



CALCOLO DEGLI AUTOVALORI E FONDAMENTI DELLA MATEMATICA NUMERICA (Localizzazione degli autovalori)

In questa In questa lezione introdurremo alcuni strumenti utili per la localizzazione degli autovalori nel piano complesso. Infatti, prima ancora di introdurre metodi numerici specifici per il calcolo degli autovalori, è opportuno dotarsi di strumenti che consentano di localizzarli nel piano complesso.

Data una matrice A , per la quale si intende determinare gli autovalori λ , è utile disporre di criteri che permettano di individuare, a priori, le regioni del piano complesso in cui tali autovalori sono situati. In tal modo, anziché effettuare la ricerca sull'intero piano complesso, è possibile circoscrivere l'indagine a specifiche aree, facilitando così l'efficienza e l'efficacia dei metodi numerici impiegati.

Procediamo quindi con l'enunciazione di alcuni risultati fondamentali sulla localizzazione degli autovalori, i quali risultano utili in qualsiasi algoritmo volto al calcolo degli autovalori. Riprendiamo per prima cosa alcune informazioni già introdotte, ma analizzandole da una nuova prospettiva.

Abbiamo visto nel precedente nucleo che il raggio spettrale di una matrice A , indicato con $\rho(A)$, è sempre minore o uguale alla norma di A , indipendentemente dalla norma considerata

$$\rho(A) \leq \|A\|.$$

In particolare, poiché il raggio spettrale corrisponde al massimo modulo degli autovalori, possiamo affermare che, per ogni autovalore λ appartenente allo spettro di A , il suo modulo è minore o uguale alla norma della matrice

$$|\lambda| \leq \|A\| \quad \forall \lambda \in \sigma(A).$$

Se siamo in grado di calcolare per la matrice A una norma $\|A\|$ (preferibilmente la più semplice possibile), che corrisponde a un numero reale positivo, allora la disuguaglianza

$$|\lambda| \leq \|A\| \quad \forall \lambda \in \sigma(A),$$

ci garantisce che tutti gli autovalori di A si trovano all'interno del cerchio del piano complesso di raggio $\|A\|$ e centro nell'origine. Pertanto, disponiamo già di uno strumento utile per delimitare con certezza la regione del piano complesso in cui ricercare gli autovalori di A .

Vediamo questo risultato di localizzazione sulla seguente matrice

$$A = \begin{bmatrix} 10 & 2 & 3 \\ -1 & 2 & -1 \\ 0 & 1 & 3 \end{bmatrix},$$

che possiede come autovalori

$$\lambda_1 = 9.69, \quad \lambda_{2,3} = 2.66 \pm i0.69.$$

È possibile calcolare una norma della matrice A , ad esempio la norma 1 (essendo il massimo delle somme dei valori assoluti di ciascuna colonna), ottenendo il valore

$$\|A\|_1 = 11.$$



Di conseguenza, tracciando nel piano complesso un cerchio di centro nell'origine e raggio pari a 11, possiamo affermare con certezza che tutti gli autovalori di A si trovano all'interno della regione definita da

$$|\lambda| \leq 11.$$

A titolo di verifica, si osserva che l'autovalore reale $\lambda_1 = 9.69$ appartiene a questa regione, così come gli altri due autovalori complessi λ_2, λ_3 risultano situati all'interno del medesimo cerchio, di raggio 11, nel piano complesso.

Il risultato appena illustrato rappresenta un primo strumento utile per la localizzazione degli autovalori. Tuttavia, non è necessariamente uno strumento affinato in quanto il cerchio di raggio pari alla norma $\|A\|$ può risultare eccessivamente ampio rispetto alla reale distribuzione degli autovalori. In altre parole, gli autovalori potrebbero essere situati ben all'interno di questo cerchio. Quindi in qualche maniera siamo indotti a cercare gli autovalori in una regione troppo più grande di quella in cui questi sono effettivamente contenuti.

Introduciamo ora un secondo risultato di localizzazione degli autovalori, più raffinato rispetto al precedente. Consideriamo una matrice A e definiamo due matrici A_S e A_{SS} nel modo seguente

$$A_S = \frac{1}{2}(A + A^T), \quad A_{SS} = \frac{1}{2}(A - A^T),$$

dove A^T denota la trasposta di A .

La matrice A_S è detta parte simmetrica di A , mentre la matrice A_{SS} , indicata con il pedice "SS" stante per *skew-symmetric*, è la parte antisimmetrica di A . Osserviamo le seguenti proprietà:

- A_S è una matrice simmetrica, ovvero $A_S^T = A_S$. Di conseguenza, essa ammette autovalori reali, ovvero $\lambda \in \mathbb{R}$;
- A_{SS} è una matrice antisimmetrica, ovvero $A_{SS}^T = -A_{SS}$. Per tale motivo, i suoi autovalori sono immaginari puri $\lambda \in \mathbb{I}$, ossia con parte reale nulla.

