Alma Mater Studiorum · Università di Bologna

SCUOLA DI SCIENZE

Corso di Laurea in Matematica

Analisi spettrale della teoria perturbativa matriciale

Tesi di Laurea in Algebra Lineare Applicata

Relatrice:
Chiar.ma Prof.ssa
Valeria Simoncini

Presentata da:
Andrea Iacco

III Sessione
Anno Accademico 2022/2023

Introduzione

Chiunque, cominciando il proprio percorso matematico, si trova immediatamente a fare i conti con l'Algebra Lineare. I suoi strumenti e i suoi risultati infatti permeano tutte le discipline matematiche, costruendo dal nulla una teoria potente e applicabile potenzialmente ad ogni Scienza. Le protagoniste assolute di questo settore della matematica sono senz'altro le matrici, che possono rappresentare una vasta famiglia di oggetti: da vettori a funzioni lineari, da superfici nello spazio a grafi. Dopo pochi mesi di studio impariamo che le matrici più studiate, quelle quadrate, possiedono delle proprietà "spettrali" che permettono di semplificarle notevolmente. Vedendo queste matrici come operatori lineari ci basta infatti conoscere le catene di autospazi generalizzati e le autocoppie per ottenere una caratterizzazione completa, applicando le riduzioni alla forma canonica di Jordan. Conseguentemente l'insieme delle forme canoniche di Jordan può essere visto come un quoziente dello spazio di tutte le matrici quadrate.

Sappiamo quindi che due operatori lineari che hanno la stessa forma di Jordan si comportano in modo totalmente analogo, ma a questo punto sorge spontanea una domanda: se sono date due matrici i cui elementi sono molto vicini ma non identici, quanto è diversa la loro forma di Jordan? È possibile valutare la vicinanza di due matrici dalle loro rispettive proprietà spettrali, che le descrivono in modo così completo?

Possiamo esprimere il problema nel modo seguente: data una matrice quadrata A, e una matrice "piccola" E, quanto le proprietà spettrali di A sono diverse dalle proprietà spettrali di $A + E$?

Si potrebbe pensare che la soluzione sia intuitiva, ossia che le proprietà spettrali di matrici quasi identiche siano molto simili, ma la risposta non è così immediata. Si scopre infatti che queste proprietà, così importanti per lo studio di una matrice, sono anche fragili ed estremamente sensibili alle perturbazioni. E così una matrice con un solo autovalore di molteplicità n può trovarsi ad avere n autovalori semplici variando un solo specifico elemento, con ciascuno di questi autovalori che dista da quello originario molto di più di quanto l'elemento sia stato perturbato. Studiare come queste variazioni hanno

luogo, e in che condizioni si riesce a contenerle, è quindi di fondamentale importanza, sia dal punto di vista teorico, in quanto ci permette di descrivere ulteriormente lo spettro di una matrice, e quindi di conseguenza la matrice stessa, sia da quello applicativo, in tutti quegli algoritmi (per esempio di Data Mining) che si basano sullo spettro di una matrice per elaborare dei risultati.

In questo lavoro cercheremo di analizzare i principali risultati ottenuti studiando questo problema, esponendo alcune sfaccettature della cosiddetta teoria perturbativa matriciale. La trattazione si concentrerà principalmente sulle variazioni delle autocoppie di una data matrice, che chiameremo A, ogni volta che viene perturbata sommandola con una matrice E. Nel Capitolo 1 richiameremo alcuni concetti e definizioni di base, oltre a citare alcuni risultati tangenziali alla teoria che ci saranno utili durante l'esposizione. Dal Capitolo 2 entreremo nel vivo della teoria perturbativa, iniziando ad analizzare le perturbazioni di classi di matrici sempre più grandi. Inizialmente, supporremo che le matrici abbiano una struttura molto ricca, e vedremo che in questo caso le autocoppie saranno molto più resistenti alle perturbazioni. Procederemo poi a richiedere delle ipotesi sulla struttura delle matrici sempre meno stringenti, osservando come il problema diventi via via sempre più malcondizionato, fino ad abbandonarle nell'ultimo Capitolo, dove verrà esposta la teoria nel caso generale.

I risultati che otterremo saranno di due tipi. Quando possibile, cercheremo di trovare una maggiorazione per la distanza tra autovalori e autovettori corrispondenti prima e dopo la perturbazione. Questo approccio sarà ottimale solo nei casi "ben strutturati", ossia quando la quantità che maggiora è abbastanza piccola da fornire un'informazione utile sul problema. Proseguendo con la trattazione, e perdendo quindi struttura, dovremo passare ad un approccio diverso, che prova ad esprimere le autocoppie perturbate come funzioni della perturbazione, evidenziando il legame con le autocoppie originali. Questo secondo approccio in generale darà dei risultati meno certi, ma funzionerà anche nei casi meno strutturati, fornendo delle stime accurate, anche se pessimiste.

In ogni Capitolo sarà necessario sfruttare gli strumenti di molte aree della matematica per costruire degli apparati specifici, che poi useremo per ottenere i risultati desiderati. In particolare nel Capitolo 2 vedremo la trattazione nel caso di matrici Hermitiane, nel Capitolo 3 supporremo che le matrici siano prima normali, poi diagonalizzabili, e nel Capitolo 4 affronteremo la teoria applicata a matrici non strutturate.

Indice

Introduzione **i**

1 Richiami, preliminari e notazioni **1**

2 Matrici Hermitiane **7**

2.1 Autovalori: la disuguaglianza di Weyl 7

2.2 Autovettori: gli angoli di Davis-Kahan 13

3 Matrici Normali e Diagonalizzabili **23**

3.1 Matrici Normali 23

3.2 Matrici Diagonalizzabili 30

4 Matrici non Strutturate **39**

4.1 La teoria perturbativa di Lidskii 39

4.2 I diagrammi di Newton 50

Conclusioni **61**

Bibliografia **63**

Capitolo 1

Richiami, preliminari e notazioni

In questo Capitolo diamo delle definizioni di base ed esponiamo alcuni risultati di Algebra Lineare, che ci saranno utili nel corso della trattazione, senza dimostrazione.
Il nostro oggetto di studio saranno le matrici, e in particolare le loro autocoppie. In generale, assumeremo che le matrici che usiamo siano definite nel campo $\mathbb{C}$. Partiamo dal definire particolari tipi di matrici.

Definizione 1.1. *Sia $A \in \mathbb{C}^{m\times n}$ una matrice. Diremo che A è:*

i. ***Normale*** *se $AA^H = A^HA$;*

ii. ***Hermitiana*** *(**Simmetrica**) se $A = A^H$ ($A = A^T$);*

iii. ***Unitaria*** *se è quadrata e vale $AA^H = A^HA = I$;*

iv. ***Non singolare*** *se $m = n$ ed esiste una matrice B tale che $AB = BA = I$. In questo caso denoteremo $B = A^{-1}$ e diremo che B è la* matrice inversa *di A;*

v. ***Diagonale*** *se è quadrata e i suoi unici elementi non nulli sono sulla diagonale principale;*

vi. ***Doppiamente stocastica*** *se è quadrata, ha elementi reali non negativi e, detto $\hat{1}_k$ il vettore di lunghezza k ed elementi tutti uguali a 1, si ha $A\hat{1}_n = \hat{1}_n$ e $A^T\hat{1}_n = \hat{1}_n$;*

vii. ***Di permutazione*** *se le sue colonne sono, nell'ordine, $I_{\pi(1)}, \ldots, I_{\pi(n)}$, con $\pi \in \mathfrak{S}_n$ elemento dell'insieme delle permutazioni di n elementi e I la matrice identità.*

Ora esponiamo qualche risultato di base di Algebra Lineare, per caratterizzare le matrici normali, la segnatura delle matrici Hermitiane, e le matrici doppiamente stocastiche.

Teorema 1.2 (Spettrale). *Una matrice $A \in \mathbb{C}^{n\times n}$ è normale se e solo se è diagonalizzabile con trasformazioni unitarie, ossia se esiste una matrice unitaria Q tale che*

$$A = Q\Lambda Q^H,$$

dove Λ è una matrice diagonale. Se A è Hermitiana, allora gli elementi di Λ sono tutti numeri reali.

Teorema 1.3 (Sylvester-Jacobi). *Sia $A \in \mathbb{C}^{n\times n}$ una matrice Hermitiana, e siano $\lambda_1, \ldots, \lambda_n$ i suoi autovalori (che per il Teorema 1.2 sono reali). Siano $\pi(A)$, $\nu(A)$, $\zeta(A)$ rispettivamente il numero di autovalori positivi, negativi, e nulli di A. Allora per ogni matrice non singolare X valgono*

$$\pi(X^H AX) = \pi(A);$$

$$\nu(X^H AX) = \nu(A);$$

$$\zeta(X^H AX) = \zeta(A).$$

Teorema 1.4 (Birkhoff). *L'insieme di tutte le matrici doppiamente stocastiche di ordine n è l'inviluppo convesso di tutte le matrici di permutazione di ordine n, ossia data S una matrice doppiamente stocastica, si può scrivere*

$$S = \sum_{i=1}^{n!} \sigma_i P_i, \qquad \sum_{i=1}^{n!} \sigma_i = 1, \quad \sigma_i \geq 0, \quad i = 1, \ldots, n!,$$

dove ogni P_i è una matrice di permutazione.

Enunciamo anche un risultato particolarmente di rilievo per gran parte dell'Analisi Numerica, che dà un metodo per il calcolo dei vari autovalori di una matrice Hermitiana. Lo riprenderemo in seguito, al Capitolo 2.

Teorema 1.5 (Courant-Fisher). *Sia $A \in \mathbb{C}^{n\times n}$ una matrice Hermitiana, con autovalori $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. Allora si ha*

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

o equivalentemente

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

o anche

$$\min_{\substack{U_k\leq\mathbb{C}^n \\ \dim(U_k)=k}} \max_{x\in U_k} \frac{x^H Ax}{x^H x} = \lambda_k.$$

Ora definiamo dei tipi particolari di decomposizioni di matrici, con una struttura qualsiasi.

Teorema 1.6. *Sia $A \in \mathbb{C}^{m\times n}$ e $q = \min\{m,n\}$. Allora esistono due matrici unitarie, $U \in \mathbb{C}^{m\times m}$, $V \in \mathbb{C}^{n\times n}$ e una matrice $\Sigma \in \mathbb{R}^{m\times n}$, con diagonale principale $(\sigma_1,\dots,\sigma_q)$, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_q \geq 0$ e nulla al di fuori di questa diagonale, tali che $A = U\Sigma V^H$. I numeri $\sigma_1,\dots,\sigma_q$ sono detti* ***valori singolari*** *di A.*

Definizione 1.7 (Decomposizione polare)**.** *Sia $A \in \mathbb{C}^{m\times n}$. Allora, se $m \geq n$, una decomposizione polare destra di A è data dalla scrittura $A = UP$, con $U \in \mathbb{C}^{m\times n}$ matrice con colonne ortonormali e $P \in \mathbb{C}^{n\times n}$ matrice semidefinita positiva. Se $m \leq n$ una decomposizione polare sinistra di A è data dalla scrittura $A = PU$, con $P \in \mathbb{C}^{m\times m}$ matrice semidefinita positiva e $U \in \mathbb{C}^{m\times n}$ matrice con righe ortonormali.*

Teorema 1.8. *Sia $A \in \mathbb{C}^{m\times n}$, con $m \leq n$ ($m \geq n$). Allora esiste una decomposizione polare sinistra (destra) per A. La matrice P è unica ed è uguale a $(AA^H)^{1/2}$ ($(A^HA)^{1/2}$), mentre la matrice U è univocamente determinata se e solo se A è non singolare.*

È d'obbligo una precisazione sul Teorema appena formulato. La matrice AA^H è Hermitiana, quindi può essere decomposta tramite il Teorema spettrale nel prodotto $X\Lambda X^H$, dove Λ è una matrice diagonale di autovalori reali. Inoltre AA^H è semidefinita positiva, e quindi Λ avrà solo elementi non negativi sulla sua diagonale. Per $(AA^H)^{1/2}$ intenderemo quindi una matrice che si decompone in XDX^H, dove D è una matrice diagonale che ha le radici quadrate reali degli elementi di Λ sulla diagonale. Analogamente interpreteremo $(A^HA)^{1/2}$.

Definizione 1.9 (Cofattore)**.** *Sia $A \in \mathbb{C}^{n\times n}$ una matrice quadrata. Allora definiamo il cofattore dell'elemento di posto (i,j) di A come la quantità*

$$C_{i,j} := (-1)^{i+j}\det(M_{i,j}),$$

dove $M_{i,j}$ è la matrice ottenuta da A rimuovendo la sua i-sima riga e j-sima colonna.

Possiamo ora definire delle particolari norme, che ci serviranno nella Sezione 2.2.

Definizione 1.10 (Norma di Ky Fan)**.** *Sia $A \in \mathbb{C}^{m\times n}$ una matrice di valori singolari $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_q \geq 0$. con $q = \min\{m,n\}$. Fissato $\nu \in \{1,\dots,q\}$ La ν-sima norma di Ky Fan di A è definita come*

$$\|A\|_\nu := \sum_{k=1}^{\nu} \sigma_k.$$

Definizione 1.11 (Norma invariante per trasformazioni unitarie)**.** *Sia $\|\cdot\|$ una norma matriciale nello spazio delle matrici $\mathbb{C}^{m\times n}$. Se accade*

$$\|UAV\| = \|A\| \quad \forall A \in \mathbb{C}^{m\times n}$$

per ogni scelta di $U \in \mathbb{C}^{m\times m}$ e $V \in \mathbb{C}^{n\times n}$ matrici unitarie, allora diremo che la norma $\|\cdot\|$ è invariante per trasformazioni unitarie.

Definizione 1.12 (Norma matriciale indotta)**.** *Siano $\|\cdot\|_1$ una norma sullo spazio vettoriale $\mathbb{C}^n$ e $\|\cdot\|_2$ una norma sullo spazio vettoriale $\mathbb{C}^m$. Allora definiamo la norma indotta da $\|\cdot\|_1$ e $\|\cdot\|_2$ sullo spazio delle matrici $\mathbb{C}^{m\times n}$ come la funzione*

$$A \mapsto \|A\| = \max_{0\neq x\in\mathbb{C}^n} \frac{\|Ax\|_2}{\|x\|_1}.$$

Definizione 1.13 (Norma di Frobenius)**.** *Sia $A \in \mathbb{C}^{m\times n}$ una matrice elementi $A_{i,j}$. Definiamo la norma di Frobenius di A come il numero*

$$\|A\|_F := \sqrt{\sum_{i,j} |A_{i,j}|^2}.$$

Lemma 1.14. *Sia $A \in \mathbb{C}^{m\times n}$ una matrice qualsiasi. Allora $\|A\| \leq \|A\|_F$ per ogni norma matriciale indotta $\|\cdot\|$. Inoltre tutte le norme matriciali indotte e la norma di Frobenius sono invarianti per trasformazioni unitarie.*

Diamo ora un risultato importante, che permetterà di semplificare molte dimostrazioni nel Capitolo 2, e alcuni risultati generali sulle norme matriciali, che ci serviranno nel Capitolo 3.

Teorema 1.15 (Ky Fan)**.** *Siano A, B due matrici in $\mathbb{C}^{m\times n}$. Allora vale la disuguaglianza $\|A\|_\nu \leq \|B\|_\nu$ per ogni $\nu = 1, \ldots, q$ se e solo se vale la disuguaglianza $\|A\| \leq \|B\|$ per ogni norma matriciale in $\mathbb{C}^{m\times n}$ invariante per trasformazioni unitarie.*

Proposizione 1.16. *Sia $A \in \mathbb{C}^{n\times n}$ una matrice normale, e $\|\cdot\|_2$ la norma matriciale indotta dalla norma vettoriale euclidea. Allora, se $\lambda_1, \ldots, \lambda_n$ sono gli autovalori di A, si ha $\|A\|_2 = \max_i |\lambda_i|$.*

Ci saranno poi utili due risultati fondamentali di Analisi Complessa, per lo studio dei polinomi caratteristici.

Teorema 1.17 (Funzioni implicite)**.** *Sia $f : U \to \mathbb{C}^m$ una funzione analitica, con U un intorno del punto $(x_0, y_0) \in \mathbb{C}^n$, dove $n \geq m$ e denotiamo i vettori di $\mathbb{C}^n$ come (x, y),*

$x \in \mathbb{C}^{n-m}, y \in \mathbb{C}^m$. *Supponiamo* $f(x_0, y_0) = 0$ *e* $\det\left(\frac{\partial f}{\partial y}\right)\big|_{(x_0,y_0)} \neq 0$, *dove con* $\frac{\partial f}{\partial y}\big|_{(x_0,y_0)}$ *abbiamo indicato la sottomatrice*

$$\begin{pmatrix} \frac{\partial f_1}{\partial y_1}(x_0, y_0) & \cdots & \frac{\partial f_1}{\partial y_m}(x_0, y_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial y_1}(x_0, y_0) & \cdots & \frac{\partial f_m}{\partial y_m}(x_0, y_0) \end{pmatrix}$$

della matrice Jacobiana di f *calcolata in* (x_0, y_0). *Allora esiste* $\varepsilon > 0$ *e palle aperte* $B(x_0, \varepsilon) = \{x : \|x - x_0\| < \varepsilon\}$, $B(y_0, \varepsilon) = \{y : \|y - y_0\| < \varepsilon\}$ *tali che il prodotto* $B(x_0, \varepsilon) \times B(y_0, \varepsilon)$ *è contenuto in* U, *ed esiste una funzione* $g : B(x_0, \varepsilon) \to B(y_0, \varepsilon)$, *analitica, per cui vale*

$$g(x_0) = y_0$$

e

$$\{(x, y) \in B(x_0, \varepsilon) \times B(y_0, \varepsilon) : f(x, y) = 0\} = \{(x, g(x)) : x \in B(x_0, \varepsilon)\}.$$

Teorema 1.18 (Rouché)**.** *Siano* $f, g : U \to \mathbb{C}$ *due funzioni analitiche su* $U \subset \mathbb{C}^n$, *dove* U *è una regione limitata e semplicemente connessa di* $\mathbb{C}^n$ *che ha come bordo una curva semplice* ∂U. *Se accade* $|g(z)| < |f(z)|$ *per ogni* $z \in \partial U$, *allora le funzioni* f *e* $f + g$ *hanno lo stesso numero di zeri in* $\operatorname{int} U$, *contati con la loro molteplicità.*

In particolare da questo ultimo Teorema segue subito una dipendenza continua delle radici di un polinomio dai suoi coefficienti, come espresso nel Teorema:

Teorema 1.19. *Sia* $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$ *un polinomio monico di grado* n, $n \geq 1$ *di cui denotiamo le* m *radici distinte come* $\lambda_1 \ldots, \lambda_m$, *con rispettive molteplicità* $\alpha_1, \ldots, \alpha_m$. *Allora, se* $\varepsilon > 0$ *è tale che i dischi* $\overline{B(\lambda_i, \varepsilon)}$, $i = 1, \ldots, m$ *sono tutti disgiunti, esiste un* $\delta = \delta(\varepsilon) > 0$ *tale che per ogni polinomio monico* q *di grado* n, $q(z) = z^n + b_{n-1}z^{n-1} + \cdots + b_1 z + b_0$, *con*

$$|a_i - b_i| < \delta, \quad i = 1, \ldots, m,$$

q *ha esattamente* α_i *radici, contate con molteplicità, in* $B(\lambda_i, \varepsilon)$, *con* $i = 1, \ldots, m$.

Enunciamo infine un Teorema classico di teoria perturbativa matriciale.

Teorema 1.20 (Bauer-Fike)**.** *Sia* $Q \in \mathbb{C}^{n \times n}$ *una matrice non singolare,* A *una matrice,* $\|\cdot\|$ *una norma matriciale invariante per trasformazioni unitarie e* $\tilde{A} = A + E$, *dove* E *è una matrice con* $\|E\|$ *piccola. Se* $\tilde{\lambda}$ *è un autovalore di* $\tilde{A}$ *ma non di* A *allora si ha*

$$\left\|Q(A - \tilde{\lambda}I)^{-1}Q^{-1}\right\|^{-1} \leq \left\|QEQ^{-1}\right\|.$$

Questo Teorema dà una prima limitazione sulla distanza degli autovalori perturbati dallo spettro di A, ma come vedremo meglio nel Capitolo 3 questa risulta spesso esageratamente larga, perché prende in considerazione l'intero spettro della matrice e lo maggiora con una quantità che dipende anche dalla matrice Q. Il nostro approccio, nei Capitoli che seguono, sarà quando possibile di prendere in considerazione le singole autocoppie, in modo da avere risultati più precisi sulla perturbazione di ognuna di esse.

Aggiungiamo infine una precisazione notazionale: per tutti i Capitoli seguenti, A indicherà la matrice di partenza non perturbata, mentre $\tilde{A}$ sarà il risultato della perturbazione. In generale scriveremo $\tilde{A} = A + E$, assumendo che E sia una matrice di norma sufficientemente piccola, oppure $\tilde{A} = A + \varepsilon E$, non ponendo in questo caso alcuna condizione su E, ma chiedendo che $\varepsilon > 0$ sia un numero reale piccolo. Questa seconda scrittura ci sarà utile quando vorremo esplicitare la norma della perturbazione e controllarla con uno scalare.

Capitolo 2

Matrici Hermitiane

Possiamo ora addentrarci nello studio della teoria perturbativa matriciale, trovando delle stime delle variazioni di autovalori e autovettori della matrice che viene perturbata. Partiamo dal caso migliore che possa capitare: quello in cui la matrice di partenza sia Hermitiana, e venga perturbata da una matrice Hermitiana. Questo ci permetterà di scartare situazioni "patologiche" e ottenere dei risultati ottimali. Per tutto questo Capitolo, A, $\tilde{A}$ e E denoteranno matrici Hermitiane, e quindi varrà per loro il Teorema 1.2. Inoltre, assumeremo $\tilde{A} = A+E$, e denoteremo con $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ gli autovalori di A. Nello studio del problema cercheremo innanzitutto una stima abbastanza precisa di dove si trovino gli autovalori perturbati, per poi sviluppare una teoria basata su rotazioni di sottospazi vettoriali per valutare la variazione degli autovettori.

2.1 Autovalori: la disuguaglianza di Weyl

In questa circostanza riusciamo ad ottenere un intervallo di appartenenza degli autovalori di $\tilde{A}$ che dipenderà solo dagli spettri di A e di E, noto come disuguaglianza di Weyl. Iniziamo dimostrando un risultato fondamentale, di "interlacing", sullo spettro delle sottomatrici principali di A.

Teorema 2.1 (Cauchy)**.** *Siano A Hermitiana di ordine n e B una sua sottomatrice principale di ordine $n-k$. Allora, se $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_{n-k}$ sono gli autovalori di B, si ha*

$$\lambda_i \geq \mu_i \geq \lambda_{i+k}, \qquad i = 1, \ldots, n-k.$$

Dimostrazione. Partiamo dal caso $k = 1$. Senza perdere di generalità supponiamo che B sia la sottomatrice principale di testa di A (se così non fosse ci basterebbe considerare

PAP^H, con P matrice di permutazione opportuna). Allora avremo

$$A = \begin{pmatrix} B & a \\ a^H & \alpha \end{pmatrix}.$$

Se per assurdo la tesi fosse falsa, allora potremmo trovare un i per cui $\mu_i > \lambda_i$ oppure $\lambda_{i+1} > \mu_i$. Consideriamo il più piccolo indice per cui ciò accade. Nel caso sia $\mu_i > \lambda_i$ scegliamo τ con $\mu_i > \tau > \lambda_i$ e $\tau \notin \mathrm{Spec}(B)$. Allora si ha

$$\begin{aligned} H &:= \begin{pmatrix} B - \tau I & 0 \\ 0 & \alpha - \tau - a^H (B - \tau I)^{-1} a \end{pmatrix} \\ &= \begin{pmatrix} I & 0 \\ -a^H (B - \tau I)^{-1} & 1 \end{pmatrix} \begin{pmatrix} B - \tau I & a \\ a^H & \alpha - \tau \end{pmatrix} \begin{pmatrix} I & -(B - \tau I)^{-1} a \\ 0 & 1 \end{pmatrix} \\ &= X(A - \tau I)X^H \end{aligned}$$

Allora per il Teorema 1.3 di Sylvester-Jacobi H ha lo stesso numero di autovalori positivi di $A - \tau I$, cioè $i - 1$. Ma per definizione H deve avere almeno i autovalori positivi, quelli di $B - \tau I$. Abbiamo quindi trovato una contraddizione. Il caso $\lambda_{i+1} > \mu_i$ si sviluppa in modo totalmente analogo, e porta allo stesso assurdo.

Ora, se C è una sottomatrice principale di ordine $n-2$ con autovalori $\nu_1 \geq \cdots \geq \nu_{n-2}$, possiamo considerarla come una sottomatrice di ordine $(n-1)-1$ di una matrice B come nel caso precedente, e quindi concludere che

$$\mu_i \geq \nu_i \geq \mu_{i+1} \quad \Rightarrow \quad \lambda_i \geq \nu_i \geq \lambda_{i+2}, \qquad i = 1, \ldots, n-1.$$

Abbiamo quindi dimostrato il Teorema per $k = 2$. Iterando questo ragionamento per k generico si conclude la dimostrazione. □

Questo risultato ci permette di ricavare una interessante proprietà per gli autovalori di A, che si è rivelata molto versatile nello studio delle caratteristiche delle matrici Hermitiane.

Teorema 2.2 (Wielandt)**.** *Sia $A \in \mathbb{C}^{n \times n}$ una matrice Hermitiana. Allora, dati $1 \leq i_1 < i_2 < \cdots < i_k \leq n$, si ha*

$$\lambda_{i_1} + \lambda_{i_2} + \cdots + \lambda_{i_k} = \max_{\substack{\mathcal{X}_{i_1} < \mathcal{X}_{i_2} < \cdots < \mathcal{X}_{i_k} \leq \mathbb{C}^n \\ \dim(\mathcal{X}_{i_j}) = i_j}} \min_{\substack{X = (x_{i_1}, x_{i_2}, \ldots, x_{i_k}),\ x_{i_j} \in \mathcal{X}_{i_j} \\ X^H X = I}} \mathrm{tr}(X^H A X)$$

Dimostrazione. Iniziamo dimostrando che esiste una successione di sottospazi vettoriali $\mathcal{X}_{i_1} < \mathcal{X}_{i_2} < \cdots < \mathcal{X}_{i_k} \leq \mathbb{C}^n$ tale che, comunque presa $X = (x_{i_1}, x_{i_2}, \ldots, x_{i_k})$, con $x_{i_j} \in \mathcal{X}_{i_j}$ per ogni j, si abbia $\mathrm{tr}(X^H A X) \geq \sum_{j=1}^{k} \lambda_{i_j}$. Questo si ottiene considerando $\mathcal{X}_{i_j} = \mathrm{span}\{v_1, v_2 \ldots, v_{i_j}\}$, con v_j autovettore per λ_j (ricordiamo che nella nostra

notazione ogni λ_i è collegato a un solo autovettore, dato che autovalori di molteplicità m compaiono in m copie nella scrittura $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$). Con questa scelta di sottospazi si ha che, preso un vettore x_{i_j} in $\mathcal{X}_{i_j}$ con $x_{i_j}^H x_{i_j} = 1$, si avrà $x_{i_j}^H A x_{i_j} \geq \lambda_{i_j}$, da cui $\mathrm{tr}(X^H AX) = \sum_{j=1}^k x_{i_j}^H A x_{i_j} \geq \sum_{j=1}^k \lambda_{i_j}$.

Rimane ora da dimostrare

$$\lambda_{i_1} + \lambda_{i_2} + \ldots + \lambda_{i_k} \geq \max_{\substack{\mathcal{X}_{i_1} < \mathcal{X}_{i_2} < \cdots < \mathcal{X}_{i_k} \leq \mathbb{C}^n \\ \dim(\mathcal{X}_{i_j}) = i_j}} \min_{\substack{X = (x_{i_1}, x_{i_2}, \ldots, x_{i_k}),\, x_{i_j} \in \mathcal{X}_{i_j} \\ X^H X = I}} \mathrm{tr}(X^H AX).$$

Dimostriamolo per induzione su n.

Nel caso $n = 1$ si ha $k = n$, e il teorema risulta essere banalmente vero in quanto $X^H AX$ è una matrice simile ad A, e come tale ne conserva la traccia. Sia allora $n > 1$ e $k < n$. Sia $\mathcal{X}_{i_1} < \mathcal{X}_{i_2} < \cdots < \mathcal{X}_{i_k} \leq \mathbb{C}^n$ una catena di sottospazi vettoriali qualsiasi, con $\dim(\mathcal{X}_{i_j}) = i_j$. Cerchiamo una matrice con colonne ortonormali $X = (x_{i_1}, x_{i_2}, \ldots, x_{i_k})$, $x_{i_j} \in \mathcal{X}_{i_j}$ per ogni j, tale che $\mathrm{tr}(X^H AX) \leq \lambda_{i_1} + \lambda_{i_2} + \ldots + \lambda_{i_k}$.

Caso 1: $i_k < n$

Scegliamo un sottospazio $\hat{\mathcal{X}}_{n-1}$ di dimensione $n-1$ che contiene $\mathcal{X}_{i_k}$. Definiamo ora una matrice con colonne ortonormali, $Z = (z_1, \ldots, z_{n-1})$, tale che $\{z_1, \ldots, z_{i_j}\}$ sia una base per $\mathcal{X}_{i_j}$, e $\{z_1, \ldots, z_{n-1}\}$ sia una base per $\hat{\mathcal{X}}_{n-1}$. Definiamo $B := Z^H AZ$. Allora B è una sottomatrice principale di una matrice simile ad A, e quindi per il Teorema 2.1, chiamando i suoi autovalori $\mu_1 \geq \ldots \geq \mu_{n-1}$ si ha

$$\mu_i \leq \lambda_i, \qquad i = 1, \ldots, n-1.$$

Definiamo ora, per ogni j,

$$\mathcal{Y}_{i_j} := \{Z^H x \,:\, x \in \mathcal{X}_{i_j}\}.$$

Per come è stata definita la matrice Z, si ha che preso y in $\mathcal{Y}_{i_j}$, $x = Zy$ è in $\mathcal{X}_{i_j}$. Inoltre, $y^H By = x^H Ax$. Possiamo quindi applicare l'ipotesi induttiva a B, e trovare una matrice di colonne ortonormali $Y = (y_{i_1}, \ldots, y_{i_k})$, con $y_{i_j} \in \mathcal{Y}_{i_j}$ per ogni $j = 1, \ldots, k$, tale che

$$\mathrm{tr}(Y^H BY) \leq \sum_{j=1}^k \mu_{i_j}.$$

Definendo $X = ZY$ otterremo dunque

$$\mathrm{tr}(X^H AX) = \mathrm{tr}(Y^H BY) \leq \sum_{j=1}^k \mu_{i_j} \leq \sum_{j=1}^k \lambda_{i_j},$$

che era ciò che si voleva.

Caso 2: $i_k = n$

Siccome abbiamo supposto $k < n$ esisterà un l tale che $i_l + 1 < i_{l+1}$. Consideriamo allora $\mathcal{X}_{i_l} < \mathcal{X}_{i_{l+1}}$. Presa $\{v_1, \dots, v_{i_l}\}$ una base di $\mathcal{X}_{i_l}$ essa potrà essere completata a una base di $\mathcal{X}_{i_{l+1}}$, con l'aggiunta di almeno due vettori. Sia quindi $\{v_1, \dots, v_{i_l}, w_{i_l+1}, \dots, w_{i_{l+1}}\}$ una base per $\mathcal{X}_{i_{l+1}}$ ottenuta in questo modo. Completiamola a una base di $\mathbb{C}^n$, $\mathcal{B}$, e consideriamo lo spazio vettoriale $\hat{\mathcal{X}}_{n-1} := \operatorname{span}(\mathcal{B} \setminus \{w_{i_l+1}\})$. Consideriamo ora la catena di sottospazi

$$\mathcal{X}_{i_1} < \cdots < \mathcal{X}_{i_l} < \mathcal{X}_{i_{l+1}} \cap \hat{\mathcal{X}}_{n-1} < \mathcal{X}_{i_{l+2}} \cap \hat{\mathcal{X}}_{n-1} < \cdots < \hat{\mathcal{X}}_{n-1}.$$

Questa catena può essere riscritta come

$$\mathcal{V}_{i_1} < \mathcal{V}_{i_2} < \cdots < \mathcal{V}_{i_l} < \mathcal{V}_{i_{l+1}-1} < \cdots < \mathcal{V}_{i_k-1},$$

dove

$$\mathcal{V}_p = \begin{cases} \mathcal{X}_p & \text{se } p \leq i_l \\ \mathcal{X}_{p+1} \cap \hat{\mathcal{X}}_{n-1} & \text{se } p > i_l \end{cases}.$$

Per la scelta di $\hat{\mathcal{X}}_{n-1}$ questa è una catena consistente con l'enunciato del teorema, e rientra nel caso 1 già analizzato. È quindi possibile trovare $X = (x_1, \dots, x_k)$, con $x_j \in \mathcal{V}_{i_j} = \mathcal{X}_{i_j}$ se $j \leq l$ e $x_j \in \mathcal{V}_{i_j-1} \subset \mathcal{X}_{i_j}$ se $j > l$, tale che X abbia colonne ortonormali e valga

$$\operatorname{tr}(X^H A X) \leq \sum_{j=1}^{k} \lambda_{i_j}.$$

Questo dimostra la disuguaglianza voluta.

Rimane infine da giustificare la scelta della terminologia "minimo" e "massimo" (invece di "estremo superiore" ed "estremo inferiore") nell'enunciato del teorema. Ci serve quindi trovare una specifica scelta di una catena di sottospazi, e una particolare matrice X con le proprietà elencate nell'enunciato, tale per cui entrambe le disuguaglianze sono raggiunte. Ma questa scelta è già stata fatta: la catena di sottospazi sarà quella illustrata all'inizio della dimostrazione, e la matrice X sarà quella ottenuta tramite la costruzione della seconda parte della dimostrazione. □

Il Teorema di Wielandt appena dimostrato ha applicazioni non solo alla teoria perturbativa che stiamo studiando, ma in generale in molte aree dell'Algebra Lineare Numerica. Fissando $k = 1$, per esempio, si ricava immediatamente dall'enunciato il Teorema 1.5 di min-max di Courant-Fischer, uno strumento di base non solo per la teoria perturbativa, ma anche per la risoluzione dei problemi agli autovalori e per la creazione di algoritmi di analisi dati. Per quanto riguarda la nostra trattazione, il teorema si rivela essere fondamentale per fissare un intervallo entro cui ogni autovalore di A possa essere perturbato.

La stima riguarda il singolo autovalore, ed evita quindi di essere inutilmente pessimista, come invece accade per le stime sulla variazione della norma matriciale.

Teorema 2.3. *Siano A una matrice Hermitiana e $\tilde{A} = A + E$, con E matrice di perturbazione Hermitiana di autovalori $\epsilon_1 \geq \epsilon_2 \geq \cdots \geq \epsilon_n$. Siano $1 \leq i_1 < \cdots < i_k \leq n$ interi distinti. Allora*

$$\lambda_{i_1} + \cdots + \lambda_{i_k} + \epsilon_{n-k+1} + \cdots + \epsilon_n \leq \tilde{\lambda}_{i_1} + \cdots + \tilde{\lambda}_{i_k} \leq \lambda_{i_1} + \cdots + \lambda_{i_k} + \epsilon_1 + \cdots + \epsilon_k.$$

Dimostrazione. Poiché $\tilde{A}$ è una matrice Hermitiana, sia $\mathcal{X}_{i_1} < \mathcal{X}_{i_2} < \cdots < \mathcal{X}_{i_k} \leq \mathbb{C}^n$ una catena di sottospazi per cui vale

$$\sum_{j=1}^{k} \tilde{\lambda}_{i_j} = \min_{\substack{X=(x_{i_1}, x_{i_2}, \ldots, x_{i_k}),\, x_{i_j} \in \mathcal{X}_{i_j} \\ X^H X = I}} \operatorname{tr}(X^H \tilde{A} X),$$

che sappiamo esistere grazie al Teorema 2.2. Sfruttando lo stesso Teorema, sia poi $X = (x_{i_1}, \ldots, x_{i_k})$, con $x_{i_j} \in \mathcal{X}_{i_j}$ per ogni j per cui vale

$$\sum_{j=1}^{k} \lambda_{i_j} \geq \operatorname{tr}(X^H A X).$$

Allora otteniamo subito

$$\tilde{\lambda}_{i_1} + \cdots + \tilde{\lambda}_{i_k} \leq \operatorname{tr}[X^H (A + E) X] \leq \lambda_{i_1} + \cdots + \lambda_{i_k} + \operatorname{tr}(X^H E X).$$

Ma $X^H E X$ è una sottomatrice principale di una matrice simile a E, e quindi, se $\nu_1 \geq \cdots \geq \nu_k$ sono gli autovalori di $X^H E X$, possiamo applicare il Teorema 2.1 e ottenere

$$\epsilon_i \geq \nu_i, \qquad i = 1, \ldots, k.$$

Sommando lungo i otteniamo

$$\operatorname{tr}(X^H E X) \leq \epsilon_1 + \cdots + \epsilon_k,$$

da cui la seconda disuguaglianza del Teorema.

Per ottenere la prima disuguaglianza ci basterà riscrivere $A = \tilde{A} - E$ e procedere seguendo gli stessi passaggi del paragrafo precedente, scambiando i ruoli di A e $\tilde{A}$. L'unica differenza, in questo caso, è che la disuguaglianza che ci serve per maggiorare $-\operatorname{tr}(X^H E X)$ è

$$\nu_i \geq \epsilon_{i+(n-k)}, \qquad i = 1, \ldots, k,$$

che avevamo ricavato ancora una volta nel Teorema 2.1. Alla fine dei passaggi il risultato che si ottiene è

$$\lambda_{i_1} + \cdots + \lambda_{i_k} \leq \tilde{\lambda}_{i_1} + \cdots + \lambda_{i_k} - \epsilon_{n-k+1} - \cdots - \epsilon_n,$$

che conclude la dimostrazione. □

Il Teorema appena dimostrato si specializza immediatamente per ottenere la disuguaglianza che stavamo cercando. Basta infatti porre $k = 1$ e otterremo il seguente Corollario, dovuto a Weyl.

Corollario 2.4 (Weyl)**.** *Per $i = 1, \ldots, n$, nelle notazioni precedenti, si ha*

$$\tilde{\lambda}_i \in [\lambda_i + \epsilon_n, \lambda_i + \epsilon_1].$$

Spesso il risultato viene espresso come una disuguaglianza, che prende in esame gli spettri interi della matrici A e $\tilde{A}$, e li mette in relazione con la norma della perturbazione. Si ottiene quindi il seguente corollario, noto come disuguaglianza di Weyl.

Corollario 2.5 (Disuguaglianza di Weyl)**.** *Nelle notazioni precedenti si ha*

$$\max_{i \in \{1,\ldots,n\}} \{|\tilde{\lambda}_i - \lambda_i|\} \leq \|E\|_2$$

Il risultato, che si ottiene immediatamente ricordando $\|E\|_2 = \max\{|\epsilon_1|, |\epsilon_n|\}$, ci fornisce una stima abbastanza precisa sulla variazione complessiva dello spettro della matrice di partenza. Questa forma della disuguaglianza di Weyl è quindi più debole della precedente, ma risulta comunque essere la più usata e riconoscibile, dato che la stima dell'errore che ci fornisce dipende comunque solo dalla norma della perturbazione, che è ragionevole assumere piccola.

2.2 Autovettori: gli angoli di Davis-Kahan

Lo studio della perturbazione degli autovettori risulta sempre più delicato e complesso, questo soprattutto perché in presenza di un autovalore multiplo nella matrice di partenza gli autovettori non sono univocamente definiti. Ciò che rimane definito però è un sottospazio invariante per la matrice di partenza, vale a dire un sottospazio $\mathcal{X}$ di $\mathbb{C}^n$, che ha come base le colonne di una matrice $X = (x_1, \dots, x_k)$, per cui accade $Ax_i \in \mathcal{X}$, $i = 1, \dots, k$. Le somme dirette degli autospazi di una matrice saranno infatti sottospazi invarianti di A per definizione. Ha quindi senso generalizzare il problema dello studio della perturbazione degli autovettori di A allo studio della perturbazione dei suoi sottospazi invarianti, che è una questione più ampia, ma comunque più facilmente controllabile di quella di partenza. In questa Sezione lavoreremo sempre con norme matriciali indotte $\|\cdot\|$, che sappiamo essere invarianti per trasformazioni unitarie dal Lemma 1.14.

Sia quindi $\mathcal{X}$ un sottospazio di $\mathbb{C}^n$ invariante rispetto ad A. Possiamo rappresentarlo attraverso una proiezione lineare che agisce su $\mathbb{C}^n$, rappresentata tramite una matrice P. In questo modo ogni vettore $x \in \mathbb{C}^n$ può essere visto come una coppia di vettori (x_0, x_1), con $Px = x_0$ e $(I-P)x =: \tilde{P}x = x_1$. Sia ora k la dimensione di $\mathcal{X}$. Definiamo una matrice $G_0 \in \mathbb{C}^{n\times k}$ che come colonne ha una base ortonormale di $\mathcal{X}$. Abbiamo immediatamente $G_0(\mathbb{C}^k) = P(\mathbb{C}^n)$. Possiamo quindi definire un'altra matrice $G_1 \in \mathbb{C}^{n\times(n-k)}$ le cui colonne sono un completamento delle colonne di G_0 a una base ortonormale di $\mathbb{C}^n$. Anche in questo caso, $G_1(\mathbb{C}^{n-k}) = \tilde{P}(\mathbb{C}^n)$. Possiamo verificare immediatamente che valgono le relazioni

$$G_0 G_0^H = P$$

e

$$G_1 G_1^H = \tilde{P}.$$

Facendo ora passaggi analoghi per un sottospazio $\mathcal{Y}$, di dimensione h relativamente a $\mathbb{C}^n$ e invariante rispetto a $\tilde{A}$, definiamo le matrici di proiezione Q e $\tilde{Q}$, e le matrici con colonne ortonormali H_0 e H_1. Ciò che vorremmo riuscire a calcolare è di quanto si discosta $P(\mathbb{C}^n)$ da $Q(\mathbb{C}^n)$. Il metodo che utilizzeremo per fare questa operazione sarà quello di stabilire delle misure tra i due sottospazi analoghe a degli angoli (sono infatti dette "angoli canonici"), e analizzare la loro ampiezza. Questo metodo risulta essere fondato anche da un punto di vista geometrico: calcolare una misura angolare tra i due sottospazi significa per noi capire di quanto si devono ruotare i vettori delle colonne di G_0 per poter raggiungere i vettori delle colonne di H_0. Nel nostro caso, con matrici Hermitiane, e in cui consideriamo che $\mathcal{X}$ e $\mathcal{Y}$ siano autospazi, una base normale è data proprio dagli

autovettori che li generano. I risultati che daremo in questa Sezione stabiliscono delle maggiorazioni sulle ampiezze di queste rotazioni, applicandole poi al caso particolare in cui i sottospazi invarianti siano generati l'uno da autovettori di A, e l'altro dalle loro perturbazioni.

Cerchiamo di dare una formalizzazione matematica di quanto appena detto. Dal momento che stiamo cercando una trasformazione unitaria tra i due sottospazi possiamo dire che ciò che stiamo cercando è in realtà una matrice unitaria V tale che

$$VP = QV,$$

che ha come conseguenze dirette $V\tilde{P} = \tilde{Q}V$ e $VG_jG_j^H = H_jH_j^HV$. Notiamo immediatamente che perché questa V esista dobbiamo imporre che i due sottospazi abbiano la stessa dimensione, ovvero $k = h$. Per facilitare la trattazione descriviamo tutte le matrici su cui dobbiamo lavorare come delle matrici a blocchi, nelle coordinate (G_0, G_1). Avremo quindi

$$A = \begin{pmatrix} G_0 & G_1 \end{pmatrix} \begin{pmatrix} A_0 & 0 \\ 0 & A_1 \end{pmatrix} \begin{pmatrix} G_0^H \\ G_1^H \end{pmatrix}, \qquad P = \begin{pmatrix} G_0 & G_1 \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} G_0^H \\ G_1^H \end{pmatrix}. \tag{2.1}$$

Per quanto riguarda la matrice $\tilde{A} = A + E$ diamo la rappresentazione in entrambi i sistemi di coordinate (G_0, G_1) e (H_0, H_1), ottenendo così

$$\tilde{A} = \begin{pmatrix} G_0 & G_1 \end{pmatrix} \begin{pmatrix} A_0 + E_0 & B^H \\ B & A_1 + E_1 \end{pmatrix} \begin{pmatrix} G_0^H \\ G_1^H \end{pmatrix} = \begin{pmatrix} H_0 & H_1 \end{pmatrix} \begin{pmatrix} \Lambda_0 & 0 \\ 0 & \Lambda_1 \end{pmatrix} \begin{pmatrix} H_0^H \\ H_1^H \end{pmatrix}. \tag{2.2}$$

La rappresentazione a cui siamo più interessati tuttavia è quella di V, per cui cerchiamo una forma particolare che sarà alla base della nostra definizione di angolo canonico. Scriviamo

$$\begin{pmatrix} C_0 & -S_1 \\ S_0 & C_1 \end{pmatrix} = \begin{pmatrix} G_0^H \\ G_1^H \end{pmatrix} V \begin{pmatrix} G_0 & G_1 \end{pmatrix}$$

appunto mettendoci nell'ottica di ottenere una matrice simile a quelle di rotazione bidimensionali: $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

Per come sono definite, C_0 e C_1 risultano essere quadrate, mentre S_0 e S_1 rettangolari. Inoltre, dall'unitarietà di V, otteniamo

$$\begin{aligned} V^HV &= \begin{pmatrix} C_0^HC_0 + S_0^HS_0 & -C_0^HS_1 + S_0^HC_1 \\ -S_1^HC_0 + C_1^HS_0 & S_1^HS_1 + C_1^HC_1 \end{pmatrix} = VV^H = \begin{pmatrix} C_0C_0^H + S_1S_1^H & C_0S_0^H - S_1C_1^H \\ S_0C_0^H - C_1S_1^H & S_0S_0^H + C_1C_1^H \end{pmatrix} \\ &= \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}. \end{aligned} \tag{2.3}$$

Ci serve ora trovare una quantità che leghi i due sottospazi e sia indipendente da V, per poter definire gli angoli canonici tra i sottospazi. Fissato $j = 0, 1$, notiamo che possiamo scrivere

$$C_j C_j^H = G_j^H V G_j G_j^H V^H G_j = G_j^H H_j H_j^H V V^H G_j = G_j^H H_j H_j^H G_j,$$

$$S_j S_j^H = G_{1-j}^H V G_j G_j^H V^H G_{1-j} = G_{1-j}^H H_j H_j^H V V^H G_{1-j} = G_{1-j}^H H_j H_j^H G_{1-j}.$$

Queste matrici risultano indipendenti da V, e possiamo quindi definire i nostri angoli a partire da esse. Utilizzeremo i valori singolari di C_j, che sono le radici degli autovalori di $C_j C_j^H$. Inoltre, dato che le colonne di V sono ortonormali, tali valori singolari dovranno essere minori o uguali a 1. È quindi ben definita la matrice

$$\Theta_j = \arccos(C_j C_j^H)^{1/2},$$

dove con questa scrittura intendiamo che, effettuando la decomposizione spettrale $C_j C_j^H = X D X^H$, la matrice Θ_j è definita, nella base degli autovettori di $C_j C_j^H$, come una matrice diagonale, che sulla diagonale presenta gli arcocoseni delle radici degli elementi corrispondenti in D.

Abbiamo quindi effettivamente interpretato la matrice C_j come una matrice di coseni di angolo, e di conseguenza, potremo interpretare anche le S_j come dei seni di angolo. Supponiamo infatti S_0 alta e consideriamo la relazione $C_0^H C_0 + S_0^H S_0 = I$ che abbiamo trovato nell'equazione (2.3). Se X è una matrice unitaria per cui $C_0^H C_0 = X D X^H$, con D matrice diagonale, allora dovrà essere anche $S_0^H S_0 = X \Lambda X^H$, con Λ matrice diagonale. Otteniamo quindi che la somma degli autovalori di $C_0^H C_0$ (che sono uguali agli autovalori di $C_0 C_0^H$) con gli autovalori di $S_0^H S_0$ è sempre pari a 1. Quindi, se abbiamo interpretato i valori singolari di C_0 come coseni di angolo, sarà immediato interpretare i valori singolari di S_0 come seni degli stessi angoli. Se fosse S_0 larga allora la costruzione è la stessa, ma $S_0^H S_0$ avrà come autovalori, oltre ai quadrati dei valori singolari di S_0, anche degli zeri. In questo caso associeremo anche questi valori ai seni degli angoli di Θ_0. Analogamente varrà per C_1 e S_1. Non possiamo però scrivere $\sin \Theta_j = S_j$, perché la matrice Θ_j è quadrata mentre S_j non lo è. Ciò che possiamo scrivere è $\|\sin \Theta_j\| = \|S_j\|$. Le matrici S_j e $\sin \Theta_j$ hanno infatti gli stessi valori singolari non nulli per costruzione, e quindi le stesse norme di Ky Fan $\|\cdot\|_\nu$. Ci basta quindi applicare il Teorema 1.15 di Ky Fan per ottenere che l'uguaglianza è valida per ogni norma invariante per matrici unitarie. Abbiamo quindi definito i nostri angoli, che dipendono solo dai due sottospazi invarianti. Uniamoli nel seguente operatore:

$$\Theta := \begin{pmatrix} \Theta_0 & 0 \\ 0 & \Theta_1 \end{pmatrix}.$$

Diamo ora una proprietà per le matrici $\sin\Theta_0$ e $\sin\Theta_1$, che si rivelerà essere utile nei Teoremi di perturbazione, e che si verifica facilmente attraverso un calcolo diretto.

Osservazione 2.6. Per ogni norma invariante per trasformazioni unitarie, valgono

$$\left\|\tilde{Q}P\right\| = \|\sin\Theta_0\|,$$

$$\left\|\tilde{P}Q\right\| = \|\sin\Theta_1\|.$$

In questo momento, date le definizioni separate di Θ_0 e Θ_1, la matrice V sembrerebbe analoga a una matrice di funzioni goniometriche che dipendono da due angoli diversi. Vogliamo renderla ancora più simile a una matrice di rotazione bidimensionale. Diamo allora la seguente definizione

Definizione 2.7 (Rotazione Diretta)**.** *Una soluzione unitaria* $V = \begin{pmatrix} C_0 & -S_1 \\ S_0 & C_1 \end{pmatrix}$ *dell'equazione* $VP = QV$ *si dice una "rotazione diretta" tra gli spazi* $P(\mathbb{C}^n)$ *e* $Q(\mathbb{C}^n)$ *se vale:*

i. $C_0 \geq 0$

ii. $C_1 \geq 0$

iii. $S_1 = S_0^H$

Dove con $C_j \geq 0$ *intendiamo che la matrice è semidefinita positiva.*

Vale la seguente proposizione:

Proposizione 2.8. *Una rotazione diretta tra i due spazi invarianti* $P(\mathbb{C}^n)$ *e* $Q(\mathbb{C}^n)$ *esiste se e solo se si ha*

$$\dim P(\mathbb{C}^n) \cap \tilde{Q}(\mathbb{C}^n) = \dim \tilde{P}(\mathbb{C}^n) \cap Q(\mathbb{C}^n).$$

Nel seguito, assumeremo che questa proprietà valga, in modo da poter scegliere una rotazione diretta tra i sottospazi e ottenere risultati più significativi.

Dimostrazione. Dimostriamo solo una delle implicazioni della proposizione, quella che dà le condizioni per l'esistenza di una rotazione diretta. La condizione $\dim P(\mathbb{C}^n) \cap \tilde{Q}(\mathbb{C}^n) = \dim \tilde{P}(\mathbb{C}^n) \cap Q(\mathbb{C}^n)$ è equivalente a $\dim \mathrm{Ker}(C_0) = \dim \mathrm{Ker}(C_0^H)$. Sia infatti $x_0 \in \mathrm{Ker}(C_0)$, e consideriamo, in coordinate (G_0, G_1) il vettore $x = \begin{pmatrix} x_0 \\ 0 \end{pmatrix}$. Si ha che $x \in P(\mathbb{C}^n)$, e quindi per definizione $Vx \in Q(\mathbb{C}^n)$. Ma

$$Vx = \begin{pmatrix} C_0 x_0 \\ S_0 x_0 \end{pmatrix} = \begin{pmatrix} 0 \\ S_0 x_0 \end{pmatrix} \in \tilde{P}(\mathbb{C}^n),$$

da cui $Vx \in \tilde{P}(\mathbb{C}^n) \cap Q(\mathbb{C}^n)$. Preso poi un vettore y in tale intersezione, le proprietà di V portano subito a scrivere $V^{-1}y \in P(\mathbb{C}^n)$, e quindi $V^{-1}y$ si può scrivere nella forma $\begin{pmatrix} a \\ 0 \end{pmatrix}$ con $a \in \mathrm{Ker}(C_0)$. Possiamo quindi dire che $\mathrm{Ker}(C_0) = V^{-1}(\tilde{P}(\mathbb{C}^n) \cap Q(\mathbb{C}^n))$. Analogamente si dimostra $\mathrm{Ker}(C_0^H) = V^{-1}(P(\mathbb{C}^n) \cap \tilde{Q}(\mathbb{C}^n))$.

Effettuiamo ora la decomposizione polare della matrice C_0 ottenendo così (dal Teorema 1.8) $C_0 = (C_0 C_0^H)^{1/2} Z_0$. Se $\mathrm{Ker}(C_0) \neq 0$ allora Z_0 non è unicamente determinata, ma è una matrice unitaria (perché C_0 è quadrata). Inoltre, essa porterà $\mathrm{Ker}(C_0)$ in $\mathrm{Ker}(C_0^H)$ per costruzione. Sia ora $x_0 \in \mathrm{Ker}(C_0)$. Allora

$$C_1^H S_0 x_0 = G_1^H V^H G_1 G_1^H V G_0 x_0.$$

Poiché $C_0 x_0 = G_0^H V G_0 x_0 = 0$, allora $V G_0 x_0 = G_1 G_1^H (V G_0 x_0)$, quindi

$$G_1^H V^H (G_1 G_1^H V G_0 x_0) = G_1^H V^H V G_0 x_0 = G_1^H G_0 x_0 = 0.$$

Inoltre, sia $\{v_1, \dots, v_r\}$ una base ortonormale di $\mathrm{Ker}(C_0)$. Esprimendo i vettori di $\mathbb{C}^n$ nelle coordinate (G_0, G_1) otteniamo

$$S_0(v_1, \dots, v_r) = G_1^H V G_0 (v_1, \dots, v_r) = (0, I_{n-k})\, V \begin{pmatrix} I_k \\ 0 \end{pmatrix} (v_1, \dots, v_r) = (0, I_{n-k})\, V \begin{pmatrix} v_1 & \cdots & v_r \\ 0 & \cdots & 0 \end{pmatrix}.$$

Dato che V è unitaria, essa manderà basi ortonormali in basi ortonormali. Quindi, ricordando quanto abbiamo visto sopra, possiamo concludere

$$S_0(v_1, \dots, v_r) = (0, I_{n-k}) \begin{pmatrix} 0 & \cdots & 0 \\ w_1 & \cdots & w_r \end{pmatrix} = (w_1, \dots, w_r),$$

una r-upla di vettori ortonormali di $\mathbb{C}^{n-k}$. Abbiamo così dimostrato che S_0 manda basi ortonormali di $\mathrm{Ker}(C_0)$ in basi ortonormali di $\mathrm{Ker}(C_1^H)$. Analogamente si dimostra che S_1 manda basi ortonormali di $\mathrm{Ker}(C_1)$ in basi ortonormali di $\mathrm{Ker}(C_0^H)$. Decomponendo polarmente C_1 in $C_1 = (C_1 C_1^H)^{1/2} Z_1$ abbiamo ancora una volta una libertà di scelta per Z_1, relativamente a come essa agisca sui vettori di $\mathrm{Ker}(C_1)$. Scegliamo Z_1 in modo che in questo nucleo si comporti come $S_0 Z_0^{-1} S_1$, che per quello che abbiamo dimostrato è una scelta ammissibile che conserva l'unitarietà di Z_1, e manda $\mathrm{Ker}(C_1)$ in $\mathrm{Ker}(C_1^H)$. Definiamo allora l'operatore Z, che si esprime nelle coordinate (G_0, G_1) come $Z = \begin{pmatrix} Z_0 & 0 \\ 0 & Z_1 \end{pmatrix}$. Sia $U = VZ^{-1}$. U è unitaria perché prodotto di matrici unitarie, e soddisfa $UP = QU$, infatti sviluppando abbiamo

$$\begin{aligned} VZ^{-1}PZ = QV &\Leftrightarrow V(G_0 \quad G_1) \begin{pmatrix} Z_0^{-1} & 0 \\ 0 & Z_1^{-1} \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} Z_0 & 0 \\ 0 & Z_1 \end{pmatrix} \begin{pmatrix} G_0^H \\ G_1^H \end{pmatrix} = QV \\ &\Leftrightarrow VP = QV \end{aligned}.$$

Inoltre, per la scelta delle Z_j, i punti i. e ii. della definizione di rotazione diretta per U sono rispettati. Rimane da dimostrare la proprietà iii. Per semplicità di scrittura, rinominiamo le C_j e le S_j in modo che corrispondano alle componenti di U nelle coordinate (G_0, G_1). Si verifica immediatamente, tenendo conto della scelta effettuata per Z_1, che per ogni vettore $x \in P(\mathbb{C}^n) \cap \tilde{Q}(\mathbb{C}^n)$ o $x \in \tilde{P}(\mathbb{C}^n) \cap Q(\mathbb{C}^n)$, vale $U^2 x = -x$. Ora, da (2.3), vale

$$C_0 S_1 = S_0^H C_1 \qquad C_0 S_0^H = S_1 C_1, \tag{2.4}$$

da cui si ricava $C_0^2 S_1 = S_1 C_1^2$. Allora per ogni polinomio, e quindi per ogni funzione analitica f, si avrà $f(C_0^2) S_1 = S_1 f(C_1^2)$. Scegliendo f come la funzione radice quadrata avremo $C_0 S_1 = S_1 C_1$, che confrontato con la prima delle (2.4) ci dice che S_1 e S_0^H coincidono in $\mathrm{Im}(C_1)$. Sia dunque $x_1 \perp \mathrm{Im}(C_1)$, e $x = \begin{pmatrix} 0 \\ x_1 \end{pmatrix}$.

Abbiamo $x \in \tilde{P}(\mathbb{C}^n) \cap Q(\mathbb{C}^n)$, infatti $U^H x = \begin{pmatrix} S_0^H x_1 \\ 0 \end{pmatrix} \in P(\mathbb{C}^n)$. Ma allora per quanto detto sopra si ha $U^2 x = -x$, da cui $Ux = -U^H x$. Ma $Ux = \begin{pmatrix} -S_1 x_1 \\ 0 \end{pmatrix} \in P(\mathbb{C}^n)$, da cui S_0^H e S_1 coincidono in $\mathrm{Im}(C_1)^\perp$, e quindi coincidono ovunque. □

In questa circostanza, quindi, risulta evidente che $\sin\Theta_0$ e $\sin\Theta_1$ hanno gli stessi valori singolari.

Gli angoli canonici risultano così ben definiti, e la trasformazione unitaria tra i due sottospazi è quasi completamente equiparata ad una rotazione. Andiamo più nello specifico nel nostro problema, considerando i sottospazi invarianti come gli autospazi che vogliamo analizzare. Da ora quindi vedremo $\mathcal{X}$ come lo spazio generato da m autovettori di A (scelti in modo che siano ortogonali tra di loro e di norma unitaria). In questo modo, nelle rappresentazioni che abbiamo dato sopra, avremo che A_0 è una matrice diagonale, la cui diagonale contiene gli m autovalori corrispondenti ai generatori di $\mathcal{X}$; inoltre, l'operatore G_0 sarà rappresentabile come una matrice che ha come colonne gli m autovettori in questione. Per trovare la distanza in termini di autovalori e autovettori tra A e $\tilde{A}$ considereremo la matrice

$$R := \tilde{A} G_0 - G_0 A_0, \tag{2.5}$$

a cui diamo nome di "residuo", e la cui norma ci indica quanto le autocoppie di A che stiamo prendendo in considerazione si discostino dall'essere autocoppie di $\tilde{A}$. Dalla rappresentazione della perturbazione E si evince che $R = EG_0$.

Ci serve ora dare una proprietà importante per le norme matriciali indotte, ossia la compatibilità.

Lemma 2.9. *Sia $\|\cdot\|$ una norma matriciale indotta, definita su $\mathbb{C}^{m\times n}$. Allora prese $V \in \mathbb{C}^{n\times n}$ e $W \in \mathbb{C}^{m\times m}$ due contrazioni (cioè $\|Vx\| \leq \|x\|$ per ogni vettore $x \in \mathbb{C}^n$ e $\|Wy\| \leq \|y\|$ per ogni vettore $y \in \mathbb{C}^m$) si ha, per ogni matrice $K \in \mathbb{C}^{m\times n}$, che*

$$\|WKV\| \leq \|K\| .$$

Una norma con questa proprietà si dice "compatibile".

Dimostrazione. Sia $K \in \mathbb{C}^{m\times n}$. Allora $\|W(KVx)\| \leq \|KVx\|$ per ogni vettore $x \in \mathbb{C}^n$, da cui $\|WKV\| \leq \|KV\|$. Inoltre si ha

$$\max_{0\neq x\in\mathbb{C}^n} \frac{\|KVx\|}{\|x\|} \leq \|K\| \max_{0\neq x\in\mathbb{C}^n} \frac{\|Vx\|}{\|x\|} \leq \|K\| ,$$

da cui la tesi. □

Ora che abbiamo costruito tutto l'apparato necessario, siamo quasi pronti a dare la stima di perturbazione. Partiamo da alcuni risultati preliminari, per poi esporre il Teorema che esprime maggiorazioni sugli angoli canonici, e quindi sulle differenze tra autovettori.

Proposizione 2.10. *Siano $\mathcal{A}$ e $\mathcal{B}$ due spazi di Banach, e A e B due operatori lineari nei rispettivi spazi tali che $\|A\|_{\mathcal{A}} \leq \alpha$ e $\|B^{-1}\|_{\mathcal{B}} \leq (\alpha+\delta)^{-1}$ con $\alpha \geq 0$, $\delta > 0$. Allora sia $\|\cdot\|$ una norma indotta sullo spazio delle funzioni lineari da $\mathcal{A}$ a $\mathcal{B}$. Se si ha $BX - XA = C$ allora vale*

$$\|C\| \geq \delta \|X\| .$$

Dimostrazione. Per le proprietà delle norme indotte sugli spazi vettoriali, e applicando le disuguaglianze dell'enunciato si ha che $\|XA\| \leq \alpha \|X\|$ e

$$(\alpha+\delta)^{-1} \|BX\| \geq \left\|B^{-1}\right\|_{\mathcal{B}} \|BX\| \geq \|X\|$$

da cui $\|BX\| \geq (\alpha+\delta)\|X\|$. Per le disuguaglianze triangolari $\|BX\| - \|XA\| \leq \|C\|$, da cui si ottiene subito la tesi. □

Lemma 2.11. *Siano $\mathcal{P}$ e $\mathcal{Q}$ due matrici di proiezione lineare, con $\tilde{\mathcal{P}} = I - \mathcal{P}$ e $\tilde{\mathcal{Q}} = I - \mathcal{Q}$. Se $\|\mathcal{P}K\mathcal{Q}\| \leq \|\mathcal{P}L\mathcal{Q}\|$ e $\left\|\tilde{\mathcal{P}}K\tilde{\mathcal{Q}}\right\| \leq \left\|\tilde{\mathcal{P}}L\tilde{\mathcal{Q}}\right\|$, allora*

$$\left\|\mathcal{P}K\mathcal{Q} + \tilde{\mathcal{P}}K\tilde{\mathcal{Q}}\right\| \leq \left\|\mathcal{P}L\mathcal{Q} + \tilde{\mathcal{P}}L\tilde{\mathcal{Q}}\right\|$$

per ogni norma invariante per trasformazioni unitarie.

Dimostrazione. Per il Teorema 1.15 di Ky Fan è sufficiente dimostrare il risultato per tutte le norme $\|\cdot\|_\nu$ con $\nu = 1, \dots, n$, dove n è il numero di valori singolari degli operatori in questione (e dipende solo dalle dimensioni del dominio e del codominio). Fissiamo allora ν intero compreso tra 1 e n. Se $\sigma_1 \dots \sigma_n$ sono i valori singolari di $\mathcal{P}K\mathcal{Q}$, $\tilde{\sigma}_1, \dots, \tilde{\sigma}_n$ quelli di $\tilde{\mathcal{P}}K\tilde{\mathcal{Q}}$, $\tau_1, \dots, \tau_n$ quelli di $\mathcal{P}L\mathcal{Q}$, $\tilde{\tau}_1, \dots, \tilde{\tau}_n$ quelli di $\tilde{\mathcal{P}}L\tilde{\mathcal{Q}}$. Allora l'ipotesi dice che $\sigma_1 + \dots + \sigma_\nu \leq \tau_1 + \dots + \tau_\nu$ e $\tilde{\sigma}_1 + \dots + \tilde{\sigma}_\nu \leq \tilde{\tau}_1 + \dots + \tilde{\tau}_\nu$. Calcoliamo allora i valori singolari di $\mathcal{P}K\mathcal{Q} + \tilde{\mathcal{P}}K\tilde{\mathcal{Q}}$. Essi sono le radici degli autovalori di

$$(\mathcal{P}K\mathcal{Q} + \tilde{\mathcal{P}}K\tilde{\mathcal{Q}})^H(\mathcal{P}K\mathcal{Q} + \tilde{\mathcal{P}}K\tilde{\mathcal{Q}}) = \mathcal{Q}^H K^H \mathcal{P} K \mathcal{Q} + \tilde{\mathcal{Q}}^H K^H \tilde{\mathcal{P}} K \tilde{\mathcal{Q}}.$$

Gli autovalori di questi due addendi sono rispettivamente il quadrato dei valori singolari di $\mathcal{P}K\mathcal{Q}$ e di $\tilde{\mathcal{P}}K\tilde{\mathcal{Q}}$. Ora, siano x_1 e x_2 autovettori rispettivamente di $\mathcal{Q}^H K^H \mathcal{P} K \mathcal{Q}$ e di $\tilde{\mathcal{Q}}^H K^H \tilde{\mathcal{P}} K \tilde{\mathcal{Q}}$, di autovalori λ e $\tilde{\lambda}$, con $\mathcal{Q}x_1 \neq 0$, $\tilde{\mathcal{Q}}x_2 \neq 0$. Allora $x = \mathcal{Q}x_1$ e $y = \tilde{\mathcal{Q}}x_2$ sono autovettori della somma, di autovalori λ e $\tilde{\lambda}$. Se invece v fosse autovettore di $\mathcal{Q}^H K^H \mathcal{P} K \mathcal{Q}$ con $\mathcal{Q}v = 0$, allora $v = \tilde{\mathcal{Q}}v \neq 0$, e v risulta essere autovettore della somma solo se è autovettore di $\tilde{\mathcal{Q}}^H K^H \tilde{\mathcal{P}} K \tilde{\mathcal{Q}}$. Quindi lo spettro della matrice somma è dato dall'unione degli spettri delle matrici addende, ed è facile vedere che non vi sono ulteriori autovalori. I valori singolari di $\mathcal{P}K\mathcal{Q} + \tilde{\mathcal{P}}K\tilde{\mathcal{Q}}$ sono quindi l'unione dei valori singolari di $\mathcal{P}K\mathcal{Q}$ e $\tilde{\mathcal{P}}K\tilde{\mathcal{Q}}$, e quindi, grazie alle ipotesi dell'enunciato, si ottiene immediatamente che

$$\left\|\mathcal{P}K\mathcal{Q} + \tilde{\mathcal{P}}K\tilde{\mathcal{Q}}\right\| \leq \left\|\mathcal{P}L\mathcal{Q} + \tilde{\mathcal{P}}L\tilde{\mathcal{Q}}\right\|.$$

□

Lemma 2.12. *Siano $\mathcal{P}$ e $\mathcal{Q}$ due matrici di proiezione lineare. Allora per ogni norma invariante per matrici unitarie si ha*

$$\left\|\mathcal{P}K\mathcal{Q} + \tilde{\mathcal{P}}K\tilde{\mathcal{Q}}\right\| \leq \|K\|.$$

Dimostrazione. Si ha, ricordando che $(\mathcal{P} - \tilde{\mathcal{P}})$ e $(\mathcal{Q} - \tilde{\mathcal{Q}})$ sono matrici unitarie,

$$\begin{aligned} 2\left\|\mathcal{P}K\mathcal{Q} + \tilde{\mathcal{P}}K\tilde{\mathcal{Q}}\right\| &= 2\|\mathcal{P}K\mathcal{Q} + (I - \mathcal{P})K(I - \mathcal{Q})\| = \|4\mathcal{P}K\mathcal{Q} + 2K - 2\mathcal{P}K - 2K\mathcal{Q}\| \\ &= \|K + (2\mathcal{P} - I)K(2\mathcal{Q} - I)\| = \left\|K + (\mathcal{P} - \tilde{\mathcal{P}})K(\mathcal{Q} - \tilde{\mathcal{Q}})\right\| \\ &\leq \|K\| + \left\|(\mathcal{P} - \tilde{\mathcal{P}})K(\mathcal{Q} - \tilde{\mathcal{Q}})\right\| = 2\|K\| \end{aligned}$$

da cui la tesi. □

Siamo ora pronti per dare il Teorema sugli angoli canonici, dovuto a Davis e Kahan.

Teorema 2.13 (Davis-Kahan, Teorema del seno)**.** *Nelle notazioni in coordinate date precedentemente, siano $[\beta, \alpha]$ un intervallo reale, $\delta > 0$, A_0 il blocco definito nell'equazione 2.1 e Λ_1 il blocco definito nell'equazione 2.2. Supponiamo che lo spettro di A_0 sia*

compreso in $[\beta, \alpha]$ *e lo spettro di* Λ_1 *non intersechi l'intervallo* $]\beta - \delta, \alpha + \delta[$ *(oppure che lo spettro di* Λ_1 *sia compreso in* $[\beta, \alpha]$ *e lo spettro di* A_0 *non intersechi* $]\beta - \delta, \alpha + \delta[$*). Allora, per ogni norma invariante per trasformazioni unitarie, si ha*

$$\delta \left\| \sin \Theta_0 \right\| \leq \left\| R \right\|.$$

Dimostrazione. Aggiungendo un multiplo dell'identità a A otteniamo di traslare gli spettri di A_0 e Λ_1 senza influire su R. Possiamo quindi assumere che $0 \leq \alpha = -\beta$. Per la proprietà di compatibilità (dato che $\|H_1\|$=1) possiamo scrivere

$$\|R\| = \left\|R^H\right\| \geq \left\|R^H H_1\right\|.$$

Ricordando $R = (A + E)G_0 + G_0 A_0$, con A_0 diagonale, possiamo scrivere
$R^H H_1 = G_0^H H_1 \Lambda_1 - A_0 G_0^H H_1$. Allora, dato che $\left\|\Lambda_1^{-1}\right\| \leq (\alpha + \delta)^{-1}$ e $\|A_0\| \leq \alpha$ per le ipotesi sullo spettro, possiamo applicare la Proposizione 2.10 e scrivere

$$\left\|G_0^H E H_1\right\| = \left\|R^H H_1\right\| \geq \delta \left\|G_0^H H_1\right\|.$$

Ma $-S_1$=$G_0^H V G_1 = G_0^H H_1 H_1^H V G_1$, e si ha

$$(H_1^H V G_1)(H_1^H V G_1)^H = H_1^H V G_1 G_1^H V^H H_1 = H_1^H H_1 H_1^H V V^H H_1 = I$$

$$(H_1^H V G_1)^H (H_1^H V G_1) = G_1^H V^H H_1 H_1^H V G_1 = G_1^H V^H V G_1 G_1^H G_1 = I.$$

Quindi i valori singolari di $G_0^H H_1$ sono gli stessi di S_1, dal momento che le due matrici differiscono solo per la moltiplicazione con una matrice unitaria. Essi coincideranno quindi con quelli di $\sin \Theta_1$ che sono uguali a quelli di $\sin \Theta_0$ per la Proposizione 2.8. Otteniamo quindi

$$\|R\| \geq \left\|R^H H_1\right\| \geq \delta \left\|\sin \Theta_0\right\|,$$

che era quello che volevamo. □

Il teorema quindi ci dice che, a patto di scegliere un cluster di autovalori ravvicinati tra di loro e distanti dalle perturbazioni degli altri autovalori di A, la perturbazione dei relativi autovettori sarà interpretabile come una rotazione, l'ampiezza del cui angolo è limitata superiormente dal residuo. Ricordando che la norma di questo residuo dipende solo dalla matrice di perturbazione e da G_0 (che ha norma matriciale pari a 1), la perturbazione risulterà essere piccola, dando un grande valore di ottimalità per questa stima. Osserviamo come questo teorema funzioni meglio proprio nel caso in cui consideriamo un autovalore multiplo. In questo caso infatti lo spettro di A_0 sarà limitato a un punto, e quindi sarà molto facile separarlo dallo spettro di Λ_1 in accordo con l'enunciato del

Teorema. Il risultato risulta quindi di fondamentale importanza, applicandosi proprio al caso che di solito è più problematico nello studio della perturbazione degli autovettori.

Diamo ora una Corollario del Teorema, che ci permette di sfruttare il resto delle strutture di A e $\tilde{A}$.

Corollario 2.14 (Teorema del seno: versione simmetrica)**.** *Siano $\delta > 0$, $[\beta, \alpha]$ un intervallo, A_0 e A_1 i blocchi definiti nell'equazione 2.1 e Λ_0, Λ_1 i blocchi definiti nell'equazione 2.2. Assumiamo che gli spettri di A_0 e Λ_1 siano separati come nel Teorema 2.13, e che anche gli spettri di A_1 e Λ_0 siano separati allo stesso modo. Allora, per ogni norma invariante per trasformazioni unitarie,*

$$\delta \left\| \sin \Theta \right\| \leq \left\| E \right\| .$$

Dimostrazione. Applichiamo il Teorema 2.13 a A_0 e Λ_1. Otteniamo

$$\delta \left\| P\tilde{Q} \right\| = \delta \left\| \sin \Theta_0 \right\| = \delta \left\| G_0^H H_1 \right\| \leq \left\| G_0^H E H_1 \right\| = \left\| PE\tilde{Q} \right\| .$$

Riapplicando lo stesso Teorema a A_1 e Λ_0 otterremo invece

$$\delta \left\| \tilde{P}Q \right\| = \delta \left\| \sin \Theta_1 \right\| = \delta \left\| G_1^H H_0 \right\| \leq \left\| G_1^H E H_0 \right\| = \left\| \tilde{P}EQ \right\| .$$

Il risultato segue quindi applicando il Lemma 2.11 e il Lemma 2.12. □

Capitolo 3

Matrici Normali e Diagonalizzabili

Iniziamo ora ad aleggerire le condizioni poste sul problema iniziale, analizzando dei casi intermedi tra quello del Capitolo precedente, e la situazione generale. Nel fare ciò, quello che riusciremo a dire sulle perturbazioni di autovalori e autovettori della matrice A diventerà più incerto, e il problema inizierà a mostrare il suo malcondizionamento. Nella prima Sezione analizzeremo il caso in cui A sia una matrice normale, con $\tilde{A} = A + E$. In questo caso, anche se continua a valere il Teorema 1.2, gli autovalori di A potrebbero non essere reali, per cui sarà necessario individuare delle distanze appropriate tra di essi. Cercheremo di ottenere risultati simili al caso Hermitiano, aggiungendo occasionalmente condizioni sulle matrici E e $\tilde{A}$ per lavorare in questo senso. Passeremo poi alla teoria generale per matrici A diagonalizzabili, confrontando i risultati dati dal Teorema 1.20 di Bauer-Fike a quelli dati da Teoremi più specifici, che considerano le singole autocoppie della matrice. In questo secondo contesto, utilizzeremo la convenzione $\tilde{A} = A+\varepsilon E$, come indicata nel Capitolo 1.

3.1 Matrici Normali

Iniziamo lo studio della perturbazione delle matrici normali. Come accennato nel paragrafo precedente, dovremo definire delle distanze appropriate tra gli autovalori di A, in quanto essi potrebbero essere complessi. Definiamo tali norme per matrici generiche A e $\tilde{A}$.

Definizione 3.1. *Siano $A \in \mathbb{C}^{n\times n}$ e $\tilde{A} \in \mathbb{C}^{n\times n}$ due matrici qualsiasi, rispettivamente di autovalori $\lambda_1, \dots, \lambda_n$ e $\tilde{\lambda}_1, \dots, \tilde{\lambda}_n$. Definiamo:*

- *la **variazione spettrale di $\tilde{A}$ rispetto ad** A:*

$$\mathrm{sv}_A(\tilde{A}) := \max_i \min_j \left|\tilde{\lambda}_i - \lambda_j\right|;$$

- *la* ***distanza di Hausdorff*** *tra le due matrici:*

$$\mathrm{hd}(A, \tilde{A}) := \max\{\mathrm{sv}_A(\tilde{A}), \mathrm{sv}_{\tilde{A}}(A)\};$$

- *la* ***distanza di corrispondenza ottimale*** *tra le due matrici:*

$$\mathrm{md}_\infty(A, \tilde{A}) := \min_{\pi \in \mathfrak{S}_n} \left(\max_i \left| \tilde{\lambda}_{\pi(i)} - \lambda_i \right| \right),$$

  dove $\mathfrak{S}_n$ *è il gruppo simmetrico delle permutazioni di* n *elementi;*

- *la* ***distanza di corrispondenza euclidea*** *tra le due matrici:*

$$\mathrm{md}_2(A, \tilde{A}) := \min_{\pi \in \mathfrak{S}_n} \left(\sqrt{\sum_i \left| \tilde{\lambda}_{\pi(i)} - \lambda_i \right|^2} \right),$$

  con $\mathfrak{S}_n$ *definito come sopra.*

Cerchiamo di dare un'interpretazione geometrica di queste distanze. La variazione spettrale $\mathrm{sv}_A(\tilde{A})$ evidenzia quanto al massimo si discosta lo spettro di $\tilde{A}$ nel suo complesso dallo spettro di A. Si tratta di un parametro facilmente calcolabile (conoscendo gli autovalori), e può rivelarsi utile per dare un'interpretazione delle differenze tra gli autovalori delle due matrici, ma spesso risulta troppo ottimista nel caso della teoria perturbativa, poiché non calcola effettivamente la distanza degli autovalori perturbati da quelli originali, ma considera solamente gli insiemi $\mathrm{Spec}(A)$ e $\mathrm{Spec}(\tilde{A})$. La distanza di Hausdorff si comporta allo stesso modo, è necessario introdurla per dare alla variazione spettrale tutte le proprietà di una metrica. La distanza di corrispondenza ottimale cerca di risolvere i problemi della distanza di Hausdorff, provando a legare gli autovalori delle due matrici in coppie, dove ogni autovalore è presente in una e una sola coppia. Si tratta di un valore molto più complicato da calcolare (complessità dovuta al fatto che $\#\mathfrak{S}_n = n!$), ma descrive in modo molto più accurato le variazioni degli autovalori, dando come risultato la più grande distanza tra un autovalore perturbato e quello in coppia con esso (che è ragionevole supporre che sia il corrispondente autovalore non perturbato, quando E ha norma piccola). La distanza di corrispondenza euclidea si comporta allo stesso modo di quella ottimale, ma somma tutti i contributi dovuti alla perturbazione di ogni autovalore di A. Per questo risulta $\mathrm{sv}_A(\tilde{A}) \leq \mathrm{hd}(A, \tilde{A}) \leq \mathrm{md}_\infty(A, \tilde{A}) \leq \mathrm{md}_2(A, \tilde{A})$.

Applichiamo ora queste norme, che sono state definite per matrici generiche, al problema della perturbazione matriciale, e cerchiamo di trovare delle maggiorazioni. Ovviamente, date le considerazioni del paragrafo precedente, lavoreremo soprattutto per trovare delle limitazioni alle distanze di corrispondenza. Iniziamo dimostrando la Proposizione seguente.

Proposizione 3.2. *Sia A una matrice normale ed E una matrice qualsiasi, con $\tilde{A} = A + E$. Allora*

$$\mathrm{sv}_A(\tilde{A}) \leq \|E\|_2 .$$

Dimostrazione. Dal momento che A è normale sia X una matrice unitaria per cui vale $A = X\Lambda X^H$, con Λ matrice diagonale. Sia ora $\tilde{\lambda}$ un autovalore di $\tilde{A}$. Se $\tilde{A}$ è anche autovalore di A allora si ha $\min_j \left|\tilde{\lambda} - \lambda_j\right| = 0$. Se invece $\tilde{\lambda} \notin \mathrm{Spec}(A)$ allora possiamo utilizzare il Teorema 1.20 di Bauer-Fike, ponendo $Q = X^H$ e scegliendo come norma matriciale quella indotta dalla norma euclidea. Infatti così facendo otteniamo la disuguaglianza

$$\left\|X^H(A - \tilde{\lambda}I)^{-1}X\right\|_2^{-1} \leq \left\|X^H E X\right\|_2$$

e, dato che la norma matriciale euclidea è invariante per trasformazioni unitarie, avremo

$$\left\|(A - \tilde{\lambda}I)^{-1}\right\|_2^{-1} \leq \|E\|_2 .$$

Ora, dato che A è normale e $\tilde{\lambda}$ non è autovalore di A, anche $(A - \tilde{\lambda}I)^{-1}$ sarà normale, di autovalori $\frac{1}{\lambda_i - \tilde{\lambda}}$, $i = 1, \ldots, n$. Per la Proposizione 1.16 vale quindi

$$\left\|(A - \tilde{\lambda})^{-1}\right\|_2^{-1} = \frac{1}{\max_i \left|\frac{1}{\lambda_i - \tilde{\lambda}}\right|} = \min_i \left|\lambda_i - \tilde{\lambda}\right| .$$

Abbiamo così dimostrato che, in ogni caso, preso un qualsiasi autovalore di $\tilde{A}$ ci sarà un autovalore di A che dista da esso al più $\|E\|_2$, e quindi che $\mathrm{sv}_A(\tilde{A}) \leq \|E\|_2$. □

Il risultato ricorda molto la disuguaglianza di Weyl, ma limita solo la variazione spettrale, che abbiamo visto fornirci decisamente meno informazioni per quanto riguarda la perturbazione dei singoli autovalori. Il nostro scopo in questa Sezione sarà di espandere questa disuguaglianza alle distanze di corrispondenza, che ci permetterebbe di ottenere risultati totalmente analoghi a quelli del Capitolo precedente. Il Corollario 2.5 aveva come ipotesi che entrambe le matrici A ed E fossero Hermitiane. Questo ci porta a domandarci se richiedere che anche E sia normale, oltre a chiedere che lo sia A, possa portare a ottenere i risultati voluti. Purtroppo in questo caso ciò non aiuterebbe perché supporre che A ed E siano normali non implica che anche $\tilde{A}$ lo sarà, come invece avveniva per il caso Hermitiano. Quello che possiamo fare è supporre che A e $\tilde{A}$ siano normali, anche se questa condizione è decisamente meno limpida della precedente, perché non permette di descrivere chiaramente le limitazioni su E.

Osserviamo subito come supporre che anche $\tilde{A}$ sia normale permetta di espandere la Proposizione 3.2 nel seguente Corollario:

Corollario 3.3. *Siano A e $\tilde{A}$ due matrici normali, con $\tilde{A} = A + E$. Allora*

$$\mathrm{hd}(A, \tilde{A}) \leq \|E\|_2 .$$

Possiamo però ottenere molto di più, come si vede dal Teorema seguente.

Teorema 3.4 (Hoffman-Wielandt)**.** *Siano A e $\tilde{A}$ matrici normali, con $\tilde{A} = A + E$. Allora si ha*

$$\mathrm{md}_2(A, \tilde{A}) \leq \|E\|_F .$$

Dimostrazione. Dal momento che A e $\tilde{A}$ sono normali, siano U e V matrici unitarie per cui $A = U\Lambda U^H$ e $\tilde{A} = V\tilde{\Lambda}V^H$ sono matrici diagonali, con $\Lambda = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$ e $\tilde{\Lambda} = \mathrm{Diag}(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n)$. Poiché la norma di Frobenius è invariante per trasformazioni unitarie scriveremo

$$\|E\|_F = \left\|A - \tilde{A}\right\|_F = \left\|\Lambda W - W\tilde{\Lambda}\right\|_F ,$$

con $W = U^H V$ un'altra matrice unitaria. Sia ora φ la funzione dallo spazio delle matrici unitarie a $\mathbb{R}$, $\varphi(W) = \left\|\Lambda W - W\tilde{\Lambda}\right\|_F$. Se mostriamo che questa funzione ha minimo quando viene calcolata su una matrice di permutazione adeguata P_π, allora abbiamo finito, perché in tal caso si ha

$$\mathrm{md}_2^2(A, \tilde{A}) = \mathrm{md}_2^2(\Lambda, \tilde{\Lambda}) \leq \sum_i \left|\lambda_i - \tilde{\lambda}_{\pi(i)}\right|^2 \leq \left\|\Lambda W - W\tilde{\Lambda}\right\|_F^2 = \|E\|_F^2 .$$

Se la matrice W ha elementi $W_{i,j}$, allora avremo $\|E\|_F^2 = \sum_{i,j} \left|\lambda_i - \tilde{\lambda}_j\right|^2 |W_{i,j}|^2$. La matrice X che ha come elementi $|W_{i,j}|^2$ è una matrice doppiamente stocastica. Definiamo la funzione ψ dallo spazio delle matrici doppiamente stocastiche a $\mathbb{R}$ tale che, presa S una matrice doppiamente stocastica di elementi $S_{i,j}$, si ha

$$\psi(S) = \sum_{i,j} \left|\lambda_i - \tilde{\lambda}_j\right|^2 S_{i,j}.$$

Questa è una funzione lineare, ed è chiaro che $\inf \varphi \geq \inf \psi$, dato che φ calcolata sulle matrici unitarie assume le immagini che assume anche ψ calcolata su delle particolari matrici doppiamente stocastiche (definite come X). Ora, dal Teorema 1.4 sappiamo che ogni matrice doppiamente stocastica è combinazione lineare di matrici di permutazione, con coefficienti reali, non negativi, e che sommano a 1. Sia quindi P_π la matrice su cui la funzione ψ assume il valore minimo quando ristretta all'insieme delle $n!$ matrici di permutazione. Allora, per linearità di ψ, data una qualsiasi matrice doppiamente stocastica S si ha che $\psi(P_\pi) \leq \psi(S)$. Questo significa che ψ assume minimo in P_π, e quindi, dato che le matrici di permutazione sono unitarie, anche φ ha minimo in quella stessa matrice. □

Questo Teorema ci dà esattamente il risultato che volevamo, limita tutte le norme che abbiamo definito ed è il più simile alla disuguaglianza di Weyl. Vi è però il problema della difficile interpretazione delle condizioni di partenza, dal momento che non sappiamo quali perturbazioni conserverebbero la normalità di A e sarebbero quindi ammissibili per ottenere questo risultato. Il Teorema quindi, sebbene fornisca delle informazioni molto incoraggianti sulla perturbazione degli autovalori, risulta molto difficilmente applicabile. Proviamo allora a vedere cosa possiamo dire rinunciando alla condizione che $\tilde{A}$ sia normale. In tal caso non possiamo dare un Teorema così generale, ma possiamo comunque ottenere una disuguaglianza simile utilizzando risultati di base.

Teorema 3.5. *Sia A una matrice normale e $\tilde{A}$ una matrice qualsiasi, con $\tilde{A} = A + E$. Se $\|E\|$ è più piccola della metà della distanza tra una qualsiasi coppia di autovalori distinti di A, allora vale*

$$\mathrm{md}_\infty(A, \tilde{A}) \leq \|E\| .$$

Dimostrazione. Chiamiamo $\mu_1, \ldots, \mu_k$ tutti gli autovalori distinti di A, e $\varepsilon = \|E\|$. Dalla Proposizione 3.2 tutti gli autovalori di $\tilde{A}$ sono contenuti nell'unione dei dischi $\overline{B(\mu_j, \varepsilon)}$, $j = 1, \ldots, k$, che per ipotesi sono a due a due disgiunti. Dimostriamo che, se μ_j ha molteplicità m_j, allora nel disco $\overline{B(\mu_j, \varepsilon)}$ sono presenti esattamente m_j autovalori di $\tilde{A}$ contati con la loro molteplicità. Una volta dimostrato questo il risultato segue immediatamente, poiché per ogni autovalore di $\tilde{A}$, esso si troverà in uno dei dischi, e allora l'autovalore di A ad esso più vicino dovrà essere il centro del disco stesso, a una distanza non superiore a ε. Poiché in ogni disco c'è lo stesso numero di autovalori per A e per $\tilde{A}$, ogni coppia formata dalla definizione della distanza di corrispondenza ottimale avrà quindi autovalori distanti meno di ε, e allora $\mathrm{md}_\infty(A, \tilde{A}) \leq \varepsilon$, che era esattamente quello che volevamo.

Sia allora φ una funzione dall'intervallo $[0, 1]$ allo spazio delle matrici, con

$$\varphi(t) = (1 - t)A + t\tilde{A}.$$

Ovviamente φ è una funzione continua e vale $\varphi(0) = A$ e $\varphi(1) = \tilde{A}$. Calcoliamo immediatamente $\|A - \varphi(t)\| = \left\|A - (1 - t)A - t\tilde{A}\right\| = \left\|tA - t\tilde{A}\right\| = t\varepsilon$, quindi anche tutti gli autovalori di $\varphi(t)$ sono contenuti nell'unione dei dischi. Dal momento che gli autovalori di una matrice sono le radici del polinomio caratteristico, che ha come coefficienti somme e prodotti degli elementi della matrice stessa, possiamo scrivere un polinomio $p(\lambda, t)$ i cui coefficienti sono funzioni continue di t. Sfruttando il Teorema 1.19 e il fatto che $[0, 1]$ è un insieme compatto possiamo ricavare che, mentre t varia tra 0 e 1, gli autovalori di $\varphi(t)$ descrivono delle curve continue, che partono dai centri dei dischi (gli autovalori di

A), e li collegano agli autovalori di $\tilde{A}$. Dal momento che queste curve sono continue e tutti i dischi sono disgiunti, ogni curva dovrà rimanere completamente all'interno del suo disco. Da ogni centro di un disco $\overline{B(\mu_k, \varepsilon)}$ devono partire m_j curve. Infatti, se così non fosse, allora per continuità nemmeno A avrebbe esattamente m_j autovalori in $\overline{B(\mu_k, \varepsilon)}$, che è assurdo. Questo comporta che ogni $\varphi(t)$ ha esattamente m_j autovalori in ogni disco $\overline{B(\mu_k, \varepsilon)}$, e quindi anche $\tilde{A}$. □

Questo Teorema ha un raggio di applicazione limitato, perché funziona solo se la matrice non risulta essere "troppo perturbata", ma comunque dà una maggiorazione simile a quella delle matrici Hermitiane su una distanza di corrispondenza, apre la strada per un utilizzo della topologia nella risoluzione del problema di perturbazione matriciale, e soprattutto ci porta a pensare agli autovalori come funzioni continue degli elementi delle loro matrici.

Anche per quanto riguarda gli autovettori possiamo ottenere risultati totalmente analoghi a quelli delle matrici Hermitiane, e in questo caso la rimodellazione delle distanze per tenere conto di eventuali autovalori complessi è presto risolta. Riprendendo la notazione della Sezione 2.2 consideriamo un autospazio di A, $\mathcal{X}$, e un autospazio di $\tilde{A}$, $\mathcal{Y}$ (che sarà l'autospazio generato dalla perturbazione degli autovettori che generano $\mathcal{X}$). Se esprimiamo questi due sottospazi attraverso delle proiezioni lineari su $\mathbb{C}^n$ (chiamiamo P la matrice relativa alla proiezione su $\mathcal{X}$ e Q la matrice della proiezione su $\mathcal{Y}$), e supponiamo che i due sottospazi abbiano dimensione k possiamo ridefinire le matrici $G_0 \in \mathbb{C}^{n\times k}$, $G_1 \in \mathbb{C}^{n\times(n-k)}$, $H_0 \in \mathbb{C}^{n\times k}$ e $H_1 \in \mathbb{C}^{n\times(n-k)}$, dove $G_0(\mathbb{C}^k) = P(\mathbb{C}^n)$, $G_1(\mathbb{C}^{n-k}) = (I-P)(\mathbb{C}^n)$, $H_0(\mathbb{C}^k) = Q(\mathbb{C}^n)$ e $H_1(\mathbb{C}^{n-k}) = (I-Q)(\mathbb{C}^n)$. Inoltre G_0 ha come colonne una base ortonormale di $\mathcal{X}$, G_1 ha come colonne un completamento della base scelta per G_0 a una base ortonormale di $\mathbb{C}^n$ e H_0 e H_1 sono costruite allo stesso modo sul sottospazio $\mathcal{Y}$. Esprimendo la matrice A nelle coordinate (G_0, G_1) e la matrice $\tilde{A}$ nelle coordinate (H_0, H_1) riotteniamo le scritture

$$A = (G_0, G_1)\begin{pmatrix} A_0 & 0 \\ 0 & A_1 \end{pmatrix}\begin{pmatrix} G_0^H \\ G_1^H \end{pmatrix}, \qquad \tilde{A} = (H_0, H_1)\begin{pmatrix} \Lambda_0 & 0 \\ 0 & \Lambda_1 \end{pmatrix}\begin{pmatrix} H_0^H \\ H_1^H \end{pmatrix}.$$

Ridefiniamo ora la matrice residuo

$$R := \tilde{A}G_0 - G_0A_0,$$

che calcola quanto le autocoppie di $\mathcal{X}$ si discostino dall'essere autocoppie di $\tilde{A}$, e la matrice degli angoli canonici tra i due sottospazi,

$$\Theta := \begin{pmatrix} \Theta_0 & 0 \\ 0 & \Theta_1 \end{pmatrix},$$

che indica di quanto bisogna ruotare i vettori di $\mathcal{X}$ per ottenere quelli di $\mathcal{Y}$. L'unico momento nell'esposizione della teoria di Davis e Kahan in cui il fatto che gli autovalori di A e $\tilde{A}$ fossero reali è stato utilizzato era nella formulazione del Teorema 2.13 del seno, in cui chiedevamo che gli autovalori facessero parte di un determinato intervallo. Per adattare il Teorema 2.13 alla nostra nuova circostanza ci basterà quindi riformularlo come segue:

Teorema 3.6 (Teorema del seno, versione normale)**.** *Siano A e $\tilde{A}$ matrici normali, con $\rho, \delta > 0$, $\alpha \in \mathbb{C}$, A_0 e Λ_1 i blocchi definiti sopra. Supponiamo che lo spettro di A_0 sia compreso nella palla aperta $B(\alpha, \rho)$ e lo spettro di Λ_1 si trovi tutto al di fuori della palla aperta $B(\alpha, \rho + \delta)$ (oppure che lo spettro di Λ_1 sia compreso nella palla aperta $B(\alpha, \rho)$ e lo spettro di A_0 si trovi tutto al di fuori della palla aperta $B(\alpha, \rho + \delta)$). Allora per ogni norma invariante per trasformazioni unitarie si ha*

$$\delta \left\| \sin \Theta_0 \right\| \leq \|R\| .$$

Dimostrazione. La dimostrazione ricalca esattamente quella del Teorema 2.13. Questa volta, trasleremo A in modo da centrare la palla $B(\alpha, \rho)$ nell'origine. Dato che le matrici A e $\tilde{A}$ sono normali, seguendo la costruzione scritta sopra otterremo che A_0 e Λ_1 sono normali. Per loro quindi vale la Proposizione 1.16 e, dalle ipotesi, avremo $\|A_0\| < \rho$ e $\left\|\Lambda_1^{-1}\right\| \leq (\rho+\delta)^{-1}$. È quindi ancora possibile applicare la Proposizione 2.10, e concludere utilizzando gli stessi ragionamenti. □

Allo stesso modo, bastano poche modifiche per poter riusare la versione simmetrica del Teorema, la cui dimostrazione è identica a quella del Corollario 2.14.

Corollario 3.7 (Teorema del seno, versione simmetrica nel caso normale)**.** *Siano A e $\tilde{A}$ matrici normali, $\rho, \delta > 0$, $\alpha \in \mathbb{C}$, A_0, A_1, Λ_0 e Λ_1 i blocchi definiti sopra. Allora, se A_0 e Λ_1 sono separati come nel Teorema 3.6 e A_1 è separato da Λ_0 allo stesso modo, vale, per ogni norma invariante per trasformazioni unitarie,*

$$\delta \left\| \sin \Theta \right\| \leq \|E\| .$$

Notiamo come per questi Teoremi abbiamo nuovamente supposto che $\tilde{A}$ fosse normale. Questo è necessario per poter riutilizzare i risultati del Capitolo precedente, perché tutta la teoria di Davis e Kahan si basava sul fatto che entrambe le matrici A e $\tilde{A}$ fossero ortogonalmente diagonalizzabili. Questa volta, quindi, non possiamo permetterci di alleggerire le ipotesi.

3.2 Matrici Diagonalizzabili

Passiamo ora allo studio della perturbazione di matrici diagonalizzabili, ossia supponiamo che esista una matrice non singolare X per cui accada $A = XDX^{-1}$, con D matrice diagonale. La matrice X conterrà sulle sue colonne gli autovettori destri di A, mentre la matrice X^{-1} conterrà sulle sue righe i trasposti e coniugati degli autovettori sinistri di A. In generale, per queste matrici non riusciremo a ricavare risultati come quelli ottenuti nelle Sezioni precedenti, perché non possiamo più contare sul fatto che X sia unitaria. Togliendo questa ipotesi (equivalente ad A matrice normale, come lo era nei casi già analizzati), il numero di condizionamento della matrice degli autovettori non ha più alcuna limitazione, e questo malcondizionamento si propaga in tutti gli aspetti del problema che stiamo analizzando. Vediamo un esempio di questo fatto nell'applicazione del Teorema 1.20 di Bauer-Fike, che in questa circostanza assume una forma particolare.

Proposizione 3.8 (Bauer-Fike, versione per matrici diagonalizzabili)**.** *Sia $A \in \mathbb{C}^{n\times n}$ una matrice diagonalizzabile, con $A = XDX^{-1}$ e D matrice diagonale. Sia E una matrice qualsiasi che supponiamo di norma piccola, e $\tilde{A} = A + E$. Allora per ogni autovalore $\tilde{\lambda}$ di $\tilde{A}$ esiste un autovalore λ di A tale che*

$$\left|\tilde{\lambda} - \lambda\right| \leq \|X\|_2 \left\|X^{-1}\right\|_2 \|E\|_2 .$$

Dimostrazione. Se fosse $\tilde{\lambda} \in \mathrm{Spec}(A)$ allora si potrebbe scegliere $\lambda = \tilde{\lambda}$ e non ci sarebbe niente da dimostrare. Supponiamo quindi che $\tilde{\lambda}$ non sia autovalore di A. Ci basta allora utilizzare il Teorema 1.20 di Bauer-Fike generale, ponendo $Q = X^{-1}$ e utilizzando come norma matriciale quella indotta dalla norma euclidea. Infatti così facendo otteniamo

$$\left\|X^{-1}(A - \tilde{\lambda}I)^{-1}X\right\|_2^{-1} \leq \left\|X^{-1}EX\right\|_2 \leq \|X\|_2 \left\|X^{-1}\right\|_2 \|E\|_2 .$$

Ma $X^{-1}(A - \tilde{\lambda}I)^{-1}X = \left(X^{-1}(A - \tilde{\lambda}I)X\right)^{-1} = \left(D - \tilde{\lambda}I\right)^{-1}$ e dalla Proposizione 1.16, detti $\lambda_1, \ldots, \lambda_n$ gli autovalori di A, abbiamo

$$\left\|X^{-1}(A - \tilde{\lambda}I)^{-1}X\right\|_2^{-1} = \frac{1}{\max_i \left|\frac{1}{\lambda_i - \tilde{\lambda}}\right|} = \min_i \left|\lambda_i - \tilde{\lambda}\right| ,$$

per cui abbiamo dimostrato che esiste un autovalore $\lambda_{\bar{i}}$ di A per cui vale la disuguaglianza voluta. □

Vediamo chiaramente come in questo caso la maggiorazione comprende, oltre a $\|E\|_2$, anche il numero di condizionamento della matrice X, $\kappa_2(X) = \|X\|_2 \|X^{-1}\|_2$. Capiamo quindi che una stima per la variazione degli autovalori "alla Weyl" non è più sufficiente

in questo caso, perché la presenza di $\kappa_2(X)$ rende il risultato decisamente meno informativo. Inoltre, come già menzionato nel Capitolo 1, questa stima è uniforme per tutti gli autovalori, non tenendo conto del fatto che alcuni potrebbero venir perturbati molto meno di altri.

L'approccio che dovremo tenere da ora in avanti sarà quello di soffermarci su una singola autocoppia della matrice A, e provare a capire come una qualsiasi perturbazione possa agire su di essa. Nella dimostrazione del Teorema 3.5 avevamo dimostrato una certa continuità degli autovalori rispetto agli elementi della matrice. Un ragionamento del tutto analogo è applicabile anche al caso generale, in cui non si suppone alcuna struttura per A o $\tilde{A}$. Infatti in ogni caso, definendo la funzione φ come nel Teorema 3.5, otterremo polinomi caratteristici con coefficienti dati da funzioni continue (anzi, analitiche) in t, e le condizioni del Teorema 1.19 risulteranno comunque soddisfatte. Otteniamo quindi una continuità generale degli autovalori dai coefficienti della loro matrice, che possiamo sfruttare per descrivere meglio il comportamento degli autovalori perturbati. Assumendo che la matrice di perturbazione abbia una norma piccola uguale a $\varepsilon > 0$, questa continuità ci permette infatti di scrivere gli autovalori perturbati come funzioni continue di ε, e se riusciamo a capire la natura di questa dipendenza capiremo allora anche come una qualsiasi perturbazione di norma ε incida sugli autovalori di A. D'ora in avanti cercheremo quindi di esplicitare la dipendenza degli autovalori perturbati in questo modo e, per poter controllare meglio la norma della matrice di perturbazione, assumeremo che E sia una matrice qualsiasi, e che $\tilde{A} = A + \varepsilon E$, con $\varepsilon > 0$ un numero supposto piccolo.

Iniziamo a costruire l'apparato che ci servirà per dimostrare i risultati di questa Sezione. Innanzitutto, dal momento che stiamo supponendo A diagonalizzabile, avremo che le colonne della matrice X, che chiamiamo $\hat{x}_1, \dots, \hat{x}_n$ sono un set completo di autovettori destri per A. Denotiamo ora $Y^H = X^{-1}$, e chiamiamo le sue righe $\hat{y}_1^H \dots, \hat{y}_n^H$. I vettori $\hat{y}_1, \dots, \hat{y}_n$ sono un set completo di autovettori sinistri per A. Sostituiamo ora questi vettori con le loro forme normalizzate, ponendo quindi $x_i = \hat{x}_i / \|\hat{x}_i\|$ e $y_i = \hat{y}_i / \|\hat{y}_i\|$ per ogni $i = 1, \dots, n$. Poiché avevamo $Y^H X = I$, e abbiamo moltiplicato ogni vettore per uno scalare, avremo ancora $y_i^H x_j = 0$ se $i \neq j$. Mentre accadeva $\hat{y}_i^H \hat{x}_i = 1$ per ogni $i = 1, \dots, n$, però, questo non accade più con i vettori normalizzati. Definiamo quindi le quantità

$$s_i := y_i^H x_i, \qquad i = 1, \dots, n.$$

Contestualmente, definiremo anche le quantità

$$\beta_{i,j} := y_i^H E x_j, \qquad i, j = 1, \dots, n,$$

che ci saranno utili per verificare la perturbazione degli autovettori. Notiamo che, comunque, tutti gli s_i, $i = 1, \ldots, n$ sono non nulli. Per poter proseguire ci servono ora due Teoremi, di base per l'Algebra Lineare Numerica e la localizzazione degli autovalori di una matrice. Nel seguito di questa Sezione, indicheremo gli elementi di A come $a_{i,j},\ i, j = 1, \ldots, n$.

Teorema 3.9 (Primo di Gerschgorin)**.** *Sia $A \in \mathbb{C}^{n\times n}$ una matrice qualsiasi. Allora ogni autovalore di A giace in uno dei dischi di centro $a_{i,i}$ e raggio $\sum_{\substack{j=1\\ j\neq i}}^{n} |a_{i,j}|$, $i = 1, \ldots n$, chiamati **dischi di Gerschgorin per righe**, e in uno dei dischi di centro $a_{i,i}$ e raggio $\sum_{\substack{j=1\\ j\neq i}}^{n} |a_{j,i}|$, $i = 1, \ldots n$, chiamati **dischi di Gerschgorin per colonne**.*

Dimostrazione. Sia λ un autovalore di A. Allora esisterà un suo autovettore destro x, non nullo per definizione, per cui accade $Ax = \lambda x$. Sia $\hat{i} \in \{1, \ldots, n\}$ l'argomento di $\max_i |x_i|$. Sostituendo il vettore x con lo stesso vettore diviso per $x_{\hat{i}}$ otterremo un altro autovettore destro per λ di coordinate tutte minori o uguali a 1 in modulo, e la cui $\hat{i}$-sima coordinata è pari a 1. Allora, da $Ax = \lambda x$, otteniamo

$$\sum_{j=1}^{n} a_{\hat{i},j} x_j = \lambda x_{\hat{i}} = \lambda,$$

da cui

$$\left|\lambda - a_{\hat{i},\hat{i}}\right| \leq \sum_{\substack{j=1\\ j\neq\hat{i}}}^{n} \left|a_{\hat{i},j} x_j\right| \leq \sum_{\substack{j=1\\ j\neq\hat{i}}}^{n} \left|a_{\hat{i},j}\right|,$$

e abbiamo trovato così il disco di Gerschgorin per righe in cui giace λ. Per ottenere lo stesso risultato per i dischi per colonne basterà applicare lo stesso ragionamento a un autovettore sinistro di λ. □

Il Teorema ci indica una limitazione in cui può stare ogni autovalore di A, a patto di calcolarne un relativo autovettore destro e sinistro. Altre informazioni sulla posizione degli autovalori sono indicati dal secondo Teorema, di seguito.

Teorema 3.10 (Secondo di Gerschgorin)**.** *Se s dei dischi per righe di cui al Teorema 3.9 formano un dominio connesso e isolato dagli altri dischi, allora ci sono esattamente s autovalori di A in questo dominio connesso.*

Dimostrazione. Scriviamo $A = D + M$, dove abbiamo indicato con D la matrice che contiene la diagonale di A sulla sua diagonale principale ed è nulla altrove. Consideriamo ora le matrici $A_t := D + tM$, con $0 \leq t \leq 1$. Per $t = 0$ questa è una matrice diagonale, i cui autovalori, essendo gli elementi $a_{i,i},\ i = 1, \ldots, n$ sono i centri dei dischi. Nel dominio connesso di cui all'enunciato del Teorema, quindi, ci saranno sicuramente s di questi

autovalori, contati con molteplicità. Ora, sappiamo che gli autovalori di A_t in questa circostanza sono funzioni continue di t e, al variare di t, formano delle curve continue sul piano complesso. Inoltre, per il Teorema 3.9, tali curve dovranno essere totalmente contenute nell'unione di tutti i dischi di Gerschgorin per righe. Poiché in partenza c'erano s autovalori nell'unione degli s dischi di cui all'enunciato del Teorema, ed essi formavano un dominio isolato dagli altri dischi, per continuità le curve originate in questo dominio non possono fuoriuscire da esso, e nuove curve non possono entrarci. Ogni matrice A_t avrà quindi esattamente s autovalori contati con la loro molteplicità nell'unione degli s dischi, e questo varrà anche per $A = A_1$. □

Questo Teorema, oltre a darci informazioni importanti sulla localizzazione degli autovalori, utilizza gli stessi principi che avevamo usato per dimostrare la continuità degli autovalori perturbati. Non stupisce, quindi, che sia applicabile anche al nostro problema di perturbazione.

Siamo ora pronti per esporre i risultati di perturbazione, soffermandoci su una specifica autocoppia (λ, x).

Sia allora $A = \hat{X}D\hat{X}^{-1}$, con $D = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$, e supponiamo $\lambda = \lambda_1$. Sostituiamo ora a $\hat{X}$ la matrice X le cui colonne sono le colonne normalizzate di $\hat{X}$. La sua inversa sarà data dalla matrice Y^H le cui righe sono y_i^H/s_i, $i = 1, \ldots, n$, dove gli y_i sono gli autovettori sinistri normalizzati che avevamo introdotto sopra. Ovviamente, dato che X e Y continuano a contenere un insieme completo di autovettori destri e sinistri per A, avremo ancora $Y^HAX = D$. Calcolando Y^HEX otteniamo la matrice

$$\begin{pmatrix} \beta_{1,1}/s_1 & \beta_{1,2}/s_1 & \cdots & \beta_{1,n}/s_1 \\ \beta_{2,1}/s_2 & \beta_{2,2}/s_2 & \cdots & \beta_{2,n}/s_2 \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n,1}/s_n & \beta_{n,2}/s_n & \cdots & \beta_{n,n}/s_n \end{pmatrix}.$$

La matrice $\tilde{A} = A + \varepsilon E$ è quindi simile alla matrice

$$\begin{pmatrix} \lambda_1 + \varepsilon\beta_{1,1}/s_1 & \varepsilon\beta_{1,2}/s_1 & \cdots & \varepsilon\beta_{1,n}/s_1 \\ \varepsilon\beta_{2,1}/s_2 & \lambda_2 + \varepsilon\beta_{2,2}/s_2 & \cdots & \varepsilon\beta_{2,n}/s_2 \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon\beta_{n,1}/s_n & \varepsilon\beta_{n,2}/s_n & \cdots & \lambda_n + \varepsilon\beta_{n,n}/s_n \end{pmatrix}$$

e dal Teorema 3.9 possiamo dire che tutti i suoi autovalori sono contenuti in uno dei dischi di centro $\lambda_i + \varepsilon\beta_{i,i}/s_i$ e raggio $\varepsilon \sum_{j \neq i} |\beta_{i,j}/s_i|$. Possiamo quindi scrivere, per ogni

autovalore $\tilde{\lambda}$ di $\tilde{A}$, che

$$\tilde{\lambda} = \lambda + \varepsilon\beta_{i,i}/s_i + c\varepsilon \sum_{j\neq i} |\beta_{i,j}/s_i| \,,$$

per un opportuno $i \in \{1, \ldots, n\}$, e dove c è un numero complesso di norma minore o uguale a 1. Cerchiamo ora di valutare la perturbazione dell'autovalore λ_1.

Caso 1: λ_1 è un autovalore semplice

In questo caso, poiché diminuendo ε i centri dei vari dischi si avvicinano agli autovalori di A e i loro raggi, dipendendo da ε, diventano più piccoli, esisterà un ε abbastanza piccolo da rendere isolato il disco dato dalla prima riga di $Y^H\tilde{A}X$. Dal Teorema 3.10 quindi, questo disco conterrà un solo autovalore, che per continuità deve essere quello ottenuto dalla perturbazione di λ_1. Consideriamo ora la matrice $M = \mathrm{Diag}(m, \underbrace{1, \ldots, 1}_{n-1 \text{ volte}})$ e la matrice $M^{-1}Y^H\tilde{A}XM$, che è simile a $\tilde{A}$ e ha quindi gli stessi autovalori. Rispetto alla matrice $Y^H\tilde{A}X$ abbiamo solo così moltiplicato la prima colonna per m e diviso la prima riga per m. Poniamo $m = k/\varepsilon$. La matrice diventa allora

$$\begin{pmatrix} \lambda_1 + \varepsilon\beta_{1,1}/s_1 & \varepsilon^2\beta_{1,2}/ks_1 & \varepsilon^2\beta_{1,3}/ks_1 & \cdots & \varepsilon^2\beta_{1,n}/ks_1 \\ k\beta_{2,1}/s_2 & \lambda_2 + \varepsilon\beta_{2,2}/s_2 & \varepsilon\beta_{2,3}/s_2 & \cdots & \varepsilon\beta_{2,n}/s_2 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ k\beta_{n,1}/s_n & \varepsilon\beta_{n,2}/s_n & \varepsilon\beta_{n,3}/s_n & \cdots & \lambda_n + \varepsilon\beta_{n,n}/s_n \end{pmatrix}.$$

Aumentando k otteniamo quindi di diminuire il raggio del disco corrispondente alla prima riga, mentre i raggi di tutti gli altri dischi aumentano e i centri restano invariati. Se riusciamo ad ottenere un k indipendente da ε (che assumiamo sempre appartenere a un intorno di 0) tale che il disco relativo alla prima riga della matrice sia ancora isolato dagli altri, allora avremo dimostrato che possiamo scrivere la perturbazione di λ_1 come

$$\tilde{\lambda}_1 = \varepsilon\beta_{1,1}/s_1 + O(\varepsilon^2).$$

Questa è un'informazione essenziale, perché ci dice che la perturbazione è di ordine ε, con il coefficiente è pari a $\beta_{1,1}/s_1 = y_1^H E x_1 / y_1^H x_1$, che dipende quindi dalla posizione reciproca degli autovettori destro e sinistro di λ_1: tanto più essi saranno ortogonali tra loro, tanto la perturbazione sarà accentuata.

Ma un k del genere si trova subito, basta infatti scegliere k tale che

$$|k\beta_{i,1}/s_i| \leq \frac{1}{p} |\lambda_1 - \lambda_i| \qquad \forall\, i = 2, 3, \ldots, n,$$

con p un qualsiasi numero naturale maggiore di 2. Per ε che tende a 0, infatti, i centri dei dischi tendono agli autovalori di A, il raggio del disco di λ_1 tende a 0, e il raggio di

tutti gli altri dischi tende a $|k\beta_{i,1}/s_i|$. Inoltre, la dipendenza dei centri dei dischi e dei raggi dei dischi diversi dal primo da ε è dello stesso ordine, mentre il raggio del primo disco aumenta con ε^2 (quindi meno velocemente, in un intorno di 0). Se k soddisfa queste disuguaglianze possiamo quindi concludere che esisterà un ε abbastanza piccolo da garantire la condizione che stavamo cercando. Possiamo allora scegliere

$$k = \min_i |(\lambda_1 - \lambda_i)s_i/p\beta_{i,1}| .$$

Abbiamo così dimostrato che, in questo caso, la perturbazione degli autovalori ha ordine paragonabile a quello della perturbazione della matrice, descrivendola precisamente in relazione agli autovettori destro e sinistro di λ_1.

Per quanto riguarda la perturbazione dell'autovettore (destro) relativo, sia esso $x_1(\varepsilon)$ (si tratta quindi dell'autovettore corrispondente a $\tilde{\lambda}_1$ per la matrice $\tilde{A}$). Esso è unico a meno di moltiplicazione per scalare, perché per le costruzioni fatte per trovare $\tilde{\lambda}_1$, esso risulta essere semplice. Chiamiamo ora $z_1(\varepsilon)$ l'autovettore per $\tilde{\lambda}_1$ nella matrice $Y^H\tilde{A}X$, e supponiamo che $z_1(\varepsilon)$ sia stato diviso per la sua componente di modulo massimo, come avevamo fatto nella dimostrazione del Teorema 3.9. Avremo intanto

$$x_1(\varepsilon) = Xz_1(\varepsilon) = \alpha_1(\varepsilon)x_1 + \alpha_2(\varepsilon) + \cdots + \alpha_n(\varepsilon)x_n,$$

dove le $\alpha_i(\varepsilon),\ i = 1, \ldots, n$ sono le componenti di $z_1(\varepsilon)$. Dimostriamo ora che, sempre supponendo che ε sia sufficientemente piccolo, si ha $\alpha_1(\varepsilon) = 1$ e tutte le altre componenti abbiano modulo strettamente minore di 1. Consideriamo infatti una componente generica di $z_1(\varepsilon)$, $\alpha_i(\varepsilon)$. Allora, da $\tilde{\lambda}_1 z_1(\varepsilon) = Y^H\tilde{A}Xz_1(\varepsilon)$ abbiamo

$$\tilde{\lambda}_1\alpha_i(\varepsilon) = \lambda_i\alpha_i(\varepsilon) + \frac{\varepsilon}{s_i}\sum_{j=1}^{n}\beta_{i,j}\alpha_j(\varepsilon) \tag{3.1}$$

da cui, se fosse $|\alpha_i| = 1$, ricaviamo $\left|\tilde{\lambda}_1 - \lambda_i\right| = \frac{\varepsilon}{|s_i|}\left|\sum_{j=1}^{n}\beta_{i,j}\alpha_j(\varepsilon)\right|$. Poiché per ε che tende a 0 $\tilde{\lambda}_1$ tende a λ_1, e le componenti $\alpha_i(\varepsilon)$ sono limitate (dalla normalizzazione di $z_1(\varepsilon)$ hanno modulo non superiore a 1), abbiamo che in questa circostanza il primo membro dell'equazione tende a $|\lambda_1 - \lambda_i|$, mentre il secondo tende a 0. Per la semplicità di λ_1 questo ha senso solo se $i = 1$.
Vogliamo ora studiare il comportamento delle altre componenti di $z_1(\varepsilon)$ in modo più preciso. Riprendiamo quindi l'equazione (3.1) per $i \neq 1$. Possiamo ricavare

$$\left|\tilde{\lambda}_1 - \lambda_i\right| |\alpha_i(\varepsilon)| \leq \frac{\varepsilon}{|s_i|}\sum_{j=1}^{n}|\beta_{i,j}| .$$

Se consideriamo $\left|\tilde{\lambda}_1 - \lambda_i\right| \geq \frac{1}{p}\left|\lambda_1 - \lambda_i\right|$, condizione che per la nostra costruzione nello studio dell'autovalore perturbato è verificata, allora si ha

$$|\alpha_i(\varepsilon)| \leq \frac{p\varepsilon \sum_{j=1}^{n} |\beta_{i,j}|}{|s_i|\,|\lambda_1 - \lambda_i|}.$$

Abbiamo così ottenuto che le coordinate diverse dalla prima di $z_1(\varepsilon)$ sono di ordine ε. Se consideriamo che $z_1(0)$ è il primo vettore della base canonica questo stabilisce anche una continuità dell'autovettore visto come funzione di ε, in un intorno di 0. Possiamo ora riprendere l'equazione (3.1) un'ultima volta, e trovare

$$(\tilde{\lambda}_1(\varepsilon) - \lambda_i)\alpha_i(\varepsilon) = \frac{\varepsilon\beta_{i,1}}{s_i} + \frac{\varepsilon}{s_i}\sum_{j=2}^{n} \beta_{i,j}\alpha_j(\varepsilon).$$

Per quanto abbiamo dimostrato il secondo addendo del termine di destra è un $O(\varepsilon^2)$. Possiamo quindi scrivere ogni coordinata in serie nel modo seguente:

$$\alpha_i(\varepsilon) = \varepsilon\frac{\beta_{i,1}}{s_i(\tilde{\lambda}_1 - \lambda_i)} + O(\varepsilon^2),$$

espressione analoga a quella ottenuta per l'autovalore.

Caso 2: λ_1 è un autovalore multiplo, di molteplicità m_1

Supponiamo che gli autovalori in D siano ordinati in modo che i primi m_1 elementi della diagonale siano tutti gli autovalori uguali a λ_1. In questo caso non riusciamo a isolare il disco di Gerschgorin per righe relativo alla prima riga della matrice, ma possiamo isolare gli m_1 dischi corrispondenti alle prime m_1 righe da tutti gli altri, scegliendo come prima un ε abbastanza piccolo. Applicando ancora il Teorema 3.10, concludiamo quindi che nell'unione di questi dischi saranno presenti m_1 autovalori perturbati, ognuno dei quali sarà quindi una perturbazione di λ_1. Sappiamo già che ognuno di questi autovalori si può scrivere come

$$\tilde{\lambda}_1^{(l)} = \lambda_1 + \varepsilon\beta_{1,1}/s_1 + c^{(l)}R^{(l)}, \qquad l = 1, \ldots, m_1,$$

nelle notazioni precedenti, con $R^{(l)}$ il raggio del disco relativo alla l-esima riga. Non riusciamo più in questo caso a ridurre i raggi a un $O(\varepsilon^2)$, ma possiamo comunque applicare un ragionamento analogo a quello del caso precedente, usando la matrice $M = \mathrm{Diag}(\underbrace{\varepsilon/k, \ldots, \varepsilon/k}_{m_1 \text{ volte}}, \underbrace{1, \ldots, 1}_{n-m_1 \text{ volte}})$ e lavorando a blocchi. Scriveremo la matrice $M^{-1}Y^H\tilde{A}XM$ come

$$\left(\begin{array}{c|c} \mathrm{Diag}(\lambda_1, \ldots, \lambda_{m_1}) + \varepsilon P & k^{-1}\varepsilon^2 Q \\ \hline kR & \mathrm{Diag}(\lambda_{m_1+1}, \ldots, \lambda_n) + \varepsilon S \end{array}\right).$$

Applicando ragionamenti analoghi a quelli del caso precedente troviamo un k indipendente da ε che isoli i primi m_1 dischi da tutti gli altri, riducendo i loro raggi. In questo modo siamo riusciti a ridurre alcuni dei termini nell'espressione del raggio a un $O(\varepsilon^2)$, ma rimangono comunque dei termini in ε (quelli presenti nel blocco in alto a sinistra). Otterremo quindi l'espressione:

$$\tilde{\lambda}_1^{(l)} = \lambda_1 + \varepsilon \left(\beta_{l,l}/s_l + c \sum_{\substack{j=1 \\ j \neq l}}^{m_1} |\beta_{l,j}/s_l| \right) + O(\varepsilon^2), \qquad l = 1, \ldots, m_1.$$

Il problema ha un risultato simile a quello del caso precedente, ma c'è una variazione dovuta al coefficiente della ε che, contenendo una dipendenza da molti più coefficienti β e s, è molto più soggetto all'eventuale malcondizionamento degli autovettori. Esploreremo meglio questo malcondizionamento nel Capitolo 4, trattando il caso generale.

Per quanto riguarda gli autovettori assumiamo che gli autovalori perturbati siano semplici, per poter riapplicare i ragionamenti del caso precedente. Possiamo ancora utilizzare l'equazione (3.1) ma ricaveremo solo che le componenti $\alpha_{m_1+1}(\varepsilon) \ldots, \alpha_n(\varepsilon)$ hanno norma strettamente minore di 1, e dipendono in modo continuo da ε, con una dipendenza di ordine ε e lo stesso coefficiente. Non riusciamo con questa strada a dire niente sulle componenti $\alpha_1(\varepsilon), \ldots, \alpha_{m_1}(\varepsilon)$. Per avere una descrizione più precisa e completa di questo caso servirà una teoria molto più robusta, che esporremo nel Capitolo 4.

Capitolo 4

Matrici non Strutturate

Avendo ora concluso lo studio della perturbazione di matrici strutturate iniziamo ad esporre la teoria generale, non facendo quindi alcuna ipotesi sulle proprietà di A. I risultati che otterremo in questo Capitolo saranno pessimisti, perché non abbiamo più alcuna garanzia che ci porti a lavorare al di fuori dei casi più malcondizionati. In questi casi, come per le matrici diagonalizzabili, otterremo delle espansioni delle autocoppie perturbate in termini del parametro ε. In questa Sezione però le potenze di ε che compariranno saranno frazionari, mostrando a pieno il malcondizionamento del problema. Inizieremo dalla teoria sviluppata da Lidskii, le cui limitazioni la rendono applicabile nella maggioranza delle situazioni, per poi coprire una teoria che fa uso dei cosiddetti "diagrammi di Newton", che coprirà molti dei casi rimanenti. Anche in questo Capitolo, come nella Sezione 3.2, porremo che E sia una matrice qualsiasi, e che la matrice di perturbazione sia εE, con ε sufficientemente piccolo.

4.1 La teoria perturbativa di Lidskii

Per esporre la teoria di Lidskii, per autovalori e autovettori, ci serve intanto costruire un apparato adeguato, a partire dalle matrici A ed E. Dal momento che stiamo cercando di analizzare la perturbazione di un autovalore qualsiasi di A, e dei suoi autovettori, focalizziamoci ancora su un preciso autovalore, $\lambda \in \mathrm{Spec}(A)$, che può avere molteplicità qualsiasi.
Decomponiamo quindi A nella sua forma di Jordan:

$$\begin{pmatrix} J & \\ & \hat{J} \end{pmatrix} = \begin{pmatrix} Q \\ \hat{Q} \end{pmatrix} A \begin{pmatrix} P & \hat{P} \end{pmatrix},$$

dove J è una matrice quadrata che contiene tutti i blocchi di Jordan relativi all'autovalore λ, $QAP = J$, e $\begin{pmatrix} Q \\ \hat{Q} \end{pmatrix} \begin{pmatrix} P & \hat{P} \end{pmatrix} = I$. Consideriamo le matrici J, Q, e P. Possiamo spezzare ulteriormente $J = \mathrm{Diag}(\Gamma_1^1, \ldots, \Gamma_1^{r_1}, \ldots, \Gamma_q^1, \ldots, \Gamma_q^{r_q})$, dove, per $j = 1, \ldots, q$, $\Gamma_j^1, \ldots, \Gamma_j^{r_j}$ sono r_j blocchi di Jordan relativi all'autovalore λ, di ordine n_j. Scegliamo di ordinare $n_1 > \cdots n_2 > \cdots > n_q$. Quindi, fissato j, si avrà

$$\Gamma_j^1 = \cdots = \Gamma_j^{r_j} = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}.$$

Concordemente, partizioniamo ulteriormente

$$P = \begin{pmatrix} P_1^1 & \cdots & P_1^{r_1} & \cdots & P_q^1 & \cdots & P_q^{r_q} \end{pmatrix} \quad Q = \begin{pmatrix} Q_1^1 \\ \vdots \\ Q_1^{r_1} \\ \vdots \\ Q_q^1 \\ \vdots \\ Q_q^{r_q} \end{pmatrix}.$$

Risulta chiaro che la prima colonna di ogni P_j^k (che chiameremo x_j^k) corrisponde ad un autovettore destro per λ, mentre le ultime righe delle Q_j^k (le y_j^k) corrispondono agli autovettori sinistri. Fissato j e chiamate X_j le matrici che hanno rispettivamente come colonne tutti gli autovettori x_j^k, $k = 1, \ldots, r_j$, e Y_j le matrici che hanno come righe gli autovettori y_j^k, $k = 1, \ldots, r_j$, possiamo quindi definire delle matrici di ordine crescente:

$$W_s = \begin{pmatrix} Y_1 \\ \vdots \\ Y_s \end{pmatrix} \quad Z_s = \begin{pmatrix} X_1 & \cdots & X_s \end{pmatrix}, \qquad s = 1, \ldots q.$$

In questo modo abbiamo modo di poter controllare specificamente i blocchi di Jordan relativi all'autovalore λ, senza dover portarci dietro anche gli altri. Rimane da definire una quantità che dipenda dalla matrice di perturbazione. Definiamo allora q matrici, sempre di ordine crescente:

$$\Phi_s = W_s E Z_s, \quad \mathcal{I}_s = \begin{pmatrix} 0 & 0 \\ 0 & I_{r_s} \end{pmatrix}, \qquad s = 1, \ldots, q,$$

dove I_j è la matrice identità di ordine j, e la dimensione del blocco di zeri in $\mathcal{I}_s$ è scelta in modo che $\mathcal{I}_s$ e Φ_s abbiano la stessa dimensione $f_s = \sum_{j=1}^{s} r_j$. Osserviamo come la

definizione di Φ_s sia quasi ricorsiva, dal momento che preso $s > 1$ si ha che Φ_{s-1} è un blocco in alto a sinistra di Φ_s.

Siamo ora pronti per dimostrare i due Teoremi di Lidskii, che, a patto di avere delle condizioni sulle Φ_s, descrivono in modo molto puntuale la perturbazione delle autocoppie.

Teorema 4.1 (Lidskii, autovalori)**.** *Nelle notazioni precedenti, sia $j \in \{1, \dots, q\}$ fissato, e supponiamo che, per $j > 1$, si abbia Φ_{j-1} non singolare. Allora ci sono $r_j n_j$ autovalori di $\tilde{A} = A + \varepsilon E$ la cui espansione al primo ordine in termini di ε è*

$$\lambda_j^{(k)(l)}(\varepsilon) = \lambda + (\xi_j^{(k)})^{1/n_j} \varepsilon^{1/n_j} + o(\varepsilon^{1/n_j}), \qquad l = 1, \dots, n_j,\ k = 1, \dots r_j,$$

dove le $\xi_j^{(k)}$ sono le soluzioni dell'equazione nella variabile ξ

$$\det(\Phi_j - \xi \mathcal{I}_j) = 0.$$

Variare quale soluzione tra le r_j totali dell'equazione si usa nell'espansione dà la variabilità di $\tilde{\lambda}^{(k)(l)}$ relativamente al parametro k, mentre le n_j radici che si ottengono dell'esponente $1/n_j$ di ξ presente nell'espansione dà la variabilità degli autovalori perturbati rispetto al parametro l.

Inoltre, se le $\xi_j^{(k)}$ sono tutte diverse, allora l'espansione degli autovalori perturbati diventa, localmente, della forma

$$\lambda_j^{(k)(l)}(\varepsilon) = \lambda + (\xi_j^{(k)})^{1/n_j} \varepsilon^{1/n_j} + \sum_{s=2}^{\infty} a_{js}^{(k)(l)} \varepsilon^{s/n_j}, \qquad l = 1, \dots, n_j,\ k = 1, \dots r_j.$$

Dimostrazione. Senza perdere di generalità possiamo assumere che A abbia solo un autovalore, ovvero che $\hat{J}$ sia vuota. Questo infatti è lo scopo primario della nostra costruzione, e se così non fosse, ci basterebbe applicare appropriate proiezioni per considerare solo l'autospazio voluto. Chiamata $\tilde{E} := P^{-1}EP$, vogliamo quindi studiare le radici ω del polinomio caratteristico

$$\det(C(\omega, \varepsilon)) = \det(\omega I - J - \varepsilon \tilde{E}).$$

Possiamo quindi considerare questo come un polinomio nelle due variabili ω e ε, e ci è quindi consentito effettuare i cambiamenti di variabile

$$\begin{cases} z = \varepsilon^{1/n_j} \\ \mu = \frac{\omega - \lambda}{z} \end{cases}.$$

L'equazione caratteristica diventa quindi

$$\det(\mathcal{P}(\mu, z)) = \det((\lambda + \mu z) I - J - z^{n_j} \tilde{E}) = 0,$$

rinominando il polinomio $\det(\mathcal{P}(\mu,z)) = \det(C(\lambda+\mu z, z^{n_j}))$. Noi stiamo lavorando con l'assunzione di fondo che ε sia un numero di norma molto ridotta, quindi le radici di $\det(\mathcal{P}(\mu,z))$ andranno cercate in un intorno di $z=0$. Consideriamo le seguenti matrici, partizionate nello stesso modo di J, Q e P:

$$L(z) = \mathrm{Diag}(L_1^1, \ldots, L_1^{r_1}, \ldots, L_q^1, \ldots, L_q^{r_q})$$

$$R(z) = \mathrm{Diag}(R_1^1, \ldots, R_1^{r_1}, \ldots, R_q^1, \ldots, R_q^{r_q}),$$

dove

$$\begin{aligned} L_i^1(z) = \cdots = L_i^{r_i}(z) &= \mathrm{Diag}(z^{-1}, z^{-2}, \ldots, z^{-n_i}) && \text{se } i \geq j \\ L_i^1(z) = \cdots = L_i^{r_i}(z) &= \mathrm{Diag}(\underbrace{1, \ldots, 1}_{n_i - n_j \text{ volte}}, z^{-1}, z^{-2}, \ldots, z^{-n_j}) && \text{se } i < j \end{aligned}$$

e

$$\begin{aligned} R_i^1(z) = \cdots = R_i^{r_i}(z) &= \mathrm{Diag}(1, z, z^2, \ldots, z^{n_i-1}) && \text{se } i \geq j \\ R_i^1(z) = \cdots = R_i^{r_i}(z) &= \mathrm{Diag}(\underbrace{1, \ldots, 1}_{n_i - n_j \text{ volte}}, 1, z, z^2, \ldots, z^{n_j-1}) && \text{se } i < j. \end{aligned}$$

Per le nostre costruzioni iniziali avremo $n_i \geq n_j$ se e solo se $i \leq j$. Ora definiamo una nuova matrice, $F(\mu,z) = L(z)\mathcal{P}(\mu,z)R(z)$, e la quantità

$$\mathcal{Q}(\mu,z) = \det F(\mu,z).$$

Per $z \neq 0$ $L(z)$ e $R(z)$ sono non singolari, quindi in tal caso avremo

$$\det \mathcal{P}(\mu,z) = 0 \Leftrightarrow \mathcal{Q}(\mu,z) = 0.$$

Possiamo quindi cercare gli autovalori della matrice perturbata come radici di $\mathcal{Q}(\mu,z)$. Dimostriamo innanzitutto che $\mathcal{Q}(\mu,z)$ è un polinomio nelle variabili μ e z. Per fare questo è utile spezzare la matrice $F(\mu,z) = G(\mu,z) + H(z)$, ponendo

$$G(\mu,z) = L(z)((\lambda+\mu z)I - J)R(z)$$

e

$$H(z) = -z^{n_j} L(z)\tilde{E}(z)R(z)$$

e dividere ulteriormente ogni blocco di Jordan nella forma $\Gamma_s^k = \lambda I + N_s$, con

$$N_s = \begin{pmatrix} 0 & 1 & & \\ & \cdot & \cdot & \\ & & \cdot & 1 \\ & & & 0 \end{pmatrix}.$$

Dei calcoli diretti mostrano subito che valgono le proprietà

$$\begin{aligned} &L_i^k(z)N_iR_i^k(z) = N_i && \text{per } i = 1,\dots,q, \quad k = 1,\dots,r_i \\ &L_i^k R_i^k = z^{-1}I && \text{per } i \geq j \\ &L_i^k R_i^k = \mathrm{Diag}(\underbrace{1,\dots,1}_{n_i-n_j \text{ volte}}, z^{-1}, z^{-1}, \dots, z^{-1}) && \text{per } i < j. \end{aligned}$$

Per ottenere $G(\mu, z)$ moltiplichiamo tre matrici diagonali a blocchi aventi blocchi delle stesse rispettive dimensioni, quindi la matrice risultante deve anch'essa presentare la stessa struttura.
Consideriamo uno dei blocchi diagonali di $G(\mu, z)$, $G_i^k(\mu, z) = L_i^k(z)(\mu z I - N_i)R_i^k(z)$. Applicando le proprietà osservate sopra possiamo vedere che

$$G_i^k(\mu, z) = \begin{cases} \mu I - N_i & \text{se } i \geq j \\ \mathrm{Diag}(\underbrace{\mu z, \dots, \mu z}_{n_i-n_j \text{ volte}}, \mu, \dots, \mu) - N_i & \text{se } i < j \end{cases}.$$

Quindi, μ e z sono presenti solo a esponenti positivi in $G(\mu, z)$. Per quanto riguarda $H(z)$, la presenza del fattore moltiplicativo z^{n_j} assicura che la matrice non contenga potenze negative di z, e la variabile μ non compare in nessun fattore tra $L(z)$, $R(z)$ e $\tilde{E}$, quindi non compare nemmeno in $H(z)$. Abbiamo quindi dimostrato che $F(\mu, z) = G(\mu, z) + H(z)$ contiene solo potenze non negative di μ e z, e quindi che $\mathcal{Q}(\mu, z)$ è effettivamente un polinomio in queste variabili.

Esaminiamo ora il caso $z = 0$. Per quanto riguarda $G(\mu, 0)$ abbiamo già un'analisi completa di tutti i suoi elementi, dati dall'analisi dei blocchi di cui sopra. Avremo quindi

$$G_i^k(\mu, 0) = \begin{cases} \mu I - N_i & \text{se } i \geq j \\ \mathrm{Diag}(\underbrace{0,\dots,0}_{n_i-n_j \text{ volte}}, \mu, \dots, \mu) - N_i & \text{se } i < j \end{cases}.$$

Per studiare $H(0)$, invece, è opportuno effettuare altre divisioni in blocchi, in modo sempre coerente con le divisioni iniziali di J, Q, e P. Scriviamo allora

$$\tilde{E}_{j_1 j_2}^{k_1 k_2}, \qquad j_i = 1,\dots,q, \quad k_i = 1,\dots,r_{j_i}, \quad i = 1, 2,$$

e usiamo questa scrittura per identificare il blocco di $\tilde{E}$ che giace sugli stessi indici di riga di $\Gamma_{j_1}^{k_1}$ e sugli stessi indici di colonna di $\Gamma_{j_2}^{k_2}$. Possiamo quindi definire

$$H_{j_1 j_2}^{k_1 k_2}(z) = -z^{n_j} L_{j_1}^{k_1}(z) \tilde{E}_{j_1 j_2}^{k_1 k_2} R_{j_2}^{k_2}(z)$$

e osservare che per ogni valore di j_1 e j_2 si avrà comunque che tutte le righe di $H_{j_1 j_2}^{k_1 k_2}(0)$, tranne l'ultima, sono identicamente nulle. Questo perché qualunque sia il valore di j_1,

$-z^{n_j}L_{j_1}^{k_1}$ presenterà sempre sulla sua diagonale potenze positive di z, a parte al più nell'ultimo elemento, in cui potrebbe essere presente $z^0 = 1$. Il prodotto $-z^{n_j}L_{j_1}^{k_1}\tilde{E}$ ha quindi elementi moltiplicati per una potenza non negativa di z in tutte le righe, tranne al più l'ultima, in cui z potrebbe avere esponente 0. Il prodotto di questa matrice appena ottenuta per $R_{j_2}^{k_2}$ non rimuove alcuna potenza di z, e in più assicura che il primo elemento dell'ultima riga di $H_{j_1j_2}^{k_1k_2}(z)$ rimanga invariato, in quanto il primo elemento della diagonale di $R_{j_2}^{k_2}$ è sempre 1. Se $j_2 \geq j$ poi, l'unico elemento pari a 1 della diagonale di $R_{j_2}^{k_2}$ sarà il primo, e gli altri saranno potenze positive di z. In questo caso, quindi, gli elementi dal secondo in poi dell'ultima riga di $H_{j_1j_2}^{k_1k_2}(z)$, verranno moltiplicati per una potenza di z. In caso $j_2 < j$, gli elementi dell'ultima riga di $H_{j_1j_2}^{k_1k_2}(z)$ che vengono moltiplicati per z saranno solo gli ultimi $(n_j - 1)$-simi, perché $R_{j_2}^{k_2}$ presenta degli 1 sui primi $(n_{j_2} - n_j + 1)$ elementi della sua diagonale. Ricapitolando, abbiamo ottenuto che la matrice $H_{j_1j_2}^{k_1k_2}(0)$ ha la seguente forma:

$$
\begin{array}{ll}
0 & \text{se } j_1 > j \\
\begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \\ -\beta_{j_1j_2}^{(k_1)(k_2)} & 0 & \cdots & 0 \end{pmatrix} & \text{se } j_1 \leq j, \quad j_2 > j \\
\begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ -\beta_{j_1j_2}^{(k_1)(k_2)} & * & \cdots & * & 0 & \cdots & 0 \end{pmatrix} & \text{se } j_1 \leq j, \quad j_2 \leq j,
\end{array}
$$

dove il numero di asterischi nell'ultimo caso è $n_{j_2} - n_j$, e si tratta di elementi diversi da 0 che non avranno ruolo nel resto della dimostrazione.

Ora osserviamo che, dato che le $-\beta_{j_1j_2}^{(k_1)(k_2)}$ sono ottenute moltiplicando il primo elemento dell'ultima riga di $\tilde{E}_{j_1j_2}^{k_1k_2}$ per degli 1 sulle diagonali di $L_{j_1}^{k_1}$ e $R_{j_2}^{k_2}$, e il segno negativo è derivante dalla moltiplicazione per $-z^{n_j}$, si ha che $\beta_{j_1j_2}^{(k_1)(k_2)} = \tilde{E}_{j_1j_2}^{k_1k_2}(n_{j_1}, 1)$, che per definizione di $\tilde{E} = QEP$ è l'elemento $y_{j_1}^{k_1}Ex_{j_2}^{k_2}$. Quindi, scelto $t = \max(j_1, j_2)$, si ha che $\beta_{j_1j_2}^{(k_1)(k_2)}$ è elemento di $\Phi_t = Y_tEX_t$. Di conseguenza ogni elemento di Φ_j è l'opposto del primo elemento dell'ultima riga di una delle matrici $H_{j_1j_2}^{k_1k_2}(0)$. Questa osservazione sarà essenziale per provare il risultato principale del Teorema.

Utilizzando la stessa indicizzazione di cui sopra, consideriamo ora i blocchi diagonali

di dimensione n_j di $F(\mu,0)$ (cioè ottenuti ponendo $j_1 = j_2 = j$ e $k_1 = k_2 = k$). Abbiamo

$$F_{jj}^{kk}(\mu,0) = \begin{pmatrix} \mu & -1 & 0 & \cdots & 0 \\ 0 & \mu & -1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \mu & -1 \\ -\beta_{jj}^{(k)(k)} & 0 & \cdots & 0 & \mu \end{pmatrix}.$$

Inoltre, per le proprietà esposte sopra, i blocchi di $F(\mu,0)$ che si trovano negli stessi indici di colonna di ogni $F_{jj}^{kk}(\mu,0)$ possono contenere come solo elemento diverso da 0 quello nell'angolo in basso a sinistra (il $-\beta_{j_1 j}^{k_1 k}$). Infatti quei blocchi si ottengono variando gli indici j_1 e k_1 e lasciando immutati $j_2 = j$ e $k_2 = k$, quindi, il numero di elementi diversi da 0 nella forma di $H_{j_1 j}^{k_1 k}$ presentata sopra rimarrebbe sempre $n_{j_2} - n_j = n_j - n_j = 0$. Per i blocchi $F_{jj}^{kk}(\mu,0)$, quindi, possiamo pensare di sommare alla prima colonna la seconda moltiplicata per μ, poi la terza moltiplicata per μ^2, e così via fino all'ultima (la n_j-sima), moltiplicata per μ^{n_j-1}. Questa operazione non cambia il determinante $\mathcal{Q}(\mu,0)$ per le proprietà del determinante, e, per l'osservazione appena fatta, ha come solo risultato quello di sostituire i blocchi $F_{jj}^{kk}(\mu,0)$ con blocchi

$$\begin{pmatrix} 0 & -1 & 0 & \cdots & 0 \\ 0 & \mu & -1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \mu & -1 \\ -\beta_{jj}^{(k)(k)} + \mu^{n_j} & 0 & \cdots & 0 & \mu \end{pmatrix}.$$

Possiamo ora dimostrare che

$$\mathcal{Q}(\mu,0) = \pm\mu^{\alpha}\det(\Phi_j - \mu^{n_j}\mathcal{I}_j),$$

per un appropriato $\alpha > 0$. Questo ci permetterà di concludere la dimostrazione del risultato principale del Teorema. La dimostrazione segue da un calcolo diretto, reso possibile ora che conosciamo in modo puntuale la struttura a blocchi di $F(\mu,0)$. La strategia da utilizzare per il calcolo consiste nell'andare a cercare le righe della matrice che presentano un solo elemento diverso da 0 (che si trova nelle prime o nelle ultime righe dei blocchi diagonali, dopo averli eventualmente sostituiti nella maniera descritta appena sopra). Sviluppando il determinante su quelle righe si trovano matrici più piccole ma con struttura esattamente analoga alla precedente, che avranno ancora righe con un solo elemento diverso da 0. Ripetendo il procedimento si arriverà alla fine a dover calcolare il determinante di una matrice composta da soli elementi $-\beta_{j_1 j_2}^{(k_1)(k_2)}$ e $-\beta_{jj}^{(k)(k)} + \mu^{n_j}$. Questi

elementi sono gli opposti di elementi delle matrici Φ_j e $\Phi_j - \mu^{n_j} I$, e poiché i blocchi $F^{kk}_{\bar{j}\bar{j}}$ con $\bar{j} > j$ non presentano elementi β (perché $H^{kk}_{\bar{j}\bar{j}}(0) = 0$), gli unici elementi non 0 nelle righe successive a quelle dei blocchi F^{kk}_{jj} sono nei blocchi diagonali, e risulta che tutti gli elementi di $F(\mu, 0)$ con indice di riga o di colonna colonna m maggiore di f_j vengono tutti eliminati nel processo di sviluppo del determinante. Il blocco fatto dai $-\beta^{(k)(k)}_{jj} - \mu^{n_j}$ nella matrice finale si sposta quindi in basso a destra, giustificando la presenta del termine "$-\mu^{n_j}\mathcal{I}_j$" nella formula che vogliamo dimostrare. La presenza del fattore $\pm\mu^\alpha$ è dovuta al fatto che l'elemento secondo il quale si fanno gli sviluppi successivi può essere -1, oppure μ. È quindi necessario aggiungere questa eventuale correzione moltiplicativa di fattore $\pm\mu^\alpha$ per tenere conto di questa variabilità.

Visualizziamo meglio la strategia in azione con un esempio mirato, proposto da Moro, Burke e Overton [8].

Siano $q = 3, \quad j = 2, \quad n_1 = 4, \quad n_2 = 3, \quad n_3 = 2, \quad r_1 = 1, \quad r_2 = 2, \quad r_3 = 1$. Allora $\mathcal{Q}(\mu, 0)$, in questa circostanza, è il determinante della matrice

$$\left(\begin{array}{cccc|cccc|cc|cc}
0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \mu & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \mu & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-\beta^{(1)(1)}_{11} & * & 0 & \mu & -\beta^{(1)(1)}_{12} & 0 & 0 & -\beta^{(1)(2)}_{12} & 0 & 0 & -\beta^{(1)(1)}_{13} & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \mu & -1 & 0 & 0 & 0 & 0 & 0 \\
-\beta^{(1)(1)}_{21} & * & 0 & 0 & -\beta^{(1)(1)}_{22} + \mu^3 & 0 & \mu & -\beta^{(1)(2)}_{22} & 0 & 0 & -\beta^{(1)(1)}_{23} & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mu & -1 & 0 & 0 \\
-\beta^{(2)(1)}_{21} & * & 0 & 0 & -\beta^{(2)(1)}_{22} & 0 & 0 & -\beta^{(2)(2)}_{22} + \mu^3 & 0 & \mu & -\beta^{(2)(1)}_{23} & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mu & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mu
\end{array}\right).$$

Sviluppando rispetto alla prima riga, poi rispetto alla quinta, all'ottava, alla dodicesima, otteniamo che il determinante in questione è il prodotto del fattore μ con il determinante di una matrice con 8 righe e 8 colonne. Sviluppando rispetto alla prima, poi alla quarta, alla sesta, e all'ottava riga di questa nuova matrice otteniamo che $\mathcal{Q}(\mu, 0)$ è il prodotto del fattore μ^2 con il determinante di una matrice con 4 righe e 4 colonne. Sviluppando rispetto alla prima riga di questa ultima matrice otteniamo come risultato finale

$$\mu^2 \det \begin{pmatrix}
-\beta^{(1)(1)}_{11} & -\beta^{(1)(1)}_{12} & -\beta^{(1)(2)}_{12} \\
-\beta^{(1)(1)}_{21} & -\beta^{(1)(1)}_{22} + \mu^3 & -\beta^{(1)(2)}_{22} \\
-\beta^{(2)(1)}_{21} & -\beta^{(2)(1)}_{22} & -\beta^{(2)(2)}_{22} + \mu^3
\end{pmatrix},$$

che è proprio $+\mu^2 \det(\Phi_2 + \mu^3 \mathcal{I}_2)$.

Possiamo finalmente utilizzare la nostra ipotesi di non singolarità di Φ_{j-1}, e concludere la dimostrazione della prima parte del Teorema. Infatti, considerando μ^{n_j} come una variabile, possiamo vedere la matrice $\Phi_j + \mu^{n_j}\mathcal{I}_j$ come la matrice a blocchi $\begin{pmatrix} \Phi_j^{11} & \Phi_j^{12} \\ \Phi_j^{21} & \Phi_j^{22} + \mu^{n_j} I_{r_j} \end{pmatrix}$ da cui ricaviamo che, a patto che $\det(\Phi_j^{11}) \neq 0$, $\mathcal{Q}(\mu, 0)$ è un polinomio di grado r_j non nullo in μ^{n_j}. Ma per come abbiamo costruito le matrici Φ, Φ_j^{11} è proprio Φ_{j-1}, e quindi l'ipotesi di non singolarità ci permette di verificare l'uguaglianza. Nel caso $j = 1$ questo ultimo risultato è banale, perché il determinante voluto corrisponde al calcolo di un polinomio caratteristico. Consideriamo ora il polinomio $\mathcal{Q}(\mu, z)$, e vediamolo come polinomio nelle variabili μ i cui coefficienti sono funzione (continua) di z. Poiché le radici di un polinomio dipendono in modo continuo dai suoi coefficienti (si veda il Teorema 1.19) e stiamo lavorando in un intorno di $z = 0$, questo ci garantisce l'esistenza di $r_j \cdot n_j$ radici di $\mathcal{Q}(\mu, z)$,

$$\mu_j^{(k)(l)} = (\xi_j^{(k)})^{1/n_j} + o(1), \qquad k = 1, \dots, r_j, \quad l = 1, \dots, n_j,$$

e ritornando alle variabili originali (λ, ε), otteniamo

$$\lambda_j^{(k)(l)} = \lambda + (\xi_j^{(k)})^{1/n_j} \varepsilon^{1/n_j} + o(\varepsilon^{1/n_j}), \qquad k = 1, \dots, r_j, \quad l = 1, \dots, n_j,$$

che era quello che volevamo.

L'ultima parte del Teorema si ricava immediatamente, perché l'ipotesi aggiuntiva che le $\xi_j^{(k)}$ siano tutte distinte al variare di k ci permette di applicare il Teorema delle funzioni implicite 1.17 e ricavare che le $\mu_j^{(k)(l)}$ sono funzioni analitiche di z, e quindi si espandono in serie di potenze. □

Questo Teorema, e la sua dimostrazione, forniscono risultati estremamente potenti sulla perturbazione degli spettri delle matrici, e la condizione Φ_{j-1} non singolare risulta essere soddisfatta nella maggioranza dei casi, rendendo il risultato quasi totalmente generale. Ciò che possiamo osservare subito, e che deriva immediatamente dai risultati, è che nel caso generale il problema della perturbazione degli spettri risulta essere molto malcondizionato, perché la potenza alla quale viene elevato il parametro ε è frazionaria. Questo ce lo aspettavamo, ma ora possiamo dare un'informazione molto più puntuale su quanto effettivamente sia malcondizionato il problema. Sappiamo infatti che l'esponente della ε dipende solo dalla dimensione e dal numero dei blocchi di Jordan dell'autovalore perturbato: più la matrice si avvicina ad essere diagonale (abbassando n_j) più gli spettri diventeranno stabili, e ciò accade anche diminuendo il numero di blocchi di Jordan uguali (abbassando r_j), portando l'autovalore più vicino alla semplicità. Questo risolve a pieno

ciò che ci era rimasto dalla Sezione 3.2, e nel caso di matrici diagonalizzabili conferma i risultati che avevamo già trovato utilizzando i Teoremi di Gerschgorin.

Un'altra importante informazione del Teorema consiste nel consentirci di visualizzare subito quali siano gli elementi della matrice di perturbazione $\varepsilon\tilde{E} = \varepsilon P^{-1}EP$ che contribuiscono maggiormente a modificare gli spettri. Questo a patto di conoscere una base per la forma di Jordan di A, data dalle colonne di P. La dimostrazione infatti mostra chiaramente come gli unici elementi che effettivamente formano il coefficiente di ε^{1/n_j} siano quelli negli angoli in basso a sinistra dei blocchi $\tilde{E}_{j_1 j_2}^{k_1 k_2}$ (quelli indicati con $\beta_{j_1 j_2}^{(k_1)(k_2)}$). Tutti gli altri elementi hanno un ruolo decisamente minore nella modifica degli autovalori, formando solo i coefficienti delle ε con esponenti più grandi, e saranno rilevanti solo se qualcuna delle $\xi_j^{(k)}$ è nulla (e quindi solo se Φ_j è singolare). In generale, basta quindi conoscere solo pochi elementi della matrice di perturbazione per avere tutte le informazioni sul problema, pagando il costo di dover calcolare preventivamente la forma di Jordan di A e le matrici Φ_j e Φ_{j-1}.

Teorema 4.2 (Lidskii, autovettori)**.** *Nelle notazioni precedenti, sia Φ_s non singolare per ogni $s = 1, \dots q$. Sia $j \in \{1, \dots, q\}$ fissato. Se le soluzioni dell'equazione*

$$\det(\Phi_j - \xi\mathcal{I}_j) = 0$$

sono tutte distinte, allora si può scegliere ε abbastanza piccolo per cui gli autovalori perturbati di cui al Teorema 4.1 siano tutti semplici. Inoltre gli autovettori destri associati a tali autovalori possono essere espressi tramite una espansione, in termini di ε, del tipo

$$v_j^{(k)(l)}(\varepsilon) = u_j^{(k)} + \sum_{s=1}^{\infty} w_{j\cdot s}^{(k)(l)} \varepsilon^{s/n_j}, \qquad l = 1, \dots, n_j,\ k = 1, \dots, r_j,$$

dove

$$u_j^{(k)} = \sum_{p=1}^{f_j} c_j^{(k)(p)} x_j^p,$$

e il vettore

$$c_j^k = \begin{pmatrix} c_j^{(k)(1)} \\ \vdots \\ c_k^{(k)(f_j)} \end{pmatrix}$$

è un elemento di $\mathrm{Ker}(\Phi_s - \xi_j^{(k)}\mathcal{I}_s)$.

Dimostrazione. Supponiamo ancora che A abbia solo un autovalore. Per quanto riguarda la prima parte del Teorema, chiamiamo la particolare j dell'enunciato $\bar{j}$. Il fatto che le

$\xi_{\bar{j}}^k$ siano tutte distinte al variare di k assicura che, a patto di scegliere un valore di ε sufficientemente piccolo, gli $r_{\bar{j}} \cdot n_{\bar{j}}$ autovalori perturbati del Teorema 4.1 sono tutti distinti tra loro. Scegliendo un altro valore per j, poi, le condizioni del Teorema 4.1 rimangono soddisfatte, in quanto abbiamo supposto che le Φ_s, $s = 1, \ldots, q$ siano tutte non singolari, e quindi si ottengono sviluppi in serie per tutti gli altri autovalori di $\tilde{A}$. Non abbiamo strumenti per dire che tutti gli autovalori siano distinti, e in generale questo sarà falso, ma dato che otteniamo un finito set di autovalori perturbati, al variare del valore di j, e che ogni set dipende da una potenza diversa di ε (per ogni set si avrà che l'espansione al primo ordine contiene il termine ε^{1/n_j}, continuando a variare $j = 1, \ldots, q$), sarà possibile scegliere un valore di ε abbastanza piccolo in modo che nessun autovalore perturbato ottenuto con $j \neq \bar{j}$ coincida con gli autovalori ottenuti con $j = \bar{j}$. Dal momento che questi sono tutti gli autovalori dell'operatore perturbato (perché sono $n = \sum_{j=1}^{q} r_j n_j$, che è la dimensione delle matrici A e $\tilde{A}$), abbiamo così dimostrato che gli autovalori ottenuti per $j = \bar{j}$ sono semplici.

Consideriamo ora, fissato un autovalore $\lambda_{\bar{j}}^{(k)(l)}$, la matrice $\tilde{A} - \lambda_{\bar{j}}^{(k)(l)} I$. Un suo autovettore si può trovare considerando i cofattori degli elementi di una delle righe della matrice stessa, scelta appositamente. Infatti, scegliamo una riga p della matrice. Il vettore che ha come primo elemento il cofattore dell'elemento di posto $(p, 1)$ di $\tilde{A} - \lambda_{\bar{j}}^{(k)(l)} I$, come secondo elemento il cofattore dell'elemento di posto $(p, 2)$ e così via è non nullo. Infatti, dato che l'autovalore è semplice, la matrice $\tilde{A} - \lambda_{\bar{j}}^{(k)(l)} I$ ha un nucleo di dimensione 1, e questo significa che esiste almeno un minore di ordine $n-1$ non nullo. Basterà allora scegliere la riga p tale che il minore non nullo si ottenga rimuovendo quella riga e una delle colonne. Chiamato v questo vettore, si avrà che moltiplicando ogni riga di $\tilde{A} - \lambda_{\bar{j}}^{(k)(l)} I$ per v diversa dalla $\bar{j}$-sima, si ottiene un multiplo del determinante di una matrice con due righe uguali, che è 0. Moltiplicando la $\bar{j}$-sima riga di $\tilde{A} - \lambda_{\bar{j}}^{(k)(l)} I$ per v si ottiene invece un multiplo del determinante della matrice stessa, che è sempre 0. Ma per semplicità dell'autovalore questo è l'unico autovettore relativo, a meno di moltiplicazione per scalare. Abbiamo quindi dimostrato che le componenti dell'autovettore che stiamo cercando sono funzioni analitiche di ε. Ora rinominiamo le quantità, tenendo conto che $\tilde{A} - \lambda_{\bar{j}}^{(k)(l)} I = A + \varepsilon E - \lambda_{\bar{j}}^{(k)(l)} I$, in modo simile a come avevamo fatto per la dimostrazione precedente.

$$\begin{cases} z = \varepsilon^{1/n_{\bar{j}}} \\ \mu_{\bar{j}}^{(k)(l)} = \frac{\lambda_{\bar{j}}^{(k)(l)} - \lambda}{z} \end{cases}$$

e siano $\mathcal{P}$, F, L, R definite come sopra. Definiamo

$$F_{\bar{j}}^{kl}(z) = F(\mu_{\bar{j}}^{(k)(l)}(z), z),$$

e sia $e_{\bar{j}}^{kl}(z)$ un vettore nel nucleo di questa ultima matrice. Allora

$$F_{\bar{j}}^{kl}(z)e_{\bar{j}}^{kl}(z) = L(z)\mathcal{P}(\mu_{\bar{j}}^{(k)(l)}(z), z)R(z)e_{\bar{j}}^{kl}(z) = 0.$$

Ma per $z \neq 0$ $L(z)$ è non singolare, da cui $R(z)e_{\bar{j}}^{kl}(z)$ è un autovettore destro di $J + z^{n_{\bar{j}}}\tilde{E}$ associato a $\lambda_{\bar{j}}^{(k)(l)}$. Per l'analiticità che abbiamo mostrato prima, quindi, nell'espansione in serie di ε dell'autovettore il primo termine, $R(0)e_{\bar{j}}^{kl}(0)$, sarà un autovettore relativo a λ, e quindi combinazione lineare degli $x_{\bar{j}}^{p}$. Per ottenere la formula sui coefficienti $c_{\bar{j}}^{(k)(p)}$ si applicano costruzioni e considerazioni completamente analoghe a quelle della dimostrazione precedente al sistema lineare $F(0)e_{\bar{j}}^{kl}(0) = 0$. □

Anche in questo caso otteniamo informazioni essenziali sia sul condizionamento del problema, sia sugli elementi della matrice di perturbazione che effettivamente vanno a modificare gli autovettori di partenza. La condizione che imponiamo in questo contesto, però, è più forte della precedente, e ci è necessaria per poter ottenere risultati così simili a quelli che avevamo ottenuto per gli autovalori. Come già notato, infatti, la perturbazione degli autovettori è sempre un problema più complicato da analizzare, a causa dell'eventuale presenza di autovalori multipli che comportano la presenza di autovettori di direzione non ben definita. Le condizioni che abbiamo posto all'inizio dell'enunciato ci permettono di bypassare questo problema di buona definizione degli autovettori, perché ci consentono di lavorare con autovalori perturbati semplici, che hanno autovettori ben delineati, a meno di una moltiplicazione per scalare. Risolviamo così anche il problema che incorreva nello studio della perturbazione degli autovettori che avevamo svolto nella Sezione 3.2, trovando un'espansione precisa che ci permette di confermare i risultati che avevamo trovato.

4.2 I diagrammi di Newton

Vogliamo dare un nuovo ultimo approccio che permetta di trovare dei risultati anche nel caso in cui qualche Φ_s sia singolare. Per fare questo ci avvarremo, come accennato all'inizio del Capitolo, dei diagrammi di Newton, che danno una rappresentazione grafica delle radici dei polinomi caratteristici che ci interessano.

I diagrammi di Newton sono intrinsecamente legati a dei polinomi, e pertanto nel nostro caso andranno definiti a partire dal polinomio caratteristico della matrice perturbata. In generale, sia p un polinomio in due variabili, che scriviamo come

$$p(\lambda, \varepsilon) = \lambda^n + \alpha_1(\varepsilon)\lambda^{n-1} + \cdots + \alpha_n(\varepsilon),$$

dove i coefficienti delle potenze di λ sono funzioni analitiche nella variabile ε, non tutte nulle. Riscriviamo il polinomio come

$$p(\lambda, \varepsilon) = \lambda^n + (\hat{\alpha}_1 \varepsilon^{a_1} + \cdots)\lambda^{n-1} + \cdots + (\hat{\alpha}_{n-1}\varepsilon^{a_{n-1}} + \cdots)\lambda + (\hat{\alpha}_n \varepsilon^{a_n} + \cdots),$$

dove le $\hat{\alpha}_j$ sono i coefficienti relativi alla potenza più bassa di ε nel polinomio α_j, che abbiamo chiamato a_j. È possibile dimostrare che per un siffatto polinomio, visto nella sola variabile λ, ogni radice può essere espressa come la somma di una opportuna serie a esponenti non negativi e frazionari di ε [10]. Ipotizziamo quindi che una radice del polinomio sia

$$\lambda = \mu_1 \varepsilon^{\beta_1} + \mu_2 \varepsilon^{\beta_2} + \cdots,$$

dove i coefficienti μ_j sono non nulli e gli esponenti delle ε sono ordinati in modo da avere $\beta_1 < \beta_2 < \cdots$. Stiamo allora ipotizzando che

$$\begin{aligned} 0 \;= &(\mu_1^n \varepsilon^{n\beta_1} + \cdots) + (\hat{\alpha}_1 \varepsilon^{a_1} + \cdots)(\mu_1^{n-1} \varepsilon^{(n-1)\beta_1} + \cdots) + \cdots \\ &+ (\hat{\alpha}_{n-1}\varepsilon^{a_{n-1}} + \cdots)(\mu_1 \varepsilon^{\beta_1} + \cdots) + (\hat{\alpha}_n \varepsilon^{a_n} + \cdots), \end{aligned}$$

dove in ogni fattore sostituito alle potenze di λ è stato evidenziato il termine con l'esponente più piccolo per ε. Tutti i coefficienti di tutte le potenze di ε nel membro di destra, quindi, devono annullarsi, in particolare il coefficiente del termine con esponente più piccolo. Poiché tutti i coefficienti presenti sono non nulli per ipotesi, questo sarà possibile solo se nel membro a destra sono presenti almeno due addendi con lo stesso esponente per la ε. L'esponente più piccolo, per costruzione, andrà cercato tra i numeri

$$n\beta_1, a_1 + (n-1)\beta_1, a_2 + (n-2)\beta_1, \ldots, a_{n-1} + \beta_1, a_n. \tag{4.1}$$

Ora, in un piano cartesiano, tracciamo i punti $A_n = (n, a_n)$, $A_{n-1} = (n-1, a_{n-1})$, $\ldots$, $A_1 = (1, a_1)$, e il punto $A_0 = (0, 0)$, che corrisponderebbe appunto all'esponente più piccolo della variabile ε nel coefficiente di λ^n. Se per un k tra 1 e n succedesse che il coefficiente di λ^{n-k} è il polinomio nullo (in ε), allora quel punto verrebbe scartato. Possiamo tradurre il problema di trovare un esponente β_1 appropriato per una radice del polinomio in una costruzione cartesiana su questo piano. Introduciamo un nuovo parametro variabile: un angolo τ, e poniamo $\beta_1 = \tan\tau$. Sul piano cartesiano, τ identifica una direzione, e quindi un fascio di rette oblique. Possiamo osservare che, fissato τ, la quantità $a_k + (n-k)\tan\tau$ corrisponde alla lunghezza del segmento tagliato sulla retta verticale $x = n$ dalla retta $y = 0$ e la retta obliqua del fascio identificato da τ passante per A_k. Poiché abbiamo osservato che β_1 deve essere tale da far comparire almeno due volte uno dei numeri in (4.1), l'osservazione precedente ci dice che un τ ammissibile dovrà identificare un fascio di rette per cui almeno due dei punti A_k giacciono su una stessa

retta r del fascio. Inoltre, poiché τ deve anche identificare la più piccola tra le quantità di (4.1), abbiamo come condizione aggiuntiva che nessun punto tra quelli che abbiamo costruito sul piano può giacere al di sotto della retta r.
Una strategia per trovare tutti i τ (e quindi i β_1) ammissibili sarà di partire dal punto A_0 e dalla semiretta orizzontale per quel punto che forma un angolo nullo con l'asse x, fare ruotare tale semiretta in senso antiorario finché non tocca uno dei punti $A_k,\ k \geq 1$, e chiamare tale punto $A_{\bar{k}}$. In caso la semiretta identificata in questo modo toccasse almeno tre punti, considereremo come $A_{\bar{k}}$ il punto di ascissa maggiore tra quelli che appartengono alla retta. Possiamo poi spostarci su $A_{\bar{k}}$ e ripetere il procedimento, fino ad arrivare al punto A_n, e le inclinazioni delle rette così ottenute saranno i possibili valori di τ che stavamo cercando. Definiamo, per semplicità, un vettore ξ le cui coordinate sono, in ordine e a partire da A_0, le ascisse dei punti origine delle varie semirette.

Il diagramma di Newton del polinomio $p(\lambda, \varepsilon)$ sarà quindi dato dai punti $A_{\xi(k)}$ così costruiti, e dai segmenti S_k che connettono i punti $A_{\xi(k)}$ e $A_{\xi(k+1)}$. Per ogni segmento S_k abbiamo un β_1 determinato, e questo β_1 porta immediatamente ad avere una condizione per trovare i possibili valori di μ_1. Essi dovranno soddisfarre la seguente equazione, nella variabile μ:

$$\sum_{(j,a_j)\in S_k} \mu^{n-j}\hat{\alpha}_j = 0. \tag{4.2}$$

L'equazione avrà sempre una soluzione, il che significa che ogni coefficiente angolare di un segmento del diagramma di Newton è associato a una espansione per radici di $p(\lambda, \varepsilon)$.

Notiamo inoltre che, per costruzione, risulterà che l'angolo τ relativo a uno dei segmenti S_k (e quindi il β_1 corrispondente) sarà strettamente minore dell'angolo relativo a un qualsiasi altro segmento $S_{k+h},\ h > 0$. Avendo ora una descrizione dei termini di grado più basso possibili nello sviluppo delle radici di $p(\lambda, \varepsilon)$ si potrebbe continuare in modo analogo per trovare condizioni simili su β_2 e μ_2, in modo ricorsivo, ma l'analisi al primo ordine risulta già sufficiente per i nostri scopi.

Ora che abbiamo stabilito la situazione di base ritorniamo al nostro caso specifico, dove $p(\lambda, \varepsilon)$ è il polinomio caratteristico di una matrice A perturbata da una matrice εE. Come per la Sezione precedente, assumeremo anche qui che A abbia un solo autovalore, di molteplicità algebrica m e la cui struttura di Jordan sia spezzabile nei blocchi descritti nella Sezione 4.1 (Con $J = \mathrm{Diag}(\Gamma_1^1, \ldots, \Gamma_1^{r_1}, \ldots, \Gamma_q^1, \ldots, \Gamma_q^{r_q})$, dove Γ_i^k è un blocco di Jordan di dimensione n_i, e $n_1 > n_2 > \cdots > n_q$). Reintroduciamo anche la notazione $f_k = \sum_{j=1}^{k} r_j$. A meno di traslare tutto il sistema, assumeremo anche che questo unico

autovalore di A sia 0.

Partendo dalla matrice A e da una matrice qualsiasi E, vogliamo ora trovare un particolare sistema di punti e di segmenti tra di essi nel piano cartesiano, dipendente solo da A, tale che il diagramma di Newton del polinomio di $A+\varepsilon E$ non possa contenere alcuna parte al di sotto di questo sistema. Chiameremo il sistema di punti in questione "inviluppo di Newton", e ci rivelerà importanti informazioni per collegare tutte queste costruzioni al problema di perturbazione di autovalori e autovettori che stiamo studiando. La nostra strategia sarà di fissare $l > 0$ e cercare il più grande $k = k(l)$ per cui possa esistere una perturbazione di A tale che $a_{k(l)} = l$. Questo corrisponde effettivamente a cercare il punto "più a destra" di altezza fissata, appartenente ad un diagramma. Enunciamo il Teorema:

Teorema 4.3. *Nelle notazioni introdotte in precedenza, per ogni $l = 1, \dots, f_q$, il corrispondente $k(l)$ è pari alla somma delle dimensioni degli l blocchi di Jordan Γ_j^k di dimensione più grande. Più precisamente, denotando $f_0 = 0$, abbiamo che, fissato $j \in \{1, \dots, q\}$, se $l = f_{j-1} + \rho$ con $0 < \rho \leq r_j$ allora si ha*

$$k(l) = r_1 n_1 + r_2 n_2 + \cdots + r_{j-1} n_{j-1} + \rho n_j.$$

Inoltre il coefficiente di ε^l nel polinomio $\alpha_{k(l)}$ è pari a $(-1)^l$ moltiplicato per la somma di tutti i minori principali di Φ_j che corrispondono a determinanti di sottomatrici di Φ_j di dimensione l e che contengono Φ_{j-1}. Nel caso $j = 1$ questo corrisponde a considerare tutti i minori principale di dimensione l di Φ_1. In particolare, per $l = f_j$, il coefficiente $\hat{\alpha}_{k(l)}$ è $(-1)^l \det(\Phi_j)$.

Dimostrazione. La dimostrazione segue da un calcolo diretto di $\det(\lambda I - A - \varepsilon E)$ (dove qui λ è una variabile), e da alcune considerazioni che ci permettono di individuare i termini di nostro interesse. Innanzitutto, la matrice di cui calcoliamo il determinante è $\lambda I - J - \varepsilon \tilde{E}$, dove ricordiamo $J = P^{-1}AP$ e $\tilde{E} = P^{-1}EP$. La matrice ha dimensioni $m \times m$, e il calcolo del suo determinante prevede di moltiplicare tra loro m elementi di questa matrice, ognuno scelto da una riga e da una colonna differenti, e sommare o sottrarre tra loro tutti i prodotti derivanti da scelte diverse per gli m fattori. Se vogliamo che uno di questi prodotti sia dell'ordine di ε^l, allora dovremo scegliere l elementi della matrice che contengono una ε e $m - l$ elementi che non sono moltiplicati per ε. Per la struttura della nostra matrice questi elementi sono λ oppure -1, perché derivano da λI, oppure da $-J$ (ricordiamo che abbiamo supposto che A abbia un solo autovalore pari a 0). Scegliamo gli $m - l$ elementi diversi da un coefficiente moltiplicato per ε. Dal momento che il $k = k(l)$ che stiamo cercando deve essere il più grande possibile, dovremo

cercare di includere quante meno λ possibile tra questi $m-l$ fattori. Cerchiamo allora di massimizzare il numero di fattori -1. Supponiamo che, per una scelta ammissibile di m fattori, vi siano γ fattori pari a -1. Questi fattori condizionano la scelta di tutti gli altri, dal momento che sceglierne uno ci impedisce di scegliere qualsiasi altro elemento dalla stessa riga o colonna. Ora siamo interessati al numero di fattori λ che vengono esclusi dalla scelta dei fattori -1. Questo dipende unicamente dal numero di blocchi di Jordan da cui scegliamo i -1. Infatti, se un fattore -1 è il primo che scegliamo in un certo blocco, esso escluderà in ogni caso due fattori λ, mentre se non è il primo, può essere scelto in modo da escludere un solo ulteriore fattore λ (basta sceglierlo in una riga accanto a quella del primo fattore -1 scelto in quel blocco). Supponiamo che i fattori -1 siano stati scelti da δ blocchi di Jordan distinti. Allora verranno esclusi almeno $\gamma+\delta$ fattori λ da quelli che si possono scegliere. Ma comunque, nella combinazione ammissibile che stiamo considerando, dato che ci sono $m-l$ fattori che non contengono il termine ε e γ di questi sono -1, i restanti $m-l-\gamma$ devono essere fattori λ. In totale, nella matrice $\lambda I - J - \varepsilon\tilde{E}$, ci sono m fattori λ disponibili. La somma del minimo di fattori esclusi e di quelli inclusi non può superare questa soglia. Otteniamo quindi

$$\begin{aligned} &(\gamma+\delta)+(m-l-\gamma)\leq m \\ \Rightarrow\ &\delta \leq l. \end{aligned}$$

Abbiamo così dimostrato che i fattori -1 possono provenire al massimo da l blocchi di Jordan.

Ora, nel caso fosse $l=f_j$ per qualche j, questo comporta che ci sia una sola possibile scelta ammissibile. Bisognerà infatti scegliere i fattori -1 dai primi l blocchi di Jordan più grandi, ottenendo così $\gamma = r_1(n_1-1)+\cdots+r_j(n_j-1)$ fattori -1, e scegliere poi tutte le λ rimanenti, che provengono dai restanti blocchi di Jordan. Otteniamo così

$$\begin{aligned} &r_1(n_1-1)+\cdots+r_j(n_j-1)+r_{j+1}n_{j+1}+\cdots+r_qn_q \\ =&\sum_{i=1}^{q} r_in_i - \sum_{i=1}^{j} r_i \\ =&\, m-l \end{aligned}$$

fattori che non contengono una ε, e quindi l fattori che la contengono. Questa è l'unica scelta possibile per la limitazione del numero di blocchi di Jordan da cui possono provenire i fattori -1. Se si prendessero nello stesso numero ma in modo diverso bisognerebbe aumentare il numero di blocchi di Jordan da cui essi provengono, e scegliendo un numero minore di -1 si dovrebbero aumentare le λ presenti nel prodotto, riducendo quindi

$k = k(l)$. La nostra soluzione è quindi l'unica ottimale, e ci permette di ottenere

$$k(l) = m - \left(\sum_{i=j+1}^{q} r_i n_i \right) = \sum_{i=1}^{j} r_i n_i,$$

come volevamo. Inoltre, eliminando le righe e le colonne corrispondenti a questa scelta per i fattori -1 e i fattori λ, la matrice rimanente ha come elementi solo quei $\beta_i^{(p)}$ che avevamo definito nella dimostrazione del Teorema 4.1, e che avevamo dimostrato essere gli elementi della matrice Φ_j, moltiplicati per un fattore $-\varepsilon$. Questa matrice ha anche dimensione f_j, quindi è dimostrata la seconda parte dell'enunciato secondo cui il coefficiente $\hat{\alpha}_{k(l)}$ deve essere $(-1)^l \det(\Phi_j)$. Per visualizzare meglio questo fatto, vediamolo con un esempio in cui $n_1 = 3$, $n_2 = 2$, $r_1 = 2$, $r_2 = 1$ e $l = f_1$. Nell'esempio, indicheremo con un asterisco gli elementi della matrice $\tilde{E}$ che non sono delle $\beta_i^{(p)}$, coloreremo in rosso i fattori -1 scelti, e in blu i fattori λ scelti. Con queste impostazioni, il polinomio caratteristico è il determinante della seguente matrice:

$$\left(\begin{array}{ccc|ccc|cc}
\lambda - \varepsilon * & -1 - \varepsilon * & -\varepsilon * & -\varepsilon * & -\varepsilon * & -\varepsilon * & -\varepsilon * & -\varepsilon * \\
-\varepsilon * & \lambda - \varepsilon * & -1 - \varepsilon * & -\varepsilon * & -\varepsilon * & -\varepsilon * & -\varepsilon * & -\varepsilon * \\
-\varepsilon \beta_{11}^{(1)(1)} & -\varepsilon * & \lambda - \varepsilon * & -\varepsilon \beta_{11}^{(1)(2)} & -\varepsilon * & -\varepsilon * & -\varepsilon \beta_{12}^{(1)(1)} & -\varepsilon * \\
\hline
-\varepsilon * & -\varepsilon * & -\varepsilon * & \lambda - \varepsilon * & -1 - \varepsilon * & -\varepsilon * & -\varepsilon * & -\varepsilon * \\
-\varepsilon * & -\varepsilon * & -\varepsilon * & -\varepsilon * & \lambda - \varepsilon * & -1 - \varepsilon * & -\varepsilon * & -\varepsilon * \\
-\varepsilon \beta_{11}^{(2)(1)} & -\varepsilon * & -\varepsilon * & -\varepsilon \beta_{11}^{(2)(2)} & -\varepsilon * & \lambda - \varepsilon * & -\varepsilon \beta_{12}^{(2)(1)} & -\varepsilon * \\
\hline
-\varepsilon * & -\varepsilon * & -\varepsilon * & -\varepsilon * & -\varepsilon * & -\varepsilon * & \lambda - \varepsilon * & -1 - \varepsilon * \\
-\varepsilon \beta_{21}^{(1)(1)} & -\varepsilon * & -\varepsilon * & -\varepsilon \beta_{21}^{(1)(2)} & -\varepsilon * & -\varepsilon * & -\varepsilon \beta_{22}^{(1)(1)} & \lambda - \varepsilon *
\end{array}\right).$$

Eliminando le righe e le colonne corrispondenti agli elementi colorati rimane, appunto, la matrice

$$\begin{pmatrix} -\varepsilon \beta_{11}^{(1)(1)} & -\varepsilon \beta_{11}^{(1)(2)} \\ -\varepsilon \beta_{11}^{(2)(1)} & -\varepsilon \beta_{11}^{(2)(2)} \end{pmatrix} = -\varepsilon \Phi_1.$$

Nel caso fosse $\rho < r_{j+1}$ si applicano ragionamenti analoghi, ma non c'è più un solo modo per scegliere i fattori -1 e λ. Bisognerà sempre scegliere tutti i -1 dei primi f_{j-1} blocchi di Jordan, ma poi si possono scegliere liberamente ρ blocchi tra quelli di dimensione n_j, e prendere tutti i fattori -1 da essi. Ancora una volta, una volta scelti i -1, si completerà scegliendo tutte le λ rimanenti, ed eliminando le righe e le colonne relative ai -1 e ai λ si ottengono ancora elementi della matrice $-\varepsilon\Phi_j$. Questa volta però non si otterrà tutta la matrice $-\varepsilon\Phi_j$, ma solo una sua sottomatrice principale che contiene $-\varepsilon\Phi_{j-1}$, ancora una volta in accordo con l'enunciato del Teorema. □

Ora possiamo definire in modo più puntuale quello che abbiamo chiamato inviluppo di Newton.

Definizione 4.4. *Data una matrice A siano $\mathcal{P}_j = (k(f_j), f_j)$, e siano S_j i segmenti che giacciono su rette di coefficiente angolare $1/n_j$ che connettono $\mathcal{P}_{j-1}$ a $\mathcal{P}_j$, con $j = 1, \dots, q$. Definiamo l'inviluppo di Newton di A come il diagramma ottenuto tracciando sul piano cartesiano i punti $\mathcal{P}_j$ e i segmenti S_j.*

È facile verificare che l'inviluppo di Newton sia davvero un diagramma, dal momento che l'unica condizione che dobbiamo porre è che i coefficienti di tutti gli $\varepsilon^{a_{k(l)}}$ siano non nulli. Per quanto abbiamo visto nel Teorema 4.3 ci basta quindi avere una matrice E per cui $\det(\Phi_j) \neq 0$ per ogni j. Questo ci permette di fare un'osservazione molto importante: l'unico caso in cui l'inviluppo di Newton di una matrice A non coincide con il diagramma di Newton di un polinomio caratteristico dato da una sua perturbazione è proprio quando c'è qualche j per cui $\det(\Phi_j) = 0$. Questo ci apre la strada per dimostrare risultati simili a quelli della Sezione 4.1 in questi ultimi casi, che erano gli unici non ricoperti dalla teoria di Lidskii. Dovremo però richiedere una piccola condizione sulla struttura dei diagrammi, che risulta facilmente soddisfatta quasi sempre. Enunciamo, quindi, questo ultimo Teorema.

Teorema 4.5. *Fissiamo $j \in \{1, \dots, q\}$, e siano $0 \leq \gamma \leq r_j$ e $0 \leq \delta \leq r_{j+1}$. Nelle notazioni introdotte precedentemente, supponiamo che i punti $\mathcal{Q}^j_\gamma = (k(f_j - \gamma), f_j - \gamma)$ e $\hat{\mathcal{Q}}^j_\delta = (k(f_j + \delta), f_j + \delta)$ appartengano al diagramma di Newton di A perturbata con matrice E, e che i punti $\mathcal{Q}^j_s$ per $s = \gamma - 1, \gamma - 2, \dots, 1$, $\mathcal{P}_j$ e $\hat{\mathcal{Q}}^j_t$ per $t = 1, \dots, \delta - 1$ non appartengano al diagramma. Definiamo $p = \gamma n_j + \delta n_{j+1}$ e $\sigma = \frac{\gamma+\delta}{p}$. Se*

$$(\sigma n_j - 1)\gamma \leq \min(\sigma, 1 - \sigma), \tag{4.3}$$

allora ci sono p autovalori perturbati che si sviluppano nella serie

$$\lambda^{(l)}(\varepsilon) = \lambda + \eta^{1/p}\varepsilon^\sigma + o(\varepsilon^\sigma), \qquad l = 1, \dots, p,$$

con $\eta \neq 0$. Se inoltre vale una delle seguenti condizioni:

i. la disuguaglianza in (4.3) è stretta;

ii. $(\sigma n_j - 1)\gamma < \sigma$ e $\delta = n_{j+1} = 1$,

allora vale

$$\eta = -\frac{\hat{\alpha}_{k(f_j+\delta)}}{\hat{\alpha}_{k(f_j-\gamma)}}.$$

Dimostrazione. Dal momento che abbiamo supposto che $\mathcal{Q}^j_\gamma$ e $\hat{\mathcal{Q}}^j_\delta$ siano punti del diagramma di Newton di $A + \varepsilon E$, la porzione di diagramma che si trova nell'intervallo $[k(f_j - \gamma), k(f_j + \delta)]$ comprenderà punti e segmenti che si trovano tutti al di sotto della

corda che connette $\mathcal{Q}_\gamma^j$ e $\hat{\mathcal{Q}}_\delta^j$, o che giacciono sulla corda stessa. D'altronde, per definizione i due punti appartengono anche all'inviluppo di Newton di A, quindi i punti e i segmenti del diagramma che sono in quell'intervallo dovranno trovarsi al di sopra degli elementi dell'inviluppo presenti nell'intervallo stesso, o coincidere con i punti dell'inviluppo. Notiamo che il coefficiente $\sigma = \frac{\gamma+\delta}{p}$ dell'enunciato è pari al coefficiente angolare della corda che connette i due punti $\mathcal{Q}_\gamma^j$ e $\hat{\mathcal{Q}}_\delta^j$. Se riusciamo a dimostrare che la porzione del diagramma di Newton contenuta in $[k(f_j-\gamma), k(f_j+\delta)]$ giace tutta sulla corda, allora avremo finito, perché abbiamo già visto che l'inclinazione dei segmenti del diagramma di Newton corrisponde all'esponente più basso di una espansione a esponenti frazionari di ε per una radice del polinomio caratteristico di $\tilde{A}$. Per dimostrare questo fatto osserviamo innanzitutto che le ipotesi sui punti $\mathcal{Q}_s^j$, $s = \gamma-1, \gamma-2, \ldots, 1$, $\mathcal{P}_j$ e $\hat{\mathcal{Q}}_t^j$, $t = 1, \ldots, \delta-1$ ci assicurano che, nell'intervallo $[k(f_j-\gamma), k(f_j+\delta)]$, il diagramma di Newton e l'inviluppo di Newton coincidono solo nei punti $\mathcal{Q}_\gamma^j$ e $\hat{\mathcal{Q}}_\delta^j$. Dimostriamo poi che non esistono punti a coordinate intere nell'interno del triangolo descritto dai due segmenti dell'inviluppo di Newton e dalla corda. Per fare questo ci basta vedere che i punti $(k(f_j)-1, f_j)$ e $(k(f_j)+1, f_j+1)$ non siano in questa regione. Infatti questo condiziona l'ampiezza dell'angolo tra la corda e il segmento tra $\mathcal{Q}_\gamma^j$ e $\mathcal{P}_j$, e l'ampiezza dell'angolo tra la corda e il segmento tra $\mathcal{P}_j$ e $\hat{\mathcal{Q}}_\delta^j$, in modo che preso comunque un punto sulla corda e tracciata la retta orizzontale passante per quel punto, il segmento individuato dall'intersezione di questa retta con il triangolo non possa contenere punti del reticolo $\mathbb{Z} \times \mathbb{Z}$.

Ma la condizione sul punto $(k(f_j)-1, f_j)$ implica che il coefficiente angolare della retta che connette questo punto a $\mathcal{Q}_\gamma^j$ debba essere maggiore o uguale a σ, quindi

$$\begin{aligned}
\frac{f_j - (f_j-\gamma)}{k(f_j) - 1 - (k(f_j-\gamma))} &\geq \sigma \\
\frac{\gamma}{\gamma n_j - 1} &\geq \sigma \\
\gamma &\geq \sigma\gamma n_j - \sigma \\
\gamma(\sigma n_j - 1) &\leq \sigma.
\end{aligned}$$

Similmente la condizione sul punto $(k(f_j)+1, f_j+1)$ comporta che il coefficiente angolare della retta tra $\mathcal{Q}_\gamma^j$ e questo punto sia maggiore o uguale di σ, da cui

$$\begin{aligned}
\frac{(f_j+1) - (f_j-\gamma)}{(k(f_j)+1) - k(f_j-\gamma)} &\geq \sigma \\
\frac{1+\gamma}{1+\gamma n_j} &\geq \sigma \\
1+\gamma &\geq \sigma(1+\gamma n_j) \\
\gamma(\sigma n_j - 1) &\leq 1-\sigma,
\end{aligned}$$

Mettendo queste due condizioni insieme otteniamo proprio quello che compare nell'enunciato del Teorema.

Il numero p di autovalori (e quindi l'esponente $1/p$ del fattore η nell'enunciato) derivano dall'equazione (4.2). In questa equazione la variabile di grado più basso ha grado $m - k(f_j + \delta) = (r_{j+1} - \delta)n_{j+1} + \sum_{i=j+2}^{q} r_i n_i$, mentre la variabile di grado più alto ha grado $m - k(f_j - \gamma) = \gamma n_j + \sum_{i=j+1}^{q} r_i n_i$. Poiché sappiamo che il coefficiente non è nullo, possiamo semplificare i fattori comuni e ottenere quindi un'equazione di grado $\gamma n_j + \delta n_{j+1}$, che è appunto p.

In caso accadesse $\delta = n_{j+1} = 1$, allora $(k(f_j)+1, f_j+1)$ sarebbe $\hat{\mathcal{Q}}_\delta^j$, quindi la seconda disuguaglianza diventerebbe ridondante. In caso le disuguaglianze dell'enunciato fossero entrambe strette, allora avremmo così dimostrato che gli unici punti del diagramma di Newton che possono giacere sulla corda sono gli estremi, e quindi possiamo ricavare il risultato su η immediatamente dall'equazione (4.2). □

Vediamo il risultato di questo Teorema in azione su un esempio mirato, che la teoria di Lidskii non riuscirebbe a risolvere.

Sia quindi $A = \left(\begin{array}{ccc|cc} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & & \\ \hline & & & 0 & 1 \\ & & & & 0 \end{array}\right)$ con matrice di perturbazione $\varepsilon E = \left(\begin{array}{ccc|cc} & & & & \\ & & & & \varepsilon \\ & & & & \\ \hline & & & & \\ \varepsilon & & & & \end{array}\right)$.

In questo caso calcoliamo

$$\Phi_1 = 0, \qquad \Phi_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

per cui abbiamo una matrice singolare. Abbiamo

$$A + \varepsilon E = \left(\begin{array}{ccc|cc} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & \varepsilon & \\ \hline & & & 0 & 1 \\ \varepsilon & & & & 0 \end{array}\right),$$

per cui si calcola subito il polinomio caratteristico pari a $p(\lambda, \varepsilon) = \lambda^5 - \varepsilon^2$. In questo caso gli autovalori perturbati sono ovvi, ma per applicare il Teorema 4.5 tracciamo il diagramma di Newton di $p(\lambda, \varepsilon)$ e l'inviluppo di Newton di A. Otteniamo il sistema di Figura 4.1

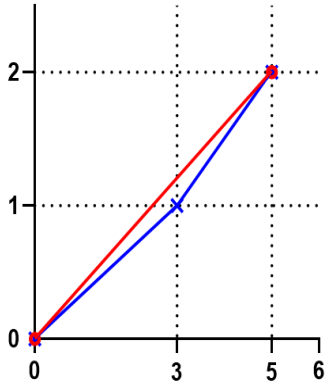

Figura 4.1: *Diagramma di Newton di $p(\lambda, \varepsilon)$, in rosso, e inviluppo di Newton di A, in blu*

Applicando il Teorema 4.5, le cui condizioni sono banalmente verificate, otteniamo $\gamma = 1$, $\delta = 1$, $j = 1$, $\mathcal{Q}_\gamma^j = (0,0)$, $\hat{\mathcal{Q}}_\delta^j = (5,2)$, $\mathcal{P}_j = (3,1)$, $p = 1\cdot 3 + 1\cdot 2 = 5$, $\sigma = \frac{1+1}{p} = \frac{2}{5}$. Inoltre, poiché valgono le disuguaglianze strette dell'enunciato del Teorema, possiamo anche calcolare

$$\eta = -\frac{\hat{\alpha}_5}{\hat{\alpha}_0} = 1.$$

Dal teorema, i 5 autovalori perturbati saranno

$$\tilde{\lambda}^l = 0 + 1^{1/5}\varepsilon^{2/5} + o(\varepsilon^{2/5}), \qquad l = 1, \dots 5,$$

che è ciò che si verificava banalmente dal polinomio caratteristico.

Per quanto riguarda gli autovettori non possiamo purtroppo guadagnare molte informazioni, perché la non singolarità delle matrici Φ_i ci è ancora necessaria per garantire la semplicità degli autovalori perturbati, ma è possibile semplificare le ipotesi del Teorema 4.2 richiedendo che, fissata j che soddisfa le ipotesi dello stesso Teorema, solo le matrici $\Phi_1, \dots, \Phi_j$ siano non singolari. Infatti il coefficiente angolare dei segmenti a destra di $\mathcal{P}_j$, che sotto queste ipotesi fa parte sia dell'inviluppo di Newton sia del diagramma di Newton, sarà strettamente maggiore di $1/n_j$ indipendentemente dal rango delle matrici Φ_i, $i = j+1 \dots, q$, e quindi gli autovalori ricavati da questi segmenti si potranno in ogni caso separare da quelli del segmento S_j, ancora garantendo la semplicità degli autovalori perturbati.

Conclusioni

Ciò che abbiamo esposto nei Capitoli precedenti mostra quanto studio ci sia stato nel contesto della teoria di perturbazione matriciale, e quanto interesse ci sia nel trovare risultati il più precisi possibile del grado di perturbazione delle autocoppie di una matrice. La teoria generale risulta quasi totalmente sviscerata, ma la ricerca è ancora decisamente attiva, con l'attenzione maggiore posta su casi particolari. Ciò potrebbe a prima vista risultare strano, perché è naturale pensare che, quando si conoscono i risultati nel caso generale, imporre delle condizioni aggiuntive al problema in partenza non possa che rendere le cose più facili, utilizzando i risultati già ottenuti che valgono in una sfera più ampia. Come spesso accade in matematica, però, questa risulterebbe essere un'analisi superficiale, non tenendo conto del fatto che aggiungere condizioni permette quasi sempre di lavorare in un modo che sarebbe stato impensabile o inaccessibile altrimenti, e ottenere quindi delle nuove teorie, molto più specifiche e più ricche. Le nuove teorie richiederanno quasi sempre uno studio maggiore, e daranno risultati migliori, che non sarebbero stati prevedibili con i risultati generali.

Nella Sezione 2.1 siamo riusciti a coinvolgere la matrice E nell'analisi della perturbazione e a collegare la perturbazione dello spettro di A con gli autovalori di E. Tutto questo non si sarebbe potuto ricavare in alcun modo utilizzando solo i risultati del Capitolo 4, e anzi, ha comportato la costruzione di una nuova teoria che sfocia ben oltre la perturbazione matriciale (come nel caso del Teorema 1.5 di Courant-Fisher). Anche il punto di vista può cambiare molto specializzando il problema: nella Sezione 2.2 ci siamo potuti mettere in un'ottica completamente geometrica, ed è risultata estremamente efficace per lo studio della perturbazione degli autovettori, dando una visione completamente nuova al problema stesso e facendone risaltare la rilevanza in ulteriori campi della matematica. Nella Sezione 3.1 abbiamo similmente coinvolto la topologia per studio della continuità degli autovalori perturbati, utilizzando questo strumento per estendere il più possibile i risultati del Capitolo precedente. Questo ultimo approccio porta ancora una volta a studiare il problema in un modo totalmente nuovo, e può essere utilizzato per trovare ancora

molti altri risultati oltre a quelli che abbiamo esposto. Ovviamente ci sono anche casi in cui ciò che otteniamo specializzando il problema è effettivamente quello che avremmo ottenuto nel caso generale: si veda per esempio la Sezione 3.2, che scrive autovalori e autovettori come espansioni del parametro ε, lo stesso risultato ottenuto ripetutamente nel Capitolo 4. In questo caso lo scopo della specializzazione del problema è stato proprio quello di semplificare la teoria, ma si è trattato comunque di un passaggio fondamentale, che ci ha permesso di introdurci nell'ottica del problema non strutturato in un modo decisamente più immediato e ha utilizzato Teoremi che comunque non sarebbero stati affrontati nel Capitolo seguente (come il Teorema 3.9 e il Teorema 3.10, di Gerschgorin). Abbiamo inoltre ricavato qualche informazione in più sugli autovettori, esprimendo in modo esplicito i coefficienti dei termini in ε che nel Capitolo 4 vengono lasciati indicati.

Quello che abbiamo voluto esplorare in questo lavoro non è quindi solo la teoria e i risultati in sé, ma come questi possano provenire da punti di vista così variegati e con studi così diversi. Tentativi di ulteriori approcci al problema tramite strumenti di altre discipline matematiche sono ancora in atto e si stanno rivelando efficaci per risolvere altri casi particolari del problema (come per esempio il caso in cui A ed E siano normali rispetto ad un altro prodotto interno dello spazio vettoriale su cui agiscono). Questo ci fa capire quanto la collaborazione e la coesione interna siano essenziali in matematica, per permettere un avanzamento comune e ottenere risultati ancora più potenti, che possano aprire la strada verso nuovi studi e metodi di ricerca.

Bibliografia

[1] G. W. Stewart, J. Sun, "Matrix Perturbation Theory", *Academic Press*, Computer Science and Scientific Computing, San Diego 1990

[2] C. Davis, W. M. Kahan, "The Rotation of Eigenvectors by a Perturbation. III", *SIAM Journal on Numerical Analysis*, Vol. 7, N. 1, 1970, pp. 1-46

[3] M. S. Moslehian, "Ky Fan inequalities", *Taylor & Francis Online*, Linear and Multilinear Algebra, Vol. 60, Is. 11-12, 2012, pp. 1313-1325

[4] J. H. Wilkinson, "The Algebraic Eigenvalue Problem", *Oxford University Press*, New York 1965

[5] R. Bhatia, "Matrix Analysis", *Springer Science + Business Media*, Graduate Texts in Mathematics, Vol. 169, New York 1997

[6] R. Bhatia, "Perturbation bounds for matrix eigenvalues", *Longman Scientific & Technical*, Pitman Research Notes in Mathematics Series, Vol. 162, Harlow 1987

[7] G. Buffington, "Polar Decomposition of a Matrix", Dispense del corso *Advanced Linear Algebra*, University of Puget Sound, Tacoma 2014

[8] J. Moro, J. V. Burke, M. L. Overton, "On the Lidskii-Vishik-Lyusternik Perturbation Theory for Eigenvalues of matrices with arbitrary Jordan structure", *SIAM Journal on Matrix Analysis and Applications*, Vol. 18, Is. 4, 1997, pp. 793-817

[9] A. Greenbaum, R. Li, M. L. Overton, "First-Order Perturbation Theory for Eigenvalues and Eigenvectors", *SIAM Review*, Vol. 62, Is. 2, 2020, pp. 463-482

[10] H. Baumgärtel, "Analytic Perturbation Theory for Matrices and Operators", *Birkhäuser Verlag*, Operator Theory: Advances and Applications, Vol. 15, Basilea 1985

[11] V. Simoncini, "Dispense del corso di Calcolo Numerico, Modulo di Algebra Lineare Numerica", Università di Bologna, 2022

[12] V. Simoncini, "Metodi Matriciali per il Data Science", Dispense del corso *Matematica Computazionale*, Università di Bologna, 2023

[13] G. Cupini, "Appunti di lezione", Dispense del corso *Analisi Matematica 2*, Modulo 1, Università di Bologna, 2022

[14] A. Alexanderian, "On continuous dependence of roots of polynomials on coefficients", North Carolina State University, 2013