

Appunti di Algebra Lineare e Matrici

Basilio Bona
Dipartimento di Automatica e Informatica
Politecnico di Torino

Internal Report: DAUIN/BB-2003-09-01

Capitolo 1

Matrici e vettori

Il lettore interessato può fare riferimento a numerosi libri che trattano le matrici e l'algebra vettoriale; in lingua italiana posso suggerire i testi di base [8, 9], in inglese un classico per la teoria della matrici è rappresentato da [4], mentre per l'algebra lineare consiglio il testo di Strang [10]. Le lezioni videoregistrate di quest'ultimo sono visibili alla pagina Web del MIT OpenCourseWare, all'indirizzo [1].

1.1 Definizioni

Con il termine *matrice* si definisce un insieme composto da elementi ordinati in m righe e n colonne, che viene indicato da una delle notazioni seguenti:

$$\mathbf{A} = \mathbf{A}_{m \times n} = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Gli elementi a_{ij} possono essere variabili reali, quando $a_{ij} \in \mathbb{R}$ o variabili complesse, quando $a_{ij} \in \mathbb{C}$; allora si scrive $\mathbf{A} \in \mathbb{R}^{m \times n}$ oppure $\mathbf{A} \in \mathbb{C}^{m \times n}$. In questa Appendice consideriamo di solito matrici reali, salvo quando espressamente dichiarato. Se $n = 1$, abbiamo una matrice particolare, detta *vettore colonna* o semplicemente *vettore*.

Data una matrice $\mathbf{A}_{m \times n}$ si definisce *matrice trasposta* la matrice ottenuta scambiando le righe e le colonne

$$\mathbf{A}_{n \times m}^{\top} = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}$$

Vale la proprietà che $(\mathbf{A}^{\top})^{\top} = \mathbf{A}$.

Una matrice si dice quadrata se $m = n$. Se una matrice è quadrata, anche la sua trasposta è quadrata. Una matrice quadrata $n \times n$ si dice *triangolare superiore* se $a_{ij} = 0$ per $i > j$

$$\mathbf{A}_{n \times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

Se una matrice quadrata è triangolare superiore, la sua trasposta è *triangolare inferiore*

$$\mathbf{A}_{n \times n}^T = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{12} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix}$$

Se una matrice \mathbf{K} ha elementi complessi $k_{ij} = a_{ij} + jb_{ij}$, (dove $j = \sqrt{-1}$) si indica con $\overline{\mathbf{K}}$ la matrice coniugata, ossia quella che ha elementi $\overline{k}_{ij} = a_{ij} - jb_{ij}$.

Data una matrice complessa \mathbf{K} , si chiama *matrice aggiunta* \mathbf{K}^* la matrice trasposta coniugata, $\mathbf{K}^* = \overline{\mathbf{K}}^T = \overline{\mathbf{K}^T}$. Alcuni testi indicano questa matrice con il simbolo \mathbf{K}^\dagger oppure con il simbolo \mathbf{K}^H .

Una matrice reale quadrata si dice *simmetrica* se $\mathbf{A} = \mathbf{A}^T$, una matrice complessa si dice *autoaggiunta* o *hermitiana* se $\mathbf{K} = \mathbf{K}^*$. In una matrice reale simmetrica vi sono al più $\frac{n(n+1)}{2}$ elementi indipendenti.

Una matrice quadrata si dice *diagonale* se $a_{ij} = 0$ per $i \neq j$

$$\mathbf{A}_{n \times n} = \text{diag}(a_i) = \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{bmatrix}$$

Una matrice diagonale è sempre simmetrica.

Una matrice quadrata \mathbf{A} si dice *antisimmetrica* se $\mathbf{A} = -\mathbf{A}^T$; dati i vincoli imposti da questa relazione, la matrice antisimmetrica ha la seguente struttura

$$\mathbf{A}_{n \times n} = \begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ -a_{12} & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{1n} & -a_{2n} & \cdots & 0 \end{bmatrix}$$

In una matrice antisimmetrica vi sono al più $\frac{n(n-1)}{2}$ elementi indipendenti e non nulli. Vedremo in seguito alcune importanti proprietà delle matrici antisimmetriche 3×3 .

1.2 Operazioni sulle matrici

Le matrici sono elementi di un'algebra (*lineare*) detta *algebra della matrici*.

Ricordiamo che, in generale, un'algebra è uno *spazio lineare* (vettoriale) con l'aggiunta di un operatore (prodotto) bilineare.¹

Uno spazio lineare è una struttura matematica in cui sono definiti gli *elementi* dello spazio, che indicheremo, per il caso che stiamo trattando, con il simbolo maiuscolo grassetto \mathbf{A} , e alcune altre condizioni, qui di seguito elencate:

- 1) è definita un'operazione di *somma*, indicata con il simbolo $+$; la somma deve essere commutativa. Esiste un elemento *neutro* rispetto alla somma detto \mathbf{O} . Esso prende il nome di *elemento nullo* (relativamente alla somma).
- 2) per ogni elemento \mathbf{A} dello spazio, data una variabile α reale o complessa², esiste l'operazione di prodotto per α , tale che $\alpha\mathbf{A}$ appartiene ancora allo spazio. Inoltre, date due variabili scalari α e β ,
 - (a) vale la proprietà associativa rispetto al prodotto degli scalari: $\alpha(\beta\mathbf{A}) = (\alpha\beta)\mathbf{A}$;
 - (b) vale la proprietà distributiva rispetto alla somma: $\alpha(\mathbf{A} + \mathbf{B}) = \alpha\mathbf{A} + \alpha\mathbf{B}$;
 - (c) vale la proprietà distributiva rispetto al prodotto per scalare: $(\alpha + \beta)\mathbf{A} = \alpha\mathbf{A} + \beta\mathbf{A}$;

Nel caso particolare delle matrici queste proprietà generali prendono le forme particolari descritte nel seguito.

Prodotto per scalare

$$\alpha\mathbf{A} = \alpha \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{m1} & \alpha a_{m2} & \cdots & \alpha a_{mn} \end{bmatrix}$$

¹per *bilinearità* si intende la linearità rispetto a entrambi gli operandi.

²ci limiteremo a considerare il caso di α reale.

Somma di matrici

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{bmatrix}$$

Per poter essere sommate, le matrici devono avere le stesse dimensioni.

Valgono le seguenti proprietà, che sono state genericamente affermate nella condizione 1) precedente:

$$\begin{aligned} \mathbf{A} + \mathbf{O} &= \mathbf{A} \\ \mathbf{A} + \mathbf{B} &= \mathbf{B} + \mathbf{A} \\ (\mathbf{A} + \mathbf{B}) + \mathbf{C} &= \mathbf{A} + (\mathbf{B} + \mathbf{C}) \\ (\mathbf{A} + \mathbf{B})^\top &= \mathbf{A}^\top + \mathbf{B}^\top \end{aligned}$$

L'elemento neutro o nullo \mathbf{O} prende il nome di *matrice nulla*. L'operazione differenza viene definita con l'ausilio dello scalare $\alpha = -1$:

$$\mathbf{A} - \mathbf{B} = \mathbf{A} + (-1)\mathbf{B}.$$

Prodotto di matrici

Indicheremo per ora l'operatore prodotto con il simbolo \cdot , ma nell'uso comune esso viene quasi sempre omesso, come si usa fare per il simbolo di prodotto tra grandezze scalari.

L'operazione si effettua con la ben nota regola "*riga per colonna*": il generico elemento c_{ij} della matrice prodotto $\mathbf{C}_{m \times p} = \mathbf{A}_{m \times n} \cdot \mathbf{B}_{n \times p}$ vale

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$$

La proprietà di bilinearità del prodotto tra matrici è garantita, in quanto si verifica immediatamente che, dato uno scalare generico α , vale la seguente identità:

$$\alpha(\mathbf{A} \cdot \mathbf{B}) = (\alpha\mathbf{A}) \cdot \mathbf{B} = \mathbf{A} \cdot (\alpha\mathbf{B})$$

Valgono le seguenti proprietà, anch'esse genericamente fissate nelle condizioni 2) (a)-(c) precedenti:

$$\begin{aligned} \mathbf{A} \cdot \mathbf{B} \cdot \mathbf{C} &= (\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{C}) \\ \mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) &= \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C} \\ (\mathbf{A} + \mathbf{B}) \cdot \mathbf{C} &= \mathbf{A} \cdot \mathbf{C} + \mathbf{B} \cdot \mathbf{C} \\ (\mathbf{A} \cdot \mathbf{B})^\top &= \mathbf{B}^\top \cdot \mathbf{A}^\top \end{aligned}$$

In generale si verifica quanto segue:

- il prodotto tra matrici non è commutativo: $\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A}$, salvo in casi particolari;
- $\mathbf{A} \cdot \mathbf{B} = \mathbf{A} \cdot \mathbf{C}$ non implica $\mathbf{B} = \mathbf{C}$, salvo in casi particolari;
- $\mathbf{A} \cdot \mathbf{B} = \mathbf{O}$ non implica che sia $\mathbf{A} = \mathbf{O}$ oppure $\mathbf{B} = \mathbf{O}$, salvo in casi particolari.

Esiste un elemento neutro rispetto al prodotto, che prende il nome di *matrice identità* e viene indicata con \mathbf{I}_n oppure semplicemente \mathbf{I} quando non ci sono ambiguità nella dimensione; data una matrice rettangolare $\mathbf{A}_{m \times n}$ si ha

$$\mathbf{A}_{m \times n} = \mathbf{I}_m \mathbf{A}_{m \times n} = \mathbf{A}_{m \times n} \mathbf{I}_n.$$

Oltre a queste operazioni fondamentali, esistono altre funzioni su matrici che elencheremo brevemente nel seguito.

Potenza di matrice

Data una matrice quadrata $\mathbf{A} \in \mathbb{R}^{n \times n}$, la potenza k -esima di matrice vale

$$\mathbf{A}^k = \mathbf{A} \cdot \mathbf{A} \cdots \mathbf{A} \quad k \text{ volte}$$

Una matrice si dice *idempotente* se

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

Traccia

La traccia di una matrice quadrata $\mathbf{A}_{n \times n}$ è la somma dei suoi elementi diagonali

$$\text{tr}(\mathbf{A}) = \sum_{k=1}^n a_{kk}$$

La traccia di una matrice soddisfa le seguenti proprietà

$$\begin{aligned} \text{tr}(\alpha \mathbf{A} + \beta \mathbf{B}) &= \alpha \text{tr}(\mathbf{A}) + \beta \text{tr}(\mathbf{B}) \\ \text{tr}(\mathbf{AB}) &= \text{tr}(\mathbf{BA}) \\ \text{tr}(\mathbf{A}) &= \text{tr}(\mathbf{A}^\top) \\ \text{tr}(\mathbf{A}) &= \text{tr}(\mathbf{T}^{-1} \mathbf{A} \mathbf{T}) \quad \text{per } \mathbf{T} \text{ non singolare (vedi oltre)} \end{aligned}$$

Determinante

Data la matrice $\mathbf{A} \in \mathbb{R}^{n \times n}$, indichiamo con $\mathbf{A}^{(ij)}$ la matrice quadrata di dimensioni $(n-1) \times (n-1)$ ottenuta cancellando la i -esima riga e la j -esima colonna della matrice \mathbf{A} . Ad esempio, data la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & -5 & 3 & 2 \\ -6 & 4 & \boxed{9} & -7 \\ 7 & -4 & -8 & 2 \\ 0 & -9 & -2 & -3 \end{bmatrix}$$

si ha

$$\mathbf{A}^{(23)} = \begin{bmatrix} 1 & -5 & 2 \\ 7 & -4 & 2 \\ 0 & -9 & -3 \end{bmatrix}$$

Si definisce *minore* di ordine p di una matrice $\mathbf{A}_{m \times n}$ il determinante D_p di una sottomatrice quadrata ottenuta selezionando p righe e p colonne qualsiasi di $\mathbf{A}_{m \times n}$. Esistono tanti minori quante sono le scelte possibili di p su m righe e p su n colonne.