Questa decomposizione è utile in quanto consente di analizzare separatamente le componenti simmetriche e antisimmetriche della matrice A , e in particolare di raffinare le tecniche di localizzazione degli autovalori. Infatti, utilizzando le matrici A_S e A_{SS} per decomporre la matrice A in

$$A = A_S + A_{SS},$$

ovvero, per scomporre la matrice A nella somma di una matrice simmetrica A_S , che ha autovalori reali, e una matrice antisimmetrica A_{SS} , con autovalori immaginari puri. A questo punto, possiamo sfruttare le informazioni relative agli autovalori delle matrici A_S e A_{SS} . In particolare, considerando una matrice A arbitraria, è possibile affermare che la parte reale di ciascun autovalore di A è compresa tra il minimo e il massimo degli autovalori della parte simmetrica A_S della matrice. Formalmente

$$\lambda_{\min}(A_S) \leq \operatorname{Re}(\lambda(A)) \leq \lambda_{\max}(A_S).$$

Questo risultato è estremamente utile, in quanto consente di limitare la ricerca degli autovalori di una matrice A a una fascia verticale del piano complesso, delimitata lateralmente dagli estremi reali degli autovalori della matrice simmetrica associata A_S . Tale fascia rappresenta una stima più raffinata rispetto al semplice cerchio centrato nell'origine di raggio $\|A\|$, e costituisce un secondo strumento di localizzazione degli autovalori, più preciso del precedente.

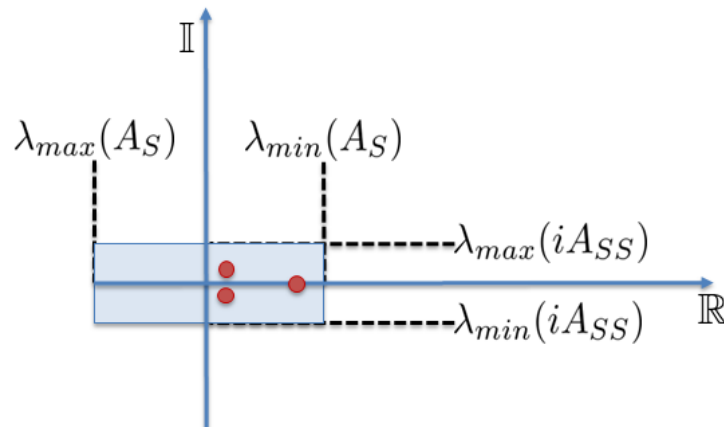


Figura 1: Rappresentazione grafica delle seguenti relazioni di localizzazione degli autovalori utilizzando le matrici A_S e A_{SS} .

In maniera simile, se prendiamo la parte immaginaria del generico autovalore di A , essa è compresa fra il minimo autovalore della matrice iA_{SS} (con A_{SS} parte antisimmetrica di A) e il massimo autovalore di iA_{SS}

$$\lambda_{\min}(iA_{SS}) \leq \text{Im}(\lambda(A)) \leq \lambda_{\max}(iA_{SS}).$$

Quindi, consideriamo ora la matrice A_{SS} e moltiplichiamola per i , l'unità immaginaria. Si ottiene così la matrice iA_{SS} , la quale risulta essere hermitiana (ossia uguale al proprio coniugato trasposto) e, di conseguenza, ammette autovalori reali. Questo consente di ottenere una caratterizzazione utile degli autovalori della matrice originale A .

Questi doppie disuguaglianze per la parte reale degli autovalori di A

$$\lambda_{\min}(A_S) \leq \text{Re}(\lambda(A)) \leq \lambda_{\max}(A_S)$$

e per la parte immaginaria degli autovalori di A

$$\lambda_{\min}(iA_{SS}) \leq \text{Im}(\lambda(A)) \leq \lambda_{\max}(iA_{SS}),$$

risultano particolarmente utili, in quanto forniscono un'indicazione più precisa riguardo alla localizzazione degli autovalori nel piano complesso. In particolare, consentono di stimare separatamente l'ampiezza della parte reale e della parte immaginaria degli autovalori, offrendo così una rappresentazione più concreta della regione in cui essi possono trovarsi.

