7.1 Lanczosev algoritem za lastne vrednosti

Če je A simetrična, se simetričnost pri Arnoldijevem algoritmu prenese na zgornjo Hessenbergovo matriko H, ki postane simetrična in tridiagonalna.

$v_1 = r_0/\|r_0\|$, $\beta_0 = 0$, $v_0 = 0$
$j = 1, 2, \ldots, k$
  $z = Av_j - \beta_{j-1}v_{j-1}$
  $\alpha_j = v_j^T z$
  $z = z - \alpha_j v_j$
  $\beta_j = \|z\|$
  če je $\beta_j = 0$, prekini računanje
  $v_{j+1} = z/\beta_j$

Dobimo $AV_k = V_k T_k + \beta_k v_{k+1} e_k^T$, kjer je

$$T_k = \begin{bmatrix} \alpha_1 & \beta_1 & & \\ \beta_1 & \alpha_2 & \ddots & \\ & \ddots & \ddots & \beta_{k-1} \\ & & \beta_{k-1} & \alpha_k \end{bmatrix}.$$

Lastne vrednosti matrike T_k so Ritzeve vrednosti.

Ocene Lanczosevega algoritma

Ritzeve vrednosti pri k se prepletajo z Ritzevimi vrednosti pri $k+1$, saj se lastne vrednosti T_k prepletajo z lastnimi vrednostmi T_{k+1}. Če so $\theta_k^{(k)} < \cdots < \theta_1^{(k)}$ Ritzeve vrednosti, ki jih dobimo v k-tem koraku Lanczosevega algoritma, potem velja

$$\theta_{k+1}^{(k+1)} < \theta_k^{(k)} < \theta_k^{(k+1)} < \cdots < \theta_2^{(k+1)} < \theta_1^{(k)} < \theta_1^{(k+1)}.$$

Zaradi tega Ritzeve vrednosti monotono konvergirajo proti lastnim vrednostim matrike A, pri čemer najprej skonvergirajo zunanje lastne vrednosti.

Izrek 1. *Za Ritzeve vrednosti, ki jih dobimo v k-tem koraku Lanczosevega algoritma, velja:*

1. *Obstaja k lastnih vrednosti matrike A, da je $|\theta_i^{(k)} - \lambda_i| \le \beta_k$ za $i = 1, \ldots, k$.*
2. *Če je $z = V_k s$ Ritzev vektor za Ritzevo vrednost θ, potem velja*

$$\min_i |\lambda_i - \theta| \le \|Az - \theta z\|_2 = \beta_k |e_k^T s|.$$

3. $\theta_i^{(k)} = \max_{\dim(S)=i} \ \min_{0 \ne x \in S \subset \mathcal{K}_k(A, v_1)} \rho(A, x)$.

Povezava z ortogonalnimi polinomi

Vektorjem $v_1, \ldots, v_k$, ki jih dobimo z Lanczosevo metodo kot ON bazo $\mathcal{K}_k(A, v_1)$, pripadajo t.i. *Lanczosevi polinomi* $P_0, \ldots, P_{k-1}$, da je $v_i = P_{i-1}(A)v_1$.

Lanczosevi polinomi zadoščajo tričlenski rekurzivni formuli

$$\beta_i P_i(A) = (A - \alpha_i I)P_{i-1}(A) - \beta_{i-1}P_{i-2}(A).$$

Če definiramo skalarni produkt polinomov kot

$$[P, Q] = \langle P(A)v_1, Q(A)v_1 \rangle,$$

potem je $[P_i, P_j] = 0$ za $i \neq j$.

Lema 1. *Ničle polinoma P_k so lastne vrednosti matrike T_k oziroma Ritzeve vrednosti.*

Ocene o hitrosti konvergence

Naj bodo $\lambda_n \leq \cdots \leq \lambda_1$ lastne vrednosti A in $x_n, \ldots, x_1$ pripadajoči ON lastni vektorji.

Lema 2. *Če je P_i spektralni projektor, ki pripada λ_i, potem v primeru $P_i v_1 \neq 0$ velja*

$$\tan\varphi(x_i, \mathcal{K}_m(A, v_1)) = \min_{p \in \mathbb{P}_{m-1},\ p(\lambda_i)=1} \|p(A)y_i\| \tan\varphi(x_i, v_1),$$

kjer je

$$y_i = \begin{cases} \frac{(I-P_i)v_1}{\|(I-P_1)v_1\|}, & \text{v primeru } (I-P_i)v_1 \neq 0 \\ 0, & \text{sicer} \end{cases}$$

Izrek 2. *Pri enakih predpostavkah kot zgoraj velja*

$$\tan\varphi(x_i, \mathcal{K}_m(A, v_1)) \leq \frac{\kappa_i}{T_{m-i}(1 + 2\gamma_i)} \tan\varphi(x_i, v_1),$$

kjer je

$$\kappa_1 = 1, \quad \kappa_i = \prod_{j=1}^{i-1} \frac{\lambda_j - \lambda_n}{\lambda_j - \lambda_i} \text{ za } i > 1, \quad \text{in} \quad \gamma_i = \frac{\lambda_i - \lambda_{i+1}}{\lambda_{i+1} - \lambda_n}.$$

Konvergenca $\theta_i^{(k)}$

Naj bodo $\lambda_n \leq \cdots \leq \lambda_1$ lastne vrednosti A in $x_n, \ldots, x_1$ pripadajoči ON lastni vektorji.

Naj bodo $\theta_m^{(m)} \leq \cdots \leq \theta_1^{(m)}$ Ritzeve vrednosti A glede na $\mathcal{K}_m(A, v_1)$.

Lema 3. *Velja*

$$0 \leq \lambda_1 - \theta_1^{(m)} \leq (\lambda_1 - \lambda_n) \min_{p \in \mathbb{P}_{m-1},\ p(\lambda_i)=1} \ \max_{\lambda \in [\lambda_n, \lambda_2]} |p(\lambda)|^2 \tan \varphi(x_i, v_1)^2.$$

Izrek 3. *Velja*

$$0 \leq \lambda_i - \theta_i^{(m)} \leq (\lambda_1 - \lambda_n) \left(\frac{\kappa_i^{(m)} \tan \varphi(x_i, v_1))}{T_{m-i}(1 + 2\gamma_i)} \right)^2,$$

kjer je

$$\kappa_1^{(m)} = 1, \quad \kappa_i^{(m)} = \prod_{j=1}^{i-1} \frac{\theta_j^{(m)} - \lambda_n}{\theta_j^{(m)} - \lambda_i} \text{ za } i > 1, \quad \text{in } \gamma_i = \frac{\lambda_i - \lambda_{i+1}}{\lambda_{i+1} - \lambda_n}.$$

Praktične ocene

Naj bodo $\lambda_n \le \cdots \le \lambda_1$ lastne vrednosti A in $x_n, \ldots, x_1$ pripadajoči ON lastni vektorji.

Naj bodo $\theta_m^{(m)} \le \cdots \le \theta_1^{(m)}$ Ritzeve vrednosti A glede na $\mathcal{K}_m(A, v_1)$ in $u_m, \ldots, u_1$ pripadajoči Ritzevi vektorji, kjer je $u_i = V_m y_i$.

Označimo $r_j = Au_j - \theta_j^{(m)} u_j$.

Vemo že, da velja $\min_i |\lambda_i - \theta_j^{(m)}| \le \|r_j\|_2 = \beta_m |e_m^T y_j|$.

Če je $\theta_i^{(m)}$ približek za λ_i, potem označimo $\text{gap}(\theta_i^{(m)}) = \min\limits_{j \ne i} |\theta_i^{(m)} - \lambda_j|$.

Velja

$$|\lambda_i - \theta_i^{(m)}| \le \frac{\|r_i\|_2^2}{\text{gap}(\theta_i^{(m)})},$$

in

$$\sin \varphi(u_i, x_i) \le \frac{\|r_i\|_2^2}{\text{gap}(\theta_i^{(m)})}.$$

V praksi lahko $\text{gap}(\theta_i^{(m)})$ ocenimo z

$$\text{gap}(\theta_i^{(m)}) \approx \min_{j \ne i} (|\theta_i^{(m)} - \theta_j^{(m)}| - \|r_j\|).$$

Harmonične Ritzeve vrednosti

θ je harmonična Ritzeva vrednost in u harmonični Ritzev vektor, če je

$$Au - \theta u \perp A\mathcal{K}_m(A, v_1).$$

Harmonične Ritzeve vrednosti so inverzi Ritzevih vrednosti A^{-1} glede na $A\mathcal{K}_m(A, v_1)$.

Lema 4. *Po k korakih Lanczosa dobimo $AV_k = V_kT_k + \beta_k v_{k+1}e_k^T = V_{k+1}T_{k+1,k}$. Za harmonično Ritzevo vrednost θ velja, da je θ^{-1} lastna vrednost posplošenega problema lastnih vrednosti*

$$T_k y = \theta^{-1} T_{k+1,k}^T T_{k+1,k} y.$$

Če uporabimo premik σ, je θ harmonična Ritzeva vrednost in u harmonični vektor, če je

$$Au - \theta u \perp (A - \sigma I)\mathcal{K}_m(A, v_1).$$

θ je harmonična Ritzeva vrednost, če je $(\theta - \sigma)^{-1}$ Ritzeva vrednost $(A - \sigma I)^{-1}$ glede na $(A - \sigma I)\mathcal{K}_m(A, v_1)$

To je ekvivalentno $V_k^T(A - \sigma I)^2 V_k y = (\theta - \sigma)V_k^T(A - \sigma I)V_k y$ oziroma

$$(\theta - \sigma)^{-1}(T_{k+1,k}^T T_{k+1,k} - 2\sigma T_k + \sigma^2 I)y = (T_k - \sigma I)y.$$

Lehmannovi intervali

Naj bo σ izbrana točka, v okolici katere iščemo lastne vrednosti. Lastne vrednosti A, ki naj bodo vse enostavne, uredimo kot

$$\lambda_{-r} < \cdots < \lambda_{-1} < \sigma < \lambda_1 < \cdots < \lambda_{n-r}$$

