## Pomožni rezultati - Rayleighov kvocient

Naj bo $A$ simetrična matrika z lastnimi vrednostmi $\lambda_n \le \cdots \le \lambda_1$.

Ker so pri simetričnih matrikah lastni vektorji realni, Rayleighov kvocient računamo le za realne vektorje. Za $x \neq 0$ je

$$\rho(x, A) = \frac{x^T A x}{x^T x}$$

in veljajo naslednje lastnosti:

a) Za $\alpha \neq 0$ je $\rho(x, A) = \rho(\alpha x, A)$.

b) Za lastni vektor $x_i$ je $\rho(x_i, A) = \lambda_i$.

c) Če je $x$ približek za lastni vektor, je $\rho(x, A)$ najboljša aproksimacija za lastno vrednost v smislu, da je

$$\min_{\sigma \in \mathbb{R}} \|Ax - \sigma x\|_2$$

dosežen pri $\sigma = \rho(x, A)$.

d) Za vsak $x \neq 0$ velja

$$\lambda_n \le \rho(x, A) \le \lambda_1.$$

# Pomožni rezultati - Minimaks izrek

**Izrek 1. [Courant–Fischerjev minimaks izrek]** *Za simetrično matriko $A$ z lastnimi vrednostmi $\lambda_n \le \cdots \le \lambda_1$ velja za $i = 1, \ldots, n$:*

$$\lambda_i = \min_{\substack{S \subset \mathbb{R}^n \\ \dim(S)=n-i+1}} \max_{\substack{x\in S \\ x\neq 0}} \rho(x, A) = \max_{\substack{R \subset \mathbb{R}^n \\ \dim(R)=i}} \min_{\substack{x\in R \\ x\neq 0}} \rho(x, A).$$

pause

**Posledica 1.** *Če sta $A$ in $E$ simetrični matriki in so $\lambda_n \le \cdots \le \lambda_1$ lastne vrednosti $A$, $\widehat{\lambda}_n \le \cdots \le \widehat{\lambda}_1$ pa lastne vrednosti $A + E$, potem za $i = 1, \ldots, n$ velja*

$$\lambda_i + \lambda_n(E) \le \widehat{\lambda}_i \le \lambda_i + \lambda_1(E).$$

**Posledica 2. [Weylov izrek]** *Če sta $A$ in $E$ simetrični matriki in so $\lambda_n \le \cdots \le \lambda_1$ lastne vrednosti $A$, $\widehat{\lambda}_n \le \cdots \le \widehat{\lambda}_1$ pa lastne vrednosti $A + E$, potem velja*

$$|\lambda_i - \widehat{\lambda}_i| \le \|E\|_2 \qquad \text{za} \quad i = 1, \ldots, n.$$

**Izrek 2.** *Če je $A$ simetrična matrika, $\|x\|_2 = 1$ in $\beta$ približek za lastno vrednost, potem obstaja lastna vrednost $\lambda_i$ matrike $A$, da je $|\beta - \lambda_i| \le \|Ax - \beta x\|_2$.*

# Pomožni rezultati - Prepletanje lastnih vrednosti

**Izrek 3. [Cauchyjev izrek o prepletanju]** *Če je $A$ simetrična matrika in je $A_r$ njena vodilna podmatrika velikosti $r \times r$ za $r = 1, \ldots, n$, potem za $k = 1, \ldots, n-1$ velja*

$$\lambda_{k+1}(A_{k+1}) \leq \lambda_k(A_k) \leq \lambda_k(A_{k+1}) \leq \cdots \leq \lambda_2(A_{k+1}) \leq \lambda_1(A_k) \leq \lambda_1(A_{k+1}).$$

# Pomožni rezultati - Sturmovo zaporedje

Naj bo

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

nerazcepna tridiagonalna simetrična matrika (torej $b_i \neq 0$ za vsak $i$). Če s $T_r$ označimo njeno vodilno $r \times r$ podmatriko in definiramo $f_r(\lambda) = \det(T_r - \lambda I)$, potem z razvijanjem po zadnji vrstici pridemo do rekurzivne formule

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

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

**Izrek 4.** *Polinomi $f_0, \ldots, f_n$ tvorijo Sturmovo zaporedje, kar pomeni, da zadoščajo naslednjim trem točkam:*

*1) $f_0(\lambda) \neq 0$ za vsak $\lambda$.*
*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 3.** *Nerazcepna tridiagonalna simetrična matrika ima enostavne lastne vrednosti.*

## 3.2 Lanczosev algoritem za lastne vrednosti

Standardni Lanczosev algoritem je

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

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, ki jih označimo z $\theta_k^{(k)} < \cdots < \theta_1^{(k)}$.