La Figura 1 fornisce una rappresentazione grafica delle seguenti relazioni di localizzazione degli autovalori

$$\begin{aligned} \lambda_{\min}(A_S) &\leq \text{Re}(\lambda(A)) \leq \lambda_{\max}(A_S), \\ \lambda_{\min}(iA_{SS}) &\leq \text{Im}(\lambda(A)) \leq \lambda_{\max}(iA_{SS}), \end{aligned}$$

dove sull'asse reale sono riportati i valori $\lambda_{\min}(A_S)$ e $\lambda_{\max}(A_S)$, mentre sull'asse immaginario compaiono $\lambda_{\min}(iA_{SS})$ e $\lambda_{\max}(iA_{SS})$. Il rettangolo così individuato nel piano complesso rappresenta quindi una regione che certamente contiene tutti gli autovalori della matrice A . Si ottiene, in tal modo, una stima dell'estensione della parte reale e della parte immaginaria degli autovalori della matrice considerata.



Esaminiamo ora come il risultato appena discusso si applichi alla matrice A considerata nell'esempio precedente

$$A = \begin{bmatrix} 10 & 2 & 3 \\ -1 & 2 & -1 \\ 0 & 1 & 3 \end{bmatrix},$$

che possiede come autovalori

$$\lambda_1 = 9.69, \quad \lambda_{2,3} = 2.66 \pm i0.69.$$

La matrice $A_S = \frac{1}{2}(A + A^T)$, parte simmetrica di A , è

$$A_S = \begin{bmatrix} 10 & 0.5 & 1.5 \\ 0.5 & 2 & 0 \\ 1.5 & 0 & 3 \end{bmatrix},$$

mentre la parte antisimmetrica di A , $A_{SS} = \frac{1}{2}(A - A^T)$, è data da

$$A_{SS} = \begin{bmatrix} 0 & 1.5 & 1.5 \\ -1.5 & 0 & -1 \\ -1.5 & 1 & 0 \end{bmatrix}.$$

Calcolando gli autovalori delle matrici A_S e A_{SS} , troviamo che

$$\lambda_{1,S} = 1.9575, \quad \lambda_{2,S} = 2.7058, \quad \lambda_{3,S} = 10.3367,$$

$$\lambda_{1,SS} = 0, \quad \lambda_{2,SS} = i2.3452, \quad \lambda_{3,SS} = -i2.3452.$$

Osserviamo che A_S è una matrice simmetrica e, in quanto tale, possiede autovalori reali. La matrice A_{SS} , ovvero la parte antisimmetrica di A , ha invece tre autovalori che, in generale, sono numeri immaginari puri. In particolare, uno di essi è nullo, mentre gli altri due, essendo complessi coniugati, sono opposti l'uno dell'altro.

Possiamo quindi calcolare il minimo e il massimo degli autovalori della parte simmetrica A_S

$$\lambda_{\min}(A_S) = 1.9575,$$

$$\lambda_{\max}(A_S) = 10.3367,$$

i quali forniscono un intervallo per la parte reale degli autovalori della matrice A . Analogamente, calcolando il minimo e il massimo degli autovalori della matrice iA_{SS} (ottenuta moltiplicando per l'unità immaginaria i la parte antisimmetrica A_{SS}), otteniamo

$$\lambda_{\min}(iA_{SS}) = -2.3452,$$

$$\lambda_{\max}(iA_{SS}) = 2.3452,$$

che delimitano l'intervallo per la parte immaginaria degli autovalori di A . Ne consegue che i tre autovalori della matrice A sono sicuramente contenuti all'interno del rettangolo del piano complesso individuato da questi quattro estremi.

Abbiamo dunque osservato che, conoscendo gli autovalori della parte simmetrica A_S e della parte antisimmetrica A_{SS} di una matrice A , è possibile determinare immediatamente un rettangolo nel piano complesso che contiene tutti gli autovalori della matrice A .

Introduciamo ora un ulteriore risultato, ancor più raffinato, per la localizzazione degli autovalori: il cosiddetto teorema di Gershgorin.

Per prima cosa, definiamo i cosiddetti cerchi di riga e cerchi di colonna di una matrice. Consideriamo una matrice arbitraria A e fissiamo una riga i . Definiamo l'insieme dei punti $z \in \mathbb{C}$ tali che il modulo della



differenza tra z e l'elemento diagonale a_{ii} sia minore o uguale alla somma dei moduli degli elementi (extradiagonali) a_{ij} della riga i -esima escluso l'elemento diagonale, quindi $j \neq i$, cioè

$$|z - a_{ii}| \leq \sum_{j=1(j \neq i)}^n |a_{ij}|.$$

Questa regione è un cerchio, del piano complesso, che ha centro a_{ii} e raggio dato dal termine di destra $\sum_{j=1(j \neq i)}^n |a_{ij}|$

$$R_i = \left\{ z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{j=1(j \neq i)}^n |a_{ij}| \right\},$$

Definiamo R_i un cerchio riga della matrice A .

