Numerička analiza

7. predavanje

Autor: Nela Bosner

Predavač: Nela Bosner

nela@math.hr

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

PMF – Matematički odjel, Zagreb

Sadržaj predavanja

- Rješavanje linearnih sustava: iteracije iz Krylovljevih potprostora:
  - Aproksimacije iz Krylovljevih potprostora i prekondicioniranje.
- Metoda konjugiranih gradijenata (CG):
  - Metoda najbržeg silaska i njena konvergencija.
  - Metoda konjugiranih smjerova i njena konvergencija.
  - Metoda konjugiranih gradijenata (CG) i njena konvergencija.
  - Prekondicionirana metoda konjugiranih gradijenata
  - Numerička stabilnost metode konjugiranih gradijenata

Sadržaj predavanja (nastavak)

- GMRES metoda:
  - Arnoldijeva metoda
  - Izvod GMRES metode
  - Svojstva GMRES metode i njena konvergencija
- Primjeri konvergencije metoda:
  - Primjer konvergencije CG metode.
  - Primjer konvergencije GMRES metode.

Rješavanje linearnih sustava: iteracije iz Krylovljevih potprostora

Krylovljevi potprostori

Prisjetimo se najprije rezultata iz linearne algebre, koji tvrdi da:

- svaka matrica poništava svoj karakteristični i minimalni polinom.

Za $A \in \mathbb{C}^{n\times n}$ i $b \in \mathbb{C}^n$, to možemo zapisati na sljedeći način

$$p_A(A) = a_0 I + a_1 A + \cdots + a_{n-1} A^{n-1} + a_n A^n = 0,$$

gdje je $p_A(\lambda) = \det(A - \lambda I) = \sum_{i=0}^n a_i \lambda^i$ karakteristični polinom matrice A.

- Slično možemo napisati da je $\mu_A(A) = 0$, gdje je $\mu_A(\lambda)$ minimalni polinom stupnja manjeg ili jednakog n.

Krylovljevi potprostori (nastavak)

- Linearni sustavi najčešće se rješavaju kada je matrica sutava regularna, jer je tada rješenje jedinstveno.
- U tom slučaju nula ne može biti korijen karakterističnog polinoma, pa je $a_0 \neq 0$.
- Odavde jednostavnim računom možemo dobiti da vrijedi

$$
\begin{aligned}
& -\frac{1}{a_0}(a_1 I + \cdots + a_{n-1} A^{n-2} + a_n A^{n-1}) \cdot A \\
& = A \cdot \left(-\frac{1}{a_0} \right) (a_1 I + \cdots + a_{n-1} A^{n-2} + a_n A^{n-1}) = I,
\end{aligned}
$$

to jest,

$$
A^{-1} = -\frac{1}{a_0}(a_1 I + \cdots + a_{n-1} A^{n-2} + a_n A^{n-1}).
$$

Krylovljevi potprostori (nastavak)

- Budući da rješenje sustava $Ax = b$ možemo zapisati kao $x = A^{-1}b$, uz uvažavanje prethodnog zapisa za A^{-1} možemo zaključiti da je
$$x = -\frac{a_1}{a_0} b - \cdots - \frac{a_{n-1}}{a_0} A^{n-2} b - \frac{a_n}{a_0} A^{n-1} b,$$
odnosno
$$x \in \operatorname{span}\{b, Ab, \ldots, A^{n-1}b\} = \mathcal{K}_n(A, b).$$
- Prostor koji se pojavljuje na desnoj strani zovemo Krylovljevim prostorom matrice A i inicijalnog vektora b.
- Upravo odavde dobivamo ideju za metode rješavanja sustava linearnih jednadžbi koje bi se temeljile na aproksimacijama iz Krylovljevih potprostora.

Aproksimacije iz Krylovljevih potprostora

- Naime kako je vektor b jedini vektor direktno vezan za problem rješavanja sutava $Ax = b$, čini nam se prirodno uzeti neki multipl od b kao početnu aproksimaciju rješenja, tj.

$$x_0 \in \operatorname{span}\{b\}.$$

- Nakon toga računamo produkt Ab i zahtijevamo da nam je sljedeća aproksimacija jednaka nekoj linearnoj kombinaciji od b i Ab, tj.

$$x_1 \in \operatorname{span}\{b, Ab\}.$$

- Taj se proces nastavlja, tako da nam aproksimacija u k-tom koraku zadovoljava

$$x_k \in \operatorname{span}\{b, Ab, \ldots, A^k b\}, \quad k = 1, 2, \ldots.$$

Aproksimacije iz Krylovljevih potprostora (nast.)

- Metoda, naravno, mora odrediti kriterij po kojem biramo vektore iz pojedinih Krylovljevih potprostora u svakom koraku,
- Prije svega promotrimo jednu vrlo jednostavnu činjenicu. Ako je x_k aproksimacija rješenja u k-tom koraku neke iterativne metode za rješavanje linearnih sustava, i ako u $(k+1)$-om koraku uzmemo

$$x_{k+1} = x_k + A^{-1}(b - Ax_k),$$

  tada je $x_{k+1} = A^{-1}b$ rješenje sustava.
- Budući da je računanje vektora $A^{-1}(b - Ax_k)$ ekvivalentno problemu rješavanja polaznog sustava, korekciju vektora x_k radit ćemo pomoću neke njegove aproksimacije.

Aproksimacije iz Krylovljevih potprostora (nast.)

- Ta aproksimacija mora
  - se lakše računati i
  - mora nas držati unutar Krylovljevih potprostora.
- Zato, najprije pretpostavimo da polazimo od iteracije jednostavnog oblika

$$x_{k+1} = x_k + \alpha_k (b - A x_k),$$

gdje je α_k parametar koji određuje točku na pravcu koji
  - prolazi kroz točku x_k i
  - pruža se u smjeru vektora $r_k = b - A x_k$.
- Tada je

$$\begin{aligned} x_1 &= x_0 + \alpha_0 r_0 \in x_0 + \operatorname{span}\{r_0\}, \\ x_2 &= x_0 + \alpha_0 r_0 + \alpha_1 (r_0 - \alpha_0 A r_0) \in x_0 + \operatorname{span}\{r_0, A r_0\}, \\ &\vdots \end{aligned}$$

Aproksimacije iz Krylovljevih potprostora (nast.)

- Na temelju iteracije lako se vidi da je za $k = 0, 1, 2, \ldots$

$$x_k \in x_0 + \text{span}\,\{r_0, Ar_0, A^2 r_0, \ldots, A^{k-1} r_0\} = x_0 + \mathcal{K}_k(A, r_0),$$
$$r_k \in r_0 + \text{span}\,\{Ar_0, A^2 r_0, A^3 r_0, \ldots, A^k r_0\} = r_0 + A\mathcal{K}_k(A, r_0),$$

  pri čemu je $r_k = b - Ax_k$ rezidual k-te iteracije, a sa $e_k = A^{-1} r_k = A^{-1} b - x_k$ označavamo grešku.
- Prostor koji se pojavljuje kod definiranja iteracije je također Krylovljev potprostor, ali određen matricom A i vektorom r_0, koji ne mora biti jednak vektoru b.
- Primijetimo da je za $x_0 = 0$, $r_0 = b$, pa je $x_k \in \mathcal{K}_k(A, b)$.

Aproksimacije iz Krylovljevih potprostora (nast.)

- Međutim, ako prema uvodnom razmatranju napišemo da je $A^{-1} = q_{k-1}(A)$, gdje je $q_{k-1}(\lambda)$ polinom stupnja $(k-1)$ za neki $k \leq n$, tada
  - za bilo koji r_0,
  - u prostoru $x_0 + \mathcal{K}_k(A, r_0)$ možemo naći vektor oblika $x_0 + q_{k-1}(A)r_0 = x_0 + A^{-1}r_0 = A^{-1}b$ koji je rješenje linearnog sustava $Ax = b$.
- Općenito, iteracija iterativne metode bit će oblika

$$x_k = x_{k-1} + z_{k-1}, \quad \text{ili} \quad x_k = x_0 + w_k$$

pri čemu će z_{k-1} i w_{k-1} biti birani tako da $x_k \in x_0 + \mathcal{K}_k(A, r_0)$ i da zadovoljava neki uvjet optimalnosti, najčešće vezan uz normu reziduala ili neku normu greške.

Prekondicioniranje

- Kao što ćemo kasnije vidjeti, sve metode ovakvog oblika konvergiraju vrlo brzo ukoliko je matrica sustava A blizu identiteti. (Jasno za iteracije oblika $x_{k+1} = x_k + \alpha_k(b - Ax_k)$.)
- Naravno, to se događa vrlo rijetko, ali mi možemo cijeli sustav $Ax = b$ zamijeniti sa modificiranim sustavom

$$M^{-1}Ax = M^{-1}b \quad \text{ili} \quad AM^{-1}y = b, \; x = M^{-1}y,$$

  čija bi matrica bila bliža identiteti od matrice A.
- Ovdje se radi o lijevom odnosno desnom prekondicioniranju.

Prekondicioniranje (nastavak)

- Ako je M hermitska i pozitivno definitna matrica, tada možemo izvesti simetrično prekondicioniranje i riješiti modificirani linearni sustav

$$L^{-1}AL^{-*}y = L^{-1}b, \quad x = L^{-*}y,$$

  pri čemu je $M = LL^*$.

- Matrica L može biti
  - hermitski drugi korijen od M,
  - ili donje trokutasti faktor Choleskog od M,
  - ili bilo koja druga matrica koja zadovoljava $M = LL^*$.

  U svakom slučaju bitno je samo to da je računanje $M^{-1}z$ za bilo koji vektor z jednostavnije (odnosno da je jednostavnije riješiti sustav sa matricom sustava M).

Prekondicioniranje (nastavak)

- Dakle, ako matricu prekondicioniranja M možemo izabrati tako
  - da se linearni sustav sa matricom sustava M može jednostavno riješiti
  - i da $M^{-1}A$ ili AM^{-1} ili $L^{-1}AL^{-*}$ aproksimiraju identitetu,

  tada primjenom iterativne metode na prekondicionirane sustave možemo dobiti još bolju tehniku za računanje rješenja sustava.
- Pravi smisao tvrdnje da "prekondicionirana matrica aproksimira identitetu", ovisi o iterativnoj metodi koju smo koristili, i o njenim svojstvima.

Metoda konjugiranih gradijenata (CG)

Metoda najbržeg silaska

- To je iterativna metode iz Krylovljevih potprostora, za rješavanje sustava $Ax = b$, $A \in \mathbb{R}^{n \times n}$, $b, x \in \mathbb{R}^n$, pri čemu je matrica sustava A simetrična pozitivno definitna:
  - $A^T = A$
  - $y^T A y > 0$ za svaki $y \in \mathbb{R}^n$, $y \neq 0$
- Ideja odabira parametra α_k u k-toj iteraciji metode $x_{k+1} = x_k + \alpha_k r_k$ je ta da se minimizira neka norma greške $e_{k+1} = e_k - \alpha_k r_k$.
- Problem je što nam je greška jednako tako nepoznata kao i samo rješenje, pa norma $\| \cdot \|_2$ ne dolazi u obzir.
- Ono što možemo izračunati je rezidual $r_{k+1} = Ae_{k+1} = r_k - \alpha_k A r_k$.

Metoda najbržeg silaska (nastavak)

- Za simetričnu pozitivno definitnu matricu A ima smisla definirati A-normu $\|\cdot\|_A$

$$\|v\|_A = \sqrt{\langle v, v \rangle_A} = \sqrt{v^T A v}.$$

- Uvjerite se da je $\|\cdot\|_A$ zaista norma na $\mathbb{R}^n$!
- Definirajmo sada funkciju $f : \mathbb{R} \longrightarrow \mathbb{R}$ kao

$$\begin{aligned} f(\alpha_k) &= e_{k+1}^T A e_{k+1} \\ &= \alpha_k^2 r_k^T A r_k - 2\alpha_k r_k^T A e_k + e_k^T A e_k \\ &= \alpha_k^2 r_k^T A r_k - 2\alpha_k r_k^T r_k + e_k^T A e_k. \end{aligned}$$

- Traženje minimuma funkcije $f(\alpha_k)$ je ekvivalentno traženju minimuma $\|e_{k+1}\|_A$.

