

Dispense Algebra Lineare e Geometria
01RKC

21 febbraio 2023

Indice

1	Operazioni fra matrici: somma e prodotto per scalari	5
1.1	Matrici a coefficienti in \mathbb{R}	5
1.2	Matrici quadrate	8
1.3	Somma e prodotto per scalare	11
2	Operazioni fra matrici: prodotto righe per colonne	15
2.1	Prodotto di matrici	15
2.2	Complessità computazionale del prodotto di matrici	20
2.3	Matrici invertibili e inversa di una matrice quadrata	22
2.4	Algebra lineare sui numeri complessi	24
3	Sistemi lineari e loro rappresentazione matriciale	25
3.1	Sistemi di equazioni lineari	25
3.2	Matrici (fortemente) ridotte per righe	28
4	Operazioni elementari di riga e rango di una matrice	31
4.1	Operazioni elementari di riga	31
4.2	Interpretazione matriciale	35
4.3	Rango di una matrice	37
5	Risoluzione di sistemi lineari ed equazioni matriciali	39
5.1	Risoluzione di sistemi	39
5.2	Equazioni matriciali	43
5.3	Calcolo dell'inversa di una matrice	47
5.4	Operazioni elementari di colonna	48
6	Determinanti	51
6.1	La definizione di determinante	51
6.2	Proprietà e calcolo di determinanti	56
6.3	Ancora sull'inversa di matrici	61
6.4	Il metodo di Cramer	63
7	Vettori applicati	65
7.1	Vettori geometrici	65
7.2	Sistemi di coordinate nel piano	67

7.3	Sistemi di coordinate nello spazio	69
7.4	Prime operazioni sui vettori applicati	71
7.5	Caratterizzazione dei vettori paralleli e complanari	75
8	Operazioni fra vettori applicati	79
8.1	Prodotto scalare	79
8.2	Prodotto vettoriale	84
8.3	Prodotto misto	88
9	Equazioni parametriche di rette e piani nello spazio	91
9.1	Equazioni parametriche di rette	91
9.2	Equazioni parametriche di piani	98
10	Equazioni cartesiane di rette e piani nello spazio	103
10.1	Equazioni cartesiane di piani	103
10.2	Equazioni cartesiane di rette	108
10.3	Posizioni relative di rette e piani	112
11	Minima distanza	119
11.1	Minima distanza tra insiemi	119
11.2	Distanza di un punto da una retta o da un piano	120
11.3	Distanza di un piano da una retta o da un piano	125
11.4	Distanza fra due rette	127
12	Spazi vettoriali e loro sottospazi	131
12.1	Spazi vettoriali	131
12.2	Sottospazi vettoriali	134
12.3	Alcune operazioni notevoli fra sottospazi	138
13	Dipendenza e indipendenza lineare	143
13.1	Combinazioni lineari	143
13.2	Dipendenza lineare	148
14	Basi di uno spazio vettoriale finitamente generato	153
14.1	Il metodo degli scarti	153
14.2	Basi di uno spazio vettoriale	155
15	Dimensione di uno spazio vettoriale finitamente generato	161
15.1	Dimensione di uno spazio vettoriale	161
15.2	Dimensione di sottospazi	163
15.3	Rango e dimensione	167
15.4	La formula di Grassmann	174

16 Applicazioni lineari fra spazi vettoriali	177
16.1 Applicazioni lineari	177
16.2 Immagine e nucleo di un'applicazione lineare	182
16.3 Applicazioni lineari iniettive e suriettive	183
16.4 Isomorfismi	186
17 Applicazioni lineari e loro rappresentazione matriciale	189
17.1 Applicazioni lineari tra spazi di dimensione finita	189
17.2 Matrice di un'applicazione lineare	192
17.3 Endomorfismi	198
18 Autovalori, autovettori e autospazi	203
18.1 Autovalori, autovettori e autospazi di endomorfismi	203
18.2 Autovalori, autovettori e autospazi di matrici	205
19 Diagonalizzazione	211
19.1 Diagonalizzazione di matrici	211
19.2 Diagonalizzazione di matrici simmetriche	215
19.3 Il teorema di Cayley–Hamilton	216
20 Spazi vettoriali con prodotto scalare	221
20.1 Prodotti scalari	221
20.2 Basi ortonormali	224
20.3 Matrici ortogonali	228
20.4 Diagonalizzazione ortogonale per matrici simmetriche	230
21 Forme quadratiche	233
21.1 Forme quadratiche reali	233
21.2 Forme quadratiche definite	234
22 Trasformazioni nel piano	241
22.1 Ellisse, iperbole, parabola	241
22.2 Rototraslazioni nel piano	247
23 Coniche	251
23.1 Coniche e loro riduzione a forma canonica	251
23.2 Determinazione del tipo di una conica	256
24 Sfere e circonferenze nello spazio	263
24.1 Sfere nello spazio	263
24.2 Circonferenze nello spazio	265
24.3 Intersezione di due sfere	270
25 Trasformazioni nello spazio e quadriche	275
25.1 Quadriche e loro riduzione a forma canonica	275
25.2 Classificazione delle quadriche	278

Operazioni fra matrici: somma e prodotto per scalari

1.1 Matrici a coefficienti in \mathbb{R}

Definizione 1.1.

Siano dati $m, n \in \mathbb{Z}$ positivi. Una *matrice* $m \times n$ a coefficienti in \mathbb{R} è un insieme di mn numeri reali disposti su m righe ed n colonne circondata da parentesi tonde. Tali numeri sono detti *entrate*, *coefficienti*, o *componenti* della matrice. L'insieme di tutte le matrici $m \times n$ a coefficienti in \mathbb{R} si indica con $\mathbb{R}^{m,n}$.

Vi sono alcuni casi particolari che vale la pena evidenziare:

- la matrice $0_{m,n} \in \mathbb{R}^{m,n}$ avente tutte le entrate nulle viene detta *matrice nulla*;
- se $m = n$, cioè se il numero delle righe è uguale al numero delle colonne, la matrice è detta *matrice quadrata*;
- se $m = 1$ la matrice è detta *matrice riga*, mentre se $n = 1$ *matrice colonna*;
- se $m = n = 1$, cioè quando ci sono una sola riga e una sola colonna (e, quindi, anche una sola entrata), si preferisce identificare $\mathbb{R}^{1,1}$ con \mathbb{R} .

Esempio 1.2. Diamo alcuni esempi di matrici.

$$A = \begin{pmatrix} 1 & \pi \\ -3/19 & \sqrt{21} \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{3,2}, \quad B = \begin{pmatrix} 1 & -3/19 & 0 \\ \pi & \sqrt{21} & 0 \end{pmatrix} \in \mathbb{R}^{2,3},$$
$$C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{2,2}, \quad D = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} \in \mathbb{R}^{3,1}, \quad E = (1 \ 0) \in \mathbb{R}^{1,2}.$$

La matrice C è quadrata, D è una matrice colonna ed E una matrice riga. Le matrici

$$0_{2,2} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad 0_{2,3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

sono rispettivamente le matrici nulle 2×2 e 2×3 . ♠

☠ La tabella

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & -1 & & 0 \\ & & 17 & \end{pmatrix}$$

non è una matrice.

⚠ Testi diversi utilizzano notazioni diverse per indicare una matrice! Ad esempio la matrice dell'Esempio 1.2 può essere scritta come

$$\begin{pmatrix} 1 & \pi \\ -3/19 & \sqrt{21} \\ 0 & 0 \end{pmatrix}, \quad \begin{bmatrix} 1 & \pi \\ -3/19 & \sqrt{21} \\ 0 & 0 \end{bmatrix}, \quad \left\| \begin{array}{cc} 1 & \pi \\ -3/19 & \sqrt{21} \\ 0 & 0 \end{array} \right\|.$$

La notazione

$$\left\{ \begin{array}{cc} 1 & \pi \\ -3/19 & \sqrt{21} \\ 0 & 0 \end{array} \right\}$$

è da evitare, perché può creare confusione con quella usata per gli insiemi.

Sia $A \in \mathbb{R}^{m,n}$ una matrice. Ad ogni sua entrata a rimangono associati due numeri interi positivi, gli indici i e j della riga e della colonna al cui incrocio si trova a : i e j vengono detti rispettivamente *indice di riga* e *indice di colonna* dell'entrata a e a si dice *entrata in posizione* (i, j) .

Spesso, per indicare nelle formule l'entrata in posizione (i, j) si scrive $a_{i,j}$. Similmente la matrice A si può indicare con la notazione

$$A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}.$$

Tale notazione significa che A è la matrice le cui entrate sono i numeri $a_{i,j}$ con gli indici i e j che variano da 1 a m e da 1 a n rispettivamente.

Se A è quadrata con $m = n$ si può scrivere anche

$$A = (a_{i,j})_{1 \leq i,j \leq n}.$$

Quando le dimensioni della matrice sono fissate, a volte conviene indicare le entrate con lettere distinte. Per esempio, una matrice 2×2 generica può essere indifferentemente indicata con

$$(a_{i,j})_{\substack{1 \leq i \leq 2 \\ 1 \leq j \leq 2}}, \quad (a_{i,j})_{1 \leq i,j \leq 2}, \quad \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}, \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Esempio 1.3. Si considerino le prime due matrici dell'Esempio 1.2

$$A = \begin{pmatrix} 1 & \pi \\ -3/19 & \sqrt{21} \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{3,2}, \quad B = \begin{pmatrix} 1 & -3/19 & 0 \\ \pi & \sqrt{21} & 0 \end{pmatrix} \in \mathbb{R}^{2,3}.$$

L'entrata $(1, 2)$ di A è $a_{1,2} = \pi$. Le entrate $(3, 1)$ e $(3, 2)$ di A sono $a_{3,1} = a_{3,2} = 0$. Invece le entrate $(3, 3)$ e $(2, 3)$ non esistono.

Analogamente, le entrate $(3, 1)$, $(3, 2)$, $(3, 3)$ di B non esistono. Invece le entrate $(1, 2)$ e $(2, 3)$ di B sono $b_{1,2} = -3/19$ e $b_{2,3} = 0$. ♠

Definizione 1.4 (Opposto di una matrice).

Sia data $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathbb{R}^{m,n}$. L'opposto di A è la matrice di $\mathbb{R}^{m,n}$, indicata con $-A$, la cui entrata (i, j) coincide con l'opposto dell'entrata (i, j) della matrice A , per $i = 1, \dots, m$ e $j = 1, \dots, n$.

Se $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathbb{R}^{m,n}$ si scrive in simboli $-A = (-a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathbb{R}^{m,n}$.

Esempio 1.5. Per le matrici A e B degli Esempi 1.2 e 1.3, vale che

$$-A = \begin{pmatrix} -1 & -\pi \\ 3/19 & -\sqrt{21} \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{3,2} \quad -B = \begin{pmatrix} -1 & 3/19 & 0 \\ -\pi & -\sqrt{21} & 0 \end{pmatrix} \in \mathbb{R}^{2,3}. \quad \spadesuit$$

Definizione 1.6 (Uguaglianza tra matrici).

Siano date $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathbb{R}^{m,n}$ e $B = (b_{i,j})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}} \in \mathbb{R}^{p,q}$. A e B si dicono *uguali* se valgono le seguenti proprietà:

- (U1) A e B hanno lo stesso numero di righe e colonne, cioè se $m = p$ ed $n = q$;
- (U2) le entrate di A e B aventi la stessa posizione nelle due matrici coincidono, cioè se $a_{i,j} = b_{i,j}$ per ogni $i = 1, \dots, m = p$ e $j = 1, \dots, n = q$.

Le due matrici A e B dell'Esempio 1.3 sono, perciò, diverse. Ciononostante, sono legate da un'ovvia relazione: l'entrata (i, j) di A coincide con l'entrata (j, i) di B .

Definizione 1.7 (Matrice trasposta).

Sia data $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathbb{R}^{m,n}$. La *trasposta* di A è la matrice di $\mathbb{R}^{n,m}$, indicata con tA , la cui entrata (j, i) coincide con l'entrata (i, j) della matrice A per $i = 1, \dots, m$ e $j = 1, \dots, n$.

Se $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathbb{R}^{m,n}$ scriveremo in simboli ${}^tA = (a_{j,i})_{\substack{1 \leq j \leq n \\ 1 \leq i \leq m}} \in \mathbb{R}^{n,m}$.

Esempio 1.8. Per le matrici A e B dell'Esempio 1.3 vale $B = {}^tA$ e $A = {}^tB$.

La trasposta di una matrice riga è una matrice colonna e viceversa: per esempio se

$$C = (\pi \quad 0 \quad -2 \quad 1/3) \in \mathbb{R}^{1,4}$$

allora

$${}^tC = \begin{pmatrix} \pi \\ 0 \\ -2 \\ 1/3 \end{pmatrix} \in \mathbb{R}^{4,1}.$$

Può accadere che la trasposta di una matrice coincida con la matrice stessa: per esempio se

$$D = \begin{pmatrix} 2 & e & -\pi \\ e & 0 & 5 \\ -\pi & 5 & \sqrt{2} \end{pmatrix} \in \mathbb{R}^{3,3}$$

allora ${}^tD = D$. Nel prossimo paragrafo studieremo meglio questo caso. ♠

⚠️ Testi diversi utilizzano notazioni diverse per denotare la trasposta. Non solo tA , ma anche A^t , ${}^\top A$, A^\top ; più raramente si trova anche A_{-1} .

Proposizione 1.9 (Proprietà della trasposta).

Valgono le seguenti proprietà:

- (T1) per ogni matrice A , si ha che $A \in \mathbb{R}^{m,n}$ se e solo se ${}^tA \in \mathbb{R}^{n,m}$;
- (T2) per ogni $A \in \mathbb{R}^{m,n}$ si ha che ${}^t({}^tA) = A$ (la trasposizione è un'operazione involutiva);
- (T3) per ogni $A \in \mathbb{R}^{m,n}$ si ha che ${}^t(-A) = -({}^tA)$ (la trasposizione è compatibile con il passaggio all'opposto).

Dimostrazione. Dimostriamo separatamente le varie proprietà. Per le uguaglianze tra matrici usiamo la Definizione 1.6. Osserviamo innanzitutto che in tutte le uguaglianze da dimostrare le dimensioni delle matrici a destra e a sinistra dell'uguale coincidono, quindi dobbiamo solo controllare che anche le entrate al posto (i, j) coincidano.

- (T1) è una conseguenza diretta della definizione: il numero di righe di A è uguale al numero di colonne di tA e similmente il numero di colonne di A è uguale al numero di righe di tA .
- (T2) L'entrata in posizione (i, j) di ${}^t({}^tA)$ coincide per definizione con l'entrata in posizione (j, i) di tA , che a sua volta coincide con l'entrata in posizione (i, j) di A . Le due matrici quindi coincidono.
- (T3) L'entrata in posizione (i, j) di ${}^t(-A)$ coincide con l'entrata (j, i) di $-A$, quindi è $-a_{j,i}$. D'altra parte l'entrata (i, j) di $-({}^tA)$ coincide con l'opposto dell'entrata (i, j) di tA , cioè $-a_{i,j}$. □

1.2 Matrici quadrate

In questo paragrafo descriveremo alcune classi notevoli di matrici quadrate.

Definizione 1.10 (Diagonale di una matrice quadrata).

Sia data $A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathbb{R}^{n,n}$. La *diagonale* di A è l'insieme ordinato delle entrate di posizione (i, i) di A .

Esempio 1.11. Si consideri la matrice

$$A = \begin{pmatrix} 1 & -17 & 4 \\ \pi & 0 & 8 \\ 2 & -3/4 & -e \end{pmatrix} \in \mathbb{R}^{3,3}.$$

La diagonale di A è la successione ordinata $(1, 0, -e)$ (e non $(1, -e)$ o $(-e, 1, 0)$ o altro). ♠

Definizione 1.12 (Matrici diagonali).

Sia data $A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathbb{R}^{n,n}$. La matrice A si dice *diagonale* se tutte le entrate al di fuori della diagonale sono nulle, cioè se $a_{i,j} = 0$ per ogni $i \neq j$.

Esempio 1.13. Le matrici

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -e \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -17 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \spadesuit$$

sono rispettivamente diagonale e non diagonale.

Osserviamo che una matrice diagonale può essere descritta indicando solo la sua diagonale: per esempio la matrice A dell'Esempio 1.13 si può indicare come $A = \text{diag}(1, 0, -e)$.

Vi sono due casi particolari di matrici diagonali che vale la pena evidenziare:

- La matrice nulla $0_{n,n}$ è diagonale.
- La matrice diagonale $n \times n$ avente tutte le entrate diagonali uguali a 1 è detta *matrice identità di ordine n* e si indica con I_n . In simboli, l'entrata in posizione (i, j) di I_n coincide con $\delta_{i,j}$, il cosiddetto *simbolo delta di Kronecker*, cioè

$$\delta_{i,j} = \begin{cases} 1 & \text{se } i = j, \\ 0, & \text{se } i \neq j. \end{cases}$$

Ad esempio, le matrici identità I_n , con $n \leq 4$, sono

$$I_1 = (1), \quad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Definizione 1.14 (Matrici triangolari).

Sia data $A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathbb{R}^{n,n}$.

- A si dice *triangolare superiore* se tutte le entrate al di sotto della diagonale sono nulle, cioè se $a_{i,j} = 0$ quando $i > j$.
- A si dice *triangolare inferiore* se le sue entrate al di sopra della diagonale si annullano, ovvero se $a_{i,j} = 0$ quando $i < j$.
- A si dice *strettamente triangolare superiore* se è triangolare superiore e inoltre le sue entrate sulla diagonale si annullano, ovvero se $a_{i,j} = 0$ quando $i \geq j$.
- A si dice *strettamente triangolare inferiore* se è triangolare inferiore e inoltre le sue entrate sulla diagonale si annullano, ovvero se $a_{i,j} = 0$ quando $i \leq j$.

Osserviamo che ogni matrice strettamente triangolare superiore (inferiore) è anche triangolare superiore (inferiore), ma non vale il viceversa. Si noti anche che la matrice nulla $0_{n,n}$ è sia (strettamente) triangolare superiore che inferiore.

Esempio 1.15. Si considerino le due matrici

$$A = \begin{pmatrix} 1 & -17 & 4 \\ 0 & 0 & 8 \\ 0 & 0 & -e \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 0 \\ \pi & 0 & 0 \\ 2 & -3/4 & 0 \end{pmatrix}.$$

A è triangolare superiore, ma non lo è strettamente, mentre B è strettamente triangolare inferiore. ♠

Definizione 1.16 (Matrici simmetriche e antisimmetriche).

Sia data $A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathbb{R}^{n,n}$.

- A si dice *simmetrica* se coincide con la sua trasposta, cioè se ${}^tA = A$ e quindi $a_{i,j} = a_{j,i}$ per ogni $i, j = 1, \dots, n$.
- A si dice *antisimmetrica* se coincide con l'opposto della sua trasposta, cioè se ${}^tA = -A$ e quindi $a_{i,j} = -a_{j,i}$ per ogni $i, j = 1, \dots, n$.

Esempio 1.17. In $\mathbb{R}^{3,3}$ si considerino le matrici

$$A = \begin{pmatrix} 1 & -17 & 4 \\ -17 & 0 & 8 \\ 4 & 8 & -e \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -17 & 4 \\ 4 & 0 & 8 \\ -17 & 8 & -e \end{pmatrix}.$$

La matrice A è simmetrica, mentre B non lo è perché $b_{1,2} = -17 \neq 4 = b_{2,1}$.

Consideriamo adesso

$$C = \begin{pmatrix} 0 & -17 & 4 \\ 17 & 0 & 8 \\ -4 & -8 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & -17 & 4 \\ 17 & 0 & 8 \\ -4 & -8 & 0 \end{pmatrix}.$$

La matrice C è antisimmetrica, mentre D non lo è perché $d_{1,1} = 1 \neq 0$. ♠

Si noti che in una matrice simmetrica le entrate in posizione simmetrica al di fuori della diagonale sono uguali; in una matrice antisimmetrica invece le entrate al di fuori della diagonale sono opposte e le entrate diagonali sono necessariamente nulle, dal momento che devono soddisfare l'uguaglianza $a_{i,i} = -a_{i,i}$.

Ogni matrice diagonale, in particolare la matrice nulla $0_{n,n}$, è simmetrica. Invece non possono essere simmetriche le matrici triangolari superiori e inferiori che non siano diagonali.

L'unica matrice diagonale, o triangolare (superiore e inferiore), o simmetrica che sia anche antisimmetrica è la matrice nulla $0_{n,n}$.

1.3 Somma e prodotto per scalare

In questo paragrafo definiremo due importanti operazioni sulle matrici.

Definizione 1.18 (Somma di matrici).

Siano date $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ e $B = (b_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ due matrici in $\mathbb{R}^{m,n}$. Definiamo *somma di A e B* la matrice di $\mathbb{R}^{m,n}$, indicata con $A + B$, la cui entrata in posizione (i, j) è $a_{i,j} + b_{i,j}$.

⚠ La somma è definita solo per matrici aventi le stesse dimensioni!

Esempio 1.19. In $\mathbb{R}^{3,2}$ si ha

$$\begin{pmatrix} 1 & -1 \\ 5 & 7 \\ 3 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 3 & -4 \\ -1/2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 8 & 3 \\ 5/2 & 2 \end{pmatrix}. \quad \spadesuit$$

La somma di matrici è un'operazione che associa a due matrici delle stesse dimensioni una terza matrice, ancora delle stesse dimensioni. Tale operazione soddisfa una serie di proprietà che ora elencheremo.

Proposizione 1.20 (Proprietà della somma di matrici).

Valgono le seguenti proprietà:

- (S1) per ogni $A, B \in \mathbb{R}^{m,n}$, si ha $A + B = B + A$ (proprietà commutativa);
- (S2) per ogni $A, B, C \in \mathbb{R}^{m,n}$, si ha $A + (B + C) = (A + B) + C$ (proprietà associativa);
- (S3) la matrice nulla è l'unico elemento neutro per la somma, cioè è l'unica matrice tale che $0_{m,n} + A = A + 0_{m,n} = A$ per ogni $A \in \mathbb{R}^{m,n}$ (esistenza e unicità dell'elemento neutro);
- (S4) per ogni $A \in \mathbb{R}^{m,n}$, $-A$ è l'unico elemento opposto di A , cioè è l'unica matrice tale che $A + (-A) = 0_{m,n}$ (esistenza e unicità dell'opposto).

Inoltre:

- (ST) per ogni $A, B \in \mathbb{R}^{m,n}$, si ha ${}^t(A + B) = {}^tA + {}^tB$ (compatibilità con la trasposta).

Dimostrazione. Per dimostrare le uguaglianze tra matrici usiamo la Definizione 1.6; come in precedenza, in tutte le uguaglianze da dimostrare le dimensioni delle matrici a destra e a sinistra dell'uguale coincidono, quindi dobbiamo solo controllare che anche le entrate al posto (i, j) coincidano.

- (S1) L'elemento in posizione (i, j) di $A + B$ è $a_{i,j} + b_{i,j}$. La tesi segue dal fatto che l'elemento in posizione (i, j) di $B + A$ è $b_{i,j} + a_{i,j} = a_{i,j} + b_{i,j}$, perché la somma tra numeri è commutativa.
- (S2) L'elemento in posizione (i, j) di $A + (B + C)$ è $a_{i,j} + (b_{i,j} + c_{i,j})$, mentre quello di $(A + B) + C$ è $(a_{i,j} + b_{i,j}) + c_{i,j}$. L'affermazione segue allora dal fatto che la somma tra numeri è associativa.

- (S3) La verifica che $0_{m,n} + A = A + 0_{m,n} = A$ è immediata. Supponiamo quindi che esista un secondo elemento neutro, ovvero che esista una matrice $X \in \mathbb{R}^{m,n}$ tale che $X + A = A$ per ogni matrice $A \in \mathbb{R}^{m,n}$. Questo significa che per ogni entrata (i, j) vale l'uguaglianza $x_{i,j} + a_{i,j} = a_{i,j}$ e quindi che $x_{i,j} = 0$ per ogni $i = 1, \dots, m$ e $j = 1, \dots, n$ e cioè $X = 0_{m,n}$.
- (S4) Di nuovo, la verifica che $A + (-A) = 0_{m,n}$ è immediata. Supponiamo quindi che esista un secondo elemento opposto, ovvero che esista una matrice $Y \in \mathbb{R}^{m,n}$ tale che $A + Y = 0_{m,n}$. Questo significa che per ogni entrata (i, j) vale l'uguaglianza $a_{i,j} + y_{i,j} = 0$ e quindi che $y_{i,j} = -a_{i,j}$ per ogni $i = 1, \dots, m$ e $j = 1, \dots, n$ e cioè $Y = -A$.
- (ST) Al posto (i, j) di ${}^t(A + B)$ troviamo l'elemento in posizione (j, i) di $A + B$, cioè $a_{j,i} + b_{j,i}$, che coincide con la somma dell'elemento in posizione (i, j) di tA più l'elemento in posizione (i, j) di tB e quindi con l'elemento in posizione (i, j) di ${}^tA + {}^tB$. \square

Se $A, B \in \mathbb{R}^{m,n}$, si scrive $A - B$ invece di $A + (-B)$. Passiamo ora a definire il prodotto di una matrice per un numero reale.

Definizione 1.21 (Prodotto per scalare).

Siano dati $\alpha \in \mathbb{R}$ e $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathbb{R}^{m,n}$. Definiamo *prodotto dello scalare α per la matrice A* la matrice di $\mathbb{R}^{m,n}$, indicata con αA , la cui entrata in posizione (i, j) è $\alpha a_{i,j}$.

Il numero α nel prodotto αA viene spesso detto anche *scalare*.

Esempio 1.22. Si ha

$$2 \begin{pmatrix} 3 & 2 & 5 \\ 1 & -7 & 0 \end{pmatrix} = \begin{pmatrix} 6 & 4 & 10 \\ 2 & -14 & 0 \end{pmatrix}. \quad \spadesuit$$

Il prodotto per scalare è, quindi, un'operazione che associa a una coppia formata da uno scalare e una matrice $m \times n$ un'altra matrice che ha le stesse dimensioni $m \times n$. Anche tale operazione soddisfa una serie di proprietà.

Proposizione 1.23 (Proprietà del prodotto per scalare).

Valgono le seguenti proprietà:

- (P1) per ogni $A \in \mathbb{R}^{m,n}$, si ha $1A = A$ (esistenza dell'elemento neutro);
- (P2) per ogni $A \in \mathbb{R}^{m,n}$ e per ogni $\alpha_1, \alpha_2 \in \mathbb{R}$, si ha $\alpha_1(\alpha_2 A) = (\alpha_1 \alpha_2)A$ (proprietà associativa);
- (SP1) per ogni $A \in \mathbb{R}^{m,n}$ e $\alpha_1, \alpha_2 \in \mathbb{R}$, si ha $(\alpha_1 + \alpha_2)A = \alpha_1 A + \alpha_2 A$ (proprietà distributiva rispetto alla somma di scalari);
- (SP2) per ogni $A, B \in \mathbb{R}^{m,n}$ e $\alpha \in \mathbb{R}$, si ha $\alpha(A + B) = \alpha A + \alpha B$ (proprietà distributiva rispetto alla somma di matrici).

Inoltre:

- (PT) per ogni $A \in \mathbb{R}^{m,n}$ e $\alpha \in \mathbb{R}$, si ha ${}^t(\alpha A) = \alpha({}^tA)$ (compatibilità con la trasposta);

(LP) per ogni $A \in \mathbb{R}^{m,n}$ e $\alpha \in \mathbb{R}$, risulta $\alpha A = 0_{m,n}$ se e solo se vale $\alpha = 0$ oppure $A = 0_{m,n}$ (legge di annullamento del prodotto per scalare).

Dimostrazione. Per dimostrare le uguaglianze tra matrici usiamo la Definizione 1.6; come in precedenza, in tutte le uguaglianze da dimostrare le dimensioni delle matrici a destra e a sinistra dell'uguale coincidono, quindi dobbiamo solo controllare che le entrate al posto (i, j) coincidano. Come nella dimostrazione della Proposizione 1.20, molte proprietà del prodotto per scalare si basano su proprietà simili a quelle della somma e del prodotto tra numeri.

(P1) L'entrata (i, j) di $1A$ è $1a_{i,j} = a_{i,j}$.

(P2) Si ha che $\alpha_1(\alpha_2 a_{i,j}) = (\alpha_1 \alpha_2) a_{i,j}$, che è l'elemento in posizione (i, j) della matrice a destra.

(SP1) L'entrata (i, j) di $(\alpha_1 + \alpha_2)A$ è $(\alpha_1 + \alpha_2)a_{i,j} = \alpha_1 a_{i,j} + \alpha_2 a_{i,j}$, che è l'elemento (i, j) di $\alpha_1 A + \alpha_2 A$.

(SP2) Si ha che $\alpha(a_{i,j} + b_{i,j}) = \alpha a_{i,j} + \alpha b_{i,j}$, da cui l'uguaglianza delle matrici.

(PT) L'elemento in posizione (i, j) di ${}^t(\alpha A)$ è $\alpha a_{j,i}$ e lo stesso vale per $\alpha({}^t A)$.

(LP) Se il prodotto $\alpha A = 0_{m,n}$ significa che tutte le entrate della matrice αA sono nulle, cioè che per ogni (i, j) il prodotto $\alpha a_{i,j} = 0$; per la legge di annullamento del prodotto tra numeri questo implica che o $\alpha = 0$, oppure $a_{i,j} = 0$ per ogni indice (i, j) e quindi A è necessariamente la matrice nulla. \square

In generale possiamo dire che le operazioni di somma tra matrici e prodotto per scalare si comportano in modo non troppo diverso dalla somma e dal prodotto tra due numeri cui siamo abituati. Nel prossimo capitolo descriveremo una terza operazione, detta *prodotto tra matrici*, che invece non avrà tali caratteristiche.

Operazioni fra matrici: prodotto righe per colonne

2.1 Prodotto di matrici

In questo paragrafo introdurremo una nuova operazione detta *prodotto tra matrici*, o anche *prodotto righe per colonne*. Procederemo per gradi: definiremo prima il prodotto di una matrice riga per una matrice colonna, e solo in seguito daremo la definizione generale.

Definizione 2.1 (Prodotto di una matrice riga per una matrice colonna).

Siano date $R = (r_{1,j})_{1 \leq j \leq p} \in \mathbb{R}^{1,p}$ e $C = (c_{i,1})_{1 \leq i \leq p} \in \mathbb{R}^{p,1}$. Il *prodotto delle matrici* R e C (nell'ordine indicato) è la matrice di $\mathbb{R}^{1,1}$, indicata con RC , definita da

$$RC = (r_{1,1} \quad r_{1,2} \quad \dots \quad r_{1,p}) \begin{pmatrix} c_{1,1} \\ c_{2,1} \\ \vdots \\ c_{p,1} \end{pmatrix} = (r_{1,1}c_{1,1} + r_{1,2}c_{2,1} + \dots + r_{1,p}c_{p,1}) \in \mathbb{R}^{1,1}.$$

In particolare, il prodotto di una matrice riga per una matrice colonna è una matrice 1×1 e si può identificare con uno scalare.

Esempio 2.2. Siano

$$A = (1 \quad 2 \quad -1) \in \mathbb{R}^{1,3}, \quad B = \begin{pmatrix} 3 \\ 2 \\ -5 \end{pmatrix} \in \mathbb{R}^{3,1},$$

allora

$$AB = (1 \quad 2 \quad -1) \begin{pmatrix} 3 \\ 2 \\ -5 \end{pmatrix} = (1 \cdot 3 + 2 \cdot 2 + (-1) \cdot (-5)) = (12) \in \mathbb{R}^{1,1}.$$

Se, invece,

$$C = (1 \quad 3 \quad -5 \quad 1) \in \mathbb{R}^{1,4}, \quad D = \begin{pmatrix} 4 \\ 1 \\ 2 \\ 3 \end{pmatrix} \in \mathbb{R}^{4,1},$$

si ha

$$CD = (1 \quad 3 \quad -5 \quad 1) \begin{pmatrix} 4 \\ 1 \\ 2 \\ 3 \end{pmatrix} = (1 \cdot 4 + 3 \cdot 1 + (-5) \cdot 2 + 1 \cdot 3) = 0_{1,1},$$

cioè il prodotto CD è nullo nonostante nessuna delle due matrici che stiamo moltiplicando lo sia. ♠

Passiamo ora alla definizione generale del prodotto tra matrici; dal momento che abbiamo imparato a fare il prodotto di una riga per una colonna, l'idea è quella di pensare la prima matrice come unione delle sue righe e la seconda come unione delle sue colonne, moltiplicandole nel modo descritto nella seguente definizione.

Definizione 2.3 (Prodotto tra matrici).

Siano date $A = (a_{i,h})_{\substack{1 \leq i \leq m \\ 1 \leq h \leq p}} \in \mathbb{R}^{m,p}$ e $B = (b_{h,j})_{\substack{1 \leq h \leq p \\ 1 \leq j \leq n}} \in \mathbb{R}^{p,n}$. Per ogni $i = 1, \dots, m$ sia

$$R_i = (a_{i,1} \quad a_{i,2} \quad \dots \quad a_{i,p}) \in \mathbb{R}^{1,p}$$

la i -esima riga di A . Similmente per ogni $j = 1, \dots, n$ sia

$$C_j = \begin{pmatrix} b_{1,j} \\ b_{2,j} \\ \vdots \\ b_{p,j} \end{pmatrix} \in \mathbb{R}^{p,1}$$

la j -esima colonna di B .

Il *prodotto delle matrici* A e B è la matrice di $\mathbb{R}^{m,n}$, indicata con AB , la cui entrata in posizione (i, j) è $R_i C_j$.

Osservazione 2.4. Mettendo insieme le Definizioni 2.1 e 2.3, se $A = (a_{i,h})_{\substack{1 \leq i \leq m \\ 1 \leq h \leq p}}$ e $B = (b_{h,j})_{\substack{1 \leq h \leq p \\ 1 \leq j \leq n}}$, allora l'entrata in posizione (i, j) della matrice prodotto AB è

$$\sum_{k=1}^p a_{i,k} b_{k,j}.$$

Si ricordi che $1 \leq i \leq m$ e $1 \leq j \leq n$.

Esempio 2.5. Riprendiamo le matrici A e B dell'Esempio 2.2. Poiché $A \in \mathbb{R}^{1,3}$ e $B \in \mathbb{R}^{3,1}$ è definito il prodotto BA e risulta

$$\begin{aligned} BA &= \begin{pmatrix} 3 \\ 2 \\ -5 \end{pmatrix} (1 \ 2 \ -1) \\ &= \begin{pmatrix} 3 \cdot 1 & 3 \cdot 2 & 3 \cdot (-1) \\ 2 \cdot 1 & 2 \cdot 2 & 2 \cdot (-1) \\ (-5) \cdot 1 & (-5) \cdot 2 & (-5) \cdot (-1) \end{pmatrix} = \begin{pmatrix} 3 & 6 & -3 \\ 2 & 4 & -2 \\ -5 & -10 & 5 \end{pmatrix}. \quad \spadesuit \end{aligned}$$

⚠ Si noti che, affinché il prodotto di due matrici sia definito, è necessario e sufficiente che il numero di colonne della prima coincida con il numero di righe della seconda! Inoltre se $A \in \mathbb{R}^{m,p}$ e $B \in \mathbb{R}^{p,n}$ allora $AB \in \mathbb{R}^{m,n}$.

Quindi il prodotto tra matrici non sempre si può fare. Per esempio se

$$A = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \in \mathbb{R}^{2,1}, \quad B = \begin{pmatrix} 1 & 3 & -1 \\ 0 & 2 & 1 \end{pmatrix} \in \mathbb{R}^{2,3},$$

non è definito né AB né BA .

Anche se due matrici si possono moltiplicare in un certo ordine, non è detto che si possano moltiplicare nell'ordine opposto. Per esempio, se

$$C = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \in \mathbb{R}^{2,2}, \quad D = \begin{pmatrix} 1 & 3 & -1 \\ 0 & 2 & 1 \end{pmatrix} \in \mathbb{R}^{2,3},$$

solo il prodotto $CD \in \mathbb{R}^{2,3}$ è definito, mentre DC non si può fare. Se facciamo i conti abbiamo

$$\begin{aligned} CD &= \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 3 & -1 \\ 0 & 2 & 1 \end{pmatrix} = \\ &= \begin{pmatrix} 1 \cdot 1 + 1 \cdot 0 & 1 \cdot 3 + 1 \cdot 2 & 1 \cdot (-1) + 1 \cdot 1 \\ 1 \cdot 1 + (-1) \cdot 0 & 1 \cdot 3 + (-1) \cdot 2 & 1 \cdot (-1) + (-1) \cdot 1 \end{pmatrix} = \begin{pmatrix} 1 & 5 & 0 \\ 1 & 1 & -2 \end{pmatrix}. \end{aligned}$$

Un altro caso che si può verificare è che due matrici si possono moltiplicare in entrambi gli ordini, ma i due prodotti hanno dimensioni diverse. Ad esempio, se

$$E = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 2 & 1 \end{pmatrix} \in \mathbb{R}^{2,3}, \quad F = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 2 & -2 \end{pmatrix} \in \mathbb{R}^{3,2},$$

si verifichi che

$$EF = \begin{pmatrix} 3 & -1 \\ 4 & -3 \end{pmatrix} \in \mathbb{R}^{2,2}, \quad FE = \begin{pmatrix} -1 & 2 & 1 \\ 1 & 1 & 1 \\ 4 & -2 & 0 \end{pmatrix} \in \mathbb{R}^{3,3}.$$

Anche quando entrambi i prodotti sono definiti e hanno le stesse dimensioni, non è detto che coincidano (cioè il prodotto tra matrici non è commutativo). Se

$$G = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 0 & 3 \\ 2 & 0 \end{pmatrix},$$

entrambi i prodotti GH e HG sono definiti e hanno dimensione 2×2 , tuttavia

$$GH = \begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix} = HG.$$

Un'osservazione interessante è che la non commutatività del prodotto di matrici non permette di utilizzare gli usuali "prodotti notevoli" nel fare i calcoli di prodotti di matrici. Per illustrare questa osservazione diamo la seguente definizione.

Definizione 2.6 (Potenze di una matrice quadrata).

Sia data $A \in \mathbb{R}^{n,n}$. Si pone $A^1 = A$: per ogni intero $p \geq 2$ la *potenza p -esima* della matrice A è la matrice $A^p = A^{p-1}A$.

In particolare $A^2 = AA$, $A^3 = A^2A = AAA$ e così via.

Esempio 2.7. Non è vero che, data una coppia di matrici A, B per cui abbiano senso entrambe le espressioni matriciali $(A + B)^2$ e $A^2 + 2AB + B^2$, queste coincidano. Come controesempio consideriamo le matrici

$$E_{1,1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad E_{1,2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

e calcoliamo

$$(E_{1,1} + E_{1,2})^2 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

Invece

$$\begin{aligned} E_{1,1}^2 + 2E_{1,1}E_{1,2} + E_{1,2}^2 &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}^2 + 2 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^2 = \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

In maniera simile si può verificare che, in generale, $A^2 - B^2 \neq (A + B)(A - B)$, $A^3 \pm B^3 \neq (A \pm B)(A^2 \mp AB + B^2)$ e così via. ♠

Nonostante quanto visto sopra, il prodotto di matrici soddisfa molte proprietà importanti.

Proposizione 2.8 (Proprietà del prodotto tra matrici).

Valgono le seguenti proprietà:

(PM1) per ogni $A \in \mathbb{R}^{m,p}$, $B \in \mathbb{R}^{p,q}$, $C \in \mathbb{R}^{q,n}$ si ha $A(BC) = (AB)C$ (proprietà associativa);

- (PM2) la matrice identità $I_p \in \mathbb{R}^{p,p}$ è l'unico elemento neutro per il prodotto, cioè l'unica matrice tale che $AI_p = A$ e $I_p B = B$ per ogni $A \in \mathbb{R}^{m,p}$ e $B \in \mathbb{R}^{p,n}$ (esistenza e unicità dell'elemento neutro);
- (PMP) per ogni $A \in \mathbb{R}^{m,p}$ e $B \in \mathbb{R}^{p,n}$ e $\alpha \in \mathbb{R}$, si ha $\alpha(AB) = (\alpha A)B = A(\alpha B)$ (compatibilità col prodotto per scalare);
- (PMS1) per ogni $A, B \in \mathbb{R}^{m,p}$ e $C \in \mathbb{R}^{p,n}$, si ha $(A + B)C = AC + BC$ (proprietà distributiva a destra);
- (PMS2) per ogni $A \in \mathbb{R}^{m,p}$ e $B, C \in \mathbb{R}^{p,n}$, si ha $A(B + C) = AB + AC$ (proprietà distributiva a sinistra);
- (PMT) per ogni $A \in \mathbb{R}^{m,p}$ e $B \in \mathbb{R}^{p,n}$, si ha ${}^t(AB) = {}^t B {}^t A$ (compatibilità con la trasposta).

Dimostrazione. Per dimostrare le uguaglianze tra matrici usiamo la Definizione 1.6; come in precedenza, in tutte le uguaglianze da dimostrare le dimensioni delle matrici a destra e a sinistra dell'uguale coincidono, quindi dobbiamo solo controllare che coincidano anche le entrate al posto (i, j) .

(PM1) In posizione (i, j) della matrice $A(BC)$ troviamo

$$\sum_{k=1}^p a_{i,k} \left(\sum_{\ell=1}^q b_{k,\ell} c_{\ell,j} \right) = \sum_{k=1}^p \sum_{\ell=1}^q a_{i,k} b_{k,\ell} c_{\ell,j} = \sum_{\ell=1}^q \left(\sum_{k=1}^p a_{i,k} b_{k,\ell} \right) c_{\ell,j}$$

che coincide con l'entrata in posizione (i, j) della matrice $(AB)C$.

(PM2) L'identità $AI_p = A$ è immediata dalla definizione. Per dimostrare l'unicità, supponiamo che esista un'altra matrice quadrata $X = (x_{i,j})_{1 \leq i,j \leq p} \in \mathbb{R}^{p,p}$ tale che $AX = A$ per ogni matrice $A \in \mathbb{R}^{m,p}$. In particolare per la matrice

$$A_1 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{m,p}$$

deve valere che

$$A_1 X = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,p} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} = A_1,$$

per cui necessariamente $x_{1,1} = 1$ e $x_{1,j} = 0$ per $2 \leq j \leq p$. Ripetendo il ragionamento con la matrice

$$A_2 = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{m,p}$$

ricaviamo che $x_{2,2} = 1$ e $x_{2,j} = 0$ per $j \neq 2$, e così via, fino a dimostrare che $x_{i,j} = \delta_{i,j}$, ovvero che $X = I_p$.

Una dimostrazione analoga, che omettiamo, vale per il prodotto BI_p .

(PMP) In posizione (i, j) troviamo l'elemento:

$$\alpha \left(\sum_{k=1}^p a_{i,k} b_{k,j} \right) = \sum_{k=1}^p (\alpha a_{i,k}) b_{k,j} = \sum_{k=1}^p a_{i,k} (\alpha b_{k,j}).$$

(PMS1) In posizione (i, j) della matrice $(A + B)C$ troviamo l'elemento:

$$\sum_{k=1}^p (a_{i,k} + b_{i,k}) c_{k,j} = \sum_{k=1}^p (a_{i,k} c_{k,j}) + (b_{i,k} c_{k,j}) = \sum_{k=1}^p a_{i,k} c_{k,j} + \sum_{k=1}^p b_{i,k} c_{k,j}$$

che coincide con l'entrata in posizione (i, j) della matrice $AC + BC$.

(PMS2) In posizione (i, j) della matrice $A(B + C)$ troviamo l'elemento:

$$\sum_{k=1}^p a_{i,k} (b_{k,j} + c_{k,j}) = \sum_{k=1}^p (a_{i,k} b_{k,j}) + (a_{i,k} c_{k,j}) = \sum_{k=1}^p a_{i,k} b_{k,j} + \sum_{k=1}^p a_{i,k} c_{k,j}$$

che coincide con l'entrata in posizione (i, j) della matrice $AB + AC$.

(PMT) L'entrata (i, j) di ${}^t B {}^t A$ è

$$\sum_{k=1}^p b_{k,i} a_{j,k} = \sum_{k=1}^p a_{j,k} b_{k,i},$$

che è proprio l'entrata (i, j) di ${}^t(AB)$. □

⚠ Osserviamo che quando trasponiamo un prodotto l'ordine di moltiplicazione si inverte! Non potrebbe essere altrimenti, visto che se $A \in \mathbb{R}^{m,p}$ e $B \in \mathbb{R}^{p,n}$ e quindi è definito il prodotto AB , si ha che ${}^t A \in \mathbb{R}^{p,m}$ e ${}^t B \in \mathbb{R}^{n,p}$, quindi è definito il prodotto ${}^t B {}^t A$.

2.2 Complessità computazionale del prodotto di matrici

Ora che sappiamo moltiplicare due matrici, possiamo osservare che eseguendo un prodotto seguiamo una serie di istruzioni successive, ovvero un *algoritmo*, e che quindi possiamo facilmente far fare queste operazioni a un computer. Ovviamente vorremmo che questo computer moltiplicasse le matrici in modo più veloce possibile. Quanto veloce possiamo farlo lavorare?

Prendiamo il primo caso interessante di matrici quadrate, quello della moltiplicazione di matrici 2×2 . Se

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}, \quad B = \begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix},$$

l'algoritmo per calcolare $C = AB$ ci dice che

$$\begin{aligned} c_{1,1} &= a_{1,1} b_{1,1} + a_{1,2} b_{2,1}, & c_{1,2} &= a_{1,1} b_{1,2} + a_{1,2} b_{2,2}, \\ c_{2,1} &= a_{2,1} b_{1,1} + a_{2,2} b_{2,1}, & c_{2,2} &= a_{2,1} b_{1,2} + a_{2,2} b_{2,2}, \end{aligned}$$

e quindi l'algoritmo ci chiede di eseguire 4 addizioni e 8 moltiplicazioni.

Dal punto di vista della memoria del computer addizioni e moltiplicazioni sono molto diverse, in quanto le prime non usano praticamente memoria, mentre le seconde sì. La *complessità computazionale* di un algoritmo si misura quindi col numero di moltiplicazioni che vanno eseguite, e quanto scritto sopra si può tradurre dicendo che il costo computazionale di moltiplicare due matrici 2×2 con l'algoritmo del prodotto riga per colonna è di 8 moltiplicazioni. È possibile fare meglio?

Nel 1969 il matematico tedesco V. Strassen, mentre cercava di dimostrare che non era possibile usare meno di 8 moltiplicazioni, trovò un algoritmo che invece ne usava solo 7. Tale algoritmo viene oggi proprio chiamato *algoritmo di Strassen*.

Vediamo di cosa si tratta. Definiamo le seguenti quantità

$$\begin{aligned} I &= (a_{1,1} + a_{2,2})(b_{1,1} + b_{2,2}), \\ II &= (a_{2,1} + a_{2,2})b_{1,1}, \\ III &= a_{1,1}(b_{1,2} - b_{2,2}), \\ IV &= a_{2,2}(-b_{1,1} + b_{2,1}), \\ V &= (a_{1,1} + a_{1,2})b_{2,2}, \\ VI &= (-a_{1,1} + a_{2,1})(b_{1,1} + b_{1,2}), \\ VII &= (a_{1,2} - a_{2,2})(b_{2,1} + b_{2,2}). \end{aligned}$$

Si verifichi che

$$\begin{aligned} c_{1,1} &= I + IV - V + VII, & c_{1,2} &= II + IV, \\ c_{2,1} &= III + V, & c_{2,2} &= I + III - II + VI. \end{aligned}$$

Questa scrittura che a noi sembra inutilmente complicata, a un computer permette di eseguire la moltiplicazione di matrici molto più velocemente!

Non solo: se moltiplichiamo riga per colonna due matrici $n \times n$, dobbiamo eseguire $n^3 - n^2$ addizioni e soprattutto n^3 moltiplicazioni. Però, se abbiamo matrici 4×4 possiamo dividerle in blocchi 2×2 e usare l'algoritmo di Strassen per ogni blocco, riducendo il numero di moltiplicazioni da $4^3 = 64$ a $7^2 = 49$. Se abbiamo matrici di dimensione $2^k \times 2^k$ Strassen ci permette di ridurre il numero di moltiplicazioni da 8^k a 7^k ; se le matrici hanno una dimensione che non è una potenza di 2 possiamo sempre aggiungere blocchi di zeri e ricondurci al caso $2^k \times 2^k$.

Si può dimostrare che, mentre l'algoritmo classico permette di eseguire un prodotto di matrici $n \times n$ usando n^3 moltiplicazioni, Strassen permette di usarne $n^{\log_2(7)} \simeq n^{2,81}$. L'algoritmo di Strassen (e le sue varianti) sono effettivamente usati ancora oggi nei casi in cui la dimensione delle matrici da moltiplicare sia molto grande.

È sicuramente possibile fare meglio: infatti nel 1990 gli informatici sono riusciti a scendere a $n^{2,375}$, utilizzando anche altri strumenti. Nondimeno, il valore minimo di ω tale che la complessità computazionale sia dell'ordine di n^ω è tuttora oggetto di indagine.

2.3 Matrici invertibili e inversa di una matrice quadrata

Quando si ha a che fare con un prodotto è naturale porsi il problema dell'esistenza dell'elemento inverso. Nel caso del prodotto tra scalari sappiamo che ogni numero reale a non nullo ammette un inverso a^{-1} , ovvero un altro numero reale non nullo univocamente determinato tale che $a \cdot a^{-1} = a^{-1} \cdot a = 1$.

Vediamo cosa succede nel caso delle matrici; in questa trattazione ci ridurremo a considerare *matrici quadrate*.

Definizione 2.9 (Matrici invertibili).

Sia data $A \in \mathbb{R}^{n,n}$. A è detta *invertibile* se esiste $B \in \mathbb{R}^{n,n}$ tale che $AB = BA = I_n$.

Esempio 2.10. La matrice

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

è invertibile. Infatti la matrice

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

è tale che $BA = AB = I_2$. ♠

⚠ Contrariamente a quanto accade per i numeri reali, non tutte le matrici quadrate non nulle sono invertibili!

Esempio 2.11. La matrice

$$C = \begin{pmatrix} 1 & 2 \\ -2 & -4 \end{pmatrix}$$

non è invertibile. Infatti se esistesse una matrice

$$D = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}^{2,2}$$

tale che $CD = DC = I_2$, calcolando il prodotto CD , avremmo che

$$CD = \begin{pmatrix} 1 & 2 \\ -2 & -4 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a + 2c & b + 2d \\ -2a - 4c & -2b - 4d \end{pmatrix}.$$

Se però esistessero $a, c \in \mathbb{R}$ tali che $a + 2c = 1$, allora si dovrebbe anche avere che $-2a - 4c = -2(a + 2c) = -2 \neq 0$, quindi il prodotto CD non può essere uguale alla matrice identità per alcun valore di $a, b, c, d \in \mathbb{R}$.

Esiste un'interessante classe di matrici quadrate che non sono invertibili, le matrici nilpotenti: una matrice $A \in \mathbb{R}^{n,n}$ è detta *nilpotente* se esiste un intero $p \geq 1$ tale che $A^p = 0_{n,n}$. Per esempio la matrice

$$E_{1,2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

è tale che $E_{1,2}^2 = 0_{2,2}$, e quindi nilpotente.

Mostriamo che nessuna matrice nilpotente A può avere un'inversa. Se esistesse una matrice B tale che $AB = I_n$ allora

$$A^{p-1} = A^{p-1}I_n = A^{p-1}AB = A^pB = 0_{n,n}B = 0_{n,n}.$$

Poiché $A^{p-1} = 0_{n,n}$, sostituendo A^{p-1} ad A^p in quanto scritto sopra, otterremo $A^{p-2} = 0_{n,n}$. Ripetendo il ragionamento otterremo alla fine $AB = 0_{n,n}$, in contraddizione con l'ipotesi che A sia invertibile. ♠

Finora abbiamo parlato di “una” inversa di una matrice invertibile. Vediamo adesso che invece “la” matrice inversa, come succede nel caso degli scalari, è univocamente determinata.

Proposizione 2.12 (Proprietà delle matrici invertibili).

Valgono le seguenti proprietà:

- (MI1) per ogni $A \in \mathbb{R}^{n,n}$ si ha che $B \in \mathbb{R}^{n,n}$ è tale che $AB = I_n$ se e solo se $BA = I_n$;
- (MI2) per ogni $A \in \mathbb{R}^{n,n}$ se esistono due matrici $B, C \in \mathbb{R}^{n,n}$ tali che $AB = BA = I_n$ e $AC = CA = I_n$, allora $B = C$.

Dimostrazione. L'unicità nella proprietà (MI2) è conseguenza della catena di uguaglianze

$$B = BI_n = B(AC) = (BA)C = I_nC = C.$$

Omettiamo la dimostrazione della prima proprietà (MI1). □

Grazie alla Proposizione 2.12 siamo finalmente pronti a dare la seguente definizione.

Definizione 2.13 (Matrice inversa).

Sia data $A \in \mathbb{R}^{n,n}$ invertibile. L'unica matrice $B \in \mathbb{R}^{n,n}$ tale che $AB = BA = I_n$ viene detta *inversa di A* e viene denotata con A^{-1} . In tal caso si pone anche $A^0 = I_n$ e $A^{-p} = (A^{-1})^p$ per ogni intero $p \geq 1$.

Siano A e $B \in \mathbb{R}^{n,n}$ due matrici invertibili. Cosa si può dire circa l'invertibilità della trasposta tA ? e circa l'invertibilità del prodotto AB ?

Proposizione 2.14 (Inversa della trasposta e inversa del prodotto).

Valgono le seguenti proprietà:

- (MIT) sia $A \in \mathbb{R}^{n,n}$, allora A è invertibile se e solo se la trasposta tA è invertibile e si ha che $({}^tA)^{-1} = {}^t(A^{-1})$;
- (MIP) siano $A, B \in \mathbb{R}^{n,n}$, allora A e B sono invertibili se e solo se il prodotto AB è invertibile e si ha che $(AB)^{-1} = B^{-1}A^{-1}$.

Dimostrazione.

- (MIT) La matrice trasposta tA è invertibile se e solo se esiste una matrice X tale che ${}^tAX = I_n$. Trasponendo entrambi i membri dell'equazione troviamo che questo equivale a $({}^tX)A = {}^tI_n$, e visto che ${}^tI_n = I_n$ questo equivale a sua volta all'invertibilità di A . Inoltre $({}^tX) = A^{-1}$, e cioè $({}^tA)^{-1} = {}^t(A^{-1})$. (Osserviamo che abbiamo usato le proprietà della Proposizione 2.12!)
- (MIP) Supponiamo prima che A e B siano entrambe invertibili con inversa A^{-1} e B^{-1} rispettivamente. Allora

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n,$$

e similmente $(B^{-1}A^{-1})(AB) = I_n$, cioè AB è invertibile con inversa $(AB)^{-1} = B^{-1}A^{-1}$. Viceversa, supponiamo che AB sia invertibile, ovvero che esista una matrice X tale che $(AB)X = X(AB) = I_n$. Allora vale che $A(BX) = I_n$, cioè A è invertibile con inversa $A^{-1} = BX$, e similmente $(XA)B = I_n$, cioè B è invertibile con inversa $B^{-1} = XA$. \square

2.4 Algebra lineare sui numeri complessi

Concludiamo questo capitolo con un'osservazione importante: le nozioni e le operazioni definite finora sulle matrici a coefficienti in \mathbb{R} si possono estendere a matrici a coefficienti nel campo dei numeri complessi \mathbb{C} . Poiché $\mathbb{R} \subseteq \mathbb{C}$, segue che $\mathbb{R}^{m,n} \subseteq \mathbb{C}^{m,n}$.

Per questo motivo, da adesso in poi, nelle definizioni e negli enunciati delle proposizioni spesso sostituirò al simbolo \mathbb{R} il simbolo \mathbb{K} che indicherà, o il campo reale \mathbb{R} o il campo complesso \mathbb{C} .

Più un generale \mathbb{K} potrà anche indicare un qualsiasi campo, cioè un qualsiasi insieme non vuoto in cui siano definite due operazioni binarie chiamate somma e prodotto che abbiano le stesse proprietà formali della somma o prodotto dei numeri reali e complessi (commutatività, associatività, distributività, esistenza di un elemento neutro, esistenza degli opposti, esistenza degli inversi di elementi non nulli).

Esempio 2.15. Si considerino le matrici in $\mathbb{C}^{2,2}$

$$A = \begin{pmatrix} 1 - 2i & 2 - i \\ 1 - i & 1 - i \end{pmatrix} \quad B = \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}.$$

A e B possono essere moltiplicate, e risulta

$$AB = \begin{pmatrix} 1 - 2i & 2 - i \\ 1 - i & 1 - i \end{pmatrix} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 2 & 2 \end{pmatrix}.$$

Si noti che mentre $A, B \in \mathbb{C}^{2,2} \setminus \mathbb{R}^{2,2}$, il prodotto AB è in $\mathbb{R}^{2,2}$ (e, quindi, anche in $\mathbb{C}^{2,2}$). \spadesuit

Sistemi lineari e loro rappresentazione matriciale

3.1 Sistemi di equazioni lineari

Definizione 3.1 (Sistemi di equazioni lineari e loro soluzioni).

- Un'equazione lineare nelle n incognite x_1, x_2, \dots, x_n a coefficienti in \mathbb{K} è un'equazione della forma

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b, \quad (3.1.1)$$

dove $a_j, b \in \mathbb{K}$ sono numeri fissati, e $j = 1, \dots, n$.

- Un sistema di m equazioni lineari nelle n incognite x_1, x_2, \dots, x_n a coefficienti in \mathbb{K} è un insieme di m equazioni lineari della forma

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,n}x_n &= b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \dots + a_{2,n}x_n &= b_2 \\ &\vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \dots + a_{m,n}x_n &= b_m \end{aligned} \quad (3.1.2)$$

dove $a_{i,j}, b_i \in \mathbb{K}$ sono numeri fissati, $i = 1, \dots, m$ e $j = 1, \dots, n$.

- Una soluzione dell'equazione (3.1.1) è una successione ordinata $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ di elementi di \mathbb{K} per cui vale l'identità numerica $a_1\bar{x}_1 + a_2\bar{x}_2 + \dots + a_n\bar{x}_n = b$.
- Similmente, una soluzione del sistema (3.1.2) è una soluzione che soddisfa ciascuna equazione del sistema.

Esempio 3.2. L'equazione

$$2x_1 - x_2 = 2$$

è lineare in 2 incognite: le coppie ordinate $(1, 0)$, $(0, -2)$, $(-1, -4)$ sono sue soluzioni. Invece le coppie ordinate $(0, 1)$, $(-2, 0)$, $(-4, -1)$ non sono soluzioni.

Il sistema

$$\begin{aligned} 3x + y - z &= 0 \\ x + y - 3z &= -5 \\ x + y &= 1 \end{aligned}$$

è lineare di 3 equazioni in 3 incognite. Una sua soluzione è la successione ordinata $(1/2, 1/2, 2)$. ♠

Convenzionalmente, a un sistema della forma (3.1.2), viene aggiunta una parentesi graffa che indica quali equazioni devono essere considerate: perciò un sistema di m equazioni lineari nelle n incognite x_1, x_2, \dots, x_n a coefficienti in \mathbb{K} viene spesso denotato con

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \dots + a_{2,n}x_n = b_2 \\ \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \dots + a_{m,n}x_n = b_m. \end{cases}$$

⚠ Le due scritture qui sotto hanno dunque significati diversi:

$$\begin{cases} 2x_1 - x_2 = 3 \\ x_1 + x_2 = 0 \\ -x_1 + x_2 = -2, \end{cases} \quad \begin{cases} 2x_1 - x_2 = 3 \\ x_1 + x_2 = 0 \\ -x_1 + x_2 = -2 \end{cases}$$

Quando il numero di incognite è basso spesso per indicare le incognite si usano le lettere distinte $x, y, z, t \dots$ o anche $a, b, c \dots$ invece di x_i .

Definizione 3.3 (Matrici associate a un sistema lineare).

Dato un sistema lineare di m equazioni in n incognite del tipo (3.1.2) a esso sono naturalmente associate le due matrici

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Le matrici A e B vengono dette rispettivamente *matrice dei coefficienti* del sistema (o anche *matrice incompleta* del sistema) e *matrice dei termini noti* del sistema. La matrice

$$(A|B) = \left(\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \dots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} & b_m \end{array} \right)$$

è detta *matrice completa* del sistema. La linea verticale serve a ricordare che ciò che si trova alla sua sinistra è la matrice incompleta del sistema e ciò che si trova alla sua destra è la colonna dei termini noti; la sua utilità (o necessità) sarà chiara più avanti.

La notazione matriciale è molto vantaggiosa e si può usare per descrivere in forma compatta il sistema (3.1.2); possiamo infatti porre

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

e questo ci permette di scrivere il sistema (3.1.2) in forma matriciale come

$$AX = B.$$

In particolare, con tale formalismo, le soluzioni del sistema (3.1.2) sono esattamente le matrici numeriche

$$\bar{X} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_n \end{pmatrix}$$

tali che valga l'identità matriciale $A\bar{X} = B$.

Esempio 3.4. Consideriamo il sistema

$$\begin{cases} 2x_1 - x_2 = 3 \\ x_1 + x_2 = 0 \\ -x_1 + x_2 = -2. \end{cases}$$

La matrice completa del sistema è

$$(A|B) = \left(\begin{array}{cc|c} 2 & -1 & 3 \\ 1 & 1 & 0 \\ -1 & 1 & -2 \end{array} \right).$$

Il sistema può essere scritto in forma matriciale $AX = B$ come

$$\begin{pmatrix} 2 & -1 \\ 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ -2 \end{pmatrix}.$$

Si noti che ${}^t(1 \ -1)$ è soluzione del sistema poiché

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

Invece ${}^t(-1 \ 1)$ non è soluzione del sistema: infatti

$$\begin{pmatrix} 2 & -1 \\ 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 0 \\ 2 \end{pmatrix} \neq \begin{pmatrix} 3 \\ 0 \\ -2 \end{pmatrix}. \quad \spadesuit$$

Introduciamo adesso un po' di terminologia.

Definizione 3.5 (Sistemi omogenei e sistemi compatibili).

Un sistema di m equazioni in n incognite del tipo (3.1.2) si dice:

- *omogeneo* se $B = 0_{m,1}$, *non omogeneo* se $B \neq 0_{m,1}$;
- *compatibile* se ha soluzioni, *non compatibile* o *incompatibile* se non ha soluzioni.

Inoltre ogni sistema lineare di m equazioni in n incognite $AX = B$ ha un suo *sistema omogeneo associato* della forma $AX = 0_{m,1}$.

Esempio 3.6. Il sistema dell'Esempio 3.4 è, ovviamente, non omogeneo; inoltre è compatibile poiché ammette ${}^t(1 \ -1)$ come soluzione.

Se $A \in \mathbb{K}^{m,n}$ allora ogni sistema omogeneo $AX = 0_{m,1}$ è compatibile, in quanto $0_{n,1}$ è sempre una sua soluzione.

Invece il sistema

$$\begin{pmatrix} 2 & -1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ -2 \end{pmatrix}$$

è non omogeneo e non può essere compatibile: infatti se (\bar{x}_1, \bar{x}_2) fosse una sua soluzione si dovrebbe avere contemporaneamente $\bar{x}_1 + \bar{x}_2 = 0$ e $\bar{x}_1 + \bar{x}_2 = -2$, cosa che, ovviamente, è impossibile. ♠

3.2 Matrici (fortemente) ridotte per righe

Avendo introdotto le nozioni di sistema di equazioni lineari e relativa soluzione, ci poniamo il problema di descrivere un metodo efficiente per determinare la totalità delle soluzioni di un dato sistema.

A tale scopo iniziamo considerando un semplice esempio.

Esempio 3.7. Si consideri il sistema

$$\begin{cases} x_1 + x_2 + 2x_3 + x_5 = 1 \\ -x_2 + x_4 + x_7 = 0 \\ 3x_2 + x_6 + 3x_7 = -2, \end{cases}$$

che possiamo scrivere in forma matriciale come

$$\begin{pmatrix} 1 & 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 3 & 0 & 0 & 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}.$$

Osserviamo che questo sistema ha una proprietà “notevole”: in ogni sua equazione c’è almeno un’incognita (quale?) con coefficiente 1 che non figura in nessuna delle altre equazioni.

Questo ci permette di risolvere il sistema in maniera veloce; in ogni equazione si sceglie un’incognita che non figura nelle rimanenti equazioni del sistema, e la si esplicita in funzione delle altre incognite dell’equazione stessa.

Per esempio dalla prima equazione ricaviamo $x_1 = 1 - x_2 - 2x_3 - x_5$, dalla seconda $x_4 = x_2 - x_7$ e dalla terza $x_6 = -2 - 3x_2 - 3x_7$.

Le soluzioni del sistema sono quindi necessariamente tutte e sole le matrici colonna in $\mathbb{R}^{7,1}$ della forma

$$X = \begin{pmatrix} 1 - x_2 - 2x_3 - x_5 \\ x_2 \\ x_3 \\ x_2 - x_7 \\ x_5 \\ -2 - 3x_2 - 3x_7 \\ x_7 \end{pmatrix}$$

al variare di x_2, x_3, x_5, x_7 in \mathbb{R} .

Per esempio, se scegliamo i valori $x_2 = x_3 = x_5 = x_7 = 0$, si ottiene la soluzione ${}^t(1 \ 0 \ 0 \ 0 \ 0 \ -2 \ 0)$. Invece, scelti $x_2 = 1, x_3 = 0, x_5 = -2, x_7 = 1$, si ottiene la soluzione ${}^t(2 \ 1 \ 0 \ 0 \ -2 \ -8 \ 1)$, e così via (le soluzioni sono quindi infinite!). ♠

La proprietà descritta nell’Esempio 3.7, cioè il fatto che in ogni equazione del sistema c’è un’incognita che non figura in nessuna delle altre equazioni, può essere tradotta in linguaggio matriciale osservando che in ogni riga della matrice incompleta c’è almeno un’entrata uguale a 1 che è l’unica entrata non nulla nella sua colonna.

A tutti i sistemi la cui matrice incompleta ha questa proprietà si può applicare il metodo descritto nell’Esempio 3.7 ottenendo facilmente la soluzione generale.

Per questo motivo è utile dare un nome a tale tipo di matrici.

Definizione 3.8 (Matrici fortemente ridotte per righe).

Sia data $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathbb{K}^{m,n}$. La matrice A si dice *fortemente ridotta per righe* se vale la seguente proprietà: se la riga di indice i_0 contiene entrate non nulle, allora esiste una sua entrata a_{i_0, j_0} che vale 1 e tale che $a_{i, j_0} = 0$ per ogni $i \neq i_0$.

Se la matrice A è fortemente ridotta per righe, su ogni riga non nulla si può scegliere esattamente un'entrata a_{i_0, j_0} che vale 1 e tale che $a_{i, j_0} = 0$ per ogni $i \neq i_0$. Tale entrate vengono spesso chiamate *pivot* (della riga in cui esse si trovano).

Esempio 3.9. La matrice

$$A_1 = \begin{pmatrix} \boxed{1} & -2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 5 & -1 & \boxed{1} & 2 \\ 0 & 4 & \boxed{1} & 0 & 2 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

è fortemente ridotta per righe, e ne abbiamo indicato i pivot, mentre non lo è la matrice

$$A_2 = \begin{pmatrix} 1 & -2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 5 & -1 & 2 & 2 \\ 0 & 4 & 1 & 0 & 2 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad \spadesuit$$

Spesso è utile lavorare con matrici aventi delle proprietà simili a quelle descritte sopra, ma un po' più deboli.

Definizione 3.10 (Matrici ridotte per righe).

Sia data $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathbb{K}^{m,n}$. La matrice A si dice *ridotta per righe* se vale la seguente proprietà: se la riga di indice $i_0 < m$ contiene entrate non nulle, allora esiste una sua entrata $a_{i_0, j_0} \neq 0$ tale che $a_{i, j_0} = 0$ per ogni $i > i_0$.

Se la matrice A è ridotta per righe, su ogni riga non nulla di indice $i_0 < m$ si può scegliere esattamente un'entrata $a_{i_0, j_0} \neq 0$ e tale che $a_{i, j_0} = 0$ per ogni $i > i_0$. Similmente anche sulla riga di indice m , se non nulla, si può scegliere $a_{m, j_0} \neq 0$. Le entrate così scelte vengono spesso chiamate *pivot* (della riga in cui esse si trovano).

Chiaramente ogni matrice fortemente ridotta per righe è anche ridotta per righe, ma non vale il viceversa, come mostra il seguente esempio.

Esempio 3.11. Consideriamo le matrici

$$A_1 = \begin{pmatrix} \boxed{1} & -2 & 3 & 0 & 0 & 3 & 0 \\ 0 & 2 & 0 & \boxed{5} & -1 & 2 & 2 \\ 0 & 4 & \boxed{1} & 0 & 2 & 0 & 4 \\ 0 & \boxed{5} & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & -2 & 3 & 0 & 0 & 3 & 0 \\ 0 & 2 & 0 & 5 & -1 & 2 & 2 \\ 0 & 4 & 1 & 0 & 2 & 0 & 4 \\ 4 & 5 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

La matrice A_1 è ridotta per righe, ma non fortemente ridotta per righe. Invece la matrice A_2 non è ridotta per righe, dunque non lo è neanche fortemente. \spadesuit

Per concludere questo capitolo, osserviamo che esistono delle analoghe nozioni di *matrici (fortemente) ridotte per colonne*, ma per non complicarci troppo la vita noi ci limiteremo, per il momento, a considerare solo le riduzioni per righe.

Operazioni elementari di riga e rango di una matrice

4.1 Operazioni elementari di riga

Nel capitolo precedente abbiamo visto un metodo per risolvere un sistema lineare la cui matrice dei coefficienti sia fortemente ridotta per righe, o anche solo ridotta per righe. Avere a che fare con un tale sistema è però abbastanza raro. Un'idea può allora essere quella di trasformare il sistema che vogliamo risolvere in un nuovo sistema che abbia *le stesse soluzioni* e che abbia una matrice fortemente ridotta per righe, risolvendo poi quest'ultimo invece di quello di partenza.

Definizione 4.1 (Sistemi equivalenti).

Due sistemi di equazioni (non necessariamente lineari) si dicono *equivalenti* se hanno lo stesso insieme di soluzioni.

Esempio 4.2. Consideriamo il sistema

$$\begin{cases} x + y + z = 1 \\ x - y + 2z = -3 \end{cases} \quad (4.1.1)$$

avente matrice completa

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -1 & 2 & -3 \end{array} \right),$$

che chiaramente non è ridotta per righe.

Però possiamo sostituire al posto della seconda equazione in (4.1.1) la somma delle due equazioni, ottenendo il nuovo sistema

$$\begin{cases} x + y + z = 1 \\ 2x + 3z = -2 \end{cases} \quad (4.1.2)$$

avente matrice completa

$$\left(\begin{array}{ccc|c} 1 & \boxed{1} & 1 & 1 \\ \boxed{2} & 0 & 3 & -2 \end{array} \right),$$

che, invece, è ridotta per righe.

Questa sostituzione lascia invariate le soluzioni del sistema di partenza. Infatti supponiamo che ${}^t(x_0 \ y_0 \ z_0)$ sia una soluzione del sistema (4.1.1) e, quindi, che valgano le uguaglianze

$$\begin{cases} x_0 + y_0 + z_0 - 1 = 0 \\ x_0 - y_0 + 2z_0 + 3 = 0. \end{cases}$$

In tal caso

$$\begin{cases} x_0 + y_0 + z_0 - 1 = 0 \\ 2x_0 + 3z_0 + 2 = (x_0 + y_0 + z_0 - 1) + (x_0 - y_0 + 2z_0 + 3) = 0, \end{cases}$$

è, quindi, ${}^t(x_0 \ y_0 \ z_0)$ è anche soluzione del sistema (4.1.2).

Viceversa, il sistema (4.1.1) si può ottenere dal sistema (4.1.2) sostituendo alla sua seconda equazione la differenza della seconda equazione meno la prima, quindi con un ragionamento completamente analogo a quello fatto sopra deduciamo che le soluzioni di (4.1.2) sono anche soluzioni di (4.1.1). Concludiamo che l'insieme delle soluzioni di (4.1.1) coincide con l'insieme delle soluzioni di (4.1.2).

Ora che ci siamo assicurati di poter risolvere il sistema (4.1.2) invece del più difficile sistema (4.1.1), procediamo a trovarne le soluzioni. Dividiamo la seconda equazione per 2 ottenendo il sistema

$$\begin{cases} x + y + z = 1 \\ x + 3z/2 = -1 \end{cases} \quad (4.1.3)$$

la cui matrice completa è

$$\left(\begin{array}{ccc|c} 1 & \boxed{1} & 1 & 1 \\ \boxed{1} & 0 & 3/2 & -1 \end{array} \right).$$

È facile osservare che il nuovo sistema (4.1.3) ha ancora le stesse soluzioni del sistema di partenza (4.1.1). Come ultimo passaggio sottraiamo dalla prima equazione di (4.1.3) la seconda, ottenendo il sistema

$$\begin{cases} y - z/2 = 2 \\ x + 3z/2 = -1 \end{cases} \quad (4.1.4)$$

avente matrice completa associata

$$\left(\begin{array}{ccc|c} 0 & \boxed{1} & -1/2 & 2 \\ \boxed{1} & 0 & 3/2 & -1 \end{array} \right),$$

che è fortemente ridotta per righe e che, di nuovo, ha le stesse soluzioni di tutti gli altri sistemi precedenti. Risolvendo il sistema (4.1.4) come spiegato nell'Esempio 3.7, otteniamo che l'insieme di tutte le soluzioni dei sistemi di questo esempio è

$$\left\{ \left(\begin{array}{c} -1 - 3z/2 \\ 2 + z/2 \\ z \end{array} \right) \mid z \in \mathbb{R} \right\}. \quad \spadesuit$$

$\triangle!$ Due sistemi equivalenti non hanno necessariamente lo stesso numero di equazioni. Per esempio, si verifichi che i due seguenti sistemi sono equivalenti:

$$\begin{cases} x + y + z = 0 \\ z = 1 \end{cases} \quad \begin{cases} x + y + z = 0 \\ x + y = -1 \\ z = 1. \end{cases}$$

Siamo interessati a capire se possiamo ripetere in generale quanto fatto nell'Esempio 4.2, ovvero se esistono delle operazioni che possiamo fare sulle equazioni di un sistema che lo trasformino in un sistema equivalente con matrice associata fortemente ridotta per righe.

Per enunciare il risultato generale introduciamo la seguente definizione.

Definizione 4.3 (Operazioni elementari di riga e matrici equivalenti per righe).

Sia data $A \in \mathbb{K}^{m,n}$. Le *operazioni elementari di riga* su A sono:

- (E1) sommare a una riga di A un multiplo di un'altra riga di A ;
- (E2) moltiplicare una riga di A per una costante non nulla $\alpha \in \mathbb{K}$;
- (E3) scambiare due righe di A .

Due matrici $A, \hat{A} \in \mathbb{K}^{m,n}$ tali che esiste una successione finita di operazioni elementari di riga che trasforma A in \hat{A} si dicono *equivalenti per righe*.

Per comodità introduciamo anche la seguente notazione per le tre operazioni:

- (E1) $R_i \rightarrow R_i + \alpha R_{i_0}$ denota l'operazione di sommare alla riga di indice i la riga di indice $i_0 \neq i$ moltiplicata per $\alpha \in \mathbb{K}$;
- (E2) $R_i \rightarrow \alpha R_i$ denota l'operazione di moltiplicare la riga di indice i per $\alpha \in \mathbb{K}$, $\alpha \neq 0$;
- (E3) $R_i \leftrightarrow R_{i_0}$ denota l'operazione di scambiare le righe indici i e i_0 .

Il risultato fondamentale è il seguente.

Proposizione 4.4.

Sia data $A \in \mathbb{K}^{m,n}$. Allora A è equivalente per righe a una matrice $\hat{A} \in \mathbb{K}^{m,n}$ (fortemente) ridotta per righe.

Dimostrazione. Sia $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$, e supponiamo che $A \neq 0_{m,n}$ (altrimenti non c'è nulla da dimostrare).

Sia i_0 il più piccolo indice per cui esiste $a_{i_0, j_0} \neq 0$. Moltiplicando la riga di indice i_0 per a_{i_0, j_0}^{-1} :

$$A \xrightarrow{R_{i_0} \rightarrow a_{i_0, j_0}^{-1} R_{i_0}} A'$$

trasformiamo la matrice A in una nuova matrice A' avente l'entrata 1 in posizione (i_0, j_0) .

Per ogni $i \neq i_0$ sostituiamo la riga di indice i con la sua somma alla riga di indice i_0 moltiplicata per a_{i,j_0}

$$A' \xrightarrow{R_i \rightarrow R_i - a_{i,j_0} R_{i_0}} A''$$

trasformiamo la matrice A' in una nuova matrice $A'' = (a''_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ la cui colonna di indice j_0 contiene un'unica entrata non nulla che vale 1 in posizione (i_0, j_0) .

A questo punto si presentano due possibilità. Nel primo caso tutte le righe di indice $i \neq i_0$ sono nulle, quindi scambiando la riga di indice i_0 con la riga di indice 1

$$A'' \xrightarrow{R_1 \leftrightarrow R_{i_0}} A'''$$

trasformiamo A'' in una nuova matrice $A''' = (a'''_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ la cui colonna di indice j_0 contiene un'unica entrata non nulla che vale 1 in posizione $(1, j_0)$ e tale che $a'''_{i,j} = 0$ per ogni $i > 1$. Quindi $\widehat{A} = A'''$ è fortemente ridotta per righe.

Nel secondo caso ripetiamo lo stesso procedimento con il più piccolo indice $i_1 > i_0$ per cui esiste $a''_{i_1, j_1} \neq 0$. Poiché $a''_{i, j_0} = 0$ per $i \neq i_0$ segue che $j_1 \neq j_0$.

In questo modo dopo al più m passi (uno per ogni riga) otteniamo una matrice \widehat{A} fortemente ridotta per righe. \square

Esempio 4.5. Trasformiamo la seguente matrice in matrice fortemente ridotta per righe, usando operazioni elementari.

$$A = \begin{pmatrix} 3 & 6 & 0 & 3 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 2 & 1 \end{pmatrix} \xrightarrow{R_1 \rightarrow (1/3)R_1} \begin{pmatrix} 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 2 & 1 \end{pmatrix}$$

$$\xrightarrow{\substack{R_2 \rightarrow R_2 - R_1 \\ R_3 \rightarrow R_3 - R_1}} \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & -2 & 2 & 0 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 - 2R_2} \begin{pmatrix} \boxed{1} & 2 & 0 & 1 \\ 0 & -1 & \boxed{1} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

che è fortemente ridotta per righe. \spadesuit

Esempio 4.6. Come nell'Esempio 4.5, trasformiamo la seguente matrice in matrice fortemente ridotta per righe, usando operazioni elementari.

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 3 & -2 & 1 & 2 \\ 2 & 4 & -1 & 2 & 3 \\ 1 & -1 & 2 & 5 & 7 \end{pmatrix} \xrightarrow{R_2 \rightarrow R_2 - R_1} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & -3 & 0 & 1 \\ 2 & 4 & -1 & 2 & 3 \\ 1 & -1 & 2 & 5 & 7 \end{pmatrix}$$

$$\xrightarrow{\substack{R_3 \rightarrow R_3 - 2R_1 \\ R_4 \rightarrow R_4 - R_1}} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & -3 & 0 & 1 \\ 0 & 2 & -3 & 0 & 1 \\ 0 & -2 & 1 & 4 & 6 \end{pmatrix} \xrightarrow{\substack{R_3 \rightarrow R_3 - R_2 \\ R_4 \rightarrow R_4 + R_2}} \begin{pmatrix} \boxed{1} & 1 & 1 & 1 & 1 \\ 0 & \boxed{2} & -3 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{-2} & 4 & 7 \end{pmatrix} = A',$$

dove A' è ridotta per righe. Possiamo proseguire:

$$\begin{aligned}
 A' &\xrightarrow[\substack{R_4 \rightarrow -R_4/2}]{R_2 \rightarrow R_2/2} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & -3/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -2 & -7/2 \end{pmatrix} \xrightarrow[\substack{R_1 \rightarrow R_1 - R_4}]{R_2 \rightarrow R_2 + 3R_4/2} \begin{pmatrix} 1 & 1 & 0 & 3 & 9/2 \\ 0 & 1 & 0 & -3 & -19/4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -2 & -7/2 \end{pmatrix} \\
 &\xrightarrow{R_1 \rightarrow R_1 - R_2} \begin{pmatrix} 1 & 0 & 0 & 6 & 37/4 \\ 0 & 1 & 0 & -3 & -19/4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -2 & -7/2 \end{pmatrix} \xrightarrow{R_3 \leftrightarrow R_4} \begin{pmatrix} \boxed{1} & 0 & 0 & 6 & 37/4 \\ 0 & \boxed{1} & 0 & -3 & -19/4 \\ 0 & 0 & \boxed{1} & -2 & -7/2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \widehat{A},
 \end{aligned}$$

ottenendo infine una matrice fortemente ridotta per righe. ♠

4.2 Interpretazione matriciale

Le operazioni elementari di riga sono legate in maniera interessante e utile al prodotto di matrici. In particolare sfrutteremo tale legame più avanti trattando le nozioni di determinante di una matrice e di segno di una forma quadratica.

Introduciamo la seguente definizione.

Definizione 4.7 (Matrici elementari).

Una matrice $E \in \mathbb{K}^{m,m}$ si dice elementare se è ottenuta dalla matrice $I_m \in \mathbb{K}^{m,m}$ mediante un'operazione elementare di riga.

In queste dispense utilizzeremo talvolta, con ovvio significato dei simboli, la seguente notazione per le matrici elementari sopra introdotte:

$$E_{R_i \rightarrow R_i + \alpha R_{i_0}}, \quad E_{R_i \rightarrow \alpha R_i}, \quad E_{R_i \leftrightarrow R_{i_0}}.$$

Si noti che $E_{R_i \rightarrow \alpha R_i}$ è diagonale. La matrice $E_{R_i \rightarrow R_i + \alpha R_{i_0}}$ è triangolare superiore se $i < i_0$, inferiore se $i > i_0$. Invece $E_{R_i \leftrightarrow R_{i_0}}$ non è triangolare né superiore, né inferiore se $i \neq i_0$.

Il legame tra operazioni elementari e prodotto di matrici è dato dalla seguente proposizione.

Proposizione 4.8.

Sia data $A \in \mathbb{K}^{m,n}$. Sia $A' \in \mathbb{K}^{m,n}$ ottenuta da A con un'operazione elementare di riga e sia $E \in \mathbb{K}^{m,m}$ la matrice elementare ottenuta con la stessa operazione da I_m . Allora

$$A' = EA.$$

Dimostrazione. La dimostrazione consiste in una verifica diretta nei vari casi.

Visto che un'operazione elementare coinvolge al massimo due righe per volta è sufficiente verificare l'enunciato nel caso $m = 2$.

Nel seguito verificheremo l'enunciato nel caso A' sia ottenuta da A con un'operazione elementare di tipo (E1), lasciando al lettore le verifiche per le operazioni elementari di tipo (E2) ed (E3). In tale caso $E \in \mathbb{K}^{2,2}$ e sia

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \end{pmatrix}.$$

Allora se consideriamo $R_2 \rightarrow R_2 + \alpha R_1$ risulta

$$E = E_{R_2 \rightarrow R_2 + \alpha R_1} = \begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix}, \quad A' = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} + \alpha a_{1,1} & a_{2,2} + \alpha a_{1,2} & \cdots & a_{2,n} + \alpha a_{1,n} \end{pmatrix}.$$

Si verifichi che $A' = EA$ in questo caso. □

Esempio 4.9. Riprendiamo la matrice dell'Esempio 4.5. In tal caso $m = 3$ e risulta

$$\begin{aligned} E_{R_1 \rightarrow (1/3)R_1} &= \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, & E_{R_2 \rightarrow R_2 - R_1} &= \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\ E_{R_3 \rightarrow R_3 - R_1} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}, & E_{R_3 \rightarrow R_3 - 2R_2} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}. \end{aligned}$$

Si ha

$$\begin{aligned} A_1 &= E_{R_1 \rightarrow (1/3)R_1} A = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 6 & 0 & 3 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 2 & 1 \end{pmatrix}, \\ A_2 &= E_{R_2 \rightarrow R_2 - R_1} A' = E_{R_2 \rightarrow R_2 - R_1} E_{R_1 \rightarrow (1/3)R_1} A \\ &= \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 1 & 0 & 2 & 1 \end{pmatrix}, \\ A_3 &= E_{R_3 \rightarrow R_3 - R_1} A'' = E_{R_3 \rightarrow R_3 - R_1} E_{R_2 \rightarrow R_2 - R_1} E_{R_1 \rightarrow (1/3)R_1} A \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 1 & 0 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & -2 & 2 & 0 \end{pmatrix}, \\ A_4 &= E_{R_3 \rightarrow R_3 - 2R_2} A''' = E_{R_3 \rightarrow R_3 - 2R_2} E_{R_3 \rightarrow R_3 - R_1} E_{R_2 \rightarrow R_2 - R_1} E_{R_1 \rightarrow (1/3)R_1} A \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & -2 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Si noti che A_4 è esattamente la matrice A' nell'Esempio 4.5. ♠

4.3 Rango di una matrice

È chiaro che le operazioni di riga che si possono fare per trasformare una matrice in una (fortemente) ridotta non sono univocamente determinate, anzi a ogni passo è possibile fare scelte diverse! Di conseguenza da ogni matrice, con operazioni elementari di riga, si potranno ottenere varie matrici fortemente ridotte per righe, anche molto diverse: infatti, a ogni passo, bisogna fare una scelta del pivot.

Però si può dimostrare che, indipendentemente dalle operazioni fatte, il numero di righe contenenti entrate non nulle non cambia.

Proposizione 4.10.

Sia data $A \in \mathbb{K}^{m,n}$. Se A' e A'' sono due matrici ridotte per righe equivalenti per righe ad A . Allora i numeri di righe di A' e di A'' contenenti entrate non nulle coincidono.

Dimostrazione. Omettiamo la dimostrazione. Vedremo in seguito che tale numero dipende solo da A (coincide con un numero chiamato “dimensione dello spazio riga di A ”) e non dalla riduzione operata. \square

Si può facilmente immaginare che tale numero riveste un’importanza particolare nell’algebra delle matrici, pertanto merita un nome particolare.

Definizione 4.11 (Rango di una matrice).

Sia data $A \in \mathbb{K}^{m,n}$. Se A' è una matrice ridotta per righe equivalente per righe ad A , il numero di righe di A' contenenti entrate non nulle viene detto *rango* di A e indicato con il simbolo $\text{rk}(A)$.

Proposizione 4.12.

Sia data $A \in K^{m,n}$. Allora $\text{rk}(A) \leq \min\{m, n\}$.

Dimostrazione. Per definizione $\text{rk}(A) \leq m$. Inoltre $\text{rk}(A)$ coincide con il numero di pivot di una forma ridotta per righe di A : ognuno di essi si trova necessariamente in una colonna diversa, quindi si ha anche $\text{rk}(A) \leq n$. \square

Esempio 4.13. Le matrici degli Esempi 4.5 e 4.6 hanno rispettivamente $\text{rk}(A) = 2$ e $\text{rk}(A) = 3$. \spadesuit

Esempio 4.14. Studiamo il rango della matrice

$$C = \begin{pmatrix} 1 & k & -1 & 2 \\ 2 & 2 & -2 & 2k \\ 1 & 1 & -k & 2k \end{pmatrix} \in \mathbb{R}^{3,4}$$

al variare del parametro $k \in \mathbb{R}$. Riduciamo C per righe

$$C \xrightarrow{R_2 \rightarrow (1/2)R_2} \begin{pmatrix} 1 & k & -1 & 2 \\ 1 & 1 & -1 & k \\ 1 & 1 & -k & 2k \end{pmatrix} \xrightarrow{\substack{R_2 \rightarrow R_2 - R_1 \\ R_3 \rightarrow R_3 - R_1}} \begin{pmatrix} 1 & k & -1 & 2 \\ 0 & 1 - k & 0 & k - 2 \\ 0 & 1 - k & 1 - k & 2k - 2 \end{pmatrix}$$

$$\xrightarrow{R_3 \rightarrow R_3 - R_2} \begin{pmatrix} 1 & k & -1 & 2 \\ 0 & 1 - k & 0 & k - 2 \\ 0 & 0 & 1 - k & k \end{pmatrix} = C'.$$

A questo punto abbiamo due possibilità: o $k = 1$, o $k \neq 1$. Se $k = 1$, la matrice ottenuta è

$$C' = \begin{pmatrix} 1 & 1 & -1 & 2 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 + R_2} \begin{pmatrix} \boxed{1} & 1 & -1 & 2 \\ 0 & 0 & 0 & \boxed{-1} \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

che ha rango 2. Altrimenti se $k \neq 1$, possiamo dividere per $k - 1$ che è diverso da zero, ottenendo

$$C' \xrightarrow{\substack{R_2 \rightarrow (1/(k-1))R_2 \\ R_3 \rightarrow (1/(k-1))R_3}} \begin{pmatrix} \boxed{1} & k & -1 & 2 \\ 0 & \boxed{1} & 0 & \frac{k-2}{k-1} \\ 0 & 0 & \boxed{1} & \frac{k}{k-1} \end{pmatrix}.$$

Quindi il rango è uguale a 3 se $k \neq 1$. ♠

⚠ Testi diversi utilizzano notazioni diverse per denotare il rango di una matrice, non solo $\text{rk}(A)$, ma anche $\text{rg}(A)$, $r(A)$, $\rho(A)$, $N(A)$.

Risoluzione di sistemi lineari ed equazioni matriciali

5.1 Risoluzione di sistemi

Sia $AX = B$ un sistema di equazioni lineari, con la sua matrice completa associata $(A|B)$. Per la Proposizione 4.4 sappiamo di poter trasformare con operazioni elementari di riga la matrice A in una nuova matrice A' fortemente ridotta per righe: con le stesse operazioni elementari si ottiene una nuova matrice $(A'|B')$ corrispondente a un nuovo sistema di equazioni lineari $A'X = B'$, diverso dal precedente ma a esso equivalente. Risolvendo, qualora compatibile, tale sistema con il metodo descritto nell'Esempio 4.2 otteniamo l'insieme delle soluzioni del sistema di partenza.

Lemma 5.1.

Siano date $A, A' \in \mathbb{K}^{m,n}$, $B, B' \in \mathbb{K}^{m,1}$. Sia $(A|B)$ la matrice completa associata al sistema lineare di m equazioni in n incognite $AX = B$ e sia $(A'|B')$ matrice equivalente per righe a $(A|B)$. Allora i sistemi $AX = B$ e $A'X = B'$ sono equivalenti.

Dimostrazione. L'enunciato segue dal fatto che ogni operazione elementare non cambia l'insieme delle soluzioni di un sistema. Questo è ovvio per le operazioni (E2) ed (E3). Nel caso dell'operazione (E1) è sufficiente imitare quanto fatto nell'Esempio 4.2. \square

Siamo finalmente pronti a enunciare e dimostrare il principale risultato sulla teoria dei sistemi di equazioni lineari, detto Teorema di Rouché–Capelli.

Proposizione 5.2 (Teorema di Rouché–Capelli).

Siano date $A \in \mathbb{K}^{m,n}$ e $B \in \mathbb{K}^{m,1}$. Si consideri il sistema lineare di m equazioni in n incognite $AX = B$, con matrice completa associata $(A|B)$. Allora

- (i) il sistema è compatibile se e solo se $\text{rk}(A) = \text{rk}(A|B)$;
- (ii) se il sistema è compatibile, le sue soluzioni dipendono da $n - \text{rk}(A)$ parametri liberi;
- (iii) se il sistema è compatibile e X_0 è una sua soluzione fissata, allora ogni altra sua soluzione X è della forma $X = X_0 + Y$, dove Y appartiene all'insieme delle soluzioni del sistema omogeneo associato $AX = 0_{m,1}$.

Dimostrazione.

- (i) Per quanto visto sopra possiamo assumere senza perdere in generalità che A sia una matrice fortemente ridotta per righe. Si possono presentare due situazioni per le righe della matrice completa $(A|B)$.

Il primo caso è che esista una riga di A , diciamo quella di indice i , con entrate tutte nulle che si prolunga in $(A|B)$ a una riga con entrate non tutte nulle: chiaramente l'entrata non nulla deve essere l' i -esima entrata di B , cioè $b_i \neq 0$. In particolare quindi $\text{rk}(A) \neq \text{rk}(A|B)$. Ciò significa che nel nostro sistema figura un'equazione della forma $0 = b_i$ che, per l'ipotesi $b_i \neq 0$, non ha soluzioni e cioè il sistema è incompatibile.

Nel secondo caso ogni riga di A con entrate tutte nulle si prolunga in $(A|B)$ a una riga con entrate tutte nulle (è anche possibile che queste righe non esistano). In questo caso possiamo risolvere il sistema come spiegato nell'Esempio 3.7, e il sistema è compatibile: inoltre i numeri di righe di A e di $(A|B)$ contenenti entrate non nulle coincidono, cioè $\text{rk}(A) = \text{rk}(A|B)$.

- (ii) Quanto visto nella parte (i) dimostra anche l'affermazione (ii): infatti possiamo esprimere le incognite i cui coefficienti sono i pivot (e che sono in totale $\text{rk}(A)$) in funzione delle rimanenti incognite (in totale $n - \text{rk}(A)$), cui possiamo dare valori arbitrari.
- (iii) Sia $X_0 \in \mathbb{R}^{n,1}$ una soluzione fissata del sistema, cioè $AX_0 = B$, e sia $X \in \mathbb{R}^{n,1}$ un'altra soluzione del sistema, quindi tale che $AX = B$. Sia $Y = X - X_0$, allora

$$AY = A(X - X_0) = AX + A(-X_0) = AX + A(-1)X_0 = AX - AX_0 = B - B = 0_{m,1},$$

quindi Y è soluzione del sistema omogeneo associato.

