Нумеричка математика

Наташа Ћировић

Београд, 2018

др Наташа Ћировић, доцент
Електротехнички факултет, Универзитет у Београду
и-мејл: natasa@etf.bg.ac.rs

НУМЕРИЧКА МАТЕМАТИКА
електронски уџбеник

Рецензенти
Проф. др Синиша Н. Јешић
Редовни професор

Проф. др Ненад Цакић
Редовни професор

Лектор
Весна Опарушић

Наставно-научно веће Електротехничког факултета одобрило је објављивање овог уџбеника одлуком број 358/3 од 21.03.2018. године на 824. седници.

Издавач
Електротехнички факултет, Универзитет у Београду
Булевар Краља Александра 73, 11000 Београд, Србија

Штампа: Наташа Ћировић, Београд, 2018
Тираж: 50 примерака

ISBN: 978-86-7225-068-8

Предговор

Савремена електротехника и рачунарство подразумевају значајне примене нумеричке математике.

Овај уџбеник развијан је током више од 10 година, колико аутор учествује у настави нумеричке математике у склопу више предмета на основним академским студијама Електротехничког факултета Универзитета у Београду. Градиво обухваћено овом књигом покрива курс Нумеричка математика, изборни курс на модулима Електроника, Физичка електроника, Сигнали и системи, Енергетика и Телекомуникације и информационе технологије, као и део курса Нумеричка анализа и дискретна математика на модулима Рачунарска техника и информатика, Сигнали и системи и на студијском програму Софтверско инжењерство.

Материјал прати теоријске поставке наставних јединица које се обрађују на предавањима, као и задатке који прате ове наставне јединице и раде се на часовима вежби. Додатно, за сваку наставну јединицу дати су задаци с резултатима који прате лабораторијске вежбе из ових предмета.

Текст је припремљен у систему LaTeX 2_ε, а слике у пакету GeoGebra (http://www.geogebra.org).

Аутор се захваљује свим професорима и колегама који су учествовали у досадашњем извођењу наставе на овом предмету и дали свој допринос унапређењу наставе у овој области. Такође, посебно је значајна улога студената који су својим питањима и сугестијама значајно утицали на коначни облик овог уџбеника.

Наташа А. Ћировић

Фебруар, 2018
Београд

Садржај

Предговор **iii**

1 Појам грешке **1**
1.1 Врсте грешака . . . 1
1.1.1 Апсолутна и релативна грешка . . . 4
1.1.2 Заокруживање и сигурне цифре . . . 5
1.1.3 Условљеност проблема и стабилност алгоритма . . . 8

2 Нелинеарне једначине **10**
2.1 Локализација решења једначине . . . 10
2.2 Метода половљења интервала . . . 11
2.3 Метода просте итерације . . . 15
2.4 Њутнова метода . . . 20
2.5 Метода сечице . . . 25

3 Системи линеарних једначина **29**
3.1 Норма матрица . . . 30
3.2 Директне методе . . . 32
3.3 Гаусова метода елиминације . . . 33
3.4 LU декомпозиција . . . 37
3.5 Пивотирање . . . 48
3.6 Итеративне методе . . . 54

4 Интерполација **63**
4.1 Полиномска интерполација . . . 64
4.2 Лагранжова интерполација . . . 67
4.3 Оцена грешке интерполације . . . 69
4.4 Њутнова интерполација . . . 78

4.5 Инверзна интерполација 88

5 Нумеричко диференцирање 92

5.1 Диференцирање помоћу интерполационих полинома 92

5.2 Диференцирање помоћу Тејлоровог реда 95

6 Нумеричка интеграција 106

6.1 Њутн-Коутсове квадратурне формуле 108

6.1.1 Основна трапезна формула 108

6.1.2 Основна Симпсонова формула 109

6.1.3 Трапезна формула 111

6.1.4 Симпсонова формула 113

6.1.5 Грешка заокруживања 115

6.2 Ромбергова интеграција 120

6.2.1 Ричардсонова екстраполација 120

6.2.2 Ромбергова интеграција 121

6.3 Формуле за нумеричку интеграцију 127

6.4 Решавање несвојствених интеграла 132

Литература 137

Глава 1

Појам грешке

Нумеричка математика бави се решавањем нумеричких проблема користећи само једноставне аритметичке операције и развојем ефикасних алгоритама којима се из датих података добијају нумерички резултати. Развијени алгоритми имлементирају се у одабраном програмском језику. Поред тога, постоје многи програмски пакети који имплементирају различите нумеричке алгоритме (нпр. Matlab, Octave, Maple, Mathematica, Scilab, и други). Алгоритам за решавање бирамо у зависности од природе проблема.

Основна карактеристика нумеричке математике је неизбежно присуство грешака, које не можемо уклонити, већ се само можемо надати да могу бити „довољно мале” и да имамо поуздану процену њихове величине.

1.1 Врсте грешака

Неотклоњиве грешке

При решавању реалног проблема формирамо математички модел који га описује. Формирани модел често представља идеализовану поставку, која занемарује мање значајне податке или описује проблем под специјалним условима. На пример, проблеми који су по природи нелинеарни некада се моделују линеарним моделима.

Решавањем приближног проблема у односу на реалан добијамо решење које одступа од тачног решења реалног проблема. Ово су *грешке у математичком моделу*.

Још један извор грешака су *грешке у улазним подацима.* До оваквих грешака може доћи када су улазни подаци добијени као резултат експерименталних мерења на инструментима који никада нису апсолутно прецизни.

Овакве грешке не можемо избећи, већ евентуално смањити (нпр. прецизнијим мерењима). На нивоу нумеричких метода, чиме се овде бавимо, не можемо утицати на њих. Ипак, треба да будемо свесни постојања оваквих грешака зато што утичу на тачност решења која добијамо применом нумеричких метода.

Грешке методе

Математички модел који решавамо често нема аналитичко решење. У том случају, примењујемо адекватну нумеричку методу за решавање проблема.

Нумеричку методу можемо добити *дискретизацијом* – заменом континуалних функција, модела или једначина одговарајућим дискретним облицима. На пример, изводе замењујемо коначним разликама, а бесконачне суме коначним.

Нелинеарне проблеме често решавамо применом *итеративних процеса*, при чему добијамо низ приближних решења. Уколико бисмо могли да поновимо итеративни процес бесконачно много пута, добили бисмо тачно решење. У пракси, заустављамо се на коначном броју итерација, па добијамо приближно решење које одступа од тачног. Повећањем броја итерација можемо утицати на повећање тачности.

Пре него што дамо један познат пример нумеричке методе, која настаје заменом бесконачне суме коначном, подсетимо се неких појмова из математичке анализе функција једне променљиве.

Тејлоров ред: За бесконачно диференцијабилну функцију f на интервалу који садржи тачке x и a, важи развој у Тејлоров ред:

$$f(x) = \sum_{k=0}^{+\infty} \frac{f^{(k)}(a)}{k!}\,(x-a)^k.$$

Уколико задржимо само првих неколико чланова Тејлоровог реда, добијамо апроксимацију функције f, представљену Тејлоровим полиномом:

$$f(x) = f(a) + f'(a)\,(x-a) + \frac{f''(a)}{2!}\,(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}\,(x-a)^n + R_n(x),$$

при чему је $R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}$ грешка апроксимације, где је ξ нека тачка која припада интервалу који садржи тачке x и a.

Пример 1. Функција $\sin x$ може се развити у Тејлоров ред облика

$$\sin x = \sum_{k=0}^{+\infty}(-1)^k \frac{x^{2k+1}}{(2k+1)!}.$$

Проценимо приближне вредности функције $\sin x$ у тачки $\frac{\pi}{6}$ апроксимацијама Тејлоровим полиномом степена 1, 3, 5 и 7 у околини тачке $a = 0$:

$$\begin{aligned}
\sin x &\approx T_1(x) = x \\
\sin x &\approx T_3(x) = x - \frac{x^3}{3!} \\
\sin x &\approx T_5(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} \\
\sin x &\approx T_7(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!}
\end{aligned}$$

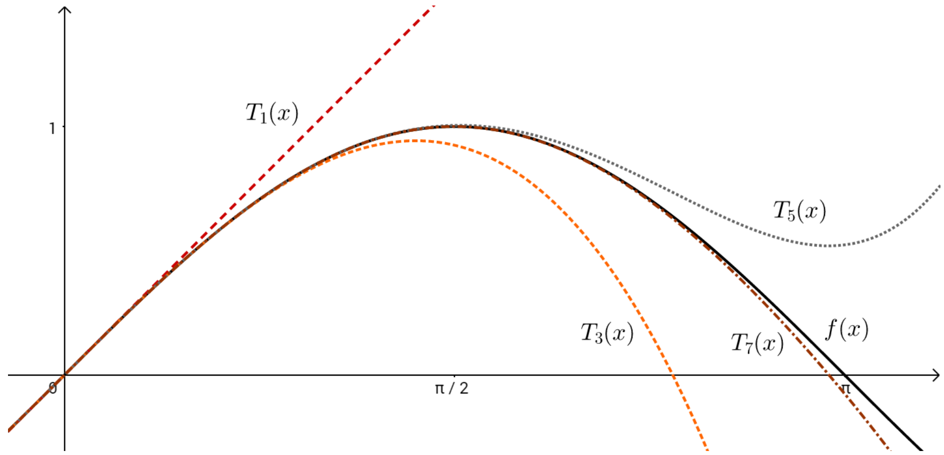
Сада је

$$\begin{aligned}
T_1\left(\frac{\pi}{6}\right) &= \frac{\pi}{6} \approx 0,5235987756 \\
T_3\left(\frac{\pi}{6}\right) &= \frac{\pi}{6} - \frac{\pi^3}{3!\,6^3} \approx 0,4996741794 \\
T_5\left(\frac{\pi}{6}\right) &= \frac{\pi}{6} - \frac{\pi^3}{3!\,6^3} + \frac{\pi^5}{5!\,6^5} \approx 0,5000021326 \\
T_7\left(\frac{\pi}{6}\right) &= \frac{\pi}{6} - \frac{\pi^3}{3!\,6^3} + \frac{\pi^5}{5!\,6^5} - \frac{\pi^7}{7!\,6^7} \approx 0,4999999919
\end{aligned}$$

С обзиром на то да је тачна вредност $\sin\frac{\pi}{6} = 0,5$, грешке ових апроксимација, $R_n(x) = \sin x - T_n(x)$, износе:

n	T_n	$\lvert R_n\rvert$
1	$0,5235987756$	$0,0235987756$
3	$0,4996741794$	$0,0003258206$
5	$0,5000021326$	$0,0000021326$
7	$0,4999999919$	$0,0000000081$

Слика 1.1 приказује графике функције $\sin x$ и њених апроксимација Тејлоровим полиномима. Поређењем графика можемо визуелно проценити квалитет апроксимација функције, који расте с повећањем степена



Слика 1.1: Апроксимације функције $\sin(x)$ Тејлоровим полиномима

полинома и с приближавањем тачке у којој апроксимирамо функцију тачки 0, у којој формирамо Тејлоров полином.

Грешке заокруживања

Свако израчунавање које обухвата реалне (ирационалне) бројеве подразумева постојање грешке заокруживања. Рачунари раде само с бројевима записаним са коначним бројем децимала. На пример, $\sqrt{2}$ представља број, који помножен самим собом, даје број 2. Може се доказати да је $\sqrt{2}$ ирационалан број, тј. да запис броја има бесконачно много децимала. Рачунајући помоћу рачунара, морамо да се ограничимо на запис с коначним бројем децимала, па самим тим долази до грешке. Овакве грешке називају се *грешке заокруживања*.

1.1.1 Апсолутна и релативна грешка

Ако је x реалан број, а $\bar{x}$ његова приближна вредност, тада је *апсолутна грешка* $|x - \bar{x}|$.

Квалитет добијеног приближног решења не зависи само од вредности апсолутне грешке, већ и од величине тачне вредности. На пример, приближно решење $25034,83$ с апсолутном грешком $0,01$ могло би бити довољно добро, али јасно је да $0,002$ с апсолутном грешком од $0,01$ не би било ни од какве користи. Ово се решава разматрањем *релативне* *грешке* $\dfrac{|x-\bar{x}|}{|x|}$, која је дефинисана за $x \neq 0$. Када није позната тачна вредност x, релативну грешку можемо приближно проценити дељењем са $|\bar{x}|$.

Иако је апсолутна грешка једнака за претходна два примера, релативне грешке су неупоредиве: $\frac{0{,}01}{25034{,}83} = 3,99 \cdot 10^{-7}$ и $\frac{0{,}01}{0{,}002} = 5$.

Тачност израчунавања нам говори у којој мери приближно решење одступа од тачног решења постављеног проблема и исказује се вредношћу грешке израчунавања.

Ако је задата тачност ε са којом смо одредили приближно решење $\bar{x}$ неког проблема, можемо закључити да тачно решење x припада сегменту $[\bar{x}-\varepsilon, \bar{x}+\varepsilon]$, тј. $|x-\bar{x}| \leq \varepsilon$.

1.1.2 Заокруживање и сигурне цифре

Нека је реалан број x у декадном запису

$$x = \pm(\alpha_1 10^n + \alpha_2 10^{n-1} + \ldots + \alpha_k 10^{n-k+1} + \alpha_{k+1} 10^{n-k} + \ldots).$$

При заокруживању броја на k цифара примењујемо следећа правила:

- α_k се не мења ако је $\alpha_{k+1} < 5$ или $\alpha_{k+1} = 5$ и α_k је парно,
- α_k се повећава за 1 ако је $\alpha_{k+1} > 5$ или $\alpha_{k+1} = 5$ и α_k је непарно.

Значајне цифре броја су све цифре почевши од пре ненула цифре слева. За значајну цифру α_k кажемо да је *сигурна цифра у ужем смислу* ако је

$$|x-\bar{x}| \leq \frac{1}{2} \cdot 10^{n-k+1}.$$

Пример 2. Нека је $x = 0,2591821$ број дат с тачношћу $0,5 \cdot 10^{-4}$. То значи да је апсолутна грешка $|x-\bar{x}| \leq 0,5 \cdot 10^{-4}$, тј. да су прве 4 значајне цифре сигурне, и то су 2,5,9 и 1. Све цифре десно од последње сигурне цифре треба одбацити, пошто нису сигурне. При одбацивању несигурних цифара, последња сигурна цифра се мења тако да буде сигурна. Након одбацивања несигурних цифара, имамо да је $\bar{x} = 0,2592$. Заиста, сада је $|x-\bar{x}| = 0,0000179 < 0,5 \cdot 10^{-4}$.

Уколико је познато да приближан број има k сигурних цифара, тада за апсолутну грешку важи да је $|x-\bar{x}| \leq 10^{n-k+1}$.

Пример 3. Нека је $\bar{x} = 0,2592$ приближна вредност броја x са 4 сигурне цифре. Одатле можемо да закључимо да $x \in [0,25915; 0,25925]$, тј. цифре приближне и тачне вредности на 5. позицији се разликују за највише 5. За апсолутну грешку важи и да је $|x-\bar{x}| \leq 0,00005 = \frac{1}{2}\cdot 10^{-4}$, па се и кроз процену апсолутне грешке види да имамо 4 значајне цифре.

Процену броја сигурних цифара можемо извршити и на основу оцене релативне грешке. У том случају, за број са k сигурних цифара (у ужем смислу) важиће да је

$$\frac{|x-\bar{x}|}{|x|} \leq \frac{1}{2}\cdot 10^{-k+1} = 5\cdot 10^{-k}.$$

Пример 4. Нека је $\bar{x} = 10\,000$ приближна, а $x = 10\,002$ тачна вредност броја. Тада је апсолутна грешка $|x-\bar{x}| = 2$, а релативна грешка износи $\frac{|x-\bar{x}|}{x} \approx 2\cdot 10^{-4} \leq 5\cdot 10^{-4}$, па закључујемо да приближан број има 4 значајне цифре.

Апсолутна грешка указује на број сигурних децималних цифара приближног броја, а релативна грешка на укупан број сигурних цифара.

Размотримо кроз примере на који начин основне аритметичке операције утичу на број сигурних цифара у вредности резултата.

Пример 5. Извршимо сабирање

$$2,3 + 3,45 = 5,75.$$

Иако смо записали 2 децималне цифре у збиру, само прва је сигурна. Заиста, први сабирак може припадати интервалу $[2,26\,;2,34]$, па збир може да буде $2,26+3,45 = 5,71$, а такође и $2,34+3,45 = 5,79$. Овде има смисла задржати само прву децималну цифру, колико их има и сабирак са мање децималних цифара.

Пример 6. Извршимо множење

$$1,23\cdot 4,5 = 5,535.$$

Иако смо у резултату записали 4 цифре, само прве две су сигурне. У овом случају задржаћемо само прве две сигурне цифре, колико их има и чинилац с мање сигурних цифара.

Напомена. При операцијама сабирања или одузимања задржавамо исти број децималних позиција као број с најмање сигурних цифара на децималним позицијама. При операцијама множења или дељења задржавамо исти број сигурних цифара као број који их има најмање. За бројеве који учествују у операцијама за које знамо да су потпуно тачни, сматрамо да имају бесконачно много сигурних цифара.

Пример 7. Решавамо систем линеарних једначина за непознату y:

$$0,2037\,x + 0,3122\,y = 0,8472$$
$$0,4082\,x + 0,6247\,y = 0,9745$$

Најпре ћемо извршити израчунавање са 3 значајне цифре, затим са 4, и на крају са 10 значајних цифара.

- *Израчунавање са 3 значајне цифре.* Заокружујемо све бројеве у почетном проблему на 3 цифре, као и све међурезултате.

$$0,204\,x + 0,312\,y = 0,847$$
$$0,408\,x + 0,625\,y = 0,974$$

  Множимо прву једначину коефицијентом $-0,408/0,204 = -2,00$ и сабирамо са другом једначином, при чему друга једначина постаје:

$$0 \cdot x + 0,001\,y = -0,720,$$

  тј. $y = -720$.

- *Израчунавање са 4 значајне цифре.* Задржавамо све значајне цифре и међурезултате заокружујемо на 4 децимале. Множимо прву једначину коефицијентом $-0,4082/0,2037 \approx -2,004$ и сабирамо са другом једначином, при чему друга једначина постаје (након заокруживања на 4 децималне цифре):

$$0 \cdot x - 0,0009\,y = -0,7233,$$

  тј. $y = 803,7$.

- *Израчунавање са 10 значајних цифрара.* Задржавамо све значајне цифре и међурезултате заокружујемо на 10 децимала. Множимо прву једначину коефицијентом $-0,4082/0,2037 \approx -2,0039273441$ и сабирамо са другом једначином, при чему друга једначина постаје (након заокруживања на 10 децималних цифара):

$$0 \cdot x - 0,0009261168\, y = -0,7232272459,$$

тј. $y = 780,9244426837$.

Напомена. Препорука је да се заокруживање врши након последњег израчунавања, а не након сваког корака.

1.1.3 Условљеност проблема и стабилност алгоритма

Важну улогу у нумеричкој математици имају појмови *условљености проблема* и *стабилности алгоритма*, тј. нумеричке методе.

Проблем је *лоше условљен* уколико мале промене у улазним подацима производе велике промене у резултату. Може се десити да за такав проблем не постоји алгоритам који може да произведе задовољавајуће решење. Понекад такав проблем можемо решити одређеном модификацијом (нпр. трансформацијом система линеарних једначина).

Пример 8. Систем једначина:

$$\begin{gathered} x + 4\, y = 5 \\ x + 4,00001\, y = 5,00001 \end{gathered}$$

има решење $x = 1$, $y = 1$. Посматрајмо сада мало измењене једначине:

$$\begin{gathered} x + 4\, y = 5 \\ x + 3,99999\, y = 5,00002 \end{gathered}$$

Решење овог система једначина је $x = 13$, $y = -2$. Уколико прихватимо да су ове вредности решења полазног система, добили бисмо да десна страна једнакости има вредности 5 и $4,99998$, што је блиско полазним вредностима.

Проблеми који су лоше условљени немају једноставна решења. Понекад је у формулисању проблема могуће извршити одређене модификације, којима се овај проблем превазилази. Условљеност проблема не зависи од избора методе коју примењујемо.

Стабилност нумеричког алгоритма односи се на то у којој мери је метода прецизна, тј. колико приближно решење које добијамо применом методе одступа од тачног решења проблема који решавамо. Алгоритам може имати различите нивое нумеричке стабилности, пошто се многа израчунавања могу вршити на више начина, који су алгебарски еквивалентни, али могу произвести нумерички различите резултате.

Пример 9. [10] Функцију e^x можемо представити Маклореновим редом, за све коначне вредности x,

$$e^x = \sum_{n=0}^{+\infty} \frac{x^n}{n!}.$$

Показује се да за позитивне вредности променљиве x и за одговарајући број сабирака Маклореновог реда можемо добити задовољавајуће приближне вредности за функцију e^x. Међутим, за негативне вредности x може се десити да добијемо вредности које значајно одступају од тачних, и чак могу да се разликују по знаку. Овај проблем може се превазићи модификацијом алгоритма за негативне вредности x, на следећи начин:

$$e^{-x} = \frac{1}{e^x} = \frac{1}{\sum_{n=0}^{+\infty} \frac{x^n}{n!}}.$$

Глава 2

Нелинеарне једначине

Решавамо нелинеарну једначину облика:

$$f(x) = 0. \tag{2.1}$$

Решења ове једначине називају се и *нуле функције* f.

У овом поглављу ћемо подразумевати да је функција f непрекидна на интервалу $[a, b]$, што означавамо са $f \in C[a, b]$.

2.1 Локализација решења једначине

Тачно решење x^* нелинеарне једначине апроксимирамо приближним решењем $\bar{x}$, добијеним адекватно одабраном методом. Без обзира на то коју нумеричку методу користимо, најпре је потребно одредити колико решења постоји и локализовати свако од њих.

Локализација решења је процес одређивања интервала $[a, b]$ који садржи решење x^* једначине $f(x) = 0$. Након што локализујемо решења, можемо да применимо адекватну нумеричку методу којом одређујемо приближно решење $\bar{x} \in [a, b]$ једначине. Да би интервал $[a, b]$ садржао решење једначине $f(x) = 0$ потребно је да буде испуњен услов:

$$f(a)\, f(b) < 0. \tag{2.2}$$

Локализацију решења можемо извршити помоћу скице графика функције, на основу које одређујемо интервал $[a, b]$ који садржи решење x^* једначине. Такође, можемо одредити вредност функције f у неколико тачака и на основу тога проценити на ком интервалу функција мења знак. Уколико једначина има више од једног решења, примењујемо одабрану методу на различитим интервалима на којима смо лоцирали решење.

Став 2.1 *Ако је функција f непрекидна на интервалу $[a, b]$ и ако важи $f(a)\, f(b) < 0$, тада на интервалу (a, b) једначина $f(x) = 0$ има бар једно решење.*

Став 2.2 *Ако је функција f непрекидна на интервалу $[a, b]$, $f(a)\, f(b) < 0$ и ако је f монотона функција, тада на интервалу (a, b) једначина $f(x) = 0$ има тачно једно решење.*

Теорема 2.1 (Теорема о процени грешке приближним израчунавањем решења једначине) *Нека је $f \in C[a, b]$ диференцијабилна функција. Ако је $x^* \in [a, b]$ тачно решење једначине $f(x) = 0$, а $\bar{x}$ њено приближно решење, и ако је $0 < m_1 \leq \min\limits_{x\in[a,b]} |f'(x)| < \infty$, тада важи процена:*

$$|x^* - \bar{x}| \leq \frac{|f(\bar{x})|}{m_1} \tag{2.3}$$

Доказ. Ако на одсечку $(x^*, \bar{x})$ $[(\bar{x}, x^*)]$ применимо Лагранжову теорему о средњој вредности, имамо да је

$$\frac{f(\bar{x}) - f(x^*)}{\bar{x} - x^*} = f'(\xi), \quad \xi \in (x^*, \bar{x}). \qquad \left[\frac{f(x^*) - f(\bar{x})}{x^* - \bar{x}} = f'(\xi), \quad \xi \in (\bar{x}, x^*)\right]$$

С обзиром на то да је $f(x^*) = 0$, важи да је

$$|x^* - \bar{x}| = \left|\frac{f(\bar{x})}{f'(\xi)}\right| \leq \frac{|f(\bar{x})|}{m_1}.$$

□

2.2 Метода половљења интервала

Метода половљења интервала је метода постепене претраге, у којој се интервал на коме је лоцирано решење полови у сваком кораку. Уколико функција мења знак на интервалу, процењује се знак функције у средишњој тачки интервала. За приближну вредност решења једначине узима се средишња тачка подинтервала на коме функција мења знак. Процес се понавља како бисмо добили прецизније процене решења.

Наводимо алгоритам за методу половљења интервала:

КОРАК 1: Лоцирати решење једначине на интервалу $[a_0, b_0]$, тако да је испуњен услов $f(a_0)\, f(b_0) < 0$.

КОРАК 2: Одређује се приближна вредност решења са

$$\bar{x}_0 = \frac{a_0 + b_0}{2}.$$

КОРАК 3: Вршимо процену која половина интервала садржи решење једначине:

- Ако је $f(\bar{x}_0) = 0$,
  тада је $x^* = \frac{a_0+b_0}{2}$ решење једначине.
- Ако је $f(a_0)\, f(\bar{x}_0) < 0$,
  презначавамо $a_1 := a_0, \quad b_1 := \bar{x}_0$
  и враћамо се на КОРАК 2 са a_1 и b_1.
- Ако је $f(\bar{x}_0)\, f(b_0) < 0$,
  презначавамо $a_1 := \bar{x}_0, \quad b_1 := b_0$
  и враћамо се на КОРАК 2 са a_1 и b_1.

Овим поступком формирамо низ уметнутих интервала

$$\ldots [a_2, b_2] \subset [a_1, b_1] \subset [a_0, b_0],$$

при чему сваки од њих садржи решење x^* једначине.

Процена грешке

Потребно је проценити када стајемо с поступком. Уколико је тачност с којом треба да одредимо решење једначине ε, због поступка којим одређујемо приближно решење, важи да је:

$$|x^* - \bar{x}_n| \leq \frac{b_n - a_n}{2} \leq \frac{b_0 - a_0}{2^{n+1}} \leq \varepsilon.$$

Из услова $\frac{b_0-a_0}{2^{n+1}} \leq \varepsilon$, одређује се број итерација n половљења интервала потребних да би се постигла захтевана тачност.

1. Методом половљења интервала решити једначину

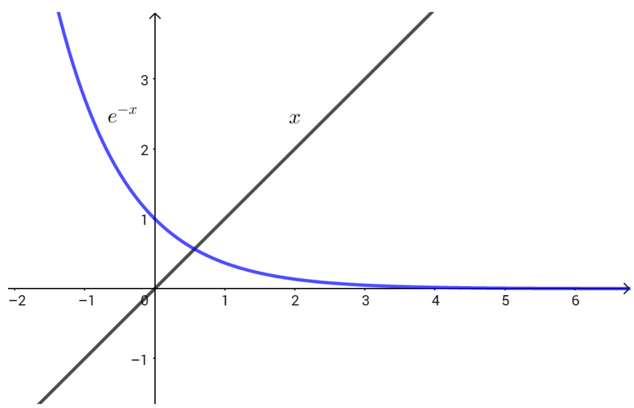
$$f(x) = e^{-x} - x = 0$$

са тачношћу $\varepsilon = 5 \cdot 10^{-2}$.

Решење. Најпре процењујемо колико решења постоји. Да бисмо извршили процену, представимо задату једначину у измењеном облику:

$$e^{-x} = x.$$

Пресечне тачке графика функција e^{-x} и x су нуле функције f.



Слика 2.1: Локализација решења једначине $e^{-x} - x = 0$

Скицирањем графика ових функција процењујемо да задата једначина има једно решење, које припада интервалу $[0, 1]$. Заиста, вредности функције f у крајевима интервала су супротног знака:

$$f(0)\, f(1) = 1 \cdot (-0,632) < 0.$$

Формирамо низ уметнутих интервала, при чему почетни интервал означимо са $[a_0, b_0] = [0, 1]$. Проценимо најпре колико је потребно уметнутих интервала како бисмо постигли захтевану тачност $\varepsilon = 5 \cdot 10^{-2}$. На основу процене $\dfrac{b_0 - a_0}{2^{n+1}} \leq \varepsilon$, имамо да је

$$\frac{1-0}{2^{n+1}} \leq 5 \cdot 10^{-2} \quad \Rightarrow \quad 2^{n+1} \geq \frac{1}{5} \cdot 10^2 = 20 \quad \Rightarrow \quad n \geq 3,322$$

Дакле, довољна су 4 половљења интервала да би се постигла тражена тачност.

n	a_n	b_n	$\bar{x}_n$	$f(a_n)$	$f(b_n)$	$f(\bar{x}_n)$
0	0	1	$0,500$	1	$-0,632$	$0,106$
1	$0,5$	1	$0,75$	$0,106$	$-0,632$	$-0,278$
2	$0,5$	$0,75$	0.625	$0,106$	$-0,278$	$-0,090$
3	$0,5$	$0,625$	$0,5625$	$0,106$	$-0,090$	$0,007$
4	$0,5625$	$0,625$	$0,59375$	$0,007$	$-0,090$	$-0,041$

Након 4 половљења интервала добили смо да је приближно решење $\bar{x}_n = 0,59375$. С обзиром на постигнуту тачност решења, од значаја су нам само прве две цифре, па можемо да пишемо:

$$\bar{x}_n = 0,59.$$

■

Задаци за вежбу

1. Израчунати вредност $\sqrt[3]{5}$ применом методе половљења интервала са 3 сигурне цифре.

Резултат. Број $\sqrt[3]{5}$ је реално решење једначине $x^3 - 5 = 0$.
$\sqrt[3]{5} \approx 1,71$.

2. Методом половљења интервала са тачношћу $5 \cdot 10^{-3}$, решити једначину

$$e^x - 2 = \cos e^x - 2.$$

Резултат. $x \approx 1,004$

3. Методом половљења интервала са тачношћу $5 \cdot 10^{-3}$, решити једначину

$$x = 2 \sin x.$$

Резултат. Користити непарност функције.
$x_{1,2} \approx \pm 1,894; \quad x_3 = 0.$

4. Методом половљења интервала са тачношћу $5 \cdot 10^{-3}$, одредити најмање позитивно решење једначине

$$x = \operatorname{tg} x.$$

Резултат. $x \approx 4,496$

2.3 Метода просте итерације

Проблем решавања једначине $f(x) = 0$, тј. одређивање нула функције f, може да се преформулише на следећи начин:

$$g(x) = x,$$

за неку функцију g. Оваква модификација проблема није јединствена. На овај начин проблем се своди на одређивање *непокретне тачке функције* $g(x)$, тј. тачке x^* која је решење једначине $x^* = g(x^*)$. Оваква формулација проблема доводи нас до *методе просте итерације*. Најпре размотримо услове под којима функција $g(x)$ има непокретну тачку и када је она јединствена.

Теорема 2.2 *Нека је функција g непрекидна на интервалу $[a, b]$. Ако је $a \le g(x) \le b$ за свако $x \in [a, b]$, тада g има непокретну тачку на интервалу $[a, b]$. Додатно, ако је g диференцијабилна на (a, b) и ако постоји константа $0 < k < 1$, таква да за свако $x \in (a, b)$ важи*

$$|g'(x)| \le k,$$

тада g има јединствену непокретну тачку.

Доказ. Ако је $g(a) = a$ или $g(b) = b$, тада g има непокретну тачку у једном од крајева интервала. У супротном, мора важити да је $g(a) > a$ и $g(b) < b$. Тада је функција $h(x) = g(x) - x$ непрекидна на $[a, b]$, важи да је

$$h(a) = g(a) - a > 0, \qquad h(b) = g(b) - b < 0,$$

па постоји тачка $x^* \in (a, b)$ за коју важи да је $h(x^*) = 0$, тј. $g(x^*) = x^*$.

Претпоставимо да је $|g'(x)| \le k < 1$ и да су $x^*, y^* \in (a, b)$ непокретне тачке функције g.

Ако претпоставимо да је $x^* < y^*$, на основу Лагранжове теореме о средњој вредности следи да постоји $\xi \in (x^*, y^*)$ тако да је $\frac{g(x^*)-g(y^*)}{x^*-y^*} = g'(\xi)$. Одатле следи

$$|x^* - y^*| = |g(x^*) - g(y^*)| = |g'(\xi)|\,|x^* - y^*| \le k\,|x^* - y^*| < |x^* - y^*|,$$

што је контрадикција, па закључујемо да је $x^* = y^*$. □

Ако је g непрекидна функција за коју знамо да има непокретну тачку на интервалу $[a, b]$, можемо пробати да одредимо ту тачку израчунавајући вредности $g(x)$, за неко $x \in [a, b]$. Некада није могуће одредити тачну вредност непокретне тачке. Наредна теорема даје услове и начин на који се можемо произвољно приближити тој вредности.

Теорема 2.3 *Нека је функција g непрекидна на $[a, b]$ и $a \leq g(x) \leq b$ за свако $x \in [a, b]$. Додатно, нека је g диференцијабилна на (a, b) и нека постоји константа $0 < k < 1$ таква да је $|g'(x)| \leq k$, за свако $x \in [a, b]$. Тада за произвољно $x_0 \in [a, b]$ низ дефинисан са $x_n = g(x_{n-1}), (n \in \mathbb{N})$, конвергира ка јединственој непокретној тачки $x^* \in [a, b]$ функције g.*

Доказ. На основу претходне теореме закључујемо да функција g има јединствену непокретну тачку x^* на интервалу $[a, b]$. Такође, због тога што $g(x) \in [a, b]$ за свако $x \in [a, b]$, низ $\{x_n\}_{n=0}^{+\infty}$ је добро дефинисан и припада интервалу $[a, b]$. На основу услова теореме и Лагранжове теореме о средњој вредности важи да је

$$|x_n - x^*| = |g(x_{n-1}) - g(x^*)| = |g'(\xi_n)|\,|x_{n-1} - x^*| \leq k\,|x_{n-1} - x^*|,$$

где $\xi_n \in (a, b)$. Индукцијом добијамо да је

$$|x_n - x^*| \leq k\,|x_{n-1} - x^*| \leq k^2\,|x_{n-2} - x^*| \leq \ldots \leq k^n\,|x_0 - x^*|.$$

Из услова $0 < k < 1$ следи да је

$$\lim_{n\to\infty} |x_n - x^*| \leq \lim_{n\to\infty} k^n\,|x_0 - x^*| = 0,$$

тј. $\lim\limits_{n\to\infty} x_n = x^*$. □

С обзиром на то да формулација проблема $f(x) = 0$ у облик $g(x) = x$ није јединствена, заправо говоримо о фамилији метода. *Метода просте итерације* дефинисана итеративним поступком

$$x_n = g(x_{n-1}), \quad (n \in \mathbb{N}), \qquad x_0 \in [a, b] \tag{2.4}$$

обухвата и одабир одговарајуће функције g за интервал $[a, b]$.

Може се показати да важи следећа последица претходне теореме:

Последица 2.1 *Ако функција g испуњава услове Теореме 2.3, тада важе следеће оцене за грешку методе:*

$$|x_n - x^*| \leq k^n \max\{x_0 - a, b - x_0\}$$

и

$$|x_n - x^*| \leq \frac{k^n}{1-k}|x_1 - x_0|, \quad (\forall x \in (a, b))\,.$$

Уколико је потребно да израчунамо приближну вредност решења постављене једначине са тачношћу ε, можемо одредити потребан број итерација n из оцене:

$$\frac{k^n}{1-k}|x_1 - x_0| \leq \varepsilon. \tag{2.5}$$

Алтернативно, можемо зауставити итеративни процес када се поклопе вредности две суседне итерације.

2. Методом просте итерације одредити сва решења једначине

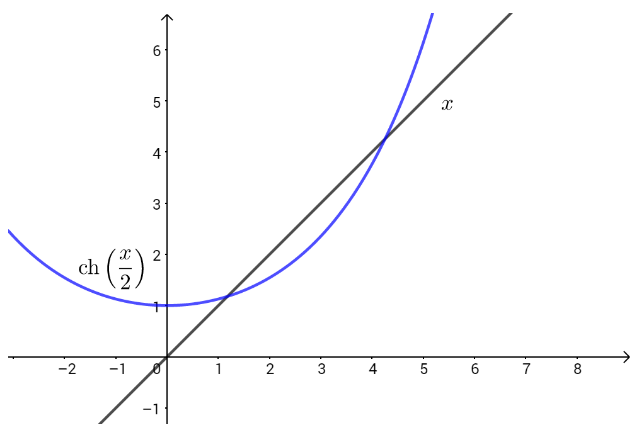
$$f(x) = \mathrm{ch}\left(\frac{x}{2}\right) - x = 0$$

са тачношћу $\varepsilon = 5 \cdot 10^{-3}$.

