

Appunti di Analisi Numerica

Pietro Battiston
Dipartimento di Matematica
Università degli Studi di Pisa

18 febbraio 2007

Capitolo 1

Introduzione

1.1 Introduzione a questi appunti

Questo documento riporta alcuni appunti che ho preso nel corso di Analisi Numerica del prof. Bini (corso di laurea in Matematica, Università di Pisa), nel primo semestre dell'anno 2005-2006. In vari punti però ho riorganizzato gli argomenti, anche consultando i libri di testo, in un modo che mi sembrava più semplice o più logico; questo per dire di non prendere niente per oro colato, e che mi assumo la responsabilità di qualsiasi bestialità presente tra queste pagine.

Ho cominciato a scriverli perché avevo dato varie volte, senza passarlo, lo scritto di Analisi Numerica, e non mi decidevo mai a studiare per bene 'sta materia, così ho pensato che copiando e riorganizzando gli appunti sarei stato finalmente costretto a capirla.

Alla sessione dopo però sono riuscito a dare l'esame, e bene, per cui ho semplicemente mollato questi appunti dov'ero arrivato. Contengono quindi solo i *primissimi* argomenti visti a lezione, ma ho pensato che magari a qualcuno potevano fare comodo comunque.

NON penso che svilupperò mai oltre questi appunti (se ti piacerebbe farlo, sono dispostissimo a passarti i sorgenti); ciononostante, se trovi errori, scrivimi pure, sarò felice di correggerli: battiston@mail.dm.unipi.it.

1.2 Introduzione all'Analisi Numerica

L'Analisi Numerica è una branca della matematica che affronta problemi di molti diversi tipi (algebra lineare, analisi...), ma da un punto di vista computazionale.

Dove spesso in matematica viene ritenuto sufficiente, o comunque fondamentale, un teorema che indichi l'esistenza di una data soluzione ad un problema, l'Analisi Numerica cerca di fornire metodi per trovare effettivamente tale soluzione, con un occhio di riguardo all'ottimizzazione dei seguenti fattori:

- tempo di esecuzione (numero e complessità delle operazioni necessarie)
- spazio necessario all'esecuzione (risultati di calcoli intermedi che è necessario memorizzare)

- possibilità di discretizzazione (si cerca di fare in modo tale che l'approssimazione dei dati in ingresso o dei risultati intermedi non implichi una perdita di precisione rilevante nell'output)

L'Analisi Numerica è una branca della matematica, e *non* dell'informatica. Ciononostante si può intuire da questi tre punti, e sarà ancora più evidente andando avanti, che c'è una relazione molto stretta con il mondo dei computer, le cui CPU sono macchine incapaci di calcolo simbolico, con cui cioè ogni problema va discretizzato.

Di seguito sono riportati due esempi semplici e noti di problemi per cui si è riusciti a trovare algoritmi sempre migliori, abbattendone il costo computazionale (la notazione utilizzata, di cui per ora si può intuire il senso, sarà spiegata nel paragrafo successivo):

1. *Moltiplicazione di due numeri interi a e b , di n cifre ognuno.*

L'algoritmo standard, che ogni bambino impara alle elementari, consiste nel moltiplicare il numero a per ogni cifra del numero b , moltiplicare i vari risultati per opportune potenze di 10 e quindi sommarli.

Ogni volta che moltiplichiamo a per una cifra di b , stiamo eseguendo però n operazioni, dato che stiamo moltiplicando ogni cifra di a per quella cifra di b . Il costo totale dell'algoritmo è quindi $O(n^2)$.

Sebbene potrebbe sembrare strano che un algoritmo di questa semplicità possa essere computazionalmente più complesso di altri, è così: il costo dell'algoritmo di *Karatsuba-Offman* ha un costo $O(n^{\log_2 3})$, ed un algoritmo basato sulla FFT (*Fast Fourier Transform*) abbassa il costo a $O(n \log n^2)$ o addirittura a $O(n \log n \log(\log n))$ ¹

2. *Triangolarizzazione di una matrice quadrata di dimensione $n \times n$.*

L'algoritmo di gran lunga più utilizzato per risolvere questo problema a mano è il cosiddetto *metodo di Gauß*. Tale algoritmo però oltre a problemi di stabilità (concetto che si affronterà più avanti), presenta un costo di $O(n^3)$. Invece è stato dimostrato che lo stesso problema si può risolvere con $O(n^{\log_2 7})$ o addirittura (per n molto grande) $O(n^{2.38})$.

