Metoda sukcesivne nadrelaksacije (SOR) *

DAVOR MENON†
Odjel za matematiku, Osijek

8. srpnja 2005.

Sažetak

Prošlo je preko pedeset godina od D. M. Youngova rada o metodi sukcesivne nadrelaksacije (eng. *successive overrelaxation method*). Ova metoda se danas pojavljuje u svim udžbenicima u kojima se obrađuju iterativne metode.
U ovom radu se naglasak stavlja na SOR metodu, ali obrađene su i srodne iterativne metode poput Jacobijeve metode i Gauss-Seidelove metode. Iterativne metode se koriste pri rješavanju velikih rijetkih sustava linearnih jednadžbi.

AMS: 65R10; 65N22; 65F10

Ključne riječi: stacionarne iterativne metode, Jacobijeva metoda, Gauss-Seidelova metoda, SOR metoda

Abstract

It has been over fifty years since D.M. Young's original work on the successive overrelaxation (SOR)methods. This fundamental method now appears in all textbooks containing an introductory discussion of iterative solution methods.
In this survey, the accent is at SOR method, but we considered some other related methods. Iterative methods are used for solving large sparse linear systems.

AMS: 65R10; 65N22; 65F10

Key words: stationary iterative methods, Jacobi method, Gauss-Seidel method, SOR method

---

*Seminarski rad iz kolegija *Računarski praktikum 2*

† *dmenon@mathos.hr; www.mathos.hr/~dmenon*

1 Uvod

U svojem dosadašnjem radu smo vjerojatno mnogo puta naišli na problem rješavanja sustava linearnih jednadžbi

$$\mathbf{A}x = \mathbf{b}. \tag{1}$$

Svi smo upoznati s direktnim metodama rješavanja sustava (1). U slučajevima kada je broj jednadžbi i nepoznanica malen, ovaj problem smo u stanju riješiti vrlo brzo i pouzdano. U slučaju da je riječ o vrlo velikim sustavima linearnih jednadžbi dolazi do vrlo velikih problema, čak i uz današnja računala ovaj problem može postati vrlo težak.

Stoga su se vremenom razvile mnoge različite metode za rješavanje sustava (1), a u ovom radu ćemo se osvrnuti na grupu metoda koje se nazivaju *stacionarne iterativne metode*. Kao što samo ime kaže, iterativnim metodama pokušavamo poboljšati neku početnu aproksimaciju, tako da u svakoj iteraciji greška bude što manja. Primjenom odgovarajućeg kriterija zaustavljanja, nakon određenog broja iteracija, dobiveni vektor smatrat ćemo dovoljno dobrom aproksimacijom.

Postavlja se pitanje zašto se pored prilično djelotvorne metode Gaussovih eliminacija (vidi [3]) za rješavanje linearnih sustava razvio i veliki broj iterativnih metoda? Opće je poznato da rješavanje linearnog sustava pomoću Gaussovih eliminacija, u općenitom slučaju, zahtijeva zalihu u memoriji za skladištenje svih n^2 elemenata matrice $\mathbf{A}$ i broj potrebnih operacija reda n^3. Matrice koje se pojavljuju u praksi često imaju posebna svojstva. Najčešće se to svodi na veliki broj elemenata koji su jednaki nuli, pa govorimo o *rijetko popunjenim* matricama. Također je često raspored netrivijalnih elemenata pravilno raspoređen, pa govorimo o *vrpčastim* matricama. Gaussove eliminacije obično su u stanju samo djelomično iskoristiti ta svojstva. Ako primijenimo Gaussove eliminacije na takve matrice, u rezultirajućim trokutastim faktorima pojaviti će se mnogi netrivijalni elementi na mjestima gdje je originalna matrica imala nule. Nadalje, vrlo često se rijetka popunjenost matrice može iskoristiti pri množenju matrice s vektorom.

Zato su se razvile iterativne metode za rješavanje sustava (1) koje koriste samo množenje matrice i vektora, uz mali broj još dodatnih operacija. Takve metode mogu prilično nadmašiti Gaussove eliminacije i u memoriji i po broju izvršenih operacija.

Osnovni nedostatak iterativnih metoda je u tome što je klasa sustava linearnih jednadžbi za koje pojedine metode konvergiraju dosta uska. No, s druge strane, sreća je sto u praksi mnogi sustavi linearnih jednadžbi spadaju u klasu za koju iterativne metode konvergiraju.

2 Spektralni radijus

Spektralni radijus, ili najveća apsolutna vrijednost svojstvenih vrijednosti matrice, je od velike važnosti u analizi određenih iterativnih metoda. Sa $\sigma(\mathbf{A})$ označavamo skup svojstvenih vrijednosti od $\mathbf{A}$.

Definicija 2.1 *Spektralni radijus matrice* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *je*

$$\rho(\mathbf{A}) = \max_{\lambda\in\sigma(\mathbf{A})} |\lambda|.$$

Teorem 2.1 *Ako je* $\|\cdot\|$ *bilo koja matrična norma i* $\mathbf{A} \in \mathbb{R}^{n\times n}$*, tada je*

$$\rho(\mathbf{A}) \leq \|\mathbf{A}\|.$$

Dokaz: Neka je λ svojstvena vrijednost od $\mathbf{A}$, za koju je $|\lambda| = \rho(\mathbf{A})$ i neka je $\mathbf{v} \neq 0$ odgovarajući svojstveni vektor. Neka je $\mathbf{V} \in \mathbb{R}^{n\times n}$ matrica kojoj je svaki stupac jednak $\mathbf{v}$. Tada je $\mathbf{AV} = \lambda\mathbf{V}$ i vrijedi

$$|\lambda| \cdot \|\mathbf{V}\| = \|\lambda\mathbf{V}\| = \|\mathbf{AV}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{V}\|.$$

Budući da je $\mathbf{V} \neq 0$, imamo da je $\|\mathbf{V}\| > 0$, odakle slijedi da je $\rho(\mathbf{A}) \leq \|\mathbf{A}\|$. $\square$

Teorem 2.2 *Neka je* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *i* $\varepsilon > 0$ *dani broj. Tada postoji matrična norma* $\|\cdot\|$ *inducirana nekom vektorskom normom takva da je*

$$\|\mathbf{A}\|_{\varepsilon,\mathbf{A}} \leq \rho(\mathbf{A}) + \varepsilon.$$

Dokaz: Neka je $\mathbf{S}^{-1}\mathbf{AS} = \mathbf{J}$ Jordanova forma matrice $\mathbf{A}$. Definirajmo

$$\mathbf{D}_\varepsilon = \mathrm{diag}(1, \varepsilon, \ldots, \varepsilon^{n-1}),$$

i

$$\mathbf{J}_\varepsilon = \mathbf{D}_\varepsilon^{-1}\mathbf{J}\mathbf{D}_\varepsilon,$$

tako da vrijedi

$$\begin{aligned} (\mathbf{J}_\varepsilon)_{i,i} &= \varepsilon^{-(i-1)}\mathbf{J}_{i,i}\varepsilon^{i-1} = \mathbf{J}_{i,i} \\ (\mathbf{J}_\varepsilon)_{i,i+1} &= \varepsilon^{-(i-1)} \cdot 1 \cdot \varepsilon^{i} = \varepsilon. \end{aligned}$$

Matričnu normu, za proizvoljnu matricu $\mathbf{B} \in \mathbb{R}^{n\times n}$, definirat ćemo na sljedeći način.

$$\|\mathbf{B}\|_{\varepsilon,\mathbf{A}} = \|\mathbf{D}_\varepsilon^{-1}\mathbf{S}^{-1}\mathbf{B}\mathbf{S}\mathbf{D}_\varepsilon\|_1,$$

pri čemu je odgovarajuća vektorska norma koja ju inducira za proizvoljan vektor $\mathbf{v}$, oblika

$$\|\mathbf{v}\|_{\varepsilon,\mathbf{A}} = \|\mathbf{D}_\varepsilon^{-1}\mathbf{S}^{-1}\mathbf{v}\|_1,$$

što se lako može provjeriti iz niza jednakosti

$$\begin{aligned}
\max_{\mathbf{v}\neq 0} \frac{\|\mathbf{D}_\varepsilon^{-1}\mathbf{S}^{-1}\mathbf{B}\mathbf{v}\|_1}{\|\mathbf{D}_\varepsilon^{-1}\mathbf{S}^{-1}\mathbf{v}\|_1} &= \max_{\mathbf{w}\neq 0} \frac{\|\mathbf{D}_\varepsilon^{-1}\mathbf{S}^{-1}\mathbf{B}\mathbf{S}\mathbf{D}_\varepsilon\mathbf{w}\|_1}{\|\mathbf{w}\|_1} \\
&= \|\mathbf{D}_\varepsilon^{-1}\mathbf{S}^{-1}\mathbf{B}\mathbf{S}\mathbf{D}_\varepsilon\|_1 \\
&= \|\mathbf{B}\|_{\varepsilon,\mathbf{A}},
\end{aligned}$$

pri čemu je $\mathbf{w} = \mathbf{D}_\varepsilon^{-1}\mathbf{S}^{-1}\mathbf{v}$. Napokon, vrijedi

$$\|\mathbf{A}\|_{\varepsilon,\mathbf{A}} = \|\mathbf{J}_\varepsilon\|_1 \leq \rho(\mathbf{A}) + \varepsilon.$$

□

Teorem 2.3 *Neka je* $\mathbf{A} \in \mathbb{R}^{n\times n}$. *Tada je* $\lim_{k\to\infty} \mathbf{A}^k = 0$ *ako i samo ako je* $\rho(\mathbf{A}) < 1$.

