4. Simetrični problem lastnih vrednosti Rayleighova iteracija in bisekcija

Bor Plestenjak

NLA

16. marec 2010

Rayleighova iteracija

Če inverzno iteracijo kombiniramo z Rayleighovim kvocientom, dobimo:

Rayleighova iteracija

izberi $z_0 \neq 0$
$k = 0, 1, \ldots$
  $\sigma_k = \rho(z_k, A)$
  reši $(A - \sigma_k I)y_{k+1} = z_k$
  $z_{k+1} = \frac{y_{k+1}}{\|y_{k+1}\|_2}$

Namesto fiksnega premika σ pri inverzni iteraciji uporabljamo Rayleighov kvocient, ki je najboljši približek za lastno vrednost danega vektorja.

Konvergenca Rayleighove iteracije

Lema

Naj bo normiran vektor z_0 približek za lastni vektor x_1. Če za simetrično matriko A z lastnimi vrednostmi $|\lambda_n| \leq \cdots \leq |\lambda_2| < |\lambda_1|$ izvedemo en korak potenčne metode z začetnim vektorjem z_0, potem velja

$$\|z_1 \pm x_1\|_2 \leq \frac{|\lambda_2|}{|\lambda_1|}\|z_0 - x_1\|_2,$$

kjer predznak $\pm$ izberemo tako, da je norma razlike manjša.

Konvergenca Rayleighove iteracije

Lema

Naj bo normiran vektor z_0 približek za lastni vektor x_1. Če za simetrično matriko A z lastnimi vrednostmi $|\lambda_n| \leq \cdots \leq |\lambda_2| < |\lambda_1|$ izvedemo en korak potenčne metode z začetnim vektorjem z_0, potem velja

$$\|z_1 \pm x_1\|_2 \leq \frac{|\lambda_2|}{|\lambda_1|} \|z_0 - x_1\|_2,$$

kjer predznak $\pm$ izberemo tako, da je norma razlike manjša.

Posledica

Naj bo A simetrična matrika in $|\lambda_i - \sigma| \ll |\lambda_j - \sigma|$ za $j \neq i$, $j = 1, \ldots, n$. Če izvedemo en korak inverzne iteracije z začetnim vektorjem z_0, potem velja

$$\|z_1 - x_i\|_2 = \mathcal{O}(|\lambda_i - \sigma| \cdot \|z_0 - x_1\|_2).$$

Konvergenca Rayleighove iteracije 2

Lema

Naj bo A simetrična matrika in naj bo normiran vektor z približek za lastni vektor x_k. Potem velja

$$|\lambda_k - \rho(z, A)| \leq 2\|A\|_2 \|z - x_k\|_2^2.$$

Konvergenca Rayleighove iteracije 2

Lema

Naj bo A simetrična matrika in naj bo normiran vektor z približek za lastni vektor x_k. Potem velja

$$|\lambda_k - \rho(z, A)| \leq 2\|A\|_2 \|z - x_k\|_2^2.$$

Izrek

Naj bo A simetrična matrika. Potem ima Rayleighova iteracija v bližini enostavne lastne vrednosti kubični red konvergence.

Zgled za kubično konvergenco Rayleighove iteracije

Naj bo $A = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$, $\lambda_1 > \lambda_2$ in $z_r = \begin{bmatrix} c_r \\ s_r \end{bmatrix}$, $\|z_r\|_2^2 = c_r^2 + s_r^2 = 1$. Pri enem koraku Rayleighove iteracije dobimo

$$\sigma_r = z_r^T A z_r = c_r^2 \lambda_1 + s_r^2 \lambda_2.$$

Iz sistema

$$\begin{bmatrix} \lambda_1 - c_r^2\lambda_1 - s_r^2\lambda_2 & 0 \\ 0 & \lambda_2 - c_r^2\lambda_1 - s_r^2\lambda_2 \end{bmatrix} y_{r+1} = z_r$$

dobimo

$$y_{r+1} = \frac{1}{(\lambda_1 - \lambda_2)c_r^2 s_r^2} \begin{bmatrix} c_r^3 \\ -s_r^3 \end{bmatrix},$$

od tod pa

$$z_{r+1} = \frac{1}{\sqrt{c_r^6 + s_r^6}} \begin{bmatrix} c_r^3 \\ -s_r^3 \end{bmatrix}.$$

Od tod v primeru $s_r \neq c_r$ sledi kubična konvergenca proti e_1 oziroma e_2.

