RENATO LEONI

# Algebra lineare

## per le applicazioni statistiche

UNIVERSITÀ DI FIRENZE
DIPARTIMENTO DI STATISTICA "G. PARENTI"
FIRENZE, 2007

Questo lavoro riproduce, con alcuni cambiamenti di forma e contenuto, il volume:

Renato Leoni, *Algebra lineare per le applicazioni statistiche*, Dipartimento Statistico della Università degli Studi di Firenze, Serie Didattica n.13, 1994.

# INDICE

# PREFAZIONE

Questo volume è principalmente rivolto agli studenti delle Facoltà di Economia e di Statistica e ai partecipanti ai Corsi di dottorato in Statistica.

In esso sono presentate quelle nozioni di Algebra lineare correntemente impiegate nelle applicazioni statistiche.

Allo scopo di non appesantire eccessivamente l'esposizione e nel tentativo di conferirle semplicità e concretezza, abbiamo evitato di trattare aspetti e problematiche, anche rilevanti, ma che ci avrebbero inevitabilmente portati a dare al volume un taglio che non riteniamo del tutto adatto alle esigenze di coloro ai quali il volume stesso è rivolto.

È ovvio che, nella scelta degli argomenti e nel modo di proporli, ci è stata di guida l'esperienza acquisita in molteplici anni di insegnamento universitario.

Ma è anche naturale che il volume rifletta, almeno in parte, gli interessi culturali e di ricerca in cui al momento ci sentiamo maggiormente impegnati.

Il lettore troverà così, in aggiunta ai consueti argomenti di Algebra lineare, la presentazione di temi che sono di interesse prevalente, anche se non esclusivo, di specifici ambiti della Statistica.

In particolare, sono state introdotte – di solito nella parte riservata ai Complementi che accompagna ciascun capitolo – nozioni che hanno larga applicazione nel campo della cosiddetta Analisi statistica multidimensionale.

A quest'ultimo tema sono dedicati altri lavori che ci offrono tra l'altro l'opportunità di trattare, nella sede adatta e in modo non superficiale, alcuni tra i principali e significativi impieghi dell'Algebra lineare in ambito statistico.

Desideriamo infine rivolgere un caldo ringraziamento ai numerosi amici e colleghi che ci hanno aiutato fornendoci preziosi consigli e suggerimenti.

R. L.

# Algebra lineare

## per le applicazioni statistiche

# Cap. I
# SPAZI VETTORIALI

## 1 GENERALITÀ SUI VETTORI

**1.1** Chiamiamo *vettore* di *ordine* n ogni n-pla ordinata di numeri reali [1]. Tali numeri sono detti *elementi* o *componenti* del vettore. Possono essere disposti secondo una linea verticale (detta *colonna*) od orizzontale (detta *riga*) e sono racchiusi tra parentesi quadre [2]. Più precisamente, nel primo caso si parla di *vettore colonna*, nel secondo di *vettore riga*. Per esempio,

$$\begin{bmatrix} 3 \\ 4 \end{bmatrix} \;,\; \begin{bmatrix} 5 & 6 \end{bmatrix}$$

sono due vettori – rispettivamente, colonna e riga – di ordine 2.

Seguendo una notazione assai diffusa, denoteremo i vettori colonna con lettere minuscole in grassetto ($\mathbf{x}, \mathbf{y}, \ldots$) e i vettori riga con lettere minuscole in grassetto e munite di apice ($\mathbf{z}', \mathbf{w}', \ldots$).

Rappresenteremo, inoltre, simbolicamente un generico vettore colonna $\mathbf{x}$ di ordine n con

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

---

AVVERTENZA. In tutto questo volume, a meno di esplicita indicazione contraria, m , n , p , q , r , t , v denotano numeri naturali maggiori o eguali a 1.

(1) Ometteremo, talvolta, di indicare l'ordine di un vettore.

(2) Alcuni, invece delle parentesi (quadre o tonde), usano le doppie sbarre verticali.

e un generico vettore riga $\mathbf{z}'$ di ordine n con

$$\left[ z_1 \cdots z_n \right]$$

dove a ciascun elemento $x_i$ e $z_i$ $(i = 1, \dots, n)$ è associato un indice che designa la sua posizione, rispettivamente, nella colonna o riga di appartenenza.

Talvolta, i vettori $\mathbf{x}$ e $\mathbf{z}'$ di cui sopra sono rappresentati, più compattamente, con $[x_i]_{(n,1)}$ e $[z_i]_{(1,n)}$.

Nel seguito di questo capitolo faremo esclusivo riferimento ai vettori colonna – denominati, senza ulteriori aggettivazioni, *vettori* – avvertendo tuttavia che quanto diremo a questo proposito vale anche per i vettori riga.

**1.2** Due vettori $\mathbf{x} = [x_i]_{(n,1)}$ e $\mathbf{y} = [y_i]_{(n,1)}$ sono *eguali*, e si scrive $\mathbf{x} = \mathbf{y}$, se e soltanto se $x_i = y_i$ per $i = 1, \dots, n$.

In altri termini, due vettori dello stesso ordine sono eguali se e soltanto se gli elementi corrispondenti sono eguali.

**1.3** Alcuni vettori ricorrono assai spesso nella trattazione che segue e conviene indicarli esplicitamente fin da questo momento.

Il vettore di ordine n le cui componenti sono tutte nulle è chiamato *vettore zero* o *vettore nullo* ed è indicato con il simbolo $\mathbf{0}$. Per esempio,

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

è il vettore nullo di ordine 2.

Il vettore di ordine n le cui componenti sono tutte eguali a 0 tranne la i-esima $(1 \le i \le n)$ posta eguale a 1 è chiamato *vettore canonico* ed è indicato con il simbolo $\mathbf{u}_i$. Per esempio,

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad , \quad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

sono i vettori canonici di ordine 2.

**OSSERVAZIONE 1.** Come è noto, fissato nel piano un sistema di assi cartesiani (non necessariamente ortogonali) e stabilito per ciascun asse un verso e una unità di misura (eventualmente la stessa per i due assi), si istituisce una corrispondenza biunivoca tra i vettori di ordine 2 e i punti del piano, e anche tra i vettori di ordine 2 e i segmenti orientati del piano aventi come punto di applicazione l'origine degli assi. Analogamente, fissato nello spazio un sistema di assi cartesiani (non necessariamente ortogonali) e stabilito per ciascun asse un verso e una unità di misura (eventualmente la stessa per i tre assi), si istituisce una corrispondenza biunivoca tra i vettori di ordine 3 e i punti dello spazio, e anche tra i vettori di ordine 3 e i segmenti orientati dello spazio aventi come punto di applicazione l'origine degli assi.

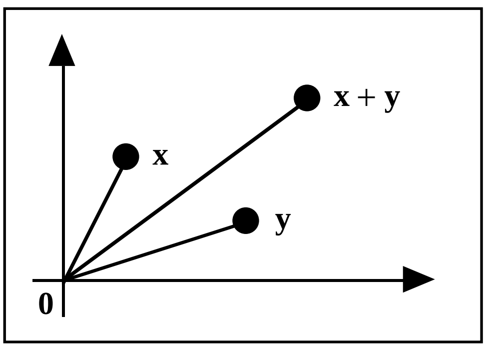
La possibilità di effettuare una rappresentazione grafica dei vettori di ordine 2 o 3, ora richiamata, sarà talvolta utilizzata nel corso di questo volume, senza ulteriori commenti, nell'intento di illustrare il significato geometrico dei concetti di volta in volta introdotti.

## 2 ADDIZIONE E MOLTIPLICAZIONE PER UN NUMERO REALE. SOTTRAZIONE

**2.1** Dati due vettori $\mathbf{x} = [x_i]_{(n,1)}$ e $\mathbf{y} = [y_i]_{(n,1)}$, si chiama *somma* di $\mathbf{x}$ e $\mathbf{y}$ il vettore, che indichiamo con $\mathbf{x} + \mathbf{y}$, i cui elementi sono espressi ordinatamente dai numeri $x_i + y_i$ per $i = 1, \dots, n$. Si ha, cioè,

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix}.$$

Fig. 1



**ESEMPIO 1.** Dati i vettori

$$\mathbf{x} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad , \quad \mathbf{y} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

si ha

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix} .$$ ▮

L'operazione in questione riceve la denominazione di *addizione*.

Come si può facilmente verificare, essa è *commutativa* e *associativa*; vale a dire, qualunque siano i vettori $\mathbf{x}, \mathbf{y}, \mathbf{z}$ di ordine n,

(i) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$

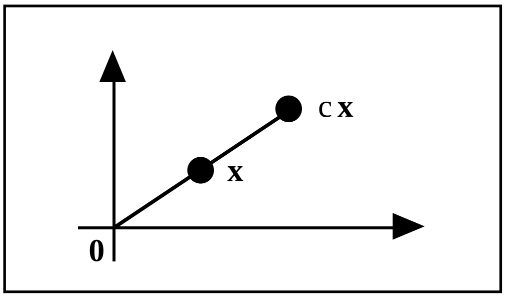
(ii) $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$.

Quest'ultima proprietà consente di indicare, senza possibilità di equivoci, la somma di $\mathbf{x}, \mathbf{y}, \mathbf{z}$ con $\mathbf{x} + \mathbf{y} + \mathbf{z}$, evitando l'uso delle parentesi, e di estendere l'addizione a un numero qualsiasi, purché finito, di vettori di ordine n.

**2.2** Dati un numero reale c e un vettore $\mathbf{x} = [x_i]_{(n,1)}$, si chiama *prodotto* di c per $\mathbf{x}$ o di $\mathbf{x}$ per c il vettore, che indichiamo con $c\mathbf{x}$ o $\mathbf{x}c$, i cui elementi sono espressi ordinatamente dai numeri $cx_i$ per $i = 1, \ldots, n$. Si ha, cioè,

$$c\mathbf{x} = \begin{bmatrix} cx_1 \\ \vdots \\ cx_n \end{bmatrix} = \mathbf{x}c .$$

Fig. 2



**ESEMPIO 2.** Dati il numero reale c = 2 e il vettore

$$\mathbf{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

risulta

$$c\mathbf{x} = 2\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} 2 = \mathbf{x}c\,. \qquad \blacksquare$$

L'operazione ora definita riceve la denominazione di *moltiplicazione per un numero reale*.

Come si può facilmente verificare, essa gode, qualunque siano i vettori **x** e **y** di ordine n e i numeri reali a e b, delle seguenti proprietà:

(i) $(ab)\mathbf{x} = a(b\mathbf{x})$

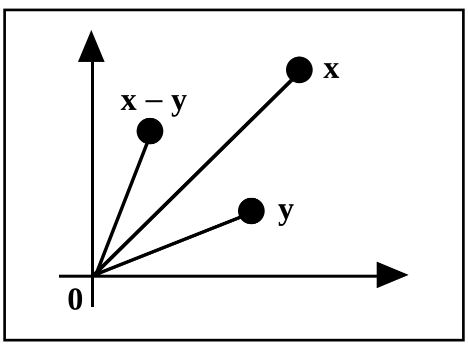
(ii) $a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y}$

(iii) $(a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}$.

**2.3** Dati due vettori $\mathbf{x} = [x_i]_{(n,1)}$ e $\mathbf{y} = [y_i]_{(n,1)}$, posto $(-\mathbf{y}) = (-1)\mathbf{y}$, si chiama *differenza* di **x** e **y** il vettore $\mathbf{x} + (-\mathbf{y}) = \mathbf{x} + (-1)\mathbf{y}$ che scriviamo, più semplicemente, $\mathbf{x} - \mathbf{y}$. Si ha, cioè,

$$\mathbf{x} - \mathbf{y} = \begin{bmatrix} x_1 - y_1 \\ \vdots \\ x_n - y_n \end{bmatrix}.$$

Fig. 3



**ESEMPIO 3.** Dati i vettori

$$\mathbf{x} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad , \quad \mathbf{y} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

risulta

$$\mathbf{x} - \mathbf{y} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} .$$

∎

L'operazione ora definita, derivata dalle operazioni fondamentali di addizione e di moltiplicazione per un numero reale, riceve la denominazione di *sottrazione* e consente di stabilire, qualunque siano i vettori $\mathbf{x}, \mathbf{y}, \mathbf{z}$ di ordine n e i numeri reali a e b, alcuni utili risultati di immediata verifica:

(i) $\mathbf{x} - \mathbf{x} = \mathbf{0}$ (iv) $\mathbf{x} - (\mathbf{y} - \mathbf{z}) = (\mathbf{x} - \mathbf{y}) + \mathbf{z}$

(ii) $-(-\mathbf{x}) = \mathbf{x}$ (v) $a(\mathbf{x} - \mathbf{y}) = a\mathbf{x} - a\mathbf{y}$

(iii) $-(\mathbf{x} - \mathbf{y}) = \mathbf{y} - \mathbf{x}$ (vi) $(a - b)\mathbf{x} = a\mathbf{x} - b\mathbf{x}$.

## 3 COMBINAZIONI LINEARI. DIPENDENZA E INDIPENDENZA LINEARE

**3.1** Dati p vettori $\mathbf{x}_1, \ldots, \mathbf{x}_p$ di ordine n e p numeri reali $a_1, \ldots, a_p$ (3), si chiama *combinazione lineare* di $\mathbf{x}_1, \ldots, \mathbf{x}_p$ con *coefficienti* $a_1, \ldots, a_p$ il vettore

$$\mathbf{x} = a_1\mathbf{x}_1 + \ldots + a_p\mathbf{x}_p .$$

È ovvio che ogni combinazione lineare di p vettori di ordine n è essa stessa un vettore di ordine n.

**ESEMPIO 4.** Dati i vettori

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad , \quad \mathbf{x}_2 = \begin{bmatrix} 3 \\ 3 \end{bmatrix} \quad , \quad \mathbf{x}_3 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

(3) Anche se non sempre appare necessario, conveniamo fin da ora, per ragioni che risulteranno chiare dal seguito di questa esposizione, di considerare insiemi di vettori del tipo $\mathbf{x}_1, \ldots, \mathbf{x}_p$ o di numeri reali del tipo $a_1, \ldots, a_p$ come insiemi *ordinati* sulla base della successione degli indici a essi apposti.

e i numeri reali $a_1 = 2, a_2 = 1, a_3 = 1$, il vettore

$$\mathbf{x} = \begin{bmatrix} 5 \\ 8 \end{bmatrix}$$

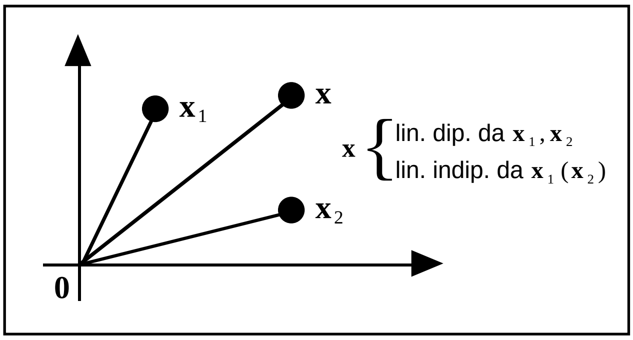
è combinazione lineare di $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ con coefficienti $a_1, a_2, a_3$. ▮

Dati un vettore $\mathbf{x}$ e p vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$ di ordine n, se $\mathbf{x}$ può essere espresso come combinazione lineare di $\mathbf{x}_1, \dots, \mathbf{x}_p$ – vale a dire se esistono p numeri reali $a_1, \dots, a_p$ tali che

$$\mathbf{x} = a_1\mathbf{x}_1 + \dots + a_p\mathbf{x}_p$$

– si dice che $\mathbf{x}$ è *linearmente dipendente* da $\mathbf{x}_1, \dots, \mathbf{x}_p$; altrimenti, si dice che $\mathbf{x}$ è *linearmente indipendente* da $\mathbf{x}_1, \dots, \mathbf{x}_p$.

Fig. 4




**ESEMPIO 5.** Dati i vettori

$$\mathbf{x} = \begin{bmatrix} 0 \\ 3 \end{bmatrix} \quad , \quad \mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad , \quad \mathbf{x}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$\mathbf{x}$ risulta linearmente dipendente da $\mathbf{x}_1, \mathbf{x}_2$ poiché esistono due numeri reali $a_1 = 2, a_2 = -1$ tali che $\mathbf{x} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2$.

**OSSERVAZIONE 2.** Qualunque siano i p vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$ di ordine n, il vettore $\mathbf{0}$ di ordine n risulta linearmente dipendente da $\mathbf{x}_1, \dots, \mathbf{x}_p$.

**OSSERVAZIONE 3.** Ogni vettore $\mathbf{x}$ di ordine n è linearmente dipendente dagli n vettori canonici $\mathbf{u}_1, \dots, \mathbf{u}_n$ di ordine n, potendosi scrivere

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 \mathbf{u}_1 + \dots + x_n \mathbf{u}_n \ .$$

In particolare, il vettore $\mathbf{u}$ di ordine n le cui componenti sono tutte eguali a 1 risulta linearmente dipendente dagli n vettori canonici $\mathbf{u}_1, \dots, \mathbf{u}_n$ di ordine n ($\mathbf{u} = \mathbf{u}_1 + \dots + \mathbf{u}_n$).

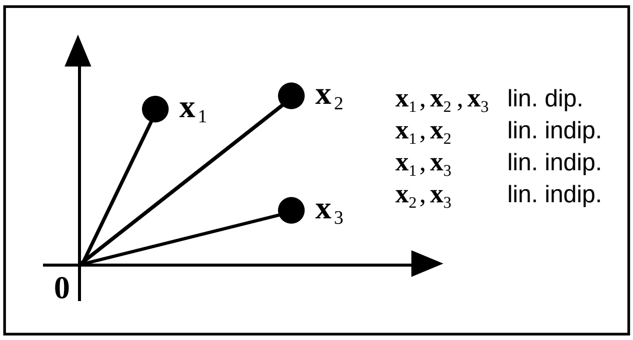
**OSSERVAZIONE 4.** Dati un vettore $\mathbf{x} \neq \mathbf{0}$ e p vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$ di ordine n, il problema di accertare se esistono p numeri reali $a_1, \dots, a_p$ tali che $\mathbf{x} = a_1 \mathbf{x}_1 + \dots + a_p \mathbf{x}_p$ sarà esaminato nel Cap. IV, nell'ambito della teoria dei sistemi di equazioni lineari.

**3.2** Dati p vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$ di ordine n, si dice che essi sono *linearmente dipendenti* oppure che formano un *insieme linearmente dipendente* se è possibile trovare p numeri reali $a_1, \dots, a_p$ *non tutti nulli* tali che

$$a_1 \mathbf{x}_1 + \dots + a_p \mathbf{x}_p = \mathbf{0} \ .$$

Invece, qualora la relazione precedente sia soddisfatta soltanto se $a_j = 0$ per $j = 1, \dots, p$, si dice che $\mathbf{x}_1, \dots, \mathbf{x}_p$ sono *linearmente indipendenti* oppure che formano un *insieme linearmente indipendente*.

Fig. 5



**ESEMPIO 6.** I vettori

$$\mathbf{x}_1 = \begin{bmatrix} 2 \\ 6 \end{bmatrix} \quad , \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

sono linearmente dipendenti poiché, posto $a_1 = 1$ e $a_2 = -2$, risulta che $a_1\mathbf{x}_1 + a_2\mathbf{x}_2 = \mathbf{0}$.

**OSSERVAZIONE 5.** Gli n vettori canonici $\mathbf{u}_1, \dots, \mathbf{u}_n$ di ordine n sono, come facilmente si verifica, linearmente indipendenti.

**OSSERVAZIONE 6.** Dati p vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$ di ordine n, se questi sono linearmente indipendenti, allora $\mathbf{x}_j \neq \mathbf{0}$ per $j = 1, \dots, p$.

**OSSERVAZIONE 7.** Dati p vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$ di ordine n, il problema di accertare se esistono p numeri reali $a_1, \dots, a_p$ non tutti nulli tali che $a_1\mathbf{x}_1 + \dots + a_p\mathbf{x}_p = \mathbf{0}$ sarà esaminato nel Cap. IV, nell'ambito della teoria dei sistemi di equazioni lineari. ▮

Si noti esplicitamente che le definizioni di dipendenza e indipendenza lineare qui sopra introdotte si applicano anche al caso in cui si consideri un unico vettore **x**.

In questa circostanza, dire che **x** è linearmente dipendente significa affermare che esiste un numero reale $a \neq 0$ tale che $a\mathbf{x} = \mathbf{0}$ e ciò si realizza se e soltanto se $\mathbf{x} = \mathbf{0}$.

Analogamente, dire che **x** è linearmente indipendente significa affermare che la relazione $a\mathbf{x} = \mathbf{0}$ si verifica se e soltanto se $a = 0$ e ciò equivale a dire che $\mathbf{x} \neq \mathbf{0}$.

**3.3** Nei teoremi che seguono stabiliremo alcuni semplici proprietà che discendono direttamente dalle definizioni di combinazione lineare e di dipendenza e indipendenza lineare, proprietà a cui faremo spesso riferimento nel prosieguo di questa esposizione.

**TEOREMA 1**

> Si considerino $p \geq 2$ vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$ di ordine n. Se uno di essi è combinazione lineare dei rimanenti, allora $\mathbf{x}_1, \dots, \mathbf{x}_p$ sono linearmente dipendenti.
>
> Viceversa, se $\mathbf{x}_1, \dots, \mathbf{x}_p$ sono linearmente dipendenti, allora almeno uno di essi è combinazione lineare dei rimanenti.

*Dim*. Posto che sia, per esempio,

$$\mathbf{x}_1 = a_2\mathbf{x}_2 + \dots + a_p\mathbf{x}_p \ ,$$

poiché l'espressione precedente può essere scritta nella forma

$$a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \dots + a_p\mathbf{x}_p = \mathbf{0}$$

dove almeno uno dei coefficienti è diverso da zero ($a_1 = -1$), la prima parte del teorema è dimostrata.

Per quanto riguarda la seconda parte del teorema, si osservi che, per ipotesi, esistono p numeri reali $a_1, a_2, \dots, a_p$ non tutti nulli tali che

$$a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \dots + a_p\mathbf{x}_p = \mathbf{0} \ .$$

Posto che sia, per esempio, $a_1 \neq 0$, ne consegue che

$$\mathbf{x}_1 = -\left(\frac{a_2}{a_1}\mathbf{x}_2 + \dots + \frac{a_p}{a_1}\mathbf{x}_p\right)$$

e cioè che almeno uno dei vettori, in questo caso $\mathbf{x}_1$, è linearmente dipendente dai rimanenti.

**TEOREMA 2**

> Si considerino p+1 vettori $\mathbf{x}_1, \dots, \mathbf{x}_p, \mathbf{x}_{p+1}$ di ordine n linearmente dipendenti. Se $\mathbf{x}_1, \dots, \mathbf{x}_p$ sono linearmente indipendenti, allora $\mathbf{x}_{p+1}$ si può esprimere come combinazione lineare di $\mathbf{x}_1, \dots, \mathbf{x}_p$.

*Dim*. Per ipotesi, esistono p+1 numeri reali $a_1, \dots, a_p, a_{p+1}$ non tutti nulli tali che

$$a_1\mathbf{x}_1 + \dots + a_p\mathbf{x}_p + a_{p+1}\mathbf{x}_{p+1} = \mathbf{0}\ .$$

Ora, $a_{p+1}$ non può essere nullo altrimenti la relazione precedente si potrebbe scrivere

$$a_1\mathbf{x}_1 + \dots + a_p\mathbf{x}_p = \mathbf{0}$$

con qualche $a_j$ $(1 \le j \le p)$ diverso da zero, contro l'ipotesi che $\mathbf{x}_1, \dots, \mathbf{x}_p$ siano linearmente indipendenti.

Quindi,

$$\mathbf{x}_{p+1} = -\left(\frac{a_1}{a_{p+1}}\mathbf{x}_1 + \dots + \frac{a_p}{a_{p+1}}\mathbf{x}_p\right)\ .$$

**OSSERVAZIONE 8.** Supposto che $\mathbf{x}_{p+1}$ si possa esprimere come combinazione lineare di $\mathbf{x}_1, \dots, \mathbf{x}_p$, allora $\mathbf{x}_1, \dots, \mathbf{x}_p, \mathbf{x}_{p+1}$ sono linearmente dipendenti (Cfr. il Teorema 1), ma non è vero, in generale, che $\mathbf{x}_1, \dots, \mathbf{x}_p$ siano linearmente indipendenti.

Per esempio, dati i vettori

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad , \quad \mathbf{x}_2 = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \quad , \quad \mathbf{x}_3 = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

$\mathbf{x}_3$ è combinazione lineare di $\mathbf{x}_1$ e $\mathbf{x}_2$ $(\mathbf{x}_3 = \mathbf{x}_1 + \mathbf{x}_2)$ e quindi $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ sono linearmente dipendenti, ma $\mathbf{x}_1$ e $\mathbf{x}_2$ non sono linearmente indipendenti, risultando $2\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}$.

**TEOREMA 3**

Sia X un insieme formato da p vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$ di ordine n linearmente indipendenti. Allora, ogni sottoinsieme non vuoto di X è costituito da vettori linearmente indipendenti.

*Dim*. Possiamo limitarci a considerare il caso in cui sia $p \geq 2$ e a supporre che il sottoinsieme Y di X sia costituito dai $p_1$ $(1 \leq p_1 < p)$ vettori $\mathbf{x}_1, \dots, \mathbf{x}_{p_1}$ di X.

Se Y fosse formato da vettori linearmente dipendenti, esisterebbero $p_1$ numeri reali $a_1, \dots, a_{p_1}$ non tutti nulli tali che

$$a_1 \mathbf{x}_1 + \dots + a_{p_1} \mathbf{x}_{p_1} = \mathbf{0}\ .$$

Ma, in tal caso, presi i numeri reali $a_{p_1+1}, \dots, a_p$ tutti eguali a zero, la relazione

$$a_1 \mathbf{x}_1 + \dots + a_{p_1} \mathbf{x}_{p_1} + a_{p_1+1} \mathbf{x}_{p_1+1} + \dots + a_p \mathbf{x}_p = \mathbf{0}$$

risulterebbe soddisfatta con qualche $a_j$ $(1 \leq j \leq p)$ diverso da zero, contraddicendo l'ipotesi che $\mathbf{x}_1, \dots, \mathbf{x}_p$ siano linearmente indipendenti.

**OSSERVAZIONE 9.** Sia X un insieme formato da $p \geq 2$ vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$ di ordine n. Supposto che ogni sottoinsieme proprio di X sia costituito da vettori linearmente indipendenti, non è vero, in generale, che X sia formato da vettori linearmente indipendenti.

Per esempio, sia X formato dai vettori

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \ , \ \ \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \ , \ \ \mathbf{x}_3 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Come facilmente si verifica, ogni sottoinsieme proprio di X è formato da vettori linearmente indipendenti.

Tuttavia, X è costituito da vettori linearmente dipendenti, risultando $-\mathbf{x}_1 - \mathbf{x}_2 + \mathbf{x}_3 = \mathbf{0}$.

**TEOREMA 4**

Sia X un insieme formato da p vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$ di ordine n linearmente dipendenti. Allora, ogni insieme finito di vettori di ordine n che contenga X è costituito da vettori linearmente dipendenti.

*Dim.* Consideriamo l'insieme Z contenente X, costituito da $q > p$ vettori $\mathbf{x}_1, \ldots, \mathbf{x}_p, \mathbf{x}_{p+1}, \ldots, \mathbf{x}_q$ di ordine n. Per ipotesi, sussiste la relazione

$$a_1 \mathbf{x}_1 + \ldots + a_p \mathbf{x}_p = \mathbf{0}$$

con $a_1, \ldots, a_p$ numeri reali non tutti nulli. Ne consegue che, presi i numeri reali $a_{p+1}, \ldots, a_q$ tutti eguali a zero, la relazione

$$a_1 \mathbf{x}_1 + \ldots + a_p \mathbf{x}_p + a_{p+1} \mathbf{x}_{p+1} + \ldots + a_q \mathbf{x}_q = \mathbf{0}$$

risulta soddisfatta con qualche $a_k$ $(1 \leq k \leq q)$ diverso da zero, e ciò prova il teorema.

## 4 SPAZI VETTORIALI

**4.1** Sia S un insieme non vuoto di vettori di ordine n e R l'insieme dei numeri reali. Si dice che S costituisce uno *spazio vettoriale* di *ordine* n (4) se, per tutti gli $\mathbf{x}, \mathbf{y} \in S$ e ogni $a \in R$, accade che $\mathbf{x} + \mathbf{y} \in S$ e $a\mathbf{x} \in S$.

In altri termini, un insieme non vuoto S di vettori di ordine n costituisce uno spazio vettoriale se è *chiuso* rispetto alle operazioni di addizione e di moltiplicazione per un numero reale.

Dalla definizione ora data risulta immediatamente che, se S è uno spazio vettoriale di ordine n, il vettore nullo di ordine n vi appartiene.

**ESEMPIO 7.** Un semplice esempio di spazio vettoriale di ordine n è offerto dall'insieme costituito dal solo vettore nullo di ordine n; lo chiameremo *spazio nullo* e lo indicheremo con il simbolo $S_0$.

**ESEMPIO 8.** L'insieme costituito da *tutti* i vettori di ordine n è uno spazio vettoriale di ordine n; esso riceve la denominazione di *spazio numerico* (*reale*) ed è indicato con il simbolo $R^n$.

**ESEMPIO 9.** L'insieme S costituito da tutti i vettori di ordine n aventi la prima

---

(4) Ometteremo, talvolta, di indicare l'ordine di uno spazio vettoriale.

componente eguale a zero è uno spazio vettoriale di ordine n. Infatti, la somma di due vettori appartenenti a S è ancora un vettore di S e altrettanto accade del prodotto di un numero reale per un vettore appartenente a S.

Al contrario, l'insieme costituito da tutti i vettori di ordine n aventi la prima componente eguale a un numero $c \neq 0$ *non* è, come si può facilmente verificare, uno spazio vettoriale.

**4.2** Dati p vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$ di ordine n, si consideri l'insieme S di tutte le combinazioni lineari di $\mathbf{x}_1, \dots, \mathbf{x}_p$, vale a dire l'insieme i cui elementi sono vettori del tipo

$$a_1 \mathbf{x}_1 + \dots + a_p \mathbf{x}_p$$

con $a_1, \dots, a_p$ variabili in R.

Considerate due combinazioni lineari

$$\mathbf{x} = b_1 \mathbf{x}_1 + \dots + b_p \mathbf{x}_p \quad , \quad \mathbf{y} = c_1 \mathbf{x}_1 + \dots + c_p \mathbf{x}_p$$

di $\mathbf{x}_1, \dots, \mathbf{x}_p$, la loro somma

$$\mathbf{x} + \mathbf{y} = (b_1 + c_1)\mathbf{x}_1 + \dots + (b_p + c_p)\mathbf{x}_p$$

è ancora una combinazione lineare di $\mathbf{x}_1, \dots, \mathbf{x}_p$ e, quindi, appartiene a S.

Considerati poi un numero reale c e una combinazione lineare

$$\mathbf{x} = d_1 \mathbf{x}_1 + \dots + d_p \mathbf{x}_p$$

di $\mathbf{x}_1, \dots, \mathbf{x}_p$, il prodotto

$$c\mathbf{x} = (cd_1)\mathbf{x}_1 + \dots + (cd_p)\mathbf{x}_p$$

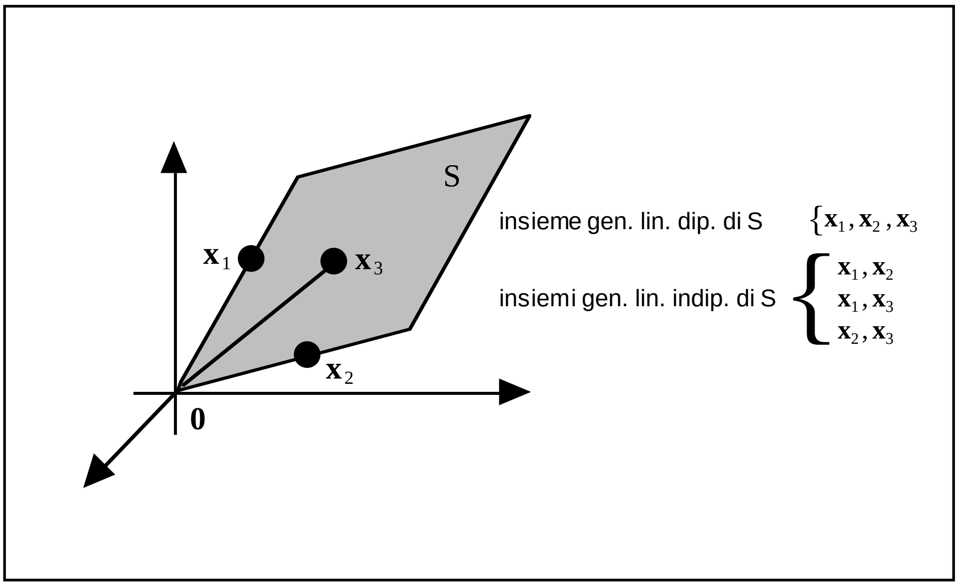
è ancora una combinazione lineare di $\mathbf{x}_1, \dots, \mathbf{x}_p$ e, come tale, appartiene a S.

Pertanto, l'insieme S, costituito da tutte le combinazioni lineari di $\mathbf{x}_1, \dots, \mathbf{x}_p$, è uno spazio vettoriale di ordine n.

Tale spazio si dice *generato* da $\mathbf{x}_1, \dots, \mathbf{x}_p$; si dice anche che $\mathbf{x}_1, \dots, \mathbf{x}_p$ formano un *insieme generatore* di S.

In particolare, qualora $\mathbf{x}_1, \ldots, \mathbf{x}_p$ siano linearmente indipendenti, si dice che essi formano un *insieme generatore linearmente indipendente* di S [5].

Fig. 6




**ESEMPIO 10.** Lo spazio vettoriale generato dal vettore

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

è costituito da tutti i vettori della forma $c\,\mathbf{x}_1$ con c variabile in R.

**OSSERVAZIONE 10.** Gli n vettori canonici $\mathbf{u}_1, \ldots, \mathbf{u}_n$ di ordine n formano un insieme generatore linearmente indipendente di $R^n$ poiché ogni vettore $\mathbf{x} \in R^n$ può essere espresso come combinazione lineare di $\mathbf{u}_1, \ldots, \mathbf{u}_n$ e questi sono linearmente indipendenti (Cfr. le Osservazioni 3 e 5).

**OSSERVAZIONE 11.** Supponiamo che $\mathbf{x}_1, \ldots, \mathbf{x}_p$ formino un insieme generatore di uno spazio vettoriale S e che $\mathbf{x}_1, \ldots, \mathbf{x}_p, \mathbf{x}_{p+1}$ formino un insieme gene-

---

(5) Nella figura che segue, lo spazio vettoriale S è raffigurato come una porzione di un piano P passante per l'origine degli assi: in realtà, S deve intendersi rappresentato da tutto il piano P. Una avvertenza simile vale ovviamente in situazioni analoghe.

ratore di uno spazio vettoriale Z.

Vogliamo mostrare che, se il vettore $\mathbf{x}_{p+1}$ è linearmente dipendente da $\mathbf{x}_1, \ldots, \mathbf{x}_p$, allora S = Z.

È ovvio che, potendosi esprimere ogni $\mathbf{x}$ appartenente a S nella forma

$$\mathbf{x} = a_1\mathbf{x}_1 + \ldots + a_p\mathbf{x}_p + 0\mathbf{x}_{p+1} ,$$

$\mathbf{x}$ appartiene anche a Z.

Reciprocamente, sia

$$\mathbf{z} = b_1\mathbf{x}_1 + \ldots + b_p\mathbf{x}_p + b_{p+1}\mathbf{x}_{p+1}$$

un generico vettore appartenente a Z.

Poiché, per ipotesi,

$$\mathbf{x}_{p+1} = c_1\mathbf{x}_1 + \ldots + c_p\mathbf{x}_p ,$$

possiamo esprimere $\mathbf{z}$ nella forma

$$\mathbf{z} = (b_1 + b_{p+1}c_1)\mathbf{x}_1 + \ldots + (b_p + b_{p+1}c_p)\mathbf{x}_p$$

e, quindi, $\mathbf{z}$ appartiene anche a S.

**ESEMPIO 11.** Dati i vettori

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad , \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad , \quad \mathbf{x}_3 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

lo spazio vettoriale generato da $\mathbf{x}_1, \mathbf{x}_2$ è identico allo spazio vettoriale generato da $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ $(\mathbf{x}_3 = \mathbf{x}_1 + \mathbf{x}_2)$.

**TEOREMA 5**

> Sia S lo spazio vettoriale generato da p vettori $\mathbf{x}_1, \ldots, \mathbf{x}_p$ di ordine n linearmente indipendenti. Comunque si scelgano p+1 vettori $\mathbf{z}_1, \ldots, \mathbf{z}_p, \mathbf{z}_{p+1}$ in S, questi sono linearmente dipendenti.

*Dim*. Si osservi anzitutto che se $\mathbf{z}_1, \ldots, \mathbf{z}_p$ fossero linearmente dipendenti lo sarebbero anche $\mathbf{z}_1, \dots, \mathbf{z}_p, \mathbf{z}_{p+1}$ (Cfr. il Teorema 4); non sarà, pertanto, restrittivo supporre che $\mathbf{z}_1, \dots, \mathbf{z}_p$ siano linearmente indipendenti.

Poiché $\mathbf{z}_1 \in S$, possiamo scrivere

$$\mathbf{z}_1 = b_1\mathbf{x}_1 + b_2\mathbf{x}_2 + \dots + b_p\mathbf{x}_p$$

dove – essendo $\mathbf{z}_1 \neq \mathbf{0}$ (Cfr. l'Osservazione 6) – i coefficienti $b_1, \dots, b_p$ non sono tutti nulli.

Salvo rinumerare $\mathbf{x}_1, \dots, \mathbf{x}_p$, possiamo supporre che sia $b_1 \neq 0$, e allora

$$\mathbf{x}_1 = \frac{1}{b_1}\mathbf{z}_1 - \frac{b_2}{b_1}\mathbf{x}_2 - \dots - \frac{b_p}{b_1}\mathbf{x}_p \ .$$

Ma, ogni $\mathbf{x} \in S$ si può esprimere come combinazione lineare di $\mathbf{x}_1, \dots, \mathbf{x}_p$ con coefficienti

$$a_1, \dots, a_p$$

e, quindi, risulta che

$$\begin{aligned}
\mathbf{x} &= a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \dots + a_p\mathbf{x}_p \\
&= a_1\left(\frac{1}{b_1}\mathbf{z}_1 - \frac{b_2}{b_1}\mathbf{x}_2 - \dots - \frac{b_p}{b_1}\mathbf{x}_p\right) + a_2\mathbf{x}_2 + \dots + a_p\mathbf{x}_p \\
&= \frac{a_1}{b_1}\mathbf{z}_1 + \left(a_2 - a_1\frac{b_2}{b_1}\right)\mathbf{x}_2 + \dots + \left(a_p - a_1\frac{b_p}{b_1}\right)\mathbf{x}_p \ .
\end{aligned}$$

Questo vuole dire che $\mathbf{z}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$ generano S, poiché ogni $\mathbf{x} \in S$ si può esprimere come combinazione lineare di $\mathbf{z}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$ con coefficienti

$$c_1 = \frac{a_1}{b_1} \ , \ c_2 = \left(a_2 - a_1\frac{b_2}{b_1}\right) \ , \ \dots \ , \ c_p = \left(a_p - a_1\frac{b_p}{b_1}\right) \ .$$

A sua volta, poiché $\mathbf{z}_2 \in S$, possiamo scrivere

$$\mathbf{z}_2 = d_1\mathbf{z}_1 + d_2\mathbf{x}_2 + \dots + d_p\mathbf{x}_p$$

dove – essendo $\mathbf{z}_1$ e $\mathbf{z}_2$ linearmente indipendenti (Cfr. il Teorema 3) – i coefficienti $d_2, \dots, d_p$ non sono tutti nulli.

Salvo rinumerare $\mathbf{z}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$, possiamo supporre che sia $d_2 \neq 0$, e allora

$$\mathbf{x}_2 = -\frac{d_1}{d_2}\mathbf{z}_1 + \frac{1}{d_2}\mathbf{z}_2 - \ldots - \frac{d_p}{d_2}\mathbf{x}_p \,.$$

Ma, ogni $\mathbf{x} \in S$ si può esprimere come combinazione lineare di $\mathbf{z}_1$, $\mathbf{x}_2, \ldots, \mathbf{x}_p$ con coefficienti

$$c_1, \ldots, c_p$$

e, quindi, risulta che

$$\begin{aligned}
\mathbf{x} &= c_1\mathbf{z}_1 + c_2\mathbf{x}_2 + \ldots + c_p\mathbf{x}_p \\
&= \frac{a_1}{b_1}\mathbf{z}_1 + (a_2 - a_1\frac{b_2}{b_1})\mathbf{x}_2 + \ldots + (a_p - a_1\frac{b_p}{b_1})\mathbf{x}_p \\
&= \frac{a_1}{b_1}\mathbf{z}_1 + (a_2 - a_1\frac{b_2}{b_1})(-\frac{d_1}{d_2}\mathbf{z}_1 + \frac{1}{d_2}\mathbf{z}_2 - \ldots - \frac{d_p}{d_2}\mathbf{x}_p) + \ldots + (a_p - a_1\frac{b_p}{b_1})\mathbf{x}_p \\
&= \left\{\frac{a_1}{b_1} - (a_2 - a_1\frac{b_2}{b_1})\frac{d_1}{d_2}\right\}\mathbf{z}_1 + \left\{(a_2 - a_1\frac{b_2}{b_1})\frac{1}{d_2}\right\}\mathbf{z}_2 - \ldots \\
&\quad - \left\{(a_2 - a_1\frac{b_2}{b_1})\frac{d_p}{d_2} - (a_p - a_1\frac{b_p}{b_1})\right\}\mathbf{x}_p \,.
\end{aligned}$$

Questo vuole dire che $\mathbf{z}_1, \mathbf{z}_2, \mathbf{x}_3, \ldots, \mathbf{x}_p$ generano S, poiché ogni $\mathbf{x} \in S$ si può esprimere come combinazione lineare di $\mathbf{z}_1, \mathbf{z}_2, \mathbf{x}_3, \ldots, \mathbf{x}_p$ con coefficienti

$$e_1 = \left\{\frac{a_1}{b_1} - (a_2 - a_1\frac{b_2}{b_1})\frac{d_1}{d_2}\right\} \;,\; e_2 = \left\{(a_2 - a_1\frac{b_2}{b_1})\frac{1}{d_2}\right\} \;,\; \ldots \;,$$
$$e_p = \left\{(a_2 - a_1\frac{b_2}{b_1})\frac{d_p}{d_2} - (a_p - a_1\frac{b_p}{b_1})\right\} \,.$$

Proseguendo nel ragionamento, si conclude che $\mathbf{z}_1, \ldots, \mathbf{z}_p$ generano S.

Ma, poiché $\mathbf{z}_{p+1} \in S$, $\mathbf{z}_{p+1}$ può essere espresso come combinazione lineare di $\mathbf{z}_1, \ldots, \mathbf{z}_p$. Ne consegue, tenuto conto del Teorema 1, che $\mathbf{z}_1, \ldots, \mathbf{z}_p, \mathbf{z}_{p+1}$ sono linearmente dipendenti.

**OSSERVAZIONE 12.** Dati n+1 vettori $\mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{x}_{n+1}$ di ordine n, poiché ciascuno di essi può essere espresso come combinazione lineare degli n vettori canonici $\mathbf{u}_1, \ldots, \mathbf{u}_n$ di ordine n (Cfr. l'Osservazione 3), $\mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{x}_{n+1}$ sono linearmente dipendenti.

## 5 BASI E DIMENSIONE DI UNO SPAZIO VETTORIALE

**5.1** Sia S uno spazio vettoriale non nullo di ordine n[(6)]. Considerati p vettori $\mathbf{x}_1, \dots, \mathbf{x}_p \in S$, si dice che $\mathbf{x}_1, \dots, \mathbf{x}_p$ formano una *base* (o *riferimento*) di S se ogni vettore $\mathbf{x} \in S$ è esprimibile *in modo univoco* come combinazione lineare di $\mathbf{x}_1, \dots, \mathbf{x}_p$.

In tal caso, i coefficienti $a_1, \dots, a_p$ dell'espressione $\mathbf{x} = a_1\mathbf{x}_1 + \dots + a_p\mathbf{x}_p$ sono anche detti le *coordinate* di $\mathbf{x}$ rispetto alla suddetta base di S.

Nel teorema seguente ci proponiamo di chiarire la relazione esistente tra il concetto di base e quello di insieme generatore linearmente indipendente di S.

**TEOREMA 6**

> Sia S uno spazio vettoriale non nullo di ordine n. Una base di S è un insieme generatore linearmente indipendente di S.
>
> Viceversa, un insieme generatore linearmente indipendente di S è una base di S.

*Dim*. Supposto che $\mathbf{x}_1, \dots, \mathbf{x}_p$ costituiscano una base di S, è intanto ovvio che $\mathbf{x}_1, \dots, \mathbf{x}_p$ formano un insieme generatore di S. Inoltre, poichè

$$0\mathbf{x}_1 + \dots + 0\mathbf{x}_p = \mathbf{0}\ ,$$

e, per definizione di base, $\mathbf{0}$ è esprimibile in modo univoco come combinazione lineare di $\mathbf{x}_1, \dots, \mathbf{x}_p$, tali vettori risultano linearmente indipendenti.

Viceversa, supposto che $\mathbf{x}_1, \dots, \mathbf{x}_p$ formino un insieme generatore linearmente indipendente di S, per ogni $\mathbf{x} \in S$ esistono p numeri reali $a_1, \dots, a_p$ tali che

$$\mathbf{x} = a_1\mathbf{x}_1 + \dots + a_p\mathbf{x}_p\ .$$

Nell'ipotesi poi che esistano altri p numeri reali $b_1, \dots, b_p$ tali che

(6) Per quanto riguarda lo spazio nullo, si veda la nota (9).

$$\mathbf{x} = b_1\mathbf{x}_1 + \ldots + b_p\mathbf{x}_p \ ,$$

si ha che

$$(a_1 - b_1)\mathbf{x}_1 + \ldots + (a_p - b_p)\mathbf{x}_p = \mathbf{0} \ .$$

Ma, essendo $\mathbf{x}_1, \ldots, \mathbf{x}_p$ linearmente indipendenti, risulta che $a_1 = b_1, \ldots, a_p = b_p$, ovvero che $\mathbf{x}$ è esprimibile in modo univoco come combinazione lineare di $\mathbf{x}_1, \ldots, \mathbf{x}_p$.

**OSSERVAZIONE 13.** Gli n vettori canonici $\mathbf{u}_1, \ldots, \mathbf{u}_n$ di ordine n formano una base di $R^n$ in quanto costituiscono un insieme generatore linearmente indipendente di $R^n$ (Cfr. l'Osservazione 10). Tale base è detta base *canonica* o *naturale* di $R^n$.

Si noti che, tenuto conto dell'Osservazione 3, le componenti di un generico vettore $\mathbf{x} \in R^n$ possono essere interpretate come le coordinate di $\mathbf{x}$ rispetto alla suddetta base di $R^n$.

**ESEMPIO 12.** Si consideri lo spazio vettoriale S generato dai vettori

$$\mathbf{x}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \ , \ \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} .$$

Tali vettori sono, come facilmente si verifica, linearmente indipendenti e, pertanto, costituiscono una base di S.

Al contrario, i vettori

$$\mathbf{x}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \ , \ \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \ , \ \mathbf{x}_3 = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

*non* costituiscono una base di S.

Infatti, pur formando $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ un insieme generatore di S, essi sono linearmente dipendenti ($\mathbf{x}_1 + \mathbf{x}_2 - \mathbf{x}_3 = \mathbf{0}$) e, quindi, non consentono di esprimere in modo univoco ogni elemento $\mathbf{x} \in S$ come combinazione lineare di $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$.

Per esempio, il vettore

$$\mathbf{x} = \begin{bmatrix} 8 \\ 6 \end{bmatrix}$$

può essere espresso come combinazione lineare di $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ secondo i coefficienti $a_1 = 1, a_2 = 1, a_3 = 1$ oppure per mezzo dei coefficienti $a_1 = 0$, $a_2 = 0, a_3 = 2$.

**OSSERVAZIONE 14.** Se uno spazio vettoriale non nullo S di ordine n possiede una base costituita da p vettori, allora ogni altra base di S è formata da p vettori.

Supponiamo, infatti, che $\mathbf{x}_1, \dots, \mathbf{x}_p$ e $\widetilde{\mathbf{x}}_1, \dots, \widetilde{\mathbf{x}}_q$ siano due basi di S.

Se fosse $q > p$, poiché ciascun $\widetilde{\mathbf{x}}_k$ $(k = 1, \dots, q)$ può essere espresso come combinazione lineare di $\mathbf{x}_1, \dots, \mathbf{x}_p$, tenuto conto del Teorema 5, $\widetilde{\mathbf{x}}_1, \dots, \widetilde{\mathbf{x}}_q$ sarebbero linearmente dipendenti, in contrasto con l'ipotesi che questi formino una base di S.

Non potendo essere $q > p$, deve essere $q \leq p$. Ma, con un ragionamento analogo a quello ora svolto, si perviene alla conclusione che non può essere $q < p$; dunque, $q = p$.

**5.2** Dato uno spazio vettoriale non nullo S di ordine n, supponiamo che S contenga p vettori linearmente indipendenti e che, comunque si scelgano p+1 vettori in S, questi risultino linearmente dipendenti; si dice allora che S ha *dimensione* p e si scrive $\dim(S) = p$ (7) (8).

Più brevemente, possiamo dire che la dimensione di S è il massimo numero di vettori linearmente indipendenti esistenti in S.

**OSSERVAZIONE 15.** Si noti che uno spazio vettoriale non nullo S di ordine n ha dimensione compresa tra 1 e n. In effetti, S contiene almeno un vettore $\mathbf{x} \neq \mathbf{0}$ e, quindi, $\dim(S) \geq 1$; inoltre, comunque si scelgano n+1 vettori in S, questi risultano linearmente dipendenti (Cfr. l'Osservazione 12) e, pertanto, $\dim(S) \leq n$. ▮

---

(7) Ometteremo, talvolta, di indicare la dimensione di uno spazio vettoriale.

(8) Per quanto riguarda lo spazio nullo, si veda la nota (9).

Nel teorema seguente ci proponiamo di stabilire il legame esistente tra il concetto di base e quello di dimensione.

**TEOREMA 7**

> Sia S uno spazio vettoriale non nullo di ordine n. Se dim(S) = p, allora S ha una base formata da p vettori.
>
> Viceversa, se S ha una base costituita da p vettori, allora dim(S) = p.

*Dim*. Se dim(S) = p, S contiene p vettori $\mathbf{x}_1, \ldots, \mathbf{x}_p$ linearmente indipendenti; inoltre, qualunque sia $\mathbf{x}$ appartenente a S, i p+1 vettori $\mathbf{x}_1, \ldots, \mathbf{x}_p, \mathbf{x}$ sono linearmente dipendenti. Ne consegue (Cfr. il Teorema 2) che $\mathbf{x}$ si può esprimere come combinazione lineare di $\mathbf{x}_1, \ldots, \mathbf{x}_p$ e, quindi, che $\mathbf{x}_1, \ldots, \mathbf{x}_p$ costituiscono un insieme generatore linearmente indipendente di S, ovvero una base di S.

Viceversa, supponiamo che $\mathbf{x}_1, \ldots, \mathbf{x}_p$ formino una base di S. Comunque si scelgano p+1 vettori in S, questi risultano linearmente dipendenti e ciò prova che dim(S) = p.

**COROLLARIO.** Poiché uno spazio vettoriale non nullo S di ordine n ha comunque dimensione finita compresa tra 1 e n (Cfr. l'Osservazione 15), S possiede *sempre* una base [9].

**5.3** Vogliamo adesso mostrare la possibilità di costruire una base di uno spazio vettoriale di ordine n e dimensione p>1 partendo da un insieme di vettori linearmente indipendenti di numerosità minore di p (*completamento di una base*).

---

(9) Si assume come base dello spazio nullo di ordine n l'insieme costituito dal vettore zero di ordine n (benché questo non sia linearmente indipendente). Per definizione, si pone poi la dimensione dello spazio nullo pari a zero.

**TEOREMA 8**

Sia S uno spazio vettoriale di ordine n e dimensione $p>1$. Posto che $\mathbf{x}_1, \dots, \mathbf{x}_{p1}$ siano $p_1$ $(1 \le p_1 < p)$ vettori linearmente indipendenti appartenenti a S, è possibile trovare $p-p_1$ vettori $\mathbf{x}_{p1+1}, \dots, \mathbf{x}_p$ in S in modo tale che $\mathbf{x}_1, \dots, \mathbf{x}_{p1}, \mathbf{x}_{p1+1}, \dots, \mathbf{x}_p$ formino una base di S.

*Dim.* Indichiamo con $S_1$ lo spazio vettoriale generato da $\mathbf{x}_1, \dots, \mathbf{x}_{p1}$.

Poiché $\mathbf{x}_1, \dots, \mathbf{x}_{p1}$ non formano una base di S, esisterà qualche elemento $\mathbf{x}_{p1+1} \in S$ tale che $\mathbf{x}_{p1+1} \notin S_1$, cioè non esprimibile come combinazione lineare di $\mathbf{x}_1, \dots, \mathbf{x}_{p1}$.

Ne consegue che $\mathbf{x}_1, \dots, \mathbf{x}_{p1}, \mathbf{x}_{p1+1}$ costituiscono un insieme di vettori linearmente indipendenti.

Ripetendo il procedimento quante volte occorre, otteniamo in definitiva un insieme di p vettori linearmente indipendenti che formano una base di S.

**COROLLARIO.** Ogni insieme di vettori linearmente indipendenti appartenenti a uno spazio vettoriale S di ordine n e dimensione p è *parte* di una base di S.

**TEOREMA 9**

Il completamento di una base di uno spazio vettoriale S di ordine n e dimensione $p>1$ può avvenire in infiniti modi.

*Dim.* Si considerino nuovamente i $p_1+1$ vettori linearmente indipendenti $\mathbf{x}_1, \dots, \mathbf{x}_{p1}, \mathbf{x}_{p1+1}$ di cui al teorema precedente e il vettore

$$\mathbf{z} = b_1 \mathbf{x}_1 + \dots + b_{p1} \mathbf{x}_{p1} + b_{p1+1} \mathbf{x}_{p1+1}$$

con $b_1, \dots, b_{p1}, b_{p1+1}$ numeri reali arbitrari e $b_{p1+1} \neq 0$.

Vogliamo anzitutto dimostrare che anche $\mathbf{x}_1, \dots, \mathbf{x}_{p1}, \mathbf{z}$ sono linearmente indipendenti, vale a dire che la relazione

$$a_1\mathbf{x}_1+\ldots+a_{p1}\mathbf{x}_{p1}+a_{p1+1}\mathbf{z}=\mathbf{0}$$

implica che sia $a_1=\ldots=a_{p1}=a_{p1+1}=0$.

Infatti, la relazione precedente può essere scritta nella forma

$$a_1\mathbf{x}_1+\ldots+a_{p1}\mathbf{x}_{p1}+a_{p1+1}(b_1\mathbf{x}_1+\ldots+b_{p1}\mathbf{x}_{p1}+b_{p1+1}\mathbf{x}_{p1+1})=\mathbf{0}$$

ovvero nell'altra

$$(a_1+a_{p1+1}b_1)\mathbf{x}_1+\ldots+(a_{p1}+a_{p1+1}b_{p1})\mathbf{x}_{p1}+a_{p1+1}b_{p1+1}\mathbf{x}_{p1+1}=\mathbf{0}\,.$$

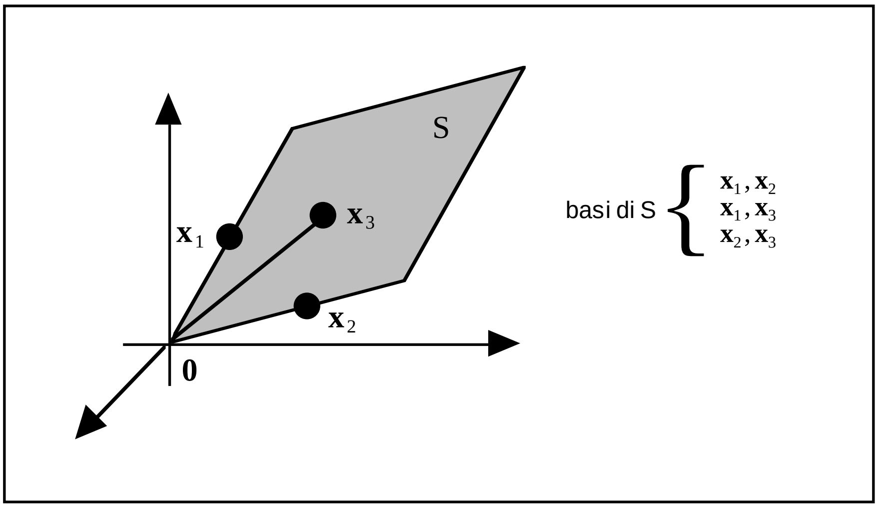
Ma, avendo supposto che $\mathbf{x}_1,\ldots,\mathbf{x}_{p1},\mathbf{x}_{p1+1}$ siano linearmente indipendenti e $b_{p1+1}\neq 0$, risulta $a_{p1+1}=0$ e anche $a_1=\ldots=a_{p1}=0$; quindi, $\mathbf{x}_1,\ldots,\mathbf{x}_{p1},\mathbf{z}$ sono linearmente indipendenti.

Dato che $\mathbf{z}$ è arbitrario, da quanto precede segue subito l'affermazione del teorema.

**OSSERVAZIONE 16.** Ogni spazio vettoriale S di ordine n e dimensione p ammette *infinite* basi.

Se $p>1$, ciò è un'immediata conseguenza del teorema precedente; se $p=1$, ogni vettore $\mathbf{x}\neq\mathbf{0}$ appartenente a S forma chiaramente una base di S.

Fig. 7



## 6 SOTTOSPAZI VETTORIALI

**6.1** Sia S uno spazio vettoriale e supponiamo, per concretezza, che S non sia lo spazio nullo.

Considerato un sottoinsieme non vuoto $S_1$ di vettori appartenenti a S, si dice che $S_1$ costituisce un *sottospazio vettoriale* di S – o, più semplicemente, un *sottospazio* di S – se, per tutti gli $\mathbf{x}, \mathbf{y} \in S_1$ e ogni $a \in R$, accade che $\mathbf{x} + \mathbf{y} \in S_1$ e $a\mathbf{x} \in S_1$.

Tale denominazione è giustificata dal fatto che *ogni sottospazio di* S *è esso stesso uno spazio vettoriale*, come risulta immediatamente dalla definizione data.

Ovviamente, sia S sia $S_0$ (spazio nullo di ordine eguale a quello di S) sono sottospazi di S (sottospazi *impropri*).

**ESEMPIO 13.** Dati p vettori $\mathbf{x}_1, \ldots, \mathbf{x}_p$ appartenenti a $R^n$, lo spazio vettoriale da essi generato è, chiaramente, un sottospazio di $R^n$.

**ESEMPIO 14.** Lo spazio vettoriale generato dai vettori

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad , \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

è un sottospazio di $R^2$ (coincidente con $R^2$).

**ESEMPIO 15.** Lo spazio vettoriale $S_1$ generato dai vettori

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad , \quad \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

è un sottospazio di $R^3$.

Lo spazio vettoriale $S_2$ generato dal vettore $\mathbf{x}_1$ è un sottospazio di $S_1$ e, quindi, è anche un sottospazio di $R^3$.

Lo spazio vettoriale $S_3$ generato dal vettore $\mathbf{x}_2$ è un sottospazio di $S_1$ e, quindi, è anche un sottospazio di $R^3$.

**TEOREMA 10**

> Dato un sottospazio $S_1$ di uno spazio vettoriale non nullo S, è $\dim(S_1) \leq \dim(S)$.

*Dim.* Qualora $S_1$ sia un sottospazio improprio di S, il teorema risulta banalmente vero. Infatti, posto $\dim(S) = p$, se $S_1 = S_0$, allora $0 = \dim(S_1) < \dim(S) = p$; se $S_1 = S$, allora $\dim(S_1) = p = \dim(S)$.

Nel caso poi in cui $S_1$ sia un sottospazio proprio di S, $S_1$ possiede una base formata da $p_1 = \dim(S_1)$ vettori i quali appartengono anche a S ma, chiaramente, non costituiscono una base di S. Quindi, deve risultare $p_1 = \dim(S_1) < \dim(S) = p$.

**ESEMPIO 16.** Riprendendo in esame gli spazi vettoriali definiti nell'Esempio 15, si verifica facilmente che: (i) $\dim(S_1) = 2 < \dim(R^3) = 3$; (ii) $\dim(S_2) = 1 < \dim(S_1) = 2$; (iii) $\dim(S_3) = 1 < \dim(S_1) = 2$.

**6.2** Siano $S_1$ e $S_2$ sottospazi di uno spazio vettoriale non nullo S. Si chiama *intersezione* di $S_1$ e $S_2$ l'insieme, denotato con il simbolo $S_1 \cap S_2$, costituito da tutti quei vettori che appartengono sia a $S_1$ sia a $S_2$.

**TEOREMA 11**

> Dati due sottospazi $S_1$ e $S_2$ di uno spazio vettoriale non nullo S, anche $S_1 \cap S_2$ è un sottospazio di S.

*Dim.* $S_1 \cap S_2$ contiene il vettore nullo dato che questo appartiene a $S_1$ e a $S_2$; pertanto, in ogni caso, $S_1 \cap S_2$ non è vuoto.

Considerati due vettori $\mathbf{z}_1, \mathbf{z}_2 \in S_1 \cap S_2$, ciascuno di essi, per definizione di intersezione, appartiene a $S_1$ e a $S_2$. Ne segue che, essendo $S_1$ e $S_2$ spazi vettoriali, $\mathbf{z}_1 + \mathbf{z}_2$ appartiene a $S_1$ e a $S_2$; quindi, $\mathbf{z}_1 + \mathbf{z}_2 \in S_1 \cap S_2$.

Analogamente, per tutti gli $\mathbf{z} \in S_1 \cap S_2$ e ogni $a \in R$, risulta che $a\mathbf{z} \in S_1 \cap S_2$ e questo completa la dimostrazione del teorema.

**ESEMPIO 17.** Sia $S_1$ il sottospazio di $R^3$ generato dai vettori

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad , \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} .$$

Analogamente, sia $S_2$ il sottospazio di $R^3$ generato dai vettori

$$\mathbf{y}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad , \quad \mathbf{y}_2 = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} .$$

Indicando con $\mathbf{z}$ un generico vettore appartenente a $S_1 \cap S_2$, poiché deve essere $\mathbf{z} \in S_1$ e $\mathbf{z} \in S_2$, possiamo scrivere $(a_1, a_2, b_1, b_2 \in R)$

$$\mathbf{z} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 \quad , \quad \mathbf{z} = b_1\mathbf{y}_1 + b_2\mathbf{y}_2$$

da cui

$$a_1\mathbf{x}_1 + a_2\mathbf{x}_2 - b_1\mathbf{y}_1 - b_2\mathbf{y}_2 = \mathbf{0} .$$

Ora, la relazione precedente è soddisfatta ponendo $a_1 = 2c, a_2 = 0$, $b_1 = -2c, b_2 = c$ ($c \in R$, arbitrario); pertanto,

$$\mathbf{z} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 = b_1\mathbf{y}_1 + b_2\mathbf{y}_2 = c\begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} ,$$

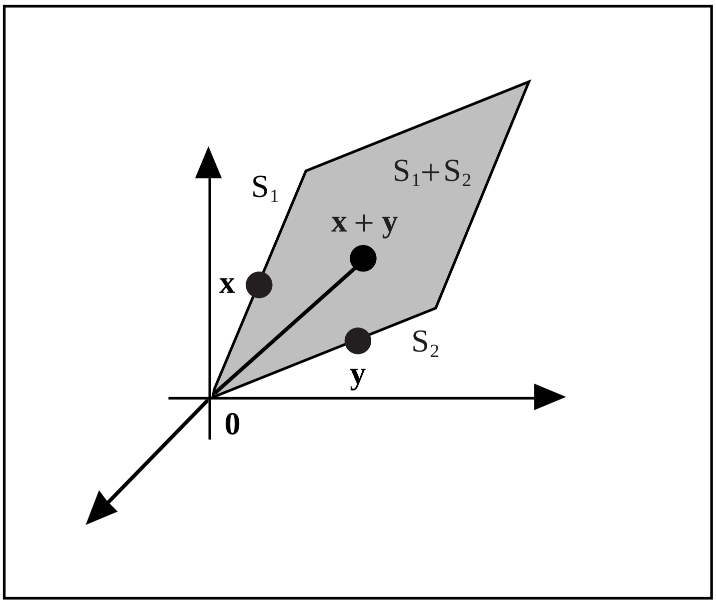
ovvero $S_1 \cap S_2$ è costituito da tutti i vettori della forma

$$c\begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$$

con c variabile in R.

**6.3** Siano $S_1$ e $S_2$ sottospazi di uno spazio vettoriale non nullo S. Si chiama *somma* di $S_1$ e $S_2$ l'insieme, denotato con $S_1+S_2$, costituito da tutti quei vettori della forma $\mathbf{x} + \mathbf{y}$ con $\mathbf{x} \in S_1$ e $\mathbf{y} \in S_2$.

Fig. 8




**TEOREMA 12**

Dati due sottospazi $S_1$ e $S_2$ di uno spazio vettoriale non nullo S, anche $S_1+S_2$ è un sottospazio di S.

*Dim*. Chiaramente, $S_1+S_2$ contiene il vettore nullo e quindi, in ogni caso, $S_1+S_2$ non è vuoto.

Supposto poi che $\mathbf{z}_1 \in S_1+S_2$ e $\mathbf{z}_2 \in S_1+S_2$, esisteranno vettori $\mathbf{x}_1, \mathbf{x}_2 \in S_1$ e $\mathbf{y}_1, \mathbf{y}_2 \in S_2$ tali che $\mathbf{z}_1 = \mathbf{x}_1 + \mathbf{y}_1$ e $\mathbf{z}_2 = \mathbf{x}_2 + \mathbf{y}_2$. Ne consegue che

$$\mathbf{z}_1 + \mathbf{z}_2 = (\mathbf{x}_1 + \mathbf{y}_1) + (\mathbf{x}_2 + \mathbf{y}_2) = (\mathbf{x}_1 + \mathbf{x}_2) + (\mathbf{y}_1 + \mathbf{y}_2) \in S_1+S_2 \,.$$

Analogamente, per ogni $a \in R$, risulta che

$$a\mathbf{z}_1 = a(\mathbf{x}_1 + \mathbf{y}_1) \in S_1+S_2 \,.$$

Pertanto, $S_1+S_2$ è un sottospazio di S.

**ESEMPIO 18.** Siano $S_1$ e $S_2$, sottospazi di $R^3$, definiti come nell'Esempio 17. Indicando con **x** e **y** due generici vettori appartenenti, rispettivamente, a $S_1$ e a $S_2$, risulta $(c_1, c_2, d_1, d_2 \in R)$

$$\mathbf{x} + \mathbf{y} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + d_1\mathbf{y}_1 + d_2\mathbf{y}_2 \ .$$

Chiaramente, al variare di $c_1, c_2, d_1, d_2$ in R, $c_1\mathbf{x}_1 + c_2\mathbf{x}_2$ descrive $S_1$ e, a sua volta, $d_1\mathbf{y}_1 + d_2\mathbf{y}_2$ descrive $S_2$.

Al tempo stesso, poiché $\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}_1, \mathbf{y}_2$ formano, come si può facilmente verificare, un insieme generatore di $R^3$, $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + d_1\mathbf{y}_1 + d_2\mathbf{y}_2$ descrive $R^3$, cosicché, in questo caso, $S_1 + S_2 = R^3$.

**6.4** Siano $S_1$ e $S_2$ sottospazi di uno spazio vettoriale non nullo S. Se $S_1 \cap S_2 = S_0$, $S_1 + S_2$ è detta *somma diretta* e la si scrive $S_1 \oplus S_2$.

Ovviamente, $S_1 \oplus S_2$ è un sottospazio di S.