Viceversa sia $Y \in \mathbb{R}^{n,1}$ una soluzione del sistema omogeneo associato, cioè $AY = 0_{m,1}$. Sia $X = Y + X_0$, allora

$$AX = A(Y + X_0) = AY + AX_0 = 0_{m,1} + B = B,$$

e quindi X è soluzione del sistema di partenza, come volevamo dimostrare. \square

Esempio 5.3. Si consideri il sistema

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 3 & -2 & 1 & 2 \\ 2 & 4 & -1 & 2 & 3 \\ 1 & -1 & 2 & 5 & 7 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Le matrici incompleta e completa del sistema sono rispettivamente

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 3 & -2 & 1 & 2 \\ 2 & 4 & -1 & 2 & 3 \\ 1 & -1 & 2 & 5 & 7 \end{pmatrix}, \quad (A|B_1) = \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 3 & -2 & 1 & 2 & 0 \\ 2 & 4 & -1 & 2 & 3 & 1 \\ 1 & -1 & 2 & 5 & 7 & 0 \end{array} \right).$$

Osserviamo che nell'Esempio 4.6 abbiamo già ridotto A alla matrice ridotta per righe

$$A' = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & -3 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 4 & 7 \end{pmatrix}.$$

Utilizziamo le stesse operazioni di riga alla matrice completa $(A|B_1)$: rifacciamo cioè gli stessi passaggi ma teniamo conto anche dell'effetto delle operazioni su B_1 .

$$\begin{aligned} (A|B_1) &\xrightarrow{R_2 \rightarrow R_2 - R_1} \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 2 & -3 & 0 & 1 & 0 \\ 2 & 4 & -1 & 2 & 3 & 1 \\ 1 & -1 & 2 & 5 & 7 & 0 \end{array} \right) \xrightarrow{\substack{R_3 \rightarrow R_3 - 2R_1 \\ R_4 \rightarrow R_4 - R_1}} \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 2 & -3 & 0 & 1 & 0 \\ 0 & 2 & -3 & 0 & 1 & 1 \\ 0 & -2 & 1 & 4 & 6 & 0 \end{array} \right) \\ &\xrightarrow{\substack{R_3 \rightarrow R_3 - R_2 \\ R_4 \rightarrow R_4 + R_2}} \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 2 & -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -2 & 4 & 7 & 0 \end{array} \right) = (A'|B'_1). \end{aligned}$$

Si noti che $\text{rk}(A) = \text{rk}(A') = 3$ mentre $\text{rk}(A|B_1) = \text{rk}(A'|B'_1) = 4$: da questo deduciamo che il sistema $AX = B_1$ è incompatibile. ♠

Esempio 5.4. Si consideri ora il sistema

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 3 & -2 & 1 & 2 \\ 2 & 4 & -1 & 2 & 3 \\ 1 & -1 & 2 & 5 & 7 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Le matrici incompleta e completa del sistema sono rispettivamente

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 3 & -2 & 1 & 2 \\ 2 & 4 & -1 & 2 & 3 \\ 1 & -1 & 2 & 5 & 7 \end{pmatrix}, \quad (A|B_2) = \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 3 & -2 & 1 & 2 & 0 \\ 2 & 4 & -1 & 2 & 3 & 0 \\ 1 & -1 & 2 & 5 & 7 & 1 \end{array} \right),$$

quindi la matrice incompleta è di nuovo quella dell'Esempio 4.6. Utilizzando le stesse operazioni elementari e tenendo di nuovo conto del loro effetto su B_2 , troviamo

$$\begin{aligned} (A|B_2) &\xrightarrow{R_2 \rightarrow R_2 - R_1} \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 2 & -3 & 0 & 1 & 0 \\ 2 & 4 & -1 & 2 & 3 & 0 \\ 1 & -1 & 2 & 5 & 7 & 1 \end{array} \right) \xrightarrow{\substack{R_3 \rightarrow R_3 - 2R_1 \\ R_4 \rightarrow R_4 - R_1}} \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 2 & -3 & 0 & 1 & 0 \\ 0 & 2 & -3 & 0 & 1 & 0 \\ 0 & -2 & 1 & 4 & 6 & 1 \end{array} \right) \\ &\xrightarrow{\substack{R_3 \rightarrow R_3 - R_2 \\ R_4 \rightarrow R_4 + R_2}} \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 2 & -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 4 & 7 & 1 \end{array} \right) = (A'|B'_2). \end{aligned}$$

Poiché $\text{rk}(A) = \text{rk}(A') = 3 = \text{rk}(A'|B_2) = \text{rk}(A|B_2)$ segue che il sistema in esame è compatibile. Perciò proseguiamo con le operazioni elementari di riga riducendo ulteriormente la matrice completa del sistema, in modo da trovarne le soluzioni.

Di nuovo, nell'Esempio 4.6 avevamo già ridotto A a una forma fortemente ridotta per righe, ora applichiamo le stesse operazioni di riga alla matrice completa $(A|B_2)$:

$$\begin{aligned}
 (A'|B_2) &\xrightarrow{\substack{R_2 \rightarrow R_2/2 \\ R_4 \rightarrow -R_4/2}} \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & -\frac{3}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -2 & -\frac{7}{2} & -\frac{1}{2} \end{array} \right) \\
 &\xrightarrow{\substack{R_2 \rightarrow R_2 + 3R_4/2 \\ R_1 \rightarrow R_1 - R_4}} \left(\begin{array}{ccccc|c} 1 & 1 & 0 & 3 & \frac{9}{2} & \frac{1}{2} \\ 0 & 1 & 0 & -3 & -\frac{19}{4} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -2 & -\frac{7}{2} & -\frac{1}{2} \end{array} \right) \\
 &\xrightarrow{R_1 \rightarrow R_1 - R_2} \left(\begin{array}{ccccc|c} 1 & 0 & 0 & 6 & \frac{37}{4} & \frac{1}{2} \\ 0 & 1 & 0 & -3 & -\frac{19}{4} & -\frac{3}{4} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -2 & -\frac{7}{2} & -\frac{1}{2} \end{array} \right) \\
 &\xrightarrow{R_3 \leftrightarrow R_4} \left(\begin{array}{ccccc|c} 1 & 0 & 0 & 6 & \frac{37}{4} & \frac{5}{4} \\ 0 & 1 & 0 & -3 & -\frac{19}{4} & -\frac{3}{4} \\ 0 & 0 & 1 & -2 & -\frac{7}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) = (\widehat{A}|\widehat{B}_2).
 \end{aligned}$$

Quindi il sistema $AX = B_2$ è equivalente al sistema $\widehat{A}X = \widehat{B}_2$, che è compatibile perché $\text{rk}(A) = \text{rk}(\widehat{A}) = 3 = \text{rk}(\widehat{A}|\widehat{B}_2) = \text{rk}(A|B_2)$.

In particolare le incognite corrispondenti ai pivot sono a , b , c e si ha

$$\begin{cases} a = \frac{5}{4} - 6d - \frac{37}{4}e \\ b = -\frac{3}{4} + 3d + \frac{19}{4}e \\ c = -\frac{1}{2} + 2d + \frac{7}{2}e \end{cases}$$

L'insieme delle soluzioni del sistema di partenza è quindi

$$\left\{ \left(\begin{array}{c} \frac{5}{4} - 6d - \frac{37}{4}e \\ -\frac{3}{4} + 3d + \frac{19}{4}e \\ -\frac{1}{2} + 2d + \frac{7}{2}e \\ d \\ e \end{array} \right) \mid d, e \in \mathbb{R} \right\}.$$

Per esempio, se scegliamo i valori $d = e = 0$ otteniamo la soluzione particolare $X_0 = {}^t(5/4 \ -3/4 \ -1/2 \ 0 \ 0)$. Si noti che ogni altra soluzione è della forma

$$\begin{pmatrix} 5/4 \\ -3/4 \\ -1/2 \\ 0 \\ 0 \end{pmatrix} + d \begin{pmatrix} -6 \\ 3 \\ 2 \\ 1 \\ 0 \end{pmatrix} + e \begin{pmatrix} -37/4 \\ 19/4 \\ 7/2 \\ 0 \\ 1 \end{pmatrix}$$

al variare di d e e in \mathbb{R} . Si verifichi che le soluzioni del sistema $AX = 0_{m,1}$ sono tutte e sole le matrici della forma

$$d \begin{pmatrix} -6 \\ 3 \\ 2 \\ 1 \\ 0 \end{pmatrix} + e \begin{pmatrix} -37/4 \\ 19/4 \\ 7/2 \\ 0 \\ 1 \end{pmatrix}$$

al variare di d e e in \mathbb{R} , come previsto dal Teorema di Rouché-Capelli. ♠

5.2 Equazioni matriciali

Negli Esempi 5.3 e 5.4 abbiamo studiato due sistemi $AX_1 = B_1$ e $AX_2 = B_2$ aventi la stessa matrice incompleta A ; questo tipo di problema si presenta in varie situazioni (vedremo in seguito il problema del calcolo della matrice inversa). È evidente che è inutile ripetere le stesse operazioni per ciascun sistema, e che sarebbe più conveniente risolvere i sistemi simultaneamente.

La risoluzione “simultanea” dei due sistemi indicati sopra equivale a considerare le soluzioni dell’equazione matriciale $AX = B$ ove X e B sono rispettivamente una matrice incognita e una numerica aventi colonna di indice j pari a X_j e B_j rispettivamente.

Vediamo più in dettaglio di cosa si tratta nella seguente definizione.

Definizione 5.5 (Equazioni matriciali e loro soluzioni).

Siano date $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathbb{K}^{m,n}$ e $B = (b_{i,h})_{\substack{1 \leq i \leq m \\ 1 \leq h \leq p}} \in \mathbb{K}^{m,p}$ matrici. Un’equazione matriciale lineare con matrice incompleta A e matrice dei termini noti B è un’equazione della forma

$$AX = B, \tag{5.2.1}$$

dove X è una matrice incognita $n \times p$.

La matrice

$$(A|B) = \left(\begin{array}{cccc|ccc} a_{1,1} & a_{1,2} & \dots & a_{1,n} & b_{1,1} & \dots & b_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} & b_{2,1} & \dots & b_{2,p} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} & b_{m,1} & \dots & b_{m,p} \end{array} \right)$$

viene detta *matrice completa* dell’equazione (5.2.1).

L’equazione (5.2.1) si dice *omogenea* se $B = 0_{m,p}$, *non omogenea* $B \neq 0_{m,p}$. Una *soluzione dell’equazione* (5.2.1) è una matrice numerica $\bar{X} \in \mathbb{K}^{n,p}$ per cui vale l’identità numerica $A\bar{X} = B$; se esiste una soluzione, l’equazione (5.2.1) si dice *compatibile*, altrimenti *incompatibile*.

L'equazione (5.2.1) può essere pensata come sistema di mp equazioni, una per ogni entrata di B , in np incognite, una per ogni entrata di X .

Si noti però che la riga di indice i di A definisce esattamente p equazioni di tale grande sistema, una per ogni entrata della riga di indice i della matrice B .

Fissato un tale i , l'entrata $a_{i,j}$ moltiplica nelle equazioni considerate tutte le entrate $x_{h,j}$ di X per $j = 1, \dots, p$. Indicata con X_h la riga di indice h di X , possiamo allora pensare all'equazione (5.2.1) come un sistema di m equazioni corrispondenti alle m righe di $(A|B)$ nelle n incognite matriciali della forma

$$\begin{cases} a_{1,1}X_1 + a_{1,2}X_2 + \dots + a_{1,n}X_n = (b_{1,1} & \dots & b_{1,p}) \\ a_{2,1}X_1 + a_{2,2}X_2 + \dots + a_{2,n}X_n = (b_{2,1} & \dots & b_{2,p}) \\ \vdots \\ a_{m,1}X_1 + a_{m,2}X_2 + \dots + a_{m,n}X_n = (b_{m,1} & \dots & b_{m,p}). \end{cases}$$

Ne segue che il metodo di soluzione delle equazioni matriciali è totalmente analogo a quello dei sistemi di equazioni lineari (che ne sono un caso particolare quando la matrice dei termini noti si riduce a un'unica colonna). Infatti esso si basa sulla riduzione della matrice completa $(A|B)$ con operazioni elementari di riga che continuano ad avere senso anche per incognite di tipo matriciale.

Esempio 5.6. Si consideri l'equazione matriciale

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} X = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad (5.2.2)$$

la cui matrice completa è

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 2 \\ 2 & 1 & 0 & 1 \end{array} \right),$$

corrispondente al sistema

$$\begin{cases} X_1 + 2X_2 = (1 & 2) \\ 2X_1 + X_2 = (0 & 1). \end{cases}$$

Trasformando $(A|B)$ con operazioni elementari di riga otteniamo

$$\begin{aligned} (A|B) &\xrightarrow{R_2 \rightarrow R_2 - 2R_1} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 2 \\ 0 & -3 & -2 & -3 \end{array} \right) \\ &\xrightarrow{R_2 \rightarrow -\frac{1}{3}R_2} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 2 \\ 0 & 1 & \frac{2}{3} & 1 \end{array} \right) \xrightarrow{R_1 \rightarrow R_1 - 2R_2} \left(\begin{array}{cc|cc} 1 & 0 & -\frac{1}{3} & 0 \\ 0 & 1 & \frac{2}{3} & 1 \end{array} \right). \end{aligned}$$

Pertanto l'equazione (5.2.2) è equivalente a

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} X = \begin{pmatrix} -1/3 & 0 \\ 2/3 & 1 \end{pmatrix},$$

ovvero al sistema a incognite matriciali

$$\begin{cases} X_1 = (-1/3 & 0) \\ X_2 = (2/3 & 1) \end{cases}$$

che, come unica soluzione, ha ovviamente la matrice

$$\begin{pmatrix} -1/3 & 0 \\ 2/3 & 1 \end{pmatrix}. \quad \spadesuit$$

Esempio 5.7. Si consideri l'equazione matriciale

$$\begin{pmatrix} 1 & 2 & 1 \\ 1 & -1 & 1 \end{pmatrix} X = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}, \quad (5.2.3)$$

la cui matrice completa è

$$\left(\begin{array}{ccc|cc} 1 & 2 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 & 0 \end{array} \right).$$

L'equazione (5.2.3) equivale al sistema

$$\begin{cases} X_1 + 2X_2 + X_3 = (1 & -1) \\ X_1 - X_2 + X_3 = (1 & 0). \end{cases}$$

Trasformando $(A|B)$ con operazioni elementari di riga otteniamo

$$\begin{aligned} (A|B) &\xrightarrow{R_2 \rightarrow R_2 - R_1} \left(\begin{array}{ccc|cc} 1 & 2 & 1 & 1 & -1 \\ 0 & -3 & 0 & 0 & 1 \end{array} \right) \\ &\xrightarrow{R_2 \rightarrow -R_2/3} \left(\begin{array}{ccc|cc} 1 & 2 & 1 & 1 & -1 \\ 0 & 1 & 0 & 0 & -\frac{1}{3} \end{array} \right) \xrightarrow{R_1 \rightarrow R_1 - 2R_2} \left(\begin{array}{ccc|cc} 1 & 0 & 1 & 1 & -\frac{1}{3} \\ 0 & 1 & 0 & 0 & -\frac{1}{3} \end{array} \right). \end{aligned}$$

Pertanto l'equazione (5.2.3) è equivalente a

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

ovvero a

$$\begin{cases} X_1 + X_3 = (1 & -1/3) \\ X_2 = (0 & -1/3). \end{cases}$$

Quindi l'insieme delle soluzioni dell'equazione (5.2.3) è

$$\left\{ \left(\begin{array}{cc} 1 - x_{3,1} & -1/3 - x_{3,2} \\ 0 & -1/3 \\ x_{3,1} & x_{3,2} \end{array} \right) \mid x_{3,1}, x_{3,2} \in \mathbb{R} \right\}.$$

In particolare le soluzioni dipendono da $1 = 3 - 2 = n - \text{rk}(A)$ righe libere. ♠

Anche per equazioni matriciali vale il Teorema di Rouché–Capelli. Lo enunciamo omettendone la dimostrazione in quanto totalmente analoga a quella della Proposizione 5.2.

Proposizione 5.8 (Teorema di Rouché–Capelli per equazioni matriciali).

Siano date $A \in \mathbb{K}^{m,n}$ e $B \in \mathbb{K}^{m,p}$. Si consideri l'equazione matriciale $AX = B$, con matrice completa associata $(A|B)$. Allora

- (i) l'equazione è compatibile se e solo se $\text{rk}(A) = \text{rk}(A|B)$;
- (ii) se l'equazione è compatibile, allora le matrici $n \times p$ che sono sue soluzioni dipendono da $n - \text{rk}(A)$ righe libere;
- (iii) se l'equazione è compatibile e X_0 è una sua soluzione fissata, allora ogni altra sua soluzione X è della forma $X = X_0 + Y$, dove Y appartiene all'insieme dell'equazione matriciale omogenea associata $AX = 0_{m,p}$.

Esempio 5.9. Si considerino i sistemi degli Esempi 5.3 e 5.4; invece di risolverli separatamente consideriamo l'equazione matriciale

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 3 & -2 & 1 & 2 \\ 2 & 4 & -1 & 2 & 3 \\ 1 & -1 & 2 & 5 & 7 \end{pmatrix} \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \\ d_1 & d_2 \\ e_1 & e_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

La matrice completa associata è

$$(A|B) = \left(\begin{array}{ccccc|cc} 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 3 & -2 & 1 & 2 & 0 & 0 \\ 2 & 4 & -1 & 2 & 3 & 1 & 0 \\ 1 & -1 & 2 & 5 & 7 & 0 & 1 \end{array} \right).$$

Con operazioni elementari per riga, tenendo conto dei già citati Esempi 5.3 e 5.4, possiamo trasformarla nella matrice

$$(A|B) = \left(\begin{array}{ccccc|cc} 1 & 0 & 0 & 6 & \frac{37}{4} & 0 & \frac{1}{2} \\ 0 & 1 & 0 & -3 & -\frac{19}{4} & 0 & 0 \\ 0 & 0 & 1 & -2 & -\frac{4}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{array} \right).$$

Deduciamo che l'equazione è incompatibile perché $\text{rk}(A) = 3 < 4 = \text{rk}(A|B)$ (infatti ogni sua soluzione darebbe una soluzione del sistema avente come colonna dei termini noti la prima colonna di B , che è incompatibile, come già visto nell'Esempio 5.3). ♠

5.3 Calcolo dell'inversa di una matrice

Un caso particolarmente interessante di equazione matriciale è quello della forma $AX = I_n$ ove $A \in \mathbb{K}^{n,n}$ è una matrice quadrata. Infatti A è invertibile se e solo se il sistema $AX = I_n$ ha soluzione e in tal caso l'unica soluzione è $X = A^{-1}$.

Proposizione 5.10.

Sia data $A \in \mathbb{K}^{n,n}$. La matrice A è invertibile se e solo se $\text{rk}(A) = n$.

Dimostrazione. Chiedere che A sia invertibile equivale a chiedere che $AX = I_n$ abbia soluzione. La Proposizione 5.8 indica che questo avviene se e solo se $\text{rk}(A) = \text{rk}(A|I_n)$: quest'ultima matrice è fortemente ridotta per righe e il suo rango è esattamente $\text{rk}(I_n) = n$. \square

La proposizione precedente ci fornisce un metodo per verificare che $A \in \mathbb{K}^{n,n}$ sia invertibile e calcolare A^{-1} , in caso essa esista.

Tale metodo consiste nello scrivere la matrice completa $(A|I_n)$. Con trasformazioni elementari di riga ridurremo tale matrice prima a una matrice $U = (A'|A'')$ in cui A' è ridotta per righe, controllando se $\text{rk}(A') = n$, ovvero se A è invertibile. Se ciò è verificato continueremo con operazioni elementari di riga arrivando a ottenere una matrice fortemente ridotta $V = (A''|A''')$. Su ogni riga di A'' ci deve essere un'entrata pari a 1, poiché $\text{rk}(A'') = \text{rk}(A') = \text{rk}(A) = n$: poiché ci sono n colonne su ogni riga tutte le entrate sono nulle eccetto una che vale 1 e che si trova sempre in una colonna diversa. Quindi, semplicemente con permutazioni di riga, si può ulteriormente trasformare V in una nuova matrice della forma $(I_n|A''')$. A questo punto si osservi che l'equazione di partenza è equivalente a $I_n X = A'''$, dunque $A^{-1} = A'''$.

Esempio 5.11. Si consideri la matrice

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

Vogliamo stabilire se A è invertibile e, in caso affermativo, determinarne l'inversa.

A tale scopo scriviamo la matrice $(A|I_3)$ trasformandola, come spiegato sopra, con operazioni elementari di riga nella nuova matrice $U = (A'|A'')$ in maniera tale che A' sia ridotta per righe:

$$(A|I_3) \xrightarrow{R_3 \rightarrow R_3 + R_1} \left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 3 & 5 & 0 & 1 & 0 & 1 \end{array} \right) \xrightarrow{R_3 \rightarrow R_3 - 3R_1} \left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 & -3 & 1 \end{array} \right) = U.$$

Notiamo che i passaggi fatti sono sufficienti per calcolare che $\text{rk}(A) = 3$, e dunque che A è invertibile, perciò ha senso continuare il calcolo di A^{-1} trasformando U nella

matrice $V = (A''|A''')$ con A'' fortemente ridotta per righe:

$$U \xrightarrow{R_3 \rightarrow R_3/2} \left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} \end{array} \right) \xrightarrow{\substack{R_2 \rightarrow R_2 - R_3 \\ R_1 \rightarrow R_1 - 2R_3}} \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 0 & 3 & -1 \\ 1 & 0 & 0 & -\frac{1}{2} & \frac{5}{2} & -\frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} \end{array} \right) \\ \xrightarrow{R_1 \rightarrow -R_1} \left(\begin{array}{ccc|ccc} -1 & 0 & 1 & 0 & -3 & 1 \\ 1 & 0 & 0 & -\frac{1}{2} & \frac{5}{2} & -\frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} \end{array} \right) \xrightarrow{R_1 \rightarrow R_1 + R_2} \left(\begin{array}{ccc|ccc} 0 & 0 & 1 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 & -\frac{1}{2} & \frac{5}{2} & -\frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} \end{array} \right) = V.$$

A questo punto A'' è fortemente ridotta per righe. Possiamo allora procedere a riordinare le righe di V :

$$V \xrightarrow{R_1 \leftrightarrow R_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{2} & \frac{5}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{1}{2} & -\frac{3}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} \end{array} \right) \xrightarrow{R_2 \leftrightarrow R_3} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{2} & \frac{5}{2} & -\frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} \\ 0 & 0 & 1 & -\frac{1}{2} & -\frac{3}{2} & \frac{1}{2} \end{array} \right) = (A''''|I_3).$$

Concludiamo che

$$A^{-1} = \frac{1}{2} \begin{pmatrix} -1 & 5 & -1 \\ 1 & -3 & 1 \\ -1 & -1 & 1 \end{pmatrix}. \quad \spadesuit$$

5.4 Operazioni elementari di colonna

Così come sono state definite le nozioni di operazione elementare di riga e di matrice (fortemente) ridotta per righe, possono essere introdotte analoghe nozioni di operazione elementare di colonna e di matrice (fortemente) ridotta per colonne. Giusto per fissare la notazione, scriveremo:

- (E1) $C_i \rightarrow C_i + \alpha C_{i_0}$ denota l'operazione di sommare alla colonna di indice i la riga di indice $i_0 \neq i$ moltiplicata per $\alpha \in \mathbb{K}$;
- (E2) $C_i \rightarrow \alpha C_i$ denota l'operazione di moltiplicare la colonna di indice i per $\alpha \in \mathbb{K}$, $\alpha \neq 0$;
- (E3) $C_i \leftrightarrow C_{i_0}$ denota l'operazione di scambiare le colonne indici i e i_0 .

Si noti che possiamo descrivere matricialmente anche questo tipo di operazioni, definendo la nozione di matrice elementare anche rispetto alle operazioni elementari di colonna. Precisamente vale il seguente risultato.

Proposizione 5.12.


Sia data $A \in \mathbb{K}^{m,n}$. Sia $A' \in \mathbb{K}^{m,n}$ ottenuta da A con un'operazione elementare di colonna e sia $E \in \mathbb{K}^{n,n}$ la matrice elementare ottenuta con la stessa operazione di colonna da I_n . Allora

$$A' = AE.$$

Dimostrazione. La dimostrazione è analoga a quella della Proposizione 4.8 e, per questo, è lasciata al lettore. \square

Con ovvio significato dei simboli si verifichi che

$$\begin{aligned} E_{C_i \rightarrow C_i + \alpha C_{i_0}} &= E_{R_{i_0} \rightarrow R_{i_0} + \alpha R_i} = {}^t E_{R_i \rightarrow R_i + \alpha R_{i_0}} = {}^t E_{C_{i_0} \rightarrow C_{i_0} + \alpha C_i}, \\ E_{C_i \rightarrow \alpha C_i} &= E_{R_i \rightarrow \alpha R_i} = {}^t E_{R_i \rightarrow \alpha R_i} = {}^t E_{C_i \rightarrow \alpha C_i}, \\ E_{C_i \leftrightarrow C_{i_0}} &= E_{R_i \leftrightarrow R_{i_0}} = {}^t E_{R_i \leftrightarrow R_{i_0}} = {}^t E_{C_i \leftrightarrow C_{i_0}}. \end{aligned}$$

 **Attenzione!** Nei paragrafi precedenti abbiamo introdotto la nozione di operazione elementare di riga e l'abbiamo utilizzata per risolvere equazioni matriciali della forma $AX = B$, con $A \in \mathbb{K}^{m,n}$ e $B \in \mathbb{K}^{m,p}$. Questo è possibile perché all'equazione matriciale $AX = B$ sostituiamo $(EA)X = (EB)$, dove E è elementare di ordine m , e tale seconda equazione è chiaramente equivalente alla precedente (perché?).

È chiaro che nella risoluzione delle equazioni matriciali non ha senso (anzi porta a sicuri errori) utilizzare le operazioni elementari di colonna. Infatti è evidente che le equazioni $AX = B$, con $A \in \mathbb{K}^{m,n}$ e $B \in \mathbb{K}^{m,p}$, e $(AE)X = (BE)$, dove E è elementare di ordine n , non sono equivalenti, anche perché in generale la seconda non esiste nemmeno!

Determinanti

6.1 La definizione di determinante

In questo capitolo affronteremo da un punto di vista prettamente operativo la nozione di determinante, descrivendone le proprietà e i metodi di calcolo, senza entrare nei dettagli e omettendo quasi tutte le dimostrazioni.

Definizione 6.1 (Sottomatrici).

Siano dati $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathbb{K}^{m,n}$ e $p, q \in \mathbb{Z}$ positivi tali che $1 \leq p \leq m$ e $1 \leq q \leq n$.

Una *sottomatrice* $p \times q$ di A è una matrice in $\mathbb{K}^{p,q}$ ottenuta da A considerando solo le entrate che si trovano all'intersezione di p righe e q colonne. In alternativa una sottomatrice $p \times q$ può essere pensata come ottenuta da A cancellando $m - p$ righe e $n - q$ colonne fissate.

Esempio 6.2. Si consideri la matrice

$$A = \begin{pmatrix} 1 & 2 & -3 \\ -i & 0 & -3 \end{pmatrix} \in \mathbb{C}^{2,3}.$$

Le sottomatrici 2×2 di A sono

$$\begin{pmatrix} 1 & 2 \\ -i & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & -3 \\ -i & -3 \end{pmatrix}, \quad \begin{pmatrix} 2 & -3 \\ 0 & -3 \end{pmatrix}.$$

Invece la matrice

$$\begin{pmatrix} 2 & -3 \\ -i & -3 \end{pmatrix}$$

non è sottomatrice di A .

Si consideri poi la matrice

$$B = \begin{pmatrix} 1 & 1 & 0 \\ -\pi & 0 & -2/5 \\ \pi & 5 & 5 \end{pmatrix} \in \mathbb{R}^{3,3}.$$

Le sottomatrici 2×3 di B sono

$$\begin{pmatrix} 1 & 1 & 0 \\ -\pi & 0 & -2/5 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 \\ \pi & 5 & 5 \end{pmatrix}, \quad \begin{pmatrix} -\pi & 0 & -2/5 \\ \pi & 5 & 5 \end{pmatrix}. \quad \spadesuit$$

Daremo adesso una definizione induttiva del determinante.

Definizione 6.3 (Determinante).

Sia data $A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathbb{K}^{n,n}$. Il *determinante* di A è il numero di \mathbb{K} , indicato con $\det(A)$ o $|A|$, definito come segue.

- Se $n = 1$, $\det(a_{1,1}) = a_{1,1}$.
- Se $n \geq 2$,

$$\det(A) = a_{1,1}A_{1,1} + a_{1,2}A_{1,2} + \cdots + a_{1,n}A_{1,n} = \sum_{j=1}^n a_{1,j}A_{1,j}, \quad (6.1.1)$$

ove $A_{i,j}$ indica il prodotto di $(-1)^{i+j}$ per il determinante della sottomatrice ottenuta da A cancellando la riga di indice i e la colonna di indice j .

La formula sopra viene detta *sviluppo di Laplace secondo la prima riga*, mentre il numero $A_{i,j}$ viene detto *complemento algebrico* o *cofattore* dell'entrata $a_{i,j}$.

Osservazione 6.4. In base alla definizione, se $A = (a_{i,j})_{1 \leq i,j \leq n} \in K^{n,n}$ è triangolare inferiore, il calcolo del suo determinante si riduce al prodotto delle entrate sulla diagonale. Infatti

$$\begin{aligned} \begin{vmatrix} a_{1,1} & 0 & 0 & \cdots & 0 \\ a_{2,1} & a_{2,2} & 0 & \cdots & 0 \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \cdots & a_{n,n} \end{vmatrix} &= a_{1,1} \begin{vmatrix} a_{2,2} & 0 & \cdots & 0 \\ a_{3,2} & a_{3,3} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,2} & a_{n,3} & \cdots & a_{n,n} \end{vmatrix} \\ &= a_{1,1} a_{2,2} \begin{vmatrix} a_{3,3} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ a_{n,3} & \cdots & a_{n,n} \end{vmatrix} = \cdots = a_{1,1} a_{2,2} a_{3,3} \cdots a_{n,n}. \end{aligned}$$

Per esempio

$$\begin{vmatrix} 3 & 0 & 0 \\ \sqrt{327} & -5 & 0 \\ 117 & \pi & 1/15 \end{vmatrix} = 3 \cdot (-5) \cdot 1/15 = -1$$

e, quando $i > i_0$,

$$\det(E_{R_i \rightarrow R_i + \alpha R_{i_0}}) = 1.$$

Ciò vale anche per le matrici diagonali. In particolare $\det(I_n) = 1$ e

$$\det(E_{C_i \rightarrow \alpha C_i}) = \det(E_{R_i \rightarrow \alpha R_i}) = \alpha.$$

Osserviamo che il costo computazionale del calcolo di un determinante diviene sempre più oneroso man mano l'ordine della matrice quadrata cresce: infatti per calcolare il determinante di una matrice $A \in K^{n,n}$ bisogna calcolare n determinanti di matrici d'ordine $n-1$, e per ciascuno di essi $n-1$ determinanti di matrici quadrate d'ordine $n-2$, e così via. In totale quindi per calcolare il determinante di una matrice $A \in K^{n,n}$ è necessario moltiplicare

$$n! = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot 2 \cdot 1$$

determinanti di matrici quadrate d'ordine 1 (cioè entrate di A) e sommarli dopo averli moltiplicati per un'opportuna potenza di -1 .

Diamo alcuni esempi di determinanti di matrici quadrate di dimensione piccola.

Esempio 6.5. Si consideri una matrice quadrata generale di dimensione 2

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

Calcolando i cofattori risulta

$$A_{1,1} = (-1)^{1+1} \det(a_{2,2}) = a_{2,2} \quad A_{1,2} = (-1)^{1+2} \det(a_{2,1}) = -a_{2,1},$$

quindi concludiamo che

$$\det(A) = \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} = a_{1,1}A_{1,1} + a_{1,2}A_{1,2} = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}.$$

Per esempio

$$\begin{aligned} \begin{vmatrix} 1 & 2 \\ -i & 0 \end{vmatrix} &= 1 \cdot 0 - 2 \cdot (-i) = 0 + 2i = 2i, \\ \begin{vmatrix} 1 & 2 \\ 4 & 1 \end{vmatrix} &= 1 \cdot 1 - 2 \cdot 4 = 1 - 8 = -7, \\ \begin{vmatrix} 1 & 2 \\ -2 & -4 \end{vmatrix} &= 1 \cdot (-4) - 2 \cdot (-2) = -4 + 4 = 0. \end{aligned}$$

Si consideri adesso una matrice quadrata generale di dimensione 3

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

Calcolando i cofattori risulta

$$\begin{aligned} A_{1,1} &= (-1)^{1+1} \begin{vmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} = a_{2,2}a_{3,3} - a_{2,3}a_{3,2}, \\ A_{1,2} &= (-1)^{1+2} \begin{vmatrix} a_{2,1} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{vmatrix} = -a_{2,1}a_{3,3} + a_{2,3}a_{3,1}, \\ A_{1,3} &= (-1)^{1+3} \begin{vmatrix} a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{vmatrix} = a_{2,1}a_{3,2} - a_{2,2}a_{3,1}. \end{aligned}$$

Concludiamo che

$$\begin{aligned}
 |A| &= \begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} = a_{1,1}A_{1,1} + a_{1,2}A_{1,2} + a_{1,3}A_{1,3} \\
 &= a_{1,1}(a_{2,2}a_{3,3} - a_{2,3}a_{3,2}) + a_{1,2}(-a_{2,1}a_{3,3} + a_{2,3}a_{3,1}) + a_{1,3}(a_{2,1}a_{3,2} - a_{2,2}a_{3,1}),
 \end{aligned}$$

quindi

$$\begin{aligned}
 |A| &= a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} \\
 &\quad - a_{1,3}a_{2,2}a_{3,1} - a_{1,1}a_{2,3}a_{3,2} - a_{1,2}a_{2,1}a_{3,3}.
 \end{aligned} \tag{6.1.2}$$

Per esempio, consideriamo la matrice A dell'Esempio 5.11. Si ha

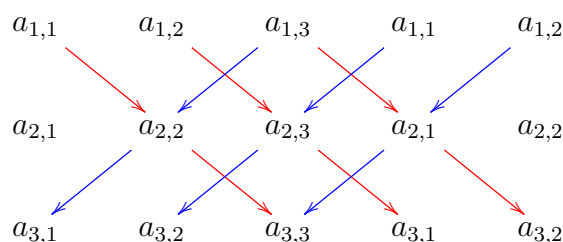
$$\begin{aligned}
 \begin{vmatrix} 1 & 2 & -1 \\ 1 & 1 & 0 \\ 2 & 3 & 1 \end{vmatrix} &= 1 \cdot 1 \cdot 1 + 2 \cdot 0 \cdot 2 + (-1) \cdot 1 \cdot 3 - (-1) \cdot 1 \cdot 2 - 0 \cdot 3 \cdot 1 - 2 \cdot 1 \cdot 1 \\
 &= 1 + 0 - 3 + 2 - 0 - 2 = -2.
 \end{aligned}$$



La formula (6.1.2) viene spesso chiamata *regola di Sarrus* ed esiste un metodo per ricordarsela, che consiste nel ricopiare le prime due colonne della matrice sulla destra in questo modo:

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} \begin{matrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{matrix}$$

quindi calcolare il determinante come la differenza tra la somma dei prodotti dei termini collegati dalle frecce rosse che partono dall'alto a sinistra dirette verso il basso a destra, meno la somma dei prodotti dei termini collegati dalle frecce blu che partono dall'alto a destra dirette verso il basso a sinistra:



Il lettore verifichi che in questo modo si riottiene la formula (6.1.2).

☠ **Attenzione!** La regola di Sarrus vale solo per matrici quadrate di ordine 3. Se ad esempio volessimo ripetere il procedimento per matrici 4×4 , otterremmo una somma algebrica di otto prodotti. Avevamo però già osservato che per calcolare il determinante di una matrice di ordine n è necessario sommare $n!$ prodotti, e ovviamente $4! = 24$ è ben più grande di 8.

Un altro aiuto mnemonico riguarda il segno per cui bisogna moltiplicare il determinante della sottomatrice cofattore (i, j) nel calcolo di $A_{i,j}$: alla matrice

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} & \dots \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} & \dots \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} & \dots \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

si può associare la “matrice di segni”

$$\begin{pmatrix} + & - & + & - & \dots \\ - & + & - & + & \dots \\ + & - & + & - & \dots \\ - & + & - & + & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

così per determinare il segno $(-1)^{i+j}$ è sufficiente guardare la corrispondente entrata di posizione (i, j) della matrice dei segni.

Esempio 6.6. Grazie a quanto visto sopra è facile convincersi che

$$\det(E_{R_i \leftrightarrow R_{i_0}}) = \det(E_{C_i \leftrightarrow C_{i_0}}) = -1$$

se $i \neq i_0$. Per esempio si consideri la matrice

$$E_{R_1 \leftrightarrow R_n} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \end{pmatrix} \in K^{n,n}.$$

Allora

$$\det(E_{R_1 \leftrightarrow R_n}) = (-1)^{1+n} = \begin{vmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \end{vmatrix} = (-1)^{1+n} (-1)^{1+2} \dots (-1)^{1+2} (-1)^{1+1},$$

dove il fattore $(-1)^{1+2}$ è ripetuto esattamente $n - 2$ volte. Concludiamo che

$$\det(E_{R_1 \leftrightarrow R_n}) = (-1)^{1+n+3(n-2)} = (-1)^{4n-5} = -1.$$



6.2 Proprietà e calcolo di determinanti

Per semplificare il calcolo dei determinanti è utile seguire un metodo che si basa sul loro buon comportamento rispetto alle operazioni elementari di riga. Tratteremo di tale argomento in questo paragrafo.

Premettiamo due risultati molto importanti di cui omettiamo la dimostrazione.

Proposizione 6.7.

Sia data $A \in \mathbb{K}^{n,n}$. Risulta $\det({}^tA) = \det(A)$.

Esempio 6.8. Quanto visto nell'Osservazione 6.4 può essere quindi esteso anche a matrici triangolari superiori: infatti la trasposta di una matrice triangolare superiore è triangolare inferiore e matrici trasposte hanno lo stesso determinante.

Per esempio

$$\begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 \\ 3 & 2 & 1 & 0 & 0 \\ 4 & 3 & 2 & 1 & 0 \\ 5 & 4 & 3 & 2 & 1 \end{vmatrix} = 1.$$

Inoltre

$$E_{R_i \rightarrow R_i + \alpha R_{i_0}} = 1.$$

per qualsiasi scelta di i e i_0 . ♠

Proposizione 6.9 (Teorema di Binet).

Siano date $A, B \in K^{n,n}$. Risulta

$$\det(AB) = \det(A) \det(B).$$

☠ **Attenzione!** $\det(A + B) \neq \det(A) + \det(B)$.

Esempio 6.10. Si considerino le matrici

$$E_{1,1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad E_{2,2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Chiaramente $\det(E_{1,1}) = \det(E_{2,2}) = 0$ e $\det(E_{1,1} + E_{2,2}) = \det(I_2) = 1$; in particolare $\det(E_{1,1}) + \det(E_{2,2}) \neq \det(E_{1,1} + E_{2,2})$. ♠

La seguente proposizione è un'importante conseguenza immediata dei due enunciati precedenti.

Proposizione 6.11.

Sia data $A \in \mathbb{K}^{n,n}$. Valgono le seguenti proprietà:

- (D1) se A' è ottenuta da A sommando a una riga (colonna) un multiplo di un'altra, allora vale l'uguaglianza $\det(A') = \det(A)$;
- (D2) se A' è ottenuta da A moltiplicando una riga (colonna) per una costante $\alpha \in \mathbb{K}$, allora vale l'uguaglianza $\det(A') = \alpha \det(A)$;
- (D3) se A' è ottenuta da A scambiando due righe (colonne) diverse, allora vale l'uguaglianza $\det(A') = -\det(A)$.

Dimostrazione. Ricordiamo che ogni matrice $A' \in \mathbb{K}^{n,n}$ ottenuta da $A \in \mathbb{K}^{n,n}$ tramite un'operazione elementare di riga o colonna si ottiene anche moltiplicando A , rispettivamente a sinistra o destra, per una matrice elementare $E \in \mathbb{K}^{n,n}$ ottenuta dalla matrice identità $I_n \in \mathbb{K}^{n,n}$ tramite la stessa operazione, come spiegato nelle Proposizioni 4.8 e 5.12.

In particolare tenendo conto dell'Osservazione 6.4 e degli Esempi 6.8 e 6.8 se

$$A \xrightarrow{R_i \rightarrow R_i + \alpha R_{i_0}} A'$$

allora $\det(A') = \det(A)$; se

$$A \xrightarrow{R_i \rightarrow \alpha R_i} A'$$

allora $\det(A') = \alpha \det(A)$; se

$$A \xrightarrow{R_i \leftrightarrow R_{i_0}} A'$$

allora $\det(A') = -\det(A)$. □

Corollario 6.12 (Sviluppo del determinante con la regola di Laplace).

Sia data $A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathbb{K}^{n,n}$. Risulta

$$\det(A) = a_{i,1}A_{i,1} + a_{i,2}A_{i,2} + \cdots + a_{i,n}A_{i,n} = a_{1,j}A_{1,j} + a_{2,j}A_{2,j} + \cdots + a_{n,j}A_{n,j}.$$

Dimostrazione. La dimostrazione si basa su un uso ripetuto della proprietà (D3) della Proposizione 6.11 e viene lasciata al lettore. □

La formula dell'enunciato del Corollario 6.12 può essere scritta in forma più compatta come

$$\det(A) = \sum_{i=1}^n a_{i,j}A_{i,j} = \sum_{j=1}^n a_{i,j}A_{i,j}.$$

Tali formule vengono dette rispettivamente *sviluppo di Laplace secondo la i -esima riga* e *sviluppo di Laplace secondo la j -esima colonna*.

Passiamo ora a descrivere un metodo più efficiente di calcolo del determinante di una matrice quadrata $A \in \mathbb{K}^{n,n}$ anch'esso basato sulla Proposizione 6.11. L'idea generale è la seguente.

Prima si trasforma con operazioni elementari di tipo (E1) la matrice A in una matrice A' ridotta per righe. In particolare, in base alla Proposizione 4.8,

$$A' = E_1 \cdots E_s A,$$

dove le matrici E_i sono elementari e rappresentano operazioni elementari (di riga) di tipo (E1), quindi tutte con determinante 1 (si veda l'Esempio 6.8). Pertanto il Teorema di Binet implica

$$\det(A) = \det(E_1) \cdots \det(E_s) \det(A) = \det(E_1 \cdots E_s A) = \det(A').$$

Poi si trasforma A' con scambi di colonna in una matrice triangolare superiore A'' . In particolare, in base alla Proposizione 5.12,

$$A'' = A' E'_1 \cdots E'_h,$$

dove le matrici E'_j sono elementari e rappresentano operazioni elementari (di colonna) di tipo (E3), quindi hanno tutte con determinante -1 (si veda l'Esempio 6.6). Pertanto il Teorema di Binet implica

$$\begin{aligned} \det(A) &= \det(A') = (-1)^h \det(E'_1) \cdots \det(E'_h) \det(A') \\ &= (-1)^h \det(E'_1 \cdots E'_h A') = (-1)^h \det(A''). \end{aligned}$$

Si noti che, per la generica matrice $A \in \mathbb{K}^{n,n}$, tale metodo permette di ridurre il numero di operazioni da $n \cdot n!$ a circa $2n^3/3$.

Esempio 6.13. Si consideri la matrice quadrata di dimensione 3 dell'Esempio 6.5. Possiamo ripetere il calcolo del determinante osservando che

$$\begin{aligned} & \left| \begin{array}{ccc} 1 & 2 & -1 \\ 1 & 1 & 0 \\ 2 & 3 & 1 \end{array} \right| \xrightarrow{\underline{R_3 \rightarrow R_3 + R_1}} \left| \begin{array}{ccc} 1 & 2 & -1 \\ 1 & 1 & 0 \\ 3 & 5 & 0 \end{array} \right| \xrightarrow{\underline{R_3 \rightarrow R_3 - 3R_2}} \left| \begin{array}{ccc} 1 & 2 & -1 \\ 1 & 1 & 0 \\ 0 & 2 & 0 \end{array} \right| \\ & \xrightarrow{\underline{C_1 \leftrightarrow C_3}} - \left| \begin{array}{ccc} -1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 2 & 0 \end{array} \right| \xrightarrow{\underline{C_2 \leftrightarrow C_3}} (-1)^2 \left| \begin{array}{ccc} -1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{array} \right| \\ & = (-1)^2 \cdot (-1) \cdot 1 \cdot 2 = -2. \quad \spadesuit \end{aligned}$$

Osserviamo che anche la proprietà (D1) ha un'importante conseguenza: consideriamo $A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathbb{K}^{n,n}$, e indichiamo con R_i la sua riga di indice i , cioè la matrice

$$R_i = (a_{i,1} \quad a_{i,2} \quad \dots \quad a_{i,n}).$$

Supponiamo, per semplicità, che esistano $\alpha_1, \dots, \alpha_{n-1} \in K$ tali che l'ultima riga si scrive come

$$R_n = \alpha_1 R_1 + \cdots + \alpha_{n-1} R_{n-1}.$$

Allora con operazioni elementari di riga (E1) della forma

$$A \xrightarrow{R_n \rightarrow R_n - \alpha_1 R_1 - \cdots - \alpha_{n-1} R_{n-1}} A'$$

si ottiene una nuova matrice A' avente la riga di indice n nulla. Sviluppando il determinante secondo l'ultima riga si ottiene allora $\det(A) = \det(A') = 0$.

Un discorso analogo vale sostituendo nel ragionamento sopra la parola riga con la parola colonna.

Un caso ancor più particolare è quello in cui la matrice di cui dobbiamo calcolare il determinante ha due righe o due colonne uguali; allora la Proposizione 6.11 ci dice subito che tale determinante è necessariamente nullo.

Esempio 6.14. Si consideri la matrice

$$A = \begin{pmatrix} 1 & -17 & 2 \\ -1 & 17 & 171/35 \\ 2 & -34 & \sqrt{\pi} \end{pmatrix}.$$

Poiché le colonne C_j di A sono legate dalla relazione $C_2 = -17C_1 + 0C_3$ segue da quanto visto sopra che $\det(A) = 0$.

Per esercizio si verifichi la nullità di $\det(A)$ calcolandolo con uno qualsiasi dei metodi sopra descritti. ♠

Esempio 6.15 (Determinante di Vandermonde). Un'altra applicazione interessante della Proposizione 6.11 è il calcolo dei cosiddetti determinanti di matrici di Vandermonde o, più semplicemente, determinanti di Vandermonde, che si utilizzano nel contesto dell'interpolazione polinomiale.

Dati $x_1, x_2, \dots, x_n \in \mathbb{R}$, la *matrice di Vandermonde relativa a x_1, x_2, \dots, x_n* è

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-2} & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-2} & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \dots & x_3^{n-2} & x_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_{n-1} & x_{n-1}^2 & \dots & x_{n-1}^{n-2} & x_{n-1}^{n-1} \\ 1 & x_n & x_n^2 & \dots & x_n^{n-2} & x_n^{n-1} \end{pmatrix}.$$

Il determinante della matrice di Vandermonde è detto *determinante di Vandermonde relativo a x_1, x_2, \dots, x_n* e si denota con $V(x_1, x_2, \dots, x_n)$; per calcolarlo in funzione dei numeri x_1, x_2, \dots, x_n si può procedere con operazioni elementari. Per esempio se $n = 4$, allora

$$\begin{aligned}
V(x_1, x_2, x_3, x_4) & \xrightarrow{\underline{\underline{C_4 \rightarrow C_4 - x_1 C_3}}} \begin{vmatrix} 1 & x_1 & x_1^2 & 0 \\ 1 & x_2 & x_2^2 & x_2^3 - x_1 x_2^2 \\ 1 & x_3 & x_3^2 & x_3^3 - x_1 x_3^2 \\ 1 & x_4 & x_4^2 & x_4^3 - x_1 x_4^2 \end{vmatrix} = \begin{vmatrix} 1 & x_1 & x_1^2 & 0 \\ 1 & x_2 & x_2^2 & (x_2 - x_1)x_2^2 \\ 1 & x_3 & x_3^2 & (x_3 - x_1)x_3^2 \\ 1 & x_4 & x_4^2 & (x_4 - x_1)x_4^2 \end{vmatrix} \\
& \xrightarrow{\underline{\underline{C_3 \rightarrow C_3 - x_1 C_2}}} \begin{vmatrix} 1 & x_1 & 0 & 0 \\ 1 & x_2 & (x_2 - x_1)x_2 & (x_2 - x_1)x_2^2 \\ 1 & x_3 & (x_3 - x_1)x_2 & (x_3 - x_1)x_3^2 \\ 1 & x_4 & (x_4 - x_1)x_2 & (x_4 - x_1)x_4^2 \end{vmatrix} \\
& \xrightarrow{\underline{\underline{C_2 \rightarrow C_2 - x_1 C_1}}} \begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & x_2 - x_1 & (x_2 - x_1)x_2 & (x_2 - x_1)x_2^2 \\ 1 & x_3 - x_1 & (x_3 - x_1)x_2 & (x_3 - x_1)x_3^2 \\ 1 & x_4 - x_1 & (x_4 - x_1)x_2 & (x_4 - x_1)x_4^2 \end{vmatrix} \\
& = \begin{vmatrix} x_2 - x_1 & (x_2 - x_1)x_2 & (x_2 - x_1)x_2^2 \\ x_3 - x_1 & (x_3 - x_1)x_3 & (x_3 - x_1)x_3^2 \\ x_4 - x_1 & (x_4 - x_1)x_4 & (x_4 - x_1)x_4^2 \end{vmatrix} \\
& = (x_2 - x_1)(x_3 - x_1)(x_4 - x_1) \begin{vmatrix} 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ 1 & x_4 & x_4^2 \end{vmatrix}.
\end{aligned}$$

L'ultimo determinante è ancora di Vandermonde, precisamente è $V(x_2, x_3, x_4)$ ma di dimensione più piccola, cioè $n = 3$. Ripetendo il ragionamento otteniamo alla fine che

$$V(x_1, x_2, x_3, x_4) = (x_2 - x_1)(x_3 - x_1)(x_4 - x_1)(x_3 - x_2)(x_4 - x_2)(x_4 - x_3).$$

In generale lo stesso procedimento induttivo permette di dimostrare la formula

$$V(x_1, x_2, \dots, x_n) = \prod_{\substack{i,j=1,\dots,n \\ i>j}} (x_i - x_j). \quad (6.2.1)$$

Per esempio abbiamo

$$\begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 & 16 \\ 1 & -2 & 4 & -8 & 16 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{vmatrix} = V(1, 2, -2, -1, 0),$$

quindi

$$\begin{aligned}
& V(1, 2, -2, -1, 0) \\
& = (2-1)(-2-1)(-1-1)(-1)(-2-2)(-1-2)(-2)(-1+2)(2)(1) = 288.
\end{aligned}$$

Si noti che la formula (6.2.1) permette di affermare che $V(x_1, x_2, \dots, x_n) = 0$ se e solo se gli x_1, x_2, \dots, x_n non sono tutti distinti. ♠

6.3 Ancora sull'inversa di matrici

In questo paragrafo illustreremo il legame fra la nozione di determinante, di invertibilità e di inversa di una matrice quadrata $A \in \mathbb{K}^{n,n}$.

Definizione 6.16 (Matrice aggiunta).

Sia data $A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathbb{K}^{n,n}$. Definiamo *aggiunta di A* la matrice $\text{adj}(A)$ la cui entrata di posizione (i, j) è il complemento algebrico $A_{j,i}$ dell'entrata $a_{j,i}$.

Data una matrice $A \in \mathbb{K}^{n,n}$, consideriamo il prodotto di A per la sua aggiunta: $B = A \text{adj}(A)$. Per definizione, la sua entrata $b_{i,j}$ di indice (i, j) è il prodotto della riga di indice i di A per la colonna di indice j di $\text{adj}(A)$, cioè

$$b_{i,j} = (a_{i,1} \quad a_{i,2} \quad \dots \quad a_{i,n}) \begin{pmatrix} A_{j,1} \\ A_{j,2} \\ \vdots \\ A_{j,n} \end{pmatrix} = a_{i,1}A_{j,1} + a_{i,2}A_{j,2} + \dots + a_{i,n}A_{j,n}. \quad (6.3.1)$$

Se $i = j$ la formula (6.3.1) implica che $b_{i,i} = \det(A)$.

Consideriamo ora l'entrata $b_{i,j}$ con $i \neq j$: per fissare le idee scegliamo $i = 2$ e $j = 1$ (gli altri casi sono analoghi). Si ha

$$\begin{aligned} b_{2,1} &= a_{2,1}A_{1,1} + a_{2,2}A_{1,2} + \dots \\ &= a_{2,1}(-1)^{1+1} \begin{vmatrix} a_{2,2} & a_{2,3} & \dots \\ a_{3,2} & a_{3,3} & \dots \\ \vdots & \vdots & \ddots \end{vmatrix} + a_{2,2}(-1)^{2+1} \begin{vmatrix} a_{2,1} & a_{2,3} & \dots \\ a_{3,1} & a_{3,3} & \dots \\ \vdots & \vdots & \ddots \end{vmatrix} + \dots \\ &= \begin{vmatrix} a_{2,1} & a_{2,2} & a_{2,3} & \dots \\ a_{2,1} & a_{2,2} & a_{2,3} & \dots \\ a_{3,1} & a_{3,2} & a_{3,3} & \dots \\ \vdots & \vdots & \ddots & \ddots \end{vmatrix} \xrightarrow{R_1 \rightarrow R_1 - R_2} \begin{vmatrix} 0 & 0 & 0 & \dots \\ a_{2,1} & a_{2,2} & a_{2,3} & \dots \\ a_{3,1} & a_{3,2} & a_{3,3} & \dots \\ \vdots & \vdots & \ddots & \ddots \end{vmatrix} = 0 \end{aligned}$$

Concludiamo che

$$b_{i,j} = \begin{cases} \det(A) & \text{se } i = j \\ 0 & \text{se } i \neq j. \end{cases}$$

In maniera simile otteniamo anche $\text{adj}(A)A = \det(A)I_n$ utilizzando la formula (6.3.1) Abbiamo quindi dimostrato la seguente proposizione.

Proposizione 6.17.

Sia data $A \in \mathbb{K}^{n,n}$. Risulta

$$A \text{adj}(A) = \text{adj}(A)A = \det(A)I_n.$$

Esempio 6.18. Consideriamo ancora la matrice dell'Esempio 6.5 (e dell'Esempio 6.13)

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

Possiamo calcolare

$$\begin{array}{lll} A_{1,1} = 1, & A_{1,2} = -1, & A_{1,3} = 1, \\ A_{2,1} = -5, & A_{2,2} = 3, & A_{2,3} = 1 \\ A_{3,1} = 1, & A_{3,2} = -1, & A_{3,3} = -1, \end{array}$$

quindi

$$\text{adj}(A) = \begin{pmatrix} 1 & -5 & 1 \\ -1 & 3 & -1 \\ 1 & 1 & -1 \end{pmatrix}.$$

Si può verificare direttamente che $A \text{adj}(A) = \text{adj}(A)A = -2I_3$. ♠

Dalla Proposizione 6.17 otteniamo il seguente importante corollario.

Corollario 6.19.

Sia data $A \in \mathbb{K}^{n,n}$. La matrice A è invertibile se e solo se $\det(A) \neq 0$. In tal caso si ha $\det(A^{-1}) = (\det(A))^{-1}$ e inoltre

$$A^{-1} = \frac{1}{\det(A)} \text{adj}(A). \quad (6.3.2)$$

Dimostrazione. Supponiamo che A sia invertibile; cioè che esista la matrice inversa A^{-1} che, come sappiamo, soddisfa la relazione $AA^{-1} = I_n$. Calcolando il determinante di ambo i membri di tale relazione, dal Teorema di Binet segue che

$$1 = \det(AA^{-1}) = \det(A) \det(A^{-1}),$$

dunque $\det(A^{-1}) \neq 0$ e si ha $\det(A^{-1}) = (\det(A))^{-1}$.

Viceversa sia $\det(A) \neq 0$. Allora

$$A \left(\frac{1}{\det(A)} \text{adj}(A) \right) = \frac{1}{\det(A)} A \text{adj}(A) = \frac{1}{\det(A)} (\det(A)I_n) = I_n,$$

cioè A è invertibile e vale la formula (6.3.2). □

Osservazione 6.20. È utile riassumere a questo punto le seguenti importanti proprietà equivalenti che valgono per ogni matrice quadrata $A \in \mathbb{K}^{n,n}$:

- (i) A è invertibile;
- (ii) $\det(A) \neq 0$;
- (iii) $\text{rk}(A) = n$.

L'equivalenza fra le affermazioni (i) e (ii) è conseguenza del Corollario 6.19 L'equivalenza fra le affermazioni (i) e (iii) è dimostrata dalla Proposizione 5.10.

Esempio 6.21. Sia ancora una volta A la matrice degli Esempi 5.11, 6.5, 6.13. Nell'Esempio 6.18 abbiamo calcolato la sua aggiunta

$$\text{adj}(A) = \begin{pmatrix} 1 & -5 & 1 \\ -1 & 3 & -1 \\ 1 & 1 & -1 \end{pmatrix},$$

quindi

$$A^{-1} = \begin{pmatrix} -1/2 & 5/2 & -1/2 \\ 1/2 & -3/2 & 1/2 \\ -1/2 & -1/2 & 1/2 \end{pmatrix},$$

che coincide con quanto avevamo calcolato nell'Esempio 5.11 risolvendo l'equazione matriciale $AX = I_3$. ♠

6.4 Il metodo di Cramer

Consideriamo un sistema di Cramer, ovvero un'equazione matriciale $AX = B$ con $A \in \mathbb{K}^{n,n}$ matrice invertibile; sappiamo che tale equazione ha come unica soluzione la matrice $\bar{X} = A^{-1}B$. Supponiamo ora che $B \in \mathbb{K}^{n,1}$, ovvero che l'equazione $AX = B$ sia, di fatto, un sistema di equazioni lineari.

Il fatto che la matrice $\bar{X} = {}^t(\bar{x}_1 \ \bar{x}_2 \ \dots \ \bar{x}_n)$ sia soluzione significa che vale l'identità di matrici numeriche

$$\bar{x}_1 C_1 + \bar{x}_2 C_2 + \dots + \bar{x}_n C_n = B$$

dove, come già fatto in precedenza, indichiamo con C_j la j -esima colonna di A , cioè $C_j = {}^t(a_{1,j} \ a_{2,j} \ \dots \ a_{n,j})$.

Sia A_j la matrice ottenuta da A sostituendo a C_j la matrice B . Si noti che sottraendo alla j -esima colonna di A_j le matrici $\bar{x}_j C_h$ per $h = 1, \dots, n$ e $h \neq j$, otteniamo una matrice \hat{A}_j che ha tutte le colonne uguali a quelle di A tranne la j -esima che coincide con $\bar{x}_j C_j$.

Tenendo conto della Proposizione 6.11 (precisamente di (D1) e (D2)) segue l'uguaglianza $\det(A_j) = \bar{x}_j \det(A)$.

Proposizione 6.22 (Metodo di Cramer).

Siano date $A \in \mathbb{K}^{n,n}$ invertibile e $B \in \mathbb{K}^{n,1}$. Si indichino $\Delta_j = \det(A_j)$ i determinanti delle matrici ottenute da A sostituendo la colonna B alla j -esima colonna. Sia poi $\Delta = \det(A)$. Allora il sistema $AX = B$ ha come unica soluzione la matrice

$$\frac{1}{\Delta} {}^t(\Delta_1 \ \Delta_2 \ \dots \ \Delta_n).$$

Esempio 6.23. Si consideri il sistema lineare

$$\begin{pmatrix} 1 & 2 & -1 \\ 1 & 1 & 0 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix},$$

avente come matrice incompleta la matrice A dell'Esempio 5.11. Nell'Esempio 6.5 abbiamo calcolato $\delta = \det(A) = -2$. Inoltre

$$\Delta_1 = \begin{vmatrix} 1 & 2 & -1 \\ 2 & 1 & 0 \\ 3 & 3 & 1 \end{vmatrix} = -6, \quad \Delta_2 = \begin{vmatrix} 1 & 1 & -1 \\ 1 & 2 & 0 \\ 2 & 3 & 1 \end{vmatrix} = 2, \quad \Delta_3 = \begin{vmatrix} 1 & 2 & 1 \\ 1 & 1 & 2 \\ 2 & 3 & 3 \end{vmatrix} = 0.$$

Quindi l'unica soluzione di tale sistema è

$$-\frac{1}{2} \begin{pmatrix} -6 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix}.$$



Vettori applicati

7.1 Vettori geometrici

In questo capitolo inizieremo a studiare enti geometrici quali *punti*, *segmenti (orientati)*, *rette*, *piani* sia nel piano S_2 che nello spazio S_3 con l'ordinaria struttura euclidea della distanza: in tali S_n cioè valgono gli assiomi della geometria euclidea.

Per fare questo introdurremo la nozione di vettore applicato (nel piano e nello spazio ordinari), per poi passare dalla geometria all'algebra dei vettori: daremo cioè una struttura "algebrica" a tali enti geometrici che ci permetterà di definire operazioni quali somma e prodotto e di algebrizzare problemi di tipo geometrico.

Per arrivare alla nozione di vettore occorre innanzi tutto ricordare cosa si intende per segmento. In ciò che segue useremo la notazione S_n , intendendo sia il piano ($n = 2$) che lo spazio ($n = 3$) che anche un generale spazio euclideo n -dimensionale, qualsiasi cosa questo voglia dire.

Definizione 7.1 (Segmenti).

Siano A e B punti dello spazio S_n , $n = 2, 3$.

- Se $A \neq B$, esiste un'unica retta r passante per A e B ; in tal caso i due punti dividono r in tre parti: una semiretta di origine A , una semiretta di origine B , e una parte limitata di retta, che verrà detto *segmento di estremi A e B* .
- Se, invece, i due punti coincidono allora la retta r non è più univocamente individuata; continueremo tuttavia a parlare del segmento di estremi A e B intendendo con ciò l'unico punto $A = B$, detto *segmento degenero, o nullo, di estremi A e B* .

Il segmento di estremi A e B si denota con il simbolo \overline{AB} .

Possiamo adesso introdurre il concetto di vettore applicato, che ci permetterà poi di definire i sistemi di coordinate.

Definizione 7.2 (Vettori applicati).

Sia O un punto fissato dello spazio euclideo S_n . Un *vettore \vec{v} applicato in O* è un segmento, eventualmente degenero, avente un estremo in O , detto *estremo vincolato*, e un altro estremo P , detto *estremo libero*.

- Il vettore applicato in O e di estremo libero P è indicato con il simbolo \overrightarrow{OP} ;

- se $O = P$ e il segmento è degenere, scriviamo $\vec{0}$ invece di \overrightarrow{OO} .
- L'insieme di tutti i vettori di S_n applicati in O è indicato con il simbolo $V_n(O)$.

Le scritture \vec{v} o \overrightarrow{OP} ricordano che il segmento è “orientato”; in questo senso le due scritture \overrightarrow{OP} e \overrightarrow{PO} , pur rappresentando lo stesso segmento, rappresentano vettori applicati diversi.

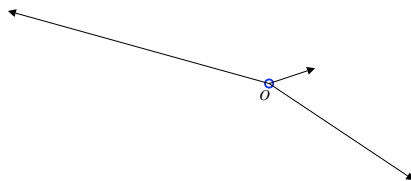


Figura 7.1

A ogni vettore geometrico sono associati alcuni dati importanti.

Definizione 7.3 (Direzione, verso, modulo di un vettore).

Sia dato $\overrightarrow{OP} \in V_n(O)$.

- Se $\overrightarrow{OP} \neq \vec{0}$, definiamo *direzione di \overrightarrow{OP}* la retta passante per i punti O e P ; la direzione del vettore nullo $\vec{0}$ è indeterminata.
- Se $\overrightarrow{OP} \neq \vec{0}$, definiamo *verso di \overrightarrow{OP}* la semiretta di origine O e contenente il punto P ; il verso del vettore nullo $\vec{0}$ è indeterminato.
- Fissata un'unità di misura u in S_n , definiamo *modulo di \overrightarrow{OP}* la lunghezza di \overrightarrow{OP} rispetto all'unità di misura u . Il modulo (o la lunghezza) del vettore \overrightarrow{OP} si indica con $|\overrightarrow{OP}|$. I vettori di $V_n(O)$ di modulo 1 vengono detti *versori*.

Osserviamo che $|\vec{v}| \geq 0$ e risulta $|\vec{v}| = 0$ se e solo se $\vec{v} = \vec{0}$.

⚠ Con la convenzione su direzione e verso del vettore nullo, è chiaro che ogni vettore applicato rimane completamente individuato da direzione, verso e modulo. Si noti perciò che due vettori non nulli coincidono se e solo se hanno stessa direzione, stesso verso e stessa lunghezza.

Definizione 7.4 (Vettori paralleli, concordi, discordi, complanari).

Siano dati $\overrightarrow{OP}, \overrightarrow{OQ}, \overrightarrow{OR} \in V_n(O)$.

- Il vettore \overrightarrow{OP} è *parallelo* a \overrightarrow{OQ} , e scriviamo $\overrightarrow{OP} \parallel \overrightarrow{OQ}$, se i punti O, P e Q sono allineati.
- Se $\overrightarrow{OP} \parallel \overrightarrow{OQ}$ sono non nulli, diciamo che i due vettori sono *concordi* se hanno lo stesso verso, *discordi* se hanno versi distinti.

- I tre vettori \overrightarrow{OP} , \overrightarrow{OQ} e \overrightarrow{OR} sono *complanari* se lo sono i quattro punti O , P , Q e R .

Osserviamo che se $\overrightarrow{OP}, \overrightarrow{OQ} \in V_n(O)$ sono non nulli essi sono paralleli se e solo se hanno la stessa direzione. Tuttavia la Definizione 7.4 si estende anche al caso in cui uno dei due sia nullo: ad esempio se $\overrightarrow{OP} = \vec{0}$, ciò significa che $P = O$, dunque i tre punti O , P e Q risultano ovviamente essere allineati. Quindi, in base alla definizione sopra, il vettore nullo $\vec{0}$ è parallelo a ogni vettore in $V_n(O)$. Un discorso analogo può essere fatto per la nozione di complanarità.

Si noti inoltre che ogni vettore non nullo $\vec{v} \in V_n(O)$ ha esattamente due versori a esso paralleli, uno concorde e uno discorde.

7.2 Sistemi di coordinate nel piano

Possiamo ora introdurre la nozione di sistema di riferimento cartesiano ortogonale nel piano S_2 e nello spazio S_3 . A tale scopo fissiamo una volta per tutte un'unità di misura u in S_n , $n = 2, 3$. Cominciamo dal caso del piano.

Definizione 7.5 (Sistemi di riferimento nel piano).

Un *sistema di riferimento* (cartesiano ortogonale) $O\vec{i}\vec{j}$ in S_2 è definito dai seguenti enti:

- (SRP1) un punto $O \in S_2$ detto *origine* del sistema di riferimento;
- (SRP2) un versore \vec{i} applicato in O ;
- (SRP3) un versore \vec{j} applicato in O tale che il versore \vec{i} si sovrappone al versore \vec{j} con una rotazione di $\pi/2$ radianti intorno a O in senso antiorario.

Nella Figura 7.2 è riportato un esempio di sistema di riferimento nel senso della Definizione 7.5.

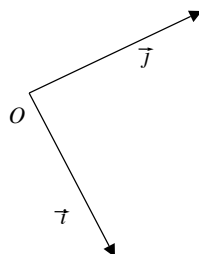


Figura 7.2

Le direzioni dei vettori \vec{i} e \vec{j} vengono dette rispettivamente *asse delle ascisse* (o delle x) e *asse delle ordinate* (o delle y). I versi di \vec{i} e di \vec{j} vengono detti rispettivamente *semiasse positivo delle ascisse* e *semiasse positivo delle ordinate*; similmente i versi opposti a quelli di \vec{i} e di \vec{j} vengono detti rispettivamente *semiasse negativo delle ascisse* e *semiasse negativo delle ordinate*.

Si noti che un sistema di riferimento può anche essere descritto partendo dagli assi e fissando poi i vettori: per questo motivo spesso parleremo di sistema di riferimento Oxy anche se, in questo caso, la notazione è ambigua in quanto non fa riferimento all'unità di misura u .

Procediamo adesso a stabilire, grazie a un sistema di riferimento fissato, una biiezione fra i punti del piano e le coppie ordinate di numeri reali. Sia $P \in S_2$ un qualsiasi punto del piano (eventualmente appartenente agli assi) e si considerino le rette r_x ed r_y passanti per P e parallele rispettivamente all'asse delle ascisse e all'asse delle ordinate: allora la retta r_x intersecherà l'asse delle ordinate in un punto corrispondente al numero reale y_P (detto *ordinata di P*), mentre la retta r_y intersecherà l'asse delle ascisse in un punto corrispondente al numero x_P (detto *ascissa di P*). Al punto P resta associata la coppia ordinata (x_P, y_P) .

In tale corrispondenza il punto di coordinate $(0, 0)$ è l'origine O , i punti di ascissa nulla sono i punti dell'asse delle y , i punti di ordinata nulla sono i punti dell'asse delle x e i punti con coordinate entrambe diverse da zero sono i punti non appartenenti agli assi coordinati.

Nella Figura 7.3 riportiamo gli assi delle ascisse e delle ordinate e le coordinate di un punto P nel sistema di riferimento introdotto nella Figura 7.2. Sono anche riportati i punti U_x e U_y estremi liberi dei vettori \vec{i} e \vec{j} rispettivamente.

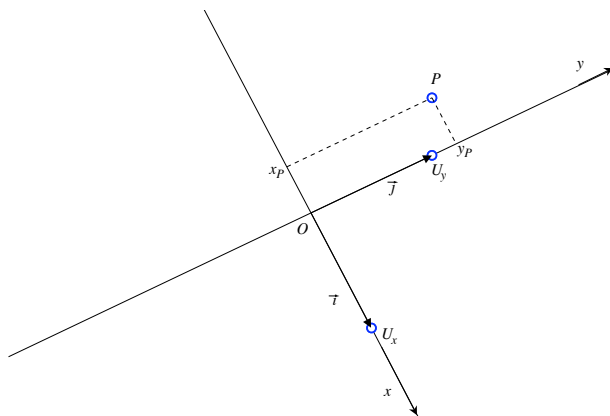


Figura 7.3

⚠ Si noti che la coppia di coordinate dipende dalla scelta del sistema di riferimento $O\vec{i}\vec{j}$ e dalla scelta dell'unità di misura u . Uno stesso punto P può corrispondere a coppie numeriche molto diverse in sistemi di riferimento differenti!

Con abuso di notazione, noi spesso scriveremo $P = (x_P, y_P)$: questa non è un'uguaglianza nel senso usuale (non può esserlo perché l'oggetto a sinistra dell'uguale è

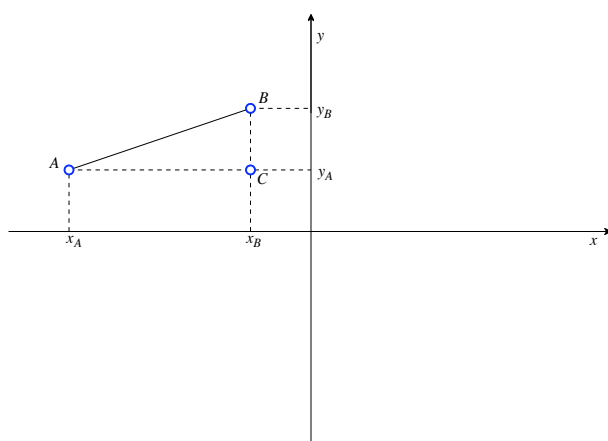


Figura 7.4

un punto del piano e l'oggetto a destra è un elemento del prodotto cartesiano \mathbb{R}^2), ma va letta come *il punto* $P \in S_2$ *che rispetto al sistema di riferimento fissato* $O\vec{i}\vec{j}$ *ha ascissa* x_P *e ordinata* y_P .

Osservazione 7.6. Avere un sistema di coordinate nel piano permette di fare alcuni conti velocemente. Per esempio è possibile calcolare la lunghezza $|\overline{AB}|$ del segmento \overline{AB} di estremi A e B in termini di tali coordinate. Infatti siano $A = (x_A, y_A)$ e $B = (x_B, y_B)$, come nella Figura 7.4.

Applicando il Teorema di Pitagora al triangolo Δ_{ABC} , che è rettangolo in C , si deduce la seguente formula per la lunghezza di un segmento (e quindi per il modulo di un vettore)

$$|\overline{AB}| = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2}.$$

7.3 Sistemi di coordinate nello spazio

Procediamo il nostro studio con il caso dello spazio S_3 .

Definizione 7.7 (Sistemi di riferimento nello spazio S_3).

Un *sistema di riferimento* (cartesiano ortogonale) $O\vec{i}\vec{j}\vec{k}$ in S_3 è definito dai seguenti enti:

- (SRS1) un punto $O \in S_3$ detto *origine* del sistema di riferimento;
- (SRS2) due versori \vec{i} e \vec{j} applicati in O e fra loro perpendicolari;
- (SRS3) un versore \vec{k} applicato in O , perpendicolare al piano contenente \vec{i} e \vec{j} e tale che la terna $\vec{i}, \vec{j}, \vec{k}$ sia orientata come l'indice, il medio e il pollice della mano destra (regola della mano destra).

Per i sistemi di riferimento nello spazio valgono le stesse considerazioni che abbiamo fatto nel caso precedente. Le direzioni di \vec{i} , \vec{j} e \vec{k} vengono dette *asse delle ascisse* (o delle x), *asse delle ordinate* (o delle y) e *asse delle quote* (o delle z) rispettivamente.

In particolare è possibile stabilire una biiezione fra i punti dello spazio e le terne ordinate di numeri reali. Sia infatti $P \in S_3$ un qualsiasi punto del piano (eventualmente appartenente agli assi) e si considerino i piani π_x , π_y e π_z passanti per P e paralleli rispettivamente al piano yz (contenente gli assi delle ordinate e delle quote), xz (contenente gli assi delle ascisse e delle quote) e xy (contenente gli assi delle ascisse e delle ordinate): allora i piani π_x , π_y e π_z intersecheranno gli assi in punti corrispondenti a numeri reali x_P , y_P e z_P (detti rispettivamente *ascissa*, *ordinata* e *quota di P*). Al punto P rimane associata la terna ordinata (x_P, y_P, z_P) .

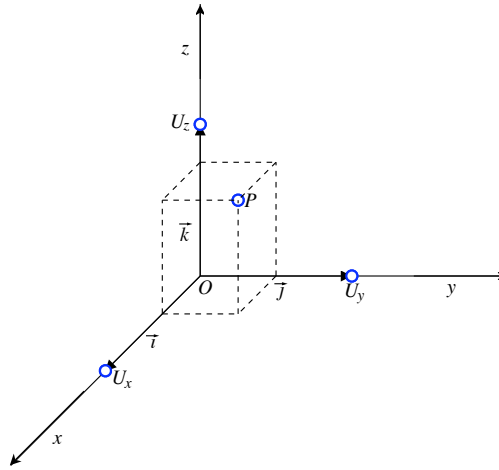


Figura 7.5

Nella Figura 7.5 riportiamo gli assi delle ascisse, delle ordinate e delle quote e le coordinate di un punto P in un sistema di riferimento in S_3 . Sono anche riportati i punti U_x , U_y e U_z estremi liberi dei versori \vec{i} , \vec{j} e \vec{k} rispettivamente.

Il motivo per cui viene adottata la convenzione della regola della mano destra è che se immaginiamo un osservatore che si trova nel semiasse positivo delle quote e guarda in basso, verso il piano delle ascisse e ordinate, la convenzione scelta fa sì che egli veda un piano con un sistema di coordinate come definito nel paragrafo precedente.

Osservazione 7.8. Come visto per segmenti nel piano, dato un segmento $\overline{AB} \subseteq S_3$, se $A = (x_A, y_A, z_A)$, $B = (x_B, y_B, z_B)$, possiamo calcolare la sua lunghezza grazie al Teorema di Pitagora, ottenendo

$$|\overline{AB}| = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2}. \quad (7.3.1)$$

7.4 Prime operazioni sui vettori applicati

D'ora innanzi ci limiteremo a considerare lo spazio ordinario S_3 e vettori applicati in esso, $V_3(O)$. I risultati e le definizioni corrispondenti per vettori nel piano ordinario S_2 (e quindi in $V_2(O)$) sono del tutto analoghi. Per questioni di semplicità i disegni saranno quasi sempre fatti nel piano.

In S_3 fissiamo un sistema di riferimento cartesiano ortogonale $O\vec{i}\vec{j}\vec{k}$; è chiaro che ogni vettore \vec{OP} rimane completamente individuato dal suo estremo libero e quindi da una terna di coordinate (x_P, y_P, z_P) . Spesso identificheremo il vettore \vec{OP} con la matrice colonna

$$\begin{pmatrix} x_P \\ y_P \\ z_P \end{pmatrix}.$$

Si faccia attenzione a non confondere la terna di coordinate di un punto P con la matrice colonna associata al vettore applicato \vec{OP} .

Vogliamo ora introdurre due operazioni fra vettori applicati di uno stesso ambiente (del piano o dello spazio): la somma di vettori applicati e il prodotto di un vettore applicato per uno scalare. Iniziamo con la prima.

Definizione 7.9 (Somma di vettori).

Siano $\vec{v}, \vec{w} \in V_3(O)$ vettori corrispondenti alle matrici

$$\begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}, \quad \begin{pmatrix} w_x \\ w_y \\ w_z \end{pmatrix}$$

rispettivamente. Definiamo *somma di \vec{v} e \vec{w}* il vettore $\vec{u} = \vec{v} + \vec{w}$ corrispondente alla matrice

$$\begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} + \begin{pmatrix} w_x \\ w_y \\ w_z \end{pmatrix} = \begin{pmatrix} v_x + w_x \\ v_y + w_y \\ v_z + w_z \end{pmatrix}.$$

La Definizione 7.9 sembra dipendere dal sistema di riferimento scelto $O\vec{i}\vec{j}\vec{k}$ ma in realtà, ciò è solo apparente. Consideriamo infatti due vettori non nulli qualsiasi \vec{v} e \vec{w} : se tali vettori sono generali non saranno paralleli e si avrà una situazione come quella rappresentata nella Figura 7.6.

Si noti che i triangoli Ov_xA e BFC sono rettangoli in v_x e F rispettivamente, hanno i cateti orizzontali $\overline{Ov_x}$ e \overline{BF} di lunghezze $|v_x|$ e $|(v_x + w_x) - w_x|$ e i cateti verticali $\overline{v_xA}$ e \overline{FC} di lunghezze $|v_y|$ e $|(v_y + w_y) - w_y|$. Concludiamo allora che $|\overline{Ov_x}| = |\overline{BF}|$ e $|\overline{v_xA}| = |\overline{FC}|$, quindi i triangoli Ov_xA e BFC sono congruenti: da ciò deduciamo che i loro angoli interni in O e B sono uguali, perciò i segmenti \overline{OA} e \overline{BC} sono paralleli. Ragionando in maniera simile sui triangoli CEA e Bw_yO deduciamo similmente che anche i segmenti \overline{AC} e \overline{OB} sono paralleli. Concludiamo che il quadrilatero $OACB$ è un parallelogramma.

Abbiamo quindi una via geometrica per calcolare la somma di due vettori applicati in O , detta *regola del parallelogramma*: se si considera il parallelogramma

avente i due lati con vertice O coincidenti con i vettori applicati \vec{v} e \vec{w} , il vettore $\vec{v} + \vec{w}$ viene a coincidere con la diagonale uscente da O . Di conseguenza la definizione non dipende dal sistema di riferimento scelto.

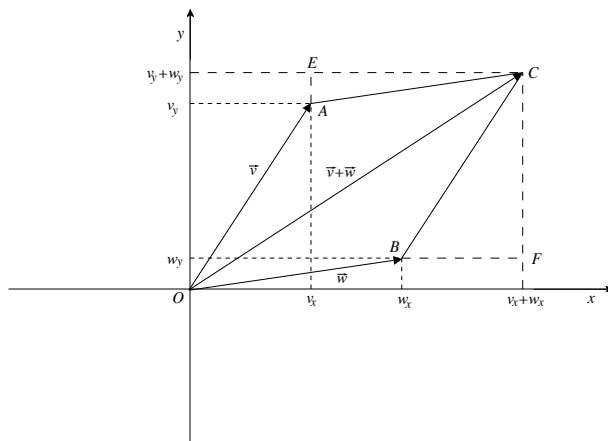


Figura 7.6

Si verifica facilmente che tale osservazione può essere estesa anche al caso di coppie di vettori non generali (per esempio, vettori paralleli o tali che uno dei due sia nullo), quindi la nostra definizione di somma non dipende, in realtà, dal sistema di riferimento scelto ma solo dai vettori.

Introduciamo adesso una seconda operazione sui vettori applicati.

Definizione 7.10 (Prodotto per uno scalare).

Siano $\alpha \in \mathbb{R}$ e $\vec{v} \in V_3(O)$ un vettore corrispondente alla matrice

$$\begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}.$$

Definiamo *prodotto del vettore \vec{v} per lo scalare α* il vettore $\vec{u} = \alpha\vec{v}$ corrispondente alla matrice

$$\alpha \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = \begin{pmatrix} \alpha v_x \\ \alpha v_y \\ \alpha v_z \end{pmatrix}.$$

Notiamo che anche la Definizione 7.10 sembra dipendere dal sistema di riferimento scelto, ma di nuovo tale dipendenza è solo apparente. Consideriamo infatti uno scalare α non nullo, per esempio $\alpha > 1$, e un vettore non nullo qualsiasi \vec{v} : si avrà una situazione come quella rappresentata nella Figura 7.7.

I triangoli Ov_xA e $O\alpha v_xB$ sono rettangoli e hanno i cateti orizzontali $\overline{Ov_x}$ e $\overline{O\alpha v_x}$ di lunghezze $|v_x|$ e $|\alpha v_x|$ e i cateti verticali $\overline{v_xA}$ e $\overline{\alpha v_xB}$ di lunghezze $|v_y|$ e $|\alpha v_y|$. Concludiamo che i triangoli Ov_xA e $O\alpha v_xB$ sono simili. In particolare hanno

gli angoli corrispondenti uguali, quindi le loro ipotenuse sono parallele. Inoltre. I segmenti \overline{OA} e \overline{OB} hanno lunghezze proporzionali secondo il fattore α .

Poiché abbiamo dato una descrizione geometrica di $\alpha\vec{v}$, la nostra definizione di prodotto non dipende dal sistema di riferimento scelto ma solo dal vettore e dallo scalare. Precisamente se o $\vec{v} = \vec{0}$ oppure $\alpha = 0$ allora $\alpha\vec{v} = \vec{0}$; se invece $\vec{v} \neq \vec{0}$ e $\alpha \neq 0$ la direzione di $\alpha\vec{v}$ coincide con la direzione di \vec{v} , $\alpha\vec{v}$ è concorde con \vec{v} se $\alpha > 0$ e discorde se $\alpha < 0$ e $|\alpha\vec{v}| = |\alpha||\vec{v}|$.

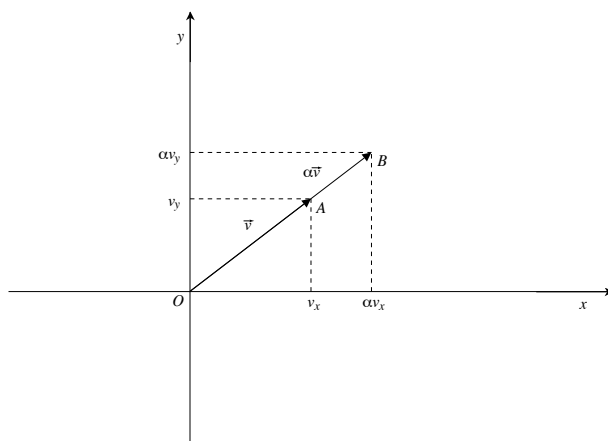


Figura 7.7

Da quanto visto sopra e dalle Proposizioni 1.20 e 1.23 che valgono per le matrici, ricaviamo la seguente lista di proprietà per le operazioni definite sopra, che valgono sia nel piano $V_2(O)$ che nello spazio $V_3(O)$.

Proposizione 7.11 (Proprietà di somma di vettori e prodotto per scalare).

Sia $n = 2, 3$. In $V_n(O)$ valgono le seguenti proprietà:

- (SVG1) *per ogni $\vec{v}, \vec{w} \in V_n(O)$, si ha $\vec{v} + \vec{w} = \vec{w} + \vec{v}$ (proprietà commutativa);*
- (SVG2) *per ogni $\vec{u}, \vec{v}, \vec{w} \in V_n(O)$, si ha $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$ (proprietà associativa);*
- (SVG3) *il vettore nullo è l'unico elemento neutro per la somma, cioè è l'unico vettore tale che $\vec{0} + \vec{v} = \vec{v}$, per ogni $\vec{v} \in V_n(O)$ (esistenza e unicità dell'elemento neutro);*
- (SVG4) *per ogni $\vec{v} \in V_n(O)$, $-\vec{v}$ è l'unico elemento opposto di \vec{v} , cioè è l'unico vettore tale che $\vec{v} + (-\vec{v}) = \vec{0}$ (esistenza e unicità dell'elemento opposto);*
- (PVG1) *per ogni $\vec{v} \in V_n(O)$, si ha $1\vec{v} = \vec{v}$ (esistenza dell'elemento neutro);*
- (PVG2) *per ogni $\alpha_1, \alpha_2 \in \mathbb{R}$ e $\vec{v} \in V_n(O)$, si ha $\alpha_1(\alpha_2\vec{v}) = (\alpha_1\alpha_2)\vec{v}$ (proprietà associativa);*

(SPVG1) per ogni $\alpha_1, \alpha_2 \in \mathbb{R}$ e $\vec{v} \in V_n(O)$, si ha $(\alpha_1 + \alpha_2)\vec{v} = \alpha_1\vec{v} + \alpha_2\vec{v}$ (proprietà distributiva rispetto alla somma di scalari);

(SPVG2) per ogni $\alpha \in \mathbb{R}$ e $\vec{v}, \vec{w} \in V_n(O)$, si ha $\alpha(\vec{v} + \vec{w}) = \alpha\vec{v} + \alpha\vec{w}$ (proprietà distributiva rispetto alla somma di vettori).

Si noti che se $\vec{v} \in V_3(O)$ corrisponde, rispetto al sistema di riferimento fissato $O\vec{i}\vec{j}\vec{k}$, alla matrice $(v_x \ v_y \ v_z)$, in base alla definizione di somma e prodotto per scalari data sopra si ha la seguente proprietà, illustrata nella Figura 7.8.

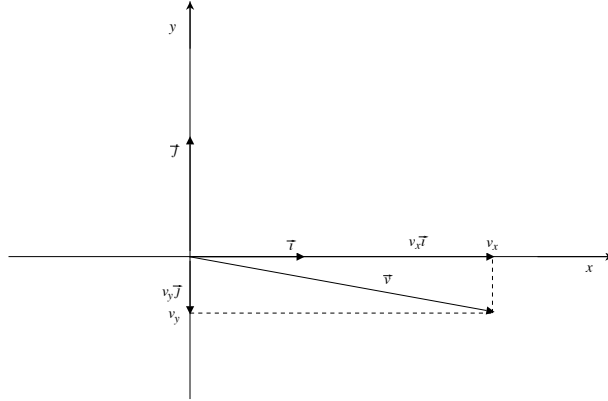


Figura 7.8

Proposizione 7.12.

Sia dato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ in S_3 . Per ogni vettore $\vec{v} \in V_3(O)$ esistono unici $v_x, v_y, v_z \in \mathbb{R}$ tali che

$$\vec{v} = v_x\vec{i} + v_y\vec{j} + v_z\vec{k}. \quad (7.4.1)$$

Nella scrittura (7.4.1) si seguono le usuali convenzioni algebriche, in particolare la decomposizione di un vettore geometrico secondo i tre versori ci permette di lavorare con i vettori trattandoli come dei polinomi lineari in $\vec{i}, \vec{j}, \vec{k}$, e si ha

$$\begin{aligned} (v_x\vec{i} + v_y\vec{j} + v_z\vec{k}) + (w_x\vec{i} + w_y\vec{j} + w_z\vec{k}) &= (v_x + w_x)\vec{i} + (v_y + w_y)\vec{j} + (v_z + w_z)\vec{k}, \\ \alpha(v_x\vec{i} + v_y\vec{j} + v_z\vec{k}) &= (\alpha v_x)\vec{i} + (\alpha v_y)\vec{j} + (\alpha v_z)\vec{k}. \end{aligned}$$

Esempio 7.13. Fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ in S_3 , si considerino i due vettori $\vec{v} = \vec{i} + 3\vec{j} - \vec{k}$ e $\vec{w} = -2\vec{i} - \vec{j} + \vec{k}$ di $V_3(O)$. Allora

$$\vec{v} + \vec{w} = (\vec{i} + 3\vec{j} - \vec{k}) + (-2\vec{i} - \vec{j} + \vec{k}) = (1 - 2)\vec{i} + (3 - 1)\vec{j} + (-1 + 1)\vec{k} = -\vec{i} + 2\vec{j}.$$

Osserviamo che il vettore $\vec{v} + \vec{w}$ si trova nel piano xy , dunque è perpendicolare a \vec{k} . Calcoliamo anche

$$\begin{aligned} \vec{v} - 2\vec{w} &= (\vec{i} + 3\vec{j} - \vec{k}) - 2(-2\vec{i} - \vec{j} + \vec{k}) \\ &= (1 + 4)\vec{i} + (3 + 2)\vec{j} + (-1 - 2)\vec{k} = 5\vec{i} + 5\vec{j} - 3\vec{k}. \end{aligned} \quad \spadesuit$$

Ricordando quanto detto, per le matrici risulta $-\vec{v} = (-1)\vec{v}$. Ne segue una nozione di differenza di vettori applicati: se \vec{v} e $\vec{w} \in V_3(O)$, scriveremo $\vec{v} - \vec{w}$ in luogo di $\vec{v} + (-\vec{w})$. Anche per la differenza $\vec{v} - \vec{w}$ c'è un'interpretazione geometrica; si ricordi che $\vec{v} + \vec{w}$ è la diagonale uscente da O del parallelogramma avente lati \vec{v} e \vec{w} . Allora la differenza $\vec{v} - \vec{w}$ è la diagonale uscente da O del parallelogramma avente lati \vec{v} e $-\vec{w}$, come mostrato in Figura 7.9.

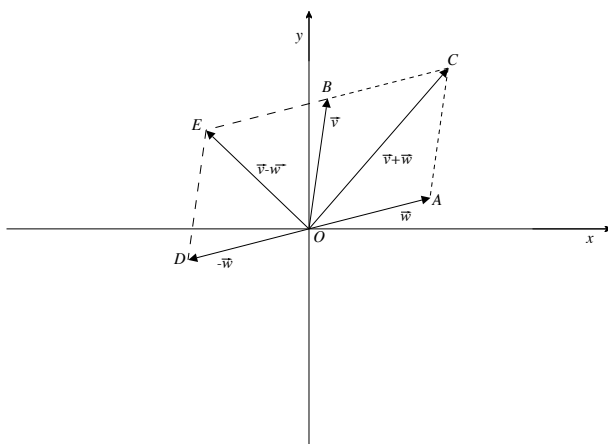


Figura 7.9

Poiché i due parallelogrammi $OACB$ e $OBED$ sono congruenti, deduciamo che $\vec{v} - \vec{w} = \overrightarrow{OB} - \overrightarrow{OA}$ e, come segmento, è parallelo e congruente alla diagonale \overline{AB} orientata da A a B .

Nel seguito scriveremo $B - A$ in luogo di $\overrightarrow{OB} - \overrightarrow{OA}$: si noti che non stiamo facendo una differenza di punti (che non ha senso), ma stiamo semplicemente definendo un nuovo simbolo per indicare un particolare vettore di $V_3(O)$.

Il simbolo $B - A$ è utile per ricordare che se $A = (x_A, y_A, z_A)$ e $B = (x_B, y_B, z_B)$ allora

$$B - A = \overrightarrow{OB} - \overrightarrow{OA} = (x_B - x_A)\vec{i} + (y_B - y_A)\vec{j} + (z_B - z_A)\vec{k}.$$

7.5 Caratterizzazione dei vettori paralleli e complanari

Concludiamo questo capitolo chiarendo il rapporto fra le operazioni di somma di vettori e prodotto di un vettore per uno scalare e le nozioni di parallelismo e complanarità di vettori.

Un'interessante interpretazione algebrica della nozione geometrica di parallelismo è data dal seguente risultato.

Proposizione 7.14.

Siano dati $\vec{v}, \vec{w} \in V_3(O)$, con $\vec{v} \neq \vec{0}$. Allora $\vec{v} \parallel \vec{w}$ se e solo se esiste $\alpha \in \mathbb{R}$ tale che $\vec{w} = \alpha\vec{v}$.

Dimostrazione. Supponiamo $\vec{w} = \alpha\vec{v}$: se $\vec{w} = \vec{0}$ allora per definizione $\vec{w} \parallel \vec{v}$. Se $\vec{w} \neq \vec{0}$, dall'interpretazione geometrica del prodotto per scalari, segue subito che $\vec{w} \parallel \vec{v}$.

Viceversa, supponiamo che $\vec{w} \parallel \vec{v}$. Se $\vec{w} = \vec{0}$ allora $\vec{w} = 0\vec{v}$. Sia $\vec{w} \neq \vec{0}$: se \vec{v}, \vec{w} sono concordi definiamo $\alpha = |\vec{w}|/|\vec{v}|$, se sono discordi $\alpha = -|\vec{w}|/|\vec{v}|$. Allora $\alpha\vec{v}$ è un vettore avente stessa direzione di \vec{v} (e di \vec{w}), verso coincidente con quello di \vec{v} se \vec{v}, \vec{w} sono concordi, opposto se \vec{v}, \vec{w} sono discordi (quindi ha verso coincidente con quello di \vec{w}) e modulo pari a $|\alpha| |\vec{v}| = (|\vec{w}|/|\vec{v}|)|\vec{v}| = |\vec{w}|$. Concludiamo che $\vec{w} = \alpha\vec{v}$. \square

Applichiamo tale proprietà al caso dei versori di un vettore non nullo $\vec{v} \neq \vec{0}$. Sia $\text{vers}(\vec{v})$ il versore a esso parallelo e concorde: allora si deve avere $\text{vers}(\vec{v}) = \alpha\vec{v}$ con $\alpha > 0$: calcolando i moduli si ottiene $1 = |\text{vers}(\vec{v})| = \alpha|\vec{v}|$ da cui si deduce che $\text{vers}(\vec{v}) = \vec{v}/|\vec{v}|$. Ovviamente $-\vec{v} = (-1)\text{vers}(\vec{v})$ è il versore a esso parallelo e discorde.

Proposizione 7.15.

Siano dati $\vec{u}, \vec{v}, \vec{w} \in V_3(O)$, con $\vec{v} \not\parallel \vec{w}$. Allora $\vec{u}, \vec{v}, \vec{w}$ sono complanari se e solo se esistono $\alpha, \beta \in \mathbb{R}$ tali che $\vec{u} = \alpha\vec{v} + \beta\vec{w}$.

Dimostrazione. Supponiamo sia $\vec{u} = \alpha\vec{v} + \beta\vec{w}$: in base all'interpretazione geometrica della somma abbiamo che $\vec{u}, \vec{v}, \vec{w}$ giacciono in uno stesso piano.

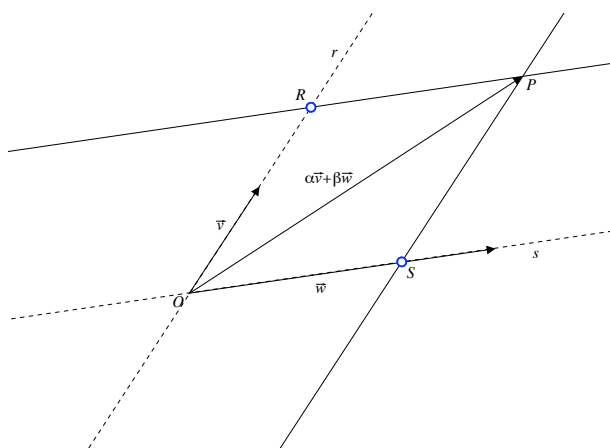


Figura 7.10

Viceversa supponiamo che $\vec{u}, \vec{v}, \vec{w}$ giacciono in uno stesso piano: supponiamo che P sia l'estremo libero di \vec{u} . Le rette per P parallele alle direzioni r ed s di \vec{v} e \vec{w} rispettivamente, intersecano s ed r rispettivamente in due punti S e R tali che il quadrilatero $OSPR$ è un parallelogramma.

Per costruzione $\vec{OS} \parallel \vec{w}$ e $\vec{OR} \parallel \vec{v}$, quindi esistono $\alpha, \beta \in \mathbb{R}$ tali che $\vec{OS} = \beta\vec{w}$ e $\vec{OR} = \alpha\vec{v}$: allora il significato geometrico di somma di vettori applicati ci permette di scrivere $\vec{OP} = \alpha\vec{v} + \beta\vec{w}$. \square

Dalle Proposizioni 7.14 e 7.15 segue il seguente risultato.

Proposizione 7.16.

