Numerička analiza

6. predavanje

Autor: Nela Bosner

Predavač: Nela Bosner

nela@math.hr

web.math.hr/~nela/nad.html

PMF – Matematički odjel, Zagreb

Sadržaj predavanja

- Numerička analiza Gaussovih eliminacija:
  - Pomoćni rezultati za grešku unazad određenih izraza (supstitucija unaprijed i unazad).
  - Greška unazad po komponentama LR faktorizacije.
  - Greška unazad po komponentama Gaussovih eliminacija.
  - Greška unazad po normi Gaussovih eliminacija s parcijalnim pivotiranjem.
- Primjer nestabilnosti Gaussovih eliminacija s parcijalnim pivotiranjem:
- Rješavanje linearnih sustava: iterativne metode:
  - Uvod u iterativne metode.

Sadržaj predavanja (nastavak)

- Standardne iteracije, i nužan i dovoljan uvjet konvergencije iteracija.
- Jacobijeva metoda i njena konvergencija.
- Gauss-Seidelova metoda i njena konvergencija.
- JOR metoda i njena konvergencija.
- SOR metoda i njena konvergencija.

- Primjer konvergencije iterativnih metoda

Numerička analiza Gaussovih eliminacija

Pomoćni rezultati

Prvo ćemo započeti sa nekoliko pomoćnih rezultata.

Lema 1. Ako imamo $|\delta_i| \leq u$ (u je jedinična greška zaokruživanja) i $\rho_i = \pm 1$ za $i = 1, \ldots, n$, pri čemu je $nu < 1$, tada vrijedi

$$\prod_{i=1}^{n} (1 + \delta_i)^{\rho_i} = 1 + \theta_n,$$

uz ocjenu

$$|\theta_n| \leq \frac{nu}{1 - nu} =: \gamma_n.$$

Dokaz. Dokaz se provodi matematičkom indukcijom.

Pomoćni rezultati (nastavak)

Baza indukcije: Za $n = 1$ vrijedi da je

- ako je $\rho_1 = 1$

$$\prod_{i=1}^{1} (1+\delta_i)^{\rho_i} = 1 + \delta_1, \quad |\delta_1| \leq u \leq \frac{u}{1-u},$$

- ako je $\rho_1 = -1$

$$\prod_{i=1}^{1} (1+\delta_i)^{\rho_i} = \frac{1}{1+\delta_1} = 1 + \theta_1,$$

  odakle je

$$\theta_1 = \frac{1}{1+\delta_1} - 1 = \frac{-\delta_1}{1+\delta_1}, \quad |\theta_1| \leq \frac{u}{1-u}.$$

Pomoćni rezultati (nastavak)

Korak indukcije: Pretpostavimo da tvrdnja leme vrijedi za $n-1$.

- Ako je $\rho_n = 1$, tada

$$\prod_{i=1}^{n} (1+\delta_i)^{\rho_i} = (1+\theta_{n-1})(1+\delta_n) = 1+\theta_n,$$

odakle je

$$\theta_n = \delta_n + (1+\delta_n)\theta_{n-1},$$

i vrijedi

$$|\theta_n| \leq u + (1+u)\frac{(n-1)u}{1-(n-1)u}$$

Pomoćni rezultati (nastavak)

$$|\theta_n| = \frac{u(1-(n-1)u) + (1+u)(n-1)u}{1-(n-1)u}$$
$$= \frac{nu}{1-(n-1)u} \leq \frac{nu}{1-nu} = \gamma_n.$$

- Ako je $\rho_n = -1$, tada

$$\prod_{i=1}^{n} (1+\delta_i)^{\rho_i} = \frac{1+\theta_{n-1}}{1+\delta_n} = 1+\theta_n,$$

odakle je

$$\theta_n = \frac{\theta_{n-1} - \delta_n}{1+\delta_n},$$

Pomoćni rezultati (nastavak)

i vrijedi

$$
\begin{aligned}
|\theta_n| &\le \frac{\frac{(n-1)u}{1-(n-1)u} + u}{1-u} \\
&= \frac{(n-1)u + u - (n-1)u^2}{(1-u)(1-(n-1)u)} \\
&= \frac{nu - (n-1)u^2}{1 - u - (n-1)u + (n-1)u^2} \le \frac{nu}{1-nu} = \gamma_n.
\end{aligned}
$$

■

Pomoćni rezultati (nastavak)

Lema 2. Pretpostavimo da se $y = (c - \sum_{i=1}^{k-1} a_i b_i)/b_k$ računa u aritmetici pomičnog zareza na sljedeći način

$$
\begin{aligned}
& s = c \\
& \text{za } i = 1, \ldots, k-1 \\
& \qquad s = s - a_i b_i \\
& \text{kraj} \\
& y = s/b_k
\end{aligned}
$$

Tada izračunati $\hat{y}$ zadovoljava

$$b_k \hat{y}(1 + \theta_k) = c - \sum_{i=1}^{k-1} a_i b_i (1 + \theta_i),$$

gdje su $|\theta_i| \le \gamma_i$.

Pomoćni rezultati (nastavak)

Dokaz. Analizirajmo prvo kako se računa $\hat{s} = f\ell(c - \sum_{i=1}^{k-1} a_i b_i)$ korak po korak.

- Za $i = 1$ imamo

$$
\begin{aligned}
\hat{s}_1 &= f\ell(c - a_1 b_1) = (c - a_1 b_1 (1 + \epsilon_1))(1 + \delta_1) \\
&= c(1 + \delta_1) - a_1 b_1 (1 + \epsilon_1)(1 + \delta_1),
\end{aligned}
$$

  gdje je $|\epsilon_1| \leq u$ greška množenja, a $|\delta_1| \leq u$ greška zbrajanja.

- Za $i = 2$ imamo

$$
\begin{aligned}
\hat{s}_2 &= f\ell(\hat{s}_1 - a_2 b_2) = (\hat{s}_1 - a_2 b_2 (1 + \epsilon_2))(1 + \delta_2) \\
&= (c(1 + \delta_1) - a_1 b_1 (1 + \epsilon_1)(1 + \delta_1) - a_2 b_2 (1 + \epsilon_2))(1 + \delta_2)
\end{aligned}
$$

Pomoćni rezultati (nastavak)

$$\hat{s}_2 = c(1+\delta_1)(1+\delta_2) - a_1 b_1 (1+\epsilon_1)(1+\delta_1)(1+\delta_2) - \\ - a_2 b_2 (1+\epsilon_2)(1+\delta_2),$$

gdje je $|\epsilon_2| \leq u$ greška množenja, a $|\delta_2| \leq u$ greška zbrajanja.

- ... i tako nastavimo dalje. Možemo zaključiti da je

$$\hat{s} = \hat{s}_{k-1} = c(1+\delta_1)\cdots(1+\delta_{k-1}) - \\ - \sum_{i=1}^{k-1} a_i b_i (1+\epsilon_i)(1+\delta_i)\cdots(1+\delta_{k-1}),$$

gdje su $|\epsilon_i|,\ |\delta_i| \leq u$.

Pomoćni rezultati (nastavak)

- Preostaje nam još samo zadnje dijeljenje

$$\hat{y} = f\ell(\hat{s}/b_k) = \frac{\hat{s}(1+\epsilon_k)}{b_k}, \quad |\epsilon_k| \le u,$$

odakle slijedi dijeljenjem sa $(1+\delta_1)\cdots(1+\delta_{k-1})(1+\epsilon_k)$

$$\begin{aligned} b_k\hat{y}\,&\frac{1}{(1+\delta_1)\cdots(1+\delta_{k-1})(1+\epsilon_k)} \\ &= c - \sum_{i=1}^{k-1} a_i b_i \frac{1+\epsilon_i}{(1+\delta_1)\cdots(1+\delta_{i-1})}. \end{aligned}$$

Rezultat slijedi primjenom Leme 1.

Pomoćni rezultati (nastavak)

Lema 3. Ako se $y = (c - \sum_{i=1}^{k-1} a_i b_i)/b_k$ računa u aritmetici pomičnog zareza bez obzira na poredak operacija, tada je

$$b_k \hat{y}(1 + \theta_k^{(0)}) = c - \sum_{i=1}^{k-1} a_i b_i (1 + \theta_k^{(i)}),$$

gdje su $|\theta_k^{(i)}| \leq \gamma_k$ za sve i. Ako je $b_k = 1$, u tom slučaju nema posljednjeg dijeljenja i vrijedi $|\theta_k^{(i)}| \leq \gamma_{k-1}$ za sve i.

Glavni rezultati

Pomoćni rezultat izražen u Lemi 3 koristit ćemo za analizu grešaka zaokruživanja kod računanja pojedinih elemenata matrica L i R iz LR faktorizacije

- Analizirajmo prvo računanje elementa $r_{ij} = a_{ij}^{(i)}$, za $i \leq j$, matrice R u egzaktnoj aritmetici.

$$
\begin{aligned}
a_{ij}^{(1)} &= a_{ij} \\
a_{ij}^{(2)} &= a_{ij}^{(1)} - m_{i1} a_{1j}^{(1)} = a_{ij} - \ell_{i1} r_{1j} \\
a_{ij}^{(3)} &= a_{ij}^{(2)} - m_{i2} a_{2j}^{(2)} = a_{ij} - \ell_{i1} r_{1j} - \ell_{i2} r_{2j} \\
&\vdots \\
r_{ij} &= a_{ij}^{(i)} = a_{ij}^{(i-1)} - m_{i,i-1} a_{i-1,j}^{(i-1)} \\
&= a_{ij} - \ell_{i1} r_{1j} - \cdots - \ell_{i,i-1} r_{i-1,j}
\end{aligned}
$$

Glavni rezultati (nastavak)

- Za elemente izračunate matrice $\hat{R}$ prema Lemi 3 slijedi:

$$\hat{r}_{ij}(1+\theta_i^{(i)}) = a_{ij} - \sum_{k=1}^{i-1} \hat{\ell}_{ik}\hat{r}_{kj}(1+\theta_i^{(k)}),$$

pri čemu su $|\theta_i^{(k)}| \leq \gamma_{i-1} \leq \gamma_i$ za sve k.

- Analizirajmo sada računanje elementa $l_{ij} = m_{ij}$, za $i > j$, matrice L u egzaktnoj aritmetici. Koristit ćemo i prethodnu analizu za $a_{ij}^{(s)}$.

$$\ell_{ij} = m_{ij} = \frac{a_{ij}^{(j)}}{a_{jj}^{(j)}} = \frac{a_{ij} - \ell_{i1}r_{1j} - \cdots - \ell_{i,j-1}r_{j-1,j}}{r_{jj}}$$

Glavni rezultati (nastavak)

- Za elemente izračunate matrice $\hat{L}$ prema Lemi 3 slijedi:

$$\hat{\ell}_{ij}\hat{r_{jj}}(1+\Theta_j^{(j)}) = a_{ij} - \sum_{k=1}^{j-1} \hat{\ell}_{ik}\hat{r}_{kj}(1+\Theta_j^{(k)}),$$

  pri čemu su $|\Theta_j^{(k)}| \leq \gamma_j$ za sve k.

- Ovi rezultati čine okosnicu greške unazad za LR faktorizaciju.

