Università degli Studi di Lecce
Facoltà di Ingegneria

Giuseppe De Cecco
Raffaele Vitolo

NOTE

DI

CALCOLO MATRICIALE

Versione del 20 gennaio 2007

ANNO ACCADEMICO 2003-2004

Indirizzo degli autori.
Giuseppe De Cecco, Raffaele Vitolo,
Università di Lecce, Dipartimento di Matematica,
via per Arnesano, 73100 Lecce
giuseppe.dececco@unile.it
raffaele.vitolo@unile.it

INDICE

INTRODUZIONE

Tutti sono d'accordo che la qualità della vita e il benessere della società nei paesi industrializzati con condizioni umane per il lavoro, l'aspettativa di una lunga vita, cibo ed energia sufficiente, protezione contro l'inclemenza della natura, facili trasporti, divertimenti di ogni genere si basano su ritrovati di tecnologia. Ciò che viene dimenticato sta nel fatto che le basi di questi ritrovati furono messe qualche tempo fa dagli scienziati i quali furono spinti dalla curiosità e non dalle promesse del mercato.

S. Thing, Nobel per la Fisica 1976)

Queste note, in forma non ancora completa e definitiva, costituiscono il contenuto del corso "Calcolo Matriciale" per la laurea specialistica in Ingegneria, Classe dell'Informazione.

Le matrici, come tabelle per rappresentare in forma abbreviata le sostituzioni lineari (oggi chiamate trasformazioni lineari) fecero la loro apparizione in matematica in lavori di Gauss (1777-1855); il termine "matrice" (dal latino *mater matris* = ventre, il punto di origine, la parte più importante) fu introdotto nel 1850 da J. Sylvester (1814-1897). Questi e il suo amico A. Cayley (1821-1895) diedero notevoli contributi alla teoria delle matrici e dei determinanti. Il concetto di determinante invece è più antico, risale a G.W. Leibnitz (1678), mentre il vocabolo è dovuto a Gauss (1801).

La scelta degli argomenti per questo corso è stata dettata dalle più usuali applicazioni delle matrici, ma le applicazioni possibili della matematica non sono prevedibili a priori. Ad esempio J. Kepler ha usato dopo 1800 anni la teoria delle coniche elaborata da Apollonio (III sec a. C.); il fisico A. Carnack e l'ingegnere G.N. Nounsfield nel 1979 ricevettero il premio Nobel per la medicina per gli studi sulla TAC, che si basa sulla teoria della trasformata di A. Radon, introdotta da questi nel 1917 per questioni legate alla Teoria della Misura; la teoria dei numeri (una delle parti più ostiche della matematica) fu usata da A. Turing, professore di matematica a Cambridge, per decifrare nel 1941 il codice ENIGMA usato dai Nazisti durante la II guerra mondiale; l'algebra astratta, in particolare lo studio dei campi finiti, interviene in modo essenziale nella costruzione di tutti i sistemi di sicurezza legati a trasmissione di dati, per esempio i codici correttori di errori.

Non vi è nulla di più pratico di una buona teoria osservava Boltzmann (1844–1906), ma già secoli prima Leonardo (1452–1519), l'ingegnere per antonomasia del Rinascimento,

così si esprimeva: *Studia prima la scienzia, e poi segui la pratica nata da essa scienzia ... Quelli che s'innamoran di pratica senza scienzia son come 'l nocchier ch'entra in navilio senza timone o bussola, che mai ha la certezza dove si vada.*

Come si vede nella storia della matematica, anche l'interesse pratico può essere lo stimolo per creare gli oggetti astratti, matematici, ai quali si rivolgerà l'attenzione, spesso la contemplazione artistica. Ma la maggior parte dei risultati sono raggiunti, come diceva C.G. Jacobi, *per l'onore dello spirito umano.*

Il "bisogno" derivante dal desiderio di comprendere la natura, il funzionamento delle cose, ha agito come motore anche dell'inventiva matematica. E' necessaria, perciò, una maggiore collaborazione tra ingegneri e matematici, molti grandi ingegneri sono stati anche grandi matematici e viceversa: Archimede, Erone, Lagrange, Fourier, Monge solo per citarne alcuni.

Limitandoci ai lavori di Informatica e Teoria dei sistemi, più vicini agli interessi degli specializzandi a cui sono rivolte queste dispense, possiamo dire che non c'è parte della matematica che a priori possa venire esclusa e dichiarata che "non serve".

Ci si imbatte così in uno dei grandi misteri della matematica: come può una scienza che sembra slegata dalla realtà avere strettissimi legami, quasi inaspettati, con la realtà fisica? In una conferenza tenuta il 27 gennaio 1921, A. Einstein così si esprimeva: *A questo punto si presenta un enigma che in tutte le età ha agitato le menti dei ricercatori: come è possibile che le matematiche, le quali dopo tutto sono un prodotto del pensiero umano, dipendente dall'esperienza, siano così ammirevolmente adatte agli oggetti della realtà?*

La matematica non è una collezione di formule pronte per l'uso, ma è un'attività creativa, intreccio tra tradizione e spirito innovativo. La parola *ingegnere* proviene da "ingegno", potenza creativa dello spirito umano, che costituisce la massima espressione del talento e dell'intelligenza.

Ci auguriamo che i lettori ne tengano conto nell'esercizio della loro professione.

Giuseppe De Cecco
Raffaele Vitolo

Ringraziamenti. Ringraziamo l'Università di Lecce, la Facoltà di Ingegneria ed il Dipartimento di Matematica per aver concesso l'uso delle proprie strutture ed apparecchiature. Ringraziamo i professori Distante e Ricci ed il dott. Bandiera per i prezioni consigli, ed anche tutti gli studenti che hanno contribuito a questo testo correggendone molti errori e sviste.

Ringraziamo Luca Greco ed Andrea Russo per aver portato a nostra conoscenza il libro [37].

Queste note sono state scritte in LaTeX2e con l'estensione `amsmath` della American Mathematical Society e le estensioni `graphicx, stmaryrd, wrapfig, easybmat`. Le figure sono state realizzate con l'estensione `pstricks`.

CAPITOLO 1

LE MATRICI

NOTA: per le principali definizioni e le operazioni tra matrici, rimandiamo il lettore al testo [16].

1.1 Premesse

Le matrici sono molto utili in matematica, permettono di semplificare espressioni considerando una tabella di numeri come un singolo ente. Alcuni esempi dimostrano il vasto campo di applicazione delle matrici.

- Si è visto [16] che ogni sistema lineare si può scrivere con la notazione matriciale come $AX = B$, dove A, X e B sono opportune matrici.

- Le matrici sono usate nei problemi di decisione, nei quali è necessario procedere ad una scelta tra diversi modi di agire. Ad ogni relazione $R \subset S \times S$ si può associare una matrice (matrice *booleana*), ponendo 1 se $(a, b) \in R$ e 0 altrimenti. Se, ad esempio, $S = \{a, b, c, d\}$ ed R è la relazione schematizzata dalla figura,

$$R \stackrel{\text{def}}{=} \quad \begin{array}{ccc} a & \longrightarrow & a \\ b & \longrightarrow & b \\ c & \longrightarrow & c \\ d & \longrightarrow & d \end{array} \qquad \Rightarrow \qquad \begin{array}{c} \\ a \\ b \\ c \\ d \end{array} \begin{array}{c} \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

- Sono largamente usate nella grafica al computer; infatti, ogni immagine è una matrice di pixel con migliaia o milioni di righe e colonne.

- `Google`, il notissimo motore di ricerca in Internet, usa il calcolo degli autovalori di matrici di grandissime dimensioni per stabilire le correlazioni tra pagine web ogni volta che viene effettuata una ricerca.

L'importanza della teoria delle matrici non si esaurisce, però, nella funzione di una utile e semplice rappresentazione di fenomeni complessi, ma ha radici più profonde legate alle strutture algebriche di cui le matrici possono dotarsi.

Ricordiamo, innanzitutto, che l'insieme $\mathbb{K}^{m,n}$ delle matrici di m righe ed n colonne, a coefficienti nel campo $\mathbb{K}$, è uno spazio vettoriale rispetto alla somma di matrici ed al prodotto di una matrice per uno scalare. Inoltre $\dim \mathbb{K}^{m,n} = m \cdot n$.

Alla nomenclatura già introdotta in [16] sulle matrici, è necessario aggiungere e generalizzare alcune definizioni.

Sia $A = (a_{ij}) \in \mathbb{K}^{n,n}$. Allora si dicono *minori principali di* A le sottomatrici

$$A_k \stackrel{\text{def}}{=} (a_{ij})_{1 \le i,j \le k} \quad \text{per} \quad k = 1, \ldots, n; \tag{1.1.1}$$

graficamente,

$$A = A_n = \begin{pmatrix} \boxed{A_1} & & \\ & \boxed{A_2} & \\ & & \ddots \\ & & & A_n \end{pmatrix}.$$

Le nozioni di 'matrice triangolare' e 'matrice diagonale' sono estese a matrici rettangolari: se $A \in \mathbb{K}^{m,n}$ allora

1. A è *triangolare superiore* se $a_{ij} = 0$ per $i > j$;
2. A è *triangolare inferiore* se $a_{ij} = 0$ per $i < j$;
3. A è *diagonale* se $a_{ij} = 0$ per $i \neq j$;
4. A è *di Hessenberg superiore* se $a_{ij} = 0$ per $i > j + 1$;
5. A è *di Hessenberg inferiore* se $a_{ij} = 0$ per $j > i + 1$;
6. A è *tridiagonale* se $a_{ij} = 0$ per $j > i + 1$ ed $i > j + 1$;
7. A è di permutazione se si ottiene da I con scambi di righe e colonne.

Il vettore $(a_{11}, \ldots, a_{kk})$, con $k = \min\{m, n\}$, si dice *diagonale principale* di A. I vettori $(a_{1,1+h}, \ldots, a_{k,k+h})$, $(a_{1+h,1}, \ldots, a_{k+h,k})$[1] si dicono, rispettivamente, *h-esima sopradiagonale* ed *h-esima sottodiagonale*. Una matrice tridiagonale ha gli elementi al di fuori della diagonale, della prima sopradiagonale e della prima sottodiagonale nulli, ed è simultaneamente di Hessenberg superiore ed inferiore.

Siano $n > m$, e $\lambda_1, \ldots, \lambda_m \in \mathbb{K}$. Allora la matrice diagonale $A = (a_{ij}) \in \mathbb{K}^{n,m}$ tale che $a_{ij} = 0$ se $i \neq j$ e $a_{ii} = \lambda_i$ per $i = 1, \ldots, m$ si indica con $\mathrm{diag}(\lambda_1, \ldots, \lambda_m)$. La definizione si ripete banalmente per il caso $n < m$.

[1] Se c'è possibilità di equivoco, al posto di a_{ij} scriveremo $a_{i,j}$.

Esercizio 1.1. *Sia $\mathcal{T}_U^{m,n}$ l'insieme delle matrici triangolari superiori in $\mathbb{K}^{m,n}$. Dimostrare che $\mathcal{T}_U^{m,n}$ è un sottospazio vettoriale di $\mathbb{K}^{m,n}$, trovarne la dimensione ed una base. Ripetere l'esercizio per l'insieme delle matrici triangolari inferiori $\mathcal{T}_L^{m,n}$.*

Una matrice $A \in \mathbb{K}^{n,n}$ si dice *unipotente superiore (inferiore)* se è triangolare superiore (inferiore) e $a_{ii} = 1$ per $i = 1, \ldots, n$.

Esercizio 1.2. *Dimostrare che le matrici unipotenti costituiscono un sottogruppo di $GL(n, \mathbb{K})$. Dimostrare che* non *costituiscono un sottospazio vettoriale di $\mathbb{K}^{n,n}$.*

Osservazione 1.1. Una matrice tridiagonale è detta anche *matrice di Jacobi* di ordine n. Se

$$J_n = \begin{pmatrix} a_1 & b_1 & 0 & \ldots & 0 \\ c_1 & a_2 & b_2 & \ldots & 0 \\ 0 & c_2 & a_3 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & b_{n-1} \\ 0 & \ldots & \ldots & c_{n-1} & a_n \end{pmatrix}$$

si vede facilmente che

$$\det J_n = a_n \det J_{n-1} - b_{n-1}c_{n-1} \det J_{n-2} \qquad n \geq 3.$$

Se $a_i = 1 = b_i$, $c_i = -1$, ponendo $\det J_k = F_k$ si ha

$$F_n = F_{n-1} + F_{n-2}, \qquad \underline{\text{successione di Fibonacci.}}$$

Posto $F_0 = 1 = F_1$ si ottiene la successione

$$1, 1, 2, 3, 5, 8, 13, 21, \ldots$$

Questa è la famosa successione introdotta da Leonardo da Pisa (o Leonardus Pisanus), detto anche il Fibonacci, vissuto nel secolo XII; egli ha introdotto in Europa tra il 1100 ed il 1200 il sistema di numerazione posizionale indiano, a lui pervenuto tramite gli arabi. La successione fu introdotta nel *Liber Abaci* (1202) per risolvere il seguente problema:

> Quante coppie di conigli possono essere prodotti dalla coppia iniziale in un anno supponendo che ogni mese ogni coppia produca una nuova coppia in grado di riprodursi a sua volta dal secondo mese?

Il matematico francese Binet (1786–1856) scoprì una formula, nota già ad Eulero (1707–1783), che permette di calcolare direttamente un qualsiasi numero di Fibonacci:

$$F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right)$$

dove $\Phi = (1+\sqrt{5})/2 = 1,6180339887\ldots$ è il valore del raggio aureo, che è un numero irrazionale, mentre F_n è sempre un intero, malgrado l'aspetto della formula precedente.

1.2 Norme sulle matrici

Definizione 1.1. Sia V uno spazio vettoriale su campo $\mathbb{K}$, con $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$. Si dice *norma* su V un'applicazione

$$\|\cdot\| : V \to \mathbb{R}$$

tale che

1. $\|\vec{x}\| \geq 0$ per ogni $\vec{x} \in V$, e $\|\vec{x}\| = 0$ se e solo se $\vec{x} = \vec{0}$;
2. $\|c\vec{x}\| = |c|\|\vec{x}\|$ per ogni $\vec{x} \in V$, $c \in \mathbb{K}$;
3. $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$ per ogni $\vec{x}$, $\vec{y} \in V$.

Lo spazio $(V, \|\cdot\|)$ si dice *spazio normato*. Si osservi che uno spazio euclideo (V, g) è anche uno spazio normato rispetto alla norma

$$\|\vec{x}\|_g \stackrel{\text{def}}{=} \sqrt{g(\vec{x}, \vec{x})}$$

indotta dal prodotto scalare g. Viceversa, non tutte le norme provengono da un prodotto scalare, come si può dimostrare con esempi. Tuttavia, se vale l'*identità del parallelogramma*

$$\|\vec{x} + \vec{y}\|^2 + \|\vec{x} - \vec{y}\|^2 = 2(\|\vec{x}\|^2 + \|\vec{y}\|^2),$$

si può definire un prodotto scalare g ponendo

$$g(\vec{x}, \vec{y}) \stackrel{\text{def}}{=} \frac{1}{2}(\|\vec{x} + \vec{y}\|^2 - \|\vec{x}\|^2 - \|\vec{y}\|^2).$$

Si noti che la dimostrazione di quest'ultima affermazione *non è banale*.

Ora, si vogliono introdurre norme sullo spazio vettoriale $\mathbb{K}^{m,n}$. Come è noto, $\dim \mathbb{K}^{m,n} = \dim \mathbb{K}^{mn}$, quindi $\mathbb{K}^{m,n} \cong \mathbb{K}^{mn}$. Un particolare isomorfismo, interessante per le applicazioni, è quello indotto dall'ordine lessicografico

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \stackrel{\varphi}{\mapsto} (a_{11}, \ldots, a_{1n}, \ldots, a_{m1}, \ldots, a_{mn}).$$

Quest'isomorfismo permette di trasportare su $\mathbb{K}^{m,n}$ molti concetti familiari in $\mathbb{K}^{mn}$. Ad esempio, se g è un prodotto scalare in $\mathbb{K}^{mn}$, si ottiene un prodotto scalare g' su $\mathbb{K}^{m,n}$ ponendo

$$g'(A, B) \stackrel{\text{def}}{=} g(\varphi(A), \varphi(B)).$$

In tal modo viene introdotto in $\mathbb{K}^{m,n}$ un concetto di lunghezza, di misura, di "vicinanza" tra matrici, indispensabile nella teoria dell'approssimazione. Così, se $\mathbb{K} = \mathbb{C}$ e g è l'usuale prodotto scalare euclideo, possiamo introdurre in $\mathbb{K}^{m,n}$ una *norma* per le matrici ponendo

$$\|A\|_F \stackrel{\text{def}}{=} \sqrt{\sum_{i,j} |a_{ij}|^2} \qquad \underline{\text{norma di Frobenius}} \text{ o } \underline{\text{euclidea}}.$$

Invece, per $A \in \mathbb{K}^{n,n}$, la norma

$$\|A\|_p \stackrel{\text{def}}{=} \left(\sum_{i,j} |a_{ij}|^p \right)^{1/p} \qquad \underline{\text{norma } L_p}$$

non proviene da un prodotto scalare per $p \neq 2$. Per $p = 1$ la norma precedente è detta anche *norma del tassista*, o *norma di Manhattan*:

$$\|A\|_1 = \sum_{i,j=1}^{n} |a_{ij}|.$$

Per $p = 2$ si ha la norma di Frobenius.

Per $p = \infty$ si definisce la *norma del massimo*:

$$\|A\|_\infty = \max_{1 \leq i,j \leq n} |a_{ij}|.$$

Esercizio 1.3. *Definita la circonferenza unitaria in* $\mathbb{R}^2$ *dotato della norma* $\|\cdot\|_p$ *come*

$$S^1 = \{\vec{x} \in \mathbb{R}^2 \mid \|\vec{x}\|_p = 1\},$$

disegnare S^1 *nei casi* $p = 1$ *e* $p = \infty$.

Si possono generare molte norme di matrici nel modo seguente.

Definizione 1.2. Sia $\|\cdot\|$ una fissata norma in $\mathbb{K}^n$, con $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$. Si definisce *norma naturale* (o norma indotta dalla norma vettoriale $\|\cdot\|$) della matrice $A \in \mathbb{K}^{n,n}$ il numero

$$|||A||| \stackrel{\text{def}}{=} \sup_{\vec{x} \neq \vec{0}} \frac{\|A\vec{x}\|}{\|\vec{x}\|}$$

(che indichiamo con $|||\cdot|||$ per distinguerla dalle norme vettoriali).

Poiché per ogni $\vec{x} \neq \vec{0}$ si può definire $\vec{u} = \vec{x}/\|\vec{x}\|$ in modo che $\|\vec{u}\| = 1$, la definizione precedente è equivalente a

$$|||A||| = \max_{\|\vec{u}\|=1} \|A\vec{u}\|.$$

Se $A = I$ si ha $|||I||| = 1$. Inoltre, dalla definizione si ha $\|A\vec{x}\| \leq |||A|||\, \|\vec{x}\|$.

Si osservi che la norma di Frobenius non è subordinata ad una norma vettoriale, poiché $\|I\|_F = \sqrt{n}$.

La norma naturale di $A = (a_{ij})$ indotta da $\|\cdot\|_1$ è

$$|||A|||_1 = \max_{1 \leq j \leq n} \sum_{i=1}^{n} |a_{ij}|,$$

ed è detta brevemente *massima somma di colonna* di A.

La norma naturale di $A = (a_{ij})$ indotta da $\|\cdot\|_\infty$ è

$$|||A|||_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|,$$

ed è detta brevemente *massima somma di riga* di A.

Osservazione 1.2. Ogni norma naturale su $\mathbb{C}^{n,n}$ soddisfa la proprietà

$$|||AB||| \le |||A|||\ |||B|||, \tag{1.2.2}$$

detta di *submoltiplicatività*.

Non tutte le norme di matrici soddisfano tale proprietà: ad esempio, le matrici

$$A = B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

non soddisfano la proprietà (1.2.2) se si considera $\|\cdot\|_\infty$. La norma di Frobenius, tuttavia, soddisfa la proprietà (1.2.2). Verificare queste affermazioni per esercizio (in caso di difficoltà, è possibile trovare queste dimostrazioni in qualsiasi libro di analisi numerica). Molti testi, nella definizione di norma di una matrice, richiedono, oltre alle proprietà della definizione 1.1, anche la submoltiplicatività, che è utile, come evidenziato nella successiva osservazione. □

Osservazione 1.3. Se $|||\cdot|||$ è una norma di matrice verificante la submoltiplicatività, allora esiste una norma vettoriale $\|\cdot\|$ su $\mathbb{C}^n$ tale che

$$\|AX\| \le |||A|||\ \|X\| \qquad \forall A \in \mathbb{C}^{n,m}, \quad \forall X \in \mathbb{C}^{m,1}.$$

Basta considerare la matrice M avente come prima colonna X e le altre nulle, e porre per definizione $\|X\| = |||M|||$.

Osservazione 1.4. La definizione 1.2 si estende in modo ovvio al caso di matrici rettangolari $A \in \mathbb{C}^{m,n}$ valutando $\|\vec{x}\|$ con la norma in $\mathbb{C}^n$ ed $\|A\vec{x}\|$ con quella in $\mathbb{C}^m$. □

Osservazione 1.5. In $(\mathbb{R}^{n,n}, g')$, dove g' è indotta dall'usuale prodotto scalare in $\mathbb{R}^{n^2}$, l'espressione della disuguaglianza di Schwarz è

$$|g'(A,B)| \le \sqrt{g'(A,A)}\sqrt{g'(B,B)} \quad \Rightarrow \quad |\sum_{i,j} a_{ij}b_{ij}| \le \left(\sum_{i,j} a_{ij}^2\right)^{1/2} \left(\sum_{i,j} b_{ij}^2\right)^{1/2}.$$

Ponendo $B = \mathrm{Id}$ si ha

$$|\operatorname{tr} A| \le \left(\sum_{i,j} a_{ij}^2\right)^{1/2} \sqrt{n},$$

da cui la nota disuguaglianza di Newton, utile in Calcolo Numerico,

$$\|A\|_{g'} = \left(\sum_{i,j} a_{ij}^2\right)^{1/2} \geq \frac{1}{\sqrt{n}} |\operatorname{tr} A|. \ \square$$

1.3 Matrici a blocchi

In diversi problemi le matrici in considerazione vengono suddivise in *blocchi* (o sottomatrici) sia perché certi sottoinsiemi hanno un significato speciale, sia per motivi di comodità di calcolo. Infatti, quando si lavora con matrici di grandi dimensioni è importante poter riguardare ciascuna matrice come insieme di matrici di dimensioni inferiori. Ad esempio, se $A \in \mathbb{K}^{m,n}$ si può scrivere

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1r} \\ A_{21} & A_{22} & \cdots & A_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ A_{s1} & A_{s2} & \cdots & A_{sr} \end{pmatrix}$$

dove A_{ij} sono a loro volta matrici. Naturalmente, $A_{i1}, \ldots, A_{ir}$ hanno lo stesso numero di righe, mentre $A_{1j}, \ldots, A_{sj}$ hanno lo stesso numero di colonne.

Esempio 1.1.

$$A = \left(\begin{array}{cc:ccc:c} 1 & 2 & -1 & 0 & 3 & 1 \\ 0 & 1 & 2 & -1 & 0 & 0 \\ 2 & 0 & 1 & 0 & -1 & 0 \\ \hdashline 0 & 4 & -1 & 0 & 0 & 0 \\ 3 & 0 & 0 & 1 & 0 & 0 \end{array}\right), \qquad A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{23} & A_{33} \end{pmatrix}$$

dove

$$A_{11} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 2 & 0 \end{pmatrix}, \quad A_{12} = \begin{pmatrix} -1 & 0 & 3 \\ 2 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}, \ldots$$

Si osservi che, nel caso di un sistema lineare $AX = B$, la matrice completa risulta essere la matrice a blocchi $\tilde{A} = (A \,\vdots\, B)$.

Se $A_{ij} = O$ per $i > j$ $(i < j)$ allora A è detta matrice *triangolare superiore (inferiore) a blocchi*.

Se $A_{ij} = O$ per $i \neq j$ allora A è detta matrice *diagonale a blocchi*.

Si dimostra facilmente che la moltiplicazione tra due matrici a blocchi è equivalente ad una matrice i cui blocchi sono la moltiplicazione 'righe di blocchi' per 'colonne di blocchi'. Supponiamo, infatti, che una matrice $A \in \mathbb{K}^{m,n}$ sia ripartita nella matrice riga delle sue

colonne, o nella matrice colonna delle sue righe:

$$A = \begin{pmatrix} C_1 & \cdots & C_n \end{pmatrix} = \begin{pmatrix} R_1 \\ \vdots \\ R_m \end{pmatrix}.$$

Se $B \in \mathbb{K}^{n,p}$, con $B = (C'_1 \cdots C'_p)$, risulta $AB = (AC'_1 \dots AC_p)$, e

$$AB = \begin{pmatrix} R_1C'_1 & \cdots & R_1C'_p \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ R_mC'_1 & \cdots & R_mC'_p \end{pmatrix}$$

Se $X \in \mathbb{K}^{n,1}$, $X = (x_1, \dots, x_n)^T$, risulta $AX = x_1C_1 + \cdots + x_nC_n$.

Le precedenti considerazioni si generalizzano facilmente al caso di matrici partizionate a blocchi di dimensione arbitraria (con il vincolo della moltiplicabilità tra blocchi).

Proposizione 1.1. *Siano $A \in \mathbb{K}^{m,n}$, $B \in \mathbb{K}^{n,p}$ matrici partizionate in blocchi $A_{ik} \in \mathbb{K}^{m_i,n_k}$, $B_{kj} \in \mathbb{K}^{n_k,p_j}$, dove $i = 1, \dots, s$, $k = 1, \dots, r$ $j = 1, \dots, t$. Allora la matrice prodotto AB è partizionata nei blocchi $(AB)_{ij} \in \mathbb{K}^{m_i,p_j}$, dove*

$$(AB)_{ij} = \sum_{k=1}^{r} A_{ik}B_{kj}.$$

Sia, ora $A \in \mathbb{K}^{n,n}$. Se M è una sottomatrice quadrata di A si dice *minore complementare di M* il minore M' ottenuto da A togliendo tutte le righe e colonne di A che concorrono a formare M. Si chiama *cofattore di M* l'espressione $(-1)^{s_M} \det M'$, dove $s_M = i_1 + \cdots + i_k + j_1 + \cdots + j_k$ ed $i_1, \dots, i_k$ sono le righe di A che concorrono a formare M, $j_1, \dots, j_k$ sono le colonne di A che concorrono a formare M. Si noti che queste definizioni generalizzano l'ordinaria nozione di minore complementare e cofattore rispetto ad un elemento di A.

Teorema 1.1 (Laplace). *Sia $A \in \mathbb{K}^{n,n}$, e si fissino k righe (o colonne) di A. Allora $\det A$ è uguale alla somma dei prodotti dei determinanti di tutti i minori di ordine k estratti da tali righe (o colonne) per i rispettivi cofattori.*

Corollario 1.1. *Sia $A \in \mathbb{K}^{n,n}$ una matrice triangolare (superiore o inferiore) a blocchi, con blocchi diagonali quadrati. Allora si ha*

$$\det A = \prod_{i=1}^{r} \det A_{ii}$$

Esempio 1.2.

- Si consideri la matrice quadrata di ordine $n + m$ così definita

$$M = \begin{pmatrix} A & B \\ O & C \end{pmatrix} \qquad \begin{matrix} A \in \mathbb{K}^{n,n} & B \in \mathbb{K}^{n,m} \\ O \in \mathbb{K}^{m,n} & C \in \mathbb{K}^{m,m} \end{matrix}$$

Allora $\det M = \det A \det C$; quindi M è invertibile se e solo se A e C lo sono. In tal caso risulta

$$M^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}BC^{-1} \\ O & C^{-1} \end{pmatrix}$$

(suggerimento: si consideri una matrice M' a blocchi e si richieda $MM' = I$).

- Si consideri la matrice quadrata di ordine $2n$ così definita

$$N = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \qquad A, B, C, D \in \mathbb{K}^{n,n}.$$

In generale, $\det N \neq \det(AD - BC)$. Ma se D è invertibile si ha

$$\det N = \det(AD - BD^{-1}CD) :$$

infatti

$$\det N = \det N \det \begin{pmatrix} I & O \\ -D^{-1}C & I \end{pmatrix} = \det \left(\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} I & O \\ -D^{-1}C & I \end{pmatrix} \right),$$

da cui il risultato. Inoltre, se $CD = DC$ (ad esempio, se $C = \lambda I$ con $\lambda \in \mathbb{R}$) si ha

$$\det N = \det(AD - BC).$$

1.4 Forme canoniche delle matrici

Come è noto, la matrice di un endomorfismo rispetto ad una base di autovettori è diagonale, ma non sempre è possibile ottenere una matrice con un'espressione così semplice. La teoria delle forme canoniche si occupa della ricerca di basi rispetto alle quali la matrice di un endomorfismo assuma forme particolarmente semplici, usualmente a blocchi soddisfacenti determinate caratteristiche.

Per fare questo, invece del concetto, troppo restrittivo, di 'autospazio', si introduce il concetto più generale di 'sottospazio invariante'.

Sia V uno spazio vettoriale su un campo $\mathbb{K}$, e sia W un suo sottospazio. Si dice che W è *invariante* (o *stabile*) rispetto ad un endomorfismo $f\colon V \to V$ se $f(W) \subset W$ (o, equivalentemente, se $f|_W$ è un endomorfismo di W).

Sia $\dim V = n$, $\dim W = k$ e $\mathcal{B}'$ una base di W. Se $\mathcal{B}$ è una base di V ottenuta da $\mathcal{B}'$ per completamento, allora

$$\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} A & B \\ O & C \end{pmatrix}, \qquad \begin{matrix} A \in \mathbb{K}^{k,k} & B \in \mathbb{K}^{k,n-k} \\ O \in \mathbb{K}^{n-k,k} & C \in \mathbb{K}^{n-k,n-k} \end{matrix}$$

Naturalmente, $A = \mathcal{M}_{\mathcal{B}'}^{\mathcal{B}'}(f|_W)$.

Ovviamente, per quanto studiato sugli autovalori, gli autospazi sono sottospazi invarianti rispetto ad f, ma non tutti i sottospazi invarianti sono autospazi rispetto ad un qualche autovalore.

Inoltre, non è detto che V sia decomponibile in somma diretta di sottospazi *propri* invarianti. Se questo succede, la matrice di f rispetto ad una base di vettori dei sottospazi invarianti è una matrice diagonale a blocchi. Se, in particolare, V è somma diretta degli autospazi (e quindi f è un endomorfismo semplice), allora i blocchi diagonali assumono la forma $A_{ii} = \lambda_i I$, quindi la matrice è diagonale.

Esempio 1.3. Sia $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ tale che la sua matrice rispetto alla base canonica sia $A = \left(\begin{smallmatrix} 2 & 1 \\ 0 & 2 \end{smallmatrix}\right)$. Provare che l'unico sottospazio di dimensione 1 invariante rispetto ad f è l'autospazio $V(2)$.

Il seguente teorema assicura l'esistenza di un'ampia classe di endomorfismi che ammettono una matrice triangolare.

Teorema 1.2. *Sia $f\colon V \to V$ un endomorfismo di uno spazio vettoriale V su campo $\mathbb{K}$ di dimensione n. Supponiamo che il polinomio caratteristico di f si decomponga in fattori di primo grado in $\mathbb{K}$. Allora esistono $n+1$ sottospazi vettoriali V_0, V_1,..., V_n tali che*

1. *$f(V_h) \subset V_h$ per $h = 0, 1, 2, \ldots, n$;*
2. *$\dim V_h = h$;*
3. *$\{\vec{0}\} = V_0 \subset V_1 \subset \cdots \subset V_n = V$.*

La successione di sottospazi $V_0 \subset V_1 \subset \cdots \subset V_n$ si dice successione di sottospazi *a bandiera*. Una base

$$\mathcal{B} = \{\vec{e}_1, \ldots, \vec{e}_n\} \quad \text{tale che} \quad V_h = \mathcal{L}(\vec{e}_1, \ldots, \vec{e}_h)$$

è detta base *a bandiera*. Allora per l'invarianza si ha

$$\begin{aligned} f(\vec{e}_1) &= t_{11}\vec{e}_1, \\ f(\vec{e}_2) &= t_{12}\vec{e}_1 + t_{22}\vec{e}_2, \\ &\cdots \\ f(\vec{e}_n) &= t_{1n}\vec{e}_1 + \cdots + t_{nn}\vec{e}_n, \end{aligned}$$

e quindi

$$\Delta \stackrel{\text{def}}{=} \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} t_{11} & t_{12} & \ldots & t_{1n} \\ 0 & t_{22} & \ldots & t_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & t_{nn} \end{pmatrix}$$

è una matrice triangolare superiore. Ovviamente gli autovalori di Δ sono t_{ii}.

Una matrice $A \in \mathbb{K}^{n,n}$ si dice *triangolarizzabile*[2] se è simile ad una matrice triangolare. Allora il teorema precedente si può formulare nel seguente modo.

Teorema 1.3. *Sia $A \in \mathbb{K}^{n,n}$ tale che il suo polinomio caratteristico* $\det(A - \lambda I)$ *si decomponga in fattori di primo grado in $\mathbb{K}$. Allora A è triangolarizzabile.*

Ovviamente, se $\mathbb{K}$ è algebricamente chiuso (come $\mathbb{C}$), il teorema vale per qualsiasi matrice in $\mathbb{K}^{n,n}$.

Osservazione 1.6. Siano A, $\tilde{A} \in \mathbb{K}^{n,n}$ due matrici simili, ovvero $\tilde{A} = B^{-1}AB$ per una certa matrice invertibile B. Allora, se $h > 0$ si ha

$$\tilde{A}^h = (B^{-1}AB)^h = B^{-1}A^hB,$$

quindi $\tilde{A}^h$ è simile ad A^h. In particolare, se A è invertibile e simile ad una matrice triangolare Δ (come nel caso $\mathbb{K} = \mathbb{C}$), risulta

$$\Delta^{-1} = (B^{-1}AB)^{-1} = B^{-1}A^{-1}B,$$

cioè Δ^{-1} è simile ad A^{-1}, ma l'inversa di una matrice triangolare si calcola facilmente. Si noti che gli autovalori di Δ^h sono le potenze h-esime degli autovalori di Δ. □

Come vedremo più avanti, si può dire di più sulle matrici su di un campo algebricamente chiuso. Più precisamente proveremo il seguente teorema.

Teorema 1.4. *Sia $\mathbb{K}$ un campo algebricamente chiuso. Ogni matrice $A \in \mathbb{K}^{n,n}$ è simile ad una matrice J, detta* forma di Jordan[3] *di A, avente le seguenti proprietà:*

1. *J è triangolare superiore;*
2. *lungo la diagonale principale di J compaiono gli autovalori di A ripetuti con la loro molteplicità;*
3. *gli $n-1$ elementi di posizione $(i, i+1)$ (che si trovano 'sopra' la diagonale di J) sono uguali a 0 o ad 1; gli elementi di posizione $(i, i+s)$ con $s > 1$ sono nulli.*

Esempio 1.4. La seguente matrice è in forma di Jordan:

$$J = \left(\begin{array}{cc:ccc:c} -2 & 1 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & 3 \end{array}\right)$$

[2] Non si confonda questa nozione con la trangolarizzazione dei *sistemi lineari*!
[3] Camille Jordan (1838–1922)

Naturalmente, gli autovalori di J sono -2 con molteplicità algebrica 2, 0 con molteplicità algebrica 3 e 3 con molteplicità algebrica 1.

Si chiama *blocco di Jordan* di ordine k associato ad un autovalore λ di A la matrice quadrata di ordine k del tipo

$$\begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \lambda & 1 \\ 0 & \dots & \dots & \dots & 0 & \lambda \end{pmatrix}$$

Ogni forma di Jordan di una matrice A che ha gli autovalori distinti $\lambda_1, \dots, \lambda_r$ ha lungo la diagonale principale blocchi di Jordan.

1.5 Matrici complesse

Nel caso in cui $\mathbb{K} = \mathbb{C}$, alle operazioni che sono state già introdotte in [16] possono essere aggiunte altre operazioni, che verranno descritte in questo paragrafo.

Poiché il campo $\mathbb{C}$ è algebricamente chiuso, ogni matrice quadrata $A \in \mathbb{C}^{n,n}$ ammette la forma canonica di Jordan. Sarà dimostrato nel seguito che, nel caso in cui A soddisfi un'ulteriore ipotesi (che è una generalizzazione al campo complesso della simmetria richiesta per le matrici reali), la forma canonica di Jordan diventa diagonale.

Definizione 1.3. Sia $A = (a_{ij}) \in \mathbb{C}^{m,n}$. Allora si dice

1. *coniugata di* A la matrice $\overline{A} \stackrel{\text{def}}{=} (\overline{a_{ij}})$;
2. *aggiunta di* A la matrice $A^* \stackrel{\text{def}}{=} \overline{A}^T = \overline{A^T}$.

Si dimostrano facilmente le seguenti proprietà:

1. $A^{**} = A$
2. $(A + B)^* = A^* + B^*$ e $(\lambda A)^* = \overline{\lambda} A^*$ per $\lambda \in \mathbb{C}$;
3. $(AB)^* = B^* A^*$;
4. $\det \overline{A} = \overline{\det A}$, $\det A^* = \overline{\det A}$;
5. se A è invertibile, $(A^*)^{-1} = (A^{-1})^*$.

Definizione 1.4. Sia $A = (a_{ij}) \in \mathbb{C}^{n,n}$. Allora A si dice

1. *hermitiana* se $A^* = A$;
2. *antihermitiana* se $A^* = -A$;

3. *unitaria* se $AA^* = A^*A = I$;

4. *di fase* se A è unitaria e diagonale;

5. *normale* se $AA^* = A^*A$.

Si noti che se $A \in \mathbb{R}^{n,n} \subset \mathbb{C}^{n,n}$, allora A è hermitiana se e solo se A è simmetrica, ed A è unitaria se e solo se A è ortogonale.

Esercizio 1.4. *Dimostrare che*

- A *hermitiana* $\Rightarrow a_{ii} \in \mathbb{R}$, $\det A \in \mathbb{R}$;
- A *unitaria* $\Rightarrow |\det A| = 1$;

provare che non vale l'implicazione opposta.

Esempio 1.5.

- Sia
$$A = \begin{pmatrix} -1 & 3-i \\ 3+i & -2 \end{pmatrix}, \qquad \overline{A} = \begin{pmatrix} -1 & 3+i \\ 3-i & -2 \end{pmatrix},$$
dunque A è hermitiana. Essendo $\det A = -8$, A non è unitaria.

- Se $\alpha \in \mathbb{R}$, sia
$$B = \begin{pmatrix} e^{\alpha i} & 0 \\ 0 & e^{-\alpha i} \end{pmatrix}.$$
Essendo $e^{\alpha i} = \cos\alpha + i \sin\alpha$, risulta $\overline{e^{\alpha i}} = e^{-\alpha i}$, dunque $\overline{B} = B^*$, quindi B è hermitiana se e solo se $\alpha = k\pi$, con $k \in \mathbb{Z}$. Ma $BB^* = I$, dunque B è unitaria.

Si pone
$$U(n, \mathbb{C}) \stackrel{\text{def}}{=} \{A \in \mathbb{C}^{n,n} \mid A \text{ unitaria}\} \subset \mathbb{C}^{n,n}.$$

Esercizio 1.5. *Provare che $U(n, \mathbb{C})$ è un gruppo rispetto al prodotto di matrici. È l'analogo del gruppo $O(n, \mathbb{R})$ del gruppo delle matrici ortogonali. Provare, inoltre, che l'insieme delle matrici di fase costituisce un sottogruppo di $U(n, \mathbb{C})$.*

Le matrici unitarie vengono molto usate nell'analisi numerica per trasformare altre matrici in modo da lasciare invariata la loro norma, così da ridurre l'introduzione di errori durante queste trasformazioni. La giustificazione matematica di questa affermazione sta nella seguente proposizione.

Proposizione 1.2. *Sia $A \in \mathbb{C}^{m,n}$, e siano $P \in U(m, \mathbb{C})$, $Q \in U(n, \mathbb{C})$ matrici unitarie. Allora*
$$\|PA\|_F = \|A\|_F = \|AQ\|_F$$

DIMOSTRAZIONE. Se $A = (C_1, \ldots, C_n)$, dove C_j è la j-esima colonna di A, risulta

$$\|PA\|_F^2 = \|(PC_1, \ldots, PC_n)\|_F^2 = \sum_j \|PC_j\|_F^2 = \sum_j \|C_j\|_F^2 = \|A\|_F^2,$$

tenendo conto del fatto che (cfr. Appendice)

$$\|PC_j\|_F^2 = (PC_j)^*(PC_j) = C_j^*(P^*P)C_j = \|C_j\|_F^2.$$

Analogamente per l'altra uguaglianza. QED

Esercizio 1.6.

- *Provare che* $\|A\|_F = \sqrt{\mathrm{tr}\,(AA^*)}$.
- *Provare che se* λ *è un autovalore di* A, *allora* $\overline{\lambda}$ *è un autovalore di* $\overline{A}$, *e quindi di* A^*.

Il teorema 1.3 ammette una formulazione più specifica per quanto riguarda la base che realizza la triangolarizzazione nel caso delle matrici complesse. La proprietà, data senza dimostrazione, è espressa dal seguente teorema.

Teorema di Schur[4]. *Sia* $A \in \mathbb{C}^{n,n}$. *Allora esiste una matrice* <u>*unitaria*</u> U *tale che*

$$U^*AU = \Delta,$$

ove Δ *è triangolare superiore.*

DIMOSTRAZIONE. Si usi l'induzione sull'ordine n della matrice. Il primo passo è banale. Se $A \in \mathbb{C}^{n,n}$ sia X_1 un autovettore di A di modulo 1. Si consideri una base ortonormale di $\mathbb{C}^n$ del tipo $\mathcal{B} = \{X_1, Y_2, \ldots, Y_n\}$. La funzione lineare f_A definita da $f_A(X) = AX$ assume, nella base $\mathcal{B}$, la forma

$$Q^*AQ = A' = \begin{pmatrix} \lambda_1 & B \\ O & C \end{pmatrix}$$

dove λ_1 è l'autovalore di X_1 e Q è unitaria. Per l'ipotesi induttiva, $V^*CV = D$, dove D è triangolare superiore e V è unitaria. Introdotta la matrice unitaria

$$Z = \begin{pmatrix} 1 & O \\ O & V \end{pmatrix}$$

risulta

$$Z^*Q^*AQZ = \begin{pmatrix} \lambda_1 & B' \\ O & D \end{pmatrix},$$

da cui la tesi, ponendo $U = QZ$. QED

[4]Issai Schur (1875–1941)

Essendo $U^* = U^{-1}$, A è simile alla matrice triangolare Δ, dunque A e Δ hanno gli stessi autovalori.

Corollario 1.2. *Se $A \in \mathbb{R}^{n,n} \subset \mathbb{C}^{n,n}$ ha tutti gli autovalori reali, allora esiste una matrice ortogonale $U \in O(n, \mathbb{R})$ tale che $U^T AU = \Delta$.*

Dal precedente teorema si ottiene una conseguenza fondamentale.

Teorema spettrale I. *Se $A \in \mathbb{C}^{n,n}$ è hermitiana, allora esiste una matrice unitaria $U \in \mathbb{C}^{n,n}$ tale che*

$$U^* AU = D,$$

dove $D \in \mathbb{C}^{n,n}$ è diagonale.

Dimostrazione. Per il teorema di Schur si ha $U^* AU = \Delta$. Ma

$$\Delta^* = (U^* AU)^* = U^* A^* U = U^* AU = \Delta.$$

Dunque Δ è diagonale. QED

Il teorema è detto 'spettrale' poiché fa intervenire lo *spettro* di una matrice, ossia l'insieme dei suoi autovalori.

Il teorema continua a valere per le matrici reali sostituendo le parole "hermitiana" ed "unitaria" con "simmetrica" ed "ortogonale", e si riduce al risultato già noto [16]. Inoltre, il teorema vale anche per le matrici unitarie, nella forma seguente.

Teorema spettrale II. *Sia $A \in \mathbb{C}^{n,n}$. Allora A è normale se e solo se esiste una matrice unitaria $U \in \mathbb{C}^{n,n}$ tale che*

$$U^* AU = D,$$

dove $D \in \mathbb{C}^{n,n}$ è diagonale.

Dimostrazione. Sia $A = UDU^*$, e si ponga $D = (d_{ij})$. Allora

$$AA^* = (UDU^*)(UDU^*)^* = U(DD^*)U^*,$$
$$A^*A = (UDU^*)^*(UDU^*) = U(D^*D)U^*,$$

ma DD^* è una matrice diagonale i cui elementi diagonali sono $|d_{ii}|^2$, da cui la tesi.

Viceversa, per il teorema di Schur si ha $\Delta = U^* AU$, e da $AA^* = A^*A$ si ha

$$\Delta\Delta^* = U^* AUU^* A^* U = U^* AA^* U = U^* A^* AU = \Delta^*\Delta,$$

cioè Δ è normale.

Utilizzando l'induzione sull'ordine n della matrice, si dimostra che Δ è diagonale. Il primo passo è ovvio. Posto $\Delta = (t_{ij})$, ricordando che Δ è triangolare superiore, risulta per l'elemento di posto $(1,1)$ in $\Delta\Delta^* = \Delta^*\Delta$

$$|t_{11}|^2 + |t_{12}|^2 + \cdots + |t_{1n}|^2 = |t_{11}|^2 \quad \Rightarrow \quad t_{12} = \cdots = t_{1n} = 0,$$

dunque

$$\Delta = \begin{pmatrix} t_{11} & O \\ O & \Delta_1 \end{pmatrix}, \qquad \Delta^* = \begin{pmatrix} \overline{t_{11}} & O \\ O & \Delta_1^* \end{pmatrix}.$$

Essendo $\Delta\Delta^* = \Delta^*\Delta$, si prova facilmente che anche Δ_1 è normale, dunque diagonale per l'ipotesi induttiva. Da questo si ha la tesi. QED

Il teorema spettrale I nel caso di una matrice simmetrica è detto anche "teorema degli assi principali" per le applicazioni che esso ha nella diagonalizzazione delle forme quadratiche, e quindi nella ricerca della forma canonica delle quadriche.

Osservazione 1.7. Gli endomorfismi simmetrici di spazi vettoriali euclidei reali sono gli unici che si diagonalizzano con basi ortonormali. Infatti, se $f\colon V \to V$ è un endomorfismo simmetrico dello spazio vettoriale euclideo V e $\mathcal{B}$ è una base ortonormale, allora

$$\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f) = U^T D U,$$

con D diagonale ed $U \in O(n, \mathbb{R})$. Viceversa, se f è diagonalizzabile in $\mathbb{R}$ tramite basi ortonormali, ossia $\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f) = U^T D U$, risulta

$$(\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f))^T = \left(U^T D U\right)^T = \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f),$$

dunque la matrice è simmetrica. □

Osservazione 1.8. Ogni matrice quadrata $A \in \mathbb{C}^{n,n}$ si scompone in modo unico così:

$$A = H + K, \qquad H \stackrel{\text{def}}{=} \frac{1}{2}(A + A^*), \quad K \stackrel{\text{def}}{=} \frac{1}{2}(A - A^*), \quad A^* = H - K. \tag{1.5.3}$$

dove H è hermitiana e K è antihermitiana. Gli autovalori di una matrice hermitiana sono reali, gli autovalori di una matrice antihermitiana sono immaginari (verificarlo!). Se si fanno corrispondere le matrici hermitiane ai numeri reali e le matrici antihermitiane ai numeri immaginari, la scomposizione (1.5.3) corrisponde alla forma algebrica dei numeri complessi: $z = a + ib$. Allora le matrici unitarie possono essere fatte corrispondere ai numeri complessi di modulo 1. □

1.6 Matrici definite positive

Definizione 1.5. Una matrice $A \in \mathbb{C}^{n,n}$ è *definita positiva* se è hermitiana e

$$X^* A X > 0 \qquad \forall\, X \in \mathbb{C}^{n,1} \setminus \{O\}.$$

Si confronti la precedente definizione con la teoria sviluppata nel paragrafo A.1.

Lemma 1.1. *Una matrice $A \in \mathbb{C}^{n,n}$ hermitiana è* definita positiva *se e solo se i suoi minori principali (definizione (1.1.1)) sono definiti positivi.*

DIMOSTRAZIONE. Si noti, innanzitutto, che i minori principali di A sono hermitiani anch'essi. Per ogni $k = 1, \dots, n$ si può decomporre A a blocchi come segue:

$$A = \begin{pmatrix} A_k & P \\ Q & R \end{pmatrix}.$$

Sia $X \in \mathbb{C}^{n,1}$ tale che $X = (Y^* \; O)^*$ ed $Y \in \mathbb{C}^{k,1}$. Se $Y \neq O$, risulta $X \neq O$ e

$$X^* A X = Y^* A_k Y,$$

da cui la tesi. QED

Criterio di Sylvester. *Una matrice $A \in \mathbb{C}^{n,n}$ hermitiana è* definita positiva *se e solo se i suoi minori principali (definizione (1.1.1)) soddisfano le condizioni*

$$\det A_k > 0 \qquad k = 1, \dots, n.$$

DIMOSTRAZIONE. Usando il precedente lemma, risulta (per induzione) che gli autovalori di A sono tutti reali positivi, e quindi, poiché A_k è diagonalizzabile $\det A_k = \lambda_1 \cdots \lambda_k > 0$. QED

Ovviamente, per $k = n$ il teorema precedente dà $\det A = \det A_n > 0$.

Esempio 1.6. Sia

$$A = \begin{pmatrix} 1 & -2 & -1 \\ -2 & 6 & 0 \\ -1 & 0 & 5 \end{pmatrix}.$$

A è simmetrica e $A_1 = (1)$, $A_2 = \left(\begin{smallmatrix} 1 & -2 \\ -2 & 6 \end{smallmatrix}\right)$, $A_3 = A$, dunque A è definita positiva.

