Numerička analiza
11. predavanje

Autori: Sanja Singer i Nela Bosner
Predavač: Nela Bosner

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

PMF – Matematički odjel, Zagreb

Sadržaj predavanja

- Lokacija svojstvenih vrijednosti:
  - Geršgorinov teorem.
  - Cassinijevi ovali.
- Varijacijske karakterizacije svojstvenih vrijednosti hermitskih matrica:
  - Rayleigh–Ritzov teorem.
  - Courant–Fischerov teorem.
  - Weylov teorem.
- Multiplikativne perturbacije hermitskih matrica:
  - Sylvesterov teorem.
  - Teorem Ostrowskog.

Lokacija svojstvenih vrijednosti

Geršgorinov teorem

Neka je A kvadratna matrica. Ako je možemo napisati kao

$$A = D + \varepsilon B,$$

pri čemu je

- D dijagonalna,
- B ima nul-dijagonalu,
- a ε je neki mali broj,

onda se svojstvene vrijednosti matrice A nalaze u maloj okolini oko dijagonalnih elemenata matrice D.

Ovakovo razmišljanje može se formalizirati Geršgorinovim teoremom.

Geršgorinov teorem (nastavak)

Teorem (Geršgorin). Neka je A kvadratna matrica reda n i neka je

$$R_i'(A) = \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|, \quad 1 \leq i \leq n.$$

$R_i'(A)$ je suma apsolutnih vrijednosti elemenata u i-tom retku, bez dijagonalnog.

Sve svojstvene vrijednosti matrice A nalaze se u uniji od n diskova

$$G(A) := \bigcup_{i=1}^{n} \{z \in \mathbb{C} \mid |z - a_{ii}| \leq R_i'(A)\}.$$

Nadalje, ako unija k tih diskova tvori povezano područje koje nema presjeka s ostalih $n-k$ diskova, tada se točno k svojstvenih vrijednosti nalazi unutar tog područja.

Geršgorinov teorem (nastavak)

Dokaz. Neka je λ bilo koja svojstvena vrijednost od A i neka je $Ax = \lambda x$. Među elementima svojstvenog vektora x postoji najveći element po apsolutnoj vrijednosti, takav da je

$$|x_p| \geq |x_i|, \quad i = 1, \ldots, n,$$

i mora biti $x_p \neq 0$, zbog $x \neq 0$. Pretpostavka $Ax = \lambda x$ daje

$$\lambda x_p = [\lambda x]_p = [Ax]_p = \sum_{j=1}^{n} a_{pj} x_j,$$

ili, ekvivalentno,

$$x_p(\lambda - a_{pp}) = \sum_{\substack{j=1 \\ j \neq p}}^{n} a_{pj} x_j.$$

Geršgorinov teorem (nastavak)

Iz ove jednadžbe, korištenjem nejednakosti trokuta, dobivamo

$$|x_p||\lambda - a_{pp}| = \left| \sum_{\substack{j=1 \\ j \neq p}}^{n} a_{pj} x_j \right| \leq \sum_{\substack{j=1 \\ j \neq p}}^{n} |a_{pj} x_j| = \sum_{\substack{j=1 \\ j \neq p}}^{n} |a_{pj}||x_j|$$

$$\leq |x_p| \sum_{\substack{j=1 \\ j \neq p}}^{n} |a_{pj}| = |x_p| R_p'(A).$$

Zbog $x_p \neq 0$, slijedi da je $|\lambda - a_{pp}| \leq R_p'(A)$, za taj indeks p.

Drugi dio teorema bavi se lokacijom nultočaka, ako postoji presjek više krugova (v. Horn–Johnson, Matrix Analysis). ■

Geršgorinov teorem (nastavak)

Geršgorinov teorem može se primijeniti i na matricu A^T. Definiramo

$$C'_j(A) = \sum_{\substack{i=1 \\ i \neq j}}^{n} |a_{ij}|, \quad 1 \leq i \leq n.$$

$C'_j(A)$ je suma apsolutnih vrijednosti elemenata u j-tom stupcu od A, bez dijagonalnog.