Ad ogni riga i , corrisponde un cerchio R_i . Pertanto, per una matrice A $n \times n$, otteniamo n cerchi di riga in totale. Indichiamo con S_R l'unione di tutti questi cerchi, ovvero

$$S_R = R_1 \cup R_2 \cup \dots \cup R_n.$$

Si può dimostrare, ed è questo il primo enunciato del teorema di Gershgorin, che tutti gli autovalori della matrice A appartengono all'unione dei cerchi riga, ossia

$$\lambda(A) \in S_R.$$

Riprendendo l'esempio fatto in precedenza sulla matrice A

$$A = \begin{bmatrix} 10 & 2 & 3 \\ -1 & 2 & -1 \\ 0 & 1 & 3 \end{bmatrix},$$

che possiede come autovalori

$$\lambda_1 = 9.69, \quad \lambda_{2,3} = 2.66 \pm i0.69.$$

Tale matrice A genera i seguenti cerchi riga

$$R_1 = \{z \in \mathbb{C} : |z - 10| \leq 5\},$$

$$R_2 = \{z \in \mathbb{C} : |z - 2| \leq 2\},$$

$$R_3 = \{z \in \mathbb{C} : |z - 3| \leq 1\},$$

Il cerchio R_1 ha centro nel punto $(10,0)$ ed un raggio pari a $|2| + |3| = |5|$, ossia la somma dei valori assoluti degli elementi della prima riga situati al di fuori della diagonale principale. Il cerchio R_2 ha centro in $(2,0)$ ed un raggio pari a $|-1| + |-1| = |2|$. Infine, il cerchio R_3 ha centro in $(3,0)$ ed un raggio pari a $|0| + |1| = |1|$.

La Figura 2 illustra i tre cerchi riga R_1, R_2, R_3 e gli autovalori $\lambda_1, \lambda_2, \lambda_3$ della matrice A di partenza. Si può affermare con certezza che tutti gli autovalori della matrice A si trovano certamente all'interno dell'unione dei cerchi riga, ovvero

$$S_R = R_1 \cup R_2 \cup R_3.$$

In particolare, in questo caso, gli autovalori stanno nell'unione di R_1 e R_2 , dato che R_3 è un sottoinsieme di R_2 .

Poiché il calcolo degli autovalori di una matrice A coincide con quello degli autovalori della sua trasposta A^T , risulta evidente che l'utilizzo dei cerchi riga per la localizzazione degli autovalori conduce agli stessi risultati che si otterrebbero impiegando i cosiddetti cerchi colonna.

Seguendo lo stesso paradigma dei cerchi riga, introduciamo ora i cerchi colonna. Per ciascun indice di colonna j , consideriamo il cerchio C_j costituito dall'insieme dei punti $z \in \mathbb{C}$ tali che

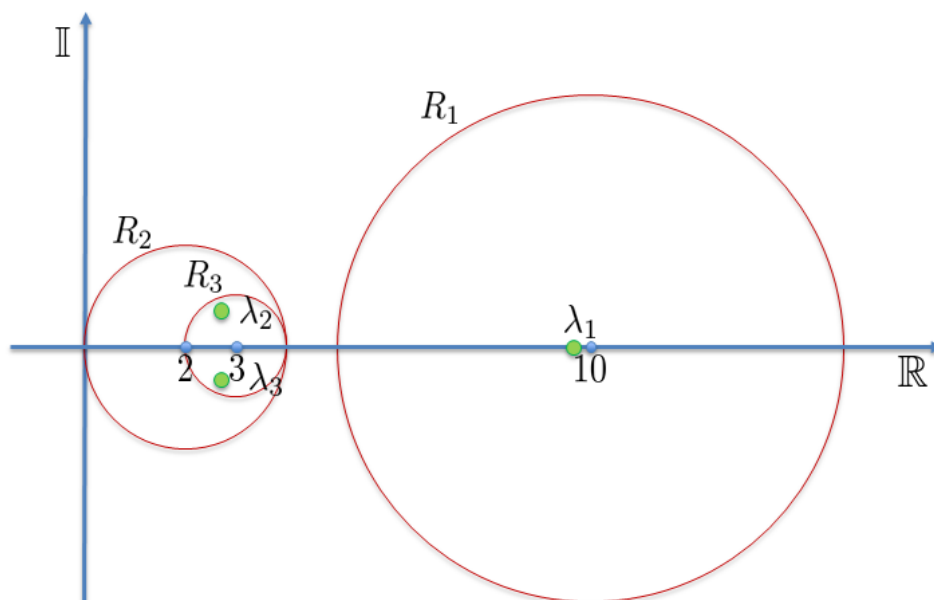


Figura 2: Grafico dei tre cerchi riga R_1, R_2, R_3 e degli autovalori $\lambda_1, \lambda_2, \lambda_3$ della matrice A dell'esempio considerato.