Esempio 1.7. Si consideri una funzione $f\colon \mathbb{R}^n \to \mathbb{R}$ di classe $\mathcal{C}^2$ e la sua matrice hessiana in $p \in \mathbb{R}^n$:

$$H(p) = (h_{ij}(p)) = \left(\frac{\partial^2 f}{\partial x^i \partial x^j}(p) \right).$$

La matrice $H(p)$ è simmetrica reale e dà informazioni sui punti critici di f. Più particolarmente, se p è un punto critico (ossia $\nabla f(p) = \vec{0}$), allora p è detto *non degenere* quando $\det H(p) \neq 0$. Il numero degli autovalori strettamente negativi di $H(p)$ è detto *indice del punto critico* p di f, $\operatorname{ind}_p(f)$. Ne segue

$$\begin{aligned} &p \text{ minimo locale di } f \quad \Leftrightarrow \quad \operatorname{ind}_p(f) = 0, \\ &p \text{ massimo locale di } f \quad \Leftrightarrow \quad \operatorname{ind}_p(f) = n. \end{aligned}$$

Se $0 < \operatorname{ind}_p(f) < n$, il punto p è un punto di sella e la forma associata è indefinita.

Le matrici definite positive intervengono spesso in problemi di ottimizzazione e trovano molteplici applicazioni nelle scienze e nell'ingegneria. Spesso, infatti, si tratta di massimizzare o minimizzare una forma quadratica soggetta a qualche vincolo. Il seguente teorema è di aiuto in questi casi.

Teorema 1.5. *Sia $A \in \mathbb{R}^{n,n}$ simmetrica. Si consideri la forma quadratica $Q(X) = X^T AX$. Siano λ_m e λ_M il minimo ed il massimo autovalore di A. Allora:*

1. $\max\{Q(X) \mid \|X\| \leq 1\} = \lambda_M$,
2. $\min\{Q(X) \mid \|X\| \leq 1\} = \lambda_m$.

Inoltre, se $\lambda_h \in \mathbb{C}$ sono gli autovalori di A per $h = 1, \ldots, n$, allora $\lambda_h = Q(F_h)$, dove F_h è un vettore nell'autospazio $V(\lambda_h)$.

Questo teorema è molto importante nelle applicazioni che coinvolgono problemi di vibrazione, dall'aerodinamica alla fisica delle particelle. I valori massimi e minimi descritti nel teorema vengono spesso determinati utilizzando tecniche di calcolo in più variabili.

1.7 Invertibilità di una matrice

Sia $A \in \mathbb{C}^{n,n}$. È illustrato un criterio per l'invertibilità di A che non richiede il calcolo del determinante, però la condizione è solo sufficiente.

Teorema 1.6 (Teorema della diagonale dominante). *Sia $A \in \mathbb{C}^{n,n}$. Condizione sufficiente affinché A sia invertibile è che ogni elemento della diagonale abbia modulo maggiore della somma degli elementi che si trovano sulla stessa riga, ossia*

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|, \qquad i = 1, \ldots, n.$$

Dimostrazione. Si supponga A non invertibile. Allora esiste $V \in \mathbb{C}^{n,1} \setminus \{O\}$ tale che $AV = O$. Se v_i è la componente di V con modulo massimo, $|v_i| > 0$ e

$$0 = (AV)_i = a_{i1}v_1 + \cdots + a_{ii}v_i + \cdots + a_{in}v_n$$

implica

$$|a_{ii}||v_i| = \Big|\sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij}v_j\Big| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}||v_i|,$$

da cui la tesi. QED

Osservazione 1.9. Ovviamente, vale un criterio analogo per le colonne di A. □

Osservazione 1.10. Si noti che il calcolo del determinante è molto dispendioso da un punto di vista computazionale: occorrono $n!$ moltiplicazioni di n oggetti ed n somme. Il criterio precedente consente di arrivare alle stesse conclusioni (quando possibile!) utilizzando solamente n somme di n elementi (una per ogni riga). □

1.8 Diagonalizzazione simultanea di matrici

Siano A, $B, \in \mathbb{C}^{n,n}$ due matrici diagonalizzabili. Ci chiediamo in quali casi sia possibile *diagonalizzarle simultaneamente*, ovvero in quali casi esista una base che sia costituita da autovettori sia di A sia di B, e quindi esista una matrice $C \in \mathbb{C}^{n,n}$ invertibile tale che

$$C^{-1}AC = D, \qquad C^{-1}BC = D',$$

dove D e D' sono le forme diagonali di A e B rispettivamente.

Questo problema è molto importante in meccanica quantistica ed in teoria delle rappresentazioni. La risposta è data dal seguente teorema.

Teorema 1.7. *Siano A, $B, \in \mathbb{C}^{n,n}$ due matrici diagonalizzabili. Esse sono simultaneamente diagonalizzabili se e solo se sono permutabili.*

Dimostrazione. Se sono simultaneamente diagonalizzabili, allora

$$AB = CDC^{-1}CD'C^{-1} = CDD'C^{-1} = CD'DC^{-1} = BA.$$

Viceversa, supponiamo che A e B siano diagonalizzabili e che $AB = BA$. Indichiamo con $V(\lambda_1), \dots, V(\lambda_r)$ gli autospazi relativi ad A, per cui

$$h_i = \dim V(\lambda_i) = m_i.$$

dove m_i è la molteplicità algebrica di λ_i. Per $\vec{v} \in V(\lambda_i)$ si ha

$$A(B\vec{v}) = B(A\vec{v}) = B(\lambda_i\vec{v}) = \lambda_i(B\vec{v}) \quad \Rightarrow \quad B\vec{v} \in V(\lambda_i),$$

quindi l'applicazione lineare associata a B muta ogni sottospazio $V(\lambda_i)$ di A in sé.

Sia ora $\vec{w}$ un autovettore per B, dunque $B\vec{w} = \mu\vec{w}$. Poiché $\vec{w} \in V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_r)$ si ha $\vec{w} = \vec{z}_1 + \cdots + \vec{z}_r$, con $\vec{z}_i \in V(\lambda_i)$ per ogni i, in modo unico. Ora, $B\vec{w} = B\vec{z}_1 + \cdots B\vec{z}_r$ e $B\vec{z}_i \in V(\lambda_i)$ per ogni i. Poiché la decomposizione $\mu\vec{w} = \mu\vec{z}_1 + \cdots + \mu\vec{z}_r$ è unica si ha

$$B\vec{z}_i = \mu\vec{z}_i \qquad i = 1, \dots, r.$$

Se per un certo indice i_0 si ha $\vec{z}_{i_0} \neq 0$, questo è un autovettore di A poiché $\vec{z}_{i_0} \in V(\lambda_{i_0})$, ed un autovettore di B poiché $B\vec{z}_{i_0} = \mu\vec{z}_{i_0}$.

Applicando il procedimento indicato sopra ad una base $\{\vec{w}_1, \dots, \vec{w}_n\}$ di autovettori di B (certo esistente poiché B è diagonalizzabile) otteniamo almeno n vettori che sono simultaneamente autovettori di A e B. Poiché i vettori $\vec{w}_i$ sono combinazione lineare degli autovettori simultanei, lo spazio da essi generato è uguale a $\mathcal{L}(\vec{w}_1, \dots, \vec{w}_n) = V$; quindi da essi si può estrarre una base. QED

Esercizio 1.7. *Si considerino le seguenti matrici A e B e si provi che sono simulta-*

neamente diagonalizzabili. Si determini inoltre una base comune di autovettori

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ -1 & 0 & 3 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 3 & 0 \\ -2 & 0 & 3 \end{pmatrix}$$

SOLUZIONE. Si vede facilmente che $AB = BA$; si osservi che gli autovalori delle matrici sono gli elementi sulla diagonale principale, e che risulta

$$V_A(2) = \mathcal{L}((1,0,1),(0,1,0)), \quad V_A(3) = \mathcal{L}((0,0,1)),$$
$$V_B(3) = \mathcal{L}((0,1,0),(0,0,1)), \quad V_B(1) = \mathcal{L}((1,1,1)),$$

dunque A e B sono diagonalizzabili. L'autovettore $\vec{w} = (0,0,1)$ di B si può decomporre univocamente come $\vec{w} = \vec{z}_1 + \vec{z}_2$, dove $\vec{z}_1 \in V_A(2)$ e $\vec{z}_2 \in V_A(3)$. In particolare, come è ovvio, $\vec{w} \in V_A(3)$. Si consideri ora $\vec{w} = (0,1,0)$; procedendo come prima si ha $\vec{w} \in V_A(2)$. Infine, preso $\vec{w} = (1,1,1)$, si ha $\vec{w} = \vec{z}_1 + \vec{z}_2$, con $\vec{z}_1 = (1,0,1) + (0,1,0)$ e $\vec{z}_2 = \vec{0}$. Dunque la base cercata è

$$(0,1,0), \quad (0,0,1), \quad (1,1,1).$$

□

1.9 Esercizi di riepilogo

Esercizio 1.8. *Sia $\mathcal{TGL}_U(n,\mathbb{K})$ l'insieme delle matrici triangolari superiori invertibili in $\mathbb{K}^{n,n}$. Dimostrare che $\mathcal{TGL}_U(n,\mathbb{K})$ è un sottogruppo di $GL(n,\mathbb{K})$. Ripetere l'esercizio per l'insieme delle matrici triangolari inferiori invertibili.*

(Suggerimento: si dimostri che l'inversa di una matrice triangolare superiore invertibile è triangolare superiore e che il prodotto di due matrici triangolari superiori è triangolare superiore.)

Esercizio 1.9. *Siano X, $Y \in \mathbb{C}^n$. Dimostrare che la matrice XY^* ha rango 1. (Suggerimento: per ogni $Z \in \mathbb{C}^n$ si ha $(XY^*)Z = X(Y^*Z)$.)*

Esercizio 1.10. *Siano X, $Y \in \mathbb{C}^n$. Dimostrare che la matrice XY^* ha traccia Y^*X.*

Esercizio 1.11. *Siano X, $Y \in \mathbb{C}^n$. Dimostrare che*

1. $\det(I + XY^*) = 1 + Y^*X$;
2. *se* $1 + Y^*X \neq 0$, *allora*

$$(I + XY^*)^{-1} = I - \frac{XY^*}{1 + Y^*X}$$

(Suggerimento: osservato che

$$\begin{pmatrix} I & O \\ -Y^* & 1 \end{pmatrix} \begin{pmatrix} I & -X \\ Y^* & 1 \end{pmatrix} = \begin{pmatrix} I & -X \\ O & 1 + Y^*X \end{pmatrix}$$

applicare il teorema di Binet.)

Esercizio 1.12. *Sia $A \in \mathbb{C}^{n,n}$ con $\det A \neq 0$, e siano X, $Y \in \mathbb{C}^n$. Dimostrare che:*

1. $\det(A + XY^*) = 0$ *se e solo se* $Y^*A^{-1}X = -1$*;*

2. *se* $\det(A + XY^*) \neq 0$*, allora vale la seguente* formula di Sherman–Morrison*:*

$$(A + XY^*)^{-1} = A^{-1} - \frac{A^{-1}XY^*A^{-1}}{1 + Y^*A^{-1}X}.$$

(Suggerimento: si usi l'esercizio precedente.ale)

Esercizio 1.13. *Dimostrare che le matrici hermitiane costituiscono un sottospazio vettoriale reale di* $\mathbb{C}^{n,n}$ *ma* non *un sottospazio complesso. Trovarne la dimensione.*

Esercizio 1.14. *Provare che, se* $A \in \mathbb{C}^{n,n}$ *è invertibile, allora* A *non può essere nilpotente.*

Esercizio 1.15. *Provare che se* $A \in \mathbb{R}^{n,n}$ *è simmetrica e nilpotente, allora* $A = O$*.*

Esercizio 1.16. *Determinare tutte le matrici* $A \in \mathbb{R}^{2,2}$ *tali che* $A^2 = O$*.*

Esercizio 1.17. *Sfruttando l'esercizio precedente, risulvere la seguente equazione matriciale con incognita* $X \in \mathbb{R}^{2,2}$*:*

$$X^2 - 2X + I = O.$$

Soluzione.

$$X = \begin{pmatrix} 1 & 0 \\ z & 1 \end{pmatrix},\ \forall z \in \mathbb{R}; \quad X = \begin{pmatrix} 1 & y \\ 0 & 1 \end{pmatrix},\ \forall y \in \mathbb{R};$$
$$X = \begin{pmatrix} 1-h & -h^2/k \\ k & 1+h \end{pmatrix},\ \forall h \in \mathbb{R},\ \forall k \in \mathbb{R} \setminus \{0\}.$$

□

Esercizio 1.18. *Posto*

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

provare che la matrice

$$R(t) = \frac{1 - t^2}{1 + t^2} I + \frac{2t}{1 + t^2} J$$

rappresenta, per ogni $t \in \mathbb{R}$*, una rotazione, e determinarne l'angolo* $\varphi(t)$*.*

Esercizio 1.19. *Se* A*,* $B \in \mathbb{C}^{n,n}$*, allora* AB *e* BA *hanno gli stessi autovalori.*

(Suggerimento: se $\lambda = 0$ *è un autovalore di* AB*, allora* $0 = \det(AB) = \det(BA)$*, quindi* λ *è un autovalore di* BA*. Se* $\lambda \neq 0$ *è un autovalore di* AB*, esiste un vettore* $X_0 \neq O$ *per cui* $ABX_0 = \lambda X_0$*, allora* $BX_0 \neq 0$ *e* λ *è un autovalore per* BA*. Si noti che in questo caso le matrici possono essere non simili: ad esempio si consideri*

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix};$$

si ha $AB = O$ *ma* $BA \neq O$*.)*

Esercizio 1.20. *Stabilire per quali eventuali valori dei parametri a, b, c, $d \in \mathbb{C}$ sono simili le matrici*

$$B = \begin{pmatrix} a & 1 & b \\ 0 & i & c \\ 0 & 0 & d \end{pmatrix}, \quad B' = \begin{pmatrix} 1 & -1 & -1+i \\ i & -i & -i \\ -1 & 1 & 1+i \end{pmatrix}.$$

SOLUZIONE. Condizione necessaria (ma non sufficiente) affinché B e B' siano simili è che $P_B(\lambda) = P_{B'}(\lambda)$. Ora

$$\begin{aligned} P_B(\lambda) &= (a-\lambda)(i-\lambda)(d-\lambda) = 0 \quad \Rightarrow \quad \lambda = a, i, d, \\ P_{B'}(\lambda) &= \lambda(\lambda - i)(\lambda - 2 + i) = 0 \quad \Rightarrow \quad \lambda = 0, i, 2-i, \end{aligned}$$

quindi $a = 0$, $d = 2 - i$, oppure $a = 2 - i$, $d = 0$. In entrambi i casi B e B' hanno autovalori distinti, e quindi sono diagonalizzabili e simili alla matrice $I = \operatorname{diag}(0, i, 2-i)$, e dunque simili tra loro. □

CAPITOLO 2

FUNZIONI MATRICIALI

2.1 Polinomi matriciali

Sia $f\colon \mathbb{R} \to \mathbb{R}$ una funzione reale di una variabile reale. Si cercherà di dare significato all'espressione $f(X)$ quando X è una matrice.

Si consideri dapprima il caso in cui f sia un polinomio del tipo

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0.$$

Se $A \in \mathbb{R}^{m,m}$ si pone

$$p(A) \stackrel{\text{def}}{=} a_n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I \in \mathbb{R}^{m,m}.$$

Esempio 2.1. Sia $p(x) = x^3 - 7x + 1$ e $A = \left(\begin{smallmatrix} 1 & 2 \\ 0 & 2 \end{smallmatrix}\right)$. Allora $A^2 = \left(\begin{smallmatrix} 1 & 6 \\ 0 & 4 \end{smallmatrix}\right)$, $A^3 = \left(\begin{smallmatrix} 1 & 14 \\ 0 & 8 \end{smallmatrix}\right)$, e $p(A) = \left(\begin{smallmatrix} -5 & 0 \\ 0 & -5 \end{smallmatrix}\right) = -5I$. □

Si osservi che $p\colon \mathbb{R}^{m,m} \to \mathbb{R}^{m,m}$ *non è* un endomorfismo, mentre a $p(A)$ si può associare l'endomorfismo

$$f_{p(A)}\colon \mathbb{R}^m \to \mathbb{R}^m,\ \vec{x} \mapsto p(A)\vec{x}.$$

Se $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_m)$ è una matrice diagonale, allora $A^h = \operatorname{diag}(\lambda_1^h, \ldots, \lambda_m^h)$ e $p(A) = \operatorname{diag}(p(\lambda_1), \ldots, p(\lambda_m))$. Quindi, in questo caso, se λ_i sono radici di $p(x) = 0$, allora $p(A) = O$. Nel caso di matrici non diagonali, invece, le cose si complicano notevolmente.

Esempio 2.2. Sia

$$p(x) = x^2 = 0.$$

Nei numeri reali l'unica soluzione è $x = 0$. Se $X = \left(\begin{smallmatrix} x & y \\ z & t \end{smallmatrix}\right)$, l'equazione $p(X) = O$ ammette infinite soluzioni, poiché esistono divisori dello 0.

Ad esempio, si verifica che $X_1 = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, $X_2 = \left(\begin{smallmatrix} 0 & 0 \\ z & 0 \end{smallmatrix}\right)$ (per ogni $z \in \mathbb{R}$), $X_3 = \left(\begin{smallmatrix} 0 & y \\ 0 & 0 \end{smallmatrix}\right)$ (per ogni $y \in \mathbb{R}$), ed $X_4 = \left(\begin{smallmatrix} -h & -h^2/k \\ k & h \end{smallmatrix}\right)$ (per $h \in \mathbb{R}$ e $k \neq 0$) sono soluzioni, diverse da 0, dell'equazione $p(X) = O$

2.2 Polinomi annullatori

Data una matrice A, si ricercano polinomi a coefficienti in un campo $\mathbb{K}$ che si annullino in A. Tali polinomi si dicono *annullatori* di A.

Lemma 2.1. *Esistono polinomi annullatori di A.*

DIMOSTRAZIONE. Sia $A \in \mathbb{R}^{n,n}$, e si considerino le potenze $I, A, A^2, \ldots, A^{n^2}$. Queste n^2+1 matrici sono dipendenti poiché $\dim \mathbb{R}^{n,n} = n^2$; allora esiste una loro combinazione lineare a coefficienti non tutti nulli che dà la matrice nulla

$$a_{n^2}A^{n^2} + a_{n^2-1}A^{n^2-1} + \cdots + a_1 A + a_0 I = O.$$

Un polinomio annullatore di A (di grado $\leq n^2$)è quindi

$$p(x) = a_{n^2}x^{n^2} + a_{n^2-1}x^{n^2-1} + \cdots + a_1 x + a_0.$$

QED

Sia A una matrice simile ad una matrice diagonale D. Allora

$$A = B^{-1}DB \quad \Rightarrow \quad A^k = B^{-1}D^k B,$$

quindi se $p(x)$ è un polinomio si ha

$$p(A) = B^{-1}p(D)B.$$

Considerando il polinomio caratteristico $P_A(x)$ di A, risulta

$$P_A(A) = B^{-1}P_A(D)B = B^{-1}P_D(D)B = O,$$

poiché $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, dove λ_i sono gli autovalori di A (e di D).

Ciò che è sorprendente è che il risultato precedente vale per ogni matrice A, anche non diagonalizzabile. Per il teorema di Schur, basterebbe dimostrarlo per matrici triangolari.

Teorema 2.1 (Cayley–Hamilton). *Ogni matrice $A \in \mathbb{K}^{n,n}$ è radice del suo polinomio caratteristico.*

Saranno ora analizzate alcune conseguenze del teorema di Cayley–Hamilton. Innanzitutto, il polinomio caratteristico ha coefficienti determinati da A. Infatti

$$P_A(\lambda) = \det(A - \lambda I) = (-1)^n\lambda^n + (-1)^{n-1}b_{n-1}\lambda^{n-1} + \cdots + (-1)b_1\lambda + b_0,$$

dove

$$\begin{aligned}
b_{n-1} &= \operatorname{tr} A = a_{11} + \cdots + a_{nn},\\
b_{n-2} &= \sum_{1\le i<j\le n} \begin{vmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{vmatrix},\\
b_{n-3} &= \sum_{1\le i<j<k\le n} \begin{vmatrix} a_{ii} & a_{ij} & a_{ik} \\ a_{ji} & a_{jj} & a_{jk} \\ a_{ki} & a_{kj} & a_{kk} \end{vmatrix},\\
&\cdots\\
b_0 &= \det A.
\end{aligned}$$

Se $b_0 = \det A \neq 0$, poiché per il teorema di Cayley-Hamilton si ha $P_A(A) = O$, risulta

$$A^{-1}P_A(A) = (-1)^n A^{n-1} + (-1)^{n-1} b_{n-1} A^{n-2} + \cdots + (-1) b_1 I + b_0 A^{-1} = O,$$

da cui

$$A^{-1} = \frac{1}{b_0}\Big((-1)^{n-1} A^{n-1} + (-1)^{n-2} b_{n-1} A^{n-2} + \cdots + (-1) b_2 A + b_1 I\Big),$$

ossia è possibile trovare A^{-1} come polinomio di grado $n-1$ della matrice A.

Esercizio 2.1. *Trovare la matrice inversa di*

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

SOLUZIONE. A è invertibile perché $\det A = -2 \neq 0$. Si ha

$$P_A(\lambda) = \det(A - \lambda I) = -\lambda^3 + 4\lambda^2 - 5\lambda - 2,$$

dunque $A^{-1} = -1/2(A^2 - 4A + 5I)$. □

Esercizio 2.2. *Calcolare, tramite il teorema di Cayley–Hamilton, l'espressione* $A^4 - 11A^2 + 22A$ *(usando la matrice* A *dell'esercizio precedente).*

SOLUZIONE. Da $P_A(A) = O$ segue $A^3 = 4A^2 - 5A - 2I$. Allora

$$\begin{aligned}
A^4 - 11A^2 + 22A = AA^3 - 11A^2 + 22A = A(4A^2 - 5A - 2I) - 11A^2 + 22A = \\
= 4A^3 - 16A^2 + 20A = 4(4A^2 - 5A - 2I) - 16A^2 + 20A = -8I.
\end{aligned}$$

□

Osservazione 2.1. Dunque, per calcolare le potenze A^k, con $k > n$, basta conoscere il polinomio caratteristico di A e le potenze $A^2, \ldots A^{n-1}$ (senza calcolare gli autovalori).

2.3 Polinomio minimo

Il lemma 2.1 assicura l'esistenza di un polinomio annulatore di A di grado minore od uguale a n^2, mentre il teorema di Cayley–Hamilton abbassa il grado ad n. Il seguente teorema prova che il grado potrebbe ulteriormente abbassarsi.

Teorema 2.2. *Sia $A \in \mathbb{K}^{n,n}$. Allora esiste un solo polinomio $\mu_A(\lambda)$, detto* polinomio minimo di A, *tale che:*

1. *$\mu_A(\lambda)$ è monico (cioè, il coefficiente del termine di grado massimo è uguale ad 1);*
2. *$\mu_A(A) = O$;*
3. *se $p(\lambda)$ è un polinomio tale che $p(A) = O$ allora $\operatorname{gr} p(\lambda) \geq \operatorname{gr} \mu_A(\lambda)$.*

Corollario 2.1. *Sia $A \in \mathbb{K}^{n,n}$. Il polinomio $\mu_A(\lambda)$ è un fattore di $P_A(\lambda)$.*

Dimostrazione. Infatti, si ha $P_A(\lambda) = \mu_A(\lambda)s(\lambda) + r(\lambda)$, con $\operatorname{gr} r(\lambda) \leq \operatorname{gr} \mu_A(\lambda)$. Poiché risulta

$$O = P_A(A) = \mu_A(A)s(A) + r(A) = r(A),$$

segue che $r(\lambda) = 0$ per la minimalità di $\mu_A(\lambda)$. QED

Teorema 2.3. *I polinomi $P_A(\lambda)$ e $\mu_A(\lambda)$ hanno gli stessi zeri.*

Dimostrazione. Poiché $P_A(\lambda) = \mu_A(\lambda)s(\lambda)$, ne segue che gli zeri di $\mu_A(\lambda)$ devono essere zeri di $P_A(\lambda)$. Viceversa, ogni zero di $P_A(\lambda)$ è uno zero di $\mu_A(\lambda)$. Infatti, se $\tilde{\lambda}$ è uno zero di $P_A(\lambda)$, allora

$$AX = \tilde{\lambda}X \quad \Rightarrow \quad A^kX = \tilde{\lambda}^kX \quad \Rightarrow \quad \mu_A(A)X = \mu_A(\tilde{\lambda})X,$$

ma $\mu_A(A) = O$, dunque $\mu_A(\tilde{\lambda}) = 0$. QED

Ma $\mu_A(\lambda)$ è monico, quindi

$$\mu_A(\lambda) = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_r)^{n_r},$$

dove $\lambda_1, \ldots, \lambda_r$ sono gli r autovalori distinti di A e $\sum_{i=1}^{r} n_i \leq n$. Ovviamente, se la matrice A ha gli autovalori distinti, risulta $P_A(\lambda) = (-1)^n \mu_A(\lambda)$.

Infine, si prova il seguente teorema.

Teorema 2.4. *Una matrice $A \in \mathbb{K}^{n,n}$ è diagonalizzabile in $\mathbb{K}$ se e solo se le radici di $\mu_A(\lambda) = 0$ sono in $\mathbb{K}$ ed hanno molteplicità algebrica 1 (ossia sono n radici distinte).*

Osservazione 2.2. Da quanto sopra detto segue che anche l'espressione di A^{-1} come polinomio in A di grado minore di n non è unica. Infatti si possono ripetere per $\mu_A(\lambda)$ le stesse argomentazioni fatte per $P_A(\lambda)$ (anche per il calcolo delle potenze di A).

Esempio 2.3. Sia

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix};$$

$P_A(\lambda) = (\lambda - 1)^3(\lambda - 2)$. Si verifica facilmente che A annulla anche il polinomio $(\lambda - 1)^2(\lambda - 2)$, mentre non anulla il polinomio $(\lambda - 1)(\lambda - 2)$. Dunque

$$\mu_A(\lambda) = (\lambda - 1)^2(\lambda - 2).$$

Poiché la molteplicità algebrica di 1 rispetto a $\mu_A(\lambda)$ è 2, la matrice A non è diagonalizzabile. □

Il teorema 2.4 dà un criterio puramente algebrico per la diagonalizzabilità di una matrice A, ma bisogna trovare $\mu_A(\lambda)$. Esiste un algoritmo per il calcolo del polinomio minimo che non richiede il calcolo esplicito degli autovalori.

2.4 Funzioni di matrici definite mediante serie di potenze

Sia $f\colon U \subset \mathbb{C} \to \mathbb{C}$ una funzione analitica della variabile z

$$f(z) = \sum_{k=0}^{\infty} a_k z^k,$$

e sia ρ il raggio di convergenza della serie. Se $X = (x_{ij}) \in \mathbb{C}^{n,n}$ si ha la serie formale

$$\sum_{k=0}^{\infty} a_k X^k. \tag{2.4.1}$$

Se tale serie converge, il suo limite si indicherà $f(X)$. Ebbene

Teorema 2.5. *Condizione sufficiente per l'assoluta convergenza della serie (2.4.1) è che la matrice X sia nilpotente o che tutti gli autovalori di X siano in modulo minori di ρ.*

Naturalmente, se ρ è infinito, è ben definita $f(A)$ per ogni A.

Osservazione 2.3. Se $A = B^{-1}\Delta B$, allora $A^k = B^{-1}\Delta^k B$ e $f(A) = B^{-1}f(\Delta)B$. Quindi se $\Delta = \operatorname{diag}(d_1, \ldots, d_n)$, risulta $f(\Delta) = \operatorname{diag}(f(d_1), \ldots, f(d_n))$ ed $f(A)$ si calcola facilmente.

Definizione 2.1. La funzione $e^{\cdot}\colon \mathbb{C}^{n,n} \to \mathbb{C}^{n,n}$ definita da

$$e^A \stackrel{\text{def}}{=} \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

è detta *esponenziale di* A.

Si noti inoltre che A^k per $k \geq n$, grazie al teorema di Cayley–Hamilton, è esprimibile come combinazione lineare delle potenze A^h con $h < n$. Naturalmente, se A è una matrice nilpotente di ordine h, quindi $A^h = O$, allora

$$e^{tA} = \sum_{k=0}^{h-1} \frac{A^k}{k!} t^k.$$

Poiché la serie e^A è rapidamente convergente, il metodo della serie è quello che viene comunemente usato per il calcolo di e^A con il computer.

Analogamente, si definiscono il *seno* ed il *coseno* di una qualsiasi matrice A:

$$\sin A \stackrel{\text{def}}{=} \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} A^{2k+1},$$
$$\cos A \stackrel{\text{def}}{=} \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} A^{2k},$$

mentre, se tutti gli autovalori λ_h di A soddisfano alla limitazione $|\lambda_h| < 1$, si può definire

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k.$$

Esempio 2.4.

- Sia $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Allora $A^k = \operatorname{diag}(\lambda_1^k, \ldots, \lambda_n^k)$, quindi
$$e^A = \operatorname{diag}(e^{\lambda_1}, \ldots, e^{\lambda_n}).$$
- Sia $A = t\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. Allora $A^k = t^k \left(\begin{smallmatrix} 1 & k \\ 0 & 1 \end{smallmatrix}\right)$ e
$$\sin A = \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k+1}}{(2k+1)!} \begin{pmatrix} 1 & 2k+1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \sin t & t\cos t \\ 0 & \sin t \end{pmatrix}.$$
- Sia $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Allora $A^2 = I$ e si ha
$$e^{tA} = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}.$$

Si noti che, se

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_r \end{pmatrix}$$

allora, come si vede facilmente,

$$e^{tA} = \begin{pmatrix} e^{tA_1} & & & \\ & e^{tA_2} & & \\ & & \ddots & \\ & & & e^{tA_r} \end{pmatrix}.$$

2.5 Proprietà dell'esponenziale di matrici

In questo paragrafo saranno esposte alcune proprietà dell'esponenziale di matrici, molto utili per le applicazioni.

Proposizione 2.1. *Siano A, $B \in \mathbb{C}^{n,n}$. Allora:*

1. *se $AB = BA$, si ha $e^{A+B} = e^A e^B$;*
2. *$(e^A)^{-1} = e^{-A}$;*
3. *$\det(e^A) = e^{\operatorname{tr} A}$,*

quindi e^A è sempre una matrice invertibile.

Dimostrazione. Poiché $AB = BA$, vale la seguente identità:

$$(A+B)^k = \sum_{r=0}^{k} \binom{k}{r} A^{k-r} B^r,$$

da cui la conclusione (1) eseguendo il prodotto delle serie assolutamente convergenti e^A ed e^B. Il secondo punto è banale, osservando che $e^O = I$. Per il terzo punto, osservando che A è simile ad una matrice triangolare Δ, risulta

$$e^A = B^{-1} e^{\Delta} B \quad \text{e} \quad \det(e^A) = \det(e^{\Delta}).$$

Pertanto bisogna dimostrare la proprietà solo per il caso delle matrici triangolari. Questo compito è lasciato al lettore. QED

Esempio 2.5. Sia $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Essa è diagonalizzabile ed $A = BDB^{-1}$, dove

$$D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix},$$

Dunque

$$e^A = Be^D B^{-1} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} e & 0 \\ 0 & e^{-1} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} \cosh 1 & \sinh 1 \\ \sinh 1 & \cosh 1 \end{pmatrix}.$$

Infine $\det(e^A) = 1$, essendo $e^{\operatorname{tr} A} = e^0 = 1$.

Esercizio 2.3. *Provare che se $\vec{v}$ è un autovettore di A relativo all'autovalore λ, allora $\vec{v}$ è anche autovettore della matrice e^A relativo all'autovalore e^λ.*

2.6 Esponenziale di matrici ed equazioni differenziali

Dati $A \in \mathbb{C}^{n,n}$ e $C \in \mathbb{C}^{n,1}$, si consideri la funzione

$$X : \mathbb{R} \to \mathbb{C}^n, \quad X(t) = e^{At}C.$$

Se X è derivabile, si ha

$$\frac{dX}{dt} = Ae^{At}C = AX.$$

Pertanto, dato il sistema di equazioni differenziali ordinarie del primo ordine

$$\frac{dX}{dt} = AX, \tag{2.6.2}$$

un modo compatto per scrivere le soluzioni del sistema è

$$X(t) = e^{At}C,$$

con $C \in \mathbb{C}^n$ costante.

L'esponenziale di una matrice interviene in modo essenziale nello studio dei sistemi di equazioni differenziali ordinarie a coefficienti costanti. Questi sistemi schematizzano l'evoluzione di un sistema autonomo (ossia, non dipendente dal tempo in modo esplicito), ed i problemi di stabilità delle loro soluzioni hanno particolare importanza nel campo dell'Ingegneria, per esempio nell'Automatica [31].

Osservazione 2.4. Nei problemi di stabilità per le soluzioni dei sistemi di equazioni differenziali lineari a coefficienti costanti è importante avere un criterio che stabilisca sotto quali ipotesi sugli autovalori di A risulti

$$\lim_{t \to +\infty} e^{tA} = O.$$

Sussiste, in proposito, il seguente teorema: *se tutti gli autovalori di A hanno la parte reale negativa, allora è verificata la precedente relazione.* In tal caso la matrice A si dice *stabile.*

Esercizio 2.4. *Dare la soluzione generale, in forma compatta, del sistema di equazioni differenziali lineari*

$$\frac{dx}{dt} = x + 2y, \qquad \frac{dy}{dt} = 2x + y.$$

(Suggerimento: si tratta di calcolare e^{tA} dove $A = \left(\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right)$. Poiché A è diagonalizzabile, il calcolo non presenta difficoltà.

Posto $C^T = (c_1\ c_2)$ *si ha*

$$\begin{cases} x(t) = (c_1 + c_2)e^{3t} + (c_1 - c_2)e^{-t} = k_1 e^{3t} + k_2 e^{-t} \\ y(t) = (c_1 + c_2)e^{3t} + (c_2 - c_1)e^{-t} = k_1 e^{3t} - k_2 e^{-t} \end{cases}$$

dove k_1 *e* k_2 *sono costanti. Eliminando il parametro* t *si ottiene*

$$(x + y)(x - y)^3 = 16 k_1 k_2^3 = \textit{costante}.$$

Le orbite sono curve algebriche di ordine 4 *aventi per asintoto le rette* $x = \pm y$*. L'origine è un punto singolare di equilibrio instabile. Si provi a disegnare la precedente famiglia di curve usando un programma di calcolo numerico o simbolico.)*

CAPITOLO 3

LA FORMA CANONICA DI JORDAN

Sia $f\colon V \to V$ un endomorfismo di uno spazio vettoriale di dimensione finita n su un campo $\mathbb{K}$. Supponiamo che $\mathbb{K}$ contenga tutti gli zeri del polinomio caratteristico $P_f(\lambda)$ di f, cioè $P_f(\lambda)$ sia completamente riducibile in $\mathbb{K}$ (nel prodotto di fattori lineari).

Sappiamo già che non tutti gli endomorfismi ammettono una forma diagonale, che possiamo considerare la più semplice per rappresentare un endomorfismo rispetto ad una base.

Vogliamo trovare una forma canonica che più si avvicini a quella diagonale e che coincida con essa quando f è semplice. Ciò è realizzato dalla cosiddetta *forma canonica di Jordan*.

Infatti, sempre nell'ipotesi in cui $\mathbb{K}$ contenga tutti gli zeri del polinomio caratteristico di f, si dimostra che esiste una base opportuna $\mathcal{B}$ di V, detta *base di Jordan*, rispetto cui si ha

$$\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f) = D + N,$$

dove D è una matrice diagonale, N è una matrice nilpotente (cioè tale che esiste $k \in \mathbb{N}$ tale che $N^k = O$) ed inoltre $DN = ND$.

Dunque, ogni matrice quadrata nel campo complesso è simile ad una matrice del tipo $D + N$; mentre nel campo reale una matrice A è simile ad una matrice in forma canonica di Jordan se e solo se il polinomio caratteristico di A ha tutte le radici reali.

Naturalmente, se f è semplice allora $N = O$ e si ritorna a quanto noto. Più precisamente sia

$$P_f(\lambda) = (-1)^n(\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_r)^{m_r}$$

dove $\lambda_1, \ldots, \lambda_r$ sono gli zeri distinti del polinomio caratteristico ed m_i è la molteplicità algebrica di λ_i. Naturalmente $\sum_{i=1}^{r} m_i = n$.

Nel caso più generale in una base di Jordan si ha

$$J = \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} J_1 & O & \ldots & O \\ O & J_2 & \ldots & O \\ \ldots & \ldots & \ldots & \ldots \\ O & O & \ldots & J_r \end{pmatrix}, \quad J_i \stackrel{\text{def}}{=} J(\lambda_i) = \begin{pmatrix} J_{\lambda_i}^{(1)} & O & \ldots & O \\ O & J_{\lambda_i}^{(2)} & \ldots & O \\ \ldots & \ldots & \ldots & \ldots \\ O & O & \ldots & J_{\lambda_i}^{(h_i)} \end{pmatrix} \in \mathbb{K}^{m_i, m_i}$$

dove $h_i = \dim V(\lambda_i)$. Ogni *blocco elementare* $J_{\lambda_i}^{(j)}$, di ordine $k_i^{(j)}$ tale che $1 \leq k_i^{(j)} \leq m_i$ è dato da

$$J_{\lambda_i}^{(j)} \stackrel{\text{def}}{=} \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_i & 1 \\ 0 & 0 & \dots & 0 & \lambda_i \end{pmatrix} = \lambda_i I + N_j, \quad \text{con} \quad N_j \stackrel{\text{def}}{=} \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

La matrice N_j è nilpotente di indice $k_i^{(j)}$ ed ovviamente commuta con $\lambda_i I$. Ovviamente N_j ha come autovalore 0 con molteplicità $k_i^{(j)}$. La matrice N_j è detta *blocco nilpotente di ordine* $k_i^{(j)}$.

Si noti che i blocchi elementari di solito vengono scritti secondo l'ordine decrescente. Se f è semplice, allora $J_{\lambda_i}^{(j)} = (\lambda_i) \in \mathbb{K}^{1,1}$.

Da quanto precede si deduce facilmente come dovrà essere una base di Jordan. Sia λ_i un autovalore di f di molteplicità algebrica m_i. Una *catena di Jordan* per f di lunghezza $k_i^{(j)}$, con $1 \leq k_i^{(j)} \leq m_i$, è un insieme di vettori

$$\{\vec{v}_h^{(i)} \mid h = 1, \dots, k_i^{(j)}\} \quad \text{tale che} \quad f(\vec{v}_h^{(i)}) = \lambda_i \vec{v}_h^{(i)} + \vec{v}_{h-1}^{(i)}, \quad h = 1, \dots, k_i^{(j)},$$

dove è stato posto $\vec{v}_0 = \vec{0}$, quindi $f(\vec{v}_1^{(i)}) = \lambda_i \vec{v}_1^{(i)}$.

Una *base di Jordan* per f è una base di V composta dall'unione di un certo numero di catene di Jordan per f a due a due disgiunte. Una base di autovettori è una base di Jordan composta di catene di Jordan di lunghezza 1. È ovvio che il concetto di base di Jordan generalizza quello di base di autovettori.

Esempio 3.1. Si consideri la matrice

$$A = \begin{pmatrix} 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Si ha $\lambda_1 = 0$ con $m_1 = 4$, dunque $r = 1$. Dal calcolo si ottiene $\dim V(0) = 2$, dunque i blocchi elementari sono 2. Ovviamente A non è diagonalizzabile. Per stabilire la dimensione dei blocchi di Jordan (che potrebbero essere entrambi di ordine 2, o di ordine 1 e di ordine 3) bisogna studiare la forma canonica degli endomorfismi nilpotenti, essendo la nostra matrice nilpotente. Infatti

$$P_A(\lambda) = \lambda^4 \quad \Rightarrow \quad O = P_A(A) = A^4$$

Esercizio 3.1. *Provare che una matrice A è nilpotente se e solo se ha come unico autovalore $\lambda = 0$.*

3.1 Gli endomorfismi nilpotenti e la loro forma canonica

Sia $t\colon V \to V$ un endomorfismo nilpotente di indice k. Allora si ha

$$\operatorname{Ker} t \subset \operatorname{Ker} t^2 \subset \cdots \subset \operatorname{Ker} t^k = V$$

e tutte le inclusioni sono proprie se $t \neq 0$. Infatti, per qualunque endomorfismo (anche non nilpotente), si ha $\operatorname{Ker} t^h \subseteq \operatorname{Ker} t^{h+1}$; se per un certo $h < k$ risultasse $\operatorname{Ker} t^h = \operatorname{Ker} t^{h+1}$, allora si avrebbe anche $\cdots \operatorname{Ker} t^{h+2} \subset \operatorname{Ker} t^{h+1} = \operatorname{Ker} t^h$ (come è facile verificare), il che implicherebbe $V = \operatorname{Ker} t^h$, contraddicendo l'ipotesi che l'indice di t sia k.

Teorema 3.1. *Sia $t\colon V \to V$ un endomorfismo nilpotente di indice k di uno spazio vettoriale V di dimensione n su campo $\mathbb{K}$. Allora t si può rappresentare con una matrice nilpotente a blocchi del tipo standard N_j.*

Dimostrazione. La dimostrazione è di tipo costruttivo; si tratta di trovare una base rispetto alla quale t assume la forma richiesta.

Passo 1. Siano $\vec{v}_1, \ldots, \vec{v}_{p_1}$ vettori indipendenti in $V \setminus \operatorname{Ker} t^{k-1}$ che costituiscono una base del supplementare di $\operatorname{Ker} t^{k-1}$ in V. Allora $\{t(\vec{v}_1), \ldots, t(\vec{v}_{p_1})\} \subset \operatorname{Ker} t^{k-1}$ e si vede che una qualunque combinazione lineare a coefficienti non nulli di tali vettori non appartiene a $\operatorname{Ker} t^{k-2}$. Dunque

$$\{t(\vec{v}_1), \ldots, t(\vec{v}_{p_1})\} \subset \operatorname{Ker} t^{k-1} \setminus \operatorname{Ker} t^{k-2};$$

inoltre tali vettori sono indipendenti.

Passo 2. Sia $\{t(\vec{v}_1), \ldots, t(\vec{v}_{p_1}), \vec{v}_{p_1+1}, \ldots, \vec{v}_{p_2}\}$ una base di un supplementare di $\operatorname{Ker} t^{k-2}$ in $\operatorname{Ker} t^{k-1}$. Ragionando come sopra si prova che

$$\{t^2(\vec{v}_1), \ldots, t^2(\vec{v}_{p_1}), t(\vec{v}_{p_1+1}), \ldots, t(\vec{v}_{p_2})\} \subset \operatorname{Ker} t^{k-2} \setminus \operatorname{Ker} t^{k-3};$$

inoltre, tali vettori sono indipendenti.

Passo $k-1$. Si ottiene la base di $\operatorname{Ker} t$

$$\{t^{k-1}(\vec{v}_1), \ldots, t^{k-1}(\vec{v}_{p_1}), t^{k-2}(\vec{v}_{p_1+1}), \ldots, t(\vec{v}_{p_{k-1}}), \vec{v}_{p_{k-1}+1}, \ldots, \vec{v}_{p_k}\}.$$

Verrà ora dimostrato che un riordinamento di tutti i vettori sopra considerati fornisce la base di Jordan cercata. Posto $n_h = \dim(\operatorname{Ker} t^h)$, risulta

$$\begin{cases} p_1 = n_k - n_{k-1} = n - n_{k-1} \\ p_2 = n_{k-1} - n_{k-2} \geq p_1 \\ \cdots \\ p_{k-1} = n_2 - n_1 \geq p_{k-1} \\ p_k = n_1 \geq p_{k-1} \end{cases}$$

Si consideri la seguente base di V:

$$\begin{aligned}
p_1 \text{ righe di vettori: } &\begin{cases} t^{k-1}(\vec{v}_1), t^{k-2}(\vec{v}_1), \dots, t(\vec{v}_1), \vec{v}_1 \\ \cdots \\ t^{k-1}(\vec{v}_{p_1}), t^{k-2}(\vec{v}_{p_1}), \dots, t(\vec{v}_{p_1}), \vec{v}_{p_1} \end{cases} \\
p_2 - p_1 \text{ righe di vettori: } &\begin{cases} t^{k-2}(\vec{v}_{p_1+1}), t^{k-3}(\vec{v}_{p_1+1}), \dots, t(\vec{v}_{p_1+1}), \vec{v}_{p_1+1} \\ \cdots \\ t^{k-2}(\vec{v}_{p_2}), t^{k-3}(\vec{v}_{p_2}), \dots, t(\vec{v}_{p_2}), \vec{v}_{p_2} \end{cases} \\
&\cdots \\
p_k - p_{k-1} \text{ righe di vettori: } &\begin{cases} \vec{v}_{p_{k-1}+1} \\ \cdots \\ \vec{v}_{p_k} \end{cases}
\end{aligned}$$

Se $\vec{w}$ è un vettore di questa base, allora $t(\vec{w}) \in \operatorname{Ker} t^i \smallsetminus \operatorname{Ker} t^{i-1}$ per un solo valore di i tale che $1 \leq i \leq k$. Dunque, ogni riga costituita da i vettori produce un blocco nilpotente N_i nella matrice di t. Più in dettaglio, risultano

$$\begin{aligned}
&p_1 \text{ blocchi di ordine } k, \\
&p_2 - p_1 \text{ blocchi di ordine } k-1, \\
&\cdots \\
&p_k - p_{k-1} \text{ blocchi di ordine } 1.
\end{aligned}$$

Dunque, il totale dei blocchi nilpotenti è

$$p_1 + (p_2 - p_1) + \cdots + (p_k - p_{k-1}) = p_k = n_1 = \dim(\operatorname{Ker} t),$$

ed il numero dei vettori della base è

$$p_1 k + (p_2 - p_1)(k-1) + \cdots + (p_k - p_{k-1}) = p_1 + p_2 + \cdots + p_k = n_k = n = \dim V.$$

La matrice di t rispetto a questa base è

$$\begin{pmatrix} N_1 & & \\ & N_2 & \\ & & \\ & & N_{n_1} \end{pmatrix}$$

QED

Si noti che il numero dei blocchi nilpotenti di ordine $k-l$ è

$$p_{l+1} - p_l = n_{k-(l+1)+1} - n_{k-(l+1)} - n_{k-l+1} + n_{k-l} = 2n_{k-l} - n_{k-l-1} - n_{k-l+1};$$

ovvero, il numero dei blocchi nilpotenti di ordine $i = k-l$ è

$$2n_i - n_{i-1} - n_{i+1},$$

posto $n_i = n$ per $i \geq k$ ed $n_i = 0$ per $i \leq 0$.

Esempio 3.2. Si consideri $t\colon \mathbb{R}^5 \to \mathbb{R}^5$ rappresentato rispetto alla base canonica da

$$A = \begin{pmatrix} -5 & -5 & 0 & 2 & 0 \\ 5 & 5 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 5 & 5 & 0 & 0 & 0 \end{pmatrix}$$

Risulta $A^2 = O$, dunque A ha indice 2, quindi

$$\{\vec{0}\} \subset \operatorname{Ker} t \subset \operatorname{Ker} t^2 = V.$$

Essendo $\operatorname{Ker} t = \{(x_1, -x_1, x_3, 0, x_5) \mid x_1, x_3, x_5 \in \mathbb{R}\}$, si ha

$$n_1 = \dim \operatorname{Ker} t = 3 \qquad n_2 = \dim \operatorname{Ker} t^2 = 5$$
$$p_1 = n_2 - n_1 = 2 \qquad p_2 = n_1 = 3$$

Si considerino due vettori indipendenti in $V \setminus \operatorname{Ker} t$, ad esempio

$$\vec{v}_1 = (0, 0, 0, 1, 0), \qquad \vec{v}_2 = (1, 1, 0, 0, 0).$$

Si ha

$$t(\vec{v}_1) = (2, -2, 1, 0, 0), \qquad t(\vec{v}_2) = (-10, 10, 0, 0, 10).$$

Siccome $\{t(\vec{v}_1), t(\vec{v}_2)\} \subset \operatorname{Ker} t$, si complementa questo insieme ad una base di $\operatorname{Ker} t$ mediante, ad esempio, il vettore $\vec{v}_3 = (0, 0, 1, 0, 0)$. La base cercata è

$$\{t(\vec{v}_1), \vec{v}_1, t(\vec{v}_2), \vec{v}_2, \vec{v}_3\}.$$

Rispetto a questa base, t ha la matrice

$$J = \left(\begin{array}{cc:cc:c} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Ovviamente, J è simile ad A, e risulta $J^2 = O$.

I blocchi di ordine 2 sono $2n_2 - n_1 - n_3 = n_2 - n_1 = 2$, ed analogamente si vede che c'è un solo blocco di ordine 1.

3.2 Endomorfismi con un solo autovalore

Sia

$$P_f(\lambda) = (-1)^n(\lambda - \lambda_1)^n$$

il polinomio caratteristico di f. Dal teorema di Cayley–Hamilton si ha

$$P_f(A) = P_f(f) = 0 \quad \Rightarrow \quad (f - \lambda_1 \operatorname{Id})^n = 0.$$

Quindi $t = f - \lambda_1 \operatorname{Id}$ è nilpotente con indice di nilpotenza uguale a $k \leq n$. Sia $\mathcal{B}$ la base rispetto alla quale t è rappresentato dalla matrice N a blocchi nilpotenti, allora

$$\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f) = \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(t + \lambda_1 \operatorname{Id}) = N + \lambda_1 I,$$

che è la forma di Jordan di f.

3.3 Endomorfismi con più autovalori

Supponiamo che l'endomorfismo f abbia più autovalori distinti, ma, in ogni caso, che il suo polinomio caratteristico si decomponga in fattori lineari.

$$P_f(\lambda) = (-1)^n(\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_r)^{m_r}.$$

Definizione 3.1. Si dice *autospazio generalizzato* per f il sottospazio di V

$$E(\lambda_i) \stackrel{\text{def}}{=} \operatorname{Ker}(f - \lambda_i \operatorname{Id})^{m_i}.$$

Ovviamente, per $m_i = 1$ (ma non solo in questo caso) l'autospazio generalizzato è anche un autospazio nel senso usuale del termine.