Glavni rezultati (nastavak)

Teorem 1. Ako se Gaussove eliminacije primijenjene na $A \in \mathbb{R}^{n\times n}$ izvedu do kraja, tada izračunati LR faktori $\hat{L}, \hat{R} \in \mathbb{R}^{n\times n}$ zadovoljavaju

$$\hat{L}\hat{R} = A + \Delta A, \qquad |\Delta A| \le \gamma_n |\hat{L}||\hat{R}|.$$

Dokaz. Iz prethodne analize izračunatih elemenata $\hat{r}_{ij}$ i $\hat{\ell}_{ij}$ možemo zaključiti:

- za $i \le j$, uz izjednačavanje $\hat{\ell}_{ii} = 1$, imamo

$$a_{ij} = \left(\sum_{k=1}^{i} \hat{\ell}_{ik}\hat{r}_{kj} \right) + \sum_{k=1}^{i} \hat{\ell}_{ik}\hat{r}_{kj}\theta_i^{(k)} = (\hat{L}\hat{R})_{ij} + \sum_{k=1}^{i} \hat{\ell}_{ik}\hat{r}_{kj}\theta_i^{(k)}$$

Glavni rezultati (nastavak)

- za $i > j$ imamo

$$a_{ij} = \left(\sum_{k=1}^{j} \hat{\ell}_{ik}\hat{r}_{kj}\right) + \sum_{k=1}^{j} \hat{\ell}_{ik}\hat{r}_{kj}\Theta_j^{(k)} = (\hat{L}\hat{R})_{ij} + \sum_{k=1}^{j} \hat{\ell}_{ik}\hat{r}_{kj}\Theta_j^{(k)}$$

- Dakle, za $\Delta A = A - \hat{L}\hat{R}$ vrijedi

$$|(\Delta A)_{ij}| = \begin{cases} \left|\sum_{k=1}^{i} \hat{\ell}_{ik}\hat{r}_{kj}\theta_i^{(k)}\right| \leq \gamma_i \sum_{k=1}^{i} |\hat{\ell}_{ik}||\hat{r}_{kj}| \leq \gamma_n(|\hat{L}||\hat{R}|)_{ij}, & \text{za } i \leq j \\ \left|\sum_{k=1}^{j} \hat{\ell}_{ik}\hat{r}_{kj}\Theta_j^{(k)}\right| \leq \gamma_j \sum_{k=1}^{j} |\hat{\ell}_{ik}||\hat{r}_{kj}| \leq \gamma_n(|\hat{L}||\hat{R}|)_{ij}, & \text{za } i > j \end{cases}$$

$\blacksquare$

Glavni rezultati (nastavak)

Teorem 2. Neka je $A \in \mathbb{R}^{n\times n}$ i pretpostavimo da smo Gaussovim eliminacijama dobili izračunate LR faktore $\hat{L}$ i $\hat{R}$, i izračunato rješenje $\hat{x}$ sustava $Ax = b$. Tada je

$$(A + \Delta A)\hat{x} = b, \qquad |\Delta A| \leq \gamma_{3n}|\hat{L}||\hat{R}|.$$

Dokaz. Prema Teoremu 1 je

$$\hat{L}\hat{R} = A + \Delta A_1, \qquad |\Delta A_1| \leq \gamma_n|\hat{L}||\hat{R}|.$$

Za rješavanje trokutastih sustava koristimo

$$\begin{aligned} Ly &= b, \qquad && \text{supstitucija unaprijed,} \\ Rx &= y, \qquad && \text{povratna supstitucija,} \end{aligned}$$

Glavni rezultati (nastavak)

pri čemu se svaki element rješenja trokutastog sustava računa

$$y_i = b_i - \sum_{j=1}^{i-1} \ell_{ij}\, y_j, \quad (y_1 = b_1),\ i = 2, \ldots, n,$$

$$x_i = \frac{1}{r_{ii}} \left(y_i - \sum_{j=i+1}^{n} r_{ij}\, x_j \right), \quad \left(x_n = \frac{y_n}{r_{nn}} \right)\ i = n-1, \ldots, 1.$$

Prema Lemi 3 za izračunate vrijednsorti vrijedi

$$\hat{y}_i(1 + \theta_i^{(i)}) = b_i - \sum_{j=1}^{i-1} \hat{\ell}_{ij}\, \hat{y}_j(1 + \theta_i^{(j)}), \quad , \ {\scriptstyle i=2,\ldots,n},$$

$$\hat{r}_{ii}\hat{x}_i(1 + \Theta_{n-i+1}^{(i)}) = \hat{y}_i - \sum_{j=i+1}^{n} \hat{r}_{ij}\, \hat{x}_j(1 + \Theta_{n-i+1}^{(j)}), \quad {\scriptstyle i=n-1,\ldots,1},$$

Glavni rezultati (nastavak)

pri čemu je $\theta_i^{(j)} \leq \gamma_{i-1} \leq \gamma_i$, a $\Theta_{n-i+1}^{(j)} \leq \gamma_{n-i+1}$ za svaki j. Zato izračunata rješenja $\hat{y}$ i $\hat{x}$ zadovoljavaju

$$
\begin{aligned}
(\hat{L} + \Delta L)\hat{y} &= b, & |\Delta L| &\leq \gamma_n |\hat{L}|, \\
(\hat{R} + \Delta R)\hat{x} &= \hat{y}, & |\delta R| &\leq \gamma_n |\hat{U}|.
\end{aligned}
$$

jer su $(\Delta L)_{ij} = \hat{\ell}_{ij}\theta_i^{(j)}$ i $(\Delta R)_{ij} = \hat{r}_{ij}\Theta_{n-i+1}^{(j)}$. Prema tome dobivamo

$$
\begin{aligned}
b &= (\hat{L} + \Delta L)(\hat{R} + \Delta R)\hat{x} \\
&= (A + \Delta A_1 + \hat{L}\Delta R + \Delta L\hat{R} + \Delta L\Delta R)\hat{x} \\
&= (A + \Delta A)\hat{x},
\end{aligned}
$$

Glavni rezultati (nastavak)

gdje je

$$\begin{aligned}|\Delta A| &= |\Delta A_1 + \hat{L}\Delta R + \Delta L\hat{R} + \Delta L\Delta R| \\ &\le |\Delta A_1| + |\hat{L}||\Delta R| + |\Delta L||\hat{R}| + |\Delta L||\Delta R| \\ &\le (3\gamma_n + \gamma_n^2)|\hat{L}||\hat{R}|.\end{aligned}$$

Uz uvijet $nu < 1$ dobivamo konačnu ocjenu:

$$\begin{aligned}3\gamma_n + \gamma_n^2 &= \frac{3nu}{1-nu} + \frac{n^2u^2}{(1-nu)^2} = \frac{3nu - 3n^2u^2 + n^2u^2}{(1-nu)^2} \\ &= \frac{3nu - 2n^2u^2}{1 - 2nu + n^2u^2} \le \frac{3nu}{1-2nu} \le \gamma_{3n}.\end{aligned}$$

$\blacksquare$

Interpretacija rezultata

Idealno, mi bi željeli dobiti rezultat oblika

$$|\Delta A| \leq \mathcal{O}(n) u |A|,$$

što predstavlja perturbaciju elemenata matrice A s malom relativnom greškom. Takva ograda greške postoji ako $\hat{L}$ i $\hat{R}$ zadovoljavaju

$$|\hat{L}||\hat{R}| = |\hat{L}\hat{R}|.$$

To vrijedi npr. za nenegativne matrice (po elementima). U tom slučaju imamo

$$|\hat{L}||\hat{R}| = |\hat{L}\hat{R}| = |A + \Delta A_1| \leq |A| + \gamma_n |\hat{L}||\hat{R}|,$$

odakle slijedi

Interpretacija rezultata (nastavak)

$$|\hat{L}||\hat{R}| \leq \frac{1}{1-\gamma_n}|A|.$$

Ako to ubacimo u rezultat Teorema 2 dobit ćemo

$$(A + \Delta A)\hat{x} = b, \quad |\Delta A| \leq \frac{\gamma_{3n}}{1-\gamma_n}|A|, \quad (\hat{L}, \hat{R} > 0).$$

To znači da $\hat{x}$ ima malu relativnu grešku u nazad po komponentama.

Jedna klasa matrica koja ima nenegativne LR faktore se definira na sljedeći način:

- Matrica $A \in \mathbb{R}^{n \times n}$ je totalno pozitivna (nenegativna) ako je determinanta svake njene kvadratne podmatrice pozitivna (nenegativna).

Interpretacija rezultata (nastavak)

Jedna važna činjenica koja slijedi iz Teorema 1 i Teorema 2 je ta da je stabilnost Gaussovih eliminacija određena

- ne veličinom multiplikatora,
- već veličinom elemenata matrice $|\hat{L}||\hat{R}|$.

Ta matrica može imati male elemente kada su multiplikatori $\hat{\ell}_{ij}$ veliki, i obratno.

Za bolje razumijevanje stabilnosti Gaussovih eleiminacija potrebno je gledati ocjene po normi.

- Za Gaussove eliminacije bez pivotiranja, omjer $\| \, |\hat{L}||\hat{R}| \, \| / \|A\|$ može biti proizvoljno velik.
- Npr. za matricu $\left[\begin{smallmatrix} \epsilon & 1 \\ 1 & 1 \end{smallmatrix}\right]$ taj omjer je reda veličine ϵ^{-1}.

Interpretacija rezultata (nastavak)

- Ako se koristi parcijalno pivotiranje, tada je $|\ell_{ij}| \leq 1$ za sve $i \geq j$. Već smo prije spomenuli da je u tom slučaju

  $$|r_{ij}| \leq 2^{i-1} \max_{k \leq i} |a_{kj}|,$$

  pa prema tome za parcijalno pivotiranje L ima male elemente, a elementi od R su relativno ograničeni elementima od A.

Sada ponovo u igru ulazi pivotni rast:

$$\rho_n = \frac{\max_{i,j,k} |a_{ij}^{(k)}|}{\max_{i,j} |a_{ij}|}.$$

Ako ubacimo ogradu $|r_{ij}| = |a_{ij}^{(i)}| \leq \rho_n \max_{i,j} |a_{ij}|$ u rezultat Teorema 2, dobivamo sljedeći klasični teorem.

Glavni rezultati (nastavak)

Teorem 3. (Wilkinson). Neka je $A \in \mathbb{R}^{n\times n}$ i pretpostavimo da Gaussove eliminacije s parcijalnim pivotiranjem daju izračunato rješenje $\hat{x}$ sustava $Ax = b$. Tada je

$$(A + \Delta A)\hat{x} = b, \qquad \|\Delta A\|_\infty \leq n^2 \gamma_{3n} \rho_n \|A\|_\infty.$$

Dokaz. Računamo normu greške iz Teorema 2:

$$\begin{aligned}
\|\Delta A\|_\infty &= \| \, |\Delta A| \, \|_\infty \leq \gamma_{3n} \| \, |\hat{L}||\hat{R}| \, \|_\infty \\
&= \gamma_{3n} \max_{i=1,\dots,n} \sum_{j=1}^{n} (|\hat{L}||\hat{R}|)_{ij} = \gamma_{3n} \max_{i=1,\dots,n} \sum_{j=1}^{n} \sum_{k=1}^{\min\{i,j\}} |\hat{\ell}_{ik}||\hat{r}_{kj}| \\
&\leq \gamma_{3n} \max_{i=1,\dots,n} \sum_{j=1}^{n} \sum_{k=1}^{\min\{i,j\}} 1 \cdot \rho_n \max_{s,t} |a_{st}|
\end{aligned}$$

Glavni rezultati (nastavak)

$$
\begin{aligned}
\|\Delta A\|_\infty &\le \gamma_{3n}\rho_n \|A\|_\infty \max_{i=1,\ldots,n} \sum_{j=1}^{n} \min\{i,j\} \\
&\le \gamma_{3n}\rho_n \|A\|_\infty \max_{i=1,\ldots,n} ni \\
&\le n^2 \gamma_{3n}\rho_n \|A\|_\infty .
\end{aligned}
$$

■

Napomena. U dokazu ovog teorema korišten je mali ilegalni manevar:

- koristili smo ograde za elemente od $\hat{L}$ i $\hat{R}$ koje striktno vrijede za elemente egzaktnih matrica L i R.