Решење. Запишимо једначину у еквивалентном облику

$$\mathrm{ch}\left(\frac{x}{2}\right) = x,$$

при чему је $g(x) = \mathrm{ch}\left(\frac{x}{2}\right)$. Скицирањем графика функција $g(x)$ и x можемо проценити да једначина има 2 решења, која припадају интервалима $[1, 2]$ и $[4, 5]$.



Слика 2.2: Локализација решења једначине $\mathrm{ch}\left(\frac{x}{2}\right) - x = 0$

Најпре одредимо приближну вредност $\bar{x}_1 \in [1, 2]$. Формирамо итеративну методу

$$x_{n+1} = \mathrm{ch}\,(0,5\,x_n).$$

Важи да је $1 \leq g(x) \leq 2$ за свако $x \in [1,2]$.

За $x \in [1,2]$, важи да је $|g'(x)| = |0,5\,\mathrm{sh}\left(\frac{x}{2}\right)| \leq 0,5876 < 1$, имајући у виду да је ово растућа функција, па на основу Теореме 2.3 можемо применити ову итеративну методу за одређивање решења $\bar{x}_1$, са $x_0 = 1$. За критеријум заустављања можемо применити оцену (2.5), за $k = 0,6$, тј.

$$\frac{0,6^n}{1-0,6}\,|1,1276 - 1| \leq 5 \cdot 10^{-3}.$$

За $n = 9$ испуњена је ова оцена, па стајемо после 9 итерација.

n	$x_n(= g(x_{n-1}))$
0	1
1	$1,1276$
2	$1,1632$
3	$1,1740$
4	$1,1773$
5	$1,1783$
6	$1,1786$
7	$1,1787$
8	$1,1788$
9	$1,1788$

С обзиром на постигнуту тачност, прве 3 децимале су сигурне, па је $\bar{x}_1 = 1,179$. Иако смо унапред одредили број итерација, видимо да су се вредности итерација поклопиле на 4 децимале у 8. и 9. итерацији.

Одредимо приближну вредност другог решења $\bar{x}_2 \in [4,5]$. Како је вредност $|g'(x)| > 1$ на интервалу $[4,5]$, не можемо искористити исту формулацију проблема како бисмо применили методу просте итерације. Преформулишимо једначину на следећи начин:

$$x = \frac{x^2}{\mathrm{ch}\left(\frac{x}{2}\right)}$$

и означимо са $g_1(x) = \frac{x^2}{\mathrm{ch}\left(\frac{x}{2}\right)}$. Важи да је $4 \leq g_1(x) \leq 5$ за свако $x \in [4,5]$. Може се показати да је $|g_1'(x)|$ опадајућа функција на интервалу $[4,5]$ и да је $|g_1'(x)| < 0,38034$. Формираћемо итеративни процес:

$$x_{n+1} = \frac{x_n^2}{\mathrm{ch}\left(\frac{x_n}{2}\right)}$$

са почетном тачком итерација $x_0 = 4$. Израчунавањем са 4 децимале добијамо:

n	$x_n(= g_1(x_{n-1}))$
0	4
1	$4,2528$
2	$4,2536$
3	$4,2536$

Због поклапања вредности на 4 децимале у 2. и 3. итерацији, а имајући у виду захтевану тачност, стајемо после 3. итерације. Ако применимо критеријум заустављања (2.5), $\frac{0,4^n}{1-0,4}|4,2528-4| \leq 5 \cdot 10^{-3}$, добијамо да је ова оцена испуњена за $n = 5$.

С обзиром на постигнуту тачност, прве 3 децимале су сигурне, па је $\bar{x}_2 = 4,254$.

■

Задаци за вежбу

5. Методом просте итерације са тачношћу 10^{-5}, решити једначину

$$x^3 - 2x - 5 = 0.$$

Резултат. $x \approx 2,09455$

6. Методом просте итерације са тачношћу 10^{-4}, решити једначину

$$x + 2\sin \pi x = 0.$$

Резултат. $x_1 \approx 1,2060, \quad x_2 \approx 1,6820$

7. Методом просте итерације са тачношћу 10^{-4}, решити једначину

$$e^x = 3x^2.$$

Резултат. $x_1 \approx -0,4590, \quad x_2 \approx 0,9100, \quad x_3 \approx 3,7330$

2.4 Њутнова метода

Решавамо једначину $f(x) = 0$. Нека је решење лоцирано на интервалу $[a, b]$, при чему је функција f два пута непрекидно-диференцијабилна на интервалу $[a, b]$, у ознаци $f \in C^2[a, b]$ и нека је $x_0 \in [a, b]$ почетна апроксимација решења. На основу Тејлоровог развоја функције за тачку x у околини x_0 можемо да запишемо:

$$f(x) = f(x_0) + f'(x_0)\,(x - x_0) + \frac{f''(\xi)}{2!}\,(x - x_0)^2,$$

за неко ξ између x и x_0. Узмимо да је $x = x^*$ за које је $f(x^*) = 0$. Ако би функција f била линеарна, важило би да је други извод свугде 0, па би решење једначине у том случају било $0 = f(x_0) + f'(x_0)\,(x^* - x_0)$, одакле бисмо добили да је $x^* = x_0 - f(x_0)/f'(x_0)$. За нелинеарну функцију f узећемо да је следећа апроксимација решења дата истом формулом, тј.

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

Узимајући за нову апроксимацију решења тачку x_1, понављамо поступак и на овај начин добијамо итеративни процес:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \tag{2.6}$$

Овај израз представља *Њутнову методу*, која се још назива и *метода тангенте*, због њене геометријске интерпретације.

Геометријска интерпретација Њутнове методе

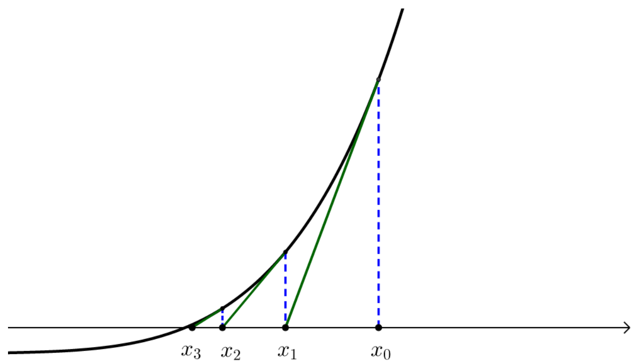
Поставимо у тачки $(x_n, f(x_n))$ тангенту t на график функције $f(x)$, чија једначина гласи:

$$t: \quad f_1(x) - f(x_n) = f'(x_n)\,(x - x_n).$$

Ако сада решавамо једначину $f_1(x) = 0$ и означимо њено решење са x_{n+1}, добијамо да је $f_1(x_{n+1}) - f(x_n) = f'(x_n)\,(x_{n+1} - x_n)$, а како је $f_1(x_{n+1}) = 0$, имамо да важи

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Заправо, тачка x_{n+1} представља пресечну тачку тангенте t и апсцисе.



Слика 2.3: Геометријска интерпретација Њутнове методе

Услови конвергенције Њутнове методе

Наредна теорема даје одговор на питања под којим условима Њутнова метода конвергира, на који начин бирамо почетну тачку итерације и која је оцена грешке методе.

Теорема 2.4 (Теорема о конвергенцији Њутнове методе) *Нека је функција $f \in C^2[a, b]$ и нека важи да је:*

1. $f(a)\, f(b) < 0$,
2. $f'(x) \neq 0$ *за свако* $x \in [a, b]$,
3. f'' *не мења знак на интервалу* $[a, b]$.

Уколико за почетну тачку итерације одаберемо $x_0 \in [a, b]$ *тако да је*

$$f(x_0)\, f''(x_0) > 0, \tag{2.7}$$

тада итеративни низ дефинисан Њутновом методом конвергира ка јединственом решењу x^* *једначине* $f(x) = 0$. *При томе важи следећа апсолутна оцена грешке:*

$$|x^* - x_n| \leq \frac{M_2}{2m_1} (x_n - x_{n-1})^2,$$

где је $0 < m_1 \leq |f'(x)|$ *и* $|f''(x)| \leq M_2$ *за свако* $x \in [a, b]$.

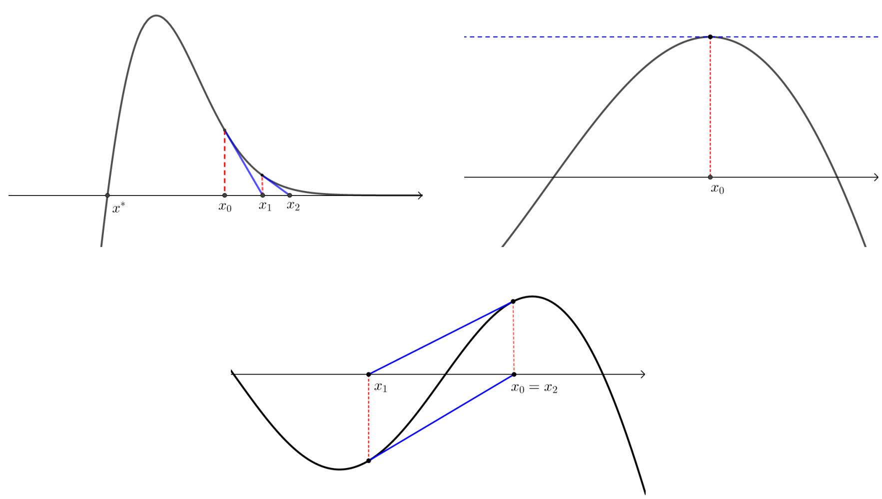
Уколико је потребно да израчунамо приближну вредност решења постављене једначине са тачношћу ε, можемо применити следећи критеријум заустављања итерација

$$\frac{M_2}{2m_1}(x_n - x_{n-1})^2 \leq \varepsilon,$$

односно:

$$|x_n - x_{n-1}| \leq \sqrt{\frac{2m_1\varepsilon}{M_2}}. \tag{2.8}$$

На сликама испод приказани су случајеви када Њутнова метода не конвергира. Приметимо да овај проблем углавном можемо решити погоднијим одабиром почетне тачке итерације x_0.



Модификована Њутнова метода

Можемо заменити $f'(x_n)$ са $f'(x_0)$ у свакој итерацији Њутновог итеративног процеса, тј.

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_0)}.$$

Оваква модификација успорава процес конвергенције.

3. Њутновом методом одредити сва позитивна решења једначине

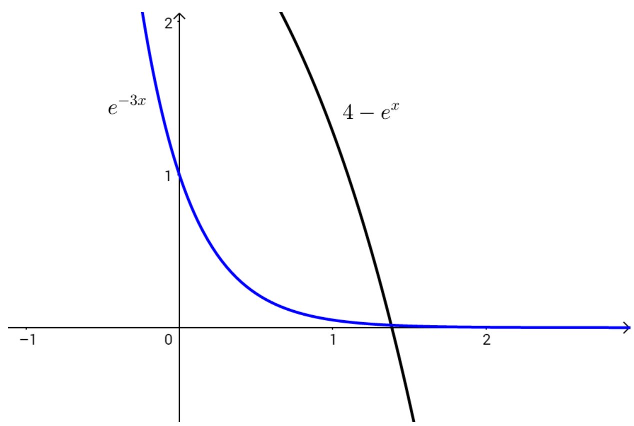
$$f(x) = e^x + e^{-3x} - 4 = 0$$

са тачношћу $\varepsilon = 0.5 \cdot 10^{-4}$.

Решење. Најпре је потребно да локализујемо решења постављене једначине. Записаћемо једначину у облику:

$$4 - e^x = e^{-3x}.$$

Скицирањем графика функција с леве и десне стране једнакости можемо проценити у ком се интервалу налазе позитивна решења.



Слика 2.4: Локализација позитивног решења једначине $e^x + e^{-3x} - 4 = 0$

На основу скице можемо закључити да се једино позитивно решење једначине налази на интервалу $[1, 2]$. Заиста, важи да је

$$f(1)\, f(2) = (-1,23193) \cdot 3,39153 < 0.$$

Проверимо остале услове Теореме 2.4. Први и други извод функције f су:

$$f'(x) = e^x - 3e^{-3x}$$
$$f''(x) = e^x + 9e^{-3x}$$

и оба су позитивна за $x \in [1, 2]$. Први извод је растућа функција, различита од нуле на $[1, 2]$, па важи да је

$$\min_{[1,2]} |f'(x)| = f'(1) = 2,56892 = m_1.$$

За други извод довољно је да одредимо горње ограничење:

$$\max_{x\in[1,2]} |f''(x)| = \max_{x\in[1,2]} |e^x + 9e^{-3x}| \leq e^2 + 9e^{-3} = 7,83714 = M_2.$$

Итеративни процес је дат са:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

За почетну вредност итерације x_0 бирамо тачку $x_0 = 2$ за коју је испуњен услов (2.7):

$$f(2)f''(2) > 0.$$

Критеријум заустављања итерација одређен је са (2.8):

$$|x_n - x_{n-1}| \leq \sqrt{\frac{2m_1\varepsilon}{M_2}} = \sqrt{\frac{2\cdot 2,56892 \cdot 0,5\cdot 10^{-4}}{7,83714}} = 0,00573$$

Формирамо табелу:

n	x_n	$f(x_n)$	$f'(x_n)$
0	2	3,39153	7,38162
1	1,54054	0,67695	4,63760
2	1,39457	0,04848	3,98751
3	1,38241	0,00030	3,93707
4	1,38233		

Критеријум за заустављање итерација је испуњен после 4 итерације:

$$|x_4 - x_3| = |1,38241 - 1,38233| = 0,00008 < 0,00573.$$

Приближно решење полазне једначине је:

$$\bar{x} = 1,38233.$$

■

Задаци за вежбу

8. Њутновом методом са тачношћу 10^{-4}, решити једначину

$$e^x = 2x^2.$$

Резултат. $x_1 \approx -0,5398$; $x_2 \approx 1,4880$; $x_3 \approx 2,6179$

9. Њутновом методом са тачношћу 10^{-4}, одредити највеће негативно решење једначине

$$e^x = \sin\frac{\pi x}{3}.$$

Резултат. $x \approx -3,0454$

10. Њутновом методом са тачношћу 10^{-5}, одредити сва позитивна решења једначине

$$\ln x + x^2 - 3 = 0.$$

Резултат. $x \approx 1,59214$

11. Њутновом методом са тачношћу 10^{-5}, одредити сва решења једначине

$$\cos x = x^2.$$

Резултат. $x_{1,2} \approx \pm 0,82413$

12. Њутновом методом са тачношћу 10^{-4}, одредити сва решења једначине

$$x^2 + \frac{1}{x+1} - 3x = 0.$$

Резултат. $x_1 \approx -1,1987$; $x_2 \approx 0,2865$; $x_3 \approx 2,9122$

2.5 Метода сечице

Њутнова метода захтева израчунавање вредности првог извода функције, што у неким случајевима може значајно отежати израчунавање. Једна од модификација Њутнове методе којом се избегава израчунавање извода функције је *метода сечице*.

На основу Тејлоровог развоја функције у околини тачке x_n, имамо да је:

$$f(x_{n-1}) = f(x_n) + f'(x_n)(x_{n-1} - x_n) + o((x_{n-1} - x_n)^2),$$

па важи апроксимација првог извода:

$$f'(x_n) \approx \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}.$$

Заменом у Њутнову методу добијамо $x_{n+1} = x_n - \frac{f(x_n)}{\frac{f(x_n)-f(x_{n-1})}{x_n - x_{n-1}}}$, тј.

$$x_{n+1} = x_n - f(x_n)\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}. \tag{2.9}$$

Геометријска интерпретација

Сечица кроз тачке $(x_{n-1}, f(x_{n-1}))$ и $(x_n, f(x_n))$ је облика:

$$f_1(x) - f(x_n) = \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}\,(x - x_n).$$

Ако је $f_1(x) = 0$ за неко x, добијамо да је:

$$x_{n+1} = x_n - f(x_n)\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}.$$

Услови конвергенције методе сечице

Нека је решење једначине $f(x) = 0$ лоцирано на интервалу $[a, b]$, на коме f' и f'' не мењају знак и важи да је $f(a)\,f(b) < 0$. Тада можемо применити методу сечице. За почетну апроксимацију $x_0 \in [a, b]$ можемо одабрати тачку за коју је испуњен услов $f(x_0)\,f''(x_0) > 0$, а за $x_1 \in [a, b]$ тачку за коју важи да је $f(x_0)\,f(x_1) < 0$. Најчешће за почетне тачке апроксимације x_0 и x_1 можемо узети крајње тачке интервала $[a, b]$.

Критеријум за заустављање итеративног процеса (2.9) је

$$|x_n - x_{n-1}| \le \frac{m_1\,\varepsilon}{M_1 - m_1}, \tag{2.10}$$

при чему је $m_1 = \min_{x\in[a,b]} |f'(x)|$, $M_1 = \max_{x\in[a,b]} |f'(x)|$.

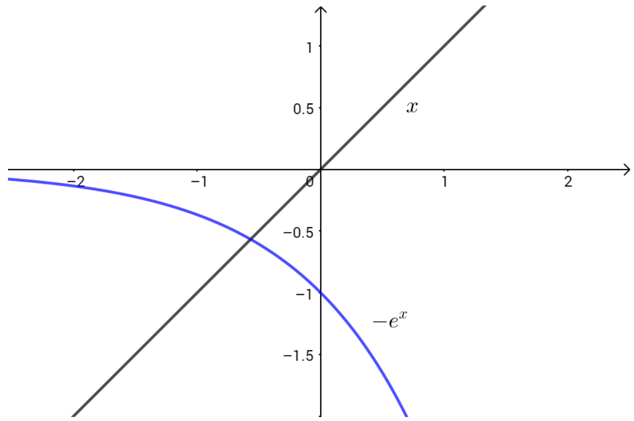
Иако су испуњени сви наведени услови, може се десити да метода сечице не конвергира. Само у „довољно малој околини” нуле x^* функције f гарантована је конвергенција ка решењу. Пре примене Њутнове методе и методе сечице можемо применити методу половљења интервала, како бисмо смањили дужину интервала на коме лоцирамо решење и тиме повећали ефикасност ових метода.

4. Методом сечице решити једначину

$$f(x) = x + e^x = 0$$

са тачношћу $\varepsilon = 0.5 \cdot 10^{-4}$.

Решење. Скицирањем графика функција x и $-e^x$ лоцирамо решење на интервалу $[-1, 0]$.



Слика 2.5: Локализација решења једначине $x + e^x = 0$

Заиста, на крајевима интервала функција има вредности различитог знака, $f(-1)\, f(0) = -0,63212 \cdot 1 < 0$. Први и други извод функције $f(x)$ су:

$$f'(x) = 1 + e^x,$$
$$f''(x) = e^x.$$

Обе функције $f'(x)$ и $f''(x)$ су позитивне на интервалу $[-1, 0]$ и не мењају знак. За тачку 0 важи да је $f(0)\, f''(0) > 0$, па примењујемо методу сечице:

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

са почетним апроксимацијама:

$$x_0 = 0, \quad x_1 = -1, \quad f(0) = 1, \quad f(-1) = -0,63212$$

Критеријум заустављања је (2.10), а како је $f''(x) > 0$, f' је растућа функција и достиже минималну и максималну вредност у крајевима интервала:

$$m_1 = \min_{[-1,0]} |f'(x)| = f'(-1) = 1,36788,$$
$$M_1 = \max_{[-1,0]} |f'(x)| = f'(0) = 2$$

Сада је:

$$|x_{n+1}-x_n| \leq \frac{m_1\,\varepsilon}{M_1-m_1} = \frac{1,36788\cdot\ 0.5\cdot 10^{-4}}{2-1,36788} = 0,00011$$

Рачунајући итерације према (2.9) формирамо табелу:

n	x_n	$f(x_n)$
0	0	1
1	-1	$-0,63212$
2	$-0,61270$	$-0,07081$
3	$-0,56384$	$0,00518$
4	$-0,56717$	$-0,00004$
5	$-0,56714$	

Како је критеријум заустављања испуњен:

$$|x_5-x_4| = 0,00003 < 0,00011$$

приближно решење једначине које испуњава тражену тачност је

$$\bar{x} = -0,56714.$$

■

Задаци за вежбу

13. Методом сечице са тачношћу 10^{-4}, решити једначину

$$e^x = 2x^2.$$

Резултат. $x_1 \approx -0,5398$; $x_2 \approx 1,4880$; $x_3 \approx 2,6179$

14. Методом сечице са тачношћу 10^{-4}, одредити највеће негативно решење једначине

$$e^x = \sin\frac{\pi x}{3}.$$

Резултат. $x \approx -3,0454$

15. Методом сечице са тачношћу 10^{-4}, одредити сва позитивна решења једначине

$$\sin x = x^3 + 0,1.$$

Резултат. $x_1 \approx 0,1012$; $x_2 \approx 0,8736$

Глава 3

Системи линеарних једначина

Решавамо систем од n линеарних једначина са n непознатих:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= b_n \end{aligned} \tag{3.1}$$

где су коефицијенти a_{ij} и константе b_j познате, а x_i непознате. Овај систем једначина можемо записати у матричном облику:

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

тј. $\mathbf{A}\mathbf{x} = \mathbf{b}$, при чему је $\mathbf{A}$ матрица димензије $n \times n$, чији су елементи коефицијенти $[a_{ij}]_{i,j=1}^{n}$, $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$ колона-матрица непознатих и $\mathbf{b} = [b_1, b_2, \ldots, b_n]^T$ слободна колона-матрица. Посебно ће нам користити запис у облику проширене матрице коефицијената:

$$[\mathbf{A}, \mathbf{b}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} & b_n \end{array}\right]$$

Постоје две класе метода за решавање система линеарних једначина, директне и итеративне методе. *Директне методе* карактерише примена трансформација полазног система једначина у еквивалентан систем који има исти скуп решења, а може се директним израчунавањем решити на једноставнији начин. *Итеративне методе* базирају се на погађању решења система једначина и итеративном побољшању решења све док није испуњен критеријум заустављања итеративног процеса. Примена итеративних метода захтева да буду испуњени одређени услови конвергенције итеративног процеса.

Пре него што се посветимо нумеричким методама, уводимо појам норме, који користимо при процени грешке израчунавања и провери услова конвергенције.

3.1 Норма матрица

Дефиниција 3.1 *Норма колона-матрица на скупу $\mathbb{R}^n$ (скуп свих реалних n-димензионалних колона-матрица, $\mathbf{x} = [x_0, \ldots, x_n]^T$) је функција $\|\cdot\| : \mathbb{R}^n \mapsto [0, +\infty)$ која испуњава следеће особине, за свако $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ и $a \in \mathbb{R}$:*

(i) $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0} = [0, \ldots, 0]^T$

(ii) $\|a\mathbf{x}\| = |a|\,\|\mathbf{x}\|$

(iii) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$

Вредност функције $\|\cdot\|$ на колона-матрици $\mathbf{x}$ називамо *нормом колона-матрице* $\|\mathbf{x}\|$. Неки примери норми колона-матрица су:

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i|, \qquad (l_1)$$

$$\|\mathbf{x}\|_2 = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} \qquad (l_2)$$

$$\|\mathbf{x}\|_\infty = \max_{i=1,\ldots,n} |x_i|, \qquad (l_\infty)$$

Дефиниција 3.2 *Норма матрице на скупу* $\mathbb{R}^{n\times n}$ *(скуп свих реалних матрица димензије* $n \times n$*) је функција* $\|\cdot\| : \mathbb{R}^{n\times n} \mapsto [0, +\infty)$ *која испуњава следеће особине, за свако* $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ *и* $a \in \mathbb{R}$*:*

(i) $\|\mathbf{A}\| = 0 \iff \mathbf{A} = \mathbf{0}$

(ii) $\|a\,\mathbf{A}\| = |a|\,\|\mathbf{A}\|$

(iii) $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$

(iv) $\|\mathbf{A} \cdot \mathbf{B}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{B}\|$

Ако су $a_{i,j}$ елементи матрице $\mathbf{A}$ и $b_{i,j}$ елементи матрице $\mathbf{B}$, важе следеће оцене:

(v) $|a_{i,j}| \leq \|\mathbf{A}\|$ за свако $i, j \in \{1, \ldots, n\}$

(vi) $|a_{i,j}| \leq |b_{i,j}|$ за свако $i, j \in \{1, \ldots, n\} \Longrightarrow \|\mathbf{A}\| \leq \|\mathbf{B}\|$

Означимо са $\mathbf{a}_j$ j-ту колону, а са $\mathbf{a}_i$ i-ту врсту матрице $\mathbf{A}$. Неки примери норми матрица су:

$$\|\mathbf{A}\|_1 = \max_{j=1,\ldots,n} \|\mathbf{a}_j\|_1 = \max_{j=1,\ldots,n} \sum_{i=1}^{n} |a_{i,j}|,$$

$$\|\mathbf{A}\|_\infty = \max_{i=1,\ldots,n} \|\mathbf{a}_i^T\|_1 = \max_{i=1,\ldots,n} \sum_{j=1}^{n} |a_{i,j}|,$$

$$\|\mathbf{A}\|_F = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2 \right)^{1/2}$$

Норма колона-матрице или матрице нам даје информацију о њиховој величини. Аналогно појму апсолутне и релативне грешке бројева, уводимо одговарајуће појмове за матрицу. Означимо са $\bar{\mathbf{x}}$ апроксимацију колона-матрице $\mathbf{x}$, а са $\bar{\mathbf{A}}$ апроксимацију матрице $\mathbf{A}$, тада је:

Апсолутна грешка апроксимације:

$$\|\bar{\mathbf{x}} - \mathbf{x}\|, \qquad \|\bar{\mathbf{A}} - \mathbf{A}\|$$

Релативна грешка апроксимације:

$$\frac{\|\bar{\mathbf{x}} - \mathbf{x}\|}{\|\mathbf{x}\|}, \qquad \frac{\|\bar{\mathbf{A}} - \mathbf{A}\|}{\|\mathbf{A}\|}$$

3.2 Директне методе

Директне методе за решавање система линеарних једначина базирају се на примени елементарних трансформација, којима се полазни систем $\mathbf{Ax} = \mathbf{b}$ трансформише у њему еквивалентан систем у облику из кога можемо директно одредити решења система.

Систем линеарних једначина $\mathbf{Ux} = \mathbf{b}$, где је $\mathbf{U}$ горње троугаона матрица (сви елементи испод главне дијагонале су 0):

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ 0 & u_{22} & \dots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & u_{nn} \end{bmatrix}, \quad u_{ii} \neq 0, i = 1, \dots, n$$

решава се поступком *решавања уназад*:

$$x_n = \frac{b_n}{u_{nn}}, \qquad x_i = \frac{1}{u_{ii}} \left(b_i - \sum_{k=i+1}^{n} u_{ik} x_k \right), \quad i = n-1, \dots, 1. \tag{3.2}$$

Слично, систем линеарних једначина $\mathbf{Lx} = \mathbf{b}$, где је $\mathbf{L}$ доње троугаона матрица (сви елементи изнад главне дијагонале су 0):

$$\mathbf{L} = \begin{bmatrix} l_{11} & 0 & \dots & 0 \\ l_{21} & l_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \dots & l_{nn} \end{bmatrix}, \quad l_{ii} \neq 0, i = 1, \dots, n$$

решава се поступком *решавања унапред*:

$$x_1 = \frac{b_1}{l_{11}}, \qquad x_i = \frac{1}{l_{ii}} \left(b_i - \sum_{k=1}^{i-1} l_{ik} x_k \right), \quad i = 2, \dots, n. \tag{3.3}$$

За неколико директних метода наводимо облик еквивалентног система на који сводимо полазни систем једначина:

- Гаусова метода елиминације $\rightarrow \mathbf{Ux} = \mathbf{c}$,
- Гаус-Жорданова метода елиминације $\rightarrow \mathbf{Ix} = \mathbf{c}$,
- LU декомпозиција $\rightarrow \mathbf{LUx} = \mathbf{b}$,

где је $\mathbf{I}$ јединична матрица. Након свођења полазног система на један од ових облика, примењујемо поступак решавања унапред или уназад.

3.3 Гаусова метода елиминације

Решавамо систем линеарних једначина $\mathbf{Ax} = \mathbf{b}$, задат проширеном матрицом:

$$[\mathbf{A}, \mathbf{b}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & b_n \end{array}\right].$$

Претпоставимо да је $a_{11} \neq 0$. Најпре трансформишемо проширену матрицу тако да се сви елементи у првој колони, испод a_{11}, анулирају након трансформације. Коефицијентом $m_{i1} = \dfrac{a_{i1}}{a_{11}}$ множимо прву врсту и одузимамо од i-те врсте проширене матрице $[\mathbf{A}, \mathbf{b}]$, тј.

$$a_{ik}^{(2)} = a_{ik} - m_{i1}\, a_{1k}, \quad b_i^{(2)} = b_i - m_{i1}\, b_1, \quad (i, k = 2, \dots, n).$$

На овај начин добијамо еквивалентан систем:

$$[\mathbf{A}^{(2)}, \mathbf{b}^{(2)}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ 0 & a_{22}^{(2)} & \dots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{n2}^{(2)} & \dots & a_{nn}^{(2)} & b_n^{(2)} \end{array}\right]$$

Примењујемо исти поступак, сада на матрици димензије $n-1$, од друге до n-те једначине. Претпоставимо да је $a_{22}^{(2)} \neq 0$. Коефицијентом $m_{i2} = \dfrac{a_{i2}^{(2)}}{a_{22}^{(2)}}$ множимо другу врсту и одузимамо од i-те врсте проширене матрице $[\mathbf{A}^{(2)}, \mathbf{b}^{(2)}]$, тј.

$$a_{ik}^{(3)} = a_{ik}^{(2)} - m_{i2}\, a_{2k}^{(2)}, \quad b_i^{(3)} = b_i^{(2)} - m_{i2}\, b_2^{(2)}, \quad (i, k = 3, \dots, n).$$

Сада добијамо еквивалентан систем:

$$[\mathbf{A}^{(3)}, \mathbf{b}^{(3)}] = \left[\begin{array}{ccccc|c} a_{11} & a_{12} & a_{13} & \dots & a_{1n} & b_1 \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \dots & a_{2n}^{(2)} & b_2^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \dots & a_{3n}^{(3)} & b_3^{(3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & a_{n3}^{(3)} & \dots & a_{nn}^{(3)} & b_n^{(3)} \end{array}\right]$$

Понављањем поступка полазни систем сводимо на:

$$[\mathbf{A}^{(n)}, \mathbf{b}^{(n)}] = \left[\begin{array}{cccccc|c} a_{11} & a_{12} & a_{13} & \ldots & a_{1,n-1} & a_{1n} & b_1 \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \ldots & a_{2,n-1}^{(2)} & a_{2n}^{(2)} & b_2^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \ldots & a_{3,n-1}^{(3)} & a_{3n}^{(3)} & b_3^{(3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & a_{n-1,n-1}^{(n-1)} & a_{n-1,n}^{(n-1)} & b_n^{(n-1)} \\ 0 & 0 & 0 & \ldots & 0 & a_{nn}^{(n)} & b_n^{(n)} \end{array}\right]$$

На овај начин смо полазни систем једначина свели на њему еквивалентан горње троугаони систем, који се директно решава поступком решавања уназад (3.2).

Коефицијенте m_{ij} које користимо при трансформацији полазног система називамо *мултипликаторима*.

Гаусову методу елиминације можемо описати следећим алгоритмом, при чему проверавамо да ли је испуњено $a_{ii}^{(i)} \neq 0, i = 1, \ldots, n$.

КОРАК 1: Свођење полазне матрице $\mathbf{A}$ система на горње троугаону матрицу. Постављамо $i = 1$.

КОРАК 1.1: Ако је $a_{11} = 0$, враћамо информацију да метода не може бити примењена. СТОП. Ако је $a_{11} \neq 0$, прелазимо на Корак 1.2.

КОРАК 1.2: Коефицијентом $m_{i1} = \dfrac{a_{i1}}{a_{11}}$ множимо прву врсту и одузимамо од i-те врсте проширене матрице.

КОРАК 1.3: Повећавамо i за 1, враћамо се на корак 1.1 за систем димензије $n - i$ до свођења на горње троугаону матрицу.

