Appendice A

Capitolo 2

A.1 Alcune questioni di base sulle matrici

Si assume che il lettore sia familiare con i concetti di base della teoria delle matrici. Qui si premettono solo alcune notazioni. Sia $\mathbb{R}$ l'insieme dei numeri reali. Si denota con $\mathbb{R}^{m\times n}$ lo spazio vettoriale di tutte le matrici reali di dimensioni $m \times n$:

$$A \in \mathbb{R}^{m\times n} \Longleftrightarrow A \;=\; [a_{ij}] \;=\; \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \quad a_{ij} \in \mathbb{R}.$$

Se $m = n$, la matrice è detta **quadrata di ordine** n.

Gli **elementi diagonali** di una matrice A quadrata di ordine n sono gli elementi $a_{11}, a_{22}, ..., a_{nn}$. Una **matrice diagonale** di ordine n è una matrice quadrata di ordine n i cui elementi non diagonali sono uguali a zero e, se con $d_1, d_2, ..., d_n$ si indicano gli elementi diagonali, essa è denotata con $\texttt{diag}[d_1, d_2, ..., d_n]$.

Si denota con I la **matrice identità** (o **matrice identica**) di ordine n, cioè la matrice diagonale di ordine n i cui elementi diagonali sono tutti uguali a 1.

Data $A \in \mathbb{R}^{m\times n}$, si denota con A^T la **trasposta** di A, cioè la matrice $C = A^T \in \mathbb{R}^{n\times m}$ tale che $c_{ij} = a_{ji}$.

Data $A \in \mathbb{R}^{n\times n}$, si denota con $\texttt{det}(A)$ il suo **determinante**.

Si denota con $\mathbb{R}^n$ lo spazio vettoriale dei vettori reali di dimensione n:

$$x \in \mathbb{R}^n \Longleftrightarrow x \;=\; [x_i] \;=\; \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad x_i \in \mathbb{R}.$$

Si osservi che si è identificato $\mathbb{R}^n$ con $\mathbb{R}^{n\times 1}$ e quindi gli elementi di $\mathbb{R}^n$ sono detti vettori **colonna**. D'altra parte, gli elementi di $\mathbb{R}^{1\times n}$ sono vettori **riga**:

$$x \in \mathbb{R}^{1\times n} \Longleftrightarrow x \;=\; (x_1, \cdots, x_n).$$

Se $x \in \mathbb{R}^n$, allora $y = x^T$ è un vettore riga.
Dati x, $y \in \mathbb{R}^n$, si denota con $\langle x, y \rangle$ il **prodotto interno (prodotto scalare)** di x e y, cioè:

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i.$$

Per indicare il prodotto interno di due vettori, x e y, si utilizzerà spesso anche la seguente notazione:

$$x^T \cdot y = \sum_{i=1}^{n} x_i y_i.$$

A.1.1 Matrice di permutazione

Una matrice di **permutazione** P di ordine n è una matrice che moltiplicata a sinistra per una qualsiasi matrice A avente n righe provoca lo scambio di due o piú righe di A. Una matrice di permutazione P di ordine n moltiplicata a destra per una qualsiasi matrice A avente n colonne provoca lo scambio di due o piú colonne di A. La struttura di una matrice di permutazione ha:

- elementi uguali a 0 o 1;
- solamente un elemento uguale a 1 in ogni riga ed in ogni colonna.

In pratica una matrice di permutazione P si ottiene dalla matrice identica (dello stesso ordine) scambiandone le righe o le colonne.

Esempi di matrici di permutazione sono i seguenti:

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

Si osservi inoltre che se P è una matrice di permutazione si ha $PP^T = I$.

A.2 Inversa di una matrice

Nel tentativo di estendere alle matrici (ed ai vettori) concetti associati all'aritmetica dei numeri, si introduce la nozione di *inversa* (o *reciproca*) di una matrice in virtù dell'esistenza dell'elemento neutro rispetto al prodotto matriciale, ovvero all'esistenza della matrice identica, che rappresenta l'analogo del numero 1, nella moltiplicazione tra numeri. Poiché, assegnato un numero $a \in \Re$, $a \neq 0$, si definisce *inverso* di a il numero

$b \in \Re$ tale che $ab = ba = 1$, dato da $b = 1/a$, appare naturale chiedersi se assegnata una matrice $A \in \Re^{m\times m}$ esiste una matrice $B \in \Re^{m\times m}$ tale che

$$AB = BA = I_m.$$

Se una tale matrice esiste essa viene detta *inversa di* A e denotata con A^{-1} (analogo di $1/a$).[1]

♣ **Esempio A.1.** Assegnata:

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

si vuole determinare se esiste, ed in tal caso, quale è l'espressione dell'inversa di A.
Applicando la definizione si ha che per essere $B \in \Re^{2\times 2}$ inversa di A deve verificarsi che

$$AB = BA = I_2$$

Pertanto, si puó valutare dapprima se esiste una matrice

$$C = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

con a, b, c, d da determinare in maniera tale che

$$AC = \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Ció è verificato se

$$a - c = 1;\; b - d = 0;\; a + 2c = 0;\; b + 2d = 1.$$

Per determinare i valori degli elementi di C che soddisfano le precedenti 4 equazioni bisogna risolvere i due seguenti sistemi di equazioni, rispettivamente in a e c, e b e d:

$$\left\{\begin{array}{ccccc} a & - & c & = & 1 \\ a & + & 2c & = & 0 \end{array}\right. ; \quad \left\{\begin{array}{ccccc} b & - & d & = & 0 \\ b & + & 2d & = & 1 \end{array}\right.$$

I due sistemi possono essere risolti applicando il metodo di addizione e sottrazione.
Considerato il primo sistema,

- si sottrae la II equazione dalla I e si ottiene

$$-3c = 1 \quad \text{da cui} \quad c = -1/3;$$

- si sostituisce il valore ottenuto per c nella I equazione e si ha

$$a + 1/3 = 1 \quad \text{da cui} \quad a = 1 - 1/3 = 2/3.$$

Analogamente, considerato il secondo sistema,

[1] Se esiste l'inversa di A essa è unica, infatti supponendo che esistano due matrici B e C tali che $BA = AB = I_m$ e $CA = AC = I_m$, moltiplicando ambo i membri dell'uguaglianza $BA = I_m$ per C si ha $(BA)C = C$ ma $(BA)C = B(AC) = B$ (proprietà associativa) da cui necessariamente $B = C$.

- si sottrae la II equazione dalla I e si ottiene

$$-3d = -1 \quad \text{da cui} \quad d = 1/3;$$

- si sostituisce il valore ottenuto per d nella I equazione e si ha

$$b - 1/3 = 0 \quad \text{da cui} \quad b = 1/3.$$

In conclusione C esiste ed è data da

$$C = \begin{pmatrix} 2/3 & 1/3 \\ -1/3 & 1/3 \end{pmatrix}$$

(C è detta *inversa destra*).
Analogamente si verifica se esiste

$$E = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$$

con a', b', c', d' da determinare in maniera tale che:

$$EA = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Ció è verificato se

$$a' + b' = 1; \; -a' + 2b' = 0; \; c' + d' = 0; \; -c' + 2d' = 1.$$

che conduce alla risoluzione dei due seguenti sistemi:

$$\left\{ \begin{array}{rcrcl} a' & + & b' & = & 1 \\ -a' & + & 2b' & = & 0 \end{array} \right. \quad ; \quad \left\{ \begin{array}{rcrcl} c' & + & d' & = & 0 \\ -c' & + & 2d' & = & 1 \end{array} \right.$$

Applicando il metodo di addizione e sottrazione si ottiene che E esiste ed è data da:

$$E = \begin{pmatrix} 2/3 & 1/3 \\ -1/3 & 1/3 \end{pmatrix}$$

(E è detta *inversa sinistra*).
Pertanto,

$$C = E = A^{-1}$$

inversa di A. ♣

A.3 Norme Vettoriali e Matriciali

In molte applicazioni, tra le quali la risoluzione di sistemi di equazioni lineari, i dati che entrano in gioco sono vettori e matrici, pertanto risulta necessario quantificare la distanza tra vettori e tra matrici e, quindi, misurare l'ordine di grandezza di tali quantità. Nel caso dei numeri reali, la quantità che misura la grandezza di un numero è il valore assoluto; ad esempio $|7| > |5|, |-3| > |-1.5|$. Quello che si farà nell'immediato seguito è introdurre i concetti di **norme** di vettori e matrici, le quali costituiscono uno strumento per ottenere le suddette misure; si definirà, dunque, qualcosa di analogo al valore assoluto, che, però, consenta di misurare, con un singolo numero reale, la "grandezza" di un vettore e di una matrice.

♣ **Esempio A.2.** Si consideri il vettore del piano $\underline{y} = (3, 4)$.