**ESEMPIO 19.** Sia $S_1$ il sottospazio di $R^2$ generato dal vettore

$$\mathbf{x}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} .$$

Analogamente, sia $S_2$ il sottospazio di $R^2$ generato dal vettore

$$\mathbf{y}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} .$$

Si verifica facilmente che $S_1 + S_2 = R^2$. Indicando poi con **z** un generico vettore appartenente a $S_1 \cap S_2$, poiché deve essere $\mathbf{z} \in S_1$ e $\mathbf{z} \in S_2$, possiamo scrivere $(a, b \in R)$

$$\mathbf{z} = a\mathbf{x}_1 \quad , \quad \mathbf{z} = b\mathbf{y}_1$$

da cui

$$a\mathbf{x}_1 - b\mathbf{y}_1 = \mathbf{0} \ .$$

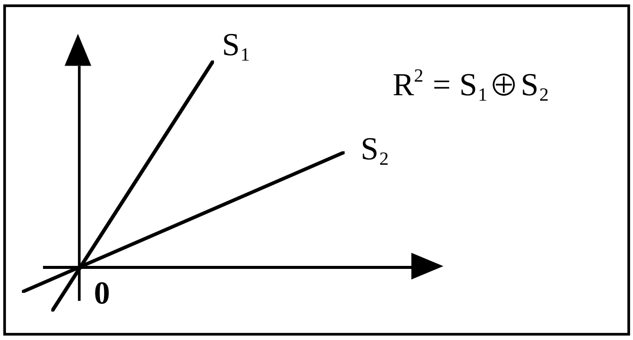
Ma, essendo $\mathbf{x}_1$ e $\mathbf{y}_1$ linearmente indipendenti, la relazione che precede è soddisfatta soltanto qualora sia $a = b = 0$. Pertanto, $\mathbf{z} = \mathbf{0}$ e $S_1 \cap S_2 = S_0$ e, dunque, $R^2 = S_1 \oplus S_2$.

**ESEMPIO 20.** Riprendiamo in esame i sottospazi $S_1$ e $S_2$ di $R^3$ considerati negli Esempi 17 e 18.

Abbiamo già dimostrato che $S_1 + S_2 = R^3$ e che $S_1 \cap S_2$ è diverso da $S_0$; quindi, $R^3$ *non* è somma diretta di $S_1$ e $S_2$.

**6.5** Dati $S_1$ e $S_2$, sottospazi di uno spazio vettoriale non nullo S, nel caso che S sia somma diretta di $S_1$ e $S_2$, si dice che $S_1$ e $S_2$ sono *sottospazi complementari* o *supplementari* (l'uno rispetto all'altro) in S [10].

Fig. 9




Il teorema che segue fornisce una condizione necessaria e sufficiente affinché $S_1$ e $S_2$ siano sottospazi complementari in S.

**TEOREMA 13**

Dati due sottospazi $S_1$ e $S_2$ di uno spazio vettoriale non nullo S, condizione necessaria e sufficiente affinché sia $S = S_1 \oplus S_2$ è che ciascun vettore appartenente a S si possa esprimere in modo univoco come somma di un vettore appartenente a $S_1$ e di un vettore appartenente a $S_2$.

(10) Si osservi che la nozione di sottospazio complementare è concettualmente distinta da quella di insieme complementare a cui si fa riferimento nella teoria degli insiemi.

*Dim*. Si supponga che S sia somma diretta di $S_1$ e $S_2$. Allora, per definizione, $S = S_1 + S_2$, cioè ciascun vettore $\mathbf{x} \in S$ è somma di un vettore $\mathbf{y} \in S_1$ e di un vettore $\mathbf{z} \in S_2$.

Supposto che $\mathbf{x}$ sia esprimibile come somma di altri due vettori $\widetilde{\mathbf{y}}$ e $\widetilde{\mathbf{z}}$ appartenenti, rispettivamente, a $S_1$ e a $S_2$, avremo che $\mathbf{y} + \mathbf{z} = \widetilde{\mathbf{y}} + \widetilde{\mathbf{z}}$.

Quest'ultima relazione può essere scritta nella forma $\mathbf{y} - \widetilde{\mathbf{y}} = \widetilde{\mathbf{z}} - \mathbf{z}$ dove, essendo $S_1$ e $S_2$ sottospazi, $\mathbf{y} - \widetilde{\mathbf{y}} \in S_1$ e $\widetilde{\mathbf{z}} - \mathbf{z} \in S_2$. Ma, stante la definizione di somma diretta, deve essere anche $S_1 \cap S_2 = S_0$ e, perciò, $\mathbf{y} - \widetilde{\mathbf{y}} = \mathbf{0}$ e $\widetilde{\mathbf{z}} - \mathbf{z} = \mathbf{0}$. Ne consegue che $\mathbf{y} = \widetilde{\mathbf{y}}$ e $\mathbf{z} = \widetilde{\mathbf{z}}$, ovvero che $\mathbf{x}$ può essere espresso univocamente come somma di $\mathbf{y}$ e $\mathbf{z}$.

Si supponga, adesso, che ogni $\mathbf{x} \in S$ possa essere espresso univocamente come somma di un vettore $\mathbf{y} \in S_1$ e di un vettore $\mathbf{z} \in S_2$.

Intanto, è $S = S_1 + S_2$. Supposto poi che $\mathbf{x} \in S_1 \cap S_2$, $\mathbf{x}$ può essere scritto sia come somma di $\mathbf{x} \in S_1$ e $\mathbf{0} \in S_2$, sia come somma di $\mathbf{0} \in S_1$ e $\mathbf{x} \in S_2$.

Ma essendo $\mathbf{x}$, per ipotesi, esprimibile univocamente come somma di due vettori, l'uno appartenente a $S_1$ e l'altro appartenente a $S_2$, ne consegue che $\mathbf{x} = \mathbf{0}$, e cioè che $S_1 \cap S_2 = S_0$ (11). ∎

Dato un sottospazio $S_1$ di uno spazio vettoriale non nullo S, vogliamo dimostrare l'esistenza di (almeno) un sottospazio $S_2$ di S tale che $S_1$ e $S_2$ siano sottospazi complementari in S.

**TEOREMA 14**

> Dato un sottospazio $S_1$ di uno spazio vettoriale non nullo S, esiste in S (almeno) un sottospazio $S_2$ tale che $S = S_1 \oplus S_2$.

*Dim*. Qualora $S_1$ sia un sottospazio improprio di S, il teorema risulta banalmente vero. Infatti, come facilmente si verifica, se $S_1 = S_0$, allora $S_2 = S$; se $S_1 = S$, allora $S_2 = S_0$.

(11) Il lettore è invitato a interpretare gli Esempi 19 e 20 in funzione del teorema ora dimostrato.

Supposto dunque che sia $0 < \dim(S_1) = p_1 < \dim(S) = p$, $S_1$ ha una base costituita da $p_1$ vettori $\mathbf{x}_1, \dots, \mathbf{x}_{p1}$ i quali, per il fatto che $S_1$ è un sottospazio di S, sono parte di una base di S (Cfr. il Corollario al Teorema 8).

Considerati $p-p_1$ vettori $\mathbf{x}_{p1+1}, \dots, \mathbf{x}_p$ che completano una base di S e dei quali il Teorema 8 assicura l'esistenza, tali vettori formano una base di uno spazio $S_2$, sottospazio di S.

Vogliamo dimostrare che

$$S = S_1 + S_2 \quad , \quad S_1 \cap S_2 = S_0 \, .$$

È ovvio che $S = S_1+S_2$. Indicando poi con $\mathbf{z}$ un generico vettore appartenente a $S_1 \cap S_2$, risulta

$$\mathbf{z} = a_1\mathbf{x}_1 + \dots + a_{p1}\mathbf{x}_{p1} \quad , \quad \mathbf{z} = a_{p1+1}\mathbf{x}_{p1+1} + \dots + a_p\mathbf{x}_p$$

da cui

$$a_1\mathbf{x}_1 + \dots + a_{p1}\mathbf{x}_{p1} - a_{p1+1}\mathbf{x}_{p1+1} - \dots - a_p\mathbf{x}_p = \mathbf{0} \, .$$

Ma, poiché $\mathbf{x}_1, \dots, \mathbf{x}_{p1}, \mathbf{x}_{p1+1}, \dots, \mathbf{x}_p$ sono linearmente indipendenti, la relazione precedente implica che sia $a_1 = \dots = a_{p1} = a_{p1+1} = \dots = a_p = 0$; ne consegue che $\mathbf{z} = \mathbf{0}$ e, quindi, che $S_1 \cap S_2 = S_0$.

**ESEMPIO 21.** Sia $S_1$ il sottospazio di $R^2$ la cui base è data dal vettore

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} .$$

Completata la base di $R^2$ mediante il vettore

$$\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} ,$$

questo forma la base di un sottospazio $S_2$ complementare di $S_1$ in $R^2$. ▮

Ci proponiamo, adesso, di stabilire la relazione esistente tra la dimensione di uno spazio vettoriale non nullo S e la dimensione di due suoi sottospazi $S_1$ e $S_2$, tali che $S = S_1 \oplus S_2$.

**TEOREMA 15**

Dati due sottospazi $S_1$ e $S_2$ di uno spazio vettoriale non nullo S, tali che $S = S_1 \oplus S_2$, risulta

$$\dim(S) = \dim(S_1) + \dim(S_2) \ .$$

*Dim*. Qualora sia $S_1 = S_0$ ($S_2 = S$) oppure $S_2 = S_0$ ($S_1 = S$), il teorema risulta banalmente vero.

Supponiamo dunque che siano

$$1 \le p_1 = \dim(S_1) < \dim(S) \quad , \quad 1 \le p_2 = \dim(S_2) < \dim(S) \ .$$

Se $\mathbf{x}_1, \dots, \mathbf{x}_{p1}$ formano una base di $S_1$ e, analogamente, $\mathbf{x}_{p1+1}, \dots, \mathbf{x}_{p1+p2}$ formano una base di $S_2$, per ogni $\mathbf{y} \in S_1$ e ogni $\mathbf{z} \in S_2$, si ha

$$\mathbf{y} = a_1 \mathbf{x}_1 + \dots + a_{p1} \mathbf{x}_{p1} \quad , \quad \mathbf{z} = a_{p1+1} \mathbf{x}_{p1+1} + \dots + a_{p1+p2} \mathbf{x}_{p1+p2} \ .$$

Ma, poiché $S = S_1 \oplus S_2$, ogni vettore $\mathbf{x} \in S$ è esprimibile in modo univoco come somma di $\mathbf{y}$ e $\mathbf{z}$.

Questo prova che $\mathbf{x}_1, \dots, \mathbf{x}_{p1}, \mathbf{x}_{p1+1}, \dots, \mathbf{x}_{p1+p2}$ costituiscono una base di S e che $\dim(S) = p_1 + p_2 = \dim(S_1) + \dim(S_2)$ (12).

(12) Reciprocamente, se

$$\dim(S) = \dim(S_1) + \dim(S_2) \quad , \quad S_1 \cap S_2 = S_0$$

allora $S = S_1 \oplus S_2$. Il lettore è invitato a dimostrare tale asserzione.

# Cap. I
# COMPLEMENTI

**1**

Dati tre vettori $\mathbf{x} = [x_i]_{(n,1)}$, $\mathbf{y} = [y_i]_{(n,1)}$, $\mathbf{z} = [z_i]_{(n,1)}$, l'operazione di addizione gode della cosiddetta *proprietà di cancellazione* espressa dalla proposizione seguente: $\mathbf{x} + \mathbf{z} = \mathbf{y} + \mathbf{z}$ *se e soltanto se* $\mathbf{x} = \mathbf{y}$.

Infatti, $\mathbf{x} + \mathbf{z} = \mathbf{y} + \mathbf{z}$ equivale a $x_i + z_i = y_i + z_i$ per $i = 1, \dots, n$. D'altro canto, per la proprietà di cancellazione dei numeri reali, $x_i + z_i = y_i + z_i$ se e soltanto se $x_i = y_i$ $(i = 1, \dots, n)$.

**2**

Dati un numero reale c e un vettore $\mathbf{x} = [x_i]_{(n,1)}$, se $c\mathbf{x} = \mathbf{0}$, allora è $c = 0$ oppure $\mathbf{x} = \mathbf{0}$.

Infatti, $c\mathbf{x} = \mathbf{0}$ equivale a $cx_i = 0$ per $i = 1, \dots, n$. D'altro canto, per una nota proprietà dei numeri reali, $cx_i = 0$ se e soltanto se risulta $c = 0$ oppure $x_i = 0$ $(i = 1, \dots, n)$.

Con un ragionamento simile si dimostra poi che se $c\mathbf{x} = c\mathbf{y}$ con $c \neq 0$, allora $\mathbf{x} = \mathbf{y}$.

**3**

Sia S uno spazio vettoriale e supponiamo, per concretezza, che S non sia lo spazio nullo.

Considerati due sottospazi $S_1$ e $S_2$ di S, si chiama *unione* di $S_1$ e $S_2$ l'insieme, denotato con il simbolo $S_1 \cup S_2$, costituito da tutti quei vettori che appartengono a $S_1$ o a $S_2$. Ovviamente, $S_1 \cup S_2$ è un sottoinsieme di S; *non* è, tuttavia, in generale un sottospazio di S.

Per esempio, siano $S_1$ e $S_2$ i sottospazi di $R^2$ generati, rispettivamente, dai vettori

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad , \quad \mathbf{y}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} .$$

È facile verificare che la somma di $\mathbf{x}_1$ e $\mathbf{y}_1$ – pur appartenendo $\mathbf{x}_1$ e $\mathbf{y}_1$ a $S_1 \cup S_2$ – non appartiene a $S_1 \cup S_2$, come la definizione di sottospazio richiede.

**4**

Vogliamo estendere le definizioni di intersezione, somma, somma diretta di due sottospazi al caso in cui si considerino $t > 2$ sottospazi $S_1, \dots, S_t$ di uno spazio vettoriale non nullo S.

Si chiama *intersezione* di $S_1, \dots, S_t$ l'insieme – denotato con il simbolo $S_1 \cap \dots \cap S_t$ oppure con l'altro

$$\bigcap_{k=1}^{t} S_k$$

– costituito da tutti quei vettori ciascuno dei quali appartiene a ogni $S_k$ per $k = 1, \dots, t$.

Si chiama *somma* di $S_1, \dots, S_t$ l'insieme – denotato con il simbolo $S_1 + \dots + S_t$ oppure con l'altro

$$\sum_{k=1}^{t} S_k$$

– costituito da tutti quei vettori della forma $\mathbf{x}_1 + \dots + \mathbf{x}_t$ con $\mathbf{x}_k \in S_k$ per $k = 1, \dots, t$.

Si dimostra facilmente che sia l'intersezione sia la somma di $S_1, \dots, S_t$ sono sottospazi di S.

La somma di $S_1, \dots, S_t$ è detta *somma diretta* – e la si scrive $S_1 \oplus \dots \oplus S_t$ – se, per ogni $h = 1, \dots, t$,

$$S_h \cap \sum_{k \neq h} S_k = S_0 \; .$$

Ovviamente, $S_1 \oplus \ldots \oplus S_t$ è un sottospazio di S.

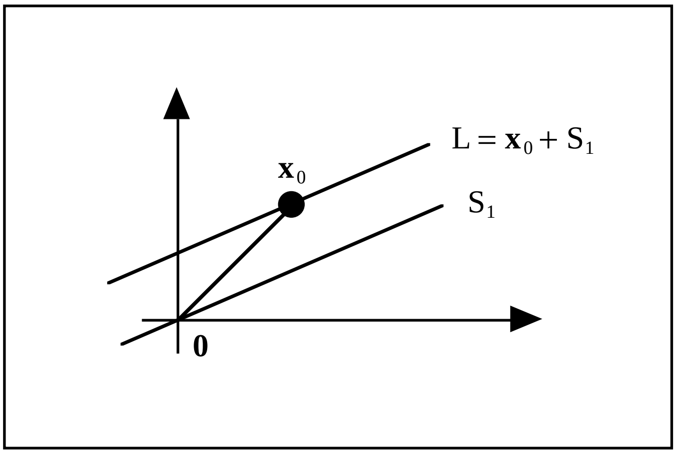
Si dimostra che S è somma diretta di $S_1, \ldots, S_t$ se e soltanto se ciascun vettore $\mathbf{x} \in S$ può essere espresso univocamente come somma di $\mathbf{x}_1, \ldots, \mathbf{x}_t$ con $\mathbf{x}_k \in S_k$ per $k = 1, \ldots, t$.

**5**

Siano S uno spazio vettoriale di dimensione p e $S_1$ un suo sottospazio. Dato un vettore $\mathbf{x}_0 \in S$, il sottoinsieme L di S costituito da tutti i vettori della forma $\mathbf{x} = \mathbf{x}_0 + \mathbf{y}$, ottenuti facendo variare $\mathbf{y}$ in $S_1$, si chiama *varietà lineare* di *direzione* o *giacitura* $S_1$ e si scrive $L = \mathbf{x}_0 + S_1$.

Si dice anche che L è ottenuta da $S_1$ mediante la *traslazione* $\mathbf{x}_0$.

La *dimensione* di L – indicata con dim (L) – è, per definizione, eguale alla dimensione di $S_1$.




Si noti che, essendo $\mathbf{0} \in S_1$, $\mathbf{x}_0 \in L$.

Inoltre, dato un vettore $\tilde{\mathbf{y}} \in S_1$, poiché la varietà lineare $\tilde{L} = \mathbf{x}_0 + \tilde{\mathbf{y}} + S_1$ coincide evidentemente con la varietà lineare $L = \mathbf{x}_0 + S_1$, $\mathbf{x}_0$ si può sostituire con $\mathbf{x}_0 + \tilde{\mathbf{y}}$, ovvero $\mathbf{x}_0$ è *arbitrario* in L.

Ovviamente, ogni sottospazio $S_1$ di S è una varietà lineare (ottenuta prendendo $\mathbf{x}_0 \in S_1$ e, in particolare, $\mathbf{x}_0 = \mathbf{0}$).

Non è, tuttavia, vero il contrario: in effetti, come facilmente si verifica, una varietà lineare è un sottospazio se e soltanto se $\mathbf{x}_0 \in S_1$.

Tenuto conto di ciò e del fatto che $\mathbf{x}_0$ è arbitrario in L, possiamo anche dire che una varietà lineare è un sottospazio se e soltanto se il vettore nullo vi appartiene.

Una varietà lineare di dimensione h = 1, ottenuta mediante la traslazione $\mathbf{x}_0$, è anche detta *retta passante per* $\mathbf{x}_0$.

A sua volta, una varietà lineare di dimensione h = 2, ottenuta mediante la traslazione $\mathbf{x}_0$, è anche detta *piano passante per* $\mathbf{x}_0$.

Infine, una varietà lineare di dimensione $h = p - 1$, ottenuta mediante la traslazione $\mathbf{x}_0$, è talvolta denominata *iperpiano passante per* $\mathbf{x}_0$ (1).

**6**

Una definizione più generale di spazio vettoriale è la seguente.

Siano S un insieme (non vuoto) di oggetti di natura qualsiasi e R l'insieme dei numeri reali.

Si dice che S è uno *spazio vettoriale su* R se, definita una operazione di *addizione* (+) tra gli elementi di S e una operazione di *moltiplicazione per un numero reale* (·) tra gli elementi di R e quelli di S, il risultato di queste operazioni è ancora un elemento di S e inoltre, per tutti gli $\alpha, \beta, \gamma \in S$ e tutti gli $a, b \in R$, si ha:

1. $\alpha + \beta = \beta + \alpha$
2. $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$
3. esiste un (unico) elemento $\theta \in S$ tale che $\alpha + \theta = \alpha$
4. esiste un (unico) elemento $(-\alpha) \in S$ tale che $\alpha + (-\alpha) = \theta$
5. $a \cdot (\alpha + \beta) = a \cdot \alpha + a \cdot \beta$
6. $(a + b) \cdot \alpha = a \cdot \alpha + b \cdot \alpha$
7. $(ab) \cdot \alpha = a \cdot (b \cdot \alpha)$
8. $1 \cdot \alpha = \alpha$.

(1) Per ulteriori considerazioni si rinvia al volume di E. Sernesi citato in Bibliografia.

Esempi elementari di spazi vettoriali secondo la definizione ora data – oltre allo spazio vettoriale definito nel paragrafo 4 di questo capitolo – sono i seguenti [(2)]:

(a) l'insieme $\mathcal{P}_p[x]$ dei polinomi a coefficienti reali di grado inferiore o eguale a p;

(b) l'insieme $\mathcal{P}[x]$ di *tutti* i polinomi a coefficienti reali.

Dati p elementi $\alpha_1, \dots, \alpha_p$ di S e p numeri reali $c_1, \dots, c_p$, si chiama *combinazione lineare* di $\alpha_1, \dots, \alpha_p$ con *coefficienti* $c_1, \dots, c_p$ l'elemento $\alpha \in S$ espresso da

$$\alpha = c_1 \cdot \alpha_1 + \dots + c_p \cdot \alpha_p \ .$$

Dati p elementi $\alpha_1, \dots, \alpha_p$ di S, si dice che essi sono *linearmente dipendenti* se esistono p numeri reali $c_1, \dots, c_p$ *non tutti nulli* tali che

$$c_1 \cdot \alpha_1 + \dots + c_p \cdot \alpha_p = \theta \ .$$

Se, al contrario, la relazione precedente implica che sia $c_1 = \dots = c_p = 0$, allora $\alpha_1, \dots, \alpha_p$ si dicono *linearmente indipendenti*.

Dati p elementi $\alpha_1, \dots, \alpha_p$ di S, si dice che essi formano una *base* (*finita*) di S se ogni $\alpha \in S$ è esprimibile *in modo univoco* come combinazione lineare di $\alpha_1, \dots, \alpha_p$.

Si può dimostrare che se S ha una base formata da p elementi, allora ogni altra base di S è composta da p elementi.

Il numero p è detto la *dimensione* (*finita*) di S e si scrive $\dim(S) = p$.

Se S è formato dal solo elemento $\theta$ (*spazio nullo*) si pone, per definizione, $\dim(S) = 0$.

Per esempio, l'insieme $\mathcal{P}_p[x]$ dei polinomi a coefficienti reali di grado inferiore o eguale a p è uno spazio vettoriale di dimensione $p+1$ poiché ha

(2) Nell'uno e nell'altro caso, le operazioni di addizione tra polinomi e di moltiplicazione di un numero reale per un polinomio sono definite nel modo usuale. L'elemento $\theta$ è dato dal polinomio a coefficienti tutti nulli.

una base costituita da $\alpha_0 \equiv 1, \alpha_1 \equiv x, ..., \alpha_p \equiv x^p$.

Al contrario, come si può facilmente dimostrare, l'insieme $\mathcal{P}[x]$ di tutti i polinomi non possiede una base finita e, in tal caso, si dice che è uno spazio vettoriale di *dimensione infinita*.

*Nel seguito, ci occuperemo esclusivamente di spazi vettoriali di dimensione finita o nulla.*

*Avvertiamo inoltre che, almeno sia altrimenti specificato, con tale dizione intenderemo riferirci a spazi vettoriali così come sono stati definiti nel paragrafo 4 del presente capitolo.*

**7**

Due spazi vettoriali S e Z su R si dicono *isomorfi* se tra essi è possibile stabilire una corrispondenza biunivoca ($\leftrightarrow$) (3) la quale verifichi le proprietà seguenti $(\alpha, \beta, \gamma \in S; \tilde{\alpha}, \tilde{\beta}, \tilde{\gamma} \in Z; c \in R)$ (4) :

(i) se $\alpha \leftrightarrow \tilde{\alpha}$ e $\beta \leftrightarrow \tilde{\beta}$, allora $(\alpha + \beta) \leftrightarrow (\tilde{\alpha} + \tilde{\beta})$;

(ii) se $\gamma \leftrightarrow \tilde{\gamma}$, allora $(c \cdot \gamma) \leftrightarrow (c \cdot \tilde{\gamma})$.

Si noti che, qualora due spazi vettoriali S e Z siano isomorfi, essi risultano «indistinguibili» per quanto attiene allo loro struttura di spazio vettoriale, nel senso che ogni proposizione o proprietà di S deducibile mediante gli assiomi che definiscono S medesimo può essere «tradotta» in una analoga proposizione o proprietà di Z, e viceversa.

Sia adesso S uno spazio vettoriale su R di dimensione p.

Se $\alpha_1, ..., \alpha_p$ costituiscono una base di S, a ogni elemento $\alpha \in S$ risultano univocamente associati p numeri reali $c_1, ..., c_p$, ossia un vettore $\mathbf{c} \in R^p$.

Viceversa, a ogni vettore $\mathbf{c} \in R^p$ formato da p numeri reali $c_1, ..., c_p$ risulta univocamente associato un elemento $\alpha \in S$.

Sussistendo una corrispondenza biunivoca tra S e $R^p$ e risultando inoltre verificate le condizioni (i) e (ii) elencate in precedenza, si conclude che S e

(3) Ovvero, una funzione che sia al tempo stesso iniettiva e surgettiva.

(4) I simboli utilizzati per denotare le operazioni di addizione e moltiplicazione per un numero reale hanno, in generale, un diverso significato in S e Z.

$R^p$ sono isomorfi e, quindi, che tutti gli spazi vettoriali su R di una stessa dimensione sono isomorfi [5].

Si noti che tale isomorfismo dipende dalla base prescelta di S o, come anche si dice, non è «canonico».

(5) Si può facilmente dimostrare che, reciprocamente, tutti gli spazi vettoriali su R isomorfi hanno una stessa dimensione.

# Cap. II
# MATRICI

## 1 GENERALITÀ SULLE MATRICI

**1.1** Chiamiamo *matrice* di *ordine* $(n, p)$ [1] o, più semplicemente, *matrice* $(n, p)$ ogni quadro di $n \cdot p$ numeri reali. Tali numeri sono detti *elementi* o *componenti* della matrice, sono disposti su n *righe* e p *colonne* e racchiusi tra parentesi quadre [2].

Per esempio,

$$\begin{bmatrix} 2 & 1 & 5 \\ 0 & 3 & 6 \end{bmatrix}$$

è una matrice di ordine $(2, 3)$.

Seguendo una notazione assai diffusa, denoteremo le matrici con lettere maiuscole in grassetto $(\mathbf{X}, \mathbf{Y}, ...)$.

Fanno eccezione a questa regola le matrici di ordine $(n, 1)$ e $(1, p)$ – corrispondenti, rispettivamente, a un vettore colonna di ordine n e a un vettore riga di ordine p – per le quali si userà di preferenza la simbologia introdotta nel capitolo precedente.

Impiegheremo la scrittura $\mathbf{X}_{(n, p)}$ per evidenziare che una matrice $\mathbf{X}$ è di ordine $(n, p)$.

Rappresenteremo, inoltre, simbolicamente una generica matrice $\mathbf{X}$ di

---

(1) Si legge «n *per* p». Ometteremo, talvolta, di indicare l'ordine di una matrice.

(2) Alcuni, invece delle parentesi (quadre o tonde), usano le doppie sbarre verticali.

ordine $(n, p)$ con

$$\begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \cdots & \cdots & \cdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix}$$

dove a ciascun elemento $x_{ij}$ $(i = 1, \ldots, n; j = 1, \ldots, p)$ sono associati due indici, il primo dei quali si riferisce alla riga e il secondo alla colonna di appartenenza.

Talvolta, la matrice **X** di cui sopra è rappresentata, più compattamente, con $[x_{ij}]_{(n,p)}$ oppure con $[x_{ij}]$.

**1.2** Due matrici $\mathbf{X} = [x_{ij}]_{(n,p)}$ e $\mathbf{Y} = [y_{ij}]_{(n,p)}$ sono *eguali*, e si scrive $\mathbf{X} = \mathbf{Y}$, se e soltanto se $x_{ij} = y_{ij}$ per $i = 1, \ldots, n$ e $j = 1, \ldots, p$.

In altri termini, due matrici dello stesso ordine sono eguali se e soltanto se gli elementi corrispondenti sono eguali.

**1.3** Indicheremo, adesso, alcuni tipi particolari di matrici che ricorrono frequentemente in molte questioni di algebra lineare.

(a) La matrice di ordine $(n, p)$ i cui elementi sono tutti eguali a zero è detta *matrice nulla* (o *matrice zero*) ed è indicata con il simbolo **O**.

Per esempio, la matrice

$$\mathbf{O} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

è la matrice nulla di ordine $(2, 3)$.

(b) Se una matrice possiede lo stesso numero n di righe e di colonne si dice che è una *matrice quadrata* di ordine $(n, n)$ [(3)].

Gli elementi di una matrice quadrata $\mathbf{X} = [x_{ij}]_{(n,n)}$ che hanno il primo e il secondo indice eguali $(i = j)$ – vale a dire, $x_{11}, \ldots, x_{nn}$ – formano la cosiddetta *diagonale principale* di **X**.

(3) O, più semplicemente, una matrice (quadrata) di ordine n.

Se $x_{ii} = c_i$ per $i = 1, \dots, n$ e $x_{ij} = 0$ per $i \neq j$, la matrice in questione è detta *diagonale*.

Poiché una matrice diagonale $\mathbf{X}$ è caratterizzata dagli elementi $c_1, \dots, c_n$ che compaiono sulla diagonale principale, essendo tutti gli altri eguali a zero, si usa talvolta scrivere

$$\mathbf{X} = \mathrm{diag}(c_1, \dots, c_n) \ .$$

Per esempio, la matrice

$$\mathbf{X} = \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix} = \mathrm{diag}(6,1)$$

è una matrice diagonale di ordine $(2,2)$.

In particolare, se in una matrice diagonale $\mathbf{X}$ risulta $x_{ii} = c$ per $i = 1, \dots, n$, $\mathbf{X}$ è detta *matrice scalare*.

(c) Se in una matrice scalare di ordine $(n,n)$ gli elementi della diagonale principale sono tutti eguali a 1, la matrice in questione è detta *matrice unità* (o *matrice identica*) ed è usualmente indicata con il simbolo **I**.

Per esempio, la matrice

$$\mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

è la matrice unità di ordine $(2,2)$.

## 2 ADDIZIONE E MOLTIPLICAZIONE PER UN NUMERO REALE. SOTTRAZIONE

**2.1** Date due matrici $\mathbf{X} = [x_{ij}]_{(n,p)}$ e $\mathbf{Y} = [y_{ij}]_{(n,p)}$, si chiama *somma* di $\mathbf{X}$ e $\mathbf{Y}$ la matrice, che indichiamo con $\mathbf{X} + \mathbf{Y}$, i cui elementi sono espressi ordinatamente dai numeri $x_{ij} + y_{ij}$ per $i = 1, \dots, n$ e $j = 1, \dots, p$. Si ha, cioè,

$$\mathbf{X} + \mathbf{Y} = \begin{bmatrix} x_{11} + y_{11} & \cdots & x_{1p} + y_{1p} \\ \cdots & \cdots & \cdots \\ x_{n1} + y_{n1} & \cdots & x_{np} + y_{np} \end{bmatrix} .$$

**ESEMPIO 1.** Date le matrici

$$\mathbf{X} = \begin{bmatrix} 2 & 1 \\ 7 & 0 \\ 3 & 1 \end{bmatrix} \quad , \quad \mathbf{Y} = \begin{bmatrix} 1 & 3 \\ 0 & 7 \\ 1 & 0 \end{bmatrix}$$

si ha

$$\mathbf{X} + \mathbf{Y} = \begin{bmatrix} 2 & 1 \\ 7 & 0 \\ 3 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 3 \\ 0 & 7 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 7 & 7 \\ 4 & 1 \end{bmatrix} .$$ ∎

Come risulta immediatamente dalla definizione data, l'operazione in questione – che riceve la denominazione di *addizione* – è *commutativa* e *associativa*; vale a dire, qualunque siano le matrici $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ di ordine $(n, p)$, si ha

(i) $\mathbf{X} + \mathbf{Y} = \mathbf{Y} + \mathbf{X}$

(ii) $(\mathbf{X} + \mathbf{Y}) + \mathbf{Z} = \mathbf{X} + (\mathbf{Y} + \mathbf{Z})$.

Quest'ultima proprietà consente di indicare, senza possibilità di equivoci, la somma di $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ con $\mathbf{X} + \mathbf{Y} + \mathbf{Z}$, evitando l'uso delle parentesi, e di estendere l'addizione a un numero qualsiasi, purché finito, di matrici di ordine $(n, p)$.

**2.2** Dati un numero reale c e una matrice $\mathbf{X} = [x_{ij}]_{(n,p)}$, si chiama *prodotto di* c *per* $\mathbf{X}$ o *prodotto di* $\mathbf{X}$ *per* c la matrice, che indichiamo con $c\mathbf{X}$ o $\mathbf{X}c$, i cui elementi sono espressi ordinatamente dai numeri $cx_{ij}$ per $i = 1, \ldots, n$ e $j = 1, \ldots, p$. Si ha, cioè,

$$c\mathbf{X} = \begin{bmatrix} cx_{11} & \cdots & cx_{1p} \\ \cdots & \cdots & \cdots \\ cx_{n1} & \cdots & cx_{np} \end{bmatrix} = \mathbf{X}c .$$

**ESEMPIO 2.** Dati il numero reale c = 2 e la matrice

$$\mathbf{X} = \begin{bmatrix} 1 & 0 & 3 \\ 4 & 1 & 0 \end{bmatrix}$$

risulta

$$c\mathbf{X} = 2\begin{bmatrix} 1 & 0 & 3 \\ 4 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 6 \\ 8 & 2 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 3 \\ 4 & 1 & 0 \end{bmatrix} 2 = \mathbf{X}c\,. \qquad \blacksquare$$

L'operazione ora definita riceve la denominazione di *moltiplicazione per un numero reale*.

Come si può facilmente verificare, essa gode, qualunque siano le matrici $\mathbf{X}$ e $\mathbf{Y}$ di ordine $(n, p)$ e i numeri reali a e b, delle seguenti proprietà:

(i) $(ab)\mathbf{X} = a(b\mathbf{X})$

(ii) $a(\mathbf{X} + \mathbf{Y}) = a\mathbf{X} + a\mathbf{Y}$

(iii) $(a + b)\mathbf{X} = a\mathbf{X} + b\mathbf{X}$.

**2.3** Date due matrici $\mathbf{X} = [x_{ij}]_{(n,p)}$ e $\mathbf{Y} = [y_{ij}]_{(n,p)}$, posto $(-\mathbf{Y}) = (-1)\mathbf{Y}$, si chiama *differenza* di $\mathbf{X}$ e $\mathbf{Y}$ la matrice $\mathbf{X} + (-\mathbf{Y}) = \mathbf{X} + (-1)\mathbf{Y}$ che scriviamo, più semplicemente, $\mathbf{X} - \mathbf{Y}$. Si ha, cioè,

$$\mathbf{X} - \mathbf{Y} = \begin{bmatrix} x_{11} - y_{11} & \cdots & x_{1p} - y_{1p} \\ \cdots & \cdots & \cdots \\ x_{n1} - y_{n1} & \cdots & x_{np} - y_{np} \end{bmatrix}.$$

**ESEMPIO 3.** Date le matrici

$$\mathbf{X} = \begin{bmatrix} 3 & 4 & 7 \\ 2 & 6 & 2 \end{bmatrix} \quad , \quad \mathbf{Y} = \begin{bmatrix} 3 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix}$$

si ha

$$\mathbf{X} - \mathbf{Y} = \begin{bmatrix} 0 & 2 & 7 \\ 0 & 6 & 0 \end{bmatrix}. \qquad \blacksquare$$

L'operazione ora definita, derivata dalle operazioni di addizione e di moltiplicazione per un numero reale, riceve la denominazione di *sottrazione* e consente di stabilire, qualunque siano le matrici $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ di ordine $(n, p)$ e i numeri reali a e b, alcuni utili risultati di immediata verifica:

(i) $\mathbf{X} - \mathbf{X} = \mathbf{O}$

(ii) $-(-\mathbf{X}) = \mathbf{X}$

(iii) $-(\mathbf{X} - \mathbf{Y}) = \mathbf{Y} - \mathbf{X}$

(iv) $\mathbf{X} - (\mathbf{Y} - \mathbf{Z}) = (\mathbf{X} - \mathbf{Y}) + \mathbf{Z}$

(v) $a(\mathbf{X} - \mathbf{Y}) = a\mathbf{X} - a\mathbf{Y}$

(vi) $(a - b)\mathbf{X} = a\mathbf{X} - b\mathbf{X}$.

## 3 MOLTIPLICAZIONE DI MATRICI

Siano **X** e **Y** due matrici tali che il numero di colonne della prima matrice sia eguale al numero di righe della seconda: per esempio, siano $\mathbf{X} = [x_{ij}]_{(n,p)}$ e $\mathbf{Y} = [y_{jk}]_{(p,q)}$.

Due matrici siffatte si dicono *conformabili* rispetto all'operazione di *moltiplicazione di matrici* che passiamo ora a definire.

Si definisce *prodotto di* **X** *per* **Y**, nell'ordinamento scritto, e si indica con **XY**, la matrice **Z** di ordine $(n, q)$ il cui generico elemento $z_{ik}$ è dato da (4)

$$z_{ik} = \sum_{j=1}^{p} x_{ij} y_{jk} .$$

Si ha, cioè,

$$\mathbf{Z} = [z_{ik}] = \left[ \sum_{j=1}^{p} x_{ij} y_{jk} \right] .$$

Il prodotto così definito si dice eseguito *riga per colonna*. Si dice anche che la matrice **X** *premoltiplica* la matrice **Y** o che la matrice **Y** *postmoltiplica* la matrice **X**.

**ESEMPIO 4.** Date le matrici

(4) Come si vedrà nel seguito (Cfr. il punto 4 dei Complementi al presente capitolo), questo non è l'unico modo di definire l'operazione di moltiplicazione di matrici.

$$\mathbf{X} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} \quad , \quad \mathbf{Y} = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}$$

risulta

$$\mathbf{XY} = \begin{bmatrix} 0 & 1 \\ 1 & 4 \\ 1 & 5 \end{bmatrix} .$$

**OSSERVAZIONE 1.** Data una matrice $\mathbf{X}_{(n,p)}$, la matrice nulla **O** e quella unità **I**, ciascuna di ordine appropriato, verificano le ovvie proprietà:

(a) $\mathbf{O}_{(q,n)}\mathbf{X}_{(n,p)} = \mathbf{O}_{(q,p)} \quad , \quad \mathbf{X}_{(n,p)}\mathbf{O}_{(p,q)} = \mathbf{O}_{(n,q)}$

(b) $\mathbf{I}_{(n,n)}\mathbf{X}_{(n,p)} = \mathbf{X}_{(n,p)} \quad , \quad \mathbf{X}_{(n,p)}\mathbf{I}_{(p,p)} = \mathbf{X}_{(n,p)}$ .

**OSSERVAZIONE 2.** La regola pratica che consente di verificare se due matrici sono conformabili rispetto all'operazione di moltiplicazione – e, in caso affermativo, di stabilire l'ordine della matrice risultante – può essere esposta nei seguenti termini.

Date due matrici **X** e **Y** – di ordine, rispettivamente, $(n,p)$ e $(p,q)$ – si scrivono, in successione, le espressioni che indicano l'ordine di ciascuna matrice; se le due matrici sono conformabili, i numeri posti all'interno dell'espressione $(n,p)(p,q)$ – i quali indicano, rispettivamente, il numero di colonne della prima matrice e di righe della seconda matrice – devono essere eguali; in tal caso, si cancellano e l'espressione risultante esprime l'ordine della matrice prodotto di quelle date: $(n,p)(p,q) = (n,q)$. ▮

Tenendo presente l'Osservazione che precede, si ha che:

(i) Il prodotto di un vettore riga per un vettore colonna, entrambi di ordine n, è una matrice di ordine $(1,1)$ [(5)].

(5) Si osservi che, se in una data espressione matriciale contenente l'indicazione di una successione di operazioni (ovviamente, tutte eseguibili sulla base delle regole che per esse sono state stabilite) accade che uno dei termini sia una matrice di ordine (1,1), è lecito trattare tale matrice come un numero reale.

(ii) Il prodotto di un vettore colonna di ordine n per un vettore riga di ordine p è una matrice di ordine $(n, p)$.

(iii) Il prodotto di una matrice di ordine $(n, p)$ per un vettore colonna di ordine p è un vettore colonna di ordine n.

(iv) Il prodotto di un vettore riga di ordine p per una matrice di ordine $(p, n)$ è un vettore riga di ordine n.

Si osservi che, nella moltiplicazione, ha importanza essenziale l'ordinamento delle matrici **X** e **Y**. Infatti, due matrici che sono conformabili rispetto a tale operazione se considerate nell'ordinamento **X** e **Y** possono non esserlo se considerate nell'ordinamento inverso **Y** e **X**. Di conseguenza, può darsi che si possa eseguire, secondo la definizione data, il prodotto di **X** per **Y** e che, invece, non si possa eseguire il prodotto di **Y** per **X**.

Condizione necessaria e sufficiente affinché esistano sia **XY** sia **YX** è che **X** e **Y** abbiano, rispettivamente, ordini $(n, p)$ e $(p, n)$. In particolare, i prodotti **XY** e **YX** esistono entrambi se **X** e **Y** sono matrici quadrate dello stesso ordine. Tuttavia, anche quando le operazioni **XY** e **YX** sono eseguibili, risulta in genere $\mathbf{XY} \neq \mathbf{YX}$, vale a dire, l'*operazione di moltiplicazione di matrici non è commutativa* (6).

Per esempio, date le matrici

$$\mathbf{X} = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} \quad , \quad \mathbf{Y} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

risulta

$$\mathbf{XY} = \begin{bmatrix} 1 & 3 \\ 3 & 3 \end{bmatrix} \quad , \quad \mathbf{YX} = \begin{bmatrix} 4 & 2 \\ 3 & 0 \end{bmatrix} .$$

e cioè $\mathbf{XY} \neq \mathbf{YX}$.

Può accadere, inoltre, come è facile verificare attraverso esempi numerici (Cfr. il punto 3 dei Complementi a questo capitolo), che sia $\mathbf{XY} = \mathbf{O}$ senza

(6) Se accade che $\mathbf{XY} = \mathbf{YX}$, le matrici sono dette *permutabili*.

che né $\mathbf{X}$ né $\mathbf{Y}$ siano eguali a $\mathbf{O}$ (*esistenza di divisori* $(\mathbf{X},\mathbf{Y})$ *dello zero* $(\mathbf{O})$), oppure che sia $\mathbf{XY}=\mathbf{XZ}$ con $\mathbf{Y}\neq\mathbf{Z}$ e $\mathbf{X}\neq\mathbf{O}$ (7).

L'operazione di moltiplicazione di matrici gode tuttavia della proprietà associativa e di quella distributiva rispetto all'addizione di matrici, che passiamo adesso a considerare.

Per quanto riguarda l'*associatività*, date tre matrici

$$\mathbf{X}=[x_{ij}]_{(n,p)} \quad , \quad \mathbf{Y}=[y_{jk}]_{(p,q)} \quad , \quad \mathbf{Z}=[z_{ks}]_{(q,t)}$$

è immediato verificare che

$$\begin{aligned}(\mathbf{XY})\mathbf{Z} &= \left[\sum_{k=1}^{q}\left(\sum_{j=1}^{p} x_{ij}y_{jk}\right)z_{ks}\right] = \left[\sum_{k=1}^{q}\sum_{j=1}^{p} x_{ij}y_{jk}z_{ks}\right] \\ &= \left[\sum_{j=1}^{p} x_{ij}\left(\sum_{k=1}^{q} y_{jk}z_{ks}\right)\right] = \mathbf{X}(\mathbf{YZ})\ .\end{aligned}$$

La proprietà suddetta consente di indicare, senza possibilità di equivoci, il prodotto di $\mathbf{X},\mathbf{Y},\mathbf{Z}$ con $\mathbf{XYZ}$, evitando l'uso delle parentesi.

**ESEMPIO 5.** Date le matrici

$$\mathbf{X}=\begin{bmatrix}1&0&0\\0&2&3\\1&1&0\end{bmatrix} \quad , \quad \mathbf{Y}=\begin{bmatrix}0&2\\1&0\\1&0\end{bmatrix} \quad , \quad \mathbf{Z}=\begin{bmatrix}1&1\\0&2\end{bmatrix}$$

risulta

$$(\mathbf{XY})\mathbf{Z}=\left(\begin{bmatrix}1&0&0\\0&2&3\\1&1&0\end{bmatrix}\begin{bmatrix}0&2\\1&0\\1&0\end{bmatrix}\right)\begin{bmatrix}1&1\\0&2\end{bmatrix}=\begin{bmatrix}0&2\\5&0\\1&2\end{bmatrix}\begin{bmatrix}1&1\\0&2\end{bmatrix}=\begin{bmatrix}0&4\\5&5\\1&5\end{bmatrix}$$

e, analogamente,

$$\mathbf{X}(\mathbf{YZ})=\begin{bmatrix}1&0&0\\0&2&3\\1&1&0\end{bmatrix}\left(\begin{bmatrix}0&2\\1&0\\1&0\end{bmatrix}\begin{bmatrix}1&1\\0&2\end{bmatrix}\right)=\begin{bmatrix}1&0&0\\0&2&3\\1&1&0\end{bmatrix}\begin{bmatrix}0&4\\1&1\\1&1\end{bmatrix}=\begin{bmatrix}0&4\\5&5\\1&5\end{bmatrix}. \quad \blacksquare$$

(7) L'argomento sarà ripreso al punto 9 dei Complementi al Cap. III.

Da quanto precede si trae immediatamente la possibilità di definire la moltiplicazione di un qualsiasi numero finito di matrici, prese in una determinata sequenza, purché beninteso ognuna sia conformabile con quella che segue nell'ordinamento stabilito.

In particolare, se le matrici considerate sono quadrate e dello stesso ordine $(n, n)$, il loro prodotto è ancora una matrice quadrata di ordine $(n, n)$.

Passando poi a considerare la *distributività rispetto all'addizione di matrici*, date tre matrici

$$\mathbf{X} = [x_{ij}]_{(n,p)} \quad , \quad \mathbf{Y} = [y_{ij}]_{(n,p)} \quad , \quad \mathbf{Z} = [z_{jk}]_{(p,q)}$$

e verificato che $\mathbf{X}$ e $\mathbf{Y}$ sono dello stesso ordine mentre sia $\mathbf{X}$ e $\mathbf{Z}$ sia $\mathbf{Y}$ e $\mathbf{Z}$ sono conformabili rispetto all'operazione di moltiplicazione, risulta che

$$(\mathbf{X}+\mathbf{Y})\mathbf{Z} = \left[\sum_{j=1}^{p} (x_{ij}+y_{ij})\ z_{jk}\right] = \left[\sum_{j=1}^{p} x_{ij}\ z_{jk} + \sum_{j=1}^{p} y_{ij}\ z_{jk}\right] = \mathbf{XZ}+\mathbf{YZ}\,.$$

Analogamente, nell'ipotesi che sia $\mathbf{Z}$ e $\mathbf{X}$ sia $\mathbf{Z}$ e $\mathbf{Y}$ siano conformabili rispetto all'operazione di moltiplicazione, si dimostra che

$$\mathbf{Z}(\mathbf{X}+\mathbf{Y}) = \mathbf{ZX}+\mathbf{ZY}\,.$$

È ovvio che, in genere, $(\mathbf{X}+\mathbf{Y})\mathbf{Z} \neq \mathbf{Z}(\mathbf{X}+\mathbf{Y})$.

**ESEMPIO 6.** Date le matrici

$$\mathbf{X} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad , \quad \mathbf{Y} = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} \quad , \quad \mathbf{Z} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

risulta

$$\begin{aligned}(\mathbf{X}+\mathbf{Y})\mathbf{Z} &= \left(\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}\right)\begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 2 \\ 7 & 2 & 5 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \mathbf{XZ}+\mathbf{YZ}\,.\end{aligned}$$

**OSSERVAZIONE 3.** Sussiste la relazione

$$(\mathbf{X}+\mathbf{Y})(\mathbf{Z}+\mathbf{V}) = \mathbf{XZ}+\mathbf{YZ}+\mathbf{XV}+\mathbf{YV}\,. \qquad \blacksquare$$

Infine, dati un numero reale c e due matrici **X** e **Y**, conformabili rispetto all'operazione di moltiplicazione, risulta immediatamente che

$$c(\mathbf{XY}) = (c\mathbf{X})\mathbf{Y} = \mathbf{X}(c\mathbf{Y})\,.$$

## 4 INVERSIONE

Data una matrice quadrata **X**, si chiama *matrice inversa* o, più semplicemente, *inversa* di **X** ogni matrice, che indichiamo con $\mathbf{X}^{-1}$, tale che

$$\mathbf{XX}^{-1} = \mathbf{X}^{-1}\mathbf{X} = \mathbf{I}\,.$$

Supposto che $\mathbf{X}^{-1}$ esista, vogliamo anzitutto mostrare che essa è *unica*.

Infatti, nell'ipotesi che esistano due matrici **S** e **T** tali che

$$\mathbf{XS} = \mathbf{SX} = \mathbf{I} \quad , \quad \mathbf{XT} = \mathbf{TX} = \mathbf{I}$$

premoltiplicando entrambi i membri di $\mathbf{XS} = \mathbf{I}$ per **T**, si ottiene

$$\mathbf{TXS} = \mathbf{TI} = \mathbf{T}\,;$$

ma $\mathbf{TX} = \mathbf{I}$ e, quindi, $\mathbf{S} = \mathbf{T}$.

Si noti anche che $\mathbf{X}^{-1}$ è dello stesso ordine di **X** e che $(\mathbf{X}^{-1})^{-1} = \mathbf{X}$.

**ESEMPIO 7.** Come si verifica facilmente, le matrici

$$\mathbf{X} = \begin{bmatrix} 2 & 5 \\ 3 & 8 \end{bmatrix} \quad , \quad \mathbf{X}^{-1} = \begin{bmatrix} 8 & -5 \\ -3 & 2 \end{bmatrix}$$

sono ciascuna l'inversa dell'altra. $\blacksquare$

Come si è osservato, il passaggio da una matrice **X** alla sua inversa $\mathbf{X}^{-1}$ – che riceve la denominazione di *inversione* – richiede che **X** sia quadrata.

Tuttavia, il fatto che **X** sia quadrata costituisce una condizione necessaria ma non sufficiente per l'esistenza dell'inversa di **X**.

Supponiamo infatti, come esempio banale, che sia

$$\mathbf{X} = \begin{bmatrix} 0 & x_{12} \\ 0 & x_{22} \end{bmatrix}.$$

È ovvio che, in tal caso, non esiste alcuna matrice **S** che premoltiplicata per **X** fornisce la matrice unità, poiché l'elemento posto all'incrocio della prima riga con la prima colonna del prodotto **S X** risulta, qualunque sia **S**, eguale a zero.

Le condizioni necessarie e sufficienti affinché esista l'inversa di una matrice **X** e il relativo metodo di calcolo saranno discussi in un capitolo successivo, dopo che avremo introdotto la nozione di determinante.

Per il momento, ci limitiamo a porre in evidenza la seguente proprietà, che è una conseguenza pressoché immediata della definizione data di inversa di una matrice.

Date due matrici (quadrate e dello stesso ordine) **X** e **Y** e supposto che esistano le inverse $\mathbf{X}^{-1}$ e $\mathbf{Y}^{-1}$, l'inversa del prodotto di **X** per **Y**, cioè $(\mathbf{X}\mathbf{Y})^{-1}$, è eguale al prodotto di $\mathbf{Y}^{-1}$ per $\mathbf{X}^{-1}$, cioè a $\mathbf{Y}^{-1}\mathbf{X}^{-1}$.

Infatti, tenendo presente la definizione di inversa, si ha che

$$\begin{aligned} (\mathbf{Y}^{-1}\mathbf{X}^{-1})\,\mathbf{X}\mathbf{Y} &= \mathbf{Y}^{-1}(\mathbf{X}^{-1}\mathbf{X})\,\mathbf{Y} &&= \mathbf{Y}^{-1}\mathbf{Y} &&= \mathbf{I} \\ \mathbf{X}\mathbf{Y}(\mathbf{Y}^{-1}\mathbf{X}^{-1}) &= \mathbf{X}(\mathbf{Y}\mathbf{Y}^{-1})\,\mathbf{X}^{-1} &&= \mathbf{X}\mathbf{X}^{-1} &&= \mathbf{I} \end{aligned}$$

e, quindi,

$$(\mathbf{X}\mathbf{Y})^{-1} = \mathbf{Y}^{-1}\mathbf{X}^{-1}\ .$$

**OSSERVAZIONE 4.** Tenendo presente l'associatività della moltiplicazione di matrici, la proprietà ora indicata può essere estesa a un numero qualsiasi, purché finito, di matrici.

Per esempio,

$$(\mathbf{X}\mathbf{Y}\mathbf{Z}\mathbf{V})^{-1} = ((\mathbf{X}\mathbf{Y}\mathbf{Z})\mathbf{V})^{-1} = \mathbf{V}^{-1}((\mathbf{X}\mathbf{Y})\mathbf{Z})^{-1} = \mathbf{V}^{-1}\mathbf{Z}^{-1}(\mathbf{X}\mathbf{Y})^{-1} = \mathbf{V}^{-1}\mathbf{Z}^{-1}\mathbf{Y}^{-1}\mathbf{X}^{-1}\ .$$

## 5 TRASPOSIZIONE

Data una matrice $\mathbf{X} = [x_{ij}]_{(n,p)}$, la matrice $\mathbf{X}' = [x_{ji}]_{(p,n)}$, ottenuta da $\mathbf{X}$ scambiando le righe con le colonne, è detta la *trasposta* di $\mathbf{X}$.

L'apice indica l'operazione di *trasposizione*.

**ESEMPIO 8.** Data la matrice

$$\mathbf{X} = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 4 \end{bmatrix}$$

risulta

$$\mathbf{X}' = \begin{bmatrix} 1 & 1 & 0 \\ 3 & 2 & 4 \end{bmatrix}. \qquad \blacksquare$$

L'operazione di trasposizione gode di alcune importanti proprietà che vogliamo, adesso, porre in evidenza.

Anzitutto, com'è immediato verificare,

$$(\mathbf{X}')' = \mathbf{X} \quad , \quad (c\mathbf{X})' = c\mathbf{X}' \quad , \quad (\mathbf{X}+\mathbf{Y})' = \mathbf{X}' + \mathbf{Y}' .$$

Inoltre, date due matrici

$$\mathbf{X} = [x_{ij}]_{(n,p)} \quad , \quad \mathbf{Y} = [y_{jk}]_{(p,q)}$$

si ha che

$$(\mathbf{X}\mathbf{Y})' = \mathbf{Y}'\mathbf{X}'$$

come risulta subito dal confronto di

$$(\mathbf{X}\mathbf{Y})' = \begin{bmatrix} \sum_{j=1}^{p} x_{1j}y_{j1} & \cdots & \sum_{j=1}^{p} x_{nj}y_{j1} \\ \cdots & \cdots & \cdots \\ \sum_{j=1}^{p} x_{1j}y_{jq} & \cdots & \sum_{j=1}^{p} x_{nj}y_{jq} \end{bmatrix}$$

e

$$\mathbf{Y}'\mathbf{X}' = \begin{bmatrix} \sum_{j=1}^{p} y_{j1}x_{1j} & \cdots & \sum_{j=1}^{p} y_{j1}x_{nj} \\ \cdots & \cdots & \cdots \\ \sum_{j=1}^{p} y_{jq}x_{1j} & \cdots & \sum_{j=1}^{p} y_{jq}x_{nj} \end{bmatrix}$$

**ESEMPIO 9.** Date le matrici

$$\mathbf{X} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \quad , \quad \mathbf{Y} = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 2 \end{bmatrix}$$

si ha

$$(\mathbf{XY})' = \begin{bmatrix} 2 & 3 \\ 1 & 3 \\ 2 & 2 \end{bmatrix} = \mathbf{Y}'\mathbf{X}' .$$

**OSSERVAZIONE 5.** Tenendo presente l'associatività della moltiplicazione di matrici, la proprietà sopra indicata può essere estesa a un numero qualsiasi, purché finito, di matrici.

Per esempio,

$$(\mathbf{XYZV})' = ((\mathbf{XYZ})\mathbf{V})' = \mathbf{V}'((\mathbf{XY})\mathbf{Z})' = \mathbf{V}'\mathbf{Z}'(\mathbf{XY})' = \mathbf{V}'\mathbf{Z}'\mathbf{Y}'\mathbf{X}' . \qquad \blacksquare$$

Infine, se esiste l'inversa di una matrice $\mathbf{X}$, risulta

$$(\mathbf{X}^{-1})' = (\mathbf{X}')^{-1} .$$

Infatti, dalla duplice identità $\mathbf{X}^{-1}\mathbf{X} = \mathbf{I} = \mathbf{X}\mathbf{X}^{-1}$, eseguendo la trasposizione, si ottiene

$$(\mathbf{X}^{-1}\mathbf{X})' = \mathbf{X}'(\mathbf{X}^{-1})' = \mathbf{I} = (\mathbf{X}\mathbf{X}^{-1})' = (\mathbf{X}^{-1})'\mathbf{X}'$$

e cioè che $(\mathbf{X}^{-1})'$ è l'inversa di $\mathbf{X}'$ la quale si scrive appunto $(\mathbf{X}')^{-1}$.

Da quanto precede è ovvio che, anche quando $\mathbf{X}$ rappresenta una matrice quadrata, in genere $\mathbf{X}' \neq \mathbf{X}$. Tuttavia, il caso in cui si abbia $\mathbf{X}' = \mathbf{X}$ è assai importante e merita una considerazione a parte.

Le matrici, ovviamente quadrate, per le quali sussiste l'eguaglianza con la propria trasposta sono dette *simmetriche* in quanto, com'è facile verificare, hanno eguali gli elementi posti simmetricamente rispetto alla diagonale principale.

In altri termini, data una matrice $\mathbf{X} = [x_{ij}]_{(n,n)}$, $\mathbf{X}' = \mathbf{X}$ se e soltanto se $x_{ij} = x_{ji}$ per $i, j = 1, \dots, n$.

Per esempio, la matrice

$$\mathbf{X} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}$$

è una matrice simmetrica di ordine $(2,2)$.

Si osservi che, in generale, il prodotto di due matrici simmetriche non è una matrice simmetrica. Vale a dire, se $\mathbf{X}' = \mathbf{X}$ e $\mathbf{Y}' = \mathbf{Y}$, allora

$$(\mathbf{XY})' = \mathbf{Y}'\mathbf{X}' = \mathbf{YX}$$

ma, generalmente, $\mathbf{YX} \neq \mathbf{XY}$ e, quindi, $\mathbf{XY}$ non rappresenta una matrice simmetrica.

Vale, invece, l'importante proprietà per cui sia il prodotto di una matrice per la sua trasposta sia il prodotto della trasposta di una matrice per la matrice stessa sono una matrice simmetrica. Infatti,

$$(\mathbf{XX}')' = (\mathbf{X}')'\mathbf{X}' = \mathbf{XX}' \quad , \quad (\mathbf{X}'\mathbf{X})' = \mathbf{X}'(\mathbf{X}')' = \mathbf{X}'\mathbf{X}\,.$$

In genere, $\mathbf{XX}' \neq \mathbf{X}'\mathbf{X}$ ma, se $\mathbf{X}' = \mathbf{X}$, allora $\mathbf{XX}' = \mathbf{X}'\mathbf{X}$.

**ESEMPIO 10.** Data la matrice

$$\mathbf{X} = \begin{bmatrix} -1 & 2 \\ 0 & 1 \\ 3 & 7 \end{bmatrix}$$

si ha

$$\mathbf{XX}' = \begin{bmatrix} -1 & 2 \\ 0 & 1 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} -1 & 0 & 3 \\ 2 & 1 & 7 \end{bmatrix} = \begin{bmatrix} 5 & 2 & 11 \\ 2 & 1 & 7 \\ 11 & 7 & 58 \end{bmatrix},$$

mentre

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} -1 & 0 & 3 \\ 2 & 1 & 7 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 0 & 1 \\ 3 & 7 \end{bmatrix} = \begin{bmatrix} 10 & 19 \\ 19 & 54 \end{bmatrix}. \qquad \blacksquare$$

Infine, sussiste l'ulteriore proprietà secondo la quale se $\mathbf{X}$ è una matrice simmetrica ed esiste l'inversa $\mathbf{X}^{-1}$ anche quest'ultima è una matrice simmetrica. Infatti, poiché

$$(\mathbf{X}^{-1})' = (\mathbf{X}')^{-1} ,$$

se $\mathbf{X}' = \mathbf{X}$, allora

$$(\mathbf{X}^{-1})' = \mathbf{X}^{-1} .$$

## 6 ELEVAZIONE A POTENZA

Data una matrice quadrata $\mathbf{X}$, si chiama *potenza di* $\mathbf{X}$ *con esponente* n, e si scrive $\mathbf{X}^n$, il prodotto di n matrici eguali a $\mathbf{X}$. Si pone poi $\mathbf{X}^0 = \mathbf{I}$.

**ESEMPIO 11.** Data la matrice

$$\mathbf{X} = \begin{bmatrix} 1 & -2 \\ 3 & 7 \end{bmatrix}$$

risulta

$$\mathbf{X}^2 = \begin{bmatrix} 1 & -2 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 3 & 7 \end{bmatrix} = \begin{bmatrix} -5 & -16 \\ 24 & 43 \end{bmatrix}. \qquad \blacksquare$$

Si dimostra facilmente che l'operazione di *elevazione a potenza* possiede le seguenti proprietà

$$\mathbf{X}^{m+n} = \mathbf{X}^m\mathbf{X}^n = \mathbf{X}^n\mathbf{X}^m \quad , \quad \mathbf{X}^{m\,n} = (\mathbf{X}^m)^n = (\mathbf{X}^n)^m$$

quali che siano i numeri naturali m e n (0 compreso).

Se poi sussiste l'inversa di $\mathbf{X}$, posto $\mathbf{X}^{-n} = (\mathbf{X}^{-1})^n$, si ha (Cfr. l'Osservazione 4)

$$\mathbf{X}^{-n} = (\mathbf{X}^{-1})^n = \mathbf{X}^{-1}\mathbf{X}^{-1} \ldots \ \mathbf{X}^{-1} = (\mathbf{X}\mathbf{X} \ldots \mathbf{X})^{-1} = (\mathbf{X}^n)^{-1} .$$

Una particolare rilevanza è assunta dalle cosiddette *matrici idempotenti*, cioè a dire da quelle matrici (quadrate) $\mathbf{X}$ tali che

$$\mathbf{X}^2 = \mathbf{X} .$$

Per esempio, la matrice

$$\mathbf{X} = \begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix}$$

è idempotente, poiché

$$\mathbf{X}^2 = \begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix} = \mathbf{X} .$$

Le matrici idempotenti godono delle seguenti proprietà elementari:

(i) Se $\mathbf{X}$ è idempotente ed esiste la sua inversa $\mathbf{X}^{-1}$, allora $\mathbf{X} = \mathbf{I}$. Infatti, se $\mathbf{X}^2 = \mathbf{X}$, allora

$$\mathbf{I} = \mathbf{X}^{-1}\mathbf{X} = \mathbf{X}^{-1}\mathbf{X}^2 = \mathbf{X}^{-1}\mathbf{X}\mathbf{X} = \mathbf{X} .$$

(ii) Se $\mathbf{X}$ è idempotente, anche $\mathbf{I} - \mathbf{X}$ è idempotente. Infatti (Cfr. l'Osservazione 3),

$$(\mathbf{I} - \mathbf{X})^2 = (\mathbf{I} - \mathbf{X})(\mathbf{I} - \mathbf{X}) = \mathbf{I} - \mathbf{X} - \mathbf{X} + \mathbf{X}^2 = \mathbf{I} - \mathbf{X} .$$

(iii) Se $\mathbf{X}$ e $\mathbf{Y}$ sono idempotenti e $\mathbf{X}\mathbf{Y} = \mathbf{Y}\mathbf{X}$, allora $\mathbf{X}\mathbf{Y}$ è idempotente. Infatti,

$$(\mathbf{X}\mathbf{Y})^2 = \mathbf{X}(\mathbf{Y}\mathbf{X})\mathbf{Y} = \mathbf{X}(\mathbf{X}\mathbf{Y})\mathbf{Y} = \mathbf{X}^2\mathbf{Y}^2 = \mathbf{X}\mathbf{Y} .$$

(iv) La matrice $\mathbf{Z} = \mathbf{X}(\mathbf{Y}\mathbf{X})^{-1}\mathbf{Y}$ è idempotente. Infatti,

$$\mathbf{Z}^2 = \mathbf{X}(\mathbf{Y}\mathbf{X})^{-1}\mathbf{Y}\mathbf{X}(\mathbf{Y}\mathbf{X})^{-1}\mathbf{Y} = \mathbf{X}(\mathbf{Y}\mathbf{X})^{-1}\mathbf{Y} = \mathbf{Z} .$$

(v) La matrice $\mathbf{P} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ è simmetrica e idempotente. Infatti,

$$\mathbf{P}' = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \mathbf{P} \quad , \quad \mathbf{P}^2 = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \mathbf{P} .$$

## 7 TRACCIA

Data una matrice quadrata $\mathbf{X} = [x_{ij}]_{(n,n)}$, si definisce *traccia* di $\mathbf{X}$, e si scrive $\mathrm{tr}\mathbf{X}$, la somma degli elementi posti sulla diagonale principale di $\mathbf{X}$. In simboli,

$$\mathrm{tr}\mathbf{X} = \sum_{i=1}^{n} x_{ii} \, .$$

Tenuto conto della definizione data, si verifica immediatamente che, qualunque siano le matrici

$$\mathbf{X} = [x_{ij}]_{(n,n)} \quad , \quad \mathbf{Y} = [y_{ij}]_{(n,n)}$$

e il numero reale c, risulta

(a) $\mathrm{tr}(\mathbf{X}+\mathbf{Y}) = \sum_{i=1}^{n} (x_{ii} + y_{ii}) = \mathrm{tr}\mathbf{X} + \mathrm{tr}\mathbf{Y}$

(b) $\mathrm{tr}(c\mathbf{X}) = \sum_{i=1}^{n} (cx_{ii}) = c\,\mathrm{tr}\mathbf{X}$

(c) $\mathrm{tr}\mathbf{X} = \sum_{i=1}^{n} x_{ii} = \mathrm{tr}\mathbf{X}'$.

In particolare, $\mathrm{tr}(c\mathbf{I}) = cn$.

Inoltre, date due matrici

$$\mathbf{X} = [x_{ij}]_{(n,p)} \quad , \quad \mathbf{Y} = [y_{ji}]_{(p,n)}$$

tali che entrambi i prodotti $\mathbf{XY}$ e $\mathbf{YX}$ esistono, vale l'importante proprietà

$$\mathrm{tr}(\mathbf{XY}) = \sum_{i=1}^{n} \Big(\sum_{j=1}^{p} x_{ij}y_{ji}\Big) = \sum_{j=1}^{p} \Big(\sum_{i=1}^{n} y_{ji}x_{ij}\Big) = \mathrm{tr}(\mathbf{YX}) \, .$$

In particolare, se $\mathbf{Y} = \mathbf{X}'$, allora

$$\mathrm{tr}(\mathbf{XX}') = \mathrm{tr}(\mathbf{X}'\mathbf{X}) = \sum_{i=1}^{n} \sum_{j=1}^{p} x_{ij}^2 \, .$$

**ESEMPIO 12.** Date le matrici

$$\mathbf{X} = \begin{bmatrix} 1 & 0 \\ -1 & 2 \\ 0 & 3 \end{bmatrix} \quad , \quad \mathbf{Y} = \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$

risulta

$$\mathbf{XY} = \begin{bmatrix} 0 & -1 & -2 \\ 2 & 1 & 0 \\ 3 & 0 & -3 \end{bmatrix} \quad , \quad \mathbf{YX} = \begin{bmatrix} 1 & -8 \\ 1 & -3 \end{bmatrix}$$

e, quindi, $\mathrm{tr}(\mathbf{XY}) = \mathrm{tr}(\mathbf{YX}) = -2$.

## 8 MATRICI A BLOCCHI

**8.1** Data una matrice $\mathbf{X} = [x_{ij}]_{(n,p)}$, consideriamone la consueta rappresentazione mediante il quadro ordinato degli n·p elementi che la compongono.