Si vede facilmente che $E(\lambda_i)$ è un sottospazio invariante rispetto ad f ed inoltre

$$\dim E(\lambda_i) = m_i, \qquad V = E(\lambda_1) \oplus \cdots \oplus E(\lambda_r).$$

Posto

$$f_i \stackrel{\text{def}}{=} f|_{E(\lambda_i)} \colon E(\lambda_i) \to E(\lambda_i),$$

f_i risulta essere un endomorfismo con un solo autovalore λ_i. Quindi

$$f_i = t_i + \lambda_i \operatorname{Id} = t_i + s_i,$$

avendo posto $s_i \stackrel{\text{def}}{=} \lambda_i \operatorname{Id} \colon E(\lambda_i) \to E(\lambda_i)$ e $t_i \stackrel{\text{def}}{=} f_i - \lambda_i \operatorname{Id}$. L'endomorfismo s_i è un endomorfismo semplice con il solo autovalore λ_i, l'endomorfismo t_i è un endomorfismo nilpotente con indice $k_i \leq m_i$. È possibile estendere gli endomorfismi f_i, t_i, s_i ad endomorfismi dello spazio V definendo f_i, t_i, s_i come le applicazioni nulle sui sottospazi $E(\lambda_j)$, per $j \neq i$. Risulta

$$f = f_1 + \cdots + f_r = (t_1 + \cdots + t_r) + (s_1 + \cdots + s_r) = t + s,$$

avendo posto $t \stackrel{\text{def}}{=} t_1 + \cdots + t_r$ ed $s \stackrel{\text{def}}{=} s_1 + \cdots + s_r$. Si osservi che $ts = st$.

Si scelga ora come base $\mathcal{B}$ di V l'unione delle basi dei sottospazi $E(\lambda_i)$ rispetto a cui ogni f_i assume la forma canonica di Jordan. Rispetto alla base $\mathcal{B}$, f è rappresentato nella forma canonica di Jordan e la sua matrice è costituita da r blocchi disposti lungo la diagonale di ordine $m_1, \ldots, m_r$ relativi agli autovalori $\lambda_1, \ldots, \lambda_r$ ognuno dei quali contiene almeno un blocco elementare di ordine $k_1, \ldots, k_r$.

Osservazione 3.1.

1. Si ha $V(\lambda_i) \subseteq E(\lambda_i)$ per ogni i, ed f è semplice se e solo se $V(\lambda_i) = E(\lambda_i)$. In tal caso, infatti, la molteplicità algebrica coincide con quella geometrica. Inoltre, per ogni i risulta $f_i = \lambda_i \operatorname{Id}$, $t_i = 0$ e $k_i = 1$.

2. Se f è un endomorfismo di uno spazio vettoriale complesso rappresentato da una matrice $A \in \mathbb{R}^{n,n}$, la sua forma canonica di Jordan conterrà elementi complessi. Tuttavia esiste una rappresentazione mediante una matrice ad elementi reali detta *forma di Jordan reale*. Questa si ottiene considerando che gli autovalori complessi sono presenti a coppie coniugate, dunque se $\vec{z} = \vec{u} + i\vec{v}$ è un autovettore relativo all'autovalore $\lambda = a+ib$, il vettore $\bar{\vec{z}} = \vec{u} - i\vec{v}$ è un autovettore relativo all'autovalore $\bar{\lambda} = a - ib$. Pertanto, anziché utilizzare i vettori $\vec{z}$ e $\bar{\vec{z}}$, si utilizzeranno i vettori

$$\vec{u} = \operatorname{Re}(\vec{z}), \qquad \vec{v} = \operatorname{Im}(\vec{z}).$$

3. Sia $f \colon V \to V$ e sia $A = \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f)$ la matrice di f rispetto ad una base $\mathcal{B}$ di V. Se $\mathcal{B}'$ è una base di Jordan, allora

$$\mathcal{M}_{\mathcal{B}'}^{\mathcal{B}'}(f) = D + N \quad \Rightarrow \quad A = B^{-1}(D+N)B = B^{-1}DB + B^{-1}NB,$$

dove $B^{-1}DB$ è una matrice diagonalizzabile e $B^{-1}NB$ è una matrice nilpotente. □

Esempio 3.3. Ridurre a forma di Jordan la matrice

$$A = \begin{pmatrix} 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Si vede subito che $P_A(\lambda) = \lambda^4$ (cioè A è nilpotente) e che $\dim V(0) = 2$. Dunque, ci sono due blocchi elementari di Jordan. Poiché

$$n_1 = \dim(\operatorname{Ker} f) = 2, \qquad n_2 = \dim(\operatorname{Ker} f^2) = 4,$$

si hanno

$$\begin{aligned} 2n_1 - n_0 - n_2 = 0 &\quad \text{blocchi di ordine 1,} \\ 2n_2 - n_1 - n_3 = 2 &\quad \text{blocchi di ordine 2,} \\ 2n_3 - n_2 - n_4 = 0 &\quad \text{blocchi di ordine 3,} \\ 2n_4 - n_3 - n_5 = 0 &\quad \text{blocchi di ordine 4.} \end{aligned}$$

Quindi la forma canonica di Jordan è

$$J = \left(\begin{array}{cc:cc} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Una base di Jordan è

$$\vec{v}_1 = (0,1,0,0), \quad \vec{v}_2 = (0,0,0,1), \quad t(\vec{v}_1) = (1,0,3,0), \quad t(\vec{v}_2) = (2,0,-1,0).$$

Esempio 3.4. Trovare una base di Jordan per l'endomorfismo $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ rappresentato nella base canonica dalla matrice

$$A = \begin{pmatrix} 3 & -5 & -1 \\ 0 & 0 & 0 \\ 1 & -3 & 1 \end{pmatrix}.$$

Gli autovalori sono $\lambda_1 = 0$ con $m_1 = 1$ e $\lambda_2 = 2$ con $m_2 = 2$. Si ha

$$E(0) = V(0) = \operatorname{Ker} f = \mathcal{L}((2,1,1)).$$

Inoltre, essendo $E(2) = \operatorname{Ker}(f - 2\operatorname{Id})^2$, e poiché

$$(A - 2I)^2 = \begin{pmatrix} 0 & 8 & 0 \\ 0 & 4 & 0 \\ 0 & 4 & 0 \end{pmatrix},$$

risulta $E(2) = \mathcal{L}((1,0,0),(0,0,1))$, e $V = E(0) \oplus E(2)$. Inoltre $f_1\colon E(0) \to E(0)$ è l'endomorfismo nullo ed $f_2\colon E(2) \to E(2)$ si scrive come $f_2 = t_2 + 2\operatorname{Id}$, ove t_2 è nilpotente di indice $k_2 \le m_2 = 2$. Indicata con $\mathcal{B}_2 = \{(1,0,0),(0,0,1)\}$ la base di $E(2)$ trovata, si ha

$$f_2((1,0,0)) = (3,0,1), \quad f_2((0,0,1)) = (-1,0,1),$$

dunque

$$\mathcal{M}_{\mathcal{B}_2}^{\mathcal{B}_2}(f_2) = A_2 = \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix}, \qquad \mathcal{M}_{\mathcal{B}_2}^{\mathcal{B}_2}(t_2) = T_2 = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}.$$

Poiché $\operatorname{Ker} t_2 = \mathcal{L}((1,1))$ e $\operatorname{Ker} t_2^2 = \mathbb{R}^2$, t_2 ha indice 2. Si prenda $\vec{v} = (1,0) \in \operatorname{Ker} t_2^2 \setminus \operatorname{Ker} t_2$. Una base $\mathcal{B}_2'$ di $E(2)$ è data da $\{\vec{v}, t_2(\vec{v}) = (1,1)\}$. Rispetto questa base si ha

$$\mathcal{M}_{\mathcal{B}_2'}^{\mathcal{B}_2'}(t_2) = T_2' = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \mathcal{M}_{\mathcal{B}_2'}^{\mathcal{B}_2'}(f_2) = A_2' = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}.$$

In conclusione,

$$A \sim A' = \left(\begin{array}{cc:c} 2 & 1 & 0 \\ 0 & 2 & 0 \\ \hdashline 0 & 0 & 0 \end{array}\right) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} = N + D.$$

Quali sono le matrici P tali che $A' = P^{-1}AP$? Basta scrivere nella base canonica le coordinate dei vettori della base $\mathcal{B}_2'$. Risulta $\vec{v} = (1,0,0)$, e $t_2(\vec{v}) = (1,0,1)$. Insieme al vettore $(2,1,1) \in E(0)$ questi vettori costituiscono una base di Jordan per A. La matrice del cambiamento di base (rispetto alla base canonica) è

$$P = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

3.4 Casi particolari

Caso $n = 2$. Sia $A \in \mathbb{K}^{2,2}$, con $P_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)$. Allora si distinguono i seguenti casi.

1. $\lambda_1 \neq \lambda_2$. Dunque A è diagonalizzabile.
2. $\lambda_1 = \lambda_2$. Si ha $1 \leq h_1 \leq 2$. Nel caso in cui $h_1 = 2$, A è diagonalizzabile. Si consideri il caso $h_1 = 1$. Risulta $t_1 \neq 0$, $\dim \operatorname{Ker} t_1 = 1$, $\dim \operatorname{Ker} t_1^2 = 2$. Presi $\vec{v}_1 \in V \setminus \operatorname{Ker} t_1$ e $t_1(\vec{v}_1)$, rispetto tale base la matrice di t_1 è $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, dunque
$$A \sim \begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}.$$

Caso $n = 3$. Sia $A \in \mathbb{K}^{3,3}$, con $P_A(\lambda) = -(\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3)$. Allora si distinguono i seguenti casi.

1. λ_1, λ_2, λ_3 distinti; in questo caso, A è diagonalizzabile.
2. $\lambda_1 = \lambda_2 \neq \lambda_3$. Si ha $1 \leq h_1 \leq 2$ e $h_3 = 1$. Nel caso in cui $h_1 = 2$, A è diagonalizzabile. Nel caso $h_1 = 1$, procedendo come al punto precedente si arriva alla forma di Jordan
$$A \sim \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

Si confronti questo caso con l'esempio 3.4.

3. $\lambda_1 = \lambda_2 = \lambda_3 = \lambda$. Se $h_1 = 3$ allora $A = \lambda I$. Sia $h_1 = 2$. Procedendo come sopra risulta
$$A \sim \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}.$$
Se $h_1 = 1$ risulta $\dim \operatorname{Ker} t_1 = 1$, $\dim \operatorname{Ker} t_1^2 = 2$, $\dim \operatorname{Ker} t_1^3 = 3$. Preso $\vec{v} \in V \setminus \operatorname{Ker} t_1^2$, l'insieme $\mathcal{B} = \{t_1^2(\vec{v}), t_1(\vec{v}), \vec{v}\}$ è una base di Jordan rispetto alla quale si ha la forma canonica
$$A \sim \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

3.5 Riepilogo

Sia $A \in \mathbb{C}^{n,n}$. I passi necessari per ottenere la forma canonica di Jordan sono i seguenti.

1. *Trovare lo spettro di A, ossia l'insieme degli autovalori di A.* Si denotino con $\lambda_1, \ldots, \lambda_r$ gli autovalori distinti di A, con $m_1, \ldots, m_r$ ed $h_1, \ldots, h_r$ le rispettive molteplicità algebriche e geometriche.

2. La forma canonica di A è la matrice J che ha sulla diagonale principale i blocchi $J_1, \ldots, J_r$, dove $J_i \in \mathbb{C}^{m_i, m_i}$, e ciascun J_i è costituito dai blocchi diagonali $J_i^{(1)}, \ldots, J_i^{(h_i)}$.

3. È poi necessario trovare l'ordine $k_i^{(j)}$ dei blocchi $J_i^{(j)}$. Se $1 \leq k_i^{(j)} \leq m_i$ allora i blocchi di ordine $k = k_i^{(j)}$ sono
$$2n_k - n_{k-1} - n_{k+1},$$
dove $n_k = \dim \operatorname{Ker}(A - \lambda_i I)^k$. Oppure, posto $r_k \stackrel{\text{def}}{=} \operatorname{rg}(A - \lambda_i I)^k$ il numero dei blocchi di ordine k è
$$r_{k+1} + r_{k-1} - 2r_k;$$
per ottenere questa relazione, si usi l'identità $n_k + r_k = n$.

Da quanto detto, segue un importante risultato. Si denoti con $\operatorname{spec}(A)$ l'insieme degli autovalori di una matrice quadrata A.

Teorema 3.2. *Due matrici A, $B \in \mathbb{C}^{n,n}$ sono simili se e solo se*

1. $\operatorname{spec}(A) = \operatorname{spec}(B)$;

2. $\operatorname{rg}(A - \lambda I)^k = \operatorname{rg}(B - \lambda I)^k$ *per ogni* $k \in \mathbb{N}$ *e per ogni* $\lambda \in \operatorname{spec}(A) = \operatorname{spec}(B)$

3.6 Esponenziale di matrici e forma canonica di Jordan

In questo paragrafo per il calcolo dell'esponenziale delle matrici sarà sfruttato il seguente lemma di immediata dimostrazione.

Lemma 3.1. *Siano A, $A' \in \mathbb{K}^{n,n}$ due matrici simili, con $A' = B^{-1}AB$. Allora*

$$e^{A'} = B^{-1}e^{A}B.$$

Ora, si supponga che la matrice A ammetta la forma canonica di Jordan $J = D + N$ (dunque A e J sono simili). Poiché $DN = ND$, risulta

$$e^{J} = e^{D}e^{N}.$$

Pertanto,

$$D = \begin{pmatrix} \lambda_1 & & \\ & \vdots & \\ & & \lambda_n \end{pmatrix} \quad \Rightarrow \quad e^{D} = \begin{pmatrix} e^{\lambda_1} & & \\ & \vdots & \\ & & e^{\lambda_n} \end{pmatrix}.$$

Inoltre, poiché N è nilpotente di indice k, $N^k = O$, dunque

$$e^{N} = I + N + \frac{N^2}{2!} + \cdots + \frac{N^{k-1}}{(k-1)!}.$$

Se A è diagonalizzabile, allora $N = O$, dunque $k = 1$ e $e^N = I$. Se A non è diagonalizzabile, si ha $k > 1$.

Esempio 3.5. Sia $N = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. Essendo $N^2 = O$, si ha

$$e^{N} = I + N = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad e^{D+N} = \begin{pmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^{\lambda_1} & e^{\lambda_1} \\ 0 & e^{\lambda_2} \end{pmatrix}.$$

Esempio 3.6. Sia

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

Risulta

$$N^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad N^3 = O.$$

Dunque

$$e^N = I + N + \frac{N^2}{2} = \begin{pmatrix} 1 & 1 & 1/2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix},$$

$$e^{D+N} = \begin{pmatrix} e^{\lambda_1} & 0 & 0 \\ 0 & e^{\lambda_2} & 0 \\ 0 & 0 & e^{\lambda_3} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1/2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} e^{\lambda_1} & e^{\lambda_2} & (e^{\lambda_3})/2 \\ 0 & e^{\lambda_2} & e^{\lambda_2} \\ 0 & 0 & e^{\lambda_3} \end{pmatrix}.$$

Osservazione 3.2. In generale, se

$$N = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

è tale che $N^k = O$, allora si ha

$$e^N = \begin{pmatrix} 1 & \frac{1}{1!} & \frac{1}{2!} & \dots & \frac{1}{(k-2)!} & \frac{1}{(k-1)!} \\ 0 & 1 & \frac{1}{1!} & \dots & \frac{1}{(k-3)!} & \frac{1}{(k-2)!} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

Esempio 3.7. Trovare l'esponenziale della matrice

$$A = \begin{pmatrix} 3 & -5 & -1 \\ 0 & 0 & 0 \\ 1 & -3 & 1 \end{pmatrix}$$

La forma canonica di Jordan ed altre quantità sono state ricavate nell'esempio 3.4. Si ha $e^A = Pe^J P^{-1}$, e

$$e^{A'} = e^{D+N} = \begin{pmatrix} e^2 & 0 & 0 \\ 0 & e^2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} e^2 & e^2 & 0 \\ 0 & e^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

dunque

$$e^A = \begin{pmatrix} 2e^2 & 2 - 3e^2 & -e^2 \\ 0 & 1 & 0 \\ e^2 & 1 - 2e^2 & 0 \end{pmatrix}$$

3.7 Esercizi di riepilogo

Esercizio 3.2. *Si consideri l'endomorfismo* $f\colon \mathbb{R}^4 \to \mathbb{R}^4$ *rappresentato rispetto alla base canonica dalla matrice*

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$

1. *Provare che* f *è nilpotente (si consiglia di usare il teorema di Cayley–Hamilton).*

2. *Dare la forma canonica di Jordan di* A*.*

3. *Provare che* $\vec{v}_{4-h} = f^h(\vec{e}_1)$ *è una base di Jordan di* A*.*

Soluzione.

1. Si vede facilmente che $P_A(\lambda) = \lambda^4$, dunque per il teorema di Hamilton–Cayley si ha $P_A(A) = -A^4 = O$, cioè A è nilpotente.

2. Troviamo $V(0) = \operatorname{Ker} f = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_2 = 0, x_4 = 0, x_1 + x_2 = 0\} = \mathcal{L}((0, 0, 1, 0))$. Poiché $\dim V(0) = 1$, si ha un solo blocco di Jordan di ordine 4.

3. Ponendo $h = 0$, 1, 2, 3 si ha

$$\vec{v}_4 = \vec{e}_1, \quad \vec{v}_3 = \vec{e}_4, \quad \vec{v}_2 = \vec{e}_1 - \vec{e}_2, \quad \vec{v}_1 = -\vec{e}_3.$$

Quindi, tenendo conto della matrice A, si ha

$$f(\vec{v}_1) = \vec{0}, \quad f(\vec{v}_2) = \vec{v}_1, \quad f(\vec{v}_3) = \vec{v}_2, \quad f(\vec{v}_4) = \vec{v}_3.$$

La matrice associata ad f rispetto alla base $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$ è proprio la forma di Jordan di A.

□

Esercizio 3.3. *Calcolare* e^B *dove*

$$B = \begin{pmatrix} 1 & -3 & 6 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

utilizzando direttamente la definizione.

Soluzione. Poiché $B^2 = B$ si ha

$$e^B = \sum_{k=0}^{\infty} \frac{B^k}{k!} = I + \sum_{k=1}^{\infty} \frac{B^k}{k!} = I + (\sum_{k=1}^{\infty} \frac{1}{k!})B = I + (e-1)B.$$

□

Esercizio 3.4. *Si consideri l'endomorfismo $f\colon \mathbb{R}^5 \to \mathbb{R}^5$ rappresentato (rispetto alla base canonica) dalla matrice*

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

1. *Provare che f è nilpotente.*
2. *Dare la forma canonica di Jordan.*

SOLUZIONE. 1. Ricordiamo che f è nilpotente se e solo se ha come unico autovalore $\lambda = 0$. Si vede facilmente che $P_A(\lambda) = -\lambda^5 = 0$.

2. Troviamo $V(0) = \mathcal{L}((1,1,0,0,0))$, quindi $\dim V(0) = 1$. Si ha un solo blocco di Jordan di ordine 5.

□

Esercizio 3.5. *Sia data la matrice*

$$A = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}.$$

Provare che $e^A = (e-1)A + I$.

SOLUZIONE. Come è noto, $e^{-A} = (e^A)^{-1}$. Ora $A^2 = A$, quindi

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} = I + (e-1)A.$$

Essendo $\det(e^A) = e$, segue

$$e^{-A} = \begin{pmatrix} e^{-1} & 1 - e^{-1} \\ 0 & 1 \end{pmatrix}.$$

□

Esercizio 3.6. *Sia $A \in \mathbb{R}^{n,n}$ una matrice nilpotente. Provare che*

1. *$I - A$ è invertibile.*
2. *$(I - A)^{-1} = I + A + \cdots + A^n$.*

SOLUZIONE. Se A è nilpotente, esiste un indice $q \geq n$ tale che $A^q = 0$ (e $A^h = 0$ per $h > q$). Quindi

$$I = I - A^q = (I + A + A^2 + \cdots + A^{q-1})(I - A) \quad \Rightarrow$$
$$(I - A)^{-1} = I + A + \cdots + A^{q-1} + A^q + \cdots + A^n.$$

□

Esercizio 3.7. *Si consideri l'endomorfismo $f\colon \mathbb{R}^5 \to \mathbb{R}^5$ rappresentato (rispetto alla base canonica) dalla matrice*

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

1. *Provare che f è nilpotente.*
2. *Dare la forma canonica di Jordan.*

SOLUZIONE.

1. Ricordiamo che f è nilpotente se e solo se ha come unico autovalore $\lambda = 0$. Si vede facilmente che $P_A(\lambda) = -\lambda^5 = 0$.
2. Troviamo $V(0) = \mathcal{L}((1,0,0,0,0))$, quindi $\dim V(0) = 1$. Si ha un solo blocco di Jordan di ordine 5.

□

Esercizio 3.8. *Si consideri la matrice*

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

1. *Provare che A è normale e rappresenta una matrice di permutazione.*
2. *Calcolare* $\sin\left(\frac{\pi}{2}A^4\right)$.

SOLUZIONE. 1. Ricordiamo che A è normale se e solo se $AA^* = A^*A$. Eseguendo i calcoli si conclude nel senso voluto. Inoltre A è una amtrice di permutazione poiché si ottiene permutando le colonne della matrice identica.

2. Gli autovalori di A sono $\lambda = \pm 1, \pm i$. Quindi A è simile alla matrice diagonale $D = \operatorname{diag}(1,-1,i,-i)$. Ora $A = B^{-1}DB$ implica $A^k = B^{-1}D^kB$, quindi, essendo $D^4 = I$,

$$\sin\left(\frac{\pi}{2}A^4\right) = \sin\left(\frac{\pi}{2}B^{-1}D^4B\right) = \sin\left(\frac{\pi}{2}I\right) = \sin\frac{\pi}{2}I = I.$$

□

Esercizio 3.9. *Si consideri per ogni $\alpha \in \mathbb{C}$ la matrice*

$$A = \begin{pmatrix} 1+\alpha & -\alpha \\ \alpha & 1-\alpha \end{pmatrix}.$$

1. *Per ogni $k \in \mathbb{N}$ determinare gli interi p e q tali che $A^k = pI + qA$.*

2. Calcolare e^A.

3. Calcolare la forma canonica di Jordan.

SOLUZIONE. 1. Si vede facilmente che

$$A^k = \begin{pmatrix} 1+k\alpha & -k\alpha \\ k\alpha & 1-k\alpha \end{pmatrix} = pI + qA \quad \Rightarrow \quad \begin{cases} p = 1-k, \\ q = k. \end{cases}$$

2. Tenendo conto che $A^k = (1-k)I + kA$ si ha

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} = \left[\sum_{k=0}^{\infty} \frac{1}{k!}\right] I - \left[\sum_{k=0}^{\infty} \frac{k}{k!}\right] I + \left[\sum_{k=0}^{\infty} \frac{k}{k!}\right] A.$$

Ora,

$$\sum_{k=0}^{\infty} \frac{1}{k!} = e; \quad \sum_{k=0}^{\infty} \frac{k}{k!} = \sum_{k=1}^{\infty} \frac{1}{(k-1)!} = e,$$

quindi

$$e^A = eA.$$

3. Gli autovalori di A sono $\lambda = 1$ con molteplicità algebrica 2. Determiniamo ora l'autospazio $V(1)$, dato dalle soluzioni dell'equazione

$$\alpha x - \alpha y = 0.$$

Ora, se $\alpha = 0$ si ha $V(1) = \mathbb{C}^2$ e quindi $A = I$ è diagonale; se $\alpha \neq 0$ si ha $V(1) = \{(x, x) \mid x \in \mathbb{C}\}$, quindi $\dim V(1) = 1$ ed A non è diagonalizzabile. In questo caso la forma di Jordan è

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

Una base di Jordan è data da $\vec{e'}_1 = (\alpha, \alpha)$, $\vec{e'}_2 = (1, 0)$, quindi la matrice del cambiamento di base è

$$B = \begin{pmatrix} \alpha & 1 \\ \alpha & 0 \end{pmatrix} \quad \Rightarrow \quad B^{-1} = \begin{pmatrix} 0 & 1/\alpha \\ 1 & -1 \end{pmatrix}.$$

Si verifica facilmente che

$$A = BJB^{-1}.$$

□

CAPITOLO 4

LA GEOMETRIA PER LA GRAFICA AL COMPUTER

4.1 Le trasformazioni geometriche 2D

Sia α un piano euclideo, e $\Phi\colon \alpha \to \alpha$ un'applicazione biunivoca. Allora Φ si dice *trasformazione* del piano α in sé. Se $G \subset \alpha$, allora $\Phi(G) \subset \alpha$ è detto il trasformato di G tramite Φ. Se Φ e Φ^{-1} sono continue, le figure G e $\Phi(G)$ sono dette *omeomorfe* o *topologicamente equivalenti*.

Nella grafica computerizzata sono particolarmente importanti le trasformazioni affini e le loro composizioni. Si ricordi che Φ è un'applicazione affine se $\Phi = \tau \circ f$, dove f è un'applicazione lineare e τ una traslazione. Se $P = (x\ y)^T$, dove (x, y) sono le coordinate di $P \in \alpha$ (rispetto ad un sistema di riferimento $(O, \mathcal{B})$ scelto) e $P' = \Phi(P)$, allora, com'è noto,

$$P' = AP + B, \qquad A \stackrel{\text{def}}{=} \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(\Phi) = \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f), \quad \tau(Q) = Q + B. \tag{4.1.1}$$

Se $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ con $\det(A) \neq 0$ e $B = (x_0, y_0)^T$ esplicitamente si ha

$$\begin{cases} x' = ax + by + x_0 \\ y' = cx + dy + y_0 \end{cases}$$

Osservazione 4.1. Se P e P' sono rappresentati dai vettori riga (e non vettori colonna), come spesso accade nella grafica al computer, l'applicazione (4.1.1) va scritta

$$P' = P^T A^T + B.$$

Tuttavia, noi non useremo questa notazione. □

Siano P_1 e P_2 due punti del piano e

$$P = (1 - t)P_1 + tP_2, \qquad t \in [0, 1],$$

cioè $P \in P_1P_2$ (questo risulta da $P - P_1 = t(P_2 - P_1)$). Allora, considerando la funzione Φ introdotta sopra, risulta

$$P' = (1 - t)AP_1 + tAP_2 + B = (1 - t)(P_1' - B) + t(P_2' - B) + B = (1 - t)P_1' + tP_2',$$

cioè $P' \in P'_1P'_2$, dove $\Phi(P_1) = P'_1$ e $\Phi(P_2) = P'_2$.

Questa proprietà è di grande importanza nelle elaborazioni grafiche, perchè consente di calcolare l'immagine trasformata di un poligono senza trasformare tutti i suoi infiniti punti, ma solo i vertici. Tenendo conto che ogni figura è approssimabile con un'opportuna figura poligonale, il vantaggio è ovvio.

Si osservi, inoltre, che le trasformazioni affini conservano il parallelismo, cioè se r ed r' sono rette parallele e Φ è una trasformazione affine, allora la retta $\Phi(r)$ è parallela alla retta $\Phi(r')$.

Nella grafica al computer una trasformazione geometrica modifica la locazione di un pixel dentro un'immagine; più precisamente, se $P(x, y)$ è un pixel dell'immagine di partenza (*input*), allora $\Phi(P(x, y)) = P'(x', y')$ è un pixel dell'immagine di arrivo (*output*). Poiché un pixel non è un punto, ma un quadratino, e le coordinate in generale non sono numeri interi, sono poi necessarie delle interpolazioni per assegnare il valore del pixel (ad esempio il livello di grigio).

Le trasformazioni geometriche sono necessarie in diverse applicazioni, sia per correggere eventuali distorsioni geometriche introdotte durante l'acquisizione dell'immagine (per esempio immagini acquisite mentre gli oggetti oppure i sensori sono in movimento, acquisizioni da satellite), oppure per introdurre effetti geometrici visivi.

Naturalmente, si possono considerare anche trasformazioni non lineari (ad esempio *warping transformations*); il tipo di trasformazione da utilizzare dipende dal contesto applicativo. Per esempio, nella ricostruzione (*restoration*) di un'immagine, per eliminare semplici aberrazioni ottiche, introdotte in fase di acquisizione, si utilizzano trasformazioni spaziali lineari; invece, si usano trasformazioni non lineari nel caso di registrazione di più immagini acquisite in tempo diversi in condizione di instabilità non lineare dei sensori.

Le trasformazioni geometriche più usate sono composizioni (anche iterate) delle seguenti trasformazioni affini elementari: *traslazioni, trasformazioni di scala, rotazioni, ribaltamenti*. Le matrici corrispondenti a queste trasformazioni vengono dette *matrici di trasformazione*.

1. Traslazione. $P' = P + B$. Qui $A = I$.

2. Trasformazione di scala. $P' = KP$, dove

$$K = \begin{pmatrix} k_x & 0 \\ 0 & k_y \end{pmatrix},$$

dove k_x (risp. k_y) è il fattore di scala lungo l'asse x (risp. y). Per direzioni diverse dalle direzioni degli assi coordinati, il fattore di scala varia. Se $k_x = k_y$ si ha una *similitudine*, usualmente detta *ingrandimento* se $k_x > 1$ o *rimpicciolimento* se $0 < k_x < 1$. Si osservi che

$$\begin{pmatrix} k_x & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & k_y \end{pmatrix} = \begin{pmatrix} k_x & 0 \\ 0 & k_y \end{pmatrix}$$

Se G è una circonferenza di centro O e raggio 1, allora $\Phi(G)$ è un'ellisse di semiassi $a = k_x$ e $b = k_y$. Naturalmente, se $k_x = k_y$, si avrà una circonferenza di raggio r.

3. Rotazione [16]. $P' = R(\varphi)P$, dove

$$R(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

Si tratta di una rotazione di angolo φ intorno all'origine O. Questa trasformazione conserva l'orientazione: le matrici che intervengono hanno determinante positivo. Si rammenti che $R(\varphi) \cdot R(\psi) = R(\varphi + \psi)$, e che $R(-\varphi) = R(\varphi)^T = R(\varphi)^{-1}$.

4. Ribaltamento [16]. $P' = S(\varphi)P$, dove

$$S(\varphi) = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}.$$

Si tratta di un ribaltamento intorno alla retta di equazione $y = \mathrm{tg}(\varphi/2)$. Questa trasformazione non conserva l'orientazione: le matrici che intervengono hanno determinante negativo. Si rammenti che $S(\varphi) \cdot S(\psi) = R(\varphi - \psi)$.

Osservazione 4.2. Le precedenti trasformazioni si estendono ad un piano proiettivo come segue. Si rappresenti P con le coordinate proiettive non omogenee $(x, y, 1)$ al posto di (x, y) (si veda l'appendice A.2). Allora

$$P' = TP = \begin{pmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} x + x_0 \\ y + y_0 \\ 1 \end{pmatrix} = \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix}$$

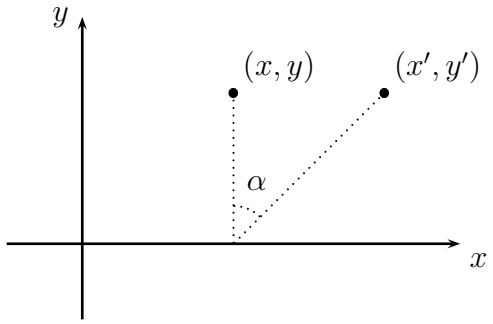
$$P' = KP = \begin{pmatrix} k_x & 0 & 0 \\ 0 & k_y & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} k_x x \\ k_y y \\ 1 \end{pmatrix} = \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix}$$

$$P' = R(\varphi)P = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} x\cos\varphi - y\sin\varphi \\ x\sin\varphi + y\cos\varphi \\ 1 \end{pmatrix} = \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} \quad \square$$

Due trasformazioni usate frequentemente, che sono composizioni delle precedenti trasformazioni, sono le seguenti.

1. Asimmetria (*skew*). $P' = AP$ dove

$$A = \begin{pmatrix} 1 & \operatorname{tg}\alpha \\ 0 & 1 \end{pmatrix}:$$



2. Forbice (*shear*). $P' = AP$, dove

$$A = \begin{pmatrix} 1 & a \\ b & 1 \end{pmatrix}.$$

Osservazione 4.3. Con una scelta opportuna di riferimenti ortonormali, un'affinità si può riportare sempre alla forma

$$x' = ax, \qquad y' = by.$$

In tal caso, se G è una figura e $\mathcal{A}(G)$ indica l'area di G, allora vale

$$\mathcal{A}(\Phi(G)) = ab\mathcal{A}(G).$$

Naturalmente, se $ab = 1$, le due figure sono equivalenti. Ad esempio, se G è il cerchio unitario, $\Phi(G)$ è un'ellisse e $\mathcal{A}(\Phi(G)) = \pi\, ab$. □

Osservazione 4.4. Se la matrice A è ortogonale, si ha un'isometria, che si può sempre rappresentare nella forma canonica $x' = x$, $y' = y$, cioè esiste un movimento che porta l'una nell'altra. Si ricordi che la composizione di movimenti non è commutativa. □

Come si diceva prima, ogni movimento si può ottenere componendo le trasformazioni elementari di cui sopra. Ad esempio, la rotazione di un angolo φ di un oggetto intorno ad un punto $(x_0, y_0) \neq (0, 0)$ si può considerare come la composizione delle seguenti trasformazioni elementari:

1. traslazione tale che $(x_0, y_0) \mapsto (0, 0)$, rappresentata da

$$T = \begin{pmatrix} 1 & 0 & -x_0 \\ 0 & 1 & -y_0 \\ 0 & 0 & 1 \end{pmatrix};$$

2. rotazione dell'angolo φ assegnato intorno all'origine, rappresentata dalla matrice $R(\varphi)$ (osservazione 4.2);

3. traslazione inversa $(0,0) \mapsto (x_0, y_0)$, rappresentata da

$$T^{-1} = \begin{pmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{pmatrix}.$$

La rotazione cercata è rappresentata dalla matrice $A = T^{-1}RT$.

Nei problemi concreti, date le figure G e $G' \stackrel{\text{def}}{=} \Phi(G)$, non è nota l'espressione analitica di Φ, che va trovata tenendo conto del modo in cui è acquisita l'immagine. Per esempio, durante l'acquisizione di un'immagine da satellite, pur conoscendo la cinematica della piattaforma spaziale, il suo assetto può variare leggermente introducendo distorsioni imprevedibili dell'immagine acquisita.

Un modo per risolvere il problema consiste nell'individuare nella scena alcuni *punti di controllo* (detti anche *punti fiduciali*) che siano ben visibili in G e G'. Se supponiamo che Φ sia affine sono sufficienti 3 punti di controllo per calcolare i 6 coefficienti che individuano la trasformazione (cioè, i coefficienti di A e di B).

Nel caso di immagini da satellite, con le trasformazioni affini si possono eliminare gli errori di rotazione e di cambiamento di scala (introdotti quando l'altezza del satellite varia rispetto all'orbita regolare prevista).

Come sopra detto, mediante una trasformazione affine Φ un triangolo va in un triangolo, un rettangolo in un parallelogramma (si osservi che gli angoli non vengono conservati, in generale)a meno che Φ non sia una trasformazione di scala (similitudine) che conserva la forma della figura. Quindi, se un rettangolo è trasformato in un quadrilatero (che non è un parallelogramma) non è possibile usare una trasformazione affine, ma una trasformazione più generale.

Nel caso particolare, basta una trasformazione *proiettiva* che dipende da 9 parametri omogenei, quindi 8 non omogenei. Se (x_1, x_2, x_3) sono le coordinate omogenee di P e (x'_1, x'_2, x'_3) quelle di P', la trasformazione proiettiva è del tipo

$$\begin{cases} x'_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3, \\ x'_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3, \\ x'_3 = a_{31}x_1 + a_{32}x_2 + a_{33}x_3. \end{cases}$$

Esercizio 4.1. *Si provi che, alterando per un fattore di proporzionalità i coefficienti della matrice, la trasformazione non muta.*

Nelle coordinate non omogenee $\xi \stackrel{\text{def}}{=} x_1/x_3$ ed $\eta \stackrel{\text{def}}{=} x_2/x_3$ la trasformazione non è più lineare:

$$\xi' = \frac{a_{11}\xi + a_{12}\eta + a_{13}}{a_{31}\xi + a_{32}\eta + a_{33}}, \qquad \eta' = \frac{a_{21}\xi + a_{22}\eta + a_{23}}{a_{31}\xi + a_{32}\eta + a_{33}}.$$

Osservazione 4.5. Ovviamente, si possono considerare anche trasformazioni

$$x' = f_1(x, y), \qquad y' = f_2(x, y),$$

dove f_i sono polinomi di ordine superiore al primo. Ad esempio la trasformazione

$$\begin{cases} x' = a_{11}x + a_{12}y + a_{13}xy + a_{14} \\ y' = a_{21}x + a_{22}y + a_{23}xy + a_{24} \end{cases}$$

porta 4 punti di controllo A, B, C, $D \in G$ nei punti A', B', C', $D' \in G'$ poiché dipende da 8 parametri essenziali. Come caso particolare si consideri

$$x' = xy + 1, \qquad y' = x.$$

Allora il segmento di equazioni parametriche $x = t$, $y = t$ $(0 \leq t \leq 1)$ si trasforma nell'arco di parabola di equazioni parametriche $x' = t^2 + 1$, $y' = t$. □

4.2 Le trasformazioni geometriche 3D

Si considerano sistemi di riferimento cartesiano $Oxyz$ di tipo *destro* o *sinistro*. Nella grafica al computer si preferisce il destro per la rappresentazione di un oggetto ed il sinistro per il riferimento solidale con l'osservatore. Il verso positivo degli angoli è l'antiorario per un sistema destro e quello orario per un sistema sinistro.

Spesso risulta utile anche utilizzare coordinate cilindriche o sferiche [16].

Usando coordinate proiettive non omogenee $(x, y, z, 1)$, le matrici delle trasformazioni affini dello spazio hanno la seguente struttura a blocchi:

$$\begin{pmatrix} A & B \\ O & 1 \end{pmatrix},$$

dove il blocco A pilota le trasformazioni di scala e di rotazione, mentre B pilota le traslazioni. Anche le trasformazioni affini dello spazio conservano l'allineamento ed il parallelismo. Le trasformazioni elementari 3D sono le seguenti.

- Traslazione. $P' = P + B$, rappresentata dalla matrice

$$\begin{pmatrix} 1 & 0 & 0 & x_0 \\ 0 & 1 & 0 & y_0 \\ 0 & 0 & 1 & z_0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.2.2}$$

  dove $B = (x_0, y_0, z_0)$ in coordinate affini.

- Scala. $P' = KP$, dove

$$\begin{pmatrix} k_x & 0 & 0 & 0 \\ 0 & k_y & 0 & 0 \\ 0 & 0 & k_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{4.2.3}$$

  dove k_x, k_y, k_z sono fattori di scala positivi.

- Rotazione intorno all'asse x. $P' = R_x(\varphi)P$, dove

$$R_x(\varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi & 0 \\ 0 & \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.2.4}$$

- Rotazione intorno all'asse y. $P' = R_y(\varphi)P$, dove

$$R_y(\varphi) = \begin{pmatrix} \cos\varphi & 0 & -\sin\varphi & 0 \\ 0 & 1 & 0 & 0 \\ \sin\varphi & 0 & \cos\varphi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.2.5}$$

- Rotazione intorno all'asse z. $P' = R_z(\varphi)P$, dove

$$R_z(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 & 0 \\ \sin\varphi & \cos\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.2.6}$$

- Simmetria rispetto al piano xy. $P' = S_{xy}P$, dove

$$S_{xy} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.2.7}$$

  Analogamente per le simmetrie rispetto agli altri piani coordinati.

Osservazione 4.6.

- La composizione di rotazioni intorno allo stesso asse è commutativa.
- La composizione di rotazioni intorno ad assi differenti non è una rotazione elementare.
- La composizione di rotazioni elementari intorno ad assi differenti non è commutativa. □

4.3 Quaternioni e rotazioni 3D

I quaternioni furono introdotti da W. R. Hamilton nel 1843 nel tentativo di generalizzare il campo dei numeri complessi per analogia con il passaggio dai reali ai complessi.

Dieci anni furono necessari ad Hamilton per rendersi conto che per effettuare questa generalizzazione sarebbe stato necessario usare uno spazio di dimensione 4 anziché uno spazio di dimensione 3.

I quaternioni hano innumerevoli applicazioni in Matematica, Fisica ed Ingegneria. Qui sarà descritta un'applicazione di particolare interesse ingegneristico che riguarda la parametrizzazione del gruppo $SO(3)$ delle rotazioni nello spazio euclideo $3D$ mediante quaternioni unitari. Questa parametrizzazione permette di passare da una descrizione di $SO(3)$ mediante funzioni trigonometriche ad una descrizione che usa polinomi di secondo grado. Il metodo permette una più agevole calcolabilità delle matrici di rotazione, ed è usato in grafica $3D$ ovunque ci sia la necessità di calcolare un gran numero di matrici di rotazione (per esempio quando si muove il punto di vista in una scena). Altri utilizzi riguardano la robotica ed il controllo dei robot manipolatori [35].

I quaternioni possono essere introdotti equivalentemente in modo vettoriale o in modo matriciale.

4.3.a Quaternioni come vettori

Si consideri lo spazio vettoriale $(\mathbb{R}^4, +, \cdot)$. Si denoti

$$\mathbf{1} = (1,0,0,0), \quad \mathbf{i} = (0,1,0,0), \quad \mathbf{j} = (0,0,1,0), \quad \mathbf{k} = (0,0,0,1).$$

In $\mathbb{R}^4$ si definisce una nuova operazione

$$\bullet \colon \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}^4$$

in questo modo: $\bullet$ è l'unica operazione, bilineare rispetto alla somma ed al prodotto per scalari, distributiva a destra ed a sinistra, tale che

$$\begin{gathered}\mathbf{1}\bullet\mathbf{1} = \mathbf{1}, \quad \mathbf{1}\bullet\mathbf{i} = \mathbf{i} = \mathbf{i}\bullet\mathbf{1}, \quad \mathbf{1}\bullet\mathbf{j} = \mathbf{j} = \mathbf{j}\bullet\mathbf{1}, \quad \mathbf{1}\bullet\mathbf{k} = \mathbf{k} = \mathbf{k}\bullet\mathbf{1},\\ \mathbf{i}\bullet\mathbf{i} = -\mathbf{1}, \quad \mathbf{j}\bullet\mathbf{j} = -\mathbf{1}, \quad \mathbf{k}\bullet\mathbf{k} = -\mathbf{1},\\ \mathbf{i}\bullet\mathbf{j} = \mathbf{k}, \quad \mathbf{j}\bullet\mathbf{k} = \mathbf{i}, \quad \mathbf{k}\bullet\mathbf{i} = \mathbf{j},\\ \mathbf{j}\bullet\mathbf{i} = -\mathbf{k}, \quad \mathbf{k}\bullet\mathbf{j} = -\mathbf{i}, \quad \mathbf{i}\bullet\mathbf{k} = -\mathbf{j}.\end{gathered}$$

Si noti che la proprietà di essere distributiva a sinistra ed a destra ha la seguente forma per ogni q, p, $r \in \mathbb{R}^4$:

$$q \bullet (p + r) = q \bullet p + q \bullet r, \quad (q + p) \bullet r = q \bullet r + p \bullet r.$$

Proposizione 4.1. $(\mathbb{R}^4, +, \bullet)$ *è un* corpo*, ossia è un insieme le cui operazioni soddisfano le proprietà*

1. $(\mathbb{R}^4, +, \bullet)$ *è un anello (*non *commutativo) dotato di unità;*
2. $(\mathbb{R}^4 \setminus \{0\}, \bullet)$ *è un gruppo (*non *commutativo);*

3. per ogni q, p, $r \in \mathbb{R}^4$ vale la proprietà distributiva (a sinistra ed a destra).

La dimostrazione della precedente proposizione è molto semplice, e consiste nella verifica del fatto che $\bullet$ è associativa, che $\mathbf{1}$ è l'elemento neutro, e che esiste un elemento inverso di ogni elemento non nullo (si veda più avanti per una sua espressione).

Definizione 4.1. L'insieme $\mathbb{R}^4$ dotato delle operazioni $+$ e $\bullet$ è detto *corpo dei quaternioni*, ed è denotato da $\mathbb{H}$.

Sia $q \in \mathbb{H}$, Si scrive

$$q = q_0\mathbf{1} + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k} = q_0 + \vec{q}.$$

Si dice che q_0 è la *parte reale* di q, e $\vec{q} = q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}$ è la *parte immaginaria* di q. Si definisce il *coniugato* di q come il quaternione $\bar{q} = q_0 - \vec{q}$. Si dice che un quaternione q è *reale* se $q = \bar{q}$ (ovvero se la parte immaginaria è nulla); si dice che un quaternione è *immaginario* se $q = -\bar{q}$ (ovvero se la parte reale è nulla). L'insieme dei quaternioni immaginari si denota con $\mathbb{H}_i$.

Si definisce la *norma* di q come il numero reale non negativo $\|q\| = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2}$. Si noti che, in generale, $q \bullet p \neq p \bullet q$. Valgono le seguenti proprietà.

1. $q \bullet \bar{q} = \|q\|^2$;
2. $\|p \bullet q\| = \|p\|\|q\| = \|q \bullet p\|$;
3. $\overline{p + q} = \bar{p} + \bar{q}$;
4. $\overline{q \bullet p} = \bar{p} \bullet \bar{q}$;
5. $q^{-1} = \bar{q}/\|q\|^2$.

Si notino le analogie e le differenze (date dalla mancanza di commutatività) di $\mathbb{H}$ con $\mathbb{C}$. In realtà si può dimostrare che le seguenti mappe iniettive commutano con le operazioni, quindi sono *morfismi* delle rispettive strutture algebriche:

$$\mathbb{R} \overset{rc}{\to} \mathbb{C} \overset{ch}{\to} \mathbb{H} \tag{4.3.8}$$

dove $rc(x) = x + i0$, $ch(x + iy) = x + y\mathbf{i} + 0\mathbf{j} + 0\mathbf{k}$. Si noti che i precedenti morfismi commutano anche con il modulo ed il coniugio.

Esempio 4.1. Siano $q = 1 - 2\mathbf{j} + \mathbf{k}$, $p = -2 + \mathbf{i} + 3\mathbf{k} \in \mathbb{H}$. Allora il prodotto $q \bullet p$ tra i due quaternioni vale

$$\begin{aligned} q \bullet p &= (1 - 2\mathbf{j} + \mathbf{k}) \bullet (-2 + \mathbf{i} + 3\mathbf{k}) \\ &= -2 + \mathbf{i} + 3\mathbf{k} + 4\mathbf{j} - 2\mathbf{j} \bullet \mathbf{i} - 6\mathbf{j} \bullet \mathbf{k} - 2\mathbf{k} + \mathbf{k} \bullet \mathbf{i} + 3\mathbf{k} \bullet \mathbf{k} \\ &= -5 - 5\mathbf{i} + 5\mathbf{j} + 3\mathbf{k} \end{aligned}$$

Esercizio 4.2. *Siano dati due quaternioni*

$$q = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}, \quad p = p_0 + p_1\mathbf{i} + p_2\mathbf{j} + p_3\mathbf{k}.$$

1. Dimostrare che

$$q \bullet p = (-p_3q_3 - p_2q_2 - p_1q_1 + p_0q_0) + (-p_2q_3 + p_3q_2 + p_0q_1 + p_1q_0)\mathbf{i} + (p_1q_3 + p_0q_2 - p_3q_1 + p_2q_0)\mathbf{j} + (p_0q_3 - p_1q_2 + p_2q_1 + p_3q_0)\mathbf{k}.$$

2. Dimostrare che

$$q \bullet p - p \bullet q = (-2p_2q_3 + 2p_3q_2)\mathbf{i} + (2p_1q_3 - 2p_3q_1)\mathbf{j} + (-2p_1q_2 + 2p_2q_1)\mathbf{k}.$$

4.3.b Quaternioni come matrici

Sia

$$\tilde{\mathbb{H}} = \left\{ A \in \mathbb{C}^{2,2} \mid A = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix},\ a, b \in \mathbb{C} \right\}.$$

Si noti che $\tilde{\mathbb{H}}$ è un sottospazio vettoriale *reale* di $\mathbb{C}^{2,2}$ di dimensione (reale) 4. Si può dimostrare che $\tilde{\mathbb{H}}$ è chiuso rispetto al prodotto di matrici. Se $A \in \tilde{\mathbb{H}}$,

$$\det A = \det \begin{pmatrix} a_0 + ia_1 & a_2 + ia_3 \\ -a_2 + ia_3 & a_0 - ia_1 \end{pmatrix} = a_0^2 + a_1^2 + a_2^2 + a_3^2,$$

dunque $A \in \mathbb{H}$ è una matrice invertibile se e solo se $A \neq O$. Si può verificare che

$$A^{-1} = \frac{A^*}{\|A\|^2},$$

dove $\|A\|$ è la norma di Frobenius (si noti che $\det A = \|A\|^2$).

Si introducano le seguenti matrici, dette *matrici di Pauli*:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Queste matrici sono usate in meccanica quantistica per formulare le equazioni di una particella non relativistica dotata di *spin*.

Una base privilegiata di $\tilde{\mathbb{H}}$ è $\{Id, I, J, K\}$, dove $I = i\sigma_3$, $J = i\sigma_2$, $K = i\sigma_1$.

Proposizione 4.2. *Si ha il seguente isomorfismo tra* $\mathbb{H}$ *ed* $\tilde{\mathbb{H}}$*:*

$$\mathbb{H} \to \tilde{\mathbb{H}}, \quad q = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k} \mapsto q_0Id + q_1I + q_2J + q_3K, \tag{4.3.9}$$

o, equivalentemente

$$\mathbb{H} \to \tilde{\mathbb{H}}, (q_0, q_1, q_2, q_3) \mapsto \begin{pmatrix} q_0 + iq_1 & q_2 + iq_3 \\ -q_2 + iq_3 & q_0 - iq_1 \end{pmatrix}. \tag{4.3.10}$$

Per la dimostrazione basta verificare che la applicazione definita sopra trasforma l'operazione $\bullet$ nell'operazione di prodotto tra matrici. Si confronti l'isomorfismo precedente con l'isomorfismo (7.3.2) nel caso dei numeri complessi.

Esercizio 4.3. *Si dimostri che se il seguente isomorfismo trasforma q in A, allora $\bar{q}$ è trasformato nella matrice A^*.*

4.3.c Quaternioni e rotazioni $3D$

D'ora in avanti sarà usata la lettera $\mathbb{H}$ per indicare una delle due formulazioni dei quaternioni. Si noti che le operazioni di coniugato, norma, ecc. possono essere utilizzate anche nela formulazione matriciale in virtù del precedente isomorfismo.

Il sottospazio $SU(2) \subset \mathbb{H}$ è l'insieme dei *quaternioni unitari*, ossia dei quaternioni q tali che $\|q\| = 1$.