Si definiscono *minori principali di ordine k* di una matrice $\mathbf{A}_{m \times n}$ i determinanti D_k , con $k = 1, \dots, \min\{m, n\}$, ottenuti selezionando le prime k righe e k colonne della matrice $\mathbf{A}_{m \times n}$.

Si definisce *minore complementare* D_{rc} di un generico elemento a_{rc} di una matrice quadrata $\mathbf{A}_{n \times n}$ il determinante di $\mathbf{A}^{(rc)}$

$$D_{rc} = \det(\mathbf{A}^{(rc)}).$$

Si definisce *complemento algebrico* o *cofattore* (in inglese *cofactor*) di un elemento a_{rc} di una matrice quadrata $\mathbf{A}_{n \times n}$ il prodotto

$$A_{rc} = (-1)^{r+c} D_{rc}$$

Una volta definito il complemento algebrico si può definire il *determinante* di \mathbf{A} .

Fissata una qualsiasi riga i , si ha la definizione per riga:

$$\det(\mathbf{A}) = \sum_{k=1}^n a_{ik} (-1)^{i+k} \det(\mathbf{A}^{(ik)}) = \sum_{k=1}^n a_{ik} A_{ik}$$

oppure, fissata una qualsiasi colonna j , si ha la definizione per colonna:

$$\det(\mathbf{A}) = \sum_{k=1}^n a_{kj} (-1)^{k+j} \det(\mathbf{A}^{(kj)}) = \sum_{k=1}^n a_{kj} A_{kj}$$

Poiché le precedenti definizioni sono ricorsive e coinvolgono i determinanti di minori via via più piccoli, occorre definire il determinante della matrice 1×1 , che vale $\det(a_{ij}) = a_{ij}$.

In generale si hanno le proprietà seguenti:

- $\det(\mathbf{A} \cdot \mathbf{B}) = \det(\mathbf{A}) \det(\mathbf{B})$;
- $\det(\mathbf{A}^\top) = \det(\mathbf{A})$;
- $\det(k\mathbf{A}_{n \times n}) = k^n \det(\mathbf{A}_{n \times n})$;
- se si effettua un numero s di scambi tra righe o tra colonne della matrice \mathbf{A} ottenendo la matrice \mathbf{A}_s , si ha $\det(\mathbf{A}_s) = (-1)^s \det(\mathbf{A})$;
- se la matrice \mathbf{A} ha due righe o due colonne uguali o proporzionali, si ha $\det(\mathbf{A}) = 0$;
- se la matrice \mathbf{A} ha una riga o una colonna ottenibile da una combinazione lineare di altre righe o colonne, si ha $\det(\mathbf{A}) = 0$;
- se la matrice $\mathbf{A}_{n \times n}$ è triangolare superiore o inferiore, si ha $\det(\mathbf{A}_{n \times n}) = \prod_{i=1}^n a_{ii}$;
- se la matrice $\mathbf{A}_{n \times n}$ è triangolare a blocchi, con p blocchi \mathbf{A}_{ii} sulla diagonale, si ha $\det(\mathbf{A}_{n \times n}) = \prod_{i=1}^p \det \mathbf{A}_{ii}$;

Una matrice \mathbf{A} si dice *singolare* se $\det(\mathbf{A}) = 0$.

Rango

Si definisce *rango* (o *caratteristica*) della matrice $\mathbf{A}_{m \times n}$ il numero $\rho(\mathbf{A}_{m \times n})$ definito come il massimo intero p per cui esiste almeno un minore D_p non nullo.

Valgono le seguenti proprietà:

- $\rho(\mathbf{A}) \leq \min\{m, n\}$;
- se $\rho(\mathbf{A}) = \min\{m, n\}$, la matrice \mathbf{A} si dice *a rango pieno*;
- $\rho(\mathbf{A} \cdot \mathbf{B}) \leq \min\{\rho(\mathbf{A}), \rho(\mathbf{B})\}$.
- $\rho(\mathbf{A}) = \rho(\mathbf{A}^\top)$;
- $\rho(\mathbf{A} \cdot \mathbf{A}^\top) = \rho(\mathbf{A}^\top \cdot \mathbf{A}) = \rho(\mathbf{A})$;

D'ora in poi il prodotto tra matrici $\mathbf{A} \cdot \mathbf{B}$ sarà indicato semplicemente come \mathbf{AB} .

Matrice aggiunta

Data una matrice quadrata $\mathbf{A} \in \mathbb{R}^{n \times n}$, si definisce *matrice aggiunta* di \mathbf{A} la matrice quadrata $\mathbf{Adj}(\mathbf{A}) = \{\alpha_{ij}\}$ i cui elementi α_{ij} sono definiti come

$$\alpha_{ij} = (-1)^{i+j} D_{ji}$$

ossia quella matrice che ha come elemento di riga i e colonna j il minore complementare del corrispondente elemento a_{ji} di riga j e colonna i .

Matrice inversa

Data una matrice quadrata $\mathbf{A} \in \mathbb{R}^{n \times n}$ si dice *invertibile* o *non singolare* se esiste la matrice *inversa* $\mathbf{A}_{n \times n}^{-1}$ tale che

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

La matrice è invertibile se e solo se $\rho(\mathbf{A}) = n$, ossia è di rango pieno; ciò equivale ad avere $\det(\mathbf{A}) \neq 0$.

L'inversa si ottiene come:

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \text{Adj}(\mathbf{A})$$

Valgono le seguenti proprietà:

- $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$;
- $(\mathbf{A}^\top)^{-1} = (\mathbf{A}^{-1})^\top$.

Si definisce matrice *ortonormale* la matrice quadrata per cui $\mathbf{A}^{-1} = \mathbf{A}^\top$. Per queste matrici vale quindi l'identità

$$\mathbf{A}^\top \mathbf{A} = \mathbf{A} \mathbf{A}^\top = \mathbf{I}$$

Date due matrici quadrate di pari dimensioni \mathbf{A} e \mathbf{B} , vale la seguente identità

$$(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$

Esiste un importante risultato, chiamato *Lemma d'inversione*, che stabilisce quanto segue: se \mathbf{A} e \mathbf{C} sono matrici quadrate invertibili e \mathbf{B} e \mathbf{D} sono matrici di dimensioni opportune, allora

$$(\mathbf{A} + \mathbf{B}\mathbf{C}\mathbf{D})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{D}\mathbf{A}^{-1}\mathbf{B} + \mathbf{C}^{-1})^{-1}\mathbf{D}\mathbf{A}^{-1}$$

La matrice $(\mathbf{D}\mathbf{A}^{-1}\mathbf{B} + \mathbf{C}^{-1})$ deve essere anch'essa invertibile.

Se la matrice quadrata $\mathbf{A}(t)$ è composta da elementi $a_{ij}(t)$ tutti derivabili nel tempo t , allora la derivata della matrice vale

$$\frac{d}{dt}\mathbf{A}(t) = \dot{\mathbf{A}}(t) = \left[\frac{d}{dt}a_{ij}(t) \right] = [\dot{a}_{ij}(t)]$$

Se la matrice quadrata $\mathbf{A}(t)$ ha rango $\rho(\mathbf{A}(t)) = n$ per ogni valore del tempo t , allora la derivata della sua inversa vale

$$\frac{d}{dt}\mathbf{A}(t)^{-1} = -\mathbf{A}^{-1}(t)\dot{\mathbf{A}}(t)\mathbf{A}(t)^{-1}$$

La matrice inversa, quando esiste, permette risolvere la trasformazione lineare

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

in funzione dell'incognita \mathbf{x} , come

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}.$$

Decomposizione di matrice

Data una matrice quadrata \mathbf{A} , è sempre possibile decomporla in una somma di due matrici, come segue:

$$\mathbf{A} = \mathbf{A}_s + \mathbf{A}_a \quad (1.1)$$

dove

$$\mathbf{A}_s = \frac{1}{2}(\mathbf{A} + \mathbf{A}^\top)$$

è una matrice simmetrica e

$$\mathbf{A}_a = \frac{1}{2}(\mathbf{A} - \mathbf{A}^\top)$$

è una matrice antisimmetrica (vedi Sezione 1.4).

Data una matrice reale di dimensioni qualsiasi $\mathbf{A} \in \mathbb{R}^{m \times n}$, risultano simmetriche entrambe le matrici seguenti

$$\begin{aligned} \mathbf{A}^\top \mathbf{A} &\in \mathbb{R}^{n \times n} \\ \mathbf{A} \mathbf{A}^\top &\in \mathbb{R}^{m \times m} \end{aligned}$$

Trasformazioni di similarità

Data una matrice quadrata $\mathbf{A} \in \mathbb{R}^{n \times n}$ e una matrice quadrata non singolare $\mathbf{T} \in \mathbb{R}^{n \times n}$, la matrice $\mathbf{B} \in \mathbb{R}^{n \times n}$ ottenuta come

$$\mathbf{B} = \mathbf{T}^{-1} \mathbf{A} \mathbf{T} \quad \text{oppure} \quad \mathbf{B} = \mathbf{T} \mathbf{A} \mathbf{T}^{-1}$$

si dice *similare* ad \mathbf{A} e la trasformazione si dice *di similarità*. Se la matrice \mathbf{A} è simile alla matrice diagonale $\mathbf{\Lambda} = \text{diag}(\lambda_i)$

$$\mathbf{A} = \mathbf{T} \mathbf{\Lambda} \mathbf{T}^{-1}$$

si può scrivere

$$\mathbf{A} \mathbf{T} = \mathbf{T} \mathbf{\Lambda}$$

e se indichiamo con \mathbf{t}_i la i -esima colonna di \mathbf{T} , ossia

$$\mathbf{T} = [\mathbf{t}_1 \quad \cdots \quad \mathbf{t}_n]$$

avremo la nota formula che lega autovalori e autovettori (vedi Paragrafo 1.2.1)

$$\mathbf{A} \mathbf{t}_i = \lambda_i \mathbf{t}_i$$

e quindi potremo dire che le costanti λ_i sono gli autovalori di \mathbf{A} e i vettori \mathbf{t}_i sono gli autovettori di \mathbf{A} , in generale non normalizzati.

1.2.1 Autovalori e autovettori

Data una matrice quadrata $\mathbf{A}_{n \times n}$, si chiamano *autovalori della matrice* (in inglese *eigenvalue*) le soluzioni λ_i (reali o complesse) dell'equazione caratteristica

$$\det(\lambda \mathbf{I} - \mathbf{A}) = 0$$

$\det(\lambda \mathbf{I} - \mathbf{A})$ è un polinomio in λ , detto *polinomio caratteristico*.

Se gli autovalori sono tutti distinti, si chiamano *autovettori* (in inglese *eigenvector*) i vettori \mathbf{u}_i che soddisfano l'identità

$$\mathbf{A}\mathbf{u}_i = \lambda_i\mathbf{u}_i$$

Se gli autovalori non sono tutti distinti, si ottengono *autovettori generalizzati*, la cui determinazione va oltre gli scopi di questa Appendice.