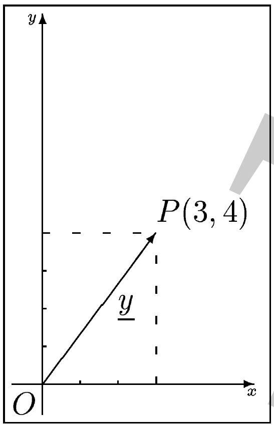

Figura A.1: Rappresentazione geometrica della distanza di un punto dall'origine

Introdotto nel piano un sistema di assi cartesiani ortogonali, la lunghezza del vettore $\underline{y}$ è data da:

$$l = \sqrt{3^2 + 4^2} = \sqrt{25} = 5$$

che, come illustrato in fig. A.1, rappresenta la distanza del punto P di coordinate $(3, 4)$ dall'origine degli assi.

♣

In generale, la lunghezza di un vettore del piano $\underline{x} = (x_1, x_2)^T$ è data da:

$$l = \sqrt{x_1^2 + x_2^2} \geq 0,$$

che indica la distanza del punto P di coordinate (x_1, x_2) dall'origine del riferimento. Tale numero è detto *norma euclidea* (o *norma due*) del vettore $\underline{x}$ e si indica con

$$l = \sqrt{x_1^2 + x_2^2} = \|\underline{x}\|_2.$$

Tale definizione si può estendere ad un generico vettore di $\Re^n$:

$$\underline{x} = (x_1\ x_2\ \ldots\ x_n)^T,$$

definendo lunghezza di $\underline{x}$ la quantità:

$$l = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} = \|\underline{x}\|_2,$$

che fornisce un numero reale non negativo che misura la "grandezza" del vettore $\underline{x}$ indicando la distanza del punto P di coordinate $(x_1, x_2, \ldots, x_n)$ dall'origine di un sistema

di riferimento di assi cartesiani introdotto nello spazio $\Re^n$.
Con la norma euclidea è possibile, quindi, avere indicazioni sulla grandezza di un vettore. È possibile tuttavia, stimare la grandezza di un vettore definendo altre norme vettoriali. Un'altra norma, molto utilizzata, è la cosiddetta *norma del massimo* (o *norma infinito*), denotata con $\|\underline{x}\|_\infty$, che identifica la grandezza del vettore $\underline{x}$ con il massimo tra i valori assoluti delle sue componenti, ovvero:

$$\|\underline{x}\|_\infty = \max_{1 \le i \le n} |x_i|.$$

♣ **Esempio A.3.**

$$\underline{x} = (-3.\ 1.5\ \ -2.\ 4.\ 0.\ \ -7.8)^T,$$

$$\|\underline{x}\|_\infty = \max_{1 \le i \le 6} |x_i| = 7.8\ \ .$$

♣

Infine, un altro modo per stimare la grandezza di un vettore è calcolare la somma dei valori assoluti di tutte le sue componenti. Tale quantità è denominata *norma uno*, denotata con $\|\underline{x}\|_1$, ed è data da:

$$\|\underline{x}\|_1 = \sum_{i=1}^{n} |x_i|.$$

♣ **Esempio A.4.** Per il vettore $\underline{x}$ dell'esempio precedente si ha:

$$\|\underline{x}\|_1 = \sum_{i=1}^{n} |x_i| = 18.3\ \ .$$

♣

In conclusione, finora sono state introdotte tre differenti stime della grandezza di un vettore, cioè tre norme vettoriali. Tali norme, ed affianco ad esse altre, posseggono proprietà analoghe a quelle del valore assoluto di un numero. Vale, dunque, la seguente definizione:

Definizione A.1. (Norma vettoriale)
Una **norma vettoriale***, che si denoterà con il simbolo* $\|\cdot\|_\alpha$*, è una funzione* $\|\cdot\|_\alpha : \mathbb{R}^n \to \mathbb{R}$ *che gode delle seguenti proprietà:*

1. $\|x\|_\alpha \ge 0,\ \forall x \in \mathbb{R}^n$; $\qquad \|x\|_\alpha = 0 \iff x = 0$;

2. $\|c\,x\|_\alpha = |c|\;\|x\|_\alpha,\ \forall x \in \mathbb{R}^n,\ \forall c \in \mathbb{R}$;

3. $\|x+y\|_\alpha \leq \|x\|_\alpha + \|y\|_\alpha,\ \forall x, y \in \mathbb{R}^n$.

La prima proprietà garantisce, in particolare, che le tre norme introdotte restituiscano un numero positivo o nullo o, meglio, che tutti i vettori, tranne quello nullo, hanno lunghezza positiva; questa proprietà vale, invero, per una qualsiasi norma vettoriale.
La seconda proprietà afferma che un vettore avente le componenti che sono multiple, secondo uno stesso fattore α, delle componenti di un altro vettore, ha una lunghezza pari ad α volte la lunghezza di tale vettore, e stabilisce anche che il vettore $\underline{x}$ ed il vettore $-\underline{x}$ hanno la stessa lunghezza.
Infine, la terza proprietà, chiamata *diseguaglianza triangolare*, in maniera analoga a quanto accade per il valore assoluto, stabilisce che la norma della somma di due vettori è minore od uguale alla somma delle singole norme.
Le norme descritte negli esempi sono alcune tra le più note ed utilizzate norme vettoriali e sono, ricapitolando, le seguenti:

$\|x\|_1 = \sum_{i=1}^n \lvert x_i\rvert$	**norma 1 o norma del valore assoluto**
$\|x\|_2 = \sqrt{\sum_{i=1}^n \lvert x_i\rvert^2}$	**norma 2 o norma Euclidea**
$\|x\|_\infty = \max_{1\leq i\leq n} \lvert x_i\rvert$	**norma ∞ o norma del massimo**

Il problema che ci si pone ora, in maniera analoga a quanto fatto per i vettori, è di valutare, con un numero, la grandezza di una matrice, ossia definire una *norma matriciale*. L'idea è di introdurla a partire da una norma vettoriale.

♣ **Esempio A.5.** Per il vettore $\underline{x} = (3\ 4)^T$ si ha che $\|\underline{x}\|_2 = \sqrt{3^2+4^2} = 5$.
Assegnata la matrice:

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

si ha che

$$A\underline{x} = \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 10 \\ 13 \end{pmatrix} = \underline{y},$$

e

$$\|\underline{y}\|_2 = \sqrt{10^2+13^2} = \sqrt{269} = 16.4012\ldots\ .$$

Considerata ora la seguente matrice:

$$B = \begin{pmatrix} 20 & 10 \\ 30 & 10 \end{pmatrix},$$

si ha che

$$B\underline{x} = \begin{pmatrix} 20 & 10 \\ 30 & 10 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 100 \\ 130 \end{pmatrix} = \underline{z},$$

e

$$\|\underline{z}\|_2 = \sqrt{100^2 + 130^2} = \sqrt{26900} = 164.012\ldots \; .$$

La quantità

$$\frac{\|A\underline{x}\|_2}{\|\underline{x}\|_2} = \frac{16.4012}{25} = 0.656049$$

misura quanto la matrice A è "grande" rispetto ad $\underline{x}$; analogamente la quantità

$$\frac{\|B\underline{x}\|_2}{\|\underline{x}\|_2} = \frac{164.012}{25} = 6.56049$$

misura quanto la matrice B è "grande" rispetto ad $\underline{x}$. Confrontando le due quantità si deduce che la matrice B è piú "grande" di A rispetto ad $\underline{x}$. ♣

In generale, fissato un vettore $\underline{x}$, la "grandezza" di una matrice A viene a dipendere dalla "grandezza" del vettore $A\underline{x}$; quindi, per introdurre una quantità che fornisca una misura della grandezza di A è necessario far variare il vettore $\underline{x}$. Una stima della grandezza di A si ottiene pertanto, considerando il

$$\max_{\underline{x} \neq 0} \frac{\|A\underline{x}\|}{\|\underline{x}\|},$$

dove $\|\cdot\|$ è una norma vettoriale. Tale quantità è un numero non negativo, denominato *norma di* A (o, più precisamente, *norma matriciale indotta da una norma vettoriale*) ed è, generalmente, indicata con $\|A\|$, ovvero

$$\|A\| = \max_{\underline{x} \neq 0} \frac{\|A\underline{x}\|}{\|\underline{x}\|}.$$

Vale, in realtà, la seguente definizione

Definizione A.2. (Norma matriciale indotta)
Sia $\|\cdot\|_\alpha$ una norma vettoriale. La norma matriciale:

$$\|A\|_\beta = \sup_{x \neq 0} \frac{\|Ax\|_\alpha}{\|x\|_\alpha}, \quad \forall A \in \mathbb{R}^{n \times n} \tag{A.1}$$

è detta **indotta** *dalla norma vettoriale $\|\cdot\|_\alpha$.*

Da questa definizione si deduce che la $\|A\|$ stima il *potere amplificante* della matrice.

Più in generale, le proprietà 1-3 della norma vettoriale sussistono anche per la norma matriciale; in particolare consentono di fornire la seguente:

Definizione A.3. (Norma matriciale)
Una **norma matriciale***, che si denoterà con il simbolo* $\|\cdot\|_\beta$*, è una funzione* $\|\cdot\|_\beta : \mathbb{R}^{n\times n} \to \mathbb{R}$ *che gode delle seguenti proprietà:*

1. $\|A\|_\beta \geq 0$, $\forall A \in \mathbb{R}^{n\times n}$; $\qquad \|A\|_\beta = 0 \iff A = 0$;
2. $\|c\,A\|_\beta = |c|\,\|A\|_\beta$, $\forall A \in \mathbb{R}^{n\times n}$, $\forall c \in \mathbb{R}$;
3. $\|A+B\|_\beta \leq \|A\|_\beta + \|B\|_\beta$, $\forall A, B \in R^{n\times n}$.