Korolar. Sve svojstvene vrijednosti kvadratne matrice A leže u uniji od n diskova

$$G(A^T) := \bigcup_{j=1}^{n} \left\{ z \in \mathbb{C} \mid |z - a_{jj}| \leq C'_j(A) \right\}.$$

Ako je unija od k diskova povezana i odvojena od ostatka, unutar nje se nalazi točno k svojstvenih vrijednosti. ■

Geršgorinov teorem (nastavak)

Iz prethodnih rezultata slijedi da se sve svojstvene vrijednosti matrice A nalaze u presjeku $G(A) \cap G(A^T)$.

Postoji nekoliko poboljšanja ovog rezultata. Prvi od njih se dobiva korištenjem sličnosti:

- Za regularan S, matrice A i $S^{-1}AS$ imaju jednake svojstvene vrijednosti.
- Za S je najbolje/najlakše uzeti dijagonalnu matricu

$$S = D = \mathrm{diag}(p_1, p_2, \ldots, p_n), \quad p_i > 0.$$

- Elementi slične matrice su $(D^{-1}AD)_{ij} = a_{ij} \cdot p_j / p_i$.
- Primijenimo Geršgorinov teorem na $D^{-1}AD$.

Geršgorinov teorem (nastavak)

Korolar. Neka je A kvadratna matrica i neka su $p_1, \ldots, p_n$ pozitivni realni brojevi. Sve svojstvene vrijednosti matrice A leže u području

$$G(D^{-1}AD) := \bigcup_{i=1}^{n} \left\{ z \in \mathbb{C} \mid |z - a_{ii}| \le \frac{1}{p_i} \sum_{\substack{j=1 \\ j \ne i}}^{n} p_j |a_{ij}| \right\},$$

i u

$$G((D^{-1}AD)^T) := \bigcup_{j=1}^{n} \left\{ z \in \mathbb{C} \mid |z - a_{jj}| \le p_j \sum_{\substack{i=1 \\ i \ne j}}^{n} \frac{1}{p_i} |a_{ij}| \right\}.$$

■

Cassinijevi ovali

Teorem (Cassinijevi ovali). Svaka svojstvena vrijednost matrice A leži unutar barem jednog od $n(n-1)/2$ Cassinijeva ovala

$$|\lambda - a_{kk}||\lambda - a_{\ell\ell}| \leq R_k'(A) \cdot R_\ell'(A) = \Bigg(\sum_{\substack{j=1 \\ j \neq k}}^{n} |a_{kj}| \Bigg) \cdot \Bigg(\sum_{\substack{j=1 \\ j \neq \ell}}^{n} |a_{\ell j}| \Bigg),$$

i unutar barem jednog ovala

$$|\lambda - a_{kk}||\lambda - a_{\ell\ell}| \leq C_k'(A) \cdot C_\ell'(A) = \Bigg(\sum_{\substack{i=1 \\ i \neq k}}^{n} |a_{ik}| \Bigg) \cdot \Bigg(\sum_{\substack{i=1 \\ i \neq \ell}}^{n} |a_{i\ell}| \Bigg),$$

gdje je $1 \leq k, \ell \leq n$ i $k \neq \ell$. ■

Varijacijska karakterizacija svojstvenih vrijednosti hermitskih matrica

Rayleigh–Ritzov teorem

Prisjetimo se, hermitske matrice

- imaju samo realne svojstvene vrijednosti,
- mogu se dijagonalizirati unitarnim sličnostima,
- bez obzira jesu li svojstvene vrijednosti različite ili ne, uvijek postoji ortogonalna baza svojstvenih vektora.

U ovom i sljedećem poglavlju, svojstvene su vrijednosti bilo koje hermitske matrice A, reda n, uvijek poredane tako da vrijedi:

$$\lambda_{\min} := \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n =: \lambda_{\max}.$$

Najmanja i najveća svojstvena vrijednost lako se mogu karakterizirati kao rješenje problema uvjetne minimizacije, odnosno, maksimizacije.

