2. Algebra delle matrici

2.1 Introduzione

Nel capitolo III si vedrà che una variabile casuale multi-dimensionale è semplicemente un insieme ordinato

$$[X_1, \ldots, X_p]$$

di variabili casuali singole. Con il termine « ordinato » si intende che la variabile casuale che descrive l'i-esimo aspetto di ciascuna unità campionaria estratta dalla popolazione, appare sempre nella i-esima posizione della sequenza nello schema. Il numero di p variabili casuali è sempre specificato e rimarrà invariato durante lo svolgimento del problema o dell'analisi in corso.

Per esempio si supponga che le componenti della variabile siano i pesi di femmine di ratto appartenenti ad una particolare razza, misurati alla nascita e a 10, 20 e 30 giorni di età.

In tal caso si potrebbero descrivere i pesi mediante la variabile casuale $[X_1, X_2, X_3, X_4]$, con determinata distribuzione nello spazio quadri-dimensionale. Se tali pesi fossero rilevati in un campione di N ratti, le osservazioni potrebbero essere riassunte nello schema

$$\begin{bmatrix} x_{11} & \cdots & x_{14} \\ \cdots & \cdots & \cdots \\ x_{N1} & \cdots & x_{N4} \end{bmatrix}$$

le cui righe corrispondono a ratti diversi.

Tali disposizioni lineari e rettangolari di numeri sono note rispettivamente come *vettori* e *matrici* e le regole per il loro uso costituiscono quella parte dei sistemi lineari nota come *algebra delle matrici*.

Tale algebra è il linguaggio virtuale dell'analisi multivariata, in particolare nella sua parte più usata e più importante, basata sulla distribuzione normale multi-dimensionale. In effetti, è piuttosto inconcepibile che si siano potute sviluppare tecniche senza la convenienza, la facilità e l'eleganza delle matrici.

In questo capitolo riassumeremo alcune delle proprietà, operazioni, e teoremi dell'algebra delle matrici, necessarie per la trattazione successiva.

Si possono trovare poi altri risultati e dimostrazioni di alcuni teoremi nei riferimenti bibliografici citati alla fine del capitolo.

2.2 Definizioni

Si supponga di avere a disposizione alcuni elementi, il cui comportamento sia in accordo con alcuni « set » di assiomi come per esempio i numeri reali o i numeri complessi.

Si definisce pertanto come « *matrice* », lo schema rettangolare ordinato di elementi:

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1c} \\ \cdots & \cdots & \cdots \\ a_{r1} & \cdots & a_{rc} \end{bmatrix} \tag{1}$$

Il termine generale della matrice verrà scritto come a_{ij}, dove il primo indice sarà sempre riferito alla i-esima riga, ed il secondo alla j-esima colonna. Le dimensioni di una matrice sono importanti e si indicherà con $r \times c$ una matrice a r righe e c colonne.

In contrapposizione al termine matrice chiameremo *scalari* i numeri e le variabili unidimensionali di uso comune. Una quantità scalare è naturalmente una matrice 1×1. Nei paragrafi iniziali di questo capitolo considereremo le matrici come insiemi costituiti da elementi scalari reali. In seguito tratteremo matrici i cui elementi sono essi stessi matrici a dimensioni inferiori. Un *vettore* è una matrice ad una sola riga o colonna.

Scriveremo abitualmente il *vettore colonna* ad n-componenti:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \tag{2}$$

in grassetto. Analogamente il *vettore-riga*

$$\mathbf{x}' = [x_1, \ldots, x_n] \tag{3}$$

consiste in una sola riga di n elementi.

Ogni vettore specifica le coordinate di un punto nello spazio euclideo ad n-dimensioni ed il legame con il concetto fisico o analitico di un vettore è immediatamente evidente non appena si consideri tale punto come il secondo estremo del segmento di retta avente origine nell'origine degli assi coordinati. L'apice del vettore-riga significa che $\mathbf{x}'$ è il *transposto* del vettore-colonna $\mathbf{x}$.

In generale se $\mathbf{A}$ è una qualsiasi matrice $r \times c$, la trasposta di $\mathbf{A}$ è la

matrice $\mathbf{A}'$ di dimensioni $c \times r$ ottenuta scambiando tra di loro le righe con le colonne.

$$\mathbf{A}' = \begin{bmatrix} a_{11} & \cdots & a_{1c} \\ \cdots & \cdots & \cdots \\ a_{r1} & \cdots & a_{rc} \end{bmatrix}' = \begin{bmatrix} a_{11} & \cdots & a_{r1} \\ \cdots & \cdots & \cdots \\ a_{1c} & \cdots & a_{rc} \end{bmatrix} \tag{4}$$

Se una matrice è quadrata ed uguale alla sua trasposta, essa è detta matrice *simmetrica.*

In tal caso $a_{ij} = a_{ji}$ per qualunque coppia degli indici i, j.

Per esempio la:

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

è simmetrica, mentre non lo è la:

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

Per brevità frequentemente verranno omessi gli elementi simmetrici inferiori. Gli elementi a_{ii} di una matrice quadrata, occupano quelle che sono chiamate le *posizioni diagonali principali.*

La somma di tali elementi diagonali è chiamata la *traccia* di $\mathbf{A}$, e sarà indicata mediante la notazione

$$\operatorname{tr} \mathbf{A} = \sum_{i=1}^{n} a_{ii} \tag{5}$$

Alcune matrici sono particolarmente importanti.

La matrice *identità*:

$$\mathbf{I} = \begin{bmatrix} 1 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & 1 \end{bmatrix} \tag{6}$$

è una matrice quadrata con 1 in ciascuna posizione della diagonale principale e zero in tutte le altre posizioni.

La *matrice diagonale* $p \times p$

$$\mathbf{D}(a_i) = \begin{bmatrix} a_1 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & a_p \end{bmatrix} \tag{7}$$

ha gli elementi $a_1, \ldots, a_p$ nelle posizioni della sua diagonale principale e zero in tutte le altre posizioni. Alcuni degli elementi a_i possono essere

nulli. In seguito la notazione diag. $(\mathbf{A}) = \mathbf{D}(a_{ii})$ indicherà la matrice diagonale costruita dalla matrice quadrata **A**. La matrice *triangolare* $p \times p$ è del tipo

$$\begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1p} \\ 0 & t_{22} & \cdots & t_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & t_{pp} \end{bmatrix} \tag{8}$$

La *matrice nulla* di dimensioni $r \times c$

$$\mathbf{0} = \begin{bmatrix} 0 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & 0 \end{bmatrix} \tag{9}$$

ha lo zero in ciascuna delle sue posizioni.

Talvolta sarà necessario anche il vettore

$$\mathbf{j}' = [1, \ldots, 1] \tag{10}$$

e la matrice

$$\mathbf{E} = \begin{bmatrix} 1 & \cdots & 1 \\ \cdots & \cdots & \cdots \\ 1 & \cdots & 1 \end{bmatrix} \tag{11}$$

con l'unità in ciascuna posizione.

2.3 Operazioni elementari con matrici e vettori

Le operazioni di addizione, sottrazione e moltiplicazione, dell'aritmetica ordinaria scalare, possono essere estese alle matrici qualora si seguano determinate regole. La matrice quoziente è più complessa e sarà trattata in un paragrafo successivo.

Eguaglianza. Due matrici **A** e **B**, $r \times c$, si dicono uguali se e soltanto se

$$a_{ij} = b_{ij} \tag{1}$$

qualunque sia la coppia di indici i, j.

Addizione. La somma di due matrici aventi le stesse dimensioni è la matrice di elementi pari alla somma degli elementi corrispondenti. Se

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1c} \\ \cdots & \cdots & \cdots \\ a_{r1} & \cdots & a_{rc} \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} b_{11} & \cdots & b_{1c} \\ \cdots & \cdots & \cdots \\ b_{r1} & \cdots & b_{rc} \end{bmatrix}$$

allora

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1c} + b_{1c} \\ \cdots & \cdots & \cdots \\ a_{r1} + b_{r1} & \cdots & a_{rc} + b_{rc} \end{bmatrix} \tag{2}$$

Per sottrarre una matrice da un'altra di eguali dimensioni, si costruisce la matrice delle differenze dei singoli elementi. Per le operazioni definite valgono le seguenti proprietà:

$$\begin{aligned} \mathbf{A} + \mathbf{B} &= \mathbf{B} + \mathbf{A} \\ \mathbf{A} + (\mathbf{B} + \mathbf{C}) &= (\mathbf{A} + \mathbf{B}) + \mathbf{C} \\ \mathbf{A} - (\mathbf{B} - \mathbf{C}) &= \mathbf{A} - \mathbf{B} + \mathbf{C} \end{aligned} \tag{3}$$

Se invece le matrici non hanno le stesse dimensioni, non sono definibili le operazioni di somma o differenza.

Moltiplicazione per uno scalare. La matrice **A** viene moltiplicata per lo scalare c, qualora si moltiplichi ciascun elemento di **A** per c.

$$c\mathbf{A} = \begin{bmatrix} ca_{11} & \cdots & ca_{1k} \\ \cdots & \cdots & \cdots \\ ca_{r1} & \cdots & ca_{rk} \end{bmatrix} \tag{4}$$

Moltiplicazioni di Matrici. Per poter definire la matrice prodotto di **AB** è necessario che il numero delle *colonne* di **A** sia eguale al numero delle *righe* di **B**.

Le dimensioni di tali matrici si dicono *conformi*. Se **A** è di dimensioni $p \times r$ e **B** è di dimensioni $r \times q$, allora l'elemento di indici i, j del prodotto $\mathbf{C} = \mathbf{AB}$ è calcolato come

$$c_{ij} = \sum_{k=1}^{r} a_{ik} b_{kj} \tag{5}$$

Tale generico elemento è la somma dei prodotti degli elementi dell'i-esima riga di **A** con i corrispondenti della j-esima colonna di **B**. Le dimensioni di **AB** sono evidentemente $p \times q$. Per esempio se

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 6 & 5 & 4 \\ -1 & 1 & -1 \\ 0 & 2 & 0 \end{bmatrix}$$

allora

$$\mathbf{AB} = \begin{bmatrix} 1(6) + 2(-1) + 3(0) & 1(5) + 2(1) + 3(2) & 1(4) + 2(-1) + 3(0) \\ -1(6) + 0(-1) + 1(0) & -1(5) + 0(1) + 1(2) & -1(4) + 0(-1) + 1(0) \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 13 & 2 \\ -3 & -3 & -4 \end{bmatrix}$$

Il prodotto **BA** non può essere definibile se le due righe di **A** non sono conformi alle 3 colonne di **B**.

Per il prodotto di matrici valgono le proprietà distributiva ed associativa:

$$\begin{aligned} \mathbf{A}(\mathbf{B} + \mathbf{C}) &= \mathbf{AB} + \mathbf{AC} \\ \mathbf{A}(\mathbf{BC}) &= \mathbf{AB}(\mathbf{C}) \end{aligned} \tag{6}$$

Per la moltiplicazione non vale la proprietà commutativa ed in generale non è vero che $\mathbf{AB} = \mathbf{BA}$. Per questo motivo è fondamentale l'ordine di moltiplicazione; e noi parleremo del prodotto **AB** come di una *moltiplicazione a sinistra* di **B** con **A** o di una *moltiplicazione a destra* di **A** con **B**.