Če je $\theta^{(k)}$ harmonična Ritzeva vrednost glede na $(A-\sigma I)\mathcal{K}_k(A,v_1)$, potem preko minimaks izreka za $(A-\sigma I)^{-1}$ sledi

$$\theta_{-1}^{(k)} < \lambda_{-1} \qquad \text{in} \qquad \lambda_1 < \theta_1^{(k)}.$$

Izrek 4. [Lehmann] *Naj bodo*

$$\theta_{-s}^{(k)} < \cdots < \theta_{-1}^{(k)} < \sigma < \theta_1(k) < \cdots < \theta_{k-s}^{(k)}$$

harmonične Ritzeve vrednosti matrike A, glede na $(A-\sigma I)\mathcal{K}_k(A,v_1)$.

Vsak interval $[\sigma, \theta_j^{(k)}]$, kjer je $j=1,\ldots,k-s$, vsebuje vsaj j lastnih vrednosti matrike A. Enako, vsak interval $[\theta_{-j}^{(k)}, \sigma]$, kjer je $j=1,\ldots,s$, vsebuje vsaj j lastnih vrednosti matrike A.

Ko k narašča, $\theta_{-j}^{(k)}$ monotono naraščajoče konvergira k λ_{-j} in podobno $\theta_j^{(k)}$ monotono padajoče konvergira k λ_j

Lanczos brez reortogonalizacije

Če izvajamo osnovno Lanczosevo metodo brez reortogonalizacije, pride do izgube ortogonalnosti in dobimo navidezne večkratne lastne vrednosti.

Opazimo, da se ortogonalnost baze za podprostor Krilova poslabša takrat, ko kak izmed Ritzevih vektorjev skonvergira do lastnega vektorja.

Izrek 5. [Paige] *Če izvajamo Lanczosev algoritem v aritmetiki z osnovno zaokrožitveno napako ϵ, potem v k-tem koraku za izračunanje Ritzeve vektorje $u_1, \ldots, u_k$ velja*

$$u_i^T v_{k+1} = \frac{\mathcal{O}(\epsilon \|A\|)}{\|r_i\|_2},$$

kjer je $r_i = Au_i - \theta_i u_i$, oziroma

$$u_i^T v_{k+1} = \frac{\mathcal{O}(\epsilon \|A\|)}{\beta_k |e_k^T y_i|},$$

kjer je $u_i = V_k y_i$.

Dokaz je v Demmelu, temelji pa na formuli

$$\widetilde{\beta}_j \widetilde{v}_{j+1} + f_j = (A - \widetilde{\alpha}_j I)\widetilde{v}_j - \widetilde{\beta}_{j-1}\widetilde{v}_{j-1},$$

kjer je $\|f_j\| = \mathcal{O}(\epsilon\|A\|)$ posledica zaokrožitvenih napak.

Lanczos s polno reortogonalizacijo

izberi začetni vektor v_1 z $\|v_1\|_2 = 1$, $\beta_0 = 0$, $v_0 = 0$
$j = 1, 2, \ldots, k$
$z = Av_j$
$\alpha_j = v_j^T z$
$z = z - \alpha_j v_j - \beta_{j-1} v_{j-1}$
$\beta_j = \|z\|_2$
če je $\beta_j = 0$, potem prekini računanje
$v_{j+1} = z/\beta_j$

izberi začetni vektor v_1 z $\|v_1\|_2 = 1$
$j = 1, 2, \ldots, k$
$z = Av_j$
$\alpha_j = v_j^T z_j$
$z = z - \sum_{k=1}^{j} (z^T v_j) v_j$
$z = z - \sum_{k=1}^{j} (z^T v_j) v_j$
$\beta_j = \|z\|_2$
če je $\beta_j = 0$, potem prekini računanje
$v_{j+1} = z/\beta_j$

Lanczos s selektivno reortogonalizacijo

izberi začetni vektor v_1 z $\|v_1\|_2 = 1$, $\beta_0 = 0$, $v_0 = 0$
$j = 1, 2, \ldots, k$
    $z = Av_j$
    $\alpha_j = v_j^T z$
    $z = z - \alpha_j v_j - \beta_{j-1} v_{j-1}$
    $\beta_j = \|z\|_2$
    če je $\beta_j = 0$, potem prekini računanje
    $v_{j+1} = z/\beta_j$

izberi začetni vektor v_1 z $\|v_1\|_2 = 1$, $\beta_0 = 0$, $v_0 = 0$
$j = 1, 2, \ldots, k$
    $z = Av_j$
    $\alpha_j = v_j^T z$
    $z = z - \alpha_j v_j - \beta_{j-1} v_{j-1}$
    za vse $i = 1, \ldots, k$, kjer je $\beta_k |e_k^T y_i| \leq \sqrt{\epsilon} \|T_k\|$
        $z = z - (u_{i,k}^T z) u_{i,k}$
    $\beta_j = \|z\|_2$
    če je $\beta_j = 0$, potem prekini računanje
    $v_{j+1} = z/\beta_j$