Geometricamente gli autovettori rappresentano quelle particolari “direzioni” nello spazio \mathbb{R}^n , in cui si applica la trasformazione lineare rappresentata da \mathbf{A} , che si trasformano in sé stesse; sono quindi le direzioni invarianti rispetto alla trasformazione \mathbf{A} e gli autovalori forniscono le rispettive costanti di “scalamento” lungo queste direzioni.

L'insieme degli autovalori di una matrice \mathbf{A} sarà indicato con $\Lambda(\mathbf{A})$ oppure con $\{\lambda_i(\mathbf{A})\}$; l'insieme degli autovettori di \mathbf{A} sarà indicato con $\{\mathbf{u}_i(\mathbf{A})\}$. In generale, essendo gli autovettori delle rappresentazioni di direzioni invarianti rispetto alla trasformazione rappresentata da \mathbf{A} , essi sono definiti a meno di una costante, ossia possono o meno essere normalizzati; tuttavia è convenzione tacita che essi abbiano norma unitaria, salvo quando altrimenti dichiarato.

Proprietà degli autovalori

Data una matrice \mathbf{A} e i suoi autovalori $\{\lambda_i(\mathbf{A})\}$, sarà

$$\{\lambda_i(\mathbf{A} + c\mathbf{I})\} = \{(\lambda_i(\mathbf{A}) + c)\}$$

Data una matrice \mathbf{A} e i suoi autovalori $\{\lambda_i(\mathbf{A})\}$, sarà

$$\{\lambda_i(c\mathbf{A})\} = \{(c\lambda_i(\mathbf{A}))\}$$

Data una matrice triangolare (superiore o inferiore)

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}, \quad \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

i suoi autovalori sono gli elementi sulla diagonale $\{\lambda_i(\mathbf{A})\} = \{a_{ii}\}$; lo stesso vale per una matrice diagonale.

Data una matrice $\mathbf{A}_{n \times n}$ e i suoi autovalori $\{\lambda_i(\mathbf{A})\}$, sarà

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

e

$$\text{tr}(\mathbf{A}) = \sum_{i=1}^n \lambda_i$$

Data una qualunque trasformazione invertibile, rappresentata dalla matrice \mathbf{T} , gli autovalori di \mathbf{A} sono invarianti alle *trasformazioni di similarità*

$$\tilde{\mathbf{A}} = \mathbf{T}^{-1} \mathbf{A} \mathbf{T}$$

ossia

$$\{\lambda_i(\tilde{\mathbf{A}})\} = \{\lambda_i(\mathbf{A})\}$$

Se costruiamo una matrice di trasformazione \mathbf{M} ordinando per colonne gli autovettori normalizzati $\mathbf{u}_i(\mathbf{A})$

$$\mathbf{M} = [\mathbf{u}_1 \quad \cdots \quad \mathbf{u}_n]$$

allora la trasformazione di similarità fornisce una matrice diagonale

$$\mathbf{A} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \mathbf{M}^{-1} \mathbf{A} \mathbf{M}$$

La matrice \mathbf{M} si chiama *matrice modale*.

Se la matrice \mathbf{A} è simmetrica, i suoi autovalori sono tutti reali e si ha l'identità

$$\mathbf{A} = \mathbf{M}^T \mathbf{A} \mathbf{M}$$

In questo caso la matrice \mathbf{M} è ortonormale.

1.2.2 Decomposizione ai valori singolari

Data una matrice $\mathbf{A} \in \mathbb{R}^{m \times n}$ qualsiasi, di rango $r = \rho(\mathbf{A}) \leq s$ con $s = \min\{m, n\}$, essa si può fattorizzare secondo la *decomposizione ai valori singolari*, come segue:

$$\mathbf{A} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \quad (1.2)$$

dove:

- \mathbf{U} è una matrice $(m \times m)$ ortonormale

$$\mathbf{U} = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_m]; \quad \mathbf{U}\mathbf{U}^\top = \mathbf{I}_m$$

contenente per colonne gli autovettori \mathbf{u}_i della matrice $\mathbf{A}\mathbf{A}^\top$

- \mathbf{V} è una matrice $n \times n$ ortonormale

$$\mathbf{V} = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_n]; \quad \mathbf{V}\mathbf{V}^\top = \mathbf{I}_n$$

contenente per colonne gli autovettori \mathbf{v}_i della matrice $\mathbf{A}^\top\mathbf{A}$

- $\mathbf{\Sigma}$ è una matrice $(m \times n)$ con la seguente struttura

$$\text{se } m < n \quad \mathbf{\Sigma} = [\mathbf{\Sigma}_s \ \mathbf{O}]$$

$$\text{se } m = n \quad \mathbf{\Sigma} = \mathbf{\Sigma}_s$$

$$\text{se } m > n \quad \mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_s \\ \mathbf{O} \end{bmatrix}.$$

La matrice $\mathbf{\Sigma}_s = \text{diag}(\sigma_i)$ è diagonale di dimensioni $s \times s$ e contiene sulla diagonale i valori singolari, definiti come segue.

- $\sigma_i(\mathbf{A}) \geq 0$ sono detti *valori singolari* e coincidono con le radici quadrate non negative degli autovalori della matrice simmetrica $\mathbf{A}^\top\mathbf{A}$:

$$\sigma_i(\mathbf{A}) = \sqrt{\lambda_i(\mathbf{A}^\top\mathbf{A})} = \sqrt{\lambda_i(\mathbf{A}\mathbf{A}^\top)} \quad \sigma_i \geq 0$$

ordinati in ordine decrescente

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_s \geq 0$$

se $r < s$ vi sono r valori singolari positivi, mentre i restanti $s - r$ sono nulli:

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0; \quad \sigma_{r+1} = \cdots = \sigma_s = 0$$

Alternativamente, possiamo descrivere la decomposizione della matrice \mathbf{A} nel modo seguente, che è del tutto analogo a quello dato in (1.2), ma mette in evidenza i soli valori singolari positivi:

$$\mathbf{A} = \underbrace{[\mathbf{P} \ \tilde{\mathbf{P}}]}_{\mathbf{U}} \underbrace{\begin{bmatrix} \mathbf{\Sigma}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}}_{\mathbf{\Sigma}} \underbrace{\begin{bmatrix} \mathbf{Q}^\top \\ \tilde{\mathbf{Q}}^\top \end{bmatrix}}_{\mathbf{V}^\top} = \mathbf{P}\mathbf{\Sigma}_r\mathbf{Q}^\top = \sum_{i=1}^s \sigma_i \mathbf{p}_i \mathbf{q}_i^\top \quad (1.3)$$

dove

- \mathbf{P} è una matrice ortonormale $m \times r$, $\tilde{\mathbf{P}}$ è una matrice ortonormale $m \times (m-r)$;
- \mathbf{Q} è una matrice ortonormale $n \times r$, \mathbf{Q}^\top è una matrice ortonormale $n \times (n-r)$;
- $\mathbf{\Sigma}_r$ è una matrice diagonale $r \times r$ che contiene sulla diagonale i soli valori singolari positivi σ_i , $i = 1, \dots, r$.

Il rango r della matrice \mathbf{A} è pari al numero $r \leq s$ di valori singolari non nulli.

Un altro modo ancora di definire la decomposizione SVD di $\mathbf{A}_{m \times n}$ è il seguente:

$$\mathbf{A}_{m \times n} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^\top \quad (1.4)$$

dove ora

- \mathbf{U} è una matrice ($m \times n$) ortonormale

$$\mathbf{U} = [\mathbf{u}_1 \quad \dots \quad \mathbf{u}_n]$$

I vettori colonna \mathbf{u}_i di \mathbf{U} sono chiamati *vettori singolari sinistri* di \mathbf{A} , e corrispondono agli n autovettori della matrice $\mathbf{A}^\top \mathbf{A}$, cioè $\mathbf{A}^\top \mathbf{A} \mathbf{u}_i = \sigma_i^2 \mathbf{u}_i$.

- \mathbf{V} è una matrice $n \times n$ ortonormale

$$\mathbf{V} = [\mathbf{v}_1 \quad \dots \quad \mathbf{v}_n]$$

I vettori colonna \mathbf{v}_i di \mathbf{V} sono chiamati *vettori singolari destri* di \mathbf{A} , e corrispondono agli m autovettori della matrice $\mathbf{A} \mathbf{A}^\top$, cioè $\mathbf{A} \mathbf{A}^\top \mathbf{v}_i = \sigma_i^2 \mathbf{v}_i$.

- $\mathbf{\Sigma}$ è una matrice ($n \times n$) diagonale $\mathbf{\Sigma} = \text{diag}(\sigma_i)$ e contiene sulla diagonale i valori singolari, definiti come descritto sopra.

Si fa notare che:

- La decomposizione singolare esiste sempre ed è unica, a meno
 1. di identiche permutazioni delle colonne di \mathbf{U} , \mathbf{V} , $\mathbf{\Sigma}$,
 2. combinazioni lineari di colonne di \mathbf{U} e \mathbf{V} con uguali valori singolari.
- data una matrice $\mathbf{A} \in \mathbb{R}^{m \times n}$ qualsiasi, le due matrici $\mathbf{A}^\top \mathbf{A}$ e $\mathbf{A} \mathbf{A}^\top$ sono simmetriche, hanno gli stessi valori singolari positivi e differiscono soltanto per il numero di valori singolari nulli.

Altre proprietà della decomposizione singolare

- le colonne \mathbf{u}_i che corrispondono a valori singolari positivi sono una base per lo spazio immagine (*range*) di \mathbf{A} .
- le colonne \mathbf{v}_i che corrispondono a valori singolari positivi sono una base per lo spazio nullo (*kernel* o *null-space*) di \mathbf{A} .

1.3 Vettori e spazi vettoriali

Un vettore può essere semplicemente interpretato come una particolare matrice colonna di dimensioni $n \times 1$:

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

Esso gioca un ruolo fondamentale in molti campi dell'ingegneria, dalla modellistica dei sistemi dinamici alla cinematica dei corpi rigidi nello spazio tridimensionale, dall'elettromagnetismo alla statica – per citarne solo alcuni – permettendo di passare agevolmente dalla rappresentazione algebrica di un fenomeno alla sua rappresentazione geometrica e viceversa.

In generale, un vettore è un elemento di uno spazio vettoriale. Nella Sezione 1.2 abbiamo già fatto conoscenza con gli spazi vettoriali delle matrici. Ora ne daremo una definizione più completa.

1.3.1 Spazio vettoriale

Come detto sopra, gli elementi di uno spazio vettoriale rappresentano entità assai utili nello studio di molti settori dell'ingegneria e della fisica classica e moderna [3].

Dato un campo³ qualsiasi \mathcal{F} , lo *spazio vettoriale* (in inglese *vector space*) $\mathbb{V}(\mathcal{F})$, è l'insieme di quegli elementi, chiamati *vettori*, che soddisfano le seguenti proprietà assiomatiche:

1. esiste l'operazione $+$, detta *somma vettoriale*, tale che $\{\mathbb{V}(\mathcal{F}); +\}$ forma un gruppo abeliano; l'elemento identità è chiamato $\mathbf{0}$;
2. per ogni $\alpha \in \mathcal{F}$ e ogni $\mathbf{v} \in \mathbb{V}(\mathcal{F})$, esiste un vettore $\alpha\mathbf{v} \in \mathbb{V}(\mathcal{F})$; inoltre per ogni $\alpha, \beta \in \mathcal{F}$ e ogni $\mathbf{v}, \mathbf{w} \in \mathbb{V}(\mathcal{F})$ valgono le seguenti proprietà:

$$\begin{aligned} \text{associativa rispetto al prodotto per scalare:} & \quad \alpha(\beta\mathbf{v}) = (\alpha\beta)\mathbf{v} \\ \text{identità rispetto al prodotto per scalare:} & \quad 1(\mathbf{v}) = \mathbf{v}; \quad \forall \mathbf{v} \\ \text{distributiva rispetto alla somma vettoriale:} & \quad \alpha(\mathbf{v} + \mathbf{w}) = \alpha\mathbf{v} + \alpha\mathbf{w} \\ \text{distributiva rispetto al prodotto per scalare:} & \quad (\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v} \end{aligned} \quad (1.5)$$

Uno spazio vettoriale è detto *reale* se $\mathcal{F} = \mathbb{R}$, è detto *complesso* se $\mathcal{F} = \mathbb{C}$.