Dokaz: Prvo pretpostavimo da je $\lim_{k\to\infty} \mathbf{A}^k = 0$. Neka je λ svojstvena vrijednost od $\mathbf{A}$ sa svojstvenim vektorom $\mathbf{v} \neq 0$. Budući da je $\mathbf{A}^k\mathbf{v} = \lambda^k\mathbf{v}$, za proizvoljnu normu $\|\cdot\|$ slijedi

$$0 = \lim_{k\to\infty} \|\lambda^k\mathbf{v}\| = \|\mathbf{v}\| \lim_{k\to\infty} |\lambda^k|.$$

Kako je $\|\mathbf{v}\| > 0$, vrijedi da je $\lim_{k\to\infty} |\lambda^k| = 0$, odakle slijedi da mora biti $|\lambda| < 1$ i to za proizvoljnu svojstvenu vrijednost od $\mathbf{A}$. Dakle, vrijedi i $\rho(\mathbf{A}) < 1$.

Obrnuto, pretpostavimo da je $\rho(\mathbf{A}) < 1$ i tada prema Teoremu 2.2 postoji matrična norma $\|\cdot\|$, takva da je za $0 < \varepsilon < 1 - \rho(\mathbf{A})$

$$\|\mathbf{A}\| \leq \rho(\mathbf{A}) + \varepsilon < 1.$$

Odatle slijedi da je

$$\lim_{k\to\infty} \|\mathbf{A}^k\| \leq \lim_{k\to\infty} \|\mathbf{A}\|^k = 0.$$

Budući da su na $\mathbb{R}^{n\times n}$ sve matrične norme ekvivalentne, znači da je i $\lim_{k\to\infty} \|\mathbf{A}^k\|_F = 0$, odakle slijedi da svi elementi od $\mathbf{A}^k$ moraju težiti k nuli. □

Korolar 2.1 *Neka je $\|\cdot\|$ matrična norma na $\mathbb{R}^{n\times n}$. Tada je*

$$\rho(\mathbf{A}) = \lim_{k\to\infty} \|\mathbf{A}^k\|^{1/k}$$

za svaki $\mathbf{A} \in \mathbb{R}^{n\times n}$.

Dokaz: Budući da je prema definiciji i Teoremu 2.1 $\rho(\mathbf{A})^k = \rho(\mathbf{A}^k) \leq \|\mathbf{A}^k\|$, imamo da je $\rho(\mathbf{A}) \leq \|\mathbf{A}^k\|^{1/k}$ za sve $k = 1, 2, \ldots$ Za proizvoljni $\varepsilon > 0$, matrica $\tilde{\mathbf{A}} = [\rho(\mathbf{A})+\varepsilon]^{-1}\mathbf{A}$ ima spektralni radijus strogo manji od jedan, pa je prema Teoremu 2.3 $\lim\limits_{k\to\infty} \|\tilde{\mathbf{A}}^k\| = 0$. Dakle, postoji broj K koji ovisi o ε, takav da je za svaki $k \geq K$, $\|\tilde{\mathbf{A}}^k\| < 1$. što je ekvivalentno tvrdnji $\|\mathbf{A}^k\| \leq [\rho(\mathbf{A}) + \varepsilon]^k$ ili $\|\mathbf{A}^k\|^{1/k} \leq \rho(\mathbf{A}) + \varepsilon$ za sve $k \leq K$. Prema tome imamo

$$\rho(\mathbf{A}) \leq \|\mathbf{A}^k\|^{1/k} \leq \rho(\mathbf{A}) + \varepsilon$$

za sve $k \leq K$, a budući da to vrijedi za sve $\varepsilon > 0$, slijedi da $\lim\limits_{k\to\infty} \|\mathbf{A}^k\|^{1/k}$ postoji i da je jedak $\rho(\mathbf{A})$. □

3 Opći iterativni postupak

Neka je dana regularna matrica $\mathbf{A} \in \mathbb{R}^{n\times n}$ i neka su $\mathbf{N}, \mathbf{P} \in \mathbb{R}^{n\times n}$ takve da je

$$\mathbf{A} = \mathbf{N} + \mathbf{P}.$$

Tada dani sustav $\mathbf{A}\mathbf{x} = \mathbf{b}$ možemo zapisati u obliku

$$\mathbf{N}\mathbf{x} = -\mathbf{P}\mathbf{x} + \mathbf{b},$$

pri čemu je matrica $\mathbf{N}$ regularna. Ako nam je dana aproksimacija $\mathbf{x}^{(k)}$, novu aproksimaciju $\mathbf{x}^{(k+1)}$ možemo dobiti kao rješenje jednadžbe

$$\mathbf{N}\mathbf{x}^{(k+1)} = -\mathbf{P}\mathbf{x}^{(k)} + \mathbf{b}, \quad k = 0, 1, 2, \ldots \tag{2}$$

U tom slučaju imamo iteracije oblika

$$\mathbf{x}^{(k+1)} = -\mathbf{N}^{-1}\mathbf{P}\mathbf{x}^{(k)} + \mathbf{N}^{-1}\mathbf{b}$$

kod kojih se traži fiksna točka (vidi [3]). Ovako dobiveni iterativni postupak možemo općenito zapisati kao

$$\mathbf{x}^{(k+1)} = \mathbf{M}\mathbf{x}^{(k)} + \mathbf{d},$$

gdje je $\mathbf{M} = \mathbf{N}^{-1}\mathbf{P}$ i $\mathbf{d} = \mathbf{N}^{-1}\mathbf{b}$.

Teorem 3.1 *Neka je* $\mathbf{M} \in \mathbb{R}^{n\times n}$ *matrica takva da za neku matričnu i s njom kompatibilnu vektorsku normu vrijedi*

$$\|\mathbf{My} - \mathbf{Mz}\| \leq \alpha\|\mathbf{y} - \mathbf{z}\|, \qquad \textit{za svaki } \mathbf{y}, \mathbf{z} \in \mathbb{R}^n,$$

pri čemu je $\alpha \in [0, 1)$. *Tada sustav* $\mathbf{x} = \mathbf{Mx} + \mathbf{d}, \quad \mathbf{d} \in \mathbb{R}^n$, *ima jedinstveno rješenje* $\mathbf{x}^* \in \mathbb{R}^n$ *i to rješenje je granična vrijednost niza vektora* $\mathbf{x}^{(0)}$, $\mathbf{x}^{(1)}$, $\mathbf{x}^{(2)}$,*... određenog prema*

$$\mathbf{x}^{(k+1)} = \mathbf{Mx}^{(k)} + \mathbf{d}, \quad k = 0, 1, 2, \ldots \tag{3}$$

za proizvoljan $\mathbf{x}^{(0)} \in \mathbb{R}^n$.

Kako za kompatibilne vektorske i matrične norme vrijedi

$$\|\mathbf{My} - \mathbf{Mz}\| \leq \|\mathbf{M}\|\|\mathbf{y} - \mathbf{z}\|, \quad \mathbf{y}, \mathbf{z} \in \mathbb{R}^n,$$

slijedi

Teorem 3.2 *Neka je za neku matričnu normu* $\|\mathbf{M}\| \geq 1$. *Tada sustav* $\mathbf{x} = \mathbf{Mx} + \mathbf{d}$ *ima jedinstveno rješenje koje je granična vrijednost iterativnog niza (3), za proizvoljan* $\mathbf{x}^{(0)}$.

Dokaz: Uzmemo $\alpha = \|\mathbf{M}\|$ i tvrdnja teorema slijedi direktno iz Teorema 3.1. □

Uvjet $\|\mathbf{M}\| < 1$ je samo dovoljan za konvergenciju iterativnog postupka (3) za proizvoljni $\mathbf{x}^{(0)}$, tj. ako za neku normu vrijedi $\|\mathbf{M}\| \geq 1$ to ne znači da iterativni postupak ne konvergira.