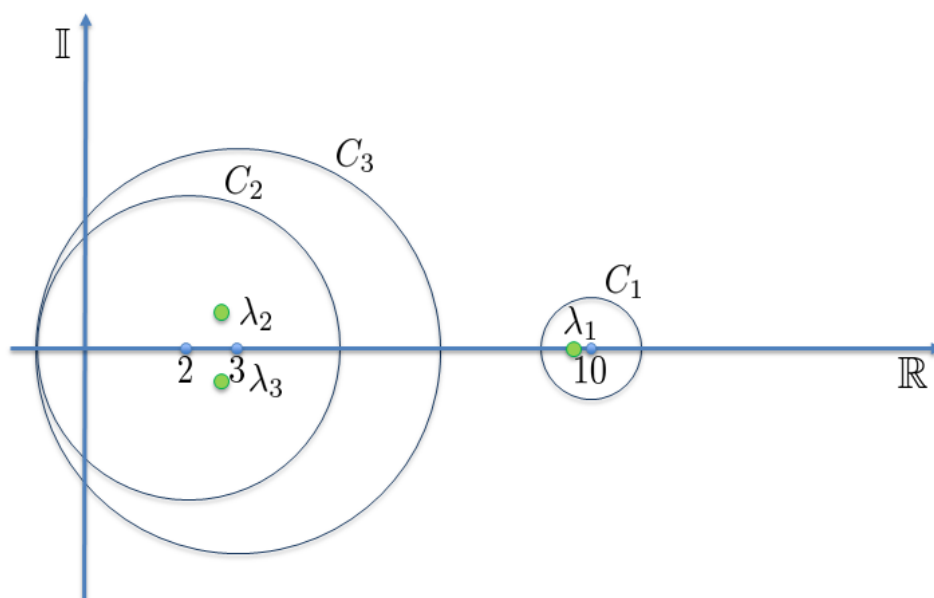


Figura 3: Grafico dei tre cerchi colonna C_1, C_2, C_3 e degli autovalori $\lambda_1, \lambda_2, \lambda_3$ della matrice A dell'esempio considerato.



$$C_j = \left\{ z \in \mathbb{C} : |z - a_{jj}| \leq \sum_{i=1(i \neq j)}^n |a_{ij}| \right\}.$$

In altre parole, il cerchio C_j ha centro nell'elemento diagonale a_{jj} e raggio pari alla somma dei moduli degli elementi extradiagonali della colonna j (ossia con somma fatta sulla colonna e per i che va da 1 a n , con $i \neq j$).

Indichiamo con S_C l'unione dei cerchi colonna

$$S_C = C_1 \cup C_2 \cup \dots \cup C_n.$$

Si può dimostrare che tutti gli autovalori della matrice A appartengono all'unione dei cerchi colonna, ovvero

$$\lambda(A) \in S_C.$$

La Figura 3 illustra i tre cerchi colonna C_1, C_2, C_3 e gli autovalori $\lambda_1, \lambda_2, \lambda_3$ della matrice A considerata nell'esempio precedente. Si può affermare con certezza che gli autovalori della matrice A si trovano all'interno dell'unione di tali cerchi

$$S_C = C_1 \cup C_2 \cup C_3.$$

In particolare, nel caso specifico, gli autovalori risultano compresi nell'unione di C_1 e C_3 , poiché C_2 è un sottoinsieme di C_3 .

Disponiamo dunque di due strumenti per la localizzazione degli autovalori: i cerchi riga e i cerchi colonna. Poiché gli autovalori devono appartenere sia all'unione dei cerchi riga sia all'unione dei cerchi colonna, si deduce che essi si trovano nell'intersezione di questi due insiemi. Tale intersezione, essendo a priori un sottoinsieme degli insiemi di partenza, è uguale o più piccola rispetto a ciascuno di essi. Pertanto, l'intersezione tra l'unione dei cerchi riga e l'unione dei cerchi colonna definisce una regione del piano complesso più restrittiva, entro la quale è possibile circoscrivere in modo più preciso la ricerca degli autovalori. La formalizzazione di questo concetto è la seguente

$$\lambda(A) \in S_R \cap S_C,$$

ovvero, gli autovalori della matrice A appartengono all'intersezione dell'unione dei cerchi riga S_R e dell'unione dei cerchi colonna S_C . Questa è la formulazione dal primo teorema di Gershgorin.

La Figura 4 mostra questo risultato per la matrice A dell'esempio precedentemente studiato, dove viene rappresenta l'unione dei cerchi riga S_R , l'unione dei cerchi colonna S_C e la relativa intersezione $S_R \cap S_C$. Gli autovalori $\lambda_1, \lambda_2, \lambda_3$ sono contenuti in $S_R \cap S_C$.