Supponiamo poi di suddividere la matrice in questione in due o più *blocchi* tracciando un certo numero di linee orizzontali e/o verticali, come nell'esempio che segue:

$$\left[\begin{array}{cc|c} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ \hline x_{41} & x_{42} & x_{43} \end{array}\right].$$

Le matrici

$$\mathbf{X}_{11} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{bmatrix} \quad , \quad \mathbf{X}_{12} = \begin{bmatrix} x_{13} \\ x_{23} \\ x_{33} \end{bmatrix} \quad , \quad \mathbf{X}_{21} = [x_{41}\ x_{42}] \quad , \quad \mathbf{X}_{22} = [x_{43}]$$

in tal modo definite sono dette *sottomatrici* di $\mathbf{X}$ e $\mathbf{X}$ è detta *ripartita* in $\mathbf{X}_{11}, \mathbf{X}_{12}, \mathbf{X}_{21}, \mathbf{X}_{22}$.

Si dice anche che la matrice $\mathbf{X}$ è ottenuta per *accostamento* di $\mathbf{X}_{11}, \mathbf{X}_{12}$, $\mathbf{X}_{21}, \mathbf{X}_{22}$ e si scrive

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} \\ \mathbf{X}_{21} & \mathbf{X}_{22} \end{bmatrix}.$$

Come si può osservare, la matrice **X**, scritta in tale forma, è una matrice i cui elementi sono essi stessi delle matrici.

Tali elementi hanno le seguenti caratteristiche: gli elementi che si trovano su una stessa riga hanno lo stesso numero di righe e, eventualmente, un diverso numero di colonne; gli elementi che si trovano su una stessa colonna hanno invece lo stesso numero di colonne e, eventualmente, un diverso numero di righe.

Ovviamente, quanto ora osservato con riferimento a uno specifico esempio ha carattere del tutto generale.

**OSSERVAZIONE 6.** Data una matrice $\mathbf{X} = [x_{ij}]_{(n,p)}$, torna talvolta utile rappresentare **X** nella forma

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 \cdots \mathbf{x}_p \end{bmatrix}$$

dove $(j = 1, \dots, p)$

$$\mathbf{x}_j = \begin{bmatrix} x_{1j} \\ \vdots \\ x_{nj} \end{bmatrix},$$

oppure nella forma

$$\mathbf{X} = \begin{bmatrix} \boldsymbol{x}'_1 \\ \vdots \\ \boldsymbol{x}'_n \end{bmatrix}$$

dove $(i = 1, \dots, n)$

$$\boldsymbol{x}'_i = \begin{bmatrix} x_{i1} \cdots x_{ip} \end{bmatrix}.$$

È ovvio che, nell'uno e nell'altro caso, siamo di fronte a una particolare ripartizione di **X** in sottomatrici.

**OSSERVAZIONE 7.** Dato uno spazio vettoriale S di ordine n e dimensione p, consideriamo una base di S costituita da p vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$ di ordine n.

Un generico vettore $\mathbf{x} \in S$ può allora essere scritto univocamente nella forma

$$\mathbf{x} = a_1\mathbf{x}_1 + \ldots + a_p\mathbf{x}_p = \begin{bmatrix} \mathbf{x}_1 \cdots \mathbf{x}_p \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix} = \mathbf{X}\mathbf{a} \ .$$

Il vettore **a** di componenti $a_1, \ldots, a_p$ è detto *vettore coordinato* di **x** rispetto alla suddetta base di S.

**8.2** Le operazioni di addizione e sottrazione di matrici (dello stesso ordine) ripartite in blocchi si eseguono secondo le regole già esposte per matrici i cui elementi sono numeri reali, purché le matrici in questione siano state ripartite in sottomatrici mediante una stessa suddivisione.

**8.3** Supponiamo che due matrici **X** e **Y**, conformabili rispetto alla operazione di moltiplicazione, siano ripartite nel modo seguente:

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{11} & \cdots & \mathbf{X}_{1t} \\ \cdots & \cdots & \cdots \\ \mathbf{X}_{v1} & \cdots & \mathbf{X}_{vt} \end{bmatrix} \ , \quad \mathbf{Y} = \begin{bmatrix} \mathbf{Y}_{11} & \cdots & \mathbf{Y}_{1q} \\ \cdots & \cdots & \cdots \\ \mathbf{Y}_{t1} & \cdots & \mathbf{Y}_{tq} \end{bmatrix} .$$

Supponiamo, inoltre, che $\mathbf{X}_{uh}$ e $\mathbf{Y}_{hk}$ $(u = 1, \ldots, v; h = 1, \ldots, t; k = 1, \ldots, q)$ siano tali che il loro prodotto $\mathbf{X}_{uh}\mathbf{Y}_{hk}$ esiste per ogni u e ogni k.

Ne consegue che

$$\mathbf{Z} = \mathbf{X}\mathbf{Y} = \begin{bmatrix} \mathbf{Z}_{uk} \end{bmatrix}_{(v,q)}$$

dove

$$\mathbf{Z}_{uk} = \sum_{h=1}^{t} \mathbf{X}_{uh}\mathbf{Y}_{hk} \ .$$

**ESEMPIO 13.** Date le matrici

$$\mathbf{X}_{11} = \begin{bmatrix} 0 & 1 \\ 4 & 1 \end{bmatrix} \ , \quad \mathbf{X}_{12} = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \ , \quad \mathbf{Y}_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \ , \ \mathbf{Y}_{21} = \begin{bmatrix} 1 & 2 \end{bmatrix}$$

e costruite le matrici a blocchi

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} \end{bmatrix} \quad , \quad \mathbf{Y} = \begin{bmatrix} \mathbf{Y}_{11} \\ \mathbf{Y}_{21} \end{bmatrix}$$

risulta

$$\mathbf{XY} = \mathbf{X}_{11}\mathbf{Y}_{11} + \mathbf{X}_{12}\mathbf{Y}_{21} = \begin{bmatrix} 0 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} + \begin{bmatrix} 3 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 9 \\ 6 & 7 \end{bmatrix}.$$

# Cap. II
# COMPLEMENTI

**1**

Con un ragionamento del tutto analogo a quello svolto ai punti 1 e 2 dei Complementi al Cap. I, si dimostrano le proprietà seguenti:

a) $\mathbf{X} + \mathbf{Z} = \mathbf{Y} + \mathbf{Z}$ se e soltanto se $\mathbf{X} = \mathbf{Y}$ (*proprietà di cancellazione*).

b) Se $c\mathbf{X} = \mathbf{O}$, allora è $c = 0$ oppure $\mathbf{X} = \mathbf{O}$.

c) Se $c\mathbf{X} = c\mathbf{Y}$, $c \neq 0$, allora $\mathbf{X} = \mathbf{Y}$.

**2**

Come si verifica facilmente, il prodotto di una matrice **X** per una matrice diagonale **D** è una matrice le cui colonne sono quelle di **X** ciascuna moltiplicata per il corrispondente elemento diagonale di **D**.

Analogamente, il prodotto di una matrice diagonale **D** per una matrice **X** è una matrice le cui righe sono quelle di **X** ciascuna moltiplicata per il corrispondente elemento diagonale di **D**.

**3**

Date le matrici

$$\mathbf{X} = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & 0 \\ -1 & 4 & 0 \end{bmatrix} \quad , \quad \mathbf{Y} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 4 & 9 \end{bmatrix}$$

risulta

$$\mathbf{XY} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

pur essendo $\mathbf{X} \neq \mathbf{O}$, $\mathbf{Y} \neq \mathbf{O}$.

Date le matrici

$$\mathbf{X} = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & 0 \\ -1 & 4 & 0 \end{bmatrix} \quad , \quad \mathbf{Y} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & -1 \\ 2 & 2 & 2 \end{bmatrix} \quad , \quad \mathbf{Z} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & -1 \\ 1 & 2 & 1 \end{bmatrix}$$

risulta

$$\mathbf{XY} = \begin{bmatrix} 3 & 4 & 1 \\ 2 & 3 & 2 \\ 3 & 2 & -7 \end{bmatrix} = \mathbf{XZ}$$

pur essendo $\mathbf{Y} \neq \mathbf{Z}$ e $\mathbf{X} \neq \mathbf{O}$.

**4**

Date due matrici

$$\mathbf{A} = [a_{ij}]_{(n,p)} \quad , \quad \mathbf{X} = [x_{hk}]_{(m,q)}$$

si chiama *prodotto tensoriale* (o *prodotto di Kronecker*) di $\mathbf{A}$ per $\mathbf{X}$ la matrice di ordine $(nm, pq)$, che indichiamo con $\mathbf{A}\otimes\mathbf{X}$, definita da

$$\mathbf{A}\otimes\mathbf{X} = \begin{bmatrix} a_{11}\mathbf{X} & \cdots & a_{1p}\mathbf{X} \\ \cdots & \cdots & \cdots \\ a_{n1}\mathbf{X} & \cdots & a_{np}\mathbf{X} \end{bmatrix}.$$

Tenuto conto di tale definizione, è possibile dimostrare, con relativa facilità, che l'operazione di *moltiplicazione tensoriale* gode delle seguenti proprietà elementari (1):

(1) Si suppone, ovviamente, che le matrici considerate siano di ordine tale che le operazioni di addizione, moltiplicazione, inversione, traccia – ove ricorrano – siano eseguibili.

(1) $ab(\mathbf{X}\otimes\mathbf{Y}) = (a\mathbf{X})\otimes(b\mathbf{Y}) = (ab\mathbf{X})\otimes\mathbf{Y} = \mathbf{X}\otimes(ab\mathbf{Y})$

(2) $\mathbf{A}\otimes(\mathbf{B}\otimes\mathbf{X}) = (\mathbf{A}\otimes\mathbf{B})\otimes\mathbf{X}$

(3) $(\mathbf{A}\otimes\mathbf{X})(\mathbf{B}\otimes\mathbf{Y}) = (\mathbf{AB})\otimes(\mathbf{XY})$

(4) $\mathbf{A}\otimes(\mathbf{B} + \mathbf{X}) = (\mathbf{A}\otimes\mathbf{B}) + (\mathbf{A}\otimes\mathbf{X})$

(5) $(\mathbf{A} + \mathbf{B})\otimes\mathbf{X} = (\mathbf{A}\otimes\mathbf{X}) + (\mathbf{B}\otimes\mathbf{X})$

(6) $(\mathbf{A}\otimes\mathbf{X})' = \mathbf{A}'\otimes\mathbf{X}'$

(7) $(\mathbf{A}\otimes\mathbf{X})^{-1} = \mathbf{A}^{-1}\otimes\mathbf{X}^{-1}$

(8) $\mathrm{tr}(\mathbf{A}\otimes\mathbf{X}) = (\mathrm{tr}\mathbf{A})(\mathrm{tr}\mathbf{X})$.

**5**

Considerata una matrice $\mathbf{X} = \left[\, \mathbf{x}_1 \cdots \mathbf{x}_p \,\right]$ di ordine $(n\,,p)$, si definisce $\mathrm{vec}\mathbf{X}$ il vettore colonna di ordine $(np\,,1)$ avente come elementi $\mathbf{x}_1\,,\dots\,,\mathbf{x}_p$. In simboli,

$$\mathrm{vec}\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_p \end{bmatrix}.$$

Tenuto conto di tale definizione, è possibile dimostrare, con relativa facilità, che l'*operatore vec* gode delle seguenti proprietà elementari (2):

(1) $\mathrm{vec}(a\mathbf{X} + b\mathbf{Y}) = a(\mathrm{vec}\,\mathbf{X}) + b(\mathrm{vec}\mathbf{Y})$

(2) $\mathrm{vec}(\mathbf{XYZ}) = (\mathbf{Z}'\otimes\mathbf{X})(\mathrm{vec}\,\mathbf{Y})$

(3) $\mathbf{XYZ} = \left[(\mathrm{vec}\mathbf{Z})'\otimes\mathbf{X}\right]\left[\mathbf{I}\otimes(\mathrm{vec}\mathbf{Y})\right]$

(4) $\mathrm{tr}(\mathbf{XY}) = (\mathrm{vec}\mathbf{X}')'(\mathrm{vec}\mathbf{Y})$.

**6**

Si riconosce immediatamente che l'insieme delle matrici di ordine $(n\,,p)$, con le operazioni di addizione e di moltiplicazione per un numero reale date in precedenza, è uno spazio vettoriale su R nel senso della definizione di cui al punto 6 dei Complementi al Cap. I.

---

(2) Si suppone, ovviamente, che le matrici considerate siano di ordine tale che le operazioni di addizione, moltiplicazione e traccia – ove ricorrano – siano eseguibili.

Una base di tale spazio vettoriale è formata dalle n·p matrici di ordine (n , p)

$$\mathbf{U}_{11}, \dots, \mathbf{U}_{1p}, \dots, \mathbf{U}_{n1}, \dots, \mathbf{U}_{np}$$

ciascuna delle quali è costituita da elementi tutti nulli tranne l'elemento $u_{ij}$ posto eguale a 1.

Naturalmente, la dimensione dello spazio vettoriale in questione è n·p.

# Cap. III
# DETERMINANTI

## 1 FUNZIONE DETERMINANTE

**1.1** Considerato l'insieme $M_n$ di tutte le matrici quadrate di ordine $(n,n)$, si definisce *funzione determinante* ogni funzione che associa a ciascun elemento $\mathbf{X} = [\mathbf{x}_1 \cdots \mathbf{x}_j \cdots \mathbf{x}_n]$ di $M_n$ un numero reale – che si chiama *determinante* di $\mathbf{X}$ e si indica con $\det\mathbf{X}$ – in modo tale che siano soddisfatte le proprietà seguenti :

$(a_1)$ Per ogni $c \in R$, risulta

$$\det[\mathbf{x}_1 \cdots c\mathbf{x}_j \cdots \mathbf{x}_n] = c\det[\mathbf{x}_1 \cdots \mathbf{x}_j \cdots \mathbf{x}_n] .$$

$(a_2)$ Se $\mathbf{x}_j = \mathbf{y}_j + \mathbf{z}_j$, si ha

$$\det[\mathbf{x}_1 \cdots \mathbf{y}_j + \mathbf{z}_j \cdots \mathbf{x}_n] = \det[\mathbf{x}_1 \cdots \mathbf{y}_j \cdots \mathbf{x}_n] + \det[\mathbf{x}_1 \cdots \mathbf{z}_j \cdots \mathbf{x}_n] .$$

$(b)$ Se $\mathbf{x}_h = \mathbf{x}_j$ $(h \neq j)$, allora

$$\det[\mathbf{x}_1 \cdots \mathbf{x}_h \cdots \mathbf{x}_j \cdots \mathbf{x}_n] = 0 .$$

$(c)$ Se $\mathbf{I} = [\mathbf{u}_1 \cdots \mathbf{u}_n]$ rappresenta la matrice unità di ordine $(n,n)$, allora

$$\det[\mathbf{u}_1 \cdots \mathbf{u}_n] = 1 .$$

Le proprietà $(a_1)$ e $(a_2)$, considerate congiuntamente, esprimono la cosiddetta *linearità* di una funzione determinante rispetto a ciascuno dei vettori colonna della matrice **X**.

Si noti che, posto

$$\mathbf{x}_j = \sum_{k=1}^{t} c_k \mathbf{z}_k ,$$

dalla proprietà di linearità segue che

$$\det\left[\mathbf{x}_1 \cdots \sum_{k=1}^{t} c_k \mathbf{z}_k \cdots \mathbf{x}_n\right] = \sum_{k=1}^{t} c_k \det[\mathbf{x}_1 \cdots \mathbf{z}_k \cdots \mathbf{x}_n] .$$

**ESEMPIO 1.** Sia

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = [\mathbf{x}_1 \ \ \mathbf{x}_2]$$

un generico elemento dell'insieme $M_2$ di tutte le matrici quadrate di ordine $(2,2)$.

Proponiamo quale funzione determinante su $M_2$ la funzione definita da

$$\det[\mathbf{x}_1 \ \ \mathbf{x}_2] = x_{11}x_{22} - x_{12}x_{21} .$$

Vogliamo dimostrare che valgono per la suddetta funzione le proprietà a cui deve soddisfare una funzione determinante.

$(a_1)$ $\det[c\mathbf{x}_1 \ \ \mathbf{x}_2] = (cx_{11})x_{22} - x_{12}(cx_{21}) = c(x_{11}x_{22} - x_{12}x_{21})$
$= c\det[\mathbf{x}_1 \ \ \mathbf{x}_2]$ .

Analogamente, $\det[\mathbf{x}_1 \ \ c\mathbf{x}_2] = c\det[\mathbf{x}_1 \ \ \mathbf{x}_2]$.

$(a_2)$ $\det[\mathbf{y}_1+\mathbf{z}_1 \ \ \mathbf{x}_2] = (y_{11}+z_{11})x_{22} - x_{12}(y_{21}+z_{21})$
$= y_{11}x_{22} - x_{12}y_{21} + z_{11}x_{22} - x_{12}z_{21}$
$= \det[\mathbf{y}_1 \ \ \mathbf{x}_2] + \det[\mathbf{z}_1 \ \ \mathbf{x}_2]$ .

Analogamente, $\det[\mathbf{x}_1 \ \ \mathbf{y}_2+\mathbf{z}_2] = \det[\mathbf{x}_1 \ \ \mathbf{y}_2] + \det[\mathbf{x}_1 \ \ \mathbf{z}_2]$.

(b) $\det[\mathbf{x}_1 \ \ \mathbf{x}_1] = x_{11}x_{21} - x_{11}x_{21} = 0 = x_{12}x_{22} - x_{12}x_{22} = \det[\mathbf{x}_2 \ \ \mathbf{x}_2]$ .

(c) $\det[\mathbf{u}_1 \ \ \mathbf{u}_2] = (1)(1) - (0)(0) = 1$ .

**1.2** Nel paragrafo successivo dimostreremo che, qualunque sia l'insieme $M_n$, una funzione determinante esiste ed è unica.

Il nostro obiettivo adesso è porre in evidenza alcune proprietà di una funzione determinante, proprietà che discendono direttamente da quelle indicate all'inizio della sezione precedente.

(d) Se si addiziona al j-esimo vettore colonna di $\left[\mathbf{x}_1 \cdots \mathbf{x}_j \cdots \mathbf{x}_n\right]$ una combinazione lineare dei rimanenti vettori colonna, il determinante non cambia. In simboli,

$$\det\left[\mathbf{x}_1 \cdots \mathbf{x}_j + \sum_{k \neq j} c_k \mathbf{x}_k \cdots \mathbf{x}_n\right] = \det\left[\mathbf{x}_1 \cdots \mathbf{x}_j \cdots \mathbf{x}_n\right].$$

Infatti, per la linearità di una funzione determinante, il primo membro della espressione precedente può essere scritto nella forma

$$\det\left[\mathbf{x}_1 \cdots \mathbf{x}_j \cdots \mathbf{x}_n\right] + \sum_{k \neq j} c_k \det\left[\mathbf{x}_1 \cdots \mathbf{x}_k \cdots \mathbf{x}_n\right]$$

e ciascun termine della sommatoria al secondo membro, per la proprietà (b), vale zero.

(e) Se gli n vettori colonna di $\left[\mathbf{x}_1 \cdots \mathbf{x}_j \cdots \mathbf{x}_n\right]$ sono linearmente dipendenti, allora [(1)]

$$\det\left[\mathbf{x}_1 \cdots \mathbf{x}_j \cdots \mathbf{x}_n\right] = 0 .$$

Infatti, per $n = 1$,

$$\det[0] = \det[0x] = 0\det[x] = 0 .$$

Per $n > 1$, almeno un vettore, diciamo $\mathbf{x}_j$, può essere espresso come combinazione lineare dei rimanenti; cioè,

$$\mathbf{x}_j = \sum_{k \neq j} c_k \mathbf{x}_k .$$

Quindi,

(1) Vale anche, come dimostreremo al punto 1 dei Complementi al Cap. IV, l'affermazione reciproca.

$$\det\left[\mathbf{x}_1 \cdots \mathbf{x}_j \cdots \mathbf{x}_n\right] = \det\left[\mathbf{x}_1 \cdots \sum_{k\neq j} c_k \mathbf{x}_k \cdots \mathbf{x}_n\right] = \sum_{k\neq j} c_k \det\left[\mathbf{x}_1 \cdots \mathbf{x}_k \cdots \mathbf{x}_n\right] = 0 .$$

(f) Se $\left[\mathbf{x}_1 \cdots \mathbf{x}_h \cdots \mathbf{x}_j \cdots \mathbf{x}_n\right]$ è ottenuta da $\left[\mathbf{x}_1 \cdots \mathbf{x}_j \cdots \mathbf{x}_h \cdots \mathbf{x}_n\right]$ scambiando $\mathbf{x}_h$ con $\mathbf{x}_j$ ($h\neq j$), allora

$$\det\left[\mathbf{x}_1 \cdots \mathbf{x}_j \cdots \mathbf{x}_h \cdots \mathbf{x}_n\right] = -\det\left[\mathbf{x}_1 \cdots \mathbf{x}_h \cdots \mathbf{x}_j \cdots \mathbf{x}_n\right] .$$

Infatti,

$$\begin{aligned}
\det\left[\mathbf{x}_1 \cdots \mathbf{x}_j \cdots \mathbf{x}_h \cdots \mathbf{x}_n\right] &= -\det\left[\mathbf{x}_1 \cdots \mathbf{x}_j \cdots (-\mathbf{x}_h) \cdots \mathbf{x}_n\right] \\
&= -\det\left[\mathbf{x}_1 \cdots \mathbf{x}_j \cdots (\mathbf{x}_j - \mathbf{x}_h) \cdots \mathbf{x}_n\right] \\
&= -\det\left[\mathbf{x}_1 \cdots \mathbf{x}_j - (\mathbf{x}_j - \mathbf{x}_h) \cdots (\mathbf{x}_j - \mathbf{x}_h) \cdots \mathbf{x}_n\right] \\
&= -\det\left[\mathbf{x}_1 \cdots \mathbf{x}_h \cdots (\mathbf{x}_j - \mathbf{x}_h) \cdots \mathbf{x}_n\right] \\
&= -\det\left[\mathbf{x}_1 \cdots \mathbf{x}_h \cdots \mathbf{x}_j \cdots \mathbf{x}_n\right] .
\end{aligned}$$

**ESEMPIO 2.** Proseguendo nell'Esempio 1, verifichiamo le ulteriori proprietà (d),(e),(f).

(d)
$$\begin{aligned}
\det\left[\mathbf{x}_1 + c\mathbf{x}_2 \;\; \mathbf{x}_2\right] &= (x_{11} + cx_{12})x_{22} - x_{12}(x_{21} + cx_{22}) \\
&= x_{11}x_{22} - x_{12}x_{21} + cx_{12}x_{22} - cx_{12}x_{22} \\
&= \det\left[\mathbf{x}_1 \;\; \mathbf{x}_2\right] .
\end{aligned}$$

Analogamente, $\det\left[\mathbf{x}_1 \;\; \mathbf{x}_2 + c\mathbf{x}_1\right] = \det\left[\mathbf{x}_1 \;\; \mathbf{x}_2\right]$.

(e) Posto $\mathbf{x}_1 = c\mathbf{x}_2$, si ha

$$\det\left[c\mathbf{x}_2 \;\; \mathbf{x}_2\right] = (cx_{12})x_{22} - x_{12}(cx_{22}) = 0 .$$

(f)
$$\det\left[\mathbf{x}_2 \;\; \mathbf{x}_1\right] = x_{12}x_{21} - x_{11}x_{22} = -\det\left[\mathbf{x}_1 \;\; \mathbf{x}_2\right] .$$

## 2 ESISTENZA E UNICITÀ DI UNA FUNZIONE DETERMINANTE

**2.1** Dimostreremo, anzitutto, l'*esistenza* di una funzione determinante utilizzando il procedimento di induzione (2).

(2) I dettagli delle dimostrazioni contenute in questo paragrafo possono essere omessi a una prima lettura.

Per $n=1$, il dominio di una funzione determinante è costituito dall'insieme di tutte le matrici di ordine $(1,1)$. Posto, per ogni $x \in R$, $\det[x] = x$, è subito visto che valgono le proprietà $(a_1),(a_2),(b),(c)$.

Per $n>1$, assumiamo l'esistenza di una funzione determinante su $M_{n-1}$, vale a dire sull'insieme di tutte le matrici quadrate di ordine $(n-1, n-1)$. Fissato l'indice di riga i $(1 \le i \le n)$, sia

$$(1) \qquad \det\mathbf{X} = (-1)^{i+1} x_{i1} \det\mathbf{X}_{i1} + \ldots + (-1)^{i+n} x_{in} \det\mathbf{X}_{in}$$

dove $\mathbf{X}_{ik}$ $(k = 1, \ldots, n)$ è la matrice di ordine $(n-1, n-1)$ ottenuta da $\mathbf{X} = [x_{ij}]_{(n,n)}$ cancellando la riga i-esima e la colonna k-esima.

Vogliamo dimostrare che la (1) soddisfa le proprietà che caratterizzano una funzione determinante.

$(a_1)$ Dalla (1) si ha che

$$\det[\mathbf{x}_1 \cdots c\mathbf{x}_j \cdots \mathbf{x}_n] = (-1)^{i+1} x_{i1} \det\mathbf{X}^*_{i1} + \ldots + (-1)^{i+j} c x_{ij} \det\mathbf{X}^*_{ij} \\ + \ldots + (-1)^{i+n} x_{in} \det\mathbf{X}^*_{in}$$

dove $\mathbf{X}^*_{ik}$ è la matrice ottenuta da $[\mathbf{x}_1 \cdots c\mathbf{x}_j \cdots \mathbf{x}_n]$ cancellando la riga i-esima e la colonna k-esima.

Per ipotesi di induzione, abbiamo che

$$\det\mathbf{X}^*_{ik} = c\det\mathbf{X}_{ik} \quad \text{per } k \ne j \quad , \quad \det\mathbf{X}^*_{ik} = \det\mathbf{X}_{ik} \quad \text{per } k=j\ .$$

Quindi,

$$\det[\mathbf{x}_1 \cdots c\mathbf{x}_j \cdots \mathbf{x}_n] = c\{(-1)^{i+1} x_{i1} \det\mathbf{X}_{i1} + \ldots + (-1)^{i+j} x_{ij} \det\mathbf{X}_{ij} \\ + \ldots + (-1)^{i+n} x_{in} \det\mathbf{X}_{in}\} \\ = c\det[\mathbf{x}_1 \cdots \mathbf{x}_j \cdots \mathbf{x}_n]$$

e la proprietà $(a_1)$ risulta soddisfatta.

$(a_2)$ Posto $\mathbf{x}_j = \mathbf{y}_j + \mathbf{z}_j$, indichiamo con $\overline{\mathbf{X}}_{ik}$ la matrice ottenuta da $[\mathbf{x}_1 \cdots \mathbf{y}_j \cdots \mathbf{x}_n]$ cancellando la riga i-esima e la colonna k-esima; analogamente, indichiamo con $\widetilde{\mathbf{X}}_{ik}$ la matrice ottenuta da $[\mathbf{x}_1 \cdots \mathbf{z}_j \cdots \mathbf{x}_n]$ cancellando la riga i-esima e la colonna k-esima.

Per ipotesi di induzione, risulta

$$\det \mathbf{X}_{ik} = \det \overline{\mathbf{X}}_{ik} + \det \widetilde{\mathbf{X}}_{ik} \quad \text{per } k \neq j\,,$$
$$\det \mathbf{X}_{ik} = \det \overline{\mathbf{X}}_{ik} = \det \widetilde{\mathbf{X}}_{ik} \quad \text{per } k = j\,.$$

Pertanto, dalla (1) si ha

$$\begin{aligned}
&\det\left[\mathbf{x}_1 \cdots \mathbf{y}_j + \mathbf{z}_j \cdots \mathbf{x}_n\right] = (-1)^{i+1} x_{i1}\ \{\det \overline{\mathbf{X}}_{i1} + \det \widetilde{\mathbf{X}}_{i1}\} \\
&+ \ldots + (-1)^{i+j}(y_{ij} + z_{ij})\ \det \mathbf{X}_{ij} + \ldots + (-1)^{i+n} x_{in}\ \{\det \overline{\mathbf{X}}_{in} + \det \widetilde{\mathbf{X}}_{in}\} \\
&= \{(-1)^{i+1} x_{i1}\ \det \overline{\mathbf{X}}_{i1} + \ldots + (-1)^{i+j} y_{ij}\ \det \overline{\mathbf{X}}_{ij} + \ldots + (-1)^{i+n} x_{in}\ \det \overline{\mathbf{X}}_{in}\} \\
&+ \{(-1)^{i+1} x_{i1}\ \det \widetilde{\mathbf{X}}_{i1} + \ldots + (-1)^{i+j} z_{ij}\ \det \widetilde{\mathbf{X}}_{ij} + \ldots + (-1)^{i+n} x_{in}\ \det \widetilde{\mathbf{X}}_{in}\} \\
&= \det\left[\mathbf{x}_1 \cdots \mathbf{y}_j \cdots \mathbf{x}_n\right] + \det\left[\mathbf{x}_1 \cdots \mathbf{z}_j \cdots \mathbf{x}_n\right].
\end{aligned}$$

e la proprietà $(a_2)$ risulta soddisfatta.

(b) Supponiamo che sia $\mathbf{x}_h = \mathbf{x}_j$ con $h<j$. Per ogni $k \neq h$ e $k \neq j$, la matrice $\mathbf{X}_{ik}$ ha due colonne eguali e quindi, per ipotesi di induzione, $\det \mathbf{X}_{ik} = 0$.

Pertanto, dalla (1) si ha

$$\det\left[\mathbf{x}_1 \cdots \mathbf{x}_h \cdots \mathbf{x}_j \cdots \mathbf{x}_n\right] = (-1)^{i+h} x_{ih}\ \det \mathbf{X}_{ih} + (-1)^{i+j} x_{ij}\ \det \mathbf{X}_{ij}\,.$$

Dobbiamo dimostrare che il secondo membro della espressione precedente vale zero.

A questo fine, si osservi intanto che $x_{ih} = x_{ij}$.

Inoltre, scambiando in sequenza la $(j-1)$-esima colonna di $\mathbf{X}_{ik}$ con ciascuna delle $j-h-1$ colonne che la precedono, si passa da $\mathbf{X}_{ih}$ a $\mathbf{X}_{ij}$.

Ricordando che, valendo per ipotesi di induzione la proprietà (f), a ciascun scambio muta il segno del determinante, si ha

$$\det \mathbf{X}_{ij} = (-1)^{j-h-1} \det \mathbf{X}_{ih} = (-1)^{-j+h+1} \det \mathbf{X}_{ih}$$

da cui

$$\begin{aligned}
\det\left[\mathbf{x}_1 \cdots \mathbf{x}_h \cdots \mathbf{x}_j \cdots \mathbf{x}_n\right] &= (-1)^{i+h} x_{ih} \det \mathbf{X}_{ih} + (-1)^{i+j} x_{ih} (-1)^{-j+h+1} \det \mathbf{X}_{ih} \\
&= (-1)^{i+h} x_{ih} \det \mathbf{X}_{ih} + (-1)^{i+h+1} x_{ih} \det \mathbf{X}_{ih} \\
&= 0
\end{aligned}$$

e la proprietà (b) risulta dimostrata.

(c) La (1) fornisce direttamente

$$\det \mathbf{I} = (-1)^{i+i}(1)\det \mathbf{I}_{ii} = 1 \ .$$

**2.2** Passando adesso a esaminare l'*unicità* di una funzione determinante, consideriamo su $M_n$ due funzioni determinanti $\det_1$ e $\det_2$. Dovendo valere per entrambe la proprietà (c), risulta

$$\det_1 \mathbf{I} = 1 = \det_2 \mathbf{I} \ .$$

Inoltre, indicata con $\mathbf{S} = [\mathbf{u}_1 \cdots \mathbf{u}_j \cdots \mathbf{u}_h \cdots \mathbf{u}_n]$ la matrice che si ottiene da $\mathbf{I} = [\mathbf{u}_1 \cdots \mathbf{u}_h \cdots \mathbf{u}_j \cdots \mathbf{u}_n]$ scambiando i vettori $\mathbf{u}_h$ e $\mathbf{u}_j$ ($h \neq j$), per la proprietà (f) si ha

$$\det_1 \mathbf{S} = -1 = \det_2 \mathbf{S} \ .$$

Infine, sia $\mathbf{P}$ la matrice ottenuta da $\mathbf{I}$ attraverso una permutazione dei vettori $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Poiché $\mathbf{P}$ può essere vista come la matrice ottenuta da $\mathbf{I}$ attraverso una successione finita di scambi tra coppie di vettori e poiché, d'altra parte, a ogni scambio muta simultaneamente il segno di $\det_1 \mathbf{P}$ e $\det_2 \mathbf{P}$, fermo restando il valore assoluto eguale a 1, ne consegue che

$$\det_1 \mathbf{P} = \det_2 \mathbf{P} \ .$$

Ciò premesso, consideriamo il j-esimo ($j = 1, \ldots, n$) vettore colonna di $\mathbf{X}$ espresso nella forma

$$\mathbf{x}_j = x_{1j}\mathbf{u}_1 + \ldots + x_{nj}\mathbf{u}_n = \sum_{i=1}^{n} x_{ij}\mathbf{u}_i \ .$$

Allora, come si può verificare,

$$(*) \quad \det_1[\mathbf{x}_1 \cdots \mathbf{x}_n] = \det_1\left[\sum_{\sigma(1)=1}^{n} x_{\sigma(1)1}\,\mathbf{u}_{\sigma(1)} \cdots \sum_{\sigma(n)=1}^{n} x_{\sigma(n)n}\,\mathbf{u}_{\sigma(n)}\right]$$
$$= \sum_{\sigma(1)=1}^{n} \ldots \sum_{\sigma(n)=1}^{n} x_{\sigma(1)1} \ldots x_{\sigma(n)n} \det_1[\mathbf{u}_{\sigma(1)} \cdots \mathbf{u}_{\sigma(n)}] \ .$$

Ora, degli $n^n$ termini che costituiscono la somma scritta all'ultimo membro di questa espressione ce ne sono $n!$ nei quali i vettori $\mathbf{u}_{\sigma(1)}, \dots, \mathbf{u}_{\sigma(n)}$ sono una permutazione dei vettori $\mathbf{u}_1, \dots, \mathbf{u}_n$.

Per essi, quindi, ricordando quanto detto in precedenza, vale la relazione

$$\det_1\left[\mathbf{u}_{\sigma(1)} \cdots \mathbf{u}_{\sigma(n)}\right] = \det_2\left[\mathbf{u}_{\sigma(1)} \cdots \mathbf{u}_{\sigma(n)}\right].$$

Nei rimanenti $n^n - n!$ termini, la matrice $\left[\mathbf{u}_{\sigma(1)} \cdots \mathbf{u}_{\sigma(n)}\right]$ ha due colonne eguali. Per questi dunque, in virtù della proprietà (b) di una funzione determinante,

$$\det_1\left[\mathbf{u}_{\sigma(1)} \cdots \mathbf{u}_{\sigma(n)}\right] = 0 = \det_2\left[\mathbf{u}_{\sigma(1)} \cdots \mathbf{u}_{\sigma(n)}\right].$$

Possiamo, pertanto, scrivere l'ultimo membro della (∗) nella forma

$$\sum_{\sigma(1)=1}^{n} \dots \sum_{\sigma(n)=1}^{n} x_{\sigma(1)1} \dots x_{\sigma(n)n} \det_2\left[\mathbf{u}_{\sigma(1)} \cdots \mathbf{u}_{\sigma(n)}\right].$$

Ma, quest'ultima espressione rappresenta, com'è immediato riconoscere, $\det_2\left[\mathbf{x}_1 \cdots \mathbf{x}_n\right]$.

Quindi,

$$\det_1\left[\mathbf{x}_1 \cdots \mathbf{x}_n\right] = \det_2\left[\mathbf{x}_1 \cdots \mathbf{x}_n\right].$$

## 3 CALCOLO DI UN DETERMINANTE

**3.1** Si osservi anzitutto che l'espressione $(1 \le i \le n; n \ge 2)$

$$\det\mathbf{X} = (-1)^{i+1} x_{i1} \det\mathbf{X}_{i1} + \dots + (-1)^{i+n} x_{in} \det\mathbf{X}_{in},$$

considerata nel paragrafo precedente, fornisce un algoritmo per il calcolo del determinante di una matrice $\mathbf{X} = \left[x_{ij}\right]_{(n,n)}$ in funzione dei determinanti delle matrici $\mathbf{X}_{ij}$ ottenute da $\mathbf{X}$ cancellando la riga i-esima $(1 \le i \le n)$ e la colonna j-esima $(j = 1, \dots, n)$.

Si dice, in tal caso, che si è proceduto allo *sviluppo del determinante di* $\mathbf{X}$ *secondo gli elementi della riga i-esima.*

Tale sviluppo, per l'unicità della funzione determinante, non dipende dalla riga considerata.

I termini $(-1)^{i+j} \det \mathbf{X}_{ij}$ $(i,j = 1, \dots, n)$ sono chiamati *cofattori* o *complementi algebrici* di $[x_{ij}]$.

**ESEMPIO 3.** Data una matrice

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix}$$

si ottiene

$$\begin{aligned}
\det \mathbf{X} &= + x_{11} \det \begin{bmatrix} x_{22} & x_{23} \\ x_{32} & x_{33} \end{bmatrix} - x_{12} \det \begin{bmatrix} x_{21} & x_{23} \\ x_{31} & x_{33} \end{bmatrix} + x_{13} \det \begin{bmatrix} x_{21} & x_{22} \\ x_{31} & x_{32} \end{bmatrix} \\
&= - x_{21} \det \begin{bmatrix} x_{12} & x_{13} \\ x_{32} & x_{33} \end{bmatrix} + x_{22} \det \begin{bmatrix} x_{11} & x_{13} \\ x_{31} & x_{33} \end{bmatrix} - x_{23} \det \begin{bmatrix} x_{11} & x_{12} \\ x_{31} & x_{32} \end{bmatrix} \\
&= + x_{31} \det \begin{bmatrix} x_{12} & x_{13} \\ x_{22} & x_{23} \end{bmatrix} - x_{32} \det \begin{bmatrix} x_{11} & x_{13} \\ x_{21} & x_{23} \end{bmatrix} + x_{33} \det \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \\
&= + x_{11} x_{22} x_{33} + x_{12} x_{23} x_{31} + x_{13} x_{21} x_{32} - x_{11} x_{23} x_{32} - x_{13} x_{22} x_{31} - x_{12} x_{21} x_{33} .
\end{aligned}$$

**3.2** Ricordando la (*) del paragrafo precedente, possiamo scrivere

$$\det[\mathbf{x}_1 \cdots \mathbf{x}_n] = \sum_{\sigma(1)=1}^{n} \dots \sum_{\sigma(n)=1}^{n} x_{\sigma(1)1} \dots x_{\sigma(n)n} \det[\mathbf{u}_{\sigma(1)} \cdots \mathbf{u}_{\sigma(n)}]$$

e osservare che vale l'identità

$$\begin{aligned}
&\sum_{\sigma(1)=1}^{n} \dots \sum_{\sigma(n)=1}^{n} x_{\sigma(1)1} \dots x_{\sigma(n)n} \det[\mathbf{u}_{\sigma(1)} \cdots \mathbf{u}_{\sigma(n)}] \\
= &\sum_{\sigma(1)=1}^{n} \dots \sum_{\sigma(n)=1}^{n} x_{1\sigma(1)} \dots x_{n\sigma(n)} \det[\mathbf{u}_{\sigma(1)} \cdots \mathbf{u}_{\sigma(n)}] .
\end{aligned}$$

Ma lo scambio dei primi con i secondi indici attuato nella espressione precedente equivale allo scambio delle righe con le colonne di **X**.

Ne consegue che il *determinante di* **X** *è eguale al determinante della trasposta di* **X** *e che tutte le proprietà che competono alla funzione determinante considerata come funzione dei vettori colonna di* **X** *sono valide anche per i vettori riga di* **X**.

In particolare, si ha che $(1 \leq j \leq n; n \geq 2)$

$$\det \mathbf{X} = (-1)^{1+j} x_{1j} \det \mathbf{X}_{1j} + \ldots + (-1)^{n+j} x_{nj} \det \mathbf{X}_{nj} . \tag{2}$$

La (2) costituisce lo *sviluppo del determinante di* **X** *secondo gli elementi della colonna j-esima*.

**ESEMPIO 4.** Data la matrice **X** considerata nell'Esempio 3, risulta

$$\begin{aligned}
\det \mathbf{X} &= + x_{11} \det \begin{bmatrix} x_{22} & x_{23} \\ x_{32} & x_{33} \end{bmatrix} - x_{21} \det \begin{bmatrix} x_{12} & x_{13} \\ x_{32} & x_{33} \end{bmatrix} + x_{31} \det \begin{bmatrix} x_{12} & x_{13} \\ x_{22} & x_{23} \end{bmatrix} \\
&= - x_{12} \det \begin{bmatrix} x_{21} & x_{23} \\ x_{31} & x_{33} \end{bmatrix} + x_{22} \det \begin{bmatrix} x_{11} & x_{13} \\ x_{31} & x_{33} \end{bmatrix} - x_{32} \det \begin{bmatrix} x_{11} & x_{13} \\ x_{21} & x_{23} \end{bmatrix} \\
&= + x_{13} \det \begin{bmatrix} x_{21} & x_{22} \\ x_{31} & x_{32} \end{bmatrix} - x_{23} \det \begin{bmatrix} x_{11} & x_{12} \\ x_{31} & x_{32} \end{bmatrix} + x_{33} \det \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \\
&= + x_{11}x_{22}x_{33} + x_{12}x_{23}x_{31} + x_{13}x_{21}x_{32} - x_{11}x_{23}x_{32} - x_{13}x_{22}x_{31} - x_{12}x_{21}x_{33} . \quad \blacksquare
\end{aligned}$$

Si osservi, infine, che la somma dei prodotti degli elementi della riga i-esima di $\mathbf{X} = [x_{ij}]_{(n,n)}$ per i cofattori degli elementi della riga h-esima $(h \neq i)$ vale zero (3). In simboli,

$$x_{i1} (-1)^{h+1} \det \mathbf{X}_{h1} + \ldots + x_{in} (-1)^{h+n} \det \mathbf{X}_{hn} = 0 . \tag{3}$$

In effetti, il primo membro della (3) rappresenta il determinante, sviluppato secondo gli elementi della riga h-esima, della matrice ottenuta da **X** sostituendo alla riga h-esima la riga i-esima. Dunque, tale matrice ha due righe eguali e, perciò, il suo determinante vale zero.

Per esempio, data la matrice

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix},$$

moltiplicando ciascun elemento della prima riga per i cofattori della seconda riga e sommando, si ha

(3) Naturalmente, un'osservazione analoga vale quando si considerino, in luogo delle righe, le colonne di **X**.

$$x_{11}(-x_{12}) + x_{12}(x_{11}) = 0 \ .$$

## 4 ALCUNI TEOREMI SUI DETERMINANTI

**TEOREMA 1**

Data una matrice a blocchi ($\mathbf{A}$ di ordine $(r,r)$)

$$\mathbf{X}_{(n,n)} = \begin{bmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{C} & \mathbf{B} \end{bmatrix}$$

si ha

$$\det\mathbf{X} = (\det\mathbf{A})(\det\mathbf{B}) \ .$$

*Dim.* La dimostrazione utilizza il procedimento di induzione nell'ordine r di $\mathbf{A} = \left[a_{ij}\right]_{(r,\,r)}$.

Per $r = 1$, si ha $\det\mathbf{X} = a_{11}\det\mathbf{B}$ e il teorema è senz'altro vero.

Per $r > 1$, supponiamo che il teorema sia vero quando la matrice $\mathbf{A}$ è di ordine $(r-1, r-1)$ e dimostriamone la validità per $\mathbf{A}$ di ordine $(r,r)$.

Infatti, sviluppando il determinante di $\mathbf{X}$ secondo gli elementi della prima riga si ottiene

$$\begin{aligned} \det\mathbf{X} &= (-1)^{1+1} a_{11}(\det\mathbf{A}_{11})(\det\mathbf{B}) + \ldots + (-1)^{1+r} a_{1r}(\det\mathbf{A}_{1r})(\det\mathbf{B}) \\ &= \left\{(-1)^{1+1} a_{11}\det\mathbf{A}_{11} + \ldots + (-1)^{1+r} a_{1r}\det\mathbf{A}_{1r}\right\}\det\mathbf{B} \\ &= (\det\mathbf{A})(\det\mathbf{B}) \ . \end{aligned}$$

**TEOREMA 2**

Data una matrice a blocchi ($\mathbf{O}$, $\mathbf{B}$, $\mathbf{A}$, $\mathbf{C}$ di ordine $(r,r)$)

$$\mathbf{Z} = \begin{bmatrix} \mathbf{O} & \mathbf{A} \\ \mathbf{B} & \mathbf{C} \end{bmatrix}$$

si ha

$$\det\mathbf{Z} = (-1)^r(\det\mathbf{A})(\det\mathbf{B}) \ .$$

*Dim.* Si consideri anzitutto la matrice

$$\mathbf{Z}_1 = \begin{bmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{C} & \mathbf{B} \end{bmatrix}$$

il cui determinante, per il Teorema 1, vale

$$\det \mathbf{Z}_1 = (\det \mathbf{A})(\det \mathbf{B}) \ .$$

Poiché si può passare da $\mathbf{Z}_1$ a $\mathbf{Z}$ con r scambi (la prima colonna con la (r+1)-esima, ecc.), si ha

$$\det \mathbf{Z} = (-1)^r \det \mathbf{Z}_1 = (-1)^r (\det \mathbf{A})(\det \mathbf{B}) \ .$$

**COROLLARIO.** Qualora sia ($\mathbf{O}$ , $\mathbf{I}$ , $\mathbf{A}$ , $\mathbf{C}$ di ordine (r,r))

$$\mathbf{Y} = \begin{bmatrix} \mathbf{O} & \mathbf{A} \\ -\mathbf{I} & \mathbf{C} \end{bmatrix}$$

il teorema precedente fornisce immediatamente

$$\det \mathbf{Y} = (-1)^r (\det \mathbf{A})(-1)^r (\det \mathbf{I}) = \det \mathbf{A} \ .$$

**TEOREMA 3**

Date due matrici quadrate $\mathbf{A}$,$\mathbf{B}$ dello stesso ordine, risulta

$$\det(\mathbf{AB}) = (\det \mathbf{A})(\det \mathbf{B}) \ .$$

*Dim.* Considerata l'identità

$$\begin{bmatrix} [\mathbf{A} \ \ \mathbf{O}] + \mathbf{A}[-\mathbf{I} \ \ \mathbf{B}] \\ -\mathbf{I} \qquad\qquad\qquad \mathbf{B} \end{bmatrix} = \begin{bmatrix} \mathbf{O} & \mathbf{AB} \\ -\mathbf{I} & \mathbf{B} \end{bmatrix}$$

è immediato verificare che il determinante della matrice al primo membro, per la proprietà (d) della funzione determinante, è eguale al determinante della matrice

$$\begin{bmatrix} \mathbf{A} & \mathbf{O} \\ -\mathbf{I} & \mathbf{B} \end{bmatrix}$$

e questo, per il Teorema 1, vale $(\det\mathbf{A})(\det\mathbf{B})$.

D'altro canto, il determinante della matrice al secondo membro, per il Corollario al Teorema 2, vale $\det(\mathbf{AB})$, cosicché il teorema risulta provato.

**COROLLARIO.** Si ha

$$\det(\mathbf{AB}) = (\det\mathbf{A})(\det\mathbf{B}) = (\det\mathbf{B})(\det\mathbf{A}) = \det(\mathbf{BA}) \ .$$

## 5 INVERSA DI UNA MATRICE

Ci proponiamo, adesso, di esprimere la condizione necessaria e sufficiente per l'esistenza dell'inversa di una matrice $\mathbf{X} = [x_{ij}]_{(n,n)}$ e di indicare un procedimento di calcolo di tale inversa.

A questo fine, occorre anzitutto introdurre il concetto di *matrice aggiunta* o, più semplicemente, di *aggiunta* di $\mathbf{X}$, denotata con $\widehat{\mathbf{X}}$, e così definita:

- se $n = 1$, $\widehat{\mathbf{X}} = [1]$;
- se $n > 1$, $\widehat{\mathbf{X}}$ è la trasposta della matrice i cui elementi sono i cofattori di $\mathbf{X}$.

Ciò premesso, vogliamo in primo luogo dimostrare che tra $\mathbf{X}$ e la sua aggiunta $\widehat{\mathbf{X}}$ sussiste la relazione

$$\mathbf{X}\widehat{\mathbf{X}} = (\det\mathbf{X})\mathbf{I} \ . \tag{4}$$

Per $n = 1$, la verifica è immediata.

Per $n > 1$, poiché il primo membro della (4) può essere scritto nella forma

$$\begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \cdots & \cdots & \cdots \\ x_{n1} & \cdots & x_{nn} \end{bmatrix} \begin{bmatrix} \widehat{x}_{11} & \cdots & \widehat{x}_{n1} \\ \cdots & \cdots & \cdots \\ \widehat{x}_{1n} & \cdots & \widehat{x}_{nn} \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{n} x_{1k}\widehat{x}_{1k} & \cdots & \sum_{k=1}^{n} x_{1k}\widehat{x}_{nk} \\ \cdots & \cdots & \cdots \\ \sum_{k=1}^{n} x_{nk}\widehat{x}_{1k} & \cdots & \sum_{k=1}^{n} x_{nk}\widehat{x}_{nk} \end{bmatrix},$$

ricordando la (1) e la (3), si trae subito la (4).

In modo del tutto analogo, si dimostra poi che sussiste la relazione

(5) $$\widehat{\mathbf{X}}\mathbf{X} = (\det\mathbf{X})\mathbf{I} .$$

**ESEMPIO 5.** Data la matrice

$$\mathbf{X} = \begin{bmatrix} -1 & 0 & 0 \\ 3 & -2 & 4 \\ 7 & 1 & 5 \end{bmatrix}$$

risulta

$$\widehat{\mathbf{X}} = \begin{bmatrix} -14 & 0 & 0 \\ 13 & -5 & 4 \\ 17 & 1 & 2 \end{bmatrix} .$$

Inoltre ($\det\mathbf{X} = 14$),

$$\begin{bmatrix} -1 & 0 & 0 \\ 3 & -2 & 4 \\ 7 & 1 & 5 \end{bmatrix}\begin{bmatrix} -14 & 0 & 0 \\ 13 & -5 & 4 \\ 17 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 14 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 14 \end{bmatrix}$$

e

$$\begin{bmatrix} -14 & 0 & 0 \\ 13 & -5 & 4 \\ 17 & 1 & 2 \end{bmatrix}\begin{bmatrix} -1 & 0 & 0 \\ 3 & -2 & 4 \\ 7 & 1 & 5 \end{bmatrix} = \begin{bmatrix} 14 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 14 \end{bmatrix} .$$

**TEOREMA 4**

Data una matrice $\mathbf{X}$ di ordine $(n,n)$, se $\det\mathbf{X} \neq 0$, l'inversa di $\mathbf{X}$ esiste ed è espressa da

(6) $$\mathbf{X}^{-1} = (\det\mathbf{X})^{-1}\widehat{\mathbf{X}} .$$

Reciprocamente, se esiste l'inversa di $\mathbf{X}$, allora $\det\mathbf{X} \neq 0$.

*Dim*. Per dimostrare la prima parte del teorema, basta verificare che il secondo membro della (6) è tale che

$$(\det\mathbf{X})^{-1}\widehat{\mathbf{X}}\mathbf{X} = \mathbf{X}(\det\mathbf{X})^{-1}\widehat{\mathbf{X}} = \mathbf{I}\ .$$

Ma la duplice relazione che precede non è che un modo diverso di scrivere la (4) e la (5) e, quindi, la prima parte del teorema è dimostrata.

Supposto poi che esista l'inversa di **X**, per il Corollario al Teorema 3, si ha

$$1 = \det\mathbf{I} = \det(\mathbf{X}^{-1}\mathbf{X}) = \det(\mathbf{X}\mathbf{X}^{-1}) = (\det\mathbf{X})(\det\mathbf{X}^{-1})$$

da cui $\det\mathbf{X} \neq 0$. ∎

Le matrici, ovviamente quadrate, il cui determinante è diverso da zero sono dette *non singolari* o *invertibili*; viceversa, le matrici il cui determinante è eguale a zero sono dette *singolari* o *non invertibili*.

**ESEMPIO 6.** Data la matrice di cui all'Esempio 5, risulta

$$\mathbf{X}^{-1} = 14^{-1}\begin{bmatrix} -14 & 0 & 0 \\ 13 & -5 & 4 \\ 17 & 1 & 2 \end{bmatrix} \ ,\ \ (\mathbf{X}^{-1})^{-1} = \begin{bmatrix} -1 & 0 & 0 \\ 3 & -2 & 4 \\ 7 & 1 & 5 \end{bmatrix}.$$

# Cap. III
# COMPLEMENTI

**1**

Dati un numero reale c e una matrice $\mathbf{X}$ di ordine $(n,n)$, tenuto conto della proprietà $(a_1)$ della funzione determinante, si ha

$$\det(c\mathbf{X}) = c^n \det\mathbf{X} .$$

**2**

Data una matrice $\mathbf{X}$ di ordine $(n,n)$, risulta (1)

$$\det\widehat{\mathbf{X}} = (\det\mathbf{X})^{n-1} .$$

Infatti, dalla relazione $\widehat{\mathbf{X}}\mathbf{X} = (\det\mathbf{X})\mathbf{I}$, si ottiene

$$\det(\widehat{\mathbf{X}}\mathbf{X}) = (\det\widehat{\mathbf{X}})(\det\mathbf{X}) = \det((\det\mathbf{X})\mathbf{I}) = (\det\mathbf{X})^n .$$

**3**

Dati un numero reale $c \neq 0$ e una matrice non singolare $\mathbf{X}$ di ordine $(n,n)$, posto $\mathbf{Y} = c\mathbf{X}$, si ha

$$\mathbf{Y}^{-1} = c^{-1}\mathbf{X}^{-1} .$$

Infatti,

(1) Se $\mathbf{X}$ è la matrice nulla di ordine (1,1), il secondo membro della espressione che segue perde di significato, mentre il primo membro assume il valore 1.

$$\mathbf{Y}^{-1} = (\det\mathbf{Y})^{-1}\widehat{\mathbf{Y}} = (c^n\det\mathbf{X})^{-1}c^{n-1}\widehat{\mathbf{X}} = c^{-1}(\det\mathbf{X})^{-1}\widehat{\mathbf{X}} = c^{-1}\mathbf{X}^{-1} .$$

**4**

Data una matrice (detta *diagonale a blocchi*, *quasi diagonale*)

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{11} & \cdots & \mathbf{O} \\ \cdots & \cdots & \cdots \\ \mathbf{O} & \cdots & \mathbf{X}_{tt} \end{bmatrix}$$

dove $\mathbf{X}_{11}, \dots, \mathbf{X}_{tt}$ sono $t \geq 2$ matrici quadrate non necessariamente dello stesso ordine, risulta

$$\det\mathbf{X} = (\det\mathbf{X}_{11}) \dots (\det\mathbf{X}_{tt}) .$$

**5**

Data una matrice a blocchi

$$\mathbf{X} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}$$

dove $\mathbf{D}$ è invertibile, proponiamoci di calcolarne il determinante.

Per il Teorema 1, si ha

$$\det\begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{Z} & \mathbf{I} \end{bmatrix} = (\det\mathbf{I})(\det\mathbf{I}) = 1$$

qualunque sia $\mathbf{Z}$.

In particolare, se $\mathbf{Z} = -\mathbf{D}^{-1}\mathbf{C}$, allora

$$\det\begin{bmatrix} \mathbf{I} & \mathbf{O} \\ -\mathbf{D}^{-1}\mathbf{C} & \mathbf{I} \end{bmatrix} = 1 .$$

Pertanto, ricordando il Teorema 3, risulta

$$\det\mathbf{X} = \left(\det\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}\right)\left(\det\begin{bmatrix} \mathbf{I} & \mathbf{O} \\ -\mathbf{D}^{-1}\mathbf{C} & \mathbf{I} \end{bmatrix}\right) = \det\left(\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}\begin{bmatrix} \mathbf{I} & \mathbf{O} \\ -\mathbf{D}^{-1}\mathbf{C} & \mathbf{I} \end{bmatrix}\right)$$

$$= \det\begin{bmatrix} \mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C} & \mathbf{B} \\ \mathbf{O} & \mathbf{D} \end{bmatrix} = \det(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})\det\mathbf{D} .$$

Analogamente, se $\mathbf{A}$ è invertibile, si dimostra che

$$\det\mathbf{X} = \det(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})\det\mathbf{A} .$$

**6**

Data una matrice a blocchi, invertibile,

$$\mathbf{X} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}$$

supponiamo che $\mathbf{A}$ sia non singolare.

Posto (2)

$$\mathbf{H} = \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B} ,$$

si verifica agevolmente che la matrice

$$\mathbf{S} = \begin{bmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}\mathbf{H}^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{H}^{-1} \\ -\mathbf{H}^{-1}\mathbf{C}\mathbf{A}^{-1} & \mathbf{H}^{-1} \end{bmatrix}$$

soddisfa alle condizioni $\mathbf{SX} = \mathbf{XS} = \mathbf{I}$ e che, pertanto, $\mathbf{S} = \mathbf{X}^{-1}$.

Analogamente, se $\mathbf{D}$ è non singolare, posto

$$\mathbf{K} = \mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C} ,$$

si dimostra che la matrice

$$\mathbf{T} = \begin{bmatrix} \mathbf{K}^{-1} & -\mathbf{K}^{-1}\mathbf{B}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}\mathbf{K}^{-1} & \mathbf{D}^{-1}\mathbf{C}\mathbf{K}^{-1}\mathbf{B}\mathbf{D}^{-1} + \mathbf{D}^{-1} \end{bmatrix}$$

è l'inversa di $\mathbf{X}$.

**7**

Date tre matrici non singolari $\mathbf{A}$, $\mathbf{C}$, $\mathbf{A}+\mathbf{BCD}$, risulta (3)

---

(2) Si noti che, avendo supposto $\mathbf{X}$ invertibile, per quanto detto al punto precedente $\mathbf{H}$ è non singolare.

(3) La verifica è lasciata per esercizio al lettore.

$$(\mathbf{A}+\mathbf{BCD})^{-1}=\mathbf{A}^{-1}-\mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1}+\mathbf{DA}^{-1}\mathbf{B})^{-1}\mathbf{DA}^{-1} .$$

**8**

Date due matrici

$$\mathbf{X}=\left[x_{ij}\right]_{(n,n)} \quad , \quad \mathbf{Y}=\left[y_{hk}\right]_{(p,p)}$$

si dimostra che

$$\det(\mathbf{Y}\otimes\mathbf{X}) = (\det\mathbf{Y})^p(\det\mathbf{X})^n .$$

**9**

Abbiamo già osservato nel paragrafo 3 del Cap. II che può accadere che sia $\mathbf{XY} = \mathbf{O}$ senza che né $\mathbf{X}$ né $\mathbf{Y}$ siano eguali a $\mathbf{O}$ (*esistenza di divisori* ($\mathbf{X}$, $\mathbf{Y}$) *dello zero* ($\mathbf{O}$)), oppure che sia $\mathbf{XY}=\mathbf{XZ}$ con $\mathbf{Y}\neq\mathbf{Z}$ e $\mathbf{X}\neq\mathbf{O}$.

Ci proponiamo adesso di riprendere l'argomento stabilendo i seguenti teoremi.

**TEOREMA**

> Siano $\mathbf{X}$, $\mathbf{Y}$ due matrici quadrate tali che $\mathbf{XY}=\mathbf{O}$. Allora, $\mathbf{X}$ o $\mathbf{Y}$ sono eguali a $\mathbf{O}$, oppure $\mathbf{X}$ e $\mathbf{Y}$ sono entrambe singolari.

*Dim*. Supponiamo che sia $\mathbf{X}\neq\mathbf{O}$, $\mathbf{Y}\neq\mathbf{O}$ e che almeno una delle due matrici, diciamo $\mathbf{X}$, sia invertibile.

Allora, $\mathbf{XY}=\mathbf{O}$ implica che $\mathbf{X}^{-1}\mathbf{XY}=\mathbf{O}$, e cioè che $\mathbf{Y} = \mathbf{O}$, contro l'ipotesi.

**TEOREMA**

> Se $\mathbf{X}$ è invertibile, $\mathbf{XY}=\mathbf{XZ}$ oppure $\mathbf{YX}=\mathbf{ZX}$ implica che sia $\mathbf{Y} = \mathbf{Z}$.

*Dim*. Infatti, $\mathbf{XY} = \mathbf{XZ}$ implica che $\mathbf{X}^{-1}\mathbf{XY}=\mathbf{X}^{-1}\mathbf{XZ}$, e cioè che $\mathbf{Y} = \mathbf{Z}$.

In modo analogo, si dimostra che $\mathbf{YX} = \mathbf{ZX}$ implica che sia $\mathbf{Y} = \mathbf{Z}$.

# Cap. IV
# SISTEMI DI EQUAZIONI LINEARI

## 1 GENERALITÀ SUI SISTEMI DI EQUAZIONI LINEARI. IL TEOREMA DI CRAMER

**1.1** Un *sistema di* n *equazioni lineari in* p *incognite* è un oggetto del tipo

$$
\begin{array}{c}
x_{11}a_1 + \ldots + x_{1p}a_p = y_1 \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
x_{n1}a_1 + \ldots + x_{np}a_p = y_n
\end{array}
\tag{1}
$$

in cui $a_1, \ldots, a_p$ rappresentano le *incognite*, mentre $x_{11}, \ldots, x_{np}$ e $y_1, \ldots, y_n$ denotano numeri reali *assegnati* che ricevono la denominazione, rispettivamente, di *coefficienti* e *termini noti* del sistema.

Posto

$$
\mathbf{X} = \begin{bmatrix} x_{11} \cdots x_{1p} \\ \cdots\cdots\cdots \\ x_{n1} \cdots x_{np} \end{bmatrix} \quad , \quad \mathbf{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix} \quad , \quad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}
$$

il sistema (1) può essere scritto, più compattamente, nella forma

$$
\mathbf{Xa} = \mathbf{y} \, . \tag{1'}
$$

Qualora sia $\mathbf{y} \neq \mathbf{0}$ – nel caso, cioè, in cui i termini noti non siano tutti nulli – il sistema è detto *non omogeneo*; altrimenti, è detto *omogeneo*.

Si dice poi *vettore soluzione* (o, più semplicemente, *soluzione*) del sistema ogni vettore $\overline{\mathbf{a}} = \left[\overline{a}_j\right]_{(p,1)}$ tale che la (1') risulti soddisfatta.

Il sistema è *compatibile* se esiste almeno un vettore soluzione; *incompatibile* nel caso contrario.

*Risolvere* il sistema significa decidere se esso è compatibile o incompatibile e, nel primo caso, calcolarne tutte le soluzioni.

È ovvio che un sistema omogeneo ammette, in ogni caso, la soluzione *banale* $\overline{\mathbf{a}} = \mathbf{0}$.

**1.2** Prima di affrontare in tutta generalità i vari problemi connessi con la soluzione di un sistema di equazioni lineari, conviene considerare un caso particolare dimostrando il seguente teorema.

**TEOREMA 1** (*Teorema di Cramer*)

Dato un sistema di n equazioni lineari in n incognite $\mathbf{X}\mathbf{a} = \mathbf{y}$, se la matrice dei coefficienti $\mathbf{X}$ è invertibile, il sistema ammette un'unica soluzione fornita dall'espressione

$$\overline{\mathbf{a}} = \mathbf{X}^{-1}\mathbf{y} \, . \tag{2}$$

*Dim*. Basta osservare che il vettore $\overline{\mathbf{a}} = \mathbf{X}^{-1}\mathbf{y}$, la cui esistenza e unicità sono conseguenza dell'ipotesi fatta sulla matrice $\mathbf{X}$, è tale che

$$\mathbf{X}\,\overline{\mathbf{a}} = \mathbf{X}\mathbf{X}^{-1}\mathbf{y} = \mathbf{y} \, .$$

**COROLLARIO.** Qualora sia $\mathbf{y} = \mathbf{0}$, vale a dire nel caso in cui il sistema sia omogeneo, l'*unica* soluzione ammissibile è quella banale costituita dal vettore $\overline{\mathbf{a}} = \mathbf{0}$.

**ESEMPIO 1.** Dato il sistema non omogeneo

$$\begin{bmatrix} 2 & -10 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

risulta

$$\overline{\mathbf{a}} = \begin{bmatrix} 2 & -10 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 12^{-1} \begin{bmatrix} 1 & 10 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 12^{-1} \begin{bmatrix} 13 \\ -1 \end{bmatrix}.$$

**OSSERVAZIONE 1.** Utilizzando la (6) del Cap. III, la (2) diventa

$$\overline{\mathbf{a}} = \mathbf{X}^{-1}\mathbf{y} = (\det\mathbf{X})^{-1}\widehat{\mathbf{X}}\mathbf{y}\,. \tag{2'}$$

Inoltre, come si verifica facilmente, quest'ultima espressione può essere scritta nella forma

$$\overline{\mathbf{a}} = (\det\mathbf{X})^{-1} \begin{bmatrix} \det\mathbf{X}_{1\,(\mathbf{y})} \\ \vdots \\ \det\mathbf{X}_{n\,(\mathbf{y})} \end{bmatrix} \tag{2''}$$

dove $\mathbf{X}_{i\,(\mathbf{y})}$ rappresenta la matrice ottenuta da $\mathbf{X}$ sostituendovi alla colonna i-esima il vettore dei termini noti $\mathbf{y}$.

**ESEMPIO 2.** Considerato nuovamente il sistema di cui all'Esempio 1, utilizzando la (2''), si ha

$$\overline{\mathbf{a}} = 12^{-1} \begin{bmatrix} \det\begin{bmatrix} 3 & -10 \\ 1 & 1 \end{bmatrix} \\ \det\begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix} \end{bmatrix} = 12^{-1} \begin{bmatrix} 13 \\ -1 \end{bmatrix}.$$

## 2 RANGO DI UNA MATRICE

**2.1** Sia data una matrice $\mathbf{X}$ di ordine $(n,p)$ e supponiamo, per il momento, che essa non sia la matrice nulla.

Consideriamo poi l'insieme $I(\mathbf{X})$ costituito da tutte le sottomatrici *quadrate* di $\mathbf{X}$, ottenute cancellando un certo numero di righe e/o di colonne non necessariamente consecutive di $\mathbf{X}$, e da $\mathbf{X}$ stessa nel caso che questa sia quadrata.

Tale insieme conterrà almeno un elemento $\mathbf{X}_*$ di ordine *massimo* $(r,r)$ con determinante diverso da zero. Il numero r si chiama *rango* di **X** e si indica con $r(\mathbf{X})$.

Quindi, dire che la matrice **X** ha rango $r(\mathbf{X}) = r$ significa affermare che:

(i) esiste almeno un elemento $\mathbf{X}_* \in I(\mathbf{X})$ di ordine massimo $(r,r)$ il cui determinante è diverso da zero;

(ii) se $r < \min\{n,p\}$, ogni elemento di $I(\mathbf{X})$ di ordine $(s,s)$ con $s > r$ ha determinante nullo.

Qualora sia $\mathbf{X} = \mathbf{O}$, si dice che **X** ha rango zero e si scrive $r(\mathbf{X}) = 0$.

**ESEMPIO 3.** Data la matrice

$$\mathbf{X} = \begin{bmatrix} -1 & 0 & 2 & 1 \\ 0 & 1 & 1 & -1 \\ 2 & 0 & -4 & -2 \end{bmatrix}$$

si verifica facilmente che il determinante di ciascuno dei seguenti elementi di $I(\mathbf{X})$ di ordine $(3,3)$

$$\begin{bmatrix} -1 & 0 & 2 \\ 0 & 1 & 1 \\ 2 & 0 & -4 \end{bmatrix}, \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & -1 \\ 2 & 0 & -2 \end{bmatrix}, \begin{bmatrix} -1 & 2 & 1 \\ 0 & 1 & -1 \\ 2 & -4 & -2 \end{bmatrix}, \begin{bmatrix} 0 & 2 & 1 \\ 1 & 1 & -1 \\ 0 & -4 & -2 \end{bmatrix}$$

è eguale a zero.

D'altro canto, **X** contiene almeno un elemento di $I(\mathbf{X})$ di ordine $(2,2)$ il cui determinante è diverso da zero: per esempio, la sottomatrice

$$\mathbf{X}_{*1} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

che ha determinante $-1$; pertanto, $r(\mathbf{X}) = 2$. ▮

Sono ovvie proprietà del rango di una matrice le seguenti:

1. Il rango di una matrice **X** di ordine $(n,p)$ è un numero compreso tra 0 e il più piccolo dei due numeri n e p; in simboli, $0 \le r(\mathbf{X}) \le \min\{n,p\}$.

