AUTOVALORI ED AUTOVETTORI DI MATRICI

Luca Gemignani

23 Marzo 2020

DEFINIZIONI

Def:
Sia $\lambda \in \mathbb{C}$ *e sia* $A \in \mathbb{C}^{n \times n}$. λ *è detto autovalore di* A *se esiste* $\mathbf{x} \in \mathbb{C}^n$, $\mathbf{x} \neq 0$ *tale che* $A\mathbf{x} = \lambda \mathbf{x}$. *Il vettore* $\mathbf{x}$ *è detto autovettore (destro) relativo all'autovalore* λ.

Def:
Sia $\lambda \in \mathbb{C}$ *e sia* $A \in \mathbb{C}^{n \times n}$. λ *è autovalore di* A *se e solo se* $p(\lambda) = \det(A - \lambda I) = 0$.

Def:
Il polinomio $p(z) = \det(A - zI) = (-1)^n z^n + p_{n-1} z^{n-1} + \ldots + p_0$ *è detto polinomio caratteristico di* A

Dal Teorema Fondamentale dell'Algebra segue che $p(z)$ ha n zeri $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ eventualmente coincidenti

MOLTEPLICITA'

Siano $\lambda_{i_1}, \ldots, \lambda_{i_k}$ gli autovalori di A a due a due distinti. Allora

$$p(z) = (-1)^n \prod_{\ell=1}^{k} (z - \lambda_{i_\ell})^{\sigma_\ell}, \quad \sum_{\ell=1}^{k} \sigma_\ell = n.$$

Def:

σ_i *è detta molteplicità algebrica dell'autovalore* λ_i

Def:

$\tau_i = \mathrm{dimKer}(A - \lambda_i I)$ *è detta molteplicità geometrica dell'autovalore* λ_i

Teo:

Sia λ *autovalore di* A *con molteplicità algebrica e geometrica rispettivamente uguale a* σ *e* τ*. Allora vale* $\sigma \geq \tau$*.*

ESEMPI

Determinare autovalori, autovettori, molteplicità algebriche e geometriche delle seguenti matrici:

- $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$;
- $A = (a_{i,j})$ di ordine n con $a_{i+1,i} = 1$ e $a_{i,j} = 0$ altrimenti;
- $A = (a_{i,j})$ di ordine n con $a_{i+1,i} = 1$, $a_{1,n} = \alpha \in \mathbb{R}$ e $a_{i,j} = 0$ altrimenti.

DIAGONALIZZABILITA'

Def:

$A \in \mathbb{C}^{n\times n}$ si dice diagonalizzabile se esiste $S \in \mathbb{C}^{n\times n}$ invertibile tale che $S^{-1}AS = D$ è diagonale ($D = (d_{i,j})$ con $d_{i,j} = 0$ se $i \neq j$)

Oss:

$$\det(D - zI) = \det(S^{-1}AS - zI) = \det(S^{-1}AS - zS^{-1}S) =$$
$$= \det(S^{-1}(A - zI)S) = \det S^{-1} \det(A - zI) \det S = \det(A - zI)$$

Quindi A e D hanno lo stesso polinomio caratteristico e quindi gli stessi autovalori con le stesse molteplicità algebriche. Riguardo le molteplicità geometriche per D vale $\sigma = \tau$ per ogni autovalore distinto. E per le molteplità geometriche?

Oss:

$$S^{-1}AS = D \iff AS = SD \iff AS\mathbf{e}_j = SD\mathbf{e}_j \forall j \iff A\mathbf{s}_j = d_j\mathbf{s}_j \forall j$$

Quindi le colonne di S sono autovettori di A e A e D hanno gli stessi autovalori con le stesse molteplicità algebriche e geometriche.

ESEMPI

Discutere la diagonalizzabilità delle seguenti matrici e nel caso determinare la matrice S degli autovettori.

- $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$;
- $A = (a_{i,j})$ di ordine n con $a_{i+1,i} = 1$ e $a_{i,j} = 0$ altrimenti;
- $A = (a_{i,j})$ di ordine n con $a_{i+1,i} = 1$, $a_{1,n} = \alpha \in \mathbb{R}$ e $a_{i,j} = 0$ altrimenti.

CLASSI DI MATRICI DIAGONALIZZABILI

- matrici con autovalori distinti $\sigma_i = \tau_i = 1$, $1 \leq i \leq n$;
- (teorema spettrale) matrici simmetriche reali, i.e., $A = A^T$, $A \in \mathbb{R}^{n \times n}$;
- (teorema spettrale) matrici hermitiane, i.e., $A = A^H = \bar{A}^T$, $A \in \mathbb{C}^{n \times n}$;

Oss:
Dal teorema spettrale segue anche che le matrici simmetriche hanno autovalori ed autovettori reali.

CALCOLO DI AUTOVALORI E AUTOVETTORI

- Il calcolo è difficile perchè gli autovalori NON si esprimono come funzioni razionali dei coefficienti della matrice. Quindi NON esistono algoritmi che in un numero finito di operazioni aritmetiche restituiscono gli autovalori. Gli algoritmi di calcolo sono ITERATIVI: generano successioni di approssimazioni che al limite tendono agli autovalori.