Teorem 3.3 *Neka je niz vektora* $\mathbf{x}^{(0)}$, $\mathbf{x}^{(1)}$, $\mathbf{x}^{(2)}$, *... generiran s*

$$\mathbf{x}^{(k+1)} = \mathbf{Mx}^{(k)} + \mathbf{d}, \quad k = 0, 1, 2, \ldots,$$

gdje je $\mathbf{M} \in \mathbb{R}^{n\times n}$, $\mathbf{x} \in \mathbb{R}^n$. *Dani niz konvergira prema jedinstvenom rješenju* $\mathbf{x}^*$*sustava* $\mathbf{x} = \mathbf{Mx} + \mathbf{d}$ *za svaki početni vektor* $\mathbf{x}^{(0)}$ *ako i samo ako je* $\rho(\mathbf{M}) < 1$.

Dokaz: Najprije pretpostavimo da je $\rho(\mathbf{M}) < 1$. Tada postoji $\varepsilon > 0$ takav da je $\rho(\mathbf{M}) + \varepsilon < 1$. Prema Teoremu 2.2 postoji matrična norma $\|\cdot\|$ takva da je

$$\|\mathbf{M}\| < \rho(\mathbf{M}) + \varepsilon < 1,$$

pa prema Teoremu 3.2 slijedi da metoda konvergira.

Pretpostavimo sada obrat, to jest, da niz $\mathbf{x}^{(k)}$ konvergira za svako $\mathbf{x}^{(0)}$ i $\mathbf{d}$. Sljedeće, pretpostavimo da je u tom slučaju $\rho(\mathbf{M}) \geq 1$. To znači da postoji svojstvena vrijednost λ od $\mathbf{M}$, takva da je $|\lambda| \geq 1$. Neka je $\mathbf{d} \neq 0$ svojstveni vektor od $\mathbf{M}$, koji pripada λ i $\mathbf{x}^{(0)} = 0$. Tada imamo

$$\begin{aligned}
\mathbf{x}^{(1)} &= \mathbf{M}\mathbf{x}^{(0)} + \mathbf{d} = \mathbf{M} \cdot \mathbf{0} + \mathbf{d} = \mathbf{d} \\
\mathbf{x}^{(2)} &= \mathbf{M}\mathbf{x}^{(1)} + \mathbf{d} = \mathbf{M}\mathbf{d} + \mathbf{d} = (\lambda + 1)\mathbf{d} \\
\mathbf{x}^{(3)} &= \mathbf{M}\mathbf{x}^{(2)} + \mathbf{d} = (\lambda + 1)\mathbf{M}\mathbf{d} + \mathbf{d} = (\lambda^2 + \lambda + 1)\mathbf{d} \\
&\vdots \\
\mathbf{x}^{(k)} &= \mathbf{M}\mathbf{x}^{(k-1)} + \mathbf{d} = (\lambda^{k-1} + \cdots + \lambda + 1)\mathbf{d}.
\end{aligned}$$

Prema tome za iteraciju u k-tom koraku vrijedi

$$\mathbf{x}^{(k)} = \begin{cases} \frac{\lambda^k - 1}{\lambda - 1}\mathbf{d}, & \lambda \neq 1 \\ k\mathbf{d}, & \lambda = 1 \end{cases}$$

što u oba slučaja divergira kada k teži ka beskonačnosti. Dakle, našli smo niz koji za određene $\mathbf{x}^{(0)}$ i $\mathbf{d}$ divergira, što je u kontradikciji sa pretpostavkom. Znači, da mora biti $\rho(\mathbf{M}) < 1$. $\square$

4 Jacobijeva metoda

Neka je dan sustav linearnih jednadžbi

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{4}$$

gdje je matrica $\mathbf{A} \in \mathbb{R}^{n \times n}$ regularna i

$$a_{ii} \neq 0, \quad i = 1, 2, \ldots, n.$$

Jacobijev iterativni postupak za rješavanje sustava (4) definiran je sa

$$\mathbf{x}^{(k+1)} = -(\mathbf{L} + \mathbf{U})\mathbf{x}^{(k)} + \mathbf{D}^{-1}\mathbf{b}, \quad k = 0, 1, 2, \ldots,$$

odnosno

$$\mathbf{x}^{(k+1)} = \mathbf{B}_J\mathbf{x}^{(k)} + \mathbf{d}, \quad k = 0, 1, 2, \ldots,$$

gdje je

$$\begin{aligned}
\mathbf{D} &= \operatorname{diag}(a_{11}, a_{22}, \ldots, a_{nn}), \\
\mathbf{B}_J &= -(\mathbf{L} + \mathbf{U}), \\
\mathbf{d} &= \mathbf{D}^{-1}\mathbf{b}.
\end{aligned} \tag{5}$$

kako je $\mathbf{D}^{-1} = \operatorname{diag}(a_{11}^{-1}, a_{22}^{-1}, \ldots, a_{nn}^{-1})$, to je

$$\mathbf{B}_J = \begin{bmatrix} 0 & -\frac{a_{12}}{a_{11}} & \cdots & -\frac{a_{1n}}{11} \\ -\frac{21}{22} & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & -\frac{a_{n-1,n}}{a_{n-1,n-1}} \\ -\frac{a_{n1}}{a_{nn}} & \cdots & -\frac{a_{n,n-1}}{a_{nn}} & 0 \end{bmatrix}, \quad \mathbf{d} = \begin{bmatrix} \frac{b_1}{a_{11}} \\ \frac{b_2}{a_{22}} \\ \vdots \\ \frac{b_n}{a_{nn}} \end{bmatrix}.$$

Prema tome, komponente vektora $\mathbf{x}^{(k+1)}$ izračunavaju se prema sljedećoj formuli

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j^{(k)} \right), \quad i = 1, \ldots, n. \tag{6}$$

Primijetimo da kod skladištenja vektora u memoriji, novi vektor $\mathbf{x}^{(k+1)}$ ne može se prepisati preko starog vektora $\mathbf{x}^{(k)}$ sve dok sve komponente vektora $\mathbf{x}^{(k+1)}$ nisu izračunate. Dakle, prije kraja jedne iteracije, oba vektora moramo posebno skladištiti.

U nastavku će nas zanimati konvergencija Jacobijeve metode, međutim, samo za matrice s posebnim svojstvima postoje rezultati koji govore o tome. Stoga definirajmo pojam *strogo dijagonalne dominantnosti.*

Definicija 4.1 *Kažemo da je matrica* $\mathbf{A}$ *strogo dijagonalno dominantna ako vrijedi*

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

Za takve matrice vrijedi sljedeći teorem.

Teorem 4.1 *Neka je* $\mathbf{A}$ *strogo dijagonalno dominantna matrica. Tada Jacobijeva metoda konvergira za svaku početnu aproksimaciju* $\mathbf{x}^{(0)}$.

Dokaz: Prema Definiciji 4.1 i (5) vrijedi

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

pa prema Teoremu 3.3, Jacobijeva metoda konvergira za svaku početnu aproksimaciju $\mathbf{x}^{(0)}$. $\square$

5 Gauss-Seidelova metoda

Jacobijevu metodu možemo modificirati tako da tekuće aproksimacije prethodnih komponenti od $\mathbf{x}^{(k+1)}$ koristimo da bi se izračunala njegova sljedeća komponenta. Time se omogućava da komponente od $\mathbf{x}^{(k+1)}$ prebrišu stare vrijednosti komponenata od $\mathbf{x}^{(k)}$, čim se izračunaju. Stoga nam nije potreban dodatni pomoćni vektor za pamćenje komponenti prethodne iteracije. S takvim modifikacijama Jacobijeve metode, dobivamo Gauss-Seidelovu metodu.

Dakle, dan je sustav jednadžbi

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{7}$$

gdje je matrica $\mathbf{A} \in \mathbb{R}^{n\times n}$ regularna i

$$a_{ii} \neq 0, \quad i = 1, 2, \ldots, n.$$

Gauss-Seidelov iterativni postupak za rješavanje sustava (7) definiran je sa

$$\mathbf{x}^{(k+1)} = -\mathbf{L}\mathbf{x}^{(k+1)} - \mathbf{U}\mathbf{x}^{(k)} + \mathbf{D}^{-1}\mathbf{b}, \quad k =, 0, 1, 2, \ldots,$$

odnosno sa

$$\mathbf{x}^{(k+1)} = \mathbf{B}_{GS}\mathbf{x}^{(k)} + \mathbf{d}, \quad k =, 0, 1, 2, \ldots,$$

gdje je

$$\begin{aligned} \mathbf{B}_{GS} &= -(\mathbf{I}+\mathbf{L})^{-1}\mathbf{U} \\ \mathbf{d} &= \mathbf{D}^{-1}(\mathbf{I}+\mathbf{L})^{-1}\mathbf{b}. \end{aligned}$$

Prema tome komponente vektora $\mathbf{x}^{(k+1)}$ izračunavaju se na sljedeći način

$$x_i^{(k+1)} = \frac{1}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}\right), \quad i = 1, 2, \ldots, n.$$

Stoga kažemo da je Jacobijeva metoda *totalno koračna*, dok se za Gauss-Seidelovu metodu kaže da je *jednokoračna*.