1.2.1 Costo degli algoritmi e comportamento asintotico

La notazione $O(f(n))$ (*O grande di f*), dove f è una funzione, e tipicamente un elevamento a potenza, verrà utilizzata molto spesso, e si deve interpretare come *dell'ordine di f*. Ad esempio se il costo computazionale di un dato algoritmo è $O(n^2)$, ci si può aspettare che per un input di dimensione n l'algoritmo preveda

¹In realtà mentre per il primo algoritmo stiamo considerando due numeri ognuno di n cifre *decimali*, negli altri due n indica il numero di cifre di a e b se scritti in notazione *binaria*. Si verifica però facilmente che, nonostante questa difformità di notazione, l'algoritmo di Karatsuba-Offman e quello basato sulla FFT sono effettivamente più efficienti di quello standard.

circa² n^2 operazioni elementari³.

In realtà la notazione si riferisce al *comportamento asintotico* della funzione, ovvero a *quanto velocemente cresce per n che tende a infinito*. Questa si rivela generalmente una buona scelta, perché si è maggiormente interessati all'efficienza di un algoritmo quando l'input del problema da risolvere è molto grande. Infatti può capitare che, nel risolvere problemi di piccole dimensioni, un algoritmo sia leggermente più veloce di un altro più efficiente, ma solitamente le differenze sono trascurabili.

È importante notare che questo tipo di notazione ci permette di non considerare tutti quei costi che sono asintoticamente trascurabili rispetto al costo complessivo dell'algoritmo. Ad esempio, nel caso dell'algoritmo standard per la moltiplicazione di due numeri, nel computo delle operazioni elementari abbiamo trascurato il costo delle moltiplicazioni per le potenze del 10, sfruttando il fatto che $O(n^2) = O(n^2) + n$, e poi il costo delle somme degli n risultati parziali (somme che nel complesso hanno un costo pari a n^2), sfruttando il fatto che $O(n^2) + O(n^2) = O(n^2)$.

Il costo di un algoritmo viene spesso detto:

- *lineare*, se è $O(n)$
- *logaritmico*, se è $O(\log_a n)$ (qualunque sia la base a)
- *quadratico*, se è $O(n^2)$
- *cubico* se è $O(n^3)$...

Notiamo che, fatte salve situazioni molto particolari in cui i dati di input hanno certe proprietà note, il costo di un algoritmo non può mai essere meno che lineare.

1.3 Programma

Come si è già detto, gli strumenti dell'Analisi Numerica possono essere sfruttati in settori molto diversi della matematica. In questo corso verranno trattati i seguenti argomenti:

1. analisi dell'errore
2. algebra lineare numerica (ad esempio il metodo di Gauß, la forma normale di Shur...)
3. risoluzione di equazioni non lineari
4. interpolazione, approssimazione (in particolare un'occhiata alla Fast Fourier Transform)

²In realtà esiste una definizione di $O(f(n))$ assolutamente rigorosa:

$$g(x) = O(f(x)) \iff \lim_{x \rightarrow +\infty} \frac{g(x)}{f(x)} \in \mathbb{R}$$

³Per *operazione elementare* si intende solitamente un'addizione, una moltiplicazione, una sottrazione o una divisione, anche se nella realtà queste stesse operazioni possono avere costi diversi per un elaboratore

Capitolo 2

Analisi dell'errore

2.1 I numeri a virgola mobile

Ogni numero $x \in \mathbb{R}, x \neq 0$ può essere rappresentato nella seguente forma:

$$x = \text{segno}(x) \cdot \beta^p \sum_{i=1}^{+\infty} d_i \beta^{-i}$$

con $\beta \in \mathbb{R}, \beta \geq 2, p \in \mathbb{Z}, 0 \leq d_i \leq \beta - 1, \forall i, d_1 \neq 0$ ed inoltre $\forall k \in \mathbb{N} \exists i \geq k : d_i \neq \beta - 1$.

La funzione $\text{segno}(x)$, come dice il nome, vale 1 se $x > 0$ e -1 se $x < 0$.

