnu hodnotu z globálnych chýb v uzloch x_i , $i = 1, 2, \dots, n$

$$E(h) = \max_{i=1,2,\dots,n} |y_i - y(x_i)|.$$

Veličina $E(h)$ charakterizuje presnosť numerickeho riešenia vychádzajúceho z bodu $[x_0, y_0]$ a je závislá na veľkosti kroku h (čím menší krok h , tým je chyba menšia).

Rád chyby numerickej metódy.

V definícii Definícia 8.8 sme uviedli, že lokálna chyba numerickej metódy je rádu $O(h^{p+1})$ a globálna chyba numerickej metódy je rádu $O(h^p)$. Čo to v praxi znamená?

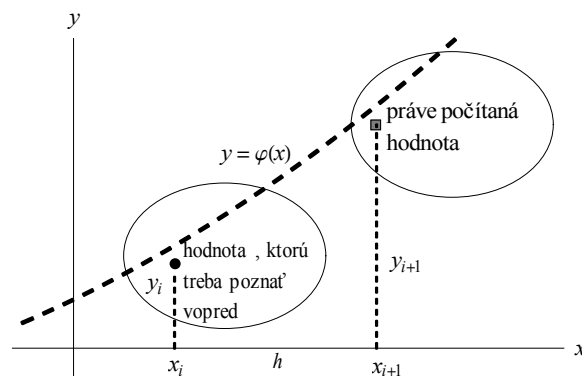
Napr. ak pri numerickom riešení CU zvolíme krok $h = 0,1 = 10^{-1}$ a numerickú metódu 3. rádu, potom:

- lokálna chyba tejto numerickej metódy je rádu $O((10^{-1})^4) = O(10^{-4})$, t. j. pri výpočte numerického riešenia sa dopúšťame lokálnej chyby na 4. desatinnom mieste – v desiatich tisícinách,
- globálna chyba tejto numerickej metódy je rádu $O((10^{-1})^3) = O(10^{-3})$, t. j. pri výpočte numerického riešenia sa dopúšťame globálnej chyby na 3. desatinnom mieste – v tisícinách.

Podrobnejšie pozri podkapitolu 8.3.5.

8.2.3 ROZDELENIE NUMERICKÝCH METÓD NA RIEŠENIE CU

1. Jednokrokové metódy – metódy typu Runge-Kutta



Obr. 8.12

Sú to metódy, ktoré na výpočet približnej hodnoty y_{i+1}

- vyžadujú poznať vopred len jednu predchádzajúcu hodnotu y_i ,
- používajú hodnoty funkcie $f(x, y)$ (smernice dotyčníc) v jednotlivých uzloch $x_i, i = 1, \dots, n$, a v bodoch medzi uzlami.

Patria sem metódy:

- Eulerova metóda,
- 1. a 2. (tzv. Heunova metóda) modifikovaná Eulerova metóda,
- metóda Runge – Kutta 3. rádu,
- metóda Runge – Kutta 4. rádu.

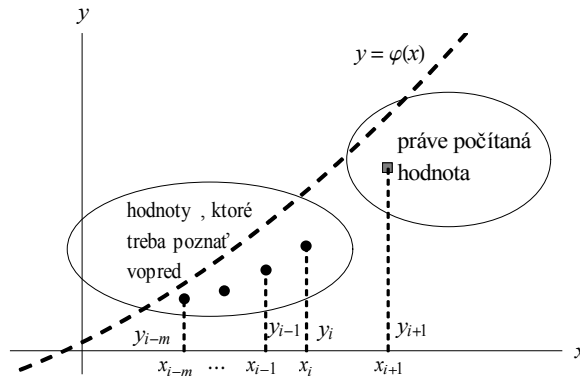
V minulosti sa používali ďalšie metódy typu Runge – Kutta 4. rádu, napr. Gillova metóda, 3/8 metóda. Tieto mali malé výhody pri realizácii na počítači, ale ich hlavným nedostatkom bolo, že neumožňovali jednoduchý odhad lokálnej chyby.

V roku 1958 R. H. Merson odvodil jednoduchú metódu typu Runge – Kutta, tiež 4. rádu (pridaním 1 smernice k_5 navyše), ktorá umožňovala jednoduchý odhad lokálnej chyby.

Na Slovensku metódy typu Runge – Kutta vyšších rádov skúmal A. Huťa.

V súčasnosti sa veľmi dobre osvedčujú metódy typu Runge – Kutta – Fehlberg, ktoré pri jednom kroku používajú výpočet 6-tich smerníc).

2. Viackrokové metódy



Obr. 8.13

Sú to metódy, realizáciou ktorých na výpočet približnej hodnoty y_{i+1}

- potrebujeme poznať niekoľko (viac ako jednu) predchádzajúcich približných hodnôt $y_i, y_{i-1}, \dots, y_{i-m}$,
- používajú hodnoty funkcie $f(x, y)$ (smernice dotyčníc) len v jednotlivých uzloch x_i , $i = 1, \dots, n$.

Patria sem metódy

- Adams – Bashfortove,
- Adams – Moultonove,
- metóda numerickej integrácie,
- metóda neurčitých koeficientov,
- metódy typu prediktor – korektor.

8.3 JEDNOKROKOVÉ METÓDY RIEŠENIA CAUCHYHO ÚLOHY

Riešme numericky Cauchyho úlohu: $y' = f(x, y)$, $y(x_0) = y_0$ na intervale $\langle a, b \rangle$.

Majme ekvidištantné delenie intervalu $\langle a, b \rangle$ s krokom $h = x_{i+1} - x_i$: $a = x_0 < x_1 < \dots < x_n = b$.

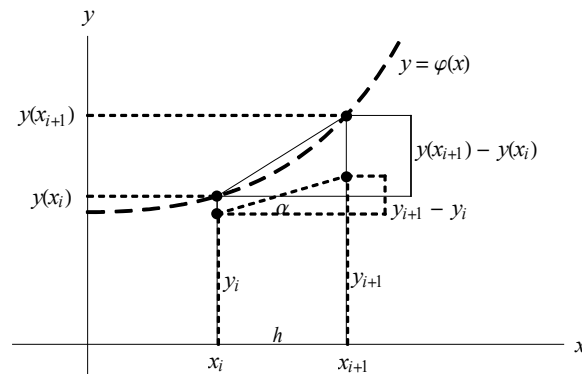
Pre DR v uzle x_i platí: $y'(x_i) = f(x_i, y(x_i))$. Pretože presné riešenie DR nepoznáme, deriváciu $y'(x_i)$ nahradíme pomernou diferenciou $\frac{y(x_{i+1}) - y(x_i)}{x_{i+1} - x_i}$ a presné hodnoty $y(x_i)$, $y(x_{i+1})$

ich aproximáciami y_i , y_{i+1} , potom (Obr. 8.14)

$$y'(x_i) = f(x_i, y(x_i)) \approx \frac{y_{i+1} - y_i}{h} = f(x_i, y_i). \quad (8.1)$$

Diferenciálnu rovnicu $y' = f(x, y)$ nahradíme diferenčnou rovnicou

$$\frac{y_{i+1} - y_i}{h} = f(x_i, y_i), \quad i = 0, 1, \dots, n-1. \quad (8.2)$$



Obr. 8.14

Úpravou (8.2) dostávame

$$y_{i+1} = y_i + f(x_i, y_i) \cdot h \quad \text{resp.} \quad y_{i+1} = y_i + f(x_i, y_i) \cdot (x_{i+1} - x_i) \quad (8.3)$$

Rovnica (8.3) predstavuje rovnicu priamky so smernicou $f(x_i, y_i) = \frac{y_{i+1} - y_i}{h}$ vedenej bodmi $X_i[x_i, y_i]$ a $X_{i+1}[x_{i+1}, y_{i+1}]$ (Obr. 8.15).

Vychádzajúc z uvedených skutočností, môžeme potom **jednokrokovú metódu** riešenia **Cauchyho úlohy** všeobecne zapísať iteračným vzťahom

$$y_{i+1} = y_i + h \cdot S(x_i, y_i, h), \quad i = 0, 1, 2, \dots, n-1 \quad (8.4)$$

kde:

- y_i je hodnota približného riešenia v uzle x_i (poznáme),
- y_{i+1} je hodnota približného riešenia v uzle x_{i+1} (chceme vypočítať),
- $h = x_{i+1} - x_i$ je krok metódy,
- $S(x_i, y_i, h)$ je smerová alebo prírastková funkcia.

Zo vzťahu (8.4) dostávame :

- $y_{i+1} - y_i = h \cdot S(x_i, y_i, h) = \Delta y_i$, kde Δy_i je prírastok y_i v uzle x_{i+1} , ktorý udáva, o koľko sa zmení hodnota y_i , keď sa posunieme o 1 krok dopredu $x_{i+1} = x_i + h$.

- $S(x_i, y_i, h) = \frac{y_{i+1} - y_i}{h} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} = \text{tg}(\alpha) = k_i$, $i = 0, 1, \dots, n-1$, kde k_i je smernica priamky

vedenej bodmi $X_i[x_i, y_i]$ a $X_{i+1}[x_{i+1}, y_{i+1}]$, α je uhol uvedenej priamky s osou x . Funkcia

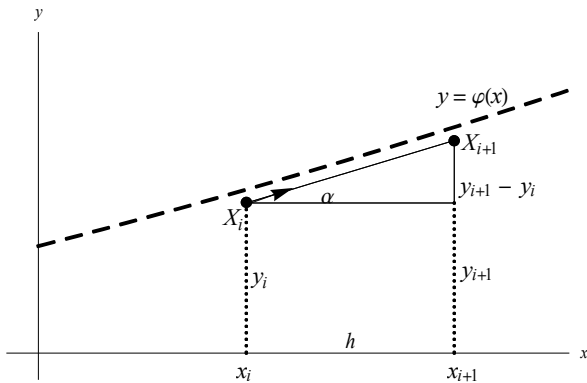
$S(x_i, y_i, h)$ udáva smer postupu od bodu $X_i[x_i, y_i]$ do bodu $X_{i+1}[x_{i+1}, y_{i+1}]$. (Obr. 8.15).

Geometrická interpretácia procesu numerického riešenia CU jedнокrokovou metódou

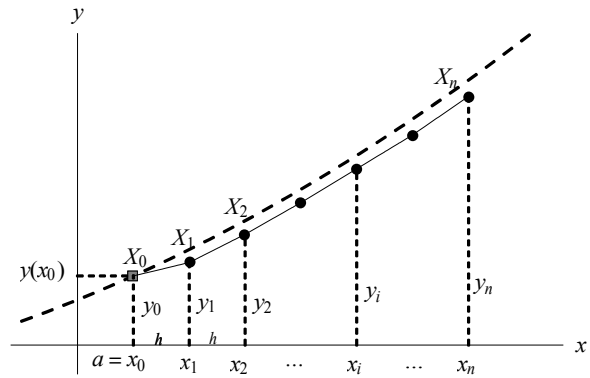
Hľadané partikulárne riešenie $y = \varphi(x)$ aproximujeme lomenou čiarou (Obr. 8.16)

$$X_0 X_1 X_2 \dots X_n = X_0 X_1 \cup X_1 X_2 \cup \dots \cup X_{n-1} X_n,$$

kde smernica k_i úsečky $X_i X_{i+1}$ je daná hodnotou $S(x_i, y_i, h)$.



Obr. 8.15



Obr. 8.16

Podme určiť smernicu $k_i = S(x_i, y_i, h)$ priamky $X_i X_{i+1}$. Integrujme DR $y' = f(x, y)$ na intervale $\langle x_i, x_{i+1} \rangle$. Dostávame:

$$\begin{aligned} y'(x) &= f(x, y(x)) \\ \int_{x_i}^{x_{i+1}} y'(x) dx &= \int_{x_i}^{x_{i+1}} f(x, y(x)) dx \\ y(x_{i+1}) - y(x_i) &= \int_{x_i}^{x_{i+1}} f(x, y(x)) dx \\ y(x_{i+1}) &= y(x_i) + \int_{x_i}^{x_{i+1}} f(x, y(x)) dx, \end{aligned}$$

kde x_i, x_{i+1} sú hranice intervalu (reálne čísla) a x je nezávisle premenná.