КОРАК 2: Решавамо горње троугаону матрицу $[\mathbf{A}^{(n)}, \mathbf{b}^{(n)}]$

$$x_n = \frac{b_n^{(n)}}{a_{nn}^{(n)}} \qquad x_i = \frac{1}{a_{ii}^{(i)}} \left(b_i^{(i)} - \sum_{k=i+1}^{n} a_{ik}^{(i)} x_k \right), \quad (i = n-1, \ldots, 1)$$

5. Гаусовом методом елиминације решити систем једначина, рачунајући са 6 децимала.

$$\begin{aligned} 2x_1 + 3x_2 + 4x_3 - 2x_4 &= 1 \\ x_1 - 2x_2 + 4x_3 - 3x_4 &= 2 \\ 4x_1 + 3x_2 - x_3 + x_4 &= 2 \\ 3x_1 - 4x_2 + 2x_3 - 2x_4 &= 5 \end{aligned} \tag{3.4}$$

Решење. Формирамо проширену матрицу постављеног система једначина:

$$[\mathbf{A}, \mathbf{b}] = \left[\begin{array}{rrrr|r} 2 & 3 & 4 & -2 & 1 \\ 1 & -2 & 4 & -3 & 2 \\ 4 & 3 & -1 & 1 & 2 \\ 3 & -4 & 2 & -2 & 5 \end{array}\right]$$

Први корак је да елиминишемо непознату x_1 из прве колоне у 2, 3. и 4. врсти. Формирамо мултипликаторе $m_{21} = \frac{1}{2}, \quad m_{31} = \frac{4}{2}, \quad m_{41} = \frac{3}{2}$. Множимо прву врсту одговарајућим мултипликатором и одузимамо од i-те врсте $(i = 2, 3, 4)$ проширене матрице $[\mathbf{A}, \mathbf{b}]$, добијамо еквивалентан систем:

$$[\mathbf{A}^{(2)}, \mathbf{b}^{(2)}] = \left[\begin{array}{cccc|c} 2 & 3 & 4 & -2 & 1 \\ 0 & -3,5 & 2 & -2 & 1,5 \\ 0 & -3 & -9 & 5 & 0 \\ 0 & -8,5 & -4 & 1 & 3,5 \end{array}\right]$$

Сада су мултипликатори $m_{32} = \frac{-3}{-3,5} = \frac{6}{7}, \quad m_{42} = \frac{-8,5}{-3,5} = \frac{17}{7}$. Множимо другу врсту одговарајућим мултипликатором и одузимамо од i-те врсте $(i = 3, 4)$ проширене матрице $[\mathbf{A}^{(2)}, \mathbf{b}^{(2)}]$, добијамо еквивалентан систем:

$$[\mathbf{A}^{(3)}, \mathbf{b}^{(3)}] = \left[\begin{array}{cccc|c} 2 & 3 & 4 & -2 & 1 \\ 0 & -3,5 & 2 & -2 & 1,5 \\ 0 & 0 & -10,714286 & 6,714286 & -1,285714 \\ 0 & 0 & -8,857143 & 5,857143 & -0,142857 \end{array}\right]$$

Последњи мултипликатор који формирамо је $m_{43} = \frac{-8,857143}{-10,714286}$. Елиминишимо непознату x_3 у 4. врсти множећи 3. врсту овим мултипликатором и одузимајући од 4. врсте, након чега добијамо:

$$[\mathbf{A}^{(4)}, \mathbf{b}^{(4)}] = \left[\begin{array}{cccc|c} 2 & 3 & 4 & -2 & 1 \\ 0 & -3,5 & 2 & -2 & 1,5 \\ 0 & 0 & -10,714286 & 6,714286 & -1,285714 \\ 0 & 0 & 0 & 0,306667 & 0,920000 \end{array}\right]$$

Сада овај систем можемо решити поступком решавања уназад:

$$\begin{aligned} x_4 &= \frac{0,920000}{0,306667} = 3,000000 \\ x_3 &= \frac{1}{-10,714286}\left(-1,285714 - 6,714286 \cdot x_4\right) = 2 \\ x_2 &= \frac{1}{-3,5}\left(1,5 - 2 \cdot x_3 + 2 \cdot x_4\right) = -1 \\ x_1 &= \frac{1}{2}\left(1 - 3 \cdot x_2 - 4 \cdot x_3 + 2 \cdot x_4\right) = 1 \end{aligned}$$

■

Задаци за вежбу

16. Гаусовом методом елиминације решити систем једначина, рачунајући са 5 децимала.

$$\begin{aligned} 100x_1 - 24x_2 + 48x_3 - 23x_4 &= 39 \\ 5x_1 + 100x_2 - 44x_3 - 31x_4 &= 72 \\ 10x_1 - 3x_2 + 100x_3 + 55x_4 &= 56 \\ -12x_1 + 7x_2 - 11x_3 + 100x_4 &= 47 \end{aligned}$$

Резултат. $x_1 = 0,61549; x_2 = 0,95546; x_3 = 0,24967; x_4 = 0,50444$

17. Гаусовом методом елиминације решити систем једначина, рачунајући са 5 децимала.

$$\begin{aligned} x_1 + 0,42x_2 + 0,54x_3 + 0,66x_4 &= 0,3 \\ 0,42x_1 + x_2 + 0,32x_3 + 0,44x_4 &= 0,5 \\ 0,54x_1 + 0,32x_2 + x_3 + 0,22x_4 &= 0,7 \\ 0,66x_1 + 0,44x_2 + 0,22x_3 + x_4 &= 0,9 \end{aligned}$$

Резултат. $x_1 = -1,25779; x_2 = 0,04349; x_3 = 1,03917; x_4 = 1,48239$

3.4 LU декомпозиција

Матрица $\mathbf{A}$ може се представити у облику производа $\mathbf{A} = \mathbf{LU}$, где је $\mathbf{L}$ доње троугаона, а $\mathbf{U}$ горње троугаона матрица. Како бисмо извели овакву декомпозицију матрице $\mathbf{A}$, размотримо поново Гаусову методу елиминације, сада на линеарном систему димензије 4:

$$[\mathbf{A}, \mathbf{b}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & a_{14} & b_1 \\ a_{21} & a_{22} & a_{23} & a_{24} & b_2 \\ a_{31} & a_{32} & a_{33} & a_{34} & b_3 \\ a_{41} & a_{42} & a_{43} & a_{44} & b_4 \end{array}\right].$$

Формирамо мултипликаторе $m_{i1} = \dfrac{a_{i1}}{a_{11}}, (i = 2, 3, 4)$ које користимо за трансформацију полазног система на следећи начин

$$a_{ik}^{(2)} = a_{ik} - m_{i1}\, a_{1k}, \quad b_i^{(2)} = b_i - m_{i1}\, b_1 \tag{3.5}$$

и добијамо еквивалентан систем

$$[\mathbf{A}^{(2)}, \mathbf{b}^{(2)}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & a_{14} & b_1 \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & a_{24}^{(2)} & b_2^{(2)} \\ 0 & a_{32}^{(2)} & a_{33}^{(2)} & a_{34}^{(2)} & b_3^{(2)} \\ 0 & a_{42}^{(2)} & a_{43}^{(2)} & a_{44}^{(2)} & b_4^{(2)} \end{array}\right].$$

Формирамо мултипликаторе $m_{i2} = \dfrac{a_{i2}^{(2)}}{a_{22}^{(2)}}, (i = 3, 4)$ које користимо за трансформацију претходних система на следећи начин

$$a_{ik}^{(3)} = a_{ik}^{(2)} - m_{i2}\, a_{2k}^{(2)}, \quad b_i^{(3)} = b_i^{(2)} - m_{i2}\, b_2^{(2)} \tag{3.6}$$

и добијамо еквивалентан систем

$$[\mathbf{A}^{(3)}, \mathbf{b}^{(3)}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & a_{14} & b_1 \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & a_{24}^{(2)} & b_2^{(2)} \\ 0 & 0 & a_{33}^{(3)} & a_{34}^{(3)} & b_3^{(3)} \\ 0 & 0 & a_{43}^{(3)} & a_{44}^{(3)} & b_4^{(3)} \end{array}\right].$$

Коначно формирамо и мултипликатор $m_{43} = \dfrac{a_{43}^{(3)}}{a_{33}^{(3)}}$ којим трансформишемо систем на следећи начин:

$$a_{44}^{(4)} = a_{44}^{(3)} - m_{43}\, a_{34}^{(3)}, \quad b_4^{(4)} = b_4^{(3)} - m_{43}\, b_3^{(3)} \tag{3.7}$$

у еквивалентан систем чија је матрица коефицијената горње троугаона

$$[\mathbf{A}^{(4)}, \mathbf{b}^{(4)}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & a_{14} & b_1 \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & a_{24}^{(2)} & b_2^{(2)} \\ 0 & 0 & a_{33}^{(3)} & a_{34}^{(3)} & b_3^{(3)} \\ 0 & 0 & 0 & a_{44}^{(4)} & b_4^{(4)} \end{array}\right].$$

Од мултипликатора m_{ij} формирамо матрицу на следећи начин:

$$\mathbf{M}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -m_{21} & 1 & 0 & 0 \\ -m_{31} & 0 & 1 & 0 \\ -m_{41} & 0 & 0 & 1 \end{bmatrix}.$$

За овако дефинисану матрицу $\mathbf{M}_1$, имајући у виду (3.5), важи да је:

$$\mathbf{A}^{(2)} = \mathbf{M}_1\,\mathbf{A}.$$

Слично, за матрицу:

$$\mathbf{M}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -m_{32} & 1 & 0 \\ 0 & -m_{42} & 0 & 1 \end{bmatrix}$$

имајући у виду (3.6), важи да је:

$$\mathbf{A}^{(3)} = \mathbf{M}_2\,\mathbf{A}^{(2)}.$$

Коначно, за матрицу:

$$\mathbf{M}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -m_{43} & 1 \end{bmatrix}$$

имајући у виду (3.7), важи да је:

$$\mathbf{U} = \mathbf{A}^{(4)} = \mathbf{M}_3\,\mathbf{A}^{(3)} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & a_{24}^{(2)} \\ 0 & 0 & a_{33}^{(3)} & a_{34}^{(3)} \\ 0 & 0 & 0 & a_{44}^{(4)} \end{bmatrix}$$

Другим речима, производ $\mathbf{U} = \mathbf{M}_3\mathbf{M}_2\mathbf{M}_1\mathbf{A}$ је горње троугаона матрица коју смо добили применом Гаусове методе елиминације. Ако означимо $\mathbf{M}_1^{-1}\mathbf{M}_2^{-1}\mathbf{M}_3^{-1} = \mathbf{L}$, тада важи:

$$\mathbf{LU} = \mathbf{M}_1^{-1}\mathbf{M}_2^{-1}\mathbf{M}_3^{-1}\mathbf{M}_3\mathbf{M}_2\mathbf{M}_1\mathbf{A} = \mathbf{A}.$$

Како је инверзна матрица доње троугаоне матрице такође доње троугаона, и производ доње троугаоних матрица је доње троугаона, закључујемо да је и сама матрица $\mathbf{L}$ доње троугаона. Овим смо показали да за матрицу $\mathbf{A}$ постоји $\mathbf{LU}$ декомпозиција.

Још је потребно да одредимо матрицу $\mathbf{L}$. Наиме, за матрице облика $\mathbf{M}_k, (k = 1, 2, 3)$ важи да је њихова инверзна матрица једнака полазној матрици, с тим што су ненула елементи ван дијагонале супротног знака, па је:

$$\mathbf{L} = \mathbf{M}_1^{-1}\mathbf{M}_2^{-1}\mathbf{M}_3^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ m_{21} & 1 & 0 & 0 \\ m_{31} & 0 & 1 & 0 \\ m_{41} & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & m_{32} & 1 & 0 \\ 0 & m_{42} & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & m_{43} & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ m_{21} & 1 & 0 & 0 \\ m_{31} & m_{32} & 1 & 0 \\ m_{41} & m_{42} & m_{43} & 1 \end{bmatrix}$$

Приметимо да на основу (3.5), (3.6) и (3.7), важи да је:

$$\mathbf{b}^{(4)} = \mathbf{M}_3\mathbf{M}_2\mathbf{M}_1\mathbf{b}. \tag{3.8}$$

Овакав поступак може се поновити за произвољну димензију матрице $\mathbf{A}$, па можемо да уопштимо закључак:

Ако применом Гаусове методе елиминације на матрицу $\mathbf{A}$ димензије n добијемо горње троугаону матрицу $\mathbf{U}$, тада је $\mathbf{A} = \mathbf{LU}$, где је $\mathbf{L}$ доње троугаона матрица са јединицама на дијагонали, а вредности испод дијагонале су мултипликатори $m_{ij}, (i > j)$.

Решавање система једначина применом LU декомпозиције

Уколико решавамо систем једначина димензије n облика $\mathbf{Ax} = \mathbf{b}$, примењујемо изведену $\mathbf{LU}$ декомпозицију применом следећег алгоритма:

КОРАК 1:	Крећемо од познате декомпозиције: $\mathbf{A} = \mathbf{LU}$, вршимо замену: $(\mathbf{LU})\mathbf{x} = \mathbf{b}$, тј. $\mathbf{L}(\mathbf{Ux}) = \mathbf{b}$.
КОРАК 2:	Уводимо нову колона-матрицу непознатих $\mathbf{y} = \mathbf{Ux}$.
КОРАК 3:	Решавамо доње троугаони систем $\mathbf{Ly} = \mathbf{b}$, поступком решавања унапред добијамо $\mathbf{y}$.
КОРАК 4:	Решавамо горње троугаони систем $\mathbf{Ux} = \mathbf{y}$, поступком решавања уназад добијамо $\mathbf{x}$.

6. Методом $\mathbf{LU}$ декомпозиције решити системе линеарних једначина $\mathbf{Ax} = \mathbf{b}_k$, $(k = 1, 2)$ где је:

$$\mathbf{A} = \begin{bmatrix} 2 & 3 & 4 & -2 \\ 1 & -2 & 4 & -3 \\ 4 & 3 & -1 & 1 \\ 3 & -4 & 2 & -2 \end{bmatrix}, \qquad \mathbf{b}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \\ 5 \end{bmatrix}, \qquad \mathbf{b}_2 = \begin{bmatrix} 8 \\ 6 \\ 3 \\ 9 \end{bmatrix}.$$

Решење. Приметимо да је $\mathbf{Ax} = \mathbf{b}_1$ систем решен у 5. задатку. Зато презначимо $\mathbf{b} = \mathbf{b}_1$. Из поступка решавања претходног задатка можемо да запишемо $\mathbf{LU}$ декомпозицију полазне матрице коефицијената система:

$$\begin{aligned} \mathbf{A} = \mathbf{LU} &= (\mathbf{M}_1^{-1}\mathbf{M}_2^{-1}\mathbf{M}_3^{-1}) \cdot \mathbf{A}^{(4)} \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0,5 & 1 & 0 & 0 \\ 2 & \frac{6}{7} & 1 & 0 \\ 1,5 & \frac{17}{7} & \frac{8,857143}{10,714286} & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 & 4 & -2 \\ 0 & -3,5 & 2 & -2 \\ 0 & 0 & -10,714286 & 6,714286 \\ 0 & 0 & 0 & 0,306667 \end{bmatrix} \end{aligned}$$

На основу (3.8) важи да је:

$$\mathbf{y} = \mathbf{L}^{-1}\mathbf{b} = (\mathbf{M}_3\mathbf{M}_2\mathbf{M}_1)\mathbf{b} = \mathbf{b}^{(4)}.$$

С обзиром на то да смо у 5. задатку већ извршили Гаусову методу елиминације и одредили колона-матрицу $\mathbf{b}^{(4)} = [1; 1,5; -1,285714; 0,92]^T$, можемо то и да искористимо. Дакле, важи да је:

$$\mathbf{y} = [1; 1,5; -1,285714; 0,92]^T$$

Сада решавамо горње троугаони систем $\mathbf{Ux} = \mathbf{y}$, тј.

$$[\mathbf{U}, \mathbf{y}] = [\mathbf{A}^{(4)}, \mathbf{b}^{(4)}] = \left[\begin{array}{cccc|c} 2 & 3 & 4 & -2 & 1 \\ 0 & -3,5 & 2 & -2 & 1,5 \\ 0 & 0 & -10,714286 & 6,714286 & -1,285714 \\ 0 & 0 & 0 & 0,306667 & 0,920000 \end{array}\right]$$

У 5. задатку одредили смо решење горње троугаоног система $\mathbf{Ux} = \mathbf{y}$, тј. $\mathbf{A}^{(4)}\mathbf{x} = \mathbf{b}^{(4)}$, и то је

$$\mathbf{x}_1 = [3, 2, -1, 1]^T.$$

Решимо сада систем $\mathbf{Ax} = \mathbf{b}_2$. Већ нам је позната $\mathbf{LU}$ декомпозиција матрице $\mathbf{A}$, па ћемо то и да искористимо. Пошто је сада колона-матрица слободног коефицијента промењена, није нам познато решење $\mathbf{y}$ доње троугаоног система $\mathbf{Ly} = \mathbf{b}_2$, тј. записано помоћу проширене матрице система:

$$[\mathbf{L}, \mathbf{b}_2] = \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 8 \\ 0,5 & 1 & 0 & 0 & 6 \\ 2 & \frac{6}{7} & 1 & 0 & 3 \\ 1,5 & \frac{17}{7} & \frac{8,857143}{10,714286} & 1 & 9 \end{array}\right]$$

Поступком решавања унапред имамо да је:

$$\begin{aligned} y_1 &= 8 \\ y_2 &= 6 - 0,5\, y_1 = 2 \\ y_3 &= 3 - 2\, y_1 - \frac{6}{7}\, y_2 = -14,714286 \\ y_4 &= 9 - 1,5\, y_1 - \frac{17}{7}\, y_2 - \frac{8,857143}{10,714286}\, y_3 = 4,306667 \end{aligned}$$

Поступком решавања уназад решавамо горње троугаони систем $\mathbf{Ux} = \mathbf{y}$:

$$[\mathbf{U}, \mathbf{y}] = \left[\begin{array}{cccc|c} 2 & 3 & 4 & -2 & 8 \\ 0 & -3,5 & 2 & -2 & 2 \\ 0 & 0 & -10,714286 & 6,714286 & -14,714286 \\ 0 & 0 & 0 & 0,306667 & 4,306667 \end{array}\right]$$

$$x_4 = \frac{4,306667}{0,306667} = 14,043464$$
$$x_3 = \frac{1}{-10,714286}(-14,714286 - 6,714286\, x_4) = 10,173904$$
$$x_2 = \frac{1}{-3,5}(2 + 2\, x_4 - 2\, x_3) = -2,782606$$
$$x_1 = \frac{1}{2}(8 + 2\, x_4 - 4\, x_3 - 3\, x_2) = 1,869565$$

■

Дулитлов алгоритам

Наводимо Дулитлов алгоритам за **LU** декомпозицију матрице **A**, независно од примене Гаусове методе елиминације. Матрица **L** је доње троугаона са јединицама на дијагонали, а матрица **U** горње троугаона.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ l_{21} & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \dots & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ 0 & u_{22} & \dots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & u_{nn} \end{bmatrix}$$

Множењем матрица **L** и **U** добијамо елементе матрице **A**:

$$a_{1j} = u_{1j}, \quad j = 1, 2, \dots, n$$
$$a_{ij} = \sum_{k=1}^{j} l_{ik} u_{kj}, \quad j < i$$
$$a_{ij} = \sum_{k=1}^{i-1} l_{ik} u_{kj} + u_{ij}, \quad j \geq i$$

Одатле добијамо *Дулитлов алгоритам*:

КОРАК 1.1: $l_{ii} = 1, i = 1, 2, \dots, n$, дијагонала матрице **L**

КОРАК 1.2: $u_{1j} = a_{1j}, j = 1, 2, \dots, n$, прва врста матрице **U**

КОРАК 2: $l_{i1} = \frac{a_{i1}}{u_{11}}, i = 2, \dots, n$, прва колона матрице **L**

КОРАК 3.1: За $i = 2$: $u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj}$, $(j = i, \dots, n)$
(i-та врста матрице **U**)

КОРАК 3.2: За $j = 2$: $l_{ij} = \frac{1}{u_{ii}}\left(a_{ij} - \sum_{k=1}^{j-1} l_{ik}u_{kj}\right)$, $(i = j+1, \dots, n)$ (j-та колона матрице $\mathbf{L}$)

КОРАК 4: Повећавамо i и j за 1, враћамо се на корак 3.1.

Краутов алгоритам

Наводимо Краутов алгоритам за **LU** декомпозицију матрице **A**. За разлику од претходне варијанте **LU** декомпозиције, у овом случају горње троугаона матрица **U** има све јединице на дијагонали, а матрица **L** је опет доње троугаона, али без ограничења за вредности дијагоналних елемената.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & \dots & 0 \\ l_{21} & l_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \dots & l_{nn} \end{bmatrix} \begin{bmatrix} 1 & u_{12} & \dots & u_{1n} \\ 0 & 1 & \dots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

Множењем матрица **L** и **U** добијамо елементе матрице **A**, одакле се добијају коефицијенти матрица **L** и **U**:

КОРАК 1.1: $u_{ii} = 1, i = 1, 2, \dots, n$, дијагонала матрице $\mathbf{U}$

КОРАК 1.2: $l_{i1} = a_{i1}, i = 1, 2, \dots, n$, прва колона матрице $\mathbf{L}$

КОРАК 2: $u_{1j} = \frac{a_{1j}}{l_{11}}, j = 2, \dots, n$, прва врста матрице $\mathbf{U}$

КОРАК 3.1: За $j = 2$: $l_{ij} = a_{ij} - \sum_{k=1}^{j-1} l_{ik}u_{kj}$, $(i = j, \dots, n)$ (j-та колона матрице $\mathbf{L}$)

КОРАК 3.2: За $i = 2$: $u_{ij} = \frac{1}{l_{ii}}\left(a_{ij} - \sum_{k=1}^{i-1} l_{ik}u_{kj}\right)$, $(j = i+1, \dots, n)$ (i-та врста матрице $\mathbf{U}$)

КОРАК 4: Повећавамо j и i за 1, враћамо се на корак 3.1.

7. Методом **LU** декомпозиције решити систем једначина:

$$
\begin{aligned}
3x_1 + x_2 - x_3 + 2x_4 &= 6 \\
-5x_1 + x_2 + 3x_3 - 4x_4 &= -12 \\
2x_1 + x_3 - x_4 &= 1 \\
x_1 - 5x_2 + 3x_3 - 3x_4 &= 3
\end{aligned}
$$

применом Краутовог алгоритма, рачунајући са 6 децимала.

Решење. Представимо полазни систем једначина проширеном матрицом:

$$
[\mathbf{A}, \mathbf{b}] = \left[\begin{array}{rrrr|r} 3 & 1 & -1 & 2 & 6 \\ -5 & 1 & 3 & -4 & -12 \\ 2 & 0 & 1 & -1 & 1 \\ 1 & -5 & 3 & -3 & 3 \end{array}\right].
$$

LU декомпозицију матрице **A** представићемо у јединственој матрици облика (матрица **U** има све јединице на дијагонали):

$$
[\mathbf{L} \backslash \mathbf{U}] = \begin{bmatrix} l_{11} & u_{12} & u_{13} & u_{14} \\ l_{21} & l_{22} & u_{23} & u_{24} \\ l_{31} & l_{32} & l_{33} & u_{24} \\ l_{41} & l_{42} & l_{43} & l_{44} \end{bmatrix}
$$

Одредимо елементе матрица **L** и **U**, наизменично по врстама и колонама:

$$
I. \quad l_{11} = a_{11}, \quad l_{21} = a_{21}, \quad l_{31} = a_{31}, \quad l_{41} = a_{41}
$$

$$
II. \quad u_{ii} = 1
$$

$$
III. \quad u_{12} = \frac{a_{12}}{l_{11}} = 0,333333, \quad u_{13} = \frac{a_{13}}{l_{11}} = -0,333333, \quad u_{14} = \frac{a_{14}}{l_{11}} = 0,666667
$$

$$
IV. \quad l_{22} = a_{22} - l_{21}u_{12} = 2,666665 \quad l_{32} = a_{32} - l_{31}u_{12} = -0,666666
$$

$$
l_{42} = a_{42} - l_{41}u_{12} = -5,333333
$$

$$
V. \quad u_{23} = \frac{1}{l_{22}}(a_{23} - l_{21}u_{13}) = 0,500001 \quad u_{24} = \frac{1}{l_{22}}(a_{24} - l_{21}u_{14}) = -0,250000
$$

$$
VI. \quad l_{33} = a_{33} - l_{31}u_{13} - l_{32}u_{23} = 2,000000
$$

$$
l_{43} = a_{43} - l_{41}u_{13} - l_{42}u_{23} = 6,000005
$$

$$VII. \quad u_{34} = \frac{1}{l_{33}}\left(a_{34} - l_{31}u_{14} - l_{32}u_{24}\right) = -1,250000$$

$$VIII. \quad l_{44} = a_{44} - l_{41}u_{14} - l_{42}u_{24} - l_{43}u_{34} = 2,500000$$

$$[\mathbf{L}\backslash\mathbf{U}] = \begin{bmatrix} 3 & 0,333333 & -0,333333 & 0,666667 \\ -5 & 2,666665 & 0,500001 & -0,250000 \\ 2 & -0,666666 & 2,000000 & -1,250000 \\ 1 & -5,333333 & 6,000000 & 2,500000 \end{bmatrix}$$

Сада можемо да применимо поступак за решавање унапред доње троугаоног система $\mathbf{Ly} = \mathbf{b}$, датог проширеном матрицом

$$[\mathbf{L}, \mathbf{b}] = \left[\begin{array}{rrrr|r} 3 & 0 & 0 & 0 & 6 \\ -5 & 2,666665 & 0 & 0 & -12 \\ 2 & -0,666666 & 2,000000 & 0 & 1 \\ 1 & -5,333333 & 6,000000 & 2,500000 & 3 \end{array}\right]$$

чије је решење колона-матрица $\mathbf{y} = [2; -0,750000; -1,750000; 3,000000]^T$. Решавањем уназад горње троугаоног система $\mathbf{Ux} = \mathbf{y}$, датог проширеном матрицом

$$[\mathbf{U}, \mathbf{y}] = \left[\begin{array}{lrrr|r} 1 & 0,333333 & -0,333333 & 0,666667 & 2 \\ 0 & 1 & 0,500001 & -0,250000 & -0,750000 \\ 0 & 0 & 1 & -1,250000 & -1,750000 \\ 0 & 0 & 0 & 1 & 3,000000 \end{array}\right]$$

добијамо решење полазног система:

$$\mathbf{x} = [0,999999; -1,000002; 2,000000; 3,000000]^T.$$

■

Одређивање инверзне матрице LU декомпозицијом

LU декомпозиција нам омогућава да решавамо читаву фамилију система линеарних једначина са истом матрицом коефицијената система $\mathbf{A}$, а различитом слободном колона-матрицом, $\mathbf{Ax} = \mathbf{b}_k$. Овакав приступ можемо применити при одређивању инверзне матрице $\mathbf{A}^{-1}$.

Илуструјмо ово на примеру матрице димензије 4. Наиме, за инверзну матрицу важи:

$$\mathbf{A}\,\mathbf{A}^{-1} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \mathbf{I}$$

Означимо са $\mathbf{x}_j$ j-ту колону непознате инверзне матрице $\mathbf{A}^{-1}$, а са $\mathbf{e}_j$ j-ту колону јединичне матрице $\mathbf{I}$. Инверзну матрицу можемо одредити решавањем 4 система једначина:

$$\mathbf{A}\mathbf{x}_j = \mathbf{e}_j, \qquad (j = 1, 2, 3, 4).$$

Применом методе **LU** декомпозиције једном ћемо извршити декомпозицију матрице **A**, а затим 4 пута примењујемо поступке решавања унапред и уназад с различитим слободним колона-матрицама. Ако бисмо решавали проблем Гаусовом методом елиминације, морали бисмо да применимо целу методу 4 пута, што је доста захтевније.

Одређивање детерминанте матрице LU декомпозицијом

Како је $\mathbf{A} = \mathbf{L}\mathbf{U}$, при чему су **L** и **U** троугаоне матрице и матрица **L** има све јединице на главној дијагонали, важи да је

$$\det \mathbf{A} = \det \mathbf{L} \cdot \det \mathbf{U} = 1 \cdot \det \mathbf{U} = u_{11} \dots u_{nn}.$$

8. Методом **LU** декомпозиције израчунати детерминанту матрице:

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

Решење. Применимо Дулитлов алгоритам за **LU** декомпозицију матрице **A**.

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 3 \\ 2 & 1 & -1 & 1 \\ 3 & -1 & -1 & 2 \\ -1 & 2 & 3 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & 0 & 3 & 13 \\ 0 & 0 & 0 & -13 \end{bmatrix}$$

Сада је

$$\det \mathbf{A} = \det \mathbf{L} \cdot \det \mathbf{U} = 1 \cdot \det \mathbf{U} = 1 \cdot (-1) \cdot 3 \cdot (-13) = 39.$$

■

Задаци за вежбу

18. Методом **LU** декомпозиције решити систем једначина, рачунајући са 6 децимала.

$$\begin{aligned} 100x_1 - 24x_2 + 48x_3 - 23x_4 &= 39 \\ 5x_1 + 100x_2 - 44x_3 - 31x_4 &= 72 \\ 10x_1 - 3x_2 + 100x_3 + 55x_4 &= 56 \\ -12x_1 + 7x_2 - 11x_3 + 100x_4 &= 47 \end{aligned}$$

Резултат. $x_1 = 0,615488; x_2 = 0,955458; x_3 = 0,249673; x_4 = 0,504441$

19. Методом **LU** декомпозиције решити систем једначина, рачунајући са 6 децимала.

$$\begin{aligned} 8,467x_1 + 5,137x_2 + 3,141x_3 + 2,063x_4 &= 29,912 \\ 5,137x_1 + 6,421x_2 + 2,617x_3 + 2,003x_4 &= 25,058 \\ 3,141x_1 + 2,617x_2 + 4,128x_3 + 1,628x_4 &= 16,557 \\ 2,063x_1 + 2,003x_2 + 1,628x_3 + 3,446x_4 &= 12,690 \end{aligned}$$

Резултат. $x_1 = 1,874134; x_2 = 1,596978; x_3 = 1,141352; x_4 = 1,093092$

20. Методом **LU** декомпозиције, рачунајући са 6 децимала, израчунати детерминанту матрице **A**.

$$\mathbf{A} = \begin{bmatrix} 2 & 2 & 1 & 4 & -1 & 3 \\ 1 & 2 & -1 & 3 & 4 & 3 \\ 2 & 1 & 3 & -1 & 2 & -1 \\ 2 & 3 & -4 & 2 & 1 & 5 \\ 1 & 1 & 1 & 3 & 2 & 2 \\ -1 & -1 & 2 & -1 & 2 & 3 \end{bmatrix}$$

Резултат. $\det(\mathbf{A}) = -538$

3.5 Пивотирање

Услов за примену Гаусове методе елиминације је $a_{kk}^{(k)} \neq 0, (k = 1, \dots, n)$, а то неће увек бити испуњено.

Дефиниција 3.3 *Матрица* $\mathbf{A} = [a_{ij}]_{i,j=1}^{n}$ *се назива* дијагонално доминантном *уколико важи да је*

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

а строго дијагонално доминантном *уколико је неједнакост у претходном изразу строга.*

У случају када је матрица **A** строго дијагонално доминантна, Гаусова метода елиминације се може применити, пошто овај услов обезбеђује да се на дијагонали неће наћи нуле.

Пример 1. Применимо Гаусову методу елиминације на систем линеарних једначина $\mathbf{Ax} = \mathbf{b}$, са проширеном матрицом:

$$[\mathbf{A}, \mathbf{b}] = \left[\begin{array}{cc|c} 0 & 1 & 1 \\ 1 & 1 & 2 \end{array}\right]$$

На основу изложеног алгоритма, први корак би био да формирамо коефицијент $m_{21} = 1/0$, што није могуће. Проверимо шта се дешава ако се мало одмакнемо од 0. Посматрајмо систем:

$$[\mathbf{A}', \mathbf{b}'] = \left[\begin{array}{cc|c} \varepsilon & 1 & 1 \\ 1 & 1 & 2 \end{array}\right]$$

где је $\varepsilon > 0$ произвољан фиксиран број близак нули. Када применимо Гаусову методу елиминације, добијамо:

$$[\mathbf{A}'^{(2)}, \mathbf{b}'^{(2)}] = \left[\begin{array}{cc|c} \varepsilon & 1 & 1 \\ 0 & 1-\varepsilon^{-1} & 2-\varepsilon^{-1} \end{array}\right]$$

Ако је ε јако мало, тада је ε^{-1} јако велико, и у том случају се $1-\varepsilon^{-1}$ и $2-\varepsilon^{-1}$ заокружују на $-\varepsilon^{-1}$. (Ако је $\varepsilon = 10^{-10}$, тада је $\varepsilon^{-1} = 10^{10}$, па су ови бројеви приближно једнаки.) Имајући то у виду, применом поступка решавања уназад, добијамо да је

$$x_2 = \frac{2-\varepsilon^{-1}}{1-\varepsilon^{-1}} \approx \frac{-\varepsilon^{-1}}{-\varepsilon^{-1}} = 1, \qquad x_1 = \varepsilon^{-1}(1-x_2) \approx 0.$$

Уколико заменимо места једначинама у полазном систему:

$$[\mathbf{A}, \mathbf{b}] = \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 & 1 \end{array}\right]$$

можемо да применимо Гаусову методу елиминације, при чему добијамо тачна решења $x_1 = x_2 = 1$. Уопште, ако су коефицијенти које формирамо током алгоритма $m_{ij} = \dfrac{a_{ij}^{(j)}}{a_{jj}^{(j)}}$ такви да је дијагонални елемент много мањи од елемента у бројиоцу, резултат ће имати велику вредност, што уводи велику грешку у израчунавање. Дакле, Гаусова метода елиминације може бити нестабилна метода, у зависности од проблема на који се примењује. Овакве ситуације могу се избећи избором *пивота* или *главног елемента*.

Претходни пример показао је да је у појединим случајевима довољно да заменимо места једначинама у систему да бисмо могли да адекватно применимо Гаусову методу елиминације. У таквим случајевима, проблем решавамо избором пивота - елемента матрице система $\mathbf{A}$ с максималном апсолутном вредношћу. Претпоставимо да смо полазни систем $\mathbf{Ax} = \mathbf{b}$ после $(k-1)$ корака свели на решавање система од $n-k+1$ једначине са $n-k+1$ непознатом.

Стратегија потпуног пивотирања подразумева да од свих елемената у матрици димензије $n-k+1$, од k-те до n-те једначине одаберемо пивот, тј. елемент $a_{rs}^{(k)}$ за који важи:

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

и затим поступно заменимо k-ту и r-ту врсту и k-ту и s-ту колону. Стратегија потпуног пивота захтева велики број операција поређења и замене места врстама и колонама, па није препоручљива, осим у случајевима када је прецизност од изузетног значаја, а не и време потребно за извођења методе.

Најчешће је довољно применити *стратегију парцијалног пивотирања*, која подразумева да пивот тражимо само међу елементима k-те колоне, чиме се избегава замена места колона, али не и врста. У том случају, за пивот бирамо елемент $a_{rk}^{(k)}$ за који важи:

$$\left|a_{rk}^{(k)}\right| = \max_{i=k,\ldots,n} \left|a_{ik}^{(k)}\right|.$$

Наводимо алгоритам за *Гаусову методу елиминације с парцијалним пивотирањем*:

КОРАК 1: Свођење полазне матрице $\mathbf{A}$ система на горње троугаону матрицу. Поставимо $i = 1$.

КОРАК 1.1: Одредити $j \in \{1, \ldots, n\}$ за које је $|a_{1j}|$ максимално.

КОРАК 1.2: Ако је $1 \neq j$, заменити врсте 1 и j у проширеној матрици система.

КОРАК 1.3: Коефицијентом $m_{k1} = \dfrac{a_{k1}}{a_{11}}$ множимо прву врсту и одузимамо од k-те врсте проширене матрице.

КОРАК 1.4: Повећавамо i за 1, враћамо се на корак 1.1 за систем димензије $n - i$ до свођења на горње троугаону матрицу.

КОРАК 2: Решавамо горње троугаону матрицу $[\mathbf{U}, \mathbf{c}] = [\mathbf{A}^{(n)}, \mathbf{b}^{(n)}]$

$$x_i = \tfrac{1}{u_{ii}} \left(b_i - \textstyle\sum_{k=i+1}^{n} u_{ik} x_k\right), \quad (i = n, n-1, \ldots, 1)$$

Током примене Гаусове методе елиминације с избором пивота није неопходно мењати положај врста у проширеној матрици, с тим што у том случају морамо водити рачуна да из исте врсте бирамо пивот само једном и врсте из којих смо бирали пивот више не трансформишемо. Да бисмо могли да применимо такав алгоритам за рачунарско израчунавање, морамо памтити из које смо врсте бирали пивот у ком кораку. То радимо записом у колона-матрици, чија свака врста представља редни број пивота. Ова колона-матрица има иницијалне вредности $p = [1, 2, \dots, n]^T$. Сваки пут када вршимо замену места врстама, та замена се врши и на врстама колона-матрице p. Ову колона-матрицу користимо затим и у поступку решавања уназад матрице која је слична горње троугаоној, да бисмо знали којим редоследом решавамо појединачне једначине.