In S_3 sia fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$. Si considerino A, B, C, D quattro punti in S_3 , di coordinate $A = (x_A, y_A, z_A)$, $B = (x_B, y_B, z_B)$, $C = (x_C, y_C, z_C)$, $D = (x_D, y_D, z_D)$.

i) I punti A, B e C sono allineati se e solo se

$$\text{rk} \begin{pmatrix} x_B - x_A & y_B - y_A & z_B - z_A \\ x_C - x_A & y_C - y_A & z_C - z_A \end{pmatrix} \leq 1;$$

ii) i punti A, B, C e D sono complanari se e solo se

$$\text{rk} \begin{pmatrix} x_B - x_A & y_B - y_A & z_B - z_A \\ x_C - x_A & y_C - y_A & z_C - z_A \\ x_D - x_A & y_D - y_A & z_D - z_A \end{pmatrix} \leq 2.$$

Dimostrazione.

i) Se $A = B = C$ non c'è nulla da verificare: supponiamo perciò $B - A \neq \vec{0}$. I vettori $B - A$ e $C - A$ sono paralleli se e solo se i punti A, B, C sono allineati. Per la Proposizione 7.14 questo avviene se e solo se esiste $t \in \mathbb{R}$ tale che

$$(x_C - x_A)\vec{i} + (y_C - y_A)\vec{j} + (z_C - z_A)\vec{k} = t((x_B - x_A)\vec{i} + (y_B - y_A)\vec{j} + (z_B - z_A)\vec{k}),$$

cioè se e solo se con l'operazione elementare $R_2 \rightarrow R_2 - tR_1$ la matrice sopra si trasforma in

$$\begin{pmatrix} x_B - x_A & y_B - y_A & z_B - z_A \\ 0 & 0 & 0 \end{pmatrix},$$

cioè se e solo se ha rango al più 1.

ii) Se i punti A, B, C, D sono allineati non c'è nulla da verificare: supponiamo perciò $B - A \nparallel C - A$. I vettori $B - A, C - A$ e $D - A$ sono complanari se e solo se i punti A, B, C, D lo sono. Per la Proposizione 7.15 questo accade se e solo se esistono $t, u \in \mathbb{R}$ tale che

$$\begin{aligned} (x_D - x_A)\vec{i} + (y_D - y_A)\vec{j} + (z_D - z_A)\vec{k} \\ = t((x_B - x_A)\vec{i} + (y_B - y_A)\vec{j} + (z_B - z_A)\vec{k}) + u((x_C - x_A)\vec{i} + (y_C - y_A)\vec{j} + (z_C - z_A)\vec{k}), \end{aligned}$$

cioè se e solo se con l'operazione elementare $R_3 \rightarrow R_3 - tR_1 - uR_2$ la matrice sopra si trasforma in

$$\begin{pmatrix} x_B - x_A & y_B - y_A & z_B - z_A \\ x_C - x_A & y_C - y_A & z_C - z_A \\ 0 & 0 & 0 \end{pmatrix},$$

cioè se e solo se ha rango al più 2. □

Operazioni fra vettori applicati

8.1 Prodotto scalare

Definizione 8.1 (Prodotto scalare).

Sia fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ in S_3 . Dati $\vec{v}, \vec{w} \in V_3(O)$, $\vec{v} = v_x\vec{i} + v_y\vec{j} + v_z\vec{k}$ e $\vec{w} = w_x\vec{i} + w_y\vec{j} + w_z\vec{k}$, il *prodotto scalare* di \vec{v} e \vec{w} è il numero

$$\langle \vec{v}, \vec{w} \rangle = v_x w_x + v_y w_y + v_z w_z.$$

Confrontando la definizione di prodotto scalare con la Formula (7.3.1) otteniamo

$$|\vec{v}| = \sqrt{\langle \vec{v}, \vec{v} \rangle}.$$

⚠ Testi diversi utilizzano notazioni diverse per denotare il prodotto scalare di due vettori, non solo $\langle \vec{v}, \vec{w} \rangle$, ma anche $\vec{v} \cdot \vec{w}$ e (\vec{v}, \vec{w}) .

Esempio 8.2. Si considerino i vettori $\vec{i} = 1\vec{i} + 0\vec{j} + 0\vec{k}$, $\vec{j} = 0\vec{i} + 1\vec{j} + 0\vec{k}$, $\vec{k} = 0\vec{i} + 0\vec{j} + 1\vec{k}$. Si ha

$$\begin{aligned} \langle \vec{i}, \vec{i} \rangle &= \langle \vec{j}, \vec{j} \rangle = \langle \vec{k}, \vec{k} \rangle = 1, \\ \langle \vec{i}, \vec{j} \rangle &= \langle \vec{j}, \vec{i} \rangle = \langle \vec{i}, \vec{k} \rangle = \langle \vec{k}, \vec{i} \rangle = \langle \vec{j}, \vec{k} \rangle = \langle \vec{k}, \vec{j} \rangle = 0. \end{aligned}$$

Si considerino poi i due vettori $\vec{v} = \vec{i} + \vec{j} + \vec{k}$ e $\vec{w} = 3\vec{i} - \vec{j} - \vec{k}$. Si ha

$$\langle \vec{v}, \vec{w} \rangle = 1 \cdot 3 + 1 \cdot (-1) + 1 \cdot (-1) = 1. \quad \spadesuit$$

Dalle proprietà del prodotto tra matrici seguono immediatamente una serie di proprietà del prodotto scalare.

Proposizione 8.3 (Proprietà del prodotto scalare).

Valgono le seguenti proprietà:

(PSVG1) per ogni $\vec{v}, \vec{w} \in V_n(O)$, si ha $\langle \vec{v}, \vec{w} \rangle = \langle \vec{w}, \vec{v} \rangle$ (proprietà commutativa);

(PSVG2) per ogni $\vec{u}, \vec{v}, \vec{w} \in V_n(O)$, si ha $\langle \vec{u}, \vec{v} + \vec{w} \rangle = \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{w} \rangle$ (proprietà distributiva rispetto alla somma di vettori);

(PSVG3) per ogni $\alpha \in \mathbb{R}$ e $\vec{v}, \vec{w} \in V_n(O)$, si ha $\alpha \langle \vec{v}, \vec{w} \rangle = \langle \alpha \vec{v}, \vec{w} \rangle$;

(PSVG4) per ogni $\vec{v} \in V_n(O) \setminus \{\vec{0}\}$, si ha $\langle \vec{v}, \vec{v} \rangle > 0$ (il prodotto scalare è definito positivo).

Le proprietà (PSVG2) e (PSVG3) vengono solitamente dette *proprietà di bilinearità del prodotto scalare*.

Si noti che le componenti di un vettore rispetto a un sistema di riferimento si descrivono facilmente in termini di prodotti scalari. Infatti se $\vec{v} = v_x \vec{i} + v_y \vec{j} + v_z \vec{k}$ si definiscono i *coefficienti di Fourier di \vec{v} rispetto al sistema di riferimento $O\vec{i}\vec{j}\vec{k}$* come

$$v_x = \langle \vec{v}, \vec{i} \rangle, \quad v_y = \langle \vec{v}, \vec{j} \rangle, \quad v_z = \langle \vec{v}, \vec{k} \rangle,$$

da cui otteniamo la decomposizione

$$\vec{v} = \langle \vec{v}, \vec{i} \rangle \vec{i} + \langle \vec{v}, \vec{j} \rangle \vec{j} + \langle \vec{v}, \vec{k} \rangle \vec{k}.$$

Questo ci fornisce un metodo relativamente semplice per determinare la decomposizione di un vettore rispetto a un nuovo sistema di riferimento.

La seguente interpretazione geometrica mostra che il prodotto scalare è indipendente dal sistema di riferimento scelto e dipende solo dai vettori coinvolti. A tale scopo introduciamo la definizione di angolo fra vettori, illustrato qui sotto nella Figura 8.1.

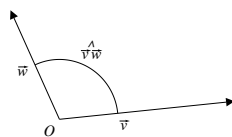


Figura 8.1

Definizione 8.4 (Angolo tra due vettori).

Sia fissato un sistema di riferimento in S_3 . Siano $\vec{v}, \vec{w} \in V_3(O)$ non nulli:

- se $\vec{v} \not\parallel \vec{w}$ definiamo *angolo tra \vec{v} e \vec{w}* la misura in radianti $\widehat{v\vec{w}} \in]0, \pi[$ dell'angolo del piano individuato dai vettori \vec{v} e \vec{w} e avente per lati le semirette contenenti i due vettori \vec{v} e \vec{w} ;
- se $\vec{v} \parallel \vec{w}$ sono concordi definiamo $\widehat{v\vec{w}} = 0$;
- se $\vec{v} \parallel \vec{w}$ sono discordi definiamo $\widehat{v\vec{w}} = \pi$.

Con tale nozione possiamo enunciare la seguente proposizione.

Proposizione 8.5.

Sia fissato un sistema di riferimento in S_3 . Dati $\vec{v}, \vec{w} \in V_3(O)$, o almeno uno dei due vettori è nullo e si ha $\langle \vec{v}, \vec{w} \rangle = 0$, oppure sono entrambi non nulli e si ha $\langle \vec{v}, \vec{w} \rangle = |\vec{v}| |\vec{w}| \cos(\widehat{v\vec{w}})$.

Dimostrazione. Siano $\vec{v} = v_x\vec{i} + v_y\vec{j} + v_z\vec{k}$ e $\vec{w} = w_x\vec{i} + w_y\vec{j} + w_z\vec{k}$. Se almeno uno dei due vettori, ad esempio \vec{v} , è nullo, allora applicando la definizione di prodotto scalare

$$\langle \vec{v}, \vec{w} \rangle = 0 \cdot w_x + 0 \cdot w_y + 0 \cdot w_z = 0.$$

Supponiamo quindi che i vettori siano entrambi non nulli; se sono anche paralleli, in base alla definizione, è facile verificare la tesi. Supponiamo allora che i vettori $\vec{v}, \vec{w} \in V_3(O)$ non siano paralleli, come illustrato nella Figura 8.2.

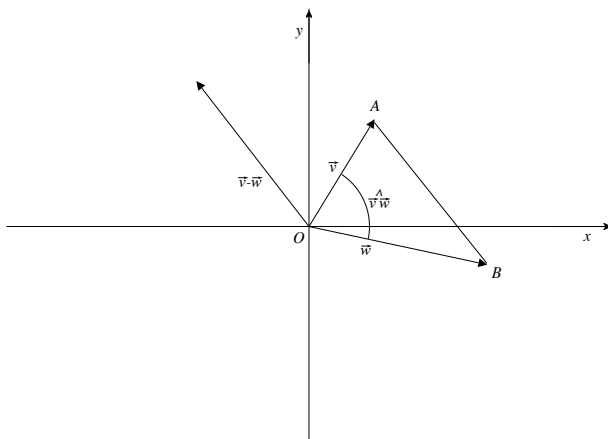


Figura 8.2

Per il Teorema di Carnot, ricordando il significato geometrico di differenza di vettori applicati, si ha che $|\vec{v} - \vec{w}|^2 = |\vec{v}|^2 + |\vec{w}|^2 - 2|\vec{v}||\vec{w}|\cos(\widehat{v\vec{w}})$, quindi

$$|\vec{v}||\vec{w}|\cos(\widehat{v\vec{w}}) = \frac{1}{2}(|\vec{v}|^2 + |\vec{w}|^2 - |\vec{v} - \vec{w}|^2).$$

Ricordando la formula (7.3.1),

$$\begin{aligned} |\vec{v}|^2 &= x_v^2 + y_v^2 + z_v^2, & |\vec{w}|^2 &= x_w^2 + y_w^2 + z_w^2, \\ |\vec{v} - \vec{w}|^2 &= (x_v - x_w)^2 + (y_v - y_w)^2 + (z_v - z_w)^2. \end{aligned}$$

Sostituendo queste espressioni nella formula sopra otteniamo $|\vec{v}||\vec{w}|\cos(\widehat{v\vec{w}}) = \langle \vec{v}, \vec{w} \rangle$. \square

Dalla Proposizione 8.5 segue immediatamente che $\langle \vec{i}, \vec{i} \rangle = \langle \vec{j}, \vec{j} \rangle = \langle \vec{k}, \vec{k} \rangle = 1$, $\langle \vec{i}, \vec{j} \rangle = \langle \vec{j}, \vec{i} \rangle = \langle \vec{i}, \vec{k} \rangle = \langle \vec{k}, \vec{i} \rangle = \langle \vec{j}, \vec{k} \rangle = \langle \vec{k}, \vec{j} \rangle = 0$ come già verificato nell'Esempio 8.2.

Facciamo ora alcune considerazioni; una prima osservazione importantissima è che il prodotto $\langle \vec{v}, \vec{w} \rangle$ è nullo se e solo se o almeno uno dei due vettori è nullo oppure se $\cos(\widehat{v\vec{w}}) = 0$, cioè se $\widehat{v\vec{w}} = \pi/2$.

Definizione 8.6 (Vettori perpendicolari).

Siano dati $\vec{v}, \vec{w} \in V_3(O)$. I vettori \vec{v}, \vec{w} si dicono *perpendicolari* (o *ortogonali*) se $\langle \vec{v}, \vec{w} \rangle = 0$. Se \vec{v} è perpendicolare a \vec{w} scriviamo $\vec{v} \perp \vec{w}$.

Osservazione 8.7. Se $\vec{v}, \vec{w} \in V_3(O)$ sono entrambi non nulli,

$$\cos(\widehat{\vec{v}\vec{w}}) = \frac{\langle \vec{v}, \vec{w} \rangle}{|\vec{v}||\vec{w}|},$$

quindi il segno del prodotto scalare $\langle \vec{v}, \vec{w} \rangle$ è esattamente il segno di $\cos(\widehat{\vec{v}\vec{w}})$: in particolare $\langle \vec{v}, \vec{w} \rangle > 0$ se e solo se $\widehat{\vec{v}\vec{w}} \in [0, \pi/2[$, cioè se e solo se \vec{v} e \vec{w} formano un angolo acuto, mentre $\langle \vec{v}, \vec{w} \rangle < 0$ se e solo se $\widehat{\vec{v}\vec{w}} \in]\pi/2, \pi]$, cioè se e solo se \vec{v} e \vec{w} formano un angolo ottuso.

Esempio 8.8. Si considerino i vettori \vec{v} e \vec{w} dell'Esempio 8.2. Poiché $\langle \vec{v}, \vec{w} \rangle = 1$ i due vettori formano un angolo acuto: precisamente essendo $|\vec{v}| = \sqrt{3}$, $|\vec{w}| = \sqrt{11}$ si ha $\cos(\widehat{\vec{v}\vec{w}}) = 1/\sqrt{33}$.

Invece i vettori $\vec{v}' = 3\vec{i} + \vec{j} - \vec{k}$ e $\vec{w}' = \vec{i} - \vec{j} + 2\vec{k}$ sono perpendicolari: infatti $\langle \vec{v}', \vec{w}' \rangle = 3 \cdot 1 + 1 \cdot (-1) + (-1) \cdot 2 = 0$. ♠

Osservazione 8.9 (Disuguaglianze: Cauchy–Schwartz e triangolare). Dal momento che la funzione coseno è limitata in modulo da 1, come corollario immediato della Proposizione 8.5 osserviamo che, dati i vettori $\vec{v}, \vec{w} \in V_3(O)$, vale la cosiddetta *disuguaglianza di Cauchy–Schwartz*

$$|\langle \vec{v}, \vec{w} \rangle| \leq |\vec{v}||\vec{w}|,$$

in cui si ha l'uguaglianza se e solo se i vettori \vec{v} e \vec{w} sono paralleli.

Inoltre, applicando le proprietà (PSVG1) e (PSVG2) della Proposizione 8.3 si ha che

$$\begin{aligned} |\vec{v} + \vec{w}|^2 &= \langle \vec{v} + \vec{w}, \vec{v} + \vec{w} \rangle = \langle \vec{v}, \vec{v} + \vec{w} \rangle + \langle \vec{w}, \vec{v} + \vec{w} \rangle \\ &= \langle \vec{v}, \vec{v} \rangle + \langle \vec{v}, \vec{w} \rangle + \langle \vec{w}, \vec{v} \rangle + \langle \vec{w}, \vec{w} \rangle = |\vec{v}|^2 + 2\langle \vec{v}, \vec{w} \rangle + |\vec{w}|^2 \\ &\leq |\vec{v}|^2 + 2|\langle \vec{v}, \vec{w} \rangle| + |\vec{w}|^2 \leq |\vec{v}|^2 + 2|\vec{v}||\vec{w}| + |\vec{w}|^2 = (|\vec{v}| + |\vec{w}|)^2. \end{aligned}$$

Poiché $|\vec{v} + \vec{w}|$ e $|\vec{v}| + |\vec{w}|$ sono quantità non negative, la catena di disuguaglianze sopra dimostra la *disuguaglianza triangolare*

$$|\vec{v} + \vec{w}| \leq |\vec{v}| + |\vec{w}|,$$

in cui vale l'uguaglianza se e solo se i vettori \vec{v} e \vec{w} sono paralleli e concordi oppure almeno uno di loro è nullo.

Una conseguenza diretta della disuguaglianza triangolare è che per ogni \vec{v} e \vec{w} vale anche

$$||\vec{v}| - |\vec{w}|| \leq |\vec{v} - \vec{w}|,$$

che equivale a dire che in un triangolo ogni lato è maggiore della differenza degli altri due.

Osservazione 8.10 (Proiezione ortogonale). Il prodotto scalare è legato anche a un'altra nozione geometrica: la proiezione ortogonale. Osservando la Figura 8.3 è facile vedere che, dati due vettori $\vec{v}, \vec{w} \in V_3(O)$, la quantità $\langle \vec{v}, \vec{w} \rangle$ coincide con il prodotto della lunghezza di \vec{v} per la lunghezza della proiezione di \vec{w} lungo la direzione di \vec{v} (o viceversa).

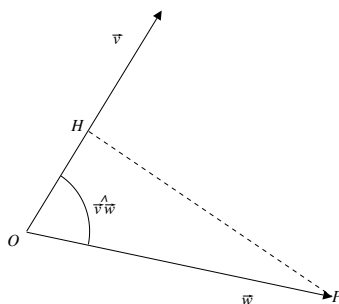


Figura 8.3

In particolare se si desidera determinare il vettore $\vec{w}_{\parallel} = \overrightarrow{OH}$ proiezione di \vec{w} lungo la direzione di \vec{v} è sufficiente applicare la formula

$$\begin{aligned} \vec{w}_{\parallel} &= |\vec{w}| \cos(\widehat{v\vec{w}}) \text{vers}(\vec{v}) = \frac{|\vec{w}||\vec{v}| \cos(\widehat{v\vec{w}})}{|\vec{v}|} \frac{\vec{v}}{|\vec{v}|} \\ &= \frac{\langle \vec{v}, \vec{w} \rangle}{|\vec{v}|} \frac{\vec{v}}{|\vec{v}|} = \frac{\langle \vec{v}, \vec{w} \rangle}{|\vec{v}|^2} \vec{v} = \frac{\langle \vec{v}, \vec{w} \rangle}{\langle \vec{v}, \vec{v} \rangle} \vec{v}. \end{aligned}$$

Il vettore \vec{w}_{\parallel} è detto *proiezione ortogonale di \vec{w} lungo la direzione di \vec{v}* .

Si noti che il triangolo Δ_{OHP} è rettangolo in H , dunque, ricordando il significato geometrico di differenza di vettori, il vettore

$$\vec{w}_{\perp} = \vec{w} - \vec{w}_{\parallel} = \vec{w} - \frac{\langle \vec{v}, \vec{w} \rangle}{\langle \vec{v}, \vec{v} \rangle} \vec{v}$$

è perpendicolare a \vec{w}_{\parallel} : in particolare questo ci dà un metodo per decomporre un vettore lungo due direzioni perpendicolari di cui una fissata, poiché $\vec{w} = \vec{w}_{\parallel} + \vec{w}_{\perp}$.

Esempio 8.11. Si considerino i vettori $\vec{v} = \vec{i} + \vec{j} - \vec{k}$ e $\vec{w} = \vec{i} - 2\vec{j} + 3\vec{k}$; vogliamo determinare la proiezione \vec{w}_{\parallel} di \vec{w} lungo la direzione di \vec{v} . Si ha

$$\vec{w}_{\parallel} = \frac{\langle \vec{i} - 2\vec{j} + 3\vec{k}, \vec{i} + \vec{j} - \vec{k} \rangle}{\langle \vec{i} + \vec{j} - \vec{k}, \vec{i} + \vec{j} - \vec{k} \rangle} (\vec{i} + \vec{j} - \vec{k}) = -\frac{4}{3} (\vec{i} + \vec{j} - \vec{k}).$$

Possiamo decomporre $\vec{w} = \vec{w}_{\parallel} + \vec{w}_{\perp}$, dove

$$\vec{w}_{\perp} = \vec{w} - \vec{w}_{\parallel} = \frac{1}{3} (7\vec{i} - 2\vec{j} + 5\vec{k}) \perp \vec{w}_{\parallel}. \quad \spadesuit$$

8.2 Prodotto vettoriale

Definizione 8.12 (Prodotto vettoriale).

Sia fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ in S_3 . Dati $\vec{v}, \vec{w} \in V_3(O)$, $\vec{v} = v_x\vec{i} + v_y\vec{j} + v_z\vec{k}$ e $\vec{w} = w_x\vec{i} + w_y\vec{j} + w_z\vec{k}$, il *prodotto vettoriale* di \vec{v} e \vec{w} è il vettore di $V_3(O)$

$$\vec{v} \times \vec{w} = (v_y w_z - v_z w_y) \vec{i} - (v_x w_z - v_z w_x) \vec{j} + (v_x w_y - v_y w_x) \vec{k}. \quad (8.2.1)$$

La formula (8.2.1) sopra è un po' difficile da ricordare. Si noti che i coefficienti di $\vec{v} \times \vec{w}$ sono esattamente i determinanti delle sottomatrici 2×2 della matrice

$$\begin{pmatrix} v_x & v_y & v_z \\ w_x & w_y & w_z \end{pmatrix},$$

con segni alternati!

Un artificio utile per ricordarsi la formula (8.2.1) può, quindi, essere quello di utilizzare proprio la nozione di determinante scrivendo

$$\vec{v} \times \vec{w} = \begin{vmatrix} v_y & v_z \\ w_y & w_z \end{vmatrix} \vec{i} - \begin{vmatrix} v_x & v_z \\ w_x & w_z \end{vmatrix} \vec{j} + \begin{vmatrix} v_x & v_y \\ w_x & w_y \end{vmatrix} \vec{k}.$$

Spesso, ricordando la definizione di determinante di una matrice 3×3 , si utilizza anche la notazione

$$\vec{v} \times \vec{w} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{vmatrix}.$$

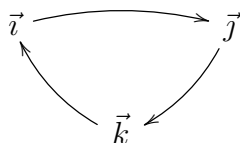
⚠ Testi diversi utilizzano notazioni diverse per denotare il prodotto vettoriale, non solo $\vec{v} \times \vec{w}$, ma anche $\vec{v} \wedge \vec{w}$ e, più raramente, $[\vec{v}, \vec{w}]$.

Esempio 8.13. Consideriamo i vettori dell'Esempio 8.2. Poiché $\vec{i} = 1\vec{i} + 0\vec{j} + 0\vec{k}$, $\vec{j} = 0\vec{i} + 1\vec{j} + 0\vec{k}$, $\vec{k} = 0\vec{i} + 0\vec{j} + 1\vec{k}$ si ha

$$\begin{aligned} \vec{i} \times \vec{i} &= \begin{vmatrix} 0 & 0 \\ 0 & 0 \end{vmatrix} \vec{i} - \begin{vmatrix} 1 & 0 \\ 1 & 0 \end{vmatrix} \vec{j} + \begin{vmatrix} 0 & 1 \\ 0 & 1 \end{vmatrix} \vec{k} = \vec{0}, \\ \vec{i} \times \vec{j} &= \begin{vmatrix} 0 & 0 \\ 1 & 0 \end{vmatrix} \vec{i} - \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} \vec{j} + \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \vec{k} = \vec{k}, \\ \vec{i} \times \vec{k} &= \begin{vmatrix} 0 & 0 \\ 0 & 1 \end{vmatrix} \vec{i} - \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \vec{j} + \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} \vec{k} = -\vec{j}. \end{aligned}$$

Il lettore verifichi in modo analogo che $\vec{j} \times \vec{j} = \vec{k} \times \vec{k} = \vec{0}$, $\vec{j} \times \vec{k} = \vec{i}$, $\vec{k} \times \vec{j} = -\vec{i}$, $\vec{j} \times \vec{i} = -\vec{k}$, $\vec{k} \times \vec{i} = \vec{j}$.

Un buon metodo per ricordarsi le relazioni appena viste è quello di considerare la figura



dove procedendo “in senso orario” si trovano relazioni col segno +, mentre procedendo in senso anti-orario si trovano relazioni col segno -.

Si considerino poi i due vettori $\vec{v} = \vec{i} + \vec{j} + \vec{k}$ e $\vec{w} = 3\vec{i} - \vec{j} - \vec{k}$. Allora

$$\vec{v} \times \vec{w} = \begin{vmatrix} 1 & 1 \\ -1 & -1 \end{vmatrix} \vec{i} - \begin{vmatrix} 1 & 1 \\ 3 & -1 \end{vmatrix} \vec{j} + \begin{vmatrix} 1 & 1 \\ 3 & -1 \end{vmatrix} \vec{k} = 0\vec{i} - 4\vec{j} - 4\vec{k}.$$

Si verifichi per esercizio che $\vec{w} \times \vec{v} = 4\vec{j} + 4\vec{k} = -\vec{v} \times \vec{w}$. ♠

Una conseguenza quasi immediata della definizione algebrica di prodotto vettoriale è la seguente Proposizione.

Proposizione 8.14 (Proprietà del prodotto vettoriale).

Valgono le seguenti proprietà:

- (PV1) per ogni $\vec{v}, \vec{w} \in V_3(O)$, si ha $\vec{v} \times \vec{w} = -\vec{w} \times \vec{v}$ (proprietà anticommutativa);
- (PV2) per ogni $\vec{u}, \vec{v}, \vec{w} \in V_3(O)$, si ha $(\vec{u} + \vec{v}) \times \vec{w} = \vec{u} \times \vec{w} + \vec{v} \times \vec{w}$ (proprietà distributiva a destra rispetto alla somma);
- (PV3) per ogni $\vec{u}, \vec{v}, \vec{w} \in V_3(O)$, si ha $\vec{u} \times (\vec{v} + \vec{w}) = \vec{u} \times \vec{v} + \vec{u} \times \vec{w}$ (proprietà distributiva a sinistra rispetto alla somma);
- (PV4) per ogni $\alpha \in \mathbb{R}$ e $\vec{v}, \vec{w} \in V_3(O)$, si ha $\alpha(\vec{v} \times \vec{w}) = (\alpha\vec{v}) \times \vec{w} = \vec{v} \times (\alpha\vec{w})$.

☠ Il prodotto vettoriale non è associativo: per esempio

$$(\vec{i} \times \vec{i}) \times \vec{j} = \vec{0} \neq -\vec{j} = \vec{i} \times \vec{k} = \vec{i} \times (\vec{i} \times \vec{j}).$$

Le proprietà della Proposizione 8.14 e la tabella di moltiplicazione dei vettori \vec{i} , \vec{j} , \vec{k} determinata nell'Esempio 8.13 ci permettono di calcolare il prodotto vettoriale di due vettori anche senza ricordarne la definizione.

Esempio 8.15. Si considerino i vettori $\vec{v} = 2\vec{i} + \vec{j} - 3\vec{k}$ e $\vec{w} = \vec{i} + \vec{j} + \vec{k}$. Allora

$$\begin{aligned} \vec{v} \times \vec{w} &= (2\vec{i} + \vec{j} - 3\vec{k}) \times (\vec{i} + \vec{j} + \vec{k}) \\ &= 2\vec{i} \times \vec{i} + 2\vec{i} \times \vec{j} + 2\vec{i} \times \vec{k} + \vec{j} \times \vec{i} + \vec{j} \times \vec{j} + \vec{j} \times \vec{k} - 3\vec{k} \times \vec{i} - 3\vec{k} \times \vec{j} - 3\vec{k} \times \vec{k}. \end{aligned}$$

Ricordando le relazioni tra \vec{i} , \vec{j} e \vec{k} , segue che

$$\vec{v} \times \vec{w} = 2\vec{k} - 2\vec{j} - \vec{k} + \vec{i} - 3\vec{j} + 3\vec{i} = 4\vec{i} - 5\vec{j} + \vec{k}. \quad \spadesuit$$

Come per il prodotto scalare, abbiamo definito la nozione di prodotto vettoriale utilizzando le componenti dei vettori in termini di coordinate rispetto a un certo sistema di riferimento. In realtà il prodotto vettoriale, proprio come quello scalare, è indipendente dal sistema di riferimento scelto e dipende solo dai vettori coinvolti; per dimostrare questa affermazione daremo la sua interpretazione geometrica.

Proposizione 8.16.

Sia fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ in S_3 . Dati i vettori $\vec{v}, \vec{w} \in V_3(O)$, o $\vec{v} \parallel \vec{w}$, e si ha $\vec{v} \times \vec{w} = \vec{0}$, oppure $\vec{v} \not\parallel \vec{w}$ e $\vec{v} \times \vec{w}$ è definito come segue:

- (i) la sua direzione è perpendicolare al piano contenente i due vettori \vec{v} e \vec{w} ;
- (ii) il suo verso è tale che la terna ordinata $(\vec{v}, \vec{w}, \vec{v} \times \vec{w})$ sia orientata secondo la regola della mano destra;
- (iii) $|\vec{v} \times \vec{w}| = |\vec{v}||\vec{w}|\sin(\widehat{\vec{v}\vec{w}})$.

Dimostrazione. Siano $\vec{v} = v_x\vec{i} + v_y\vec{j} + v_z\vec{k}$ e $\vec{w} = w_x\vec{i} + w_y\vec{j} + w_z\vec{k}$. Se $\vec{v} \parallel \vec{w}$ allora in base alla definizione è facile verificare la tesi. Supponiamo allora che i vettori siano entrambi non nulli e iniziamo a dimostrare (iii). Si ha che

$$|\vec{v} \times \vec{w}|^2 = v_y^2 w_z^2 - 2v_y w_z v_z w_y + v_z^2 w_y^2 + v_z^2 w_x^2 - 2v_z^2 w_x^2 v_x w_z + v_y^2 w_z^2 + v_x^2 w_z^2 + v_x^2 w_y^2 - 2v_x w_y v_y w_x + v_y^2 w_x^2.$$

Tenendo conto della definizione geometrica del prodotto scalare, si ha

$$\begin{aligned} |\vec{v}|^2 |\vec{w}|^2 \sin^2(\widehat{\vec{v}\vec{w}}) &= |\vec{v}|^2 |\vec{w}|^2 - |\vec{v}|^2 |\vec{w}|^2 \cos^2(\widehat{\vec{v}\vec{w}}) = |\vec{v}|^2 |\vec{w}|^2 - \langle \vec{v}, \vec{w} \rangle^2 \\ &= v_y^2 w_z^2 - 2v_y w_z v_z w_y + v_z^2 w_y^2 + v_z^2 w_x^2 - 2v_z^2 w_x^2 v_x w_z \\ &\quad + v_y^2 w_z^2 + v_x^2 w_z^2 + v_x^2 w_y^2 - 2v_x w_y v_y w_x + v_y^2 w_x^2. \end{aligned}$$

Poiché i quadrati delle quantità non negative $|\vec{v} \times \vec{w}|$ e $|\vec{v}||\vec{w}|\sin(\widehat{\vec{v}\vec{w}})$ coincidono, segue (iii).

Dimostriamo adesso (ii), cioè che $\vec{v} \times \vec{w} \perp \vec{v}, \vec{w}$. A tale scopo è sufficiente verificare che $\langle \vec{v}, \vec{v} \times \vec{w} \rangle = \langle \vec{w}, \vec{v} \times \vec{w} \rangle = 0$. Risulta, ad esempio,

$$\langle \vec{v}, \vec{v} \times \vec{w} \rangle = \langle \vec{v}, \begin{vmatrix} v_y & v_z \\ w_y & w_z \end{vmatrix} \vec{i} - \begin{vmatrix} v_x & v_z \\ w_x & w_z \end{vmatrix} \vec{j} + \begin{vmatrix} v_x & v_y \\ w_x & w_y \end{vmatrix} \vec{k} \rangle = v_x \begin{vmatrix} v_y & v_z \\ w_y & w_z \end{vmatrix} - v_y \begin{vmatrix} v_x & v_z \\ w_x & w_z \end{vmatrix} + v_z \begin{vmatrix} v_x & v_y \\ w_x & w_y \end{vmatrix} = 0.$$

Si verifichi in modo analogo che $\langle \vec{w}, \vec{v} \times \vec{w} \rangle = 0$.

Omettiamo la dimostrazione di (i) in quanto coinvolge la nozione di “matrice di passaggio fra due basi” che sarà introdotta solo più avanti nel corso. \square

La Proposizione 8.16 è bene illustrata dalla Figura 8.4.

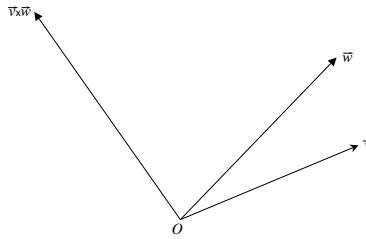


Figura 8.4

Alla luce della Proposizione 8.16 discende immediatamente che

$$\begin{aligned} \vec{i} \times \vec{j} = \vec{k} = -\vec{j} \times \vec{i}, & \quad \vec{j} \times \vec{k} = \vec{i} = -\vec{k} \times \vec{j}, & \quad \vec{k} \times \vec{i} = \vec{j} = -\vec{i} \times \vec{k}, \\ \vec{i} \times \vec{i} = \vec{j} \times \vec{j} = \vec{k} \times \vec{k} = \vec{0}, & & \end{aligned}$$

come già verificato nell'Esempio 8.13.

Come nel caso del prodotto scalare anche l'annullarsi del prodotto vettoriale $\vec{v} \times \vec{w}$ dà informazioni sulla posizione relativa dei due vettori \vec{v} e \vec{w} : infatti $\vec{v} \times \vec{w} = \vec{0}$ se e solo se $|\vec{v} \times \vec{w}| = 0$ se e solo se o almeno uno dei due vettori è nullo oppure se $\sin(\widehat{\vec{v}\vec{w}}) = 0$, cioè se $\widehat{\vec{v}\vec{w}} = 0, \pi$, cioè se e solo se $\vec{v} \parallel \vec{w}$.

⚠ Utilizzare questo metodo per verificare il parallelismo di vettori è però senza dubbio molto più oneroso che applicare la Proposizione 7.16.

Osservazione 8.17 (Prodotto vettoriale e area). Il prodotto vettoriale, o meglio il suo modulo, è legato alla nozione di area. Si considerino tre punti non allineati $A, B, C \in S_3$ e si consideri il triangolo Δ_{ABC} avente tali punti come vertici, come nella Figura 8.5.

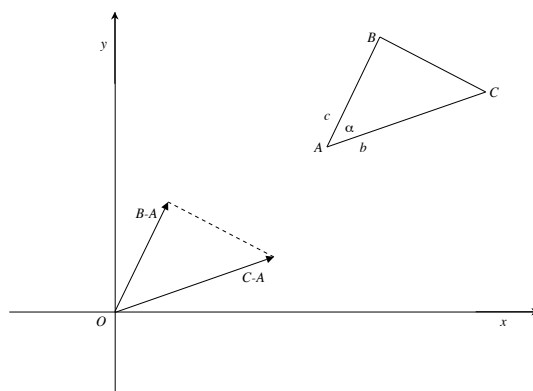


Figura 8.5

Allora è noto dalla trigonometria elementare che la sua area è

$$\text{Area}(ABC) = \frac{1}{2}(bc \sin \alpha),$$

ove b e c sono le lunghezze dei lati opposti ai vertici B e C rispettivamente e α è l'angolo interno con vertice A .

D'altra parte il triangolo Δ_{ABC} è congruente al triangolo avente lati $B - A$ e $C - A$, quindi $b = |B - A|$, $c = |C - A|$ e $\alpha = \widehat{(B - A)(C - A)}$. Concludiamo allora che

$$\begin{aligned} \text{Area}(\Delta_{ABC}) &= \frac{1}{2}bc \sin \alpha = \frac{1}{2}|B - A||C - A| \sin \widehat{(B - A)(C - A)} \\ &= \frac{1}{2}|(B - A) \times (C - A)|. \end{aligned} \quad (8.2.2)$$

Per esempio se consideriamo un triangolo nel piano di vertici A, B, C , con coordinate $A = (x_A, y_A)$, $B = (x_B, y_B)$, $C = (x_C, y_C)$ otteniamo

$$\begin{aligned} (B - A) \times (C - A) &= \begin{vmatrix} x_B - x_A & y_B - y_A \\ x_C - x_A & y_C - y_A \end{vmatrix} \vec{k} \\ &= ((x_B - x_A)(y_C - y_A) - (x_C - x_A)(y_B - y_A)) \vec{k}, \end{aligned}$$

da cui la formula, nota ad alcuni,

$$\text{Area}(\Delta_{ABC}) = \frac{1}{2} \left| \begin{vmatrix} x_B - x_A & y_B - y_A \\ x_C - x_A & y_C - y_A \end{vmatrix} \right| = \frac{1}{2} \left| \begin{vmatrix} x_A & y_A & 1 \\ x_B & y_B & 1 \\ x_C & y_C & 1 \end{vmatrix} \right|. \quad (8.2.3)$$

⚠ La formula (8.2.3) è valida solo nel caso di triangoli nel piano! Per triangoli nello spazio si deve applicare la formula (8.2.2) per ottenere il valore corretto dell'area.

Esempio 8.18. Si considerino i punti $A = (1, 1, 1)$, $B = (2, -1, 3)$, $C = (-1, 0, 1)$. Allora $B - A = \vec{i} - 2\vec{j} + 2\vec{k}$, $C - A = -2\vec{i} - \vec{j}$: tali vettori non sono proporzionali, quindi i tre punti non sono allineati, quindi definiscono un triangolo. Poiché

$$(B - A) \times (C - A) = \begin{vmatrix} -2 & 2 \\ -1 & 0 \end{vmatrix} \vec{i} - \begin{vmatrix} 1 & 2 \\ -2 & 0 \end{vmatrix} \vec{j} + \begin{vmatrix} 1 & -2 \\ -2 & -1 \end{vmatrix} \vec{k} = 2\vec{i} - 4\vec{j} - 5\vec{k}$$

si ha

$$\text{Area}(\Delta_{ABC}) = \frac{1}{2} |2\vec{i} - 4\vec{j} - 5\vec{k}| = \frac{1}{2} \sqrt{4 + 16 + 25} = \frac{3}{2} \sqrt{5}.$$

Si noti, a conferma del fatto che la formula (8.2.3) vale solo nel piano, che

$$\frac{1}{2} \left| \begin{vmatrix} 1 & 1 & 1 \\ 2 & -1 & 3 \\ 1 & 0 & 1 \end{vmatrix} \right| = \frac{1}{2} \neq \text{Area}(\Delta_{ABC}). \quad \spadesuit$$

8.3 Prodotto misto

Definizione 8.19 (Prodotto misto).

Sia fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ in S_3 . Dati $\vec{u}, \vec{v}, \vec{w} \in V_3(O)$, il *prodotto misto* di \vec{u} , \vec{v} e \vec{w} è il numero

$$\langle \vec{u}, \vec{v} \times \vec{w} \rangle.$$

Se $\vec{u} = u_x\vec{i} + u_y\vec{j} + u_z\vec{k}$, $\vec{v} = v_x\vec{i} + v_y\vec{j} + v_z\vec{k}$ e $\vec{w} = w_x\vec{i} + w_y\vec{j} + w_z\vec{k}$, tenendo conto delle formule che permettono di calcolare il prodotto scalare e il prodotto vettoriale, otteniamo

$$\langle \vec{u}, \vec{v} \times \vec{w} \rangle = u_x \begin{vmatrix} v_y & v_z \\ w_y & w_z \end{vmatrix} - u_y \begin{vmatrix} v_x & v_z \\ w_x & w_z \end{vmatrix} + u_z \begin{vmatrix} v_x & v_y \\ w_x & w_y \end{vmatrix} = \begin{vmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{vmatrix}.$$

Osserviamo che se $\vec{v} \parallel \vec{w}$ allora $\langle \vec{u}, \vec{v} \times \vec{w} \rangle = 0$: in tal caso \vec{u} , \vec{v} , \vec{w} sono ovviamente complanari.

Se invece $\vec{v} \not\parallel \vec{w}$, allora $\vec{v} \times \vec{w} \neq \vec{0}$ e ricordando la Proposizione 8.16 deduciamo che $\langle \vec{u}, \vec{v} \times \vec{w} \rangle = 0$ se e solo se \vec{u} è perpendicolare a $\vec{v} \times \vec{w}$, cioè se e solo se \vec{u} , \vec{v} e \vec{w} sono complanari.

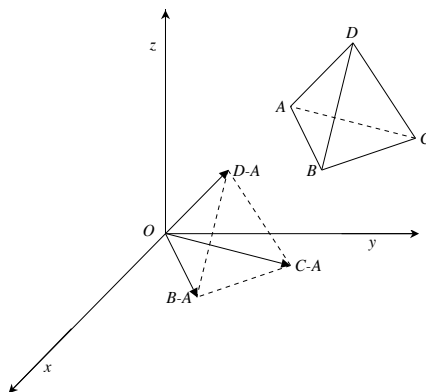


Figura 8.6

Osservazione 8.20 (Prodotto misto e volume). Il prodotto misto, o meglio il suo modulo, è legato alla nozione di volume. Si considerino quattro punti non complanari $A, B, C, D \in S_3$ e consideriamo il tetraedro T_{ABCD} avente tali punti come vertici, come nella Figura 8.6.

Allora è noto dalla geometria elementare che il suo volume è

$$\text{Volume}(T_{ABCD}) = \frac{1}{3} \text{Area}(\Delta_{ABC})h,$$

ove h è l'altezza relativa al vertice D (si veda la Figura 8.6).

D'altra parte sappiamo che

$$\text{Area}(\Delta_{ABC}) = \frac{1}{2} |(B - A) \times (C - A)|,$$

e h non è altro che la lunghezza della proiezione del vettore $D - A$ lungo la direzione perpendicolare al triangolo Δ_{ABC} , che è la direzione del vettore $(B - A) \times (C - A)$: quindi $h = |D - A| |\cos \varphi|$ ove φ è l'angolo formato da $D - A$ e $(B - A) \times (C - A)$. In particolare

$$\begin{aligned} \text{Volume}(ABCD) &= \frac{1}{3} \frac{1}{2} |D - A| |(B - A) \times (C - A)| |\cos \varphi| = \\ &= \frac{1}{6} |\langle D - A, (B - A) \times (C - A) \rangle|. \end{aligned}$$

Esempio 8.21. Si considerino i punti $A = (1, 1, 1)$, $B = (2, -1, 3)$, $C = (-1, 0, 1)$, $D = (3, 3, 3)$. Allora $B - A = \vec{i} - 2\vec{j} + 2\vec{k}$, $C - A = -2\vec{i} - \vec{j}$, $D - A = 2\vec{i} + 2\vec{j} + 2\vec{k}$: poiché

$$\begin{pmatrix} 1 & -2 & 2 \\ -2 & -1 & 0 \\ 2 & 2 & 2 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 - R_1} \begin{pmatrix} 1 & -2 & 2 \\ -2 & -1 & 0 \\ 1 & 4 & 0 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 + 4R_2} \begin{pmatrix} 1 & -2 & 2 \\ -2 & -1 & 0 \\ -7 & 0 & 0 \end{pmatrix},$$

segue che i quattro punti non sono complanari, dunque definiscono un tetraedro e si ha

$$\begin{aligned}\text{Volume}(T_{ABCD}) &= \frac{1}{6} |\langle D - A, (B - A) \times (C - A) \rangle| = \\ &= \frac{1}{6} \left| \begin{vmatrix} 2 & 2 & 2 \\ 1 & -2 & 2 \\ -2 & -1 & 0 \end{vmatrix} \right| = \frac{1}{6} \left| \begin{vmatrix} 1 & -2 & 2 \\ -2 & -1 & 0 \\ -7 & 0 & 0 \end{vmatrix} \right| = \frac{1}{6} |-14| = \frac{7}{3}. \quad \spadesuit\end{aligned}$$

Equazioni parametriche di rette e piani nello spazio

9.1 Equazioni parametriche di rette

In questo capitolo daremo una descrizione parametrica di oggetti geometrici quali rette (nel piano e nello spazio) e piani nello spazio. Nel prossimo capitolo ci occuperemo di definire gli stessi oggetti tramite equazioni lineari, ovvero della loro descrizione cartesiana

Cominciamo a caratterizzare rette nello spazio. Sia $r \subseteq S_3$ una tale retta: osserviamo che r è sempre parallela a un'unica retta passante per l'origine r' e rimane completamente individuata da essa e da un punto qualsiasi $R \in r$. Ciò significa che dare r' equivale a dare un qualsiasi vettore non nullo \vec{v} avente r' come direzione, come si può vedere nella Figura 9.1. In altre parole, una retta nello spazio è completamente individuata dalla sua direzione $\vec{v} \neq \vec{0}$ e da un suo punto qualsiasi R .

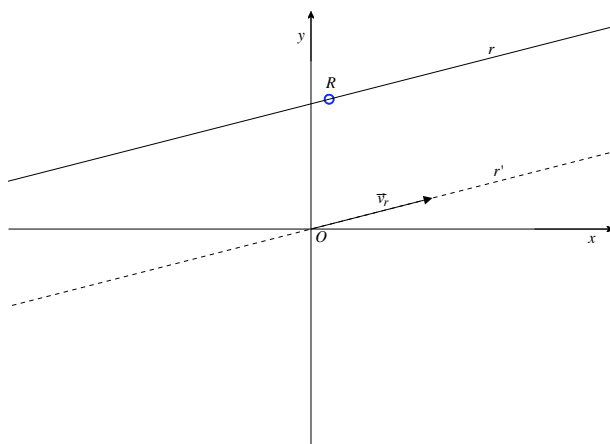


Figura 9.1

Cerchiamo un modo per caratterizzare i punti della retta r : per definizione, per ogni punto $P \in r$ vale che

$$\overrightarrow{OP} - \overrightarrow{OR} = P - R;$$

sommando ad ambo i membri \overrightarrow{OR} segue che

$$\overrightarrow{OP} = \overrightarrow{OR} + (P - R).$$

Poiché $P - R$ è parallelo al segmento \overline{PR} , dunque a r , ha direzione r' ; quindi esso è parallelo al vettore \vec{v} menzionato sopra, come mostrato nella Figura 9.2.

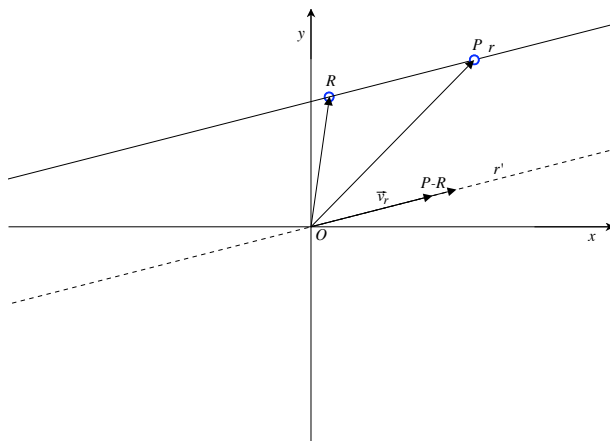


Figura 9.2

Dalla caratterizzazione di vettori paralleli mostrata nella Proposizione 7.14 segue che $P - R \parallel \vec{v}$ se e solo se esiste un numero reale, che chiameremo t , tale che $P - R = t\vec{v}$. Segue quindi che un punto dello spazio $P \in S_3$ giace sulla retta r se e solo se

$$\overrightarrow{OP} = \overrightarrow{OR} + t\vec{v}, \quad (9.1.1)$$

per qualche valore di $t \in \mathbb{R}$.

Fissiamo un sistema di riferimento cartesiano ortogonale $O\vec{i}\vec{j}\vec{k}$ in S_3 . Supponiamo che in tale sistema il punto R abbia coordinate $R = (x_0, y_0, z_0)$, quindi

$$\overrightarrow{OR} = x_0\vec{i} + y_0\vec{j} + z_0\vec{k},$$

e $\vec{v} = l\vec{i} + m\vec{j} + n\vec{k}$: indicando con (x, y, z) le coordinate del punto generico $P \in S_3$ si ha $\overrightarrow{OP} = x\vec{i} + y\vec{j} + z\vec{k}$, dunque l'equazione (9.1.1) diventa

$$x\vec{i} + y\vec{j} + z\vec{k} = x_0\vec{i} + y_0\vec{j} + z_0\vec{k} + t(l\vec{i} + m\vec{j} + n\vec{k}),$$

per qualche valore di $t \in \mathbb{R}$, oppure, eguagliando le componenti dei due vettori lungo gli assi coordinati,

$$\begin{cases} x = x_0 + lt \\ y = y_0 + mt \\ z = z_0 + nt, \end{cases} \quad t \in \mathbb{R}. \quad (9.1.2)$$

Le equazioni (9.1.2) sono dette *equazioni parametriche della retta r passante per il punto $R = (x_0, y_0, z_0)$ e parallela al vettore $\vec{v} = l\vec{i} + m\vec{j} + n\vec{k} \neq \vec{0}$* .

Viceversa, supponiamo di avere fissato nello spazio S_3 un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$. Dati numeri reali fissati x_0, y_0, z_0, l, m, n , si consideri il luogo r dei punti $P = (x, y, z)$ dello spazio le cui coordinate sono della forma

$$\begin{cases} x = x_0 + lt \\ y = y_0 + mt \\ z = z_0 + nt, \end{cases} \quad t \in \mathbb{R}.$$

Allora, preso $t = 0$, segue che il punto $R = (x_0, y_0, z_0) \in r$. Se poi l, m, n non sono tutti nulli, esistono in R infiniti altri punti P tali che

$$x\vec{i} + y\vec{j} + z\vec{k} = x_0\vec{i} + y_0\vec{j} + z_0\vec{k} + t(l\vec{i} + m\vec{j} + n\vec{k}),$$

ovvero tali che $\overrightarrow{OP} = \overrightarrow{OR} + t\vec{v}$ ove $\vec{v} = l\vec{i} + m\vec{j} + n\vec{k}$. Tali punti descrivono quindi la retta passante per il punto R sopra definito e parallela al vettore non nullo \vec{v} .

Concludiamo che, fissato in S_3 un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$, ogni retta può essere descritta mediante un sistema di equazioni della forma (9.1.2) con l, m, n non tutti nulli e, viceversa, ogni sistema di equazioni della forma (9.1.2) con l, m, n non tutti nulli rappresenta una retta nello spazio.

Esempio 9.1. In S_3 sia fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$, $A = (1, 2, 3) \in S_3$ e $\vec{v} = 2\vec{i} - 3\vec{k} \in V_3(O)$. Delle equazioni parametriche della retta r di S_3 parallela al vettore \vec{v} e passante per A sono date da

$$\begin{cases} x = 1 + 2t \\ y = 2 \\ z = 3 - 3t, \end{cases} \quad t \in \mathbb{R}. \quad (9.1.3)$$

Ci chiediamo quale fra i punti $B = (3, 2, 0)$ e $C = (-1, 2, 1)$ di S_3 appartenga alla retta r . Per rispondere bisogna capire se esistono valori di $t \in \mathbb{R}$ per cui le coordinate di B e C possono essere scritte nella forma data dall'equazione (9.1.2), ovvero se e quale fra i due sistemi

$$\begin{cases} 3 = 1 + 2t \\ 2 = 2 \\ 0 = 3 - 3t, \end{cases} \quad \begin{cases} -1 = 1 + 2t \\ 2 = 2 \\ 1 = 3 - 3t \end{cases}$$

abbia soluzione, per qualche $t \in \mathbb{R}$.

Consideriamo il primo dei due sistemi. Dalla prima equazione si ricava $t = 1$, valore che sostituito nelle equazioni seguenti le soddisfa identicamente, quindi $B \in r$.

Similmente, dalla prima equazione del secondo sistema si ricava $t = -1$, valore che sostituito nella terza equazione dà l'identità numerica $1 = 6$ che, ovviamente, non è verificata: concludiamo che $C \notin r$. ♠

⚠ Le equazioni parametriche di una retta non sono uniche! Infatti esse dipendono sia dalla scelta di un suo punto, che di un vettore non nullo a essa parallelo. Di punti su una retta ne esistono infiniti, così come sono infiniti i vettori a essa paralleli. Questa infinità di possibili scelte ci permette di rappresentare una stessa retta con sistemi di equazioni parametriche anche molto diversi: per tale motivo non si dovrebbe mai scrivere “le equazioni parametriche di r sono...”, bensì “delle equazioni parametriche di r sono...”.

Esempio 9.2. In S_3 sia fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ e si considerino la retta r dell'Esempio 9.1 e la retta r' di equazioni parametriche

$$\begin{cases} x = 2 - 2t \\ y = 0 \\ z = 3t, \end{cases} \quad t \in \mathbb{R}.$$

Si ha che $r \parallel r'$: infatti r è parallela al vettore $\vec{v} = 2\vec{i} - 3\vec{k}$, mentre r' è parallela a $\vec{w} = -2\vec{i} + 3\vec{k}$, che sono paralleli fra loro. ♠

Esempio 9.3. In S_3 sia fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ e si consideri la retta s di equazioni parametriche

$$\begin{cases} x = 3 - 4t \\ y = 2 \\ z = 6t, \end{cases} \quad t \in \mathbb{R}.$$

Tale retta passa per il punto $B = (3, 2, 0)$ ed è parallela al vettore $\vec{w} = -4\vec{i} + 6\vec{k}$.

Ricordando l'Esempio 9.1 segue che s ha in comune con la retta r di equazioni (9.1.3) il punto B ed è a essa parallela, perché $\vec{w} = -4\vec{i} + 6\vec{k} = -2(2\vec{i} - 3\vec{k}) = -2\vec{v}$. Due rette parallele e incidenti devono coincidere, cioè $s = r$: questa uguaglianza non è immediatamente deducibile dall'analisi dei sistemi di equazioni che definiscono r ed s . ♠

⚠ Più in generale, dal confronto delle equazioni parametriche di due rette, si può dedurre la loro *posizione relativa*. Vedremo in dettaglio come fare nel paragrafo 10.4.

Ricordiamo che due rette $r, s \subseteq S_3$ possono essere *complanari*, cioè contenute in uno stesso piano, e in particolare

- *coincidenti*,
- *incidenti* in un unico punto,
- *parallele distinte*;


oppure due rette possono essere *non complanari* (dette anche *sghembe*), cioè non essere nè parallele, nè incidenti.

Osservazione 9.4. Le equazioni (9.1.2) della retta r possono essere pensate come leggi orarie del moto di un punto P lungo la retta r con posizione iniziale $R = (x_0, y_0, z_0)$ e velocità costante $\vec{v} = l\vec{i} + m\vec{j} + n\vec{k}$. Questo punto di vista può essere molto utile nell'affrontare problemi di incidenza fra rette date tramite equazioni parametriche, come nell'esempio seguente.

Esempio 9.5. Si consideri la retta r dell'Esempio 9.1 e la retta s di equazioni parametriche

$$\begin{cases} x = t - 1 \\ y = 2 - t \\ z = 6 + t, \end{cases} \quad t \in \mathbb{R}. \quad (9.1.4)$$

La retta s è parallela al vettore $\vec{w} = \vec{i} - \vec{j} + \vec{k}$; poiché invece r è parallela a $\vec{v} = 2\vec{i} - 3\vec{j}$, deduciamo che $r \nparallel s$. Ci domandiamo allora se r ed s siano incidenti oppure sghembe.

 Un primo approccio (sbagliato) che può venire in mente è il seguente: conosciamo le coordinate del punto generico su r e su s in funzione di un parametro, quindi basta eguagliare tali coordinate e vedere se il sistema così ottenuto ha soluzione o no: se sì allora $r \cap s \neq \emptyset$, se no $r \cap s = \emptyset$. Bisogna fare però attenzione a come si traduce praticamente tale approccio. Infatti se semplicemente eguagliamo le equazioni (9.1.3) alle equazioni (9.1.4) otteniamo

$$\begin{cases} 1 + 2t = t - 1 \\ 2 = 2 - t \\ 3 - 3t = 6 + t, \end{cases}$$

che, come è facile verificare, non ha soluzione, dunque sembrerebbe che $r \cap s = \emptyset$.

Invece si noti che $(-1, 2, 6) \in r \cap s$: infatti si ottiene per $t = -1$ dalle equazioni (9.1.3) e per $t = 0$ dalle equazioni (9.1.4). Cosa abbiamo sbagliato?

L'errore sta nel fatto che in realtà non ci siamo domandati se i punti in moto sulle due rette r ed s potranno mai passare per uno stesso punto, bensì se ciò accade esattamente nello stesso istante!

Il modo per non sbagliare è quello di misurare i tempi in modo diverso sulle due rette utilizzando, ad esempio, il parametro t su r e il parametro t' su s : in questo modo il problema si traduce nel sistema

$$\begin{cases} 1 + 2t = t' - 1 \\ 2 = 2 - t' \\ 3 - 3t = 6 + t'. \end{cases}$$

Dalla seconda equazione otteniamo $t' = 0$: sostituendo nelle rimanenti otteniamo $t = -1$. Come visto sopra i valori $t' = 0$ su s e $t = -1$ su r danno lo stesso punto $(-1, 2, 6)$.

Se invece consideriamo la retta u di equazioni parametriche

$$\begin{cases} x = t \\ y = 2 - t \\ z = 6 + t, \end{cases} \quad t \in \mathbb{R}.$$

ancora $r \not\parallel u$ e, inoltre, è facile verificare che $r \cap u = \emptyset$, poiché il sistema

$$\begin{cases} 1 + 2t = t' \\ 2 = 2 - t' \\ 3 - 3t = 6 + t', \end{cases} \quad t' \in \mathbb{R}.$$

non è compatibile: concludiamo che r e u sono sghembe. ♠

Per concludere questa trattazione delle equazioni parametriche di rette nello spazio, ricordiamo che è noto dalla geometria euclidea che per descrivere una retta r è sufficiente individuare due punti distinti A e B che le appartengono; in tal caso è facile ricondursi al caso precedente di un punto e un vettore parallelo. Infatti un punto, per esempio A , lo abbiamo; per costruire un vettore parallelo a r basta considerare $B - A$, come si vede nella Figura 9.3.

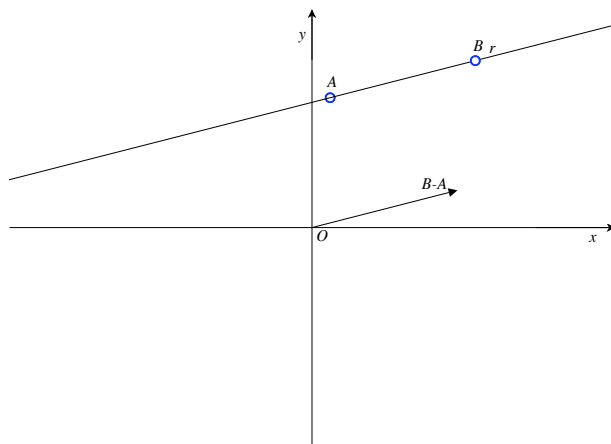


Figura 9.3

Se fissiamo un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ in S_3 in cui i due punti hanno coordinate $A = (x_A, y_A, z_A)$ e $B = (x_B, y_B, z_B)$, allora sostituendo nell'equazione (9.1.2) otteniamo le *equazioni parametriche della retta r passante per i due punti A e B* :

$$\begin{cases} x = x_A + (x_B - x_A)t \\ y = y_A + (y_B - y_A)t \\ z = z_A + (z_B - z_A)t, \end{cases} \quad t \in \mathbb{R},$$

che, talvolta, si scrivono sinteticamente anche come $P = (1 - t)A + tB$.

Si noti che anche in S_3 un punto $P = (x, y)$ appartiene ad \overline{AB} se e solo se

$$\begin{cases} x = (1-t)x_A + tx_B \\ y = (1-t)y_A + ty_B \\ z = (1-t)z_A + tz_B, \end{cases} \quad t \in [0, 1].$$

Ad esempio il punto medio M del segmento \overline{AB} ha coordinate corrispondenti a $t = 1/2$, cioè

$$M = \left(\frac{x_A + x_B}{2}, \frac{y_A + y_B}{2}, \frac{z_A + z_B}{2} \right).$$

Esempio 9.6. Fissiamo un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ in S_3 . Siano $A = (1, 2, -3)$, $B = (2, 1, 1)$: chiaramente $A \neq B$, quindi esiste unica una retta r contenente A e B le cui equazioni parametriche sono

$$\begin{cases} x = 1 + t \\ y = 2 - t \\ z = -3 + 4t, \end{cases} \quad t \in \mathbb{R}. \quad \spadesuit$$

Osservazione 9.7 (Equazioni parametriche di rette nel piano). Spesso si identifica S_2 col piano di S_3 di equazione $z = 0$. In particolare si può pensare ai punti di S_2 come punti di S_3 con quota nulla e ai vettori di $V_2(O)$ come a vettori di $V_3(O)$ con componente nulla nella direzione di \vec{k} .

Con questa identificazione ogni retta $r \subset S_2$, pensata come retta in S_3 ha delle equazioni parametriche della forma

$$\begin{cases} x = x_0 + lt \\ y = y_0 + mt, \\ z = 0 \end{cases} \quad t \in \mathbb{R}.$$

Nel momento in cui si affronta un problema esclusivamente in S_2 , spesso si omette l'equazione banale $z = 0$. Pertanto si scrive talvolta solo

$$\begin{cases} x = x_0 + lt \\ y = y_0 + mt, \end{cases} \quad t \in \mathbb{R}. \quad (9.1.5)$$

Le equazioni (9.1.5) sono dette *equazioni parametriche della retta r passante per il punto $R = (x_0, y_0)$ e parallela al vettore $\vec{v} = l\vec{i} + m\vec{j}$.*

Ad esempio, la retta $r \subset S_2$ parallela al vettore $\vec{v} = (2, 4)$ e passante per il punto $R = (-1, -1)$ ha equazioni parametriche

$$\begin{cases} x = -1 + 2t \\ y = -1 + 4t, \end{cases} \quad t \in \mathbb{R}.$$

Il lettore verifichi che r è anche descritta dalle equazioni

$$\begin{cases} x = 1 + t' \\ y = 3 + 2t', \end{cases} \quad t' \in \mathbb{R}.$$

Anche in questo caso possiamo descrivere una retta $r \subset S_2$ a partire da due punti distinti A e B che le appartengono. Se fissiamo un sistema di riferimento $O\vec{i}\vec{j}$ in S_2 nel quale $A = (x_A, y_A)$ e $B = (x_B, y_B)$, otteniamo le *equazioni parametriche della retta r passante per i due punti $A = (x_A, y_A)$ e $B = (x_B, y_B)$* :

$$\begin{cases} x = x_A + (x_B - x_A)t \\ y = y_A + (y_B - y_A)t, \end{cases} \quad t \in \mathbb{R}.$$

Per passare dalle equazioni precedenti all'usuale rappresentazione delle rette nel piano mediante una singola equazione è sufficiente osservare che $P = (x, y)$ è un punto di r corrispondente al valore $t \in \mathbb{R}$ del parametro se e solo se

$$t = \frac{x - x_A}{x_B - x_A} = \frac{y - y_A}{y_B - y_A},$$

cioè se e solo se le coordinate (x, y) soddisfano l'equazione

$$\frac{x - x_A}{x_B - x_A} = \frac{y - y_A}{y_B - y_A}.$$

Un discorso simile si può ovviamente fare a partire dalle equazioni (9.1.5).

9.2 Equazioni parametriche di piani

Sia $\alpha \subseteq S_3$ un piano nello spazio euclideo: esso è sempre parallelo a un'unico piano passante per l'origine α' e rimane dunque completamente individuato da esso e da un punto qualsiasi $A \in \alpha$, come si può vedere nella Figura 9.4. Per descrivere α è quindi necessario descrivere α' .

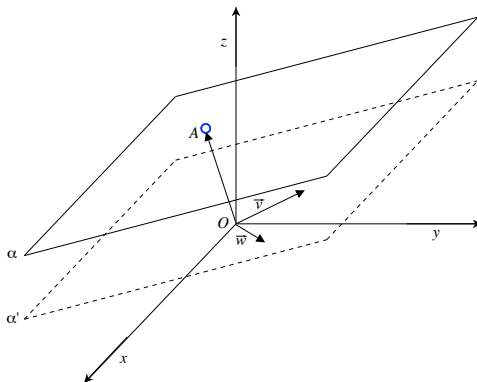


Figura 9.4

Siano \vec{v} e \vec{w} due vettori contenuti in α' e non paralleli: dalla Proposizione 7.15, che caratterizza i vettori complanari, segue che un punto P' appartiene ad α' se e solo se esistono $t, u \in \mathbb{R}$ tali che $\overrightarrow{OP'} = t\vec{v} + u\vec{w}$.

Sia ora $P \in \alpha$ un punto qualsiasi. Per definizione si ha

$$\overrightarrow{OP} - \overrightarrow{OA} = P - A,$$

quindi $\overrightarrow{OP} = \overrightarrow{OA} + (P - A)$. Poiché $P - A$ è parallelo al segmento \overline{PA} , dunque a α , esso è contenuto in α' , quindi esistono, per quanto osservato sopra, dei valori di $t, u \in \mathbb{R}$ tali che $P - A = t\vec{v} + u\vec{w}$. In conclusione un punto $P \in S_3$ giace sul piano α se e solo se

$$\overrightarrow{OP} = \overrightarrow{OA} + t\vec{v} + u\vec{w} \quad (9.2.1)$$

per un qualche $t, u \in \mathbb{R}$, come si può osservare nella Figura 9.5.

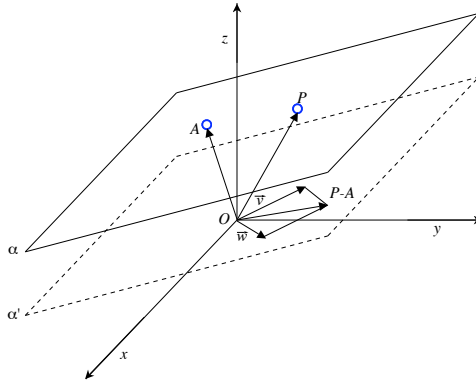


Figura 9.5

Fissiamo ora un sistema di riferimento cartesiano ortogonale $O\vec{i}\vec{j}\vec{k}$ in S_3 , nel quale $A = (x_A, y_A, z_A)$, sicché $\overrightarrow{OA} = x_A\vec{i} + y_A\vec{j} + z_A\vec{k}$, $\vec{v} = v_x\vec{i} + v_y\vec{j} + v_z\vec{k}$, e $\vec{w} = w_x\vec{i} + w_y\vec{j} + w_z\vec{k}$. Se indichiamo con (x, y, z) le coordinate del punto generico $P \in S_3$ si ha $\overrightarrow{OP} = x\vec{i} + y\vec{j} + z\vec{k}$, dunque l'equazione (9.2.1) diventa

$$x\vec{i} + y\vec{j} + z\vec{k} = x_A\vec{i} + y_A\vec{j} + z_A\vec{k} + t(v_x\vec{i} + v_y\vec{j} + v_z\vec{k}) + u(w_x\vec{i} + w_y\vec{j} + w_z\vec{k}), \quad t, u \in \mathbb{R},$$

o, eguagliando le componenti dei due vettori lungo gli assi coordinati,

$$\begin{cases} x = x_A + v_x t + w_x u \\ y = y_A + v_y t + w_y u \\ z = z_A + v_z t + w_z u, \end{cases} \quad t, u \in \mathbb{R}. \quad (9.2.2)$$

Le equazioni (9.2.2) sono dette *equazioni parametriche del piano α passante per il punto $A = (x_A, y_A, z_A)$ e parallelo ai vettori $\vec{v} = v_x\vec{i} + v_y\vec{j} + v_z\vec{k}$, $\vec{w} = w_x\vec{i} + w_y\vec{j} + w_z\vec{k}$.*

Viceversa, supponiamo di avere fissato in S_3 un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$. Dati numeri reali fissati $x_A, y_A, z_A, v_x, v_y, v_z, w_x, w_y, w_z$, si consideri il luogo α dei punti dello spazio $P = (x, y, z)$ le cui coordinate sono della forma

$$\begin{cases} x = x_A + v_x t + w_x u \\ y = y_A + v_y t + w_y u \\ z = z_A + v_z t + w_z u, \end{cases}$$

al variare di due parametri $t, u \in \mathbb{R}$. Procedendo come nel caso della retta, è facile verificare che tale luogo è il piano α passante per il punto $A = (x_A, y_A, z_A)$ e parallelo ai vettori $\vec{v} = v_x \vec{i} + v_y \vec{j} + v_z \vec{k}$, $\vec{w} = w_x \vec{i} + w_y \vec{j} + w_z \vec{k}$.

Esempio 9.8. In S_3 sia fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$. Siano poi dati $A = (1, 2, 3) \in S_3$ e $\vec{v} = 2\vec{i} - 3\vec{k}$, $\vec{w} = \vec{i} + \vec{j} + \vec{k}$ in $V_3(O)$. I vettori \vec{v} e \vec{w} non sono paralleli, quindi i dati individuano un piano α le cui equazioni parametriche sono date da

$$\begin{cases} x = 1 + 2t + u \\ y = 2 + u \\ z = 3 - 3t + u, \end{cases} \quad t, u \in \mathbb{R}. \quad (9.2.3)$$

Si noti che la retta r dell'Esempio 9.1 di equazioni parametriche

$$\begin{cases} x = 1 + 2t \\ y = 2 \\ z = 3 - 3t, \end{cases} \quad t \in \mathbb{R},$$

giace su α : infatti i suoi punti si ottengono ponendo $u = 0$ in (9.2.3). ♠

Per concludere il paragrafo, osserviamo come ricondurci al caso precedente se ci viene chiesto di descrivere un piano α a partire da tre suoi punti A, B e C non allineati, cosa che è noto essere possibile dalla geometria euclidea. Un punto, per esempio A , l'abbiamo: per costruire due vettori paralleli a α basta considerare $B - A$ e $C - A$. Se rispetto al sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ fissato in S_3 i punti hanno coordinate $A = (x_A, y_A, z_A)$, $B = (x_B, y_B, z_B)$, $C = (x_C, y_C, z_C)$, allora

$$\begin{aligned} B - A &= (x_B - x_A)\vec{i} + (y_B - y_A)\vec{j} + (z_B - z_A)\vec{k}, \\ C - A &= (x_C - x_A)\vec{i} + (y_C - y_A)\vec{j} + (z_C - z_A)\vec{k}. \end{aligned}$$

Sostituendo nell'equazione (9.2.2) otteniamo le *equazioni parametriche del piano α passante per i tre punti A, B, C* :

$$\begin{cases} x = x_A + (x_B - x_A)t + (x_C - x_A)u \\ y = y_A + (y_B - y_A)t + (y_C - y_A)u \\ z = z_A + (z_B - z_A)t + (z_C - z_A)u, \end{cases} \quad t, u \in \mathbb{R},$$

oppure

$$\begin{cases} x = (1 - t - u)x_A + tx_B + ux_C \\ y = (1 - t - u)y_A + ty_B + uy_C \\ z = (1 - t - u)z_A + tz_B + uz_C \end{cases} \quad t, u \in \mathbb{R}. \quad (9.2.4)$$

Talvolta le equazioni sopra si scrivono sinteticamente nella forma

$$P = (1 - t - u)A + tB + uC \quad t, u \in \mathbb{R}.$$

Se poi vogliamo descrivere le coordinate dei punti del triangolo Δ_{ABC} è sufficiente che ci limitiamo a considerare i punti le cui coordinate si possono esprimere tramite la formula (9.2.4) con $t, u \in [0, 1]$ e $t + u \leq 1$, cioè $P = (x, y, z) \in \Delta_{ABC}$ se e solo se

$$\begin{cases} x = (1 - t - u)x_A + tx_B + ux_C \\ y = (1 - t - u)y_A + ty_B + uy_C \\ z = (1 - t - u)z_A + tz_B + uz_C \end{cases} \quad t, u, t + u \in [0, 1],$$

o, equivalentemente, se

$$\begin{cases} x = \lambda x_A + \mu x_B + \nu x_C \\ y = \lambda y_A + \mu y_B + \nu y_C \\ z = \lambda z_A + \mu z_B + \nu z_C \end{cases} \quad \lambda, \mu, \nu \geq 0, \lambda + \mu + \nu = 1.$$

Esempio 9.9. In un sistema di riferimento fissato $O\vec{i}\vec{j}\vec{k}$ in S_3 si considerino i tre punti $A = (1, 2, -3)$, $B = (2, 1, 1)$, $C = (2, 2, 2)$. Chiaramente $A \neq B$, quindi esiste unico un piano α contenente A , B e C le cui equazioni parametriche si ottengono utilizzando la formula (9.2.4)

$$\begin{cases} x = 1 + t + u \\ y = 2 - t \\ z = -3 + 4t + 5u, \end{cases} \quad t, u \in \mathbb{R}. \quad \spadesuit$$

Equazioni cartesiane di rette e piani nello spazio

10.1 Equazioni cartesiane di piani

Sia $\alpha \subseteq S_3$ un piano. Nel capitolo precedente abbiamo visto che α può essere individuato da un suo punto qualsiasi A e da due direzioni a esso parallele; è però anche possibile individuare lo stesso piano da un suo punto A e da una retta r per l'origine a esso perpendicolare. Supponiamo che la retta r sia a sua volta individuata da un suo vettore non nullo \vec{v} ; allora se P è un qualsiasi punto del piano α , si ha $P - A \perp \vec{v}$ e viceversa, come si può vedere qui sotto nella Figura 10.1.

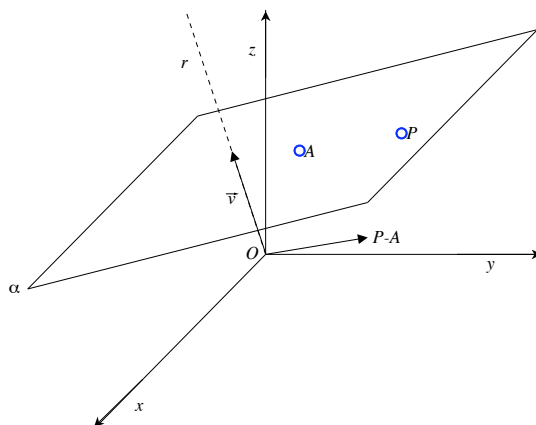


Figura 10.1

Segue allora dalla Proposizione 8.5 e dalla Definizione 8.6 che i punti $P \in \alpha$ sono tutti e soli i punti dello spazio tali che

$$\langle \vec{v}, P - A \rangle = 0. \quad (10.1.1)$$

Fissiamo adesso un sistema di riferimento ortogonale $O\vec{i}\vec{j}\vec{k}$ in S_3 : otteniamo coordinate $A = (x_A, y_A, z_A)$ e $\vec{v} = a\vec{i} + b\vec{j} + c\vec{k}$. Se indichiamo con (x, y, z) le coordinate del punto generico $P \in S_3$, allora vale che

$$P - A = (x - x_A)\vec{i} + (y - y_A)\vec{j} + (z - z_A)\vec{k}$$

e l'equazione (10.1.1) diventa

$$a(x - x_A) + b(y - y_A) + c(z - z_A) = 0.$$

Sviluppando i prodotti e ponendo $d = ax_A + by_A + cz_A$ otteniamo un'equazione della forma

$$ax + by + cz = d. \quad (10.1.2)$$

L'equazione (10.1.2) viene chiamata *equazione cartesiana del piano α passante per il punto $A = (x_A, y_A, z_A)$ e perpendicolare al vettore $\vec{v} = a\vec{i} + b\vec{j} + c\vec{k}$* .

Esempio 10.1. In S_3 sia fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$. Siano poi dati $A = (1, 2, 3) \in S_3$ e $\vec{v} = 2\vec{i} + \vec{j} - 3\vec{k} \in V_3(O)$. L'equazione cartesiana del piano α di S_3 perpendicolare al vettore \vec{v} è data da $2(x - 1) + 1(y - 2) - 3(z - 3) = 0$, cioè

$$2x + y - 3z = -5. \quad (10.1.3)$$

Ci chiediamo se qualcuno fra i punti $B = (1, 1, 1)$ e $C = (-2, 2, 1)$ di S_3 appartenga al piano α . Per sincerarcene basta osservare che ${}^t(1 \ 1 \ 1)$ non è soluzione dell'equazione (10.1.3), quindi $B \notin \alpha$, mentre ${}^t(-2 \ 2 \ 1)$ lo è, quindi $C \in \alpha$. ♠

Viceversa, supponiamo di avere fissato in S_3 un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$. Fissati numeri reali a, b, c, d , con $a\vec{i} + b\vec{j} + c\vec{k} \neq \vec{0}$, si consideri il luogo α dei punti P dello spazio le cui coordinate (x, y, z) soddisfano l'equazione

$$ax + by + cz = d.$$

Sia ${}^t(x_A \ y_A \ z_A)$ una soluzione di tale equazione e sia $A = (x_A, y_A, z_A) \in S_3$. Allora $d = ax_A + by_A + cz_A$, sicché l'equazione di cui sopra diviene

$$a(x - x_A) + b(y - y_A) + c(z - z_A) = 0,$$

ovvero, posto $\vec{v} = a\vec{i} + b\vec{j} + c\vec{k}$,

$$\langle \vec{v}, P - A \rangle = 0.$$

Quindi α è il piano passante per A e perpendicolare a \vec{v} .

Concludiamo che, fissato in S_3 un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$, ogni piano può essere descritto mediante un'equazione della forma (10.1.2) con a, b, c non simultaneamente nulli e, viceversa, ogni equazione della forma (10.1.2) con a, b, c non simultaneamente nulli rappresenta un piano.

Si noti che dato un piano α rappresentato tramite un'equazione della forma (10.1.2), è facile determinarne un punto (basta scegliere una soluzione dell'equazione) e un vettore a esso perpendicolare (basta considerare il vettore $a\vec{i} + b\vec{j} + c\vec{k}$). Di conseguenza, se di piani ne abbiamo due, è facile stabilire essi sono paralleli o meno.

Esempio 10.2. In S_3 sia fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ e si considerino il piano α dell'Esempio 10.1 e il piano β di equazione $x + y + z = 1$.

I piani α e β sono perpendicolari ai vettori $\vec{v} = 2\vec{i} + \vec{j} - 3\vec{k}$ e $\vec{w} = \vec{i} + \vec{j} + \vec{k}$ rispettivamente: poiché $\vec{v} \nparallel \vec{w}$ segue che α e β non sono paralleli, in particolare si intersecano in una retta.

Sia γ il piano d'equazione $-2x - y + 3z + 1 = 0$: un vettore perpendicolare a γ è $-\vec{v} = -2\vec{i} - \vec{j} + 3\vec{k}$, quindi α e γ sono paralleli. Inoltre il sistema

$$\begin{cases} 2x + y - 3z = -5 \\ -2x - y + 3z + 1 = 0 \end{cases}$$

non può avere soluzione, quindi $\alpha \cap \gamma = \emptyset$: concludiamo che α e γ sono paralleli e distinti.

Infine sia δ il piano d'equazione $-2x - y + 3z - 5 = 0$. Anche α e δ sono paralleli, ma il sistema

$$\begin{cases} 2x + y - 3z = -5 \\ -2x - y + 3z - 5 = 0 \end{cases}$$

ha soluzione: concludiamo che $\alpha = \delta$. ♠

L'Esempio 10.2 mette in luce un legame fra le posizioni relative di due piani nello spazio e le soluzioni del sistema dato dalle loro equazioni cartesiane: vediamo tale legame in dettaglio.

Proposizione 10.3 (Posizioni relative di due piani).

In S_3 sia fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$. Dati i piani α e α' rispettivamente di equazioni

$$\alpha : ax + by + cz = d, \quad \alpha' : a'x + b'y + c'z = d',$$

siano

$$A = \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix}, \quad (A|B) = \begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \end{pmatrix}.$$

Allora:

- (i) $\alpha = \alpha'$ se e solo se $\text{rk}(A) = 1 = \text{rk}(A|B)$;
- (ii) α e α' sono paralleli e distinti se e solo se $\text{rk}(A) = 1, \text{rk}(A|B) = 2$;
- (iii) α e α' si intersecano lungo una retta se e solo se $\text{rk}(A) = 2 = \text{rk}(A|B)$.

Dimostrazione. Come già osservato i punti d'intersezione corrispondono alle soluzioni del sistema

$$\begin{cases} ax + by + cz = d \\ a'x + b'y + c'z = d' \end{cases} \quad (10.1.4)$$

I piani α e α' sono paralleli e distinti se e solo se $\alpha \cap \alpha' = \emptyset$, cioè se e solo se il sistema (10.1.4) non ha soluzioni. Ciò può accadere se e solo se $\text{rk}(A) = 1, \text{rk}(A|B) = 2$.

I piani α e α' sono coincidenti se e solo se il sistema (10.1.4) è equivalente alla sola equazione $ax + by + cz = d$, ovvero se e solo se $\text{rk}(A) = 1 = \text{rk}(A|B)$.

Infine i piani α e α' si intersecano lungo una retta se e solo se non sono paralleli coincidenti o distinti, ovvero se e solo se $\text{rk}(A) = 2 = \text{rk}(A|B)$. □

Analoghi risultati si possono dimostrare nel caso di tre o più piani.

Vogliamo ora studiare come distinguere le posizioni relative che possono assumere un piano α e una retta r nello spazio S_3 . A tale scopo ricordiamo che un piano α e una retta r in S_3

- possono essere *incidenti in un unico punto*,
- r può essere contenuta in α ,
- oppure r può non avere punti in comune con α :

(in questi due ultimi casi si dice che r e α sono *paralleli*).

Lo studio della posizione relativa di un piano di cui è nota l'equazione cartesiana e di una retta di cui è noto un sistema di equazioni parametriche è immediato, come si può vedere dall'esempio seguente.

Esempio 10.4. In S_3 sia fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ e si consideri ancora una volta il piano α dell'Esempio 10.1, di equazione $2x + y - 3z = -5$; un vettore perpendicolare ad α è $\vec{v} = 2\vec{i} + \vec{j} - 3\vec{k}$. Siano r' , r'' , r''' le rette di equazioni parametriche

$$r' : \begin{cases} x = t - 1 \\ y = t \\ z = 1 - t, \end{cases} \quad r'' : \begin{cases} x = t \\ y = t + 2 \\ z = 1 + t, \end{cases} \quad r''' : \begin{cases} x = -1 + t \\ y = t - 3 \\ z = t, \end{cases}$$

al variare di $t \in \mathbb{R}$, rispettivamente.

Iniziamo a determinare l'intersezione $\alpha \cap r'$. Osserviamo che se $P \in r'$ le sue coordinate sono della forma $(t - 1, t, 1 - t)$ per un qualche $t \in \mathbb{R}$: affinché $P \in \alpha$ allora $\begin{pmatrix} t - 1 & t & 1 - t \end{pmatrix}$ deve essere soluzione dell'equazione di α , cioè si deve avere

$$2(t - 1) + t - 3(1 - t) = -5,$$

da cui si deduce $6t = 0$, ovvero $t = 0$, che corrisponde al punto di coordinate $(-1, 0, 1)$.

Si noti che verificare che α e r' non sono paralleli senza determinarne il punto di intersezione è immediato: un vettore parallelo a r' è $\vec{v}' = \vec{i} + \vec{j} - \vec{k}$ ed è facile convincersi che $\alpha \parallel r'$ se e solo se $\vec{v} \perp \vec{v}'$: poiché $\langle \vec{v}, \vec{v}' \rangle = 6$ segue che $\vec{v} \not\perp \vec{v}'$, dunque $\alpha \not\parallel r'$.

Passiamo a determinare $\alpha \cap r''$. Procediamo come sopra sostituendo le equazioni parametriche di r'' nell'equazione di α : si ottiene

$$2t + (t + 2) - 3(1 + t) = -5,$$

da cui si deduce $-1 = -5$, che non ha soluzioni, perciò $\alpha \cap r'' = \emptyset$.

Concludiamo esaminando l'intersezione $\alpha \cap r'''$. Procedendo come sopra si ottiene

$$2(-1 + t) + (t - 3) - 3t = -5,$$

da cui si deduce $0 = 0$. Ogni $t \in \mathbb{R}$ è soluzione di tale equazione, quindi ogni punto di r''' è in α , perciò la retta r''' giace sul piano α , $r''' \subset \alpha$.

Nei secondi due casi risulta che il piano e la retta sono paralleli, cosa che si poteva dedurre direttamente. Infatti un vettore parallelo a r'' e r''' è $\vec{v}'' = \vec{i} + \vec{j} + \vec{k}$. Poiché $\langle \vec{v}, \vec{v}'' \rangle = 0$ segue che $\vec{v} \perp \vec{v}''$, dunque $\alpha \parallel r'', r'''$. ♠

Concludiamo il paragrafo ricordando che è noto dalla geometria euclidea che è possibile descrivere un piano α dando tre punti non allineati A, B, C che gli appartengono, come nella Figura 10.2.

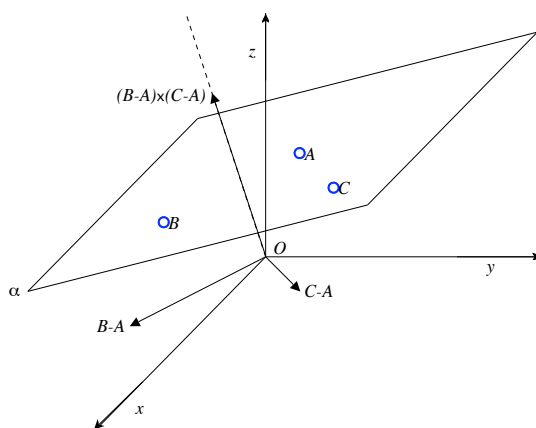


Figura 10.2

Ci possiamo ricondurre al caso precedente osservando che i punti A, B, C, P sono complanari se e solo se tali sono i vettori $P - A, B - A, C - A$. Nel paragrafo 8.3 abbiamo visto che tre vettori sono complanari se e solo se il loro prodotto misto si annulla; la condizione di complanarità si traduce quindi nell'equazione

$$\langle P - A, (B - A) \times (C - A) \rangle = 0.$$

Fissiamo ora un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ in S_3 nel quale i punti A, B, C abbiano coordinate $A = (x_A, y_A, z_A), B = (x_B, y_B, z_B), C = (x_C, y_C, z_C)$, quindi

$$\begin{aligned} B - A &= (x_B - x_A)\vec{i} + (y_B - y_A)\vec{j} + (z_B - z_A)\vec{k}, \\ C - A &= (x_C - x_A)\vec{i} + (y_C - y_A)\vec{j} + (z_C - z_A)\vec{k}. \end{aligned}$$

Un punto generico P di coordinate (x, y, z) giace sul piano α se e solo se

$$\begin{vmatrix} x - x_A & y - y_A & z - z_A \\ x_B - x_A & y_B - y_A & z_B - z_A \\ x_C - x_A & y_C - y_A & z_C - z_A \end{vmatrix} = 0. \quad (10.1.5)$$

Esempio 10.5. In S_3 sia fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ e si considerino i punti $A = (1, 1, 0), B = (1, 0, 1), C = (0, 1, 1)$. Allora $B - A = -\vec{j} + \vec{k}, C - A = -\vec{i} + \vec{k}$:

poiché $B - A$ e $C - A$ non sono proporzionali, i tre punti A, B, C non sono allineati. Per determinare il piano α individuato da A, B, C si può applicare la formula (10.1.5). Poiché

$$\begin{vmatrix} x-1 & y-1 & z \\ 0 & -1 & 1 \\ -1 & 0 & 1 \end{vmatrix} = -x + 1 - y + 1 - z,$$

l'equazione del piano α è $x + y + z = 2$. ♠

10.2 Equazioni cartesiane di rette

Come già osservato nel paragrafo precedente, due piani non paralleli in S_3 , α e α' , si intersecano lungo una retta r , come nella Figura 10.3; viceversa, ogni retta r può essere descritta in questo modo come intersezione di una coppia qualsiasi di piani distinti che la contengono.

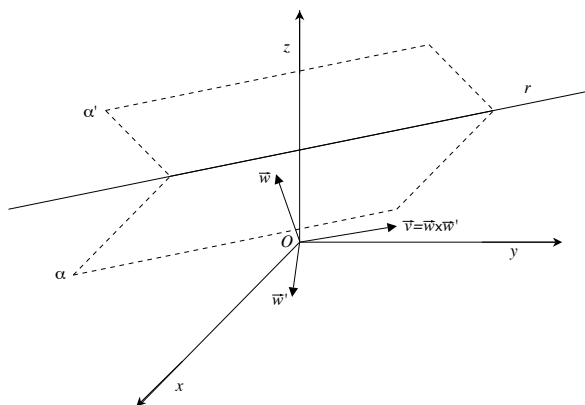


Figura 10.3

Se fissiamo un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ in S_3 , i piani α e α' possono essere descritti mediante due equazioni cartesiane, diciamo

$$ax + by + cz = d, \quad a'x + b'y + c'z = d'$$

rispettivamente, soddisfacenti per la Proposizione 10.3 alla condizione

$$\text{rk} \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} = 2.$$

Se ciò accade le equazioni

$$\begin{cases} ax + by + cz = d \\ a'x + b'y + c'z = d' \end{cases} \quad (10.2.1)$$

vengono dette *equazioni cartesiane della retta r intersezione dei due piani α e α'* .

Si noti che, poiché $\vec{w} = a\vec{i} + b\vec{j} + c\vec{k}$ e $\vec{w}' = a'\vec{i} + b'\vec{j} + c'\vec{k}$ sono rispettivamente perpendicolari a α e α' , allora $\vec{v} = \vec{w} \times \vec{w}'$ è parallelo a $r = \alpha \cap \alpha'$; in altre parole, per ottenere il vettore direzione di una retta intersezione di due piani è sufficiente prendere il prodotto vettoriale dei due vettori perpendicolari ai due piani.

Esempio 10.6. Abbiamo visto nell'Esempio 10.2 che il sistema

$$\begin{cases} 2x + y - 3z = -5 \\ x + y + z = 1 \end{cases}$$

definisce una retta r perché

$$\text{rk} \begin{pmatrix} 2 & 1 & -3 \\ 1 & 1 & 1 \end{pmatrix} = 2.$$

Verifichiamo se qualcuno fra i punti $A = (0, 0, 0)$, $B = (-1, 0, 1)$, $C = (1, 0, 0)$, $D = (-2, 2, 1)$ giace sulla retta r . Per fare ciò basta sostituire le coordinate di tali punti nelle due equazioni del sistema: se entrambe le equazioni sono soddisfatte il punto appartiene alla retta, se almeno una delle due equazioni non è soddisfatta allora il punto non vi appartiene. Con questo in mente è facile verificare che $A, B, C \notin r$: invece $D \in r$.

Un vettore parallelo a r è

$$\vec{v} = (2\vec{i} + \vec{j} - 3\vec{k}) \times (\vec{i} + \vec{j} + \vec{k}) = 4\vec{i} - 5\vec{j} + \vec{k}.$$

In particolare r è la retta per $D = (-2, 2, 1)$ parallela a $\vec{v} = 4\vec{i} - 5\vec{j} + \vec{k}$, quindi un sistema di equazioni parametriche per r è

$$\begin{cases} x = -2 + 4t \\ y = 2 - 5t \\ z = 1 + t, \end{cases} \quad t \in \mathbb{R}. \quad \spadesuit$$

⚠ Le equazioni cartesiane di una retta r non sono univocamente determinate: esse dipendono dalla scelta di una coppia di piani per r .

Nel paragrafo 9.1 abbiamo mostrato che ogni retta può essere rappresentata mediante le sue equazioni parametriche, dunque viene spontaneo chiedersi quale sia, fissata una retta $r \subseteq S_3$, il legame fra questi due metodi di rappresentarla e come si possa passare dall'uno all'altro.

Supponiamo ad esempio che la retta r sia definita dalle equazioni parametriche

$$\begin{cases} x = x_0 + lt \\ y = y_0 + mt \\ z = z_0 + nt, \end{cases} \quad t \in \mathbb{R},$$

con l, m, n tutti non nulli. Esplicitando il parametro t troviamo

$$t = \frac{x - x_0}{l} = \frac{y - y_0}{m} = \frac{z - z_0}{n},$$

da cui si ricavano tre equazioni lineari non contenenti il parametro t :

$$m(x - x_0) = l(y - y_0), \quad n(x - x_0) = l(z - z_0), \quad n(y - y_0) = m(z - z_0),$$

che rappresentano tre piani che chiameremo α , β e γ rispettivamente. Tali piani, per come sono stati ottenuti, contengono la retta r . Si noti che i vettori

$$\vec{u} = m\vec{i} - l\vec{j}, \quad \vec{v} = n\vec{i} - l\vec{j}, \quad \vec{w} = n\vec{j} - m\vec{k}$$

sono rispettivamente perpendicolari a α , β , γ . Poiché tali vettori non sono proporzionali a coppie, ciascuna coppia di piani costituisce un insieme di equazioni cartesiane per la retta r .

Per esercizio, verificare che tale metodo continua a valere anche se qualcuno fra l, m, n è nullo (al massimo possono esserlo due fra di essi).

Esempio 10.7. Riprendiamo la retta r dell'Esempio 9.1, le cui equazioni parametriche sono date da

$$\begin{cases} x = 1 + 2t \\ y = 2 \\ z = 3 - 3t, \end{cases} \quad t \in \mathbb{R}.$$

Procedendo come spiegato sopra si ha, formalmente,

$$t = \frac{x - 1}{2} = \frac{y - 2}{0} = \frac{z - 3}{-3}$$

(dove si è adottata la convenzione standard che se in una frazione il denominatore è zero, anche il numeratore lo è).

Eliminando i denominatori si hanno le tre equazioni

$$y - 2 = 0, \quad -3(x - 1) = 2(z - 3), \quad y - 2 = 0,$$

cioè

$$y - 2 = 0, \quad 3x + 2z = 3, \quad y - 2 = 0.$$

Quindi un sistema di equazioni cartesiane per r è

$$\begin{cases} y - 2 = 0 \\ 3x + 2z = 3. \end{cases} \quad \spadesuit$$

Viceversa supponiamo di avere un sistema di equazioni cartesiane di una retta r della forma (10.2.1). Se risolviamo tale sistema, poiché

$$\text{rk} \begin{pmatrix} a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} = 2,$$

le sue soluzioni dipendono da un unico parametro $t \in \mathbb{R}$ e sono quindi della forma

$$\left\{ \begin{pmatrix} x_0 + lt \\ y_0 + mt \\ z_0 + nt \end{pmatrix} \mid t \in \mathbb{R} \right\}.$$

Quindi il generico punto della retta r è $(x, y, z) = (x_0 + lt, y_0 + mt, z_0 + nt)$, cioè

$$\begin{cases} x = x_0 + lt \\ y = y_0 + mt \\ z = z_0 + nt, \end{cases} \quad t \in \mathbb{R},$$

sono equazioni parametriche per la retta r .