³per le definizioni di *campo* e *gruppo*, vedere [2, Appendice A]

Esempio classico di spazio vettoriale reale è quello rappresentato da n -ple di reali, $\mathbb{V}_n(\mathbb{R}) = \mathbb{R}^n$; in questi casi un elemento (*vettore*) viene rappresentato *per componenti*

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}, \quad \mathbf{v} \in \mathbb{R}^n, \quad v_i \in \mathbb{R}$$

Poiché le proprietà (1.5) inducono una struttura *lineare* sullo spazio \mathbb{V} , esso viene indicato anche con il termine di *spazio vettoriale lineare* o semplicemente *spazio lineare* (in inglese *linear vector space* o semplicemente *linear space*).

Il termine “vettore” ha un significato molto ampio, nel senso che essa include entità matematiche in apparenza anche molto diverse. Ad esempio, non solo le n -ple di reali, ma anche le sequenze infinite di numeri reali o complessi, le funzioni continue che assumono valori reali nell’intervallo $[a, b]$, i polinomi a coefficienti complessi definiti nell’intervallo $[a, b]$ ecc. [7].

Come si può notare, tra gli assiomi dello spazio vettoriale non compare alcuna operazione di prodotto.

Per questo motivo la *struttura* dello spazio vettoriale, ossia l’insieme di proprietà che derivano dagli assiomi, non permette di definire concetti geometrici quali l’*angolo* o la *distanza*, che invece sono impliciti nella definizione puramente geometrica di vettore. Per consentire di definire tali concetti è necessario dotare lo spazio vettoriale di una *struttura quadratica* o *metrica*. L’introduzione di una metrica in uno spazio vettoriale genera un’algebra che rende possibile l’esecuzione di calcoli su oggetti geometrici. La metrica più comune è quella indotta dalla definizione di prodotto scalare.

Prima di passare alla definizione di questo prodotto, riassumiamo brevemente alcune proprietà delle funzioni lineari.

1.3.2 Funzioni lineari

Dati due spazi vettoriali $\mathbb{U}(\mathcal{F})$ e $\mathbb{V}(\mathcal{F})$, che per comodità assumiamo definiti entrambi sul campo \mathcal{F} , una funzione $\mathcal{L} : \mathbb{U} \rightarrow \mathbb{V}$ si dice *lineare*, se per ogni $\mathbf{a}, \mathbf{b} \in \mathbb{U}$ e $\lambda \in \mathcal{F}$ valgono i seguenti assiomi:

$$\begin{aligned} \mathcal{L}(\mathbf{a} + \mathbf{b}) &= \mathcal{L}(\mathbf{a}) + \mathcal{L}(\mathbf{b}) = \mathcal{L}\mathbf{a} + \mathcal{L}\mathbf{b} \\ \mathcal{L}(\lambda\mathbf{a}) &= \lambda\mathcal{L}(\mathbf{a}) = \lambda\mathcal{L}\mathbf{a} \end{aligned} \tag{1.6}$$

La funzione lineare $\mathcal{L} : \mathbb{U} \rightarrow \mathbb{V}$ viene anche chiamata *operatore lineare*, *trasformazione lineare*, *applicazione lineare* oppure *endomorfismo* (in inglese *endomorphism*).

L’insieme di tutte le funzioni lineari $\mathcal{L} : \mathbb{U} \rightarrow \mathbb{V}$ forma, a sua volta, uno spazio lineare vettoriale $\mathbb{L}(\mathcal{F})$.

L'insieme delle funzioni lineari $\mathcal{L} : \mathbb{U} \rightarrow \mathbb{U}$ forma un anello, indicato con il simbolo $\text{End}(\mathbb{U})$.

Ricordiamo infine che qualsiasi funzione o trasformazione lineare da \mathbb{U} a \mathbb{V} è rappresentabile con una matrice quadrata $\mathbf{A} \in \mathbb{R}^{m \times n}$, dove m e n sono le dimensioni (vedere più oltre la definizione di dimensione) rispettivamente di \mathbb{V} e \mathbb{U} .

Indipendenza lineare – Base – Dimensione

Dati n vettori qualsiasi $\mathbf{a}_i \in \mathbb{V}(\mathcal{F})$, un vettore generico $\mathbf{v} \in \mathbb{V}(\mathcal{F})$ è detto *combinazione lineare* di $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ se esso può essere scritto come

$$\mathbf{v} = v_1 \mathbf{a}_1 + v_2 \mathbf{a}_2 + \dots + v_n \mathbf{a}_n$$

con $v_i \in \mathcal{F}$. L'insieme di vettori $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ è detto *linearmente indipendente* se nessun \mathbf{a}_i può essere scritto come combinazione lineare dei restanti \mathbf{a}_j , $j \neq i$. In altre parole, l'unica soluzione dell'equazione

$$v_1 \mathbf{a}_1 + v_2 \mathbf{a}_2 + \dots + v_n \mathbf{a}_n = \mathbf{0}$$

è quella con $v_1 = v_2 = \dots = v_n = 0$. In caso contrario si dice che \mathbf{a}_i è *linearmente dipendente* dagli altri vettori.

Nella combinazione lineare $\mathbf{v} = v_1 \mathbf{a}_1 + v_2 \mathbf{a}_2 + \dots + v_n \mathbf{a}_n$, se tutti i vettori \mathbf{a}_i sono linearmente indipendenti, allora gli scalari v_i sono unici e prendono il nome di *coordinate* o *componenti* di \mathbf{v} .

Le combinazioni lineari di vettori linearmente indipendenti $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k\}$, con $k \leq n$, formano un sottospazio $\mathbb{S}(\mathcal{F}) \subseteq \mathbb{V}(\mathcal{F})$. Si dice che questo sottospazio è *coperto* o *descritto* (in inglese *spanned*) da $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k\}$.

Ogni insieme di vettori $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ che risulti linearmente indipendente, forma una *base* in \mathbb{V} . Tutte le basi in \mathbb{V} hanno lo stesso numero di elementi (nel nostro caso n), che prende il nome di *dimensione* dello spazio e si indica con $\dim(\mathbb{V})$.

1.3.3 Matrici come rappresentazione di operatori lineari

Dati due spazi vettoriali $\mathcal{X} \subseteq \mathbb{R}^n$ e $\mathcal{Y} \subseteq \mathbb{R}^m$, aventi rispettivamente dimensioni n e m , e dati due generici vettori $\mathbf{x} \in \mathcal{X}$ e $\mathbf{y} \in \mathcal{Y}$, la più generica trasformazione lineare tra gli spazi si può rappresentare attraverso l'operatore matriciale $\mathbf{A} \in \mathbb{R}^{m \times n}$, come segue:

$$\mathbf{y} = \mathbf{A}\mathbf{x}; \quad \mathbf{x} \in \mathbb{R}^n; \quad \mathbf{y} \in \mathbb{R}^m.$$

Quindi una matrice può essere sempre interpretata come un operatore che prende un vettore dello spazio di “partenza” \mathcal{X} e lo trasforma in un vettore dello spazio di “arrivo” \mathcal{Y} . Qualunque trasformazione lineare ha (almeno) una matrice che la

rappresenta e, di converso, qualunque matrice è la rappresentazione di una qualche trasformazione lineare.

Si definisce *spazio immagine* (in inglese *range*) della trasformazione \mathbf{A} il sottospazio di \mathcal{Y} definito dalla seguente proprietà:

$$\mathcal{R}(\mathbf{A}) = \{\mathbf{y} \mid \mathbf{y} = \mathbf{A}\mathbf{x}, \mathbf{x} \in \mathcal{X}\}; \quad \mathcal{R}(\mathbf{A}) \subseteq \mathcal{Y}$$

Si definisce *spazio nullo* (in inglese *kernel* o *null-space*) della trasformazione \mathbf{A} il sottospazio di \mathcal{X} definito dalla seguente proprietà:

$$\mathcal{N}(\mathbf{A}) = \{\mathbf{x} \mid \mathbf{0} = \mathbf{A}\mathbf{x}, \mathbf{x} \in \mathcal{X}\}; \quad \mathcal{N}(\mathbf{A}) \subseteq \mathcal{X}$$

Lo spazio nullo rappresenta cioè tutti quei vettori di \mathcal{X} che vengono trasformati nel vettore nullo (l'origine) di \mathcal{Y} .

Le dimensioni dello spazio immagine e dello spazio nullo si chiamano, rispettivamente, *rango* $\rho(\mathbf{A})$ (che abbiamo già definito in 1.2) e *nullità* $\nu(\mathbf{A})$:

$$\rho(\mathbf{A}) = \dim(\mathcal{R}(\mathbf{A})); \quad \nu(\mathbf{A}) = \dim(\mathcal{N}(\mathbf{A})).$$

Se \mathcal{X} e \mathcal{Y} hanno dimensioni finite, e questo è il nostro caso in quanto $\mathcal{X} \subseteq \mathbb{R}^n$ e $\mathcal{Y} \subseteq \mathbb{R}^m$, allora valgono le seguenti relazioni:

$$\begin{aligned} \mathcal{N}(\mathbf{A}) &= \mathcal{R}(\mathbf{A}^\top)^\perp \\ \mathcal{R}(\mathbf{A}) &= \mathcal{N}(\mathbf{A}^\top)^\perp \\ \mathcal{N}(\mathbf{A})^\perp &= \mathcal{R}(\mathbf{A}^\top) \\ \mathcal{R}(\mathbf{A})^\perp &= \mathcal{N}(\mathbf{A}^\top) \end{aligned}$$

dove il simbolo \perp indica il *complemento ortogonale* al (sotto-)spazio corrispondente. Ricordiamo che $\{\mathbf{0}\}^\perp = \mathbb{R}$.

Vale anche la seguente decomposizione ortogonale degli spazi \mathcal{X} e \mathcal{Y} (vedere [9] e [10])

$$\begin{aligned} \mathcal{X} &= \mathcal{N}(\mathbf{A}) \oplus \mathcal{N}(\mathbf{A})^\perp = \mathcal{N}(\mathbf{A}) \oplus \mathcal{R}(\mathbf{A}^\top) \\ \mathcal{Y} &= \mathcal{R}(\mathbf{A}) \oplus \mathcal{R}(\mathbf{A})^\perp = \mathcal{R}(\mathbf{A}) \oplus \mathcal{N}(\mathbf{A}^\top) \end{aligned}$$

dove il simbolo \oplus indica la *somma diretta* tra due sottospazi.