Uzevši u obzir navedene razlike između Jacobijeve i Gauss-Seidelove metode, razumno je očekivati da Gauss-Seidelova metoda konvergira brže od Jacobijeve. U većini slučajeva to je točno, ali ne uvijek.

Primjer 5.1 *Neka je*

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

dana matrica. Tada matrice $\mathbf{B}_J$ i $\mathbf{B}_{GS}$ glase

$$\mathbf{B}_J = \begin{bmatrix} 0 & 2 & -2 \\ 1 & 0 & 1 \\ 2 & 2 & 0 \end{bmatrix} \; i \quad \mathbf{B}_{GS} = \begin{bmatrix} 0 & 2 & -2 \\ 0 & 2 & -1 \\ 0 & 8 & -6 \end{bmatrix}.$$

Slijedi

$$\rho(\mathbf{B}_J) = 0, \quad \rho(\mathbf{B}_{GS}) = 2(1+\sqrt{2}),$$

što znači da Jacobijeva metoda konvergira, a Gauss-Seidelova divergira (vidi [2]).

Primjer 5.2 *Neka je*

$$\mathbf{A} = \begin{bmatrix} 1 & \frac{1}{2} & -\frac{1}{2} \\ -1 & 1 & -1 \\ -\frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix}$$

dana matrica. Tada za matrice $\mathbf{B}_J$ i $\mathbf{B}_{GS}$ vrijedi

$$\rho(\mathbf{B}_J) = 1, \quad \rho(\mathbf{B}_{GS}) = \frac{1}{2},$$

pa se sada lako vidi da Gauss-Seidelova metoda konvergira, a Jacobijeva metoda divergira.

O konvergenciji Gauss-Seidelove metode govore sljedeći teoremi.

Teorem 5.1 *Neka je $\mathbf{A}$ strogo dijagonalno dominantna matrica ili ireducibilna dijagonalno dominantna matrica. Tada Gauss-Seidelova metoda konvergira za proizvoljan početni vektor $\mathbf{x}^{(0)}$.*

Teorem 5.2 *Ako je matrica $\mathbf{A} \in \mathbb{R}^{n\times n}$ simetrična i pozitivno definitna, onda Gauss-Seidelova metoda za rješavanje sustava (7) konvergira za svaki početni vektor $\mathbf{x}^{(0)}$.*

Teorem 5.3 *Neka je $\mathbf{A}$ trodijagonalna matrica. Tada ili i Jacobijeva i Gauss-Seidelova metoda za rješavanje sustava (7) konvergiraju ili obje divergiraju. Ako obje konvergiraju, Gauss-Seidelova metoda konvergira brže.*

6 SOR metoda

Začetnik metode sukcesivne nadrelaksacije (engl. *successive overrelaxation method*), skraćeno SOR, je svakako D. M. Young koji je 1950. u svojoj doktorskoj disertaciji ukazao na prednosti dobivene u nekim slučajevima korištenjem fiksnog relaksacijskog parametra ω (vidi [4]). Svakako treba naglasiti da se pojam relaksacijskog parametra mogao naći i u ranijim radovima.

Od svih stacionarnih iterativnih metoda, SOR metoda se najčešće koristi. U današnje vrijeme su stacionarne iterativne metode većinom zamijenjene *nestacionarnim iterativnim metodama*. No, mnoge stacionarne iterativne metode se koriste kod prekondicioniranja. Prekondicioniranje je bilo kakvo modificiranje originalnog linearnog sustava, koje na neki način olakšava rješavanje danog sustava. Odnosno, želi se postojeći sustav transformirati u ekvivalentan sustav koji ima npr. bolja spektralna svojstva. I do sada smo vidjeli kako je spektralni radijus kod stacionarnih iterativnih metoda vrlo važan. Njegova važnost je vrlo velika i kod nestacionarnih iterativnih metoda.

SOR metoda je modifikacija Gauss-Seidelove metoda sa svrhom poboljšanja konvergencije. Cilj tih modifikacija nije u tome da proširi klasu matrica za koje iterativna metoda konvergira, već da ubrza konvergenciju. Kasnije ćemo na primjerima vidjeti koliko se značajno konvergencija može ubrzati.

Ponovno promatramo sustav

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{8}$$

gdje je $\mathbf{A} \in \mathbb{R}^{n\times n}$ uz pretpostavku

$$a_{ii} \neq 0, \quad i = 1, 2, \ldots, n.$$

Matricu $\mathbf{A}$ rastavimo na

$$\mathbf{A} = \mathbf{N} + \mathbf{P}$$

gdje je

$$\begin{aligned} \mathbf{N} &= \frac{1}{\omega}\mathbf{D} + \tilde{\mathbf{L}}, \\ \mathbf{P} &= \frac{\omega - 1}{\omega}\mathbf{D} + \tilde{\mathbf{U}}. \end{aligned}$$

Pri tome su matrice $\mathbf{D}$, $\tilde{\mathbf{L}}$ i $\tilde{\mathbf{U}}$ definirane kao $\mathbf{D} = \mathrm{diag}(a_{11}, a_{22}, \ldots, a_{nn})$, $\tilde{\mathrm{L}}$ je strogo donji trokutasti dio, a $\tilde{\mathrm{U}}$ strogo gornji trokutasti dio. Dok je ω realan broj različit od nule koji zovemo relaksacijski parametar.

Sada iz (8) dobivamo

$$\mathbf{Nx} = -\mathbf{Px} + \mathbf{b}$$

odnosno

$$\begin{aligned} \mathbf{D}\left(\frac{1}{\omega}\mathbf{I} + \mathbf{L}\right)\mathbf{x}^{(k+1)} &= -\mathbf{D}\left(\frac{\omega-1}{\omega}\mathbf{I} + \mathbf{U}\right)\mathbf{x}^{(k)} + \mathbf{b} \\ (\mathbf{I} + \omega\mathbf{L})\,\mathbf{x}^{(k+1)} &= ((1-\omega)\,\mathbf{I} - \omega\mathbf{U})\,\mathbf{x}^{(k)} + \mathbf{D}^{-1}\omega\mathbf{b}, \end{aligned}$$

pa je

$$\mathbf{x}^{(k+1)} = (\mathbf{I}+\omega\mathbf{L})^{-1}((1-\omega)\mathbf{I}-\omega\mathbf{U})\mathbf{x}^{(k)} + (\mathbf{I}+\omega\mathbf{L})^{-1}\mathbf{D}^{-1}\omega\mathbf{b}, \quad k = 0, 1, 2, \dots$$

Uvedimo oznake

$$\begin{aligned} \mathbf{B}_\omega &= (\mathbf{I}+\omega\mathbf{L})^{-1}((1-\omega)\mathbf{I}-\omega\mathbf{U}) \\ \mathbf{d} &= (\mathbf{I}+\omega\mathbf{L})^{-1}\mathbf{D}^{-1}\omega\mathbf{b}, \end{aligned}$$

te slijedi

$$\mathbf{x}^{(k+1)} = \mathbf{B}_\omega\mathbf{x}^{(k)} + \mathbf{d}, \quad k = 0, 1, 2, \dots$$

Slijedi

$$x_i^{(k+1)} = (1-\omega)x_i^{(k)} - \frac{\omega}{a_{ii}}\left(\sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} + \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}\right) + \omega\frac{b_i}{a_{ii}}, \quad i = 1, 2, \dots, n.$$

Razlikujemo tri slučaja koja ovise o broju ω. Izbor $\omega = 1$ daje Gauss-Seidelovu metodu, $\omega < 1$ se zove podrelaksacijski parametar (engl. *underrelaxation*), a $\omega > 1$ nadrelaksacijski parametar (engl. *overrelaxation*).

Teorem 6.1 (Kahan) *Za proizvoljnu matricu* $\mathbf{A}$ *vrijedi*

$$\rho(\mathbf{B}_\omega) \geq |\omega - 1|$$

za svaki ω.