Glavni rezultati (nastavak)

- Umjesto toga smo mogli redefinirati pivotni rast preko izračunatih elemenata $\hat{a}_{ij}^{(k)}$, ali tada bi svaka ograda za pivotni rast uključivala i jediničnu grešku zaokruživanja.
- Slično, mi možemo samo garantirati $|\hat{\ell}_{ij}| \leq 1 + \mathcal{O}(u)$.
- Ipak ovaj propust u korektnosti dokaza je bezopasan za naše potrebe.

Sličan rezultat može se postići i za Gaussove eliminacije bez pivotiranja, samo sa drugačijom konstantom.

- Povratna stabilnost po normi Gaussovih eliminacija sa ili bez pivotiranja ovisi o pivotnom rastu!

Primjer nestabilnosti Gaussovih eliminacija s parcijalnim pivotiranjem

Primjer rubnog problema

- Numerička analiza Gaussovih eliminacija sa parcijalnim pivotiranjem nam je potvrdila da stabilnost eliminacija ovisi o pivotnom rastu.
- Već od prije znamo da je pivotni rast za parcijalno pivotiranje manji ili jednak 2^{n-1}.
- Isto tako znamo da se ta gornja ograda pivotnog rasta od 2^{n-1} postiže za "umjetno" konstruiran Wilkinsonov primjer.
- U ovom primjeru promotrit ćemo rubni problem, kod kojeg standardna numerička metoda korištena za njegovo rješavanje proizvodi matricu sa eksponencijalnim pivotnim rastom.

Primjer rubnog problema (nastavak)

Promotrimo općeniti rubni problem u dvije različite točke:

$$y' = M(t)y + q(t), \quad B_a y(a) + B_b y(b) = \beta, \quad y(t) \in \mathbb{R}^n,$$

i jedan konkretni problem definiran sljedećim parametrima

$$n = 2, \quad M(t) \equiv \begin{bmatrix} -\frac{1}{6} & 1 \\ 1 & -\frac{1}{6} \end{bmatrix}, \quad a = 0,\ b = 60, \quad B_a = B_b = I.$$

Vrijednosti $q(t)$ i β nisu važne za naše razmatranje.

- Ovaj problem je dobro uvjetovan u smislu da rješenje y je neosjetljivo na perturbacije u ulaznim podacima $M(t)$, $q(t)$, B_a, B_b i β.

Numeričko rješavanje rubnog problema

Standardan algoritam za rješavanje problema ovog oblika je višestruka metoda gađanja.

- Najprije se segment $[a, b]$ razdijeli na N podsegmenata, definiranjem mreže

$$a = t_0 < t_1 < \cdots < t_N = b$$

pri čemu su

$$t_i = a + ih, \qquad i = 0, 1, \ldots, N, \quad h = \frac{b-a}{N}.$$

- Na svakom podsegmentu $[t_i, t_{i+1}]$, potrebno je riješiti inicijalne probleme

$$Y_i' = M(t)Y_i, \qquad Y_i(t_i) = I,$$

$$y_{pi} = M(t)y_{pi} + q(t), \qquad y_{pi}(t_i) = 0.$$

Numeričko rješavanje rubnog problema (nast.)

- Ako sa s_i definiramo vrijednost pravog rješenja y u t_i, tada možemo provjeriti da je

$$y(t) = Y_i(t)s_i + y_{pi}(t), \qquad t \in [t_i, t_{i+1}],\ i = 0, \ldots, N-1.$$

- Budući da $y(t)$ mora biti neprekidna u točkama mreže, dobivamo

$$Y(t_{i+1})s_i + y_{pi}(t_{i+1}) = s_{i+1}, \qquad i = 0, \ldots, N-1.$$

- Štoviše, zbog rubnih uvjeta vrijedi

$$B_a s_0 + B_b s_N = \beta.$$

- Zadnje dvije jednakosti definiraju sustav linearnih jednadžbi čije je rješenje $s_0, \ldots, s_N$.

Numeričko rješavanje rubnog problema (nast.)

Matrica sustava ima opći oblik

$$A = \begin{bmatrix} B_a & & & & B_b \\ -Y_0(t_1) & I & & & 0 \\ & -Y_1(t_2) & I & & \vdots \\ & & \ddots & \ddots & 0 \\ & & & -Y_{N-1}(t_N) & I \end{bmatrix}.$$

Za naš konkretan primjer imamo

$$Y_i(t_{i+1}) = e^{hM}$$

Numeričko rješavanje rubnog problema (nast.)

U tom slučaju matrica sustava ima konkretan oblik

$$A = \begin{bmatrix} I & & & & I \\ -e^{hM} & I & & & 0 \\ & -e^{hM} & I & & \vdots \\ & & \ddots & \ddots & 0 \\ & & & -e^{hM} & I \end{bmatrix}.$$

Pretpostavimo da je h dovoljno mali, tako da su svi elementi od e^{hM} po modulu ≤ 1. To je moguće jer

$$e^{hM} = I + hM + \mathcal{O}(h^2) \approx \begin{bmatrix} 1 - h/6 & h \\ h & 1 - h/6 \end{bmatrix}.$$

Numeričko rješavanje rubnog problema (nast.)

Ako se Gaussove eliminacije s parcijalnim pivotiranjem primijene na matricu A, tada neće doći niti do jedne izmjene redaka (nema pivotiranja), i dobit ćemo sljedeću LR faktorizaciju

$$
A = \begin{bmatrix} I & & & & \\ -e^{hM} & I & & & 0 \\ & -e^{hM} & I & & \vdots \\ & & \ddots & \ddots & 0 \\ & & & -e^{hM} & \bar{L} \end{bmatrix} \begin{bmatrix} I & & & & & I \\ & I & & & & e^{hM} \\ & & I & & & e^{2hM} \\ & & & \ddots & & \vdots \\ & & & & I & e^{(60-h)M} \\ & & & & & \bar{R} \end{bmatrix},
$$

gdje je $\bar{L}\bar{R} = (I + e^{60M})$, $\bar{L}$ je donje trokutasta a $\bar{R}$ je gornje trokutasta.

Numeričko rješavanje rubnog problema (nast.)

U prethodnoj faktorizaciji vidljivo je da elementi u posljednjem stupcu matrice R eksponencijalno rastu.

- Za konkretan izbor $N = 200$ tj. $h = 0.3$ imamo zadovoljen uvjet da su svi elementi od e^{hM} po modulu ≤ 1:

$$e^{0.3M} = \begin{bmatrix} 0.9944 & 0.2897 \\ 0.2897 & 0.9944 \end{bmatrix}$$

- Najveći element u matrici R je $\approx 2.59 \cdot 10^{21}$.
- Detalje pogledajte u MATLAB simulaciji.

Potpuno pivotiranje, s druge strane, daje stabilnu faktorizaciju: najveći element u izračunatom faktoru R je 10.665545.

Rješavanje linearnih sustava: iterativne metode

Uvod u iterativne metode

- Vidjeli smo da rješenje linearnog sustava $Ax = b$ općenito ne možemo izračunati apsolutno točno.
- Također, u nekim primjenama niti točno rješenje $x = A^{-1}b$ nije puno bolje od neke dovoljno dobre aproksimacije $\hat{x}$, gdje obično $\hat{x}$ zadovoljava sustav $(A + \delta A)\hat{x} = b$, blizak polaznom.
- Nadalje, Gaussove eliminacije, koje su jednostavno konačan niz formula koje vode rješenju sustava, ne garantiraju idealnu točnost.
- Osim toga, u praksi moramo biti svjesni da je računalo omeđeno ne samo u pitanju numeričke točnosti nego i u još dva važna aspekta:
  - raspoloživom memorijskom prostoru,
  - vremenu izvršavanja.

Uvod u iterativne metode (nastavak)

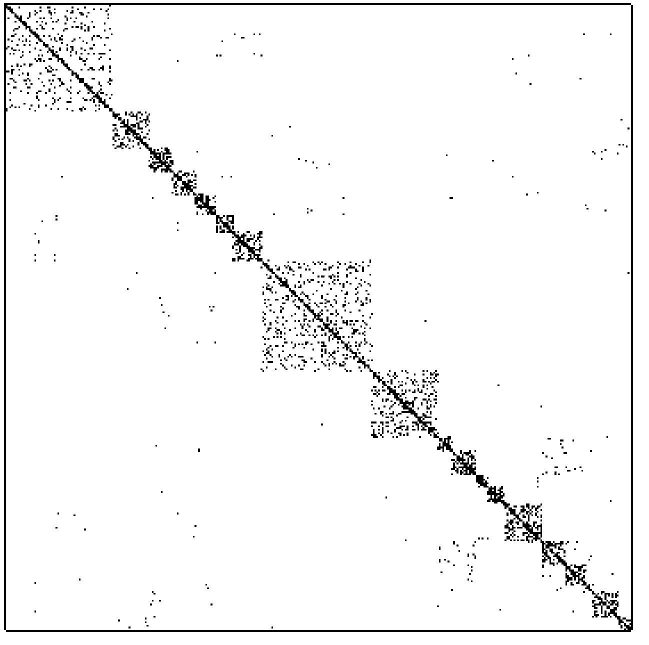
- Moderne primjene matematike zahtijevaju rješenja sustava velikih dimenzija, npr. $n > 10^5$.
- Lako se uvjeriti da je u takvim primjerima proces Gaussovih eliminacija često praktično neupotrebljiv.
- Jer, matrica dimenzije $n = 10^5$ zahtijeva $n^2 = 10^{10}$ lokacija u memoriji, svaka barem 4 bajta (veličina reprezentacije realnog broja u jednostrukoj preciznosti) — samo spremanje matrice koeficijenata u memoriju računala predstavalja poteškoću.
- Ponekad matricu držimo na vanjskoj, sporoj memoriji (datoteka na disku) i onda možemo dijelove matrice učitavati dio po dio u radnu memoriju.
- Da ne govorimo o vremenu izvršavanja Gaussovih eliminacija, čija je složenost $\mathcal{O}(n^3)$!

Uvod u iterativne metode (nastavak)

- U puno važnih primjena matrica sustava A je velike dimenzije, ali je rijetko popunjena. To znači da je velika većina elemenata od A jednaka nuli, a elementi koji nisu nula obično su
  - pravilno raspoređni po matrici
  - ili čak imaju i pravilno raspoređene numeričke vrijednosti.

Uvod u iterativne metode (nastavak)

Jedan primjer rijetko popunjene matrice dan je na slici.



- U svakom retku broj elemenata $\neq 0$ unaprijed je poznat i broj takvih elemenata je puno manji od dimenzije n.
- To znači da je npr. računanje produkta Av složenosti puno manje od $2n^2 - n$, što je potrebno za punu matricu.