Metoda najbržeg silaska (nastavak)

- Funkcija $f(\alpha_k)$ je kvadratna funkcija po varijabli α_k, i parametar uz α_k^2 je $r_k^T A r_k \geq 0$, što znači da funkcija poprima minimum u tjemenu, koje je jedina nultočka derivacije funkcije f'.

$$0 = f'(\alpha_k) = 2\alpha_k r_k^T A r_k - 2 r_k^T r_k,$$

  odakle slijedi da se minimalna A-norma greške e_{k+1} postiže za

$$\alpha_k = \frac{r_k^T r_k}{r_k^T A r_k}.$$

- Zbog ovakvog odabira parametra α_k vrijedi da je r_{k+1} okomit na r_k, tj.

$$r_k^T r_{k+1} = r_k^T r_k - \alpha_k r_k^T A r_k = r_k^T r_k - \frac{r_k^T r_k}{r_k^T A r_k} r_k^T A r_k = 0.$$

Metoda najbržeg silaska (nastavak)

- Važno je još primijetiti, da tako dugo dok nismo našli egzaktno rješenje ($r_k \neq 0$), α_k je strogo veći od nule.
- Zbog okomitosti r_{k+1} i r_k slijedi

$$\begin{aligned} \|e_k\|_A^2 &= e_{k+1}^T A e_{k+1} + 2\alpha_k r_k^T A e_{k+1} + \alpha_k^2 r_k^T A r_k \\ &= \|e_{k+1}\|_A^2 + \alpha_k^2 \|r_k\|_A^2 > \|e_{k+1}\|_A^2, \end{aligned}$$

  odakle se vidi da se A-norma greške smanjuje u svakom koraku.
- Ova metoda, zbog načina odabira parametra, zove se metoda najbržeg silaska.

Algoritam metode najbržeg silaska

Algoritam (Metoda najbržeg silaska).

```
x_0 zadan;  k = 0;
r_0 = b − A x_0;
dok nije zadovoljen kriterij zaustavljanja radi {
    α_k = (r_k^T r_k) / (r_k^T A r_k);
    x_{k+1} = x_k + α_k r_k;
    r_{k+1} = r_k − α_k A r_k;
    k = k + 1;
}
```

x_0 zadan; $k = 0$;

$r_0 = b - Ax_0$;

dok nije zadovoljen kriterij zaustavljanja radi $\{$

$\quad \alpha_k = \frac{r_k^T r_k}{r_k^T A r_k}$;

$\quad x_{k+1} = x_k + \alpha_k r_k$;

$\quad r_{k+1} = r_k - \alpha_k A r_k$;

$\quad k = k + 1$;

$\}$

Konvergencija metode najbržeg silaska

- Iz iteracije možemo dobiti rekurziju za e_k (već smo ju vidjeli), koja glasi

$$e_{k+1} = e_k - \alpha_k r_k = e_k - \alpha_k A e_k = (I - \alpha_k A) e_k,$$

koju na drugačiji način možemo zapisati kao

$$e_{k+1} = p_1(A) e_k,$$

pri čemu je p_1 polinom prvog stupnja, takav da je $p_1(\lambda) = 1 - \alpha_k \lambda$.

- Prema gore navedenome, za grešku e_{k+1} onda vrijedi

$$\|e_{k+1}\|_A = \min_{p_1 \in \mathbb{P}_1, p_1(0)=1} \|p(A) e_k\|_A,$$

gdje se minimum uzima po svim polinomima p prvog stupnja, za koje je $p(0) = 1$, odnosno koji su oblika $p(\lambda) = 1 - \alpha\lambda$. — to nam je i bio uvjet za izbor α_k

Konvergencija metode najbržeg silaska (nast.)

- Za nastavak analize trebat će nam još sljedeće svojstvo A-norme za proizvoljni vektor u:

$$\begin{aligned} \|u\|_A &= (u^T A u)^{1/2} = (u^T A^{1/2} A^{1/2} u)^{1/2} \\ &= ((A^{1/2} u)^T A^{1/2} u)^{1/2} = \|A^{1/2} u\|_2. \end{aligned}$$

- Vidjet ćemo da simetričnu matricu A možemo dijagonalizirati ortogonalnom matricom: $A = U \Lambda U^T$, $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, gdje je $U^T U = U U^T = I$, i $\lambda_i \in \mathbb{R}$ za $i = 1, \ldots, n$ su realne svojstvene vrijednosti.
- Za simetričnu pozitivno definitnu matricu još vrijedi da su $\lambda_i > 0$ za $i = 1, \ldots, n$, tj. sve svojstvene vrijednosti su pozitivne.

Konvergencija metode najbržeg silaska (nast.)

- Tada možemo definirati $A^{1/2}$ simetrični kvadratni korijen matrice A, oblika $A^{1/2} = U\Lambda^{1/2}U^T$, koji komutira sa A i bilo kojim polinomom od A.
- Zato vrijedi

$$
\begin{aligned}
\|e_{k+1}\|_A &= \min_{p_1 \in \mathbb{P}_1, p_1(0)=1} \|A^{1/2} p_1(A) e_k\|_2 \\
&= \min_{p_1 \in \mathbb{P}_1, p_1(0)=1} \|U p_1(\Lambda) U^T A^{1/2} e_k\|_2 \\
&\leq \min_{p_1 \in \mathbb{P}_1, p_1(0)=1} \|p_1(\Lambda)\|_2 \|e_k\|_A \\
&= \min_{p_1 \in \mathbb{P}_1, p_1(0)=1} \max_{j=1,\ldots,n} |p_1(\lambda_j)| \|e_k\|_A.
\end{aligned}
$$

Konvergencija metode najbržeg silaska (nast.)

- Može se pokazati da se taj minimum postiže za polinom

$$p_1(\lambda) = 1 - \frac{2}{\lambda_{min} + \lambda_{max}} \lambda$$

i iznosi

$$|p_1(\lambda_{min})| = |p_1(\lambda_{max})| = \frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} + \lambda_{min}} = \frac{\kappa_2(A) - 1}{\kappa_2(A) + 1} < 1,$$

pri čemu je λ_{min} minimalna, a λ_{max} maksimalna svojstvena vrijednost matrice A, te $\kappa_2(A) = \lambda_{max}/\lambda_{min}$ uvjetovanost matrice A.

Konvergencija metode najbržeg silaska (nast.)

- Dakle dobivamo rezultat

$$\|e_{k+1}\|_A \leq \left(\frac{\kappa(A) - 1}{\kappa(A) + 1} \right) \|e_k\|_A,$$

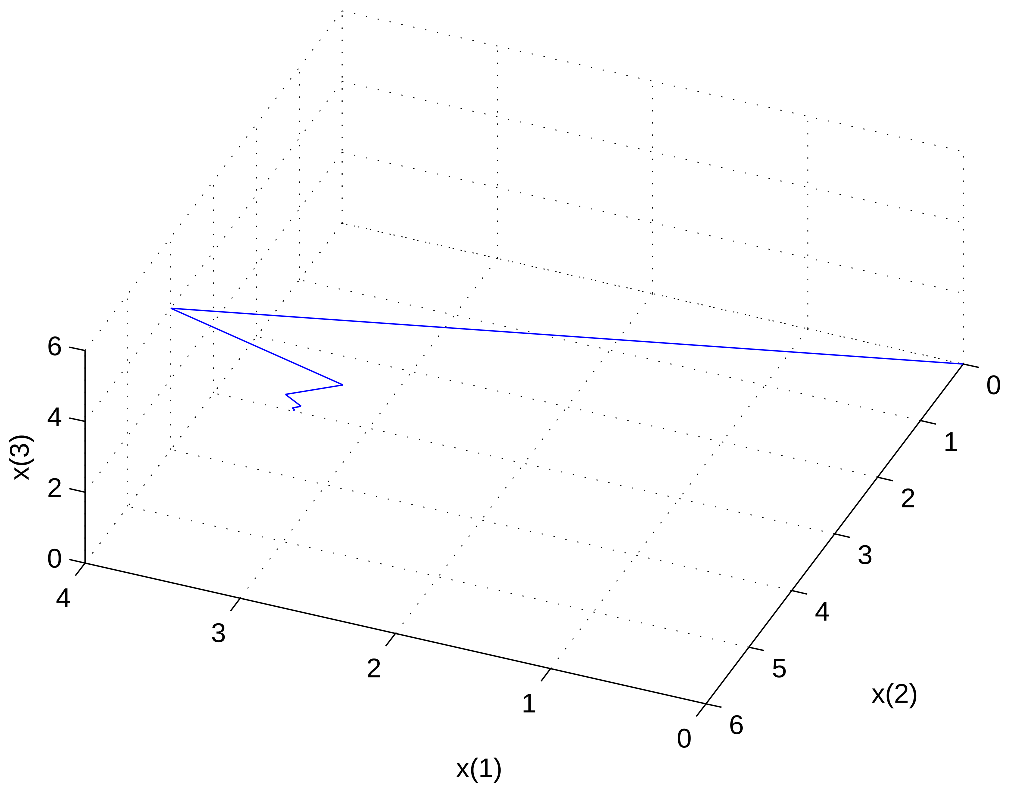
i kao konačni oblik

$$\|e_k\|_A \leq \left(\frac{\kappa(A) - 1}{\kappa(A) + 1} \right)^k \|e_0\|_A.$$

- Iz njega možemo zaključiti da za loše uvjetovane matrice možemo imati sporu konvergenciju jer je u tom slučaju $\frac{\kappa(A)-1}{\kappa(A)+1} \approx 1$.
- Ova metoda često dosta sporo konvergira, jer radi korake u smjeru kojim je neki raniji korak već prošao.

Konvergencija metode najbržeg silaska (nast.)

Iteracije metode najbrzeg silaska za $A \in \mathbb{R}^{3\times 3}$.



Metoda konjugiranih smjerova

- Da bi se izbjegla spora konvergencija metode najbržeg silaska, unaprijed odabiremo skup A-ortogonalnih vektora, odnosno smjerove traganja d_0, $d_1, \ldots, d_{n-1}$.
- Dva vektora d_i i d_j su A-ortogonalna ili konjugirana ako vrijedi da je

$$\langle d_i, d_j \rangle_A = d_j^T A d_i = 0.$$

- Lagano se može provjeriti da su A-ortogonalni vektori linearno nezavisni.
- Znači u svakom koraku biramo točku

$$x_{k+1} = x_k + \alpha_k d_k$$

s minimalnom A-normom greške.

Metoda konjugiranih smjerova (nastavak)

- Dakle, u svakom smjeru d_k napravit ćemo točno jedan korak, takav da ćemo poništiti komponentu vektora greške e_k u smjeru Ad_k.
- Nakon n koraka bit ćemo gotovi!
- U $(k+1)$-om koraku e_{k+1} je jednak početnoj grešci kojoj su odstranjene sve komponente u smjerovima $Ad_0, \ldots, Ad_k$, odnosno
- e_{k+1} je A-ortogonalan na $d_0, \ldots, d_k$.
- A-ortogonalnost između e_{k+1} i d_k je ekvivalentna nalaženju točke minimuma duž smjera traganja d_k, kao i u metodi najbržeg silaska.

Metoda konjugiranih smjerova (nastavak)

- Ponovo ćemo derivirati po α_k funkciju
$$g(\alpha_k) = e_{k+1}^T A e_{k+1}$$
i izjednačiti je s nulom, samo što je u tom slučaju r_{k+1} okomit na d_k.
- Ako opet uvrstimo da je $r_{k+1} = r_k - \alpha_k A d_k$ i $e_{k+1} = e_k - \alpha_k d_k$
$$g(\alpha_k) = \alpha_k^2 d_k^T A d_k - 2\alpha_k d_k^T A e_k + e_k^T A e_k,$$
slijedi da je
$$g'(\alpha_k) = 2\alpha_k d_k^T A d_k - 2 d_k^T r_k$$

Metoda konjugiranih smjerova (nastavak)

odakle ćemo izjednačavanjem s nulom dobit izraz za α_k

$$\alpha_k = \frac{d_k^T r_k}{d_k^T A d_k}.$$

- Ovako dobivena metoda naziva se metoda konjugiranih smjerova.
- Da je e_{k+1} A-ortogonalan na $d_0, \ldots, d_k$, može se sad dokazati matematičkom indukcijom:
  - $d_0^T A e_1 = d_0^T A e_0 - \alpha_0 d_0^T A d_0 = d_0^T r_0 - \frac{d_0^T r_0}{d_0^T A d_0} d_0^T A d_0 = 0$ iz definicije parametra α_0
  - Ako pretpostavimo da je $d_i^T A e_k = 0$ za $i = 0, \ldots, k-1$, tada imamo

Metoda konjugiranih smjerova (nastavak)

$$d_k^T A e_{k+1} = d_k^T A e_k - \alpha_k d_k^T A d_k = d_k^T r_k - \frac{d_k^T r_k}{d_k^T A d_k} d_k^T A d_k = 0$$

$$d_i^T A e_{k+1} = d_i^T A e_k - \alpha_k d_i^T A d_k = 0, \quad i = 0, \ldots, k-1$$

iz definicije parametra α_k, pretpostavke indukcije i A-ortogonalnosti vektora d_i.