Dokaz: Pošto su svojstvene vrijednosti $\lambda_1, \dots, \lambda_n$ od $\mathbf{B}_\omega$ nultočke karakterističnog polinoma, vrijedi

$$\prod_{j=1}^{n} \lambda_j = \det \mathbf{B}_\omega.$$

Odavde, zbog pravila za množenje determinanti i pošto su $\mathbf{I} + \omega\mathbf{L}$ i $(1-\omega)\mathbf{I} - \omega\mathbf{U}$ trokutaste matrice, slijedi

$$\prod_{j=1}^{n} \lambda_j = \det(\mathbf{I}+\omega\mathbf{L})^{-1}\det((1-\omega)\mathbf{I}-\omega\mathbf{U}) = (1-\omega)^n.$$

Nadalje slijedi

$$\rho(\mathbf{B}_\omega) \geq |\omega - 1|.$$

$\square$

Teorem 6.2 (Ostrowski) *Neka je* $\mathbf{A} \in \mathbb{R}^{n \times n}$ *simetrična pozitivno definitna matrica i pretpostavimo da je* $0 < \omega < 2$. *Tada SOR metoda konvergira prema jedinstvenom rješenju od (8).*

Dokaz: Neka je λ proizvoljna svojstvena vrijednost matrice $\mathbf{B}_\omega$ i neka je $\mathbf{y} \neq 0$ odgovarajući svojstveni vektor. Tada vrijedi

$$((1-\omega)\mathbf{I} - \omega\mathbf{U})\,\mathbf{y} = \lambda(\mathbf{I} + \omega\mathbf{L})\mathbf{y}. \tag{9}$$

Za lijevu stranu jednakosti (9), direktnim računom dobivamo

$$(2-\omega)\mathbf{I} - \omega\mathbf{A} - \omega(\mathbf{U} - \mathbf{L}) = 2\left((1-\omega)\mathbf{I} - \omega\mathbf{U}\right),$$

a za desnu stranu vrijedi

$$(2-\omega)\mathbf{I} - \omega\mathbf{A} - \omega(\mathbf{U} - \mathbf{L}) = 2\left(\mathbf{I} + \omega\mathbf{L}\right).$$

Poslije skalarnog množenja sa $\mathbf{y}$, iz (9) dobivamo

$$\lambda = \frac{(2-\omega)d - \omega a + \imath\omega s}{(2-\omega)d + \omega a + \imath\omega s}$$

gdje je

$$a := (\mathbf{A}\mathbf{y}, \mathbf{y}), \quad d := (\mathbf{D}\mathbf{y}, \mathbf{y}), \quad s := \imath(\mathbf{U}\mathbf{y} - \mathbf{L}\mathbf{y}, \mathbf{y}).$$

Pošto je $\mathbf{A}$ pozitivno definitna, imamo $a > 0$ i $d > 0$ i pošto je $\mathbf{A}$ simetrična, $s \in \mathbb{R}$. Iz

$$|(2-\omega)d - \omega a| < |(2-\omega)d + \omega a|$$

za $0 < \omega < 2$ sada možemo zaključiti da je $|\lambda < 1|$. Stoga konvergencija SOR metode za $0 < \omega < 2$ slijedi iz Teorema 3.3. $\square$

Teorem 6.3 *SOR metoda ne konvergira za* $\omega \leq 0$ *i* $\omega \geq 2$.

Dokaz: Očito je da za matricu koraka

$$\mathbf{B}_\omega = (\mathbf{I} + \omega\mathbf{L})^{-1}((1-\omega)\mathbf{I} - \omega\mathbf{U})$$

SOR metode vrijedi

$$\begin{aligned} \det \mathbf{B}_\omega &= (\det(\mathbf{I} + \omega\mathbf{L}))^{-1} \det((1-\omega)\mathbf{I} - \omega\mathbf{U}) \\ &= (\det \mathbf{I})^{-1}(1-\omega)^n \det \mathbf{I} \\ &= (1-\omega)^n, \end{aligned}$$

jer su $\mathbf{I}+\omega\mathbf{L}$ i $(1-\omega)\mathbf{I}-\omega\mathbf{U}$ donja i gornja trokutasta matrica, a $\mathbf{L}$ i $\mathbf{U}$ strogo donja odnosno strogo gornja trokutasta matrica. Kako je

$$\rho(\mathbf{B}_\omega)^n \geq \prod_{j=1}^{n} |\lambda_i| = |\det \mathbf{B}_\omega| = |1-\omega|^n,$$

gdje su λ_j svojstvene vrijednosti matrice $\mathbf{B}_\omega$, slijedi

$$\rho(\mathbf{B}_\omega) \geq |1-\omega|.$$

Prema Teoremu 3.3 SOR metoda će konvergirati ako i samo ako je $\rho(\mathbf{B}_\omega) < 1$, pa ako imamo da je $|1-\omega| \geq 1$, tada SOR metoda neće konvergirati. Ovaj uvjet biti će ispunjen ako i samo ako je $\omega \leq 0$ i $\omega \geq 2$. $\square$

6.1 Određivanje relaksacijskog parametra

Određivanje optimalnog relaksacijskog parametra ω_{opt} je u pravilu dosta komplicirano. U mnogim praktičnim slučajevima određivanje optimalnog relaksacijskog parametra bi zahtijevao više rada računala nego provođenje iterativnog postupka. Srećom, pokazalo se da je i relativno dobra procjena vrijednosti relaksacijskog parametra i više nego dovoljna da se konvergencija iterativne metode značajno ubrza. Nadalje, za neke klase matrica se zna vrijednost optimalnog relaksirajućeg parametra, dok se za neke druge klase matrica eksperimentalno utvrdila vrijednost relaksirajućeg parametra koja značajno ubrzava konvergenciju.

Pokazati ćemo kako se relaksacijski parametar ω računa za klasu konzistentno uređenih matrica (engl. *consistently ordered matrices*).

Definicija 6.1 *Matrica $\mathbf{T}$ ima "svojstvo $\mathcal{A}$" ako postoji matrica permutacija $\mathbf{P}$ takva da je*

$$\mathbf{P}\mathbf{T}\mathbf{P}^T = \begin{bmatrix} \mathbf{D}_1 & \mathbf{M}_1 \\ \mathbf{M}_2 & \mathbf{D}_2 \end{bmatrix}, \tag{10}$$

gdje su $\mathbf{D}_1$ i $\mathbf{D}_2$ dijagonalne matrice.

Mnoge matrice dobivene diskretizacijom parcijalnih diferencijalnih jednadžbi zadovoljavaju "svojstvo $\mathcal{A}$". Na primjer, takva matrica je matrica koja dolazi iz aproksimacije s pet točaka Poissonove jednadžbe na kvadratu, pomoću konačnih diferencija. Ako čvorove na mreži numeriramo na crveno-crni način (vidi [4]), poput polja na šahovskoj ploči, tada dobivamo matricu oblika (10). Dakle, isplati se posebno promatrati matrice sa "svojstvom $\mathcal{A}$" i njihovo ponašanje u SOR procesu.

Definicija 6.2 *Neka je* $\mathbf{B}_J(\alpha) = -\alpha\mathbf{L} - \alpha^{-1}\mathbf{U}$. *Tada je* $\mathbf{B}_J(1) = \mathbf{B}_J$ *matrica koraka za Jacobijevu metodu.*

Najvažnija činjenica o matricama sa "svojstvom $\mathcal{A}$" dana je sljedećim teoremom.

Teorem 6.4 *Za svaku* $n \times n$ *matricu sa "svojstvom* $\mathcal{A}$*", za koju je* $a_{ii} \neq 0, \quad i = 1, 2, \ldots, n$, *postoji matrica permutacija* $\mathbf{P}$ *takva da je dekompozicija* $\tilde{\mathbf{A}} = \mathbf{D}(\mathbf{I} + \mathbf{L} + \mathbf{U})$, *permutirane matrice* $\tilde{\mathbf{A}} = \mathbf{P}\mathbf{A}\mathbf{P}^T$ *ima sljedeće svojstvo:*

Svojstvene vrijednosti matrice

$$\mathbf{B}_J(\alpha) = -\alpha\mathbf{L} - \alpha^{-1}\mathbf{U}, \quad \alpha \in \mathbb{C}, \quad \alpha \neq 0,$$

ne ovise o α.