Potom na základe predpokladu $y(x_i) \approx y_i, y(x_{i+1}) \approx y_{i+1}$ dostávame východiskový vzťah na riešenie CU

$$y_{i+1} = y_i + \int_{x_i}^{x_{i+1}} f(x, y(x)) dx. \quad (8.5)$$

Z porovnania vzťahov (8.4) a (8.5) pre smerovú funkciu $S(x_i, y_i, h)$ platí:

$$S(x_i, y_i, h) = \frac{1}{h} \int_{x_i}^{x_{i+1}} f(x, y(x)) dx.$$

Smerová funkcia $S(x_i, y_i, h)$ závisí len od výpočtu $\int_{x_i}^{x_{i+1}} f(x, y(x)) dx$, pretože $1/h$ je konštanta:

Použitie jednokrokových alebo viackrokových riešenia DR závisí od spôsobu výpočtu daného integrálu:

- ak $\int_{x_i}^{x_{i+1}} f(x, y(x))$ integrujeme numericky, dostávame jednokrokové metódy riešenia DR,
- ak $f(x, y(x))$ aproximujeme a následne vypočítame presnú hodnotu určitého integrálu aproximačnej funkcie, dostaneme základ pre viackrokové metódy riešenia DR.

Integrál $\int_{x_i}^{x_{i+1}} f(x, y(x))$ budeme počítať numericky metódami uvedenými v kapitole 7. Podľa druhu použitej metódy dostaneme jednotlivé jednokrokové numerické metódy riešenia CU.

8.3.1 EULEROVA METÓDA

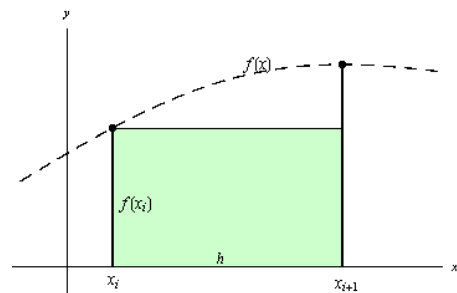
Na výpočet integrálu použijeme **obdĺžnikovú metódu (Typ 1)**:

$$\int_{x_i}^{x_{i+1}} f(x) dx \approx h \cdot f(x_i).$$

Potom

$$\int_{x_i}^{x_{i+1}} f(x, y(x)) dx \approx h \cdot f(x_i, y(x_i)) \wedge y(x_i) \approx y_i$$

$$\int_{x_i}^{x_{i+1}} f(x, y(x)) dx \approx h f(x_i, y_i), \quad (8.6)$$



Obr. 8.17

kde $f(x_i, y_i) = y'(x_i) = k_i$ je smernica priamky $X_i X_{i+1}$. Dosadením integrálu (8.6) do vzťahu (8.5) dostávame vzorec pre výpočet y_{i+1} prostredníctvom **Eulerovej** metódy:

$$y_{i+1} = h \cdot f(x_i, y_i) = y_i + h \cdot k. \quad (8.7)$$

Postup výpočtu hodnoty y_{i+1} z hodnôt x_i, y_i :

- vypočítame smernicu dotyčnice ku grafu presného riešenia v bode $X_i[x_i, y_i]$: $k = f(x_i, y_i)$,
- vypočítame y_{i+1} podľa vzťahu (8.7). Vzhľadom na jednoduchosť výpočtu, obidva kroky môžeme zlúčiť do jedného.

Algoritmus výpočtu súradníc x_i, y_i na intervale $\langle a, b \rangle$

Vstupné údaje: funkcia $f(x, y)$, x_0, y_0

Pre $i = 0, 1, \dots, n-1$ počítame:

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

$i = 0$	$i = 1$	$i = 2$
$x_1 = x_0 + h$	$x_2 = x_1 + h$	$x_3 = x_2 + h$
$k = f(x_0, y_0)$	$k = f(x_1, y_1)$	$k = f(x_2, y_2)$
$y_1 = y_0 + h \cdot k$	$y_2 = y_1 + h \cdot k$	$y_3 = y_2 + h \cdot k$
⇓	⇓	⇓
$[x_1, y_1]$	$[x_2, y_2]$	$[x_3, y_3] \dots$ atď.

Tabuľka 8.2. Proces výpočtu súradníc x_i, y_i

Schéma Eulerovej metódy pre $\langle x_i, x_{i+1} \rangle$	Geometrická interpretácia metódy
$x_{i+1} = x_i + h$ $k = f(x_i, y_i)$ $y_{i+1} = y_i + h \cdot k$ <p style="text-align: center;">⇓</p> $[x_{i+1}, y_{i+1}]$ $i = 0, 1, \dots, n - 1$	<p style="text-align: center;">Obr. 8.18</p>

Geometrická interpretácia procesu Eulerovej metódy

Z bodu $X_i[x_i, y_i]$ do bodu $X_{i+1}[x_{i+1}, y_{i+1}]$ sa dostaneme po priamke so smernicou $k = S(x_i, y_i, h) = f(x_i, y_i)$ (Obr. 8.18).

Ak spojíme body $X_i, X_{i+1}, i = 0, 1, \dots, n$ grafu numerického riešenia úsečkami, dostaneme tzv. Eulerov polygón $X_0X_1X_2 \dots X_n$ (Obr. 8.16).

Poznámka. K výsledku Eulerovej metódy môžeme prísť aj postupom uvedeným na začiatku tejto podkapitoly (pozri 8.1)

Eulerova metóda je veľmi nepresná metóda a na praktické použitie nevhodná. Používa sa len pre malý počet uzlov x_i a veľmi malý krok h .

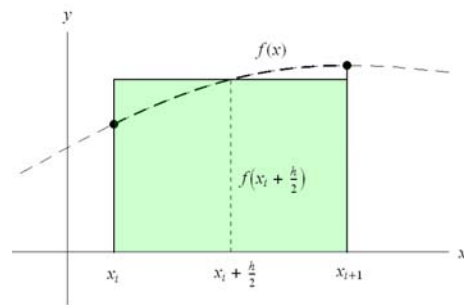
8.3.2 1. MODIFIKOVANÁ EULEROVA METÓDA

Na výpočet integrálu použijeme **obdĺžnikovú metódu (Typ 2)**:

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

Potom

$$\int_{x_i}^{x_{i+1}} f(x, y(x)) dx \approx h f\left(x_i + \frac{h}{2}, y\left(x_i + \frac{h}{2}\right)\right). \quad (8.8)$$



Obr. 8.19

Pretože presnú hodnotu $y(x_i + \frac{h}{2})$ partikulárneho riešenia v bode $(x_i + \frac{h}{2})$ nepoznáme, na jej výpočet použijeme numerickú, Eulerovu metódu, t. j. $y\left(x_i + \frac{h}{2}\right) \approx y_i + \frac{h}{2} f(x_i, y_i)$.

Po dosadení do (8.8) a následne do (8.5) bude

$$y_{i+1} = y_i + h f\left(x_i + \frac{h}{2}, y_i + \frac{h}{2} f(x_i, y_i)\right) \quad (8.9)$$

a po označení smerníc $f(x_i, y_i) = k_1$ a $f(x_i + \frac{h}{2}, y_i + \frac{h}{2} \cdot k_1) = k_2$ dostaneme

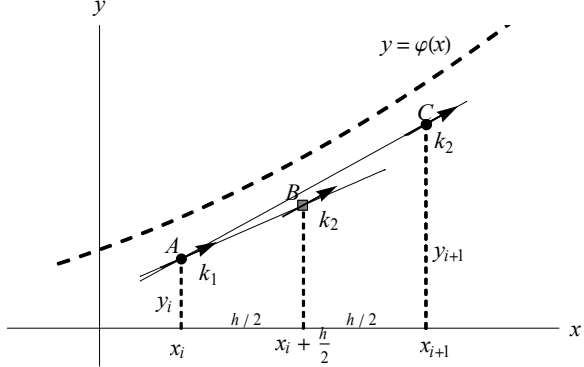
$$y_{i+1} = y_i + h \cdot k_2. \quad (8.10)$$

Postup výpočtu hodnoty y_{i+1} z hodnôt x_i, y_i :

- vypočítame smernicu k_1 dotyčnice ku grafu presného riešenia v bode $A = [x_i, y_i]$, kde $k_1 = f(x_i, y_i)$ a x_i je začiatkový bod intervalu $\langle x_i, x_{i+1} \rangle$,
- vypočítame smernicu k_2 dotyčnice ku grafu presného riešenia v bode $B[x_i + \frac{h}{2}, y_i + \frac{h}{2} k_1]$, kde $k_2 = f(x_i + \frac{h}{2}, y_i + \frac{h}{2} \cdot k_1)$ a $(x_i + \frac{h}{2})$ je bod v strede intervalu $\langle x_i, x_{i+1} \rangle$,
- vypočítame y_{i+1} podľa zjednodušeného vzťahu (8.10): $y_{i+1} = y_i + h \cdot k_2$.

Geometrická interpretácia procesu 1. modifikovanej Eulerovej metódy (Obr. 8.20):

- z bodu $A[x_i, y_i]$ do bodu $B[x_i + \frac{h}{2}, y_i + \frac{h}{2} k_1]$ sa dostaneme po priamke so smernicou k_1 ,
- z bodu $A[x_i, y_i]$ do bodu $C[x_{i+1}, y_{i+1}]$ sa dostaneme po priamke so smernicou $k_2 = S(x_i, y_i, h)$.

Schéma 1. modifikovanej Eulerovej metódy pre interval $\langle x_i, x_{i+1} \rangle$	Geometrická interpretácia metódy
$x_{i+1} = x_i + h$ $k_1 = f(x_i, y_i)$ $k_2 = f\left(x_i + \frac{h}{2}, y_i + \frac{h}{2} \cdot k_1\right)$ $y_{i+1} = y_i + h \cdot k_2$ \Downarrow $[x_{i+1}, y_{i+1}]$ $i = 0, 1, \dots, n-1$	 <p style="text-align: center;">Obr. 8.20</p>

Algoritmus výpočtu súradníc x_{i+1}, y_{i+1} na intervale $\langle a, b \rangle$

Vstupné údaje: funkcia $f(x, y)$, x_0, y_0

Pre $i = 0, 1, \dots, n-1$ počítame:

a) x_{i+1} : $x_{i+1} = x_i + h$

b) y_{i+1} : $k_1 = f(x_i, y_i)$

$$k_2 = f\left(x_i + \frac{h}{2}, y_i + \frac{h}{2} \cdot k_1\right)$$

$$y_{i+1} = y_i + h \cdot k_2$$

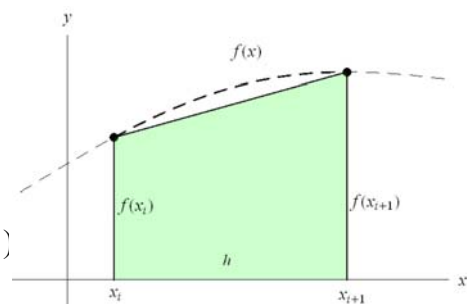
8.3.3 2. MODIFIKOVANÁ EULEROVA METÓDA (HEUNOVA METÓDA)

Na výpočet integrálu použijeme **lichobežníkovú metódu**:

$$\int_{x_i}^{x_{i+1}} f(x) dx \approx \frac{h}{2} \cdot (f(x_i) + f(x_{i+1}))$$

Potom

$$\int_{x_i}^{x_{i+1}} f(x, y(x)) dx \approx \frac{h}{2} (f(x_i, y(x_i)) + f(x_{i+1}, y(x_{i+1}))) \quad (8.11)$$