Algoritam metode konjugiranih smjerova

Algoritam (Metoda konjugiranih smjerova).

```
x_0 zadan;
A-ortogonalni vektori d_0, d_1, ..., d_{n-1} zadani;
k = 0;
r_0 = b - A x_0;
dok nije zadovoljen kriterij zaustavljanja radi {
    α_k = (d_k^T r_k) / (d_k^T A d_k);
    x_{k+1} = x_k + α_k d_k;
    r_{k+1} = r_k - α_k A d_k;
    k = k + 1;
}
```

x_0 zadan;

A-ortogonalni vektori $d_0, d_1, \ldots, d_{n-1}$ zadani;

$k = 0$;

$r_0 = b - Ax_0$;

dok nije zadovoljen kriterij zaustavljanja radi $\{$

$\quad \alpha_k = \dfrac{d_k^T r_k}{d_k^T A d_k}$;

$\quad x_{k+1} = x_k + \alpha_k d_k$;

$\quad r_{k+1} = r_k - \alpha_k A d_k$;

$\quad k = k + 1$;

$\}$

Svojstva metode konjugiranih smjerova

Teorem. Za metodu konjugiranih smjerova vrijede sljedeća svojstva:

$$d_j^T A d_i = 0 \quad (i \neq j)$$

$$d_j^T r_i = d_j^T A e_i = 0 \quad (j < i)$$

$$d_i^T r_0 = d_i^T r_1 = \cdots = d_i^T r_i.$$

Skalar α_i može se zato napisati kao

$$\alpha_k = \frac{d_k^T r_0}{d_k^T A d_k}.$$

Dokaz. Prve dvije tvrdnje smo već pokazali, a treća slijedi iz svojstva A-ortogonalnosti vektora d_i:

Svojstva metode konjugiranih smjerova (nast.)

$$\begin{aligned} d_i^T r_i &= d_i^T r_{i-1} - \alpha_{i-1} d_i^T A d_{i-1} \\ &= d_i^T r_{i-1} = d_i^T r_{i-2} - \alpha_{i-2} d_i^T A d_{i-2} = d_i^T r_{i-2} \\ &= \cdots = d_i^T r_0 \end{aligned}$$

Poslijedica formule za α_k iz ovog Teorema je ta da se aproksimacije x_1,x_2,... od x mogu izračunati bez računanja reziduala r_1,r_2,...,. Ovisne su samo o početnom rezidualu r_0.

Konvergencija metode konjugiranih smjerova

Teorem. Metoda konjugiranih smjerova je m-koračna metoda ($m \leq n$), u smislu da je u m-tom koraku aproksimacija x_m jednaka rješenju $x = A^{-1}b$.

Dokaz. Neka je m najmanji cijeli broj takav da se $e_0 = x - x_0$ nalazi u $\text{span}\{d_0, \ldots, d_{m-1}\}$. Očito je $m \leq n$, budući da su vektori d_0,d_1,... linearno nezavisni, pa ih maksimalno može biti n. Zatim, izaberimo skalare a_0,...,a_{m-1} takve da je

$$e_0 = a_0 d_0 + \cdots + a_{m-1} d_{m-1}.$$

Odavde slijedi

$$x = x_0 + a_0 d_0 + \cdots + a_{m-1} d_{m-1}.$$

Konvergencija metode konjugiranih smjerova

Nadalje,

$$r_0 = b - Ax_0 = A(x - x_0) = a_0 A d_0 + \cdots + a_{m-1} A d_{m-1}.$$

Koristeći se činjenicom da su smjerovi traženja d_i medjusobno konjugirani i jednakošću za α_k iz prethodnog Teorema, računanjem skalarnog produkta $d_i^T r_0$ dobivamo

$$d_i^T r_0 = a_i d_i^T A d_i$$

odnosno

$$a_i = \frac{d_i^T r_0}{d_i^T A d_i} = \alpha_i.$$

Konvergencija metode konjugiranih smjerova

Budući da je, primjenom indukcije na iteraciju za x_k

$$x_m = x_0 + \alpha_0 d_0 + \cdots + \alpha_{m-1} d_{m-1},$$

možemo zaključiti da tvrdnja $x = x_m$ vrijedi. $\blacksquare$

Metoda konjugiranih gradijenata [CG]

- Preostaje nam još pronaći odgovarajuće vektore $d_0, d_1, \ldots, d_{n-1}$.
- Skup A-ortogonalnih smjerova $\{d_i\}$ možemo dobiti primjenom Gramm–Schmidtove metode A-ortogonalizacije na niz linearno nezavisnih vektora $u_0,\ldots,u_{n-1}$ sa skalarnim produktom $\langle \cdot, \cdot \rangle_A$.
- Dakle, A-ortogonalne vektore možemo dobiti kao

$$d_k = u_k + \sum_{i=0}^{k-1} \beta_{ki} d_i,$$

pri čemu su koeficijenti oblika

$$\beta_{ki} = -\frac{d_i^T A u_k}{d_i^T A d_i}.$$

Metoda konjugiranih gradijenata [CG] (nast.)

- Konkretan odabir vektora $u_0, \ldots, u_{n-1}$ vodi nas do metode konjugiranih gradijenata (CG).
- Metoda konjugiranih gradijenata (CG) je, zapravo, metoda konjugiranih smjerova za $u_i = r_i$.
- Činjenica da su vektori r_i dobiveni metodom konjugiranih smjerova linearno nezavisni, može se provjeriti uz pomoć prethodnih teorema.
- Ponovo vrijedi

$$\operatorname{span}\{d_0, d_1, \ldots, d_{k-1}\} = \operatorname{span}\{r_0, r_1, \ldots, r_{k-1}\},$$

  i budući da je r_k ortogonalan na prethodne smjerove traganja vrijedi

$$r_i^T r_j = 0, \qquad i \neq j.$$

Metoda konjugiranih gradijenata [CG] (nast.)

- Promatramo sljedeći skalarni produkt
$$r_{i+1}^T r_k = r_i^T r_k - \alpha_i d_i^T A r_k,$$
- odakle vrijedi
$$d_i^T A r_k = \frac{1}{\alpha_i}(r_i^T r_k - r_{i+1}^T r_k).$$
Za $i < k - 1$ lijeva strana ove jednakosti je jednaka 0, pa su $\beta_{ki} = 0$ za $i = 0, 1, \ldots, k - 2$, a za $\beta_k = \beta_{k,k-1}$ zbog izraza za α_{k-1} i prethodnih teorema vrijedi
$$\beta_k = -\frac{d_{k-1}^T A r_k}{d_{k-1}^T A d_{k-1}} = \frac{r_k^T r_k}{\alpha_{k-1} d_{k-1}^T A d_{k-1}}$$
$$= \frac{r_k^T r_k}{d_{k-1}^T r_{k-1}} = \frac{r_k^T r_k}{r_{k-1}^T r_{k-1}}$$

Metoda konjugiranih gradijenata [CG] (nast.)

- Također, zbog istih teorema možemo i α_k napisati u ljepšem obliku

$$\alpha_k = \frac{r_k^T r_k}{d_k^T A d_k},$$

  odakle se vidi, da ukoliko nismo našli egzaktno rješenje u k-tom koraku, α_k je pozitivan.
- Ovime smo u potpunosti definirali metodu konjugiranih gradijenata.

Algoritam metode konjugiranih gradijenata

Algoritam (Metoda konjugiranih gradijenata).

x_0 `zadan;`
$k = 0$;
$d_0 = r_0 = b - Ax_0$;
`dok nije zadovoljen kriterij zaustavljanja radi {`

$\alpha_k = \frac{r_k^T r_k}{d_k^T A d_k}$;
$x_{k+1} = x_k + \alpha_k d_k$;
$r_{k+1} = r_k - \alpha_k A d_k$;
$\beta_{k+1} = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}$;
$d_{k+1} = r_{k+1} + \beta_{k+1} d_k$;
$k = k + 1$;

`}`

Svojstva metode konjugiranih gradijenata

- Za primjenu metode konjugiranih gradijenata ne trebamo pristupati pojedinim elementima matrice.
- Dovoljno je znati djelovanje matrice na vektor $A \cdot y$ — često se zadaje kao funkcija od y.

Konvergencija metode konjugiranih gradijenata

Teorem. Greška e_k dobivena u k-tom koraku metode konjugiranih gradijenata ima najmanju A-normu na prostoru

$$e_0 + \mathrm{span}\{Ae_0, A^2e_0, \ldots, A^ke_0\}.$$

Dokaz. Ako gledamo kako su definirani vektori e_k i d_k u ovoj metodi, tada možemo zaključiti sljedeće: $d_0 = r_0 = Ae_0$, pa je

$$e_1 = e_0 - \alpha_0 r_0 = e_0 - \alpha_0 Ae_0 \in e_0 + \mathrm{span}\{Ae_0\}.$$

Pretpostavimo da vrijedi da je $e_k \in \mathrm{span}\{e_0, Ae_0, \ldots, A^ke_0\}$ i $d_{k-1} \in \mathrm{span}\{Ae_0, A^2e_0, \ldots, A^ke_0\}$, tada za e_{k+1} imamo

$$e_{k+1} = e_k - \alpha_k d_k = e_k - \alpha_k\beta_k d_{k-1} - \alpha_k Ae_k \in \mathrm{span}\{e_0, \ldots, A^{k+1}e_0\}.$$

Konvergencija metode konjugiranih gradijenata

Matematičkom indukcijom možemo zaključiti da je koeficijent uz e_0 uvijek jednak 1, odnosno

$$e_{k+1} \in e_0 + \operatorname{span}\{Ae_0, A^2 e_0, \ldots, A^{k+1} e_0\}$$

i zbog $d_k = Ae_k + \beta_k d_{k-1}$

$$d_k \in \operatorname{span}\{Ae_0, A^2 e_0, \ldots, A^{k+1} e_0\}.$$

Odavde slijedi da je

- $\operatorname{span}\{d_0, d_1, \ldots, d_k\} = \operatorname{span}\{Ae_0, \ldots, A^{k+1} e_0\}$,
- a kako je e_{k+1} A-ortogonalan na $\operatorname{span}\{d_0, d_1, \ldots, d_k\}$,

onda slijedi da je e_{k+1} vektor u prostoru $e_0 + \operatorname{span}\{Ae_0, A^2 e_0, \ldots, A^{k+1} e_0\}$ sa najmanjom A-normom.

Konvergencija metode konjugiranih gradijenata

Za $k = n - 1$ zbog linearne nezavisnosti vektora $\{d_0, \ldots, d_{n-1}\}$ slijedi da je $e_n = 0$. ∎

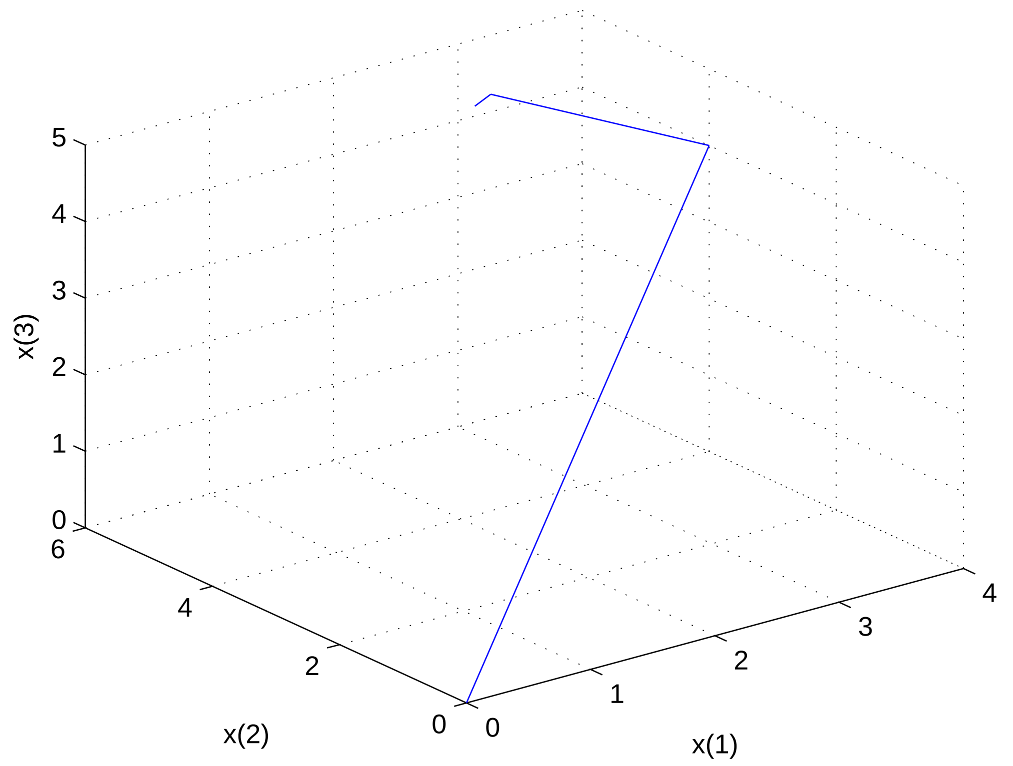
Dakle, metoda konjugiranih gradijenata konvergira u najviše n koraka.