1.3.4 Inversa generalizzata

Data una matrice reale qualsiasi $\mathbf{A} \in \mathbb{R}^{m \times n}$, con $m \neq n$, la matrice inversa non risulta definita. Tuttavia, è possibile definire una classe di matrici, dette *pseudo-inverse*, *inverse generalizzate* o *1-inverse* \mathbf{A}^- , che soddisfano la seguente relazione:

$$\mathbf{A}\mathbf{A}^-\mathbf{A} = \mathbf{A}$$

Se la matrice \mathbf{A} ha rango pieno, ossia $\rho(\mathbf{A}) = \min\{m, n\}$, è possibile definire due classi di matrici inverse generalizzate particolari

- se $m < n$ (ossia $\rho(\mathbf{A}) = m$), l'inversa destra di \mathbf{A} è quella matrice $\mathbf{A}_d \in \mathbb{R}^{n \times m}$ per cui

$$\mathbf{A}\mathbf{A}_d = \mathbf{I}_{m \times m}$$

- se $n < m$ (ossia $\rho(\mathbf{A}) = n$), l'inversa sinistra di \mathbf{A} è quella matrice $\mathbf{A}_s \in \mathbb{R}^{n \times m}$ per cui

$$\mathbf{A}_s\mathbf{A} = \mathbf{I}_{n \times n}$$

Tra le molte inverse destre e sinistre concepibili, due sono particolarmente importanti:

- *pseudo-inversa destra* ($m < n$):

$$\mathbf{A}_d^+ = \mathbf{A}^\top (\mathbf{A}\mathbf{A}^\top)^{-1}$$

che rappresenta una particolare inversa destra. Si può dimostrare che quando $\rho(\mathbf{A}) = m$ allora $(\mathbf{A}\mathbf{A}^\top)^{-1}$ esiste.

- *pseudo-inversa sinistra* ($n < m$):

$$\mathbf{A}_s^+ = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top$$

che rappresenta una particolare inversa sinistra. Si può dimostrare che quando $\rho(\mathbf{A}) = n$ allora $(\mathbf{A}^\top \mathbf{A})^{-1}$ esiste.

Questa particolare pseudo-inversa sinistra prende anche il nome di *pseudo-inversa di Moore-Penrose*. In generale, anche se $\mathbf{A}^\top \mathbf{A}$ non risulta invertibile, si può sempre definire una pseudo-inversa di Moore-Penrose, che è quella matrice \mathbf{A}^+ che soddisfa le seguenti relazioni:

$$\begin{aligned} \mathbf{A}\mathbf{A}^+\mathbf{A} &= \mathbf{A} \\ \mathbf{A}^+\mathbf{A}\mathbf{A}^+ &= \mathbf{A}^+ \\ (\mathbf{A}\mathbf{A}^+)^\top &= \mathbf{A}\mathbf{A}^+ \\ (\mathbf{A}^+\mathbf{A})^\top &= \mathbf{A}^+\mathbf{A} \end{aligned} \tag{1.7}$$

Queste due pseudo-inverse coincidono con la matrice inversa quando \mathbf{A} è quadrata e ha rango pieno:

$$\mathbf{A}^{-1} = \mathbf{A}_d^+ = \mathbf{A}_s^+$$

La trasformazione lineare associata alla matrice $\mathbf{A} \in \mathbb{R}^{m \times n}$

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \tag{1.8}$$

con $\mathbf{x} \in \mathbb{R}^n$ e $\mathbf{y} \in \mathbb{R}^m$, è equivalente ad un sistema di m equazioni lineari in n incognite, i cui coefficienti sono dati dagli elementi di \mathbf{A} ; questo sistema lineare può non ammettere soluzioni, ammetterne una sola o ammetterne un numero infinito [10].

Se vogliamo utilizzare le pseudo-inverse definite sopra per risolvere il sistema lineare in (1.8), dobbiamo distinguere due casi, sempre nell'ipotesi che il rango di \mathbf{A} sia pieno:

- $m < n$: abbiamo più incognite che equazioni; allora, tra le infinite soluzioni possibili $\mathbf{x} \in \mathbb{R}^n$, scegliamo convenzionalmente quella che ha norma $\|\mathbf{x}\|$ minima, che viene individuata da

$$\tilde{\mathbf{x}} = \mathbf{A}_d^+ \mathbf{y} = \mathbf{A}^\top (\mathbf{A} \mathbf{A}^\top)^{-1} \mathbf{y}$$

Tutte le altre possibili soluzioni di $\mathbf{y} = \mathbf{A} \mathbf{x}$ si ottengono come

$$\bar{\mathbf{x}} = \tilde{\mathbf{x}} + \mathbf{v} = \mathbf{A}_d^+ \mathbf{y} + \mathbf{v}$$

dove $\mathbf{v} \in \mathcal{N}(\mathbf{A})$ è un vettore dello spazio nullo di \mathbf{A} , che ha dimensioni $n - m$. Queste possibili soluzioni si possono anche esprimere in una forma alternativa, ovvero

$$\bar{\mathbf{x}} = \mathbf{A}_d^+ \mathbf{y} + (\mathbf{I} - \mathbf{A}_d^+ \mathbf{A}) \mathbf{w}$$

dove $\mathbf{w} \in \mathbb{R}^n$ è un vettore $n \times 1$ qualsiasi. La matrice $\mathbf{I} - \mathbf{A}_d^+ \mathbf{A}$ proietta \mathbf{w} nello spazio nullo di \mathbf{A} , trasformandolo in $\mathbf{v} \in \mathcal{N}(\mathbf{A})$; essa prende il nome di *matrice di proiezione*.

- $n < m$: abbiamo più equazioni che incognite; allora non esistono soluzioni esatte alla $\mathbf{y} = \mathbf{A} \mathbf{x}$, ma solo soluzioni approssimate, tali per cui esiste un errore $\mathbf{e} = \mathbf{A} \mathbf{x} - \mathbf{y} \neq \mathbf{0}$. Tra queste possibili soluzioni approssimate si sceglie convenzionalmente quella che minimizza la norma dell'errore, ossia

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{A} \mathbf{x} - \mathbf{y}\|$$

Essa vale

$$\hat{\mathbf{x}} = \mathbf{A}_s^+ \mathbf{y} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{y}$$

e geometricamente consiste nella proiezione ortogonale di \mathbf{y} sul complemento ortogonale di $\mathcal{N}(\mathbf{A})$, ovvero sul sottospazio $\mathcal{R}(\mathbf{A}^\top)$. L'errore di approssimazione, detto anche *errore di proiezione*, vale in questo caso

$$\hat{\mathbf{e}} = (\mathbf{I} - \mathbf{A} \mathbf{A}_s^+) \mathbf{y}$$

e la sua norma è minima, come detto sopra.

La somiglianza tra la matrice di proiezione $\mathbf{I} - \mathbf{A}_d^+ \mathbf{A}$ e la matrice che fornisce l'errore di proiezione $\mathbf{I} - \mathbf{A} \mathbf{A}_s^+$ verrà spiegata nella Sezione 1.3.6.

Per calcolare le inverse generalizzate, si può utilizzare la decomposizione ai valori singolari; in particolare, ricordando la (1.3), la pseudo-inversa vale

$$\mathbf{A}^+ = \mathbf{V} \begin{bmatrix} \Sigma_r^{-1} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} \mathbf{U}^\top = \mathbf{Q} \Sigma_r^{-1} \mathbf{P}^\top.$$

1.3.5 Prodotti tra vettori

Abbiamo visto nella Sezione 1.3.1 che la definizione assiomatica di spazio vettoriale non comprenda anche la definizione di un prodotto, mentre per calcolare con enti geometrici è spesso necessario introdurre sia un prodotto sia una *struttura quadratica* o *metrica*, che possa fornire la “misura” degli elementi che si manipolano; una delle metriche più comuni è quella derivata dal prodotto scalare o interno tra vettori.

Prodotto scalare e norma di vettore

Dati due vettori reali $\mathbf{a}, \mathbf{b} \in \mathbb{V}(\mathbb{R})$, il prodotto *scalare* o *interno* (in inglese *scalar* o *inner product*) $\mathbf{a} \cdot \mathbf{b}$ è un numero reale che può venire definito sia in modo geometrico sia in modo analitico (per componenti):

$$\text{definizione geometrica: } \mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\| \|\mathbf{b}\| \cos \theta \quad (1.9)$$

$$\text{definizione analitica: } \mathbf{a} \cdot \mathbf{b} = \sum_k a_k b_k = \mathbf{a}^\top \mathbf{b} \quad (1.10)$$

dove $\|\mathbf{a}\|$ è la lunghezza del vettore \mathbf{a} e θ , ($0^\circ \leq \theta \leq 180^\circ$) è l'angolo compreso tra \mathbf{a} e \mathbf{b} . Alcuni testi indicano il prodotto scalare con il simbolo $\langle \mathbf{a}, \mathbf{b} \rangle$.

Il prodotto scalare soddisfa le seguenti proprietà:

$$\begin{aligned} \text{è distributivo rispetto alla somma:} & \quad (\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c} \\ \text{è distributivo rispetto al prodotto per scalare:} & \quad \alpha(\mathbf{a} \cdot \mathbf{b}) = (\alpha\mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (\alpha\mathbf{b}) \\ \text{è commutativo:} & \quad \mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a} \\ \text{è positivo:} & \quad \mathbf{a} \cdot \mathbf{a} > 0, \quad \forall \mathbf{a} \neq \mathbf{0} \end{aligned} \quad (1.11)$$

La definizione geometrica di prodotto scalare implica di aver preventivamente definito il concetto di angolo e di lunghezza, mentre nell'approccio analitico la lunghezza ovvero la *norma* (in inglese *norm*) può essere definita come grandezza derivata dal prodotto scalare

$$\|\mathbf{a}\| = \sqrt{\mathbf{a} \cdot \mathbf{a}} = \sqrt{\sum_k a_k^2} = \sqrt{\mathbf{a}^\top \mathbf{a}} \quad (1.12)$$

e l'angolo tra \mathbf{a} e \mathbf{b} come

$$\theta = \cos^{-1} \left(\frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \|\mathbf{b}\|} \right)$$

Si chiamano spazi euclidei o cartesiani quelli per i quali è definita la norma euclidea (1.12).

La norma euclidea derivata dal prodotto scalare non è l'unica norma possibile per “misurare le dimensioni” di un vettore. In generale la norma di un vettore deve soddisfare le seguenti tre proprietà (assiomi della norma):

1. $\|\mathbf{x}\| > 0$ per ogni $\mathbf{x} \neq \mathbf{0}$, $\|\mathbf{x}\| = 0$ se e solo se $\mathbf{x} = \mathbf{0}$;
2. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (diseguaglianza triangolare);
3. $\|\alpha\mathbf{x}\| = |\alpha| \|\mathbf{x}\|$ per ogni scalare α e ogni \mathbf{x} .

Generalizzando la norma euclidea, possiamo definire la p -norma $\|\cdot\|_p$ come

$$\|\mathbf{x}\|_p = \left(\sum_k |a_k|^p \right)^{1/p}$$

Tra le p -norme più usate nei vari campi della matematica e dell'ingegneria ricordiamo:

la norma 2 o euclidea	$p = 2$	$\ \mathbf{x}\ _2 = \sqrt{\mathbf{x}^\top \mathbf{x}}$
la norma 1	$p = 1$	$\ \mathbf{x}\ _1 = \sum_k a_k $
la norma ∞ o max-norma	$p = \infty$	$\ \mathbf{x}\ _\infty = \max_k x_k $

Si definisce *versore* un generico vettore diviso per la sua norma

$$\mathbf{u} = \frac{\mathbf{x}}{\|\mathbf{x}\|}; \quad \|\mathbf{u}\| = 1.$$

Il prodotto scalare \cdot opera tra due vettori e genera uno scalare, mentre in generale vorremmo poter definire un prodotto più generale che operi su due vettori e generi come risultato ancora un vettore.