9. Гаусовом методом елиминације с парцијалним пивотирањем, решити систем једначина:

$$\begin{aligned} 2x_1 - 6x_2 + 4x_3 - 2x_4 &= 8 \\ x_1 - 3x_2 + 4x_3 + 3x_4 &= 6 \\ 4x_1 + 3x_2 - 2x_3 + 3x_4 &= 3 \\ x_1 - 4x_2 + 3x_3 + 3x_4 &= 9 \end{aligned}$$

Решење. Проширена матрица постављеног система је:

$$[\mathbf{A}, \mathbf{b}] = \left[\begin{array}{rrrr|r} 2 & -6 & 4 & -2 & 8 \\ 1 & -3 & 4 & 3 & 6 \\ 4 & 3 & -2 & 3 & 3 \\ 1 & -4 & 3 & 3 & 9 \end{array}\right]$$

За прву колону бирамо пивот - по апсолутној вредности највећи елемент у првој колони, налази се у 3. врсти, па 1. и 3. врста мењају места:

$$[\mathbf{A}', \mathbf{b}'] = \left[\begin{array}{rrrr|r} 4 & 3 & -2 & 3 & 3 \\ 1 & -3 & 4 & 3 & 6 \\ 2 & -6 & 4 & -2 & 8 \\ 1 & -4 & 3 & 3 & 9 \end{array}\right]$$

Прву врсту у којој се сада налази пивот не мењамо, а све остале трансформишемо тако да се елементи у 1. колони анулирају. Мултипликатори којима множимо k-ту врсту, $(k = 2, 3, 4)$, су

$$m_{21} = \frac{a'_{21}}{a'_{11}} = \frac{1}{4}, \quad m_{31} = \frac{a'_{31}}{a'_{11}} = \frac{1}{2}, \quad m_{41} = \frac{a'_{41}}{a'_{11}} = \frac{1}{4}.$$

Сада 2, 3. и 4. врсту проширене матрице множимо одговарајућим мултипликатором и одузимамо од 1. врсте:

$$a_{ij}^{(2)} = a'_{ij} - m_{i1}\, a'_{1j}, \quad b_i^{(2)} = b'_i - m_{i1}\, b'_1$$

и добијамо:

$$[\mathbf{A}^{(2)}, \mathbf{b}^{(2)}] = \left[\begin{array}{rrrr|r} 4 & 3 & -2 & 3 & 3 \\ 0 & -3,75 & 4,5 & 2,25 & 5,25 \\ 0 & -7,5 & 5 & -3,5 & 6,5 \\ 0 & -4,75 & 3,5 & 2,25 & 8,25 \end{array}\right]$$

Понављамо поступак на систему од 3 једначине. Бирамо пивот међу елементима 2. колоне у 2, 3. и 4. врсти, и то је елемент $a_{32}^{(2)} = -7,5$. Мењамо места 2. и 3. врсти:

$$[\mathbf{A'}^{(2)}, \mathbf{b'}^{(2)}] = \left[\begin{array}{rrrr|r} 4 & 3 & -2 & 3 & 3 \\ 0 & -7,5 & 5 & -3,5 & 6,5 \\ 0 & -3,75 & 4,5 & 2,25 & 5,25 \\ 0 & -4,75 & 3,5 & 2,25 & 8,25 \end{array}\right]$$

Другу врсту у којој се налази пивот не мењамо, а 3. и 4. трансформишемо тако да се елементи у 2. колони анулирају. Мултипликатори којима множимо k-ту врсту, $(k = 3, 4)$, су

$$m_{32} = \frac{a'^{(2)}_{32}}{a'^{(2)}_{22}} = \frac{1}{2}, \quad m_{42} = \frac{a'^{(2)}_{42}}{a'^{(2)}_{22}} = \frac{19}{30}.$$

Сада 3. и 4. врсту проширене матрице множимо одговарајућим мултипликатором и одузимамо од 2. врсте, $i = 3, 4$:

$$a_{ij}^{(3)} = a'^{(2)}_{ij} - m_{i2}\, a'^{(2)}_{2j}, \quad b_i^{(3)} = b'^{(2)}_i - m_{i2}\, b'^{(2)}_2$$

и добијамо:

$$[\mathbf{A}^{(3)}, \mathbf{b}^{(3)}] = \left[\begin{array}{rrrr|r} 4 & 3 & -2 & 3 & 3 \\ 0 & -7,5 & 5 & -3,5 & 6,5 \\ 0 & 0 & 2 & 4 & 2 \\ 0 & 0 & \frac{1}{3} & 4\frac{7}{15} & 4\frac{2}{15} \end{array}\right]$$

Понављамо поступак на систему од 2 једначине, 3. и 4. Како је сада пивот у 3. врсти, не мењамо места врстама. Трећу врсту у којој се налази пивот не мењамо, а 4. врсту трансформишемо тако да се елементи у

3. колони анулирају. Мултипликатор којим множимо 4. врсту је $m_{43} = \frac{a_{43}^{(3)}}{a_{33}^{(3)}} = \frac{1}{6}$. Сада 4. врсту проширене матрице множимо овим мултипликатором и одузимамо од 3. врсте:

$$a_{4j}^{(4)} = a_{4j}^{(3)} - m_{43}\, a_{3j}^{(3)}, \quad b_4^{(4)} = b_4^{(3)} - m_{43}\, b_3^{(3)}$$

и добијамо:

$$[\mathbf{A}^{(4)}, \mathbf{b}^{(4)}] = \left[\begin{array}{rrrr|r} 4 & 3 & -2 & 3 & 3 \\ 0 & -7,5 & 5 & -3,5 & 6,5 \\ 0 & 0 & 2 & 4 & 2 \\ 0 & 0 & 0 & 3\frac{4}{5} & 3\frac{4}{5} \end{array}\right]$$

Поступком решавања уназад добијамо решење система једначина:

$$\begin{aligned} x_4 &= \frac{3\frac{4}{5}}{3\frac{4}{5}} = 1 \\ x_3 &= \frac{1}{2}(2 - 4\,x_4) = -1 \\ x_2 &= \frac{1}{-7,5}(6,5 + 3,5\,x_4 - 5\,x_3) = -2 \\ x_1 &= \frac{1}{4}(3 - 3\,x_4 + 2\,x_3 - 3\,x_2) = 1 \end{aligned}$$

тј. $\mathbf{x} = [1, -2, -1, 1]^T$.

■

Уколико примењујемо Гаусову методу елиминације с избором пивота, да бисмо применили **LU** декомпозицију морамо водити рачуна о промени места врстама (и колонама у случају потпуног пивотирања). Тада, уместо декомпозиције полазне матрице **A**, вршимо декомпозицију матрице **PA**, где је **P** матрица пермутација чији су елементи 0 и 1, која врши замену врста матрице **A** приликом множења. На примеру матрице димензије 3, замена 2. и 3. врсте матрице **A** реализује се множењем матрицом пермутација $\mathbf{P} = \left[\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{smallmatrix}\right]$.

У случају да смо применили одабир пивота и вршили замене врстама матрице **A**, то морамо узети у обзир када одређујемо детерминанту матрице **A**. Тада ће важити:

$$\det \mathbf{A} = (-1)^m u_{11} \ldots u_{nn},$$

где је m број замена врста.

Задаци за вежбу

21. Гаусовом методом елиминације с парцијалним пивотирањем решити систем једначина, рачунајући са 6 децимала.

$$\begin{aligned} -x_1 + x_2 \qquad - 3x_4 &= 4 \\ x_1 \qquad + 3x_3 + x_4 &= 0 \\ x_2 - x_3 - x_4 &= 3 \\ 3x_1 \qquad + x_3 + 2x_4 &= 1 \end{aligned}$$

Резултат. $x_1 = 1; x_2 = 2; x_3 = 0; x_4 = -1$

22. Гаусовом методом елиминације с парцијалним пивотирањем решити систем једначина, рачунајући са 6 децимала.

$$\begin{aligned} x_1 + 3x_2 + 2x_3 + x_4 &= 2 \\ 4x_1 + 2x_2 + x_3 + 2x_4 &= 2 \\ 2x_1 + x_2 + 2x_3 + 3x_4 &= 1 \\ x_1 + 2x_2 + 4x_3 + x_4 &= 1 \end{aligned}$$

Резултат. $x_1 = 0,137255; x_2 = 0,686275; x_3 = -0,156863; x_4 = 0,117647$

23. Гаусовом методом елиминације с парцијалним пивотирањем решити систем једначина, рачунајући са 6 децимала.

$$\begin{aligned} 2x_1 - x_2 + 3x_3 + 7x_4 &= 15 \\ 4x_1 + 4x_2 \qquad + 7x_4 &= 11 \\ 2x_1 + x_2 + x_3 + 3x_4 &= 7 \\ 6x_1 + 5x_2 + 4x_3 + 17x_4 &= 31 \end{aligned}$$

Резултат. $x_1 = 1; x_2 = 0; x_3 = 2; x_4 = 1$

3.6 Итеративне методе

Итеративне методе базирају се на одабиру почетне процене решења система линеарних једначина облика $\mathbf{Ax} = \mathbf{b}$ и итеративног побољшања овакве процене.

Представићемо Јакобијеву и Гаус-Зајделову итеративну методу, које се базирају на идеји да из i-те једначине представимо непознату x_i на следећи начин:

$$
\begin{aligned}
x_1 &= \phantom{-\frac{a_{21}}{a_{22}}x_1} -\frac{a_{12}}{a_{11}}x_2 - \frac{a_{13}}{a_{11}}x_3 \dots - \frac{a_{1n}}{a_{11}}x_n + \frac{b_1}{a_{11}} \\
x_2 &= -\frac{a_{21}}{a_{22}}x_1 \phantom{-\frac{a_{12}}{a_{11}}x_2} - \frac{a_{23}}{a_{22}}x_3 \dots - \frac{a_{2n}}{a_{22}}x_n + \frac{b_2}{a_{22}} \\
&\vdots \\
x_n &= -\frac{a_{n1}}{a_{nn}}x_1 - \frac{a_{n2}}{a_{nn}}x_2 - \frac{a_{n3}}{a_{nn}}x_3 \dots \phantom{- \frac{a_{1n}}{a_{11}}x_n} + \frac{b_n}{a_{nn}}
\end{aligned}
\tag{3.9}
$$

под условом $a_{ii} \neq 0, (i = 1, \dots, n)$.

Јакобијева итеративна метода

Јакобијева итеративна метода настаје из облика система (3.9) на следећи начин:

$$
\begin{aligned}
x_1^{(k+1)} &= \phantom{-\frac{a_{21}}{a_{22}}x_1^{(k)}} -\frac{a_{12}}{a_{11}}x_2^{(k)} - \frac{a_{13}}{a_{11}}x_3^{(k)} \dots - \frac{a_{1n}}{a_{11}}x_n^{(k)} + \frac{b_1}{a_{11}} \\
x_2^{(k+1)} &= -\frac{a_{21}}{a_{22}}x_1^{(k)} \phantom{-\frac{a_{12}}{a_{11}}x_2^{(k)}} - \frac{a_{23}}{a_{22}}x_3^{(k)} \dots - \frac{a_{2n}}{a_{22}}x_n^{(k)} + \frac{b_2}{a_{22}} \\
&\vdots \\
x_n^{(k+1)} &= -\frac{a_{n1}}{a_{nn}}x_1^{(k)} - \frac{a_{n2}}{a_{nn}}x_2^{(k)} - \frac{a_{n3}}{a_{nn}}x_3^{(k)} \dots \phantom{- \frac{a_{1n}}{a_{11}}x_n^{(k)}} + \frac{b_n}{a_{nn}}
\end{aligned}
\tag{3.10}
$$

У матричном облику Јакобијеву итеративну методу можемо записати:

$$\mathbf{x}^{(k+1)} = \mathbf{B}\mathbf{x}^{(k)} + \mathbf{c}, \tag{3.11}$$

где је

$$
\mathbf{B} = \begin{bmatrix} 0 & -\frac{a_{12}}{a_{11}} & \dots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & \dots & -\frac{a_{2n}}{a_{22}} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & \dots & 0 \end{bmatrix}, \qquad \mathbf{c} = \begin{bmatrix} \frac{b_1}{a_{11}} \\ \frac{b_2}{a_{22}} \\ \vdots \\ \frac{b_n}{a_{nn}} \end{bmatrix}
$$

При томе је почетна колона-матрица итерације $\mathbf{x}^{(0)}$ произвољно изабрана.

Приметимо да је ово иста метода као метода итерације за решавање нелинеарних једначина из секције 2.3, само сада примењена на систем линеарних једначина. Ово видимо ако запишемо Јакобијеву методу у облику:

$$\mathbf{x}^{(k+1)} = G(\mathbf{x}^{(k)}) = \mathbf{B}\mathbf{x}^{(k)} + \mathbf{c},$$

где је G пресликавање колона-матрица.

Гаус-Зајделова итеративна метода

Јакобијева итеративна метода подразумева да се у $(k+1)$ итерацији израчунавају вредности $x_1^{(k+1)}, \ldots, x_n^{(k+1)}$ користећи претходно израчунате вредности у k-тој итерацији $x_1^{(k)}, \ldots, x_n^{(k)}$. Међутим, посматрајући облик итеративног поступка (3.10), видимо да у $(k+1)$ итерацији најпре одређујемо вредност $x_1^{(k+1)}$, а затим $x_2^{(k+1)}$, користећи процену за прву непознату $x_1^{(k)}$. Овај поступак можемо убрзати коришћењем *последње израчунате процене* за непознату $x_1^{(k+1)}$ уместо процене из претходне итерације $x_1^{(k)}$. Примењујући овај поступак за све променљиве у $(k+1)$ итерацији добијамо *Гаус-Зајделову итеративну методу*:

$$\begin{aligned}
x_1^{(k+1)} &= \phantom{-\frac{a_{21}}{a_{22}}x_1^{(k+1)}} -\frac{a_{12}}{a_{11}}x_2^{(k)} - \frac{a_{13}}{a_{11}}x_3^{(k)} \ldots - \frac{a_{1n}}{a_{11}}x_n^{(k)} + \frac{b_1}{a_{11}} \\
x_2^{(k+1)} &= -\frac{a_{21}}{a_{22}}x_1^{(k+1)} \qquad - \frac{a_{23}}{a_{22}}x_3^{(k)} \ldots - \frac{a_{2n}}{a_{22}}x_n^{(k)} + \frac{b_2}{a_{22}} \\
&\vdots \\
x_n^{(k+1)} &= -\frac{a_{n1}}{a_{nn}}x_1^{(k+1)} - \frac{a_{n2}}{a_{nn}}x_2^{(k+1)} - \frac{a_{n3}}{a_{nn}}x_3^{(k+1)} \ldots \qquad + \frac{b_n}{a_{nn}}
\end{aligned} \tag{3.12}$$

У матричном облику за Гаус-Зајделову итеративну методу важи:

$$\begin{aligned}
\mathbf{x}^{(k+1)} &= \mathbf{B}_1\mathbf{x}^{(k+1)} + \mathbf{B}_2\mathbf{x}^{(k)} + \mathbf{c} \\
(I - \mathbf{B}_1)\mathbf{x}^{(k+1)} &= \mathbf{B}_2\mathbf{x}^{(k)} + \mathbf{c} \\
\mathbf{x}^{(k+1)} &= (I - \mathbf{B}_1)^{-1}\mathbf{B}_2\mathbf{x}^{(k)} + (I - \mathbf{B}_1)^{-1}\mathbf{c}
\end{aligned}$$

где је

$$\mathbf{B}_1 = \begin{bmatrix} 0 & 0 & \dots & 0 \\ -\frac{a_{21}}{a_{22}} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & \dots & 0 \end{bmatrix}$$

$$\mathbf{B}_2 = \begin{bmatrix} 0 & -\frac{a_{12}}{a_{11}} & \dots & -\frac{a_{1n}}{a_{11}} \\ 0 & 0 & \dots & -\frac{a_{2n}}{a_{22}} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}, \qquad \mathbf{c} = \begin{bmatrix} \frac{b_1}{a_{11}} \\ \frac{b_2}{a_{22}} \\ \vdots \\ \frac{b_n}{a_{nn}} \end{bmatrix}$$

При томе је почетна колона-матрица итерације $\mathbf{x}^{(0)}$ произвољно изабрана.

Услови конвергенције и критеријуми заустављања

Подсетимо се услова *строго дијагоналне доминантности*, тј. да је сваки дијагонални елемент матрице по апсолутној вредности већи од збира апсолутних вредности преосталих коефицијената у истој врсти матрице, тј.

$$|a_{ii}| > \sum_{j=1, j\neq i}^{n} |a_{ij}|, \qquad (i = 1, \dots, n), \tag{3.13}$$

Теорема 3.1 *Ако је матрица* $\mathbf{A}$ *коефицијената система* $\mathbf{Ax} = \mathbf{b}$ *стриктно дијагонално доминантна, тада Јакобијева и Гаус-Зајделова метода конвергирају ка решењу система, за произвољно одабрану почетну вредност* $\mathbf{x}^{(0)}$.

Приметимо да је услов строго дијагоналне доминантности само довољан услов, не и потребан. Наиме, постоје матрице које не испуњавају овај услов, а ове методе ипак конвергирају.

Услов строго дијагоналне доминантности обезбеђује да при трансформацији полазног система дијагонални коефицијенти којима делимо, $a_{ii}, (i = 1, \dots n)$, буду различити од 0.

Уколико овај услов није испуњен, потребно је да се елементарним трансформацијама полазни систем трансформише у њему еквивалентан, који испуњава овај услов.

Иако се може чинити да је услов строго дијагоналне доминантности доста рестриктиван, многи примењени проблеми га ипак испуњавају. То су најчешће системи великих димензија са матрицама система које су ретке (имају много већи број нула елемената). Овакви системи често се јављају код анализе кола, или код нумеричког решавања обичних диференцијалних једначина са граничним условима, као и парцијалних диференцијалних једначина.

У општем случају, Гаус-Зајделова метода доста брже конвергира у односу на Јакобијеву методу (под одређеним условима и дупло брже). С друге стране, код решавања система великих димензија (више хиљада) има смисла применити паралелно програмирање за израчунавање решења, а за то је примереније користити Јакобијеву методу.

Ако је захтевана тачност ε, критеријум заустављања дат је у односу на норму разлике две узастопне итерације. Најчешће узимамо l_∞ норму и стајемо када је испуњен услов

$$\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|_\infty = \max_{i=1,\dots,n} |x_i^{(k)} - x_i^{(k-1)}| \le \varepsilon. \tag{3.14}$$

У случају када је тачност $\varepsilon = 0,5\cdot 10^{-m}$, услов (3.14) значи да се вредности две суседне итерације поклапају на m децимала (Пример 3 у Глави 1).

Други критеријум заустављања који можемо користити јесте:

$$\frac{\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|_\infty}{\|\mathbf{x}^{(k)}\|_\infty} < \varepsilon.$$

10. Итеративном методом решити систем једначина

$$\begin{aligned}
1,02x_1 - 0,05x_2 - 0,10x_3 &= 0,795\\
-0,11x_1 + 1,03x_2 - 0,05x_3 &= 0,849\\
-0,11x_1 - 0,12x_2 + 1,04x_3 &= 1,398
\end{aligned}$$

са тачношћу $\varepsilon = 0,5 \cdot 10^{-4}$.

Решење. Задатак ћемо решити Јакобијевом и Гаус-Зајделовом итеративном методом. Матрица коефицијената система једначина је строго дијагонално доминантна, тј. испуњен је услов

$$\begin{aligned}
|1,02| &> |-0,05| + |-0,10|\\
|1,03| &> |-0,11| + |-0,05|\\
|1,04| &> |-0,11| + |-0,12|
\end{aligned}$$

па ће обе итеративне методе конвергирати ка решењу. Трансформишимо полазни систем на следећи начин:

$$
\begin{aligned}
x_1 &= \phantom{\frac{0,11}{1,03}x_1 +{}} \frac{0,05}{1,02}x_2 + \frac{0,10}{1,02}x_3 + \frac{0,795}{1,02}\\
x_2 &= \frac{0,11}{1,03}x_1 \phantom{{}+\frac{0,12}{1,04}x_2} + \frac{0,05}{1,03}x_3 + \frac{0,849}{1,03}\\
x_3 &= \frac{0,11}{1,04}x_1 + \frac{0,12}{1,04}x_2 \phantom{{}+\frac{0,05}{1,03}x_3} + \frac{1,398}{1,04}
\end{aligned}
$$

Из претходног облика система једначина формираћемо итеративни процес за обе методе.

Јакобијева итеративна метода: Итеративни процес формирамо на следећи начин:

$$
\begin{aligned}
x_1^{(k+1)} &= \frac{0,05}{1,02}x_2^{(k)} + \frac{0,10}{1,02}x_3^{(k)} + \frac{0,795}{1,02}\\
x_2^{(k+1)} &= \frac{0,11}{1,03}x_1^{(k)} + \frac{0,05}{1,03}x_3^{(k)} + \frac{0,849}{1,03}\\
x_3^{(k+1)} &= \frac{0,11}{1,04}x_1^{(k)} + \frac{0,12}{1,04}x_2^{(k)} + \frac{1,398}{1,04}
\end{aligned}
$$

са почетном вредношћу $\mathbf{x}^{(0)} = [0,0,0]^T$. С обзиром на задату тачност, стајемо са итерацијама када се вредности 2 суседне итерације поклопе на 4 децимале. Како бисмо умањили утицај грешке заокруживања на резултат, с обзиром на дељење, можемо или да у сваком израчунавању делимо са вредношћу a_{ii}, или да након дељења заокружујемо вредности на мало већи број децимала него што је потребно. Формирамо табелу итерација:

k	x_1	x_2	x_3
0	0	0	0
1	0,7794	0,8243	1,3442
2	0,9516	0,9728	1,5218
3	0,9763	0,9998	1,5571
4	0,9811	1,0041	1,5628
5	0,9818	1,0049	1.5638
6	0,9820	1,0050	1,5640
7	0,9820	1,0051	1,5641
8	0,9820	1,0051	1,5641

Јакобијевом итеративном методом смо после 8 итерација добили решење полазног система једначина $\bar{x}_1 = 0,9820$, $\bar{x}_2 = 1,0051$, $\bar{x}_3 = 1,5641$.

Гаус-Зајделова итеративна метода: Итеративни процес формирамо на следећи начин:

$$
\begin{aligned}
x_1^{(k+1)} &= \qquad\qquad\qquad\qquad \frac{0,05}{1,02}x_2^{(k)} + \frac{0,10}{1,02}x_3^{(k)} + \frac{0,795}{1,02} \\
x_2^{(k+1)} &= \frac{0,11}{1,03}x_1^{(k+1)} \qquad\qquad\quad + \frac{0,05}{1,03}x_3^{(k)} + \frac{0,849}{1,03} \\
x_3^{(k+1)} &= \frac{0,11}{1,04}x_1^{(k+1)} + \frac{0,12}{1,04}x_2^{(k+1)} \qquad\qquad + \frac{1,398}{1,04}
\end{aligned}
$$

са почетном вредношћу $\mathbf{x}^{(0)} = [0,0,0]^T$. У овој методи за израчунавање текуће приближне вредности непознате користимо последње израчунате приближне вредности за сваку непознату, без обзира на то у којој су итерацији одређене. Формирамо табелу итерација:

k	x_1	x_2	x_3
0	0	0	0
1	0,7794	0,9075	1,5314
2	0,9740	1,0026	1,5629
3	0,9818	1,0050	1,5640
4	0,9820	1,0051	1,5641
5	0,9820	1,0051	1,5641

Гаус-Зајделовом итеративном методом смо после 5 итерација добили исто решење као и Јакобијевом методом, $\bar{x}_1 = 0,9820$, $\bar{x}_2 = 1,0051$, $\bar{x}_3 = 1,5641$.

■

11. Гаус-Зајделовом итеративном методом решити систем једначина:

$$
\begin{aligned}
2x_1 + 8x_2 + 3x_3 + x_4 &= -2 \\
2x_2 - x_3 + 4x_4 &= 4 \\
7x_1 - 2x_2 + x_3 + 2x_4 &= 3 \\
-x_1 \qquad + 5x_3 + 2x_4 &= 5
\end{aligned}
$$

са тачношћу $\varepsilon = 0,5 \cdot 10^{-4}$.

Решење. Приметимо да коефицијенти система не испуњавају услов строге дијагоналне доминантности. Да бисмо могли да применимо Гаус-Зајделову методу потребно је да извршимо трансформацију полазног система елементарним трансформацијама у еквивалентан систем који испуњава овај услов. У овом случају биће довољно да променимо места једначинама:

$$\begin{aligned}
7x_1 - 2x_2 + x_3 + 2x_4 &= 3\\
2x_1 + 8x_2 + 3x_3 + x_4 &= -2\\
-x_1 \qquad\quad + 5x_3 + 2x_4 &= 5\\
2x_2 - x_3 + 4x_4 &= 4
\end{aligned}$$

Из i-те једначине изразићемо променљиву x_i и формирамо Гаус-Зајделов итеративни процес:

$$\begin{aligned}
x_1^{(k+1)} &= \qquad\qquad\quad \frac{2}{7}x_2^{(k)} - \frac{1}{7}x_3^{(k)} - \frac{2}{7}x_4^{(k)} + \frac{3}{7}\\
x_2^{(k+1)} &= -\frac{2}{8}x_1^{(k+1)} \qquad\quad - \frac{3}{8}x_3^{(k)} - \frac{1}{8}x_4^{(k)} - \frac{2}{8}\\
x_3^{(k+1)} &= \frac{1}{5}x_1^{(k+1)} + \qquad\qquad\quad - \frac{2}{5}x_4^{(k)} + 1\\
x_4^{(k+1)} &= \qquad\quad -\frac{2}{4}x_2^{(k+1)} + \frac{1}{4}x_3^{(k+1)} \qquad + 1
\end{aligned}$$

са почетном вредношћу $\mathbf{x}^{(0)} = [0,0,0,0]^T$. С обзиром на задату тачност, стајемо са итерацијама када се вредности 2 суседне итерације поклопе на 4 децимале. Формирамо табелу итерација:

k	x_1	x_2	x_3	x_4
0	0	0	0	0
1	$0,4286$	$-0,25$	1	1
2	$-0,0714$	$-0,7322$	$0,5857$	$1,5125$
3	$-0,2964$	$-0,5846$	$0,3357$	$1,3762$
4	$-0,1796$	$-0,5030$	$0,4136$	$1,3549$
5	$-0,1613$	$-0,5341$	$0,4258$	$1,3735$
6	$-0,1773$	$-0,5370$	$0,4151$	$1,3723$
7	$-0,1762$	$-0,5332$	$0,4158$	$1,3706$
8	$-0,1748$	$-0,5336$	$0,4168$	$1,3710$
9	$-0,1751$	$-0,5339$	$0,4166$	$1,3711$
10	$-0,1752$	$-0,5338$	$0,4165$	$1,3710$
11	$-0,1752$	$-0,5338$	$0,4166$	$1,3710$
12	$-0,1752$	$-0,5338$	$0,4166$	$1,3710$

Гаус-Зајделовом итеративном методом после 12 итерација добијамо поклапање вредности на 4 децимале и решење $\bar{x}_1 = -0,1752$, $\bar{x}_2 = -0,5338$, $\bar{x}_3 = 0,4166$, $\bar{x}_4 = 1,3710$.

■

Задаци за вежбу

24. Јакобијевом и Гаус-Зајделовом итеративном методом решити систем једначина, рачунајући са 5 децимала.

$$\begin{aligned}
10x_1 + 1,05x_2 + 1,65x_3 + 1,27x_4 &= 1,75\\
1,05x_1 + \quad 9x_2 + 1,10x_3 + 1,41x_4 &= 2,3\\
1,65x_1 + 1,10x_2 + \quad 7x_3 + 1,6x_4 &= 3\\
1,27x_1 + 1,41x_2 + 1,6x_3 + \quad 6x_4 &= 6
\end{aligned}$$

Резултат. $x_1 = 0,01577; x_2 = 0,08455; x_3 = 0,20052; x_4 = 0,92332$

25. Јакобијевом и Гаус-Зајделовом итеративном методом решити систем једначина, са тачношћу $0,5 \cdot 10^{-4}$.

$$\begin{aligned}
10x_1 + 2x_2 - x_3 &= 27\\
-3x_1 - 6x_2 + 2x_3 &= 61,5\\
x_1 + x_2 + 5x_3 &= 0,7
\end{aligned}$$

Резултат. $x_1 = 0,6536; x_2 = 9,3059; x_3 = -1,8519$

Глава 4

Интерполација

Интерполација представља посебан случај апроксимације функција и основа је за многе друге методе које представљамо у оквиру овог курса. У овом одељку радимо само са функцијама једне променљиве, мада се тема може уопштити и на функције више променљивих.

Пример 1. Претпоставимо да су нам познате вредности функције у две тачке:

x	$0,32$	$0,53$
$f(x)$	$1,27$	$2,04$

Уколико желимо да проценимо вредност ове функције у тачки $0,38$, могли бисмо да претпоставимо да је $f(0,38)$ блиска вредности функције у најближој тачки у којој нам је позната вредност функције, $0,32$, па бисмо узели да је $f(0,38) = 1,27$. Иако је ово решење најједноставније, поставља се питање да ли на основу познатих података можемо да извршимо квалитетнију процену. С обзиром на то да су нам познате 2 тачке кроз које пролази график функције f, разумно би било претпоставити да ће вредност линеарне функције $l(x)$, чији график пролази кроз задате тачке, бити блиска вредности полазне функције f у тачки $0,38$.

Интерполација подразумева да полазну функцију заменимо неком другом, интерполационом функцијом, чије се вредности поклапају са вредностима полазне функције на дискретном скупу тачака.

Може се десити да је функција позната само својим вредностима у дискретним тачкама, не и аналитичким обликом, нпр. ако су у питању резултати изведеног мерења у појединим временским тренуцима. С друге стране, можда радимо с компликованом функцијом, задатом у експлицитном или имплицитном облику. Уколико је израчунавање вредности овакве функције захтевно, можемо да је интерполирамо и да уместо израчунавања вредности полазне функције радимо с интерполационом функцијом. У овом случају имамо слободу избора тачака у којима ћемо вршити интерполацију.

Интерполациону функцију можемо користити за предвиђање вредности функције у тачки која не припада дискретном скупу тачака у којима вршимо интерполацију. Такође, можемо заменити полазну функцију њеном интерполационом функцијом при диференцирању или интеграцији. Добра интерполациона функција треба да буде таква да лако вршимо израчунавања, али такође да буде „довољно” блиска почетној функцији.

Ми ћемо се бавити полиномском интерполацијом, када је интерполациона функција полином одговарајућег степена.

4.1 Полиномска интерполација

Претпоставимо да је функција $f : [x_0, x_n] \mapsto \mathbb{R}$ непрекидна и да су познате њене вредности у $(n+1)$ тачки $x_0 < x_1 < \ldots < x_n$, $f(x_i) = f_i$, $i = 0, \ldots, n$. Означимо полином n-тог степена са

$$p_n(x) = \sum_{i=0}^{n} a_i x^i = a_0 + a_1 x + \ldots + a_n x^n. \tag{4.1}$$

Да би овај полином био *интерполациони полином за функцију* f на интервалу $[x_0, x_n]$, потребно је да испуњава услове:

$$p_n(x_i) = f_i, \qquad i = 0, 1, \ldots, n. \tag{4.2}$$

Тачке $x_i, i = 0, 1, \ldots, n$, називају се *чворови интерполације*. Геометријска интерпретација интерполације функције јесте да треба одредити полином степена n чији график пролази кроз тачке $(x_i, f_i), i = 0, \ldots, n$.

Пример 2. Формирајмо интерполациони полином првог степена $p_1(x) = a_0 + a_1 x$ за познате вредности у чворовима x_0, x_1 на основу табеле вредности:

x	0	1	2
$f(x)$	5	8	9

Из интерполационих услова (4.2) одређујемо непознате коефицијенте a_0 и a_1, интерполационог полинома p_1:

$$\begin{aligned} p_1(x_0) &= a_0 + a_1 \cdot 0 = 5 \\ p_1(x_1) &= a_0 + a_1 \cdot 1 = 8 \end{aligned}$$

Решавањем овог система једначина добијамо да је $a_0 = 5$, $a_1 = 3$, па је

$$p_1(x) = 5 + 3x.$$

Формирајмо сада интерполациони полином $p_2(x) = a_0 + a_1 x + a_2 x^2$ степена 2, на основу вредности у сва 3 позната чвора. (Овде су a_0, a_1, a_2 нови непознати коефицијенти.) Формирамо систем на основу интерполационих услова:

$$\begin{aligned} p_2(x_0) &= a_0 + a_1 \cdot 0 + a_2 \cdot 0^2 = 5 \\ p_2(x_1) &= a_0 + a_1 \cdot 1 + a_2 \cdot 1^2 = 8 \\ p_2(x_2) &= a_0 + a_1 \cdot 2 + a_2 \cdot 2^2 = 9 \end{aligned}$$

Решавањем овог система једначина добијамо коефицијенте полинома $a_0 = 5$, $a_1 = 4$ $a_2 = -1$, тј. интерполациони полином 2. степена за дате вредности функције је

$$p_2(x) = 5 + 4\,x - x^2.$$

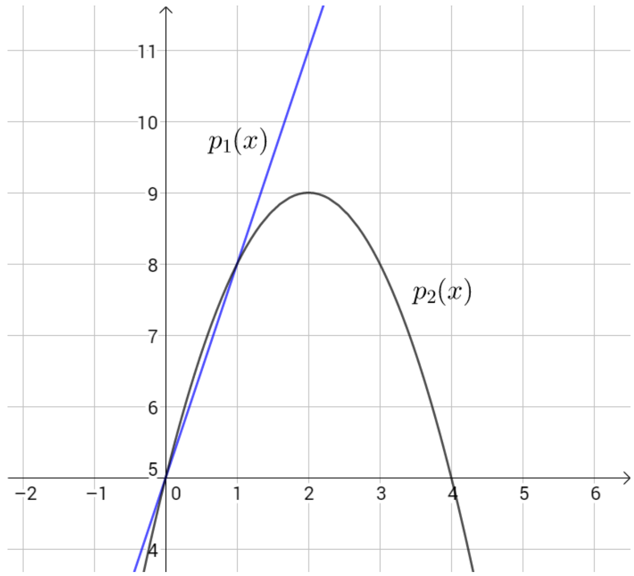
Упоредимо графике интерполационих полинома p_1 и p_2.

Приметимо да графици оба интерполациона полинома морају да садрже тачке $(0,5)$ и $(1,8)$. Вредност полинома $p_1(2) = 11$ разликује се од вредности задате функције за 2. Иако су оба полинома интерполациони за исте вредности функције, приметимо да се, нпр. у тачки $0,5$, вредности полинома такође разликују, $p_1(0,5) = 6,5$ и $p_2(0,5) = 6,75$.

Јединственост интерполационог полинома

Теорема 4.1 *Нека је функција $f : [x_0, x_n] \mapsto \mathbb{R}$ непрекидна и задата својим вредностима у $(n+1)$ тачки $x_0 < x_1 < \ldots < x_n$, са $f_i = f(x_i)$, $i = 0, \ldots, n$. Тада постоји јединствен полином облика (4.1), тј.*

$$p_n(x) = \sum_{i=0}^{n} a_i x^i = a_0 + a_1 x + \ldots + a_n x^n$$



Слика 4.1: Интерполациони полиноми p_1 и p_2, пример 2

такав да је испуњен услов (4.2), тј.

$$p_n(x_i) = f_i, \qquad i = 0, 1, \ldots, n.$$

Доказ. Ако у једнакости (4.1) заменимо $n+1$ услов (4.2), добијамо систем линеарних једначина за одређивање коефицијената $a_0, \ldots, a_n$.

$$\begin{aligned}
a_0 + a_1 x_0 + \ldots + a_{n-1} x_0^{n-1} + a_n x_0^n &= f_0 \\
a_0 + a_1 x_1 + \ldots + a_{n-1} x_1^{n-1} + a_n x_1^n &= f_1 \\
&\vdots \\
a_0 + a_1 x_n + \ldots + a_{n-1} x_n^{n-1} + a_n x_n^n &= f_n
\end{aligned}$$

односно у матричном запису:

$$\begin{bmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^n \\ 1 & x_1 & x_1^2 & \ldots & x_1^n \\ \vdots & \vdots & \ddots & \vdots & \\ 1 & x_n & x_n^2 & \ldots & x_n^n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_n \end{bmatrix}$$

Детерминанта матрице коефицијената система је *Вандермондова детерминанта*, чија је вредност $\prod_{i,j=0,i>j}^{n}(x_i - x_j)$ различита од 0, па систем има јединствено решење за коефицијенте интерполационог полинома. □

Пошто применом различитих метода формирамо интерполационе полиноме различитог облика, јединственост интерполационог полинома гарантује да ће, без обзира на примењену методу и облик, резултат бити исти. Ипак, због грешке заокруживања може доћи до мањих одступања у резултату.

Доказ претходне теореме нам даје и начин да одредимо коефицијенте, па самим тим и интерполациони полином. Међутим, Вандермондова матрица може бити лоше условљена, посебно са повећањем степена полинома или повећањем дужине интервала $[x_0, x_n]$, и захтева велики број израчунавања уколико бисмо применили Гаусову методу елиминације. За одређивање интерполационог полинома користићемо другачије методе, које су флексибилније и одређују полином у облику који је погоднији за рад.