Uvod u iterativne metode (nastavak)

- Ponekad je poznat način kako su generirani elementi matrice (npr. poznato da je $a_{ij} = f_{ij}(\cdots)$, gdje su $f_{ij}(\cdots)$ poznate funkcije nekih parametara) pa uvijek možemo generirati dijelove matrice.
- Moguće je i da je u konkretnoj aplikaciji jedino zadano kako A djeluje kao linearni operator – postoji potprogram koji za zadani v računa Av.
- Ako je to jedini način kako doći do matrice A, kako onda riješiti $Ax = b$?

Uvod u iterativne metode (nastavak)

Prethodna diskusija nas motivira da potražimo i drugačije pristupe za rješavanje linearnog sustava $Ax = b$.

- Primijetimo da ne moramo nužno težiti pronalaženju egzaktnog rješenja – umjesto toga želimo dovoljno dobru aproksimaciju $\hat{x}$.
- Zato ima smisla pokušati konstruirati niz $x^{(0)}, x^{(1)}, \ldots, x^{(k)}, \ldots$ vektora iz $\mathbb{R}^n$ sa sljedećim svojstvima:
  - za svaki k formula za računanje $x^{(k)}$ je jednostavna;
  - $x^{(k)}$ teži prema $x = A^{-1}b$ i za neki k (obično $k \ll n$) je $x^{(k)}$ prihvatljiva aproksimacija za x.

Općenito o iterativnim metodama

Pretpostavimo da je A regularna matrica reda n.

- Iterativna metoda koja pronalazi približno rješenje sustava $Ax = b$ zadana je
  - početnim vektorom $x^{(0)}$
  - i načinom na koji se računa aproksimacija $x^{(k)}$ u k-tom koraku iteracije.
- U praksi se gotovo isključivo koriste iterativne metode prvog reda, koje iz jednog prethodnog vektora $x^{(k)}$ nalaze sljedeću aproksimaciju $x^{(k+1)}$.
- Kriterij zaustavljanja sličan je kao kod svih metoda "limes" tipa – tj. kad je $x^{(k)}$ dovoljno dobra aproksimacija za pravo rješenje x, stani!.
- Naravno, i tu postoji problem, jer pravi x ne znamo.

Općenito o iterativnim metodama (nastavak)

- Zbog toga, koristimo svojstvo da je konvergentan niz Cauchyjev, tj. da susjedni članovi niza moraju postati po volji bliski.
- Standardno, uzima se da su $x^{(k+1)}$ i $x^{(k)}$ dovoljno bliski ako je
$$\|x^{(k+1)} - x^{(k)}\| \leq \varepsilon,$$
gdje je ε neka unaprijed zadana točnost (reda veličine u, odnosno $n \cdot u$), a $\| \, \|$ neka vektorska norma.
- Drugi način za određivanje kriterija zaustavljanje je kada je norma reziduala dovoljno mala.
- Međutim, vidjeli smo da je taj kriterij nepouzdan za loše uvjetovane matrice.

Standardne iteracije

Iterativnu metodu pokušavamo definirati izrazom

$$x^{(k+1)} = Tx^{(k)} + c, \qquad k = 0, 1, 2, \ldots, \quad x^{(0)} \text{ zadan},$$

gdje je $T \in \mathbb{R}^{n \times n}$ matrica iteracije i $c \in \mathbb{R}^n$.

Da bismo ostvarili takvu iterativnu metodu, potrebno je pažljivo rastaviti matricu A.

Definicija. Rastav matrice A je par matrica (M, N) (obje reda n) za koje vrijedi

- $A = M - N$,
- M je regularna.

Obično se M bira tako da se može lako invertirati: dijagonalna ili trokutasta.

Standardne iteracije (nastavak)

Bilo koji rastav matrice A generira iterativnu metodu na sljedeći način:

$$Ax = Mx - Nx = b \implies Mx = Nx + b,$$

pa zbog regularnosti od M izlazi

$$x = M^{-1}Nx + M^{-1}b.$$

Ako označimo $T = M^{-1}N$, $c = M^{-1}b$ onda je prethodna relacija ekvivalentna s

$$x = Tx + c.$$

Time smo definirali našu iterativnu metodu

$$x^{(k+1)} = Tx^{(k)} + c, \quad k \in \mathbb{N}_0.$$

Konvergencija standardnih iteracija

Pravo rješenje x je fiksna točka iteracione funkcije

$$f(x) = Tx + c.$$

To znači da u analizi konvergencije možemo koristiti poznate teoreme o fiksnoj točki (poput Banachovog u potpunim prostorima).

Lema 1. Niz iteracija $\big(x^{(k)}\big)$, $k \in \mathbb{N}_0$, generiran relacijom $x^{(k+1)} = Tx^{(k)} + c, \ k \in \mathbb{N}_0$ konvergira prema rješenju linearnog sustava $Ax = b$ za sve početne vektore $x^{(0)}$ i sve desne strane b, ako je

$$\|T\| < 1,$$

pri čemu je $\|\ \|$ proizvoljna operatorska norma.

Konvergencija standardnih iteracija (nastavak)

Dokaz. Oduzmimo $x = Tx + c$ od $x^{(k+1)} = Tx^{(k)} + c$. Dobivamo

$$x^{(k+1)} - x = T\big(x^{(k)} - x\big),$$

pa uzimanjem norme dobivamo

$$\big\|x^{(k+1)} - x\big\| \leq \|T\| \, \big\|x^{(k)} - x\big\| \leq \cdots \leq \|T\|^{k+1} \, \big\|x^{(0)} - x\big\|.$$

No, zbog $\|T\| < 1$ slijedi $\|T\|^{k+1} \to 0$ za $k \to \infty$, odakle izlazi

$$\big\|x - x^{(k+1)}\big\| \to 0,$$

pa iz neprekidnosti norme slijedi

$$x^{(k+1)} \to x \quad \text{za svaki } x^{(0)}.$$

Konvergencija standardnih iteracija (nastavak)

No, može se dobiti i nešto bolji rezultat, korištenjem veze spektralnog radijusa i operatorske norme matrice.

Lema 2.

- Za sve operatorske norme $\|\ \|$ vrijedi

$$\rho(T) \leq \|T\|.$$

- Za sve T i sve $\varepsilon > 0$ postoji operatorska norma $\|\ \|_*$ takva da je

$$\|T\|_* \leq \rho(T) + \varepsilon.$$

Norma $\|\ \|_*$ ovisi i o T, i o ε.

Konvergencija standardnih iteracija (nastavak)

Dokaz. Da bismo dokazali da je $\rho(T) \leq \|T\|$ za bilo koju operatorsku normu, uzet ćemo $\rho(T) = |\lambda|$, i pripadni svojstveni vektor od λ, kojeg označavamo sa x. Tada vrijedi

$$\|T\| = \max_{y \neq 0} \frac{\|Ty\|}{\|y\|} \geq \frac{\|Tx\|}{\|x\|} = \frac{\|\lambda x\|}{\|x\|} = \frac{|\lambda| \, \|x\|}{\|x\|} = |\lambda| = \rho(T).$$

S druge strane, da bismo konstruirali operatorsku normu $\| \ \|_*$ takvu da je $\|T\|_* \leq \rho(T) + \varepsilon$, svedimo matricu T na njenu Jordanovu formu.

- Znamo da postoji regularna matrica S za koju je $J = S^{-1}TS$ Jordanova forma matrice T.
- Jedinice iznad dijagonale u J "kvare" spektralni radijus, pa je ideja da umjesto njih dobijemo zadani $\varepsilon > 0$.

Konvergencija standardnih iteracija (nastavak)

- To se postiže dijagonalnom matricom sličnosti. Neka je

$$D_\varepsilon = \operatorname{diag}(1, \varepsilon, \varepsilon^2, \ldots, \varepsilon^{n-1}).$$

Tada je $(SD_\varepsilon)^{-1} T (SD_\varepsilon) = D_\varepsilon^{-1} J D_\varepsilon =$

$$= \left[\begin{array}{cccc:cccc:c}
\lambda_1 & \varepsilon & & & & & & & \\
 & \ddots & \ddots & & & & & & \\
 & & \ddots & \varepsilon & & & & & \\
 & & & \lambda_1 & & & & & \\
\hdashline
 & & & & \lambda_2 & \varepsilon & & & \\
 & & & & & \ddots & \ddots & & \\
 & & & & & & \ddots & \varepsilon & \\
 & & & & & & & \lambda_2 & \\
\hdashline
 & & & & & & & & \ddots
\end{array}\right].$$

Konvergencija standardnih iteracija (nastavak)

Uzmimo kao vektorsku normu

$$\|x\|_* = \|(SD_\varepsilon)^{-1}x\|_\infty$$

(pokažite da je to vektorska norma!), koja generira operatorsku normu. U toj operatorskoj normi vrijedi

$$\begin{aligned}
\|T\|_* &= \max_{x\neq 0} \frac{\|Tx\|_*}{\|x\|_*} = \max_{x\neq 0} \frac{\|(SD_\varepsilon)^{-1}Tx\|_\infty}{\|(SD_\varepsilon)^{-1}x\|_\infty} \\
&= \max_{y\neq 0} \frac{\|(SD_\varepsilon)^{-1}T(SD_\varepsilon)y\|_\infty}{\|y\|_\infty} \\
&= \|(SD_\varepsilon)^{-1}T(SD_\varepsilon)\|_\infty \leq \max_i |\lambda_i| + \varepsilon = \rho(T) + \varepsilon.
\end{aligned}$$

Konvergencija standardnih iteracija (nastavak)

Konačno, prethodne dvije leme daju potpunu karakterizaciju konvergencije iterativnih metoda.

Teorem. Niz iteracija $\left(x^{(k)}\right)$, $k \in \mathbb{N}_0$, generiran relacijom $x^{(k+1)} = Tx^{(k)} + c$, $k \in \mathbb{N}_0$ konvergira prema rješenju linearnog sustava $Ax = b$ za sve početne vektore $x^{(0)}$ i sve desne strane b, ako i samo ako je

$$\rho(T) < 1,$$

pri čemu je $\rho(T)$ spektralni radijus matrice $T = M^{-1}N$. Uočite da $\rho(T)$ ovisi samo o A i njenom rastavu, a ne o b.

Dokaz. Ako je $\rho(T) \geq 1$, izaberimo startnu aproksimaciju $x^{(0)}$ tako da je $x^{(0)} - x$ svojstveni vektor koji pripada svojstvenoj vrijednosti λ, $\rho(T) = |\lambda|$.

Konvergencija standardnih iteracija (nastavak)

Tada vrijedi

$$\begin{aligned}\big(x^{(k+1)} - x\big) &= T\big(x^{(k)} - x\big) = \cdots = T^{k+1}\big(x^{(0)} - x\big) \\ &= \lambda^{k+1}\big(x^{(0)} - x\big),\end{aligned}$$

pa, očito $(k+1)$-a potencija broja koji je po apsolutnoj vrijednosti veći ili jednak 1, ne može težiti u 0.

S druge strane, ako je $\rho(T) < 1$, onda možemo izabrati $\varepsilon > 0$ takav da je

$$\rho(T) + \varepsilon < 1,$$

a zatim po Lemi 2 i postoji operatorska norma $\|\ \|_*$ takva da vrijedi

$$\|T\|_* \leq \rho(T) + \varepsilon < 1.$$

Konvergencija standardnih iteracija (nastavak)

Budući da je $\| \; \|_*$ operatorska norma, ponovno, po prvom dijelu Leme 2 slijedi da je

$$\rho(T) \leq \|T\|_*.$$

Zadnje dvije relacije zajedno daju da je

$$\|T\|_* < 1,$$

pa primjenom Leme 1 dobivamo traženi rezultat. ■

Lema 1 i prethodni teorem, zapravo nam daju i brzinu konvergencije.