2. Se una matrice $\mathbf{X}$ è di ordine $(n,n)$, $\mathbf{X}^{-1}$ esiste se e soltanto se $r(\mathbf{X}) = n$. In tal caso, si dice anche che $\mathbf{X}$ è *di pieno rango* (1).
3. Data una matrice $\mathbf{Z} = [\mathbf{X}\ \ \mathbf{Y}]$, ottenuta per accostamento di due matrici $\mathbf{X}$ e $\mathbf{Y}$, è $r(\mathbf{X}) \leq r(\mathbf{Z})$.
4. Qualunque sia la matrice $\mathbf{X}$, è $r(\mathbf{X}) = r(\mathbf{X}')$.

**2.2** Il teorema seguente pone in evidenza il legame esistente tra il rango di una matrice non nulla $\mathbf{X}$ e il massimo numero di vettori (colonna o riga) linearmente indipendenti di $\mathbf{X}$.

**TEOREMA 2**

Data una matrice $\mathbf{X}$ di ordine $(n,p)$ e rango r, in $\mathbf{X}$ esistono r vettori colonna (riga) linearmente indipendenti.

Inoltre, se $r<p$ $(r<n)$, ogni altro vettore colonna (riga) di $\mathbf{X}$ può essere espresso come combinazione lineare degli r vettori colonna (riga) linearmente indipendenti.

*Dim*. È sufficiente dimostrare il teorema per i vettori colonna $\mathbf{x}_1, \dots, \mathbf{x}_p$ di $\mathbf{X}$, sussistendo una dimostrazione del tutto analoga per i vettori riga.

Per quanto riguarda la prima parte del teorema, per ipotesi, esiste almeno un elemento $\mathbf{X}_* \in I(\mathbf{X})$ di ordine massimo $(r,r)$, il cui determinante è diverso da zero.

Supponiamo, senza perdita di generalità, che sia

$$\mathbf{X}_* = \mathbf{X}_{*1} = \begin{bmatrix} x_{11} & \cdots & x_{1r} \\ \cdots & \cdots & \cdots \\ x_{r1} & \cdots & x_{rr} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{*1} & \cdots & \mathbf{x}_{*r} \end{bmatrix}$$

e domandiamoci per quali valori di $c_1, \dots, c_r$ si abbia

$$c_1 \mathbf{x}_{*1} + \dots + c_r \mathbf{x}_{*r} = \mathbf{0} \ .$$

---

(1) Le denominazioni matrice *non singolare*, *invertibile*, *di pieno rango* sono equivalenti.

La relazione ora scritta rappresenta un sistema omogeneo la cui matrice dei coefficienti $\mathbf{X}_{*1}$ ha determinante diverso da zero.

Ne consegue che l'unica soluzione ammissibile per tale sistema è quella banale (Cfr. il Corollario al Teorema 1) e, quindi, che $\mathbf{x}_{*1}, \ldots, \mathbf{x}_{*r}$ sono linearmente indipendenti.

Ciò comporta, a sua volta, che anche $\mathbf{x}_1, \ldots, \mathbf{x}_r$ siano linearmente indipendenti, e la prima parte del teorema risulta dimostrata.

Per quanto riguarda la seconda parte del teorema, supposto $r < p$, si consideri la matrice

$$\mathbf{D} = \left[\begin{array}{ccc|c} x_{11} & \cdots & x_{1r} & x_{1j} \\ \cdots & \cdots & \cdots & \cdots \\ x_{r1} & \cdots & x_{rr} & x_{rj} \\ \hline x_{i1} & \cdots & x_{ir} & x_{ij} \end{array}\right]$$

dove $1 \leq i \leq n$ e $r+1 \leq j \leq p$.

Se $i \leq r$, $\mathbf{D}$ possiede due righe eguali e perciò $\det\mathbf{D} = 0$; se $i > r$, $\mathbf{D}$ rappresenta una sottomatrice di $\mathbf{X}$ di ordine $(r+1, r+1)$ e, quindi, è ancora $\det\mathbf{D} = 0$.

Il cofattore dell'elemento $x_{ih}$ $(1 \leq h \leq r)$ di $\mathbf{D}$ dipende da h e da j ma non da i e possiamo indicarlo con $D_{jh}$, mentre il cofattore di $x_{ij}$ è

$$(-1)^{r+1+r+1}\det\mathbf{X}_{*1} = \det\mathbf{X}_{*1} .$$

Quindi, sviluppando il $\det\mathbf{D}$ secondo gli elementi dell'ultima riga, si ha

$$x_{i1}D_{j1} + \ldots + x_{ir}D_{jr} + x_{ij}\det\mathbf{X}_{*1} = \det\mathbf{D} = 0 .$$

Facendo assumere all'indice i, in successione, i valori $1, \ldots, n$, dalla relazione precedente si ottiene

$$\begin{array}{c} x_{11}D_{j1} + \ldots + x_{1r}D_{jr} + x_{1j}\det\mathbf{X}_{*1} = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_{n1}D_{j1} + \ldots + x_{nr}D_{jr} + x_{nj}\det\mathbf{X}_{*1} = 0 \end{array}$$

da cui, posto $(j = r+1, \ldots, p;\ h = 1, \ldots, r)$

$$b_{jh} = -\frac{D_{jh}}{\det \mathbf{X}_{*1}},$$

si ha infine che

$$b_{j1}\mathbf{x}_1 + \dots + b_{jr}\mathbf{x}_r = \mathbf{x}_j$$

e cioè che ciascun vettore $\mathbf{x}_j$ $(j = r+1, \dots, p)$ può essere espresso come combinazione lineare degli r vettori colonna $\mathbf{x}_1, \dots, \mathbf{x}_r$.

**ESEMPIO 4.** Consideriamo nuovamente la matrice **X** di cui all'Esempio 3. Abbiamo già posto in evidenza che $r(\mathbf{X}) = 2$ e che la sottomatrice $\mathbf{X}_{*1}$ ivi indicata risulta invertibile.

Verifichiamo ora, seguendo lo schema della dimostrazione appena svolta, che ciascuno degli ultimi due vettori colonna di **X** si può esprimere come combinazione lineare dei primi due.

Infatti, sviluppando il determinante delle sottomatrici

$$\begin{bmatrix} -1 & 0 & 2 \\ 0 & 1 & 1 \\ -1 & 0 & 2 \end{bmatrix}, \quad \begin{bmatrix} -1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} -1 & 0 & 2 \\ 0 & 1 & 1 \\ 2 & 0 & -4 \end{bmatrix}$$

secondo gli elementi dell'ultima riga, si ottiene

$$\begin{aligned} -1(-2) + 0(1) + 2(-1) &= 0 \\ 0(-2) + 1(1) + 1(-1) &= 0 \\ 2(-2) + 0(1) - 4(-1) &= 0 \end{aligned}$$

da cui

$$(-2)\begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ -4 \end{bmatrix}.$$

In modo del tutto analogo si verifica poi che

$$(-1)\begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} + (-1)\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}.$$

**2.3** Data una matrice $\mathbf{X}$ di ordine (n,p), siano $\mathcal{R}(\mathbf{X})$ e $\widetilde{\mathcal{R}}(\mathbf{X})$ gli spazi vettoriali generati, rispettivamente, dai p vettori colonna e dagli n vettori riga di $\mathbf{X}$.

Se $\mathbf{X} = \mathbf{O}$, $\mathcal{R}(\mathbf{X})$ e $\widetilde{\mathcal{R}}(\mathbf{X})$ sono spazi vettoriali nulli e, quindi,

$$\dim(\mathcal{R}(\mathbf{X})) = \dim(\widetilde{\mathcal{R}}(\mathbf{X})) = \mathrm{r}(\mathbf{X}) = 0 \ .$$

Se $\mathbf{X} \neq \mathbf{O}$, tenuto conto del Teorema 2, si riconosce subito che

$$\dim(\mathcal{R}(\mathbf{X})) = \dim(\widetilde{\mathcal{R}}(\mathbf{X})) = \mathrm{r}(\mathbf{X}) = \mathrm{r} \ .$$

La dimensione (comune) di $\mathcal{R}(\mathbf{X})$ e $\widetilde{\mathcal{R}}(\mathbf{X})$ riceve talvolta la denominazione, rispettivamente, di *rango per colonne* e *rango per righe* di $\mathbf{X}$.

Qualora sia $\mathrm{r}(\mathbf{X}) = \mathrm{p}$ ($\mathrm{p} \leq \mathrm{n}$), si dice che $\mathbf{X}$ è di *pieno rango per colonne*.

Analogamente, nel caso in cui sia $\mathrm{r}(\mathbf{X}) = \mathrm{n}$ ($\mathrm{n} \leq \mathrm{p}$), si dice che $\mathbf{X}$ è di *pieno rango per righe*.

## 3 SOLUZIONE DEI SISTEMI DI EQUAZIONI LINEARI

**3.1** Siamo adesso in grado di affrontare con tutta generalità il problema della soluzione di un sistema di equazioni lineari.

Sussiste anzitutto il seguente teorema che fornisce un criterio per decidere se un dato sistema di equazioni lineari è compatibile.

**TEOREMA 3** (*Teorema di Rouché-Capelli*)

Condizione necessaria e sufficiente affinché un sistema di n equazioni lineari in p incognite $\mathbf{Xa} = \mathbf{y}$ compatibile è che la matrice $\mathbf{X}$ e la cosiddetta matrice *completa* $\mathbf{Z} = [\mathbf{X} \ \ \mathbf{y}]$ abbiano eguale rango.

*Dim*. Supponiamo che il sistema $\mathbf{Xa} = \mathbf{y}$ sia compatibile, ovvero che il vettore $\mathbf{y}$ sia combinazione lineare di $\mathbf{x}_1, \dots, \mathbf{x}_\mathrm{p}$ ($\mathbf{y} \in \mathcal{R}(\mathbf{X})$).

Ne consegue che i due insiemi di vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$ e $\mathbf{x}_1, \dots, \mathbf{x}_p, \mathbf{y}$ generano lo stesso spazio vettoriale (Cfr. l'Osservazione 11 del Cap. I) e, quindi, che le matrici **X** e **Z** hanno eguale rango.

Reciprocamente, supponiamo che le matrici **X** e **Z** abbiano eguale rango.

Qualora sia $r(\mathbf{X}) = r$, per il Teorema 2 − applicato alla matrice **X** − in **X** esistono r vettori colonna, diciamo $\mathbf{x}_1, \dots, \mathbf{x}_r$, linearmente indipendenti.

Sempre per il Teorema 2 − applicato alla matrice **Z** − il vettore **y** risulta combinazione lineare di $\mathbf{x}_1, \dots, \mathbf{x}_r$; ma, in tal caso, **y** è anche combinazione lineare di $\mathbf{x}_1, \dots, \mathbf{x}_p$ ($\mathbf{y} \in \mathcal{R}(\mathbf{X})$), ovvero il sistema $\mathbf{Xa} = \mathbf{y}$ è compatibile.

**OSSERVAZIONE 2.** Si noti esplicitamente che se fosse $\mathbf{X} = \mathbf{O}$ e $\mathbf{y} = \mathbf{0}$, qualunque vettore appartenente a $R^p$ sarebbe un soluzione del sistema.

Invece, qualora fosse $\mathbf{X} = \mathbf{O}$ e $\mathbf{y} \neq \mathbf{0}$, il sistema sarebbe incompatibile.

**OSSERVAZIONE 3.** Il Teorema 3 suggerisce un procedimento che può essere seguito per accertare la compatibilità di un sistema di equazioni lineari.

Tale procedimento, tuttavia, dipendendo dalla valutazione del rango delle matrici **X** e **Z** e, in definitiva, dal calcolo di un certo numero di determinanti, può risultare assai oneroso non appena l'ordine delle matrici coinvolte superi limiti anche modesti.

**ESEMPIO 5.** Dato il sistema non omogeneo

$$\begin{bmatrix} 2 & -3 \\ 1 & 7 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} -4 \\ 15 \\ 11 \end{bmatrix}$$

si verifica facilmente che

$$r(\mathbf{X}) = r\left( \begin{bmatrix} 2 & -3 \\ 1 & 7 \\ 3 & 4 \end{bmatrix} \right) = 2 \quad , \quad r(\mathbf{Z}) = r\left( \begin{bmatrix} 2 & -3 & -4 \\ 1 & 7 & 15 \\ 3 & 4 & 11 \end{bmatrix} \right) = 2$$

e, quindi, il sistema è compatibile.

Per contro, considerato il sistema non omogeneo

$$\begin{bmatrix} 2 & -3 \\ 1 & 7 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} -4 \\ 15 \\ 10 \end{bmatrix}$$

poiché

$$r(\mathbf{X}) = r\left( \begin{bmatrix} 2 & -3 \\ 1 & 7 \\ 3 & 4 \end{bmatrix} \right) = 2 \quad , \quad r(\mathbf{Z}) = r\left( \begin{bmatrix} 2 & -3 & -4 \\ 1 & 7 & 15 \\ 3 & 4 & 10 \end{bmatrix} \right) = 3$$

il sistema è incompatibile.

**3.2** Supponiamo di aver accertato che il sistema di n equazioni lineari in p incognite $\mathbf{Xa} = \mathbf{y}$ sia compatibile e che $r(\mathbf{X}) = r = r(\mathbf{Z})$.

Supponiamo inoltre, senza perdita di generalità e con la notazione stabilita in precedenza, che sia

$$\mathbf{X}_{*1} = \begin{bmatrix} x_{11} \cdots x_{1r} \\ \cdots\cdots\cdots \\ x_{r1} \cdots x_{rr} \end{bmatrix} .$$

Circa le soluzioni del sistema $\mathbf{Xa} = \mathbf{y}$, per ragioni di chiarezza espositiva, conviene distinguere le situazioni seguenti.

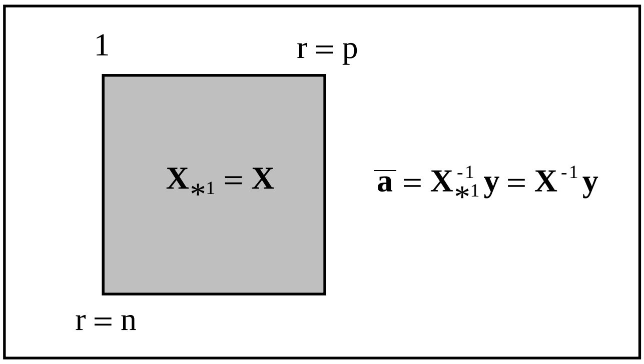
(a) Supponiamo che sia $r = n = p$. In tal caso, la matrice $\mathbf{X}_{*1}$ coincide con la matrice $\mathbf{X}$ di ordine $(n,n)$, vale a dire si ha

$$\mathbf{X}_{*1} = \begin{bmatrix} x_{11} \cdots x_{1n} \\ \cdots\cdots\cdots \\ x_{n1} \cdots x_{nn} \end{bmatrix} = \mathbf{X} .$$

Pertanto, in accordo con il Teorema 1, il sistema $\mathbf{Xa} = \mathbf{y}$ ammette un'unica soluzione data da

$$\overline{\mathbf{a}} = \mathbf{X}_{*1}^{-1}\mathbf{y} = \mathbf{X}^{-1}\mathbf{y} . \tag{2}$$

Quanto ora detto è riassunto nel riquadro che segue.




**ESEMPIO 6.** Si consideri il sistema lineare non omogeneo

$$\begin{bmatrix} 3 & 1 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}.$$

Poiché

$$\mathbf{X}_{*1} = \begin{bmatrix} 3 & 1 \\ 2 & -1 \end{bmatrix} = \mathbf{X}$$

risulta

$$\overline{\mathbf{a}} = (-5)^{-1}\begin{bmatrix} -1 & -1 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} 6 \\ 2 \end{bmatrix} = (-5)^{-1}\begin{bmatrix} -8 \\ -6 \end{bmatrix}.$$

(b) Supponiamo che sia $r = n < p$. In tal caso, la matrice $\mathbf{X}_{*1}$ è una sotto-matrice di ordine $(n,n)$ della matrice $\mathbf{X}$ di ordine $(n,p)$.

Pertanto, fatte le posizioni

$$\mathbf{X}_{*1} = \begin{bmatrix} x_{11} \cdots x_{1n} \\ \cdots\cdots\cdots \\ x_{n1} \cdots x_{nn} \end{bmatrix}, \quad \mathbf{X}_2 = \begin{bmatrix} x_{1,n+1} \cdots x_{1p} \\ \cdots\cdots\cdots\cdots \\ x_{n,n+1} \cdots x_{np} \end{bmatrix}, \quad \mathbf{a}_1 = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} a_{n+1} \\ \vdots \\ a_p \end{bmatrix}$$

si ha

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{*1} & \mathbf{X}_2 \end{bmatrix}$$

e il sistema $\mathbf{Xa} = \mathbf{y}$ può essere riscritto nella forma

$$\begin{bmatrix} \mathbf{X}_{*1} & \mathbf{X}_2 \end{bmatrix} \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{bmatrix} = \mathbf{X}_{*1}\mathbf{a}_1 + \mathbf{X}_2\mathbf{a}_2 = \mathbf{y}$$

da cui

$$\mathbf{X}_{*1}\mathbf{a}_1 = \mathbf{y} - \mathbf{X}_2\mathbf{a}_2 \, .$$

Scelto arbitrariamente il vettore $\mathbf{a}_2$, quest'ultimo sistema, per il Teorema 1, ammette un'unica soluzione, dipendente da $\mathbf{a}_2$, data da

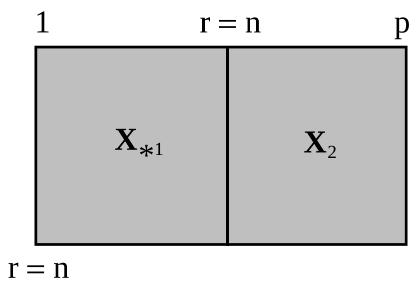
(3) $$\overline{\mathbf{a}}_1 = \mathbf{X}_{*1}^{-1}\mathbf{y} - \mathbf{X}_{*1}^{-1}\mathbf{X}_2\mathbf{a}_2 \, .$$

Quindi, in corrispondenza di ogni $\mathbf{a}_2$, mediante la (3), si ottiene un vettore $\overline{\mathbf{a}}_1$ che, unitamente al vettore $\mathbf{a}_2$ medesimo, costituisce una soluzione di $\mathbf{Xa} = \mathbf{y}$.

In definitiva, il sistema $\mathbf{Xa} = \mathbf{y}$ ammette infinite soluzioni della forma ($\mathbf{0}$: vettore colonna di ordine $p-n$; $\mathbf{I}$: matrice unità di ordine $(p-n, p-n)$)

(4) $$\overline{\mathbf{a}} = \begin{bmatrix} \mathbf{X}_{*1}^{-1}\mathbf{y} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} -\mathbf{X}_{*1}^{-1}\mathbf{X}_2 \\ \mathbf{I} \end{bmatrix} \mathbf{a}_2 \, .$$

Quanto ora detto è riassunto nel riquadro che segue.




$$\overline{\mathbf{a}} = \begin{bmatrix} \mathbf{X}_{*1}^{-1}\mathbf{y} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} -\mathbf{X}_{*1}^{-1}\mathbf{X}_2 \\ \mathbf{I} \end{bmatrix} \mathbf{a}_2$$

**ESEMPIO 7.** Si consideri il sistema lineare non omogeneo

$$\begin{bmatrix} 3 & 1 & 2 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix} .$$

Poiché

$$\mathbf{X}_{*1} = \begin{bmatrix} 3 & 1 \\ 2 & -1 \end{bmatrix} \quad , \quad \mathbf{X}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

risulta

$$\begin{aligned} \overline{\mathbf{a}}_1 &= (-5)^{-1} \begin{bmatrix} -1 & -1 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 6 \\ 2 \end{bmatrix} - (-5)^{-1} \begin{bmatrix} -1 & -1 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} a_3 \\ &= (-5)^{-1} \begin{bmatrix} -8 \\ -6 \end{bmatrix} + (-5)^{-1} \begin{bmatrix} 3 \\ 1 \end{bmatrix} a_3 \end{aligned}$$

e

$$\overline{\mathbf{a}} = \begin{bmatrix} (-5)^{-1}\,(-8) \\ (-5)^{-1}\,(-6) \\ 0 \end{bmatrix} + \begin{bmatrix} (-5)^{-1}\,(3) \\ (-5)^{-1}\,(1) \\ 1 \end{bmatrix} a_3 \,.$$

(c) Supponiamo che sia $r = p < n$. In tal caso, la matrice $\mathbf{X}_{*1}$ è una sotto-matrice di ordine $(p,p)$ della matrice $\mathbf{X}$ di ordine $(n,p)$.

Fatte le posizioni

$$\mathbf{X}_{*1} = \begin{bmatrix} x_{11} \cdots x_{1p} \\ \cdots\cdots\cdots \\ x_{p1} \cdots x_{pp} \end{bmatrix} \;,\; \mathbf{X}_3 = \begin{bmatrix} x_{p+1,1} \cdots x_{p+1,p} \\ \cdots\cdots\cdots\cdots \\ x_{n1} \cdots x_{np} \end{bmatrix} \;,\; \mathbf{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix}$$

$$\mathbf{y}_1 = \begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix} \quad , \quad \mathbf{y}_3 = \begin{bmatrix} y_{p+1} \\ \vdots \\ y_n \end{bmatrix}$$

risulta

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{*1} \\ \mathbf{X}_3 \end{bmatrix}$$

e il sistema $\mathbf{Xa} = \mathbf{y}$ può essere riscritto nella forma

$$\begin{bmatrix} \mathbf{X}_{*1} \\ \mathbf{X}_3 \end{bmatrix} \mathbf{a} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_3 \end{bmatrix} .$$

In altri termini, il sistema in questione può essere ripartito nei sistemi

$$\mathbf{X}_{*1}\mathbf{a} = \mathbf{y}_1 \quad , \quad \mathbf{X}_3\mathbf{a} = \mathbf{y}_3 \ .$$

Ciò premesso – e posto che due sistemi di equazioni lineari, nelle stesse incognite, si dicono *equivalenti* se hanno le medesime soluzioni – mostriamo, anzitutto, che i sistemi $\mathbf{X}\,\mathbf{a} = \mathbf{y}$ e $\mathbf{X}_{*1}\mathbf{a} = \mathbf{y}_1$ sono equivalenti.

In effetti, come è subito visto, ogni soluzione $\overline{\mathbf{a}}$ di $\mathbf{X}\mathbf{a} = \mathbf{y}$ è anche soluzione di $\mathbf{X}_{*1}\mathbf{a} = \mathbf{y}_1$.

Viceversa, sia $\overline{\mathbf{a}}$ una soluzione di $\mathbf{X}_{*1}\mathbf{a} = \mathbf{y}_1$, ovvero sia $\mathbf{X}_{*1}\overline{\mathbf{a}} = \mathbf{y}_1$.

Fissata l'attenzione sulla matrice completa $\mathbf{Z}$ che è della forma

$$\mathbf{Z} = \begin{bmatrix} x_{11} & \cdots & x_{1p} & y_1 \\ \cdots & \cdots & \cdots & \cdots \\ x_{p1} & \cdots & x_{pp} & y_p \\ x_{p+1,1} & \cdots & x_{p+1,p} & y_{p+1} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & \cdots & x_{np} & y_n \end{bmatrix},$$

per il Teorema 2, ciascuno degli ultimi $n-p$ vettori riga di tale matrice è combinazione lineare dei primi p vettori riga, vale a dire si ha ($h = p+1, \dots, n$)

$$[\, x_{h1} \cdots x_{hp} \quad y_h \,] = d_{h1}[\, x_{11} \cdots x_{1p} \quad y_1 \,] + \dots + d_{hp}[\, x_{p1} \cdots x_{pp} \quad y_p \,] \ .$$

Ma le $n-p$ relazioni precedenti, posto

$$\widetilde{\mathbf{D}} = \begin{bmatrix} d_{p+1,1} & \cdots & d_{p+1,p} \\ \cdots & \cdots & \cdots \\ d_{n1} & \cdots & d_{np} \end{bmatrix},$$

possono essere scritte, più compattamente, nella forma

$$[\mathbf{X}_3 \quad \mathbf{y}_3] = \widetilde{\mathbf{D}}\,[\mathbf{X}_{*1} \quad \mathbf{y}_1]$$

da cui

$$\mathbf{X}_3 = \widetilde{\mathbf{D}}\mathbf{X}_{*1} \quad , \quad \mathbf{y}_3 = \widetilde{\mathbf{D}}\mathbf{y}_1 \ .$$

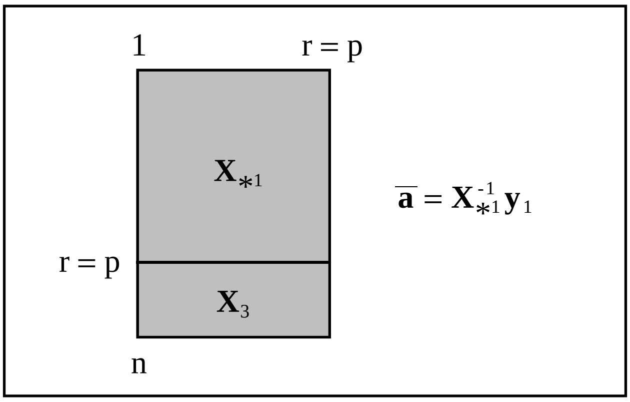
Premoltiplicando ambo i membri di $\mathbf{X}_{*1}\overline{\mathbf{a}} = \mathbf{y}_1$ per la matrice $\widetilde{\mathbf{D}}$, si ottiene

$\widetilde{\mathbf{D}}\mathbf{X}_{*1}\overline{\mathbf{a}} = \widetilde{\mathbf{D}}\mathbf{y}_1$ da cui, tenuto conto di quanto sopra, si desume subito che anche il sistema $\mathbf{X}_3\mathbf{a} = \mathbf{y}_3$ risulta soddisfatto.

Avendo mostrato che i sistemi $\mathbf{Xa} = \mathbf{y}$ e $\mathbf{X}_{*1}\mathbf{a} = \mathbf{y}_1$ sono equivalenti, possiamo concentrare l'attenzione sul secondo sistema il quale, per il Teorema 1, ammette un'unica soluzione data da

$$\overline{\mathbf{a}} = \mathbf{X}_{*1}^{-1}\mathbf{y}_1 \ . \tag{5}$$

Quanto ora detto è riassunto nel riquadro che segue.




**ESEMPIO 8.** Si consideri il sistema lineare non omogeneo

$$\begin{bmatrix} 3 & 1 \\ 2 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \\ 14/5 \end{bmatrix}.$$

Poiché

$$\mathbf{X}_{*1} = \begin{bmatrix} 3 & 1 \\ 2 & -1 \end{bmatrix} \ , \quad \mathbf{X}_3 = \begin{bmatrix} 1 & 1 \end{bmatrix}$$

risulta

$$\overline{\mathbf{a}} = (-5)^{-1}\begin{bmatrix} -1 & -1 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} 6 \\ 2 \end{bmatrix} = (-5)^{-1}\begin{bmatrix} -8 \\ -6 \end{bmatrix}.$$

(d) Supponiamo che sia r < n e r < p. In tal caso, la matrice $\mathbf{X}_{*1}$ è una sottomatrice di ordine $(r,r)$ della matrice $\mathbf{X}$ di ordine $(n,p)$.

Fatte le posizioni

$$\mathbf{X}_{*1} = \begin{bmatrix} x_{11} \cdots x_{1r} \\ \cdots\cdots\cdots \\ x_{r1} \cdots x_{rr} \end{bmatrix} , \ \mathbf{X}_2 = \begin{bmatrix} x_{1,r+1} \cdots x_{1p} \\ \cdots\cdots\cdots\cdots \\ x_{r,r+1} \cdots x_{rp} \end{bmatrix} , \ \mathbf{X}_3 = \begin{bmatrix} x_{r+1,1} \cdots x_{r+1,p} \\ \cdots\cdots\cdots\cdots \\ x_{n1} \cdots x_{np} \end{bmatrix}$$

$$\mathbf{a}_1 = \begin{bmatrix} a_1 \\ \vdots \\ a_r \end{bmatrix} , \ \mathbf{a}_2 = \begin{bmatrix} a_{r+1} \\ \vdots \\ a_p \end{bmatrix} , \ \mathbf{y}_1 = \begin{bmatrix} y_1 \\ \vdots \\ y_r \end{bmatrix} , \ \mathbf{y}_3 = \begin{bmatrix} y_{r+1} \\ \vdots \\ y_n \end{bmatrix}$$

risulta

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{*1} \ \mathbf{X}_2 \\ \mathbf{X}_3 \end{bmatrix}$$

e il sistema $\mathbf{Xa} = \mathbf{y}$ può essere riscritto nella forma

$$\begin{bmatrix} \mathbf{X}_{*1}\mathbf{a}_1 + \mathbf{X}_2\mathbf{a}_2 \\ \mathbf{X}_3\mathbf{a} \end{bmatrix} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_3 \end{bmatrix} .$$

In altri termini, il sistema in questione può essere ripartito nei sistemi

$$\mathbf{X}_{*1}\mathbf{a}_1 + \mathbf{X}_2\mathbf{a}_2 = \mathbf{y}_1 \ , \ \mathbf{X}_3\mathbf{a} = \mathbf{y}_3 \ .$$

Ora, con un ragionamento del tutto simile a quello svolto in precedenza a proposito della situazione descritta in (c), si può facilmente mostrare che i sistemi $\mathbf{Xa} = \mathbf{y}$ e $\mathbf{X}_{*1}\mathbf{a}_1 + \mathbf{X}_2\mathbf{a}_2 = \mathbf{y}_1$ sono equivalenti.

Dunque, possiamo concentrare l'attenzione sul secondo sistema il quale, per il Teorema 1, ammette un'unica soluzione, dipendente da $\mathbf{a}_2$, data da

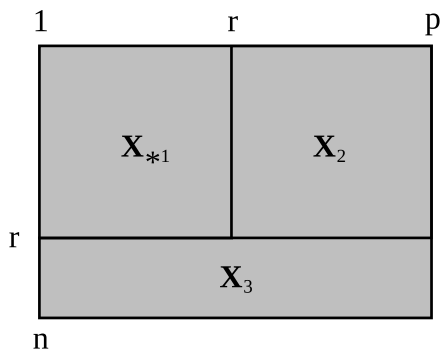
$$\bar{\mathbf{a}}_1 = \mathbf{X}_{*1}^{-1}\mathbf{y}_1 - \mathbf{X}_{*1}^{-1}\mathbf{X}_2\mathbf{a}_2 \ . \tag{6}$$

Pertanto, in corrispondenza di ogni $\mathbf{a}_2$, mediante la (6), si ottiene un vettore $\bar{\mathbf{a}}_1$ che, unitamente al vettore $\mathbf{a}_2$ medesimo, costituisce una soluzione di $\mathbf{Xa} = \mathbf{y}$.

In definitiva, il sistema $\mathbf{X}\,\mathbf{a} = \mathbf{y}$ ammette infinite soluzioni della forma ($\mathbf{0}$: vettore colonna di ordine $p-r$; $\mathbf{I}$: matrice unità di ordine $(p-r, p-r)$)

$$\overline{\mathbf{a}} = \begin{bmatrix} \mathbf{X}_{*1}^{-1}\mathbf{y}_1 \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} -\mathbf{X}_{*1}^{-1}\mathbf{X}_2 \\ \mathbf{I} \end{bmatrix} \mathbf{a}_2 \,. \tag{7}$$

Quanto ora detto è riassunto nel riquadro che segue.




$$\overline{\mathbf{a}} = \begin{bmatrix} \mathbf{X}_{*1}^{-1}\mathbf{y}_1 \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} -\mathbf{X}_{*1}^{-1}\mathbf{X}_2 \\ \mathbf{I} \end{bmatrix} \mathbf{a}_2$$

**ESEMPIO 9.** Si consideri il sistema lineare non omogeneo

$$\begin{bmatrix} 3 & 1 & 2 \\ 2 & -1 & 1 \\ 7 & -1 & 4 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \\ 10 \end{bmatrix}.$$

Poiché

$$\mathbf{X}_{*1} = \begin{bmatrix} 3 & 1 \\ 2 & -1 \end{bmatrix} \;,\quad \mathbf{X}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \;,\quad \mathbf{X}_3 = \begin{bmatrix} 7 & -1 & 4 \end{bmatrix}$$

risulta

$$\begin{aligned} \overline{\mathbf{a}}_1 &= (-5)^{-1} \begin{bmatrix} -1 & -1 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 6 \\ 2 \end{bmatrix} - (-5)^{-1} \begin{bmatrix} -1 & -1 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} a_3 \\ &= (-5)^{-1} \begin{bmatrix} -8 \\ -6 \end{bmatrix} + (-5)^{-1} \begin{bmatrix} 3 \\ 1 \end{bmatrix} a_3 \end{aligned}$$

e

$$\overline{\mathbf{a}} = \begin{bmatrix} (-5)^{-1}\,(-8) \\ (-5)^{-1}\,(-6) \\ 0 \end{bmatrix} + \begin{bmatrix} (-5)^{-1}\,(3) \\ (-5)^{-1}\,(1) \\ 1 \end{bmatrix} a_3 \,.$$

**OSSERVAZIONE 4.** Si sottolinea che un sistema di n equazioni lineari in p incognite $\mathbf{Xa} = \mathbf{y}$ ammette una soluzione unica se e soltanto se $r(\mathbf{X}) = r(\mathbf{Z}) = p$ (situazioni descritte in (a) e (c)).

Il sistema ammette, invece, infinite soluzioni se e soltanto se $r(\mathbf{X}) = r(\mathbf{Z}) < p$ (situazioni descritte in (b) e (d)).

**3.3** Quanto detto nella sezione precedente ha carattere del tutto generale; in particolare, vale qualora il sistema considerato sia omogeneo.

In tal caso, può facilmente dedursi quanto segue.

(a') Se $r = n = p$, l'unica soluzione ammissibile del sistema, in accordo con il Corollario al Teorema 1, è quella banale costituita dal vettore $\overline{\mathbf{a}} = \mathbf{0}$.

(b') Se $r = n < p$, il sistema ammette infinite soluzioni della forma ($\mathbf{I}$: matrice unità di ordine $(p-n, p-n)$)

$$\overline{\mathbf{a}} = \begin{bmatrix} -\mathbf{X}_{*1}^{-1}\mathbf{X}_2 \\ \mathbf{I} \end{bmatrix} \mathbf{a}_2 \,. \tag{8}$$

(c') Se $r = p < n$, l'unica soluzione ammissibile del sistema è quella banale costituita dal vettore $\overline{\mathbf{a}} = \mathbf{0}$.

(d') Se $r < n$ e $r < p$, il sistema ammette infinite soluzioni della forma ($\mathbf{I}$: matrice unità di ordine $(p-r, p-r)$)

$$\overline{\mathbf{a}} = \begin{bmatrix} -\mathbf{X}_{*1}^{-1}\mathbf{X}_2 \\ \mathbf{I} \end{bmatrix} \mathbf{a}_2 \,. \tag{9}$$

**OSSERVAZIONE 5.** Si sottolinea che un sistema omogeneo di n equazioni lineari in p incognite $\mathbf{Xa} = \mathbf{0}$ ammette soluzioni *non banali* – cioè, diverse dalla soluzione $\overline{\mathbf{a}} = \mathbf{0}$ – se e soltanto se $r(\mathbf{X}) < p$.

**3.4** Vogliamo, adesso, porre in evidenza alcune importanti proprietà possedute dall'insieme (eventualmente costituito dal solo vettore nullo) delle soluzioni di un sistema lineare omogeneo $\mathbf{Xa} = \mathbf{0}$.

Tale insieme riceve abitualmente la denominazione di *nucleo* (o *kernel*) di $\mathbf{X}$ e sarà indicato con il simbolo $\mathcal{N}(\mathbf{X})$.

**TEOREMA 4**

> Data una matrice $\mathbf{X}$ di ordine $(n,p)$, il nucleo di $\mathbf{X}$ è uno spazio vettoriale la cui dimensione, detta *nullità* di $\mathbf{X}$, è $p-r(\mathbf{X})$.

*Dim*. Siano $\overline{\mathbf{a}}$ e $\overline{\mathbf{b}}$ appartenenti a $\mathcal{N}(\mathbf{X})$ e c un qualsiasi numero reale.

Allora, come è subito visto, anche $\overline{\mathbf{a}}+\overline{\mathbf{b}}$ e $c\overline{\mathbf{a}}$ appartengono a $\mathcal{N}(\mathbf{X})$ e, quindi, quest'ultimo è uno spazio vettoriale (di ordine p).

Se $r(\mathbf{X}) = 0$, il nucleo di $\mathbf{X}$ coincide evidentemente con $R^p$ e la nullità di $\mathbf{X}$ è p.

Se $r(\mathbf{X}) = p$, il nucleo di $\mathbf{X}$ si identifica con lo spazio nullo (di ordine p) e la nullità di $\mathbf{X}$ è 0.

Posto $0 < r(\mathbf{X}) < p$, vogliamo dimostrare che $\mathcal{N}(\mathbf{X})$ possiede una base formata da $p-r(\mathbf{X})$ vettori e indicare esplicitamente quali sono i vettori che compongono tale base.

A questo fine, è sufficiente rilevare che ogni soluzione $\overline{\mathbf{a}}$ del sistema $\mathbf{Xa} = \mathbf{0}$ – ovvero ciascun vettore appartenente a $\mathcal{N}(\mathbf{X})$ – è esprimibile come combinazione lineare dei $p-r(\mathbf{X})$ vettori colonna di ordine p della matrice ($\mathbf{I}$: matrice unità di ordine $(p-r(\mathbf{X}),p-r(\mathbf{X}))$)

$$\mathbf{C} = \begin{bmatrix} -\mathbf{X}_{*1}^{-1}\mathbf{X}_2 \\ \mathbf{I} \end{bmatrix}.$$

D'altro canto, questi ultimi risultano linearmente indipendenti e, pertanto, formano una base di $\mathcal{N}(\mathbf{X})$.

**ESEMPIO 10.** Dato il sistema lineare omogeneo

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \mathbf{0}$$

tale che $r(\mathbf{X}) = 1$, si ha ($\mathbf{X}_{*1} = [1]$ , $\mathbf{X}_2 = [\,2 \quad 3\,]$)

$$\overline{\mathbf{a}}_1 = [\,\overline{a}_1\,] = -[\,2 \quad 3\,]\begin{bmatrix} a_2 \\ a_3 \end{bmatrix} = -2a_2 - 3a_3 \,.$$

Inoltre, poiché

$$\mathbf{C} = \begin{bmatrix} -2 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix},$$

per ogni $a_2, a_3$,

$$\overline{\mathbf{a}} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} a_2 + \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} a_3 \in \mathcal{N}(\mathbf{X}) \,.$$

**OSSERVAZIONE 6.** Si noti esplicitamente che l'insieme delle soluzioni di un sistema non omogeneo $\mathbf{X}\mathbf{a} = \mathbf{y}$ *non* è uno spazio vettoriale, come risulta dal fatto che tale insieme non contiene il vettore nullo.

## 4 ALCUNI TEOREMI SUL RANGO DI UNA MATRICE

Vogliamo, adesso, completare il quadro delle nozioni fondamentali sul rango di una matrice dimostrando i seguenti teoremi.

**TEOREMA 5**

Siano $\mathbf{X}$ e $\mathbf{Y}$ due matrici tali che il prodotto $\mathbf{XY}$ esiste e sia p il numero di colonne di $\mathbf{Y}$ (e, quindi, di $\mathbf{XY}$). Allora, il rango di $\mathbf{XY}$ è minore o eguale al più piccolo tra i numeri $r(\mathbf{X})$ e $r(\mathbf{Y})$. In simboli,

$$r(\mathbf{XY}) \le \min\,\{r(\mathbf{X})\,,\, r(\mathbf{Y})\}\,.$$

*Dim*. Sia $\mathcal{N}(\mathbf{Y})$ il nucleo di $\mathbf{Y}$; analogamente, sia $\mathcal{N}(\mathbf{XY})$ il nucleo di $\mathbf{XY}$.

Per ogni $\overline{\mathbf{a}} \in \mathcal{N}(\mathbf{Y})$, risulta anche $\overline{\mathbf{a}} \in \mathcal{N}(\mathbf{XY})$, cosicché $\mathcal{N}(\mathbf{Y})$ è un sotto-spazio di $\mathcal{N}(\mathbf{XY})$.

Ne consegue (Cfr. il Teorema 10 del Cap. I e il Teorema 4) che $p - r(\mathbf{Y}) \leq p - r(\mathbf{XY})$, da cui

$$\text{(a)} \qquad r(\mathbf{XY}) \leq r(\mathbf{Y})\ .$$

Allo stesso modo si dimostra poi che $r(\mathbf{Y}'\mathbf{X}') \leq r(\mathbf{X}')$. Ma, poiché $r(\mathbf{X}') = r(\mathbf{X})$ e $r(\mathbf{Y}'\mathbf{X}') = r(\mathbf{XY})$, si ha

$$\text{(b)} \qquad r(\mathbf{XY}) \leq r(\mathbf{X})\ .$$

La (a) e la (b) insieme permettono di concludere secondo quanto affermato nell'enunciato del teorema.

**ESEMPIO 11.** Date le matrici

$$\mathbf{X} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \ , \quad \mathbf{Y} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

entrambe di rango 2, risulta

$$r(\mathbf{XY}) = r\left( \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right) = 1\ .$$

**TEOREMA 6**

> Il rango di una matrice rimane invariato se questa è premoltiplicata o postmoltiplicata per la sua trasposta.

*Dim*. Data una matrice $\mathbf{X}$ di ordine $(n,p)$, vogliamo dimostrare che $r(\mathbf{X}) = r(\mathbf{X}'\mathbf{X}) = r(\mathbf{XX}')$.

Per quanto riguarda l'eguaglianza $r(\mathbf{X}) = r(\mathbf{X}'\mathbf{X})$, è sufficiente provare che il nucleo di $\mathbf{X}$ coincide con il nucleo di $\mathbf{X}'\mathbf{X}$; in tal caso, infatti, tenuto conto del Teorema 4, si ha che $\mathrm{p} - \mathrm{r}(\mathbf{X}) = \mathrm{p} - \mathrm{r}(\mathbf{X}\mathbf{X}')$, da cui $\mathrm{r}(\mathbf{X}) = \mathrm{r}(\mathbf{X}'\mathbf{X})$.

Ora, per ogni $\overline{\mathbf{a}} \in \mathcal{N}(\mathbf{X})$, risulta anche $\overline{\mathbf{a}} \in \mathcal{N}(\mathbf{X}'\mathbf{X})$, cosicché $\mathcal{N}(\mathbf{X})$ è un sottospazio di $\mathcal{N}(\mathbf{X}'\mathbf{X})$.

Sia poi $\overline{\mathbf{b}} \in \mathcal{N}(\mathbf{X}'\mathbf{X})$, vale a dire $\mathbf{X}'\mathbf{X}\overline{\mathbf{b}} = \mathbf{0}$; allora, $\overline{\mathbf{b}}'\mathbf{X}'\mathbf{X}\overline{\mathbf{b}} = 0$, ovvero $(\mathbf{X}\overline{\mathbf{b}})'\mathbf{X}\overline{\mathbf{b}} = 0$. Ma, quest'ultima relazione implica che sia $\mathbf{X}\overline{\mathbf{b}} = \mathbf{0}$ in quanto l'espressione $(\mathbf{X}\overline{\mathbf{b}})'\mathbf{X}\overline{\mathbf{b}}$ rappresenta la somma dei quadrati degli elementi del vettore $\mathbf{X}\overline{\mathbf{b}}$; quindi, $\overline{\mathbf{b}} \in \mathcal{N}(\mathbf{X})$.

Ne segue che $\mathcal{N}(\mathbf{X}'\mathbf{X})$ è un sottospazio di $\mathcal{N}(\mathbf{X})$ e, in definitiva, che $\mathcal{N}(\mathbf{X}) = \mathcal{N}(\mathbf{X}'\mathbf{X})$.

Con una dimostrazione analoga si prova che $\mathrm{r}(\mathbf{X}) = \mathrm{r}(\mathbf{X}\mathbf{X}')$.

**TEOREMA 7**

Il rango di una matrice rimane invariato se questa è premoltiplicata per una matrice di pieno rango per colonne.

*Dim*. Siano $\mathbf{X}$ una matrice di ordine $(\mathrm{n},\mathrm{p})$ e $\mathbf{Y}$ una matrice di ordine $(\mathrm{m},\mathrm{n})$ e rango n. Vogliamo dimostrare che $\mathrm{r}(\mathbf{X}) = \mathrm{r}(\mathbf{Y}\mathbf{X})$.

A questo fine, è sufficiente provare che il nucleo di $\mathbf{X}$ coincide con il nucleo di $\mathbf{YX}$; in tal caso, infatti, tenuto conto del Teorema 4, si ha che $\mathrm{p} - \mathrm{r}(\mathbf{X}) = \mathrm{p} - \mathrm{r}(\mathbf{Y}\mathbf{X})$, da cui $\mathrm{r}(\mathbf{X}) = \mathrm{r}(\mathbf{Y}\mathbf{X})$.

Ora, per ogni $\overline{\mathbf{a}} \in \mathcal{N}(\mathbf{X})$, è anche $\overline{\mathbf{a}} \in \mathcal{N}(\mathbf{Y}\mathbf{X})$, cosicché $\mathcal{N}(\mathbf{X})$ è un sottospazio di $\mathcal{N}(\mathbf{Y}\mathbf{X})$.

Viceversa, per ogni $\overline{\mathbf{b}} \in \mathcal{N}(\mathbf{Y}\mathbf{X})$, è anche $\overline{\mathbf{b}} \in \mathcal{N}(\mathbf{X})$, in quanto, essendo $\mathbf{Y}'\mathbf{Y}$ invertibile [(2)],

$$\{\mathbf{Y}\mathbf{X}\overline{\mathbf{b}} = \mathbf{0}\} \Rightarrow \{\mathbf{Y}'\mathbf{Y}\mathbf{X}\overline{\mathbf{b}} = \mathbf{0}\} \Rightarrow \{\mathbf{X}\overline{\mathbf{b}} = \mathbf{0}\}\ .$$

Ne consegue che $\mathcal{N}(\mathbf{Y}\mathbf{X})$ è un sottospazio di $\mathcal{N}(\mathbf{X})$ e, in definitiva, che $\mathcal{N}(\mathbf{X}) = \mathcal{N}(\mathbf{Y}\mathbf{X})$.

---

(2) La matrice $\mathbf{Y}'\mathbf{Y}$ di ordine $(\mathrm{n},\mathrm{n})$, per il Teorema 6, è tale che $\mathrm{r}(\mathbf{Y}'\mathbf{Y}) = \mathrm{r}(\mathbf{Y}) = \mathrm{n}$.

**OSSERVAZIONE 7.** Il Teorema 7 fornisce una condizione sufficiente ma non necessaria affinché il rango di una matrice rimanga invariato quando questa è premoltiplicata per una matrice.

Per esempio, date le matrici

$$\mathbf{X} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad , \quad \mathbf{Y} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

è $\mathrm{r}(\mathbf{X}) = 1$ e $\mathrm{r}(\mathbf{Y}) = 2$, ma

$$\mathrm{r}(\mathbf{YX}) = \mathrm{r}\left( \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right) = 1 \ .$$

**OSSERVAZIONE 8.** Date due matrici $\mathbf{X}$ di ordine $(n,p)$ e $\mathbf{Z}$ di ordine $(p,q)$ e rango q, risulta, in generale, $\mathrm{r}(\mathbf{X}) \neq \mathrm{r}(\mathbf{XZ})$.

Per esempio, date le matrici

$$\mathbf{X} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} \quad , \quad \mathbf{Z} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

è $\mathrm{r}(\mathbf{X}) = 1$ e $\mathrm{r}(\mathbf{Z}) = 1$, ma

$$\mathrm{r}(\mathbf{XZ}) = \mathrm{r}\left( \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) = \mathrm{r}\left( \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right) = 0 \ .$$

**TEOREMA 8**

Il rango di una matrice rimane invariato se questa è premoltiplicata o postmoltiplicata per una matrice invertibile.

*Dim*. Siano $\mathbf{X}$ una matrice di ordine $(n,p)$ e $\mathbf{Y}$ e $\mathbf{Z}$ due matrici invertibili di ordine, rispettivamente, $(n,n)$ e $(p,p)$.

Tenuto conto del Teorema 7, si ha

$$\mathrm{r}(\mathbf{YX}) = \mathrm{r}(\mathbf{X}) \quad , \quad \mathrm{r}(\mathbf{XZ}) = \mathrm{r}(\mathbf{Z}'\mathbf{X}') = \mathrm{r}(\mathbf{X}') = \mathrm{r}(\mathbf{X}) \ .$$

## 5 CAMBIAMENTI DI BASE IN UNO SPAZIO VETTORIALE

Sia S uno spazio vettoriale di ordine n e dimensione p.

Considerata una base di S costituita da p vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$ di ordine n, scriviamo un generico vettore $\mathbf{x} \in \mathrm{S}$ nella forma (Cfr. l'Osservazione 7 del Cap. II)

$$\mathbf{x} = a_1\mathbf{x}_1 + \dots + a_p\mathbf{x}_p = \left[\mathbf{x}_1 \cdots \mathbf{x}_p\right]\begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix} = \mathbf{Xa} \ .$$

Consideriamo poi p vettori $\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_p$ appartenenti a S.

Poiché risulta

$$\tilde{\mathbf{x}}_1 = a_{11}\mathbf{x}_1 + \dots + a_{p1}\mathbf{x}_p = \left[\mathbf{x}_1 \cdots \mathbf{x}_p\right]\begin{bmatrix} a_{11} \\ \vdots \\ a_{p1} \end{bmatrix} = \mathbf{Xa}_1$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\tilde{\mathbf{x}}_p = a_{1p}\mathbf{x}_1 + \dots + a_{pp}\mathbf{x}_p = \left[\mathbf{x}_1 \cdots \mathbf{x}_p\right]\begin{bmatrix} a_{1p} \\ \vdots \\ a_{pp} \end{bmatrix} = \mathbf{Xa}_p$$

possiamo scrivere

$$\tilde{\mathbf{X}} = \left[\tilde{\mathbf{x}}_1 \cdots \tilde{\mathbf{x}}_p\right] = \left[\mathbf{x}_1 \cdots \mathbf{x}_p\right]\left[\mathbf{a}_1 \cdots \mathbf{a}_p\right] = \mathbf{XA} \, .$$

Ma, tenuto conto del Teorema 7,

$$\mathrm{r}(\tilde{\mathbf{X}}) = \mathrm{r}(\mathbf{XA}) = \mathrm{r}(\mathbf{A}) \ .$$

Dunque, condizione necessaria e sufficiente affinché i p vettori $\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_p$ costituiscano a loro volta una base di S è che la matrice **A**, di ordine (p,p), sia non singolare.

Qualora tale condizione sia verificata, un generico vettore $\mathbf{x} \in S$ può essere scritto univocamente nella forma

$$\mathbf{x} = \tilde{a}_1 \tilde{\mathbf{x}}_1 + \ldots + \tilde{a}_p \tilde{\mathbf{x}}_p = \begin{bmatrix} \tilde{\mathbf{x}}_1 & \cdots & \tilde{\mathbf{x}}_p \end{bmatrix} \begin{bmatrix} \tilde{a}_1 \\ \vdots \\ \tilde{a}_p \end{bmatrix} = \widetilde{\mathbf{X}} \tilde{\mathbf{a}} \; .$$

Pertanto, premoltiplicando il primo e l'ultimo membro dell'identità

$$\mathbf{X}\mathbf{a} = \mathbf{x} = \widetilde{\mathbf{X}}\tilde{\mathbf{a}} = \mathbf{X}\mathbf{A}\tilde{\mathbf{a}}$$

per $\mathbf{X}'$ e osservando che $\mathbf{X}'\mathbf{X}$ è invertibile (3), il legame che sussiste tra i vettori coordinati $\mathbf{a}$ e $\tilde{\mathbf{a}}$ di $\mathbf{x}$, rispetto alle basi di S formate da $\mathbf{x}_1, \ldots, \mathbf{x}_p$ e $\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_p$, risulta espresso dalla relazione

$$\mathbf{a} = \mathbf{A}\tilde{\mathbf{a}}$$

ovvero da

$$\tilde{\mathbf{a}} = \mathbf{A}^{-1}\mathbf{a} \; .$$

**OSSERVAZIONE 9.** Date le matrici $\widetilde{\mathbf{X}}$ e $\mathbf{X}$ di cui sopra, la matrice $\mathbf{A}$ del cambiamento di base può essere facilmente ottenuta mediante l'espressione

$$\mathbf{A} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\widetilde{\mathbf{X}} \, .$$

(3) La matrice $\mathbf{X}'\mathbf{X}$ di ordine $(p, p)$, per il Teorema 6, è tale che $r(\mathbf{X}'\mathbf{X}) = r(\mathbf{X}) = p$.

# Cap. IV
# COMPLEMENTI

**1**

Nel capitolo precedente abbiamo posto in evidenza che *se i vettori colonna* (*riga*) *che compongono una matrice* $\mathbf{X}$ *di ordine* $(n,n)$ *sono linearmente dipendenti*, *allora* $\det\mathbf{X} = 0$.

Vogliamo mostrare che sussiste la proprietà reciproca di quella ora ricordata.

In effetti, escludendo il caso banale in cui sia $\mathbf{X} = \mathbf{O}$, se $\det\mathbf{X} = 0$, allora $r(\mathbf{X}) = p < n$; pertanto (Cfr. il Teorema 2), in $\mathbf{X}$ esistono p vettori colonna (riga) linearmente indipendenti e ogni altro vettore colonna (riga) di $\mathbf{X}$ può essere espresso come combinazione lineare di tali p vettori colonna (riga) linearmente indipendenti. Quindi, i vettori colonna (riga) di $\mathbf{X}$ sono linearmente dipendenti.

**2**

**TEOREMA**

> Date due matrici $\mathbf{Y},\mathbf{X}$ di ordine, rispettivamente, $(m,n)$ e $(n,p)$, condizione necessaria e sufficiente affinché sia $r(\mathbf{YX}) = r(\mathbf{X})$ è che risulti $\mathcal{N}(\mathbf{YX}) = \mathcal{N}(\mathbf{X})$.

*Dim*. Se $\mathcal{N}(\mathbf{YX}) = \mathcal{N}(\mathbf{X})$, allora (Cfr. il Teorema 4) $p - r(\mathbf{YX}) = p - r(\mathbf{X})$ e, quindi, $r(\mathbf{YX}) = r(\mathbf{X})$.

Per dimostrare la proposizione reciproca, osserviamo intanto che, per ogni $\overline{\mathbf{a}} \in \mathcal{N}(\mathbf{X})$, risulta anche $\overline{\mathbf{a}} \in \mathcal{N}(\mathbf{YX})$, cosicché $\mathcal{N}(\mathbf{X})$ è un sottospazio di $\mathcal{N}(\mathbf{YX})$.

D'altra parte, se $r(\mathbf{YX}) = r(\mathbf{X})$, $\mathcal{N}(\mathbf{X})$ e $\mathcal{N}(\mathbf{YX})$ hanno la stessa dimensione $(\{r(\mathbf{YX}) = r(\mathbf{X})\} \equiv \{p - r(\mathbf{YX}) = p - r(\mathbf{X})\})$ e, quindi, coincidono.

**3**

Data una matrice $\mathbf{X} = \left[\mathbf{x}_1 \cdots \mathbf{x}_p\right]$ di ordine $(n,p)$ e rango r, vogliamo mostrare la possibilità di scomporre $\mathbf{X}$ nel prodotto di due matrici (*fattorizzazione* di $\mathbf{X}$).

Poiché $r(\mathbf{X}) = r$, lo spazio vettoriale (di ordine n) generato dai vettori colonna di $\mathbf{X}$ ha dimensione r.

Supposto che $\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_r$ siano una base di tale spazio vettoriale, ogni vettore colonna $\mathbf{x}_j$ $(j = 1, \ldots, p)$ di $\mathbf{X}$ si può esprimere nella forma

$$\mathbf{x}_j = h_{1j}\tilde{\mathbf{x}}_1 + \ldots + h_{rj}\tilde{\mathbf{x}}_r = \left[\tilde{\mathbf{x}}_1 \cdots \tilde{\mathbf{x}}_r\right]\begin{bmatrix} h_{1j} \\ \vdots \\ h_{rj} \end{bmatrix} = \tilde{\mathbf{X}}\mathbf{h}_j$$

ovvero

$$\mathbf{X} = [\,\mathbf{x}_1 \cdots \mathbf{x}_p\,] = \left[\tilde{\mathbf{x}}_1 \cdots \tilde{\mathbf{x}}_r\right][\,\mathbf{h}_1 \cdots \mathbf{h}_p\,] = \tilde{\mathbf{X}}\mathbf{H}\ .$$

Le matrici $\tilde{\mathbf{X}}, \mathbf{H}$ – di ordine, rispettivamente, $(n,r)$ e $(r,p)$ – attuano la desiderata fattorizzazione e, per il Teorema 7, $r(\mathbf{H}) = r(\tilde{\mathbf{X}}\mathbf{H}) = r(\mathbf{X}) = r \le p$.

Si osservi che la fattorizzazione $\mathbf{X} = \tilde{\mathbf{X}}\mathbf{H}$ non è unica. In effetti, se $\mathbf{A}$ è una matrice non singolare di ordine $(r,r)$, allora

$$\mathbf{X} = \tilde{\mathbf{X}}\mathbf{A}\mathbf{A}^{-1}\mathbf{H} = \tilde{\mathbf{Z}}\tilde{\mathbf{H}}\ .$$

**4**

Supposto che $\mathbf{X}$ sia una matrice di ordine $(n,p)$ e rango $r = n < p$, si consideri il sistema non omogeneo $\mathbf{Xa} = \mathbf{y}$.

Come abbiamo visto, tale sistema ammette infinite soluzioni della forma (**0**: vettore colonna di ordine $p-n$; **I**: matrice unità di ordine $(p-n,p-n)$)

$$\bar{\mathbf{a}} = \begin{bmatrix} \mathbf{X}_{*1}^{-1}\mathbf{y} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} -\mathbf{X}_{*1}^{-1}\mathbf{X}_2 \\ \mathbf{I} \end{bmatrix} \mathbf{a}_2$$

ovvero – posto

$$\bar{\mathbf{b}}_1 = \begin{bmatrix} \mathbf{X}_{*1}^{-1}\mathbf{y} \\ \mathbf{0} \end{bmatrix} \ , \ \mathbf{C} = \begin{bmatrix} -\mathbf{X}_{*1}^{-1}\mathbf{X}_2 \\ \mathbf{I} \end{bmatrix}$$

– della forma

$$\bar{\mathbf{a}} = \bar{\mathbf{b}}_1 + \mathbf{C}\mathbf{a}_2 \ .$$

Chiaramente, $\bar{\mathbf{b}}_1$ è una soluzione particolare del sistema $\mathbf{Xa} = \mathbf{y}$, ottenuta ponendo $\mathbf{a}_2 = \mathbf{0}$.

A sua volta, $\mathbf{Ca}_2$ fornisce una soluzione, dipendente da $\mathbf{a}_2$, del sistema omogeneo associato $\mathbf{Xa} = \mathbf{0}$.

D'altro canto, poiché i vettori colonna della matrice **C** costituiscono una base di $\mathcal{N}(\mathbf{X})$ che è un sottospazio di $R^p$ di dimensione $p-r$, possiamo anche dire (Cfr. il punto 5 dei Complementi al Cap. I) che le soluzioni del sistema non omogeneo $\mathbf{Xa} = \mathbf{y}$ costituiscono una varietà lineare di traslazione $\bar{\mathbf{b}}_1$ e direzione $\mathcal{N}(\mathbf{X})$.

Considerazioni analoghe a quelle ora svolte possono ovviamente farsi nel caso in cui sia $r < n$ e $r < p$.

# Cap. V
# TRASFORMAZIONI LINEARI

## 1 GENERALITÀ SULLE TRASFORMAZIONI LINEARI

Dati uno spazio vettoriale S di ordine n e dimensione p e uno spazio vettoriale Z di ordine m e dimensione q, si chiama *trasformazione* (*applicazione*, *operatore*) *lineare di* S *in* Z (o anche *omomorfismo di* S *in* Z) ogni funzione T che associa a ciascun vettore $\mathbf{x} \in S$ un vettore $T(\mathbf{x}) \in Z$ in modo tale che siano soddisfatte le proprietà seguenti ($\mathbf{x}_1, \mathbf{x}_2 \in S$; $a \in R$):

(i) $T(\mathbf{x}_1 + \mathbf{x}_2) = T(\mathbf{x}_1) + T(\mathbf{x}_2)$

(ii) $T(a\mathbf{x}_1) = aT(\mathbf{x}_1)$.

Il vettore $T(\mathbf{x})$ è detto *immagine di* $\mathbf{x}$.

**ESEMPIO 1.** La funzione T definita ponendo, per ogni $\mathbf{x} \in R^3$,

$$T(\mathbf{x}) = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

è, come subito si verifica, una trasformazione lineare di $R^3$ in $R^2$.

**OSSERVAZIONE 1.** Dati t vettori $\mathbf{x}_1, \dots, \mathbf{x}_t$ appartenenti a S e t numeri reali $a_1, \dots, a_t$, qualunque sia la trasformazione lineare T di S in Z, risulta

$$T(a_1\mathbf{x}_1 + \dots + a_t\mathbf{x}_t) = a_1T(\mathbf{x}_1) + \dots + a_tT(\mathbf{x}_t) \ .$$

**OSSERVAZIONE 2.** Qualunque sia la trasformazione lineare T di S in Z, l'immagine del vettore $\mathbf{0} \in S$ è il vettore $\mathbf{0} \in Z$ [1]. Infatti,

$$T(\mathbf{0}) = T(0\mathbf{x}) = 0T(\mathbf{x}) = \mathbf{0} .$$

**OSSERVAZIONE 3.** Chiaramente, la funzione che associa a ogni $\mathbf{x} \in S$ il vettore $\mathbf{0} \in Z$ è una trasformazione lineare di S in Z.

Essa riceve la denominazione di *trasformazione* (*lineare*) *nulla* ed è indicata con il simbolo O.

## 2 TRASFORMAZIONI LINEARI E BASI

**2.1** Supponiamo che i vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$ e $\mathbf{z}_1, \dots, \mathbf{z}_q$ formino due prefissate basi, rispettivamente, di S e Z.

Data una trasformazione lineare T di S in Z e posto $(j = 1, \dots, p)$

$$T(\mathbf{x}_j) = t_{1j}\mathbf{z}_1 + \dots + t_{qj}\mathbf{z}_q = \begin{bmatrix} \mathbf{z}_1 \cdots \mathbf{z}_q \end{bmatrix} \begin{bmatrix} t_{1j} \\ \vdots \\ t_{qj} \end{bmatrix} = \mathbf{Z}\mathbf{t}_j ,$$

per ogni

$$\mathbf{x} = a_1\mathbf{x}_1 + \dots + a_p\mathbf{x}_p = \begin{bmatrix} \mathbf{x}_1 \cdots \mathbf{x}_p \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix} = \mathbf{X}\mathbf{a}$$

appartenente a S, tenuto conto dell'Osservazione 1, si ha che

$$(1) \qquad \begin{aligned} T(\mathbf{x}) &= T(a_1\mathbf{x}_1 + \dots + a_p\mathbf{x}_p) &&= a_1T(\mathbf{x}_1) + \dots + a_pT(\mathbf{x}_p) \\ &= \begin{bmatrix} T(\mathbf{x}_1) \cdots T(\mathbf{x}_p) \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix} &&= \begin{bmatrix} \mathbf{Z}\mathbf{t}_1 \cdots \mathbf{Z}\mathbf{t}_p \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix} \\ &= \mathbf{Z} \begin{bmatrix} \mathbf{t}_1 \cdots \mathbf{t}_p \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix} &&= \mathbf{Z}\mathbf{T}\mathbf{a} . \end{aligned}$$

---

(1) Il simbolo **0** denota sia il vettore nullo di S (di ordine n) sia il vettore nullo di Z (di ordine m).

Mediante la (1) l'immagine $\mathrm{T}(\mathbf{x})$ di $\mathbf{x}$ risulta espressa in funzione del vettore coordinato $\mathbf{a}$ di $\mathbf{x}$ e della matrice $\mathbf{T}$ di ordine $(q,p)$.

Quest'ultima matrice, univocamente associata alla trasformazione lineare T e alle suddette basi di S e Z, riceve la denominazione di *matrice rappresentativa* o, più semplicemente, *matrice* di T rispetto a tali basi.

A sua volta, il vettore

$$\mathbf{c} = \mathbf{Ta}$$

è il vettore coordinato di $\mathrm{T}(\mathbf{x})$ rispetto alla base di Z.

**OSSERVAZIONE 4.** Date le matrici $\left[\, \mathrm{T}(\mathbf{x}_1) \cdots \mathrm{T}(\mathbf{x}_p) \,\right]$ e $\mathbf{Z}$, dalla relazione $\left[\, \mathrm{T}(\mathbf{x}_1) \cdots \mathrm{T}(\mathbf{x}_p) \,\right] = \mathbf{ZT}$ – premoltiplicando ambo i membri per $(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$ – si ha che (2)

$$\mathbf{T} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\left[\, \mathrm{T}(\mathbf{x}_1) \cdots \mathrm{T}(\mathbf{x}_p) \,\right].$$

**ESEMPIO 2.** Rispetto alle basi naturali di $R^3$ e $R^2$, la matrice rappresentativa della trasformazione lineare T di $R^3$ in $R^2$ di cui all'Esempio 1 è

$$\mathbf{T} = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 0 & 0 \end{bmatrix}.$$

Invece, rispetto alle basi di $R^3$ e $R^2$ formate, rispettivamente, da

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \ , \ \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \ , \ \mathbf{x}_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

e

$$\mathbf{z}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \ , \ \mathbf{z}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

– risultando

$$\mathrm{T}(\mathbf{x}_1) = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \ , \ \mathrm{T}(\mathbf{x}_2) = \begin{bmatrix} 6 \\ 1 \end{bmatrix} \ , \ \mathrm{T}(\mathbf{x}_3) = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$$

(2) La matrice $(\mathbf{Z}'\mathbf{Z})$ di ordine (q , q), per il Teorema 6 del Cap. IV, è tale che $r(\mathbf{Z}'\mathbf{Z}) = r(\mathbf{Z}) = q$.

– la matrice rappresentativa della stessa trasformazione lineare T è

$$\mathbf{T} = \left( \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 6 & 7 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -4 & -5 \\ 1 & 5 & 6 \end{bmatrix}.$$

**2.2** Supponiamo, come in precedenza, che i vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$ e $\mathbf{z}_1, \dots, \mathbf{z}_q$ formino due prefissate basi, rispettivamente, di S e Z.

Scelti *arbitrariamente* p vettori $\overline{\mathbf{z}}_1, \dots, \overline{\mathbf{z}}_p$ appartenenti a Z e posto $(j = 1, \dots, p)$

$$\overline{\mathbf{z}}_j = t_{1j}\mathbf{z}_1 + \dots + t_{qj}\mathbf{z}_q = \begin{bmatrix} \mathbf{z}_1 \cdots \mathbf{z}_q \end{bmatrix} \begin{bmatrix} t_{1j} \\ \vdots \\ t_{qj} \end{bmatrix} = \mathbf{Z}\mathbf{t}_j\,,$$

per ogni

$$\mathbf{x} = a_1\mathbf{x}_1 + \dots + a_p\mathbf{x}_p = \begin{bmatrix} \mathbf{x}_1 \cdots \mathbf{x}_p \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix} = \mathbf{X}\mathbf{a}$$

appartenente a S, si definisca la funzione T ponendo

$$(1')\quad T(\mathbf{x}) = a_1\overline{\mathbf{z}}_1 + \dots + a_p\overline{\mathbf{z}}_p = a_1\mathbf{Z}\mathbf{t}_1 + \dots + a_p\mathbf{Z}\mathbf{t}_p = \mathbf{Z}\begin{bmatrix} \mathbf{t}_1 \cdots \mathbf{t}_p \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix} = \mathbf{Z}\mathbf{T}\mathbf{a}\,.$$

Si riconosce subito che T è una trasformazione lineare di S in Z univocamente associata alla matrice **T** (3) e che la matrice rappresentativa di T, rispetto alle suddette basi di S e Z, è proprio **T**.