Per esempio sia

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 3 & -1 \\ -1 & 1 \end{bmatrix}$$

allora

$$\mathbf{AB} = \begin{bmatrix} 1 & 1 \\ 3 & 1 \end{bmatrix} \qquad \mathbf{BA} = \begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix}$$

I due prodotti non sono affatto eguali nè viene conservata nella moltiplicazione la simmertria delle matrici originali. Il prodotto di una qualsiasi matrice con la matrice identità di egual dimensioni lascia la matrice stessa invariata. La moltiplicazione a sinistra con la matrice diagonale di elementi $d_1, \ldots, d_r$ ha l'effetto di moltiplicare ciascun elemento dell'i-esima riga per d_i:

$$\mathbf{D}(d_i)\mathbf{A} = \begin{bmatrix} d_1 a_{11} & \cdots & d_1 a_{1c} \\ \cdots & \cdots & \cdots \\ d_r a_{r1} & \cdots & d_r a_{rc} \end{bmatrix} \tag{7}$$

La moltiplicazione a destra con una siffatta matrice diagonale $c \times c$ moltiplica ciascun elemento della j-esima colonna per d_j. Una matrice può essere considerata come operante una trasformazione lineare di vettori dallo spazio di questi ad un altro.

Se **x** ha m componenti ed **y** ha n componenti, è possibile ottenere una trasformazione dal sistema di coordinate ad m dimensioni degli elementi di **x** nello spazio ad n-dimensioni di quelli di **y**, mediante la forma matriciale del tipo

$$\begin{bmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{bmatrix} \tag{8}$$

$$= \mathbf{Ax}$$

Una trasformazione ad un terzo « set » di variabili rappresentato dal vettore-colonna **z** a p-componenti, potrebbe essere rappresentata come

$$\begin{aligned} \mathbf{z} &= \mathbf{By} \\ &= \mathbf{BAx} \end{aligned} \tag{9}$$

e così il prodotto di due o più matrici può essere considerato la matrice risultante da una successione di trasformazioni lineari.

Esempio 2.1 - È spesso necessario nelle scienze sociali ridurre parecchi e disparati punteggi, raccolti su un certo numero di individui ad una scala con origine ed unità di misura comuni.

Se x_{ij} è il punteggio del i-esimo individuo per la j-esima misura e $\bar{x}_j$ e s_j sono la media e la deviazione standard del campione per quella misura, una trasformazione comune è il *punteggio z*.

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}$$

Le osservazioni trasformate possono essere convenientemente calcolate con alcune semplici operazioni matriciali.

Si scrivano i punteggi originali come *matrici-dati* di dimensione $N \times p$

$$\mathbf{X} = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \cdots & \cdots & \cdots \\ x_{N1} & \cdots & x_{Np} \end{bmatrix}$$

e si formi la matrice diagonale

$$\mathbf{D} = \begin{bmatrix} \frac{1}{s_1} & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & \frac{1}{s_p} \end{bmatrix}$$

delle deviazioni standard del campione.

Se introduciamo la matrice di dimensione $N \times N$

$$\mathbf{E} = \begin{bmatrix} 1 & \cdots & 1 \\ \cdots & \cdots & \cdots \\ 1 & \cdots & 1 \end{bmatrix}$$

con 1 in ogni posizione, la matrice **Z** di dimensioni $N \times p$ dei punteggi standardizzati può essere calcolata come

$$\mathbf{Z} = \left(\mathbf{I} - \frac{1}{N}\mathbf{E}\right)\mathbf{XD}$$

La moltiplicazione a destra di **E** per **X** ha l'effetto di sommare ogni colonna della matrice dei dati.

Prodotto interno dei vettori

Il prodotto interno di due vettori con lo stesso numero di elementi si definisce come la somma dei prodotti degli elementi corrispondenti.

$$\mathbf{x}'\mathbf{y} = [x_1, \ldots, x_p] \begin{bmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_p \end{bmatrix} \tag{10}$$

$$= \sum_{i=1}^{p} x_i y_i$$

Poichè il prodotto interno è uno scalare, $\mathbf{x}'\mathbf{y} = \mathbf{y}'\mathbf{x}$. Il prodotto interno di $\mathbf{x}$ con se stesso è la somma dei quadrati degli elementi di $\mathbf{x}$.

I prodotti interni hanno importanti interpretazioni geometriche. Il prodotto interno di $\mathbf{x}$ con se stesso è chiamato *lunghezza al quadrato* di $\mathbf{x}$, poichè è il quadrato della distanza dall'origine del sistema di coordinate p-dimensionali dal punto specificato dagli elementi di $\mathbf{x}$.

Più generalmente la distanza fra i punti con coordinate date da $\mathbf{x}$ ed $\mathbf{y}$ è

$$d = \sum_{i=1}^{p} (x_i - y_i)^2]^{1/2} \tag{11}$$

Il coseno dell'angolo 0 fra i vettori $\mathbf{x}$ ed $\mathbf{y}$ è

$$\cos\ 0 = \frac{\mathbf{x}'\mathbf{y}}{(\mathbf{x}'\mathbf{x})^{1/2}\,(\mathbf{y}'\mathbf{y})^{1/2}} \tag{12}$$

La divisione dei vettori per le loro rispettive lunghezze viene chiamata *normalizzazione*: si può facilmente notare che, via via che i vettori normalizzati tendono a sovrapporsi, il loro prodotto interno tende all'unità.

In modo analogo si vede che i vettori ad angolo retto l'uno rispetto all'altro, hanno prodotto interno eguale a zero.

Esempio 2.2 - Nello spazio tridimensionale i vettori

$$\mathbf{x}' = [1,0,0]$$
$$\mathbf{y}' = [0,1,0]$$
$$\mathbf{z}' = [0,0,1]$$

possono essere assunti come specificanti i 3 assi coordinati.

Il vettore

$$\mathbf{e}' = [1,1,1]$$

determina un angolo avente coseno comune $\sqrt{3}/3$ rispetto a $\mathbf{x}'$, $\mathbf{y}'$ e $\mathbf{z}'$ e caratterizza

la retta *equiangolata* nello spazio tridimensionale. I numeri $\sqrt{3}/3$ sono chiamati *coseni direttori* della retta. In modo analogo i vettori

$$\mathbf{u}' = [1,1,0]$$
$$\mathbf{v}' = [1, -1, 0]$$

creano un angolo di 90° nel piano xy dello spazio. $\mathbf{u}'$ e $\mathbf{v}'$ hanno ciascuno un angolo di quarantacinque gradi con $\mathbf{x}'$, ed i loro angoli con $\mathbf{y}'$ sono rispettivamente di 45° e —45°.

Trasposta di una matrice prodotto

Se la matrice **A** è conforme, per la moltiplicazione a destra, con un'altra matrice **B**, si può facilmente verificare che la transposta del loro prodotto è uguale al prodotto delle loro trasposte prese nell'ordine opposto.

$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}' \tag{13}$$

In generale se $\mathbf{A}_1, \ldots, \mathbf{A}_k$ sono conformi:

$$(\mathbf{A}_1 \cdot \cdot \cdot \mathbf{A}_k)' = \mathbf{A}'_k \cdot \cdot \cdot \mathbf{A}'_1 \tag{14}$$

Ci serviremo spesso delle proprietà ora citate nei capitoli seguenti.

2.4 Determinante di una matrice quadrata

Associata ad ogni matrice quadrata, vi è un'unica quantità scalare, chiamata suo *determinante*.

La definizione formale del determinante di una matrice **A** di dimensioni $n \times n$ è la somma

$$\sum \cdot \, {}_{P} \, \cdot \sum (-1)^{\alpha} a_{1j_1} a_{2j_2} \cdot \cdot \cdot a_{nj_n} \tag{1}$$

di tutti i prodotti che si ottengono prendendo come fattori uno ed uno solo elemento da ogni riga e da ogni colonna, e moltiplicandoli per —1 se il numero di spostamenti nella permutazione degli indici $j_1, j_2 \ldots j_n$ rispetto all'ordine standard 1, 2, . . ., n è dispari.

La somma viene calcolata dall'insieme P di tutte le $n!$ permutazioni degli indici di colonna. Il numero delle inversioni α in una particolare permutazione coincide con il numero totale di volte in cui un elemento è seguito da numeri che normalmente nell'ordine naturale 1, 2 . . . n. lo precederebbero.

Il determinante della matrice **A** è scritto come $|\mathbf{A}|$. I determinanti delle tre matrici quadrate di minori dimensioni sono per la definizione del tipo:

$$
\begin{aligned}
|a_{11}| &= a_{11} \\
\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} &= a_{11}a_{22} - a_{12}a_{21} \\
\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= \begin{aligned}[t] & a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ & - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \end{aligned}
\end{aligned}
\tag{2}
$$

È più conveniente calcolare i determinanti di matrici a maggiori dimensioni, con altri metodi.

Si definisce come « *minore* » dell'elemento a_{ij} della matrice **A** il determinante della matrice che si ottiene cancellando la i-esima riga e la j-esima colonna di **A**.

Il *complemento algebrico* di a_{ij} è il minore moltiplicato per $(-1)^{i+j}$ e si scriverà come A_{ij}. Si può dimostrare che il determinante di una matrice quadrata A può essere espressa in termini dei complementi algebrici degli elementi di una determinata riga o colonna nel modo seguente:

$$
\begin{aligned}
|\mathbf{A}| &= a_{i1}A_{i1} + \cdots + a_{in}A_{in} \qquad i = 1, \ldots, n \\
&= a_{1j}A_{1j} + \cdots + a_{nj}A_{nj} \qquad j = 1, \ldots, n
\end{aligned}
\tag{3}
$$

In una semplice applicazione di questa regola, si può notare che la matrice dei complementi algebrici di una tipica matrice 2×2 è

$$
\begin{bmatrix} a_{22} & -a_{21} \\ -a_{12} & a_{11} \end{bmatrix}
\tag{4}
$$

ed il valore del suo determinante può essere calcolato immediatamente.

Esempio 2.3 - Si voglia calcolare il determinante della matrice

$$
\begin{bmatrix} 3 & 0 & 0 & 0 \\ 1 & 2 & 0 & 1 \\ -1 & -3 & 2 & -1 \\ 5 & 4 & 3 & 2 \end{bmatrix}
$$

Si vede immediatamente che la prima riga contiene un solo elemento diverso da zero, così che sarà necessario calcolare solo il complemento algebrico dell'elemento (1.1). Se noi sviluppiamo tale complemento algebrico secondo i complementi algebrici della sua prima riga, il determinante originale sarà eguale a:

$$
3\left(2\begin{vmatrix} 2 & -1 \\ 3 & 2 \end{vmatrix} - 0\begin{vmatrix} -3 & -1 \\ 4 & 2 \end{vmatrix} + 1\begin{vmatrix} -3 & 2 \\ 4 & 3 \end{vmatrix}\right) = -9
$$

Il metodo dei complementi algebrici è conveniente solo per piccole matrici, o per matrici che contengono molti elementi nulli. Considereremo una tecnica di maggior praticità nel paragrafo 2.6.

Le seguenti proprietà dei determinanti risulteranno molto utili nei prossimi capitoli:

1) Il determinante di una matrice diagonale è uguale al prodotto degli elementi diagonali. Analogamente il determinante di una matrice triangolare è dato dal prodotto dei suoi elementi diagonali.

2) Se gli elementi di una riga o colonna di una matrice **A** di dimensioni $n \times n$ sono moltiplicati tutti per uno stesso valore scalare c, il determinante della nuova matrice è eguale a $c|\mathbf{A}|$.

Se invece ogni elemento è moltiplicato per c, allora $|c\mathbf{A}| = c^n|\mathbf{A}|$.