Tale rappresentazione viene detta *rappresentazione in base dei numeri reali* (ma dai!) ed è facile verificare che, fissati x e β , è unica. Osserviamo solo che l'ultima condizione (detta *condizione di normalizzazione*) è necessaria per evitare che, ad esempio, in base decimale, il numero 2 possa essere raffigurato sia come 2 che come $1, \overline{9}$.¹

Il numero p viene detto *esponente*, $\sum_{i=1}^{+\infty} d_i \beta^{-i}$ (o, talvolta, la successione dei b_i) *mantissa*.

A titolo di esempio, osserviamo che, in base 10, $\frac{1}{3} = 1 \cdot 10^0 \sum_{i=1}^{+\infty} 3 \beta^{-i}$.

Osservazione: basi diverse hanno numeri periodici diversi. Ad esempio $\frac{1}{10}$ è periodico in base 2.

2.2 Cosa succede nei computer

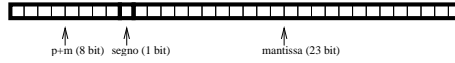
Per memorizzare i numeri in questa forma nei computer abbiamo alcune restrizioni: la mantissa non è affatto infinita, p non è qualunque ma $-m \leq p \leq M$. Inoltre, in generale, $\beta = 2$.

$\{\pm \beta^p \sum_{i=1}^t d_i \beta^{-i} \mid -m \leq p \leq M\} \cup \{0\} = \mathcal{F}(\beta, t, m, M)$ viene così detto *l'insieme dei numeri di macchina su base β con t cifre*.

Questa rappresentazione dei numeri viene detta *in virgola mobile* (in inglese *floating point*).

¹Da un punto di vista formale, sarebbe stata equivalente questa condizione: $\exists k : \forall i > kb_i = \beta - 1$. A parte il fatto evidente che tale convenzione ci avrebbe reso la vita più scomoda, conviene scegliere l'altra anche perché come vedremo più avanti, in un computer le cifre non sono affatto infinite.

Tipicamente in un computer lo spazio occupato da un numero in virgola mobile è di 4 *bytes*, ovvero 32 *bit*. In questo caso la suddivisione dello spazio è la seguente:



I primi 8 bit rappresentano, in notazione binaria, il numero (positivo) $p+m$, con $m = 2^7 - 1 = 127$. In questo modo, abbiamo $-m = -125 \leq p \leq 128 = M^2$. Un numero in questa forma non potrà quindi essere più grande (in valore assoluto) di $2^{128} \approx 3,40 \cdot 10^{38}$. Potrà invece essere più piccolo (sempre in valore assoluto) 2^{-125} , semplicemente perché le cifre della mantissa possono essere tutte 0 tranne l'ultima (in modo tale da rappresentare il numero 2^{-147}).

A differenza di quanto si potrebbe pensare, il nono bit, che indica il segno, è *on* (1) per un numero negativo e *off* (0) per un positivo.

Lo zero, unico numero per cui non abbiamo ancora previsto una rappresentazione, viene associato alla configurazione di tutti i bit nulli.

I numeri floating point di 4 byte vengono detti *a precisione semplice*. Le possibili precisioni, definite da uno standard IEEE, sono:

Precisione	β	t	m	M	byte utilizzati	cifre (in $\beta = 10$)
semplice	2	23	125	128	4	≈ 8
doppia	2	52	1021	1024	8	≈ 16
estesa	2	63	16381	16384	20	≈ 20

Tabella 2.1: Gli standard dell'IEEE per la rappresentazione di numeri floating point

² $m = 125$, e non 127, perché alcune configurazioni di bit sono riservate per i *NaN* (*Not A Number*), ovvero i numeri troppo grandi o troppo piccoli per essere raffigurati.

Capitolo 3

Elementi di Algebra Lineare Numerica

3.1 Introduzione

In questo capitolo avremo molto a che fare con matrici e vettori. Solitamente lo spazio vettoriale in cui ci muoveremo sarà \mathbb{C}^n , talvolta \mathbb{R}^n . Per lo spazio delle matrici $n \times n$ a coefficienti in \mathbb{C} (in \mathbb{R}) utilizzeremo la notazione $\mathbb{C}^{n \times n}$ ($\mathbb{R}^{n \times n}$), introdotta da *Golub*.

Alternativamente scriveremo $A = (a_{ij})$, $a_{ij} \in \mathbb{C}$, dove a_{ij} è ovviamente il j -esimo coefficiente della i -esima riga.

3.2 I tre teoremi di Gerschgorin

Gerschgorin, un matematico russo di origine ebrea, ha formulato tre teoremi fondamentali per la localizzazione degli autovalori di matrici complesse con un basso costo computazionale:

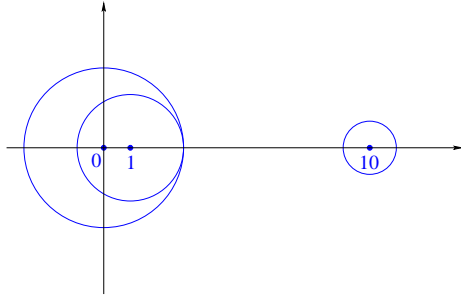
3.2.1 Il primo teorema di Gerschgorin

Definizione 1. Sia $A = (a_{ij})$, $a_{ij} \in \mathbb{C}$. Si dice i -esimo cerchio di Gerschgorin di A :

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

Esempio 1. Di seguito vediamo un esempio di matrice e i cerchi di Gerschgorin associati disegnati nel piano complesso:

$$\begin{pmatrix} 10 & 1 & 0 \\ -3 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$



Teorema 1 (Primo teorema di Gerschgorin). *Siano $\lambda_1 \dots \lambda_n$ gli autovalori di A . Allora $\forall j \in \{1 \dots n\}$ si ha:*

$$\lambda_j \in \bigcup_{i=1}^n K_i$$

Esempio 2. *Sfruttando il primo teorema di Gerschgorin, è facile verificare che la matrice quadrata:*

$$\begin{pmatrix} 3 & 1 & 0 & \dots & 0 \\ 1 & 3 & 1 & \dots & 0 \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 3 & 1 \\ 0 & 0 & \dots & 1 & 3 \end{pmatrix}$$

è singolare. Infatti $\forall \lambda_i, |\lambda_i - 3| \leq 2 \Rightarrow |\lambda_i| \geq 1$.

Dimostrazione. Sia $v \in \mathbb{C}^n : Av = \lambda v$ un generico autovettore, e sia v_i la sua i -esima coordinata.

Sia $k : |v_k| = \max_i |v_i|$, ovvero v_k è la componente di modulo massimo. Allora:

$$\begin{aligned} \sum_{j=1}^n a_{kj} v_j &= \lambda v_k \Rightarrow (a_{kk} - \lambda) v_k = - \sum_{\substack{j=1 \\ j \neq k}}^n a_{kj} v_j \Rightarrow \\ \Rightarrow a_{kk} - \lambda &= - \sum_{\substack{j=1 \\ j \neq k}}^n a_{kj} \frac{v_j}{v_k} \Rightarrow |a_{kk} - \lambda| \leq \sum_{\substack{j=1 \\ j \neq k}}^n |a_{kj}| \end{aligned}$$

□

3.2.2 Matrici companion (di Frobenius)

Definizione 2. *Vengono dette matrici companion o matrici di Frobenius le matrici quadrate della seguente forma:*

$$F = \begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \dots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -a_{n-1} \end{pmatrix}$$

Tali matrici sono interessanti perché il loro determinante ha una caratteristica molto particolare:

$$\det(F - \lambda I) = (-1)^n \sum_{i=0}^n \lambda^i a_i$$

Ovvero se troviamo gli autovalori della matrice F (che sono proprio i valori che annullano il determinante), troviamo automaticamente gli zeri del polinomio associato, cioè quello che ha per coefficienti $a_1 \cdots a_n$. Questa constatazione, associata al primo teorema di Gerschgorin, ci è utile quando abbiamo bisogno di localizzare gli zeri di un polinomio.

3.2.3 Il secondo teorema di Gerschgorin

Teorema 2 (Secondo teorema di Gerschgorin). *Siano $K_1 \cdots K_n$ i cerchi di Gerschgorin di una matrice $A \in \mathbb{C}^{n \times n}$. Allora, se esistono due insiemi M_1, M_2 tali che $M_1 \cap M_2 = \emptyset$ e:*

$$\bigcup_{i=1}^n K_i = M_1 \cup M_2$$

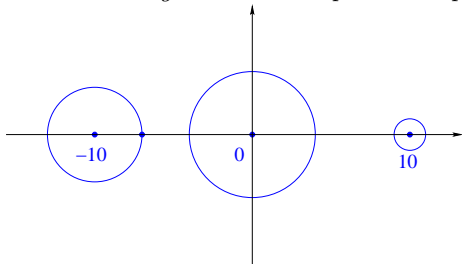
allora in M_1 si trovano tanti autovalori quanti sono i cerchi che ne fanno parte. In particolare,

$\exists i : \forall j K_i \cap K_j = \emptyset \Rightarrow K_i$ contiene esattamente un autovalore.

Esempio 3. *Sia A la seguente matrice:*

$$\begin{pmatrix} 10 & 1 & 0 \\ -1 & 0 & 3 \\ -2 & 1 & -10 \end{pmatrix}$$

È semplice dimostrare che gli autovalori di A sono tutti reali. Disegniamo i cerchi di Gerschgorin di A sul piano complesso:



Il polinomio caratteristico della matrice ha coefficienti reali quindi per ogni $z \in \mathbb{C}$ che ne è soluzione deve essere soluzione anche $\bar{z} \in \mathbb{C}$. Se per assurdo avessimo un autovalore non reale, nello stesso cerchio di Gerschgorin troveremmo il coniugato, e quindi due autovalori nello stesso cerchio (notiamo che, dato i , $\forall z \in K_i, \bar{z} \in K_i$). Ma per il secondo teorema di Gerschgorin sappiamo che questo non può succedere, perché i 3 cerchi di Gerschgorin sono disgiunti.

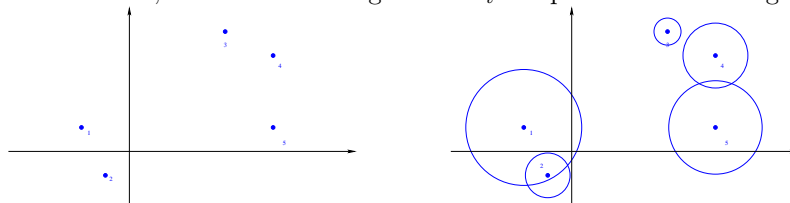
Dimostrazione. Sia $A = (a_{ij}), a_{ij} \in \mathbb{C}$. Definiamo le seguenti matrici $n \times n$:

$$D = \text{diag}(a_{11} \cdots a_{nn}) = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

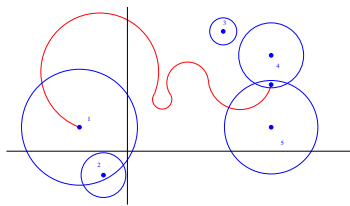
$$A_t = D + t(A - D) = \begin{pmatrix} a_{11} & ta_{12} & \cdots & ta_{1n} \\ ta_{21} & a_{22} & \cdots & ta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ta_{n1} & ta_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Osserviamo che $A_0 = D, A_1 = A, A_2 = 2A - D \dots$
 Notiamo inoltre che $\forall i, j$ si ha che ta_{ij} è una funzione continua e che, siccome gli zeri di un polinomio sono funzioni continue dei coefficienti, gli autovalori di A_t sono ancora funzioni continue di t .

Per $t = 0$, i cerchi di Gerschgorin di A_t comprendono soltanto gli autovalori:



Per la continuità degli autovalori, se un autovalore è in un dato insieme $M_1 = \bigcup_{i \in I} K_i, I \subset \{1 \cdots n\}$ per $t = 0$ e in $M_2 = \bigcup_{i \notin I} K_i$ per $t = 1$, con $M_1 \cap M_2 = \emptyset$, allora necessariamente ci sarà un istante \bar{t} per cui l'autovalore si trova all'esterno sia di M_1 che di M_2 (se siamo in vena di pignoleggiare, notiamo che questo ci è garantito perché sono entrambi chiusi in \mathbb{R}).