Kubična konvergenca QR iteracije za simetrično matriko

Naj bo $T_k = \begin{bmatrix} a_1^{(k)} & b_1^{(k)} & & & \\ b_1^{(k)} & a_2^{(k)} & b_2^{(k)} & & \\ & \ddots & \ddots & \ddots & \\ & & b_{n-2}^{(k)} & a_{n-1}^{(k)} & b_{n-1}^{(k)} \\ & & & b_{n-1}^{(k)} & a_n^{(k)} \end{bmatrix}.$

QR iteracija z Rayl. premikom

$T_0 = T$
$k = 0, 1, \dots$
$\quad \sigma_k = a_n^{(k)}$
$\quad T_k - \sigma_k I = Q_k R_k$
$\quad T_{k+1} = R_k Q_k + \sigma_k I$

Rayleighova iteracija

izberi $z_0 \neq 0$
$k = 0, 1, \dots$
$\quad \sigma_k = \rho(z_k, T)$
$\quad$ reši $(T - \sigma_k I) y_{k+1} = z_k$
$\quad z_{k+1} = \frac{y_{k+1}}{\|y_{k+1}\|_2}$

Lema

Če je T simetrična in delamo QR iteracijo z Rayleighovimi premiki, potem so premiki σ_k enaki Rayleighovim kvocientom, ki jih dobimo pri Rayleighovi iteraciji, če za začetni vektor vzamemo $z_0 = e_n$.

Sturmovo zaporedje

Naj bo

$$T = \begin{bmatrix} a_1 & b_1 & & & \\ b_1 & a_2 & b_2 & & \\ & \ddots & \ddots & \ddots & \\ & & b_{n-2} & a_{n-1} & b_{n-1} \\ & & & b_{n-1} & a_n \end{bmatrix}$$

ireducibilna tridiagonalna simetrična matrika.

Naj bo T_r njena vodilna $r \times r$ podmatrika in $f_r(\lambda) = \det(T_r - \lambda I)$.

Izpeljemo lahko rekurzivno formulo

$$f_{r+1}(\lambda) = (a_{r+1} - \lambda) f_r(\lambda) - b_r^2 f_{r-1}(\lambda)$$

za $r = 0, \ldots, n-1$, ki se začne z $f_0(\lambda) \equiv 1$ in $f_1(\lambda) = a_1 - \lambda$.

Porabimo $4n + \mathcal{O}(1)$ osnovnih operacij (če shranimo b_i^2).

Sturmovo zaporedje

$$\begin{aligned} f_0(\lambda) &\equiv 1 \\ f_1(\lambda) &= a_1 - \lambda \\ f_{r+1}(\lambda) &= (a_{r+1} - \lambda) f_r(\lambda) - b_r^2 f_{r-1}(\lambda) \quad \text{za } r = 0, \ldots, n-1 \end{aligned}$$

Izrek

Polinomi $f_0, \ldots, f_n$ tvorijo Sturmovo zaporedje, kar pomeni, da zadoščajo naslednjim trem točkam:

1) $f_0(\lambda) \neq 0$ *za vsak* λ.
2) *Če je* $f_r(\lambda_0) = 0$ *za* $r < n$, *potem je* $f_{r-1}(\lambda_0) f_{r+1}(\lambda_0) < 0$.
3) *Če je* $f_n(\lambda_0) = 0$, *potem je* $f_{n-1}(\lambda_0) f_n'(\lambda_0) < 0$.

Posledica

Ireducibilna tridiagonalna simetrična matrika ima enostavne lastne vrednosti.

Štetje lastnih vrednosti

Pri fiksnem λ_0 označimo z $u(\lambda_0)$ število ujemanj predznaka v zaporedju $f_0(\lambda_0), \ldots, f_n(\lambda_0)$. Pri tem vsako notranjo ničlo štejemo za eno ujemanje, ničlo na koncu pa ne. Primeri:

$$\begin{aligned} u(++--+) &= 2, \\ u(++0-+) &= 2, \\ u(++0-+0) &= 2. \end{aligned}$$

Lema

Število $u(\lambda_0)$ je enako številu lastnih vrednosti matrike T, ki so strogo večje od λ_0.

