

Ričardsonova metoda (dvosložna, NESTACIONARNA)

← prethodna
lekcija

Dokija se iz $x_{n+1} = x_n - \tau (Ax_n - b)$ varijacijom
parametra τ , tj: $x_{n+1} = x_n - \tau_{n+1} (Ax_n - b)$.

Metoda je nestacionarna jer $G(x_n)$ zavisi od n , tj.
u svakoj iteraciji se koristi drugačija vrednost za τ_n .

Zadatak je odrediti $\tau_1, \tau_2, \tau_3, \dots$.

Sledeće je slično, ali ne, isto kao u prethodnoj lekciji.

$$C_n = \tau_n \cdot I$$

$$\begin{aligned} x_n - x^* &= x_{n-1} - \tau_n (Ax_{n-1} - \overbrace{Ax^*}^b) - x^* \\ &= (I - \tau_n A) x_{n-1} + \tau_n Ax^* - x^* \\ &= (I - \tau_n A) x_{n-1} + (\tau_n A - I) x^* \\ &= (I - \tau_n A) (x_{n-1} - x^*) \\ &\quad \underbrace{(I - \tau_{n-1} A)}_{\text{razlika je što ovi članovi proizvođa sada nisu isti}} (x_{n-2} - x^*) \\ &= \dots \\ &= \prod_{j=1}^n (I - \tau_j A) (x_0 - x^*) \end{aligned}$$

razlika je što ovi članovi
proizvođa sada nisu
isti

$$\Rightarrow \|x_n - x^*\|_2 \leq \left\| \prod_{j=1}^n (I - \tau_j A) \right\|_2 \cdot \|x_0 - x^*\|_2$$

Uvodimo iste pretpostavke: A je Hermitseova i poz. definitna!

$\Rightarrow \prod_{j=1}^n (I - \tau_j A)$ će takođe biti Hermitseova čije su sop. vrednosti:

$$\lambda_i \left(\prod_{j=1}^n (I - \tau_j A) \right) = \prod_{j=1}^n (1 - \tau_j \lambda_i(A))$$

$$\left\| \prod_{j=1}^n (I - \tau_j A) \right\|_2 = \max_i \left| \lambda_i \left(\prod_{j=1}^n (I - \tau_j A) \right) \right|$$

$$= \max_i \left| \prod_{j=1}^n (1 - \tau_j \lambda_i(A)) \right|$$

$$\leq \max_{\lambda \in [m, M]} \left| \prod_{j=1}^n (1 - \tau_j \lambda) \right| \quad \begin{aligned} M &= \max \lambda_i(A) \\ m &= \min \lambda_i(A) \end{aligned}$$

$$M \geq m > 0$$

(preskočeni su neki koraci koji su isti kao u prethodnom)

⇒ ocena greške:

$$\|x_n - x^*\|_2 \leq \max_{\lambda \in [m, M]} \left| \prod_{j=1}^n (1 - \tau_j \lambda) \right| \cdot \|x_0 - x^*\|_2$$

Zadatak izbora iteracionih parametara kod nestacionarne je da pri fiksanom n izaberemo $\tau_1, \dots, \tau_n \in \mathbb{R}$ tako da posle n iteracija greška bude što manja. τ_j treba rešiti: $\min_{\tau_1, \dots, \tau_n} \max_{\lambda \in [m, M]} \left| \prod_{j=1}^n (1 - \tau_j \lambda) \right|$

Napomena: Mi ne znamo unapred koliko iteracija bi bilo potrebno da bi se postigla tražena tačnost, dok metoda zahteva da znamo unapred n . Komentar na tu temu će biti na kraju, za sada pretpostavimo da je n fiksirano.

$$\min_{\tau_1, \dots, \tau_n} \max_{\lambda \in [m, M]} \left| \prod_{j=1}^n (1 - \tau_j \lambda) \right|$$

ovo je polinom n -tog stepena po λ
 $Q(\lambda)$

treba da odredimo $Q(\lambda)$ takvo da mu maksimum

na intervalu $[m, M]$ bude minimalni

(da mu odstupanje od nule bude minimalno)

("naprijubijem" uz x -osu)

Da bi ovaj problem rešili, prisjetimo se Čebiševljevih polinoma.

Čebiševjevi polinomi

(detaljno rešeno na NA12)

$$T_n(x) = \cos(n \cdot \arccos x), \quad n=1, 2, \dots, \quad -1 \leq x \leq 1$$

$$\rho(x) = \frac{1}{\sqrt{1-x^2}}, \quad (T_n, T_m) = \int_{-1}^1 \frac{T_n(x) \cdot T_m(x)}{\sqrt{1-x^2}} dx = \begin{cases} 0, & m \neq n \\ \neq 0, & m = n \end{cases}$$

II zapis: $T_{n+1}(x) = 2T_n(x) \cdot x - T_{n-1}(x), \quad T_0(x) = 1, T_1(x) = x$

III zapis: $T_n(x) = \frac{1}{2} \left[(x + \sqrt{x^2 - 1})^n + (x - \sqrt{x^2 - 1})^n \right]$

Teorema. Normiran Čebiševjev polinom $\overline{T}_n(x) = 2^{-n} T_n(x) = x^n + \dots$ najmanje odstopa od nule na $[-1, 1]$ medu svimi polinomima n -tog stepena sa koeficijentom 1 uz x^n .