Rayleigh–Ritzov teorem (nastavak)

Teorem (Rayleigh–Ritz). Neka je A hermitska matrica sa svojstvenim vrijednostima poredanim od najmanje do najveće. Tada je

$$\lambda_1 x^* x \leq x^* A x \leq \lambda_n x^* x,$$

za sve $x \in \mathbb{C}^n$. Nadalje,

$$\lambda_{\max} = \lambda_n = \max_{x \neq 0} \frac{x^* A x}{x^* x} = \max_{x^* x = 1} x^* A x,$$

$$\lambda_{\min} = \lambda_1 = \min_{x \neq 0} \frac{x^* A x}{x^* x} = \min_{x^* x = 1} x^* A x.$$

Rayleigh–Ritzov teorem (nastavak)

Dokaz. Budući da je A hermitska, postoji unitarna matrica U koja dijagonalizira A

$$A = U\Lambda U^*, \quad \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

Za proizvoljni $x \in \mathbb{C}^n$ vrijedi

$$x^* A x = (x^* U)\Lambda(U^* x) = \sum_{i=1}^{n} \lambda_i |(U^* x)_i|^2.$$

U ovoj linearnoj kombinaciji svojstvenih vrijednosti λ_i, za koeficijente vrijedi $|(U^* x)_i|^2 \geq 0$. Iz $\lambda_{\min} \leq \lambda_i \leq \lambda_{\max}$ slijedi

$$\lambda_{\min} \sum_{i=1}^{n} |(U^* x)_i|^2 \leq x^* A x \leq \lambda_{\max} \sum_{i=1}^{n} |(U^* x)_i|^2.$$

Rayleigh–Ritzov teorem (nastavak)

Budući da je U unitarna, onda je

$$\sum_{i=1}^{n} |(U^* x)_i|^2 = (x^* U)(U^* x) = x^* x.$$

Dakle, dokazali smo da je

$$\lambda_{\min} x^* x \leq x^* A x \leq \lambda_{\max} x^* x.$$

Primijetite da je ocjena oštra (tj. ne može se poboljšati) i dostiže se ako je x svojstveni vektor matrice A za $\lambda_{\min}$, odnosno, $\lambda_{\max}$. ∎

Rayleigh–Ritzov teorem daje varijacijsku karakterizaciju dvije rubne svojstvene vrijednosti. Što je s ostalima?

Courant–Fischerov teorem

Ideja. Promatrajmo sve vektore x koji su ortogonalni na svojstveni vektor u_1.

Za takve x-ove, zbog $u_1^* x = 0$, vrijedi sljedeća modifikacija:

$$x^* A x = \sum_{i=1}^{n} \lambda_i |(U^* x)_i|^2 = \sum_{i=1}^{n} \lambda_i |u_i^* x|^2 = \sum_{i=2}^{n} \lambda_i |u_i^* x|^2.$$

Zbog nenegativnosti ove linearne kombinacije vrijednosti $\lambda_2, \ldots, \lambda_n$ i njihovog poretka, imamo

$$\sum_{i=2}^{n} \lambda_i |u_i^* x|^2 \geq \lambda_2 \sum_{i=2}^{n} |u_i^* x|^2 = \lambda_2 \sum_{i=1}^{n} |(U^* x)_i|^2 = \lambda_2 x^* x.$$

Nejednakost će postati jednakost, ako izaberemo $x = u_2$.

Courant–Fischerov teorem (nastavak)

Dakle, dobili smo varijacijsku karakterizaciju za drugu najmanju svojstvenu vrijednost λ_2:

$$\min_{\substack{x \neq 0 \\ x \perp u_1}} \frac{x^* A x}{x^* x} = \min_{\substack{x^* x = 1 \\ x \perp u_1}} x^* A x = \lambda_2.$$

Dodatno, ako na lijevoj strani ovog izraza,

- umjesto svojstvenog vektora u_1 za vrijednost λ_1,
- uzmemo bilo koji vektor $w_1 \in \mathbb{C}^n$,

onda se taj izraz može samo smanjiti, tj. njegov maksimum se dotiže baš za $w_1 = u_1$. Dokaz ide projekcijom w_1 na u_1.