Teorem. U svakom koraku metode konjugiranih gradijenata, duljina (2-norma) vektora greške $e_k = x - x_k$ se reducira, pri čemu je $A^{-1}b = x = x_m$, za neki $m \leq n$.

Dokaz. Tehnički! ∎

Konvergencija metode konjugiranih gradijenata

Iteracije metode konjugiranih gradijenata za $A \in \mathbb{R}^{3\times 3}$.



Konvergencija metode konjugiranih gradijenata

Sada nas zanima kolika je točno A-norma greške u svakom koraku metode konjugiranih gradijenata.

- Zbog relacije $e_k \in e_0 + \operatorname{span}\{Ae_0, A^2e_0, \ldots, A^ke_0\}$, greška u k-tom koraku metode ima oblik
$$e_k = e_0 + \sum_{i=1}^{k} \psi_i A^i e_0 = \left(I + \sum_{i=1}^{k} \psi_i A^i \right) e_0.$$
Koeficijenti ψ_i su u linearnoj vezi sa koeficijentima α_i i β_i, a metoda konjugiranih gradijenata bira ψ_j takve da oni minimiziraju $\|e_k\|_A$.
- Tada, izraz za grešku možemo izraziti kao
$$e_k = p_k(A)e_0,$$
gdje je p_k polinom k-tog stupnja kod kojeg zahtijevamo da je $p_k(0) = 1$.

Konvergencija metode konjugiranih gradijenata

Kako je matrica A simetrična i pozitivno definitna, tada matricu možemo zapisati kao produkt matrica $A = U\Lambda U^T$, pri čemu su

- za $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, $\lambda_1, \ldots, \lambda_n$ svojstvene vrijednosti od A i
- $U^T U = U U^T = I$ je ortogonalna.
- $A^{1/2}$ je simetrični drugi korijen od A i vrijedi $A^{1/2} = U\Lambda^{1/2}U^T$, pa komutira sa A.

Zbog toga slijedi

$$\begin{aligned} \|e_k\|_A &= \min_{p_k \in \mathbb{P}_k, p_k(0)=1} \|p_k(A)e_0\|_A \\ &= \min_{p_k \in \mathbb{P}_k, p_k(0)=1} \|A^{1/2}p_k(A)e_0\|_2 \end{aligned}$$

Konvergencija metode konjugiranih gradijenata

$$\begin{aligned}
\|e_k\|_A &= \min_{p_k \in \mathbb{P}_k, p_k(0)=1} \|U p_k(\Lambda) U^T A^{1/2} e_0\|_2 \\
&\leq \min_{p_k \in \mathbb{P}_k, p_k(0)=1} \|p_k(\Lambda)\|_2 \|e_0\|_A \\
&= \min_{p_k \in \mathbb{P}_k, p_k(0)=1} \max_{i=1,\ldots,n} |p_k(\lambda_i)| \|e_0\|_A.
\end{aligned}$$

- Općenito, polinom n-tog stupnja može biti jednoznačno određen ako je zadan na $n+1$ točaka.
- Ako uzmemo da je $p_n(0) = 1$ i da u svim svojstvenim vrijednostima od A (njih n) p_n poprima vrijednost 0, tada je to polinom koji minimizira gornju ogradu, pri čemu je onda e_n jednak nuli.

Konvergencija metode konjugiranih gradijenata

- U egzaktnoj aritmetici broj iteracija k koje su potrebne da se izračuna egzaktno rješenje je manji ili jednak broju različitih svojstvenih vrijednosti, jer možemo naći polinom stupnja k koji je
  - u 0 jednak 1, i
  - u svim različitim svojstvenim vrijednostima jednak 0.
- Za takav polinom je $\max_{i=1,\ldots,n} |p_k(\lambda_i)| = 0$, pa iz prethodne ocjene slijedi da je $e_k = 0$.
- Time se vidi da je metoda konjugiranih gradijenata brža ako A ima

višestrukih svojstvenih vrijednosti.

Konvergencija metode konjugiranih gradijenata

- Sve svojstvene vrijednosti od A smještene su u segmentu $[\lambda_{min}, \lambda_{max}]$, pri čemu je $\lambda_{min} > 0$ najmanja, a $\lambda_{max} > 0$ najveća svojstvena vrijednost.
- Budući da je ograda na A-normu greške manja ili jednaka minimizaciji pa maksimizaciji po cijelom intervalu $[\lambda_{min}, \lambda_{max}]$, jednostavniji pristup je minimizacija po tom segmentu nego po konačnom broju točaka.
- Polinomi s kojima se to postiže bazirani su na Čebiševljevim polinomima.
- Čebiševljev polinom stupnja k je

$$T_k(\omega) = \frac{1}{2}\left[(\omega + \sqrt{\omega^2 - 1})^k + (\omega - \sqrt{\omega^2 - 1})^k\right]$$

$$T_k(\omega) = \cos(k \arccos(\omega)), \quad \omega \in [-1, 1].$$

Konvergencija metode konjugiranih gradijenata

- Čebiševljevi polinomi imaju sljedeća svojstva:

$$|T_k(\omega)| \leq 1, \qquad \omega \in [-1, 1],$$

rapidno osciliraju između -1 i 1, odnosno

$$T_k\left(\cos\left(\frac{i\pi}{k}\right)\right) = (-1)^i, \qquad i = 0, 1, \ldots, k,$$

i k nultočaka polinoma T_k moraju se nalaziti između $k+1$ ekstrema od T_k u segmentu $[-1, 1]$.

- Najbitnije mu je svojstvo to da između svih normiranih polinoma stupnja k ($a_k = 1$) polinom $2^{1-k} T^k$ ima najmanju ∞-normu na intervalu $[-1, 1]$, koja iznosi 2^{1-k}.

Konvergencija metode konjugiranih gradijenata

- U ovom slučaju, polinom k-tog stupnja koji
  - ima vrijednost 1 u ishodištu, i
  - ima minimalnu ∞-normu na $[\lambda_{min}, \lambda_{max}]$

  je dan sa

$$p_k(\lambda) = \frac{T_k\left(\frac{\lambda_{max}+\lambda_{min}-2\lambda}{\lambda_{max}-\lambda_{min}}\right)}{T_k\left(\frac{\lambda_{max}+\lambda_{min}}{\lambda_{max}-\lambda_{min}}\right)}.$$

- Argument brojnika je takav da segment $[\lambda_{min}, \lambda_{max}]$ prebacuje u segment $[-1, 1]$.
- Slijedi da za $z \in [\lambda_{min}, \lambda_{max}]$ apsolutna vrijednost brojnika je ograničena sa 1.
- Nazivnik je određen tako da zadovolji svojstvo $p_k(0) = 1$.

Konvergencija metode konjugiranih gradijenata

- Znači, imamo sljedeće

$$\|e_k\|_A \leq T_k \left(\frac{\lambda_{max} + \lambda_{min}}{\lambda_{max} - \lambda_{min}} \right)^{-1} \|e_0\|_A = T_k \left(\frac{\kappa(A) + 1}{\kappa(A) - 1} \right)^{-1} \|e_0\|_A,$$

i nakon što izračunamo vrijednost Čebiševljevog polinoma, dobivamo

$$\|e_k\|_A \leq 2 \left[\left(\frac{\sqrt{\kappa(A)} + 1}{\sqrt{\kappa(A)} - 1} \right)^k + \left(\frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1} \right)^k \right]^{-1} \|e_0\|_A$$

Konvergencija metode konjugiranih gradijenata

odnosno

$$\|e_k\|_A \leq 2 \left(\frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1} \right)^k \|e_0\|_A.$$

pri čemu je $\kappa(A) = \lambda_{max}/\lambda_{min}$ broj uvjetovanosti matrice A.

- Ovo je slabija ograda od rezultata sa višestrukim svojstvenim vrijednostima!
- Loša konvergencija se ostvaruje za loše uvjetovane matrice.

Prekondicionirana metode konjugiranih grad.

- Kada rješavamo simetrični sustav, želimo da se to svojstvio zadrži i nakon prekondicioniranja.
- Zato biramo matricu prekondicioniranja M koja je simetrična i pozitivno definitna.
- Ako matricu M faktoriziramo kao $M = LL^T$, tada algoritam možemo primijeniti na sustav

  $$L^{-1}AL^{-T}\bar{x} = L^{-1}b,$$

  koji i dalje ostaje simetrični, a svojstvene vrijednosti matrice $L^{-1}AL^{-T}$ su iste kao i kod $M^{-1}A$.
- Matricu M također biramo i uz kriterij da $L^{-1}AL^{-T}$ ima
  - što sabitije svojstene vrijednosti, i
  - što manju uvjetovanost.

Prekondicionirana metode konjugiranih grad.

- Ako sada označimo sve veličine vezane uz prekondicionirani sustav sa $\bar{}$, a veličine vezane uz početni sustav $Ax = b$ sa standardnim oznakama, tada uz pomoć relacija

  $$x_k = L^{-T} \bar{x}_k, \quad r_k = L \bar{r}_k, \quad d_k = L^{-T} \bar{d}_k,$$

  metodu konjugiranih gradijenata primijenjenu na prekondicionirani sustav možemo transformirati u algoritam, koji ne ovisi o faktorizaciji matrice prekondicioniranja M.

Algoritam prekondicionirane metode konj. grad.

Algoritam (Prekondicionirana metoda konj. grad.).

x_0 `zadan;`
$k = 0$;
$r_0 = b - Ax_0$;
`riješi` $Mp_0 = r_0$; $d_0 = p_0$;
`dok nije zadovoljen kriterij zaustavljanja radi {`

$$\alpha_k = \frac{r_k^T p_k}{d_k^T A d_k};$$

$$x_{k+1} = x_k + \alpha_k d_k;$$

$$r_{k+1} = r_k - \alpha_k A d_k;$$

`riješi` $Mp_{k+1} = r_{k+1}$;

$$\beta_{k+1} = \frac{r_{k+1}^T p_{k+1}}{r_k^T p_k};$$

$$d_{k+1} = p_{k+1} + \beta_{k+1} d_k;$$

$$k = k + 1;$$

`}`

Numerička stabilnost metode konj. grad.

Teorem. Razlika između pravog reziduala $b - A\hat{x}_k$, za izračunatu aproksimaciju $\hat{x}_k$, i izračunatog vektora $\hat{r}_k$ zadovoljava nejednakost

$$\frac{\|b - A\hat{x}_k - \hat{r}_k\|_2}{\|A\|_2 \|x\|_2} \leq$$

$$\leq (u + \mathcal{O}(u^2)) \left[k + 1 + (1 + c + k(10 + 2c)) \left(2 + \frac{\|x_0\|_2}{\|x\|_2} \right) \right],$$

gdje je $c = \sqrt{n}\gamma_n$.

Napomena. Promotrimo sada što se događa kada je algoritam blizu konvergencije,

- budući da u većini slučajeva možemo očekivati da $\hat{r}_k$ teži k 0 kada $k \to \infty$.

Numerička stabilnost metode konj. grad.