Ma al variare di $t \in [0, 1]$, i cerchi di Gerschgorin aumentano il proprio raggio, fino a diventare effettivamente i cerchi di A per $t = 1$. Quindi se due gruppi di cerchi di A_t sono disgiunti per $t = 1$, lo saranno anche per ogni $t \in [0, 1]$. Perciò se un autovalore ad un dato istante \bar{t} è esterno ad $M_1 \cup M_2$, è anche esterno all'unione dei cerchi di Gerschgorin di A_t . Sappiamo, dal primo teorema di Gerschgorin, che questo è impossibile. L'osservazione (ovvia) che all'istante $t = 0$ ogni k -upla di cerchi di Gerschgorin riguardanti lo stesso autovalore (di molteplicità k) contiene esattamente k autovalori (contati con molteplicità) conclude la dimostrazione.

□

3.2.4 Matrici irriducibili e il terzo teorema di Gerschgorin

Sia data la permutazione $\sigma : \{1 \cdots n\} \rightarrow \{1 \cdots n\}$.

Definiamo il simbolo δ_{ij} :

$$\delta_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{altrimenti} \end{cases}$$

Osserviamo che $I = (\delta_{ij})$, dove I è la matrice *identità*.

Definizione 3. Chiamiamo matrice di permutazione associata a σ la matrice $\Pi = (\delta_{\sigma(i),j})$.

Osservazioni:

1. ΠA , con $A \in \mathbb{C}^{n \times n}$, sarà una matrice avente le righe di A scambiate di posto.
2. Πx , con $x \in \mathbb{C}^n$, sarà una matrice avente le componenti di x scambiate di posto.
3. $\Pi \Pi^t = \Pi^t \Pi = I$

Definizione 4 (Riducibilità). Una matrice $A \in \mathbb{C}^{n \times n}$ si dice *riducibile* se esiste una matrice di permutazione Π tale che:

$$\Pi A \Pi^t = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$$

dove A_{11} e A_{22} sono due matrici quadrate di dimensioni rispettivamente $k \times k$ e $(n - k) \times (n - k)$.

$\Pi A \Pi^t \Pi$ viene detta, come forse alcuni dei lettori più fantasiosi avranno già immaginato, matrice *ridotta*.

Il concetto di riducibilità si può rivelare utile, ad esempio, nel risolvere sistemi quadrati di equazioni lineari. Infatti:

$$Ax = b \Rightarrow \Pi A \Pi^t \Pi x = \Pi b \Rightarrow \begin{cases} A_{11} \hat{x}_1 = \hat{b}_1 - A_{12} \hat{x}_2 \\ A_{22} \hat{x}_2 = \hat{b}_2 \end{cases}$$

Dove $\hat{x} = \Pi x$.

Mentre il problema iniziale era composto da n equazioni in n incognite, quello finale, equivalente, è composto da sole $n - k$ equazioni in $n - k$ incognite. Se inizialmente il costo computazionale era di n^3 operazioni, l'abbiamo quindi abbassato a $k^3 + (n - k)^3$.

Matrici irriducibili e grafi fortemente connessi

È interessante spendere due parole sulle relazioni tra matrici e *grafi diretti*¹ *associati*. In questo paragrafo, con (i) si intende l' i -esimo nodo del grafo, con (i, j) l'arco (orientato) che collega (i) a (j) .

¹Con grafo diretto intendiamo un grafo i cui archi sono orientati, per cui l'arco (i, j) NON è la stessa cosa dell'arco (j, i) .

Definizione 5. Data una matrice $A = (a_{ij})$, con $a_{ij} \in \mathbb{C}$, chiamiamo grafo associato ad A il grafo di n nodi in cui l'arco (i, j) esiste se e solo se $a_{ij} \neq 0$.

Esempio 4. Una semplice matrice irriducibile e il suo grafo associato:

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \quad \begin{array}{ccccccc} \textcircled{1} & \rightleftarrows & \textcircled{2} & \rightleftarrows & \textcircled{3} & \rightleftarrows & \cdots & \rightleftarrows & \textcircled{n} \end{array}$$

Definizione 6. Un grafo con n nodi viene detto fortemente connesso se $\forall i, j \in \{1 \cdots n\}$ esiste un cammino (una successione di archi) che va dal nodo (i) al nodo (j) .

Teorema 3. Una matrice è irriducibile se e solo se il suo grafo diretto associato è fortemente connesso.