Brzina konvergencije standardnih iteracija

Definicija. Za niz aproksimacija $x^{(k)}$, $k \in \mathbb{N}_0$, reći ćemo da konvergira prema x s redom p ako je

$$\|x^{(k+1)} - x\| \leq a\|x^{(k)} - x\|^p, \quad a \in \mathbb{R}_0^+.$$

Ako je $p = 1$ (tzv. linearna konvergencija), mora biti $a < 1$ (tzv. geometrijska konvergencija s faktorom a).

U slučaju konvergentnih iterativnih metoda, konvergencija je linearna, a faktor je $\mathcal{O}(\rho(T)) < 1$, tj. za $q > 0$ takav da je $(1 + q)\rho(T) < 1$ i za $\varepsilon = q\rho(T)$ vrijedi

$$\|x^{(k+1)} - x\|_* \leq (1 + q)\rho(T)\|x^{(k)} - x\|_*.$$

tako da spektralni radijus možemo smatrati ocjenom stope konvergencije iterativne metode.

Izbor rastava matrice

Naravno, sljedeći cilj nam je odgovoriti na pitanje kako odrediti rastav matrice $A = M - N$ koji zadovoljava

- $Tx = M^{-1}Nx$ i $c = M^{-1}b$ se lako računaju,
- $\rho(T)$ je malen?

Odmah nam se nameću neka jednostavna rješenja za ova dva suprotna cilja.

- Na primjer, izaberemo li $M = I$, M je regularna, i Tx i c se lako računaju, ali nije jasno da smo zadovoljili da je $\rho(T) < 1$.
- S druge strane, izbor $N = 0$ je izvrsan za drugi cilj $(\rho(T) = 0)$, ali nije dobar za prvi cilj, jer je $c = A^{-1}b$, tj. dobivamo polazni problem kojeg treba riješiti $(x = c)$.

Izbor rastava matrice (nastavak)

- Dakle, rastav koji bi uvijek dobro radio nije lako konstruirati.
- Međutim tu će nam pomoći praksa.
- Matrice koje se javljaju u praksi su ili pozitivno definitne ili (strogo) dijagonalno dominantne, pa će za takve tipove matrica biti mnogo lakše konstruirati iterativne metode i pokazati da one konvergiraju.

Uvedimo sljedeću notaciju. Pretpostavimo da A nema nula na dijagonali. Tada A možemo zapisati kao

$$A = L + D + R,$$

pri čemu je

- L strogo donji trokut od A
- D dijagonala od A
- R strogo gornji trokut od A

Jacobijeva metoda

Ako je matrica M iz rastava matrice A dijagonalna, tada se dobivena iterativna metoda naziva Jacobijevom metodom.

- To je jedna od najlakših metoda jer moramo invertirati dijagonalnu matricu.
- Kod Jacobijeve metode je

$$M_J = D, \qquad N_J = -(L + R),$$

odakle slijedi da je u iteracijama oblika

$$x^{(k+1)} = T_J x^{(k)} + c_J$$

$$T_J = -D^{-1}(L + R), \qquad c_J = D^{-1}b.$$

Jacobijeva metoda (nastavak)

Promatrajući jednadžbu $M_J x^{(k+1)} = N_J x^{(k)} + b$

$$Dx^{(k+1)} = -(L+R)x^{(k)} + b$$

po komponentama dobivamo

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

Odavde direktno slijedi kako se računa i-ta komponenta $(k+1)$-ve iteracije:

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

Algoritam Jacobijeve metode

Algoritam (Jacobijeva metoda).

$x^{(0)}$ `zadan;` $k = 0$;
`dok nije zadovoljen kriterij zaustavljanja radi {`
    `za` $i = 1$ `do` n `radi{`
        $x_i^{(k+1)} = b_i$;
        `za` $j = 1$ `do` $i - 1$ `radi {`
            $x_i^{(k+1)} = x_i^{(k+1)} - a_{ij} \cdot x_j^{(k)}$;
        }
        `za` $j = i + 1$ `do` n `radi {`
            $x_i^{(k+1)} = x_i^{(k+1)} - a_{ij} \cdot x_j^{(k)}$;
        }
        $x_i^{(k+1)} = x_i^{(k+1)} / a_{ii}$;
        $k = k + 1$;
    }
}

Detalji implementacije Jacobijeve metode

- Primijetimo da kod skladištenja vektora u memoriji, novi vektor $x^{(k+1)}$ ne može se prepisati preko starog vektora $x^{(k)}$ sve dok sve komponenete vektora $x^{(k+1)}$ nisu izračunate.
- Dakle, prije kraja jedne iteracije oba vektora moramo posebno skladištiti.
- Dalje, uočimo da prolaz kroz petlju po i u ovom algoritmu možemo izvršiti bilo kojim poretkom — ne nužno sekvencijalnim.
- Komponente novog vektora $x^{(k+1)}$ ovise samo o komponentama starog vektora $x^{(k)}$, a ne i o nekim novim komponentama. Zbog toga je Jacobijeva metoda idealna za paralelno računanje, jer pojedine komponente novog vektora možemo računati potpuno nezavisno.

Konvergencija Jacobijeve metode

Teorem Ako je matrica A strogo dijagonalno dominantna u smislu

$$|a_{ii}| > \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|, \quad i = 1, \ldots, n,$$

tada Jacobijeva metoda konvergira za svaku početnu iteraciju $x^{(0)}$ prema rješenju sustava $Ax = b$.

Dokaz. Iz pretpostavke teorema vrijedi sljedeće

$$\|T_J\|_\infty = \max_{i=1,\ldots,n} \sum_{j=1}^{n} |(T_J)_{ij}| = \max_{i=1,\ldots,n} \frac{1}{|a_{ii}|} \sum_{j\neq i} |a_{ij}| < 1,$$

pa prema Teoremu o konvergenciji iterativnih metoda tj. prema Lemi 1, Jacobijeve iteracije konvergiraju za svaku početnu iteraciju. ■

Gauss–Seidelova metoda

Ako je matrica M iz rastava matrice A donje trokutasta, tada se dobivena iterativna metoda naziva Gauss–Seidelovom metodom.

- Kod Gauss–Seidelove metode je

$$M_{GS} = D + L, \qquad N_{GS} = -R,$$

odakle slijedi da je u iteracijama oblika

$$x^{(k+1)} = T_{GS} x^{(k)} + c_{GS}$$

$$T_{GS} = -(D + L)^{-1} R, \qquad c_{GS} = (D + L)^{-1} b.$$

Gauss–Seidelova metoda (nastavak)

Promatrajući jednadžbu $M_{GS}x^{(k+1)} = N_{GS}x^{(k)} + b$

$$(D + L)x^{(k+1)} = -Rx^{(k)} + b$$

po komponentama dobivamo

$$\sum_{j=1}^{i} a_{ij}x_j^{(k+1)} = -\sum_{j=i+1}^{n} a_{ij}x_j^{(k)} + b_i, \quad i = 1, \ldots, n.$$

Odavde slijedi kako se računa i-ta komponenta $(k + 1)$-ve iteracije:

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

Algoritam Gauss–Seidelove metode

Algoritam (Gauss–Seidelova metoda).

```
x^(0) zadan; k = 0;
dok nije zadovoljen kriterij zaustavljanja radi {
    za i = 1 do n radi{
        x_i^(k+1) = b_i;
        za j = 1 do i − 1 radi {
            x_i^(k+1) = x_i^(k+1) − a_ij · x_j^(k+1);
        }
        za j = i + 1 do n radi {
            x_i^(k+1) = x_i^(k+1) − a_ij · x_j^(k);
        }
        x_i^(k+1) = x_i^(k+1)/a_ii;
        k = k + 1;
    }
}
```

Detalji implementacije Gauss-Seidelove metode

- Kod računanja tekuće komponente nove iteracije $x^{(k+1)}$ koriste se sve do sada izračunate komponente od $x^{(k+1)}$.
- Time se omogućava da komponente od $x^{(k+1)}$ prebrišu stare vrijednosti komponenata od $x^{(k)}$, čim se izračunaju, tako da ne moramo pamtiti dva vektora, već samo jedan.
- S druge strane, redosljed računanja komponenti iteracije kod Gauss–Seidelovog algoritma je bitan. U algoritmu smo uzeli prirodni redosljed — od prve prema zadnjoj komponenti (petlja po i od 1 do n).
- To ne znači da je to i jedini mogući redosljed.
- U principu, možemo uzeti i bilo koji drugi redosljed, tj. bilo koju drugu od mogućih $n!$ permutacija jednadžbi. No, zbog sekvencijalnosti popravaka, rezultat nije isti, tj. dobivamo drugačiju iterativnu metodu.

Detalji implementacije Gauss-Seidelove metode

- U praksi se katkad i koriste drugačiji redosljedi, ali za sasvim posebne linearne sustave koji nastaju diskretizacijom parcijalnih diferencijalnih jednadžbi.
- Na primjer, za Laplaceovu jednadžbu u dvije dimenzije koristi se tzv. "crveno–crni" poredak (engl. "red–black" ordering), koji naliči šahovskoj ploči s crvenim i crnim poljima, s tim da se prvo računaju sva polja crvene boje, a zatim sva polja crne boje. Zbog posebne strukture sustava, ovaj poredak dozvoljava i efikasnu paralelizaciju.

Konvergencija Gauss–Seidelove metode

Teorem Ako je A strogo dijagonalno dominantna matrica po recima, onda i Jacobijeva i Gauss–Seidelova metoda konvergiraju i vrijedi

$$\|T_{GS}\|_\infty \leq \|T_J\|_\infty < 1.$$

Napomena. Dakle, Gauss–Seidelova metoda u najgorem slučaju konvergira barem tako brzo kao Jacobijeva metoda u najogorem slučaju. To ne znači da će Gauss–Seidelova metoda konvergirati brže nego Jacobijeva za bilo koji problem $Ax = b$.

Dokaz. Prema Teoremu za konvergenciju Jacobijeve metode znamo da je $\|T_J\|_\infty < 1$.

Konvergencija Gauss–Seidelove metode (nast.)

Za potrebe dokaza definirajmo:

$$\tilde{L} = -D^{-1}L, \qquad \tilde{R} = -D^{-1}R.$$

Tada su

$$T_J = -D^{-1}(L + R) = \tilde{L} + \tilde{R},$$
$$T_{GS} = -(D + L)^{-1}R = -(I + D^{-1}L)^{-1}D^{-1}R = (I - \tilde{L})^{-1}\tilde{R}.$$

Želimo dokazati da vrijedi

$$\|T_{GS}\|_\infty = \|\,|T_{GS}|\,e\|_\infty \leq \|\,|T_J|\,e\|_\infty = \|T_J\|_\infty,$$

pri čemu je e vektor sa svim komponentama jednakim 1, tj. $e = [1, \ldots, 1]^T$.

Konvergencija Gauss–Seidelove metode (nast.)