Abbiamo quindi ottenuto un miglioramento significativo, poiché ora possiamo concentrare la ricerca degli autovalori esclusivamente nella regione $S_R \cap S_C$. Ciò comporta una notevole riduzione dell'area del piano complesso in cui è necessario cercare gli autovalori, rappresentando un risultato importante nel contesto della localizzazione degli autovalori.

È possibile approfondire ulteriormente questo risultato esaminando il modo in cui gli autovalori sono raggruppati. A tal proposito, il secondo teorema di Gershgorin afferma che, considerando i cerchi riga R_i , e suddividendoli in due insiemi mediante l'unione di m cerchi e l'unione degli altri $n - m$ cerchi

$$S_1 = \bigcup_{i=1}^m R_i, \quad S_2 = \bigcup_{i=m+1}^n R_i,$$

ed inoltre accade che l'intersezione tra S_1 e S_2 è vuota

$$S_1 \cap S_2 = \emptyset,$$

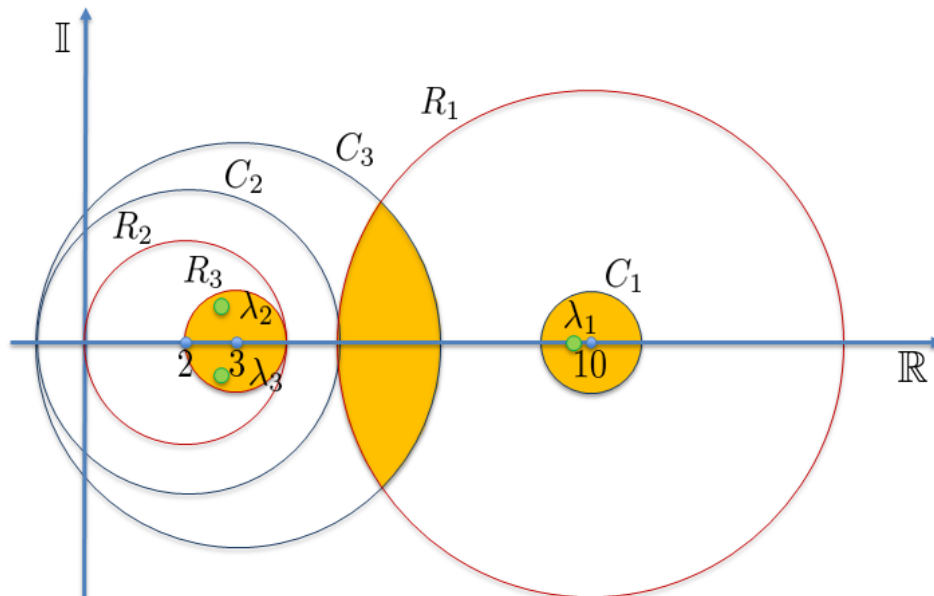


Figura 4: Rappresentazione dell'unione dei cerchi riga S_R , l'unione dei cerchi colonna S_C e la relativa intersezione $S_R \cap S_C$ (mostrato in giallo), per la matrice A dell'esempio. Gli autovalori $\lambda_1, \lambda_2, \lambda_3$ sono contenuti in $S_R \cap S_C$.

ovvero, se i cerchi selezionati definiscono due regioni disgiunte del piano complesso ($S_1 \cap S_2 = \emptyset$), allora possiamo affermare con certezza che la regione S_1 contiene esattamente m autovalori della matrice A , mentre la regione S_2 contiene i restanti $n - m$ autovalori.

Questo risultato, noto come secondo teorema di Gershgorin, fornisce non solo un criterio di localizzazione degli autovalori all'interno dell'unione dei cerchi riga (o dei cerchi colonna), ma permette anche di determinare quanti autovalori si trovano in ciascuna regione, a condizione che tali regioni siano disgiunte.

I risultati finora esposti sono risultati di localizzazione degli autovalori, in quanto permettono di individuare regioni del piano complesso all'interno delle quali ricercare gli autovalori di una matrice. Prima di introdurre algoritmi specifici per il calcolo numerico degli autovalori, è opportuno soffermarsi brevemente su un aspetto fondamentale: l'analisi di stabilità del problema del calcolo degli autovalori. Ogniquale si impiega un metodo numerico per risolvere un problema, è essenziale valutare quanto tale metodo sia sensibile a eventuali perturbazioni nei dati iniziali. In altre parole, ci si chiede, nel caso specifico, come variano gli autovalori di una matrice A al variare (anche minimo) dei suoi elementi. Questo tipo di analisi è particolarmente rilevante perché, come vedremo nel seguito di questo nucleo, le operazioni, nell'ambito del calcolo numerico, vengono effettuate mediante rappresentazioni approssimate dei numeri reali, tipicamente utilizzando quella che è chiamata rappresentazione floating point dei numeri macchina. Di conseguenza, i valori memorizzati e manipolati durante l'esecuzione di un algoritmo non corrispondono esattamente a quelli teorici, e introducono inevitabilmente errori numerici. Alla luce di ciò, è necessario comprendere se il problema del calcolo degli autovalori è ben



condizionato, ovvero se piccole perturbazioni nei coefficienti della matrice comportano soltanto piccole variazioni negli autovalori. In caso contrario, il problema risulterebbe mal condizionato e quindi potenzialmente instabile dal punto di vista computazionale. Tale analisi prende quindi il nome di analisi di stabilità o, più precisamente, di studio del condizionamento del problema della determinazione degli autovalori.