Una norma matriciale è inoltre caratterizzata da proprietà, alcune delle quali sono espresse dalle seguenti definizioni:

Definizione A.4. (Norma matriciale moltiplicativa)
Una norma matriciale $\|\cdot\|_\beta$ *è detta* **moltiplicativa** *se*

$$\|AB\|_\beta \leq \|A\|_\beta\,\|B\|_\beta, \;\; \forall A, B \in R^{n\times n}$$

In generale si utilizzano solo norme moltiplicative.

Definizione A.5. (Norma matriciale compatibile)
Sia $\|\cdot\|_\alpha$ *una norma vettoriale. Una norma matriciale* $\|\cdot\|_\beta$ *si dice* **compatibile** *con* $\|\cdot\|_\alpha$ *se, per ogni* $A \in \mathbb{R}^{n\times n}$ *e per ogni* $x \in \mathbb{R}^n$*, si ha:*

$$\|Ax\|_\alpha \leq \|A\|_\beta \|x\|_\alpha$$

Norme matriciali moltiplicative e compatibili possono essere ottenute a partire dalle norme vettoriali e si definiscono *indotte* da esse.

È chiaro che in corrispondenza della norma vettoriale scelta si ottengono diverse espressioni per la relativa norma matriciale.
Ad esempio, scelta la norma vettoriale del massimo, si ha la seguente espressione per la corrispondente norma matriciale:

$$\|A\|_\infty = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{i,j}|, \quad A = (a_{i,j}) \in \Re^{n\times n}$$

Infatti,

$$\|A\underline{x}\|_\infty = \max_{1\leq i\leq n} \left| \sum_{j=1}^{n} a_{i,j} x_j \right| \leq \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{i,j}||x_j| \leq \max_{1\leq i\leq n} \big((\max_{1\leq j\leq n} |x_j|) \sum_{j=1}^{n} |a_{i,j}| \big) =$$

$$= \max_{1\le j\le n} |x_j| \max_{1\le i\le n} \sum_{j=1}^{n} |a_{i,j}| = \|\underline{x}\|_\infty \max_{1\le i\le n} \sum_{j=1}^{n} |a_{i,j}|$$

da cui

$$\frac{\|A\underline{x}\|_\infty}{\|\underline{x}\|_\infty} \le \max_{1\le i\le n} \sum_{j=1}^{n} |a_{i,j}|, \quad \text{per ogni } \underline{x} \in \Re^n,$$

e quindi

$$\max_{1\le i\le n} \sum_{j=1}^{n} |a_{i,j}| = \max_{\underline{x}\neq 0} \frac{\|A\underline{x}\|_\infty}{\|\underline{x}\|_\infty} = \|A\|_\infty.$$

Analogamente si può dimostrare che, scelta la norma vettoriale uno, la corrispondente norma matriciale è data da

$$\|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^{n} |a_{i,j}|, \quad A = (a_{i,j}) \in \Re^{n\times n}.$$

Alcune tra le più note ed utilizzate norme matriciali sono le seguenti, le quali sono indotte dalle norme vettoriali $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$:

$\|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^n \|a_{ij}\|$	**norma 1**
$\|A\|_2 = \sqrt{\rho(A^T A)}$	**norma 2**
$\|A\|_\infty = \max_{1\le i\le n} \sum_{j=1}^n \|a_{ij}\|$	**norma ∞**

Infine, data una norma matriciale $\|\cdot\|_\beta$ ed una matrice non singolare $B \in \mathbb{R}^{n\times n}$, si definisce la norma matriciale $\|\cdot\|_{\beta,B}$ nel modo seguente:

$$\|A\|_{\beta,B} = \|B^{-1}AB\|_\beta\ , \quad \forall A \in \mathbb{R}^{n\times n}.$$

Si dimostra ora che tutte le norme vettoriali sono equivalenti.

Teorema A.1. *Siano $\|\cdot\|_\alpha$ e $\|\cdot\|_{\alpha'}$ due norme vettoriali qualsiasi. Esistono due costanti $c_2 \ge c_1 > 0$ tali che:*

$$c_1\|x\|_\alpha \le \|x\|_{\alpha'} \le c_2\|x\|_\alpha, \ \ \forall x \in \mathbb{R}^n. \tag{A.2}$$

Dimostrazione Si osservi innanzitutto che è sufficiente dimostrare che la (A.2) è vera se una delle due norme è la norma Euclidea. Infatti, se valgono le due relazioni:

$$d_1\|x\|_\alpha \leq \|x\|_2 \leq d_2\|x\|_\alpha, \quad {d_1}'\|x\|_{\alpha'} \leq \|x\|_2 \leq {d_2}'\|x\|_{\alpha'}$$

con $d_2 \geq d_1 > 0$ e ${d_2}' \geq {d_1}' > 0$, allora la (A.2) è verificata con $c_1 = d_1/{d_2}'$ e $c_2 = d_2/{d_1}'$

Si ricorda ora la nota disuguaglianza di Cauchy-Schwarz:

$$\left|\sum_{i=1}^{n} x_i y_i\right| \leq \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2} \left(\sum_{i=1}^{n} |y_i|^2\right)^{1/2}, \tag{A.3}$$

che vale per ogni x, $y \in \mathbb{R}^n$. Sia e^i, $i = 1, ..., n$, l'i-mo vettore unitario (cioè il vettore le cui componenti sono uguali a zero, tranne l'i-ma che è uguale a 1) e sia $\|\cdot\|$ una qualsiasi norma vettoriale. Dalla (A.3) e dalla *3.* della Definizione A.1 si ha:

$$\|x\| = \left\|\sum_{i=1}^{n} x_i e^i\right\| \leq \sum_{i=1}^{n} |x_i|\|e^i\| \leq \beta\|x\|_2, \quad \beta = \left(\sum_{i=1}^{n} \|e^i\|^2\right)^{1/2}, \tag{A.4}$$

da cui si ottiene il primo membro della (A.2) con $c_1 = 1/\beta$. Inoltre, dalla (A.4) si ha:

$$|\|x\| - \|y\|| \leq \|x - y\| \leq \beta\|x - y\|_2,$$

da cui si evince che $\|\cdot\|$ è una funzione continua rispetto alla norma Euclidea. Di conseguenza, considerata la sfera unitaria in $\mathbb{R}^n$, $S = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$, che è un insieme chiuso e limitato, la funzione $\|\cdot\|$ ha un minimo, necessariamente diverso da zero, in S. Quindi, $\|x\| \geq \alpha > 0$, con $x \in S$. Sia $x \in \mathbb{R}^n$, $x \neq 0$. Allora, il vettore $x/\|x\|_2 \in S$, da cui $\|x\| \geq \alpha\|x\|_2$. Quindi, anche il secondo membro della (A.2) è vero con $c_2 = 1/\alpha$. ■

Un risultato analogo sussiste anche per le norme matriciali. In particolare si ha:

Teorema A.2. *Siano $\|\cdot\|_\beta$ e $\|\cdot\|_{\beta'}$ due norme matriciali indotte, rispettivamente, dalle norme vettoriali $\|\cdot\|_\alpha$ e $\|\cdot\|_{\alpha'}$. Allora si ha:*

$$d_1\|A\|_\beta \leq \|A\|_{\beta'} \leq d_2\|A\|_\beta, \quad \forall A \in \mathbb{R}^{n\times n}, \tag{A.5}$$

dove $d_1 = c_1/c_2$ e $d_2 = c_2/c_1$, con c_1 e c_2 della relazione (A.2).

Dimostrazione Utilizzando la definizione di norma matriciale indotta (Definizione A.2), dalla (A.2) si ha:

$$\|A\|_{\beta'} = \sup_{x\neq 0} \frac{\|Ax\|_{\alpha'}}{\|x\|_{\alpha'}} \leq \sup_{x\neq 0} \frac{c_2\|Ax\|_\alpha}{c_1\|x\|_\alpha} = (c_2/c_1)\|A\|_\beta$$

da cui il secondo membro della (A.5). Allo stesso modo, si dimostra che vale il primo membro della (A.5) con $d_1 = c_1/c_2$. ■

A.4 Esistenza di soluzioni di sistemi lineari - Regola di Cramer

Il ruolo fondamentale che i sistemi lineari giocano nella risoluzione dei problemi scientifici, comporta la necessità di determinare se un assegnato sistema ammette (o non ammette) soluzioni.

♣ **Esempio A.6.** Assegnate le due rette del piano:

$$\begin{array}{ccccc} x & + & y & = & 2; \\ x & - & 2y & = & -1, \end{array}$$

dalla loro rappresentazione grafica (fig. A.2) si deduce che esse si intersecano nel solo punto P, di coordinate $(1, 1)$.

