

Numerične metode

izročki predavanj

Aljaž Zalar

Fakulteta za računalništvo in informatiko
Univerza v Ljubljani

Verzija 07.10.2022

Literatura

Osnovna vira:

- ▶ Bojan Orel, *Osnove numerične matematike*, Založba FE in FRI.
- ▶ Bor Plestenjak: *Razširjen uvod v numerične metode*, DMFA založništvo.

Tuji viri:

- ▶ K. Atkinson, W. Han: *Elementary Numerical Analysis*, 3rd edition, John Wiley & Sons, Inc., New Jersey, 2003.
- ▶ R.L. Burden, J.D. Faires, A.M. Burden: *Numerical Analysis*, 10th edition, Cengage Learning, Boston, 2016.
- ▶ G.H. Golub, C.F. Van Loan: *Matrix Computations*, 3rd edition, Johns Hopkins Univ. Press, Baltimore, 1996.
- ▶ D.R. Kincaid, E.W. Cheney: *Numerical Analysis, Mathematics of Scientific Computing*, 3rd edition, Brooks/Cole, Pacific Grove, 2002.
- ▶ L.N. Trefethen, D. Bau: *Numerical Linear Algebra*, SIAM, Philadelphia, 1997.

Obveznosti

Potek predmeta:

- ▶ Predavanja: 3 ure na teden.
- ▶ Vaje: 2 uri na teden.

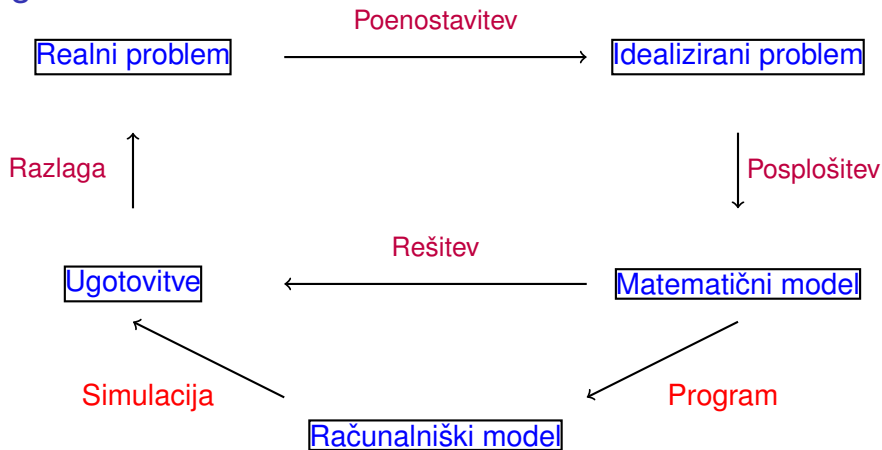
Ocena:

- ▶ 3 domače naloge.
- ▶ Pisni izpit.
- ▶ Ustni izpit.

Programska oprema:

- ▶ *Matlab*: Licenca dostopna za študente UL.
- ▶ *Octave*: Prosto dostopna alternativa Matlabu.

Vloga numerične matematike



Numerična matematika ima ključno vlogo pri pretvorbi matematičnega modela v računalniškega, reševanju tega modela in razlagi rešitev s stališča napak.

Vsebina predmeta

1. Računanje in vloga napak pri numerični matematiki
2. Reševanje sistemov linearnih enačb
 - ▶ Gausova eliminacija in LU razcep - cena in problemi
 - ▶ Pivotiranje
 - ▶ Iterativne metode - Jacobijeva in Gauss-Seidlova iteracija
3. Reševanje (sistemov) nelinearnih enačb in optimizacija
 - ▶ Tangentna oz. Newtonova metoda
 - ▶ Metoda fiksne točke
 - ▶ Newtonova optimizacijska metoda
4. Aproksimacija in interpolacija
 - ▶ Lagrangeov in Newtonov interpolacijski polinom
 - ▶ Aproksimacija po metodi najmanjših kvadratov
 - ▶ QR razcep za predoločene sisteme

5. Numerično odvajanje in integriranje

- ▶ Trapezna metoda
- ▶ Simpsonova metoda
- ▶ Rombergova metoda

6. Numerično reševanje diferencialnih enačb

- ▶ Eulerjeva metoda
- ▶ Runge-Kutta metode

Prvo poglavje:

Uvod v numerično računanje

- ▶ Numerično računanje
- ▶ Predstavljiva števila
- ▶ Zaokrožitvene napake
- ▶ Katastrofalno seštevanje/odštevanje
- ▶ Primeri (ne)stabilnega računanja

Numerično in simbolno računanje

Numerično računanje:

- ▶ Takoj v formulo vstavljamo **števila**
- ▶ Pridemo do numeričnega rezultata - **numerične rešitve**

Simbolno računanje:

- ▶ **simboli** predstavljajo števila
- ▶ izraz preoblikujemo s simbolnim računanjem do novega simbolnega izraza - **analitična rešitev**

Primer

- ▶ *Numerično:*

$$\frac{(17.36)^2 - 1}{17.36 + 1} = 16.36; \quad 0.25, 0.33333 \dots (?), 3.14159 \dots (?)$$

- ▶ *Simbolno:*

$$\frac{x^2 - 1}{x + 1} = x - 1; \quad \frac{1}{4}, \frac{1}{3}, \pi, \tan 83$$

Numerično in simbolno računanje

Primer

```
1 >> x=rand; (x^2-1)/(x+1) - (x-1)
2
3 ans=1.387778780781446e-17
```

Analitično bi rezultat moral biti 0, vendar zaradi numeričnih napak dobimo majhno napako.

Kaj zanima numerično matematiko?

Metoda . . . matematična konstrukcija, s katero rešujemo problem

Algoritem . . . koraki metode

Implementacija . . . zapis algoritma v izbranem jeziku

Kaj pomeni 'biti numerično dober'?

majhna sprememba podatkov \Rightarrow majhna napaka rezultata

Tipična vprašanja numerične matematike:

- ▶ Ali je problem občutljiv?
- ▶ Ali je metoda 'dobra'?
- ▶ Ali je algoritem robusten - deluje na širokem spektru problemov?
- ▶ Ali je implementacija hitra - časovna in prostorska zahtevnost?

Občutljivih problemov NM ne more rešiti

Problem je občutljiv, če se ob majhni spremembi začetnih podatkov točen rezultat zelo spremeni.

Občutljivost je odvisna le od narave problema in ne od izbrane numerične metode.

Primer (presečišča premic)

Sistem in njegova perturbacija

$$x + y = 2 \quad \rightarrow \quad x + y = 1.9999$$

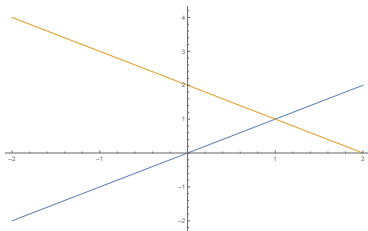
$$x - y = 0 \quad \rightarrow \quad x - y = 0.0002$$

ima rešitvi $x = y = 1$ oz. $x = 1.00005$ in $y = 0.99985$. Problem je neobčutljiv, saj je šlo za spremembo za isti velikostni razred.

Sistem in njegova perturbacija

$$\begin{aligned}x + 0.99y &= 1.99 & \rightarrow & \quad x + 0.99y = 1.9899 \\0.99x + 0.98y &= 1.97 & \rightarrow & \quad 0.99x + 0.98y = 1.9701\end{aligned}$$

ima rešitvi $x = y = 1$ oz. $x = 2.97$ in $y = -0.99$. Problem je občutljiv, saj je majhna sprememba začetnih podatkov povzročila veliko spremembo rezultata.



Na čem temeljijo numerične metode?

- ▶ **Matrike nadomestimo z enostavnejšimi** (upoštevamo samo diagonalni ali zgornjetrikotni del).
- ▶ **Nelinearne probleme nadomestimo z linearnimi** (linearna aproksimacija v točki).
- ▶ **Neskončne procese nadomestimo s končnimi** (uporabimo Taylorjev polinom) .
- ▶ **Neskončno razsežne prostore nadomestimo s končno razsežnimi** (funkcije nadomestimo s polinomi).
- ▶ **Diferencialne enačbe nadomestimo z algebraičnimi** (znebimo se vseh parcialnih odvodov iz enačb).

Zakaj sploh potrebujemo numerično matematiko?

Znanost, ki temelji na matematičnih izračunih, je neposredno odvisna od NM.

Nekatere katastrofe so se zgodile zaradi slabega numeričnega računanja (<http://www-users.math.umn.edu/~arnold//disasters/>):

- ▶ *Nesreča Misije Patriot, Zalivska vojna 1991, Savdska Arabija, 28 žrtev: slaba analiza zaokrožitvenih napak.*

Čas zadetka iraške rakete, usmerjene na Savdsko Arabijo, je bil računat na vsako desetino sekunde v 24-bitnem sistemu. Ker velja

$$\frac{1}{10} = 2^{-4} + 2^{-5} + 2^{-8} + 2^{-9} + 2^{-12} + 2^{-13} + 2^{-16} + 2^{-17} + 2^{-20} + 2^{-21} +$$

$\underbrace{+2^{-24} + 2^{-25} + 2^{-28} + \dots}_{\text{zanemarimo}}$

je vsako desetinko sekunde napaka $9.5 \cdot 10^{-8}$ s. Po 100 urah računanja je bila napaka $9.5 \cdot 10^{-8}$ s $\cdot 100 \cdot 60 \cdot 60 \cdot 10 = 0.34$ s. Ker je hitrost rakete 1.676 m/s, je bila pozicija rakete za več kot 500 m napačno predvidena in je ta ušla radarjem.

- ▶ *Eksplozija rakete Ariana 5, Francoska Gvajana, 1996: posledica prekoračitve obsega števil.*

https://www.youtube.com/watch?v=PK_yguLapgA

<https://www.youtube.com/watch?v=W3YJeoYgozw>

Ob prenovi rakete so 'pozabili' nadgraditi uporabljen številski sistem, ki je horizontalno hitrost meril v 16-bitnem sistemu (1 bit porabimo za predznak). Največja hitrost v tem sistemu je

$$2^0 + 2^1 + \dots + 2^{13} + 2^{14} = \frac{2^{15} - 1}{2 - 1} = 32767.$$

Ker je prenovljena raketa po 37 sekundah preseгла to hitrost, je prišlo do zaustavitve motorjev...

- ▶ *Potop naftne ploščadi Sleipner A, Stavanger, Norveška, 1991, milijarda dolarjev škode: nenatančna obdelava obremenitev pri reševanju PDE-jev.*

<https://www.youtube.com/watch?v=eGdiPs4THW8>

Ponovitev predstavljivih števil

Števila shranjujemo v obliki

$$x = \pm 0.d_1d_2d_3 \dots d_m \times \beta^e,$$

kjer je

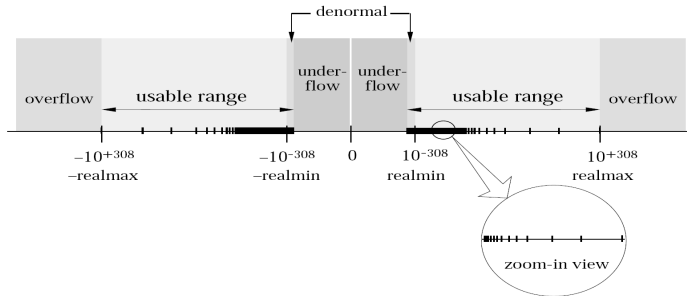
- ▶ β naravno število (v računalništvu $\beta = 2$),
- ▶ $d_1d_2d_3 \dots d_m$ mantisa, e eksponent.

Primer (baza 10)

- ▶ 1000.12345 zapišemo kot $+(0.100012345)_{10} \times 10^4$.
- ▶ 0.000812345 zapišemo kot $+(0.812345)_{10} \times 10^{-3}$.

Prekoračitev in podkoračitev

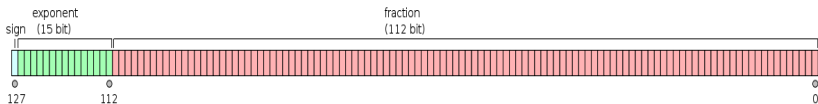
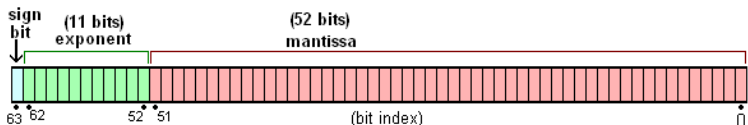
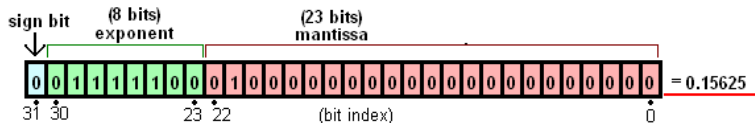
Floating Point Number Line



- ▶ izračuni preblizu 0 lahko povzročijo **podkoračitev**
- ▶ preveliki izračuni lahko povzročijo **prekoračitev**
- ▶ prekoračitev je v splošnem hujši problem

Različne natančnosti

- ▶ *IEEE Enojna natančnost*: števila so predstavljena z 32 biti.
- ▶ *IEEE Dvojna natančnost*: števila so predstavljena z 64 biti.
- ▶ *Multiprecision Computing Toolbox for MATLAB*: Omogoča računanje v višjih natančnostih. Dostopno na naslovu <https://www.advanpix.com/>



Kaj so zaokrožitvene napake?