Obr. 8.21

Keďže presné hodnoty $y(x_i)$ a $y(x_{i+1}) = y(x_i + h)$

partikulárneho riešenia v uzloch x_i a $x_{i+1} = x_i + h$ nepoznáme, na výpočet ich približných hodnôt použijeme opäť Eulerovu metódu:

$$y(x_i) \approx y_i \Rightarrow f(x_i, y(x_i)) \approx f(x_i, y_i) \wedge f(x_i, y_i) = y'_i = k_1$$

$$y(x_i + h) \approx y_{i+1}, \text{ pričom } y_{i+1} = y_i + h \cdot f(x_i, y_i) = y_i + h \cdot k_1$$

a po dosadení do (8.11) a ďalej do (8.5) bude:

$$\int_{x_i}^{x_{i+1}} f(x, y(x)) dx \approx \frac{h}{2} \cdot (k_1 + f(x_i + h, y_i + h \cdot k_1))$$

$$\int_{x_i}^{x_{i+1}} f(x, y(x)) dx \approx \frac{h}{2} \cdot (k_1 + k_2)$$

$$y_{i+1} = y_i + h \cdot \frac{k_1 + k_2}{2}. \quad (8.12)$$

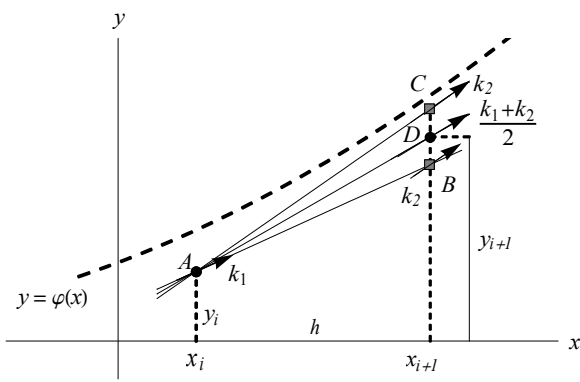
Postup výpočtu hodnoty y_{i+1} z hodnôt x_i, y_i :

- vypočítame smernicu k_1 dotyčnice ku grafu presného riešenia v bode $A[x_i, y_i]$, kde $k_1 = f(x_i, y_i)$ a x_i je začiatkový bod $\langle x_i, x_{i+1} \rangle$,
- vypočítame smernicu k_2 dotyčnice ku grafu presného riešenia v bode $B[x_i + h, y_i + hk_1]$, kde $k_2 = f(x_i + h, y_i + h \cdot k_1)$ a $x_i + h = x_{i+1}$ je koncový bod $\langle x_i, x_{i+1} \rangle$,
- vypočítame y_{i+1} podľa zjednodušeného vzťahu (8.12).

Geometrická interpretácia procesu 2. modifikovanej Eulerovej metódy (Obr. 8.22):

- z bodu $A[x_i, y_i]$ do bodu $B[x_i + h, y_i + hk_1]$ sa dostaneme po priamke so smernicou k_1 ,
- z bodu $A[x_i, y_i]$ do bodu $D[x_{i+1}, y_{i+1}]$ sa dostaneme po priamke so smernicou

$$\frac{k_1 + k_2}{2} = S(x_i, y_i, h).$$

Schéma 2. modifikovanej Eulerovej metódy pre $\langle x_i, x_{i+1} \rangle$	Geometrická interpretácia metódy
$x_{i+1} = x_i + h$ $k_1 = f(x_i, y_i)$ $k_2 = f(x_i + h, y_i + h \cdot k_1)$ $y_{i+1} = y_i + h \cdot \frac{(k_1 + k_2)}{2}$ <p style="text-align: center;">⇓</p> $[x_{i+1}, y_{i+1}]$ $i = 0, 1, \dots, n - 1$	 <p style="text-align: center;">Obr. 8.22</p>

Algoritmus výpočtu súradníc x_{i+1}, y_{i+1} na intervale $\langle a, b \rangle$

 Vstupné údaje: funkcia $f(x, y)$, x_0, y_0

 Pre $i = 0, 1, \dots, n-1$ počítame:

a) $x_{i+1}: x_{i+1} = x_i + h$

 b) y_{i+1} :

– $k_1 = f(x_i, y_i)$

– $k_2 = f(x_i + h, y_i + h \cdot k_1)$

– $y_{i+1} = y_i + h \cdot \frac{(k_1 + k_2)}{2}$

Poznámka.

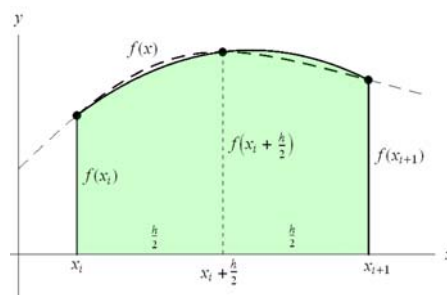
- Smerovú funkciu $S(x_i, y_i, h) = \frac{(1 \cdot k_1 + 1 \cdot k_2)}{2}$ počítame ako vážený priemer zo smerníc k_1 a k_2 , pričom každá smernica má váhu 1.
- Bod D je zároveň stred dvojice bodov A, C . Do bodu $C[x_i + h, y_i + hk_2]$ sa z bodu $A[x_i, y_i]$ dostaneme po priamke so smernicou k_2 .

8.3.4 METÓDA RUNGE – KUTTA 4. RÁDU

 Na výpočet integrálu použijeme **Simpsonovu metódu**:

$$\int_{x_i}^{x_{i+1}} f(x) dx \approx \frac{h}{3} \cdot \left(f(x_i) + 4 \cdot f\left(x_i + \frac{h}{2}\right) + f(x_{i+1}) \right)$$

V našom prípade polynóm $P_2(x)$, ktorého grafom je parabola, preložíme cez uzly $x_i, x_{i+1} = x_i + h$ a bod medzi nimi $\left(x_i + \frac{h}{2}\right) \in \langle x_i, x_{i+1} \rangle$, pričom krok bude $\frac{h}{2}$.



Obr. 8.23

Potom

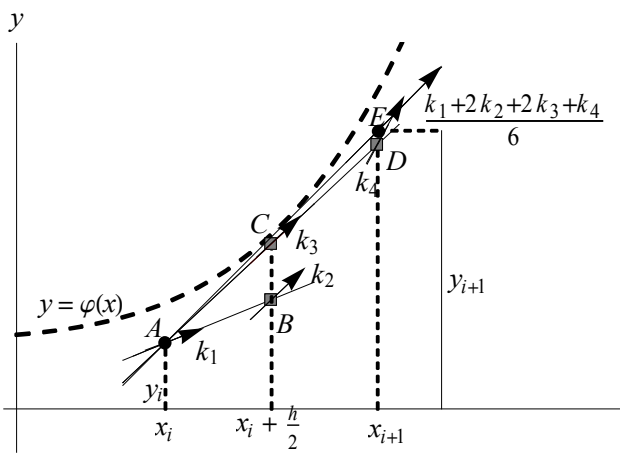
$$\int_{x_i}^{x_{i+1}} f(x, y(x)) dx \approx \frac{1}{3} \cdot \frac{h}{2} \cdot \left(f\left(x_i, y(x_i)\right) + 4 \cdot f\left(x_i + \frac{h}{2}, y\left(x_i + \frac{h}{2}\right)\right) + f\left(x_i + h, y(x_i + h)\right) \right). \quad (8.13)$$

Dosadením (8.13) do (8.5) bude

$$y_{i+1} = y_i + \frac{h}{6} \cdot \left(f(x_i, y(x_i)) + 2 \cdot f\left(x_i + \frac{h}{2}, y\left(x_i + \frac{h}{2}\right)\right) + 2 \cdot f\left(x_i + \frac{h}{2}, y\left(x_i + \frac{h}{2}\right)\right) + f(x_i + h, y(x_i + h)) \right).$$

Keďže presné hodnoty $y(x_i + \frac{h}{2})$ a $y(x_{i+1}) = y(x_i + h)$ partikulárneho riešenia v bodoch $\left(x_i + \frac{h}{2}\right)$ a $x_{i+1} = x_i + h$ nepoznáme, na výpočet ich približných hodnôt použijeme opäť Eule-

rovu metódu $y(x_i + \frac{h}{2}) \approx y_i + \frac{h}{2} \cdot f(x_i, y_i)$, $y(x_i + h) \approx y_i + h \cdot f(x_i, y_i)$. Ako výsledok dostaneme vzorce pre metódu Runge – Kutta 4. rádu:

Schéma metódy Runge – Kutta 4. rádu pre $\langle x_i, x_{i+1} \rangle$	Geometrická interpretácia metódy
$x_{i+1} = x_i + h$ $k_1 = f(x_i, y_i)$ $k_2 = f(x_i + \frac{h}{2}, y_i + \frac{h}{2} \cdot k_1)$ $k_3 = f(x_i + \frac{h}{2}, y_i + \frac{h}{2} \cdot k_2)$ $k_4 = f(x_i + h, y_i + h \cdot k_3)$ $y_{i+1} = y_i + h \cdot \frac{(k_1 + 2 \cdot k_2 + 2 \cdot k_3 + k_4)}{6}$ \Downarrow $[x_{i+1}, y_{i+1}]$ $i = 0, 1, \dots, n - 1$	 <p style="text-align: center;">Obr. 8.24</p>

Postup výpočtu hodnoty y_{i+1} z hodnôt x_i, y_i :

- vypočítame smernicu k_1 dotyčnice ku grafu presného riešenia v bode $A[x_i, y_i]$, kde $k_1 = f(x_i, y_i)$ a x_i je začiatkový bod $\langle x_i, x_{i+1} \rangle$ (pozri Obr. 8.24),
- vypočítame smernicu k_2 dotyčnice ku grafu presného riešenia v bode $B[x_i + \frac{h}{2}, y_i + \frac{h}{2} k_1]$, kde $k_2 = f(x_i + \frac{h}{2}, y_i + \frac{h}{2} \cdot k_1)$ a $(x_i + \frac{h}{2})$ je bod v strede $\langle x_i, x_{i+1} \rangle$,
- vypočítame smernicu k_3 dotyčnice ku grafu presného riešenia v bode $C[x_i + \frac{h}{2}, y_i + \frac{h}{2} k_2]$, kde $k_3 = f(x_i + \frac{h}{2}, y_i + \frac{h}{2} \cdot k_2)$ a $(x_i + \frac{h}{2})$ je bod v strede $\langle x_i, x_{i+1} \rangle$,
- vypočítame smernicu k_4 dotyčnice ku grafu presného riešenia v bode $D[x_i + h, y_i + h k_3]$, kde $k_4 = f(x_i + h, y_i + h \cdot k_3)$ a $x_i + h = x_{i+1}$ je koncový bod $\langle x_i, x_{i+1} \rangle$,
- vypočítame y_{i+1} podľa zjednodušeného vzťahu $y_{i+1} = y_i + h \cdot \frac{(k_1 + 2 \cdot k_2 + 2 \cdot k_3 + k_4)}{6}$.

Geometrická interpretácia procesu metódy Runge – Kutta 4. rádu (Obr. 8.24):

- z bodu $A[x_i, y_i]$ do bodu $B[x_i + \frac{h}{2}, y_i + \frac{h}{2}k_1]$ sa dostaneme po priamke so smernicou k_1 ,
- z bodu $A[x_i, y_i]$ do bodu $C[x_i + \frac{h}{2}, y_i + \frac{h}{2}k_2]$ sa dostaneme po priamke so smernicou k_2 ,
- z bodu $A[x_i, y_i]$ do bodu $D[x_i + h, y_i + hk_3]$ sa dostaneme po priamke so smernicou k_3 ,
- z bodu $A[x_i, y_i]$ do bodu $E[x_{i+1}, y_{i+1}]$ sa dostaneme po priamke so smernicou

$$\frac{(k_1 + 2 \cdot k_2 + 2 \cdot k_3 + k_4)}{6} = S(x_i, y_i, h).$$

Algoritmus výpočtu súradníc x_{i+1}, y_{i+1} na intervale $\langle a, b \rangle$

Vstupné údaje: funkcia $f(x, y)$, x_0, y_0

Pre $i = 0, 1, \dots, n-1$ počítame:

a) $x_{i+1}: x_{i+1} = x_i + h$

b) y_{i+1} :

– $k_1 = f(x_i, y_i)$

– $k_2 = f(x_i + \frac{h}{2}, y_i + \frac{h}{2} \cdot k_1)$

– $k_3 = f(x_i + \frac{h}{2}, y_i + \frac{h}{2} \cdot k_2)$

– $k_4 = f(x_i + h, y_i + h \cdot k_3)$

– $y_{i+1} = y_i + h \cdot \frac{(k_1 + 2 \cdot k_2 + 2 \cdot k_3 + k_4)}{6}$

Poznámka. Smerovú funkciu $S(x_i, y_i, h) = \frac{(1 \cdot k_1 + 2 \cdot k_2 + 2 \cdot k_3 + 1 \cdot k_4)}{6}$ vypočítame ako vážený priemer zo smerníc k_1, k_2, k_3, k_4 , pričom smernice k_1, k_4 majú váhu 1 a smernice k_2, k_3 majú váhu 2.