Na sličan se način mogu izvesti i općenitije karakterizacije.

Courant–Fischerov teorem (nastavak)

Teorem (Courant–Fischer). Neka je A hermitska matrica sa svojstvenim vrijednostima λ_i, poredanim od najmanje do najveće

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n.$$

Neka je k cijeli broj za koji vrijedi $1 \leq k \leq n$. Tada je

$$\min_{w_1, w_2, \ldots, w_{n-k} \in \mathbb{C}^n} \max_{\substack{x \neq 0, x \in \mathbb{C}^n \\ x \perp w_1, w_2, \ldots, w_{n-k}}} \frac{x^* A x}{x^* x} = \lambda_k,$$

$$\max_{w_1, w_2, \ldots, w_{k-1} \in \mathbb{C}^n} \min_{\substack{x \neq 0, x \in \mathbb{C}^n \\ x \perp w_1, w_2, \ldots, w_{k-1}}} \frac{x^* A x}{x^* x} = \lambda_k.$$

■

Courant–Fischerov teorem (nastavak)

Courant–Fischerov teorem možemo iskazati i na alternativan način. Uočimo sljedeće:

- Kod prve karakterizacije svojstvene vrijednosti λ_k biramo $x \perp w_1, w_2, \ldots, w_{n-k}$ koji daje maksimalan Rayleighev kvocijent $\frac{x^* A x}{x^* x}$.
- Ako definiramo potprostore $\mathcal{N} = \operatorname{span}\{w_1, \ldots, w_{n-k}\}$ i $\mathcal{M} = \mathcal{N}^\perp$, tada biramo $x \in \mathcal{M}$.
- Dalje, vrijedi $\dim(\mathcal{N}) \leq n - k$ i $\dim(\mathcal{M}) = n - \dim(\mathcal{N}) \geq n - (n - k) = k$.
- Ako su $w_1, \ldots, w_{n-k}$ linearno nezavisni, tada skup po kojem tražimo maksimalni Rayleighev kvocijent ima dimenziju $\dim(\mathcal{M}) = k$.

Courant–Fischerov teorem (nastavak)

- Ako su $w_1, \ldots, w_{n-k}$ linearno zavisni, tada skup po kojem tražimo maksimalni Rayleighev kvocijent ima dimenziju $\dim(\mathcal{M}) > k$. U tom slučaju maksimalni Rayleighev kvocijent po takvom skupu je veći ili jednak maksimalnom Rayleighevom kvocijentu po bilo kojem njegovom k-dimenzionalnom podskupu.
- Budući da mi tražimo maksimalne Rayleigheve kvocijente za sve moguće izbore skupa $\mathcal{N}$, i onda među njima tražim najmanji, isplati se uzimati samo linearno nezavisne $w_1, \ldots, w_{n-k}$.
- Osim toga, odabir skupa $\mathcal{N}$ pa biranje $x \in \mathcal{M} = \mathcal{N}^\perp$ je ekvivalentno direktnom odabiru skupa $\mathcal{M}$, takvog da po prethodnim razmatranjima ima dimenziju $\dim(\mathcal{M}) = k$, i onda opet biramo $x \in \mathcal{M}$.

Courant–Fischerov teorem (nastavak)

- Za drugu karakterizaciju svojstvene vrijednosti λ_k možemo napraviti analognu analizu.

Ovime smo dobili alternativan iskaz teorema.

Courant–Fischerov teorem (nastavak)

Teorem (Courant–Fischer). Neka je A hermitska matrica sa svojstvenim vrijednostima λ_i, poredanim od najmanje do najveće

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n.$$

Neka je k cijeli broj za koji vrijedi $1 \leq k \leq n$. Tada je

$$\min_{\substack{\mathcal{M} \subset \mathbb{C}^n \\ \dim(\mathcal{M}) = k}} \max_{x \in \mathcal{M}, x \neq 0} \frac{x^* A x}{x^* x} = \lambda_k,$$

$$\max_{\substack{\mathcal{M} \subset \mathbb{C}^n \\ \dim(\mathcal{M}) = n-k+1}} \min_{x \in \mathcal{M}, x \neq 0} \frac{x^* A x}{x^* x} = \lambda_k.$$

■

Courant–Fischerov teorem (nastavak)

Ako sa $u_1, \ldots, u_n$ označimo svojstvene vektore matrice A koje redom odgovaraju svojstvenim vrijednostima $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$, tada možemo uočiti sljedeće.