- ▶ Večine realnih števil ne moremo predstaviti v strojni aritmetiki \Rightarrow **zaokrožujemo** in delamo **zaokrožitvene napake**.
- ▶ IEEE standard... **zaokroži x do najbližjega predstavljivega števila $\text{fl}(x)$** . Naj bosta

$$x_- \leq x \leq x_+$$

najbližji predstavljivi števili števila x . Potem je

$$\text{fl}(x) = \begin{cases} x_-, & \text{če je } x \text{ bližje } x_-, \\ x_+, & \text{če je } x \text{ bližje } x_+. \end{cases}$$

- ▶ Kako velika je napaka? Recimo, da je x bližje x_- :

$$\begin{aligned} x &= (0.1b_2b_3 \dots b_mb_{m+1})_2 \times 2^e, \\ x_- &= (0.1b_2b_3 \dots b_m)_2 \times 2^e, \\ x_+ &= ((0.1b_2b_3 \dots b_m)_2 + 2^{-m}) \times 2^e, \end{aligned}$$

$$\text{fl}(x) = x(1 + \delta), |\delta| < 2^{-m}$$

Absolutna napaka:

$$x - x_- \leq \frac{x_+ - x_-}{2} = 2^{e-m-1}.$$

Relativna napaka:

$$\frac{x - x_-}{x} \leq \frac{2^{e-m-1}}{1/2 \times 2^e} \leq \underbrace{2^{-m}}_u \dots \text{osnovna zaokrožitvena napaka}$$

Torej je

$$x_- = x_- - x + x \geq -ux + x = x(1 - u).$$

Podobno

$$x_+ \leq x(1 + u).$$

Sledi

$$\boxed{\text{fl}(x) = x(1 + \delta)}, \quad \text{kjer je } |\delta| < u.$$

Kako računamo s predstavljivimi števili?

Za **predstavljivi** števili x, y in katerokoli od osnovnih operacij $\odot \in \{+, -, \cdot, :\}$ število $x \odot y$ ni nujno predstavljivo. Po zgornjem pa velja

$$\boxed{\text{fl}(x \odot y) = (x \odot y)(1 + \delta)}, \quad \text{kjer je } |\delta| \leq u.$$

Seštevanje numerično **ni asociativna operacija**, tj.

$$\boxed{(a + b) + c \neq a + (b + c)} :$$

Primer

```
1 >> a=rand;b=rand;c=rand;((a+b)+c)-(a+(b+c))
2
3 ans=-2.220446049250313e-16
```

Seštevamo od manjših k večjim številom

$$\begin{aligned}(a + b) + c &= \text{fl}(\text{fl}(a + b) + c) = \text{fl}((a + b)(1 + \delta_1) + c) \\ &= [(a + b)(1 + \delta_1) + c](1 + \delta_2) \\ &= [(a + b + c) + (a + b)\delta_1](1 + \delta_2) \\ &= (a + b + c) \left[1 + \frac{a + b}{a + b + c} \delta_1(1 + \delta_2) + \delta_2 \right]\end{aligned}$$

Podobno

$$a + (b + c) = (a + b + c) \left[1 + \frac{b + c}{a + b + c} \delta_3(1 + \delta_4) + \delta_4 \right].$$

Če pozabimo na člena $\delta_1\delta_2$ in $\delta_3\delta_4$ (Zakaj to lahko naredimo?), dobimo

$$(a + b) + c = (a + b + c)(1 + \epsilon_3) \quad \text{kjer je} \quad \epsilon_3 \approx \frac{a + b}{a + b + c} \delta_1 + \delta_2,$$

$$a + (b + c) = (a + b + c)(1 + \epsilon_4) \quad \text{kjer je} \quad \epsilon_4 \approx \frac{b + c}{a + b + c} \delta_3 + \delta_4.$$

Sklep: Ko seštevamo števila, je za čim manjšo napako najbolje začeti z najmanjšim in prištevati večje.