8.3.5 METÓDA RUNGE – KUTTA 3. RÁDU

Ešte uvedieme (bez odvodenia) poslednú často používanú metódu – metódu Runge – Kutta 3. rádu:

Postup výpočtu hodnoty y_{i+1} z hodnôt x_i, y_i :

- vypočítame smernicu k_1 dotyčnice ku grafu presného riešenia v bode $A[x_i, y_i]$, kde

$$k_1 = f(x_i, y_i) \text{ a } x_i \text{ je začiatkový bod } \langle x_i, x_{i+1} \rangle,$$

- vypočítame smernicu k_2 dotyčnice ku grafu presného riešenia v bode $B[x_i + \frac{h}{3}, y_i + \frac{h}{3}k_1]$,
kde $k_2 = f(x_i + \frac{h}{3}, y_i + \frac{h}{3} \cdot k_1)$ a $(x_i + \frac{h}{3})$ je bod ležiaci v jednej tretine $\langle x_i, x_{i+1} \rangle$,
- vypočítame smernicu k_3 dotyčnice ku grafu presného riešenia v bode
 $C[x_i + \frac{2h}{3}, y_i + \frac{2h}{3}k_2]$, kde $k_3 = f(x_i + \frac{2}{3} \cdot h, y_i + \frac{2}{3} \cdot h \cdot k_2)$ a $(x_i + \frac{2h}{3})$ je bod ležiaci
v dvoch tretinách $\langle x_i, x_{i+1} \rangle$,
- vypočítame y_{i+1} podľa vzťahu $y_{i+1} = y_i + h \cdot \frac{(k_1 + 3 \cdot k_3)}{4}$.

Geometrická interpretácia procesu metódy Runge – Kutta 3. rádu:

- z bodu $A[x_i, y_i]$ do bodu $B[x_i + \frac{h}{3}, y_i + \frac{h}{3}k_1]$ sa dostaneme po priamke so smernicou k_1 ,
- z bodu $A[x_i, y_i]$ do bodu $C[x_i + \frac{2h}{3}, y_i + \frac{2h}{3}k_2]$ sa dostaneme po priamke so smernicou
 k_2 ,
- z bodu $A[x_i, y_i]$ do bodu $D[x_{i+1}, y_{i+1}]$ sa dostaneme po priamke so smernicou
 $\frac{(k_1 + 3 \cdot k_3)}{4} = S(x_i, y_i, h)$.

Algoritmus výpočtu súradníc x_{i+1}, y_{i+1} na intervale $\langle a, b \rangle$

Vstupné údaje: funkcia $f(x, y)$, x_0, y_0

Pre $i = 0, 1, \dots, n-1$ počítame:

a) $x_{i+1}: x_{i+1} = x_i + h$

b) y_{i+1} :

– $k_1 = f(x_i, y_i)$

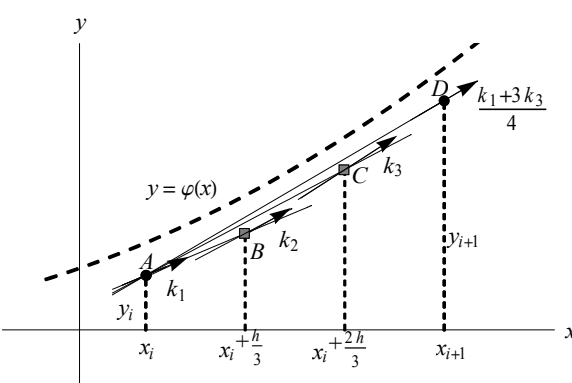
– $k_2 = f(x_i + \frac{h}{3}, y_i + \frac{h}{3} \cdot k_1)$

– $k_3 = f(x_i + \frac{2}{3} \cdot h, y_i + \frac{2}{3} \cdot h \cdot k_2)$

– $y_{i+1} = y_i + h \cdot \frac{(k_1 + 3 \cdot k_3)}{4}$

Poznámka. Smerovú funkciu $S(x_i, y_i, h) = \frac{(1 \cdot k_1 + 3 \cdot k_3)}{4}$ vypočítame ako vážený priemer zo smerníc

k_1, k_3 , pričom smernica k_1 má váhu 1 a smernica k_3 má váhu 3.

Schéma metódy Runge – Kutta 3. rádu pre $\langle x_i, x_{i+1} \rangle$	Geometrická interpretácia metódy
$x_{i+1} = x_i + h$ $k_1 = f(x_i, y_i)$ $k_2 = f\left(x_i + \frac{h}{3}, y_i + \frac{h}{3} \cdot k_1\right)$ $k_3 = f\left(x_i + \frac{2}{3} \cdot h, y_i + \frac{2}{3} \cdot h \cdot k_2\right)$ $y_{i+1} = y_i + h \cdot \frac{(k_1 + 3 \cdot k_3)}{4}$ <p style="text-align: center;">⇓</p> $[x_{i+1}, y_{i+1}]$ $i = 0, 1, \dots, n-1$	 <p style="text-align: center;">Obr. 8.25</p>

8.3.6 PRESNOŠŤ JEDNOKROKOVÝCH METÓD

Metódy typu Runge – Kutta sú metódy vždy konvergentné a numericky stabilné. Z uvedených metód najpresnejšia je metóda Runge – Kutta 4. rádu a najmenej presná je Eulerova metóda. Väčšiu presnosť približného riešenia DR niektorou z metód dosiahneme zmenšením kroku h numerickej metódy. Praktický význam majú vzorce numerickej metódy až do rádu $p = 4$. Pre rád numerickej metódy $p > 4$ sa vzorce používajú zriedka, aj keď sú teoreticky presnejšie. Podobne ako pri integráloch, vkladáním ďalších bodov medzi uzly – počítáním ďalších smerníc sa presnosť metódy zvyšuje, avšak nárastom počtu operácií sa zvyšuje náročnosť výpočtov ako aj rastie vplyv šírenia zaokrúhľovacích chýb.

Rád metódy	Metóda typu Runge-Kutta	Chyba metódy (rádu)		
		lokálna	globálna	$E(h)$
$p = 1$	Eulerova	$O(h^2)$	$O(h^1)$	$O(h^1)$
$p = 2$	1.+2. modif. Eulerova	$O(h^3)$	$O(h^2)$	$O(h^2)$
$p = 3$	Runge – Kutta 3. rádu	$O(h^4)$	$O(h^3)$	$O(h^3)$
$p = 4$	Runge – Kutta 4. rádu	$O(h^5)$	$O(h^4)$	$O(h^4)$

Tabuľka 8.3

Presnú hodnotu lokálnej a globálnej chyby vyčíslujeme iba v prípade, keď poznáme presné riešenie DR. Inak na ich určenie používame odhad. Na stanovenie hrubého odhadu chyby metódy stačí rád metódy p , ktorý nesie informáciu o rýchlosti konvergenencie, t. j. o vplyve kroku na rád chyby.

Pre lokálnu chybu platí $d_i = O(h^{p+1})$, t. j. existuje také kladné reálne číslo c a reálne číslo x_0 , že pre všetky $x > x_0$, $x \in \langle x_0, b \rangle$: $d_i \leq c \cdot h^{p+1}$, a teda $d_i \sim h^{p+1}$.

Pre globálnu chybu platí $e_i = O(h^p)$, t. j. existuje také kladné reálne číslo m a reálne číslo x_0 , že pre všetky $x > x_0$, $x \in \langle x_0; b \rangle$: $e_i \leq m \cdot h^{p+1}$, a teda $e_i \sim h^p$.

Za predpokladu, že najväčšia globálna chyba je v poslednom uzle x_n , (platí pre jednokrokové metódy) chybu metódy stanovíme ako globálnu chybu v poslednom uzle x_n : $e_n = O(h^p)$.

Z uvedeného vyplýva, že ak pri hľadaní približného riešenia na intervale $\langle a, b \rangle$ požadujeme iba „rádovo“ zaručiť presnosť metódy ε , t. j. $e_n \leq O(\varepsilon)$, stačí vykonať výber vhodnej numerickej metódy a voľbu vhodného kroku h numerickej metódy. Ako? Ukážeme na príklade.

Nech je pri výpočte riešenia DR numericou metódou požadovaná presnosť $\varepsilon = 10^{-3}$. Znamená to, že potrebujeme zabezpečiť, aby výpočet približného riešenia DR prebehol tak, že $e_n \leq O(\varepsilon)$, $e_n = |y_i - y(x_i)| = O(h^p)$, t. j. $O(h^p) \leq O(\varepsilon = 10^{-3})$, $i = 1, 2, \dots, n$.

Predpokladajme, že chceme výpočet realizovať **metódou 2. rádu**, ktorej globálna chyba je rádu $O(h^2)$.

Keď zvolíme krok:

- $h = 0,2 = 2 \cdot 10^{-1}$, potom $e_n = O(h^p) = O((10^{-1})^2) = O(10^{-2})$, $O(10^{-2}) > O(10^{-3}) \Rightarrow$ krok $h = 0,2$ nevyhovuje pre požadovanú presnosť $O(\varepsilon = 10^{-3})$
- $h = 0,02 = 2 \cdot 10^{-2}$, potom $e_n = O((10^{-2})^2) = O(10^{-4})$, $O(10^{-4}) < O(10^{-3}) \Rightarrow$ krok $h = 0,2$ vyhovuje pre požadovanú presnosť $O(\varepsilon = 10^{-3})$.

Keby sme chceli ten istý výpočet realizovať **metódou 4. rádu**, ktorej globálna chyba je rádu $O(h^4)$, a opäť zvolíme najskôr krok $h = 0,2 = 2 \cdot 10^{-1}$, potom $e_n = O((10^{-1})^4) = O(10^{-4})$, $O(10^{-4}) < O(10^{-3})$ a teda už krok $h = 0,2$ vyhovuje pre požadovanú presnosť $O(\varepsilon = 10^{-3})$.

Pre jednokrokové metódy odhad globálnej chyby, resp. odhad metódy je možné pomerne ľahko určiť metódou polovičného kroku.

Veta 8.3 Metóda polovičného kroku. Nech h_1 a h_2 sú dva rôzne kroky numerickej metódy, pričom $h_1 = q \cdot h_2$, $q > 1$. Nech $y_j^{h_1}$ a $y_j^{h_2}$, sú hodnoty numerických riešení v uzle $x_{j=J}$ pre kroky h_1 a h_2 , kde $J = q \cdot j$. Označme $y_j^{h_1} = y^1$ a $y_j^{h_2} = y^2$. Potom pre odhad globálnej chyby v bode x_j a krok h_2 platí:

$$e_j(h_2) = \frac{|y_j^{h_2} - y_j^{h_1}|}{q^p - 1}, \quad (8.14)$$

v zjednodušenom zápise:

$$e_j = \frac{|y^2 - y^1|}{q^p - 1}, \quad (8.15)$$

pričom $O(h^p)$ je rád numerickej metódy.