**ESEMPIO 3.** Supponiamo che

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad , \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

(3) Ovvero, univocamente associata ai p vettori $\overline{\mathbf{z}}_1, \dots, \overline{\mathbf{z}}_p$ appartenenti a Z tali che

$$T(\mathbf{x}_1) = \overline{\mathbf{z}}_1, \dots, T(\mathbf{x}_p) = \overline{\mathbf{z}}_p.$$

siano una base di uno spazio vettoriale S e che

$$\mathbf{z}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

sia una base di uno spazio vettoriale Z.

Considerata la matrice

$$\mathbf{T} = \begin{bmatrix} 2 & 3 \end{bmatrix},$$

per ogni

$$\mathbf{x} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

appartenente a S, la funzione T definita da

$$T(\mathbf{x}) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 2 & 3 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

è una trasformazione lineare di S in Z univocamente associata a **T** (4).

**OSSERVAZIONE 5.** Tenuto conto di quanto mostrato in questa e nella sezione precedente, possiamo affermare che esiste una corrispondenza biunivoca, dipendente dalle basi prescelte di S e Z, tra l'insieme delle trasformazioni lineari di S in Z e l'insieme delle matrici di ordine appropriato.

**2.3** Siano **a** il vettore coordinato di **x** rispetto a una base di S costituita dai vettori $\mathbf{x}_1, \ldots, \mathbf{x}_p$ e **c** il vettore coordinato di T(**x**) rispetto a una base di Z formata dai vettori $\mathbf{z}_1, \ldots, \mathbf{z}_q$.

---

(4) Ovvero, univocamente associata ai vettori

$$\bar{\mathbf{z}}_1 = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} \quad , \quad \bar{\mathbf{z}}_2 = \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}$$

appartenenti a Z.

Come si è visto, tali vettori coordinati sono legati dalla relazione

$$\mathbf{c} = \mathbf{T}\mathbf{a}$$

dove **T** denota la matrice rappresentativa di T rispetto alle suddette basi di S e Z.

Considerate due nuove basi di S e Z costituite, rispettivamente, dai vettori $\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_p$ e $\tilde{\mathbf{z}}_1, \dots, \tilde{\mathbf{z}}_q$ e posto

$$\begin{aligned} \widetilde{\mathbf{X}} &= [\tilde{\mathbf{x}}_1 \cdots \tilde{\mathbf{x}}_p] = [\mathbf{x}_1 \cdots \mathbf{x}_p][\mathbf{a}_1 \cdots \mathbf{a}_p] = \mathbf{X}\mathbf{A} \\ \widetilde{\mathbf{Z}} &= [\tilde{\mathbf{z}}_1 \cdots \tilde{\mathbf{z}}_q] = [\mathbf{z}_1 \cdots \mathbf{z}_q][\mathbf{b}_1 \cdots \mathbf{b}_q] = \mathbf{Z}\mathbf{B} \end{aligned}$$

dove **A** e **B** sono due matrici non singolari di ordine appropriato, i vettori coordinati $\tilde{\mathbf{a}}$ di **x** e $\tilde{\mathbf{c}}$ di T(**x**), relativi a quest'ultime basi, si trasformano in

$$\tilde{\mathbf{a}} = \mathbf{A}^{-1}\mathbf{a} \quad , \quad \tilde{\mathbf{c}} = \mathbf{B}^{-1}\mathbf{c} \, .$$

Ne consegue subito che

$$\tilde{\mathbf{c}} = (\mathbf{B}^{-1}\mathbf{T}\mathbf{A})\tilde{\mathbf{a}} = \widetilde{\mathbf{T}}\,\tilde{\mathbf{a}} \, ,$$

ovvero che la matrice **T** della trasformazione lineare T, relativa alle basi di S e Z formate dai vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$ e $\mathbf{z}_1, \dots, \mathbf{z}_q$, quando tali basi mutano nel modo qui sopra indicato, si trasforma in

$$\widetilde{\mathbf{T}} = \mathbf{B}^{-1}\mathbf{T}\mathbf{A} \, . \tag{2}$$

Viceversa, due matrici $\widetilde{\mathbf{T}}$ e **T**, legate da una relazione del tipo espresso dalla (2), si possono sempre interpretare come matrici della *stessa* trasformazione lineare T, relative a due opportune coppie di basi di S e Z.

**OSSERVAZIONE 6.** Date le matrici $\widetilde{\mathbf{X}}, \mathbf{X}$ e $\widetilde{\mathbf{Z}}, \mathbf{Z}$ di cui sopra, le matrici **A** e **B** possono essere ottenute mediante le espressioni (Cfr. l'Osservazione 9 del Cap IV)

$$\mathbf{A} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\widetilde{\mathbf{X}} \quad , \quad \mathbf{B} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\widetilde{\mathbf{Z}} \, .$$

**ESEMPIO 4.** Riprendendo l'Esempio 2, abbiamo visto che, relativamente alle basi di $R^3$ e $R^2$ formate, rispettivamente, da

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \;,\; \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \;,\; \mathbf{x}_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

e

$$\mathbf{z}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \;,\; \mathbf{z}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

la matrice rappresentativa della trasformazione lineare T ivi considerata era

$$\mathbf{T} = \begin{bmatrix} 0 & -4 & -5 \\ 1 & 5 & 6 \end{bmatrix} .$$

Rispetto alle basi di $R^3$ e $R^2$ formate, rispettivamente, da

$$\widetilde{\mathbf{x}}_1 = \begin{bmatrix} 3 \\ 3 \\ 2 \end{bmatrix} \;,\; \widetilde{\mathbf{x}}_2 = \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix} \;,\; \widetilde{\mathbf{x}}_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

e

$$\widetilde{\mathbf{z}}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \;,\; \widetilde{\mathbf{z}}_2 = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

– risultando

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \;,\; \mathbf{B} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

– la matrice rappresentativa della stessa trasformazione lineare T diviene

$$\widetilde{\mathbf{T}} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 0 & -4 & -5 \\ 1 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} -30 & -29 & -16 \\ 21 & 20 & 11 \end{bmatrix} .$$

**OSSERVAZIONE 7.** Tenuto conto del Teorema 8 del Cap. IV, si ha che

$$r(\widetilde{\mathbf{T}}) = r(\mathbf{B}^{-1}\mathbf{T}\mathbf{A}) = r(\mathbf{T}) .$$

## 3 NUCLEO E IMMAGINE DI UNA TRASFORMAZIONE LINEARE

**3.1** Si chiama *nucleo* (o *kernel*) di una trasformazione lineare T di S in Z l'insieme − che denotiamo con $\mathcal{N}(T)$ − dei vettori $\overline{\mathbf{x}} \in S$ tali che $T(\overline{\mathbf{x}}) = \mathbf{0}$.

Chiaramente, $\mathcal{N}(T)$ è un spazio vettoriale in quanto, per tutti gli $\overline{\mathbf{x}}, \overline{\mathbf{y}} \in \mathcal{N}(T)$ e ogni $a \in R$, risulta

$$T(\overline{\mathbf{x}} + \overline{\mathbf{y}}) = T(\overline{\mathbf{x}}) + T(\overline{\mathbf{y}}) = \mathbf{0} \quad , \quad T(a\overline{\mathbf{x}}) = aT(\overline{\mathbf{x}}) = \mathbf{0} \ .$$

La $\dim(\mathcal{N}(T))$ è anche detta *nullità* di T.

Indichiamo, al solito, con **T** la matrice della trasformazione lineare T rispetto a due basi di S e Z costituite, rispettivamente, dai vettori colonna delle matrici **X** e **Z**.

La determinazione di $\mathcal{N}(T)$ si risolve, in tutta evidenza, nella ricerca di quei vettori $\overline{\mathbf{a}} = [a_j]_{(p,1)}$ tali che $\mathbf{T}\,\overline{\mathbf{a}} = \mathbf{0}$, ovvero di quei vettori $\overline{\mathbf{a}}$ che appartengono a $\mathcal{N}(\mathbf{T})$.

Allora, ricordando quanto detto nell'ambito della dimostrazione del Teorema 4 del Cap. IV, si deduce facilmente quanto segue.

Se $r(\mathbf{T}) = 0$ − vale a dire se T è la trasformazione nulla − allora il nucleo di T è S e la nullità di T è p.

Se $r(\mathbf{T}) = p$, allora il nucleo di T è lo spazio nullo di S e la nullità di T è 0.

Se $0 < r(\mathbf{T}) = r < p$, supponiamo, senza perdita di generalità, che i primi r vettori colonna e riga di $\mathbf{T} = [t_{ij}]_{(q,p)}$ siano linearmente indipendenti.

Posto

$$\mathbf{T}_{*1} = \begin{bmatrix} t_{11} \cdots t_{1r} \\ \cdots\cdots\cdots \\ t_{r1} \cdots t_{rr} \end{bmatrix} \quad , \quad \mathbf{T}_2 = \begin{bmatrix} t_{1,r+1} \cdots t_{1p} \\ \cdots\cdots\cdots\cdots \\ t_{r,r+1} \cdots t_{rp} \end{bmatrix} \quad , \quad \mathbf{C}_{(p,p-r)} = \begin{bmatrix} -\mathbf{T}_{*1}^{-1}\mathbf{T}_2 \\ \mathbf{I} \end{bmatrix}$$

una base del nucleo di **T** è formata dai vettori colonna di $\mathbf{C}_{(p,p-r)}$.

Pertanto, il nucleo di T è costituito da tutte le combinazioni lineari dei vettori colonna della matrice $\mathbf{X}\mathbf{C}_{(p,p-r)}$ e la sua nullità è $p - r$.

Si noti esplicitamente che in quest'ultimo caso $\mathcal{N}(T)$ è un sottospazio proprio di S, mentre nei primi due casi $\mathcal{N}(T)$ è un sottospazio improprio di S.

**3.2** Si chiama *immagine di* S di una trasformazione lineare T l'insieme – che denotiamo con $\mathcal{R}(T)$ – costituito dalle immagini $T(\mathbf{x})$ di ciascun $\mathbf{x} \in S$.

Chiaramente, $\mathcal{R}(T)$ è uno spazio vettoriale dato che, per ogni $T(\mathbf{x})$, $T(\mathbf{y}) \in Z$ e ogni $a \in R$, risulta

$$T(\mathbf{x}) + T(\mathbf{y}) = T(\mathbf{x} + \mathbf{y}) \in \mathcal{R}(T) \quad , \quad aT(\mathbf{x}) = T(a\mathbf{x}) \in \mathcal{R}(T) \ .$$

La $\dim(\mathcal{R}(T))$ è anche detta *rango* di T.

Come è immediato riconoscere, la determinazione di $\mathcal{R}(T)$ si riduce, in buona sostanza, nella ricerca di quei vettori che – risultando combinazioni lineari dei vettori colonna di **T** – appartengono a $\mathcal{R}(T)$.

Allora, tenuto conto di quanto detto nella sezione 2.3 del Cap. IV, si deduce facilmente quanto segue.

Se $r(\mathbf{T}) = 0$ – vale a dire se T è la trasformazione nulla – allora $\mathcal{R}(T)$ è lo spazio nullo di Z e il rango di T è 0.

Se $r(\mathbf{T}) = p$, vogliamo mostrare che i p vettori $T(\mathbf{x}_1), \dots, T(\mathbf{x}_p)$ – cioè le immagini dei vettori che formano una base di S – costituiscono una base di $\mathcal{R}(T)$ e, quindi, che il rango di T è p.

A questo fine, osserviamo intanto che, per ogni

$$\mathbf{x} = a_1\mathbf{x}_1 + \dots + a_p\mathbf{x}_p$$

appartenente a S, si ha che

$$T(\mathbf{x}) = a_1 T(\mathbf{x}_1) + \dots + a_p T(\mathbf{x}_p) \ ,$$

ovvero che $T(\mathbf{x})$ è esprimibile come combinazione lineare dei p vettori $T(\mathbf{x}_1), \dots, T(\mathbf{x}_p)$.

Per dimostrare che $T(\mathbf{x}_1), \dots, T(\mathbf{x}_p)$ sono linearmente indipendenti, supponiamo che esistano p numeri reali $d_1, \dots, d_p$ tali che

$$d_1 T(\mathbf{x}_1) + \dots + d_p T(\mathbf{x}_p) = \mathbf{0} \ .$$

Ciò implica che sia

$$T(d_1\mathbf{x}_1 + \dots + d_p\mathbf{x}_p) = \mathbf{0}$$

e, quindi, che il vettore $d_1\mathbf{x}_1 + ... + d_p\mathbf{x}_p$ appartenga al nucleo di T.

Ma, coincidendo in questo caso $\mathcal{N}(T)$ con lo spazio nullo di S, risulta

$$d_1\mathbf{x}_1 + ... + d_p\mathbf{x}_p = \mathbf{0} .$$

D'altra parte, l'indipendenza di $\mathbf{x}_1, ..., \mathbf{x}_p$ implica che $d_1, ..., d_p$ siano tutti nulli e ciò prova l'indipendenza di $T(\mathbf{x}_1), ..., T(\mathbf{x}_p)$.

Infine, qualora sia $0 < r(\mathbf{T}) = r < p$, supponiamo che i vettori $\overline{\mathbf{x}}_1, ..., \overline{\mathbf{x}}_{p-r}$, $\overline{\mathbf{x}}_{p-r+1}, ..., \overline{\mathbf{x}}_p$ formino una base di S ottenuta mediante il completamento di una base del nucleo di T costituita dai vettori $\overline{\mathbf{x}}_1, ..., \overline{\mathbf{x}}_{p-r}$.

Con un ragionamento analogo a quello appena svolto, si può facilmente provare che gli r vettori $T(\overline{\mathbf{x}}_{p-r+1}), ..., T(\overline{\mathbf{x}}_p)$ formano una base di $\mathcal{R}(T)$ e, quindi, che il rango di T è r.

**ESEMPIO 5.** Consideriamo la trasformazione lineare T di $R^3$ in $R^2$ la cui matrice, rispetto alle basi naturali di $R^3$ e $R^2$, è

$$\mathbf{T} = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 0 & 0 \end{bmatrix} .$$

Come è immediato verificare, una base di $\mathcal{N}(T)$ è costituita dal vettore

$$\overline{\mathbf{x}}_1 = \begin{bmatrix} 0 \\ -3 \\ 1 \end{bmatrix} .$$

A sua volta, una base di $\mathcal{R}(T)$ si può ottenere attraverso il completamento della base di $\mathcal{N}(T)$ mediante, per esempio, i vettori

$$\overline{\mathbf{x}}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad , \quad \overline{\mathbf{x}}_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

e, successivamente, trasformando quest'ultimi nei vettori

$$T(\overline{\mathbf{x}}_2) = \begin{bmatrix} 6 \\ 1 \end{bmatrix} \quad , \quad T(\overline{\mathbf{x}}_3) = \begin{bmatrix} 7 \\ 1 \end{bmatrix} .$$

**OSSERVAZIONE 8.** Poiché la dimensione di $\mathcal{R}(T)$ è eguale a r(**T**), ricordando quanto detto nella sezione che precede a proposito della dimensione di $\mathcal{N}(T)$, possiamo concludere che

$$\dim(S) = \dim(\mathcal{N}(T)) + \dim(\mathcal{R}(T)) \ .$$

**OSSERVAZIONE 9.** Dato che $\dim(\mathcal{N}(T)) \geq 0$, risulta che

$$\dim(\mathcal{R}(T)) \leq \dim(S) \ .$$

**3.3** Si dice che una trasformazione lineare T di S in Z è *iniettiva* se, per tutti gli $\mathbf{x}, \mathbf{y} \in S$, $T(\mathbf{x}) = T(\mathbf{y})$ implica che sia $\mathbf{x} = \mathbf{y}$, ovvero $\mathbf{x} \neq \mathbf{y}$ implica che sia $T(\mathbf{x}) \neq T(\mathbf{y})$.

A sua volta, si dice che una trasformazione lineare T di S in Z è *surgettiva* se $\mathcal{R}(T) = Z$.

Ciò premesso, ci proponiamo di dimostrare il seguente teorema.

**TEOREMA 1**

Una trasformazione lineare T di S in Z

a) è iniettiva se e soltanto se la dimensione di $\mathcal{N}(T)$ è eguale a zero;

b) è surgettiva se e soltanto se $\dim(\mathcal{R}(T)) = \dim(Z)$.

*Dim*. a) Se T è iniettiva, il nucleo di T è costituito dal solo vettore $\mathbf{0} \in S$ e, quindi, $\dim(\mathcal{N}(T)) = 0$. Viceversa, supposto che sia $\dim(\mathcal{N}(T)) = 0$, se $T(\mathbf{x}) = T(\mathbf{y})$ per tutti gli $\mathbf{x}, \mathbf{y} \in S$, allora $T(\mathbf{x} - \mathbf{y}) = T(\mathbf{x}) - T(\mathbf{y}) = \mathbf{0}$ e, pertanto, il vettore $\mathbf{x} - \mathbf{y}$ appartiene a $\mathcal{N}(T)$; ma, essendo $\mathbf{0}$ l'unico elemento di $\mathcal{N}(T)$, risulta $\mathbf{x} = \mathbf{y}$ e, quindi, T è iniettiva.

b) Poiché $\mathcal{R}(T)$ è un sottospazio (proprio o improprio) di Z, $\mathcal{R}(T)$ e Z coincidono se e soltanto se hanno la stessa dimensione, vale a dire se e soltanto se $\dim(\mathcal{R}(T)) = \dim(Z)$ (5).

---

(5) Si noti che una trasformazione lineare di S in Z al tempo stesso iniettiva e surgettiva stabilisce un isomorfismo tra i due spazi vettoriali (Cfr. il punto 7 dei Complementi al Cap. I).

## 4 TRASFORMAZIONI LINEARI DI UNO SPAZIO VETTORIALE IN SE STESSO

Supposto che sia Z = S, la funzione T che associa a ciascun vettore $\mathbf{x} \in S$ un vettore $T(\mathbf{x}) \in S$ in modo tale che siano soddisfatte le proprietà (i) e (ii) di cui al paragrafo 1, si chiama *trasformazione* (*applicazione*, *operatore*) *lineare di* S *in se stesso* (o anche *endomorfismo di* S).

**OSSERVAZIONE 10.** Considerata la funzione che associa a ogni $\mathbf{x} \in S$ il vettore $c\mathbf{x} \in S$, si verifica facilmente che questa è una trasformazione lineare di S in se stesso. Essa riceve la denominazione di *omotetia* di *rapporto* c.

Per c = 1, si ha la *trasformazione* (*lineare*) *identica*, indicata con il simbolo I. ▮

Si applicano, ovviamente, a questo caso le considerazioni svolte in precedenza circa le trasformazioni lineari di S in Z.

In particolare, indichiamo con **A** la matrice del passaggio da una base formata da $\mathbf{x}_1, \dots, \mathbf{x}_p$, vettori colonna della matrice **X**, a una base costituita da $\widetilde{\mathbf{x}}_1, \dots, \widetilde{\mathbf{x}}_p$, vettori colonna della matrice $\widetilde{\mathbf{X}} = \mathbf{XA}$.

Allora, la matrice **T** della trasformazione lineare T di S in se stesso, relativa alla prima base, si trasforma in

$$\widetilde{\mathbf{T}} = \mathbf{A}^{-1}\mathbf{TA}\,. \tag{3}$$

In generale, due matrici quadrate $\widetilde{\mathbf{T}}$ e $\mathbf{T}$, dello stesso ordine, legate da una relazione del tipo scritto in (3), si dicono *simili* e il legame tra $\widetilde{\mathbf{T}}$ e $\mathbf{T}$ espresso dalla (3) si chiama *trasformazione per similitudine*.

Possiamo pertanto affermare che la matrice **T** della trasformazione lineare T di S in se stesso, quando cambia la base, è soggetta a una trasformazione per similitudine.

Viceversa, due matrici simili si possono sempre interpretare come matrici della *stessa* trasformazione lineare T di S in se stesso, relative a due opportune basi di S.

**ESEMPIO 6.** Relativamente alla base naturale di $R^2$, sia

$$\mathbf{T} = \begin{bmatrix} 2 & -3 \\ 1 & -1 \end{bmatrix}$$

la matrice di una trasformazione lineare T di $R^2$ in se stesso.

Rispetto alla base di $R^2$ formata da

$$\tilde{\mathbf{x}}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad , \quad \tilde{\mathbf{x}}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

la matrice della stessa trasformazione lineare T diventa

$$\tilde{\mathbf{T}} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 2 & -3 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -1 & -7 \\ 1 & 3 \end{bmatrix} .$$

**OSSERVAZIONE 11.** Ovviamente, $r(\tilde{\mathbf{T}}) = r(\mathbf{T})$.

Inoltre,

(a) $\quad tr(\tilde{\mathbf{T}}) = tr(\mathbf{A}^{-1}\mathbf{TA}) = tr(\mathbf{TAA}^{-1}) = tr(\mathbf{T})$

(b) $\quad det(\tilde{\mathbf{T}}) = det(\mathbf{A}^{-1}\mathbf{TA}) = det(\mathbf{A}^{-1})\,det(\mathbf{T})\,det(\mathbf{A}) = det(\mathbf{T})$.

## 5 AUTOVALORI E AUTOVETTORI

**5.1** Si considerino una trasformazione lineare T di uno spazio vettoriale S di ordine n e dimensione p in se stesso e un sottospazio $S_1$ di S.

Si dice che $S_1$ è *invariante* rispetto a T se, per ogni $\mathbf{x} \in S_1$, anche $T(\mathbf{x}) \in S_1$.

Data la trasformazione lineare T, ci proponiamo di determinare i sottospazi di S di dimensione eguale a 1 invarianti rispetto a T.

Formalmente, ciò equivale a considerare l'equazione

$$T(\mathbf{x}) = \lambda \mathbf{x} \tag{4}$$

e a ricercare quei numeri $\tilde{\lambda} \in R$ e quei vettori $\tilde{\mathbf{x}} \neq \mathbf{0}$ tali che la (4) risulti soddisfatta.

Qualora siffatti numeri e vettori esistano, essi ricevono la denominazione, rispettivamente, di *autovalori* (*radici caratteristiche*, *radici latenti*, *valori propri*) e *autovettori* (*vettori caratteristici*, *vettori latenti*, *vettori propri*) di T.

**ESEMPIO 7.** Si consideri la trasformazione lineare T di $R^2$ in se stesso definita ponendo, per ogni $\mathbf{x} \in R^2$,

$$T(\mathbf{x}) = \begin{bmatrix} 1 & 4 \\ 9 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Si verifica facilmente che il numero reale $\tilde{\lambda} = 7$ e il vettore

$$\tilde{\mathbf{x}} = c \begin{bmatrix} 2 \\ 3 \end{bmatrix},$$

dove c è un numero reale qualsiasi purché diverso da zero, soddisfano la equazione $T(\mathbf{x}) = \lambda\mathbf{x}$ e sono pertanto, rispettivamente, un autovalore e un autovettore di T.

**ESEMPIO 8.** Si consideri la trasformazione lineare T di $R^2$ in se stesso definita ponendo, per ogni $\mathbf{x} \in R^2$,

$$T(\mathbf{x}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

In tal caso, com'è subito visto, non esiste un numero reale $\tilde{\lambda}$ e un vettore $\tilde{\mathbf{x}} \neq \mathbf{0}$ tali che l'equazione

$$\begin{bmatrix} -x_2 \\ x_1 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

risulti soddisfatta e, pertanto, la trasformazione lineare in questione *non* ammette autovalori e autovettori.

**5.2** Il procedimento mediante il quale si determinano gli autovalori e gli autovettori di una trasformazione lineare T di uno spazio vettoriale S in se stesso può essere esposto nei seguenti termini.

Si consideri una base di S costituita da p vettori $\mathbf{x}_1, \ldots, \mathbf{x}_p$ e sia $\mathbf{X} = [\mathbf{x}_1 \cdots \mathbf{x}_p]$.

Siano, inoltre, $\mathbf{T}$ la matrice rappresentativa di T e $\mathbf{a}$ il vettore coordinato di un generico vettore $\mathbf{x} \in S$ rispetto alla suddetta base di S.

La (4) può essere allora espressa nella forma

(4') $$\mathbf{XTa} = \lambda \mathbf{Xa}$$

oppure nell'altra

(4'') $$\mathbf{Ta} = \lambda \mathbf{a} .$$

Supponiamo, adesso, che esistano un numero $\tilde{\lambda} \in R$ e un vettore $\tilde{\mathbf{a}} \neq \mathbf{0}$ che soddisfino la (4').

È ovvio che questi soddisfano anche la (4''), e viceversa.

Ne consegue che $\tilde{\lambda}$ è un autovalore di T e che $\tilde{\mathbf{a}}$ è il vettore coordinato rispetto alla base di S costituita dai vettori colonna della matrice $\mathbf{X}$ dell'autovettore $\tilde{\mathbf{x}}$ associato a tale autovalore.

Per determinare gli autovalori e gli autovettori della trasformazione lineare T occorrerà, pertanto, ricercare quei numeri $\tilde{\lambda} \in R$ e quei vettori $\tilde{\mathbf{a}} \neq \mathbf{0}$ tali che la (4'') risulti soddisfatta.

A questo fine, si osservi anzitutto che la (4'') può essere riscritta nella forma

(4''') $$(\mathbf{T} - \lambda \mathbf{I})\,\mathbf{a} = \mathbf{0} .$$

Ora, condizione necessaria e sufficiente affinché il sistema lineare omogeneo di cui sopra ammetta soluzioni non banali è che sia

(5) $$\det(\mathbf{T} - \lambda \mathbf{I}) = 0 .$$

Il primo membro della (5) – com'è subito visto sviluppandone il determinante – è un polinomio di grado p in $\lambda$, detto *polinomio caratteristico* di $\mathbf{T}$.

A sua volta, la (5) riceve la denominazione di *equazione caratteristica* di $\mathbf{T}$.

Fissata una base di S e quindi la matrice $\mathbf{T}$ rappresentativa rispetto a tale base della trasformazione lineare T, la ricerca degli autovalori di T conduce dunque alla ricerca delle radici (reali) dell'equazione caratteristica di **T**.

Il passo successivo consiste nel mostrare che le radici dell'equazione caratteristica di **T** sono indipendenti dalla base prescelta di S.

In effetti, sia **A** la matrice del cambiamento di base di S.

Tenuto conto della (3), si ha che

$$\begin{aligned}\det(\mathbf{T}-\lambda\mathbf{I}) &= \det(\mathbf{A}^{-1})\det(\mathbf{T}-\lambda\mathbf{I})\det(\mathbf{A}) = \det(\mathbf{A}^{-1}(\mathbf{T}-\lambda\mathbf{I})\mathbf{A})\\ &= \det(\mathbf{A}^{-1}\mathbf{T}\mathbf{A}-\lambda\mathbf{I}) \qquad\qquad = \det(\widetilde{\mathbf{T}}-\lambda\mathbf{I})\end{aligned}$$

ovvero che matrici simili, avendo lo stesso polinomio caratteristico, hanno anche una medesima equazione caratteristica ed eguali radici.

Poiché matrici simili identificano una stessa trasformazione lineare T, le radici dell'equazione caratteristica di **T** coincidono con gli autovalori di T.

Supponiamo, adesso, che la (5) ammetta l'autovalore $\widetilde{\lambda}$. Sostituendo nella (4''') $\widetilde{\lambda}$ a $\lambda$, l'insieme delle soluzioni del sistema lineare omogeneo risultante

$$(\mathbf{T}-\widetilde{\lambda}\mathbf{I})\,\mathbf{a} = \mathbf{0}$$

è uno spazio vettoriale di ordine p la cui dimensione (Cfr. il Teorema 4 del Cap. IV) è $p - r(\mathbf{T}-\widetilde{\lambda}\mathbf{I})$ e, premoltiplicando ciascun vettore $\widetilde{\mathbf{a}}\neq\mathbf{0}$ appartenente a tale spazio vettoriale per **X**, si ottiene un autovettore $\widetilde{\mathbf{x}}$ di T associato all'autovalore $\widetilde{\lambda}$.

Lo spazio vettoriale generato da tutti gli autovettori associati a $\widetilde{\lambda}$ riceve generalmente la denominazione di *autospazio* di $\widetilde{\lambda}$ e la sua dimensione, pari a $p - r(\mathbf{T}-\widetilde{\lambda}\mathbf{I})$, è anche detta *molteplicità geometrica* di $\widetilde{\lambda}$.

Ovviamente, considerato un cambiamento della base di S determinato da una matrice **A**, ogni soluzione $\widetilde{\mathbf{a}}\neq\mathbf{0}$ del sistema $(\mathbf{T}-\widetilde{\lambda}\mathbf{I})\,\mathbf{a} = \mathbf{0}$, risultando

$$\begin{aligned}&\{\widetilde{\mathbf{a}} : (\mathbf{T}-\widetilde{\lambda}\mathbf{I})\,\widetilde{\mathbf{a}} = \mathbf{0}\} \qquad\qquad = \{\widetilde{\mathbf{a}} : \mathbf{A}^{-1}(\mathbf{T}-\widetilde{\lambda}\mathbf{I})\,\widetilde{\mathbf{a}} = \mathbf{0}\}\\ &= \{\widetilde{\mathbf{a}} : (\mathbf{A}^{-1}\mathbf{T}\mathbf{A}-\widetilde{\lambda}\mathbf{A}^{-1}\mathbf{A})\mathbf{A}^{-1}\,\widetilde{\mathbf{a}} = \mathbf{0}\} = \{\widetilde{\mathbf{a}} : (\widetilde{\mathbf{T}}-\widetilde{\lambda}\mathbf{I})\,\mathbf{A}^{-1}\,\widetilde{\mathbf{a}} = \mathbf{0}\}\ ,\end{aligned}$$

si trasforma, al cambiamento di base, in $\mathbf{A}^{-1}\,\widetilde{\mathbf{a}}$.

**OSSERVAZIONE 12.** L'equazione caratteristica di $\mathbf{T}$ ammette l'autovalore $\tilde{\lambda}$ = 0 se e soltanto se $\mathbf{T}$ è singolare.

Infatti, se $\mathbf{T}$ è singolare, esiste un vettore $\tilde{\mathbf{a}} \neq \mathbf{0}$ tale che $\mathbf{T}\tilde{\mathbf{a}} = \mathbf{0}$, e questa ultima relazione si scrive anche $\mathbf{T}\tilde{\mathbf{a}} = \tilde{\lambda}\tilde{\mathbf{a}}$ con $\tilde{\lambda} = 0$.

Viceversa, se $\tilde{\lambda} = 0$, allora $\det\mathbf{T} = 0$ e, quindi, $\mathbf{T}$ è singolare.

**ESEMPIO 9.** Si consideri la trasformazione lineare T di $R^2$ in se stesso la cui matrice rispetto alla base formata da

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \;,\; \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

è

$$\mathbf{T} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}.$$

L'equazione caratteristica di $\mathbf{T}$ – vale a dire, l'equazione

$$\det\left( \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) = \det \begin{bmatrix} 1-\lambda & -2 \\ 0 & 1-\lambda \end{bmatrix} = 0$$

– ammette l'autovalore (unico) $\tilde{\lambda} = 1$.

Poiché $r(\mathbf{T} - \tilde{\lambda}\mathbf{I}) = 1$, le soluzioni del sistema

$$\begin{bmatrix} 0 & -2 \\ 0 & 0 \end{bmatrix} \mathbf{a} = \mathbf{0}$$

formano uno spazio vettoriale di dimensione 1 e il vettore

$$\tilde{\mathbf{a}} = b \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

dove b è un numero reale qualsiasi purché diverso da zero, è un vettore non nullo che appartiene a tale spazio vettoriale.

Ne consegue che

$$\tilde{\mathbf{x}} = [\mathbf{x}_1 \;\; \mathbf{x}_2]\, \tilde{\mathbf{a}} = b \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

è un autovettore di T corrispondente all'autovalore $\widetilde{\lambda} = 1$.

La dimensione dell'autospazio corrispondente a tale autovalore è pari 1.

**5.3** Facendo seguito alle considerazioni esposte nelle due sezioni precedenti, vogliamo anzitutto dimostrare il seguente teorema.

**TEOREMA 2**

> Autovettori corrispondenti ad autovalori *distinti* di una trasformazione lineare T di uno spazio vettoriale S in se stesso sono linearmente indipendenti.

*Dim.* Supponiamo che T ammetta t autovalori distinti $\widetilde{\lambda}_{*1}, \widetilde{\lambda}_{*2}, \dots, \widetilde{\lambda}_{*t}$ e consideriamo t autovettori $\widetilde{\mathbf{x}}_{*1}, \widetilde{\mathbf{x}}_{*2}, \dots, \widetilde{\mathbf{x}}_{*t}$ associati a tali autovalori.

La dimostrazione utilizza il procedimento di induzione su t.

Se $t = 1$ il teorema è senz'altro vero. Sia adesso $t > 1$ e supponiamo che il teorema sia vero per ogni insieme di $t-1$ autovettori.

Considerata la relazione

$$\text{(a)} \qquad c_1 \widetilde{\mathbf{x}}_{*1} + c_2 \widetilde{\mathbf{x}}_{*2} + \dots + c_t \widetilde{\mathbf{x}}_{*t} = \mathbf{0}$$

dove $c_1, c_2, \dots, c_t$ sono t numeri reali per il momento arbitrari, vogliamo dimostrare che la (a) implica che sia $c_1 = c_2 = \dots = c_t = 0$.

Applicando a entrambi i membri della relazione precedente la trasformazione lineare T, si ottiene

$$T(c_1 \widetilde{\mathbf{x}}_{*1} + c_2 \widetilde{\mathbf{x}}_{*2} + \dots + c_t \widetilde{\mathbf{x}}_{*t}) = c_1 T(\widetilde{\mathbf{x}}_{*1}) + c_2 T(\widetilde{\mathbf{x}}_{*2}) + \dots + c_t T(\widetilde{\mathbf{x}}_{*t}) = \mathbf{0}$$

da cui

$$\text{(b)} \qquad c_1 \widetilde{\lambda}_{*1} \widetilde{\mathbf{x}}_{*1} + c_2 \widetilde{\lambda}_{*2} \widetilde{\mathbf{x}}_{*2} + \dots + c_t \widetilde{\lambda}_{*t} \widetilde{\mathbf{x}}_{*t} = \mathbf{0}\ .$$

Moltiplicando la (a) per $\widetilde{\lambda}_{*1}$ e sottraendo il risultato dalla (b), si ha

$$\text{(c)} \qquad c_2 (\widetilde{\lambda}_{*2} - \widetilde{\lambda}_{*1}) \widetilde{\mathbf{x}}_{*2} + \dots + c_t (\widetilde{\lambda}_{*t} - \widetilde{\lambda}_{*1}) \widetilde{\mathbf{x}}_{*t} = \mathbf{0}\ .$$

Ma, essendo per ipotesi $\widetilde{\mathbf{x}}_{*2}, \dots, \widetilde{\mathbf{x}}_{*t}$ linearmente indipendenti e $(\widetilde{\lambda}_{*s} - \widetilde{\lambda}_{*1}) \neq 0$ per s = 2, ..., t, la (c) implica che sia $c_2 = \dots = c_t = 0$.

Ne segue, come si rileva subito dalla (a), che $c_1 \widetilde{\mathbf{x}}_{*1} = \mathbf{0}$ e, quindi, che anche $c_1 = 0$.

**COROLLARIO.** Una trasformazione lineare T di uno spazio vettoriale S di dimensione p in se stesso ammette al più p autovalori distinti.

Infatti, se esistessero p+1 autovalori distinti di T, allora, per il Teorema 2, S conterrebbe p+1 autovettori linearmente indipendenti, contro l'ipotesi che sia $\dim(S) = p$. ▮

Si dice che l'autovalore $\widetilde{\lambda}_{*s}$ $(1 \leq s \leq t)$ ha *molteplicità algebrica* (o, più semplicemente, *molteplicità*) $m_s$ se questa è la sua molteplicità quale radice del polinomio caratteristico di T.

Se $m_s = 1$, $\widetilde{\lambda}_{*s}$ è detto autovalore *semplice*.

Ciò premesso, dimostriamo il seguente teorema.

**TEOREMA 3**

Data una trasformazione lineare T di uno spazio vettoriale S di dimensione p in se stesso, sia $\mathbf{T}$ la matrice di T rispetto a una base di S.

Supponiamo, inoltre, che T ammetta gli autovalori distinti $\widetilde{\lambda}_{*1}, \dots, \widetilde{\lambda}_{*t}$ di molteplicità algebriche $m_1, \dots, m_t$ tali che

$$\sum_{s=1}^{t} m_s = p\ .$$

Condizione necessaria e sufficiente affinché esista una base di S formata da autovettori di T è che, per ogni s = 1, ..., t, sia

$$r(\mathbf{T} - \widetilde{\lambda}_{*s}\mathbf{I}) = p - m_s\ , \tag{6}$$

ovvero che la molteplicità algebrica di $\widetilde{\lambda}_{*s}$ sia eguale alla sua molteplicità geometrica.

*Dim.* Chiaramente, la (6) esprime una condizione necessaria e sufficiente affinché lo spazio vettoriale generato dalle soluzioni del sistema

$$(\mathbf{T} - \tilde{\lambda}_{*s}\mathbf{I})\,\mathbf{a}_{*s} = \mathbf{0}$$

abbia dimensione pari a $m_s$.

Pertanto, scelti $m_s$ vettori linearmente indipendenti ($1 \le s \le t$)

$$\tilde{\mathbf{a}}_{*s,1}\,,\,\dots\,,\,\tilde{\mathbf{a}}_{*s,m_s}$$

appartenenti a tale spazio vettoriale, premoltiplicando ciascuno di essi per la matrice $\mathbf{X}$ i cui vettori colonna formano una base di S, si ottengono $m_s$ autovettori, a loro volta linearmente indipendenti,

$$\tilde{\mathbf{x}}_{*s,1} = \mathbf{X}\,\tilde{\mathbf{a}}_{*s,1}\,,\,\dots\,,\,\tilde{\mathbf{x}}_{*s,m_s} = \mathbf{X}\,\tilde{\mathbf{a}}_{*s,m_s}$$

i quali formano una base dell'autospazio di $\tilde{\lambda}_{*s}$.

Poiché autovettori associati ad autovalori distinti sono linearmente indipendenti (Cfr. il Teorema 2), riunendo le basi degli autospazi associati agli autovalori $\tilde{\lambda}_{*1}, \dots, \tilde{\lambda}_{*t}$, si ottiene una base di S formata da autovettori di T.

**COROLLARIO.** Qualora la trasformazione lineare T di S in se stesso ammetta p autovalori e questi siano semplici, esiste una base di S formata da autovettori di T.

**ESEMPIO 10.** Si consideri la trasformazione lineare T di $R^2$ in se stesso la cui matrice rispetto alla base formata da

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad , \quad \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

è

$$\mathbf{T} = \begin{bmatrix} 2+\sqrt{2} & \sqrt{2} \\ -1 & 1-\sqrt{2} \end{bmatrix}.$$

L'equazione caratteristica di $\mathbf{T}$ – vale a dire, l'equazione

$$\det\left(\begin{bmatrix} 2+\sqrt{2} & \sqrt{2} \\ -1 & 1-\sqrt{2} \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \det\begin{bmatrix} 2+\sqrt{2}-\lambda & \sqrt{2} \\ -1 & 1-\sqrt{2}-\lambda \end{bmatrix} = 0$$

– ammette gli autovalori $\tilde{\lambda}_{*1} = 3$ e $\tilde{\lambda}_{*2} = 0$, di molteplicità $m_1 = m_2 = 1$.

Poiché $r(\mathbf{T} - \tilde{\lambda}_{*1}\mathbf{I}) = 1$, le soluzioni del sistema

$$\begin{bmatrix} -1+\sqrt{2} & \sqrt{2} \\ -1 & -2-\sqrt{2} \end{bmatrix} \mathbf{a}_{*1} = \mathbf{0}$$

formano uno spazio vettoriale di dimensione 1 e il vettore

$$\tilde{\mathbf{a}}_{*1,1} = b \begin{bmatrix} -\sqrt{2} \\ -1+\sqrt{2} \end{bmatrix},$$

dove b è un numero reale qualsiasi purché diverso da zero, è un vettore non nullo che appartiene a tale spazio vettoriale.

Ne consegue che

$$\tilde{\mathbf{x}}_{*1,1} = \left[\mathbf{x}_1 \;\; \mathbf{x}_2\right] \tilde{\mathbf{a}}_{*1,1} = b \begin{bmatrix} -\sqrt{2} \\ -1 \end{bmatrix}$$

è un autovettore di T corrispondente all'autovalore $\tilde{\lambda}_{*1} = 3$.

Analogamente, poiché $r(\mathbf{T} - \tilde{\lambda}_{*2}\mathbf{I}) = 1$, le soluzioni del sistema

$$\begin{bmatrix} 2+\sqrt{2} & \sqrt{2} \\ -1 & 1-\sqrt{2} \end{bmatrix} \mathbf{a}_{*2} = \mathbf{0}$$

formano uno spazio vettoriale di dimensione 1 e il vettore

$$\tilde{\mathbf{a}}_{*2,1} = c \begin{bmatrix} -\sqrt{2} \\ 2+\sqrt{2} \end{bmatrix},$$

dove c è un numero reale qualsiasi purché diverso da zero, è un vettore non nullo che appartiene a tale spazio vettoriale.

Ne consegue che

$$\tilde{\mathbf{x}}_{*2,1} = [\mathbf{x}_1 \ \ \mathbf{x}_2]\, \tilde{\mathbf{a}}_{*2,1} = c \begin{bmatrix} -\sqrt{2} \\ 2 \end{bmatrix}$$

è un autovettore di T corrispondente all'autovalore $\tilde{\lambda}_{*2} = 0$.

Ovviamente, tenuto conto del Corollario al Teorema 3, gli autovettori di cui sopra formano una base di $R^2$.

**5.4** Supponiamo, come nell'enunciato del Teorema 3, che la trasformazione lineare T di uno spazio vettoriale S di dimensione p in se stesso ammetta t autovalori distinti $\tilde{\lambda}_{*1}, \dots, \tilde{\lambda}_{*t}$ di molteplicità $m_1, \dots, m_t$ con $m_1 + \dots + m_t = p = \dim(S)$.

Indichiamo con $\tilde{\lambda}_1, \dots, \tilde{\lambda}_p$ l'insieme di tali autovalori (eventualmente non tutti distinti) e con $\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_p$ l'insieme dei p autovettori a essi associati.

In una base di S formata da $\mathbf{x}_1, \dots, \mathbf{x}_p$, la relazione $(j = 1, \dots, p)$

$$T(\tilde{\mathbf{x}}_j) = \tilde{\lambda}_j \tilde{\mathbf{x}}_j$$

può essere espressa nella forma

$$\mathbf{T}\, \tilde{\mathbf{a}}_j = \tilde{\lambda}_j\, \tilde{\mathbf{a}}_j$$

da cui – fatte le posizioni

$$\tilde{\mathbf{A}} = [\tilde{\mathbf{a}}_1 \cdots \tilde{\mathbf{a}}_p] \quad , \quad \tilde{\mathbf{D}} = \begin{bmatrix} \tilde{\lambda}_1 & \cdots & 0 \\ \dots & \dots & \dots \\ 0 & \cdots & \tilde{\lambda}_p \end{bmatrix}$$

– si ottiene

$$\mathbf{T}\tilde{\mathbf{A}} = \tilde{\mathbf{A}}\tilde{\mathbf{D}}\,. \tag{7}$$

Ciò premesso, è immediato rilevare che, qualora i vettori $\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_p$ siano linearmente indipendenti, la matrice $\tilde{\mathbf{A}}$ è di pieno rango.

Pertanto,

$$\widetilde{\mathbf{A}}^{-1}\mathbf{T}\widetilde{\mathbf{A}} = \widetilde{\mathbf{D}} \,. \tag{8}$$

Ciò significa che, assunta $\widetilde{\mathbf{A}}$ quale matrice del passaggio dalla base formata da $\mathbf{x}_1, \dots, \mathbf{x}_p$ alla base costituita da $\widetilde{\mathbf{x}}_1, \dots, \widetilde{\mathbf{x}}_p$, la matrice della trasformazione lineare T, in quest'ultima base, è

$$\widetilde{\mathbf{D}} = \mathrm{diag}(\widetilde{\lambda}_1, \dots, \widetilde{\lambda}_p) \,.$$

Reciprocamente, se esiste una matrice $\widetilde{\mathbf{A}}$ tale che la (8) sia soddisfatta, allora $\widetilde{\mathbf{a}}_1, \dots, \widetilde{\mathbf{a}}_p$ sono i vettori coordinati degli autovettori corrispondenti agli autovalori $\widetilde{\lambda}_1, \dots, \widetilde{\lambda}_p$ della trasformazione lineare T associata alla matrice **T**.

**OSSERVAZIONE 13.** È importante sottolineare che, posto che la trasformazione lineare T ammetta p autovalori, la possibilità di ridurre a forma diagonale la matrice **T** di T mediante una trasformazione per similitudine dipende criticamente dal fatto che esistano p autovettori linearmente indipendenti.

Se i p autovalori di T sono tutti *semplici*, allora (Cfr. il Corollario al Teorema 3) è possibile trovare p autovettori linearmente indipendenti e **T** risulta diagonalizzabile attraverso una trasformazione per similitudine.

Se i p autovalori di T *non* sono tutti *semplici*, affinché **T** sia riducibile a forma diagonale mediante una trasformazione per similitudine occorre che la molteplicità algebrica e la molteplicità geometrica di ciascun autovalore siano eguali (Cfr. il Teorema 3).

**ESEMPIO 11.** Riprendendo l'Esempio 10, si consideri la trasformazione lineare T di $R^2$ in se stesso la cui matrice rispetto alla base formata da

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad , \quad \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

è

$$\mathbf{T} = \begin{bmatrix} 2+\sqrt{2} & \sqrt{2} \\ -1 & 1-\sqrt{2} \end{bmatrix}.$$

Come si è visto, tale trasformazione ammette gli autovalori $\tilde{\lambda}_1 = \tilde{\lambda}_{*1} = 3$ e $\tilde{\lambda}_2 = \tilde{\lambda}_{*2} = 0$.

Quindi, $\mathbf{T}$ risulta diagonalizzabile attraverso una trasformazione di similitudine.

In effetti, la matrice del passaggio dalla base formata da $\mathbf{x}_1, \mathbf{x}_2$ alla base costituita dagli autovettori

$$\tilde{\mathbf{x}}_1 = \tilde{\mathbf{x}}_{*1,1} = \begin{bmatrix} -\sqrt{2} \\ -1 \end{bmatrix} \quad , \quad \tilde{\mathbf{x}}_2 = \tilde{\mathbf{x}}_{*1,1} = \begin{bmatrix} -\sqrt{2} \\ 2 \end{bmatrix}$$

è, come subito si verifica,

$$\tilde{\mathbf{A}} = \begin{bmatrix} -\sqrt{2} & -\sqrt{2} \\ -1+\sqrt{2} & 2+\sqrt{2} \end{bmatrix}$$

e risulta

$$\begin{bmatrix} -\sqrt{2} & -\sqrt{2} \\ -1+\sqrt{2} & 2+\sqrt{2} \end{bmatrix}^{-1} \begin{bmatrix} 2+\sqrt{2} & \sqrt{2} \\ -1 & 1-\sqrt{2} \end{bmatrix} \begin{bmatrix} -\sqrt{2} & -\sqrt{2} \\ -1+\sqrt{2} & 2+\sqrt{2} \end{bmatrix}$$
$$= \frac{1}{-3\sqrt{2}} \begin{bmatrix} 2+\sqrt{2} & \sqrt{2} \\ 1-\sqrt{2} & -\sqrt{2} \end{bmatrix} \begin{bmatrix} 2+\sqrt{2} & \sqrt{2} \\ -1 & 1-\sqrt{2} \end{bmatrix} \begin{bmatrix} -\sqrt{2} & -\sqrt{2} \\ -1+\sqrt{2} & 2+\sqrt{2} \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix}.$$

**ESEMPIO 12.** Rispetto alla base naturale di $R^2$, si consideri la trasformazione lineare T di $R^2$ in se stesso la cui matrice è

$$\mathbf{T} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Com'è subito visto, l'equazione caratteristica di $\mathbf{T}$ ammette un unico autovalore $\tilde{\lambda}_{*1} = 0$ di molteplicità $m_1 = 2$ e l'autospazio di $\tilde{\lambda}$ ha dimensione pari a 1.

Pertanto, **T** *non* è diagonalizzabile attraverso una trasformazione di similitudine.

**ESEMPIO 13.** Rispetto alla base naturale di $R^3$, si consideri la trasformazione lineare T di $R^3$ in se stesso la cui matrice è

$$\mathbf{T} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} .$$

Come si può facilmente verificare, l'equazione caratteristica di **T** ammette gli autovalori $\widetilde{\lambda}_{*1} = 3$ di molteplicità $m_1 = 1$ e $\widetilde{\lambda}_{*2} = 0$ di molteplicità $m_2 = 2$. Inoltre, l'autospazio di $\widetilde{\lambda}_{*1}$ ha dimensione pari a 1 e l'autospazio di $\widetilde{\lambda}_{*2}$ ha dimensione pari a 2.

Pertanto, **T** è diagonalizzabile attraverso una trasformazione di similitudine.

**OSSERVAZIONE 14.** Nell'ipotesi che la matrice $\widetilde{\mathbf{A}}$ sia di pieno rango, si possono facilmente dimostrare le seguenti proprietà (6).

(a) Il determinante di **T** è eguale al prodotto degli autovalori della trasformazione lineare T. Infatti,

$$\det \mathbf{T} = \det(\widetilde{\mathbf{A}}^{-1}\mathbf{T}\widetilde{\mathbf{A}}) = \det \widetilde{\mathbf{D}} = \prod_{j=1}^{p} \widetilde{\lambda}_j .$$

(b) La traccia di **T** è eguale alla somma degli autovalori della trasformazione lineare T. Infatti,

$$\mathrm{tr}\,\mathbf{T} = \mathrm{tr}(\widetilde{\mathbf{A}}^{-1}\mathbf{T}\widetilde{\mathbf{A}}) = \mathrm{tr}\,\widetilde{\mathbf{D}} = \sum_{j=1}^{p} \widetilde{\lambda}_j .$$

(c) Il rango di **T** è eguale al numero degli autovalori diversi da zero della trasformazione lineare T, come risulta subito dal fatto che

$$\mathrm{r}(\mathbf{T}) = \mathrm{r}(\widetilde{\mathbf{A}}^{-1}\mathbf{T}\widetilde{\mathbf{A}}) = \mathrm{r}(\widetilde{\mathbf{D}}) .$$

(6) Le proprietà (a) e (b) sussistono anche nel caso in cui tale matrice non sia di pieno rango.

## 6 TRASFORMAZIONI IDEMPOTENTI. PROIETTORI

**6.1** Siano S uno spazio vettoriale di ordine n e dimensione p e T una trasformazione lineare di S in se stesso. Si dice che T è una *trasformazione* (*lineare*) *idempotente* di S se, per ogni $\mathbf{x} \in S$, risulta

$$\text{(9)} \qquad T^2(\mathbf{x}) = T(T(\mathbf{x})) = T(\mathbf{x}) .$$

Poiché, con le notazioni ormai consuete, è

$$T(\mathbf{x}) = \mathbf{XTa} \quad , \quad T^2(\mathbf{x}) = T(T(\mathbf{x})) = \mathbf{XTTa} = \mathbf{XT}^2\mathbf{a}$$

la condizione necessaria e sufficiente affinché valga la (9) è espressa – come facilmente si riconosce – dalla relazione

$$\text{(9')} \qquad \mathbf{T}^2 = \mathbf{T} .$$

In particolare, sono idempotenti sia la trasformazione nulla O ($\mathbf{T} = \mathbf{O}$) sia la trasformazione identica I ($\mathbf{T} = \mathbf{I}$).

**OSSERVAZIONE 15.** Indichiamo con $\mathbf{A}$ la matrice del passaggio da una base costituita dai vettori colonna di $\mathbf{X}$ a una base formata dai vettori colonna di $\widetilde{\mathbf{X}}$. Allora, la matrice $\mathbf{T}$ si trasforma in $\widetilde{\mathbf{T}} = \mathbf{A}^{-1}\mathbf{TA}$ e $\mathbf{T}^2$ si trasforma in $\widetilde{\mathbf{T}}^2 = \mathbf{A}^{-1}\mathbf{TA}\mathbf{A}^{-1}\mathbf{TA} = \mathbf{A}^{-1}\mathbf{TA} = \widetilde{\mathbf{T}}$.

Ciò significa che una trasformazione idempotente è caratterizzata dal fatto che, rispetto a una *qualunque* base di S, la matrice a essa associata è idempotente.

**OSSERVAZIONE 16.** Scritta la (9') nella forma

$$\mathbf{T}(\mathbf{I} - \mathbf{T}) = \mathbf{O} ,$$

è subito visto che o $\mathbf{T} = \mathbf{O}$ e $(\mathbf{I} - \mathbf{T}) = \mathbf{I}$ o $\mathbf{T} = \mathbf{I}$ e $(\mathbf{I} - \mathbf{T}) = \mathbf{O}$, oppure sia $\mathbf{T}$ sia $\mathbf{I} - \mathbf{T}$ sono singolari (Cfr. il punto 9 dei Complementi al Cap. III).

**ESEMPIO 14.** Si consideri la trasformazione lineare T di $R^2$ in se stesso la cui matrice, rispetto alla base naturale di $R^2$, è

$$\mathbf{T} = \begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix}.$$

Come subito si verifica, **T** è idempotente.

Considerata poi la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & -1/2 \\ -1 & 1 \end{bmatrix}$$

del passaggio dalla base naturale di $R^2$ alla base formata da

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad , \quad \mathbf{x}_2 = \begin{bmatrix} -1/2 \\ 1 \end{bmatrix}$$

**T** si trasforma in

$$\tilde{\mathbf{T}} = \begin{bmatrix} 1 & -1/2 \\ -1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1/2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

e la matrice che compare a secondo membro di quest'ultima espressione è chiaramente idempotente (7).

Si noti anche che le matrici **T** e $\mathbf{I} - \mathbf{T}$ sono entrambe singolari, così come le matrici $\tilde{\mathbf{T}}$ e $\mathbf{I} - \tilde{\mathbf{T}}$.

**6.2** Siano S uno spazio vettoriale di dimensione p e $S_1$ e $S_2$ due suoi sottospazi tali che $S = S_1 \oplus S_2$.

Com'è noto (Cfr. il Teorema 13 del Cap. I), ogni $\mathbf{x} \in S$ ammette un'unica rappresentazione della forma $\mathbf{x} = \mathbf{y} + \mathbf{z}$ con $\mathbf{y} \in S_1$ e $\mathbf{z} \in S_2$.

Considerata la funzione T che associa a ogni $\mathbf{x} \in S$ il vettore $\mathbf{y} \in S_1$, vogliamo anzitutto mostrare che T è una trasformazione lineare di S in se stesso.

(7) Come avremo modo di porre in evidenza nel seguito (Cfr. la sezione 6.3 e l'Esempio 15), il fatto che la matrice risultante sia del tipo indicato non è casuale.

Infatti, per tutti gli $\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2 \in S$ tali che

$$\begin{aligned} \mathbf{x} &= \mathbf{y} + \mathbf{z} && \text{con } \mathbf{y} \in S_1 \text{ e } \mathbf{z} \in S_2 \\ \mathbf{x}_1 &= \mathbf{y}_1 + \mathbf{z}_1 && \text{con } \mathbf{y}_1 \in S_1 \text{ e } \mathbf{z}_1 \in S_2 \\ \mathbf{x}_2 &= \mathbf{y}_2 + \mathbf{z}_2 && \text{con } \mathbf{y}_2 \in S_1 \text{ e } \mathbf{z}_2 \in S_2 \end{aligned}$$

e ogni $a \in R$, si ha

(i) $T(\mathbf{x}_1 + \mathbf{x}_2) = T((\mathbf{y}_1 + \mathbf{y}_2) + (\mathbf{z}_1 + \mathbf{z}_2)) = \mathbf{y}_1 + \mathbf{y}_2 = T(\mathbf{x}_1) + T(\mathbf{x}_2)$

(ii) $T(a\mathbf{x}) = T(a\mathbf{y} + a\mathbf{z}) = a\mathbf{y} = aT(\mathbf{x})$.

Inoltre, si verifica facilmente che la trasformazione lineare T qui sopra definita gode delle seguenti proprietà:

(a) $T(\mathbf{y}) = T(\mathbf{y} + \mathbf{0}) = \mathbf{y}$ per tutti gli $\mathbf{y} \in S_1$

(b) $T(\mathbf{z}) = T(\mathbf{0} + \mathbf{z}) = \mathbf{0}$ per tutti gli $\mathbf{z} \in S_2$.

Tali proprietà determinano univocamente T, nel senso che ogni altra trasformazione lineare $T_1$ che possieda le medesime proprietà – risultando

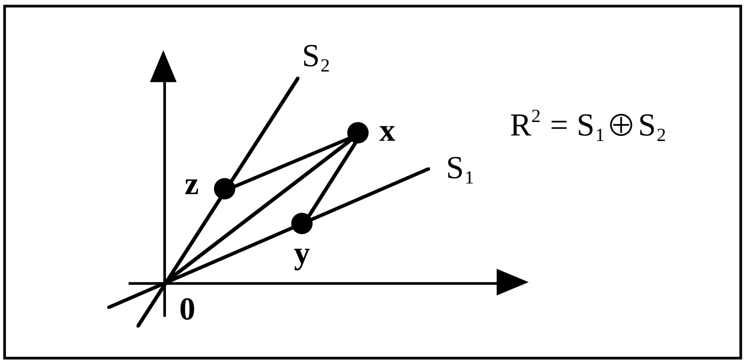
$$T_1(\mathbf{x}) = T_1(\mathbf{y} + \mathbf{z}) = T_1(\mathbf{y}) + T_1(\mathbf{z}) = \mathbf{y} + \mathbf{0} = \mathbf{y} = T(\mathbf{x})$$

– coincide necessariamente con T.

La trasformazione lineare T in questione è anche detta *proiettore* di S *su* $S_1$ *lungo* $S_2$.

A sua volta, il vettore $\mathbf{y}$ è detto *proiezione* di $\mathbf{x}$ *su* $S_1$ *lungo* $S_2$.

Fig. 1



Il seguente teorema stabilisce la relazione che sussiste tra proiettori e trasformazioni idempotenti.

**TEOREMA 4**

Una trasformazione lineare T di uno spazio vettoriale S di dimensione p in se stesso è un proiettore se e soltanto se T è idempotente.

*Dim*. Supponiamo che T sia idempotente. Vogliamo dimostrare che esistono due sottospazi $S_1$ e $S_2$ di S per i quali valgono le proprietà (a) e (b) indicate in precedenza e tali che $S = S_1 \oplus S_2$.

Sia $S_1$ definito come l'immagine $\mathcal{R}(T)$ di S. Se $\mathbf{y} \in S_1$, allora $\mathbf{y} = T(\mathbf{x})$ per qualche $\mathbf{x} \in S$; quindi,

$$T(\mathbf{y}) = T(T(\mathbf{x})) = T^2(\mathbf{x}) = T(\mathbf{x}) = \mathbf{y}$$

e la proprietà (a) risulta soddisfatta.

Sia $S_2$ definito come il nucleo $\mathcal{N}(T)$ di T. Se $\mathbf{z} \in S_2$, allora $T(\mathbf{z}) = \mathbf{0}$; quindi, la proprietà (b) risulta soddisfatta.

Occorre, adesso, fare vedere che $S = S_1 \oplus S_2$. A questo fine, si consideri, per ogni $\mathbf{x} \in S$, l'identità

$$\mathbf{x} = T(\mathbf{x}) + (\mathbf{x} - T(\mathbf{x})) \ .$$

È ovvio che $S = S_1 + S_2$. Inoltre, se per qualche $\mathbf{x} \in S$ accade che $\mathbf{x}$ appartenga sia a $S_1$ sia a $S_2$, deve risultare da un lato $T(\mathbf{x}) = \mathbf{x}$, dall'altro $T(\mathbf{x}) = \mathbf{0}$. Ma ciò implica che sia $\mathbf{x} = \mathbf{0}$ e, quindi, che $S_1 \cap S_2 = S_0$.

Viceversa, supponiamo che T sia un proiettore di S su $S_1$ lungo $S_2$. Allora, per ogni $\mathbf{x} \in S$ tale che $\mathbf{x} = \mathbf{y} + \mathbf{z}$ con $\mathbf{y} \in S_1$ e $\mathbf{z} \in S_2$, risulta

$$T^2(\mathbf{x}) = T(T(\mathbf{x})) = T(T(\mathbf{y}) + T(\mathbf{z})) = T(\mathbf{y}) = \mathbf{y} = T(\mathbf{x})$$

e, quindi, T è idempotente.

**6.3** Considerati uno spazio vettoriale S di dimensione p e due suoi sottospazi $S_1$ e $S_2$ tali che $S = S_1 \oplus S_2$, nella sezione precedente abbiamo definito il proiettore T di S su $S_1$ lungo $S_2$.

Supponiamo adesso che $\mathbf{x}_1, \ldots, \mathbf{x}_{p1}$ siano una base di $S_1 = \mathcal{R}(T)$ e, a loro volta, $\mathbf{x}_{p1+1}, \ldots, \mathbf{x}_p$ siano una base di $S_2 = \mathcal{N}(T)$.

Rispetto alla base di S formata da $\mathbf{x}_1, \ldots, \mathbf{x}_{p1}, \mathbf{x}_{p1+1}, \ldots, \mathbf{x}_p$, per ogni $\mathbf{x} \in S$, possiamo scrivere

$$\mathbf{x} = a_1 \mathbf{x}_1 + \ldots + a_{p1} \mathbf{x}_{p1} + a_{p1+1} \mathbf{x}_{p1+1} + \ldots + a_p \mathbf{x}_p$$

da cui

$$\begin{aligned} T(\mathbf{x}) &= T(a_1 \mathbf{x}_1 + \ldots + a_{p1} \mathbf{x}_{p1} + a_{p1+1} \mathbf{x}_{p1+1} + \ldots + a_p \mathbf{x}_p) \\ &= a_1 T(\mathbf{x}_1) + \ldots + a_{p1} T(\mathbf{x}_{p1}) + a_{p1+1} T(\mathbf{x}_{p1+1}) + \ldots + a_p T(\mathbf{x}_p) \\ &= a_1 \mathbf{x}_1 + \ldots + a_{p1} \mathbf{x}_{p1} + a_{p1+1} \mathbf{0} + \ldots + a_p \mathbf{0} \\ &= a_1 \mathbf{x}_1 + \ldots + a_{p1} \mathbf{x}_{p1} + 0\, \mathbf{x}_{p1+1} + \ldots + 0\, \mathbf{x}_p \\ &= [\mathbf{x}_1 \cdots \mathbf{x}_{p1}\ \mathbf{x}_{p1+1} \cdots \mathbf{x}_p] \begin{bmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_{p1} \\ a_{p1+1} \\ \vdots \\ a_p \end{bmatrix} . \end{aligned}$$

Quindi, la matrice rappresentativa di T, relativa alla base di S qui sopra indicata, è

$$\mathbf{D} = \begin{bmatrix} \mathbf{I}_{(p1,p1)} & \mathbf{O}_{(p1,p2)} \\ \mathbf{O}_{(p2,p1)} & \mathbf{O}_{(p2,p2)} \end{bmatrix} .$$

Ovviamente,

$$r(\mathbf{D}) = p_1 = \dim(S_1) = \mathrm{tr}(\mathbf{D}) \ , \ r(\mathbf{I} - \mathbf{D}) = p_2 = \dim(S_2) = \mathrm{tr}(\mathbf{I} - \mathbf{D}) .$$

Ricordando che il rango e la traccia di una matrice rimangono invariati rispetto a una trasformazione per similitudine, si conclude poi che la matrice $\mathbf{T}$ del proiettore T, rispetto a una base qualsiasi di S, è tale che

$$r(\mathbf{T}) = p_1 = \dim(S_1) = \mathrm{tr}(\mathbf{T}) \ , \ r(\mathbf{I} - \mathbf{T}) = p_2 = \dim(S_2) = \mathrm{tr}(\mathbf{I} - \mathbf{T}) .$$

**OSSERVAZIONE 18.** Chiaramente, ciascun vettore della base di $S_1 = \mathcal{R}(T)$ è un autovettore di S corrispondente all'autovalore $\widetilde{\lambda}_{*1} = 1$ e ogni vettore della base di $S_2 = \mathcal{N}(T)$ è un autovettore di S associato all'autovalore $\widetilde{\lambda}_{*2} = 0$.

**ESEMPIO 15.** Si consideri la trasformazione idempotente T di $R^2$ la cui matrice, rispetto alla base naturale di $R^2$, è (Cfr. l'Esempio 14)

$$\mathbf{T} = \begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix}.$$

Come facilmente si verifica, una base di $S_1 = \mathcal{R}(R^2)$ è fornita dal vettore

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

e una base di $S_2 = \mathcal{N}(T)$ è data dal vettore

$$\mathbf{x}_2 = \begin{bmatrix} -1/2 \\ 1 \end{bmatrix}.$$

Si noti che $R^2 = S_1 \oplus S_2 = \mathcal{R}(R^2) \oplus \mathcal{N}(T)$.

Inoltre, rispetto alla base di $R^2$ costituita dai vettori $\mathbf{x}_1, \mathbf{x}_2$ di cui sopra, risulta che la matrice di T è

$$\mathbf{D} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

**6.4** Date la trasformazione identica I e una trasformazione idempotente T di uno spazio vettoriale S, si consideri la funzione $\overline{T} = (I - T)$ che associa a ciascun $\mathbf{x} \in S$ il vettore $\overline{T}(\mathbf{x}) = (I - T)(\mathbf{x})$.

Si verifica subito che $\overline{T}$ è una trasformazione lineare e idempotente.

Infatti, si ha anzitutto che $(\mathbf{x}, \mathbf{y} \in S;\ a \in R)$

$$\begin{aligned} \overline{T}(\mathbf{x} + \mathbf{y}) &= (I - T)(\mathbf{x} + \mathbf{y}) = \mathbf{x} - T(\mathbf{x}) + \mathbf{y} - T(\mathbf{y}) = (I - T)(\mathbf{x}) + (I - T)(\mathbf{y}) \\ \overline{T}(a\mathbf{x}) &= (I - T)(a\mathbf{x}) = a\mathbf{x} - aT(\mathbf{x}) = a(I - T)(\mathbf{x}) . \end{aligned}$$

Inoltre, risulta che

$$\overline{\mathrm{T}}^2(\mathbf{x}) = (\mathrm{I} - \mathrm{T})\,(\mathrm{I} - \mathrm{T})\,(\mathbf{x})) = (\mathrm{I} - \mathrm{T})(\mathbf{x})\ .$$

Come si riconosce subito, la funzione $\overline{\mathrm{T}}$ in tal modo definita costituisce il *proiettore* di S *su* $S_2$ *lungo* $S_1$.

A sua volta, il vettore **z** è detto la *proiezione* di **x** *su* $S_2$ *lungo* $S_1$.

Naturalmente, se, rispetto a una prefissata base di S, **T** denota la matrice associata a T, **I** – **T** è la matrice associata a $\overline{\mathrm{T}}$ .

Inoltre, il rango e la nullità di $\overline{\mathrm{T}}$ sono eguali, rispettivamente, alla nullità e al rango di T.

# Cap. V
# COMPLEMENTI

**1**

Supponiamo che T sia una trasformazione lineare di S in Z e che $\mathbf{x}_1, \dots, \mathbf{x}_t$ siano t vettori appartenenti a S.

(a) Se $\mathbf{x}_1, \dots, \mathbf{x}_t$ sono linearmente dipendenti anche $T(\mathbf{x}_1), \dots, T(\mathbf{x}_t)$ sono linearmente dipendenti. Infatti, dalla relazione

$$a_1\mathbf{x}_1 + \dots + a_t\mathbf{x}_t = \mathbf{0} \ ,$$

con $a_1, \dots, a_t$ numeri reali non tutti nulli, si ha

$$T(a_1\mathbf{x}_1 + \dots + a_t\mathbf{x}_t) = a_1T(\mathbf{x}_1) + \dots + a_tT(\mathbf{x}_t) = \mathbf{0} \ .$$

Ovviamente, qualora $T(\mathbf{x}_1), \dots, T(\mathbf{x}_t)$ siano linearmente indipendenti anche $\mathbf{x}_1, \dots, \mathbf{x}_t$ sono linearmente indipendenti.