Esempio 10.8. Riprendiamo ancora l'Esempio 10.1: sia r la retta avente equazioni cartesiane

$$\begin{cases} 2x + y - 3z = -5 \\ x + y + z = 1, \end{cases} \quad (10.2.2)$$

e analizziamo il sistema lineare associato. Si ha

$$\begin{aligned} \begin{pmatrix} 2 & 1 & -3 & -5 \\ 1 & 1 & 1 & 1 \end{pmatrix} &\xrightarrow{R_2 \rightarrow R_2 - R_1} \begin{pmatrix} 2 & 1 & -3 & -5 \\ -1 & 0 & 4 & 6 \end{pmatrix} \\ &\xrightarrow{R_1 \rightarrow R_1 + 2R_2} \begin{pmatrix} 0 & 1 & 5 & 7 \\ -1 & 0 & 4 & 6 \end{pmatrix} \xrightarrow{R_2 \rightarrow -R_2} \begin{pmatrix} 0 & 1 & 5 & 7 \\ 1 & 0 & -4 & -6 \end{pmatrix}. \end{aligned}$$

Quindi il sistema (10.2.2) è equivalente al sistema

$$\begin{cases} y + 5z = 7 \\ x - 4z = -6, \end{cases}$$

le cui soluzioni sono

$$\left\{ \begin{pmatrix} 4z - 6 \\ 7 - 5z \\ z \end{pmatrix} \mid z \in \mathbb{R} \right\}.$$

Posto $t = z$ otteniamo allora delle equazioni parametriche di r

$$\begin{cases} x = 4t - 6 \\ y = 7 - 5t \\ z = t, \end{cases} \quad t \in \mathbb{R}$$

Tali equazioni sono diverse da quelle che avevamo trovato nell'Esempio 10.6, pur rappresentando la stessa retta. Per verificarlo si osservi che il seguente sistema ha infinite soluzioni:

$$\begin{cases} -2 + 4t = 4t' - 6 \\ 2 - 5t = 7 - 5t' \\ 1 + t = t'. \end{cases} \quad \spadesuit$$

Osservazione 10.9 (Equazioni cartesiane di rette nel piano). Come già sottolineato nell'Osservazione 9.7, spesso si identifica S_2 col piano di S_3 di equazione $z = 0$. In particolare si può pensare ai punti di S_2 come punti di S_3 con quota nulla e ai vettori di $V_2(O)$ come a vettori di $V_3(O)$ con componente nulla nella direzione di \vec{k} .

Con questa identificazione ogni retta $r \subset S_2$, pensata come retta in S_3 ha delle equazioni cartesiane della forma

$$\begin{cases} ax + by = c \\ z = 0 : \end{cases}$$

si noti che in questa scrittura $z = 0$ e $ax + by = c$ sono le equazioni cartesiane di due piani contenenti r .

Nel momento in cui si affronta un problema esclusivamente in S_2 , spesso si omette l'equazione banale $z = 0$. Pertanto si scrive talvolta solo

$$ax + by = c. \quad (10.2.3)$$

Come visto in precedenza un vettore parallelo a r è $(a\vec{i} + b\vec{j}) \times \vec{k} = b\vec{i} - a\vec{j}$.

Osserviamo che, fissato in S_2 un sistema di riferimento $O\vec{i}\vec{j}$, non solo ogni retta può essere descritta mediante un'equazione della forma (10.2.3) con a, b non simultaneamente nulli ma, viceversa, è anche possibile dimostrare (e la dimostrazione è lasciata al lettore) che ogni equazione della forma (10.2.3) con a, b non simultaneamente nulli rappresenta una retta.

10.3 Posizioni relative di rette e piani

Lo studio della posizione relativa di una retta di cui sono note le equazioni cartesiane e di una retta di cui è noto un sistema di equazioni parametriche è immediato, ed è illustrato nell'esempio seguente. Ricordiamo che avevamo analizzato i vari casi di rette coincidenti, incidenti, parallele distinte, sghembe, nel capitolo precedente.

Esempio 10.10. In S_3 sia fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ e si consideri la retta r dell'Esempio 10.2 di equazioni cartesiane

$$\begin{cases} 2x + y - 3z = -5 \\ x + y + z = 1. \end{cases}$$

Sia poi s la retta di equazioni parametriche

$$\begin{cases} x = t - 1 \\ y = -3 + t \\ z = t, \end{cases} \quad t \in \mathbb{R}.$$

Le rette r ed s sono parallele ai vettori $\vec{v} = 4\vec{i} - 5\vec{j} + \vec{k}$ e $\vec{w} = \vec{i} + \vec{j} + \vec{k}$ rispettivamente, quindi $r \nparallel s$. Ci domandiamo se siano incidenti: a tale scopo si può procedere come nell'Esempio 10.4, sostituendo le equazioni di s dentro le equazioni dei due piani che definiscono r . Si ottiene in tal modo il sistema

$$\begin{cases} 2(t-1) + (-3+t) - 3t = -5 \\ (t-1) + (-3+t) + t = 1 \end{cases}$$

che è equivalente a

$$\begin{cases} 0 = 0 \\ 3t = 5. \end{cases}$$

L'unica soluzione di tale sistema è $t = 5/3$; il punto corrispondente sulla retta s è $(2/3, -4/3, 5/3)$ che, per costruzione, appartiene anche a r . ♠

Affrontiamo ora lo studio delle posizioni reciproche di una retta e un piano di cui siano note le equazioni cartesiane. Il seguente risultato è un'applicazione immediata del Teorema di Rouché–Capelli (Proposizione 5.2).

Proposizione 10.11 (Posizioni relative di un piano e una retta).

In S_3 sia fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$. Dati la retta r e il piano α di equazioni cartesiane

$$r : \begin{cases} ax + by + cz = d \\ a'x + b'y + c'z = d', \end{cases} \quad \alpha : a''x + b''y + c''z = d'',$$

siano

$$A = \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix}, \quad (A|B) = \begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{pmatrix}.$$

Allora:

- (i) $r \subseteq \alpha$ se e solo se $\text{rk}(A) = 2 = \text{rk}(A|B)$;
- (ii) $r \cap \alpha = \emptyset$ se e solo se $\text{rk}(A) = 2$, $\text{rk}(A|B) = 3$;
- (iii) r e α si intersecano in un punto se e solo se $\text{rk}(A) = 3 = \text{rk}(A|B)$.

Esempio 10.12. In S_3 sia fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ e si considerino la retta r e il piano α rispettivamente di equazioni cartesiane

$$r : \begin{cases} 2x + y - 3z = -5 \\ x + y + z = 1, \end{cases} \quad \alpha : 3z - 2x - y + 1 = 0.$$

Facendo operazioni elementari sulla matrice completa $(A|B)$ del sistema

$$\begin{cases} 3z - 2x - y + 1 = 0 \\ 2x + y - 3z = -5 \\ x + y + z = 1, \end{cases}$$

otteniamo

$$\left(\begin{array}{ccc|c} -2 & -1 & 3 & -1 \\ 2 & 1 & -3 & -5 \\ 1 & 1 & 1 & 1 \end{array} \right) \xrightarrow{\substack{R_2 \leftrightarrow R_2 - R_1 \\ R_3 \leftrightarrow R_3 + R_1}} \left(\begin{array}{ccc|c} -2 & -1 & 3 & -1 \\ 0 & 0 & 0 & -4 \\ -1 & 0 & 4 & 0 \end{array} \right),$$

da cui deduciamo che $r \cap \alpha = \emptyset$. ♠

Abbiamo visto nel paragrafo precedente che una stessa retta può essere rappresentata da molti sistemi di equazioni cartesiane, anche assai diversi fra di loro. Pertanto è utile capire come variano le equazioni cartesiane di una retta quando cambiamo la coppia di piani che la rappresenta.

A tale scopo consideriamo una retta r le cui equazioni cartesiane sono della forma (10.2.1). Sia poi α il piano di equazione cartesiana $a''x + b''y + c''z = d''$. Affinché $r \subseteq \alpha$ occorre e basta che ogni punto di r appartenga ad α , ovvero che ogni soluzione del sistema (10.2.1) sia anche soluzione dell'equazione $a''x + b''y + c''z = d''$. Questa condizione equivale a chiedere che il sistema (10.2.1) sia equivalente a

$$\begin{cases} ax + by + cz = d \\ a'x + b'y + c'z = d' \\ a''x + b''y + c''z = d''. \end{cases}$$

Sappiamo che almeno uno fra i coefficienti a , b e c è non nullo. Supponiamo sia $a \neq 0$: posto $\alpha' = a'/a$ e $\alpha'' = a''/a$ consideriamo le trasformazioni

$$\left(\begin{array}{ccc|c} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{array} \right) \xrightarrow{\substack{R_2 \leftrightarrow R_2 - \alpha' R_1 \\ R_3 \leftrightarrow R_3 - \alpha'' R_1}} \left(\begin{array}{ccc|c} a & b & c & d \\ 0 & b' - \alpha'b & c' - \alpha'c & d' - \alpha'd \\ 0 & b'' - \alpha''b & c'' - \alpha''c & d'' - \alpha''d \end{array} \right).$$

D'altra parte sappiamo anche che

$$\text{rk} \begin{pmatrix} a & b & c \\ 0 & b' - \alpha'b & c' - \alpha'c \end{pmatrix} = \text{rk} \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} = 2,$$

quindi o $b' - \alpha'b \neq 0$ o $c' - \alpha'c \neq 0$. Supponiamo sia $b' - \alpha'b \neq 0$: posto $\beta = (b'' - \alpha''b)/(b' - \alpha'b)$ possiamo considerare la trasformazione

$$\left(\begin{array}{ccc|c} a & b & c & d \\ 0 & b' - \alpha'b & c' - \alpha'c & d' - \alpha'd \\ 0 & b'' - \alpha''b & c'' - \alpha''c & d'' - \alpha''d \end{array} \right) \xrightarrow{R_3 \leftrightarrow R_3 - \beta R_2} \left(\begin{array}{ccc|c} a & b & c & d \\ 0 & b' - \alpha'b & c' - \alpha'c & d' - \alpha'd \\ 0 & 0 & c'' - \alpha''c - \beta(c' - \alpha'c) & d'' - \alpha''d - \beta(d' - \alpha'd) \end{array} \right).$$

Dalla Proposizione 10.11 deduciamo che $r \subseteq \alpha$ se e solo se

$$\begin{aligned} & \text{rk} \begin{pmatrix} a & b & c \\ 0 & b' - \alpha'b & c' - \alpha'c \\ 0 & 0 & c'' - \alpha''c - \beta(c' - \alpha'c) \end{pmatrix} \\ &= \text{rk} \left(\begin{array}{ccc|c} a & b & c & d \\ 0 & b' - \alpha'b & c' - \alpha'c & d' - \alpha'd \\ 0 & 0 & c'' - \alpha''c - \beta(c' - \alpha'c) & d'' - \alpha''d - \beta(d' - \alpha'd) \end{array} \right) = 2, \end{aligned}$$

cioè se e solo se

$$\begin{aligned} c'' - (\alpha'' - \beta\alpha')c - \beta c' &= c'' - \alpha''c - \beta(c' - \alpha'c) = 0, \\ d'' - (\alpha'' - \beta\alpha')d - \beta d' &= d'' - \alpha''d - \beta(d' - \alpha'd) = 0. \end{aligned}$$

Ciè se e solo se

$$a''x + b''y + c''z - d'' = \lambda(ax + by + cz - d) + \mu(a'x + b'y + c'z - d')$$

ove $\lambda = \alpha'' - \beta\alpha'$ e $\mu = \beta$. Concludiamo che i piani α contenenti r sono tutti e solo quelli la cui equazione si può scrivere nella forma

$$\lambda(ax + by + cz - d) + \mu(a'x + b'y + c'z - d') = 0 \quad (10.3.1)$$

per qualche coppia $\lambda, \mu \in \mathbb{R}$. Questo risultato viene spesso chiamato *metodo del fascio (di piani)*.

Esempio 10.13. Si consideri la retta r dell'Esempio 10.6; essendo parallela al vettore $\vec{v} = 4\vec{i} - 5\vec{j} + \vec{k}$, essa è anche parallela alla retta s di equazioni parametriche

$$\begin{cases} x = 4t \\ y = -5t \\ z = t, \end{cases} \quad t \in \mathbb{R}.$$

Inoltre il punto $(0, 0, 0) \notin r$, quindi $r \neq s$; pertanto r ed s sono contenute in un unico piano α di cui vogliamo determinare un'equazione cartesiana.

Osserviamo che tale piano deve contenere r : il metodo del fascio ci permette allora di dedurre che l'equazione di α deve essere della forma dell'equazione (10.3.1), cioè

$$\lambda(2x + y - 3z + 5) + \mu(x + y + z - 1) = 0. \quad (10.3.2)$$

Dobbiamo determinare $\lambda, \mu \in \mathbb{R}$ in maniera tale che il piano avente equazione 10.3.2 contenga la retta s : a tale scopo, essendo $r \parallel s$, è sufficiente scegliere $\lambda, \mu \in \mathbb{R}$ in modo tale che il piano avente equazione 10.3.2 contenga almeno un punto di $s \setminus r$, per esempio $(0, 0, 0)$. Concludiamo che $5\lambda = \mu$ ovvero le equazioni cercate sono $\lambda(7x + 6y + 2z) = 0$. Si noti che tutte queste equazioni sono proporzionali, quindi definiscono lo stesso piano; dunque possiamo fissare λ , per esempio $\lambda = 1$. ♠

Concludiamo il paragrafo studiando l'interessante caso dello studio delle posizioni relative di due rette descritte tramite le loro equazioni cartesiane.

Proposizione 10.14 (Posizioni relative di due rette nello spazio).

In S_3 sia fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$. Date le rette r e r' di equazioni cartesiane

$$r : \begin{cases} ax + by + cz = d \\ a'x + b'y + c'z = d' \end{cases}, \quad r' : \begin{cases} a''x + b''y + c''z = d'' \\ a'''x + b'''y + c'''z = d''' \end{cases},$$

siano

$$A = \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \\ a''' & b''' & c''' \end{pmatrix}, \quad (A|B) = \begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \\ a''' & b''' & c''' & d''' \end{pmatrix}.$$

Allora:

- (i) $r = r'$ se e solo se $\text{rk}(A) = 2 = \text{rk}(A|B)$;
- (ii) r e r' sono parallele e distinte se e solo se $\text{rk}(A) = 2$, $\text{rk}(A|B) = 3$;
- (iii) r ed r' si intersecano in un punto se e solo se $\text{rk}(A) = 3 = \text{rk}(A|B)$;
- (iv) r ed r' sono sghembe se e solo se $\text{rk}(A) = 3$, $\text{rk}(A|B) = 4$.

Dimostrazione. Si considerino le due rette r_O e r'_O rispettivamente di equazioni

$$r_O : \begin{cases} ax + by + cz = 0 \\ a'x + b'y + c'z = 0 \end{cases}, \quad r'_O : \begin{cases} a''x + b''y + c''z = 0 \\ a'''x + b'''y + c'''z = 0 \end{cases}.$$

Tali rette passano per l'origine O . Inoltre r_O è parallela al vettore $(a\vec{i} + b\vec{j} + c\vec{k}) \times (a'\vec{i} + b'\vec{j} + c'\vec{k})$, dunque $r \parallel r_O$; similmente $r' \parallel r'_O$.

Poiché $A \in \mathbb{R}^{4,3}$ segue che $2 \leq \text{rk}(A) \leq 3$. Chiaramente la condizione $r \parallel r'$ equivale a $r_O = r'_O$. Ciò equivale a dire che ogni piano contenente r'_O contiene anche r_O . Tenendo conto di come si possono scrivere le equazioni dei piani contenenti una retta data (cioè del metodo del fascio), tale condizione equivale a $\text{rk}(A) = 2$. Di conseguenza deduciamo anche che $r \not\parallel r'$ se e solo se $\text{rk}(A) = 3$.

Tenendo conto che $\text{rk}(A|B) \leq \text{rk}(A) + 1$ (perché aggiungiamo una sola colonna), segue la classificazione data nell'enunciato. \square

Esempio 10.15. In S_3 sia fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ e si considerino le rette r e r' rispettivamente di equazioni cartesiane

$$r : \begin{cases} 2x + y - 3z = -5 \\ x + y + z = 1 \end{cases}, \quad r' : \begin{cases} 3x - y + 3z = 2 \\ x - y + z = 1 \end{cases}.$$

Facendo operazioni elementari sulla matrice completa

$$(A|B) = \left(\begin{array}{ccc|c} 2 & 1 & -3 & -5 \\ 1 & 1 & 1 & 1 \\ 3 & -1 & 3 & 2 \\ 1 & -1 & 1 & 1 \end{array} \right)$$

del sistema

$$\begin{cases} 2x + y - 3z = -5 \\ x + y + z = 1 \\ 3x - y + 3z = 2 \\ x - y + z = 1, \end{cases}$$

otteniamo

$$\begin{aligned} (A|B) &\xrightarrow{\substack{R_2 \rightarrow R_2 - R_1 \\ R_3 \rightarrow R_3 + R_1 \\ R_4 \rightarrow R_4 + R_1}} \left(\begin{array}{ccc|c} 2 & 1 & -3 & -5 \\ -1 & 0 & 4 & 6 \\ 5 & 0 & 0 & -3 \\ 3 & 0 & -2 & -4 \end{array} \right) \\ &\xrightarrow{R_4 \rightarrow R_4 + R_2/2} \left(\begin{array}{ccc|c} 2 & 1 & -3 & -5 \\ -1 & 0 & 4 & 6 \\ -5 & 0 & 0 & -3 \\ 5/2 & 0 & 0 & -1 \end{array} \right) \xrightarrow{R_4 \rightarrow R_4 + R_3/2} \left(\begin{array}{ccc|c} 2 & 1 & -3 & -5 \\ -1 & 0 & 4 & 6 \\ -5 & 0 & 0 & -3 \\ 0 & 0 & 0 & -5/2 \end{array} \right), \end{aligned}$$

da cui deduciamo che r e r' sono sghembe. ♠

Minima distanza

11.1 Minima distanza tra insiemi

Definizione 11.1.

In S_n , $n = 2, 3$, sia fissata un'unità di misura u . Dati due punti $A, B \in S_n$, definiamo *distanza fra A e B*, e scriviamo $d(A, B)$, la lunghezza del segmento \overline{AB} rispetto all'unità di misura u .

Abbiamo già visto nella Lezione 7 come calcolare la lunghezza di un segmento avente come estremi due punti $A, B \in S_3$ di cui si conoscono le coordinate rispetto a un fissato sistema di riferimento $O\vec{i}\vec{j}\vec{k}$: se $A = (x_A, y_A, z_A)$ e $B = (x_B, y_B, z_B)$, allora la formula (7.3.1) afferma che

$$d(A, B) = |\overline{AB}| = |B - A| = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2}.$$

Vogliamo estendere la nozione di distanza a una qualsiasi coppia di sottoinsiemi $X, Y \subseteq S_n$. Osserviamo subito che, in generale, X e Y conterranno più punti, quindi avremo più segmenti aventi un estremo in X e uno in Y .

È possibile perciò associare agli insiemi X e Y un insieme di numeri reali

$$D_{X,Y} = \{ d(x, y) \mid x \in X, y \in Y \}.$$

Poiché $d(x, y) \geq 0$ per ogni coppia di punti $x, y \in S_n$ segue che $D_{X,Y} \subseteq [0, +\infty[$. In particolare $D_{X,Y}$ è un insieme limitato inferiormente, quindi risulta avere un estremo inferiore per la completezza di \mathbb{R} .

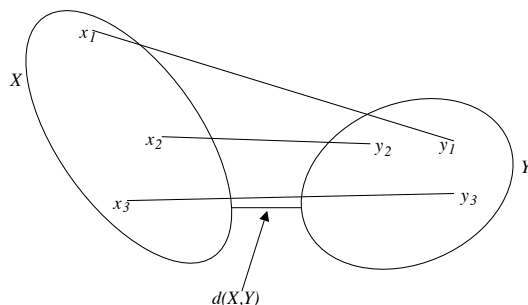


Figura 11.1

Definizione 11.2 (Minima distanza).

In S_n , $n = 2, 3$, sia fissata un'unità di misura u . Dati due sottoinsiemi $X, Y \subseteq S_n$, definiamo *minima distanza fra X e Y* il numero

$$d(X, Y) = \inf_{x \in X, y \in Y} D_{X, Y}.$$

Verrebbe naturale pensare alla minima distanza fra X e Y come alla “minima” delle distanze fra i punti di X e i punti di Y , ma in generale ciò non è vero.

Esempio 11.3. In S_2 si considerino gli insiemi $X = \{ (-1/n, 0) \mid n \in \mathbb{Z}, n > 0 \}$ e $Y = \{(0, 0)\}$. Osserviamo che per ogni $n \in \mathbb{Z}, n > 0$ si ha

$$d((-1/n, 0), (0, 0)) = \frac{1}{n} > 0,$$

quindi

$$D_{X, Y} = \left\{ \frac{1}{n} \mid n \in \mathbb{Z}, n > 0 \right\}.$$

Concludiamo allora che

$$d(X, Y) = \inf_{n \in \mathbb{Z}, n > 0} \left\{ \frac{1}{n} \mid n \in \mathbb{Z}, n > 0 \right\} = 0,$$

ma, chiaramente, $X \cap Y = \emptyset$ ♠

Si noti che se i due insiemi X e Y sono formati da un punto ciascuno, ad esempio $X = \{A\}$ e $Y = \{B\}$, allora

$$d(X, Y) = d(A, B).$$

Inoltre, se i due insiemi hanno intersezione non vuota, $X \cap Y \neq \emptyset$, allora hanno minima distanza nulla. Infatti scelto $A \in X \cap Y$ risulta

$$0 \leq d(X, Y) \leq d(A, A) = 0,$$

sicché $d(X, Y) = 0$.

⚠ Il viceversa non è vero, cioè se $d(X, Y) = 0$ non è detto che $X \cap Y \neq \emptyset$, come si vede dall'Esempio 11.3.

11.2 Distanza di un punto da una retta o da un piano

In questo paragrafo spiegheremo come calcolare la minima distanza $d(X, Y)$ nei due casi in cui $X = \{P_0\}$ e Y è una retta o un piano. A tale scopo ci restringeremo al

caso di sottoinsiemi dello spazio S_3 e supporremo, d'ora innanzi, di aver fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$, nel quale $P_0 = (x_0, y_0, z_0)$.

Cominciamo dal caso in cui Y è un piano α di equazione $ax + by + cz = d$: dobbiamo quindi valutare la quantità

$$d(P_0, \alpha) = \inf_{P \in \alpha} \{ d(P_0, P) \}.$$

Si consideri la retta r passante per P_0 , perpendicolare ad α e sia H il punto di intersezione di r con α , come illustrato nella Figura 11.2.

Se $P = (x, y, z) \in \alpha$, il triangolo di vertici P_0 , H e P risulta essere rettangolo in H . È noto dalla geometria elementare che, in un triangolo rettangolo, la lunghezza dell'ipotenusa è sempre maggiore o eguale alla lunghezza di ogni suo cateto: in particolare $d(P_0, H) \leq d(P_0, P)$ per ogni $P \in \alpha$. D'altra parte fra i punti che intervengono nella definizione di $d(P_0, \alpha)$ c'è anche H , quindi risulta

$$d(P_0, H) \leq d(P_0, \alpha) = \inf_{P \in \alpha} \{ d(P_0, P) \} \leq d(P_0, H).$$

Ne deduciamo che $d(P_0, \alpha) = d(P_0, H)$: in questo caso l'estremo inferiore è un minimo.

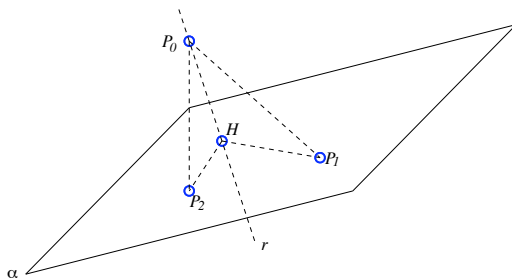


Figura 11.2

Passiamo ora a determinare una formula esplicita per calcolare $d(P_0, \alpha)$; abbiamo appena visto che si tratta di calcolare la minima distanza $d(P_0, H)$ e questo numero può essere facilmente calcolato osservando che coincide con la lunghezza della proiezione del vettore

$$P_0 - P = (x_0 - x)\vec{i} + (y_0 - y)\vec{j} + (z_0 - z)\vec{k}$$

lungo una direzione r ortogonale a α .

Ricordiamo dall'Osservazione 8.10 che, dato un vettore \vec{w} e una retta r per l'origine, la proiezione ortogonale \vec{w}_{\parallel} , di \vec{w} lungo r è

$$\vec{w}_{\parallel} = \frac{\langle \vec{w}, \vec{v} \rangle}{|\vec{v}|^2} \vec{v},$$

ove \vec{v} è un qualsiasi vettore non nullo parallelo a r : dunque

$$|\vec{w}_{\parallel}| = \frac{|\langle \vec{w}, \vec{v} \rangle|}{|\vec{v}|}.$$

Nel nostro caso $a\vec{i} + b\vec{j} + c\vec{k}$ è perpendicolare ad α , perciò

$$\begin{aligned} d(P_0, H) = |P_0 - H| &= \frac{|a(x_0 - x) + b(y_0 - y) + c(z_0 - z)|}{\sqrt{a^2 + b^2 + c^2}} \\ &= \frac{|ax_0 + by_0 + cz_0 - ax - by - cz|}{\sqrt{a^2 + b^2 + c^2}}. \end{aligned}$$

D'altra parte $P \in \alpha$, quindi le sue coordinate soddisfano $ax + by + cz = d$: in conclusione otteniamo

$$d(P_0, \alpha) = d(P_0, H) = \frac{|ax_0 + by_0 + cz_0 - d|}{\sqrt{a^2 + b^2 + c^2}}. \quad (11.2.1)$$

Esempio 11.4. In S_3 sia fissato sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ e si considerino il punto $P_0 = (1, 2, 3)$ e il piano α di equazione cartesiana $3x - 2y - z = 0$. Allora, utilizzando la formula (11.2.1), si ha

$$d(P_0, \alpha) = \frac{|3 - 4 - 3|}{\sqrt{9 + 4 + 1}} = \frac{4}{\sqrt{14}}.$$

Ovviamente $d(P_0, \alpha)$ può anche essere ottenuta come $d(P_0, H) = |H - P_0|$. Nel nostro caso la retta r passante per P_0 e perpendicolare ad α ha equazioni parametriche

$$\begin{cases} x = 1 + 3t \\ y = 2 - 2t \\ z = 3 - t, \end{cases} \quad t \in \mathbb{R}.$$

Quindi $H = (x_H, y_H, z_H)$, punto di intersezione di r con α , corrisponde alla soluzione dell'equazione $3(1 + 3t) - 2(2 - 2t) - (3 - t) = 0$, cioè a $t = 2/7$, in particolare $H = (13/7, 10/7, 19/7)$. Quindi $H - P_0 = (6\vec{i} - 4\vec{j} - 2\vec{k})/7$, da cui segue

$$d(P_0, \alpha) = d(P_0, H) = |H - P_0| = \sqrt{\frac{8}{7}} = \frac{4}{\sqrt{14}}. \quad \spadesuit$$

Come secondo caso consideriamo la minima distanza tra un punto P_0 e una retta r in S_3 , per cui dobbiamo valutare la quantità

$$d(P_0, r) = \inf_{P \in r} \{ d(P_0, P) \}.$$

Sia s la retta passante per P_0 , perpendicolare e incidente a r e sia H il punto di intersezione di r con s . Come nel caso precedente deduciamo che $d(P_0, r) = d(P_0, H)$

e, di nuovo, l'estremo inferiore è un minimo. Anche in questo caso cerchiamo formula per determinare $d(P_0, H)$.

La determinazione della retta s può non essere facile: tale retta deve infatti essere perpendicolare e, simultaneamente, incidente a r . Abbiamo due modi per calcolare s .

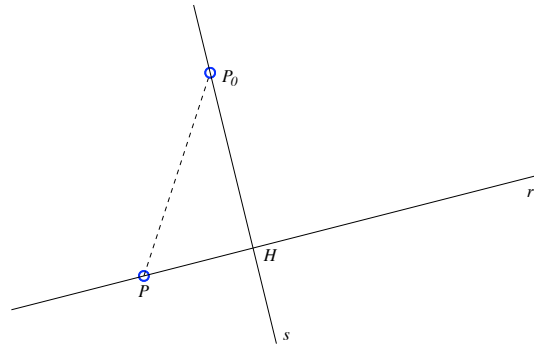


Figura 11.3

Un primo modo consiste nell'osservare che s è sempre contenuta nel piano α passante per P_0 e perpendicolare a r , sicché $r \cap s = r \cap \alpha$. La determinazione di un'equazione di α è, in generale, semplice, così come calcolarne l'intersezione H con r : a questo punto $d(P_0, r) = d(P_0, H)$.

Alternativamente, siano P_1 e P_2 punti arbitrari su r . Osserviamo dalla Figura 11.4 che il segmento $\overline{P_0H}$ è esattamente l'altezza di un parallelogramma avente lato obliquo $\overline{P_0P_1}$ e base $\overline{P_1P_2}$.

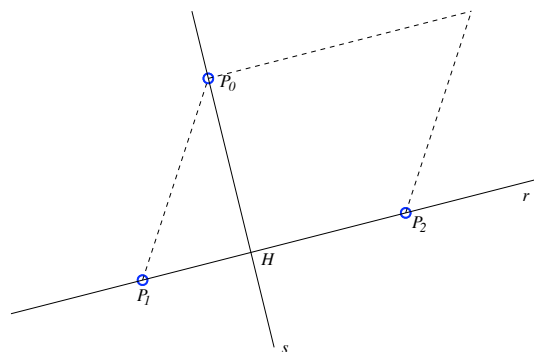


Figura 11.4

Se \mathcal{A} è la misura dell'area di tale parallelogramma, allora

$$d(P_0, r) = d(P_0, H) = |\overline{P_0H}| = \frac{\mathcal{A}}{|\overline{P_1P_2}|}.$$

Ricordiamo quanto visto nell'Osservazione 8.17: se abbiamo un parallelogramma di cui conosciamo tre vertici consecutivi P_0, P_1, P_2 la sua area \mathcal{A} si può determinare con la formula (8.2.2):

$$\mathcal{A} = |(P_0 - P_1) \times (P_1 - P_2)|.$$

Nel nostro caso, se la retta r è data tramite equazioni parametriche

$$\begin{cases} x = x_1 + lt \\ y = y_1 + mt \\ z = z_1 + nt, \end{cases} \quad t \in \mathbb{R},$$

presi ad esempio P_1 e P_2 i punti corrispondenti rispettivamente a $t = 0$ e $t = 1$, cioè $P_1 = (x_1, y_1, z_1)$ e $P_2 = (x_1 + l, y_1 + m, z_1 + n)$, segue che

$$d(P_0, r) = \frac{|((x_0 - x_1)\vec{i} + (y_0 - y_1)\vec{j} + (z_0 - z_1)\vec{k}) \times (l\vec{i} + m\vec{j} + n\vec{k})|}{\sqrt{l^2 + m^2 + n^2}}. \quad (11.2.2)$$

Esempio 11.5. In S_3 sia fissato sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ e si considerino il punto $P_0 = (1, 2, 3)$ e la retta r di equazioni parametriche

$$\begin{cases} x = 3 + t \\ y = t - 2 \\ z = 2t - 3, \end{cases} \quad t \in \mathbb{R}.$$

Un vettore parallelo a r è $\vec{v} = \vec{i} + \vec{j} + 2\vec{k}$, quindi il piano α passante per P_0 e perpendicolare a r ha equazione cartesiana $x + y + 2z = 9$. Il punto d'intersezione $H = (x_H, y_H, z_H)$ di r con α corrisponde alla soluzione dell'equazione

$$(3 + t) + (t - 2) + 2(2t - 3) = 9,$$

cioè $t = 7/3$, sicché $H = (16/3, 1/3, 5/3)$. Risulta $H - P_0 = (13\vec{i} - 5\vec{j} - 4\vec{k})/3$, da cui segue

$$d(P_0, r) = d(P_0, H) = |H - P_0| = \sqrt{\frac{210}{9}}.$$

Calcoliamo di nuovo la stessa minima distanza utilizzando la formula (11.2.2). Poiché $P_0 - P_1 = -2\vec{i} + 4\vec{j} + 6\vec{k}$, si ha

$$d(P_0, r) = \frac{|(-2\vec{i} + 4\vec{j} + 6\vec{k}) \times (\vec{i} + \vec{j} + 2\vec{k})|}{\sqrt{1 + 1 + 4}} = \frac{|2\vec{i} + 10\vec{j} - 6\vec{k}|}{\sqrt{6}} = \sqrt{\frac{140}{6}}. \quad \spadesuit$$

11.3 Distanza di un piano da una retta o da un piano

In questo paragrafo spiegheremo come calcolare la minima distanza $d(X, Y)$ nel caso in cui X sia un piano fissato α . Come al solito, supporremo di aver fissato sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ nello spazio S_3 e che $ax + by + cz = d$ sia l'equazione di α .

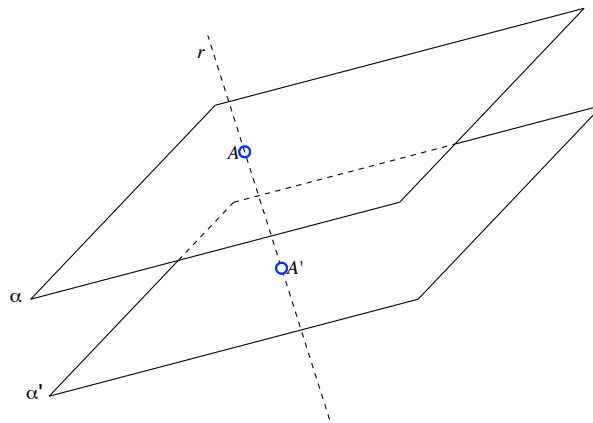


Figura 11.5

Iniziamo a esaminare il caso in cui Y sia un secondo piano α' , di equazione cartesiana $a'x + b'y + c'z = d'$. Ovviamente se $\alpha \nparallel \alpha'$ allora $\alpha \cap \alpha' \neq \emptyset$, dunque $d(\alpha, \alpha') = 0$. Possiamo perciò ridurci a esaminare il caso in cui $\alpha \parallel \alpha'$, come in Figura 11.5. In particolare possiamo supporre $a = a'$, $b = b'$ e $c = c'$, ovvero che α e α' abbiano rispettivamente equazioni

$$ax + by + cz = d, \quad ax + by + cz = d'.$$

Innanzitutto al variare di $A = (x_A, y_A, z_A) \in \alpha$ si ha $ax_A + by_A + cz_A = d$. Dalla formula (11.2.1) deduciamo allora che

$$d(A, \alpha') = \frac{|ax_A + by_A + cz_A - d'|}{\sqrt{a^2 + b^2 + c^2}} = \frac{|d - d'|}{\sqrt{a^2 + b^2 + c^2}}$$

è indipendente da A .

Consideriamo ora i due insiemi

$$D_{\alpha, \alpha'} = \{ d(A, A') \mid A \in \alpha, A' \in \alpha' \}, \quad D_{A, \alpha'} = \{ d(A, A') \mid A' \in \alpha' \}.$$

Chiaramente $D_{A, \alpha'} \subseteq D_{\alpha, \alpha'}$ per ogni $A \in \alpha$, quindi

$$d(\alpha, \alpha') = \inf_{A \in \alpha, A' \in \alpha'} D_{\alpha, \alpha'} \leq \inf_{A \in \alpha} D_{A, \alpha'} = d(A, \alpha') = \frac{|d - d'|}{\sqrt{a^2 + b^2 + c^2}}.$$

Sia ora $H_A \in \alpha'$ la proiezione ortogonale di A . Per quanto visto nel paragrafo precedente sappiamo che

$$\frac{|d - d'|}{\sqrt{a^2 + b^2 + c^2}} = d(A, \alpha') = d(A, H_A) \leq d(A, A'),$$

qualsiasi sia $A' \in \alpha'$, quindi

$$\frac{|d - d'|}{\sqrt{a^2 + b^2 + c^2}} \leq \inf_{A \in \alpha, A' \in \alpha'} D_{\alpha, \alpha'} = d(\alpha, \alpha').$$

Le disuguaglianze dimostrate sopra implicano allora

$$d(\alpha, \alpha') = \frac{|d - d'|}{\sqrt{a^2 + b^2 + c^2}}. \quad (11.3.1)$$

Si noti che la quantità $|d - d'|$ coincide con la minima distanza tra α e α' solamente se $\sqrt{a^2 + b^2 + c^2} = 1$.

Esempio 11.6. In S_3 sia fissato sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ e si considerino i piani α , α' e α'' rispettivamente di equazioni cartesiane

$$\alpha : x + y + z = 1, \quad \alpha' : x - y + 2z = -1, \quad \alpha'' : y - 2z - x = -1.$$

Iniziamo a considerare i due piani α e α' . Poiché

$$\text{rk} \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \end{pmatrix} = 2,$$

segue che $\alpha \not\parallel \alpha'$, dunque $d(\alpha, \alpha') = 0$.

Consideriamo ora i due piani α' e α'' . Poiché

$$\text{rk} \begin{pmatrix} -1 & 1 & -2 \\ 1 & -1 & 2 \end{pmatrix} = 1,$$

segue che $\alpha' \parallel \alpha''$. Per determinare la minima distanza $d(\alpha', \alpha'')$ utilizzando la formula (11.3.1) si devono prima modificare le equazioni date di α' e α'' in modo che abbiano lo stesso primo membro: possiamo allora supporre che le due equazioni siano rispettivamente

$$\alpha' : x - y + 2z = -1, \quad \alpha'' : x - y + 2z = 1,$$

quindi la formula (11.3.1) ci dà

$$d(\alpha', \alpha'') = \frac{|-1 - 1|}{\sqrt{1 + 1 + 4}} = \frac{2}{\sqrt{6}}. \quad \spadesuit$$

In maniera simile si può procedere nel caso in cui si debba calcolare la minima distanza fra una retta r e un piano α . Anche in questo caso o $r \not\parallel \alpha$ e $d(r, \alpha) = 0$ oppure $r \parallel \alpha$ e $d(r, \alpha) = d(P_0, \alpha)$ per un qualsiasi punto $P_0 \in r$. Quindi, in questo secondo caso, si può applicare la formula (11.2.1).

11.4 Distanza fra due rette

Concludiamo il capitolo analizzando l'ultimo caso rimanente, cioè il caso in cui X e Y siano due rette nello spazio, diciamo r ed r' . Di nuovo, supponiamo di aver fissato sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ nello spazio S_3 .

Dobbiamo distinguere due casi principali: il caso in cui r ed r' siano complanari e quello in cui siano rette sghembe

Il primo caso si divide a sua volta in due sottocasi: o le rette sono incidenti in un punto, quindi $r \cap r' \neq \emptyset$ e $d(r, r') = 0$, oppure sono parallele, cioè $r \parallel r'$ e dunque $d(r, r') = d(P_0, r')$ per un qualsiasi punto $P_0 \in r$, come nella Figura 11.6.

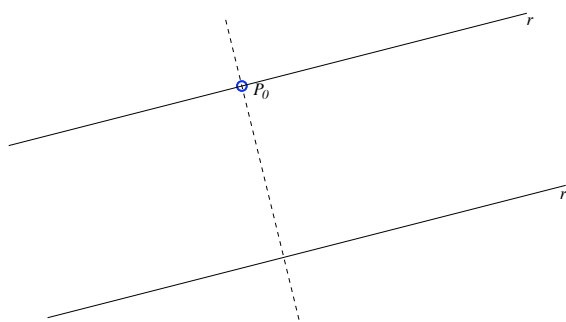


Figura 11.6

Esempio 11.7. Sia fissato un sistema riferimento $O\vec{i}\vec{j}\vec{k}$ in S_3 e si considerino la retta r di equazioni parametriche

$$r : \begin{cases} x = 3 + t \\ y = t - 2 \\ z = 2t - 3, \end{cases} \quad t \in \mathbb{R},$$

la retta r' di equazioni cartesiane

$$r' : \begin{cases} x + y - z = 4 \\ x - y = 5 \end{cases}$$

e il piano α di equazione $3x + y - 2z = 0$.

Il vettore $\vec{v}_r = \vec{i} + \vec{j} + 2\vec{k}$ è parallelo alla retta r , mentre il vettore $\vec{v}_\alpha = 3\vec{i} + \vec{j} - \vec{k}$ è perpendicolare al piano α . Poiché $\langle \vec{v}_r, \vec{v}_\alpha \rangle = 0$ segue che $r \parallel \alpha$. Pertanto, preso $P_0 = (3, -2, -3) \in r$, dalla formula (11.2.1) si ha

$$d(r, \alpha) = d(P_0, \alpha) = \frac{13}{\sqrt{14}}.$$

Prendiamo adesso in considerazione la retta r' : un vettore parallelo a essa è

$$\vec{v}_{r'} = (\vec{i} + \vec{j} - \vec{k}) \times (\vec{i} - \vec{j}) = -\vec{i} - \vec{j} - 2\vec{k} = \vec{v}_r,$$

quindi $r' \parallel r$ e, dunque, $r' \parallel \alpha$. Per calcolare $d(r', \alpha)$ possiamo procedere in due modi: possiamo determinare un punto $P_0 \in r'$, per esempio $P_0 = (5, 0, 1)$ e poi calcolare $d(r', \alpha) = d(P_0, \alpha)$. Oppure possiamo osservare che il piano α' di equazione $3x + y - 2z - 13 = 2(x + y - z - 4) + (x - y - 5) = 0$ contiene r' ed è parallelo a α : perciò

$$d(r', \alpha) = d(\alpha', \alpha) = \frac{|9 - 2 + 6|}{\sqrt{9 + 1 + 4}} = \frac{13}{\sqrt{14}}. \quad \spadesuit$$

Esempio 11.8. In S_3 sia fissato sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ e si considerino le due rette r ed r' rispettivamente di equazioni parametriche

$$r : \begin{cases} x = 3 + t \\ y = t - 2 \\ z = 2t - 3, \end{cases} \quad r' : \begin{cases} x = 1 + t' \\ y = 2 + t' \\ z = 3 + 2t', \end{cases}$$

al variare di $t, t' \in \mathbb{R}$. Le due rette sono entrambe parallele al vettore $\vec{v} = \vec{i} + \vec{j} + 2\vec{k}$, quindi sono parallele fra loro. Poiché $P_0 = (1, 2, 3) \in r'$ segue dall'Esempio 11.5 che

$$d(r, r') = d(r, P_0) = \sqrt{\frac{70}{3}}. \quad \spadesuit$$

Il secondo caso è chiaramente più interessante, e va analizzato in dettaglio: supponiamo quindi che le due rette r e r' siano sghembe, con vettori direzione \vec{v}_r e $\vec{v}_{r'}$ rispettivamente.

Osserviamo che esiste un unico piano α contenente r e parallelo a r' . Infatti poiché $r \not\parallel r'$, segue che $\vec{v}_r \not\parallel \vec{v}_{r'}$, sicché il vettore $\vec{w} = \vec{v}_r \times \vec{v}_{r'}$ è non nullo. Ogni piano contenente r e parallelo a r' deve essere parallelo sia a \vec{v}_r che a $\vec{v}_{r'}$, quindi deve essere perpendicolare a \vec{w} : dovendo intersecare r , tale piano è univocamente determinato. In modo simile si dimostra l'esistenza e l'unicità di un piano α' contenente r' e parallelo a r . Poiché $r \subseteq \alpha$ e $r' \subseteq \alpha'$ si ha $D_{r,r'} \subseteq D_{\alpha,\alpha'}$, quindi vale la disuguaglianza

$$d(\alpha, \alpha') \leq d(r, r'). \quad (11.4.1)$$

La situazione è illustrata nella Figura 11.7.

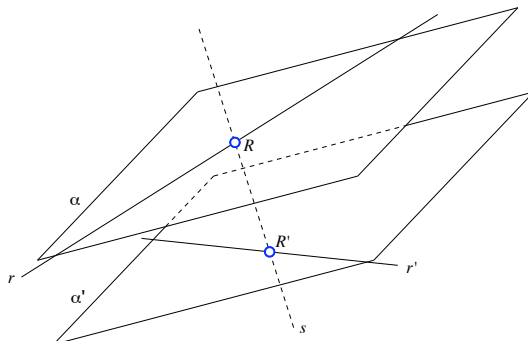


Figura 11.7

Dimostreremo che in realtà nella formula (11.4.1) vale l'uguaglianza, e quindi che il calcolo della minima distanza $d(r, r')$ si riduce al più semplice calcolo della minima distanza fra i due piani paralleli α e α' .

Proposizione 11.9 (Esistenza e unicità della retta di minima distanza).

Siano r ed r' due rette sghembe in S_3 . Esiste un'unica retta s incidente e perpendicolare sia a r che a r' . I punti R ed R' dove essa interseca rispettivamente le rette r ed r' soddisfano la condizione $d(r, r') = d(R, R') = d(\alpha, \alpha')$.

Dimostrazione. Ricordiamo innanzi tutto che α e α' sono ortogonali allo stesso vettore non nullo $\vec{w} = \vec{v}_r \times \vec{v}_{r'}$. In particolare \vec{v}_r , $\vec{v}_{r'}$ e \vec{w} non sono complanari.

Si consideri il piano β contenente r e perpendicolare ad α (e quindi perpendicolare al vettore $\vec{v}_r \times \vec{w}$): per costruzione $r = \alpha \cap \beta$. Similmente si consideri il piano β' contenente r' e perpendicolare ad α' (quindi perpendicolare a $\vec{v}_{r'} \times \vec{w}$): per costruzione $r' = \alpha' \cap \beta'$.

Se i piani β e β' fossero paralleli, tali sarebbero $\vec{v}_r \times \vec{w}$ e $\vec{v}_{r'} \times \vec{w}$. Il primo vettore è perpendicolare al piano individuato da \vec{v}_r e \vec{w} , il secondo al piano individuato da $\vec{v}_{r'}$ e \vec{w} . Deduciamo allora che \vec{v}_r , $\vec{v}_{r'}$ e \vec{w} dovrebbero essere complanari, in contraddizione con quanto osservato all'inizio. Quindi β e β' non sono paralleli, pertanto $s = \beta \cap \beta'$ è una retta.

Si noti che i piani β e β' sono ortogonali sia ad α che ad α' , quindi anche s lo è. In particolare $s \subseteq \beta$ è ortogonale a $r = \alpha \cap \beta \subseteq \beta$, dunque interseca r in un unico punto R . Similmente s interseca R' in un unico punto R' . Risulta allora che R' è la proiezione ortogonale di R su α' , quindi

$$d(\alpha, \alpha') = d(R, R') = d(R, r') \geq d(r, r') \geq d(\alpha, \alpha').$$

Pertanto

$$d(R, R') = d(\alpha, \alpha') = d(r, r').$$

Quanto visto dimostra l'esistenza della retta s descritta nella tesi.

Per quanto riguarda la sua unicità, notiamo che se u è una qualsiasi retta avente le proprietà indicate nella tesi, allora u è parallela al vettore \vec{w} , quindi ortogonale sia ad α che ad α' . In particolare u deve essere contenuta nei piani β e β' , quindi deve coincidere con la loro intersezione che è s . \square

La retta s la cui esistenza e unicità è stata dimostrata nella Proposizione 11.9 viene detta *retta di minima distanza fra r ed r'* . Si noti che anche in questo caso l'estremo inferiore che definisce la minima distanza $d(r, r')$ è un minimo.

Esempio 11.10. In S_3 sia fissato sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ e si considerino le rette r ed r' rispettivamente di equazioni parametriche

$$r : \begin{cases} x = 3 + t \\ y = t - 2 \\ z = 2t - 3, \end{cases} \quad r' : \begin{cases} x = -t' \\ y = t' - 2 \\ z = 2t' + 1, \end{cases}$$

al variare di $t, t' \in \mathbb{R}$.

La retta r è parallela al vettore $\vec{v}_r = \vec{i} + \vec{j} + 2\vec{k}$, mentre la retta r' è parallela al vettore $\vec{v}_{r'} = -\vec{i} + \vec{j} + 2\vec{k}$, quindi $r \nparallel r'$. Inoltre il sistema

$$\begin{cases} 3 + t = -t' \\ t - 2 = t' - 2 \\ 2t - 3 = 2t' + 1, \end{cases}$$

non ha soluzione. Concludiamo che r ed r' sono sghembe.

Vogliamo calcolare la minima distanza $d(r, r')$: cerchiamo quindi i due piani paralleli α e α' . Essi sono entrambi ortogonali al vettore

$$\vec{v}_r \times \vec{v}_{r'} = -2\vec{j} + \vec{k},$$

quindi hanno equazione

$$\alpha : -2y + z = d, \quad \alpha' : -2y + z = d'.$$

Imponendo il passaggio di α per un punto di r , ad esempio $(1, 0, 1)$, troviamo $d = 1$; similmente, imponendo il passaggio di α' per un punto di r' , ad esempio $(-1, -1, 3)$, troviamo $d' = 5$.

La minima distanza tra le due rette è dunque

$$d(r, r') = d(\alpha, \alpha') = \frac{|1 - 5|}{\sqrt{0 + 4 + 1}} = \frac{4}{\sqrt{5}}.$$

Osserviamo che un'altra possibilità per calcolare la minima distanza tra due rette sghembe è quella di determinare esplicitamente i punti $R \in r$ ed $R' \in r'$ della Proposizione 11.9: per fare ciò dobbiamo imporre che valga sia $R - R' \perp \vec{v}_r$ sia $R - R' \perp \vec{v}_{r'}$. Abbiamo

$$R - R' = (3 + t + t')\vec{i} + (t - t')\vec{j} + (2t - 2t' - 4)\vec{k},$$

dunque le due condizioni di perpendicolarità si traducono nel sistema

$$\begin{cases} 6t - 4t' = 5 \\ 4t - 6t' = 11, \end{cases}$$

la cui unica soluzione è data da $t = -7/10$ e $t' = -23/10$ corrispondente ai due punti

$$R = \left(\frac{23}{10}, -\frac{27}{10}, -\frac{44}{10} \right), \quad R' = \left(\frac{23}{10}, -\frac{43}{10}, -\frac{36}{10} \right).$$

Segue che

$$d(r, r') = d(R, R') = \frac{\sqrt{0 + 256 + 64}}{10} = \frac{4}{\sqrt{5}}.$$

Per completezza, osserviamo che la retta s di minima distanza fra r ed r' , che è la retta passante per R e R' , ha equazioni parametriche

$$\begin{cases} x = 23/10 \\ y = -27/10 - 16t \\ z = -44/10 + 8t, \end{cases} \quad t \in \mathbb{R}. \quad \spadesuit$$

Spazi vettoriali e loro sottospazi

12.1 Spazi vettoriali

Da questo capitolo \mathbb{K} denoterà sempre un campo, cioè un qualsiasi insieme non vuoto in cui siano definite due operazioni binarie chiamate somma e prodotto che abbiano le stesse proprietà formali della somma o prodotto dei numeri reali e complessi (commutatività, associatività, distributività, esistenza di un elemento neutro, esistenza degli opposti, esistenza degli inversi di elementi non nulli): per esempio $\mathbb{K} = \mathbb{R}, \mathbb{C}$.

La nozione di *spazio vettoriale* generalizza quanto visto nelle lezioni precedenti: l'insieme $\mathbb{K}^{m,n}$ delle matrici $m \times n$ a coefficienti in \mathbb{K} (paragrafo 1.3), $V_2(O)$ e $V_3(O)$ dei vettori geometrici applicati del piano e dello spazio (paragrafo 7.4).

Definizione 12.1 (Spazi vettoriali).

Uno *spazio vettoriale su un campo* \mathbb{K} è un insieme non vuoto V munito di due applicazioni

$$\begin{aligned} s_V: V \times V &\longrightarrow V \\ (v_1, v_2) &\longrightarrow v_1 + v_2 = s_V(v_1, v_2), \end{aligned}$$

$$\begin{aligned} p_V: \mathbb{K} \times V &\longrightarrow V \\ (\alpha, v) &\longrightarrow \alpha v = p_V(\alpha, v), \end{aligned}$$

dette rispettivamente *somma* e *prodotto*, per cui valgono i seguenti assiomi:

- (S1) per ogni $v_1, v_2 \in V$, si ha $v_1 + v_2 = v_2 + v_1$ (*la somma è commutativa*);
- (S2) per ogni $v_1, v_2, v_3 \in V$, si ha $v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3$ (*la somma è associativa*);
- (S3) esiste un elemento neutro per la somma, cioè un elemento $0_V \in V$ tale che $0_V + v = v$, per ogni $v \in V$;
- (S4) per ogni $v \in V$ esiste un elemento opposto di v , cioè $w \in V$ tale che $v + w = 0_V$;
- (P1) per ogni $v \in V$, si ha $1v = v$;
- (P2) per ogni $\alpha_1, \alpha_2 \in \mathbb{K}$ e $v \in V$, si ha $\alpha_1(\alpha_2 v) = (\alpha_1 \alpha_2)v$;

(SP1) per ogni $\alpha_1, \alpha_2 \in \mathbb{K}$ e $v \in V$, si ha $(\alpha_1 + \alpha_2)v = \alpha_1v + \alpha_2v$;

(SP2) per ogni $\alpha \in \mathbb{K}$ e $v_1, v_2 \in V$, si ha $\alpha(v_1 + v_2) = \alpha v_1 + \alpha v_2$.

Gli elementi di V vengono detti *vettori*, quelli di \mathbb{K} *scalari*.

In uno spazio vettoriale vale la seguente legge di annullamento del prodotto.

Proposizione 12.2 (Legge di annullamento del prodotto per scalare).

Sia V uno spazio vettoriale su un campo \mathbb{K} . Dati $\alpha \in \mathbb{K}$ e $v \in V$, allora $\alpha v = 0_V$ se e solo se è soddisfatta almeno una delle condizioni $\alpha = 0$, $v = 0_V$.

Dimostrazione. Se $\alpha = 0$ allora $0 + 0 = 0$: moltiplicando ambo i membri per v e sfruttando la proprietà (SP2) si ha $0v + 0v = 0v$. Sia $w \in V$ tale che $0v + w = 0_V$ (ne esiste almeno uno da (S4)). Allora

$$0v = 0v + (0v + w) = (0v + 0v) + w = 0v + w = 0_V.$$

Per dimostrare $\alpha 0_V = 0_V$ si procede nello stesso modo. Viceversa, supponiamo $\alpha v = 0_V$. Se $\alpha \neq 0$ allora

$$v = 1v = (\alpha^{-1}\alpha)v = \alpha^{-1}(\alpha v) = \alpha^{-1}0_V = 0_V,$$

che conclude la dimostrazione. □

Osservazione 12.3. In uno spazio vettoriale V su \mathbb{K} l'elemento neutro rispetto alla somma è unico. Infatti se $0'_V$ fosse un altro elemento neutro, dall'assioma (S3) seguirebbero le due relazioni

$$0_V + 0'_V = 0'_V, \quad 0'_V + 0_V = 0_V;$$

poiché la somma è commutativa per l'assioma (S1), segue che $0'_V = 0_V$.

Similmente, l'elemento opposto w indicato nell'assioma (S4) è unico e coincide con $(-1)v$: infatti dalla relazione $0 = 1 + (-1)$ abbiamo che

$$0_V = 0v = (1 + (-1))v = v + (-1)v,$$

cioè $(-1)v$ è un opposto di v . Se w fosse un altro elemento con la stessa proprietà, da (S4) seguirebbe $v + w = 0_V = v + (-1)v$ e dunque

$$w = w + 0_V = w + (v + (-1)v) = (w + v) + (-1)v = 0_V + (-1)v = (-1)v.$$

D'ora innanzi indicheremo l'unico opposto di $v \in V$ con il simbolo $-v$.

⚠ Della definizione di spazio vettoriale sono parte integrante sia le due operazioni di somma e prodotto, sia il campo di definizione, ovvero l'insieme degli scalari. A uno stesso insieme possiamo dare strutture di spazio vettoriale modificando una o entrambe le operazioni, o modificando il campo degli scalari. Per esempio osserviamo che \mathbb{R} è spazio vettoriale su \mathbb{R} con le operazioni di somma e prodotto che in esso sono definite. In modo simile \mathbb{C} è uno spazio vettoriale su \mathbb{C} . Notiamo però che se ci restringiamo a considerare solo il prodotto di numeri complessi per numeri reali, \mathbb{C} viene ad essere uno spazio vettoriale su \mathbb{R} .

Procediamo ora a dare alcuni esempi di spazi vettoriali.

Esempio 12.4. Si consideri l'insieme $\mathbb{K}^{m,n}$ delle matrici $m \times n$ a coefficienti in \mathbb{K} . Nella Lezione 1 abbiamo definito delle operazioni di somma e prodotto per uno scalare in tale insieme: precisamente se $\alpha \in \mathbb{K}$ e

$$A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}, \quad B = (b_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}},$$

abbiamo posto

$$\begin{aligned} (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} + (b_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} &= (a_{i,j} + b_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}, \\ \alpha (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} &= (\alpha a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}. \end{aligned}$$

Nelle Proposizioni 1.20 e 1.23 abbiamo mostrato che esse soddisfano tutti gli otto assiomi della Definizione 12.1, dunque $\mathbb{K}^{m,n}$ con tali operazioni è uno spazio vettoriale su \mathbb{K} . ♠

Esempio 12.5. Consideriamo l'insieme $V_2(O)$ dei vettori geometrici nel piano. In un precedente capitolo abbiamo definito per via geometrica la somma e il prodotto per uno scalare reale. Poi nella Proposizione 7.11 abbiamo visto che tali operazioni soddisfano tutti gli otto assiomi della Definizione 12.1, dunque $V_2(O)$ con tali operazioni è uno spazio vettoriale su \mathbb{R} .

La stessa cosa vale per i vettori geometrici nello spazio, $V_3(O)$, e le operazioni di somma e prodotto su di essi definiti in precedenza. ♠

Esempio 12.6. Sia $I \subseteq \mathbb{R}$ un intervallo non vuoto e consideriamo l'insieme \mathbb{R}^I delle funzioni a valori reali $f: I \rightarrow \mathbb{R}$. Osserviamo subito che $\mathbb{R}^I \neq \emptyset$: infatti la funzione costantemente nulla $0_I: I \rightarrow \mathbb{R}$ definita da $0_I(x) = 0$ per ogni $x \in I$ vi appartiene.

Introduciamo delle operazioni di somma e prodotto come segue. Se $\varphi, \psi \in \mathbb{R}^I$, poniamo

$$\begin{aligned} \varphi + \psi: I &\longrightarrow \mathbb{R} \\ x &\longrightarrow (\varphi + \psi)(x) = \varphi(x) + \psi(x). \end{aligned}$$

Se $\varphi \in \mathbb{R}^I$ e $\alpha \in \mathbb{R}$, poniamo

$$\begin{aligned} \alpha\varphi: I &\longrightarrow \mathbb{R} \\ x &\longrightarrow (\alpha\varphi)(x) = \alpha\varphi(x). \end{aligned}$$

Per verificare che con queste operazioni \mathbb{R}^I è uno spazio vettoriale su \mathbb{R} bisogna verificare che esse soddisfano gli assiomi della Definizione 12.1. Cominciamo con l'assioma (S1): date $\varphi, \psi \in \mathbb{R}^I$, poiché per ogni $x \in I$ i numeri reali $\varphi(x) + \psi(x)$ e $\psi(x) + \varphi(x)$ coincidono, segue, per definizione, che $\varphi + \psi = \psi + \varphi$.

Coloro i quali, in modo analogo, verificheranno gli altri assiomi della Definizione 12.1, si renderanno conto che tutto dipende dal fatto che il codominio delle funzioni considerate è esso stesso uno spazio vettoriale. Più in generale è vero che se X è un insieme e V è uno spazio vettoriale su \mathbb{K} , allora l'insieme V^X delle funzioni $\varphi: X \rightarrow V$ a valori in V , dotato delle operazioni naturali di somma di funzioni e prodotto di una funzione per scalare è, a sua volta, uno spazio vettoriale su \mathbb{K} . ♠

Esempio 12.7. Indichiamo con \mathbb{K}^n l'insieme delle n -uple ordinate di elementi di un campo \mathbb{K} : un elemento di \mathbb{K}^n è quindi una sequenza ordinata (x_1, x_2, \dots, x_n) di n elementi $x_1, x_2, \dots, x_n \in \mathbb{K}$. Possiamo definire due operazioni di somma e prodotto:

$$\begin{aligned}(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) &= (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n), \\ \alpha(x_1, x_2, \dots, x_n) &= (\alpha x_1, \alpha x_2, \dots, \alpha x_n),\end{aligned}$$

per ogni scelta di $(x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n) \in \mathbb{K}^n$ e $\alpha \in \mathbb{K}$.

Osserviamo che gli elementi di \mathbb{K}^n si possono identificare naturalmente con le matrici in $\mathbb{K}^{n,1}$ e, in tale identificazione, le operazioni introdotte in \mathbb{K}^n vengono a corrispondere con le operazioni di somma di matrici e di prodotto di una matrice per uno scalare definite in $\mathbb{K}^{n,1}$. Segue che \mathbb{K}^n è uno spazio vettoriale.

Per questo motivo d'ora innanzi non distingueremo più tra \mathbb{K}^n e $\mathbb{K}^{n,1}$; quindi scriveremo (x_1, \dots, x_n) per indicare la matrice colonna

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

In particolare \mathbb{K}^1 è identificato con $\mathbb{K}^{1,1} = \mathbb{K}$. ♠

12.2 Sottospazi vettoriali

Definizione 12.8 (Sottospazi vettoriali).

Sia V uno spazio vettoriale su un campo \mathbb{K} . Un sottoinsieme $W \subseteq V$ si dice *sottospazio vettoriale di V* o, brevemente, *sottospazio di V* se esso, munito delle due applicazioni ristrette $(s_V)|_{W \times W}$ e $(p_V)|_{\mathbb{K} \times W}$, è uno spazio vettoriale su \mathbb{K} .

Analizziamo la definizione di sottospazio: una prima osservazione è che deve valere $W \neq \emptyset$. Inoltre vogliamo che le applicazioni $(s_V)|_{W \times W}$ e $(p_V)|_{\mathbb{K} \times W}$ rendano W uno spazio vettoriale su \mathbb{K} . Poiché, a priori, si ha che

$$(s_V)|_{W \times W}: W \times W \longrightarrow V, \quad (p_V)|_{\mathbb{K} \times W}: \mathbb{K} \times W \longrightarrow V,$$

ciò significa che tali applicazioni devono avere di fatto immagine contenuta in W , ovvero la somma di due elementi di W deve essere ancora un elemento di W (un sottoinsieme con tale proprietà si dice *chiuso rispetto alla somma*) e il prodotto di uno scalare per un elemento di W deve essere anch'esso un elemento di W (un sottoinsieme con tale proprietà si dice *chiuso rispetto al prodotto per scalari*).

Supponiamo che ciò accada. Poiché gli assiomi (S1), (S2), (P1), (P2), (SP1), (SP2) della Definizione 12.1 valgono per ogni scelta dei vettori $v, v_1, v_2, v_3 \in V$ e degli scalari $\alpha, \alpha_1, \alpha_2 \in \mathbb{K}$, necessariamente esse valgono anche se ci restringiamo ai soli elementi in W .

Rimangono da verificare gli assiomi (S3) ed (S4). Per verificare (S3) si osservi che W è non vuoto, quindi contiene almeno un vettore w : poiché W è chiuso rispetto al prodotto per scalari si ha anche che $0_V = 0 w \in W$, quindi in W c'è un elemento neutro per la somma (lo stesso di V). Infine, per verificare (S4), si osservi che se $w \in W \subseteq V$, dall'Osservazione 12.3 segue $-w = (-1) w \in W$, ovvero se W contiene un elemento, allora contiene anche il suo opposto.

Le considerazioni appena fatte dimostrano il seguente criterio per verificare se un sottoinsieme è un sottospazio o meno.

Proposizione 12.9.

Sia V uno spazio vettoriale su un campo \mathbb{K} . Dato un sottoinsieme $W \subseteq V$, allora W è sottospazio vettoriale di V se e solo se è non vuoto ed è chiuso rispetto alla somma e al prodotto per scalari definiti in V .

Se W è sottospazio di V allora, come visto sopra, $0_V \in W$ e W è chiuso rispetto alla somma e al prodotto per scalari definiti in V . Viceversa se $0_V \in W$ e W è chiuso rispetto alla somma e al prodotto per scalari definiti in V , allora W è sottospazio di V per il precedente criterio. Questo dimostra il seguente criterio, simile al precedente.

Proposizione 12.10.

Sia V uno spazio vettoriale su \mathbb{K} . Dato un sottoinsieme $W \subseteq V$, allora W è sottospazio di V se e solo se $0_V \in W$ ed è chiuso rispetto alla somma e al prodotto per scalari definiti in V .

Diamo ora alcuni esempi di sottospazi, utilizzando i criteri enunciati sopra.

Esempio 12.11. In \mathbb{R}^2 si consideri l'insieme

$$W_1 = \{ (x, y) \in \mathbb{R}^2 \mid y = 0 \}.$$

Gli elementi di W_1 sono tutte e sole le coppie (x, y) con $y = 0$ e $x \in \mathbb{R}$. In particolare, se si identifica \mathbb{R}^2 con il piano ordinario con fissato un sistema di riferimento $O\vec{i}\vec{j}$, l'insieme W_1 rappresenta l'asse delle ascisse.

Si noti che $0_{\mathbb{R}^2} = (0, 0) \in W_1$. Siano poi $(x', 0), (x'', 0) \in W_1$: allora

$$(x', 0) + (x'', 0) = (x' + x'', 0 + 0) = (x' + x'', 0) \in W_1.$$

Infine siano $\alpha \in \mathbb{R}$ e $(x, 0) \in W_1$: allora

$$\alpha(x, 0) = (\alpha x, \alpha 0) = (\alpha x, 0) \in W_1.$$

Concludiamo che W_1 è un sottospazio di \mathbb{R}^2 .

Sempre in \mathbb{R}^2 si consideri ora l'insieme

$$W_2 = \{ (x, y) \in \mathbb{R}^2 \mid x + y = 0 \}.$$

Gli elementi di W_2 sono tutte e sole le coppie del tipo $(x, -x)$ con $x \in \mathbb{R}$. Se si identifica \mathbb{R}^2 con il piano ordinario con fissato un sistema di riferimento $O\vec{i}\vec{j}$, l'insieme W_2 rappresenta la bisettrice del secondo e quarto quadrante.

In particolare $0_{\mathbb{R}^2} = (0, 0) = (0, -0) \in W_2$. Siano poi $(x', -x'), (x'', -x'') \in W_2$: allora $(x', -x') + (x'', -x'') = (x' + x'', -x' - x'') = (x' + x'', -(x' + x'')) \in W_2$. Infine siano $\alpha \in \mathbb{R}$ e $(x, -x) \in W_2$: allora $\alpha(x, -x) = (\alpha x, \alpha(-x)) = (\alpha x, -(\alpha x)) \in W_2$. Concludiamo che W_2 è un sottospazio di \mathbb{R}^2 .

Siano adesso $a, b \in \mathbb{R}$. In \mathbb{R}^2 si consideri l'insieme

$$W_{a,b} = \{ (x, y) \in \mathbb{R}^2 \mid ax + by = 0 \}.$$

Se si identifica \mathbb{R}^2 con il piano ordinario con fissato un sistema di riferimento $O\vec{i}\vec{j}$, l'insieme $W_{a,b}$ rappresenta o l'intero piano oppure una retta per l'origine.

Si ha $a \cdot 0 + b \cdot 0 = 0$, dunque $0_{\mathbb{R}^2} = (0, 0) \in W_{a,b}$. Se poi $(x', y'), (x'', y'') \in W_{a,b}$, allora $ax' + by' = ax'' + by'' = 0$, dunque

$$a(x' + x'') + b(y' + y'') = ax' + ax'' + by' + by'' = (ax' + by') + (ax'' + by'') = 0 + 0 = 0,$$

cioè $(x', y') + (x'', y'') = (x' + x'', y' + y'')$ è un elemento di $W_{a,b}$.

Infine siano $\alpha \in \mathbb{R}$ e $(x, y) \in W_{a,b}$: allora $ax + by = 0$ e

$$a(\alpha x) + b(\alpha y) = \alpha ax + \alpha by = \alpha(ax + by) = \alpha \cdot 0 = 0,$$

cioè $\alpha(x, y) = (\alpha x, \alpha y)$ è anch'esso un elemento di $W_{a,b}$.

Concludiamo che $W_{a,b}$ è un sottospazio di \mathbb{R}^2 . Si noti che i sottospazi W_1 e W_2 sono casi particolari di questo esempio: precisamente W_1 si ottiene considerando $a = 1, b = 0$, W_2 considerando $a = b = 1$. ♠

Esempio 12.12. Sia $A \in \mathbb{K}^{m,n}$ una matrice, \mathbb{K} ; in $\mathbb{K}^{n,p}$ si consideri l'insieme

$$W = \{ X \in \mathbb{K}^{n,p} \mid AX = 0_{m,p} \}.$$

Chiaramente la matrice nulla $0_{n,p} \in \mathbb{K}^{n,p}$ appartiene a W . Siano poi $X', X'' \in W$: ciò significa che $AX' = AX'' = 0_{m,p}$. Allora

$$A(X' + X'') = AX' + AX'' = 0_{m,p} + 0_{m,p} = 0_{m,p},$$

cioè $X' + X'' \in W$. Siano infine $\alpha \in \mathbb{K}$ e $X \in W$: ciò significa che $AX = 0_{m,p}$. Allora

$$A(\alpha X) = \alpha AX = \alpha 0_{m,p} = 0_{m,p},$$

cioè $\alpha X \in W$. Concludiamo che W è un sottospazio di \mathbb{R}^2 .

Si noti che il sottospazio $W_{a,b}$ dell'Esempio 12.11 (e quindi anche W_1 e W_2) è un caso particolare di questo esempio: infatti siamo nel caso $m = p = 1, n = 2, A = (a \ b)$.

Abbiamo quindi dimostrato che l'insieme delle soluzioni di una qualsiasi equazione matriciale omogenea è un sottospazio di $\mathbb{K}^{n,p}$: vedremo più avanti che vale anche il viceversa, ovvero che ogni sottospazio di $\mathbb{K}^{n,p}$ è insieme delle soluzioni di una qualche (sempre più d'una) equazione matriciale omogenea. ♠

Esempio 12.13. Sia $I \subseteq \mathbb{R}$ un intervallo non vuoto che, per comodità, supponiamo aperto, diciamo $I =]a, b[$. Nell'Esempio 12.6 abbiamo verificato che l'insieme \mathbb{R}^I delle funzioni $\varphi: I \rightarrow \mathbb{R}$ è uno spazio vettoriale.

Si consideri adesso il sottoinsieme

$$\mathcal{C}^0(I) = \{ \varphi \in \mathbb{R}^I \mid \varphi \text{ continua in tutto } I \}.$$

Dal momento che la funzione identicamente nulla $0_I: I \rightarrow \mathbb{R}$ è continua (come tutte le funzioni costanti), $0_I \in \mathcal{C}^0(I)$. Inoltre se $\varphi, \psi \in \mathcal{C}^0(I)$ la loro somma $\varphi + \psi$ è ancora in $\mathcal{C}^0(I)$. Infine se $\varphi \in \mathcal{C}^0(I)$ e $\alpha \in \mathbb{R}$ allora anche $\alpha\varphi$ è in $\mathcal{C}^0(I)$. Concludiamo che $\mathcal{C}^0(I)$ è un sottospazio di \mathbb{R}^I .

Più in generale, per ogni $p \geq 1$, consideriamo l'insieme

$$\mathcal{C}^p(I) = \{ \varphi \in \mathbb{R}^I \mid \varphi \text{ derivabile } p \text{ volte in tutto } I \text{ con } \varphi^{(p)} \in \mathcal{C}^0(I) \}.$$

Da quanto visto nel corso di Analisi Matematica I segue che $\mathcal{C}^p(I)$ è sottospazio di \mathbb{R}^I per ogni p . Si noti che $\mathcal{C}^p(I)$ è sottospazio di ogni $\mathcal{C}^q(I)$ con $0 \leq q \leq p$. ♠

Esempio 12.14. Siano $I \subseteq \mathbb{R}$ un intervallo aperto non vuoto, $x_0 \in I$ e $\psi \in \mathbb{R}^I$ non nulla in $I \setminus \{x_0\}$. Si consideri il sottoinsieme

$$\mathcal{F}_{x_0, \psi} = \left\{ \varphi \in \mathbb{R}^I \mid \lim_{x \rightarrow x_0} \frac{\varphi(x)}{\psi(x)} \in \mathbb{R} \right\}.$$

Il lettore dimostri che $\mathcal{F}_{x_0, \psi}$ è un sottospazio di \mathbb{R}^I con le solite operazioni di somma tra funzioni e prodotto per uno scalare. ♠

Diamo ora alcuni esempi di insiemi che non sono sottospazi: in particolare mostreremo che possono valere due delle tre proprietà, senza che necessariamente debba valere anche la terza.

Esempio 12.15. Siano $a, b, c \in \mathbb{R}$ con $c \neq 0$. In \mathbb{R}^2 si consideri l'insieme

$$W = \{ (x, y) \in \mathbb{R}^2 \mid ax + by = c \}.$$

Se si identifica \mathbb{R}^2 con il piano ordinario con fissato un sistema di riferimento $O\vec{i}\vec{j}$, l'insieme W o è vuoto oppure rappresenta una retta non passante per l'origine.

Chiaramente $a \cdot 0 + b \cdot 0 + c = c \neq 0$, dunque $0_{\mathbb{R}^2} = (0, 0) \notin W$: in particolare W non è un sottospazio di \mathbb{R}^2 in forza della Proposizione 12.10.

Più in generale se $A \in \mathbb{K}^{m,n}$ e $B \in \mathbb{K}^{m,p} \setminus \{0_{m,p}\}$, allora l'insieme

$$W = \{ X \in \mathbb{K}^{n,p} \mid AX = B \}$$

non può essere sottospazio di $\mathbb{K}^{n,p}$ perché non contiene mai $0_{n,p}$.

Ricordiamo però che se l'equazione matriciale $AX = B$ ha almeno una soluzione X_0 , allora l'insieme delle sue soluzioni è

$$\{ X_0 + Y \mid AY = 0_{m,p} \}.$$

Quindi, pur non essendo W un sottospazio, può essere ottenuto "traslando" un sottospazio (l'insieme delle soluzioni dell'equazione matriciale omogenea associata) di una stessa soluzione fissata X_0 . ♠

Esempio 12.16. In \mathbb{R}^2 si consideri l'insieme

$$W = \{ (x, y) \in \mathbb{R}^2 \mid xy = 0 \}.$$

Gli elementi di W sono tutte e sole le coppie aventi almeno una componente nulla. Se si identifica \mathbb{R}^2 con il piano ordinario con fissato un sistema di riferimento $O\vec{i}\vec{j}$, l'insieme W rappresenta l'unione degli assi coordinati.

Chiaramente $00 = 0$, dunque $0_{\mathbb{R}^2} = (0, 0) \in W$. Siano quindi $\alpha \in \mathbb{R}$ e $(x, y) \in W$: allora $xy = 0$, dunque

$$(\alpha x)(\alpha y) = \alpha^2 xy = \alpha^2 0 = 0,$$

cioè $\alpha(x, y) = (\alpha x, \alpha y)$ è anch'esso un elemento di W . Nonostante ciò, W non è un sottospazio perché non è chiuso rispetto alla somma. Infatti $(1, 0)$ e $(0, 1) \in W$, ma $(1, 1) = (1, 0) + (0, 1) \notin W$. ♠

Esempio 12.17. In \mathbb{R}^2 si consideri l'insieme

$$W = \{ (x, y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0 \}.$$

Gli elementi di W sono tutte e sole le coppie aventi entrambe le componenti non negative. Se si identifica \mathbb{R}^2 con il piano ordinario con fissato un sistema di riferimento $O\vec{i}\vec{j}$, allora l'insieme W rappresenta il primo quadrante.

Chiaramente $0 \geq 0$, dunque $0_{\mathbb{R}^2} = (0, 0) \in W$. Se poi $(x', y'), (x'', y'') \in W$, allora $x', y', x'', y'' \geq 0$, dunque $x' + x'', y' + y'' \geq 0$, cioè $(x', y') + (x'', y'') = (x' + x'', y' + y'')$ è un elemento di W .

Nonostante ciò, W non è un sottospazio perché non è chiuso rispetto al prodotto per scalari: infatti $(1, 1) \in W$, ma $-1(1, 1) = (-1, -1) \notin W$. ♠

Osservazione 12.18. Se V è uno spazio vettoriale su \mathbb{K} allora esso contiene come sottoinsiemi sia se stesso sia $\{0_V\}$. Chiaramente entrambi questi sottoinsiemi sono sottospazi di V . Tali sottospazi, che ci sono sempre, vengono detti *sottospazi banali di V* .

12.3 Alcune operazioni notevoli fra sottospazi

Sia V uno spazio vettoriale su un campo \mathbb{K} e siano W' e W'' due suoi sottospazi. Ci domandiamo se, tramite le operazioni insiemistiche usuali (differenza, intersezione, unione), si ottengono ancora sottospazi.

Una prima osservazione più o meno banale è che la differenza $W' \setminus W''$ non è mai un sottospazio. Infatti $0_V \in W', W''$ dunque $0_V \notin W' \setminus W''$.

Il caso dell'intersezione $W' \cap W''$ è, invece, diverso.

Proposizione 12.19.

Sia V uno spazio vettoriale su \mathbb{K} . Dati due sottospazi $W', W'' \subseteq V$, allora $W' \cap W''$ è un sottospazio di V .

Dimostrazione. Poiché $0_V \in W', W''$ si ha $0_V \in W' \cap W''$.

Siano ora $w_1, w_2 \in W' \cap W''$: ciò significa che $w_1, w_2 \in W'$, dunque $w_1 + w_2 \in W'$ (perché W' è un sottospazio), e $w_1, w_2 \in W''$, dunque $w_1 + w_2 \in W''$ (perché anche W'' è un sottospazio), sicché $w_1 + w_2 \in W' \cap W''$ ovvero $W' \cap W''$ è chiuso rispetto alla somma.

Siano infine $\alpha \in \mathbb{K}$ e $w \in W' \cap W''$: ciò significa che $w \in W'$, dunque $\alpha w \in W'$ (perché W' è un sottospazio), e $w \in W''$, dunque $\alpha w \in W''$ (perché anche W'' è un sottospazio), sicché $\alpha w \in W' \cap W''$ ovvero $W' \cap W''$ è chiuso rispetto al prodotto per scalari.

Concludiamo che $W' \cap W''$ è un sottospazio di V in virtù della Proposizione 12.10. \square

Esempio 12.20. Siano $a', b', c', a'', b'', c'' \in \mathbb{R}$. In \mathbb{R}^3 siano dati i sottoinsiemi

$$\begin{aligned} W' &= \{ (x, y, z) \in \mathbb{R}^3 \mid a'x + b'y + c'z = 0 \}, \\ W'' &= \{ (x, y, z) \in \mathbb{R}^3 \mid a''x + b''y + c''z = 0 \}. \end{aligned}$$

Il lettore verifichi per esercizio che W' e W'' sono sottospazi.

L'intersezione $W' \cap W''$ è l'insieme delle terne $(x, y, z) \in \mathbb{R}^3$ che appartengono sia a W' che a W'' , quindi delle terne $(x, y, z) \in \mathbb{R}^3$ che soddisfano simultaneamente le equazioni $a'x + b'y + c'z = a''x + b''y + c''z = 0$, ovvero $W' \cap W''$ è l'insieme delle soluzioni del sistema omogeneo

$$\begin{cases} a'x + b'y + c'z = 0 \\ a''x + b''y + c''z = 0. \end{cases}$$

Si noti che tale insieme è un sottospazio di \mathbb{R}^3 sia per la Proposizione 12.19, perché è intersezione di sottospazi, sia dall'Esempio 12.12. \spadesuit

La Proposizione 12.19 vale per l'intersezione di un numero qualsiasi (anche infinito) di sottospazi, come spiegato di seguito.

Proposizione 12.21.

Sia V uno spazio vettoriale su un campo \mathbb{K} . Dati sottospazi $W_j \subseteq V$ per $j \in J$, opportuno insieme di indici, allora l'intersezione $\bigcap_{j \in J} W_j$ è un sottospazio di V .

Esempio 12.22. Sia $I \subseteq \mathbb{R}$ un intervallo non vuoto, che per comodità supponiamo aperto, diciamo $I =]a, b[$. Nell'Esempio 12.13 abbiamo dimostrato che l'insieme $\mathcal{C}^p(I)$ delle funzioni $\varphi: I \rightarrow \mathbb{R}$ derivabili fino all'ordine p con derivata p -esima continua è un sottospazio di \mathbb{R}^I (o $\mathcal{C}^0(I)$). Definiamo

$$\mathcal{C}^\infty(I) = \bigcap_{p \geq 0} \mathcal{C}^p(I).$$

Per quanto detto sopra $\mathcal{C}^\infty(I)$ è un sottospazio di \mathbb{R}^I (o di $\mathcal{C}^p(I)$ per $p \geq 0$). I suoi elementi sono le funzioni definite su I aventi derivate di ogni ordine. \spadesuit

Passiamo ora a esaminare il caso dell'unione.

Proposizione 12.23.

Sia V uno spazio vettoriale su un campo \mathbb{K} . Dati due sottospazi $W', W'' \subseteq V$, allora $W' \cup W''$ è un sottospazio di V se e solo se o $W' \subseteq W''$, e in tal caso $W' \cup W'' = W''$, oppure $W'' \subseteq W'$, e in tal caso $W' \cup W'' = W'$.

Dimostrazione. È chiaro che se $W' \subseteq W''$ allora $W' \cup W'' = W''$, mentre se $W'' \subseteq W'$ allora $W' \cup W'' = W'$, dunque $W' \cup W''$ è banalmente un sottospazio in questi due casi.

Viceversa supponiamo che $W' \not\subseteq W''$ e $W'' \not\subseteq W'$. Allora esistono vettori $w' \in W' \setminus W''$ e $w'' \in W'' \setminus W'$. Si consideri $w = w' + w''$: se fosse $w \in W' \cup W''$ allora o $w \in W'$ o $w \in W''$. Nel primo caso si dovrebbe avere $w'' = w - w' \in W'$, nel secondo $w' = w - w'' \in W''$, in contrasto con la scelta fatta. Concludiamo che $w \notin W' \cup W''$. \square

In sostanza l'unione non è “quasi mai” un sottospazio. Dal momento che l'unione di due insiemi $X', X'' \subseteq X$ può essere definita come il più piccolo sottoinsieme di X contenente sia X' che X'' , possiamo definire per analogia una nuova operazione tra sottospazi che generalizza l'unione.

Definizione 12.24 (Somma di sottoinsiemi).

Sia V uno spazio vettoriale su un campo \mathbb{K} . Dati due sottoinsiemi $W', W'' \subseteq V$ non vuoti, definiamo *somma di W' e W''* l'insieme

$$W' + W'' = \{ v \in V \mid \exists w' \in W', w'' \in W'' : v = w' + w'' \}.$$

Osservazione 12.25. Se $0_V \in W''$ allora $w' + 0_V \in W' + W''$ per ogni $w' \in W'$, dunque $W' \subseteq W' + W''$. In particolare se $0_V \in W', W''$ (ed è questo il caso se sono entrambi sottospazi), allora l'unione $W' \cup W'' \subseteq W' + W''$.

Esempio 12.26. Siano $A \in \mathbb{K}^{m,n}$ e $B \in \mathbb{K}^{m,p}$ matrici, \mathbb{K} . Supponiamo che l'equazione matriciale $AX = B$ abbia almeno una soluzione $X_0 \in \mathbb{K}^{n,p}$; allora

$$\{ X \in \mathbb{K}^{n,p} \mid AX = B \} = \{ X_0 \} + \{ X \in \mathbb{K}^{n,p} \mid AX = 0_{m,p} \}.$$

Osserviamo che $0_{n,p} \in \{ X \in \mathbb{K}^{n,p} \mid AX = 0_{m,p} \}$ e, infatti,

$$X_0 \in \{ X_0 \} + \{ X \in \mathbb{K}^{n,p} \mid AX = 0_{m,p} \}. \quad \spadesuit$$

Proposizione 12.27.

Sia V uno spazio vettoriale su \mathbb{K} . Dati due sottospazi $W', W'' \subseteq V$, allora $W' + W''$ è un sottospazio di V ; inoltre se $W \subseteq V$ è un qualsiasi sottospazio contenente W' e W'' , allora $W' + W'' \subseteq W$.

Dimostrazione. Poiché $0_V \in W', W''$ si ha $0_V + 0_V \in W' + W''$.

Siano ora $v_1, v_2 \in W' + W''$: ciò significa che esistono $w'_1, w'_2 \in W'$ e $w''_1, w''_2 \in W''$ tali che $v_1 = w'_1 + w''_1$ e $v_2 = w'_2 + w''_2$. Allora

$$\begin{aligned} v_1 + v_2 &= (w'_1 + w''_1) + (w'_2 + w''_2) = w'_1 + w''_1 + w'_2 + w''_2 \\ &= w'_1 + w'_2 + w''_1 + w''_2 = (w'_1 + w'_2) + (w''_1 + w''_2) \in W' + W'', \end{aligned}$$

ovvero $W' + W''$ è chiuso rispetto alla somma.

Siano $\alpha \in \mathbb{K}$ e $v \in W' + W''$: ciò significa che esistono $w' \in W'$ e $w'' \in W''$ tali che $v = w' + w''$. Allora

$$\alpha v = \alpha(w' + w'') = \alpha w' + \alpha w'' = (\alpha w') + (\alpha w'') \in W' + W'',$$

ovvero $W' + W''$ è chiuso rispetto al prodotto per scalari.

Concludiamo che $W' + W''$ è un sottospazio di V grazie alla Proposizione 12.10.

Infine sia $W \subseteq V$ un sottospazio contenente W' e W'' : se $v \in W' + W''$ esistono $w' \in W' \subseteq W$ e $w'' \in W'' \subseteq W$ tali che $v = w' + w''$, sicché $v \in W$ (perché W è un sottospazio, quindi è chiuso rispetto alla somma), da cui si deduce che $W' + W'' \subseteq W$. \square

Spesso si riassume la Proposizione 12.27 dicendo che “se W' e W'' sono sottospazi di V allora $W' + W''$ è il più piccolo sottospazio di V contenente $W' \cup W''$ (rispetto alla relazione di ordine parziale sull'insieme dei sottospazi di V dato dall'inclusione)”.

Dipendenza e indipendenza lineare

13.1 Combinazioni lineari

Si ricordi che \mathbb{K} indica un campo: per esempio $\mathbb{K} = \mathbb{R}, \mathbb{C}$.

Definizione 13.1.

Sia V uno spazio vettoriale su un campo \mathbb{K} . Dati $v_1, \dots, v_n \in V$, un vettore $v \in V$ si dice *combinazione lineare* di v_1, \dots, v_n se esistono scalari $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ tali che

$$v = \alpha_1 v_1 + \dots + \alpha_n v_n = \sum_{i=1}^n \alpha_i v_i.$$

- Il sottoinsieme di V costituito dai vettori che sono combinazione lineare di v_1, \dots, v_n si indica con $\mathcal{L}(v_1, \dots, v_n)$.
- L'insieme $\mathcal{L}(v_1, \dots, v_n)$ si dice *generato* da v_1, \dots, v_n e i vettori v_1, \dots, v_n si dicono *generatori* di $\mathcal{L}(v_1, \dots, v_n)$.
- Se esistono $v_1, \dots, v_n \in V$ tali che $V = \mathcal{L}(v_1, \dots, v_n)$, V si dice *finitamente generato*.

Per capire meglio il concetto di combinazione lineare prendiamo in considerazione alcuni esempi.

Esempio 13.2. Sia V uno spazio vettoriale su un campo \mathbb{K} e $v_1, \dots, v_n \in V$ vettori fissati. Allora

$$0_V = 0v_1 + \dots + 0v_n,$$

ovvero il vettore nullo è sempre combinazione lineare di un qualsiasi insieme di vettori. ♠

Esempio 13.3. Fissato nello spazio ordinario S_3 un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$, per la Proposizione 7.12 ogni vettore geometrico $\vec{v} \in V_3(O)$ si può decomporre come

$$\vec{v} = a\vec{i} + b\vec{j} + c\vec{k}$$

per opportuni $a, b, c \in \mathbb{R}$. Quindi lo spazio $V_3(O)$ coincide con $\mathcal{L}(\vec{i}, \vec{j}, \vec{k})$, ovvero $\vec{i}, \vec{j}, \vec{k}$ sono generatori di $V_3(O)$: in particolare $V_3(O)$ è finitamente generato. ♠

Esempio 13.4. In \mathbb{R}^3 si considerino i vettori $e_1 = (1, 0, 0)$ ed $e_2 = (0, 1, 0)$. Allora $v' = (3, 2, 1) \notin \mathcal{L}(e_1, e_2)$, cioè v' non è combinazione lineare di e_1, e_2 . Infatti dall'equazione vettoriale

$$(3, 2, 1) = v' = \alpha e_1 + \beta e_2 = \alpha(1, 0, 0) + \beta(0, 1, 0),$$

confrontando le componenti nelle stessa posizione dei vettori al primo e al secondo membro otteniamo il sistema

$$\begin{cases} \alpha = 3 \\ \beta = 2 \\ 0 = 1, \end{cases}$$

che, ovviamente, non ammette soluzioni. Quindi e_1 ed e_2 non sono generatori di \mathbb{R}^3 .

Invece $v'' = (3, 2, 0) \in \mathcal{L}(e_1, e_2)$ cioè v'' è combinazione lineare di e_1, e_2 . Infatti l'equazione vettoriale

$$(3, 2, 0) = v'' = \alpha e_1 + \beta e_2 = \alpha(1, 0, 0) + \beta(0, 1, 0),$$

si traduce nel sistema

$$\begin{cases} \alpha = 3 \\ \beta = 2 \\ 0 = 0, \end{cases}$$

avente soluzione $(\alpha, \beta) = (3, 2)$.

Più in generale gli elementi di $\mathcal{L}(e_1, e_2)$, cioè i vettori che sono combinazione lineare di e_1 ed e_2 , sono tutti e soli i vettori del tipo $(\alpha, \beta, 0)$ al variare di $\alpha, \beta \in \mathbb{R}$, cioè

$$\mathcal{L}(e_1, e_2) = \{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_3 = 0 \}.$$

Verificare per esercizio che $\mathcal{L}(e_1, e_2)$ è un sottospazio di \mathbb{R}^3 . ♠

Esempio 13.5. In \mathbb{R}^3 si considerino $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ e $e_3 = (0, 0, 1)$. Ogni vettore di \mathbb{R}^3 è combinazione lineare di e_1, e_2, e_3 : infatti

$$(x_1, x_2, x_3) = x_1(1, 0, 0) + x_2(0, 1, 0) + x_3(0, 0, 1),$$

quindi $\mathbb{R}^3 = \mathcal{L}(e_1, e_2, e_3)$, cioè e_1, e_2, e_3 generano \mathbb{R}^3 , che risulta essere \mathbb{R}^3 è finitamente generato.

Più in generale si considerino i vettori e_1, \dots, e_n in \mathbb{K}^n così definiti: e_i ha tutte le componenti nulle la i -esima che vale 1. Allora e_1, \dots, e_n sono generatori di \mathbb{K}^n .

Infatti scelto $x = (x_1, \dots, x_n) \in \mathbb{K}^n$ si ha

$$\begin{aligned} & x_1(1, 0, 0, 0, \dots, 0, 0) + \\ & x_2(0, 1, 0, 0, \dots, 0, 0) + \\ & x_3(0, 0, 1, 0, \dots, 0, 0) + \\ & \quad \vdots \\ & x_{n-1}(0, 0, 0, \dots, 1, 0) + \\ & x_n(0, 0, 0, 0, \dots, 0, 1) = \\ & \hline & (x_1, x_2, x_3, \dots, x_{n-1}, x_n) \end{aligned}$$

cioè $x = x_1e_1 + \dots + x_n e_n$. Quindi $\mathbb{K}^n = \mathcal{L}(e_1, \dots, e_n)$ è finitamente generato. ♠

Esempio 13.6. In $\mathbb{C}^{2,2}$ si considerino i vettori

$$E_{1,1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_{1,2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad E_{2,1} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad E_{2,2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Si ha che

$$\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} = a_{1,1} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + a_{1,2} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + a_{2,1} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + a_{2,2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

quindi $\mathbb{C}^{2,2} = \mathcal{L}(E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2})$, ovvero $E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2}$ sono generatori di $\mathbb{C}^{2,2}$, che risulta essere $\mathbb{C}^{2,2}$ finitamente generato.

Più in generale, in $\mathbb{K}^{m,n}$ si consideri per ogni coppia di indici (i, j) , con $1 \leq i \leq m$ e $1 \leq j \leq n$, la matrice $E_{i,j}$ avente tutte le entrate nulle tranne quella in posizione (i, j) che vale 1. Allora le matrici $E_{i,j}$ con $1 \leq i \leq m$ e $1 \leq j \leq n$ sono generatori di $\mathbb{K}^{m,n}$, cioè $\mathbb{K}^{m,n} = \mathcal{L}(E_{i,j} | 1 \leq i \leq m, 1 \leq j \leq n)$. In particolare $\mathbb{K}^{m,n}$ è finitamente generato. ♠

Osservazione 13.7. In tutti gli esempi presi in considerazione finora la scrittura di un fissato vettore come combinazione lineare di generatori dati è unica. Non è detto che questo accada sempre, cioè uno stesso vettore può essere scritto in più di un modo come combinazione lineare di un fissato insieme di vettori.

Per esempio, se consideriamo $v_1 = (1, 1, -1)$, $v_2 = (2, -3, 0)$, $v_3 = (-3, 2, 1)$ in \mathbb{R}^3 , allora

$$0 \cdot v_1 + 0 \cdot v_2 + 0 \cdot v_3 = (0, 0, 0) = 1 \cdot v_1 + 1 \cdot v_2 + 1 \cdot v_3.$$

Il lettore verifichi per esercizio che il vettore $0_{\mathbb{R}^3}$ si scrive in infiniti modi come combinazione lineare dei vettori v_1, v_2, v_3 .

Vediamo adesso che non tutti gli spazi vettoriali sono finitamente generati.

Esempio 13.8. Si consideri l'insieme $\mathbb{K}[x]$ dei polinomi nella variabile x a coefficienti in un campo \mathbb{K} . Si ricordi che ogni polinomio si può scrivere in maniera unica come

$$p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots + a_nx^n + \dots \quad (13.1.1)$$

dove gli esponenti della variabile x sono ordinati in maniera strettamente crescente (in particolare sono tutti diversi) e i coefficienti $a_i \in \mathbb{K}$. Tale scrittura viene detta *forma ridotta del polinomio*.

Mostriamo che $\mathbb{K}[x]$ è uno spazio vettoriale su \mathbb{K} con l'usuale operazione di somma di polinomi e di prodotto di un polinomio per una costante in \mathbb{K} .

Infatti $\mathbb{K}[x] \neq \emptyset$ poiché contiene tutti i polinomi costanti, in particolare il polinomio nullo. Inoltre le operazioni di somma e prodotto si possono definire a partire dalla successione dei coefficienti del polinomio, utilizzando le operazioni di somma e prodotto definite in \mathbb{K} . Per esempio, se il polinomio $p(x)$ ha forma ridotta data da (13.1.1), allora

$$\alpha p(x) = \alpha a_0 + \alpha a_1x + \alpha a_2x^2 + \alpha a_3x^3 + \cdots + \alpha a_nx^n + \dots$$

Poiché la somma e il prodotto di elementi di \mathbb{K} soddisfano gli otto assiomi della definizione 12.1, anche le operazioni di somma e prodotto in $\mathbb{K}[x]$ soddisfano le stesse proprietà.

Si noti che in $\mathbb{K}[x]$ si possono considerare polinomi formati da un solo addendo, cioè polinomi della forma $p(x) = ax^n$, dove $a \in \mathbb{K}$ ed $n \in \mathbb{Z}$, $n \geq 0$: tali polinomi particolari si dicono *monomi nell'indeterminata x* .

Se $a \neq 0$, si definisce *grado* di ax^n il numero $\deg(ax^n) = n$, mentre se $a = 0$ si pone $\deg(0) = -\infty$. Un polinomio nell'indeterminata x è, dunque, una qualsiasi somma finita di monomi: il grado $\deg(p)$ di un polinomio $p(x)$ è il grado massimo dei monomi con coefficiente non nullo che compaiono quando è scritto in forma ridotta. Si noti che

$$\deg(p + q) \leq \max\{\deg(p), \deg(q)\}, \quad \deg(\alpha p) \leq \deg(p)$$

per ogni $\alpha \in \mathbb{K}$ e per ogni $p(x), q(x) \in \mathbb{K}[x]$.