Lemma 4.1. *La seguente applicazione $\mathbb{R}^3 \to \mathbb{H}_i$,*

$$\vec{v} = (x, y, z) \to q_{\vec{v}} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = \begin{pmatrix} xi & y + zi \\ -y + zi & -xi \end{pmatrix} = i \begin{pmatrix} x & z - iy \\ z + iy & -x \end{pmatrix}$$

è un isomorfismo lineare (su $\mathbb{R}$).

Dimostrazione. La linearità dell'applicazione è immediata, il fatto che è un isomorfismo segue dall' iniettività (banale) e dal fatto che la dimensione di entrambi gli spazi è 3. QED

Teorema 4.1. *L'applicazione*

$$\mathcal{CK}\colon SU(2) \to SO(3), \quad q \mapsto \mathcal{CK}(q)$$

dove $\mathcal{CK}(q)\colon \mathbb{R}^3 \to \mathbb{R}^3$, $T_q(\vec{v}) = qq_{\vec{v}}q^{-1}$, è un morfismo di gruppi con nucleo uguale a $\{\pm Id\}$.

Dimostrazione. Poiché $\overline{q_{\vec{v}}} = -q_{\vec{v}}$ (essendo $q_{\vec{v}}$ immaginario) e $q^{-1} = \bar{q}$ (essendo q unitario) risulta, per ogni $\vec{v} \in \mathbb{R}^3$,

$$\overline{qq_{\vec{v}}q^{-1}} = \overline{q^{-1}}\,\overline{q_{\vec{v}}}\,\bar{q} = -qq_{\vec{v}}q^{-1},$$

quindi il quaternione $qq_{\vec{v}}q^{-1}$ è immaginario. Per questo, esiste un vettore $\vec{v'} = (x', y', z')$ tale che $q_{\vec{v'}} = qq_{\vec{v}}q^{-1}$, dunque la definizione è ben posta.

È facile verificare che l'applicazione $\mathcal{CK}(q)$ è lineare.

Si noti che $\det(q_{\vec{v}}) = x^2 + y^2 + z^2 = x'^2 + y'^2 + z'^2 = \det(q_{\vec{v'}})$, per la regola di Binet applicata al prodotto $qq_{\vec{v}}q^{-1}$. Pertanto $\mathcal{CK}(q)$ è un'applicazione ortogonale. Rimane solo da verificare che questa trasformazione abbia determinante 1. Scrivendo

$$\begin{pmatrix} x'i & y' + z'i \\ -y' + z'i & -x'i \end{pmatrix} = \begin{pmatrix} q_0 + iq_1 & q_2 + iq_3 \\ -q_2 + iq_3 & q_0 - iq_1 \end{pmatrix} \begin{pmatrix} xi & y + zi \\ -y + zi & -xi \end{pmatrix} \begin{pmatrix} q_0 - iq_1 & -q_2 - iq_3 \\ q_2 - iq_3 & q_0 + iq_1 \end{pmatrix}$$

si ha

$$\begin{aligned}
x' &= (-q_3^2 - q_2^2 + q_1^2 + q_0^2)x + (-2q_0q_3 + 2q_1q_2)y + (2q_1q_3 + 2q_0q_2)z \\
y' &= (2q_0q_3 + 2q_1q_2)x + (-q_3^2 + q_2^2 - q_1^2 + q_0^2)y + (2q_2q_3 - 2q_0q_1)z \\
z' &= (2q_1q_3 - 2q_0q_2)x + (2q_2q_3 + 2q_0q_1)y + (q_3^2 - q_2^2 - q_1^2 + q_0^2)z
\end{aligned}$$

da cui si ottiene la matrice di $\mathcal{CK}(q)$ rispetto alla base canonica $\mathcal{C}$ di $\mathbb{R}^3$

$$\mathcal{M}_{\mathcal{C}}^{\mathcal{C}}(\mathcal{CK}(q)) = \begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_1q_2 + q_0q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{pmatrix}, \qquad (4.3.11)$$

dove si ha $q_0^2+q_1^2+q_2^2+q_3^2 = 1$. Si verifica (meglio con l'ausilio di un programma di calcolo simbolico, come `maxima`, liberamente disponibile in `http://maxima.sourceforge.net`) che il determinante di questa matrice è 1, che implica che la matrice è un elemento di $SO(3)$.

Richiedendo in (4.3.11) che la matrice sia uguale all'identità si ha che la parte immaginaria del quaternione q si annulla e si ha $q_0^2 = 1$, da cui si ottiene l'affermazione sul nucleo di $\mathcal{CK}$. QED

Definizione 4.2. Il morfismo di gruppi $\mathcal{CK}\colon SU(2) \to SO(3)$ definito nel teorema precedente è detto *parametrizzazione di Cayley–Klein* del gruppo $SO(3)$.

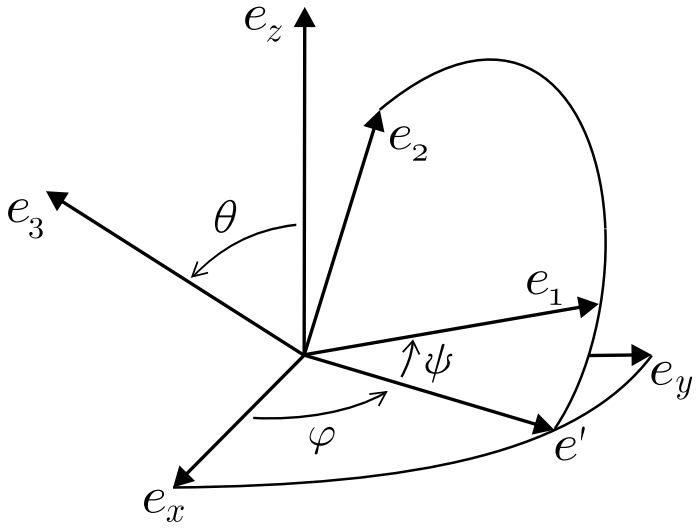


Figura 4.1. Angoli di Eulero

Si confronti la parametrizzazione (4.3.11) con la parametrizzazione di $SO(3)$ data dagli angoli di Eulero (figura 4.1; si veda per la definizione degli angoli di Eulero [21], o un qualunque libro di meccanica razionale). Questa parametrizzazione permette di ottenere una matrice di rotazione generica (ossia un elemento di $SO(3)$) mediante la rotazione

di tre angoli φ, θ, ψ rispetto a tre assi in successione. Più precisamente, si osservi che ogni matrice di $SO(3)$ è un'applicazione lineare che trasforma una base ortonormale data $\{\vec{e}_x, \vec{e}_y, \vec{e}_z\}$ in un'altra base ortonormale $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$. Quest'ultima può essere individuata dalla seguente successione di trasformazioni ortogonali.

1. una prima trasformazione ortogonale che porta la base ortonormale $\{\vec{e}_x, \vec{e}_y, \vec{e}_z\}$ nella base ortonormale $\{\vec{e'}_x, \vec{e'}_y, \vec{e'}_z\}$. La trasformazione è una rotazione attorno all'asse $\vec{e}_z$ (dunque $\vec{e}_z = \vec{e'}_z$) di un angolo φ in modo tale che il trasformato $\vec{e'}_x$ dell'asse $\vec{e}_x$ giaccia nel piano individuato da $\{\vec{e}_1, \vec{e}_2\}$. La matrice di questa trasformazione è
$$\begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}; \tag{4.3.12}$$

2. una seconda trasformazione ortogonale che porta la base ortonormale $\{\vec{e'}_x, \vec{e'}_y, \vec{e'}_z\}$ nella base ortonormale $\{\vec{e''}_x, \vec{e''}_y, \vec{e''}_z\}$. La trasformazione è una rotazione intorno all'asse $\vec{e'}_x$ (dunque $\vec{e'}_x = \vec{e''}_x$) di un angolo θ in modo tale che il trasformato $\vec{e''}_y$ dell'asse $\vec{e'}_y$ giaccia nel piano individuato da $\{\vec{e}_1, \vec{e}_2\}$. La matrice di questa trasformazione è
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}; \tag{4.3.13}$$

3. una terza trasformazione ortogonale che porta la base ortonormale $\{\vec{e''}_x, \vec{e''}_y, \vec{e''}_z\}$ nella base ortonormale $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$. La trasformazione è una rotazione intorno all'asse $\vec{e''}_z$ (dunque $\vec{e''}_z = \vec{e}_3$) che porta $\{\vec{e''}_x, \vec{e''}_y\}$ a coincidere con $\{\vec{e}_1, \vec{e}_2\}$. La matrice di questa trasformazione è
$$\begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.3.14}$$

La generica matrice di $SO(3)$ si ottoene, dunque, come il prodotto delle precedenti:

$$\begin{pmatrix} \cos\varphi\cos\psi - \sin\varphi\sin\psi\cos\theta & -\sin\varphi\cos\psi\cos\theta - \cos\varphi\sin\psi & \sin\varphi\sin\theta \\ \cos\varphi\sin\psi\cos\theta + \sin\varphi\cos\psi & \cos\varphi\cos\psi\cos\theta - \sin\varphi\sin\psi & -\cos\varphi\sin\theta \\ \sin\psi\sin\theta & \cos\psi\sin\theta & \cos\theta \end{pmatrix}. \tag{4.3.15}$$

È possibile ricostruire dalla precedente matrice un quaternione unitario che la descrive. In particolare, ponendo $q = q_0 Id + q_1 I + q_2 J + q_3 K$ si ha

$$q_0 = \cos\frac{\psi+\varphi}{2}\cos\frac{\theta}{2}, \qquad q_1 = \cos\frac{\psi-\varphi}{2}\sin\frac{\theta}{2}, \tag{4.3.16}$$

$$q_2 = \sin\frac{\psi-\varphi}{2}\sin\frac{\theta}{2}, \qquad q_3 = \sin\frac{\psi+\varphi}{2}\cos\frac{\theta}{2}. \tag{4.3.17}$$

Questa corrispondenza è un'inversa della parametrizzazione quaternionica di $SO(3)$. Da qui si può dimostrare che ogni quaternione unitario q può essere scritto come

$$q = \cos(\alpha/2) + \sin(\alpha/2)q_{\vec{v}}. \tag{4.3.18}$$

Quindi, *ogni quaternione unitario q ammette la seguente interpretazione geometrica: esso rappresenta una rotazione di un angolo α attorno ad un asse $\vec{v}$.* Si noti che i quaternioni unitari $\pm q$ rappresentano la stessa rotazione. Questo poichè l'asse di rotazione rappresentato dai vettori $\pm\vec{v}$ è lo stesso.

Si può dimostrare che la parametrizzazione di Cayley-Klein per una serie di ragioni è molto più efficiente degli angoli di Eulero in problemi come, ad esempio:

1. la rotazione del punto di vista ('camera') dalla posizione (φ, θ, ψ) alla posizione $(\varphi', \theta', \psi')$ [37];

2. la composizione di due o più rotazioni: moltiplicare due quaternioni comporta l'esecuzione di 16 moltiplicazioni, mentre moltiplicare due matrici di $SO(3)$ ne comporta 27.

Esercizio 4.4. *Moltiplicare tra loro i quaternioni $q = 3+2\mathbf{i}-\mathbf{j}+2\mathbf{k}$ e $p = -1-2\mathbf{j}+\mathbf{k}$; verificare che $q \bullet p \neq p \bullet q$.*

SOLUZIONE. Si ha $q \bullet p = -7 + \mathbf{i} - 7\mathbf{j} - 3\mathbf{k}$, $p \bullet q = -7 - 5\mathbf{i} - 3\mathbf{j} + 5\mathbf{k}$. □

Esercizio 4.5. *Verificare che il prodotto dei quaternioni $x = -\mathbf{i} + \mathbf{j} + 4\mathbf{k}$ e $y = -3 + 3\mathbf{j} - \mathbf{k}$ è uguale al prodotto delle corrispondenti matrici in $\tilde{\mathbb{H}}$.*

SOLUZIONE. Basta osservare che

$$x = \begin{pmatrix} -i & 1+4i \\ -1+4i & i \end{pmatrix}, \quad y = \begin{pmatrix} -3 & 3-i \\ -3-i & -3 \end{pmatrix}.$$

□

Esercizio 4.6. *Trovare tre quaternioni corrispondenti alle tre rotazioni usate nella parametrizzazione di $SO(3)$ con gli angoli di Eulero.*

SOLUZIONE. Si può iniziare ricavando il quaternione q_φ indotto dalla rotazione (4.3.12). Uguagliando le matrici (4.3.11) e (4.3.12) si ottengono 9 equazioni nelle 4 incognite q_0, q_1, q_2, q_3, componenti il quaternione q_φ. Delle 9, 4 sono

$$q_1 q_3 = \pm q_0 q_2, \qquad q_2 q_3 = \pm q_0 q_1,$$

che implicano $q_1 = q_2 = 0$ oppure $q_0 = q_3 = 0$ (ma quest'ultima si riconduce ad un caso banale). Richiedendo $q_1 = q_2 = 0$ e ricordando che q_φ è unitario si ha

$$q_\varphi = \begin{pmatrix} \cos\frac{\varphi}{2} & i\sin\frac{\varphi}{2} \\ i\sin\frac{\varphi}{2} & \cos\frac{\varphi}{2} \end{pmatrix}. \tag{4.3.19}$$

Analogamente si trova che

$$q_\theta = \begin{pmatrix} \cos\frac{\theta}{2} + i\sin\frac{\theta}{2} & 0 \\ 0 & \cos\frac{\theta}{2} + i\sin\frac{\theta}{2} \end{pmatrix}, \quad q_\psi = \begin{pmatrix} \cos\frac{\psi}{2} & i\sin\frac{\psi}{2} \\ i\sin\frac{\psi}{2} & \cos\frac{\psi}{2} \end{pmatrix}. \tag{4.3.20}$$

□

Esercizio 4.7. *Dimostrare le formule* (4.3.16) *e* (4.3.17)*.*

SOLUZIONE. Basta calcolare il prodotto $q_\psi q_\theta q_\varphi$ delle matrici dell'esercizio precedente. □

Esercizio 4.8. *Trovare il quaternione unitario corrispondente a* $p = -1 - 2\mathbf{j} + \mathbf{k}$*. Individuare la matrice di* $SO(3)$ *descritta da* p*. Individuare i tre angoli di Eulero della corrispondente matrice di rotazione.*

SOLUZIONE. Il quaternione unitario corrispondente a p è

$$q = \frac{p}{\|p\|} = \frac{1}{\sqrt{6}}(-1 - 2\mathbf{j} + \mathbf{k}).$$

Per trovare la matrice di $SO(3)$ corrispondente a q si utilizzi la matrice (4.3.11) sostituendo i valori dei coefficienti di q.

Per trovare gli angoli di Eulero si può utilizzare la matrice di rotazione trovata sopra ed uguagliarla alla (4.3.15). Ovviamente si possono subito trovare valori di θ utilizzando il tezo elemento della terza riga, poi gli altri elementi usando la terza riga e la terza colonna. Verificare il risultato su tutti gli elementi. Un procedimento alternativo può essere quello di usare le formule (4.3.16) e (4.3.17). □

Esercizio 4.9. *Si interpreti geometricamente la matrice di rotazione corrispondente ad un quaternione unitario del tipo* $w = a + b\mathbf{i} \in \mathbb{C} \subset \mathbb{H}$*.*

Esercizio 4.10. (Esercizio non banale – si pensi alle radici di un numero complesso per analogia.) *Descrivere l'insieme delle soluzioni (se esistono) dell'equazione* $x^2 = z$*, ove* $z = 2\mathbf{i} - \mathbf{j} + \mathbf{k}$ *ed* $x \in \mathbb{H}$ *è una incognita.*

4.4 Trasformazioni parallele e trasformazioni prospettiche

Il modello di mondo che usualmente adoperiamo è uno spazio euclideo tridimensionale, mentre lo schermo di un computer è una porzione di piano bidimensionale. Ogni processo di rappresentazione di uno spazio tridimensionale su uno spazio bidimensionale è chiamato una *proiezione*. Praticamente, si tratta di avere una visione bidimensionale di un oggetto tridimensionale che sia il più possibile realistica od almeno significativa.

Esistono molte vie possibile, ma due sono le tecniche principali di trasformazione di vista (3D *viewing*): le *proiezioni parallele* e le *proiezioni prospettiche*.

Le proiezioni parallele sono particolari trasformazioni che, fissata una direzione visuale $\vec{u}$, ad un punto P dell'oggetto fanno corrispondere un punto P' dello schermo ottenuto come intersezione del piano dello schermo con la retta r per P ed avente la direzione $\vec{u}$.

Si noti che tutti i punti di r hanno come proiezione lo stesso punto P'. La proiezione è detta *ortografica* se $\vec{u}$ è perpendicolare allo schermo, altrimenti è detta *obliqua*.

Rappresentazioni di questo tipo sono familiari agli ingegneri, che spesso usano un insieme di proiezioni ortografiche (sui piani coordinati), per esempio nelle proiezioni planimetriche di tipo architettonico.

Analiticamente, si tratta di scrivere il *cilindro* che proietta l'oggetto su di un piano, parallelamente alla direzione data.

Queste proiezioni, che danno informazioni precise sulle mutue distanze tra i punti non appartenenti a rette parallele ad $\vec{u}$, non danno una vista realistica dell'oggetto.

Le proiezioni prospettiche forniscono, invece, il realismo, rendono l'immagine bidimensionale come una fotografia; la grandezza degli oggetti dipende dalla distanza δ dall'osservatore, che viene assimilato ad un punto, dal quale escono i raggi di luce riflessi dall'oggetto.

Analiticamente, si tratta di scrivere il *cono* che proietta l'oggetto su di un piano avente il vertice E nell'osservatore. Se l'osservatore si sposta all'infinito, il cono diventa un cilindro e la proiezione prospettica diventa proiezione parallela.

Accanto a proiezioni prospettiche con un punto di fuga, ci sono proiezioni prospettiche con 2 o 3 punti di fuga.

Nella proiezione centrale (con un punto di fuga), l'osservatore si trova su uno degli assi coordinati, la direzione visuale è quella che va dall'osservatore al centro del sistema di riferimento dell'oggetto, il piano di proiezione è spesso ortogonale a questa direzione. La proiezione avviene in due passi:

1. si trasforma il sistema $Oxyz$ nel quale è rappresentato l'oggetto nel sistema $Ex_ey_ez_e$ dell'occhio dell'osservatore;

2. si trasforma quest'ultimo nel sistema del piano dello schermo x_sy_s.

Passo 1. 1. Si trasla il punto di vista nell'origine di xyz. Indicata con δ la distanza dell'osservatore dall'oggetto, la traslazione è rappresentata dalla matrice

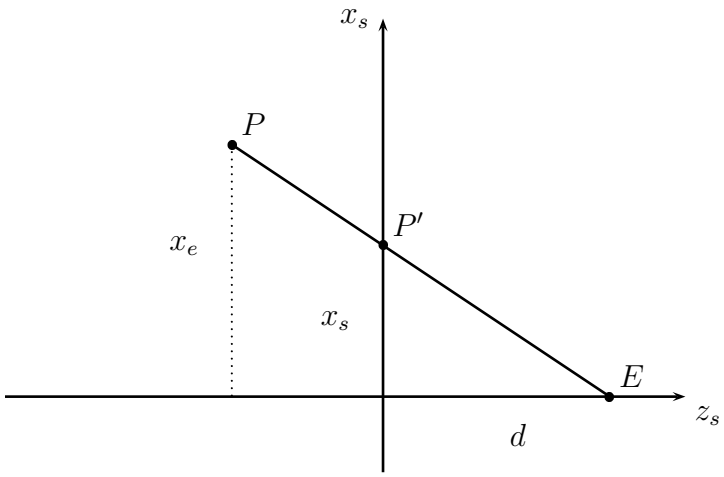
$$T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -\delta \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

2. Si effettua una simmetria rispetto al piano x_ey_e per avere un sistema sinistro, tramite la matrice S_{xy} (4.2.7).

Quindi, questo passo è dato dalla matrice ST:

$$x_e = x, \quad y_e = y, \quad z_e = \delta - z.$$

Passo 2. Si procede alla proiezione sul piano dello schermo x_sy_s, ortogonale all'asse z, avente distanza (focale) d dall'osservatore. Indichiamo con P la proiezione sul piano x_sz_s dell'oggetto (punto).



Guardando alle proiezioni ortogonali sul piano xz si ha

$$\frac{x_s}{d} = \frac{x_e}{z_e} \quad \Rightarrow \quad x_s = \frac{dx_e}{z_e} = \frac{dx}{\delta - z}.$$

Allo stesso modo, nel piano yz si ha $y_s = dy/(\delta - z)$. Tale proiezione si chiama *ad un punto di fuga*, poiché il piano di proiezione è perpendicolare alla direzione di vista e le rette parallele a x e quelle parallele ad y rimangono tali anche nella proiezione, mentre le rette parallele a z convergono ad un punto (il cosiddetto *punto di fuga*).

Se il punto di vista è arbitrario, la proiezione avviene lo stesso seguendo i passi (1) e (2), ma il passo (1) è più complicato: non basta, infatto, una traslazione per portare $Oxyz$ in $Ex_ey_ez_e$. In generale, la matrice che rappresenta il movimento è

$$M = \begin{pmatrix} -\sin\theta & \cos\theta & 0 & 0 \\ -\cos\theta\cos\varphi & -\sin\theta\cos\varphi & \sin\varphi & 0 \\ -\cos\theta\sin\varphi & -\sin\theta\sin\varphi & -\cos\varphi & \delta \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

dove θ, φ sono angoli che permettono di vedere l'oggetto da diverse posizioni; se $\varphi = 0$ e $\theta = -\pi/2$ si riottiene la matrice ST della proiezione centrale. Quindi

$$\begin{pmatrix} x_e \\ y_e \\ z_e \\ 1 \end{pmatrix} = M \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad \text{oppure} \quad (x_e\ y_e\ z_e\ 1)^T = M(x\ y\ z\ 1)^T.$$

Il passo (2) è uguale a quello della proiezione centrale

$$x_s = \frac{dx_e}{z_e}, \qquad y_s = \frac{dy_e}{z_e}.$$

Variando δ varia la grandezza dell'oggetto, variando φ e θ (che si possono considerare coordinate sferiche) si avranno prospettive diverse, variando, invece, d si avrà una modifica della grandezza dell'immagine ma non della prospettiva.

4.5 Nota storica

Nelle proiezioni prospettiche interviene in modo essenziale la geometria proiettiva, che storicamente nasce proprio dalla teoria e pratica della prospettiva sviluppatasi nel Rinascimento italiano [15].

Già Euclide (III sec. a. C.) nell'opera "Ottica" costruisce una teoria scientifica della visione, che dal punto di vista della struttura interna può considerarsi una parte della Geometria. Significativo è il teorema $VIII$, che stabilisce che *la differenza apparente tra due grandezze uguali viste da distanze differenti è determinata non dal rapporto tra queste distanze, ma dai rapporti tra gli angoli visivi corrispondenti.*

I Greci avrebbero, quindi, costruito una *perspectiva naturalis* che è una prospettiva angolare, che portò poi alla *perspectiva artificialis*, che è una prospettiva lineare, derivata puramente da considerazioni geometriche.

Per dipingere, i pittori del Rinascimento furono condotti alla Matematica, convinti che la Matematica è l'essenza del mondo reale, poiché l'universo è ordinato e spiegabile razionalmente in termini geometrici.

Leon Battista Alberti (1404–1472) diceva anche il primo requisito per il pittore è conoscere la Geometria: gli occhi del pittore si trovano davanti alla tela come davanti ad una finestra aperta sul mondo. Si osservi che lo stesso termine 'finestra' (*window*) è usato attualmente per denominare la porzione di schermo in cui è rappresentato il 'mondo'.

I problemi che si trovano ad affrontare gli informatici per rappresentare sullo schermo gli oggetti tridimensionali sono analoghi a quelli incontrati dai pittori per rappresentare sulla tela il mondo che li circondava.

Notevoli contributi alla teoria della prospettiva sono stati dati proprio da architetti ed ingegneri come L. B. Alberti, G. Desargues (1593–1662), J.V. Poncelet (1788–1867), G. Monge (1746–1823), col quale la prospettiva finisce di essere una 'geometria per pittori' e diventa 'geometria degli ingegneri', geometria per disegnatori tecnici che, dalle proiezioni si propongono di ricostruire l'oggetto raffigurato in forma e grandezza.

La singolare bellezza della geometria proiettiva e l'elegante armonia delle sue dimostrazioni fecero sì che essa diventasse lo studio favorito dei geometri del XIX secolo, in particolare di J. Steiner (1796–1863), considerato come il più grande studioso di geometria dei tempi moderni (paragonabile ad Apollonio nei tempi antichi). Steiner nutriva un intenso disprezzo per i metodi analitici, egli era del parere che i calcoli fossero sostituto del pensiero, mentre la geometria avrebbe dovuto studiare l'attività creativa. Non si può negare, però, che i metodi analitici hanno contribuito in maniera determinante allo sviluppo della geometria proiettiva, in particolare l'uso delle coordinate omogenee scoperte contemporaneamente da quattro matematici: J. Plücker (1801–1868), K.F. Feuerbach (1800–1834), A.F. Möbius (1790–1868) ed E. Bobillier (1797–1832).

Il nuovo punto di vista (analitico) fu poi sviluppato da M. Chasles (1793–1890) e da F. Klein (1849–1925), che nel celebre "Programma" (università di Erlangen, dicembre 1872) fonda la geometria sullo studio delle proprietà che sono invarianti rispetto ad un fissato gruppo di trasformazioni.

È così possibile costruire infinite geometrie a seconda del gruppo di trasformazioni scelto.

«. . . la forma matematica si conosce
attraverso le sue trasformazioni.
Si potrebbe dire dell'essere matematico:
dimmi come ti trasformano e ti dirò chi sei.»
(G. Bachelard, Le nouvel esprit scientifique, 1949)

4.6 Curve di Bézier

Le curve di Bézier sono curve algebriche di terzo grado usate per approssimare archi di curve qualunque di cui siano noti i punti iniziale e finale ed i vettori tangenti in questi due punti. Queste curve sono usate nella grafica per disegnare, rappresentando oggetti reali od immaginari. Le curve di Bézier ammettono un analogo di dimensione superiore, le superficie di Bézier, per le quali si rimanda al testo [18].

Il motivo per cui si preferiscono curve di terzo grado è che le curve di secondo grado sono 'poco versatili' e le curve di quarto grado sono difficilmente 'controllabili'.

Siano P_0, $P_3 \in \mathbb{R}^3$ i due estremi di una curva di Bézier. I vettori tangenti in P_0 e P_3 sono assegnati tramite due punti P_1, $P_2 \in \mathbb{R}^3$, usando i vettori $P_1 - P_0$ e $P_3 - P_2$ nel modo che segue.

Sia $p(t) = a_3t^3 + a_2t^2 + a_1t + a_0 \in \mathbb{R}_3[t]$. Richiedendo che

$$p(0) = P_0, \quad p(1) = P_3, \quad p'(0) = 3(P_1 - P_0), \quad p'(1) = 3(P_3 - P_2),$$

dove il coefficiente 3 è inserito per motivi di comodità, si ottiene la curva

$$p(t) = (1-t)^3P_0 + 3t(1-t)^2P_1 + 3t^2(1-t)P_2 + t^3P_3. \tag{4.6.21}$$

Definizione 4.3. La curva (4.6.21) si dice *curva di Bézier di grado* 3 determinata dai punti P_0, P_1, P_2, $P_3 \in \mathbb{R}^3$.

Proposizione 4.3. *Le curve di Bézier di grado* 3 *risultano essere medie pesate dei polinomi di Bernstein di grado* 3 *(paragrafo A.4).*

Dimostrazione. Segue la definizione dei polinomi di Bernstein e dalle proprietà (1) e (3) del lemma A.1. QED

Da questo fatto segue che le proprietà generali dei polinomi di Bernstein dànno interessanti proprietà delle curve di Bézier. Quindi, è bene generalizzare la definizione 4.3 ad un numero arbitrario di punti per poi analizzare le proprietà generali di queste curve.

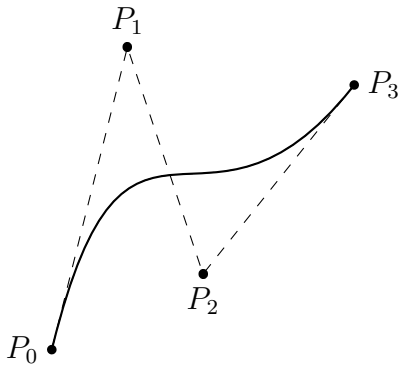


Figura 4.2. Una curva di Bézier

Definizione 4.4. Siano $P_0, \ldots, P_n \in \mathbb{R}^3$ e $t \in [0,1]$. Si definiscano

$$P_i^0 \stackrel{\text{def}}{=} P_i,$$
$$P_i^r(t) \stackrel{\text{def}}{=} (1-t)P_i^{r-1}(t) + tP_{i+1}^{r-1}(t),$$

per $r = 1, \ldots, n$, $i = 0, \ldots, n-r$.

La curva descritta da $P_0^n(t)$ è detta *curva di Bézier di grado n*.

L'algoritmo sopra descritto è dovuto a P. de Casteljau (1959), che lavorava per la Citröen, ma i suoi risultati rimasero coperti dal segreto industriale. I lavori di P. Bézier sono stati pubblicati nel '70 e servirono allo sviluppo di un sistema di CAD per l'industria automobilistica: Bézier lavorava, infatti, alla Renault.

Osservazione 4.7. Una curva di Bézier di grado n non è piana, in generale, se $n > 2$.

La relazione tra i polinomi di Bernstein e le curve di Bézier è descritta dal seguente teorema.

Teorema 4.2. *Per $r \in \{0, \ldots, n\}$ si ha*

$$P_i^r(t) = \sum_{j=0}^{r} P_{i+j} B_j^r(t).$$

DIMOSTRAZIONE. Per induzione su r si ha

$$\begin{aligned} P_i^r(t) &= (1-t)P_i^{r-1}(t) + tP_{i+1}^{r-1}(t) \\ &= (1-t)\sum_{k=i}^{i+r-1} P_k B_{k-i}^{r-1}(t) + t\sum_{k=i+1}^{i+r} P_k B_{k-i-1}^{r-1}(t) \\ &= \sum_{k=i}^{i+r} P_k((1-t)B_{k-i}^{r-1}(t) + tB_{k-i-1}^{r-1}(t)) \\ &= \sum_{j=0}^{r} P_{i+j} B_j^r(t). \end{aligned}$$

QED

Corollario 4.1. $P_0^n(t) = \sum_{j=0}^n P_j B_j^n(t)$.

Da questa relazione seguono una serie di proprietà delle curve di Bézier, di seguito elencate. Sia $P_0^n(t)$ una curva di Bézier.

1. **Invarianza per affinità.** Sia $\Phi\colon \mathbb{R}^3 \to \mathbb{R}^3$ una trasformazione affine, cioè $\Phi(X) = AX + B$, allora
$$\Phi(P_0^n(t)) = \sum_j \Phi(P_j) B_j^n(t).$$
Infatti
$$\begin{aligned} \sum_j \Phi(P_j) B_j^n(t) &= \sum_j (AP_j + B) B_j^n(t) \\ &= A(\sum_j P_j B_j^n(t)) + B \\ &= AP_0^n(t) + B \\ &= \Phi(P_0^n(t)) \end{aligned}$$
2. **Simmetria.** $\sum_j P_j B_j^n(t) = \sum_j P_{n-j} B_{n-j}^n(t) = \sum_j P_{n-j} B_j^n(1-t)$.
3. **Invarianza per riparametrizzazione affine.** Sia
$$\varphi\colon [a,b] \to [0,1],\ u \mapsto t = \frac{u-a}{b-a}$$
un cambiamento affine di parametro. Allora
$$\begin{aligned} \sum_j P_j B_j^n(t) &= \sum_j P_j B_j^n\left(\frac{u-a}{b-a}\right) \\ &= \sum_j P_j \tilde{B}_j^n(u), \end{aligned}$$

dove per $u \in [a, b]$ si è posto

$$\tilde{B}_j^n(u) \stackrel{\text{def}}{=} \binom{n}{j} \frac{(b-u)^{n-j}(u-a)^j}{(b-a)^n}.$$

4. **Precisione lineare.** Si supponga che

$$P_j = \left(1 - \frac{j}{n}\right) P_0 + \frac{j}{n} P_n,$$

per $j = 1, \ldots, n-1$; ossia, si supponga che i punti siano uniformemente distribuiti nel segmento di estremi P_0 e P_n. Allora la curva $P_0^n(t)$ generata dagli n punti dati è il segmento di estremi P_0 e P_n.

5. **Proprietà dell'inviluppo convesso.** $P_0^n(t)$ giace nell'inviluppo convesso dei punti $P_0, \ldots, P_n$, dove per inviluppo convesso si intende l'insieme delle combinazioni lineari

$$P = k_0 P_0 + \cdots + k_n P_n, \quad \sum_i k_i = 1.$$

Intuitivamente, si può pensare l'inviluppo convesso come il più piccolo poligono (o poliedro, se la curva non è piana) convesso che contiene i punti dati (anche non come vertici). Tale poligono (o poliedro) è detto *poligono (o poliedro) di controllo*. Le curve di Bézier, in un certo senso, "assomigliano" al poligono di controllo, perciò sono uno strumento adatto per disegnare curve. Più precisamente il calcolatore può disegnare le curve con una successione di operazioni matematiche semplici, senza usare l'espressione parametrica della curva, usando una procedura introdotta da de Casteljau che sfrutta la proprietà dell'inviluppo convesso. Infatti, dati i punti di controllo P_0, P_1, P_2, P_3, denotato con P_{ij} il punto medio tra P_i e P_j, si calcolano P_{01}, P_{12} e P_{23}, e successivamente i punti medi P_{012} tra P_{01} e P_{12}, e P_{123} tra P_{12} e P_{23} ed infine P_{0123} tra P_{012} e P_{123}. Il punto P_{0123} è sulla curva di Bézier definita dai punti P_0, P_1, P_2, P_3, come si verifica facilmente, e la curva risulta divisa in due parti dal punto P_{0123}, come si osserva dalla figura. Si può verificare che i punti di controllo P_0, P_{01}, P_{012}, P_{0123} determinano una nuova curva di Bézier che coincide con il ramo di sinistra della curva di Bézier determinata da P_1, P_2, P_3, P_4. Ripetendo questo procedimento si guinge rapidamente al limite della risoluzione grafica richiesta; il calcolatore ora può disegnare tutti i punti del tipo P_{0123} trovati con semplici operazioni di somma vettoriale e divisione per 2.

La proprietà dell'inviluppo convesso è estremamente utile anche per il *clipping* [18]: la rimozione di linee indesiderate, o la colorazione di certe aree del piano o dello spazio sono semplificate controllando prima se le aree giacciono all'interno o all'esterno del poliedro, controllo semplice da realizzare per un poliedro di tipo convesso. Un altro campo di utilizzo di questa proprietà è quello del controllo delle collisioni di oggetti in grafica 3D, con un meccanismo simile a quanto spiegato sopra.

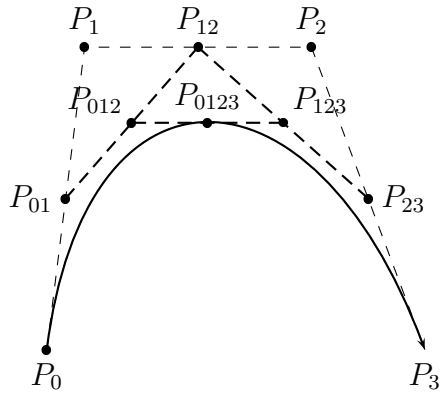


Figura 4.3. Suddivisioni di una curva di Bézier

6. **Invarianza per combinazioni baricentriche.** È chiaro che

$$\sum_{j=0}^{n}(\alpha P_j + \beta Q_j)B_j^n(t) = \alpha\sum_{j=0}^{n} P_j B_j^n(t) + \beta\sum_{j=0}^{n} Q_j B_j^n(t),$$

quindi se α e β sono coordinate baricentriche ($\alpha + \beta = 1$) nel poligono di Bézier, allora α e β risultano coordinate baricentriche per le curve corrispondenti.

In altre parole, noi possiamno costruire la media pesata di due curve di Bézier o facendo la media pesata dei corrispondenti punti sulla curva o facendo prima la media pesata dei corrispondenti valori di controllo, e poi trovare la curva.

7. **Controllo pseudo-locale.** Il polinomio di Bernstein B_i^n ha un solo massimo per $t = i/n$. Ciò ha applicazioni nel disegno: se si toglie solo il vertice P_i del poligono di controllo allora la curva risente in maniera sensibile di questo cambiamento solo intorno al valore i/n del parametro.

Esempio 4.2. Il seguente codice in PostScript^{TM} mostra come i caratteri PostScript^{TM} siano disegnati tramite curve di Bézier. Questo procedimento conserve di risparmiare moltissima memoria rispetto ai caratteri "bitmapped", poiché l'archivio del carattere conserva solo le informazioni relative ai punti di controllo.

```
  /Times-Roman findfont
   40 scalefont
   setfont
   newpath
   0 0 moveto
   (O) false charpath
%  stroke
  { [ 3 1 roll (moveto) ] == }
```

```
{ [ 3 1 roll (lineto) ] == }
{ [ 7 1 roll (curveto) ] == }
{ [ (closepath) ] == }
pathforall
```

Il codice dell'esempio fa produrre ad un interprete PostScript^{TM} (come Ghostscript) i punti di controllo delle curve di Bézier che delimitano i margini della lettera 'O' dal carattere Times-Roman, ingrandendolo a 40 punti tipografici. Se si vuole stampare la lettera sullo schermo si levi il carattere di commento di fronte all'istruzione `stroke`. Il risultato del precedente codice è il seguente:

```
[14.4401855 27.0002441 (moveto)]
[6.80029297 27.0002441 2.00024414 21.2570801 2.00024414 13.2399902 (curveto)]
[2.00024414 9.47998 3.04199219 5.84008789 4.80029297 3.52001953 (curveto)]
[7.12036133 1.0 10.6804199 -1.0 14.2004395 -1.0 (curveto)]
[22.0803223 -1.0 27.0002441 4.78076172 27.0002441 13.0800781 (curveto)]
[27.0002441 17.0400391 26.0197754 20.4399414 24.1604 22.8398438 (curveto)]
[21.6403809 25.5998535 18.2802734 27.0002441 14.4401855 27.0002441 (curveto)]
[(closepath)]
[14.4401855 25.0 (moveto)]
[16.2802734 25.0 18.1203613 24.4960938 19.5603027 23.5998535 (curveto)]
[21.7202148 21.6398926 23.0002441 17.8798828 23.0002441 13.119873 (curveto)]
[23.0002441 10.7597656 22.4643555 7.99975586 21.6401367 5.91967773 (curveto)]
[21.2800293 4.91967773 20.6000977 3.91967773 19.6801758 2.99975586 (curveto)]
[18.2802734 1.59985352 16.4802246 1.0 14.3601074 1.0 (curveto)]
[12.5200195 1.0 10.7199707 1.67602539 9.32006836 2.83984375 (curveto)]
[7.23999 4.67993164 5.99975586 8.7199707 5.99975586 13.1599121 (curveto)]
[5.99975586 17.2399902 7.08789062 21.119873 8.7199707 22.9997559 (curveto)]
[10.2800293 24.7197266 12.2402344 25.0 14.4401855 25.0 (curveto)]
[(closepath)]
[28.8798828 -0.0 (moveto)]
```

L'istruzione `x1 y1 x2 y2 x3 y3 curveto` disegna una curva di Bézier a partire dal punto corrente (ossia il punto attualmente nello *stack* di PostScript), con punti di controllo $(x1, y1)$ e $(x2, y2)$ e con secondo estremo $(x3, y3)$. L'istruzione `moveto` serve per cambiare il punto corrente. Le istruzioni contenenti `roll` servono per mostrare il contenuto dello *stack*.

Si noti che non tutti i caratteri rivelano i propri punti di controllo: Adobe prevede un meccanismo per la 'chiusura' dei caratteri in questo senso. Se si desidera approfondire il linguaggio PostScript^{TM} si consulti il sito `http://www.adobe.com/` per guide di riferimento, manuali e specifiche dei caratteri.

L'esempio è stato elaborato a partire dal testo [9], liberamente disponibile in Internet.

Esercizio 4.11. *Si dimostri che le curve di Bézier di grado* 2 *sono parabole.*

Esercizio 4.12. *Si provi che*

$$(B_i^n(t))' = n(B_{i-1}^{n-1}(t) - B_i^{n-1}(t)).$$

Esercizio 4.13. *Si provi che il punto P_{0123}, definito nel procedimento descritto dalla figura 4.3 giace nella curva di Bézier definita da P_0, P_1, P_2, P_3.*

CAPITOLO 5

I SISTEMI LINEARI

Sia $A = (a_{ij}) \in \mathbb{K}^{m,n}$ e $B \in \mathbb{K}^{m,1}$. Il problema di determinare gli $X \in \mathbb{K}^{n,1}$ tali che

$$AX = B \tag{5.0.1}$$

è detto *sistema lineare*.

Questo capitolo è un'introduzione ai metodi per calcolare *tutte* le soluzioni di un sistema lineare. Questi metodi sono detti *metodi diretti*.

I metodi considerati in questo capitolo sono essenzialmente metodi di fattorizzazione del sistema dato in due sistemi lineari più semplici (di solito a matrice triangolare).

5.1 Metodi di fattorizzazione: premesse

I metodi di fattorizzazione qui considerati hanno tutti origine dalla seguente osservazione: *se la matrice di un sistema è facilmente invertibile, il sistema è di facile risoluzione.* Infatti, si verificano i casi seguenti:

1. se $m = n$ ed A è invertibile, allora

$$AX = B \quad \Rightarrow \quad X = A^{-1}B;$$

2. se $m < n$ ed A ha rango m, è possibile, scambiando opportunamente le righe e le colonne di A, ricondursi ad una matrice $A' = (C\ P_1 \cdots P_{n-m})$ con $C \in \mathbb{K}^{m,m}$ invertibile. In questo sistema le incognite $x_{m+1},\ldots,x_n$ sono parametri, dunque la loro determinazione è arbitraria. Quindi si ottiene una sola soluzione $X' = (x_1, \ldots, x_m)$

$$(C\ P_1 \cdots P_{n-m})X = B \quad \Rightarrow \quad X' = C^{-1}(B - x_{m+1}P_1 - \cdots - x_nP_{n-m}) \tag{5.1.2}$$

per ogni valore dei parametri.

Il caso $m > n$ può non essere preso in considerazione poiché in ogni sistema di questo tipo ci sono sicuramente $m - n$ equazioni che sono combinazioni lineari delle altre e possono essere scartate.

Un caso di matrice facilmente invertibile è il caso di una matrice triangolare superiore (inferiore) $A \in \mathbb{K}^{n,n}$ tale che $\det A = a_{11} \cdots a_{nn} \neq 0$. Dunque, gli elementi diagonali non sono mai nulli, e, indicata con $Z \in \mathbb{K}^{n,n}$ la matrice inversa incognita, si può facilmente vedere che $AZ = I$ se e solo se Z è triangolare superiore (inferiore). Il sistema $AZ = I$ si risolve *ricorsivamente all'indietro (in avanti, con gli opportuni cambiamenti)*:

$$\begin{aligned} 1 &= a_{nn} z_{nn} \quad \Rightarrow \quad z_{nn} = \frac{1}{a_{nn}}; \\ &\dots \\ 0 &= \sum_{k=i}^{j} a_{ik} z_{kj} \quad \Rightarrow \quad z_{ij} = -\frac{1}{a_{ii}} \sum_{k=i+1}^{j} a_{ik} z_{kj}. \end{aligned}$$

Ovviamente, gli elementi vengono ricavati nell'ordine z_{nn}, $z_{n-1,n-1}$, $z_{n-1,n}$, ... Si può dimostrare che il numero di operazioni che si effettuano è dell'ordine di $n^3/6$ [5].

Analogamente, un sistema con matrice triangolare superiore (inferiore) invertibile può essere facilmente risolto *ricorsivamente all'indietro (in avanti, con gli opportuni cambiamenti)* come segue:

$$\begin{aligned} x_n &= \frac{b_n}{a_{nn}}, \\ &\dots \\ x_i &= \frac{1}{a_{ii}} \left(b_i - \sum_{j=i+1}^{n} a_{ij} x_j \right). \end{aligned} \tag{5.1.3}$$

I metodi studiati in questo capitolo sono diretti ad ottenere da un sistema arbitrario con matrice incompleta $A \in \mathbb{C}^{n,n}$ una coppia di sistemi la cui matrice sia facilmente invertibile, mediante la fattorizzazione di A come

$$A = CD \quad \Rightarrow \quad CDX = B \quad \Rightarrow \quad CY = B,\ DX = Y. \tag{5.1.4}$$

I due sistemi $CY = B$ e $DX = Y$ vengono poi risolti in successione con i metodi sopra descritti. Il campo scelto è quello complesso $\mathbb{C}$ perchè permette di descrivere la maggior parte dei problemi di natura ingegneristica[1]

Le fattorizzazioni che verranno studiate nei paragrafi seguenti sono:

1. la fattorizzazione $\mathcal{LU}$, dove L è unipotente inferiore ed U è triangolare superiore[2]; essa è associata al metodo di Gauss.

2. la fattorizzazione $\mathcal{QR}$, dove Q è una matrice unitaria ed R è una matrice triangolare superiore; essa è associata al metodo di Householder.

[1] Non tutti: ad esempio, in crittografia hanno notevole importanza i cosiddetti *campi finiti* $\mathbb{Z}_p$ (p numero primo), che hanno un numero finito p di elementi.

[2] L viene dall'inglese *lower triangular*, ed U da *upper triangular*.

3. la fattorizzazione $\mathcal{LL}^*$, dove L è una matrice triangolare inferiore con $a_{ii} > 0$ per $i = 1, \ldots, n$; essa è associata al metodo di Cholesky.

L'idea è quella di raggiungere la fattorizzazione cercata mediante trasformazioni elementari che non cambiano le soluzioni del sistema. Tali trasformazioni sono realizzabili come particolari matrici, e concorreranno alla formazione della fattorizzazione.

Verranno ora analizzate le condizioni di esistenza ed unicità delle precedenti fattorizzazioni.

Teorema 5.1. *Sia $A \in \mathbb{C}^{n,n}$. Si supponga che i minori principali di A siano tutti invertibili, esclusa al più A; in altre parole,* $\det A_k \neq 0$ *per* $1 \leq k \leq n-1$. *Allora esiste un'unica coppia di matrici L ed U, con L unipotente inferiore ed U triangolare superiore, tale che*

$$A = LU.$$

Dimostrazione. Si proceda per induzione. Per $n = 1$ basta porre $L = (1)$ ed $U = (a_{11})$. Sia valido il teorema per $n = k-1$, e si consideri il caso $n = k$. Posto

$$A_k = \begin{pmatrix} A_{k-1} & D \\ C^* & \alpha \end{pmatrix},$$

con $\alpha \in \mathbb{C}$ e D, $C \in \mathbb{C}^{n,1}$. sia $A_{k-1} = L_{k-1}U_{k-1}$. Si ponga

$$L_k \stackrel{\text{def}}{=} \begin{pmatrix} L_{k-1} & O \\ U^* & 1 \end{pmatrix}, \qquad U_k \stackrel{\text{def}}{=} \begin{pmatrix} U_{k-1} & V \\ O & \beta \end{pmatrix},$$

dove $\beta \in \mathbb{C}$ e U, $V \in \mathbb{C}^{n,1}$ sono quantità incognite. Allora

$$A_k = L_k U_k \quad \Leftrightarrow \quad \begin{cases} D = L_{k-1}V, \\ C^* = U^* U_{k-1}, \\ \alpha = U^* V + \beta, \end{cases}$$

da cui risultano U e V univocamente determinate poichè L_{k-1} e U_{k-1} sono invertibili, e $\beta = \alpha - U^*V$. QED

Il metodo di Gauss è stato applicato in [7] anche a matrici che non soddisfacevano le ipotesi del precedente teorema. In realtà, vale il seguente risultato, che generalizza il precedente.

Teorema 5.2. *Sia $A \in \mathbb{C}^{n,n}$. Allora, esiste una matrice di permutazione Π tale che la matrice ΠA ammette una fattorizzazione $\mathcal{LU}$.*

Dimostrazione. L'affermazione sarà qui dimostrata solo per il caso in cui $\mathrm{rg}(A) = n-1$. In questo caso è possibile dimostrare che esiste una matrice di permutazione Π tale che ΠA soddisfa le ipotesi del teorema 5.1. Infatti, esiste sicuramente in A una sottomatrice $M' \in \mathbb{C}^{n-1,n-1}$ tale che $\det M' \neq 0$. Dunque, esiste una matrice di permutazione Π' tale che $(\Pi' A)_{n-1} = M'$. Continuando, esiste una sottomatrice $M'' \in \mathbb{C}^{n-2,n-2}$

di $(\Pi' A)_{n-1}$ tale che $\det M'' \neq 0$: se questo non si verificasse, un qualunque sviluppo di Laplace di $\det(\Pi' A)_{n-1}$ sarebbe uguale a 0. Ragionando come prima, ed applicando induttivamente il ragionamento si ottiene la tesi. QED

Si noti che la fattorizzazione $\mathcal{LU}$ nel caso generale ha poco interesse per le osservazioni fatte all'inizio di questo paragrafo (5.1.2).

Più in là (paragrafi 5.8 e 5.11) si dimostrerà, con metodi costruttivi, che la fattorizzazione $\mathcal{QR}$ esiste qualunque sia la matrice di partenza e che la fattorizzazione LL^* esiste ed è unica sotto certe ipotesi.

Per quanto riguarda l'unicità della fattorizzazione $\mathcal{QR}$, se $S \in \mathbb{C}^{n,n}$ è diagonale ed unitaria (si veda il paragrafo 1.5), allora

$$QR = QSS^*R,$$

dove QS è unitaria e S^*R è triangolare superiore. Quindi, la decomposizione non è univocamente determinata. Se la matrice A è non singolare, le fattorizzazioni $\mathcal{QR}$ sono definite *esattamente* a meno di fattori di fase.

Teorema 5.3. *Sia $A \in \mathbb{C}^{n,n}$ tale che $\det A \neq 0$. Allora, se A ammette due fattorizzazioni $\mathcal{QR}$: $A = QR = Q'R'$, esiste una matrice di fase $S \in \mathbb{C}^{n,n}$ tale che*

$$Q' = QS, \qquad R' = S^*R.$$

Dimostrazione. Infatti, si ha

$$Q^*Q' = RR'^{-1},$$

dunque la matrice unitaria Q^*Q' è triangolare superiore, quindi è diagonale. Si ponga $S = Q^*Q'$. QED

Corollario 5.1. *Sia $A \in \mathbb{C}^{n,n}$ tale che $\det A \neq 0$. Allora esiste un'unica fattorizzazione $\mathcal{QR}$ $A = QR$ tale che $r_{ii} > 0$.*

5.2 Trasformazioni sulle matrici

In questo paragrafo verranno descritte le trasformazioni che conducono alle fattorizzazioni del paragrafo precedente.

Sia $\tilde{A} \in \mathbb{C}^{m,n}$. Denotiamo con $\tilde{A} = (R_i)$ (risp. $\tilde{A} = (C_j)$), per $i = 1, \ldots, m$ (risp. $j = 1, \ldots, n$), il partizionamento di $\tilde{A}$ nelle sue righe (o colonne). Il seguente teorema è di facile dimostrazione.

Teorema 5.4. *Siano $A \in \mathbb{C}^{m,n}$, $B \in \mathbb{C}^{n,1}$. Allora ognuna delle seguenti operazioni sulla matrice $\tilde{A} = (A \vdots B)$*

1. *scambio di due righe;*

2. moltiplicazione di una riga per uno scalare non nullo;

3. sostituzione di una riga R_i con la somma $R_i + \lambda R_h$, dove $i \neq h$ e $\lambda \in \mathbb{C}$.

trasforma la matrice in una matrice completa di un sistema lineare equivalente al primo.

Osservazione 5.1. Il risultato precedente non vale per le colonne, in virtù del diverso significato che hanno le righe e le colonne della matrice di un sistema. Ma lo scambio delle due colonne j e j' nella matrice incompleta dà un sistema lineare che ha soluzioni dove la j-esima componente è scambiata con la j'-esima rispetto al sistema di partenza. □

Quindi, le trasformazioni che vengono usate concretamente sono le trasformazioni del teorema 5.4 sulle righe e lo scambio di colonne nella matrice incompleta (tenendo conto dell'osservazione precedente).

Le trasformazioni precedenti sono realizzate madiante una moltiplicazione per opportune matrici, come segue. Sia $\mathcal{C} = \{\vec{e}_1, \dots, \vec{e}_n\}$ la base canonica di $\mathbb{C}^n$.

1. Lo scambio di righe di A è dato da $\Pi_{jj'}A$, dove $\Pi_{jj'} \stackrel{\text{def}}{=} \mathcal{M}_{\mathcal{C}}^{\mathcal{C}}(f)$, essendo

$$f\colon \mathbb{C}^n \to \mathbb{C}^n, \quad (\vec{e}_1, \dots, \vec{e}_j, \dots, \vec{e}_{j'}, \dots \vec{e}_n) \mapsto (\vec{e}_1, \dots, \vec{e}_{j'}, \dots, \vec{e}_j, \dots \vec{e}_n). \tag{5.2.5}$$

   Si noti che lo scambio di colonne di A è dato da $A\Pi_{jj'}$, dove $\Pi_{jj'}$ è uno scambio di righe.

2. La moltiplicazione di una riga per $\lambda \in \mathbb{C} \setminus \{0\}$ è data da $M_i(\lambda)A$, dove si pone $M_i(\lambda) \stackrel{\text{def}}{=} \mathcal{M}_{\mathcal{C}}^{\mathcal{C}}(g)$, essendo

$$g\colon \mathbb{C}^n \to \mathbb{C}^n, \quad (\vec{e}_1, \dots, \vec{e}_j, \dots \vec{e}_n) \mapsto (\vec{e}_1, \dots, \lambda\vec{e}_j, \dots \vec{e}_n). \tag{5.2.6}$$

3. La sostituzione di una riga con una somma (teorema 5.4) è data da $S_{ij}(\lambda)A$, dove $S_{ij}(\lambda) \stackrel{\text{def}}{=} \mathcal{M}_{\mathcal{C}}^{\mathcal{C}}(h)$, essendo

$$h\colon \mathbb{C}^n \to \mathbb{C}^n, \quad (\vec{e}_1, \dots, \vec{e}_n) \mapsto (\vec{e}_1, \dots, \vec{e}_i + \lambda\vec{e}_j, \dots \vec{e}_n). \tag{5.2.7}$$

Definizione 5.1. Le matrici delle applicazioni lineari del tipo di f (5.2.5) sono dette *matrici di scambio*. Il prodotto di matrici di scambio dà le *matrici di permutazione*.

Esempio 5.1.

1. Una matrice $\Pi_{jj'}$ di scambio di righe ha l'espressione

$$\Pi_{jj'} = \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix}, \tag{5.2.8}$$

in altri termini, $\Pi_{jj'}$ è ottenuta da I scambiandone le colonne j e j'. Le matrici di permutazione sono tutte e sole le matrici che rappresentano le applicazioni lineari del tipo

$$(\vec{e}_1, \ldots, \vec{e}_n) \mapsto (\vec{e}_{\sigma_1}, \ldots, \vec{e}_{\sigma_n}),$$

e sono ottenute permutando le colonne (o le righe) di I.

2. Una matrice $M_i(\lambda)$ di moltiplicazione di una riga per uno scalare ha la forma

$$M_i(\lambda) = \begin{pmatrix} 1 & 0 & \ldots\ldots\ldots & 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots & & & \\ 0 & \ldots & 1 & 0 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & \lambda & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 0 & 1 & \ldots & 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots & & & \\ 0 & \ldots\ldots\ldots\ldots\ldots & & 1 \end{pmatrix}. \tag{5.2.9}$$

3. Una matrice $S_{ij}(\lambda)$ di sostituzione di una riga con una somma ha la forma

$$S_{ij}(\lambda) = \begin{pmatrix} 1 & \ldots & 0 & \ldots\ldots\ldots & 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots & & & & \\ 0 & \ldots & 1 & \ldots\ldots\ldots & 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots & & & & \\ 0 & \ldots & \lambda & \ldots & 1 & \ldots & 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots & & & & \\ 0 & \ldots & 0 & \ldots\ldots\ldots & 1 \end{pmatrix} \tag{5.2.10}$$

se $i < j$. Invece, se $i > j$ il coefficiente λ si trova 'al di sopra' della diagonale.

Esercizio 5.1. *Presa la seguente matrice A*

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ -3 & 2 & 8 & 5 \\ 2 & 0 & -1 & 2 \end{pmatrix},$$

realizzare la somma della prima riga moltiplicata per -2 con la terza calcolando la matrice $S_{31}(-2)$ ed eseguendo $S_{31}(-2)A$.

5.3 Matrici elementari

In pratica, ognuna delle fattorizzazioni (paragrafo 5.1) della matrice incompleta A di un sistema lineare viene raggiunta mediante la moltiplicazione successiva per matrici che rappresentano trasformazioni che non cambiano le soluzioni del sistema. Le matrici che vengono usate sono opportune composizioni delle matrici precedenti. A questo scopo, introduciamo una vasta classe di matrici, comprendenti le precedenti.

Definizione 5.2. Siano $\sigma \in \mathbb{C}$ ed X, $Y \in \mathbb{C}^{n,1} \setminus \{O\}$. La matrice

$$E(\sigma, X, Y) \stackrel{\text{def}}{=} I - \sigma X Y^* \in \mathbb{C}^{n,n}$$

si dice *matrice elementare*.

Si denoti con $\mathbb{E}^{n,n} \subset \mathbb{C}^{n,n}$ l'insieme delle matrici elementari.

Una interessante proprietà delle matrici elementari è la seguente.

Proposizione 5.1. *Siano σ, $\tau \in \mathbb{C}$, ed X, $Y \in \mathbb{C}^{n,1}$. Allora vale*

$$E(\sigma, X, Y)E(\tau, X, Y) = E(\gamma, X, Y)$$

dove $\gamma = \sigma + \tau - \sigma\tau(Y^*X)$.

La dimostrazione è lasciata al lettore. Tale proprietà afferma che il prodotto di matrici elementari è una matrice elementare. Viceversa, è sempre possibile decomporre una matrice elementare come prodotto di matrici elementari in più modi. Ad esempio, dato $\gamma \in \mathbb{C}$ e scelto $\sigma \in \mathbb{C}$ tale che $\sigma(Y^*X) \neq 1$, si ha

$$\tau = \frac{\gamma - \sigma}{1 - \sigma(Y^*X)}.$$

Naturalmente, se $E(\gamma, X, Y)$ è invertibile, lo sono anche i fattori.

Una conseguenza immediata della precedente proprietà è la seguente.

Proposizione 5.2. *Una matrice elementare $E(\sigma, X, Y) \in \mathbb{E}^{n,n}$ è invertibile se e solo se $\sigma Y^*X \neq 1$.*

Dimostrazione. Per $\sigma = 0$ tutte le matrici elementari valgono I.

Per $\sigma \neq 0$, se la condizione $\sigma Y^*X \neq 1$ è soddisfatta, applicando la precedente proposizione si trova un $\tau \in \mathbb{C}$ tale che $E(\tau, X, Y) = E(\sigma, X, Y)^{-1}$. Viceversa, usando l'esercizio 1.12 si vede che le matrici $E(\sigma, X, Y)$ tali che $\sigma X^*Y = 1$ hanno determinante nullo, quindi non sono invertibili. QED

Esercizio 5.2. *Trovare τ nella precedente dimostrazione.*

Proposizione 5.3. *Siano U, $V \in \mathbb{C}^n$ con $U \neq O$, $V \neq O$. Allora esiste $E(\sigma, X, Y) \in \mathbb{E}^{n,n}$ tale che*

1. *$E(\sigma, X, Y)U = V$;*
2. *$\det E(\sigma, X, Y) \neq 0$.*

Dimostrazione. La condizione è verificata scegliendo un qualunque Y in modo che $Y^*U \neq 0$. Infatti, in tal caso si possono scegliere σ ed X in modo che

$$\sigma X = \frac{U - V}{Y^*U}.$$

Inoltre, se Y è scelto in modo che anche $Y^*V \neq 0$, la matrice $E(\sigma, X, Y)$ ha determinante diverso da zero *(dimostrarlo!)*. QED

Proposizione 5.4. *Le matrici delle trasformazioni descritte nel teorema 5.4 sono matrici elementari.*

DIMOSTRAZIONE. Si ha:

1. $\Pi_{jj'} = E(1, X, X)$ dove $X = (0, \ldots, \underset{j}{1}, 0 \ldots, \underset{j'}{-1}, 0, \ldots 0)^*$.
2. $M_i(\lambda) = E(1 - \lambda, E_i, E_i)$, dove E_i è l'i-esimo elemento della base dello spazio dei vettori colonna.
3. $S_{ij}(\lambda) = E(-\lambda, E_i, E_j)$.

QED

5.4 Fattorizzazione mediante matrici elementari

Sia $A \in \mathbb{C}^{n,n}$. Usando la proposizione 5.3 è possibile fattorizzare A come $A = EU$, dove E è un prodotto di matrici elementari ed U è una matrice triangolare superiore. La dimostrazione di quest'affermazione è costruttiva e consiste nel seguente algoritmo.

Passo 1**.** Si pone $A^{(1)} \stackrel{\text{def}}{=} A$, e si partiziona $A^{(1)}$ a blocchi così:

$$A^{(1)} = \begin{pmatrix} C^{(1)} & D^{(1)} \\ T & F^{(1)} \end{pmatrix}, \qquad C^{(1)} \in \mathbb{C},\ F^{(1)} \in \mathbb{C}^{n-1,n-1},\ D^{(1)} \in \mathbb{C}^{1,n-1},\ T \in \mathbb{C}^{n-1,1}.$$

Per la proposizione 5.3, esiste una matrice elementare $E^{(1)}$ tale che

$$E^{(1)} \begin{pmatrix} C^{(1)} \\ T \end{pmatrix} = \begin{pmatrix} C^{(2)} \\ O \end{pmatrix},$$

dove $C^{(2)} \in \mathbb{C}$. Risulta

$$E^{(1)} A^{(1)} = \begin{pmatrix} C^{(2)} & D^{(2)} \\ O & F^{(2)} \end{pmatrix},$$

dove $D^{(2)} \in \mathbb{C}^{1,n-1}$ e $F^{(2)} \in \mathbb{C}^{n-1,n-1}$ sono opportune matrici. Si ponga $A^{(2)} \stackrel{\text{def}}{=} E^{(1)} A^{(1)}$.

Passo k**.** Sia

$$A^{(k)} = \begin{pmatrix} C^{(k)} & D^{(k)} \\ O & F^{(k)} \end{pmatrix},$$

con $C^{(k)} \in \mathcal{T}_U^{k-1,k-1}$, $D^{(k)} \in \mathbb{C}^{k-1,n-k+1}$, $F^{(k)} \in \mathbb{C}^{n-k+1,n-k+1}$. Allora, per la proposizione 5.3, esiste una matrice elementare $\tilde{E}$ tale che

$$\tilde{E} F^{(k)} = \begin{pmatrix} \alpha & T \\ O & F^{(k+1)} \end{pmatrix},$$

dove $\alpha \in \mathbb{C}$, $T \in \mathbb{C}^{(1,n-k)}$, $F^{(k+1)} \in \mathbb{C}^{n-k,n-k}$. Si ponga

$$E^{(k)} \stackrel{\text{def}}{=} \begin{pmatrix} I & O \\ O & \tilde{E} \end{pmatrix};$$

la matrice $E^{(k)}$ è una matrice elementare. Per rendersi conto di questo, si osservi che, se $\tilde{E} = E(\sigma, \tilde{X}, \tilde{Y})$, si ha $E^{(k)} = E(\sigma, X, Y)$, dove $X = (O\ \tilde{X})^T \in \mathbb{C}^{n,1}$ ed $Y = (O\ \tilde{Y})^T \in \mathbb{C}^{n,1}$. Si ponga $A^{(k+1)} \stackrel{\text{def}}{=} E^{(k)} A^{(k)}$; si ottiene

$$A^{(k+1)} = \begin{pmatrix} C^{(k+1)} & D^{(k+1)} \\ O & F^{(k+1)} \end{pmatrix},$$

con $C^{(k+1)} \in \mathcal{T}_U^{k,k}$, $D^{(k+1)} \in \mathbb{C}^{k,n-k}$, $F^{(k+1)} \in \mathbb{C}^{n-k,n-k}$.

Al passo $n-1$ si ottiene una matrice $A^{(n)}$ triangolare superiore tale che

$$A^{(n)} = E^{(n-1)} \cdots E^{(1)} A^{(1)} \quad \Rightarrow \quad A = E A^{(n)},$$

dove si è posto $E \stackrel{\text{def}}{=} \left(E^{(1)}\right)^{-1} \cdots \left(E^{(n-1)}\right)^{-1}$. La matrice E può assumere forme diverse a seconda delle particolari matrici elementari scelte per realizzare l'algoritmo. Nei prossimi paragrafi verranno introdotte due particolari classi di matrici elementari: le matrici elementari di Gauss e le matrici elementari di Householder.

5.5 Il metodo di Gauss

Il *metodo di Gauss* è stato introdotto in [16]. Si può dimostrare [25] che il metodo di Gauss è il metodo che comporta il minor numero di calcoli tra i metodi che utilizzano esclusivamente combinazioni di righe e di colonne.

L'obbiettivo di questo paragrafo è realizzare il metodo di Gauss come fattorizzazione $\mathcal{LU}$, ottenendo questa fattorizzazione mediante successive moltiplicazioni per matrici elementari. A questo scopo, si introduce una sottoclasse delle matrici elementari.

Proposizione 5.5. *Sia $U \in \mathbb{C}^{n,1}$ tale che $u_i \neq 0$ per un certo $i \in \{1, \ldots, n\}$. Si ponga $U_i \stackrel{def}{=} u_1 E_1 + \cdots + u_i E_i$, dove $\{E_k\}_{1 \leq k \leq n}$ sono le matrici della base canonica di $\mathbb{C}^{n,1}$. Allora, esiste un'unica matrice elementare $E(\sigma, X, E_i)$ tale che*

$$E(\sigma, X, E_i) U = U_i = \begin{pmatrix} u_1 \\ \vdots \\ u_i \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

DIMOSTRAZIONE. Dalla dimostrazione della proposizione 5.3, posto $V = U_i$, ed essendo $E_i^* U = u_i \neq 0$, si ha

$$\sigma X = \left(0, \dots, 0, \frac{u_{i+1}}{u_i}, \dots, \frac{u_n}{u_i}\right)^T,$$

dunque

$$E(\sigma, X, E_i) = \begin{pmatrix} 1 & \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & \dots & 0 \\ 0 & \dots & -u_{(i+1)}/u_i & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & -u_n/u_i & \dots & 1 \end{pmatrix}$$

QED

Definizione 5.3. Nelle ipotesi della precedente proposizione, la matrice elementare determinata univocamente) si dice *matrice elementare di Gauss*.

È facile dimostrare che $E(\sigma, X, E_i) = S_{i+1,i}(-u_{i+1}/u_i) \cdots S_{n,i}(-u_n/u_i)$. Da $\sigma E_i^* X = 0$ risulta che la matrice è invertibile e

$$E(\sigma, X, E_i)^{-1} = \begin{pmatrix} 1 & \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & \dots & 0 \\ 0 & \dots & u_{(i+1)}/u_i & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & u_n/u_i & \dots & 1 \end{pmatrix}.$$

In particolare, le matrici elementari di Gauss sono unipotenti inferiori.

Esempio 5.2. Sia $U = (2, 2, 3)^T \in \mathbb{C}^{3,1}$. Allora $u_1 = 2 \neq 0$, $\sigma X = (0, 1, 3/2)^T$ e

$$E(\sigma, X, E_1) = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -3/2 & 0 & 1 \end{pmatrix}.$$

Teorema 5.5. *Sia $A \in \mathbb{C}^{n,n}$. Se A soddisfa le ipotesi del teorema 5.1, allora l'unica fattorizzazione $\mathcal{LU}$ si può ottenere mediante l'algoritmo descritto nel paragrafo 5.4, e risulta*

$$L = \left(E^{(1)}\right)^{-1} \cdots \left(E^{(n-1)}\right)^{-1}, \qquad U = A^{(n)}.$$

DIMOSTRAZIONE. Si indichi con $a_{ij}^{(k)}$ l'elemento generico della matrice $A^{(k)}$. La ma-

trice elementare di Gauss $E^{(k)}$ ha la forma

$$E^{(k)} = E(\sigma, X, E_k) = \begin{pmatrix} 1 & \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & \dots & 0 \\ 0 & \dots & -a^{(k)}_{k+1k}/a^{(k)}_{kk} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & -a^{(k)}_{nk}/a^{(k)}_{kk} & \dots & 1 \end{pmatrix},$$

dove σ ed X sono tali che $E(\sigma, X, E_k)C_k = (C_k)_k$ (qui, C_k sta per la colonna k-esima di $A^{(k)}$, e $(C_k)_k$ è il vettore ottenuto da C_k come nella dimostrazione della proposizione 5.5). Questo ha senso solo se $a^{(k)}_{kk} \neq 0$. Per $k = 1$ questo è vero per ipotesi. Al passo k il minore principale $(A^{(k)})_k$ di $A^{(k)}$ è tale che $\det(A^{(k)})_k = \det C^{(k)} a^{(k)}_{kk}$. Ma $\det(A^{(k)})_k \neq 0$ poiché il determinante delle matrici elementari di Gauss è 1 e

$$\begin{aligned} \det(A^{(k)})_k &= \det\left(E^{(k-1)} \cdots E^{(1)} A\right)_k \\ &= \det\left((E^{(k-1)})_k \cdots (E^{(1)})_k A_k\right) \\ &= \det A_k, \end{aligned}$$

da cui la tesi. QED

Osservazione 5.2. Si noti che la matrice L *non è*, in generale, una matrice elementare: le matrici $E^{(k)}$ ed $E^{(k+1)}$ sono elementari ma definite tramite coppie diverse di vettori.

L'algoritmo presentato nel paragrafo 5.4 usato con le matrici elementari di Gauss al fine di risolvere un sistema lineare $AX = B$ prende il nome di *metodo di Gauss*. Ovviamente, il metodo di Gauss è completato dalla risoluzione dei due sistemi triangolari in cui è fattorizzato il sistema $AX = B$ (paragrafo 5.1, equazione (5.1.4)). Di questi sistemi triangolari, il sistema $LY = B$ è risolto durante la fattorizzazione, applicando l'algoritmo sopra descritto alla matrice completa $(A|B)$, ed il sistema $UX = Y$ può essere agevolmente risolto con il metodo dato con l'equazione 5.1.3 (si veda l'esempio 5.3).

È possibile determinare il numero di operazioni compiute durante il metodo di Gauss: esso è dell'ordine (asintotico) di $n^3/3$ [2].

Osservazione 5.3. Risulta

$$L = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ a^{(1)}_{21}/a^1_{11} & 1 & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & \dots & 0 \\ \vdots & \vdots & \vdots & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & a^{(k)}_{k+1k}/a^{(k)}_{kk} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ a^{(1)}_{n1}/a^1_{11} & \vdots & \vdots & a^{(k)}_{nk}/a^{(k)}_{kk} & \vdots & 1 \end{pmatrix}.$$

Si noti che il prodotto delle matrici elementari di Gauss che formano L non è commutativo: $E^{(1)}E^{(2)} \neq E^{(2)}E^{(1)}$. □

Esempio 5.3. Si consideri il sistema lineare $AX = B$, dove

$$A = \begin{pmatrix} 5 & 0 & 2 & 0 \\ 0 & 5 & 0 & 2 \\ 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 12 \\ 9 \\ 5 \\ 4 \end{pmatrix}.$$

È conveniente applicare le matrici elementari anche al termine noto. Dunque, risulta

$$E^{(1)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2/5 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (A^{(2)}|B^{(2)}) = \begin{pmatrix} 5 & 0 & 2 & 0 & 12 \\ 0 & 5 & 0 & 2 & 9 \\ 0 & 0 & 1/5 & 0 & 1/5 \\ 0 & 2 & 0 & 1 & 4 \end{pmatrix}$$

$$E^{(2)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -2/5 & 0 & 1 \end{pmatrix}, \quad (A^{(3)}|B^{(3)}) = \begin{pmatrix} 5 & 0 & 2 & 0 & 12 \\ 0 & 5 & 0 & 2 & 9 \\ 0 & 0 & 1/5 & 0 & 1/5 \\ 0 & 0 & 0 & 1/5 & 2/5 \end{pmatrix}$$

$$E^{(3)} = I, \qquad (A^{(4)}|B^{(4)}) = (A^{(3)}|B^{(3)}).$$

Dunque

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 2/5 & 0 & 1 & 0 \\ 0 & 2/5 & 0 & 1 \end{pmatrix}, \quad U = A^{(4)} = A^{(3)}.$$

Il sistema $UX = B^{(4)}$ così ottenuto si può risolvere "all'indietro" usando il metodo (5.1.3).

Osservazione 5.4. Il metodo di Gauss può essere applicato anche a più sistemi lineari simultaneamente purché abbiano la stessa matrice incompleta. In questo caso si procede come sopra, usando al posto del vettore B una matrice $B \in \mathbb{C}^{n,s}$. Questo è utile nel caso del calcolo dell'inversa, dove il termine noto prende la forma $B = I$ [7]. □

Osservazione 5.5 (Metodo del *pivot*). Sia $A \in \mathbb{C}^{n,n}$ tale che $\det A \neq 0$. Allora una fattorizzazione $\Pi A = LU$ è ottenibile modificando il metodo di Gauss come segue. Al primo passo, esiste nella prima colonna di A un elemento non nullo; si scambi la corrispondente riga con la prima riga di A e si applichi il primo passo del metodo di Gauss. Al passo k si proceda allo stesso modo, operando gli scambi sempre sulla matrice di partenza. □

Esempio 5.4. Si consideri la matrice

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 2 & -1 \\ -1 & 1 & 1 \end{pmatrix}.$$

Si osserva subito che la matrice non soddisfa le ipotesi del teorema 5.1. Risulta:

$$E^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad A^{(2)} = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 0 & -5 \\ 0 & 2 & 3 \end{pmatrix}.$$

È necessario scambiare le ultime due righe. Risulta $\Pi_{23}A = LU$, dove

$$L = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 3 \\ 0 & 0 & -5 \end{pmatrix}.$$

Si noti che la matrice L è ottenuta scambiando i coefficienti relativi alle righe 2 e 3 nelle corrispondenti matrici elementari.

Infine, si noti che scambiando le prime due righe di A si ottiene una matrice che soddisfa le ipotesi del teorema 5.1. Tuttavia, la fattorizzazione $\mathcal{LU}$ ottenuta da $\Pi_{12}A$ è diversa da quella precedente (verificarlo!).

Osservazione 5.6. Il metodo di Gauss può anche essere applicato a sistemi per i quali la matrice incompleta è del tipo $A \in \mathbb{C}^{m,n}$, con $m \geq n$ e $\mathrm{rg}(A) = n$. In questo caso, effettuando un ulteriore passo del metodo si arriva ad una matrice $A^{(n+1)}$ tale che $(A^{(n+1)})_{ij} = 0$ per $i > j$. Le righe $n+1, \ldots, m$ di $A^{(n+1)}$ sono tutte nulle, quindi possono essere scartate, e si ottiene la fattorizzazione $\mathcal{LU}$ della sottomatrice di A (o di una sua permutazione secondo l'osservazione 5.5) contenente le prime n righe. □

Osservazione 5.7. Il metodo di Gauss può essere usato per calcolare il rango di una matrice $A \in \mathbb{C}^{m,n}$. Infatti, se al passo k si incontra una colonna nulla, si va al passo successivo senza calcolare la matrice elementare relativa. Con un numero di passi uguale od inferiore a $\min\{m, n\}$ si ottiene una matrice a scalini. □

Osservazione 5.8. Una variante molto nota del metodo di Gauss è il *metodo del massimo pivot parziale*. Questo è una modifica del metodo di Gauss, e consiste nella selezione, al passo k, della riga con l'elemento di massimo modulo in $F^{(k)}$. Ovviamente questo metodo consente di ottenere una matrice L tale che $\|L\|_\infty = 1$. Per maggiori informazioni sull'analisi dell'errore si consulti [5, 13, 22, 28, 29].

5.6 Caso particolare: le matrici tridiagonali

Esistono problemi nei quali è necessario risolvere un sistema lineare $AX = B$ dove A è una matrice tridiagonale. Questo capita nei metodi per la ricerca degli autovalori di una matrice (come il metodo $\mathcal{QR}$), e in molti metodi per la soluzione di sistemi di equazioni differenziali discretizzate. Questi sistemi si risolvono velocemente con la fattorizzazione $\mathcal{LU}$.

Proposizione 5.6. *Sia $A \in \mathbb{C}^{n,n}$ una matrice tridiagonale soddisfacente alle condi-*

zioni del teorema 5.1. Allora L ed U hanno la forma

$$L = \begin{pmatrix} 1 & 0 & \dots & 0 \\ \beta_1 & 1 & \dots & 0 \\ & \ddots & & \\ 0 & \dots & \beta_{n-1} & 1 \end{pmatrix}, \qquad U = \begin{pmatrix} \alpha_1 & a_{12} & \dots & 0 \\ & \ddots & & \\ 0 & \dots & \alpha_{n-1} & a_{n-1n} \\ 0 & \dots & 0 & \alpha_n \end{pmatrix}$$

Dimostrazione. La dimostrazione è costruttiva, e consiste nel confronto tra i coefficienti di LU e di A. Si ottiene il sistema

$$\begin{cases} \alpha_1 = a_1, \\ \beta_i \alpha_i = a_{(i+1)i}, \quad i = 1, \dots, n-1, \\ \beta_i a_{i(i+1)} + \alpha_{i+1} = a_{(i+1)(i+1)}, \quad i = 1, \dots, n-1. \end{cases}$$

Nel sistema si risolve prima la seconda equazione e poi la terza per $i = 1, \dots, n-1$. È necessaria la condizione $\alpha_i \neq 0$ per risolvere la seconda equazione. In realtà si vede subito che

$$\det A = \det L \det U = \alpha_1 \cdots \alpha_n,$$

dunque la condizione $\alpha_i \neq 0$ per $i = 1, \dots, n$ equivale alla condizione sui minori di A richiesta dal teorema 5.1. QED

Si noti che il numero di operazioni necessario alla fattorizzazione $\mathcal{LU}$ è $2n-2$ in questo caso. Si noti, inoltre, che la verifica $\alpha_i \neq 0$ può essere eseguita durante il calcolo.

Esempio 5.5. Sia

$$A = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 3 & 5 & -3 & 0 \\ 0 & 4 & 14 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}.$$

Risulta, seguendo i calcoli della precedente dimostrazione,

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 0 & -4 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad U = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & -1 & -3 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}.$$

5.7 Il metodo di Gauss–Jordan

Il metodo di Gauss–Jordan[3] è una variante del metodo di Gauss nella quale al passo k si annullano *tutti* gli elementi sulla colonna $k+1$ tranne l'elemento sulla diagonale. Tramite questo metodo si ottiene una fattorizzazione $A = ED$, dove E non è una matrice unipotente inferiore come nel metodo di Gauss, ma D è una matrice diagonale.

[3] Wilhelm Jordan, 1842–1899

Proposizione 5.7. *Sia $U \in \mathbb{C}^{n,1}$ tale che $u_i \neq 0$ per un certo $i \in \{1, \ldots, n\}$. Allora, esiste un'unica matrice elementare $E(\sigma, X, E_i)$ tale che*

$$E(\sigma, X, E_i)U = u_i E_i = \begin{pmatrix} 0 \\ \vdots \\ u_i \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

La dimostrazione è simile a quella della proposizione 5.5. Le matrici elementari del tipo della precedente proposizione sono dette *matrici elementari di Gauss–Jordan*.

Ora, si consideri l'algoritmo descritto nel paragrafo 5.4, e si utilizzi come matrice $E^{(k)}$ la matrice elementare di Gauss–Jordan $E(\sigma, X, E_{k+1})$ che trasforma la colonna k-esima C_{k+1} di $A^{(k)}$ nel vettore che ha la sola componente non nulla $a_{kk}^{(k)}$ al posto k. Si ha

$$E^{(k)} = \begin{pmatrix} 1 & 0 & \ldots & -a_{1k}^{(k)}/a_{kk}^{(k)} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & -a_{k-1k}^{(k)}/a_{kk}^{(k)} & \ldots & 0 \\ 0 & \ldots & \ldots & 1 & \ldots & 0 \\ 0 & \ldots & \ldots & -a_{k+1k}^{(k)}/a_{kk}^{(k)} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & -a_{nk}^{(k)}/a_{kk}^{(k)} & \ldots & 1 \end{pmatrix}.$$

La matrice $A^{(k+1)}$ è tale che il blocco $C^{(k+1)}$ è diagonale.

Le condizioni di fattorizzabilità sono le stesse descritte nel teorema 5.1 e nell'osservazione 5.5. Il sistema diagonale si risolve con una versione semplificata del metodo usato per i sistemi triangolari. Il numero di operazioni effettuato, approssimabile con $n^3/2$ [2], è superiore a quello del metodo di Gauss, ma c'e' un caso particolare dove questi due metodi danno il risultato con le stesse operazioni, ed è il calcolo della matrice inversa.

Esempio 5.6. Si inverte la matrice

$$A = \begin{pmatrix} 2 & 4 & 1 \\ 4 & 1 & 0 \\ -2 & 2 & 1 \end{pmatrix}$$

col metodo di Gauss–Jordan. I passi necessari forniscono le matrici

$$E^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad (A^{(2)}|B^{(2)}) = \begin{pmatrix} 2 & 4 & 1 & 1 & 0 & 0 \\ 0 & -7 & -2 & -2 & 1 & 0 \\ 0 & 6 & 2 & 1 & 0 & 1 \end{pmatrix},$$

$$E^{(2)} = \begin{pmatrix} 1 & 4/7 & 0 \\ 0 & 1 & 0 \\ 0 & 6/7 & 1 \end{pmatrix}, \quad (A^{(3)}|B^{(3)}) = \begin{pmatrix} 2 & 0 & -1/7 & -1/7 & 4/7 & 0 \\ 0 & -7 & -2 & -2 & 1 & 0 \\ 0 & 0 & 2/7 & -5/7 & 6/7 & 1 \end{pmatrix},$$

$$E^{(3)} = \begin{pmatrix} 1 & 0 & 1/2 \\ 0 & 1 & 7 \\ 0 & 0 & 1 \end{pmatrix}, \quad (A^{(4)}|B^{(4)}) = \begin{pmatrix} 2 & 0 & 0 & -1/2 & 1 & 1/2 \\ 0 & -7 & 0 & -7 & 7 & 7 \\ 0 & 0 & 2/7 & -5/7 & 6/7 & 1 \end{pmatrix}.$$

La matrice inversa si trova risolvendo il sistema diagonale con termine noto matriciale $(A^{(4)}|B^{(4)})$, ossia dividendo le righe per i coefficienti sulla diagonale di $A^{(4)}$. Si ottiene

$$A^{-1} = \begin{pmatrix} -1/4 & -1/2 & 1/4 \\ 1 & -1 & -1 \\ -5/2 & 3 & 7/2 \end{pmatrix}.$$

5.8 Il metodo di Householder

Malgrado il fatto che il metodo di Gauss sia "ottimale" rispetto al numero delle operazioni, è interessante studiare anche altri metodi che in particolari circostanze forniscono il risultato più rapidamente o con migliore approssimazione.

Il metodo di Householder sfrutta l'algoritmo presentato nel paragrafo 5.4 impiegando una particolare classe di matrici elementari, introdotta di seguito.

Definizione 5.4. Siano $\beta \in \mathbb{R}$, $V \in \mathbb{C}^{n,1} \setminus \{O\}$. Se la matrice elementare

$$P \stackrel{\text{def}}{=} E(\beta, V, V) = I - \beta V V^*$$

è unitaria, allora P si dice *matrice elementare di Householder*.

Se P è una matrice elementare di Householder, allora valgono le seguenti proprietà, di verifica immediata:

1. P è hermitiana;
2. $P^{-1} = P$;
3. $\beta = 2/\|V\|^2$;
4. P è la simmetria ortogonale rispetto allo spazio vettoriale $\mathcal{L}(V)^\perp$ (per dimostrare questa proprietà basta provare che $PX = P(kV+Y) = -kV+Y$, dove $Y \in \mathcal{L}(V)^\perp$).

Il seguente teorema assicura l'esistenza di matrici elementari di Householder che realizzano l'algoritmo del paragrafo 5.4.

Teorema 5.6. *Dato $X \in \mathbb{C}^{n,1} \setminus \{O\}$, esiste una matrice elementare di Householder P tale che $PX = \alpha E_1$, con $\alpha \in \mathbb{C} \setminus \{0\}$.*

DIMOSTRAZIONE. Le quantità da determinare sono α e V.

1. *Determinazione di α.* Per il fatto che P è unitaria ed hermitiana devono essere verificate, rispettivamente, le seguenti equazioni:

$$\begin{cases} \|PX\| = \|X\| = |\alpha|, \\ X^*PX \in \mathbb{R} \quad \Rightarrow \quad X^*\alpha E_1 \in \mathbb{R} \quad \Rightarrow \quad \overline{x_1}\alpha \in \mathbb{R}. \end{cases}$$

La prima equazione dà il modulo di α, l'argomento viene dalla seconda equazione, e deve essere tale che $\arg(\overline{x_1}) + \arg(\alpha) = k\pi$, con $k \in \mathbb{Z}$. Posto

$$\theta = \begin{cases} \frac{x_1}{|x_1|} & \text{se } x_1 \neq 0, \\ 1 & \text{se } x_1 = 0, \end{cases}$$

risulta $\alpha = \pm\|X\|\theta$.

2. *Determinazione di V.* Si ha

$$PX = \alpha E_1 \quad \Rightarrow \quad (\beta V^*X)V = X - \alpha E_1.$$

Ponendo $V = X - \alpha E_1$ si ottiene la soluzione del problema. Infatti $\|V\|^2 = 2\|X\|(\|X\| + |x_1|)$, e

$$\beta = \frac{1}{V^*X} = \frac{1}{\|X\|^2 + \|X\||x_1|}.$$

Risulta

$$V = \begin{pmatrix} \theta(|x_1| + \|X\|) \\ x_2 \\ \vdots \\ x_n. \end{pmatrix}$$

QED

Il teorema successivo collega il precedente teorema alla fattorizzazione $\mathcal{QR}$ introdotta nel paragrafo 5.1.

Teorema 5.7. *Sia $A \in \mathbb{C}^{n,n}$. Allora esiste una fattorizzazione $A = QR$, dove Q è unitaria ed R è triangolare superiore.*

DIMOSTRAZIONE. La dimostrazione consiste nell'applicazione dell'algoritmo descritto nel paragrafo 5.4. La differenza col metodo di Gauss è che l'algoritmo è sempre applicabile in questo caso. Al passo k si pone

$$\theta^{(k)} = \begin{cases} \frac{a_{kk}^{(k)}}{|a_{kk}^{(k)}|} & \text{se } a_{kk}^{(k)} \neq 0, \\ 1 & \text{se } a_{kk}^{(k)} = 0. \end{cases}$$

Indicata con $F_{c1}^{(k)}$ la prima colonna di $F^{(k)}$, posto

$$V^{(k)} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \theta^{(k)}(\|F_{c1}^{(k)}\| + |a_{kk}^{(k)}|) \\ a_{k+1k}^{(k)} \\ \vdots \\ a_{nk}^{(k)} \end{pmatrix}, \quad \beta^{(k)} = \frac{1}{\|F_{c1}^{(k)}\|^2 + \|F_{c1}^{(k)}\||a_{kk}^{(k)}|},$$

si pone $E^{(k)} = P^{(k)} = E(\beta^{(k)}, V^{(k)}, V^{(k)})$.

Risulta:

$$R = A^{(n)}, \qquad Q = {P^{(1)}}^{-1} \cdots {P^{(n-1)}}^{-1} = P^{(1)} \cdots P^{(n-1)}.$$

QED

L'algoritmo presentato nel paragrafo 5.4 usato con le matrici elementari di Householder al fine di risolvere un sistema lineare $AX = B$ prende il nome di *metodo di Householder*. Valgono, per il metodo di Householder e relativamente alla soluzione del sistema con la fattorizzazione $\mathcal{QR}$, considerazioni simili a quanto detto per il metodo di Gauss. Inoltre, valgono conclusioni simili a quelle delle osservazioni 5.4, 5.6, 5.7. Invece, per quanto riguarda l'utilizzo di metodi del massimo pivot, si può dimostrare che il metodo di Householder non è influenzato dalla scelta di un elemento pivot con particolari proprietà [5].

Si può dimostrare che il numero di operazioni effettuato per ottenere la fattorizzazione $\mathcal{QR}$ è approssimabile (asintoticamente) da $2n^3/3$ [2].

Esempio 5.7. Trovare la fattorizzazione $\mathcal{QR}$ della matrice

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & 0 \\ 3 & 0 & -1 \end{pmatrix}.$$

Si ha $P^{(k)} = I - \beta^{(k)} V^{(k)} V^{(k)^*}$. I calcoli verranno effettuati in maniera approssimata. Per

$k = 1$ vale

$$\theta^{(1)} = 1, \quad \beta^{(1)} = 0,056365, \quad V^{(1)} = \begin{pmatrix} 4,7417 \\ 2 \\ 3 \end{pmatrix},$$

$$P^{(1)} = \begin{pmatrix} -0,26726 & -0,53452 & -0,80178 \\ -0,53452 & 0,77454 & -0,33819 \\ -0,80178 & -0,33819 & 0,49272 \end{pmatrix},$$

$$A^{(2)} = P^{(1)}A^{(1)} = \begin{pmatrix} -3,7417 & -1,0690 & 1,0690 \\ 0 & -0,29450 & 0,87271 \\ 0 & -1,9418 & 0,30906 \end{pmatrix}.$$

Nella prima colonna di $A^{(2)}$ i valori nulli lo sono a meno di errori di arrotondamento. Proseguendo:

$$\theta^{(2)} = -1, \quad \beta^{(2)} = 0,22544, \quad V^{(2)} = \begin{pmatrix} 0 \\ -2.2585 \\ -1,9418 \end{pmatrix},$$

$$P^{(2)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -0,14994 & -0,98869 \\ 0 & -0,98869 & 0,14995 \end{pmatrix},$$

$$A^{(3)} = P^{(2)}A^{(2)} = \begin{pmatrix} -3.7417 & -1.0690 & 1.0690 \\ 0 & 1.9640 & -0,43643 \\ 0 & 0 & -0,81650 \end{pmatrix}.$$

Dunque, $R = A^{(3)}$ e

$$Q = \begin{pmatrix} -0,26726 & 0,87286 & 0,40825 \\ -0,53452 & 0,21823 & -0,81649 \\ -0,80178 & -0,43644 & 0,40825 \end{pmatrix}.$$

Si può verificare che $A = QR$ con un errore dell'ordine di 10^{-6}.

5.9 Calcolo di basi ortonormali con il metodo $\mathcal{QR}$

Siano dati m vettori in $\mathbb{C}^n$ in forma di m colonne di una matrice $A \in \mathbb{C}^{n,m}$. Si supponga che gli m vettori dati siano indipendenti. Si consideri il problema di estrarre dai vettori considerati una base ortonormale. È semplice modificare l'algoritmo del paragrafo 5.4 in modo che possa essere applicato alla matrice rettangolare $A \in \mathbb{C}^{n,m}$. In questo modo si può calcolare una fattorizzazione $\mathcal{QR}$ della matrice A. Si ottiene $A = QR$, dove

$$Q \in U(n, \mathbb{C}), \qquad R = \begin{pmatrix} R_1 \\ O \end{pmatrix}, \quad R_1 \in \mathbb{C}^{m,m}. \tag{5.9.11}$$

Proposizione 5.8. *Nelle ipotesi precedenti, Q è una base ortonormale per* $\operatorname{Im} f_A$*, dove $f_A\colon \mathbb{C}^m \to \mathbb{C}^n$ è definita da $f_A(X) = AX$.*

Ovviamente, il problema della ricerca di una base ortonormale può anche essere risolto con il metodo di Gram–Schmidt. Questo metodo presenta un'efficienza di calcolo superiore di un fattore 2 rispetto al metodo $\mathcal{QR}$ presentato sopra. Tuttavia, si può dimostrare che il metodo di Gram–Schmidt è numericamente instabile [22]. Anche modifiche di questo metodo (*modified Gram–Schmidt algorithm*, si veda [22]) presentano il difetto di essere numericamente instabili se i vettori della base ortonormale sono vicini in norma. Si può dimostrare che il metodo $\mathcal{QR}$ è il più stabile [22, 6].

Osservazione 5.9. Si noti che, in linea di principio, sarebbe possibile realizzare una fattorizzazione $\mathcal{QR}$ utilizzando il metodo di Gram–Schmidt. Ma questa, ovviamente, soffrirebbe dei problemi descritti sopra.

5.10 Il metodo di Givens

Il *metodo di Givens* è stato introdotto come metodo per annullare selettivamente singoli elementi di una matrice $A \in \mathbb{C}^{n,n}$ mediante matrici unitarie. Esso consiste nell'applicazione del teorema 5.8 alla soluzione di sistemi lineari.

Definizione 5.5. Siano i, $j \in \mathbb{N}$, con $1 \leq i < j \leq n$. Si dice *matrice di Givens* una matrice $G_{ij} = (g_{hk}) \in \mathbb{C}^{n,n}$ con elementi tali che

$$g_{hk} = \begin{cases} 1 & \text{se } h = k \text{ e } h \neq i,\ h \neq j; \\ \cos\varphi & \text{se } h = k = i \text{ o } h = k = j; \\ \psi \sin\varphi & \text{se } h = j,\ k = i; \\ -\overline{\psi} \sin\varphi & \text{se } h = i,\ k = j; \\ 0 & \text{negli altri casi,} \end{cases}$$

dove $\varphi \in \mathbb{R}$ e $\psi \in \mathbb{C}$, $|\psi| = 1$.

In altri termini G_{ij} è la matrice

$$G_{ij} \stackrel{\text{def}}{=} \begin{pmatrix} 1 & 0 & \dots & & & & & & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \cos\varphi & 0 & \dots & -\overline{\psi}\sin\varphi & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \psi\sin\varphi & 0 & \dots & \cos\varphi & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & & & & & & & 1 \end{pmatrix},$$

ovvero la matrice (g_{hk}) che ha per sottomatrice la matrice

$$\begin{pmatrix} g_{ii} & g_{ij} \\ g_{ji} & g_{jj} \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\overline{\psi}\sin\varphi \\ \psi\sin\varphi & \cos\varphi \end{pmatrix},$$

e nella quale gli altri elementi coincidono con gli elementi della matrice identità.

Le matrici di Givens sono unitarie, e generalizzano a $\mathbb{C}^n$ le matrici di rotazione di $\mathbb{R}^2$. Se $U \in \mathbb{R}^{2,2}$ è una matrice ortogonale, allora

- se $\det U = 1$, allora U è una matrice di Givens;
- se $\det U = -1$, allora U è una matrice di Householder.

Ovviamente, se $A \in \mathbb{C}^{n,n}$, allora G_{ij} è una trasformazione elementare sulle righe di A. Il prossimo teorema individua una sottoclasse di queste trasformazioni particolarmente utile.

Teorema 5.8. *Sia $A \in \mathbb{C}^{n,n}$. Siano a_{ii}, $a_{ji} \neq 0$. Allora esiste una matrice di Givens G_{ij} tale che $(G_{ij}A)_{ji} = 0$.*

Dimostrazione. La dimostrazione è costruttiva. Posto

$$\psi = -\frac{a_{ji}}{a_{ii}} \left| \frac{a_{ii}}{a_{ji}} \right|,$$

$$\cos\varphi = \frac{|a_{ii}|}{\sqrt{|a_{ii}|^2 + |a_{ji}|^2}}, \quad \sin\varphi = \frac{|a_{ji}|}{\sqrt{|a_{ii}|^2 + |a_{ji}|^2}},$$

si ha

$$(G_{ij}A)_{ji} = \psi \sin\varphi a_{ii} + \cos\varphi a_{ji} = 0.$$

QED

Ovviamente, è possibile realizzare una fattorizzazione $\mathcal{QR}$ con le matrici di Givens. Ad esempio, per $n = 4$ tale fattorizzazione è ottenuta come segue

$$R = G_{34}G_{24}G_{23}G_{14}G_{13}G_{12}A, \quad Q^* = G_{34}G_{24}G_{23}G_{14}G_{13}G_{12}.$$

Tuttavia, il numero di operazioni effettuato è approssimabile con $4n^3/3$ [2], quindi è maggiore del numero di operazioni effettuate con le matrici elementari di Householder. Questa considerazione rende l'impiego delle matrici di Givens utile in casi particolari, come nel caso di matrici sparse (cioè con molti coefficienti nulli).

Esempio 5.8. Data la matrice $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & -1 \end{smallmatrix}\right)$, si determina la matrice di Givens G_{12} tale che $G_{12}A$ è triangolare superiore. Seguendo la dimostrazione del precedente teorema, sarà

$$G_{12} = \begin{pmatrix} \cos\varphi & -\overline{\psi}\sin\varphi \\ \psi\sin\varphi & \cos\varphi \end{pmatrix},$$

dove $\psi = -1$, $\cos\varphi = 1/\sqrt{10}$ e $\sin\varphi = 3/\sqrt{10}$, dunque

$$G_{12} = \begin{pmatrix} 1/\sqrt{10} & 3/\sqrt{10} \\ -3/\sqrt{10} & 1/\sqrt{10} \end{pmatrix}.$$

5.11 Il metodo di Cholesky

In questo paragrafo sarà dimostrato che ogni matrice hermitiana A definita positiva ammette una fattorizzazione che si calcola in maniera relativamente più rapida rispetto agli altri metodi. Questa fattorizzazione è alla base del metodo di Cholesky per la soluzione di un sistema lineare del tipo $AX = B$.

Teorema 5.9. *Sia $A \in \mathbb{C}^{n,n}$ una matrice hermitiana e definita positiva. Allora esiste un'unica matrice triangolare inferiore $L = (l_{ij}) \in \mathbb{C}^{n,n}$, con $l_{ii} > 0$ per $i = 1, \ldots, n$, tale che $A = LL^*$.*

Dimostrazione. La dimostrazione è costruttiva. Si ha

$$a_{ij} = \sum_{k=1}^{j} l_{ik} l_{kj}^* = \sum_{k=1}^{\min\{i,j\}} l_{ik} \overline{l}_{jk},$$

dove il secondo passaggio è dovuto al fatto che L è triangolare inferiore. Sarà risolta, ora, l'equazione precedente per $i \geq j$ (poiché A è hermitiana). Dalle equazioni

$$\begin{cases} a_{jj} = \sum_{k=1}^{n} |l_{jk}|^2, & j = 1, \ldots, n, \\ a_{ij} = \sum_{k=1}^{j} l_{ik} \overline{l}_{jk}, & i = j+1, \ldots, n,\ j = 1, \ldots, n-1, \end{cases} \tag{5.11.12}$$

si ricava direttamente

$$\begin{cases} l_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} |l_{jk}|^2}, & j = 1, \ldots, n \\ l_{ij} = \frac{1}{l_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik} \overline{l}_{jk} \right), & i = j+1, \ldots, n,\ j = 1, \ldots, n-1. \end{cases} \tag{5.11.13}$$

QED

L'algoritmo presentato nel precedente teorema può essere utilizzato al fine di risolvere un sistema lineare $AX = B$; esso prende il nome di *metodo di Cholesky*.

Il numero di operazioni da eseguire per trovare la fattorizzazione LL^* è approssimabile da $n^3/6$ [2]. Esso è inferiore al metodo di Gauss, ma il metodo non è applicabile a tutte le matrici.

Osservazione 5.10. Gli elementi della matrice L vengono trovati a partire da l_{11} incrementando prima l'indice di riga e poi quello di colonna. Ad esempio, se $n = 3$, l'ordine in cui sono ricavati gli elementi di L è il seguente:

$$\begin{pmatrix} 1 & & \\ 2 & 4 & \\ 3 & 5 & 6 \end{pmatrix}. \square$$

Osservazione 5.11. Dalla prima equazione in (5.11.12) si ottiene $|l_{jk}| \leq \sqrt{a_{ii}}$, dunque $\|L\|_\infty \leq \sqrt{\|A\|_\infty}$. $\square$

Esempio 5.9. Si fattorizzi con il metodo LL^* la matrice

$$A = \begin{pmatrix} 1 & 1+i & -1 \\ 1-i & 6 & 1-i \\ -1 & 1+i & 12 \end{pmatrix}.$$

Come prima cosa deve essere verificato il fatto che A è hermitiana e definita positiva. Usando la dimostrazione del teorema 5.9 si ottiene

$$l_{11} = \sqrt{a_{11}} = 1,\ l_{21} = \frac{1}{l_{11}}(a_{21}) = 1-i,\ l_{31} = \frac{1}{l_{11}}(a_{31}) = -1$$
$$l_{22} = \sqrt{a_{22} - |l_{21}|^2} = \sqrt{6 - |1-i|^2} = 2,\ l_{32} = \frac{1}{l_{22}}(a_{32} - l_{31}\overline{l_{21}}) = 1+i,$$
$$l_{33} = \sqrt{a_{33} - |l_{31}|^2 - |l_{32}|^2} = 3.$$

5.12 Esercizi di riepilogo

Esercizio 5.3. *Trovare l'inversa della matrice*

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{pmatrix}$$

con il metodo descritto nel paragrafo 5.1.

Esercizio 5.4. *Trovare, con il metodo di Gauss–Jordan, l'inversa della matrice*

$$A = \begin{pmatrix} 1 & 2 & -1 \\ -1 & 0 & 1 \\ 2 & 0 & 2 \end{pmatrix}$$

Risultato:

$$A^{-1} = \begin{pmatrix} 0 & -1/2 & 1/4 \\ 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/4 \end{pmatrix}.$$

Esercizio 5.5. *Si trovi la fattorizzazione $\mathcal{LU}$ della seguente matrice, ricorrendo, se necessario, al metodo del pivot*

$$A = \begin{pmatrix} 1 & 1 & -1 & 0 \\ 2 & -1 & 1 & 0 \\ 1 & -2 & 2 & 0 \\ -1 & 1 & 1 & 2 \end{pmatrix}.$$

Esercizio 5.6. *Se $U \in \mathbb{C}^n$ è un vettore di lunghezza unitaria, allora $E(2, U, U)$ è una matrice di Householder.*

Esercizio 5.7. *Sia data la matrice*

$$D = \begin{pmatrix} 1 & -i & 2 \\ i & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}.$$

1. *Dire quali fattorizzazioni di D sono realizzabili (tra le fattorizzazioni studiate $\mathcal{LU}$, $\mathcal{QR}$, $\mathcal{LL}^*$) e quali no, motivando le risposte. In caso affermativo, dire quante fattorizzazioni esistono.*

2. *Trovare la matrice elementare di Householder P tale che $PC_1 = \alpha E_1$, dove $\alpha \in \mathbb{C} \setminus 0$ ed $E_1 = (1, 0, 0)^T$.*

3. *Trovare una fattorizzazione $\Pi D = LU$, dove Π è un'opportuna matrice di permutazione.*

Soluzione.

1. I minori principali hanno determinanti $\det D_1 = 1$, $\det D_2 = 0$, $\det D_3 = -4$. Pertanto la matrice D non soddisfa le condizioni del teorema 5.5, ma esiste una matrice di permutazione Π tale che ΠA ammette una fattorizzazione $\mathcal{LU}$. Inoltre, per il calcolo precedente risulta che anche se D è hermitiana, non è definita positiva, pertanto non si può fattorizzare usando il metodo $\mathcal{LL}^*$. Una fattorizzazione $\mathcal{QR}$ esiste ed è univocamente determinata a meno di matrici di fase (secondo il teorema 5.3).

2. Banale.

3. Si può permutare la prima riga con la terza, ottenendo

$$\Pi D = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} D = \begin{pmatrix} 2 & 0 & 1 \\ i & 1 & 0 \\ 1 & -i & 2 \end{pmatrix}.$$

   In tal caso risulta

$$\Pi D = LU = \begin{pmatrix} 1 & 0 & 0 \\ 0.5i & 1 & 0 \\ 0.5 & -i & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & -0.5i \\ 0 & 0 & 2 \end{pmatrix}.$$

   Ovviamente, questa non è l'unica permutazione di D che si può fattorizzare!

□

Esercizio 5.8. *Sia data la matrice*

$$B = \begin{pmatrix} 1 & 2 & 0 & 1 \\ -1 & -2 & 1 & 2 \\ 0 & -1 & -3 & 1 \\ 1 & 2 & 0 & -1 \end{pmatrix}.$$

1. *Sotto quali condizioni la matrice B ammette una fattorizzazione LU?*

2. *Si calcoli l'inversa di B con il metodo di Gauss–Jordan, ricorrendo, se necessario, ad una permutazione delle righe.*

SOLUZIONE. 1. La matrice B ammette una fattorizzazione $\mathcal{LU}$ se i minori principali di ordine 1, 2 e 3 sono non singolari. Questo non è vero per B.

2. Per quanto detto sopra, si calcola l'inversa di ΠB col metodo di Gauss–Jordan, ove Π è la permutazione che scambia la seconda riga con la terza. Si ha:

$$\left(\begin{array}{cccc|cccc} 1 & 2 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & -3 & 1 & 0 & 1 & 0 & 0 \\ -1 & -2 & 1 & 2 & 0 & 0 & 1 & 0 \\ 1 & 2 & 0 & -1 & 0 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{cccc|cccc} 1 & 2 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & -3 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 & -1 & 0 & 0 & 1 \end{array}\right)$$

$$\to \left(\begin{array}{cccc|cccc} 1 & 0 & -6 & 3 & 1 & 2 & 0 & 0 \\ 0 & 1 & 3 & -1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 3 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 & -1 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 21 & 7 & 2 & 6 & 0 \\ 0 & 1 & 0 & -10 & -3 & -1 & -3 & 0 \\ 0 & 0 & 1 & 3 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 & -1 & 0 & 0 & 1 \end{array}\right)$$

$$\to \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -7/2 & 2 & 6 & 21/2 \\ 0 & 1 & 0 & 0 & 2 & -1 & -3 & -5 \\ 0 & 0 & 1 & 0 & -1/2 & 0 & 1 & 3/2 \\ 0 & 0 & 0 & 1 & 1/2 & 0 & 0 & -1/2 \end{array}\right).$$

Si noti che, in questo caso, il metodo di Gauss–Jordan ed il metodo di Gauss procedono in modi molto simili. Posto

$$X = \begin{pmatrix} -7/2 & 2 & 6 & 21/2 \\ 2 & -1 & -3 & -5 \\ -1/2 & 0 & 1 & 3/2 \\ 1/2 & 0 & 0 & -1/2 \end{pmatrix}$$

si nota che $X = (\Pi B)^{-1}$, da cui si ricava B^{-1}.

□

Esercizio 5.9. *Sia data la matrice*

$$M = \begin{pmatrix} 1 & 0 & 2 & 1 \\ -1 & 0 & -2 & 1 \\ 2 & 1 & 0 & -1 \\ 1 & 1 & -3 & 0 \end{pmatrix}.$$

1. *Quali fattorizzazioni della matrice M sono possibili, secondo le teorie studiate, al fine di risolvere un eventuale sistema lineare avente M come matrice dei coefficienti?*

2. *Si calcoli l'inversa di M con il metodo di Gauss–Jordan, ricorrendo, se necessario, ad una permutazione delle righe.*

SOLUZIONE. 1. Essendo $\det M_2 = 0$, le condizioni per la fattorizzabilità $\mathcal{LU}$ non sono verificate. Ma $\det M = -2$, quindi esiste una matrice di permutazioni Π tale che ΠM ammetta una fattorizzazione $\mathcal{LU}$. Una fattorizzazione $\mathcal{LL}^*$ non è possibile perché la matrice M non è simmetrica. Una fattorizzazione $\mathcal{QR}$ è sempre possibile, in questo caso è anche unica sotto le condizioni del teorema 5.3.

2. Si proceda come per l'esercizio 5.8.

□

Esercizio 5.10. *Si determinino gli autovalori e gli autovettori della matrice XY^* per $X, Y \in \mathbb{C} \setminus \{O\}$.*

SOLUZIONE. Si ha

$$(XY^*)Z = X(Y^*Z),$$

dove $Y^*Z = Y \cdot Z \in \mathbb{C}$. Pertanto la matrice XY^* ha rango 1 in quanto l'immagine è $\mathcal{L}(X)$, e ci sono gli autovalori $\lambda_1 = Y^*X$ e $\lambda_2 = 0$ Se $\lambda_1 \neq \lambda_2$, i corrispondenti autospazi sono

$$V(Y^*X) = \mathcal{L}(X), \quad V(0) = \mathcal{L}(Y)^\perp,$$

e la matrice XY^* è diagonalizzabile. Se $\lambda_1 = \lambda_2 = 0$ allora $V(0) = \mathcal{L}(Y)^\perp$, ma la matrice non è diagonalizzabile poiché $XY^* \neq O$ per l'ipotesi iniziale. □

Esercizio 5.11. *Si dica per quali valori α, $\beta \in \mathbb{C}$ esiste ed è unica la fattorizzazione LL^* della matrice*

$$M = \begin{pmatrix} \alpha & \beta & 0 \\ -\beta & \alpha & \beta \\ 0 & -\beta & \alpha \end{pmatrix},$$

e la si calcoli per tali valori.

SOLUZIONE. Affinché la matrice sia hermitiana deve essere $\alpha \in \mathbb{R}$ e $\beta = ib$, con $b \in \mathbb{R}$. Affinché sia definita positiva deve essere $\alpha > \sqrt{2}|b|$. Applicando il procedimento iterativo si ottiene

$$L = \sqrt{\alpha} \begin{pmatrix} 1 & 0 & 0 \\ -\frac{\beta}{\alpha} & \frac{\sqrt{\alpha^2 - b^2}}{\alpha} & 0 \\ 0 & -\frac{\beta}{\sqrt{\alpha^2 - b^2}} & \frac{\sqrt{\alpha^2 - 2b^2}}{\sqrt{\alpha^2 - b^2}} \end{pmatrix}.$$

□

CAPITOLO 6

LA RICERCA DEGLI AUTOVALORI

6.1 Premesse

Data una matrice hermitiana $A \in \mathbb{C}^{n,n}$, si dice *quoziente di Rayleigh* rispetto ad un vettore $X \in \mathbb{C}^n \setminus \{O\}$ l'espressione

$$R_A(X) \stackrel{\text{def}}{=} \frac{X^*AX}{X^*X}. \tag{6.1.1}$$

Poiché $A = A^*$ si ha $\overline{R_A(X)} = R_A(X)$, cioè $R_A(X)$ è reale. Posto $Y = X/\|X\|_2$, cioè $\|Y\|_2 = 1$, ne segue che i quozienti di Rayleigh non sono altro che i valori che la forma quadratica $Q(X) = ({}^*X)AX$ assume sulla sfera unitaria di $\mathbb{C}^n$.

La seguente proposizione mostra l'utilizzo del quoziente di Rayleigh.

Proposizione 6.1. *Sia $A \in \mathbb{C}^{n,n}$, e sia X un autovettore di A. Allora Il quoziente di Rayleigh di A rispetto ad X è un autovalore di A corrispondente ad X.*

DIMOSTRAZIONE. Infatti, se $AX = \lambda X$, risulta $X^*AX = X^*\lambda X$. QED

Poiché gli autovalori di A sono numeri reali, essi si possono ordinare in modo non crescente $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Un risultato dovuto a Courant e Fischer stabilisce che

$$\lambda_1 = \max\{R_A(X) \mid X \in \mathbb{C}^n,\ X \neq O\}$$
$$\lambda_n = \min\{R_A(X) \mid X \in \mathbb{C}^n,\ X \neq O\}$$

È possibile trovare facilmente autovalori di una matrice che sia un'espressione polinomiale di un'altra matrice di cui siano noti gli autovalori.

Proposizione 6.2. *Sia $q(z) \in \mathbb{C}_n[z]$ un polinomio. Sia $A \in \mathbb{C}^{n,n}$, λ un autovalore di A ed X un suo autovettore. Allora $q(\lambda)$ è un autovalore di $q(A)$.*

6.2 Localizzazione degli autovalori

La localizzazione degli autovalori è importante in quanto il calcolo numerico degli autovalori è molto difficile. Infatti, i metodi studiati per la soluzione dei sistemi lineari

non si applicano direttamente, come si vedrà nel paragrafo 6.3. Una prima soluzione del problema dela localizzazione degli autovalori è rappresentata dai teoremi di Gershgorin[1].

Teorema 6.1 (Primo teorema di Gershgorin). *Sia $A \in \mathbb{C}^{n,n}$. Posto*

$$C_i \stackrel{def}{=} \{z \in \mathbb{C} \mid |z - a_{ii}| \leq r_i\}, \quad r_i \stackrel{def}{=} \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \tag{6.2.2}$$

$$C \stackrel{def}{=} \bigcup_{i=1}^{n} C_i, \tag{6.2.3}$$

allora $\lambda_k \in C$ per $k = 1, \ldots, n$.

Analogamente, posto

$$D_i \stackrel{def}{=} \{z \in \mathbb{C} \mid |z - a_{ii}| \leq c_i\}, \quad c_i \stackrel{def}{=} \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ji}| \tag{6.2.4}$$

$$D \stackrel{def}{=} \bigcup_{i=1}^{n} D_i, \tag{6.2.5}$$

allora $\lambda_k \in D$ per $k = 1, \ldots, n$.

Dunque, $\lambda_k \in C \cap D$ per $k = 1, \ldots, n$.

Dimostrazione. Si applichi il teorema della diagonale dominante alla matrice $A' = A - \lambda I$. Se risultasse

$$|a_{ii} - \lambda| = |a'_{ii}| > \sum_{j \neq i} |a'_{ij}| = \sum_{j \neq i} |a_{ij}| = r_i,$$

allora la matrice A' risulterebbe invertibile, quindi λ non potrebbe essere un autovalore.

Si proceda in modo analogo per le colonne. QED

Definizione 6.1. Nelle ipotesi del primo teorema di Gershgorin, i cerchi $C_i \subset C$ ed i cerchi $D_i \subset D$ si dicono *cerchi di Gershgorin*.

Si osservi che se A è diagonale, i cerchi si riducono a punti.

Un secondo teorema permette di affinare la localizzazione degli autovalori.

Teorema 6.2 (Secondo teorema di Gershgorin). *Nelle ipotesi del primo teorema di Gershgorin, se m cerchi di Gershgorin $C_{i_1}, \ldots, C_{i_m}$ formano un insieme connesso[2] $S = \cup_{h=1}^{m} C_{i_h}$ disgiunto dai rimanenti $n - m$ cerchi C_i, allora S contiene esattamente m autovalori di A, contati ciascuno secondo la propria molteplicità algebrica.*

[1] Semyon A. Gershgorin, 1901–1933

[2] si veda [12] per la definizione di 'insieme connesso'

DIMOSTRAZIONE. Si ponga $A = D + R$ dove D è la diagonale principale di A. Si consideri la curva $t \mapsto A(t) \stackrel{\text{def}}{=} D + tR$. Gli autovalori della matrice $A(t)$ sono funzioni continue dei coefficienti di $A(t)$, poiché i coefficienti del polinomio caratteristico sono funzioni continue di t e gli zeri di un polinomio sono funzioni continue dei propri coefficienti. Per $t \in [0,1]$ ci sono m funzioni continue $\lambda_{i_1}(t), \ldots, \lambda_{i_m}(t)$ tali che $\lambda_{i_h}(0) = a_{i_h i_h}$ e $\lambda_{i_h}(1)$ è un autovalore di $A = A(1)$ (per $h = 1, \ldots, m$). Queste funzioni continue sono curve in S per $t \in [0,1]$ poiché i loro punti iniziali sono in S, e poiché l'immagine dell'insieme connesso $[0,1]$, essendo connessa, non può 'uscire da S'. QED

È possibile applicare i precedenti teoremi al problema della localizzazione degli zeri di un polinomio. Infatti, come è facilmente dimostrabile, dato un polinomio monico

$$p(x) = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1} + x^n \in \mathbb{C}_n[x] \tag{6.2.6}$$

esiste una matrice A_p, detta *matrice compagna* di $p(x)$, tale che $P_{A_p}(x) = p(x)$; tale matrice è così definita:

$$A_p \stackrel{\text{def}}{=} \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{pmatrix}.$$

Applicando alla matrice compagna i precedenti risultati si ottiene il seguente teorema.

Teorema 6.3 (Cauchy). *Gli zeri del polinomio* (6.2.6) *si trovano nel disco di centro* $0 \in \mathbb{C}$ *e raggio*

$$M = \max\{|a_0|, 1 + |a_1|, \ldots, 1 + |a_{n-1}|\}.$$

DIMOSTRAZIONE. Si consideri la matrice compagna A_p del polinomio $p(x)$. Applicando il primo teorema di Gershgorin (per colonne) risulta

$$\begin{aligned} &D_1 = \{z \in \mathbb{C} \mid |z| \leq |a_0|\}, \\ &D_2 = \{z \in \mathbb{C} \mid |z| \leq 1 + |a_1|\}, \\ &\ldots \\ &D_{n-1} = \{z \in \mathbb{C} \mid |z| \leq 1 + |a_{n-2}|\}, \\ &D_n = \{z \in \mathbb{C} \mid |z + a_{n-1}| \leq 1\} \subset \{z \in \mathbb{C} \mid |z| \leq 1 + |a_{n-1}|\} = D'_n, \end{aligned}$$

poiché per la disuguaglianza triangolare risulta

$$||z| - |a_{n-1}|| \leq |z + a_{n-1}| \leq 1.$$

La dimostrazione termina osservando che i dischi D_i $(i = 1, \ldots, n-1)$ e D'_n sono concentrici ed il loro centro è l'origine. QED

Nota 6.1. Un altro risultato di localizzione, meno fine dei precedenti, è costituito dal teorema di Hirsch 7.7.

Esempio 6.1. Si consideri $p(x) = x^3 - 2x^2 + x - 2 = 0$. Allora $M = \max\{2, 2, 3\} = 3$, quindi le radici si trovano nel disco di raggio 3 del piano complesso. Infatti, con un semplice calcolo si può verificare che le radici sono 2, i, $-i$.

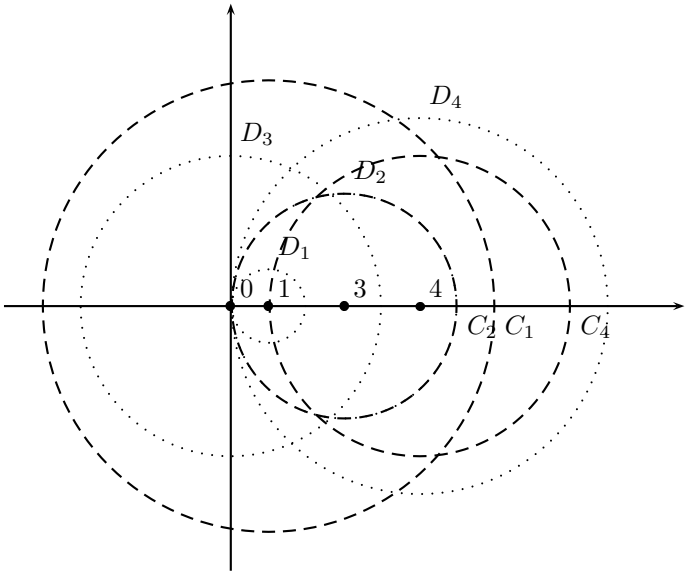
Esercizio 6.1. *Trovare le regioni del piano complesso in cui si possono trovare gli autovalori della matrice*

$$A = \begin{pmatrix} 1 & -1 & 2 & 3 \\ 0 & 3 & -1 & 2 \\ 0 & 0 & 0 & 0 \\ 1 & 2 & 1 & 5 \end{pmatrix}.$$

SOLUZIONE. Si ha

$$r_1 = 6,\ r_2 = 3,\ r_3 = 0,\ r_4 = 4, \quad d_1 = 1,\ d_2 = 3,\ d_3 = 4,\ d_4 = 5.$$

I cerchi di Gershgorin



(si noti che $C_2 = D_2$) si trovano tutti all'interno del cerchio $\{z \in \mathbb{C} \mid |z| \leq 9\}$. Dal primo teorema di Gershgorin si ha una stima ancora più precisa: gli autovalori di A devono trovarsi nell'insieme

$$(C_1 \cup C_4) \cap (D_3 \cup D_4).$$

Infatti, dal calcolo risulta che gli autovalori di A sono $\lambda = 0$ e $\lambda = 9$. □

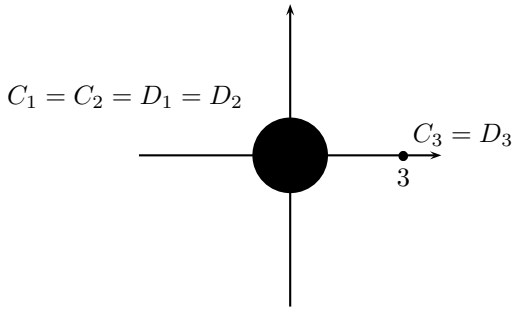
Esercizio 6.2. *Si localizzino gli autovalori della matrice*

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

SOLUZIONE. Si ha

$$C_1 = \{z \in \mathbb{C} \mid |z| \leq 1\} = C_2 = D_1 = D_2,$$
$$C_3 = \{z \in \mathbb{C} \mid |z-3| = 0\} = \{3\} = D_3$$
$$C = C_1 \cup \{3\} = D_1 \cup \{3\} = D.$$

Per il secondo teorema di Gershgorin, $C_1 \cup C_2$ contiene 2 autovalori (si verifica che sono $\lambda = \pm i$), e C_3 contiene un solo autovalore, necessariamente $\lambda = 3$.



□

6.3 Il metodo $\mathcal{QR}$ per la ricerca degli autovalori

La ricerca degli autovalori è un problema la cui soluzione è assai complessa nel caso in cui l'ordine della matrice non sia piccolo. Lo scopo di questo paragrafo è far vedere come il metodo di Householder sia usato a questo scopo, a scapito della sua minore efficienza computazionale rispetto al metodo di Gauss.

Sia $A \in \mathbb{C}^{n,n}$ diagonalizzabile, con $D = T^{-1}AT$ diagonale. Si può dimostrare [2] che l'errore che si commette nella ricerca degli autovalori di A è limitato dalla costante $\mu(T) = \|T\|\|T^{-1}\|$, detta *numero di condizionamento di* T, dove $\|\cdot\|$ è una norma matriciale tale che $\|D\| = \max_{i=1,\ldots,n}\{|d_{ii}|\}$. È evidente che la scelta di T unitaria permette di minimizzare tale valore: ad esempio, $\|T\|_\infty$, $\|T^{-1}\|_\infty \leq 1$.

Sarà descritto, ora, l'algoritmo che è il nucleo del metodo $\mathcal{QR}$: il metodo stesso è composto anche di altri passi, per i quali si rimanda a [5]. Si osservi che una matrice simile ad una matrice diagonalizzabile è diagonalizzabile: se $B = S^{-1}AS$ ed A è diagonalizzabile come nell'enunciato precedente, allora $D = (S^{-1}T)^{-1}B(S^{-1}T)$, e $\mu(S^{-1}T) \leq \|S\|\|S^{-1}\|\|T\|\|T^{-1}\|$.

Si ponga $A^{(1)} \stackrel{\text{def}}{=} A$. Al passo k, se la decomposizione $\mathcal{QR}$ di $A^{(k)}$ è $A^{(k)} = Q^{(k)}R^{(k)}$, si pone $A^{(k+1)} \stackrel{\text{def}}{=} R^{(k)}Q^{(k)}$. Risulta $A^{(k+1)} = Q^{(k)*}A^{(k)}Q^{(k)}$. Il metodo è iterativo, ossia fornisce una successione che si dimostra essere convergente ad una matrice triangolare superiore sotto opportune ipotesi.

Lemma 6.1. *Sia A_k una successione di matrici in $\mathbb{C}^n$ convergente ad I. Se $A_k = Q_kR_k$ è una fattorizzazione $\mathcal{QR}$, allora esiste una successione di matrici di fase S_k tali che*

$$\lim_{k\to+\infty} Q_kS_k = I = \lim_{k\to+\infty} S_kQ_k, \quad \lim_{k\to+\infty} R_kS_k^* = I = \lim_{k\to+\infty} S_k^*R_k$$

DIMOSTRAZIONE. Poiché gli elementi di Q_k hanno modulo limitato, si ha

$$O = \lim_{k\to+\infty}(Q_kR_k - I) = \lim_{k\to+\infty} Q_k(R_k - Q_k^*) = \lim_{k\to+\infty}(R_k - Q_k^*).$$

Essendo R_k triangolari superiori, risulta $\lim_{k\to+\infty} q_{kij} = 0$ per $i < j$; pertanto $q_{kii} \neq 0$ per $k > N$. Posto $\varphi_i^{(k)} = q_{kii}/|q_{kii}|$ la tesi è dimostrata definendo S_k come le matrici tali che $s_{ii} = \varphi_i^{(k)}$. QED

Teorema 6.4. *Sia $A \in \mathbb{C}^{n,n}$ con autovalori λ_i, $i = 1, \ldots, n$, distinti in modulo, cioè*

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > 0.$$

Sia $X \in \mathbb{C}^{n,n}$ tale che $A = XDX^{-1}$, con D diagonale. Supponiamo che esista una fattorizzazione $\mathcal{LU}$ di X^{-1}. Allora esiste una successione di matrici di fase $\{S_k\}_{k\in\mathbb{N}}$ tale che

$$\lim_{k\to+\infty} S_k^* R^{(k)} S_{k-1} = T = \lim_{k\to+\infty} S_{k-1}^* A^{(k)} S_{k-1},$$
$$\lim_{k\to+\infty} S_{k-1}^* Q^{(k)} S_k = I,$$

dove T è triangolare superiore.

Se A è hermitiana, allora T è diagonale.

DIMOSTRAZIONE. La dimostrazione si basa sul confronto di due fattorizzazioni $\mathcal{QR}$ della matrice A^k (potenza k-es. di A, da non confondere con $A^{(k)}$!).

Prima fattorizzazione $\mathcal{QR}$. Si ha

$$A^k = H_kU_k, \qquad H_k = Q^{(1)}\cdots Q^{(k)}, \quad U_k = R^{(k)}\cdots R^{(1)}.$$

Infatti, per induzione, si ha $A = Q^{(1)}R^{(1)}$, e

$$\begin{aligned} H_{k+1}U_{k+1} &= Q^{(1)}\cdots Q^{(k+1)}R^{(k+1)}\cdots R^{(1)} \\ &= Q^{(1)}\cdots Q^{(k)}A^{(k+1)}R^{(k)}\cdots R^{(1)} \\ &= Q^{(1)}\cdots A^{(k)}Q^{(k)}R^{(k)}\cdots R^{(1)} \\ &= AQ^{(1)}\cdots Q^{(k)}R^{(k)}\cdots R^{(1)} \\ &= AA^k = A^{k+1}. \end{aligned}$$

Seconda fattorizzazione $\mathcal{QR}$. Posto $X^{-1} = LU$, si ha

$$A^k = XD^kLU = XD^kLD^{-k}D^kU,$$

dove risulta

$$(D^k L D^{-k})_{ij} = \begin{cases} l_{ij} \left(\frac{\lambda_i}{\lambda_j}\right)^k & i > j \\ 1 & i = j \\ 0 & i < j \end{cases}$$

Pertanto si può scrivere $D^k L D^{-k} = I + E_k$, dove $\lim_{k\to+\infty} E_k = O$. Se $X = QR$ è una fattorizzazione $\mathcal{QR}$ di X, si può scrivere

$$A^k = QR(I + E_k)D^k U = Q(I + RE_kR^{-1})RD^kU = (QP_k)(T_kRD^kU),$$

dove $I + RE_kR^{-1} = P_kT_k$ è una fattorizzazione $\mathcal{QR}$.

Confronto tra le due fattorizzazioni. Usando il teorema 5.3, esistono matrici di fase $\Phi_k \in U(n, \mathbb{C})$ tali che

$$H_k = (QP_k)\Phi_k^*, \quad U_k = \Phi_k(T_kRD^kU).$$

Pertanto si ha

$$\begin{aligned} Q^{(k)} &= H_{k-1}^{-1}H_k = \Phi_{k-1}P_{k-1}^*P_k\Phi_k^*, \\ R^{(k)} &= U_kU_{k-1}^{-1} = \Phi_kT_kRDR^{-1}T_{k-1}^{-1}\Phi_{k-1}^*. \end{aligned}$$

Applicando il lemma 6.1 alla successione P_kT_k si ottengono matrici di fase F_k tali che $\lim_{k\to+\infty} P_kF_k = I$, $\lim_{k\to+\infty} F_k^*T_k = I$. Si ponga $S_k \stackrel{\text{def}}{=} \Phi_kF_k$. Si ottiene la tesi con $T \stackrel{\text{def}}{=} RDR^{-1}$. QED

Per ridurre il numero delle operazioni effettuate con il metodo $\mathcal{QR}$, si effettua una preliminare riduzione della matrice A, secondo l'algoritmo seguente. Esiste una matrice di Householder Q tale che $(QA)_{i1} = 0$ per $i > 2$. Si ripete il metodo sulla seconda colonna di QA, e così via fino ad ottenere una matrice di Hessenberg superiore. È facile rendersi conto del fatto che, se A è hermitiana, allora la matrice di Hessenberg superiore così ottenuta è, in realtà, una matrice tridiagonale.

Altri metodi permettono di valutare il numero di iterazioni del metodo $\mathcal{QR}$ necessarie ad ottenere una matrice triangolare con la precisione richiesta. Varianti del metodo considerano l'utilizzo di matrici di Givens, ed anche matrici elementari di Gauss. Per maggiori informazioni, si veda [5, 22].

CAPITOLO 7

DECOMPOSIZIONE AI VALORI SINGOLARI ED APPLICAZIONI

7.1 Premesse

Si consideri una matrice $A \in \mathbb{C}^{m,n}$. I seguenti teoremi caratterizzano le situazioni in cui $\mathrm{rg}(A) = n$ o $\mathrm{rg}(A) = m$.

Teorema 7.1. *Sia A una matrice di $\mathbb{C}^{m,n}$. Le seguenti affermazioni sono equivalenti.*

1. *$AX = O$ ha l'unica soluzione $X = O$.*
2. *Le colonne di A sono linearmente indipendenti.*
3. *$\mathrm{rg}(A) = n$.*
4. *$A^*A \in \mathbb{C}^{n,n}$ è invertibile, quindi $\mathrm{rg}(A^*A) = n$.*

DIMOSTRAZIONE. Le equivalenze tra i primi tre punti sono già note [16].
(4) $\Rightarrow$ (1) Se $AX = O$, allora $A^*AX = A^*O = O$ e quindi l'unica soluzione è $X = O$ per la (4).
(1) $\Rightarrow$ (4) Sia $(A^*A)X = O$ e poniamo $AX = (y_1 \dots y_m)^T$. Allora $|y_1|^2 + \cdots + |y_m|^2 = (AX)^*(AX) = X^*(A^*AX) = O$, dunque $y_i = 0$ per $i = 1, \dots, m$, da cui $AX = O$ e l'unica soluzione $X = O$ per la (1). Quindi $(A^*A)X = O$ implica $X = O$ e $\mathrm{rg}\,(A^*A) = n$. QED

Teorema 7.2. *Sia A una matrice di $\mathbb{C}^{m,n}$. Le seguenti affermazioni sono equivalenti.*

1. *$AX = B$ ammette soluzione per ogni $B \in \mathbb{C}^m$.*
2. *Le colonne di A generano $\mathbb{C}^m$.*
3. *$\mathrm{rg}(A) = m$.*
4. *$AA^* \in \mathbb{C}^{m,m}$ è invertibile, quindi $\mathrm{rg}(AA^*) = m$.*

DIMOSTRAZIONE. Le equivalenze tra i primi tre punti sono già note. Si consideri (3) $\Leftrightarrow$ (4). Posto $C = A^* \in \mathbb{C}^{n,m}$, l'equivalenza è dimostrata applicando il teorema precedente alla matrice C per la quale $\mathrm{rg}\, C = \mathrm{rg}\, A = m$. QED

Proposizione 7.1. *Sia $A \in \mathbb{C}^{m,n}$. Allora le matrici A^*A e AA^* sono hermitiane e semidefinite positive.*

DIMOSTRAZIONE. La prima affermazione è ovvia. Si consideri una base ortonormale di autovettori $\{Z_i\}$ di A^*A con associati autovalori $\{\lambda_i\}$ per $i = 1, \ldots, n$. Si ha

$$(AZ_i) \cdot (AZ_j) = (AZ_i)^* AZ_j = \lambda_j Z_i^* Z_j = \lambda_j Z_i \cdot Z_j,$$

e per $i = j$ risulta $\|AZ_i\|^2 = \lambda_i \geq 0$. Si proceda in maniera analoga per AA^*. QED

Si può dimostrare in generale che se $A \in \mathbb{C}^{m,n}$ e $B \in \mathbb{C}^{n,m}$, allora le matrici AB e BA hanno gli stessi autovalori non nulli, e con le stesse molteplicità geometriche. Ne segue che A^*A e AA^* hanno gli stessi autovalori non nulli, e con le stesse molteplicità geometriche. Mentre può accadere che 0 sia un autovalore per A^*A e non per AA^*, o viceversa; inoltre $\operatorname{rg}(A^*A) = \operatorname{rg}(AA^*) = \operatorname{rg} A = \operatorname{rg} A^*$[1].

Definizione 7.1. Sia $A \in \mathbb{C}^{m,n}$, e siano $\{\lambda_i\}_{i=1,\ldots,n}$ gli autovalori di A^*A (o AA^*). I numeri $\sigma_i \stackrel{\text{def}}{=} \sqrt{\lambda_i}$ sono detti *valori singolari* di A.

Si osservi che se A è hermitiana, allora $A^*A = A^2$, quindi i numeri σ_i sono i valori assoluti degli autovalori di A.

Esercizio 7.1. *Mostrare che se $A \in \mathbb{C}^{n,n}$ è non singolare, allora A^*A è definita positiva.*

7.2 Decomposizione ai valori singolari

Sia $A \in \mathbb{C}^{m,n}$, con $\operatorname{rg} A = r$. Ordinati i valori singolari non nulli di A così: $\sigma_1 \geq \cdots \geq \sigma_r$, si pone

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & \ldots\ldots & 0 \\ \ldots\ldots\ldots\ldots & & & \\ 0 & \ldots & \sigma_r & \ldots & 0 \\ 0 & \ldots\ldots\ldots\ldots & & 0 \end{pmatrix} \in \mathbb{R}^{m,n} \subset \mathbb{C}^{m,n}. \tag{7.2.1}$$

Teorema 7.3. *Sia $A \in \mathbb{C}^{m,n}$, con $\operatorname{rg} A = r$. Allora la matrice A si fattorizza nel modo seguente*

$$A = P\Sigma Q^*, \qquad P \in \mathbb{C}^{m,m}, \quad Q \in \mathbb{C}^{n,n},$$

dove P e Q sono unitarie e non univocamente determinate, mentre Σ è univocamente determinata.

DIMOSTRAZIONE. Sia $\{Q_i\}_{i=1,\ldots,n}$ una base ortonormale di autovettori di A^*A associati agli autovalori $\{\lambda_i\}_{i=1,\ldots,n}$. La matrice $Q = (Q_1 \cdots Q_n)$ è una matrice unitaria. Posto

$$P_i \stackrel{\text{def}}{=} \frac{1}{\sigma_i} AQ_i \qquad i = 1, \ldots, r,$$

[1] In generale $\operatorname{rg}(A+B) \leq \operatorname{rg} A + \operatorname{rg} B$, $\operatorname{rg}(AB) \leq \min\{\operatorname{rg} A, \operatorname{rg} B\}$

P_i risulta un versore poiché $\|AQ_i\|^2 = \sigma_i^2$. Si completi l'insieme ortonormale $\{P_i\}_{i=1,\dots,r}$ ad una base ortonormale $\{P_i\}_{i=1,\dots,m}$ di $\mathbb{C}^m$, e si ponga $P = (P_1 \cdots P_m)$. La matrice P è unitaria e risulta $A = P\Sigma Q^*$. Infatti, $P\Sigma = (\sigma_1 P_1 \cdots \sigma_r P_r\ 0 \cdots 0)$. e $AQ = (AQ_1 \cdots AQ_n) = (\sigma_1 P_1 \cdots \sigma_r P_r\ 0 \cdots 0)$ da cui la tesi. QED

Considerata l'applicazione lineare $f_A \colon \mathbb{C}^n \to \mathbb{C}^m$ indotta da A [16], il teorema assicura che si possono scegliere in $\mathbb{C}^n$ e $\mathbb{C}^m$ delle basi ortonormali rispetto alle quali l'applicazione f_A ha come matrice associata Σ, che è di tipo diagonale (nel senso specificato in 1.1).

Esempio 7.1. Sia

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \end{pmatrix}. \quad \text{Allora} \quad A^T A = \begin{pmatrix} 5 & 4 & 0 \\ 4 & 5 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

e gli autovalori sono $\lambda_1 = 9$, $\lambda_2 = 1$, $\lambda_3 = 0$. Si noti che $AA^T = \left(\begin{smallmatrix} 5 & 4 \\ 4 & 5 \end{smallmatrix}\right)$, con gli stessi autovalori non nulli di $A^T A$. Si ha $\sigma_1 = 3$, $\sigma_2 = 1$,

$$Q_1 = \frac{1}{\sqrt{2}}(1\ 1\ 0)^T, \quad Q_2 = \frac{1}{\sqrt{2}}(-1\ 1\ 0)^T, \quad Q_3 = (0\ 0\ 1)^T.$$

Usando $P_i = (1/\sigma_i)AQ_i$ si ottiene

$$P_1 = \frac{1}{\sqrt{2}}(1\ 1)^T, \quad P_2 = \frac{1}{\sqrt{2}}(1\ -1)^T,$$

quindi

$$P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad Q = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}.$$

La decomposizione non è unica.

Osservazione 7.1.

1. La decomposizione ai valori singolari (SVD) si può considerare come un'estensione del teorema degli assi principali a matrici non quadrate. Essa è molto utile in svariate applicazioni dell'algebra lineare, per esempio nella compressione e nella trasformazione di dati.

   Per le matrici reali quadrate la decomposizione risale a Sylvester nel 1889, la prima dimostrazione nel caso generale risale al 1939 ed è dovuta ad Eckart e Young.

2. La decomposizione ai valori singolari è utile per stimare quanto un sistema lineare sia sensibile agli arrotondamenti dei coefficienti. Se σ_i sono i valori singolari, la frazione $c = \sigma_1/\sigma_r \geq 1$ è detta *indice di condizionamento* della matrice A: valori alti di c segnalano un'alta pericolosità degli errori dovuti agli arrotondamenti.

3. L'interpretazione geometrica dei valori singolari è la seguente. Sia $f\colon \mathbb{R}^n \to \mathbb{R}^m$ l'applicazione lineare associata ad A. Allora, per SVD, esistono due basi ortonormali $\mathcal{B} = \{\vec{u}_i\}$ e $\mathcal{B}' = \{\vec{u'}_i\}$ di $\mathbb{R}^n$ tali che

$$\begin{aligned} f(\vec{u}_i) &= \sigma_i \vec{u}_i && i = 1, \dots, r \\ f(\vec{u}_i) &= 0 && i = r+1, \dots, n \end{aligned}$$

Se (x_i) sono le coordinate rispetto a $\mathcal{B}$ in $\mathbb{R}^n$ e (y_j) quelle rispetto a $\mathcal{B}'$ in $\mathbb{R}^m$, le equazioni di f sono

$$\begin{aligned} y_i &= \sigma_i x_i && i = 1, \dots, r \\ y_i &= 0 && i = r+1, \dots, m \end{aligned}$$

Quindi la sfera unitaria S (di dimensione $r-1$) di $\mathbb{R}^r \subset \mathbb{R}^n$ si muta nell'ellissoide $f(S)$ (di dimensione $r-1$) dello spazio $\mathbb{R}^r \subset \mathbb{R}^m$. I valori singolari sono le lunghezze dei vari assi dell'ellissoide, mentre l'indice di condizionamento è legato all'eccentricità dell'ellissoide.

4. Si vede facilmente che A ed A^* hanno gli stessi valori singolari non nulli:

$$A = P\Sigma Q^* \quad \Rightarrow \quad A^* = Q\Sigma^* P^* = Q\Sigma^T P^*. \ \square$$

Esercizio 7.2. *Provare che i valori singolari di A^{-1} sono i reciproci di quelli di A.*

Esercizio 7.3. *Se A è una matrice quadrata, mostrare che $|\det A|$ è il prodotto dei valori singolari di A.*

7.3 Applicazioni

7.3.a Sistemi lineari generali

Si consideri il sistema lineare di m equazioni in n incognite ($m \geq n$) espresso da $AX = B$, con $A \in \mathbb{C}^{m,n}$, $X \in \mathbb{C}^{n,1}$, $B \in \mathbb{C}^{m,1}$. Se $A = P\Sigma Q^*$, ponendo

$$Y = Q^* X, \qquad \tilde{B} = P^* B,$$

il problema considerato si spezza in due problemi:

1. risolvere il sistema $\Sigma Y = \tilde{B}$, con Σ matrice diagonale;

2. risolvere il sistema $Y = Q^* X$, con Q matrice unitaria.

1. Se $\{\sigma_i\}$ $(i = 1, \dots, n)$ sono i valori singolari di A, posto $Y = (y_1 \cdots y_n)^T$ e $\tilde{B} = (\tilde{b_1} \cdots \tilde{b_m})^T$ si ha

$$\begin{aligned} \sigma_j y_j &= \tilde{b}_j && \text{se } j \le n \text{ e } \sigma_j \neq 0, \\ 0 y_j &= \tilde{b}_j && \text{se } j \le n \text{ e } \sigma_j = 0, \\ 0 &= \tilde{b}_j && \text{se } j > n. \end{aligned}$$

   Il secondo insieme dei valori è vuoto se $r = \operatorname{rg} A = n$; il terzo insieme è vuoto se $n = m$. Dunque

$$AX = B \text{ compatibile} \quad \Leftrightarrow \quad \tilde{b}_j = 0 \text{ per } j > r.$$

   Se $r < n$, allora la y_j associata ad un σ_j nullo può assumere valore arbitrario (per $j = r + 1, \dots, n$).

2. Poiché Q è invertibile, si ha $X = QY$, dipendente da $n - r$ parametri.

Esempio 7.2. Si consideri il sistema lineare $AX = B$ dove

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} -5 \\ -4 \end{pmatrix}.$$

Tenendo conto dell'esempio 7.1 si ha

$$\Sigma Y = \tilde{B} \quad \Rightarrow \quad \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -9/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix},$$

da cui

$$Y = \frac{1}{\sqrt{2}}(-3 - 1h\sqrt{2})^T, \qquad X = (-1 - 2h)^T.$$

7.3.b La pseudoinversa di Moore–Penrose

Sia $A \in \mathbb{C}^{m,n}$ una matrice che ammette la decomposizione ai valori singolari $A = P\Sigma Q^*$. Siano $\sigma_1 \ge \cdots \ge \sigma_r$ i valori singolari non nulli di A. Se

$$D \stackrel{\text{def}}{=} \begin{pmatrix} \sigma_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \sigma_r \end{pmatrix}, \quad \Sigma = \begin{pmatrix} D & O \\ O & O \end{pmatrix} \in \mathbb{C}^{m,n},$$

si pone

$$\Sigma^+ \stackrel{\text{def}}{=} \begin{pmatrix} D^{-1} & O \\ O & O \end{pmatrix} \in \mathbb{C}^{n,m}.$$

Definizione 7.2. La matrice

$$A^+ \stackrel{\text{def}}{=} Q\Sigma^+ P^* \in \mathbb{C}^{n,m}$$

si dice *pseudoinversa di Moore–Penrose* di A.

Esempio 7.3. Considerando la matrice A dell'esempio 7.1, risulta

$$\Sigma^+ = \begin{pmatrix} 1/3 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad A^+ = \begin{pmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \\ 0 & 0 \end{pmatrix}.$$

Il nome 'pseudoinversa' è giustificato dal fatto che, se A è invertibile, allora $A^+ = A^{-1}$, come si può facilmente verificare. Infatti, in generale risulta

$$AA^+ = (P\Sigma Q^*)(Q\Sigma^+ P^*) = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} P^*,$$
$$A^+A = (Q\Sigma^+ P^*)(P\Sigma Q^*) = Q \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q^*.$$

Se A è invertibile allora $\operatorname{rg} A = n$, dunque $r = n$, quindi $AA^+ = PP^* = I$, $A^+A = QQ^* = I$.

La pseudoinversa può essere univocamente caratterizzata mediante la richiesta di opportune proprietà (come quelle dell'esercizio 7.5).

Teorema 7.4. *Sia $A \in \mathbb{C}^{n,m}$. Allora la matrice A^+ è univocamente caratterizzata dalle seguenti proprietà*

$$AA^+A = A \qquad AA^+ = (AA^+)^*$$
$$A^+AA^+ = A^+ \qquad A^+A = (A^+A)^*$$

Inoltre,

$$(A^+)^+ = A \qquad (A^+)^* = (A^*)^+$$
$$(cA)^+ = c^{-1}A^+ \qquad (O_{n,m})^+ = O_{m,n}$$

$$A = A^* \quad \Rightarrow \quad A^+ = (A^+)^*,$$
$$A^2 = A = A^* \quad \Rightarrow \quad A^+ = A,$$

e se $\operatorname{rg}(A) = k$, $A = BC$, con $B \in \mathbb{C}^{m,k}$, $C \in \mathbb{C}^{k,n}$, allora $A^+ = C^+B^+$, dove

$$C^+ = C^*(CC^*)^{-1}, \qquad B^+ = (B^*B)^{-1}B^*.$$

Si noti che alle decomposizioni $\mathcal{LU}$ e $\mathcal{QR}$ è possibile applicare la proprietà (8). Inoltre, esistono fattorizzazioni $A = BC$ dove $A^+ \neq C^+B^+$ nel caso in cui non sia verificata l'ipotesi sul rango di A.

Osservazione 7.2. Se $A \in \mathbb{C}^{m,n}$ ha rango massimo n, allora essa ammette un'inversa sinistra, ossia esiste una matrice $A' \in \mathbb{C}^{n,m}$ tale che $A'A = I \in \mathbb{C}^{n,n}$. In tal caso, per il teorema 7.1, A^*A è invertibile e

$$A' = (A^*A)^{-1}A^* = A^+,$$

ed analogamente nel caso in cui esista un'inversa destra. □

Osservazione 7.3. Sia $AX = B$ un sistema lineare (anche non compatibile), e sia

$$A = P\Sigma Q^* \in \mathbb{C}^{m,n}$$

la decomposizione ai valori singolari di A. Sia $\mathrm{rg}(A) = r = \mathrm{rg}(\Sigma)$. Si scriva $\mathbb{C}^{n,1} = \mathbb{C}^{r,1} \times \mathbb{C}^{n-r,1}$. Allora, $Y \in \mathbb{C}^{n,1}$ si scrive come

$$Y = \begin{pmatrix} Y' \\ Y'' \end{pmatrix} \quad \text{dove} \quad Y' \in \mathbb{C}^{r,1}, \quad Y'' \in \mathbb{C}^{n-r,1}.$$

Ora

$$A^+B = Q(\Sigma^+P^*B) = QY,$$

avendo posto $Y \stackrel{\text{def}}{=} \Sigma^+P^*B$. Eseguendo i calcoli, ci si rende conto che la struttura di Σ^+ fa sì che $Y'' = O$. Dunque

$$X = A^+B \quad \Leftrightarrow \quad X = QY \quad \text{con} \quad Y'' = O.$$

Dunque, la pseudoinversa di una matrice può servire a trovare una soluzione di un sistema lineare non quadrato, anche non compatibile (come si vedrà nel paragrafo 7.4). □

Esempio 7.4. Considerando l'esempio 7.2 si ha

$$X = A^+B = (-1 \ -2)^T$$

soluzione del sistema $X = QY$ dove $Y = (y_1, y_2, 0)^T$.

7.3.c La decomposizione polare

Teorema 7.5. *Se $A \in \mathbb{C}^{n,n}$, allora A ammette la decomposizione*

$$A = GR,$$

dove $G \in \mathbb{C}^{n,n}$ è hermitiana e semidefinita positiva, ed $R \in \mathbb{C}^{n,n}$ è unitaria.

Dimostrazione. Sia $A = P\Sigma Q^*$ una decomposizione ai valori singolari di A. Inserendo $I = P^*P$ in tale prodotto si ottiene la fattorizzazione

$$A = (P\Sigma P^*)(PQ^*).$$

Si verifica facilmente che $G = P\Sigma P^* = P\Sigma P^{-1}$ (quindi G è simile a Σ) ed $R = PQ^*$ hanno le proprietà richieste. QED

Si vede facilmente che se $A = GR$, allora $A^+ = R^* G^+$.

Osservazione 7.4. Analogamente, ponendo

$$A = PQ^* Q\Sigma Q^*$$

si conclude che A si può decomporre anche nel seguente modo: $A = RH$ dove H è hermitiana e semidefinita positiva ed R è la matrice unitaria prima considerata.

Si noti che, se $\det A \neq 0$, allora G è definita positiva. La decomposizione del teorema precedente prende il nome di *decomposizione polare*. Questo nome dipende dal fatto che essa è analoga alla scrittura tramite coordinate polari di un numero complesso $z = \rho\, e^{i\theta}$, dove $\rho \in \mathbb{R}_+$ ed $e^{i\theta}$ è un numero complesso di modulo 1 (con $\theta \in [0, 2\pi[$) la cui moltiplicazione è la rotazione del piano $\mathbb{R}^2$ dell'angolo θ.

Quest'analogia può essere formalizzata come segue. Sia

$$\mathcal{C} \stackrel{\text{def}}{=} \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \mid a, b \in \mathbb{R} \right\} \subset \mathbb{R}^{2,2}.$$

Allora l'applicazione lineare

$$f \colon \mathbb{C} \to \mathcal{C},\ a + ib \mapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \tag{7.3.2}$$

è un isomorfismo di spazi vettoriali reali (interpretando $\mathbb{C}$ come spazio vettoriale reale) tale che $f(zw) = f(z)f(w)$. Ossia, f trasforma la moltiplicazione di numeri complessi in moltiplicazione di matrici. Posto $a = \rho\cos\theta$ e $b = \rho\sin\theta$ si ha

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} \rho & 0 \\ 0 & \rho \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

dove $\left(\begin{smallmatrix} \rho & 0 \\ 0 & \rho \end{smallmatrix}\right)$ è una matrice semidefinita positiva (definita positiva se $z \neq 0$) e $\left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$ è una matrice ortogonale (corrispondente alla rotazione di angolo θ).