- Jednom kada se $\hat{r}_k$ reducira ispod određenog stupnja, kada je vrlo mali, aporoksimacija rješenja $\hat{x}_{k+1}$ ostaje približno nepromijenjna u odnosu na prethodnu iteraciju.
- To je zbog toga što je norma izraza, koji mijenja vrijednost aproxsimacije $\hat{x}_k$ u $\hat{x}_{k+1}$, u bliskoj vezi sa veličinom vektora $\hat{r}_k$, kao što se može vidjeti iz

$$\alpha_k \hat{d}_k = A^{-1}(\hat{r}_k - \hat{r}_{k+1} + \eta_{k+1}), \quad \eta_{k+1} \text{ je greška zaokruživanja.}$$

- Iz ovoga slijedi da kada vektori $\hat{r}_k$, a prema tome i $\alpha_k \hat{d}_{k-1}$, konvergiraju ka nuli, i kada su iteracije algoritma prošle točku u kojoj $\|A^{-1}\hat{r}_k\|_2$ dostiže $\mathcal{O}(u)\|\hat{x}_k\|_2$, pravi rezidual $b - A\hat{x}_k$ neće ovisiti o k, kao što pretpostavlja ocjena Teorema, već će ostati skoro konstantan.

Numerička stabilnost metode konj. grad.

- Na žalost, svojstvo ortogonalnosti reziduala, koje se koristi za dobivanje redukcije euklidske norme greške može se u potpunosti izgubiti u aritmetici konačne preciznosti.
- Međutim, može se napraviti analogna analiza i u tom slučaju, što nam omogućava sličan zaključak i za aritmetiku konačne preciznosti.

GMRES metoda

GMRES metoda

- GMRES metoda (Generalized minimal residual algorithm) ili generalizirani algoritam minimalnog reziduala koristi modificirani Gram–Schmidtov postupak kako bi konstruirao ortonormiranu bazu za niz Krylovljevih potprostora

$$\mathcal{K}_k(A, r_0) = \mathrm{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}.$$

- Kada se modificirani Gram–Schmidtov postupak primijeni na ovakav prostor tada govotimo o Arnoldijevoj metodi.

Algoritam Arnoldijeve metode

Algoritam (Arnoldijeva metoda).

q_1 sa $\|q_1\|_2 = 1$ zadan;
za $j = 1$ do $n-1$ radi $\{$
    $\tilde{q}_{j+1} = Aq_j$;
    za $i = 1$ do j radi $\{$
        $h_{ij} = q_i^T \tilde{q}_{j+1}$;
        $\tilde{q}_{j+1} := \tilde{q}_{j+1} - h_{ij} q_i$;
    $\}$
    $h_{j+1,j} = \|\tilde{q}_{j+1}\|_2$;
    $q_{j+1} = \dfrac{\tilde{q}_{j+1}}{h_{j+1,j}}$;
$\}$

Arnoldijeva metoda

- Ako je Q_k $n \times k$ matrica čije stupce čine vektori ortonormirane baze $q_1, \ldots, q_k$, tada se Arnoldijeva iteracija može napisati u matričnom obliku

$$AQ_k = Q_k H_k + h_{k+1,k} q_{k+1} \xi_k^T = Q_{k+1} H_{k+1,k}.$$

- Ovdje je H_k $k \times k$ gornja Hessenbergova matrica kojoj je (i, j)-ti element jednak $h_{i,j}$ za $j = 1, \ldots, k$, $i = 1, \ldots, \min\{j + 1, k\}$, a svi ostali elementi su jednaki 0.
- Vektor ξ_k je k-ti jedinični vektor $[0 \ \ldots \ 0 \ 1]^T$.
- $(k + 1) \times k$ matrica $H_{k+1,k}$ je matrica kojoj je gornji $k \times k$ blok jednak H_k, a zadnji redak jednak nuli, osim na poziciji $(k + 1, k)$, na kojoj je element jednak $h_{k+1,k}$.

Arnoldijeva metoda (nastavak)

$$Q_k = [\ q_1 \quad q_2 \quad \cdots \quad q_k\ ]$$

$$H_{k+1,k} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & \cdots & \cdots & h_{1k} \\ h_{21} & h_{22} & h_{23} & \cdots & \cdots & h_{2k} \\ 0 & h_{32} & h_{33} & \cdots & \cdots & h_{3k} \\ 0 & 0 & h_{43} & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & h_{k,k-1} & h_{kk} \\ \hdashline 0 & 0 & \cdots & 0 & 0 & h_{k+1,k} \end{bmatrix} \left.\vphantom{\begin{matrix} a\\a\\a\\a\\a\\a \end{matrix}}\right\} H_k$$

Arnoldijeva metoda (nastavak)

- Rekurzivna definicija za matricu Q_{k+1} može se u matričnom obliku napisati kao

$$[q_1 \quad AQ_k] = Q_{k+1}R_{k+1},$$

pri čemu je R_{k+1} $(k+1) \times (k+1)$ gornje trokutasta matrica

$$R_{k+1} = [\xi_1 \quad H_{k+1,k}],$$

a $\xi_1 = [1, 0, \ldots, 0]^T$.

- Prema tome Arnoldijev algoritam od prvog do $(k+1)$-og koraka možemo smatrati rekurzivnom QR faktorizacijom matrice $[q_1 \quad AQ_k]$.

GMRES metoda (nastavak)

- U GMRES metodi, aproksimacija rješenja u k-tom koraku tada ima oblik

$$x_k = x_0 + Q_k y_k$$

za neki k-dimenzionalni vektor y_k, tj.

  - x_k se dobiva tako da se početnoj aproksimaciji x_0 doda neki vektor iz Krylovljevog potprostora.

- Sada tražimo aproksimaciju za koju $r_k = r_0 - AQ_k y_k$ ima minimalnu euklidsku normu, i za nju vektor y_k mora biti rješenje problema najmanjih kvadrata

$$\min_{y\in\mathbb{R}^k} \|r_0 - AQ_k y\|_2 = \min_{y\in\mathbb{R}^k} \|r_0 - Q_{k+1}H_{k+1,k}y\|_2$$

$$= \min_{y\in\mathbb{R}^k} \|Q_{k+1}(\beta\xi_1 - H_{k+1,k}y)\|_2 = \min_{y\in\mathbb{R}^k} \|\beta\xi_1 - H_{k+1,k}y\|_2,$$

GMRES metoda (nastavak)

gdje je $\beta = \|r_0\|_2$, ξ_1 prvi jedinični $(k+1)$-dimenzionalni vektor $[1\ 0\ \ldots\ 0]^T$.

- Druga jednakost je ostvarena uz korištenje činjenice da je $Q_{k+1}\xi_1 = q_1$ prvi vektor ortonormirane baze koji je jednak $q_1 = r_0/\beta$.

Glavni koraci GMRES algoritma su sljedeći:

- Dana je početna iteracija x_0, izračunaj $r_0 = b - Ax_0$ i $q_1 = r_0/\|r_0\|_2$.
- Za k=1,2,...
  - Izračunaj q_{k+1} i h_{ik}, $i = 1, \ldots, k+1$ pomoću Arnoldijevog algoritma.
  - Nađi $x_k = x_0 + Q_k y_k$, gdje je y_k rješenje problema najmanjih kvadrata $\min_y \|\beta\xi_1 - H_{k+1,k}y\|_2$.

GMRES metoda (nastavak)

- Kao što ćemo vidjeti standardna metoda za rješavanje problema najmanjih kvadrata sa matricom punog ranga je QR faktorizacija $(k+1) \times k$ matrice $H_{k+1,k}$.
- U jednom od sljedećih teorema uvjerit ćemo se da je $H_{k+1,k}$ punog ranga za
  - regularnu matricu A, i
  - u slučaju kada x_k još nije rješenje sustava.

  $H_{k+1,k}$ se svodi na produkt
  - $(k+1) \times (k+1)$ ortogonalne matrice $G^{(k)T}$ i
  - $(k+1) \times k$ gornje trokutaste matrice $R^{(k)}$.
- Gornji $k \times k$ blok matrice $R^{(k)}$ je gornje trokutasti, a zadnji redak je jednak nul-vektoru.

GMRES metoda (nastavak)

$$H_{k+1,k} = G^{(k)T} R^{(k)} = G^{(k)T} \begin{bmatrix} R^{(k)}_{k \times k} \\ 0 \end{bmatrix}$$

- Ta QR faktorizacija može se ostvariti upotrebom Givensovih rotacija jer imamo samo jedan netrivijalan element ispod dijagonale.
- Budući da je

$$\begin{aligned} \|\beta\xi_1 - H_{k+1,k} y\|_2 &= \|G^{(k)T}(\beta G^{(k)}\xi_1 - R^{(k)} y)\|_2 \\ &= \|\beta G^{(k)}\xi_1 - R^{(k)} y\|_2, \end{aligned}$$

i da je $(k+1)$-a koordinata vektora $R^{(k)}y$ jednaka nuli, rješenje y_k se tada može dobiti rješavanjem gornje trokutastog sustava

GMRES metoda (nastavak)

$$R^{(k)}_{k\times k}y = \beta(G^{(k)}\xi_1)_{k\times 1},$$

gdje je $R^{(k)}_{k\times k}$ gornji $k \times k$ blok od $R^{(k)}$, a $(G^{(k)}\xi_1)_{k\times 1}$ je dio od k gornjih elemenata prvog stupca od $G^{(k)}$.

- Važno je primijetiti da je u tom slučaju apsolutna vrijednost zadnjeg elementa $(k+1)$-dimenzionalnog vektora $\beta G^{(k)}\xi_1$ jednaka euklidskoj normi reziduala u k-tom koraku GMRES metode jer

$$\|b - Ax_k\|_2 = \|\beta G^{(k)}\xi_1 - R^{(k)}y_k\|_2,$$

a $\beta G^{(k)}\xi_1 - R^{(k)}y_k$ je jednak nuli u svim komponentama osim u zadnjoj, $(k+1)$-oj, koja je jednaka zadnjem elementu od $\beta G^{(k)}\xi_1$.

GMRES metoda (nastavak)

Promotrimo sada samu QR faktorizaciju matrice $H_{k+1,k}$.

- Ako nam je dana QR faktorizacija matrice $H_{k,k-1}$, tada želimo izračunati QR faktorizaciju sljedeće matrice $H_{k+1,k}$ sa što manje utrošenog posla.
- Da bismo to postigli najprije označimo sa G_i Givensovu rotaciju koja rotira ravninu razapetu sa jediničnim vektorima ξ_i i ξ_{i+1} za kut θ_i:

$$G_i = \begin{bmatrix} I & & & \\ & c_i & s_i & \\ & -s_i & c_i & \\ & & & I \end{bmatrix},$$

gdje je $c_i = \cos(\theta_i)$ i $s_i = \sin(\theta_i)$.

GMRES metoda (nastavak)

- Dimenzija matrice G_i kao i dimenzija drugog identičnog bloka ovisi o kontekstu u kojem se koristi.
- Pretpostavimo da su rotacije G_i, $i = 1, \ldots, k$ prethodno bile upotrebljene na matrici $H_{k,k-1}$ tako da je

$$(G_{k-1}G_{k-2}\cdots G_1)H_{k,k-1} = R^{(k-1)} = \begin{bmatrix} x & x & \ldots & x \\ & x & \ldots & x \\ & & \ddots & \vdots \\ & & & x \\ 0 & 0 & \ldots & 0 \end{bmatrix},$$

gdje x-evi označavaju netrivijalne elemente.

GMRES metoda (nastavak)

- Tada je
$$G^{(k-1)} = G_{k-1} G_{k-2} \cdots G_1.$$
- Da bismo dobili matricu $R^{(k)}$, prvo trebamo izmnožiti zadnji stupac od $H_{k+1,k}$ sa prethodnim rotacijama, jer se prvih $k-1$ stupaca te matrice poklapa sa $\begin{bmatrix} H_{k,k-1} \\ 0 \end{bmatrix}$.
- Trokutasti faktor matrice u k-tom koraku je zbog toga jednak trokutastom faktoru matrice iz $(k-1)$-og koraka, samo kojemu je također dodan još jedan stupac, pa zato je dovoljno obraditi samo novi, k-ti stupac matrice $H_{k+1,k}$.

GMRES metoda (nastavak)

- Time dobivamo

$$(G_{k-1}G_{k-2}\cdots G_1)H_{k+1,k} = \begin{bmatrix} x & x & \dots & x & x \\ & x & \dots & x & x \\ & & \ddots & \vdots & \vdots \\ & & & x & x \\ 0 & 0 & \dots & 0 & d \\ 0 & 0 & \dots & 0 & h \end{bmatrix},$$

  čiji je $(k+1,k)$-i element h upravo $h_{k+1,k}$ budući da na taj element ne utječu prethodne rotacije.