È possibile definire quest'ultimo prodotto anche nel caso generale di vettori appartenenti a spazi n -dimensionali, ma per i fini che ci proponiamo è sufficiente limitarsi al caso tridimensionale, in cui possiamo definire il cosiddetto prodotto vettoriale.

Prodotto vettoriale o esterno

Nel caso particolare di due vettori tridimensionali

$$\mathbf{x} = [x_1 \ x_2 \ x_3]^\top \quad \text{e} \quad \mathbf{y} = [y_1 \ y_2 \ y_3]^\top, \quad \text{con} \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^3$$

il prodotto *vettoriale* o *esterno*⁴ (in inglese *outer* o *external* o *vector product*) $\mathbf{x} \times \mathbf{y}$ è un vettore che soddisfa le relazioni seguenti:

$$\mathbf{z} = \mathbf{x} \times \mathbf{y} = \begin{bmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{bmatrix} \quad (1.13)$$

⁴ Per il prodotto esterno utilizzeremo il simbolo \times , che è molto comune nella letteratura anglosassone; testi italiani usano più spesso il simbolo \wedge .

La (1.13) può essere scritta come prodotto della matrice antisimmetrica $\mathbf{S}(\mathbf{x})$ per il vettore \mathbf{y} :

$$\mathbf{x} \times \mathbf{y} = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix} \mathbf{y} = \mathbf{S}(\mathbf{x})\mathbf{y}$$

Le proprietà delle matrici antisimmetriche e i loro utilizzi sono descritte nel paragrafo 1.4.

La norma del prodotto esterno vale

$$\|\mathbf{z}\| = \|\mathbf{x}\| \|\mathbf{y}\| \sin \theta \quad (1.14)$$

dove θ è l'angolo tra i due vettori \mathbf{x} e \mathbf{y} misurato sul piano xy definito da questi ultimi; la direzione di \mathbf{z} è ortogonale al piano, il verso è dato dall'applicazione della regola della mano destra⁵, per portare \mathbf{x} su \mathbf{y} compiendo la rotazione di angolo minimo.

Il prodotto vettoriale soddisfa le seguenti proprietà:

$$\begin{aligned} \text{è non commutativo o anticommutativo:} & \quad \mathbf{x} \times \mathbf{y} = -(\mathbf{y} \times \mathbf{x}) \\ \text{è distributivo rispetto alla somma:} & \quad \mathbf{x} \times (\mathbf{y} + \mathbf{z}) = (\mathbf{x} \times \mathbf{y}) + (\mathbf{x} \times \mathbf{z}) \\ \text{è distributivo rispetto al prodotto per scalare:} & \quad \alpha(\mathbf{x} \times \mathbf{y}) = (\alpha\mathbf{x}) \times \mathbf{y} = \mathbf{x} \times (\alpha\mathbf{y}) \\ \text{è non associativo:} & \quad \mathbf{x} \times (\mathbf{y} \times \mathbf{z}) \neq (\mathbf{x} \times \mathbf{y}) \times \mathbf{z} \end{aligned} \quad (1.15)$$

Dati tre vettori tridimensionali \mathbf{x} , \mathbf{y} , \mathbf{z} , si definisce prodotto *triplo* il prodotto esterno triplo non associativo, ossia:

$$\begin{aligned} \mathbf{x} \times (\mathbf{y} \times \mathbf{z}) &= (\mathbf{x} \cdot \mathbf{z})\mathbf{y} - (\mathbf{x} \cdot \mathbf{y})\mathbf{z} \\ (\mathbf{x} \times \mathbf{y}) \times \mathbf{z} &= (\mathbf{x} \cdot \mathbf{z})\mathbf{y} - (\mathbf{y} \cdot \mathbf{z})\mathbf{x} \end{aligned} \quad (1.16)$$

Dati tre vettori \mathbf{x} , \mathbf{y} , \mathbf{z} , vale inoltre la seguente relazione:

$$(\mathbf{x} \times \mathbf{y}) \cdot \mathbf{z} = -(\mathbf{z} \times \mathbf{y}) \cdot \mathbf{x} \quad (1.17)$$

Il prodotto vettoriale, come abbiamo visto, è definito solo in spazi vettoriali di dimensione 3; la generalizzazione di questo prodotto a spazi n -dimensionali, con $n > 3$, richiede di introdurre le algebre di Clifford [5], cosa che va oltre gli scopi di questa Appendice.

⁵*Regola della mano destra*: se con le dita della mano destra si accompagna il movimento che porta \mathbf{x} a sovrapporsi a \mathbf{y} secondo il percorso a minimo angolo, allora il pollice si allinea lungo $\mathbf{x} \times \mathbf{y}$; questa è la stessa regola che definisce i tre versori di un sistema di riferimento cartesiano destrorso.

Prodotto diadico

Dati due vettori $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, si definisce *prodotto diadico* il seguente

$$\mathbf{x} \diamond \mathbf{y} = \mathbf{x} \mathbf{y}^\top = \mathbf{D}(\mathbf{x}, \mathbf{y}) = \begin{bmatrix} x_1 y_1 & \cdots & x_1 y_n \\ \vdots & x_i y_i & \vdots \\ x_n y_1 & \cdots & x_n y_n \end{bmatrix}$$

Si fa notare che alcuni testi anglosassoni chiamano questo prodotto *external product*, generando confusione con il prodotto vettoriale.

Risulta immediato dimostrare che il prodotto non è commutativo, in quanto, in generale $\mathbf{x} \mathbf{y}^\top \neq \mathbf{y} \mathbf{x}^\top$, essendo

$$\mathbf{D}(\mathbf{x}, \mathbf{y}) = \mathbf{D}^\top(\mathbf{y}, \mathbf{x})$$

risulta quindi $\mathbf{x} \diamond \mathbf{y} \neq \mathbf{y} \diamond \mathbf{x}$.

La matrice \mathbf{D} che si ottiene come risultato del prodotto risulta sempre avere rango $\rho(\mathbf{D}) = 1$, qualunque sia la dimensione n dei vettori di partenza.

Una proprietà utile che lega il prodotto vettoriale triplo e il prodotto diadico per vettori tridimensionali è la seguente

$$\begin{aligned} \mathbf{x} \times (\mathbf{y} \times \mathbf{z}) &= [(\mathbf{x} \cdot \mathbf{z}) \mathbf{I} - \mathbf{z} \diamond \mathbf{x}] \mathbf{y} \\ (\mathbf{x} \times \mathbf{y}) \times \mathbf{z} &= [(\mathbf{x} \cdot \mathbf{z}) \mathbf{I} - \mathbf{x} \diamond \mathbf{z}] \mathbf{y} \end{aligned}$$

È interessante sottolineare che, mentre il prodotto esterno al primo termine delle equazioni precedenti, è stato definito solo per vettori tridimensionali, i prodotti al secondo termine si possono calcolare indipendentemente dalle dimensioni dello spazio vettoriale.

Altri prodotti

Poiché il prodotto interno è un prodotto tra vettori che fornisce uno scalare ed il prodotto esterno non è associativo, nasce la necessità di definire un prodotto $\mathbf{a}\mathbf{b}$ tra vettori che obbedisca alla maggior parte delle regole della moltiplicazione “ordinaria”, ovvero possedga almeno le proprietà di essere associativo e distributivo, mentre la commutatività non è essenziale. Si richiede anche che, nel fare il prodotto, venga preservata la norma, ossia $\|\mathbf{a}\mathbf{b}\| = \|\mathbf{a}\| \|\mathbf{b}\|$.

Sono stati definiti in passato prodotti tra vettori che soddisfano questi requisiti. Di solito essi vengono trascurati nei testi elementari di algebra vettoriale. Tra questi, un qualche interesse per l'applicazione alla cinematica teorica e alla computer vision, oltreché nella fisica quantistica, rivestono il prodotto di Hamilton e il prodotto di Clifford.

Prodotto di Hamilton

Il prodotto di Hamilton trova la sua giustificazione nell'ambito della definizione di prodotto tra *quaternioni*. Qui ci limitiamo a definire tale prodotto come quel vettore \mathbf{c} ottenuto dai vettori \mathbf{a} e \mathbf{b} nel modo seguente

$$\mathbf{c} = \mathbf{a}\mathbf{b} = -\mathbf{a} \cdot \mathbf{b} + \mathbf{a} \times \mathbf{b} \quad (1.18)$$

Questo prodotto ha solo più un significato storico, in quanto presenta la spiacevole caratteristica di fornire un numero negativo come risultato del prodotto di un vettore per se stesso

$$\mathbf{a}\mathbf{a} = -\mathbf{a} \cdot \mathbf{a} + \mathbf{a} \times \mathbf{a} = -\|\mathbf{a}\|^2 \quad (1.19)$$

Esso fu presto abbandonato in favore di altri più semplici e utili, come i precedenti prodotti interno ed esterno, oppure più generali dal punto di vista geometrico, come il prodotto di Clifford.

Prodotto di Clifford

È stato dimostrato che un prodotto vettoriale che permetta di soddisfare gli stessi assiomi del prodotto tra due numeri reali, ossia la distributività, l'associatività e la commutatività, non esiste per spazi vettoriali con dimensioni $n \geq 3$; se si lascia cadere l'assioma della commutatività, si può definire il prodotto di Clifford, dal nome del matematico inglese William Clifford (1845-79) che per primo lo introdusse. Esso consente di estendere a spazi vettoriali \mathbb{R}^n , con $n > 3$, il prodotto esterno definito in 1.3.5.

Limitiamoci in un primo momento, per semplicità, al piano \mathbb{R}^2 : dati due vettori $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j}$, e $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j}$, il prodotto di Clifford risulta essere definito come:

$$\mathbf{a}\mathbf{b} = a_1b_1 + a_2b_2 + (a_1b_2 - a_2b_1)\mathbf{e}_{12} = \mathbf{a} \cdot \mathbf{b} + (a_1b_2 - a_2b_1)\mathbf{e}_{12} \quad (1.20)$$

dove \mathbf{e}_{12} prende il nome di *bivettore*. Esso è definito come l'area dotata di segno del parallelogrammo compreso tra \mathbf{i} e \mathbf{j} ; in un certo senso è analogo al prodotto esterno $\mathbf{i} \times \mathbf{j}$, salvo il fatto che quest'ultimo è interpretato come vettore ortogonale al piano in cui sono contenuti \mathbf{i} e \mathbf{j} , mentre il primo è da interpretarsi come una "pezza" (in inglese *patch*) nel medesimo piano, come illustrato in Fig. 1.1.

L'estensione allo spazio \mathbb{R}^3 si ottiene assumendo che sia verificata per il prodotto la seguente identità:

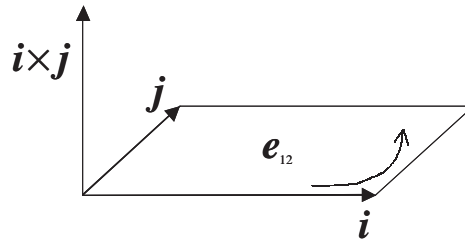
$$\mathbf{c}\mathbf{c} = \mathbf{c}^2 = \mathbf{c} \cdot \mathbf{c} \quad (1.21)$$

se poi consideriamo $\mathbf{c} = \mathbf{a} + \mathbf{b}$, otteniamo:

$$(\mathbf{a} + \mathbf{b})(\mathbf{a} + \mathbf{b}) = (\mathbf{a} + \mathbf{b}) \cdot (\mathbf{a} + \mathbf{b}) \quad (1.22)$$

da cui segue

$$\mathbf{a}\mathbf{b} + \mathbf{b}\mathbf{a} = 2\mathbf{a} \cdot \mathbf{b} \quad (1.23)$$

Figura 1.1: Il bivettore e_{12} nel piano \mathbb{R}^3 .

e quindi

$$ab = 2a \cdot b - ba \quad (1.24)$$

Il lettore interessato può fare riferimento al testo [6] per ulteriori approfondimenti.