- Kod prve karakterizacije svojstvene vrijednosti λ_k minimum se postiže za $\mathcal{M} = \mathrm{span}\{u_1, u_2 \ldots, u_k\}$.
- Kod druge karakterizacije svojstvene vrijednosti λ_k maksimum se postiže za $\mathcal{M} = \mathrm{span}\{u_k, u_{k+1} \ldots, u_n\}$.

Primjene varijacijske karakterizacije

Teorem (Cauchyev teorem ispreplitanja). Neka je A $n \times n$ hermitska matrica sa nepadajućim svojstvenim vrijednostima

$$\lambda_1(A) \le \lambda_2(A) \le \cdots \le \lambda_n(A).$$

$k \times k$ matrica B, gdje je $k \le n$, naziva se kompresija od A ako postoji $n \times k$ ortonormalna matrica Q takva da je $B = Q^* A Q$. Tada, ako sa $\lambda_i(B)$ označimo nepadajuće svojstvene vrijednosti od B

$$\lambda_1(B) \le \lambda_2(B) \le \cdots \le \lambda_k(B),$$

za sve $i \le k$ vrijedi

$$\lambda_i(A) \le \lambda_i(B) \le \lambda_{i+n-k}(A).$$

Primjene varijacijske karakterizacije (nastavak)

Dokaz. Neka su sa $v_1, \ldots, v_k$ označeni svojstveni vektori matrice B koje redom odgovaraju svojstvenim vrijednostima $\lambda_1(B), \ldots, \lambda_k(B)$. Neka je za $i \leq k$ definiran i-dimenzionalni potprostor $\mathcal{M}_i = \operatorname{span}\{v_1, \ldots, v_i\}$. Tada prema Courant–Fischerovom teoremu znamo da vrijedi

$$\lambda_i(B) = \max_{x \in \mathcal{M}_i, x \neq 0} \frac{x^* B x}{x^* x} = \max_{x \in \mathcal{M}_i, x \neq 0} \frac{x^* Q^* A Q x}{x^* Q^* Q x} = \max_{y \in Q\mathcal{M}_i, y \neq 0} \frac{y^* A y}{y^* y}$$

$$\geq \min_{\substack{\mathcal{M} \subset \mathbb{C}^n \\ \dim(\mathcal{M}) = i}} \max_{x \in \mathcal{M}, x \neq 0} \frac{x^* A x}{x^* x} = \lambda_i(A).$$

Ako za $i \leq k$ definiram $k - i + 1$-dimenzionalni potprostor $\mathcal{M}_{k-i+1} = \operatorname{span}\{v_i, \ldots, v_k\}$. Tada opet prema Courant–Fischerovom teoremu znamo da vrijedi

Primjene varijacijske karakterizacije (nastavak)

$$\begin{aligned}
\lambda_i(B) &= \min_{x\in\mathcal{M}_{k-i+1},x\neq 0} \frac{x^*Bx}{x^*x} = \min_{x\in\mathcal{M}_{k-i+1},x\neq 0} \frac{x^*Q^*AQx}{x^*Q^*Qx} \\
&= \min_{y\in Q\mathcal{M}_{k-i+1},y\neq 0} \frac{y^*Ay}{y^*y} \\
&\leq \max_{\substack{\mathcal{M}\subset\mathbb{C}^n \\ \dim(\mathcal{M})=k-i+1}} \min_{x\in\mathcal{M},x\neq 0} \frac{x^*Ax}{x^*x} = \lambda_{i+n-k}(A),
\end{aligned}$$

jer je $k - i + 1 = n - (i + n - k) + 1$. ■

Primjene varijacijske karakterizacije (nastavak)