Napake pri numeričnem računanju

- ▶ Neodstranljiva napaka $D_n \dots$ nenatančni začetni podatki.
- ▶ Napaka metode $D_m \dots$ npr. neskončni proces aproksimiramo s končnim.
- ▶ Zaokrožitvena napaka $D_z \dots$ računanje s približki in zaokroževanje.

Celotna napaka D je

$$D = D_n + D_m + D_z.$$

Stabilnost meri kakovost metode

Stabilnost metode preverimo z **analizo zaokrožitvenih napak**.

Vrste napak (x naj bo točna vrednost, \bar{x} pa približek zanjo):

▶ Prva delitev:

▶ **Absolutna napaka**: $\bar{x} - x$.

▶ **Relativna napaka**: $\frac{\bar{x} - x}{x}$.

▶ Druga delitev:

▶ **Direktna napaka**: Numerična napaka rezultata.

▶ **Obratna napaka**: Koliko je potrebno spremeniti začetne podatke, da dobimo izračunan rezultat.

Velja

$$|\text{direktna napaka}| \approx \text{občutljivost} \times |\text{obratna napaka}|.$$

Izračunana vrednost je blizu pravi, če rešujemo neobčutljiv problem z obratno stabilno metode.

Odštevanje in seštevanje sta lahko 'katastrofalni'

odštevanje dveh približno enakih števil

seštevanje dveh približno nasprotnih števil

$$a = x.xxxx\ xxxx\ xxx1 \overbrace{ssss\dots}^{\text{izguba}}$$

$$b = x.xxxx\ xxxx\ xxx0 \overbrace{tttt\dots}^{\text{izguba}}$$

$$\begin{array}{r} \text{Potem} \\ \begin{array}{r} \overbrace{x.xxx\ xxxx\ xxx1}^{\text{končna natančnost}} \\ - \overbrace{x.xxx\ xxxx\ xxx0}^{\text{končna natančnost}} \\ \hline = 0.000\ 0000\ 0001 \quad \text{???? ????} \\ = 1. \underbrace{\text{???? ????}}_{\text{izguba natančnosti}} \cdot \beta^{-m} \end{array} \end{array}$$

S ponavljanjem se napake seštevajo.

Primer katastrofalnega odštevanja

Iščemo rešitve kvadratne enačbe

$$x^2 + 2ax + b = 0, \quad \text{kjer je } a > 0 \text{ in } a^2 > b.$$

Rešitev z manjšo absolutno vrednostjo je

$$x_2 = \frac{-2a + \sqrt{4a^2 - 4b}}{2} = -a + \sqrt{a^2 - b}.$$

1 $k_1 := a^2$

2 $k_2 := k_1 - b$

3 $k_3 := \sqrt{k_2}$

4 $k_4 := -a + k_3$

Če je a^2 veliko večji od b , potem ima lahko korak 4 veliko napako. Možna rešitev:

$$x_2 = (-a + \sqrt{a^2 - b}) \cdot \frac{a + \sqrt{a^2 - b}}{a + \sqrt{a^2 - b}} = \frac{-b}{a + \sqrt{a^2 - b}}.$$

```
1  $k_1 := a^2$   
2  $k_2 := k_1 - b$   
3  $k_3 := \sqrt{k_2}$   
4  $k_4 := a + k_3$   
5  $k_5 := \frac{-b}{k_4}$ 
```

```
1 >> a = 10000; b=-1;  
2 >> x = -a+sqrt(a^2 - b)  
3 x = 5.0000000055588316e-05  
4  
5 >> x^2 + 2 * a * x +b  
6 ans = 1.361766321927860e-08  
7  
8 >> x = -b/(a+sqrt(a^2-b))  
9 x = 4.999999987500000e-05  
10  
11 >> x^2 + 2 * a * x +b  
12 ans = -9.011402890989895e-17
```

Koda primera: [klik](#)

Računanje s stabilnejšo obliko

- ▶ Izračun vrednosti funkcije

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

ni stabilen za velike x , ker je $\sqrt{x+1} \approx \sqrt{x}$. Tej težavi se lahko izognemo:

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

Koda primera: [klik](#)