Iniziamo introducendo il concetto di stima a priori. Con questo termine si intende una previsione teorica dell'effetto di una perturbazione, senza eseguire calcoli espliciti, volta a stimare in anticipo quanto possano variare gli autovalori in conseguenza di una modifica della matrice.

Per introdurre il concetto, analizziamo il caso più semplice: la perturbazione di una matrice diagonale. Se la matrice è diagonale, il calcolo degli autovalori risulta immediato, in quanto questi coincidono esattamente con gli elementi presenti sulla diagonale principale. Supponiamo quindi di perturbare tale matrice diagonale, e analizziamo l'effetto di questa perturbazione sulla posizione degli autovalori

$$\tilde{A} = A + E, \quad \text{con } A \text{ matrice diagonale,}$$

dove E è una matrice di perturbazione che supponiamo avere elementi diagonali nulli

$$e_{ii} = 0, \quad i = 1, \dots, n.$$

Inoltre, gli altri elementi di E sono a priori diversi da zero. Per gli autovalori della matrice perturbata \tilde{A} , possiamo affermare che

$$|\lambda_i(\tilde{A}) - \lambda_i(A)| \leq \sum_{j=1}^n |e_{ij}|, \quad i = 1, \dots, n.$$

Questa relazione consente di comprendere immediatamente quale possa essere, in prima approssimazione, l'entità della variazione tra gli autovalori esatti della matrice originaria e quelli ottenuti a seguito della perturbazione.

Consideriamo, a titolo esemplificativo, il seguente esempio. Sia data una matrice diagonale A e una matrice di perturbazione E

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad E = \begin{pmatrix} 0 & \epsilon \\ \epsilon & 0 \end{pmatrix},$$

dove ϵ è un numero reale positivo sufficientemente piccolo, e i cui elementi sono nulli sulla diagonale principale e pari a ϵ sulla controdiagonale. La matrice A ha autovalori

$$\lambda_1 = 1, \quad \lambda_2 = 2,$$

mentre la matrice perturbata E ha autovalori

$$\tilde{\lambda}_{1,2} = \frac{3 \mp \sqrt{1 + \epsilon^2}}{2}.$$

Calcolando la differenza tra $\lambda_{1,2}$ e $\tilde{\lambda}_{1,2}$ possiamo osservare se la seguente disuguaglianza $|\lambda_i(\tilde{A}) - \lambda_i(A)| \leq \sum_{j=1}^n |e_{ij}|$, ($i = 1, \dots, n$), risulta soddisfatta. In effetti, si trova facilmente che

$$\begin{aligned} |\lambda_1 - \tilde{\lambda}_1| &= \frac{\sqrt{1 + \epsilon^2} - 1}{2} < \frac{\sqrt{(1 + \epsilon)^2} - 1}{2} = \frac{\epsilon}{2}, \\ |\lambda_2 - \tilde{\lambda}_2| &= \left| \frac{4 - 3\sqrt{1 + \epsilon^2}}{2} \right| < |\lambda_1 - \tilde{\lambda}_1| \leq \frac{\epsilon}{2}, \end{aligned}$$

che verifica la stima



$$|\lambda_i(\tilde{A}) - \lambda_i(A)| \leq \sum_{j=1}^n |e_{ij}| = \epsilon, \quad i = 1, \dots, n.$$

Ovviamente questo è un caso molto semplice di una matrice 2×2 , ma questo risultato vale per una matrice di ordine arbitrario.