Verifichiamo che $\mathbb{K}[x]$ non è finitamente generato. Consideriamo un qualsiasi insieme finito di polinomi $p_1(x), \dots, p_n(x) \in \mathbb{K}[x]$ e sia

$$d = \max\{\deg(p_i) \mid i = 1, \dots, n\}.$$

Allora se $p(x) \in \mathcal{L}(p_1(x), \dots, p_n(x)) \subseteq \mathbb{K}[x]$ segue che $\deg(p) \leq d$, dunque, comunque si scelgano $p_1(x), \dots, p_n(x) \in \mathbb{K}[x]$, i polinomi di grado strettamente maggiore di d non sono loro combinazione lineare.

Altri esempi di spazi non finitamente generati sono gli spazi di funzioni $\mathbb{R}^I, \mathcal{C}^p(I)$, $p \geq 0$, ove $I \subseteq \mathbb{R}$ è un intervallo aperto. Infatti, per esempio, si consideri $\mathbb{R}[x]$ come sottospazio di $\mathbb{R}^{\mathbb{R}}, \mathcal{C}^p(\mathbb{R})$ con la naturale identificazione di un polinomio con la funzione polinomiale corrispondente: vedremo più avanti che se V è uno spazio vettoriale finitamente generato allora tali sono tutti i suoi sottospazi. Poiché $\mathbb{R}[x]$ non è finitamente generato allora né $\mathbb{R}^{\mathbb{R}}$, né $\mathcal{C}^p(\mathbb{R})$ possono esserlo. ♠

In tutti gli esempi presi in esame sopra, l'insieme generato da un certo numero di vettori di uno spazio vettoriale V viene a essere, di fatto, un sottospazio di V , eventualmente coincidente con V stesso: questo è un fatto del tutto generale.

Proposizione 13.9.

Sia V uno spazio vettoriale su un campo \mathbb{K} . Dati $v_1, \dots, v_n \in V$, allora l'insieme $\mathcal{L}(v_1, \dots, v_n)$ è un sottospazio di V , inoltre

$$\mathcal{L}(v_1, \dots, v_n) = \mathcal{L}(v_1) + \dots + \mathcal{L}(v_n). \quad (13.1.2)$$

Dimostrazione. Innanzi tutto osserviamo che gli elementi di $\mathcal{L}(v_1, \dots, v_n)$ sono tutti e soli i vettori di V della forma $\alpha_1 v_1 + \dots + \alpha_n v_n$. In particolare $\alpha_i v_i \in \mathcal{L}(v_i)$ per ogni $i = 1, \dots, n$. L'identità (13.1.2) segue da queste due osservazioni.

Verifichiamo che $\mathcal{L}(v_1, \dots, v_n)$ è un sottospazio utilizzando la Proposizione 12.10. Dall'Esempio 13.2 deduciamo $0_V \in \mathcal{L}(v_1, \dots, v_n)$. Siano $v', v'' \in \mathcal{L}(v_1, \dots, v_n)$: quindi

$$v' = \alpha'_1 v_1 + \dots + \alpha'_n v_n, \quad v'' = \alpha''_1 v_1 + \dots + \alpha''_n v_n,$$

da cui deduciamo

$$v' + v'' = \alpha'_1 v_1 + \dots + \alpha'_n v_n + \alpha''_1 v_1 + \dots + \alpha''_n v_n = (\alpha'_1 + \alpha''_1) v_1 + \dots + (\alpha'_n + \alpha''_n) v_n \in \mathcal{L}(v_1, \dots, v_n).$$

Infine se $\alpha \in \mathbb{K}$ e $v \in \mathcal{L}(v_1, \dots, v_n)$, possiamo supporre $v = \alpha_1 v_1 + \dots + \alpha_n v_n$, dunque

$$\alpha v = \alpha(\alpha_1 v_1 + \dots + \alpha_n v_n) = (\alpha \alpha_1) v_1 + \dots + (\alpha \alpha_n) v_n \in \mathcal{L}(v_1, \dots, v_n).$$

Concludiamo che $\mathcal{L}(v_1, \dots, v_n)$ è un sottospazio. □

Per tale motivo, d'ora in avanti chiameremo $\mathcal{L}(v_1, \dots, v_n)$ *il sottospazio di V generato dai vettori v_1, \dots, v_n* .

Mostriamo adesso che se conosciamo gli insiemi di generatori di due sottospazi, siamo anche in grado di determinare un insieme di generatori per la loro somma.

Proposizione 13.10.

Sia V uno spazio vettoriale su un campo \mathbb{K} . Dati sottospazi $W', W'' \subseteq V$ finitamente generati, anche il sottospazio somma $W' + W''$ è finitamente generato. Inoltre se

$$W' = \mathcal{L}(w'_1, \dots, w'_{n'}), \quad W'' = \mathcal{L}(w''_1, \dots, w''_{n''}),$$

allora

$$W' + W'' = \mathcal{L}(w'_1, \dots, w'_{n'}, w''_1, \dots, w''_{n''}).$$

Dimostrazione. Utilizzando ripetutamente la formula (13.1.2) si ottiene

$$\begin{aligned} \mathcal{L}(w'_1, \dots, w'_{n'}) + \mathcal{L}(w''_1, \dots, w''_{n''}) &= \mathcal{L}(w'_1) + \dots + \mathcal{L}(w'_{n'}) + \mathcal{L}(w''_1) + \dots + \mathcal{L}(w''_{n''}) \\ &= \mathcal{L}(w'_1, \dots, w'_{n'}, w''_1, \dots, w''_{n''}), \end{aligned}$$

e ciò dimostra completamente l'enunciato. □

Concludiamo che un insieme di generatori per la somma di sottospazi finitamente generati è l'unione di insiemi di generatori per ciascuno dei sottospazi.

13.2 Dipendenza lineare

Definizione 13.11.

Sia V uno spazio vettoriale su un campo \mathbb{K} . I vettori $v_1, \dots, v_n \in V$ si dicono *linearmente dipendenti* se esistono $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ non tutti nulli tali che

$$\alpha_1 v_1 + \dots + \alpha_n v_n = 0_V.$$

In caso contrario i vettori v_1, \dots, v_n si dicono *linearmente indipendenti*.

Quindi dei vettori v_1, \dots, v_n sono linearmente indipendenti se, per ogni scelta di scalari $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ non tutti nulli, risulta

$$\alpha_1 v_1 + \dots + \alpha_n v_n \neq 0_V.$$

Un altro modo per definire la nozione di indipendenza lineare è il seguente: i vettori v_1, \dots, v_n sono linearmente indipendenti se l'equazione in V

$$\alpha_1 v_1 + \dots + \alpha_n v_n = 0_V$$

ha $(\alpha_1, \dots, \alpha_n) = 0_{\mathbb{K}^n} \in \mathbb{K}^n$ come unica soluzione, mentre sono linearmente dipendenti se ha soluzioni non nulle.

Esempio 13.12. Sia V uno spazio vettoriale su un campo \mathbb{K} e si considerino vettori v_1, \dots, v_n . Allora

$$1 \cdot 0_V + 0v_1 + 0v_2 + \dots + 0v_n = 0_V,$$

cioè $0_V, v_1, v_2, \dots, v_n$ sono linearmente dipendenti. In particolare, un qualsiasi insieme di vettori che contiene il vettore nullo è costituito da vettori linearmente dipendenti.

Più in generale siano $v_1, \dots, v_n \in V$ linearmente dipendenti e $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ non tutti nulli tali che $\alpha_1 v_1 + \dots + \alpha_n v_n = 0_V$: pertanto nella relazione di dipendenza lineare

$$\alpha_1 v_1 + \dots + \alpha_n v_n + 0v_{n+1} + \dots + 0v_m = 0_V$$

non tutti gli scalari sono nulli, quindi anche $v_1, \dots, v_n, v_{n+1}, \dots, v_m$ sono linearmente dipendenti, qualsiasi sia la scelta dei vettori $v_{n+1}, \dots, v_m \in V$. ♠

Esempio 13.13. Fissato nello spazio ordinario S_3 un sistema di riferimento cartesiano ortogonale $O\vec{i}\vec{j}\vec{k}$, i tre versori $\vec{i}, \vec{j}, \vec{k}$ sono linearmente indipendenti in $V_3(O)$: infatti vale che $\vec{0} = a\vec{i} + b\vec{j} + c\vec{k}$ se e solo se $0 = |\vec{0}| = |a\vec{i} + b\vec{j} + c\vec{k}| = \sqrt{a^2 + b^2 + c^2}$, ovvero se e solo se $a = b = c = 0$. ♠

Esempio 13.14. I tre vettori $v_1 = (1, 1, -1)$, $v_2 = (2, -3, 0)$, $v_3 = (-3, 2, 1)$ sono linearmente dipendenti in \mathbb{R}^3 poiché

$$1(1, 1, -1) + 1(2, -3, 0) + 1(-3, 2, 1) = (0, 0, 0).$$

Invece i vettori $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$ sono linearmente indipendenti poiché

$$\alpha_1(1, 0, 0) + \alpha_2(0, 1, 0) + \alpha_3(0, 0, 1) = (\alpha_1, \alpha_2, \alpha_3),$$

che è nullo se e solo se $\alpha_1 = \alpha_2 = \alpha_3 = 0$.

Più in generale, in \mathbb{K}^n i vettori e_1, \dots, e_n sono linearmente indipendenti. Lo stesso vale per i vettori $E_{i,j}$, $1 \leq i \leq m$, $1 \leq j \leq n$ in $\mathbb{K}^{m,n}$. ♠

Osservazione 13.15. Le definizioni di dipendenza e indipendenza lineare non dipendono dall'ordine dei vettori: se dei vettori sono linearmente dipendenti o indipendenti in un fissato ordine, lo sono in qualsiasi altro.

Sia V uno spazio vettoriale su un campo \mathbb{K} e siano $v_1, \dots, v_n \in V$. Ci domandiamo cosa significhi che uno, due o tre vettori siano linearmente dipendenti.

- Dire che il vettore v_1 è linearmente dipendente equivale ad affermare che esiste $\alpha_1 \in \mathbb{K} \setminus \{0\}$ tale che $\alpha_1 v_1 = 0_V$. Per la legge di annullamento del prodotto ciò significa che $v_1 = 0_V$: quindi *un vettore è linearmente dipendente se e solo se è nullo*.
- Dire che v_1 e v_2 sono linearmente dipendenti equivale ad affermare che esistono scalari $\alpha_1, \alpha_2 \in \mathbb{K}$ non contemporaneamente nulli tali che

$$\alpha_1 v_1 + \alpha_2 v_2 = 0_V.$$

Supponiamo che sia $\alpha_2 \neq 0$. Allora, sommando ad ambo i membri della relazione sopra l'opposto di $\alpha_1 v_1$ e moltiplicando quindi ambo i membri per α_2^{-1} (che esiste perché $\alpha_2 \neq 0$), si ottiene

$$v_2 = -\frac{\alpha_1}{\alpha_2} v_1,$$

cioè uno dei vettori è multiplo dell'altro (cioè è sua combinazione lineare). Viceversa, se ciò accade, diciamo

$$v_2 = \lambda v_1$$

per qualche $\lambda \in \mathbb{K}$, allora sommando ad ambo i membri l'opposto di λv_1 si ottiene la relazione di dipendenza lineare

$$(-\lambda)v_1 + 1v_2 = 0_V,$$

i cui coefficienti non sono tutti nulli (infatti $1 \neq 0$): quindi *due vettori sono linearmente dipendenti se e solo se uno dei due è multiplo dell'altro*.

- Dire che v_1, v_2 e v_3 sono linearmente dipendenti equivale ad affermare che esistono $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{K}$ non contemporaneamente nulli tali che

$$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = 0_V.$$

Supponiamo che sia $\alpha_3 \neq 0$. Allora sommando ad ambo i membri della relazione sopra l'opposto di $\alpha_1 v_1 + \alpha_2 v_2$ e moltiplicando quindi ambo i membri α_3^{-1} (che esiste perché $\alpha_3 \neq 0$), si ottiene

$$v_3 = -\frac{\alpha_1}{\alpha_3} v_1 - \frac{\alpha_2}{\alpha_3} v_2,$$

cioè uno dei vettori è combinazione lineare dei rimanenti. Viceversa, se ciò accade, diciamo

$$v_3 = \lambda_1 v_1 + \lambda_2 v_2$$

per qualche $\lambda_1, \lambda_2 \in \mathbb{K}$, sommando ad ambo i membri l'opposto di $\lambda_1 v_1 + \lambda_2 v_2$ si ottiene la relazione di dipendenza lineare

$$(-\lambda_1)v_1 + (-\lambda_2)v_2 + 1v_3 = 0_V,$$

i cui coefficienti non sono tutti nulli (infatti $1 \neq 0$): quindi *tre vettori sono linearmente dipendenti se e solo se uno di loro è combinazione lineare dei rimanenti*.

Esempio 13.16. In \mathbb{R}^3 consideriamo i vettori $v_1 = (1, 1, -1)$, $v_2 = (2, -3, 0)$ e $v_3 = (-3, 2, 1)$ dell'Esempio 13.14: i tre vettori sono linearmente dipendenti e si ha

$$(-1)(1, 1, -1) + (-1)(2, -3, 0) = (-3, 2, 1). \quad \spadesuit$$

Il ragionamento che abbiamo appena fatto, con le modifiche del caso, è alla base della dimostrazione del risultato seguente.

Proposizione 13.17.

Siano V uno spazio vettoriale su un campo \mathbb{K} . Dati $v_1, \dots, v_n \in V$ con $n \geq 2$, si ha:

- (i) i vettori v_1, \dots, v_n sono linearmente dipendenti se e solo se uno di loro è combinazione lineare dei rimanenti;
- (ii) i vettori v_1, \dots, v_n sono linearmente dipendenti se e solo se uno di loro è combinazione lineare di quelli che lo precedono nell'ordine fissato.

Dimostrazione.

- (i) I vettori v_1, \dots, v_n sono linearmente dipendenti se e solo se esistono scalari $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ non contemporaneamente nulli tali che

$$\alpha_1 v_1 + \dots + \alpha_n v_n = 0_V;$$

se, per fissare le idee, $\alpha_n \neq 0$ si ottiene

$$v_n = -\frac{\alpha_1}{\alpha_n} v_1 - \dots - \frac{\alpha_{n-1}}{\alpha_n} v_{n-1},$$

cioè uno dei vettori è combinazione lineare dei rimanenti.

Viceversa, se ciò accade, diciamo

$$v_n = \lambda_1 v_1 + \cdots + \lambda_{n-1} v_{n-1}$$

per qualche $\lambda_1, \dots, \lambda_{n-1} \in \mathbb{K}$, si ottiene la relazione di dipendenza lineare

$$(-\lambda_1)v_1 + \cdots + (-\lambda_{n-1})v_{n-1} + 1v_n = 0_V,$$

i cui coefficienti non sono tutti nulli, quindi v_1, \dots, v_n sono linearmente dipendenti.

- (ii) È chiaro che se esiste m tale che $v_m \in \mathcal{L}(v_1, \dots, v_{m-1})$ allora i vettori sono linearmente dipendenti dalla parte (i) dell'enunciato. Viceversa, supponiamo che v_1, \dots, v_n siano linearmente dipendenti e sia

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0_V$$

una loro relazione di dipendenza lineare a coefficienti $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ non tutti nulli. Poniamo

$$m = \max\{ i \in \{ 1, \dots, n \} \mid \alpha_i \neq 0 \} :$$

ciò significa che nella relazione di dipendenza lineare $\alpha_i = 0$ per $i \geq m + 1$. Quindi tale relazione di dipendenza lineare è del tipo

$$\alpha_1 v_1 + \cdots + \alpha_m v_m = 0_V$$

con $\alpha_m \neq 0$: per quanto visto sopra

$$v_m = -\frac{\alpha_1}{\alpha_m} v_1 - \cdots - \frac{\alpha_{m-1}}{\alpha_m} v_{m-1},$$

ovvero $v_m \in \mathcal{L}(v_1, \dots, v_{m-1})$. □

Basi di uno spazio vettoriale finitamente generato

14.1 Il metodo degli scarti

Sia dato uno spazio vettoriale V su un campo \mathbb{K} , per esempio $\mathbb{K} = \mathbb{R}, \mathbb{C}$, e siano $v_1, \dots, v_n \in V$ vettori fissati. Quanto visto nel capitolo precedente ci suggerisce il seguente algoritmo per stabilire se tali vettori sono linearmente indipendenti oppure dipendenti. Tale procedimento, detto *metodo degli scarti*, ci permette di ricavare da un qualsiasi insieme di vettori un sottoinsieme di vettori che siano anche linearmente indipendenti.

Passo (1). Se $v_1 = 0_V$, allora si scarta v_1 e si riparte, rinominando tutti i vettori rimanenti come v_1, v_2, \dots ; se invece $v_1 \neq 0_V$, andiamo al passo (2).

Passo (2). Se $v_2 = \lambda v_1$, allora si scarta v_2 e si riparte, rinominando tutti i vettori rimanenti come v_2, v_3, \dots ; se invece $v_2 \neq \lambda v_1$, andiamo al passo (3).

⋮

Passo (i). Se $v_i \in \mathcal{L}(v_1, \dots, v_{i-1})$, allora si scarta v_i e si riparte, rinominando tutti i vettori rimanenti come v_i, v_{i+1}, \dots ; se invece $v_i \notin \mathcal{L}(v_1, \dots, v_{i-1})$, andiamo al passo (i+1).

Ovviamente l'algoritmo di cui sopra non produce nessun risultato se $V = \{0_V\}$.

Si noti che, se a un certo passo, risulta $v_i \in \mathcal{L}(v_1, \dots, v_{i-1})$, allora scartando tale vettore v_i non si cambia il sottospazio, cioè $\mathcal{L}(v_1, \dots, v_{i-1}) = \mathcal{L}(v_1, \dots, v_i)$. Infatti, chiaramente, ogni vettore che è combinazione lineare di v_1, \dots, v_{i-1} lo è anche di v_1, \dots, v_i , cioè $\mathcal{L}(v_1, \dots, v_{i-1}) \subseteq \mathcal{L}(v_1, \dots, v_i)$.

Viceversa sia $v_i = \lambda_1 v_1 + \dots + \lambda_{i-1} v_{i-1}$: se $v \in \mathcal{L}(v_1, \dots, v_i)$, si ha

$$\begin{aligned} v &= \alpha_1 v_1 + \dots + \alpha_{i-1} v_{i-1} + \alpha_i v_i \\ &= \alpha_1 v_1 + \dots + \alpha_{i-1} v_{i-1} + \alpha_i (\lambda_1 v_1 + \dots + \lambda_{i-1} v_{i-1}) \\ &= (\alpha_1 + \alpha_i \lambda_1) v_1 + \dots + (\alpha_{i-1} + \alpha_i \lambda_{i-1}) v_{i-1} \in \mathcal{L}(v_1, \dots, v_{i-1}), \end{aligned}$$

dunque $\mathcal{L}(v_1, \dots, v_{i-1}) \supseteq \mathcal{L}(v_1, \dots, v_i)$. In particolare, con il metodo degli scarti a ogni passo i vettori considerati possono eventualmente diminuire, ma il sottospazio che essi generano rimane invariato.

L'utilità di tale algoritmo sarà chiara tra poco, quando parleremo di basi.

Esempio 14.1. Nello spazio \mathbb{R}^4 si considerino $v_1 = (1, 2, -1, 1)$, $v_2 = (2, 1, 1, 0)$, $v_3 = (0, 0, 0, 0)$, $v_4 = (1, 1, 0, 0)$, $v_5 = (4, 4, 0, 1)$. Vogliamo trovare un sottoinsieme di $\{v_1, v_2, v_3, v_4, v_5\}$ costituito da vettori linearmente indipendenti e che generano il sottospazio $\mathcal{L}(v_1, v_2, v_3, v_4, v_5)$: a tale scopo applichiamo il metodo degli scarti.

Passo (1). Innanzi tutto $v_1 \neq 0_{\mathbb{R}^4}$, quindi possiamo considerarlo come primo elemento del sottoinsieme cercato e passare a v_2 .

Passo (2). Si ha $v_2 \notin \mathcal{L}(v_1)$: infatti se esistesse $\alpha \in \mathbb{R}$ tale che $v_2 = \alpha v_1$, confrontando l'ultima componente dei due vettori ricaveremmo $\alpha = 0$ e quindi $v_2 = 0_{\mathbb{R}^4}$. Dunque possiamo prendere v_2 come secondo elemento del nostro sottoinsieme e passare a v_3 .

Passo (3). Chiaramente $v_3 \in \mathcal{L}(v_1, v_2)$ (in quanto $v_3 = 0v_1 + 0v_2$), quindi lo scartiamo e passiamo a v_4 .

Passo (4). Se $v_4 \in \mathcal{L}(v_1, v_2)$, esisterebbero $\alpha_1, \alpha_2 \in \mathbb{R}$ tali che $v_4 = \alpha_1 v_1 + \alpha_2 v_2$ cioè

$$\alpha_1(1, 2, -1, 1) + \alpha_2(2, 1, 1, 0) = (\alpha_1 + 2\alpha_2, 2\alpha_1 + \alpha_2, -\alpha_1 + \alpha_2, \alpha_1) = (1, 1, 0, 0).$$

Confrontando l'ultima componente dei due vettori segue $\alpha_1 = 0$, dunque, dalla penultima, $\alpha_2 = 0$, sicché dovrebbe essere $v_4 = 0_{\mathbb{R}^4}$. Quindi possiamo considerare v_4 come terzo elemento del nostro sottoinsieme e passare a v_5 .

Passo (5). In questo caso si verifica facilmente che $v_5 \in \mathcal{L}(v_1, v_2, v_4)$: infatti si ha che $v_1 + v_2 + v_4 = v_5$, per tale motivo il vettore v_5 è da scartare.

Concludiamo che il sottoinsieme cercato è $\{v_1, v_2, v_4\}$: in particolare risulta $\mathcal{L}(v_1, v_2, v_4) = \mathcal{L}(v_1, v_2, v_3, v_4, v_5) \subset \mathbb{R}^4$. Si ha $e_1 = (1, 0, 0, 0) \notin \mathcal{L}(v_1, v_2, v_4)$: infatti la relazione

$$\alpha_1(1, 2, -1, 1) + \alpha_2(2, 1, 1, 0) + \alpha_3(1, 1, 0, 0) = (1, 0, 0, 0)$$

si traduce nel sistema

$$\begin{cases} \alpha_1 + 2\alpha_2 + \alpha_3 = 1 \\ 2\alpha_1 + \alpha_2 + \alpha_3 = 0 \\ -\alpha_1 + \alpha_2 = 0 \\ \alpha_1 = 0 \end{cases}$$

che è incompatibile.

Il lettore verifichi che se si applica l'algoritmo a partire dall'ultimo vettore, si ottiene un altro sottoinsieme, sempre costituito da tre vettori: non cambia invece il sottospazio generato, che è sempre $\mathcal{L}(v_1, v_2, v_3, v_4, v_5)$. ♠

14.2 Basi di uno spazio vettoriale

Vogliamo ora confrontare le nozioni di generatori e indipendenza lineare introdotte in precedenza. A tale scopo sia V uno spazio vettoriale su un campo \mathbb{K} e siano $v_1, \dots, v_n \in V$ suoi generatori linearmente indipendenti. Allora se $v \in V$ è un qualsiasi vettore esistono scalari $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ tali che

$$v = \alpha_1 v_1 + \dots + \alpha_n v_n,$$

dunque v_1, \dots, v_n, v sono vettori linearmente dipendenti dalla Proposizione 13.17.

D'altro canto, per la stessa Proposizione, nessuno dei vettori v_i è combinazione lineare dei rimanenti, quindi i rimanenti vettori non sono più generatori di V .

In particolare, rispetto alla relazione di inclusione, gli insiemi di generatori linearmente indipendenti di uno spazio vettoriale finitamente generato sono i sottoinsiemi *minimali* di generatori e i sottoinsiemi *massimali* di vettori linearmente indipendenti.

Per questo motivo (e per molti altri che saranno chiari nel seguito) tali sottoinsiemi rivestono una particolare importanza e meritano un nome particolare.

Definizione 14.2 (Basi di uno spazio vettoriale).

Sia V uno spazio vettoriale su un campo \mathbb{K} . Dati $v_1, \dots, v_n \in V$, l'insieme ordinato $\mathcal{B} = (v_1, \dots, v_n)$ si dice *base* di V se:

(B1) $V = \mathcal{L}(v_1, \dots, v_n)$, cioè v_1, \dots, v_n sono generatori di V ;

(B2) v_1, \dots, v_n sono linearmente indipendenti.

Chiaramente, perché si possa parlare di base secondo la definizione data, lo spazio vettoriale V deve essere finitamente generato e $V \neq \{0_V\}$.

Un'importante proprietà dei sistemi di generatori linearmente indipendenti, quindi anche delle basi, è la seguente.

Proposizione 14.3.

Sia V uno spazio vettoriale su un campo \mathbb{K} . Data una base $\mathcal{B} = (v_1, \dots, v_n)$ di V , per ogni vettore $v \in V$ esiste unico $(\alpha_1, \dots, \alpha_n) \in \mathbb{K}^n$ tale che

$$v = \alpha_1 v_1 + \dots + \alpha_n v_n.$$

Dimostrazione. Poiché $V = \mathcal{L}(v_1, \dots, v_n)$, esiste $(\alpha_1, \dots, \alpha_n) \in \mathbb{K}^n$ tale che $v = \alpha_1 v_1 + \dots + \alpha_n v_n$. Per dimostrarne l'unicità supponiamo che sia anche $v = \beta_1 v_1 + \dots + \beta_n v_n$ per qualche altro $(\beta_1, \dots, \beta_n) \in \mathbb{K}^n$. Allora

$$\alpha_1 v_1 + \dots + \alpha_n v_n = v = \beta_1 v_1 + \dots + \beta_n v_n,$$

dunque

$$(\alpha_1 - \beta_1)v_1 + \dots + (\alpha_n - \beta_n)v_n = 0_V;$$

poiché v_1, \dots, v_n sono linearmente indipendenti, segue che $\alpha_1 - \beta_1 = \dots = \alpha_n - \beta_n = 0$, cioè $\alpha_i = \beta_i$ per ogni $i = 1, \dots, n$. \square

Definizione 14.4 (Componenti di un vettore rispetto a una base).

Sia V uno spazio vettoriale su un campo \mathbb{K} . Data una base $\mathcal{B} = (v_1, \dots, v_n)$ di V definiamo *componenti* di $v \in V$ rispetto alla base \mathcal{B} l'unico $[v]_{\mathcal{B}} = (\alpha_1, \dots, \alpha_n) \in \mathbb{K}^n$ tale che

$$v = \alpha_1 v_1 + \dots + \alpha_n v_n.$$

Osservazione 14.5. Sia V uno spazio vettoriale su un campo \mathbb{K} e una sua base $\mathcal{B} = (v_1, \dots, v_n)$. Chiaramente $[0_V]_{\mathcal{B}} = 0_{\mathbb{K}^n} \in \mathbb{K}^n$.

Siano poi $v', v'' \in V$ con

$$[v']_{\mathcal{B}} = (\alpha'_1, \dots, \alpha'_n) \in \mathbb{K}^n, \quad [v'']_{\mathcal{B}} = (\alpha''_1, \dots, \alpha''_n) \in \mathbb{K}^n;$$

poiché ciò significa che

$$v' = \alpha'_1 v_1 + \dots + \alpha'_n v_n, \quad v'' = \alpha''_1 v_1 + \dots + \alpha''_n v_n,$$

risulta anche

$$v' + v'' = \alpha'_1 v_1 + \dots + \alpha'_n v_n + \alpha''_1 v_1 + \dots + \alpha''_n v_n = (\alpha'_1 + \alpha''_1) v_1 + \dots + (\alpha'_n + \alpha''_n) v_n,$$

ovvero

$$[v' + v'']_{\mathcal{B}} = (\alpha'_1 + \alpha''_1, \dots, \alpha'_n + \alpha''_n) \in \mathbb{K}^n.$$

Concludiamo che *le componenti rispetto a una base della somma di due vettori sono la somma delle componenti rispetto alla stessa base dei due vettori.*

In maniera simile il lettore verifichi che se $\lambda \in \mathbb{K}$ e $v \in V$ allora

$$[\lambda v]_{\mathcal{B}} = \lambda [v]_{\mathcal{B}}.$$

Per capire meglio il concetto di base e componenti di un vettore rispetto a essa prendiamo in considerazione alcuni esempi.

Esempio 14.6. Fissato nello spazio ordinario un sistema di riferimento cartesiano ortogonale $O\vec{i}\vec{j}\vec{k}$, la terna ordinata $\mathcal{B} = (\vec{i}, \vec{j}, \vec{k})$ è una base di $V_3(O)$ in forza degli Esempi 13.5 e 13.13. ♠

Esempio 14.7. In \mathbb{R}^3 si considerino $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ e $e_3 = (0, 0, 1)$. Allora $\mathcal{C} = (e_1, e_2, e_3)$ è una base di \mathbb{R}^3 dagli Esempi 13.5 e 13.14. Tale base viene detta *base canonica di \mathbb{R}^3* . Se $(x, y, z) \in \mathbb{R}^3$, allora chiaramente

$$[(x, y, z)]_{\mathcal{C}} = (x, y, z).$$

Invece i vettori e_1, e_2 non possono formare in nessun ordine una base di \mathbb{R}^3 , poiché, pur essendo linearmente indipendenti, non sono generatori di \mathbb{R}^3 (si veda l'Esempio 13.4): infatti

$$\mathcal{L}(e_1, e_2) = \{ (x, y, z) \in \mathbb{R}^3 \mid z = 0 \}.$$

In particolare $\mathcal{B} = (e_1, e_2)$ è una base del sottospazio $\mathcal{L}(e_1, e_2) \subseteq \mathbb{R}^3$, ma non di \mathbb{R}^3 . Poiché si ha $(x, y, 0) = xe_1 + ye_2$, allora

$$[(x, y, 0)]_{\mathcal{B}} = (x, y) \in \mathbb{R}^2.$$

Ovviamente anche $\mathcal{B}' = (e_2, e_1)$ è base dello stesso sottospazio ma si osservi che

$$[(x, y, 0)]_{\mathcal{B}'} = (y, x) \in \mathbb{R}^2.$$

Si noti che anche i vettori e_1, e_2 e $e = (1, 1, 0)$ sono generatori dello stesso sottospazio $\mathcal{L}(e_1, e_2) \subseteq \mathbb{R}^3$, poiché si ha

$$(x, y, 0) = xe_1 + ye_2 + 0e = (x - y)e_1 + 0e_2 + ye = 0e_1 + (y - x)e_2 + xe,$$

ma in nessun ordine essi possono formare una base di tale sottospazio essendo linearmente dipendenti: infatti $e = e_1 + e_2$.

Più in generale sia \mathbb{K} un campo. Allora l'insieme ordinato $\mathcal{C} = (e_1, \dots, e_n)$ di vettori di \mathbb{K}^n è una base, detta *base canonica di \mathbb{K}^n* . Scelto $x = (x_1, \dots, x_n) \in \mathbb{K}^n$, si ha

$$\begin{aligned} & x_1(1, 0, 0, 0, \dots, 0, 0) + \\ & x_2(0, 1, 0, 0, \dots, 0, 0) + \\ & x_3(0, 0, 1, 0, \dots, 0, 0) + \\ & \quad \vdots \\ & x_{n-1}(0, 0, 0, 0, \dots, 1, 0) + \\ & x_n(0, 0, 0, 0, \dots, 0, 1) = \\ & \hline & (x_1, x_2, x_3, \dots, x_{n-1}, x_n) \end{aligned}$$

risulta $x = x_1e_1 + \dots + x_n e_n$, quindi

$$[(x_1, \dots, x_n)]_{\mathcal{C}} = (x_1, \dots, x_n) \in \mathbb{K}^n. \quad \spadesuit$$

Esempio 14.8. In $\mathbb{C}^{2,2}$ si considerino i vettori

$$E_{1,1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_{1,2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad E_{2,1} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad E_{2,2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

L'insieme ordinato $\mathcal{B} = (E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2})$ è una base di $\mathbb{C}^{2,2}$ (si vedano gli Esempi 13.6 e 13.14). Inoltre, se

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

risulta

$$[A]_{\mathcal{B}} = (a_{1,1}, a_{1,2}, a_{2,1}, a_{2,2}) \in \mathbb{C}^4,$$

cioè la base \mathcal{B} permette in un certo senso di “stirare” gli elementi di $\mathbb{C}^{2,2}$ trasformandoli in elementi di \mathbb{C}^4 .

Più in generale nello spazio vettoriale $\mathbb{K}^{m,n}$ l’insieme ordinato

$$\mathcal{B} = (E_{i,j} | 1 \leq i \leq m, 1 \leq j \leq n)$$

è una base, e se $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$, si ha $[A]_{\mathcal{B}} = (a_{i,j} | 1 \leq i \leq m, 1 \leq j \leq n) \in \mathbb{K}^{m,n}$. ♠

Il metodo degli scarti ci permette anche di dimostrare la seguente proposizione.

Proposizione 14.9.

Sia $V \neq \{0_V\}$ uno spazio vettoriale su un campo \mathbb{K} . Dati $v_1, \dots, v_n \in V$, si ha:

- (i) se v_1, \dots, v_n sono generatori di V (e, quindi, V è finitamente generato), allora esiste una base \mathcal{B} di V i cui elementi sono vettori dell’insieme $\{v_1, \dots, v_n\}$;
- (ii) se V è finitamente generato e v_1, \dots, v_n sono linearmente indipendenti allora esiste una base \mathcal{B} di V i cui primi n vettori sono v_1, \dots, v_n .

Dimostrazione.

- (i) L’affermazione è conseguenza immediata del metodo degli scarti partendo da un insieme di generatori, che possiamo supporre non nulli poiché $V \neq \{0_V\}$.
- (ii) Sappiamo che esiste un sistema di generatori di V , ad esempio $V = \mathcal{L}(u_1, \dots, u_m)$. Chiaramente

$$V = \mathcal{L}(u_1, \dots, u_m) \subseteq \mathcal{L}(v_1, \dots, v_n, u_1, \dots, u_m) \subseteq V$$

quindi anche $v_1, \dots, v_n, u_1, \dots, u_m$ sono generatori di V : se applichiamo il metodo degli scarti all’insieme (ordinato) di generatori $v_1, \dots, v_n, u_1, \dots, u_m$ nessuno dei primi n vettori può essere scartato (perché nessuno di loro dipende da quelli che lo precedono essendo i vettori v_1, \dots, v_n linearmente indipendenti). Quindi l’insieme che si ottiene dopo aver applicato il metodo degli scarti a $v_1, \dots, v_n, u_1, \dots, u_m$ contiene tutti i vettori v_1, \dots, v_n e, per quanto osservato nel paragrafo precedente, esso è ancora un sistema di generatori di $V = \mathcal{L}(v_1, \dots, v_n, u_1, \dots, u_m)$: fissato un ordine otteniamo allora una base contenente i vettori dati. \square

Spesso l’enunciato precedente si riassume dicendo che *se V è uno spazio vettoriale su \mathbb{K} finitamente generato e non nullo, allora da ogni insieme di generatori si può estrarre una base e ogni insieme di vettori linearmente indipendenti può essere completato a una base.*

Inoltre la Proposizione 14.9 ha come immediata conseguenza il corollario seguente.

Corollario 14.10.

Sia $V \neq \{0_V\}$ uno spazio vettoriale finitamente generato su un campo \mathbb{K} . Allora esistono basi di V .

Esempio 14.11. In \mathbb{R}^4 si considerino i vettori $v_1 = (1, 2, -1, 1)$, $v_2 = (2, 1, 1, 0)$, $v_3 = (0, 0, 0, 0)$, $v_4 = (1, 1, 0, 0)$, $v_5 = (4, 4, 0, 1)$. Vogliamo trovare un sottoinsieme di $\{v_1, v_2, v_3, v_4, v_5\}$ costituito da vettori linearmente indipendenti. Applicando il metodo degli scarti abbiamo costruito nell'Esempio 14.1 la base $\mathcal{B} = (v_1, v_2, v_4)$ del sottospazio $\mathcal{L}(v_1, v_2, v_3, v_4, v_5)$ di \mathbb{R}^4 .

In particolare i vettori v_1, v_2, v_4 sono linearmente indipendenti in \mathbb{R}^4 , dunque è possibile completarli a una base. Consideriamo la base canonica $\mathcal{C} = (e_1, e_2, e_3, e_4)$ di \mathbb{R}^4 e applichiamo il metodo degli scarti all'insieme ordinato $v_1, v_2, v_4, e_1, e_2, e_3, e_4$: verifichiamo che $e_1 \notin \mathcal{L}(v_1, v_2, v_4)$. Infatti abbiamo visto nell'Esempio 14.1 che l'equazione vettoriale

$$\alpha_1(1, 2, -1, 1) + \alpha_2(2, 1, 1, 0) + \alpha_3(1, 1, 0, 0) = (1, 0, 0, 0)$$


non ha soluzione. Concludiamo che esiste una base di \mathbb{R}^4 i cui primi elementi sono v_1, v_2, v_4, e_1 .

Osserviamo poi che $e_2, e_3, e_4 \in \mathcal{L}(v_1, v_2, v_4, e_1)$: infatti

$$e_2 = 0(1, 2, -1, 1) + 0(2, 1, 1, 0) + (1, 1, 0, 0) - (1, 0, 0, 0),$$

$$e_3 = 0(1, 2, -1, 1) + (2, 1, 1, 0) - (1, 1, 0, 0) - (1, 0, 0, 0),$$

$$e_4 = (1, 2, -1, 1) + (2, 1, 1, 0) - 3(1, 1, 0, 0) + 0(1, 0, 0, 0),$$

quindi l'insieme $\mathcal{B}' = (v_1, v_2, v_4, e_1)$ è una base di \mathbb{R}^4 , diversa dalla base canonica $\mathcal{C} = (e_1, e_2, e_3, e_4)$. 

Dimensione di uno spazio vettoriale finitamente generato

15.1 Dimensione di uno spazio vettoriale

Abbiamo visto nel capitolo precedente come l'esistenza di una base in uno spazio vettoriale V su un campo \mathbb{K} , per esempio $\mathbb{K} = \mathbb{R}, \mathbb{C}$, permetta di sostituire a V , che può essere "complicato" da trattare, uno spazio \mathbb{K}^n , in maniera che le operazioni si conservino tramite tale identificazione: ci poniamo il problema di capire se tale numero n (che chiameremo tra poco *dimensione* dello spazio V) è in qualche modo un attributo intrinseco di V .

La definizione di dimensione è giustificata dal seguente risultato fondamentale.

Lemma 15.1 (Lemma di Steinitz).

Sia V uno spazio vettoriale su un campo \mathbb{K} . Se $v_1, \dots, v_n \in V$ sono generatori e $w_1, \dots, w_m \in V$ sono linearmente indipendenti, allora vale $m \leq n$.

Dimostrazione. Poiché i vettori $v_1, \dots, v_n \in V$ sono generatori di V , esistono scalari $a_{i,j} \in \mathbb{K}$, $1 \leq i \leq m$, $1 \leq j \leq n$, tali che

$$\begin{cases} a_{1,1}v_1 + a_{1,2}v_2 + a_{1,3}v_3 + \cdots + a_{1,n-1}v_{n-1} + a_{1,n}v_n = w_1 \\ a_{2,1}v_1 + a_{2,2}v_2 + a_{2,3}v_3 + \cdots + a_{2,n-1}v_{n-1} + a_{2,n}v_n = w_2 \\ a_{3,1}v_1 + a_{3,2}v_2 + a_{3,3}v_3 + \cdots + a_{3,n-1}v_{n-1} + a_{3,n}v_n = w_3 \\ \vdots \\ a_{m,1}v_1 + a_{m,2}v_2 + a_{m,3}v_3 + \cdots + a_{m,n-1}v_{n-1} + a_{m,n}v_n = w_m. \end{cases} \quad (15.1.1)$$

Dalla famiglia di identità vettoriali (15.1.1) otteniamo una matrice $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathbb{K}^{m,n}$; ogni operazione elementare su A equivale a un'analoga operazione sulle identità di (15.1.1).

Con un numero finito di operazioni elementari di riga di tipo E1 possiamo trasformare A in una nuova matrice ridotta per righe $A' = (a'_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathbb{K}^{m,n}$ corrispondente a delle nuove identità vettoriali aventi come termini noti certe combinazioni lineari dei vettori w_1, \dots, w_m .

Poiché ogni identità (15.1.1) contiene un vettore w_i e i vettori w_1, \dots, w_m sono linearmente indipendenti, dopo tali operazioni il termine noto dell' i -esima identità conterrà ancora il vettore w_i con coefficiente 1.

Se per assurdo fosse $m > n$, almeno una riga di A' (per fissare le idee diciamo quella di indice m) sarebbe nulla, quindi avremmo una relazione di dipendenza lineare del tipo

$$0_V = 0v_1 + 0v_2 + \cdots + 0v_n = b_1w_1 + \cdots + b_{m-1}w_{m-1} + 1w_m$$

che non è possibile se, come stiamo supponendo, i vettori w_1, \dots, w_m sono linearmente indipendenti. Concludiamo che $m \leq n$. \square

Corollario 15.2.

Sia V uno spazio vettoriale su un campo \mathbb{K} . Date due basi $\mathcal{B} = (v_1, \dots, v_n)$ e $\mathcal{D} = (w_1, \dots, w_m)$ di V , allora $m = n$.

Dimostrazione. Poiché v_1, \dots, v_n sono generatori di V e w_1, \dots, w_m sono linearmente indipendenti, per il Lemma 15.1 risulta $m \leq n$. Similmente, poiché w_1, \dots, w_m sono generatori di V e v_1, \dots, v_n sono linearmente indipendenti, ancora per il Lemma 15.1 risulta anche $m \geq n$. \square

Quindi in uno spazio vettoriale finitamente generato e non nullo tutte le basi hanno lo stesso numero di elementi. Questa osservazione ci permette di introdurre la seguente definizione.

Definizione 15.3 (Dimensione di uno spazio vettoriale).

Sia V uno spazio vettoriale finitamente generato su un campo \mathbb{K} . La *dimensione* di V è il numero $\dim_{\mathbb{K}}(V)$ definito come segue:

- se $V = \{0_V\}$, poniamo $\dim_{\mathbb{K}}(V) = 0$;
- se $V \neq \{0_V\}$, $\dim_{\mathbb{K}}(V)$ è il numero di elementi di una qualsiasi sua base;

Nel caso in cui il campo \mathbb{K} sia evidente dal contesto, si scrive semplicemente $\dim(V)$.

Se lo spazio V non è finitamente generato, allora poniamo per convenzione $\dim_{\mathbb{K}}(V) = \infty$, e V è detto spazio vettoriale di dimensione infinita.

Esempio 15.4. Dall'Esempio 14.7 segue che la base canonica $\mathcal{C} = (e_1, e_2, e_3)$ è una base di \mathbb{R}^3 , quindi $\dim_{\mathbb{R}}(\mathbb{R}^3) = 3$.

Più in generale, sia \mathbb{K} un campo. Allora la base canonica $\mathcal{C} = (e_1, \dots, e_n)$ di \mathbb{K}^n (si veda ancora l'Esempio 14.7) è formata da n vettori, dunque $\dim_{\mathbb{K}}(\mathbb{K}^n) = n$. \spadesuit

Esempio 15.5. Dall'Esempio 14.8 segue che lo spazio vettoriale $\mathbb{C}^{2,2}$ ha come base $\mathcal{B} = (E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2})$, quindi $\dim_{\mathbb{C}}(\mathbb{C}^{2,2}) = 4$.

Più in generale, se \mathbb{K} è un campo, allora $\dim_{\mathbb{K}}(\mathbb{K}^{m,n}) = mn$. \spadesuit

Esempio 15.6. Sia V uno spazio vettoriale su \mathbb{C} . Poiché $\mathbb{R} \subseteq \mathbb{C}$, allora V è anche uno spazio vettoriale su \mathbb{R} : l'operazione di somma rimane la stessa, quella di prodotto per uno scalare si ottiene restringendo quella definita in V a $\mathbb{R} \times V \subseteq \mathbb{C} \times V$.

Per esempio \mathbb{C} è uno spazio vettoriale su \mathbb{C} , $1 \in \mathbb{C}$ è linearmente indipendente e genera \mathbb{C} come spazio vettoriale su \mathbb{C} , poiché $a+bi = (a+bi)1$, dunque $\dim_{\mathbb{C}}(\mathbb{C}) = 1$, come già sappiamo.

D'altra parte \mathbb{C} è anche spazio vettoriale su \mathbb{R} . Da questo punto di vista 1 non è più generatore di \mathbb{C} su \mathbb{R} : infatti se lo fosse ogni suo elemento sarebbe combinazione lineare di 1 a coefficienti in \mathbb{R} e, in questo modo, possiamo ottenere solo i numeri complessi con parte immaginaria nulla. Per generare \mathbb{C} su \mathbb{R} occorrono almeno due elementi: per esempio 1 e i sono generatori di \mathbb{C} su \mathbb{R} , poiché $a+bi = (a)1 + (b)i$ per ogni $a, b \in \mathbb{R}$. Inoltre essi sono linearmente indipendenti: infatti se $a, b \in \mathbb{R}$, risulta $a+bi = 0$ se e solo se $a = b = 0$. Concludiamo che $\dim_{\mathbb{R}}(\mathbb{C}) = 2$.

Più in generale si può dimostrare che se $\dim_{\mathbb{C}}(V) = n$, allora $\dim_{\mathbb{R}}(V) = 2n$. \spadesuit

Concludiamo il paragrafo con la seguente conseguenza della Proposizione 14.9.

Proposizione 15.7.

Sia $V \neq \{0_V\}$ uno spazio vettoriale su un campo \mathbb{K} . Dati $v_1, \dots, v_n \in V$, se $\dim_{\mathbb{K}}(V) = n$ le seguenti affermazioni sono equivalenti:

- (i) $v_1, \dots, v_n \in V$ sono generatori di V ;
- (ii) $v_1, \dots, v_n \in V$ sono linearmente indipendenti;
- (iii) (v_1, \dots, v_n) è base di V .

Dimostrazione. Se i vettori $v_1, \dots, v_n \in V$ sono generatori di V scartando eventualmente alcuni di loro otteniamo una base di V : poiché ogni base di V è costituita da n vettori, non è necessario scartare niente, cioè v_1, \dots, v_n sono già linearmente indipendenti.

Viceversa, se i vettori $v_1, \dots, v_n \in V$ sono linearmente indipendenti, aggiungendo eventualmente altri vettori otteniamo una base di V : poiché ogni base di V è costituita da n vettori, non è necessario aggiungere niente, cioè v_1, \dots, v_n sono già generatori di V .

Quindi (i) vale se e solo se vale (ii), dunque sono equivalenti a (iii). \square

Esempio 15.8. In \mathbb{R}^4 si considerino i vettori $v_1 = (1, 2, \pi, 0)$, $v_2 = (3/2, -117, 0, \sqrt{2})$, $v_3 = (0, 1, 0, -3/7)$, $v_4 = (0, 3, 0, 0)$. La relazione di dipendenza lineare

$$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4 = 0_{\mathbb{R}^4}$$

tra di loro si traduce immediatamente nel sistema omogeneo

$$\begin{cases} \alpha_1 + \frac{3}{2}\alpha_2 = 0 \\ 2\alpha_1 - 117\alpha_2 + \alpha_3 + 3\alpha_4 = 0 \\ \pi\alpha_1 = 0 \\ \sqrt{2}\alpha_2 - \frac{3}{7}\alpha_3 = 0. \end{cases}$$

È facile vedere che l'unica soluzione di tale sistema è $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$, cioè i vettori v_1, v_2, v_3, v_4 sono linearmente indipendenti. Per la Proposizione 15.7 concludiamo che essi sono anche generatori di \mathbb{R}^4 senza doverlo verificare direttamente. In particolare $\mathcal{B} = (v_1, v_2, v_3, v_4)$ è una base di \mathbb{R}^4 . \spadesuit

15.2 Dimensione di sottospazi

Sia V uno spazio vettoriale finitamente generato su un campo \mathbb{K} e sia $W \subseteq V$ un suo sottospazio: in questo paragrafo ci occuperemo di rispondere alle domande naturali. W è finitamente generato? Se lo è, ci sono legami tra $\dim_{\mathbb{K}}(W)$ e $\dim_{\mathbb{K}}(V)$?

Proposizione 15.9.

Sia V uno spazio vettoriale finitamente generato su un campo \mathbb{K} . Allora ogni sottospazio $W \subseteq V$ è finitamente generato. Inoltre risulta $\dim_{\mathbb{K}}(W) \leq \dim_{\mathbb{K}}(V)$ e $\dim_{\mathbb{K}}(W) = \dim_{\mathbb{K}}(V)$ se e solo se $W = V$.

Dimostrazione. Se $W = \{ 0_V \}$ la tesi è ovviamente verificata.

Supponiamo che $W \neq \{ 0_V \}$, sicché $V \neq \{ 0_V \}$: sia $n = \dim_{\mathbb{K}}(V) \geq 1$. Innanzi tutto osserviamo che vettori di W linearmente indipendenti in W lo sono anche in V . Pertanto ogni insieme di vettori linearmente indipendenti in W contiene al massimo n elementi a due a due distinti.

Se $\{ w_1, \dots, w_m \} \subseteq W$ è un insieme massimale di vettori linearmente indipendenti in W , per ogni $w \in W$ i vettori w_1, \dots, w_m, w sono linearmente dipendenti, dunque $w \in \mathcal{L}(w_1, \dots, w_m)$ per la Proposizione 14.9. In particolare $W \subseteq \mathcal{L}(w_1, \dots, w_m) \subseteq W$ e, di conseguenza, $W = \mathcal{L}(w_1, \dots, w_m)$ che, dunque, è finitamente generato con base (w_1, \dots, w_m) , da cui $m = \dim_{\mathbb{K}}(W) \leq \dim_{\mathbb{K}}(V) = n$.

Se $m = n$ allora w_1, \dots, w_m sono n vettori linearmente indipendenti in $W \subseteq V$, dunque sono generatori di V per la Proposizione 15.7. Segue che $W = \mathcal{L}(w_1, \dots, w_m) = V$. \square

Esempio 15.10. In \mathbb{R}^4 si considerino i vettori $v_1 = (1, 2, -1, 1)$, $v_2 = (2, 1, 1, 0)$, $v_3 = (0, 0, 0, 0)$, $v_4 = (1, 1, 0, 0)$, $v_5 = (4, 4, 0, 1)$. Calcoliamo la dimensione del sottospazio $W = \mathcal{L}(v_1, v_2, v_3, v_4, v_5) \subseteq \mathbb{R}^4$. Abbiamo visto nell'Esempio 14.11 che $\mathcal{B} = (v_1, v_2, v_4)$ è una base di $W \subset \mathbb{R}^4$, quindi $\dim(W) = 3 < 4$: in particolare $W \subsetneq \mathbb{R}^4$. \spadesuit

Esempio 15.11. Sia $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathbb{K}^{m,n}$. Abbiamo verificato nell'Esempio 12.14 che l'insieme W delle soluzioni in $\mathbb{K}^{n,p}$ del sistema omogeneo $AX = 0_{m,p}$ è un sottospazio. Poiché $\mathbb{K}^{n,p}$ è finitamente generato, è lecito calcolare la dimensione di W . Poiché la dimensione di W dipende solo dal sottospazio W e non dalla matrice A , possiamo supporre che A sia fortemente ridotta per righe e, per fissare le idee, possiamo anche supporre che il suo pivot sulla i -esima riga si trovi nella colonna i -esima.

Quindi, posto $r = \text{rk}(A)$, il sistema $AX = 0_{m,p}$ è della forma

$$\begin{pmatrix} 1 & 0 & \dots & 0 & a_{1,r+1} & a_{1,r+2} & \dots & a_{1,n} \\ 0 & 1 & \dots & 0 & a_{2,r+1} & a_{2,r+2} & \dots & a_{2,n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & a_{r,r+1} & a_{r,r+2} & \dots & a_{r,n} \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix} X = 0_{m,p}. \quad (15.2.1)$$

Indicando con X_i la i -esima riga di X , segue che le soluzioni del sistema (15.2.1) dipendono dai parametri (vettoriali) liberi X_{r+1}, \dots, X_n .

Per semplicità limitiamoci, d'ora in poi, a studiare il caso $p = 1$, cioè $X \in \mathbb{K}^{n,1}$. Le soluzioni del sistema (15.2.1) sono allora tutte e sole le matrici colonna della

forma

$$\begin{pmatrix} -a_{1,r+1}X_{r+1} - a_{1,r+2}X_{r+2} - \cdots - a_{1,n}X_n \\ -a_{2,r+1}X_{r+1} - a_{2,r+2}X_{r+2} - \cdots - a_{2,n}X_n \\ \vdots \\ -a_{r,r+1}X_{r+1} - a_{r,r+1}X_{r+2} - \cdots - a_{r,n}X_n \\ X_{r+1} \\ X_{r+2} \\ \vdots \\ X_n \end{pmatrix} = \sum_{j=r+1}^n C_j X_j,$$

dove

$$C_{r+1} = \begin{pmatrix} -a_{1,r+1} \\ -a_{2,r+1} \\ \vdots \\ -a_{r,r+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad C_{r+2} = \begin{pmatrix} -a_{1,r+2} \\ -a_{2,r+2} \\ \vdots \\ -a_{r,r+2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots \quad C_n = \begin{pmatrix} -a_{1,n} \\ -a_{2,n} \\ \vdots \\ -a_{r,n} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

In particolare $W = \mathcal{L}(C_{r+1}, C_{r+2}, \dots, C_n)$. Inoltre la relazione di dipendenza lineare

$$\alpha_{r+1}C_{r+1} + \alpha_{r+2}C_{r+2} + \cdots + \alpha_n C_n = 0_{\mathbb{K}^{n,1}}$$

si traduce in un sistema le cui ultime $n - r$ equazioni sono

$$\alpha_{r+1} = \alpha_{r+2} = \cdots = \alpha_n = 0.$$

Concludiamo che $C_{r+1}, C_{r+2}, \dots, C_n$ sono generatori linearmente indipendenti di W che, quindi, ha dimensione $\dim_{\mathbb{K}}(W) = n - r = n - \text{rk}(A)$. In particolare il rango di una matrice dipende solo dalla matrice stessa e non dalle operazioni elementari fatte su di essa per calcolarlo, come già anticipato senza dimostrazione nella Lezione 3!

Se, invece, $p \geq 2$, il lettore verifichi che $\dim_{\mathbb{K}}(W) = (n - \text{rk}(A))p$ seguendo lo stesso procedimento. ♠

Esempio 15.12. Si consideri il sottoinsieme di $\mathbb{K}^{n,n}$ costituito dalle matrici triangolari superiori:

$$TS_n(\mathbb{K}) = \{ A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathbb{K}^{n,n} \mid a_{i,j} = 0 \text{ se } i > j \}.$$

Osserviamo che $TS_n(\mathbb{K})$ è un sottospazio di $\mathbb{K}^{n,n}$; infatti la matrice nulla $0_{n,n} \in TS_n(\mathbb{K})$. Inoltre se $A = (a_{i,j})_{1 \leq i,j \leq n}$ e $B = (b_{i,j})_{1 \leq i,j \leq n} \in TS_n(\mathbb{K})$ e $\lambda \in \mathbb{K}$, le entrate di indici (i, j) di $A + B$ e di λA sono $a_{i,j} + b_{i,j}$ e $\lambda a_{i,j}$, che sono dunque nulle se $i > j$, cioè $A + B$ e λA sono triangolari superiori.

Poiché $\mathbb{K}^{n,n}$ è finitamente generato, tale deve essere $TS_n(\mathbb{K})$: calcoliamone la dimensione. Si noti che $E_{i,j} \in TS_n(\mathbb{K})$ per ogni $i, j = 1, \dots, n$ con $i \leq j$. Tali matrici sono linearmente indipendenti e possiamo scrivere una qualsiasi matrice

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots \\ 0 & a_{2,2} & a_{2,3} & \cdots \\ 0 & 0 & a_{3,3} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

come combinazione lineare

$$A = a_{1,1}E_{1,1} + a_{1,2}E_{1,2} + a_{1,3}E_{1,3} + \cdots + a_{2,2}E_{2,2} + a_{2,3}E_{2,3} + \cdots + a_{3,3}E_{3,3} + \cdots,$$

ovvero $TS_n(\mathbb{K}) = \mathcal{L}(E_{i,j} \mid i, j = 1, \dots, n, i < j)$.

Dunque, fissato un ordine, tali matrici formano una base di $TS_n(\mathbb{K})$: in particolare

$$\dim_{\mathbb{K}}(TS_n) = n + (n-1) + (n-2) + \cdots + 2 + 1 = n(n+1)/2.$$

Si considerino i sottoinsiemi di $\mathbb{K}^{n,n}$ costituiti dalle matrici triangolari inferiori, strettamente triangolari superiori, strettamente triangolari inferiori, cioè rispettivamente

$$\begin{aligned} TI_n(\mathbb{K}) &= \{ A = (a_{i,j})_{1 \leq i, j \leq n} \in \mathbb{K}^{n,n} \mid a_{i,j} = 0 \text{ se } i < j \}, \\ STS_n(\mathbb{K}) &= \{ A = (a_{i,j})_{1 \leq i, j \leq n} \in \mathbb{K}^{n,n} \mid a_{i,j} = 0 \text{ se } i \geq j \}, \\ STI_n(\mathbb{K}) &= \{ A = (a_{i,j})_{1 \leq i, j \leq n} \in \mathbb{K}^{n,n} \mid a_{i,j} = 0 \text{ se } i \leq j \}. \end{aligned}$$

Il lettore dimostri che $STS_n(\mathbb{K})$, $TI_n(\mathbb{K})$ e $STI_n(\mathbb{K})$ sono sottospazi di $\mathbb{K}^{n,n}$, determinandone delle basi e verificando che

$$\dim_{\mathbb{K}}(TI_n) = n(n+1)/2, \quad \dim_{\mathbb{K}}(STS_n) = \dim_{\mathbb{K}}(STI_n) = n(n-1)/2. \quad \spadesuit$$

Esempio 15.13. L'esempio precedente è un caso particolarmente fortunato in cui l'intersezione di una base dello spazio vettoriale V con il sottospazio W , dà una base di W : purtroppo non è sempre così. Si consideri il sottoinsieme di $\mathbb{K}^{n,n}$ costituito dalle matrici simmetriche, cioè

$$Sim_n(\mathbb{K}) = \{ A \in \mathbb{K}^{n,n} \mid {}^tA = A \}.$$

Osserviamo che $Sim_n(\mathbb{K})$ è un sottospazio di $\mathbb{K}^{n,n}$. Infatti $0_{n,n} \in Sim_n(\mathbb{K})$; se $A, B \in Sim_n(\mathbb{K})$ e $\lambda \in \mathbb{K}$, per le proprietà che legano la trasposizione alla somma di matrici e al prodotto di matrici per scalari, si ha

$$A + B = {}^tA + {}^tB = {}^t(A + B), \quad \lambda A = \lambda {}^tA = {}^t(\lambda A),$$

cioè $A + B, \lambda A \in Sim_n(\mathbb{K})$.

Poiché $\mathbb{K}^{n,n}$ è finitamente generato, tale deve essere $Sim_n(\mathbb{K})$. Per calcolare la sua dimensione restringiamoci, per semplicità, al caso $n = 2$, $\mathbb{K} = \mathbb{R}$: tratteremo il caso generale più avanti. In questo caso sappiamo che una base di $\mathbb{R}^{2,2}$ è data da $\mathcal{B} = (E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2})$. Da un lato osserviamo che non ogni matrice di $\mathbb{R}^{2,2}$ è simmetrica, dunque sicuramente

$$\dim_{\mathbb{R}}(Sim_2(\mathbb{R})) \leq 3 < 4 = \dim_{\mathbb{R}}(\mathbb{R}^{2,2}).$$

Dall'altro gli elementi della base \mathcal{B} che sono in $Sim_2(\mathbb{R})$ sono $E_{1,1}$ ed $E_{2,2}$, che sono linearmente indipendenti, dunque $\dim_{\mathbb{R}}(Sim_2(\mathbb{R})) \geq 2$.

Se fosse $\dim_{\mathbb{R}}(Sim_2(\mathbb{R})) = 2$, per la Proposizione 15.7 seguirebbe che $E_{1,1}$ ed $E_{2,2}$ sarebbero generatori di $(Sim_2(\mathbb{R}))$. Tuttavia poiché

$$a_{1,1}E_{1,1} + a_{2,2}E_{2,2} = \begin{pmatrix} a_{1,1} & 0 \\ 0 & a_{2,2} \end{pmatrix}$$

è chiaro che $E_{1,1}$ ed $E_{2,2}$ non sono sufficienti a generare $Sim_2(\mathbb{R})$: concludiamo che $\dim_{\mathbb{R}}(Sim_2(\mathbb{R})) \geq 3$ e, quindi, $\dim_{\mathbb{R}}(Sim_2(\mathbb{R})) = 3$.

Per costruire una base di $Sim_2(\mathbb{R})$ è sufficiente trovare allora tre matrici simmetriche linearmente indipendenti: due, $E_{1,1}$ ed $E_{2,2}$, le abbiamo già, quindi basta determinare una terza matrice simmetrica che non sia in $\mathcal{L}(E_{1,1}, E_{2,2})$, cioè che non sia diagonale. Posto allora

$$\bar{E}_{1,2} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

segue che $\mathcal{B}' = (E_{1,1}, E_{2,2}, \bar{E}_{1,2})$ è base di $Sim_2(\mathbb{R})$. ♠

⚠ Vedremo più avanti che $\dim_{\mathbb{K}}(Sim_n(\mathbb{K})) = \dim_{\mathbb{K}}(TS_n(\mathbb{K})) = n(n+1)/2$. Si noti però che $Sim_n(\mathbb{K}) \neq TS_n(\mathbb{K})$ pur avendo la stessa dimensione, infatti non è vero che due sottospazi della stessa dimensione coincidono!

15.3 Rango e dimensione

Proposizione 15.14.

Siano \mathbb{K} un campo e $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathbb{K}^{m,n}$. Posto

$$\begin{aligned} v_i &= (a_{i,1}, \dots, a_{i,n}) \in \mathbb{K}^n, & i &= 1, \dots, m, \\ w_j &= (a_{1,j}, \dots, a_{m,j}) \in \mathbb{K}^m, & j &= 1, \dots, n, \end{aligned}$$

risulta

$$\text{rk}(A) = \dim_{\mathbb{K}}(\mathcal{L}(v_1, \dots, v_m)) = \dim_{\mathbb{K}}(\mathcal{L}(w_1, \dots, w_n)) = \text{rk}({}^t A).$$

Dimostrazione. Nel paragrafo 4.2 abbiamo definito il rango di una matrice $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathbb{K}^{m,n}$ come il numero di righe non nulle di una matrice $A' = (a'_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathbb{K}^{m,n}$ ridotta per righe ed equivalente per righe ad A .

Per dimostrare la prima parte della tesi è sufficiente far vedere che ogni tipo di operazione elementare di riga, pur cambiando le righe di A , non muta il sottospazio che esse generano. Poiché esse coinvolgono al massimo due righe per volta, possiamo ridurci a studiare il caso $m = 2$.

È chiaro che $\mathcal{L}(v_1, v_2, \dots) = \mathcal{L}(v_2, v_1, \dots)$, cioè lo scambio di righe (operazione elementare E3) non muta la dimensione.

Poiché $v = \alpha_1 v_1 + \alpha_2 v_2 + \dots$ se e solo se per ogni $\lambda \in \mathbb{K} \setminus \{0\}$ si ha $v = \alpha_1 \lambda^{-1}(\lambda v_1) + \alpha_2 v_2 + \dots$, segue che $\mathcal{L}(v_1, v_2, \dots) = \mathcal{L}(\lambda v_1, v_2, \dots)$, cioè la moltiplicazione di una riga per uno scalare non nullo (operazione elementare E2) non muta la dimensione.

Infine $v = \alpha_1 v_1 + \alpha_2 v_2 + \dots$ se e solo se per ogni $\lambda \in \mathbb{K}$ si ha $v = \alpha_1(v_1 + \lambda v_2) + (\alpha_2 - \lambda \alpha_1)v_2 + \dots$ e quindi $\mathcal{L}(v_1, v_2, \dots) = \mathcal{L}(v_1 + \lambda v_2, v_2, \dots)$, cioè sommare a una riga un multiplo di un'altra (operazione elementare E1) non muta la dimensione.

Posto $v'_i = (a'_{i,1}, \dots, a'_{i,n}) \in \mathbb{K}^n$, $i = 1, \dots, m$, concludiamo che

$$\dim_{\mathbb{K}}(\mathcal{L}(v_1, \dots, v_m)) = \dim_{\mathbb{K}}(\mathcal{L}(v'_1, \dots, v'_m)) = \text{rk}(A') = \text{rk}(A).$$

Poiché le righe di ${}^t A$ coincidono con le colonne di A , il ragionamento fatto sopra dimostra anche che $\dim_{\mathbb{K}}(\mathcal{L}(w_1, \dots, w_n)) = \text{rk}({}^t A)$.

Sia $r = \text{rk}(A)$. Allora è possibile estrarre dall'insieme v_1, \dots, v_m un sottoinsieme di r vettori linearmente indipendenti che generano $\mathcal{L}(v_1, \dots, v_m)$: per fissare le idee supponiamo che essi siano v_1, \dots, v_r . Allora abbiamo delle relazioni della forma

$$\begin{cases} v_i = v_i & i = 1, \dots, r \\ v_{r+1} = \alpha_{r+1,1}v_1 + \alpha_{r+1,2}v_2 + \alpha_{r+1,3}v_3 + \dots + \alpha_{r+1,r}v_r \\ \vdots \\ v_m = \alpha_{m,1}v_1 + \alpha_{m,2}v_2 + \alpha_{m,3}v_3 + \dots + \alpha_{m,r}v_r. \end{cases}$$

Eguagliando le componenti j -esime al primo e al secondo membro di tali equazioni, otteniamo

$$\begin{cases} a_{i,j} = a_{i,j} & i = 1, \dots, r \\ a_{r+1,j} = \alpha_{r+1,1}a_{1,j} + \alpha_{r+1,2}a_{2,j} + \alpha_{r+1,3}a_{3,j} + \dots + \alpha_{r+1,r}a_{r,j} \\ \vdots \\ a_{m,j} = \alpha_{m,1}a_{1,j} + \alpha_{m,2}a_{2,j} + \alpha_{m,3}a_{3,j} + \dots + \alpha_{m,r}a_{r,j}. \end{cases}$$

Posto

$$\begin{aligned} u_1 &= (1, 0, 0, \dots, 0, \alpha_{r+1,1}, \dots, \alpha_{m,1}), \\ u_2 &= (0, 1, 0, \dots, 0, \alpha_{r+1,2}, \dots, \alpha_{m,2}), \\ u_3 &= (0, 0, 1, \dots, 0, \alpha_{r+1,3}, \dots, \alpha_{m,3}), \\ &\vdots \\ u_r &= (0, 0, 0, \dots, r, \alpha_{r+1,r}, \dots, \alpha_{m,r}), \end{aligned}$$

segue allora che $w_j = a_{1,j}u_1 + a_{2,j}u_2 + a_{3,j}u_3 + \dots + a_{r,j}u_r$, cioè $\mathcal{L}(w_1, \dots, w_n) \subseteq \mathcal{L}(u_1, \dots, u_r)$, dunque

$$\text{rk}({}^t A) = \dim_{\mathbb{K}}(\mathcal{L}(w_1, \dots, w_n)) \subseteq \dim_{\mathbb{K}}(\mathcal{L}(u_1, \dots, u_r)) \leq r = \text{rk}(A).$$

Sostituendo ora A con ${}^t A$ e ripetendo il ragionamento otteniamo $\text{rk}(A) = \text{rk}({}^{tt} A) \leq \text{rk}({}^t A)$ sicché, in conclusione, risulta $\text{rk}(A) = \text{rk}({}^t A)$ e, perciò, la tesi è completamente dimostrata. \square

Osservazione 15.15. Quasi sempre la definizione di rango di una matrice viene data utilizzando la proprietà sopra, che non dipende dalla riduzione operata, dicendo

che “il rango di $A \in \mathbb{K}^{m,n}$ è la dimensione del suo spazio riga, cioè del sottospazio di \mathbb{K}^n generato dalle sue righe, e del suo spazio colonna, cioè del sottospazio di \mathbb{K}^m generato dalle sue colonne”.

Poiché tale dimensione si calcola a partire da una base dello spazio riga, cioè da un insieme massimale di vettori linearmente indipendenti dello spazio riga, talvolta si dice anche che “il rango di $A \in \mathbb{K}^{m,n}$ è il massimo numero di sue righe o colonne linearmente indipendenti”.

Vogliamo ora utilizzare la Proposizione 15.14, descrivendo un altro metodo per determinare basi di uno spazio vettoriale V finitamente generato su un campo \mathbb{K} a partire da un suo insieme di generatori o di suoi vettori linearmente indipendenti.

A tale scopo fissiamo una base $\mathcal{B} = (w_1, \dots, w_n)$ di V e vettori $v_1, \dots, v_m \in V$. Supponiamo che $[v_i]_{\mathcal{B}} = (a_{i,1}, \dots, a_{i,n}) \in \mathbb{K}^n$ per $i = 1, \dots, m$. Poiché

$$[\alpha_1 v_1 + \dots + \alpha_m v_m]_{\mathcal{B}} = \alpha_1 [v_1]_{\mathcal{B}} + \dots + \alpha_m [v_m]_{\mathcal{B}},$$

allora v_1, \dots, v_m sono linearmente dipendenti se e solo la stessa proprietà vale per $[v_1]_{\mathcal{B}}, \dots, [v_m]_{\mathcal{B}} \in \mathbb{K}^n$. Quindi $\dim_{\mathbb{K}}(\mathcal{L}(v_1, \dots, v_m))$ coincide con il rango della matrice $A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ avente come riga i -esima le componenti del vettore v_i rispetto alla base \mathcal{B} fissata.

Illustriamo questa osservazione con un paio di esempi.

Esempio 15.16. In \mathbb{R}^5 siano dati i vettori $v_1 = (1, 1, 2, -1, 3)$, $v_2 = (2, -1, -1, 0, 1)$, $v_3 = (0, 2, -1, 0, 1)$. Fissata la base canonica di \mathbb{R}^5 , la matrice A definita sopra è

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

che è ridotta per righe. Quindi v_1, v_2, v_3 sono linearmente indipendenti. ♠

Esempio 15.17. Si considerino in $\mathbb{C}^{2,2}$ le matrici

$$\begin{aligned} A_1 &= \begin{pmatrix} 1 & 1+i \\ i-2 & i^{-1} \end{pmatrix}, & A_2 &= \begin{pmatrix} i/2 & -3 \\ i-2 & 0 \end{pmatrix}, \\ A_3 &= \begin{pmatrix} 0 & -1+2i \\ 3i-2 & 0 \end{pmatrix}, & A_4 &= \begin{pmatrix} 0 & 1+2i \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

Sia $\mathcal{B} = (E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2})$ la base di $\mathbb{C}^{2,2}$ definita nell'Esempio 14.8. Allora

$$\begin{aligned} [A_1]_{\mathcal{B}} &= (1, 1+i, i-2, i^{-1}), \\ [A_2]_{\mathcal{B}} &= (i/2, -3, i-2, 0), \\ [A_3]_{\mathcal{B}} &= (0, -1+2i, 3i-2, 0), \\ [A_4]_{\mathcal{B}} &= (0, 1+2i, 0, 0) \end{aligned}$$

e la dimensione $\dim_{\mathbb{C}}(\mathcal{L}(A_1, A_2, A_3, A_4))$ coincide con il rango della matrice

$$\begin{pmatrix} 1 & 1+i & i-2 & i^{-1} \\ i/2 & -3 & i-2 & 0 \\ 0 & -1+2i & 3i-2 & 0 \\ 0 & 1+2i & 0 & 0 \end{pmatrix}$$

che è ridotta per righe: pertanto $\dim_{\mathbb{C}}(\mathcal{L}(A_1, A_2, A_3, A_4)) = 4 = \dim_{\mathbb{C}}(\mathbb{C}^{2,2})$. Quindi i vettori A_1, A_2, A_3, A_4 sono linearmente indipendenti e (A_1, A_2, A_3, A_4) è una base di $\mathbb{C}^{2,2}$. ♠

Quanto visto suggerisce un metodo per calcolare la dimensione del sottospazio generato da un insieme di vettori v_1, \dots, v_m linearmente indipendenti di uno spazio vettoriale V , determinarne una base e completarla a base di V .

Si può considerare la matrice A le cui righe sono le componenti dei vettori v_1, \dots, v_m rispetto a una fissata base \mathcal{B} : per calcolare la dimensione e una base del sottospazio $\mathcal{L}(v_1, \dots, v_m)$ basta allora ridurre per righe A ottenendo una matrice A' ridotta per righe e poi considerare i vettori di V le cui componenti rispetto a \mathcal{B} sono dati da tali righe.

Per completare tale base a base di V , in forza della Proposizione 15.14, basta aggiungere alle righe non nulle della matrice A' esattamente $\dim_{\mathbb{K}}(V) - \text{rk}(A)$ righe non nulle in modo che la matrice finale abbia rango $\dim_{\mathbb{K}}(V)$ e poi considerare i vettori di V aventi quelle righe come componenti rispetto alla base \mathcal{B} .

Esempio 15.18. È facile verificare che i vettori v_1, v_2, v_3, e_2, e_3 dell'Esempio 15.16 sono linearmente indipendenti. Infatti la matrice A definita sopra è

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

che è ridotta per righe e ha rango 5. In particolare $(v_1, v_2, v_3, e_2, e_3)$ è base di \mathbb{R}^5 .

A un'analogia conclusione si giungerebbe scegliendo in luogo di e_2, e_3 altri vettori, purché la matrice risultante A rimanga ridotta per righe: per esempio possiamo scegliere $v_4 = (0, 1, 0, 0, -\pi)$ e $v_5 = (0, 0, 0, 0, 17/33)$. ♠

Qualora la matrice A non fosse ridotta per righe, si potrebbe procedere analogamente riducendola prima per righe e poi studiando la matrice A' così ottenuta: infatti ogni operazione di riga sulla matrice A equivale a un'operazione sull'insieme dei vettori v_1, \dots, v_m che non cambia lo spazio $\mathcal{L}(v_1, \dots, v_m)$ ma solo l'insieme dei suoi generatori.

Esempio 15.19. In $\mathbb{R}^{2,2}$ si considerino le matrici

$$A_1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & -2 \\ 3 & 1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix}.$$

Vogliamo stabilire se A_1, A_2, A_3 sono linearmente indipendenti o meno e, nel caso lo siano, trovare una base di $\mathbb{R}^{2,2}$ che li contenga.

Si consideri la base $\mathcal{B} = (E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2})$ di $\mathbb{R}^{2,2}$: allora

$$[A_1]_{\mathcal{B}} = (1, 1, 1, 1), \quad [A_2]_{\mathcal{B}} = (1, -2, 3, 1), \quad [A_3]_{\mathcal{B}} = (2, -1, 2, 1).$$

La matrice avente come righe le componenti di tali matrici rispetto alla base \mathcal{B} è

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

Con operazioni elementari di riga otteniamo

$$A \xrightarrow{\substack{R_2 \rightarrow R_2 - R_1 \\ R_3 \rightarrow R_3 - R_1}} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & -3 & 2 & 0 \\ 1 & -2 & 1 & 0 \end{pmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -2 & 1 & 0 \\ 0 & -3 & 2 & 0 \end{pmatrix} = A'.$$

Le righe di A' sono le componenti rispetto a \mathcal{B} delle matrici

$$B_1 = A_1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 1 & -2 \\ 1 & 0 \end{pmatrix}, \quad B_3 = \begin{pmatrix} 0 & -3 \\ 2 & 0 \end{pmatrix},$$

che sono linearmente indipendenti, perchè le loro componenti rispetto a \mathcal{B} sono le righe non nulle di una matrice ridotta per righe, cioè A' .

Per determinare una base di $\mathbb{R}^{2,2}$ contenente A_1, A_2, A_3 , è sufficiente determinare una matrice $A_4 \notin \mathcal{L}(A_1, A_2, A_3)$. A tale scopo basta determinare una matrice A_4 tale che

$$\begin{aligned} [A_4]_{\mathcal{B}} &\notin \mathcal{L}((1, 1, 1, 1), (1, -2, 3, 1), (2, -1, 2, 1)) \\ &= \mathcal{L}((1, 1, 1, 1), (1, -2, 1, 0), (0, -3, 2, 0)). \end{aligned}$$

Per fare questo è sufficiente determinare una matrice 4×4 di rango 4 e le cui prime tre righe coincidano con le righe di A' , per esempio

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

Quindi scegliamo $A_4 = E_{1,2}$ e $\mathcal{B}' = (A_1, A_2, A_3, A_4)$ è la base richiesta di $\mathbb{R}^{2,2}$. ♠

Esempio 15.20. Si consideri l'insieme $\mathbb{K}[x]$ dei polinomi nell'indeterminata x a coefficienti in \mathbb{K} : nell'Esempio 13.8 abbiamo visto che esso è uno spazio vettoriale non finitamente generato. Il motivo è che il grado di un polinomio può essere "grande" a piacere; per ovviare a tale problema, per ogni $n \geq 0$, si può introdurre il sottoinsieme

$$\mathbb{K}[x]_n = \{ p(x) \in \mathbb{K}[x] \mid \deg(p) \leq n \}.$$

Osserviamo che $\mathbb{K}[x]_n$ è un sottospazio di $\mathbb{K}[x]$: infatti il polinomio nullo (che, per definizione, ha grado $-\infty$) è in $\mathbb{K}[x]_n$ per ogni $n \geq 0$. Inoltre la somma di due polinomi di grado non maggiore di n è ancora un polinomio di grado non maggiore di n e il prodotto di un polinomio di grado non maggiore di n per uno scalare $\lambda \in \mathbb{K}$ è ancora un polinomio di grado non maggiore di n .

Rispetto a $\mathbb{K}[x]$, il sottospazio $\mathbb{K}[x]_n$ ha il vantaggio di essere finitamente generato. Infatti se $p(x) \in \mathbb{K}[x]_n$ si ha

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1} + a_nx^n$$

ove $a_i \in \mathbb{K}$ per $i = 0, \dots, n$. Concludiamo che

$$\mathbb{K}[x]_n = \mathcal{L}(1, x, x^2, \dots, x^{n-1}, x^n) \subseteq \mathbb{K}[x].$$

Ci domandiamo quale sia la dimensione di $\mathbb{K}[x]_n$ (come spazio vettoriale su \mathbb{K}). Poiché per generare $\mathbb{K}[x]_n$ bastano $n + 1$ suoi elementi, precisamente i monomi $1, x, x^2, \dots, x^{n-1}, x^n$, risulta $\dim_{\mathbb{K}}(\mathbb{K}[x]_n) \leq n + 1$. Inoltre tali monomi sono linearmente indipendenti: infatti una loro combinazione lineare non è altro che un polinomio e, per definizione, un polinomio è nullo se e solo se tutti i suoi coefficienti sono nulli! Concludiamo che $\mathcal{B} = (1, x, x^2, \dots, x^{n-1}, x^n)$ è una base di $\mathbb{K}[x]_n$, quindi $\dim_{\mathbb{K}}(\mathbb{K}[x]_n) = n + 1$.

Ogni altra base di $\mathbb{K}[x]_n$ è costituita da $n + 1$ vettori. Per verificare che $n + 1$ polinomi in $\mathbb{K}[x]_n$ formano una base basta verificare che sono linearmente indipendenti per la Proposizione 15.7. Per esempio sia $a \in \mathbb{K}$ e si consideri l'insieme ordinato di $n + 1$ polinomi

$$\mathcal{B}' = (1, x - a, (x - a)^2, \dots, (x - a)^{n-1}, (x - a)^n).$$

Tali polinomi sono linearmente indipendenti poiché nessuno di loro è combinazione lineare dei precedenti: per convincersene basta osservare che il grado dell' h -esimo polinomio è $h - 1$.

Quindi se $p(x) \in \mathbb{K}[x]_n$ esistono scalari a_0, \dots, a_n tali che

$$p(x) = a_0 + a_1(x - a) + a_2(x - a)^2 + \cdots + a_{n-1}(x - a)^{n-1} + a_n(x - a)^n :$$

questo è lo sviluppo di Taylor di $p(x)$ di punto iniziale a e, nel corso di Analisi, viene solitamente osservato che $a_0 = p(a)$ e $a_h = (d^h p(x)/dx^h)(a)$ per $h \geq 1$. ♠

Esempio 15.21. In $\mathbb{R}[x]_3$ si considerino i polinomi

$$\begin{aligned} p_1(x) &= 1 + x + x^2 + x^3, & p_2(x) &= 1 - 2x + 3x^2, & p_3(x) &= x - 2x^3, \\ p_4(x) &= 2 + 4x^2 - x^3, & p_5(x) &= 3x - 2x^2 + x^3 : \end{aligned}$$

vogliamo stabilire se questi 5 polinomi sono linearmente indipendenti e, nel caso non lo siano, trovare una base di $W = \mathcal{L}(p_1(x), p_2(x), p_3(x), p_4(x), p_5(x)) \subseteq \mathbb{R}[x]_3$.

Osserviamo subito che $\dim(\mathbb{R}[x]_3) = 4$, perciò cinque polinomi devono essere linearmente dipendenti per il Lemma 15.1.

Si consideri la base $\mathcal{B} = (1, x, x^2, x^3)$ di $\mathbb{R}[x]_3$. Allora

$$\begin{aligned} [p_1(x)]_{\mathcal{B}} &= (1, 1, 1, 1), & [p_2(x)]_{\mathcal{B}} &= (1, -2, 3, 0), & [p_3(x)]_{\mathcal{B}} &= (0, 1, 0, -2), \\ [p_4(x)]_{\mathcal{B}} &= (2, 0, 4, -1), & [p_5(x)]_{\mathcal{B}} &= (0, 3, -2, 1). \end{aligned}$$

La matrice avente come righe le componenti dei polinomi dati rispetto alla base \mathcal{B} è

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

Con operazioni elementari di riga otteniamo

$$\begin{aligned} A &\xrightarrow{\substack{R_2 \rightarrow R_2 - R_1 \\ R_4 \rightarrow R_4 - 2R_1}} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & -3 & 2 & -1 \\ 0 & 1 & 0 & -2 \\ 0 & -2 & 2 & -3 \\ 0 & 3 & -2 & 1 \end{pmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & -3 & 2 & -1 \\ 0 & -2 & 2 & -3 \\ 0 & 3 & -2 & 1 \end{pmatrix} \\ &\xrightarrow{\substack{R_3 \rightarrow R_3 + 3R_2 \\ R_4 \rightarrow R_4 + 2R_2 \\ R_5 \rightarrow R_5 - 3R_2}} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 2 & -7 \\ 0 & 0 & 2 & -7 \\ 0 & 0 & -2 & 7 \end{pmatrix} \xrightarrow{\substack{R_4 \rightarrow R_4 - R_3 \\ R_5 \rightarrow R_5 + R_3}} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 2 & -7 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = A'. \end{aligned}$$

Le righe di A' sono le componenti rispetto a \mathcal{B} dei polinomi

$$q_1(x) = 1 + x + x^2 + x^3, \quad q_2(x) = x - 2x^3, \quad q_3(x) = 2x^2 - 7x^3$$

che sono linearmente indipendenti, perchè le loro componenti rispetto a \mathcal{B} sono le righe di una matrice ridotta per righe, cioè A' .

Inoltre si ha $W = \mathcal{L}(p_1(x), p_2(x), p_3(x), p_4(x), p_5(x)) = \mathcal{L}(q_1(x), q_2(x), q_3(x))$. In particolare la terna $(q_1(x), q_2(x), q_3(x))$ è una base di W , dunque $\dim(W) = 3$ e $W \subset \mathbb{R}[x]_3$.

Chi si voglia convincere dell'uguaglianza

$$\mathcal{L}(p_1(x), p_2(x), p_3(x), p_4(x), p_5(x)) = \mathcal{L}(q_1(x), q_2(x), q_3(x))$$

direttamente osservi che

$$q_1(x) = p_1(x), \quad q_2(x) = p_3(x), \quad q_3(x) = p_2(x) - p_1(x) + 3p_3(x)$$

quindi $q_1(x), q_2(x), q_3(x) \in \mathcal{L}(p_1(x), p_2(x), p_3(x), p_4(x), p_5(x))$; ne deduciamo l'inclusione $\mathcal{L}(q_1(x), q_2(x), q_3(x)) \subseteq \mathcal{L}(p_1(x), p_2(x), p_3(x), p_4(x), p_5(x))$.

Viceversa,

$$\begin{aligned} p_1(x) &= q_1(x), & p_2(x) &= q_3(x) + q_1(x) - 3q_2(x), & p_3(x) &= q_2(x), \\ p_4(x) &= q_3(x) + 2q_1(x) - 2q_2(x), & p_5(x) &= -q_3(x) + 3q_2(x) \end{aligned}$$

quindi $p_1(x), p_2(x), p_3(x), p_4(x), p_5(x) \in \mathcal{L}(q_1(x), q_2(x), q_3(x))$: in particolare abbiamo l'inclusione $\mathcal{L}(q_1(x), q_2(x), q_3(x)) \supseteq \mathcal{L}(p_1(x), p_2(x), p_3(x), p_4(x), p_5(x))$.

Per esercizio il lettore completi l'insieme $q_1(x), q_2(x), q_3(x)$ a base di $\mathbb{R}[x]_3$. ♠

15.4 La formula di Grassmann

Concludiamo il capitolo con un ultimo risultato importante.

Proposizione 15.22 (Formula di Grassmann).

Sia V uno spazio vettoriale su un campo \mathbb{K} . Se $U, W \subseteq V$ sono sottospazi finitamente generati allora

$$\dim_{\mathbb{K}}(U) + \dim_{\mathbb{K}}(W) = \dim_{\mathbb{K}}(U \cap W) + \dim_{\mathbb{K}}(U + W). \quad (15.4.1)$$

Dimostrazione. Se U e W sono finitamente generati lo stesso vale per $U + W$, dalla Proposizione 13.10. Poiché $U \cap W$ è sottospazio di W che è finitamente generato, esso stesso è finitamente generato.

Sia (v_1, \dots, v_n) una base di $U \cap W$: allora $U \cap W = \mathcal{L}(v_1, \dots, v_n)$. Poiché $v_1, \dots, v_n \in U, W$ sono vettori linearmente indipendenti, per la Proposizione 14.9 li possiamo completare a basi di U e W rispettivamente; in altre parole, esistono $u_1, \dots, u_p \in U$ e $w_1, \dots, w_q \in W$ tali che $(v_1, \dots, v_n, u_1, \dots, u_p)$ sia una base di U e $(v_1, \dots, v_n, w_1, \dots, w_q)$ sia una base di W . In particolare $u_1, \dots, u_p, w_1, \dots, w_q \notin U \cap W = \mathcal{L}(v_1, \dots, v_n)$.

Ancora dalla Proposizione 13.10 i vettori $v_1, \dots, v_n, u_1, \dots, u_p, w_1, \dots, w_q$ sono generatori di $U + W$. Verifichiamo che sono linearmente indipendenti: si consideri la relazione di dipendenza lineare

$$\alpha_1 v_1 + \dots + \alpha_n v_n + \beta_1 u_1 + \dots + \beta_p u_p + \gamma_1 w_1 + \dots + \gamma_q w_q = 0_V.$$

Poiché $\alpha_1 v_1 + \dots + \alpha_n v_n + \beta_1 u_1 + \dots + \beta_p u_p \in U$ e $-(\gamma_1 w_1 + \dots + \gamma_q w_q) \in W$ coincidono, essi appartengono a $U \cap W$: in particolare esistono $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tali che

$$\alpha_1 v_1 + \dots + \alpha_n v_n + \beta_1 u_1 + \dots + \beta_p u_p = -(\gamma_1 w_1 + \dots + \gamma_q w_q) = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

Dall'uguaglianza $-(\gamma_1 w_1 + \dots + \gamma_q w_q) = \lambda_1 v_1 + \dots + \lambda_n v_n$ segue

$$\lambda_1 v_1 + \dots + \lambda_n v_n + \gamma_1 w_1 + \dots + \gamma_q w_q = 0_V :$$

poiché $(v_1, \dots, v_n, w_1, \dots, w_q)$ è una base di W si ha $\lambda_1 = \dots = \lambda_n = \gamma_1 = \dots = \gamma_q = 0$, perciò $\alpha_1 v_1 + \dots + \alpha_n v_n + \beta_1 u_1 + \dots + \beta_p u_p = 0_V$: d'altra parte $(v_1, \dots, v_n, u_1, \dots, u_p)$ è una base di U , dunque si ha anche $\alpha_1 = \dots = \alpha_n = \beta_1 = \dots = \beta_p = 0$. Concludiamo che $\dim_{\mathbb{K}}(U) = n + p$, $\dim_{\mathbb{K}}(W) = n + q$, $\dim_{\mathbb{K}}(U \cap W) = n$ e $\dim_{\mathbb{K}}(U + W) = n + p + q$, che implica la formula (15.4.1). \square

Osservazione 15.23. Se riscriviamo la Formula di Grassmann (15.4.1) come

$$\dim_{\mathbb{K}}(U + W) = \dim_{\mathbb{K}}(U) + \dim_{\mathbb{K}}(W) - \dim_{\mathbb{K}}(U \cap W),$$

segue che la dimensione della somma di due sottospazi è sempre minore o uguale alla somma delle dimensioni. L'uguaglianza vale se e solo se risulta $\dim_{\mathbb{K}}(U \cap W) = 0$, se e solo se $U \cap W = \{0_V\}$. Quando questo accade, si parla di *somma diretta* di due sottospazi, e si indica con $U \oplus W$ (invece che solo $U + W$).

Esempio 15.24. Si considerino $v_1 = (1, 2, 3, 4)$, $v_2 = (1, -1, 0, 5)$, $v_3 = (2, 1, 3, 3)$, $v_4 = (1, 1, 1, 1)$, $v_5 = (1, -1, 2, -5)$ in \mathbb{R}^4 e i sottospazi

$$V = \mathcal{L}(v_1, v_2, v_3), \quad W = \mathcal{L}(v_4, v_5) \subseteq \mathbb{R}^4.$$

Il lettore verifichi per esercizio che $\dim(V) = 3$ e $\dim(W) = 2$. Allora $V \cap W$ contiene vettori non nulli. Infatti applicando la formula di Grassmann

$$\dim(V \cap W) = \dim(V) + \dim(W) - \dim(V + W) = 3 + 2 - \dim(V + W).$$

Poiché $V + W \subseteq \mathbb{R}^4$ si ha $\dim(V + W) \leq \dim(\mathbb{R}^4) = 4$, quindi $\dim(V \cap W) \geq 1$, quindi $V \cap W \neq \{0_{\mathbb{R}^4}\}$. ♠

Esempio 15.25. Osserviamo che ogni matrice $A \in \mathbb{K}^{n,n}$ può essere scritta come somma di una matrice triangolare superiore e di una matrice strettamente triangolare inferiore, cioè $\mathbb{K}^{n,n} = TS_n(\mathbb{K}) + STI_n(\mathbb{K})$. Poiché $TS_n(\mathbb{K}) \cap STI_n(\mathbb{K}) = 0_{n,n}$, dalla Proposizione 15.22 segue (si veda l'Esempio 15.12)

$$\begin{aligned} \dim_{\mathbb{K}}(STI_n(\mathbb{K})) &= \dim_{\mathbb{K}}(TS_n(\mathbb{K}) \cap STI_n(\mathbb{K})) + \dim_{\mathbb{K}}(TS_n(\mathbb{K}) + STI_n(\mathbb{K})) \\ &\quad - \dim_{\mathbb{K}}(TS_n(\mathbb{K})) = 0 + n^2 - n(n+1)/2 = n(n-1)/2. \quad \spadesuit \end{aligned}$$

Applicazioni lineari fra spazi vettoriali

16.1 Applicazioni lineari

Si ricordi che \mathbb{K} indica un campo: per esempio $\mathbb{K} = \mathbb{R}, \mathbb{C}$.

Definizione 16.1 (Applicazioni lineari).

Siano V e W spazi vettoriali su un campo \mathbb{K} . Un'applicazione $f: V \rightarrow W$ si dice \mathbb{K} -lineare se:

(AL1) per ogni $v_1, v_2 \in V$ si ha $f(v_1 + v_2) = f(v_1) + f(v_2)$;

(AL2) per ogni $\alpha \in \mathbb{K}$ e $v \in V$ si ha $f(\alpha v) = \alpha f(v)$.

Nel caso sia chiaro dal contesto quale sia il campo \mathbb{K} , si parla semplicemente di *applicazione lineare*.

Siano V e W spazi vettoriali su un campo \mathbb{K} . Sia $f: V \rightarrow W$ un'applicazione \mathbb{K} -lineare. Allora per ogni $v \in V$ si ha

$$\begin{aligned} f(0_V) &= f(0 \cdot 0_V) = 0f(0_V) = 0_W, \\ f(-v) &= f((-1)v) = (-1)f(v) = -f(v). \end{aligned}$$

Più in generale, un'applicazione $f: V \rightarrow W$ è \mathbb{K} -lineare se e solo se fissati $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ e $v_1, \dots, v_n \in V$, si ha

$$f(\alpha_1 v_1 + \dots + \alpha_n v_n) = \alpha_1 f(v_1) + \dots + \alpha_n f(v_n). \quad (16.1.1)$$

Infatti se vale (16.1.1) allora (AL1) è verificata prendendo $n = 2$ e $\alpha_1 = \alpha_2 = 1$, mentre (AL2) è verificata prendendo $n = 1$ e $\alpha_1 = \alpha$. Viceversa, se f è \mathbb{K} -lineare, allora possiamo verificare (16.1.1) per induzione su $n \geq 1$. A tale scopo osserviamo che $n = 1$ allora $f(\alpha_1 v_1) = \alpha_1 f(v_1)$ per la proprietà (AL2). Se poi $n \geq 2$ allora ponendo $w' = \alpha_1 v_1 + \dots + \alpha_{n-1} v_{n-1}$ e $w'' = \alpha_n v_n$ abbiamo

$$\begin{aligned} f(\alpha_1 v_1 + \dots + \alpha_n v_n) &= f(w' + w'') = f(w') + f(w'') \\ &= \alpha_1 f(v_1) + \dots + \alpha_{n-1} f(v_{n-1}) + f(w'') \\ &= \alpha_1 f(v_1) + \dots + \alpha_n f(v_n), \end{aligned}$$

dove le uguaglianze seguono rispettivamente dalla definizione di w' e w'' , dalla proprietà (AL1), dall'ipotesi induttiva e dalla proprietà (AL2).

Se $U \subseteq V$ è un sottospazio del dominio della funzione, allora $f|_U: U \rightarrow W$ è lineare. Se $W \subseteq U$ è un sottospazio del codominio, allora l'applicazione $\widehat{f}: V \rightarrow U$, ottenuta componendo f con l'inclusione $W \subseteq U$, è lineare.

Il lettore verifichi che l'applicazione nulla $0_{V,W}: V \rightarrow W$, definita da $v \mapsto 0_W$, e l'applicazione identità $id_V: V \rightarrow V$, definita da $v \mapsto v$, sono \mathbb{K} -lineari.