- ▶ Vrsto

$$\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \dots + \frac{1}{n(n+1)},$$

ki se sešteje v $\frac{n}{n+1}$ (dokaz: indukcija), je bolje numerično računati vzvratno kot

$$\frac{1}{n \cdot (n+1)} + \frac{1}{(n-1) \cdot n} + \dots + \frac{1}{1 \cdot 2}.$$

Koda primera: [klik](#)

- Vrednost integrala $I_n = \int_0^1 x^n e^{-x} dx$ se lahko rekurzivno (integracija per partes) izračuna kot

$$I_n = -\frac{1}{e} + nI_{n-1}, \quad I_0 = 1 - \frac{1}{e}.$$

Če iz formule izrazimo I_{n-1} , dobimo

$$I_{n-1} = \frac{1}{n}I_n + \frac{1}{ne}.$$

Izkaže se, da je druga formula boljša, pri čemer za začetni približek I_N (pri velikem N) lahko vzamemo karkoli. Zakaj?

Koda primera: [klik](#)

Seštevanje in odštevanje v splošnem nista relativno direktno stabilni operaciji

$x, y \in \mathbb{R}$. Računamo približek \bar{p} za $p = x + y$.

$$\begin{aligned}\bar{p} &= \text{fl}(\text{fl}(x) + \text{fl}(y)) = \text{fl}(x(1 + \delta_1) + y(1 + \delta_2)) \\ &= (x(1 + \delta_1) + y(1 + \delta_2))(1 + \delta_3) \\ &= x(1 + \delta_1)(1 + \delta_3) + y(1 + \delta_2)(1 + \delta_3) \\ &= x + y + x(\delta_1 + \delta_3 + \delta_1\delta_3) + y(\delta_2 + \delta_3 + \delta_2\delta_3)\end{aligned}$$

kjer je $|\delta_i| \leq u$, $i = 1, 2, 3$. Relativna napaka je

$$\frac{|\bar{p} - p|}{|p|} \leq \frac{|x(\delta_1 + \delta_3 + \delta_1\delta_3) + y(\delta_2 + \delta_3 + \delta_2\delta_3)|}{|x + y|}.$$

Torej:

Če je $x + y$ blizu 0, potem je $\frac{|\bar{p} - p|}{|p|}$ veliko.

Množenje (in deljenje) je relativno direktno stabilna operacija

$x, y \in \mathbb{R}$. Računamo približek \bar{p} za $p = x \cdot y$.

$$\begin{aligned}\bar{p} &= \text{fl}(\text{fl}(x) \cdot \text{fl}(y)) = \text{fl}(x(1 + \delta_1) \cdot y(1 + \delta_2)) \\ &= x(1 + \delta_1) \cdot y(1 + \delta_2)(1 + \delta_3) \\ &= xy(1 + \delta_1 + \delta_2 + \delta_3 + \text{produkti več } \delta),\end{aligned}$$

kjer je $|\delta_i| \leq u$, $i = 1, 2, 3$. Relativna napaka je

$$\boxed{\frac{|\bar{p} - p|}{|p|} \leq \frac{|xy||\delta_1 + \delta_2 + \delta_3 + \mathcal{O}(u^2)|}{|xy|} = |\delta_1 + \delta_2 + \delta_3 + \mathcal{O}(u^2)|}.$$

Torej:

Relativna napaka $\frac{|\bar{p} - p|}{|p|}$ ni odvisna od velikosti produkta xy .

Večina numeričnih metod ni relativno direktno stabilnih

Vse numerične metode, kjer sta vključeni

operaciji $+$ / $-$

in kot rezultat lahko dobimo npr. vrednost 0 ali nekje po poti kot vmesno vrednost skoraj singularno matriko, **niso relativno direktno stabilne**, tj. v rezultatu je lahko veliko relativna napaka.

Zato moramo vedno premisliti:

1. V katerih primerih so zgodi velika napaka?
2. Kako nestabilne primere preoblikovati v stabilne?

Primeri takih operacij:

- ▶ Računanje vrednosti polinoma.
- ▶ Računanje skalarnega produkta.
- ▶ Reševanje linearnega sistema.
- ▶ :

Drugo poglavje:

Linearni sistemi

$$Ax = b$$

- ▶ Direktne metode za reševanje
 - ▶ LU razcep
 - ▶ Pivotna rast $\rho(A)$
- ▶ Iterativne metode za reševanje
 - ▶ Jacobi, Gauss-Seidel, SOR, SSOR, konjugirani gradienti

Direktne metode

$$Ax = b$$

- ▶ Gaussova-eliminacija
- ▶ LU razcep
- ▶ Pivotiranje
- ▶ Pivotna rast

Reševanje kvadratnih linearnih sistemov

Linearni sistem n enačb z n neznankami x_1, \dots, x_n je oblike

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2, \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= b_n, \end{aligned}$$

kjer so a_{ij}, b_j realna števila.