Per una matrice $A \in \mathbb{R}^{n \times n}$ generica, ma assunta simmetrica, vale il seguente risultato noto come Teorema di Bauer-Fike. Sia $E \in \mathbb{R}^{n \times n}$ una matrice di perturbazione, e sia $A + E$ la matrice perturbata. Se $\mu \in \sigma(A + E)$ è un autovalore della matrice perturbata, allora, per ogni norma matriciale $\|\cdot\|$, esiste un autovalore $\lambda \in \sigma(A)$ tale che

$$\min_{\lambda \in \sigma(A)} |\lambda - \mu| \leq \|E\|_p.$$

Questo risultato ci dice che, se perturbo la matrice A di una matrice E che ha una norma piccola (quindi con elementi di perturbazione piccoli) allora la perturbazione sugli autovalori è anch'essa piccola.

$$\min_{\lambda \in \sigma(A)} |\lambda - \mu| \leq \|E\|_p.$$

Vediamo come applicare questo risultato alla seguente matrice simmetrica

$$A = \begin{pmatrix} 1 & 3 & 4 \\ 3 & 1 & 2 \\ 4 & 2 & 1 \end{pmatrix},$$

alla quale corrisponde il polinomio caratteristico

$$P_A(\lambda) = \lambda^3 - 3\lambda^2 - 26\lambda - 20,$$

e il cui spettro, corrispondente all'insieme degli autovalori di A , è

$$\sigma(A) = \{7.0747, -3.1879, -0.88679\},$$

che contiene tutti autovalori reali, in quanto la matrice A è simmetrica. Se perturbiamo la matrice A con una matrice di perturbazione E de tipo

$$E = \begin{pmatrix} \eta & \eta & \eta \\ \eta & \eta & \eta \\ \eta & \eta & \eta \end{pmatrix}, \quad \eta = 0.1, \quad \|E\|_1 = 0.3,$$

ovvero, con E che ha elementi tutti uguali a $\eta=0.1$ e con norma 1 $\|E\|_1 = 0.3$. Ricordiamo che la norma matriciale $\|\cdot\|_1$ è definita come il massimo, al variare delle colonne, della somma dei valori assoluti degli elementi lungo ciascuna colonna. Analogamente, la norma infinito $\|\cdot\|_\infty$ è definita come il massimo, al variare delle righe, della somma dei valori assoluti degli elementi lungo ciascuna riga. In particolare, nel caso in cui la matrice A sia simmetrica, si ha che infinito $\|A\|_1 = \|A\|_\infty$ poiché la struttura simmetrica garantisce che le somme per righe e per colonne coincidano. Di conseguenza avremo potuto considerare anche la norma infinito.

Gli autovalori della matrice perturbata $A + E$ sono

$$\sigma(A + E) = \{7.3722, -3.1869, -0.88530\},$$

i quali sono abbastanza simili a quelli della matrice A

$$\sigma(A) = \{7.0747, -3.1879, -0.88679\}.$$

Nella tabella seguente sono riportati gli autovalori della matrice perturbata $\mu \in \sigma(A + E)$ a confronto con la minima distanza dagli autovalori della matrice A

$\mu \in \sigma(A + E)$	$\min_{\lambda \in \sigma(A)} \lambda - \mu $
7.3722	0.29755
-3.1869	$9.590 \cdot 10^{-4}$



-0.88530	$1.486 \cdot 10^{-3}$
----------	-----------------------

Per quanto riguarda il primo autovalore, si osserva un errore pari a circa 0.29, che risulta inferiore a 0.3, ovvero al limite superiore fornito dalla stima del teorema di Bauer-Fike. Per gli altri due autovalori, le distanze tra gli autovalori perturbati e quelli originari risultano essere ancora più contenute, con errori dell'ordine rispettivamente di un millesimo e di un centesimo. Pertanto, i risultati ottenuti sono pienamente coerenti con quanto previsto dal teorema di Bauer-Fike, confermandone la validità nel fornire una stima a priori dell'effetto delle perturbazioni sulla posizione degli autovalori.

Il teorema di Bauer-Fike si estende anche al caso in cui la matrice A non sia simmetrica, purché sia normale. Una matrice A dice normale se

$$AA^T = A^T A.$$

dove A^T denota la trasposta di A . La matrice AA^T è una matrice simmetrica definita positiva. Tuttavia, non è in generale vero che, per una matrice arbitraria A , valga l'uguaglianza $AA^T = A^T A$. Nel caso in cui tale uguaglianza sia soddisfatta, la matrice A si dice matrice normale. Si osserva inoltre che, se la matrice A è simmetrica, allora essa è necessariamente anche normale

$$A \text{ simmetrica} \Rightarrow A \text{ normale},$$

dato che in questo caso

$$AA^T = A^T A = A^2.$$

In questa prime lezioni sugli autovalori abbiamo esaminato diversi risultati teorici di rilievo che permettono di localizzare gli autovalori nel piano complesso. Inoltre, abbiamo analizzato la sensibilità del problema della determinazione degli autovalori, in particolare per matrici simmetriche o normali, rispetto a perturbazioni sulla matrice stessa. Il teorema di Bauer-Fike afferma che, perturbando una matrice A con una matrice E , l'errore commesso sugli autovalori corrispondenti è limitato dalla dimensione della matrice di perturbazione, misurata tramite una qualunque norma della matrice E .