Dokaz: Po Definiciji 6.1 postoji matrica permutacija $\mathbf{P}$ takva da je

$$\mathbf{P}\mathbf{A}\mathbf{P}^T = \begin{bmatrix} \mathbf{D}_1 & \mathbf{M}_1 \\ \mathbf{M}_2 & \mathbf{D}_2 \end{bmatrix} = \mathbf{D}(\mathbf{I} + \mathbf{L} + \mathbf{U}),$$

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 \end{bmatrix}, \quad \mathbf{L} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{D}_2^{-1}\mathbf{M}_2 & \mathbf{0} \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} \mathbf{0} & \mathbf{D}_1^{-1}\mathbf{M}_1 \\ \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

Ovdje su $\mathbf{D}_1$ i $\mathbf{D}_2$ regularne dijagonalne matrice. Za $\alpha \neq 0$ sada imamo

$$\mathbf{B}_J(\alpha) = -\begin{bmatrix} \mathbf{0} & \alpha^{-1}\mathbf{D}_1^{-1}\mathbf{M}_1 \\ \alpha\mathbf{D}_2^{-1}\mathbf{M}_2 & \mathbf{0} \end{bmatrix} = -\mathbf{S}_\alpha \begin{bmatrix} \mathbf{0} & \mathbf{D}_1^{-1}\mathbf{M}_1 \\ \mathbf{D}_2^{-1}\mathbf{M}_2 & \mathbf{0} \end{bmatrix} \mathbf{S}_\alpha^{-1} = \mathbf{S}_\alpha \mathbf{B}_J(1) \mathbf{S}_\alpha^{-1},$$

gdje je $\mathbf{S}_\alpha$ regularna dijagonalna matrica

$$\mathbf{S}_\alpha = \begin{bmatrix} \mathbf{I}_1 & \mathbf{0} \\ \mathbf{0} & \alpha\mathbf{I}_2 \end{bmatrix},$$

a $\mathbf{I}_1$ i $\mathbf{I}_2$ jedinične matrice. Matrice $\mathbf{B}_J(\alpha)$ i $\mathbf{B}_J(1)$ su slične, pa slijedi da imaju jednake svojstvene vrijednosti. $\square$

Definicija 6.3 *Neka je* $\mathbf{A}$ *bilo koja matrica takva da je* $\mathbf{A} = \mathbf{D} + \tilde{\mathbf{L}} + \tilde{\mathbf{U}}$ *i*

$$\mathbf{B}_J(\alpha) = -\alpha\mathbf{D}^{-1}\tilde{\mathbf{L}} - \alpha^{-1}\mathbf{D}^{-1}\tilde{\mathbf{U}} = -\alpha\mathbf{L} - \alpha^{-1}\mathbf{U}.$$

Ako svojstvene vrijednosti matrice $\mathbf{B}_J(\alpha)$ *ne ovise o* α, *tada matricu* $\mathbf{A}$ *zovemo* ***konzistentno uređena matrica****.*

Teorem 6.5 (Young, Varga) *Neka je* $\mathbf{A}$ *konzistentno uređena matrica i* $\omega \neq 0$. *Tada vrijedi:*

(a) Svojstvene vrijednosti matrice $\mathbf{B}_J = -\mathbf{L} - \mathbf{U}$ *pojavljuju se u* $\pm$ *parovima.*

(b) Ako je μ *svojstvena vrijednost matrice* $\mathbf{B}_J$ *i*

$$(\lambda + \omega - 1)^2 = \lambda \omega^2 \mu^2, \tag{11}$$

tada je λ *svojstvena vrijednost matrice* $\mathbf{B}_\omega$.

(c) Obratno, ako je $\lambda \neq 0$ *svojstvena vrijednost matrice* $\mathbf{B}_\omega$ *i vrijedi (11), tada je* μ *svojstvena vrijednost matrice* $\mathbf{B}_J$.

Dokaz:

(a) Konzistentna uređenost povlači da svojstvene vrijednosti matrice $\mathbf{B}_J(\alpha)$ ne ovise o α, pa matrice $\mathbf{B}_J = \mathbf{B}_J(1)$ i $\mathbf{B}_J(-1) = \mathbf{B}_J(1)$ imaju iste svojstvene vrijednosti. Stoga se one pojavljuju u $\pm$ parovima.

(b) Jer je $\det(\mathbf{I} + \omega \mathbf{L}) = 1$ za svaki ω, slijedi

$$\begin{aligned}
\det(\lambda \mathbf{I} - \mathbf{B}_\omega) &= \det\left((\mathbf{I} + \omega \mathbf{L})(\lambda \mathbf{I} - \mathbf{B}_\omega)\right) \\
&= \det\left(\lambda \mathbf{I} + \lambda \omega \mathbf{L} - (1-\omega)\mathbf{I} + \omega \mathbf{U}\right) \\
&= \det\left((\lambda + \omega - 1)\mathbf{I} + \lambda \omega \mathbf{L} + \omega \mathbf{U}\right).
\end{aligned} \tag{12}$$

Neka je μ svojstvena vrijednost matrice $\mathbf{B}_J = -\mathbf{L} - \mathbf{U}$ i λ rješenje jednakosti (11). Tada je

$$\lambda + \omega - 1 = \sqrt{\lambda}\omega\mu \quad \text{ili} \quad \lambda + \omega - 1 = -\sqrt{\lambda}\omega\mu.$$

Zbog *(a)*, a bez smanjenja općenitosti možemo pretpostaviti da je

$$\lambda + \omega - 1 = \sqrt{\lambda}\omega\mu.$$

Ako je $\lambda = 0$, tada je $\omega = 1$, pa iz (12) slijedi

$$\det(0 \cdot \mathbf{I} - \mathbf{B}_1)) \det(\mathbf{U}) = 0,$$

tj. λ je svojstvena vrijednost matrice $\mathbf{B}_\omega$. Ako je $\lambda \neq 0$ iz (12) slijedi da je

$$\begin{aligned}
\det(\lambda \mathbf{I} - \mathbf{B}_\omega) &= \det\left((\lambda + \omega - 1)\mathbf{I} + \sqrt{\lambda}\omega\left(\sqrt{\lambda}\mathbf{L} + \frac{1}{\sqrt{\lambda}}\mathbf{U}\right)\right) \\
&= (\sqrt{\lambda}\omega)^n \det\left(\mu \mathbf{I} + \left(\sqrt{\lambda}\mathbf{L} + \frac{1}{\sqrt{\lambda}}\mathbf{U}\right)\right) \\
&= (\sqrt{\lambda}\omega)^n \det(\mu \mathbf{I} + (\mathbf{L} + \mathbf{U})) \\
&= 0,
\end{aligned} \tag{13}$$

jer matrica $\mathbf{B}_J(\sqrt{\lambda}) = -\sqrt{\lambda}\mathbf{L} - (1/\sqrt{\lambda})\mathbf{U}$ ima iste svojstvene vrijednosti kao matrica $\mathbf{B}_J = -\mathbf{L} - \mathbf{U}$ i μ je svojstvena vrijednost matrice $\mathbf{B}_J$.

(c) Neka je $\lambda \neq 0$ svojstvena vrijednost matrice $\mathbf{B}_\omega$, i μ broj koji zadovoljava (11), odnosno $\lambda+\omega-1 = \pm\omega\sqrt{\lambda}\mu$. Zbog *(a)*, dovoljno je dokazati da je broj μ uz $\lambda + \omega - 1 = \omega\sqrt{\lambda}\mu$ svojstvena vrijednost matrice $\mathbf{B}_J$. Ovo, međutim, slijedi odmah iz (13).

□

Korolar 6.1 *Neka je* $\mathbf{A}$ *konzistentno uređena matrica. Tada za matricu koraka* $\mathbf{B}_{GS} = \mathbf{B}_1 = -(\mathbf{I}+\mathbf{L})^{-1}\mathbf{U}$ *Gauss-Seidelove metode imamo*

$$\rho(\mathbf{B}_{GS}) = (\rho(\mathbf{B}_J))^2 .$$

Ovo znači da je Gauss-Seidelova metoda dva puta brža od Jacobijeve.

Dokaz: Izbor $\omega = 1$ je ekvivalentan Gauss-Seidelovoj metodi, pa je $\lambda^2 = \lambda\mu^2$ ili $\lambda = \mu^2$. □

Sljedeći teorem nam daje ω_{opt} koji minimizira $\rho(\mathbf{B}_\omega)$.