Ovu relaciju možemo lako pokazati ako dokažemo jaču komponentnu nejednakost

$$|(I - \tilde{L})^{-1}\tilde{R}| \cdot e = |T_{GS}| \cdot e \leq |T_J| \cdot e = (|\tilde{L}| + |\tilde{R}|) \cdot e.$$

Krenimo slijeva i iskoristimo relaciju trokuta. Vrijedi

$$|(I - \tilde{L})^{-1}\tilde{R}| \cdot e \leq |(I - \tilde{L})^{-1}| \, |\tilde{R}| \cdot e.$$

Ako je $\rho(\tilde{L}) < 1$, (a je, jer je $\rho(\tilde{L}) = 0$!), onda je $I - \tilde{L}$ regularna i možemo $(I - \tilde{L})^{-1}$ razviti u red

$$(I - \tilde{L})^{-1} = I + \tilde{L} + \tilde{L}^2 + \cdots + \tilde{L}^n + \cdots$$

Konvergencija Gauss–Seidelove metode (nast.)

Primijetimo da je $\tilde{L}$ strogo donje trokutasta pa je i nilpotentna matrica, i vrijedi $\tilde{L}^k = 0$ za $k \geq n$. Time prethodni red postaje

$$(I - \tilde{L})^{-1} = I + \tilde{L} + \tilde{L}^2 + \cdots + \tilde{L}^{n-1}.$$

Korištenjem prethodnih razmatranja dobivamo

$$\begin{aligned} |(I - \tilde{L})^{-1} \tilde{R}| \cdot e &\leq \left| \sum_{i=0}^{n-1} \tilde{L}^i \right| |\tilde{R}| \cdot e \leq \sum_{i=0}^{n-1} |\tilde{L}|^i \, |\tilde{R}| \cdot e \\ &\leq (I - |\tilde{L}|)^{-1} \, |\tilde{R}| \cdot e. \end{aligned}$$

Još moramo pokazati da je

$$(I - |\tilde{L}|)^{-1} \, |\tilde{R}| \cdot e \leq (|\tilde{L}| + |\tilde{R}|) \cdot e.$$

Konvergencija Gauss–Seidelove metode (nast.)

Budući da su svi članovi u redu za $(I - |\tilde{L}|)^{-1}$ nenegativni, dovoljno je pokazati da je

$$|\tilde{R}| \cdot e \leq (I - |\tilde{L}|)\,(|\tilde{L}| + |\tilde{R}|) \cdot e = (|\tilde{L}| + |\tilde{R}| - |\tilde{L}|^2 - |\tilde{L}|\,|\tilde{R}|) \cdot e,$$

odnosno

$$0 \leq (|\tilde{L}| - |\tilde{L}|^2 - |\tilde{L}|\,|\tilde{R}|) \cdot e = |\tilde{L}|\,(I - |\tilde{L}| - |\tilde{R}|) \cdot e.$$

Ponovno, budući da su svi elementi $|\tilde{L}|$ nenegativni, prethodna nejednakost bit će ispunjena ako je

$$0 \leq (I - |\tilde{L}| - |\tilde{R}|) \cdot e,$$

odnosno

$$(|\tilde{L}| + |\tilde{R}|) \cdot e \leq e.$$

Konvergencija Gauss–Seidelove metode (nast.)

Budući da je $|T_J| = |\tilde{L} + \tilde{R}| = |\tilde{L}| + |\tilde{R}|$, jer se elementi $\tilde{L}$ i $\tilde{R}$ nigdje ne zbrajaju, onda je posljednja nejednakost ekvivalentna s

$$|T_J| \cdot e \leq e.$$

Ovu posljednju nejednakost možemo dokazati jer je

$$\| \, |T_J| \cdot e\|_\infty = \|T_J\|_\infty < 1.$$

Čitanjem dokaza u obratnom redosljedu, dobivamo tražena svojstva. ∎

Konvergencija Gauss–Seidelove metode (nast.)

Teorem. Ako je matrica A hermitska (simetrična) i pozitivno definitna, tada Gauss–Seidelova metoda konvergira za svaku početnu iteraciju $x^{(0)}$.

Dokaz. Neka je λ proizvoljna svojstvena vrijednost matrice T_{GS} i neka je $y \neq 0$ odgovarajući svojstveni vektor. Tada imamo

$$-Ry = \lambda(D + L)y,$$

pa ako toj jednakosti dodamo λRy s lijeve i desne strane, vrijedi

$$-(1 - \lambda)Ry = \lambda(D + L + R)y = \lambda Ay.$$

S druge strane je, ponovo zbog $-Ry = \lambda(D + L)y$

$$Ay = (D + L)y + Ry = (1 - \lambda)(D + L)y.$$

Konvergencija Gauss–Seidelove metode (nast.)

Pomnožimo sada prvu prethodnu jednakost skalarno sa λy, a drugu sa y, i oduzmimo prvu jednadžbu od druge, tada vrijedi

$$(1 - |\lambda|^2)\langle Ay, y\rangle = (1 - \lambda)(\langle Dy, y\rangle + \langle Ly, y\rangle + \langle Ry, \lambda y\rangle).$$

Promatramo sada samo zadnji izraz u prethodnoj jednakosti. Zbog hermitičnosti matrice A vrijedi da je $R^* = L$, zbog pozitivne definitnosti $D > 0$, a iz prve jednakosti dokaza vrijedi da je $\lambda Ly = -\lambda Dy - Ry$, pa imao

$$\begin{aligned} \langle Ry, \lambda y\rangle &= \langle y, \lambda Ly\rangle = -\langle y, \lambda Dy\rangle - \langle y, Ry\rangle \\ &= -\langle Dy, \lambda y\rangle - \langle Ly, y\rangle. \end{aligned}$$

Sada zadnji izraz ubacimo u gornju jednakost, čime dobivamo

Konvergencija Gauss–Seidelove metode (nast.)

$$\begin{aligned}(1-|\lambda|^2)\langle Ay,y\rangle &= (1-\lambda)(\langle Dy,y\rangle + \langle Ly,y\rangle - \langle Dy,\lambda y\rangle - \langle Ly,y\rangle)\\ &= (1-\lambda)(1-\bar{\lambda})\langle Dy,y\rangle,\end{aligned}$$

odakle zbog $\langle Ay,y\rangle > 0$ i zbog $\lambda \neq 1$ (budući da bi iz $-(1-\lambda)Ry = \lambda Ay$ za $\lambda = 1$ slijedilo da je $Ay = 0$, što je kontradiktorno sa svojstvom pozitivne definitnosti matrice A), slijedi

$$1-|\lambda|^2 = \frac{\langle Dy,y\rangle}{\langle Ay,y\rangle}|1-\lambda|^2 > 0.$$

Znači za svaku svojstvenu vrijednost λ od T_{GS} vrijedi da je $|\lambda| < 1$, pa to vrijedi i za svojstvenu vrijednost koja ima maksimalnu apsolutnu vrijednost, odnosno $\rho(T_{GS}) < 1$, iz čega prema Teoremu o konvergenciji iterativnih metoda slijedi da Gauss–Seidelova metoda konvergira za svaki $x^{(0)}$. ∎

Relaksacijske metode

- Kad jednom znamo konstruirati iterativni proces, nameće se vrlo jednostavna ideja za poboljšanje njezine konvergencije, uvođenjem jednog realnog parametra.
- Nove aproksimacije možemo računati u dva koraka.
  - Prvo iz $x^{(k)}$ nađemo (jednostavnu) pomoćnu sljedeću aproksimaciju $x_*^{(k+1)}$, a zatim
  - za "pravu" novu aproksimaciju $x^{(k+1)}$ uzmemo težinsku sredinu prethodne aproksimacije $x^{(k)}$ i pomoćne nove aproksimacije $x_*^{(k+1)}$

$$x^{(k+1)} = (1 - \omega)x^{(k)} + \omega x_*^{(k+1)} = x^{(k)} + \omega(x_*^{(k+1)} - x^{(k)}),$$

gdje je ω težinski parametar kojeg možemo birati.

Relaksacijske metode (nastavak)

- Očito, za $\omega = 1$ dobivamo $x^{(k+1)} = x_*^{(k+1)}$, pa je ova metoda proširenje metode za nalaženje pomoćnih aproksimacija.
- Obično se uzima $\omega \in \mathbb{R}$ i $\omega \neq 0$, da ne dobijemo stacionaran niz.
- Ideja za ovakav postupak dolazi iz općih metoda za rješavanje jednadžbi i minimizaciju funkcionala.
- Pomoćna aproksimacija $x_*^{(k+1)}$ daje smjer korekcije prethodne aproksimacije $x^{(k)}$ u kojem treba ići da bismo se približili pravom rješenju sustava (smanjili rezidual $r(x) = Ax - b$ u nekoj normi).

Relaksacijske metode (nastavak)

- Ako je $x_*^{(k+1)} - x^{(k)}$ dobar smjer korekcije, onda možemo dodati i izbor duljine koraka ω u smjeru te korekcije, tako da dobijemo što bolji $x^{(k+1)}$. (Ovakav pristup ćemo vidjeti kod metoda najbržeg silaska i konjugiranih gradijenata.)
- Općenito očekujemo da je $\omega > 0$, tj. da idemo u smjeru vektora korekcije, a ne suprotno od njega, a stvarno želimo dobiti $\omega > 1$, tako da se još više maknemo u dobrom smjeru i približimo pravom rješenju.
- Naziv "relaksacija" dolazi upravo iz minimizacijskih metoda, a odnosi se na sve iterativne metode koje koriste neki oblik minimizacije ili pokušaja minimizacije reziduala. U tom smislu, često se koristi i tradicionalni naziv "relaksacijski parametar" za ω.

Relaksacijske metode (nastavak)

- Obzirom na vrijednost parametra ω, imamo tri različita slučaja.
  - Ako je $\omega = 1$, onda se metoda svodi na pomoćnu metodu i to je tzv. obična ili standardna relaksacija.
  - Ako je $\omega < 1$, onda takvu metodu zovemo podrelaksacija (engl. "underrelaxation"),
  - a ako je $\omega > 1$ onda metodu zovemo pre- ili nad-relaksacija (engl. "overrelaxation").
- U općem slučaju, ω se posebno računa u svakoj pojedinoj iteraciji, tako da dobijemo što bolji $x^{(k+1)}$.
- Postupak se svodi na jednodimenzionalnu optimizaciju, a ovisi o kriteriju optimalnosti za mjerenje "kvalitete" aproksimacija.

Relaksacijske metode (nastavak)

- Srećom, za neke klase linearnih sustava, koje su izrazito bitne u praksi, unaprijed se može dobro odabrati optimalni ili skoro optimalni parametar ω za maksimalno ubrzanje konvergencije iterativnih metoda,
- i to tako da isti ω vrijedi za sve iteracije.
- U većini slučajeva dobivamo $\omega > 1$ za optimalni ω, pa se takve metode standardno zovu "OverRelaxation" i skraćeno označavaju s OR.
- Obzirom na to da se ω zadaje ili bira unaprijed, a zatim koristi za sve iteracije, metoda je ovisna o jednom parametru i standardno koristimo oznaku $\mathrm{OR}(\omega)$.

JOR metoda – Jacobi overrelaxation

Ako se u iteraciji

$$x^{(k+1)} = (1-\omega)x^{(k)} + \omega x_*^{(k+1)} = x^{(k)} + \omega(x_*^{(k+1)} - x^{(k)}),$$

pomoćna nova aproksimacija $x_*^{(k+1)}$ računa po Jacobijevoj metodi, dobivamo Jacobijevu nadrelaksaciju ili JOR metodu.

Dakle, i-ta komponenta $(k+1)$-ve iteracije je oblika:

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

Kao i kod obične Jacobijeve metode, petlju po i možemo izvršiti bilo kojim redom, uz isti rezultat, pa se i ovdje komponente novog vektora $x^{(k+1)}$ mogu paralelno računati.

JOR metoda – Jacobi overrelaxation (nastavak)

Vektorski oblik iteracija u JOR(ω) metodi je

$$x^{(k+1)} = (1-\omega)x^{(k)} + \omega(T_J x^{(k)} + c_J) := T_{JOR(\omega)} x^{(k)} + c_{JOR(\omega)}$$

pa je

$$\begin{aligned} T_{JOR(\omega)} &= (1-\omega)I + \omega T_J = (1-\omega)I - \omega D^{-1}(L+R), \\ c_{JOR(\omega)} &= \omega c_J = \omega D^{-1} b. \end{aligned}$$

Pripadni rastav matrice A je

$$A = \frac{1}{\omega} D - \left(\frac{1-\omega}{\omega} D - L - R \right),$$

tj. $M = \frac{1}{\omega} D$, $N = \frac{1-\omega}{\omega} D - L - R$. Naravno, za $\omega = 1$ dobivamo Jacobijevu metodu.

JOR metoda – Jacobi overrelaxation (nastavak)

- Iz oblika matrice $T_{JOR(\omega)}$ vidimo da se optimalni parametar ω koji minimizira njen spektralni radijus može odrediti unaprijed, prije početka iteracija.
- Drugim riječima, treba koristiti isti ω u svim iteracijama, naravno, pod uvjetom da imamo konvergenciju.

Algoritam JOR metode

Algoritam (JOR metoda).

```
x^(0) i ω zadani; k = 0;
dok nije zadovoljen kriterij zaustavljanja radi {
    za i = 1 do n radi{
        pom = (1 − ω) x_i^(k); x_i^(k+1) = b_i;
        za j = 1 do i − 1 radi {
            x_i^(k+1) = x_i^(k+1) − a_ij · x_j^(k);
        }
        za j = i + 1 do n radi {
            x_i^(k+1) = x_i^(k+1) − a_ij · x_j^(k);
        }
        x_i^(k+1) = pom + ω · x_i^(k+1) / a_ii;
        k = k + 1;
    }
}
```

Konvergencija JOR metode

Teorem. Ako Jacobijeva metoda konvergira za svaku početnu iteraciju x_0, onda za proizvoljno $\omega \in \langle 0, 1]$ konvergira i JOR(ω) metoda i to za svako x_0.

Dokaz. Neka je λ proizvoljna svojstvena vrijednost Jacobijeve matrice iteracija T_J. Tada zbog konvergencije Jacobijeve metode za svako x_0 slijedi

$$|\lambda| < 1.$$

$T_{JOR(\omega)}$ je oblika $T_{JOR(\omega)} = (1 - \omega)I + \omega T_J$, pa je njezina svojstvena vrijednost μ oblika

$$\mu = 1 - \omega + \omega\lambda.$$

Konvergencija JOR metode (nastavak)

Neka je $\lambda = \alpha + \iota\beta$, za $\alpha, \beta \in \mathbb{R}$, i $\iota^2 = -1$.

- Tada zbog $|\lambda| < 1$ slijedi da je $\alpha^2 + \beta^2 < 1$ i $\alpha^2 < 1$.
- Zbog prethodno rečenog i zbog činjenice da je $1 - \omega \geq 0$, vrijedi

$$\begin{aligned} |\mu|^2 &= (1 - \omega + \omega\alpha)^2 + \omega^2\beta^2 \\ &\leq (1 - \omega)^2 + 2\omega(1 - \omega)|\alpha| + \omega^2(\alpha^2 + \beta^2) \\ &< (1 - \omega)^2 + 2\omega(1 - \omega) + \omega^2 = (1 - \omega + \omega)^2 = 1. \end{aligned}$$

Odavde slijedi da je $\rho(T_{JOR(\omega)}) < 1$, pa prema Teoremu o konvergenciji iterativnih metoda, JOR(ω) metoda konvergira za svaku početnu iteraciju. ∎

Konvergencija JOR metode (nastavak)

Teorem. Neka je A hermitska (simetrična) pozitivno definitna matrica i neka Jacobijeva metoda konvergira. Tada konvergira i JOR(ω) metoda za

$$0 < \omega < \frac{2}{1-\lambda} \leq 2,$$

gdje je $\lambda \leq 0$ najmanja svojstvena vrijednost Jacobijeve matrice iteracija T_J.

Dokaz. Neka su λ_i, $i = 1, \ldots, n$ svojstvene vrijednosti matrice T_J.

- Zbog konvergencije Jacobijeve metode slijedi da je $|\lambda_i| < 1$ za sve $i = 1, \ldots, n$.

Konvergencija JOR metode (nastavak)

- Kako su, zbog sličnosti matrice $T_J = -D^{-1}(L + L^*)$ sa hermitskom matricom, sve λ_i realni brojevi i

$$0 = \mathrm{tr}(T_J) = \sum_{i=1}^{n} \lambda_i$$

slijedi da nisu sve λ_i pozitivne, te vrijedi

$$\lambda = \min_{i=1,\ldots,n} \lambda_i \leq 0.$$

Ponovo vrijedi da su svojstvene vrijednosti matrice $T_{JOR(\omega)}$ oblika

$$\mu_i = 1 - \omega(1 - \lambda_i), \qquad i = 1, \ldots, n.$$

Konvergencija JOR metode (nastavak)

Kako je prema uvjetima teorema

$$0 < \omega(1-\lambda) < 2, \quad \text{i} \quad \lambda_i < 1, \quad \text{za } i = 1, \ldots, n$$

slijedi

$$0 < \omega(1-\lambda_i) < \omega(1-\lambda) < 2,$$

odnosno

$$-1 < \mu_i = 1 - \omega(1-\lambda_i) < 1,$$

odakle je $\rho(T_{JOR(\omega)}) < 1$, pa prema Teoremu o konvergenciji iterativnih metoda JOR(ω) metoda konvergira za svaku početnu iteraciju. $\blacksquare$

Konvergencija JOR metode (nastavak)

Teorem. JOR(ω) metoda ne konvergira za

$$\omega < 0, \quad \text{i} \quad \omega \geq 2.$$

Dokaz. Iz definicije matrice $T_{JOR(\omega)} = (1 - \omega)I + \omega T_J$ vidimo da je

$$\mathrm{tr}(T_{JOR(\omega)}) = n(1 - \omega) = \sum_{i=1}^{n} \mu_i,$$

pri čemu su μ_i svojstvene vrijednosti matrice $T_{JOR(\omega)}$, imamo

$$n|1 - \omega| \leq \sum_{i=1}^{n} |\mu_i| \leq n\rho(T_{JOR(\omega)}).$$

Konvergencija JOR metode (nastavak)

Dakle,

$$\rho(T_{JOR(\omega)}) \geq |1 - \omega|,$$

odakle se vidi da je za $\omega < 0$ i $\omega \geq 2$ $\rho(G_{JOR,\omega}) \geq 1$, iz čega, prema Teoremu o konvergenciji iterativnih metoda slijedi tvrdnja teorema. ∎

SOR metoda – Successive overrelaxation

- Relacija

$$x^{(k+1)} = (1-\omega)x^{(k)} + \omega x_*^{(k+1)} = x^{(k)} + \omega(x_*^{(k+1)} - x^{(k)}),$$

  koristi ideju težinske sredine ili duljine koraka na nivou vektorskih aproksimacija $x^{(k)}$.
- To prirodno odgovara Jacobijevoj metodi i paralelnom računanju.
- Međutim, potpuno istu ideju možemo koristiti i za poboljšanje svake pojedine varijable $x_i^{(k)}$, tj. pojedinačnih komponenti vektora $x^{(k)}$, što odgovara "Gauss–Seidelovskom" pristupu.

SOR metoda – Successive overrelaxation (nast.)

Dakle, nova aproksimacija i-te varijable ima oblik

$$x_i^{(k+1)} = (1-\omega)x_i^{(k)} + \omega x_{i,*}^{(k+1)} = x_i^{(k)} + \omega(x_{i,*}^{(k+1)} - x_i^{(k)}), \quad {\scriptstyle i=1,\ldots,n,}$$

gdje je $x_{i,*}^{(k+1)}$ neka pomoćna nova aproksimacija i-te varijable, koju računamo tog trenutka kad nam treba, za svaki i.

- SOR metoda (engl. "Successive OverRelaxation" ili ponovljena nadrelaksacija) je proširenje ili poboljšanje Gauss–Seidelove metode u smislu da se pomoćna nova aproksimacija $x_{i,*}^{(k+1)}$ računa po Gauss–Seidelovoj metodi, pa ju označavamo s $x_{i,GS}^{(k+1)}$.
- Tada gornja relacija za i-tu komponentu nove aproksimacije u SOR(ω) metodi ima oblik

$$x_i^{(k+1)} = (1-\omega)x_i^{(k)} + \omega x_{i,GS}^{(k+1)}, \qquad i = 1, \ldots, n.$$

SOR metoda – Successive overrelaxation (nast.)

Dakle, i-ta komponenta $(k+1)$-ve iteracije je oblika:

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

Iz gornje jednadžbe odmah vidimo da je za $i = 1, \ldots, n$

$$a_{ii}x_i^{(k+1)} + \omega\sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} = (1-\omega)a_{ii}x_i^{(k)} - \omega\sum_{j=i+1}^{n} a_{ij}x_j^{(k)} + \omega b_i, .$$

Ove jednadžbe po komponentama možemo zapisati u vektorskom obliku

$$(D+\omega L)x^{(k+1)} = \big((1-\omega)D - \omega R\big)x^{(k)} + \omega b,$$

SOR metoda – Successive overrelaxation (nast.)

odakle slijedi

$$T_{SOR(\omega)} = (D + \omega L)^{-1}((1 - \omega)D - \omega R),$$
$$c_{SOR(\omega)} = \omega(D + \omega L)^{-1} b.$$

I ovdje vidimo da se dobra vrijednost za ω, ako postoji, može odrediti unaprijed za sve iteracije.

Pripadni rastav matrice A je

$$A = \frac{1}{\omega} D + L - \left(\frac{1 - \omega}{\omega} D - R \right),$$

tj. $M = \frac{1}{\omega} D + L$, $N = \frac{1-\omega}{\omega} D - R$. Naravno, za $\omega = 1$ dobivamo Gauss–Seidelovu metodu.