Esempio 5. Si verifica facilmente che A_1 e A_2 , (matrici con importanti proprietà e dette rispettivamente derivata seconda nel discreto e operatore di Laplace) sono irriducibili:

$$A_1 = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 2 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} B & -I & 0 & \cdots & 0 & 0 \\ -I & B & -I & \cdots & 0 & 0 \\ 0 & -I & B & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & B & -I \\ 0 & 0 & 0 & \cdots & -I & B \end{pmatrix}$$

$$\text{dove } B = \begin{pmatrix} 4 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 4 & -1 & \cdots & 0 & 0 \\ 0 & -1 & 4 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 4 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 4 \end{pmatrix}, \quad I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Dimostrazione. Abbiamo la nostra matrice $A = (a_{ij})$ e la matrice di permutazione Π , associata alla permutazione σ , tale che $B = \Pi A \Pi^t$ è una matrice ridotta.

Osserviamo che $a_{i,j} \neq 0 \Leftrightarrow b_{\sigma(i), \sigma(j)} \neq 0$, e che quindi nel grafo associato ad A esiste l'arco (i, j) se e solo se in quello associato a B esiste l'arco $(\sigma i, \sigma j)$ (ovvero

i due grafi sono identici a meno di rinumerazione dei nodi. Sappiamo che:

$$B = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}, A_{11} \in \mathbb{C}^{k \times k}.$$

Sia P l'insieme dei nodi $\{(1) \dots (k)\}$ del grafo associato a B , e sia Q l'insieme dei restanti $n - k$ nodi. Si verifica facilmente, utilizzando la definizione di grafo associato, che non esiste nessun arco che va da un nodo in Q ad un nodo in P . Il grafo non è quindi fortemente connesso.

Viceversa, se il grafo associato ad una matrice B non è connesso, esistono due nodi $(i), (j)$ tali che non esiste nessun cammino che va da (i) a (j) . Sia P l'insieme dei nodi raggiungibili da i (i compreso ovviamente) e Q l'insieme degli altri nodi (che comprende almeno j); possiamo rinumerare i nodi in modo tale che $(i) = (1)$, $P = \{(1), (2), \dots, (k)\}$, $(j) = (k+1)$, $Q = \{(k+1), (k+2), \dots, (n)\}$. Se per assurdo ci fosse un arco che connette un nodo $(p) \in P$ ad un nodo di $(q) \in Q$, aggiungendo al cammino che unisce (i) a (p) l'arco (p, q) otterremo un cammino da (i) a (q) , e questo è assurdo. Quindi $\forall i \in \{1 \dots k\}, j \in \{k+1 \dots n\}$ abbiamo $b_{ij} = 0$, ovvero la matrice B è ridotta. \square

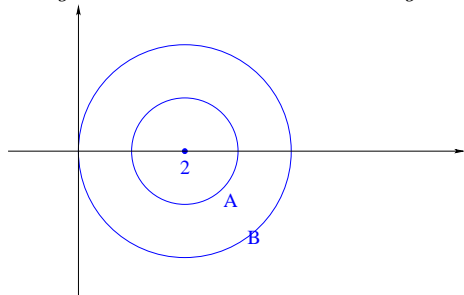
Il terzo teorema di Gerschgorin

Teorema 4. Sia $A \in \mathbb{C}^{n \times n}$ irriducibile, e sia λ un suo autovalore. Allora $\lambda \in \bigcup_i \partial K_i \Rightarrow \lambda \in \bigcap_i \partial K_i$ ²

Esempio 6. Supponiamo di voler sapere se questa matrice è singolare:

$$\begin{pmatrix} 2 & -1 & 0 & \dots & 0 & 0 \\ -1 & 2 & -1 & \dots & 0 & 0 \\ 0 & -1 & 2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 2 & -1 \\ 0 & 0 & 0 & \dots & -1 & 2 \end{pmatrix}$$

Disegniamo i suoi cerchi di Gerschgorin:



K_1 e K_n coincidono con A , mentre $K_2 \dots K_{n-1}$ coincidono con B . Quindi 0 non può essere un autovalore, perché è sulla frontiera di K_1 e K_n ma non sulla frontiera di tutti gli altri cerchi di Gerschgorin.

²Dato un insieme U , si denota con ∂U la frontiera di U . Nel caso dei cerchi di Gerschgorin, $\partial K_i \subset K_i$.

Un'applicazione interessante del teorema riguarda le matrici *dominanti diagonali*.

Definizione 7 (Dominanza diagonale). *Una matrice A si dice dominante diagonale se:*

$$\forall i, |a_{ii}| \geq \sum_{j \neq i} a_{ij}$$

$$\exists i : |a_{ii}| > \sum_{j \neq i} a_{ij}$$

A si dice dominante in senso stretto se:

$$\forall i, |a_{ii}| > \sum_{j \neq i} a_{ij}$$

Il primo teorema di Gerschgorin ci dice che se una matrice è dominante diagonale in senso stretto, allora è non singolare. Per il terzo teorema di Gerschgorin, se una matrice irriducibile è dominante diagonale, allora è non singolare.

Dimostrazione. Per dimostrare il terzo teorema di Gerschgorin, ci ricollegiamo alla dimostrazione del primo. Avevamo una matrice A con un autovalore λ , e sapevamo:

$$\exists \hat{k} : a_{\hat{k}\hat{k}} - \lambda = - \sum_{\substack{j=1 \\ j \neq \hat{k}}}^n a_{\hat{k}j} \frac{v_j}{v_{\hat{k}}} \quad (3.1)$$

Osserviamo che:

$$\lambda \in \partial \bigcup_{i=0}^n K_i \Rightarrow \forall k : \lambda \in K_k, \lambda \in \partial K_k \quad (3.2)$$

e che:

$$\lambda \in \partial K_k \Leftrightarrow |a_{kk} - \lambda| = \sum_{\substack{i=1 \\ i \neq k}}^n |a_{ki}| \quad (3.3)$$

Dalle equazioni 3.1, 3.2 e 3.3 otteniamo:

$$|a_{ii} - \lambda| = \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \Rightarrow \forall i, \left| \frac{v_i}{v_k} \right| = 1 \quad (3.4)$$

Ma allora *tutte* le componenti di x hanno modulo massimo, quindi nell'equazione 3.1 possiamo sostituire a \hat{k} ogni $i \in \{1 \cdots n\}$. Inoltre, sempre perché $\forall j, \left| \frac{v_j}{v_k} \right| = 1$, dall'equazione 3.1 possiamo sempre ($\forall j$) derivare la 3.4, ovvero l'autovalore è sulla frontiera di tutti i cerchi di Gerschgorin. \square

3.3 Forma normale di Schur

È noto che la matrice A^t viene detta *trasposta* di A ; è probabilmente quasi altrettanto noto che \bar{A} viene detta *coniugata* di A (si ottiene sostituendo \bar{a}_{ij} ad

ogni a_{ij}). Si utilizza la seguente notazione³ per indicare la matrice *coniugata trasposta* (o *trasposta coniugata*) di A :

$$A^h = \overline{(A^t)} = (\overline{A})^t \quad (3.5)$$

Definizione 8 (Matrici unitarie). *Una matrice $U \in \mathbb{C}^{n \times n}$ si dice unitaria se:*

$$U^h U = U U^h = I \quad (3.6)$$

(notiamo che se $U \in \mathbb{R}^{n \times n}$ viene detta ortonormale).

Definizione 9 (Forma normale di Schur). *Sia data $A \in \mathbb{C}^{n \times n}$. Una matrice $T \in \mathbb{C}^{n \times n}$ triangolare superiore viene detta forma normale di Schur di A se esiste U unitaria tale che:*

$$A = U T U^h \quad (3.7)$$

Teorema 5. *Ogni matrice $A \in \mathbb{C}^{n \times n}$ ha una forma normale di Schur.*

Dimostrazione. Dimostriamolo per induzione sulla dimensione del lato di A :

1. passo 1: se A è una matrice 1×1 è già in forma triangolare, quindi $U = U^h = 1$.
2. passo n : sia $v_1 \in \mathbb{C}^n : A v_1 = \lambda v_1, v_1^h v_1 = 1$ (ovvero prendiamo un autovettore e lo normalizziamo). Mettiamo v come prima colonna di U , e completiamolo ad una base di vettori ortonormali $v_2 \cdots v_n$, che saranno le altre colonne di U .

□

³La lettera h viene utilizzata in onore del matematico francese dell'800 *Charles Hermite*