Numerička analiza
26. predavanje

Saša Singer

singer@math.hr
web.math.hr/~singer

PMF – Matematički odjel, Zagreb

Sadržaj predavanja

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

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 poglavlju, svojstvene su vrijednosti 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

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

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

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

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

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

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

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

■

Primjene varijacijske karakterizacije

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

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

$$
\begin{aligned}
\lambda_k(A+B) &= \min_{w_1, w_2, \ldots, w_{n-k} \in \mathbb{C}^n} \; \max_{\substack{x \neq 0, x \in \mathbb{C}^n \\ x \perp w_1, w_2, \ldots, w_{n-k}}} \frac{x^*(A+B)x}{x^*x} \\
&= \min_{w_1, w_2, \ldots, w_{n-k} \in \mathbb{C}^n} \; \max_{\substack{x \neq 0, x \in \mathbb{C}^n \\ x \perp w_1, w_2, \ldots, w_{n-k}}} \left(\frac{x^*Ax}{x^*x} + \frac{x^*Bx}{x^*x} \right) \\
&\geq \min_{w_1, w_2, \ldots, w_{n-k} \in \mathbb{C}^n} \; \max_{\substack{x \neq 0, x \in \mathbb{C}^n \\ x \perp w_1, w_2, \ldots, w_{n-k}}} \left(\frac{x^*Ax}{x^*x} + \lambda_1(B) \right) \\
&= \lambda_k(A) + \lambda_1(B).
\end{aligned}
$$

Na isti način dobivamo i gornju ogradu. ■

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

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

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

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

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

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

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

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

Geršgorinov teorem

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

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

i u

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

■

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 $k \neq \ell$. ■