1. Studio del condizionamento per suggerire criteri di arrresto;
2. Teoremi di Localizzazione per suggerire approssimazioni iniziali

CONDIZIONAMENTO I

Sia λ autovalore semplice $\sigma = \tau = 1$ di A con autovettore $\mathbf{x}$ e sia $B = A + \epsilon F$ la matrice perturbata. Si mostra che $\exists \mathbf{x}(\epsilon) \colon I \to C^n,\ \boldsymbol{\lambda}(\epsilon) \colon I \to C$ con $\mathbf{x}(\epsilon) \doteq \mathbf{x} + \epsilon \mathbf{z}$ e $\lambda(\epsilon) \doteq \lambda + \epsilon\eta$ tali che

$$B\mathbf{x}(\epsilon) = \lambda(\epsilon)\mathbf{x}(\epsilon), \quad \forall \epsilon \in I.$$

Segue che

$$(A + \epsilon F)(\mathbf{x} + \epsilon \mathbf{z}) \doteq (\lambda + \epsilon\eta)(\mathbf{x} + \epsilon \mathbf{z})$$

da cui

$$A\mathbf{x} + \epsilon A\mathbf{z} + \epsilon F\mathbf{x} \doteq \lambda\mathbf{x} + \epsilon\lambda\mathbf{z} + \epsilon\eta\mathbf{x}$$

e quindi

$$A\mathbf{z} + F\mathbf{x} \doteq \lambda\mathbf{z} + \eta\mathbf{x}$$

CONDIZIONAMENTO II

Consideriamo un autovettore sinistro di A relativo all'autovalore λ cioè $\mathbf{y}^H A = \lambda \mathbf{y}^H$, $\mathbf{y} \neq \mathbf{0}$. Otteniamo allora

$$\mathbf{y}^H A\mathbf{z} + \mathbf{y}^H F\mathbf{x} \doteq \lambda \mathbf{y}^H \mathbf{z} + \eta \mathbf{y}^H \mathbf{x}$$

e quindi

$$\mathbf{y}^H F\mathbf{x} \doteq \eta \mathbf{y}^H \mathbf{x}, \quad \eta \doteq \frac{\mathbf{y}^H F\mathbf{x}}{\mathbf{y}^H \mathbf{x}}$$

Si conclude che

$$|\lambda(\epsilon) - \lambda| \doteq |\frac{\mathbf{y}^H F\mathbf{x}}{\mathbf{y}^H \mathbf{x}}||\epsilon| \leq (1/|\mathbf{y}^H \mathbf{x}|) \parallel \mathbf{y} \parallel_2 \parallel \mathbf{x} \parallel_2 \parallel \epsilon F \parallel_2$$

da cui per autovettori normalizzati (norma-2=1)

$$|\lambda(\epsilon) - \lambda| \leq (1/|\mathbf{y}^H \mathbf{x}|) \parallel \epsilon F \parallel_2$$

La perturbazione sulla matrice è amplificata da $\mathcal{K}(\lambda) = (1/|\mathbf{y}^H \mathbf{x}|)$

LOCALIZZAZIONE

Teo:

(Teorema di Gerschgorin) Sia $A = (a_{i,j}) \in \mathbb{C}^{n \times n}$ *e siano*

$$\mathcal{K}_i = \{z \in \mathbb{C} \colon |z - a_{i,i}| \leq \sum_{j=1, j \neq i}^{n} |a_{i,j}|\}, \quad 1 \leq i \leq n.$$

Allora λ *autovalore di* $A \Rightarrow \lambda \in \cup \mathcal{K}_i$.

Oss:

$\mathcal{K} = \{z \in \mathbb{C} \colon |z - a| \leq r\}$ *è un cerchio di centro* $a \in \mathbb{C}$ *e raggio* $r \geq 0$. *Il teorema sopra è detto anche teorema dei cerchi di Gerschgorin.* $\mathcal{K}$ *viene rappresentato nel piano di Gauss.*

DIMOSTRAZIONE

Partiamo dalla relazione autovalore/autovettore.

$$Ax = \lambda \mathbf{x}, \mathbf{x} \neq \mathbf{0} \iff \sum_{j=1}^{n} a_{i,j} x_j = \lambda x_i, \; 1 \leq i \leq n.$$

Scegliamo p: $|x_p| = \| \mathbf{x} \|_\infty \neq 0$. Vale

$$\sum_{j=1, j \neq p}^{n} a_{p,j} x_j = (\lambda - a_{p,p}) x_p \to | \sum_{j=1, j \neq p}^{n} a_{p,j} x_j | = |(\lambda - a_{p,p}) x_p|$$

da cui

$$|\lambda - a_{p,p}| |x_p| \leq \sum_{j=1, j \neq p}^{n} |a_{p,j}| |x_j|$$

e quindi dividendo per $|x_p|$ segue che

$$|\lambda - a_{p,p}| \leq \sum_{j=1, j \neq p}^{n} |a_{p,j}| |x_j| / |x_p| \leq \sum_{j=1, j \neq p}^{n} |a_{p,j}| \to \lambda \in \mathcal{K}_p.$$

ESTENSIONI E APPLICAZIONI

- A e A^T hanno gli stessi autovalori. Se applico il teorema ad A^T trovo che gli autovalori di A appartengono a $\cup \mathcal{H}_i$ con $\mathcal{H}_i = \{z \in \mathbb{C} \colon |z - a_{i,i}| \leq \sum_{j=1, j \neq i}^{n} |a_{j,i}|\}$ cerchi "per colonne". Segue che gli autovalori appartengono a $\cup \mathcal{H}_i \cap \cup \mathcal{K}_i$.
- Il teorema di Gerschgorin fornisce uno strumento per lo studio dell'invertibilità di una matrice.

ESERCIZIO: Si studi il condizionamento in norma 2 della matrice tridiagonale di ordine n con elementi diagonali 3 ed elementi soprsa e sotto diagonali uguali ad 1.