Algoritam SOR metode

Algoritam (SOR metoda).

```
x^(0) i ω zadani; k = 0;
dok nije zadovoljen kriterij zaustavljanja radi {
    za i = 1 do n radi{
        pom = (1 − ω) x_i^(k);  x_i^(k+1) = b_i;
        za j = 1 do i − 1 radi {
            x_i^(k+1) = x_i^(k+1) − a_ij · x_j^(k+1);
        }
        za j = i + 1 do n radi {
            x_i^(k+1) = x_i^(k+1) − a_ij · x_j^(k);
        }
        x_i^(k+1) = pom + ω · x_i^(k+1) / a_ii;
        k = k + 1;
    }
}
```

Konvergencija SOR metode

Teorem. SOR(ω) metoda ne konvergira za

$$\omega \leq 0, \quad \text{i} \quad \omega \geq 2.$$

Dokaz. Očito je da za SOR(ω) matricu iteracija vrijedi

$$\begin{aligned} \det(T_{SOR(\omega)}) &= (\det(D + \omega L))^{-1} \det((1-\omega)D - \omega R) = \\ &= (\det(D))^{-1}(1-\omega)^n \det(D) = (1-\omega)^n, \end{aligned}$$

jer su $D + \omega L$ i $(1-\omega)D - \omega R$ donja i gornja trokutasta matrica. Kako je

$$(\rho(T_{SOR(\omega)}))^n \geq \prod_{i=1}^{n} |\lambda_i| = |\det(T_{SOR(\omega)})| = |1-\omega|^n,$$

Konvergencija SOR metode (nastavak)

gdje su λ_i svojstvene vrijednosti matrice $T_{SOR(\omega)}$, onda je

$$\rho(T_{SOR(\omega)}) \geq |1 - \omega|.$$

Prema Teoremu o konvergenciji iterativnih metoda SOR(ω) metoda konvergirat će ako i samo ako je $\rho(T_{SOR(\omega)}) < 1$, pa ako imamo da je $|1 - \omega| \geq 1$, tada SOR metoda neće konvergirati. Ovaj uvjet bit će ispunjen ako i samo ako je $\omega \leq 0$ i $\omega \geq 2$. ■

Konvergencija SOR metode (nastavak)

Teorem. Neka je A hermitska (simetrična) pozitivno definitna matrica. Tada $\mathrm{SOR}(\omega)$ metoda konvergira za $\omega \in \langle 0, 2 \rangle$.

Dokaz. Neka je λ proizvoljna svojstvena vrijednost $\mathrm{SOR}(\omega)$ matrice $T_{SOR(\omega)}$ i neka je $y \neq 0$ odgovarajući svojstveni vektor. Tada vrijedi

$$G_{SOR(\omega)} y = \lambda y,$$

odnosno

$$((1 - \omega)D - \omega R)y = \lambda(D + \omega L)y.$$

Za lijevu stranu prethodne jednakosti, direktnim računom dobivamo

$$2((1 - \omega)D - \omega R) = (2 - \omega)D - \omega A + \omega(L - R),$$

Konvergencija SOR metode (nastavak)

a za desnu stranu vrijedi

$$2(D + \omega L) = (2 - \omega)D + \omega A + \omega(L - R).$$

Koristeći se time, uz činjenicu da je zbog hermitičnosti matrice A, $L = R^*$, poslije skalarnog množenja sa y, iz prve jednakosti dokaza dobivamo

$$\begin{aligned}(2 - \omega)\langle Dy, y\rangle - \omega\langle Ay, y\rangle + \omega\langle (R^* - R)y, y\rangle \\ = \lambda((2 - \omega)\langle Dy, y\rangle + \omega\langle Ay, y\rangle + \omega\langle (R^* - R))y, y\rangle).\end{aligned}$$

Budući da je A pozitivno definitna matrica, vrijedi da je $a_{ii} > 0$ za svako $i = 1, \ldots, n$, a kako je $y \neq 0$ imamo

$$\langle Ay, y\rangle > 0, \qquad \langle Dy, y\rangle > 0.$$

Konvergencija SOR metode (nastavak)

Matrica $R^* - R$ je antihermitska pa je $\langle (R^* - R)y, y \rangle$ čisto imaginarni broj ili nula. Definirajmo

$$d = \langle Dy, y \rangle, \quad a = \langle Ay, y \rangle, \quad \iota u = \langle (R^* - R)y, y \rangle,$$

gdje je $\iota = \sqrt{-1}$. Sada prethodna jednakost glasi

$$(2 - \omega)d - \omega a + \iota\omega u = \lambda((2 - \omega)d + \omega a + \iota\omega u),$$

odakle je

$$\lambda = \frac{(2 - \omega)d - \omega a + \iota\omega u}{(2 - \omega)d + \omega a + \iota\omega u}.$$

Imaginarni dijelovi brojnika i nazivnika u izrazu za λ su jednaki, pa je $|\lambda| < 1$ ako i samo ako je

Konvergencija SOR metode (nastavak)

$$|(2-\omega)d + \omega a| > |(2-\omega)d - \omega a|,$$

što je, nakon kvadriranja, ekvivalentno sa

$$4\omega(2-\omega)da > 0.$$

Nadalje, budući da su $d > 0$ i $a > 0$, to će vrijediti ako i samo ako je

$$\omega(2-\omega) > 0,$$

što se postiže ako i samo ako je

$$\omega \in \langle 0, 2 \rangle.$$

Za takav izbor ω, vrijedi $|\lambda| < 1$ za proizvoljni λ, pa će vrijediti i $\rho(T_{SOR(\omega)}) < 1$.

Konvergencija SOR metode (nastavak)

Dakle, prema Teoremu o konvergenciji iterativnih metoda SOR(ω) metoda konvergira za svaku početnu iteraciju. ■

Primjer konvergencije iterativnih metoda

Primjer konvergencije iterativnih metoda

Promatramo sustav sa matricom

$$A = \begin{bmatrix} 1 & 0.01 & \cdots & 0.01 & 0.01 \\ 0.02 & 2 & \cdots & 0.02 & 0.02 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0.99 & 0.99 & \cdots & 99 & 0.99 \\ 1 & 1 & \cdots & 1 & 100 \end{bmatrix},$$

i vektorom desne strane b takvim da je egzaktno rješenje jednako $x = [\; 1 \quad \cdots \quad 1 \;]^T$.

Primjer konvergencije iterativnih metoda (nast.)

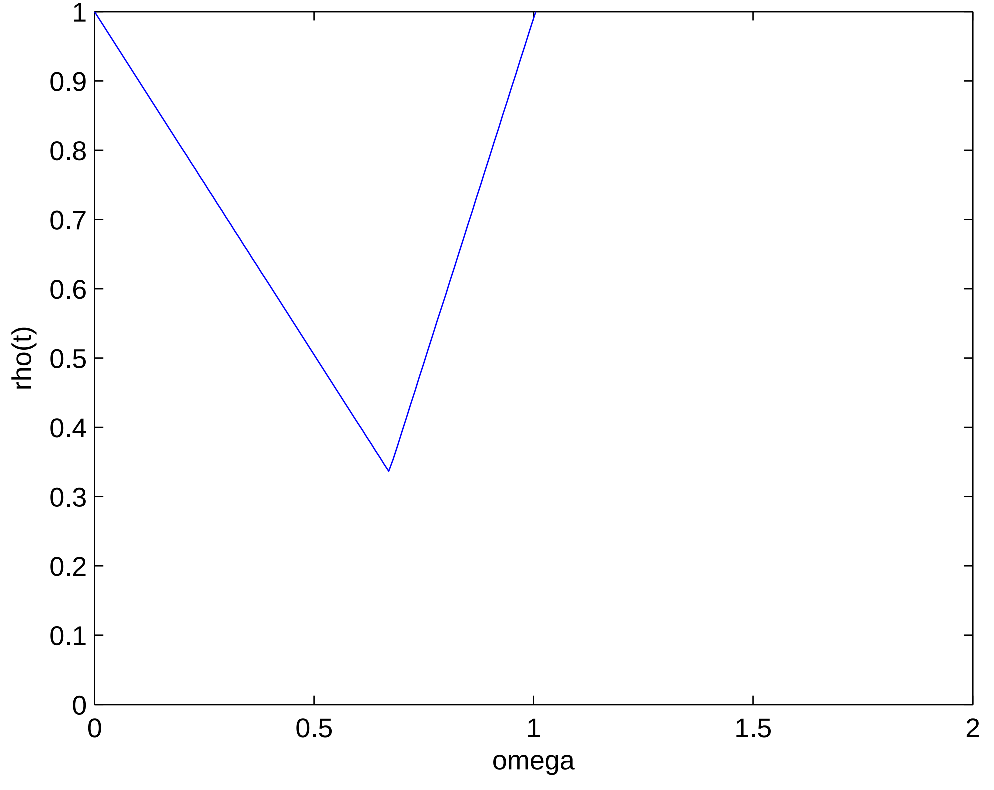
Spektralni radijusi matrica iteracija su sljedeći:

- Za Jacobijevu metodu : $\rho(G_J) = 0.99$.
- Za Gauss–Seidelovu metodu : $\rho(G_{GS}) = 0.2144$.
- Za JOR metodu eksperimentalno je utvrđeno da optimalni parametar iznosi $\omega_{JOR} = 0.67$: $\rho(G_{JOR(0.67)}) = 0.3367$.
- Za SOR metodu eksperimentalno je utvrđeno da optimalni parametar iznosi $\omega_{SOR} = 0.9$: $\rho(G_{SOR(0.9)}) = 0.1713$.

Pogledajte MATLAB program!

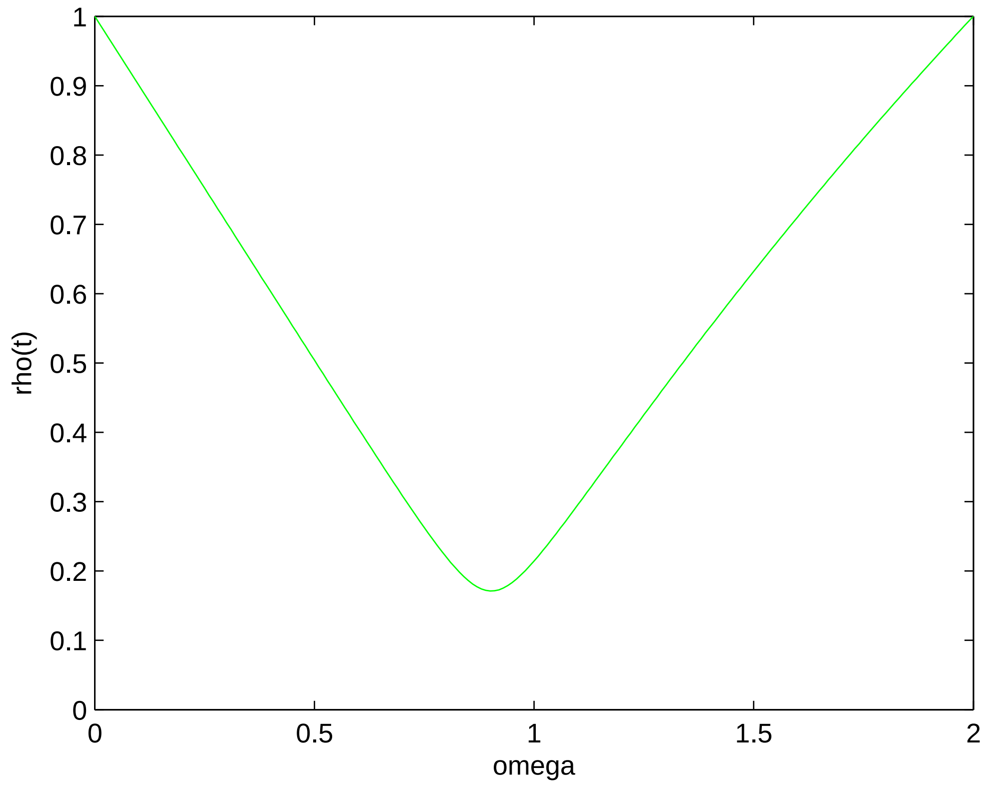
Primjer konvergencije iterativnih metoda (nast.)

Spektralni radijusi matrice iteracije JOR metode



Primjer konvergencije iterativnih metoda (nast.)

Spektralni radijusi matrice iteracije SOR metode



Primjer konvergencije iterativnih metoda (nast.)

Konvergencija iterativnih metoda