Teorem (Weyl). Neka su A i B hermitske matrice istog reda n, i neka su $\lambda_i(A)$, $\lambda_i(B)$, $\lambda_i(A+B)$ svojstvene vrijednosti u rastućem poretku. Za svaki $k = 1, \ldots, n$ vrijedi

$$\lambda_k(A) + \lambda_1(B) \leq \lambda_k(A+B) \leq \lambda_k(A) + \lambda_n(B).$$

Dokaz. Krenimo od Rayleigh–Ritzovog teorema za matricu B. Za proizvoljni $x \in \mathbb{C}^n$, $x \neq 0$, imamo ogradu

$$\lambda_1(B) \leq \frac{x^* B x}{x^* x} \leq \lambda_n(B).$$

Sad iskoristimo Courant–Fischerov teorem za matricu $A + B$.

Primjene varijacijske karakterizacije (nastavak)

Za bilo koji $k = 1, \ldots, n$, iz prve karakterizacije imamo redom

$$\begin{aligned}
\lambda_k(A+B) &= \min_{\substack{\mathcal{M}\subset\mathbb{C}^n \\ \dim(\mathcal{M})=k}} \ \max_{x\in\mathcal{M}, x\neq 0} \frac{x^*(A+B)x}{x^*x} \\
&= \min_{\substack{\mathcal{M}\subset\mathbb{C}^n \\ \dim(\mathcal{M})=k}} \ \max_{x\in\mathcal{M}, x\neq 0} \left(\frac{x^*Ax}{x^*x} + \frac{x^*Bx}{x^*x} \right) \\
&\geq \min_{\substack{\mathcal{M}\subset\mathbb{C}^n \\ \dim(\mathcal{M})=k}} \ \max_{x\in\mathcal{M}, x\neq 0} \left(\frac{x^*Ax}{x^*x} + \lambda_1(B) \right) \\
&= \lambda_k(A) + \lambda_1(B).
\end{aligned}$$

Na isti način dobivamo i gornju ogradu. ∎

Multiplikativne perturbacije hermitskih matrica

Kongruencija i inercija

Definicija. Neka su A i B $n \times n$ hermitske matrice i neka je S regularna matrica, takva da je $B = SAS^*$. Kažemo da je B kongruentna sa A. Kongruencija je relacije ekvivalencije.

Definicija. Definiramo inerciju od hermitske matrice A kao

$$i(A) = (i_+(A), i_0(A), i_-(A)),$$

gdje su

$$\begin{aligned} i_+(A) &= \text{broj pozitivnih svojstvenih vrijednosti od A} \\ i_0(A) &= \text{broj svojstvenih vrijednosti od A jednakih 0} \\ i_-(A) &= \text{broj negativnih svojstvenih vrijednosti od A} \end{aligned}$$

Inercija hermitske matrice

Za hermitsku matricu A napišimo njenu spektralnu dekompoziciju (svojstvene vrijednosti u rastućem poretku)

$$A = U\Lambda U^*, \qquad \Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n), \quad U^*U = UU^* = I.$$

Dijagonalnu matricu Λ dalje raspišemo kao

$$\Lambda = |\Lambda|^{1/2} I(A) |\Lambda|^{1/2},$$

gdje je

$$I(A) = \begin{bmatrix} -I_{i_-(A)} & 0 & 0 \\ 0 & O_{i_0(A)} & 0 \\ 0 & 0 & I_{i_+(A)} \end{bmatrix}$$

pri čemu je I_k matrica identitete dimenzije k, a O_ℓ nul-matrica dimenzije ℓ.