V matrični obliki ga zapišemo kot

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \dots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}}_A \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_x = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}}_b.$$

Geometrijski pomen sistema $Ax = b$

Naj bodo $a_{(1)}, a_{(2)}, \dots, a_{(n)}$ stolpci matrike A , tj.,

$$a_{(i)} := \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix} \in \mathbb{R}^n$$

Linearna kombinacija vektorjev $a_{(1)}, a_{(2)}, \dots, a_{(n)}$ je vsak vektor oblike

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix}, \quad (1)$$

kjer so $x_i \in \mathbb{R}$ realna števila.

Zanima nas, ali obstaja linearna kombinacija (1), ki je enaka vektorju b .

Sistem $Ax = b$ z vidika numerične matematike

- ▶ Kako **drago** je reševanje sistema $Ax = b$?
cena=število osnovnih računskih operacij (+, -, ·, :).
- ▶ Kateri **problemi** in **napake** se pojavijo med reševanjem $Ax = b$?
Ali obstajajo slabe matrike? Kako take matrike identificirati?
- ▶ Za katere matrike se da **enostavno** in **poceni** rešiti tak sistem?

Ponovitev Gaussove eliminacije (GE)

Cilj je pretvoriti sistem v zgornjetrikotnega, nato pa ga rešiti z obratno substitucijo.

Primer

Rešujemo $Ax = b$, kjer sta

$$A = \begin{bmatrix} -3 & 2 & -1 \\ 6 & -6 & 7 \\ 3 & -4 & 4 \end{bmatrix}, \quad b = \begin{bmatrix} -1 \\ -7 \\ -6 \end{bmatrix}.$$

Tvorimo *razširjen sistem*

$$\tilde{A} = [A \mid b] = \left[\begin{array}{ccc|c} -3 & 2 & -1 & -1 \\ 6 & -6 & 7 & -7 \\ 3 & -4 & 4 & -6 \end{array} \right]$$

Prištejemo 2-kratnik prve vrstice drugi in 1-kratnik prve vrstice tretji.

$$\tilde{A}_{(1)} = \left[\begin{array}{ccc|c} -3 & 2 & -1 & -1 \\ 0 & -2 & 5 & -9 \\ 0 & -2 & 3 & -7 \end{array} \right]$$

Primer

Odštejemo 1-kratnik druge vrstice od tretje

$$\tilde{A}_{(2)} = \left[\begin{array}{ccc|c} -3 & 2 & -1 & -1 \\ 0 & -2 & 5 & -9 \\ 0 & 0 & -2 & 2 \end{array} \right]$$

Rešimo z *obratno substitucijo*

$$x_3 = \frac{2}{-2} = -1,$$

$$x_2 = \frac{1}{-2} (-9 - 5x_3) = 2,$$

$$x_1 = \frac{1}{-3} (-1 - 2x_2 + x_3) = 2.$$

V nadaljevanju bomo:

1. Prešteli število potrebnih računskih operacij za Gaussovo eliminacijo (GE).
2. GE bomo zapisali s pomočjo matričnih množenj.
3. Ukvarjali se bomo s stabilnostjo GE.

Algoritem GE in cena GE

```
1   $-n \times n$  matrika  $A = [a_{ij}]_{ij}$  in  $n \times 1$  vektor  $b = [b_i]_i$ 
2  -preoblikujemo  $[A|b]$  v zgornjetrikotno z GE
3
4  for  $k = 1 \dots n - 1$ 
5      for  $i = k + 1 \dots n$ 
6           $xmult = a_{ik} / a_{kk}$ 
7           $a_{ik} = 0$ 
8          for  $j = k + 1 \dots n$ 
9               $a_{ij} = a_{ij} - (xmult) a_{kj}$ 
10             end
11              $b_i = b_i - (xmult) b_k$ 
12         end
13     end
```

Izrek

Število računskih operacij (+, -, ·, :) za prevedbo matrike A in razširjene matrike $[A|b]$ v zgornjetrikotno obliko je

$$\frac{2}{3}n^3 + \mathcal{O}(n^2).$$