Esempio 16.2. L'applicazione

$$\begin{aligned} f: \mathbb{R}^3 &\longrightarrow \mathbb{R}^2 \\ (x, y, z) &\longrightarrow (3x + y - z, x - y + 2z) \end{aligned}$$

è lineare. Infatti, se $\alpha \in \mathbb{R}$ e $(x, y, z) \in \mathbb{R}^3$, risulta

$$\begin{aligned} f(\alpha(x, y, z)) &= f(\alpha x, \alpha y, \alpha z) \\ &= (3(\alpha x) + (\alpha y) - (\alpha z), (\alpha x) - (\alpha y) + 2(\alpha z)) \\ &= (\alpha(3x + y - z), \alpha(x - y + 2z)) \\ &= \alpha(3x + y - z, x - y + 2z) = \alpha f(x, y, z). \end{aligned}$$

Inoltre, se $(x', y', z'), (x'', y'', z'') \in \mathbb{R}^3$, risulta

$$\begin{aligned} f((x', y', z') + (x'', y'', z'')) &= f(x' + x'', y' + y'', z' + z'') \\ &= (3(x' + x'') + (y' + y'') - (z' + z''), (x' + x'') - (y' + y'') + 2(z' + z'')) \\ &= (3x' + y' - z' + 3x'' + y'' - z'', x' - y' + 2z' + x'' - y'' + 2z'') \\ &= (3x' + y' - z', x' - y' + 2z') + (3x'' + y'' - z'', x'' - y'' + 2z'') \\ &= f(x', y', z') + f(x'', y'', z''). \end{aligned}$$

Per esercizio verificare in modo analogo che anche la seguente applicazione è lineare:

$$\begin{aligned} g: \mathbb{C}^2 &\longrightarrow \mathbb{C}^3 \\ (x, y) &\longrightarrow (3x + y, x - y, -x + 2y) \end{aligned} \spadesuit$$

Esempio 16.3. Si consideri l'applicazione

$$\begin{aligned} f: \mathbb{R}^3 &\longrightarrow \mathbb{R}^2 \\ (x, y, z) &\longrightarrow (3x + y - z, x - y + 2z + 1). \end{aligned}$$

Se f fosse lineare, da quanto osservato sopra, si dovrebbe avere $f(0, 0, 0) = (0, 0)$. Poiché però $f(0, 0, 0) = (0, 1)$, segue che f non è lineare. \spadesuit

Esempio 16.4. Sia \mathbb{K} un campo e sia $A \in \mathbb{K}^{m,n}$. L'applicazione

$$\begin{aligned}\mu_A: \mathbb{K}^n &\longrightarrow \mathbb{K}^m \\ X &\longrightarrow AX\end{aligned}$$

è lineare. Infatti per qualsiasi $\alpha \in \mathbb{K}$ e $X, X', X'' \in \mathbb{K}^n$ risulta

$$\begin{aligned}\mu_A(\alpha X) &= A(\alpha X) = \alpha(AX) = \alpha\mu_A(X), \\ \mu_A(X' + X'') &= A(X' + X'') = AX' + AX'' = \mu_A(X') + \mu_A(X'').\end{aligned}$$

Verifichiamo ora che ogni applicazione lineare $f: \mathbb{K}^n \rightarrow \mathbb{K}^m$ è della forma μ_A per un'unica $A \in \mathbb{K}^{m,n}$. Sia $\mathcal{B} = (E_{1,1}, \dots, E_{n,1})$ la base definita nell'Esempio 14.8 e siano $A_j = f(E_{j,1}) \in \mathbb{K}^m$ per $j = 1, \dots, n$: definiamo allora A come la matrice $m \times n$ avente A_j come colonna j -esima. Per ogni $X = (x_{1,1}, \dots, x_{n,1}) \in \mathbb{K}^n$ risulta

$$f(X) = f\left(\sum_{j=1}^n x_{j,1} E_{j,1}\right) = \sum_{j=1}^n x_{j,1} f(E_{j,1}) = \sum_{j=1}^n x_{j,1} A_j = AX = \mu_A(X),$$

quindi $f = \mu_A$ come applicazioni.

Si noti che la matrice costruita sopra è univocamente determinata: infatti se $A' \in \mathbb{K}^{m,n}$ fosse un'altra matrice, cioè $f = \mu_{A'}$, avremmo $A_j = f(E_{j,1}) = A'E_{j,1}$ che è la j -esima colonna di A' , dunque $A = A'$.

Osserviamo che l'applicazione nulla $\mathbb{K}^n \rightarrow \mathbb{K}^m$ è $\mu_{0_{m,n}}$, mentre l'applicazione identità $\mathbb{K}^n \rightarrow \mathbb{K}^n$ è μ_{I_n} .

Si verifichi che l'applicazione f dell'Esempio 16.2 si ottiene come caso particolare prendendo $\mathbb{K} = \mathbb{R}$ e

$$A = \begin{pmatrix} 3 & 1 & -1 \\ 1 & -1 & 2 \end{pmatrix} \in \mathbb{R}^{2,3}.$$

Similmente, si verifichi che l'applicazione g dell'Esempio 16.3 coincide con μ_B , dove

$$B = \begin{pmatrix} 3 & 1 \\ 1 & -1 \\ -1 & 2 \end{pmatrix} = {}^t A \in \mathbb{C}^{3,2} \quad \spadesuit$$

Esempio 16.5. Sia V uno spazio vettoriale su un campo \mathbb{K} di dimensione n e sia \mathcal{B} una sua base. L'applicazione

$$\begin{aligned}[\cdot]_{\mathcal{B}}: V &\longrightarrow \mathbb{K}^n \\ v &\longrightarrow [v]_{\mathcal{B}}\end{aligned}$$

è lineare. Infatti se $\alpha \in \mathbb{K}$ e $v, v_1, v_2 \in V$, per quanto visto nell'Osservazione 14.5 risulta

$$[\alpha v]_{\mathcal{B}} = \alpha[v]_{\mathcal{B}}, \quad [v_1 + v_2]_{\mathcal{B}} = [v_1]_{\mathcal{B}} + [v_2]_{\mathcal{B}}. \quad \spadesuit$$

Esempio 16.6. Siano \mathbb{K} un campo e $a \in \mathbb{K}$. L'applicazione

$$\begin{aligned} v_a: \mathbb{K}[x] &\longrightarrow \mathbb{K} \\ p(x) &\longrightarrow p(a) \end{aligned}$$

di valutazione di un polinomio in un punto a è lineare. Infatti si ha che

$$\begin{aligned} v_a((p' + p'')(x)) &= (p' + p'')(a) = p'(a) + p''(a) = v_a(p'(x)) + v_a(p''(x)), \\ v_a((\alpha p)(x)) &= (\alpha p)(a) = \alpha p(a) = \alpha v_a(p(x)). \end{aligned}$$

Poichè la restrizione a sottospazi conserva la linearità, anche le applicazioni ristrette $v_{a|\mathbb{K}[x]_n}: \mathbb{K}[x]_n \rightarrow \mathbb{K}$ sono lineari per ogni $n \geq 0$. ♠

Esempio 16.7. Sia $\vec{v}_0 \in V_3(O)$ un vettore fissato. Verificare per esercizio che l'applicazione *prodotto scalare per \vec{v}_0* , definita da

$$\begin{aligned} \langle \cdot, \vec{v}_0 \rangle: V_3(O) &\longrightarrow \mathbb{R} \\ \vec{v} &\longrightarrow \langle \vec{v}, \vec{v}_0 \rangle, \end{aligned}$$

e l'applicazione *prodotto vettoriale per \vec{v}_0* , definita da

$$\begin{aligned} \cdot \times \vec{v}_0: V_3(O) &\longrightarrow V_3(O) \\ \vec{v} &\longrightarrow \vec{v} \times \vec{v}_0, \end{aligned}$$

sono entrambe \mathbb{R} -lineari grazie alle proprietà di bilinearità dei prodotti scalare e vettoriale (si veda la Proposizione 8.3). ♠

Esempio 16.8. Sia $I =]a, b[\subseteq \mathbb{R}$ un intervallo aperto non vuoto. Nell'Esempio 12.22 abbiamo osservato che l'insieme $\mathcal{C}^p(I)$ è uno spazio vettoriale su \mathbb{R} per ogni $p \geq 0$. Uno dei risultati dell'analisi delle funzioni di una variabile reale è che l'applicazione

$$\begin{aligned} D: \mathcal{C}^p(I) &\longrightarrow \mathcal{C}^{p-1}(I) \\ \varphi(x) &\longrightarrow D\varphi(x) = \frac{d\varphi}{dx}(x) \end{aligned}$$

è lineare per ogni $p \geq 1$. ♠

Esempio 16.9. Sia $z = a + bi \in \mathbb{C}$, $a, b \in \mathbb{R}$. Ricordiamo che il coniugato di un numero complesso z è, per definizione, il numero complesso $\bar{z} = a - bi$ e consideriamo l'applicazione f di *coniugio* su \mathbb{C} , cioè $f: \mathbb{C} \rightarrow \mathbb{C}$ è definita da $z \mapsto \bar{z}$. Allora f è chiaramente additiva, infatti

$$f(z' + z'') = \overline{z' + z''} = \bar{z}' + \bar{z}'' = f(z') + f(z'').$$

Invece se $\alpha \in \mathbb{C}$ si ha $f(\alpha z) = \overline{\alpha z}$ che coincide con $\alpha f(z) = \alpha \bar{z}$ se e solo se $\alpha \in \mathbb{R} \subseteq \mathbb{C}$. Quindi il coniugio non è un'applicazione \mathbb{C} -lineare, ma solo \mathbb{R} -lineare. ♠

Concludiamo questo paragrafo con la seguente proposizione.

Proposizione 16.10.

Siano U, V e W spazi vettoriali su un campo \mathbb{K} . Date le applicazioni \mathbb{K} -lineari $f, g: V \rightarrow W$ e $h: W \rightarrow U$, allora la composizione $h \circ f: V \rightarrow U$ e la somma $f + g: V \rightarrow W$ sono applicazioni lineari; lo stesso vale per il prodotto per uno scalare $\lambda f: V \rightarrow W$, per ogni $\lambda \in \mathbb{K}$.

Dimostrazione. Per ogni $v, v', v'' \in V$ risulta

$$\begin{aligned}(h \circ f)(v' + v'') &= h(f(v' + v'')) = h(f(v') + f(v'')) \\ &= h(f(v')) + h(f(v'')) = (h \circ f)(v') + (h \circ f)(v''), \\ (h \circ f)(\alpha v) &= h(f(\alpha v)) = h(\alpha f(v)) = \alpha h(f(v)) = \alpha((h \circ f)(v));\end{aligned}$$

$$\begin{aligned}(f + g)(v' + v'') &= f(v' + v'') + g(v' + v'') = f(v') + f(v'') + g(v') + g(v'') = \\ &= f(v') + g(v') + f(v'') + g(v'') = (f + g)(v') + (f + g)(v''), \\ (f + g)(\alpha v) &= f(\alpha v) + g(\alpha v) = \alpha f(v) + \alpha g(v) = \alpha(f(v) + g(v)) = \alpha((f + g)(v));\end{aligned}$$

$$\begin{aligned}(\lambda f)(v' + v'') &= \lambda f(v' + v'') = \lambda(f(v') + f(v'')) = \lambda f(v') + \lambda f(v'') = (\lambda f)(v') + (\lambda f)(v''), \\ (\lambda f)(\alpha v) &= \lambda f(\alpha v) = \lambda \alpha f(v) = \alpha(\lambda f(v)) = \alpha((\lambda f)(v)).\end{aligned} \quad \square$$

Osservazione 16.11. Siano \mathbb{K} un campo, $A, B \in \mathbb{K}^{m,n}$ e $\lambda \in \mathbb{K}$: allora si ha che

$$\mu_A + \mu_B = \mu_{A+B}, \quad \lambda \mu_A = \mu_{\lambda A}.$$

Infatti, per ogni $X \in \mathbb{K}^n$, si ha

$$\begin{aligned}(\mu_A + \mu_B)(X) &= \mu_A(X) + \mu_B(X) = AX + BX = (A + B)X = \mu_{A+B}(X), \\ (\lambda \mu_A)(X) &= \lambda(\mu_A(X)) = \lambda(AX) = (\lambda A)X = \mu_{\lambda A}(X).\end{aligned}$$

Siano $A \in \mathbb{K}^{m,n}$, $B \in \mathbb{K}^{n,p}$ e si considerino $\mu_B: \mathbb{K}^p \rightarrow \mathbb{K}^n$, $\mu_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$. Allora per ogni $X \in \mathbb{K}^p$ risulta

$$\mu_A \circ \mu_B(X) = \mu_A(\mu_B(X)) = \mu_A(BX) = A(BX) = (AB)X = \mu_{AB}(X) :$$

concludiamo che $\mu_A \circ \mu_B = \mu_{AB}$.

Esempio 16.12. Con le notazioni dell'Esempio 16.8 possiamo definire l'applicazione di derivazione n -esima $D^n = D \circ D^{n-1}: \mathcal{C}^\infty(I) \rightarrow \mathcal{C}^\infty(I)$.

Più in generale, se $a_0, \dots, a_n \in \mathcal{C}^\infty(I)$, allora

$$P(D) = a_0 D^n + a_1 D^{n-1} + \dots + a_{n-1} D + a_n$$

è \mathbb{R} -lineare. $P(D)$ si dice *operatore differenziale lineare d'ordine n* . ♠

16.2 Immagine e nucleo di un'applicazione lineare

Ricordiamo che se $\varphi: X \rightarrow Y$ è un'applicazione fra due insiemi si definisce *immagine di φ* l'insieme

$$\text{im}(\varphi) = \{ y \in Y \mid \text{esiste } x \in X \text{ tale che } \varphi(x) = y \}.$$

Proposizione 16.13.

Siano V e W spazi vettoriali su un campo \mathbb{K} . Data un'applicazione \mathbb{K} -lineare $f: V \rightarrow W$, valgono le seguenti proprietà:

- (i) $\text{im}(f)$ è un sottospazio di W ;
- (ii) se $w \in \text{im}(f)$ e $v_0 \in V$ è tale che $f(v_0) = w$, allora

$$f^{-1}(w) = \{ v \in V \mid f(v) = w \} = \{ v_0 + v' \mid v' \in f^{-1}(0_W) \} = \{ v_0 \} + f^{-1}(0_W);$$
- (iii) se $V = \mathcal{L}(v_1, \dots, v_n)$, allora $\text{im}(f) = \mathcal{L}(f(v_1), \dots, f(v_n))$. In particolare se V è finitamente generato, lo stesso vale per $\text{im}(f)$.

Dimostrazione.

- (i) Sappiamo che $0_W \in \text{im}(f)$; inoltre se $w, w', w'' \in \text{im}(f)$ e $\alpha \in \mathbb{K}$ allora esistono $v, v', v'' \in V$ tali che $f(v) = w, f(v') = w', f(v'') = w''$. Segue allora che

$$\begin{aligned} w' + w'' &= f(v') + f(v'') = f(v' + v'') \in \text{im}(f), \\ \alpha w &= \alpha f(v) = f(\alpha v) \in \text{im}(f). \end{aligned}$$

- (ii) Sia $v \in f^{-1}(w)$. Allora $f(v - v_0) = f(v) - f(v_0) = w - w = 0_W$, quindi esiste $v' \in f^{-1}(0_W)$ tale che $v = v_0 + v'$, dunque $f^{-1}(w) \subseteq \{ v_0 \} + f^{-1}(0_W)$. Viceversa, se $v' \in f^{-1}(0_W)$, risulta che $f(v_0 + v') = f(v_0) + f(v') = w + 0_W = w$, dunque si ha anche $\{ v_0 \} + f^{-1}(0_W) \subseteq f^{-1}(w)$. Concludiamo che $f^{-1}(w) = \{ v_0 \} + f^{-1}(0_W)$.
- (iii) Sia $V = \mathcal{L}(v_1, \dots, v_n)$: se $v = \sum_{i=1}^n \alpha_i v_i \in V$, allora si ha che

$$f(v) = f\left(\sum_{i=1}^n \alpha_i v_i\right) = \sum_{i=1}^n \alpha_i f(v_i) \in \mathcal{L}(f(v_1), \dots, f(v_n)),$$

dunque $\text{im}(f) \subseteq \mathcal{L}(f(v_1), \dots, f(v_n))$. Viceversa, se $w \in \mathcal{L}(f(v_1), \dots, f(v_n))$, allora

$$w = \sum_{i=1}^n \alpha_i f(v_i) = f\left(\sum_{i=1}^n \alpha_i v_i\right) \in \text{im}(f),$$

dunque vale anche l'inclusione opposta $\mathcal{L}(f(v_1), \dots, f(v_n)) \subseteq \text{im}(f)$. Concludiamo quindi che $\text{im}(f) = \mathcal{L}(f(v_1), \dots, f(v_n))$. \square

Una conseguenza immediata della Proposizione 16.13 è che se $U \subseteq V$ è un sottospazio, lo stesso vale per $f(U)$: infatti $f(U) = \text{im}(f|_U)$.

Più interessante è osservare che la controimmagine del vettore nullo 0_W riveste un ruolo fondamentale per le applicazioni \mathbb{K} -lineari: l'affermazione (ii) della Proposizione 16.13 ci dice infatti che le controimmagini $f^{-1}(w)$ tramite f dei vettori $w \in \text{im}(f)$ si ottengono tutte da quella $f^{-1}(0_W)$ di 0_W per traslazione di una controimmagine particolare.

Per questo motivo $f^{-1}(0_W)$ merita un proprio nome.

Definizione 16.14 (Nucleo di un'applicazione lineare).

Siano V e W spazi vettoriali su un campo \mathbb{K} . Data un'applicazione \mathbb{K} -lineare $f: V \rightarrow W$, si definisce *nucleo di f* l'insieme

$$\ker(f) = f^{-1}(0_W) = \{ v \in V \mid f(v) = 0_W \}.$$

16.3 Applicazioni lineari iniettive e suriettive

Ricordiamo le seguenti due definizioni, valide per un'applicazione $\varphi: X \rightarrow Y$ fra due insiemi di qualsiasi tipo. L'applicazione φ si dice *iniettiva* se, dati $x', x'' \in X$ con $x' \neq x''$, si ha $\varphi(x') \neq \varphi(x'')$. L'applicazione φ si dice invece *suriettiva* se $\text{im}(\varphi) = Y$.

Il seguente risultato è un'immediata conseguenza della Proposizione 16.13.

Proposizione 16.15.

Siano V e W spazi vettoriali su un campo \mathbb{K} . Data un'applicazione \mathbb{K} -lineare $f: V \rightarrow W$, allora f è iniettiva se e solo se $\ker(f) = \{ 0_V \}$.

Dimostrazione. Siano $f: V \rightarrow W$ un'applicazione lineare iniettiva e $v \in \ker(f)$; allora risulta $f(v) = 0_W = f(0_V)$, quindi per definizione di iniettività $v = 0_V$.

Viceversa, supponiamo che $\ker(f) = \{ 0_V \}$ e siano $v_1, v_2 \in V$ tali che $f(v_1) = f(v_2)$; allora $0_W = f(v_1) - f(v_2) = f(v_1 - v_2)$, quindi $v_1 - v_2 \in \ker(f)$, $v_1 - v_2 = 0_V$, cioè $v_1 = v_2$. \square

Il nucleo ha anche un'altra importante proprietà: infatti, così come nella Proposizione 16.13 abbiamo dimostrato che l'immagine di un'applicazione lineare è un sottospazio, lo stesso vale per il nucleo.

Proposizione 16.16.

Siano V e W spazi vettoriali su un campo \mathbb{K} . Data un'applicazione \mathbb{K} -lineare $f: V \rightarrow W$, allora $\ker(f)$ è un sottospazio di V .

Dimostrazione. Abbiamo già osservato che ogni applicazione lineare $f: V \rightarrow W$ manda 0_V in 0_W , quindi $0_V \in \ker(f)$. Inoltre se $v, v', v'' \in \ker(f)$ e $\alpha \in \mathbb{K}$ allora

$$f(v' + v'') = f(v') + f(v'') = 0_W, \quad f(\alpha v) = \alpha f(v) = \alpha 0_W = 0_W. \quad \square$$

Esempio 16.17. Sia \mathbb{K} un campo e sia $A \in \mathbb{K}^{m,n}$. Si consideri l'applicazione lineare $\mu_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$ dell'Esempio 16.4.

Lo spazio $\ker(\mu_A)$ è formato dalle $X \in \mathbb{K}^n$ tali che $0_{\mathbb{K}^m} = \mu_A(X) = AX$. In particolare $\ker(\mu_A)$ è un sottospazio di \mathbb{K}^n di dimensione $n - \text{rk}(A)$.

Concludiamo che l'applicazione μ_A è iniettiva se e solo se $n = \text{rk}(A)$: necessariamente allora $n \leq m$.

Ad esempio, l'applicazione f definita nell'Esempio 16.4 non può essere iniettiva: infatti tale applicazione coincide con μ_A , dove

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

ha, chiaramente, rango 2. Quindi $\dim(\ker(f)) = 3 - \text{rk}(A) = 1$; più in dettaglio, si ha che $\ker(f) = \mathcal{L}(1, -7, -4)$.

Invece l'applicazione g , sempre dell'Esempio 16.4, è iniettiva per lo stesso motivo: infatti è definita dalla matrice $B = {}^tA$, dunque $\dim(\ker(g)) = 2 - \text{rk}(A) = 0$, da cui segue $\ker(g) = \{ (0, 0) \}$.

Per quanto riguarda l'immagine $\text{im}(\mu_A)$, poiché i vettori e_1, \dots, e_n sono generatori di \mathbb{K}^n la Proposizione 16.13 implica che

$$\text{im}(\mu_A) = \mathcal{L}(\mu_A(e_1), \dots, \mu_A(e_n)) = \mathcal{L}(A_1, \dots, A_n) \subseteq \mathbb{K}^m$$

dove $A_1, \dots, A_n \in \mathbb{K}^m$ sono le colonne della matrice A .

In particolare $\text{im}(\mu_A)$ è un sottospazio di \mathbb{K}^m di dimensione $\text{rk}(A)$; quindi μ_A è suriettiva se e solo se $m = \text{rk}(A)$: necessariamente allora $n \geq m$.

Per esempio, tra le applicazioni f e g sopra f è suriettiva, mentre g non lo è. ♠

Ricordiamo dalla Lezione 5 che le soluzioni di un sistema lineare $AX = B$ sono della forma $X_0 + Y$, ove X_0 è una soluzione particolare e Y è soluzione del sistema omogeneo associato $AY = 0$: osserviamo che ciò è coerente con la Proposizione 16.13, punto (ii).

Un'interessante conseguenza di questo legame fra dimensione dell'immagine e rango è la seguente: se $A \in \mathbb{K}^{m,n}$, $B \in \mathbb{K}^{n,p}$, risulta

$$\text{rk}(AB) = \dim(\text{im}(\mu_{AB})) = \dim(\text{im}(\mu_A \circ \mu_B)).$$

Poiché $\text{im}(\mu_A \circ \mu_B) \subseteq \text{im}(\mu_A)$, concludiamo che $\text{rk}(AB) \leq \dim(\text{im}(\mu_A)) = \text{rk}(A)$. In maniera analoga si ha che

$$\begin{aligned} \text{rk}(AB) &= \text{rk}({}^t(AB)) = \text{rk}({}^tB {}^tA) = \dim(\text{im}(\mu_{{}^tB} {}^tA)) \\ &= \dim(\text{im}(\mu_{{}^tB} \circ \mu_{{}^tA})) \leq \dim(\text{im}(\mu_{{}^tB})) = \text{rk}({}^tB) = \text{rk}(B). \end{aligned}$$

Concludiamo quindi che

$$\text{rk}(AB) \leq \min\{\text{rk}(A), \text{rk}(B)\}.$$

Alla disuguaglianza qui sopra si può anche giungere direttamente osservando che le righe (rispettivamente colonne) di AB sono combinazione lineare delle righe (rispettivamente colonne) di B (rispettivamente A).

Se poi $n = p$ e la matrice $B \in \mathbb{K}^{n,n}$ è invertibile, allora

$$\text{rk}(A) = \text{rk}((AB)B^{-1}) \leq \text{rk}(AB) \leq \text{rk}(A),$$

dunque $\text{rk}(A) = \text{rk}(AB)$.

☠ La condizione che la matrice B sia invertibile è fondamentale perché valga l'uguaglianza tra $\text{rk}(A)$ e $\text{rk}(AB)$: basti pensare al caso di due matrici (chiaramente non invertibili) A e B non nulle tali che $AB = 0_{m,p}$.

Esempio 16.18. Sia $I =]a, b[\subseteq \mathbb{R}$ un intervallo aperto non vuoto e si consideri l'applicazione $D: \mathcal{C}^p(I) \rightarrow \mathcal{C}^{p-1}(I)$ di derivazione definita nell'Esempio 16.12. Allora $\ker(D)$ è costituito dalle funzioni $\varphi: I \rightarrow \mathbb{R}$ tali che $\varphi' = 0$ come funzioni. Segue allora dal Teorema di Lagrange che una tale φ deve essere costante su I . Concludiamo che D non è iniettiva.

Per quanto riguarda la suriettività, si osservi che $\text{im}(D)$ è costituito dalle funzioni $\psi: I \rightarrow \mathbb{R}$ tali che esiste $\varphi \in \mathcal{C}^p(I)$ per cui $\varphi' = \psi$ ovvero dalle primitive di ψ : in particolare D è suriettiva, in quanto è noto che le primitive di una funzione di $\mathcal{C}^{p-1}(I)$ sono tutte elementi in $\mathcal{C}^p(I)$.

Concludiamo che D è suriettiva e $D^{-1}\psi$ è l'insieme di tutte le primitive di ψ che, usualmente, si indica con il simbolo $\int \psi(x)dx$ e viene chiamato integrale indefinito di ψ .

Ricordiamo dalla teoria che due primitive di una funzione differiscono, su ogni intervallo contenuto nel dominio, per una costante: molto spesso si scrive

$$\int \psi(x)dx = \varphi(x) + c.$$

Come abbiamo visto sopra le funzioni costanti su I formano $\ker(D)$.

Più in generale si consideri l'operatore differenziale lineare

$$P(D) = a_0 D^n + a_1 D^{n-1} + \dots + a_{n-1} D + a_n: \mathcal{C}^\infty(I) \longrightarrow \mathcal{C}^\infty(I),$$

anch'esso introdotto nell'Esempio 16.12. Il nucleo $\ker(P(D))$ è l'insieme delle soluzioni $\varphi \in \mathcal{C}^\infty(I)$ dell'*equazione differenziale lineare omogenea*

$$a_0 \varphi^{(n)} + a_1 \varphi^{(n-1)} + \dots + a_{n-1} \varphi' + a_n \varphi = 0. \quad (16.3.1)$$

Ovviamente se $b \in \mathcal{C}^\infty(I)$ allora $P(D)^{-1}(b)$ è l'insieme delle soluzioni $\varphi \in \mathcal{C}^\infty(I)$ dell'*equazione differenziale lineare non omogenea*

$$a_0 \varphi^{(n)} + a_1 \varphi^{(n-1)} + \dots + a_{n-1} \varphi' + a_n \varphi = b. \quad (16.3.2)$$

La Proposizione 16.13, punto (ii) ci permette d'affermare che le soluzioni dell'equazione (16.3.2) si ottengono tutte sommando a una soluzione particolare la totalità delle soluzioni dell'equazione omogenea associata (16.3.1), ovvero l'integrale generale dell'equazione (16.3.2) è somma di un integrale particolare più l'integrale generale dell'equazione omogenea associata (16.3.1). ♠

Esempio 16.19. Siano $I \subseteq \mathbb{R}$ un intervallo aperto non vuoto, $x_0 \in I$ e $g \in \mathbb{R}^I$ non nulla in $I \setminus \{x_0\}$. Si consideri il sottospazio

$$\mathcal{F}_{x_0, g} = \left\{ f \in \mathbb{R}^I \mid \lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} \in \mathbb{R} \right\}.$$

definito nell'Esempio 12.14.

Si consideri poi l'applicazione

$$\begin{aligned} \varphi : \mathcal{F}_{x_0, g} &\longrightarrow \mathbb{R} \\ f &\longrightarrow \lim_{x \rightarrow x_0} \frac{f(x)}{g(x)}, \end{aligned}$$

che è chiaramente lineare per le proprietà dei limiti. Si ha che

$$\ker(\varphi) = \{ f \in \mathcal{F}_{x_0, g} \mid \varphi(f) = 0 \} = \left\{ f \in \mathcal{F}_{x_0, g} \mid \lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = 0 \right\} = o(g),$$

ovvero il nucleo dell'applicazione φ coincide con le funzioni trascurabili rispetto a g per $x \rightarrow x_0$, il cosiddetto *o piccolo* di g per $x \rightarrow x_0$. In particolare, φ non è iniettiva.

Data una qualsiasi $f \in \mathcal{F}_{x_0, g}$, possiamo definire le funzioni "equivalenti" a essa come gli elementi di $\mathcal{F}_{x_0, g}$ che in un intorno del punto x_0 si differenziano da f per un elemento di $o(g)$. ♠

Esempio 16.20. Nell'Esempio 16.7 sono state definite le applicazioni lineari prodotto scalare $\langle \cdot, \vec{v}_0 \rangle$ e prodotto vettoriale $\cdot \times \vec{v}_0$ per un vettore v_0 fissato. Possiamo calcolarne i rispettivi nuclei:

$$\begin{aligned} \ker(\langle \cdot, \vec{v}_0 \rangle) &= \{ \vec{v} \in V_3(O) \mid \langle \vec{v}, \vec{v}_0 \rangle = 0 \}, \\ \ker(\cdot \times \vec{v}_0) &= \{ \vec{v} \in V_3(O) \mid \vec{v} \times \vec{v}_0 = \vec{0} \} = \begin{cases} \mathcal{L}(\vec{v}_0), & \text{se } \vec{v}_0 \neq \vec{0} \\ V_3(O) & \text{se } \vec{v}_0 = \vec{0}. \end{cases} \end{aligned}$$

Passiamo ora a studiare la suriettività delle due applicazioni. Si ha

$$\text{im}(\langle \cdot, \vec{v}_0 \rangle) = \{ \alpha \in \mathbb{R} \mid \exists \vec{v} \in V_3(O) \text{ tale che } \alpha = \langle \vec{v}, \vec{v}_0 \rangle \}.$$

Poiché $\dim_{\mathbb{R}}(\mathbb{R}) = 1$, segue che $\langle \cdot, \vec{v}_0 \rangle$ è suriettiva se e solo se $\text{im}(\langle \cdot, \vec{v}_0 \rangle) \neq \{ 0 \}$. Se $\vec{v}_0 = \vec{0}$ chiaramente $\text{im}(\langle \cdot, \vec{v}_0 \rangle) = \{ 0 \}$, quindi l'applicazione non è suriettiva: invece se $\vec{v}_0 \neq \vec{0}$, poiché $\langle \vec{v}_0, \vec{v}_0 \rangle = |\vec{v}_0|^2 \neq 0$, segue che l'applicazione è suriettiva.

Per il prodotto vettoriale invece risulta

$$\begin{aligned} \text{im}(\cdot \times \vec{v}_0) &= \{ \vec{w} \in V_3(O) \mid \exists \vec{v} \in V_3(O) \text{ tale che } \vec{w} = \vec{v} \times \vec{v}_0 \} \\ &\subseteq \vec{v}_0^\perp = \{ \vec{w} \in V_3(O) \mid \langle \vec{w}, \vec{v}_0 \rangle = 0 \} \end{aligned}$$

Vedremo in seguito che $\text{im}(\cdot \times \vec{v}_0) = \vec{v}_0^\perp$ se $\vec{v}_0 \neq \vec{0}$. ♠

16.4 Isomorfismi

Ricordiamo che un'applicazione $\varphi: X \rightarrow Y$ fra due insiemi di qualsiasi tipo si dice *biiettiva* se è sia iniettiva che suriettiva. Ciò significa che se $y \in Y$, esiste (suriettività) ed è unico (iniettività) $x \in X$ tale che $\varphi(x) = y$. In tal caso è quindi possibile definire l'applicazione inversa $\varphi^{-1}: Y \rightarrow X$, ponendo $\varphi^{-1}(y) = x$.

Di più, è possibile dimostrare che φ è biiettiva se e solo se esiste un'applicazione $\psi: Y \rightarrow X$ tale che $\psi \circ \varphi = id_X$ e $\varphi \circ \psi = id_Y$ e risulta $\psi = \varphi^{-1}$.

Definizione 16.21 (Isomorfismi).

Siano V e W spazi vettoriali su un campo \mathbb{K} . Un'applicazione $f: V \rightarrow W$ si dice *isomorfismo* se è lineare e biiettiva; in tal caso V e W si dicono *isomorfi* e si scrive $V \cong W$.

Esempio 16.22. Siano V uno spazio vettoriale su un campo \mathbb{K} e $\mathcal{B} = (v_1, \dots, v_n)$ una sua base. Si consideri l'applicazione $[\cdot]_{\mathcal{B}}: V \rightarrow \mathbb{K}^n$, che abbiamo dimostrato essere lineare nell'Esempio 16.5. Chiaramente essa è suriettiva: infatti per qualsiasi $(\alpha_1, \dots, \alpha_n) \in \mathbb{K}^n$ si ha $[\sum_{i=1}^n \alpha_i v_i]_{\mathcal{B}} = (\alpha_1, \dots, \alpha_n)$. Inoltre $[\cdot]_{\mathcal{B}}$ è anche iniettiva: infatti se $v \in \ker([\cdot]_{\mathcal{B}})$ allora $v = \sum_{i=1}^n 0v_i = 0_V$.

Concludiamo perciò che $[\cdot]_{\mathcal{B}}$ è un isomorfismo. ♠

Esempio 16.23. Siano \mathbb{K} un campo e $A \in \mathbb{K}^{m,n}$. Da quanto visto nel paragrafo precedente, segue che l'applicazione μ_A è un isomorfismo se e solo se $n - \text{rk}(A) = 0$ e $m = \text{rk}(A)$ ovvero se e solo se $n = m = \text{rk}(A)$, cioè se e solo se A è invertibile. Dalla catena di uguaglianze

$$\mu_{A^{-1}} \circ \mu_A = \mu_{A^{-1}A} = \mu_{I_n} = \text{id}_{\mathbb{K}^n}, \quad \mu_A \circ \mu_{A^{-1}} = \mu_{AA^{-1}} = \mu_{I_n} = \text{id}_{\mathbb{K}^n}$$

si deduce che, in tal caso, $\mu_{A^{-1}} = \mu_A^{-1}$.

In particolare $\mathbb{K}^n \cong \mathbb{K}^m$ se e solo se $n = m$. ♠

Esempio 16.24. Dalla discussione fatta nell'Esempio 16.20 sulle applicazioni lineari prodotto scalare $\langle \cdot, \vec{v}_0 \rangle$ e prodotto vettoriale $\cdot \times \vec{v}_0$ deduciamo che tali applicazioni non sono mai isomorfismi. ♠

Siano adesso V e W spazi vettoriali su un campo \mathbb{K} . Se $f: V \rightarrow W$ è un isomorfismo, allora l'applicazione inversa $f^{-1}: W \rightarrow V$ è definita nel modo seguente: $f^{-1}(w) = v$ se e solo se $f(v) = w$. Concludiamo il paragrafo dimostrando che anche l'applicazione inversa è un isomorfismo.

Proposizione 16.25.

Siano V e W spazi vettoriali su un campo \mathbb{K} . Se $f: V \rightarrow W$ è un isomorfismo, allora anche f^{-1} è un isomorfismo.

Dimostrazione. Siano $w', w'' \in W$ e poniamo $v' = f^{-1}(w')$, $v'' = f^{-1}(w'')$, $\hat{v} = f^{-1}(w' + w'')$. Risulta che

$$f(\hat{v}) = w' + w'' = f(v') + f(v'') = f(v' + v''),$$

quindi

$$f^{-1}(w' + w'') = \hat{v} = v' + v'' = f^{-1}(w') + f^{-1}(w'').$$

Sia ora $w \in W$, $\alpha \in \mathbb{K}$ e poniamo $v = f^{-1}(w)$, $\tilde{v} = f^{-1}(\alpha w)$. Risulta che

$$f(\tilde{v}) = \alpha w = \alpha f(v) = f(\alpha v),$$

quindi

$$f^{-1}(\alpha w) = \tilde{v} = \alpha v = \alpha f^{-1}(w). \quad \square$$

Osservazione 16.26. Si noti che $V \cong V$ in quanto $id_V: V \rightarrow V$ è lineare e biiettiva. La Proposizione 16.25 mostra che se $V \cong W$, allora $W \cong V$. Infine la composizione di applicazioni biiettive è, ovviamente, biiettiva: segue quindi che se $U \cong V$ e $V \cong W$ allora $U \cong W$. Concludiamo che la relazione d'isomorfismo fra spazi vettoriali su \mathbb{K} è una relazione di equivalenza.

Applicazioni lineari e loro rappresentazione matriciale

17.1 Applicazioni lineari tra spazi di dimensione finita

Siano V e W spazi vettoriali su un campo \mathbb{K} , per esempio $\mathbb{K} = \mathbb{R}, \mathbb{C}$, ed $f: V \rightarrow W$ un'applicazione \mathbb{K} -lineare. Supponiamo che V sia finitamente generato: allora sappiamo che $\text{im}(f)$ è a sua volta finitamente generato per il punto (iii) della Proposizione 16.13.

Per questo motivo, qualora il dominio di un'applicazione lineare sia finitamente generato, a patto di cambiare opportunamente il codominio si può supporre che anch'esso sia finitamente generato.

Il seguente risultato afferma che, per descrivere un'applicazione lineare definita su uno spazio vettoriale finitamente generato, è sufficiente avere un numero finito di informazioni.

Proposizione 17.1.

Siano V e W spazi vettoriali finitamente generati su un campo \mathbb{K} . Dati una base $\mathcal{B} = (v_1, \dots, v_n)$ di V e $w_1, \dots, w_n \in W$ esiste un'unica applicazione \mathbb{K} -lineare $f: V \rightarrow W$ tale che $f(v_i) = w_i$ per $i = 1, \dots, n$.

Dimostrazione. Supponiamo che una tale f esista. Essendo f lineare se $v \in V$ e $[v]_{\mathcal{B}} = (x_1, \dots, x_n)$ si deve avere

$$f(v) = f\left(\sum_{i=1}^n x_i v_i\right) = \sum_{i=1}^n x_i f(v_i) = \sum_{i=1}^n x_i w_i, \quad (17.1.1)$$

da cui segue immediatamente l'unicità

Verifichiamo che la formula (17.1.1) definisce un'applicazione \mathbb{K} -lineare f soddisfacente la tesi. Chiaramente $[v_i]_{\mathcal{B}} = e_i \in \mathbb{K}^n$, dunque $f(v_i) = w_i$, $i = 1, \dots, n$. Se poi $\alpha \in \mathbb{K}$ si ha $[\alpha v]_{\mathcal{B}} = (\alpha x_1, \dots, \alpha x_n)$, dunque

$$f(\alpha v) = \sum_{i=1}^n (\alpha x_i) w_i = \alpha \left(\sum_{i=1}^n x_i w_i \right) = \alpha f(v).$$

Se $v', v'' \in V$ e $[v']_{\mathcal{B}} = (x'_1, \dots, x'_n)$, $[v'']_{\mathcal{B}} = (x''_1, \dots, x''_n)$, allora $[v' + v'']_{\mathcal{B}} = (x'_1 + x''_1, \dots, x'_n + x''_n)$ e risulta

$$f(v' + v'') = \sum_{i=1}^n (x'_i + x''_i) w_i = \sum_{i=1}^n x'_i w_i + \sum_{i=1}^n x''_i w_i = f(v') + f(v'').$$

Concludiamo che f è un'applicazione \mathbb{K} -lineare. □

Dalla Proposizione 17.1 deduciamo che se due applicazioni lineari coincidono sugli elementi di una base del dominio, esse coincidono ovunque, come spiegato nel seguente risultato.

Corollario 17.2.

Siano V e W spazi vettoriali finitamente generati su un campo \mathbb{K} . Data una base $\mathcal{B} = (v_1, \dots, v_n)$ di V si ha:

- (i) se $f, g: V \rightarrow W$ sono \mathbb{K} -lineari e tali che $f(v_i) = g(v_i)$ per $i = 1, \dots, n$, allora $f = g$;
- (ii) se $h: V \rightarrow W$ è \mathbb{K} -lineare e tale che $h(v_i) = 0_W$ per $i = 1, \dots, n$, allora $h = 0_{V,W}$.

Diamo ora alcuni esempi di applicazione della Proposizione e del Corollario di cui sopra.

Esempio 17.3. In \mathbb{R}^3 siano dati $v_1 = (1, 1, 1)$, $v_2 = (1, 0, -1)$, $v_3 = (2, 0, -1)$. Fissiamo poi $w_1 = (1, 1)$, $w_2 = (-1, 0)$, $w_3 = (0, 2)$ in \mathbb{R}^2 . Ci domandiamo se esiste un'applicazione lineare $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ tale che $f(v_i) = w_i$, $i = 1, 2, 3$.

Si verifichi che $\mathcal{B} = (v_1, v_2, v_3)$ è una base di \mathbb{R}^3 : dalla Proposizione 17.1 segue l'esistenza e l'unicità dell'applicazione lineare f .

Poiché

$$(x, y, z) = y v_1 + (-x + 3y - 2z) v_2 + (x - 2y + z) v_3,$$

segue che

$$\begin{aligned} f(x, y, z) &= f(yv_1 + (-x + 3y - 2z)v_2 + (x - 2y + z)v_3) \\ &= yw_1 + (-x + 3y - 2z)w_2 + (x - 2y + z)w_3 \\ &= (x - 2y + 2z, 2x - 3y + 2z). \end{aligned} \spadesuit$$

Esempio 17.4. In \mathbb{R}^3 siano dati $v_1 = (1, 1, 1)$, $v_2 = (1, 0, -1)$, $v_3 = (2, 0, -1)$, $v_4 = (4, 1, -1)$. Fissiamo poi $w_1 = (1, 1)$, $w_2 = (-1, 0)$, $w_3 = (0, 2)$, $w_4 = (0, 3)$ in \mathbb{R}^2 . Ci domandiamo se esiste un'applicazione lineare $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ tale che $f(v_i) = w_i$, $i = 1, 2, 3, 4$.

È chiaro che v_1, v_2, v_3, v_4 non possono formare una base di \mathbb{R}^3 , quindi non possiamo applicare direttamente la Proposizione 17.1.

Tuttavia, dall'esempio precedente, sappiamo che esiste un'unica applicazione lineare $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ tale che $f(v_i) = w_i$ per $i = 1, 2, 3$: si deve stabilire se vale anche $f(v_4) = w_4$. A tale scopo, o utilizziamo la formula già ottenuta per f , oppure osserviamo che $v_4 = v_1 + v_2 + v_3$: dunque affinché $f(v_4) = w_4$ si deve avere la relazione

$$w_4 = f(v_4) = f(v_1 + v_2 + v_3) = f(v_1) + f(v_2) + f(v_3) = w_1 + w_2 + w_3,$$

che è di immediata verifica.

Cosa si può affermare se si sostituisce w_4 con $w'_4 = (1, 1)$? ♠

Prendiamo adesso in considerazione il caso degli isomorfismi, sempre tra spazi vettoriali di dimensione finita.

Proposizione 17.5.

Siano V e W spazi vettoriali finitamente generati su un campo \mathbb{K} . Allora $V \cong W$ se e solo se $\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(W)$.

Dimostrazione. Siano $f: V \rightarrow W$ un isomorfismo e $\mathcal{B} = (v_1, \dots, v_n)$ e $\mathcal{D} = (w_1, \dots, w_m)$ basi di V e W rispettivamente. Consideriamo il seguente diagramma

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \uparrow [\cdot]_{\mathcal{B}}^{-1} & & \downarrow [\cdot]_{\mathcal{D}} \\ \mathbb{K}^n & \xrightarrow{g} & \mathbb{K}^m \end{array} \quad (17.1.2)$$

dove la mappa g è data dalla composizione

$$g = [\cdot]_{\mathcal{D}} \circ f \circ [\cdot]_{\mathcal{B}}^{-1}: \mathbb{K}^n \rightarrow \mathbb{K}^m.$$

Essendo composizione di applicazioni lineari biettive, g è lineare e biettiva, quindi è un isomorfismo. Poiché $g = \mu_A$ per una qualche $A \in \mathbb{K}^{m,n}$, come mostrato nell'Esempio 16.4, deduciamo che $n = m$ dall'Esempio 16.23.

Viceversa, supponiamo che $\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(W)$ e siano $\mathcal{B} = (v_1, \dots, v_n)$ e $\mathcal{D} = (w_1, \dots, w_n)$ basi di V e W rispettivamente. Possiamo considerare il diagramma (17.1.2) nel caso particolare in cui $n = m$ e $g = id_{\mathbb{K}^n}$ è l'applicazione identità; invertendo il senso delle frecce verticali troviamo

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \downarrow [\cdot]_{\mathcal{B}} & & \uparrow [\cdot]_{\mathcal{D}}^{-1} \\ \mathbb{K}^n & \xrightarrow{id_{\mathbb{K}^n}} & \mathbb{K}^n \end{array}$$

Allora $V \cong W$, poiché l'applicazione

$$f = [\cdot]_{\mathcal{D}}^{-1} \circ id_{\mathbb{K}^n} \circ [\cdot]_{\mathcal{B}}: V \rightarrow W,$$

essendo composizione di isomorfismi, è essa stessa un isomorfismo. \square

Esempio 17.6. Riprendiamo in considerazione i sottospazi $TS_n(\mathbb{K})$ (matrici triangolari superiori) e $Sim_n(\mathbb{K})$ (matrici simmetriche) di $\mathbb{K}^{n,n}$ ove \mathbb{K} è un campo. Nell'Esempio 15.12 abbiamo calcolato che $\dim_{\mathbb{K}}(TS_n(\mathbb{K})) = n(n+1)/2$.

Sia $A \in \mathbb{K}^{n,n}$; allora $A + {}^t A \in Sim_n(\mathbb{K})$ ed è definita l'applicazione

$$\begin{aligned} f: TS_n(\mathbb{K}) &\longrightarrow Sim_n(\mathbb{K}) \\ A &\longrightarrow A + {}^t A. \end{aligned}$$

Più in dettaglio, se $A = (a_{i,j})_{1 \leq i,j \leq n}$, allora $f(A) = B = (b_{i,j})_{1 \leq i,j \leq n} \in Sim_n(\mathbb{K})$ dove

$$b_{i,j} = \begin{cases} a_{i,j} & \text{se } i < j, \\ 2a_{i,i} & \text{se } i = j, \\ a_{j,i} & \text{se } i > j. \end{cases} \quad (17.1.3)$$

Chiaramente f è \mathbb{K} -lineare: infatti se $\alpha \in \mathbb{K}$ e $A, A', A'' \in TS_n(\mathbb{K})$ risulta

$$f(\alpha A) = \alpha A + {}^t(\alpha A) = \alpha A + \alpha {}^t A = \alpha(A + {}^t A) = \alpha f(A),$$

$$\begin{aligned} f(A' + A'') &= A' + A'' + {}^t(A' + A'') = A' + A'' + {}^t A' + {}^t A'' \\ &= A' + {}^t A' + A'' + {}^t A'' = f(A') + f(A''). \end{aligned}$$

Dimostriamo che f è un isomorfismo. Data $B = (b_{i,j})_{1 \leq i,j \leq n} \in Sim_n(\mathbb{K})$, risulta $B = f(A)$ con $A = (a_{i,j})_{1 \leq i,j \leq n} \in TS_n(\mathbb{K})$ definita da

$$a_{i,j} = \begin{cases} b_{i,j} & \text{se } i < j, \\ b_{i,i}/2 & \text{se } i = j, \\ 0 & \text{se } i > j, \end{cases}$$

quindi f è suriettiva.

Inoltre f è iniettiva, cioè $\ker(f) = \{ 0_{n,n} \}$: se $A = (a_{i,j})_{1 \leq i,j \leq n} \in TS_n(\mathbb{K})$ è tale che $f(A) = 0_{n,n}$, dalla relazione (17.1.3) segue che $a_{i,j} = 0 = 2a_{i,i}$, ovvero $A = 0_{n,n}$.

La Proposizione 17.5 garantisce dunque che

$$\dim_{\mathbb{K}}(Sim_n(\mathbb{K})) = \dim_{\mathbb{K}}(TS_n(\mathbb{K})) = \frac{n(n+1)}{2},$$

come anticipato nell'Esempio 15.13.

Similmente si consideri l'insieme delle matrici antisimmetriche di $\mathbb{K}^{n,n}$

$$Alt_n(\mathbb{K}) = \{ A \in \mathbb{K}^{n,n} \mid {}^t A = -A \}.$$

Si verifichi che $Alt_n(\mathbb{K})$ è un sottospazio di $\mathbb{K}^{n,n}$ e che l'applicazione

$$\begin{aligned} g: STS_n(\mathbb{K}) &\longrightarrow Alt_n(\mathbb{K}) \\ A &\longrightarrow A - {}^t A \end{aligned}$$

è un isomorfismo: in particolare

$$\dim_{\mathbb{K}}(Alt_n(\mathbb{K})) = \dim_{\mathbb{K}}(STS_n(\mathbb{K})) = \frac{n(n-1)}{2}. \quad \spadesuit$$

17.2 Matrice di un'applicazione lineare

Siano V e W spazi vettoriali finitamente generati su un campo \mathbb{K} e $\mathcal{B} = (v_1, \dots, v_n)$, $\mathcal{D} = (w_1, \dots, w_m)$ basi di V e W rispettivamente. Come abbiamo visto nella dimostrazione della Proposizione 17.5, invece di studiare direttamente una certa applicazione \mathbb{K} -lineare $f: V \rightarrow W$, può risultare più agevole comporla con opportuni isomorfismi con spazi vettoriali "semplici" come \mathbb{K}^n e \mathbb{K}^m e studiare al suo posto

l'applicazione \mathbb{K} -lineare composta $\mathbb{K}^n \rightarrow \mathbb{K}^m$ utilizzando quanto visto negli Esempi 16.4 e 16.23.

Nella dimostrazione della Proposizione 17.5 abbiamo visto che a un'applicazione \mathbb{K} -lineare $f: V \rightarrow W$ possiamo associare il diagramma (17.1.2), cioè

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \uparrow [\cdot]_{\mathcal{B}}^{-1} & & \downarrow [\cdot]_{\mathcal{D}} \\ \mathbb{K}^n & \xrightarrow{g} & \mathbb{K}^m \end{array}$$

dove, per definizione, g è l'applicazione \mathbb{K} -lineare $g = [\cdot]_{\mathcal{B}}^{-1} \circ f \circ [\cdot]_{\mathcal{D}}$.

Dall'Esempio 16.4 segue l'esistenza di una matrice $A \in \mathbb{K}^{m,n}$ tale che l'applicazione g del diagramma sia della forma μ_A , la moltiplicazione per A :

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \uparrow [\cdot]_{\mathcal{B}}^{-1} & & \downarrow [\cdot]_{\mathcal{D}} \\ \mathbb{K}^n & \xrightarrow{\mu_A} & \mathbb{K}^m \end{array} \quad (17.2.1)$$

La matrice A dipende sia dall'applicazione f che dalle basi \mathcal{B} e \mathcal{D} . Si noti che le colonne di A sono $\mu_A(E_{j,1}) = ([\cdot]_{\mathcal{D}} \circ f \circ [\cdot]_{\mathcal{B}}^{-1})(E_{j,1})$: poiché $[\cdot]_{\mathcal{B}}^{-1}(E_{j,1}) = v_j$, segue che le colonne di A non sono altro che $([\cdot]_{\mathcal{D}} \circ f)(v_j) = [f(v_j)]_{\mathcal{D}}$: cioè la j -esima colonna di A è formata dalle componenti rispetto alla base fissata nel codominio di f (disposte in colonna!) del j -esimo vettore della base fissata nel dominio di f .

Definizione 17.7 (Matrice di un'applicazione lineare).

Siano V e W spazi vettoriali finitamente generati su un campo \mathbb{K} . Date due basi $\mathcal{B} = (v_1, \dots, v_n)$ e $\mathcal{D} = (w_1, \dots, w_m)$ di V e W rispettivamente e un'applicazione \mathbb{K} -lineare $f: V \rightarrow W$, definiamo *matrice di f rispetto alle basi \mathcal{B} e \mathcal{D}* la matrice $M_{\mathcal{D}}^{\mathcal{B}}(f)$ avente per colonne le componenti delle immagini dei vettori di \mathcal{B} rispetto a \mathcal{D} .

⚠ Nel caso in cui $V = \mathbb{K}^n$ e $W = \mathbb{K}^m$ e $f: \mathbb{K}^n \rightarrow \mathbb{K}^m$ sia un'applicazione lineare, dall'Esempio 16.4 sappiamo che esiste $A \in \mathbb{K}^{m,n}$ tale che $f = \mu_A$. Allora A non è altro che la matrice di f rispetto alle basi canoniche \mathcal{C}_n di \mathbb{K}^n e \mathcal{C}_m di \mathbb{K}^m nel senso della definizione data sopra. Quindi $M_{\mathcal{C}_m}^{\mathcal{C}_n}(\mu_A) = A$.

Si noti che se $f: V \rightarrow W$ è un'applicazione \mathbb{K} -lineare, allora

$$\begin{aligned} M_{\mathcal{D}}^{\mathcal{B}}(f)[v]_{\mathcal{B}} &= \mu_{[M_{\mathcal{D}}^{\mathcal{B}}(f)]}([v]_{\mathcal{B}}) = ([\cdot]_{\mathcal{D}} \circ f \circ [\cdot]_{\mathcal{B}}^{-1})([v]_{\mathcal{B}}) \\ &= ([\cdot]_{\mathcal{D}} \circ f)([\cdot]_{\mathcal{B}}^{-1})([v]_{\mathcal{B}}) = ([\cdot]_{\mathcal{D}} \circ f)(v) = [f(v)]_{\mathcal{D}}. \end{aligned}$$

Si noti la "cancellazione" in croce

$$M_{\mathcal{D}}^{\mathcal{B}}(f)[v]_{\mathcal{B}} = [f(v)]_{\mathcal{D}}.$$

Esempio 17.8. Sia W uno spazio vettoriale di dimensione 2 su \mathbb{R} , dotato di una base $\mathcal{D} = (w_1, w_2)$, e si consideri l'applicazione

$$\begin{aligned} f: \mathbb{R}^3 &\longrightarrow W \\ (x, y, z) &\longrightarrow xw_1 + (2y + z)w_2. \end{aligned}$$

Detta $\mathcal{C} = (e_1, e_2, e_3)$ la base canonica di \mathbb{R}^3 , la matrice $M_{\mathcal{D}}^{\mathcal{C}}(f)$ è una matrice 2×3 le cui colonne sono $[f(e_i)]_{\mathcal{D}}$. Si ha che

$$f(e_1) = w_1, \quad f(e_2) = 2w_2, \quad f(e_3) = w_2,$$

quindi le coordinate di tali vettori sono

$$[f(e_1)]_{\mathcal{D}} = (1, 0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad [f(e_2)]_{\mathcal{D}} = (0, 2) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \quad [f(e_3)]_{\mathcal{D}} = (0, 1) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

e la matrice cercata è

$$M_{\mathcal{D}}^{\mathcal{C}}(f) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \end{pmatrix}. \quad \spadesuit$$

Esempio 17.9. Si consideri l'applicazione

$$\begin{aligned} f: \mathbb{R}[x]_1 &\longrightarrow \mathbb{R}[x]_2 \\ p(x) &\longrightarrow xp(x). \end{aligned}$$

Se $\alpha \in \mathbb{R}$ e $p_1(x), p_2(x) \in \mathbb{R}[x]_1$, si ha

$$\begin{aligned} f(\alpha p(x)) &= x(\alpha p(x)) = \alpha(xp(x)) = \alpha f(p(x)), \\ f(p_1(x) + p_2(x)) &= x(p_1(x) + p_2(x)) = xp_1(x) + xp_2(x) = f(p_1(x)) + f(p_2(x)), \end{aligned}$$

quindi f è \mathbb{R} -lineare.

Consideriamo le basi $\mathcal{B} = (1, x)$ in $\mathbb{R}[x]_1$ e $\mathcal{D} = (1, x, x^2)$ in $\mathbb{R}[x]_2$ e determiniamo $M_{\mathcal{D}}^{\mathcal{B}}(f)$. Risulta

$$[f(1)]_{\mathcal{D}} = [x]_{\mathcal{D}} = (0, 1, 0) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad [f(x)]_{\mathcal{D}} = [x^2]_{\mathcal{D}} = (0, 0, 1) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

quindi

$$M_{\mathcal{D}}^{\mathcal{B}}(f) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Si noti che $[f(a + bx)]_{\mathcal{D}} = M_{\mathcal{D}}^{\mathcal{B}}(f)[a + bx]_{\mathcal{B}}$. ♠

Osservazione 17.10. Alla luce della definizione di matrice associata a un'applicazione lineare, la Proposizione 16.13 può essere interpretata dicendo che dati V e W spazi vettoriali su un campo \mathbb{K} e $f: V \rightarrow W$ un'applicazione \mathbb{K} -lineare, l'immagine $\text{im}(f)$ è il sottospazio di W generato dai vettori le cui componenti sono le colonne della matrice $M_{\mathcal{D}}^{\mathcal{B}}(f)$, dove $\mathcal{B} = (v_1, \dots, v_n)$ e $\mathcal{D} = (w_1, \dots, w_m)$ sono basi di V e W rispettivamente.

Osservazione 17.11. Una conseguenza diretta della definizione e dell'Osservazione 16.11 è che la matrice associata a un'applicazione lineare si "comporta bene" rispetto alla composizione di applicazioni: siano infatti U, V, W spazi vettoriali finitamente generati su un campo \mathbb{K} di dimensioni p, n, m rispettivamente, dotati di basi $\mathcal{A}, \mathcal{B}, \mathcal{D}$ rispettivamente. Siano $g: U \rightarrow V, f: V \rightarrow W$ due applicazioni \mathbb{K} -lineari con matrici associate $M_{\mathcal{B}}^{\mathcal{A}}(g)$ e $M_{\mathcal{D}}^{\mathcal{B}}(f)$. Allora la matrice associata alla composizione $f \circ g: U \rightarrow W$ rispetto alle basi \mathcal{A} di U e \mathcal{D} di W è

$$M_{\mathcal{D}}^{\mathcal{A}}(f \circ g) = M_{\mathcal{D}}^{\mathcal{B}}(f)M_{\mathcal{B}}^{\mathcal{A}}(g),$$

come si può dedurre anche dal seguente diagramma commutativo

$$\begin{array}{ccccc} & & \text{f} \circ \text{g} & & \\ & \text{U} & \xrightarrow{\text{g}} & \text{V} & \xrightarrow{\text{f}} & \text{W} \\ & \uparrow [\cdot]_{\mathcal{A}}^{-1} & & \uparrow [\cdot]_{\mathcal{B}} \downarrow [\cdot]_{\mathcal{B}}^{-1} & & \downarrow [\cdot]_{\mathcal{D}} \\ \mathbb{K}^p & \xrightarrow{\mu_{M_{\mathcal{B}}^{\mathcal{A}}(g)}} & \mathbb{K}^n & \xrightarrow{\mu_{M_{\mathcal{D}}^{\mathcal{B}}(f)}} & \mathbb{K}^m. \\ & & \mu_{M_{\mathcal{D}}^{\mathcal{A}}(f \circ g)} & & \end{array}$$

Un'altra conseguenza immediata è che se $f: V \rightarrow W$ è un isomorfismo, allora

$$M_{\mathcal{B}}^{\mathcal{D}}(f^{-1}) = (M_{\mathcal{D}}^{\mathcal{B}}(f))^{-1}.$$

Posto $A = M_{\mathcal{D}}^{\mathcal{B}}(f)$, il diagramma (17.2.1) implica $\mu_A = [\cdot]_{\mathcal{D}} \circ f \circ [\cdot]_{\mathcal{B}}^{-1}$. Otteniamo quindi la seguente serie di affermazioni equivalenti dove si è posto $v = [\cdot]_{\mathcal{B}}^{-1}(X)$:

$$\begin{aligned} X \in \ker(\mu_A) \subseteq \mathbb{K}^n &\Leftrightarrow \mu_A(X) = 0_{\mathbb{K}^m} \\ &\Leftrightarrow ([\cdot]_{\mathcal{D}} \circ f \circ [\cdot]_{\mathcal{B}}^{-1})(X) = 0_{\mathbb{K}^m} \\ &\Leftrightarrow [f(v)]_{\mathcal{D}} = ([\cdot]_{\mathcal{D}} \circ f)(v) = 0_{\mathbb{K}^m}, \\ &\Leftrightarrow f(v) = ([\cdot]_{\mathcal{D}}^{-1} \circ [\cdot]_{\mathcal{D}} \circ f)(v) = [\cdot]_{\mathcal{D}}^{-1}(0_{\mathbb{K}^m}) = 0_W \\ &\Leftrightarrow v \in \ker(f) \subseteq V. \end{aligned}$$

Restringendo $[\cdot]_{\mathcal{B}}^{-1}$ a $\ker(\mu_A)$, otteniamo un isomorfismo fra $\ker(\mu_A)$ e $\ker(f)$. Similmente, restringendo $[\cdot]_{\mathcal{D}}$ a $\text{im}(f)$ otteniamo un isomorfismo fra $\text{im}(f)$ e $\text{im}(\mu_A)$. Dalla Proposizione 15.14 segue allora che

$$\begin{aligned} \dim_{\mathbb{K}}(\text{im}(f)) &= \dim_{\mathbb{K}}(\text{im}(\mu_A)) = \text{rk}(A), \\ \dim_{\mathbb{K}}(\ker(f)) &= \dim_{\mathbb{K}}(\ker(\mu_A)) = n - \text{rk}(A). \end{aligned}$$

In particolare abbiamo immediatamente il seguente risultato: esso non è altro che una riscrittura del Teorema di Rouché–Capelli (Proposizione 5.2).

Proposizione 17.12 (Teorema della dimensione).

Siano V e W spazi vettoriali finitamente generati su un campo \mathbb{K} . Data un'applicazione \mathbb{K} -lineare $f: V \rightarrow W$ si ha

$$\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(\ker(f)) + \dim_{\mathbb{K}}(\operatorname{im}(f)).$$

Corollario 17.13.

Siano V e W spazi vettoriali finitamente generati su un campo \mathbb{K} . Data un'applicazione \mathbb{K} -lineare $f: V \rightarrow W$ si ha:

- (i) se f è iniettiva, allora $\dim_{\mathbb{K}}(V) \leq \dim_{\mathbb{K}}(W)$;
- (ii) se f è suriettiva, allora $\dim_{\mathbb{K}}(V) \geq \dim_{\mathbb{K}}(W)$.

Quanto visto sopra ci permette di studiare un'applicazione lineare tra spazi vettoriali di dimensione finita in modo più facile, studiandola cioè tramite la sua matrice rispetto a basi fissate (che possiamo scegliere “comode”).

Esempio 17.14. Il lettore verifichi che l'applicazione

$$\begin{aligned} f: \mathbb{R}[x]_2 &\longrightarrow \mathbb{R}^{2,2} \\ a + bx + cx^2 &\longrightarrow \begin{pmatrix} a + b & a + c \\ b - c & b - c \end{pmatrix} \end{aligned}$$

è lineare.

Fissiamo le basi $\mathcal{B} = (1, x, x^2)$ e $\mathcal{D} = (E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2})$ in $\mathbb{R}[x]_2$ e $\mathbb{R}^{2,2}$ rispettivamente. Poiché

$$\begin{aligned} [f(1)]_{\mathcal{D}} &= \left[\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \right]_{\mathcal{D}} = (1, 1, 0, 0), & [f(x)]_{\mathcal{D}} &= \left[\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \right]_{\mathcal{D}} = (1, 0, 1, 1), \\ [f(x^2)]_{\mathcal{D}} &= \left[\begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} \right]_{\mathcal{D}} = (0, 1, -1, -1), \end{aligned}$$

segue che la matrice di f rispetto alle basi \mathcal{B} e \mathcal{D} è

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

Possiamo calcolare

$$\begin{aligned} \dim(\ker(f)) &= \dim(\ker(\mu_A)) = 3 - \operatorname{rk}(A) = 1, \\ \dim(\operatorname{im}(f)) &= \dim(\operatorname{im}(\mu_A)) = \operatorname{rk}(A) = 2, \end{aligned}$$

da cui deduciamo che f non è né iniettiva né suriettiva.

Se vogliamo studiare $\ker(f)$ e $\operatorname{im}(f)$ più in dettaglio, possiamo studiare $\ker(\mu_A)$ e $\operatorname{im}(\mu_A)$. Risolvendo il sistema $AX = 0_{4,1}$ otteniamo

$$\ker(\mu_A) = \{ (a, -a, -a) \mid a \in \mathbb{R} \},$$

quindi

$$\begin{aligned} \ker(f) &= \{ p(x) \in \mathbb{R}[x]_2 \mid [p(x)]_{\mathcal{B}} \in \ker(\mu_A) \} \\ &= \{ a - ax - ax^2 \mid a \in \mathbb{R} \} = \mathcal{L}(1 - x - x^2). \end{aligned}$$

In particolare $\dim_{\mathbb{R}}(\ker(f)) = 1$, come già visto sopra.

Una conseguenza immediata è che $0_{2,2} = f(1 - x - x^2) = f(1) - f(x) - f(x^2)$ ovvero $f(x^2) = f(1) - f(x)$, quindi

$$\operatorname{im}(f) = \mathcal{L}(f(1), f(x), f(x^2)) = \mathcal{L}(f(1), f(x)).$$

Poiché $f(x) \notin \mathcal{L}(f(1))$, segue che $\dim_{\mathbb{R}}(\operatorname{im}(f)) = 2$, come già visto sopra. ♠

Esempio 17.15. Nell'Esempio 16.20 abbiamo verificato che, se $\vec{v}_0 \in V_3(O)$ è un vettore fissato, risulta

$$\begin{aligned} \operatorname{im}(\cdot \times \vec{v}_0) &= \{ \vec{w} \in V_3(O) \mid \exists \vec{v} \in V_3(O) \text{ tale che } \vec{w} = \vec{v} \times \vec{v}_0 \} \\ &\subseteq \vec{v}_0^\perp = \{ \vec{w} \in V_3(O) \mid \langle \vec{w}, \vec{v}_0 \rangle = 0 \} \end{aligned}$$

Se $\vec{v}_0 = \vec{0}$, allora $\operatorname{im}(\cdot \times \vec{v}_0) = \mathcal{L}(\vec{0})$; se, invece, $\vec{v}_0 \neq \vec{0}$, ancora l'Esempio 16.20 ci permette di affermare che $\dim_{\mathbb{R}}(\ker(\cdot \times \vec{v}_0)) = 1$, dunque la Proposizione 17.12 implica $\dim_{\mathbb{R}}(\operatorname{im}(\cdot \times \vec{v}_0)) = 2$. D'altra parte \vec{v}_0^\perp è un sottospazio di $V_3(O)$ (il lettore lo verifichi per esercizio) non contenente \vec{v}_0 : poiché $\dim_{\mathbb{R}}(V_3(O)) = 3$ segue che $\dim_{\mathbb{R}}(\vec{v}_0^\perp) \leq 2$. Essendo $\operatorname{im}(\cdot \times \vec{v}_0) \subseteq \vec{v}_0^\perp$ segue allora che deve valere l'uguaglianza, come anticipato.

Si noti che, fissato un sistema di riferimento $0\vec{i}\vec{j}\vec{k}$, risulta $\vec{v}_0 = a\vec{i} + b\vec{j} + c\vec{k}$. La matrice di $\cdot \times \vec{v}_0$ rispetto alla stessa base $\mathcal{B} = (\vec{i}, \vec{j}, \vec{k})$ fissata nel dominio e codominio è

$$A = \begin{pmatrix} 0 & c & -b \\ -c & 0 & a \\ b & -a & 0 \end{pmatrix}. \quad \spadesuit$$

Esempio 17.16. Si considerino $v_1 = (1, 2, 1)$, $v_2 = (1, 0, 1)$, $v_3 = (1, 0, 2)$ in \mathbb{R}^3 e

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 1 \\ 0 & 3 \end{pmatrix}$$

in $\mathbb{R}^{2,2}$. Poiché risulta

$$\text{rk} \begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 2 \end{pmatrix} = \text{rk} \begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} = 3,$$

segue che v_1, v_2, v_3 sono linearmente indipendenti, quindi per la Proposizione 15.7 si ha che $\mathcal{B} = (v_1, v_2, v_3)$ è una base di \mathbb{R}^3 .

Per la Proposizione 17.1 esiste un'unica applicazione lineare $f: \mathbb{R}^3 \rightarrow \mathbb{R}^{2,2}$ tale che $f(v_i) = A_i$, $i = 1, 2, 3$. Vogliamo studiare tale applicazione: a tale scopo scriviamone la matrice rispetto a basi opportunamente scelte nel dominio e nel codominio.

Nel dominio abbiamo varie scelte possibili, ad esempio la base canonica \mathcal{C} . Per semplificare al massimo la forma della matrice e, di conseguenza, i calcoli, la scelta migliore è, però, quella di prendere la base $\mathcal{B} = (v_1, v_2, v_3)$.

Anche nel codominio possiamo fare molte scelte lecite: potremmo ad esempio prendere la base $\mathcal{E} = (E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2})$, ma per semplificare al massimo i conti una scelta migliore è $\mathcal{D} = (A_1, A_2, E_{1,2}, E_{2,1})$ (verificare, per esercizio, che \mathcal{D} è base di $\mathbb{R}^{2,2}$).

Si ha

$$\begin{aligned} f(v_1) = A_1 &= 1A_1 + 0A_2 + 0E_{1,2} + 0E_{2,1} &\Rightarrow [f(v_1)]_{\mathcal{D}} &= (1, 0, 0, 0), \\ f(v_2) = A_2 &= 0A_1 + 1A_2 + 0E_{1,2} + 0E_{2,1} &\Rightarrow [f(v_2)]_{\mathcal{D}} &= (0, 1, 0, 0), \\ f(v_3) = A_3 &= 1A_1 + 1A_2 + 0E_{1,2} + 0E_{2,1} &\Rightarrow [f(v_3)]_{\mathcal{D}} &= (1, 1, 0, 0), \end{aligned}$$

quindi

$$M = M_{\mathcal{D}}^{\mathcal{B}}(f) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

È facile vedere dalla matrice che gli elementi e_1, e_2 della base canonica di \mathbb{R}^4 formano una base di $\text{im}(\mu_M) \subseteq \mathbb{R}^4$. Quindi $A_1 = [\cdot]_{\mathcal{D}}^{-1}(e_1)$, $A_2 = [\cdot]_{\mathcal{D}}^{-1}(e_2)$, formano una base (A_1, A_2) di $\text{im}(f)$.

Inoltre $\ker(\mu_M)$ è generato dal singolo vettore $e = (1, 1, -1)$: dal momento che $(1, 2, 0) = v_1 + v_2 - v_3 = [\cdot]_{\mathcal{B}}^{-1}(e)$, segue che $\ker(f)$ è generato dal vettore $(1, 2, 0)$.

Per esercizio si calcoli $M_{\mathcal{E}}^{\mathcal{C}}(f)$: si verifichi che $M_{\mathcal{E}}^{\mathcal{C}}(f) \neq M_{\mathcal{D}}^{\mathcal{B}}(f)$ e che procedendo come fatto sopra con la matrice $M_{\mathcal{E}}^{\mathcal{C}}(f)$ in luogo di $M_{\mathcal{D}}^{\mathcal{B}}(f)$ si riottengono gli stessi risultati. ♠

17.3 Endomorfismi

Proposizione 17.17.

Siano V e W spazi vettoriali finitamente generati su un campo \mathbb{K} con $\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(W)$. Data un'applicazione \mathbb{K} -lineare $f: V \rightarrow W$ le seguenti affermazioni sono equivalenti:

- f è iniettiva,
- f è suriettiva,
- f è un isomorfismo.

Dimostrazione. Per definizione, se f è un isomorfismo è sia iniettiva che suriettiva.

Supponiamo f sia iniettiva, quindi che $\dim_{\mathbb{K}}(\ker(f)) = 0$: per la Proposizione 17.12 e per l'ipotesi segue allora che $\dim_{\mathbb{K}}(W) = \dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(\text{im}(f))$, quindi per la Proposizione 15.9 $\text{im}(f) = W$, cioè f è anche suriettiva, e quindi è un isomorfismo.

Se f è suriettiva, allora $\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(W) = \dim_{\mathbb{K}}(\text{im}(f))$, quindi $\dim_{\mathbb{K}}(\ker(f)) = 0$ per la Proposizione 17.12, cioè f è anche iniettiva, quindi è un isomorfismo. \square

Vediamo ora un esempio che illustra l'utilità della precedente proposizione.

Esempio 17.18. Si consideri l'applicazione

$$\begin{aligned} f: \mathbb{C}^3 &\longrightarrow \mathbb{C}[x]_2 \\ (a, b, c) &\longrightarrow a + (a + b)x + (a + b + c)x^2. \end{aligned}$$

Si ha che $(a, b, c) \in \ker(f)$ se e solo se $a + (a + b)x + (a + b + c)x^2 = 0$ se e solo se $a = a + b = a + b + c = 0$, cioè se e solo se $(a, b, c) = (0, 0, 0)$: quindi f è iniettiva. Grazie alla Proposizione 17.17 possiamo concludere che f è un isomorfismo senza doverne studiare la suriettività.

A un analogo risultato si poteva arrivare osservando che la matrice di f rispetto alla base canonica di \mathbb{C}^3 e alla base $(1, x, x^2)$ di $\mathbb{C}[x]_2$ è

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

che è invertibile. ♠

Consideriamo adesso un'applicazione lineare da uno spazio vettoriale in se stesso.

Definizione 17.19 (Endomorfismi).

Sia V uno spazio vettoriale su un campo \mathbb{K} . Un *endomorfismo* di V è un'applicazione \mathbb{K} -lineare $f: V \rightarrow V$.

Chiaramente la Proposizione 17.17 si applica, in particolare, agli endomorfismi di spazi vettoriali finitamente generati. Tuttavia essa non è valida se si lavora con uno spazio non finitamente generato: esistono endomorfismi suriettivi ma non iniettivi, o viceversa iniettivi ma non suriettivi, come mostrano i seguenti esempi.

Esempio 17.20. Sia $I =]a, b[\subseteq \mathbb{R}$ un intervallo aperto non vuoto. Nell'Esempio 16.8 abbiamo osservato che l'applicazione $D: \mathcal{C}^\infty(I) \rightarrow \mathcal{C}^\infty(I)$ è lineare, quindi è un endomorfismo. Tale applicazione non è iniettiva, ma è suriettiva per un ben noto risultato di analisi. ♠

Esempio 17.21. Sia \mathbb{K} un campo e si consideri l'applicazione

$$\begin{aligned} f: \mathbb{K}[x] &\longrightarrow \mathbb{K}[x] \\ p(x) &\longrightarrow xp(x). \end{aligned}$$

Si verifichi che f è lineare (e quindi un endomorfismo). Chiaramente f è iniettiva, ma non è suriettiva, perché i polinomi costanti non sono in $\text{im}(f)$. ♠

Un caso particolarmente importante di endomorfismo è l'identità in uno spazio vettoriale finitamente generato V su un campo \mathbb{K} .

Definizione 17.22 (Matrice del cambiamento di base).

Sia V uno spazio vettoriale finitamente generato su un campo \mathbb{K} . Date due basi $\mathcal{B} = (v_1, \dots, v_n)$ e $\mathcal{D} = (w_1, \dots, w_n)$ di V la matrice $M_{\mathcal{D}}^{\mathcal{B}}(id_V)$ avente per colonne le componenti dei vettori di \mathcal{B} rispetto a \mathcal{D} è detta *matrice del cambiamento di base* da \mathcal{B} a \mathcal{D} .

Osserviamo subito che per ogni vettore $v \in V$ vale la relazione

$$[v]_{\mathcal{D}} = M_{\mathcal{D}}^{\mathcal{B}}(id_V)[v]_{\mathcal{B}}.$$

Un modo per ricordarsi la definizione è vedere la relazione sopra come una specie di "cancellazione" in croce:

$$[v]_{\mathcal{D}} = M_{\mathcal{D}}^{\mathcal{B}}(id_V)[v]_{\mathcal{B}}.$$

Si noti anche che vale la relazione

$$M_{\mathcal{D}}^{\mathcal{B}}(id_V) = (M_{\mathcal{B}}^{\mathcal{D}}(id_V))^{-1}.$$

Esempio 17.23. Si consideri lo spazio vettoriale $V = \mathbb{R}^3$, sia $\mathcal{C} = (e_1, e_2, e_3)$ la sua base canonica e siano $\mathcal{B} = (v_1, v_2, v_3)$ e $\mathcal{D} = (w_1, w_2, w_3)$ altre due basi, dove

$$\begin{aligned} v_1 &= \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, & v_2 &= \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, & v_3 &= \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \\ w_1 &= \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, & w_2 &= \begin{pmatrix} 1 \\ -2 \\ -3 \end{pmatrix}, & w_3 &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \end{aligned}$$

Il lettore verifichi che si tratta in effetti di vettori indipendenti.

Le matrici di cambiamento di base più semplici da scrivere sono quelle di passaggio dalle basi \mathcal{B} e \mathcal{D} a quella canonica, semplicemente perché abbiamo già le componenti dei loro vettori rispetto a \mathcal{C} . Si ha che

$$M_{\mathcal{C}}^{\mathcal{B}}(id_V) = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & -2 \\ 0 & 1 & 1 \end{pmatrix}, \quad M_{\mathcal{C}}^{\mathcal{D}}(id_V) = \begin{pmatrix} -1 & 1 & 1 \\ 0 & -2 & 1 \\ 1 & -3 & 1 \end{pmatrix}.$$

Per calcolare le matrici di cambiamento di base dalla canonica \mathcal{C} alle nuove basi \mathcal{B} e \mathcal{D} dobbiamo invece fare un po' più fatica. Cominciamo da $M_{\mathcal{B}}^{\mathcal{C}}(id_V)$: possiamo direttamente la definizione e calcolare le componenti $[e_i]_{\mathcal{B}}$ esplicitamente, oppure possiamo usare l'Osservazione 17.11 e ricordarci che

$$M_{\mathcal{B}}^{\mathcal{C}}(id_V) = (M_{\mathcal{C}}^{\mathcal{B}}(id_V))^{-1} = \begin{pmatrix} 3 & -2 & -4 \\ -1 & 1 & 2 \\ 1 & -1 & -1 \end{pmatrix}.$$

Il lettore verifichi il calcolo dell'inversa e trovi la matrice $M_{\mathcal{D}}^{\mathcal{C}}(id_V)$.

Passiamo adesso a calcolare la matrice di cambiamento di base $M_{\mathcal{B}}^{\mathcal{D}}(id_V)$: di nuovo, grazie all'Osservazione 17.11 si ha che

$$\begin{aligned} M_{\mathcal{B}}^{\mathcal{D}}(id_V) &= M_{\mathcal{B}}^{\mathcal{C}}(id_V)M_{\mathcal{C}}^{\mathcal{D}}(id_V) \\ &= \begin{pmatrix} 3 & -2 & -4 \\ -1 & 1 & 2 \\ 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} -1 & 1 & 1 \\ 0 & -2 & 1 \\ 1 & -3 & 1 \end{pmatrix} = \begin{pmatrix} -7 & 19 & -3 \\ 3 & -9 & 2 \\ -2 & -36 & -1 \end{pmatrix}. \end{aligned}$$

Lasciamo al lettore la verifica che

$$\begin{aligned} M_{\mathcal{D}}^{\mathcal{B}}(id_V) &= M_{\mathcal{D}}^{\mathcal{C}}(id_V)M_{\mathcal{C}}^{\mathcal{B}}(id_V) \\ &= (M_{\mathcal{C}}^{\mathcal{D}}(id_V))^{-1}M_{\mathcal{C}}^{\mathcal{B}}(id_V) = \begin{pmatrix} -3/2 & 1/2 & 11/2 \\ -1/2 & 1/2 & 5/2 \\ 0 & 2 & 3 \end{pmatrix}. \quad \spadesuit \end{aligned}$$

⚠ Testi diversi utilizzano notazioni e nomi diversi per la matrice di cambiamento di base, ad esempio $A_{\mathcal{B} \rightarrow \mathcal{D}}$, $[T]_{\mathcal{B}\mathcal{D}}$ e molte altre. Segnaliamo al lettore che purtroppo alcuni testi usano anche la notazione $M_{\mathcal{B}}^{\mathcal{D}}$ per intendere la matrice di cambio di base da \mathcal{D} a \mathcal{B} , ovvero esattamente l'inversa della nostra.

Autovalori, autovettori e autospazi

18.1 Autovalori, autovettori e autospazi di endomorfismi

Abbiamo visto che per descrivere un'applicazione lineare fra spazi vettoriali finitamente generati la cosa più semplice è quella di considerare la sua matrice rispetto a certe basi fissate. Si ricordi che \mathbb{K} indica un campo: per esempio $\mathbb{K} = \mathbb{R}, \mathbb{C}$.

Un caso particolarmente interessante è quello di un endomorfismo f di uno spazio vettoriale finitamente generato V . In tal caso può essere utile riuscire a determinare una base $\mathcal{B} = (v_1, \dots, v_n)$ tale che la matrice $M_{\mathcal{B}}^{\mathcal{B}}(f)$ sia particolarmente semplice, per esempio diagonale: in tal caso si dice proprio che l'endomorfismo f è *semplice*.

Ricordando come è definita $M_{\mathcal{B}}^{\mathcal{B}}(f)$, questo significa chiedere l'esistenza di $\lambda_i \in \mathbb{K}$ tali che $[f(v_i)]_{\mathcal{B}} = \lambda_i e_i$, cioè

$$f(v_i) = \lambda_i [\cdot]_{\mathcal{B}}^{-1}(e_i) = \lambda_i v_i.$$

Purtroppo ciò in generale non accade. Tuttavia i vettori aventi tale proprietà per un fissato endomorfismo sono di particolare interesse e importanza: per tale motivo introduciamo la seguente definizione.

Definizione 18.1 (Autovalori, autovettori e autospazi di un endomorfismo).

Sia V uno spazio vettoriale su un campo \mathbb{K} . Sia $f: V \rightarrow V$ un endomorfismo.

- Uno scalare $\lambda \in \mathbb{K}$ si dice *autovalore di f* se esiste $v \in V \setminus \{0_V\}$ tale che $f(v) = \lambda v$.
- Se $\lambda \in \mathbb{K}$ è un autovalore di f , ogni vettore $v \in V \setminus \{0_V\}$ tale che $f(v) = \lambda v$ si dice *autovettore di f relativo a λ* .
- Se $\lambda \in \mathbb{K}$ è un autovalore di f , l'insieme

$$E_f(\lambda) = \{ v \in V \mid f(v) = \lambda v \}$$

viene detto *autospazio di f relativo a λ* .

Se $A \in \mathbb{K}^{n,n}$ si parla di autovalori, autovettori e autospazi di A in luogo di μ_A .

In particolare l'autospazio $E_f(\lambda)$ è l'unione dell'insieme degli autovettori di f relativi a λ e l'insieme $\{0_V\}$.

Siano V uno spazio vettoriale su un campo \mathbb{K} , $f: V \rightarrow V$ un endomorfismo e λ un suo autovalore. Si noti che, nella definizione di autovalore, l'esistenza di $v \in V$ non nullo tale che $f(v) = \lambda v$ è necessaria affinché la definizione sia ben posta: infatti se cancellassimo la condizione $v \neq 0_V$ ogni $\lambda \in \mathbb{K}$ sarebbe autovalore di f ! Questa condizione di non nullità implica $E_f(\lambda) \neq \{0_V\}$ per ogni autospazio.

Proposizione 18.2.

Sia V uno spazio vettoriale su un campo \mathbb{K} . Dato un endomorfismo $f: V \rightarrow V$, allora $\lambda \in \mathbb{K}$ è autovalore di f se e solo se $f - \lambda id_V: V \rightarrow V$ non è iniettivo.

In tal caso $E_f(\lambda) = \ker(f - \lambda id_V)$: in particolare $E_f(\lambda)$ è un sottospazio di V .

Dimostrazione. Uno scalare $\lambda \in \mathbb{K}$ è autovalore di f se e solo se esiste $v \in V \setminus \{0_V\}$ tale che

$$\begin{aligned} f(v) = \lambda v &\Leftrightarrow (f - \lambda id_V)(v) = f(v) - \lambda v = 0_V \\ &\Leftrightarrow 0_V \neq v \in \ker(f - \lambda id_V). \end{aligned} \quad \square$$

In particolare, se $A \in \mathbb{K}^{n,n}$ è una matrice, gli autovalori di A sono i $\lambda \in \mathbb{K}$ per cui

$$\begin{aligned} \ker(\mu_A - \lambda id_{\mathbb{K}^n}) &= \ker(\mu_A - \lambda \mu_{I_n}) = \ker(\mu_{A - \lambda I_n}) \\ &= \{ X \in \mathbb{K}^{n,1} \mid (A - \lambda I_n)X = 0_{n,1} \} \neq \{0_{n,1}\} \end{aligned}$$

e, in tal caso, $E_A(\lambda) = \{ X \in \mathbb{K}^n \mid (A - \lambda I_n)X = 0_{n,1} \}$. Quindi gli autovalori di A sono i numeri $\lambda \in \mathbb{K}$ tali che il sistema $(A - \lambda I_n)X = 0_{n,1}$ abbia soluzioni non banali, cioè tali che $\text{rk}(A - \lambda I_n) \leq n - 1$.

Esempio 18.3. Si consideri la matrice

$$A = \begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix}.$$

Verifichiamo se qualcuna fra le entrate a di A è suo autovalore. Per la Proposizione 18.2, si tratta di determinare per quale entrata a di A risulta $\text{rk}(A - aI_2) \leq 1$.

Poiché

$$\begin{aligned} \text{rk}(A - I_2) &= \text{rk} \begin{pmatrix} 0 & 2 \\ 3 & -5 \end{pmatrix} = 2, & \text{rk}(A - 2I_2) &= \text{rk} \begin{pmatrix} -1 & 2 \\ 3 & -6 \end{pmatrix} = 1, \\ \text{rk}(A - 3I_2) &= \text{rk} \begin{pmatrix} -2 & 2 \\ 3 & -7 \end{pmatrix} = 2, & \text{rk}(A + 4I_2) &= \text{rk} \begin{pmatrix} 5 & 2 \\ 3 & 0 \end{pmatrix} = 2, \end{aligned}$$

deduciamo che l'unica entrata di A che sia anche suo autovalore è 2.

Verifichiamo che anche -5 è autovalore di A . Infatti

$$\text{rk}(A + 5I_2) = \text{rk} \begin{pmatrix} 6 & 2 \\ 3 & 1 \end{pmatrix} = 1. \quad \spadesuit$$

☠ Da questo esempio ricaviamo alcune osservazioni. Non è detto che gli autovalori di una matrice vadano cercati fra le entrate della matrice stessa (-5 non è entrata di A). In particolare non è detto che gli autovalori di una matrice siano le sue entrate diagonali della matrice stessa (2 e -5 non sono entrate diagonali di A).

Vedremo in seguito che la ricerca degli autovalori di una matrice o di un endomorfismo è un problema assai più sottile, talvolta difficile o anche impossibile da risolvere in maniera esatta!

Esempio 18.4. Sia $I =]a, b[\subseteq \mathbb{R}$ un intervallo aperto e sia $D: \mathcal{C}^\infty(I) \rightarrow \mathcal{C}^\infty(I)$ l'endomorfismo definito nell'Esempio 16.8. Allora $\lambda \in \mathbb{R}$ è autovalore se e solo se esiste $\varphi \in \mathcal{C}^\infty(I)$, $\varphi \neq 0$, tale che $D\varphi = \lambda\varphi$.

Dall'analisi sappiamo che ogni funzione $\varphi(x) = \exp(\lambda x)$ soddisfa tale relazione (è un' *autofunzione* di D): in particolare ogni $\lambda \in \mathbb{R}$ è autovalore di D . Studiando le equazioni differenziali si avrà modo di verificare che $E_D(\lambda) = \mathcal{L}(\exp(\lambda x))$.

Discorsi analoghi si possono applicare a ogni operatore differenziale lineare della forma

$$P(D) = a_0 D^n + a_1 D^{n-1} + \dots + a_{n-1} D + a_n: \mathcal{C}^\infty(I) \rightarrow \mathcal{C}^\infty(I),$$

con $a_0, \dots, a_n \in \mathcal{C}^\infty(I)$. ♠

18.2 Autovalori, autovettori e autospazi di matrici

Restringiamoci ora allo studio degli autovalori, autovettori e autospazi di endomorfismi di spazi di dimensione finita. In tal caso è sufficiente limitarsi a determinare gli autovalori, autovettori, autospazi della corrispondente matrice rispetto a un'opportuna base: per tale motivo d'ora innanzi ci limiteremo al caso di matrici.

Siano \mathbb{K} un campo e $A \in \mathbb{K}^{n,n}$: abbiamo visto che un autovalore di A è uno scalare $\lambda \in \mathbb{K}$ tale che $\text{rk}(A - \lambda I_n) \leq n - 1$ o, equivalentemente, un elemento $\lambda \in \mathbb{K}$ tale che $\det(A - \lambda I_n) = 0$.

Sviluppando tale determinante con la regola di Laplace, si può verificare facilmente che

$$p_A(t) = \det(A - tI_n) = (-1)^n t^n + a_1 t^{n-1} + a_2 t^{n-2} + \dots + a_{n-1} t + a_n$$

è un polinomio di grado esattamente n , dove i coefficienti a_i sono polinomi nelle entrate di A : in particolare $p_A(t) \in \mathbb{K}[t]$.

Ricordiamo che ogni polinomio di grado n a coefficienti in \mathbb{C} può essere fattorizzato in un prodotto di n polinomi di grado 1, cioè

$$p_A(t) = (-1)^n (t - \lambda_1) \cdots (t - \lambda_n),$$

dove $\lambda_1, \dots, \lambda_n$ sono le radici in \mathbb{C} , non necessariamente distinte, del polinomio $p_A(t)$.

Esistono vari legami tra i coefficienti a_i e le radici λ_j . Per esempio il termine noto a_n di $p_A(t)$ è esattamente il prodotto di $(-1)^n$ per il prodotto dei termini noti dei binomi $t - \lambda_j$: d'altra parte $a_n = p_A(0) = \det(A - 0I_n) = \det(A)$, quindi

$$\det(A) = p_A(0) = a_n = \lambda_1 \cdots \lambda_n.$$

Invece il coefficiente a_1 di t^{n-1} è esattamente il prodotto di $(-1)^{n+1}$ per la somma dei termini noti dei binomi $t - \lambda_j$: d'altra parte, sviluppando il determinante, ci si rende conto che tale coefficiente è anche il prodotto di $(-1)^{n-1}$ per la somma $Tr(A)$ delle entrate diagonali di A (tale somma viene detta *traccia di A*) cioè

$$(-1)^{n-1}Tr(A) = a_1 = (-1)^{n+1}\lambda_1 + \cdots + \lambda_n,$$

da cui segue

$$Tr(A) = \lambda_1 + \cdots + \lambda_n.$$

Definizione 18.5 (Polinomio caratteristico).

Siano \mathbb{K} un campo e $A \in \mathbb{K}^{n,n}$. Il polinomio $p_A(t) = \det(A - tI_n)$ è chiamato *polinomio caratteristico di A*. L'equazione $p_A(t) = 0$ è detta *equazione caratteristica di A*.

Proposizione 18.6. *Siano \mathbb{K} un campo e $A \in \mathbb{K}^{n,n}$. Allora gli autovalori in \mathbb{K} di A sono le radici in \mathbb{K} del polinomio caratteristico di A . In particolare A ha al più n autovalori.*

Da quanto visto sopra deduciamo che, se $A \in \mathbb{R}^{n,n}$ con n dispari, essa ha almeno un autovalore in \mathbb{R} . Invece se $A \in \mathbb{C}^{n,n}$ ogni radice del suo polinomio caratteristico è autovalore di A su \mathbb{C} .

Esempio 18.7. Gli autovalori dipendono non solo dall'endomorfismo, ma anche dal campo \mathbb{K} . Si consideri ad esempio la matrice

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^{2,2} \subseteq \mathbb{C}^{2,2}.$$

Gli autovalori di A sono le radici di

$$p_A(t) = \begin{vmatrix} -t & 1 \\ -1 & -t \end{vmatrix} = t^2 + 1.$$

In particolare A non ha autovalori in \mathbb{R} , mentre ha autovalori $\pm i \in \mathbb{C}$ come matrice su \mathbb{C} . Per calcolare l'autospazio $E_A(i) \subseteq \mathbb{C}^2$ dobbiamo risolvere il sistema

$$\begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} :$$

quindi $E_A(i) = \mathcal{L}((1, i))$. Similmente, per determinare l'autospazio $E_A(-i) \subseteq \mathbb{C}^2$ si deve risolvere il sistema

$$\begin{pmatrix} i & 1 \\ -1 & i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} :$$

quindi $E_A(-i) = \mathcal{L}((1, -i))$. ♠

Esempio 18.8. Ritorniamo all'Esempio 18.3: gli autovalori della matrice

$$A = \begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix} \in \mathbb{R}^{2,2}$$

sono le radici (in \mathbb{R}) del polinomio

$$p_A(t) = \begin{vmatrix} 1-t & 2 \\ 3 & -4-t \end{vmatrix} = t^2 + 3t - 10 = (t-2)(t+5),$$

quindi 2 e -5 . Per calcolare l'autospazio $E_A(2) \subseteq \mathbb{R}^2$ dobbiamo risolvere il sistema

$$\begin{pmatrix} -1 & 2 \\ 3 & -6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} :$$

quindi $E_A(2) = \mathcal{L}((2, 1))$. Similmente, per determinare l'autospazio $E_A(-5) \subseteq \mathbb{R}^2$ si deve risolvere il sistema

$$\begin{pmatrix} 6 & 2 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} :$$

quindi $E_A(-5) = \mathcal{L}((1, -3))$. ♠

Esempio 18.9. Si consideri la matrice

$$A = \begin{pmatrix} -1 & 2 & 2 \\ 1 & -2 & 1 \\ 3 & 3 & 0 \end{pmatrix} \in \mathbb{R}^{3,3}.$$

Gli autovalori di A sono le radici (in \mathbb{R}) del polinomio

$$p_A(t) = \begin{vmatrix} -1-t & 2 & 2 \\ 1 & -2-t & 1 \\ 3 & 3 & -t \end{vmatrix} = -(t+3)^2(t-3),$$

cioè ± 3 . Per calcolare l'autospazio $E_A(3)$ risolviamo il sistema

$$\begin{pmatrix} -4 & 2 & 2 \\ 1 & -5 & 1 \\ 3 & 3 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} :$$

quindi $E_A(3) = \mathcal{L}((2, 1, 3)) \subseteq \mathbb{R}^3$. Similmente, per determinare l'autospazio $E_A(-3)$ risolviamo il sistema

$$\begin{pmatrix} 2 & 2 & 2 \\ 1 & 1 & 1 \\ 3 & 3 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} :$$

quindi $E_A(-3) = \mathcal{L}((-1, 1, 0), (-1, 0, 1)) \subseteq \mathbb{R}^3$. ♠

Esempio 18.10. Si consideri la matrice

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{3,3}$$

che ha polinomio caratteristico

$$p_A(t) = \begin{vmatrix} 2-t & 0 & 0 \\ 1 & 1-t & 1 \\ 0 & 0 & 1-t \end{vmatrix} = -(t-1)^2(t-2).$$

Gli autovalori di A sono 1 e 2: calcoliamone gli autospazi relativi. Per $E_A(1)$ risolviamo il sistema

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

il cui spazio delle soluzioni è $E_A(1) = \mathcal{L}((0, 1, 0)) \subseteq \mathbb{R}^3$. Similmente, calcoliamo $E_A(2)$ a partire dal sistema

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} :$$

il suo spazio delle soluzioni è $E_A(2) = \mathcal{L}((1, 1, 0)) \subseteq \mathbb{R}^3$. ♠

Esempio 18.11. Si consideri la matrice

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 2 \end{pmatrix} \in \mathbb{R}^{3,3}$$

che ha polinomio caratteristico

$$p_A(t) = \begin{vmatrix} 1-t & 2 & 0 \\ 0 & 1-t & 1 \\ -1 & 0 & 2-t \end{vmatrix} = -t(t^2 - 4t + 5).$$

La matrice A ha un unico autovalore in \mathbb{R} . Calcoliamo il corrispondente autospazio $E_A(0) = \ker(\mu_A)$: a tale scopo risolviamo il sistema

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

il cui spazio delle soluzioni è $E_A(0) = \mathcal{L}(2, -1, 1) \subseteq \mathbb{R}^3$. ♠

Si noti che in tutti i casi sopra esaminati la dimensione di un certo autospazio $E_A(\lambda)$ è limitata dalla molteplicità di λ come radice del polinomio caratteristico $p_A(t)$, ma non è necessariamente uguale a essa.