Ak $h_2 = h_1 / 2$, t. j. $h_1 = 2h_2$, potom $e_J(h_2) = \frac{y_J^{h_2} - y_J^{h_1}}{2^p - 1}$, resp. $e_J = \frac{|y^2 - y^1|}{2^p - 1}$, pričom $O(h^p)$ je rád numerickej metódy.

8.4 RIEŠENÉ PRÍKLADY

PRÍKLAD č. 1

Vyšetrite existenciu a jednoznačnosť riešenia nasledujúcich úloh

- a) $x \cdot y' - e^{\frac{x}{y}} = 0, y(1) = 0$
- b) $y' = \sqrt{xy}, y(1) = 0$
- c) $y' = \ln xy, y(1) = 1$
- d) $\sqrt{x+2} \cdot y' = y, y(1) = 2$ alebo $y(-3) = 0$

RIEŠENIE

a) Rovnicu upravíme na tvar $y' = f(x, y): y' = \frac{1}{x} e^{\frac{x}{y}} \Rightarrow f(x, y) = \frac{1}{x} e^{\frac{x}{y}}$.

Definičný obor tejto funkcie je $D(f(x, y)) = \{(x, y) \in E_2; x \neq 0; y \neq 0\} = R - \{0\} \times R - \{0\}$.

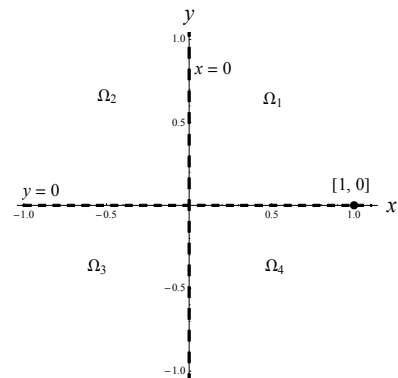
Funkcia $f(x, y)$ je definovaná a spojitá na oblastiach:

$$\Omega_1 = (0, \infty) \times (0, \infty)$$

$$\Omega_2 = (-\infty, 0) \times (0, \infty)$$

$$\Omega_3 = (-\infty, 0) \times (-\infty, 0)$$

$$\Omega_4 = (0, \infty) \times (-\infty, 0)$$



Obr. 8.26

Začiatočná podmienka hovorí, že riešenie má vychádzať z bodu $[1; 0]$, ktorý nepatrí do $D(f(x, y))$.

Keďže funkcia v ňom nie je definovaná, nemôže byť v ňom ani spojitá. Nemáme zaručenú existenciu riešenia tejto úlohy.

b) V tomto prípade $f(x, y) = \sqrt{xy}$, ktorej

$$D(f(x, y)) = \{(x, y) \in E_2; x \cdot y \geq 0\} = ((-\infty, 0) \times (-\infty, 0)) \cup ((0, \infty) \times (0, \infty))$$

Funkcia $f(x, y)$ je definovaná a spojitá na oblastiach

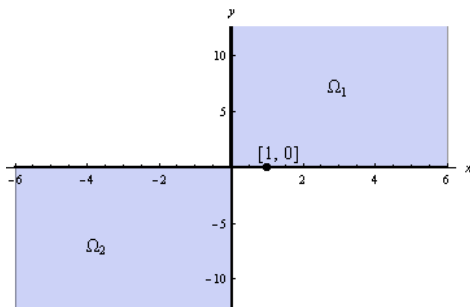
$$\Omega_1 = (-\infty, 0) \times (-\infty, 0),$$

$$\Omega_2 = (0, \infty) \times (0, \infty).$$

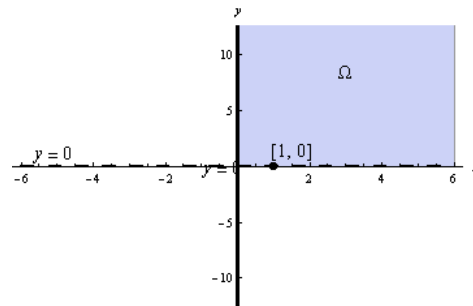
Bod $[1; 0]$ leží v oblasti Ω_2 , čo zaručuje existenciu riešenia úlohy.

$\frac{\partial f(x,y)}{\partial y} = \frac{\sqrt{x}}{2\sqrt{y}}$ je definovaná $D(\frac{\partial f(x,y)}{\partial y}) = \{[x,y] \in E_2: x \geq 0, y > 0\} = \langle 0, \infty \rangle \times (0, \infty)$

a spojitá na oblasti $\Omega = \langle 0, \infty \rangle \times (0, \infty)$. Pretože v bode $[1; 0]$ nie je definovaná, nemôže tam byť spojitá. Nemáme zaručenú jednoznačnosť riešenia úlohy (dá sa ukázať, že úloha má dve riešenia).



Obr. 8.27. $D(f)$ a spojitost' $f(x,y)$



Obr. 8.28. $D(f)$ a spojitost' $\frac{\partial f(x,y)}{\partial y}$

c) V treťom prípade $f(x,y) = \ln(xy)$ a bod $[x_0, y_0]$ má súradnice $[1; 1]$.

Definičný obor tejto funkcie je

$$D(f(x,y)) = \{[x,y] \in E_2; xy > 0\} = ((-\infty, 0) \times (-\infty, 0)) \cup ((0, \infty) \times (0, \infty))$$

a funkcia $f(x,y)$ je definovaná a spojitá na oblastiach

$$\Omega_1 = (-\infty, 0) \times (-\infty, 0),$$

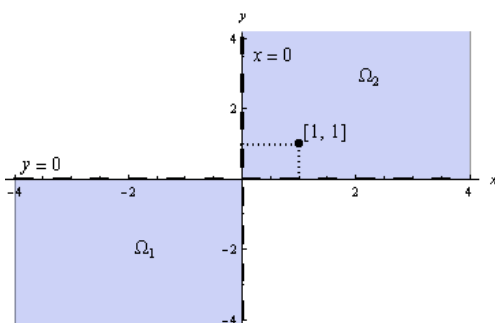
$$\Omega_2 = (0, \infty) \times (0, \infty).$$

Bod $[1; 1]$ leží v oblasti Ω_2 , čo zaručuje existenciu riešenia úlohy.

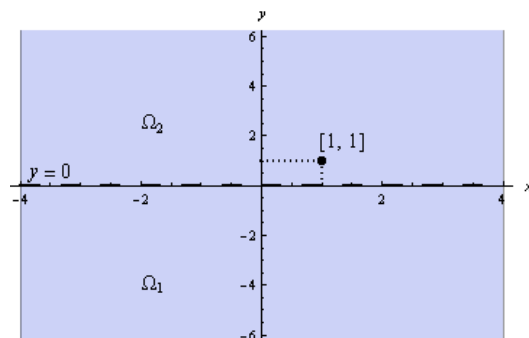
$$\frac{\partial f(x,y)}{\partial y} = \frac{1}{x \cdot y} \cdot x = \frac{1}{y}. \text{ Jej } D(\frac{\partial f(x,y)}{\partial y}) = \{[x,y] \in E_2; : x \in \mathbb{R}, y \neq 0\} = \mathbb{R} \times \mathbb{R} - \{0\}.$$

$\frac{\partial f(x,y)}{\partial y}$ je definovaná a spojitá na oblastiach $\Omega_1 = \mathbb{R} \times (-\infty, 0)$, $\Omega_2 = \mathbb{R} \times (0, \infty)$. Bod $[1; 1]$

leží v oblasti Ω_2 , čo zaručuje jednoznačnosť riešenia úlohy.



Obr. 8.29. $D(f)$ a spojitost' $f(x,y)$



Obr. 8.30. $D(f)$ a spojitost' $\frac{\partial f(x,y)}{\partial y}$

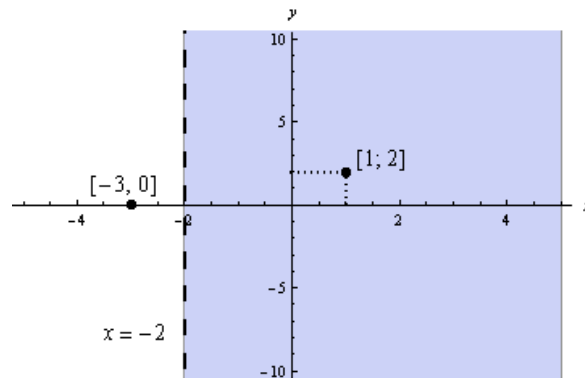
d) Funkcia $f(x, y) = \frac{y}{\sqrt{x+2}}$ a jej $D(f(x, y)) = \{[x, y] \in E_2; x+2 > 0, y \in R\} = (-2, \infty) \times R$,

t. j. $f(x, y)$ je definovaná a spojitá na oblasti $\Omega = (-2, \infty) \times R$. Začiatkový bod $[1; 2]$ leží v oblasti Ω , čo zaručuje existenciu riešenia úlohy:

$$\frac{\partial f(x, y)}{\partial y} = \frac{1}{\sqrt{x+2}}, \text{ jej } D\left(\frac{\partial f(x, y)}{\partial y}\right) = \{[x, y] \in E_2; x+2 > 0, y \in R\} = (-2, \infty) \times R = \Omega.$$

V bode $[1; 2]$ je definovaná a spojitá, čo zaručuje jednoznačnosť riešenia úlohy.

Bod $[x_0, y_0] = [-3; 0]$ neleží v oblasti $\Omega = (-2, \infty) \times R$. Keďže funkcia $f(x, y)$ nie je v ňom definovaná, nemôže tam byť ani spojitá. Nemáme zaručenú existenciu riešenia tejto úlohy.



Obr. 8.31. $D(f)$ a spojitosť $f(x, y)$ a $\frac{\partial f(x, y)}{\partial y}$

V nasledujúcej úlohe uvádzame riešenie diferenciálnej rovnice, ktorej partikulárne riešenie poznáme. Predpokladáme, že čitateľ takto lepšie pochopí význam numerických metód riešenia tých diferenciálnych rovníc, ktorých partikulárne riešenie nevieme vypočítať.

PRÍKLAD č. 2

Je daná diferenciálna rovnica $(x+1) \cdot y' - y + 2x = 0$, $y(0) = 1$, $x \in \langle 0, 1 \rangle$, ktorej partikulárne riešenie je $y(x) = 1 + 3x - 2 \ln(1+x) - 2x \ln(1+x)$.

- Vyšetrite existenciu a jednoznačnosť jej riešenia.
- Nájdite jej približné riešenie s krokom $h = 0,1$ Eulerovou metódou a zobrazte ho.
- Nájdite jej približné riešenie s krokom $h = 0,1$ metódou Runge – Kutta 3. rádu a zobrazte ho.
- Ak viete, že presné všeobecné riešenie DR je $y(x, c) = c(1+x) - 2(1+x)\left(\frac{1}{1+x} + \ln(1+x)\right)$, vypočítajte presné partikulárne riešenie. Obe riešenia zobrazte.
- Pre každú z metód vypočítajte chybu metódy. Porovnajte presné partikulárne a obe numerické riešenia graficky a výpočtom. Ktorá z metód je presnejšia?

RIEŠENIE

Ide o Cauchyho úlohu, obyčajnú diferenciálnu rovnicu so začiatočnou podmienkou.

Pristúpme k riešeniu našej úlohy $(x + 1) \cdot y' - y + 2x = 0, y(0) = 1, x \in \langle 0, 1 \rangle$.