Teorem 6.6 *Neka je* $\mathbf{A}$ *konzistentno uređena matrica, neka matrica* $\mathbf{B}_J$ *ima realne svojstvene vrijednosti, i neka je* $\mu = \rho(\mathbf{B}_J) < 1$. *Tada je*

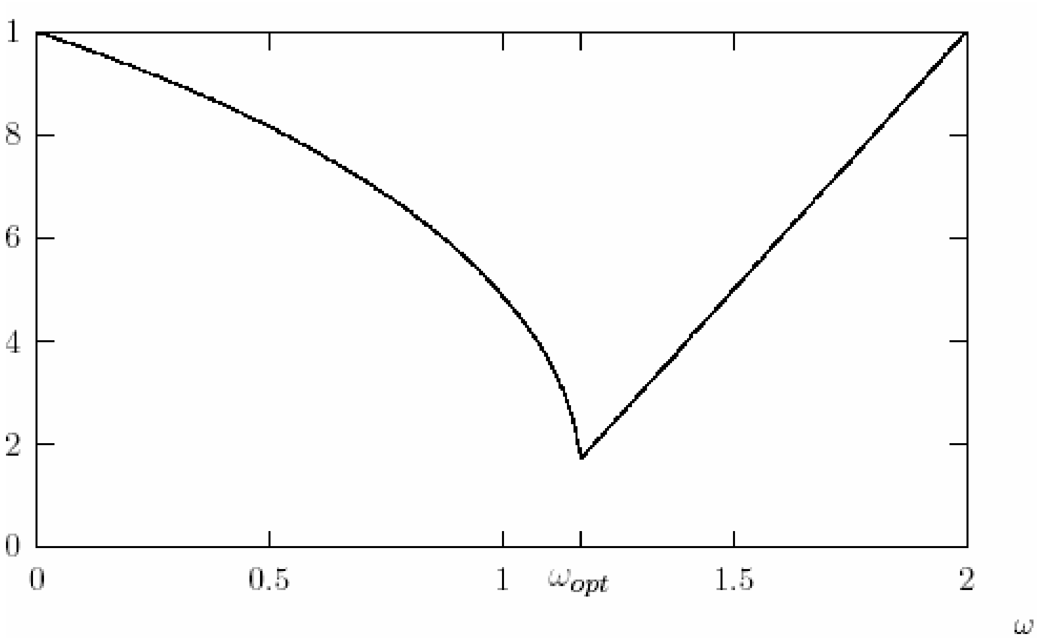
$$\begin{aligned}
\omega_{opt} &= \frac{2}{1+\sqrt{1-\mu^2}}, \\
\rho(\mathbf{B}_{\omega_{opt}}) &= \omega_{opt} - 1 = \frac{\mu^2}{\left(1+\sqrt{1-\mu^2}\right)^2}, \\
\rho(\mathbf{B}_\omega) &= \begin{cases} \omega - 1, & \omega_{opt} \leq \omega \leq 2 \\ 1-\omega+\frac{1}{2}\omega^2\mu^2+\omega\mu\sqrt{1-\omega+\frac{1}{4}\omega^2\mu^2}, & 0 < \omega \leq \omega_{opt} \end{cases},
\end{aligned}$$

(vidi sliku 1).

Dokaz: (vidi [1]) □

7 Ocjena greške

Računala su strojevi koji će ponavljati jednu te istu radnju proizvoljno puta, a današnja su računala vrlo brza i moćna. No, u praksi se može obaviti samo



Slika 1: Spektralni radijus matrice $\mathbf{B}_\omega$

konačan broj iteracija, pa je vrlo važno znati kada je aproksimacija dovoljno dobra.

Proces se zaustavlja poslije konačnog broja koraka k, pa je vektor $\mathbf{x}^{(k)}$ tada približno rješenje danog sustava. Pitanje koje se prirodno nameće je ocjena veličine tako nastale pogreške $\|\mathbf{x}^* - \mathbf{x}^{(k)}\|$. U tu svrhu promatrati ćemo sustav

$$\mathbf{x} = \mathbf{M}\mathbf{x} + \mathbf{d}, \tag{14}$$

odnosno iterativni postupak

$$\mathbf{x}^{(k+1)} = \mathbf{M}\mathbf{x}^{(k)} + \mathbf{d}, \quad k = 0, 1, 2, \ldots \tag{15}$$

Teorem 7.1 *Neka za matricu* $\mathbf{M} \in \mathbb{R}^{n\times n}$ *vrijedi*

$$\|\mathbf{M}\mathbf{y} - \mathbf{M}\mathbf{z}\| \le \alpha\|\mathbf{y} - \mathbf{z}\|, \quad \textit{za svaki } \mathbf{y}, \mathbf{z} \in \mathbb{R}^{n\times n},$$

za neki $\alpha \in [0, 1)$. *Ako je (15) za proizvoljni* $\mathbf{x}^{(0)} \in \mathbb{R}^n$, *onda je*

$$\|\mathbf{x}^* - \mathbf{x}^{(k)}\| \le \frac{\alpha}{1-\alpha}\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|, \quad k = 1, 2, \ldots, \tag{16}$$

gdje je $\mathbf{x}^*$ *točno rješenje sustava (14).*

Dokaz: Rješenje $\mathbf{x}^*$ sustava (14) postoji i jedinstveno je i $\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}^*$ (Teorem 3.1). Za svaki prirodan broj m i k vrijedi

$$\begin{aligned}
\|\mathbf{x}^{(k+m)} - \mathbf{x}^{(k+m-1)}\| &= \|\mathbf{M}\mathbf{x}^{(k+m-1)} - \mathbf{M}\mathbf{x}^{(k+m-2)}\| \\
&\leq \alpha\|\mathbf{x}^{(k+m-1)} - \mathbf{x}^{(k+m-2)}\| \\
&\vdots \\
&\leq \alpha^m\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|.
\end{aligned}$$

Nadalje imamo

$$\begin{aligned}
\|\mathbf{x}^{(k)} - \mathbf{x}^{(k+m)}\| &\leq \|\mathbf{x}^{(k)} - \mathbf{x}^{(k+1)}\| + \|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k+2)}\| + \cdots + \|\mathbf{x}^{(k+m-1)} - \mathbf{x}^{(k+m)}\| \\
&\leq \alpha\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| + \alpha^2\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| + \cdots + \alpha^m\|\mathbf{x}^{(k)} - \mathbf{x}^{(k+1)}\| \\
&\leq \alpha\frac{1-\alpha^m}{1-\alpha}\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|.
\end{aligned}$$

Kako je $\lim_{k\to\infty} \mathbf{x}^{(k+m)} = \mathbf{x}^*$ i $\alpha < 1$ slijedi

$$\|\mathbf{x}^{(k)} - \mathbf{x}^*\| \leq \frac{\alpha}{1-\alpha}\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|, \quad k = 1, 2, \ldots$$

$\square$

Teorem 7.2 *Neka je*

$$\|M\| < 1,$$

$$\mathbf{x}^{(k+1)} = \mathbf{M}\mathbf{x}^{(k)} + \mathbf{d}, \quad k = 0, 1, 2, \ldots$$

i

$$\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}^*.$$

Tada je

$$\|\mathbf{x}^* - \mathbf{x}^{(k)}\| \leq \frac{\|\mathbf{M}\|}{1-\|\mathbf{M}\|}\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|, \quad k = 1, 2, \ldots$$

Dokaz: Dokaz slijedi neposredno iz Teorema 7.1 i

$$\|\mathbf{M}\mathbf{x} - \mathbf{M}\mathbf{y}\| \leq \|\mathbf{M}\|\|\mathbf{x} - \mathbf{y}\|$$

sa $\alpha = \|\mathbf{M}\|$. $\square$

Dobivene ocjene pogreške $\|\mathbf{x}^* - \mathbf{x}^{(k)}\|$ mogu se koristiti za svaki iterativni postupak oblika (15) bez obzira na način na koji smo iz sustava $\mathbf{A}\mathbf{x} = \mathbf{b}$ dobili ekvivalentan sustav oblika (14).

Do sada smo pretpostavljali da se sva računanja izvode točno, odnosno da nema pogrešaka pri zaokruživanju. U praksi uvijek dolazi do pogreške zaokruživanja, jer računala imaju svoja ograničenja. Vrlo važna stvar kod iterativnih metoda je da greška zaokruživanja ne utječe mnogo na njih. Razlog tome je što iterativne metode u svakoj iteraciji nastoje poboljšati trenutnu aproksimaciju i nije važno koliko je točna ta aproksimacija. Problem se javlja kada iterativna metoda konvergira previše sporo, što ima za posljedicu da zbog pogreške zaokruživanja aproksimacija $\mathbf{x}^{(k)}$ bude jednaka sljedećoj aproksimaciji $\mathbf{x}^{(k)}$.

Primjer 7.1 *Neka je dan sustav*

$$\begin{aligned} 1.00000x_1 + 0.70710x_2 &= 0.29290 \\ 0.70710x_1 + 0.50000x_2 &= 0.20711. \end{aligned}$$

Matrica koraka $\mathbf{B}_{GS}$ *Gauss-Seidelovog iterativnog postupka ima spektralni radijus* $\rho(\mathbf{B}_{GS}) = 0.99998$. *Polazeći od* $\mathbf{x}^{(0)} = [0.00000, 0.00000]^T$ *i računajući s pet mjesta iza decimalne točke dobivamo*

$$\lim_{k\to\infty} = \mathbf{x}^{(0)},$$

što nije točno, jer rješenje danog sustava glasi $\mathbf{x}^* = [0.26267\ldots, 0.04275\ldots]^T$. *Očito, spora konvergencija Gauss-Seidelovog iterativnog procesa i pogreške zaokruživanja su uzrok tome.*

Sljedeći teorem govori o iterativnom postupku kod kojeg se u svakom koraku pojavljuje pogreška koja je posljedica pogreške zaokruživanja.