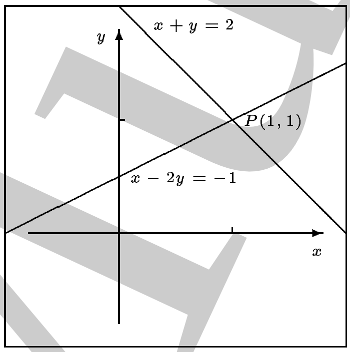

Figura A.2: Rappresentazione grafica di due rette del piano incidenti

Ciò significa che il sistema costituito dalle due equazioni delle rette "ammette una sola soluzione" data da $x = 1, y = 1$.

Considerate le due equazioni:

$$\begin{array}{ccccc} x & + & y & = & 2; \\ 2x & + & 2y & = & 4, \end{array}$$

si ha che la rappresentazione grafica delle rette evidenzia che tali rette sono coincidenti (fig. A.3), ovvero hanno infiniti punti di intersezione.

Pertanto il sistema costituito dalle equazioni delle due rette "ammette infinite soluzioni" (ovvero esistono infinite coppie di valori x, y che soddisfano il sistema).

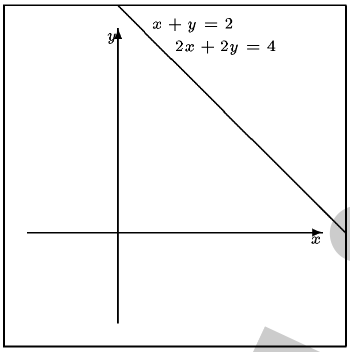

Figura A.3: Rappresentazione grafica di due rette del piano coincidenti

Infine, le due rette:

$$\begin{array}{ccccc} x & + & y & = & 2; \\ 2x & + & 2y & = & 6 \end{array}$$

risultano parallele, come si deduce dalla loro rappresentazione grafica (fig. A.4), il che vuol dire che esse non hanno punti di intersezione e pertanto il sistema di equazioni relativo "non ammette soluzioni". ♣

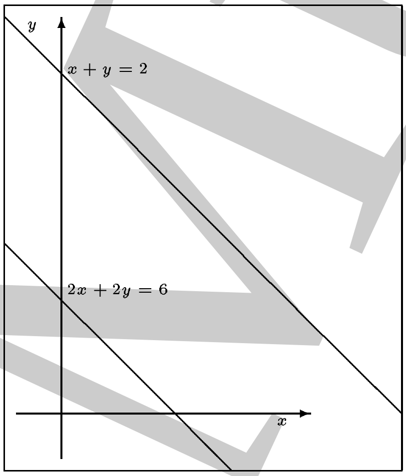

Figura A.4: Rappresentazione grafica di due rette del piano parallele

Piú in generale, un sistema di ordine n può:

- non ammettere soluzioni, in tal caso il sistema è detto *incompatibile* (o *inconsistente*);
- ammettere soluzioni, in tal caso il sistema è detto *compatibile* (o *consistente*). Un sistema compatibile o ammette una sola soluzione, e si dice *determinato*, oppure ammettere infinite soluzioni, e si dice *indeterminato*.

Come stabilire l'esistenza di soluzioni di un assegnato sistema lineare?

♣ **Esempio A.7.** Al sistema

$$\begin{cases} x + y = 2 \\ x - 2y = -1 \end{cases}$$

si applichi il metodo di addizione e sottrazione:

- sottraendo la II equazione dalla I si ottiene

$$3y = 3 \Longrightarrow y = 1$$

  e quindi, si ottiene il sistema equivalente

$$\begin{cases} x + y = 2 \\ y = 1 \end{cases}$$

- sostituendo il valore di y nella I equazione si ha

$$x + 1 = 2 \Longrightarrow x = 2 - 1 = 1,$$

e quindi, il sistema ha una sola soluzione $(x = 1, y = 1)$.

Analogamente, si applichi il metodo di addizione e sottrazione al sistema

$$\begin{cases} x + y = 2 \\ 2x + 2y = 4 \end{cases}$$

- dividendo per 2 la II equazione si ha

$$x + y = 2,$$

- sottrendo l'equazione ottenuta dalla I si ha

$$0y = 0,$$

e quindi il sistema equivalente che si ottiene

$$\begin{cases} x + y = 2 \\ 0y = 0 \end{cases}$$

si riduce ad una sola equazione in due incognite; pertanto esso ha infinite soluzioni. Infatti, assegnando un valore a scelta ad una delle due incognite si ottiene un valore per l'altra. Se, ad esempio, si sceglie di assegnare un valore arbitrario alla y si ha che le infinite coppie del tipo $(x = 2 - y, y)$ soddisfano il sistema.

Infine, si applichi il metodo di addizione e sottrazione al sistema

$$\begin{cases} x + y = 2 \\ 2x + 2y = 6 \end{cases}$$

- dividendo per 2 la II equazione si ha

$$x + y = 3,$$

- sottraendo l'equazione ottenuta dalla I, si ha il sistema equivalente

$$\begin{cases} x + y = 2 \\ \quad\;\; 0y = -1 \end{cases}$$

che, chiaramente, non ammette soluzioni, in quanto nessun numero moltiplicato per 0 restituisce -1. ♣

Da quanto visto, si deduce che, nel caso di un sistema di 2 equazioni in 2 incognite che è incompatibile (nessuna soluzione) o compatibile ma indeterminato (infinite soluzioni), la matrice dei coefficienti del sistema equivalente, ottenuto applicando il metodo di addizione e sottrazione, ha un elemento nullo sulla diagonale. Infatti,

$$\begin{cases} x + y = 2 \\ \quad\;\; 0y = 0 \end{cases} \quad \text{infinite soluzioni} \quad \Longleftrightarrow \quad A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

$$\begin{cases} x + y = 2 \\ \quad\;\; 0y = -1 \end{cases} \quad \text{nessuna soluzione} \quad \Longleftrightarrow \quad A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

Tale situazione non si verifica se il sistema ammette una sola soluzione. Nel §2.5 è illustrato un criterio per stabilire l'esistenza di soluzioni di un sistema di ordine n, che costituisce una generalizzazione di quanto ora illustrato per un sistema di ordine 2.

♣ **Esempio A.8.** Si consideri la matrice dei coefficienti del sistema

$$\begin{cases} x + y = 2 \\ 2x + 2y = 4 \end{cases}$$

che, come visto ammette infinite soluzioni,

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

Se si calcola la quantità d, differenza tra il prodotto degli elementi della diagonale principale ed il prodotto degli elementi della diagonale secondaria, ovvero

$$d = 1 \cdot 2 - 1 \cdot 2,$$

si ha $d = 0$.
Ció chiaramente si verifica anche per la matrice dei coefficienti del sistema

$$\begin{cases} x + y = 2 \\ 2x + 2y = 6 \end{cases}$$

che non ammette soluzioni.
Considerata, invece, la matrice dei coefficienti del sistema

$$\begin{cases} x + y = 2 \\ x - 2y = -1 \end{cases}$$

che ammette una sola soluzione, si ha

$$d = 1 \cdot (-2) - 1 \cdot 1 = -3.$$

♣

L' esempio precedente sembra mostrare che, considerato un sistema di ordine 2 con matrice dei coefficienti

$$\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix},$$

si ha che il numero

$$d = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}$$

è diverso da zero se il sistema ammette una sola soluzione (nel caso $d = 0$ il sistema o non ammette soluzioni o ne ammette infinite).
Tale risultato sussiste per qualsiasi sistema di ordine 2 e puó essere generalizzato ad un qualsiasi sistema si ordine n, dopo aver opportunamente definito (cos'è) d in relazione ad una matrice di dimensione n.
Ciò che si sta per definire è un singolo numero associato ad una matrice $A \in \Re^{n\times n}$, denominato *determinante* di A e denotato con $detA$.
Il determinante di una matrice di dimensione n sarà definito in termini di determinanti di matrici di ordine $(n-1)$ e così via, fina ad arrivare a matrici di dimensione 1, $A = (\alpha)$, per cui si definisce $detA = \alpha$.
A tale scopo è necessario dare dapprima la segunte definizione:

Definizione A.6. (Minori complementari e complementi algebrici)
Assegnata una matrice $A = (a_{i,j}) \in \Re^{n\times n}$ si dice

- minore complementare dell'elemento $a_{i,j}$, *e si denota con $M_{i,j}$, il determinante della matrice di ordine $(n-1)$ che si ottiene da A cancellando la i-ma riga e la j-ma colonna;*
- complemento algebrico dell'elemento $a_{i,j}$, *e si denota con $A_{i,j}$, il numero $(-1)^{i+j}M_{i,j}$.*

♣ **Esempio A.9.** Considerata

$$A = \begin{pmatrix} 1 & 3 & 4 \\ 5 & 7 & 8 \\ 9 & 10 & 2 \end{pmatrix},$$

il minore complementare dell'elemento $a_{1,1}$ è il determinante della matrice che si ottiene da A cancellando la prima riga e la prima colonna, ovvero

$$M_{1,1} = detB, \text{ con } B = \begin{pmatrix} 7 & 8 \\ 10 & 2 \end{pmatrix},$$

mentre il complemento algebrico di $a_{1,1}$ è dato da

$$A_{1,1} = (-1)^{1+1}M_{1,1} = M_{1,1}.$$

♣

È possibile, ora, dare la definizione di determinante di una matrice.