Najskôr rovnicu upravíme na tvar $y' = f(x, y)$:

$$(x + 1)y' - y + 2x = 0 \Rightarrow y' = \frac{y - 2x}{x + 1} \Rightarrow f(x, y) = \frac{y - 2x}{x + 1}.$$

a) *Vyšetrte existenciu a jednoznačnosť jej riešenia.*

- *Existencia riešenia:*

$$D(f(x, y)) = \{[x, y] \subset E_2; x + 1 \neq 0; y \in R\} = R - \{-1\} \times R.$$

Funkcia $f(x, y)$ je definovaná a spojitá na oblastiach $\Omega_1 = (-\infty, -1) \times R$ a $\Omega_2 = (-1, \infty) \times R$. Vzhľadom na začiatočnú podmienku $[x_0, y_0] = [0, 1] \in \Omega_2$, partikulárne riešenie $y(x)$ bude ležať v oblasti Ω_2 . Partikulárne riešenie budeme hľadať na intervale $x \in \langle 0, 1 \rangle \subset (-1, \infty)$ a v oblasti $\langle 0, 1 \rangle \times R \subset \Omega_2$, kde je $f(x, y)$ spojitá \Rightarrow v okolí bodu $[x_0, y_0] = [0, 1]$ je existencia riešenia úlohy zaručená. (Obr. 8.32)

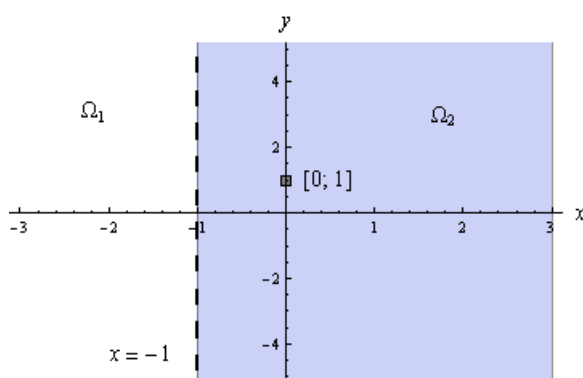
- *Jednoznačnosť riešenia:*

$$\frac{\partial f(x, y)}{\partial y} = \frac{1}{x + 1}, D\left(\frac{\partial f(x, y)}{\partial y}\right) = \{[x, y] \subset E_2; x + 1 \neq 0; y \in R\} = R - \{-1\} \times R.$$

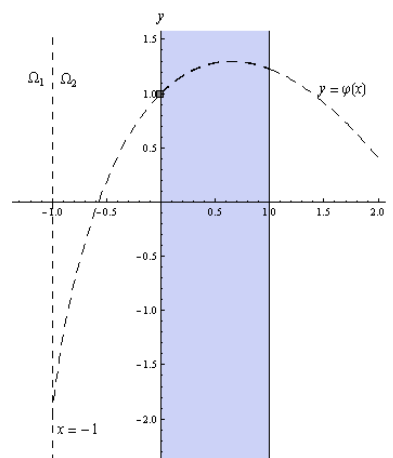
Pretože $\frac{\partial f(x, y)}{\partial y}$ je definovaná a spojitá na oblastiach $\Omega_1 = (-\infty, -1) \times R, \Omega_2 = (-1, \infty) \times R$, je

spojitá aj na oblasti $\langle 0, 1 \rangle \times R \subset \Omega_2$ a aj v bode $[0, 1] \in \Omega_2$, čo v okolí bodu $[x_0, y_0] = [0, 1]$ zaručuje jednoznačnosť riešenia úlohy (Obr. 8.32).

Záver. Bodom $[0, 1]$ je dané jediné riešenie Cauchyho úlohy (Obr. 8.33).



Obr. 8.32



Obr. 8.33

b) Eulerovou metódou nájdite približné riešenie diferenciálnej rovnice na intervale $\langle 0, 1 \rangle$ s krokom $h = 0,1$ a zobrazte ho.

Výpočet približného riešenia DR budeme realizovať na základe vzorcov pre danú metódu.

- Počet iterácií, ktoré potrebujeme vypočítať, ako aj počet čiastkových intervalov, na ktoré

bude rozdelený interval $\langle 0;1 \rangle$, je $n = \frac{b-a}{h} = \frac{1-0}{0,1} = 10$.

- Počet uzlov je $n + 1 = 11$.

Eulerova metóda:

$$\begin{aligned} x_{i+1} &= x_i + h \\ k &= f(x_i, y_i) \Rightarrow [x_{i+1}, y_{i+1}] \\ y_{i+1} &= y_i + h \cdot k \end{aligned} \quad i=0,1,\dots,n-1$$

Vstupné údaje: $f(x, y) = \frac{y-2x}{x+1}$, $x_0 = 0, y_0 = 0, n = 10$

Postup:

$$[x_0, y_0] = [0, 1] \Rightarrow$$

$$i = 0: \quad x_1 = x_0 + h = 0 + 0,1 = 0,1$$

$$k = f(x_0, y_0) = \frac{y_0 - 2x_0}{x_0 + 1} = f(0, 1) = \frac{1 - 2 \cdot 0}{0 + 1} = 1$$

$$y_1 = y_0 + h \cdot k = 1 + 0,1 \cdot 1 = 1,1$$

$$[x_1, y_1] = [0,1 ; 1,1]$$

$$i = 1: \quad x_2 = x_1 + h = 0,1 + 0,1 = 0,2$$

$$k = f(x_1, y_1) = \frac{y_1 - 2x_1}{x_1 + 1} = f(0,1; 1,1) = \frac{1,1 - 2 \cdot 0,1}{0,1 + 1} = 0,818$$

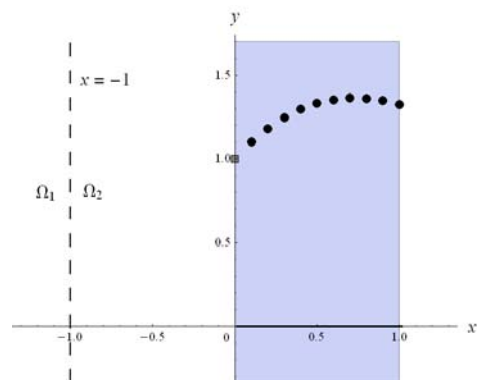
$$y_2 = y_1 + h \cdot k = 1,1 + 0,1 \cdot 0,8182 = 1,1818$$

$$[x_2, y_2] = [0,2 ; 1,1818]$$

Analogicky vykonáme výpočet aj pre ostatné uzly intervalu $\langle 0, 1 \rangle$.

i	x_i	y_i
0	0	1
1	0.1	1.1
2	0.2	1.18182
3	0.3	1.24697
4	0.4	1.29674
5	0.5	1.33222
6	0.6	1.35437
7	0.7	1.36401
8	0.8	1.3619
9	0.9	1.34867
10	1.	1.32491

Tabuľka 8.4. Numerické riešenie CU



Obr. 8.34. Graf numerického riešenia

Záver. Numerickým riešením danej diferenciálnej rovnice Eulerovou metódou na intervale $\langle 0,1 \rangle$ s krokom $h = 0,1$ je funkcia daná tabuľkou (Tabuľka 8.4). Graf numerického riešenia je na Obr. 8.34.

c) *Nájdite približné riešenie diferenciálnej rovnice na intervale $\langle 0, 1 \rangle$ s krokom $h = 0,1$ metódou Runge – Kutta 3. rádu a zobrazte ho.*

Výpočet približného riešenia DR realizujeme na základe vzorcov pre príslušnú metódu.

- Počet čiastkových intervalov (= počet počítaných iterácií): $n = \frac{b-a}{h} = \frac{1-0}{0,1} = 10$.
- Počet uzlov je $n + 1 = 11$.

Metóda Runge – Kutta 3. rádu:

$$\begin{aligned}
 x_{i+1} &= x_i + h \\
 k_1 &= f(x_i, y_i) \\
 k_2 &= f\left(x_i + \frac{h}{3}, y_i + \frac{h}{3} \cdot k_1\right) \\
 k_3 &= f\left(x_i + \frac{2}{3} \cdot h, y_i + \frac{2}{3} \cdot h \cdot k_2\right) \Rightarrow [x_{i+1}, y_{i+1}] \\
 & \qquad \qquad \qquad i = 0, 1, \dots, n-1 \\
 y_{i+1} &= y_i + h \cdot \frac{(k_1 + 3 \cdot k_3)}{4}
 \end{aligned}$$

Vstupné údaje: $f(x, y) = \frac{y-2x}{x+1}$, $x_0 = 0, y_0 = 0, n = 10$

Postup:

$$[x_0, y_0] = [0, 1] \Rightarrow$$

$$i = 0: \quad x_1 = x_0 + h = 0 + 0,1 = 0,1$$

$$k_1 = f(x_0, y_0) = \frac{y_0 - 2x_0}{x_0 + 1} = f(0, 1) = \frac{1 - 2 \cdot 0}{0 + 1} = 1$$

$$k_2 = f\left(x_0 + \frac{h}{3}, y_0 + \frac{h}{3} k_1\right) = f(0,0333; 1,0333) = \frac{1,0333 - 2 \cdot 0,0333}{0,0333 + 1} = 0,93548$$

$$k_3 = f\left(x_0 + \frac{2h}{3}, y_0 + \frac{2h}{3} k_2\right) = f(0,0666; 1,06236) = \frac{1,06236 - 2 \cdot 0,0666}{0,0666 + 1} = 0,87097$$

$$y_1 = y_0 + h \cdot \frac{(k_1 + 3 \cdot k_3)}{4} = 1 + 0,1 \cdot \frac{1 + 3 \cdot 0,87097}{4} = 1,0903$$

$$[x_1, y_1] = [0,1; 1,0903];$$

$$i = 1: \quad x_2 = x_1 + h = 0,1 + 0,1 = 0,2$$

$$k_1 = f(x_1, y_1) = f(0,1) = \frac{1,0903 - 2 \cdot 0,1}{0,1 + 1} = 0,80938$$

$$k_2 = f\left(x_1 + \frac{h}{3}, y_1 + \frac{h}{3}k_1\right) = f(0,1333; 1,1173) = \frac{1,1173 - 2 \cdot 0,1333}{0,1333 + 1} = 0,75056$$

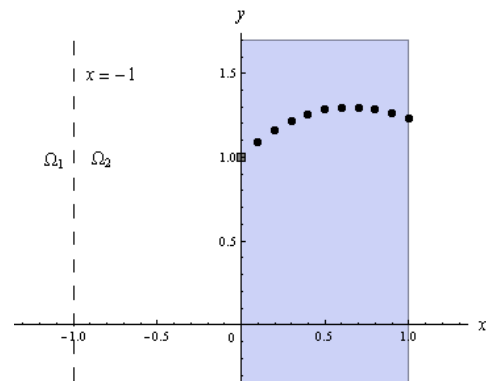
$$k_3 = f\left(x_1 + \frac{2h}{3}, y_1 + \frac{2h}{3}k_2\right) = f(0,1667; 1,1404) = \frac{1,1404 - 2 \cdot 0,1667}{0,1667 + 1} = 0,69174$$

$$y_2 = y_1 + h \cdot \frac{(k_1 + 3 \cdot k_3)}{4} = 1,0903 + 0,1 \cdot \frac{0,80938 + 3 \cdot 0,69174}{4} = 1,1624$$

$$[x_2, y_2] = [0,2; 1,1624].$$

Analogicky vykonáme výpočet aj pre ostatné uzly intervalu $\langle 0, 1 \rangle$.

i	x_i	y_i
0	0	1
1	0.1	1.09032
2	0.2	1.16244
3	0.3	1.21787
4	0.4	1.25789
5	0.5	1.28362
6	0.6	1.29601
7	0.7	1.29589
8	0.8	1.284
9	0.9	1.26098
10	1.	1.22744



Tabuľka 8.5. Numerické riešenie CU

Obr. 8.35. Graf numerického riešenia

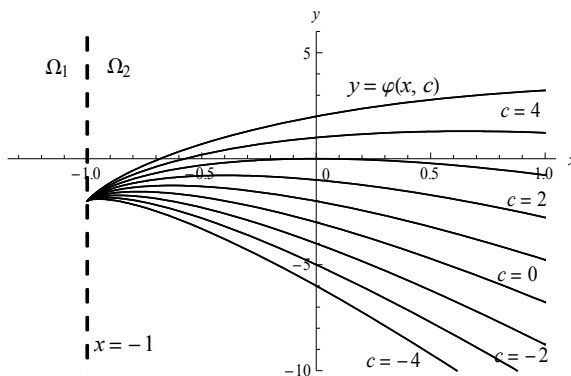
Záver. Numerickým riešením danej diferenciálnej rovnice metódou Runge – Kutta 3. rádu na intervale $\langle 0,1 \rangle$ s krokom $h = 0,1$ je funkcia daná tabuľkou (Tabuľka 8.5). Graf numerického riešenia je na Obr. 8.35.

d) Ak viete, že presné všeobecné riešenie DR je $y(x, c) = c(1+x) - 2(1+x)\left(\frac{1}{1+x} + \ln(1+x)\right)$, vypočítajte presné partikulárne riešenie. Obe riešenia zobrazte.