4.2 Лагранжова интерполација

Лагранжов интерполациони полином

Нека је функција $f : [x_0, x_n] \mapsto \mathbb{R}$ непрекидна и задата својим вредностима у $(n+1)$ тачки $x_0 < x_1 < \ldots < x_n$, са $f_i = f(x_i)$, $i = 0, \ldots, n$. Одредимо најпре полином $l_i(x)$ степена n који испуњава услов:

$$l_i(x_j) = \begin{cases} 0, & j \neq i \\ 1, & j = i \end{cases}, \quad i, j = 0, 1, \ldots, n$$

То значи да график полинома $l_i(x)$ сече x осу у тачкама $x_0, \ldots, x_{i-1}, x_{i+1}$, $\ldots, x_n$. Како се полином $l_i(x)$ анулира у овим тачкама и степена је n, мора бити облика:

$$l_i(x) = c_i(x - x_0)\ldots(x - x_{i-1})(x - x_{i+1})\ldots(x - x_n).$$

Константу c_i одређујемо из услова $l_i(x_i) = 1$, па заменом у претходни полином добијамо да је

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

Сада можемо да одредимо и интерполациони полином степена n:

$$\begin{aligned} L_n(x) &= \sum_{i=0}^{n} l_i(x)\, f_i \\ &= \sum_{i=0}^{n} \frac{(x-x_0)(x-x_1)\ldots(x-x_{i-1})(x-x_{i+1})\ldots(x-x_n)}{(x_i-x_0)(x_i-x_1)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots(x_i-x_n)}\, f_i. \end{aligned}$$

Полином $L_n(x)$ назива се *Лагранжов интерполациони полином.* Заиста, испуњени су услови (4.2):

$$L_n(x_j) = \sum_{i=0}^{n} l_i(x_j)\, f_i = l_j(x_j)\, f_j = f_j.$$

Други начин записа Лагранжовог интерполационог полинома

Лагранжов интерполациони полином можемо записати на други начин, који је ефикаснији за израчунавање. Уведимо следећу ознаку за полином степена $n+1$:

$$\Pi_{n+1}(x) = (x-x_0)(x-x_1)\ldots(x-x_n).$$

Искористићемо и први извод овог полинома, у тачки x_i:

$$\Pi'_{n+1}(x_i) = (x_i-x_0)(x_i-x_1)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots(x_i-x_n).$$

Користећи ове ознаке, Лагранжов интерполациони полином можемо записати на следећи начин:

$$L_n(x) = \Pi_{n+1}(x) \sum_{i=0}^{n} \frac{1}{(x-x_i)\Pi'_{n+1}(x_i)}\, f_i. \tag{4.3}$$

У случају када је потребно да израчунамо вредност Лагранжовог интерполационог полинома у конкретној тачки x, а не и да формирамо сам полином, можемо користити следећу шему за израчунавање.

$$\begin{array}{cccc|l} (x-x_0) & (x_0-x_1) & \ldots & (x_0-x_n) & D_0(x) \\ (x_1-x_0) & (x-x_1) & \ldots & (x_1-x_n) & D_1(x) \\ & \ldots & & & \vdots \\ (x_n-x_0) & (x_n-x_1) & \ldots & (x-x_n) & D_n(x) \\ \hline & & & & \Pi_{n+1}(x) \end{array} \tag{4.4}$$

Овде је са $D_i(x)$ означен производ елемената у врсти, а са $\Pi_{n+1}(x)$ производ елемената на дијагонали. Користећи ове ознаке, вредност интерполационог полинома у тачки x израчунавамо на следећи начин:

$$L_n(x) = \Pi_{n+1}(x) \sum_{i=0}^{n} \frac{f_i}{D_k(x)}.$$

4.3 Оцена грешке интерполације

До сада смо претпостављали да је функција коју интерполирамо непрекидна, како би њена апроксимација полиномима имала смисла. Да бисмо могли да се бавимо оценом грешке, тј. колико интерполациони полином одступа од вредности функције, потребно је да претпоставимо да функција има и одговарајућу глаткост.

Грешку интерполације изразићемо користећи L_∞ *норму* непрекидних функција на интервалу $[a, b]$, дефинисану на следећи начин:

$$\|f\|_\infty = \max_{x\in[a,b]} |f(x)|.$$

Теорема 4.2 *Нека је $f \in C^{n+1}[x_0, x_n]$ задата својим вредностима у $(n+1)$ тачки $x_0 < x_1 < \ldots < x_n$, са $f_i = f(x_i)$, $i = 0, \ldots, n$, и нека је p_n одговарајући интерполациони полином. Тада за свако $x \in [x_0, x_n]$ постоји тачка $\xi = \xi(x) \in [x_0, x_n]$ таква да је*

$$f(x) - p_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{i=0}^{n} (x - x_i). \tag{4.5}$$

Додатно, важи следећа ***оцена за горње ограничење грешке:***

$$\|f - p_n\|_\infty \le \frac{1}{(n+1)!} \max_{t\in[x_0,x_n]} |f^{(n+1)}(t)| \max_{s\in[x_0,x_n]} \prod_{i=0}^{n} |s - x_i|. \tag{4.6}$$

Доказ. За $x = x_i$ грешка је 0, па тврђење тривијално важи. За $x \neq x_i, i = 0, \ldots, n$, посматрајмо функцију

$$\varphi(t) = f(t) - p_n(t) - \lambda \Pi_{n+1}(t),$$

где константу λ одређујемо тако да буде испуњен услов $\varphi(x) = 0$. То значи да ће функција φ имати барем $n+2$ нуле, $x, x_0, \ldots, x_n$. На основу Ролове теореме о средњој вредности закључујемо да постоји $\xi \in (x_0, x_n)$ за које важи

$$\varphi^{(n+1)}(\xi) = 0.$$

Приметимо да је $p_n^{(n+1)}(t) = 0$ (полином степена n) и $\Pi_{n+1}^{(n+1)}(t) = (n+1)!$ (полином степена $n+1$ са водећим коефицијентом 1). Из дефиниције функције φ, узимајући $(n+1)$ извод и замењујући t са ξ, имамо да важи:

$$0 = \varphi^{(n+1)}(\xi) = f^{(n+1)}(\xi) - \lambda(n+1)!.$$

Одавде је $\lambda = \frac{f^{(n+1)}(\xi)}{(n+1)!}$. Када уврстимо ово у израз за функцију $\varphi(\xi)$, добијамо израз (4.5).

Како тачка $\xi = \xi(x)$ није позната, можемо узети ограничење грешке (4.5) за све вредности x на интервалу $[x_0, x_n]$, користећи L_∞ норму за $(n+1)$ извод функције f:

$$\|f^{(n+1)}\|_\infty = \max_{t\in[x_0,x_n]} |f^{(n+1)}(t)|.$$

Слично, можемо узети и максимум полинома $|\Pi_{n+1}(x)|$ по интервалу $[x_0, x_n]$, чиме добијамо оцену за горње ограничење грешке (4.6). □

За конкретну вредност $x \in [x_0, x_n]$ користимо ***оцену грешке у тачки*** x, у ознаци $R_n(x)$, уз увођење следеће ознаке $\|f^{(n+1)}\|_\infty \leq M_{n+1}$, у облику:

$$R_n(x) = |f(x) - p_n(x)| \leq \frac{M_{n+1}}{(n+1)!}|\Pi_{n+1}(x)| \tag{4.7}$$

Ове оцене грешке важе без обзира на то у ком је облику исказан интерполациони полином, што је у складу с његовом јединственошћу.

12. Функција $f(x) = \cos x$ је табелирана својим вредностима у чворовима $x_0 = 0; x_1 = 0,6; x_2 = 0,9$.

1. Конструисати Лагранжов интерполациони полином 2. степена за функцију $f(x)$ за чворове интерполације $x_0 = 0; x_1 = 0,6; x_2 = 0,9$.

2. Одредити оцену за горње ограничење грешке.

3. Израчунати $L_2(0,45)$ и одредити оцену грешке у тачки $R_2(0,45)$.

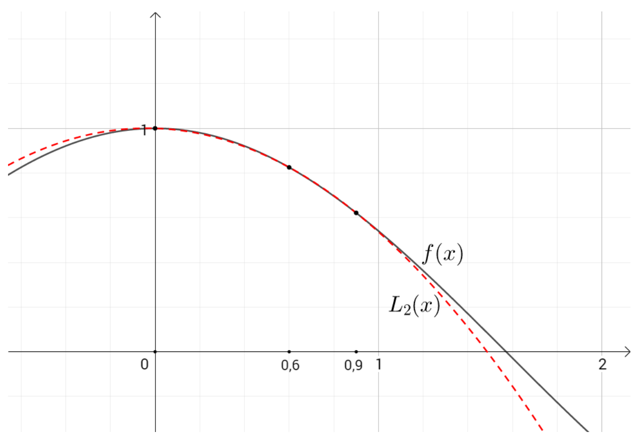
4. Упоредити $R_2(0,45)$ са стварном грешком $|f(0,45) - L_2(0,45)|$.

Решење. Сва израчунавања тригонометријских функција се врше у радијанима. Функција $f(x) = \cos x$ је дата следећом табелом:

x	0	$0,6$	$0,9$
$\cos x$	1	$0,8253356149$	$0,6216099683$

1. Лагранжов интерполациони полином 2. степена за функцију $f(x) = \cos x$ је облика:

$$\begin{aligned}
L_2(x) &= \frac{(x-0,6)(x-0,9)}{(0-0,6)(0-0,9)} \cdot 1 + \frac{(x-0)(x-0,9)}{(0,6-0)(0,6-0,9)} \cdot 0,8253356149 \\
&+ \frac{(x-0)(x-0,6)}{(0,9-0)(0,9-0,6)} \cdot 0,6216099683 \\
&= \frac{(x-0,6)(x-0,9)}{0,6 \cdot 0,9} + \frac{x(x-0,9)}{0,6(0,6-0,9)} \cdot 0,8253356149 \\
&+ \frac{x(x-0,6)}{0,9(0,9-0,6)} \cdot 0,6216099683
\end{aligned}$$



Слика 4.2: Интерполациони полином $L_2(x)$ и функција $f(x) = \cos x$

2. Одредимо оцену за горње ограничење грешке на основу (4.6):

$$\|f - L_2\|_\infty \leq \frac{1}{3!} \max_{t\in[0;0,9]} |f'''(t)| \cdot \max_{x\in[0;0,9]} |(x-0)(x-0,6)(x-0,9)|.$$

Ову оцену ћемо одредити у два корака.

Како је $f'''(x) = \sin x$, а $\sin x$ је растућа и позитивна функција на интервалу $[0; 0,9]$, важи да је:

$$\max_{t\in[0;0,9]} |f'''(t)| \leq \sin 0,9 < 0,79 = M_3.$$

Одредићемо $\max_{x\in[0;0,9]}|x(x-0,6)(x-0,9)|$. Екстремне вредности функције $g(x)=x(x-0,6)(x-0,9)$ су нуле првог извода $g'(x)=3\,(x^2-x+0,18)$ тј. тачке $x_1=0,2354248689$ и $x_2=0,7645751311$. Ове тачке су кандидати и за екстремне вредности за функцију $|g(x)|$. Како је $|g(x_1)|=0,0570405$ и $|g(x_2)|=0,0170405$, имамо да је

$$\max_{x\in[0;0,9]}|x(x-0,6)(x-0,9)|<0,06.$$

Сада је:

$$\|f-L_2\|_\infty\le\frac{1}{6}\cdot 0,79\cdot 0,06=7,9\cdot 10^{-3}.$$

3. Вредност полинома у тачки $x=0,45$ износи

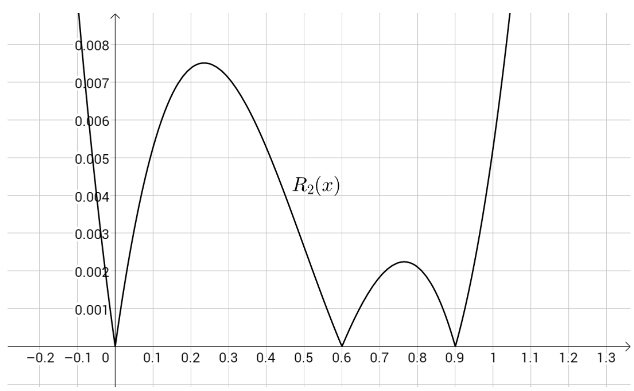
$$L_2(0,45)=0,8981000747.$$

У тачки $x=0,45$ оцена грешке је на основу (4.7):

$$R_2(x)\le\frac{M_3}{3!}\,|x(x-0,6)(x-0,9)|,$$

где је $M_3=0,79$, тј.

$$R_2(0,45)\le\frac{0,79}{6}|0,45\,(0,45-0,6)\,(0,45-0,9)|=4\cdot 10^{-3}.$$



Слика 4.3: Оцена грешке интерполације $R_2(x)$.

На основу оцене грешке закључујемо да су прве две децималне цифре сигурне, тј. $L_2(0,45)=0,898$ и

$$\cos 0,45=L_2(0,45)\pm R_2(0,45)=0,898\pm 0,004$$

4. Стварна грешка износи

$$|f(0,45) - L_2(0,45)| = |\cos 0,45 - 0,898| = |0,900 - 0,898| = 2 \cdot 10^{-3},$$

што је мање од процењене оцене грешке.

■

13. Функција $f(x)$ је задата табелом:

x	0	$0,5$	1	$1,5$
$f(x)$	1	$0,8$	$0,5$	2

Формирати Лагранжове интерполационе полиноме:

1. $L_{2,1}(x)$ у тачкама $x_0 = 0; x_1 = 0,5; x_2 = 1;$
2. $L_{2,2}(x)$ у тачкама $x_0 = 0,5; x_1 = 1; x_2 = 1,5;$
3. $L_3(x)$.

Израчунати вредност формираних интерполационих полинома у тачки $0,8$ и упоредити их.

Решење.

1. Лагранжов интерполациони полином $L_{2,1}(x)$ за функцију $f(x)$ је облика:

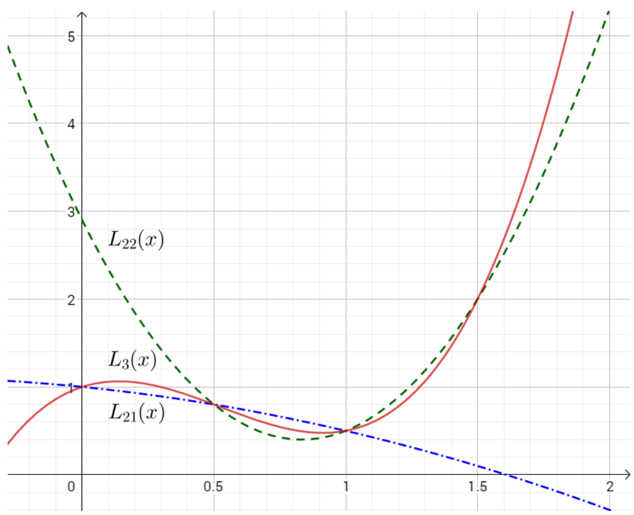
$$\begin{aligned} L_{2,1}(x) = & \frac{(x-0,5)(x-1)}{(0-0,5)(0-1)} \cdot 1 + \frac{(x-0)(x-1)}{(0,5-0)(0,5-1)} \cdot 0,8 \\ & + \frac{(x-0)(x-0,5)}{(1-0)(1-0,5)} \cdot 0,5 \end{aligned}$$

2. Лагранжов интерполациони полином $L_{2,2}(x)$ за функцију $f(x)$ је облика:

$$\begin{aligned} L_{2,2}(x) = & \frac{(x-1)(x-1,5)}{(0,5-1)(0,5-1,5)} \cdot 0,8 + \frac{(x-0,5)(x-1,5)}{(1-0,5)(1-1,5)} \cdot 0,5 \\ & + \frac{(x-0,5)(x-1)}{(1,5-0,5)(1,5-1)} \cdot 2 \end{aligned}$$

3. Лагранжов интерполациони полином $L_3(x)$ за функцију $f(x)$ је облика:

$$L_3(x) = \frac{(x-0,5)(x-1)(x-1,5)}{(-0,5)(-1)(-1,5)} \cdot 1 + \frac{x(x-1)(x-1,5)}{0,5(0,5-1)(0,5-1,5)} \cdot 0,8$$
$$+ \frac{x(x-0,5)(x-1,5)}{1(1-0,5)(1-1,5)} \cdot 0,5 + \frac{x(x-0,5)(x-1)}{1,5(1,5-0,5)(1,5-1)} \cdot 2$$



Слика 4.4: Лагранжови интерполациони полиноми

Вредности одговарајућих интерполационих полинома у тачки 0,8 дате су у табели:

x	$0,8$
$L_{2,1}(x)$	$0,6320$
$L_{2,2}(x)$	$0,4040$
$L_3(x)$	$0,5104$

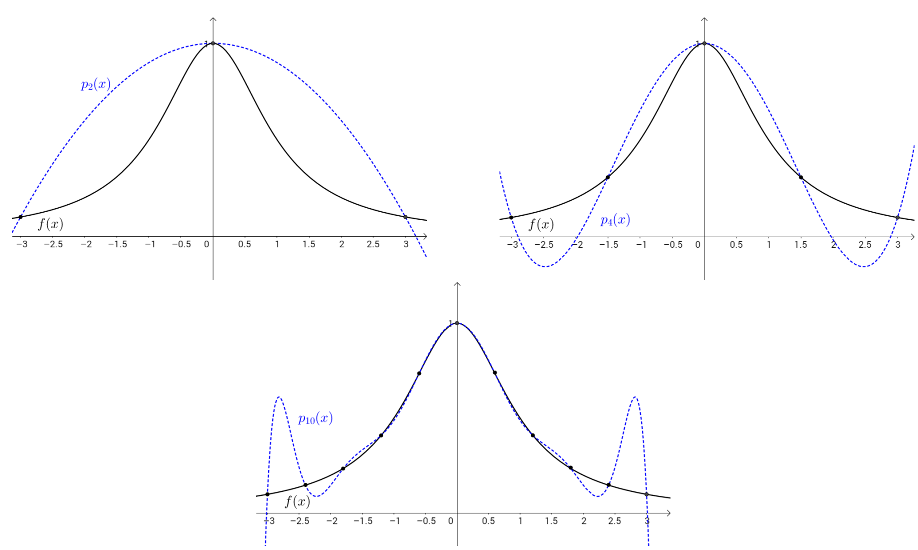
■

Приметимо да у претходном задатку применом Лагранжове интерполације нисмо могли да оценимо грешку, пошто функција није позната у свом аналитичком облику. Додатно, вредности 3 различита полинома у истој тачки се значајно разликују, већ на првој децимали.

Један од недостатака Лагранжове интерполације је потреба да знамо аналитички облик функције како бисмо оценили грешку. Такође, да бисмо формирали интерполациони полином вишег степена, морамо да поновимо цео алгоритам. Ови недостаци могу се исправити применом Невиловог алгоритма, који се користи за одређивање вредности Лагранжовог интерполационог полинома и оцене грешке у конкретној тачки.

Осцилације интерполационих полинома

Иако бисмо очекивали да интерполациони полиноми вишег степена тачније апроксимирају функцију, ово није правило. Овакви полиноми, поред тога што захтевају више израчунавања, често могу веома одступати од функције коју интерполирају, посебно близу крајева интервала на коме се врши интерполација, због генерално већих осцилација у вредностима у односу на полиноме нижих степена. На слици 4.5 приказани су графици интерполационих полинома степена 2, 4 и 10 за функцију $\frac{1}{1+x^2}$ и упоређени са графиком саме функције.



Слика 4.5: Интерполациони полиноми p_2, p_4, p_{10} за функцију $f(x) = \frac{1}{1+x^2}$

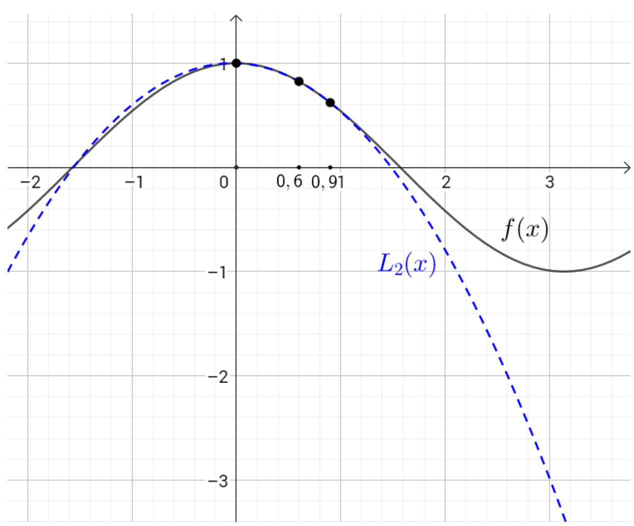
Један од начина на који можемо да решимо овај проблем јесте формирање део-по-део полинома нижег степена, који у тачкама додира испуњавају одређене услове глаткости. На пример, можемо извршити линеарну интерполацију (полиномом првог степена) за свака 2 чвора, и на тај начин бисмо као резултат добили изломљену линију која спаја тачке интерполације. Овакав начин интерполације назива се сплајн интерполација.

Напомена. У случају када је функција задата својим вредностима у више од 4 или 5 чворова, у пракси ћемо формирати интерполационе полиноме степена 3 или 4. У том случају ће бити потребно да одаберемо чворове које ћемо користити за формирање интерполационог полинома. Препорука је да тачка x_0 буде одабрана тако да $x \in (x_0, x_1)$.

Екстраполација

Екстраполација је процес одређивања вредности функције $f(x)$ у тачки која не припада интервалу на коме нам је позната вредност функције ($x \notin [x_0, x_n]$).

Ако се вратимо на пример из задатка 12, вршили смо интерполацију полиномом 2. степена функције $\cos x$ на интервалу $[0; 0,9]$. На слици 4.6 приказан је график ове функције и интерполационог полинома на ширем интервалу, $[-1; 5]$. Видимо да са удаљавањем од интервала на коме смо извршили интерполацију одступање вредности значајно расте.



Слика 4.6: Екстраполација функције $\cos x$

Оцене грешке за интерполацију не важе у случају екстраполације.

У том случају, не постоји начин да се контролише грешка, што захтева посебан опрез ако се ипак одлучимо за екстраполацију.

Задаци за вежбу

26. Функцију $f(x) = \sin \pi x$ табелирати у чворовима $x_0 = 0, x_1 = \frac{1}{6}, x_2 = \frac{1}{2}$.

1. Конструисати Лагранжов интерполациони полином 2. степена $L_2(x)$ за функцију $f(x)$ у формираним чворовима.
2. Одредити оцену за горње ограничење грешке.
3. Израчунати $L_2(0,4)$ и одредити оцену грешке у тачки $R_2(0,4)$.
4. Упоредити $R_2(0,4)$ са стварном грешком $|f(0,4) - L_2(0,4)|$.

Резултат. 2. $\|f - L_2\|_\infty \leq 0,050543$;
3. $L_2(0,4) = 0,920000$, $R_2(0,4) = 0,048232$;
4. $|f(0,4) - L_2(0,4)| = 0,031056$.

27. Функцију $f(x) = e^{-x}$ табелирати у чворовима $x_0 = 1, x_1 = 2, x_2 = 3$.

1. Конструисати Лагранжов интерполациони полином 2. степена $L_2(x)$ за функцију $f(x)$ у формираним чворовима.
2. Одредити оцену за горње ограничење грешке.
3. Израчунати $L_2(1,5)$ и одредити оцену грешке у тачки $R_2(1,5)$.
4. Упоредити $R_2(1,5)$ са стварном грешком $|f(1,5) - L_2(1,5)|$.

Резултат. 2. $\|f - L_2\|_\infty \leq 0,002360$;
3. $L_2(1,5) = 0,233233$, $R_2(1,5) = 0,022992$;
4. $|f(1,5) - L_2(1,5)| = 0,010103$.

28. Функција $f(x)$ је задата табелом:

x	1	2	3	5
$f(x)$	1	5	14	81

Формирати Лагранжове интерполационе полиноме:

1. $L_2(x)$ у чворовима $x_0 = 1, x_1 = 2, x_2 = 3$;
2. $L_3(x)$.

Израчунати вредност формираних интерполационих полинома у тачки $2,8$ и упоредити их.

Резултат. $L_2(x) = 11,8000; L_3(x) = 11,3920$.

4.4 Њутнова интерполација

Задржаћемо се на случају када су чворови интерполације равномерно распоређени, тј. *еквидистантни*:

$$x_i - x_{i-1} = h, \qquad i = 1, \ldots, n$$

при чему разлику свака два чвора h називамо *корак интерполације*. Приметимо да у том случају важи да је $x_i = x_0 + i\,h, (i = 1, \ldots, n)$. Уводимо појам коначних разлика, које ћемо користити при формирању још једног облика интерполационог полинома.

Коначне разлике

Под претпоставком да су нам познате вредности функције у еквидистантним тачкама, $f(x_i) = f_i$, са кораком интерполације h, дефинишемо *коначну разлику унапред*:

$$\Delta f_i = f_{i+1} - f_i.$$

Коначне разлике вишег реда дефинишемо рекурзивно, примењујући линеарност оператора Δ:

$$\begin{aligned}
\Delta^2 f_i &= \Delta(\Delta f_i) = \Delta(f_{i+1} - f_i) = \Delta f_{i+1} - \Delta f_i \\
&= f_{i+2} - 2f_{i+1} + f_i \\
\Delta^3 f_i &= \Delta(f_{i+2} - 2f_{i+1} + f_i) \\
&= f_{i+3} - f_{i+2} - 2f_{i+2} + 2f_{i+1} + f_{i+1} - f_i \\
&= f_{i+3} - 3f_{i+2} + 3f_{i+1} - f_i \\
&\ldots \\
\Delta^k f_i &= \Delta(\Delta^{k-1} f_i)
\end{aligned}$$

Вредности функције f_i представљају коначне разлике нултог реда. Може се показати, математичком индукцијом, да важи:

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

Ради једноставнијег коришћења, коначне разлике записујемо у *табели коначних разлика*:

x_0	f_0	Δf_0	$\Delta^2 f_0$	$\Delta^3 f_0$	$\cdots$
x_1	f_1	Δf_1	$\Delta^2 f_1$	$\vdots$	
x_2	f_2	Δf_2	$\vdots$	$\Delta^3 f_{n-3}$	
x_3	f_3	$\vdots$	$\Delta^2 f_{n-2}$		
$\vdots$	$\vdots$	Δf_{n-1}			
x_n	f_n				

Пропагација грешке у табели коначних разлика

У случају када постоје грешке у вредностима функције, може доћи до њиховог акумулирања у вредностима коначних разлика, посебно вишег реда. До оваквих грешака може доћи ако су вредности функције добијене из мерења. Претпоставимо да је вредност функције f_2 дата са грешком ε, када нам је вредност функције позната у 6 тачака, и формирајмо табелу коначних разлика.

x_0	f_0	Δf_0	$\Delta^2 f_0 + \varepsilon$	$\Delta^3 f_0 - 3\,\varepsilon$	$\Delta^4 f_0 + 6\,\varepsilon$
x_1	f_1	$\Delta f_1 + \varepsilon$	$\Delta^2 f_1 - 2\,\varepsilon$	$\Delta^3 f_0 + 3\,\varepsilon$	$\Delta^4 f_1 - 4\,\varepsilon$
x_2	$f_2 + \varepsilon$	$\Delta f_2 - \varepsilon$	$\Delta^2 f_2 + \varepsilon$	$\Delta^3 f_2 - \varepsilon$	
x_3	f_3	Δf_3	$\Delta^2 f_3$		
x_4	f_4	Δf_4			
x_5	f_5				

Овакво увећање грешке заокруживања значи да су операције коначних разлика виших редова осетљиве на грешке у полазним подацима, па је то још један разлог да интерполацију вршимо с интерполационим полиномима нижег степена. Додатно, иако је интерполациони полином истог степена над истим подацима јединствен, услед грешака у улазним подацима и грешке заокруживања, може се десити да, приликом формирања различитих облика интерполационог полинома, ипак дође до одређених одступања.

Њутнов интерполациони полином за интерполацију унапред

Нека је функција $f : [x_0, x_n] \mapsto \mathbb{R}$ непрекидна и задата својим вредностима у $(n+1)$ еквидистантној тачки $x_0 < x_1 < \ldots < x_n$, са $f_i = f(x_i)$, $i = 0, \ldots, n$. Формираћемо интерполациони полином степена n у следећем облику:

$$\begin{aligned} p_n(x) = a_0 + a_1 (x - x_0) + a_2 (x - x_0)(x - x_1) \\ + \ldots + a_n (x - x_0) \ldots (x - x_{n-1}). \end{aligned}$$

Коефицијенте $a_0, a_1, \ldots, a_n$ ћемо одредити из услова $p_n(x_i) = f_i, i = 0, 1, \ldots, n$. Замењујући вредности за x добијамо:

$$x = x_0 \qquad \Longrightarrow \qquad a_0 = f_0$$

$$x = x_1 \quad \Rightarrow \quad f_0 + a_1(x_1 - x_0) = f_1 \qquad \Longrightarrow \qquad a_1 = \frac{f_1 - f_0}{h} = \frac{\Delta f_0}{h}$$

$$x = x_2 \quad \Rightarrow \quad f_0 + \frac{\Delta f_0}{h}(x_2 - x_0) + a_2(x_2 - x_0)(x_2 - x_1) = f_2$$

$$\Longrightarrow \qquad a_2 = \frac{f_2 - 2f_1 + f_0}{2h^2} = \frac{\Delta^2 f_0}{2h^2}$$

Индукцијом добијамо да је

$$a_k = \frac{\Delta^k f_0}{k!\, h^k}, \quad k = 0, 1, \ldots, n.$$

На овај начин добијамо:

$$\begin{aligned} p_n(x) = f_0 + \frac{\Delta f_0}{h}(x - x_0) + \frac{\Delta^2 f_0}{2!\, h^2}(x - x_0)(x - x_1) + \ldots \\ + \frac{\Delta^n f_0}{n!\, h^n}(x - x_0) \cdots (x - x_{n-1}). \end{aligned}$$

Уводећи смену $\dfrac{x - x_0}{h} = u$, уобичајно је да записујемо *Њутнов интерполациони полином за интерполацију унапред* (или I Њутнов интерполациони полином) у следећем облику:

$$N_n^I(x) = f_0 + \Delta f_0\, u + \frac{\Delta^2 f_0}{2!} u\,(u-1) + \ldots + \frac{\Delta^n f_0}{n!} u\,(u-1) \cdots (u-n+1). \quad (4.8)$$

Напомена. Уколико смо формирали I Њутнов интерполациони полином степена $k-1$, тада полином степена k можемо добити користећи следећу везу:

$$N_k^I(x) = N_{k-1}^I(x) + \frac{\Delta^k f_0}{k!} u\,(u-1)\cdots(u-k+1).$$

Овде се види предност облика Њутновог интерполационог полинома у односу на Лагранжову интерполацију. Код Лагранжове интерполације, уколико желимо да формирамо Лагранжове интерполационе полиноме различитих степена, морамо поново да изведемо цео поступак. Код Њутновог интерполационог полинома довољно је додати одговарајуће сабирке на већ формиран полином.

Облик оцене грешке за I Њутнов интерполациони полином можемо да изведемо из општег облика грешке интерполације (4.5):

$$f(x) - N_n^I(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)\,(x-x_1)\cdots(x-x_n).$$

Користећи уведену смену, имамо да важи:

$$x - x_k = x_0 - u\,h - (x_0 - k\,h) = (u-k)\,h.$$

Заменом у израз за грешку добијамо:

$$f(x) - N_n^I(x) = \frac{1}{(n+1)!}\,u\,(u-1)\cdots(u-n)\,h^{n+1} f^{(n+1)}(\xi).$$

Може се показати да важи следећа апроксимација за вредност извода функције:

$$f^{(n+1)}(\xi) \approx \frac{\Delta^{n+1} f_0}{h^{n+1}}.$$

Овом релацијом, узимајући горње ограничење за оцену грешке, добијамо да је:

$$R_n(x) \leq \frac{|\Delta^{n+1} f|}{(n+1)!}\,|u\,(u-1)\cdots(u-n)|, \tag{4.9}$$

где за $|\Delta^{n+1} f|$ узимамо максималну од апсолутних вредности коначних разлика $(n+1)$ реда.

Напомена. У случају када је функција задата својим вредностима у $(m+1)$ чвору, за $m > n$, где је n степен интерполационог полинома, препорука је да I Њутнов интерполациони полином формирамо од чвора x_0, за који важи да је $x \in (x_0, x_0 + h)$.

14. Функцију $f(x) = \frac{1}{x}$ табелирати на интервалу $[3,1; 3,9]$ са кораком $h = 0,1$, заокружено на 6 децимала. Користећи коначне разлике 3. реда израчунати $f(3,44)$. Проценити грешку интерполације.

Решење. Формирајмо табелу коначних разлика функције $f(x) = \frac{1}{x}$. При томе ћемо коначне разлике 4. реда користити за оцену грешке.

x	f	Δf	$\Delta^2 f$	$\Delta^3 f$	$\Delta^4 f$
3,1	0,322581	−0,010081	0,000611	−0,000053	0,000003
3,2	0,3125	−0,009470	0,000558	−0,000050	0,000010
3,3	0,303030	−0,008912	0,000508	−0,000040	0,000000
3,4	0,294118	−0,008404	0,000468	−0,000040	0,000008
3,5	0,285714	−0,007936	0,000428	−0,000032	0,000000
3,6	0,277778	−0,007508	0,000396	−0,000032	
3,7	0,270270	−0,007112	0,000364		
3,8	0,263158	−0,006748			
3,9	0,256410				

Пошто ћемо формирати интерполациони полином 3. степена, потребне су нам вредности у 4 узастопне тачке. За $x = 3,44$ узимамо вредности почевши од тачке $x_0 = 3,4$. Сада је $u = \frac{x-x_0}{h} = \frac{3,44-3,4}{0,1} = 0,4$. За формирање I Њутновог интерполационог полинома користимо вредности функције и коначних разлика које су у табели подвучене. Заменом у израз (4.8) добијамо:

$$\begin{aligned} N_3^I(3,44) &= 0,294118 - 0,008404 \cdot 0,4 + \frac{0,000468}{2!} \cdot 0,4 \cdot (0,4-1) \\ &\quad + \frac{-0,000040}{3!} \cdot 0,4 \cdot (0,4-1) \cdot (0,4-2), \\ N_3^I(3,44) &= 0,290698. \end{aligned}$$

За оцену грешке интерполације користимо израз (4.9), где за $|\Delta^4 f|$ узимамо максималну од апсолутних вредности коначних разлика 4. реда.

$$\begin{aligned} R_3(3,44) &\le \frac{|\Delta^4 f|}{4!} \left|u\,(u-1)\,(u-2)\,(u-3)\right| \\ &= \frac{0,000010}{4!} \left|0,4\,(0,4-1)\,(0,4-2)\,(0,4-3)\right| = 0,4 \cdot 10^{-6} \approx 0 \end{aligned}$$

■

Њутнов интерполациони полином за интерполацију уназад

С обзиром на претходну напомену, имајући у виду да се број коначних разлика смањује са повећањем реда коначне разлике, за тачку x која се налази ближе чвору x_n може се десити да немамо довољно коначних разлика за формирање I Њутновог интерполационог полинома одговарајућег степена. Због тога ћемо извести верзију овог интерполационог полинома која користи вредности с краја табеле коначних разлика.

Под истим условима, формираћемо интерполациони полином степена n у мало другачијем облику:

$$\begin{aligned} p_n(x) &= a_0 + a_1\,(x-x_n) + a_2\,(x-x_n)\,(x-x_{n-1}) \\ &\quad + \ldots + a_n\,(x-x_n)\ldots(x-x_1). \end{aligned}$$

Коефицијенте $a_0, a_1, \ldots, a_n$ ћемо одредити из услова $p_n(x_i) = f_i, (i = 0, 1, \ldots, n)$. Замењујући вредности за x добијамо:

$$x = x_n \qquad \Longrightarrow \qquad a_0 = f_n$$

$$x = x_{n-1} \quad \Rightarrow \quad f_n + a_1(x_{n-1} - x_n) = f_{n-1} \qquad \Longrightarrow \qquad a_1 = \frac{\Delta f_{n-1}}{h}$$

$$x = x_{n-2} \quad \Rightarrow \quad f_n + \frac{\Delta f_{n-1}}{h}(x_{n-2} - x_n) + a_2(x_{n-2} - x_n)(x_{n-2} - x_{n-1}) = f_{n-2}$$

$$\Longrightarrow \qquad a_2 = \frac{f_n - 2f_{n-1} + f_{n-2}}{2h^2} = \frac{\Delta^2 f_{n-2}}{2h^2}$$

Индукцијом добијамо да је

$$a_k = \frac{\Delta^k f_{n-k}}{k!\,h^k}, \quad k = 0, 1, \ldots, n.$$

На овај начин долазимо до:

$$\begin{aligned} p_n(x) = f_n + \frac{\Delta f_{n-1}}{h}\,(x-x_n) &+ \frac{\Delta^2 f_{n-2}}{2!\,h^2}(x-x_n)\,(x-x_{n-1}) + \ldots \\ &+ \frac{\Delta^n f_0}{n!\,h^n}(x-x_n)\cdots(x-x_1). \end{aligned}$$

Уводећи смену $\frac{x-x_n}{h} = v$, уобичајно је да записујемо *Њутнов интерполациони полином за интерполацију уназад* (или II Њутнов интерполациони полином) у следећем облику:

$$N_{II}(x) = f_n + \Delta f_{n-1}\,v + \frac{\Delta^2 f_{n-2}}{2!}v\,(v+1) + \ldots + \frac{\Delta^n f_0}{n!}v\,(v+1)\cdots(v+n-1). \tag{4.10}$$

Слично као код I Њутновог интерполационог полинома, оцена грешке интерполације II Њутновим интерполационим полиномом је облика:

$$R_n(x) \leq \frac{|\Delta^{n+1} f|}{(n+1)!} \left| v\,(v+1) \cdots (v+n) \right|, \tag{4.11}$$

где за $|\Delta^{n+1} f|$ узимамо максималну од апсолутних вредности коначних разлика $(n+1)$ реда.