Inercija hermitske matrice (nastavak)

Matricu A tada možemo napisati kao produkt

$$A = ZI(A)Z^*, \qquad Z = U|\Lambda|^{1/2}.$$

Ako umjesto $|\Lambda|^{1/2}$ definiramo dijagonalu matricu $L = \mathrm{diag}(\ell_1, \ldots, \ell_n)$ sa

$$\ell_i = \begin{cases} |\lambda_i|^{1/2} & \lambda_i \neq 0 \\ 1 & \lambda_i = 0 \end{cases}$$

tada A ponovo možemo napisati kao produkt

$$A = SI(A)S^*, \qquad S = UL.$$

pri čemu je S regularna. Dakle možemo zaključiti da je A kongruentna sa $I(A)$. Štoviše, ako je matrica A kongruentna sa nekom dijagonalnom matricom D, tada je broj negativnih i pozitivnih dijagonalnih elemenata u D uvijek jednak, o čemu govori Sylvesterov teorem.

Sylvester ov teorem

Teorem. Neka su A i B hermitske matrice. Tada A i B su kongruentne ako i samo ako je $i(A) = i(B)$.

Dokaz.

- Pretpostavimo da je $i(A) = i(B)$. Tada su A i B kongruentne sa $I(A) = I(B)$, pa su i A i B međusobno kongruentne (relacija ekvivalencije).
- Neka su A i B kongruentne, tj. neka je $A = SBS^*$ za regularnu matricu S. Odmah slijedi da je $\mathrm{rang}(A) = \mathrm{rang}(B)$, što povlači $i_0(A) = i_0(B)$.
    - Neka su $\lambda_1(A) \leq \cdots \leq \lambda_{i_-(A)} < 0$ negativne svojstvene vrijednosti od A, i $v_1, \ldots, v_{i_-(A)}$ pripadni ortonormirani svojstveni vektori.

Sylvesterov teorem (nastavak)

- Označimo $S_-(A) = \mathrm{span}\{v_1, \ldots, v_{i_-(A)}\}$ sa $\dim(S_-(A)) = i_-(A)$.
- Tada za proizvoljni $x \in S_-(A)$, $x \neq 0$ slijedi da je $x = \sum_{i=1}^{i_-(A)} \xi_i v_i$, što povlači

$$x^* A x = \left(\sum_{i=1}^{i_-(A)} \xi_i v_i \right)^* \left(\sum_{i=1}^{i_-(A)} \lambda_i \xi_i v_i \right) = \sum_{i=1}^{i_-(A)} \lambda_i |\xi_i|^2 < 0.$$

- Za $y = S^* x$ dalje slijedi

$$y^* B y = x^* S B S^* x = x^* A x < 0,$$

pri čemu je $y \in S^* S_-(A)$ proizvoljni element potprostora dimenzije $i_-(A)$.

Sylvesterov teorem (nastavak)

- Dakle, možemo zaključiti da je

$$\max_{y \in S^* S_-(A), y \neq 0} \frac{y^* A y}{y^* y} < 0$$

pa mora vrijediti i za minimum po svim potprostorima dimenzije $i_-(A)$. Prema Courant–Fischerovom teoremu imamo

$$\lambda_{i_-(A)}(B) = \min_{\substack{\mathcal{M} \subset \mathbb{C}^n \\ \dim(\mathcal{M}) = i_-(A)}} \max_{y \in \mathcal{M}, y \neq 0} \frac{y^* A y}{y^* y} < 0$$

gdje su $\lambda_1(B) \leq \lambda_2(B) \leq \cdots \leq \lambda_n(B)$ svojstvene vrijednosti od B.

- Napokon možemo zaključiti da je $i_-(B) \geq i_-(A)$.

Sylvesterov teorem (nastavak)

- Analogno se dokazuje i $i_-(B) \leq i_-(A)$ odakle slijedi jednakost.
- Jednako se dokazuje i $i_+(B) = i_+(A)$.

$\blacksquare$

Teorem Ostrowskog

Teorem (Ostrowski). Neka su A $n \times n$ hermitska matrica i S $n \times n$ regularna matrica. Tada za svaki $k = 1, \ldots, n$ postoji $\theta_k \in [\lambda_1(SS^*), \lambda_n(SS^*)]$ takav da je

$$\lambda_k(SAS^*) = \theta_k \lambda_k(A),$$

pri čemu se uzima da su svojstvene vrijednosti svih matrica uređene u nepadajućem poretku.