(b) Se $\mathbf{x}_1, \dots, \mathbf{x}_t$ sono linearmente indipendenti, nulla si può dire in generale circa le loro immagini $T(\mathbf{x}_1), \dots, T(\mathbf{x}_t)$; queste possono risultare linearmente indipendenti o linearmente dipendenti. Tuttavia, se il nucleo di T è costituito dallo spazio nullo, si può facilmente dimostrare che $T(\mathbf{x}_1), \dots, T(\mathbf{x}_t)$ sono linearmente indipendenti.

**2**

Rispetto a una prefissata base di uno spazio vettoriale S, sia $\mathbf{T}$ la matrice di una trasformazione lineare T di S in se stesso.

Definita la trasformazione lineare T' di S in se stesso la cui matrice, rispetto alla suddetta base di S, è data da **T**', poiché risulta

$$\det(\mathbf{T} - \lambda\mathbf{I}) = \det(\mathbf{T} - \lambda\mathbf{I})' = \det(\mathbf{T}' - \lambda\mathbf{I}) ,$$

le trasformazioni lineari T e T' ammettono i medesimi autovalori.

**3**

Rispetto a una prefissata base di uno spazio vettoriale S di dimensione p, sia **T** la matrice di una trasformazione lineare T di S in se stesso.

Supposto **T** invertibile, si definisca la trasformazione lineare $T^{-1}$ di S in se stesso la cui matrice, rispetto alla suddetta base di S, è data da $\mathbf{T}^{-1}$.

Se $\tilde{\lambda} \neq 0$ è un autovalore di T, allora $\tilde{\lambda}^{-1} \neq 0$ è un autovalore di $T^{-1}$.

Infatti,

$$\begin{aligned} 0 &= \det(-\tilde{\lambda}\mathbf{I} + \mathbf{T}) \qquad = (-\tilde{\lambda}^{-1})^p \det\mathbf{T}^{-1}\det(-\tilde{\lambda}\mathbf{I} + \mathbf{T}) \\ &= (-\tilde{\lambda}^{-1})^p \det(-\tilde{\lambda}\mathbf{T}^{-1} + \mathbf{I}) = \det\big((-\tilde{\lambda}^{-1})(-\tilde{\lambda}\mathbf{T}^{-1} + \mathbf{I})\big) \\ &= \det(\mathbf{T}^{-1} - \tilde{\lambda}^{-1}\mathbf{I}) . \end{aligned}$$

Si dimostra poi che il sistema $(-\tilde{\lambda}\mathbf{I} + \mathbf{T})\mathbf{a} = \mathbf{0}$ e l'altro $(\mathbf{T}^{-1} - \tilde{\lambda}^{-1}\mathbf{I})\mathbf{a} = \mathbf{0}$ ammettono la medesima soluzione $\tilde{\mathbf{a}}$.

**4**

Date le matrici **X** di ordine $(p_1, p_2)$ e **Y** di ordine $(p_2, p_1)$, vogliamo anzitutto mostrare che, qualunque sia $\lambda \in R$, risulta (**I** di ordine appropriato)

$$\lambda^{p_2}\det(\lambda\mathbf{I} - \mathbf{XY}) = \lambda^{p_1}\det(\lambda\mathbf{I} - \mathbf{YX}) .$$

Infatti, considerata l'identità (**O** di ordine appropriato)

$$\begin{bmatrix} \lambda\mathbf{I} - \mathbf{XY} & \mathbf{X} \\ \mathbf{O} & \lambda\mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{Y} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{Y} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \lambda\mathbf{I} & \mathbf{X} \\ \mathbf{O} & \lambda\mathbf{I} - \mathbf{YX} \end{bmatrix} ,$$

prendendo il determinante di entrambi i membri, segue direttamente quanto asserito.

Ciò premesso, posto $p_1 \leq p_2$, supponiamo che $\mathbf{XY}$ sia la matrice rappresentativa, rispetto a una prefissata base di $R^{p_1}$, di una trasformazione lineare $T_1$ di $R^{p_1}$ in se stesso.

Analogamente, supponiamo che $\mathbf{YX}$ sia la matrice rappresentativa, rispetto a una prefissata base di $R^{p_2}$, di una trasformazione lineare $T_2$ di $R^{p_2}$ in se stesso.

Allora, nell'ipotesi che $T_1$ ammetta complessivamente h $(0 < h \leq p_1)$ autovalori diversi da zero, $T_2$ ammette i medesimi h autovalori diversi da zero.

Infatti, se $\tilde{\lambda} \neq 0$ è un autovalore di $T_1$, si ha che

$$\{\det(\tilde{\lambda}\mathbf{I} - \mathbf{XY}) = 0\} \Rightarrow \{\tilde{\lambda}^{p_2} \det(\tilde{\lambda}\mathbf{I} - \mathbf{XY}) = 0\}$$

$$\Rightarrow \{\tilde{\lambda}^{p_1} \det(\tilde{\lambda}\mathbf{I} - \mathbf{YX}) = 0\} \Rightarrow \{\det(\tilde{\lambda}\mathbf{I} - \mathbf{YX}) = 0\} .$$

Inoltre, come si può agevolmente verificare, se $\tilde{\lambda} = 0$ è un autovalore di $T_1$ di molteplicità pari a $p_1 - h$, allora $\tilde{\lambda} = 0$ è un autovalore di $T_2$ di molteplicità pari a $p_2 - h$.

**5**

Facendo seguito alla considerazioni esposte al punto precedente, supponiamo che $\tilde{\mathbf{a}}, \tilde{\mathbf{b}}$ siano i vettori coordinati degli autovettori di $T_1, T_2$ associati a un autovalore $\tilde{\lambda} \neq 0$.

Vogliamo mostrare che $\tilde{\mathbf{a}}, \tilde{\mathbf{b}}$ sono legati dalle relazioni

$$\mathbf{Y}\tilde{\mathbf{a}} = \tilde{\mathbf{b}} \quad , \quad \mathbf{X}\tilde{\mathbf{b}} = \tilde{\mathbf{a}} .$$

Infatti, se $\tilde{\mathbf{a}}$ è una soluzione non nulla dell'equazione $\mathbf{XYa} = \tilde{\lambda}\mathbf{a}$, allora $\mathbf{Y}\tilde{\mathbf{a}}$ è una soluzione non nulla ($\tilde{\lambda}\tilde{\mathbf{a}} \neq \mathbf{0}$ implica $\mathbf{Y}\tilde{\mathbf{a}} \neq \mathbf{0}$) dell'equazione $\mathbf{YXYa} = \tilde{\lambda}\mathbf{Ya}$, cioè $\mathbf{Y}\tilde{\mathbf{a}} = \tilde{\mathbf{b}}$, e viceversa.

**6**

Siano S uno spazio vettoriale di ordine n e dimensione p e Z uno spazio vettoriale di ordine m e dimensione q.

Considerate due trasformazioni lineari $T_1$ e $T_2$ di S in Z e un numero reale c, si definiscono la *somma* $T_1 + T_2$ e il *prodotto* $cT_1$ ponendo, per ogni $\mathbf{x} \in S$ e ogni $c \in R$,

$$(T_1 + T_2)(\mathbf{x}) = T_1(\mathbf{x}) + T_2(\mathbf{x}) \quad , \quad (cT_1)(\mathbf{x}) = cT_1(\mathbf{x}) \ .$$

Chiaramente, l'insieme delle trasformazioni lineari di S in Z, con le operazioni di *addizione* e di *moltiplicazione per un numero reale* qui sopra definite, costituisce uno spazio vettoriale su R (Cfr. il punto 6 dei Complementi al Cap. I).

Tale spazio vettoriale è solitamente indicato con $\mathrm{Hom}(S, Z)$. Nel caso particolare in cui sia Z = S, in luogo di $\mathrm{Hom}(S, S)$ si usa di preferenza il simbolo $\mathrm{End}(S)$ (1).

Si noti che, fissate le basi di S e Z, sussiste una corrispondenza biunivoca tra $\mathrm{Hom}(S, Z)$ e l'insieme M delle matrici di ordine $(q, p)$ (Cfr. l'Osservazione 5).

Ma, poiché a quest'ultimo insieme M può essere attribuita una struttura di spazio vettoriale su R di dimensione $q \cdot p$ (Cfr. il punto 6 dei Complementi al Cap. II), si riconosce subito che $\mathrm{Hom}(S, Z)$ e M sono isomorfi (Cfr. il punto 7 dei Complementi al Cap. I) e che

$$\dim(\mathrm{Hom}(S, Z)) = q \cdot p \ .$$

(1) Hom sta per *omomorfismo* e End sta per *endomorfismo*.

# Cap. VI
# FORME BILINEARI E QUADRATICHE

## 1 FORME BILINEARI

Dato uno spazio vettoriale S di ordine n e dimensione p, si chiama *forma bilineare su* S ogni funzione B che associa a ciascuna coppia ordinata formata dai vettori $\mathbf{x},\mathbf{y}\in S$ un numero reale $B(\mathbf{x},\mathbf{y})$ in modo tale che siano verificate le proprietà seguenti ($\mathbf{x},\mathbf{y},\mathbf{v},\mathbf{z}\in S$; $a,b\in R$):

(i) $\quad B(\mathbf{x}+\mathbf{v},\mathbf{y}) = B(\mathbf{x},\mathbf{y}) + B(\mathbf{v},\mathbf{y})$ , $B(a\mathbf{x},\mathbf{y}) = a\, B(\mathbf{x},\mathbf{y})$

(ii) $\quad B(\mathbf{x},\mathbf{y}+\mathbf{z}) = B(\mathbf{x},\mathbf{y}) + B(\mathbf{x},\mathbf{z})$ , $B(\mathbf{x},b\mathbf{y}) = b\, B(\mathbf{x},\mathbf{y})$ .

Le proprietà di cui sopra esprimono la *linearità* di una forma bilineare nei riguardi, rispettivamente, del primo e del secondo argomento della funzione.

**ESEMPIO 1.** Come si verifica facilmente, la funzione B definita ponendo, per ogni coppia ordinata formata dai vettori $\mathbf{x},\mathbf{y}\in R^2$,

$$B(\mathbf{x},\mathbf{y}) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

è una forma bilineare su $R^2$.

## 2 FORME BILINEARI E BASI

**2.1** Data una forma bilineare B su S, supponiamo che i vettori $\mathbf{x}_1,\dots,\mathbf{x}_p$ costituiscano una prefissata base di S.

Considerata la coppia ordinata formata dai vettori

$$\mathbf{x} = a_1\mathbf{x}_1 + ... + a_p\mathbf{x}_p = \begin{bmatrix}\mathbf{x}_1 \cdots \mathbf{x}_p\end{bmatrix}\begin{bmatrix}a_1\\ \vdots\\ a_p\end{bmatrix} = \mathbf{Xa}$$

$$\mathbf{y} = b_1\mathbf{x}_1 + ... + b_p\mathbf{x}_p = \begin{bmatrix}\mathbf{x}_1 \cdots \mathbf{x}_p\end{bmatrix}\begin{bmatrix}b_1\\ \vdots\\ b_p\end{bmatrix} = \mathbf{Xb}$$

appartenenti a S, risulta

(1)
$$\begin{aligned} B(\mathbf{x},\mathbf{y}) &= B(a_1\mathbf{x}_1 + ... + a_p\mathbf{x}_p\,,\, b_1\mathbf{x}_1 + ... + b_p\mathbf{x}_p) \\ &= \sum_{k=1}^{p}\sum_{j=1}^{p} a_k\, B(\mathbf{x}_k,\mathbf{x}_j)\, b_j = \sum_{k=1}^{p}\sum_{j=1}^{p} a_k c_{kj} b_j \\ &= \begin{bmatrix}a_1 \cdots a_p\end{bmatrix}\begin{bmatrix}c_{11} \cdots c_{1p}\\ \cdots\cdots\cdots\\ c_{p1} \cdots c_{pp}\end{bmatrix}\begin{bmatrix}b_1\\ \vdots\\ b_p\end{bmatrix} = \mathbf{a}'\mathbf{C}\mathbf{b}\,. \end{aligned}$$

Mediante la (1) il valore $B(\mathbf{x},\mathbf{y})$ assunto dalla forma bilineare B, in corrispondenza della coppia ordinata formata dai vettori di cui sopra, è espresso in funzione dei vettori coordinati $\mathbf{a}$ di $\mathbf{x}$ e $\mathbf{b}$ di $\mathbf{y}$ nonché della matrice $\mathbf{C}$.

Quest'ultima matrice, univocamente associata a B e alla suddetta base di S, riceve la denominazione di *matrice rappresentativa* o, più semplicemente, *matrice* di B rispetto a tale base di S.

**ESEMPIO 2.** Rispetto alla base naturale di $R^2$, la matrice della forma bilineare B definita nell'Esempio 1 è

$$\mathbf{C} = \begin{bmatrix}1 & 1\\ 0 & -1\end{bmatrix}.$$

Invece, rispetto alla base di $R^2$ costituita dai vettori

$$\mathbf{x}_1 = \begin{bmatrix}1\\ 1\end{bmatrix} \quad , \quad \mathbf{x}_2 = \begin{bmatrix}2\\ 1\end{bmatrix}$$

– risultando

$$B(\mathbf{x}_1,\mathbf{x}_1) = 1 \quad , \quad B(\mathbf{x}_1,\mathbf{x}_2) = 2 \quad , \quad B(\mathbf{x}_2,\mathbf{x}_1) = 3 \quad , \quad B(\mathbf{x}_2,\mathbf{x}_2) = 5$$

– la matrice rappresentativa della stessa forma bilineare B è

$$\mathbf{C} = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}.$$

**2.2** Supponiamo, come in precedenza, che i vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$ formino una prefissata base di S.

Scelti *arbitrariamente* i numeri reali $c_{11}, \dots, c_{1p}, \dots, c_{p1}, \dots, c_{pp}$, sia

$$\mathbf{C} = \begin{bmatrix} c_{11} \cdots c_{1p} \\ \cdots\cdots\cdots \\ c_{p1} \cdots c_{pp} \end{bmatrix}.$$

Considerata la coppia ordinata formata dai vettori

$$\mathbf{x} = a_1\mathbf{x}_1 + \dots + a_p\mathbf{x}_p = \begin{bmatrix} \mathbf{x}_1 \cdots \mathbf{x}_p \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix} = \mathbf{Xa}$$

$$\mathbf{y} = b_1\mathbf{x}_1 + \dots + b_p\mathbf{x}_p = \begin{bmatrix} \mathbf{x}_1 \cdots \mathbf{x}_p \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_p \end{bmatrix} = \mathbf{Xb}$$

appartenenti a S, si definisca la funzione B ponendo

$$(1') \qquad B(\mathbf{x},\mathbf{y}) = \mathbf{a}'\mathbf{C}\mathbf{b}.$$

Si riconosce subito che B è una forma bilineare su S univocamente associata alla matrice **C** (1) e che, rispetto alla suddetta base di S, la matrice rappresentativa di B è proprio **C**.

**OSSERVAZIONE 1.** Tenuto conto di quanto mostrato in questa e nella sezione precedente, possiamo affermare che esiste una corrispondenza biunivoca,

---

(1) Ovvero, univocamente associata ai numeri reali $c_{11}, \dots, c_{1p}, \dots, c_{p1}, \dots, c_{pp}$ tali che

$$B(\mathbf{x}_1, \mathbf{x}_1) = c_{11}, \dots, B(\mathbf{x}_1, \mathbf{x}_p) = c_{1p}, \dots, B(\mathbf{x}_p, \mathbf{x}_1) = c_{p1}, \dots, B(\mathbf{x}_p, \mathbf{x}_p) = c_{pp}.$$

dipendente dalla prescelta base di S, tra l'insieme delle forme bilineari su S e l'insieme delle matrici di ordine appropriato.

**2.3** Supponiamo adesso che $\widetilde{\mathbf{x}}_1, \dots, \widetilde{\mathbf{x}}_p$ formino una nuova base di S.

Posto

$$\mathbf{x} = \widetilde{a}_1 \widetilde{\mathbf{x}}_1 + \dots + \widetilde{a}_p \widetilde{\mathbf{x}}_p = \begin{bmatrix} \widetilde{\mathbf{x}}_1 \cdots \widetilde{\mathbf{x}}_p \end{bmatrix} \begin{bmatrix} \widetilde{a}_1 \\ \vdots \\ \widetilde{a}_p \end{bmatrix} = \widetilde{\mathbf{X}}\,\widetilde{\mathbf{a}}$$

$$\mathbf{y} = \widetilde{b}_1 \widetilde{\mathbf{x}}_1 + \dots + \widetilde{b}_p \widetilde{\mathbf{x}}_p = \begin{bmatrix} \widetilde{\mathbf{x}}_1 \cdots \widetilde{\mathbf{x}}_p \end{bmatrix} \begin{bmatrix} \widetilde{b}_1 \\ \vdots \\ \widetilde{b}_p \end{bmatrix} = \widetilde{\mathbf{X}}\,\widetilde{\mathbf{b}}$$

risulta

$$\begin{aligned} \text{(2)} \qquad B(\mathbf{x},\mathbf{y}) &= B(\widetilde{a}_1\widetilde{\mathbf{x}}_1 + \dots + \widetilde{a}_p\widetilde{\mathbf{x}}_p\,,\, \widetilde{b}_1\widetilde{\mathbf{x}}_1 + \dots + \widetilde{b}_p\widetilde{\mathbf{x}}_p) \\ &= \sum_{k=1}^{p}\sum_{j=1}^{p} \widetilde{a}_k\, B(\widetilde{\mathbf{x}}_k, \widetilde{\mathbf{x}}_j)\, \widetilde{b}_j = \sum_{k=1}^{p}\sum_{j=1}^{p} \widetilde{a}_k \widetilde{c}_{kj} \widetilde{b}_j \\ &= \begin{bmatrix} \widetilde{a}_1 \dots \widetilde{a}_p \end{bmatrix} \begin{bmatrix} \widetilde{c}_{11} \cdots \widetilde{c}_{1p} \\ \dots\dots\dots \\ \widetilde{c}_{p1} \cdots \widetilde{c}_{pp} \end{bmatrix} \begin{bmatrix} \widetilde{b}_1 \\ \vdots \\ \widetilde{b}_p \end{bmatrix} = \widetilde{\mathbf{a}}'\widetilde{\mathbf{C}}\widetilde{\mathbf{b}}\,. \end{aligned}$$

Indichiamo con **A** la matrice del passaggio dalla base iniziale di S formata da $\mathbf{x}_1, \dots, \mathbf{x}_p$, vettori colonna della matrice **X**, alla nuova base costituita da $\widetilde{\mathbf{x}}_1, \dots, \widetilde{\mathbf{x}}_p$, vettori colonna della matrice $\widetilde{\mathbf{X}} = \mathbf{X}\mathbf{A}$.

Allora, è subito visto che, dall'eguaglianza

$$\mathbf{a}'\mathbf{C}\,\mathbf{b} = \widetilde{\mathbf{a}}'\widetilde{\mathbf{C}}\,\widetilde{\mathbf{b}}$$

– essendo $\mathbf{a} = \mathbf{A}\widetilde{\mathbf{a}}$ e $\mathbf{b} = \mathbf{A}\widetilde{\mathbf{b}}$ (Cfr. il paragrafo 5 del Cap. IV) – si ottiene

$$\widetilde{\mathbf{a}}'\mathbf{A}'\mathbf{C}\mathbf{A}\widetilde{\mathbf{b}} = \widetilde{\mathbf{a}}'\widetilde{\mathbf{C}}\widetilde{\mathbf{b}}$$

ovvero

$$\widetilde{\mathbf{a}}'(\mathbf{A}'\mathbf{C}\mathbf{A} - \widetilde{\mathbf{C}})\,\widetilde{\mathbf{b}} = 0\,.$$

Ora, quest'ultima relazione deve valere, in particolare, per ciascuna coppia ordinata formata dai vettori canonici $\mathbf{u}_k, \mathbf{u}_j \in R^p$ $(k, j = 1, \dots, p)$.

Ne consegue che

$$\mathbf{A}'\mathbf{C}\mathbf{A} - \widetilde{\mathbf{C}} = \mathbf{O}$$

ossia che

$$\widetilde{\mathbf{C}} = \mathbf{A}'\mathbf{C}\mathbf{A}\,. \tag{3}$$

In generale, due matrici quadrate $\widetilde{\mathbf{C}}$ e $\mathbf{C}$, dello stesso ordine, legate mediante una matrice non singolare $\mathbf{A}$ da una relazione del tipo scritto in (3), si dicono *congruenti* e il legame tra $\widetilde{\mathbf{C}}$ e $\mathbf{C}$ espresso dalla (3) si chiama *trasformazione per congruenza*.

Possiamo pertanto concludere che la matrice $\mathbf{C}$ della forma bilineare B, quando cambia la base, è soggetta a una trasformazione per congruenza.

Viceversa, due matrici $\widetilde{\mathbf{C}}$ e $\mathbf{C}$, legate da una relazione del tipo scritto in (3), si possono sempre interpretare come matrici della *stessa* forma bilineare B, relativamente a due opportune basi di S.

Essendo poi

$$r(\widetilde{\mathbf{C}}) = r(\mathbf{A}'\mathbf{C}\mathbf{A}) = r(\mathbf{C})\ ,$$

possiamo definire il *rango* di una forma bilineare B come il rango comune a ciascuna delle matrici rappresentative di B.

**ESEMPIO 3.** Rispetto alla base naturale di $R^2$, sia

$$\mathbf{C} = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$$

la matrice di una forma bilineare B su $R^2$.

Rispetto alla base di $R^2$ formata dai vettori

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \ , \ \ \mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

la matrice della stessa forma bilineare B, come subito si verifica, diventa

$$\widetilde{\mathbf{C}} = \begin{bmatrix} 4 & 5 \\ 6 & 7 \end{bmatrix}.$$

**OSSERVAZIONE 2.** Si sottolinea che, rispetto a una prefissata base di uno spazio vettoriale S, sia una trasformazione lineare di S in se stesso sia una forma bilineare su S sono rappresentate da una matrice quadrata, diciamo $\mathbf{E}$.

Tuttavia, a un cambiamento della base determinato da una matrice $\mathbf{A}$, nel primo caso $\mathbf{E}$ si trasforma per similitudine ($\mathbf{E} \to \mathbf{A}^{-1}\mathbf{E}\,\mathbf{A}$), nel secondo caso si trasforma per congruenza ($\mathbf{E} \to \mathbf{A}'\mathbf{E}\,\mathbf{A}$) e, in generale, $\mathbf{A}^{-1}\mathbf{E}\,\mathbf{A} \neq \mathbf{A}'\mathbf{E}\,\mathbf{A}$, salvo nel caso in cui $\mathbf{A}$ sia tale che $\mathbf{A}^{-1} = \mathbf{A}'$.

## 3 FORME BILINEARI SIMMETRICHE E FORME QUADRATICHE

**3.1** Una forma bilineare B su uno spazio vettoriale S di ordine n e dimensione p si dice *simmetrica* se, per tutti gli $\mathbf{x},\mathbf{y} \in \mathrm{S}$, risulta

$$\mathrm{B}(\mathbf{x},\mathbf{y}) = \mathrm{B}(\mathbf{y},\mathbf{x}) \tag{4}$$

ovvero se il valore da essa assunto non dipende dall'ordine con cui si susseguono i suoi argomenti.

Ora, con le notazioni introdotte in precedenza, si ha che

$$\mathrm{B}(\mathbf{x},\mathbf{y}) = \mathbf{a}'\mathbf{C}\,\mathbf{b} \quad , \quad \mathrm{B}(\mathbf{y},\mathbf{x}) = \mathbf{b}'\mathbf{C}\,\mathbf{a} = \mathbf{a}'\mathbf{C}'\mathbf{b}.$$

Pertanto, condizione necessaria e sufficiente affinché valga la (4) è che, rispetto a una *qualunque* base di S, la matrice di B sia simmetrica.

**OSSERVAZIONE 3.** Ovviamente, sussiste una corrispondenza biunivoca, dipendente dalla base prescelta di S, tra l'insieme delle forme bilineari simmetriche su S e l'insieme delle matrici simmetriche di ordine appropriato. ▮

Data la forma bilineare simmetrica B su S, sia $\mathbf{C}$ la matrice di B rispetto a una base di S costituita dai vettori $\mathbf{x}_1, \ldots, \mathbf{x}_p$.

Supponiamo poi di operare un cambiamento di base e che, rispetto a una nuova base di S formata dai vettori $\widetilde{\mathbf{x}}_1, \ldots, \widetilde{\mathbf{x}}_p$, la matrice $\widetilde{\mathbf{C}}$ di B sia diagonale, vale a dire sia

$$\widetilde{\mathbf{C}} = \mathrm{diag}(\widetilde{c}_{11}, \ldots, \widetilde{c}_{pp}) \ .$$

Si dice, in tal caso, che **C** è stata ridotta a forma diagonale o canonica.

**OSSERVAZIONE 4.** Il problema della riduzione a forma diagonale della matrice **C** della forma bilineare simmetrica B equivale, evidentemente, a quello di trovare una matrice non singolare **A** tale che **A'CA** sia diagonale. ▮

Nella dimostrazione del teorema seguente è indicato un procedimento per ridurre a forma diagonale la matrice di una forma bilineare simmetrica.

**TEOREMA 1**

Sia B una forma bilineare simmetrica su uno spazio vettoriale S di ordine n e dimensione p.

Sia, inoltre, $\mathbf{C} = [c_{kj}]_{(p,p)}$ la matrice di B rispetto a una base di S costituita dai vettori $\mathbf{x}_1, \ldots, \mathbf{x}_p$.

Se $p = 1$, rispetto a ogni base di S, la matrice di B è diagonale.

Se $p > 1$, supposto che i p determinanti

$$\Delta_1 = \det[c_{11}] = \det(\mathbf{C}^{(1)}) \quad , \ \ldots \ , \quad \Delta_p = \det\begin{bmatrix} c_{11} \cdots c_{1p} \\ \cdots\cdots\cdots \\ c_{p1} \cdots c_{pp} \end{bmatrix} = \det(\mathbf{C}^{(p)})$$

siano tutti diversi da zero, è possibile individuare una base di S formata dai vettori $\widetilde{\mathbf{x}}_1, \ldots, \widetilde{\mathbf{x}}_p$, rispetto alla quale la matrice di B è data da ($\Delta_0 = 1$)

$$\widetilde{\mathbf{C}} = \mathrm{diag}\left(\Delta_1/\Delta_0 \ , \Delta_2/\Delta_1 \ , \ldots , \Delta_p/\Delta_{p-1}\right) .$$

*Dim.* Se $p = 1$, la conclusione del teorema è evidente in quanto la matrice rappresentativa di B è di ordine $(1,1)$.

Se $p > 1$, consideriamo i p vettori

$$\widetilde{\mathbf{x}}_1 = \mathbf{x}_1 = \frac{c_{11}^{(1)}}{\Delta_0}\mathbf{x}_1$$

$$\widetilde{\mathbf{x}}_2 = \frac{c_{12}^{(2)}}{\Delta_1}\mathbf{x}_1 + \mathbf{x}_2 = \frac{c_{12}^{(2)}}{\Delta_1}\mathbf{x}_1 + \frac{c_{22}^{(2)}}{\Delta_1}\mathbf{x}_2$$

..........................................................................................................................

$$\widetilde{\mathbf{x}}_p = \frac{c_{1p}^{(p)}}{\Delta_{p-1}}\mathbf{x}_1 + \ldots + \frac{c_{(p-1)p}^{(p)}}{\Delta_{p-1}}\mathbf{x}_{p-1} + \mathbf{x}_p = \frac{c_{1p}^{(p)}}{\Delta_{p-1}}\mathbf{x}_1 + \ldots + \frac{c_{(p-1)p}^{(p)}}{\Delta_{p-1}}\mathbf{x}_{p-1} + \frac{c_{pp}^{(p)}}{\Delta_{p-1}}\mathbf{x}_p$$

dove $c_{11}^{(1)} = 1$ e $c_{hs}^{(u)}$ $(u = 2, \ldots, p; h,s = 1, \ldots, u)$ indica il cofattore dell'elemento $c_{hs}$ in $\mathbf{C}^{(u)}$.

Tenuto conto dello sviluppo di un determinante secondo gli elementi di una riga e del fatto che la somma dei prodotti degli elementi di una riga per i cofattori di una riga diversa vale zero, si può verificare che $(k,j = 1, \ldots, p)$

$$B(\widetilde{\mathbf{x}}_k, \widetilde{\mathbf{x}}_j) = 0 \;\text{ se }\; k \neq j \quad , \quad B(\widetilde{\mathbf{x}}_k, \widetilde{\mathbf{x}}_j) = \Delta_j/\Delta_{j-1} \;\text{ se }\; k=j$$

e quindi, rispetto alla base di S costituita dai vettori $\widetilde{\mathbf{x}}_1, \ldots, \widetilde{\mathbf{x}}_p$ di cui sopra, la matrice $\mathbf{C}$ della forma bilineare simmetrica B assume la forma diagonale indicata nell'enunciato del teorema.

**OSSERVAZIONE 5.** La possibilità di ridurre a forma diagonale la matrice $\mathbf{C}$ di una forma bilineare simmetrica B su S attraverso il procedimento qui sopra esposto dipende criticamente, oltreché da B, dalla base iniziale di S (2).

Si noti inoltre che, mutando la base di partenza, si giunge a una diversa forma diagonale di $\mathbf{C}$.

**OSSERVAZIONE 6.** La matrice del passaggio da una base formata dai vettori $\mathbf{x}_1, \ldots, \mathbf{x}_p$ alla base definita nella dimostrazione del Teorema 1 e costituita dai vettori $\widetilde{\mathbf{x}}_1, \ldots, \widetilde{\mathbf{x}}_p$ è

$$\mathbf{A} = \begin{bmatrix} c_{11}^{(1)}/\Delta_0 & c_{12}^{(2)}/\Delta_1 & \cdots & c_{1p}^{(p)}/\Delta_{p-1} \\ 0 & c_{22}^{(2)}/\Delta_1 & \cdots & c_{2p}^{(p)}/\Delta_{p-1} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & c_{pp}^{(p)}/\Delta_{p-1} \end{bmatrix}.$$

(2) Si veda, al riguardo, il punto 1 dei Complementi a questo capitolo.

Ovviamente, **A** è non singolare, come si verifica facilmente in considerazione del fatto che il suo determinante è dato dal prodotto degli elementi situati sulla diagonale principale e che questi sono tutti eguali a 1.

**ESEMPIO 4.** Rispetto alla base naturale di $R^3$, si consideri la forma bilineare simmetrica B su $R^3$ la cui matrice è

$$\mathbf{C} = \begin{bmatrix} 2 & 3/2 & 2 \\ 3/2 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} .$$

Come si può facilmente verificare – poiché

$$\Delta_1 = 2 \neq 0 \quad , \quad \Delta_2 = -1/4 \neq 0 \quad , \quad \Delta_3 = -17/4 \neq 0$$

– i vettori

$$\tilde{\mathbf{x}}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad , \quad \tilde{\mathbf{x}}_2 = \begin{bmatrix} -3/4 \\ 1 \\ 0 \end{bmatrix} \quad , \quad \tilde{\mathbf{x}}_3 = \begin{bmatrix} 8 \\ -12 \\ 1 \end{bmatrix}$$

formano una base di $R^3$ rispetto alla quale risulta

$$\tilde{\mathbf{C}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1/8 & 0 \\ 0 & 0 & 17 \end{bmatrix} .$$

**3.2** Data la forma bilineare simmetrica B su S, si consideri la funzione Q che associa, a ogni $\mathbf{x} \in S$, il numero reale $Q(\mathbf{x}) = B(\mathbf{x}, \mathbf{x})$.

Si dice, in tal caso, che Q è la *forma quadratica su* S *determinata da* (o *associata a*) B.

Sia, adesso, $\mathbf{X} = [\, \mathbf{x}_1 \cdots \mathbf{x}_p \,]$ la matrice i cui vettori colonna formano una base di S; siano, inoltre, $\mathbf{x} = \mathbf{X}\mathbf{a}$ e $\mathbf{y} = \mathbf{X}\mathbf{b}$ due vettori appartenenti a S.

Poiché il valore assunto dalla forma bilineare simmetrica B, in corrispondenza dei vettori di cui sopra, è dato da

$$B(\mathbf{x}, \mathbf{y}) = \mathbf{a}'\mathbf{C}\,\mathbf{b} \tag{1''}$$

dove $\mathbf{C}$ è la matrice simmetrica di ordine $(p,p)$ rappresentativa di B rispetto alla suddetta base di S, il valore assunto dalla forma quadratica Q determinata da B, in corrispondenza del vettore $\mathbf{x} \in S$, è dato da

$$(5) \qquad \mathrm{Q}(\mathbf{x}) = \mathbf{a}'\mathbf{C}\,\mathbf{a}\ .$$

In altri termini, rispetto a una stessa base di S, sia la forma bilineare simmetrica B sia la forma quadratica Q a essa associata sono rappresentate dalla medesima matrice simmetrica $\mathbf{C}$.

Reciprocamente, siano $\mathbf{X} = [\,\mathbf{x}_1 \cdots \mathbf{x}_p\,]$ la matrice i cui vettori colonna formano una base di S e $\mathbf{C}$ una matrice simmetrica di ordine $(p,p)$.

Poiché alla matrice $\mathbf{C}$ è possibile associare, mediante la (1"), la forma bilineare simmetrica B, alla stessa matrice $\mathbf{C}$ risulta associata, mediante la (5), la forma quadratica Q determinata da B.

**OSSERVAZIONE 7.** Tenuto conto di quanto ora mostrato, possiamo affermare che esiste una corrispondenza biunivoca tra l'insieme delle forme bilineari simmetriche e l'insieme delle forme quadratiche a esse associate.

**3.3** Data una forma bilineare simmetrica B su S, sia Q la forma quadratica a essa associata.

Si dice che Q è *definita positiva* (*definita negativa*) se, per ogni $\mathbf{x} \in S$ diverso dal vettore nullo, risulta $\mathrm{Q}(\mathbf{x}) > 0$ ($\mathrm{Q}(\mathbf{x}) < 0$).

Si dice poi che Q è *semidefinita positiva* (*semidefinita negativa*) o anche *definita non negativa* (*definita non positiva*) se $\mathrm{Q}(\mathbf{x}) \geq 0$ ($\mathrm{Q}(\mathbf{x}) \leq 0$) per ogni $\mathbf{x} \in S$ e $\mathrm{Q}(\mathbf{x}) = 0$ per qualche $\mathbf{x} \neq \mathbf{0}$.

Si dice infine che Q è *indefinita* se essa assume su S sia valori positivi sia valori negativi.

La stessa terminologia è correntemente impiegata nei riguardi della matrice (simmetrica) $\mathbf{C}$ associata, nella base prescelta di S, alla forma quadratica Q.

**3.4** Il concetto di forma quadratica definita positiva gioca un ruolo importante in molte questioni di algebra lineare.

Tra l'altro, come vedremo nel capitolo successivo, una forma bilineare simmetrica tale che la forma quadratica a essa associata sia definita positiva è alla base della definizione di spazio euclideo.

Il teorema seguente fornisce un criterio – basato sullo studio della matrice $\mathbf{C}$ associata, in una data base di S, alla forma quadratica Q determinata da una forma bilineare simmetrica B – per riconoscere quando Q è definita positiva.

**TEOREMA 2**

Siano B una forma bilineare simmetrica su uno spazio vettoriale S di ordine n e dimensione p e Q la forma quadratica a essa associata.

Sia inoltre $\mathbf{C}$ la matrice di B e quindi di Q rispetto a una base di S formata da p vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$.

Condizione necessaria e sufficiente affinché Q sia definita positiva è che i p determinanti

$$\Delta_1 = \det[c_{11}] = \det(\mathbf{C}^{(1)}) \quad , \ \dots \ , \quad \Delta_p = \det\begin{bmatrix} c_{11} \cdots c_{1p} \\ \cdots\cdots\cdots \\ c_{p1} \cdots c_{pp} \end{bmatrix} = \det(\mathbf{C}^{(p)})$$

siano tutti positivi.

*Dim.* Supposto che Q sia definita positiva, vogliamo anzitutto dimostrare che non può essere $(j = 1, \dots, p)$

$$\Delta_j = \det\begin{bmatrix} c_{11} \cdots c_{1j} \\ \cdots\cdots\cdots \\ c_{j1} \cdots c_{jj} \end{bmatrix} = \det\begin{bmatrix} B(\mathbf{x}_1, \mathbf{x}_1) & \cdots & B(\mathbf{x}_1, \mathbf{x}_j) \\ \cdots & \cdots & \cdots \\ B(\mathbf{x}_j, \mathbf{x}_1) & \cdots & B(\mathbf{x}_j, \mathbf{x}_j) \end{bmatrix} = \det(\mathbf{C}^{(j)}) = 0 \, .$$

Infatti, se fosse $\Delta_j = 0$, i vettori riga che compongono la matrice $\mathbf{C}^{(j)}$ sarebbero linearmente dipendenti (Cfr. quanto detto al punto 1 dei Complementi al Cap. IV).

Esisterebbero, pertanto, j numeri reali $c_1, \dots, c_j$ non tutti nulli tali che

$$\sum_{k=1}^{j} c_k \left[ B(\mathbf{x}_k, \mathbf{x}_1) \cdots B(\mathbf{x}_k, \mathbf{x}_j) \right] = \sum_{k=1}^{j} \left[ B(c_k \mathbf{x}_k, \mathbf{x}_1) \cdots B(c_k \mathbf{x}_k, \mathbf{x}_j) \right]$$
$$= \left[ \sum_{k=1}^{j} B(c_k \mathbf{x}_k, \mathbf{x}_1) \cdots \sum_{k=1}^{j} B(c_k \mathbf{x}_k, \mathbf{x}_j) \right] = \mathbf{0}_{(1,j)}$$

ovvero tali che $(h = 1, \dots, j)$

$$B\big(\sum_{k=1}^{j} c_k \mathbf{x}_k, \mathbf{x}_h\big) = 0 \, .$$

Ma, moltiplicando ambo i membri di quest'ultima relazione per $c_h$ e sommando rispetto all'indice h, otterremmo

$$\sum_{h=1}^{j} c_h \, B\big(\sum_{k=1}^{j} c_k \mathbf{x}_k, \mathbf{x}_h\big) = B\Big(\sum_{k=1}^{j} c_k \mathbf{x}_k \, , \sum_{h=1}^{j} c_h \mathbf{x}_h\Big) = B(\mathbf{x}, \mathbf{x}) = Q(\mathbf{x}) = 0$$

dove $\mathbf{x}$ è diverso dal vettore nullo, in contrasto con l'ipotesi che Q sia definita positiva.

Avendo dimostrato che $\Delta_1, \Delta_2, \dots, \Delta_p$ sono tutti diversi da zero, esiste una base di B, e quindi di Q, costituita dai vettori $\widetilde{\mathbf{x}}_1, \dots, \widetilde{\mathbf{x}}_p$ definiti nella dimostrazione del Teorema 1, rispetto alla quale – per ogni

$$\mathbf{x} = \widetilde{a}_1 \widetilde{\mathbf{x}}_1 + \dots + \widetilde{a}_p \widetilde{\mathbf{x}}_p \in S$$

– risulta

$$Q(\mathbf{x}) = \sum_{j=1}^{p} \frac{\Delta_j}{\Delta_{j-1}} \widetilde{a}_j^2 \, . \tag{6}$$

D'altra parte, poiché Q è definita positiva, dall'espressione ora scritta segue che $(j = 1, \dots, p;\ \Delta_0 = 1)$

$$\frac{\Delta_j}{\Delta_{j-1}} > 0$$

e, quindi, che $\Delta_1, \Delta_2, \dots, \Delta_p$ sono tutti positivi.

Viceversa, supposto che $\Delta_1, \Delta_2, \dots, \Delta_p$ siano tutti positivi, il Teorema 1 assicura che esiste una base di S rispetto alla quale, per ogni $\mathbf{x} \in S$, vale la (6); pertanto, Q è definita positiva.

**ESEMPIO 5.** Rispetto alla base naturale di $R^3$, si consideri la forma bilineare simmetrica B su $R^3$ la cui matrice è

$$\mathbf{C} = \begin{bmatrix} 7 & -2 & 0 \\ -2 & 6 & -2 \\ 0 & -2 & 5 \end{bmatrix} .$$

Poiché risulta

$$\Delta_1 = 7 > 0 \quad , \quad \Delta_2 = 38 > 0 \quad , \quad \Delta_3 = 162 > 0$$

la forma quadratica Q associata a B è definita positiva.

In effetti, rispetto alla base di $R^3$ costituita da

$$\tilde{\mathbf{x}}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad , \quad \tilde{\mathbf{x}}_2 = \begin{bmatrix} 2/7 \\ 1 \\ 0 \end{bmatrix} \quad , \quad \tilde{\mathbf{x}}_3 = \begin{bmatrix} 2/19 \\ 7/19 \\ 1 \end{bmatrix}$$

la matrice di B e di Q è

$$\tilde{\mathbf{C}} = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 38/7 & 0 \\ 0 & 0 & 162/38 \end{bmatrix}$$

e questa è chiaramente definita positiva.

**OSSERVAZIONE 8.** Nella formulazione del teorema precedente, i determinanti $\Delta_1, \Delta_2, \dots, \Delta_p$ dipendono dai vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$ che formano la base iniziale considerata.

Per contro, il fatto che la forma quadratica Q sia definita positiva è del tutto indipendente da tale base.

Ciò significa tra l'altro che, permutando i vettori che compongono la base prescelta di S, i determinanti che risultano, diciamo $\overline{\Delta}_1, \overline{\Delta}_2, \dots, \overline{\Delta}_p$, benché diversi dai precedenti, sono ancora tutti positivi.

Consideriamo, per esempio, la forma quadratica Q su S la cui matrice, rispetto a una base formata da $\mathbf{x}_1, \mathbf{x}_2$, è

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}.$$

Supposto che Q sia definita positiva, risulta

$$\Delta_1 = \det\left[c_{11}\right] = c_{11} > 0 \quad , \quad \Delta_2 = \det\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} > 0 .$$

Analogamente, rispetto a una base di S ottenuta permutando i vettori $\mathbf{x}_1, \mathbf{x}_2$, deve essere

$$\overline{\Delta}_1 = \det\left[c_{22}\right] = c_{22} > 0 \quad , \quad \overline{\Delta}_2 = \det\begin{bmatrix} c_{22} & c_{21} \\ c_{12} & c_{11} \end{bmatrix} > 0 .$$

Da quanto precede, si desume in particolare che gli elementi che compaiono sulla diagonale principale della matrice rappresentativa di Q, rispetto a una qualsiasi base di S, sono tutti positivi (3).

(3) Chiaramente, questa è una condizione necessaria ma non sufficiente affinché una forma quadratica Q sia definita positiva.

# Cap. VI
# COMPLEMENTI

**1**

Si consideri la forma bilineare simmetrica B su $R^3$ la cui matrice rispetto alla base naturale di $R^3$ è

$$\mathbf{C} = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix} .$$

Poiché

$$\Delta_1 = 1 \quad , \quad \Delta_2 = 0 \quad , \quad \Delta_3 = -9$$

il metodo di riduzione a forma diagonale di cui al Teorema 1 non può essere applicato.

Viceversa, rispetto alla base di $R^3$ costituita da

$$\mathbf{x}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad , \quad \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad , \quad \mathbf{x}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

la matrice di B risulta

$$\mathbf{C}_1 = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}$$

e, essendo

$$\overline{\Delta}_1 = 1 \quad , \quad \overline{\Delta}_2 = -3 \quad , \quad \overline{\Delta}_3 = -9 \; ,$$

è possibile ottenere, mediante il procedimento di cui al Teorema 1, una base di $R^3$ formata dai vettori

$$\widetilde{\mathbf{x}}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad , \quad \widetilde{\mathbf{x}}_2 = \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix} \quad , \quad \widetilde{\mathbf{x}}_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

rispetto alla quale si ha

$$\widetilde{\mathbf{C}} = \mathrm{diag}(1 \, , -3 \, , 3) \; .$$

**2**

Sia B una forma bilineare simmetrica su $R^p$ e Q la forma quadratica a essa associata.

Sia inoltre $\mathbf{C}$ la matrice di B e quindi di Q rispetto alla base naturale di $R^p$.

Supposto che Q sia definita positiva, la matrice $\mathbf{C}$ è invertibile e, pertanto, esiste la sua inversa $\mathbf{C}^{-1}$.

Vogliamo mostrare che la forma quadratica associata alla matrice (simmetrica) $\mathbf{C}^{-1}$ è definita positiva.

A questo fine, osserviamo intanto che $\mathbf{C}^{-1}$ può essere scritta nella forma

$$\mathbf{C}^{-1} = (\mathbf{C}')^{-1} = (\mathbf{C}^{-1})' = (\mathbf{C}^{-1})'\mathbf{C}\,\mathbf{C}^{-1} \; .$$

Ma, per ogni $\mathbf{x} \in R^p$ diverso dal vettore nullo, risulta $\mathbf{C}^{-1}\mathbf{x} \neq \mathbf{0}$, cosicché

$$\mathbf{x}'(\mathbf{C}^{-1})'\mathbf{C}\,\mathbf{C}^{-1}\mathbf{x} = \mathbf{x}'\mathbf{C}^{-1}\mathbf{x} > 0$$

e quindi la forma quadratica associata alla matrice $\mathbf{C}^{-1}$ è definita positiva.

**3**

Rispetto alla base naturale di $R^n$, sia $\mathbf{M}$ la matrice di ordine $(n,n)$ rappresentativa di una forma quadratica definita positiva su $R^n$.

Analogamente, rispetto alla base naturale di $R^p$, sia $\mathbf{Q}$ la matrice di ordine

$(p,p)$ rappresentativa di una forma quadratica definita positiva su $R^p$.

Data la matrice $\mathbf{Y}$ di ordine $(n,p)$ e posto

$$\mathbf{V} = \mathbf{Y}'\mathbf{M}\mathbf{Y} \quad , \quad \mathbf{W} = \mathbf{Y}\mathbf{Q}\mathbf{Y}'$$

ci proponiamo di svolgere qui di seguito alcune considerazioni riguardanti tali matrici che risultano di ordine, rispettivamente, $(p,p)$ e $(n,n)$.

($i$) La matrice $\mathbf{V}$, che riceve correntemente la denominazione di *matrice di Gram* costruita a partire dai p vettori colonna $\mathbf{y}_1, \dots, \mathbf{y}_p$ che compongono $\mathbf{Y}$ e da $\mathbf{M}$, è simmetrica e può essere interpretata, rispetto alla base naturale di $R^p$, come rappresentativa di una forma quadratica su $R^p$.

Se $\mathbf{y}_1, \dots, \mathbf{y}_p$ sono linearmente indipendenti, e cioè se $r(\mathbf{Y}) = p$, per ogni $\mathbf{x} \in R^p$ diverso dal vettore nullo risulta $\mathbf{Yx} \neq \mathbf{0}$; ne consegue che

$$0 < (\mathbf{Yx})'\mathbf{M}(\mathbf{Yx}) = \mathbf{x}'\mathbf{Y}'\mathbf{MYx} = \mathbf{x}'\mathbf{Vx}$$

e quindi $\mathbf{V}$ è definita positiva.

Viceversa, se $\mathbf{V}$ è definita positiva, per ogni $\mathbf{x} \in R^p$ diverso dal vettore nullo risulta $\mathbf{x}'\mathbf{Vx} > 0$; ma ciò implica che

$$0 < \mathbf{x}'\mathbf{Vx} = \mathbf{x}'\mathbf{Y}'\mathbf{MYx} = (\mathbf{Yx})'\mathbf{M}(\mathbf{Yx})$$

ovvero che $\mathbf{y}_1, \dots, \mathbf{y}_p$ siano linearmente indipendenti, e cioè che sia $r(\mathbf{Y}) = p$.

Se $\mathbf{y}_1, \dots, \mathbf{y}_p$ sono linearmente dipendenti, e cioè se $r(\mathbf{Y}) < p$, esiste qualche $\mathbf{x} \in R^p$ diverso dal vettore nullo tale che $\mathbf{Yx} = \mathbf{0}$; ne consegue che

$$0 = (\mathbf{Yx})'\mathbf{M}(\mathbf{Yx}) = \mathbf{x}'\mathbf{Y}'\mathbf{MYx} = \mathbf{x}'\mathbf{Vx}$$

e quindi $\mathbf{V}$ è semidefinita positiva.

Reciprocamente, se $\mathbf{V}$ è semidefinita positiva, per qualche $\mathbf{x} \in R^p$ diverso dal vettore nullo risulta $\mathbf{x}'\mathbf{Vx} = 0$; ma ciò implica che

$$0 = \mathbf{x}'\mathbf{Vx} = \mathbf{x}'\mathbf{Y}'\mathbf{MYx} = (\mathbf{Yx})'\mathbf{M}(\mathbf{Yx})$$

ovvero che $\mathbf{y}_1, \dots, \mathbf{y}_p$ siano linearmente dipendenti, e cioè che sia $r(\mathbf{Y}) < p$.

In definitiva, $\mathbf{V}$ è

- definita positiva se e soltanto se $\mathbf{y}_1, \dots, \mathbf{y}_p$ sono linearmente indipendenti, ovvero se e soltanto se $r(\mathbf{Y}) = p$;
- semidefinita positiva se e soltanto $\mathbf{y}_1, \dots, \mathbf{y}_p$ sono linearmente dipendenti, ovvero se e soltanto se $r(\mathbf{Y}) < p$.

(*ii*) Posto $\mathbf{Z} = \mathbf{Y}'$, risulta $\mathbf{W} = \mathbf{Y}\mathbf{Q}\mathbf{Y}' = \mathbf{Z}'\mathbf{Q}\mathbf{Z}$ ed è ovvio che considerazioni analoghe a quelle svolte qui sopra nei riguardi della matrice $\mathbf{V}$ possono farsi a proposito di $\mathbf{W}$, vale a dire relativamente alla *matrice di Gram* costruita a partire dagli n vettori colonna $\mathbf{z}_1, \dots, \mathbf{z}_n$ che compongono $\mathbf{Z}$ e da $\mathbf{Q}$.

Ne consegue che $\mathbf{W}$ è

- definita positiva se e soltanto se $\mathbf{z}_1, \dots, \mathbf{z}_n$ sono linearmente indipendenti, ovvero se e soltanto se $r(\mathbf{Z}) = r(\mathbf{Y}) = n$;
- semidefinita positiva se e soltanto se $\mathbf{z}_1, \dots, \mathbf{z}_n$ sono linearmente dipendenti, ovvero se e soltanto se $r(\mathbf{Z}) = r(\mathbf{Y}) < n$.

(*iii*) Le matrici $\mathbf{V}$ e $\mathbf{W}$ hanno lo stesso rango.

Infatti, poiché $\mathbf{M}$ e $\mathbf{Q}$ sono definite positive, esistono due matrici (invertibili) $\mathbf{H}$ e $\mathbf{K}$ tali che

$$\mathbf{H}'\mathbf{M}\mathbf{H} = \widetilde{\mathbf{M}} = \operatorname{diag}(\tilde{m}_1, \dots, \tilde{m}_n) \quad , \quad \mathbf{K}'\mathbf{Q}\mathbf{K} = \widetilde{\mathbf{Q}} = \operatorname{diag}(\tilde{q}_1, \dots, \tilde{q}_p)$$

dove $\tilde{m}_i > 0$ per $i = 1, \dots, n$ e $\tilde{q}_j > 0$ per $j = 1, \dots, p$.

Pertanto, posto

$$\widetilde{\mathbf{M}}^{1/2} = \operatorname{diag}\left(\tilde{m}_1^{1/2}, \dots, \tilde{m}_n^{1/2}\right) \quad , \quad \widetilde{\mathbf{Q}}^{1/2} = \operatorname{diag}\left(\tilde{q}_1^{1/2}, \dots, \tilde{q}_p^{1/2}\right),$$

risulta

$$\begin{aligned} r(\mathbf{V}) &= r(\mathbf{Y}'(\mathbf{H}^{-1})'\mathbf{H}'\mathbf{M}\mathbf{H}\mathbf{H}^{-1}\mathbf{Y}) = r(\mathbf{Y}'(\mathbf{H}^{-1})'\widetilde{\mathbf{M}}^{1/2}\widetilde{\mathbf{M}}^{1/2}\mathbf{H}^{-1}\mathbf{Y}) = r(\widetilde{\mathbf{M}}^{1/2}\mathbf{H}^{-1}\mathbf{Y}) = r(\mathbf{Y}) \\ r(\mathbf{W}) &= r(\mathbf{Z}'(\mathbf{K}^{-1})'\mathbf{K}'\mathbf{Q}\mathbf{K}\mathbf{K}^{-1}\mathbf{Z}) = r(\mathbf{Z}'(\mathbf{K}^{-1})'\widetilde{\mathbf{Q}}^{1/2}\widetilde{\mathbf{Q}}^{1/2}\mathbf{K}^{-1}\mathbf{Z}) = r(\widetilde{\mathbf{Q}}^{1/2}\mathbf{K}^{-1}\mathbf{Z}) = r(\mathbf{Z}) \end{aligned}$$

da cui, essendo $r(\mathbf{Y}) = r(\mathbf{Z})$, segue che $r(\mathbf{V}) = r(\mathbf{W})$.

($iv$) A completamento di quanto precede, si osservi che dal fatto che $\mathbf{V}$ sia definita positiva, segue che $\det(\mathbf{V}) > 0$.

Analogamente, in conseguenza del fatto che $\mathbf{V}$ sia semidefinita positiva – poiché $r(\mathbf{V}) = r(\mathbf{Y}) < p$ – risulta che $\det(\mathbf{V}) = 0$.

**4**

Vogliamo adesso stabilire alcune regole di derivazione di funzioni di vettori e matrici che risultano utili in molteplici applicazioni.

Sia f una funzione, a valori reali, degli elementi del vettore di ordine n

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

La derivata (prima) di f rispetto a $\mathbf{x}$ è definita ponendo

$$\frac{\partial f}{\partial \mathbf{x}} = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{bmatrix}.$$

Analogamente, sia f una funzione, a valori reali, degli elementi della matrice di ordine $(n, p)$

$$\mathbf{X} = \begin{bmatrix} x_{11} & \dots & x_{1p} \\ \dots & \dots & \dots \\ x_{n1} & \cdots & x_{np} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 \cdots \mathbf{x}_p \end{bmatrix}.$$

La derivata (prima) di f rispetto a $\mathbf{X}$ è definita ponendo

$$\frac{\partial f}{\partial \mathbf{X}} = \begin{bmatrix} \dfrac{\partial f}{\partial x_{11}} & \cdots & \dfrac{\partial f}{\partial x_{1p}} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial f}{\partial x_{n1}} & \cdots & \dfrac{\partial f}{\partial x_{np}} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f}{\partial \mathbf{x}_1} & \cdots & \dfrac{\partial f}{\partial \mathbf{x}_p} \end{bmatrix}.$$

Ciò premesso, è possibile dimostrare che:

(1) $f(\mathbf{x}) = \mathbf{a}'\mathbf{x}$ , $\dfrac{\partial f}{\partial \mathbf{x}} = \mathbf{a}$

(2) $f(\mathbf{x}) = \mathbf{x}'\mathbf{C}\mathbf{x}$ , $\dfrac{\partial f}{\partial \mathbf{x}} = (\mathbf{C} + \mathbf{C}')\mathbf{x}$

(3) $f(\mathbf{x}) = \mathbf{C}\mathbf{x}$ , $\dfrac{\partial f}{\partial \mathbf{x}} = \mathbf{C}'$

(4) $f(\mathbf{X}) = \mathrm{tr}(\mathbf{C}\mathbf{X}) = \mathrm{tr}(\mathbf{X}\mathbf{C})$ , $\dfrac{\partial f}{\partial \mathbf{X}} = \mathbf{C}'$

(5) $f(\mathbf{X}) = \mathrm{tr}(\mathbf{X}'\mathbf{C}\mathbf{X})$ , $\dfrac{\partial f}{\partial \mathbf{X}} = (\mathbf{C} + \mathbf{C}')\mathbf{X}$

(6) $f(\mathbf{X}) = \mathrm{tr}(\mathbf{X}'\mathbf{C}\mathbf{X}\mathbf{B})$ , $\dfrac{\partial f}{\partial \mathbf{X}} = \mathbf{C}\mathbf{X}\mathbf{B} + \mathbf{C}'\mathbf{X}\mathbf{B}'$

(7) $f(\mathbf{X}) = \log(\det(\mathbf{X}'\mathbf{C}\mathbf{X}))$ con $\mathbf{C} = \mathbf{C}'$, $\dfrac{\partial f}{\partial \mathbf{X}} = 2\mathbf{C}\mathbf{X}(\mathbf{X}'\mathbf{C}\mathbf{X})^{-1}$.

# Cap. VII
# SPAZI EUCLIDEI

## 1 GENERALITÀ

Considerato uno spazio vettoriale S di ordine n e dimensione p, supponiamo che su S sia definita una forma bilineare simmetrica G tale che la forma quadratica a essa associata sia *definita positiva*.

Si dice, in questo caso, che S è uno *spazio euclideo*.

La forma bilineare simmetrica G riceve abitualmente la denominazione di *metrica* (*euclidea*) di S [(1)]. A sua volta, il numero reale $G(\mathbf{x}, \mathbf{y})$ – vale a dire il valore assunto da G in corrispondenza dei vettori $\mathbf{x}, \mathbf{y} \in S$ – è chiamato *prodotto scalare* (o *interno*) di $\mathbf{x}$ e $\mathbf{y}$.

**OSSERVAZIONE 1.** Per tutti gli $\mathbf{x}, \mathbf{y} \in R^p$, si definisca la funzione G ponendo

$$G(\mathbf{x}, \mathbf{y}) = \mathbf{x}'\mathbf{y} .$$

Si riconosce subito che G è una forma bilineare simmetrica tale che la forma quadratica a essa associata è definita positiva e, quindi, G può essere assunta quale metrica di $R^p$.

Essa riceve la denominazione di *metrica* (*euclidea*) *standard*.

Si noti che, rispetto a una base di $R^p$ formata dai vettori colonna di una

(1) Nel seguito, a meno di una diversa indicazione, assumeremo, senza ulteriore specificazione, che S sia di ordine n e dimensione p e che G denoti la metrica di S.

matrice $\mathbf{X}$, la matrice rappresentativa di G è $\mathbf{G} = \mathbf{X}'\mathbf{X}$. Invece, rispetto alla base naturale di $R^p$, è $\mathbf{G} = \mathbf{I}$.

## 2 ORTOGONALITÀ

**2.1** Due vettori $\mathbf{x}$ e $\mathbf{y}$, appartenenti a uno spazio euclideo S, sono detti *ortogonali* (o *perpendicolari*), l'uno rispetto all'altro, e si scrive $\mathbf{x}\perp\mathbf{y}$, se $G(\mathbf{x}, \mathbf{y}) = 0$.

Chiaramente, l'unico vettore di S che è ortogonale a se stesso è il vettore nullo.

**ESEMPIO 1.** Supponiamo $R^2$ munito della metrica standard.

Dati i vettori

$$\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad , \quad \mathbf{y} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

risulta

$$G(\mathbf{x}, \mathbf{y}) = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix} = 0$$

e, quindi, $\mathbf{x}$ e $\mathbf{y}$ sono ortogonali.

**ESEMPIO 2.** Supponiamo $R^2$ munito della metrica definita ponendo, per tutti gli $\mathbf{x}, \mathbf{y} \in R^2$,

$$G(\mathbf{x}, \mathbf{y}) = \mathbf{x}' \begin{bmatrix} 1 & 5 \\ 5 & 26 \end{bmatrix} \mathbf{y} \; .$$

Dati i vettori $\mathbf{x}$ e $\mathbf{y}$ di cui all'Esempio precedente, risulta

$$G(\mathbf{x}, \mathbf{y}) = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 5 \\ 5 & 26 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix} = 35$$

e, quindi, $\mathbf{x}$ e $\mathbf{y}$ *non* sono ortogonali rispetto alla metrica qui considerata.

**OSSERVAZIONE 2.** Dato uno spazio euclideo S, il vettore nullo di S è ortogonale a ogni vettore appartenente a S.

**2.2** Un vettore $\mathbf{y}$ appartenente a uno spazio euclideo S è *unitario* se $G(\mathbf{y}, \mathbf{y}) = 1$.

Un vettore unitario è anche detto *versore*.

Si noti che ogni vettore *non nullo* $\mathbf{x} \in S$ può sempre essere trasformato in un vettore unitario (*normalizzazione*) moltiplicandolo per il reciproco del numero reale (positivo) $\sqrt{G(\mathbf{x}, \mathbf{x})}$ .

Si dice che $t \geq 2$ vettori $\mathbf{x}_1, \dots, \mathbf{x}_t$ appartenenti a uno spazio euclideo S formano un *insieme ortogonale* se

$$G(\mathbf{x}_r, \mathbf{x}_s) = 0 \quad \text{per } r,s = 1, \dots, t \text{ e } r \neq s \ .$$

Si dice, poi, che t vettori $\mathbf{y}_1, \dots, \mathbf{y}_t$ appartenenti a uno spazio euclideo S formano un *insieme ortonormale* se

$$G(\mathbf{y}_s, \mathbf{y}_s) = 1 \quad \text{per } s = 1, \dots, t$$

e inoltre, qualora sia $t \geq 2$,

$$G(\mathbf{y}_r, \mathbf{y}_s) = 0 \quad \text{per } r,s = 1, \dots, t \text{ e } r \neq s \ .$$

Un insieme ortogonale (ortonormale) che sia una base di S riceve la denominazione di *base ortogonale* (*ortonormale*).

**OSSERVAZIONE 3.** Un insieme ortogonale costituito da $t \geq 2$ vettori *non nulli* $\mathbf{x}_1, \dots, \mathbf{x}_t$ può essere sempre trasformato in un insieme ortonormale.

**TEOREMA 1**

> Un insieme *ortonormale* formato da t vettori $\mathbf{y}_1, \dots, \mathbf{y}_t$ appartenenti a uno spazio euclideo S è un insieme linearmente indipendente.

*Dim*. Se $t = 1$, l'insieme in questione è formato da un unico vettore il quale, essendo unitario, è certamente diverso dal vettore nullo e, quindi, è linearmente indipendente.

Se $t \geq 2$, si osservi anzitutto che, per ipotesi,

$$\begin{aligned} &G(\mathbf{y}_s, \mathbf{y}_s) = 1 \quad \text{per } s = 1, \dots, t \\ &G(\mathbf{y}_r, \mathbf{y}_s) = 0 \quad \text{per } r,s = 1, \dots, t \text{ e } r \neq s \ . \end{aligned}$$

Considerati t numeri reali $a_1, \dots, a_s, \dots, a_t$ tali che

$$a_1\mathbf{y}_1 + \dots + a_s\mathbf{y}_s + \dots + a_t\mathbf{y}_t = \mathbf{0}$$

ma per il resto arbitrari, poiché $(s = 1, \dots, t)$

$$\begin{aligned} 0 = G(\mathbf{0}, \mathbf{y}_s) &= G(a_1\mathbf{y}_1 + \dots + a_s\mathbf{y}_s + \dots + a_t\mathbf{y}_t , \mathbf{y}_s) \\ &= a_1 G(\mathbf{y}_1, \mathbf{y}_s) + \dots + a_s G(\mathbf{y}_s, \mathbf{y}_s) + \dots + a_t G(\mathbf{y}_t, \mathbf{y}_s) \\ &= a_s G(\mathbf{y}_s, \mathbf{y}_s) = a_s \ , \end{aligned}$$

$\mathbf{y}_1, \dots, \mathbf{y}_t$ sono linearmente indipendenti.

**COROLLARIO.** Ogni insieme *ortonormale* costituito da t vettori $\mathbf{y}_1, \dots, \mathbf{y}_t$ appartenenti a uno spazio euclideo S è parte di una base di S.

**2.3** Dato uno spazio euclideo S, ci proponiamo adesso di indicare un procedimento costruttivo che consenta di pervenire a una base ortonormale di S.

Se $\dim(S) = p = 1$, allora ogni base di S è formata da un unico vettore ed è ovvio che normalizzando tale vettore si ottiene una base ortonormale di S medesimo.

Se $\dim(S) = p \geq 2$, sia

$$\mathbf{G} = \left[G(\mathbf{x}_k, \mathbf{x}_j)\right]_{(p,\,p)} = \left[g_{kj}\right]_{(p,\,p)}$$

la matrice della metrica G di S rispetto a una base di S formata dai vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$.

Per il Teorema 2 del Cap. VI, i determinanti

$$\Delta_1 = \det[g_{11}] = \det\left(\mathbf{G}^{(1)}\right) \quad , \ \ldots \ , \quad \Delta_p = \det\begin{bmatrix} g_{11} \cdots g_{1p} \\ \cdots\cdots\cdots \\ g_{p1} \cdots g_{pp} \end{bmatrix} = \det\left(\mathbf{G}^{(p)}\right)$$

sono tutti positivi.

Tenuto conto del Teorema 1 del Cap. VI, ne consegue che i vettori

$$\widetilde{\mathbf{x}}_1 = \mathbf{x}_1 = \frac{g_{11}^{(1)}}{\Delta_0}\mathbf{x}_1$$

$$\widetilde{\mathbf{x}}_2 = \frac{g_{12}^{(2)}}{\Delta_1}\mathbf{x}_1 + \mathbf{x}_2 = \frac{g_{12}^{(2)}}{\Delta_1}\mathbf{x}_1 + \frac{g_{22}^{(2)}}{\Delta_1}\mathbf{x}_2$$

.....................................................................................................................

$$\widetilde{\mathbf{x}}_p = \frac{g_{1p}^{(p)}}{\Delta_{p-1}}\mathbf{x}_1 + \ldots + \frac{g_{(p-1)p}^{(p)}}{\Delta_{p-1}}\mathbf{x}_{p-1} + \mathbf{x}_p = \frac{g_{1p}^{(p)}}{\Delta_{p-1}}\mathbf{x}_1 + \ldots + \frac{g_{(p-1)p}^{(p)}}{\Delta_{p-1}}\mathbf{x}_{p-1} + \frac{g_{pp}^{(p)}}{\Delta_{p-1}}\mathbf{x}_p$$

– dove $g_{11}^{(1)} = 1$ e $g_{hs}^{(u)}$ $(u = 2, \ldots, p; h,s = 1, \ldots, u)$ indica il cofattore dell'elemento $g_{hs}$ in $\mathbf{G}^{(u)}$ – sono mutualmente ortogonali e formano una base di S rispetto alla quale la matrice di G è

$$\widetilde{\mathbf{G}} = \mathrm{diag}\left(\Delta_1/\Delta_0, \Delta_2/\Delta_1, \ldots, \Delta_p/\Delta_{p-1}\right) .$$

Il procedimento qui sopra descritto è noto come *processo di ortogonalizzazione di Gram-Schmidt* (2).

Supposto poi di normalizzare ciascuno dei vettori $\widetilde{\mathbf{x}}_1, \ldots, \widetilde{\mathbf{x}}_p$, si ottiene una base ortonormale rispetto alla quale la matrice di G è la matrice unità.

Il procedimento complessivo mediante il quale si passa da una base di S formata dai vettori $\mathbf{x}_1, \ldots, \mathbf{x}_p$ a una base di S medesimo costituita dai vettori ortonormali $\mathbf{y}_1, \ldots, \mathbf{y}_p$ riceve la denominazione di *processo di ortonormalizzazione di Gram-Schmidt*.

**ESEMPIO 3.** Supponiamo $R^2$ munito della metrica rappresentata, rispetto alla

---

(2) Di tale processo sarà data successivamente una interessante interpretazione in termini di *proiezioni ortogonali*.

base naturale di $R^2$, dalla matrice

$$\mathbf{G} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

Attraverso il processo di ortonormalizzazione di Gram-Schmidt, si ottiene anzitutto una base di $R^2$ formata dai vettori ortogonali

$$\tilde{\mathbf{x}}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad , \quad \tilde{\mathbf{x}}_2 = \begin{bmatrix} -1/2 \\ 1 \end{bmatrix},$$

tali cioè che

$$\begin{bmatrix} 1 & -1/2 \\ 0 & 1 \end{bmatrix}' \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1/2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix},$$

e, infine, una base di $R^2$ costituita dai vettori ortonormali

$$\mathbf{y}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \end{bmatrix} \quad , \quad \mathbf{y}_2 = \begin{bmatrix} -1/\sqrt{2} \\ \sqrt{2} \end{bmatrix},$$

tali cioé che

$$\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 0 & \sqrt{2} \end{bmatrix}' \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 0 & \sqrt{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

**2.4** Dato uno spazio euclideo S, si consideri un sottospazio $S_1$ di S. Si dice che un vettore $\mathbf{x} \in S$ è *ortogonale* (o *perpendicolare*) a $S_1$, e si scrive $\mathbf{x} \perp S_1$, se, per ogni $\mathbf{y} \in S_1$, accade che $\mathbf{x} \perp \mathbf{y}$.