Všeobecným riešením je množina funkcií, ktoré sa navzájom líšia o konštantu $c \in \mathbb{R}$

$$y(x, c) = c(1+x) - 2(1+x)\left(\frac{1}{1+x} + \ln(1+x)\right),$$

pričom toto riešenie je definované pre $x \in (-1, \infty)$.



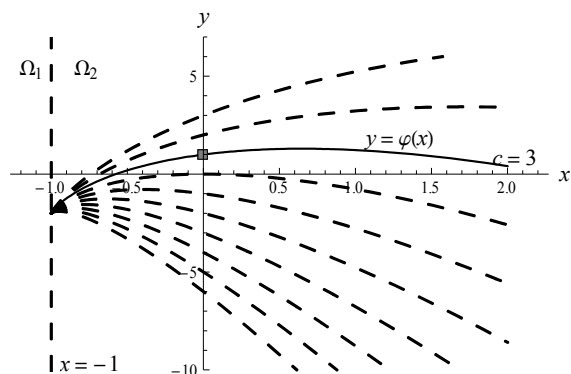
Obr. 8.36. Všeobecné riešenie DR $(x + 1) \cdot y' - y + 2x = 0$

Partikulárne riešenie dostaneme, ak do všeobecného riešenia $y = y(x, c)$ dosadíme začiatočnú podmienku $y(0) = 1$ a vypočítame konštantu c . Znamená to, že z množiny všetkých funkcií všeobecného riešenia vyberieme práve tú jednu funkciu, ktorej graf prechádza bodom $[x_0, y_0] = [0, 1]$

$$1 = c(1 + 0) - 2(1 + 0) \left(\frac{1}{1 + 0} + \ln(1 + 0) \right) \Rightarrow c = 3$$

Partikulárnym riešením Cauchyho úlohy je funkcia $y(x)$ daná predpisom

$$y(x) = 1 + 3x - 2 \ln(1 + x) - 2x \ln(1 + x)$$



Obr. 8.37. Všeobecné a partikulárne riešenie DR $(x + 1) \cdot y' - y + 2x = 0$

e) Pre každú z metód vypočítajte chybu metódy. Porovnajte presné partikulárne a obe numerické riešenia graficky a výpočtom. Ktorá z metód je presnejšia?

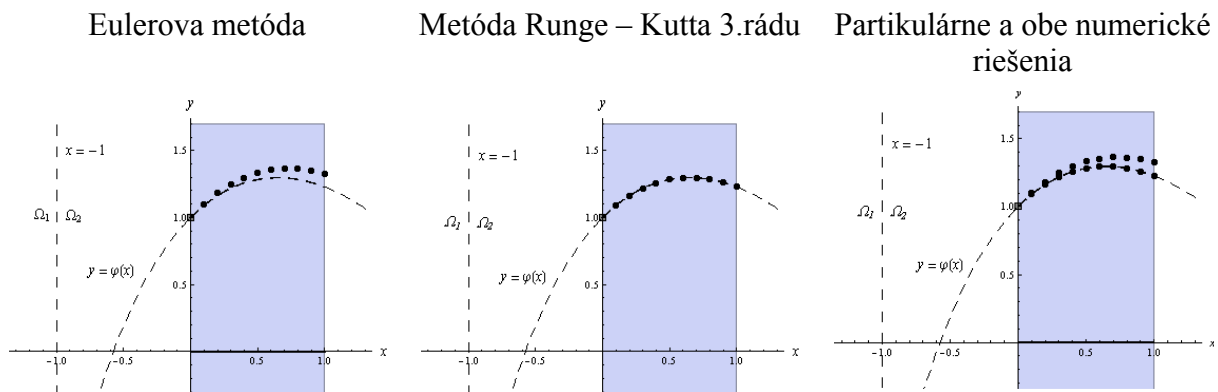
Presnejšia je tá z metód, ktorá má menšiu chybu metódy, t. j. menšiu maximálnu globálnu chybu. Globálnu chybu v každom uzle určíme podľa vzťahu $e_i = |y(x_i) - y_i|$, kde $y(x_i)$ je hodnota presného partikulárneho riešenia $y(x) = 1 + 3x - 2 \ln(1 + x) - 2x \ln(1 + x)$ v bode x_i . Vypočítané hodnoty globálnej chyby v jednotlivých uzloch sú uvedené v tabuľke Tabuľka 8.6. Maximálnu globálnu chybu dosahujú obe metódy v poslednom uzle $x_{10} = 1$.

Globálna chyba			
i	x_i	$ yE_i - y_P(x_i) $	$ yRK3_i - y_P(x_i) $
0	0	0	0
1	0.1	0.0096824	4.97621×10^{-6}
2	0.2	0.0193899	9.20396×10^{-6}
3	0.3	0.0291168	0.0000129028
4	0.4	0.0388589	0.0000162174
5	0.5	0.0486131	0.000019246
6	0.6	0.0583772	0.0000220576
7	0.7	0.0681495	0.0000247014
8	0.8	0.0779286	0.0000272135
9	0.9	0.0877134	0.0000296207
10	1.	0.0975031	0.0000319434

Tabuľka 8.6 (Použité označenie: $y_i = yE_i$ pre Eulerovu metódu, $y_i = yRK3_i$ pre metódu Runge Kutta 3. rádu.)

Chyba Eulerovej metódy je $e_n = 0,0975031 = 9,75031 \cdot 10^{-2}$, chyba metódy Runge-Kutta 3. rádu je $e_n = 0,0000319434 = 3,19434 \cdot 10^{-5}$.

Grafické porovnanie presného partikulárneho riešenia a numerických metód:



Obr. 8.38. Grafy numerických riešení

Záver. Na nájdenie približného riešenia danej DR je vhodnejšia metóda Runge – Kutta 3. rádu, pretože má menšiu chybu metódy, menšiu maximálnu globálnu chybu:

$$3,19434 \cdot 10^{-5} < 9,75031 \cdot 10^{-2}, \text{ t.j. } 9,75031 \cdot 10^{-2} \doteq 3\,052 \cdot 3,19434 \cdot 10^{-5}.$$

PRIKLAD č.3

Je daná Cauchyho úloha $y^2 y' - x + y^2 = 0, y(1) = 1$.

- Vyšetríte existenciu a jednoznačnosť jej riešenia.
- V bode $x = 1,2$ nájdite hodnotu numerického riešenia Heunovou metódou. Na výpočet použite kroky $h_1 = 0,1$ a $h_2 = 0,05$. Metódou polovičného kroku vypočítajte odhad globálnej chyby metódy v danom bode pre krok h_2 .

- c) V bode $x = 1,2$ nájdite hodnotu numerického riešenia metódou Runge – Kutta 4. rádu. Na výpočet použite kroky $h_1 = 0,1$ a $h_2 = 0,05$. Metódou polovičného kroku vypočítajte odhad globálnej chyby metódy v danom bode pre krok h_2 .
- d) Môžeme hľadané numerické riešenie v bode $x = 1,2$ v každej metóde považovať za numerické riešenie nájdené s presnosťou $\varepsilon = 10^{-5}$ vzhľadom na odhad globálnej chyby metódy v tomto bode?

RIEŠENIE

Presné partikulárne riešenie $y(x)$ tejto CU nevieme vypočítať. Vieme nájsť len jej približné partikulárne riešenie numerickými metódami.

Najskôr rovnicu upravíme na tvar $y' = f(x, y)$:

$$y^2 y' - x + y^2 = 0 \Rightarrow y' = \frac{x - y^2}{y^2} = \frac{x}{y^2} - 1 \Rightarrow f(x, y) = \frac{x}{y^2} - 1$$

a) *Vyšetrte existenciu a jednoznačnosť jej riešenia*

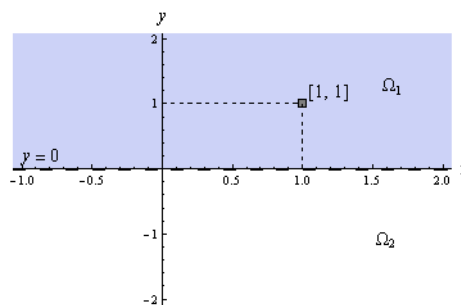
• *Existencia riešenia:*

$$D(f(x, y)) = \{[x, y] \subset E_2; x \in R; y^2 \neq 0\} = R \times R - \{0\}.$$

Funkcia $f(x, y)$ je definovaná a spojitá na oblastiach

$$\Omega_1 = R \times (0, \infty) \text{ a } \Omega_2 = R \times (-\infty, 0).$$

Partikulárne riešenie je definované pre $x \in R$ a vzhľadom na začiatočnú podmienku $[x_0, y_0] = [1, 1] \in \Omega_1$ bude ležať v oblasti $\Omega_1 \Rightarrow$ existencia riešenia CU je zaručená. (Obr. 8.39)



Obr. 8.39

• *Jednoznačnosť riešenia:*

$$\frac{\partial f(x, y)}{\partial y} = \frac{-2x}{y^3}, \quad D\left(\frac{\partial f(x, y)}{\partial y}\right) = \{[x, y] \subset E_2; x \in R; y^3 \neq 0\} = R \times R - \{0\}.$$

Pretože $\frac{\partial f(x, y)}{\partial y}$ je definovaná a spojitá na oblastiach $\Omega_1 = R \times (0, \infty)$ a $\Omega_2 = R \times (-\infty, 0)$, je

spojitá aj v bode $[1, 1] \in \Omega_1$, čo zaručuje jednoznačnosť riešenia CU (Obr. 8.39).

Záver. Pre bod $[1, 1] \in \Omega_1$ existuje práve jedno partikulárne riešenie danej DR $y = \varphi(x)$.

- b) V bode $x = 1,2$ nájdite hodnotu numerického riešenia Heunovou metódou. Na výpočet použite kroky $h_1 = 0,1$ a $h_2 = 0,05$. Metódou polovičného kroku vypočítajte odhad globálnej chyby metódy v danom bode pre krok h_2 .

Keďže hľadáme hodnotu numerického riešenia CU v bode $x = 1,2$ uvedenou metódou, hľadáme približné partikulárne riešenie CU na intervale $\langle 1; 1,2 \rangle$ s krokmi $h_1 = 0,1$ a $h_2 = 0,05$. Počet iterácií, ktoré potrebujeme vypočítať, ako aj počet čiastkových intervalov, na ktoré bude

rozdelený interval $\langle 1;1,2 \rangle$, je $n_1 = \frac{b-a}{h_1} = \frac{1,2-1}{0,1} = 2$ pre krok h_1 a $n_2 = \frac{b-a}{h_2} = \frac{1,2-1}{0,05} = 4$ pre krok h_2 . Schéma výpočtu Heunovou metódou a numerické riešenia DR pre obidva kroky sú v tabuľke Tabuľka 8.7.

Záver. Hodnota numerického riešenia danej CU v bode $x=1,2$ vypočítaného Heunovou metódou pre krok $h_1=0,1$ je $y^1=1,01781930$ a pre krok $h_2=0,05$ je $y^2=1,01740824$.

Odhad globálnej chyby metódy pre krok $h_2=0,05$ v bode $x=1,2$ vypočítame metódou polovičného kroku podľa vzťahu:

$$e_n = \frac{|y^2 - y^1|}{q^p - 1} = \frac{|1,01740824 - 1,01781930|}{2^2 - 1} = \frac{0,00041106}{3} = 0,00013702 = 1,3702 \cdot 10^{-4},$$

kde p je rád metódy. Heunova metóda je rádu $O(h^2)$, $p=2$ a pre q platí $h_1 = q \cdot h_2$, v našom prípade $q=2$.

Heunova metóda																																		
Schéma	Numerické riešenie																																	
	$h_1 = 0,1$		$h_2 = 0,05$																															
$x_{i+1} = x_i + h$ $k_1 = f(x_i, y_i)$ $k_2 = f(x_i + h, y_i + h \cdot k_1)$ $y_{i+1} = y_i + h \cdot \frac{k_1 + k_2}{2}$ $[x_{i+1}, y_{i+1}]$ $i = 0, 1; i = 0, 1, \dots, 4$	<table border="1" style="margin: auto; border-collapse: collapse;"> <thead> <tr> <th style="padding: 2px 5px;">i</th> <th style="padding: 2px 5px;">x_i</th> <th style="padding: 2px 5px;">y_i</th> </tr> </thead> <tbody> <tr> <td style="padding: 2px 5px;">0</td> <td style="padding: 2px 5px;">1</td> <td style="padding: 2px 5px;">1.00000000</td> </tr> <tr> <td style="padding: 2px 5px;">1</td> <td style="padding: 2px 5px;">1.1</td> <td style="padding: 2px 5px;">1.00500000</td> </tr> <tr> <td style="padding: 2px 5px;">2</td> <td style="padding: 2px 5px;">1.2</td> <td style="padding: 2px 5px;">1.01781930</td> </tr> </tbody> </table>		i	x_i	y_i	0	1	1.00000000	1	1.1	1.00500000	2	1.2	1.01781930	<table border="1" style="margin: auto; border-collapse: collapse;"> <thead> <tr> <th style="padding: 2px 5px;">i</th> <th style="padding: 2px 5px;">x_i</th> <th style="padding: 2px 5px;">y_i</th> </tr> </thead> <tbody> <tr> <td style="padding: 2px 5px;">0</td> <td style="padding: 2px 5px;">1</td> <td style="padding: 2px 5px;">1.00000000</td> </tr> <tr> <td style="padding: 2px 5px;">1</td> <td style="padding: 2px 5px;">1.05</td> <td style="padding: 2px 5px;">1.00125000</td> </tr> <tr> <td style="padding: 2px 5px;">2</td> <td style="padding: 2px 5px;">1.1</td> <td style="padding: 2px 5px;">1.00473653</td> </tr> <tr> <td style="padding: 2px 5px;">3</td> <td style="padding: 2px 5px;">1.15</td> <td style="padding: 2px 5px;">1.01020500</td> </tr> <tr> <td style="padding: 2px 5px;">4</td> <td style="padding: 2px 5px;">1.2</td> <td style="padding: 2px 5px;">1.01740824</td> </tr> </tbody> </table>		i	x_i	y_i	0	1	1.00000000	1	1.05	1.00125000	2	1.1	1.00473653	3	1.15	1.01020500	4	1.2	1.01740824
i	x_i	y_i																																
0	1	1.00000000																																
1	1.1	1.00500000																																
2	1.2	1.01781930																																
i	x_i	y_i																																
0	1	1.00000000																																
1	1.05	1.00125000																																
2	1.1	1.00473653																																
3	1.15	1.01020500																																
4	1.2	1.01740824																																

Tabuľka 8.7

c) V bode $x=1,2$ nájdite hodnotu numerického riešenia metódou Runge – Kutta 4. rádu. Na výpočet použite kroky $h_1=0,1$ a $h_2=0,05$. Metódou polovičného kroku vypočítajte odhad globálnej chyby metódy v danom bode pre krok h_2 .

Opäť hľadáme numerické riešenie CU v bode $x=1,2$ tentoraz metódou Runge – Kutta 4. rádu, na intervale $\langle 1;1,2 \rangle$ s krokmi $h_1=0,1$ a $h_2=0,05$.

Pre krok $h_1=0,1$:

- počet čiastkových intervalov (= počet počítaných iterácií): $n_1 = \frac{b-a}{h_1} = \frac{1,2-1}{0,1} = 2$,
- počet deliacich bodov: $n_1 + 1 = 3$.

Pre krok $h_2=0,05$:

- počet čiastkových intervalov (= počet počítaných iterácií): $n_2 = \frac{b-a}{h_2} = \frac{1,2-1}{0,05} = 4$,

- počet deliacich bodov: $n_2 + 1 = 3$.

Schéma výpočtu metódou Runge – Kutta 4. rádu a numerické riešenia DR pre obidva kroky sú v tabuľke Tabuľka 8.8.

Záver. Hodnota numerického riešenia danej CU v bode $x = 1,2$ vypočítaného metódou Runge – Kutta 4. rádu pre krok $h_1 = 0,1$ je $y^1 = 1,01729315$ a pre krok $h_2 = 0,05$ je $y^2 = 1,01729129$.

Odhad globálnej chyby metódy pre krok $h_2 = 0,05$ v bode $x = 1,2$ vypočítame metódou polovičného kroku podľa vzťahu:

$$e_n = \frac{|y^2 - y^1|}{q^p - 1} = \frac{|1,01729129 - 1,01729315|}{2^4 - 1} = \frac{1,87}{15} = 1,25 \cdot 10^{-7},$$

kde p je rád metódy. Metóda Runge – Kutta 4. rádu je rádu $O(h^4)$, $p = 4$, $q = 2$.

Metóda Runge – Kutta 4. rádu																																	
Schéma	Numerické riešenie																																
	$h_1 = 0,1$		$h_2 = 0,05$																														
$x_{i+1} = x_i + h$ $k_1 = f(x_i, y_i)$ $k_2 = f(x_i + \frac{h}{2}, y_i + \frac{h}{2} \cdot k_1)$ $k_3 = f(x_i + \frac{h}{2}, y_i + \frac{h}{2} \cdot k_2)$ $k_4 = f(x_i + h, y_i + h \cdot k_3)$ $y_{i+1} = y_i + h \cdot \frac{(k_1 + 2k_2 + 2k_3 + k_4)}{6}$ $i = 0, 1, \dots, 4$	<table border="1"> <thead> <tr> <th>i</th> <th>x_i</th> <th>y_i</th> </tr> </thead> <tbody> <tr> <td>0</td> <td>1</td> <td>1.00000000</td> </tr> <tr> <td>1</td> <td>1.1</td> <td>1.00466259</td> </tr> <tr> <td>2</td> <td>1.2</td> <td>1.01729315</td> </tr> </tbody> </table>		i	x_i	y_i	0	1	1.00000000	1	1.1	1.00466259	2	1.2	1.01729315	<table border="1"> <thead> <tr> <th>i</th> <th>x_i</th> <th>y_i</th> </tr> </thead> <tbody> <tr> <td>0</td> <td>1</td> <td>1.00000000</td> </tr> <tr> <td>1</td> <td>1.05</td> <td>1.00120795</td> </tr> <tr> <td>2</td> <td>1.1</td> <td>1.00466143</td> </tr> <tr> <td>3</td> <td>1.15</td> <td>1.01010523</td> </tr> <tr> <td>4</td> <td>1.2</td> <td>1.01729129</td> </tr> </tbody> </table>	i	x_i	y_i	0	1	1.00000000	1	1.05	1.00120795	2	1.1	1.00466143	3	1.15	1.01010523	4	1.2	1.01729129
i	x_i	y_i																															
0	1	1.00000000																															
1	1.1	1.00466259																															
2	1.2	1.01729315																															
i	x_i	y_i																															
0	1	1.00000000																															
1	1.05	1.00120795																															
2	1.1	1.00466143																															
3	1.15	1.01010523																															
4	1.2	1.01729129																															

Tabuľka 8.8

d) Môžeme hľadať hodnotu numerického riešenia v bode $x = 1.2$ pre krok $h_2 = 0,05$ v každej metóde považovať za numerické riešenie nájdené s presnosťou $\varepsilon = 10^{-5}$?

- Pre Heunovu metódu je $e_n = 1,3702 \cdot 10^{-4}$, $e_n > \varepsilon$, $\varepsilon = 10^{-5}$. Hodnotu riešenia v danom bode nájdenú Heunovou metódou nemožno považovať za numerické riešenie s presnosťou $\varepsilon = 10^{-5}$.

Pre metódu Runge – Kutta 4. rádu je $e_n = 1,25 \cdot 10^{-7}$, $e_n < \varepsilon$, $\varepsilon = 10^{-5}$. Hodnotu riešenia v danom bode nájdenú metódou Runge – Kutta 4. rádu možno považovať za numerické riešenie s presnosťou $\varepsilon = 10^{-5}$.

LITERATÚRA

- [1] BURDEN, R.L. - FAIRES, J.D.: *Numerical Analysis*. 7th. Edition. Brooks/Cole, Pacific Grove, CA, USA, 2001, pp. 841. ISBN 0-534-38216-9
- [2] BUŠA, J. - PIRČ, V. - SCHRÖTTER, Š.: *Numerické metódy, pravdepodobnosť a matematická štatistika*. FEI TU Košice: Edícia vysokoškolských učebníc, 2006. ISBN 80-8073-632-4
- [3] DEMIDOVÍČ, B.P. - MARON, I.A.: *Základy numerické matematiky*. Praha: SNTL, 1966, s. 721.
- [4] DRIENSKY, D.: *Úvod do inžinierskej pedagogiky*. Bratislava: STU, 1999.
- [5] ELIÁŠ, J.: *Matematika. Úvod do numerickej matematiky*. Bratislava: Edičné stredisko SVŠT, 1986, s. 221. ČÚKK Š-331/66
- [6] EPPERSON, J.F.: *An Introduction to Numerical Methods and Analysis*. New York: John Wiley & Sons, 2001, pp. 556. ISBN 0-471-316474
- [7] FAJMON, B. - RŮŽIČKOVÁ, I.: *Matematika 3*. Brno: UMAT FEKT VUT, 2003, s. 255.
- [8] GERALD, C.F. - WHEATLEY, P.O.: *Applied Numerical Analysis*. 6th. Edition. Addison Wesley Publishing Company, 1998, pp. 698. ISBN-10: 020187072X
- [9] GROFÍK, R. a i.: *Štatistika*. Bratislava: Príroda, 1987.
- [10] IVAN, J.: *Matematika I*. 1. vyd. Bratislava: Alfa, 1983.
- [11] IVAN, J.: *Matematika II*. 1. vyd. Bratislava: Alfa, 1989.
- [12] KUČERA, R.: *Numerické metódy*. Ostrava: VŠB-Technická univerzita, 2006. ISBN 80-248-1198-7
- [13] MIKA, S.: *Numerické metódy algebry*. Praha: SNTL, 1985, s. 176. 04-011-85
- [14] MOŠOVÁ, V.: *Numerické metódy*. Olomouc: UP-2003, s. 147. ISBN 80-244-0620-9
- [15] RALSTON, A.: *Základy numerické matematiky*. České vydání 2. Praha: Academia, 1978, s. 636. 104-21-852
- [16] RIEČANOVÁ, Z. a i.: *Numerické metódy a matematická štatistika*. Bratislava: Alfa, 1987, s. 496. 063-559-87 NMA
- [17] ŠALÁT, T. et al.: *Malá encyklopédia matematiky*. 2. vyd. Bratislava: Obzor, 1978, s. 856. 65-035-78

RNDr. Jana Gabková, PhD., RNDr. Daniela Richtáriková, PhD.

ZÁKLADY NUMERICKEJ MATEMATIKY
Teória a riešené príklady

Vydala Slovenská technická univerzita v Bratislave v Nakladateľstve STU, Bratislava,
Vazovova 5, v roku 2012.

Edícia skrípt

Rozsah 213 strán, 123 obrázkov, 57 tabuliek, 8,721 AH, 9,033 VH, 1. vydanie,
edičné číslo 5613, tlač Nakladateľstvo STU v Bratislave.

85 – 214 – 2012

ISBN 978-80-227-3658-9