3) Scambiando tra di loro due righe o due colonne di una matrice quadrata, il determinante cambia segno.

4) Discende direttamente dalla proprietà 3) che se due colonne o due righe sono uguali, il determinante è zero. In modo analogo, se righe o colonne di una matrice sono proporzionali, il determinante risulta eguale a zero.

5) Il determinante di una matrice non varia, aggiungendo il multiplo di una colonna ad un'altra colonna. Lo stesso risultato vale per le righe.

6) Se tutti gli elementi di una riga o di una colonna di una matrice quadrata, sono eguali a zero, il determinante è zero.

7) Se **A** e **B** sono due matrici ciascuna di dimensioni $n \times n$, il determinante della matrice prodotto **A B**, è uguale al prodotto dei singoli determinanti.

8) La somma dei prodotti degli elementi di una determinata riga di una matrice quadrata per i complementi algebrici di una riga *diversa* è uguale a zero.

Un identico risultato vale per le colonne.

2.5 La matrice inversa

Possiamo ora definire la matrice equivalente al quoziente tra scalari. Si definisce matrice *inversa* di una matrice quadrata **A** quella matrice $\mathbf{A}^{-1}$, i cui elementi soddisfano alle seguenti uguaglianze

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I} \tag{1}$$

La matrice $\mathbf{A}^{-1}$ può non esistere, proprio come è impossibile eseguire una divisione tra scalari con divisore zero.

In tal caso si definisce **A**, matrice *singolare*. Le matrici, per cui esistono le inverse, sono dette *non singolari*. Si possono infine calcolare, gli elementi della matrice $\mathbf{A}^{-1}$ in base a due risultati del paragrafo precedente. Si costruisca la matrice dei complementi algebrici

$$\mathbf{C} = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \cdots & \cdots & \cdots \\ A_{n1} & \cdots & A_{nn} \end{bmatrix} \tag{2}$$

chiamata *matrice aggiunta* di **A**.

Allora il prodotto interno della i-esima riga di $\mathbf{C}$ e della h-esima riga di $\mathbf{A}$ è eguale al determinante di $\mathbf{A}$, $|\mathbf{A}|$, se $i = h$ ed a zero se $i \neq h$. Si otterrà pertanto la matrice inversa desiderata prendendo, la transposta di $\mathbf{C}$ e dividendone ciascun elemento per $|\mathbf{A}|$,

$$\mathbf{A}^{-1} = \begin{bmatrix} \frac{1}{|\mathbf{A}|} A_{11} & \cdots & \frac{1}{|\mathbf{A}|} A_{n1} \\ \cdots & \cdots & \cdots \\ \frac{1}{|\mathbf{A}|} A_{1n} & \cdots & \frac{1}{|\mathbf{A}|} A_{nn} \end{bmatrix} \tag{3}$$

È evidente da questa definizione che la matrice inversa di $\mathbf{A}$ esiste se e soltanto se $|\mathbf{A}|$ è diverso da zero. Il calcolo di $\mathbf{A}^{-1}$ mediante i complementi algebrici, è inadeguato in molte applicazioni pratiche, per cui si prenderanno in considerazione altri metodi nel paragrafo 2.6.

Saranno frequentemente utilizzate le seguenti proprietà della matrice inversa.

1) L'inversa di una matrice simmetrica è anch'essa simmetrica.

2) L'inversa della matrice transposta di $\mathbf{A}$, è la matrice transposta di $\mathbf{A}^{-1}$.

3) La matrice inversa del prodotto di più matrici quadrate è uguale al prodotto delle inverse in ordine opposto:

$$(\mathbf{ABC})^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1} \tag{4}$$

4) Se c è uno scalare diverso da zero, allora $(c\mathbf{A})^{-1} = 1/c\mathbf{A}^{-1}$.

5) L'inversa di una matrice diagonale è una matrice ancora diagonale, avente per elementi i reciproci degli elementi originali.

Esempio 2.4 - L'inversa di:

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

è:

$$\mathbf{A}^{-1} = \begin{bmatrix} 1/2 & 0 & 0 & 0 \\ 0 & 1/4 & 0 & 0 \\ 0 & 0 & 1/2 & -1/2 \\ 0 & 0 & -1/4 & 3/4 \end{bmatrix}$$

2.6 Rango di una matrice

Due vettori a p-componenti si dicono *linearmente dipendenti* se gli elementi del primo vettore sono proporzionali a quelli del secondo. Per la definizione data, due vettori riga

$$\mathbf{x}' = [1, 0, -1] \qquad \mathbf{y}' = [4, 0, -4]$$

sono linearmente dipendenti, mentre:

$$\mathbf{u}' = [2, -1, 0, 7] \qquad \mathbf{v}' = [6,2,0,0]$$

sono linearmente indipendenti.

Un insieme di k vettori di eguali dimensioni è detto linearmente indipendente se non è possibile scrivere alcun vettore dello insieme dato come combinazione lineare dei rimanenti. Cioè, i vettori $\mathbf{x}_1, \ldots, \mathbf{x}_k$ costituiscono un insieme linearmente indipendente se non esiste un insieme di scalari $c_1, \ldots, c_k$ tale che per ciascuno dei vettori dati valga la relazione:

$$c_i\mathbf{x}_i = c_1\mathbf{x}_1 + \cdots + c_{i-1}\mathbf{x}_{i-1} + c_{i+1}\mathbf{x}_{i+1} + \cdots + c_k\mathbf{x}_k$$

I vettori

$$\mathbf{x}' = [1, -1, 2] \qquad \mathbf{y}' = [2, 0, -1] \qquad \mathbf{z}' = [0, -2, 5]$$

costituiscono un insieme linearmente dipendente, poichè $\mathbf{z}' = 2\mathbf{x}' - \mathbf{y}'$ mentre sono indipendenti i vettori unitari:

$$\mathbf{t}' = [1,0,0] \qquad \mathbf{u}' = [0,1,0] \qquad \mathbf{v}' = [0,0,1]$$

Esempio 2.5 - Supponiamo che un test di livello intellettivo sia costituito da sei sottogruppi. I primi tre misurano alcune abilità verbali, e la somma dei loro punteggi dà quello che è chiamato *punteggio verbale*. Gli ultimi tre riflettono determinate abilità motorie e spaziali, e il loro totale è chiamato *punteggio di preparazione*. Infine la somma dei sei tests è usata come una misura generale della abilità intellettiva.

Il test è stato provato su di un ampio numero di soggetti ed è stato proposto che i nove punteggi vengano usati con l'intento di correlare l'intelligenza ad altre misure ottenute sugli individui. D'altra parte se si considera come vettore ad N componenti ciascun campione di osservazioni relative a uno dei nove punteggi, si vede immediatamente che, solo sei dei nove vettori sono linearmente indipendenti, e che non è stata ricavata alcuna nuova informazione dai tre punteggi derivati.

Nei prossimi capitoli vedremo che l'inclusione di tali combinazioni lineari può precludere determinati tipi di analisi statistica.

Si definisce ora formalmente il grado di indipendenza lineare di un insieme di vettori, come il *rango* della matrice da essi costituita. Si supponga per il momento che il numero dei vettori k non superi la dimensione p dei vettori. Sotto tali ipotesi il rango di:

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}'_1 \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{x}'_k \end{bmatrix} \tag{1}$$

è il numero dei vettori riga linearmente indipendenti nella matrice. Il rango

può variare da zero, per qualsiasi matrice nulla, a k per ogni matrice a *pieno rango*.

Se, d'altra parte, il numero delle righe di **X** supera il numero delle colonne, il rango di **X** sarà il numero delle colonne linearmente indipendenti. In entrambi i casi si può dimostrare che il rango della matrice è unico, indipendentemente dal fatto che sia calcolato sulle righe o sulle colonne. Dalla proprietà n. 4 del paragrafo 2.4 segue che:

La matrice **A** *è di rango r se contiene al massimo un minore non nullo di dimensioni* $r \times r$, *e non contiene alcun minore diverso da zero di dimensioni maggiori di r.*

Esempio 2.6 - La matrice:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 6 & 9 & 12 \\ 4 & 3 & 2 & 1 \\ -1 & 3 & 7 & 11 \\ 8 & 6 & 4 & 2 \end{bmatrix}$$

ha rango due, perchè si può dimostrare che la seconda riga è multipla secondo il numero tre della prima, la quarta riga è eguale alla differenza tra la seconda e la terza e infine la quinta riga è il doppio della terza. Le sole righe linearmente indipendenti sono la prima e la terza; si dice che esse costituiscono la *base* per i 5 vettori riga.

Il rango gode delle seguenti importanti proprietà:

1) Il rango di $\mathbf{A}'$ è eguale a quello di **A**. Questa proprietà discende dall'equivalenza delle definizioni di rango date per riga e per colonna.

2) Il rango del prodotto $\mathbf{A}'\mathbf{A}$ è eguale a quello di **A**. Analogamente, il rango di $\mathbf{A}\mathbf{A}'$ è equivalente al rango di **A**.

3) Il rango di **A** rimane invariato per una moltiplicazione a destra o a sinistra di **A** con una matrice non singolare.

Operazioni elementari di riga e di colonna

Il rango di una matrice rimane invariato per le seguenti *operazioni elementari di riga e di colonna*:

1) Scambio di due righe (o colonne).

2) Prodotto di ciascun elemento di una riga (o colonna) per una quantità scalare costante.

3) Somma di una riga (colonna), i cui elementi siano stati moltiplicati per una quantità scalare costante, con un'altra riga (o colonna). Trasformazioni di questo tipo si possono pensare ottenute mediante matrici non singolari. Le operazioni di riga vengono realizzate mediante moltiplicazioni a

sinistra della matrice data con una operazione matriciale elementare, mentre le operazioni di colonna si ottengono da moltiplicazioni a destra.

Per la proprietà 3 del paragrafo precedente, tali trasformazioni non singolari, lasciano il rango invariato. Scegliendo opportunamente la successione delle operazioni di riga e di colonna, è possibile ridurre una matrice ad un'altra avente righe o colonne linearmente indipendenti, e quindi calcolarne il relativo rango.

Esempio 2.7 - Le matrici

$$\mathbf{E}_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{E}_2 = \begin{bmatrix} c & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{E}_3 = \begin{bmatrix} 1 & 0 & 0 \\ d & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

possono essere usate per eseguire trasformazioni elementari sulle prime due righe della matrice:

$$\mathbf{A} = \begin{bmatrix} 5 & 1 \\ -2 & 3 \\ 3 & 2 \end{bmatrix}$$

dando:

$$\mathbf{E}_1\mathbf{A} = \begin{bmatrix} -2 & 3 \\ 5 & 1 \\ 3 & 2 \end{bmatrix} \qquad \mathbf{E}_2\mathbf{A} = \begin{bmatrix} 5c & c \\ -2 & 3 \\ 3 & 2 \end{bmatrix} \qquad \mathbf{E}_3\mathbf{A} = \begin{bmatrix} 5 & 1 \\ 5d - 2 & d + 3 \\ 3 & 2 \end{bmatrix}$$

Esempio 2.8 - Si usino operazioni elementari di riga per calcolare il rango della matrice

$$\mathbf{X} = \begin{bmatrix} 1 & 0 & 2 \\ -3 & 2 & 1 \\ -1 & 2 & 5 \\ 3 & 0 & 6 \end{bmatrix}$$

Si vede immediatamente che l'ultima riga è uguale alla prima, moltiplicata per tre.