Napomena. Što je S bliža unitarnoj matrici, to je θ_k bliži 1.

Dokaz. Promotrimo matricu $A - \lambda_k(A)I$. Očito je

$$\lambda_k(A - \lambda_k(A)I) = 0.$$

Teorem Ostrowskog (nastavak)

Zbog Sylvesterovog teorema je

$$\lambda_k(S(A - \lambda_k(A)I)S^*) = 0,$$

jer se sačuvao broj pozitivnih, negativnih i nula svojstvenih vrijednosti. Dakle, vrijedi

$$\lambda_k(SAS^* - \lambda_k(A)SS^*) = 0.$$

Dalje, iz Weylovog teorem slijedi

$$\begin{aligned}\lambda_k(SAS^*)+\lambda_1(-\lambda_k(A)SS^*) \leq \lambda_k(SAS^* - \lambda_k(A)SS^*) = 0 \\ \leq \lambda_k(SAS^*) + \lambda_n(-\lambda_k(A)SS^*).\end{aligned}$$

Teorem Ostrowskog (nastavak)

Iz donje ograde prethodne nejednakosti slijedi

$$\lambda_k(SAS^*) \leq -\lambda_1(-\lambda_k(A)SS^*) = \lambda_n(\lambda_k(A)SS^*)$$
$$= \begin{cases} \lambda_k(A) \cdot \lambda_n(SS^*), & \lambda_k(A) \geq 0 \\ \lambda_k(A) \cdot \lambda_1(SS^*), & \lambda_k(A) < 0 \end{cases}$$

a iz gornje ograde slijedi

$$\lambda_k(SAS^*) \geq -\lambda_n(-\lambda_k(A)SS^*) = \lambda_1(\lambda_k(A)SS^*)$$
$$= \begin{cases} \lambda_k(A) \cdot \lambda_1(SS^*), & \lambda_k(A) \geq 0 \\ \lambda_k(A) \cdot \lambda_n(SS^*), & \lambda_k(A) < 0 \end{cases}$$

Teorem Ostrowskog (nastavak)

Za $\lambda_k(A) \geq 0$ imamo

$$\lambda_k(A) \cdot \lambda_1(SS^*) \leq \lambda_k(SAS^*) \leq \lambda_k(A) \cdot \lambda_n(SS^*),$$

a za $\lambda_k(A) < 0$

$$\lambda_k(A) \cdot \lambda_n(SS^*) \leq \lambda_k(SAS^*) \leq \lambda_k(A) \cdot \lambda_1(SS^*),$$

pa u oba slučaja za $\lambda_k(A) \neq 0$ vrijedi (za $\lambda_k(A) = 0$ je i $\lambda_k(SAS^*) = 0$)

$$\lambda_1(SS^*) \leq \frac{\lambda_k(SAS^*)}{\lambda_k(A)} \leq \lambda_n(SS^*).$$

■

Teorem Ostrowskog (nastavak)

Korolar. Neka su A $n \times n$ hermitska matrica i S $n \times n$ regularna matrica. Neka su svojstvene vrijednosti svih matrica uređene u nepadajućem poretku. Tada za svaki $k = 1, \ldots, n$ vrijedi

$$|\lambda_k(SAS^*) - \lambda_k(A)| \leq \|SS^* - I\|_2 |\lambda_k(A)|.$$

Dokaz. Iz teorema Ostrowskog je $\lambda_k(SAS^*) = \theta_k \lambda_k(A)$ i $\theta_k \in [\lambda_1(SS^*), \lambda_n(SS^*)]$, odakle je

$$\lambda_k(SAS^*) - \lambda_k(A) = (\theta_k - 1)\lambda_k(A), \quad \lambda_1(SS^*)-1 \leq \theta_k - 1 \leq \lambda_n(SS^*)-1.$$

Jednostavno možemo zaključiti da je

$$|\theta_k - 1| \leq \max\{|\lambda_1(SS^*) - 1|, |\lambda_n(SS^*) - 1| = \|SS^* - I\|_2.$$