1.3.6 Proiezioni e matrici di proiezione

Dato uno spazio vettoriale reale $\mathbb{V}(\mathbb{R}^n)$ di dimensioni n , dotato di prodotto scalare, ed un suo sottospazio $\mathbb{W}(\mathbb{R}^k)$ di dimensioni $k \leq n$, è possibile definire la *proiezione* dei vettori $\mathbf{v} \in \mathbb{V}$ sul sottospazio \mathbb{W} .

L'operatore di proiezione è definito dalla matrice quadrata di proiezione $\mathbf{P} \in \mathbb{R}^{n \times n}$, le cui colonne sono le proiezioni degli elementi della base di \mathbb{V} in \mathbb{W} . Una matrice è di proiezione se e solo se $\mathbf{P}^2 = \mathbf{P}$ ossia se essa è idempotente. La proiezione può essere *ortogonale*, oppure non ortogonale; nel primo caso la matrice \mathbf{P} è simmetrica, nel secondo caso no. Se \mathbf{P} è una matrice di proiezione, anche $\mathbf{I} - \mathbf{P}$ lo è.

Classici esempi di matrici di proiezione ortogonale sono le matrici associate alla pseudo-inversa sinistra $\mathbf{P}_1 = \mathbf{A}\mathbf{A}_s^+$ e $\mathbf{P}_2 = \mathbf{I} - \mathbf{A}\mathbf{A}_s^+$ e alla pseudo-inversa destra $\mathbf{P}_3 = \mathbf{A}_d^+\mathbf{A}$ e $\mathbf{P}_4 = \mathbf{I} - \mathbf{A}_d^+\mathbf{A}$ (vedi Sezione 1.3.4).

Dal punto di vista geometrico, \mathbf{P}_1 proietta ogni vettore $\mathbf{v} \in \mathbb{V}$ nel spazio immagine $\mathcal{R}(\mathbf{A})$ (vedi 1.3.3), mentre \mathbf{P}_2 proietta \mathbf{v} nel suo complemento ortogonale $\mathcal{R}(\mathbf{A})^\perp = \mathcal{N}(\mathbf{A}^\top)$.

1.3.7 Norme di matrice

Come per il vettore, è possibile fornire una “misura” della “grandezza” di una matrice, definendone la norma. Poiché una matrice rappresenta una trasformazione lineare tra vettori, la norma misura quando grande sia questa trasformazione, ma deve in qualche modo “normalizzare” il risultato perché questo non sia influenzato dalla “grandezza” del vettore che viene trasformato, ossia:

$$\|\mathbf{A}\| = \sup_{\|\mathbf{x}\|} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|} = \sup_{\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x}\|.$$

Data una matrice quadrata $\mathbf{A} \in \mathbb{R}^{n \times n}$ la sua norma deve soddisfare in generale i seguenti assiomi generali:

1. $\|\mathbf{A}\| > 0$ per ogni $\mathbf{A} \neq \mathbf{O}$, $\|\mathbf{A}\| = 0$ se e solo se $\mathbf{A} = \mathbf{O}$;
2. $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$ (diseguaglianza triangolare);
3. $\|\alpha\mathbf{A}\| = |\alpha| \|\mathbf{A}\|$ per ogni scalare α e ogni \mathbf{A} ;
4. $\|\mathbf{AB}\| \leq \|\mathbf{A}\| \|\mathbf{B}\|$.

Data $\mathbf{A} \in \mathbb{R}^{n \times n}$ e i suoi autovalori $\{\lambda_i(\mathbf{A})\}$, vale la seguente diseguaglianza

$$\frac{1}{\|\mathbf{A}^{-1}\|} \leq |\lambda_i| \leq \|\mathbf{A}\| \quad \forall i = 1, \dots, n$$

Fatte queste premesse, elenchiamo le norme di matrice più comunemente adottate, considerando solo matrici reali

1. norma *spettrale*:

$$\|\mathbf{A}\|_2 = \sqrt{\max_i \{\lambda_i(\mathbf{A}^T \mathbf{A})\}}$$

2. norma di *Frobenius*:

$$\|\mathbf{A}\|_F = \sqrt{\sum_i \sum_j a_{ij}^2} = \sqrt{\text{tr } \mathbf{A}^T \mathbf{A}}$$

3. massimo valore singolare:

$$\|\mathbf{A}\|_\sigma = \sqrt{\max_i \{\sigma_i(\mathbf{A})\}}$$

4. norma 1 o max-norma:

$$\|\mathbf{A}\|_1 = \max_j \sum_{i=1}^n |a_{ij}|$$

5. norma ∞ :

$$\|\mathbf{A}\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|$$

In generale $\|\mathbf{A}\|_2 = \|\mathbf{A}\|_\sigma$ e $\|\mathbf{A}\|_2^2 \leq \|\mathbf{A}\|_1 \|\mathbf{A}\|_\infty$

1.4 Matrici antisimmetriche

Abbiamo visto sopra che una matrice \mathbf{S} si dice *antisimmetrica* quando soddisfa la seguente proprietà:

$$\mathbf{S} + \mathbf{S}^T = \mathbf{O} \quad (1.25)$$

Una matrice antisimmetrica ha perciò tutti zeri sulla diagonale principale ed elementi fuori dalla diagonale che soddisfano la relazione $s_{ij} = -s_{ji}$.

Ne segue che una matrice antisimmetrica è definita da soli $\frac{n(n-1)}{2}$ parametri.

Per $n = 3$ risulta $\frac{n(n-1)}{2} = 3$, per cui la matrice antisimmetrica ha esattamente tre elementi indipendenti, che possono essere considerati come le componenti di un generico vettore tridimensionale \mathbf{v} ; la matrice viene allora indicata come $\mathbf{S}(\mathbf{v})$.

Nel seguito studieremo le proprietà delle matrici antisimmetriche per $n = 3$, in quanto sono di fondamentale importanza per definire la cinematica delle rotazioni di corpi rigidi in spazi tridimensionali.

Se $\mathbf{v} = [v_x \ v_y \ v_z]^T$ è un vettore qualsiasi, possiamo definire $\mathbf{S}(\mathbf{v})$ come l'operatore che trasforma \mathbf{v} in una matrice antisimmetrica:

$$\mathbf{S}(\mathbf{v}) = \begin{bmatrix} 0 & -v_z & v_y \\ v_z & 0 & -v_x \\ -v_y & v_x & 0 \end{bmatrix} \quad (1.26)$$

e viceversa, data una matrice antisimmetrica qualsiasi, è sempre possibile estrarre da essa un vettore \mathbf{v} .

La matrice $\mathbf{S}(\mathbf{v})$ si indica semplicemente con \mathbf{S} quando non si vuole evidenziare la dipendenza dal vettore \mathbf{v} . La proprietà di antisimmetria comporta la seguente identità:

$$\mathbf{S}^T(\mathbf{v}) = -\mathbf{S}(\mathbf{v}) = \mathbf{S}(-\mathbf{v}) \quad (1.27)$$

Le matrici antisimmetriche soddisfano la proprietà di linearità; dati due scalari $\lambda_i \in \mathbb{R}$, vale la proprietà

$$\mathbf{S}(\lambda_1 \mathbf{u} + \lambda_2 \mathbf{v}) = \lambda_1 \mathbf{S}(\mathbf{u}) + \lambda_2 \mathbf{S}(\mathbf{v}) \quad (1.28)$$

Inoltre, dati due vettori qualsiasi \mathbf{v} e \mathbf{u} , si ha la seguente importante proprietà:

$$\mathbf{S}(\mathbf{u})\mathbf{v} = \mathbf{u} \times \mathbf{v} \quad (1.29)$$

e quindi $\mathbf{S}(\mathbf{u})$ può essere interpretata come l'operatore $(\mathbf{u} \times)$ e viceversa.

Da questa proprietà e dalla anti-commutatività del prodotto esterno segue che $\mathbf{S}(\mathbf{u})\mathbf{v} = -\mathbf{S}(\mathbf{v})\mathbf{u}$.

È semplice verificare che la matrice $\mathbf{S}(\mathbf{u})\mathbf{S}(\mathbf{u}) = \mathbf{S}^2(\mathbf{u})$ è simmetrica e verifica la relazione

$$\mathbf{S}^2(\mathbf{u}) = \mathbf{v}\mathbf{v}^T - \|\mathbf{v}\|^2 \mathbf{I} \quad (1.30)$$

Autovalori e autovettori di matrici antisimmetriche Data la matrice antisimmetrica $\mathbf{S}(\mathbf{v})$ i suoi autovalori sono immaginari o nulli:

$$\lambda_1 = 0, \quad \lambda_{2,3} = \pm j \|\mathbf{v}\|$$

L'autovettore relativo all'autovalore $\lambda_1 = 0$ vale \mathbf{v} ; gli altri due sono complessi coniugati.

1.5 Matrici ortogonali e ortonormali

Una generica matrice quadrata $\mathbf{U} \in \mathbb{R}^n$ è detta *ortonormale* quando:

- la sua inversa coincide con la sua trasposta

$$\mathbf{U} \mathbf{U}^\top = \mathbf{U}^\top \mathbf{U} = \mathbf{I} \quad (1.31)$$

ovvero

$$\mathbf{U}^{-1} = \mathbf{U}^\top \quad (1.32)$$

- Le colonne di \mathbf{U} sono tra loro ortogonali e a norma unitaria, come pure le righe.
- $\|\mathbf{U}\| = 1$;
- Il determinante di \mathbf{U} ha modulo unitario:

$$|\det(\mathbf{U})| = 1 \quad (1.33)$$

perciò esso può valere +1 oppure -1.

Viene chiamata *ortogonale* quella matrice quadrata per cui vale una relazione meno forte della (1.31), ovvero

$$\mathbf{U}^\top \mathbf{U} = \begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_n \end{bmatrix}$$

con $\alpha_{ii} \neq 0$.

Il termine "ortonormale" viene riservato al caso in cui $\alpha_{ii} = 1$.

Il prodotto scalare è invariante a trasformazioni ortonormali, ossia

$$(\mathbf{U}\mathbf{x}) \cdot (\mathbf{U}\mathbf{y}) = (\mathbf{U}\mathbf{x})^\top (\mathbf{U}\mathbf{y}) = \mathbf{x}^\top \mathbf{U}^\top \mathbf{U} \mathbf{y} = \mathbf{x}^\top \mathbf{I} \mathbf{y} = \mathbf{x}^\top \mathbf{y} = \mathbf{x} \cdot \mathbf{y}$$

Se \mathbf{U} è una matrice ortonormale,⁶ di dimensioni opportune, allora $\|\mathbf{A}\mathbf{U}\| = \|\mathbf{U}\mathbf{A}\| = \|\mathbf{A}\|$.

Limitandoci al caso di matrici $\mathbf{U}_{3 \times 3}$, solo 3 dei 9 elementi che compongono la matrice sono indipendenti, in quanto le condizioni di ortonormalità tra le righe o tra le colonne definiscono 6 vincoli.