Definizione A.7. (Determinante di una matrice)

1. *Il determinante di una matrice di dimensione 1,* $A = (\alpha)$*, è dato da:* $det A = \alpha$.

2. *Il determinante di una matrice di dimensione* n*,* $A = (a_{i,j})$*, è dato da:*

$$det A = \begin{vmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{vmatrix} = a_{1,1}A_{1,1} + a_{1,2}A_{1,2} + \ldots + a_{1,n}A_{1,n} =$$

$$= \sum_{i=1}^{n} a_{1,i}A_{1,i}$$

ovvero il determinante di A *è la somma dei prodotti degli elementi della I riga,* $a_{1,1}, a_{1,2}, \ldots, a_{1,n}$*, per i rispettivi complementi algebrici* $A_{1,1}, A_{1,2} \ldots, A_{1,n}$*. (Piú in generale il* $det A$ *è uguale alla somma dei prodotti degli elementi di una riga (o di una colonna) per i rispettivi complementi algebrici).*

♣ **Esempio A.10.** Si consideri il determinante di una matrice di dimensione 2

$$\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}.$$

Per la definizione di determinante si ha:

$$det A = a_{1,1}A_{1,1} + a_{1,2}A_{1,2} = a_{1,1}det(a_{1,2}) - a_{1,2}det(a_{2,2}) = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}$$

che rappresenta la quantità d introdotta precedentemente.
Per esempio:

$$det \begin{pmatrix} -4 & -5 \\ 2 & 3 \end{pmatrix} = (-4) \cdot 3 - (-5) \cdot 2 = -12 + 10 = -2$$

♣

A questo punto è possibile ricordare il seguente ben noto risultato riguardo l'esistenza, l'unicità e l'espressione della soluzione di un sistema lineare di ordine n.

Definizione A.8. (Regola di Cramer)
Assegnato un sistema lineare di ordine n:

$$A\underline{x} = \underline{b},$$

se $detA \neq 0$, *il sistema ammette una ed una sola soluzione* $\underline{x} = (x_i), i = 1, 2, \ldots, n$ *con*

$$x_i = detB_i/detA,$$

dove B_i *è la matrice di dimensione n che si ottiene da A sostituendo la i-ma colonna con il vettore dei termini noti* $\underline{b}$.

Si consideri, ora la seguente:

Definizione A.9. (Singolarità di una matrice)
Sia $A \in \Re^{n\times n}$, *si definisce* **singolare** *se e soltanto se non ammette matrice inversa. Al contrario,* $A \in \Re^{n\times n}$, *si definisce* **non singolare** *se e soltanto se ammette inversa, ovvero esiste una matrice* $X \in \Re^{n\times n}$ *tale che* $AX = XA = I_n$ $(X = A^{-1})$.

La singolarità di una matrice $A \in \Re^{n\times n}$ è legata all'esistenza di soluzioni di un sistema lineare avente A come matrice dei coefficienti. Vale, infatti, la seguente equivalenza:

$$A \text{ non singolare} \iff A\underline{x} = \underline{b} \text{ ammette una sola soluzione.}$$

Ciò fornisce un modo per stabilire se una matrice $A \in \Re^{n\times n}$ è non singolare. Infatti, si risolve un sistema

$$A\underline{x} = \underline{b},$$

se esso ammette una sola soluzione allora si puó affermare che A è non singolare. Viceversa, assegnato un sistema di ordine n

$$A\underline{x} = \underline{b},$$

si puó pensare di risolverlo calcolando A^{-1}, se A è non singolare, e ponendo $\underline{x} = A^{-1}\underline{b}$. In tal caso, si presenta il problema di come calcolare l'inversa di una matrice. L'esempio A.1 mette in luce che per fare ció è necessario risolvere dei sistemi lineari, che è esattamente il problema di partenza!

Un **metodo per calcolare l'inversa** di una assegnata matrice A consiste nel determinare una alla volta le colonne di A^{-1} risolvendo un sistema lineare per ciascuna di esse. Infatti, ricordando che

$$AA^{-1} = I_n = (\underline{e}_1\ \underline{e}_2\ \ldots\ \underline{e}_n),$$

e considerata la decomposizione per colonne di A^{-1}, cioé

$$A^{-1} = (\underline{x}_1\ \underline{x}_2\ \ldots\ \underline{x}_n),$$

si ha che:

$$A\underline{x}_i = \underline{e}_i, \quad i = 1, \ldots, n.$$

Pertanto le colonne di A^{-1} possono essere calcolate risolvendo n sistemi lineari aventi tutti come matrice dei coefficienti la matrice A. Da quanto detto si evince anche che il calcolo dell'inversa di una matrice A (e di $A^{-1}\underline{b}$) richiede piú operazioni di quelle necessarie per la risoluzione di un sistema $A\underline{x} = \underline{b}$.

Dalla regola di Cramer si deduce anche l'equivalenza tra non singolarità di una matrice A e valore del suo determinante, ovvero:

$$A \in \Re^{n\times n} \text{ non singolare} \iff detA \neq 0$$

Infatti, se $detA \neq 0$ la regola di Cramer stabilisce che il sistema ammette una sola soluzione. Ciò assicura, per quanto detto, che la matrice A è non singolare.
La regola di Cramer sembra fornire un metodo per calcolare la soluzione di un sistema, ma in realtà l'applicazione della regola è, come illustrato nel §2.3, computazionalmente impraticabile.

A.5 Autovalori ed autovettori

Definizione A.10. (Autovalori ed autovettori)
Data $A \in \mathbb{R}^{n\times n}$, se $Ax = \lambda x$, con $x \in \mathbb{R}^n$, $x \neq 0$ e λ numero complesso, allora λ è detto un **autovalore** *di A e x un* **autovettore** *ad esso associato. L'insieme di tutti gli autovalori di A è chiamato* **spettro** *di A ed è denotato con $S(A)$.*

Si osservi che se x è un autovettore di $A \in \mathbb{R}^{n\times n}$ associato all'autovalore λ, allora si ha:

$$A(\alpha x) = \alpha Ax = \alpha\lambda x = \lambda(\alpha x), \ \ \forall \alpha \in \mathbb{R}, \ \alpha \neq 0,$$

cioè αx, con $\alpha \neq 0$, è un autovettore di A associato allo stesso autovalore λ.

Gli autovalori di una matrice sono caratterizzati dalla seguente proprietà:

Teorema A.3. *Sia $A \in \mathbb{R}^{n\times n}$. Il numero complesso λ è un autovalore di A se e solo se* $\mathtt{det}(A - \lambda I) = 0$.

Dimostrazione Dalla definizione di autovalore si ha $Ax = \lambda x$, con $x \neq 0 \iff Ax - \lambda x = 0$, e quindi se e solo se:

$$(A - \lambda I)x = 0, \ \ x \neq 0. \tag{A.6}$$

La (A.6) è un sistema di equazioni lineari omogeneo con una soluzione non nulla. Di conseguenza, la matrice $(A - \lambda I)$ è singolare e quindi il suo determinante è nullo. ■

Si osservi che $\texttt{det}(A - \lambda I)$ è un polinomio di grado n in λ, detto **polinomio caratteristico** di A. Dal Teorema fondamentale dell'Algebra e dal Teorema A.3 segue che una matrice $A \in \mathbb{R}^{n\times n}$ ha esattamente n autovalori (alcuni dei quali possono essere multipli). Inoltre, il polinomio caratteristico può essere espresso nella forma:

$$(-1)^n \texttt{det}(A - \lambda I) = \lambda^n - \left(\sum_{i=1}^{n} a_{ii}\right)\lambda^{n-1} + ... + (-1)^n \texttt{det}(A),$$

dalla quale, poichè il prodotto delle radici di un polinomio di grado n è uguale al termine noto del polinomio stesso, si ottiene che:

$$\texttt{det}(A) = \prod_{i=1}^{n} \lambda_i \tag{A.7}$$

dove $\lambda_1, ..., \lambda_n$ sono gli autovalori di A.

Definizione A.11. (Matrici simili)
Due matrici A e B quadrate di ordine n sono **simili** *se esiste una matrice P non singolare tale che $B = P^{-1}AP$.*

Teorema A.4. *Se A e B sono matrici simili, allora hanno gli stessi autovalori.*

Dimostrazione Se A e B sono simili, allora esiste una matrice P non singolare tale che $B = P^{-1}AP$. Sia λ un autovalore di A e sia x un autovettore ad esso associato. Si ha quindi $Ax = \lambda x$. Posto $v = P^{-1}x$ si ha:

$$Bv = P^{-1}APP^{-1}x = P^{-1}Ax = \lambda P^{-1}x = \lambda v,$$

da cui λ è anche un autovalore di B. ■

Definizione A.12. (Raggio spettrale)
Sia $A \in \mathbb{R}^{n\times n}$. Si definisce **raggio spettrale** *di A la quantità:*

$$\rho(A) = \max_{1\le i\le n} |\lambda_i|,$$

con λ_i i-mo autovalore di A.

Teorema A.5. *Data una matrice A di ordine n, se $\rho(A) < 1$ la matrice $I - A$ è non singolare.*

Dimostrazione Sia λ un autovalore di A e sia x un autovettore ad esso associato. Posto $\mu = 1 - \lambda$ si ha:

$$(I - A)x = x - Ax = x - \lambda x = (1 - \lambda)x = \mu x,$$

da cui μ è un autovalore della matrice $I - A$. Di conseguenza, se $\rho(A) < 1$, gli autovalori di $I - A$ sono tutti diversi da zero e, dalla (A.7), si ha che $I - A$ è non singolare. ■

A.5.1 Localizzazione degli autovalori

Molti teoremi in letteratura descrivono dove sono situati, di massima, gli autovalori di una matrice nel piano complesso. Il più famoso tra questi *teoremi di localizzazione* è il seguente [Kincaid,Cheney]:

Teorema A.6. [Teorema di Gershgorin]
Lo spettro di una matrice A, $n \times n$, è contenuto nell'unione dei seguenti n dischi, D_i, nel piano complesso:

$$D_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|\} \quad (1 \leq i \leq n)$$

Dimostrazione Sia λ un qualsiasi elemento dello spettro di A. Si scelga un vettore x tale che $Ax = \lambda x$ e $\|x\|_\infty = 1$. Sia i un indice per il quale $|x_i| = 1$. Poiché

$$(Ax)_i = \lambda x_i,$$

si ha

$$\lambda x_i = \sum_{j=1}^{n} a_{ij} x_j.$$

Quindi

$$(\lambda - a_{ii}) x_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j.$$

Prendendo i valori assoluti e sfruttando la disuguaglianza triangolare e la condizione

$$|x_j| \leq \|x\|_\infty = 1 = |x_i|$$

si ha

$$|\lambda - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| |x_j| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|.$$

Ne segue che $\lambda \in D_i$ e, dunque, la tesi. ■

A.6 Matrici simmetriche e definite positive

Definizione A.13. (Matrice simmetrica)
Una matrice $A \in \mathbb{R}^{n\times n}$ *è* **simmetrica** *se* $A = A^T$.

Definizione A.14. (Matrice ortogonale)
Una matrice $A \in \mathbb{R}^{n\times n}$ *è* **ortogonale** *se* $AA^T = I$.

Alcune proprietà delle matrici simmetriche sono stabilite dai seguenti Teoremi, che si riportano senza dimostrazione.

Teorema A.7. *Se* $A \in \mathbb{R}^{n\times n}$ *è simmetrica, allora si ha:*

$$A = UDU^T,$$

con $U \in \mathbb{R}^{n\times n}$ *ortogonale e* $D \in \mathbb{R}^{n\times n}$, $D = \mathtt{diag}[\lambda_1, ..., \lambda_n]$, *con* $\lambda_i \in S(A)$. *I vettori colonna della matrice* U *sono autovettori di* A, *che quindi ha un insieme di* n *autovettori ortonormali.*

Dal Teorema precedente segue che gli autovalori di una matrice simmetrica sono reali. Inoltre, si ha:

Teorema A.8. *Se* $A \in \mathbb{R}^{n\times n}$ *è simmetrica e se* λ_{max} *e* λ_{min} *sono il più grande e il più piccolo autovalore di* A, *rispettivamente, si ha:*

$$\lambda_{max} = \max_{x\neq 0} \frac{\langle x, Ax\rangle}{\langle x, x\rangle} = \frac{\langle v_{max}, Av_{max}\rangle}{\langle v_{max}, v_{max}\rangle};$$

$$\lambda_{min} = \min_{x\neq 0} \frac{\langle x, Ax\rangle}{\langle x, x\rangle} = \frac{\langle v_{min}, Av_{min}\rangle}{\langle v_{min}, v_{min}\rangle},$$

dove v_{max} *e* v_{min} *sono autovettori di* A *associati rispettivamente a* λ_{max} *e* λ_{min}.

Definizione A.15. (Matrice definita positiva *e* **semidefinita positiva)**
Una matrice $A \in \mathbb{R}^{n\times n}$ *è* **definita positiva** *se è simmetrica e se*

$$\langle x, Ax\rangle > 0, \ \ \forall x \in \mathbb{R}^n, \ \ x \neq 0.$$

Se $\langle x, Ax\rangle \geq 0$, $\forall x \in \mathbb{R}^n$, *allora* A *è detta* **semidefinita positiva**.

Alcune tra le caratterizzazioni delle matrici definite positive sono le seguenti. Dalla Definizione A.15 e dal Teorema A.8 si ha immediatamente:

Teorema A.9. *Una matrice* A *simmetrica è definita positiva (semidefinita positiva) se e solo se tutti i suoi autovalori sono positivi (non negativi).*

Teorema A.10. *Data una matrice $A \in \mathbb{R}^{n\times n}$, la matrice AA^T è simmetrica e semidefinita positiva. Se A è non singolare, allora AA^T è definita positiva.*

Dimostrazione Data una matrice $A \in \mathbb{R}^{n\times n}$ è immediato verificare che la matrice AA^T è simmetrica. Inoltre si ha:

$$\langle x, AA^T x\rangle = \langle A^T x, A^T x\rangle \geq 0, \;\; \forall x \in \mathbb{R}^n,$$

da cui, in base alla Definizione A.15, la matrice AA^T è semidefinita positiva.
Se A è non singolare e se $\langle x, AA^T x\rangle = 0$ si ha $A^T x = 0$ e $x = 0$, da cui la matrice AA^T è definita positiva. ■

Per una matrice simmetrica vale il seguente:

Teorema A.11. *Se $A \in \mathbb{R}^{n\times n}$ è simmetrica si ha:*

$$\|A\|_2 = \rho(A).$$

Dimostrazione Si osservi innanzitutto che se λ è un autovalore di A e x è un autovettore ad esso associato si ha:

$$A^2 x = A(Ax) = A(\lambda x) = \lambda Ax = \lambda^2 x,$$

cioè λ^2 è un autovalore di A^2. Quindi, poichè A è simmetrica, si ha:

$$\|A\|_2 = \sqrt{\rho(A^T A)} = \sqrt{\rho(A^2)} = \sqrt{\rho(A)^2} = \rho(A).$$

■

Segue, inoltre, il seguente:

Teorema A.12. *Sia A una matrice $\in \mathbb{R}^{n\times n}$ e sia $\|\cdot\|_\beta$ una norma matriciale. Si ha:*

$$\rho(A) \leq \|A\|_\beta \,.$$

Dimostrazione Sia λ un qualsiasi autovalore di A e sia x un autovettore ad esso associato. Quindi:

$$Ax = \lambda x.$$

Indicata con $\|\cdot\|_\alpha$ una norma vettoriale tale che $\|\cdot\|_\beta$ sia una norma matriciale con essa compatibile, si ha:

$$|\lambda| \, \|x\|_\alpha = \|\lambda x\|_\alpha = \|Ax\|_\alpha \leq \|A\|_\beta \|x\|_\alpha$$

da cui:

$$|\lambda| \leq \|A\|_\beta, \quad \forall \lambda \, \in S(A),$$

e quindi l'asserto. ■

Si riporta infine un Teorema dal quale segue che, nel caso in cui $\rho(A) < 1$, è sempre possibile trovare una norma matriciale $\|\cdot\|_\beta$ tale che $\|A\|_\beta < 1$. Per dimostrare tale Terorema, importante nell'ambito dell'analisi della convergenza delle successioni di matrici, oggetto del prossimo paragrafo, si richiama il seguente risultato, che può essere visto come una generalizzazione del Teorema A.7.
Data una matrice $A \in \mathbb{R}^{n\times n}$, esiste una matrice non singolare $P \in \mathbb{R}^{n\times n}$ tale che:

$$P^{-1}AP = J,$$

dove la matrice $J \in \mathbb{R}^{n\times n}$, detta **forma canonica di Jordan** di A, ha la seguente struttu ra:

$$J = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & 0 \\ 0 & 0 & \cdots & J_p \end{pmatrix},$$

cioè è una matrice **diagonale a blocchi**, dove ciascun blocco J_i, detto **blocco di Jordan**, è uno scalare o una matrice della forma:

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{pmatrix}, \tag{A.8}$$

con λ_i autovalore di A di molteplicità almeno uguale alla dimensione di J_i. Si osservi che, se tutti gli autovalori di A sono distinti, la forma canonica di Jordan di A è una matrice diagonale.

Teorema A.13. *Sia A una matrice $\in \mathbb{R}^{n\times n}$ e sia $\epsilon > 0$. Esiste una norma matriciale tale che:*

$$\|A\|_\beta \leq \rho(A) + \epsilon\ .$$

Dimostrazione Sia P una matrice non singolare tale che $P^{-1}AP = J$, dove J è la forma canonica di Jordan di A, e sia D la matrice diagonale $\mathtt{diag}[1, \epsilon, ..., \epsilon^{n-1}]$. Denotata con $J^{(\epsilon)}$ la matrice ottenuta a partire da J sostituendo con ϵ gli elementi uguali a 1 in (A.8), si verifica facilmente che $D^{-1}JD = J^{(\epsilon)}$, da cui:

$$J^{(\epsilon)} = L^{-1}AL, \quad L = PD. \tag{A.9}$$

In base alla struttura di $J^{(\epsilon)}$ si ha, considerando la norma 1 ed indicando con λ_i gli autovalori di A:

$$\|J^{(\epsilon)}\|_1 = \max_{1\leq k\leq n} \sum_{i=1}^{n} |j_{ik}| = \max_{1\leq k\leq n} |\lambda_k + \epsilon| \leq \rho(A) + \epsilon. \tag{A.10}$$

Se si considera la norma matriciale $\|\cdot\|_{1,L}$, si ha:

$$\|A\|_{1,L} = \|L^{-1}AL\|_1 = \|J^{(\epsilon)}\|_1,$$

e, quindi, dalla (A.10), l'asserto. ■

Un'immediata conseguenza del risultato appena dimostrato è la seguente:

Corollario A.1. *Data una matrice $A \in \mathbb{R}^{n\times n}$, se $\rho(A) < 1$, allora esiste una matrice non singolare L tale che:*

$$\|A\|_{1,L} < 1.$$

Dimostrazione Posto $\epsilon = (1 - \rho(A))/2$, dal Teorema A.13 si ha:

$$\|A\|_{1,L} < \rho(A) + (1 - \rho(A))/2 = (1 + \rho(A))/2,$$

da cui, se $\rho(A) < 1$, si ha $\|A\|_{1,L} < 1$. ■

A.7 Successioni di vettori e matrici

Per lo studio dei metodi iterativi per la risoluzione di sistemi di equazioni lineari è importante il concetto di successioni di vettori e matrici e le loro proprietà di convergenza.

Definizione A.16. (Successione convergente di vettori)
Una successione di vettori $\{v^{(k)}\}$, $v^{(k)} \in \mathbb{R}^n$, si dice che **converge** *ad un vettore $v \in \mathbb{R}^n$ se:*

$$\lim_{k\to\infty} v_i^{(k)} = v_i, \quad i = 1, ..., n.$$

Definizione A.17. (Successione convergente di matrici)
Una successione di matrici $\{A^{(k)}\}$, $A^{(k)} \in \mathbb{R}^{n\times n}$, si dice che **converge** *ad una matrice $A \in \mathbb{R}^{n\times n}$ se:*

$$\lim_{k\to\infty} a_{ij}^{(k)} = a_{ij}, \quad i, j = 1, ..., n.$$

Teorema A.14. *Una successione di vettori* $\{v^{(k)}\}$, $v^{(k)} \in \mathbb{R}^n$, *converge a* $v \in \mathbb{R}^n$ *se e solo se, per ogni norma vettoriale* $\|\cdot\|_\alpha$, *si ha:*

$$\lim_{k\to\infty} \|v^{(k)} - v\|_\alpha = 0.$$

Dimostrazione Per definizione di norma del massimo, è evidente che $v^{(k)} \to v$ se e solo se $\|v^{(k)} - v\|_\infty \to 0$. Poichè, in base al Teorema A.1, tutte le norme vettoriali sono equivalenti, si ha l'asserto. ∎

Utilizzando il Teorema A.2 sull'equivalenza delle norme matriciali, si dimostra in maniera simile che

Teorema A.15. *Una successione di matrici* $\{A^{(k)}\}$, $A^{(k)} \in \mathbb{R}^{n\times n}$, *converge alla matrice* $A \in \mathbb{R}^{n\times n}$ *se e solo se, per ogni norma matriciale* $\|\cdot\|_\beta$, *si ha:*

$$\lim_{k\to\infty} \|A^{(k)} - A\|_\beta = 0.$$

Teorema A.16. *Sia* A *una matrice* $\in \mathbb{R}^{n\times n}$. *Considerata la successione di potenze di* A, $\{A^k\}$, $A^k \in \mathbb{R}^{n\times n}$, *si ha che* $A^k \to 0$ *se e solo se* $\rho(A) < 1$.

Dimostrazione Se $\rho(A) < 1$, in base al Corollario A.1 esiste una norma matriciale $\|\cdot\|_\beta$ tale che $\|A\|_\beta < 1$. Poichè $\|A^k\|_\beta \leq \|A\|_\beta^k$ si ha che $\|A^k\|_\beta \to 0$. Dal Teorema A.15 segue che $A^k \to 0$. Si supponga ora che $A^k \to 0$, da cui $\|A^k\|_\beta \to 0$. Se per assurdo $\rho(A) \geq 1$, poichè $\forall k \geq 0$ gli autovalori di A^k sono dati da $\{\lambda_i^k\}_{i=1}^n$, con λ_i i-mo autovalore di A, si ha $\rho(A^k) \geq 1$. Dal Teorema A.12 si ha quindi $\|A^k\|_\beta \geq \rho(A^k) \geq 1$, $\forall k \geq 0$, ma ciò contraddice l'ipotesi. ∎

Teorema A.17. *Sia* A *una matrice* $\in \mathbb{R}^{n\times n}$. *Allora*

$$\lim_{k\to\infty} A^k v = 0, \quad \forall v \in \mathbb{R}^n \iff \rho(A) < 1.$$

Dimostrazione In base al Teorema A.16, è sufficiente dimostrare che la condizione $\lim_{k\to\infty} A^k v = 0, \ \ \forall v \in \mathbb{R}^n$ vale se e solo se $A^k \to 0$. Se $A^k \to 0$, dal Teorema A.15 si ha $\|A^k\|_\beta \to 0$. Poichè $\|A^k v\|_\alpha \leq \|A^k\|_\beta \, \|v\|_\alpha$, se $\|\cdot\|_\alpha$ è una norma vettoriale tale che $\|\cdot\|_\beta$ sia una norma matriciale con essa compatibile, segue che $\|A^k v\|_\alpha \to 0$ e, dal Teorema A.14, che $A^k v \to 0$. Si supponga ora che $\lim_{k\to\infty} A^k v = 0$ per ogni vettore v e si assuma, per assurdo, che la successione $\{A^k\}$ non converga a zero. Dal Teorema A.16 si ha $\rho(A) \geq 1$. Indicato con λ un autovalore di A tale che $|\lambda| \geq 1$ e con v un autovettore ad esso associato, si ha $\|A^k v\|_\alpha = |\lambda|^k \|v\|_\alpha$, da cui si evince che $\|A^k v\|_\alpha$ non converge a zero, e ciò contraddice l'ipotesi. ∎

A.8 Polinomi di matrici

Definizione A.18. (Polinomio di matrice)
Sia

$$p(x) = a_0 + a_1 x + ... + a_k x^k, \quad a_i \in \mathbb{R},\ 0 \le i \le k, \quad a_k \neq 0$$

un polinomio di grado k. Data la matrice $A \in \mathbb{R}^{n\times n}$, si definisce **polinomio in** *A la matrice:*

$$p(A) = a_0 + a_1 A + ... + a_k A^k.$$

Un'importante relazione tra gli autovalori di una matrice A e quelli della matrice $p(A)$ è data dal seguente:

Teorema A.18. *Sia λ un autovalore di $A \in \mathbb{R}^{n\times n}$ e sia $p(x)$ un polinomio di grado k. Allora $p(\lambda)$ è un autovalore di $p(A)$.*

Dimostrazione Sia x un autovettore associato a λ. È noto che λ^i è un autovalore di A^i, $\forall i$; si ha, quindi:

$$A^i x = \lambda^i x, \ \ \forall i. \tag{A.11}$$

Dalla (A.11) si ottiene:

$$p(A)x = a_0 x + a_1 Ax + ... + a_k A^k x = a_0 x + a_1 \lambda x + ... + a_k \lambda^k x = p(\lambda)x,$$

da cui l'asserto. ■

Utilizzando il Teorema A.18 si può stabilire un'analogia, per il calcolo algebrico simbolico, tra le matrici simmetriche e definite positive ed i numeri reali.

Teorema A.19. *Sia $A \in \mathbb{R}^{n\times n}$ simmetrica e definita positiva. Allora esiste, ed è unica, la matrice $B \in \mathbb{R}^{n\times n}$ simmetrica e definita positiva tale che $B^2 = A$. Tale matrice è detta la* **radice quadrata di** *A.*

Dimostrazione Poichè A è definita positiva, i suoi autovalori $\{\lambda_i\}_{i=1}^n$ sono tutti positivi ed è quindi possibile definire la matrice $D_0 = \mathtt{diag}[\sqrt{\lambda_1}, ..., \sqrt{\lambda_n}]$. Inoltre, poichè A è simmetrica, in base al Teorema A.7, esiste una matrice U ortogonale tale che $A = UDU^T$ e $D = D_0^2$. Allora la matrice simmetrica:

$$B = UD_0U^T \tag{A.12}$$

è tale che $B^2 = A$. Si ha infatti:

$$B^2 = UD_0U^TUD_0U^T = UD_0^2U^T = UDU^T = A.$$

Si osservi che, dalla (A.12), si ricava che B e D_0 sono simili, e quindi hanno gli stessi autovalori. Di conseguenza, gli autovalori di B sono le radici quadrate degli autovalori di A e sono positivi. Ciò mostra che B è anche definita positiva.
Si supponga ora che esista un'altra matrice C simmetrica e definita positiva tale che $C^2 = A$. Dal Teorema A.18 si ha che gli autovalori di C sono le radici quadrate positive degli autovalori di A. Infatti, considerato il polinomio $p(x) = x^2$, se λ è un autovalore di A e quindi di C^2, si ha $\lambda = p(\mu) = \mu^2$, con μ autovalore di C. Poichè $\lambda > 0$, si ha $\mu = \sqrt{\lambda} > 0$. Gli autovalori di B e di C sono quindi uguali. Inoltre, poichè C è simmetrica, si ha $C = UD_0U^T$ e $A = C^2 = UD_0^2U^T$, da cui $C = B$. ■

A.9 Alcune proprietà delle matrici simmetriche definite positive

Il teorema che segue fornisce una caratterizzazione delle matrici simmetriche definite positive.

Teorema A.20. *Sia $A \in R^{n\times n}$ simmetrica. A è definita positiva se e solo se tutti i suoi autovalori sono positivi.*

Dimostrazione E' noto che, essendo A simmetrica, tutti i suoi autovalori sono reali. Si supponga che A sia definita positiva e si considerino un generico autovalore λ_i ed un autovettore reale q_i ad esso associato, cioè tale che

$$Aq_i = \lambda_i q_i\,. \tag{A.13}$$

Moltiplicando a sinistra per q_i^T ambo i membri di (A.13), si ha

$$q_i^T A q_i = \lambda_i q_i^T q_i\,.$$

Essendo A definita positiva, risulta $q_i^T A q_i > 0$ e quindi, dato che $q_i^T q_i > 0$, si ha necessariamente $\lambda_i > 0$.

Si supponga ora che tutti gli autovalori di A siano positivi. Dato che A è simmetrica, esiste una matrice $Q \in \Re^{n\times n}$ ortogonale tale che $A = Q^T D Q$, con $D = diag\,(\lambda_1, \ldots, \lambda_n)$, dove λ_i è il generico autovalore di A. Per ogni $x \in \Re^n$, $x \neq 0$, si ha dunque

$$x^T A x = x^T Q^T D Q x = (Qx)^T D\,(Qx)\,.$$

Posto $v = Qx$ e indicata con v_i la i-ma componente di v, risulta $v \neq 0$, in quanto Q è non singolare, e quindi, essendo $\lambda_i > 0$ per $i = 1, \ldots, n$, si ha:

$$x^T A x = v^T D v = \lambda_1 v_1^2 + \ldots + \lambda_n v_n^2 > 0.$$

■

Il seguente teorema mostra alcune proprietà delle matrici simmetriche definite positive.

Teorema A.21. *Sia $A \in R^{n\times n}$ simmetrica definita positiva. Si ha:*

1. *$det(A) > 0$ e quindi A è non singolare;*

2. *tutti gli elementi diagonali di A sono positivi;*

3. *il massimo in modulo tra gli elementi di A appartiene alla diagonale principale;*

4. *ogni matrice ottenuta da A eliminando j righe, $0 < j < n$, e le corrispondenti colonne è definita positiva; in particolare, ogni sottomatrice principale A_k di A, $k = 1, \ldots, n$, è definita positiva;*

5. *$(a_{i,j})^2 < a_{i,i}a_{j,j}$, per ogni $i \neq j$.*

Dimostrazione

1. E' noto che

$$det(A) = \prod_{i=1}^{n} \lambda_i,$$

dove λ_i è il generico autovalore di A, e quindi, per il Teorema A.20, risulta $det(A) > 0$.

2. Per ogni $i = 1, \ldots, n$, considerato l'i-mo vettore unitario $e_i = (e_i^1, \ldots, e_i^n)^T$, con $e_i^i = 1$ ed $e_i^j = 0$ per $j \neq i$, si ha:

$$0 < e_i^T A e_i = a_{i,i}\,.$$

3. Per $k \neq j$, sia $x = (x_1, \ldots, x_n)^T$ il vettore cosí definito:

$$x_i = \begin{cases} 0 & i \neq j, \quad i \neq k \\ 1 & i = j \\ -1 & i = k \end{cases}.$$

Dato che $x \neq 0$ ed A è simmetrica, risulta

$$0 < x^T A x = a_{j,j} - a_{k,k} - a_{j,k} - a_{k,j} = a_{j,j} - a_{k,k} - 2a_{k,j}\,,$$

da cui:

$$2a_{k,j} < a_{j,j} - a_{k,k}\,.$$

Considerato il vettore $z = (z_1, \ldots, z_n)^T$ tale che

$$z_i = \begin{cases} 0 & i \neq j, \quad i \neq k \\ 1 & i = j \quad \text{oppure} \quad i = k \end{cases},$$

con un ragionamento analogo al precedente si ha:

$$-2a_{k,j} < a_{j,j} + a_{k,k}\,.$$

Di conseguenza, per $k \neq j$, risulta

$$|a_{k,j}| < \frac{a_{j,j} + a_{k,k}}{2} \leq \max_{1\leq i\leq n} a_{i,i}$$

e quindi

$$\max_{1\leq k,j\leq n} |a_{k,j}| \leq \max_{1\leq i\leq n} a_{i,i}\,.$$

4. Sia $\tilde{A}$ una sottomatrice di A ottenuta eliminando j righe, $0 < j < n$, e le corrispondenti colonne. Sia $v = (v_1, \ldots, v_n)^T$ tale che $v_i = 0$ se la i-ma riga e la i-ma colonna di A sono state eliminate per ottenere $\tilde{A}$, e sia $\tilde{v}$ il vettore ottenuto da v eliminando gli elementi (nulli) corrispondenti alle righe eliminate da A. Si ha:

$$v^T A v = \tilde{v}^T \tilde{A} \tilde{v}$$

e quindi, dato che A è definita positiva e $v \neq 0$, si ha:

$$0 < v^T A v = \tilde{v}^T \tilde{A} \tilde{v}.$$

Per la generalità di $\tilde{v}$, $\tilde{A}$ risulta definita positiva.

5. Si verifica che, se si applica una stessa permutazione alle righe ed alle colonne di A, si ottiene una matrice simmetrica definita positiva. Di conseguenza, in base alle proposizioni 1 e 4, risulta

$$0 < det \begin{pmatrix} a_{i,i} & a_{i,j} \\ a_{i,j} & a_{j,j} \end{pmatrix} = a_{i,i} a_{j,j} - (a_{i,j})^2.$$

■

Un'altra caratterizzazione delle matrici simmetriche definite positive è fornita dal criterio di Sylvester, di seguito riportato (senza dimostrazione).

Teorema A.22. [Criterio di Sylvester]
Una matrice simmetrica $A \in R^{n \times n}$ è definita positiva se e solo se

$$det(A_k) > 0, \quad k = 1, \ldots, n,$$

dove A_k è la sottomatrice principale di A di ordine k:

$$A_k = (a_{i,j})_{\substack{1 \le i \le k \\ 1 \le j \le k}}.$$

Bibliografia

[1] Alonso M., Finn E. - *Elementi di Fisica per l'Università* - Inter. European Eds. (1974)

[2] Coleman T., Van Loan C. - *Handbook for Matrix Computations* - SIAM (1988)

[3] Conte S.D., De Boor C. - *Elementary Numerical Analysis*

[4] Demmel J.W. - *Applied Numerical Linear Algebra* - SIAM, Philadelphia, 1997.

[5] Forsythe G.E., Moler C.B. - *Computer Solution of Linear Algebraic Equations*

[6] Gill R.E., Murray W., Wright M. - *Numerical Linear Algebra and Optimization* - Addison Wesley (1991)

[7] Golub G.H., Van Loan C.F. - *Matrix Computations* - third ed., The Johns Hopkins University Press, 1997.

[8] Henry S. - *Elementi di Matematica per lo Studio dell'Economia* - Il Mulino (1969)

[9] Hultquist P.F. - *Numerical Methods for Engineers and Computer Scientists* - The Benjamin/Cummings Pub. Comp. (1988)

[10] Jennings A. - *Matrix Computation for Engineers and Scientists* - John Wiley and Sons (1977)

[11] Livesley R.K. - *Mathematical Methods for Engineers* - Ellis Horwood Series in Mathematics and its Applications (1989)

[12] Morris J. L. - *Computational Methods in Elementary Numerical Analysis* - John Wiley & Sons (1983)

[13] Noble B., Daniel J. - *Applied Linear Algebra* - Prentice-Hall Int. Inc. (1988)

[14] Ortega J.M. - *Numerical Analysis* - SIAM (1990)

[15] Rice J. - *Matrix Computation and Mathematical Software* - Mc Graw Hill (1981)

[16] Strang G. - *Linear Algebra and Its Applications* - Academic press (1970)

[17] Trefethen L.N., Bau D. - *Numerical linear algebra* - SIAM, 1997.

[18] Vernin G., Chanan M. - *Computer Aids to Chemistry* - Ellis Horwood Series in Chemical Science (1986)