La terza riga è uguale alla seconda, più il doppio della prima. Opportune operazioni di riga riducono successivamente $\mathbf{X}$ a

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

Poichè le righe rimanenti sono linearmente indipendenti, noi concludiamo che $\mathbf{X}$ è di rango due.

Entrambi i tipi di operazioni di riga e di colonna possono essere appli-

cati contemporaneamente ad ogni matrice di dimensioni $m \times n$ per ridurla alla sua *forma canonica*

$$\begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} \tag{2}$$

avente il valore uno nelle prime r-posizioni diagonali e zero in tutte le altre posizioni. Questo significa che per ogni matrice $\mathbf{A}$, esistono matrici di trasformazione elementari di riga o di colonna $\mathbf{R}_1, \ldots, \mathbf{R}_p, \mathbf{C}_1 \ldots, \mathbf{C}_q$ tali che

$$\mathbf{R}_1 \cdots \mathbf{R}_p \mathbf{A} \mathbf{C}_1 \cdots \mathbf{C}_q \tag{3}$$

sia la forma canonica della matrice $\mathbf{A}$.

Ma se $\mathbf{A}$ è quadrata ed a pieno rango n, allora le sole operazioni di riga danno

$$\mathbf{R}_1 \cdots \mathbf{R}_p \mathbf{A} = \mathbf{I} \tag{4}$$

e:

$$\mathbf{A}^{-1} = \mathbf{R}_1 \cdots \mathbf{R}_p \mathbf{I} \tag{5}$$

Di conseguenza, quella *stessa* successione di operazioni elementari di riga che riduce la matrice $\mathbf{A}$ ad una identità, trasforma anche la matrice identità $n \times n$ nella inversa di $\mathbf{A}$. Questa proprietà delle trasformazioni elementari ci fornisce un importante strumento per calcolare le matrici inverse: scriviamo per semplicità la matrice di dimensioni $n \times 2n$ come:

$$[\mathbf{A} \quad \mathbf{I}] \tag{6}$$

ed applichiamo le operazioni elementari a ciascuna riga, con lo scopo di trasformare le prime n colonne nella matrice identità. Fatto ciò nelle ultime n colonne si leggeranno gli elementi di $\mathbf{A}^{-1}$. Per matrici simmetriche tale metodo è stato formalmente definito mediante la tecnica abbreviata di Gauss-Doolittle [3].

Esempio 2.9 - Le trasformazioni successive per l'inversione di:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 3 & 3 \\ 3 & -1 & 8 \end{bmatrix}$$

sono qui di seguito riportate:

$$\begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 2 & 3 & 3 & 0 & 1 & 0 \\ 3 & -1 & 8 & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & -1 & -2 & 1 & 0 \\ 0 & -4 & 2 & -3 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 3 & 3 & -1 & -1 \\ 0 & 1 & -1 & -2 & 1 & 0 \\ 0 & 0 & -2 & -11 & 4 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & -13,5 & 5 & 1,5 \\ 0 & 1 & 0 & 3,5 & -1 & -0,5 \\ 0 & 0 & 1 & 5,5 & -2 & -0,5 \end{bmatrix}$$

e:

$$\mathbf{A}^{-1} = \begin{bmatrix} -13,5 & 5 & 1,5 \\ 3,5 & -1 & -0,5 \\ 5,5 & -2 & -0,5 \end{bmatrix}$$

Il determinante di una matrice può anche essere calcolato mediante operazioni elementari di riga o di colonna perchè per la proprietà 5 del paragrafo 2.4, tali trasformazioni non ne alterano il valore. Il procedimento chiamato *pivotal-condensation*, consiste nel ridurre la matrice a una forma triangolare, per la quale viene calcolato il determinante come il prodotto degli elementi diagonali. Se si usano operazioni di riga, la matrice è scritta con l'elemento maggiore della prima colonna nella posizione (1,1). Tale elemento, usato come *perno* o « pivot », permette di ridurre i rimanenti elementi della colonna al valore zero. La precedente rielaborazione e riduzione può essere continuata fino ad ottenere per la matrice la forma triangolare. La scelta dell'elemento maggiore, come « *perno* » va a vantaggio dell'accuratezza numerica [12, p. 59].

Esempio 2.10 - Si valuti il determinante della matrice

$$\mathbf{A} = \begin{bmatrix} 10 & 5 & -2 & 0 \\ 6 & 3 & 2 & 1 \\ 4 & 5 & 12 & 3 \\ 5 & 1 & 3 & 8 \end{bmatrix}$$

Innanzitutto si sottraggano opportuni multipli della prima riga dalle rimanenti:

$$\begin{bmatrix} 10 & 5 & -2 & 0 \\ 0 & 0 & 3{,}2 & 1 \\ 0 & 3 & 12{,}8 & 3 \\ 0 & -1{,}5 & 4 & 8 \end{bmatrix}$$

Poichè l'elemento (2,2) è nullo, si scambino la seconda e la terza riga della matrice, e si compensi il conseguente cambiamento di segno del determinante, con la moltiplicazione di ciascun elemento della nuova seconda riga per il fattore -1:

$$\begin{bmatrix} 10 & 5 & -2 & 0 \\ 0 & -3 & -12{,}8 & -3 \\ 0 & 0 & 3{,}2 & 1 \\ 0 & -1{,}5 & 4 & 8 \end{bmatrix}$$

Una sola operazione di riga nella seconda colonna la trasforma in:

$$\begin{bmatrix} 10 & 5 & -2 & 0 \\ 0 & -3 & -12{,}8 & -3 \\ 0 & 0 & 3{,}2 & 1 \\ 0 & 0 & 10{,}4 & 9{,}5 \end{bmatrix}$$

Infine si ottiene:

$$\begin{bmatrix} 10 & 5 & -2 & 0 \\ 0 & -3 & -12{,}8 & -3 \\ 0 & 0 & 3{,}2 & 1 \\ 0 & 0 & 0 & 6{,}25 \end{bmatrix}$$

e il determinante è:

$$\begin{aligned} |\mathbf{A}| &= (10)\,(-3)\,(3{,}2)\,(6{,}25) \\ &= -600 \end{aligned}$$

2.7 Sistemi di equazioni lineari

L'insieme di equazioni nelle incognite $x_1, \ldots, x_n$

$$\begin{aligned} a_{11}x_1 + \quad \cdots \quad + a_{1n}x_n &= c_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{m1}x_1 + \quad \cdots \quad + a_{mn}x_n &= c_m \end{aligned} \tag{1}$$

è chiamato *sistema* di *m equazioni lineari, in n incognite*, e può essere scritto sinteticamente nella forma matriciale:

$$\mathbf{A}\mathbf{x} = \mathbf{c} \tag{2}$$

dove **A** è la matrice $m \times n$ dei coefficienti, $\mathbf{x}' = [x_1, \ldots, x_n]$ e $\mathbf{c}' = [c_1, \ldots, c_m]$.

Il sistema generale $\mathbf{Ax} = \mathbf{c}$ di m equazioni in n incognite ha una soluzione se e soltanto se la *matrice di dimensioni aumentate* $m \times (n + 1)$:

$$[\mathbf{A} \quad \mathbf{c}]$$

ha lo stesso rango r di **A**. In caso contrario il sistema è detto impossibile.

È necessario distinguere tra sistemi omogenei, per i quali $\mathbf{c} = \mathbf{0}$, e sistemi non omogenei. Si considereranno tre tipi di equazioni particolarmente importanti per i nostri scopi.

Sistema non omogeneo: **A** matrice *quadrata* e *non singolare*. Poichè esiste $\mathbf{A}^{-1}$ la soluzione unica del sistema è

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{c} \tag{3}$$

Questo si ricava moltiplicando a sinistra entrambi i termini dell'equazione (1) per $\mathbf{A}^{-1}$.

Esempio 2.11 - La matrice inversa del sistema

$$\begin{aligned} x_1 + x_2 - x_3 &= 1 \\ -x_1 + x_2 + x_3 &= -1 \\ x_1 - x_2 + x_3 &= 1 \end{aligned}$$

è

$$\begin{bmatrix} 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \end{bmatrix}$$

la soluzione è:

$$\begin{aligned} x_1 &= 1/2(1) + 0(-1) + 1/2(1) \\ &= 1 \\ x_2 &= 1/2(1) + 1/2(-1) + 0(1) \\ &= 0 \\ x_3 &= 0(1) + 1/2(-1) + 1/2(1) \\ &= 0 \end{aligned}$$

Sistema non omogeneo: **A** *matrice* di dimensione $m \times n$ *e di rango* r: Se è soddisfatta la condizione sui ranghi della matrice dei coefficienti e di quella dei coefficienti aumentata dei termini noti, si può calcolare una soluzione scegliendo r qualsiasi equazioni linearmente indipendenti e risolvendole per r delle n incognite in funzione delle costanti c_i e delle rimanenti $n - r$ variabili.

Esempio 2.12 - Le equazioni

$$x_1 + x_2 = 2$$
$$x_1 + x_2 = 1$$

possono essere interpretate geometricamente, come rette parallele nel piano $x_1\ x_2$.

Poichè queste non hanno alcun punto in comune, il sistema non può avere soluzione.

Il rango delle matrici dei coefficienti

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

è uno, mentre quello della matrice aumentata dei termini noti

$$\begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix}$$

è due.

Esempio 2.13 - Il sistema.

$$2x_1 + 3x_2 - x_3 = 1$$
$$x_2 + x_3 = 2$$

ha sia la matrice dei coefficienti, che la matrice dei coefficienti aumentate dei termini noti, di rango due.

Si scrivano le equazioni come

$$2x_1 + 3x_2 = 1 + x_3$$
$$x_2 = 2 - x_3$$

e si risolvano rispetto ad x_1 con il metodo di sostituzione. La soluzione generale è data dal vettore.

$$[{}^1/_2(-5 + 4x_3),\ 2 - x_3,\ x_3]$$

in cui x_3 può assumere qualunque valore.

Sistemi omogenei di equazioni.

Non esiste alcun sistema omogeneo di equazioni lineari impossibile, in quanto il rango di $[\mathbf{A}\ \ \mathbf{0}]$ è eguale a quello di $\mathbf{A}$.

Ciascun sistema omogeneo è soddisfatto dalla soluzione nulla banale $\mathbf{x}' = [0, \ldots, 0]$, e in realtà se il rango della matrice dei coefficienti è uguale al numero delle incognite n, questa è la *sola* soluzione.

Quindi un sistema omogeneo avrà una soluzione non banale se e soltanto se il rango r della matrice $\mathbf{A}$ è minore di n; e inoltre è sempre possibile trovare $n - r$ soluzioni *linearmente indipendenti del sistema*, tali che ogni loro combinazione lineare sia essa stessa una soluzione.

In applicazioni statistiche si incontrano frequentemente sistemi con tante incognite quante sono le equazioni; tali sistemi hanno soluzioni non banali se e soltanto se i determinanti dei loro coefficienti sono nulli.

Esempio 2.14 - Il sistema

$$3x_1 - 2x_2 = 0$$
$$5x_1 + x_2 = 0$$

definisce due rette del piano $x_1\, x_2$ intersecantesi nell'origine. Quindi la sola soluzione è quella banale [0,0].

Esempio 2.15 - Nel sistema

$$x_1 - x_2 + 2x_3 = 0$$
$$x_1 + 3x_2 - 2x_3 = 0$$
$$3x_1 + x_2 + 2x_3 = 0$$

la terza equazione è pari alla somma del doppio della prima equazione con la seconda, e quindi il suo rango è due.