Напомена. У случају када је функција задата својим вредностима у $(m+1)$ чвору, за $m > n$, где је n степен интерполационог полинома, препорука је да II Њутнов интерполациони полином формирамо од чвора x_n, за који важи да је $x \in (x_n - h, x_n)$.

15. За функцију задату у задатку 14 израчунати $f(3,44)$ применом II Њутновог интерполационог полинома и проценити грешку.

Решење. Формираћено II Њутнов интерполациони полином степена 3, на основу табеле коначних разлика из задатка 14. Поново наводимо већ формирану табелу:

x	f	Δf	$\Delta^2 f$	$\Delta^3 f$	$\Delta^4 f$
$3,1$	$0,322581$	$-0,010081$	$0,000611$	$-0,000053$	$0,000003$
$3,2$	$0,3125$	$-0,009470$	$0,000558$	$\underline{-0,000050}$	$0,000010$
$3,3$	$0,303030$	$-0,008912$	$\underline{0,000508}$	$-0,000040$	$0,000000$
$3,4$	$0,294118$	$\underline{-0,008404}$	$0,000468$	$-0,000040$	$0,000008$
$3,5$	$\underline{0,285714}$	$-0,007936$	$0,000428$	$-0,000032$	$0,000000$
$3,6$	$0,277778$	$-0,007508$	$0,000396$	$-0,000032$	
$3,7$	$0,270270$	$-0,007112$	$0,000364$		
$3,8$	$0,263158$	$-0,006748$			
$3,9$	$0,256410$				

Пошто формирамо интерполациони полином 3. степена, потребне су нам вредности у 4 узастопне тачке. За $x = 3,44$ узимамо вредности почевши од тачке $x_n = 3,5$. Сада је $v = \frac{x - x_n}{h} = \frac{3,44 - 3,5}{0,1} = -0,6$. За формирање II Њутновог интерполационог полинома користимо вредности функције и коначних разлика које су подвучене у табели. Заменом у израз (4.10) добијамо:

$$
\begin{aligned}
N_3^{II}(3,44) &= 0,285714 - 0,008404 \cdot (-0,6) + \frac{0,000508}{2!} \cdot (-0,6) \cdot (-0,6+1) \\
&\quad + \frac{-0,000050}{3!} \cdot (-0,6) \cdot (-0,6+1) \cdot (-0,6+2), \\
N_3^{II}(3,44) &= 0,290698.
\end{aligned}
$$

За оцену грешке интерполације користимо израз (4.11), где за $|\Delta^4 f|$ узимамо максималну од апсолутних вредности коначних разлика 4. реда.

$$
\begin{aligned}
R_3(3,44) &\leq \frac{|\Delta^4 f|}{4!} \, |v\,(v+1)\,(v+2)\,(v+3)| \\
&= \frac{0,000010}{4!} \, |-0,6 \cdot 0,4 \cdot 1,4 \cdot 2,4| = 0,3 \cdot 10^{-6} \approx 0
\end{aligned}
$$

Приметимо да је резултат исти као када смо применили I Њутнов интерполациони полином. То је и очекивано, с обзиром на то да смо доказали да је интерполациони полином над истим подацима јединствен.

■

16. Табелом је задата функција $f(x)$:

x	15	20	25	30	35	40	45	50	55
$f(x)$	0,2588	0,3420	0,4226	0,5000	0,5736	0,6428	0,7071	0,7660	0,8192

Формирајући интерполациони полином трећег степена, израчунати $f(18)$ и $f(53)$, и проценити грешку.

Решење. Формирајмо табелу коначних разлика:

x	f	Δf	$\Delta^2 f$	$\Delta^3 f$	$\Delta^4 f$
15	$\underline{0,2588}$	$\underline{0,0832}$	$\underline{-0,0026}$	$\underline{-0,0006}$	0
20	0,3420	0,0806	−0,0032	−0,0006	0
25	0,4226	0,0774	−0,0038	−0,0006	0,0001
30	0,5000	0,0736	−0,0044	−0,0005	0
35	0,5736	0,0692	−0,0049	−0,0005	0,0002
40	0,6428	0,0643	−0,0054	$\underline{-0,0003}$	
45	0,7071	0,0589	$\underline{-0,0057}$		
50	0,7660	$\underline{0,0532}$			
55	$\underline{0,8192}$				

За израчунавање вредности $f(18)$ користимо I Њутнов интерполациони полином.

$$N_3^I(x) = 0,2588 + 0,0832\,u + \frac{0,0026}{2!}\,u\,(u-1) + \frac{-0,0006}{3!}\,u\,(u-1)\,(u-2),$$

при чему је $u = \frac{x-x_0}{h}, h = 5, x_0 = 15, x = 18$, па је $u = \frac{18-15}{5} = 0.6$. Дакле, приближна вредност $f(18)$ износи:

$$N_3^I(18) = 0,3090.$$

Грешка у тачки 18 је процењена следећим изразом:

$$\begin{aligned} R_3^I(18) &\leq \frac{|\Delta^4 f|}{4!}|u(u-1)(u-2)(u-3)| = \frac{0.0002}{24}\,0,6\cdot 0,4\cdot 1,4\cdot 2,4 \\ &= 6,7\cdot 10^{-6} \approx 0 \end{aligned}$$

За приближно израчунавање вредности $f(53)$ формираћено II Њутнов интерполациони полином:

$$N_3^{II}(x) = 0,8192 + 0,0532\,v + \frac{-0,0057}{2!}\,v\,(v+1) + \frac{-0,0003}{3!}\,v\,(v+1)\,(v+2),$$

при чему је $v = \frac{x-x_8}{h}, h = 5, x_8 = 55, x = 53$, па је $v = \frac{53-55}{5} = -0,4$.

$$N_3^{II}(53) = 0,7986.$$

Грешку израчунате вредности у тачки 53 процењујемо изразом:

$$\begin{aligned} R_3^{II}(53) &\leq \frac{|\Delta^4 y|}{4!}|\,v\,(v+1)\,(v+2)\,(v+3)| = \frac{0,0002}{24}\,0,4\cdot 0,6\cdot 1,6\cdot 2,6 \\ &= 8,3\cdot 10^{-6} \approx 0. \end{aligned}$$

■

Напомена. У претходном задатку интерполациони полином N_3^I формирали смо на интервалу $[15, 30]$. Наиме, користимо само вредности функције и одговарајућих коначних разлика у овим тачкама. То значи да бисмо за тачку 32 применом формираног полинома вршили екстраполацију. Слично, интерполациони полином N_3^{II} формиран је на интервалу $[40, 55]$, а ван тог интервала вршила би се екстраполација.

Задаци за вежбу

29. Табелом је задата функција $f(x)$:

x	$2,1$	$2,2$	$2,3$	$2,4$
$f(x)$	$0,901951$	$0,978432$	$1,052661$	$1,124724$

x	$2,5$	$2,6$	$2,7$
$f(x)$	$1,194703$	$1,262688$	$1,328751$

Формирајући одговарајући Њутнов интерполациони полином трећег степена, израчунати $f(2,15)$, $f(2,35)$ и $f(2,66)$, и проценити грешку интерполације у овим тачкама.

Резултат. $f(2,15) \approx 0,940478, f(2,35) \approx 1,088959, f(2,66) \approx 1,302552$

30. Табелом је задата функција $f(x)$:

x	$0,6$	$0,7$	$0,8$	$0,9$	$1,0$
$f(x)$	$2,62188$	$2,79665$	$2,95358$	$3,09518$	$3,22383$

x	$1,1$	$1,2$	$1,3$
$f(x)$	$3,34138$	$3,44926$	$3,54864$

Формирајући одговарајући Њутнов интерполациони полином трећег степена, израчунати $f(0,65)$, $f(0,95)$ и $f(1,25)$, и проценити грешку интерполације у овим тачкама.

Резултат. $f(0,65) \approx 2,71165, f(0,95) \approx 3,16098, f(1,25) \approx 3,49994$

31. Функцију $f(x) = \ln x \cos 2x$ табелирати на интервалу $[4; 6,7]$ са кораком $h = 0,3$ на 5 децимала. Формирајући одговарајући Њутнов интерполациони полином трећег степена, израчунати $f(4,2)$, $f(5)$ и $f(6,3)$, и проценити грешку интерполације у овим тачкама на основу формуле за оцену грешке интерполације, као и стварну грешку.

Резултат. $f(4,2) \approx -0,74885, f(5) \approx -1,35587, f(6,3) \approx 1,83471$

32. Функцију $f(x) = \frac{e^x}{\sin x}$ табелирати на интервалу $[1,5; 2,5]$ са кораком $h = 0,1$ на 5 децимала. Формирајући одговарајући Њутнов интерполациони полином трећег степена, израчунати $f(1,65)$, $f(2,05)$ и $f(2,45)$, и проценити грешку интерполације у овим тачкама на основу формуле за оцену грешке интерполације, као и стварну грешку.

Резултат. $f(1,65) \approx 5,22374, f(2,05) \approx 8,75601, f(2,45) \approx 18,17854$

4.5 Инверзна интерполација

Инверзна интерполација је процес одређивања вредности независно променљиве x која одговара некој вредности зависно променљиве $y = f(x)$, када је функција задата својим вредностима у дискретним тачкама.

Инверзна интерполација може да се реализује на следеће начине:

- формирањем интерполационог полинома инверзне функције $x = f^{-1}(y)$;
- решавањем једначине $y = p_n(x)$, где је y задата вредност функције, а p_n одговарајући интерполациони полином, најчешће методом просте итерације за решавање нелинеарних једначина.

Интерполација инверзне функције делује као интуитиван избор. Међутим, он није увек и најбољи. Може се десити да функција f није монотона на интервалу интерполације; у том случају, инверзија није могућа. Додатно, најчешће вредности функције нису еквидистантне, па инверзија табеле може бити компликованија. Уколико су дискретне вредности функције монотоне, сматраћемо да је и функција задата тим вредностима такође монотона.

17. Одредити нуле функције $f(x)$ задате табелом:

x	2	$2,5$	$3,5$	4
$f(x)$	$0,9093$	$0,5985$	$-0,3508$	$-0,7568$

Решење. Треба да одредимо тачку x^* за коју важи да је $y = f(x^*) = 0$, тј. $f^{-1}(0) = x^*$, где је f^{-1} инверзна функција, задата инвертованом табелом:

y	$-0,7568$	$-0,3508$	$0,5985$	$0,9093$
$f^{-1}(y)$	4	$3,5$	$2,5$	2

С обзиром на то да су вредности функције f монотоне, можемо да извршимо овакву инверзију. У ту сврху користимо Лагранжов интерполациони полином за функцију $f^{-1}(y)$. С обзиром на то да нам је потребна само вредност полинома у конкретној тачки 0, довољно је да формирамо шему облика (4.4):

$$
\begin{array}{cccc|c}
(y-y_0) & (y_0-y_1) & (y_0-y_2) & (y_0-y_3) & D_0(y) \\
(y_1-y_0) & (y-y_1) & (y_1-y_2) & (y_1-y_3) & D_1(y) \\
(y_2-y_0) & (y_2-y_1) & (y-y_2) & (y_2-y_3) & D_2(y) \\
(y_3-y_0) & (y_3-y_1) & (y_3-y_2) & (y-y_3) & D_3(x) \\
\hline
 & & & & \Pi_4(y)
\end{array}
$$

Заменом одговарајућих вредности добијамо следећу шему:

$$
\begin{array}{cccc|l}
0,7568 & -0,4060 & -1,3553 & -1,6661 & D_0=-0,693815 \\
0,4060 & 0,3508 & -0,9493 & -1,2601 & D_1=0,170370 \\
1,3553 & 0,9493 & -0,5985 & -0,3108 & D_2=0,239323 \\
1,6661 & 1,2601 & 0,3108 & -0,9093 & D_3=-0,593327 \\
\hline
 & & & & \Pi_4(0)=0,144481
\end{array}
$$

Лагранжов интерполациони полином је облика $P_3(y)=\Pi_4(0)\sum_{k=0}^{3}\frac{y_k}{D_k}$.

$$
\begin{aligned}
P_3(0) &= 0,144481\left(\frac{4}{-0,693815}+\frac{3,5}{0,170370}+\frac{2,5}{0,239323}+\frac{2}{-0,593327}\right) \\
&= 3,15743
\end{aligned}
$$

■

18. Функција $f(x)$ је задата својим вредностима табелом:

x	10	20	30	40
$f(x)$	0,1763	0,3640	0,5774	0,8391

Решити једначину $f(x)=0,25$.

Решење. Одредићемо вредност x^* за коју је $f(x^*)=0,25$. Формирамо табелу коначних разлика:

x	$f(x)$	$\Delta f(x)$	$\Delta^2 f(x)$	$\Delta^3 f(x)$
10	0,1763	0,1877	0,0257	0,0226
20	0,3640	0,2134	0,0483	
30	0,5774	0,2617		
40	0,8391			

На основу вредности у табели, очекујемо да $x^* \in [10, 20]$, па формирамо I Њутнов интерполациони полином:

$$f(x^*) = f(x_0) + \Delta f(x_0)\, u + \frac{\Delta^2 f(x_0)}{2!}\, u\,(u-1) + \frac{\Delta^3 f(x_0)}{3!}\, u\,(u-1)\,(u-2).$$

Изразићемо u из другог сабирка са десне стране:

$$\begin{aligned}
u &= \frac{f(x^*)}{\Delta f(x_0)} - \frac{f(x_0)}{\Delta f(x_0)} - \frac{\Delta^2 f(x_0)}{\Delta f(x_0)}\frac{u(u-1)}{2} - \frac{\Delta^3 f(x_0)}{\Delta f(x_0)}\frac{u(u-1)(u-2)}{6}\\
u &= \frac{0,25}{0,1877} - \frac{0,1763}{0,1877} - \frac{0,0257}{0,1877}\,\frac{u(u-1)}{2} - \frac{0,0226}{0,1877}\,\frac{u(u-1)(u-2)}{6}\\
u &= 0,3926478 - 0,0684603\, u\,(u-1) - 0,0200675\, u\,(u-1)\,(u-2)
\end{aligned}$$

Сада решавамо нелинеарну једначину облика $u = g(u)$, коју ћемо решити методом просте итерације. Нећемо посебно проверавати услове за функцију $g(u)$.

$$u_{n+1} = 0,3926478 - 0,0684603\, u_n\,(u_n-1) - 0,0200675\, u_n\,(u_n-1)\,(u_n-2)$$

Добијамо следећи итеративни низ:

$$\begin{aligned}
u_0 &= 0\\
u_1 &= 0,3926478\\
u_2 &= 0,4012817\\
u_3 &= 0,4013878\\
u_4 &= 0,4013891\\
u_5 &= 0,4013891
\end{aligned}$$

Вредности у 4. и 5. итерацији су се поклопиле, па узимамо да је $u = 0,4013891$. Вршили смо заокруживање на већем броју децимала како бисмо умањили утицај грешке заокруживања при израчунавању. Сада је:

$$x = u\,h + x_0 = 0,4013891 \cdot 10 + 10 = 14,013891.$$

Можемо узети да је $x^* \approx 14,0139$.

■

Задаци за вежбу

33. Одредити вредност x за коју је $f(x) = 10$, при чему је $f(x)$ функција задата табелом:

x	10	15	17	20
$f(x)$	3	7	11	17

Резултат. $x \approx 16,641$

34. Одредити нуле функције $f(x)$ задате табелом:

x	$1,5$	2	$2,5$	3
$f(x)$	$1,9047$	$0,8509$	$-0,4112$	$-1,5727$

Резултат. $x \approx 2,334$

35. Функција f је задата својим вредностима у табели:

x	1	$1,1$	$1,2$	$1,3$	$1,4$
$f(x)$	$0,901951$	$0,978432$	$1,052661$	$1,124724$	$1,194703$

Формирањем одговарајућег интерполационог полинома трећег степена, решити једначину $f(x) = 1$.

Резултат. $x \approx 1,12875$

36. Функција f је задата својим вредностима у табели:

x	$0,7$	1	$1,3$	$1,6$	$1,9$
$f(x)$	$-0,00400$	$0,00471$	$0,011756$	$0,017627$	$0,022641$

Формирањем одговарајућег интерполационог полинома трећег степена одредити нуле функције.

Резултат. $x \approx 0,8297$

Глава 5

Нумеричко диференцирање

Иако знамо експлицитне формуле за одређивање извода функције, јавља се потреба и за нумеричким одређивањем извода, на пример у случају када је израз за функцију коју треба да диференцирамо компликован. Такође, могуће је да функција коју треба да диференцирамо није позната у аналитичком облику, већ само својим вредностима у дискретним тачкама. Формуле за нумеричко диференцирање често се користе за нумеричко решавање диференцијалних једначина.

5.1 Диференцирање помоћу интерполационих полинома

Општи приступ диференцирању функције јесте да извршимо интерполацију функције полиномом, па да диференцирамо интерполациони полином. На сличан начин можемо да проценимо и изводе вишег реда. Овде ћемо се задржати на интерполацији Њутновим интерполационим полиномима у еквидистантним чворовима. Грешка диференцирања је у том случају једнака изводу грешке интерполације $R(x)$.

19. Функција $f(x)$ је задата својим вредностима у табели:

x	-2	0	2	4	6	8
$f(x)$	$2,1272$	$1,5167$	$1,7044$	$3,3285$	$5,0229$	$7,2814$

Одредити координате тачке екстрема функције $f(x)$.

Решење. Апсциса тачке екстрема је нула првог извода функције, који ћемо апроксимирати првим изводом интерполационог полинома.

x	$f(x)$	$\Delta f(x)$	$\Delta^2 f(x)$	$\Delta^3 f(x)$	$\Delta^4 f(x)$
-2	$2,1272$	$-0,6085$	$0,7942$	$0,6442$	$-2,0123$
0	$1,5187$	$0,1857$	$1,4384$	$-1,3681$	$1,8619$
2	$1,7044$	$1,6241$	$0,0703$	$0,4938$	
4	$3,3285$	$1,6944$	$0,5641$		
6	$5,0229$	$2,2585$			
8	$7,2814$				

Према вредностима за $f(x_k) = f_k$ видимо да функција достиже минималну вредност у интервалу $[-2, 2]$, тј. у околини тачке 0. У покушају да смањимо грешку формираћемо I Њутнов интерполациони полином, узимајући за почетну тачку $x_1 = 0$, при чему је $u = \frac{x-x_1}{h}$ и $h = 2$.

$$N_I(x) = f_1 + \Delta f_1 u + \frac{\Delta^2 f_1}{2!} u(u-1) + \frac{\Delta^3 f_1}{3!} u(u-1)(u-2)$$

Први извод I Њутновог интерполационог полинома у овом случају је:

$$N_I'(x) = \frac{1}{h} \left[\Delta f_1 + \frac{\Delta^2 f_1}{2!}(2u-1) + \frac{\Delta^3 f_1}{3!}(3u^2 - 6u + 2) \right].$$

Довољно је да изједначимо израз у загради са 0 и заменимо вредности из табеле коначних разлика:

$$0,1857 + \frac{1,4384}{2}(2u-1) + \frac{-1,3681}{6}(3u^2 - 6u + 2) = 0.$$

Горњу једначину ћемо решити формирањем итеративног процеса по линеарном члану. Наиме, једначина је еквивалентна следећем изразу:

$$u = \frac{1}{2,8065} \left(0,68405\, u^2 + 0,9895333 \right)$$

Формирамо итеративни процес:

$$u_{k+1} = \frac{1}{2,8065} \left(0,68405\, u_k^2 + 0,9895333 \right)$$

Почетна вредност у итеративном поступку је $u_0 = 0$, а низ итерација:

$$\begin{aligned}
u_0 &= 0\\
u_1 &= 0,3526\\
u_2 &= 0,3829\\
u_3 &= 0,3883\\
u_4 &= 0,3893\\
u_5 &= 0,3895\\
u_6 &= 0,3896\\
u_7 &= 0,3896
\end{aligned}$$

Добили смо $u = u_7 = 0,3896$, па је $u = \frac{x-x_1}{h} = \frac{x-0}{2}$. Дакле, функција $f(x)$ достиже минималну вредност у тачки $x = 2\,u + 0 = 0,7792$. Одредимо минимум функције $f(x)$ користећи формиран интерполациони полином:

$$N_I(x) = 1,5187 + 0,1857u + \frac{1,4384}{2!}u(u-1) + \frac{-1,3681}{3!}u(u-1)(u-2)$$

за $u = 0,3896$. Сада је:

$$f(x) \approx N_I(x) = 1,3327.$$

■

Задаци за вежбу

37. Функција f је задата својим вредностима у табели:

x	$1,0$	$1,1$	$1,2$	$1,3$	$1,4$
$f(x)$	$1,1752$	$1,3356$	$1,5095$	$1,6984$	$1,9043$

Израчунати $f'(1,1)$ и $f''(1,1)$ формирањем интерполационих полинома 3. и 4. степена. Добијене резултате упоредити с тачним вредностима $f'(1,1) = 1,6685$ и $f''(1,1) = 1,3356$ заокружене на 4 децимале.

Резултат. $N_3'(1,1) \approx 1,6707, N_3''(1,1) \approx 1,3000,$
$N_4'(1,1) \approx 1,6694, N_4''(1,1) \approx 1,3458$

38. Одредити координате тачке екстрема функције f задате табелом:

x	$1,3$	$1,4$	$1,5$	$1,6$	$1,7$	$1,8$
$f(x)$	$1,2431$	$1,1486$	$1,2095$	$1,4606$	$1,9391$	$2,6846$

Резултат. $x \approx 1,4142; f(x) \approx 1,1471$

39. Одредити координате тачака минимума и максимума функције f задате табелом:

x	0	$0,5$	1	$1,5$	2	$2,5$
$f(x)$	0	$-0,45885$	$-0,68294$	$-0,49499$	$0,18141$	$1,30306$

x	3	$3,5$	4	$4,5$	5	$5,5$	6
$f(x)$	$2,71776$	$4,20157$	$5,51360$	$6,45506$	$6,91785$	$6,91108$	$6,55883$

Резултат. $x_1 \approx 1,0668; f(x_1) \approx -0,6876; x_2 \approx 5,1553; f(x_2) \approx 6,9472$

5.2 Диференцирање помоћу Тејлоровог реда

Разматраћемо формуле које добијамо *дискретизацијом* континуалне функције f, тј. одабирањем вредности функције у еквидистантним тачкама са *кораком дискретизације* h.

Фунцију $f(x)$, која има непрекидне изводе до $(n+1)$ реда на интервалу који садржи тачке x_0 и x_0+h, можемо представити Тејлоровим полиномом степена n

$$f(x_0+h) = f(x_0) + f'(x_0)\,h + \frac{f''(x_0)}{2}\,h^2 + \ldots + \frac{f^{(n)}(x_0)}{n!}\,h^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}\,h^{n+1},$$

где је ξ нека тачка која припада интервалу $[x_0, x_0+h]$.

Нотација $\mathcal{O}$ („велико О”): За величину r која зависи од h пишемо:

$$r = \mathcal{O}(h^s)$$

уколико постоје позитивне константе s и M тако да важи:

$$|r| \le M\,h^s$$

за свако h довољно мало.

Извод функције $f(x)$ у тачки x_0 је дефинисан са:

$$f'(x_0) = \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h}.$$

Да бисмо добили формулу која апроксимира извод функције, бирамо еквидистантне тачке у околини x_0, нпр. $x_0-h, x_0, x_0+h, \ldots$, за мало, позитивно h и конструишемо апроксимацију извода користећи Тејлоров ред у тачки x_0. Грешка методе ће се смањивати са смањењем корака h и зависиће од извода вишег реда функције f. Стога, да бисмо могли да применимо овакве апроксимације, захтевамо да функција f има довољну глаткост, тј. да има непрекидне изводе одговарајућег реда (у зависности од конкретне формуле за апроксимацију) на неком интервалу који садржи све тачке у којима процењујемо вредности функције. У овом одељку ћемо подразумевати да је функција f довољно глатка на оваквом интервалу.

Извешћемо неколико основних формула за нумеричко диференцирање помоћу Тејлоровог реда, према броју тачака које фигуришу у формули.

Нека је функција $f(x) \in C^2[a,b]$, где $x_0-h, x_0, x_0+h \in [a,b]$ за неко позитивно, довољно мало h. Користићемо следећу ознаку M_{n+1}:

$$\|f^{(n+1)}\|_\infty = \max_{x\in[a,b]} |f^{(n+1)}(x)| \le M_{n+1}.$$

Формуле за диференцирање унапред и уназад

Изведимо формулу за нумеричко диференцирање функције f у тачки x_0 која користи вредности функције f у две тачке, x_0 и x_0+h. Функција f се може представити Тејлоровим полиномом у тачки x_0:

$$f(x_0+h) = f(x_0) + f'(x_0)\,h + \frac{f''(\xi)}{2}\,h^2, \qquad \xi \in [x_0, x_0+h].$$

Одавде можемо да изразимо $f'(x_0)$:

$$f'(x_0) = \frac{f(x_0+h)-f(x_0)}{h} - \frac{f''(\xi)}{2}\,h.$$

Формула за диференцирање унапред је облика:

$$D_h = \frac{f(x_0+h)-f(x_0)}{h}. \tag{5.1}$$

Добијена апроксимација има *грешку методе* $r(h)$:

$$r(h) = |f'(x_0) - D_h| \approx \frac{|f''(\xi)|}{2}\, h,$$

за коју важи:

$$r(h) \le \frac{M_2}{2}\, h,$$

тј. $r = \mathcal{O}(h)$, па кажемо да је *тачност формуле првог реда* (степен корака h је 1). Приметимо да ће се за преполовљен корак h грешка методе такође преполовити.

Аналогно претходном, можемо извести *формулу за диференцирање уназад*, за $x = x_0 - h$:

$$f'(x_0) \approx \frac{f(x_0) - f(x_0 - h)}{h}, \tag{5.2}$$

чија је грешка методе r такође првог реда.

Обе ове формуле су двотачкасте.

Пример 1. Апроксимираћемо извод $f'(x_0)$ функције $f(x) = \sin x$ у тачки $x_0 = 1$ користећи формулу за диференцирање унапред. Тачна вредност извода $f'(x) = \cos x$ у тачки x_0 има вредност $f'(1) = \cos 1 = 0,5403023058681397174\ldots$. У табели су дате вредности апроксимације формулом за диференцирање унапред за различите кораке h и вредност апсолутне грешке за сваку апроксимацију.

h	$\frac{\sin(1+h) - \sin(1)}{h}$	$\cos 1 - \frac{\sin(1+h) - \sin(1)}{h}$
$0,1$	$0,497363752535389$	$4,29386 \cdot 10^{-2}$
10^{-2}	$0,536085981011869$	$4,21632 \cdot 10^{-3}$
10^{-3}	$0,539881480360327$	$4,20826 \cdot 10^{-4}$
10^{-5}	$0,540298098505865$	$4,20736 \cdot 10^{-6}$
10^{-7}	$0,540302264040449$	$4,18277 \cdot 10^{-8}$
10^{-9}	$0,540302358409406$	$5,25413 \cdot 10^{-8}$
10^{-10}	$0,540302247387103$	$5,84810 \cdot 10^{-8}$
10^{-12}	$0,540345546085064$	$4,32402 \cdot 10^{-5}$
10^{-14}	$0,544009282066327$	$3,70698 \cdot 10^{-3}$
10^{-15}	$0,555111512312578$	$1,48092 \cdot 10^{-2}$

Приметимо да се до корака $h = 10^{-7}$ грешка смањује линеарно у односу на h. Међутим, када наставимо да смањујемо корак h, грешка апроксимације почиње да расте, толико да је за $h = 10^{-15}$ грешка истог реда као и за $h = 10^{-1}$.

Феномен да се грешка диференцирања повећава са смањивањем корака h је карактеристичан за формуле за нумеричко диференцирање. Наиме, поред грешке методе, на тачност диференцирања утиче и грешка заокруживања. За изузетно малу вредност h формулом којом апроксимирамо први извод реализујемо одузимање блиских вредности функције, али и дељење малом вредношћу, што повећава грешку.

Претпоставимо да се вредности функције $f(x)$ могу изразити са тачношћу ε. Означимо са $\bar{f}(x)$ приближну вредност функције f у тачки x коју користимо у израчунавањима. Тада је

$$f(x) = \bar{f}(x) + e(x),$$

где је $e(x)$ грешка заокруживања вредности функције у тачки x, која је ограничена са ε, тј. $|e(x)| \leq \varepsilon$.

Формула за диференцирање унапред коју ми заправо израчунавамо користи приближне вредности функције:

$$\bar{D}_h = \frac{\bar{f}(x_0 + h) - \bar{f}(x_0)}{h}.$$

Грешка заокруживања утиче на апроксимацију првог извода на следећи начин:

$$\begin{aligned}
|D_h - \bar{D}_h| &= \left| \frac{f(x_0 + h) - f(x_0)}{h} - \frac{\bar{f}(x_0 + h) - \bar{f}(x_0)}{h} \right| \\
&= \left| \frac{e(x_0 + h) - e(x_0)}{h} \right| \\
&\leq \left| \frac{e(x_0 + h)}{h} \right| + \left| \frac{e(x_0)}{h} \right| \leq \frac{2\,\varepsilon}{h}.
\end{aligned}$$

За *грешку заокруживања формуле за диференцирање унапред* узимамо горње ограничење:

$$E(h) = \frac{2\,\varepsilon}{h}.$$

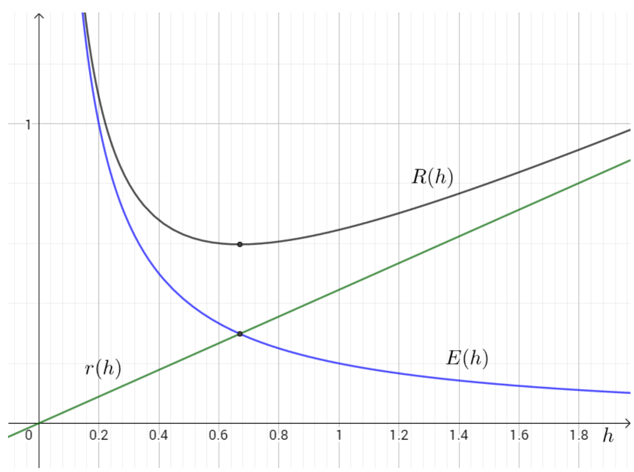
Формула за диференцирање унапред има грешку методе $r(h) = \frac{M_2}{2}\, h$, па је грешка апроксимације првог извода функције овом формулом ограничена са:

$$\begin{aligned} |f'(x_0) - \bar{D}_h| &= |(f'(x_0) - D_h) + (D_h - \bar{D}_h)| \\ &\leq |f'(x_0) - D_h| + |D_h - \bar{D}_h| \\ &\leq \frac{M_2}{2}\, h + \frac{2\,\varepsilon}{h}. \end{aligned}$$

Укупна *грешка диференцирања* формулом за диференцирање унапред је збир грешке методе и грешке заокруживања:

$$R(h) = r(h) + E(h).$$

На слици 5.1 приказани су графици функција $r(h), E(h)$ и $R(h)$, за Пример 1, са тачношћу $\varepsilon = 0,1$.



Слика 5.1: Грешке за формулу диференцирања унапред

Оптимални корак h_o је вредност корака h за коју је грешка диференцирања минимална, тј. $R(h)$ за формулу за диференцирање унапред достиже минимум у тачки

$$h_o = 2\sqrt{\frac{\varepsilon}{M_2}},$$

и то је оптимална вредност корака h за коју је укупна грешка најмања:

$$R(h_o) = 2\sqrt{M_2\,\varepsilon}.$$

Ова оцена за минималну грешку диференцирања значи да је $R(h_o) = \mathcal{O}(\varepsilon^{1/2})$, на основу чега закључујемо да формулом за диференцирање унапред можемо да израчунамо $f'(x_0)$ са двоструко мање сигурних децималних цифара него што их је дато за $f(x_0)$, при оптималном избору корака диференцирања.

Формула за централно диференцирање за први извод

Извешћемо формулу за нумеричко диференцирање првог извода функције f у тачки x_0, која користи вредност функције f у тачкама $x_0 + h$ и $x_0 - h$. Оваква формула је тротачкаста.

20. Нека је функција $f \in C^3[a,b]$, где је $x_0 - h, x_0 + h \in [a,b]$ за корак дискретизације h, и нека се вредности функције $f(x)$ могу изразити са тачношћу ε. Одредити оптималан корак h за формулу за централно диференцирање:

$$f'(x_0) \approx \frac{f(x_0+h) - f(x_0-h)}{2h}.$$

Решење. Да бисмо могли да одредимо оптималан корак h формуле за централно диференцирање, потребно је да најпре изведемо ову формулу, тј. да докажемо да она заиста и важи, и да одредимо грешку диференцирања, која представља збир грешке методе и грешке заокруживања.

Представићемо функцију f у тачкама $x_0 - h$ и $x_0 + h$ Тејлоровим полиномом одговарајућег степена у тачки x_0.

$$f(x_0+h) = f(x_0) + f'(x_0)\,h + \frac{f''(x_0)}{2}\,h^2 + \frac{f'''(\xi_1)}{3!}\,h^3, \qquad \xi_1 \in [x_0, x_0+h]$$

$$f(x_0-h) = f(x_0) - f'(x_0)\,h + \frac{f''(x_0)}{2}\,h^2 - \frac{f'''(\xi_2)}{3!}\,h^3, \qquad \xi_2 \in [x_0-h, x_0]$$

Одузимањем ових развоја добијамо:

$$f(x_0+h) - f(x_0-h) = 2\,f'(x_0)\,h + \frac{f'''(\xi_1) + f'''(\xi_2)}{6}\,h^3 \tag{5.3}$$

На основу теореме о средњој вредности постоји неко $\xi \in [x_0 - h, x_0 + h]$ тако да је

$$\frac{f'''(\xi_1) + f'''(\xi_2)}{2} = f'''(\xi).$$

Када изразимо $f'(x_0)$ из (5.3) добијамо:

$$f'(x_0) = \frac{f(x_0+h) - f(x_0-h)}{2h} - \frac{f'''(\xi)}{6}\, h^2$$

Формула за централно диференцирање је:

$$f'(x_0) \approx \frac{f(x_0+h) - f(x_0-h)}{2h}$$

са грешком методе:

$$r(h) \approx \frac{|f'''(\xi)|}{6}\, h^2.$$

за коју важи:

$$r(h) \leq \frac{M_3}{6}\, h^2.$$

тј. $r = \mathcal{O}(h^2)$, па је *тачност формуле другог реда.* Приметимо да ће се за преполовљен корак h грешка методе смањити 4 пута (за малу вредност h). Зато је формула за централно диференцирање квалитетнија од формула за диференцирање унапред и уназад, чије су грешке методе првог реда.

Означимо са $\bar{f}(x)$ приближну вредност функције f у тачки x коју користимо у израчунавањима. Тада је $f(x) = \bar{f}(x) + e(x)$, где је $e(x)$ грешка заокруживања вредности функције у тачки x, која је ограничена са ε, тј. $|e(x)| \leq \varepsilon$. Сада имамо да је укупна грешка диференцирања

$$f'(x_0) - \frac{\bar{f}(x_0+h) - \bar{f}(x_0-h)}{2h} = \frac{e(x_0+h) - e(x_0-h)}{2h} - \frac{f'''(\xi)}{6}\, h^2.$$

Грешка апроксимације првог извода функције овом формулом ограничена је са:

$$\left| f'(x_0) - \frac{\bar{f}(x_0+h) - \bar{f}(x_0-h)}{2h} \right| \leq \frac{\varepsilon}{h} + \frac{M_3}{6} h^2 = R(h).$$

$R(h)$ за формулу за централно диференцирање достиже минимум у тачки за коју је $R'(h) = -\frac{\varepsilon}{h^2} + \frac{M_3}{3} h = 0$, тј.

$$h_o = \sqrt[3]{\frac{3\varepsilon}{M_3}}.$$

■

Формула за централно диференцирање за други извод

21. Нека је функција $f \in C^4[a,b]$, где је $x_0 - h, x_0 + h \in [a,b]$ за корак дискретизације h, и нека се вредности функције $f(x)$ могу изразити са тачношћу ε. Одредити оптималан корак h за формулу

$$f''(x_0) \approx \frac{f(x_0+h) - 2\,f(x_0) + f(x_0-h)}{h^2}.$$

Решење. Да бисмо могли да одредимо оптималан корак h формуле за централно диференцирање за други извод функције f у тачки x_0, потребно је да најпре изведемо ову формулу, тј. да докажемо да она заиста и важи, и да одредимо грешку диференцирања, која представља збир грешке методе и грешке заокруживања.