**ESEMPIO 4.** Supponiamo $R^2$ munito della metrica standard.

Sia poi $S_1$, sottospazio di $R^2$, costituito da tutti i vettori della forma

$$a_1 \mathbf{x}_1 = a_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

con $a_1$ variabile in R.

Dato il vettore

$$\mathbf{x} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

risulta

$$G(\mathbf{x}, a_1\mathbf{x}_1) = a_1\left(\begin{bmatrix} -1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = 0$$

e, quindi, è $\mathbf{x} \perp S_1$.

**TEOREMA 2**

Condizione necessaria e sufficiente affinché un vettore $\mathbf{x}$ appartenente a uno spazio euclideo S sia ortogonale a un sottospazio $S_1$ di S è che $\mathbf{x}$ sia ortogonale a ciascuno dei vettori che formano una base di $S_1$.

*Dim.* Supponiamo che $\mathbf{x}_1, \dots, \mathbf{x}_{p_1}$ formino una base di $S_1$.

Allora, $\mathbf{x} \perp S_1$ implica, in particolare, che sia $\mathbf{x} \perp \mathbf{x}_h$ per ogni $h = 1, \dots, p_1$.

Viceversa, supposto

$$G(\mathbf{x}, \mathbf{x}_h) = 0 \,,$$

qualunque siano i numeri reali $a_1, \dots, a_{p_1}$, risulta

$$G(\mathbf{x}, a_h\mathbf{x}_h) = 0$$

e anche

$$G(\mathbf{x}, a_1\mathbf{x}_1 + \dots + a_{p_1}\mathbf{x}_{p_1}) = 0$$

Ma, al variare di $a_1, \dots, a_{p_1}$ in R, $a_1\mathbf{x}_1 + \dots + a_{p_1}\mathbf{x}_{p_1}$ descrive $S_1$ e, quindi, è $\mathbf{x} \perp S_1$. ▮

Due sottospazi $S_1$ e $S_2$ di uno spazio euclideo S si dicono *ortogonali* (o *perpendicolari*), e si scrive $S_1 \perp S_2$, se – comunque si scelgano $\mathbf{x} \in S_1$ e $\mathbf{y} \in S_2$ – risulta $\mathbf{x} \perp \mathbf{y}$.

**ESEMPIO 5.** Supponiamo $R^3$ munito della metrica standard.

Siano poi $S_1$ e $S_2$ sottospazi di $R^3$, costituiti rispettivamente da tutti i vettori della forma

$$a_1\mathbf{x}_1 = a_1\begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} \quad , \quad a_2\mathbf{x}_2 = a_2\begin{bmatrix} 4 \\ 1 \\ -1 \end{bmatrix}$$

con $a_1$ e $a_2$ variabili in R.

Poiché

$$G(a_1\mathbf{x}_1, a_2\mathbf{x}_2) = a_1 a_2\left(\begin{bmatrix} 1 & -1 & 3 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \\ -1 \end{bmatrix}\right) = 0$$

è $S_1 \perp S_2$.

**TEOREMA 3**

Condizione necessaria e sufficiente affinché due sottospazi $S_1$ e $S_2$ di uno spazio euclideo S siano ortogonali è che ciascun vettore appartenente a una base di $S_1$ sia ortogonale a ciascun vettore appartenente a una base di $S_2$.

*Dim*. Supponiamo che $\mathbf{x}_1, \dots, \mathbf{x}_{p1}$ e $\mathbf{x}_{p1+1}, \dots, \mathbf{x}_{p1+p2}$ formino due basi, rispettivamente, di $S_1$ e $S_2$.

Allora, $S_1 \perp S_2$ implica, in particolare, che sia $\mathbf{x}_h \perp \mathbf{x}_k$ per ogni $h = 1, \dots, p_1$ e $k = p_1+1, \dots, p_1+p_2$.

Viceversa, supposto

$$G(\mathbf{x}_h, \mathbf{x}_k) = 0 \ ,$$

qualunque siano i numeri reali $a_1, \dots, a_{p1}$ e $a_{p1+1}, \dots, a_{p1+p2}$, risulta

$$G(a_h\mathbf{x}_h, a_k\mathbf{x}_k) = 0$$

e anche

$$G(a_1\mathbf{x}_1+\ldots+a_{p1}\mathbf{x}_{p1}\,,\,a_{p1+1}\mathbf{x}_{p1+1}+\ldots+a_{p1+p2}\mathbf{x}_{p1+p2})=0\,.$$

Ma, al variare di $a_1,\ldots,a_{p1}$ e $a_{p1+1},\ldots,a_{p1+p2}$ in R,

$$a_1\mathbf{x}_1+\ldots+a_{p1}\mathbf{x}_{p1}\quad,\quad a_{p1+1}\mathbf{x}_{p1+1}+\ldots+a_{p1+p2}\mathbf{x}_{p1+p2}$$

descrivono, rispettivamente, $S_1$ e $S_2$; quindi, è $S_1 \perp S_2$. ▮

Dati uno spazio euclideo S e un suo sottospazio $S_1$, consideriamo l'insieme di *tutti* i vettori appartenenti a S che sono ortogonali a $S_1$. Tale insieme viene indicato con $S_1^{\perp}$ che si legge «$S_1$ *ortogonale* (o *perpendicolare*)».

**TEOREMA 4**

> Dati uno spazio euclideo S e un suo sottospazio $S_1$, l'insieme $S_1^{\perp}$ di tutti i vettori appartenenti a S che sono ortogonali a $S_1$ è un sottospazio di S.

*Dim*. È intanto ovvio che $S_1^{\perp}$, sottoinsieme di S, non è vuoto giacché, in ogni caso, $\mathbf{0}\in S_1^{\perp}$.

Inoltre, poiché per ogni $\hat{\mathbf{y}}\in S_1\,,\hat{\mathbf{e}}_1,\hat{\mathbf{e}}_2\in S_1^{\perp}\,,c\in R$ accade che

$$\begin{aligned} G(\hat{\mathbf{y}}\,,\hat{\mathbf{e}}_1+\hat{\mathbf{e}}_2) &= G(\hat{\mathbf{y}}\,,\hat{\mathbf{e}}_1)+G(\hat{\mathbf{y}}\,,\hat{\mathbf{e}}_2) &&= 0\\ G(\hat{\mathbf{y}}\,,c\,\hat{\mathbf{e}}_1) &= c\,G(\hat{\mathbf{y}}\,,\hat{\mathbf{e}}_1) &&= 0\end{aligned}$$

ne consegue che $\hat{\mathbf{e}}_1+\hat{\mathbf{e}}_2\in S_1^{\perp}$ e $c\hat{\mathbf{e}}_1\in S_1^{\perp}$, ossia che $S_1^{\perp}$ è un sottospazio di S.

**COROLLARIO.** Come è immediato rilevare,

$$S^{\perp}=S_0\quad,\quad S_0^{\perp}=S\quad,\quad (S_1^{\perp})^{\perp}=S_1\,.$$

**TEOREMA 5**

> Dati uno spazio euclideo S e un suo sottospazio $S_1$, risulta $S=S_1\oplus S_1^{\perp}$, ovvero, come anche si dice, $S_1$ e $S_1^{\perp}$ sono *complementi ortogonali* in S.

*Dim*. Escluso il caso banale in cui sia $S_1 = S_0$ oppure $S_1 = S$, supponiamo che $\mathbf{x}_1, \dots, \mathbf{x}_{p_1}$ $(1 \le p_1 < p)$ formino una base ortonormale di $S_1$.

È sempre possibile trovare $p - p_1$ vettori ortonormali $\mathbf{x}_{p_1+1}, \dots, \mathbf{x}_p$ di S in modo tale che $\mathbf{x}_1, \dots, \mathbf{x}_{p_1}, \mathbf{x}_{p_1+1}, \dots, \mathbf{x}_p$ formino una base ortonormale di S.

Basta, a questo fine, applicare il procedimento di ortonormalizzazione di Gram-Schmidt a partire da uno qualsiasi dei $p - p_1$ vettori che completano una base di S.

Come è allora ovvio, $\mathbf{x}_{p_1+1}, \dots, \mathbf{x}_p$ costituiscono una base ortonormale di $S_1^{\perp}$ e $S = S_1 \oplus S_1^{\perp}$.

**2.5** A conclusione del presente paragrafo, ci proponiamo di stabilire i seguenti teoremi.

**TEOREMA 6**

Dato un insieme *ortonormale* formato da t vettori $\mathbf{y}_1, \dots, \mathbf{y}_t$ appartenenti a uno spazio euclideo S, per ogni $\mathbf{x} \in S$, sussiste la relazione (*diseguaglianza di Bessel*)

$$\sum_{s=1}^{t} G^2(\mathbf{x}, \mathbf{y}_s) \le G(\mathbf{x}, \mathbf{x}) \tag{1}$$

valendo il segno di eguaglianza se e soltanto se $\mathbf{x}$ appartiene al sottospazio di S generato da $\mathbf{y}_1, \dots, \mathbf{y}_t$.

Inoltre, il vettore

$$\mathbf{x} - \sum_{s=1}^{t} G(\mathbf{x}, \mathbf{y}_s)\, \mathbf{y}_s \tag{2}$$

è ortogonale a ciascuno dei vettori $\mathbf{y}_1, \dots, \mathbf{y}_t$.

*Dim*. Escludiamo il caso banale in cui $\mathbf{x}$ sia il vettore nullo.

La (1) risulta immediatamente dal fatto che

$$0 \le G\Big(\mathbf{x} - \sum_{s=1}^{t} G(\mathbf{x}, \mathbf{y}_s)\, \mathbf{y}_s\, ,\, \mathbf{x} - \sum_{s=1}^{t} G(\mathbf{x}, \mathbf{y}_s)\, \mathbf{y}_s\Big)$$

$$= G(\mathbf{x},\mathbf{x}) - \sum_{s=1}^{t} G^2(\mathbf{x},\mathbf{y}_s) - \sum_{s=1}^{t} G^2(\mathbf{x},\mathbf{y}_s) + \sum_{s=1}^{t} G^2(\mathbf{x},\mathbf{y}_s)$$
$$= G(\mathbf{x},\mathbf{x}) - \sum_{s=1}^{t} G^2(\mathbf{x},\mathbf{y}_s) \ .$$

Dovendo inoltre valere, nella relazione precedente, il segno di eguaglianza se e soltanto se

$$\mathbf{x} - \sum_{s=1}^{t} G(\mathbf{x},\mathbf{y}_s)\,\mathbf{y}_s = \mathbf{0}$$

la prima parte del teorema è dimostrata.

La (2) infine si ottiene subito osservando che, per ogni $r = 1, \dots, t$, risulta

$$G\Big(\mathbf{x} - \sum_{s=1}^{t} G(\mathbf{x},\mathbf{y}_s)\,\mathbf{y}_s\,,\,\mathbf{y}_r\Big) = G(\mathbf{x},\mathbf{y}_s) - G(\mathbf{x},\mathbf{y}_s) = 0\ .$$

**ESEMPIO 6.** Supposto $R^3$ munito della metrica standard e considerati i vettori ortonormali

$$\mathbf{y}_1 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \frac{1}{\sqrt{14}} \quad , \quad \mathbf{y}_2 = \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix} \frac{1}{\sqrt{21}}$$

vogliamo verificare che il vettore

$$\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

soddisfa la diseguaglianza di Bessel.

Infatti,

$$G^2(\mathbf{x},\mathbf{y}_1) = \frac{25}{14} \quad , \quad G^2(\mathbf{x},\mathbf{y}_2) = \frac{1}{21} \quad , \quad G(\mathbf{x},\mathbf{x}) = 2$$

e, quindi,

$$25/14 + 1/21 = 11/6 < 2\ .$$

Si noti che, poiché $\mathbf{x}$ non appartiene al sottospazio di $R^3$ generato da $\mathbf{y}_1, \mathbf{y}_2$, la diseguaglianza di Bessel è verificata in senso stretto.

**TEOREMA 7**

Per tutti gli $\mathbf{x}, \mathbf{y}$ appartenenti a uno spazio euclideo S sussiste la relazione (*diseguaglianza di Cauchy-Schwarz*)

$$G^2(\mathbf{x}, \mathbf{y}) \leq G(\mathbf{x}, \mathbf{x})\, G(\mathbf{y}, \mathbf{y}) \tag{3}$$

valendo il segno di eguaglianza se e soltanto se $\mathbf{x}, \mathbf{y}$ sono linearmente dipendenti.

*Dim*. Escludiamo il caso banale in cui **x** o **y** siano eguali al vettore nullo.

Considerato l'insieme costituito dal vettore ortonormale $\mathbf{y}/\sqrt{G(\mathbf{y}, \mathbf{y})}$, la diseguaglianza di Bessel fornisce

$$G^2(\mathbf{x}, \frac{\mathbf{y}}{\sqrt{G(\mathbf{y}, \mathbf{y})}}) = G^2(\mathbf{x}, \mathbf{y}) \frac{1}{G(\mathbf{y}, \mathbf{y})} \leq G(\mathbf{x}, \mathbf{x})$$

da cui la (3).

**ESEMPIO 7.** Supposto $R^3$ munito della metrica standard e considerati i vettori

$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad , \quad \mathbf{y} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

vogliamo verificare la diseguaglianza di Cauchy-Schwarz.

Infatti,

$$G^2(\mathbf{x}, \mathbf{y}) = 25 \quad , \quad G(\mathbf{x}, \mathbf{x}) = 2 \quad , \quad G(\mathbf{y}, \mathbf{y}) = 14$$

e, quindi,

$$25 < 2 \cdot 14 = 28.$$

Si noti che, non essendo **x** e **y** linearmente dipendenti, la diseguaglianza di Cauchy-Schwarz è verificata in senso stretto.

**TEOREMA 8**

Data una base *ortonormale* di uno spazio euclideo S formata da p vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$, per tutti gli $\mathbf{x}, \mathbf{y} \in S$, sussistono le relazioni

(4) $$\mathbf{y} = \sum_{j=1}^{p} G(\mathbf{y}, \mathbf{x}_j)\, \mathbf{x}_j \,,$$

(5) $$G(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{p} G(\mathbf{x}, \mathbf{x}_j)\, G(\mathbf{y}, \mathbf{x}_j) \,,$$

(6) $$G(\mathbf{x}, \mathbf{x}) = \sum_{j=1}^{p} G^2(\mathbf{x}, \mathbf{x}_j) \,.$$

Il numero reale $G(\mathbf{y}, \mathbf{x}_j)$ $(j = 1, \dots, p)$ che compare nella (4) è detto *coefficiente di Fourier*. La (5) è anche nota come *identità di Parseval*.

*Dim*. Posto

$$\mathbf{y} = \sum_{j=1}^{p} a_j \mathbf{x}_j$$

si ha $(j = 1, \dots, p)$

$$G(\mathbf{y}, \mathbf{x}_j) = G\Big(\sum_{j=1}^{p} a_j \mathbf{x}_j \,, \mathbf{x}_j\Big) = a_j$$

e la (4) risulta dimostrata.

Per quanto riguarda la (5), posto

$$\mathbf{x} = \sum_{j=1}^{p} b_j \mathbf{x}_j$$

si ha

$$\mathbf{x} = \sum_{j=1}^{p} G(\mathbf{x}, \mathbf{x}_j)\, \mathbf{x}_j$$

e, quindi,

$$G(\mathbf{x}, \mathbf{y}) = G\Big(\sum_{j=1}^{p} G(\mathbf{x}, \mathbf{x}_j)\, \mathbf{x}_j \,, \sum_{j=1}^{p} G(\mathbf{y}, \mathbf{x}_j)\, \mathbf{x}_j\Big) = \sum_{j=1}^{p} G(\mathbf{x}, \mathbf{x}_j)\, G(\mathbf{y}, \mathbf{x}_j) \,.$$

Infine, se nella (5) si pone $\mathbf{y} = \mathbf{x}$ si ha subito la (6).

## 3 ANGOLI. LUNGHEZZE. DISTANZE

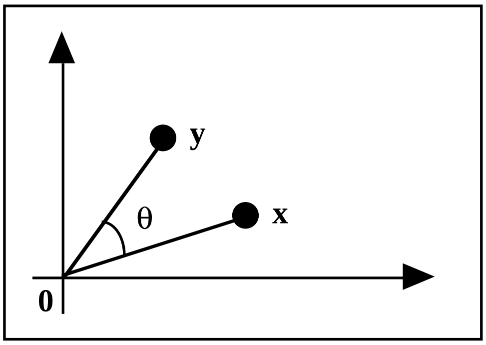
**3.1** Dato uno spazio euclideo S, dalla diseguaglianza di Cauchy-Schwarz risulta che

$$|G(\mathbf{x},\mathbf{y})| \leq \sqrt{G(\mathbf{x},\mathbf{x})}\,\sqrt{G(\mathbf{y},\mathbf{y})}\ .$$

Pertanto, per tutti gli $\mathbf{x}, \mathbf{y} \in S$ diversi dal vettore nullo, essendo il numero $G(\mathbf{x},\mathbf{y})/\sqrt{G(\mathbf{x},\mathbf{x})}\,\sqrt{G(\mathbf{y},\mathbf{y})}$ compreso tra −1 e +1, esiste un angolo $\theta$ tra 0 e $\pi$, detto *angolo* tra **x** e **y**, tale che [3]

$$\cos\theta = \frac{G(\mathbf{x},\mathbf{y})}{\sqrt{G(\mathbf{x},\mathbf{x})}\,\sqrt{G(\mathbf{y},\mathbf{y})}} = \cos(\mathbf{x},\mathbf{y})\ .$$

Fig. 1




Si osservi che la definizione data è in accordo con quanto stabilito a proposito dell'ortogonalità tra due vettori.

**3.2** Dato uno spazio vettoriale S, si chiama *norma* su S ogni funzione N che associa a ciascun vettore $\mathbf{x} \in S$ un numero reale − che riceve la denominazione di *lunghezza* (*norma*, *modulo*) di **x** e si indica con $N(\mathbf{x})$ oppure con $\|\mathbf{x}\|$ − in modo tale che siano soddisfatte le proprietà qui di seguito elencate ($\mathbf{x}, \mathbf{y} \in S, a \in R$):

(i) $N(\mathbf{x}) > 0$ se $\mathbf{x} \neq \mathbf{0}$; $N(\mathbf{x}) = 0$ se $\mathbf{x} = \mathbf{0}$

(3) Qualora **x** oppure **y** siano eguali al vettore nullo, si pone $\theta = 0$.

(ii) $\quad N(a\mathbf{x}) = |a|\, N(\mathbf{x})$

(iii) $\quad N(\mathbf{x}+\mathbf{y}) \leq N(\mathbf{x}) + N(\mathbf{y})$ (*diseguaglianza triangolare*).

Uno spazio vettoriale sul quale sia definita una norma è detto *spazio normato*.

Supposto che S sia uno spazio euclideo, vogliamo adesso mostrare che la funzione N definita ponendo, per ogni $\mathbf{x} \in S$,

$$N(\mathbf{x}) = \sqrt{G(\mathbf{x}, \mathbf{x})}$$

soddisfa le proprietà (i), (ii), (iii) ed è quindi una norma su S.

Si dice anche, in tal caso, che la struttura di spazio euclideo *induce* la struttura di spazio normato (4).

$$\boxed{\text{Struttura di spazio euclideo}} \Rightarrow \boxed{\text{Struttura di spazio normato}}$$

Per quanto riguarda la (i) e (ii), esse sono una conseguenza pressoché immediata della definizione di metrica euclidea.

Per quanto attiene alla (iii), si osservi anzitutto che

$$G(\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}) = G(\mathbf{x}, \mathbf{x}) + G(\mathbf{y}, \mathbf{y}) + 2G(\mathbf{x}, \mathbf{y}) .$$

Pertanto, tenuto conto della (3), si ha

$$\begin{aligned} G(\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}) &\leq G(\mathbf{x}, \mathbf{x}) + G(\mathbf{y}, \mathbf{y}) + 2\big(G(\mathbf{x}, \mathbf{x})\, G(\mathbf{y}, \mathbf{y})\big)^{1/2} \\ &= \big(G^{1/2}(\mathbf{x}, \mathbf{x}) + G^{1/2}(\mathbf{y}, \mathbf{y})\big)^2 . \end{aligned}$$

Quindi,

$$G(\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}) \leq G^{1/2}(\mathbf{x}, \mathbf{x}) + G^{1/2}(\mathbf{y}, \mathbf{y})$$

e la (iii) risulta così verificata.

Nel seguito supporremo che in S una norma sia stata introdotta mediante una metrica euclidea.

---

(4) Si può facilmente mostrare, attraverso esempi, che non vale la proposizione reciproca.

In questo caso, si può dimostrare quanto segue.

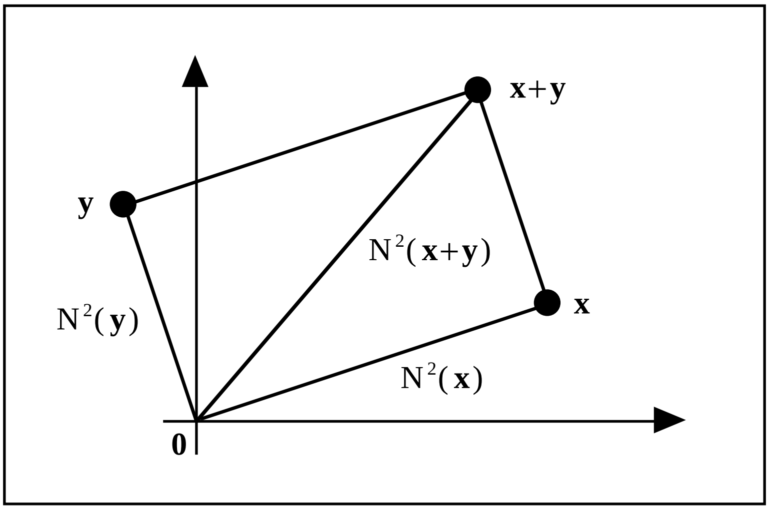
(a) Nella (iii) vale il segno di eguaglianza se e soltanto se i vettori **x** e **y** sono linearmente dipendenti.

(b) Vale la relazione (*di Pitagora*)

$$N^2(\mathbf{x}+\mathbf{y}) = N^2(\mathbf{x}) + N^2(\mathbf{y})$$

se e soltanto se **x** e **y** sono ortogonali.

Fig. 2




(c) Sussiste la relazione (*identità del parallelogramma*)

$$N^2(\mathbf{x}+\mathbf{y}) + N^2(\mathbf{x}-\mathbf{y}) = 2\left[N^2(\mathbf{x}) + N^2(\mathbf{y})\right].$$

(d) Infine, se in un qualsiasi spazio normato S vale la relazione precedente, allora S è uno spazio euclideo con una metrica definita ponendo, per tutti gli $\mathbf{x}, \mathbf{y} \in S$,

$$G(\mathbf{x},\mathbf{y}) = \frac{1}{4}\left[N^2(\mathbf{x}+\mathbf{y}) - N^2(\mathbf{x}-\mathbf{y})\right].$$

**3.3** Dato un insieme non vuoto S di oggetti di natura qualsiasi, si chiama *distanza* su S ogni funzione D che associa, a ogni $(\alpha,\beta) \in S\times S$, un numero reale $D(\alpha,\beta)$ – il quale, a sua volta, è detto *distanza* tra $\alpha$ e $\beta$ – in modo tale che siano soddisfatte le proprietà seguenti ($\alpha, \beta, \gamma \in S$):

(A) $\quad D(\alpha, \beta) > 0$ se $\alpha \neq \beta$; $D(\alpha, \beta) = 0$ se $\alpha = \beta$

(B) $\quad D(\alpha, \beta) = D(\beta, \alpha)$ (*simmetria*)

(C) $\quad D(\alpha, \gamma) \leq D(\alpha, \beta) + D(\beta, \gamma)$ (*diseguaglianza triangolare*).

Un insieme S sul quale è definita una distanza si chiama *spazio metrico*.

Supposto che S sia uno spazio normato, la funzione D definita ponendo, per tutti gli $\mathbf{x}, \mathbf{y} \in S$,

$$D(\mathbf{x}, \mathbf{y}) = N(\mathbf{x} - \mathbf{y})$$

è, come facilmente si verifica, una distanza.

Si dice anche, in tal caso, che la struttura di spazio normato *induce* la struttura di spazio metrico [5].

| Struttura di spazio normato | $\Rightarrow$ | Struttura di spazio metrico |
|---|---|---|

Nel seguito, a meno di esplicita avvertenza contraria, supporremo che in S una distanza sia stata introdotta mediante una norma e, in questo caso, valgono le relazioni ($\mathbf{x}, \mathbf{y}, \mathbf{z} \in S$; $a \in R$)

$$D(\mathbf{x}-\mathbf{z}, \mathbf{y}-\mathbf{z}) = D(\mathbf{x}, \mathbf{y}) \quad , \quad D(a\mathbf{x}, a\mathbf{y}) = |a|\, D(\mathbf{x}, \mathbf{y}) .$$

Infine, se S è uno spazio vettoriale in cui è data una distanza D soddisfacente le relazioni precedenti, allora S è uno spazio normato con una norma definita ponendo, per ogni $\mathbf{x} \in S$,

$$N(\mathbf{x}) = D(\mathbf{x}, \mathbf{0}) .$$

(5) Si può facilmente mostrare, attraverso esempi, che non vale la proposizione reciproca.

## 4 TRASFORMAZIONI AUTOAGGIUNTE

Dato uno spazio euclideo S, si consideri una trasformazione lineare T di S in se stesso tale che, per tutti gli $\mathbf{x}, \mathbf{y} \in$ S, si abbia

$$\text{(7)} \qquad G(T(\mathbf{x}), \mathbf{y}) = G(\mathbf{x}, T(\mathbf{y})) \ .$$

Si dice allora che T è una *trasformazione* (*lineare*) *autoaggiunta* (o *simmetrica*) di S.

Le trasformazioni autoaggiunte godono di importanti proprietà che saranno poste in evidenza nel seguito dell'esposizione.

Per il momento, ci limitiamo a osservare che, indicata con **G** la matrice della metrica G rispetto a una base di S costituita dai vettori colonna di una matrice $\mathbf{X} = [\mathbf{x}_1 \cdots \mathbf{x}_p]$ e considerati due vettori $\mathbf{x} = \mathbf{X}\mathbf{a}$ e $\mathbf{y} = \mathbf{X}\mathbf{b}$ appartenenti a S, possiamo scrivere la (7) nella forma

$$\mathbf{a}'\mathbf{T}'\mathbf{G}\mathbf{b} = \mathbf{a}'\mathbf{G}\mathbf{T}\mathbf{b}$$

ovvero

$$\mathbf{a}'(\mathbf{T}'\mathbf{G} - \mathbf{G}\mathbf{T})\mathbf{b} = 0 \ .$$

dove **T** è la matrice rappresentativa di T rispetto alla suddetta base di S.

Ora, quest'ultima relazione deve valere per ogni $\mathbf{a}, \mathbf{b} \in R^p$ e, in particolare, per ogni $\mathbf{u}_k, \mathbf{u}_j \in R^p$ dove $\mathbf{u}_k$ e $\mathbf{u}_j$ $(k, j = 1, \dots, p)$ indicano, rispettivamente, il k-esimo e il j-esimo vettore canonico di ordine p.

Ne consegue che

$$\mathbf{T}'\mathbf{G} - \mathbf{G}\mathbf{T} = \mathbf{O}$$

ossia che

$$\text{(8)} \qquad \mathbf{T}'\mathbf{G} = \mathbf{G}\mathbf{T} \, .$$

La (8) esprime dunque la condizione necessaria e sufficiente affinché una trasformazione lineare T sia autoaggiunta.

Tenendo presente che, rispetto a una base ortonormale, la matrice **G** della metrica G è sempre la matrice unità, possiamo anche dire che condizione necessaria e sufficiente affinché T sia autoaggiunta è che, in una qualsiasi base ortonormale di S, **T** sia una matrice simmetrica.

**OSSERVAZIONE 4.** Si noti esplicitamente che, rispetto a una base non ortonormale, la matrice **T** di una trasformazione autoaggiunta T non è necessariamente simmetrica.

Per esempio, siano

$$\mathbf{T} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad , \quad \mathbf{G} = \begin{bmatrix} 3 & 2 \\ 2 & 4 \end{bmatrix} .$$

Allora, come subito si verifica, è $\mathbf{T}'\mathbf{G} = \mathbf{G}\mathbf{T}$ pur non essendo **T** una matrice simmetrica.

**ESEMPIO 8.** Supponiamo $R^2$ munito della metrica standard.

Considerata la trasformazione lineare T di $R^2$ in se stesso la cui matrice rispetto alla base naturale (ortonormale) di $R^2$ è

$$\mathbf{T} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} ,$$

per quanto detto in precedenza, T è autoaggiunta.

Viceversa, supposto che **T** sia la matrice della trasformazione lineare $T_1$ di $R^2$ in se stesso rispetto alla base di $R^2$ formata dai vettori non ortonormali

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad , \quad \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

allora $T_1$ *non* è autoaggiunta, come risulta subito osservando che

$$\mathbf{G} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \quad , \quad \mathbf{G}\mathbf{T} = \begin{bmatrix} 4 & 5 \\ 3 & 3 \end{bmatrix} \neq \begin{bmatrix} 4 & 3 \\ 5 & 3 \end{bmatrix} = \mathbf{T}'\mathbf{G} .$$

## 5 PROIETTORI ORTOGONALI

**5.1** Siano $S_1$ e $S_1^{\perp}$ complementi ortogonali in uno spazio euclideo S, tali cioè che $S = S_1 \oplus S_1^{\perp}$.

Allora, come sappiamo, ogni $\mathbf{y} \in S$ si esprime univocamente nella forma $\mathbf{y} = \widehat{\mathbf{y}} + \widehat{\mathbf{e}}$ con $\widehat{\mathbf{y}} \in S_1$ e $\widehat{\mathbf{e}} \in S_1^{\perp}$.

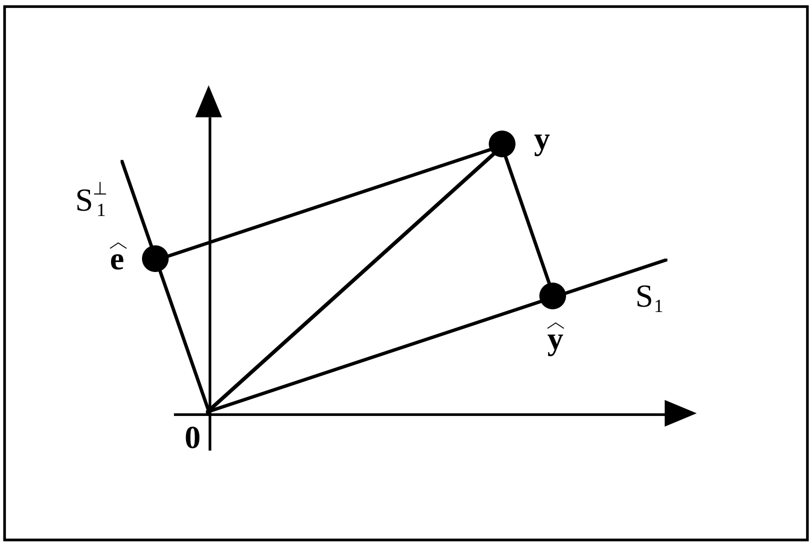
Inoltre, la funzione T (trasformazione idempotente) che associa a ogni $\mathbf{y} \in S$ il vettore $\widehat{\mathbf{y}} \in S_1$ è un proiettore di S su $S_1$ lungo $S_1^{\perp}$ (Cfr. la sezione 6.2 del Cap. V).

Analogamente, la funzione $I - T$ (trasformazione idempotente) che associa a ogni $\mathbf{y} \in S$ il vettore $\widehat{\mathbf{e}} \in S_1^{\perp}$ è un proiettore di S su $S_1^{\perp}$ lungo $S_1$ (Cfr. la sezione 6.4 del Cap. V).

I proiettori T e $I - T$ ricevono, in questo caso, la denominazione di *proiettori ortogonali*.

A loro volta, i vettori $\widehat{\mathbf{y}}$ e $\widehat{\mathbf{e}}$ sono chiamati le *proiezioni ortogonali* di $\mathbf{y}$, rispettivamente, sui sottospazi $S_1$ e $S_1^{\perp}$ di S.

Fig. 3




Scopo del seguente teorema è quello di caratterizzare i proiettori ortogonali T e $I - T$ in termini di *proiettori autoaggiunti*, vale a dire in termini di trasformazioni (idempotenti e) autoaggiunte.

**TEOREMA 9**

In uno spazio euclideo S un proiettore è ortogonale se e soltanto se esso è autoaggiunto.

*Dim*. Siano i vettori $\hat{\mathbf{y}}$ e $\hat{\mathbf{e}}$ definiti come in precedenza.

Supposto che il proiettore T sia autoaggiunto, risulta

$$G(\hat{\mathbf{y}}, \hat{\mathbf{e}}) = G(T(\hat{\mathbf{y}}), \hat{\mathbf{e}}) = G(\hat{\mathbf{y}}, T(\hat{\mathbf{e}})) = G(\hat{\mathbf{y}}, \mathbf{0}) = 0$$

e, quindi, T è ortogonale.

Viceversa, supposto che il proiettore T sia ortogonale, allora

$$\begin{aligned} G(T(\hat{\mathbf{y}}), \hat{\mathbf{e}}) &= G(T(\hat{\mathbf{y}}), T(\hat{\mathbf{e}}) + (I-T)(\hat{\mathbf{e}})) \\ &= G(T(\hat{\mathbf{y}}), T(\hat{\mathbf{e}})) + G(T(\hat{\mathbf{y}}), (I-T)(\hat{\mathbf{e}})) \\ &= G(T(\hat{\mathbf{y}}), T(\hat{\mathbf{e}})) = G(\hat{\mathbf{y}}, T(\hat{\mathbf{e}})) \end{aligned}$$

e cioè T è autoaggiunto.

Ovviamente, una dimostrazione del tutto analoga sussiste qualora in luogo di T si consideri $I-T$.

**5.2** Supponiamo che $\mathbf{x}_1, \dots, \mathbf{x}_{p_1}$ e $\mathbf{x}_{p_1+1}, \dots, \mathbf{x}_p$ siano due basi *ortonormali*, rispettivamente, di $S_1$ e $S_1^{\perp}$, sottospazi di uno spazio euclideo S.

Allora, ricordando la (4), per ogni $\mathbf{y} \in S$, possiamo scrivere

$$\mathbf{y} = \sum_{j=1}^{p} G(\mathbf{y}, \mathbf{x}_j)\, \mathbf{x}_j = \sum_{h=1}^{p_1} G(\mathbf{y}, \mathbf{x}_h)\, \mathbf{x}_h + \sum_{k=p_1+1}^{p} G(\mathbf{y}, \mathbf{x}_k)\, \mathbf{x}_k \,.$$

Quindi, le proiezioni ortogonali $\hat{\mathbf{y}}$ e $\hat{\mathbf{e}}$ di $\mathbf{y}$ sono date da

$$\hat{\mathbf{y}} = \sum_{h=1}^{p_1} G(\mathbf{y}, \mathbf{x}_h)\, \mathbf{x}_h \quad , \quad \hat{\mathbf{e}} = \sum_{k=p_1+1}^{p} G(\mathbf{y}, \mathbf{x}_k)\, \mathbf{x}_k \,.$$

Supponiamo, adesso, che $\mathbf{x}_1, \dots, \mathbf{x}_{p_1}$ e $\mathbf{x}_{p_1+1}, \dots, \mathbf{x}_p$ formino due basi *non necessariamente ortonormali*, rispettivamente, di $S_1$ e $S_1^{\perp}$, sottospazi di uno spazio euclideo S.

Scriviamo il vettore $\widehat{\mathbf{y}}$, proiezione ortogonale di $\mathbf{y} \in S$ su $S_1$, nella forma

$$\widehat{\mathbf{y}} = \widehat{a}_1 \mathbf{x}_1 + \ldots + \widehat{a}_{p1} \mathbf{x}_{p1} = \left[ \mathbf{x}_1 \cdots \mathbf{x}_{p1} \right] \begin{bmatrix} \widehat{a}_1 \\ \vdots \\ \widehat{a}_{p1} \end{bmatrix} = \mathbf{X}_1 \widehat{\mathbf{a}}_1 \ .$$

Dovendo il vettore $\widehat{\mathbf{e}} = \mathbf{y} - \widehat{\mathbf{y}}$ essere ortogonale a $S_1$, tenuto conto del Teorema 2, i coefficienti che compaiono nell'espressione precedente possono essere determinati come soluzioni del sistema di equazioni lineari – dette *equazioni normali* –

$$\begin{gathered} G(\mathbf{y} - a_1 \mathbf{x}_1 - \ldots - a_{p1} \mathbf{x}_{p1}, \mathbf{x}_1) = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ G(\mathbf{y} - a_1 \mathbf{x}_1 - \ldots - a_{p1} \mathbf{x}_{p1}, \mathbf{x}_{p1}) = 0 \end{gathered}$$

nelle incognite $a_1, \ldots, a_{p1}$.

Tale sistema – posto

$$\widetilde{\mathbf{y}} = \begin{bmatrix} G(\mathbf{x}_1, \mathbf{y}) \\ \vdots \\ G(\mathbf{x}_{p1}, \mathbf{y}) \end{bmatrix} \quad , \quad \mathbf{G}^{(p1)} = \begin{bmatrix} G(\mathbf{x}_1, \mathbf{x}_1) & \cdots & G(\mathbf{x}_1, \mathbf{x}_{p1}) \\ \cdots & \cdots & \cdots \\ G(\mathbf{x}_{p1}, \mathbf{x}_1) & \cdots & G(\mathbf{x}_{p1}, \mathbf{x}_{p1}) \end{bmatrix} \quad , \quad \mathbf{a}_1 = \begin{bmatrix} a_1 \\ \vdots \\ a_{p1} \end{bmatrix}$$

– può essere scritto nella forma

$$\widetilde{\mathbf{y}} = \mathbf{G}^{(p1)} \mathbf{a}_1$$

e, poiché $\det(\mathbf{G}^{(p1)}) > 0$ (Cfr. quanto ricordato nella sezione 2.3), il sistema in questione ammette la soluzione (unica)

$$\widehat{\mathbf{a}}_1 = (\mathbf{G}^{(p1)})^{-1} \widetilde{\mathbf{y}} \ .$$

Pertanto, si ha che

$$(9) \qquad \widehat{\mathbf{y}} = \mathbf{X}_1 (\mathbf{G}^{(p1)})^{-1} \widetilde{\mathbf{y}} \quad , \quad \widehat{\mathbf{e}} = \mathbf{y} - \widehat{\mathbf{y}} = \mathbf{y} - \mathbf{X}_1 (\mathbf{G}^{(p1)})^{-1} \widetilde{\mathbf{y}} \ .$$

**OSSERVAZIONE 5.** Supposto $R^p$ munito della metrica rappresentata, rispetto alla base naturale di $R^p$, dalla matrice **G**, risulta

(9') $\quad \hat{\mathbf{y}} = \mathbf{X}_1(\mathbf{X}_1'\mathbf{G}\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{G}\mathbf{y} \quad , \quad \hat{\mathbf{e}} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X}_1(\mathbf{X}_1'\mathbf{G}\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{G}\mathbf{y}$ .

**ESEMPIO 9.** Supposto $R^2$ munito della metrica standard, sia

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

una base di $S_1$.

Come facilmente si verifica, la proiezione ortogonale su $S_1$ di

$$\mathbf{y} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

è

$$\hat{\mathbf{y}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \frac{3}{2} .$$

**5.3** Nel teorema seguente, vogliamo stabilire alcune notevoli proprietà delle proiezioni ortogonali $\hat{\mathbf{y}} \in S_1$ e $\hat{\mathbf{e}} \in S_1^{\perp}$ di un vettore $\mathbf{y} \in S$.

**TEOREMA 10** (*Teorema della proiezione*)

Siano S uno spazio euclideo e $S_1$ un suo sottospazio. Allora, per ogni $\mathbf{y} \in S$ , $\hat{\mathbf{y}} \in S_1$ , $\hat{\mathbf{e}} \in S_1^{\perp}$ tali che $\mathbf{y} = \hat{\mathbf{y}} + \hat{\mathbf{e}}$,

(i) vale la relazione (*di Pitagora*)

$$N^2(\mathbf{y}) = N^2(\hat{\mathbf{y}}) + N^2(\hat{\mathbf{e}}) \ ;$$

(ii) il vettore $\hat{\mathbf{y}}$, proiezione ortogonale di **y** su $S_1$, è tale che la distanza di **y** da $\mathbf{z} \in S_1$ è minima per $\mathbf{z} = \hat{\mathbf{y}}$.

*Dim*. Poiché $G(\hat{\mathbf{y}}, \hat{\mathbf{e}}) = 0$, la relazione di Pitagora di cui al punto (i) è subito verificata (Cfr. anche quanto detto in (b) della sezione 3.2), risultando

$$\begin{aligned} N^2(\mathbf{y}) &= G(\hat{\mathbf{y}}+\hat{\mathbf{e}}\,,\hat{\mathbf{y}}+\hat{\mathbf{e}}) &&= G(\hat{\mathbf{y}},\hat{\mathbf{y}}) + 2G(\hat{\mathbf{y}},\hat{\mathbf{e}}) + G(\hat{\mathbf{e}}\,,\hat{\mathbf{e}}) \\ &= G(\hat{\mathbf{y}},\hat{\mathbf{y}}) + G(\hat{\mathbf{e}}\,,\hat{\mathbf{e}}) &&= N^2(\hat{\mathbf{y}}) + N^2(\hat{\mathbf{e}})\ . \end{aligned}$$

Per dimostrare la proprietà indicata al punto (ii), si consideri il vettore $(\hat{\mathbf{y}} - \mathbf{z}) \in S_1$.

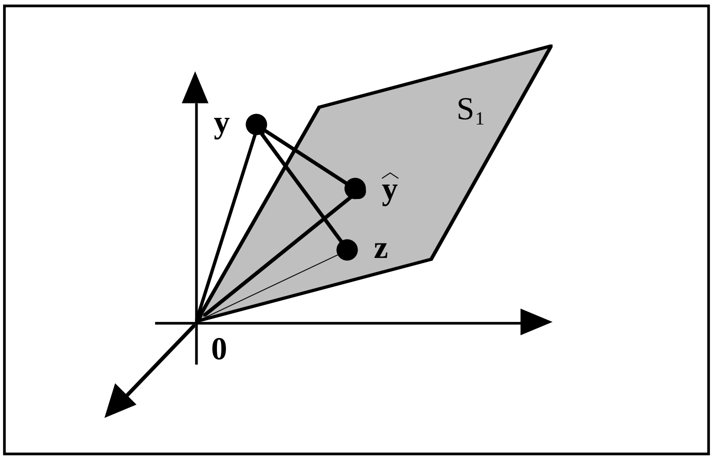
Chiaramente, tale vettore risulta ortogonale a $(\mathbf{y} - \hat{\mathbf{y}}) \in S_1^{\perp}$.

Ne consegue che

$$N^2(\mathbf{y}-\mathbf{z}) = N^2\big((\hat{\mathbf{y}}-\mathbf{z})+(\mathbf{y}-\hat{\mathbf{y}})\big) = N^2(\hat{\mathbf{y}}-\mathbf{z}) + N^2(\mathbf{y}-\hat{\mathbf{y}})$$

e, al variare di $\mathbf{z}$, quest'ultima quantità, che rappresenta il quadrato della distanza tra $\mathbf{y}$ e $\mathbf{z}$, è minima quando $\mathbf{z} = \hat{\mathbf{y}}$.

Fig. 4



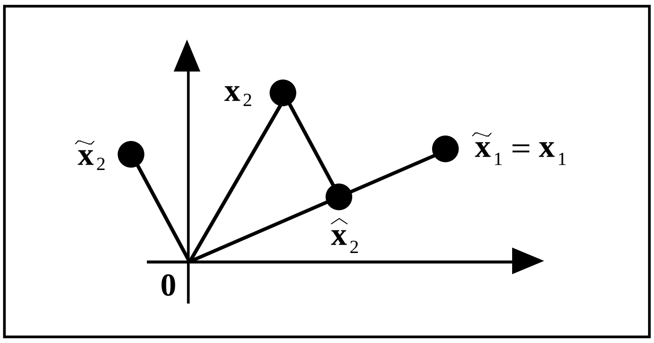

**5.4** Vogliamo adesso fornire una interpretazione del procedimento di ortogonalizzazione di Gram-Schmidt in termini di proiezioni ortogonali.

Posto $\tilde{\mathbf{x}}_1 = \mathbf{x}_1$, mostriamo anzitutto che la differenza tra $\mathbf{x}_2$ e la proiezione ortogonale $\hat{\mathbf{x}}_2$ di $\mathbf{x}_2$ nel sottospazio generato da $\tilde{\mathbf{x}}_1$ è data dal vettore $\tilde{\mathbf{x}}_2$ definito nella sezione 2.3.

Infatti,

$$\begin{aligned} \mathbf{x}_2 - \hat{\mathbf{x}}_2 &= \mathbf{x}_2 - \tilde{\mathbf{x}}_1 (\mathbf{G}^{(1)}(\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_1))^{-1}\, G(\tilde{\mathbf{x}}_1, \mathbf{x}_2) \\ &= \mathbf{x}_2 - \mathbf{x}_1\, g_{11}^{-1}\, g_{12} \qquad = \frac{g_{22}^{(2)}}{\Delta_1}\mathbf{x}_2 + \frac{g_{12}^{(2)}}{\Delta_1}\mathbf{x}_1 = \tilde{\mathbf{x}}_2 \end{aligned}$$

Fig. 5




Analogamente, come si può facilmente verificare, la differenza tra $\mathbf{x}_3$ e la proiezione ortogonale $\widehat{\mathbf{x}}_3$ di $\mathbf{x}_3$ nel sottospazio generato da $\widetilde{\mathbf{x}}_1$ e $\widetilde{\mathbf{x}}_2$ è data dal vettore $\widetilde{\mathbf{x}}_3$ definito nella sezione 2.3.

A conclusioni simili, a parte la laboriosità dei calcoli, si perviene per gli altri vettori.

Per esempio, la differenza tra $\mathbf{x}_p$ e la proiezione ortogonale $\widehat{\mathbf{x}}_p$ di $\mathbf{x}_p$ nel sottospazio generato da $\widetilde{\mathbf{x}}_1, \widetilde{\mathbf{x}}_2, \dots, \widetilde{\mathbf{x}}_{p-1}$ è data dal vettore $\widetilde{\mathbf{x}}_p$ definito nella sezione 2.3.

## 6 MATRICI ORTOGONALI. TRASFORMAZIONI ORTOGONALI

**6.1** Considerate due basi di uno spazio euclideo S formate da $\mathbf{x}_1, \dots, \mathbf{x}_p$ e $\widetilde{\mathbf{x}}_1, \dots, \widetilde{\mathbf{x}}_p$, siano $\mathbf{G}$ e $\widetilde{\mathbf{G}}$ le matrici rappresentative della metrica G relativamente alla prima e alla seconda base.

Come sappiamo (Cfr. la (3) del Cap. VI), la relazione che lega queste ultime matrici è del tipo

$$\widetilde{\mathbf{G}} = \mathbf{A}'\mathbf{G}\mathbf{A}$$

dove con $\mathbf{A}$ si è indicato, al solito, la matrice del passaggio da $\mathbf{X} = \left[\, \mathbf{x}_1 \cdots \mathbf{x}_p \,\right]$ a $\widetilde{\mathbf{X}} = \left[\, \widetilde{\mathbf{x}}_1 \cdots \widetilde{\mathbf{x}}_p \,\right]$.

Tenendo presente che, rispetto a una base ortonormale, la matrice della metrica G è la matrice unità, possiamo affermare che l'una o l'altra delle condizioni (equivalenti)

$$\mathbf{A'A} = \mathbf{I} \quad , \quad \mathbf{A'} = \mathbf{A}^{-1} \tag{10}$$

è necessaria e sufficiente affinché le basi costituite dai vettori $\mathbf{x}_1, \ldots, \mathbf{x}_p$ e $\widetilde{\mathbf{x}}_1, \ldots, \widetilde{\mathbf{x}}_p$ siano entrambe ortonormali, vale a dire affinché sia $\mathbf{G} = \mathbf{I} = \widetilde{\mathbf{G}}$.

Ogni matrice (invertibile) che soddisfa una relazione del tipo espresso nella (10) è detta matrice *ortogonale*.

Le matrici ortogonali godono delle seguenti proprietà elementari.

(a) Se $\mathbf{A}$ è una matrice ortogonale, allora $\det(\mathbf{A}) = \pm 1$. Infatti, tenuto conto della prima scritta in (10), è

$$1 = \det(\mathbf{I}) = \det(\mathbf{A'A}) = \det(\mathbf{A'})\det(\mathbf{A}) = \det^2(\mathbf{A})$$

da cui l'asserto.

(b) Se $\mathbf{A}$ è una matrice ortogonale, anche $\mathbf{A'}$ è ortogonale. Infatti, la seconda scritta in (10) può essere posta nella forma

$$\mathbf{AA'} = \mathbf{I}$$

da cui

$$(\mathbf{A'})'\mathbf{A'} = \mathbf{I} \; .$$

(c) Se $\mathbf{A}$ è una matrice ortogonale, anche $\mathbf{A}^{-1}$ è ortogonale. Infatti, dalla relazione scritta per ultimo, facendo uso della seconda scritta in (10), si ottiene

$$(\mathbf{A}^{-1})'\mathbf{A}^{-1} = \mathbf{I} \; .$$

(d) Se $\mathbf{A}$ e $\mathbf{B}$ sono due matrici ortogonali (dello stesso ordine), anche $\mathbf{AB}$ e $\mathbf{BA}$ sono ortogonali. Infatti,

$$(\mathbf{AB})'\mathbf{AB} = \mathbf{B'A'AB} = \mathbf{I} \quad , \quad (\mathbf{BA})'\mathbf{BA} = \mathbf{A'B'BA} = \mathbf{I} \; .$$

**6.2** Dato uno spazio euclideo S, si chiama *trasformazione* (*lineare*) *ortogonale* o *isometria* ogni trasformazione lineare T di S in se stesso tale che, per tutti gli $\mathbf{x}, \mathbf{y} \in S$, si abbia

$$G(\mathbf{x}, \mathbf{y}) = G(T(\mathbf{x}), T(\mathbf{y})) . \tag{11}$$

Indicata con $\mathbf{G}$ la matrice della metrica G di S rispetto a una base costituita dai vettori colonna di una matrice $\mathbf{X} = [\mathbf{x}_1 \cdots \mathbf{x}_p]$ e considerati due generici vettori $\mathbf{x} = \mathbf{Xa}$ e $\mathbf{y} = \mathbf{Xb}$ appartenenti a S, possiamo scrivere la (11) nella forma

$$\mathbf{a'Gb} = \mathbf{a'T'GTb}$$

ovvero

$$\mathbf{a'}(\mathbf{G} - \mathbf{T'GT})\mathbf{b} = 0$$

dove $\mathbf{T}$ è la matrice rappresentativa di T rispetto alla suddetta base di S.

Ora, quest'ultima relazione deve valere per ogni $\mathbf{a}, \mathbf{b} \in R^p$ e, in particolare, per ogni $\mathbf{u}_k, \mathbf{u}_j \in R^p$ dove $\mathbf{u}_k$ e $\mathbf{u}_j$ $(k,j = 1, \ldots, p)$ indicano, rispettivamente, il k-esimo e il j-esimo vettore canonico di ordine p.

Ne consegue che

$$\mathbf{G} - \mathbf{T'GT} = \mathbf{O}$$

ossia che

$$\mathbf{G} = \mathbf{T'GT} . \tag{12}$$

La (12) esprime dunque la condizione necessaria e sufficiente affinché la trasformazione lineare T sia ortogonale.

Ricordando quanto detto nella sezione precedente, possiamo anche affermare che condizione necessaria e sufficiente affinché la trasformazione lineare T sia ortogonale è che la matrice di T, rispetto a una base ortonormale di S, sia ortogonale.

Com'è ovvio, l'invarianza del prodotto scalare rispetto a una trasformazione ortogonale T implica l'invarianza di angoli, lunghezze, distanze.

## 7 AUTOVALORI E AUTOVETTORI DI UNA TRASFORMAZIONE AUTOAGGIUNTA

In precedenza (Cfr. l'Osservazione 13 del Cap. V), abbiamo sottolineato che non sempre è possibile diagonalizzare attraverso una trasformazione per similitudine la matrice **T** di una trasformazione lineare T di uno spazio vettoriale S in se stesso. Invece, nel caso in cui S sia uno spazio euclideo e T sia autoaggiunta, come mostreremo in questo paragrafo, **T** può essere ridotta a forma diagonale mediante una trasformazione per similitudine.

**TEOREMA 11**

Siano S uno spazio euclideo e T una trasformazione autoaggiunta di S.

Indichiamo con U la *sfera unitaria*, vale a dire l'insieme costituito da tutti i vettori $\mathbf{x} \in S$ per i quali risulta $G(\mathbf{x}, \mathbf{x}) = 1$.

Indichiamo poi con $\varphi$ la funzione definita ponendo, per tutti gli $\mathbf{x} \in S$,

$$\varphi(\mathbf{x}) = G(T(\mathbf{x}), \mathbf{x}) .$$

Allora, ciascun vettore $\widetilde{\mathbf{x}}_1 \in U$ (vettore unitario) in corrispondenza del quale $\varphi$ assume il valore massimo $\varphi(\widetilde{\mathbf{x}}_1)$ è un autovettore di T e $\varphi(\widetilde{\mathbf{x}}_1)$ è l'autovalore a esso associato.

*Dim*. Poiché U è un insieme chiuso e limitato in S e la funzione $\varphi$ è continua su U, per il teorema di Weierstrass, esiste almeno un vettore $\widetilde{\mathbf{x}}_1 \in U$ per il quale $\varphi$ assume il valore massimo $\varphi(\widetilde{\mathbf{x}}_1)$.

Inoltre, per un noto teorema sui massimi vincolati, esiste un (unico) numero reale $\widetilde{\lambda}_1$ (*moltiplicatore di Lagrange*) tale che

$$\frac{\partial \varphi(\widetilde{\mathbf{x}}_1)}{\partial \widetilde{\mathbf{x}}_1} = \widetilde{\lambda}_1 \frac{\partial G(\widetilde{\mathbf{x}}_1, \widetilde{\mathbf{x}}_1)}{\partial \widetilde{\mathbf{x}}_1} .$$

Sia adesso $\mathbf{T} = \mathbf{T}'$ la matrice della trasformazione autoaggiunta T rispetto a una base ortonormale di S costituita dai vettori colonna di una matrice $\mathbf{X} = [\, \mathbf{x}_1 \cdots \mathbf{x}_p \,]$.

Posto $\widetilde{\mathbf{x}}_1 = \mathbf{X}\,\widetilde{\mathbf{a}}_1$ e tenuto conto che (Cfr. il punto 4 dei Complementi al Cap. VI)

$$\frac{\partial\varphi(\widetilde{\mathbf{x}}_1)}{\partial\widetilde{\mathbf{x}}_1} = \frac{\partial(\widetilde{\mathbf{a}}_1'\mathbf{T}'\widetilde{\mathbf{a}}_1)}{\partial\widetilde{\mathbf{a}}_1} = 2\mathbf{T}\widetilde{\mathbf{a}}_1 \quad , \quad \frac{\partial G(\widetilde{\mathbf{x}}_1, \widetilde{\mathbf{x}}_1)}{\partial\widetilde{\mathbf{x}}_1} = \frac{\partial(\widetilde{\mathbf{a}}_1'\widetilde{\mathbf{a}}_1)}{\partial\widetilde{\mathbf{a}}_1} = 2\widetilde{\mathbf{a}}_1$$

si ha

$$\mathbf{T}\widetilde{\mathbf{a}}_1 = \widetilde{\lambda}_1\widetilde{\mathbf{a}}_1$$

e quindi $\widetilde{\mathbf{x}}_1 = \mathbf{X}\,\widetilde{\mathbf{a}}_1$ è un autovettore di T associato all'autovalore $\widetilde{\lambda}_1 = \varphi(\widetilde{\mathbf{x}}_1)$ $= \widetilde{\mathbf{a}}_1'\mathbf{T}\widetilde{\mathbf{a}}_1$.

**TEOREMA 12**

Siano S uno spazio euclideo e T una trasformazione autoaggiunta di S.

Indichiamo con $S^{\perp}(\widetilde{\mathbf{x}})$ il complemento ortogonale in S del sottospazio generato da un autovettore $\widetilde{\mathbf{x}}$ associato all'autovalore $\widetilde{\lambda}$ di T.

Allora, $S^{\perp}(\widetilde{\mathbf{x}})$ è un sottospazio di S di dimensione $p-1$ tale che la restrizione di T a $S^{\perp}(\widetilde{\mathbf{x}})$ è una trasformazione autoaggiunta di $S^{\perp}(\widetilde{\mathbf{x}})$.

*Dim*. Se $p=1$ non c'è niente da dimostrare. Supposto $p>1$, è ovvio che la dimensione di $S^{\perp}(\widetilde{\mathbf{x}})$ è $p-1$. Dobbiamo poi dimostrare che $T(\mathbf{x})$ appartiene a $S^{\perp}(\widetilde{\mathbf{x}})$ per ogni $\mathbf{x}$ appartenente a $S^{\perp}(\widetilde{\mathbf{x}})$, vale a dire che, per ogni $\mathbf{x}$ tale che $G(\mathbf{x}, \widetilde{\mathbf{x}}) = 0$, si ha $G(T(\mathbf{x}), \widetilde{\mathbf{x}}) = 0$.

Infatti, se $\widetilde{\lambda}$ è l'autovalore di T corrispondente all'autovettore $\widetilde{\mathbf{x}}$, allora

$$G(T(\mathbf{x}), \widetilde{\mathbf{x}}) = G(\mathbf{x}, T(\widetilde{\mathbf{x}})) = G(\mathbf{x}, \widetilde{\lambda}\widetilde{\mathbf{x}}) = \widetilde{\lambda}G(\mathbf{x}, \widetilde{\mathbf{x}}) = 0 .$$

**TEOREMA 13** (*Teorema spettrale*)

Siano S uno spazio euclideo e T una trasformazione autoaggiunta di S. Esiste allora una base ortonormale di S formata da autovettori di T (*base spettrale*).

*Dim*. La dimostrazione procede per induzione sulla dimensione di S.

Se $p=1$ il teorema è senz'altro vero. Supposto $p>1$, per il Teorema 11, esistono un autovalore $\tilde{\lambda}_1$ di T e un autovettore $\tilde{\mathbf{x}}_1$ a esso associato.

Considerato il complemento ortogonale $S^{\perp}(\tilde{\mathbf{x}}_1)$ in S del sottospazio di S generato da $\tilde{\mathbf{x}}_1$, per il Teorema 12, $S^{\perp}(\tilde{\mathbf{x}}_1)$ ha dimensione pari a $p-1$ e la restrizione di T a $S^{\perp}(\tilde{\mathbf{x}}_1)$ è una trasformazione autoaggiunta di $S^{\perp}(\tilde{\mathbf{x}}_1)$.

Per ipotesi di induzione, esiste allora una base di $S^{\perp}(\tilde{\mathbf{x}}_1)$ costituita da $p-1$ autovettori ortonormali $\tilde{\mathbf{x}}_2, \dots, \tilde{\mathbf{x}}_p$ di T i quali, insieme a $\tilde{\mathbf{x}}_1$, formano una base ortonormale di S. ▮

Dato uno spazio euclideo S con metrica G, siano **T** la matrice di una trasformazione autoaggiunta T di S e **G** la matrice di G, rispetto a una base di S costituita da p vettori $\mathbf{x}_1, \dots, \mathbf{x}_p$.

In forza del teorema ora dimostrato, esistono p autovettori ortonormali (e, quindi, linearmente indipendenti) $\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_p$ associati agli autovalori $\tilde{\lambda}_1, \dots, \tilde{\lambda}_p$ di T.

Posto

$$\mathbf{X} = \left[\mathbf{x}_1 \cdots \mathbf{x}_p\right] \quad , \quad \tilde{\mathbf{X}} = \left[\tilde{\mathbf{x}}_1 \cdots \tilde{\mathbf{x}}_p\right] = \left[\mathbf{x}_1 \cdots \mathbf{x}_p\right]\left[\tilde{\mathbf{a}}_1 \cdots \tilde{\mathbf{a}}_p\right] = \mathbf{X}\tilde{\mathbf{A}} \,,$$

$$\tilde{\mathbf{D}} = \operatorname{diag}(\tilde{\lambda}_1, \dots, \tilde{\lambda}_p)$$

risulta, dunque,

$$\tilde{\mathbf{A}}'\mathbf{G}\tilde{\mathbf{A}} = \mathbf{I} \tag{13}$$

e (Cfr. la sezione 5.4 del Cap. V)

$$\tilde{\mathbf{A}}^{-1}\mathbf{T}\tilde{\mathbf{A}} = \tilde{\mathbf{D}} \,. \tag{14}$$

La (14) indica che la matrice **T** può essere ridotta a forma diagonale attraverso una trasformazione per similitudine.

Nel seguente teorema, infine, è stabilita una importante proprietà degli autovettori associati ad autovalori distinti di una trasformazione autoaggiunta.

**TEOREMA 14**

Siano S uno spazio euclideo e T una trasformazione autoaggiunta di S. Autovettori associati ad autovalori distinti di T sono ortogonali.

*Dim*. Siano $\tilde{\lambda}_{*k}, \tilde{\lambda}_{*s}$ due autovalori distinti di T e $\tilde{\mathbf{x}}_{*k}, \tilde{\mathbf{x}}_{*s}$ due autovettori a essi associati.

Allora,

$$\begin{aligned} \tilde{\lambda}_{*k} G(\tilde{\mathbf{x}}_{*k}, \tilde{\mathbf{x}}_{*s}) &= G(\tilde{\lambda}_{*k} \tilde{\mathbf{x}}_{*k}, \tilde{\mathbf{x}}_{*s}) \\ = G(T(\tilde{\mathbf{x}}_{*k}), \tilde{\mathbf{x}}_{*s}) &= G(\tilde{\mathbf{x}}_{*k}, T(\tilde{\mathbf{x}}_{*s})) \\ = G(\tilde{\mathbf{x}}_{*k}, \tilde{\lambda}_{*s} \tilde{\mathbf{x}}_{*s}) &= \tilde{\lambda}_{*s} G(\tilde{\mathbf{x}}_{*k}, \tilde{\mathbf{x}}_{*s}) . \end{aligned}$$

Ma, essendo $\tilde{\lambda}_{*k} \neq \tilde{\lambda}_{*s}$, l'eguaglianza del primo e l'ultimo termine nella relazione precedente implica che sia $G(\tilde{\mathbf{x}}_{*k}, \tilde{\mathbf{x}}_{*s}) = 0$, ossia l'ortogonalità di $\tilde{\mathbf{x}}_{*k}, \tilde{\mathbf{x}}_{*s}$. ∎

Naturalmente, il procedimento operativo attraverso il quale − dato uno spazio euclideo S e una trasformazione autoaggiunta T di S − si perviene alla effettiva costruzione di una base ortonormale di S costituita da autovettori di T, base di cui il Teorema 13 afferma l'esistenza, è in sostanza quello stesso descritto nella sezione 5.2 del Cap. V.

Si tratta di determinare, anzitutto, in corrispondenza di ciascun autovalore distinto di T, un numero di autovettori linearmente indipendenti pari alla molteplicità dell'autovalore stesso e di ortogonalizzare, eventualmente, tali autovettori mediante il procedimento di Gram-Schmidt.

Successivamente, normalizzando ciascuno degli autovettori in questione si ottiene la base spettrale desiderata.

**OSSERVAZIONE 6.** Come facilmente si verifica, se i vettori colonna della matrice $\tilde{\mathbf{X}}_s$ di ordine $(p, m_s)$ sono autovettori ortonormali associati a un autovalore di molteplicità $m_s$, qualunque sia la matrice ortogonale $\mathbf{B}_s$ di ordine $(m_s, m_s)$, anche i vettori colonna di $\tilde{\mathbf{X}}_s \mathbf{B}_s$ sono autovettori ortonormali associati al medesimo autovalore.

Si noti, inoltre, che autovettori unitari associati a un autovalore semplice sono univocamente determinati a meno del segno.

**ESEMPIO 10.** Supponiamo $R^3$ munito della metrica rappresentata, rispetto alla base naturale di $R^3$, dalla matrice

$$\mathbf{G} = \begin{bmatrix} 14 & 20 & 29 \\ 20 & 29 & 42 \\ 29 & 42 & 61 \end{bmatrix}.$$

Consideriamo la trasformazione lineare T di $R^3$ in se stesso la cui matrice, rispetto alla suddetta base di $R^3$, è

$$\mathbf{T} = \begin{bmatrix} -6 & -9 & -13 \\ 6 & 9 & 13 \\ 0 & 0 & 0 \end{bmatrix}.$$

Come si verifica subito, T è una trasformazione autoaggiunta e ammette gli autovalori $\widetilde{\lambda}_{*1} = 3$ e $\widetilde{\lambda}_{*2} = 0$, di molteplicità $m_1 = 1$ e $m_2 = 2$.

Inoltre,

$$\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \;,\; \begin{bmatrix} 3/2 \\ -1 \\ 0 \end{bmatrix} \;,\; \begin{bmatrix} 13 \\ 0 \\ -6 \end{bmatrix}$$

sono autovettori di T associati agli autovalori di cui sopra.

Applicando a tali autovettori il procedimento di ortogonalizzazione di Gram-Schmidt e normalizzando, si ottengono poi i vettori

$$\widetilde{\mathbf{x}}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \frac{1}{\sqrt{3}} \;,\; \widetilde{\mathbf{x}}_2 = \begin{bmatrix} 3/2 \\ -1 \\ 0 \end{bmatrix} \sqrt{2} \;,\; \widetilde{\mathbf{x}}_3 = \begin{bmatrix} 1 \\ 8 \\ -6 \end{bmatrix} \frac{1}{\sqrt{6}}$$

che costituiscono una base ortonormale di $R^3$, formata da autovettori di T.

**OSSERVAZIONE 7.** Si considerino gli autovalori $\widetilde{\lambda}_1 \geq ... \geq \widetilde{\lambda}_p$ di una trasformazione autoaggiunta T di S e p autovettori ortonormali $\widetilde{\mathbf{x}}_1, ..., \widetilde{\mathbf{x}}_p$ associati a tali autovalori.

Come risulta dal Teorema 11, la funzione $\varphi$ definita ponendo, per ogni $\mathbf{x} \in S$,

$$\varphi(\mathbf{x}) = G(T(\mathbf{x}), \mathbf{x})$$

assume il valore massimo $\tilde{\lambda}_1$ in corrispondenza di un autovettore unitario $\tilde{\mathbf{x}}_1$.

Si può facilmente dimostrare che anche i successivi autovalori godono di una proprietà analoga a quella ora ricordata.

Per esempio, la funzione di cui sopra assume il valore massimo $\tilde{\lambda}_2$ in corrispondenza di un autovettore $\tilde{\mathbf{x}}_2$ tale che

$$G(\tilde{\mathbf{x}}_2, \tilde{\mathbf{x}}_2) = 1 \quad , \quad G(\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2) = 0 \, .$$

## 8 TRASFORMAZIONI AUTOAGGIUNTE E FORME BILINEARI SIMMETRICHE

**8.1** Dato uno spazio euclideo S, siano $\mathbf{G}$ la matrice della metrica G e $\mathbf{T}$ la matrice di una trasformazione autoaggiunta T ($\mathbf{T}'\mathbf{G} = \mathbf{G}\mathbf{T}$), rispetto a una base di S formata dai vettori colonna di una matrice $\mathbf{X}$.