Dokaz na NA12

$$\max_{T_1, \dots, T_n} \min_{x \in [m, M]} \underbrace{\left| \prod_{j=1}^n (1 - \zeta_j x) \right|}_{Q(x)}$$

$Q(x)$ ćemo izraziti preko Čebiševjevih polinoma T_n i linearne transformacije koja $[m, M]$ preslikava na $[-1, 1]$.

$$x \in [-1, 1], \quad \lambda \in [m, M]$$

Linearna gredna $ax + b = \lambda$:

$$\begin{aligned} -a + b &= m \\ a + b &= M \end{aligned} \quad \oplus$$

$$b = \frac{M+m}{2}, \quad a = \frac{M-m}{2}$$

$$\Rightarrow \lambda = \frac{M-m}{2} x + \frac{M+m}{2}$$

$$x = \frac{2\lambda - (M+m)}{M-m}$$

$$\Rightarrow T_n(x) = T_n\left(\frac{2\lambda - (M+m)}{M-m}\right) \equiv \overline{T}^{[m, M]}(\lambda)$$

$$\Rightarrow c = \overline{T}(\lambda), \quad c = \text{const}$$

Važi $Q(0) = \left| \prod_{j=1}^n (1 - \zeta_j \cdot 0) \right| = 1$. Iz ovog uslova odredimo c :

$$1 = Q(0) = c \cdot T_n \left(\frac{0 - (M+m)}{M-m} \right) \Rightarrow c = \frac{1}{T_n \left(-\frac{M+m}{M-m} \right)}$$

Takođe, $Q(\lambda)$ i $\overline{T}_n^{[m, M]}$ imaju iste nule:

$$Q(\lambda) = 0$$

$$c \cdot \overline{T}_n^{[m, M]}(\lambda) = 0$$

$$\Rightarrow \cos \left(n \cdot \arccos \left(\frac{2\lambda - (M+m)}{M-m} \right) \right) = 0$$

$$\Rightarrow \arccos \left(\frac{2\lambda - (M+m)}{M-m} \right) = -\frac{\pi}{2} + k\pi = \frac{(2k-1)\pi}{2}, \quad k=1, 2, \dots, n$$

$$\Rightarrow \arccos \left(\frac{2\lambda - (M+m)}{M-m} \right) = \frac{(2k-1)\pi}{2n}$$

$$\Rightarrow \frac{2\lambda - (M+m)}{M-m} = \cos \frac{(2k-1)\pi}{2n}$$

$$\Rightarrow 2\lambda = \left(\cos \frac{(2k-1)\pi}{2n} \right) \cdot (M-m) + (M+m)$$

$$\Rightarrow \lambda = \lambda_k = \frac{1}{2} \left[\cos \frac{(2k-1)\pi}{2n} (M-m) + (M+m) \right], \quad k=1, 2, \dots, n$$

$$Q(\lambda) = \prod_{j=1}^n (1 - \tau_j \lambda) = 0 \Rightarrow 1 - \tau_j \lambda = 0 \Rightarrow \tau_j = \frac{1}{\lambda}$$

$$\Rightarrow \tau_j = \frac{1}{\lambda_j} = \frac{2}{M+m + (M-m) \cos \frac{(2j-1)\pi}{2n}} \quad j=1, 2, \dots, n \quad (*)$$

Sad kada znamo $Q(\lambda)$, vratimo se oceni greške

$$\|x_n - x^*\|_2 \leq \max_{\lambda \in [m, M]} \underbrace{\left| \prod_{j=1}^n (1 - \tau_j \lambda) \right|}_{Q(\lambda)} \cdot \|x_0 - x^*\|_2$$

$$\max_{\lambda} |Q(\lambda)| = \max_{\lambda} \left| c \cdot \overline{T}_n^{[m, M]}(\lambda) \right|$$

$$= \frac{\max \left| T_n \left(\frac{2\lambda - (M+m)}{M-m} \right) \right|}{\left| T_n \left(-\frac{M+m}{M-m} \right) \right|}$$

ovo je c (izračunato na vrhu strane)

jer je cos parna

$$T_n \left(-\frac{M+m}{M-m} \right) \downarrow = T_n \left(\frac{M+m}{M-m} \right) = \frac{1}{2} \left[\left(\left(\frac{M+m}{M-m} \right) + \sqrt{\left(\frac{M+m}{M-m} \right)^2 - 1} \right)^n + \right. \\ \left. \overline{\text{ili zapis}} + \left(\frac{M+m}{M-m} \right) - \sqrt{\left(\frac{M+m}{M-m} \right)^2 - 1} \right)^n \right]$$

