Numerička matematika

Skripta iz vježbi

4. lipnja 2025.

***Disclaimer**: Ovo je radna verzija skripte iz vježbi za kolegij Numerička matematika nastala u proljeće 2024. godine. Svjesni smo da ima grešaka, tako da je njezino korištenje na "vlastitu odgovornost". Sve pronađene greške možete javiti autorima skripte kako bi se u sljedećim verzijama ispravile.*

Sadržaj

1

Greške, stabilnost, uvjetovanost

1.1 Tipovi grešaka

Neke od grešaka koje se događaju prilikom numeričkog rješavanja nekog problema su

- greške modela,
- greške u ulaznim podacima,
- greške numeričkih metoda.

Greške numeričkih metoda se dijele u dvije kategorije: *greške diskretizacije* i *greške odbacivanja*. Kao primjer greške odbacivanja, na predavanjima je analizirano korištenje Taylorovog reda u svrhu računanja aproksimacije funkcije u nekoj točki.

1.2 Greške približnog računanja i aritmetike računala

Realni brojevi u računalu su zapisani u binarnom brojevnom sustavu (standard IEEE-754). Pri tome je zapis određen sa tri veličine: predznakom, eksponentom i mantisom. Duljinu eksponenta u bitovima označavamo s w, a duljinu u bitovima mantise sa t.

Aproksimacija broja $x = \pm(1.b_{-1}b_{-2}\ldots)_2 \cdot 2^e$ spremljena u računalu se označava sa $f\ell(x) = \pm(1.b_{-1}b_{-2}\ldots b_{-t})_2 \cdot 2^e$.

Za **apsolutnu grešku** napravljenu prilikom spremanja broja x u računalo vrijedi

$$|x - f\ell(x)| \leq 2^e \cdot 2^{-(t+1)},$$

dok za **relativnu grešku** vrijedi sljedeća ocjena

$$\frac{|x - f\ell(x)|}{|x|} \leq 2^{-(t+1)} =: u.$$

u zovemo **jedinična greška zaokruživanja**.

1.2.1 Širenje grešaka u egzaktnoj aritmetici

Neka su $x, y \in \mathbb{R}$ i neka su $\tilde{x}, \tilde{y}$ odgovarajuće aproksimacije koje imaju malu relativnu grešku, tj.

$$\tilde{x} = (1 + \varepsilon_x)x, \;\; \tilde{y} = (1 + \varepsilon_y)y,$$

za malene ε_x, ε_y.

Zanima nas što se događa sa konačnim rezultatom ako u računu koristimo perturbirane podatke:

- **množenje** (bezopasno):

$$\tilde{x} * \tilde{y} = (x * y)(1 + \varepsilon_*), \;\; \varepsilon_* \approx \varepsilon_x + \varepsilon_y;$$

- **dijeljenje** (bezopasno):

$$x/y = (x/y)(1 + \varepsilon_/), \;\; \varepsilon_/ \approx \varepsilon_x - \varepsilon_y;$$

- **zbrajanje i oduzimanje** (potencijalno opasno):

$$x + \tilde{y} = (x + y)(1 + \varepsilon_+), \;\; \varepsilon_+ = \frac{x}{x+y}\varepsilon_x + \frac{y}{x+y}\varepsilon_y.$$

  - x i y istog predzanaka (bezopasno):

$$|\varepsilon_+| \leq \max\{|\varepsilon_x|, |\varepsilon_y|\};$$

  - x i y različitog predznaka (**katastrofalno kraćenje**):

$$|\varepsilon_+| \leq \frac{|x|}{|x+y|}|\varepsilon_x| + \frac{|y|}{|x+y|}|\varepsilon_y|.$$

Primjer 1.1. Aproksimirajmo broj e^{-10} početnim komadom Taylorovog reda oko nule.

Rješenje. Broj e^{-10} računamo putem početnog komada Taylorovog reda oko nule

$$e^x \approx \sum_{k=0}^{n} \frac{x^k}{k!} =: p(x),$$

sve dok posljednji član (po apsolutnoj vrijednosti) u zbroju ne padne ispod neke zadane točnosti ε

$$\left| \frac{x^{n+1}}{(n+1)!} \right| < \varepsilon.$$

Na predavanjima je izvedeno da tada za grešku odbacivanja vrijedi

$$|R_{n+1}(-10)| \leq e^{-10}\varepsilon, \;\; \frac{e^{-10} - p(-10)}{e^{-10}} \leq \varepsilon.$$

Dakle, ovakvom aproksimacijom osiguravamo malu relativnu grešku.

Pogledajmo članove Taylorovog polinoma kojim aproksimiramo vrijednost e^{-10}:

$$e^{-10} = 1 - \frac{10}{1!} + \frac{10^2}{2!} - \frac{10^3}{3!} + \ldots + (-1)^n \frac{10^n}{n!}.$$

Rezultat koji očekujemo je mali broj. S druge strane, članovi polinoma različitog predznaka i prvo rastu po apsolutnoj vrijednosti a onda padaju. Dakle, rezultat možemo dobiti oduzimanjem relativno velikih brojeva pa dolazi do katastrofalnog kraćenja.

Iako smo analizom metode zaključili da je relativna greška mala, u aritmetici računala ćemo dobiti potpuno krivi rezultat. Taj rezultat nije posljedica greške numeričke metode niti greške odbacivanja nego je posljedica grešaka aritmetike računala. △

Gornji račun možemo provjeriti u pythonu, koristimo ažurirani kod s predavanja:

```
import math;
import matplotlib.pyplot as plt;

def taylor_exp(x, eps):
    clanovi = [];
    parcijalne_sume = [];
    suma = 0.0;
    clan = 1;
    k = 0;
    while (abs(clan) > eps):
        suma = suma + clan;
        clanovi.append(clan);
        parcijalne_sume.append(suma);
        k = k + 1;
        clan = clan * x / k;

    return (suma, clanovi, parcijalne_sume);

x0 = -10;
(suma, clanovi, parcijalne_sume) = taylor_exp(-x0, 5e-16);
egzaktna_vrijednost = math.exp(x0);

print(f'Funkcija exp iz Pythona vraca (egzaktno): {egzaktna_vrijednost:.
                                    16e}.');
print(f'Vrijednost izracunata pomocu Taylorovog polinoma: {suma:.16e}.');
print(f'Apsolutna greska: {abs(suma - egzaktna_vrijednost):.16e}.');
print(f'Relativna greska: {abs(suma - egzaktna_vrijednost) / abs(
                                    egzaktna_vrijednost):.16e}.');
print(f'Taylorov polinom koristi {len(clanovi)} clanova.');
print(f'Najveci clan: {max([abs(c) for c in clanovi]):.5e}.');
print(f'Teorija daje da je greska odbacivanja manja od: {math.exp(x0) *
                                    5e-16:.5e}.');
```

Rezultat tog koda bi bio:

```
Funkcija exp iz Pythona vraca (egzaktno):  4.5399929762484854e-05.
Vrijednost izracunata pomocu Taylorovog polinoma:  2.2026465794806711e+04.
Apsolutna greska:  2.2026465749406780e+04.
```

```
Relativna greska:  4.8516519440979010e+08.
Taylorov polinom koristi 52 clanova.
Najvećci clan:  2.75573e+03.
Teorija daje da je greska odbacivanja manja od:  2.27000e-20.
```

1.3 Uvjetovanost skalarnih funkcija

Apsolutna uvjetovanost funkcije $f : \mathbb{R} \to \mathbb{R}$ klase C^2 je

$$\kappa_f^{\text{abs}}(x) = \left|f'(x)\right|,$$

a njena **relativna uvjetovanost** (za $x \neq 0$) je

$$\kappa_f^{\text{rel}}(x) = \left|\frac{xf'(x)}{f(x)}\right|.$$

Primjer 1.2. Zadana je funkcija $f(x) = x - 2$. Odredite apsolutnu i relativnu uvjetovanost te funkcije. Što možete zaključiti o točnosti rezultata za razne $x \in \mathbb{R}$?

Rješenje.

$$f'(x) = 1.$$

$$\kappa_f^{rel}(x) = \left|\frac{xf'(x)}{f(x)}\right| = \left|\frac{x}{x-2}\right|.$$

$$\kappa_f^{abs}(x) = |f'(x)| = 1.$$

U apsolutnom smislu izvrednjavanje ove funkcije nije osjetljivo ni za koji $x \in \mathbb{R}$, no u relativnom smislu kada $x \to 2$, relativna uvjetovanost problema raste, pa time zaključujemo da je problem osjetljiv kada je x blizu 2, tj. prilikom izračuna funkcije f za x koji su blizu 2 možemo očekivati veliku relativnu grešku. △

Zaključak iz gornjeg zadatka u skladu je s *opasnim kraćenjem*: prilikom oduzimanja dva bliska broja u računalu dolazi do velike relativne pogreške.

Zadatak 1.3 (1. kolokvij 2016.)**.** Za kvadratnu jednadžbu $x^2 - 8x + q = 0$ neka je f funkcija koja realnom koeficijentu q pridružuje veće realno rješenje te kvadratne jednadžbe. Neka je domena f svi parametri q za koje ta jednadžba ima realna rješenja. Odredite apsolutnu i relativnu uvjetovanost preslikavanja f i komentirajte za koje q je račun funkcije f loše uvjetovan u apsolutnom i relativnom smislu.

Rješenje. Rješenje kvadratne jednadžbe su dana formulom

$$x_{1,2} = \frac{8 \pm \sqrt{64 - 4q}}{2} = 4 \pm \sqrt{16 - q}.$$

Domena funkcije f određena je diskriminantom: $q \leq 16$. Veće rješenje dobivamo kada u gornjoj formuli uzmemo "+". Dakle, dobili smo funkciju $f : \langle -\infty, 16] \to \mathbb{R}$ definiranu s

$$f(q) = 4 + \sqrt{16 - q}.$$

Derivacija te funkcije je

$$f'(q) = \frac{1}{2\sqrt{16-q}}.$$

Sada je

$$\kappa_f^{\text{abs}}(q) = |f'(x)| = \frac{1}{2\sqrt{16-q}},$$
$$\kappa_f^{\text{rel}}(q) = \frac{|q|}{2(\sqrt{16-q})(4+\sqrt{16-q})}.$$

Obje uvjetovanosti mogu težiti u $+\infty$ samo ako im nazivnici idu u 0. Kako je izraz $4+\sqrt{16-q}$ ograničen odozdo s 4, obje uvjetovanosti imaju problem samo ako $\sqrt{16-q} \to 0$, a to je ako i samo ako $q \to 16^-$. Samo za takve q je problem loše uvjetovan. △

Zadatak 1.4 (1. kolokvij 2017., ažuriran)**.** Promotrimo funkciju

$$f(x) = 1 - \sqrt{1-x}.$$

Pretpostavimo da želimo izračunati vrijednost $f(x)$ u realnoj aritmetici računala, uz dodatnu pretpostavku da za sve y postoji neki $|\alpha_y| < \varepsilon$ takav da vrijedi $f\ell(\sqrt{x}) = \sqrt{x}(1+\alpha_y)$. Ovdje je ε jedinična greška zaokruživanja.

- Izvedite izraz za relativnu grešku izračunate vrijednosti $f\ell(f(x))$ u odnosu na egzaktnu vrijednost $f(x)$ koristeći gornju formulu i objasnite što se događa kada $x \to 0$.
- Izračunajte relativnu uvjetovanost za funkciju f u točki x i objasnite njezino ponašanje kada $x \to 0$. Je li "krivac" za lošu stabilnost iz prošlog dijela zadatka uvjetovanost funkcije?
- Predložite, ako je moguće, alternativni način za računanje $f(x)$ za male x i ukratko obrazložite zašto je takav način stabilan.

Rješenje. Kada je x blizu nule, očekujemo opasno kraćenje između članova 1 i $\sqrt{1-x}$. Potvrdimo to računom. Neka su $\varepsilon_1, \varepsilon_2, \varepsilon_3$ svi po apsolutnoj vrijednosti manji od ε. Tada vrijedi

$$\begin{aligned} f\ell(1-\sqrt{1-x}) &= \left(1 - \sqrt{(1-x)(1+\varepsilon_1)}\cdot(1+\varepsilon_2)\right)(1+\varepsilon_3) \\ &= (1-\sqrt{1-x})\left(1 - 1 + \frac{1-\sqrt{(1-x)(1+\varepsilon_1)}\cdot(1+\varepsilon_2)}{1-\sqrt{1-x}}\right)(1+\varepsilon_3) \\ &= (1-\sqrt{1-x})\left(1 + \underbrace{\frac{\sqrt{1-x}\,(1-\sqrt{1+\varepsilon_1}(1+\varepsilon_2))}{1-\sqrt{1-x}}}_{\text{neograničeno kad } x\to 0}\right)(1+\varepsilon_3). \end{aligned}$$

Naša hipoteza je potvrđena, kada $x \to 0$, relativna greška je nekontrolirana.

Izračunajmo relativnu uvjetovanost:

$$f'(x) = \frac{-1}{2\sqrt{1-x}},$$
$$\kappa_f^{\text{rel}}(x) = \left|\frac{xf'(x)}{f(x)}\right| = \left|\frac{x}{(2\sqrt{1-x})1-\sqrt{1-x})}\right|$$

Odredimo ponašanje kada $x \to 0$, a za to racionalizirajmo nazivnik:

$$\lim_{x\to 0} \kappa_f^{\text{rel}}(x) = \lim_{x\to 0} \left|\frac{x}{(2\sqrt{1-x})(1-\sqrt{1-x})}\right| \cdot \frac{1+\sqrt{1-x}}{1+\sqrt{1-x}}$$
$$= \lim_{x\to 0} \left|\frac{x(1+\sqrt{1-x})}{2(\sqrt{1-x})x}\right| = 1.$$

Dakle, krivac velike relativne greške nije u funkciji f nego u načinu njezinog izračuna.

Problem ćemo riješiti tako da izbjegnemo oduzimanje van korijena, a to ćemo postići racionalizacijom:

$$f(x) = (1-\sqrt{1-x}) \cdot \frac{1+\sqrt{1-x}}{1+\sqrt{1-x}} = \frac{x}{1+\sqrt{1-x}}.$$

△

Zadatak 1.5 (1. kolokvij 2021.)**.** Promotrimo funkciju

$$f(x) = \frac{1-\cos x}{\sin x}.$$

Pretpostavimo da želimo izračunati vrijednost $f(x)$ u realnoj aritmetici računala, uz dodatnu pretpostavku da za sve y postoje neki $|\alpha_y|, |\beta_y| < \varepsilon$ takvi da vrijedi $f\ell(\cos x) = \cos x(1+\alpha_y)$ i $f\ell(\sin x) = \sin x(1+\beta_y)$. Ovdje je ε jedinična greška zaokruživanja.

- Izvedite izraz za relativnu grešku izračunate vrijednosti $f\ell(f(x))$ u odnosu na egzaktnu vrijednost $f(x)$ koristeći gornju formulu i objasnite što se događa kada $x \to 0$.
- Izračunajte relativnu uvjetovanost za funkciju f u točki x i objasnite njezino ponašanje kada $x \to 0$. Je li "krivac" za lošu stabilnost iz prošlog dijela zadatka uvjetovanost funkcije?
- Predložite, ako je moguće, alternativni način za računanje $f(x)$ za male x i ukratko obrazložite zašto je takav način stabilan.

Rješenje. Očekujemo da će račun funkcije $f(x)$ prema gornjoj formuli biti loš za male x budući da u brojniku oduzimamo dva bliska broja, te cijeli rezultat dijelimo malim

brojem čime još više povećavamo tu grešku dobivenu opasnim kraćenjem. Uvjerimo se to i računom. Neka su $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$ svi po apsolutnoj vrijednosti manji od ε. Tada vrijedi

$$\begin{aligned} f\ell(f(x)) &= \frac{(1-\cos x(1+\varepsilon_1))(1+\varepsilon_2)}{\sin x(1+\varepsilon_3)}(1+\varepsilon_4) \\ &= f(x)\frac{(1+\varepsilon_2)(1+\varepsilon_4)}{1+\varepsilon_3}\frac{\sin x}{1-\cos x}\frac{1-\cos x(1+\varepsilon_1)}{\sin x} \\ &= f(x)\frac{(1+\varepsilon_2)(1+\varepsilon_4)}{1+\varepsilon_3}\frac{1-\cos x-\cos x\varepsilon_1)}{1-\cos x} \\ &= f(x)\frac{(1+\varepsilon_2)(1+\varepsilon_4)}{1+\varepsilon_3}\left(1+\frac{-\cos x\varepsilon_1}{1-\cos x}\right). \end{aligned}$$

Drugi sumand u zagradi se ne može ograničiti kada $x \to 0$, pa zato u tom slučaju dobivamo nekontroliranu relativnu grešku, odnosno račun je zaista nestabilan.

Izračunajmo relativnu uvjetovanost. Prvo odredimo derivaciju funkcije f:

$$f'(x) = \frac{\sin x \cdot \sin x - (1-\cos x)\cos x}{\sin^2 x} = \frac{1-\cos x}{\sin^2 x}.$$

Sada je

$$\kappa_f^{\text{rel}}(x) = \left|\frac{xf'(x)}{f(x)}\right| = \left|\frac{x}{\sin x}\right|.$$

Kad $x \to 0$, uvjetovanost teži u 1. Dakle, krivac za nestabilnost iz prvog dijela zadatka je način izračuna funkcije f, a ne funkcija sama po sebi.

Formulu ćemo poraviti tako da iskoristimo svojstva trigonometrijskih funkcija tako da izbjegnemo katastrofalno kraćenje u brojniku. Više je načina, recimo

$$f(x) = \frac{\sin(x/2)}{\cos(x/2)} \text{ ili } f(x) = \frac{\sin x}{1+\cos x}.$$

Sada su te formule stabilne kada je x blizu nule. *Nije bitno za zadatak, ali druga formula nije stabilna kada je x blizu $\frac{\pi}{2}$, ponovno zbog opasnog kraćenja, no tada možemo računati prema originalnoj formuli.* △

2

Linearni sustavi

2.1 LU faktorizacija

Primjer 2.1. Nađimo LU faktorizaciju matrice

$$A = \begin{bmatrix} 5 & 1 & 4 \\ 10 & 4 & 7 \\ -15 & 5 & -9 \end{bmatrix}$$

koristeći kompaktni zapis.

Rješenje. Na matrici provodimo Gaussove eliminacije: u k-tom koraku elementom na mjestu (k, k) poništavamo elemente ispod njega. Nova verzija matrice A na tim mjestima ispod dijagonale imat će nule. Kako želimo pamtiti i multiplikatore kojima smo poništili te retke, upravo ta mjesta iskoristit ćemo za pamćenje tih multiplikatora. To je u duhu tzv. **kompaktnog zapisa**.

Počinjemo se matricom

$$A^{(1)} = A.$$

Elementom na mjestu $(1, 1)$ (koji iznosi 5) poništavamo preostale elemente u prvom stupcu ispod njega Gaussovim eliminacijama. Za to je potrebno pomnožiti drugi redak s $-\frac{10}{5} = -2$, te treći redak s $-\frac{-15}{5} = 3$. U novom stanju matrice A u prvom stupcu ispod dijagonale pisat će nule. U kompaktnom zapisu to koristimo tako da tamo ne napišemo eksplicitno te nule, nego ta mjesta iskoristimo za zapis iskorištenih multiplikatora. Te multiplikatore zapisujemo u matrici $A^{(2)}$ i odijeljujemo ih crtama:

$$A^{(2)} = \left[\begin{array}{ccc} 5 & 1 & 4 \\ \hline -2 & 2 & -1 \\ 3 & 8 & 3 \end{array}\right].$$

Preostali elementi u matrici $A^{(2)}$ u drugom i trećem retku dobiveni su na standardan način Gaussovim eliminacijama. Nastavljamo s ovom matricom dalje: elementom na

mjestu $(2,2)$ (koji iznosi 2) poništavamo preostale elemente u drugom stupcu ispod njega, a to je samo element na mjestu $(3,2)$. Njega poništimo multiplikatorom $-\frac{8}{2} = -4$. Ponovno taj multiplikator pamtimo na mjestu na kojem bi u novoj matrici A pisala nula. Nakon ažuriranja trećeg retka dobivamo novu matricu.

$$A^{(3)} = \left[\begin{array}{ccc} 5 & 1 & 4 \\ \hline -2 & 2 & -1 \\ 3 & -4 & 7 \end{array}\right].$$

Kada je proces završen, iz gornjeg trokuta (uključujući i dijagonalu) zadnje matrice $A^{(3)}$ dobivamo gornjetrokutastu matricu U, a iz donjeg trokuta stvaramo donjetrokutastu matricu L tako da svim elementima promijenimo predznak, te na dijagonalu stavimo jedinice. Dobivamo

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & 4 & 1 \end{bmatrix} \text{ i } U = \begin{bmatrix} 5 & 1 & 4 \\ 0 & 2 & -1 \\ 0 & 0 & 7 \end{bmatrix}.$$

Provjerite da zaista vrijedi $A = LU$. △

Definicija 2.2. Neka je A kvadratna matrica reda n. Kažemo da je prikaz $A = LU$ **LU faktorizacija** matrice A ako je L donjetrokutasta matrica s jedinicama na dijagonali, a U gornjetrokutasta matrica, pri čemu su obje matrice L i U kvadratne reda n.

Pokazuje se da ne mora svaka matrica imati LU faktorizaciju, te da ako neka matrica i ima LU faktorizaciju, da rješavanje sustava ne mora biti stabilno. Rješenje za to je **LU faktorizacija s parcijalnim pivotiranjem**: prije svakog eliminiranja elemenata ispod dijagonale, na dijagonalno mjesto (k,k) dovedemo onaj element koji je najveći po apsolutnoj vrijednosti u tom stupcu ispod dijagonale (ili na dijagonali).

Na predavanjima se pokazuje da se taj algoritam može uvijek provesti.

Teorem 2.3. Neka je A kvadratna matrica reda n. Tada postoje gornjetrokutasta matrica U, donjetrokutasta matrica s jedinicama na dijagonali L i permutacijska matrica P, sve reda n, takve da je $PA = LU$. Ovaj zapis zovemo **LU faktorizacija s parcijalnim pivotiranjem**

Primjer 2.4. Riješimo sustav

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

koristeći LU s parcijalnim pivotiranjem i kompaktni zapis.

Rješenje. Iz sustava čitamo:

$$A = \begin{bmatrix} 2 & -4 & 1 \\ 1 & -2 & 2 \\ 3 & -2 & 1 \end{bmatrix} \text{ i } b = \begin{bmatrix} 6 \\ 3 \\ 5 \end{bmatrix}.$$

Kako 2 nije najveći element po apsolutnoj vrijednosti u prvom stupcu, nego je to 3 koji se nalazi u trećem retku, prvo moramo zamijeniti prvi i treći redak. To radimo množenjem slijeva s matricom permutacije $P^{(1)} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$. Dobivamo matricu

Računamo LU faktorizaciju s parcijalnim pivotiranjem

$$\check{A}^{(1)} = P^{(1)}A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & -4 & 1 \\ 1 & -2 & 2 \\ 3 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -2 & 1 \\ 1 & -2 & 2 \\ 2 & -4 & 1 \end{bmatrix}.$$

Sada na toj matrici provodimo Gaussove eliminacije u kompaktnom zapisu kao u prošlom primjeru. Dobivamo

$$\hat{A}^{(1)} = \left[\begin{array}{ccc} 3 & -2 & 1 \\ \hline -1/3 & -4/3 & 5/3 \\ -2/3 & -8/3 & 1/3 \end{array}\right].$$

U drugom koraku u gornjoj matrici opet tražimo najveći element po apsolutnoj vrijednosti u drugom stupcu ispod ili na dijagonali. Od brojeva $-4/3$ i $-4/3$ drugi ima veću apsolutnu vrijednost, pa moramo zamijeniti drugi i treći redak. To radimo množeći slijeva matricom $P^{(2)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$. Množimo cijelu matricu $\hat{A}^{(1)}$ (uključujući i crtama odijeljene multiplikatore). Dobivamo

$$\check{A}^{(2)} = P^{(2)}\hat{A}^{(1)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \hat{A}^{(1)} = \left[\begin{array}{ccc} 3 & -2 & 1 \\ \hline -2/3 & -8/3 & 1/3 \\ -1/3 & -4/3 & 5/3 \end{array}\right].$$

Na kraju u matrici $\hat{A}^{(2)}$ poništavamo element ispod dijagonale, dopisujući novi multiplikator odijeljen crtama. Dobivamo

$$\hat{A}^{(2)} = \left[\begin{array}{ccc} 3 & -2 & 1 \\ \hline -2/3 & -8/3 & 1/3 \\ -1/3 & -1/2 & 3/2 \end{array}\right].$$

Matrice L i U dobivamo kao prije. Matricu P dobivamo tako da pomnožimo sve matrice $P^{(k)}$ obrnutim redoslijedom:

$$P = P^{(2)}P^{(1)} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad L = \begin{bmatrix} 1 & 0 & 0 \\ 2/3 & 1 & 0 \\ 1/3 & 1/2 & 1 \end{bmatrix} \text{ i } U = \begin{bmatrix} 3 & -2 & 1 \\ 0 & -8/3 & 1/3 \\ 0 & 0 & 3/2 \end{bmatrix}.$$

Možete se sami uvjeriti da zaista vrijedi $PA = LU$.

Kada dobijemo LU faktorizaciju s parcijalnim pivotiranjem za neku matricu A, sustav rješavamo u 2 koraka

1. Rješavamo donjetrokutasti sustav $Ly = Pb$ suptitucijom unaprijed (krećemo od gornjih redaka prema donjim):

$$\begin{bmatrix} 1 & 0 & 0 \\ 2/3 & 1 & 0 \\ 1/3 & 1/2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \\ 3 \end{bmatrix} \implies y = \begin{bmatrix} 5 \\ 8/3 \\ 0 \end{bmatrix},$$

2. Rješavamo donjetrokutasti sustav $Ux = y$ suptitucijom unazad (krećemo od donjih redaka prema gornjim):

$$\begin{bmatrix} 3 & -2 & 1 \\ 0 & -8/3 & 1/3 \\ 0 & 0 & 3/2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 8/3 \\ 0 \end{bmatrix} \implies x = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

△

2.2 Uvjetovanost linearnih sustava i stabilnost LU

Uvjetovanost matrice A definiramo kao

$$\kappa(A) = \|A\| \cdot \|A^{-1}\|,$$

gdje je $\|\cdot\|$ operatorska norma koja je inducirana vektorskom normom $\|\cdot\|$.

Vrijede sljedeći rezultati o uvjetovanosti rješavanja sustava:

- Neka je $A \in \mathbb{M}_n$ regularna kvadratna matrica te $b, \Delta b, x$ i Δx iz $\mathbb{R}^n$ takvi da vrijedi

$$Ax = b \text{ i } A(x + \Delta x) = b + \Delta b.$$

  Tada za $x \neq 0$ vrijedi

$$\frac{\|\Delta x\|}{\|x\|} \leq \kappa(A) \frac{\|\Delta b\|}{\|b\|}.$$

- Ako je dodatno $\Delta A \in \mathbb{M}_n$ takva da vrijedi

$$Ax = b \text{ i } (A + \Delta A)(x + \Delta x) = b + \Delta b.$$

  Neka je $\varepsilon > 0$ takav da je $\|\Delta A\| \leq \varepsilon \|A\|$ i $\|\Delta b\| \leq \varepsilon \|b\|$. Za $x \neq 0$ i $\varepsilon\kappa(A) < 1$ vrijedi

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{2\varepsilon\kappa(A)}{1 - \varepsilon\kappa(A)}.$$

Izračunati $\hat{L}$ i $\hat{U}$ zadovoljavaju $\hat{L}\hat{U} = A + \Delta A$, pri čemu je

$$|\Delta A| \leq \gamma_n |\hat{L}||\hat{U}|, \quad \gamma_n = \frac{nu}{1 - nu}.$$

Idealno bi bilo kada bi $|\hat{L}||\hat{U}| = |\hat{L}\hat{U}|$ jer bi tada imali

$$|\Delta A| \leq \frac{\gamma_n}{1-\gamma_n}|A|.$$

Kako to općenito nije slučaj, promatramo omjer

$$\frac{\||\hat{L}||\hat{U}|\|}{\|A\|}.$$

Što je omjer manji, očekujemo veću točnost rješenja.

Još jedan broj koji nam je bitan kod analize stabilnosti LU faktorizacije je **pivotni rast** $\rho_n(A)$:

$$\rho_n(A) = \frac{\max_{i,j,k} |a^{(k)_{i,j}}|}{\max_{i,j} |a_{i,j}|},$$

gdje su $a_{i,j}^{(k)}$ elementi matrice $A^{(k)}$ u k-tom koraku LU faktorizacije. Cilj je da taj broj bude što manji. Njegova veličina ovisi naravno o matrici A, ali i o pivotiranju.

Zadatak 2.5. Nađite uvjetovanost matrice A i pivotni rast za matrice iz zadataka 2.1 i 2.2. Za uvjetovanost promatrajte ∞-normu matrice

$$\|A\|_\infty = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}|.$$

Rješenje. Računamo prvo $\kappa(A)$ i $\rho_n(A)$ za zadatak 2.1. Za uvjetovanost nam treba $\|A\|_\infty$ i $\|A^{-1}\|_\infty$:

$$\|A\|_\infty = \max\{10, 21, 29\} = 29. \tag{2.1}$$

A^{-1} možemo izračunati kao $A^{-1} = U^{-1}L^{-1}$.

Računamo L^{-1}. Neka je $X_L^{(0)} = \left[\begin{array}{c|c} L & I \end{array}\right]$:

$$X_L^{(1)} = \left[\begin{array}{ccc:ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 4 & 1 & 3 & 0 & 1 \end{array}\right]$$

$$X_L^{(2)} = \left[\begin{array}{ccc:ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & 11 & -4 & 1 \end{array}\right].$$

Računamo U^{-1}. Neka je $X_U^{(0)} = \left[\begin{array}{c|c} U & I \end{array}\right]$:

$$X_U^{(1)} = \left[\begin{array}{ccc:ccc} 5 & 1 & 0 & 1 & 0 & -4/7 \\ 0 & 2 & 0 & 0 & 1 & 1/7 \\ 0 & 0 & 1 & 0 & 0 & 1/7 \end{array}\right]$$

$$X_U^{(2)} = \left[\begin{array}{ccc:ccc} 5 & 0 & 0 & 1 & -1/2 & -9/14 \\ 0 & 1 & 0 & 0 & 1/2 & 1/14 \\ 0 & 0 & 1 & 0 & 0 & 1/7 \end{array}\right]$$

$$X_U^{(3)} = \left[\begin{array}{ccc:ccc} 1 & 0 & 0 & 1/5 & -1/10 & -9/70 \\ 0 & 1 & 0 & 0 & 1/2 & 1/14 \\ 0 & 0 & 1 & 0 & 0 & 1/7 \end{array}\right].$$

Sada je A^{-1}:

$$A^{-1} = U^{-1}L^{-1} = \begin{bmatrix} 1/5 & -1/10 & -9/70 \\ 0 & 1/2 & 1/14 \\ 0 & 0 & 1/7 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 11 & -4 & 1 \end{bmatrix} = \begin{bmatrix} -71/70 & 29/70 & -9/70 \\ -3/14 & 3/14 & 1/14 \\ 11/7 & -4/7 & 1/7 \end{bmatrix},$$

pa je $\|A^{-1}\|_\infty = \max\{109/70, 1/2, 16/7\} = \frac{16}{7}$. Prema tome, uvjetovanost matrice A je $\kappa(A) = \|A\|_\infty \|A^{-1}\|_\infty = 29 \cdot \frac{16}{7} = 66.285714$, a pivotni rast je $\rho_n = \frac{8}{15}$.

Prelazimo na matricu iz zadatka 2.2. $\|A\|_\infty = \max\{7, 5, 6\} = 7$. A^{-1} računamo kao $A^{-1} = U^{-1}L^{-1}P$. Računamo L^{-1}. Neka je $X_L^{(0)} = \left[\begin{array}{c|c} L & I \end{array}\right]$:

$$X_L^{(1)} = \left[\begin{array}{ccc:ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2/3 & 10 & \\ 0 & 1/2 & 1 & -1/3 & 0 & 1 \end{array}\right]$$

$$X_L^{(2)} = \left[\begin{array}{ccc:ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2/3 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1/2 & 1 \end{array}\right]$$

Računamo U^{-1}. Neka je $X_U^{(0)} = \left[\begin{array}{c|c} U & I \end{array}\right]$:

$$X_U^{(1)} = \left[\begin{array}{ccc:ccc} 3 & -2 & 0 & 1 & 0 & -2/3 \\ 0 & -8/3 & 0 & 0 & 1 & -2/9 \\ 0 & 0 & 1 & 0 & 0 & 2/3 \end{array}\right]$$

$$X_U^{(2)} = \left[\begin{array}{ccc:ccc} 3 & 0 & 0 & 1 & -\frac{3}{4} & -1/2 \\ 0 & 1 & 0 & 0 & -3/8 & 1/12 \\ 0 & 0 & 1 & 0 & 0 & 2/3 \end{array}\right]$$

$$X_U^{(2)} = \left[\begin{array}{ccc:ccc} 1 & 0 & 0 & 1/3 & -1/4 & -1/6 \\ 0 & 1 & 0 & 0 & -3/8 & 1/12 \\ 0 & 0 & 1 & 0 & 0 & 2/3 \end{array}\right].$$

Sada je

$$A^{-1} = \begin{bmatrix} 1/3 & -1/4 & -1/6 \\ 0 & -3/8 & 1/12 \\ 0 & 0 & 2/3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2/3 & 1 & 0 \\ 0 & -1/2 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} -1/6 & -1/6 & 1/2 \\ -5/12 & 1/12 & 1/4 \\ -1/3 & 2/3 & 0 \end{bmatrix},$$

pa možemo izračunati $\|A^{-1}\|_\infty = \max\{5/6, 3/4, 1\} = 1$. Sada je $\kappa(A) = 7 \cdot 1 = 7$. Pivotni rast je $\rho_n(A) = \frac{3}{4}$. △

Zadatak 2.6. Za neki **neparan** prirodan broj n neka je A_n matrica reda n dana s

$$A_n = \begin{bmatrix} 3 & 0 & 0 & \dots & & \dots & 0 & 0 \\ 6 & 1 & 1 & 0 & & & 0 & 0 \\ 0 & 2 & 3 & 0 & 0 & & & \\ & 0 & 2 & 1 & 1 & & & \\ & & \ddots & \ddots & \ddots & \ddots & & \vdots \\ & & & 0 & 2 & 1 & 1 & 0 \\ \vdots & & & & 0 & 2 & 3 & 0 \\ 0 & & & & \dots & 0 & 2 & 1 \end{bmatrix}$$

(na dijagonali se izmjenjuju brojevi 3 i 1, iznad dijagonale se izmjenjuju 0 i 1, a ispod dijagonale su sve dvojke osim u prvom stupcu gdje je broj 6).

- Za $n = 5$ odredite LU faktorizaciju bez pivotiranja matrice A_n. Zaključite kako bi ti faktori L_n i U_n izgledali za velike n (bez dokaza indukcijom). Jesu li ti faktori L_n i U_n jednaki onima koji bi bili dobiveni prilikom računa LU faktorizacije s parcijalnim pivotiranjem?
- Izračunajte pivotni rast u ovisnosti o n.
- U ovisnosti o n izračunajte omjer $\dfrac{\||L_n| \cdot |U_n|\|_2}{\|A_n\|_2}$. Izvedite konačan zaključak o stabilnosti rješavanja sustava LU faktorizacijom bez pivotiranja (pod pretpostavkom da je uvjetovanost matrice dovoljno mala).

Rješenje. Za $n = 5$ matrica A_5 je

$$A_5 = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ 6 & 1 & 1 & 0 & 0 \\ 0 & 2 & 3 & 0 & 0 \\ 0 & 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 2 & 3 \end{bmatrix}.$$

Računamo LU faktorizaciju:

$$A_5^{(1)} = \left[\begin{array}{c|cccc} 3 & 0 & 0 & 0 & 0 \\ \hline -2 & 1 & 1 & 0 & 0 \\ 0 & 2 & 3 & 0 & 0 \\ 0 & 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 2 & 3 \end{array}\right], \quad A_5^{(2)} = \left[\begin{array}{cc|ccc} 3 & 0 & 0 & 0 & 0 \\ -2 & 1 & 1 & 0 & 0 \\ \hline 0 & -2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 2 & 3 \end{array}\right],$$

$$A_5^{(3)} = \left[\begin{array}{ccc|cc} 3 & 0 & 0 & 0 & 0 \\ -2 & 1 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 \\ \hline 0 & 0 & -2 & 1 & 1 \\ 0 & 0 & 0 & 2 & 3 \end{array}\right], \quad A_5^{(4)} = \left[\begin{array}{cccc|c} 3 & 0 & 0 & 0 & 0 \\ -2 & 1 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 \\ 0 & 0 & -2 & 1 & 1 \\ \hline 0 & 0 & 0 & -2 & 1 \end{array}\right],$$

pa su L i U jednaki

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 1 \end{bmatrix}, \; U = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Odavde vidimo da će elementi ispod glavne dijagonale matrice L_n biti 2, dok će preostali elementi biti 0. S druge strane, U_n će na dijagonali imati jedinice, osim na mjestu $(1,1)$ gdje će biti 3. Iznad glavne dijagonale se izmjenjuju 0 i 1, s tim da je 0 na mjestu $(1,2)$. Svi elementi iznad su jednaki 0.

To nisu matrice kao u parcijalnom pivotiranju, što vidimo jer već treba u prvom koraku pivotirati, ili zato što su elementi ispod dijagonale kod L veći od 1.

Pivotni rast je $\frac{3}{6} = \frac{1}{2}$ za sve n.

Omjer $\dfrac{\||L_n| \cdot |U_n|\|_2}{\|A_n\|_2}$ je točno 1. Naime, matrice L i U su nenegativne, pa je $|L| \cdot |U| = L \cdot U = A$.

Dakle, u ovom slučaju nije potrebno parcijalno pivotirati, algoritam bez parcijalnog pivotiranja dovoljno je stabilan.

△

Zadatak 2.7 (1. kolokvij 2023.)**.** Za neki prirodan broj n dana je matrica

$$A_n = \begin{bmatrix} 1 & 1 & 0 & \dots & & & \dots & 0 & 1 \\ 2 & 3 & 1 & 0 & & & & 0 & 0 \\ 0 & 3 & 4 & 1 & 0 & & & & \\ & 0 & 4 & 5 & 1 & & & & \\ & & \ddots & \ddots & \ddots & & \ddots & & \vdots \\ & & & 0 & n-2 & n-1 & 1 & 0 & \\ \vdots & & & & 0 & n-1 & n & 0 & \\ 0 & & & & \dots & 0 & n & 0 & \end{bmatrix}.$$

Na toj matrici provodimo

(i) LU faktorizaciju bez pivotiranja;

(ii) LU faktorizaciju s parcijalnim privotiranjem.

Za $n = 5$ odredite faktore L i U, odnosno L, U i P za (i) i (ii), te zaključite kako bi ti faktori izgledali za proizvoljan n (slutnju nije potrebno dokazivati indukcijom). Koristeći pivotni rast (dovoljno je gledati samo konačne matrice L i U, a ne i sve međukorake) argumentirajte jesu li računi (i) i (ii) stabilni u aritmetici računala za velike n.

Rješenje. Za $n = 5$ imamo ovakvu situaciju bez pivotiranja:

$$A_5 = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 \\ 2 & 3 & 1 & 0 & 0 \\ 0 & 3 & 4 & 1 & 0 \\ 0 & 0 & 4 & 5 & 0 \\ 0 & 0 & 0 & 5 & 0 \end{bmatrix}, \quad L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 4 & 1 & 0 \\ 0 & 0 & 0 & 5 & 1 \end{bmatrix}, \quad U \begin{bmatrix} 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & -2 \\ 0 & 0 & 1 & 1 & 6 \\ 0 & 0 & 0 & 1 & -24 \\ 0 & 0 & 0 & 0 & 120 \end{bmatrix}$$

Za proizvoljan n zaključujemo da će matrica L na dijagonali imati jedinice, a ispod dijagonale će ići brojevi $2, 3, \ldots, n$. U matrici U sve do zadnjeg stupca na dijagonali i neposredno iznad nje nalazit će se jedinice, a u zadnjem stupcu će pisati $(-1)^{k+1}k!$. Zbog tih zadnjih članova pivotni rast iznosi $\dfrac{n!}{n} = (n-1)!$, te je ogroman, pa račun nije stabilan.

Ako uključimo parcijalno pivotiranje:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1/2! & -1/3! & 1/4! & -1/5! & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 2 & 3 & 1 & 0 & 0 \\ 0 & 3 & 4 & 1 & 0 \\ 0 & 0 & 4 & 5 & 0 \\ 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad P = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Zaključujemo da za općeniti n u i-tom retku matrice U $(i < n)$ piše $(i+1)$-ti redak matrice A_n, a u zadnjem retku su sve nule osim jedinice u zadnjem stupcu. U matrici L jedini netrivijalni elementi su u zadnjem retku, gdje u k-tom stupcu piše $(-1)^{k+1}/(k+1)!$. Matrica P dobivena je od jedinične tako da je prvi redak prebačen u zadnji, a svi ostali su pomaknuti za jedno mjesto gore.

U ovom slučaju pivotni rast iznosi općenito $n/n = 1$, pa bi ovaj račun trebao biti stabilan.

△

Primjer 2.8 (*Za one koji žele znati više*). Na predavanjima smo kao primjer rješavali sustav oblika

$$\begin{bmatrix} a & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

LU faktorizacijom sa i bez parcijalnog pivotiranja.

Promotrimo funkciju koje komponenti a matrice sustava pridružuje rješenje $f(a) := x_1$. Izračunajte relativnu uvjetovanost preslikavanja $f(a)$, povežite zaključak o lošoj uvjetovanosti problema s pojmom uvjetovanosti matrice. Nakon toga, izvedite izraze za $fl(f(a))$ u slučaju rješavanja sustava Gaussovim eliminacijama sa i bez parcijalnog pivotiranja, i zaključite što se događa s relativnom greškom kada $a \to 0$.

Rješenje. Riješimo originalan sustav Gaussovim eliminacijama. Kako bismo izveli izraz za $fl(f(x))$, nemojmo raditi nikakva algebarska pojednostavljenja formula, nego isključivo ono što se događa u računalu. Pomnožimo prvu jednadžbu s $-1/a$ i pridodajmo drugoj i dobivamo

$$\left(\left(-\frac{1}{a}\right)\cdot 1 + 1\right) x_2 = \left(\left(-\frac{1}{a}\right)\cdot 1 + 2\right) \implies x_2 = \frac{2-\frac{1}{a}}{1-\frac{1}{a}}.$$

Taj rezultat uvrštavamo u prvu jednadžbu sustava, odakle dobivamo

$$x_1 = \frac{1}{a}\left(1 - x_2\right) = \frac{1}{a}\cdot\left(1 - \frac{2-\frac{1}{a}}{1-\frac{1}{a}}\right).$$

Dakle, dobili smo preslikavanje

$$f(a) = \frac{1}{a}\cdot\left(1 - \frac{2-\frac{1}{a}}{1-\frac{1}{a}}\right).$$

Za potrebe računa $fl(f(a))$ ovaj izraz ne trebamo pojednostavljivati, ali za potrebe računa uvjetovanosti, pojednostavimo ga:

$$f(a) = \frac{1}{a}\cdot\left(1a - \frac{2a-1}{a-1}\right) = \frac{a-1-(2a-1)}{a(a-1)} = \frac{-a}{a(a-1)} = \frac{1}{1-a}.$$

Za relativnu uvjetovanost računamo derivaciju

$$f'(a) = \frac{1}{(1-a)^2},$$

odakle je

$$\kappa_f^{\text{rel}}(a) = \left|\frac{af'(a)}{f(a)}\right| = \left|\frac{a(1-a)}{(1-a)^2}\right| = \left|\frac{a}{1-a}\right|.$$

Vidimo da uvjetovanost preslikavanja $f(a)$ ide u bekonačnost kada $a \to 1$, što je logično jer tada matrica sustava postaje sve bliža singularnoj matrici (u izrazu za uvjetovanost matrice, izraz $\|A^{-1}\|$ teži u beskonačnost u tom slučaju).

Izvedimo izraz za stabilnost: za dovoljno male $\varepsilon_1, \ldots, \varepsilon_7$ imamo

$$\begin{aligned} fl(f(x)) &= \frac{1}{x}\left(1 - \frac{(2 - \frac{1}{x}(1+\varepsilon_1))(1+\varepsilon_2)}{(1-\frac{1}{x}(1+\varepsilon_3))(1+\varepsilon_4)}(1+\varepsilon_5)\right)(1+\varepsilon_6)(1+\varepsilon_7) = \\ &\quad \left[1+\bar{\varepsilon} := \frac{(1+\varepsilon_2)(1+\varepsilon_5)}{1+\varepsilon_4}\right] \\ &= f(x)(1+\varepsilon_6)(1+\varepsilon_7)\frac{1-x}{x}\left(1-(1+\bar{\varepsilon}) - \frac{(2x-(1+\varepsilon_1))-(x-(1+\varepsilon_3))}{x-(1+\varepsilon_3)}(1+\bar{\varepsilon})\right) \\ &= f(x)(1+\varepsilon_6)(1+\varepsilon_7)\frac{1-x}{x}\left(-\bar{\varepsilon} - \frac{x+\varepsilon_3-\varepsilon_1}{x-1-\varepsilon_3}(1+\bar{\varepsilon})\right). \end{aligned}$$

Kada $x \to 0$, u nazivniku prvog razlomka imamo izraz koji teži k nuli. Zbog preostalih "epsilonova" u zagradi zadnji faktor ne teži k nuli, pa zato $fl(f(x))$ nije ograničen. Zaključujemo da je račun nestabilan.

Sada provjerimo što se događa kada rješavamo sustav Gaussovim eliminacijama s pivotiranjem. U stvari rješavamo sustav

$$\begin{bmatrix} 1 & 1 \\ a & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Pomnožimo prvu jednadžbu s $-a$ i pridodajmo drugoj i dobivamo

$$((-a)\cdot 1 + 1)x_2 = ((-a)\cdot 2 + 1) \implies x_2 = \frac{-2a+1}{-a+1}.$$

Taj rezultat uvrštavamo u prvu jednadžbu sustava, odakle dobivamo

$$x_1 = 2 - x_2 = 2 - \frac{-2a+1}{-a+1}.$$

Dakle, dobili smo isto preslikavanje

$$f(a) = 2 - \frac{-2a+1}{-a+1}$$

kao i prije, samo zapisano u ekvivalentnom formatu.

Izvedimo izraz za stabilnost: za dovoljno male $\varepsilon_1, \ldots, \varepsilon_6$ imamo

$$\begin{aligned} fl(f(x)) &= \left(2 - \frac{(-2x(1+\varepsilon_1)+1)(1+\varepsilon_2)}{(-x(1+\varepsilon_3)+1)(1+\varepsilon_4)}(1+\varepsilon_5)\right)(1+\varepsilon_6) \\ &\quad \left[1+\bar{\varepsilon} := \frac{(1+\varepsilon_2)(1+\varepsilon_5)}{1+\varepsilon_4}\right] \\ &= f(x)(1+\varepsilon_6)(1-x)\left(2 - 2(1+\bar{\varepsilon}) + \frac{2(x(1+\varepsilon_3)-1) - 2(x(1+\varepsilon_1)+1}{x(1+\varepsilon_3)-1}(1+\bar{\varepsilon})\right) \\ &= f(x)(1+\varepsilon_6)(1-x)\left(-2\bar{\varepsilon} + \frac{2x(\varepsilon_3-\varepsilon_1)-1}{x(1+\varepsilon_3)-1}(1+\bar{\varepsilon})\right). \end{aligned}$$

Sada, kada $x \to 0$, faktor koji množi $f(x)$ teži u

$$(1+\varepsilon_6)(1-\bar{\varepsilon}).$$

Ako pretpostavimo da su svi ε_i po apsolutnoj vrijednosti manji od $\varepsilon \ll 1$, tada je

$$\begin{aligned}(1+\bar{\varepsilon}) &= \frac{(1+\varepsilon_2)(1+\varepsilon_5)}{1+\varepsilon_4} = \frac{1+\varepsilon_2+\varepsilon_5+\mathcal{O}(\varepsilon)}{1+\varepsilon_4} = (1+\varepsilon_2+\varepsilon_5)\frac{1}{1+\varepsilon_4}+O(\varepsilon)\\ &= (1+\varepsilon_2+\varepsilon_5)\left(1-\varepsilon_4+\sum_{n=2}^{\infty}(-1)^n\varepsilon_4^n\right)+\mathcal{O}(\varepsilon)\\ &= (1+\varepsilon_2+\varepsilon_5)(1-\varepsilon_4)+\mathcal{O}(\varepsilon) = 1+\varepsilon_2+\varepsilon_5-\varepsilon_4+\mathcal{O}(\varepsilon)\\ \implies & |\bar{\varepsilon}| \lessapprox 3\varepsilon.\end{aligned}$$

Slično, $(1+\varepsilon_6)(1-\bar{\varepsilon}) = (1+\bar{\bar{\varepsilon}})$, gdje je $|\bar{\bar{\varepsilon}}| \lessapprox 4\varepsilon$, pa je ovaj račun stabilan kada $x \to 0$. $\triangle$

2.3 Faktorizacija Choleskog

Simetrična matrica $A \in \mathbb{R}^{n\times n}$ je **pozitivno definitna** ako za svaki vektor $x \in \mathbb{R} \setminus \{0\}$ vrijedi $x^T Ax > 0$.

Vrijede sljedeće karakterizacije pozitivne definitnost: simetrična matrica A je pozitivno definitna ako i samo ako

- sve svojstvene vrijednosti od A su pozitivne,
- sve vodeće minore su pozitivne,
- postoji gornjetrokutasta matrica R sa pozitivnim elementima na dijagonali takva da je $A = R^T R$ (**faktorizacija Choleskog**).

Elementi matrice $R = [r_{i,j}]_{i,j=1}^n$ se računaju sljedećim formulama:

- dijagonalni elementi, $i = 1, \ldots, n$:

$$r_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} r_{ki}^2},$$

- elementi i-tog retka, $j = i+1, \ldots, n$:

$$r_{ij} = \frac{1}{r_{ii}}\left(a_{ij} - \sum_{k=1}^{i-1} r_{ki}r_{kj}\right).$$

Naglasimo da se zadnji uvjet karakterizacije simetričnih pozitivno definitnih matrica može iskazati na sljedeći način:

- simetrična matrica je pozitivno definitna ako i samo ako je gornji algoritam za traženje faktorizacije Choleskog provediv (ako prilikom algoritma nismo došli do dijeljenja s nulom ili korijenovanja negativnog broja).

Zadatak 2.9. Odredite faktorizaciju Choleskog matrice

$$A = \begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix}.$$

Rješenje. Računamo elemente prvog retka:

$$\begin{aligned} r_{11} &= \sqrt{a_{11}} = \sqrt{25} = 5, \\ r_{12} &= \frac{1}{r_{11}} a_{12} = \frac{1}{5} \cdot 15 = 3, \\ r_{13} &= \frac{1}{r_{11}} a_{13} = \frac{1}{5} \cdot (-5) = -1. \end{aligned}$$

Računamo elemente drugog retka:

$$\begin{aligned} r_{22} &= \sqrt{a_{22} - r_{12}^2} = \sqrt{18 - 9} = 3, \\ r_{23} &= \frac{1}{r_{22}} \left(a_{23} - r_{12} r_{13} \right) = \frac{1}{3} \left(0 - 3 \cdot (-1) \right) = 1. \end{aligned}$$

Računamo element trećeg retka:

$$r_{33} = \sqrt{a_{33} - r^2 13 - r_{23}^2} = \sqrt{11 - (-1)^2 - 1^2} = 3.$$

Dakle, matrica R je

$$R = \begin{bmatrix} 5 & 3 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix}.$$

△

Zadatak 2.10. Zadane su četiri matrice. Za svaku odredite je li ona pozitivno definitna. Detaljno obrazložite svoje odgovore. Za barem jednu od matrica provedite algoritam za određivanje Choleskyjeve faktorizacije.

$$A = \begin{bmatrix} 4 & 2 & -4 & 6 \\ 2 & 7 & -2 & -3 \\ -4 & -2 & 4 & -6 \\ 6 & -3 & -6 & 9 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & 2 & -4 & 6 \\ 2 & 7 & -2 & -3 \\ -4 & -2 & -4 & -6 \\ 6 & -3 & -6 & 9 \end{bmatrix}$$

$$C = \begin{bmatrix} 4 & 0 & -2 & 0 \\ 0 & 16 & -8 & 0 \\ -2 & -8 & 6 & -2 \\ 0 & 0 & -2 & 29 \end{bmatrix}, \quad D = \begin{bmatrix} 4 & 0 & -2 & 0 \\ 0 & 16 & -8 & 0 \\ -2 & -8 & 14 & -2 \\ 0 & 0 & -2 & 29 \end{bmatrix}.$$

Rješenje. Matrica A je singularna budući da je zbroj prvog i trećeg retka jednak nul-retku. Zato joj je jedna od singularnih vrijednosti jednaka nuli, pa joj nisu sve svojstvene vrijednosti pozitivne, dakle nije pozitivno definitna.

Matrica B ima jedan dijagoalni element (-4 na mjestu $(3,3)$) manji od nule, pa opet ne može biti pozitivno definitna.

Matrica D je strogo dijagonalno dominantna s pozitivnim dijagonalnim elementima:

$$\begin{aligned} 4 &> 2 = |0| + |-2| + |0| \\ 16 &> 8 = |0| + |-8| + |0| \\ 14 &> 12 = |-2| + |-8| + |-2| \\ 29 &> 2 = |0| + |0| + |-2|. \end{aligned}$$

Zato je pozitivno definitna.

Za matricu C provedimo algoritam.

$$\begin{aligned} r_{1,1} &= \sqrt{c_{1,1}} = \sqrt{4} = 2; \\ r_{1,2} &= \frac{c_{2,1}}{r_{1,1}} = 0,\ r_{1,3} = \frac{c_{3,1}}{r_{1,1}} = -1,\ r_{1,4} = \frac{c_{4,1}}{r_{1,1}} = 0; \\ r_{2,2} &= \sqrt{c_{2,2} - r_{1,2}^2} = \sqrt{16 - 0} = 4; \\ r_{2,3} &= \frac{c_{2,3} - r_{1,2}r_{1,3}}{r_{2,2}} = \frac{-8 - 0}{4} = -2,\ r_{2,4} = \frac{c_{2,4} - r_{1,2}r_{1,4}}{r_{2,2}} = \frac{0 - 0}{4} = 0; \\ r_{3,3} &= \sqrt{c_{3,3} - r_{1,3}^2 - r_{2,3}^2} = \sqrt{6 - 1 - 4} = 1; \\ r_{3,4} &= \frac{c_{3,4} - r_{1,3}r_{1,4} - r_{2,3}r_{2,4}}{r_{3,3}} = \frac{-2 - 0 - 0}{1} = -2; \\ r_{4,4} &= \sqrt{c_{4,4} - r_{1,4}^2 - r_{2,4}^2 - r_{3,4}^2} = \sqrt{29 - 0 - 0 - 4} = 5. \end{aligned}$$

Algoritam smo uspješno proveli, pa je matrica C pozitivno definitna. Njezina Choleskyjeva faktorizacija je $C = R^T R$, gdje je

$$R = \begin{bmatrix} 2 & 0 & -1 & 0 \\ 0 & 4 & -2 & 0 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 5 \end{bmatrix}.$$

△

3

Interpolacija

3.1 Interpolacija polinomom

Problem interpolacije polinomom: Neka su zadane točke (x_i, y_i), $i = 0, 1, \ldots, n$ i $x_i \neq x_j$ za $i \neq j$. Treba naći polinom $p \in \mathcal{P}_n$ takav da je $p(x_i) = y_i$ za sve $i = 0, 1, \ldots, n$. Takav polinom zovemo **interpolacijski polinom**.

Ukoliko su $x_i \neq x_j$ za $i \neq j$ interpolacijski polinom postoji i jedinstven je. Ovisno koju bazu za prostor polinoma izaberemo dobijemo:

- Prikaz u standardnoj bazi
- Lagrangeov oblik interpolacijskog polinoma
- Newtonov oblik interpolacijskog polinoma

Prikaz interpolacijskog polinoma u standardnoj bazi. Skup $\{1, x, x^2, \ldots, x^n\}$ je baza za realan vektorski prostor $\mathcal{P}_n = \left\{p : p(x) = \sum_{k=0}^{n} a_k x^k, a_i \in \mathbb{R}\right\}$. Problem nalaženja interpolacijskog polinoma p se svodi na rješavanje sustava linearnih jednadžbi

$$p(x_i) = y_i, \quad i = 0, 1, \ldots, n.$$

To je sustav $(n+1) \times (n+1)$ koji ima jedinstveno rješenje.

Lagrangeov oblik interpolacijskog polinoma. Neka su zadane točke (x_i, y_i), $i = 0, 1, \ldots, n$. Definiramo polinome $\ell_i \in \mathcal{P}_n$ za $i = 0, 1, \ldots, n$

$$\ell_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^{n} \frac{x - x_j}{x_i - x_j} = \frac{\omega_i(x)}{\omega_i(x_i)}$$

pri čemu su

$$\omega(x) = \prod_{i=0}^{n} (x - x_i), \quad \omega_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^{n} (x - x_j) = \frac{\omega(x)}{x - x_i}.$$

Za polinome ℓ_i vrijedi: $\ell_i(x_j) = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$. Dakle traženi interpolacijski polinom p možemo dobiti kao

$$p(x) = \sum_{i=0}^{n} y_i \ell_i(x), \quad i, j = 0, 1, \ldots, n. \tag{3.1}$$

Newtonov oblik interpolacijskog polinoma. Neka su zadane točke (x_i, y_i), $i = 0, 1, \ldots, n$. **Newtonov interpolacijski polinom** za zadane čvorove dan je sa

$$p(x) = f[x_0] + \sum_{i=1}^{n} f[x_0, \ldots, x_i](x - x_0) \ldots (x - x_{i-1}).$$

Brojeve $f[x_0, \ldots, x_i]$ zovemo **podijeljene razlike**.
Primijetimo da je baza zadana sa

$$\left\{1, (x - x_0), (x - x_0)(x - x_1), \ldots, \prod_{k=0}^{n-1} (x - x_k)\right\}.$$

Za vježbu dokažite da je ovo baza u $\mathcal{P}_n$.

Definicija 3.1. **Podijeljena razlika nultog reda** u čvoru x_i dana je sa

$$f[x_i] = y_i, \quad i = 0, 1, 2, \ldots, n$$

Podijeljena razlika k-tog reda ($k \geq 1$) dobiva se iz podijeljenih razlika $(k-1)$-og reda formulom

$$f[x_i, \ldots, x_{i+k}] = \frac{f[x_{i+1}, \ldots, x_{i+k}] - f[x_i, \ldots, x_{i+k-1}]}{x_{i+k} - x_i}$$

za $k = 1, \ldots, n$ i $i = 0, \ldots, n - k$.

Tablica podijeljenih razlika izgleda

x_i \ k	0	1	2	3
x_0	$f[x_0] = y_0$			
		$f[x_0, x_1]$		
x_1	$f[x_1] = y_1$		$f[x_0, x_1, x_2]$	
		$f[x_1, x_2]$		$f[x_0, x_1, x_2, x_3]$
x_2	$f[x_2] = y_2$		$f[x_1, x_2, x_3]$	
		$f[x_2, x_3]$		
x_3	$f[x_3] = y_3$			

Tablica 3.1: Tablica podijeljenih razlika

Zadatak 3.2. Nađite Newtonov interpolacijski polinom stupnja ≤ 2 koji prolazi točkama $(-1, 3)$, $(1, 5)$, $(2, 0)$.

Rješenje. Formirajmo tablicu podijeljenih razlika

x_i \ k	0	1	2
-1	3		
		$\frac{5-3}{1-(-1)} = 1$	
1	5		$\frac{-5-1}{2-(-1)} = -2$
		$\frac{0-5}{2-1} = -5$	
2	0		

Slijedi

$$p(x) = 3 + (x+1) - 2(x+1)(x-1).$$

△

3.2 Pogreške interpolacije

Neka je $f \in C^{n+1}([a,b])$, funkciju f interpoliramo u točkama $x_0, x_1, \ldots, x_n$ polinomom $p \in \mathcal{P}_n$, za kojeg vrijedi

$$p(x_i) = f(x_i) \quad i = 0, 1, \ldots, n.$$

Prava greška interpolacije u točki $x \in [x_0, x_n]$ dana je sa

$$|f(x) - p(x)|.$$

Ocjena prave greške interpolacije je

$$|f(x) - p(x)| \leq \frac{|\prod_{i=0}^{n}(x - x_i)|}{(n+1)!} M_{n+1} f$$

gdje je $M_{n+1}f = \max_{y \in [x_0, x_n]} |f^{(n+1)}(y)|$.
Pogrešku interpolacije na čitavom intervalu $[x_0, x_n]$ zovemo još i **uniformna pogreška**. Računamo je kao

$$\max_{x \in [x_0, x_n]} |f(x) - p(x)| = \max_{x \in [x_0, x_n]} \frac{|\prod_{i=0}^{n}(x - x_i)|}{(n+1)!} f^{n+1}(\xi_x).$$

Ocjena uniformne pogreške je

$$\max_{x \in [x_0, x_n]} |f(x) - p(x)| \leq \frac{|\omega(x_0, \ldots, x_n)|}{(n+1)!} M_{n+1} f$$

gdje je $\omega(x_0, \ldots, x_n) = \max_{x \in [x_0, x_n]} |(x - x_0) \ldots (x - x_n)|$.

Zadatak 3.3. Nađite interpolacijski polinom stupnja 2 za funkciju $f(x) = \ln(1 + x)$ sa čvorovima interpolacije 0, 1 i 3.

a) Nađite $p(2)$, pravu pogrešku interpolacije u točki 2 i ocjenu pogreške interpolacije u točki 2,

b) nađite ocjenu uniformne pogreške interpolacije.

Rješenje.

x_i \ k	0	1	2
0	0		
		$\ln 2$	
1	$\ln 2$		$\frac{-\ln 2}{6}$
		$\frac{\ln 4-\ln 2}{3-1} = \frac{\ln 2}{2}$	
3	$\ln 4$		

a)

$$p(x) = \ln 2 \cdot x - \frac{\ln 2}{6} x(x-1)$$

Prava pogreška u točki 2 je

$$|f(2) - p(2)| = |\ln 3 - \frac{5}{3}\ln 2| = 0.056633.$$

Da bismo odredili ocjenu prave pogreške u točki 2 treba nam $f^{(3)}$ i maksimum njene aposlutne vrijednosti na $[0,3]$.

$$f'(x) = \frac{1}{1+x}, \quad f''(x) = -\frac{1}{(x+1)^2} \quad f'''(x) = \frac{2}{(x+1)^3}.$$

Kako je f''' strogo padajuća slijedi da je

$$M_3 f = f'''(0) = 2.$$

Slijedi da je ocjena

$$|f(2) - p(2)| \leq \frac{|(2-0)(2-1)(2-3)|}{3!} 2 = \frac{2}{3} = 0.666667$$

Vidimo da je ocjena prave pogreške pesimistična (deset puta veća od prave).

b) Za ocjenu uniformne pogreške treba nam

$$\omega(0,1,3) = \max_{x\in[0,3]} |x(x-1)(x-3)| = \max_{x\in[0,3]} |x^3 - 4x^2 + 3x|.$$

Tražimo globalni maksimum

$$(x^3 - 4x^2 + 3x)' = 3x^2 - 8x + 3$$

Stacionarne točke su $x_1 = 0.451516$ i $x_2 = 2.21525$. Kako je $|\omega(x_1)| = 0.63113$, a $|\omega(x_2)| = 2.11261$, slijedi da je $\omega(0,1,3) = 2.11261$. Konačno

$$\max_{x\in[0,3]} |f(x) - p(x)| \leq \frac{|\omega(0,1,3)|}{3!} M_3 f = \frac{2.11261}{6} 2 = 0.704204.$$

△

3.3 Ekvidistantni čvorovi

Kažemo da su čvorovi interpolacije $x_0, \ldots, x_n$ ekvidistantni ako su svi susjedni čvorovi jednako udaljeni:

$$x_i = x_0 + ih, \quad h = \frac{x_n - x_0}{n} \quad i = 0, \ldots, n.$$

Za male n (cca $n \leq 5$) u ocjeni uniformne greške interpolacije član $\omega(x0, \ldots, x_n)$ tražimo kao i inače, dok za velike n možemo (bez dokaza) koristiti ocjenu

$$\omega(x_0, x_1, \ldots, x_n) = \max_{x\in[x_0,x_n]} |(x - x_0) \ldots (x - x_n)| < n! h^{n+1},$$

odakle je

$$\max_{x\in[x_0,x_n]} |f(x) - p(x)| \leq \frac{h^{n+1}}{n+1} M_{n+1} f.$$

Zadatak 3.4. Zadana je funkcija

$$f(x) = \frac{x+1}{(3-x)^2}$$

na intervalu $[1,2]$. Funkciju interpoliramo polinomom p_n sa $n+1$ čvorova na ekvidistantnoj mreži.

a) Dokažite da niz $(p_n)_n$ uniformno konvergira ka f na $[1,2]$.

b) Nađite najmanji $n \in \mathbb{N}$ za koji ocjena pogreške ne prelazi 10^{-2} na čitavom $[1,2]$.

Rješenje. Rastavimo funkciju na pracijalne razlomke

$$f(x) = \frac{4}{(3-x)^2} - \frac{1}{3-x}$$

Kako je

$$\left(\frac{1}{3-x}\right)' = \frac{1}{(3-x)^2}$$

imamo da je

$$f^{(n+1)}(x) = 4\left(\frac{1}{3-x}\right)^{(n+2)} - \left(\frac{1}{3-x}\right)^{(n+1)}$$

Indukcijom se pokaže da je

$$\left(\frac{1}{3-x}\right)^{(n)} = \frac{n!}{(3-x)^{n+1}}$$

pa je

$$f^{(n+1)}(x) = 4\frac{(n+2)!}{(3-x)^{n+3}} - \frac{(n+1)!}{(3-x)^{n+2}} = \frac{(n+1)!}{(3-x)^{n+3}}(5+4n+x)$$

pa je

$$|f^{(n+1)}(x)| = \frac{(n+1)!}{(3-x)^{n+3}}(5+4n+x) \leq (n+1)!(7+4n)$$

(najveće za $x = 2$). Koristeći formulu za ocjenu pogreške za ekvidistantnu mrežu, dobivamo

$$\max_{x\in[x_0,x_n]} |f(x) - p_n(x)| \leq \frac{h^{n+1}}{n+1}(n+1)!(7+4n) \underbrace{=}_{h=\frac{1}{n}} \frac{1}{n^{n+1}} n!(7+4n) = \frac{(n-1)!(4n+7)}{n^n}$$

Koristimo da je svaki od $(n-2)$ faktora u $(n-1)!$ (svi osim prvog) manji od odgovarajućih faktora u n^{n-2}. Zato imamo

$$\max_{x\in[x_0,x_n]} |f(x) - p_n(x)| \leq \frac{(n-1)!(4n+7)}{n^n} = \frac{1}{n}\frac{(n-1)!}{n^{n-2}}\left(4+\frac{7}{n}\right) < \frac{1}{n}\cdot 1\cdot 11 = \frac{11}{n} \to 0$$

kada n teži u ∞. Za drugi dio zadatka potrebno je riješiti nejednadžbu

$$\frac{(n-1)!(4n+7)}{n^n} \leq 10^{-2}.$$

Nema bolje strategije od uvrštavanja brojeva $n = 1, 2, \ldots$ redom dok ne naiđemo na prvi za koji je nejednakost zadovoljena. To je za $n = 9$ kada je izraz na lijevoj strani jednak 0.00447514. △

3.4 Čebiševljeva mreža

Na domeni $[-1, 1]$ definiramo Čebiševljevu mrežu kao točke $x_0, \ldots, x_n$:

$$x_k = \cos\frac{(2k+1)\pi}{2n+2}, \quad k = 0, 1, \ldots, n.$$

To su točke koje su dobivene kao nultočke Čebiševljevog polinoma prve vrste T_{n+1}.

Čebiševljevi polinomi prve vrste $(T_n(x))_{n\in\mathbb{N}}$ dani su rekurzivnom formulom

$$T_0(x) = 1, \quad T_1(x) = x$$
$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x).$$

Nekoliko prvih članova iznosi:

$$T_2(x) = 2x^2 - 1, \quad T_3(x) = 4x^3 - 3x, \quad T_4(x) = 8x^4 - 8x^2 + 1, \quad \dots$$

Za njih vrijedi

$$T_n(\cos\varphi) = \cos(n\varphi), \quad n \in \mathbb{N}_0.$$

Glavna svojstva tih polinoma su (za $n \in \mathbb{N}$):

- polinom T_n je stupnja n i vodeći koeficijent mu je 2^{n-1} (to se induktivno dokaže iz rekurzije);
- polinom T_n ima n nultočaka, koliko ih ima i jednadžba $\cos(n\varphi)$ za $\varphi \in [0, \pi]$ – to su $\cos\frac{(2k+1)\pi}{2n}$, za $k = 0, \dots, n-1$;
- prema teoremu s predavanja, za polinom $2^{-n} \cdot T_{n+1}$ se postiže najmanja vrijednost izraza $|\omega(x_0, \dots, x_n)|$ među svim normiranim polinomima stupnja $n+1$ s nultočkama $x_0, \dots, x_n$ – ona iznosi 2^{-n}.

Zbog zadnjeg svojstva za nultočke polinoma T_{n+1} dobivamo najbolju ogradu u ocjeni uniformne pogreške interpolacije, što nas motivira da na intervalu $[-1, 1]$ koristimo Čebiševljeve točke.

U slučaju da se radi o nekom drugom intervalu $[a, b]$, nultočke je potrebno pomaknuti funkcijom $l : [a, b] \to [-1, 1]$ zadanom s

$$l(x) = \frac{2}{b-a}(x - a) - 1.$$

Tada dobivamo točke

$$\tilde{x}_k = \frac{a+b}{2} + \frac{b-a}{2}\cos\frac{(2k+1)\pi}{2n+2}, \quad k = 0, 1, \dots, n,$$

i ocjenu

$$\max_{x\in[a,b]} |f(x) - p(x)| \leq 2\left(\frac{b-a}{4}\right)^{n+1} \frac{M_{n+1}f}{(n+1)!}.$$

Zadatak 3.5. Za funkciju

$$f(x) = \frac{1}{1 + 25x^2}$$

na segmentu $[-1, 1]$ nađite interpolacijski polinom stupnja 2 na ekvidistantnoj i Čebiševljevoj mreži te uniformnu ocjenu greške.

Rješenje. Ekvidistantna mreža sastoji se od čvorova $-1, 0, 1$. Tablica podijeljenih razlika je

$x_i \setminus k$	0	1	2
-1	$\frac{1}{26}$		
		$\frac{25}{26}$	
0	1		$\frac{-25}{26}$
		$\frac{-25}{26}$	
1	$\frac{1}{26}$		

pa je interpolacijski polinom dan sa

$$p_2(x) = \frac{1}{26} + \frac{25}{26}(x+1) - \frac{25}{26}x(x+1) = 1 - \frac{25}{26}x^2.$$

Na Čebiševljevoj mreži čvorovi su dani kao nultočke polinoma $T_3(x) = \cos(3 \arccos x)$, a to su

$$x_i = \cos\frac{(2i+1)\pi}{6}, \quad i = 0, 1, 2.$$

To vidimo iz formule za Čebiševljeve točke, ili računajući nultočke polinoma $T_3(x) = 4x^3 - 3x$. Zato je

$$x_0 = \cos\frac{\pi}{6} = \frac{\sqrt{3}}{2}, \quad x_1 = 0, \quad x_2 = -\frac{\sqrt{3}}{2}$$

Tablica podijeljenih razlika je dana sa

$x_i \setminus k$	0	1	2
$\frac{\sqrt{3}}{2}$	$\frac{4}{79}$		
		$-\frac{150}{79\sqrt{3}}$	
0	1		$-\frac{100}{79}$
		$\frac{150}{79\sqrt{3}}$	
$-\frac{\sqrt{3}}{2}$	$\frac{4}{79}$		

pa je interpolacijski polinom dan sa

$$r_2(x) = \frac{4}{79} - \frac{150}{79\sqrt{3}}(x - \frac{\sqrt{3}}{2}) - \frac{100}{79}x(x - \frac{\sqrt{3}}{2}) = 1 - \frac{100}{79}x^2.$$

Tražimo ocjenu pogreške, pa nam treba maksimum funkcije $|f^{(3)}(x)|$ na $[-1, 1]$. Kako je

$$\begin{aligned}
f'(x) &= -\frac{50x}{(1+25x^2)^2} \\
f''(x) &= -\frac{50(1-75x^2)}{(1+25x^2)^3} \\
f'''(x) &= -15000\frac{25x^3 - x}{(1+25x^2)^4} \\
f^{(4)} &= \frac{15000}{(1+25x^2)^5}(3125x^4 - 250x^2 + 1)
\end{aligned}$$

Stacionarne točke od f''' su

$$x_{1,2} = \pm 0.275, \quad x_{3,4} = \pm 0.06498$$

Da nađemo maksimum, pogledajmo vrijednosti od f''' u stacionarnim točkama i rubovima intervala:

$$\begin{aligned}
|f'''(x_2)| = |f'''(x_1)| &= 56.62 \\
|f'''(x_3)| = |f'''(x_4)| &= 583.57 \\
|f'''(1)| = |f'''(-1)| &= 0.787
\end{aligned}$$

Za uniformnu ocjenu greške treba naći

$$\max_{x\in[-1,1]} |(x+1)x(x-1)|.$$

Promatramo polinom $\omega_2(x) = x^3 - x$. On svoje ekstreme postiže u točkama $\pm\frac{1}{\sqrt{3}}$ i po apsolutnoj vrijednosti oni iznose $\frac{2\sqrt{3}}{9}$ pa je uniformna ocjena greške za ekvidistantnu mrežu jednaka

$$\max_{x\in[-1,1]} |f(x) - p_2(x)| \leq 37.436.$$

Za interpolacijski polinom na Čebiševljevoj mreži koristimo da je

$$\max_{x\in[-1,1]} |2^{-2}T_3| = \frac{1}{4},$$

pa je uniformna ocjena pogreške dana sa

$$\max_{x\in[-1,1]} |f(x) - r_2(x)| \leq 24.315.$$

△

U prošlom zadatku, prave maksimalne pogreške su:

- 0.646229 za ekvidistantnu mrežu (postiže se u $x = 0.404921$),
- 0.6005977 za Čebiševljevu mrežu (postiže se u $x = 0.371166$).

3.5 Hermiteova interpolacija

Neka su zadane točke (x_i, y_i) i (x_i, y_i'), $i = 0, 1, \ldots, n$. Tražimo interpolacijski polinom h takav da

$$h(x_i) = y_i, \quad h'(x_i) = y_i', \quad i = 0, 1, \ldots, n.$$

U ovom slučaju govorimo o **Hermiteovoj** interpolaciji. Takav polinom postoji i jedinstven je.

Ako ovakav interpolacijski problem rješavamo pomoću Newtonovog interpolacijskog polinoma, uočavamo da ne znamo izračunati $f[x_i, x_i]$, jer dobivamo $0/0$. No vrijedi:

$$f[x_i, x_i] = \lim_{h\to 0} f[x_i, x_i + h] = \lim_{h\to 0} \frac{f(x_i + h) - f(x_i)}{x_i + h - x_i} = f'(x_i).$$

Vrijedi i općenitiji rezultat:

$$f\underbrace{[x_i, \ldots, x_i]}_{k+1} = \frac{f^{(k)}(x_i)}{k!}.$$

Zadatak 3.6. Konstruirajte interpolacijski polinom koji zadovoljava sljedeće uvjete:

x_i	$f(x_i)$	$f'(x_i)$
1	2	3
2	6	7

Rješenje. Tablica podijeljenih razlika je

x_i \ k	0	1	2	3
1	2			
		3		
1	2		1	
		4		2
2	6		3	
		7		
2	6			

Interpolacijski polinom je

$$p(x) = 2 + 3(x-1) + (x-1)^2 + 2(x-1)^2(x-2).$$

△

Neka je $h_{2n+1}(x)$ Hermiteov interpolacijski polinom za funkciju f na mreži čvorova $x_0, \ldots, x_n$. Ako na $[x_0, x_n]$ postoji $f^{(2n+2)}$ tada je ocjena greške interpolacije za $x \in [x_0, x_n]$ jednaka

$$|f(x) - h_{2n+1}(x)| \leq \omega^2(x) \frac{M_{2n+2} f}{(2n+2)!}.$$

Zadatak 3.7. Funkciju $f(x) = \ln(x)$ aproksimiramo Hermiteovim interpolacijskim polinomom stupnja 3 na čvorovima $x_0 = 1$ i $x_1 = 2$. Nađite $h_3(1.5)$ i izračunajte pravu grešku te ocjenu greške interpolacije.

Rješenje. Pravimo tablicu podijeljenih razlika pa je

x_i \ k	0	1	2	3
1	0			
		1		
1	0		$\ln 2 - 1$	
		$\ln 2$		$3/2 - 2\ln 2$
2	$\ln 2$		$1/2 - \ln 2$	
		$1/2$		
2	$\ln 2$			

$$h_3(x) = (x-1) + (\ln 2 - 1)(x-1)^2 + \left(\frac{3}{2} - 2\ln 2\right)(x-1)^2(x-2).$$

Vrijedi $h_3(1.5) = \frac{1}{2}\ln 2 + \frac{1}{16} = 0.40907359028$ i $\ln(1.5) = 0.405465108108$, pa je prava greška $|\ln(1.5) - h_3(1.5)| = 3.60848217181 \cdot 10^{-3}$.

Kako je $f^{(4)} = -\frac{6}{x^4}$ imamo da je $M_4 f = 6$. S druge strane, $\omega(1.5)^2 = 0.5^4 = 1/16$. Dakle, ukupno imamo

$$|\ln(1.5) - h_3(1.5)| \leq \frac{1}{16} \cdot \frac{6}{4!} = 0.015624.$$

△

3.6 Po dijelovima linearna interpolacija

Neka je dana mreža

$$a = x_0 < x_1 < \ldots < x_n = b$$

na $[a, b]$. **LINEARNI SPLINE** na toj mreži, za funkciju f je ona funkcija čija je restrikcija na $[x_{i-1}, x_i]$ linearni polinom:

$$\ell_i(x) = \frac{x - x_{i-1}}{x_i - x_{i-1}} f(x_i) + \frac{x_i - x}{x_i - x_{i-1}} f(x_{i-1}), \quad \text{za } x \in [x_{i-1}, x_i].$$

Ako je funkcija $f \in C^2([a, b])$ tada je ocjena uniformne pogreške interpolacije

$$\max_{x \in [a,b]} |f(x) - \ell(x)| \leq \frac{1}{8} \max_{i=1,2,\ldots,n} (x_i - x_{i-1})^2 M_2 f.$$

Zadatak 3.8. Aproksimiramo funkciju $f(x) = \ln x$ na $[1, 100]$ po dijelovima linearnom interpolacijom. Fiksiramo traženu točnost $\varepsilon = 10^{-4}$ koju zahtijevamo na čitavom intervalu. Nađite broj čvorova da se postigne tražena točnost uz

a) ekvidistantnu mrežu s korakom h na čitavom $[1, 100]$,

b) podijelimo interval na tri dijela : $[1, 2], [2, 7], [7, 100]$ te napravimo ekvidistantnu mrežu na svakom od njih s koracima h_1, h_2 i h_3.

Rješenje.

$$f(x) = \ln x, \quad f'(x) = \frac{1}{x}, \quad f''(x) = -\frac{1}{x^2}.$$

Vidimo da na intervalu $[a, b]$, funkcija $|f''|$ postiže svoj maksimum u lijevom rubu intervala (strogo padajuća).

a) Koristeći ocjenu pogreške dobivamo da je na $[1, 100]$, uz ekvidistantnu mrežu s korakom h, pogreška interpolacije

$$\max_{x \in [1,100]} |f(x) - l(x)| \leq \frac{1}{8} h^2 |f''(1)| = \frac{h^2}{8}.$$

Tražimo h tako da je

$$\frac{h^2}{8} < \varepsilon = 10^{-4}.$$

To znači da mora biti

$$h < 10^{-2}\sqrt{8} = 0.028284.$$

$$h = \frac{b-a}{n} \implies n = \frac{b-a}{h} > \frac{99}{10^{-2}\sqrt{8}} = 3500.17$$

Prema tome broj podsegmenata mora biti $n = 3501$, tj. broj čvorova je 3502.

b) 1. interval [1,2]: $M_2 f = |f''(1)| = 1$, dakle

$$h_1^2 < 8 \cdot 10^{-4} \implies h_1 < 0.02828427.$$

$$h_1 = \frac{2-1}{n_1} \implies n_1 > \frac{1}{h_1} = 35.35$$

Dakle $n_1 = 36$

2. interval [2,7]: $M_2 f = |f''(2)| = \frac{1}{4}$, dakle

$$h_2^2 < 32 \cdot 10^{-4} \implies h_2 < 0.05656824.$$

$$h_2 = \frac{7-2}{n_2} \implies n_2 > \frac{5}{h_2} = 88.38$$

Dakle $n_2 = 89$

3. interval [7,100]: $M_2 f = |f''(7)| = \frac{1}{49}$, dakle

$$h_3^2 < 392 \cdot 10^{-4} \implies h_3 < 0.197989.$$

$$h_3 = \frac{100-7}{n_3} \implies n_3 > \frac{93}{h_3} = 469.72$$

Dakle $n_3 = 470$

Dakle ukupno nam treba $(n_1 + 1) + (n_2 + 1) + (n_3 + 1) - 2 = 596$ čvorova.

△

Zadatak 3.9. Funkciju $f(x) = \frac{1}{1+2x}$ aproksimiramo na ekvidistantnoj mreži $[0, 1]$ s 13 čvorova.

- polinomom stupnja 12;
- po dijelovima linearnom interpolacijom.

Izračunajte ocjenu uniformne pogreške u oba slučaja. Izračunajte vrijednost funkcije u $x_0 = 0.7$ i jedne od aproksimacija u istoj točki, te pravu pogrešku.

Rješenje. • Kako je $n = 12$, $h = 1/12$. Koristimo ocjenu

$$\max_{x\in[0,1]} |f(x) - p(x)| \leq \frac{h^{n+1}}{n+1} M_{n+1} f.$$

Ovdje nam treba maksimum 13. derivacije. Pogledajmo prvih nekoliko derivacija funkcije f:

$$f'(x) = \frac{-2}{(1+2x)^2}, \quad f''(x) = \frac{8}{(1+2x)^3}, \quad f'''(x) = \frac{-48}{(1+2x)^4}.$$

Zaključujemo da derivacija općenito izgleda (može se dokazati indukcijom)

$$f^{(n)} = \frac{(-1)^n 2^n n!}{(1+2x)^{(n+1)}}.$$

Na $[0,1]$ je $|f^{(13)}(x)| = f^{(13)}(x) = \frac{2^{13}13!}{(1+2x)^{14}}$. $f^{(13)}$ je na tom intervalu padajuća, pa se maksimum postiže u lijevom rubu intervala, tj. $M_{13}f = 2^{13}13!$. Uvrštavanjem u ocjenu greške dobivamo

$$\max_{x\in[0,1]} |f(x) - p(x)| \leq \frac{\frac{1}{12^{13}}}{13} 2^{13}13! = 0.036675.$$

- Za ocjenu uniformne pogreške koristimo formulu

$$\max_{x\in[a,b]} |f(x) - \ell(x)| \leq \frac{1}{8} \max_{i=1,2,\ldots,n} (x_i - x_{i-1})^2 M_2 f.$$

Kako smo na ekvidistantnoj mreži, vrijedi da je $\max_{i=1,2,\ldots,n}(x_i - x_{i-1})^2 = 1/12^2$. Kako je $|f''(x)| = f''(x)$ na $[0,1]$ i kako je $f''(x)$ padajuća na tom intervalu, zaključujemo da se maksimum postiže u lijevom rubu intervala, tj. $M_2 f = 8$. Ukupno imamo:

$$\max_{x\in[a,b]} |f(x) - \ell(x)| \leq \frac{1}{8} \cdot \frac{1}{12^2} \cdot 8 = 0.0069444.$$

Aproksimaciju ćemo izračunati koristeći linearni spline. Kako je $0.7 \in [x_8, x_9] = [8/12, 9/12] = [2/3, 3/4]$ računamo $\ell_9(0.7)$

$$\ell_9(x) = -\frac{12}{35}x + \frac{23}{35}, \quad \ell_9(0.7) = 0.41714286.$$

Kako je $f(0.7) = 0.4166666$, prava greška je $|f(0.7) - \ell_9(0.7)| = 4.7619 \cdot 10^{-4}$. △

3.7 Po dijelovima kubična interpolacija

Funkciju f aproksimiramo funkcijom φ koja je na svakom podintervalu $[x_{i-1}, x_i]$ kubični polinom, tj.

$$\varphi|_{[x_{i-1},x_i]} = p_i,$$

gdje je svaki polinom p_i stupnja 3 i zapisujemo ga relativno s obzirom na lijevu točku intervala:

$$p_i(x) = c_{0,i} + c_{1,i}(x - x_{i-1}) + c_{2,i}(x - x_{i-1})^2 + c_{3,i}(x - x_{i-1})^3.$$

Za svaki polinom p_i treba odrediti 4 koeficijenta. Uvjeti su sljedeći

$$\begin{aligned} p_k(x_{k-1}) &= f(x_{k-1}) \\ p_k(x_k) &= f(x_k) \\ p_k'(x_{k-1}) &= s_{k-1} \\ p_k'(x_k) &= s_k \end{aligned}$$

Ako je $s_k = f'(x_k)$ tada govorimo o **kubičnoj Hermiteovoj interpolaciji.**

Ako je funkcija $f \in C^4([a,b])$ i $h = \max_{k=1,\dots,n} |x_k - x_{k-1}|$ tada vrijedi sljedeća ocjena pogreške interpolacije

$$|f(x) - \varphi(x)| \leq \frac{1}{384} h^4 M_4,$$

dok za ocjenu uniformne pogreške vrijedi

$$\max_{x\in[a,b]} |f(x) - \varphi(x)| \leq \frac{(b-a)^4 M_4}{384 n^4}.$$

Zadatak 3.10. Funkciju $f(x) = 2\sqrt{1+x}$ aproksimiramo na intervalu $[0,4]$.

a) Odredi najmanji broj podintervala n tako da ocjena uniformne pogreške na $[0,4]$ ne prelazi $\varepsilon = 10^{-2}$ ako funkciju aproksimiramo po dijelovima kubnom Hermiteovom interpolacijom.

b) Za n iz podzadatka a) odredite aproksimaciju u točki $x = 0.5$ i nađite pravu pogrešku u toj točki.

Detaljno obrazložite tvrdnje vezane uz ocjenu greške.

Rješenje. a) Ekvidistantna mreža sa n podintervala u intervalu $[a,b]$ ima korak $h = (b-a)/n$. Iz zahtjeva da je ocjena manja ili jednaka ε dobivamo ocjenu za broj podintervala n

$$n \geq (b-a)\sqrt[4]{\frac{M_4}{384\varepsilon}}.$$

Funkcija f i njene derivacije su redom

$$\begin{aligned} f(x) &= 2\sqrt{1+x}, \\ f'(x) &= \frac{1}{\sqrt{1+x}}, \\ f''(x) &= -\frac{1}{2(1+x)^{3/2}}, \\ f'''(x) &= \frac{3}{4(1+x)^{5/2}}, \\ f^{(4)} &= -\frac{15}{8(x+1)^{7/2}}. \end{aligned}$$

Na intervalu [0,4] $|f^{(4)}(x)| = \frac{15}{8(x+1)^{7/2}}$ je padajuća funkcija pa se maksimum postiže u lijevom rubu, tj. $M_4 = |f^{(4)}(0)| = 15/8$.

Uz traženu točnost $\varepsilon = 10^{-2}$, za broj podintervala n mora vrijediti

$$n \geq 4\sqrt[4]{\frac{15/8}{384 \cdot 10^{-2}}} = 3.34370152488,$$

tj. potrebno nam je $n = 4$ podintervala, odnosno 5 čvorova.

b) Kako je $0.5 \in [0, 1]$ računamo kubni polinom na prvom podintervalu. Računamo podijeljene razlike Dakle, polinom $p_1(x)$ je

x_i \ k	0	1	2	3
0	2			
		1		
0	2		$2\sqrt{2}-3$	
		$2\sqrt{2}-2$		$5-\frac{7\sqrt{2}}{2}$
1	$2\sqrt{2}$		$2-\frac{3\sqrt{2}}{2}$	
		$\frac{\sqrt{2}}{2}$		
1	$2\sqrt{2}$			

$$p_1(x) = 2 + x + (2\sqrt{2} - 3)x^2 + \left(5 - \frac{7}{2}\sqrt{2}\right) x^2 (x - 1).$$

Prava greška u točki 0.5 je

$$|f(0.5) - p(0.5)| = \left|\sqrt{6} - \frac{18 + 15\sqrt{2}}{16}\right| = |2.44948974278 - 2.45082521472| = 1.3354719416 \cdot 10^{-3}.$$

△

3.8 Kubični spline

Brojeve $s_0, \ldots, s_n$ određujemo iz zahtjeva da φ ima neprekidnu drugu derivaciju

$$\begin{aligned} \varphi(x_{k-1}) &= f(x_{k-1}) \\ \varphi(x_k) &= f(x_k) \\ p_k'(x_k) &= p_{k+1}'(x_k) \\ p_k''(x_k) &= p_{k+1}''(x_k) \end{aligned}$$

Uvrštavanjem i sređivanjem dobivamo sustav jednadžbi u nepoznanicama s_k:

$$h_{k+1}s_{k-1} + 2(h_k + h_{k+1})s_k + h_k s_{k+1} = 3(h_{k+1} f[x_{k-1}, x_k] + h_k f[x_k, x_{k+1}]), \quad k = 1, \ldots, n-1$$

Dakle, imamo $n - 1$ jednadžbi i $n + 1$ nepoznanica. Jedinstvenost rješenja ćemo imati ako su zadani s_0 i s_n.

Zadatak 3.11. Za funkciju su poznati sljedeći podatci:

x_i	$f(x_i)$	$f'(x_i)$
0	-1	-3
2	1	?
3	1	?
4	2	?
6	0	-9

Potpunim kubičnim splajnom ($s_0 = f'(x_0)$, $s_n = f'(x_n)$) aproksimirajte vrijednost ove funkcije u $x = 3.5$.

Rješenje. Prvo ćemo naći vrijednosti $s_0, \ldots, s_n$ rješavanjem sustava, a onda naći Hermiteov interpolacijski polinom na odgovarajućem intervalu, u koji ćemo na kraju uvrstiti vrijednost $x = 3.5$.

U ovom primjeru je $n = 4$. Za vrijednosti h_k redom imamo: $h_1 = x_1 - x_0 = 2$, $h_2 = 1$, $h_3 = 1$, $h_4 = 2$. Za podijeljene razlike prvog reda popunjavamo tablicu kao kod Newtonovog oblika interpolacijskog polinoma, no stajemo nakon drugog stupca:

x_i \ k	0	1
0	-1	
		1
2	1	
		0
3	1	
		1
4	2	
		-1
6	0	

Zato je $n - 1 = 3$ jednadžbi koje određuju vrijednosti $s_0, \ldots, s_4$ dano s

$$\begin{aligned} 1 \cdot s_0 + 2 \cdot (1+2)s_1 + 2s_2 &= 3(1 \cdot (-1) + 2 \cdot 1) \\ 1 \cdot s_1 + 2 \cdot (1+1)s_2 + 1s_3 &= 3(1 \cdot 1 + 1 \cdot 1) \\ 2 \cdot s_2 + 2 \cdot (2+1)s_3 + 1s_4 &= 3(2 \cdot 1 + 1 \cdot 2) \end{aligned}$$

odnosno (korištenjem $s_0 = f'(x_0) = -3$ i $s_4 = f'(x_4) = -9$):

$$\begin{aligned} 6s_1 + 2s_2 &= 6 \\ s_1 + 4s_2 + s_3 &= 3 \\ 2s_2 + 6s_3 &= 12 \end{aligned}$$

Pomnožimo drugu jednadžbu sa 6, te oduzmimo prvu i treću od nje, tako ćemo se riješiti nepoznanica s_1 i s_3. Dobivamo $20s_2 = 0$, odakle je $s_2 = 0$. Iz treće jednadžbe je onda $s_3 = 2$, a iz prve $s_1 = 1$.

Kako je $3.5 \in [3,4]$, tražimo odgovarajući Hermiteov interpolacijski polinom na tom intervalu. Podatke za $f[3,3]$ i $f[4,4]$ mijenjamo vrijednostima $s_2 = 0$ i $s_3 = 2$.

x_i \ k	0	1	2	3
3	1			
		0		
3	1		1	
		1		0
4	2		1	
		2		
4	2			

Na tom intervalu splajn glasi $p(x) = 1 + (x-3)^2$, pa je $f(3.5) \approx p(3.5) = 1 + \frac{1}{4} = \frac{5}{4}$. $\triangle$

Zadatak 3.12. (DZ) Funkciju $f(x) = (x+1)\sin x$ interpoliramo potpunim kubičnim splajnom na ekvidistantnoj mreži s $n = 4$ podintervala na intervalu $[0, \pi/2]$. Potpuni kubični spline definira $s_0 = f'(x_0)$ i $s_4 = f'(x_4)$. Dodatno su zadani $s_2 = 1.9689828101$ i $s_3 = 1.7564789686$. Izračunajte vrijednost tog splajna u točki $x = \pi/6$ i pripadnu pravu pogrešku.

Rješenje. Kako je $\frac{\pi}{6} \in [x_1, x_2] = [\frac{\pi}{8}, \frac{\pi}{4}]$ potrebno je izračunati s_1. To možemo koristeći npr. prvu jednadžbu (uzeli smo u obzir da smo na ekvidistantnoj mreži pa su svi $h_i = h = \pi/8$):

$$4s_1 + s_2 = 3(f[x_0, x_1] + f[x_1, x_2]) - s_0.$$

Kako je $f'(x) = \sin x + (x+1)\cos x$ možemo izračunati $s_0 = f'(0) = 1$. Podijeljenje razlike su redom

$$f[x_0, x_1] = \frac{f(\pi/8) - f(0)}{\frac{\pi}{8}} = 1.3571787908,$$
$$f[x_1, x_2] = \frac{f(\pi/4) - f(\pi/8)}{\frac{\pi}{8}} = 1.8576674039.$$

Uvrštavanjem izračunamo $s_1 = 1.6688889435$. Sada računamo podijeljene razlike za polinom p_2 na drugom podintervalu Interpolacijski polinom onda glasi

$$\begin{aligned} p_2(x) =& 0.5329628648 + 1.6688889435(x - \pi/8) + 0.4807204020(x - \pi/8)^2 - \\ & 0.5023134940(x - \pi/8)^2(x - \pi/4). \end{aligned}$$

$p_2(\pi/6) = 0.7619102398$ pa je prava pogreška $|f(\pi/6) - p_2(\pi/6)| = 0.0001108520$. $\triangle$

x_i \ k	0	1	2	3
$\pi/8$	0.5329628648			
		1.6688889435		
$\pi/8$	0.5329628648		0.4807204020	
		1.8576674039		-0.5023134940
$\pi/4$	1.2624671485		0.2834623542	
		1.9689828101		
$\pi/4$	1.2624671485			

Zadatak 3.13. Nađite kubični splajn s koji interpolira sljedeći skup podataka (točaka): Za nalaženje splajna iskoristite "not-a-knot" rubni uvjet, tj. uvjet $p_1 = p_2$ i $p_{n-1} = p_n$.

x_i	-3	-2	2	3
y_i	1	2	2	1

Izračunajte vrijednost interpolacijskog splajna u točki $x = 0$.

Rješenje. U zadatku imamo $n = 3$ podintervala, pa slijedi da mora biti

$$p_1 = p_2 = p_3.$$

Dakle, s je obični interpolacijski polinom za zadane 4 točke. Računamo tablicu podijeljenih razlika Dakle, $s(x)$ je

x_i \ k	0	1	2	3
-3	1			
		1		
-2	2		$-1/5$	
		0		0
2	2		$-1/5$	
		-1		
3	1			

$$\begin{aligned} s(x) &= 1 + (x+3) - \frac{1}{5}(x+3)(x+2) + 0 \cdot (x+3)(x+2)(x-2) \\ &= -\frac{1}{5}x^2 + \frac{14}{5}. \end{aligned}$$

U točki $x = 0$ je $s(0) = 2.8$.

DZ. Riješite zadatak koristeći sustav. △

4

Problem najmanjih kvadrata

4.1 Diskretni problem najmanjih kvadrata

Za zadanu funkciju $f : X \to \mathbb{R}$ naći što bolju aproksimacijsku funkciju φ među svim funkcijama $\psi \in \mathcal{S}$, tj. za zadanu normu $\|\cdot\|$, tražimo $\varphi \in \mathcal{S}$ takvu da je

$$\|f - \varphi\| = \inf_{\psi \in \mathcal{S}} \|f - \psi\|.$$

Mi ćemo promatrati sljedeće:

- $X = \{t_1, t_2, \ldots, t_n\} \subset \mathbb{R}$ - **DISKRETNI PROBLEM NAJMANJIH KVADRATA**;
- $\mathcal{S} = \mathcal{P}_m$;
- norma je Euklidska u istaknutim točkama domene:

$$\|g\| = \left(\sum_{i=1}^{n} g(t_i)^2 \right)^{1/2};$$

- dimenzija m je mnogo manja od broja čvorova n.

Neka je $\varphi(x) = a_0\varphi_0(x) + a_1\varphi_1(x) + \ldots + a_m\varphi_m(x)$ linearna u $a_0, \ldots, a_m$. Sada je funkcija koju minimiziramo oblika:

$$S := S(a_0, a_1, \ldots, a_m) = \|f - \varphi\| = \sum_{k=1}^{n} (f(t_k) - \varphi(t_k))^2.$$

Nužan uvjet egzistencije ekstrema je

$$\frac{\partial S}{\partial a_i} = 0, \quad i = 0, 1, \ldots, m.$$

Ovaj sustav zovemo **sustav normalnih jednadžbi**.

Zadatak 4.1. Izmjereni su sljedeći podaci:

x_k	0	1	2	3	4
f_k	3.8	3.7	4.0	3.9	4.3

Nađite aproksimaciju pravcem metodom najmanjih kvadrata za taj skup podataka.

Rješenje. Tražimo funkciju φ oblika

$$\varphi(x) = ax + b.$$

Treba minimizirati funkciju

$$S = \sum_{k=0}^{4} \left(f_k - \varphi(x_k)\right)^2 = \sum_{k=0}^{4} \left(f_k - ax_k - b\right)^2.$$

Tražimo stacionarne točke od S:

$$\begin{aligned} \frac{\partial S}{\partial a} &= -2 \sum_{k=0}^{4} \left(f_k - ax_k - b\right) x_k = 0 \\ \frac{\partial S}{\partial b} &= -2 \sum_{k=0}^{4} \left(f_k - ax_k - b\right) \quad = 0 \end{aligned}$$

Sređivanjem dobivamo sustav 2×2:

$$\begin{aligned} a \sum_{k=0}^{4} x_k^2 + b \sum_{k=0}^{4} x_k &= \sum_{k=0}^{4} f_k x_k \\ a \sum_{k=0}^{4} x_k + 5b &= \sum_{k=0}^{4} f_k \end{aligned}$$

Napravimo tablicu koeficijenata za sustav:

x_k	f_k	x_k^2	$x_k f_k$
0	3.8	0	0.0
1	3.7	1	3.7
2	4.0	4	8.0
3	3.9	9	11.7
4	4.3	16	17.2
$\sum:$ 10	19.7	30	40.6

Sustav sada glasi:

$$30a + 10b = 40.6$$
$$10a + 5b = 19.7.$$

Rješavanjem sustava dobivamo:

$$a = 0.12, \quad b = 3.7,$$

pa tražena aproksimacijska funkcija φ glasi:

$$\varphi(x) = 0.12x + 3.7.$$

△

Zadatak 4.2. Odredite funkciju oblika $\varphi(x) = be^{ax}$ koja aproksimira podatke

x_k	1	2	3	4
f_k	7	11	17	27

metodom najmanjih kvadrata.

Rješenje. Kako funkcija φ nije linearna (u koeficijentu a), treba je linearizirati. U ovom slučaju imamo:

$$\psi(x) = \ln \varphi(x) = \ln (be^{ax}) = \ln b + ax.$$

Uvedimo oznake:

$$A = a, \quad B = \ln b.$$

Dakle linearizirana funkcija je $\psi(x) = Ax + B$. Definirati ćemo još $h_k = \ln f_k$. Sada imamo novu tablicu podataka:

x_k	1	2	3	4
h_k	$\ln 7 = 1.94591$	$\ln 11 = 2.397895$	$\ln 17 = 2.83321$	$\ln 27 = 3.29584$

Minimiziramo funkciju

$$S = S(A, B) = \sum_{k=0}^{3} (h_k - \psi(x_k))^2 = \sum_{k=0}^{3} (h_k - Ax_k - B)^2 .$$

Uvjeti za minimum daju:

$$\frac{\partial S}{\partial A} = -2 \sum_{k=0}^{3} (h_k - Ax_k - B)\, x_k = 0$$
$$\frac{\partial S}{\partial b} = -2 \sum_{k=0}^{3} (h_k - Ax_k - B) \qquad = 0$$

tj., sređivanjem dobivamo sustav

$$\begin{aligned} A\sum_{k=0}^{3} x_k^2 + B\sum_{k=0}^{3} x_k &= \sum_{k=0}^{3} h_k x_k \\ A\sum_{k=0}^{3} x_k + 4B &= \sum_{k=0}^{3} h_k \end{aligned}$$

Izračunajmo koeficijente sustava:

x_k	h_k	x_k^2	$x_k h_k$
1	1.94591	1	1.94591
2	2.39795	4	4.7959
3	2.83321	9	8.49963
4	3.29584	16	13.18336
$\sum:$ 10	10.47291	30	28.4248

Sustav sada glasi:

$$\begin{aligned} 30A + 10B &= 28.4248 \\ 10A + 4B &= 10.47291 \end{aligned}$$

Rješavanjem sustava dobivamo:

$$A = 0.448505, \quad B = 1.496965.$$

Vraćamo zamjenu:

$$a = A = 0.0.448505, \quad b = e^B = 4.46812.$$

Pa funkcija φ glasi:

$$\varphi(x) = 4.46812e^{0.448505\cdot x}.$$

△

Zadatak 4.3. U ovisnosti o zadanom skupu podataka (x_k, f_k), $k = 0, \ldots, n$, diskretnom metodom najmanjih kvadrata odredite parametre funkcije oblika

$$\varphi(x) = a_3 x^2 + a_2 x^2 + x$$

koja zadovoljava uvjet $\varphi'(-1) = 1$, te najbolje aproksimira zadane podatke.

Rješenje. `https://web.math.pmf.unizg.hr/nastava/unm/vjezbe/nm_dodzad.pdf`, str 98. △

4.1.1 Matrični zapis diskretnog problema najmanjih kvadrata

Neka je $\{\varphi_1, \ldots, \varphi_m\}$ baza za $\mathcal{S}$. Tražimo funkciju

$$\varphi(t) = \sum_{i=1}^{m} x_i \varphi_i(t)$$

koja najbolje aproksimira $f : \{t_1, \ldots, t_n\} \to \mathbb{R}$ određenu s $f(t_i) = y_i$. Neka je $A \in \mathbb{M}_{n\times m}$ matrica takva da je $a_{ij} = \varphi_j(t_i)$ i $b \in \mathbb{R}^n$ vektor takav da je $b_i = y_i$. Tada je cilj minimizirati

$$\|f - \varphi\|^2 = \|Ax - b\|_2^2.$$

Rješenje $x \in \mathbb{R}^m$ postoji i jedinstveno je te je dano kao rješenje sustava normalnih jednadžbi

$$A^T Ax = A^T b.$$

4.2 QR faktorizacija

Definicija 4.4. Neka je $A \in \mathbb{M}_{n\times m}$, $n \geq m$, matrica punog stupčanog ranga m. Njezina QR faktorizacija je rastav oblika

$$A = QR = Q \begin{bmatrix} R_0 \\ 0 \end{bmatrix},$$

pri čemu je $Q \in \mathbb{M}_n$ ortogonalna matrica te $R_0 \in \mathbb{M}_m$ gornjetrokutasta matrica s pozitivnim elementima na dijagonali. Skraćeni oblika QR faktorizacije je

$$A = Q_0 R_0,$$

pri čemu je Q_0 dobivena od prvih m stupaca matrice Q.

QR faktorizaciju možemo izračunati Gram – Schmidtovim postupkom. Označimo stupe matrice $A = \begin{bmatrix} a_1 & a_2 & \ldots & a_n \end{bmatrix}$.

Za $i = 1$, $q_1' = a_1$, $q_1 = \dfrac{q_1'}{\|q_1'\|_2}$; za $i > 1$:

$$q_j' = a_j - \sum_{j=1}^{j-1} \langle a_j, q_i \rangle q_i, \quad q_j = \frac{q_j'}{\|q_j'\|_2}.$$

Sada je $Q = \begin{bmatrix} q_1 & q_2 & \ldots & q_n \end{bmatrix}$ i $R = \begin{bmatrix} \|q_1'\| & \langle q_1, a_2 \rangle & \ldots & \langle q_1, a_n \rangle \\ 0 & \|q_2'\| & \ldots & \langle q_2, a_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \|q_n'\| \end{bmatrix}$

Zadatak 4.5. Odredite QR faktorizaciju matrice

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

Rješenje.

$$\begin{aligned} q_1' &= a_1 = \begin{bmatrix} 2 & -2 & -2 & -2 \end{bmatrix}^T \\ r_{11} &= \|q_1'\| = \sqrt{2^2 + (-2)^2 + (-2)^2 + (-2)^2} = 4 \\ q_1 &= \frac{q_1'}{\|q_1'\|} = \begin{bmatrix} 1/2 & -1/2 & -1/2 & -1/2 \end{bmatrix}^T \end{aligned}$$

$$\begin{aligned} r_{12} &= \langle q_1, a_2 \rangle = \frac{1}{2} \cdot 2 + \left(-\frac{1}{2}\right) \cdot (-2) + \left(-\frac{1}{2}\right) \cdot 0 + \left(-\frac{1}{2}\right) \cdot 0 = 2 \\ q_2' &= a_2 - r_{12} q_1 = \begin{bmatrix} 2 & -2 & 0 & 0 \end{bmatrix}^T - 2 \begin{bmatrix} 1/2 & -1/2 & -1/2 & -1/2 \end{bmatrix}^T \\ &= \begin{bmatrix} 1 & -1 & 1 & 1 \end{bmatrix}^T \\ r_{22} &= \|q_2'\| = \sqrt{1^2 + (-1)^2 + 1^2 + 1^2} = 2 \\ q_2 &= \frac{q_2'}{\|q_2'\|} = \begin{bmatrix} 1/2 & -1/2 & 1/2 & 1/2 \end{bmatrix}^T \end{aligned}$$

$$\begin{aligned} r_{13} &= \langle q_1, a_3 \rangle = \frac{1}{2} \cdot 2 + \left(-\frac{1}{2}\right) \cdot 4 + \left(-\frac{1}{2}\right) \cdot (-4) + \left(-\frac{1}{2}\right) \cdot 2 = 0 \\ r_{23} &= \langle q_2, a_3 \rangle = \frac{1}{2} \cdot 2 + \left(-\frac{1}{2}\right) \cdot 4 + \frac{1}{2} \cdot (-4) + \frac{1}{2} \cdot 2 = -2 \\ q_3' &= a_3 - r_{13} q_1 - r_{23} q_2 = \begin{bmatrix} 2 & 4 & -4 & 2 \end{bmatrix}^T + 2 \begin{bmatrix} 1/2 & -1/2 & 1/2 & 1/2 \end{bmatrix}^T \\ &= \begin{bmatrix} 3 & 3 & -3 & 3 \end{bmatrix} \\ r_{33} &= \|q_3'\| = \sqrt{3^2 + 3^2 + (-3)^2 + 3^2} = 6 \\ q_3 &= \frac{q_3'}{\|q_3'\|} = \begin{bmatrix} 1/2 & 1/2 & -1/2 & 1/2 \end{bmatrix} \end{aligned}$$

$$
\begin{aligned}
r_{14} &= \langle q_1, a_4 \rangle = \frac{1}{2} \cdot 3 + \left(-\frac{1}{2}\right) \cdot 1 + \left(-\frac{1}{2}\right) \cdot 3 + \left(-\frac{1}{2}\right) \cdot 3 = -2 \\
r_{24} &= \langle q_2, a_4 \rangle = \frac{1}{2} \cdot 3 + \left(-\frac{1}{2}\right) \cdot 1 + \frac{1}{2} \cdot 3 + \frac{1}{2} \cdot 3 = 4 \\
r_{34} &= \langle q_3, a_4 \rangle = \frac{1}{2} \cdot 3 + \frac{1}{2} \cdot 1 + \left(-\frac{1}{2}\right) \cdot 3 + \frac{1}{2} \cdot 3 = 2 \\
q_4' &= a_4 - r_{14} q_1 - r_{24} q_2 - r_{34} q_3 \\
&= \begin{bmatrix} 1 & 1 & 1 & -1 \end{bmatrix} \\
r_{44} &= \|q_4'\| = 2 \\
q_4 &= \frac{q_4'}{\|q_4'\|} = \begin{bmatrix} 1/2 & 1/2 & 1/2 & -1/2 \end{bmatrix}^T
\end{aligned}
$$

Dakle, QR faktorizacija matrice A ima sljedeći oblik:

$$
Q = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{bmatrix}, \quad R = \begin{bmatrix} 4 & 2 & 0 & -2 \\ 0 & 2 & -2 & 4 \\ 0 & 0 & 6 & 2 \\ 0 & 0 & 0 & 2 \end{bmatrix}
$$

△

QR faktorizacija se koristi za rješavanje problema najmanjih kvadrata. Kako je

$$
\|Ax - b\|_2^2 = \|R_0 x - Q_0^T b\|_2^2 + \|Q_1^T b\|_2^2
$$

vrijedi sljedeće

$$
\min_{x \in \mathbb{R}^m} \|Ax - b\|_2^2 = \|Q_1^T b\|_2^2
$$

i taj se minimum postiže za x koji je rješenje sustava $R_0 x = Q_0^T b$.

Zadatak 4.6. Matrica A dana je svojom skraćenom QR faktorizacijom:

$$
A = \underbrace{\begin{bmatrix} 0.7 & \alpha & -0.5 \\ 0.1 & 0.7 & -0.5 \\ \beta & -0.5 & 0.1 \\ 0.5 & \gamma & 0.7 \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} \delta & 2 & 3 \\ 0 & 2 & -2 \\ 0 & 0 & 1 \end{bmatrix}}_{R}.
$$

(a) Odredite sve moguće vrijednosti za realne konstante α, β, γ i δ iz gornje jednakosti ako je poznato da je element matrice A na poziciji $(1, 1)$ jednak 0.7.

(b) Odredite sva rješenja problema najmanjih kvadrata $\min \|Ax - b\|_2$ za vektor $b = e_1$ (prvi vektor kanonske baze).

(c) Odredite Frobeniusovu normu matrice A.

(d) Odredite sve moguće potpune QR faktorizacije matrice A.

Rješenje. Kako bi A na poziciji $(1,1)$ imala vrijednost 0.7, umnožak prvog retka matrice Q s prvim stupcem matrice R treba biti jednako 0.7, odakle je $\delta = 1$.

Iz ortogonalnosti prvog i trećeg stupca matrice Q slijedi

$$0 = 0.7 \cdot (-0.5) + 0.1 \cdot (-0.5) + \beta \cdot 0.1 + 0.5 \cdot 0.7 \implies \beta = 0.5$$

Iz ortogonalnosti drugog stupca s preostalima, imamo

$$\begin{aligned} 0.7\alpha + 0.07 - 0.25 + 0.5\gamma &= 0 \\ -0.5\alpha - 0.35 - 0.05 + 0.7\gamma &= 0 \end{aligned}$$

Množeći prvu jednakost s 0.5 i drugu s 0.7 pa zbrajanjem, dobivamo $0.74\gamma = 0.37$, odnosno $\gamma = 0.5$. Iz bilo koje od gornjih jednakosti dobivamo $\alpha = -0.1$.

Formula za problem najmanjih kvadrata QR faktorizacijom glasi $x = R^{-1}Q^T b$. Konkretno, treba riješiti sustav

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0.7 \\ -0.1 \\ -0.5 \end{bmatrix}$$

Povratnom supstitucijom dobivamo $x_3 = -0.5$, $2x_2 = -1.1$, tj. $x_2 = -0.55$, te $x_1 = 3.3$.

Frobeniusova norma matrice A jednaka je Frobeniusovoj normi matrice R. Iznosi $\|A\|_F = \sqrt{1^2 + 2^2 + 3^2 + 2^2 + (-2)^2 + 1^2} = \sqrt{23}$.

Zadnji stupac matrice $\tilde{Q}$ (proširenja matrice Q do kvadratne ortogonalne matrice) leži u ortogonalnom komplementu trodimenzionalnog potprostora od R^4 razapetog sa stupcima matrice Q. Kako zadnji stupac matrice $\tilde{Q}$ leži u jednodimenzionalnom prostoru i norme je 1, postoje dva međusobno suprotna izbora za taj vektor. Njih možemo naći Gram-Schmidtovim postupkom ili pogađanjem. Opcije za taj stupac su $\pm(-0.5, 0.5, 0.7, 0.1)$. Kod potpune QR faktorizacije drugom faktoru na donji dio matrice dopisujemo nule. Zato su sve moguće potpune QR faktorizacije:

$$\begin{bmatrix} 0.7 & -0.1 & -0.5 & -0.5 \\ 0.1 & 0.7 & -0.5 & 0.5 \\ 0.5 & -0.5 & 0.1 & 0.7 \\ 0.5 & 0.5 & 0.7 & -0.1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & -2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0.7 & -0.1 & -0.5 & 0.5 \\ 0.1 & 0.7 & -0.5 & -0.5 \\ 0.5 & -0.5 & 0.1 & -0.7 \\ 0.5 & 0.5 & 0.7 & 0.1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & -2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

△

4.3 SVD faktorizacija

Teorem 4.7 (Singularna dekompozicija)**.** Neka je $A \in \mathbb{M}^{n\times m}$, $n \geq m$. Tada postoje unitarne matrice $U \in \mathbb{M}_n$ i $V \in \mathbb{M}_m$, te matrica $\Sigma \in \mathbb{M}^{n\times m}$ koja na mjestima (i,i)

($1 \leq i \leq \min\{n, m\}$) ima padajući niz nenegativnih realnih vrijednosti, a na ostalim mjestima nule, takve da je

$$A = U\Sigma V^T.$$

Vrijednosti $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min\{m,n\}} \geq 0$ na dijagonali matrice Σ zovemo singularnim vrijednostima.

Stupce matrice U ponekad nazivamo lijevim, a stupce matrice V desnim singularnim vektorima.

Primijetite da, za razliku od spektralne dekompozicije, ne postoji uvjet na matricu A da bi imala singularnu dekompoziciju: može biti pravokutna, singularna, a također dozvoljava singularnu dekompoziciju i kada je kompleksna. Singularne vrijednosti uvijek su nenegativni realni brojevi.

Singularne vrijednosti za matricu su jedinstvene, ali sama faktorizacija ne mora biti (pogledajte za primjer nul-matricu).

Propozicija 4.8. Dana je matrica $A \in \mathbb{M}^{n\times m}$. Tada su netrivijalne singularne vrijednosti matrice A upravo drugi korijeni netrivijalnih svojstvenih vrijednosti matrica AA^T i A^TA.

Propozicija 4.9. Neka je $A = U\Sigma V^T$ singularna dekompozicija matrice A. Neka je točno r njenih singularnih vrijednosti različito od nula. Neka su v_i i u_i oznake za i-te stupce matrica V i U. Tada vrijedi:

a) $r(A) = r$;

b) $\operatorname{Ker} A = [\{v_{r+1}, \ldots, v_m\}]$;

c) $\operatorname{Im} A = [\{u_1, \ldots, u_r\}]$;

d) $A = U_r\Sigma_r V_r$, gdje je Σ_r gornja lijeva podmatrica matrice Σ reda r, a U_r i V_r matrice sastavljene od prvih r stupaca matrica U i V;

e) $\|A\|_F^2 = \sigma_1^2 + \ldots + \sigma_r^2$;

f) $\|A\|_2 = \sigma_1$.

Teorem 4.10 (Ekhard, Young, Mirsky)**.** Neka je $A = U\Sigma V^T$ singularna dekompozicija matrice $A \in \mathbb{M}_{n\times m}$ ranga r. Neka je $1 \leq k < r$. Tada je

$$\min_{r(X)=k} \|A - X\|_2 = \sigma_{k+1}.$$

Minimum se ostvaruje za matricu $X = U_k\Sigma_k V_k$, gdje je Σ_k gornja lijeva podmatrica matrice Σ reda k, a U_k i V_k matrice sastavljene od prvih k stupaca matrica U i V.

SVD dekompoziciju možemo koristiti za rješavanje diskretnog problema najmanjih kvadrata. Naime, kako je

$$\|Ax - b\|_2^2 = \|\Sigma_r V_r^T x - U_r^T b\|_2^2 + \|(U_r^\perp)^T b\|_2$$

vrijedi

$$\min_{x\in\mathbb{R}^m} \|Ax-b\|_2 = \|(U_r^{\perp})^T b\|_2$$

te se postiže za x koji je rješenje sustava $\Sigma_r V_r^T x = U_r^T b$.

Zadatak 4.11. Dane su matrice

$$\tilde{U} = \begin{bmatrix} 0.6 & 0.48 & 0.64 & 0 \\ -0.8 & 0.36 & 0.48 & 0 \\ 0 & -0.48 & 0.36 & 0.8 \\ 0 & 0.64 & -0.48 & 0.6 \end{bmatrix}, \quad \tilde{\Sigma} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 100 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \tilde{V} = \begin{bmatrix} 0.6 & 0.8 & 0 \\ -0.48 & 0.36 & 0.8 \\ 0.64 & -0.48 & 0.6 \end{bmatrix},$$

te matrica $A = \tilde{U}\tilde{\Sigma}\tilde{V}^*$. Poznato je da su matrice $\tilde{U}$ i $\tilde{V}$ unitarne.

1. Je li $A = \tilde{U}\tilde{\Sigma}\tilde{V}^*$ singularna dekompozicija matrice A? Ako da, obrazložite; ako ne, odredite ju.
2. Odredite Frobeniusovu normu matrice, sliku matrice A.
3. Dokažite sljedeću tvrdnju: ako je matrica $X \in \mathbb{R}^{n\times m}$ ranga 2, tada postoje vektori $v_1, v_2 \in \mathbb{R}^n$ i vektori $w_1, w_2 \in \mathbb{R}^m$, svi različiti od nul-vektora, takvi da je $X = v_1 w_1^T + v_2 w_2^T$. Vrijedi li obrat te tvrdnje?

Rješenje. 1. Ovdje se ne radi o singularnoj dekompoziciji matrice jer dijagonalne vrijednosti u matrici $\tilde{\Sigma}$ nisu padajuće po vrijednosti. Da bismo dobili SVD, trebamo permutirati vrijednosti u toj matrici. Koristimo matrice permutacije P_3 i P_4 koje permutiraju elemente u prva dva retka/stupca:

$$P_3 := \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad P_4 := \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Tada imamo

$$A = \tilde{U}\tilde{\Sigma}\tilde{V}^* = \tilde{U}P_4P_4\tilde{\Sigma}P_3P_3\tilde{V}^* = (\tilde{U}P_4)(P_4\tilde{\Sigma}P_3)(\tilde{V}P_3)^* = U\Sigma V,$$

gdje je

$$U = \tilde{U}P_4 = \begin{bmatrix} 0.48 & 0.6 & 0.64 & 0 \\ 0.36 & -0.8 & 0.48 & 0 \\ -0.48 & 0 & 0.36 & 0.8 \\ 0.64 & 0 & -0.48 & 0.6 \end{bmatrix}, \quad \Sigma = P_4\tilde{\Sigma}P_3 = \begin{bmatrix} 100 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

$$V = \tilde{V}P_3 = \begin{bmatrix} 0.8 & 0.6 & 0 \\ 0.36 & -0.48 & 0.8 \\ -0.48 & 0.64 & 0.6 \end{bmatrix}.$$

To sada jest SVD matrice A.

2. Frobeniusova norma matrice A jednaka je Frobeniusovoj normi matrice Σ i iznosi $\sqrt{100^2 + 2^2 + 1^2} = \sqrt{10005}$. Slika je razapeta s prvih r stupaca matrice U. Rang je 3 (jer je to broj netrivijalnih dijagonalnih vrijednosti matrice Σ), pa je

$$\mathrm{Im}(A) = [\{(0.48, 0.36, -0.48, 0.64)^T, (0.6, -0.8, 0, 0)^T, (0.64, 0.48, 0.36, -0.48)^T\}].$$

3. Ako je matrica ranga 2, tada se njezin SVD može zapisati kao $X = U\Sigma V^* = U_2\Sigma_2 V_2^*$, gdje su U_2 i V_2 prva dva stupca matrica U i V, dok je Σ_2 prva 2×2 podmatrica matrice Σ, te ima dvije pozitivne dijagonalne vrijednosti σ_1 i σ_2. Posebno, ako s I označimo 2×2 jediničnu matricu, te e_1 i e_2 dva kanonska vektora u $\mathbb{R}^2$, možemo pisati

$$\begin{aligned} X &= U_2\Sigma_2 V_2^T = U_2\Sigma_2 I V_2^T = U_2\Sigma_2(e_1e_1^T + e_2e_2^T)V_2^T \\ &= U_2\Sigma_2 e_1e_1^T V_2^T + U_2\Sigma_2 e_2e_2^T V_2^T = u_1\sigma_1 v_1^T + u_2\sigma_2 v_2^T = (u_1\sigma_1)v_1^T + (u_2\sigma_2)v_2^T, \end{aligned}$$

čime je prva tvrdnja dokazana. Obrat ne vrijedi: uzmimo da su vektori v_1,v_2 jednaki prvom vektoru kanonske baze za $\mathbb{R}^n$, a w_1,w_2 jednaki prvom vektoru kanonske baze za $\mathbb{R}^m$. Tada matrica X ima sve vrijednosti jednake nuli osim na mjestu $(1, 1)$ gdje se nalazi broj 2. Ta matrica očito je ranga 1.

△

4.4 Neprekidni najmanji kvadrati

Promatramo isti problem: za zadanu $f : X \to \mathbb{R}$ naći što bolju aproksimacijsku funkciju φ među svim $\psi \in \mathcal{S}$ takvu da

$$\|f - \varphi\| = \inf_{\psi \in \mathcal{S}} \|f - \psi\|.$$

Ovdje imamo sljedeće:

- $X = [a, b]$;
- $\mathcal{S}$ je npr. $\mathcal{P}_m$ (konačnodimenzionalni potprostor prostora $\mathcal{C}([a, b]; \mathbb{R})$);
- L^2-norma za neprekidnu $g : [a, b] \to \mathbb{R}$

$$\|g\|_{L^2([a,b];\mathbb{R})} := \left(\int_a^b g(t)^2 dt\right)^{1/2}.$$

- Kompatibilni skalarni produkt

$$\langle g, h\rangle_{L^2([a,b];\mathbb{R})} := \int_b^a g(t)h(t)dt.$$

Ako je $\{\varphi_1, \ldots, \varphi_m\}$ baza za $\mathcal{S}$, cilj je naći funkciju

$$\varphi(x) = \sum_{i=1}^{m} x_i \varphi_i$$

takvu da

$$\|f - \varphi\|_{L^2} = \left\| f - \sum_{i=1}^{m} x_i \varphi_i \right\|_{L^2}$$

što manja. Rješavanje ovog problema se opet svodi na rješavanje sustava normalnih jednadžbi $Mx = p$ gdje je $M \in \mathbb{M}_{m\times m}$ matrica takva da joj je na mjestu (i, j) $\langle \varphi_i, \varphi_j \rangle_{L^2}$, a $p \in \mathbb{R}^m$ vektor sa komponentama $p_i = \langle \varphi_i, f \rangle_{L^2}$.

Kada bi baza $\{\varphi_1, \ldots, \varphi_m\}$ bila ortogonalna, vandijagonalni elementi matrice M $\langle \varphi_i, \varphi_j \rangle$ bi bili jednaki 0, pa bi se matrica sustava svela na dijagonalnu. Jednu takvu bazu čine Legendreovi polinomi koji su definirani sljedećemo rekurzijom:

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= x \\ n \geq 1: \quad (n+1)P_{n+1}(x) &= (2n+1)xP_n(x) - nP_{n-1}(x) \end{aligned}$$

Niz Legendreovih polinoma je ortogonalan na $[-1, 1]$. Također vrijedi

$$\langle P_i, P_i, \rangle_{L^2} = \int_{-1}^{1} (P_k(x))^2 dx = \frac{2}{2k+1}.$$

U slučaju Legendreove baze na $[-1, 1]$ imamo

$$x_i = \frac{\langle P_i, f \rangle_{L^2}}{\langle P_i, P_i, \rangle_{L^2}}.$$

Zadatak 4.12. Koristeći Legendreovu bazu, aproksimiraj $\dfrac{1}{x^2+1}$ polinomom iz $\mathcal{P}_3$ na $[-1, 1]$.

Rješenje. Promatramo bazu za $\mathcal{P}_3$ $\{P_0, P_1, P_2, P_3\} = \left\{1, x, \frac{1}{2}(3x^2 - 1), \frac{1}{2}(5x^3 - 3x)\right\}$.

Aproksimacijska funkcija je oblika $\varphi(x) = \sum_{k=0}^{3} x_k P_k$, gdje su

$$x_k = \frac{\int_{-1}^{1} P_k(x) f(x) dx}{\int_{-1}^{1} (P_k(x))^2 dx} = \frac{2k+1}{2} I_k, \quad k = 0, 1, 2, 3.$$

Računamo redom:

$$I_0 = \int_{-1}^{1} \frac{dx}{x^2+1} = \pi/2$$
$$I_1 = \int_{-1}^{1} \frac{x}{x^2+1} dx = [\textit{neparna funkcija na simetričnom intervalu}] = 0,$$
$$I_2 = \frac{1}{2}\int_{-1}^{1} \frac{3x^2-1}{x^2+1} dx = 3-\pi,$$
$$I_3 = \frac{1}{2}\int_{-1}^{1} \frac{5x^3-3x}{x^2+1} dx = [\textit{neparna funkcija na simetričnom intervalu}] = 0.$$

Dakle,

$$\varphi(x) = \sum_{k=0}^{3} x_k P_k = \frac{1}{2}\cdot\frac{\pi}{2} + \frac{5}{2}(3-\pi)\frac{3x^2-1}{2}.$$

△

4.4.1 Trigonometrijske funkcije

Kako trigonometrijske funkcije

$$\{1, \cos x, \cos 2x, \ldots, \cos Nx, \sin x, \sin 2x, \ldots, \sin Nx\}$$

čine ortogonalnu familiju funkcija na $[-\pi, \pi]$, vrijedi da je funkcija

$$\varphi(x) = \frac{a_0}{2} + \sum_{k=1}^{N} (a_k \cos(kx) + b_k \sin(kx)),$$

gdje su

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(kx)dx, \quad b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(kx)dx$$

rješenje neprekidnog problema najmanjih kvadrata za funkciju $f : [-\pi, \pi] \to \mathbb{R}$ na skupu

$$\mathcal{S} = [\{1, \cos x, \cos 2x, \ldots, \cos Nx, \sin x, \sin 2x, \ldots, \sin Nx\}].$$

Kada $N \to \infty$ red

$$S_f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx)$$

nazivamo Fourierovim redom za funkciju f.

Teorem 4.13. Neka je funkcija $f : [-\pi, \pi] \to \mathbb{R}$ takva da su f i f' po dijelovima neprekidne funkcije. Tada za svaki $x \in [-\pi, \pi]$ Fourierov red $S_f(x)$ funkcije f konvergira, te vrijedi

$$S_f(x) = \frac{f(x^+) + f(x^-)}{2},\ x \in \langle -\pi, \pi\rangle, \quad S_f(x) = \frac{f(-\pi^+) + f(\pi^-)}{2},\ x = -\pi, \pi.$$

Posebno, u unutarnjim točkama domene u kojima f nema prekid Fourierov red ima vrijednost funkcije f .

Zadatak 4.14. Odredi Fourierov red funkcije $f : [-\pi, \pi]$, $f(x) = x^2$. Koristeći dobiveni red, dokaži identitet $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$.

Rješenje. Računamo koeficijente a_k, b_k. Primjetimo da je $b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} x^2 \sin(kx)dx = 0$ za svaki $k \in \mathbb{N}_0$ jer integriramo neparnu funkciju na simetričnom intervalu. S druge strane:

$$\begin{aligned}
a_0 &= \frac{1}{\pi}\int_{-\pi}^{\pi} x^2 dx = \frac{2\pi^2}{3},\\
a_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} x^2\cos(kx)dx = \begin{bmatrix} u = x^2 & du = 2xdx \\ dv = \cos(kx)dx & v = \frac{1}{k}\sin(kx)\end{bmatrix}\\
&= \underbrace{\frac{1}{k\pi}x^2\sin(kx)|_{-\pi}^{\pi}}_{=0} - \frac{2}{k\pi}\int_{-\pi}^{\pi} x\sin(kx)dx = \begin{bmatrix} u = x & du = dx \\ dv = \sin(kx)dx & v = -\frac{1}{k}\cos(kx)\end{bmatrix}\\
&= \frac{2}{k^2\pi}x\cos(kx)|_{-\pi}^{\pi} + \frac{2}{k^2\pi}\int_{-\pi}^{\pi}\cos(kx)dx\\
&= \frac{2}{k^2\pi}(\pi\cos(k\pi) - \pi\cos(-k\pi)) + 0\\
&= \frac{4}{k^2}\cos(k\pi) = \frac{4}{k^2}(-1)^k
\end{aligned}$$

Sada Fourierov razvoj za $f(x)$ ima oblik

$$f(x) = \frac{\pi^2}{3} + \sum_{k=1}^{\infty}\frac{4(-1)^k}{k^2}\cos(kx).$$

Da bi dokazali traženu jednakost, uvrstimo $x = \pi$:

$$\pi^2 = f(\pi) = \frac{\pi}{3} + \sum^{\infty}\frac{4(-1)^k}{k^2}\cos(k\pi) = \frac{\pi}{3} + \sum_{k=1}^{\infty}\frac{4}{k^2}.$$

Podijelimo izraz s 4 i imamo

$$\sum_{k=1}^{\infty}\frac{1}{k^2} = \frac{\pi^2}{4} - \frac{\pi^2}{12} = \frac{\pi^2}{6}.$$

△

5

Numerička integracija i derivacija

5.1 Numeričko deriviranje

Za zadanu dovoljno glatku funkciju $f : \langle a, b\rangle \to \mathbb{R}$ i točku $x_0 \in \langle a, b\rangle$ želimo odrediti aproksimaciju prve derivacije $f'(x_0)$ ili općenito k-te derivacije $f^{(k)}(x_0), k \in \mathbb{N}$.

Ideja je funkciju aproksimirati interpolacijskim polinom, a derivaciju funkcije derivacijom odgovarajućeg interpolacijskog polinoma.

Ako funkciju f aprosksimiramo interpolacijskim polinomom stupnja 1, kroz točke $a < x_0 < x_1 < b$ dobivamo sljedeće aproksimacije

- **Podijeljena razlika unaprijed**
$$f'(x_0) \approx \frac{f(x_1) - f(x_0)}{x_1 - x_0}.$$
- **Podijeljena razlika unazad**
$$f'(x_1) \approx \frac{f(x_1) - f(x_0)}{x_1 - x_0}.$$
- **Središnja podijeljena razlika**
$$f'(x') \approx \frac{f(x_1) - f(x_0)}{x_1 - x_0}, \quad x' \in \langle x_0, x_1\rangle.$$

Ocjene greške (vidi predavanja) izvode se pomoću tvrdnje koja kaže: ako je f derivabilna $(n + 2)$ puta, tada vrijedi

$$|e'(x_i)| = |f'(x_i) - p'_n(x_i)| \leq \frac{|\omega'(x_i)|}{(n + 1)!} M_{n+1}.$$

Zadatak 5.1. Izvedi ocjenu pogreške za formulu

$$f'(x_{-1}) \approx \frac{-3f(x_{-1}) + 4f(x_0) - f(x_1)}{2h},$$

gdje je f dovoljno glatka funkcija, te su točke x_{-1}, x_0 i x_1 ekvidistantne ($x_1 - x_0 = x_0 - x_{-1} = h$.)

Rješenje. Ta formula izvedena je kao vrijednost derivacije interpolacijskog polinoma p_2 u x_{-1} za funkciju f kroz točke x_{-1}, x_0, x_1.

Pretpostavimo li da koristimo formulu iz zadatka za četiri puta derivabilnu funkciju ($n = 2$), tada je za ocjenu potrebno samo odrediti $|\omega'(x_{-1})|$.

$$\omega(x) = (x - x_{-1})(x - x_0)(x - x_1) = (x - x_0)^3 - (x - x_0)h^2$$
$$\implies \omega'(x) = 3(x - x_0)^2 - h^2 \implies \omega'(x_{-1}) = 2h^2.$$

Dakle, dobili smo da za svaku 4 puta derivabilnu funkciju f vrijedi

$$\left| f'(x_{-1}) - \frac{-3f(x_{-1}) + 4f(x_0) - f(x_1)}{2h} \right| = |f'(x_i) - p_n'(x_i)| \leq \frac{|\omega'(x_i)|}{3!} M_3 = \frac{h^2}{3} M_3.$$

△

Ocjene možemo izvoditi i preko Taylorovog polinoma. Pokažimo to za formulu za aproksimaciju druge derivacije.

Zadatak 5.2. Izvedi ocjenu pogreške za formulu

$$f''(x_0) \approx \frac{f(x_{-1}) - 2f(x_0) + f(x_1)}{h^2},$$

gdje je f dovoljno glatka funkcija, te su točke x_{-1}, x_0 i x_1 ekvidistantne ($x_1 - x_0 = x_0 - x_{-1} = h$.)

Rješenje. Koristeći Taylorov razvoj za funkciju f, dobivamo

$$f(x_{-1}) = f(x_0 - h) = f(x_0) - hf'(x_0) + \frac{h^2}{2} f''(x_0) - \frac{h^3}{6} f'''(x_0) + \frac{h^4}{24} f^{(4)}(\xi_1)$$
$$f(x_1) = f(x_0 + h) = f(x_0) + hf'(x_0) + \frac{h^2}{2} f''(x_0) + \frac{h^3}{6} f'''(x_0) + \frac{h^4}{24} f^{(4)}(\xi_2).$$

Zbrajajući te dvije jednadžbe zajedno s trivijalnom jednadžbom $-2f(x_0) = -2f(x_0)$, dobivamo

$$\begin{aligned} f(x_{-1}) - 2f(x_0) + f(x_1) &= [f(x_0) - 2f(x_0) + f(x_0)] \\ &\quad + [-hf'(x_0) + hf'(x_0)] \\ &\quad + \left[\frac{h^2}{2} f''(x_0) + \frac{h^2}{2} f''(x_0) \right] \\ &\quad + \left[\frac{h^3}{6} f'''(x_0) - \frac{h^3}{6} f'''(x_0) \right] \\ &\quad + \left[\frac{h^4}{24} f^{(4)}(\xi_1) + \frac{h^4}{24} f^{(4)}(\xi_2) \right] \\ &= h^2 f''(x_0) + \frac{h^4}{24} \left(f^{(4)}(\xi_1) + f^{(4)}(\xi_2) \right). \end{aligned}$$

$$\implies \left| f''(x_0) - \frac{f(x_{-1}) - 2f(x_0) + f(x_1)}{h^2} \right| = \left| \frac{h^2}{24} \left(f^{(4)}(\xi_1) + f^{(4)}(\xi_2) \right) \right| \leq \frac{h^2}{12} M_4.$$

Gornji raspis vrijedi za svaku funkciju f koja je četiri puta derivabilna. △

Zadatak 5.3 (DZ)**.** Prilagođavajući pristup iz prošlog zadatka, nađite realne parametre A, B, C takve da formula

$$f'(x_1) \approx \frac{Af(x_{-1}) + Bf(x_0) + Cf(x_1)}{h}$$

ima što bolju ocjenu greške, te nađite tu ocjenu greške.

5.2 Numeričko integriranje

Aproksimirajući integral funkcije integralom njezinog interpolacijskog polinoma na ekvidistantnoj mreži, na predavanjima smo izveli sljedeće formule (s odgovarajućim ocjenama):

- **Formula srednje točke**:

$$\int_a^b f(x)\ dx \approx I_0(f) := hf\left(\frac{a+b}{2}\right), \quad \left| \int_a^b f(x)dx - I_0(f) \right| \leq \frac{(b-a)^3}{24} M_2;$$

- **Trapezna formula**:

$$\int_a^b f(x)\ dx \approx I_1(f) := \frac{h}{2}\left(f(a) + f(b)\right), \quad \left| \int_a^b f(x)dx - I_1(f) \right| \leq \frac{(b-a)^3}{12} M_2;$$

- **Simpsonova formula**:

$$\int_a^b f(x)\ dx \approx I_2(f) := \frac{h}{6}\left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right), \quad \left| \int_a^b f(x)dx - I_2(f) \right| \leq \frac{(b-a)^5}{2880} M_4,$$

gdje je $M_n := \max_{x\in[a,b]} |f^{(n)}(x)|$.

Kako bismo povećali točnost formula, u praksi koristimo produljene verzije ovih formula (originalan interval podijelimo na manje, te na svakom intervalu primijenimo odgovarajuću formulu). Na predavanjima smo radili sljedeće formule:

- **Produljena trapezna formula:** Ako je za neki $n \in \mathbb{N}$ $a = x_0 < x_1 < \cdots < x_n = b$ ekvidistantna mreža ($h := x_i - x_{i-1}$), tada primjenom trapezne formule na svakom intervalu $[x_{i-1}, x_i]$ dobivamo

$$\int_a^b f(x)\ dx \approx I_n^{PT}(f) := \frac{h}{2} f(a) + h \sum_{i=1}^{n-1} f(x_i) + \frac{h}{2} f(b),$$

ocjena greške:

$$\left| \int_a^b f(x)dx - I_n^{PT}(f) \right| \leq \frac{(b-a)^3}{12n^2} M_2 = \frac{(b-a)h^2}{12} M_2.$$

- **Produljena Simpsonova formula:** Ako je za neki $n \in \mathbb{N}$ $a = x_0 < x_1 < \cdots < x_{2n} = b$ ekvidistantna mreža ($h := x_i - x_{i-1}$), tada primjenom trapezne formule na svakom intervalu $[x_{2i-2}, x_{2i}]$ duljine $2h$ nakon sređivanja dobivamo

$$\int_a^b f(x)\ dx \approx I_n^{PS}(f) := \frac{h}{3} f(a) + \frac{4h}{3} \sum_{i=1}^{n} f(x_{2i-1}) + \frac{2h}{3} \sum_{i=1}^{n-1} f(x_{2i}) + \frac{h}{3} f(b),$$

ocjena greške:

$$\left| \int_a^b f(x)dx - I_n^{PS}(f) \right| \leq \frac{(b-a)^5}{2880n^4} M_4 = \frac{(b-a)h^4}{180} M_4.$$

Primjer 5.4. Koji je najmanji broj čvorova potreban da bi ocjena greške prilikom numeričke integracije bila manja od zadanog $\varepsilon > 0$ ako koristimo:

1. produljenu trapeznu formulu,
2. produljenu Simpsonovu formulu.

Rješenje. 1. Želimo:

$$R_{PT} \leq \frac{(b-a)h^2}{12} M_2 f \leq \varepsilon.$$

Uvrstimo da je $h = \frac{b-a}{n}$ $\implies$

$$\frac{(b-a)^3}{12n^2} M_2 f \leq \varepsilon.$$

Rješavanjem po n dobijemo:

$$n \geq \sqrt{\frac{(b-a)^3}{12\varepsilon} M_2 f}.$$

2. Sasvim analogno za produljenu Simpsonovu formulu dobijemo:

$$2n \geq \sqrt[4]{\frac{(b-a)^5}{180\varepsilon} M_4 f}.$$

Kod PS formule, očito biramo paran broj.

△

Zadatak 5.5. Izračunajte vrijednost integrala

$$I = \int_0^1 \frac{dx}{1+x}$$

a) produljenom trapeznom formulom,

b) produljenom Simpsonovom formulom

uz $h = 0.1$. Ocijenite pogrešku i nađite pravu grešku. Nađite najmanji h tako da pogreška integracije ne prelazi $\varepsilon = 10^{-4}$.

Rješenje. Kako je $h = \frac{b-a}{n} = \frac{1}{n} = 0.1$ zaključujemo da je $n = 10$. Napravimo tablicu čvorova i vrijednosti funkcije $f(x) = \frac{1}{1+x}$:

i	x_i	$f(x_i)$
0	0	1
1	0.1	0.90909
2	0.2	0.83333
3	0.3	0.76923
4	0.4	0.71429
5	0.5	0.66667
6	0.6	0.625
7	0.7	0.58824
8	0.8	0.55556
9	0.9	0.52632
10	1	0.5

a)

$$I_{PT} = \frac{0.1}{2}\left(f(x_0) + 2\left(f(x_1) + f(x_2) + \ldots + f(x_9)\right) + f(x_{10})\right) = 0.693773.$$

Za pogrešku nam je potreban maksimum apsolutne vrijednosti druge derivacije:

$$f''(x) = \frac{2}{(1+x)^3}.$$

To je strogo padajuća funkcija pa je

$$M_2 f = f''(0) = 2,$$

pa za grešku integracije vrijedi:

$$R_{PT} \leq 0.001667.$$

Kako je

$$\int_0^1 \frac{dx}{1+x} = \ln(1+x)\Big|_0^1 = \ln 2 = 0.69315,$$

prava pogreška iznosi

$$R_{PT} = |0.693773 - 0.69315| = 0.00062.$$

Broj intervala n da bi pogreška interpolacije bila manja od 10^{-4} je

$$n \leq \sqrt{\frac{(b-a)^3}{12\varepsilon} M_2 f} = \sqrt{\frac{1}{12 \cdot 10^{-4}} \cdot 2} = 40.82$$

pa možemo uzeti $n = 41$.

b)

$$I_{PS} = \frac{0.1}{3} \left(f(x_0) + 4 \left(f(x_1) + f(x_3) + f(x_5) + f(x_7) + f(x_9) \right) \right.$$
$$\left. + 2 \left(f(x_2) + f(x_4) + f(x_6) + f(x_8) \right) + f(x_{10}) \right) = 0.69315.$$

Prava pogreška je

$$R_{PS} = 3.0505 \cdot 10^{-6}$$

Za ocjenu pogreške nam treba maksimum apsolutne vrijednosti četvrte derivacije:

$$f^{IV}(x) = \frac{24}{(1+x)^5}.$$

Ovo je strogo padajuća funkcija pa je $M_4 f = f^{(IV)}(0) = 24$, pa je ocjena pogreške

$$R_{PS} \leq 1.3333 \cdot 10^{-5}.$$

Broj čvorova $2n$ tako da ocjena pogreške bude manja od $\varepsilon = 10^{-4}$ je

$$2n \geq 6.042,$$

pa uzimamo $2n = 8$, odnosno $n = 4$.

△

Za sve prošle formule, fiksirali smo čvorove integracije. U slučaju da ih biramo tako da polinomijalni stupanj egzaktnosti tih formula bude što viši, ispada da ih treba birati tako da budu nultočke Lagrangeovih polinoma (na $[-1, 1]$), odnosno da budu afine transformacije tih nultočaka (na $[a, b]$). Time dobivamo formule:

- $n = 1$: midpoint formula;
- $n = 2 : \int_{-1}^{1} f(x)dx \approx I_1^{GL}(f) = f\left(-\frac{\sqrt{3}}{3}\right) + f\left(\frac{\sqrt{3}}{3}\right)$;
- $n = 3 : \int_{-1}^{1} f(x)dx \approx I_2^{GL}(f) = \frac{5}{9} f\left(-\sqrt{\frac{3}{5}}\right) + \frac{8}{9} f(0) + \frac{5}{9} f\left(\sqrt{\frac{3}{5}}\right)$.

Općenito za produljene i pomaknute G-L formule:

Teorem 5.6. Neka su dani $n, m \in \mathbb{N}$, te neka su $w_1, \ldots, w_n$ i $x_1, \ldots, x_n$ težine i čvorovi Gauss-Legendreove integracijske formule $I_n^{GL}(f)$ na intervalu $[-1, 1]$. Za funkciju $f : [a, b] \to \mathbb{R}$ promatramo produljenu Gauss-Legendreovu integracijsku formulu primijenjenu na svaki od m podintervala intervala $[a, b]$duljine $h = \frac{b-a}{m}$

$$I_{m,n}^{PGL}(f) = \sum_{i=1}^{m} \sum_{k=1}^{n} \frac{h}{2} w_k f\left(\frac{h}{2} x_k + a + \frac{2i-1}{2} h\right).$$

$f \in C^{2n}([a, b])$, tada je zadovoljena ocjena

$$\left| \int_a^b f(x)dx - I_{m,n}^{PGL}(f) \right| \leq \frac{(n!)^4}{(2n+1)[(2n)!]^3} \frac{(b-a)^{2n+1}}{m^{2n}} M_{2n}.$$

Pomaknute formule dobili smo tako da smo afino pomaknuli čvorove, dok smo težine pomnožili omjerom duljine novih intervala (h) i intervala iz G-L formula (2).

Zadatak 5.7. Proizvoljnom metodom aproksimirajte $\int_{-1}^{1} e^{x/2} dx$ s točnošću $\varepsilon = 10^{-5}$. Izračunajte pravu pogrešku tom metodom.

Rješenje. Izabrat ćemo Gauss-Legendreovu formulu za $n = 3$. Naime, u slučaju da koristimo produljenu formulu s m podintervala, taj m treba zadovoljavati nejednakost

$$\frac{(n!)^4}{(2n+1)[(2n)!]^3} \frac{(b-a)^{2n+1}}{m^{2n}} M_{2n} < 10^{-5}.$$

Uzimajući u obzir da je $n = 3$, te $M_{2n} = 2^{-2n} e^{1/2}$ (sve derivacije funkcije $f(x) = e^{x/2}$ su pozitivne, pa su onda i rastuće), imamo

$$\begin{aligned} \frac{(3!)^4}{(2 \cdot 3 + 1)[(2 \cdot 3)!]^3} \frac{2^{2 \cdot 3 + 1}}{m^{2 \cdot 3}} M_6 &= \frac{6^4}{7 \cdot 720^3} \frac{2^7}{m^6} 2^{-6} e^{1/2} \\ &= \frac{6}{7 \cdot 120^3} \frac{2}{m^6} e^{1/2} \\ &= \frac{1}{70 \cdot 120^2 \cdot m^6} e^{1/2} \\ &= \frac{e^{1/2}}{1008000} m^{-6} = 1.63563618125 \cdot 10^{-6} \cdot m^{-6}, \end{aligned}$$

što je već za $m = 1$ manje od 10^{-5}. Zato je

$$\begin{aligned} I &:= \int_{-1}^{1} e^{x/2} dx = 2e^{1/2} - 2e^{-1/2} = 2.08438122197, \\ I_0 &= \frac{5}{9} f\left(-\sqrt{\frac{3}{5}}\right) + \frac{8}{9} f(0) + \frac{5}{9} f\left(\sqrt{\frac{3}{5}}\right) = 2.08438022231, \\ |I - I_0| &= 9.9966748524 \cdot 10^{-7}. \end{aligned}$$

Za usporedbu, za produljenu trapeznu formulu imali bismo $n \geq 166$, za produljenu Simpsonovu $n \geq 4$, za produljenu formulu srednje točke (Gauss-Legendreovu formulu za $n = 1$) $m \geq 118$, a za produljenu Gauss-Legendreovu formulu s $n = 2$ imali bismo $m \geq 3$. △

Zadatak 5.8. Za neki $h > 0$ dana je integracijska formula

$$\int_0^h f(x)dx \approx I(f) = w_1 f(0) + w_2 f\left(\frac{h}{3}\right),$$

s težinama w_1 i w_2 takvima da je formula što veće polinomijalnog stupnja egzaktnosti.

1. Odredite težine w_1 i w_2 te polinomijalni stupanj egzaktnosti ove formule.
2. Nađite ocjenu pogreške za ovu formulu.
3. Kako bi glasila ova integracijska formula na proizvoljnom intervalu $[a, b]$?
4. Kako bi glasila ova produljena integracijska formula ekvidistantnoj podijeli intervala $[0, 1]$?
5. Koristeći produljenu verziju ove integracijske formule s podijelom na dva intervala odredite aproksimaciju integrala $\int_0^1 e^{x^2} dx$. Odredite ocjenu pogreške u tom slučaju.

Rješenje. Za odrediti w_1 i w_2 sa što bolji stupnjem egzaktnosti potrebno je postići da je ova formula egzaktna na polinomima što višeg stupnja. Uvrštavajući $f(x) = 1$ i $f(x) = x$, dobivamo

$$h = \int_0^h 1dx = I(1) = w_1 + w_2,$$
$$\frac{h^2}{2} = \int_0^h xdx = I(x) = w_1 \cdot 0 + w_2 \cdot \frac{h}{3},$$

odakle dobivamo $w_2 = \frac{3}{2}h$ i $w_1 = -\frac{1}{2}h$. Uvrštavajući $f(x) = x^2$ vidimo da formula ne integrira egzaktno tu funkciju, pa je polinomijalni stupanj egzaktnosti jednak 1.

Po Teoremu 5.12 s predavanja, budući da $I(f)$ ima polinomijalni stupanj egzaktnosti barem 1 (što je broj čvorova umanjen za 1) ova formula izvedena je kao integral interpolacijskog polinoma stupnja 1 u čvorovima 0 i $\frac{h}{3}$. Zato za ocjenu greške za dvaput diferencijabilnu funkciju f imamo

$$|f(x) - p_1(x)| = \left|\frac{\omega(x)}{2!} f''(\xi_x)\right| \leq \frac{1}{2} M_2 |\omega(x)|\, dx,$$

$$\begin{aligned}\left|\int_0^h f(x)dx - I(f)\right| &= \left|\int_0^h f(x)dx - \int_0^h p_1(x)dx\right| \\ &\leq \int_0^h |f(x) - p_1(x)|\, dx \\ &\leq \frac{1}{2}M_2 \int_0^h |\omega(x)|.\end{aligned}$$

Za dovršetak, trebamo odrediti

$$\int_0^h |\omega(x)| = \int_0^h \left|x\left(x - \frac{h}{3}\right)\right| dx = \int_0^{\frac{h}{3}} x\left(\frac{h}{3} - x\right) dx + \int_{\frac{h}{3}}^h x\left(x - \frac{h}{3}\right) dx = ... = \frac{29}{162}h^3.$$

Zato je

$$\left|\int_0^h f(x)dx - I(f)\right| \leq \frac{29}{324}h^3 M_2.$$

Pokažimo još jedan način kako riješiti a) i b) dio zadatka, pomoću Taylorovog teorema. Za neku dvaput diferencijabilnu funkciju f definirajmo

$$F(x) := \int_0^x f(t)dt.$$

Trivijalno vidimo: $F(0) = 0$, $F(h) = \int_0^h f(x)dx$, $F'(x) = f(x)$. Iz zadnjeg svojstva F je triput diferencijabilna, pa ima Taylorov polinom stupnja 2 (s ostatkom stupnja 3). Razvijmo $F(h)$ i $f\left(\frac{h}{3}\right)$ u Taylorove polinome oko 0:

$$\begin{aligned}F(h) &= F(0) + hF'(0) + \frac{h^2}{2}F''(0) + \frac{h^3}{6}F'''(\xi_1) \\ &= hf(0) + \frac{h^2}{2}f'(0) + \frac{h^3}{6}f''(\xi_1), \\ f\left(\frac{h}{3}\right) &= f(0) + \frac{h}{3}f'(0) + \frac{h^2}{18}f''(\xi_2),\end{aligned}$$

za neke $\xi_1, \xi_2 \in \langle 0, h\rangle$. Greška koju činimo računajući integral integracijskom formulom apsolutna je vrijednost izraza

$$\begin{aligned}&\int_0^h f(x)dx - w_1 f(0) - w_2 f\left(\frac{h}{3}\right) = F(h) - w_1 f(0) - w_2 f\left(\frac{h}{3}\right) \\ &\quad = f(0)\left(h - w_1 - w_2\right) + f'(0)\left(\frac{h^2}{2} - w_2\frac{h}{3}\right) + \frac{h^3}{6}f''(\xi_1) - w_2\frac{h^2}{18}f''(\xi_2).\end{aligned}$$

Da bismo postigli što viši polinomijalni stupanj egzaktnosti, greška koja ova integracijska formula daje mora biti jednaka nuli za polinome što višeg stupnja. Za polinome prvog stupnja vrijedi $f'' = 0$, dok brojevi $f(0)$ i $f'(0)$ mogu biti proizvoljni. Zato koeficijenti uz $f(0)$ i $f'(0)$ moraju biti jednaki nuli. To su upravo one jednadžbe koje smo dobili i u

rješavanju a) zadatka prvim načinom pa ponovno dobivamo $w_2 = \frac{3}{2}h$ i $w_1 = -\frac{1}{2}h$. Za ocjenu greške sada imamo

$$\left|\int_0^h f(x)dx - w_1 f(0) - w_2 f\left(\frac{h}{3}\right)\right| = \left|\frac{h^3}{6}f''(\xi_1) - w_2\frac{h^2}{18}f''(\xi_2)\right|$$
$$\leq M_2\left|\frac{h^3}{6} + \frac{3h}{2}\cdot\frac{h^2}{18}\right| = \frac{1}{4}h^3 M_2.$$

Dobili smo drugačiju, ali i dalje točnu ocjenu greške (iako nešto lošiju).

Za c) dio zadatka, potrebno je ovu formulu napisati na proizvoljnom intervalu $[a, b]$. Za to je potrebno čvorove preslikati afino, a težine pomnožiti omjerom duljina intervala $(b-a)/h$:

$$\int_a^b f(x)dx \approx I^{[a,b]}(f) = \frac{b-a}{h}w_1 f(a) + \frac{b-a}{h}w_2 f\left(a + \frac{b-a}{h}\frac{h}{3}\right)$$
$$= \frac{a-b}{2}f(a) + \frac{3(b-a)}{2}f\left(\frac{2a+b}{3}\right).$$

Podijeljena formula dobije se zbrajanjem svih gornjih formula na intervalima oblika $[(i-1)/n, i/n]$ $(i = 1, \ldots, n)$, za neki $n \in \mathbb{N}$:

$$\int_0^1 f(x)dx \approx I_n^{[0,1]}(f) = \sum_{i=1}^n I^{[(i-1)/n,i/n]}(f) = \sum_{i=1}^n \left[\frac{-1}{2n}f\left(\frac{i-1}{n}\right) + \frac{3}{2n}f\left(\frac{3i-2}{3n}\right)\right]$$

Za funkciju $f(x) = e^{x^2}$ primjena ove formule glasi

$$\int_0^1 f(x)dx \approx I_2^{[0,1]}(f) = \frac{-1}{4}f(0) + \frac{3}{4}f\left(\frac{1}{6}\right) + \frac{-1}{4}f\left(\frac{1}{2}\right) + \frac{3}{4}f\left(\frac{4}{6}\right) = 1.36983665211.$$

Derivacije funkcije f su $f'(x) = e^{x^2}\cdot 2x$, $f''(x) = e^{x^2}\cdot(4x^2+2)$. Druga derivacija je pozitivna i rastuća kao umnožak takvih funkcija, pa je $M_2 = f''(1) = 6e$. Ocjenu pogreške produljene metode izvodimo tako da za svaki od 2 intervala duljine $h = 1/2$ primijenimo ocjenu pogreške koju smo izveli za proizvoljan integral duljine h:

$$\left|\int_0^1 f(x)dx - I_2^{[0,1]}(f)\right| \leq 2\cdot\frac{29}{324}\cdot\left(\frac{1}{2}\right)^3 M_2 = 0.364954504747,$$

(odnosno, ako koristimo ocjenu dobivenu Taylorovim polinomima: 1.01935568567).

Budući da je prava vrijednost integrala jednaka 1.46265174591, ove ocjene su prilično pesimistične. △

6

Nelinearne jednadžbe i optimizacija

6.1 Metoda bisekcije

Početne pretpostavke na f:

- $f : [a, b] \to \mathbb{R}$ neprekidna
- $f(a)f(b) < 0$ (f ima barem jednu nultočku na $[a, b]$).

Algoritam:
U svakom koraku konstruiramo interval $[a_n, b_n]$ čija je duljina jednaka polovini duljine prethodnog intervala tako da se u njemu nalazi nultočka, tj. $f(a_n)f(b_n) \leq 0$.

- **POČETAK:** duljina intervala je $d = b - a$.
- **ITERACIJA:** sve dok je $d > \varepsilon$ (ε je zadana točnost) ponavljamo:
  - $d = \frac{d}{2}$
  - $x_n = a_n + d$
  - Ako je $f(a_n)f(x_n) < 0$ onda je $a_{n+1} = a_n$, $b_{n+1} = x_n$,
    Ako je $f(a_n)f(x_n) > 0$ onda je $a_{n+1} = x_n$, $b_{n+1} = b_n$,
    Ako je $f(a_n)f(x_n) = 0$ onda je $d = 0$ (izlazimo iz petlje)
  - $n = n + 1$.
- **IZLAZ:** $y = x_{n-1}$ je tražena nultočka sa točnošću ε.

Ocjena greške:
Neka je ξ nultočka funkcije $f \in C([a, b])$, tj. $f(\xi) = 0$

$$|\xi - x_n| \leq \frac{1}{2^{n+1}}(b - a)$$

Ako je, osim toga, $f \in C^1([a,b])$

$$|\xi - x_n| \leq \frac{|f(x_n)|}{m_1}, \quad m_1 = \min_{x\in[a,b]} |f'(x)| > 0$$

Kriterij zaustavljanja iteracija uz zadanu točnost ε:

- $f \in C([a,b])$:
$$n \geq \frac{\log\left(\frac{b-a}{\varepsilon}\right)}{\log 2} - 1$$
- $f \in C^1([a,b])$:
$$|f(x_n)| \leq \varepsilon m_1$$

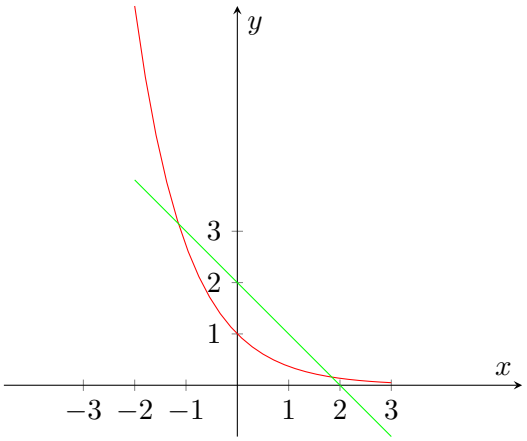
Iteracije zaustavljamo čim je jedan od kriterija zadovoljen.
Ima sigurnu konvergenciju, ali je spora! Nultočke parnog reda ne možemo naći ovom metodom.

Zadatak 6.1. Riješite, metodom raspolavljanja, jednadžbu

$$e^{-x} - 2 + x = 0,$$

s točnošću $\varepsilon = 10^{-2}$.

Rješenje. Prvo moramo lokalizirati nultočke. Tražimo x koji zadovoljavaju $e^{-x} = 2 - x$. Stoga gledamo gdje se grafovi tih funkcija sijeku:



Grafovi nam služe samo za pomoć i intuiciju, no pravi dokaz moramo provesti analitički, alatima s Matematičke analize 1 i 2. Pogledajmo derivaciju funkcije f:

$$f'(x) = -e^{-x} + 1 = \begin{cases} < 0 & \text{za } x < 0 \\ > 0 & \text{za } x > 0 \end{cases}$$

Dakle, f strogo raste na $\langle 0, +\infty\rangle$ i strogo pada na $\langle -\infty, 0\rangle$. To znači da na svakom od intervala $\langle 0, +\infty\rangle$ i $\langle -\infty, 0\rangle$ imamo najviše jednu nultočku. S druge strane, uvrštavanjem brojeva $-2, -1, 1, 2$ u funkciju f lokaliziramo dvije nultočke $x_0 \in [-2, -1]$ i $y_0 \in [1, 2]$, pa su to i sve nultočke funkcije f, tj. sva rješenja početne jednadžbe:

$$f(-2)f(-1) < 0$$
$$f(2)f(1) < 0.$$

Iz prvog kriterija zaustavljanja

$$n \geq \frac{\log \frac{1}{\varepsilon}}{\log 2} - 1 = 5.643856$$

vidimo da nam je, u oba slučaja $(b - a = 1)$, dovoljno $n = 6$ iteracija.
Izračunajmo i odgovarajuće minimume za dinamički uvjet zaustavljanja ($f \in C^1$). Kako je funkcija f' rastuća, računamo:

$$\min_{y\in[-2,-1]} |f'(y)| = |f'(-1)| = e - 1 \implies m_1 = 1.7182818$$
$$\min_{x\in[1,2]} |f'(x)| = |f'(1)| = -e^{-1} + 1 \implies m_1 = 0.632120559$$

Dakle, za nultočku $x_0 \in [-2, -1]$ uvjet zaustavljanja je

$$|f(x)| \leq 0.017182818.$$

Za nultočku $y_0 \in [1, 2]$ uvjet zaustavljanje je

$$|f(x)| \leq 0.00632120559.$$

Provodimo iteracije na $[1, 2]$:

n	a_n	b_n	x_n	$\lvert f(x_n)\rvert$	$f(a_n)f(x_n)$
0	1	2	1.5	0.2768698	$0.175015 > 0$
1	1.5	2	1.75	0.076226	$0.0211047 > 0$
2	1.75	2	1.875	0.0283549	$-0.002161 < 0$
3	1.75	1.875	1.8125	0.0242545	$0.0018488 > 0$
4	1.8125	1.875	$\boxed{1.84375}$	$0.00197298 < 0.0063212$	

Stajemo u 4. koraku. rješenje je $x_4 = 1.84375$.
Provodimo iteracije na $[-2, -1]$:

n	a_n	b_n	x_n	$\lvert f(x_n)\rvert$	$f(a_n)f(x_n)$
0	-2	-1	-1.5	0.981689	$3.326999 > 0$
1	-1.5	-1	-1.25	0.2403429	$0.235942 > 0$
2	-1.25	-1	-1.125	0.044783	$-0.010763 < 0$
3	-1.25	-1.125	-1.1875	0.0913738	$0.021961 > 0$
4	-1.1875	-1.125	-1.15625	0.021743	$0.00198678 > 0$
5	-1.15625	-1.125	$\boxed{-1.140625}$	$0.011902 < 0.017182818$	

Stajemo u 5. koraku. rješenje je $x_5 = -1.140625$.

△

6.2 Metoda jednostavne iteracije

Pretpostavimo da tražimo rješenje jednadžbe

$$x = f(x).$$

Takve točke x zovemo fiksne točke funkcije f. Definiramo jednostavnu iteracijsku funkciju

$$x_{n+1} = f(x_n), \quad n \geq 0$$

uz x_0 kao neku početnu aproksimaciju x. Naš problem je traženje nultočke $f(x) = 0$. No ako jednadžbu zapišemo u obliku $g(x) = x$, problem se svodi na traženje fiksne točke funkcije g.

Teorem 6.2. Neka je funkcija g neprekidna na $[a, b]$ i neka je

$$g([a, b]) \subset [a, b].$$

Pretpostavimo da postoji konstanta q, takva da je $0 < q < 1$ i da vrijedi

$$|g(x) - g(y)| \leq q|x - y|, \quad \forall x, y \in [a, b].$$

Ovo svojstvo kaže da je g kontrakcija na $[a, b]$. Tada funkcija g ima jedinstvenu fiksnu točku x unutar $[a, b]$. Nadalje, za proizvoljnu startnu točku $x_0 \in [a, b]$, niz iteracija

$$x_n = g(x_{n-1}), \quad n \geq 1,$$

konvergira prema x.

Ako je $g \in C([a, b])$ imamo ocjene pogreške:

1) $|\xi - x_n| \leq \frac{q^n}{1-q}|x_1 - x_0|$

2) $|\xi - x_n| \leq q^n(b - a)$.

Ako je $g \in C^1([a, b])$ i ako je $g([a, b]) \subset [a, b]$, da bismo provjerili da li je kontrakcija, dovoljno je vidjeti da je

$$M_1 = \max_{y \in [a,b]} |g'(y)| \leq q,$$

za neki $q \in \langle 0, 1\rangle$.
Za zadanu točnost $\varepsilon > 0$ potreban broj koraka može se odrediti iz odgovarajućih ocjena:

1) $n \geq \frac{\log \frac{\varepsilon(1-q)}{|x_1 - x_0|}}{\log q}$

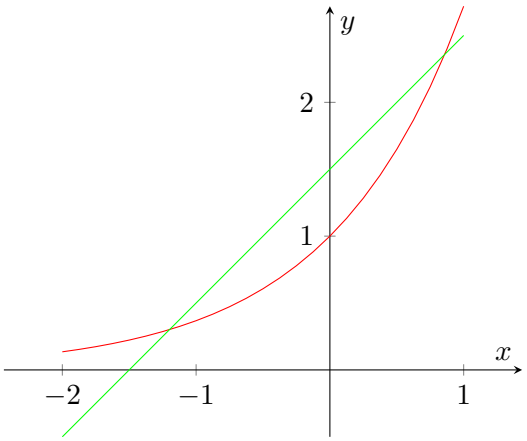
2) $n \geq \frac{\log \frac{\varepsilon}{b-a}}{\log q}$.

Dinamička ocjena pogreške je

$$|\xi - x_n| \leq \frac{q}{1-q}|x_n - x_{n-1}|.$$

Zadatak 6.3. Metodom iteracija riješite jednadžbu $e^x = x + 1.5$ s točnošću $\varepsilon = 0.005$.

Rješenje. Nacrtajmo funkcije e^x i $x + 1.5$:



Iz slike vidimo da bismo trebali dobiti dva rješenja. Jedno u segmentu $[-2, -1]$ i jedno u segmentu $[0, 1]$. To opravdavamo slično kao u prošlom zadatku, pa ovdje te argumente preskačemo.

Definiramo

$$g(x) = e^x - 1.5, \quad f(x) = e^x - x - 1.5.$$

Tada jednadžba iz zadatka glasi $g(x) = x$.
Treba provjeriti je li g kontrakcija. $g'(x) = e^x$, pa je

$$|g'(x)| < 1 \iff x < 0.$$

Dakle, $|g'(x)| < 1$ na $[-2, -1]$, ali ne i na $[0, 1]$.
Promatramo $[-2, -1]$: Vrijedi

$$f(-1)f(-2) < 0,$$

prema tome, imamo nultočku na $[-2, -1]$, tj. imamo rješenje jednadžbe. Treba provjeriti da je $g([-2, -1]) \subset [-2, -1]$. Kako je g strogo rastuća funkcija vrijedi

$$g([-2, -1]) \subset [g(-2), g(-1)] \subset [-1.37, -1.13] \subset [-2, -1].$$

Zaključujemo da je funkcija $g|_{[-2,-1]} : [-2, -1] \to [-2, -1]$ kontrakcija.

$$M_1 = \max_{x \in [-2,-1]} |g'(x)| = \max_{x \in [-2,-1]} e^x = e^{-1} = \underbrace{0.367879}_{=q} < 1,$$

pa je potrebni broj koraka (prema 2))

$$n \geq \frac{\log \frac{\varepsilon}{b-a}}{\log q} = \frac{\log \frac{5\cdot 10^{-3}}{1}}{\log e^{-1}} = 5.2983.$$

Za $n = 6$ iteracija dobit ćemo traženu točnost ε.
S algoritmom krećemo tako da odaberemo $x_0 = -1.5$ (sredina intervala).

$$x_1 = g(x_0) = -1.27686984.$$

Iz 1) dobivamo:

$$n \geq \frac{\log \frac{\varepsilon(1-q)}{|x_1-x_0|}}{\log q} = \frac{\log \frac{5\cdot 10^{-3}\cdot 0.632121}{|0.22313016|}}{\log e^{-1}} = 4.25,$$

pa vidimo da nam je dovoljno 5 iteracija.

$$\begin{aligned} x_0 &= -1.5, \\ x_1 = g(x_0) &= -1.27686984, \\ x_2 = g(x_1) &= -1.221091, \\ x_3 = g(x_2) &= -1.205092, \\ x_4 = g(x_3) &= -1.200336, \\ x_5 = g(x_4) &= -1.198907. \end{aligned}$$

Promatramo $[0, 1]$: Primijenimo ln na jednadžbu i dobijemo:

$$x = \ln(x + 1.5).$$

Definiramo funkciju $g(x) = \ln(x + 1.5)$. Provjerimo je li to konrakcija:

$$g'(x) = \frac{1}{x + 1.5},$$

$$M_1 = \max_{x\in[0,1]} |g'(x)| = |g'(0)| = \underbrace{\frac{2}{3}}_{=q} < 1,$$

$$g([0, 1]) \underbrace{\subset}_{g\ raste} [g(0), g(1)] \subset [0.405, 0.917] \subset [0, 1].$$

Dakle, g je kontrakcija. Ocjenjujemo potreban broj iteracija:

$$n \geq \frac{\log \frac{\varepsilon}{b-a}}{\log q} = \frac{\log \frac{5\cdot 10^{-3}}{1}}{\log \frac{2}{3}} = 13.067,$$

pa nam je potrebno 14 iteracija da zadovoljimo zadanu točnost. Ako uzmemo $x = 0.5$ (polovište intervala)

$$x_1 = g(x_0) = 0.693147,$$

pa iz druge ocjene dobijemo da je potrebno $n = 12$ iteracija. Računamo:

$$\begin{aligned} x_0 &= 0.5, \\ x_1 = g(x_0) &= 0.693147, \\ x_2 = g(x_1) &= 0.785338, \\ x_3 = g(x_2) &= 0.826514, \\ x_4 = g(x_3) &= 0.844371, \\ x_5 = g(x_4) &= 0.852017, \\ x_6 = g(x_5) &= 0.855273, \\ x_7 = g(x_6) &= 0.856657, \\ x_8 = g(x_7) &= 0.857243, \\ x_9 = g(x_8) &= 0.857493, \\ x_{10} = g(x_9) &= 0.857599, \\ x_{11} = g(x_{10}) &= 0.857644, \\ x_{12} = g(x_{11}) &= 0.857663. \end{aligned}$$

△

6.3 Newtonova metoda

Početne pretpostavke na funkciju f:

- $f : [a, b] \to \mathbb{R}$, $f \in C^2([a, b])$,
- $f(a)f(b) \leq 0$,
- $|f'(x)|$, $|f''(x)| > 0$, $\forall x \in [a, b]$ (f' i f'' imaju konstantan predznak)

tada Newtonova metoda konvergira prema jedinstvenoj nultočki ξ, i to za svaku početnu aproksimaciju $x_0 \in [a, b]$, za koju vrijedi

$$f(x_0)f''(x_0) > 0.$$

Algoritam:
Počevši od polazne aproksimacije $x_0 \in [a, b]$ iterativno se formira niz točaka $(x_n)_n$ koji konvergira prema nultočki ξ.

- **POČETAK:** Odaberemo početnu točku x_0 tako da $f(x_0)f''(x_0) > 0$.
- **ITERACIJA:** sve dok nije zadovoljena tražena točnost ε ponavljamo

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

Ocjena pogreške:

$$|\xi - x_n| \leq \frac{M_2}{2m_1}|x_n - x_{n-1}|^2,$$

$$M_2 = \max_{x\in[a,b]}|f''(x)|, \;\; m_1 = \min_{x\in[a,b]}|f'(x)|.$$

Kriterij zaustavljanja:

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

Napomenimo još da je uvjet da f' i f'' ne mijenaju predznak ponekad teško provjeriti, pa je dovoljno osigurati da $f'(\xi) \neq 0$ i $f''(\xi) \neq 0$ (nije višestruka nultočka).

Zadatak 6.4. Nađite najveće negativno rješenje jednadžbe

$$\operatorname{tg} x - \frac{1}{2}x^2 + x + 1 = 0$$

s točnošću $\varepsilon = 10^{-4}$. Duljina početnog intervala za nalaženje rješenja mora biti barem $1/2$.

Napomena: Detaljno obrazložite sve svoje tvrdnje vezane za lokaciju nultočke i ocjenu greške!

Rješenje. Zadatak 9.2.1. `https://web.math.pmf.unizg.hr/nastava/unm/vjezbe/nm_dodzad.pdf` △

Zadatak 6.5. Nađite najmanje rješenje jednadžbe

$$xe^{-x} + 1 = 2x\sqrt{x}$$

s točnošću $\varepsilon = 10^{-6}$. Duljina početnog intervala za nalaženje rješenja mora biti barem $1/2$.

Napomena: Detaljno obrazložite sve svoje tvrdnje vezane za lokaciju nultočke i ocjenu greške!

Rješenje. Zadatak 9.1.1. `https://web.math.pmf.unizg.hr/nastava/unm/vjezbe/nm_dodzad.pdf` △

6.4 Sustavi nelinearnih jednadžbi

Ovdje ćemo spomenuti jednu metodu za rješavanje sustava nelinearnih jednadžbi: Newtonovu metodu. Problem sustava nelinearnih jednadžbi mogu se zapisati kao problem nalaženja nultočke funkcije $f : \mathbb{R}^n \to \mathbb{R}^n$: treba naći $x \in \mathbb{R}^n$ (ili nekog podskupa tog skupa) takav da je $f(x) = 0$. Newtonova metoda za rješavanje tog problema opet je zadana u obliku iteracija i glasi

$$x^{(k+1)} = x^{(k)} - \left(D_f(x^{(k)})\right)^{-1} f(x^{(k)}).$$

Konvergencija ove metode je opet kvadratna, ali kao i u jednodimenzionalnom slučaju ima iste probleme (konvergencija je garantirana tek u maloj okolini nultočke). Za kriterij zaustavljanja uzima se

$$\max_{i=1,\dots,n} \frac{\left|x_i^{(k+1)} - x_i^{(k)}\right|}{\left|x_i^{(k)}\right|} \leq \varepsilon.$$

Zadatak 6.6. Riješi sustav jednadžbi

$$\begin{aligned} \ln(x_1^2 + x_2) + x_2 - 1 &= 0 \\ \sqrt{x_1} + x_1 x_2 &= 0 \end{aligned}$$

uz početnu aproksimaciju $(x_1^{(0)}, x_2^{(0)}) = (2.4, -0.6)$ i tri iteracije Newtonove metode.

Rješenje. Označimo s

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad f(x_1, x_2) = \begin{bmatrix} ln(x_1^2 + x_2) + x_2 - 1 \\ \sqrt{x_1} + x_1 x_2 \end{bmatrix}.$$

Newtonova metoda je dana formulom

$$x^{(k+1)} = x^{(k)} - \left(D_f(x^{(k)})\right)^{-1} f(x^{(k)})$$

gdje je

$$D_f(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} \frac{2x_1}{x_1^2+x_2} & 1 + \frac{1}{x_1^2+x_2} \\ x_2 + \frac{1}{2\sqrt{x_1}} & x_1 \end{bmatrix}.$$

Za inverz 2×2 matrice koristimo formulu

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}^{-1}$$

(identitet možete provjeriti "na ruke", a formula slijedi iz formule za inverz matrice pomoću adjuntke; primijetite da je izraz u nazivniku faktora zapravo determinanta te matrice). Sada je

$$(D_f(x))^{-1} = \frac{1}{J_f(x)} \begin{bmatrix} \frac{2x_1}{x_1^2+x_2} & -1 - \frac{1}{x_1^2+x_2} \\ -x_2 - \frac{1}{2\sqrt{x_1}} & x_1 \end{bmatrix}$$

gdje je

$$\begin{aligned} J_f(x) &= \frac{2x_1^2}{x_1^2 + x_2} - \frac{x_1^2 + x_2 + 1}{x_1^2 + x_2} \cdot \frac{1 + 2\sqrt{x_1} x_2}{2\sqrt{x_1}} \\ &= \frac{1}{x_1^2 + x_2} \left(2x_1^2 - \left(x_1^2 + x_2 + 1\right) \left(x_2 + \frac{1}{2\sqrt{x_1}} \right) \right). \end{aligned}$$

Dakle, Newtonova formula glasi

$$\begin{bmatrix} x_1^{(k+1)} \\ \\ x_2^{(k+1)} \end{bmatrix} = \begin{bmatrix} x_1^{(k)} \\ \\ x_2^{(k)} \end{bmatrix} - \frac{1}{J_f(x)} \begin{bmatrix} \frac{2x_1^{(k)}}{\left(x_1^{(k)}\right)^2 + x_2^{(k)}} & -1 - \frac{1}{\left(x_1^{(k)}\right)^2 + x_2^{(k)}} \\ \\ -x_2^{(k)} - \frac{1}{2\sqrt{x_1^{(k)}}} & x_1^{(k)} \end{bmatrix}$$

Za početnu aproksimaciju $(x_1^{(0)}, x_2^{(0)}) = (2.4, -0.6)$ dobivamo sljedeću tablicu △

k	$x_1^{(k)}$	$x_2^{(k)}$
0	2.4	-0.6
1	2.4125245	-0.6440504
2	2.4122488	-0.6438563
3	2.4122488	-0.6438563

6.5 Uvod u optimizaciju

Za probleme bezuvjetne optimizacije (traženje minimuma funkcije $f : \mathbb{R}^n \to \mathbb{R}$) na predavanjima smo radili metode spusta. To su iterativne metode u kojima se, počevši s iteracijom θ^0, svaka nova iteracija računa formulom

$$\theta^{(k+1)} = \theta^{(k)} - \eta_k d^{(k)},$$

dok nije zadovoljen neki uvjet zaustavljanja. Skalar η_k naziva se faktorom brzine učenja, a $d^{(k)}$ smjerom. Kod **metode gradijentnog spusta** za smjer se uzima $d^{(k)} = \nabla f(\theta^{(k)})$. Optimalni faktor brzine učenja nije lako odabrati. Na predavanjima se radio rezultat:

Teorem 6.7. Neka je $f : \mathbb{R}^n \to \mathbb{R}$ konveksna i dvaput diferencijabilna funkcija, te neka joj je θ_* stacionarna točka, tj. $\nabla f(\theta_*) = 0$. Dodatno, pretpostavimo da je ∇f Lipschitz neprekidna funkcija s konstantom L.Tada funkcijske vrijednosti $(f(\theta^{(k)}))_{k\geq 0}$ iz gradijentnog spusta s proizvoljnom početnom iteracijom $\theta^{(0)}$ konvergiraju ka globalnom minimumu $f(\theta_*)$ za svaki konstantni korak $\eta \leq 1/L$, te vrijedi

$$f(\theta^{(k)}) - f(\theta_*) \leq \frac{\|\theta^{(0)} - \theta_*\|^2}{2\eta k}.$$

Zadatak 6.8. Dana je funkcija

$$f(x, y) = \frac{1}{2} \begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix} + b^T \begin{bmatrix} x \\ y \end{bmatrix},$$

za neku pozitivno definitnu 2×2 matricu A i vektor b duljine 2.

(a) Za $A = \begin{bmatrix} 6 & 2 \\ 2 & 6 \end{bmatrix}$ i $b = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ provedite dva koraka gradijentnog spusta s početnom iteracijom $(0, 0)$ i korakom 0.1.

(b) *(za općenite A i b):* Dokažite da funkcija ima globalni minimum i odredite ga, te točku u kojoj se postiže u ovisnosti o A i b.

(c) *(za općenite A i b):* Za koje početne korake (x^0, y^0) postoji korak η takav da gradijentna metoda u prvom koraku pogađa globalni minimum.

Rješenje. Za ovaj izbor matrica A i b imamo

$$f(x, y) = 3x^2 + 2xy + 3y^2 + x - 2y, \quad \nabla f(x, y) = \begin{bmatrix} 6x + 2y + 1 \\ 2x + 6y - 2 \end{bmatrix}.$$

Zato je

$$\begin{bmatrix} x^{(k+1)} \\ y^{(k+1)} \end{bmatrix} = \begin{bmatrix} x^{(k)} \\ y^{(k)} \end{bmatrix} - \eta \nabla f(x, y) = \begin{bmatrix} 6x^{(k)} + 2y^{(k)} + 1 \\ 2x^{(k)} + 6y^{(k)} - 2 \end{bmatrix} = \begin{bmatrix} (1 - 6\eta) & -2\eta \\ -2\eta & (1 - 6\eta) \end{bmatrix} \begin{bmatrix} x^{(k)} \\ y^{(k)} \end{bmatrix} + \begin{bmatrix} -\eta \\ 2\eta \end{bmatrix}.$$

Za $\eta = 0.1$, imamo

$$\begin{bmatrix} x^{(k+1)} \\ y^{(k+1)} \end{bmatrix} = \begin{bmatrix} 0.4 & -0.2 \\ -0.2 & 0.4 \end{bmatrix} \begin{bmatrix} x^{(k)} \\ y^{(k)} \end{bmatrix} + \begin{bmatrix} -0.1 \\ 0.2 \end{bmatrix}.$$

Zato je $(x^0, y^0) = (0, 0)$, $(x^1, y^1) = (-0.1, 0.2)$, $(x^2, y^2) = (-0.18, 0.3)$. Kako je točka globalnog minimuma $(-0.3125, 0.4375)$, vidimo da je konvergencija dosta spora.

Općenito, imamo

$$\nabla f(x, y) = A \begin{bmatrix} x \\ y \end{bmatrix} + b, \quad Hf(x, y) = A.$$

Kako je Hesseova matrica konstantna pozitivno definitna matrica, f je konveksna pa u stacionarnoj točki postiže globalni minimum. Točka globalnog minimuma je jedinstveno rješenje sustava $A\theta_* = b$ (matrica A je regularna jer je pozitivno definitna).

Kada bi gradijentni spust završio u jednom koraku, tada bi trebalo vrijediti

$$\theta_* = \theta^{(1)} = \theta^{(0)} - \eta \nabla f(\theta^{(0)}) = \theta^{(0)} - \eta(A\theta^{(0)} + b).$$

Korištenjem distributivnosti i svojstva $A\theta_* = b$, imamo

$$\begin{aligned} 0 = A\theta_a st + b &= A(\theta^{(0)} - \eta(A\theta^{(0)} + b)) + b \\ &= A\theta^{(0)} - \eta A^2\theta^{(0)} - \eta Ab + b \\ &= (I - \eta A)A\theta^{(0)} + (I - \eta A)b \\ &= (I - \eta A)(A\theta^{(0)} + b). \end{aligned}$$

Zadnji izraz jednak je nuli kada je drugi faktor $v := A\theta^{(0)} + b$ jednak nulvektoru, što znači da je već početna iteracija točka globalnog minimuma, ili ako se taj vektor v nalazi u jezgri matrice $(I - \eta A)$ za neki pozitivni realni η. Za to treba vidjeti kada je ta matrica singularna. Očito nije singularna za $\eta = 0$, a inače možemo pisati

$$I - \eta A = -\eta \left(A - \frac{1}{\eta} I \right)$$

odakle čitamo da je singularna kada je η recipročna vrijednost svojstvene vrijednosti matrice A. Dakle, gradijentni spust s početnom iteracijom $\theta^{(0)}$ će završiti u jednom koraku ako za duljinu koraka uzmemo $1/\lambda_A$, a za početnu iteraciju vrijedi da je $A\theta^{(0)} + b$ jednak svojstvenom vektoru te svojstvene vrijednosti λ_A, ili za proizvoljnu duljinu koraka ako je početna iteracija upravo točka globalnog minimuma.

△