Definita la matrice $\mathbf{C} = \mathbf{G}\mathbf{T}$ e verificato che questa è simmetrica in quanto

$$\mathbf{C}' = \mathbf{T}'\mathbf{G} = \mathbf{G}\mathbf{T} = \mathbf{C} \, ,$$

possiamo fare corrispondere in modo univoco a T la forma bilineare simmetrica B la cui matrice, rispetto alla suddetta base di S, è $\mathbf{C} = \mathbf{G}\mathbf{T}$.

Reciprocamente, rispetto a una base di S formata dai vettori colonna di una matrice $\mathbf{X}$, sia $\mathbf{C}$ la matrice di una forma bilineare simmetrica B.

Definita la matrice $\mathbf{T} = \mathbf{G}^{-1}\mathbf{C}$ e accertato che questa soddisfa la (8) in quanto

$$\mathbf{T}'\mathbf{G} = \mathbf{C}\mathbf{G}^{-1}\mathbf{G} = \mathbf{C} = \mathbf{G}\mathbf{G}^{-1}\mathbf{C} = \mathbf{G}\mathbf{T} \, ,$$

possiamo fare corrispondere in modo univoco a B la trasformazione autoaggiunta T la cui matrice, rispetto alla base di S sopra indicata, è $\mathbf{T} = \mathbf{G}^{-1}\mathbf{C}$.

In tal modo, si viene a istituire una corrispondenza biunivoca tra trasformazioni autoaggiunte e forme bilineari simmetriche che risulta inoltre indipendente dalla base di S formata dai vettori colonna di una matrice **X**.

Infatti, considerato un cambiamento di tale base determinato da una matrice **A**, si ha che $\mathbf{C} = \mathbf{GT}$ si trasforma in

$$\widetilde{\mathbf{C}} = \mathbf{A}'\mathbf{G}\mathbf{A}\mathbf{A}^{-1}\mathbf{T}\mathbf{A} = \mathbf{A}'\mathbf{G}\mathbf{T}\mathbf{A} = \mathbf{A}'\mathbf{C}\mathbf{A}$$

ovvero per congruenza e $\mathbf{T} = \mathbf{G}^{-1}\mathbf{C}$ si trasforma in

$$\widetilde{\mathbf{T}} = \mathbf{A}^{-1}\mathbf{G}^{-1}(\mathbf{A}')^{-1}\mathbf{A}'\mathbf{C}\mathbf{A} = \mathbf{A}^{-1}\mathbf{G}^{-1}\mathbf{C}\mathbf{A} = \mathbf{A}^{-1}\mathbf{T}\mathbf{A}$$

ovvero per similitudine.

**8.2** Si consideri nuovamente la matrice $\mathbf{T} = \mathbf{G}^{-1}\mathbf{C}$ definita nella sezione precedente. Come si è visto, rispetto a una base di S formata dai vettori colonna di una matrice **X**, questa si può interpretare come la matrice di una trasformazione autoaggiunta T univocamente associata a una forma bilineare simmetrica B la cui matrice, rispetto alla suddetta base di S, è **C**.

Considerata una nuova base di S formata da p autovettori ortonormali $\widetilde{\mathbf{x}}_1, \dots, \widetilde{\mathbf{x}}_p$ associati agli autovalori $\widetilde{\lambda}_1, \dots, \widetilde{\lambda}_p$ di T e posto

$$\widetilde{\mathbf{X}} = \left[\widetilde{\mathbf{x}}_1 \cdots \widetilde{\mathbf{x}}_p\right] = \mathbf{X}\widetilde{\mathbf{A}} \quad , \quad \widetilde{\mathbf{D}} = \mathrm{diag}\left(\widetilde{\lambda}_1, \dots, \widetilde{\lambda}_p\right)$$

risulta

$$\widetilde{\mathbf{A}}^{-1}\mathbf{G}^{-1}\mathbf{C}\widetilde{\mathbf{A}} = \widetilde{\mathbf{D}} \quad , \quad \widetilde{\mathbf{A}}'\mathbf{G}\widetilde{\mathbf{A}} = \mathbf{I}$$

da cui, essendo $\widetilde{\mathbf{A}}' = \widetilde{\mathbf{A}}^{-1}\mathbf{G}^{-1}$, si ha che

$$\widetilde{\mathbf{A}}'\mathbf{C}\widetilde{\mathbf{A}} = \widetilde{\mathbf{D}} \quad , \quad \widetilde{\mathbf{A}}'\mathbf{G}\widetilde{\mathbf{A}} = \mathbf{I}\,. \tag{15}$$

In altri termini, operando un cambiamento di base mediante la matrice $\widetilde{\mathbf{A}}$ di cui sopra, la matrice **C** viene ridotta a forma diagonale.

**OSSERVAZIONE 8.** Gli autovalori di T sono, ovviamente, le radici della equazione caratteristica di $\mathbf{G}^{-1}\mathbf{C}$, ovvero sono le soluzioni della equazione

$$\det(\mathbf{G}^{-1}\mathbf{C} - \lambda\mathbf{I}) = 0 \ .$$

Ma,

$$\begin{aligned} &\{\det(\mathbf{G}^{-1}\mathbf{C} - \lambda\mathbf{I}) = 0\} &&\Leftrightarrow \{\det(\mathbf{G})\ \det(\mathbf{G}^{-1}\mathbf{C} - \lambda\mathbf{I}) = 0\} \\ \Leftrightarrow\ &\{\det\left[\mathbf{G}(\mathbf{G}^{-1}\mathbf{C} - \lambda\mathbf{I})\right] = 0\} &&\Leftrightarrow \{\det(\mathbf{C} - \lambda\mathbf{G}) = 0\} \end{aligned}$$

e, quindi, tali autovalori possono anche essere ottenuti come soluzioni della equazione

$$\det(\mathbf{C} - \lambda\mathbf{G}) = 0 \ .$$

Inoltre, si verifica facilmente che, indicando al solito con $\tilde{\lambda}_{*s}$ $(s = 1, \ldots, t)$ l'autovalore distinto s-esimo di T, i vettori coordinati degli autovettori associati a tale autovalore possono essere determinati come soluzioni dell'uno o dell'altro dei due sistemi

$$(\mathbf{G}^{-1}\mathbf{C} - \tilde{\lambda}_{*s}\mathbf{I})\tilde{\mathbf{a}}_{*s} = \mathbf{0} \quad , \quad (\mathbf{C} - \tilde{\lambda}_{*s}\mathbf{G})\tilde{\mathbf{a}}_{*s} = \mathbf{0} \ .$$

**OSSERVAZIONE 9.** Data una matrice simmetrica $\mathbf{C}$ di ordine $(p,p)$, questa si può riguardare come matrice rappresentativa, rispetto alla base naturale di $R^p$, di una forma bilineare simmetrica B su $R^p$.

Nell'ipotesi che $R^p$ sia munito della metrica standard ($\mathbf{G} = \mathbf{I}$), $\mathbf{C}$ si può anche interpretare come matrice di una trasformazione autoaggiunta T.

Pertanto, considerata una nuova base di $R^p$ formata da p autovettori ortonormali $\tilde{\mathbf{a}}_1, \ldots, \tilde{\mathbf{a}}_p$ associati agli autovalori $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_p$ di T e posto

$$\tilde{\mathbf{A}} = \left[\, \tilde{\mathbf{a}}_1 \cdots \tilde{\mathbf{a}}_p \,\right] \quad , \quad \tilde{\mathbf{D}} = \mathrm{diag}\left(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_p\right)$$

risulta

$$\tilde{\mathbf{A}}^{-1}\mathbf{C}\tilde{\mathbf{A}} = \tilde{\mathbf{D}} \quad , \quad \tilde{\mathbf{A}}'\tilde{\mathbf{A}} = \mathbf{I}$$

da cui, essendo $\tilde{\mathbf{A}}' = \tilde{\mathbf{A}}^{-1}$, si ha che

$$\tilde{\mathbf{A}}'\mathbf{C}\tilde{\mathbf{A}} = \tilde{\mathbf{D}} \ .$$

**8.3** Nel capitolo precedente, abbiamo introdotto i concetti di forma quadratica definita positiva (negativa), semidefinita positiva (semidefinita negativa), indefinita.

Ci proponiamo adesso di caratterizzare tali forme prendendo le mosse da quanto detto qui sopra.

A questo fine, si osservi anzitutto che, supposto che gli autovalori che compongono la matrice $\widetilde{\mathbf{D}}$ che compare nella (15) siano tutti positivi, la forma quadratica Q associata alla forma bilineare simmetrica B, per ogni $\mathbf{x}$ appartenente a S della forma

$$\mathbf{x} = \tilde{a}_1 \widetilde{\mathbf{x}}_1 + \ldots + \tilde{a}_p \widetilde{\mathbf{x}}_p = \widetilde{\mathbf{X}} \tilde{\mathbf{a}} \,,$$

assume il valore

$$\text{(16)} \qquad Q(\mathbf{x}) = \tilde{\mathbf{a}}' \widetilde{\mathbf{D}} \tilde{\mathbf{a}} = \sum_{j=1}^{p} \tilde{\lambda}_j \tilde{a}_j^2$$

e, quindi, Q è definita positiva.

Reciprocamente, supposto che la forma quadratica Q sia definita positiva, è chiaro che deve essere $\tilde{\lambda}_j > 0$ per ogni $j = 1, \ldots, p$.

Allo stesso modo si dimostra che condizione necessaria e sufficiente affinché Q sia

(a) definita negativa,

(b) semidefinita positiva (semidefinita negativa),

(c) indefinita

è che gli autovalori che compaiono in $\widetilde{\mathbf{D}}$ siano, rispettivamente,

(*a*) tutti negativi,

(*b*) tutti non negativi (non positivi) e almeno uno eguale a zero,

(*c*) sia positivi sia negativi.

# Cap. VII
# COMPLEMENTI

**1**

Dato un vettore $\mathbf{x} \in R^p$, la matrice simmetrica di ordine $(p,p)$

$$\mathbf{I} - 2\mathbf{x}\mathbf{x}'/\mathbf{x}'\mathbf{x}$$

è detta *matrice di Householder*.

Si verifica subito che tale matrice è ortogonale. Infatti,

$$(\mathbf{I} - 2\mathbf{x}\mathbf{x}'/\mathbf{x}'\mathbf{x})'(\mathbf{I} - 2\mathbf{x}\mathbf{x}'/\mathbf{x}'\mathbf{x}) = \mathbf{I} - 4\mathbf{x}\mathbf{x}'/\mathbf{x}'\mathbf{x} + 4\mathbf{x}\mathbf{x}'/\mathbf{x}'\mathbf{x} = \mathbf{I}\ .$$

**2**

Data una matrice simmetrica $\mathbf{C}$ di ordine $(p,p)$, questa si può interpretare, rispetto alla base naturale di $R^p$ munito della metrica standard, come matrice rappresentativa di una trasformazione autoaggiunta T (Cfr. l'Osservazione 9 di questo capitolo).

Pertanto, considerata una nuova base di $R^p$ formata da p autovettori ortonormali $\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_p$ associati agli autovalori $\tilde{\lambda}_1, \dots, \tilde{\lambda}_p$ di T e posto

$$\tilde{\mathbf{A}} = \left[\tilde{\mathbf{a}}_1 \cdots \tilde{\mathbf{a}}_p\right] \quad , \quad \tilde{\mathbf{D}} = \mathrm{diag}\left(\tilde{\lambda}_1, \dots, \tilde{\lambda}_p\right)$$

risulta

$$\mathbf{C} = \tilde{\mathbf{A}}\tilde{\mathbf{D}}\tilde{\mathbf{A}}'\ .$$

Ora, nell'ipotesi che la forma quadratica Q associata a $\mathbf{C}$ sia definita positiva o semidefinita positiva, posto

$$\widetilde{\mathbf{D}}^{1/2} = \mathrm{diag}\left(\widetilde{\lambda}_1^{1/2}, \dots, \widetilde{\lambda}_p^{1/2}\right),$$

si ha ($\widetilde{\mathbf{A}}'\widetilde{\mathbf{A}} = \mathbf{I}$)

$$\mathbf{C} = \left(\widetilde{\mathbf{A}}\widetilde{\mathbf{D}}^{1/2}\widetilde{\mathbf{D}}^{1/2}\widetilde{\mathbf{A}}'\right) = \left(\widetilde{\mathbf{A}}\widetilde{\mathbf{D}}^{1/2}\widetilde{\mathbf{A}}'\right)\left(\widetilde{\mathbf{A}}\widetilde{\mathbf{D}}^{1/2}\widetilde{\mathbf{A}}'\right).$$

La matrice (simmetrica) $\widetilde{\mathbf{A}}\widetilde{\mathbf{D}}^{1/2}\widetilde{\mathbf{A}}'$ è detta la *radice quadrata* di $\mathbf{C}$ ed è indicata con $\mathbf{C}^{1/2}$.

Ovviamente, qualora Q sia definita positiva, è subito visto che esiste anche l'inversa $\mathbf{C}^{-1/2}$ di tale radice quadrata.

**3**

Supponiamo $R^p$ munito della metrica rappresentata, rispetto alla base naturale di $R^p$, da una matrice $\mathbf{Q}$ di ordine $(p,p)$.

Analogamente, supponiamo $R^n$ munito della metrica rappresentata, rispetto alla base naturale di $R^n$, da una matrice $\mathbf{M}$ di ordine $(n,n)$.

Data una matrice $\mathbf{Y}$ di ordine $(n,p)$ e $r(\mathbf{Y}) = r \le \min\{n,p\}$ − rappresentativa, rispetto alle suddette basi di $R^p$ e $R^n$, di una trasformazione lineare di $R^p$ in $R^n$ − e posto

$$\mathbf{V} = \mathbf{Y}'\mathbf{M}\mathbf{Y} \quad , \quad \mathbf{W} = \mathbf{Y}\mathbf{Q}\mathbf{Y}'$$

ci proponiamo di svolgere alcune considerazioni circa le matrici $\mathbf{VQ}$ e $\mathbf{WM}$.

(*a*) In luogo di $\mathbf{VQ}$, si consideri la matrice simmetrica $\mathbf{Q}^{1/2}\mathbf{V}\mathbf{Q}^{1/2}$.

Rispetto alla base naturale di $R^p$ dotato della metrica standard, $\mathbf{Q}^{1/2}\mathbf{V}\mathbf{Q}^{1/2}$ si può interpretare come la matrice di una trasformazione autoaggiunta $\overline{T}_1$ di $R^p$ (Cfr. l'Osservazione 9 di questo capitolo).

Pertanto, $\overline{T}_1$ possiede p autovettori ortonormali $\widetilde{\widetilde{\mathbf{a}}}_1, \dots, \widetilde{\widetilde{\mathbf{a}}}_p$ associati agli autovalori (non necessariamente distinti) $\widetilde{\lambda}_1, \dots, \widetilde{\lambda}_p$.

Posto

$$\widetilde{\widetilde{\mathbf{A}}} = \left[\,\widetilde{\widetilde{\mathbf{a}}}_1 \cdots \widetilde{\widetilde{\mathbf{a}}}_p\,\right] \quad , \quad \widetilde{\mathbf{D}}_{(\widetilde{\lambda})} = \mathrm{diag}(\widetilde{\lambda}_1, \dots, \widetilde{\lambda}_p)$$

si ha dunque che

$$(a_1) \qquad \mathbf{Q}^{1/2}\mathbf{V}\mathbf{Q}^{1/2}\tilde{\tilde{\mathbf{A}}} = \tilde{\tilde{\mathbf{A}}}\breve{\mathbf{D}}_{(\breve{\lambda})} \quad , \quad \tilde{\tilde{\mathbf{A}}}'\tilde{\tilde{\mathbf{A}}} = \mathbf{I}$$

ovvero che ($\tilde{\mathbf{A}} = \mathbf{Q}^{-1/2}\tilde{\tilde{\mathbf{A}}}$)

$$(a_2) \qquad \mathbf{V}\mathbf{Q}\tilde{\mathbf{A}} = \tilde{\mathbf{A}}\breve{\mathbf{D}}_{(\breve{\lambda})} \quad , \quad \tilde{\mathbf{A}}'\mathbf{Q}\tilde{\mathbf{A}} = \mathbf{I} \, .$$

Riguardata la matrice $\mathbf{VQ}$ come rappresentativa, rispetto alla base naturale di $R^p$, di una trasformazione lineare $T_1$ di $R^p$ in se stesso e verificato che $T_1$ è autoaggiunta ($(\mathbf{V}\mathbf{Q})'\mathbf{Q} = \mathbf{Q}(\mathbf{V}\mathbf{Q})$), dalla ($a_2$) si deduce che i vettori colonna di $\tilde{\mathbf{A}}$ sono autovettori ortonormali di $T_1$ associati ai corrispondenti autovalori.

Si noti che dalle ($a_2$) si ottiene la relazione

$$(a_3) \qquad \tilde{\mathbf{A}}'\mathbf{Q}\mathbf{V}\mathbf{Q}\tilde{\mathbf{A}} = \tilde{\mathbf{A}}'\mathbf{Q}\tilde{\mathbf{A}}\breve{\mathbf{D}}_{(\breve{\lambda})} = \breve{\mathbf{D}}_{(\breve{\lambda})} \, .$$

Ma, tenuto conto di quanto mostrato al punto 3 (*i*) dei Complementi al Cap. VI, si riconosce facilmente che $\mathbf{QVQ} = \mathbf{Q}\mathbf{Y}'\mathbf{M}\mathbf{Y}\mathbf{Q}$ è

- definita positiva se e soltanto se $r(\mathbf{Y}) = r = p$;
- semidefinita positiva se e soltanto se $r(\mathbf{Y}) = r < p$.

Ne consegue che, nel primo caso, tutti i p autovalori $\breve{\lambda}_1, \dots, \breve{\lambda}_p$ saranno positivi; nel secondo, r autovalori $\breve{\lambda}_1, \dots, \breve{\lambda}_r$ saranno positivi e $p - r$ nulli.

($b$) In luogo di $\mathbf{WM}$, si consideri la matrice simmetrica $\mathbf{M}^{1/2}\mathbf{W}\mathbf{M}^{1/2}$.

Rispetto alla base naturale di $R^n$ dotato della metrica standard, $\mathbf{M}^{1/2}\mathbf{W}\mathbf{M}^{1/2}$ si può interpretare come la matrice di una trasformazione autoaggiunta $\bar{T}_2$ di $R^n$ (Cfr. l'Osservazione 9 di questo capitolo).

Pertanto, $\bar{T}_2$ possiede n autovettori ortonormali $\tilde{\tilde{\mathbf{b}}}_1, \dots, \tilde{\tilde{\mathbf{b}}}_n$ associati agli autovalori (non necessariamente distinti) $\breve{\mu}_1, \dots, \breve{\mu}_n$.

Posto

$$\tilde{\tilde{\mathbf{B}}} = \left[ \tilde{\tilde{\mathbf{b}}}_1 \cdots \tilde{\tilde{\mathbf{b}}}_n \right] \quad , \quad \breve{\mathbf{D}}_{(\breve{\mu})} = \mathrm{diag}\left(\breve{\mu}_1, \dots, \breve{\mu}_n\right)$$

si ha dunque che

$(b_1)$ $$\mathbf{M}^{1/2}\mathbf{W}\mathbf{M}^{1/2}\tilde{\tilde{\mathbf{B}}} = \tilde{\tilde{\mathbf{B}}}\tilde{\mathbf{D}}_{(\tilde{\mu})} \quad , \quad \tilde{\tilde{\mathbf{B}}}'\tilde{\tilde{\mathbf{B}}} = \mathbf{I}$$

ovvero che ($\tilde{\mathbf{B}} = \mathbf{M}^{-1/2}\tilde{\tilde{\mathbf{B}}}$)

$(b_2)$ $$\mathbf{W}\mathbf{M}\tilde{\mathbf{B}} = \tilde{\mathbf{B}}\tilde{\mathbf{D}}_{(\tilde{\mu})} \quad , \quad \tilde{\mathbf{B}}'\mathbf{M}\tilde{\mathbf{B}} = \mathbf{I} \,.$$

Riguardata la matrice **WM** come rappresentativa, rispetto alla base naturale di $R^n$, di una trasformazione lineare $T_2$ di $R^n$ in se stesso e verificato che $T_2$ è autoaggiunta ($(\mathbf{WM})'\mathbf{M} = \mathbf{M}(\mathbf{WM})$), dalla $(b_2)$ si deduce che i vettori colonna di $\tilde{\mathbf{B}}$ sono autovettori ortonormali di $T_2$ associati ai corrispondenti autovalori.

Si noti che dalle $(b_2)$ si ottiene la relazione

$(b_3)$ $$\tilde{\mathbf{B}}'\mathbf{M}\mathbf{W}\mathbf{M}\tilde{\mathbf{B}} = \tilde{\mathbf{B}}'\mathbf{M}\tilde{\mathbf{B}}\tilde{\mathbf{D}}_{(\tilde{\mu})} = \tilde{\mathbf{D}}_{(\tilde{\mu})} \,.$$

Ma, tenuto conto di quanto mostrato al punto 3 (*ii*) dei Complementi al Cap. VI, si riconosce facilmente che $\mathbf{MWM} = \mathbf{MYQY}'\mathbf{M}$ è

- definita positiva se e soltanto se $r(\mathbf{Y}) = r = n$;
- semidefinita positiva se e soltanto se $r(\mathbf{Y}) = r < n$.

Ne consegue che, nel primo caso, tutti gli n autovalori $\tilde{\mu}_1, \dots, \tilde{\mu}_n$ saranno positivi; nel secondo, r autovalori $\tilde{\mu}_1, \dots, \tilde{\mu}_r$ saranno positivi e $n-r$ nulli.

(*c*) A parziale completamento di quanto detto precede, si osservi anzitutto, ricordando quanto mostrato al punto 4 dei Complementi al Cap. V, che $T_1$ e $T_2$ hanno i medesimi r autovalori positivi, cioè risulta $\tilde{\lambda}_1 = \tilde{\mu}_1, \dots, \tilde{\lambda}_r = \tilde{\mu}_r$.

Ciò premesso, indichiamo con $\tilde{\mathbf{A}}_1$ la matrice di ordine $(p,r)$ i cui vettori colonna $\tilde{\mathbf{a}}_1 = \mathbf{Q}^{-1}\tilde{\tilde{\mathbf{a}}}_1, \dots, \tilde{\mathbf{a}}_r = \mathbf{Q}^{-1}\tilde{\tilde{\mathbf{a}}}_r$ sono autovettori ortonormali di $T_1$, rispetto alla metrica rappresentata da **Q**, associati agli autovalori $\tilde{\lambda}_1, \dots, \tilde{\lambda}_r$.

Allora, i vettori colonna di ($\tilde{\mathbf{D}}_1^{-1/2} = \mathrm{diag}(\tilde{\lambda}_1^{-1/2}, \dots, \tilde{\lambda}_r^{-1/2})$)

$$\tilde{\mathbf{B}}_1 = \mathbf{Y}\mathbf{Q}\tilde{\mathbf{A}}_1\tilde{\mathbf{D}}_1^{-1/2}$$

sono autovettori ortonormali di $T_2$, rispetto alla metrica rappresentata da **M**, associati ai medesimi autovalori $\tilde{\lambda}_1, \dots, \tilde{\lambda}_r$.

Infatti, dalla relazione ($\widetilde{\mathbf{D}}_1 = \mathrm{diag}(\tilde{\lambda}_1, \dots, \tilde{\lambda}_r)$)

$$\mathbf{Y'MYQ}\widetilde{\mathbf{A}}_1 = \widetilde{\mathbf{A}}_1 \widetilde{\mathbf{D}}_1 \ ,$$

premoltiplicando ambo i membri per **YQ** e postmoltiplicando per $\widetilde{\mathbf{D}}_1^{-1/2}$, si ha

$$\mathbf{YQY'M}(\mathbf{YQ}\widetilde{\mathbf{A}}_1 \widetilde{\mathbf{D}}_1^{-1/2}) = \mathbf{YQ}\widetilde{\mathbf{A}}_1 \widetilde{\mathbf{D}}_1 \widetilde{\mathbf{D}}_1^{-1/2} = (\mathbf{YQ}\widetilde{\mathbf{A}}_1 \widetilde{\mathbf{D}}_1^{-1/2})\, \widetilde{\mathbf{D}}_1 \ .$$

Reciprocamente, indichiamo con $\widetilde{\mathbf{B}}_1$ la matrice di ordine $(n,r)$ i cui vettori colonna $\tilde{\mathbf{b}}_1 = \mathbf{M}^{-1}\tilde{\tilde{\mathbf{b}}}_1, \dots, \tilde{\mathbf{b}}_r = \mathbf{M}^{-1}\tilde{\tilde{\mathbf{b}}}_r$ sono autovettori ortonormali di $T_2$, rispetto alla metrica rappresentata da **M**, associati agli autovalori $\tilde{\lambda}_1, \dots, \tilde{\lambda}_r$.

Allora i vettori colonna di

$$\widetilde{\mathbf{A}}_1 = \mathbf{Y'M}\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{D}}_1^{-1/2}$$

sono autovettori ortonormali di $T_1$, rispetto alla metrica rappresentata da **Q**, associati agli autovalori $\tilde{\lambda}_1, \dots, \tilde{\lambda}_r$.

Infatti, dalla relazione

$$\mathbf{YQY'M}\widetilde{\mathbf{B}}_1 = \widetilde{\mathbf{B}}_1 \widetilde{\mathbf{D}}_1 \ ,$$

premoltiplicando ambo i membri per **Y'M** e postmoltiplicando per $\widetilde{\mathbf{D}}_1^{-1/2}$, si ha

$$\mathbf{Y'MYQ}(\mathbf{Y'M}\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{D}}_1^{-1/2}) = \mathbf{Y'M}\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{D}}_1 \widetilde{\mathbf{D}}_1^{-1/2} = (\mathbf{Y'M}\widetilde{\mathbf{B}}_1 \widetilde{\mathbf{D}}_1^{-1/2})\, \widetilde{\mathbf{D}}_1 \ .$$

**4**

Supponiamo $R^p$ munito della metrica rappresentata, rispetto alla base naturale di $R^p$, da una matrice **Q** di ordine $(p,p)$.

Analogamente, supponiamo $R^n$ munito della metrica rappresentata, rispetto alla base naturale di $R^n$, da una matrice **M** di ordine $(n,n)$.

Data una matrice **Y** di ordine $(n,p)$ e $r(\mathbf{Y}) = r \leq \min\{n,p\}$ – rappresentativa, rispetto alle suddette basi di $R^p$ e $R^n$, di una trasformazione lineare di $R^p$ in $R^n$ – sussiste la possibilità di *scomporre* (*fattorizzare*) **Y** nel modo seguente

$$(*) \qquad \mathbf{Y} = \widetilde{\mathbf{B}} \begin{bmatrix} \widetilde{\mathbf{S}}_{(r,r)} & \mathbf{O}_{(r,p-r)} \\ \mathbf{O}_{(n-r,r)} & \mathbf{O}_{(n-r,p-r)} \end{bmatrix} \widetilde{\mathbf{A}}'$$

dove

$$\widetilde{\mathbf{A}}'\mathbf{Q}\widetilde{\mathbf{A}} = \mathbf{I}_{(p,p)} \quad , \quad \widetilde{\mathbf{B}}'\mathbf{M}\widetilde{\mathbf{B}} = \mathbf{I}_{(n,n)} \quad , \quad \widetilde{\mathbf{S}} = \mathrm{diag}(s_1, \ldots, s_r)$$

e i numeri (reali) $s_1, \ldots, s_r$ che compaiono in $\widetilde{\mathbf{S}}$ sono tali che $s_h > 0$ per $h = 1, \ldots, r$.

Questi numeri ricevono la denominazione di *valori singolari* di $\mathbf{Y}$ e la $(*)$ è detta *scomposizione* (*fattorizzazione*) *di* $\mathbf{Y}$ *mediante valori singolari*.

Dati per acquisiti i risultati di cui al punto precedente, la dimostrazione di quanto asserito procede nel modo qui sotto indicato.

Supponiamo che i vettori colonna di $\widetilde{\mathbf{A}}_1$ e $\widetilde{\mathbf{A}}_2$ siano autovettori ortonormali associati, rispettivamente, agli autovalori (positivi) $\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_r$ e all'autovalore nullo della trasformazione autoaggiunta $T_1$ rappresentata dalla matrice $\mathbf{Y}'\mathbf{MYQ}$.

Allora,

$$(i) \qquad \mathbf{Y}'\mathbf{MYQ}\widetilde{\mathbf{A}}_1 = \widetilde{\mathbf{A}}_1 \widetilde{\mathbf{D}}_1 = \widetilde{\mathbf{A}}_1 \widetilde{\mathbf{S}}^2 \quad , \quad \widetilde{\mathbf{A}}_1'\mathbf{Q}\widetilde{\mathbf{A}}_1 = \mathbf{I}_{(r,r)}$$

con

$$\widetilde{\mathbf{S}} = \mathrm{diag}\left(\widetilde{\lambda}_1^{1/2}, \ldots, \widetilde{\lambda}_r^{1/2}\right) = \widetilde{\mathbf{D}}_1^{1/2}$$

e anche

$$(ii) \qquad \mathbf{Y}'\mathbf{MYQ}\widetilde{\mathbf{A}}_2 = \mathbf{O}_{(p-r,p-r)} \quad , \quad \widetilde{\mathbf{A}}_2'\mathbf{Q}\widetilde{\mathbf{A}}_2 = \mathbf{I}_{(p-r,p-r)} \,.$$

Analogamente, supponiamo che i vettori colonna di $\widetilde{\mathbf{B}}_1$ e $\widetilde{\mathbf{B}}_2$ siano autovettori ortonormali associati, rispettivamente, agli autovalori (positivi) $\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_r$ e all'autovalore nullo della trasformazione autoaggiunta $T_2$ rappresentata dalla matrice $\mathbf{YQY}'\mathbf{M}$.

Quindi,

$$(iii) \qquad \widetilde{\mathbf{B}}_1 = \mathbf{YQ}\widetilde{\mathbf{A}}_1\widetilde{\mathbf{S}}^{-1} \quad , \quad \widetilde{\mathbf{B}}_1'\mathbf{M}\widetilde{\mathbf{B}}_1 = \mathbf{I}_{(r,r)}$$

e anche

$$(iv) \qquad \mathbf{YQY'M\widetilde{B}}_2 = \mathbf{O}_{(n-r,\,n-r)} \quad , \quad \mathbf{\widetilde{B}}_2'\mathbf{M\widetilde{B}}_2 = \mathbf{I}_{(n-r,\,n-r)} \,.$$

Ma, dalla prima delle eguaglianze in (*ii*), premoltiplicando ambo i membri per $\mathbf{\widetilde{A}}_2'\mathbf{Q}$, si ottiene

$$\mathbf{\widetilde{A}}_2'\mathbf{QY'MYQ\widetilde{A}}_2 = \mathbf{O}_{(p-r,\,p-r)}$$

da cui

$$(v) \qquad \mathbf{YQ\widetilde{A}}_2 = \mathbf{O}_{(p-r,\,p-r)} \,.$$

A sua volta, dalla prima delle eguaglianze in (*iv*), premoltiplicando ambo i membri per $\mathbf{\widetilde{B}}_2'\mathbf{M}$, si ottiene

$$\mathbf{\widetilde{B}}_2'\mathbf{MYQY'M\widetilde{B}}_2 = \mathbf{O}_{(n-r,\,n-r)}$$

da cui

$$(vi) \qquad \mathbf{Y'M\widetilde{B}}_2 = \mathbf{O}_{(n-r,\,n-r)} \,.$$

Ciò premesso – posto

$$\mathbf{\widetilde{A}} = \left[\mathbf{\widetilde{A}}_1 \; \mathbf{\widetilde{A}}_2\right] \quad , \quad \mathbf{\widetilde{B}} = \left[\mathbf{\widetilde{B}}_1 \; \mathbf{\widetilde{B}}_2\right]$$

– possiamo scrivere, con notazione semplificata,

$$\mathbf{\widetilde{B}}'\mathbf{MYQ\widetilde{A}} = \begin{bmatrix} \mathbf{\widetilde{B}}_1' \\ \mathbf{\widetilde{B}}_2' \end{bmatrix} \mathbf{MYQ}\left[\mathbf{\widetilde{A}}_1 \; \mathbf{\widetilde{A}}_2\right] = \begin{bmatrix} \mathbf{\widetilde{B}}_1'\mathbf{MYQ\widetilde{A}}_1 & \mathbf{\widetilde{B}}_1'\mathbf{MYQ\widetilde{A}}_2 \\ \mathbf{\widetilde{B}}_1'\mathbf{MYQ\widetilde{A}}_1 & \mathbf{\widetilde{B}}_2'\mathbf{MYQ\widetilde{A}}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{\widetilde{S}} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}$$

e quindi, risultando $\mathbf{\widetilde{B}}'\mathbf{M} = \mathbf{B}^{-1}$ e $\mathbf{\widetilde{A}}'\mathbf{M} = \mathbf{A}^{-1}$, si ha infine (*scomposizione piena*)

$$\mathbf{Y} = \mathbf{\widetilde{B}} \begin{bmatrix} \mathbf{\widetilde{S}} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} \mathbf{\widetilde{A}}'$$

e anche (*scomposizione ridotta*)

$$\mathbf{Y} = \widetilde{\mathbf{B}}_1 \widetilde{\mathbf{S}} \widetilde{\mathbf{A}}_1' = \sum_{j=1}^{r} s_j \widetilde{\mathbf{b}}_j \widetilde{\mathbf{a}}_j' \,.$$

Si noti che, mentre i valori singolari $s_1, \ldots, s_r$ di $\mathbf{Y}$ sono univocamente determinati, ciò non accade per i vettori colonna delle matrici $\widetilde{\mathbf{A}}$ e $\widetilde{\mathbf{B}}$. In effetti, autovettori associati ad autovalori semplici sono univocamente determinati a meno del segno, mentre autovettori associati a un autovalore di molteplicità maggiore di uno sono determinati a meno di postmoltiplicazione per una matrice ortogonale (Cfr. l'Osservazione 6 di questo capitolo).

**5**

La matrice

$$\mathbf{Y}^{+}_{(p,n)} = \widetilde{\mathbf{A}} \begin{bmatrix} \widetilde{\mathbf{S}}^{-1}_{(r,r)} & \mathbf{O}_{(r,n-r)} \\ \mathbf{O}_{(p-r,r)} & \mathbf{O}_{(p-r,n-r)} \end{bmatrix} \widetilde{\mathbf{B}}'$$

– dove $\widetilde{\mathbf{S}}$, $\widetilde{\mathbf{A}}$, $\widetilde{\mathbf{B}}$ sono definite come nel punto precedente – è detta la *pseudoinversa* di $\mathbf{Y}_{(n,p)}$.

Come si può facilmente verificare, $\mathbf{Y}^{+}$ gode delle seguenti proprietà elementari ($\mathbf{M}$, $\mathbf{Q}$ matrici definite positive di ordine appropriato):

(i) $\mathbf{Y}\mathbf{Y}^{+}\mathbf{Y} = \mathbf{Y}$ (ii) $\mathbf{Y}^{+}\mathbf{Y}\mathbf{Y}^{+} = \mathbf{Y}^{+}$

(iii) $\mathbf{M}\mathbf{Y}\mathbf{Y}^{+} = (\mathbf{Y}\mathbf{Y}^{+})'\mathbf{M}$ (iv) $\mathbf{Q}\mathbf{Y}^{+}\mathbf{Y} = (\mathbf{Y}^{+}\mathbf{Y})'\mathbf{Q}$ .

Si può, inoltre, dimostrare che:

(a) se $\mathbf{Y}$ è invertibile, allora $\mathbf{Y}^{+} = \mathbf{Y}^{-1}$

(b) se $r(\mathbf{Y}) = p$, allora $\mathbf{Y}^{+} = (\mathbf{Y}'\mathbf{M}\mathbf{Y})^{-1}\mathbf{Y}'\mathbf{M}$

(c) se $r(\mathbf{Y}) = n$, allora $\mathbf{Y}^{+} = \mathbf{Q}\mathbf{Y}'(\mathbf{Y}\mathbf{Q}\mathbf{Y}')^{-1}$.

**6**

Supponiamo $R^n$ munito della metrica rappresentata, rispetto alla base naturale di $R^n$, da una matrice $\mathbf{M}$ di ordine $(n,n)$.

Siano $\mathbf{Y}_1$ e $\mathbf{Y}_2$ due matrici di ordine, rispettivamente, $(n, p_1)$ e $(n, p_2)$ tali che $p_1 \le p_2$ , $r(\mathbf{Y}_1) = p_1$ , $r(\mathbf{Y}_2) = p_2$ .

Posto

$$\mathbf{V}_{11\,(p_1,p_1)} = \mathbf{Y}_1'\mathbf{M}\mathbf{Y}_1 \quad , \quad \mathbf{V}_{22\,(p_2,p_2)} = \mathbf{Y}_2'\mathbf{M}\mathbf{Y}_2$$
$$\mathbf{V}_{21\,(p_2,p_1)} = \mathbf{Y}_2'\mathbf{M}\mathbf{Y}_1 \quad , \quad \mathbf{V}_{12\,(p_1,p_2)} = \mathbf{Y}_1'\mathbf{M}\mathbf{Y}_2$$

ci proponiamo di svolgere alcune considerazioni circa le matrici

$$\mathbf{V}_{21} \ , \ \mathbf{V}_{12} \ , \mathbf{V} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix} \ , \ \mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21} \ , \ \mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12} \ .$$

(*1*) Per quanto riguarda la matrice $\mathbf{V}_{21}$, risulta

$$0 \le r = r(\mathbf{V}_{21}) = r(\mathbf{Y}_2'\mathbf{M}\mathbf{Y}_1) = r(\mathbf{Y}_2'\mathbf{M}^{1/2}\mathbf{M}^{1/2}\mathbf{Y}_1) \le \min\{r(\mathbf{Y}_2)\,, r(\mathbf{Y}_1)\} = p_1 \ .$$

Analogamente, poiché $\mathbf{V}_{12} = \mathbf{V}_{21}'$, si ha che

$$0 \le r = r(\mathbf{V}_{12}) \le p_1 \ .$$

Si noti che, pur essendo $\mathbf{Y}_1$ e $\mathbf{Y}_2$ di pieno rango per colonne, può risultare $r(\mathbf{V}_{21}) = r(\mathbf{V}_{12}) < p_1$, come è illustrato dal seguente esempio.

Siano

$$\mathbf{Y}_1 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & -1 \end{bmatrix} \quad , \quad \mathbf{Y}_2 = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 1 & 1 \end{bmatrix} .$$

Allora, supposto $\mathbf{M} = \mathbf{I}$, si ha

$$\mathbf{V}_{21} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \quad , \quad \mathbf{V}_{21} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

ed è $r(\mathbf{V}_{21}) = r(\mathbf{V}_{12}) = 1$.

(*2*) Definita la matrice $\mathbf{Y} = [\mathbf{Y}_1\,\mathbf{Y}_2]$, possiamo scrivere

$$\mathbf{V} = \mathbf{Y}'\mathbf{M}\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1' \\ \mathbf{Y}_2' \end{bmatrix} \mathbf{M} \begin{bmatrix} \mathbf{Y}_1 \ \mathbf{Y}_2 \end{bmatrix} .$$

Pertanto, $\mathbf{V}$ è (Cfr. il punto 3 (*i*) dei Complementi al Cap. VI)

- definita positiva se e soltanto se $r(\mathbf{Y}) = p_1 + p_2$;
- semidefinita positiva se e soltanto se $r(\mathbf{Y}) < p_1 + p_2$.

(*3*) Per quanto riguarda le matrici $\mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}$ e $\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}$, risulta anzitutto che

$$
\begin{aligned}
& r(\mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}) = r(\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}) = r(\mathbf{V}_{21}) = r \\
& = r(\mathbf{V}_{21}) \qquad\qquad = r(\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}) = r(\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}) .
\end{aligned}
$$

Inoltre, tenuto conto di quanto detto nel corso della dimostrazione dei Teotemi 7 e 6 del Cap. IV, si ha che

$$
\begin{aligned}
& \mathcal{N}(\mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}) = \mathcal{N}(\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}) = \mathcal{N}(\mathbf{V}_{21}) \\
& = \mathcal{N}(\mathbf{V}_{12}) \qquad\qquad = \mathcal{N}(\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}) = \mathcal{N}(\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12})
\end{aligned}
$$

e

$$
\begin{aligned}
& \dim(\mathcal{N}(\mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21})) = \dim(\mathcal{N}(\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21})) = \dim(\mathcal{N}(\mathbf{V}_{21})) = p_1 - r , \\
& \dim(\mathcal{N}(\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12})) = \dim(\mathcal{N}(\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12})) = \dim(\mathcal{N}(\mathbf{V}_{12})) = p_2 - r .
\end{aligned}
$$

**7**

Dati per acquisiti le notazioni e i concetti di cui al punto precedente, ci proponiamo di svolgere le seguenti considerazioni.

(a) Qualunque sia il numero reale $\lambda$, risulta

$$
\begin{aligned}
& (-\lambda)^{2p_1} \det\begin{bmatrix} -\lambda\mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & -\lambda\mathbf{V}_{22} \end{bmatrix} \\
& = (-\lambda)^{p_1+p_2} \det\mathbf{V}_{11} \det\mathbf{V}_{22} \det\left(\lambda^2 \mathbf{I}_{(p_1,\,p_1)} - \mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\right) .
\end{aligned}
$$

Infatti, la relazione ora scritta è certamente vera per $\lambda = 0$.

Qualora sia $\lambda \neq 0$, sviluppando il determinante al primo membro di tale relazione, si ottiene (Cfr. i punti 5 e 3 dei Complementi al Cap. III)

$$(-\lambda)^{2p_1} \det\begin{bmatrix} -\lambda\mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & -\lambda\mathbf{V}_{22} \end{bmatrix}$$

$$= (-\lambda)^{2p_1} \det(-\lambda\mathbf{V}_{22}) \det\left(-\lambda\mathbf{V}_{11} + \lambda^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\right)$$

$$= (-\lambda)^{2p_1} \det(-\lambda\mathbf{V}_{22}) \det\mathbf{V}_{11} \det\left(-\lambda\mathbf{I}_{(p_1,p_1)} + \lambda^{-1}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\right)$$

$$= (-\lambda)^{2p_1} \det(-\lambda\mathbf{V}_{22}) \det\mathbf{V}_{11} \det(\lambda^{-1}\mathbf{I}_{(p_1,p_1)}) \det\left(\lambda^2\mathbf{I}_{(p_1,p_1)} - \mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\right)$$

$$= (-\lambda)^{2p_1} (-\lambda)^{p_1} (-\lambda)^{-p_2} \det\mathbf{V}_{11} \det\mathbf{V}_{22} \det\left(\lambda^2\mathbf{I}_{(p_1,p_1)} - \mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\right)$$

$$= (-\lambda)^{p_1+p_2} \det\mathbf{V}_{11} \det\mathbf{V}_{22} \det\left(\lambda^2\mathbf{I}_{(p_1,p_1)} - \mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\right) .$$

(b) Se $\widetilde{\lambda}$ è una soluzione (radice) dell'equazione

$$0 = (-\lambda)^{2p_1} \det\begin{bmatrix} -\lambda\mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & -\lambda\mathbf{V}_{22} \end{bmatrix} = \det\begin{bmatrix} -\lambda\mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & -\lambda\mathbf{V}_{22} \end{bmatrix},$$

$\widetilde{\lambda}^2$ è una soluzione dell'equazione

$$\begin{aligned} 0 &= (-\lambda)^{p_1+p_2} \det\mathbf{V}_{11} \det\mathbf{V}_{22} \det\left(\lambda^2\mathbf{I}_{(p_1,p_1)} - \mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\right) \\ &= \det\left(\lambda^2\mathbf{I}_{(p_1,p_1)} - \mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\right) \end{aligned}$$

e viceversa.

(c) Chiaramente, la matrice $\mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}$ che compare nell'equazione precedente può essere interpretata come la matrice di una trasformazione autoaggiunta $\widetilde{T}_1$ di $R^{p_1}$ rispetto alla base naturale di $R^{p_1}$ dotato della metrica rappresentata dalla matrice $\mathbf{V}_{11}$ (Cfr. la sezione 8.1 di questo capitolo).

Pertanto, $\widetilde{T}_1$ ammette $p_1$ autovalori (non necessariamente distinti) $\widetilde{\lambda}_1^2, \dots, \widetilde{\lambda}_{p_1}^2$.

Esiste inoltre una matrice $\widetilde{\mathbf{A}}$ i cui vettori colonna sono autovettori della trasformazione lineare in oggetto tale che ($\widetilde{\mathbf{D}}_{(p_1,p_1)} = \mathrm{diag}(\widetilde{\lambda}_1^2, \dots, \widetilde{\lambda}_{p_1}^2)$)

$$\widetilde{\mathbf{A}}^{-1}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\widetilde{\mathbf{A}} = \widetilde{\mathbf{D}}_{(p_1,p_1)} \quad , \quad \widetilde{\mathbf{A}}'\mathbf{V}_{11}\widetilde{\mathbf{A}} = \mathbf{I}_{(p_1,p_1)}$$

e anche ($\widetilde{\mathbf{A}}' = \widetilde{\mathbf{A}}^{-1}\mathbf{V}_{11}^{-1}$)

$$\widetilde{\mathbf{A}}'\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\widetilde{\mathbf{A}} = \widetilde{\mathbf{D}}_{(p_1,p_1)} \quad , \quad \widetilde{\mathbf{A}}'\mathbf{V}_{11}\widetilde{\mathbf{A}} = \mathbf{I}_{(p_1,p_1)} .$$

Ricordando adesso che $r(\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}) = r(\mathbf{V}_{21}) = r$, ne consegue che il numero degli autovalori non nulli di $\widetilde{T}_1$ è pari a r, mentre l'autovalore nullo si presenta con molteplicità pari a $p_1 - r$.

(d) Con una dimostrazione del tutto analoga a quella condotta in (a), si prova che, qualunque sia il numero reale $\lambda$, risulta

$$(-\lambda)^{2p_1} \det\begin{bmatrix} -\lambda\mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & -\lambda\mathbf{V}_{22} \end{bmatrix}$$
$$= (-\lambda)^{p_1+p_2}\det\mathbf{V}_{11}\det\mathbf{V}_{22}\det\left(\lambda^2\mathbf{I}_{(p_2,p_2)} - \mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}\right) .$$

(e) Se $\widetilde{\lambda}$ è una soluzione (radice) dell'equazione

$$0 = (-\lambda)^{2p_2} \det\begin{bmatrix} -\lambda\mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & -\lambda\mathbf{V}_{22} \end{bmatrix} = \det\begin{bmatrix} -\lambda\mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & -\lambda\mathbf{V}_{22} \end{bmatrix},$$

$\widetilde{\lambda}^2$ è una soluzione dell'equazione

$$\begin{aligned} 0 &= (-\lambda)^{p_1+p_2}\det\mathbf{V}_{11}\det\mathbf{V}_{22}\det\left(\lambda^2\mathbf{I}_{(p_2,p_2)} - \mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}\right) \\ &= \det\left(\lambda^2\mathbf{I}_{(p_2,p_2)} - \mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}\right) \end{aligned}$$

e viceversa.

(f) Chiaramente, la matrice $\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}$ che compare nell'equazione precedente può essere interpretata come la matrice di una trasformazione autoaggiunta $\widetilde{T}_2$ di $R^{p_2}$ rispetto alla base naturale di $R^{p_2}$ dotato della metrica rappresentata dalla matrice $\mathbf{V}_{22}$ (Cfr. la sezione 8.1 di questo capitolo).

Pertanto, $\widetilde{T}_2$ ammette $p_2$ autovalori (non necessariamente distinti) $\widetilde{\lambda}_1^2, \dots, \widetilde{\lambda}_{p_2}^2$.

Esiste inoltre una matrice $\widetilde{\mathbf{B}}$ i cui vettori colonna sono autovettori della trasformazione lineare in oggetto tale che ($\widetilde{\mathbf{D}}_{(p_2,p_2)} = \mathrm{diag}(\widetilde{\lambda}_1^2, \dots, \widetilde{\lambda}_{p_2}^2)$)

$$\widetilde{\mathbf{B}}^{-1}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}\widetilde{\mathbf{B}} = \widetilde{\widetilde{\mathbf{D}}}_{(p_2,p_2)} \quad , \quad \widetilde{\mathbf{B}}'\mathbf{V}_{22}\widetilde{\mathbf{B}} = \mathbf{I}_{(p_2,p_2)}$$

e anche ($\widetilde{\mathbf{B}}' = \widetilde{\mathbf{B}}^{-1}\mathbf{V}_{22}^{-1}$)

$$\tilde{\mathbf{B}}'\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}\tilde{\mathbf{B}} = \tilde{\tilde{\mathbf{D}}}_{(p2,\,p2)} \quad , \quad \tilde{\mathbf{B}}'\mathbf{V}_{22}\tilde{\mathbf{B}} = \mathbf{I}_{(p2,\,p2)}\,.$$

Ricordando adesso che $r(\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}) = r(\mathbf{V}_{21}) = r$, ne consegue che il numero degli autovalori non nulli di $\tilde{T}_2$ è pari a r, mentre l'autovalore nullo si presenta con molteplicità pari a $p_2 - r$.

(g) Tenuto conto di quanto detto, l'equazione

$$\det\begin{bmatrix} -\lambda\mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & -\lambda\mathbf{V}_{22} \end{bmatrix} = 0$$

ammette r radici positive $\tilde{\lambda}_1, \dots, \tilde{\lambda}_r$ e r radici negative $-\tilde{\lambda}_1, \dots, -\tilde{\lambda}_r$, mentre la radice nulla si presenta con molteplicità pari a $p_1 + p_2 - 2r$.

Per esempio, siano

$$\mathbf{V}_{11} = \begin{bmatrix} 2 \end{bmatrix} \quad , \quad \mathbf{V}_{22} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \quad , \quad \mathbf{V}_{12} = \begin{bmatrix} 0 & 1 \end{bmatrix} \quad , \quad \mathbf{V}_{21} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Allora, l'equazione

$$\begin{aligned} \det\begin{bmatrix} -\lambda\mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & -\lambda\mathbf{V}_{22} \end{bmatrix} &= \det\begin{bmatrix} -\lambda 2 & 0 & 1 \\ 0 & -\lambda 2 & -\lambda 1 \\ 1 & -\lambda 1 & -\lambda 1 \end{bmatrix} \\ = \lambda(-\lambda^2 + 1) \qquad &= 0 \end{aligned}$$

ammette le radici $\tilde{\lambda}_1 = 1, -\tilde{\lambda}_1 = -1, \tilde{\lambda}_2 = 0$.

A sua volta, l'equazione

$$\det\left(\lambda^2\mathbf{I} - \mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\right) = \det(-\lambda^2 + 1) = 0$$

ammette un unico autovalore $\tilde{\lambda}_1^2 = 1$.

Infine, l'equazione

$$\begin{aligned} \det\left(\lambda^2\mathbf{I} - \mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}\right) &= \det\left(\lambda^2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & -1/2 \\ 0 & 1 \end{bmatrix}\right) \\ = \lambda^2(\lambda^2 - 1) \qquad &= 0 \end{aligned}$$

ammette gli autovalori $\tilde{\lambda}_1^2 = 1, \tilde{\lambda}_2^2 = 0$.

(h) Siano $\tilde{\lambda}_1^2, \dots, \tilde{\lambda}_r^2$ gli r autovalori non nulli delle trasformazioni autoaggiunte $\tilde{T}_1$ e $\tilde{T}_2$ rappresentate, rispettivamente, dalle matrici $\mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}$ e $\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}$.

Indichiamo con $\tilde{\mathbf{A}}_1$ la matrice di ordine $(p_1, r)$ i cui vettori colonna sono autovettori (ortonormali) di $\tilde{T}_1$ associati a $\tilde{\lambda}_1^2, \dots, \tilde{\lambda}_r^2$.

Analogamente, indichiamo con $\tilde{\mathbf{B}}_1$ la matrice di ordine $(p_2, r)$ i cui vettori colonna sono autovettori (ortonormali) di $\tilde{T}_2$ associati a $\tilde{\lambda}_1^2, \dots, \tilde{\lambda}_r^2$.

Dalla relazione ($\tilde{\mathbf{D}}_{(r,r)} = \mathrm{diag}(\tilde{\lambda}_1^2, \dots, \tilde{\lambda}_r^2)$)

$$\mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\tilde{\mathbf{A}}_1 = \tilde{\mathbf{A}}_1\tilde{\mathbf{D}}_{(r,r)},$$

premoltiplicando ambo i membri per $\mathbf{V}_{22}^{-1}\mathbf{V}_{21}$ e postmoltiplicando per $\tilde{\mathbf{D}}_{(r,r)}^{-1/2}$, si ottiene

$$\begin{aligned}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}(\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\tilde{\mathbf{A}}_1\tilde{\mathbf{D}}_{(r,r)}^{-1/2}) &= \mathbf{V}_{22}^{-1}\mathbf{V}_{21}\tilde{\mathbf{A}}_1\tilde{\mathbf{D}}_{(r,r)}\tilde{\mathbf{D}}_{(r,r)}^{-1/2} \\ &= (\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\tilde{\mathbf{A}}_1\tilde{\mathbf{D}}_{(r,r)}^{-1/2})\,\tilde{\mathbf{D}}_{(r,r)}.\end{aligned}$$

Pertanto,

$$\tilde{\mathbf{B}}_1 = \mathbf{V}_{22}^{-1}\mathbf{V}_{21}\tilde{\mathbf{A}}_1\tilde{\mathbf{D}}_{(r,r)}^{-1/2}.$$

Con una dimostrazione del tutto simile, si prova poi che

$$\tilde{\mathbf{A}}_1 = \mathbf{V}_{11}^{-1}\mathbf{V}_{12}\tilde{\mathbf{B}}_1\tilde{\mathbf{D}}_{(r,r)}^{-1/2}.$$

Indichiamo adesso con $\tilde{\mathbf{A}}_2$ la matrice di ordine $(p_1, p_1-r)$ i cui vettori colonna sono autovettori (ortonormali) di $\tilde{T}_1$ associati all'autovalore nullo di molteplicità pari a $p_1-r$.

Analogamente, indichiamo con $\tilde{\mathbf{B}}_2$ la matrice di ordine $(p_2, p_2-r)$ i cui vettori colonna sono autovettori (ortonormali) di $\tilde{T}_2$ associati all'autovalore nullo di molteplicità pari a $p_2-r$.

Si riconosce subito, in vista di quanto detto al punto 6 che precede, che ($\mathbf{O}$ : matrice nulla di ordine appropriato)

$$\begin{aligned}&\{\mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\tilde{\mathbf{A}}_2 = \mathbf{O}\} \Leftrightarrow \{\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\tilde{\mathbf{A}}_2 = \mathbf{O}\} \\ &\Leftrightarrow \{\mathbf{V}_{21}\tilde{\mathbf{A}}_2 = \mathbf{O}\}\end{aligned}$$

e

$$\{\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}\widetilde{\mathbf{B}}_2 = \mathbf{O}\} \Leftrightarrow \{\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}\widetilde{\mathbf{B}}_2 = \mathbf{O}\}$$
$$\Leftrightarrow \{\mathbf{V}_{12}\widetilde{\mathbf{B}}_2 = \mathbf{O}\}$$

ovvero che i vettori colonna di $\widetilde{\mathbf{A}}_2$ e $\widetilde{\mathbf{B}}_2$ costituiscono basi (ortonormali) del nucleo di ciascuna delle matrici qui sopra indicate.

(i) Si consideri nuovamente l'equazione

$$\det\begin{bmatrix} -\lambda\mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & -\lambda\mathbf{V}_{22} \end{bmatrix} = 0 \ .$$

Come abbiamo visto, questa ammette r radici positive $\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_r$ e r radici negative $-\widetilde{\lambda}_1, \ldots, -\widetilde{\lambda}_r$, mentre la radice nulla si presenta con molteplicità pari a $p_1 + p_2 - 2r$.

Indicata con $\widetilde{\lambda}$ una generica radice positiva o nulla dell'equazione precedente, vogliamo occuparci delle soluzioni del sistema

$$(i1) \qquad \begin{bmatrix} -\widetilde{\lambda}\mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & -\widetilde{\lambda}\mathbf{V}_{22} \end{bmatrix}\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \mathbf{0}$$

nelle incognite $\mathbf{a}, \mathbf{b}$, sotto i vincoli

$$(i2) \qquad \mathbf{a}'\mathbf{V}_{11}\mathbf{a} = 1 \quad , \quad \mathbf{b}'\mathbf{V}_{22}\mathbf{b} = 1 \ .$$

A questo fine, qualora sia $\widetilde{\lambda} > 0$, premoltiplicando ambo i membri del sistema ($i1$) per la matrice (invertibile)

$$\begin{bmatrix} -\widetilde{\lambda}\mathbf{I} & \mathbf{V}_{12}\mathbf{V}_{22}^{-1} \\ \mathbf{O} & -\widetilde{\lambda}^{-1}\mathbf{V}_{22}^{-1} \end{bmatrix}$$

si ottengono i sottosistemi

$$(i3) \qquad \left(\widetilde{\lambda}^2\mathbf{I} - \mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\right)\mathbf{a} = \mathbf{0} \quad , \quad \mathbf{b} = \widetilde{\lambda}^{-1}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{a}$$

che, complessivamente, costituiscono un sistema del tutto equivalente al sistema originario (*i1*) per $\tilde{\lambda} > 0$.

Pertanto, soluzioni del sistema (*i1*) per $\tilde{\lambda} > 0$, soddisfacenti i vincoli espressi in (*i2*), sono date

- dagli autovettori $\tilde{\mathbf{a}}_1, \ldots, \tilde{\mathbf{a}}_r$ della trasformazione autoaggiunta $\tilde{T}_1$ rappresentata dalla matrice $\mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}$ associati agli r autovalori non nulli $\tilde{\lambda}_1^2, \ldots, \tilde{\lambda}_r^2$;
- dai vettori (ortonormali) $\tilde{\mathbf{b}}_1 = \tilde{\lambda}_1^{-1}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\tilde{\mathbf{a}}_1, \ldots, \tilde{\mathbf{b}}_r = \tilde{\lambda}_r^{-1}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\tilde{\mathbf{a}}_r$.

Qualora poi sia $\tilde{\lambda} = 0$, sempre dal sistema (*i1*), si ricavano i sottosistemi

(*i4*) $$\mathbf{V}_{21}\mathbf{a} = \mathbf{0} \quad \text{o} \quad (\mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21})\,\mathbf{a} = \mathbf{0}$$

e

(*i5*) $$\mathbf{V}_{12}\mathbf{b} = \mathbf{0} \quad \text{o} \quad (\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12})\,\mathbf{b} = \mathbf{0}\,.$$

Ne consegue che soluzioni del sistema (*i1*) per $\tilde{\lambda} = 0$, soddisfacenti i vincoli espressi in (*i2*), sono date

- dagli autovettori $\tilde{\mathbf{a}}_{r+1}, \ldots, \tilde{\mathbf{a}}_{p_1}$ della trasformazione autoaggiunta $\tilde{T}_1$ rappresentata dalla matrice $\mathbf{V}_{11}^{-1}\mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21}$ associati all'autovalore nullo di molteplicità pari a $p_1 - r$;
- dagli autovettori $\tilde{\mathbf{b}}_{r+1}, \ldots, \tilde{\mathbf{b}}_{p_2}$ della trasformazione autoaggiunta $\tilde{T}_2$ rappresentata dalla matrice $\mathbf{V}_{22}^{-1}\mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}$ associati all'autovalore nullo di molteplicità pari a $p_2 - r$.

**8**

Supponiamo $R^n$ munito della metrica rappresentata, rispetto alla base naturale di $R^n$, dalla matrice $\mathbf{M}$ di ordine (n,n) e siano $\tilde{\mu}_1, \ldots, \tilde{\mu}_n$ e $\tilde{\mathbf{a}}_1, \ldots, \tilde{\mathbf{a}}_n$, rispettivamente, gli n autovalori di $\mathbf{M}$ e n autovettori associati a tali autovalori.

Analogamente, supponiamo $R^p$ munito della metrica rappresentata, rispetto alla base naturale di $R^p$, dalla matrice $\mathbf{Q}$ di ordine $(p,p)$ e siano $\tilde{\lambda}_1, \dots, \tilde{\lambda}_p$ e $\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_p$, rispettivamente, i p autovalori di $\mathbf{Q}$ e p autovettori associati a tali autovalori.

Vogliamo mostrare che anche $\mathbf{Q} \otimes \mathbf{M}$ è definita positiva.

Infatti, poiché $(i = 1, \dots, n; j = 1, \dots, p)$

$$\mathbf{M}\tilde{\mathbf{a}}_i = \tilde{\mu}_i \tilde{\mathbf{a}}_i \quad , \quad \mathbf{Q}\tilde{\mathbf{b}}_j = \tilde{\lambda}_j \tilde{\mathbf{b}}_j \, ,$$

risulta (Cfr. la proprietà (3) di cui al punto 4 dei Complementi al Cap. II)

$$(\mathbf{Q}\tilde{\mathbf{b}}_j) \otimes (\mathbf{M}\tilde{\mathbf{a}}_i) = (\mathbf{Q} \otimes \mathbf{M})(\tilde{\mathbf{b}}_j \otimes \tilde{\mathbf{a}}_i) = \tilde{\mu}_i \tilde{\lambda}_j (\tilde{\mathbf{b}}_j \otimes \tilde{\mathbf{a}}_i) \, .$$

Quindi, $\tilde{\mu}_i \tilde{\lambda}_j$ è un autovalore di $\mathbf{Q} \otimes \mathbf{M}$ e $\tilde{\mathbf{b}}_j \otimes \tilde{\mathbf{a}}_i$ è un autovettore a esso associato.

Essendo poi $\tilde{\mu}_i$ e $\tilde{\lambda}_j$ positivi anche il prodotto $\tilde{\mu}_i \tilde{\lambda}_j$ è positivo e ciò prova quanto asserito.

**9**

Il concetto di spazio euclideo e quelli a esso collegati, dati con riferimento a uno spazio vettoriale di ordine n e dimensione p, sono chiaramente applicabili a un qualsiasi spazio vettoriale sull'insieme dei numeri reali.

Sia, per esempio, lo spazio vettoriale S su R costituito dall'insieme di tutte le matrici di ordine $(n,p)$.

Si consideri la funzione G di $S \times S$ in R definita ponendo, per ogni $(\mathbf{X}, \mathbf{Y}) \in S \times S$,

$$G(\mathbf{X}, \mathbf{Y}) = \mathrm{tr}(\mathbf{X}'\mathbf{M}\mathbf{Y}\mathbf{Q})$$

dove $\mathbf{M}$ e $\mathbf{Q}$ sono matrici definite positive di ordine, rispettivamente, $(n,n)$ e $(p,p)$.

Vogliamo dimostrare che tale funzione è lineare rispetto a entrambi gli argomenti della funzione stessa, simmetrica, e che, per ogni $\mathbf{X}$ diversa dalla matrice nulla, $G(\mathbf{X}, \mathbf{Y})$ è maggiore di zero.

In altri termini, vogliamo dimostrare che la funzione G qui sopra indicata può essere assunta quale metrica su S.

Le proprietà di linearità e simmetria sono subito verificate tenendo presenti le proprietà elementari della traccia di una matrice.

Per verificare l'ultima delle proprietà sopra dette, osserviamo anzitutto (Cfr. il punto 5 dei Complementi al Cap. II) che possiamo scrivere

$$\begin{aligned} \mathrm{G}(\mathbf{X},\mathbf{Y}) &= \mathrm{tr}(\mathbf{X}'\mathbf{M}\mathbf{Y}\mathbf{Q}) &&= (\mathrm{vec}\mathbf{X})'(\mathrm{vec}\mathbf{M}\mathbf{X}\mathbf{Q}) \\ &= (\mathrm{vec}\mathbf{X})'(\mathbf{Q}\otimes\mathbf{M})(\mathrm{vec}\mathbf{X}) &&= \mathbf{x}'(\mathbf{Q}\otimes\mathbf{M})\mathbf{x} \end{aligned}$$

dove $\mathbf{x} = \mathrm{vec}\mathbf{X}$ appartiene a $\mathrm{R}^{\mathrm{np}}$ e $\mathbf{Q}\otimes\mathbf{M}$, di ordine $(\mathrm{np},\mathrm{np})$, è definita positiva (Cfr. il punto 8 che precede).

Quindi − interpretando $\mathbf{Q}\otimes\mathbf{M}$ come matrice della metrica di $\mathrm{R}^{\mathrm{np}}$ rispetto alla base naturale di $\mathrm{R}^{\mathrm{np}}$ medesimo − risulta che, per ogni $\mathbf{x}$ diverso dal vettore nullo, $\mathbf{x}'(\mathbf{Q}\otimes\mathbf{M})\mathbf{x}$ è maggiore di zero e ciò equivale, appunto, a dire che, per ogni $\mathbf{X}$ diversa dalla matrice nulla, $\mathrm{G}(\mathbf{X},\mathbf{Y})$ è maggiore di zero.

Da quanto precede segue che la funzione N definita ponendo, per ogni $\mathbf{Y}$ di ordine $(\mathrm{n},\mathrm{p})$,

$$\mathrm{N}(\mathbf{Y}) = \sqrt{\mathrm{tr}(\mathbf{Y}'\mathbf{M}\mathbf{Y}\mathbf{Q})} = \sqrt{\mathrm{tr}(\mathbf{Y}\mathbf{Q}\mathbf{Y}'\mathbf{M})}$$

è una norma sullo spazio vettoriale delle matrici di ordine $(\mathrm{n},\mathrm{p})$.

Tale norma è nota come *norma di Frobenius*.

Si osservi che, indicando con $\widetilde{\lambda}_1, \dots, \widetilde{\lambda}_r$ gli autovalori positivi di $\mathbf{Y}'\mathbf{M}\mathbf{Y}\mathbf{Q}$ o $\mathbf{Y}\mathbf{Q}\mathbf{Y}'\mathbf{M}$ (Cfr. il punto 3 che precede), risulta

$$\mathrm{N}(\mathbf{Y}) = \widetilde{\lambda}_1 + \dots + \widetilde{\lambda}_r \,.$$

# BIBLIOGRAFIA


AA. VV. — *Strutture algebriche e strutture topologiche*
Feltrinelli, Milano, 1963.

BELLMAN R. E. — *Introduction to Matrix Analysis*
McGraw-Hill, New York, 1970.

DIEUDONNÉ J. — *Algebra lineare e geometria elementare*
Feltrinelli, Milano, 1970.

FANO G. — *Metodi matematici della meccanica quantistica*
Zanichelli, Bologna, 1967.

FÜRST D. — *Introduzione all'algebra lineare*
Cedam, Padova, 1975.

FÜRST D. — *Spazi lineari, metrica, calcolo matriciale*
Cedam, Padova, 1982.

GANTMACHER F. R. — *The Theory of Matrices*
Chelsea, New York, 1959.

GEL' FAND I. M. — *Lectures on Linear Algebra*
Interscience, New York, 1961.

GRAHAM A. — *Kronecker Products and Matrix Calculus: with Applications*
Ellis Horwood Limited, Chichester, 1981.

GREUB W. — *Linear Algebra*
Springer, New York, 1975.

HADLEY G. — *Linear Algebra*
Addison-Wesley, Londra, 1961.

HALMOS P. R. — *Finite-Dimensional Vector Spaces*
D. Van Nostrand, Princeton, 1958.

HOHN F. E. — *Elementary Matrix Algebra*
Collier-McMillan, New York, 1967.

LANCASTER P. — *The Theory of Matrices*
Academic Press, Orlando, 1985.

LANG S. — *Algebra Lineare*
Boringhieri, Torino, 1972.

MEYER C. D. — *Matrix Analysis and Applied Linear Algebra*
Siam, 2000.

MIRSKY L. — *An Introduction to Matrix Algebra*
Oxford University Press, Oxford, 1961.

NERING E.D. — *Linear Algebra and Matrix Theory*
Wiley&Sons, New York, 1970.

PEASE M.C. — *Methods of Matrix Algebra*
Academic Press, New York, 1965.

SEARLE S. R. — *Matrix Algebra for the Biological Sciences*
Wiley&Sons, New York, 1966.

SERNESI E. — *Geometria 1*
Bollati Boringhieri, Torino, 1989.

SHEPARD G. C. — *Spazi vettoriali di dimensioni finite*
Cremonese, Roma, 1969.

SHILOV G. E. — *An Introduction to the Theory of Linear Spaces*
Prentice-Hall, Englewood Cliffs (N. J.), 1961.

STOLL R. R. - WONG E. T. — *Linear Algebra*
Academic Press, New York, 1968.

# INDICE ANALITICO