Sedaj lahko npr. z bisekcijo poiščemo k-to lastno vrednost. Iščemo točko λ_k, za katero velja $u(\lambda_k - \epsilon) = k$ in $u(\lambda_k + \epsilon) = k - 1$ za dovolj majhen $\epsilon > 0$.

LDL^T razcep in Sylvestrov izrek

Denimo, da za matriko

$$T = \begin{bmatrix} a_1 & b_1 & & \\ b_1 & \ddots & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ & & b_{n-1} & a_n \end{bmatrix}$$

izračunamo razcep $T - \lambda I = LDL^T$, kjer je

$$L = \begin{bmatrix} 1 & & & \\ l_1 & 1 & & \\ & \ddots & \ddots & \\ & & l_{n-1} & 1 \end{bmatrix}, \quad D = \begin{bmatrix} d_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & d_n \end{bmatrix}.$$

$$d_1 = a_1 - \lambda, \quad d_1 l_1 = b_1$$
$$l_{k-1}^2 d_{k-1} + d_k = a_k - z, \quad d_k l_k = b_k \quad \text{za } k = 2, \ldots, n-1.$$

Po Sylvestrovem izreku je število pozitivnih diagonalnih elementov D enako $u(\lambda)$.

Izračun LDL^T razcepa

$d_1 = a_1 - \lambda, \quad d_1 l_1 = b_1$
$l_{k-1}^2 d_{k-1} + d_k = a_k - z, \quad d_k l_k = b_k \quad$ za $k = 2, \ldots, n-1.$

Izračun $d_1, \ldots, d_n$

$$d_1 = a_1 - \lambda$$
$$k = 2, \ldots, n$$
$$\quad d_k = a_k - \lambda - \frac{b_{k-1}^2}{d_{k-1}}$$

Porabimo $3n + \mathcal{O}(1)$ osnovnih operacij (če shranimo b_i^2).

Lema

Če ne pride do prekoračitve ali podkoračitve, se predznaki numerično izračunanih diagonalnih elementov $\widehat{d}_1, \ldots, \widehat{d}_n$ ujemajo s predznaki točnih $d_1, \ldots, d_n$ iz LDL^T razcepa matrike $T + \delta T$, kjer je $\delta a_k = 0$ in $|\delta b_k| \leq (5/2)|b_k|u$ za $k = 1, \ldots, n$.

Povezava s Sturmovim zaporedjem

Sturmovo zaporedje:

$$
\begin{aligned}
f_0(\lambda) &\equiv 1 \\
f_1(\lambda) &= a_1 - \lambda \\
f_{k+1}(\lambda) &= (a_{k+1} - \lambda) f_k(\lambda) - b_k^2 f_{k-1}(\lambda) \quad \text{za } k = 0, \ldots, n-1
\end{aligned}
$$

LDL^T:

$$
\begin{aligned}
d_1 &= a_1 - \lambda \\
d_k &= a_k - \lambda - \frac{b_{k-1}^2}{d_{k-1}} \quad \text{za } k = 2, \ldots, n
\end{aligned}
$$

Velja $f_k(\lambda) = d_1 d_2 \ldots d_k$ oziroma $d_k = f_k(\lambda)/f_{k-1}(\lambda)$ za $k = 1, \ldots, n$.

Motivacija: Robni problem

Sturm-Liouvillov robni problem lastnih vrednosti

Na intervalu $[a, b]$ iščemo lastno vrednost λ in neničelno funkcijo y, ki zadošča

$$-(p(x)y')' + q(x)y = \lambda y,$$

in robnima pogojema $y(a) = y(b) = 0$, pri čemer za koeficienta p in q velja $p(x) > 0$ in $q(x) \geq 0$ za $x \in [a, b]$.

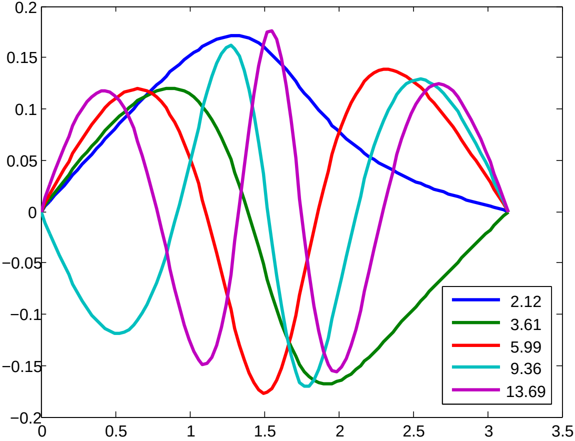
Pri diferenčni metodi interval $[a, b]$ razdelimo z ekvidistantnimi točkami $x0 = a$, $x_i = x_0 + ih$, kjer je $h = (b - a)/(n+1)$. Naj bo y_i približek za $y(x_i)$, ki ga iščemo. Enačbo v točki x_i, $i = 1, \ldots, n$, aproksimiramo z

$$\frac{p(x_i - h/2)(y_i - y_{i-1}) + p(x_i + h/2)(y_i - y_{i+1})}{h^2} + q(x_i)y_i = \lambda y_i.$$