Представићемо функцију f у тачкама $x_0 - h$ и $x_0 + h$ Тејлоровим полиномом одговарајућег степена у тачки x_0.

$$f(x_0+h) = f(x_0) + f'(x_0)\,h + \frac{f''(x_0)}{2}\,h^2 + \frac{f'''(x_0)}{3!}\,h^3 + \frac{f^{(iv)}(\xi_1)}{4!}\,h^4,$$

$$f(x_0-h) = f(x_0) - f'(x_0)\,h + \frac{f''(x_0)}{2}\,h^2 - \frac{f'''(x_0)}{3!}\,h^3 + \frac{f^{(iv)}(\xi_2)}{4!}\,h^4.$$

Ове развоје замењујемо у израз за формулу коју доказујемо:

$$\begin{aligned}
&\frac{1}{h^2}\left[f(x_0+h) - 2\,f(x_0) + f(x_0-h)\right]\\
&= \frac{1}{h^2}\Bigg[\left(f(x_0) + f'(x_0)\,h + \frac{f''(x_0)}{2}\,h^2 + \frac{f'''(x_0)}{3!}\,h^3 + \frac{f^{(iv)}(\xi_1)}{4!}\,h^4\right)\\
&- 2\,f(x_0) + \left(f(x_0) - f'(x_0)\,h + \frac{f''(x_0)}{2}\,h^2 - \frac{f'''(x_0)}{3!}\,h^3 + \frac{f^{(iv)}(\xi_2)}{4!}\,h^4\right)\Bigg]\\
&= f''(x_0) + \frac{f^{(iv)}(\xi_1) + f^{(iv)}(\xi_2)}{24}\,h^2.
\end{aligned}$$

На основу теореме о средњој вредности, постоји неко $\xi \in [x_0 - h, x_0 + h]$ тако да је

$$\frac{f^{(iv)}(\xi_1) + f^{(iv)}(\xi_2)}{2} = f^{(iv)}(\xi).$$

Када изразимо $f''(x_0)$ добијамо:

$$f''(x_0) = \frac{f(x_0+h) - 2\,f(x_0) + f(x_0-h)}{h^2} - \frac{f^{(iv)}(\xi)}{12}\,h^2$$

са грешком методе

$$r(x) \approx \frac{|f^{(iv)}(\xi)|}{12}\, h^2,$$

за коју важи:

$$r(h) \leq \frac{M_4}{12}\, h^2.$$

За грешку методе важи да је $r = \mathcal{O}(h^2)$, па је тачност формуле другог реда.

Означимо са $\bar{f}(x)$ приближну вредност функције f у тачки x коју користимо у израчунавањима. Тада је $f(x) = \bar{f}(x) + e(x)$, где је $e(x)$ грешка заокруживања вредности функције у тачки x, која је ограничена са ε, тј. $|e(x)| \leq \varepsilon$. Укупна грешка диференцирања је

$$\begin{aligned}
&\left| f''(x_0) - \frac{\bar{f}(x_0+h) - 2\,\bar{f}(x_0) + \bar{f}(x_0-h)}{h^2} \right| \\
&= \left| \frac{e(x_0+h) - 2\,e(x_0) + e(x_0-h)}{h^2} - \frac{f^{(iv)}(\xi)}{12} h^2 \right| \\
&\leq \frac{1}{h^2}\left(|e(x_0+h)| + 2\,|e(x_0)| + |e(x_0-h)|\right) + \frac{M_4}{12}\, h^2 \\
&\leq \frac{4\,\varepsilon}{h^2} + \frac{M_4}{12}\, h^2
\end{aligned}$$

Укупна грешка ове формуле је $R(h) = \frac{4\,\varepsilon}{h^2} + \frac{M_4}{12}\, h^2$ и достиже минимум у тачки за коју је $R'(h) = -\frac{8\varepsilon}{h^3} + \frac{M_4}{6} h = 0$, тј.

$$h_o = \sqrt[4]{\frac{48\,\varepsilon}{M_4}}.$$

■

22. Нека је функција $f \in C^3[a,b]$, где је $x_0, x_0+h, x_0+2h \in [a,b]$ за корак дискретизације h, и нека се вредности функције $f(x)$ могу изразити са тачношћу ε. Одредити оптималан корак h за формулу

$$f'(x_0) \approx \frac{-3\,f(x_0) + 4\,f(x_0+h) - f(x_0+2h)}{2h}.$$

Решење. Представићемо функцију f у тачкама x_0+h и x_0+2h Тејлоровим полиномом одговарајућег степена у тачки x_0.

$$f(x_0+h) = f(x_0) + f'(x_0)\,h + \frac{f''(x_0)}{2}\,h^2 + \frac{f'''(\xi_1)}{3!}\,h^3, \qquad \xi_1 \in [x_0, x_0+h]$$

$$f(x_0+2h) = f(x_0)+f'(x_0)\,2h+\frac{f''(x_0)}{2}\,4h^2+\frac{f'''(\xi_2)}{3!}\,8h^3, \qquad \xi_2 \in [x_0, x_0+2h]$$

Ове развоје замењујемо у израз за формулу коју доказујемо:

$$\begin{aligned}
&\frac{1}{2h}\left(-3\,f(x_0)+4\,f(x_0+h)-f(x_0+2h)\right)\\
&=\frac{1}{2h}\Bigg[-3\,f(x_0)+4\left(f(x_0)+f'(x_0)\,h+\frac{f''(x_0)}{2}\,h^2+\frac{f'''(\xi_1)}{3!}\,h^3\right)\\
&-\left(f(x_0)+f'(x_0)\,2h+\frac{f''(x_0)}{2}\,4h^2+\frac{f'''(\xi_2)}{3!}\,8h^3\right)\Bigg]\\
&=f'(x_0)+\frac{f'''(\xi_1)-2f'''(\xi_2)}{3}\,h^2
\end{aligned}$$

У овом случају не можемо применити теорему о средњој вредности на остатак, али ако је корак h мали, онда ће и вредности трећег извода у две блиске тачке ξ_1 и ξ_2 бити блиске, па можемо сматрати да је у питању иста тачка $\xi \in [x_0, x_0+2h]$. Када изразимо $f'(x_0)$, добијамо

$$f'(x_0)=\frac{-3\,f(x_0)+4\,f(x_0+h)-f(x_0+2h)}{2h}+\frac{f'''(\xi)}{3}\,h^2.$$

Грешка методе је одавде

$$r(h)\le\frac{M_3}{3}\,h^2.$$

Ако је $\bar{f}(x)$ приближна вредност функције f и $e(x)$ грешка заокруживања вредности функције у тачки x, тако да је $|e(x)|\le\varepsilon$, имамо да је укупна грешка диференцирања

$$\begin{aligned}
&\left|f'(x_0)-\frac{-3\,\bar{f}(x_0)+4\,\bar{f}(x_0+h)-\bar{f}(x_0+2h)}{2h}\right|\\
&=\left|\frac{-3\,e(x_0)+4\,e(x_0+h)-e(x_0+2h)}{2h}+\frac{f'''(\xi)}{3}\,h^2\right|\\
&\le\frac{8\,\varepsilon}{2h}+\frac{M_3}{3}\,h^2
\end{aligned}$$

Укупна грешка формуле је $R(h)=\frac{4\,\varepsilon}{h}+\frac{M_3}{3}\,h^2$ и достиже минимум у тачки за коју је $R'(h)=-\frac{4\varepsilon}{h^2}+\frac{2\,M_3}{3}h=0$, тј.

$$h_o=\sqrt[3]{\frac{6\,\varepsilon}{M_3}}.$$

■

Задаци за вежбу

40. Нека је функција $f \in C^2[a,b]$, где је $x_k, x_k + 2h \in [a,b]$ за корак дискретизације h, и нека се вредности функције $f(x)$ могу изразити са тачношћу ε. Одредити оптималан корак h за формулу

$$f'\left(x_k + \frac{2h}{3}\right) \approx \frac{f(x_k + 2h) - f(x_k)}{2h}.$$

Резултат. $h_o = \sqrt{\frac{3\,\varepsilon}{M_2}}$

41. Нека је функција $f \in C^5[a,b]$, где је $x_k - 2h, x_k + 2h \in [a,b]$ за корак дискретизације h, и нека се вредности функције $f(x)$ могу изразити са тачношћу ε. Одредити оптималан корак h за формулу

$$f'\left(x_k\right) \approx \frac{f(x_k - 2h) - 8f(x_k - h) + 8f(x_k + h) - f(x_k + 2h)}{12h}.$$

Резултат. $h_o = \sqrt[5]{\frac{45\,\varepsilon}{4M_5}}$

42. Нека је функција $f \in C^5[a,b]$, где је $x_k - 2h, x_k + 2h \in [a,b]$ за корак дискретизације h, и нека се вредности функције $f(x)$ могу изразити са тачношћу ε. Одредити оптималан корак h за формулу

$$f'''(x) \approx \frac{f(x + 2h) - 2f(x + h) + 2f(x - h) - f(x - 2h)}{2h^3}.$$

Резултат. $h_o = \sqrt[5]{\frac{18\,\varepsilon}{M_5}}$

Глава 6

Нумеричка интеграција

У овом поглављу бавимо се приближним израчунавањем одређеног интеграла

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

где је $f(x)$ интеграбилна функција на коначном интервалу $[a,b]$. Често се јавља потреба да израчунамо вредност одређеног интеграла функције чија примитивна функција нема експлицитан облик или је није лако одредити. Основне методе за израчунавање вредности одређеног интеграла функције f називају се *квадратурне формуле* и базиране су на следећој апроксимацији интеграла

$$\int_a^b f(x)\,dx = \sum_{i=0}^{n} A_i\, f(x_i) + R(f), \tag{6.1}$$

где су $x_0 < x_1 < \ldots < x_n$ различити чворови и $A_0, A_1, \ldots A_n$ одговарајући коефицијенти. Најчешће сви чворови леже у интервалу $[a,b]$, а коефицијенти A_i одређују се тако да формула (6.1) буде тачна за полиноме што вишег степена. Стога тачност квадратурне формуле зависи од тога колико добро можемо да апроксимирамо подинтегралну функцију $f(x)$ полиномом на интервалу $[a,b]$. Уколико подинтегрална функција има сингуларну тачку на интервалу интеграције или ако је интервал интеграције бесконачан, потребне су одговарајуће модификације методе. У оба случаја може бити од користи размотрити одговарајућу тежинску квадратурну формулу:

$$\int_a^b w(x)\, f(x)\, dx = \sum_{i=0}^{n} A_i\, f(x_i) + R(f),$$

где је $w(x) \geq 0$ *тежинска функција* која преузима сингуларитет, тако да се функција $f(x)$ може добро апроксимирати полиномом. У том случају, интервал интеграције може да остане бесконачан.

За сада ћемо се задржати на формули облика (6.1). Основна квадратурна формула за израчунавање одређеног интеграла $\int_a^b f(x)\, dx$ на коначном интервалу $[a, b]$ базира се на интеграцији интерполационих полинома. Наиме, ако интерполирамо подинтегралну функцију $f(x)$ у различитим чворовима $x_0, \ldots, x_n \in [a, b]$, Лагранжовим интерполационим полиномом $L_n(x)$ степена n, тада је $f(x) = L_n(x) + R_n(x)$ и

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

док је грешка ове формуле $R(f) \leq |\int_a^b R_n(x)\, dx|$. Лагранжов интерполациони полином је облика

$$L_n(x) = \sum_{i=0}^{n} l_i(x)\, f_i,$$

где је $f_i = f(x_i)$ и

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

Тада је

$$\int_a^b f(x)\, dx \approx \int_a^b L_n(x)\, dx = \int_a^b \sum_{i=0}^{n} l_i(x)\, f_i\, dx = \sum_{i=0}^{n} f_i \int_a^b l_i(x)\, dx.$$

Овим смо извели формулу облика (6.1) са коефицијентима:

$$A_i = \int_a^b l_i(x)\, dx. \tag{6.2}$$

Ови коефицијенти се израчунавају само једном, при конструкцији квадратурне формуле, у зависности од степена интерполационог полинома n и одабраних чворова интерполације.

6.1 Њутн-Коутсове квадратурне формуле

Квадратурне формуле (6.1) са коефицијентима (6.2) за еквидистантне чворове интерполације $x_0 < x_1 < \ldots < x_n$ са кораком $h = \frac{b-a}{n}$ називају се *Њутн-Коутсове квадратурне формуле.* У том случају важи да је $x_i = x_0 + i\,h$, и може се показати да су коефицијенти облика $A_i = (b-a)\,C_i^n$, као и да важи $\sum_{i=0}^n C_i^n = 1$ и $C_i^n = C_{n-i}^n$.

6.1.1 Основна трапезна формула

За $n = 1$ интерполирамо подинтегралну функцију линеарним полиномом у крајевима интервала интеграције $x_0 = a$ и $x_1 = b$. У том случају је

$$l_0(x) = \frac{b-x}{b-a}, \qquad l_1(x) = \frac{x-a}{b-a}.$$

Сада су коефицијенти

$$A_0 = \int_a^b l_0(x)\,dx = \int_a^b \frac{b-x}{b-a}\,dx = \frac{b-a}{2},$$
$$A_1 = \int_a^b l_1(x)\,dx = \int_a^b \frac{x-a}{b-a}\,dx = \frac{b-a}{2} = A_0.$$

Овим смо добили *трапезну формулу*

$$\int_a^b f(x)\,dx = \frac{b-a}{2}(f(a) + f(b)), \tag{6.3}$$

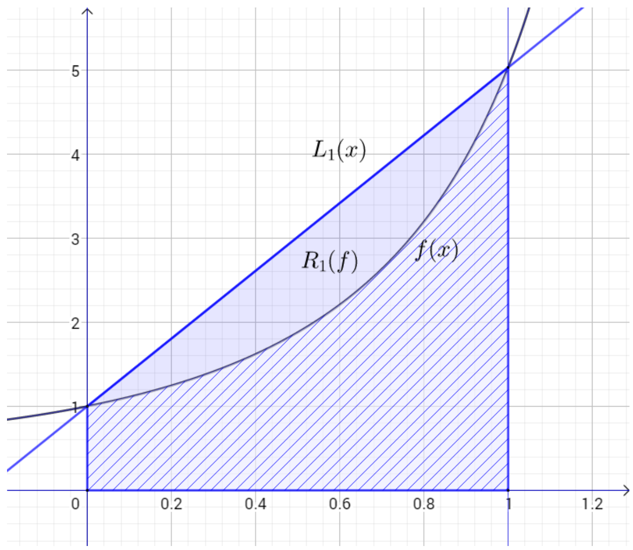
коју још називамо и *основна трапезна формула.*

За позитивну функцију трапезна формула апроксимира површину испод графика функције $f(x)$ површином трапеза.

Грешка основне трапезне формуле

Грешку основне трапезне формуле одређујемо користећи оцену грешке линеарне интерполације $f(x) - L_1(x) = \frac{f''(\xi)}{2}(x-a)\,(x-b)$, под претпоставком да је f'' непрекидна функција на интервалу $[a, b]$,

$$R_1(f) = \int_a^b \frac{f''(\xi)}{2}(x-a)\,(x-b)\,dx = -\frac{f''(\xi)}{12}\,(b-a)^3.$$



Слика 6.1: Апроксимација интеграла функције $\frac{e^x}{\cos x}$ на интервалу $[0, 1]$ основном трапезном формулом

Максимизацијом десне стране једнакости добијамо горњу границу грешке основне трапезне формуле:

$$R_1(f) \leq \frac{M_2}{12}\,(b-a)^3,$$

где је $\max\limits_{x\in[a,b]} |f''(x)| \leq M_2$.

6.1.2 Основна Симпсонова формула

За $n = 2$ интерполирамо подинтегралну функцију полиномом 2. степена, у еквидистантним тачкама са кораком $h = \frac{b-a}{2}$, $x_0 = a, x_1 = a + h$ и $x_2 = b$. У том случају је

$$l_0(x) = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}, \qquad l_1(x) = \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)},$$

$$l_2(x) = \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}.$$

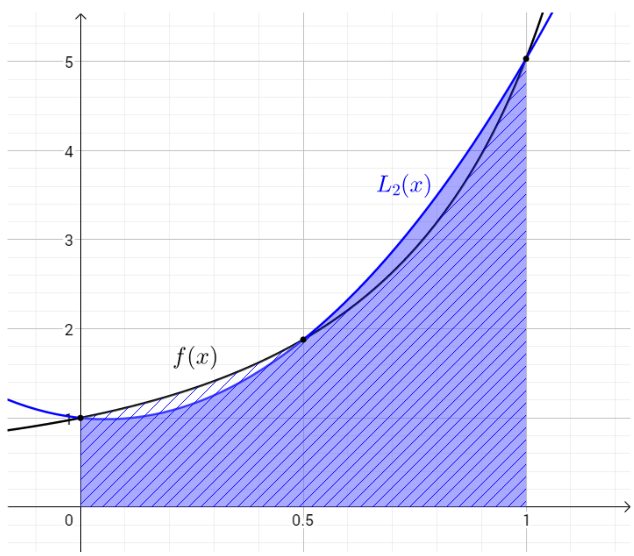
Сада су коефицијенти формуле

$$A_0 = \int_a^b l_0(x)\,dx = \frac{1}{6}\,(b-a) = A_2,$$
$$A_1 = \int_a^b l_1(x)\,dx = \frac{2}{3}\,(b-a).$$

Овим смо добили *Симпсонову формулу*, при чему је $b - a = 2h$,

$$\int_{x_0}^{x_2} f(x)\,dx = \frac{h}{3}(f(x_0) + 4\,f(x_1) + f(x_2)), \tag{6.4}$$

коју још називамо и *основна Симпсонова формула.*



Слика 6.2: Апроксимација интеграла функције $\frac{e^x}{\cos x}$ на интервалу $[0, 1]$ основном Симпсоновом формулом

Грешка основне Симпсонове формуле

Претпоставимо да је f^{IV} непрекидна функција на интервалу $[a, b]$. Грешку основне Симпсонове формуле одређујемо користећи оцену грешке квадратне интерполације

$$\int_{x_0}^{x_2} \frac{f'''(\xi)}{3!}(x-x_0)(x-x_1)(x-x_2)\,dx = \frac{f'''(\xi)}{3!}\cdot 0 = 0.$$

Добили смо да је грешка овакве интеграције једнака 0, а знамо да грешка интерполације функције која је различита од свог интерполационог полинома не може бити 0. Из оваквог резултата закључујемо да је Симпсонова формула за интеграцију тачна за све полиноме 3. степена, па је грешка 4. реда. То значи да полином под интегралом мора бити 4. степена, а то постижемо тако што сматрамо да је први чвор двострука тачка интерполације. Због тога је грешка Симпсонове формуле облика

$$R_2(f) = \int_{x_0}^{x_2} \frac{f^{IV}(\xi)}{4!}(x-x_0)^2(x-x_1)(x-x_2)\,dx = \frac{1}{90} f^{IV}(\xi)\, h^5,$$

где $\xi \in [a,b]$. Узимајући у обзир да је $b-a=2h$, добијамо горњу границу грешке основне Симпсонове формуле

$$R_2(f) \le \frac{M_4}{90}\left(\frac{b-a}{2}\right)^5,$$

где је $\max\limits_{x\in[a,b]} |f^{IV}(x)| \le M_4$.

Грешка основне Симпсонове методе може се извести и коришћењем Њутновог интерполационог полинома за интерполацију подинтегралне функције.

Сличним поступцима можемо да изведемо и Њутн-Коутсове формуле вишег степена од 2. Овакав приступ у неким случајевима може да доведе до повећања тачности интеграције. Други приступ који можемо да применимо у циљу повећања тачности базира се на смањењу дужине интервала на коме примењујемо Њутн-Коутсове формуле. На тај начин добијамо одговарајуће композитне формуле.

6.1.3 Трапезна формула

Грешка основне трапезне формуле $R_1(f) = \frac{f''(\xi)}{12}(b-a)^3$ зависи од дужине интервала интеграције $[a,b]$, па има смисла смањити интервал на коме примењујемо изведене формуле. Поделимо интервал интеграције $[a,b]$ на m подинтервала са кораком $h = \frac{b-a}{m}$, еквидистантним тачкама

$a = x_0 < x_1 < \ldots < x_m = b$, при чему је $x_i = x_0 + ih, x_i - x_{i-1} = h, (i = 1, \ldots, m)$ и означимо $f_i = f(x_i)$. Примениђемо трапезну формулу на сваки подинтервал $[x_{i-1}, x_i]$,

$$\begin{aligned}\int_a^b f(x)\,dx &= \int_{x_0}^{x_m} f(x)\,dx \\ &= \int_{x_0}^{x_1} f(x)\,dx + \int_{x_1}^{x_2} f(x)\,dx + \ldots + \int_{x_{m-1}}^{x_m} f(x)\,dx \\ &= \frac{h}{2}(f_0 + f_1) + \frac{h}{2}(f_1 + f_2) + \ldots + \frac{h}{2}(f_{m-1} + f_m) - \sum_{i=1}^{m} \frac{h^3}{12}\, f''(\xi_i),\end{aligned}$$

за $\xi_i \in [x_{i-1}, x_i], (i = 1, \ldots m)$. С обзиром на претпоставку да је f'' непрекидна функција на интервалу $[a, b]$ и $m\,h = b - a$, грешка може бити записана у облику

$$-\sum_{i=1}^{m} \frac{h^3}{12}\, f''(\xi_i) = -\frac{h^3}{12}\, m f''(\xi) = -(b-a)\,\frac{h^2}{12}\, f''(\xi),$$

где је $\xi \in [a, b]$. Овим смо добили *композитну трапезну формулу* коју називамо само трапезном формулом:

$$T(h) = \frac{h}{2}\left(f_0 + f_m + 2\sum_{i=1}^{m-1} f_i\right), \tag{6.5}$$

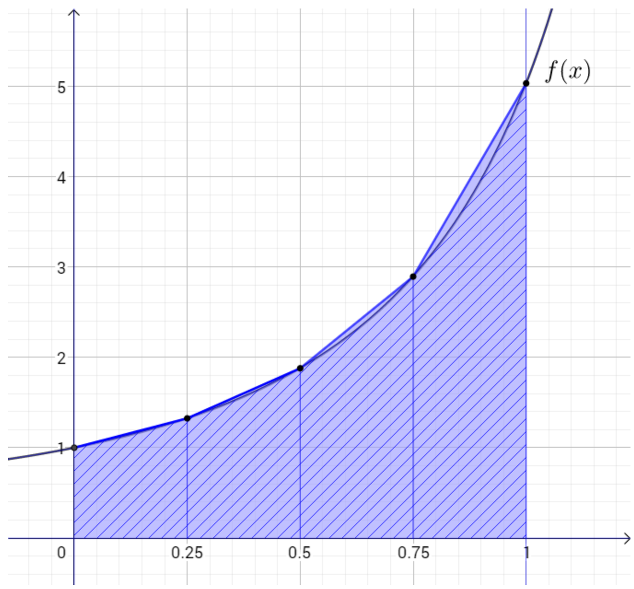
чија је оцена грешке

$$R_T = -(b-a)\,\frac{h^2}{12}\, f''(\xi). \tag{6.6}$$

Горња граница грешке трапезне формуле је

$$R_T \leq \frac{b-a}{12}\, M_2\, h^2,$$

где је $\max\limits_{x\in[a,b]} |f''(x)| \leq M_2$. За грешку трапезне формуле важи да је $R_T = \mathcal{O}(h^2)$, па је тачност формуле другог реда, а формула је тачна за полиноме до 1. степена.



Слика 6.3: Апроксимација интеграла функције $\frac{e^x}{\cos x}$ на интервалу $[0, 1]$ трапезном формулом

6.1.4 Симпсонова формула

Применићемо исти принцип као код трапезне формуле. Делимо интервал интеграције на подинтервале на којима ћемо применити основну Симпсонову формулу. Основну Симпсонову формулу добијамо квадратном интерполацијом подинтегралне функције за коју је потребна подела основног интервала на 2 подинтервала. Зато је приликом уопштавања ове формуле потребно да интервал интеграције поделимо на паран број интервала, како бисмо основну Симпсонову формулу применили на пар подинтервала.

Поделимо интервал интеграције $[a, b]$ на $2m$ подинтервала са кораком $h = \frac{b-a}{2m}$, еквидистантним тачкама $a = x_0 < x_1 < \ldots < x_{2m} = b$, при чему је $x_i = x_0 + ih, x_i - x_{i-1} = h, (i = 1, \ldots, 2m)$ и означимо $f_i = f(x_i)$.

Применићемо Симпсонову формулу на сваки пар подинтервала:

$$\begin{aligned}
\int_a^b f(x)\,dx &= \int_{x_0}^{x_{2m}} f(x)\,dx \\
&= \int_{x_0}^{x_2} f(x)\,dx + \int_{x_2}^{x_4} f(x)\,dx + \ldots + \int_{x_{2m-2}}^{x_{2m}} f(x)\,dx \\
&= \frac{h}{3}(f_0 + 4\,f_1 + f_2) + \frac{h}{3}(f_2 + 4\,f_3 + f_4) + \\
&\qquad \ldots + \frac{h}{3}(f_{2m-2} + 4\,f_{2m-1} + f_{2m}) + \sum_{i=1}^{m} \frac{1}{90} f^{IV}(\xi_{2i})\,h^5,
\end{aligned}$$

за $\xi_{2i} \in [x_{2i-2}, x_{2i}], i = 1, \ldots m$. С обзиром на претпоставку да је f^{IV} непрекидна функција на интервалу $[a, b]$ и $2mh = b - a$, грешка може бити записана у облику

$$\sum_{i=1}^{m} \frac{1}{90} f^{IV}(\xi_{2i})\,h^5 = \frac{1}{90}\,m f^{IV}(\xi)\,h^5 = (b-a)\,\frac{h^4}{180}\,f^{IV}(\xi),$$

где је $\xi \in [a, b]$.

Овим смо добили *композитну Симпсонову формулу* коју називамо само Симпсоновом формулом:

$$S(h) = \frac{h}{3}[f_0 + 4(f_1 + \ldots + f_{2m-1}) + 2\,(f_2 + \ldots + f_{2m-2}) + f_{2m}], \tag{6.7}$$

чија је оцена грешке

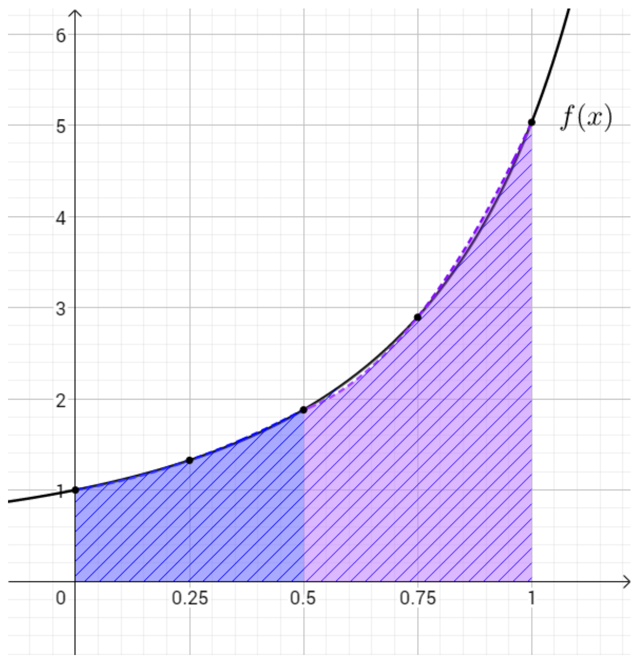
$$R_S = (b-a)\,\frac{h^4}{180}\,f^{IV}(\xi). \tag{6.8}$$

Горња граница грешке Симпсонове формуле је

$$R_S \leq \frac{b-a}{180}\,M_4\,h^4,$$

где је $\max\limits_{x\in[a,b]} |f^{IV}(x)| \leq M_4$.

За грешку Симпсонове формуле важи да је $R_S = \mathcal{O}(h^4)$, па је тачност формуле четвртог реда, а формула је тачна за полиноме до 3. степена.



Слика 6.4: Апроксимација интеграла функције $\frac{e^x}{\cos x}$ на интервалу $[0,1]$ Симпсоновом формулом

6.1.5 Грешка заокруживања

Претпоставимо да се вредности функције $f(x)$ могу изразити са тачношћу ε. Ово ће се десити увек када вршимо заокруживање вредности функције на коначан број децимала. Проценимо на који начин коришћење приближних вредности функције утиче на грешку заокруживања трапезне формуле. Означимо са $\bar{f}_k$ приближну вредност функције f у тачки x_k, коју користимо у израчунавањима. Тада је $f(x_k) = f_k = \bar{f}_k + e_k$, где је e_k грешка заокруживања вредности функције у тачки x_k која је ограничена са ε, тј. $|e_k| \le \varepsilon$.

$$
\begin{aligned}
&|T(h) - \bar{T}(h)| = \\
&= \left|\frac{h}{2}(f_0 + 2\,(f_1 + \ldots + f_{m-1}) + f_m) - \frac{h}{2}(\bar{f}_0 + 2\,(\bar{f}_1 + \ldots + \bar{f}_{m-1}) + \bar{f}_m)\right| \\
&= \frac{h}{2}\,|(e_0 + 2\,(e_1 + \ldots + e_{m-1}) + e_m)| \le \frac{h}{2}\,2m\varepsilon \\
&= (b-a)\,\varepsilon
\end{aligned}
$$

За грешку заокруживања трапезне формуле узимамо горње ограничење

$$E = (b-a)\,\varepsilon.$$

Показује се да је иста грешка заокруживања и Симпсонове формуле. Грешка заокруживања директно зависи од броја децимала с којим вршимо израчунавања. Стога, уколико је грешка заокруживања много мања од грешке методе, може се сматрати занемарљивом. У супротном, укупна грешка трапезне, односно Симпсонове методе је

$$R = R_T + E, \qquad R = R_S + E.$$

Уколико заокружујемо вредности на k децимала, имамо да је $\varepsilon = \frac{1}{2}\cdot 10^{-k}$.

23. Израчунати интеграл

$$\int_0^{\frac{1}{2}} \sin \pi x\, dx$$

применом трапезне и Симпсонове формуле са кораком $h = 0,125$ и проценити грешку интеграције за обе формуле. На који начин број децимала којима рачунамо утиче на грешку интеграције?

Решење. Најпре одредимо вредности функције $f(x) = \sin(\pi x)$ у чворовима са кораком $h = 0.125$, заокружујући на 10 децимала.

x_i	f_i
0	0
$0,125$	$0,3826834325$
$0,25$	$0,7071067813$
$0,375$	$0,9238795325$
$0,5$	1

Трапезна и Симпсонова формула дају

$$T(0,125) = \frac{0,125}{2}(f(0) + f(0,5) + 2\,(f(0,125) + f(0,25) + f(0,375))) = 0,3142087183$$

$$S(0,125) = \frac{0,125}{3}(f(0) + f(0,5) + 4\,f(0,125) + 2\,f(0,25) + 4\,f(0,375)) = 0,3183125267$$

Оцене грешке за трапезну и Симпсонову методу с овим кораком су

$$R_T \le (b-a)\frac{M_2}{12}\,h^2 \qquad R_S \le \frac{b-a}{180}\,M_4\,h^4$$

Како су одговарајући изводи

$$f''(x) = -\pi^2 \sin(\pi x), \qquad f^{IV}(x) = \pi^4 \sin(\pi x),$$

добијамо да су

$$\max_{x\in[0;0,5]} |f''(x)| = |f''(0,5)| \approx 9,8696 \le 9,9 = M_2$$
$$\max_{x\in[0;0,5]} |f^{IV}(x)| = f^{IV}(0,5) \approx 97,409 \le 97,5 = M_4,$$

па је

$$R_T \le \frac{0,5}{12} \cdot 9,9 \cdot (0,125)^2 = 6,4 \cdot 10^{-3} \qquad R_S \le \frac{0,5}{180} \cdot 97,5 \cdot (0,125)^4 = 6,6 \cdot 10^{-5}$$

С обзиром на то да смо израчунавања вршили заокруживањем на 10 децимала, тј. $\varepsilon = \frac{1}{2} \cdot 10^{-10}$, важи да је за обе методе грешка заокруживања $E = (b-a)\varepsilon = \frac{0,5-0}{2} \cdot 10^{-10} = 0,25 \cdot 10^{-10}$, што је занемарљиво у односу на оцену грешака метода. С обзиром на добијене оцене грешке, за трапезну методу би било довољно вршити заокруживања на 3 децимале, а за Симпсонову на 5 децимала.

Оцене грешке показују да је трапезном формулом постигнута тачност на прве 2 децимале, а Симпсоновом формулом са истим кораком на прве 4 децимале. Уколико упоредимо добијена решења с аналитичким решењем интеграла које износи $\frac{1}{\pi} \approx 0,3183098861\ldots$ видимо да се ове оцене уклапају у стварну грешку решења.

$\blacksquare$

24. Израчунати интеграл

$$\int_0^1 \frac{dx}{1+x}$$

Симпсоновом формулом са тачношћу $\varepsilon = 10^{-4}$.

Решење. Укупна грешка Симпсонове формуле износи $R = 10^{-4}$, а грешка заокруживања $E = \frac{1-0}{2} 10^{-4} = 0,5 \cdot 10^{-4}$ уколико заокружујемо на 4 децимале. Максимална дозвољена грешка Симпсонове методе је

$$R_S = R - E = 10^{-4} - 0,5 \cdot 10^{-4} = 0,5 \cdot 10^{-4}.$$

Да бисмо испунили захтевану тачност, тј. да грешка Симпсонове формуле R_S буде мања од захтеване тачности, потребно је да проценимо колики треба да буде корак поделе интервала интеграције h. То радимо на основу оцене грешке Симпсонове формуле (6.8):

$$R_S = \frac{b-a}{180}\, M_4\, h^4, \qquad M_4 = \max_{x\in[0,1]} |f^{IV}(x)|$$

$$f(x) = \frac{1}{1+x}, \qquad f^{IV}(x) = \frac{24}{(1+x)^5}$$

$$M_4 = \max_{[0,1]} \left| \frac{24}{(1+x)^5} \right| = 24$$

Сада је

$$R_S = \frac{24}{180}\, h^4 \le 0,5 \cdot 10^{-4} \quad \Rightarrow \quad h^4 \le \frac{180}{48} \cdot 10^{-4}.$$

На основу овог услова следи да за корак h мора да важи $h \le 0,139$ да би била испуњена тачност. За корак h поделе интервала узимамо $h = 0,1$, чиме добијамо поделу интервала интеграције на паран број подинтервала. Формирамо табелу с одговарајућим вредностима функције $f(x) = \dfrac{1}{1+x}$.

x	f_0, f_{2m}	$f_{2k-1}(*4)$	$f_{2k}(*2)$
0	1		
$0,1$		$0,9091$	
$0,2$			$0,8333$
$0,3$		$0,7692$	
$0,4$			$0,7143$
$0,5$		$0,6667$	
$0,6$			$0,6250$
$0,7$		$0,5882$	
$0,8$			$0,5556$
$0,9$		$0,5263$	
$1,0$	$0,5$		
$\sum$	$1,5000$	$3,4595$	$2,7282$

Формирамо Симпсонову формулу са кораком $0,1$:

$$S(0,1) = \frac{0,1}{3}(1,5 + 4 \cdot 3,4595 + 2 \cdot 2,7282).$$

С обзиром на то да смо унапред одредили потребан корак h који обезбеђује тражену тачност, имамо да је

$$I \approx S(0,1) = 0,6931.$$

■

Задаци за вежбу

43. Израчунати интеграл

$$I = \int_1^2 x \ln x \, dx$$

1. трапезном формулом са тачношћу $\varepsilon = 10^{-4}$;
2. Симпсоновом формулом са тачношћу $\varepsilon = 10^{-5}$.

Резултат. 1. $I \approx 0,6364$
2. $I \approx 0,63630$

44. Израчунати интеграл

$$I = \int_0^1 \sin\left(\frac{\pi}{2} x^2\right) dx$$

1. трапезном формулом са тачношћу $\varepsilon = 5 \cdot 10^{-3}$;
2. Симпсоновом формулом са тачношћу $\varepsilon = 10^{-3}$.

Резултат. 1. $I \approx 0,438$
2. $I \approx 0,438$

45. Израчунати интеграл

$$I = \int_{0,75}^{1,75} (\sin^2 x - 2x \sin x + 1) \, dx$$

1. трапезном формулом са тачношћу $\varepsilon = 10^{-3}$;
2. Симпсоновом формулом са тачношћу $\varepsilon = 10^{-4}$.

Резултат. 1. $I \approx -0,489$
2. $I \approx -0,4890$

6.2 Ромбергова интеграција

Ромбергова интеграција представља ефикасан алгоритам за нумеричко решавање интеграла, базиран на узастопној примени трапезне формуле.

6.2.1 Ричардсонова екстраполација

Ричардсонова екстраполација односи се на методе које користе две нумеричке процене вредности интеграла да би се добила трећа процена веће тачности.

У оцени грешке трапезне и Симпсонове методе фигуришу оцене 2. односно 4. извода подинтегралне функције на интервалу интеграције, што може додатно да оптерети израчунавање и није увек једноставан задатак. Зато уводимо методу која даје прецизнију апроксимацију интеграла, користећи само вредности добијене одговарајућом формулом с различитим корацима поделе интервала интеграције. Задржаћемо се на случају када корак h половимо, пошто можемо да користимо претходно израчунате вредности подинтегралне функције за нову примену Њутн-Коутсове формуле.

Претпоставићемо да подинтегрална функција има довољну глаткост, тј. да је непрекидна и да има непрекидне изводе довољно високог реда. Може се показати да се грешка трапезне формуле

$$R_T(h) = -(b-a)\,\frac{h^2}{12}\,f''(\xi)$$

може развити на следећи начин:

$$R_T(h) = c_2h^2 + c_4h^4 + \ldots + c_{2s}h^{2s} + \mathcal{O}(h^{2s+1}),$$

где константе c_i зависе од извода одговарајућег реда функције f, али не и од h.

Претпоставимо да је израчуната приближна вредност одређеног интеграла $I = \int_a^b f(x)\,dx$ трапезном формулом с корацима $h_1 = h = b-a$ и $h_2 = \frac{h}{2}$. Израчунајмо две апроксимације трапезном формулом с овим корацима:

$$T(h_1) = \frac{h_1}{2}(f(a) + f(b))$$

$$T(h_2) = \frac{h_2}{2}(f(a) + 2f(a+h_2) + f(b))$$

Одговарајуће грешке трапезне формуле са ова два корака су:

$$R_T(h_1) = I - T(h_1) = c_2h^2 + c_4h^4 + c_6h^6 + \cdots,$$
$$R_T(h_2) = I - T(h_2) = c_2(h/2)^2 + c_4(h/2)^4 + c_6(h/2)^6 + \cdots$$

Приметимо да је $R_T(h_1) = 4R_T(h_2) + \mathcal{O}(h^4)$. Заправо, важи да је

$$R_T(h_1) - 4R_T(h_2) = \frac{3}{4}c_4h^4 + \frac{15}{16}c_6h^6 + \cdots.$$

С друге стране, имамо и да је $R_T(h_1) - 4R_T(h_2) = 4T(h_2) - T(h_1) - 3I$, па следи да важи

$$I = \frac{4T(h_2) - T(h_1)}{3} - \frac{1}{4}c_4h^4 - \frac{5}{16}c_6h^6 + \cdots.$$

Овим смо добили нову апроксимацију полазног интеграла I, чија је грешка $\mathcal{O}(h^4)$:

$$I \approx \frac{4T(h_2) - T(h_1)}{3}. \tag{6.9}$$

6.2.2 Ромбергова интеграција

Претходно изведени корак Ричардсонове екстраполације можемо да уопштимо. Како бисмо добили формулу која је тачности $\mathcal{O}(h^{2s})$, користимо трапезне формуле с корацима $h, h/2, \ldots, h/2^{s-1}$ и примењујемо описану процедуру Ричардсонове екстраполације $s-1$ пута.

Ромбергова интеграција је алгоритам којим повећавамо тачност процене вредности интеграла, узастопном применом Ричардсонове екстраполације.

Означимо апроксимацију трапезном формулом $T(h_k)$ са $T_k^{(0)}$, за $h_k = \frac{b-a}{2^k}, k = 0, 1, \ldots$. На основу ових апроксимација конструисаћемо вредности $T_k^{(j)}$ $(j = 1, \ldots, k;\ k = 1, 2, \ldots)$, на следећи начин

$$T_k^{(j)} = \frac{4^j T_k^{(j-1)} - T_{k-1}^{(j-1)}}{4^j - 1}. \tag{6.10}$$

Приметимо да је за $j = 1$ ово апроксимација (6.9). Користећи добијене вредности можемо да формирамо T-табелу са вредностима апроксимација полазног интеграла I:

$$
\begin{array}{ccccccc}
T_0^{(0)} & & & & & & \\
 & \searrow & & & & & \\
T_1^{(0)} & \rightarrow & T_1^{(1)} & & & & \\
 & \searrow & & \searrow & & & \\
T_2^{(0)} & \rightarrow & T_2^{(1)} & \rightarrow & T_2^{(2)} & & \\
 & \searrow & & \searrow & & \searrow & \\
T_3^{(0)} & \rightarrow & T_3^{(1)} & \rightarrow & T_3^{(2)} & \rightarrow & T_3^{(3)} \\
\vdots & & \vdots & & \vdots & & \ddots
\end{array}
$$

Вредности у првој колони су апроксимације вредности интеграла I добијене трапезном формулом с кораком h_k. Апроксимације у 1. колони су тачности $\mathcal{O}(h^2)$, у 2. колони тачности $\mathcal{O}(h^4)$, у s. колони $\mathcal{O}(h^{2s})$.

Итеративни процес заустављамо када је испуњено $|T_j^{(j)} - T_{j-1}^{(j-1)}| \leq \varepsilon$, где је ε унапред захтевана тачност. Тада узимамо да је $I \approx T_j^{(j)}$.

Напомена. Када примењујемо Ромбергову интеграцију са унапред задатом тачношћу коју треба да постигнемо, немамо унапред одређен број k, тј. колико је апроксимација трапезном формулом потребно да израчунамо (прва колона). Табелу попуњавамо по редовима и након сваког реда процењујемо вредност $|T_j^{(j)} - T_{j-1}^{(j-1)}|$. Уколико је тачност испуњена стајемо, у супротном, израчунавамо вредности у наредном реду.

Наводимо алгоритам за Ромбергову интеграцију, са почетним кораком $h_0 = \frac{b-a}{r}$.

КОРАК 1: Бирамо почетни корак $h_0 = \frac{b-a}{r}$.

КОРАК 2: Израчунати полазну трапезну формулу

$$T_0^{(0)} = \frac{h_0}{2}\left(f(a) + f(b) + 2\sum_{i=1}^{r-1} f(a + ih_0)\right)$$

КОРАК 3.1: За $k = 1$: За $h_k = h_0/2^k$ израчунати трапезну формулу:

$$T_k^{(0)} = \frac{h_k}{2}\left(f(a) + f(b) + 2\sum_{i=1}^{2^k r-1} f(a + ih_k)\right)$$

КОРАК 3.2: Израчунати Ричардсонову екстраполацију:

$$T_k^{(j)} = \frac{4^j T_k^{(j-1)} - T_{k-1}^{(j-1)}}{4^j - 1}, \qquad (j = 1, \ldots, k)$$

КОРАК 4: Ако је $|T_j^{(j)} - T_{j-1}^{(j-1)}| > \varepsilon$, повећати k за 1, повратак на корак 3.1

Ако је $|T_j^{(j)} - T_{j-1}^{(j-1)}| \le \varepsilon$, $I \approx T_j^{(j)}$

25. Применом Ромбергове интеграције одредити апроксимацију $T_3^{(3)}$ интеграла

$$\int_0^{\frac{1}{2}} \sin \pi x \, dx$$

са почетним кораком $h_0 = 0,5$. Одредити аналитичко решење интеграла и проценити стварну грешку $|I - T_k^{(j)}|, j, k = 0, 1, 2, 3$.

Решење. Израчунавамо апроксимације трапезним формулама с корацима $h_0 = 0,5; h_1 = h_0/2 = 0,25; h_2 = h_0/4 = 0,125; h_3 = h_0/8 = 0,0625$ (вредности у првој колони T-табеле). Вредности у преосталим колонама добијамо применом Ричардсонове екстраполације (6.10).

0,25			
0,301777	0,319036		
0,314209	0,318353	0,318307	
0,317286	0,318312	0,318310	0,318310
$\mathcal{O}(h^2)$	$\mathcal{O}(h^4)$	$\mathcal{O}(h^6)$	$\mathcal{O}(h^8)$

Аналитичко решење интеграла је $\frac{1}{\pi}$. Вредности грешке $|I - T_k^{(j)}|, (j, k = 0, 1, 2, 3)$ су записане у табели:

$6,8 \cdot 10^{-2}$			
$1,6 \cdot 10^{-2}$	$7,2 \cdot 10^{-4}$		
$4,1 \cdot 10^{-3}$	$4,2 \cdot 10^{-5}$	$2,7 \cdot 10^{-6}$	
$1,0 \cdot 10^{-3}$	$2,6 \cdot 10^{-6}$	$7,0 \cdot 10^{-8}$	$2,8 \cdot 10^{-8}$

■

Напомена. [6] Трапезна формула има грешку $\mathcal{O}(h^2)$, што значи да очекујемо да је грешка $I - T(h/2)$ мања 4 пута од грешке $I - T(h)$, када преполовимо корак h. Показује се да се као индикатор грешке формуле $T(h)$ може користити грешка формуле за два узастопна корака, $|T(h) - T(2h)|$. На основу претходног, очекујемо да је однос грешака трапезне формуле са два узастопна корака, користећи ознаке из Ромбергове интеграције:

$$R_m = \frac{|T_m^{(0)} - T_{m-1}^{(0)}|}{|T_{m+1}^{(0)} - T_m^{(0)}|} \approx 4.$$

Ромбергова интеграција захтева да је подинтегрална функција довољно глатка. У супротном, количници R_m ће сувише одступати од 4. Уколико је одступање веће до 10%, а корак h_m довољно мали, не можемо успешно применити Ромбергову интеграцију за решавање интеграла. У претходном задатку услов за успешну примену ове методе је испуњен.

26. Израчунати интеграл

$$\int_0^1 \cos(x^2)\,dx$$

са тачношћу $\varepsilon = 0,5 \cdot 10^{-4}$ применом Ромбергове интеграције, с почетним кораком $h_0 = 0,5$.

Решење. Сва израчунавања заокружујемо на 5 децимала. Одредићемо приближну вредност интеграла трапезном формулом са корацима $h_0 = 0,5$ и $h_1 = 0,25$. Вредности подинтегралне функције $f(x) = \cos(x^2)$ у чворовима са кораком $0,25$ дате су у табели.

x	f_0, f_m	f_k
0	$1,00000$	
$0,25$		$0,99988$
$0,5$		$0,96891$
$0,75$		$0,84592$
$1,0$	$0,54030$	
$\sum$	$1,54030$	$2,81288$

Апроксимације интеграла трапезном формулом са корацима $h_0 = 0,5$ и $h_1 = 0,25$ су

$$T_0^{(0)} = T(0,5) = \frac{0,5}{2}(1,54030 + 2 \cdot 0,96891) = 0,86953$$
$$T_1^{(0)} = T(0,25) = \frac{0,25}{2}(1,54030 + 2 \cdot 2,81288) = 0,89576$$

Применимо Ричардсонову екстраполацију на претходно израчунате апроксимације.

$$T_1^{(1)} = \frac{4T_1^{(0)} - T_0^{(0)}}{4-1} = 0,90450$$

T-табела је сада облика:

$$\begin{array}{lll} T_0^{(0)} = 0,86953 & & \\ & \searrow & \\ T_1^{(0)} = 0,89576 & \rightarrow & 0,90450 \end{array}$$

Процена грешке је

$$|T_1^{(1)} - T_0^{(0)}| = |0,90450 - 0,86953| = 0,03497,$$

а пошто тачност није испуњена, израчунаћемо апроксимацију трапезном формулом са кораком $h_2 = h_1/2 = 0,125$.

x	f_k
$0,125$	$0,99988$
$0,375$	$0,99013$
$0,625$	$0,92467$
$0,875$	$0,72095$
$\sum$	$3,63563$

$$T_2^{(0)} = T(0,125) = \frac{0,125}{2}(1,54030 + 2 \cdot (2,81288 + 3,63563)) = 0,90233.$$

Применимо Ричардсонову екстраполацију на апроксимације трапезном формулом са корацима $h_1 = 0,25$ и $h_2 = 0,125$.

$$T_2^{(1)} = \frac{4T_2^{(0)} - T_1^{(0)}}{4-1} = 0,90452$$

T-табела је сада облика:

$$\begin{array}{lcl}
T_0^{(0)} = 0,86953 & & \\
 & \searrow & \\
T_1^{(0)} = 0,89576 & \rightarrow & 0,90450 \\
 & \searrow & \\
T_2^{(0)} = 0,90233 & \rightarrow & 0,90452
\end{array}$$

Сада можемо да применимо Ричардсонову екстраполацију на апроксимације интеграла чија је тачност $\mathcal{O}(h^4)$, чиме се добија апроксимација $T_2^{(2)}$ тачности $\mathcal{O}(h^6)$.

$$T_2^{(2)} = \frac{4^2 T_2^{(1)} - T_1^{(1)}}{4^2-1} = 0,90452$$

T-табела је сада облика:

$$\begin{array}{lclcl}
T_0^{(0)} = 0,86953 & & & & \\
 & \searrow & & & \\
T_1^{(0)} = 0,89576 & \rightarrow & 0,90450 & & \\
 & \searrow & & \searrow & \\
T_2^{(0)} = 0,90233 & \rightarrow & 0,90452 & \rightarrow & 0,90452
\end{array}$$

Процена грешке је

$$|T_2^{(2)} - T_1^{(1)}| = |0,90452 - 0,90450| = 0,2 \cdot 10^{-4} < \varepsilon.$$

Дакле, постигнута је тражена тачност, па стајемо са израчунавањем. Вредност интеграла приближно је једнака вредности $T_2^{(2)}$:

$$I \approx 0,90452.$$

■

Задаци за вежбу

46. Израчунати интеграл

$$I = \int_0^1 \frac{e^x}{1+x^2}\, dx$$

са тачношћу $\varepsilon = 0,5 \cdot 10^{-4}$ применом Ромбергове интеграције.

Резултат. $I \approx 1,2707$

47. Израчунати интеграл

$$I = \int_0^{\frac{\pi}{4}} x \mathrm{tg}^{\,2} x\, dx$$

са тачношћу $\varepsilon = 0,5 \cdot 10^{-4}$ применом Ромбергове интеграције.

Резултат. $I \approx 0,1304$

48. Израчунати интеграл

$$I = \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \frac{\sin x}{x}\, dx$$

са тачношћу $\varepsilon = 0,5 \cdot 10^{-4}$ применом Ромбергове интеграције.

Резултат. $I \approx 0,6118$

6.3 Формуле за нумеричку интеграцију

За израчунавање одређеног интеграла $\int_a^b w(x)\, f(x)\, dx$ са тежинском функцијом $w(x)$, формирамо квадратурне формуле облика

$$\int_a^b w(x)\, f(x)\, dx = \sum_{i=0}^{n} A_i\, f(x_i) + R(f), \tag{6.11}$$

где су $a \leq x_0 < x_1 < \ldots < x_n \leq b$ различити, унапред познати чворови и $A_0, A_1, \ldots, A_n$ коефицијенти које треба одредити тако да формула (6.11) буде тачна за функције $f(x) = 1, x, x^2 \ldots, x^n$ (самим тим је формула тачна и за полиноме до степена n). Другим речима, мора важити да је грешка формуле

$$R(f) = \int_a^b w(x)\, f(x)\, dx - \sum_{i=0}^{n} A_i\, f(x_i)$$

једнака нули за функције $f(x) = 1, x, x^2 \ldots, x^n$. На овај начин, добијамо систем $(n+1)$ једначине чијим решавањем одређујемо $n+1$ коефицијент $A_0, \ldots, A_n$.

Може се десити да је формула (6.11) тачна и за полиноме вишег степена од n. Претпоставимо да је грешка формуле једнака нули и за функције $x^{n+1}, \ldots, x^{n+k}$, а различита од нуле за x^{n+k+1}. Тада кажемо да је *степен тачности формуле* $n+k$, а *ред грешке* $n+k+1$. Оцена грешке квадратурне формуле (6.11), чији је ред $n+k+1$, дата је са

$$R(f) \le \frac{M_{n+k+1}}{(n+k+1)!} \left| \int_a^b w(x)\, (x-x_0)^{k+1}(x-x_1)\cdots(x-x_n)\, dx \right|. \quad (6.12)$$

За еквидистантне чворове поделе интервала интеграције и тежинску функцију $w(x) = 1$ добијамо Њутн-Коутсове квадратурне формуле.

27. Извести формулу за нумеричку интеграцију облика

$$\int_0^h f(x)\, dx = Af(0) + Bf\left(\frac{2h}{3}\right) + R(f)$$

за $h > 0$, тако да буде тачна за полиноме што већег степена, и проценити грешку интеграције $R(f)$.

Решење. Одредимо коефицијенте A и B тако да формула буде тачна за функције 1 и x, тј. за полиноме до првог степена.

$$\begin{aligned} f(x) = 1: &\qquad A + B = h \\ f(x) = x: &\qquad \frac{h^2}{2} = \frac{2h}{3}B \quad \Rightarrow \quad \boxed{B = \tfrac{3}{4}h} \\ A = h - B &\quad \Rightarrow \quad \boxed{A = \tfrac{1}{4}h} \end{aligned}$$

Овим смо извели квадратурну формулу за нумеричку интеграцију

$$\int_0^h f(x)\, dx \approx \frac{h}{4} f(0) + \frac{3h}{4} f\left(\frac{2h}{3}\right)$$

и тачна је за полиноме до првог степена. Проверимо да ли је формула тачна и за полиноме вишег степена. За $f(x) = x^2$ лева страна формуле је

$$\int_0^h x^2\,dx = \frac{h^3}{3},$$

док је десна страна

$$\frac{h}{4}\cdot 0^2 + \frac{3h}{4}\left(\frac{2h}{3}\right)^2 = \frac{h^3}{3}.$$

С обзиром на то да су лева и десна страна једнаке, мора бити $R(x^2) = 0$, па је формула тачна и за полиноме другог степена. Проверимо да ли је формула тачна за функцију $f(x) = x^3$. Сада је лева страна формуле

$$\int_0^h x^3\,dx = \frac{h^4}{4},$$

а десна

$$\frac{h}{4}\cdot 0^3 + \frac{3h}{4}\left(\frac{2h}{3}\right)^3 = \frac{2h^4}{9}.$$

Из ове две релације следи да је $R(x^3) \neq 0$. Дакле, грешка ове квадратурне формуле је трећег реда. Оцена грешке изведене формуле је облика:

$$R(f) \leq \frac{M_3}{3!}\left|\int_0^h (x-0)^2\left(x-\frac{2h}{3}\right)\,dx\right| = \frac{M_3h^4}{216}.$$

$\blacksquare$

28. Извести формулу за нумеричку интеграцију облика

$$\int_{-1}^1 f(x)\,dx = A_1f(-1) + A_2f\left(-\frac{1}{\sqrt{5}}\right) + A_3f\left(\frac{1}{\sqrt{5}}\right) + A_4f(1) + R(f)$$

тако да буде тачна за полиноме што већег степена и одредити ред грешке $R(f)$. Применом добијене формуле израчунати приближну вредност интеграла $I = \displaystyle\int_0^1 e^{-x^2}dx$.

Решење. Одредимо коефицијенте A_i тако да формула буде тачна за функције $1, x, x^2$ и x^3, тј. за полиноме до трећег степена.

$$
\begin{aligned}
f(x) = 1: &\qquad A_1 + A_2 + A_3 + A_4 = \int_{-1}^{1} dx = 2 \\
f(x) = x: &\qquad -A_1 - \frac{1}{\sqrt{5}}A_2 + \frac{1}{\sqrt{5}}A_3 + A_4 = \int_{-1}^{1} xdx = 0 \\
f(x) = x^2: &\qquad A_1 + \frac{1}{5}A_2 + \frac{1}{5}A_3 + A_4 = \int_{-1}^{1} x^2dx = \frac{2}{3} \\
f(x) = x^3: &\qquad -A_1 - \frac{1}{5\sqrt{5}}A_2 + \frac{1}{5\sqrt{5}}A_3 + A_4 = \int_{-1}^{1} x^3dx = 0
\end{aligned}
$$

Решавањем овог линеарног система једначина добијамо коефицијенте

$$A_1 = A_4 = \frac{1}{6}, \qquad A_2 = A_3 = \frac{5}{6}.$$

Овим смо извели квадратурну формулу за нумеричку интеграцију

$$\int_{-1}^{1} f(x)\,dx \approx \frac{1}{6}f(-1) + \frac{5}{6}f\left(-\frac{1}{\sqrt{5}}\right) + \frac{5}{6}f\left(\frac{1}{\sqrt{5}}\right) + \frac{1}{6}f(1)$$

Проверимо да ли је формула тачна и за полиноме вишег степена. За $f(x) = x^4$ лева страна формуле је

$$\int_{-1}^{1} x^4\,dx) = \frac{2}{5},$$

док је десна страна $\frac{2}{5}$. С обзиром на то да су лева и десна страна једнаке, мора бити $R(x^4) = 0$, па је формула тачна и за полиноме четвртог степена.

Проверимо да ли је формула тачна за функцију $f(x) = x^5$. Лева страна формуле је $\int_{-1}^{1} x^5\,dx = 0$, и десна страна такође је једнака 0.

Проверимо да ли је формула тачна за функцију $f(x) = x^6$. Сада је лева страна формуле $\int_{-1}^{1} x^6\,dx) = \frac{2}{7}$, а десна $\frac{26}{75}$. Одатле следи да је $R(x^6) \neq 0$. Дакле, ред грешке је 6.

С обзиром на то да се изведена формула односи на интервал интеграције $[-1, 1]$, да бисмо могли да применимо ову формулу за решавање интеграла I, потребно је да уведемо одговарајућу смену. Уобичајено се користи линеарна смена, у овом случају то је $x = \frac{1}{2}t + \frac{1}{2}$.

$$
\begin{aligned}
I &= \int_0^1 f(x)dx = \frac{1}{2}\int_{-1}^1 f\left(\frac{1}{2}t + \frac{1}{2}\right)dt \\
&\approx \frac{1}{2}\left(\frac{1}{6}f(0) + \frac{5}{6}f\left(\frac{1}{2} - \frac{1}{2\sqrt{5}}\right) + \frac{5}{6}f\left(\frac{1}{2} + \frac{1}{2\sqrt{5}}\right) + \frac{1}{6}f(1)\right) \\
&\approx 0,746836.
\end{aligned}
$$

■

29. Извести формулу за нумеричку интеграцију облика

$$\frac{1}{\sqrt{2h}}\int_0^{2h} \frac{f(x)}{\sqrt{x}}\,dx = A_1 f(0) + A_2 f(h) + A_3 f(2h) + R(f)$$

тако да буде тачна за полиноме што већег степена и одредити ред грешке $R(f)$.

Решење. Одредимо коефицијенте A_i тако да формула буде тачна за функције $1, x$ и x^2, тј. за полиноме до другог степена.

$$
\begin{aligned}
f(x) = 1: &\qquad A_1 + A_2 + A_3 = \frac{1}{\sqrt{2h}}\int_0^{2h} \frac{dx}{\sqrt{x}} = 2 \\
f(x) = x: &\qquad A_2 + 2\,A_3 = \frac{1}{\sqrt{2h}}\int_0^{2h} \sqrt{x}\,dx = \frac{4}{3} \\
f(x) = x^2: &\qquad A_2 + 4\,A_3 = \frac{1}{\sqrt{2h}}\int_0^{2h} x^{3/2}\,dx = \frac{8}{5}
\end{aligned}
$$

Решавањем овог линеарног система једначина добијамо коефицијенте

$$A_1 = \frac{4}{5}, \qquad A_2 = \frac{16}{15}, \qquad A_3 = \frac{2}{15}.$$

Овим смо извели квадратурну формулу за нумеричку интеграцију

$$\frac{1}{\sqrt{2h}}\int_0^{2h} \frac{f(x)}{\sqrt{x}}\,dx \approx \frac{4}{5}\,f(0) + \frac{16}{15}\,f(h) + \frac{2}{15}f(2h). \tag{6.13}$$

Провером за функцију $f(x) = x^3$ закључујемо да формула није тачна за полиноме степена већег од 2, па је ред грешке 3.

■

Задаци за вежбу

49. Извести формулу за нумеричку интеграцију облика

$$\int_0^{+\infty} e^{-x} f(x)\,dx = Af(0) + Bf(1) + Cf(2) + R(f)$$

тако да буде тачна за полиноме што већег степена и проценити грешку интеграције $R(f)$.

Резултат. $A = C = \frac{1}{2}, B = 0, \quad R(F) \leq \frac{M_3}{3}$

50. Извести формулу за нумеричку интеграцију облика

$$\int_{-1}^{1} f(x)\,dx = Af\left(-\sqrt{\frac{3}{5}}\right) + Bf(0) + Cf\left(\sqrt{\frac{3}{5}}\right) + R(f)$$

тако да буде тачна за полиноме што већег степена и проценити грешку интеграције $R(f)$.

Резултат. $A = C = \frac{5}{9}, B = -\frac{8}{9}, \quad R(F) \leq 6.4 \cdot 10^{-5} M_6$

51. Извести формулу за нумеричку интеграцију облика

$$\int_0^1 f(x)\,dx = Af(1/4) + Bf(1/2) + Cf(3/4) + R(f)$$

тако да буде тачна за полиноме што већег степена и проценити грешку интеграције $R(f)$.

Резултат. $A = C = \frac{2}{3}, B = -\frac{1}{3}, \quad R(F) \leq 3 \cdot 10^{-4} M_4$

6.4 Решавање несвојствених интеграла

Решавамо несвојствени интеграл на бесконачном интервалу

$$\int_a^{+\infty} f(x)\,dx$$

са тачношћу ε. Како је у питању конвергентан интеграл, за $\frac{\varepsilon}{2}$ постоји $M > a$ тако да је вредност „репа” интеграла „довољно” мала, тј. да је испуњено

$$\left| \int_M^{+\infty} f(x)\,dx \right| \leq \frac{\varepsilon}{2}.$$

Сада рашавамо одређени интеграл

$$\int_a^M f(x)\,dx$$

са тачношћу $\frac{\varepsilon}{2}$.

Сличан поступак примењујемо и у случају када решавамо интеграле облика $\int_{-\infty}^{b} f(x)\,dx$ или $\int_{-\infty}^{+\infty} f(x)\,dx$.

30. Израчунати интеграл применом Ромбергове интеграције

$$I = \int_1^\infty \frac{xe^{-x^2}}{2+\sin x}\,dx.$$

са тачношћу $\varepsilon = 10^{-5}$.

Решење. Одредимо $M > 1$ тако да је:

$$\left|\int_M^\infty \frac{xe^{-x^2}}{2+\sin x}\,dx\right| \le \frac{1}{2}\varepsilon = \frac{1}{2}\cdot 10^{-5}.$$

Како је $|2+\sin x| \ge 1$, имамо да важи

$$\left|\int_M^\infty \frac{xe^{-x^2}}{2+\sin x}\,dx\right| \le \int_M^\infty \left|\frac{xe^{-x^2}}{2+\sin x}\right| dx \le \int_M^\infty xe^{-x^2}\,dx$$

$$= -\frac{1}{2}e^{-x^2}\Big|_M^\infty = \frac{1}{2}e^{-M^2} \le \frac{1}{2}\cdot 10^{-5}.$$

Дакле, тражимо M за које важи да је

$$e^{-M^2} \le 10^{-5} \implies M > 3,3930702$$

Узмимо да је $M = 3,4$. Сада интеграл $\int_1^{3,4} \frac{xe^{-x^2}}{2+\sin x}\,dx$ израчунавамо са тачношћу $\frac{1}{2}\cdot 10^{-5}$. Грешка заокруживања E при израчунавању интеграла је $E = (3,4-1)\frac{1}{2}\cdot 10^{-6} = 1,2\cdot 10^{-6}$, па можемо да вршимо заокруживање на 6 децимала.

Интеграл решавамо примењујући Ромбергову интеграцију, прво са кораком $h_0 = 1,2$, а затим са кораком $h_1 = 0,6$, за $f(x) = \frac{xe^{-x^2}}{2+\sin x}$.

x	f_0, f_m	f_k
1	$0,129468$	
$2,2$		$0,006194$
$3,4$	$0,000019$	
$\sum$	$0,129487$	$0,006194$

Применом трапезне формуле имамо да је

$$T_0^{(0)} = \frac{1,2}{2}(0,129468 + 0,000019 + 2 \cdot 0,006194) = 0,085125.$$

За преполовљен корак $h_1 = 0,6$ имамо да је:

x	f_k
$1,6$	$0,041235$
$2,8$	$0,000472$
$\sum$	$0,041707$

$$\begin{aligned} T_1^{(0)} &= \frac{0,6}{2}(0,129468 + 0,000019 + 2 \cdot (0,006194 + 0,041707)) \\ &= 0,067587. \end{aligned}$$

Применимо Ричардсонову екстраполацију на претходно израчунате апроксимације:

$$T_1^{(1)} = \frac{4T_1^{(0)} - T_0^{(0)}}{4-1} = 0,061741$$

Настављајући поступак, применом алгоритма Ромбергове интеграције, користећи формуле (6.10),

$$T_k^{(j)} = \frac{4^j T_k^{(j-1)} - T_{k-1}^{(j-1)}}{4^j - 1}, \qquad (j = 1, \ldots, k;\ k = 1, 2, \ldots)$$

формирамо следећу T-табелу:

$$\begin{array}{llll} T_0^{(0)} = 0,085125 & & & \\ T_1^{(0)} = 0,067587 & 0,061741 & & \\ T_2^{(0)} = 0,063897 & 0,062667 & 0,062729 & \\ T_3^{(0)} = 0,063017 & 0,062724 & 0,062728 & 0,062728 \end{array}$$

Процена грешке је

$$|T_3^{(3)} - T_2^{(2)}| = |0,062728 - 0,062729| = 0,000001,$$

што је мање од захтеване тачности.

Вредност интеграла I је приближно једнака $T_3^{(3)}$,

$$I \approx 0,062728.$$

■

31. Са тачношћу $\varepsilon = 10^{-2}$ израчунати интеграл

$$I = \int_0^1 \frac{\cos x}{\sqrt{x}}\, dx.$$

Решење. У питању је несвојствени интеграл, пошто подинтегрална функција има сингуларитет на интервалу интеграције:

$$\lim_{x \to 0+} \frac{\cos x}{\sqrt{x}} = +\infty.$$

Овај интеграл можемо решити на више начина.

I. У питању је несвојствени интеграл неограничене функције на интервалу интеграције, у околини левог краја интервала интеграције. Како је у питању конвергентан интеграл, можемо да одредимо r тако да одбацивањем дела интеграла направимо грешку која није већа од $\frac{\varepsilon}{2}$, тј. тако да је

$$\left| \int_0^r \frac{\cos x}{\sqrt{x}}\, dx \right| \le \int_0^r \frac{|\cos x|}{\sqrt{x}}\, dx \le \int_0^r \frac{1}{\sqrt{x}}\, dx = 2\sqrt{r} \le \frac{\varepsilon}{2}.$$

Дакле, потребно је да је $r \le 6,25 \cdot 10^{-6}$. Можемо узети да је $r = 6 \cdot 10^{-6}$. Сада решавамо одређени интеграл $\int_r^1 \frac{\cos x}{\sqrt{x}}\, dx$ са тачношћу $0,5 \cdot 10^{-2}$, на пример Симпсоновом формулом.

II. У 29. задатку извели смо формулу (6.13) за нумеричку интеграцију коју можемо да применимо на решавање интеграла I за $h = \frac{1}{2}$ и $f(x) = \cos x$, тј.

$$\int_0^1 \frac{\cos x}{\sqrt{x}}\, dx \approx \frac{4}{5}\cos 0 + \frac{16}{15}\,\cos 0,5 + \frac{2}{15}\cos 1 = 1,81.$$

III. Интеграл I можемо да израчунамо увођењем смене $t = \sqrt{x}$, којом се полазни интеграл своди на својствени интеграл:

$$I = 2\int_0^1 \cos t^2 \, dt$$

чију смо вредност одредили у 26. задатку са већом тачношћу.

На сва три начина добијамо исту вредност интеграла I у односу на захтевану тачност.

■

Задаци за вежбу

52. Одговарајућом нумеричком методом израчунати интеграл

$$I = \int_5^{+\infty} \frac{\sin\frac{1}{x}}{1+x^3} \, dx$$

са тачношћу $\varepsilon = 10^{-4}$.

Резултат. $I \approx 0,0027$

53. Одговарајућом нумеричком методом израчунати интеграл

$$I = \int_0^{+\infty} \frac{1}{(1+x^2)^3} \, dx$$

са тачношћу $\varepsilon = 10^{-3}$.

Резултат. $I \approx 0,589$

Литература

[1] Acton F.S. *Numerical methods that work*, Mathematical Association of America, USA (1990) p. 549

[2] Ascher U.M., Greif C. *A First Course in Numerical Methods*, Society for Industrial and Applied Mathematics, USA (2011) p. 580

[3] Burden R., Faires J.D. *Numerical Analysis*, 9th Edition, International Edition, Cengage Learning, Inc. USA (2011) p. 888

[4] Chapra S.C., *Applied Numerical Methods with MATLAB*, 3rd Edition, McGraw-Hill Education - Europe, (2011) p. 672

[5] Cheney W., Kincaid D. *Numerical Mathematics and Computing*, 6th Edition, Thomson Brooks/Cole, USA (2008) p. 763

[6] Davis P.J., Rabinowitz P *Methods of numerical integration*, Academic Press, inc. (1984) p. 625

[7] Делић А., Дражић З., Живановић С., Ивановић М., *Збирка решених задатака из увода у нумеричку математику*, Математички факултет, Београд, Србија (2017), стр. 167

[8] Gilat A., Subramaniam V. *Numerical Methods for Engineers and Scientists : An Introduction with Applications Using MATLAB*, 3rd Edition, John Wiley & Sons Inc, USA (2013) p. 576

[9] Јешић С., *Нумеричка математика*, скрипта, (2001) стр. 40

[10] Kahaner D., Moler C., Nash S., *Numerical methods and software*, Prentice Hall, USA (1989), p. 504

[11] Радуновић Д., *Нумеричка анализа*, Академска мисао, Београд, Србија (2003), стр. 240

[12] Sauer T., *Numerical Analysis*, Pearson Education, USA (2011), str. 672

[13] Тошић Д., *Увод у нумеричку анализу: са збирком задатака и проблема*, Академска мисао, Београд, Србија (2004), стр. 355

CIP - Каталогизација у публикацији - Народна библиотека Србије, Београд

519.61/.65(075.8)

ЋИРОВИЋ, Наташа, 1978-
Нумеричка математика / Наташа Ћировић. - Београд : Универзитет, Електротехнички факултет, 2018 (Београд : Н. Ћировић). - VI, 138 стр. : граф. прикази, табеле ; 30 cm

Тираж 50. - Библиографија: стр. 137-138.

ISBN 978-86-7225-068-8

а) Нумеричка анализа
COBISS.SR-ID 270994444