Ker je $T_k$ nerazcepna ($\beta_j \neq 0$ za $j = 1, \ldots, k-1$), so vse Ritzeve vrednosti enostavne.

# Lastnosti 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

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

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

**Izrek 5.** *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 $\theta$, potem velja*

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

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

# 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 1.** *Če je $P_i$ spektralni projektor, ki pripada $\lambda_i$, potem v primeru $P_i v_1 \neq 0$ velja*

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

*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 6.** *Pri enakih predpostavkah kot zgoraj velja*

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

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

Naj bodo $\theta_k^{(k)} \le \cdots \le \theta_1^{(k)}$ Ritzeve vrednosti $A$ glede na $\mathcal{K}_k(A, v_1)$.

**Lema 2.** *Velja*

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

**Izrek 7.** *Velja*

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

*kjer je*

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

# Praktične ocene

Naj bodo $\lambda_n \leq \cdots \leq \lambda_1$ 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 $\mathcal{K}_m(A, v_1)$ 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 $\min_i |\lambda_i - \theta_j^{(m)}| \leq \|r_j\|_2 = \beta_m |e_m^T y_j|$.

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

Velja

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

in

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

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

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

# Lastnosti Lanczosevega algoritma

- časovna zahtevnost in prostorske zahteve so manjše kot pri Arnoldiju,
- več se da povedati o konvergenci,
- do izgube ortogonalnosti pride hitreje kot pri Arnoldiju,
- ne moremo ugotoviti večkratnosti lastne vrednosti,
- lahko imamo težave z navidezno konvergenco,
- pojavijo se prividi lastnih vrednosti.

# 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 8. [Paige]** *Če izvajamo Lanczosev algoritem v aritmetiki z osnovno zaokrožitveno napako $\epsilon$, potem v $k$-tem koraku za izračunane 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 Demmelovi knjigi, 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 $\beta_k |e_k^T y_i| \le \sqrt{\epsilon}\|T_k\|$
$z = z - (u_{i,k}^T z) u_{i,k}$
$\beta_j = \|z\|_2$
če je $\beta_j = 0$, potem prekini računanje
$v_{j+1} = z/\beta_j$

# Harmonične Ritzeve vrednosti

$\theta$ 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 3.** *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 $\theta$ velja, da je $\theta^{-1}$ lastna vrednost posplošenega problema lastnih vrednosti*

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

Če uporabimo premik $\sigma$, je $\theta$ harmonična Ritzeva vrednost in $u$ harmonični vektor, če je

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

$\theta$ 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 $\sigma$ 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 9. [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 $\lambda_{-j}$ in podobno $\theta_j^{(k)}$ monotono padajoče konvergira k $\lambda_j$

# Prepletanje Ritzevih in harmoničnih Ritzevih vrednosti

Če izvedemo $k$ korakov Lanczoseve metode za simetrično matriko $A$, dobimo $AV_k = V_{k+1}T_{k+1,k} = V_kT_k + \beta_k v_{k+1}e_k^T$, kjer je

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

Naj bodo $\theta_1, \ldots, \theta_k$ Ritzeve vrednosti, $\widetilde{\theta}_1, \ldots, \widetilde{\theta}_k$ pa harmonične Ritzeve vrednosti matrike $A$ glede na $V_k$. Podobno naj bodo $\theta_1(\sigma), \ldots, \theta_k(\sigma)$ Ritzeve vrednosti, $\widetilde{\theta}_1(\sigma), \ldots, \widetilde{\theta}_k(\sigma)$ pa harmonične Ritzeve vrednosti matrike $A - \sigma I$ glede na isti podprostor kot prej.

Pokazati se da, da so lastne vrednosti matrike $(k+1) \times (k+1)$ matrike

$$\begin{bmatrix} T_k - \sigma I & \beta_{k+1} e_k \\ \beta_{k+1}^T e_k^T & \beta_{k+1}^2/\delta_k \end{bmatrix},$$

kjer je $\delta_k^{-1} = e_k^T(T_k - \sigma I)^{-1}e_k$, ravno $0$ in $\widetilde{\theta}_1(\sigma), \ldots, \widetilde{\theta}_k(\sigma)$. Od tod sledi, da se Ritzeve in harmonične Ritzeve vrednosti $A - \sigma I$ prepletajo v smislu

$$\cdots < \widetilde{\theta}_{-1}(\sigma) < \theta_{-1}(\sigma) < 0 < \theta_1(\sigma) < \widetilde{\theta}_1(\sigma) < \cdots,$$

kjer so novi indeksi določeni glede na oddaljenost $0$ v pozitivni oz. negativni smeri.