Dobimo $Ay = \lambda y$, kjer je A tridiagonalna simetrična matrika.

Zgled robnega problema

$[a, b] = [0, \pi]$, $p(x) = 1 - \frac{4}{5}\sin^2 x$, $q(x) = x$ in robna pogoja $y(a) = y(b) = 0$.



Motivacija: sistem vzmeti in nihal

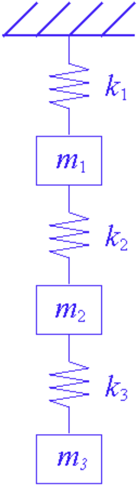
Denimo, da imamo sistem treh uteži in treh vzmeti. Po Newtonovih zakonih za odmike uteži y_i velja sistem diferencialnih enačb

$$My'' + Ky = 0,$$

kjer sta

$$M = \begin{bmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{bmatrix}, \quad K = \begin{bmatrix} k_1 + k_2 & -k_2 & 0 \\ -k_2 & k_2 + k_3 & -k_3 \\ 0 & -k_3 & k_3 \end{bmatrix},$$

M je masna matrika, K pa togostna matrika.



Motivacija: sistem vzmeti in nihal

Denimo, da imamo sistem treh uteži in treh vzmeti. Po Newtonovih zakonih za odmike uteži y_i velja sistem diferencialnih enačb

$$My'' + Ky = 0,$$

kjer sta

$$M = \begin{bmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{bmatrix}, \quad K = \begin{bmatrix} k_1 + k_2 & -k_2 & 0 \\ -k_2 & k_2 + k_3 & -k_3 \\ 0 & -k_3 & k_3 \end{bmatrix},$$

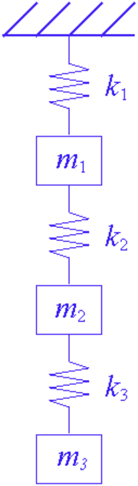
M je masna matrika, K pa togostna matrika. Sistem niha z lastno frekvenco ω, kar pomeni, da za komponente rešitve velja

$$y_k(t) = x_k e^{i\omega t},$$

kjer so x_1, x_2, x_3 amplitude. Lastno frekvenco in amplitude določimo iz

$$Kx = \omega^2 Mx,$$

saj mora veljati $y_k''(t) = -\omega^2 x_k e^{i\omega t}$.



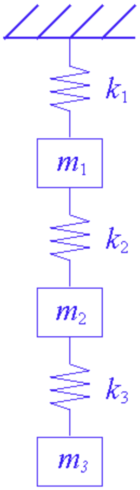
Motivacija: sistem vzmeti in nihal

Denimo, da imamo sistem treh uteži in treh vzmeti. Po Newtonovih zakonih za odmike uteži y_i velja sistem diferencialnih enačb

$$My'' + Ky = 0,$$

kjer sta

$$M = \begin{bmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{bmatrix}, \quad K = \begin{bmatrix} k_1 + k_2 & -k_2 & 0 \\ -k_2 & k_2 + k_3 & -k_3 \\ 0 & -k_3 & k_3 \end{bmatrix},$$



M je masna matrika, K pa togostna matrika. Sistem niha z lastno frekvenco ω, kar pomeni, da za komponente rešitve velja

$$y_k(t) = x_k e^{i\omega t},$$

kjer so x_1, x_2, x_3 amplitude. Lastno frekvenco in amplitude določimo iz

$$Kx = \omega^2 Mx,$$

saj mora veljati $y_k''(t) = -\omega^2 x_k e^{i\omega t}$. To prevedemo na lastni problem $Az = \lambda z$, kjer je $A = M^{-1/2} K M^{-1/2}$, $\lambda = \omega^2$ in $z = M^{1/2} x$.