Come si può vedere in [2], le matrici ortonormali $\mathbf{U}_{3 \times 3}$ sono la rappresentazione di trasformazioni geometriche di corpi rigidi nello spazio Euclideo: infatti quando $\det(\mathbf{U}) = +1$, \mathbf{U} rappresenta una **rotazione propria**, mentre quando $\det(\mathbf{U}) = -1$, \mathbf{U} rappresenta una **rotazione impropria** ovvero una **roto-riflessione**.

Se \mathbf{U} è una matrice ortonormale, vale la proprietà distributiva⁷ rispetto al prodotto esterno:

$$\mathbf{U}(\mathbf{x} \times \mathbf{y}) = (\mathbf{U}\mathbf{x}) \times (\mathbf{U}\mathbf{y}) \quad (1.34)$$

Per ogni matrice di rotazione \mathbf{U} e ogni vettore \mathbf{x} si dimostra che

$$\begin{aligned} \mathbf{U}\mathbf{S}(\mathbf{x})\mathbf{U}^\top\mathbf{y} &= \mathbf{U}(\mathbf{x} \times (\mathbf{U}^\top\mathbf{y})) \\ &= (\mathbf{U}\mathbf{x}) \times (\mathbf{U}\mathbf{U}^\top\mathbf{y}) \\ &= (\mathbf{U}\mathbf{x}) \times \mathbf{y} \\ &= \mathbf{S}(\mathbf{U}\mathbf{x})\mathbf{y} \end{aligned} \quad (1.35)$$

dove $\mathbf{S}(\mathbf{x})$ è la matrice antisimmetrica associata a \mathbf{x} ; si ricavano pertanto le relazioni seguenti:

$$\begin{aligned} \mathbf{U}\mathbf{S}(\mathbf{x})\mathbf{U}^\top &= \mathbf{S}(\mathbf{U}\mathbf{x}) \\ \mathbf{U}\mathbf{S}(\mathbf{x}) &= \mathbf{S}(\mathbf{U}\mathbf{x})\mathbf{U} \end{aligned} \quad (1.36)$$

1.6 Forme bilineari e quadratiche

Si definisce *forma bilineare* associata alla matrice $\mathbf{A} \in \mathbb{R}^{m \times n}$ la variabile scalare

$$b(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\top \mathbf{A} \mathbf{y} = \mathbf{y}^\top \mathbf{A}^\top \mathbf{x}$$

Si definisce *forma quadratica* associata alla matrice quadrata $\mathbf{A} \in \mathbb{R}^{n \times n}$ la variabile scalare

$$q(\mathbf{x}) = \mathbf{x}^\top \mathbf{A} \mathbf{x} = \mathbf{x}^\top \mathbf{A}^\top \mathbf{x}$$

Qualsiasi forma quadratica associata ad una matrice antisimmetrica $\mathbf{S}(\mathbf{y})$ è identicamente nulla, ossia

$$\mathbf{x}^\top \mathbf{S}(\mathbf{y}) \mathbf{x} \equiv 0 \quad (1.37)$$

⁶La regola vale anche nel caso più generale in cui \mathbf{U} sia una *matrice unitaria*, definita da $\mathbf{U}^* \mathbf{U} = \mathbf{I}$.

⁷ Questa proprietà **non** è generalmente vera, salvo appunto quando \mathbf{U} è ortonormale.

per ogni \mathbf{x} e ogni \mathbf{y} . La dimostrazione di questa proprietà è semplice: definendo $\mathbf{w} = \mathbf{S}(\mathbf{y})\mathbf{x} = \mathbf{y} \times \mathbf{x}$, avremo $\mathbf{x}^\top \mathbf{S}(\mathbf{y})\mathbf{x} = \mathbf{x}^\top \mathbf{w}$, ma essendo \mathbf{w} ortogonale sia a \mathbf{y} sia a \mathbf{x} per definizione, il prodotto scalare $\mathbf{x}^\top \mathbf{w}$, e quindi la forma quadratica, sarà sempre nullo.

Pertanto, ricordando la decomposizione (1.1), la forma quadratica dipende solo dalla parte simmetrica \mathbf{A}_s della matrice \mathbf{A} :

$$q(\mathbf{x}) = \mathbf{x}^\top \mathbf{A} \mathbf{x} = \mathbf{x}^\top (\mathbf{A}_s + \mathbf{A}_a) \mathbf{x} = \mathbf{x}^\top \mathbf{A}_s \mathbf{x}$$

Una matrice quadrata \mathbf{A} si dice *definita positiva* se la forma quadratica associata $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ soddisfa le condizioni

$$\begin{aligned} \mathbf{x}^\top \mathbf{A} \mathbf{x} &> 0 \quad \forall \mathbf{x} \neq \mathbf{0} \\ \mathbf{x}^\top \mathbf{A} \mathbf{x} &= 0 \quad \mathbf{x} = \mathbf{0} \end{aligned}$$

Una matrice quadrata \mathbf{A} si dice *semidefinita positiva* se la forma quadratica associata $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ soddisfa la condizione

$$\mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0 \quad \forall \mathbf{x}$$

Una matrice quadrata \mathbf{A} si dice *definita negativa* se $-\mathbf{A}$ è definita positiva; analogamente una matrice quadrata \mathbf{A} si dice *semidefinita negativa* se $-\mathbf{A}$ è semidefinita positiva.

Spesso, per indicare queste matrici si usano le notazioni seguenti:

$$\begin{aligned} \text{matrice definita positiva:} & \quad \mathbf{A} \succ 0 \\ \text{matrice semidefinita positiva:} & \quad \mathbf{A} \succeq 0 \\ \text{matrice definita negativa:} & \quad \mathbf{A} \prec 0 \\ \text{matrice semidefinita negativa:} & \quad \mathbf{A} \preceq 0 \end{aligned}$$

Condizione necessaria affinché la matrice quadrata \mathbf{A} sia definita positiva è che gli elementi sulla diagonale siano strettamente positivi.

Condizione necessaria e sufficiente affinché la matrice quadrata \mathbf{A} sia definita positiva è che tutti gli autovalori siano strettamente positivi.

Il *criterio di Sylvester* afferma che condizione necessaria e sufficiente affinché la matrice quadrata \mathbf{A} sia definita positiva è che tutti i suoi minori principali siano strettamente positivi. Una matrice definita positiva ha rango pieno ed è sempre invertibile.

La forma quadratica $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ soddisfa la relazione seguente

$$\lambda_{\min}(\mathbf{A}) \|\mathbf{x}\|^2 \leq \mathbf{x}^\top \mathbf{A} \mathbf{x} \leq \lambda_{\max}(\mathbf{A}) \|\mathbf{x}\|^2$$

dove $\lambda_{\min}(\mathbf{A})$ e $\lambda_{\max}(\mathbf{A})$ sono, rispettivamente, l'autovalore minimo e massimo di \mathbf{A} .

Una matrice \mathbf{A} semidefinita positiva ha rango $\rho(\mathbf{A}) = r < n$, ovvero possiede r autovalori strettamente positivi e $n - r$ autovalori nulli. La forma quadratica si annulla per ogni $\mathbf{x} \in \mathcal{N}(\mathbf{A})$.

Data una matrice reale di dimensioni qualsiasi $\mathbf{A} \in \mathbb{R}^{m \times n}$, abbiamo visto che sia $\mathbf{A}^\top \mathbf{A}$, sia $\mathbf{A} \mathbf{A}^\top$ sono simmetriche; inoltre abbiamo visto nella Sezione 1.2.2 che $\rho(\mathbf{A}^\top \mathbf{A}) = \rho(\mathbf{A} \mathbf{A}^\top) = \rho(\mathbf{A})$. Si può dimostrare che esse hanno sempre autovalori reali non negativi, e quindi sono definite o semi-definite positive: in particolare, se la matrice \mathbf{A} ha rango pieno,

- se $m < n$, $\mathbf{A}^\top \mathbf{A} \succeq 0$ e $\mathbf{A} \mathbf{A}^\top \succ 0$,
- se $m = n$, $\mathbf{A}^\top \mathbf{A} \succ 0$ e $\mathbf{A} \mathbf{A}^\top \succ 0$,
- se $m > n$, $\mathbf{A}^\top \mathbf{A} \succ 0$ e $\mathbf{A} \mathbf{A}^\top \succeq 0$.

Data la forma bilineare $b(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\top \mathbf{A} \mathbf{y}$, si definiscono *gradienti* le seguenti espressioni:

$$\begin{aligned} \text{gradiente rispetto a } \mathbf{x}: \quad \mathbf{grad}_x b(\mathbf{x}, \mathbf{y}) &= \left(\frac{\partial b(\mathbf{x}, \mathbf{y})}{\partial \mathbf{x}} \right)^\top = \mathbf{A} \mathbf{y} \\ \text{gradiente rispetto a } \mathbf{y}: \quad \mathbf{grad}_y b(\mathbf{x}, \mathbf{y}) &= \left(\frac{\partial b(\mathbf{x}, \mathbf{y})}{\partial \mathbf{y}} \right)^\top = \mathbf{A}^\top \mathbf{x} \end{aligned}$$

Data la forma quadratica $q(\mathbf{x}) = \mathbf{x}^\top \mathbf{A} \mathbf{x}$, si definisce *gradiente* rispetto a \mathbf{x} la seguente espressione:

$$\mathbf{grad}_x q(\mathbf{x}) = \left(\frac{\partial q(\mathbf{x})}{\partial \mathbf{x}} \right)^\top = 2\mathbf{A} \mathbf{x}.$$

Bibliografia

- [1] <http://ocw.mit.edu/18/18.06/f02/index.htm>.
- [2] B. Bona. *Modellistica dei Robot Industriali*. CELID, Torino, 2002.
- [3] C. N. Dorny. *A Vector Space Approach to Models and Optimization*. John Wiley & Sons, 1975.
- [4] F. R. Gantmacher. *The Theory of Matrices*, volume 2. Chelsea Publishing Company, 3rd edition, 1989.
- [5] P. Lounesto. *Clifford Algebras and Spinors*. Cambridge University Press, second edition, 2001.
- [6] P. Lounesto. *Clifford Algebras and Spinors*. Cambridge University Press, second edition, 2001.
- [7] D. G. Luenberger. *Optimization by Vector Space Methods*. John Wiley & Sons, 1969.
- [8] R. Monaco and A. Rèpaci. *Algebra Lineare: Vettori, Matrici, Applicazioni*. CELID, Torino, 2002.
- [9] S. Rinaldi. *Algebra Lineare*. CLUP, Milano, 1971.
- [10] G. Strang. *Introduction to Linear Algebra*. Wellesey- Cambridge Press, 1998.

Indice

1	Matrici e vettori	2
1.1	Definizioni	2
1.2	Operazioni sulle matrici	4
1.2.1	Autovalori e autovettori	11
1.2.2	Decomposizione ai valori singolari	12
1.3	Vettori e spazi vettoriali	15
1.3.1	Spazio vettoriale	15
1.3.2	Funzioni lineari	16
1.3.3	Matrici come rappresentazione di operatori lineari	17
1.3.4	Inversa generalizzata	18
1.3.5	Prodotti tra vettori	21
1.3.6	Proiezioni e matrici di proiezione	26
1.3.7	Norme di matrice	26
1.4	Matrici antisimmetriche	28
1.5	Matrici ortogonali e ortonormali	29
1.6	Forme bilineari e quadratiche	30

Elenco delle figure

1.1	Il bivettore e_{12} nel piano \mathbb{R}^3	26
-----	----------------------------------------------------------	----