Teorem 7.3 *Neka se vektori niza* $\mathbf{x}^{(0)}$, $\mathbf{x}^{(1)}$, $\mathbf{x}^{(2)} \ldots$, *koji zadovoljavaju uvjet*

$$\mathbf{x}^{(k+1)} = \mathbf{M}\mathbf{x}^{(k)} + \mathbf{d}, \quad k = 0, 1, 2, \ldots$$

gdje je $\|\mathbf{M}\| < 1$, *izračunavaju s pogreškama zaokruživanja tako da se dobiva niz* $\bar{\mathbf{x}}^{(0)}$, $\bar{\mathbf{x}}^{(1)}$, $\bar{\mathbf{x}}^{(2)} \ldots$

Ako je

$$\begin{aligned} \bar{\mathbf{x}}^{(k+1)} &= \mathbf{M}\bar{\mathbf{x}}^{(k)} + \mathbf{d} + \mathbf{r}^k, \quad k = 0, 1, 2, \ldots \\ \|\mathbf{r}^{(k)}\| &\le \varepsilon, \quad k = 0, 1, 2, \ldots \end{aligned}$$

i

$$\mathbf{x}^{(0)} = \bar{\mathbf{x}}^{(0)},$$

onda je

$$\|\bar{\mathbf{x}}^{(k)} - \mathbf{x}^{(k)}\| \le \frac{\varepsilon}{1 - \|\mathbf{M}\|}, \quad k = 0, 1, 2, \ldots$$

i

$$\|\mathbf{x}^* - \bar{\mathbf{x}}^{(k)}\| \leq \frac{1}{1-\|\mathbf{M}\|}\left(\varepsilon\frac{1+\|\mathbf{M}\|}{1-\|\mathbf{M}\|} + \|\mathbf{M}\|\|\bar{\mathbf{x}}^{(k)} - \bar{\mathbf{x}}^{(k-1)}\|\right)$$

gdje je $\mathbf{x}^* = \lim\limits_{k\to\infty} \mathbf{x}^{(k)}$ *točno rješenje sustava (14).*

8 Jacobijeva, Gauss-Seidelova i SOR metoda u programskom paketu Mathematica

Jacobijeva metoda

```
Jac[a_, b_, n_, it_, gr_] :=
  Module[
    {i, k, xs, xn},
    xs = Table[0, {i, n}];
    xn = xs;
    k = 1;
    r = -b;
    PNormaVektora[r, n, 2];
    While[
      (No > gr) && (k <= it),
      Do[
        xn[[i]] = (b[[i]] - Sum[a[[i, j]] xs[[j]], {j, i - 1}] -
                Sum[a[[i, j]] xs[[j]], {j, i + 1, n}])/a[[i, i]],
        {i, n}
        ];
      r = a.xn - b;
      PNormaVektora[r, n, 2];
      xs = xn;
      pr = k;
      k = k + 1;
      ];
    Print["It_", pr, "=", N[xn]];
    ]
```

Gauss-Seidelova metoda

```
GS[a_, b_, n_, it_, gr_] :=
  Module[
    {i, k, xs, xn},
    xn = Table[0, {i, n}];
```

```
    k = 1;
    r = -b;
    PNormaVektora[r, n, 2];
    While[
      (No > gr) && (k <= it),
      Do[
        xn[[i]] = (b[[i]] - Sum[a[[i, j]] xn[[j]], {j, i - 1}] -
                Sum[a[[i, j]] xn[[j]], {j, i + 1, n}])/a[[i, i]],
        {i, n}
        ];
      r = a.xn - b;
      PNormaVektora[r, n, 2];
      pr = k;
      k = k + 1;
      ];
    Print["It_", pr, "=", N[xn]];
    ]
```

SOR metoda

```
SOR[a_, b_, n_, it_, gr_, om_] :=
  Module[
    {i, k, xs, xn},
    xs = Table[0, {i, n}];
    xn = xs;
    k = 1;
    r = -b;
    PNormaVektora[r, n, 2];
    While[
      (No > gr) && (k <= it),
      Do[
        xn[[i]] = (1 - om)xs[[i]] + (b[[i]] -
                  Sum[a[[i, j]] xn[[j]], {j, i - 1}] -
                  Sum[a[[i, j]] xs[[j]], {j, i + 1, n}])(om/a[[i, i]]),
        {i, n}
        ];
      r = a.xn - b;
      PNormaVektora[r, n, 2];
      xs = xn;
      pr = k;
      k = k + 1;
```

```
      ];
    Print["It_", pr, "=", N[xn]];
    ]
```

Ostali moduli

P-norma vektora

```
PNormaVektora[x_, n_, p_] :=
  Module[
    {i},
    Clear[No];
    If[
      (IntegerQ[p]) && (p > 0),
      No = N[(Sum[Abs[x[[i]]]^p, {i, n}])^(1/p)]
      ];
    ]
```

Generiranje podataka

```
GenPod[n_] :=
    Module[
      {a},
      a = Table[-1, {i, 1, n - 2}];
      a = Insert[a, 0, n/2];
      A =
        Table[Switch[i - j, n/2, -1, 1, a[[i - 1]], 0, 4, -1,
            a[[i]], -n/2, -1, _, 0], {i, n}, {j, n}];
      bo = Table[1, {i, n}];
      ];
```

$\mathbf{B}_J$ matrica

```
MatricaBJ[a_, n_] :=
  Module[
    {i, j},
    BJ = Table[0, {i, n}, {j, n}];
    Do[
      Do[
        BJ[[i, j]] = -a[[i, j]]/a[[i, i]], {j, i - 1}
        ];
      Do[
```

```
        BJ[[i, j]] = -a[[i, j]]/a[[i, i]], {j, i + 1, n}
        ],
      {i, n}
      ];
    ]
```

Spektralni radijus

```
SpecRad[a_, n_] :=
  Module[
    {i},
    Sv = Eigenvalues[a];
    Ro = N[Abs[Sv[[1]]]];
    Do[
      If[
        Ro < Abs[Sv[[i]]], Ro = N[Abs[Sv[[i]]]], Ro = Ro
        ],
      {i, n}
      ];
    ]
```

Testiranje

```
velicina = 50;
GenPod[velicina];
Timing[Jac[A, bo, velicina, 200, 0.0001]]
Timing[GS[A, bo, velicina, 200, 0.0001]]
MatricaBJ[A, velicina];
Timing[SpecRad[BJ, velicina];]
om = N[2/(1 + Sqrt[1 - (Ro^2)])]
Timing[SOR[A, bo, velicina, 200, 0.0001, om]]
```

Testiranje

```
velicina = 50;
GenPod[velicina];
Timing[SOR[A, bo, velicina, 200, 0.0001, 0.5]]
Timing[SOR[A, bo, velicina, 200, 0.0001, 0.8]]
Timing[SOR[A, bo, velicina, 200, 0.0001, 1.2]]
Timing[SOR[A, bo, velicina, 200, 0.0001, 1.5]]
Timing[SOR[A, bo, velicina, 200, 0.0001, 1.8]]
MatricaBJ[A, velicina];
```

```
Timing[SpecRad[BJ, velicina];]
om = N[2/(1 + Sqrt[1 - (Ro^2)])]
Timing[SOR[A, bo, velicina, 200, 0.0001, om]]
```

Literatura

[1] N.Bosner, *Iterativne metode za rješavanje linearnih sustava*, diplomski rad, Matematički odjel, PMF, Zagreb, 2001.

[2] M.Vrzić, *Veliki rijetki sustavi*, diplomski rad, Odjel za matematiku, Osijek, 2000.

[3] R.Scitovski, *Numerička matematika*, Elektrotehnički fakultet, Osijek, 1999.

[4] D.M.Young, *Iterative Solution of Large Linear Systems*, Academic Press, New York, 1972.

[5] R.Kress, *Numerical Analysis*, Springer-Verlag, Berlin, 1998.

[6] J.M.Ortega, *Numerical Analysis*, SIAM, Philadelphia, 1990.

[7] A. Berman,R.J.Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, SIAM, Philadelphia, 1994.

[8] R.Barrett,M.Berry,T.F.Chan,J.Demmel,J.Donato,J.Dongarra, V.Eijkhont,R.Pozo, C.Rpmine,H.Von der Vorst, *Templates for the Solution of Linear Systems: Bilding Blocks for Iterative Methods*, SIAM, Philadelphia, 1994.