- Sljedeća rotacija G_k se bira tako da eliminira taj element, odnosno da vrijedi

$$-s_k d + c_k h = 0,$$

GMRES metoda (nastavak)

a to možemo postići ako stavimo da je

$$c_k = \frac{|d|}{\sqrt{d^2 + h^2}}, \quad s_k = \frac{c_k h}{d}, \qquad \text{za } d \neq 0,\ h \neq 0,$$

$$c_k = 0, \quad s_k = 1, \qquad \text{za } d = 0,$$

$$(c_k = 1, \quad s_k = 0, \qquad \text{za } h = 0).$$

- Primijetimo da ako je $h = 0$, tada je egzaktno rješenje linearnog sustava dostignuto u k-tom koraku.
- Naime, u tom slučaju je $G_k = I$, pa je zadnja $(k + 1)$-a komponenta vektora $\beta G^{(k)} \xi_1$ ostala nepromijenjena, odnosno jednaka nuli.

GMRES metoda (nastavak)

- Budući da je apsolutna vrijednost te zadnje komponente jednaka euklidskoj normi reziduala u tom koraku GMRES metode, zaključujemo da je $r_k = 0$.
- Prema tome, ako egzaktno rješenje još nije izračunato, tada je $h \neq 0$, i k-ti dijagonalni element od $R^{(k)}$ je različit od nule.
  - Za $d \neq 0$ taj dijagonalni element je jednak

$$c_k d + s_k h = \frac{d}{|d|} \sqrt{d^2 + h^2},$$

  - dok za $d = 0$ on je jednak h.

Algoritam GMRES metode

Algoritam (GMRES metoda).

x_0 zadan;

$k = 0$;

$r_0 = b - Ax_0$; $\beta = \|r_0\|_2$; $q_1 = \dfrac{r_0}{\beta}$;

$\ell = [1, 0, \ldots, 0]^T$;

dok nije zadovoljen kriterij zaustavljanja radi {

$k = k + 1$;

Izračunaj q_{k+1} i $H(i, k) = h_{ik}$ za $i = 0, \ldots, k + 1$,

koristeći Arnoldijev algoritam.

Primijeni $G_0, \ldots, G_{k-1}$ na zadnji stupac od H:

za $i = 1$ do $k - 1$ radi {

$$\begin{bmatrix} H(i,k) \\ H(i+1,k) \end{bmatrix} = \begin{bmatrix} c_i & s_i \\ -s_i & c_i \end{bmatrix} \begin{bmatrix} H(i,k) \\ H(i+1,k) \end{bmatrix};$$

}

Algoritam GMRES metode (nastavak)

```
Izračunaj k-tu Givensovu rotaciju G_k kako bi se
    poništio H(k+1,k):
```

$$c_k = \frac{|H(k,k)|}{\sqrt{H(k,k)^2 + H(k+1,k)^2}};$$

ako je $c_k \neq 0$ tada

$$s_k = c_k \frac{H(k+1,k)}{H(k,k)};$$

inače

$$s_k = 1;$$

Primijeni G_k na ℓ i na k-ti stupac od H:

$$\begin{bmatrix} \ell(k) \\ \ell(k+1) \end{bmatrix} = \begin{bmatrix} c_k & s_k \\ -s_k & c_k \end{bmatrix} \begin{bmatrix} \ell(k) \\ 0 \end{bmatrix};$$

Algoritam GMRES metode (nastavak)

```
            H(k,k) = c_k H(k,k) + s_k H(k+1,k);
            H(k+1,k) = 0;
}
ako je (ocjena norme reziduala β|ℓ(k+1)|
     dovoljno mala) tada
     Riješi gornje trokutasti sustav
          H(1:k,1:k) y_k = β ℓ(1:k).
     x_k = x_0 + Q_k y_k;
}
```

Svojstva GMRES metode

- Potpun GMRES algoritam može biti nepraktičan zbog
  - velikog zahtjeva za memorijom i
  - brojem operacija,

  ako je broj iteracija koje su potrebne za rješavanje sustava velik.
- GMRES(j) algoritam se definira tako da restartamo GMRES svakih j koraka, koristeći zadnju iteraciju kao početnu za sljedeći GMRES ciklus.
- Poznati problem vezan uz GMRES sa restartom je taj što on može stagnirati.
- Naime može se dogoditi da pri restartanju ponovno pretražujemo smjerove koje smo već u prethodnom ciklusu pretraživali.

Svojstva GMRES metode (nastavak)

Teorem. Neka su G_i, $i = 1, \ldots, k$ matrice rotacija koje su korištene za svođenje matrice $H_{k+1,k}$ na trokutasti oblik, te $R^{(k)} = G^{(k)} H_{k+1,k}$ i $g^{(k)} = \beta G^{(k)} \xi_1$ matrica i desna strana sustava dobivenog tokom rješavanja GMRES metodom. Označimo sa $R^{(k)}_{k\times k}$ $k \times k$ gornje trokutastu matricu dobivenu iz $R^{(k)} = [R^{(k)}_{ij}]$ brisanjem zadnjeg retka, i sa $g^{(k)}_{k\times 1}$ k-dimenzionalan vektor dobiven iz $g^{(k)} = [g^{(k)}_i]$ brisanjem zadnje komponente. Tada,

- Rang od AQ_k je jednak rangu od $R^{(k)}_{k\times k}$. Posebno, ako je $R^{(k)}_{k,k} = 0$ tada A mora biti singularna.
- Vektor y_k koji minimizira $\|\beta\xi_1 - H_{k+1,k}y\|_2$ dan je sa

$$y_k = (R^{(k)}_{k\times k})^{-1} g^{(k)}_{k\times 1}.$$

Svojstva GMRES metode (nastavak)

- Rezidual u k-tom koraku zadovoljava

$$r_k = b - Ax_k = Q_{k+1}(\beta\xi_1 - H_{k+1,k}y_k) = Q_{k+1}G^{(k)T}(g_{k+1}^{(k)}\xi_{k+1})$$

i, kao rezultat

$$\|r_k\|_2 = \|b - Ax_k\|_2 = |g_{k+1}^{(k)}|.$$

Dokaz.

- Vrijedi

$$AQ_k = Q_{k+1}H_{k+1,k} = Q_{k+1}G^{(k)T}G^{(k)}H_{k+1,k} = Q_{k+1}G^{(k)T}R^{(k)}.$$

Budući da je $Q_{k+1}G^{(k)T}$ ortogonalna, rang od AQ_k je jednak rangu od $R^{(k)}$, koji je opet jednak rangu od $R_{k\times k}^{(k)}$, budući da se te dvije matrice razlikuju samo za zadnji nul–redak od $R^{(k)}$.

Svojstva GMRES metode (nastavak)

Ako je pak $R^{(k)}_{k,k} = 0$ tada je rang od $R^{(k)}_{k\times k}$ manji ili jednak $k-1$, a kao rezultat je i rang od AQ_k manji ili jednak $k-1$. Kako je Q_k punog ranga, to znači da je A singularna.

- To je već uglavnom pokazano kod izvoda metode. Ako smo krenuli od regularne matrice A tada je i $R^{(k)}_{k\times k}$ regularna, pa možemo iračunati rješenje $y_k = (R^{(k)}_{k\times k})^{-1} g^{(k)}_{k\times 1}$.
- Za bilo koji $x = x_0 + Q_k y$ vrijedi

$$\begin{aligned} b - Ax &= Q_{k+1}(\beta\xi_1 - H_{k+1,k}y) \\ &= Q_{k+1}G^{(k)T}G^{(k)}(\beta\xi_1 - H_{k+1,k}y) \\ &= Q_{k+1}G^{(k)T}(g^{(k)} - R^{(k)}y). \end{aligned}$$

Svojstva GMRES metode (nastavak)

Za normu tada vrijedi

$$\|b - Ax\|_2^2 = \|g^{(k)} - R^{(k)}y\|_2^2 = |g_{k+1}^{(k)}|^2 + \|g_{k\times 1}^{(k)} - R_{k\times k}^{(k)}y\|_2^2$$

pa se minimum upravo postiže za y_k kada je drugi član gornjeg izraza jednak nuli.

Svojstva GMRES metode (nastavak)

Napomena.

- Primijetimo da, kada vektor $\beta\xi_1$ izmnožimo redom sa $G_1, G_2, \ldots, G_{k-1}$, zadnja $k+1$-a komponenta tako dobivenog vektora $g^{(k,k-1)}$ ostaje nepromijenjena, odnosno jednaka nuli.
- To znači da je on oblika

$$g^{(k,k-1)} = [g_1^{(k)}\ g_2^{(k)}\ \ \ldots\ \ g_{k-1}^{(k)}\ g_k^{(k-1)}\ \ 0]^T.$$

- U zadnjem koraku, kada taj vektor na kraju još pomnožimo i sa G_k, dobivamo da je

$$g^{(k)} = [g_1^{(k)}\ g_2^{(k)}\ \ \ldots\ \ g_{k-1}^{(k)}\ c_k g_k^{(k-1)}\ \ -s_k g_k^{(k-1)}],$$

Svojstva GMRES metode (nastavak)

odnosno dobivamo korisnu jednakost za $g_{k+1}^{(k)}$

$$g_{k+1}^{(k)} = -s_k g_k^{(k-1)}.$$

- Posebno, ako je $s_k = 0$ tada norma reziduala mora biti jednaka nuli, što znači da je egzaktno rješenje dostignuto u k-tom koraku.
- Iz Algoritma se vidi da je $s_k = 0$ samo ako je $h_{k+1,k} = 0$.
- S druge strane, jedini mogući slučaj kada k-ti korak algoritma neće biti izvediv, i kada će doći do prekida GMRES metode, je prekid Arnoldijevog algoritma kada je $\tilde{q}_{k+1} = 0$, odnosno $h_{k+1,k} = 0$.
- U tom slučaju se algoritam zaustavlja jer se sljedeći Arnoldijev vektor neće moći generirati zbog dijeljenja s nulom.

Svojstva GMRES metode (nastavak)

- Ali u tom je slučaju, kao što smo već vidjeli, rezidual je jednak nuli, pa smo dostigli egzaktno rješenje u tom koraku.
- Obrat je također istinit: ako se algoritam zaustavi u k-tom koraku sa $r_k = b - Ax_k = 0$, tada je $h_{k+1,k} = 0$.

Svojstva GMRES metode (nastavak)

Teorem. Neka je A regularna matrica. Tada se GMRES algoritam prekida u k-tom koraku ($h_{k+1,k} = 0$, $k \leq n$) ako i samo ako je aproksimacija x_k jednaka egzaktnom rješenju.

Dokaz. Nužnost smo već pokazali.

- Da bismo dokazali dovoljnost tvrdnje, pretpostavimo da je $r_k = 0$ i pri tome koristimo $g_{k+1}^{(k)} = -s_k g_k^{(k-1)}$.
- Budući da smo u k-tom koraku dostigli egzaktno rješenje, a ne u $(k-1)$-tom, znači da je $g_k^{(k,k-1)} = g_k^{(k-1)}$ različito od nule, pa onda mora biti $s_k = 0$.
- Ponovo iz Algoritma vidimo da onda mora biti $h_{k+1,k} = 0$.

$\blacksquare$

Konvergencija GMRES metode

GMRES algoritam primijenjen na općeniti linearni susutav daje u k-tom koraku rezidual koji zadovoljava, slično kao kod CG,

$$\|r_k\|_2 = \min_{p_k \in \mathbb{P}_k, p_k(0)=1} \|p_k(A) r_0\|_2,$$

gdje se minimum uzima po svim polinomima p_k stupnja k ili manje, uz svojstvo da je $p_k(0) = 1$.

Teorem. Neka je x_k aproksimacija rješenja ostvarena u k-tom koraku GMRES algoritma, i neka je $r_k = b - Ax_k$. Tada postoji $q_{k-1} \in \mathbb{P}_{k-1}$ takav da je x_k oblika

$$x_k = x_0 + q_{k-1}(A) r_0$$

i

$$\|r_k\|_2 = \min_{q_{k-1} \in \mathbb{P}_{k-1}} \|(I - A q_{k-1}(A)) r_0\|_2.$$

Konvergencija GMRES metode (nastavak)

Dokaz.