$$\begin{aligned}
\frac{M+m}{M-m} - \sqrt{\left(\frac{M+m}{M-m}\right)^2 - 1} &= \frac{M+m}{M-m} - \sqrt{\frac{(M+m)^2 - (M-m)^2}{(M-m)^2}} \\
&= \frac{M+m}{M-m} - \sqrt{\frac{(M+m+M-m)(M+m-M+m)}{(M-m)^2}} \\
&= \frac{M+m}{M-m} - \sqrt{\frac{2M \cdot 2m}{(M-m)^2}} \\
&= \frac{M+m}{M-m} - \frac{2\sqrt{Mm}}{M-m} \quad \leftarrow \text{for } M \geq m \\
&= \frac{(\sqrt{M} - \sqrt{m})^2}{M-m} \\
&= \frac{(\sqrt{M} - \sqrt{m})^2}{(\sqrt{M} - \sqrt{m})(\sqrt{M} + \sqrt{m})} \\
&= \frac{\sqrt{M} - \sqrt{m}}{\sqrt{M} + \sqrt{m}} \equiv \rho
\end{aligned}$$

$$\text{Analogous: } \frac{M+m}{M-m} + \sqrt{\left(\frac{M+m}{M-m}\right)^2 - 1} = \dots = \frac{\sqrt{M} + \sqrt{m}}{\sqrt{M} - \sqrt{m}} = \frac{1}{\rho}$$

$$\Rightarrow T_n\left(\frac{M+m}{M-m}\right) = \frac{1}{2} \left(\rho^n + \frac{1}{\rho^n} \right)$$

$$\begin{aligned}
\Rightarrow \max_{\lambda} \left| \prod_{j=1}^n (1 - \tau_j^* \lambda) \right| &= \frac{1}{|T_n\left(\frac{M+m}{M-m}\right)|} = \frac{1}{\frac{1}{2} \left(\rho^n + \frac{1}{\rho^n} \right)} = \frac{2}{\rho^n + \frac{1}{\rho^n}} \\
&= \frac{2\rho^n}{\rho^{2n} + 1} < 2\rho^n
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \|x^* - x_n\|_2 &\leq 2\rho^n \cdot \|x^* - x_0\|_2 \\
&= 2 \left(\frac{\sqrt{M} - \sqrt{m}}{\sqrt{M} + \sqrt{m}} \right)^n \cdot \|x^* - x_0\|_2
\end{aligned}$$

Uporedimo brzinu konvergencije dvoslojne stacionarne metode $x_{n+1} = x_n - \tau (Ax_n - b)$, dvoslojne nestacionarne (Richardsonove) metode $x_{n+1} = x_n - \tilde{\tau}_{n+1} (Ax_n - b)$:

$$0 < m < M \Rightarrow 0 < \frac{m}{M} < \sqrt{\frac{m}{M}} < 1$$

$$0 < \frac{\sqrt{M} - \sqrt{m}}{\sqrt{M} + \sqrt{m}} = \frac{1 - \sqrt{\frac{m}{M}}}{1 + \sqrt{\frac{m}{M}}} < \frac{1 - \frac{m}{M}}{1 + \frac{m}{M}} = \frac{M - m}{M + m} < 1$$

$\Rightarrow$ Nestacionarna konvergira brže od stacionarne.

Važna napomena:

Richardsonova metoda je nestabilna!

Mi smo odredili $\tilde{\tau}_j, j=1, \dots, n$ u (*) ali ti indeksi koje smo im dodelili nisu nužno redosled korišćenja $\tilde{\tau}_j$ u n iteracija!

Može se pokazati da ako se $\tilde{\tau}_j, j=1, \dots, n$ uzimaju redom po indeksima, polovina njih uvećava grešku, a polovina smanjuje.

Ako se uzme jedan sa početka, jedan sa kraja, npr. $\tilde{\tau}_1, \tilde{\tau}_n, \tilde{\tau}_2, \tilde{\tau}_{n-1}, \dots$, greška je pod kontrolom.

MATLAB: Ispitajte upotrebu $\tilde{\tau}_j, j=1, \dots, n$ na ova 2 različita načina (redosleda). Nakon svake iteracije štampajte grešku i u grafici predstavite $\|Ax_k - b\|$ i uporedite.

Napomena za n:

Rekli smo da je sve ovo za fiksirano n, ali mi ne znamo unapred koliko će ono biti.

Npr ako izaberemo $n=10$, x_{10} ne zadovoljava tačnost to ne znači da treba povećati n, raditi sve ispočetka!

U praksi se onda već dobijeni $\tilde{\tau}_1, \dots, \tilde{\tau}_{10}$ ciklično vrte do zadovoljenja tačnosti.