L'unica soluzione indipendente può essere determinata in funzione di x_3, risolvendo il sistema

$$x_1 - x_2 = -2x_3$$
$$x_1 + 3x_2 = 2x_3$$

La soluzione generale per il sistema originale è, in unità di x_3,

$$x_1 = -x_3$$
$$x_2 = x_3$$

Esempio 2.16 - Il sistema

$$x_1 + x_2 - x_3 - x_4 = 0$$
$$-2x_2 + 3x_3 + x_4 = 0$$
$$x_1 - x_2 + x_3 = 0$$
$$3x_1 + x_2 - 2x_4 = 0$$

è anch'esso di rango 2.

La prima e la terza equazione infatti possono essere convenientemente espresse in termini di x_3 ed x_4 dando il vettore soluzione

$$[{}^1/_2(-x_3 + x_4),\ {}^1/_2(3x_3 + x_4),\ x_3,\ x_4]$$

Se si assume x_3 eguale a zero ed x_4 eguale ad uno o alternativamente x_3 eguale ad uno ed x_4 eguale a zero, si otterranno i seguenti valori *linearmente indipendenti* del vettore soluzione:

$$[{}^1/_2,\ {}^1/_2,\ 0,\ 1]$$
$$[-{}^1/_2,\ {}^3/_2,\ 1,\ 0]$$

Poichè ogni combinazione lineare di queste soluzioni è ancora soluzione del sistema originale, la soluzione più generale sarà:

$$x_1 = {}^1/_2(a - b)$$
$$x_2 = {}^1/_2(a + 3b)$$
$$x_3 = b$$
$$x_4 = a$$

con a e b scelti arbitrariamente

Soluzioni numeriche di equazioni lineari

Il metodo diretto più comune per risolvere un sistema di rango r in r incognite, è il metodo di eliminazione di Gauss-Doolittle.

Come precedentemente accennato, tale tecnica consiste nel ridurre il sistema ad una forma triangolare mediante operazioni elementari di riga e nel risolvere poi a ritroso le nuove equazioni fino ad ottenere tutte le incognite.

Lo schema, nella sua forma usuale prevede che la matrice dei coefficienti sia simmetrica, ma sono validi anche altri procedimenti generalizzati. Si rimanda il lettore a Dwyer [3] Faddeeva [5] e Graybill [7], per una estesa discussione ed per esempi risolti relativi al metodo di Gauss-Doolittle.

Un altro utile metodo diretto di risoluzione è quello della radice quadrata. Se il sistema di equazioni originale è stato ridotto alla forma quadrata non omogenea:

$$\mathbf{Ax} = \mathbf{c}$$

a pieno rango r, si può mostrare che la matrice dei coefficienti può essere fattorizzata nel prodotto di una matrice triangolare $\mathbf{T}$ e nella sua transposta:

$$\mathbf{A} = \mathbf{TT}'$$

È innanzi tutto necessario determinare gli elementi di $\mathbf{T}$ mediante una successione di rapporti quadratici ricorrenti. Posto allora $\mathbf{y} = \mathbf{T}'\mathbf{x}$, il sistema triangolare

$$\mathbf{Ty} = \mathbf{c}$$

può essere risolto a ritroso, rispetto agli elementi di $\mathbf{y}$.

Infine, è necessario risolvere l'altro sistema triangolare

$$\mathbf{T}'\mathbf{x} = \mathbf{y}$$

per il vettore incognito originale $\mathbf{x}$. Dal Dwyer [3] e dal Faddeeva [5] sono state ricavate espressioni per gli elementi di $\mathbf{T}$ ed efficaci metodi di calcolo pratico.

Spesso è necessario calcolare esplicitamente la matrice inversa dei coefficienti, invece della soluzione delle relative equazioni.

Ciò è particolarmente vero nell'analisi della regressione, perchè le varianze e le covarianze campionarie delle stime ai minimi quadrati dei parametri della regressione, sono determinate dagli elementi della matrice inversa.

Il problema dell'inversione di una matrice può essere formulato come la soluzione di r sistemi di equazioni aventi tutti, quale matrice dei coefficienti, la matrice $\mathbf{A}$ ma con vettori costanti $\mathbf{c}$ dati dalle successive colonne della matrice identità di dimensioni $r \times r$.

La soluzione di tali sistemi può essere ottenuta mediante il metodo abbreviato di Gauss-Doolittle. Nei testi di Dwyer e Graybill si possono trovare particolari ed esempi relativi a queste procedure.

I metodi precedenti, basati su operazioni elementari di riga sono tutti *esatti*, nel senso che un numero finito di operazioni aritmetiche porterà alla soluzione con una accuratezza che dipende solo dalla precisione mantenuta a ciascuno stadio dei calcoli e naturalmente dalla esatta o approssimativa natura dei coefficienti della matrice originale.

I metodi *iterativi* partono da una approssimazione o da una ipotesi plausibile della soluzione e ripetutamente correggono questi valori di prova fino a convergere, con un determinato grado di accuratezza, alla soluzione effettiva.

Forse la tecnica iterativa più familiare per risolvere i sistemi di equazioni lineari è il metodo di Gauss-Seidel [5, 10]. La convergenza di questo algoritmo dipende in massima parte dal valore degli elementi nella matrice dei coefficienti. Le condizioni necessarie e sufficienti per la convergenza sono state studiate dal Faddeeva [5], insieme con altri metodi iterativi dotati di proprietà più generali o più sintetiche di convergenza.

2.8 Vettori e matrici ortogonali

Nel paragrafo 2.3 è stata introdotta l'espressione

$$\cos\ \theta = \frac{\mathbf{x}'\mathbf{y}}{(\mathbf{x}'\mathbf{x})^{1/2}\,(\mathbf{y}'\mathbf{y})^{1/2}} \tag{1}$$

per il coseno dell'angolo compreso tra il vettore $\mathbf{x}$ e il vettore $\mathbf{y}$, e si è visto che, se il prodotto interno dei due vettori risulta uguale a zero, i due vettori sono perpendicolari l'uno all'altro. Tali vettori sono detti « *ortogonali* » o « *ortonormali* » nel caso particolare in cui le loro lunghezze siano state normalizzate all'unità.

Una *matrice ortogonale* $\mathbf{T}$ è una matrice quadrata le cui righe costituiscono un insieme di vettori ortonormali. Dato che:

$$\begin{aligned}\mathbf{T}\mathbf{T}' &= \mathbf{T}'\mathbf{T} \\ &= \mathbf{I}\end{aligned} \tag{2}$$

l'inversa di $\mathbf{T}$ è semplicemente la sua transporta $\mathbf{T}'$. Una semplice matrice ortogonale 2×2 è

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Questa è la matrice di trasformazione per la riflessione dei punti nel piano xy attorno alla retta inclinata di 45° Una più generale matrice orto-

gonale delle stesse dimensioni è

$$\mathbf{T} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{3}$$

interpretabile come la matrice di trasformazione per una rotazione degli assi coordinati xy di un angolo θ.

Cioè se x e y fossero le coordinate di un punto rispetto ai vecchi assi il punto avrebbe le coordinate

$$u = x\cos\theta + y\sin\theta$$
$$v = -x\sin\theta + y\cos\theta$$

dopo una rotazione rigida degli assi di un angolo di θ gradi. Matrici ortogonali di maggiori dimensioni possono essere costruite mediante il processo di ortogonalizzazione di Gram-Schmidt [8, cap. 7], che parte da una riga normalizzata e procede costruendo su questa una matrice secondo l'esigenza della mutua-ortogonalità delle righe stesse. Per es. la matrice di Helmert

$$\mathbf{T} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \end{bmatrix}$$

è una matrice ortogonale particolare 3×3.

La matrice ortogonale $\mathbf{T}$ $n \times n$ si può considerare come la matrice della trasformazione lineare equivalente ad una *rotazione rigida* o ad una rotazione seguita da una *riflessione* degli n assi coordinati rispetto alla loro origine.

Le distanze tra due punti nello spazio rimangono invariate per rotazione. Infatti se operiamo la trasformazione

$$\mathbf{y} = \mathbf{Tx}$$

si ha:

$$\mathbf{x'x} = \mathbf{y'T'Ty}$$
$$= \mathbf{y'Iy}$$

da cui discende che le somme dei quadrati rimangono invariate.

Le matrici ortogonali godono delle seguenti importanti proprietà:

1. Le colonne di una matrice ortogonale sono mutuamente ortogonali
2. Il determinante di una matrice ortogonale è sempre o 1 o -1. Poichè $|\mathbf{TT'}| = |\mathbf{I}| = 1$ e $|\mathbf{T}| = |\mathbf{T'}|$, allora $|\mathbf{T}| = \pm 1$.

3. Il prodotto di matrici ortogonali della stessa dimensione è esso stesso ortogonale. Quindi una successione di rotazioni rigide e riflessioni degli assi coordinati è esprimibile come una sola rotazione rigida ed opportune riflessioni.

2.9 Forme quadratiche

Si definisce *forma quadratica* nelle variabili $x_1, \ldots, x_n$ una espressione del tipo

$$\begin{aligned} f(x_1, \ldots, x_n) &= a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 \\ &\quad + 2a_{12}x_1x_2 + \cdots + 2a_{1n}x_1x_n + \cdots + 2a_{n-1,n}x_{n-1}x_n \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j \end{aligned} \tag{1}$$

dove $a_{ij} = a_{ji}$. Alcuni degli a_{ij} possono essere nulli. Si nota immediatamente che la forma quadratica può essere scritta nella notazione matriciale come

$$\mathbf{x}'\mathbf{A}\mathbf{x}$$

dove $\mathbf{x}' = [x_1, \ldots, x_n]$ ed $\mathbf{A}$ è la matrice simmetrica dei coefficienti di dimensioni $n \times n$.

La forma quadratica più semplice è $f(x) = a_{11}x^2$, equazione di una parabola nella sola variabile x. Le forme quadratiche nelle quali gli x_i sono variabili casuali giocano un importante ruolo sia nella teoria statistica multivariata che in quella univariata.

Per esempio la somma dei quadrati delle deviazioni dalla media del campione

$$\sum_{i=1}^{N}(x_i - \bar{x})^2 = \sum_{i=1}^{N} x_i^2 - \frac{1}{N}\left(\sum x_i\right)^2 \tag{2}$$

può essere scritta come forma quadratica nelle osservazioni x_i con la matrice

$$\mathbf{A} = \begin{bmatrix} \frac{N-1}{N} & \frac{-1}{N} & \cdots & \frac{-1}{N} \\ \frac{-1}{N} & \frac{N-1}{N} & \cdots & \frac{-1}{N} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{-1}{N} & \frac{-1}{N} & \cdots & \frac{N-1}{N} \end{bmatrix} \tag{3}$$

Una matrice simmetrica $\mathbf{A}$ e la forma quadratica ad essa associata sono dette *definite positive* se $\mathbf{x}'\mathbf{A}\mathbf{x} > \mathbf{0}$ per tutti gli $\mathbf{x}$ diversi da zero.

Se $\mathbf{x}'\mathbf{A}\mathbf{x} \geq \mathbf{0}$ la forma e la sua matrice sono dette *semidefinite positive*.

Definizioni simili vengono date per la forma quadratica *definita e semidefinita negativa*.

Una forma *indefinita* può assumere sia valori positivi, nulli e negativi. Sono utili per la trattazione successiva particolari proprietà delle matrici e delle forme quadratiche definite e semidefinite positive.