- Iz definicije metode u k-tom koraku bira se aproksimacija rješenja x_k takva da za $\mathcal{K}_k(A, r_0) = \text{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}$ vrijedi

$$\|b - Ax_k\|_2 = \min_{x \in x_0 + \mathcal{K}_k(A, r_0)} \|b - Ax\|_2,$$

  (znamo da je x_k oblika $x_k = x_0 + Q_k y_k$, pri čemu stupci od Q_k čine bazu Krylovljevog potprostora $\mathcal{K}_k(A, r_0)$).
- Kako je svaki $x \in x_0 + \mathcal{K}_k(A, r_0)$ oblika $x = x_0 + \sum_{j=0}^{k-1} \alpha_j A^j r_0$, tada možemo pisati

$$x = x_0 + q_{k-1}(A) r_0, \qquad \text{za } q_{k-1} \in \mathbb{P}_{k-1}.$$

Konvergencija GMRES metode (nastavak)

- U tom slučaju je

$$\begin{aligned} r &= b - Ax = b - Ax_0 - Aq_{k-1}(A) r_0 \\ &= (I - Aq_{k-1}(A)) r_0 = p_k(A) r_0, \end{aligned}$$

  za $p_k \in \mathbb{P}_k$, sa $p_k(0) = 1$.

■

Napomena. Za GMRES metodu ne postoji "lijepi" teorem o konvergenciji za općenite matrice. Ipak, nešto se može reći za dijagonalizabilne matrice.

Konvergencija GMRES metode (nastavak)

Teorem. Pretpostavimo da je A dijagonalizabilna matrica i neka je $A = V\Lambda V^{-1}$ spektralna dekompozicija od A, gdje je $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ dijagonalna matrica svojstvenih vrijednosti, a stupci regularne matrice V su svojstveni vektori od A. Definirajmo,

$$\epsilon_k = \min_{p_k \in \mathbb{P}_k, p_k(0)=1} \max_{i=1,\ldots,n} |p_k(\lambda_i)|.$$

Tada, norma reziduala postignutog u k-tom koraku GMRES metode zadovoljava nejednakost

$$\|r_k\|_2 \le \kappa(V)\epsilon_k \|r_0\|_2,$$

gdje je $\kappa(V) = \|V\|_2 \|V^{-1}\|_2$.

Konvergencija GMRES metode (nastavak)

Dokaz. Neka je p_k bilo koji polinom stupnja manjeg ili jednakog k, koji zadovoljava uvjet $p_k(0) = 1$, i neka je $x \in x_0 + \mathcal{K}_k(A, r_0)$ vektor za koji je $r = b - Ax = p_k(A)r_0$ iz $\mathcal{K}_{k+1}(A, r_0)$. Tada

$$\|b - Ax\|_2 = \|V p_k(\Lambda) V^{-1} r_0\|_2 \le \|V\|_2 \|V^{-1}\|_2 \|p_k(\Lambda)\|_2 \|r_0\|_2.$$

Budući da je Λ dijagonalna matrica, vrijedi

$$\|p_k(\Lambda)\|_2 = \max_{i=1,\ldots,n} \|p_k(\lambda_i)\|_2.$$

Kako x_k minimizira normu reziduala na $x_0 + \mathcal{K}_k(A, r_0)$, tada za bilo koji polinom p_k sa gornjim svojstvima vrijedi

$$\|b - Ax_k\|_2 \le \|b - Ax\|_2 \le \|V\|_2 \|V^{-1}\|_2 \|r_0\|_2 \max_{i=1,\ldots,n} |p_k(\lambda_i)|.$$

Konvergencija GMRES metode (nastavak)

Sada možemo u gornjoj nejednakosti uzeti polinom p_k koji minimizira desnu stranu nejednakosti, to odgovara odabiru odgovarajućeg $x_{min} \in x_0 + \mathcal{K}_k(A, r_0)$. Time dobivamo željeni rezultat

$$\|b - Ax_k\|_2 \le \|b - Ax_{min}\|_2 \le \|V\|_2 \|V^{-1}\|_2 \|r_0\|_2 \epsilon_k.$$

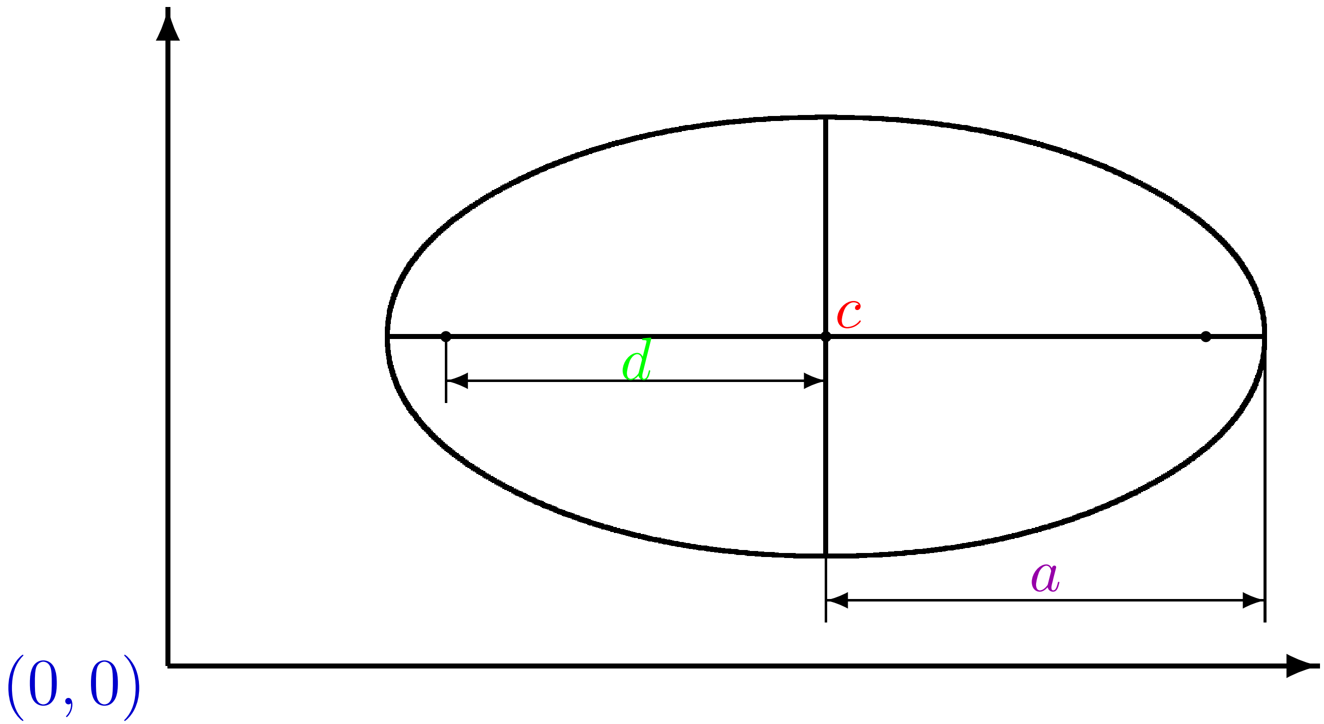
■

Napomena.

- Promatrat ćemo one matrice kojima spektar možemo smjestiti u neki pravilni geometrijski lik, koji je po mogućnosti udaljen od ishodišta (ne sadrži ga).
- Najjednostavniji slučaj je kada spektar matrice A možemo smjestititi unutar elipse $\mathbf{E}(c,d,a)$ sa centrom u c, gdje je udaljenost između fokusa jednaka $2d$, a veća poluos je a.

Konvergencija GMRES metode (nastavak)

Elipsa $\mathbf{E}(c,d,a)$, koja ne sadrži ishodište.



Konvergencija GMRES metode (nastavak)

Korolar. Neka je A dijagonalizabilna matrica, to jest neka je $A = V\Lambda V^{-1}$ gdje je $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ dijagonalna matrica svojstvenih vrijednosti. Pretpostavimo da su sve svojstvene vrijednosti od A smještene unutar elipse $\mathbf{E}(c, d, a)$ u kojoj se ne nalazi ishodište. Tada norma reziduala, postignutog u k-tom koraku GMRES metode, zadovoljava nejednakost

$$\|r_k\|_2 \leq \kappa(V) \frac{T_k\left(\frac{a}{d}\right)}{\left|T_k\left(\frac{c}{d}\right)\right|} \|r_0\|_2,$$

gdje je T_k Čebiševljev polinom stupnja k.

Dokaz. Tehnički! ∎

Konvergencija GMRES metode (nastavak)

Napomena.

- Budući da uvjetovanost $\kappa(V)$ matrice svojstvenih vektora V obično nije poznata i može biti vrlo velika, ocjena prethodnog Korolara ne mora biti precizna.
- Ona može biti korisna jedino ako znamo da je matrica A normalna, pa je tada $\kappa(V) = 1$.
- U tom slučaju dobili smo vrlo primjenljivu ocjenu konvergencije samo pomoću analize distribucije svojstvenih vrijednosti.
- U normalnom slučaju, kao i kod simetričnih matrica, problem konvergencije GMRES metode općenito svodi se na problem teorije aproksimacija: kako se dobro može aproksimirati nula na skupu kompleksnih svojstvenih vrijednosti, koristeći polinome stupnja k koji u ishodištu poprimaju vrijednost 1 (isto kao kod CG).

Konvergencija GMRES metode (nastavak)

- U ovom problemu može se dobro primijeniti informacija o tome da li su svojstvene vrijednosti dobro ili loše distribuirane u kompleksnoj ravnini.
- Svojstvene vrijednosti nakupljene u uskom području oko jedne točke c, daleke od ishodišta daju dobru distribuciju, budući da je vrijednost polinoma $(1 - z/c)^k$ mala u svim točkama kompleksne ravnine koje su bliske točki c (prethodni Korolar).
- Svojstvene vrijednosti koje su raspoređene svuda oko ishodišta su loše distribuirane, jer je nemoguće naći polinom (po principu maksimuma modula analitičke funkcije) čija je vrijednost u ishodištu jednaka 1, a u svakoj točki na nekoj zatvorenoj krivulji oko ishodišta, manja od 1.

Konvergencija GMRES metode (nastavak)

- Također polinom niskog stupnja ne može biti 1 u ishodištu, i malen po apsolutnoj vrijednosti u mnogim točkama distribuiranim svuda oko ishodišta.
- Općenito se, međutim, ponašanje GMRES metode ne može odrediti samo iz svojstava svojstvenih vrijednosti.
- Zapravo, bilo koja nerastuća krivulja može predstavljati graf normi reziduala po broju iteracija za GMRES metodu primijenjenu na nekom ne-normalnom problemu, štoviše, matrica problema može imati bilo koje svojstvene vrijednosti.

Konvergencija GMRES metode (nastavak)

Teorem. Neka je dan nerastući niz $f(0) \geq f(1) \geq \cdots \geq f(n-1) > 0$ pozitivnih brojeva i skup kompleksnih brojeva različitih od nule $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$, tada postoji matrica A sa svojstvenim vrijednostima $\lambda_1, \lambda_2, \ldots, \lambda_n$ i desna strana sustava b sa $\|b\|_2 = f(0)$ takvi da reziduali r_k u svakom koraku GMRES algoritma primijenjenog na sustav $Ax = b$ sa $x_0 = 0$, zadovoljavaju $\|r_k\|_2 = f(k)$, $k = 1, 2, \ldots, n-1$.