Osservazione 7.5. Si noti che l'isomorfismo (7.3.2) non commuta con l'isomorfismo (4.3.9). In altri termini, l'isomorfismo (7.3.2) non coincide con l'isomorfismo (4.3.9) sui numeri complessi visti come sottoinsieme dei quaternioni mediante l'inclusione $\mathbb{C} \subset \mathbb{H}$ (4.3.8). Ci si può rendere conto di questo anche osservando che la rotazione del piano complesso indotta da un numero complesso di modulo 1 mediante l'isomorfismo (7.3.2) è *diversa* dalla rotazione indotta dallo stesso numero mediante la parametrizzazione di Cayley–Klein.

7.3.d La radice quadrata di una matrice

Sia A una matrice hermitiana definita positiva. Per il teorema spettrale, $A = U^*DU$, dove D è diagonale con elementi non nulli positivi ed U unitaria. Indicata con Σ_2 la matrice delle radici quadrate dei valori singolari (uguali, in questo caso, agli autovalori di A), quindi $\Sigma_2^2 = \Sigma = D$, si pone

$$\sqrt{A} \stackrel{\text{def}}{=} U^*\Sigma_2 U;$$

$\sqrt{A}$ è detta *radice quadrata* di A. La denominazione è giustificata dal fatto che

$$(\sqrt{A})^2 = (U^*\Sigma_2 U)(U^*\Sigma_2 U) = A.$$

Analogamente si definisce $\sqrt[k]{A} \stackrel{\text{def}}{=} U^*\Sigma_k U$, dove Σ_k è la matrice costruita come Σ_2 con le radici k-esime degli autovalori al posto delle radici quadrate.

Naturalmente, i risultati riportati si possono riformulare in termini di applicazioni lineari.

7.3.e Norma spettrale

Data $A \in \mathbb{C}^{m,n}$, si chiama *norma spettrale* la norma naturale indotta dalla norma vettoriale euclidea, ossia

$$|||A|||_2 \stackrel{\text{def}}{=} \sup_{X \neq O} \frac{\|AX\|_2}{\|X\|_2}, \qquad X \in \mathbb{C}^{n,1}.$$

Teorema 7.6. *Sia $A \in \mathbb{C}^{m,n}$. Allora $|||A|||_2$ è uguale al massimo valore singolare di A, ossia*

$$|||A|||_2 = \sigma_1 = \sqrt{\lambda_1},$$

dove λ_1 è il massimo autovalore di A^*A.

Dimostrazione. Sia $A = P\Sigma Q^*$ una decomposizione ai valori singolari, con Σ come in (7.2.1). Essendo P unitaria si ha $\|PY\|_2 = \|Y\|_2$ per ogni $Y \in \mathbb{C}^{m,1}$, quindi

$$\|AX\|_2 = \|P\Sigma Q^*X\|_2 = \|\Sigma Q^*X\|_2 = \|\Sigma Z\|_2,$$

dove $Z = Q^*X = (z_1, \ldots, z_n)^T$. Al variare di X nell'insieme dei vettori di norma 1, il vettore Z assume tutti i possibili valori nell'insieme dei vettori di norma 1. Dunque

$$|||A|||_2 = \sup\{\|\Sigma Z\|_2 \mid \|Z\| = 1\}.$$

Ma

$$\|\Sigma Z\|_2^2 = \sigma_1^2|z_1|^2 + \cdots + \sigma_r^2|z_r|^2 \leq \sigma_1^2(|z_1|^2 + \cdots + |z_r|^2),$$

L'espressione $\|\Sigma Z\|_2^2$ raggiunge il suo massimo in $Z = (1, 0, \ldots, 0)^T$, pertanto

$$|||A|||_2^2 = \sigma_1^2,$$

da cui la conclusione. QED

Osservazione 7.6. Essendo P e Q matrici unitarie, si ha anche

$$\|A\|_2 = \|\Sigma\|_2 = \left(\sum_{i=1}^{r} \sigma_i^2\right)^{\frac{1}{2}}.$$

Il nome norma spettrale è giustificato osservando che, detto *raggio spettrale* di $M \in \mathbb{C}^{n,n}$ il numero reale

$$\rho(M) \stackrel{\text{def}}{=} \max\{|\lambda| \mid \lambda \text{ autovalore di } M\},$$

risulta

$$|||A|||_2 = \sqrt{\rho(A^*A)}.$$

Si noti che se A è hermitiana, allora $|||A|||_2 = \rho(A)$.

Il seguente teorema mostra che per una data matrice A ogni norma matriciale è collegata al raggio spettrale di A.

Teorema 7.7 (Hirsch). *Siano $A \in \mathbb{C}^{n,n}$, e $|||\cdot|||$ una qualunque norma matriciale che soddisfi la* (1.2.2)*. Allora $\rho(A) \leq |||A|||$.*

Dimostrazione. Sia λ un autovalore di A di modulo massimo ed $X \in \mathbb{C}^{n,1}$ un autovettore di A relativo a λ. Si consideri la matrice $M = (X_1 \dots X_n)$, con $X_i \stackrel{\text{def}}{=} X$ per $i = 1, \dots, n$. Allora

$$AM = (AX_1 \dots AX_n) = \lambda M,$$

perciò

$$\rho(A)\, |||\, M||| = |\lambda|\, |||\, M||| = |||\lambda M||| = |||AM||| \leq |||A|||\ \ |||M|||\,.$$

Poiché $X \neq O$ segue $M \neq O$ e $|||M||| > 0$, da cui la tesi. QED

Ne segue che se A è hermitiana allora $|||A|||_2 \leq |||A|||$.

Il teorema precedente giustifica anche l'interesse della norma spettrale in calcolo numerico.

7.4 Il metodo dei minimi quadrati

Il metodo dei minimi quadrati ha larga applicazione nella soluzione di sistemi lineari i cui coefficienti provengono da dati sperimentali. Per gli errori nelle misurazioni si possono trovare equazioni tra loro incompatibili, mentre teoricamente le equazioni corrispondenti ai dati, non affetti da errori, devono avere una soluzione. La soluzione secondo i minimi quadrati approssima la soluzione "ideale" del sistema non affetto da errori.

Sia $A = (a_{ij}) \in \mathbb{C}^{m,n}$, e si consideri il sistema lineare $AX = B$, con $B \in \mathbb{C}^{m,1}$. Allora, posto $U \stackrel{\text{def}}{=} \mathcal{L}(C_1, \dots, C_n)$, dove $A = (C_1, \dots, C_n)$, il sistema è (si veda [16]):

- **compatibile** se e solo se $B \in U$, o, equivalentemente, $B = B_U$;

- **incompatibile** se e solo se $B \notin U$, o, equivalentemente, $B \neq B_U$,

dove B_U è la proiezione ortogonale del vettore B sul sottospazio vettoriale U. Si introduce la funzione

$$f\colon \mathbb{C}^{n,1} \to \mathbb{R},\ f(X) \stackrel{\text{def}}{=} \|AX - B\|^2, \tag{7.4.3}$$

detta *scarto quadratico medio*. L'idea è di rendere minima la funzione f. Tenendo conto della disuguaglianza di Bessel A.1, i vettori X che minimizzano f sono quelli per cui $AX = B_U$. La ricerca di tali vettori X è detta *metodo dei minimi quadrati*. Il numero $\|B - B_U\|^2$ è detto *minimo scarto quadratico medio*.

L'insieme

$$\mathrm{Min}_A(B) \stackrel{\text{def}}{=} \{X \in \mathbb{C}^n \mid AX = B_U\}$$

è detto *insieme delle soluzioni approssimate* del sistema $AX = B$.

Teorema 7.8.

$$Z \in \mathrm{Min}_A(B) \quad \Leftrightarrow \quad A^*AZ = A^*B.$$

Dimostrazione. Se $Z \in \mathrm{Min}_A(B)$, allora $B - B_U = B - AZ \in U^\perp$. Cioè $B - AZ$ è ortogonale ad ogni vettore $AX \in U$, quindi

$$0 = (AX) \cdot (B - AZ) = (AX)^*(B - AZ) = X^*A^*(B - AZ) = X \cdot A^*(B - AZ).$$

In altre parole, $A^*(B - AZ)$ è ortogonale ad ogni $X \in \mathbb{C}^n$ (quindi anche a sé stesso), per cui

$$A^*(B - AZ) = O \quad \Rightarrow \quad A^*B = A^*AZ.$$

QED

Dal teorema segue che ogni *soluzione approssimata* Z del sistema $AX = B$ è una *soluzione esatta* del sistema lineare

$$A^*AY = A^*B$$

e viceversa. Questo sistema è detto *sistema di equazioni lineari normali*. Perciò ci sarà una sola soluzione approssimata se e solo se A^*A è invertibile, cioè se $\operatorname{rg} A = n$ (per il teorema 7.1).

Chiamasi *soluzione ai minimi quadrati* del sistema lineare $AX = B$ ogni vettore $Z \in \mathrm{Min}_A(B)$ tale che

$$\|Z\| \leq \|X\| \quad \forall X \in \mathrm{Min}_A(B).$$

Indichiamo con $W = \mathcal{L}(\overline{R_1}, \ldots, \overline{R_m})$ lo spazio delle righe della matrice $\overline{A}$. Come è noto (vedi Appendice),

$$W = \mathcal{L}(C_1^*, \ldots, C_m^*),$$

dove C_i^* sono le colonne di A^*. Vale il seguente notevole teorema [32].

Teorema 7.9. *La soluzione ai minimi quadrati del sistema lineare $AX = B$ è unica e coincide con la proiezione X_W su W di un qualunque vettore X appartenente a $\mathrm{Min}_A(B)$. Inoltre essa è data da A^+B.*

Osservazione 7.7. Le matrici $AA^+ \in \mathbb{C}^{m,m}$ e $A^+A \in \mathbb{C}^{n,n}$ sono hermitiane ed idempotenti, quindi matrici di proiezione. Più precisamente, AA^+ proietta un vettore di $\mathbb{C}^m$ su U e A^+A proietta un vettore di C^n su W. Infatti, poiché A^+B è soluzione del sistema $AX = BU$, si ha $AA^+B = BU$. Ebbene A^+ ha l'effetto di mandare prima B su U e poi $B_U = AX$ in X_W, proiezione di X su W. Infatti

$$A^+B = (A^+AA^+)B = A^+(B_U) = A^+AX = X_W.$$

Esempio 7.5. Dati i punti $P_1(0,0)$, $P_2(1,2)$, $P_3(3,4)$, $P_4(4,6)$ (non allineati), si ricerca la retta che approssima in media tali punti nel senso dei minimi quadrati. Se esistesse una retta passante per i punti $P_i(\alpha_i, \beta_i)$, $i = 1,2,3,4$, di equazione $y = a_1x + a_0$, si avrebbe

$$\begin{cases} \alpha_1 a_1 + a_0 = \beta_1, \\ \alpha_2 a_1 + a_0 = \beta_2, \\ \alpha_3 a_1 + a_0 = \beta_3, \\ \alpha_4 a_1 + a_0 = \beta_4, \end{cases} \qquad A = \begin{pmatrix} \alpha_1 & 1 \\ \alpha_2 & 1 \\ \alpha_3 & 1 \\ \alpha_4 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 3 & 1 \\ 4 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 4 \\ 6 \end{pmatrix}.$$

Questo sistema non ha soluzioni perchè i punti non sono allineati. Si cerca il vettore $B_U = a_1C_1 + a_0C_2$ tale che $(B - B_U) \perp C_1, C_2$. Ora $B_U = (a_0, a_1 + a_0, 3a_1 + a_0, 4a_1 + a_0)^T$, e

$$\begin{cases} (B - B_U) \cdot C_1 = 0, \\ (B - B_U) \cdot C_2 = 0 \end{cases} \quad \Rightarrow \quad \begin{cases} 38 - 26a_1 - 8a_0 = 0, \\ 12 - 8a_1 - 4a_0 = 0, \end{cases}$$

da cui si ottiene la soluzione $a_1 = 7/5$, $a_0 = 1/5$ per cui la retta "ottimale" (nel senso dei minimi quadrati) è $y = 7/5x + 1/5$.

Naturalmente l'esercizio si può risolvere anche trovando direttamente A^+B. Ora, poiché $\operatorname{rg} A = 2 = n$, A^+ ha la seguente espressione:

$$A^+ = (A^*A)^{-1}A^* = \frac{1}{20}\begin{pmatrix} -4 & -2 & 2 & 4 \\ 13 & 9 & 1 & -3 \end{pmatrix},$$

come si verifica facilmente. Ne segue $A^+B = (7/5, 1/5)^T$, come si voleva.

Quanto detto per trovare la retta "ottimale" si può ripetere per trovare un polinomio "ottimale" di grado al più n.

Supposto di aver raccolto m coppie di dati (α_i, β_i), $i = 1, \ldots, m$, vogliamo trovare il polinomio di grado al più m che meglio approssima i dati e sia

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0.$$

Imponendo le condizioni $p(\alpha_i) = \beta_i$ si ottiene un sistema lineare di m equazioni nelle $n+1$ incognite $a_n, \ldots, a_0$. La matrice associata al sistema è del tipo di Vandermonde:

$$A = \begin{pmatrix} \alpha_0^n & \alpha_0^{n-1} & \cdots & \alpha_0 & 1 \\ \alpha_1^n & \alpha_1^{n-1} & \cdots & \alpha_1 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_m^n & \alpha_m^{n-1} & \cdots & \alpha_m & 1 \end{pmatrix},$$

con $X^T = (a_n, \ldots, a_0)$, e $B^T = (\beta_1, \ldots, \beta_n)$. Si procede come nel caso generale trovando A^+B. Se $m = n+1$, allora $\det A \neq 0$, e quindi si avrà una sola soluzione non approssimata del sistema (si confronti con l'Appendice).

Osservazione 7.8. Si è visto che la pseudoinversa di una matrice può servire a trovare soluzioni di un sistema lineare non quadrato (osservazione 7.3). Le soluzioni ottenute sono le soluzioni ai minimi quadrati del sistema lineare dato: infatti, basta prendere $U = \mathbb{C}^r$, $U^\perp = \mathbb{C}^{n-r}$ e tener conto di quanto precedentemente esposto. □

7.5 Altri metodi per la soluzione del problema dei minimi quadrati

Metodo $\mathcal{QR}$. Sia $A \in \mathbb{C}^{n,m}$, con $n \geq m$. Sia $\operatorname{rg} A = m$. Si vuole risolvere il problema ai minimi quadrati per un sistema $AX = B$. Si trovi una fattorizzazione $\mathcal{QR}$ di A procedendo come in 5.9, con Q ed R come in (5.9.11). Allora

$$\|AX - B\|_2 = \|QRX - B\|_2 = \|Q(RX - Q^*B)\|_2 = \|RX - C\|_2,$$

dove si è posto $C = Q^*B$. Scrivendo $C = (C_1, C_2)^T$, con $C_1 \in \mathbb{C}^m$ e $C_2 \in \mathbb{C}^{n-m}$, e $R = (R_1, O)^T$, con $R_1 \in \mathbb{C}^{m,m}$, risulta

$$\begin{aligned}\min_{X\in\mathbb{C}^m} \|AX - B\|_2^2 &= \min_{X\in\mathbb{C}^m} (\|R_1X - C_1\|_2^2 + \|C_2\|_2^2) \\ &= \min_{X\in\mathbb{C}^m} (\|R_1X - C_1\|_2^2) + \min_{X\in\mathbb{C}^m} \|C_2\|_2^2 \\ &= \|C_2\|_2^2,\end{aligned}$$

poiché il sistema $R_1X = C_1$ ammette una soluzione e questa rende nulla la norma.

Metodo $\mathcal{LL}^*$. Sia $A \in \mathbb{C}^{n,m}$, con $n \geq m$. Sia $\operatorname{rg} A = m$. Allora, come è facile dimostrare, A^*A è hermitiana e definita positiva, quindi è possibile applicare il metodo $\mathcal{LL}^*$ per risolvere il problema ai minimi quadrati. Il numero di operazioni effettuate è inferiore a quello del precedente metodo, ma diventa uguale se $m = n$.

Osservazione 7.9. In generale non vi è una relazione semplice tra gli autovalori (di una matrice quadrata) ed i suoi valori singolari. In teoria i valori singolari di A si possono trovare risolvendo il problema degli autovalori per la matrice hermitiana A^*A, ma questo procedimento, in pratica, può portare ad una perdita di accuratezza. Una procedura numerica più stabile consiste nel ridurre, mediante trasformazioni ortogonali di tipo Householder, la matrice A ad una matrice bidiagonale e nell'applicare a questa il metodo $\mathcal{QR}$.

7.6 Esercizi di riepilogo

Esercizio 7.4. *Trovare la decomposizione polare di* $A = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix}$.

Esercizio 7.5. *Se $A \in \mathbb{C}^{m,n}$, verificare che*

1. *AA^+ e A^+A sono hermitiane ed idempotenti;*
2. *$AA^+A = A$ e $A^+AA^+ = A^+$.*

Esercizio 7.6. *Indicata con C^+ la pseudo-inversa di C, provare che*

$$CC^+C = C.$$

SOLUZIONE. Se $C = P\Sigma Q^*$ è una decomposizione ai valori singolari di C, la sua pseudoinversa è $C^+ = Q\Sigma^+P^*$. Segue, col simbolismo usato nelle dispense,

$$CC^+ = P\Sigma Q^*(Q\Sigma^+P^*) = P\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}P^*$$

and

$$CC^+C = (CC^+)C = P\Sigma Q^* = C.$$

□

Esercizio 7.7. *Se A è una matrice (hermitiana) definita positiva, mostrare che gli autovalori di A sono i valori singolari di A.*

SOLUZIONE. Se A è hermitiana, allora $A = A^*$ e quindi $A^*A = A^2$. Se λ è un autovalore per A, allora λ^2 è un autovalore per A^*A e $\sqrt{\lambda^2} = |\lambda| = \sigma$ è un valore singolare per A. Ma A è definita positiva, quindi $\lambda > 0$ e $\sigma = \lambda$. □

Esercizio 7.8. *Si consideri la matrice*

$$A = \frac{1}{5}\begin{pmatrix} 2 & 6 \\ 6 & 7 \end{pmatrix}.$$

1. *Determinare una decomposizione a valori singolari di A.*
2. *Determinare una forma polare di A.*

SOLUZIONE. 1. Poiché $A = A^T$, i valori singolari di A coincidono con i suoi autovalori presi in valore assoluto. Gli autovalori di A sono $\lambda_1 = 11/5$, $\lambda_2 = -2/5$, quindi $\sigma_1 = 11/5$, $\sigma_2 = 2/5$. D'altra parte gli autovalori di $A^T A = A^2$ sono $\mu_1 = 121/25$ e $\mu_2 = 4/25$.

$$V(\mu_1) = \{(x,y) \mid y = 3/2x\} = \mathcal{L}(Q_1), \quad \text{dove } Q_1 = \frac{1}{\sqrt{13}}\begin{pmatrix} 2 \\ 3 \end{pmatrix},$$
$$V(\mu_2) = \{(x,y) \mid y = -3/2x\} = \mathcal{L}(Q_2), \quad \text{dove } Q_2 = \frac{1}{\sqrt{13}}\begin{pmatrix} 3 \\ -2 \end{pmatrix}.$$

Ora, $P_1 = 1/\sigma_1 AQ_1$ e $P_2 = 1/\sigma_2 AQ_2$.

Poiché σ_1 è autovalore per A, allora $AQ_1 = \sigma_1 P_1$, dunque $P_1 = Q_1$.

Poiché σ_2 è autovalore per A, allora $AQ_2 = \sigma_2 P_2$, dunque $P_2 = -Q_2$.

Quindi

$$A = P\Sigma Q^T,$$
$$P = \frac{1}{\sqrt{13}}\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 11/5 & 0 \\ 0 & 2/5 \end{pmatrix}, \quad Q = \frac{1}{\sqrt{13}}\begin{pmatrix} 2 & 3 \\ 3 & -2 \end{pmatrix}.$$

2. $A = P\Sigma Q^T = (P\Sigma P^T)(PQ^T)$ dà una forma polare ponendo

$$P\Sigma P^T = G = \frac{1}{65}\begin{pmatrix} 62 & 54 \\ 54 & 107 \end{pmatrix} \quad \text{hermitiana e def. positiva,}$$
$$PQ^T = R = \frac{1}{13}\begin{pmatrix} -5 & 12 \\ 12 & -5 \end{pmatrix} \quad \text{unitaria.}$$

□

Esercizio 7.9. *Siano $A \in \mathbb{C}^{n,n}$ e $|||\cdot|||$ una qualunque norma matriciale indotta.*

1. *Si dimostri che vale l'implicazione*

$$\lim_{k\to+\infty} |||A^k||| = 0 \quad \Rightarrow \quad \rho(A) < 1.$$

2. *Sia $|||A||| < 1$. Si dimostri che la matrice $I + A$ è invertibile e che vale la disuguaglianza*

$$|||(I + A)^{-1}||| \leq \frac{1}{1 - |||A|||}.$$

Soluzione.

1. Risulta $|||A^k||| \geq \rho(A^k) = \rho(A)^k$ (l'ultima uguaglianza viene dal teorema di Schur).
2. Poiché $\rho(A) \leq |||A|||$ risulta $\rho(A) < 1$. Ma supponendo che 0 sia un autovalore di $I + A$ si giunge ad una contraddizione (fornire i dettagli), dunque $I + A$ è invertibile. Dalla relazione $(I + A)(I + A)^{-1} = I$ si ottiene la disuguaglianza

$$|||(I + A)^{-1}||| \leq 1 + |||A||| \, |||(I + A)^{-1}|||,$$

da cui la tesi.

□

CAPITOLO 8

MATRICI POLINOMIALI

In questo capitolo daremo le prime nozioni sulle matrici polinomiali, largamente usate nello studio dei sistemi di controllo con più ingressi ed uscite.

8.1 Premesse

Definizione 8.1. Sia $\mathbb{K}$ un campo (in particolare $\mathbb{K}$ può essere uno dei campi $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Z}_p$).

Una *matrice polinomiale* $A(z)$ è una matrice del tipo

$$A(z) = A_0 + A_1 z + \cdots + A_h z^h,$$

dove $A_p \in \mathbb{K}^{m,n}$ per $p = 0, 1, \ldots, h$.

Alternativamente si può definire una matrice polinomiale come un elemento $A(z) \in \mathbb{K}^{m,n}[z]$. Ponendo $A_p = (a_{ij}^{(p)})$ si ha

$$a_{ij}(z) = a_{ij}^{(0)} + a_{ij}^{(1)} z + \cdots + a_{ij}^{(h)} z^h,$$

cioè gli elementi della matrice $A(z)$ sono polinomi in z, in simboli $a_{ij}(z) \in \mathbb{K}[z]$. In altre parole, vale l'isomorfismo canonico

$$\mathbb{K}^{m,n}[z] \simeq \mathbb{K}[z]^{m,n},$$

dove $\mathbb{K}[z]$ è l'anello dei polinomi con coefficienti in $\mathbb{K}$.

Esempio 8.1. Si consideri la matrice $A(z) \in \mathbb{R}^{2,2}[z]$ così definita

$$A(z) = A_0 + A_1 z$$
$$A_0 = \begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix}, \qquad A_1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Allora

$$A(z) = \begin{pmatrix} z-1 & z \\ z+1 & z-1 \end{pmatrix} \in \mathbb{R}[z]^{2,2}.$$

Una matrice $A(z) \in \mathbb{K}^{n,n}[z]$ è *singolare* se il suo determinante è il polinomio nullo. In caso contrario si dice non singolare.

Esempio 8.2.

- La matrice del precedente esempio è non singolare, poiché $\det A(z) = 1 - 3z$ non è il polinomio nullo (anche se la matrice $A(1/3)$ ha determinante nullo).
- La matrice
$$B(z) = \begin{pmatrix} z-1 & z+1 \\ z-1 & z+1 \end{pmatrix}$$
è ovviamente singolare: $\det B(z) = 0$ identicamente (cioè per ogni z).

Una matrice $A(z) \in \mathbb{K}[z]^{m,n}$ ha *rango* r se i determinanti dei suoi minori di ordine $r+1$ sono polinomi identicamente nulli, mentre esiste almeno un minore di ordine r con determinante non identicamente nullo.

Indichiamo con $D_k(z)$ il massimo comun divisore[1] (monico) di tutti i determinanti dei minori di ordine k di $A(z)$. Si prova che $D_k(z)$ divide $D_{k+1}(z)$, in simboli $D_k(z) \mid D_{k+1}(z)$. I polinomi
$$d_k(z) = \frac{D_k(z)}{D_{k-1}(z)}, \qquad k \geq 1 \tag{8.1.1}$$
si chiamano *fattori invarianti di* $A(z)$. Per definizione si pone $D_0(z) = 1$.

Esempio 8.3. Si consideri il seguente blocco di Jordan di ordine r
$$A(z) = \begin{pmatrix} z-z_0 & 1 & 0 & \dots & 0 \\ 0 & z-z_0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & z-z_0 & 1 \\ 0 & \dots & 0 & 0 & z-z_0 \end{pmatrix}$$
I minori di ordine 1 sono 0, 1, $z - z_0$, quindi $d_1(z) = 1$. Così procedendo risulta $d_k(z) = 1$ per $k < r$ e $d_r(z) = (z - z_0)^r$.

La nozione di rango di una matrice si presta ad una immediata interpretazione in termini di "spazi vettoriali", anche se $\mathbb{K}[z]$ è un anello ma non un campo. Se pensiamo alle colonne della matrice $A(z) \in \mathbb{K}[z]^{m,n}$ come a vettori in $\mathbb{K}[z]^m$, il rango di $A(z)$ rappresenta la dimensione del sottospazio di $\mathbb{K}[z]^m$ generato dalle colonne di $A(z)$; analogamente si prova che $\mathrm{rg}(A(z))$ è uguale alla dimensione del sottospazio vettoriale di $\mathbb{K}[z]^n$ generato dalle righe di $A(z)$.

Esempio 8.4. Sia
$$A(z) = \begin{pmatrix} z+1 & z & z+2 \\ z^2-1 & z^2-z & z^2+z-2 \end{pmatrix}.$$

[1] Ricordiamo che $d(x)$ è massimo comun divisore di $f(x)$ e $g(x)$ se $d(x)|f(x)$, $d(x)|g(x)$ e per ogni $d'(x)$ tale che $d'(x)|f(x)$, $d'(x)|g(x)$ si ha $d'(x)|d(x)$. L'algoritmo di Euclide permette di trovare i massimi comun divisori.

Si prova facilmente che i minori di ordine 2 hanno determinante identicamente nullo, quindi il rango di $A(z)$ è 1 in $\mathbb{R}[z]$. Infatti le due righe sono proporzionali in $\mathbb{R}[z]$ (il coefficiente è $z-1$). Ovviamente le due righe non sono proporzionali rispetto ad $\mathbb{R}$.

Le colonne di $A(z)$ generano un sottospazio di $\mathbb{R}[z]^2$ di dimensione 1 su $\mathbb{R}[z]$: un generatore è il vettore $(1\ z-1)^T$. Infatti

$$\begin{pmatrix} z+1 \\ z^2-1 \end{pmatrix} = (z+1)\begin{pmatrix} 1 \\ z-1 \end{pmatrix},\ \begin{pmatrix} z \\ z^2-z \end{pmatrix} = z\begin{pmatrix} 1 \\ z-1 \end{pmatrix},\ \begin{pmatrix} z+2 \\ z^2+z-2 \end{pmatrix} = (z+2)\begin{pmatrix} 1 \\ z-1 \end{pmatrix}.$$

Le righe di $A(z)$ generano un sottospazio di $\mathbb{R}[z]^3$ di dimensione 1 su $\mathbb{R}[z]$: un generatore è il vettore $(z+1\ z\ z+2)$.

Definizione 8.2. Si definisce *matrice polinomiale elementare* una matrice di uno dei tre tipi seguenti (si confronti con (5.2.8), (5.2.9), (5.2.10)):

$$\tilde{E}_1(z) = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & 1 \end{pmatrix} = M_i(\lambda), \tag{8.1.2}$$

$$\tilde{E}_2(z) = \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} = \Pi_{ij}, \tag{8.1.3}$$

$$\tilde{E}_3(z) = \begin{pmatrix} 1 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & p(z) & \dots & 1 \end{pmatrix} = S_{ij}(p(z)), \tag{8.1.4}$$

dove $\lambda \in \mathbb{K} \setminus \{0\}$ e $p(z) \in \mathbb{K}[z]$.

Le matrici polinomiali elementari hanno alcune importanti proporietà. Innanzitutto, per $i = 1, 2, 3$, $\det \tilde{E}_i$ è una costante non nulla quindi $\tilde{E}_i$ è invertibile per ogni z e la sua inversa è dello stesso tipo.

Se $A(z) \in \mathbb{K}[z]^{p,m}$ ed $\tilde{E}_i(z) \in \mathbb{K}[z]^{m,m}$, allora $A(z)\tilde{E}_i(z) \in \mathbb{K}[z]^{p,m}$ è una matrice ottenuta da $A(z)$ mediante trasformazioni elementari sulle sue colonne; analogamente se

$\tilde{E}_i(z) \in \mathbb{K}[z]^{p,p}$ allora $\tilde{E}_i(z)A(z) \in \mathbb{K}[z]^{p,m}$ è una matrice ottenuta da $A(z)$ mediante trasformazioni elementari sulle sue righe.

8.2 Forma canonica di Smith

Siano $A(z)$ e $B(z)$ due matrici polinomiali in $\mathbb{K}[z]^{p,m}$; diciamo che $A(z)$ è *equivalente* a $B(z)$, in simboli

$$A(z) \sim B(z)$$

se $A(z)$ può essere ottenuta da $B(z)$ attraverso una successione di operazioni elementari realizzate sia sulle righe sia sulle colonne. In altre parole, $A(z) \sim B(z)$ se esistono matrici elementari $E_i \in \mathbb{K}[z]^{p,p}$, $E'_j \in \mathbb{K}[z]^{m,m}$ tali che

$$E_k(z) \cdots E_1(z)B(z)E'_1(z) \cdots E'_h(z) = A(z).$$

La relazione introdotta è una relazione d'equivalenza nell'insieme delle matrici $\mathbb{K}[z]^{p,m}$, che viene così suddivisa in classi di equivalenza in $\mathbb{K}[z]^{p,m}/\sim$. Come di consueto si desidera trovare un rappresentante distinto per ogni classe di equivalenza che sia rappresentativo. Il seguente teorema assicura l'esistenza di una matrice (rettangolare) di tipo diagonale in ogni classe d'equivalenza.

Teorema 8.1 (Forma canonica di Smith). *Ogni matrice $A(z) \in \mathbb{K}[z]^{p,m}$ è equivalente ad una matrice $D(z) \in \mathbb{K}[z]^{p,m}$ di tipo diagonale, dove*

$$D(z) = \operatorname{diag}(d_1(z), \ldots, d_r(z), 0, \ldots, 0),$$

essendo $r = \operatorname{rg}(A(z))$ e $d_k(z)$ i fattori invarianti di $A(z)$.

Gli elementi r e $d_k(z)$ $(k = 1, \ldots, r)$ sono univocamente individuati, per cui è unica la forma di Smith di una data matrice. Però le matrici che intervengono nella decomposizione non sono univocamente individuate.

Esempio 8.5. Sia

$$A(z) = \begin{pmatrix} z-1 & z \\ z+1 & z-1 \end{pmatrix}.$$

I minori di ordine 1 di $A(z)$ sono $z-1$, z, $z+1$, dunque $D_1(z) = 1$. Il minore di ordine 2 ha determinante $1-3z$, dunque $D_2(z) = 1/3 - z$. Dunque

$$d_1(z) = \frac{D_1(z)}{D_0(z)} = 1, \quad d_2(z) = \frac{D_2(z)}{D_1(z)} = 1/3 - z,$$

e la forma canonica di Smith è

$$D(z) = \begin{pmatrix} 1 & 0 \\ 0 & 1/3 - z \end{pmatrix}.$$

8.3 Matrici unimodulari

Definizione 8.3. Una matrice $U(z) \in \mathbb{K}[z]^{m,m}$ si dice *unimodulare* se è invertibile e la sua inversa $U(z)^{-1} = U^{-1}(z)$ è a sua volta una matrice polinomiale, cioè anche $U^{-1}(z) \in \mathbb{K}[z]^{m,m}$.

Le matrici unimodulari sono quindi gli elementi invertibili nell'anello non commutativo delle matrici quadrate polinomiali; cioè svolgono il ruolo di $+1$ e -1 dell'anello $\mathbb{Z}$.

Valgono le seguenti caratterizzazioni.

Teorema 8.2. *Sia $U(z) \in \mathbb{K}[z]^{m,m}$. Sono equivalenti le seguenti affermazioni:*

1. *$U(z)$ è una matrice unimodulare;*
2. *$\det U(z)$ è un polinomio costante non nullo;*
3. *la forma canonica di Smith di $U(z)$ è I_m;*
4. *$U(z)$ è esprimibile come prodotto di matrici elementari.*

Dimostrazione.

$1 \Rightarrow 2$ Da $U(z)U(z)^{-1} = I$ risulta $\det U(z) \cdot \det U(z)^{-1} = 1$. Essendo gli unici elementi invertibili in $\mathbb{K}[z]$ le costanti non nulle, si ha la tesi.

$2 \Rightarrow 3$ Considerando la forma canonica di Smith, $U(z)$ è l'unico minore di ordine m, quindi $\det U(z)$ coinciderà a meno di una costante moltiplicativa con $d_1(z) \cdots d_m(z)$. Ma essendo $\det U(z)$ una costante non nulla e $d_i(z)$ polinomi monici si ha $d_i(z) = 1$ per tutti gli i.

$3 \Rightarrow 4$ Essendo

$$E_k(z) \cdots E_1(z) U(z) E'_1(z) \cdots E'_h(z) = I_m,$$

risulta

$$U(z) = E_1(z)^{-1} \cdots E_k(z)^{-1} E'_h(z)^{-1} \cdots E'_1(z)^{-1},$$

da cui la conclusione poiché le inverse di matrici elementari sono elementari.

$4 \Rightarrow 1$ Poiché il determinante di matrici elementari è una costante, anche le inverse sono matrici polinomiali e quindi $U(z)^{-1}$ è polinomiale.

QED

Poiché ogni matrice $A(z)$ è equivalente ad una matrice $D(z)$ segue il corollario.

Corollario 8.1. *Data una matrice $A(z) \in \mathbb{K}[z]^{p,m}$, esiste una coppia di matrici unimodulari $U(z) \in \mathbb{K}[z]^{p,p}$ e $V(z) \in \mathbb{K}[z]^{m,m}$ tali che*

$$D(z) = U(z) A(z) V(z).$$

Ovviamente, poiché $U(z)^{-1}$ e $V(z)^{-1}$ sono unimodulari si ha $U(z)^{-1}D(z)V(z)^{-1} = A(z)$.

Esercizio 8.1. *Dedurre che la seguente matrice è unimodulare, verificando direttamente le proprietà 1–4 del teorema precedente:*

$$A(z) = \begin{pmatrix} 1 & -z & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & -3z & 1 & z \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

CAPITOLO 9

ESPERIMENTI NUMERICI CON `OCTAVE`

In questo capitolo saranno approfonditi mediante esempi numerici alcuni temi della teoria sviluppata in precedenza. Il programma usato è `octave`, che può essere liberamente copiato dal sito `http://www.octave.org/`. In particolare è suggerita ai lettori un'analisi del codice sorgente di `octave` per quanto riguarda gli argomenti trattati nel corso. Ci si può rendere conto di come le librerie usate per implementare gli algoritmi presentati vengano quasi tutte dal pacchetto `lapack`, reperibile in `http://netlib.org`.

Esempio 9.1. Scrivere una funzione che crei una matrice triangolare superiore di numeri casuali compresi tra 0 e 100 di ordine n.

SOLUZIONE.

```
function a=mat_sup(n)
a=zeros(n);
   for j=1:n
     for i=1:j
       a(i,j)=floor(100*rand(1,1));
     end
   end
endfunction
```

Problema 9.1. *Scrivere una funzione che crei una matrice tridiagonale di ordine n che ha 2 sulla diagonale principale e -1 sulle altre.*

Esempio 9.2. Scrivere una funzione di due variabili che risolva un sistema con matrice triangolare superiore. Iniziare la funzione con un test sulle variabili.

SOLUZIONE.

```
function x=tri_sup(a,b)
% Controlla le dimensioni di a:
[n,m] = size(a);
if m~= n
    printf('A non e'' quadrata')
    return
end
for i=1:n
    for j=1:i-1
```

```
            if a(i,j) ~=0
                printf('A non e'' triangolare superiore')
                return
            end
        end
    end
    % Controlla le dimensioni di b:
    if length(b) ~= n
        printf('B non e'' compatibile')
        return
    end
    % Risolve il sistema
    x(n) = b(n)/a(n,n);
    for i=n-1:-1:1
        somma=b(i);
        for j=i+1:n
            somma = somma - a(i,j)*x(j);
        end
        x(i) = somma/a(i,i);
    end
```

Esempio 9.3. Scrivere un programma che confronti la velocità del metodo di fattorizzazione $\mathcal{LU}$ con il metodo di fattorizzazione $\mathcal{QR}$ usando matrici quadrate casuali di dimensione sufficientemente grande.

SOLUZIONE.

```
index=0;
for n=10:100:1000
    index=index+1;
    a=rand(n,n);
    t0=cputime;
    [l,u]=lu(a);
    printf(" fattorizzazione LU per n=%d :",n)
    cputime-t0
    t1=cputime;
    [q,r]=qr(a);
    printf(" fattorizzazione QR per n=%d :",n)
    cputime-t1
end
```

Problema 9.2. *Scrivere un programma che confronti la velocità di risoluzione di un sistema triangolare superiore usando, rispettivamente, la funzione dell'esercizio precedente e la funzione di 'divisione matriciale' di* OCTAVE.

Problema 9.3. *Scrivere una funzione di due variabili (matrice incompleta e termine noto) che risolva un sistema con matrice triangolare* inferiore. *Iniziare la funzione con un test sulle variabili.*

Esempio 9.4. Scrivere una funzione di due variabili (matrice incompleta e termine noto) che risolva un sistema lineare a matrice quadrata con il metodo di eliminazione. Alla fine, si usi la funzione 'tri_sup'.

Si applichi la funzione al sistema lineare $AX = B$ dove

$$A = \begin{pmatrix} -2 & 4 & -1 & -1 \\ 4 & -9 & 0 & 5 \\ -4 & 5 & -5 & 5 \\ -8 & 8 & -23 & 20 \end{pmatrix}, \quad B = \begin{pmatrix} 12 \\ -32 \\ 3 \\ 13 \end{pmatrix}.$$

SOLUZIONE.

```
function x=ris_sistema(a,b)
%
n=size(a,1);
a=[a,b]
% Test per l'esistenza di un' unica soluzione
%
for k=1:n-1,
 rigaprimo=k;
 p(k)=k;
 if a(k,k) == 0
  for i=k+1:n,          % scelta pivot (parziale)
   if a(i,k) != 0,
    rigaprimo = i;
   else
   printf('Il sistema non ha un'unica soluzione, passo k=%d',k)
   return
   end
  end
 end
%
 if k != rigaprimo,      % Eventuale scambio righe
  for j=1:n+1,
   tmp=a(k,j); a(k,j)=a(rigaprimo,j); a(rigaprimo,j)=tmp;
  end
 end
 for i=k+1:n,           % Calcolo dei moltiplicatori e eliminazione
  aik=a(i,k)/a(k,k);
  a(i,k)=aik;          % memorizza il moltiplicatore
  for j=k+1:n+1,
   a(i,j)=a(i,j)-aik*a(k,j);
  end
 end
end
%
```

```
% Copia della parte tr. sup. in U
%
utemp=a(:,1:n);
U=triu(utemp,0);
%
% Soluzione con tri_sup
x=tri_sup(U,a(:,n+1));
% endfunction
```

Si noti che il programma non è in grado di distiinguere tra il caso in cui ci sono infinite soluzioni ed il caso in cui il sistema non è compatibile.

Problema 9.4. *Scrivere una funzione che, migliorando la precedente, sia in grado di dire se il sistema ha infinite soluzioni o nessuna.*

Problema 9.5. *Scrivere una funzione di una variabile (matrice quadrata) che calcoli il determinante con il metodo di riduzione ad una matrice triangolare. Si confronti il tempo di esecuzione con il tempo di esecuzione della funzione '**det**' e con il tempo di calcolo del determinante con il metodo di Cramer.*

Problema 9.6. *Scrivere una funzione di una variabile che calcoli l'inversa di una matrice quadrata con il metodo di Gauss–Jordan. Si confronti il tempo di esecuzione con il tempo di esecuzione della funzione '**inv**'.*

Esempio 9.5. Scrivere una funzione che calcola la fattorizzazione $\mathcal{LU}$ di una matrice quadrata (senza pivot).

SOLUZIONE.

```
%   La funzione termina con un errore
%   se c'e' un pivot < EPS*NORM(A)
function [l,u] = elleu(a)
nor=norm(a);
% Controlla le dimensioni di A:
[n,m] = size(a);
if m != n
    display('A non e'' quadrata')
    return
end
l=-eye(n);
for k=1:n-1
    if abs(a(k,k)) < eps*nor
        display('Pivot troppo piccolo')
        return
    end
    l(k+1:n,k) = -a(k+1:n,k)/a(k,k);
    a(k+1:n,k+1:n) = a(k+1:n,k+1:n)+l(k+1:n,k)*a(k,k+1:n);
end
% Immagazzina i risultati nelle matrici l e u:
l = -l;
```

```
u = zeros(n);
for i=1:n
    for j=i:n
        u(i,j) = a(i,j);
    end
end
```

Esempio 9.6. Calcolare l'inversa della matrice di Hilbert $H = (h_{ij})$, dove

$$h_{ij} = \frac{1}{i+j-1}$$

con i metodi $\mathcal{LU}$, $\mathcal{QR}$, $\mathcal{SVD}$. Valutare l'errore commesso con '**norm**(H∗**inv**(H)−I)' nei tre casi.

SOLUZIONE.

```
% ESERCIZIO: calcolare H=hilb(n) per n=5 e n=10,
%
% SOLUZIONE:
%
for n=[5 10]
    h=hilb(n);
    [l,u] = lu(h);
    [q,r] = qr(h);
    [uu,s,vv]= svd(h);
    % dimensiona le tre inverse
    hlu = zeros(n);
    hqr = zeros(n);
    hsvd = zeros(n);
    % costruisce l'inversa per colonne in ognuno dei casi
    for j = 1:n
        % costruisce il termine noto, ej
        ej = zeros(n,1);
        ej(j) = 1;
        hlu(:,j) = u\(l\ej);
        hqr(:,j) = r\(q'*ej);
        hsvd(:,j) = vv*(s\ (uu'*ej));
    end
    fprintf('Valore di N %3.0f \n',n);
    disp('Errore nella fattorizzazione LU');
    n1=norm(h*hlu -eye(n));
    n2=norm(hlu*h -eye(n));
    fprintf('%e %e \n',n1,n2);
    disp('Errore nella fattorizzazione QR');
    n1=norm(h*hqr -eye(n));
    n2=norm(hqr*h -eye(n));
    fprintf('%e %e \n',n1,n2);
```

```
    disp('Errore nella fattorizzazione SVD');
    n1=norm(h*hsvd -eye(n));
    n2=norm(hsvd*h -eye(n));
    fprintf('%e %e \n',n1,n2);
end
```

Esempio 9.7. Stimare numericamente l'andamento del rango della matrice di Hilbert.

SOLUZIONE.

```
for n=1:100
    a=hilb(n);
    r(n)=rank(a)
end
r
```

Si osservi che, in teoria, la matrice di Hilbert è non singolare, ma il rango 'percepito' dal programma è sensibilmente inferiore. È necessario, quindi, aumentare la precisione del calcolo in dipendenza dal tipo di dato.

Problema 9.7. *Chiamare la fattorizzazione $\mathcal{LU}$ di*

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

Perché L non è triangolare?

Problema 9.8. *Scrivere una funzione che calcoli la matrice Q di Householder che trasforma un dato vettore V in un vettore che ha tutte le componenti successive alla prima nulle.*

Problema 9.9. *Data la matrice*

$$A = \begin{pmatrix} 2 & 1 & -1 & 2 \\ 1 & 1 & 3-i & 1+i \\ -1 & 3+i & 0 & i \\ 2 & 1-i & -i & 0 \end{pmatrix}$$

trovare la matrice $Q^{(1)}$ con la funzione dell'esercizio precedente. Dire se esiste una fattorizzazione $\mathcal{LL}^*$ di A. Trovare (se esiste) una fattorizzazione $\mathcal{LL}^*$ di A^*A.

Esempio 9.8. Calcolare il rango di una matrice usando il metodo $\mathcal{QR}$.

SOLUZIONE. Si usi la seguente tolleranza:

```
% Sintassi: r=rango(a,tol)
% tol: precisione di calcolo richiesta.
function r=rango(a,tol)
if nargin < 2
    tol=eps*norm(a,1);
```

end

Si noti che la funzione '**qr**' calcola le matrici Q ed R a meno di una permutazione Π che organizza le righe di R in modo da avere gli elementi della diagonale decrescenti in modulo. Si tenga conto del fatto che R sarà a scalini nel caso di rango inferiore al massimo.

Esempio 9.9. Date le matrici

$$M = \begin{pmatrix} 1 & 2 \\ 1 & 3 \\ 2 & 1 \\ 3 & 3 \\ 6 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 2 \\ 1 \\ 3 \\ 4 \\ 3 \end{pmatrix},$$

risolvere con il metodo $\mathcal{QR}$ il problema ai minimi quadrati.

Soluzione.

```
function x=min_quad(a,b)
% Controlla le dimensioni di a:
[n,m] = size(a);
p=length(b);
if p != n
  printf('Termine noto non valido')
  return
end
if m > n
    printf('AX=B non ammette un'unica soluzione ai m.q.')
    return
end
%
[q,r]=qr(a);
c=zeros(n,1);
c=q'*b
for i=1:m
 for j=i:m
  r1(i,j)=r(i,j);
 end
end
tn=c(1:m)
x=r1\tn;
distanza=c(m+1:n);
d=norm(distanza);
printf("Minima distanza dalla soluzione: %d",d)
endfunction
```

Problema 9.10. *Risolvere lo stesso problema usando la fattorizzazione* $\mathcal{LL}^*$.

Esempio 9.10. Scrivere un programma che calcoli gli autovalori di una matrice usando il metodo iterativo $\mathcal{QR}$.

Soluzione. Bisogna creare la successione delle iterate $\mathcal{QR}$. Non sono state studiate le stime per la convergenza del metodo $\mathcal{QR}$; si procede in maniera diretta.

```
function x=aval_qr(a,iter)
% Controlla le dimensioni di a:
[n,m] = size(a);
if m != n
    printf('Bocciato :-)')
    return
end
%
x=zeros(n,1);
for i=1:iter
 [q,r]=qr(a);
 a=r*q;
 if norm(diag(a)-x)< eps
   printf("Tolleranza raggiunta al passo %d\n",i)
  return
 end
 x=diag(a);
end
endfunction
```

Problema 9.11. *Usando il precedente programma, scrivere una funzione che calcoli i valori singolari di una matrice non quadrata.*

Esempio 9.11. Scrivere una funzione che calcoli i cerchi di Gershgorin per una data matrice quadrata, per righe o per colonne, e li disegni.

Soluzione.

```
function [r,c]=gersh(a)
% Controlla le dimensioni di a:
[n,m] = size(a);
if m != n
    printf('Bocciato :-)')
    return
end
%
r=zeros(n,1);c=diag(a);
for i=1:n
 r(i)=sqrt(a(i,:)*a(i,:)'-(abs(a(i,i)))^2);
end
hold on
t =(0:0.1:6.3)'; X=cos(t);Y=sin(t);
for i=1:n
```

```
 for tt=1:64
  data(tt ,:) = [r(i)*X(tt)+real(c(i)),r(i)*Y(tt)+imag(c(i))];
 end
 gplot data
end
endfunction
```

Problema 9.12. *Si può usare il procedimento seguente per la riduzione del tempo di calcolo degli autovalori con il metodo seguente. Sia $A \in \mathbb{C}^{n,n}$ non singolare. Posto $A^{(1)} = A$, sia $P_1 \in \mathbb{C}^{n-1,n-1}$ la matrice elementare di Householder tale che*

$$P_1 \begin{pmatrix} a_{21} \\ a_{31} \\ \vdots \\ a_{n1} \end{pmatrix} = \begin{pmatrix} \alpha \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Si ponga

$$Q^{(1)} \stackrel{def}{=} \begin{pmatrix} 1 & O \\ O & P_1 \end{pmatrix}.$$

Allora, si dimostra facilmente che la matrice $Q^{(1)^*} A^{(1)} Q^{(1)}$ ha gli elementi della prima colonna con indice di riga maggiore di 2 nulli. Iterando il procedimento, si ottiene una matrice di Hessenberg superiore, ossia una matrice $A^{(n-2)}$ tale che $a_{ij}^{(n-2)} = 0$ se $i > j+1$. La matrice $A^{(n-2)}$ ha, ovviamente, gli stessi autovalori della matrice di partenza.

Si noti che, se A è hermitiana, allora $A^{(n-2)}$ è hermitiana.

Dimostrare numericamente che il tempo necessario alla ricerca degli autovalori di una matrice A è inferiore se questa è preventivamente ridotta in forma di Hessenberg superiore.