Una forma quadratica definita positiva ha matrici a *pieno rango*. È possibile, mediante operazioni ripetute, ridurre una forma quadratica in n variabili alla forma

$$d_1 y_1^2 + \cdots + d_n y_n^2 \tag{4}$$

contenente solo i quadrati delle nuove variabili y_i e con coefficienti $d_i > 0$. In modo analogo una forma quadratica, semidefinita positiva può essere ridotta a:

$$d_1 y_1^2 + \cdots + d_r y_r^2 \tag{5}$$

dove tutti i coefficienti sono positivi e $r \leq n$ è il rango della matrice corrispondente. Tuttavia questi non sono metodi atti a determinare convenientemente la natura della forma, per cui daremo ora una condizione necessaria e sufficiente per stabilire se una forma quadratica è definita o semidefinita positiva. Posta la successione progressiva dei *determinanti* relativi ai *minori principali* estratti dalla matrice:

$$p_1 = a_{11} \qquad p_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} \qquad \cdots \qquad p_i = \begin{vmatrix} a_{11} & \cdots & a_{1i} \\ \cdots & \cdots & \cdots \\ a_{1i} & \cdots & a_{ii} \end{vmatrix} \tag{6}$$

$$\cdots \qquad \text{e} \qquad p_n = |A|$$

si ha:

1. Condizione necessaria e sufficiente perchè la forma sia definita positiva è che $p_i > 0$ per $i = 1, \ldots, n$.

2. Condizione necessaria e sufficiente perchè la forma sia semidefinita positiva è che $p_1 > 0, \ldots, p_r > 0$, mentre i rimanenti $n - r$ p_i siano eguali a zero, dove r può essere eguale ad n .

Esempio 2.17 - La matrice

$$\begin{bmatrix} 2 & 2 & 2 \\ 2 & 3 & 1 \\ 2 & 1 & 3 \end{bmatrix}$$

ha quali determinanti dei minori principali in ordine progressivo:

$$p_1 = 2 \qquad p_2 = 2 \qquad p_3 = 0$$

La matrice e la sua forma quadratica sono pertanto semidefinite positive con rango pari a due.

Esempio 2.18 - La matrice (3) della somma dei quadrati relativi al campione è un caso particolare della matrice formata con un elemento diagonale a comune ed elementi b eguali al di fuori della diagonale. Il determinante di tale matrice di dimensione $i \times i$ è

$$(a - b)^{i-1}[a + (i - 1)b]$$

così che i successivi determinanti dei minori principali estratti da (3) sono

$$p_i = \frac{1}{N}[N - 1 - (i - 1)]$$

I primi $N - 1$ di questi sono positivi mentre l'ennesimo è sempre zero. La somma dei quadrati delle deviazioni dalla media è quindi una forma quadratica semidefinita positiva di rango $N - 1$.

2.10 Radici e vettori caratteristici di una matrice

Le *radici caratteristiche* di una matrice $\mathbf{A}$ di dimensioni $p \times p$ sono le soluzioni dell'*equazione*

$$|\mathbf{A} - \lambda\mathbf{I}| = 0 \tag{1}$$

Il determinante (1) è un polinomio di p-esimo grado in λ e quindi $\mathbf{A}$ ha esattamente p radici caratteristiche. Lo sviluppo di Laplace relativo al *determinante caratteristico*, ci conduce a scrivere *la polinoniale caratteristica* nella forma:

$$|\mathbf{A} - \lambda\mathbf{I}| = (-\lambda)^p + S_1(-\lambda)^{p-1} + S_2(-\lambda)^{p-2} + \cdots - S_{p-1}\lambda + |\mathbf{A}| \tag{2}$$

dove S_i indica la somma di tutti i determinanti relativi ai minori principali di dimensioni $i \times i$.

S_1 è evidentemente la somma degli elementi diagonali di $\mathbf{A}$ o tr $\mathbf{A}$. Dalla teoria delle equazioni polinomiali discende immediatamente che:

1. Il *prodotto* delle radici caratteristiche di $\mathbf{A}$ è uguale ad $|\mathbf{A}|$
2. La *somma* delle radici caratteristiche di $\mathbf{A}$ è eguale alla *traccia* di $\mathbf{A}$

Esempio 2.19 - La matrice

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

ha traccia 6 e i tre determinati relativi ai minori principali di dimensione 2×2 sono ciascuno eguale a

$$\begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 3$$

mentre $|\mathbf{A}| = 4$. L'equazione caratteristica di **A** è

$$\lambda^3 - 6\lambda^2 + 9\lambda - 4 = 0$$

e le sue radici sono pertanto 1,1 e 4.

In seguito saranno di uso frequente le seguenti proprietà delle radici caratteristiche:

1. Le radici caratteristiche di una matrice simmetrica ad elementi reali, sono tutte reali.

2. Le radici caratteristiche di una matrice definita positiva sono tutte positive.

3. Se una matrice simmetrica di dimensioni $n \times n$ è semidefinita positiva e di rango r, essa contiene esattamente r radici caratteristiche positive ed $n - r$ radici nulle.

4. Le radici caratteristiche diverse da zero del prodotto **AB** sono eguali alle radici non nulle di **BA**. Da ciò segue che le tracce di **AB** e **BA** sono eguali.

5. Le radici caratteristiche di una matrice diagonale sono gli elementi diagonali stessi.

Associato a ciascuna radice caratteristica λ_i della matrice quadrata **A** vi è il *vettore caratteristico* $\mathbf{x}_i$ i cui elementi soddisfano il sistema di equazioni omogeneo:

$$[\mathbf{A} - \lambda_i \mathbf{I}]\, \mathbf{x}_i = \mathbf{0} \tag{3}$$

Per la definizione data di radice caratteristica il determinante del sistema risulta eguale a zero e di conseguenza esiste sempre una soluzione non banale $\mathbf{x}_i$.

Si vede subito che gli elementi del vettore sono determinati solo a meno di un fattore di scala. Molti dei vettori caratteristici che si incontreranno in seguito saranno calcolati da matrici simmetriche. Le radici ed i vettori caratteristici di tali matrici godono delle seguenti importanti proprietà:

1. Se λ_i e λ_j sono radici caratteristiche distinte di una matrice simmetrica **A**, i vettori ad esse associate $\mathbf{x}_i$ ed $\mathbf{x}_j$ sono ortogonali. Ciò risulta evidente moltiplicando a sinistra l'espressione di definizione dei vettori stessi:

$$\mathbf{A}\mathbf{x}_i = \lambda_i \mathbf{x}_i \qquad \mathbf{A}\mathbf{x}_j = \lambda_j \mathbf{x}_j$$

per $\mathbf{x}'_i$ ed $\mathbf{x}'_j$ rispettivamente. Ma ciò implica che

$$\lambda_i \mathbf{x}'_j \mathbf{x}_i = \lambda_j \mathbf{x}'_i \mathbf{x}_j$$

e poichè $\lambda_i \neq \lambda_j$ si conclude che i vettori sono ortogonali.

2. Per ciascuna matrice reale simmetrica **A** esiste una matrice ortogonale **P** tale che

$$\mathbf{P'AP} = \mathbf{D} \tag{4}$$

dove **D** è la matrice diagonale delle radici caratteristiche di **A**. I vettori caratteristici normalizzati di **A** possono essere presi come colonne di **P**.

Se le radici caratteristiche non sono distinte, è ancora *possibile* selezionare gli elementi di vettori ad esse associati per ottenere un insieme di vettori caratteristici mutuamente ortogonali.

Le precedenti proprietà delle matrici simmetriche hanno una importante applicazione relativa alle forme quadratiche. Se si opera la trasformazione ortogonale

$$\mathbf{x} = \mathbf{Py} \tag{5}$$

sulle p variabili della forma quadratica $\mathbf{x'Ax}$ la forma diventa

$$\begin{aligned} \mathbf{x'Ax} &= \mathbf{y'P'APy} \\ &= \mathbf{y'Dy} \\ &= \lambda_1 y_1^2 + \cdots + \lambda_r y_r^2 \end{aligned} \tag{6}$$

dove i λ_i sono le radici caratteristiche della matrice dei coefficienti ed r è il rango della forma. Ogni forma quadratica reale può essere ridotta ad una somma di quadrati mediante il calcolo delle radici dei vettori caratteristici della matrice corrispondente.

Esempio 2.20 - Il vettore caratteristico relativo alla radice maggiore dell'esempio 2.19 deve soddisfare le equazioni

$$\begin{aligned} -2x_{11} + x_{12} + x_{13} &= 0 \\ x_{11} - 2x_{12} + x_{13} &= 0 \\ x_{11} + x_{12} - 2x_{13} &= 0 \end{aligned}$$

Ponendo arbitrariamente x_{13} eguale ad uno e risolvendo le prime 2 equazioni non omogenee, il vettore caratteristico corrispondente è:

$$\mathbf{x'} = [1, 1, 1]$$

ed evidentemente ogni vettore non nullo del tipo $[a,a,a]$ è un vettore caratteristico per $\lambda_3 = 4$.

Il vettore associato alla radice doppia $\lambda_1 = \lambda_2 = 1$ deve soddisfare il sistema con la sola equazione linearmente indipendente

$$x_{21} + x_{22} + x_{23} = 0$$

La soluzione più generale per essa è il vettore

$$[a, b, -a - b]$$

da cui scegliendo opportunamente a e b possiamo ottenere due vettori caratteristici della radice doppia linearmente indipendenti.

Si noti che, comunque si scelgano a e b, si otterranno vettori ortogonali al primo vettore [1,1,1].

È possibile assumere a e b nel secondo e terzo vettore in modo da conservare tale ortogonalità.

Per esempio:

$$\mathbf{x}'_1 = [0, 1, -1]$$

$$\mathbf{x}'_2 = [-2, 1, 1]$$

Esempio 2.21 - Matrici quadrate che godono della proprietà

$$\mathbf{AA} = \mathbf{A}$$

sono chiamate *equipotenti* e giocano un importantissimo ruolo nella teoria dell'analisi della varianza.

Le radici caratteristiche di una matrice equipotente sono o zero o uno, ed una forma quadratica con una siffatta matrice può essere ridotta ad una somma di r termini al quadrato. È facile verificare che la matrice (3) del paragrafo 2.9 è equipotente e che è quindi possibile operare la trasformazione:

$$\sum_{i=1}^{N} (x_i - \bar{x})^2 = y_1^2 + \cdots + y^2_{N-1}$$

nella somma dei quadrati di $N - 1$ nuove variabili casuali, la cui indipendenza statistica discende dalla ortogonalità della trasformazione.

Si rimanda la discussione dei metodi di calcolo relativi ai vettori ed alle radici caratteristiche al capitolo 7 dove verrà trattata la tecnica delle componenti principali. Si può trovare una trattazione più estesa su tali tecniche di calcolo nei trattati di Bodewig [1], Faddeeva [5] e Householder [11]. Sono stati ottenuti limiti per le radici caratteristiche da diversi algebristi. Questi risultati sono stati raccolti e riassunti da Marcus e Minc[13].

2.11 Matrici ripartite

Risulterà spesso conveniente pensare di raggruppare insieme, in base a comuni caratteristiche delle variabili ad esse associate, un certo numero di righe e di colonne di una matrice. La matrice così ottenuta può essere scritta come lo schema

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \cdots & \mathbf{A}_{1n} \\ \cdots & \cdots & \cdots \\ \mathbf{A}_{m1} & \cdots & \mathbf{A}_{mn} \end{bmatrix} \tag{1}$$

di sottomatrici $\mathbf{A}_{ij}$ contenenti r_i righe e c_j colonne, dove è ovvio che tutte le sottomatrici in una data riga di $\mathbf{A}$ devono avere lo stesso numero di righe, e ciascuna colonna deve essere costituita da matrici con lo stesso numero di colonne.

Le operazioni che coinvolgono matrici ripartite sono dello stesso tipo

di quelle relative a matrici non ripartite, purchè si tenga conto della natura non scalare degli elementi.

La somma di due matrici ripartite **A** e **B** aventi sottomatrici delle stesse dimensioni è

$$\mathbf{A}+\mathbf{B}=\begin{bmatrix}\mathbf{A}_{11}+\mathbf{B}_{11} & \cdots & \mathbf{A}_{1n}+\mathbf{B}_{1n}\\ \cdots & \cdots & \cdots\\ \mathbf{A}_{m1}+\mathbf{B}_{m1} & \cdots & \mathbf{A}_{mn}+\mathbf{B}_{mn}\end{bmatrix} \tag{2}$$

Il prodotto delle matrici ripartite **A** e **B** è

$$\mathbf{AB}=\begin{bmatrix}\sum_{j=1}^{n}\mathbf{A}_{1j}\mathbf{B}_{j1} & \cdots & \sum_{j=1}^{n}\mathbf{A}_{nj}\mathbf{B}_{jp}\\ \cdots & \cdots & \cdots\\ \sum_{j=1}^{n}\mathbf{A}_{mj}\mathbf{B}_{j1} & \cdots & \sum_{j=1}^{n}\mathbf{A}_{mj}\mathbf{B}_{jp}\end{bmatrix} \tag{3}$$

Le dimensioni delle sottomatrici che entrano in un prodotto devono essere conformi; se le matrici di **A** hanno rispettivamente colonne $c_1, \ldots, c_n$ quelle di **B** devono avere dimensioni di riga $c_1, \ldots, c_n$.

Quando devono essere esplicitati gli elementi di matrici ripartite come negli esempi numerici, è conveniente separare le sotto matrici con linee tratteggiate o spazi appropriati.

Esempio 2.22 - Sia

$$\mathbf{A}=\left[\begin{array}{ccc:c}1 & 0 & 3 & \\ -2 & 3 & 1 & \\ \hdashline 1 & 1 & 2 & \end{array}\right] \qquad \mathbf{B}=\left[\begin{array}{cc}1 & 1\\ 2 & 0\\ \hline 3 & 1\end{array}\right]$$

allora:

$$\mathbf{AB}=\left[\begin{array}{cc}10 & 4\\ 7 & -1\\ \hline 9 & 3\end{array}\right]$$

È anche possibile esprimere l'inversa di una matrice ripartita non singolare, nei termini delle sue sottomatrici. Nel caso particolarmente importante di:

$$\mathbf{A}=\begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \mathbf{A}_{21} & \mathbf{A}_{22}\end{bmatrix} \tag{4}$$

dove $\mathbf{A}_{11}$ e $\mathbf{A}_{22}$ sono entrambe quadrate, e con rispettivi minori principali di natura non singolare, si può verificare che

$$\mathbf{A}^{-1} = \tag{5}$$

$$\begin{bmatrix} (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1} & -(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1} & \mathbf{A}_{22}^{-1} + \mathbf{A}_{22}^{-1}\mathbf{A}_{21}(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \end{bmatrix}$$

Una espressione alternativa può essere ottenuta, invertendo le posizioni di $\mathbf{A}_{11}$ e $\mathbf{A}_{22}$ nella matrice originale.

È spesso necessario calcolare il determinante della matrice ripartita (4). Se $\mathbf{A}_{11}$ è non singolare si ha:

$$|\mathbf{A}| = |\mathbf{A}_{11}| \cdot |\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}| \tag{6}$$

Se $\mathbf{A}_{22}$ è non singolare si ha;

$$|\mathbf{A}| = |\mathbf{A}_{22}| \cdot |\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}| \tag{7}$$

Esempio 2.23 - Le precedenti espressioni per le matrici inverse e per i determinanti sono valutabili con calcoli pratici, se le sottomatrici sono di piccole dimensioni o con forma opportuna. Se

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

si può calcolare $\mathbf{A}^{-1}$ con una ripartizione delle matrici nei termini della prima, seconda, terza riga e colonna sino alla quinta. Allora

$$\mathbf{A}_{11}^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \mathbf{A}_{22}^{-1} = \begin{bmatrix} 1/4 & 0 & 0 \\ 0 & 1/4 & 0 \\ 0 & 0 & 1/4 \end{bmatrix}$$

$$\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} = \begin{bmatrix} -5/4 & 3/4 \\ -9/2 & 5/2 \end{bmatrix}$$

$$(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1} = \begin{bmatrix} 10 & -3 \\ 18 & -5 \end{bmatrix}$$

$$\mathbf{A}_{12}\mathbf{A}_{22}^{-1} = \begin{bmatrix} 1/4 & 1/4 & 1/4 \\ 1/2 & 1/2 & 1/2 \end{bmatrix} \qquad \mathbf{A}_{22}^{-1}\mathbf{A}_{21} = \begin{bmatrix} 3/4 & -1/4 \\ 3/4 & -1/4 \\ 3/4 & -1/4 \end{bmatrix}$$

Quindi

$$\mathbf{A}^{-1} = \begin{bmatrix} 10 & -3 & -1 & -1 & -1 \\ 18 & -5 & -2 & -2 & -2 \\ -3 & 1 & 1/2 & 1/4 & 1/4 \\ -3 & 1 & 1/4 & 1/2 & 1/4 \\ -3 & 1 & 1/4 & 1/4 & 1/2 \end{bmatrix}$$

Il determinante di **A** può essere calcolato con risultato 16 dalla (6) o (7).

2.12 Differenziazione di funzioni e di vettori

Talvolta sarà necessario conoscere le derivate parziali di alcune semplici funzioni di un vettore di variabili, rispetto agli elementi del vettore stesso. Tali derivate possono essere scritte in forma vettoriale. Si definisce l'operatore vettoriale derivata parziale

$$\frac{\partial}{\partial \mathbf{x}} = \begin{bmatrix} \dfrac{\partial}{\partial x_1} \\ \cdot \\ \cdot \\ \cdot \\ \dfrac{\partial}{\partial x_p} \end{bmatrix} \tag{1}$$

come il vettore colonna delle derivate parziali degli elementi di $\mathbf{x}' = [x_1, \ldots, x_p]$.

Una applicazione di questo operatore alla funzione scalare $f(\mathbf{x})$ produce il vettore colonna delle derivate parziali

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \begin{bmatrix} \dfrac{\partial f(\mathbf{x})}{\partial x_1} \\ \cdot \\ \cdot \\ \cdot \\ \dfrac{\partial f(\mathbf{x})}{\partial x_p} \end{bmatrix} \tag{2}$$

Incontreremo nei capitoli successivi, le seguenti funzioni e le loro derivate:

1. Se $f(\mathbf{x})$ è costante per tutti gli $\mathbf{x}$ si ha:

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \tag{3}$$

2. Se $f(\mathbf{x}) = \mathbf{a}'\mathbf{x}$, si ha:

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \begin{bmatrix} a_1 \\ \cdot \\ \cdot \\ \cdot \\ a_p \end{bmatrix} \tag{4}$$

Si noti che la forma a colonna del vettore-derivata rimane invariata se si scrive $f(\mathbf{x}) = \mathbf{x}'\mathbf{a}$.

3. Per calcolare il vettore-derivata della forma quadratica $\mathbf{x}'\mathbf{A}\mathbf{x}$, si scriva:

$$\mathbf{x}'\mathbf{A}\mathbf{x} = \sum_{i=1}^{p}\sum_{j=1}^{p} a_{ij}x_i x_j = a_{ii}x_i^2 + 2x_i \sum_{\substack{j=1 \\ j\neq i}} a_{ij}x_j + \sum_{\substack{h\neq i \\ j\neq i}}\sum a_{hj}x_h x_j$$

e si differenzi rispetto ad x_i. Se poniamo $\mathbf{a}'_i$ come la i-esima riga della matrice simmetrica $\mathbf{A}$, la derivata parziale della forma quadratica rispetto ad x_i, è:

$$2\mathbf{a}'_i\mathbf{x} = 2\sum_{j=1}^{p} a_{ij}x_j$$

ed il vettore *delle derivate parziali* è

$$\frac{\partial \mathbf{x}'\mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = 2\mathbf{A}\mathbf{x} \tag{5}$$

In particolare il vettore-derivata della somma dei quadrati $\mathbf{x}'\mathbf{x}$ è semplicemente $2\mathbf{x}$.

La derivata del determinante della matrice $\mathbf{A}$ di dimensioni $n \times n$ rispetto ai suoi elementi a_{ij} può essere calcolata dallo sviluppo di $\mathbf{A}$ in termini di complementi algebrici della i-esima riga o j-esima colonna. Mediante il primo sviluppo *si ottiene*:

$$\begin{aligned} \frac{\partial |\mathbf{A}|}{\partial a_{ij}} &= \frac{\partial}{\partial a_{ij}}(a_{i1}\mathbf{A}_{i1} + \cdots + a_{ij}A_{ij} + \cdots + a_{in}A_{in}) \\ &= A_{ij} \end{aligned} \tag{6}$$

Se $\mathbf{A}$ è simmetrica:

$$\frac{\partial |\mathbf{A}|}{\partial a_{ii}} = A_{ii} \tag{7}$$

e si può dimostrare da formule più generali per derivate di determinanti, i cui elementi sono a loro volta funzioni di altre variabili, che:

$$\frac{\partial |\mathbf{A}|}{\partial a_{ij}} = 2A_{ij} \tag{8}$$

Nella trattazione successiva sarà necessario differenziare forme quadratiche ed altri scalari rispetto agli elementi della loro matrice costituente. Se $\mathbf{X}$ è una matrice di dimensioni $m \times n$ con elemento generico x_{ij} la sua derivata rispetto a questo elemento è:

$$\frac{\partial \mathbf{X}}{\partial x_{ij}} = \mathbf{J}_{ij} \tag{9}$$

dove $\mathbf{J}_{ij}$ è la matrice di dimensioni $m \times n$ con 1 nella posizione i, j-esima e zero in tutte le altre. Se $\mathbf{X}$ è simmetrica:

$$\frac{\partial \mathbf{X}}{\partial x_{ij}} = J_{ij} + J_{ji} \qquad i \neq j \tag{10}$$

La regola per differenziare un prodotto di matrici è simile a quella per gli scalari. Si supponga che gli elementi delle matrici conformi $\mathbf{X}$ e $\mathbf{Y}$ siano funzioni $x_{ij}(z)$, $y_{ij}(z)$ di una variabile z.

Allora:

$$\frac{\partial \mathbf{XY}}{\partial z} = \frac{\partial \mathbf{X}}{\partial z}\mathbf{Y} + \mathbf{X}\frac{\partial \mathbf{Y}}{\partial z} \tag{11}$$

Questa formula fornisce un metodo per differenziare l'inversa della matrice quadrata non singolare $\mathbf{X}$.

Si scriva

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

allora

$$\frac{\partial \mathbf{I}}{\partial x_{ij}} = \mathbf{J}_{ij}\mathbf{X}^{-1} + \mathbf{X}\frac{\partial \mathbf{X}^{-1}}{\partial x_{ij}} = \mathbf{0}$$

e

$$\frac{\partial \mathbf{X}^{-1}}{\partial x_{ij}} = -\mathbf{X}^{-1}\mathbf{J}_{ij}\mathbf{X}^{-1} \tag{12}$$

Se $\mathbf{X}$ è simmetrica

$$\frac{\partial \mathbf{X}^{-1}}{\partial x_{ij}} = \begin{cases} -\mathbf{X}^{-1}\mathbf{J}_{ii}\mathbf{X}^{-1} & i = j \\ -\mathbf{X}^{-1}(\mathbf{J}_{ij} + \mathbf{J}_{ji})\mathbf{X}^{-1} & i \neq j \end{cases} \tag{13}$$

Queste e molte altre formule per la differenziazione delle matrici sono state raccolte sistematicamente da Dwyer e Mac Phail [4].

2.13 Esercizi

1. Sia:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 \\ -1 & 3 & -1 \\ 2 & 2 & 4 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 3 & 2 & -1 \\ 2 & 3 & 1 \\ -1 & 1 & 3 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 2 & 0 \\ -1 & 1 \\ 3 & 2 \end{bmatrix}$$

Svolgere le operazioni matriciali così definite

a. $\mathbf{A} + \mathbf{B}$ b. $\mathbf{A} - 2\mathbf{B}$ c. $\mathbf{A}' + \mathbf{B}$ d. $\mathbf{A} + \mathbf{C}$

e. $(\mathbf{A} + \mathbf{B})'$ f. $(3\mathbf{A}' - 2\mathbf{B})'$

2. Se:

$$\mathbf{P} = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 5 & -1 \\ 1 & -1 & 3 \end{bmatrix} \quad \mathbf{Q} = \begin{bmatrix} 1 & 2 & 2 & 1 \\ 1 & 1 & 1 & 4 \\ 1 & 2 & 2 & 1 \end{bmatrix} \quad \mathbf{R} = \begin{bmatrix} 1 & 2 \\ -5 & 2 \\ 3 & -1 \\ -2 & 2 \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix} \quad \mathbf{z} = \begin{bmatrix} -1 \\ -2 \\ -3 \\ -4 \end{bmatrix}$$

Calcolare i prodotti tra matrici così definiti:

a. $\mathbf{PQ}$ b. $\mathbf{PQR}$ c. $\mathbf{QR}'$ d. $\mathbf{yx}'$

e. $\mathbf{x}'\mathbf{y}$ f. $\mathbf{x}'\mathbf{P}\mathbf{x}$ g. $\mathbf{x}'\mathbf{P}\mathbf{y}$ h. $\mathbf{P}(\mathbf{x} + \mathbf{y})$

3. Se $\mathbf{A}$ è la matrice generale $n \times n$ e $\mathbf{j}$ e $\mathbf{E}$ hanno le forme definite nel paragrafo 2.2, interpretare le seguenti espressioni:

a. $\mathbf{j}'\mathbf{A}$ b. $\mathbf{A}\mathbf{j}$ c. diag $(\mathbf{EA})$ d. diag $(\mathbf{AE})$

e. $\mathbf{j}'\mathbf{j}$ f. $\mathbf{E}\mathbf{j}$ g. $\mathbf{E}^2$

4. Delle seguenti matrici, quali sono commutative rispetto alla moltiplicazione?

$$\mathbf{A} = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} \quad \mathbf{D} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

5. Calcolare i determinanti delle seguenti matrici usando metodi più opportuni:

a. $\begin{bmatrix} 4 & 2 & 0 \\ 5 & 3 & 0 \\ 6 & 9 & 2 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 0.8 & 0.5 \\ 0.8 & 1 & 0.6 \\ 0.5 & 0.6 & 1 \end{bmatrix}$ c. $\begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

d. $\begin{bmatrix} 1 & 4 & -1 \\ 3 & 12 & -3 \\ 0 & 35 & 7 \end{bmatrix}$ e. $\begin{bmatrix} 2 & 0 & 4 & 0 \\ 0 & 3 & 0 & 5 \\ 5 & 0 & 1 & 0 \\ 0 & 4 & 0 & 1 \end{bmatrix}$ f. $\begin{bmatrix} 2 & 0 & 1 & 1 & 1 \\ 0 & 2 & 3 & 3 & 3 \\ 1 & 3 & 1 & 0 & 0 \\ 1 & 3 & 0 & 1 & 0 \\ 1 & 3 & 0 & 0 & 1 \end{bmatrix}$

6. Per ognuna delle seguenti matrici non singolari, calcolarne l'inversa:

a. $\begin{bmatrix} 5 & 1 & -2 \\ 2 & 6 & 3 \\ -1 & 0 & 3 \end{bmatrix}$ b. $\begin{bmatrix} a & b & b \\ b & a & b \\ b & b & a \end{bmatrix}$ c. $\begin{bmatrix} 3 & 4 & 5 \\ 4 & 5 & 6 \\ 5 & 6 & 7 \end{bmatrix}$

d. $\begin{bmatrix} 5 & 0 & 0 \\ 0 & 8 & 6 \\ 0 & 6 & 5 \end{bmatrix}$ e. $\begin{bmatrix} 4 & 3 & 2 & 1 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

7. Usare operazioni elementari di riga o di colonna per calcolare il rango delle seguenti matrici:

a. $\begin{bmatrix} 1 & 0 & 2 & 1 \\ 1 & 1 & 2 & 0 \\ 1 & -1 & 2 & 2 \\ 1 & 1 & 2 & 0 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 0 & -1 & 3 & 1 \\ 2 & 1 & 1 & 0 & 1 \\ 0 & 3 & 8 & -5 & 3 \\ -1 & 2 & 6 & -2 & 3 \\ 1 & 1 & 2 & -3 & 0 \end{bmatrix}$

8. Determinare la soluzione del sistema di equazioni:

$$\begin{aligned} 2x_1 - x_2 + x_3 &= 2 \\ x_1 + 4x_2 - 2x_3 &= -1 \\ 2x_1 + 2x_2 + x_3 &= 0 \end{aligned}$$

9. Verificare che il sistema di equazioni omogenee

$$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ -x_1 + 2x_2 - x_3 &= 0 \\ 3x_1 + 3x_3 &= 0 \end{aligned}$$

ha una soluzione non banale e calcolare tutte le soluzioni linearmente indipendenti del sistema.

10. Determinare il rango del sistema di equazioni

$$\begin{aligned} 2x_1 + x_2 + x_3 &= 0 \\ x_1 + 2x_2 - x_3 + 3x_4 &= 0 \\ x_1 - x_2 + 3x_3 - 4x_4 &= 0 \\ 3x_2 - 4x_3 + 7x_4 &= 0 \end{aligned}$$

e calcolare la sua soluzione più generale.

11. Identificare i vettori dell'insieme

$$\mathbf{t}' = [1/2,\ 1/2,\ 1/2,\ 1/2]$$
$$\mathbf{u}' = [1,\ 0,\ -1,\ 0]$$
$$\mathbf{v}' = \left[\frac{\sqrt{2}}{2},\ 0,\ \frac{\sqrt{2}}{2},\ 0\right]$$

ortogonali o ortonormali, e calcolare l'angolo compreso tra ciascuna coppia di vettori.

12. Verificare che la matrice

$$\begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ -\frac{\sqrt{2}}{4} & \frac{\sqrt{6}}{4} & -\frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{4} & \frac{\sqrt{6}}{4} & \frac{\sqrt{2}}{2} \end{bmatrix}$$

È ortogonale.

13. Scrivere la matrice della forma quadratica:

$$2x_1^2 - 2x_1x_2 + x_2^2 + 4x_1x_3 - 3x_3^2$$

e determinare se è definita positiva.

14. Classificare le seguenti matrici come definite o semidefinite positive:

a. $\begin{bmatrix} 4 & 1 & 2 \\ 1 & 4 & -1 \\ 2 & -1 & 4 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$ c. $\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}$

15. Data la matrice:

$$\mathbf{A} = \begin{bmatrix} 2 & \sqrt{2} & 0 \\ \sqrt{2} & 2 & \sqrt{2} \\ 0 & \sqrt{2} & 2 \end{bmatrix}$$

calcolarne radici e vettori caratteristici, e determinare la matrice ortogonale dopo averla ridotta a forma diagonale.

16. Se

$$\mathbf{A} = \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 5 & 3 \\ 3 & 6 \end{bmatrix}$$

calcolare le somme e i prodotti delle radici caratteristiche di

a. $\mathbf{A}^{-1}\mathbf{B}$ b. $\mathbf{A}\mathbf{B}^{-1}$

17. Dovrebbe essere evidente che la somma dei quadrati

discussa nel paragrafo 2.9 assume valore minimo zero quando gli x_i sono eguali.

$$\sum_{i=1}^{N} (x_i - \bar{x})^2$$

Scrivere la somma in forma quadratica e calcolare il suo vettore-derivata per verificare quel valore minimo.

2.14 Riferimenti bibliografici

I testi di Browne [2], Hohn [8] e Perlis [14], servono sia come introduzione all'algebra delle matrici, sia per lo sviluppo matematico dell'argomento. Horst (9] ha tentato di sviluppare e di descrivere le operazioni matriciali in modo dettagliato e in termini più adatti a studiosi di sociologia. Frazer, Duncan e Collar [6] hanno invece trattato molti argomenti, in relazione sia all'algebra sia all'analisi numerica delle matrici.

Sebbene la maggior parte delle loro applicazioni riguardi la fisica e l'ingegneria, tali tecniche numeriche sono importanti anche per l'analisi multivariata.

La monografia di Roy [15] contiene nell'appendice un numero notevole di teoremi particolari sulle matrici ripartite, sulle forme quadratiche e sulle radici caratteristiche.

[1] Bodewig E.: *Matrix Calculus*, North-Holland Publishing Company, Amsterdam, 1956.

[2] Browne E. T.: *Introduction to the Theory of Determinants and Matrices*, The University of North Carolina Press, Chapel Hill, N.C., 1958.

[3] Dwyer P. S.: *Linear Computations*, John Wiley & Sons, Inc., New York, 1951.

[4] Dwyer P. S., and MacPhail M. S.: *Symbolic matrix derivatives*, in *Annals of Mathematical Statistics*, vol. 19 (1948), pp. 517-534.

[5] Faddeeva V. N.: *Computational Methods of Linear Algebra*, Dover Publications, Inc., New York, 1959.

[6] Frazer R. A., Duncan W. J., and Collar A. R.: *Elementary Matrices and Some Applications to Dynamics and Differential Equations*, Cambridge University Press, Cambridge, 1963.

[7] Graybill F. A.: *An Introduction to Linear Statistical Models*, vol. 1, McGraw-Hill Book Company, New York, 1961.

[8] Hohn F. E.: *Elementary Matrix Algebra*, 2d ed., The Macmillan Company, New York, 1964.

[9] Horst P.: *Matrix Algebra for Social Scientifich*, Holt, Rinehart and Winston, Inc., New York, 1963.

[10] Hotelling H.: *Some new methods for matrix calculation*, in *Annals of Mathematical Statistics*, vol. 14 (1943), pp. 1-33.

[11] Householder, A. S.: *Principles of Numerical Analysis*, McGraw-Hill Book Company, 1953.

[12] Macon N.: *Numerical Analysis*, John Wiley & Sons, Inc., New York, 1963.

[13] Marcus M. and Mine H.: *A Survey of Matrix Theory and Matrix Inequalities*, Allyn and Bacon, Inc., Boston, 1964.

[14] Perlis S.: *Theory of Matrices*, Addison-Wesley Publishing Company, Inc., Reading, Mass., 1952.

[15] Roy S. N.: *Some Aspects of Multivariate Analysis*, John Wiley & Sons, Inc., New York, 1957.