Definizione 18.12 (Molteplicità algebrica e molteplicità geometrica).

Siano \mathbb{K} un campo, $A \in \mathbb{K}^{n,n}$, $\lambda \in \mathbb{K}$ un suo autovalore.

- La *molteplicità algebrica* di λ è la sua molteplicità come radice di $p_A(t)$ e si denota $m_a(\lambda, A)$.
- La *molteplicità geometrica* di λ è il numero $m_g(\lambda, A) = \dim_{\mathbb{K}}(E_A(\lambda))$.

Osserviamo che, in base a quanto visto nella Lezione 17, si ha

$$m_g(\lambda, A) = n - \text{rk}(A - \lambda I_n).$$

Inoltre per definizione $m_g(\lambda, A) \geq 1$. Il seguente risultato, di cui omettiamo la dimostrazione, mette in relazione le due nozioni di molteplicità.

Proposizione 18.13. *Siano \mathbb{K} un campo, $A \in \mathbb{K}^{n,n}$, $\lambda \in \mathbb{K}$ un suo autovalore. Allora $1 \leq m_g(\lambda, A) \leq m_a(\lambda, A)$.*

Concludiamo il capitolo con la definizione di spettro di un endomorfismo.

Definizione 18.14 (Spettro di un endomorfismo).

Sia V uno spazio vettoriale su un campo \mathbb{K} , non necessariamente di dimensione finita.

Se $f: V \rightarrow V$ è un endomorfismo si definisce *spettro* di f su \mathbb{K} l'insieme

$$\text{Sp}_{\mathbb{K}}(f) = \{ \lambda \in \mathbb{K} \mid f - \lambda id_V \text{ non è invertibile} \}.$$

Se V è finitamente generato, in forza di quanto visto finora, l'insieme $\text{Sp}_{\mathbb{K}}(f)$ coincide esattamente con l'insieme degli autovalori di f . Ciò non è più vero se V non è finitamente generato, come mostra il seguente esempio.

Esempio 18.15. Sia \mathbb{K} un campo e si consideri l'applicazione

$$\begin{aligned} f: \mathbb{K}[x] &\longrightarrow \mathbb{K}[x] \\ p(x) &\mapsto xp(x). \end{aligned}$$

Risulta che

$$(f - \lambda id_V)(p(x)) = f(p(x)) - \lambda p(x) = xp(x) - \lambda p(x) = (x - \lambda)p(x),$$

che è iniettiva per ogni $\lambda \in \mathbb{K}$: concludiamo che f non ha autovalori.

D'altra parte $\text{Sp}_{\mathbb{K}}(f) = \mathbb{K}$ poiché $f - \lambda id_V$ non è mai suriettiva: infatti in $\text{im}(f - \lambda id_V)$ non ci sono polinomi di grado 0. ♠

Diagonalizzazione

19.1 Diagonalizzazione di matrici

Come visto nel precedente capitolo, in generale, data una matrice $A \in \mathbb{K}^{n,n}$ con \mathbb{K} campo, per esempio $\mathbb{K} = \mathbb{R}, \mathbb{C}$, non è immediato che esista sempre una base costituita da suoi autovettori.

Definizione 19.1 (Matrici diagonalizzabili).

Siano \mathbb{K} un campo e $A \in \mathbb{K}^{n,n}$. Allora A si dice *diagonalizzabile su \mathbb{K}* se esistono n autovettori di A linearmente indipendenti.

Quando, come spesso accade, il campo su cui si lavora è fissato, si parla semplicemente di matrice diagonalizzabile omettendo l'indicazione del campo.

In particolare la matrice $A \in \mathbb{K}^{n,n}$ è diagonalizzabile se e solo se l'endomorfismo μ_A di \mathbb{K}^n è semplice.

È chiaro che ogni matrice diagonale D è diagonalizzabile! Infatti se

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

allora $De_j = \lambda_j e_j$, dunque esistono n autovettori di D linearmente indipendenti, precisamente $e_1, \dots, e_n \in \mathbb{K}^n$.

Esempio 19.2. La matrice

$$A = \begin{pmatrix} -1 & 2 & 2 \\ 1 & -2 & 1 \\ 3 & 3 & 0 \end{pmatrix} \in \mathbb{R}^{3,3}$$

dell'Esempio 18.9 è diagonalizzabile: infatti in tale esempio abbiamo visto che A ha due autovalori ± 3 , con autospazi relativi rispettivamente $E_A(3) = \mathcal{L}((2, 1, 3)) \subseteq \mathbb{R}^3$ e $E_A(-3) = \mathcal{L}((-1, 1, 0), (-1, 0, 1)) \subseteq \mathbb{R}^3$.

Si noti che i vettori $P_1 = (2, 1, 3)$, $P_2 = (-1, 1, 0)$ e $P_3 = (-1, 0, 1)$ sono linearmente indipendenti: infatti la matrice

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

avente tali vettori come righe, ha rango 3. Concludiamo che $\mathcal{B} = (P_1, P_2, P_3)$ è una base di \mathbb{R}^3 formata da autovettori di A . ♠

Sia \mathbb{K} un campo; ricordiamo che gli autovalori di $A \in \mathbb{K}^{n,n}$ sono le radici $\lambda_1, \dots, \lambda_h \in \mathbb{K}$ del polinomio caratteristico $p_A(t)$. Inoltre a ognuno degli autovalori $\lambda_i \in \mathbb{K}$ di A rimangono associati due numeri interi non negativi, la sua molteplicità algebrica $m_a(\lambda, A)$ e la sua molteplicità geometrica $m_g(\lambda, A)$.

La somma delle molteplicità delle radici di un polinomio è al massimo al grado del polinomio stesso. Quindi, se $\lambda_1, \dots, \lambda_h$ sono a due e a due distinti, risulta

$$m_a(\lambda_1, A) + \dots + m_a(\lambda_h, A) \leq n,$$

e, se vale l'uguaglianza, tutte le radici di $p_A(t)$ devono essere in \mathbb{K} .

Quindi, se $\lambda_1, \dots, \lambda_h \in \mathbb{K}$ sono le radici di $p_A(t)$, tenendo conto della Proposizione 18.13, al massimo possiamo determinare

$$m_g(\lambda_1, A) + \dots + m_g(\lambda_h, A) \leq m_a(\lambda_1, A) + \dots + m_a(\lambda_h, A) \leq n$$

autovettori linearmente indipendenti. Se vale l'uguaglianza, tutte le radici λ di $p_A(t)$ devono essere in \mathbb{K} e si deve avere $m_g(\lambda, A) = m_a(\lambda, A)$ per ognuna di esse.

In particolare, sia nel caso in cui non tutte le radici di $p_A(t)$ sono in \mathbb{K} , che nel caso in cui lo sono, ma esiste almeno una di esse per cui $m_g(\lambda, A) < m_a(\lambda, A)$, la matrice A non è diagonalizzabile.

Esempio 19.3. Si considerino le matrici di $\mathbb{R}^{3,3}$

$$A_1 = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 2 \end{pmatrix}.$$

Nell'Esempio 18.10 abbiamo visto che A_1 ha come autovalori i numeri 1 e 2 e che $m_a(2, A_1) = 1 = m_g(2, A_1)$, $m_a(1, A_1) = 2 > 1 = m_g(1, A_1)$.

Nell'Esempio 18.11 abbiamo visto che A_2 ha come unico autovalore in \mathbb{R} il numero 0 e che $m_a(0, A_2) = 1 = m_g(0, A_2)$. Invece su \mathbb{C} tale matrice ha i numeri 0, $2+i$ e $2-i$ come autovalori e $m_a(0, A_2) = 1 = m_g(0, A_2)$, $m_a(2+i, A_2) = 1 = m_g(2+i, A_2)$, $m_a(2-i, A_2) = 1 = m_g(2-i, A_2)$.

Concludiamo che A_1 non è diagonalizzabile su \mathbb{R} . Per quanto riguarda A_2 , è evidente che essa non è diagonalizzabile su \mathbb{R} . Se, invece si pensa ad A_2 come matrice a coefficienti complessi, si verifica che $E_{A_2}(2+i) = \mathcal{L}((2i, i-1, -2))$ ed $E_{A_2}(2-i) = \mathcal{L}((2i, i+1, 2))$. Poiché risulta $E_{A_2}(0) = \mathcal{L}((2, -1, 1))$ e i tre vettori $(2i, i-1, -2)$, $(2i, i+1, 2)$, $(2, -1, 1)$ sono linearmente indipendenti, segue che A_2 , come matrice a coefficienti in \mathbb{C} , è diagonalizzabile. ♠

Proposizione 19.4.

Siano \mathbb{K} un campo e $A \in \mathbb{K}^{n,n}$. Se $\lambda_1, \dots, \lambda_h \in \mathbb{K}$ sono autovalori a due a due distinti di A e $P_i \in E_A(\lambda_i)$, $i = 1, \dots, h$, vettori non nulli, allora i P_1, \dots, P_h sono linearmente indipendenti.

Dimostrazione. Siano $\lambda_1, \dots, \lambda_h \in \mathbb{K}$ tutte le radici di $p_A(t)$ in \mathbb{K} a due a due distinte e siano $P_i \in E_A(\lambda_i)$, per $i = 1, \dots, h$ vettori non nulli. Se ci fosse una relazione di dipendenza lineare tra i vettori P_1, \dots, P_h , uno di loro sarebbe combinazione lineare di quelli che lo precedono per la Proposizione 13.17, punto (ii).

Sia q il minimo intero per cui ciò accade: allora $P_q = \alpha_1 P_1 + \dots + \alpha_{q-1} P_{q-1}$ e almeno uno fra $\alpha_1, \dots, \alpha_{q-1}$ deve essere non nullo, altrimenti $P_q = 0_{n,1}$. Segue che

$$\begin{aligned} \alpha_1 \lambda_q P_1 + \dots + \alpha_{q-1} \lambda_q P_{q-1} &= \lambda_q (\alpha_1 P_1 + \dots + \alpha_{q-1} P_{q-1}) \\ &= \lambda_q P_q = A P_q = A (\alpha_1 P_1 + \dots + \alpha_{q-1} P_{q-1}) \\ &= \alpha_1 A P_1 + \dots + \alpha_{q-1} A P_{q-1} = \alpha_1 \lambda_1 P_1 + \dots + \alpha_{q-1} \lambda_{q-1} P_{q-1}, \end{aligned}$$

da cui si ricava

$$\alpha_1 (\lambda_q - \lambda_1) P_1 + \dots + \alpha_{q-1} (\lambda_q - \lambda_{q-1}) P_{q-1} = 0_{n,1}.$$

Poiché, per ipotesi, P_1, \dots, P_{q-1} sono linearmente indipendenti e $\lambda_q - \lambda_i \neq 0$, $i = 1, \dots, q-1$, segue che deve essere $\alpha_1 = \dots = \alpha_{q-1} = 0$. \square

Concludiamo questo paragrafo con il seguente risultato fondamentale.

Proposizione 19.5.

Siano \mathbb{K} un campo e $A \in \mathbb{K}^{n,n}$. La matrice A è diagonalizzabile su \mathbb{K} se e solo se valgono le due seguenti condizioni:

- (i) tutte le radici di $p_A(t)$ sono in \mathbb{K} ;
- (ii) per ogni radice $\lambda \in \mathbb{K}$ di $p_A(t)$, risulta $m_a(\lambda, A) = m_g(\lambda, A)$.

Dimostrazione. Se la matrice A è diagonalizzabile, abbiamo già dimostrato che devono valere le affermazioni (i) e (ii).

Viceversa, supponiamo che le due condizioni siano verificate. Siano $\lambda_1, \dots, \lambda_h$ le radici a due a due distinte di $p_A(t)$ e sia $(P_{j,1}, \dots, P_{j,m_j})$ una base di $E_A(\lambda_j)$, per $j = 1, \dots, h$ (quindi $m_j = m_g(\lambda_j, A)$). Per ipotesi i vettori vale che

$$m_g(\lambda_1, A) + \dots + m_g(\lambda_h, A) = m_a(\lambda_1, A) + \dots + m_a(\lambda_h, A) = n.$$

Quindi per verificare che

$$\mathcal{B} = (P_{1,1}, \dots, P_{1,m_1}, P_{2,1}, \dots, P_{2,m_2}, \dots, P_{h,m_h})$$

è una base di \mathbb{K}^n basta verificare che tali vettori sono linearmente indipendenti.

Supponiamo per assurdo che esista una relazione di dipendenza lineare fra tali vettori, diciamo

$$\alpha_{1,1} P_{1,1} + \dots + \alpha_{1,m_1} P_{1,m_1} + \alpha_{2,1} P_{2,1} + \dots + \alpha_{2,m_2} P_{2,m_2} + \dots + \alpha_{h,m_h} P_{h,m_h} = 0_{n,1}.$$

Sappiamo che $\alpha_{j,1} P_{j,1} + \dots + \alpha_{j,m_j} P_{j,m_j} \in E_A(\lambda_j)$ per ogni $j = 1, \dots, h$, quindi deduciamo che $\alpha_{j,1} P_{j,1} + \dots + \alpha_{j,m_j} P_{j,m_j} = 0_{n,1}$, in contraddizione con il fatto che $(P_{1,j}, \dots, P_{m_j,j})$ è base di $E_A(\lambda_j)$. \square

Osservazione 19.6. Chiariamo il motivo per cui si parla di matrici diagonalizzabili. Supponiamo che \mathbb{K} sia un campo e $A \in \mathbb{K}^{n,n}$ diagonalizzabile. Siano $\lambda_1, \dots, \lambda_n$ le radici, non necessariamente distinte, di $p_A(t)$ (che sono tutte in \mathbb{K} per ipotesi) e

$P_j \in E_A(\lambda_j)$ autovettori linearmente indipendenti di A . Sia $P \in \mathbb{K}^{n,n}$ la matrice avente P_j come j -esima colonna: allora

$$AP_j = \lambda_j P_j, \quad j = 1, \dots, n.$$

Quindi $AP = PD$ ove D è la matrice diagonale avente nell'ordine $\lambda_1, \dots, \lambda_n$ come entrate diagonali. Per costruzione P è invertibile. Concludiamo che se $A \in \mathbb{K}^{n,n}$ è diagonalizzabile allora esiste $P \in \mathbb{K}^{n,n}$ invertibile tale che $P^{-1}AP = D$ sia diagonale.

Viceversa, se ciò accade, procedendo a ritroso con il ragionamento sopra, si verifica che A è diagonalizzabile, che P ha per colonne n autovettori di A linearmente indipendenti e che l'elemento j -esimo sulla diagonale di D è esattamente l'autovalore corrispondente alla colonna j -esima di P .

⚠ Quindi le matrici diagonalizzabili sono tutte e sole le matrici $A \in \mathbb{K}^{n,n}$ per cui esiste $P \in \mathbb{K}^{n,n}$ invertibile tale che $P^{-1}AP$ sia diagonale.

Una definizione importante in algebra lineare è la seguente.

Definizione 19.7 (Matrici simili).

Siano $A, B \in \mathbb{K}^{n,n}$. Allora A si dice *simile* a B , e si scrive $A \sim B$, se esiste $P \in \mathbb{K}^{n,n}$ invertibile tale che $P^{-1}AP = B$.

⚠ In base alla definizione sopra, segue che una matrice $A \in \mathbb{K}^{n,n}$ è diagonalizzabile su \mathbb{K} se e solo se è simile a una matrice diagonale in \mathbb{K} .

Esempio 19.8. Riprendiamo la matrice A dell'Esempio 18.3

$$A = \begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix} \in \mathbb{R}^{2,2}.$$

Come visto nell'Esempio 18.8, i suoi autovalori sono 2 e -5 e $E_A(2) = \mathcal{L}((2, 1))$, $E_A(-5) = \mathcal{L}((1, -3))$.

Consideriamo adesso la matrice

$$P_1 = \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix};$$

essa è invertibile e $AP_1 = P_1D$, ovvero $P_1^{-1}AP_1 = D$, con

$$D = \begin{pmatrix} 2 & 0 \\ 0 & -5 \end{pmatrix}.$$

Allo stesso risultato saremmo arrivati prendendo in luogo della matrice P_1 sopra indicata la matrice

$$P_2 = \begin{pmatrix} -2 & -2 \\ -1 & 6 \end{pmatrix}.$$

Invece presa

$$P_3 = \begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix}$$

risulta

$$P_3^{-1}AP_3 = \begin{pmatrix} -5 & 0 \\ 0 & 2 \end{pmatrix}. \quad \spadesuit$$

Esempio 19.9. Riprendiamo la matrice A dell'Esempio 18.9

$$A = \begin{pmatrix} -1 & 2 & 2 \\ 1 & -2 & 1 \\ 3 & 3 & 0 \end{pmatrix} \in \mathbb{R}^{3,3}.$$

Abbiamo visto che i suoi autovalori sono ± 3 e che gli autospazi relativi sono rispettivamente $E_A(3) = \mathcal{L}((2, 1, 3))$ e $E_A(-3) = \mathcal{L}((-1, 1, 0), (-1, 0, 1))$.

Consideriamo adesso le tre matrici

$$P_1 = \begin{pmatrix} 2 & -1 & -1 \\ 1 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}, \quad P_2 = \begin{pmatrix} -2 & 2 & 0 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{pmatrix}, \quad P_3 = \begin{pmatrix} -1 & -1 & -4 \\ 1 & 0 & -2 \\ 0 & 1 & -6 \end{pmatrix}.$$

Per $i = 1, 2, 3$, le matrici P_i sono invertibili e $P_i^{-1}AP_i = D_i$, con

$$D_1 = \begin{pmatrix} 3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{pmatrix}, \quad D_2 = \begin{pmatrix} -3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -3 \end{pmatrix}, \quad D_3 = \begin{pmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 3 \end{pmatrix}. \quad \spadesuit$$

19.2 Diagonalizzazione di matrici simmetriche

Come visto nel paragrafo precedente, il fatto che una matrice sia diagonalizzabile o meno non può essere, in generale, stabilito a priori, ma solo dopo lo studio dei suoi autospazi. Esiste però una classe di matrici la cui diagonalizzabilità è assicurata da un risultato generale di cui omettiamo la dimostrazione e su cui torneremo nelle prossime lezioni.

Proposizione 19.10.

Sia $A \in \mathbb{R}^{n,n}$ simmetrica. Allora A è diagonalizzabile su \mathbb{R} .

Per le matrici simmetriche a coefficienti reali è dunque, assicurata la diagonalizzabilità su \mathbb{R} , cioè l'esistenza di una matrice invertibile $P \in \mathbb{R}^{n,n}$ tale che $P^{-1}AP = D \in \mathbb{R}^{n,n}$ sia diagonale.

Esempio 19.11. Si consideri la matrice

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

Risulta

$$p_A(t) = \begin{vmatrix} -t & 1 & 1 \\ 1 & -t & 1 \\ 1 & 1 & -t \end{vmatrix} = -t^3 + 3t + 2 = -(t+1)^2(t-2):$$

concludiamo che gli autovalori di A sono -1 e 2 . Inoltre per la Proposizione 19.10 sappiamo che $m_a(-1, A) = m_g(-1, A) = 2$ e $m_a(2, A) = m_g(2, A) = 1$.

Per determinare l'autospazio $E_A(-1)$ risolviamo il sistema

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

ottenendo $E_A(-1) = \mathcal{L}((1, -1, 0), (1, 0, -1))$. Similmente, per determinare $E_A(2)$ risolviamo il sistema

$$\begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

quindi $E_A(2) = \mathcal{L}((1, 1, 1))$. Posto

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

risulta

$$P^{-1}AP = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}. \quad \spadesuit$$

Osservazione 19.12. Per renderci conto dell'importanza della Proposizione 19.10 osserviamo che, spesso, è assai difficile determinare esattamente gli autovalori di una matrice: può però essere utile poterne determinare la diagonalizzabilità. Per esempio la matrice

$$A = \begin{pmatrix} 1/11 & -3/2 & \pi & \sqrt{21} \\ -3/2 & 0 & 1117 & -7^{3/4} \\ \pi & 1117 & -31 & e \\ \sqrt{21} & -7^{3/4} & e & 1 \end{pmatrix}$$

è senza dubbio diagonalizzabile su \mathbb{R} perché simmetrica a coefficienti reali.

19.3 Il teorema di Cayley–Hamilton

Siano \mathbb{K} un campo e $A \in \mathbb{K}^{n,n}$ una matrice; poiché $\dim_{\mathbb{K}}(\mathbb{K}^{n,n}) = n^2$, le $n^2 + 1$ matrici

$$I_n, A, A^2, \dots, A^{n^2-1}, A^{n^2}$$

sono necessariamente linearmente dipendenti, che significa che esistono altrettanti scalari $\alpha_{n^2}, \alpha_{n^2-1}, \alpha_{n^2-2}, \dots, \alpha_1, \alpha_0 \in \mathbb{K}$ tali che

$$\alpha_0 A^{n^2} + \alpha_1 A^{n^2-1} + \dots + \alpha_{n^2-2} A^2 + \alpha_{n^2-1} A + \alpha_{n^2} I_n = 0_{n,n}.$$

Si consideri il polinomio

$$p(t) = \alpha_0 t^{n^2} + \alpha_1 t^{n^2-1} + \cdots + \alpha_{n^2-2} t^2 + \alpha_{n^2-1} t + \alpha_{n^2} \in \mathbb{K}[t] :$$

quanto osservato sopra viene spesso riassunto affermando che A è radice di $p(t)$ o, anche, che $p(A) = 0_{n,n}$.

Poiché le matrici $A, A^2, \dots, A^{n^2-1}, A^{n^2}$ non sono arbitrarie, ma sono potenze di una stessa matrice, è lecito domandarsi se non esista un polinomio di grado più basso di cui A sia radice: a questa domanda risponde il seguente risultato, di cui omettiamo la dimostrazione.

Proposizione 19.13 (Teorema di Cayley–Hamilton).

Siano \mathbb{K} un campo e $A \in \mathbb{K}^{n,n}$. Allora A è radice di $p_A(t)$.

Ciò significa che se il polinomio caratteristico di A è

$$p_A(t) = (-1)^n t^n + a_1 t^{n-1} + a_2 t^{n-2} + \cdots + a_{n-1} t + a_n,$$

allora vale

$$p_A(A) = (-1)^n A^n + a_1 A^{n-1} + a_2 A^{n-2} + \cdots + a_{n-1} A + a_n I_n = 0_{n,n}. \quad (19.3.1)$$

Tale osservazione permette di introdurre un nuovo metodo di inversione di matrici. Infatti A è invertibile se e solo se $a_n = \det(A) \neq 0$, dunque dall'equazione (19.3.1) otteniamo

$$\frac{1}{a_n} \left((-1)^{n-1} A^{n-1} - a_1 A^{n-2} - a_2 A^{n-3} + \cdots - a_{n-1} I_n \right) A = I_n,$$

cioè

$$A^{-1} = \frac{1}{a_n} \left((-1)^{n-1} A^{n-1} - a_1 A^{n-2} - a_2 A^{n-3} + \cdots - a_{n-1} I_n \right). \quad (19.3.2)$$

Esempio 19.14. Si consideri la generica matrice 2×2

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

Allora $p_A(t) = t^2 - (a+d)t + ad - bc$, dunque se $ad - bc \neq 0$, segue dalla formula (19.3.2) che

$$A^{-1} = \frac{1}{ad - bc} (-A + (a+d)I_2) = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \quad \spadesuit$$

Esempio 19.15. Si consideri la matrice

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

Abbiamo visto che $p_A(t) = -(t-3)(t+3)^2 = -t^3 - 3t^2 + 9t + 27$ nell'Esempio 18.9, quindi A ha autovalori ± 3 con $m_a(3, A) = 1$ e $m_a(-3, A) = 2$, in particolare $\det(A) = 3(-3)^2 = 27$, sicché A è invertibile: inoltre abbiamo anche visto che A è diagonalizzabile, e nell'Esempio 19.9 abbiamo trovato $P \in \mathbb{R}^{3,3}$ invertibile tale che

$$P^{-1}AP = D = \begin{pmatrix} 3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{pmatrix},$$

quindi $A = PDP^{-1}$ da cui si ottiene

$$A^2 = PDP^{-1}PDP^{-1} = PD^2P^{-1} = P(9I_3)P^{-1} = 9PP^{-1} = 9I_3,$$

perciò

$$A^{-1} = -\frac{1}{27}(A^2 - 3A - 9I_3) = \frac{1}{9}A = \frac{1}{9} \begin{pmatrix} -1 & 2 & 2 \\ 1 & -2 & 1 \\ 3 & 3 & 0 \end{pmatrix}. \quad \spadesuit$$

Un'altra interessante conseguenza del Teorema di Cayley–Hamilton riguarda le matrici nilpotenti; ricordiamo dal paragrafo 2.3 che una matrice $A \in \mathbb{K}^{n,n}$ si dice *nilpotente* se $A^N = 0_{n,n}$ per qualche intero positivo N .

Corollario 19.16.

Siano \mathbb{K} un campo e $A \in \mathbb{K}^{n,n}$. Allora A è nilpotente se e solo se ha 0 come unico autovalore e $m_a(0, A) = n$.

Dimostrazione. La matrice $A \in \mathbb{K}^{n,n}$ ha 0 come unico autovalore e $m_a(0, A) = n$ se e solo se t^n divide il polinomio caratteristico $p_A(t)$ di A cioè se e solo se $p_A(t) = (-1)^n t^n$.

Se $p_A(t) = (-1)^n t^n$, dalla Proposizione 19.13 segue allora che $0_{n,n} = p_A(A) = (-1)^n A^n$, cioè A è nilpotente.

Viceversa, supponiamo che esista un intero positivo N tale che $A^N = 0_{n,n}$. Questa è un'uguaglianza su $\mathbb{K} \subseteq \mathbb{C}$. Se $\lambda \in \mathbb{C}$ è una qualsiasi radice di $p_A(t)$, allora λ è un autovalore di A vista come matrice a entrate complesse, dunque esiste $X \in \mathbb{C}^{n,1} \setminus \{0_{n,1}\}$ tale che $AX = \lambda X$. Moltiplicando ambo i membri di tale identità per A^{N-1} otteniamo allora che $0_{n,n} = A^N X = \lambda^N X$ da cui segue che $\lambda = 0$ è l'unica radice in \mathbb{C} di $p_A(t)$, necessariamente con molteplicità n . \square

Esempio 19.17. Si consideri la matrice

$$A = \begin{pmatrix} -2 & -5 & -8 \\ 1 & 4 & 7 \\ 0 & -1 & -2 \end{pmatrix}.$$

Il suo polinomio caratteristico è

$$p_A(t) = \begin{vmatrix} -2-t & -5 & -8 \\ 1 & 4-t & 7 \\ 0 & -1 & -2-t \end{vmatrix} = -t^3,$$

quindi $A^3 = 0_{3,3}$, cioè A è nilpotente.



Spazi vettoriali con prodotto scalare

20.1 Prodotti scalari

Definizione 20.1 (Prodotto scalare).

Sia V uno spazio vettoriale su \mathbb{R} . Un *prodotto scalare* su V è un'applicazione

$$\begin{aligned}\langle \cdot, \cdot \rangle: V \times V &\longrightarrow \mathbb{R} \\ (v_1, v_2) &\mapsto \langle v_1, v_2 \rangle\end{aligned}$$

tale che valgono i seguenti assiomi:

- (PS1) per ogni $v_1, v_2 \in V$, si ha $\langle v_1, v_2 \rangle = \langle v_2, v_1 \rangle$ (*il prodotto scalare è commutativo*);
- (PS2) per ogni $v_1, v_2, v_3 \in V$, si ha $\langle v_1, v_2 + v_3 \rangle = \langle v_1, v_2 \rangle + \langle v_1, v_3 \rangle$ (*il prodotto scalare è distributivo rispetto alla somma*);
- (PS3) per ogni $\alpha \in \mathbb{R}$ e $v_1, v_2 \in V$, si ha $\alpha \langle v_1, v_2 \rangle = \langle \alpha v_1, v_2 \rangle$;
- (PS4) per ogni $v \in V \setminus \{0_V\}$, si ha $\langle v, v \rangle > 0$ (*il prodotto scalare è definito positivo*).

Osservazione 20.2. Elenchiamo alcune immediate conseguenze degli assiomi della definizione precedente.

1. Sia $v_0 \in V$ fissato. Allora l'applicazione

$$\begin{aligned}\langle \cdot, v_0 \rangle: V &\longrightarrow \mathbb{R} \\ v &\mapsto \langle v, v_0 \rangle\end{aligned}$$

è lineare. Per la commutatività del prodotto scalare segue anche la linearità dell'applicazione

$$\begin{aligned}\langle v_0, \cdot \rangle: V &\longrightarrow \mathbb{R} \\ v &\mapsto \langle v_0, v \rangle.\end{aligned}$$

Per questa doppia proprietà di linearità si dice spesso che il prodotto scalare è un'applicazione *bilineare*.

2. Per linearità, si ha che $\langle 0, v \rangle = \langle v, 0 \rangle = 0$ per ogni $v \in V$: in particolare $\langle v, v \rangle = 0$ se e solo se $v = 0$.
3. Se $W \subseteq V$ è un sottospazio, ha senso considerare la restrizione $\langle \cdot, \cdot \rangle|_{W \times W}$, che è un prodotto scalare su W .

Definizione 20.3 (Modulo di un vettore).

Sia V uno spazio vettoriale su \mathbb{R} munito di prodotto scalare $\langle \cdot, \cdot \rangle$. Per ogni $v \in V$, il numero $|v| = \sqrt{\langle v, v \rangle}$ si dice *modulo* di v . I vettori di modulo 1 si dicono *versori*.

Si noti che in uno spazio vettoriale V su \mathbb{R} munito di prodotto scalare l'unico vettore di modulo zero è il vettore nullo $\vec{0}_V$.

Esempio 20.4. Nello spazio $V_3(O)$ dei vettori applicati si può definire un prodotto scalare ponendo

$$\langle \vec{v}, \vec{w} \rangle = |\vec{v}| |\vec{w}| \cos(\widehat{\vec{v}\vec{w}}) \quad (20.1.1)$$

per ogni coppia di vettori $\vec{v}, \vec{w} \in V_3(O)$ non nulli, come avevamo visto nella Lezione 8.

Fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$, se $\vec{v} = v_x\vec{i} + v_y\vec{j} + v_z\vec{k}$ e $w = w_x\vec{i} + w_y\vec{j} + w_z\vec{k}$, è noto che

$$\langle \vec{v}, \vec{w} \rangle = v_x w_x + v_y w_y + v_z w_z. \quad \spadesuit$$

Esempio 20.5. Dati $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$ in $V = \mathbb{R}^n$, si definisce *prodotto scalare standard* il prodotto scalare

$$\langle x, y \rangle = x_1 y_1 + \dots + x_n y_n. \quad (20.1.2)$$

Si noti che in tal caso risulta

$$\langle x, y \rangle = {}^t x y = {}^t x I_n y$$

come prodotto di matrici. Il fatto che tale applicazione soddisfi gli assiomi (PS1), (PS2) e (PS3) è evidente dalla definizione. Per quanto riguarda l'assioma (PS4), si noti che

$$\langle x, x \rangle = x_1^2 + \dots + x_n^2$$

è una somma di numeri reali non negativi, come lo sono i quadrati di numeri reali, quindi è non negativa ed è nulla se e solo se tutti gli addendi sono nulli. \spadesuit

$\triangle!$ Quello standard non è l'unico possibile prodotto scalare che possiamo definire su \mathbb{R}^n come mostra l'esempio seguente!

Esempio 20.6. Per esempio si verifichi che l'applicazione

$$((x_1, x_2), (y_1, y_2)) \mapsto 3x_1y_1 + x_2y_2/2$$

è un prodotto scalare in \mathbb{R}^2 diverso dal prodotto standard. Si noti che

$$3x_1y_1 + x_2y_2/2 = (x_1 \ x_2) \begin{pmatrix} 3 & 0 \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}. \quad \spadesuit$$

Proposizione 20.7 (Disuguaglianza di Cauchy–Schwartz).

Sia V uno spazio vettoriale su \mathbb{R} munito di prodotto scalare $\langle \cdot, \cdot \rangle$. Per ogni $v, w \in V$ si ha

$$|\langle v, w \rangle| \leq |v||w| \quad (20.1.3)$$

Inoltre vale l'uguaglianza in (20.1.3) se e solo se v e w sono proporzionali.

Dimostrazione. Se w è nullo, allora l'enunciato è banalmente vero. Dimostriamo allora l'enunciato nell'ipotesi che w sia non nullo. In tal caso per ogni $t \in \mathbb{R}$ si ha

$$|v|^2 - 2t\langle v, w \rangle + t^2|w|^2 = \langle v - tw, v - tw \rangle \geq 0.$$

Il primo membro di tale trinomio non può avere radici distinte, dovendo altrimenti cambiare di segno, quindi $\langle v, w \rangle^2 - |v|^2|w|^2 \leq 0$: essendo $|v|, |w| > 0$, segue allora la disuguaglianza (20.1.3) con lo stesso metodo usato nell'Osservazione 8.9.

Rimane da dimostrare solo la seconda affermazione. Vale l'uguaglianza in (20.1.3) se e solo se l'equazione $|v|^2 - 2t\langle v, w \rangle + t^2|w|^2 = 0$ ha soluzione, ovvero se e solo se $\langle v - tw, v - tw \rangle = 0$ ha soluzione, cioè se e solo se $v = tw$. \square

Osservazione 20.8. Sia V uno spazio vettoriale su \mathbb{R} munito di prodotto scalare $\langle \cdot, \cdot \rangle$. Se $v, w \in V \setminus \{0_V\}$, allora

$$-1 \leq \frac{\langle v, w \rangle}{|v||w|} \leq 1 :$$

possiamo perciò definire l'angolo fra v e w come

$$\widehat{vw} = \arccos \left(\frac{\langle v, w \rangle}{|v||w|} \right).$$

Si ha quindi $\langle v, w \rangle = |v||w| \cos(\widehat{vw})$, che generalizza la formula (20.1.1) a prodotti scalari qualsiasi.

Osservazione 20.9. Sia V uno spazio vettoriale su \mathbb{R} munito di prodotto scalare $\langle \cdot, \cdot \rangle$. Dalla disuguaglianza (20.1.3) ricaviamo anche la *disuguaglianza triangolare* (si veda di nuovo l'Osservazione 8.9):

$$\begin{aligned} |v + w|^2 &= \langle v + w, v + w \rangle = |v|^2 + \langle v, w \rangle + \langle w, v \rangle + |w|^2 \\ &= |v|^2 + 2\langle v, w \rangle + |w|^2 \leq |v|^2 + 2|\langle v, w \rangle| + |w|^2 \\ &\leq |v|^2 + 2|v||w| + |w|^2 = (|v| + |w|)^2, \end{aligned}$$

quindi $|v + w| \leq |v| + |w|$. In modo analogo il lettore verifichi che $|v - w| \geq ||v| - |w||$.

Esempio 20.10. Sia $I = [a, b] \subseteq \mathbb{R}$ non vuoto e si consideri nello spazio $\mathcal{C}^0(I)$ l'applicazione

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx.$$

È immediato verificare che gli assiomi di prodotto scalare (PS1), (PS2), (PS3) sono soddisfatti; inoltre il teorema della permanenza del segno per funzioni continue ci assicura che lo stesso vale per l'assioma (PS4).

In questo caso la disuguaglianza di Cauchy–Schwartz diviene

$$\left| \int_a^b f(x)g(x)dx \right| \leq \sqrt{\int_a^b f(x)^2 dx} \sqrt{\int_a^b g(x)^2 dx}.$$

Per ogni $f \in \mathcal{C}^0(I)$ la quantità

$$\|f\|_2 = \sqrt{\int_a^b f(x)^2 dx}$$

viene detta *norma L^2 di f* . ♠

20.2 Basi ortonormali

Definizione 20.11.

Sia V uno spazio vettoriale su \mathbb{R} munito di prodotto scalare $\langle \cdot, \cdot \rangle$.

- I vettori $v_1, v_2 \in V$ si dicono *ortogonali* (o *perpendicolari*) se $\langle v_1, v_2 \rangle = 0$ e in tal caso si scrive $v_1 \perp v_2$.
- L'insieme $\{v_1, \dots, v_n\} \subseteq V$ si dice *ortogonale* se $v_i \perp v_j$ per $i, j \in I$ con $i \neq j$.
- L'insieme $\{v_1, \dots, v_n\} \subseteq V$ si dice *ortonormale* se è ortogonale e i v_i sono versori.
- Se V è finitamente generato, una *base ortonormale* $\mathcal{B} = (v_1, \dots, v_n)$ è una base di V tale che l'insieme $\{v_1, \dots, v_n\}$ sia ortonormale.

Dalla definizione sopra possiamo dedurre che un insieme $\{v_1, \dots, v_n\} \subseteq V$ è ortonormale se per ogni $i, j = 1, \dots, n$

$$\langle v_i, v_j \rangle = \delta_{i,j} = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases}$$

Esempio 20.12. Si fissi un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ nello spazio. Allora l'insieme $\{\vec{i}, \vec{j}, \vec{k}\}$ è ortonormale in $V_3(O)$ rispetto al prodotto scalare geometrico definito nell'Esempio 20.4. Quindi $\mathcal{B} = (\vec{i}, \vec{j}, \vec{k})$ è una base ortonormale di $V_3(O)$. ♠

Esempio 20.13. Si consideri lo spazio \mathbb{R}^n munito del prodotto scalare standard $\langle \cdot, \cdot \rangle$ definito nell'Esempio 20.5. I vettori $\{e_1, \dots, e_n\}$ formano un insieme ortonormale: perciò la base canonica $\mathcal{C} = (e_1, \dots, e_n)$ è una base ortonormale rispetto al prodotto scalare standard.

Invece $\{e_1, e_2\}$ non è ortonormale rispetto al prodotto scalare

$$((x_1, x_2), (y_1, y_2)) \mapsto 3x_1y_1 + x_2y_2/2$$

introdotto nello stesso Esempio. Infatti è vero che $e_1 \perp e_2$, ma $|e_1| = \sqrt{3}$ e $|e_2| = 1/\sqrt{2}$. Concludiamo che, rispetto a tale prodotto scalare, $\{e_1/\sqrt{3}, \sqrt{2}e_2\}$ è ortonormale. ♠

Esempio 20.14. Si consideri lo spazio V delle funzioni continue e periodiche di periodo 2π : per esempio $1, \cos px, \sin px \in V$ per ogni $p \in \mathbb{N}$ (si noti che $\sin px$ e $\cos px$ hanno periodo minimo $2\pi/p$). In V definiamo l'applicazione

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)g(x)dx;$$

è facile verificare che $\langle \cdot, \cdot \rangle$ è un prodotto scalare. Inoltre dall'analisi è noto che

$$\begin{aligned} \left\langle \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\rangle &= 1 \\ \langle \cos px, \cos qx \rangle &= \begin{cases} 1 & \text{se } p = q \neq 0, \\ 0 & \text{se } p \neq q, \end{cases} \\ \langle \sin px, \cos qx \rangle &= 0 \\ \langle \sin px, \sin qx \rangle &= \begin{cases} 1 & \text{se } p = q \neq 0, \\ 0 & \text{se } p \neq q, \end{cases} \end{aligned}$$

quindi, per ogni $N \in \mathbb{N}$, l'insieme $\left\{ \frac{1}{\sqrt{2}}, \cos px, \sin qx \right\}_{p,q=1,\dots,N}$ è ortonormale. ♠

Osserviamo che in tutti gli esempi trattati è stato possibile determinare una base ortonormale. Questo è un risultato generale che si può dimostrare in modo algoritmico, come spiegato nel seguente risultato.

Proposizione 20.15 (Metodo di ortonormalizzazione di Gram–Schmidt).

Sia $V \neq \{0_V\}$ uno spazio vettoriale finitamente generato su \mathbb{R} munito di prodotto scalare $\langle \cdot, \cdot \rangle$. Allora esistono in V basi ortonormali.

Dimostrazione. Sia $\mathcal{B} = (v_1, \dots, v_n)$ una base di V ; vogliamo costruire una base ortonormale a partire da \mathcal{B} . Procediamo nel modo seguente.

Passo (1). Il vettore u_1 è semplicemente uguale a v_1 :

$$u_1 = v_1 .$$

Passo (2). Il vettore u_2 è v_2 meno la proiezione ortogonale di v_2 nella direzione di u_1 , si veda l'Osservazione 8.10:

$$u_2 = v_2 - \frac{\langle v_2, u_1 \rangle}{|u_1|^2} u_1 .$$

Si noti che, per costruzione, $\langle u_2, u_1 \rangle = 0$ e $\mathcal{L}(u_1, u_2) = \mathcal{L}(v_1, v_2)$, quindi u_1 e u_2 sono linearmente indipendenti.

Passo (3). Il vettore u_3 è v_3 cui vengono sottratte le proiezioni ortogonali di v_3 nelle direzioni di u_1 e u_2 :

$$u_3 = v_3 - \frac{\langle v_3, u_1 \rangle}{|u_1|^2} u_1 - \frac{\langle v_3, u_2 \rangle}{|u_2|^2} u_2 .$$

Si noti che, per costruzione, $\langle u_3, u_2 \rangle = \langle u_3, u_1 \rangle = 0$ e $\mathcal{L}(u_1, u_2, u_3) = \mathcal{L}(v_1, v_2, v_3)$, quindi u_1, u_2, u_3 sono linearmente indipendenti.

⋮

Passo (i). Il vettore u_i è v_i cui vengono sottratte le proiezioni ortogonali di v_i nelle direzioni di tutti i vettori precedenti u_1, \dots, u_{i-1} :

$$u_i = v_i - \frac{\langle v_i, u_1 \rangle}{|u_1|^2} u_1 - \frac{\langle v_i, u_2 \rangle}{|u_2|^2} u_2 - \dots - \frac{\langle v_i, u_{i-1} \rangle}{|u_{i-1}|^2} u_{i-1} .$$

Si noti che, per costruzione, $\langle u_i, u_{i-1} \rangle = \dots = \langle u_i, u_1 \rangle = 0$ e $\mathcal{L}(u_1, \dots, u_i) = \mathcal{L}(v_1, \dots, v_i)$, quindi u_1, \dots, u_i sono linearmente indipendenti.

Alla fine otteniamo una base ortogonale (u_1, \dots, u_n) . Normalizzando i vettori ricaviamo la base ortonormale cercata . \square

Esempio 20.16. Dati i vettori di \mathbb{R}^4

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -1 \\ 4 \\ 4 \\ -1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 4 \\ -2 \\ -2 \\ 0 \end{pmatrix},$$

vogliamo determinare una base ortonormale dello spazio $V = \mathcal{L}(v_1, v_2, v_3) \subseteq \mathbb{R}^4$.

Il lettore verifichi per esercizio che $\mathcal{B} = (v_1, v_2, v_3)$ è una base di V . Costruiamo la nuova base ortonormale $\mathcal{B}' = (w_1, w_2, w_3)$ con l'algoritmo di Gram-Schmidt.

Passo (1). Si ha che

$$u_1 = v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Passo (2). Calcoliamo il prodotto scalare $\langle v_2, u_1 \rangle = -1 + 4 + 4 - 1 = 6$. Definiamo poi il vettore

$$u_2 = v_2 - \frac{\langle v_2, u_1 \rangle}{|u_1|^2} u_1 = \begin{pmatrix} -1 \\ 4 \\ 4 \\ -1 \end{pmatrix} - \frac{3}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -5/2 \\ 5/2 \\ 5/2 \\ -5/2 \end{pmatrix} .$$

Passo (3). Calcoliamo i due prodotti scalari $\langle v_3, u_1 \rangle = 0$ e $\langle v_3, u_2 \rangle = -20$ e definiamo il vettore

$$u_3 = v_3 - \frac{\langle v_3, u_1 \rangle}{|u_1|^2} u_1 - \frac{\langle v_3, u_2 \rangle}{|u_2|^2} u_2 = \begin{pmatrix} 4 \\ -2 \\ -2 \\ 0 \end{pmatrix} - 0 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \frac{4}{5} \begin{pmatrix} -5/2 \\ 5/2 \\ 5/2 \\ -5/2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \\ -2 \end{pmatrix}.$$

Concludiamo normalizzando i 3 vettori, quindi la base ortonormale di V cercata è

$$w_1 = \frac{u_1}{|u_1|} = \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix}, \quad w_2 = \frac{u_2}{|u_2|} = \begin{pmatrix} -1/2 \\ 1/2 \\ 1/2 \\ -1/2 \end{pmatrix}, \quad w_3 = \frac{u_3}{|u_3|} = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 0 \\ -1/\sqrt{2} \end{pmatrix}$$

Il lettore verifichi che la base \mathcal{B}' così ottenuta è effettivamente ortonormale. ♠

L'importanza e utilità delle basi ortonormali è data dalla seguente proposizione.

Proposizione 20.17.

Sia V uno spazio vettoriale su \mathbb{R} munito di prodotto scalare $\langle \cdot, \cdot \rangle$. Dato un insieme ortonormale $\{v_1, \dots, v_n\} \subseteq V$, allora:

- (i) v_1, \dots, v_n sono linearmente indipendenti;
- (ii) se V è finitamente generato e $\dim_{\mathbb{R}}(V) = n$, allora $\mathcal{B} = (v_1, \dots, v_n)$ è una base ortonormale di V e si ha $v = \langle v, v_1 \rangle v_1 + \dots + \langle v, v_n \rangle v_n$ per ogni $v \in V$.

Dimostrazione. Per definizione $\{v_1, \dots, v_n\} \subseteq V$ è un insieme di vettori ortonormali se e solo se

$$\langle v_i, v_j \rangle = \begin{cases} 0 & \text{se } i \neq j, \\ 1 & \text{se } i = j, \end{cases}$$

quindi, se $\alpha_1 v_1 + \dots + \alpha_n v_n = 0$ è una relazione di dipendenza lineare, si ha

$$0 = \langle 0, v_j \rangle = \langle \alpha_1 v_1 + \dots + \alpha_n v_n, v_j \rangle = \alpha_1 \langle v_1, v_j \rangle + \dots + \alpha_n \langle v_n, v_j \rangle = \alpha_j.$$

In particolare v_1, \dots, v_n sono linearmente indipendenti.

Se V è finitamente generato e $\dim_{\mathbb{R}}(V) = n$, per la Proposizione 15.7, segue che \mathcal{B} è una base di V . In particolare per ogni $v \in V$ esistono $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ tali che $v = \alpha_1 v_1 + \dots + \alpha_n v_n$. Quindi

$$\langle v, v_j \rangle = \langle \alpha_1 v_1 + \dots + \alpha_n v_n, v_j \rangle = \alpha_1 \langle v_1, v_j \rangle + \dots + \alpha_n \langle v_n, v_j \rangle = \alpha_j$$

per ogni $j = 1, \dots, n$. □

Il coefficiente $\langle v, v_j \rangle$ viene detto *coefficiente di Fourier* (di v rispetto a v_j).

Esempio 20.18. Si consideri lo spazio \mathbb{R}^3 munito del prodotto scalare $\langle \cdot, \cdot \rangle$ definito nell'Esempio 20.5. I tre vettori

$$v_1 = \frac{1}{3}(2, 2, 1), \quad v_2 = \frac{1}{3}(1, -2, 2), \quad v_3 = \frac{1}{3}(-2, 1, 2)$$

formano un insieme $\{v_1, v_2, v_3\}$ ortonormale, quindi $\mathcal{B} = (v_1, v_2, v_3)$ è una base di \mathbb{R}^3 .

Sia $v = (1, 1, 1) \in \mathbb{R}^3$. Allora $\langle v, v_1 \rangle = 5/3$, $\langle v, v_2 \rangle = \langle v, v_3 \rangle = 1/3$: quindi, come è anche facile verificare direttamente, risulta

$$v = \frac{5}{3}v_1 + \frac{1}{3}v_2 + \frac{1}{3}v_3. \quad \spadesuit$$

20.3 Matrici ortogonali

In questo paragrafo introduciamo un'importante famiglia di matrici quadrate invertibili.

Definizione 20.19 (Matrici ortogonali).

Una matrice $P \in \mathbb{R}^{n,n}$ si dice *ortogonale* se vale l'uguaglianza ${}^t P P = I_n$.

Osservazione 20.20. Sia $P \in \mathbb{R}^{n,n}$ ortogonale.

1. La matrice identità I_n è ortogonale; anche ogni matrice ottenuta da I_n cambiando segno a una o più delle sue entrate è ortogonale.
2. Poiché ${}^t P P = I_n$, segue che P è invertibile e $P^{-1} = {}^t P$: in particolare si ha anche $P {}^t P = I_n$. In maniera analoga si dimostra che se $P {}^t P = I_n$ allora anche ${}^t P P = I_n$, cioè P è ortogonale se e solo se $P {}^t P = I_n$.
3. Si ha $1 = \det(I_n) = \det({}^t P P) = \det({}^t P) \det(P) = \det(P)^2$, dunque

$$\det(P) = \pm 1 :$$

quanto osservato sulla matrice identità ci permette di affermare che esistono matrici di entrambi i tipi.

4. Poiché la riga i -esima di ${}^t P$ è la colonna i -esima P_i di P , allora la condizione ${}^t P P = I_n$ si può leggere dicendo che il prodotto scalare standard dell'Esempio 20.5 delle colonne P_i e P_j di P è $\delta_{i,j}$: in altre parole una matrice è ortogonale se e solo se le sue colonne sono un insieme ortonormale, rispetto al prodotto scalare standard, di n vettori di \mathbb{R}^n .
5. Identificando $\mathbb{R}^{1,n}$ con \mathbb{R}^n tramite la trasposta, si verifica in maniera simile che la condizione $P {}^t P = I_n$ equivale a dire che le righe di P formano un insieme ortonormale di n vettori di \mathbb{R}^n rispetto al prodotto scalare standard.

Le matrici ortogonali si dividono, quindi, in due classi non vuote, quelle con determinante 1 e quelle con determinante -1 .

Definizione 20.21 (Matrici ortogonali speciali e non speciali).

Una matrice ortogonale $P \in \mathbb{R}^{n,n}$ si dice *speciale* se $\det(P) = 1$, *non speciale* se $\det(P) = -1$.

Esempio 20.22. Le matrici di $\mathbb{R}^{3,3}$

$$P_1 = \frac{1}{3} \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{pmatrix}, \quad P_2 = \frac{1}{3} \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ -2 & 2 & -1 \end{pmatrix},$$

sono ortogonali. La prima è non speciale, la seconda speciale. ♠

Esempio 20.23. Determiniamo tutte le matrici ortogonali di ordine due; sia

$$P = \begin{pmatrix} p_{1,1} & p_{1,2} \\ p_{2,1} & p_{2,2} \end{pmatrix} \in \mathbb{R}^{2,2}$$

ortogonale. La condizione $P^t P = I_2$ si traduce nel sistema

$$\begin{cases} p_{1,1}^2 + p_{1,2}^2 = 1 \\ p_{1,1}p_{2,1} + p_{1,2}p_{2,2} = 0 \\ p_{2,1}^2 + p_{2,2}^2 = 1. \end{cases}$$

La prima e la terza equazione implicano l'esistenza di $\vartheta, \psi \in [0, 2\pi]$ tali che

$$p_{1,1} = \cos \vartheta, \quad p_{1,2} = -\sin \vartheta, \quad p_{2,1} = \sin \psi, \quad p_{2,2} = \cos \psi.$$

La seconda equazione è allora equivalente a

$$0 = \cos \vartheta \sin \psi - \sin \vartheta \cos \psi = \sin(\psi - \vartheta).$$

In particolare, a meno di multipli di 2π , si deve avere o $\psi = \vartheta$ ovvero $\psi = \vartheta + \pi$. Nel primo caso

$$P = \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix},$$

(e P è ortogonale speciale); nel secondo

$$P = \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ -\sin \vartheta & -\cos \vartheta \end{pmatrix}$$

(e P è ortogonale non speciale).

Concludiamo l'esempio con un'interpretazione geometrica delle matrici ortogonali speciali: consideriamo nel piano due sistemi di riferimento $O\vec{i}\vec{j}$ e $O\vec{i}'\vec{j}'$ e sia

φ l'angolo misurato in senso antiorario fra i versori \vec{i} e \vec{i}' . Allora si deve avere $\vec{i}' = a\vec{i} + b\vec{j}$, $\vec{j}' = c\vec{i} + d\vec{j}$ e si ha, per la Proposizione 20.17,

$$\begin{aligned} a &= \langle \vec{i}', \vec{i} \rangle = \cos \varphi, & b &= \langle \vec{i}', \vec{j} \rangle = \sin \varphi, \\ c &= \langle \vec{j}', \vec{i} \rangle = -\sin \varphi, & d &= \langle \vec{j}', \vec{j} \rangle = \cos \varphi. \end{aligned}$$

Se ora consideriamo $\vec{v} = x'\vec{i}' + y'\vec{j}' = x\vec{i} + y\vec{j}$, sostituendo le espressioni ottenute sopra di \vec{i}' e \vec{j}' in funzione di \vec{i} e \vec{j} , tenendo conto che (\vec{i}, \vec{j}) è una base di $V_2(O)$, si ottiene

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Concludiamo che le matrici ortogonali speciali di ordine due corrispondono alle rotazioni nel piano. Per questo spesso indichiamo con R_φ la matrice

$$\begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix}.$$

Nella figura 20.1 qui sotto vediamo un sistema $Ox'y'$ ottenuto da Oxy con una rotazione.

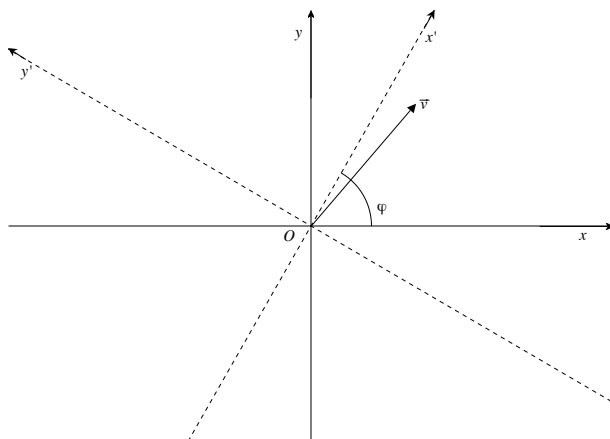


Figura 20.1

Un'analoga interpretazione può essere data per matrici ortogonali in $\mathbb{R}^{n,n}$ con $n \geq 3$. ♠

20.4 Diagonalizzazione ortogonale per matrici simmetriche

La Proposizione 19.10 afferma che ogni matrice simmetrica reale è diagonalizzabile su \mathbb{R} : il risultato principale di questa sezione è che la matrice che diagonalizza può essere scelta ortogonale, cioè tale che $P^{-1} = {}^tP$.

Sia data $A \in \mathbb{R}^{n,n}$ simmetrica e siano $\lambda_1, \dots, \lambda_h \in \mathbb{R}$ i suoi autovalori a due a due distinti. In ciascuno degli autospazi $E_A(\lambda_j)$ fissiamo una base $\mathcal{B}_j = (P_{j,1}, \dots, P_{j,m_j})$, dove $m_j = m_a(\lambda_j, A) = m_g(\lambda_j, A)$. Per la Proposizione 20.15, \mathcal{B}_j può essere supposta ortonormale rispetto al prodotto scalare standard di \mathbb{R}^n .

Osserviamo che autovettori di A relativi ad autovalori distinti sono fra loro ortogonali. Infatti, se $P_i \in E_A(\lambda_i)$ e $P_j \in E_A(\lambda_j)$ con $\lambda_i \neq \lambda_j$, allora

$$\lambda_j {}^t P_i P_j = {}^t P_i (\lambda_j P_j) = {}^t P_i (A P_j) = {}^t P_i A P_j = {}^t P_i {}^t A P_j = {}^t (A P_i) P_j = \lambda_i {}^t P_i P_j,$$

quindi $(\lambda_j - \lambda_i) {}^t P_i P_j = 0$. Poiché $\lambda_j \neq \lambda_i$, segue $\lambda_j - \lambda_i \neq 0$, dunque deve valere ${}^t P_i P_j = 0$.

Queste considerazioni ci permettono di affermare che è sempre possibile trovare una base ortonormale di autovettori di A , ovvero una matrice ortogonale $P \in \mathbb{R}^{n,n}$ tale che $P^{-1}AP$ sia diagonale.

Proposizione 20.24.

Sia $A \in \mathbb{R}^{n,n}$ simmetrica. Allora esiste $P \in \mathbb{R}^{n,n}$ ortogonale tale che $P^{-1}AP = {}^t P A P$ sia diagonale. Inoltre P può essere scelta speciale.

Illustriamo quanto detto con un esempio.

Esempio 20.25. Si consideri la matrice

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

Come già visto nell'Esempio 19.11, gli autovalori di A sono -1 e 2 , con autospazi

$$E_A(-1) = \mathcal{L}((1, -1, 0), (1, 0, -1)), \quad E_A(2) = \mathcal{L}((1, 1, 1)).$$

Per ottenere una matrice ortogonale P che diagonalizzi A è necessario determinare tre autovettori P_1, P_2, P_3 ortonormali. A tale scopo basta determinare basi ortonormali in $E_A(-1)$ e $E_A(2)$. Per quanto riguarda $E_A(2)$, è sufficiente scegliere in esso un versore, ad esempio consideriamo

$$P_1 = \frac{1}{\sqrt{3}}(1, 1, 1).$$

Per l'autospazio $E_A(-1)$ utilizziamo il metodo di Gram-Schmidt a partire dalla base $((1, -1, 0), (1, 0, -1))$. Poiché $\langle (1, -1, 0), (1, -1, 0) \rangle = 2$, scegliamo

$$P_2 = \frac{1}{\sqrt{2}}(1, -1, 0).$$

Per determinare P_3 , tenendo conto che $\langle (1, -1, 0), (1, 0, -1) \rangle = 1$, consideriamo invece un versore associato al vettore

$$(1, 0, -1) - \frac{1}{2}(1, -1, 0) = \frac{1}{2}(1, 1, -2),$$

cioè

$$P_3 = \frac{1}{\sqrt{6}}(-1, -1, 2).$$

Quindi la matrice ortogonale cercata è

$$P = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}.$$

P è non speciale e risulta

$$P^{-1}AP = {}^tPAP = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Per esercizio si verifichi invece che

$$P' = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}$$

è speciale e che ancora risulta

$$P'^{-1}AP' = {}^tP'AP' = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$



Forme quadratiche

21.1 Forme quadratiche reali

Nella Lezione 20 abbiamo visto due esempi di prodotto scalare $\langle \cdot, \cdot \rangle$ su \mathbb{R}^n e in entrambi i casi abbiamo verificato l'esistenza di una matrice $A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathbb{R}^{n,n}$ simmetrica tale che

$$\langle x, y \rangle = {}^t x A y \quad (21.1.1)$$

per ogni $x, y \in \mathbb{R}^n$.

Viceversa, data una matrice simmetrica $A \in \mathbb{R}^{n,n}$, si consideri l'applicazione

$$\langle \cdot, \cdot \rangle: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$$

definita dalla formula (21.1.1): per le proprietà dei prodotti di matrici le proprietà (PS2) e (PS3) della Definizione 20.1 sono banalmente soddisfatte. Per quanto riguarda (PS1) si noti che, poiché ${}^t x A y \in \mathbb{R}$ e A è simmetrica, segue

$$\langle x, y \rangle = {}^t x A y = {}^t ({}^t x A y) = {}^t y {}^t A x = {}^t y A x = \langle y, x \rangle.$$

L'unica condizione di non immediata verifica è (PS4), cioè che vale ${}^t x A x > 0$ per ogni $x \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$. Si noti che ${}^t x A x$ è un polinomio di grado 2 omogeneo nelle entrate di $x = (x_1, \dots, x_n)$ della forma

$$\begin{aligned} & (x_1 \quad x_2 \quad \dots \quad x_n) \begin{pmatrix} a_{1,1}x_1 + \dots + a_{1,n}x_n \\ a_{2,1}x_1 + \dots + a_{2,n}x_n \\ \vdots \\ a_{n,1}x_1 + \dots + a_{n,n}x_n \end{pmatrix} = \\ & = a_{1,1}x_1^2 + \dots + a_{1,n}x_1x_n + a_{2,1}x_2x_1 + \dots + a_{2,n}x_2x_n + \dots + a_{n,n}x_n^2. \end{aligned}$$

Poiché $a_{i,j} = a_{j,i}$ e $x_i x_j = x_j x_i$ per $i, j = 1, \dots, n$, segue che

$${}^t x A x = a_{1,1}x_1^2 + 2a_{1,2}x_1x_2 + \dots + 2a_{1,n}x_1x_n + a_{2,2}x_2^2 + 2a_{2,3}x_2x_3 + \dots + a_{n,n}x_n^2.$$

Quindi determinare se l'applicazione $(x, y) \mapsto {}^t x A y$ è un prodotto scalare equivale a studiare il segno del polinomio $q(x) = {}^t x A x$ al variare di $x \in \mathbb{R}^n$.

Definizione 21.1 (Forme quadratiche reali).

Una *forma quadratica reale* q nelle variabili x_1, \dots, x_n è un polinomio

$$q(x_1, \dots, x_n) \in \mathbb{R}[x_1, \dots, x_n]$$

omogeneo di grado due.

Possiamo riformulare quanto abbiamo visto sopra dicendo che ogni matrice simmetrica $A \in \mathbb{R}^{n,n}$ definisce una forma quadratica reale $q(x) = {}^t x A x$ in n variabili. Vale anche il viceversa: infatti, data la forma quadratica

$$q(x_1, \dots, x_n) = q_{1,1}x_1^2 + q_{1,2}x_1x_2 + \dots + q_{1,n}x_1x_n + q_{2,2}x_2^2 + q_{2,3}x_2x_3 + \dots + q_{n,n}x_n^2,$$

consideriamo la matrice $A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathbb{R}^{n,n}$ definita da

$$a_{i,j} = \begin{cases} q_{i,i} & \text{se } i = j, \\ q_{i,j}/2 & \text{se } i < j, \\ q_{j,i}/2 & \text{se } i > j. \end{cases} \quad (21.1.2)$$

Ovviamente A è simmetrica ed è immediato verificare che $x A^t x = q(x)$: inoltre A è l'unica matrice simmetrica con tale proprietà.

Definizione 21.2 (Matrice associata a una forma quadratica reale).

Sia q una forma quadratica reale nelle variabili x_1, \dots, x_n . La matrice simmetrica $A \in \mathbb{R}^{n,n}$ definita dalla formula (21.1.2) viene detta la *matrice associata a q* .

Esempio 21.3. Si consideri la forma quadratica

$$q(t, x, y, z) = 2t^2 - y^2 + 3tx + tz - xy + 2xz + 9yz.$$

La matrice associata a q è

$$A = \begin{pmatrix} 2 & 3/2 & 0 & 1/2 \\ 3/2 & 0 & -1/2 & 1 \\ 0 & -1/2 & -1 & 9/2 \\ 1/2 & 1 & 9/2 & 0 \end{pmatrix}. \quad \spadesuit$$

21.2 Forme quadratiche definite

Definizione 21.4 (Forme quadratiche definite).

Sia q una forma quadratica reale in n variabili e sia $A \in \mathbb{R}^{n,n}$ la matrice simmetrica associata. Allora:

(DP) q e A si dicono *definite positive* se $q(x) = {}^t x A x > 0$ per ogni $x \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$;

(SDP) q e A si dicono *semidefinite positive* se $q(x) = {}^t x A x \geq 0$ per ogni $x \in \mathbb{R}^n$;

(DN) q e A si dicono *definite negative* se $q(x) = {}^t x A x < 0$ per ogni $x \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}$;

(SDN) q e A si dicono *semidefinite negative* se $q(x) = {}^t x A x \leq 0$ per ogni $x \in \mathbb{R}^n$;

(ID) q e A si dicono *indefinite* nei rimanenti casi, cioè se esistono $x', x'' \in \mathbb{R}^n$ tali che $q(x') = {}^t x' A x' > 0$ e $q(x'') = {}^t x'' A x'' < 0$.

È chiaro che la forma quadratica q e la sua matrice A sono definite o semidefinite negative se e solo se $-q$ e la sua matrice $-A$ sono definite o semidefinite positive rispettivamente.

Osserviamo che alla luce della Definizione 21.4 il problema di stabilire se l'applicazione $(x, y) \mapsto {}^t x A y$ è un prodotto scalare è equivalente a quello di stabilire se A è definita positiva o meno.

Esempio 21.5. In $\mathbb{R}^{2,2}$ si considerino le matrici

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad E_{1,1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad -E_{2,2} = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}, \quad E_{1,1} - E_{2,2} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Si ha che I_2 è definita positiva, $E_{1,1}$ è semidefinita positiva, $-E_{1,1}$ è semidefinita negativa, $E_{1,1} - E_{2,2}$ è indefinita. L'affermazione per I_2 è ovvia. Per quanto riguarda $E_{1,1}$ si ha che $q(x) = x E_{1,1} {}^t x = x_1^2$, dunque $q(0, 1) = 0$: per $-E_{2,2}$ si procede in maniera analoga. Infine $q(x) = x(E_{1,1} - E_{2,2}) {}^t x = x_1^2 - x_2^2$, dunque $q(1, 0) > 0$ e $q(0, 1) < 0$. ♠

Più in generale si ha il seguente ovvio risultato.

Proposizione 21.6.

Sia data la matrice diagonale $D \in \mathbb{R}^{n,n}$ con entrate diagonali $\lambda_1, \dots, \lambda_n$. Allora:

- (i) D è definita positiva se e solo se $\lambda_i > 0$ per ogni $i = 1, \dots, n$;
- (ii) D è semidefinita positiva se e solo se $\lambda_i \geq 0$ per ogni $i = 1, \dots, n$;
- (iii) D è definita negativa se e solo se $\lambda_i < 0$ per ogni $i = 1, \dots, n$;
- (iv) D è semidefinita negativa se e solo se $\lambda_i \leq 0$ per ogni $i = 1, \dots, n$;
- (v) D è indefinita se e solo se esistono $i, j = 1, \dots, n$ tali che $\lambda_i > 0$ e $\lambda_j < 0$.

⚠ Per una matrice simmetrica qualsiasi non è possibile determinare il segno a partire dal segno dei suoi coefficienti come nel caso delle matrici diagonali. Per esempio, vedremo in seguito che la matrice

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

è definita positiva, mentre

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

è indefinita.

Per determinare il segno di una matrice simmetrica A procediamo con *cambiamenti di variabili*. Più in dettaglio, sia $P \in \mathbb{R}^{n,n}$ una matrice invertibile e si consideri la trasformazione in \mathbb{R}^n definita da

$$x = P\bar{x}$$

o, equivalentemente,

$$\bar{x} = P^{-1}x.$$

Definiamo la nuova forma quadratica $\bar{q}(\bar{x}) = q(P\bar{x})$ nella variabile \bar{x} .

Il segno delle forme quadratiche \bar{q} e q è lo stesso: infatti $\bar{x} \in \mathbb{R}^n$ soddisfa $\bar{q}(\bar{x}) > 0$ se e solo se $x = P\bar{x}$ soddisfa $q(x) > 0$, poiché per definizione

$$\bar{q}(\bar{x}) = q(P\bar{x}) = q(x).$$

Un ragionamento simile si può fare per dimostrare che $q(x) = 0$ (rispettivamente $q(x) < 0$) se e solo se $\bar{q}(\bar{x}) = 0$ (rispettivamente $\bar{q}(\bar{x}) < 0$).

Studiamo l'azione di un tale cambiamento di variabili sulla matrice A di q . Dal momento che $x = P\bar{x}$ e, quindi, ${}^t x = {}^t \bar{x} {}^t P$, si ha

$${}^t x A x = {}^t (P\bar{x}) A (P\bar{x}) = {}^t \bar{x} {}^t P A P \bar{x} = {}^t \bar{x} ({}^t P A P) \bar{x}.$$

La matrice ${}^t P A P$ è simmetrica: infatti

$${}^t ({}^t P A P) = {}^t P {}^t A P = {}^t P A P,$$

in forza della simmetria di A . Concludiamo, quindi, che la matrice di \bar{q} è $\bar{A} = {}^t P A P$.

Definizione 21.7 (Matrici congruenti).

Siano $A, \bar{A} \in \mathbb{R}^{n,n}$ simmetriche. Allora A e \bar{A} si dicono *congruenti*, e si scrive $A \approx \bar{A}$, se esiste $P \in \mathbb{R}^{n,n}$ invertibile tale che $\bar{A} = {}^t P A P$.

Esempio 21.8. Si consideri la forma quadratica

$$q(x, y, z) = 2xy + 2xz + 2yz.$$

La sua matrice associata è

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

Per capire se q e A sono o meno definite, si consideri la matrice

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

È facile verificare che P è invertibile: inoltre

$$\bar{A} = {}^tPAP = \begin{pmatrix} 1 & 1 & 0 \\ -1/2 & 1/2 & 0 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1/2 & -1 \\ 1 & 1/2 & -1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

Quindi q e A non sono definite perché, per la Proposizione 21.6, \bar{A} non è definita. ♠

L'esempio precedente mostra quanto sia semplice stabilire il segno di una matrice simmetrica $A \in \mathbb{R}^{n,n}$ se si è in grado di determinare una matrice $P \in \mathbb{R}^{n,n}$ invertibile tale che tPAP sia diagonale. Sono vari gli algoritmi che si possono utilizzare a tale scopo.

Per esempio si osservi che per ogni matrice elementare $E \in \mathbb{R}^{n,n}$, le matrici A e tEAE sono congruenti. Si noti che tale seconda matrice è ottenuta da A facendo la stessa operazione elementare sulle sue righe (tramite la moltiplicazione a sinistra per tE) e sulle sue colonne (tramite la moltiplicazione a destra per E).

Questo suggerisce un metodo abbastanza veloce per studiare il segno di una matrice simmetrica per riduzione simultanea per righe e colonne che illustriamo nei due seguenti esempi.

Esempio 21.9. Si consideri la forma quadratica

$$q(x, y, z) = x^2 - 2y^2 + 2z^2 + 3xy + 2xz + 2yz.$$

La sua matrice associata è

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

Per calcolare il segno di q e A procediamo come segue con operazioni elementari di riga e colonna.

$$A \xrightarrow[\substack{R_2 \rightarrow R_2 - 3R_1/2 \\ R_3 \rightarrow R_3 - R_1}]{\substack{R_2 \rightarrow R_2 - 3R_1/2 \\ R_3 \rightarrow R_3 - R_1}} \begin{pmatrix} 1 & 3/2 & 1 \\ 0 & -17/4 & -1/2 \\ 0 & -1/2 & 1 \end{pmatrix} \xrightarrow[\substack{C_2 \rightarrow C_2 - 3C_1/2 \\ C_3 \rightarrow C_3 - C_1}]{\substack{C_2 \rightarrow C_2 - 3C_1/2 \\ C_3 \rightarrow C_3 - C_1}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -17/4 & -1/2 \\ 0 & -1/2 & 1 \end{pmatrix}$$

$$\xrightarrow[\substack{R_3 \rightarrow R_3 - 2R_2/17}]{\substack{R_3 \rightarrow R_3 - 2R_2/17}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -17/4 & -1/2 \\ 0 & 0 & 18/17 \end{pmatrix} \xrightarrow[\substack{C_3 \rightarrow C_3 - 2C_2/17}]{\substack{C_3 \rightarrow C_3 - 2C_2/17}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -17/4 & 0 \\ 0 & 0 & 18/17 \end{pmatrix}$$

Deduciamo che q ed A non sono definite. Si noti che

$$E_{R_2 \rightarrow R_2 - 3R_1/2} = \begin{pmatrix} 1 & 0 & 0 \\ -3/2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_{R_3 \rightarrow R_3 - R_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix},$$

$$E_{R_3 \rightarrow R_3 - 2R_2/17} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2/17 & 1 \end{pmatrix}.$$

Allora, posto

$$P = E_{C_3 \rightarrow C_3 - 2C_2/17} E_{C_3 \rightarrow C_3 - C_1} E_{C_2 \rightarrow C_2 - 3C_1/2} = \begin{pmatrix} 1 & -3/2 & -14/17 \\ 0 & 1 & -2/17 \\ 0 & 0 & 1 \end{pmatrix},$$

$${}^tP = E_{R_3 \rightarrow R_3 - 2R_2/17} E_{R_3 \rightarrow R_3 - R_1} E_{R_2 \rightarrow R_2 - 3R_1/2} = \begin{pmatrix} 1 & 0 & 0 \\ -3/2 & 1 & 0 \\ -14/17 & -2/17 & 1 \end{pmatrix},$$

risulta

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -17/4 & 0 \\ 0 & 0 & 18/17 \end{pmatrix}. \quad \spadesuit$$

Esempio 21.10. Riprendiamo la forma quadratica q e la matrice A dell'Esempio 21.8, procedendo come segue con operazioni elementari di riga e colonna.

$$A \xrightarrow{R_1 \rightarrow R_1 + R_2} \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \xrightarrow{C_1 \rightarrow C_1 + C_2} \begin{pmatrix} 2 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}$$

$$\xrightarrow{\substack{R_2 \rightarrow R_2 - R_1/2 \\ R_3 \rightarrow R_3 - R_1}} \begin{pmatrix} 2 & 1 & 2 \\ 0 & -1/2 & 0 \\ 0 & 0 & -2 \end{pmatrix} \xrightarrow{\substack{C_2 \rightarrow C_2 - C_1/2 \\ C_3 \rightarrow C_3 - C_1}} \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

che coincide con quanto già visto. In particolare il lettore verifichi che

$$P = E_{C_1 \rightarrow C_1 + C_2} E_{C_2 \rightarrow C_2 - C_1/2} E_{C_3 \rightarrow C_3 - C_1}, \quad {}^tP = E_{R_3 \rightarrow R_3 - R_1} E_{R_2 \rightarrow R_2 - R_1/2} E_{R_1 \rightarrow R_1 + R_2}$$

nell'Esempio 21.8. ♠

Un altro metodo per il calcolo del segno di una forma quadratica (e della corrispondente matrice simmetrica) si basa sulla Proposizione 20.24.

Infatti sappiamo che ogni matrice simmetrica A a coefficienti reali è diagonalizzabile mediante una matrice ortogonale P : poiché, $P^{-1} = {}^tP$, segue che A e $\overline{A} = P^{-1}AP = {}^tPAP$ sono congruenti. In particolare A è congruente a una matrice diagonale avente sulla diagonale gli autovalori di A ciascuno ripetuto con la sua molteplicità algebrica. Quindi il problema della determinazione del segno di A si può ridurre alla determinazione del segno dei suoi autovalori.

Esempio 21.11. Consideriamo la matrice

$$A_1 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Il suo polinomio caratteristico è

$$p_{A_1}(t) = \begin{vmatrix} 2-t & 1 \\ 1 & 2-t \end{vmatrix} = t^2 - 4t + 3 = (t-3)(t-1),$$

quindi A_1 ha autovalori 1 e 3: in particolare, in base a quanto osservato sopra, A_1 è definita positiva.

Se invece consideriamo la matrice

$$A_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix},$$

essa ha polinomio caratteristico

$$p_{A_2}(t) = \begin{vmatrix} 1-t & 2 \\ 2 & 1-t \end{vmatrix} = t^2 - 2t - 3 = (t-3)(t+1),$$

quindi autovalori -1 e 3 . In particolare, in base a quanto osservato sopra, A_2 è indefinita. ♠

Abbiamo avuto modo di osservare che il calcolo degli autovalori di una matrice può essere particolarmente complicato poiché non vi sono modi semplici per calcolare le radici di un polinomio generale di grado n . In realtà, però, per determinare se una matrice simmetrica a coefficienti reali è definita o meno a noi non interessa il valore numerico degli autovalori ma solo il loro segno e questo può essere facilmente determinato utilizzando la cosiddetta *regola dei segni di Cartesio*.

Per illustrare tale metodo dobbiamo introdurre la nozione di variazione di una successione di numeri reali.

Definizione 21.12.

Sia $(a_0, a_1, \dots, a_{n-1}, a_n)$ una successione di numeri reali. Il *numero di variazioni* $v(a_0, a_1, \dots, a_{n-1}, a_n)$ è il numero delle coppie di interi $(i, i+k)$ con $k \geq 1$ tali che

1. $a_i a_{i+k} < 0$;
2. $a_{i+h} = 0$ per $h = 1, \dots, k-1$.

Per ogni polinomio a coefficienti reali $p(t) = a_0 t^n + a_1 t^{n-1} + \dots + a_{n-1} t + a_n \in \mathbb{R}[t]$ poniamo $v(p(t)) = v(a_0, a_1, \dots, a_{n-1}, a_n)$.

Esempio 21.13. Per esempio, $v(-t^6 + t^5 - t^2 + t + 1) = v(-1, 1, 0, 0, -1, 1, 1) = 3$. ♠

Siamo pronti per enunciare la regola dei segni di Cartesio, di cui omettiamo la dimostrazione.

Proposizione 21.14 (Regola dei segni di Cartesio).

Se tutte le radici di un polinomio $p(t) = a_0 t^n + a_1 t^{n-1} + \dots + a_{n-1} t + a_n \in \mathbb{R}[t]$ sono reali, allora $v(p(t))$ è il numero delle sue radici positive contate con la loro molteplicità.

Si osservi che tale proposizione non ci permette di dire nulla sul segno delle radici del polinomio dell'Esempio 21.13, perchè non sappiamo a priori che le sue radici sono tutte reali. Se però consideriamo il polinomio caratteristico di una matrice simmetrica a coefficienti reali, sappiamo che esso ha tutte le radici in \mathbb{R} , grazie alla Proposizione 19.10.

Esempio 21.15. Riprendiamo la matrice A dell'Esempio 21.8

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

Il suo polinomio caratteristico è $p_A(t) = -t^3 + 3t + 2$, dunque il numero di radici positive di $p_A(t)$ è $v(-1, 0, 3, 2) = 1$: poiché $p_A(t)$ non ha la radice nulla e ha tre radici reali (perché è polinomio caratteristico di una matrice simmetrica che è diagonalizzabile!), segue che $p_A(t)$ ha $3 - 1 = 2$ radici negative: infatti gli autovalori di A sono 2 e -1 con $m_a(2, A) = 1$ e $m_a(-1, A) = 2$ (si veda l'Esempio 19.11). ♠

Trasformazioni nel piano

22.1 Ellisse, iperbole, parabola

La parte finale del corso riguarda l'applicazione della teoria delle forme quadratiche e della riduzione ortogonale allo studio di alcuni oggetti geometrici detti "coniche" e "quadriche".

Cominciamo con un veloce ripasso di alcune figure geometriche che sicuramente il lettore ha già incontrato nei suoi studi: l'iperbole, l'ellisse, la parabola (che, nel seguito, chiameremo *coniche classiche*).

Nei seguenti esempi supporremo fissato un sistema di riferimento Oxy nel piano: osserviamo che il sistema è indicato con Oxy invece di $O\vec{i}\vec{j}$ semplicemente perchè in questi esempi è preferibile mettere l'accento sul nome delle coordinate di un punto piuttosto che sui versori.

Esempio 22.1 (Iperbole). Siano $a, b \in]0, +\infty[$ due numeri reali positivi. Si consideri

$$\mathcal{C} = \left\{ P = (x, y) \mid \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \right\}.$$

\mathcal{C} si dice *iperbole di semiassi a e b (in forma canonica)*. L'equazione

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \tag{22.1.1}$$

si dice *equazione canonica di \mathcal{C}* .

L'iperbole \mathcal{C} è simmetrica sia rispetto a O sia rispetto agli assi coordinati: l'origine e gli assi coordinati vengono detti *centro* (in questo caso si parla di *conica a centro*) e *assi* di \mathcal{C} . L'asse delle ascisse viene detto *asse trasverso* o *asse reale*, quello delle ordinate *asse secondario* o *asse immaginario*. I punti di intersezione di \mathcal{C} con l'asse trasverso sono detti *vertici*: sono, nel nostro caso, i punti $(\pm a, 0)$. Talvolta le coppie $(0, \pm ib)$ sono detti *vertici immaginari*.

Determiniamo i punti di intersezione di \mathcal{C} con le rette parallele all'asse delle ordinate: ciò equivale a determinare le soluzioni dei sistemi

$$\begin{cases} \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \\ x = k \end{cases}$$

al variare di $k \in \mathbb{R}$: quindi ci riduciamo a risolvere l'equazione

$$\frac{y^2}{b^2} = \frac{k^2}{a^2} - 1.$$

Tale equazione ha sempre soluzioni (reali o complesse): ha però soluzioni reali (corrispondenti quindi a punti del piano) se e solo se $|k| \geq a$, ovvero se e solo se \mathcal{C} è all'esterno della fascia $] -k, k[\times \mathbb{R}$. In particolare \mathcal{C} ha, necessariamente, due componenti distinte che non si intersecano: \mathcal{C} è, dunque, sconnessa.

Intersechiamo \mathcal{C} con le rette per l'origine: ciò equivale a determinare le soluzioni del sistema

$$\begin{cases} \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \\ y = mx \end{cases}$$

al variare di $m \in \mathbb{R}$. Quindi ci riduciamo a risolvere l'equazione

$$(b^2 - m^2 a^2)x^2 = a^2 b^2.$$

Per $m = \pm b/a$ tale equazione non ha soluzioni (reali o complesse): le rette $y = \pm bx/a$ si dicono *asintoti* di \mathcal{C} . Per ogni altro valore di m l'equazione ha soluzioni (reali o complesse): tali soluzioni sono reali (corrispondenti quindi a punti del piano) se e solo se $|m| < b/a$, ovvero \mathcal{C} è all'interno del cono delimitato dagli asintoti e contenente l'asse delle ascisse.

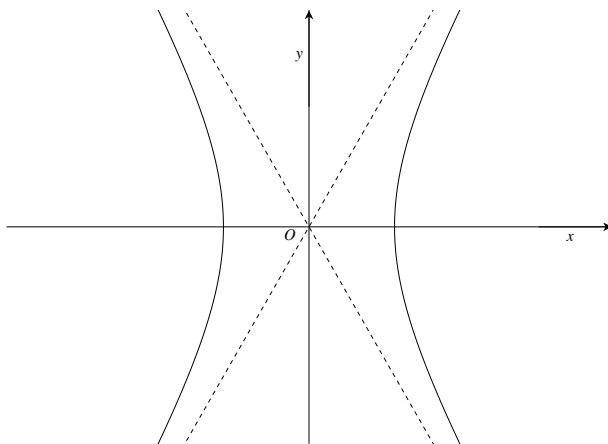


Figura 22.1

Si noti che gli asintoti di \mathcal{C} sono perpendicolari fra loro se e solo se $a = b$: in tal caso si parla di *iperbole equilatera*.

Ricordiamo le due funzioni iperboliche

$$\cosh x = (e^x + e^{-x})/2, \quad \sinh x = (e^x - e^{-x})/2.$$

Si noti che $\cosh^2 x - \sinh^2 x = 1$: quindi i punti delle due componenti connesse di \mathcal{C} possono essere descritti parametricamente dai due sistemi

$$\begin{cases} x = a \cosh t \\ y = b \sinh t \end{cases} \quad \begin{cases} x = -a \cosh t \\ y = b \sinh t, \end{cases} \quad t \in \mathbb{R}.$$

La figura 22.1 rappresenta la conica \mathcal{C} di equazione

$$\frac{x^2}{2} - \frac{y^2}{4} = 1$$

(con i suoi asintoti). ♠

Esempio 22.2 (Ellisse). Siano $a, b \in]0, +\infty[$ due numeri reali positivi. Si consideri

$$\mathcal{C} = \left\{ P = (x, y) \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \right\}.$$

\mathcal{C} si dice *ellisse di semiassi a e b (in forma canonica)*: talvolta si parla anche di *ellisse reale* per distinguerla dall'ellisse immaginaria che sarà definita in seguito. L'equazione

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \tag{22.1.2}$$

si dice *equazione canonica di \mathcal{C}* .

Come nel caso dell'iperbole, \mathcal{C} è simmetrica sia rispetto a O sia rispetto agli assi coordinati: l'origine e gli assi coordinati vengono detti *centro* (anche \mathcal{C} è una *conica a centro*) e *assi* di \mathcal{C} . I punti di intersezione di \mathcal{C} con i suoi assi vengono detti *vertici*: sono, nel nostro caso, $(\pm a, 0)$ e $(0, \pm b)$.

Intersecando con rette parallele agli assi e studiando le soluzioni delle equazioni di secondo grado che così si ottengono, è facile verificare che $\mathcal{C} \subseteq [-a, a] \times [-b, b]$.

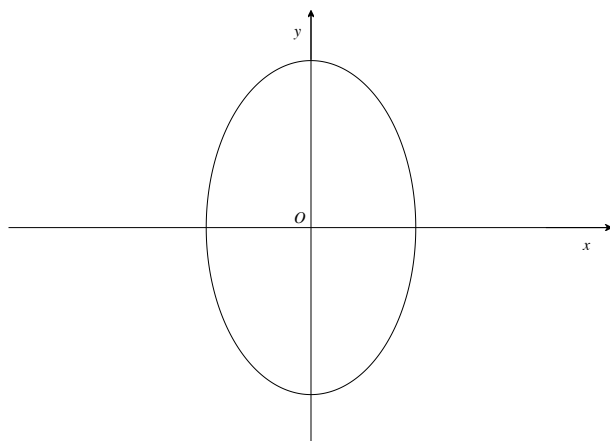


Figura 22.2

Si noti che più a si avvicina a b , più \mathcal{C} tende ad assomigliare a una circonferenza. Inoltre è facile convincersi che i suoi punti possono essere descritti parametricamente dal sistema

$$\begin{cases} x = a \cos t \\ y = b \sin t \end{cases} \quad t \in \mathbb{R}.$$

La figura 22.2 rappresenta la conica \mathcal{C} di equazione

$$\frac{x^2}{3} + \frac{y^2}{5} = 1. \quad \spadesuit$$

Esempio 22.3 (Parabola). Sia $p \in]0, +\infty[$ un numero reale positivo. Si consideri

$$\mathcal{C} = \{ P = (x, y) \mid y^2 = 2px \}.$$

\mathcal{C} si dice *parabola di parametro p (in forma canonica)*. L'equazione

$$y^2 = 2px$$

si dice *equazione canonica di \mathcal{C}* .

La parabola \mathcal{C} è simmetrica rispetto all'asse delle ascisse: tale retta viene detta *asse*. Invece non esistono punti nel piano rispetto a cui \mathcal{C} sia simmetrica (perciò \mathcal{C} non è una conica a centro). Il punto di intersezione di \mathcal{C} con l'asse delle ascisse viene detto *vertice*: esso è, nel nostro caso, O .

Intersecando con rette parallele agli assi e studiando le soluzioni delle equazioni di secondo grado che così si ottengono, è facile verificare che $\mathcal{C} \subseteq [0, +\infty[\times \mathbb{R}$.

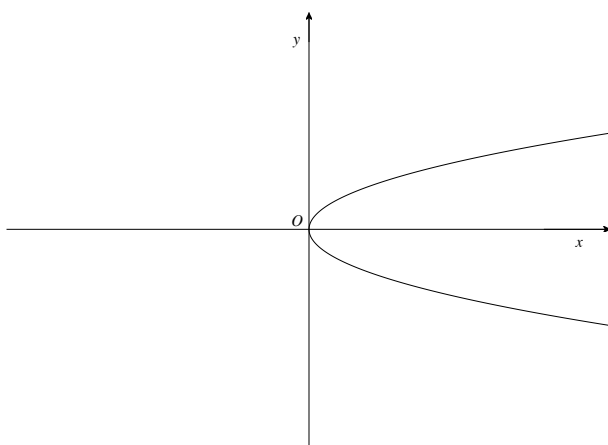


Figura 22.3

I punti di \mathcal{C} possono essere descritti parametricamente dal sistema

$$\begin{cases} x = t^2/2p \\ y = t, \end{cases} \quad t \in \mathbb{R}.$$

La figura 22.3 rappresenta la parabola \mathcal{C} di equazione

$$y^2 = x/3. \quad \spadesuit$$

Notiamo che tutte le coniche classiche sopra descritte sono luoghi geometrici definiti, rispetto a un opportuno sistema di riferimento Oxy fissato nel piano, da equazioni di grado 2 nelle variabili x, y : se due di tali equazioni differiscono per una costante moltiplicativa non nulla, esse definiscono la stessa conica classica.

In maniera analoga abbiamo visto che, rispetto a un opportuno sistema di riferimento Oxy fissato nel piano, ogni retta nel piano è individuata da un'equazione di grado 1 nelle variabili x, y (unica a meno di costanti moltiplicative non nulle). Di più abbiamo visto che, viceversa, ogni equazione di grado 1 nelle variabili x, y rappresenta una retta nel piano.

Viene allora naturale domandarsi se una tale proprietà vale anche per le equazioni di grado 2: cioè è vero o falso che ogni equazione di grado 2 nelle variabili x, y rappresenta un'ellisse, un'iperbole o una parabola? L'esempio seguente mostra che purtroppo ciò non accade.

Esempio 22.4. Si consideri la generica equazione

$$\alpha x^2 + \beta y^2 = \gamma \quad (22.1.3)$$

e si noti che le equazioni (22.1.1) e (22.1.2) hanno questa forma. Chiaramente se $\alpha = \beta = 0$ il luogo geometrico $\{ P = (x, y) \mid \alpha x^2 + \beta y^2 = \gamma \}$ o è vuoto (se $\gamma \neq 0$) o è tutto il piano (se $\gamma = 0$). Assumeremo nel seguito che $\beta \neq 0$: il motivo per cui tale scelta è possibile sarà chiarito nel prossimo paragrafo (si veda, in particolare, l'Esempio 22.5).

Ricordiamo inoltre che l'equazione (22.1.3) va intesa a meno di una costante moltiplicativa non nulla, poiché

$$\{ P = (x, y) \mid \alpha x^2 + \beta y^2 = \gamma \} = \{ P = (x, y) \mid \lambda \alpha x^2 + \lambda \beta y^2 = \lambda \gamma \}$$

se $\lambda \neq 0$.

Analizziamo i vari casi possibili.

- (i) Se $\alpha > 0, \beta < 0, \gamma = 0$ si parla di *coppia iperbolica di rette incidenti*: spesso si parla anche di *iperbole degenera*. Poiché

$$\alpha x^2 + \beta y^2 = (\sqrt{\alpha}x + \sqrt{-\beta}y)(\sqrt{\alpha}x - \sqrt{-\beta}y)$$

(ricordiamo che $\beta < 0$, dunque $\sqrt{-\beta} \in \mathbb{R}$!), in tal caso il luogo geometrico

$$\{ P = (x, y) \mid \alpha x^2 + \beta y^2 = 0 \}$$

è l'unione delle due rette di equazioni $\sqrt{\alpha}x + \sqrt{-\beta}y = 0$ e $\sqrt{\alpha}x - \sqrt{-\beta}y = 0$.

- (ii) Se $\alpha = 0, \beta, \gamma > 0$ si parla di *coppia iperbolica di rette parallele*: spesso si parla anche di *parabola degenera reale*. In tal caso il luogo geometrico

$$\{ P = (x, y) \mid \beta y^2 = \gamma \}$$

è l'unione delle due rette di equazioni $\sqrt{\beta}y = \sqrt{\gamma}$ e $\sqrt{\beta}y = -\sqrt{\gamma}$.

(iii) Se $\alpha = \gamma = 0$ si parla di *retta doppia*: in tal caso il luogo geometrico

$$\{ P = (x, y) \mid \beta y^2 = 0 \}$$

coincide con l'asse delle ascisse (ma contato con molteplicità due).

(iv) Se $\alpha, \beta > 0, \gamma = 0$ si parla di *coppia ellittica di rette incidenti*: spesso si parla anche di *ellisse degenera*. In tal caso il luogo geometrico

$$\{ P = (x, y) \mid \alpha x^2 + \beta y^2 = 0 \}$$

si riduce a un solo punto, l'origine. Si noti che, anche in questo caso il polinomio di grado 2 si scompone in prodotto di polinomi di grado 1: infatti si ha

$$\alpha x^2 + \beta y^2 = (\sqrt{\alpha}x + i\sqrt{\beta}y)(\sqrt{\alpha}x - i\sqrt{\beta}y).$$

La differenza rispetto al caso della coppia iperbolica di rette è che tali polinomi lineari hanno coefficienti complessi non reali.

(v) Se $\alpha = 0, \beta > 0, \gamma < 0$ si parla di *coppia ellittica di rette parallele*: spesso si parla anche di *parabola degenera immaginaria*. Si noti che pur essendo

$$\{ P = (x, y) \mid \beta y^2 = \gamma \} = \emptyset,$$

ancora il polinomio di grado 2 si scompone in prodotto di polinomi di grado 1: infatti si ha

$$\beta y^2 - \gamma = (\sqrt{\beta}y + i\sqrt{-\gamma})(\sqrt{\beta}y - i\sqrt{-\gamma}).$$

Anche in questo caso tali polinomi lineari hanno coefficienti complessi non reali.

(vi) Se $\alpha, \beta > 0, \gamma < 0$ si parla di *ellisse immaginaria*. Anche in questo caso il luogo geometrico

$$\{ P = (x, y) \mid \alpha x^2 + \beta y^2 = \gamma \}$$

è evidentemente vuoto.

Osserviamo che c'è una differenza sostanziale tra il caso dell'ellisse immaginaria e quello della coppia ellittica di rette parallele. Infatti, nel caso dell'ellisse immaginaria, il polinomio $\alpha x^2 + \beta y^2 = \gamma$ non si può scomporre come prodotto di due polinomi lineari, nè a coefficienti reali, nè a coefficienti complessi.

Infatti se fosse

$$\alpha x^2 + \beta y^2 - \gamma = (a'x + b'y + c')(a''x + b''y + c''),$$

dovrebbe essere $a'a'' = \alpha$. Se $\alpha \neq 0$ moltiplicando ambo i membri per l'inverso di $\alpha = a'a'' \in \mathbb{R}$ possiamo supporre $\alpha = a' = a'' = 1$.

Ci riconduciamo quindi a studiare l'uguaglianza

$$x^2 + \beta y^2 - \gamma = (x + b'y + c')(x + b''y + c'') :$$

quindi si deve avere $b'b'' = \beta$, $c'c'' = -\gamma$ e $b' + b'' = c' + c'' = b'c'' + b''c' = 0$, dunque $b'' = -b'$, $c'' = -c'$, $-2b'c' = 0$, quindi o $c' = 0$ (da cui $\gamma = 0$) o $b' = 0$ (da cui $\beta = 0$).

Ragionando in maniera analoga ci si convince che anche le equazioni dell'ellisse, dell'iperbole e della parabola non si possono scomporre come prodotto di due polinomi lineari, nè a coefficienti reali, nè a coefficienti complessi. ♠

22.2 Rototraslazioni nel piano

Dalle considerazioni fatte nel paragrafo precedente, sorge in maniera naturale il problema seguente: fissato nel piano un sistema di riferimento Oxy e dato un polinomio di grado 2 nelle variabili x, y a meno di costanti moltiplicative non nulle, diciamo

$$ax^2 + bxy + cy^2 + dx + ey + f$$

($a, b, c, d, e, f \in \mathbb{R}$, a, b, c non simultaneamente nulli), vogliamo descrivere il luogo (eventualmente vuoto)

$$\mathcal{C} = \{ P = (x, y) \mid ax^2 + bxy + cy^2 + dx + ey + f = 0 \}.$$

In particolare, se possibile, vogliamo disegnare tale luogo geometrico.

Come visto negli esempi del paragrafo 23.1, siamo in grado di dare una tale descrizione se il polinomio è di forma particolarmente semplice: ciò dipende dal fatto che il luogo \mathcal{C} ha delle particolari simmetrie rispetto al sistema di riferimento fissato Oxy .

È allora naturale porsi i seguenti due quesiti.

- (Q1) Sia $O'x'y'$ un sistema di riferimento: come sono legate le coordinate (x, y) e (x', y') di uno stesso punto nei due sistemi di riferimento?
- (Q2) Conoscendo una descrizione in termini di equazioni rispetto al sistema di riferimento Oxy di un certo luogo geometrico \mathcal{C} , esiste un modo per descriverlo con equazioni anche rispetto al sistema di riferimento $O'x'y'$?

Iniziamo a rispondere al secondo quesito. Supponiamo che sia

$$\mathcal{C} = \{ P = (x, y) \mid q(x, y) = 0 \}$$

ove $q(x, y)$ è una qualche espressione matematica funzione delle variabili x e y . Supponiamo poi che ci sia un legame fra le coordinate (x, y) e (x', y') di uno stesso punto P rispetto a i due sistemi di riferimento Oxy e $O'x'y'$ del tipo

$$\begin{cases} x = f(x', y'), \\ y = g(x', y'). \end{cases}$$

Allora $P \in \mathcal{C}$ se e solo se P ha coordinate (\bar{x}, \bar{y}) tali che $q(\bar{x}, \bar{y}) = 0$, cioè se e solo se le coordinate (\bar{x}', \bar{y}') per cui

$$\begin{cases} \bar{x} = f(\bar{x}', \bar{y}'), \\ \bar{y} = g(\bar{x}', \bar{y}') \end{cases}$$

soddisfano $q(f(\bar{x}', \bar{y}'), g(\bar{x}', \bar{y}')) = 0$. Concludiamo che è possibile descrivere \mathcal{C} anche nel sistema di riferimento $O'x'y'$ in termini di equazioni: precisamente abbiamo

$$\mathcal{C} = \{ P = (x', y') \mid q(f(x', y'), g(x', y')) = 0 \}.$$

Esempio 22.5. Sia Oxy un sistema di riferimento fissato nel piano e sia

$$\mathcal{C} = \{ P = (x, y) \mid x^2 - 2y^2 = -1 \}.$$

Si consideri ora un sistema di riferimento $Ox'y'$ ottenuto da Oxy ruotando di un angolo $\pi/2$ in senso antiorario. Allora $x' = y$ e $y' = -x$: quindi

$$\mathcal{C} = \{ P = (x', y') \mid 2x'^2 - y'^2 = 1 \}.$$

Concludiamo che \mathcal{C} è un'iperbole e possiamo anche tracciarne il disegno utilizzando quanto visto nell'Esempio 22.1.

Con un ragionamento simile deduciamo che è sempre possibile supporre $\beta \neq 0$ nell'Esempio 22.4. ♠

Passiamo ora al primo quesito. Supponiamo che il semiasse positivo delle x' formi un angolo φ (misurato in senso antiorario) con il semiasse positivo delle x e che le coordinate di O' rispetto al sistema di riferimento Oxy siano (u, v) .

Per passare dal Oxy a $O'x'y'$ si può procedere in due modi. Possiamo ruotare in senso antiorario il sistema di riferimento Oxy di un angolo φ ottenendo un sistema di riferimento "ausiliario" $O\hat{x}\hat{y}$ e poi traslare quest'ultimo in modo da portare la sua origine in O' . Oppure possiamo prima traslare Oxy portando la sua origine a coincidere con O' e ottenendo un sistema di riferimento "ausiliario" $O'\hat{x}\hat{y}$ e poi ruotare quest'ultimo di un angolo φ in senso antiorario.

In entrambi i casi abbiamo decomposto la nostra trasformazione in due trasformazioni più elementari, una *rotazione* e una *traslazione*: per questo motivo parleremo di *rototraslazione* quando ci riferiremo a trasformazioni di coordinate nel piano.

Iniziamo a ricordare cosa accade nel caso in cui non ci sia traslazione, cioè $O = O'$. Siano \vec{i}, \vec{j} e \vec{I}, \vec{J} i versori degli assi coordinati naturalmente associati a i due sistemi di riferimento Oxy e $O'x'y'$ rispettivamente. Se P è un punto di coordinate (x, y) rispetto a Oxy e (x', y') rispetto a $O'x'y'$ risulta

$$\overrightarrow{OP} = x\vec{i} + y\vec{j} = x'\vec{I} + y'\vec{J}. \quad (22.2.1)$$

Poiché (\vec{i}, \vec{j}) è una base ortonormale di $V_2(O)$ segue che

$$\vec{I} = \langle \vec{I}, \vec{i} \rangle \vec{i} + \langle \vec{I}, \vec{j} \rangle \vec{j}, \quad \vec{J} = \langle \vec{J}, \vec{i} \rangle \vec{i} + \langle \vec{J}, \vec{j} \rangle \vec{j}.$$

Ma, poiché $|\vec{i}| = |\vec{j}| = |\vec{I}| = |\vec{J}| = 1$, segue che

$$\begin{aligned}\langle \vec{I}, \vec{i} \rangle &= \cos \varphi, & \langle \vec{I}, \vec{j} \rangle &= \cos(\pi/2 - \varphi) = \sin \varphi, \\ \langle \vec{J}, \vec{i} \rangle &= \cos(\pi/2 + \varphi) = -\sin \varphi, & \langle \vec{J}, \vec{j} \rangle &= \cos \varphi,\end{aligned}$$

dunque

$$\begin{cases} \vec{I} = \cos \varphi \vec{i} + \sin \varphi \vec{j}, \\ \vec{J} = -\sin \varphi \vec{i} + \cos \varphi \vec{j} \end{cases} \quad (22.2.2)$$

Sostituendo le relazioni (22.2.2) nelle relazioni (22.2.1) si ottiene allora

$$\overrightarrow{OP} = x\vec{i} + y\vec{j} = x' \cos \varphi \vec{i} + x' \sin \varphi \vec{j} - y' \sin \varphi \vec{i} + y' \cos \varphi \vec{j}.$$

Abbiamo perciò scritto lo stesso vettore geometrico come combinazione lineare degli elementi della base (\vec{i}, \vec{j}) in due modi diversi, quindi i coefficienti di \vec{i} e \vec{j} nelle due espressioni devono coincidere ordinatamente: eguagliandoli otteniamo, come già visto ripetutamente in precedenza, le relazioni

$$\begin{cases} x = x' \cos \varphi - y' \sin \varphi \\ y = x' \sin \varphi + y' \cos \varphi, \end{cases}$$

o, in forma matriciale,

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}. \quad (22.2.3)$$

Passiamo a descrivere il caso in cui non ci sia rotazione, cioè $\varphi = 0$. Siano \vec{i}, \vec{j} e \vec{I}, \vec{J} i versori degli assi coordinati naturalmente associati ai due sistemi di riferimento Oxy e $O'x'y'$ rispettivamente. Se P è un punto di coordinate (x, y) rispetto a Oxy e (x', y') rispetto a $O'x'y'$ risulta ancora

$$\overrightarrow{OP} = x\vec{i} + y\vec{j} = \overrightarrow{OO'} + (P - O'). \quad (22.2.4)$$

Se (u, v) sono le coordinate di O' rispetto a Oxy , poiché $\overrightarrow{OP} = x'\vec{I} + y'\vec{J}$, si ha

$$\overrightarrow{OO'} = u\vec{i} + v\vec{j}, \quad P - O' = x'\vec{i} + y'\vec{j},$$

quindi la relazione (22.2.4) diviene

$$\overrightarrow{OP} = x\vec{i} + y\vec{j} = x'\vec{i} + y'\vec{j} + u\vec{i} + v\vec{j}.$$

Abbiamo perciò scritto lo stesso vettore geometrico come combinazione lineare degli elementi della base (\vec{i}, \vec{j}) in due modi diversi, quindi i coefficienti di \vec{i} e \vec{j} nelle due espressioni devono coincidere: eguagliandoli otteniamo

$$\begin{cases} x = x' + u \\ y = y' + v, \end{cases}$$

o, in forma matriciale,

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} u \\ v \end{pmatrix}. \quad (22.2.5)$$

Concludiamo il paragrafo con la seguente risposta al quesito (Q1).

Proposizione 22.6.

Nel piano siano fissati due sistemi di riferimento Oxy e $O'x'y'$. Si assuma che il semiasse positivo delle x' formi un angolo φ (misurato in senso antiorario) con il semiasse positivo delle x e che le coordinate di O' rispetto al sistema di riferimento Oxy siano (u, v) . Le coordinate (x, y) e (x', y') di uno stesso punto P in Oxy e $O'x'y'$ rispettivamente sono legate dalle seguenti relazioni:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} u \\ v \end{pmatrix}.$$

Dimostrazione. Si decomponga la rototraslazione nella traslazione che porta il vecchio sistema di riferimento Oxy in quello ausiliario $O'\hat{x}\hat{y}$ seguita dalla rotazione in senso antiorario dell'angolo φ che porta il sistema di riferimento ausiliario $O'\hat{x}\hat{y}$ in quello nuovo $O'x'y'$. Allora le formule (22.2.3) e (22.2.5) diventano rispettivamente

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} + \begin{pmatrix} u \\ v \end{pmatrix}, \quad \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

Eliminando \hat{x} e \hat{y} tra le due equazioni matriciali otteniamo la tesi. □

Coniche

23.1 Coniche e loro riduzione a forma canonica

Fissiamo nel piano un sistema di riferimento Oxy e consideriamo un polinomio di grado 2 in x, y a meno di costanti moltiplicative non nulle, diciamo

$$ax^2 + bxy + cy^2 + dx + ey + f$$

($a, b, c, d, e, f \in \mathbb{R}$). Nel capitolo precedente ci siamo posti il problema di descrivere il luogo geometrico (eventualmente vuoto)

$$\mathcal{C} = \{ P = (x, y) \mid ax^2 + bxy + cy^2 + dx + ey + f = 0 \}$$

e abbiamo visto che tale luogo può essere un'ellisse, un'iperbole, una parabola (coniche classiche) o qualcos'altro.

Per tale motivo è opportuno introdurre la seguente definizione.

Definizione 23.1 (Coniche).

Una *conica* \mathcal{C} nel piano (rispetto a un fissato sistema di riferimento Oxy) è il dato, a meno di costanti moltiplicative non nulle, di un'equazione della forma

$$ax^2 + bxy + cy^2 + dx + ey + f = 0,$$

con $a, b, c, d, e, f \in \mathbb{R}$ e a, b, c non simultaneamente nulli.

Se \mathcal{C} è la conica $ax^2 + bxy + cy^2 + dx + ey + f = 0$, molto spesso, nel seguito, utilizzeremo la locuzione *la conica \mathcal{C} di equazione $ax^2 + bxy + cy^2 + dx + ey + f = 0$* .

Nel capitolo precedente abbiamo studiato come cambia l'equazione di una conica (nel senso della definizione sopra) quando operiamo nel piano una rototraslazione: la "nuova" equazione si ottiene sostituendo alle "vecchie" variabili la loro espressione in funzione di "nuove" variabili.

Vediamo cosa si può dire nel caso specifico delle coniche. Come prima cosa osserviamo che l'equazione di cui sopra si può sempre scrivere nella forma

$$a_{1,1}x^2 + 2a_{1,2}xy + a_{2,2}y^2 + 2a_{1,3}x + 2a_{2,3}y + a_{3,3} = 0 \quad (23.1.1)$$

per opportuni $a_{i,j} \in \mathbb{R}$ (sarà chiara fra poco la maggiore convenienza di una tale notazione rispetto alla precedente). Vorremmo trovare un sistema di riferimento $O'x'y'$ nel piano con coordinate x', y' in modo tale che l'equazione (23.1.1) divenga più semplice e, soprattutto, riconoscibile: per esempio potremmo desiderare d'avere un'equazione *canonica* nel senso della seguente definizione.

Definizione 23.2 (Forma canonica di una conica).

Nel piano sia fissato un sistema di riferimento Oxy e sia \mathcal{C} una conica. Diciamo che \mathcal{C} è in *forma canonica* (o che Oxy è un *sistema di riferimento canonico* per \mathcal{C}) se l'equazione di \mathcal{C} è in una delle due forme

$$\alpha x^2 + \beta y^2 = \gamma, \quad \beta y^2 = 2\gamma x$$

per qualche $\alpha, \beta, \gamma \in \mathbb{R}$.

Cerchiamo, quindi, di capire come fare per passare dal “vecchio riferimento” Oxy a un “nuovo riferimento” $O'x'y'$ rispetto a cui \mathcal{C} sia in forma canonica.

Per individuare l'angolo di rotazione φ e l'origine O' del nuovo sistema di riferimento faremo uso della teoria delle forme quadratiche vista in precedenza. Infatti all'equazione (23.1.1) si possono associare facilmente le due forme quadratiche

$$a_{1,1}x^2 + 2a_{1,2}xy + a_{2,2}y^2, \quad a_{1,1}x^2 + 2a_{1,2}xy + a_{2,2}y^2 + 2a_{1,3}xt + 2a_{2,3}yt + a_{3,3}t^2$$

le cui matrici (simmetriche) sono, rispettivamente,

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

dette rispettivamente *matrice dei termini di secondo grado della conica* e *matrice (completa) della conica*. Osserviamo che si ha

$$(x \ y \ 1) B \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = a_{1,1}x^2 + 2a_{1,2}xy + a_{2,2}y^2 + 2a_{1,3}x + 2a_{2,3}y + a_{3,3}. \quad (23.1.2)$$

Esempio 23.3. Nel piano con fissato sistema di riferimento Oxy si consideri la conica \mathcal{C} di equazione

$$3x^2 + 2xy + 3y^2 + 2\sqrt{2}x = 0.$$

La matrice dei termini di secondo grado di \mathcal{C} e la matrice completa di \mathcal{C} sono rispettivamente

$$A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 1 & \sqrt{2} \\ 1 & 3 & 0 \\ \sqrt{2} & 0 & 0 \end{pmatrix}. \quad \spadesuit$$

Sia ora $O'x'y'$ un nuovo sistema di riferimento nel piano legato a Oxy da una certa rototraslazione della forma

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} u \\ v \end{pmatrix}$$

(si veda la Proposizione 22.6) e supponiamo che \mathcal{C} sia rappresentata rispetto a tale sistema dall'equazione

$$(x' \ y' \ 1) B' \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = 0. \quad (23.1.3)$$

Qual è allora il legame tra le matrici B e B' nelle equazioni (23.1.2) e (23.1.3)? La rototraslazione di cui sopra si può scrivere in forma compatta come

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi & u \\ \sin \varphi & \cos \varphi & v \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix}.$$

Sostituendo tale formula nelle (23.1.2) e (23.1.3) otteniamo il seguente risultato.

Proposizione 23.4.

Nel piano siano fissati due sistemi di riferimento Oxy e $O'x'y'$. Si assuma che il semiasse positivo delle x' formi un angolo φ (misurato in senso antiorario) con il semiasse positivo delle x e che le coordinate di O' rispetto al sistema di riferimento Oxy siano (u, v) . Poniamo

$$Q = \begin{pmatrix} \cos \varphi & -\sin \varphi & u \\ \sin \varphi & \cos \varphi & v \\ 0 & 0 & 1 \end{pmatrix}, \quad P = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}.$$

Sia poi \mathcal{C} una conica avente nei due sistemi di riferimento matrici complete B e B' e matrici dei termini di grado 2 A e A' rispettivamente. Allora esiste $\lambda \in \mathbb{R}$ non nullo tale che $\lambda B' = {}^tQBQ$ e $\lambda A' = {}^tPAP$.

Data una conica \mathcal{C} di equazione (23.1.1), per determinare una sua equazione canonica si può dunque procedere come segue.

Si determina prima una matrice ortogonale speciale che diagonalizzi la matrice A del complesso dei termini di secondo grado: in questo modo l'equazione di \mathcal{C} viene trasformata in una della forma

$$\widehat{a}_{1,1}\widehat{x}^2 + \widehat{a}_{2,2}\widehat{y}^2 + 2\widehat{a}_{1,3}\widehat{x} + 2\widehat{a}_{2,3}\widehat{y} + \widehat{a}_{3,3} = 0.$$

Ciò equivale a una rotazione che ci fa passare dal vecchio sistema di riferimento Oxy a un sistema di riferimento ausiliare $O\widehat{x}\widehat{y}$.

A questo punto con trasformazioni del tipo $(\widehat{x}, \widehat{y}) \mapsto (x' + a, y' + b)$ (cioè *formando i quadrati*: è il metodo con cui si risolvono le equazioni di secondo grado!) si fa in modo che scompaiano il massimo numero di monomi di grado 1 e, eventualmente, il termine noto. L'equazione risultante diviene

$$\alpha x'^2 + \beta y'^2 = \gamma,$$

se $\det(A) \neq 0$, oppure

$$\beta y'^2 = 2\gamma x',$$

se $\det(A) = 0$, dunque è in forma canonica. Ciò equivale a una traslazione che ci fa passare dal sistema di riferimento ausiliario $O\widehat{x}\widehat{y}$ al nuovo sistema di riferimento $O'x'y'$.

Illustriamo quanto visto con un esempio dettagliato.

Esempio 23.5. Consideriamo la conica \mathcal{C} dell'Esempio 23.3 di equazione

$$3x^2 + 2xy + 3y^2 + 2\sqrt{2}x = 0, \quad (23.1.4)$$

con matrici associate

$$A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 1 & \sqrt{2} \\ 1 & 3 & 0 \\ \sqrt{2} & 0 & 0 \end{pmatrix}.$$

Il polinomio caratteristico di A è

$$p_A(t) = \begin{vmatrix} 3-t & 1 \\ 1 & 3-t \end{vmatrix} = t^2 - 6t + 8 = (t-2)(t-4),$$

quindi gli autovalori di A sono 2 e 4. L'autospazio $E_A(2)$ si determina risolvendo il sistema

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} :$$

calcoliamo che $E_A(2) = \mathcal{L}((1, -1))$. Poiché sappiamo che l'autospazio relativo all'altro autovalore 4 contiene autovettori non nulli (per definizione) ortogonali agli autovettori di $E_A(2)$ (perché A è simmetrica: si veda la Proposizione 20.24), sappiamo che l'autospazio $E_A(4) = \mathcal{L}((1, 1))$.

Quindi la matrice ortogonale speciale cercata è

$$P = \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}$$

e la rotazione corrispondente è

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix} \begin{pmatrix} \widehat{x} \\ \widehat{y} \end{pmatrix}$$

o, equivalentemente,

$$\begin{cases} x = \sqrt{2}/2\widehat{x} + \sqrt{2}/2\widehat{y} \\ y = -\sqrt{2}/2\widehat{x} + \sqrt{2}/2\widehat{y}. \end{cases}$$

La conica \mathcal{C} nel sistema di riferimento ausiliario $O\widehat{x}\widehat{y}$ ha matrice dei termini di secondo grado

$$\widehat{A} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}.$$

Per determinare i monomi di grado 1 nella sua equazione relativa a $O\widehat{x}\widehat{y}$ basta sostituire nell'equazione (23.1.4) le espressioni di x e y in funzione di \widehat{x} e \widehat{y} nei monomi di grado 1. Quindi la sua equazione è

$$2\widehat{x}^2 + 4\widehat{y}^2 + 2\widehat{x} + 2\widehat{y} = 0.$$

Si noti che

$$\begin{aligned} 2\widehat{x}^2 + 2\widehat{x} &= 2(\widehat{x}^2 + \widehat{x}) = 2((\widehat{x} + 1/2)^2 - 1/4) = 2(\widehat{x} + 1/2)^2 - 1/2, \\ 4\widehat{y}^2 + 2\widehat{y} &= 4(\widehat{y}^2 + \widehat{y}/2) = 4((\widehat{y} + 1/4)^2 - 1/16) = 4(\widehat{y} + 1/4)^2 - 1/4. \end{aligned}$$

Ponendo $x' = \widehat{x} + 1/2$ e $y' = \widehat{y} + 1/4$, l'equazione di \mathcal{C} rispetto al sistema di riferimento $O'x'y'$ diviene

$$2x'^2 + 4y'^2 - 3/4 = 0,$$

che è un'equazione canonica. Moltiplicando ambo i membri per $4/3$ otteniamo infine l'equazione

$$\frac{8}{3}x'^2 + \frac{16}{3}y'^2 = 1. \quad (23.1.5)$$

Deduciamo che \mathcal{C} è un'ellisse di semiassi $\sqrt{3/8}$ e $\sqrt{3/16}$.

Le equazioni della rototraslazione che trasforma l'equazione (23.1.4) nell'equazione (23.1.5) sono

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} -3\sqrt{2}/8 \\ \sqrt{2}/8 \end{pmatrix}. \quad (23.1.6)$$

Il centro di \mathcal{C} è il punto che, del nuovo sistema di riferimento, è l'origine. Quindi è il punto che ha coordinate $(0, 0)$ rispetto a $O'x'y'$. Dalle equazioni (23.1.6) deduciamo allora che il centro di \mathcal{C} ha coordinate $(-3\sqrt{2}/8, \sqrt{2}/8)$ rispetto a Oxy .

Gli assi di \mathcal{C} sono le rette che, del nuovo sistema di riferimento, sono gli assi coordinati. Quindi sono le rette di equazioni $x' = 0$ e $y' = 0$ rispetto a $O'x'y'$. Dalle equazioni (23.1.6) deduciamo che

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1/2 \\ 1/4 \end{pmatrix},$$

dunque gli assi della conica \mathcal{C} sono le rette di equazione $\sqrt{2}x - \sqrt{2}y + 1 = 0$ e $2\sqrt{2}x + 2\sqrt{2}y + 1 = 0$ rispetto a Oxy .

In maniera simile il lettore determini le coordinate dei vertici di \mathcal{C} rispetto a Oxy .

Per disegnare \mathcal{C} , oltre alle informazioni già determinate, può essere utile calcolare le intersezioni con gli assi x e y . Per calcolare le intersezioni di \mathcal{C} con l'asse delle ordinate risolviamo il sistema

$$\begin{cases} 3x^2 + 2xy + 3y^2 + 2\sqrt{2}x = 0 \\ x = 0 \end{cases}$$

la cui unica soluzione è $(0, 0)$. Le intersezioni con l'asse delle ascisse sono invece date dalle soluzioni del sistema

$$\begin{cases} 3x^2 + 2xy + 3y^2 + 2\sqrt{2}x = 0 \\ y = 0 \end{cases}$$

cioè $(0, 0)$ e $(-2\sqrt{2}/3, 0)$.

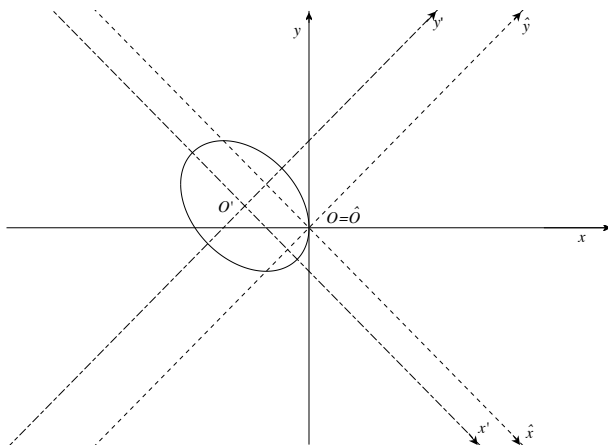


Figura 23.1

In Figura 23.1 abbiamo riportato il disegno della conica \mathcal{C} . ♠

23.2 Determinazione del tipo di una conica

Abbiamo visto finora che ogni conica \mathcal{C} ha una forma canonica e abbiamo trattato in dettaglio come trovarla nell'Esempio 23.5. Viene spontaneo domandarsi se c'è un metodo che, a priori, ci permetta di stabilire se \mathcal{C} è una conica classica o no e, nel caso \mathcal{C} sia classica se sia un'ellisse, un'iperbole o una parabola, senza dovere necessariamente operare il cambio di coordinate.

Nel seguito indicheremo con A la matrice dei termini di secondo grado di \mathcal{C} e con B la sua matrice completa.

Definizione 23.6.

Una conica \mathcal{C} si dice *degenere* se la sua equazione si decompone in un prodotto di due polinomi di grado 1 (non necessariamente distinti), *non degenere* altrimenti.

È facile rendersi conto che una conica in forma canonica è degenere se e solo se la sua matrice B' ha determinante nullo.

Poiché dalla Proposizione 23.4 segue che la matrice B di \mathcal{C} è legata a quella B' della sua forma canonica da una relazione del tipo $\lambda B' = {}^tQBQ$ per un qualche $\lambda \in \mathbb{R}$ non nullo, calcolando i determinanti di ambo i membri possiamo dedurre il seguente risultato.

Proposizione 23.7.

Una conica \mathcal{C} è degenere se e solo se la sua matrice completa B ha determinante nullo.

Supponiamo che \mathcal{C} sia non degenere: allora \mathcal{C} è un'ellisse immaginaria, oppure è una conica classica. Supponiamo di essere in questo secondo caso. Se $\alpha x'^2 + \beta y'^2 = \gamma$ è una sua equazione canonica, \mathcal{C} è un'ellisse, un'iperbole o una parabola secondoché α e β siano concordi, discordi, o uno di essi sia nullo. Indicata con

$$A' = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$$

la matrice dei termini di secondo grado della forma canonica di \mathcal{C} , di nuovo dalla Proposizione 23.4 sappiamo che $\lambda A' = {}^t P A P$ per un qualche $\lambda \in \mathbb{R}$ non nullo. Deduciamo allora che \mathcal{C} è un'ellisse, un'iperbole o una parabola secondoché A' , e dunque A , sia definita, indefinita, semidefinita.

Queste considerazioni ci permettono anche di classificare una conica calcolandone la sua equazione canonica senza necessariamente determinare la rototraslazione che la riduce in tale forma canonica.

Esempio 23.8. Si consideri la conica \mathcal{C} di equazione

$$4x^2 + 4xy + 4y + y^2 = 0.$$

Le matrici associate a \mathcal{C} sono

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & 2 & 0 \\ 2 & 1 & 2 \\ 0 & 2 & 0 \end{pmatrix}.$$

Poiché $\det(B) = -16 \neq 0$, la conica \mathcal{C} è non degenere. Poiché $\det(A) = 0$, la conica \mathcal{C} è una parabola. Se vogliamo determinarne l'equazione canonica osserviamo che le matrici Q e P definite nella Proposizione Proposizione 23.4 soddisfano le relazioni

$${}^t Q B Q = \lambda B' = \lambda \begin{pmatrix} 0 & 0 & -\gamma \\ 0 & \beta & 0 \\ -\gamma & 0 & 0 \end{pmatrix}, \quad {}^t P A P = \lambda A' = \lambda \begin{pmatrix} 0 & 0 \\ 0 & \beta \end{pmatrix},$$

per un qualche $\lambda \in \mathbb{R}$ non nullo. Poiché l'equazione di una conica e, di conseguenza, la sua matrice completa, è individuata a meno di una costante moltiplicativa, possiamo sempre supporre che sia $\lambda = 1$.

Quindi β deve essere l'autovalore non nullo di A , cioè $\beta = 5$. Confrontando i determinanti di B e B' si ottiene $-\beta\gamma^2 = -16$, cioè $\gamma = 4/\sqrt{5}$. Concludiamo che \mathcal{C} è una parabola la cui equazione canonica è

$$5y'^2 = \frac{8}{\sqrt{5}}x'.$$

Il parametro di \mathcal{C} è dunque $8/5\sqrt{5}$.

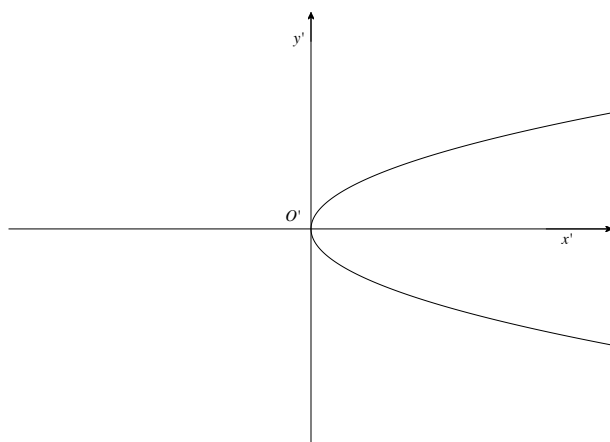


Figura 23.2

Nella Figura 23.2 riportiamo il disegno della conica \mathcal{C} . ♠

Esempio 23.9. Si consideri la conica \mathcal{C} di equazione

$$7x^2 + 8xy + y^2 + 9x - 1 = 0.$$

Le matrici associate a \mathcal{C} sono

$$A = \begin{pmatrix} 7 & 4 \\ 4 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 7 & 4 & 9/2 \\ 4 & 1 & 0 \\ 9/2 & 0 & -1 \end{pmatrix}.$$

Poiché $\det(B) = -45/4 \neq 0$, la conica \mathcal{C} è non degenera. Poiché $\det(A) = -9$, la conica \mathcal{C} è un'iperbole. Se vogliamo calcolarne l'equazione canonica osserviamo che dobbiamo determinare le matrici Q e P definite nella Proposizione 23.4 tali che

$${}^tQBQ = B' = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & -\gamma \end{pmatrix}, \quad {}^tPAP = A' = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}.$$

Quindi α e β devono essere gli autovalori di A , cioè $\alpha = 9$ e $\beta = -1$, e confrontando i determinanti di B e B' , $-\alpha\beta\gamma = -45/4$, cioè $\gamma = -5/4$. Concludiamo che \mathcal{C} è un'iperbole la cui equazione canonica è

$$9x'^2 - y'^2 = -5/4.$$

Moltiplicando ambo i membri per $-4/5$ otteniamo

$$-\frac{36}{5}x'^2 + \frac{4}{5}y'^2 = 1.$$

I semiassi di \mathcal{C} sono dunque $\sqrt{5}/6$ (quello corrispondente all'asse immaginario) e $\sqrt{5}/2$ (quello corrispondente all'asse trasverso).

Nella Figura 23.3 riportiamo il disegno di \mathcal{C} nel sistema di riferimento $Ox'y'$.

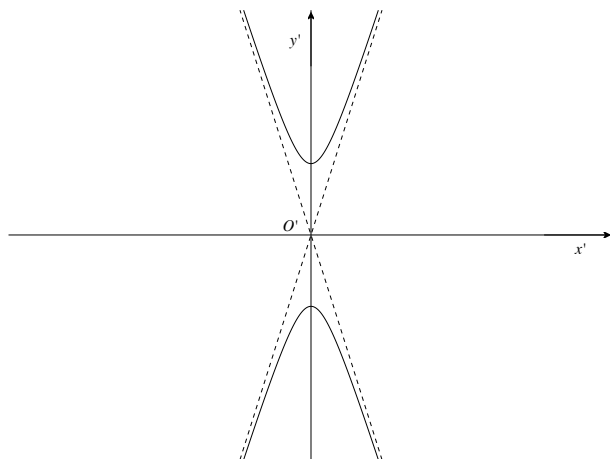


Figura 23.3

Tale ultima equazione non è l'equazione standard data nell'Esempio 22.1. Consideriamo allora la rotazione $x' = -y''$, $y' = x''$ di $\pi/2$ radianti (si veda l'Esempio 22.5) che trasforma l'equazione di \mathcal{C} nell'equazione standard

$$\frac{4}{5}x''^2 - \frac{36}{5}y''^2 = 1.$$

Nella Figura 23.4 riportiamo il disegno di \mathcal{C} nel sistema di riferimento $Ox''y''$.

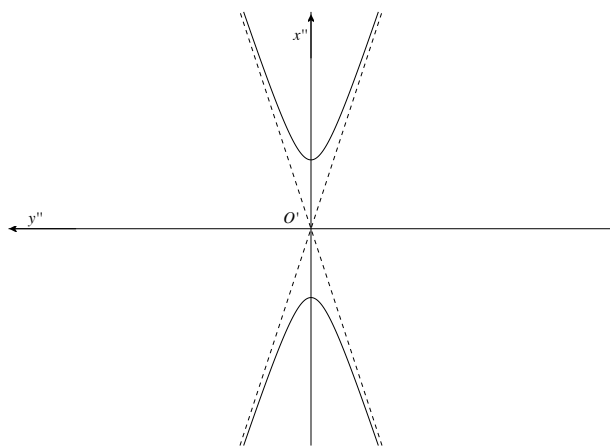


Figura 23.4



Nel caso si sia verificato che una certa conica \mathcal{C} è a centro (cioè \mathcal{C} sia un'ellisse o un'iperbole), è possibile determinarne il centro di simmetria in maniera relativamente facile.

Si ricordi che per ridurre \mathcal{C} a forma canonica si può procedere in due modi equivalenti tramite una rototraslazione. Si può procedere come nell'Esempio 23.3 facendo prima una rotazione e poi una traslazione. Oppure si può prima operare la traslazione e poi procedere con una rotazione (si veda il Paragrafo 22.2).

Consideriamo questo secondo approccio: se

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & u \\ 0 & 1 & v \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{y} \\ 1 \end{pmatrix}$$

è la traslazione e

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

è la matrice di \mathcal{C} , allora con qualche conto si osserva che la matrice di \mathcal{C} nel sistema di riferimento traslato $O'\hat{x}\hat{y}$ sarà trasformata nella matrice

$$\hat{B} = \begin{pmatrix} a_{1,1} & a_{1,2} & \hat{a}_{1,3} \\ a_{1,2} & a_{2,2} & \hat{a}_{2,3} \\ \hat{a}_{1,3} & \hat{a}_{2,3} & \hat{a}_{3,3} \end{pmatrix}$$

ove $\hat{a}_{1,3} = a_{1,1}u + a_{1,2}v + a_{1,3}$ e $\hat{a}_{2,3} = a_{2,1}u + a_{2,2}v + a_{2,3}$.

Poiché l'eventuale rotazione non ci permette di eliminare termini lineari che possono essere presenti nell'equazione di \mathcal{C} , dobbiamo determinare u e v in modo tale che

$$\begin{cases} a_{1,1}u + a_{1,2}v + a_{1,3} = 0 \\ a_{2,1}u + a_{2,2}v + a_{2,3} = 0. \end{cases}$$

Tenendo conto che \mathcal{C} è a centro, sappiamo che la matrice A dei termini di secondo grado è invertibile, quindi il sistema precedente ha un'unica soluzione (u, v) che rappresenta le coordinate del centro di \mathcal{C} .

Esempio 23.10. Riprendiamo la conica \mathcal{C} di equazione

$$3x^2 + 2xy + 3y^2 + 2\sqrt{2}x = 0$$

descritta in dettaglio nell'Esempio 23.5. Le coordinate del centro si ottengono risolvendo il sistema

$$\begin{cases} 3u + v + \sqrt{2} = 0 \\ u + 3v = 0, \end{cases}$$

quindi sono $(-3\sqrt{2}/8, \sqrt{2}/8)$. ♠

Esempio 23.11. Si consideri la conica \mathcal{C} di equazione

$$7x^2 + 8xy + y^2 + 9x - 1 = 0$$

classificata nell'Esempio 23.9. Le coordinate del centro si ottengono risolvendo il sistema

$$\begin{cases} 7u + 4v + 9/2 = 0 \\ 4u + v = 0, \end{cases}$$

quindi sono $(-1/2, 2)$.



Sfere e circonferenze nello spazio

24.1 Sfere nello spazio

In questo capitolo studieremo alcuni dei più semplici oggetti geometrici “non lineari”: circonferenze e sfere nello spazio S_3 . Analizzeremo poi in dettaglio il caso delle circonferenze in un piano qualsiasi.

Definizione 24.1 (Sfere).

Siano fissati un punto $C \in S_3$ e un numero reale positivo $\varrho \in \mathbb{R}$. Definiamo *sfera di centro C e raggio ϱ* il luogo $\mathcal{S}(C, \varrho)$ dei punti $P \in S_3$ a distanza fissata ϱ da C , ovvero tali che $d(P, C) = \varrho$.

Nella Figura 24.1 è illustrata una sfera di centro C e raggio ϱ .

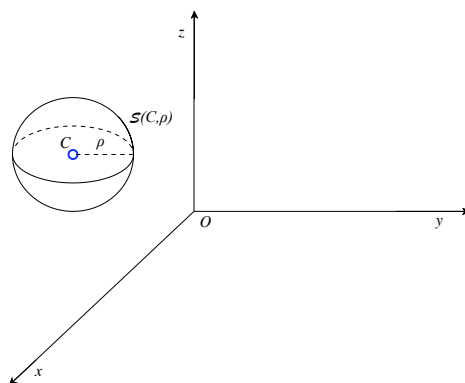


Figura 24.1

Fissiamo in S_3 un riferimento $Oxyz$ e analizziamo la condizione $d(P, C) = \varrho$: entrambe le quantità ai due membri sono positive, quindi essa è equivalente alla condizione $d(P, C)^2 = \varrho^2$. Se il punto C ha coordinate $C = (x_C, y_C, z_C) \in S_3$, dalla condizione $d(P, C)^2 = \varrho^2$ si ottiene l'equazione cartesiana della sfera nello spazio

$$(x - x_C)^2 + (y - y_C)^2 + (z - z_C)^2 = \varrho^2.$$

Svolgendo i conti si trova l'equazione cartesiana della sfera di centro $C = (x_C, y_C, z_C)$ e raggio ϱ

$$x^2 + y^2 + z^2 - 2x_Cx - 2y_Cy - 2z_Cz + x_C^2 + y_C^2 + z_C^2 - \varrho^2 = 0 : \quad (24.1.1)$$

ciò significa che

$$\mathcal{S}(C, \varrho) = \{ (x, y, z) \mid x^2 + y^2 + z^2 - 2x_Cx - 2y_Cy - 2z_Cz + x_C^2 + y_C^2 + z_C^2 - \varrho^2 = 0 \}.$$

Osserviamo che, dal momento che siamo interessati al luogo dei punti che annullano l'equazione (24.1.1) e non all'equazione stessa, possiamo sostituire a essa un qualsiasi suo multiplo non nullo: quindi, per ogni $\lambda \in \mathbb{R}$ non nullo, abbiamo anche

$$\mathcal{S}(C, \varrho) = \{ (x, y, z) \mid \lambda(x^2 + y^2 + z^2 - 2x_Cx - 2y_Cy - 2z_Cz + x_C^2 + y_C^2 + z_C^2 - \varrho^2) = 0 \}.$$

Esempio 24.2. La sfera di centro $C = (0, -2, 1)$ e raggio $\varrho = 1$ ha equazione

$$(x - 0)^2 + (y + 2)^2 + (z - 1)^2 - 1 = x^2 + y^2 + z^2 + 4y - 2z + 4 = 0.$$

Anche l'equazione

$$-2x^2 - 2y^2 - 2z^2 - 8y + 4z - 8 = 0$$

è un'equazione della stessa sfera di centro $C = (0, -2, 1)$ e raggio $\varrho = 1$ ♠

L'equazione (24.1.1) ha due caratteristiche principali: la prima è che manca dei monomi "misti" (cioè in xy , xz , yz), la seconda è che i coefficienti dei termini quadratici sono non nulli e uguali fra loro.

Viceversa, supponiamo di avere un'equazione di grado 2 con tali proprietà, quindi (a meno di dividere per il coefficiente comune dei termini quadratici) riconducibile alla forma

$$x^2 + y^2 + z^2 + \alpha x + \beta y + \gamma z + \delta = 0 : \quad (24.1.2)$$

ci domandiamo se l'equazione (24.1.2) rappresenta una sfera e, in caso affermativo, come calcolare il suo centro e il suo raggio.

Confrontando le equazioni (24.1.1) e (24.1.2), deduciamo che, affinché l'equazione (24.1.2) rappresenti una sfera, devono esistere $x_C, y_C, z_C \in \mathbb{R}$ e $\varrho \in \mathbb{R}$ positivo per cui valgono le relazioni

$$\alpha = -2x_C, \quad \beta = -2y_C, \quad \gamma = -2z_C, \quad \delta = x_C^2 + y_C^2 + z_C^2 - \varrho^2,$$

quindi

$$x_C = -\frac{\alpha}{2}, \quad y_C = -\frac{\beta}{2}, \quad z_C = -\frac{\gamma}{2}, \quad 4\varrho^2 = \alpha^2 + \beta^2 + \gamma^2 - 4\delta.$$

Queste considerazioni dimostrano il seguente risultato.

Proposizione 24.3.

Sia fissato un sistema di riferimento $Oxyz$ in S_3 . L'insieme

$$\mathcal{S} = \{ (x, y, z) \mid x^2 + y^2 + z^2 + \alpha x + \beta y + \gamma z + \delta = 0 \}$$

è una sfera se e solo se $\alpha^2 + \beta^2 + \gamma^2 - 4\delta > 0$.

Se ciò accade, risulta $\mathcal{S} = \mathcal{S}(C, \varrho)$ con

$$C = \left(-\frac{\alpha}{2}, -\frac{\beta}{2}, -\frac{\gamma}{2} \right), \quad \varrho = \frac{\sqrt{\alpha^2 + \beta^2 + \gamma^2 - 4\delta}}{2}.$$

⚠ Qualora valga la condizione $\alpha^2 + \beta^2 + \gamma^2 - 4\delta < 0$ per l'equazione (24.1.2), si dice che essa rappresenta una *sfera immaginaria* (o, anche, una *sfera di raggio immaginario o a punti immaginari*)

Invece, quando vale la condizione $\alpha^2 + \beta^2 + \gamma^2 - 4\delta = 0$ la sfera si riduce a un solo punto: in tal caso si dice che essa rappresenta una *sfera degenera*.

Esempio 24.4. Si consideri l'equazione

$$x^2 + y^2 + z^2 + 3x - 2y + 1 = 0.$$

Poiché $3^2 + (-2)^2 + 0^2 - 4 \cdot 1 = 9 > 0$, essa è l'equazione di una sfera S in S_3 . Il suo centro è $C = (-3/2, 1, 0)$, il suo raggio $\varrho = 3/2$.

Invece l'equazione

$$x^2 + y^2 + z^2 + 3x - 2y + 4 = 0$$

non rappresenta una sfera nel senso della Definizione 24.1, ma una sfera immaginaria: infatti $3^2 + (-2)^2 + 0^2 - 4 \cdot 4 = -3 < 0$. ♠

24.2 Circonferenze nello spazio

Definizione 24.5 (Circonferenze).

Siano dati un piano $\pi \subseteq S_3$, $C \in \pi$ e $\varrho \in \mathbb{R}$ positivo. Definiamo *circonferenza* $\mathcal{C}(\pi, C, \varrho)$ del piano π , di centro C e raggio ϱ il luogo dei punti $P \in \pi$ a distanza fissata ϱ da C , ovvero tali che $d(P, C) = \varrho$.

Per rappresentare la circonferenza $\mathcal{C}(\pi, C, \varrho)$ ci possono essere vari modi. Il più comodo è quello di pensarla come intersezione del piano π con $\mathcal{S}(C, \varrho)$: se π ha equazione

$$ax + by + cz = d$$

e $C = (x_C, y_C, z_C)$, otteniamo le equazioni cartesiane per $\mathcal{C}(\pi, C, \varrho)$:

$$\begin{cases} ax + by + cz = d \\ (x - x_C)^2 + (y - y_C)^2 + (z - z_C)^2 = \varrho^2. \end{cases}$$

Nella figura 24.2 è illustrata una circonferenza come intersezione del piano π con la sfera $\mathcal{S}(C, \varrho)$.

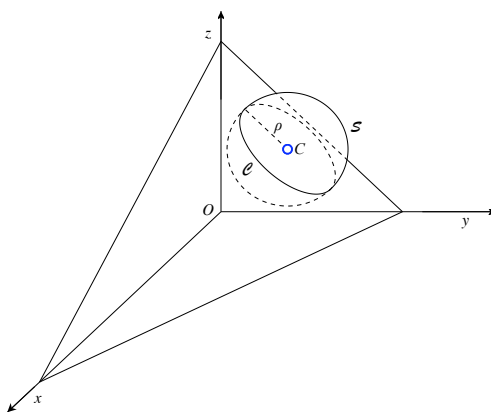


Figura 24.2

Esempio 24.6. Nel piano π di equazione $x + y + z = 3$ sia $C = (1, 1, 1)$: la circonferenza del piano π di centro C e raggio $\rho = 1$ ha equazioni cartesiane

$$\begin{cases} x + y + z = 3 \\ x^2 + y^2 + z^2 - 2x - 2y - 2z + 2 = 0. \end{cases} \spadesuit$$

Come abbiamo fatto nel caso della sfera, ci poniamo ora il problema inverso a quello della rappresentazione: dato un sistema della forma

$$\begin{cases} ax + by + cz = d \\ x^2 + y^2 + z^2 + \alpha x + \beta y + \gamma z + \delta = 0, \end{cases} \quad (24.2.1)$$

ci domandiamo se esso rappresenta una circonferenza e, in caso affermativo, come calcolare il suo centro e il suo raggio (il piano d'appartenenza è, evidentemente, quello d'equazione $ax + by + cz = d$).

Affinché il sistema (24.2.1) rappresenti una circonferenza è innanzi tutto necessario che l'equazione $x^2 + y^2 + z^2 + \alpha x + \beta y + \gamma z + \delta = 0$ rappresenti una sfera $\mathcal{S}(C, \rho)$. Se ciò accade, allora occorre e basta che il piano π di equazione $ax + by + cz = d$ e la sfera $\mathcal{S}(C, \rho)$ abbiano punti in comune: ciò accade se e solo se C ha distanza da α minore di ρ .

Sia ora $\mathcal{C} = \pi \cap \mathcal{S}(C, \rho)$: come si può osservare in Figura 24.3, il centro della circonferenza \mathcal{C} è la proiezione ortogonale C' sul piano π , mentre il raggio ρ' di \mathcal{C} soddisfa la relazione $\rho^2 = \rho'^2 + d(\pi, C)^2$, da cui si deduce

$$\rho' = \sqrt{\rho^2 - d(\pi, C)^2}. \quad (24.2.2)$$

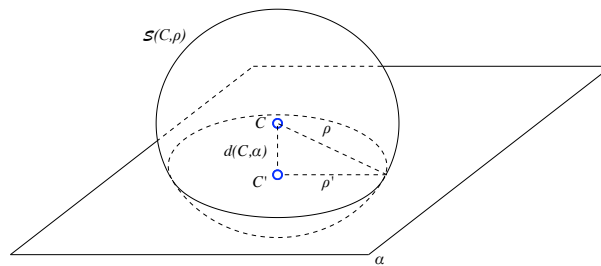


Figura 24.3

Se, invece, $d(C, \pi) > \rho$, allora il sistema (24.2.1) non ha soluzioni, cioè $\pi \cap \mathcal{S}(C, \rho) = \emptyset$. È questo il caso illustrato in Figura 24.4.

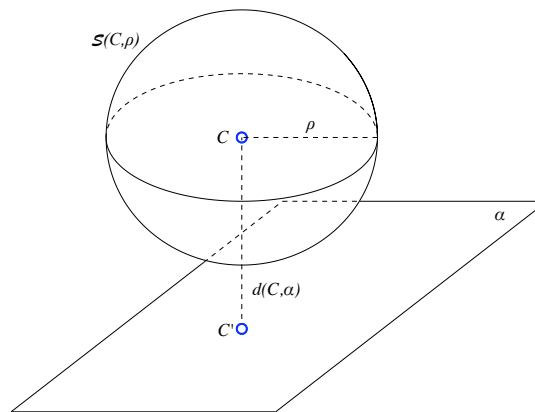


Figura 24.4

Esempio 24.7. Si consideri la famiglia di piani π_h di equazione $x + y + z = 1 + h$, $h \in \mathbb{R}$. Sia \mathcal{S} la sfera di equazione

$$x^2 + y^2 + z^2 - 2x - 2y - 2z - 1 = 0.$$

Vogliamo individuare i valori di $h \in \mathbb{R}$ tali che $\mathcal{S} \cap \pi_h$ sia una circonferenza. A tale scopo osserviamo che \mathcal{S} ha centro nel punto $C = (1, 1, 1)$ e raggio $\rho = 2$. Poiché

$$d(C, \pi_h) = \frac{|2 - h|}{\sqrt{3}},$$

segue che $\mathcal{S} \cap \pi_h$ è una circonferenza se e solo se $|2 - h| < 2\sqrt{3}$: svolgendo i calcoli ciò significa che $\mathcal{S} \cap \pi_h$ è una circonferenza se e solo se $h \in]2 - 2\sqrt{3}, 2 + 2\sqrt{3}[$.

Siano C_h e ρ_h rispettivamente il centro e il raggio di questa circonferenza, di modo che $\mathcal{C}(\pi_h, C_h, \rho_h) = \mathcal{S} \cap \pi_h$. Per determinare ρ_h utilizziamo la formula (24.2.2), ottenendo

$$\rho_h = \sqrt{2^2 - \frac{(2 - h)^2}{3}} = \sqrt{\frac{8 + 4h - h^2}{3}}.$$

Per quanto riguarda il calcolo delle coordinate del centro C_h , si noti che la retta u per C e perpendicolare a π_h ha equazioni parametriche

$$\begin{cases} x = 1 + t \\ y = 1 + t \\ z = 1 + t, \end{cases} \quad t \in \mathbb{R}.$$

Dunque

$$C_h = u \cap \pi_h = \left(\frac{1+h}{3}, \frac{1+h}{3}, \frac{1+h}{3} \right). \quad \spadesuit$$

Analizziamo adesso il caso in cui la distanza del centro della sfera dal piano è uguale al raggio della circonferenza, cioè $d(C, \pi) = \rho$: allora $\pi \cap \mathcal{S}(C, \rho)$ si riduce a un solo punto P_0 . Se $\alpha \neq \pi$ è un piano diverso da π che passa per P_0 , allora α non è perpendicolare a $P_0 - C$. Se chiamiamo C' la proiezione ortogonale di C su α allora l'intersezione $\alpha \cap \mathcal{S}(C, \rho)$ è a sua volta una circonferenza di centro C' e raggio $|P_0 - C'|$, quindi contiene infiniti punti. Concludiamo che π è l'unico piano passante per P_0 e perpendicolare a $P_0 - C$. Questo caso è illustrato in Figura 24.5.

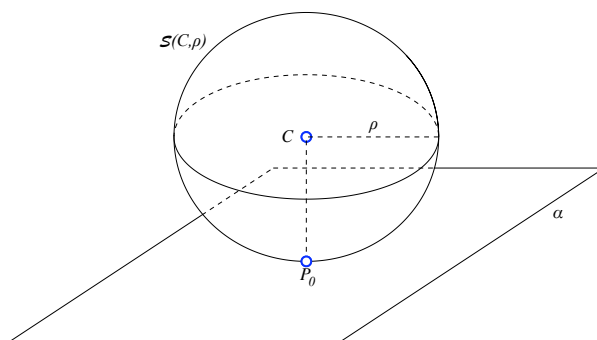


Figura 24.5

Definizione 24.8 (Piano e retta tangenti a una sfera).

Si considerino la sfera $\mathcal{S}(C, \rho) \subseteq S_3$ e un suo punto $P_0 \in \mathcal{S}(C, \rho)$:

- si definisce *piano tangente a $\mathcal{S}(C, \rho)$ nel punto P_0* l'unico piano passante per P_0 e perpendicolare a $P_0 - C$;
- una *retta tangente a $\mathcal{S}(C, \rho)$ in P_0* è una qualsiasi retta passante per P_0 e contenuta nel piano tangente a $\mathcal{S}(C, \rho)$ nel punto P_0 .

Si noti che il piano tangente è lo stesso per tutte le sfere passanti per il punto P_0 e aventi centro sulla retta per P_0 e C : infatti il centro di tali sfere ha coordinate

$$(x_0 + t(x_C - x_0), y_0 + t(y_C - y_0), z_0 + t(z_C - z_0))$$

per un opportuno $t \in \mathbb{R}$ non nullo, dunque il piano tangente in P_0 ha in tal caso equazione

$$t(x_0 - x_C)(x - x_0) + t(y_0 - y_C)(y - y_0) + t(z_0 - z_C)(z - z_0) = 0,$$

cioè, semplificando t ,

$$(x_0 - x_C)(x - x_0) + (y_0 - y_C)(y - y_0) + (z_0 - z_C)(z - z_0) = 0.$$

Sia ora \mathcal{C} una circonferenza che giace su un piano π : è immediato verificare che i piani tangenti in P_0 alle sfere \mathcal{S} contenenti \mathcal{C} passano tutti per una stessa retta $r \subseteq \pi$. Tale retta interseca \mathcal{C} solo in P_0 e ha la proprietà di essere perpendicolare al vettore $P_0 - C'$, ove $C' \in \pi$ è il centro di \mathcal{C} .

Esempio 24.9. Siano $C = (1, 1, 1)$ e $\rho = \sqrt{3}$. La sfera $\mathcal{S}(C, \rho) \subseteq S_3$ ha equazione

$$x^2 + y^2 + z^2 - 2x - 2y - 2z = 0$$

e contiene il punto $P_0 = (2, 2, 2)$. Il piano tangente a $\mathcal{S}(C, \rho)$ in P_0 ha equazione

$$(2 - 1)(x - 2) + (2 - 1)(y - 2) + (2 - 1)(z - 2) = 0,$$

che, una volta semplificata, diventa $x + y + z = 6$.

La retta

$$\begin{cases} x = 2 + t \\ y = 2 - 2t \\ z = 2 + \ell t, \end{cases} \quad t \in \mathbb{R},$$

è tangente a $\mathcal{S}(C, \rho)$ in P_0 se e solo se $\ell = 1$. ♠

Osservazione 24.10. Un caso interessante di circonferenze sono quelle contenute nel piano xy , cioè quelle le cui equazioni cartesiane sono della forma

$$\begin{cases} z = 0 \\ x^2 + y^2 + \alpha x + \beta y + \gamma z + \delta = 0. \end{cases}$$

Il sistema sopra è equivalente a

$$\begin{cases} z = 0 \\ x^2 + y^2 + \alpha x + \beta y + \delta = 0, \end{cases}$$

che rappresenta la circonferenza data come intersezione del piano xy con un cilindro circolare avente asse perpendicolare a tale piano. Spesso si parla allora della *circonferenza nel piano di equazione*

$$x^2 + y^2 + \alpha x + \beta y + \delta = 0.$$

24.3 Intersezione di due sfere

Vogliamo adesso considerare l'intersezione di due sfere in S_3 , diciamo $\mathcal{S}(C_1, \varrho_1)$ e $\mathcal{S}(C_2, \varrho_2)$: la struttura di tale intersezione $\mathcal{S}(C_1, \varrho_1) \cap \mathcal{S}(C_2, \varrho_2)$ è legata strettamente alla distanza $d(C_1, C_2)$. Più in dettaglio, si possono verificare tre casi principali.

1. Nel primo caso, illustrato in Figura 24.6, la distanza $d(C_1, C_2) > \varrho_1 + \varrho_2$ oppure $d(C_1, C_2) < |\varrho_1 - \varrho_2|$: le due sfere non possono avere punti in comune e sono, rispettivamente, esterne o interne l'una all'altra.

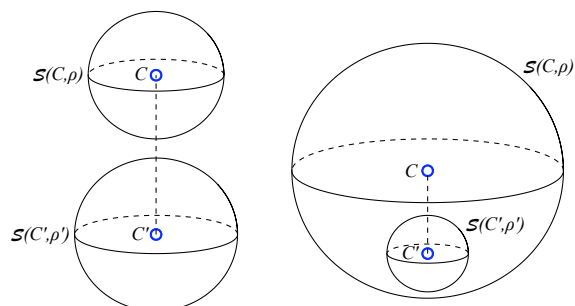


Figura 24.6

2. Nel secondo caso, illustrato in Figura 24.7, la distanza $d(C_1, C_2) = \varrho_1 + \varrho_2$ oppure $d(C_1, C_2) = |\varrho_1 - \varrho_2|$: le due sfere hanno esattamente un punto in comune. Si dicono *tangenti*, rispettivamente, esternamente o internamente.

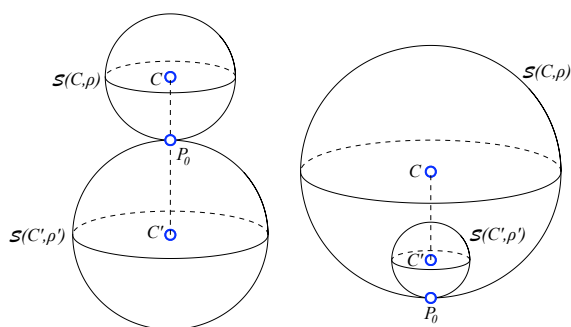


Figura 24.7

3. Nel terzo caso, illustrato in Figura 24.8, $|\varrho_1 - \varrho_2| < d(C_1, C_2) < \varrho_1 + \varrho_2$: in questo caso le due sfere hanno punti in comune. Tali punti descrivono una circonferenza \mathcal{C} avente centro sulla retta che unisce i punti C_1 e C_2 .

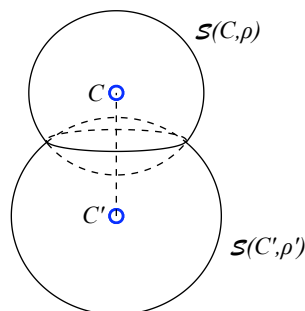


Figura 24.8

Concentriamoci sul terzo caso: vogliamo determinarne equazioni cartesiane per la circonferenza data dall'intersezione delle due sfere $\mathcal{S}(C_1, \varrho_1)$ e $\mathcal{S}(C_2, \varrho_2)$. Si osservi preliminarmente che $C_1 \neq C_2$, cioè le due sfere non sono concentriche. Se le equazioni delle due sfere sono rispettivamente

$$\begin{aligned} x^2 + y^2 + z^2 + \alpha_1 x + \beta_1 y + \gamma_1 z + \delta_1 &= 0 \\ x^2 + y^2 + z^2 + \alpha_2 x + \beta_2 y + \gamma_2 z + \delta_2 &= 0, \end{aligned} \quad (24.3.1)$$

allora tale condizione si traduce nella disuguaglianza $(\alpha_1, \beta_1, \gamma_1) \neq (\alpha_2, \beta_2, \gamma_2)$.

Le coordinate dei punti di \mathcal{C} soddisfano le due equazioni di $\mathcal{S}(C_1, \varrho_1)$ e $\mathcal{S}(C_2, \varrho_2)$, dunque soddisfano anche l'equazione ottenuta sottraendo membro a membro le due equazioni (24.3.1). Quindi le coordinate dei punti di \mathcal{C} soddisfano anche l'equazione di primo grado

$$(\alpha_1 - \alpha_2)x + (\beta_1 - \beta_2)y + (\gamma_1 - \gamma_2)z + (\delta_1 - \delta_2) = 0 :$$

poiché $(\alpha_1, \beta_1, \gamma_1) \neq (\alpha_2, \beta_2, \gamma_2)$ tale equazione rappresenta, nello spazio S_3 , un piano π che contiene \mathcal{C} . Dunque possiamo scrivere $\mathcal{C} = \pi \cap \mathcal{S}(C_i, \varrho_i)$. Si noti che il piano π è perpendicolare a

$$C_2 - C_1 = \frac{(\alpha_1 - \alpha_2)}{2}\vec{i} + \frac{(\beta_1 - \beta_2)}{2}\vec{j} + \frac{(\gamma_1 - \gamma_2)}{2}\vec{k} \neq \vec{0}.$$

Definizione 24.11 (Piano radicale).

Siano \mathcal{S}_1 ed \mathcal{S}_2 sfere non concentriche, descritte rispettivamente dalle equazioni (24.3.1). Il piano di equazione

$$(\alpha_1 - \alpha_2)x + (\beta_1 - \beta_2)y + (\gamma_1 - \gamma_2)z + (\delta_1 - \delta_2) = 0$$

è detto *piano radicale della coppia di sfere \mathcal{S}_1 ed \mathcal{S}_2* .

Osservazione 24.12. Se una delle due sfere è degenere (cioè è un punto), ad esempio $\mathcal{S}_1 = \{P_0\}$ e si ha che $P_0 \in \mathcal{S}_2$, allora è facile vedere che il piano radicale coincide con il piano tangente alla sfera \mathcal{S}_2 nel punto P_0 .

Esempio 24.13. Si considerino le due sfere \mathcal{S}_1 ed \mathcal{S}_2 rispettivamente di equazione

$$x^2 + y^2 + z^2 - 2x - 2y + 4z + 5 = 0, \quad x^2 + y^2 + z^2 + 2x + 2y - 4z + 5 = 0.$$

Il centro di \mathcal{S}_1 è $C_1 = (1, 1, -2)$ e quello di \mathcal{S}_2 è $C_2 = (-1, -1, 2)$, quindi la distanza $d(C_1, C_2) = \sqrt{24}$. Poiché $\varrho_1 = \varrho_2 = 1$ si deduce che $\mathcal{S}_1 \cap \mathcal{S}_2 = \emptyset$: di più, le sfere sono esterne l'una all'altra. ♠

Esempio 24.14. Si considerino le due sfere \mathcal{S}_1 e \mathcal{S}_2 rispettivamente di equazione

$$x^2 + y^2 + z^2 - 2x - 4z + 4 = 0, \quad x^2 + y^2 + z^2 - 2x - 2y - 2z + 2 = 0.$$

Il centro di \mathcal{S}_1 è $C_1 = (1, 0, 2)$, mentre il centro di \mathcal{S}_2 è $C_2 = (1, 1, 1)$, quindi $d(C_1, C_2) = \sqrt{2}$. Per quanto riguarda i raggi, abbiamo $\varrho_1 = \varrho_2 = 1$. Concludiamo che $\mathcal{C} = \mathcal{S}_1 \cap \mathcal{S}_2$ è una circonferenza: calcoliamone centro e raggio.

A tale scopo osserviamo prima che il piano radicale π , che contiene \mathcal{C} , ha equazione

$$y - z + 1 = 0.$$

La retta passante per C_1 e C_2 ha equazioni

$$\begin{cases} x = 1 \\ y = 1 + t \\ z = 1 - t, \end{cases} \quad t \in \mathbb{R}.$$

Concludiamo che il centro di \mathcal{C} è $C = (1, 1/2, 3/2)$. Per quanto riguarda il raggio ϱ , poiché $d(C_1, \pi) = 1/\sqrt{2}$, segue che

$$\varrho = \sqrt{\varrho_1^2 - d(C_1, \pi)^2} = 1/\sqrt{2}. \quad \spadesuit$$

Quanto visto sopra circa l'intersezione di due sfere può essere utile per la determinazione di sfere che soddisfino certe proprietà come, per esempio, contenere una circonferenza data o essere tangenti a un piano dato.

Esempio 24.15. Si consideri la circonferenza \mathcal{C} di equazioni

$$\begin{cases} x^2 + y^2 + z^2 - 7 = 0 \\ x + y + z - 3 = 0. \end{cases}$$

Una sfera contenente \mathcal{C} è perciò \mathcal{S}_1 di equazione

$$x^2 + y^2 + z^2 - 7 = 0.$$

Per quanto osservato sopra, ogni altra sfera \mathcal{S} contenente \mathcal{C} deve avere un'equazione della forma

$$x^2 + y^2 + z^2 + \alpha x + \beta y + \gamma z + \delta = 0$$

tale che

$$(x^2 + y^2 + z^2 + \alpha x + \beta y + \gamma z + \delta) - (x^2 + y^2 + z^2 - 7) = \lambda(x + y + z - 3),$$

cioè

$$x^2 + y^2 + z^2 + \alpha x + \beta y + \gamma z + \delta = x^2 + y^2 + z^2 - 7 + \lambda(x + y + z - 3)$$

per un'opportuno $\lambda \in \mathbb{R}$.

Se, per esempio, vogliamo determinare la sfera contenente \mathcal{C} e passante per il punto $P_0 = (1, 0, 0)$ dobbiamo scegliere λ tale che

$$1^2 + 0^2 + 0^2 - 7 + \lambda(1 + 0 + 0 - 3) = 0,$$

ovvero $\lambda = -3$. Pertanto la sfera cercata ha equazione

$$x^2 + y^2 + z^2 - 3x - 3y - 3z + 2 = 0. \quad \spadesuit$$

Esempio 24.16. Si consideri il piano π di equazione

$$x + y + z - 3 = 0$$

e sia $P_0 = (1, 1, 1)$: si noti che $P_0 \in \pi$. Vogliamo determinare le sfere tangenti a π in P_0 . Per fare ciò possiamo procedere in due modi.

Il primo consiste nel pensare il punto P_0 come una sfera degenera, di equazione

$$(x - 1)^2 + (y - 1)^2 + (z - 1)^2 = 0.$$

Dal momento che una sfera \mathcal{S} è tangente a π in P_0 se e solo se $\mathcal{S} \cap \pi = \{P_0\}$, una tale sfera sarà necessariamente caratterizzata dalla condizione che π sia il piano radicale tra essa e la sfera degenera $\{P_0\}$, e quindi avrà equazione

$$x^2 + y^2 + z^2 + \alpha x + \beta y + \gamma z + \delta = 0$$

tale che

$$(x^2 + y^2 + z^2 + \alpha x + \beta y + \gamma z + \delta) - ((x - 1)^2 + (y - 1)^2 + (z - 1)^2) = \lambda(x + y + z - 3),$$

cioè

$$x^2 + y^2 + z^2 + \alpha x + \beta y + \gamma z + \delta = (x - 1)^2 + (y - 1)^2 + (z - 1)^2 + \lambda(x + y + z - 3)$$

per un'opportuno $\lambda \in \mathbb{R}$.

Se, per esempio, vogliamo determinare la sfera tangente a π in P_0 e passante per $P_1 = (1, 0, 0)$ dobbiamo scegliere λ tale che

$$0^2 + (-1)^2 + (-1)^2 - 2\lambda(1 + 0 + 0 - 3) = 0,$$

ovvero $\lambda = 1$. Pertanto la sfera cercata ha equazione


$$x^2 + y^2 + z^2 - x - y - z = 0.$$

Il secondo metodo consiste nell'osservare che ogni sfera tangente a π in P_0 ha centro in un punto della retta per P_0 perpendicolare a π , cioè in un punto C_t avente coordinate $(1 + t, 1 + t, 1 + t)$, quindi ha equazione della forma

$$x^2 + y^2 + z^2 - 2(1 + t)x - 2(1 + t)y - 2(1 + t)z + 3 + 6t + t^2 - \varrho^2 = 0.$$

Poiché P_0 appartiene a tale sfera, si ha necessariamente $t^2 = \varrho^2$. A questo punto si osserva facilmente che tale equazione si può anche scrivere come

$$(x - 1)^2 + (y - 1)^2 + (z - 1)^2 + \lambda(x + y + z - 3) = 0$$

con $\lambda = -2t \in \mathbb{R}$, cioè la stessa equazione calcolata col primo metodo. 

Trasformazioni nello spazio e quadriche

25.1 Quadriche e loro riduzione a forma canonica

Fissiamo nello spazio un sistema di riferimento $Oxyz$ e consideriamo un polinomio $q(x, y, z)$ di grado 2 nelle tre variabili x, y, z a meno di costanti moltiplicative non nulle.

Sulla falsariga di quanto fatto per le coniche nel piano Oxy , vogliamo descrivere il luogo geometrico (eventualmente vuoto)

$$\mathcal{Q} = \{ P = (x, y, z) \mid q(x, y, z) = 0 \}.$$

Definizione 25.1 (Quadriche).

Una *quadrica* nello spazio (rispetto ad un fissato sistema di riferimento $Oxyz$) è il dato, a meno di costanti moltiplicative non nulle, di un'equazione della forma

$$\begin{aligned} a_{1,1}x^2 + a_{2,2}y^2 + a_{3,3}z^2 + 2a_{1,2}xy + 2a_{1,3}xz + 2a_{2,3}yz \\ + 2a_{1,4}x + 2a_{2,4}y + 2a_{3,4}z + a_{4,4} = 0 \end{aligned} \quad (25.1.1)$$

per opportuni $a_{i,j} \in \mathbb{R}$ non simultaneamente nulli.

Anche in questo caso spesso parleremo della *quadrica di equazione* (25.1.1). Come abbiamo fatto per le coniche, ad una quadrica \mathcal{Q} di equazione (25.1.1) possiamo associare due matrici simmetriche

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{1,2} & a_{2,2} & a_{2,3} \\ a_{1,3} & a_{2,3} & a_{3,3} \end{pmatrix}, \quad B = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{1,2} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{1,3} & a_{2,3} & a_{3,3} & a_{3,4} \\ a_{1,4} & a_{2,4} & a_{3,4} & a_{4,4} \end{pmatrix},$$

dette rispettivamente *matrice dei termini di secondo grado della quadrica* e *matrice (completa) della quadrica*.

Si noti che l'equazione di \mathcal{Q} si può scrivere come prodotto

$$(x \ y \ z \ 1) B \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = 0.$$

Come per le coniche, vorremmo trovare un sistema di riferimento $O'x'y'z'$ nello spazio in modo tale che l'equazione (25.1.1) divenga più semplice e, soprattutto, riconoscibile: per esempio potremmo desiderare d'avere un'equazione canonica nel senso della seguente definizione.

Definizione 25.2 (Forma canonica di una quadrica).

Nello spazio sia fissato un sistema di riferimento $Oxyz$ e sia \mathcal{Q} una quadrica. Diciamo che \mathcal{Q} è in *forma canonica* (o che $Oxyz$ è un *sistema di riferimento canonico* per \mathcal{Q}) se l'equazione di \mathcal{Q} è della forma

$$\alpha x^2 + \beta y^2 + \gamma z^2 = \delta, \quad (25.1.2)$$

oppure

$$\beta y^2 + \gamma z^2 = 2\delta x \quad (25.1.3)$$

per qualche $\alpha, \beta, \gamma, \delta \in \mathbb{R}$.

Vogliamo passare dal “vecchio riferimento” $Oxyz$ a un “nuovo riferimento” $O'x'y'z'$ rispetto a cui \mathcal{Q} sia in forma canonica tramite una *rototraslazione*, cioè una trasformazione di coordinate del tipo

$$\begin{cases} x = p_{1,1}x' + p_{1,2}y' + p_{1,3}z' + u \\ y = p_{2,1}x' + p_{2,2}y' + p_{2,3}z' + v \\ z = p_{3,1}x' + p_{3,2}y' + p_{3,3}z' + w, \end{cases} \quad (25.1.4)$$

che ci permetta di passare dall'equazione (25.1.1) ad una tra le equazioni (25.1.2) o (25.1.3).

La trasformazione (25.1.4) può essere pensata come composta dalle due trasformazioni

$$\begin{cases} x = \hat{x} + u \\ y = \hat{y} + v \\ z = \hat{z} + w, \end{cases}$$

che rappresenta una traslazione nello spazio, e

$$\begin{cases} \hat{x} = p_{1,1}x' + p_{1,2}y' + p_{1,3}z' \\ \hat{y} = p_{2,1}x' + p_{2,2}y' + p_{2,3}z' \\ \hat{z} = p_{3,1}x' + p_{3,2}y' + p_{3,3}z', \end{cases}$$

che, invece, vogliamo rappresenti una rotazione che fissi l'origine: anche in questo caso si può dimostrare che $P = (p_{i,j})_{i,j=1,2,3}$ è una matrice ortogonale speciale e che esistono tre angoli φ, ψ, ϑ (detti angoli di Eulero) tali che le entrate di P siano somme di opportuni prodotti di funzioni tipo coseno e seno di tali angoli.

La Proposizione 23.4 può essere facilmente generalizzata al caso delle quadriche come segue.

Proposizione 25.3.

Nel piano siano fissati due sistemi di riferimento $Oxyz$ e $O'x'y'z'$ per cui vale la relazione (25.1.4) e siano

$$Q = \begin{pmatrix} p_{1,1} & p_{1,2} & p_{1,3} & u \\ p_{2,1} & p_{2,2} & p_{2,3} & v \\ p_{3,1} & p_{3,2} & p_{3,3} & w \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad P = \begin{pmatrix} p_{1,1} & p_{1,2} & p_{1,3} \\ p_{2,1} & p_{2,2} & p_{2,3} \\ p_{3,1} & p_{3,2} & p_{3,3} \end{pmatrix}.$$

Sia poi \mathcal{Q} una quadrica avente nei due sistemi di riferimento matrici complete B e B' e matrici dei termini di grado 2 A ed A' rispettivamente. Allora $B' = {}^tQBQ$ e $A' = {}^tPAP$.

Definizione 25.4.

Sia \mathcal{Q} una quadrica rappresentata dalla matrice B . Definiamo *rango di \mathcal{Q}* come il rango della matrice B . La quadrica \mathcal{Q} si dice *non degenera* se il suo rango è 4, *degenera* in caso contrario. Se \mathcal{Q} è non degenera si dice *a centro* se il determinante della matrice del complesso dei termini di secondo grado è non nullo, *paraboloide* in caso contrario.

Osservazione 25.5. Poiché valgono le relazioni $B' = {}^tQBQ$ e $A' = {}^tPAP$, segue che $\text{rk}(B) = \text{rk}(B')$ e $\text{rk}(A) = \text{rk}(A')$. Concludiamo che il fatto che una quadrica \mathcal{Q} sia degenera o meno, oppure che sia a centro o meno, non dipende dal sistema di riferimento, ma è una proprietà geometrica.

Esempio 25.6. Si consideri la quadrica \mathcal{Q} di equazione

$$-2xz + y^2 - \sqrt{2}x - 2y - \sqrt{2}z - 1 = 0,$$

con matrici associate

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

Osserviamo $\det(A) = -1$, $\det(B) = 1$: quindi \mathcal{Q} è una quadrica a centro. In particolare è non degenera. ♠

Dalla trattazione risulta chiaro che, data una quadrica \mathcal{Q} di equazione (25.1.1), per determinare una sua equazione canonica si può procedere come nel caso delle coniche.

Si determina prima una matrice ortogonale speciale che diagonalizzi la matrice A del complesso dei termini di secondo grado. In questo modo l'equazione di \mathcal{Q} viene trasformata in una della forma

$$\widehat{a}_{1,1}\widehat{x}^2 + \widehat{a}_{2,2}\widehat{y}^2 + \widehat{a}_{3,3}\widehat{z}^2 + 2\widehat{a}_{1,4}\widehat{x} + 2\widehat{a}_{2,4}\widehat{y} + 2\widehat{a}_{3,4}\widehat{z} + \widehat{a}_{3,3} = 0.$$

Questa prima trasformazione è una rotazione intorno al punto O .

Poi con trasformazioni del tipo $(\widehat{x}, \widehat{y}, \widehat{z}) \mapsto (x' + a, y' + b, z' + c)$ si fa in modo che scompaiano il maggior numero possibile di monomi di grado 1 e, eventualmente, il termine noto. L'equazione risultante è in forma canonica ed è del tipo

$$\alpha x^2 + \beta y^2 + \gamma z^2 = \delta,$$

nel caso in cui $\det(A) \neq 0$, oppure del tipo

$$\beta y^2 + \gamma z^2 = 2\delta x,$$

nel caso in cui invece $\det(A) = 0$. Osserviamo che la seconda trasformazione è una traslazione che ci fa passare dal sistema di riferimento ausiliario $O\widehat{x}\widehat{y}\widehat{z}$ al nuovo sistema di riferimento $Oxyz$.

25.2 Classificazione delle quadriche

Per classificare e disegnare le quadriche si procede come nel caso delle coniche classiche, intersecando con opportuni piani paralleli ai piani coordinati.

Quadriche a centro.

Sia \mathcal{Q} una quadrica a centro con matrici associate A e B . Poiché $\det(A) \neq 0$ e $\det(B) \neq 0$, l'equazione canonica di \mathcal{Q} deve essere della forma

$$\alpha x^2 + \beta y^2 + \gamma z^2 = \delta$$

con $\alpha, \beta, \gamma, \delta \neq 0$. Si può sempre supporre che $\delta > 0$; di seguito analizziamo i vari casi possibili.

• Se $\alpha, \beta, \gamma > 0$, allora posto $a^2 = \delta/\alpha$, $b^2 = \delta/\beta$, $c^2 = \delta/\gamma$, possiamo sostituire l'equazione (25.1.2) con

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$

In tal caso, \mathcal{Q} risulta essere simmetrica sia rispetto all'origine O , che rispetto a qualsiasi asse coordinato e rispetto a qualsiasi piano coordinato.

Per determinare l'intersezione di \mathcal{Q} con i piani paralleli al piano yz dobbiamo studiare il sistema

$$\begin{cases} \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1. \\ x = k \end{cases}$$

al variare di $k \in \mathbb{R}$. Ci riduciamo perciò a studiare la corrispondente curva di livello

$$\frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 - \frac{k^2}{a^2}$$

nel piano yz , proiezione ortogonale su tale piano della conica intersezione di \mathcal{Q} con il piano di equazione $x = k$.

Con questo in mente, l'equazione di cui sopra rappresenta un'ellisse se $|k| < a$, un punto se $|k| = a$ ed un'ellisse immaginaria se $|k| > a$.

Procedendo similmente con gli altri piani coordinati, concludiamo che il luogo rappresentato da \mathcal{Q} è contenuto nel parallelepipedo $[-a, a] \times [-b, b] \times [-c, c]$: la quadrica \mathcal{Q} è detta *ellissoide* ed è raffigurata in Figura 25.1.

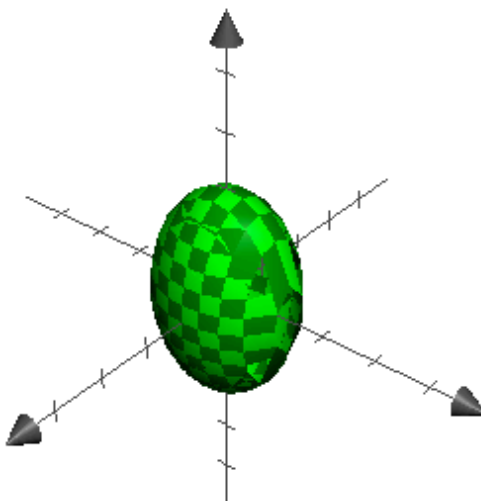


Figura 25.1

• Se $\alpha, \beta > 0$, $\gamma < 0$, allora posto $a^2 = \delta/\alpha$, $b^2 = \delta/\beta$, $c^2 = -\delta/\gamma$, possiamo sostituire l'equazione (25.1.2)

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1.$$

Come nel caso precedente, \mathcal{Q} è simmetrica rispetto all'origine O , rispetto a qualsiasi asse coordinato e rispetto a qualsiasi piano coordinato. Intersecando con i piani paralleli al piano yz otteniamo la curva di livello

$$\frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 - \frac{k^2}{a^2}$$

nel piano di equazione yz : tale conica è un'iperbole con asse reale coincidente con l'asse delle y se $|k| < a$, con asse delle z se $|k| > a$. Se $|k| = a$ la conica viene ad

essere una coppia iperbolica di rette. Un discorso simile si può fare intersecando \mathcal{Q} con piani paralleli al piano xz .

Invece, intersecando \mathcal{Q} con un piano parallelo al piano xy otteniamo la curva di livello

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 + \frac{k^2}{c^2}$$

nel piano di equazione xy : tale conica è sempre un'ellisse: la quadrica \mathcal{Q} è detta *iperboloide iperbolico* o *iperboloide ad una falda*, raffigurato in Figura 25.2.

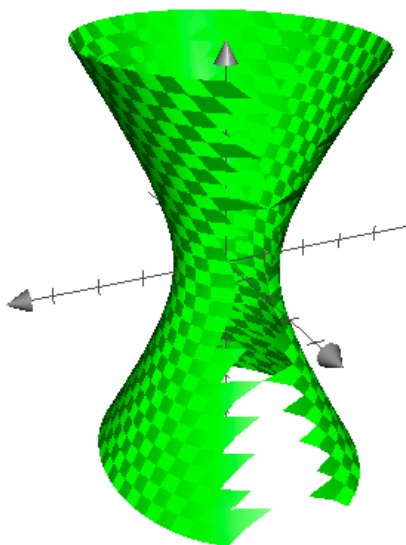


Figura 25.2

Si noti che l'identità

$$\frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 - \frac{x^2}{a^2}$$

ci permette di affermare che per ogni coppia di numeri reali non contemporaneamente nulli λ, μ si ha

$$\lambda \mu \left(\frac{y}{b} - \frac{z}{c} \right) \cdot \left(\frac{y}{b} + \frac{z}{c} \right) = \lambda \mu \left(1 - \frac{x}{a} \right) \cdot \left(1 + \frac{x}{a} \right),$$

quindi i punti dello spazio che soddisfano simultaneamente le equazioni di grado 1

$$\begin{cases} \lambda \left(\frac{y}{b} - \frac{z}{c} \right) = \mu \left(1 - \frac{x}{a} \right) \\ \mu \left(\frac{y}{b} + \frac{z}{c} \right) = \lambda \left(1 + \frac{x}{a} \right) \end{cases} \quad \begin{cases} \lambda \left(\frac{y}{b} - \frac{z}{c} \right) = \mu \left(1 + \frac{x}{a} \right) \\ \mu \left(\frac{y}{b} + \frac{z}{c} \right) = \lambda \left(1 - \frac{x}{a} \right) \end{cases}$$

soddisfano anche l'equazione di \mathcal{Q} . È facile verificare che tali sistemi descrivono due famiglie di rette nello spazio: in particolare \mathcal{Q} contiene infinite rette.

• Se $\alpha > 0$ e $\beta, \gamma < 0$, allora posto $a^2 = \delta/\alpha$, $b^2 = -\delta/\beta$, $c^2 = -\delta/\gamma$, possiamo sostituire l'equazione (25.1.2) con

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1.$$

Ancora \mathcal{Q} è simmetrica rispetto all'origine O , rispetto a qualsiasi asse coordinato e rispetto a qualsiasi piano coordinato. Intersecando con i piani paralleli al piano yz otteniamo la curva di livello

$$\frac{y^2}{b^2} + \frac{z^2}{c^2} = -1 + \frac{k^2}{a^2}$$

nel piano di equazione $x = k$: tale conica è un'ellisse se $|k| > a$, un'ellisse immaginaria se $|k| < a$. Se $|k| = a$ la conica viene ad essere una coppia ellittica di rette. In particolare \mathcal{Q} è esterna alla banda infinita $] -a, a[\times \mathbb{R} \times \mathbb{R}$ e consiste di due componenti connesse simmetriche.

Invece, intersecando \mathcal{Q} con un piano parallelo al piano xy otteniamo la curva di livello

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 + \frac{k^2}{c^2}$$

nel piano di equazione $z = k$: tale conica è sempre un'iperbole con asse reale coincidente con l'asse delle x . Un discorso simile si può fare intersecando \mathcal{Q} con piani paralleli al piano xz . La quadrica \mathcal{Q} è detta *iperboloide ellittico* o *iperboloide a due falde*, raffigurato in Figura 25.3.

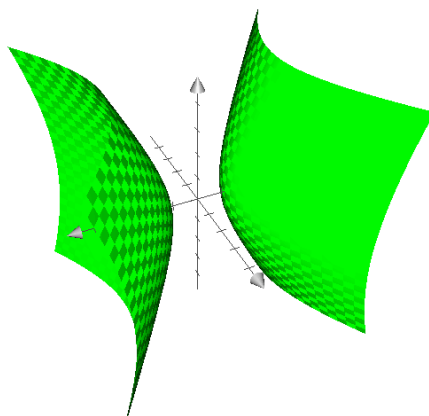


Figura 25.3

• Se $\alpha, \beta, \gamma < 0$, allora posto $a^2 = -\delta/\alpha$, $b^2 = -\delta/\beta$, $c^2 = -\delta/\gamma$, possiamo sostituire l'equazione (25.1.2)

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} + 1 = 0.$$

In tal caso, intersecando \mathcal{Q} con i piani paralleli al piano yz otteniamo la curva di livello

$$\frac{y^2}{b^2} + \frac{z^2}{c^2} = -1 - \frac{k^2}{a^2}$$

nel piano di equazione $x = k$, che è sempre un'ellisse immaginaria. La stessa cosa accade intersecando con gli altri piani coordinati: \mathcal{Q} è detta *ellissoide immaginario*.

Paraboloidi.

Sia \mathcal{Q} un paraboloido con matrici associate A e B . Poiché $\det(A) = 0$, ma $\det(B) \neq 0$, l'equazione canonica di \mathcal{Q} deve essere della forma

$$\beta y^2 + \gamma z^2 = 2\delta x$$

con $\beta, \gamma, \delta \neq 0$. Si può sempre supporre che $\delta > 0$; di seguito analizziamo i vari casi possibili.

• Se $\beta, \gamma > 0$, allora, $b^2 = \delta/\beta$, $c^2 = \delta/\gamma$, possiamo sostituire l'equazione (25.1.3) con

$$\frac{y^2}{b^2} + \frac{z^2}{c^2} = 2x.$$

\mathcal{Q} è simmetrica rispetto all'asse delle ascisse e rispetto ai piani coordinati xy e xz . Intersecandola con i piani di equazione $x = k$, al variare di $k \in \mathbb{R}$, ci riduciamo a studiare la curva di livello

$$\frac{y^2}{b^2} + \frac{z^2}{c^2} = 2k$$

nel piano di equazione $x = k$: tale conica è un'ellisse se $|k| > 0$, un'ellisse immaginaria se $|k| < 0$. Se $|k| = 0$ la conica viene ad essere una coppia ellittica di rette. In particolare \mathcal{Q} è contenuta nel semispazio $[0, +\infty[\times \mathbb{R} \times \mathbb{R}$.

Intersecando \mathcal{Q} con il piano xy otteniamo invece

$$\frac{y^2}{b^2} = 2x,$$

che è una parabola con asse coincidente con l'asse delle ascisse. Un discorso simile si può fare intersecando \mathcal{Q} con gli altri piani passanti per l'asse delle ascisse di equazione $z = \lambda y$, ottenendo

$$\frac{c^2 + \lambda^2 b^2}{b^2 c^2} y^2 = 2x,$$

che è ancora una parabola. La quadrica \mathcal{Q} è detta *paraboloido ellittico*, raffigurato in Figura 25.4.

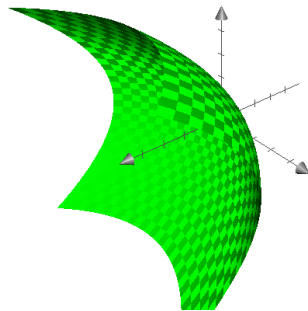


Figura 25.4

• Se vale $\beta > 0$ e $\gamma < 0$, allora posto $b^2 = \delta/\beta$ e $c^2 = -\delta/\gamma$, possiamo sostituire l'equazione (25.1.3) con

$$\frac{y^2}{b^2} - \frac{z^2}{c^2} = 2x.$$

\mathcal{Q} è simmetrica rispetto all'asse delle ascisse e rispetto ai piani coordinati xy e xz . Intersecando con i piani paralleli al piano yz otteniamo la curva di livello

$$\frac{y^2}{b^2} - \frac{z^2}{c^2} = 2k$$

nel piano di equazione $x = k$: tale conica è un'iperbole con asse reale coincidente con l'asse delle ordinate se $|k| > 0$, un'iperbole con asse reale coincidente con l'asse delle quote se $|k| < 0$. Se $|k| = 0$ la conica viene ad essere una coppia iperbolica di rette.

Intersecando \mathcal{Q} con piani contenenti l'asse delle ascisse otteniamo delle parabole il cui asse ancora coincide con l'asse delle ascisse. \mathcal{Q} in questo caso è detta *paraboloide iperbolico*, raffigurato in Figura 25.5.

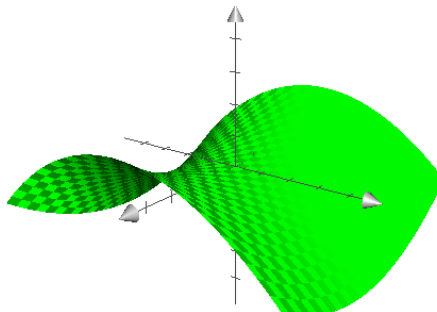


Figura 25.5

Con la stessa procedura descritta nel caso dell'iperboloide ellittico, si verifica l'identità

$$\frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 - \frac{x^2}{a^2}$$

ci permette di affermare che per ogni coppia di numeri reali non contemporaneamente nulli λ, μ si ha

$$\lambda\mu \left(\frac{y}{b} - \frac{z}{c} \right) \cdot \left(\frac{y}{b} + \frac{z}{c} \right) = \lambda\mu \left(1 - \frac{x}{a} \right) \cdot \left(1 + \frac{x}{a} \right),$$

quindi i punti dello spazio che soddisfano simultaneamente le equazioni di grado 1

$$\begin{cases} \lambda \left(\frac{y}{b} - \frac{z}{c} \right) = 2\mu \\ \mu \left(\frac{y}{b} + \frac{z}{c} \right) = \lambda x \end{cases} \quad \begin{cases} \lambda \left(\frac{y}{b} - \frac{z}{c} \right) = \mu x \\ \mu \left(\frac{y}{b} + \frac{z}{c} \right) = 2\lambda \end{cases}$$

soddisfano anche l'equazione di \mathcal{Q} per ogni scelta di $\lambda, \mu \in \mathbb{R}$ non simultaneamente nulli. In particolare \mathcal{Q} contiene infinite rette.

Quadriche degeneri.

Ci limiteremo a considerare le quadriche degeneri di rango 3.

Coni.

Consideriamo prima il caso

$$\alpha x^2 + \beta y^2 + \gamma z^2 = 0,$$

con $\alpha, \beta, \gamma \neq 0$ (la quadrica allora è detta *cono*).

• Se $\alpha, \beta > 0$ e $\gamma < 0$, \mathcal{Q} è simmetrica rispetto all'asse delle quote e rispetto ai piani coordinati xz e yz . Intersecando con i piani paralleli al piano xy otteniamo la curva di livello

$$\alpha x^2 + \beta y^2 = -\gamma k^2 :$$

tale conica è un'ellisse se $k \neq 0$, una coppia ellittica di rette se $k = 0$. Invece intersecando \mathcal{Q} con il piano xz otteniamo

$$\alpha x^2 + \gamma z^2 = 0,$$

che è una coppia iperbolica di rette passanti per l'origine O ($\gamma < 0$).

Un discorso simile si può fare intersecando \mathcal{Q} con gli altri piani passanti per l'asse delle quote di equazione $x = \lambda y$, ottenendo

$$(\alpha + \beta\lambda^2)x^2 + \gamma z^2 = 0,$$

che è ancora una coppia iperbolica di rette passanti per l'origine O . Il cono \mathcal{Q} è raffigurato in Figura 25.6.

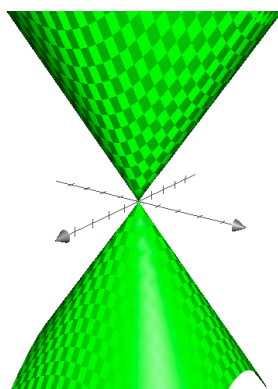


Figura 25.6

Si osservi che in questo caso \mathcal{Q} può essere pensata come l'unione delle rette passanti per l'origine che si appoggiano all'ellisse del piano $z = 1$

$$\alpha x^2 + \beta y^2 = -\gamma.$$

• Se $\alpha, \beta, \gamma > 0$, intersecando \mathcal{Q} con i piani paralleli al piano xy otteniamo la curva di livello

$$\alpha x^2 + \beta y^2 = -\gamma k^2 :$$

tale conica è un'ellisse immaginaria se $k \neq 0$, una coppia ellittica di rette se $k = 0$. La quadrica \mathcal{Q} è detta *cono immaginario*.

Cilindri ellittici o iperbolici.

In questo caso l'equazione è della forma

$$\beta y^2 + \gamma z^2 = \delta,$$

con $\beta, \gamma, \delta \neq 0$ (in questo parliamo di *cilindro ellittico* o *cilindro iperbolico*).

• Se $\beta, \gamma, \delta > 0$, allora \mathcal{Q} è simmetrica rispetto alla retta di equazioni $y = z = 0$ (cioè all'asse delle ascisse). Intersecando con i piani paralleli al piano yz di equazione $x = k$ otteniamo

$$\beta y^2 + \gamma z^2 = \delta :$$

tale conica è sempre un'ellisse, qualsiasi sia $k \in \mathbb{R}$, e \mathcal{Q} viene detta *cilindro ellittico*, raffigurato in Figura 25.7.

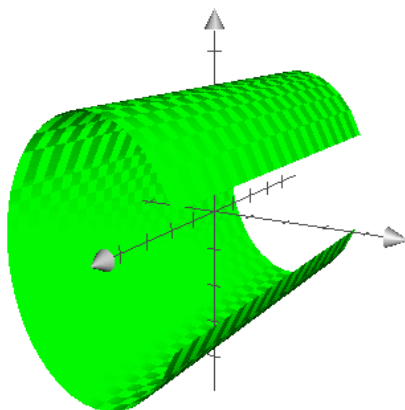


Figura 25.7

In questo caso \mathcal{Q} può essere pensata come l'unione delle rette passanti parallele all'asse delle ascisse che si appoggiano all'ellisse del piano $x = 0$

$$\beta y^2 + \gamma z^2 = \delta.$$

• Se $\beta, \gamma > 0$ e $\delta < 0$, allora intersecando \mathcal{Q} con piani paralleli al piano yz otteniamo la curva di livello

$$\beta y^2 + \gamma z^2 = \delta :$$

tale conica è sempre un'ellisse immaginaria. \mathcal{Q} è un *cilindro ellittico immaginario*.

• Se $\beta, \delta > 0$ e $\gamma < 0$, \mathcal{Q} è ancora simmetrica rispetto alla retta di equazioni $y = z = 0$ e, intersecandola con i piani di equazione $x = k$, otteniamo ancora la curva di livello

$$\beta y^2 + \gamma z^2 = \delta,$$

che, in questo caso, è sempre un'iperbole; \mathcal{Q} viene detta *cilindro iperbolico*, raffigurato in Figura 25.8.

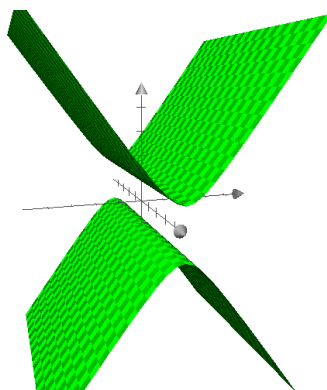


Figura 25.8

In questo caso \mathcal{Q} è l'unione delle rette passanti parallele all'asse delle ascisse che si appoggiano all'iperbole del piano $x = 0$

$$\beta y^2 + \gamma z^2 = \delta.$$

Cilindri parabolici.

In questo caso l'equazione è della forma

$$\gamma z^2 = 2\delta x,$$

con $\gamma, \delta \neq 0$ (in questo parliamo di *cilindro parabolico*).

• \mathcal{Q} è simmetrica rispetto al piano di equazione $z = 0$ (cioè al piano xy) e intersecando con i piani paralleli al piano xz otteniamo la curva di livello

$$\gamma z^2 = 2\delta x :$$

tale conica è sempre una parabola, qualsiasi sia $k \in \mathbb{R}$, e la quadrica \mathcal{Q} è un *cilindro parabolico*, raffigurato in Figura 25.9.

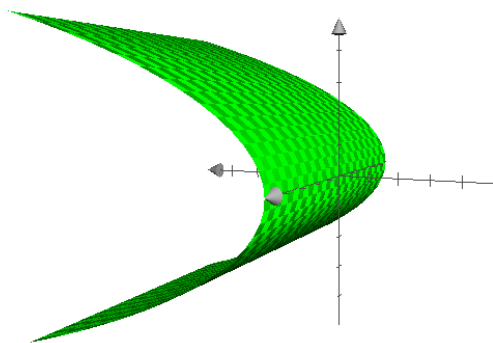


Figura 25.9

Come nei casi precedenti, \mathcal{Q} è l'unione delle rette parallele alla retta $x = z = 0$ (cioè all'asse delle ordinate) che si appoggiano alla parabola del piano $y = 0$

$$\gamma z^2 = 2\delta x.$$

Quadriche riducibili.

Per concludere la nostra lista, rimangono da analizzare le quadriche di rango al più 2, per le quali però non entreremo nei dettagli, limitandoci ad enunciare il seguente risultato.

Proposizione 25.7. *Sia \mathcal{Q} una quadrica con matrice completa associata B . La quadrica \mathcal{Q} si decompone in un prodotto di due polinomi di grado 1 (non necessariamente distinti) se e solo se $\text{rk}(B) \leq 2$.*