Esempio 9.12. Questo esempio è stato fornito dal Prof. G. Ricci e dall'Ing. F. Bandiera. Si tratta di trovare gli autovalori di una matrice C di segnali ed i corrispondenti autovettori. Questi dati corrispondono rispettivamente agli utenti in un dato dominio UMTS ed al loro segnale. La base teorica di questo esempio si trova in

Il programma. La parte interessante ai fini del corso di calcolo matriciale è quella dove si calcolano gli autovalori della matrice C. SI noti che questo esempio è eseguito su una matrice di codici data (CodiciGold31).

```
% Processing Gain
Q=31;
% Numero di utenti
K=15;
% Numero di intervalli di simbolo
% su cui valutare la matrice di correlazione campionaria
M=400;
% Rapporto segnale-rumore
SNR_dB=10;
```

```
% Multiple Access Interference
MAI_dB=0;

% Inizializzazione dei generatori di numeri casuali
% randn('state',sum(100*clock));
% rand('state',sum(100*clock));

% Codici di spreading (PN)
load CodiciGold31;
S=Gold(:,1:K);

% Matrice delle ampiezze
A=zeros(K,K);
% Ampiezza dell'utente da demodulare
A(1,1)=1;
% Ampiezza degli altri utenti
A(2:K,2:K)=sqrt(10^(0.1*MAI_dB)).*eye(K-1);
% Varianza dei campioni di rumore
Sigma_quadro=(A(1)^2)*Q*(10.^((-0.1)*SNR_dB));

%%%%% Inizio Costruzione della matrice dati %%%%%
% Bit dei vari utenti sugli M intervalli
B=sign(randn(K,M));

% segnali trasmessi
T=S*A*B;

% rumore termico
N=sqrt(Sigma_quadro/2)*(randn(Q,M)+sqrt(-1)*randn(Q,M));

% matrice dati con rumore
R=T+N;
%%%%% Fine Costruzione della matrice dati %%%%%

% matrice di correlazione campionaria
C=(1/M)*R*R';

% estrazione degli autovalori ed ordinamento
L=eig(C);
Aval=flipud(sort(L))
```

APPENDICE

A.1 Prodotti scalari hermitiani

Definizione A.1. Sia V uno spazio vettoriale sul campo $\mathbb{C}$. Si dice che una funzione

$$\beta\colon V \times V \to \mathbb{C}$$

è un *prodotto hermitiano* se:

1. $\beta(\vec{x} + \vec{x'}, \vec{y}) = \beta(\vec{x}, \vec{y}) + \beta(\vec{x'}, \vec{y})$ per ogni $\vec{x}$, $\vec{x'}$, $\vec{y} \in V$;
2. $\beta(\vec{x}, \vec{y} + \vec{y'}) = \beta(\vec{x}, \vec{y}) + \beta(\vec{x}, \vec{y'})$ per ogni $\vec{x}$, $\vec{y}$, $\vec{y'} \in V$;
3. $\beta(\alpha\vec{x}, \vec{y}) = \overline{\alpha}\beta(\vec{x}, \vec{y})$ per ogni $\vec{x}$, $\vec{y} \in V$ e $\alpha \in \mathbb{C}$;
4. $\beta(\vec{x}, \alpha\vec{y}) = \alpha\beta(\vec{x}, \vec{y})$ per ogni $\vec{x}$, $\vec{y} \in V$ e $\alpha \in \mathbb{C}$;
5. $\beta(\vec{x}, \vec{y}) = \overline{\beta(\vec{y}, \vec{x})}$.

Si noti che un prodotto hermitiano *non è* bilineare rispetto $\mathbb{C}$ per la presenza del coniugio in 3. Una funzione $\beta\colon V \times V \to \mathbb{C}$ che soddisfi le proprietà 1-4 si dice *forma sesquilineare*. La proprietà 5 dice che β è *autoconiugata*.

Naturalmente, $\beta(\vec{x}, \vec{x}) \in \mathbb{R}$ per ogni $\vec{x} \in V$. Ne segue che si possono suddividere i prodotti hermitiani in prodotti hermitiani *(semi)definiti positivi*, *(semi)definiti negativi*, *indefiniti* a seconda del segno di $\beta(\vec{x}, \vec{x})$ come per le forme bilineari simmetriche reali [16].

Definizione A.2. Sia V uno spazio vettoriale sul campo $\mathbb{C}$. Un prodotto hermitiano β si dice *prodotto scalare hermitiano* se β è definito positivo.

Esempio A.1. In $\mathbb{C}^n$ si consideri la funzione

$$\cdot\colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}, \quad (X, Y) \mapsto X \cdot Y \stackrel{\text{def}}{=} \sum_{i=1}^{n} \overline{x^i}\, y^i.$$

Questa è un prodotto scalare hermitiano su $\mathbb{C}^n$ (dimostrarlo).

Esercizio A.1. *Dire sotto quali condizioni su $A \in \mathbb{C}^{n,n}$ l'applicazione*

$$\cdot\colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}, X \cdot Y \stackrel{def}{=} X^* A Y$$

è un prodotto scalare hermitiano.

Definizione A.3. Si dice *spazio vettoriale hermitiano* uno spazio vettoriale complesso dotato di un prodotto scalare hermitiano.

Buona parte dei risultati sugli spazi vettoriali euclidei valgono senza modifiche sugli spazi vettoriali hermitiani. Lo studio della geometria di questi spazi permette di dare un'impostazione geometrica a molti problemi di analisi.

Sia (V, β) uno spazio vettoriale hermitiano. Se $\mathcal{B} = \{\vec{e}_1, \dots, \vec{e}_n\}$ è una base di V, posto

$$a_{ij} \stackrel{\text{def}}{=} \beta(\vec{e}_i, \vec{e}_j) \in \mathbb{C}, \qquad (i, j = 1, \dots, n),$$

si dice che la matrice $A = (a_{ij})$ è la matrice associata a β tramite $\mathcal{B}$. Si dimostra facilmente che A è hermitiana. Le fomule per il cambiamento di base sono simili a quelle per le forme bilineari reali.

Si definisce la *norma* di un vettore $\vec{v} \in V$ rispetto a β ponendo

$$\|\vec{v}\|_\beta \stackrel{\text{def}}{=} \sqrt{\beta(\vec{v}, \vec{v})} \geq 0$$

rispetto a β.

Si dice che due vettori $\vec{u}, \vec{v} \in V$ sono *ortogonali* se $\beta(\vec{u}, \vec{v}) = 0$, e si scrive $\vec{u} \perp_\beta \vec{v}$, o anche $\vec{u} \perp \vec{v}$ quando non esista possibilità di equivoco.

Valgono le disuguaglianze di Schwarz e di Minkowski:

$$|\beta(\vec{u}, \vec{v})| \leq \|\vec{u}\|_\beta \, \|\vec{v}\|_\beta,$$
$$\|\vec{u} + \vec{v}\|_\beta \leq \|\vec{u}\|_\beta + \|\vec{v}\|_\beta.$$

Il procedimento di ortonormalizzazione di Gram–Schmidt si può eseguire come nel caso di uno spazio vettoriale euclideo. Dunque, in ogni spazio vettoriale hermitiano esistono basi ortonormali.

Se U è un sottospazio di V, esiste ed è unico il complemento ortogonale $U^\perp$ di U, e vale $V = U \oplus U^\perp$. Dunque, dato $\vec{x} \in V$ esiste un'unica decomposizione $\vec{x} = \vec{x}_U + \vec{x}_{U^\perp}$. Se $\{\vec{e}_1, \dots, \vec{e}_n\}$ è una base ortonormale di V tale che $\{\vec{e}_1, \dots, \vec{e}_r\}$ è una base di U (una tale base esiste sempre, dimostrarlo!), allora $(\vec{x} - \vec{x}_U) \perp U$ implica

$$\beta((\vec{x} - \lambda_1\vec{e}_1 - \cdots - \lambda_r\vec{e}_r), \vec{e}_i) = 0 \quad \text{per} \quad i = 1, \dots, r$$

da cui risulta $\bar{\lambda_i} = \beta(\vec{x}, \vec{e}_i)$. I coefficienti λ_i prendono il nome di *coefficienti di Fourier*.

Proposizione A.1 (Disuguaglianza di Bessel). *Sia (V, β) uno spazio vettoriale hermitiano, ed $U \subset V$ un sottospazio. Se $\vec{x} \in V$, allora*

$$\|\vec{x}\|_\beta \geq \|\vec{x}_U\|_\beta.$$

Dimostrazione. Infatti, $\vec{x} = \vec{x}_U + \vec{x}_{U^\perp}$, dunque per il teorema di Pitagora $\|\vec{x}\|_\beta^2 = \|\vec{x}_U\|_\beta^2 + \|\vec{x}_{U^\perp}\|_\beta^2 \geq \|\vec{x}_U\|_\beta^2$. QED

In modo analogo si vede che, se $\vec{y} \in U$, allora $d(\vec{x}, \vec{y}) \geq d(\vec{x}, \vec{x}_U)$ dove $d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\|_\beta$. Dunque, il vettore $\vec{x}_U$ è il vettore di U avente minima distanza da x.

Esempio A.2. Si consideri lo spazio vettoriale $V = \mathcal{C}^0([a,b])$ costituito da tutte le funzioni continue $f\colon [a,b] \to \mathbb{C}$. L'applicazione

$$\beta\colon V \times V \to \mathbb{C},\ \beta(f,g) = \int_a^b \overline{f(x)} g(x)\, dx$$

è un prodotto scalare hermitiano. Infatti si dimostrano facilmente le proprietà della definizione A.1. Proviamo ora che β è definita positiva.

$$\beta(f,f) = \int_a^b |f(x)|^2\, dx \geq 0$$

poiché l'integrando è positivo. Rimane da provare che $\beta(f,f) = 0$ implica $f = 0$. Se f non è la funzione nulla esisterà qualche punto x_0 in cui $f(x_0) \neq 0$. Essendo f continua, ci sarà un intervallo $(x_0 - \epsilon, x_0 + \epsilon)$ in cui $f(x) \neq 0$, e quindi $\beta(f,f) \neq 0$ contro l'ipotesi.

La distanza indotta da questo prodotto scalare è

$$d(f,g) = \|f - g\| = \left(\int_a^b |f(x) - g(x)|^2\, dx \right)^{1/2},$$

detta *distanza* L^2, utilizzata nella teoria dei segnali [31].

Se $f\colon V \to W$ è un'applicazione lineare (rispetto a $\mathbb{C}$) tra i due spazi vettoriali hermitiani (V,β) e (W,γ), si dimostra che esiste una ed una sola applicazione lineare $f^*\colon W \to V$ detta *aggiunta* di f tale che

$$\gamma(f(\vec{x}), \vec{y}) = \beta(\vec{x}, f^*(\vec{y})) \qquad \forall \vec{x} \in V,\ \forall \vec{y} \in W.$$

Rispetto a basi ortonormali, la matrice dell'applicazione aggiunta è la matrice aggiunta dell'applicazione di partenza. Inoltre

$$V = \operatorname{Im} f^* \oplus \operatorname{Ker} f, \quad W = \operatorname{Im} f \oplus \operatorname{Ker} f^*,$$
$$\operatorname{Im} f^* = (\operatorname{Ker} f)^\perp, \quad \operatorname{Im} f = (\operatorname{Ker} f^*)^\perp$$

Si osservi che se A è la matrice associata ad f, allora $\operatorname{Im} f^*$ è lo spazio generato dalle righe di $\overline{A}$, matrice coniugata di A.

Un endomorfismo f di uno spazio vettoriale euclideo V con prodotto scalare g si dice *hermitiano* (o *autoaggiunto*) se $f = f^*$. In questo caso, la matrice rispetto a basi ortonormali risulta hermitiana.

Una *trasformazione unitaria* di V è un endomorfismo $f\colon V \to V$ tale che

$$\beta(f(\vec{x}), f(\vec{y})) = \beta(\vec{x}, \vec{y}).$$

La matrice di una trasformazione unitaria rispetto a basi ortonormali è una matrice unitaria.

Esercizio A.2. *Si dimostri che gli autovalori di una trasformazione autoaggiunta sono reali.*

Esercizio A.3. *Si dimostri che gli eventuali autovalori di una trasformazione unitaria sono numeri complessi di modulo* 1*.*

Esercizio A.4. *Sia data la matrice*

$$G = \begin{pmatrix} 4 & 1+3i \\ 1-3i & 3 \end{pmatrix}.$$

1. *Provare che G è associato ad un prodotto scalare hermitiano g.*
2. *Calcolare $g(2\vec{e}_1 + 3i\vec{e}_2, (2+i)\vec{e}_1 - \vec{e}_2)$.*

SOLUZIONE. 1. Si prove facilmente che $G^* = \overline{G}^T = G$. Inoltre i determinanti dei minori principali indicano che la G è definita positiva.

2. Tenendo conto delle proprietà di un prodotto scalare hermitiano si ha

$$g(2\vec{e}_1 + 3i\vec{e}_2, (2+i)\vec{e}_1 - \vec{e}_2) = 2(2+1)g(\vec{e}_1, \vec{e}_1) - 2g(\vec{e}_1, \vec{e}_2) - 3i(2+i)g(\vec{e}_2, \vec{e}_1) + 3ig(\vec{e}_2, \vec{e}_2) = -1 - 4i.$$

□

A.2 Elementi impropri e spazi proiettivi

La figura A.1 mostra una corrispondenza tra i punti della retta r e le rette del fascio passante per P_0. Tale corrispondenza è biunivoca, a meno di un'eccezione: alla retta r_0 per P_0 e parallela ad r non corrisponde alcun punto di r. Per rendere biunivoca la corrispondenza diciamo che alla retta r_0 corrisponde il punto all'infinito di r, detto anche *punto (o elemento) improprio* di r, ed indicato con P_∞. È chiaro che, se $r \parallel s$, allora s ed r avranno lo stesso punto improprio. Quindi punto improprio di una retta è sinonimo di direzione. La retta ampliata $\bar{r} \stackrel{\text{def}}{=} r \cup \{P_\infty\}$ ha un solo punto all'infinito, mentre in Analisi $\bar{\mathbb{R}}$ ha due punti all'infinito.

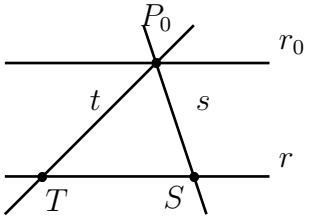


Figura A.1.

La retta affine ampliata è anche detta retta *proiettiva*: essa è in corrispondenza biunivoca con una circonferenza (figura A.2).

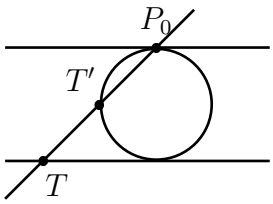


Figura A.2.

Da un punto di vista astratto, un fascio di rette è la stessa cosa di una retta proiettiva: l'unica 'differenza' è che una retta affine ampliata è generata da un punto, mentre un fascio di rette è generato da una retta. Se il fascio di rette è improprio, si può considerare come fascio di rette con centro in un punto improprio. In questo caso, la corrispondenza biunivoca con una retta è molto semplice: basta considerare le intersezioni di una retta non appartenente al fascio con le rette del fascio (figura A.3).

Dunque, se consideriamo due rette proiettive nel piano, esse si incontrano sempre; più precisamente,

$$\begin{aligned} \bar{r} \cap \bar{s} = P_\infty \quad &\Rightarrow \quad r \mathbin{/\!\!/} s, \\ \bar{r} \cap \bar{s} = P_0 \quad &\Rightarrow \quad r \text{ incidente } s. \end{aligned}$$

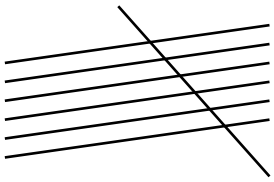


Figura A.3.

Le direzioni nel piano sono ∞^1; possiamo dire che esse costituiscono la *retta impropria* r_∞ (figura A.4). Posto $\bar{\pi} \stackrel{\text{def}}{=} \pi \cup r_\infty$ considerando un altro piano σ può verificarsi:

$$\begin{aligned} \bar{\pi} \cap \bar{\sigma} = r_\infty \quad &\Rightarrow \quad \pi \mathbin{/\!\!/} \sigma, \\ \bar{\pi} \cap \bar{\sigma} = r_0 \quad &\Rightarrow \quad \pi \text{ incidente } \sigma. \end{aligned}$$

Le usuali coordinate cartesiane (x, y) non sono adatte a rappresentare i punti impropri. A tale scopo si introducono le coordinate omogenee $(x_1, x_2, x_3) \in \mathbb{R}^3 \setminus \{(0,0,0)\}$. Se il punto P è proprio ed ha coordinate (x, y), ad esso si associa la terna (x_1, x_2, x_3) tale che $x_1/x_3 = x$, $x_2/x_3 = y$. Naturalmente $x_3 \neq 0$. È chiaro che si tratta di terne proporzionali. Più formalmente, si introduce in $\mathbb{R}^3 \setminus \{(0,0,0)\}$ la seguente relazione di equivalenza

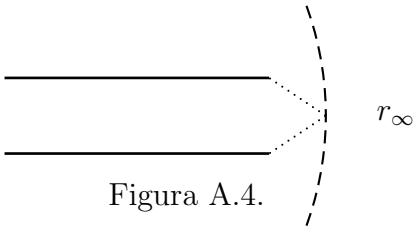


Figura A.4.

$$(x_1, x_2, x_3) \sim (x_1', x_2', x_3') \quad \Leftrightarrow \quad \text{esiste } k \in \mathbb{R} \setminus \{0\} \text{ tale che } (x_1, x_2, x_3) = k(x_1', x_2', x_3').$$

L'insieme quoziente $\mathbb{R}^3 \setminus \{(0,0,0)\}/\sim$ rispetto a tale relazione di equivalenza è detto *piano proiettivo* (reale). Ogni punto del piano proiettivo corrisponde ad una classe $P = [(x_1, x_2, x_3)]_\sim$ di vettori di coordinate proporzionali. Se $x_3 \neq 0$, allora P è proprio e le sue coordinate cartesiane (x, y) sono date dalle precedenti relazioni. Se $x_3 = 0$, allora P è improprio, ed è individuato da una direzione avente parametri direttori (l, m). Si ha

$$r \mathbin{/\!\!/} r' \quad \Leftrightarrow \quad (l, m) \sim (l', m') \quad \Rightarrow \quad (l, m, 0) \sim (l', m', 0) \quad \Rightarrow \quad P_\infty = P_\infty'.$$

Dunque, i punti impropri sono caratterizzati dall'avere $x_3 = 0$.

Le coniche ammettono un'interessante caratterizzazione usando la loro intersezione con la retta impropria. Una conica $\mathcal{C}$ irriducibile è rappresentabile come il luogo degli zeri di un polinomio di secondo grado irriducibile. Se r è una retta del piano (propria o impropria) risulta $r \cap \mathcal{C} = \{A, B\}$. Consideriamo ora la retta impropria: si ha

$$\begin{aligned} r_\infty \cap \mathcal{C} = 2 \text{ punti immaginari coniugati} \quad &\Rightarrow \quad \mathcal{C} \text{ ellisse} \\ r_\infty \cap \mathcal{C} = 2 \text{ punti reali distinti} \quad &\Rightarrow \quad \mathcal{C} \text{ iperbole} \\ r_\infty \cap \mathcal{C} = 2 \text{ punti reali coincidenti} \quad &\Rightarrow \quad \mathcal{C} \text{ parabola} \end{aligned}$$

In questa classificazione si prende come elemento di confronto la retta impropria e non una retta propria poiché in un cambiamento di riferimento *affine* i punti impropri rimangono

impropri (cioè hanno in entrambi i riferimenti la terza coordinata uguale a 0). Infatti, un cambiamento di riferimento affine (subordinato da uno proiettivo) è espresso da

$$\begin{cases} x'_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3, \\ x'_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3, \\ x'_3 = a_{33}x_3, \end{cases}$$

e ponendo come al solito $x' = x'_1/x'_3$, $y' = x'_2/x'_3$ si ha

$$\begin{cases} x' = ax + by + x_0, \\ y' = cx + dy + y_0. \end{cases}$$

A.3 Polinomio interpolatore

Sia $p(x) \in \mathbb{R}_n[x]$,

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0;$$

$n+1$ condizioni lineari indipendenti sui coefficienti individuano $p(x)$.

Teorema A.1. *Sia $\{\alpha_i\}_{0\le i\le n}$ una famiglia di numeri reali distinti e sia $\{\beta_i\}_{0\le i\le n}$ una famiglia di numeri reali. Allora, esiste un unico polinomio $P(x) \in \mathbb{R}_n[x]$ tale che*

$$P(\alpha_i) = \beta_i.$$

Dimostrazione. Le condizioni che $P(x)$ deve soddisfare sono

$$a_n\alpha_i^n + a_{n-1}\alpha_i^{n-1} + \cdots + a_1\alpha_i + a_0 = \beta_i,$$

per $0 \le i \le n$. Questo è un sistema lineare non omogeneo la cui matrice incompleta è

$$A = \begin{pmatrix} \alpha_0^n & \alpha_0^{n-1} & \cdots & \alpha_0 & 1 \\ \alpha_1^n & \alpha_1^{n-1} & \cdots & \alpha_1 & 1 \\ \alpha_n^n & \alpha_n^{n-1} & \cdots & \alpha_n & 1 \end{pmatrix}.$$

Si può dimostrare che

$$\det A = \prod_{0\le\lambda<\mu\le n} (\alpha_\mu - \alpha_\lambda);$$

questo è il *determinante di A. T. Vandermonde* (1735–1796). Per le ipotesi, $\det A \neq 0$, dunque la soluzione è unica. QED

L'interpretazione geometrica del precedente risultato è la seguente. Si dice *parabola di ordine n* la curva algebrica

$$y = a_0 + a_1 x + \cdots + a_n x^n.$$

Il teorema precedente, quindi, afferma che fissati nel piano $n+1$ punti (x_i, y_i), con $x_i \neq x_j$ se $i \neq j$, allora esiste un'unica parabola di ordine n passante per i suddetti punti.

Si osservi che una funzione $f(x)$ reale, derivabile n volte in un punto, ammette uno sviluppo di Taylor di ordine n, ed è quindi approssimabile in un intorno del punto da una parabola di ordine n.

Si consideri ora una funzione $f\colon U \subset \mathbb{R}^2 \to \mathbb{R}$. Essa, sotto condizioni di regolarità, ammette uno sviluppo di Taylor del tipo

$$f(x,y) = f(x_0,y_0) + (f_x^0(x-x_0) + f_y^0(y-y_0)) + \\ + \frac{1}{2}(f_{xx}^0(x-x_0)^2 + 2f_{xy}^0(x-x_0)(y-y_0) + f_{yy}^0(y-y_0)^2) + \cdots \quad \text{(A.3.1)}$$

La superficie Σ di equazione $z = f(x,y)$ avrà nel punto $P_0(x_0,y_0,z_0)$ il piano tangente

$$z - z_0 = f_x^0(x-x_0) + f_y^0(y-y_0).$$

Dunque, se nella (A.3.1) ci si ferma ai termini lineari, la superficie Σ è approssimata dal suo piano tangente, arrestandosi ai termini di secondo grado Σ è approssimata da un paraboloide.

Lo sviluppo (A.3.1) può essere traslato nell'origine per semplicità.

Se P_0 è un punto critico per f (quindi $f_x^0 = f_y^0 = 0$), allora il paraboloide approssimante è

$$\tilde{\Sigma}\colon z = \frac{1}{2}(f_{xx}^0 x^2 + 2f_{xy}^0 xy + f_{yy}^0 y^2). \quad \text{(A.3.2)}$$

La matrice associata alla precedente forma quadratica è la matrice hessiana

$$\mathrm{Hess}(P_0) = \begin{pmatrix} f_{xx}^0 & f_{xy}^0 \\ f_{yx}^0 & f_{yy}^0 \end{pmatrix},$$

che, come è noto, in questo caso è non degenere. Più in particolare si ha:

$$\det \mathrm{Hess}(P_0) > 0 \quad \Leftrightarrow \quad \text{Hess è definita} \quad \Leftrightarrow \quad P_0 \text{ ellittico}$$
$$\det \mathrm{Hess}(P_0) < 0 \quad \Leftrightarrow \quad \text{Hess è indefinita} \quad \Leftrightarrow \quad P_0 \text{ iperbolico}.$$

Se Hess è definita positiva, allora P_0 è un punto di minimo relativo; se Hess è definita negativa, allora P_0 è un punto di massimo relativo.

Riducendo a forma canonica la quadrica (A.3.2) si ha

$$\tilde{\Sigma}\colon \tilde{z} = \frac{1}{2}(\lambda_1 \tilde{x}^2 + \lambda_2 \tilde{y}^2).$$

Se P_0 è ellittico, le sezioni di $\tilde{\Sigma}$ con i piani $\tilde{z} = \mathit{cost}$ saranno ellissi; se P_0 è iperbolico, le sezioni di $\tilde{\Sigma}$ con i piani $\tilde{z} = \mathit{cost}$ saranno iperboli.

I numeri reali λ_1 e λ_2 sono detti *curvature principali* di $\tilde{\Sigma}$ in P_0, mentre $K(P_0) = \lambda_1\lambda_2$ è detta *curvatura gaussiana* di $\tilde{\Sigma}$ in P_0, quindi

$$\begin{aligned} P_0 \text{ ellittico} \quad &\Leftrightarrow \quad K(P_0) > 0, \\ P_0 \text{ iperbolico} \quad &\Leftrightarrow \quad K(P_0) < 0. \end{aligned}$$

Si dice *curvatura media* in P_0 il numero

$$H(P_0) = \frac{1}{2}(\lambda_1 + \lambda_2).$$

I concetti sopra esposti hanno senso non solo nei punti critici, ma anche in ogni punto della superficie, nell'ipotesi in cui esista il piano tangente ed il paraboloide approssimante. Basta considerare $\mathrm{Hess}(P_0)$ e trovare $\det(\mathrm{Hess}(P_0))$. Se $\det(\mathrm{Hess}(P_0)) = 0$ il punto si dice parabolico.

Ad esempio un ellissoide ha tutti i punti ellittici, un iperboloide iperbolico ha tutti i punti iperbolici, un cono ed un cilindro hanno tutti i punti parabolici, mentre un toro ha punti di tutti e tre i tipi.

Arrestando lo sviluppo di Taylor alla potenza n-esima si avrà un paraboloide approssimante di ordine n che è una superficie algebrica descritta da un polinomio di grado n, detto *polinomio interpolante*.

A.4 Polinomi di Bernstein

I polinomi di Bernstein sono utilizzati in Analisi Matematica, Analisi Numerica ed Informatica Grafica con scopi differenti. Tuttavia, le proprietà che li rendono così importanti in campi diversi hanno un notevole interesse matematico che prescinde dalle applicazioni.

Definizione A.4. Si dicono *polinomi di Bernstein* di grado n i polinomi

$$B_i^n(t) = \binom{n}{i} t^i (1-t)^{n-i}, \quad i = 0, \ldots, n,$$

ove, per $n = 0$, si pone $B_0^0(t) = 1$, e si pone $B_i^n(t) = 0$ per $i \notin \{0, \ldots, n\}$.

Lemma A.1. *I polinomi di Bernstein soddisfano le seguenti proprietà.*

1. $B_i^n(t) \geq 0$ *per* $t \in [0,1]$;
2. $B_i^n(t) = B_{n-i}^n(1-t)$;
3. $\sum_{i=0}^n B_i^n(t) = 1$;
4. $B_i^n(t) = (1-t)B_i^{n-1}(t) + tB_{i-1}^{n-1}(t)$;
5. $B_i^n(t)$ *ha un solo massimo (anche relativo) in* $[0,1]$, *per* $t = i/n$.

Dimostrazione.

1. Banale.
2. Immediata, tenendo conto dell'identità $\binom{n}{i} = \binom{n}{n-i}$.
3. Si usi l'identità $1 = (t + (1-t))^n$ e la formula del binomio di Newton.
4. Si usi l'identità
$$\binom{n}{i} = \binom{n-1}{i} + \binom{n-1}{i-1}.$$
5. Basta vedere che $d/dt\, B_i^n(t) = 0$ per $t = i/n$.

QED

Proposizione A.2. *I polinomi di Bernstein di grado n costituiscono una base di* $\mathbb{R}_n[t]$.

Dimostrazione. Infatti,

$$t^i = t^i((1-t)+t)^{n-i} = \sum_{j=i}^{n} \frac{\binom{n-i}{j-i}}{\binom{n}{j}} B_j^n(t) = \sum_{j=i}^{n} \frac{\binom{j}{i}}{\binom{n}{i}} B_j^n(t) \tag{A.4.3}$$

in tal caso la matrice di passaggio tra $\{1, t, \ldots, t^n\}$ e $\{B_j^n\}$ è la matrice $(\lambda_{ij}) \in \mathbb{R}^{n+1,n+1}$ tale che per i, $j = 1, \ldots, n+1$

$$\lambda_{ij} = \begin{cases} \frac{\binom{j-1}{i-1}}{\binom{n}{i-1}} & j \geq i \\ 0 & \text{altrimenti,} \end{cases}$$

che è chiaramente non singolare. QED

Esempio A.3. Si ha

$$B_0^1(t) = 1 - t, \qquad B_1^1(t) = t,$$
$$B_0^2(t) = (1-t)^2, \quad B_1^2(t) = 2t(1-t), \quad B_2^2(t) = t^2.$$

Si noti che i polinomi soddisfano la proprietà relativa al loro massimo descritta sopra. Inoltre, $B_i^n(t)$ per $i = 0, \ldots, n$ sono tutti di grado n; quindi un polinomio di grado minore od uguale ad n sarà combinazione lineare di polinomi di grado uguale ad n. Per esempio

$$3t + 5 = 5B_0^2(t) + \frac{13}{2} B_1^2(t) + 8B_2^2(t).$$

Esprimendo 1, t, t^2 tramite $B_j^2(t)$, $j = 1, 2, 3$, si ha

$$\begin{aligned} 1 &= B_0^2(t) + B_1^2(t) + B_2^2(t), \\ t &= \frac{1}{2}B_1^2(t) + B_2^2(t), \\ t^2 &= B_2^2(t). \end{aligned}$$

Esercizio A.5. *Con riferimento all'esempio precedente, si scrivano esplicitamente le matrici di passaggio dalla base monomiale a quella dei polinomi di Bernstein e viceversa.*

Nota A.1. I polinomi di Bernstein sonno stati introdotti dal matematico russo S. Bernstein per dimostrare il seguente teorema: *sia* $f\colon [0,1] \to \mathbb{R}$ *una funzione continua; allora esiste una successione* $\{s_n\}_{n\in\mathbb{N}}$ *di combinazioni lineari di polinomi di Bernstein che tende ad* f *uniformemente*. Il teorema si esprime in maniera più 'compatta' dicendo che *l'insieme dei polinomi è denso nell'insieme delle funzioni continue*, e mostra che ogni funzione continua può essere approssimata da un polinomio (si veda [30] per una dimostrazione), anche se si prova che tale approssimazione è poco efficiente dal punto di vista numerico [2].

A.5 Numeri modulari e crittografia

Il materiale qui presentato è tratto da [20].

A.5.a Congruenze

Definizione A.5. Sia $n \in \mathbb{Z} \setminus \{0\}$, e siano x, $y \in \mathbb{Z}$. Diciamo che x ed y sono congruenti modulo n ($x \equiv y \pmod{n}$) se $x - y$ è un multiplo di n (ossia $n \mid x - y$).

Osservazione A.1. La congruenza modulo n coincide con la congruenza modulo $-n$, quindi nel seguito considereremo solo congruenze con moduli positivi.

Se $n = 1$ allora tutti i numeri sono congruenti.

Esempio A.4.

$$\begin{aligned} 239 &\equiv 1 \pmod{7} \quad \text{essendo} \quad 239 - 1 = 7 \cdot 34; \\ -1845 &\equiv 3 \pmod{8} \quad \text{essendo} \quad -1845 - 3 = 8 \cdot (-231); \end{aligned}$$

La relazione di congruenza gode delle seguenti proprietà: per $n \in \mathbb{Z} \setminus \{0\}$

1. $x \equiv x \pmod{n}$ (proprietà riflessiva);
2. $x \equiv y \pmod{n} \quad \Rightarrow \quad y \equiv x \pmod{n}$ (proprietà simmetrica);
3. $x \equiv y \pmod{n}$ e $y \equiv z \pmod{n} \quad \Rightarrow \quad x \equiv z \pmod{n}$ (proprietà transitiva).

Dunque la relazione di congruenza modulo n è una relazione di *equivalenza* in $\mathbb{Z}$.

Proposizione A.3. *Se $n \in \mathbb{Z} \setminus \{0\}$ allora $x \equiv y \pmod{n}$ se e solo se la divisione di x ed y per n dà lo stesso resto.*

Dimostrazione. Siano

$$x = q_1 n + r_1, \qquad y = q_2 n + r_2.$$

Se $x \equiv y \pmod{n}$, allora $y - x = nk$, quindi $y = q_1 n + r_1 + nk = (q_1 + k)n + r_1$. Viceversa, se $r_1 = r_2$, risulta $y - x = (q_2 - q_1)n$. In entrambi i casi il risultato segue dall'unicità del quoziente e del resto nella divisione. QED

Definizione A.6. Sia $x \in \mathbb{Z}$. L'insieme degli $y \in \mathbb{Z}$ tali che $x \equiv y \pmod{n}$ si dice *classe di congruenza di x*, e si indica con $[x]_n$.

Ovviamente, $[x]_n = [r]_n$, dove r è il resto della divisione di x per n. Fissato n, ci sono n distinte classi di congruenza: $[0]_n$, $[1]_n$,... $[n-1]_n$. Esse si chiamano anche *classi dei resti*. Poniamo $\mathbb{Z}_n = \{[0]_n, [1]_n, \ldots, [n-1]_n\}$.

Esempio A.5. In $\mathbb{Z}_4 = \{[0]_4, [1]_4, [2]_4, [3]_4\}$ si ha

$$[0]_4 = \{x \in \mathbb{Z} \mid x : 4 \text{ ha resto } 0\} = \{0, 4, 8, \ldots\}$$
$$[1]_4 = \{x \in \mathbb{Z} \mid x : 4 \text{ ha resto } 1\} = \{1, 5, 9, \ldots\}$$

Dunque, $[m]_n \in \mathbb{Z}_n$ ha la forma $[m]_n = \{n\alpha + m \mid \alpha \in \mathbb{Z}\}$.

Se $[x]_n \in \mathbb{Z}_n$, ogni $y \in [x]_n$ si dice *rappresentante* di $[x]_n$. Ovviamente, x è un rappresentante di $[x]_n$, e $\{0, 1, \ldots, n-1\}$ sono rappresentanti per $\mathbb{Z}_n$.

Esempio A.6. $\{1024, 81, -18, -5\}$ è un insieme di rappresentanti di $\mathbb{Z}_4$ (verificarlo).

Valgono le seguenti proprietà:

1. $x \equiv y \pmod{n} \quad \Rightarrow \quad x + z \equiv y + z \pmod{n}$;
2. $x \equiv y \pmod{n} \quad \Rightarrow \quad xz \equiv yz \pmod{n}$.

Il viceversa della seconda proprietà non è valido: $14 \equiv 2 \pmod{12}$ ma $7 \not\equiv 1 \pmod{12}$.

Definizione A.7. Siano $[x]_n$, $[y]_n \in \mathbb{Z}_n$, ed $\alpha \in \mathbb{Z}$. Si definiscono le operazioni in $\mathbb{Z}_n$

1. $[x]_n + [y]_n = [x + y]_n$, somma,
2. $[x]_n \cdot [y]_n = [x \cdot y]_n$, prodotto,
3. $\alpha \cdot [x]_n = [\alpha \cdot x]_n$, prodotto per un intero.

Si può dimostrare facilmente che queste definizioni sono ben poste, cioè le definizioni non dipendono dai rappresentanti.

Notiamo che la somma ed il prodotto di classi rendono $(\mathbb{Z}_n, +, \cdot)$ un *anello commutativo con unità*, detto *anello delle classi di resti* o *anello dei numeri modulari*.

Esempio A.7. Le operazioni in $\mathbb{Z}_2$ e $\mathbb{Z}_3$.

$n = 2$:

$+$	$[0]_2$	$[1]_2$
$[0]_2$	$[0]_2$	$[1]_2$
$[1]_2$	$[1]_2$	$[0]_2$

$\cdot$	$[0]_2$	$[1]_2$
$[0]_2$	$[0]_2$	$[0]_2$
$[1]_2$	$[0]_2$	$[1]_2$

$n = 3$:

$+$	$[0]_3$	$[1]_3$	$[2]_3$
$[0]_3$	$[0]_3$	$[1]_3$	$[2]_3$
$[1]_3$	$[1]_3$	$[2]_3$	$[0]_3$
$[2]_3$	$[2]_3$	$[0]_3$	$[1]_3$

$\cdot$	$[0]_3$	$[1]_3$	$[2]_3$
$[0]_3$	$[0]_3$	$[0]_3$	$[0]_3$
$[1]_3$	$[0]_3$	$[1]_3$	$[2]_3$
$[2]_3$	$[0]_3$	$[2]_3$	$[1]_3$

Consideriamo ora equazioni con incognite in $\mathbb{Z}_n$. Dati $a, b \in \mathbb{Z}$ vogliamo trovare $x \in \mathbb{Z}$ tale che

1. $[a]_n + [x]_n = [b]_n$. In questo caso la soluzione è $[x]_n = [b]_n - [a]_n = [b - a]_n$;
2. $[a]_n \cdot [x]_n = [b]_n$. In questo caso la soluzione può non esistere.

Infatti, consideriamo in $\mathbb{Z}_4$ l'equazione $[2]_4 \cdot [x]_4 = [1]_4$. Si hanno i seguenti prodotti in $\mathbb{Z}_4$:

$$[2]_4[0]_4 = [0]_4, \quad [2]_4[1]_4 = [2]_4, \quad [2]_4[2]_4 = [0]_4, \quad [2]_4[3]_4 = [2]_4,$$

quindi non esiste $[x]_4 \in \mathbb{Z}_4$ tale che $[2]_4 \cdot [x]_4 = [1]_4$. Un esempio di equazione del tipo precedente che ammette soluzione in $\mathbb{Z}_4$ è $[3]_4[x]_4 = [1]_4$. In generale si ha il seguente risultato, che non sarà dimostrato.

Proposizione A.4. *L'equazione* $[a]_n[x]_n = [1]_n$ *in* $\mathbb{Z}_n$ *ha soluzione se e solo se* a *ed* n *sono* coprimi*, ovvero se e solo se* $\mathrm{MCD}(a, n) = 1$*; in tal caso, la soluzione è unica* (mod n).

Esempio A.8. L'equazione $[27]_{31} \cdot [x]_{31} = [1]_{31}$ ha soluzione in $\mathbb{Z}_{31}$ poiché 27 e 31 sono coprimi. Troviamo $x, y \in \mathbb{Z}$ tali che $27x + 31y = 1$ con l'algoritmo della divisione euclidea [8, 20]. Risulta $x = -8$ e $y = 7$, da cui $[27]_{31}^{-1} = [-8]_{31} = [23]_{31}$.

Dunque, gli elementi invertibili in $\mathbb{Z}_n$ sono tutte le classi $[a]_n$ tali che $\mathrm{MCD}(a, n) = 1$. Vale la seguente ovvia conseguenza.

Corollario A.1. *L'anello* $\mathbb{Z}_n$ *è un campo se e solo se* n *è un numero primo.*

A.5.b Applicazioni alla crittografia

$$\text{MITTENTE} \xrightarrow{\text{MESSAGGIO}} \text{RICEVENTE}$$

Il mittente vuole spedire un messaggio M al ricevente in condizioni di sicurezza. Per fare ciò, il mittente codifica il messaggio M secondo procedimenti che vedremo in seguito, trasformandolo in un messaggio N, e lo trasmette. Il ricevente lo decodifica ed ottiene il messaggio M.

L'obiettivo di questo paragrafo è applicare la teoria degli anelli $\mathbb{Z}_n$ al precedente problema.

Codice 1×1 (o codice di sostituzione monoalfabetico)

Consideriamo $\mathbb{Z}_{26}$, dove ogni classe di resto rappresenta una lettera in modo ovvio (ad esempio, $[1]_{26} = A$, $[2]_{26} = B$, ecc.). Vogliamo trasmettere il messaggio "TANTI AUGURI". Trasformiamo le lettere nei numeri corrispondenti

$$\begin{array}{cccccccccccc} T & A & N & T & I & X & A & U & G & U & R & I \\ 20 & 1 & 14 & 20 & 9 & 24 & 1 & 21 & 7 & 21 & 18 & 9 \\ -6 & 1 & -12 & -6 & 9 & -2 & 1 & -5 & 7 & -5 & -8 & 9 \end{array} \tag{A.5.4}$$

La terza stringa è stata ottenuta considerando i rappresentanti delle classi di resto aventi modulo minimo.

Scegliamo un numero invertibile in $\mathbb{Z}_{26}$, ad esempio 5 (infatti $[5]_{26}^{-1} = [-5]_{26}$ poiché $[5]_{26}[-5]_{26} = [1]_{26}$) Moltiplicando la terza stringa di (A.5.4) per 5 (modulo 26) e riconducendoci alle lettere si ottiene

$$\begin{array}{cccccccccccc} 22 & 5 & 18 & 22 & 19 & 16 & 5 & 1 & 9 & 1 & 12 & 19 \\ V & E & R & V & S & P & E & A & I & A & L & S \end{array} \tag{A.5.5}$$

Il ricevente trasformerà l'ultima stringa nella stringa numerica corrispondente, poi otterrà la stringa numerica corretta (cioè corrispondente al messaggio) moltiplicando tale stringa per -5 (modulo 26). Infatti $22 \cdot (-5) \equiv 20 \pmod{26}$, che corrisponde a T, e così via.

Ovviamente, questo codice è facilmente violabile perché ad ogni lettera del messaggio corrisponde la stessa lettera del messaggio codificato. Per ogni linguaggio sono facilmente reperibili le cosiddette *analisi di frequenza*, che dànno le percentuali di utilizzo di ogni lettera dell'alfabeto in quel linguaggio. Con questi dati è possibile capire a quale lettera corrisponda un dato simbolo nel messaggio codificato, conoscendo il linguaggio utilizzato. Inoltre le chiavi possibili sono solo 12 (perché?).

Codice matriciale 2×2 (o codice di sostituzione a blocchi)

Scegliamo una matrice $A \in \mathbb{Z}_{26}^{2,2}$ quadrata invertibile. Ad esempio

$$A = \begin{pmatrix} 8 & 13 \\ -5 & -8 \end{pmatrix}, \qquad \det A = 1 \quad \text{e} \quad A^{-1} = \begin{pmatrix} -8 & 13 \\ 5 & 8 \end{pmatrix}.$$

Accoppiamo da sinistra i numeri della stringa corrispondente al messaggio del precedente sottoparagrafo (A.5.4)

$$\begin{pmatrix} -6 \\ 1 \end{pmatrix}, \begin{pmatrix} -12 \\ -6 \end{pmatrix}, \begin{pmatrix} 9 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ -5 \end{pmatrix}, \begin{pmatrix} 7 \\ -5 \end{pmatrix}, \begin{pmatrix} -8 \\ 9 \end{pmatrix}.$$

Moltiplichiamo questi vettori per la matrice A, ottenendo

$$\begin{pmatrix} 8 & 13 \\ -5 & -8 \end{pmatrix}\begin{pmatrix} -6 \\ 1 \end{pmatrix} = \begin{pmatrix} 17 \\ 22 \end{pmatrix}, \quad \begin{pmatrix} 8 & 13 \\ -5 & -8 \end{pmatrix}\begin{pmatrix} 9 \\ -2 \end{pmatrix} = \begin{pmatrix} 20 \\ 23 \end{pmatrix}, \quad \begin{pmatrix} 8 & 13 \\ -5 & -8 \end{pmatrix}\begin{pmatrix} 7 \\ -5 \end{pmatrix} = \begin{pmatrix} 17 \\ 5 \end{pmatrix},$$
$$\begin{pmatrix} 8 & 13 \\ -5 & -8 \end{pmatrix}\begin{pmatrix} -12 \\ -6 \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \end{pmatrix}, \quad \begin{pmatrix} 8 & 13 \\ -5 & -8 \end{pmatrix}\begin{pmatrix} 1 \\ -5 \end{pmatrix} = \begin{pmatrix} 21 \\ 9 \end{pmatrix}, \quad \begin{pmatrix} 8 & 13 \\ -5 & -8 \end{pmatrix}\begin{pmatrix} -8 \\ 9 \end{pmatrix} = \begin{pmatrix} 1 \\ 20 \end{pmatrix}.$$

Otteniamo il messaggio codificato

$$\begin{array}{cccccccccccc} 17 & 22 & 8 & 4 & 20 & 23 & 21 & 9 & 17 & 5 & 1 & 20 \\ Q & V & H & D & T & W & U & I & Q & E & A & T \end{array} \tag{A.5.6}$$

Il ricevente, ottenuta la stringa alfabetica precedente, la trasforma nella corrispondente stringa numerica, poi accoppia i numeri e moltiplica i vettori così ottenuti per A^{-1}:

$$\begin{pmatrix} -8 & 13 \\ 5 & 8 \end{pmatrix}\begin{pmatrix} 17 \\ 22 \end{pmatrix} = \begin{pmatrix} 20 \\ 1 \end{pmatrix}, \dots$$

ottenendo la stringa numerica corrispondente al messaggio.

Anche questo codice non è molto sicuro. Per aumentarne la sicurezza si può passare ad un codice 3×3. Oppure, si può aumentare la lunghezza dell'alfabeto includendo, ad esempio, i simboli !, ?, ecc. passando ad un alfabeto a 31 simboli. Essendo 31 primo, risulta $\varphi(31) = 30$[1], quindi aumentano le possibili chiavi. Inoltre, è facile rendersi conto del fatto che questo codice non è sensibile alle analisi di frequenza, poiché la codifica di ogni lettera dipende da altri simboli (un altro simbolo nell'esempio considerato, 2 simboli nel caso di un codice 3×3, ecc.).

[1]Qui φ rappresenta l'*indicatore di Eulero*, ovvero la funzione che assegna ad ogni numero naturale $n \in \mathbb{N}$ il numero dei numeri naturali $1 \leq k < n$ che sono primi con n, ovvero tali che k ed n non hanno divisori in comune. Ovviamente $\varphi(p) = p - 1$ se p è primo.

BIBLIOGRAFIA

[1] S. Abeasis: Algebra lineare e geometria, Zanichelli 1990.

[2] K. E. Atkinson: An introduction to numerical analysis, 2nd ed., J.Wiley & Sons, 1989.

[3] Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, D. Sorensen: LAPACK user's guide, Society for Industrial and Applied Mathematics (1999), `http://www.netlib.org/lapack/lug/index.html`

[4] M. W. Baldoni, C. Ciliberto, G. M. Piacentini Cattaneo: Aritmetica, crittografia e codici, Springer 2006.

[5] D. Bini, M. Capovani, O. Menchi: Metodi numerici per l'algebra lineare, Zanichelli, 1989.

[6] Å. Björck: *Solving linear least square problems by Gram–Schmidt orthogonalization*, BIT **7** (1967), 1–21.

[7] G. Calvaruso, R. Vitolo: Esercizi di Geometria ed Algebra, Università di Lecce, 2001 `http://poincare.unile.it/vitolo`.

[8] A. Caranti, S. Mattarei: Teoria dei numeri e crittografia, note del corso di A. Caranti, Univ. di Trento, 2002 (reperibile in Internet).

[9] B. Casselman: Mathematical Illustrations – a manual of geometry and PostScript, Cambridge Univ. Press 1005, disponibile in Internet `http://www.math.ubc.ca/~cass/graphics/manual/`

[10] I. Cattaneo Gasparini: Strutture algebriche, operatori lineari, Roma, 1989.

[11] I. Cattaneo Gasparini, G. De Cecco: Introduzione ai metodi della geometria differenziale (per ingegneri e fisici), ed. Veschi, Roma, 1979.

[12] V. Checcucci, A. Tognoli, E. Vesentini: Lezioni di topologia generale, Feltrinelli 1972.

[13] V. Comincioli: Analisi numerica: complementi e problemi, McGraw–Hill, 1991

[14] P. de Bartolomeis: Algebra Lineare, La Nuova Italia, Firenze 1993.

[15] G. De Cecco: Le origini storiche della geometria proiettiva (appunti redatti da S. P. Ruggeri), Dipart. di Matematica "E. De Giorgi", Lecce, 1999.

[16] G. De Cecco, R. Vitolo: Note di geometria ed algebra. Dispense del corso di Geometria ed Algebra per la laurea di I livello della Facoltà di Ingegneria dell'Università di Lecce, URL: `http://poincare.unile.it/vitolo`.

[17] G. Farin: Curves and Surfaces for Computer Aided Geometric Design, A Practical Guide, Academic Press Inc.

[18] J. D. Foley, A. van Dam, S. K. Feiner, J. F. Hughes: Computer Graphics, principles and practice, Addison-Wesley, 1997.

[19] E. Fornasini, M. E. Valcher: Metodi e algoritmi per lo studio dei sistemi multivariabili (Parte I), a.a. 1997/1998, Università di Padova.

[20] E. Francot, R. Vitolo: Note di teoria dei numeri, dispense per i corsi di laurea teledidattici Net.T.Un.O., 2002. URL: `http://www.nettuno.unile.it/`

[21] H. Goldstein: Classical Mechanics, 2nd ed., Addison-Wesley 1980.

[22] G. H. Golub, C. F. Van Loan: Matrix Computations, Johs Hopkins Univ. Press, 3rd edition.

[23] W. Greub: Linear Algebra, IV ed., Springer, 1975.

[24] W. Keith Nicholson: Algebra lineare (dalle applicazioni alla teoria), McGraw–Hill, 2002.

[25] V. V. Klyuev, H. I. Kokovkin-Shcherbak: *Minimization of the number of arithmetic operations in the solution of linear algebraic systems of equations*, USSR Computational Mathematics and Mathematical Physics **5** (1965), 25–43.

[26] N. Koblitz: A course in number theory and cryptography, 2nd ed., Springer 1994.

[27] A. K. Jain: Fundamentals of digital image processing, Englewood Cliffs, Prentice Hall, 1989.

[28] A. Quarteroni, F. Saleri: Introduzione al calcolo scientifico: esercizi e problemi risolti con MATLAB, Springer, 2002.

[29] A. Quarteroni, R. Sacco, F. Saleri: Matematica numerica, Springer, 1998.

[30] C. Ravaglia: Analisi Matematica I, CUSL Bologna, 1989.

[31] G. Ricci, M. E. Valcher: Segnali e sistemi, Ed. Libreria Progetto, Padova 2001.

[32] L. Salce: Lezioni sulle matrici, ed. Decibel–Zanichelli, 1993.

[33] A. Sanini: Lezioni di Geometria, ed. Levrotto-Bella, Torino 1993; Esercizi di Geometria, ed. Levrotto-Bella, Torino 1993.

[34] L.L. Scharf: Statistical signal processing, Addison-Wesley 1990.

[35] L. Sciavicco, B. Siciliano: Robotica industriale: modellistica e controllo di manipolatori, Ed. McGraw-Hill, Milano, 2000.

[36] K. Sigmon: Matlab primer, 3rd ed., Dept. of Mathematics, Univ. of Florida, 1993. http://poincare.unile.it/vitolo/vitolo_files/didattica/calcmat/matlab.ps.gz

[37] S. Savchenko: 3Dgpl's – 3D Graphics Reference. 1995.

INDICE ANALITICO