Dokaz. Pogledajte primjere! ■

Primjeri konvergencije metoda

Primjer konvergencija CG metode

- U ovom primjeru ćemo pokazati konvergenciju metode konjugiranih gradijenata primijenjenu na dva sustava.
- Matrice oba sustava A_1 i A_2 su simetrične pozitivno definitne 100×100 matrice, sa svojstvenim vrijednostima $\lambda(A_1) \in \{1, 4, 9, \ldots, 10000\}$ i $\lambda(A_2) \in \{1, 100, 200, 300, 400, 500, 600, 700, 900, 10000\}$, pri čemu 1 i 10000 imaju kratnost 5, a ostale imaju kratnost 10.
- Dakle, A_1 ima 100 prilično različitih svojstvenih vrijednosti, dok ih A_2 ima 11.
- Matrice su
  - dobivene kao produkt $A_i = Q_i \Lambda_i Q_i^T$, $i = 1, 2$,
  - pri čemu su Λ_i dijagonalne matrice svojstvenih vrijednosti,

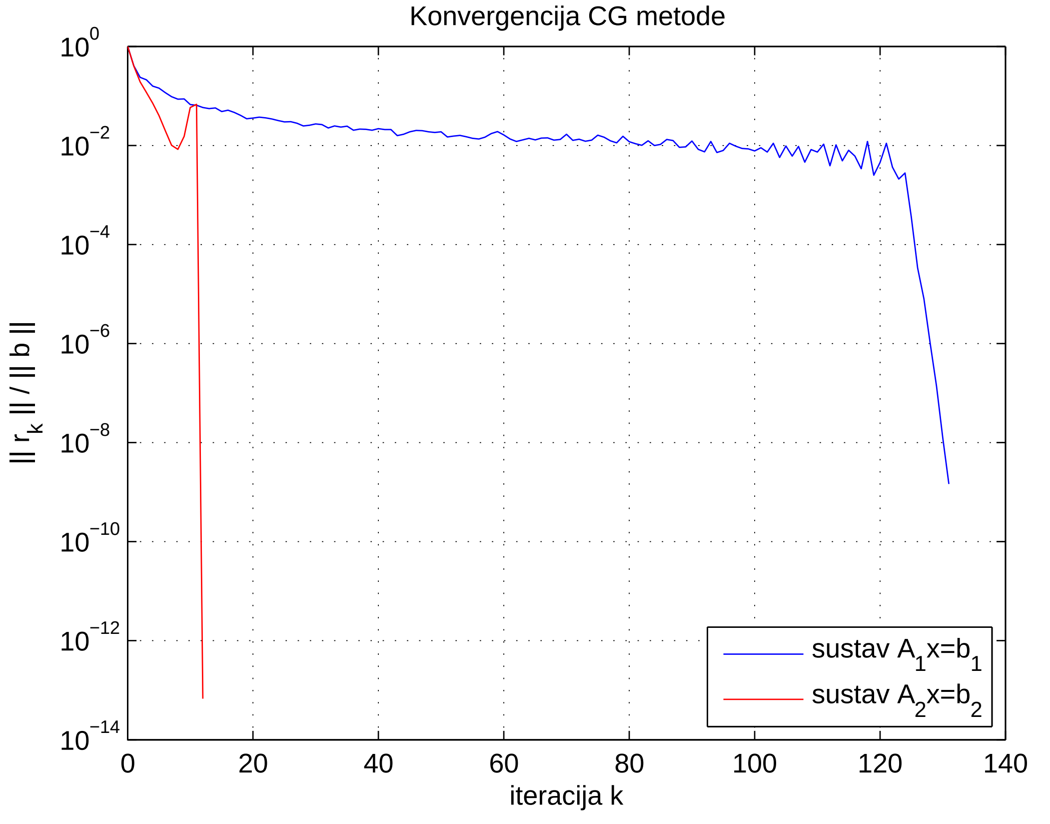
Primjer konvergencija CG metode (nastavak)

  - a Q_i slučajne ortogonalne matrice.
- Uvjetovanost obaju matrica jednaka je $\kappa(A_i) = 10^4$.
- Desne strane sustava, b_i određene su tako da rješenje oba sustava bude jednako $x = [1\ 1\ \ldots\ 1]^T$, odnosno da je $b_i = A_i \cdot x$.
- Za početnu iteraciju u oba slučaja uzet ćemo $x_0 = [0\ 0\ \ldots\ 0]^T$.
- Iteracije se trebaju zaustaviti kada je relativna norma reziduala $\|r_k\|_2/\|b\|_2$ manja od $tol = 10^{-8}$.

Pogledajte MATLAB program.

Primjer konvergencija CG metode (nastavak)

Konvergencija metode konjugiranih gradijenata za sustave $A_1 x = b_1$ i $A_2 x = b_2$:



Primjer konvergencija CG metode (nastavak)

- Prema teoriji prvi sustav treba konvergirati u 100 koraka, a drugi u 11. Odstupanja od toga dolaze zbog grešaka zaokruživanja.
- U praksi se iteracije uvijek zaustavljaju kada broj iteracija dostigne dimenziju matrice n.

Primjer konvergencija GMRES metode

Ovaj primjer ilustrira zadnji Teorem o neovisnosti konvergencije GMRES metode o svojstvenim vrijedostima ne-normalne matrice.

- Norme reziduala dobivene u nizu koraka GMRES metode su nerastući niz, budući da se reziduali minimiziraju po skupu ekspandirajućih podprostora.
- Postavlja se sljedeće pitanje: ako je dan nerastući niz pozitivnih realnih brojeva $f(0) \geq f(1) \geq \cdots \geq f(n-1) > 0$, i skup kompleksnih brojeva $\{\lambda_1 \ldots, \lambda_n\} \subset \mathbb{C}$, različitih od nule, da li tada postoje
  - $n \times n$ matrica A sa svojstvenim vrijednostima $\lambda_1, \ldots, \lambda_n$ i
  - početni vektor r_0 sa $\|r_0\|_2 = f(0)$,

Primjer konvergencija GMRES metode (nast.)

takvi da kod primjene GMRES algoritma na linearni sustav $Ax = b$, sa početnim rezidualom r_0, dobijemo aproksimacije x_k takve da je $\|r_k\|_2 = f(k)$, $k = 1, \ldots, n-1$?

- Odgovor na ovo pitanje je potvrdan.
- No, prije samog odgovora primijetimo da pretpostavka $f(n-1) > 0$ znači da GMRES algoritam ne konvergira k egzaktnom rješenju sve do zadnjeg, n-tog, koraka, kada su dimenizije potprostora $\mathcal{K}_n(A, r_0)$ i $A\mathcal{K}_n(A, r_0)$ jednake n.

Primjer konvergencija GMRES metode (nast.)

Započet ćemo najprije sa jednostavnom analizom nekih svojstava traženog rješenja.

- Budući da su reziduali generirani primjenom GMRES algoritma na linearni sustav $Ax = b$, sa početnom aproksimacijom x_0, potpuno određeni matricom A i početnim rezidualom r_0, bez smanjenja općenitosti možemo pretpostaviti da je početna aproksimacija x_0 jednaka nuli, i da je vektor desne strane sustava b zapravo početni rezidual.
- Pretpostavimo da A i b predstavljaju nepoznatu matricu i desnu stranu sustava.
- Neka je $\mathcal{W} = \{w_1, \ldots, w_n\}$ ortonormirana baza Krylovljevog prostora reziduala $A\mathcal{K}_n(A, b)$, takva da je $\mathrm{span}\{w_1, \ldots, w_j\} = A\mathcal{K}_j(A, b)$, $j = 1, \ldots, n$, i neka je $W = [w_1 \ \ldots \ w_n]$.

Primjer konvergencija GMRES metode (nast.)

- Prema Teoremu o konvergenciji, minimizacijsko svojstvo reziduala možemo preformulirati u

$$\|r_k\|_2 = \min_{u \in A\mathcal{K}_k(A,b)} \|b - u\|_2,$$

odakle se vidi da je $r_k \perp A\mathcal{K}_k(A, b)$ za $k = 0, 1, \ldots, n-1$ (kao kod izvoda CG metode), pa je on oblika

$$r_k = \sum_{j=k+1}^{n} \langle b, w_j \rangle w_j,$$

r_n je nula, a $b = r_0$ se može raspisati kao

$$b = \sum_{j=1}^{n} \langle b, w_j \rangle w_j.$$

Primjer konvergencija GMRES metode (nast.)

- Iz prethodnoga se vidi da je
$$|\langle b, w_j\rangle| = \sqrt{\|r_{j-1}\|_2^2 - \|r_j\|_2^2}.$$
- Ako je zadan nerastući pozitivan niz $f(0) \geq f(1) \geq \cdots \geq f(n-1) > 0$, definirajmo $f(n) = 0$ i nove vrijednosti $g(k)$ sa
$$g(k) = \sqrt{(f(k-1))^2 - (f(k))^2}, \quad k = 1, \ldots, n.$$
- Uvjeti $\|b\|_2 = f(0)$, $\|r_k\|_2 = f(k)$, $k = 1, 2, \ldots, n-1$ bit će tada zadovoljeni ako su koordinate od b u bazi $\mathcal{W}$ određene sa
$$W^T b = [g(1) \ \ldots \ g(n)]^T.$$

Primjer konvergencija GMRES metode (nast.)

- Zaista, u tom slučaju je $r_k = \sum_{j=k+1}^{n} \langle b, w_j \rangle w_j = \sum_{j=k+1}^{n} g(j) w_j$, pa slijedi

$$\|r_k\|_2^2 = \sum_{j=k+1}^{n} |g(j)|^2 = \sum_{j=k+1}^{n} [(f(j-1))^2 - (f(j))^2] = (f(k))^2.$$

- Neka je $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$, $\lambda_j \neq 0$, $j = 1, 2, \ldots, n$ skup točaka različitih od nule u komplesnoj ravnini.
- Promotrimo normirani polinom

$$a(z) = z^n - \sum_{j=0}^{n-1} \alpha_j z^j = (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_n).$$

- Zbog toga što su svi λ_j različiti od nule, slijedi $\alpha_0 \neq 0$.

Primjer konvergencija GMRES metode (nast.)

- Matricu A možemo smatrati linearnim operatorom na n-dimenzionalnim Hilbertovim prostorom $\mathbb{R}^n$.
- Taj operator označit ćemo sa $\mathcal{A}$, a njegova reprezantacija u standardnoj bazi $\mathcal{E} = \{\xi_1, \ldots, \xi_n\}$, pri čemu je ξ_i i-ti jedinični vektor, daje traženu matricu A:
$$\mathcal{A}^{\mathcal{E}} = A.$$
- $\mathcal{A}$ je jedinstveno određen djelovanjem na bilo koji skup vektora baze.
- Neka je $\mathcal{V} = \{v_1, \ldots, v_n\}$ bilo koja ortonormirana baza u $\mathbb{R}^n$, i neka je $V = [v_1 \ \ldots \ v_n]$.
- Neka b zadovoljava
$$V^T b = [g(1) \ \ldots \ g(n)]^T$$

Primjer konvergencija GMRES metode (nast.)

pri čemu

- ukoliko je dan b sa $\|b\|_2 = f(0)$, V može biti izabran,
- ili ako je V zadan, b može biti izabran.

- Budući da je $g(n) = f(n-1)$ različit od nule, skup vektora $\mathcal{B} = \{b, v_1, \ldots v_{n-1}\}$ je linearno nezavisan i također tvori bazu za $\mathbb{R}^n$.
- Neka je $B = [b \ v_1 \ \ldots \ v_{n-1}]$.
- Tada je operator $\mathcal{A}$ jednostavno određen jednadžbama

$$
\begin{aligned}
\mathcal{A}b &= v_1, \\
\mathcal{A}v_1 &= v_2, \\
&\vdots
\end{aligned}
$$

Primjer konvergencija GMRES metode (nast.)

$$\mathcal{A}v_{n-2} = v_{n-1},$$

$$\mathcal{A}v_{n-1} = \alpha_0 b + \alpha_1 v_1 + \cdots + \alpha_{n-1} v_{n-1}.$$

iz svojstva $A\mathcal{K}_n(A, r_0) = \mathcal{K}_n(A, r_0)$.

- U matričnoj reprezentaciji u bazi $\mathcal{B}$ matrica

$$\mathcal{A}^{\mathcal{B}} = \begin{bmatrix} 0 & \dots & 0 & \alpha_0 \\ 1 & \dots & 0 & \alpha_1 \\ & \ddots & \vdots & \vdots \\ & & 1 & \alpha_{n-1} \end{bmatrix}$$

tada ima svojstvene vrijednosti koje odgovaraju skupu Λ. (Izračunajte karakteristični polinom!)

Primjer konvergencija GMRES metode (nast.)

- Radi se o matrici koju još nazivamo pratilac polinoma.
- Na kraju, matrica A je dana sa

$$A = \mathcal{A}^{\mathcal{E}} = B\mathcal{A}^{\mathcal{B}}B^{-1}.$$

- Za naš konkretan primjer uzet ćemo:
  - skup svojstvenih vrijednosti $\lambda(A) \in \{e^{\frac{2\pi i}{100}k} : k = 0, 1, \ldots, 99\}$,
  - niz vrijednosti $f(0) = 100$, $f(1) = 99$, $f(2) = 98, \ldots, f(99) = 1$, $f(100) = 0$.
- Izračunajmo koeficijente polinoma

$$a(z) = z^{100} - \sum_{i=0}^{99} \alpha_i z^i = (z - \lambda_1(A)) \cdots (z - \lambda_{100}(A)),$$

Primjer konvergencija GMRES metode (nast.)

- u našem slučaju je $a(z) = z^{100} - 1$, odnosno

$$\alpha_0 = 1, \quad \alpha_1 = \alpha_2 = \cdots = \alpha_{99} = 0,$$

jer se radi o stotim korijenima jedinice.

Pogledajte MATLAB program.

Primjer konvergencija CG metode (nastavak)

Konvergencija GMRES metode za sustav $Ax = b$: