

# Geometria affine

# 10

## Sommario

- 10.1** Equazioni di rette e piani
- 10.2** Punti e rette
- 10.3** Punti e piani
- 10.4** Rette e rette
- 10.5** Rette e piani
- 10.6** Piani e piani
- 10.7** Sistemi di riferimento affini
- 10.8** Orientazione

## Complementi

- 10C.1** Affinità
- 10C.2** Geometria proiettiva

## Esercizi

In questo capitolo torniamo a occuparci della geometria del piano e dello spazio. Vedremo come identificare le posizioni reciproche di punti, rette e piani usando sia equazioni parametriche sia cartesiane: si tratterà di applicare in questo caso particolare la teoria sviluppata finora. Discuteremo anche i cambiamenti di coordinate affini e introdurremo il concetto di orientazione. Infine, i complementi a questo capitolo sono dedicati alle affinità e ad alcuni cenni di geometria proiettiva.

## 10.1 Equazioni di rette e piani

Vogliamo applicare quanto visto finora allo studio della geometria affine del piano e dello spazio. La *geometria affine*<sup>(1)</sup> si occupa delle posizioni reciproche di rette e piani (e triangoli, pentagoni...) nello spazio  $A^3$  e nel piano  $A^2$ . Il nostro scopo è dare criteri operativi per risolvere problemi quali “decidere quando due rette sono sghembe”, “trovare il piano passante per tre punti non allineati”, e simili.

Fissiamo una volta per tutte un sistema di riferimento affine  $RA(O, A_1, A_2, A_3)$  in  $A^3$  in modo da poter identificare  $A^3$  con  $\mathbb{R}^3$  tramite le coordinate (come indicato nel Paragrafo 2.2). In questo modo punti, rette e piani di  $A^3$  diventano sottospazi affini di  $\mathbb{R}^3$  di dimensione 0, 1 o 2 rispettivamente. Per semplicità tipografica in questo capitolo le coordinate dei punti di  $\mathbb{R}^2$  ed  $\mathbb{R}^3$  verranno talvolta scritte per riga invece che per colonna  $(1, 2) \in \mathbb{R}^2$ , o  $(x_0, y_0, z_0) \in \mathbb{R}^3$  e così via.

Come visto nel Paragrafo 6.5, un sottospazio affine  $L$  di  $\mathbb{R}^3$  può venire rappresentato tramite equazioni cartesiane o tramite equazioni parametriche. Per comodità di lettura riportiamo qui le definizioni adattate alla situazione attuale.

### Definizione 10.1

Se  $L$  è lo spazio delle soluzioni del sistema

$$Ax = b, \quad (10.1)$$

dove  $A \in M_{p,3}(\mathbb{R})$  e  $b \in \mathbb{R}^p$ , diremo che (10.1) è un’equazione cartesiana per  $L$ . La dimensione di  $L$  è  $3 - \text{rg } A$ , e il sottospazio di giacitura di  $L$  è lo spazio  $V$  delle soluzioni del sistema omogeneo  $Ax = O$ .

Come già discusso nei Paragrafi 6.4 e 6.5, di solito si scelgono le equazioni cartesiane in modo che  $\text{rg } A = p$ , cioè in modo che non vi siano equazioni inutili (il che equivale a richiedere che  $L_A$  sia surgettiva, o che il sistema  $Ax = b$  abbia soluzione qualunque  $b \in \mathbb{R}^p$ ). Quindi un piano nello spazio sarà descritto da un’equazione e una retta da due. Ma vediamolo in dettaglio.

a) *Equazione cartesiana di un piano nello spazio:* l’equazione

$$ax + by + cz = d \quad (10.2)$$

rappresenta un piano in  $\mathbb{R}^3$ , ammesso che la matrice  $A = [a \ b \ c]$  abbia rango 1 (il che equivale a richiedere che almeno uno dei coefficienti  $a, b, c$  sia non nullo). I coefficienti  $a, b$  e  $c$  si dicono *parametri di giacitura* del piano. I parametri di giacitura non sono univocamente determinati; l’equazione  $(\lambda a)x + (\lambda b)y + (\lambda c)z = (\lambda d)$  identifica lo stesso piano qualunque sia  $\lambda \in \mathbb{R}$  non nullo (vedi anche l’Esercizio 10.1).

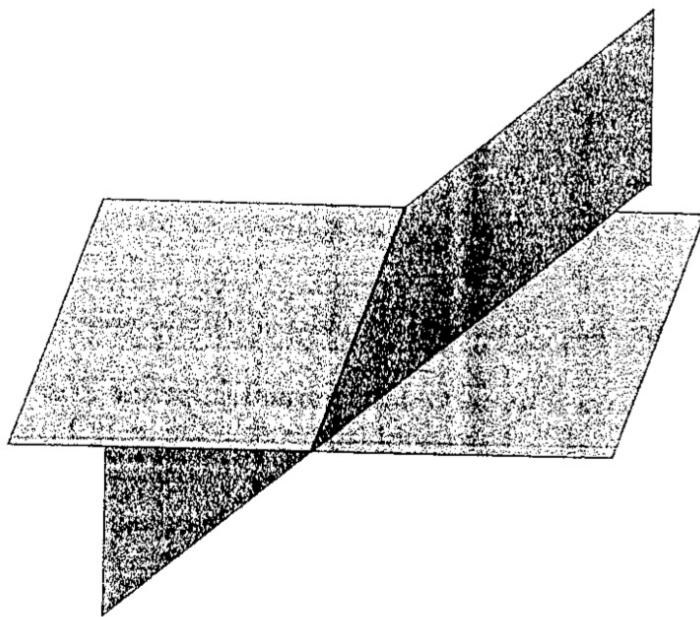
b) *Equazioni cartesiane di una retta nello spazio:* il sistema

$$\begin{cases} ax + by + cz = d, \\ a'x + b'y + c'z = d', \end{cases} \quad (10.3)$$

rappresenta una retta in  $\mathbb{R}^3$ , non appena la matrice

$$A = \begin{vmatrix} a & b & c \\ a' & b' & c' \end{vmatrix}$$

<sup>(1)</sup> L’aggettivo “affine” serve a distinguerla dalla geometria “metrica” di rette e piani che affrontiamo nel Capitolo 12, e dalla geometria “proiettiva” discussa nei complementi. Vedi anche Paragrafo 10C.1 per un’ulteriore precisazione del termine “affine”.



**Figura 10.1** - Una retta come intersezione di due piani.

ha rango 2. In questo modo abbiamo rappresentato la retta come intersezione di due piani (vedi la Figura 10.1): il piano  $ax + by + cz = d$ , e il piano  $a'x + b'y + c'z = d'$ . Come vedremo nel Paragrafo 10.6, la condizione  $\text{rg } A = 2$  equivale a richiedere che i due piani non siano paralleli. Inutile dire che la matrice  $A$  si guarda bene dall'essere unica: una retta è intersezione di infinite coppie distinte di piani.

### Osservazione 10.1

Il numero minimo di equazioni necessarie per descrivere un sottospazio affine è dato dalla dimensione dello spazio ambiente meno la dimensione del sottospazio. Dunque per descrivere una retta (dimensione 1) nel piano (dimensione 2) basta un'equazione (e non ne servono due, come accadeva nello spazio). Quindi nel piano  $\mathbb{R}^2$  l'equazione

$$ax + by = c \quad (10.4)$$

rappresenta (ammesso che almeno uno dei due coefficienti  $a$  e  $b$  sia diverso da zero) una retta. In particolare, *l'insieme rappresentato da equazioni cartesiane dipende dallo spazio ambiente*. In  $\mathbb{R}^2$ , l'equazione (10.4) rappresenta una retta; in  $\mathbb{R}^3$ , la stessa equazione rappresenta un piano.

Passiamo ora alle equazioni parametriche.

### Definizione 10.2

Se  $L$  è l'insieme dei punti  $P \in \mathbb{R}^3$  della forma

$$P = P_0 + Bt, \quad (10.5)$$

al variare di  $t \in \mathbb{R}^q$ , dove  $P_0 \in \mathbb{R}^3$  e  $B \in M_{3,q}(\mathbb{R})$ , diremo che (10.5) è un'*equazione parametrica* per  $L$ . Il sottospazio  $L$  ha dimensione  $\text{rg } B$ , e il sottospazio vettoriale  $V$  di giacitura di  $L$  ha equazione parametrica  $P = Bt$ .

Anche le equazioni parametriche vengono di solito scelte in modo che  $\text{rg } B = q$ , cioè in modo che non ci siano variabili inutili (il che equivale a richiedere che  $L_B$  sia iniettiva, cioè che ogni punto del sottospazio affine corrisponda a uno e un solo valore dei para-

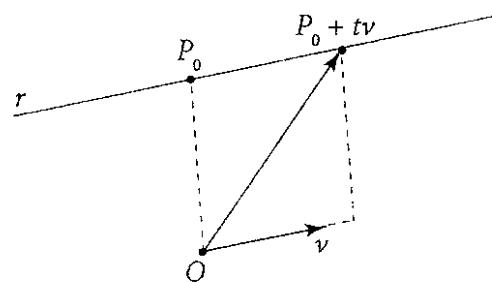


Figura 10.2 - Equazione parametrica di una retta nello spazio.

metri). Dunque questa volta il numero dei parametri dipende dalla dimensione del sottospazio, mentre il numero delle equazioni (tre in  $\mathbb{R}^3$ , due in  $\mathbb{R}^2$ ) dipende dall'ambiente. Ma vediamolo in dettaglio.

a) *Equazioni parametriche di una retta nello spazio:* le equazioni

$$\begin{cases} x = x_0 + lt, \\ y = y_0 + mt, \\ z = z_0 + nt, \end{cases} \quad \text{oppure} \quad P = P_0 + t \begin{vmatrix} l \\ m \\ n \end{vmatrix} \quad (10.6)$$

al variare di  $t \in \mathbb{R}$  rappresentano una retta nello spazio, non appena il vettore

$$v = \begin{vmatrix} l \\ m \\ n \end{vmatrix},$$

detto *vettore direttore* della retta, non è il vettore nullo (vedi la Figura 10.2). Le tre coordinate  $l$ ,  $m$  e  $n$  del vettore direttore si chiamano *parametri direttori* della retta. Chiaramente, non sono univocamente determinati; i vettori  $v$  e  $\lambda v$  sono vettori direttori della stessa retta, qualunque sia  $\lambda \in \mathbb{R}$  non nullo. Anche il punto  $P_0 = (x_0, y_0, z_0)$  si guarda bene dall'essere univocamente determinato: può essere sostituito da un qualunque altro punto della retta (vedi per esempio l'Esercizio 10.3).

b) *Equazioni parametriche di un piano nello spazio:* le equazioni

$$\begin{cases} x = x_0 + ls + l't, \\ y = y_0 + ms + m't, \\ z = z_0 + ns + n't, \end{cases} \quad \text{oppure} \quad P = P_0 + s \begin{vmatrix} l \\ m \\ n \end{vmatrix} + t \begin{vmatrix} l' \\ m' \\ n' \end{vmatrix} \quad (10.7)$$

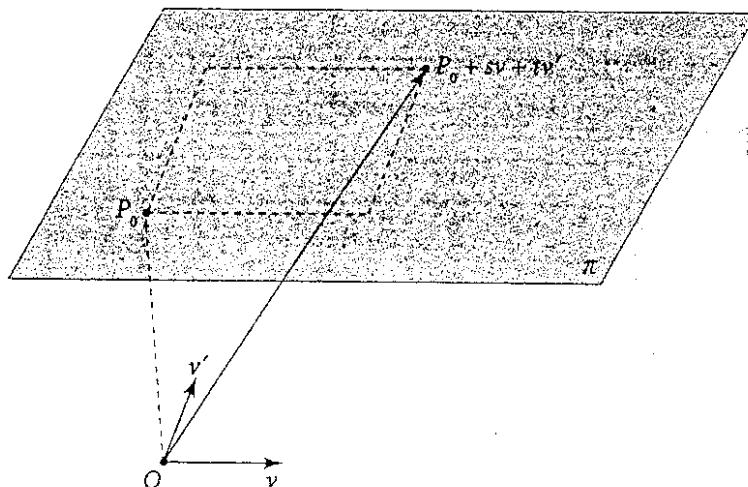


Figura 10.3 - Equazione parametrica di un piano nello spazio.

al variare di  $s, t \in \mathbb{R}$  rappresentano un piano nello spazio, non appena la matrice

$$B = \begin{vmatrix} l & l' \\ m & m' \\ n & n' \end{vmatrix}$$

ha rango 2 ovvero ha colonne linearmente indipendenti (vedi la Figura 10.3). I due vettori

$$\nu = \begin{vmatrix} l \\ m \\ n \end{vmatrix} \quad \text{e} \quad \nu' = \begin{vmatrix} l' \\ m' \\ n' \end{vmatrix},$$

che formano una base del sottospazio di giacitura del piano, sono detti *vettori di giacitura*. Non sono unici: possono essere sostituiti da una qualunque altra base del sottospazio di giacitura, e  $P_0$  può essere un qualunque punto del piano.

### Osservazione 10.2

Le equazioni parametriche di una retta nel piano non sono molto diverse da quelle di una retta nello spazio; semplicemente, sono due invece di tre. Per la precisione, sono della forma

$$\begin{cases} x = x_0 + lt, \\ y = y_0 + mt, \end{cases} \quad \text{oppure} \quad P = P_0 + t \begin{vmatrix} l \\ m \end{vmatrix}$$

al variare di  $t \in \mathbb{R}$ , sempre supponendo che il vettore  $\nu = \begin{vmatrix} l \\ m \end{vmatrix}$ , detto *vettore direttore* della retta, non sia il vettore nullo.

Tecniche per passare da equazioni cartesiane a equazioni parametriche e viceversa nel caso di rette e piani sono descritte negli Esercizi 10.4-10.7.

## 10.2 Punti e rette

Dopo aver rivisto come rappresentare punti, rette e piani, passiamo a discutere in che posizioni reciproche si possono trovare.

(A) *Una retta e un punto*. Iniziamo col caso del piano. Dato un punto  $P_0 = (x_0, y_0)$  in  $\mathbb{R}^2$ , vogliamo tutte le rette che passano per quel punto. Le equazioni parametriche sono immediate: se  $P = P_0 + tv$  sono equazioni parametriche per la retta  $r$ , ponendo  $t = 0$  vediamo che  $P_0$  appartiene alla retta. Quindi le rette che passano per  $P_0$  sono tutte e sole quelle di equazioni

$$\begin{cases} x = x_0 + tl, \\ y = y_0 + tm, \end{cases} \tag{10.8}$$

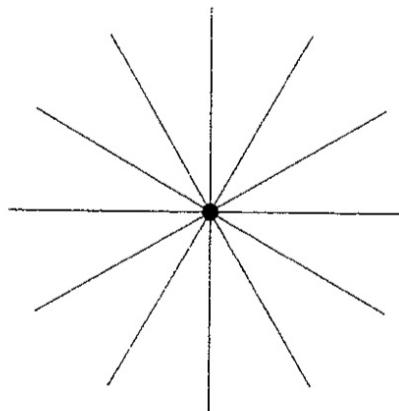
quali che siano  $l, m \in \mathbb{R}$  non entrambi nulli.

Anche le equazioni cartesiane si trovano facilmente. Vogliamo che  $P_0$  appartenga alla retta (generica) di equazione  $ax + by = c$ , cioè che le coordinate di  $P_0$  siano soluzione dell'equazione  $ax + by = c$ . Quindi  $ax_0 + by_0 = c$ , per cui l'equazione cartesiana della generica retta per  $P_0$  è

$$a(x - x_0) + b(y - y_0) = 0. \tag{10.9}$$

**Definizione 10.3**

L'insieme delle rette del piano passanti per  $P_0$  si chiama *fascio di rette* di centro  $P_0$  (vedi la Figura 10.4).



**Figura 10.4** - Fascio di rette.

Passiamo ora allo spazio. Dato un punto  $P_0 = (x_0, y_0, z_0) \in \mathbb{R}^3$ , vogliamo tutte le rette che passano per quel punto. Le equazioni parametriche si ottengono esattamente come prima, e sono

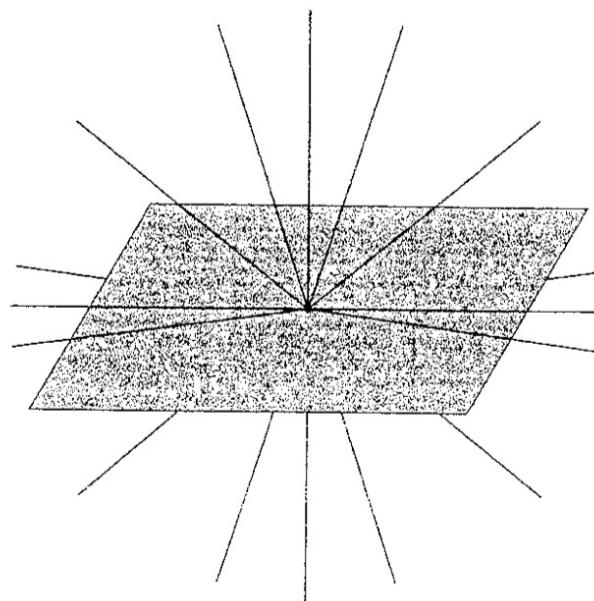
$$\begin{cases} x = x_0 + tl, \\ y = y_0 + tm, \\ z = z_0 + tn, \end{cases} \quad (10.10)$$

quali che siano  $l, m, n \in \mathbb{R}$  non tutti nulli. In modo analogo si trovano le equazioni cartesiane

$$\begin{cases} a(x - x_0) + b(y - y_0) + c(z - z_0) = 0, \\ a'(x - x_0) + b'(y - y_0) + c'(z - z_0) = 0. \end{cases} \quad (10.11)$$

**Definizione 10.4**

L'insieme delle rette dello spazio passanti per  $P_0$  si chiama *stella di rette* di centro  $P_0$  (vedi la Figura 10.5).



**Figura 10.5** - Stella di rette.

(B) *Una retta e due punti.* Com'è noto, per due punti distinti passa una e una sola retta; vogliamo trovarla. Cominciamo di nuovo nel piano. Vogliamo la retta che passa per i punti distinti  $P_0 = (x_0, y_0)$  e  $P_1 = (x_1, y_1) \in \mathbb{R}^2$ . Usando le equazioni parametriche dobbiamo trovare  $l, m \in \mathbb{R}$  tali che il sistema nell'incognita  $t$

$$\begin{cases} x_1 = x_0 + tl, \\ y_1 = y_0 + tm, \end{cases} \quad (10.12)$$

abbia soluzione. In questo caso  $l$  ed  $m$  si trovano a occhio (vero?), ma visto che utilizzeremo lo stesso procedimento molte volte in futuro vediamo di risolvere il problema una volta per tutte. Abbiamo un sistema lineare di cui vogliamo soltanto sapere se ha soluzione. È quindi naturale utilizzare il Teorema di Rouché-Capelli (Corollario 5.9), che dice che il sistema (10.12) ha soluzione se e solo se

$$1 = \operatorname{rg} \begin{vmatrix} l & x_1 - x_0 \\ m & y_1 - y_0 \end{vmatrix} = \operatorname{rg} \begin{vmatrix} l & x_1 - x_0 \\ m & y_1 - y_0 \end{vmatrix}.$$

Questo accade se e solo se la matrice sulla destra è singolare, cioè se e solo se

$$l(y_1 - y_0) - m(x_1 - x_0) = 0 \quad (10.13)$$

(è il determinante della matrice). Una soluzione di questa equazione (tutte le altre sono dei multipli, per cui ci forniscono la stessa retta) è  $l = x_1 - x_0$  e  $m = y_1 - y_0$ . Dunque la retta passante per  $P_0$  e  $P_1$  ha equazioni parametriche

$$\begin{cases} x = x_0 + (x_1 - x_0)t, \\ y = y_0 + (y_1 - y_0)t. \end{cases} \quad (10.14)$$

### Osservazione 10.3

La tecnica utilizzata per risolvere questo problema può essere riassunta come segue: la condizione posta (passaggio per il punto  $P_1$ ) si traduce in un sistema dipendente da parametri ( $l$  e  $m$ ) che deve avere soluzione; il Teorema di Rouché-Capelli fornisce allora una condizione necessaria e sufficiente in termini di ranghi di matrici; a sua volta questa condizione può essere espressa in termini di determinanti, e quindi in un certo numero di equazioni lineari nei parametri. La soluzione di queste ultime equazioni fornisce la soluzione del nostro problema.

Vediamo ora le equazioni cartesiane della retta per due punti. Richiedendo che la generica retta del fascio (10.9) passi per  $P_1$  otteniamo  $a(x_1 - x_0) + b(y_1 - y_0) = 0$ ; risolvendo troviamo che la retta passante per  $P_0$  e  $P_1$  ha equazione cartesiana

$$(y_1 - y_0)(x - x_0) - (x_1 - x_0)(y - y_0) = 0. \quad (10.15)$$

Passiamo ora ad  $\mathbb{A}^3$ ; vogliamo trovare equazioni della retta passante per i punti  $P_0 = (x_0, y_0, z_0)$  e  $P_1 = (x_1, y_1, z_1) \in \mathbb{R}^3$ . Usando le equazioni parametriche, vogliamo che il sistema

$$\begin{cases} x_1 = x_0 + tl, \\ y_1 = y_0 + tm, \\ z_1 = z_0 + tn, \end{cases}$$

abbia soluzione. Per il Teorema di Rouché-Capelli questo accade se e solo se

$$\operatorname{rg} \begin{vmatrix} l & x_1 - x_0 \\ m & y_1 - y_0 \\ n & z_1 - z_0 \end{vmatrix} = 1,$$

che quando  $x_1 \neq x_0$ , grazie al Teorema 9.13 degli orlati, avviene se (quando  $x_1 = x_0$  vedi l'Esercizio 10.15)

$$\begin{cases} l(y_1 - y_0) - m(x_1 - x_0) = 0, \\ l(z_1 - z_0) - n(x_1 - x_0) = 0. \end{cases} \quad (10.16)$$

Questo sistema ha soluzione  $l = x_1 - x_0$ ,  $m = y_1 - y_0$  e  $n = z_1 - z_0$  (o multipli), per cui la retta per  $P_0$  e  $P_1$  (vedi la Figura 10.6) ha equazioni parametriche

$$\begin{cases} x = x_0 + (x_1 - x_0)t, \\ y = y_0 + (y_1 - y_0)t, \\ z = z_0 + (z_1 - z_0)t. \end{cases} \quad (10.17)$$

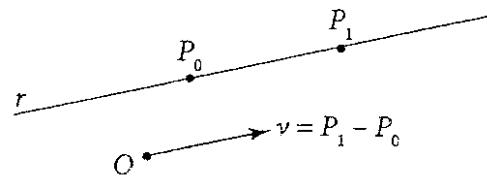


Figura 10.6 - La retta per due punti.

Per le equazioni cartesiane conviene procedere in modo un po' diverso, evitando di ricorrere al sistema (10.11) in quanto dipendente da troppi parametri. Sappiamo che un punto  $P = (x, y, z) \in \mathbb{R}^3$  appartiene alla retta passante per  $P_0$  e  $P_1$  se e solo se il vettore  $P - P_0$  è un multiplo del vettore  $P_1 - P_0$ ; questo è un altro modo di esprimere il sistema (10.17). In termini di matrici, questo accade se e solo se

$$\operatorname{rg} \begin{vmatrix} x - x_0 & x_1 - x_0 \\ y - y_0 & y_1 - y_0 \\ z - z_0 & z_1 - z_0 \end{vmatrix} = 1.$$

Usando il Teorema 9.13 degli orlati a partire da una coordinata non nulla di  $P_1 - P_0$  otteniamo le due equazioni cartesiane cercate. Per esempio, se  $x_1 \neq x_0$  si trova

$$\begin{cases} (y_1 - y_0)(x - x_0) - (x_1 - x_0)(y - y_0) = 0, \\ (z_1 - z_0)(x - x_0) - (x_1 - x_0)(z - z_0) = 0. \end{cases} \quad (10.18)$$

### Esempio 10.1

Vogliamo equazioni parametriche e cartesiane della retta passante per  $P_0 = (1, 0, 1)$  e  $P_1 = (1, 1, 2)$ . Equazioni parametriche sono date da (10.17):

$$\begin{cases} x = 1, \\ y = t, \\ z = 1 + t. \end{cases}$$

Per le equazioni cartesiane, non possiamo utilizzare (10.18) in quanto  $x_1 = 1 = x_0$ . Ma  $y_1 - y_0 = 1 \neq 0$ ; quindi il Teorema 9.13 degli orlati ci fornisce come equazioni cartesiane

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

□

## 10.3 Punti e piani

Passiamo a vedere le relazioni fra punti e piani – ovviamente, solo nello spazio.

(C) *Un piano e un punto.* Dato un punto  $P_0 = (x_0, y_0, z_0) \in \mathbb{R}^3$  vogliamo tutti i piani che passano per quel punto. Si procede esattamente come per le rette, per cui troviamo

$$\begin{cases} x = x_0 + sl + tl', \\ y = y_0 + sm + tm', \\ z = z_0 + sn + tn', \end{cases} \quad (10.19)$$

come equazioni parametriche, mentre come equazione cartesiana abbiamo

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0. \quad (10.20)$$

### Definizione 10.5

L'insieme dei piani passanti per  $P_0$  si chiama *stella di piani* di centro  $P_0$  (vedi la Figura 10.7).

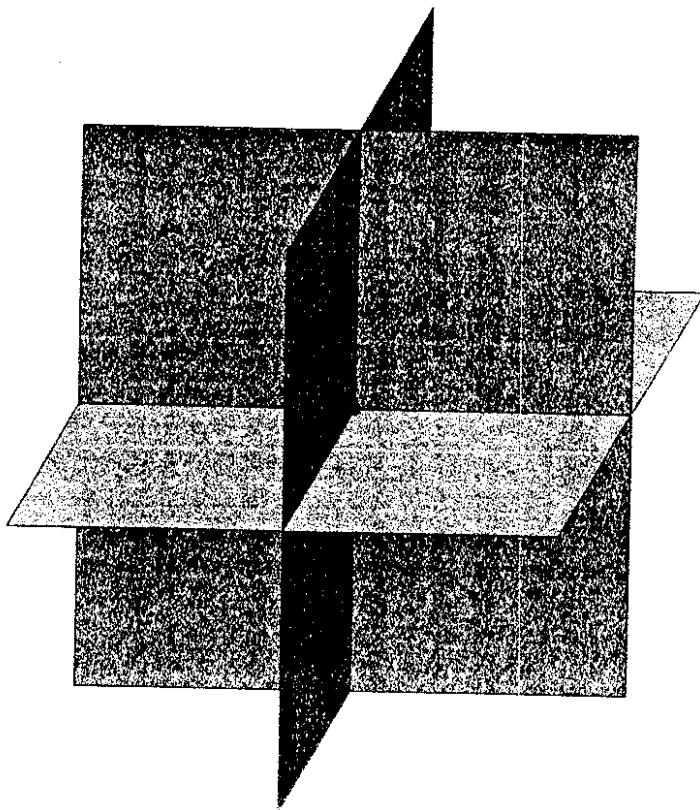


Figura 10.7 - Stella di piani.

(D) *Un piano e due punti.* Dati i due punti  $P_0 = (x_0, y_0, z_0)$  e  $P_1 = (x_1, y_1, z_1) \in \mathbb{R}^3$  con  $P_0 \neq P_1$ , vogliamo tutti i piani che li contengono (e che contengono quindi anche la retta passante per  $P_0$  e  $P_1$ ; vedi la Figura 10.8 e (H) più oltre). Per le equazioni parametriche, conviene ragionare come segue: il piano  $\pi$  contiene sia  $P_0$  sia  $P_1$  se e solo se contiene  $P_0$  ed è parallelo al vettore  $P_1 - P_0$ . Dunque possiamo prendere  $P_1 - P_0$  come uno dei vettori di giacitura del piano, e le equazioni parametriche sono

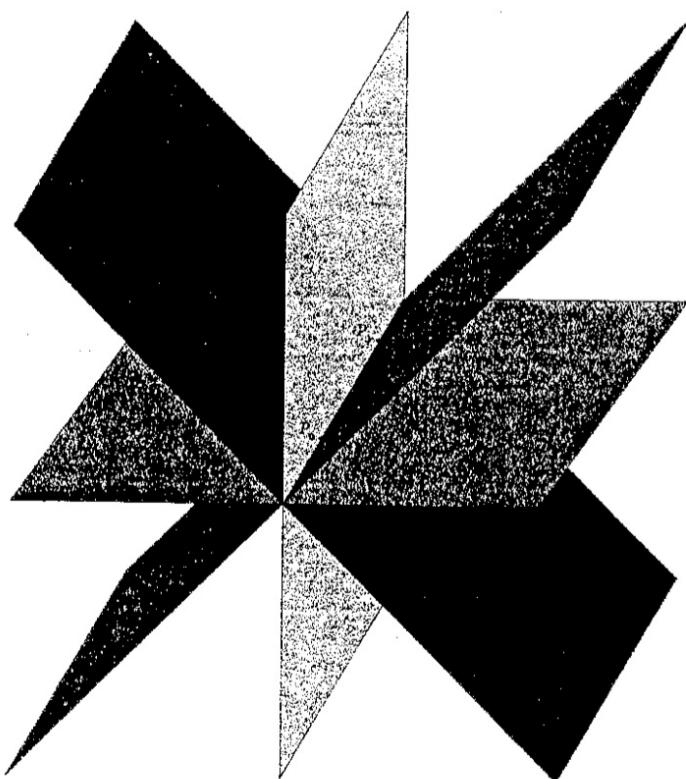
$$\begin{cases} x = x_0 + s(x_1 - x_0) + tl, \\ y = y_0 + s(y_1 - y_0) + tm, \\ z = z_0 + s(z_1 - z_0) + tn, \end{cases}$$

per tutti gli  $l, m, n \in \mathbb{R}$  tali che

$$\operatorname{rg} \begin{vmatrix} x_1 - x_0 & l \\ y_1 - y_0 & m \\ z_1 - z_0 & n \end{vmatrix} = 2.$$

I parametri di giacitura di questo piano devono soddisfare (perché?) l'equazione  $a(x_1 - x_0) + b(y_1 - y_0) + c(z_1 - z_0) = 0$ . Per esempio, se  $z_1 \neq z_0$  possiamo ricavare  $a$  e  $b$  e, dopo aver moltiplicato per  $z_1 - z_0$ , otteniamo come equazione cartesiana

$$a(z_1 - z_0)(x - x_0) + b(z_1 - z_0)(y - y_0) - (a(x_1 - x_0) + b(y_1 - y_0))(z - z_0) = 0$$



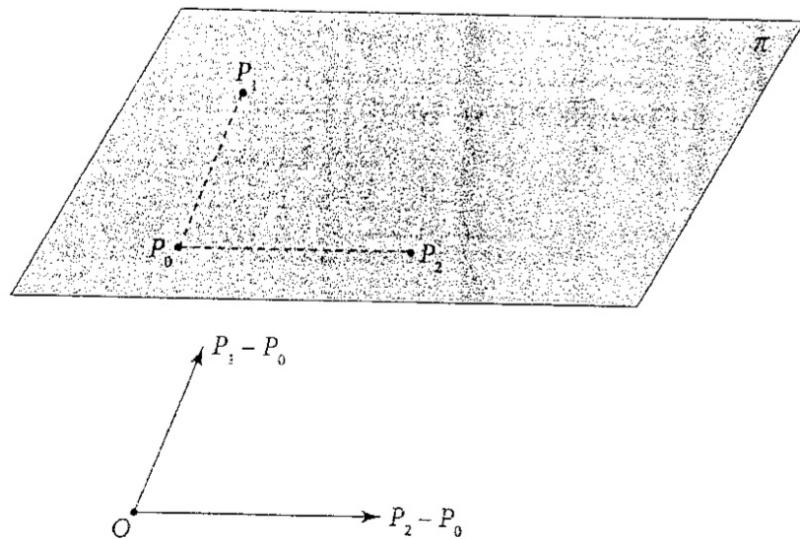
**Figura 10.8** - Piani per due punti.

**Esempio 10.2**

Vogliamo l'equazione cartesiana di tutti i piani contenenti i punti  $P_0 = (-1, 0, 1)$  e  $P_1 = (1, 1, 1)$ . I parametri di giacitura di tali piani devono soddisfare  $2a + b = 0$ ; quindi troviamo l'equazione cartesiana  $a(x + 1) - 2ay + c(z - 1) = 0$ .

(E) *Un piano e tre punti.* Dati tre punti non allineati  $P_0, P_1, P_2 = (x_2, y_2, z_2)$  nello spazio, vogliamo il piano che passa per questi tre punti. Prima di tutto notiamo che i tre punti non sono allineati se e solo se (perché?) i vettori  $P_1 - P_0$  e  $P_2 - P_0$  (vedi la Figura 1) sono linearmente indipendenti (vedi anche l'Esercizio 10.9). Dunque possiamo prendere entrambi come vettori di giacitura, e otteniamo le equazioni parametriche cercate

$$\begin{cases} x = x_0 + s(x_1 - x_0) + t(x_2 - x_0), \\ y = y_0 + s(y_1 - y_0) + t(y_2 - y_0), \\ z = z_0 + s(z_1 - z_0) + t(z_2 - z_0). \end{cases}$$



**Figura 10.9** - Il piano per tre punti.

Per le equazioni cartesiane, i parametri di giacitura sono le soluzioni del sistema

$$\begin{cases} a(x_1 - x_0) + b(y_1 - y_0) + c(z_1 - z_0) = 0, \\ a(x_2 - x_0) + b(y_2 - y_0) + c(z_2 - z_0) = 0. \end{cases}$$

Ma c'è un modo più elegante per ottenere le equazioni cartesiane. Un punto  $P \in \mathbb{R}^3$  appartiene al piano per  $P_0, P_1$  e  $P_2$  se e solo se  $P - P_0$  sta nel sottospazio di giacitura, che è generato da  $P_1 - P_0$  e  $P_2 - P_0$ . Quindi  $P = (x, y, z)$  appartiene al piano se e solo se  $P - P_0, P_1 - P_0$  e  $P_2 - P_0$  sono linearmente dipendenti, cioè se e solo se

$$\det \begin{vmatrix} x - x_0 & x_1 - x_0 & x_2 - x_0 \\ y - y_0 & y_1 - y_0 & y_2 - y_0 \\ z - z_0 & z_1 - z_0 & z_2 - z_0 \end{vmatrix} = 0.$$

Sviluppando il determinante troviamo l'equazione cartesiana cercata.

### Esempio 10.3

Cerchiamo l'equazione cartesiana del piano passante per  $P_0 = (1, 0, 0)$ ,  $P_1 = (0, 0, 1)$  e  $P_2 = (1, 1, 1)$ . Per quanto visto l'equazione è

$$\det \begin{vmatrix} x - 1 & -1 & 0 \\ y & 0 & 1 \\ z & 1 & 1 \end{vmatrix} = 0, \quad \text{cioè} \quad -x + y - z = -1.$$

## 10.4 Rette e rette

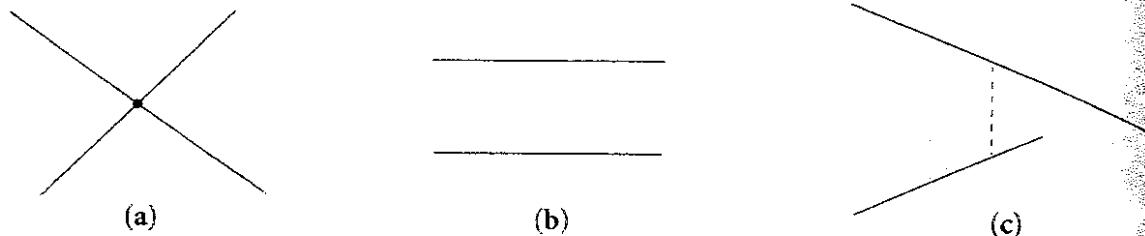
Vediamo ora le posizioni reciproche di due rette.

(F) *Due rette*. Per fissare la terminologia ci serve una definizione.

### Definizione 10.6

Due rette  $r_0$  ed  $r_1$  (sia nel piano sia nello spazio) sono *incidenti* se si intersecano in un punto; *parallele* se non si intersecano ma hanno vettori direttori paralleli (cioè linear-

mente dipendenti); *coincidenti* se sono la stessa retta; e infine *sghembe* se non sono né incidenti né parallele né coincidenti, e dunque in particolare hanno vettori direttori linearmente indipendenti (vedi la Figura 10.10).



**Figura 10.10** - (a) Rette incidenti; (b) rette parallele; (c) rette sghembe.

Vogliamo stabilire quando due rette sono incidenti, parallele o sghembe. Cominciamo supponendo che entrambe siano date in forma parametrica, la retta  $r_0$  dalle equazioni  $P = P_0 + tv_0$ , e la retta  $r_1$  dalle equazioni  $P = P_1 + sv_1$ . Se siamo nel piano,  $P_0$  e  $P_1$  saranno punti del piano, e  $v_0, v_1$  vettori di  $\mathbb{R}^2$ ; nello spazio, saranno invece rispettivamente punti dello spazio e vettori di  $\mathbb{R}^3$ . Le due rette sono parallele (o coincidenti) se e solo se i vettori direttori sono linearmente dipendenti, cioè  $\text{rg } |v_0 \ v_1| = 1$  (*condizione di parallelismo*).

Vediamo ora quando le rette si intersecano (in uno o più punti). Se  $P \in r_0 \cap r_1$ , il punto  $P$  deve potersi scrivere sia come  $P_0 + tv_0$  per un determinato valore di  $t$ , sia come  $P_1 + sv_1$  per un determinato valore di  $s$ . Quindi le due rette  $r_0$  e  $r_1$  si intersecano se e solo se il sistema  $P_0 + tv_0 = P_1 + sv_1$  ammette soluzione. Applicando il Teorema di Rouche-Capelli otteniamo

$$\text{rg } |v_0 \ v_1 \ P_0 - P_1| = \begin{cases} \text{rg } |v_0 \ v_1| & \text{le rette si intersecano;} \\ \text{rg } |v_0 \ v_1| + 1 & \text{le rette non si intersecano.} \end{cases}$$

Supponiamo che  $\text{rg } |v_0 \ v_1| = 1$ . Allora la condizione  $\text{rg } |v_0 \ v_1 \ P_0 - P_1| = 1$  implica che le due rette si intersecano, necessariamente in una retta, per cui sono *coincidenti*. Se invece  $\text{rg } |v_0 \ v_1 \ P_0 - P_1| = 2$ , non si intersecano per cui sono *parallele*.

Supponiamo che  $\text{rg } |v_0 \ v_1| = 2$ . Allora la condizione  $\text{rg } |v_0 \ v_1 \ P_0 - P_1| = 2$  implica che le due rette si intersecano, necessariamente (perché?) in un punto, per cui sono *incidenti*. Se invece  $\text{rg } |v_0 \ v_1 \ P_0 - P_1| = 3$  le due rette non sono né parallele né intersecanti, per cui sono *sghembe*.

Quanto abbiamo ottenuto può venire riassunto dalla tabella seguente.

	$\text{rg }  v_0 \ v_1 \ P_0 - P_1  = 1$	$\text{rg }  v_0 \ v_1 \ P_0 - P_1  = 2$	$\text{rg }  v_0 \ v_1 \ P_0 - P_1  = 3$
$\text{rg }  v_0 \ v_1  = 1$	Rette coincidenti	Rette parallele	<i>Impossibile</i>
$\text{rg }  v_0 \ v_1  = 2$	<i>Impossibile</i>	Rette incidenti	Rette sghembe

#### Osservazione 10.4

Nel piano la matrice  $|v_0 \ v_1 \ P_0 - P_1|$  ha solo due righe, e quindi non può mai avere ранgo 3. Per questo motivo nel piano due rette sono sempre parallele o incidenti (o coincidenti). Nello spazio, invece, la matrice  $|v_0 \ v_1 \ P_0 - P_1|$  è quadrata; quindi per stabilire se ha ранgo 3 o meno basta controllare il determinante. In particolare, le due rette sono sghembe se e solo se  $\det |v_0 \ v_1 \ P_0 - P_1| \neq 0$ .

Passiamo ora al caso in cui le due rette siano date da equazioni cartesiane, iniziando con il piano. Se  $r_0$  e  $r_1$  hanno equazione  $a_0x + b_0y = c_0$  e  $a_1x + b_1y = c_1$ , rispettivamente, esse (sono distinte e) si intersecano se e solo se il sistema

$$\begin{cases} a_0x + b_0y = c_0, \\ a_1x + b_1y = c_1, \end{cases}$$

ammette un'unica soluzione. Questo a sua volta accade, come sappiamo, se e solo se

$$\det \begin{vmatrix} a_0 & b_0 \\ a_1 & b_1 \end{vmatrix} \neq 0.$$

Dunque le due rette sono parallele (o coincidenti) se e solo se  $a_0b_1 - a_1b_0 = 0$  (*condizione di parallelismo*). Infine, le due rette sono coincidenti se e solo se sono parallele e s'intersecano, che accade se e solo se

$$\operatorname{rg} \begin{vmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \end{vmatrix} = 1,$$

ovvero se e solo se le due equazioni sono una multiplo dell'altra.

Nello spazio, invece, la retta  $r_0$  ha equazione cartesiana  $A_0x = b_0$ , e la retta  $r_1$  ha equazione cartesiana  $A_1x = b_1$ , dove  $A_0$  e  $A_1$  sono matrici  $2 \times 3$  di rango 2, e  $b_0$  e  $b_1$  sono vettori di  $\mathbb{R}^2$ . I sottospazi di giacitura delle due rette hanno rispettivamente equazione  $A_0x = O$  e  $A_1x = O$ . Le due rette sono parallele (o coincidenti) se e solo se questi due sistemi omogenei hanno le stesse soluzioni. Siccome entrambi hanno una retta di soluzioni, le due rette sono parallele se e solo se il sistema di quattro equazioni in tre incognite

$$\begin{cases} A_0x = O, \\ A_1x = O \end{cases}$$

ammette soluzioni diverse da  $x = O$ , che accade se e solo se

$$\operatorname{rg} \begin{vmatrix} A_0 \\ A_1 \end{vmatrix} = 2$$

(*condizione di parallelismo*; stiamo usando il Teorema 5.7 della dimensione). D'altro can-  
to le due rette si intersecano se e solo se il sistema

$$\begin{cases} A_0x = b_0, \\ A_1x = b_1 \end{cases} \quad (10.21)$$

ammette soluzioni, il che accade se e solo se

$$\operatorname{rg} \begin{vmatrix} A_0 & b_0 \\ A_1 & b_1 \end{vmatrix} = \operatorname{rg} \begin{vmatrix} A_0 \\ A_1 \end{vmatrix};$$

se i ranghi di queste matrici sono diversi le rette sono parallele o sghembe.

Anche stavolta possiamo riassumere quanto trovato in una tabella.

	$\operatorname{rg} \begin{vmatrix} A_0 & b_0 \\ A_1 & b_1 \end{vmatrix} = 2$	$\operatorname{rg} \begin{vmatrix} A_0 & b_0 \\ A_1 & b_1 \end{vmatrix} = 3$	$\operatorname{rg} \begin{vmatrix} A_0 & b_0 \\ A_1 & b_1 \end{vmatrix} = 4$
$\operatorname{rg} \begin{vmatrix} A_0 \\ A_1 \end{vmatrix} = 2$	Rette coincidenti	Rette parallele	<i>Impossibile</i>
$\operatorname{rg} \begin{vmatrix} A_0 \\ A_1 \end{vmatrix} = 3$	<i>Impossibile</i>	Rette incidenti	Rette sghembe

**Esempio 10.4**

Date le rette  $r_0$  ed  $r_1$  di equazioni cartesiane

$$r_0 : \begin{cases} x - y + z = 1, \\ x + y - z = 1, \end{cases} \quad r_1 : \begin{cases} x + y + z = 0, \\ 3x + y + z = 2, \end{cases}$$

vogliamo stabilire in che posizione reciproca si trovano. Siccome

$$\text{rg} \begin{vmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \\ 3 & 1 & 1 \end{vmatrix} = 3 = \text{rg} \begin{vmatrix} 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 0 \\ 3 & 1 & 1 & 2 \end{vmatrix},$$

le due rette sono incidenti. □

## 10.5 Rette e piani

Puntiamo ora i riflettori sulle relazioni fra rette e piani.

(G) *Una retta e un piano.* Consideriamo una retta  $r$  di equazioni parametriche  $P = P_0 + sv_0$  e un piano  $\pi$  di equazioni parametriche  $P = P_1 + tv_1 + t'v_2$ . Abbiamo varie possibilità (vedi la Figura 10.11).

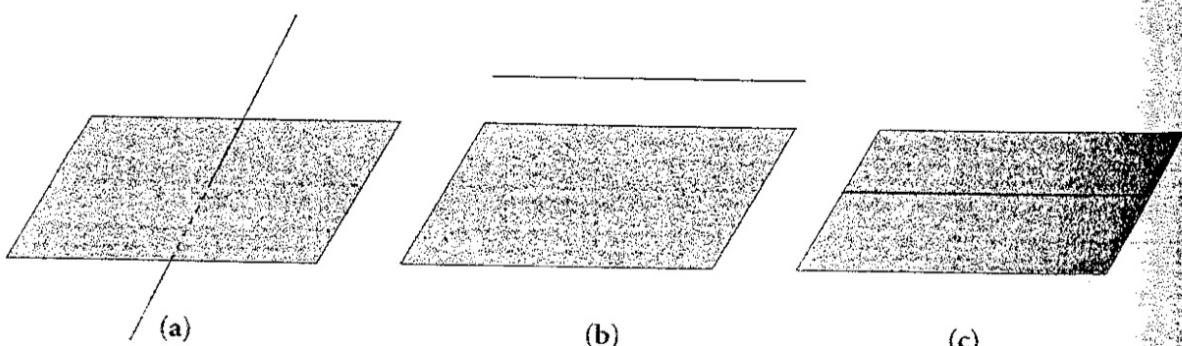
- Se  $\det |v_0 \ v_1 \ v_2| \neq 0$ , allora il sistema

$$P_0 + sv_0 = P_1 + tv_1 + t'v_2 \tag{10.22}$$

nelle incognite  $s$ ,  $t$  e  $t'$  ammette un'unica soluzione. Questo vuol dire che la retta e il piano sono *incidenti*: si intersecano in un punto.

- Se invece  $\det |v_0 \ v_1 \ v_2| = 0$ , il sistema o non ha soluzioni (nel qual caso la retta e il piano sono *paralleli*), oppure ha una retta di soluzioni (nel qual caso la retta  $r$  è *contenuta* nel piano). In particolare, retta e piano sono paralleli se e solo se

$$\text{rg} |v_0 \ v_1 \ v_2 \ P_0 - P_1| = 3 \neq 2 = \text{rg} |v_0 \ v_1 \ v_2|.$$



**Figura 10.11** - (a) Retta incidente un piano; (b) retta parallela a un piano; (c) retta contenuta in un piano.

Supponiamo invece che la retta  $r$  abbia equazioni cartesiane  $A_0x = b_0$ , e che il piano  $\pi$  abbia equazioni cartesiane  $\alpha^T x = d$ , dove  $\alpha \in \mathbb{R}^3$  è il vettore dei parametri di giacitura del piano. Retta e piano sono paralleli se e solo se il sistema quadrato

$$\begin{cases} A_0x = O, \\ \alpha^T x = 0 \end{cases}$$

ha soluzioni diverse da  $x = O$  (perché?), il che accade se e solo se  $\det \begin{vmatrix} A_0 \\ \alpha^T \end{vmatrix} = 0$ . In questo caso, la retta è contenuta nel piano se e solo se  $\text{rg} \begin{vmatrix} A_0 & b_0 \\ \alpha^T & d \end{vmatrix} = 2$ . Invece  $\det \begin{vmatrix} A_0 \\ \alpha^T \end{vmatrix} \neq 0$  se e solo se il piano e la retta sono incidenti.

**Esempio 10.5**

Consideriamo la retta  $r$  e il piano  $\pi$  di equazioni cartesiane

$$r : \begin{cases} x + y + z = 1, \\ x - y + z = 0, \end{cases} \quad \pi : 3x + y + 3z = 2.$$

Siccome

$$\text{rg} \begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 3 & 1 & 3 \end{vmatrix} = 2 = \text{rg} \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 0 \\ 3 & 1 & 3 & 2 \end{vmatrix},$$

la retta  $r$  è contenuta nel piano  $\pi$ . □

Vi è un altro caso che vale la pena considerare. Supponiamo che il piano  $\pi$  abbia equazione cartesiana  $ax + by + cz = d$ , e che la retta  $r$  abbia equazioni parametriche  $P = P_0 + t\mathbf{v}_0$ . In tal caso, la retta e il piano sono paralleli se e solo se il vettore direttore  $\mathbf{v}_0 = (l, m, n)$  della retta appartiene al sottospazio di giacitura del piano, cioè se e solo se

$$al + bm + cn = 0 \tag{10.23}$$

(condizione di parallelismo). In questo caso, la retta è contenuta nel piano se e solo se si ha inoltre  $P_0 \in \pi$ , il che accade se e solo se  $d = ax_0 + by_0 + cz_0$ .

(H) *Un piano e una retta.* Data una retta  $r$ , vogliamo trovare tutti i piani che la contengono. Cominciamo come al solito supponendo che la retta  $r$  abbia equazioni parametriche  $P = P_0 + t\mathbf{v}_0$ , con  $P_0 = (x_0, y_0, z_0)$  e  $\mathbf{v}_0 = (l_0, m_0, n_0)$ . Se un piano contiene  $r$ , il vettore  $\mathbf{v}_0$  può venire preso come un vettore di giacitura del piano. Quindi tutti i piani contenenti  $r$  sono descritti dalle equazioni parametriche

$$\begin{vmatrix} x \\ y \\ z \end{vmatrix} = \begin{vmatrix} x_0 \\ y_0 \\ z_0 \end{vmatrix} + t \begin{vmatrix} l_0 \\ m_0 \\ n_0 \end{vmatrix} + s \begin{vmatrix} l \\ m \\ n \end{vmatrix} \tag{10.24}$$

al variare di  $t, m, n \in \mathbb{R}$  in modo che si abbia

$$\text{rg} \begin{vmatrix} l_0 & l \\ m_0 & m \\ n_0 & n \end{vmatrix} = 2.$$

Supponiamo invece che la retta  $r$  sia data da equazioni cartesiane,

$$\begin{cases} ax + by + cz = d, \\ a'x + b'y + c'z = d'. \end{cases}$$

Per quanto visto in (G), il piano di equazione cartesiana  $a''x + b''y + c''z = d''$  contiene la retta  $r$  se e solo se

$$\text{rg} \begin{vmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{vmatrix} = 2,$$

cioè se e solo se esistono  $\lambda, \mu \in \mathbb{R}$  non entrambi nulli tali che

$$\begin{vmatrix} a'' \\ b'' \\ c'' \\ d'' \end{vmatrix} = \lambda \begin{vmatrix} a \\ b \\ c \\ d \end{vmatrix} + \mu \begin{vmatrix} a' \\ b' \\ c' \\ d' \end{vmatrix}.$$

Dunque tutti i piani contenenti  $r$  sono descritti dall'equazione cartesiana

$$\lambda(ax + by + cz) + \mu(a'x + b'y + c'z) = \lambda d + \mu d'$$

al variare di  $(\lambda, \mu) \in \mathbb{R}^2$  con  $(\lambda, \mu) \neq (0, 0)$ .

### Definizione 10.7

L'insieme dei piani passanti per la retta  $r$  si chiama *fascio di piani* di asse  $r$  (vedi la Figura 10.12).

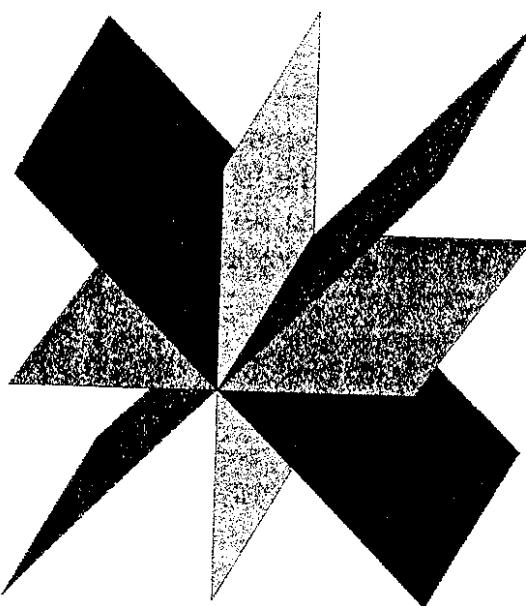


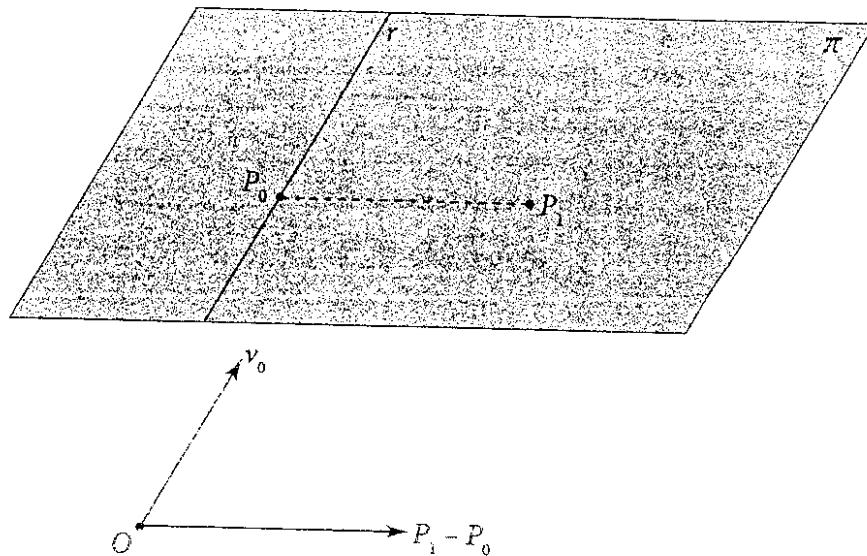
Figura 10.12 - Fascio di piani.

(I) *Un piano, una retta e un punto.* Vogliamo trovare il piano contenente una retta data  $r$  e passante per un punto  $P_1 = (x_1, y_1, z_1)$  non appartenente alla retta. Supponiamo che la retta abbia equazioni parametriche  $P = P_0 + tv_0$ , con  $P_0 = (x_0, y_0, z_0)$  e  $v_0 = (l_0, m_0, n_0)$ . Ricordando la forma delle equazioni parametriche dei piani del fascio di asse  $r$ , vediamo subito che dobbiamo trovare  $l, m, n \in \mathbb{R}$  tali che il sistema

$$P_1 = P_0 + tv_0 + sv \quad (10.25)$$

abbia soluzione, dove  $v = (l, m, n)$ . Chiaramente,  $v = P_1 - P_0$  è una possibilità (la soluzione è  $t = 0, s = 1$ ). Viceversa, supponiamo che l'equazione (10.25) sia soddisfatta per qualche  $s_0, t_0 \in \mathbb{R}$ . Siccome per ipotesi  $P_1$  non appartiene alla retta  $r$ , la differenza  $P_1 - P_0$  non può essere un multiplo di  $v_0$ , per cui  $s_0 \neq 0$ . Dividendo per  $s_0$  troviamo  $v = \lambda(P_1 - P_0) + \mu v_0$  per qualche  $\lambda, \mu \in \mathbb{R}$  con  $\lambda = 1/s_0 \neq 0$ . Inoltre, qualunque scelta di  $\lambda \in \mathbb{R}^*$  e  $\mu \in \mathbb{R}$  ci fornisce lo stesso piano, in quanto si ha  $\text{Span}(v_0, \lambda(P_1 - P_0) + \mu v_0) = \text{Span}(v_0, P_1 - P_0)$ . In conclusione, esiste un unico piano contenente  $r$  e  $P_0$ , di equazioni parametriche (vedi la Figura 10.13)

$$P = P_0 + tv_0 + s(P_1 - P_0).$$



**Figura 10.13** - Piano contenente una retta e passante per un punto.

Se invece la retta  $r$  è data dalle equazioni cartesiane

$$\begin{cases} ax + by + cz = d, \\ a'x + b'y + c'z = d', \end{cases}$$

l'unico piano del fascio di asse  $r$  che contiene il punto  $P_1$  è quello per cui i parametri  $\lambda$  e  $\mu$  soddisfano  $(ax_1 + by_1 + cz_1 - d)\lambda + (a'x_1 + b'y_1 + c'z_1 - d')\mu = 0$  (nota che, siccome  $P_1 \notin r$ , i coefficienti di questa equazione non sono entrambi nulli). Quindi il piano ha equazione cartesiana

$$(a'x_1 + b'y_1 + c'z_1 - d')(ax + by + cz - d) - (ax_1 + by_1 + cz_1 - d)(a'x + b'y + c'z - d') = 0.$$

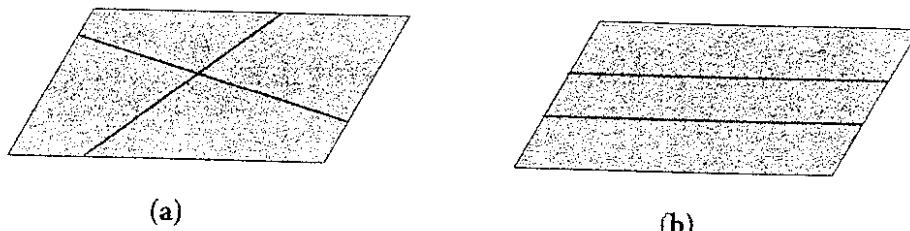
(L) *Un piano e due rette.* Vogliamo vedere quando (e se) esiste un piano contenente due rette diverse  $r_0$  ed  $r_1$ . Come vedremo, questo piano esiste sempre tranne nel caso in cui le rette siano sghembe.

Supponiamo che le equazioni parametriche delle rette siano date da  $P = P_0 + tv_0$  e  $P = P_1 + sv_1$  rispettivamente, e distinguiamo due casi (vedi la Figura 10.14).

♦  $\text{rg } |v_0 \ v_1| = 2$ : le rette sono incidenti o sghembe. Perché un piano  $\pi$  contenga sia  $r_0$  sia  $r_1$  deve contenere sia  $P_0$  sia  $P_1$ , e il sottospazio di giacitura deve contenere sia  $v_0$  sia  $v_1$ . Siccome  $v_0$  e  $v_1$  sono per ipotesi linearmente indipendenti, possiamo prenderli come vettori di giacitura di  $\pi$ , che ha quindi

$$P = P_0 + sv_0 + tv_1 \quad (10.26)$$

come equazioni parametriche. Questo piano contiene anche  $P_1$  (e quindi entrambe le rette) se e solo se il sistema  $P_1 = P_0 + sv_0 + tv_1$  ha soluzione, il che accade se e solo se  $\text{rg } |P_0 - P_1 \ v_0 \ v_1| = \text{rg } |v_0 \ v_1|$ , per cui le rette sono incidenti. Dunque se  $v_0$  e  $v_1$  sono



**Figura 10.14** - (a) Piano contenente due rette incidenti; (b) piano contenente due rette parallele.

linearmente indipendenti esiste un piano che contiene  $r_0$  ed  $r_1$  se e solo se le due rette sono incidenti, e in tal caso il piano ha (10.26) come equazioni parametriche.

- ◆  $\operatorname{rg} |\nu_0 \nu_1| = 1$ : le rette sono parallele (e distinte, per ipotesi). In questo caso è sufficiente prendere il piano  $\pi$  che contiene  $r_0$  e  $P_1$ , che ha equazioni parametriche  $P = P_0 + s\nu_0 + t(P_1 - P_0)$ ; infatti  $\pi$  contiene anche  $r_1$  (perché?).

Vediamo ora il caso in cui  $r_0$  ed  $r_1$  siano date da equazioni cartesiane

$$r_0 : \begin{cases} a_0x + b_0y + c_0z = d_0, \\ a'_0x + b'_0y + c'_0z = d'_0, \end{cases} \quad r_1 : \begin{cases} a_1x + b_1y + c_1z = d_1, \\ a'_1x + b'_1y + c'_1z = d'_1 \end{cases}$$

Il piano che le contiene entrambe (se esiste) dev'essere un piano comune ai due fasci di asse rispettivamente  $r_0$  ed  $r_1$ . Questo accade se e solo se esistono  $\lambda_0, \mu_0, \lambda_1, \mu_1 \in \mathbb{R}$  non tutti nulli tali che le due equazioni

$$\begin{aligned} \lambda_0(a_0x + b_0y + c_0z) + \mu_0(a'_0x + b'_0y + c'_0z) &= \lambda_0d_0 + \mu_0d'_0, \\ \lambda_1(a_1x + b_1y + c_1z) + \mu_1(a'_1x + b'_1y + c'_1z) &= \lambda_1d_1 + \mu_1d'_1 \end{aligned} \quad (10.27)$$

siano uguali (vedi l'Esercizio 10.1). Confrontando i coefficienti rispettivamente di  $x, y, z$  e i termini noti, vediamo che a sua volta questo succede se e solo se il sistema

$$\left| \begin{array}{cccc|c} a_0 & a'_0 & -a_1 & -a'_1 & \lambda_0 \\ b_0 & b'_0 & -b_1 & -b'_1 & \mu_0 \\ c_0 & c'_0 & -c_1 & -c'_1 & \lambda_1 \\ d_0 & d'_0 & -d_1 & -d'_1 & \mu_1 \end{array} \right| = \left| \begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \end{array} \right| \quad (10.28)$$

ammette soluzioni non nulle. Questo accade se e solo se la matrice dei coefficienti ha rango strettamente minore di 4, il che – ricordando (F) – equivale a dire che le rette non sono sghembe. Infine, il piano cercato ha come equazione cartesiana la seconda delle equazioni in (10.27), dove  $\lambda_0, \mu_0, \lambda_1$  e  $\mu_1$  sono soluzioni del sistema (10.28).

### Esempio 10.6

Vogliamo l'equazione cartesiana del piano contenente le rette incidenti studiate nell'Esempio 10.4. Risolvendo il sistema

$$\left| \begin{array}{ccccc|c} 1 & 1 & -1 & -3 & \lambda_0 \\ -1 & 1 & -1 & -1 & \mu_0 \\ 1 & -1 & -1 & -1 & \lambda_1 \\ 1 & 1 & 0 & -2 & \mu_1 \end{array} \right| = \left| \begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \end{array} \right|$$

troviamo (per esempio)  $\lambda_1 = -1$  e  $\mu_1 = 1$ , per cui il piano cercato ha equazione cartesiana  $x = 1$ . □

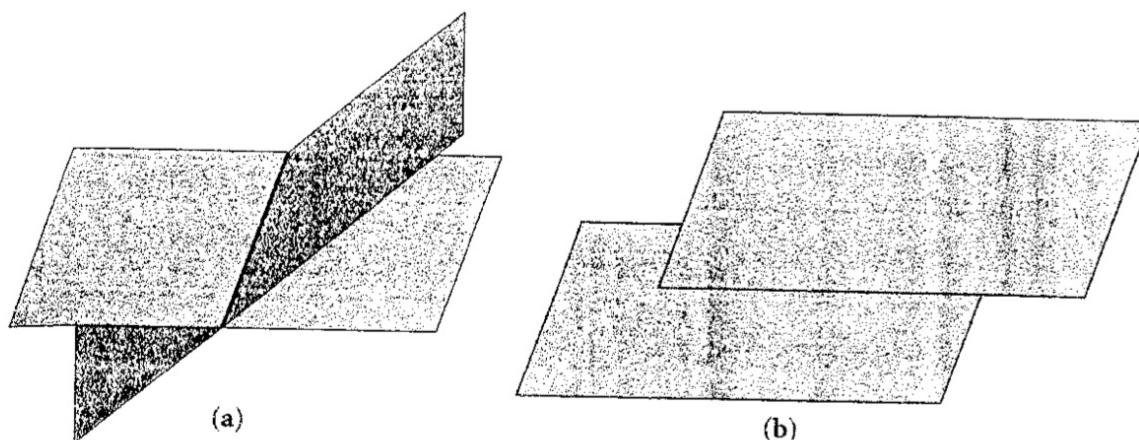
## 10.6 Piani e piani

Concludiamo con le posizioni reciproche di due piani.

(M) *Due piani*. Una definizione è di rigore.

### Definizione 10.8

Diremo che due piani  $\pi_0$  e  $\pi_1$  sono *incidenti* se si intersecano in una retta; *paralleli* se non si intersecano o coincidono, cioè sono lo stesso piano (vedi la Figura 10.15).



**Figura 10.15** - (a) Piani incidenti; (b) piani paralleli.

Supponiamo che i due piani  $\pi_0$  e  $\pi_1$  abbiano equazioni parametriche rispettivamente  $P = P_0 + sv_0 + tw_0$  e  $P = P_1 + s'v_1 + t'w_1$ . Ci sono soltanto due possibilità: se  $\text{rg} |v_0 \ w_0 \ v_1 \ w_1| = 2$  i due piani hanno lo stesso sottospazio di giacitura, e quindi sono *paralleli*. In particolare, sono coincidenti (perché?) se e solo se  $\text{rg} |v_0 \ w_0 \ v_1 \ w_1 \ P_0 - P_1| = 2$ . Se invece  $\text{rg} |v_0 \ w_0 \ v_1 \ w_1| = 3$ , i due piani sono *incidenti*; infatti in tal caso il sistema  $P_0 + sv_0 + tw_0 = P_1 + s'v_1 + t'w_1$  ha sempre una retta di soluzioni.

Supponiamo ora che i due piani abbiano equazioni cartesiane  $a_0x + b_0y + c_0z = d_0$  e  $a_1x + b_1y + c_1z = d_1$  rispettivamente. Anche stavolta ci sono solo due possibilità. Se

$$\text{rg} \begin{vmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \end{vmatrix} = 1$$

(condizione di parallelismo), i due piani hanno lo stesso sottospazio di giacitura, e quindi sono paralleli. In particolare, sono coincidenti se e solo se

$$\text{rg} \begin{vmatrix} a_0 & b_0 & c_0 & d_0 \\ a_1 & b_1 & c_1 & d_1 \end{vmatrix} = 1.$$

Se invece

$$\text{rg} \begin{vmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \end{vmatrix} = 2,$$

il sistema composto dalle equazioni dei due piani ammette una retta di soluzioni, per cui i piani sono incidenti.

Quanto visto finora ovviamente non esaurisce le questioni di geometria affine che è possibile porsi (e risolvere), ma descrive in maniera abbastanza esaustiva il tipo di tecniche che vengono utilizzate. Concludiamo questo paragrafo con un esempio di come si possa risolvere un altro problema di questo genere.

### Esempio 10.7

Vogliamo determinare equazioni parametriche e cartesiane del piano passante per il punto  $P_0 = (1, 0, 1)$  e parallelo alle rette

$$r_1 : \begin{cases} x = t, \\ y = 2t, \\ z = 1 + 3t, \end{cases} \quad r_2 : \begin{cases} 2x + z = 1, \\ 2y - z = 3. \end{cases}$$

La retta  $r_2$  ha equazioni parametriche

$$\begin{cases} x = -2s, \\ y = 2 + 2s, \\ z = 1 + 4s; \end{cases}$$

quindi il piano cercato ha equazioni parametriche

$$\begin{cases} x = 1 + t - 2s, \\ y = 2t + 2s, \\ z = 1 + 3t + 4s, \end{cases}$$

e dunque equazione cartesiana  $2(x - 1) - 10y + 6(z - 1) = 0$ , ovvero  $x - 5y + 3z = 4$ .

□

## 10.7 Sistemi di riferimento affini

Quanto fatto finora si basa sull'aver fissato una volta per tutte un sistema di riferimento affine  $RA(O, A_1, A_2, A_3)$  in  $\mathcal{A}^3$ . Ma che succede se cambiamo sistema di riferimento? Ovviamente, cambieranno le coordinate. Come abbiamo già accennato nel Paragrafo 2.2, a volte può essere utile scegliere un sistema di riferimento *ad hoc* per il problema da esaminare; sorge quindi spontaneo il problema di come mutano le coordinate passando da un sistema di riferimento a un altro. Abbiamo già esaminato la questione nel Paragrafo 8.1 nel caso degli spazi vettoriali, in cui cambiava la base ma l'origine rimaneva la stessa. Cambiando sistema di riferimento affine, invece, muoviamo anche l'origine, per cui la situazione è un po' diversa.

Ricordiamo brevemente le definizioni.

### Definizione 10.9

Un *sistema di riferimento affine*  $RA(O, A_1, A_2, A_3)$  in  $\mathcal{A}^3$  consiste in una quaterna ordinata di punti non complanari  $O, A_1, A_2, A_3 \in \mathcal{A}^3$ . Il primo punto  $O$  è l'*origine* del sistema di riferimento. Siccome i quattro punti sono non complanari, i tre vettori  $\overrightarrow{OA_1}, \overrightarrow{OA_2}$  e  $\overrightarrow{OA_3}$  formano una base dello spazio vettoriale  $\mathcal{V}_O^3$ . Le coordinate associate a questa base si chiamano *coordinate affini* rispetto al sistema di riferimento affine  $RA(O, A_1, A_2, A_3)$ .

Per l'esattezza, un punto  $P \in \mathcal{A}^3$  ha coordinate affini  $(x_1, x_2, x_3) \in \mathbb{R}^3$  se e solo se (vedi la Figura 10.16)

$$\overrightarrow{OP} = x_1 \overrightarrow{OA_1} + x_2 \overrightarrow{OA_2} + x_3 \overrightarrow{OA_3}.$$

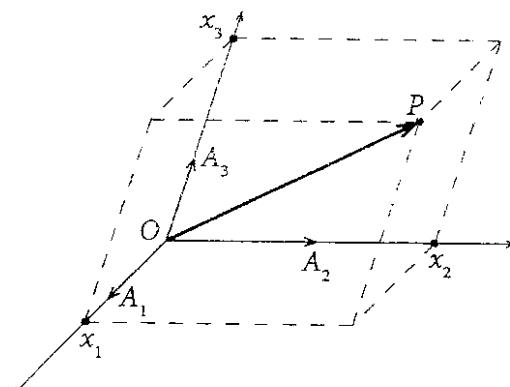


Figura 10.16 - Le coordinate affini.

Siano  $\mathcal{R} = RA(O, A_1, A_2, A_3)$  e  $\mathcal{R}' = RA(O', A'_1, A'_2, A'_3)$  due sistemi di riferimento affini in  $\mathcal{A}^3$ . Prendiamo un punto  $P \in \mathcal{A}^3$ ; vogliamo vedere che relazione intercorre fra le sue coordinate rispetto a  $\mathcal{R}$  e quelle rispetto a  $\mathcal{R}'$ .

Siano  $(c_1, c_2, c_3)$  le coordinate di  $O'$  rispetto a  $\mathcal{R}$  (ovviamente, il punto  $O'$  ha coordinate  $(0, 0, 0)$  rispetto a  $\mathcal{R}'$ ), cioè

$$\overrightarrow{OO'} = c_1 \overrightarrow{OA_1} + c_2 \overrightarrow{OA_2} + c_3 \overrightarrow{OA_3} = \begin{vmatrix} \overrightarrow{OA_1} & \overrightarrow{OA_2} & \overrightarrow{OA_3} \end{vmatrix} c,$$

dove  $c = [c_1 \ c_2 \ c_3]^T \in \mathbb{R}^3$ , e il prodotto è il solito prodotto righe per colonne.

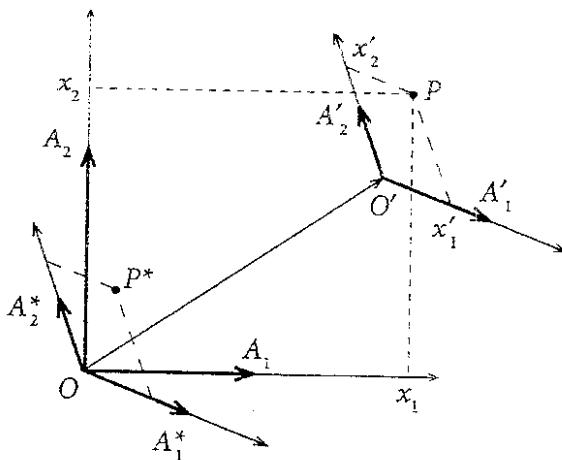


Figura 10.17 - Cambiamento di coordinate affini in  $\mathcal{A}^2$ .

Consideriamo ora il sistema di riferimento affine  $\mathcal{R}^* = RA(O, A_1^*, A_2^*, A_3^*)$  ottenuto traslando l'origine di  $\mathcal{R}'$  in  $O$  (vedi la Figura 10.17, dove per semplicità è raffigurata la situazione nel piano  $\mathcal{A}^2$ ). I punti  $A_j^*$  sono ottenuti traslando i punti  $A'_j$  tramite il vettore  $\overrightarrow{OO'}$ , cioè  $\overrightarrow{OA_j^*} = \overrightarrow{OA'_j} - \overrightarrow{OO'}$ , per  $j = 1, 2, 3$ . Siccome i quattro punti  $O', A'_1, A'_2$  e  $A'_3$  non erano complanari, anche i punti  $O, A_1^*, A_2^*$  e  $A_3^*$  non lo sono, in quanto loro traslati; ne segue che i vettori  $\{\overrightarrow{OA_1^*}, \overrightarrow{OA_2^*}, \overrightarrow{OA_3^*}\}$  sono una base dello spazio vettoriale  $\mathcal{V}_O^3$ . Quanto visto nel Paragrafo 8.1 ci fornisce allora una matrice  $B \in GL_3(\mathbb{R})$  – la matrice di cambiamento di base – tale che

$$\begin{vmatrix} \overrightarrow{OA_1^*} & \overrightarrow{OA_2^*} & \overrightarrow{OA_3^*} \end{vmatrix} = \begin{vmatrix} \overrightarrow{OA_1} & \overrightarrow{OA_2} & \overrightarrow{OA_3} \end{vmatrix} B.$$

Prendiamo allora il nostro punto  $P \in \mathcal{A}^3$ , di coordinate  $x = (x_1, x_2, x_3)$  rispetto a  $\mathcal{R}$ , e  $x' = (x'_1, x'_2, x'_3)$  rispetto a  $\mathcal{R}'$ . Il punto  $P$  viene traslato nel punto  $P^*$  tramite il vettore  $\overrightarrow{OO'}$ , in modo da avere  $\overrightarrow{OP^*} = \overrightarrow{OP} - \overrightarrow{OO'}$ . Dunque le coordinate di  $P^*$  rispetto a  $\mathcal{R}^*$  sono le stesse di  $P$  rispetto a  $\mathcal{R}'$ , per cui

$$\begin{aligned} \begin{vmatrix} \overrightarrow{OA_1} & \overrightarrow{OA_2} & \overrightarrow{OA_3} \end{vmatrix} x &= x_1 \overrightarrow{OA_1} + x_2 \overrightarrow{OA_2} + x_3 \overrightarrow{OA_3} = \overrightarrow{OP} = \overrightarrow{OO'} + \overrightarrow{OP^*} \\ &= \overrightarrow{OO'} + x'_1 \overrightarrow{OA_1^*} + x'_2 \overrightarrow{OA_2^*} + x'_3 \overrightarrow{OA_3^*} \\ &= \begin{vmatrix} \overrightarrow{OA_1} & \overrightarrow{OA_2} & \overrightarrow{OA_3} \end{vmatrix} c + \begin{vmatrix} \overrightarrow{OA_1^*} & \overrightarrow{OA_2^*} & \overrightarrow{OA_3^*} \end{vmatrix} x' \\ &= \begin{vmatrix} \overrightarrow{OA_1} & \overrightarrow{OA_2} & \overrightarrow{OA_3} \end{vmatrix} (c + Bx'). \end{aligned}$$

Quindi

$$x = Bx' + c, \quad (10.29)$$

cioè

$$\begin{cases} x_1 = b_{11}x'_1 + b_{12}x'_2 + b_{13}x'_3 + c_1, \\ x_2 = b_{21}x'_1 + b_{22}x'_2 + b_{23}x'_3 + c_2, \\ x_3 = b_{31}x'_1 + b_{32}x'_2 + b_{33}x'_3 + c_3. \end{cases} \quad (10.30)$$

Dunque un cambiamento di coordinate affini consiste in un'applicazione lineare seguita da una traslazione (combinazione che rivedremo nei complementi a questo capitolo quando parleremo di affinità). Per completezza, riportiamo qui anche il cambiamento inverso, che è dato da

$$x' = B^{-1}x - B^{-1}c.$$

### Osservazione 10.5

Fin qui abbiamo ragionato in  $\mathcal{A}^3$ , ma ovviamente la stessa argomentazione si applica in  $\mathcal{A}^2$ , con una coordinata in meno. La relazione (10.29) rimane identica, con  $B \in GL_2(\mathbb{R})$  e  $c \in \mathbb{R}^2$ , mentre il sistema (10.30) perde una riga diventando

$$\begin{cases} x_1 = b_{11}x'_1 + b_{12}x'_2 + c_1, \\ x_2 = b_{21}x'_1 + b_{22}x'_2 + c_2. \end{cases}$$

## 10.8 Orientazione

La formula (10.29) ci permette di introdurre un'ulteriore definizione.

### Definizione 10.10

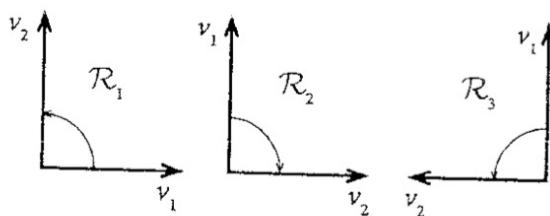
Due sistemi di riferimento affini  $\mathcal{R} = RA(O, A_1, A_2, A_3)$  e  $\mathcal{R}' = RA(O', A'_1, A'_2, A'_3)$  hanno la *stessa orientazione* (diremo anche che sono *equivalenti*) se la matrice  $B$  del cambiamento di coordinate affini da  $\mathcal{R}$  a  $\mathcal{R}'$  ha determinante positivo; diremo che  $\mathcal{R}$  e  $\mathcal{R}'$  hanno *orientazione opposta* (sono *contraversi*) se invece  $\det B < 0$ . La stessa definizione si applica nel caso del piano.

Per capire meglio di cosa stiamo parlando, consideriamo un *sistema di riferimento* affine  $\mathcal{R}_1 = RA(O, A_1, A_2)$  in  $\mathcal{A}^2$ . Rispetto a  $\mathcal{R}_1$ , la base  $\{v_1 = \overrightarrow{OA_1}, v_2 = \overrightarrow{OA_2}\}$  di  $\mathcal{V}_O^2$  ha come coordinate i vettori della base canonica  $\{e_1, e_2\}$  di  $\mathbb{R}^2$ . Se le ascisse vengono messe in orizzontale dirette verso destra e le ordinate in verticale dirette verso l'alto, per passare da  $e_1$  a  $e_2$  bisogna ruotare in senso antiorario.

Consideriamo invece il sistema di riferimento affine  $\mathcal{R}_2 = RA(O, A_2, A_1)$ , con la stessa origine (o anche un'altra: è evidente che l'orientazione non dipende dall'origine scelta) ma con i vettori di base  $\{v_1 = \overrightarrow{OA_2}, v_2 = \overrightarrow{OA_1}\}$  scambiati. Stavolta per passare dal primo al secondo vettore di base bisogna ruotare in senso orario. La matrice del cambiamento di coordinate è

$$B = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}$$

che ha determinante  $-1$ ; quindi i due sistemi  $\mathcal{R}_1$  ed  $\mathcal{R}_2$  hanno orientazione opposta.

Figura 10.18 - Orientazione in  $A^2$ .

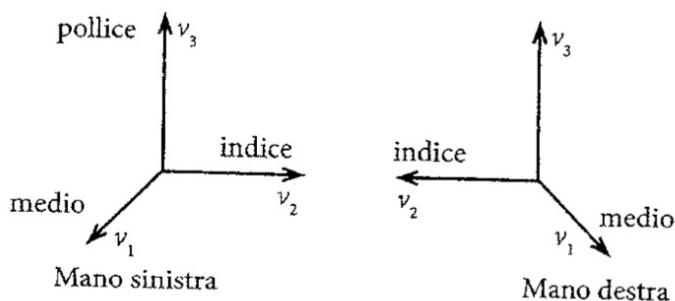
Se ora in  $\mathcal{R}_2$  sostituiamo  $A_1$  col suo opposto rispetto a  $O$  otteniamo (vedi la Figura 10.18) un sistema di riferimento affine  $\mathcal{R}_3$  con base  $\{v_1 = \overrightarrow{OA_2}, v_2 = -\overrightarrow{OA_1}\}$ . Le matrici di cambiamento di base  $B_1$  da  $\mathcal{R}_1$  a  $\mathcal{R}_3$  e  $B_2$  da  $\mathcal{R}_2$  a  $\mathcal{R}_3$  sono date da

$$B_1 = \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} \quad \text{e} \quad B_2 = \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix};$$

dunque  $\mathcal{R}_2$  e  $\mathcal{R}_3$  hanno orientazione opposta, mentre  $\mathcal{R}_1$  e  $\mathcal{R}_3$  hanno la stessa orientazione. E infatti per passare dal primo al secondo vettore di base in  $\mathcal{R}_3$  bisogna ruotare di nuovo in senso antiorario.

Riassumendo, in  $A^2$  il succo intuitivo del concetto di orientazione è: due sistemi di riferimento affine hanno la stessa orientazione se e solo se per passare dal primo al secondo vettore di base si effettua una rotazione che ha lo stesso senso in entrambi i casi.

In  $A^3$  la situazione è analoga, anche se espressa in maniera lievemente diversa. Nei disegni usuali, la base canonica  $\{e_1, e_2, e_3\}$  di  $\mathbb{R}^3$  viene rappresentata nella stessa posizione in cui si trovano, nell'ordine, il medio, l'indice e il pollice della mano sinistra se cerchiamo di tenerli perpendicolari (vedi la Figura 10.19). Allora due sistemi di riferimento affine in  $A^3$  hanno la stessa orientazione se e solo se i tre vettori della base dei due sistemi sono disposti come medio, indice e pollice (nell'ordine) della stessa mano, destra o sinistra che sia.

Figura 10.19 - Orientazione in  $A^3$ .

# Complementi 10

## 10C.1 Affinità

Possiamo ora descrivere meglio lo scopo della geometria affine: consiste nello studio delle cosiddette *proprietà affini*, cioè di quelle proprietà dei sottoinsiemi di  $\mathcal{A}^3$  (o  $\mathcal{A}^2$ , o  $\mathcal{A}^1$ ) che non dipendono dal sistema di riferimento affine usato per descriverli. Per esempio, “essere una retta”, oppure “essere parallela a un dato piano” o “intersecare una data retta” sono tutte proprietà affini; “avere la terza coordinata nulla”, invece, non è una proprietà affine (come puoi agevolmente verificare).

Una domanda abbastanza naturale a questo punto, stimolata anche dal fatto che finora ci siamo sistematicamente rifiutati di parlarne, è: la distanza fra due punti è una proprietà affine oppure no? Per rispondere, consideriamo il caso di due punti  $P_1$  e  $P_2$  in  $\mathcal{A}^1$ , cioè sulla retta. Fissiamo un sistema di riferimento affine  $\mathcal{R} = RA(O, A_1)$  in  $\mathcal{A}^1$  in modo da poter scrivere  $OP_1 = x_1 \overrightarrow{OA}_1$  e  $OP_2 = x_2 \overrightarrow{OA}_1$ . Dunque in questo sistema di riferimento la “distanza” fra  $P_1$  e  $P_2$  è data da  $|x_2 - x_1|$ . In un altro sistema di riferimento affine  $\mathcal{R}' = RA(O', A'_1)$  invece la “distanza” fra  $P_1$  e  $P_2$  sarà data da  $|x'_2 - x'_1|$ , dove stavolta  $x'_1$  e  $x'_2$  sono le coordinate affini di  $P_1$  e  $P_2$  rispetto a  $\mathcal{R}'$ . Per quanto visto nel Paragrafo 10.7 (adattato al caso di  $\mathcal{A}^1$ ) esistono  $b \in \mathbb{R}^*$  e  $c \in \mathbb{R}$  tali che  $x_1 = bx'_1 + c$  e  $x_2 = bx'_2 + c$ . Quindi non appena  $|b| \neq 1$  le “distanze”  $|x_2 - x_1|$  e  $|x'_2 - x'_1|$  risultano diverse, per cui la distanza fra due punti *non* è una proprietà affine.

Questo risultato suggerisce però un’altra possibilità. Consideriamo anche un terzo punto  $P_3$ , di coordinata  $x_3$  rispetto a  $\mathcal{R}$  e  $x'_3$  rispetto a  $\mathcal{R}'$ . Siccome abbiamo di nuovo  $x_3 = bx'_3 + c$ , vediamo subito che

$$\frac{x_3 - x_1}{x_2 - x_1} = \frac{x'_3 - x'_1}{x'_2 - x'_1}$$

non appena  $P_2 \neq P_1$ . Dunque il rapporto  $(x_3 - x_1)/(x_2 - x_1)$  è una *proprietà affine* dei tre punti  $P_1$ ,  $P_2$  e  $P_3$  presi in quest’ordine; diamogli un nome.

### Definizione 10C.1

Siano  $P_1, P_2, P_3 \in \mathcal{A}^1$  tre punti della retta. Il *rapporto semplice*  $(P_1, P_2, P_3) \in \mathbb{R} \cup \{\infty\}$  di  $P_1, P_2$  e  $P_3$  (presi in quest’ordine e con  $P_1 \neq P_2$ ) è

$$(P_1, P_2, P_3) = \frac{x_3 - x_1}{x_2 - x_1}, \quad (10C.1)$$

dove  $x_j$  è la coordinata di  $P_j$  (per  $j = 1, 2, 3$ ) rispetto a un qualsiasi sistema di riferimento affine su  $\mathcal{A}^1$  (infatti abbiamo appena visto che il quoziente (10C.1) non dipende dal sistema di riferimento in cui viene calcolato). Se invece  $P_1 = P_2 \neq P_3$ , poniamo  $(P_1, P_2, P_3) = \infty$ . Infine, il rapporto semplice  $(P_1, P_2, P_3)$  non è definito se  $P_1 = P_2 = P_3$ .

Un modo per capire il significato geometrico del rapporto semplice (e il motivo per cui non dipende dal sistema di riferimento) è contenuto nel seguente lemma.

### Lemma 10C.1

*Siano  $O, P_1, P_2, P_3 \in \mathcal{A}^1$  quattro punti della retta, con  $P_1 \neq P_2$ . Allora si ha  $(P_1, P_2, P_3) = t \in \mathbb{R}$  se e solo se*

$$\overrightarrow{OP_3} = \overrightarrow{OP_1} + t(\overrightarrow{OP_2} - \overrightarrow{OP_1}).$$

### Dimostrazione

Siccome  $P_1 \neq P_2$ , esiste un unico  $t \in \mathbb{R}$  tale che

$$\overrightarrow{OP_3} - \overrightarrow{OP_1} = t(\overrightarrow{OP_2} - \overrightarrow{OP_1}); \quad (10C.2)$$

dobbiamo quindi verificare che  $t = (P_1, P_2, P_3)$ . Fissiamo un sistema di riferimento affine  $RA(O, A_1)$  rispetto a cui  $P_j$  ha coordinata  $x_j$ , per  $j = 1, 2, 3$ . Allora (10C.2) diventa  $(x_3 - x_1)\overrightarrow{OA_1} = t(x_2 - x_1)\overrightarrow{OA_1}$ , ed è fatta.  $\square$

Studiando gli spazi vettoriali ci siamo soffermati a lungo sulle applicazioni lineari, cioè su quelle funzioni che conservano la struttura di spazio vettoriale (ovvero le operazioni di somma e prodotto per scalari). Analogamente ora siamo interessati alle funzioni che conservano le proprietà affini, come l'appartenenza o il rapporto semplice. Non esamineremo però la situazione nella sua massima generalità (in particolare, non considereremo il caso di spazi affini di dimensione maggiore di tre, o di applicazioni fra spazi affini di dimensione diversa), limitandoci invece a quanto descritto nelle seguenti definizioni.

### Definizione 10C.2

Una retta affine (piano affine, spazio affine) è semplicemente una copia di  $\mathcal{A}^1$  (rispettivamente, di  $\mathcal{A}^2$  o  $\mathcal{A}^3$ ). Per esempio, una qualunque retta contenuta in  $\mathcal{A}^3$  è una retta affine.

### Definizione 10C.3

Un'affinità fra due rette affini  $r$  ed  $r'$  è una funzione bigettiva  $\varphi: r \rightarrow r'$  che conserva i rapporti semplici, cioè tale che

$$(\varphi(P_1), \varphi(P_2), \varphi(P_3)) = (P_1, P_2, P_3)$$

per qualunque tripla di punti  $P_1, P_2, P_3 \in r$  non tutti coincidenti. Un'affinità fra due piani affini  $\pi$  e  $\pi'$  è una funzione bigettiva  $\varphi: \pi \rightarrow \pi'$  che manda rette in rette e tale che ristretta a ogni retta  $r \subset \pi$  sia un'affinità fra  $r$  e la sua immagine  $\varphi(r)$ . In particolare  $\varphi$  conserva i rapporti semplici di punti allineati. Infine, un'affinità fra due spazi affini  $\Sigma$  e  $\Sigma'$  è una funzione bigettiva  $\varphi: \Sigma \rightarrow \Sigma'$  che manda piani in piani e tale che ristretta a ogni piano  $\pi \subset \Sigma$  sia un'affinità fra  $\pi$  e la sua immagine  $\varphi(\pi)$ . In particolare,  $\varphi$  manda rette in rette e conserva i rapporti semplici di punti allineati.

Dalla definizione è chiaro che le affinità conservano alcune delle proprietà affini (l'appartenenza, il rapporto semplice, l'essere una retta); vedremo fra un istante che le conservano proprio tutte, giustificando quindi il nostro interesse per loro.

Cominciamo con le affinità fra rette.

**Proposizione 10C.2**

Sia  $\varphi: r \rightarrow r'$  un'affinità fra le rette affini  $r$  ed  $r'$ . Fissiamo un sistema di riferimento affine  $\mathcal{R} = RA(O, A_1)$  su  $r$ , e uno  $\mathcal{R}' = RA(O', A'_1)$  su  $r'$ . Allora esistono  $b \in \mathbb{R}$  e  $c \in \mathbb{R}$  tali che

$$\forall P \in r \quad y = bx + c, \quad (10C.3)$$

dove  $x$  è la coordinata di  $P$  rispetto a  $\mathcal{R}$  e  $y$  è la coordinata di  $\varphi(P)$  rispetto a  $\mathcal{R}'$ . In particolare, la funzione  $\varphi$  conserva tutte le proprietà affini. Viceversa, ogni  $\varphi: r \rightarrow r'$  che in coordinate è descritta da (10C.3) con  $b \neq 0$  è un'affinità.

**Dimostrazione**

Siano  $c, d \in \mathbb{R}$  le coordinate di  $\varphi(O)$  e  $\varphi(A_1)$  rispetto a  $\mathcal{R}'$ . Essendo  $\varphi$  iniettiva, il numero  $b = d - c$  è diverso da zero.

Prendiamo ora  $P \in r$  qualunque; siccome  $\varphi$  conserva i rapporti semplici, si ha

$$(O, A_1, P) = (\varphi(O), \varphi(A_1), \varphi(P)),$$

ovvero, usando il riferimento  $\mathcal{R}$  in  $r$  e il riferimento  $\mathcal{R}'$  in  $r'$ ,

$$\frac{x - 0}{1 - 0} = \frac{y - c}{d - c},$$

cioè  $y = bx + c$ , come voluto. In particolare, siccome (10C.3) ha la stessa espressione di un cambiamento di coordinate affini, e le proprietà affini sono esattamente quelle conservate da tali cambiamenti di coordinate, se ne deduce che  $\varphi$  conserva tutte le proprietà affini.

Infine, una  $\varphi: r \rightarrow r'$  data in coordinate da (10C.3) è chiaramente bigettiva (in quanto  $b \neq 0$ ) e conserva i rapporti semplici (esercizio), per cui è un'affinità.  $\square$

**Osservazione 10C.1**

Attenzione a non confondere affinità e cambiamenti di coordinate, nonostante abbiano la stessa espressione. In un'affinità, i sistemi di riferimento sono fissi e sono i punti che si muovono: il punto  $P$  di coordinata  $x$  va a finire nel punto  $\varphi(P)$  di coordinata  $y$ . Invece nei cambiamenti di coordinate i punti sono fermi e cambia il sistema di riferimento: il punto  $P$  ha coordinata  $x$  rispetto a  $\mathcal{R}$ , e coordinata  $x'$  rispetto a  $\mathcal{R}'$ .

Passiamo ora alle affinità del piano. Ci serve un lemma.

**Lemma 10C.3**

Sia  $\varphi: \pi \rightarrow \pi'$  un'affinità fra piani affini, e  $O, A_1, A_2 \in \pi$  tre punti non allineati. Allora anche  $\varphi(O), \varphi(A_1), \varphi(A_2) \in \pi'$  sono non allineati.

**Dimostrazione**

Supponiamo per assurdo che  $\varphi(O), \varphi(A_1)$  e  $\varphi(A_2)$  siano allineati, e consideriamo la retta  $r$  passante per  $O$  e  $A_1$ . Ora,  $\varphi$  manda rette in rette; quindi  $\varphi(r)$  è la retta che contiene  $\varphi(O)$  e  $\varphi(A_1)$ . In particolare,  $\varphi(A_2) \in \varphi(r)$ , per cui esiste  $P \in r$  tale che  $\varphi(P) = \varphi(A_2)$ . Ma  $O, A_1$  e  $A_2$  non sono allineati; quindi  $A_2 \notin r$ , e dunque  $P \neq A_2$  mentre per ipotesi  $\varphi$  è iniettiva, contraddizione.  $\square$

**Proposizione 10C.4**

Sia  $\varphi: \pi \rightarrow \pi'$  un'affinità fra i piani affini  $\pi$  e  $\pi'$ . Sia dato un sistema di riferimento affine  $\mathcal{R} = RA(O, A_1, A_2)$  su  $\pi$ , e uno  $\mathcal{R}' = RA(O', A'_1, A'_2)$  su  $\pi'$ . Allora esistono  $B \in GL_2(\mathbb{R})$  e  $c \in \mathbb{R}^2$  tali che

$$\forall P \in \pi \quad y = Bx + c, \quad (10C.4)$$

dove  $x = (x_1, x_2)$  sono le coordinate di  $P$  rispetto a  $\mathcal{R}$  e  $y = (y_1, y_2)$  le coordinate di  $\varphi(P)$  rispetto a  $\mathcal{R}'$ . In particolare,  $\varphi$  conserva tutte le proprietà affini. Viceversa, ogni  $\varphi: \pi \rightarrow \pi'$  che in coordinate è data da (10C.4) con  $B \in GL_2(\mathbb{R})$  è un'affinità.

**Dimostrazione**

Siano  $(c_1, c_2)$  le coordinate di  $\varphi(O)$  rispetto a  $\mathcal{R}'$ , e  $(d_{1j}, d_{2j})$  le coordinate di  $\varphi(A_j)$  rispetto a  $\mathcal{R}'$ , per  $j = 1, 2$ .

Un punto  $P_1 \in \pi$  avente coordinate  $(x_1, 0)$  sta sulla retta passante per  $O$  e  $A_1$ ; quindi  $\varphi(P_1)$  deve stare sulla retta passante per  $\varphi(O)$  e  $\varphi(A_1)$ . Siccome il rapporto semplice fra  $O, A_1$  e  $P_1$  è proprio  $x_1$ , si deve avere  $(\varphi(O), \varphi(A_1), \varphi(P_1)) = x_1$ . Il Lemma 10C.1 allora ci dice che  $\varphi(P_1)$  deve avere coordinate

$$\begin{vmatrix} c_1 \\ c_2 \end{vmatrix} + x_1 \begin{vmatrix} d_{11} - c_1 \\ d_{21} - c_2 \end{vmatrix}.$$

Analogamente, il punto di coordinate  $(0, x_2)$  ha per immagine il punto di coordinate

$$\begin{vmatrix} c_1 \\ c_2 \end{vmatrix} + x_2 \begin{vmatrix} d_{12} - c_1 \\ d_{22} - c_2 \end{vmatrix}.$$

Ora, l'affinità  $\varphi$  manda rette in rette per ipotesi, e conserva il parallelismo in quanto biegettiva; dunque in particolare manda parallelogrammi in parallelogrammi. Se  $P \in \pi$  ha coordinate  $(x_1, x_2)$ , allora è il quarto vertice del parallelogramma determinato da  $O$  e dai punti  $P_1$  di coordinate  $(x_1, 0)$  e  $P_2$  di coordinate  $(0, x_2)$ . Ne segue che  $\varphi(P)$  è il quarto vertice del parallelogramma di vertici  $\varphi(O)$ ,  $\varphi(P_1)$  e  $\varphi(P_2)$ , cioè

$$\overrightarrow{\varphi(O)\varphi(P)} = \overrightarrow{\varphi(O)\varphi(P_1)} + \overrightarrow{\varphi(O)\varphi(P_2)}.$$

In coordinate otteniamo (perché?)

$$\begin{vmatrix} y_1 \\ y_2 \end{vmatrix} - \begin{vmatrix} c_1 \\ c_2 \end{vmatrix} = \left( \begin{vmatrix} c_1 \\ c_2 \end{vmatrix} + x_1 \begin{vmatrix} d_{11} - c_1 \\ d_{21} - c_2 \end{vmatrix} - \begin{vmatrix} c_1 \\ c_2 \end{vmatrix} \right) + \left( \begin{vmatrix} c_1 \\ c_2 \end{vmatrix} + x_2 \begin{vmatrix} d_{12} - c_1 \\ d_{22} - c_2 \end{vmatrix} - \begin{vmatrix} c_1 \\ c_2 \end{vmatrix} \right),$$

cioè

$$\begin{vmatrix} y_1 \\ y_2 \end{vmatrix} = \begin{vmatrix} d_{11} - c_1 & d_{12} - c_1 \\ d_{21} - c_2 & d_{22} - c_2 \end{vmatrix} \begin{vmatrix} x_1 \\ x_2 \end{vmatrix} + \begin{vmatrix} c_1 \\ c_2 \end{vmatrix}.$$

Ponendo

$$B = \begin{vmatrix} d_{11} - c_1 & d_{12} - c_1 \\ d_{21} - c_2 & d_{22} - c_2 \end{vmatrix},$$

abbiamo dimostrato (10C.4), perché l'Esercizio 10.8 implica  $B \in GL_2(\mathbb{R})$  in quanto  $\varphi(O)$ ,  $\varphi(A_1)$  e  $\varphi(A_2)$  non sono allineati grazie al Lemma 10C.3.

Il fatto che  $\varphi$  conserva tutte le proprietà affini si dimostra come nel caso delle affinità fra rette. Infine, se  $\varphi: \Sigma \rightarrow \Sigma'$  è data da (10C.4), allora  $\varphi$  è bigettiva perché  $B \in GL_2(\mathbb{R})$ ; manda ovviamente rette in rette; e ristretta a qualunque retta si esprime in coordinate nella forma (10C.3). Quindi  $\varphi$  è un'affinità fra i due piani.  $\square$

Con considerazioni analoghe si risolve anche il caso dello spazio.

### Proposizione 10C.5

*Sia  $\varphi: \Sigma \rightarrow \Sigma'$  un'affinità fra gli spazi affini  $\Sigma$  e  $\Sigma'$ . Fissiamo un sistema di riferimento affine  $\mathcal{R} = RA(O, A_1, A_2, A_3)$  su  $\Sigma$ , e un secondo  $\mathcal{R}' = RA'(O', A'_1, A'_2, A'_3)$  su  $\Sigma'$ . Allora esistono  $B \in GL_3(\mathbb{R})$  e  $c \in \mathbb{R}^3$  tali che*

$$\forall P \in \Sigma \quad y = Bx + c, \quad (10C.5)$$

*dove  $x = (x_1, x_2, x_3)$  sono le coordinate di  $P$  rispetto a  $\mathcal{R}$ , e  $y = (y_1, y_2, y_3)$  sono le coordinate di  $\varphi(P)$  rispetto a  $\mathcal{R}'$ . In particolare,  $\varphi$  conserva tutte le proprietà affini. Viceversa, ogni  $\varphi: \Sigma \rightarrow \Sigma'$  che in coordinate è data da (10C.5) con  $B \in GL_3(\mathbb{R})$  è un'affinità.*

### Dimostrazione

Si procede come per la Proposizione 10C.4, usando parallelepipedi al posto dei parallelogrammi.  $\square$

## 10C.2 Geometria proiettiva

In questo paragrafo vogliamo introdurre i concetti di base della geometria proiettiva, con l'obiettivo di mostrare come nascano in modo naturale dalla geometria affine. Cominciamo innanzitutto a studiare le proiezioni nel piano affine.

### Definizione 10C.4

Siano  $r_1$  e  $r_2 \subset \mathcal{A}^2$  due rette distinte del piano, e  $r_0 \subset \mathcal{A}^2$  una terza retta, incidente sia con  $r_1$  sia con  $r_2$ . La *proiezione parallela* da  $r_1$  a  $r_2$  di *direzione*  $r_0$  è l'applicazione  $\psi_{r_0}: r_1 \rightarrow r_2$  che a ogni punto  $P$  appartenente a  $r_1$  associa l'intersezione  $\psi_{r_0}(P) \in r_2$  di  $r_2$  con la retta passante per  $P$  e parallela a  $r_0$  (vedi la Figura 10.20).

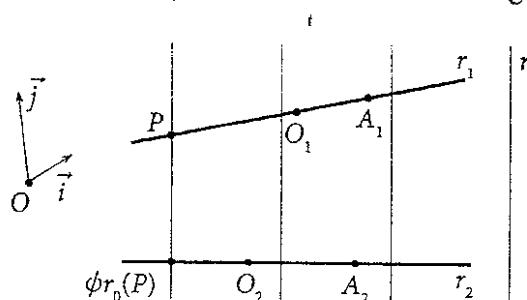


Figura 10.20 - La proiezione parallela.

### Definizione 10C.5

Siano  $r_1$  e  $r_2 \subset \mathcal{A}^2$  due rette distinte del piano, e  $P_0 \in \mathcal{A}^2$  un punto che non appartiene né a  $r_1$  né a  $r_2$ . La *proiezione centrale* da  $r_1$  a  $r_2$  di *centro*  $P_0$  è l'applicazione  $\psi_{P_0}: r_1 \rightarrow r_2$  che a ogni punto <sup>(2)</sup>  $P \in r_1$  associa l'intersezione  $\psi_{P_0}(P) \in r_2$  di  $r_2$  con la retta per  $P_0$  e  $P$  (vedi la Figura 10.21).

<sup>(2)</sup> O meglio, a quasi ogni punto; come vedremo, c'è un problema, e per risolverlo saremo portati a studiare la geometria proiettiva.

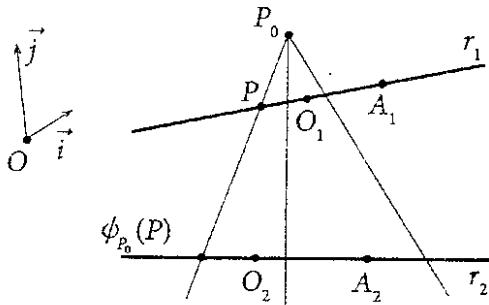


Figura 10.21 - La proiezione centrale.

Vediamo adesso come si esprimono in coordinate queste proiezioni. Fissiamo un sistema di riferimento affine  $\mathcal{R} = RA(O, \vec{i}, \vec{j})$  nel piano  $A^2$ , un sistema di riferimento affine  $\mathcal{R}_1 = RA(O_1, A_1)$  in  $r_1$  e un altro  $\mathcal{R}_2 = RA(O_2, A_2)$  in  $r_2$ . Per  $j = 1, 2$ , indichiamo con

$$O_j = \begin{vmatrix} c_{1j} \\ c_{2j} \end{vmatrix}, \quad A_j = \begin{vmatrix} d_{1j} \\ d_{2j} \end{vmatrix}$$

le coordinate rispetto a  $\mathcal{R}$  dei punti che compongono il riferimento  $\mathcal{R}_j$ . In particolare, un vettore direttore di  $r_j$  ha coordinate

$$\begin{vmatrix} e_{1j} \\ e_{2j} \end{vmatrix} = \begin{vmatrix} d_{1j} - c_{1j} \\ d_{2j} - c_{2j} \end{vmatrix},$$

per cui un punto  $P \in r_1$  di coordinata  $x$  rispetto a  $\mathcal{R}_1$  avrà rispetto a  $\mathcal{R}$  coordinate

$$\begin{vmatrix} c_{11} + e_{11}x \\ c_{21} + e_{21}x \end{vmatrix},$$

e un punto  $Q \in r_2$  di coordinata  $y$  rispetto a  $\mathcal{R}_2$  avrà rispetto a  $\mathcal{R}$  coordinate

$$\begin{vmatrix} c_{12} + e_{12}y \\ c_{22} + e_{22}y \end{vmatrix}.$$

Iniziamo con lo studiare la proiezione parallela  $\psi_{r_0}$ . Indichiamo con

$$\nu = \begin{vmatrix} \nu_1 \\ \nu_2 \end{vmatrix}$$

le coordinate rispetto a  $\mathcal{R}$  di un vettore direttore di  $r_0$ . Preso (vedi la Figura 10.20) un punto  $P \in r_1$  di coordinata  $x$  (rispetto a  $\mathcal{R}_1$ ), i punti della retta per  $P$  parallela a  $r_0$  hanno coordinate

$$\begin{vmatrix} c_{11} + e_{11}x + t\nu_1 \\ c_{21} + e_{21}x + t\nu_2 \end{vmatrix}$$

al variare di  $t \in \mathbb{R}$ , per cui la coordinata  $y$  (rispetto a  $\mathcal{R}_2$ ) del punto  $\psi_{r_0}(P)$  deve soddisfare il sistema

$$\begin{cases} c_{11} + e_{11}x + t\nu_1 = c_{12} + e_{12}y, \\ c_{21} + e_{21}x + t\nu_2 = c_{22} + e_{22}y, \end{cases}$$

ovvero

$$\begin{cases} e_{12}y - \nu_1 t = (c_{11} - c_{12}) + e_{11}x, \\ e_{22}y - \nu_2 t = (c_{21} - c_{22}) + e_{21}x. \end{cases}$$

Siccome a noi interessa solo il valore di  $y$ , possiamo applicare il Teorema di Cramer (Corollario 9.11) ottenendo

$$y = \frac{(v_1 e_{21} - v_2 e_{11})x + [v_1(c_{21} - c_{22}) - v_2(c_{11} - c_{12})]}{v_1 e_{22} - v_2 e_{12}};$$

nota che  $v_1 e_{21} - v_2 e_{11}, v_1 e_{22} - v_2 e_{12} \neq 0$  in quanto  $r_0$  non è parallela né a  $r_1$  né a  $r_2$ , que in coordinate la proiezione parallela è data da  $x \mapsto y = \alpha x + \beta$  con  $\alpha \neq 0$ , per (Proposizione 10C.2) è un'affinità fra le due rette.

Consideriamo invece (vedi la Figura 10.21) la proiezione centrale  $\psi_{P_0}$  di centro punto  $P_0$  di coordinate

$$P_0 = \begin{vmatrix} x_1 \\ x_2 \end{vmatrix}$$

rispetto a  $\mathcal{R}$ . Preso un punto  $P \in r_1$  di coordinata  $x$  (rispetto a  $\mathcal{R}_1$ ), i punti della retta per  $P$  e  $P_0$  hanno coordinate

$$\begin{vmatrix} x_1 + t(c_{11} + e_{11}x - x_1) \\ x_2 + t(c_{21} + e_{21}x - x_2) \end{vmatrix}$$

al variare di  $t \in \mathbb{R}$ , per cui la coordinata  $y$  (rispetto a  $\mathcal{R}_2$ ) del punto  $\psi_{P_0}(P)$  deve soddisfare il sistema

$$\begin{cases} e_{12}y - (c_{11} - x_1 + e_{11}x)t = x_1 - c_{12}, \\ e_{22}y - (c_{21} - x_2 + e_{21}x)t = x_2 - c_{22}. \end{cases}$$

Siccome a noi interessa solo il valore di  $y$ , possiamo di nuovo applicare il Teorema di Cramer ottenendo

$$y = \frac{[e_{11}(x_2 - c_{22}) - e_{21}(x_1 - c_{12})]x + [(x_1 - c_{12})(x_2 - c_{21}) - (x_1 - c_{11})(x_2 - c_{22})]}{(e_{11}e_{22} - e_{12}e_{21})x + [e_{12}(x_2 - c_{21}) - e_{22}(x_1 - c_{11})]}$$

Dunque la proiezione centrale è data in coordinate da

$$x \mapsto y = \frac{\alpha x + \beta}{\gamma x + \delta}, \quad (10C.6)$$

per cui non è un'affinità (a meno che  $\gamma = 0$ , che accade se e solo se  $r_1$  ed  $r_2$  sono parallele); vogliamo studiarne le proprietà, nel caso più interessante in cui  $r_1$  ed  $r_2$  non siano parallele.

Prima di tutto, la presenza di un denominatore in (10C.6) ci avvisa che la proiezione centrale non è definita su tutta la retta  $r_1$ . Geometricamente era prevedibile: siccome  $r_1$  ed  $r_2$  non sono parallele, esiste un punto  $P_1^\infty \in r_1$  (di coordinata  $x = -\delta/\gamma$ ; esercizio) tale che la retta passante per esso e  $P_0$  è parallela a  $r_2$ , per cui  $\psi_{P_0}(P_1^\infty)$  non esiste.

Dunque la proiezione centrale è definita soltanto su  $r_1 \setminus \{P_1^\infty\}$ ; vediamo se è almeno iniettiva. Di nuovo, geometricamente è immediato: supponiamo che  $P_1, P_2 \in r_1$  siano tali che  $\psi_{P_0}(P_1) = Q = \psi_{P_0}(P_2)$ . Siccome  $P_0 \neq Q$ , la retta  $s$  passante per  $P_0$  e  $Q$  è unica, interseca  $r_1$  sia in  $P_1$  sia in  $P_2$ . Se  $P_1$  fosse diverso da  $P_2$ , questo implicherebbe  $s = r_1$  e quindi  $P_0 \in r_1$  contro l'ipotesi; quindi  $P_1 = P_2$  e  $\psi_{P_0}$  è iniettiva.

Rimane da vedere se  $\psi_{P_0}$  è surgettiva. Siccome  $r_1$  ed  $r_2$  non sono parallele, esiste un punto  $P_2^\infty \in r_2$  (di coordinata  $y = \alpha/\gamma$ ; esercizio) tale che la retta per  $P_0$  e  $P_2^\infty$  sia parallela a  $r_1$ ; quindi  $P_2^\infty$  non può essere immagine di alcun punto di  $r_1$ .

Tutto ciò è piuttosto fastidioso. Siamo partiti da una tranquilla operazione geometrica e ci siamo trovati costretti a tener presente tutta una serie di eccezioni. Per vedere se è possibile evitare il problema, esaminiamo più accuratamente la situazione. Prendiamo un punto  $P \in r_1 \setminus \{P_1^\infty\}$  e facciamolo avvicinare a  $P_1^\infty$ ; l'immagine  $\psi_{P_0}(P)$  si allontana sempre di più su  $r_2$ , fino a fuggire all'infinito quando  $P$  raggiunge  $P_1^\infty$ . Analogamente, l'intersezione fra  $r_1$  e la retta per  $P_0$  e  $Q \in r_2$  tende ad allontanarsi all'infinito ogni volta che  $Q$  si avvicina a  $P_2^\infty$ . Questo suggerisce di eliminare tutte le eccezioni aggiungendo un "punto all'infinito"  $\infty_j$  a ciascuna delle rette  $r_j$  per poi definire  $\psi_{P_0}(P_1^\infty) = \infty_1$  e  $\psi_{P_0}(\infty_1) = P_2^\infty$ . In questo modo otteniamo una vera applicazione bigettiva da  $\hat{r}_1 = r_1 \cup \{\infty_1\}$  a  $\hat{r}_2 = r_2 \cup \{\infty_2\}$ .

Dunque ci siamo trovati a dover mettere un punto in più a ciascuna retta, un punto che si raggiunge andando all'infinito (in una qualunque delle due direzioni); in un certo senso, è come considerare la nostra retta una sorta di circonferenza di raggio infinito. Per capire meglio la struttura dell'oggetto che abbiamo costruito, esaminiamo più in dettaglio (per esempio)  $r_1$  supponendo per semplicità che  $\mathcal{R}$  sia stato scelto in modo che la retta abbia equazione  $y = 1$ .

Ogni punto  $P \in r_1$  determina una e una sola retta del fascio  $\mathcal{F}$  di centro l'origine  $O$ ; viceversa, ogni retta di  $\mathcal{F}$  (tranne, guarda caso, una sola, quella parallela a  $r_1$ ) interseca  $r_1$  esattamente in un punto. Dunque abbiamo una corrispondenza biunivoca fra i punti di  $r_1$  e le rette del fascio  $\mathcal{F}$  tranne una sola; associando quest'ultima al punto all'infinito  $\infty_1$  otteniamo una perfetta corrispondenza biunivoca fra  $\hat{r}_1$  ed  $\mathcal{F}$  (vedi la Figura 10.22). Ora, una retta del fascio ha equazione parametrica

$$\begin{cases} x = tv_1, \\ y = tv_2, \end{cases} \quad (10C.7)$$

con  $(v_1, v_2) \neq (0, 0)$ ; e due equazioni

$$\begin{cases} x = tv_1, \\ y = tv_2, \end{cases} \quad \begin{cases} x = tv'_1, \\ y = tv'_2, \end{cases}$$

descrivono la stessa retta se e solo se  $v'_1 = \lambda v_1$  e  $v'_2 = \lambda v_2$  per qualche  $\lambda \neq 0$ . Inoltre, la retta di equazione (10C.7) interseca  $r_1$  nel punto  $(v_1/v_2, 1)$  tranne quando  $v_2 = 0$ , caso in cui abbiamo la retta  $y = 0$  che "interseca"  $r_1$  nel punto all'infinito  $\infty_1$ .

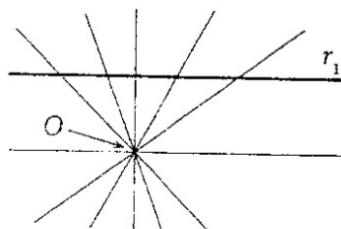


Figura 10.22 - Le coordinate omogenee.

Esprimiamo queste osservazioni in modo un po' diverso. Diremo che due vettori  $v, w \in \mathbb{R}^2 \setminus \{O\}$  sono *equivalenti*, e scriveremo  $v \sim w$ , se esiste  $\lambda \in \mathbb{R}^*$  tale che  $w = \lambda v$  (e quindi  $v = \lambda^{-1}w$ ). In altre parole,  $v \sim w$  se e solo se appartengono alla stessa retta per l'origine. È facile vedere che  $\sim$  è una relazione d'equivalenza (vedi il Paragrafo 2C.1), e ovviamente c'è una corrispondenza biunivoca fra il fascio  $\mathcal{F}$  e l'insieme quoziente  $\mathbb{R}^2 \setminus \{O\}/\sim$  (e quindi fra questo insieme ed  $\hat{r}_1$ ).

**Definizione 10C.6**

L'insieme quoziante  $\mathbb{P}^1(\mathbb{R}) = \mathbb{R}^2 \setminus \{O\} / \sim$  viene chiamato *retta proiettiva*. Se  $\pi: \mathbb{R}^2 \setminus \{O\} \rightarrow \mathbb{P}^1(\mathbb{R})$  è l'applicazione quoziante e  $(v_1, v_2) \in \mathbb{R}^2 \setminus \{O\}$ , scriveremo  $[v_1 : v_2]$  al posto di  $\pi(v_1, v_2)$ ; la coppia  $(v_1, v_2) \neq (0, 0)$  contiene le *coordinate omogenee* del punto  $[v_1 : v_2] \in \mathbb{P}^1(\mathbb{R})$ . Chiaramente si ha  $[v_1 : v_2] = [\lambda v_1 : \lambda v_2]$  per ogni  $\lambda \in \mathbb{R}^*$ . Se  $v_2 \neq 0$ , il rapporto  $v_1/v_2$  viene detto *coordinata non omogenea* del punto *proprio*  $[v_1 : v_2]$ ; il punto  $[1 : 0]$  (l'unico per cui  $v_2 = 0$ ) è detto *punto all'infinito* (o *punto improprio*) della retta proiettiva, e verrà indicato con  $\infty$ .

**Osservazione 10C.2**

Sia  $f: \mathbb{R} \rightarrow \mathbb{P}^1(\mathbb{R})$  data da  $f(x) = [x : 1]$ . Si vede subito che  $f$  è iniettiva e che  $f(\mathbb{R}) = \mathbb{P}^1(\mathbb{R}) \setminus \{\infty\}$ . Quindi in un certo senso la retta proiettiva può venire considerata come la retta reale a cui si è aggiunto un punto.

Nella corrispondenza biunivoca fra  $\hat{r}_1$  e  $\mathbb{P}^1(\mathbb{R})$  ottenuta tramite il fascio  $\mathcal{F}$ , al punto  $(x, 1) \in r_1$  viene associato il punto  $[x : 1]$ , e al punto all'infinito il punto  $[1 : 0]$ ; quindi  $\mathbb{P}^1(\mathbb{R})$  riproduce esattamente la struttura dell'insieme  $\hat{r}_1$ . In particolare, la coordinata affine  $x$  non è altro che la coordinata non omogenea della retta proiettiva; se vogliamo considerare tutti i punti di  $\hat{r}_1$  allo stesso modo, senza fastidiose eccezioni, conviene quindi utilizzare coordinate omogenee.

Torniamo ora alla nostra proiezione centrale  $\psi_{P_0}: \hat{r}_1 \rightarrow \hat{r}_2$ , espressa in coordinate non omogenee dalla (10C.6). Indicando con  $[x_1 : x_2]$  le coordinate omogenee in  $\hat{r}_1$  e con  $[y_1 : y_2]$  le coordinate omogenee in  $\hat{r}_2$ , la (10C.6) diventa

$$\frac{y_1}{y_2} = \frac{\alpha(x_1/x_2) + \beta}{\gamma(x_1/x_2) + \delta} = \frac{\alpha x_1 + \beta x_2}{\gamma x_1 + \delta x_2},$$

ovvero

$$\psi_{P_0}([x_1 : x_2]) = [\alpha x_1 + \beta x_2 : \gamma x_1 + \delta x_2], \quad (10C.8)$$

e questa espressione dà il risultato corretto (controlla anche per i punti all'infinito! Inoltre, se consideriamo la matrice (non singolare; esercizio)

$$A = \begin{vmatrix} \alpha & \beta \\ \gamma & \delta \end{vmatrix},$$

vediamo che  $\psi_{P_0}(\pi(v)) = \pi(L_A(v))$  per ogni  $v \in \mathbb{R}^2 \setminus \{O\}$ , dove  $\pi: \mathbb{R}^2 \setminus \{O\} \rightarrow \mathbb{P}^1(\mathbb{R})$  è l'applicazione quoziante. In altre parole, la proiezione centrale è ottenibile tramite un'applicazione lineare invertibile di  $\mathbb{R}^2$ .

**Definizione 10C.7**

Una *proiettività*  $\psi: \mathbb{P}^1(\mathbb{R}) \rightarrow \mathbb{P}^1(\mathbb{R})$  della retta proiettiva è un'applicazione della forma (10C.8) con  $\alpha\delta - \beta\gamma \neq 0$ , cioè tale che  $A \in GL_2(\mathbb{R})$ .

**Osservazione 10C.3**

Le affinità possono essere viste come particolari proiettività: quelle della forma (10C.8) con  $\gamma = 0$ , che mandano cioè il punto all'infinito in sé.

Dunque abbiamo completamente chiarito la struttura di una proiezione centrale: è una proiettività. Vogliamo ora esaminare brevemente quali proprietà geometriche possiede.

Abbiamo visto che le affinità conservano i rapporti semplici; e le proiettività? Sia  $\psi: \mathbb{P}^1(\mathbb{R}) \rightarrow \mathbb{P}^1(\mathbb{R})$  una proiettività, e consideriamo tre punti propri  $P_1, P_2$  e  $P_3$ , di coordinate non omogenee  $x_1, x_2$  e  $x_3$  rispettivamente. Supponiamo per il momento anche che le immagini dei  $P_j$  siano punti propri, in modo da poter utilizzare (10C.6). Allora un conto veloce mostra che

$$(\psi(P_1), \psi(P_2), \psi(P_3)) = \frac{\gamma x_2 + \delta}{\gamma x_3 + \delta} (P_1, P_2, P_3).$$

Dunque il rapporto semplice non è conservato, ma viene moltiplicato per un fattore. Viene dunque la tentazione di prendere un quarto punto  $P_4$  e considerare

$$\frac{(P_1, P_3, P_4)}{(P_2, P_3, P_4)} = \frac{x_4 - x_2}{x_3 - x_2} \Bigg/ \frac{x_4 - x_1}{x_3 - x_1}. \quad (10C.9)$$

Se riscriviamo questa formula in coordinate omogenee otteniamo la seguente definizione.

### Definizione 10C.8

Siano  $P_1, P_2, P_3, P_4 \in \mathbb{P}^1(\mathbb{R})$  quattro punti distinti. Supponiamo che le coordinate omogenee di  $P_j$  siano date da  $v_j \in \mathbb{R}^2 \setminus \{O\}$ , per  $j = 1, \dots, 4$ . Allora il *birapporto*  $[P_1, P_2, P_3, P_4]$  dei quattro punti (nell'ordine) è

$$[P_1, P_2, P_3, P_4] = \frac{\det |v_2 \ v_4|}{\det |v_2 \ v_3|} \Bigg/ \frac{\det |v_1 \ v_4|}{\det |v_1 \ v_3|}. \quad (10C.10)$$

Ponendo inoltre

$$[P_1, P_1, P_3, P_4] = [P_1, P_2, P_3, P_3] = 1, \quad [P_1, P_2, P_1, P_4] = [P_1, P_2, P_3, P_2] = 0, \\ [P_1, P_2, P_3, P_1] = [P_1, P_2, P_2, P_4] = \infty,$$

il birapporto è definito per ogni quaterna di punti in cui non ve ne siano tre uguali.

### Osservazione 10C.4

Per utilizzare le coordinate omogenee nella definizione di oggetti sulla retta proiettiva bisogna fare attenzione. Se il vettore  $v_0 \in \mathbb{R}^2 \setminus \{O\}$  fornisce coordinate omogenee per il punto  $P_0 \in \mathbb{P}^1(\mathbb{R})$  e definiamo qualcosa utilizzando le coordinate di  $v_0$ , questo qualcosa non deve cambiare se moltiplichiamo tutte le coordinate per un numero  $\lambda \neq 0$ , in quanto anche il vettore  $\lambda v_0$  fornisce coordinate omogenee dello stesso punto  $P_0$ . Per esempio, se per  $j = 1, \dots, 4$  prendiamo coordinate omogenee  $v_j$  dei punti  $P_j$  e scegliamo dei numeri  $\lambda_j \neq 0$ , la formula (10C.10) deve dare lo stesso risultato sia usando  $v_j$  sia usando  $\lambda_j v_j$  come coordinate omogenee di  $P_j$ . Fortunatamente si vede subito che abbiamo  $\det |\lambda_h v_h \ \lambda_k v_k| = \lambda_h \lambda_k \det |v_h \ v_k|$ , e quindi (10C.10) fornisce davvero lo stesso risultato per ogni valore di  $\lambda_j \in \mathbb{R}^*$ . Invece, il "rapporto semplice"  $\det |v_1 \ v_3| / \det |v_1 \ v_2|$  non è ben definito su  $\mathbb{P}^1(\mathbb{R})$  in quanto se  $\lambda_2 \neq \lambda_3$  si ha

$$\frac{\det |\lambda_1 v_1 \ \lambda_3 v_3|}{\det |\lambda_1 v_1 \ \lambda_2 v_2|} = \frac{\lambda_3}{\lambda_2} \frac{\det |v_1 \ v_3|}{\det |v_1 \ v_2|} \neq \frac{\det |v_1 \ v_3|}{\det |v_1 \ v_2|}.$$

Le proiettività conservano il birapporto. Infatti la (10C.8) ci dice che se  $v_j$  sono le coordinate omogenee di  $P_j$ , allora le coordinate omogenee di  $\psi(P_j)$  sono  $A v_j$ . Siccome (esercizio)  $\det |A v_h \ A v_k| = (\det A) \det |v_h \ v_k|$ , il birapporto è conservato.

Le proprietà del birapporto sono molte; per esempio, si può dimostrare (esercizio; Proposizione 10C.2) che un'applicazione bigettiva  $\psi: \mathbb{P}^1(\mathbb{R}) \rightarrow \mathbb{P}^1(\mathbb{R})$  è una proiezione se e solo se conserva il birapporto. Ma a noi preme procedere in una direzione diversa per giungere a definire il piano proiettivo.

Lo studio delle proiezioni centrali ci ha portato ad aggiungere a ogni retta di un punto all'infinito, trasformandole tutte in rette proiettive. Ma è necessario aggiungere un punto diverso per ogni retta? Riesaminiamo la corrispondenza fra il fascio di rette  $r_1$  che ci aveva indirizzato alla struttura proiettiva di  $\mathbb{P}^1(\mathbb{R})$ . Se prendiamo una retta  $r'_1$  parallela a  $r_1$  e ripetiamo il procedimento, vediamo che la retta di  $\mathcal{F}$  associata al punto all'infinito di  $r'_1$  è la stessa che veniva associata al punto all'infinito di  $r_1$ ; il punto all'infinito di  $r'_1$  è "lo stesso" di quello di  $r_1$ ; in un certo senso stiamo dicendo che due rette parallele si "intersecano" all'infinito.

Questa idea viene rafforzata ripetendo per il piano la costruzione che ci aveva consigliato per definire  $\mathbb{P}^1(\mathbb{R})$ . Identifichiamo il piano  $\mathbb{R}^2$  col piano affine  $\pi_0$  di  $\mathbb{R}^3$  di equazione cartesiana  $z = 1$ . A ogni punto di  $\pi_0$  corrisponde una e una sola retta della stessa fascio di rette  $r_1$  che passa per l'origine; rimangono fuori soltanto le rette del piano  $z = 0$ , che per analogia con la costruzione della retta corrispondono ai punti all'infinito del piano  $\pi_0$ . Ma se andiamo a intersecare lungo due rette parallele  $r_1$  ed  $r_2$  di  $\pi_0$ , in  $\mathcal{S}$  finiamo in entrambi i casi sulla stessa retta, l'unica parallela a  $r_1$  ed  $r_2$ . Quindi di nuovo viene naturale assegnare lo stesso punto all'infinito a rette parallele.

### Definizione 10C.9

Due vettori  $v, w \in \mathbb{R}^3 \setminus \{O\}$  si dicono *equivalenti*, indicato con  $v \sim w$ , se esiste un numero reale  $\lambda$  tale che  $w = \lambda v$ . L'insieme quoziante  $\mathbb{P}^2(\mathbb{R}) = \mathbb{R}^3 \setminus \{O\} / \sim$  si chiama *piano proiettivo*. Se  $\pi: \mathbb{R}^3 \setminus \{O\} \rightarrow \mathbb{P}^2(\mathbb{R})$  è l'applicazione quoziante, per ogni  $(v_1, v_2, v_3) \in \mathbb{R}^3 \setminus \{O\}$  scriveremo  $[v_1: v_2: v_3]$  al posto di  $\pi(v_1, v_2, v_3)$ ; la tripla di numeri  $(v_1, v_2, v_3)$  si chiameranno *coordinate omogenee* del punto  $[v_1: v_2: v_3] \in \mathbb{P}^2(\mathbb{R})$ . Se inoltre si ha  $v_3 \neq 0$ , i rapporti  $(v_1/v_3, v_2/v_3)$  sono le *coordinate non omogenee* del punto proprio  $[v_1: v_2: v_3]$ . I punti con  $v_3 = 0$  sono i *punti all'infinito* (o *punti impropri*).

### Osservazione 10C.5

Possiamo immergere  $\mathbb{R}^2$  in  $\mathbb{P}^2(\mathbb{R})$  tramite l'applicazione iniettiva  $f: \mathbb{R}^2 \rightarrow \mathbb{P}^2(\mathbb{R})$  data da  $f(x_1, x_2) = [x_1: x_2: 1]$ . L'immagine di  $f$  è composta dai punti propri di  $\mathbb{P}^2(\mathbb{R})$  che quindi formano una copia di  $\mathbb{R}^2$ . Un'altra applicazione iniettiva  $g: \mathbb{P}^1(\mathbb{R}) \rightarrow \mathbb{P}^2(\mathbb{R})$  è data da  $g([v_1: v_2]) = [v_1: v_2: 0]$ ; l'immagine di  $g$  sono i punti impropri del piano proiettivo che quindi compongono una retta proiettiva. Per questo motivo, l'insieme dei punti propri di  $\mathbb{P}^2(\mathbb{R})$  è detto *retta all'infinito*. In particolare,  $\mathbb{P}^2(\mathbb{R})$  è l'unione di un piano affine, una retta proiettiva e di un punto.

Identifichiamo il piano  $\mathbb{R}^2$  con l'insieme dei punti propri di  $\mathbb{P}^2(\mathbb{R})$ , e sia  $r_1$  una retta di  $\mathbb{R}^2$  data da  $ax_1 + bx_2 + c = 0$ . Il punto proprio  $[v_1: v_2: v_3] \in \mathbb{P}^2(\mathbb{R})$  giace in  $r_1$  se e solo se  $a(v_1/v_3) + b(v_2/v_3) + c = 0$ , cioè se e solo se  $av_1 + bv_2 + cv_3 = 0$ . Questa equazione è soddisfatta anche da uno (e da uno solo) punto improprio: il punto  $[-b: a: 0]$ , chiamato punto all'infinito di  $r_1$ . Una retta  $r'_1$  parallela a  $r_1$  ha equazione  $ax_1 + bx_2 + c' = 0$  per qualche  $c' \in \mathbb{R}$  (Esercizio 10.13), per cui il suo punto all'infinito è ancora  $[-b: a: 0]$ . Due rette parallele hanno lo stesso punto all'infinito. Quindi *due rette del piano proiettivo si intersecano sempre*.

## Esercizi

10

10.1 Dimostra che due equazioni cartesiane  $ax + by + cz = d$  e  $a'x + b'y + c'z = d'$  descrivono lo stesso piano se e solo se esiste  $\lambda \in \mathbb{R}$  non nullo tale che  $a' = \lambda a$ ,  $b' = \lambda b$ ,  $c' = \lambda c$  e  $d' = \lambda d$ .

10.2 Quale insieme viene descritto da equazioni cartesiane del tipo  $Ax = b$ , dove  $A \in GL_3(\mathbb{R})$  è una matrice invertibile,  $x \in \mathbb{R}^3$  e  $b \in \mathbb{R}^3$ ?

10.3 Dimostra che le equazioni parametriche  $x = 1 + t$ ,  $y = -2 - t$ ,  $z = 3 + 2t$ , e  $x = 4 - 2s$ ,  $y = -5 + 2s$ ,  $z = 9 - 4s$ , descrivono la stessa retta.

10.4 Se  $ax + by = c$  è l'equazione cartesiana della retta  $r$  nel piano, dimostra che  $\begin{vmatrix} -b \\ a \end{vmatrix}$  è un vettore direttore per  $r$ .

10.5 Se  $x = x_0 + lt$ ,  $y = y_0 + mt$ , sono equazioni parametriche per la retta  $r$  nel piano dimostra che  $mx - ly = mx_0 - ly_0$  è un'equazione cartesiana per  $r$ .

10.6 Se

$$\begin{cases} ax + by + cz = d, \\ a'x + b'y + c'z = d', \end{cases}$$

sono equazioni cartesiane per una retta  $r$  nello spazio, dimostra che

$$l = \det \begin{vmatrix} b & c \\ b' & c' \end{vmatrix}, \quad m = -\det \begin{vmatrix} a & c \\ a' & c' \end{vmatrix}, \quad n = \det \begin{vmatrix} a & b \\ a' & b' \end{vmatrix}$$

sono parametri direttori per  $r$ .

10.7 Se

$$\begin{cases} x = x_0 + sl + tl', \\ y = y_0 + sm + tm', \\ z = z_0 + sn + tn', \end{cases}$$

sono equazioni parametriche per un piano  $\pi$ , dimostra che

$$a = \det \begin{vmatrix} m & m' \\ n & n' \end{vmatrix}, \quad b = -\det \begin{vmatrix} l & l' \\ n & n' \end{vmatrix}, \quad c = \det \begin{vmatrix} l & l' \\ m & m' \end{vmatrix}$$

sono parametri di giacitura per  $\pi$ .

10.8 Dimostra che tre punti del piano  $P_0 = (x_0, y_0)$ ,  $P_1 = (x_1, y_1)$  e  $P_2 = (x_2, y_2)$  sono allineati se e solo se

$$\det \begin{vmatrix} x_1 - x_0 & x_2 - x_0 \\ y_1 - y_0 & y_2 - y_0 \end{vmatrix} = 0.$$

(Suggerimento: i tre punti sono allineati se e solo se esistono  $v \in \mathbb{R}^2$  e  $t_1, t_2 \in \mathbb{R}$  tali che  $P_j - P_0 = t_j v$  per  $j = 1, 2$ .)

- 10.9** Dimostra che i punti  $P_0 = (x_0, y_0, z_0)$ ,  $P_1 = (x_1, y_1, z_1)$  e  $P_2 = (x_2, y_2, z_2)$  dello spazio sono allineati se e solo se

$$\operatorname{rg} \begin{vmatrix} x_1 - x_0 & x_2 - x_0 \\ y_1 - y_0 & y_2 - y_0 \\ z_1 - z_0 & z_2 - z_0 \end{vmatrix} \leq 1.$$

- 10.10** Dimostra che i quattro punti  $P_0 = (x_0, y_0, z_0)$ ,  $P_1 = (x_1, y_1, z_1)$ ,  $P_2 = (x_2, y_2, z_2)$  e  $P_3 = (x_3, y_3, z_3)$  dello spazio sono complanari se e solo se

$$\det \begin{vmatrix} x_1 - x_0 & x_2 - x_0 & x_3 - x_0 \\ y_1 - y_0 & y_2 - y_0 & y_3 - y_0 \\ z_1 - z_0 & z_2 - z_0 & z_3 - z_0 \end{vmatrix} = 0.$$

- 10.11** Verifica che le definizioni di parallelismo fra rette e/o piani date nelle Definizioni 10.6, 10.8 e nel Paragrafo 10.5 sono casi particolari della Definizione 6.5 di parallelismo fra sottospazi affini.

- 10.12** L'insieme delle rette del piano parallele a una retta data  $r$  si chiama *fascio improprio di rette di asse r*. Se la retta  $r$  ha equazioni parametriche  $P = P_0 + t v_0$ , dimostra che le rette di equazioni parametriche  $P = Q + t v_0$  al variare di  $Q \in \mathbb{R}^2$  sono tutte e sole le rette del fascio improprio di asse  $r$ .

- 10.13** Se la retta  $r$  del piano ha equazioni cartesiane  $ax + by = c$ , dimostra che le rette di equazioni cartesiane  $ax + by = k$  al variare di  $k \in \mathbb{R}$  sono tutte e sole le rette del fascio improprio di asse  $r$  (vedi l'esercizio precedente).

- 10.14** Dimostra che il sistema

$$\begin{cases} m(x - x_0) = l(y - y_0), \\ n(y - y_0) = m(z - z_0), \end{cases}$$

al variare di  $l, m, n \in \mathbb{R}$  descrive tutte le rette della stella di centro  $P_0 = (x_0, y_0, z_0)$  escluse quelle parallele all'asse  $z$  e all'asse  $x$ .

- 10.15** Nel calcolo delle equazioni parametriche di una retta per due punti nello spazio, se  $x_1 = x_0$  il sistema (10.16) ci dice solo  $l = 0$ , lasciando  $m$  ed  $n$  liberi. Eppure anche in questo caso la retta passante per i due punti è unica, ed è data ancora da (10.17). Cosa sta succedendo, e perché?

- 10.16** Dimostra che la retta passante per i punti  $P_0 = (\pi, 1, 3)$  e  $P_1 = (\pi, 0, 2)$  dello spazio ha equazioni parametriche  $x = \pi$ ,  $y = 1 - t$ ,  $z = 3 - t$ , ed equazioni cartesiane  $x = \pi$ ,  $y - z = -2$ .

- 10.17** Trova equazioni parametriche e cartesiane del piano in  $\mathcal{A}^3$  passante per i punti  $P_0 = (0, 1, 0)$ ,  $P_1 = (-1, 0, 0)$  e  $P_2 = (0, 0, 1)$ .

- 10.18** Perché nelle tabelle del Paragrafo 10.4 due caselle descrivono situazioni impossibili?

- 10.19** Date le rette di equazioni parametriche

$$r_0 : \begin{cases} x = 3 + t, \\ y = -1 - t, \\ z = 2 + 2t, \end{cases} \quad r_1 : \begin{cases} x = 2 - t, \\ y = 1 + t, \\ z = 2 + t, \end{cases}$$

decidi se sono coincidenti, parallele, incidenti o sghembe.

- 10.20** Siano  $r_1$  ed  $r_2$  due rette del piano di equazioni cartesiane rispettivamente  $a_1x + b_1y = c_1$  e  $a_2x + b_2y = c_2$ . Dimostra che l'insieme delle rette di equazione cartesiana  $\lambda(a_1x + b_1y) + \mu(a_2x + b_2y) = \lambda c_1 + \mu c_2$  al variare di  $\lambda, \mu \in \mathbb{R}$ , con  $(\lambda, \mu) \neq (0, 0)$ , è il fascio di centro il punto d'intersezione di  $r_1$  ed  $r_2$  se le due rette s'intersecano, o il fascio improprio di asse  $r_1$  (vedi l'Esercizio 10.12) se le due rette sono parallele.
- 10.21** Date le rette di equazioni cartesiane  $x + y = 0$ ,  $x + 2y + z = 0$ , e  $x - y = -1$ ,  $x + 5y + z = 0$ , stabilisci se sono parallele, incidenti, sghembe o coincidenti.
- 10.22** Dati il piano  $\pi$  di equazioni cartesiane  $2x + 3y - z = 4$  e la retta  $r$  di equazioni cartesiane  $x + y = 0$ ,  $x - z = 1$ , determina se sono paralleli, una contenuta nell'altro o incidenti e trova l'eventuale punto di intersezione.
- 10.23** L'insieme dei piani paralleli a un piano dato  $\pi$  si chiama *fascio improprio di giacitura*  $\pi$ . Supponiamo che  $\pi$  abbia equazioni parametriche  $P = P_0 + sv_1 + tv_2$  ed equazione cartesiana  $ax + by + cz = d$ . Dimostra che i piani del fascio improprio di giacitura  $\pi$  sono descritti dalle equazioni parametriche  $P = Q + sv_1 + tv_2$  al variare di  $Q \in \mathbb{R}^3$ , e dall'equazione cartesiana  $ax + by + cz = k$  al variare di  $k \in \mathbb{R}$ .
- 10.24** Siano dati due piani  $\pi_1$  e  $\pi_2$  di equazioni cartesiane  $a_1x + b_1y + c_1z = d_1$  e  $a_2x + b_2y + c_2z = d_2$  rispettivamente. Dimostra che l'insieme dei piani di equazione cartesiana  $\lambda(a_1x + b_1y + c_1z) + \mu(a_2x + b_2y + c_2z) = \lambda d_1 + \mu d_2$  al variare di  $\lambda, \mu \in \mathbb{R}$ , con  $(\lambda, \mu) \neq (0, 0)$ , è il fascio di asse la retta d'intersezione di  $\pi_1$  e  $\pi_2$  se i due piani s'intersecano, o il fascio improprio (vedi l'esercizio precedente) di giacitura  $\pi_1$  (o  $\pi_2$ ) se i due piani sono paralleli.
- 10.25** Trova equazioni parametriche e cartesiane del piano in  $\mathcal{A}^3$  passante per il punto  $P_1 = (1, 0, 1)$  e contenente la retta di equazioni cartesiane  $x + y + z = 3$ ,  $x - y - z = 0$ .
- 10.26** Trova le equazioni cartesiane e parametriche del piano (se esiste) che contiene le rette di equazioni cartesiane

$$\begin{cases} 3x - y + z = 0, \\ x + y + z = 0, \end{cases} \quad \text{e} \quad \begin{cases} -x - y + 2z = 0, \\ x - z = 0. \end{cases}$$

- 10.27** Trova equazioni parametriche e cartesiane della retta passante per il punto  $P = (0, 7, 1)$  e parallela alla retta di equazioni  $3x + y + z = 4$ ,  $2x + y + z = 4$ .

**10.28** Sia fissato un sistema di riferimento affine  $\mathcal{R}_0$  in  $\mathcal{A}^3$ , e considera i punti di coordinate affini (rispetto a  $\mathcal{R}_0$ )

$$O' = \begin{vmatrix} 1 \\ 0 \\ 1 \end{vmatrix}, A'_1 = \begin{vmatrix} 0 \\ 1 \\ 1 \end{vmatrix}, A'_2 = \begin{vmatrix} 1 \\ 1 \\ 0 \end{vmatrix}, A'_3 = \begin{vmatrix} 1 \\ 1 \\ 1 \end{vmatrix},$$

$$O'' = \begin{vmatrix} 0 \\ 2 \\ 0 \end{vmatrix}, A''_1 = \begin{vmatrix} -1 \\ 1 \\ -1 \end{vmatrix}, A''_2 = \begin{vmatrix} 1 \\ 0 \\ 2 \end{vmatrix}, A''_3 = \begin{vmatrix} 1 \\ 0 \\ 0 \end{vmatrix}.$$

Dimostra che  $\mathcal{R}' = RA(O', A'_1, A'_2, A'_3)$  e  $\mathcal{R}'' = RA(O'', A''_1, A''_2, A''_3)$  sono sistemi di riferimento affini, trova la formula di cambiamento di coordinate affini da  $\mathcal{R}'$  a  $\mathcal{R}''$  e determina quali dei tre sistemi  $\mathcal{R}_0$ ,  $\mathcal{R}'$  ed  $\mathcal{R}''$  hanno la stessa orientazione.

**10.29** Due basi di uno spazio vettoriale  $V$  hanno la *stessa orientazione* se la matrice di cambiamento di base ha determinante positivo; hanno *orientazione opposta* altrimenti. Dimostra che “avere la stessa orientazione” è una relazione d’equivalenza sull’insieme delle basi di  $V$ , e che l’insieme quoziante contiene due elementi (vedi il Paragrafo 2C.1 per le necessarie definizioni).

# Prodotti scalari

# 11

## Sommario

- 11.1 Definizioni ed esempi
- 11.2 Disuguaglianza di Cauchy-Schwarz
- 11.3 Basi ortogonali
- 11.4 Proiezioni ortogonali
- 11.5 Prodotti scalari e matrici
- 11.6 Endomorfismi simmetrici e isometrie

## Esercizi

E finalmente giunto il momento di occuparci di distanze e angoli, ma senza dimenticarci quanto visto finora: vogliamo trattare questioni metriche con tecniche lineari, in modo da poterle inserire facilmente nel contesto degli spazi vettoriali. Questo ci porterà a introdurre il concetto di prodotto scalare che è l'argomento principale di questo capitolo. Dopo una lunga serie di definizioni ed esempi, ci concentreremo sui prodotti scalari definiti positivi, dimostrando la disuguaglianza di Cauchy-Schwarz, introducendo il concetto di base ortonormale e descrivendo il procedimento di ortogonalizzazione di Gram-Schmidt. Dopo aver definito l'ortogonale di un sottospazio e le proiezioni ortogonali, mostreremo come rappresentare un prodotto scalare rispetto a una base tramite una matrice. Infine studieremo le isometrie lineari e gli endomorfismi simmetrici.

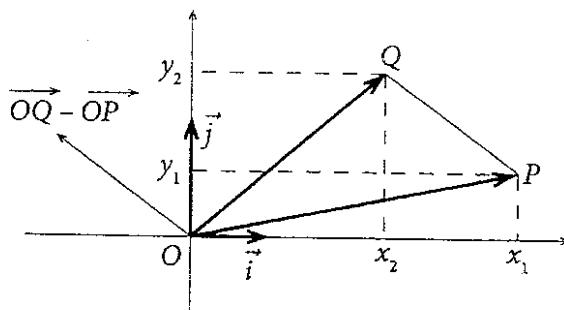
## 11.1 Definizioni ed esempi

Cominciamo a vedere come si può procedere in  $\mathcal{A}^2$ . In questo caso, dovresti già avere un'idea di cos'è la distanza fra due punti, e l'angolo fra due direzioni; il nostro obiettivo è trovare un modo per calcolarli. Prima di tutto, fissiamo un sistema di riferimento affine  $RA(O, \vec{i}, \vec{j})$ , in modo da poter lavorare in  $\mathbb{R}^2$ ; inoltre supponiamo che i vettori  $\vec{i}$  e  $\vec{j}$  siano di lunghezza 1 e formino un angolo di  $\pi/2$  radianti.

Prendiamo un punto  $P$  del piano, di coordinate  $v_1 = (x_1, y_1) \in \mathbb{R}^2$ . Il Teorema di Pitagora (Teorema 1.2) ci dice che la distanza  $d(O, P)$  di  $P$  dall'origine  $O$ , cioè la lunghezza del vettore  $\overrightarrow{OP}$ , è data da

$$d(O, P) = \sqrt{x_1^2 + y_1^2}. \quad (11.1)$$

Spesso scriveremo  $\|\overrightarrow{OP}\|$  o anche  $\|v_1\|$  per indicare la distanza  $d(O, P)$ . Il numero  $\|v_1\|$  viene detto *norma* di  $v_1$ ; la doppia sbarra serve a ricordare che si tratta di una generalizzazione del modulo di un numero reale (o di un numero complesso).



**Figura 11.1** - La distanza fra due punti.

Il Teorema di Pitagora ci dice anche (vedi la Figura 11.1) come calcolare la distanza  $d(P, Q)$  fra  $P$  e un altro punto  $Q$  del piano, di coordinate  $v_2 = (x_2, y_2) \in \mathbb{R}^2$ :

$$d(P, Q) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} = \|v_2 - v_1\| = \|\overrightarrow{OQ} - \overrightarrow{OP}\|. \quad (11.2)$$

### Osservazione 11.1

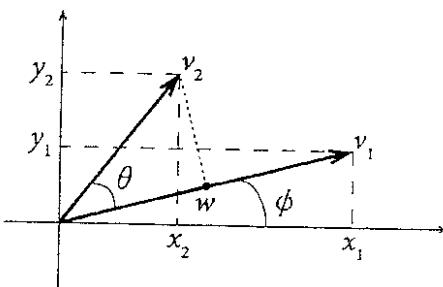
Attenzione a un particolare apparentemente secondario ma importante: le formule (11.1) e (11.2) ci forniscono l'usuale distanza euclidea del piano *soltanto perché i vettori di base del sistema di riferimento affine scelto sono di lunghezza unitaria e perpendicolari fra loro*. Se non fosse stato così, potevamo ancora identificare  $\mathcal{A}^2$  con  $\mathbb{R}^2$  come sempre, e introdurre una "distanza" fra punti con la formula (11.2); ma questa distanza non sarebbe stata quella euclidea. Per esempio, come illustrato nell'Esercizio 11.2, l'insieme dei punti di distanza uno dall'origine invece di una circonferenza potrebbe essere un'ellisse.

La formula (11.2) non sembra, almeno per ora, dare indicazioni su come esprimere le distanze tramite oggetti lineari; vediamo se con gli angoli ci andrà meglio.

### Osservazione 11.2

Per noi, l'angolo fra due direzioni (o rette, o segmenti, o vettori) sarà sempre *convesso*, e mai concavo, per cui varierà fra 0 e  $\pi$  (inclusi). In particolare, siccome il coseno nell'intervallo  $[0, \pi]$  è una funzione iniettiva, per trovare l'angolo fra due vettori basterà calcolarne il coseno.

Consideriamo due vettori in  $\mathbb{R}^2$ , di coordinate  $v_1 = (x_1, y_1)$  e  $v_2 = (x_2, y_2)$ ; vogliamo individuare l'angolo  $\theta$  che determinano. Per quanto detto sopra, sarà sufficiente trovare  $\cos \theta$  (vedi la Figura 11.2).



**Figura 11.2** - L'angolo fra due vettori.

Il calcolo di  $\cos \theta$  non è immediato; la trigonometria invece ci dice come trovare facilmente l'angolo  $\psi$  fra l'asse  $x$  e  $v_1$ , e l'angolo  $\psi + \theta$  fra l'asse  $x$  e  $v_2$ . Infatti si vede subito che

$$\begin{cases} \cos \psi = \frac{x_1}{\|v_1\|}, \\ \sin \psi = \frac{y_1}{\|v_1\|}, \end{cases} \quad \text{e} \quad \begin{cases} \cos(\psi + \theta) = \frac{x_2}{\|v_2\|}, \\ \sin(\psi + \theta) = \frac{y_2}{\|v_2\|}. \end{cases} \quad (11.3)$$

Le formule d'addizione per le funzioni trigonometriche trasformano (11.3) in

$$\begin{cases} \cos \psi \cos \theta - \sin \psi \sin \theta = \frac{x_2}{\|v_2\|}, \\ \sin \psi \cos \theta + \cos \psi \sin \theta = \frac{y_2}{\|v_2\|}, \end{cases}$$

che è un sistema lineare quadrato nelle incognite  $\cos \theta$  e  $\sin \theta$ . La matrice dei coefficienti di questo sistema ha determinante  $\cos^2 \psi + \sin^2 \psi = 1$ , per cui il sistema ha un'unica soluzione, che possiamo calcolare per esempio col Teorema di Cramer (Corollario 9.11):

$$\cos \theta = \det \begin{vmatrix} x_2/\|v_2\| & -\sin \psi \\ y_2/\|v_2\| & \cos \psi \end{vmatrix} = \det \begin{vmatrix} x_2/\|v_2\| & -y_1/\|v_1\| \\ y_2/\|v_2\| & x_1/\|v_1\| \end{vmatrix} = \frac{x_1 x_2 + y_1 y_2}{\|v_1\| \|v_2\|}.$$

Viene quindi naturale associare ai vettori  $v_1$  e  $v_2$  il numero

$$\langle v_1, v_2 \rangle = x_1 x_2 + y_1 y_2, \quad (11.4)$$

per poter poi scrivere

$$\cos \theta = \frac{\langle v_1, v_2 \rangle}{\|v_1\| \|v_2\|}.$$

Ora,  $\langle \cdot, \cdot \rangle$  è lineare in entrambe le variabili. Inoltre, anche la norma può venire definita tramite questo nuovo oggetto: infatti  $\|v\| = \sqrt{\langle v, v \rangle}$  per qualunque vettore  $v \in \mathbb{R}^2$ , come avrai cura di verificare. Dunque sembra proprio che abbiamo trovato l'oggetto lineare che stavamo cercando.

### Osservazione 11.3

Ci sono altri concetti di geometria metrica del piano che si possono esprimere facilmente usando  $\langle \cdot, \cdot \rangle$ . Per esempio, due vettori  $v_1$  e  $v_2$  sono ortogonali se e solo se formano un

angolo  $\theta$  di  $\pi/2$  radianti, e quindi se e solo se  $\cos \theta = 0$ , ovvero se e solo se  $\langle v_1, v_2 \rangle = 0$ . Analogamente, sono linearmente dipendenti se e solo se formano un angolo  $\theta$  di  $0$  o  $\pi$  radianti, e quindi se e solo se  $\cos \theta = \pm 1$ , ovvero se e solo se  $\langle v_1, v_2 \rangle = \pm \|v_1\| \|v_2\|$ . Un altro concetto che possiamo esprimere tramite  $\langle \cdot, \cdot \rangle$  è la *proiezione ortogonale* di  $v_2$  su  $v_1$ . Si tratta del vettore  $w$  ottenuto tracciando da  $v_2$  la retta perpendicolare a  $v_1$  (vedi la Figura 11.2), per cui è un multiplo di  $v_1/\|v_1\|$ , e ha lunghezza  $\|v_2\| \cos \theta$ . Dunque

$$w = (\|v_2\| \cos \theta) \frac{v_1}{\|v_1\|} = \frac{\langle v_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1.$$

Potremmo ora passare ad  $\mathcal{A}^3$  (o meglio  $\mathbb{R}^3$ ), facendo considerazioni analoghe a quelle sopra. Ma tanto vale fare il passo tutto in una volta e introdurre la seguente definizione

### Definizione 11.1

Il *prodotto scalare canonico*  $\langle \cdot, \cdot \rangle: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  su  $\mathbb{R}^n$  è la funzione data da

$$\langle v, w \rangle = v_1 w_1 + \cdots + v_n w_n = w^T \cdot v \quad (11.5)$$

per ogni  $v = (v_1, \dots, v_n)^T$ ,  $w = (w_1, \dots, w_n)^T \in \mathbb{R}^n$ , dove il prodotto all'ultimo membro è il solito prodotto righe per colonne. La *norma* associata al prodotto scalare canonico è la funzione  $\|\cdot\|: \mathbb{R}^n \rightarrow \mathbb{R}^+$  data da

$$\forall v \in \mathbb{R}^n \quad \|v\| = \sqrt{\langle v, v \rangle} = (v_1^2 + \cdots + v_n^2)^{1/2}. \quad (11.6)$$

Nota che la norma è ben definita in quanto  $\langle v, v \rangle = v_1^2 + \cdots + v_n^2 \geq 0$  per tutti vettori  $v \in \mathbb{R}^n$ .

Il prodotto scalare canonico generalizza (11.4) a dimensione qualunque e gode di una serie di proprietà importanti di verifica elementare (lasciata per esercizio):

- (a<sub>1</sub>) per ogni  $v_1, v_2$  e  $w \in \mathbb{R}^n$  si ha  $\langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle$  (*additività rispetto alla prima variabile*);
- (a<sub>2</sub>) per ogni  $v, w \in \mathbb{R}^n$  e  $\lambda \in \mathbb{R}$  si ha  $\langle \lambda v, w \rangle = \lambda \langle v, w \rangle$  (*omogeneità rispetto alla prima variabile*);
- (b<sub>1</sub>) per ogni  $v, w_1$  e  $w_2 \in \mathbb{R}^n$  si ha  $\langle v, w_1 + w_2 \rangle = \langle v, w_1 \rangle + \langle v, w_2 \rangle$  (*additività rispetto alla seconda variabile*);
- (b<sub>2</sub>) per ogni  $v, w \in \mathbb{R}^n$  e  $\lambda \in \mathbb{R}$  si ha  $\langle v, \lambda w \rangle = \lambda \langle v, w \rangle$  (*omogeneità rispetto alla seconda variabile*).

Le proprietà (a<sub>1</sub>)–(a<sub>2</sub>) si riassumono dicendo che il prodotto scalare canonico è *lineare rispetto alla prima variabile*; le (b<sub>1</sub>)–(b<sub>2</sub>) dicendo che è *lineare rispetto alla seconda variabile*; tutte e quattro dicendo che è *bilineare*. Inoltre:

- (c) per ogni  $v, w \in \mathbb{R}^n$  si ha  $\langle v, w \rangle = \langle w, v \rangle$  (*simmetria*);
- (d) si ha  $\langle v, w \rangle = 0$  per ogni  $w \in V$  se e solo se  $v = O$  (il prodotto scalare canonico è *non-degenero*);
- (e) per tutti i  $v \in \mathbb{R}^n$  con  $v \neq O$  si ha  $\langle v, v \rangle > 0$  (il prodotto scalare canonico è *definito positivo*).

Usando solo queste proprietà è possibile dimostrare che la distanza fra due punti definita da  $\|v - w\|$  e l'angolo fra due vettori definito da  $\cos \theta = \langle v, w \rangle / \|v\| \|w\|$  godono di tutti

le proprietà usuali<sup>(1)</sup>. Le parole chiave qui sono “usando solo queste proprietà”. In altri termini, non appena abbiamo un’applicazione  $\langle \cdot, \cdot \rangle$  che soddisfa le proprietà (a<sub>1</sub>)–(e) possiamo ricostruire la geometria euclidea. Ma allora, perché limitarci al solo prodotto scalare canonico? E, soprattutto, perché limitarci al solo  $\mathbb{R}^n$  e non invece farci prendere dall’ebbrezza di calcolare la distanza fra due polinomi, l’angolo fra due applicazioni lineari, la proiezione ortogonale di una matrice? Presentiamo quindi subito una serie di definizioni.

### Definizione 11.2

Sia dato uno spazio vettoriale  $V$  su  $\mathbb{R}$ . Una *forma bilineare* su  $V$  è un’applicazione  $g: V \times V \rightarrow \mathbb{R}$  tale che

$$\begin{aligned} g(v_1 + v_2, w) &= g(v_1, w) + g(v_2, w), & g(v, w_1 + w_2) &= g(v, w_1) + g(v, w_2), \\ g(\lambda v, w) &= \lambda g(v, w) = g(v, \lambda w), \end{aligned}$$

per tutti i  $v, v_1, v_2, w, w_1, w_2 \in V$  e  $\lambda \in \mathbb{R}$ .

### Definizione 11.3

Sia  $V$  uno spazio vettoriale su  $\mathbb{R}$ . Un *prodotto scalare* su  $V$  è una forma bilineare  $g$  *simmetrica*, cioè tale che  $g(v, w) = g(w, v)$  per ogni  $v, w \in V$ . In questo caso scriveremo sempre  $\langle v, w \rangle$  invece di  $g(v, w)$ .

### Definizione 11.4

Sia  $\langle \cdot, \cdot \rangle$  un prodotto scalare su uno spazio vettoriale  $V$ . Il *nucleo* di  $\langle \cdot, \cdot \rangle$  è l’insieme  $V^\perp$  dei vettori  $v \in V$  tali che  $\langle v, w \rangle = 0$  per ogni  $w \in V$ .

Il nucleo di un prodotto scalare è un sottospazio vettoriale: se  $v_1, v_2, v \in V^\perp$  e  $\lambda \in \mathbb{R}$  si ha infatti  $\langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle = 0$  e  $\langle \lambda v, w \rangle = \lambda \langle v, w \rangle = 0$  per ogni  $w \in V$ , per cui  $v_1 + v_2, \lambda v \in V^\perp$ .

### Definizione 11.5

Sia  $V$  uno spazio vettoriale su  $\mathbb{R}$ . Un prodotto scalare  $\langle \cdot, \cdot \rangle$  su  $V$  si dice *non degenere* se  $V^\perp = \{O\}$ ; *degenero* altrimenti.

Nota che  $\langle O, O \rangle = 0$  per qualunque prodotto scalare (esercizio).

### Definizione 11.6

Sia  $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{R}$  un prodotto scalare su uno spazio vettoriale  $V$ . Questo prodotto scalare è *definito positivo* se  $\langle v, v \rangle > 0$  per qualunque  $v \neq O$ ; *definito negativo* se  $\langle v, v \rangle < 0$  per qualunque  $v \neq O$ ; in entrambi i casi è necessariamente non degenere (perché?). Diremo invece che è *semidefinito positivo* se  $\langle v, v \rangle \geq 0$  per ogni  $v \in V$  ed esiste un  $v_0 \neq O$  con  $\langle v_0, v_0 \rangle = 0$ ; che è *semidefinito negativo* se  $\langle v, v \rangle \leq 0$  per ogni  $v \in V$ , ed esiste un  $v_0 \neq O$  tale che  $\langle v_0, v_0 \rangle = 0$ . Infine, diremo che il prodotto scalare è *indefinito* se non è né (semi)definito positivo né (semi)definito negativo, cioè in altre parole se esistono  $v, w \in V$  tali che  $\langle v, v \rangle > 0$  e  $\langle w, w \rangle < 0$ .

Vediamo innanzitutto qualche esempio.

<sup>(1)</sup> Nota che bisogna anche dimostrare che  $\langle v, w \rangle / \|v\| \|w\|$  è sempre in modulo minore di 1, in modo da poter essere il coseno di qualcosa. La differenza rispetto al caso di  $\mathcal{A}^2$  è che mentre prima sapevamo *a priori* cos’era l’angolo fra due vettori e dovevamo solo calcolarlo, stavolta vogliamo *definire* l’angolo fra due vettori partendo dal prodotto scalare canonico.

**Esempio 11.1**

Il prodotto scalare canonico è chiaramente un prodotto scalare definito positivo su  $\mathbb{R}^n$ .  $\square$

**Esempio 11.2**

Consideriamo l'applicazione  $\langle \cdot, \cdot \rangle: \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$  data da

$$\langle v, w \rangle = 2v_1w_1 + v_1w_2 + v_2w_1 + v_2w_2 + 3v_3w_3.$$

È facile verificare che si tratta di un prodotto scalare su  $\mathbb{R}^3$ : infatti, se fissiamo  $v$  (ovvero, consideriamo le coordinate di  $v$  come parametri), quello che resta è un polinomio di primo grado senza termini noti nelle coordinate di  $w$ , per cui è lineare in  $w$ ; analogamente fissando  $w$ , quello che resta risulta essere lineare in  $v$ . Anche la simmetria è immediata: infatti, ogni volta che è presente un termine del tipo  $v_iw_j$  è presente anche un termine del tipo  $v_jw_i$ , ed esattamente con lo stesso coefficiente. Nota che queste osservazioni si possono riassumere dicendo che questo prodotto scalare è dato da una formula del tipo  $\langle v, w \rangle = w^T S v$  dove  $S$  è la matrice simmetrica (cioè tale che  $S^T = S$ )

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

Questo prodotto scalare è definito positivo: infatti  $\langle v, v \rangle = v_1^2 + (v_1 + v_2)^2 + 3v_3^2 \geq 0$ , e inoltre  $\langle v, v \rangle = 0$  se e solo se  $v_1 = v_1 + v_2 = v_3 = 0$ , cioè se e solo se  $v = O$ .  $\square$

**Esempio 11.3**

Consideriamo invece l'applicazione  $\langle \cdot, \cdot \rangle: \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$  data da

$$\langle v, w \rangle = v_1w_1 + v_1w_2 + v_2w_1 + v_2w_2 + v_3w_3 = w^T S v,$$

dove stavolta  $S$  è la matrice

$$S = \begin{vmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}.$$

Anche in questo caso è chiaro che si tratta di un prodotto scalare; ma, contrariamente a prima, è solo semidefinito positivo. Infatti,  $\langle v, v \rangle = (v_1 + v_2)^2 + v_3^2 \geq 0$ , ma  $\langle v, v \rangle = 0$  se e solo se  $v_1 + v_2 = v_3 = 0$  che non implica  $v = O$ , in quanto esistono dei vettori non nulli tali che  $\langle v, v \rangle = 0$ ; per esempio,  $v = (1, -1, 0)^T$ .  $\square$

**Esempio 11.4**

L'applicazione  $\langle \cdot, \cdot \rangle: \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$  data da

$$\langle v, w \rangle = -2v_1w_1 + v_1w_2 + v_2w_1 - v_2w_2 - v_3w_3 = w^T S v,$$

dove stavolta  $S$  è la matrice

$$S = \begin{vmatrix} -2 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{vmatrix},$$

è un prodotto scalare, ma definito negativo. Infatti,  $\langle v, v \rangle = -v_1^2 - (v_1 - v_2)^2 - v_3^2$ .  $\square$

**Esempio 11.5**

L'applicazione  $\langle \cdot, \cdot \rangle: \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$  data da

$$\langle v, w \rangle = v_1 w_2 + v_2 w_1 + v_3 w_3 = w^T S v,$$

dove stavolta  $S$  è la matrice

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

è ancora un prodotto scalare, ma indefinito. Infatti,  $\langle v, v \rangle = 2v_1 v_2 + v_3^2$ , e per esempio abbiamo

$$\left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle = 2, \quad \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle = 0, \quad \left\langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\rangle = -2.$$

Pur essendo indefinito, questo prodotto scalare è non degenere: infatti,  $\langle v, w \rangle = 0$  per tutti i  $w \in \mathbb{R}^3$  se e solo se  $v_1 = v_2 = v_3 = 0$ , cioè se e solo se  $v = O$ .  $\square$

**Osservazione 11.4**

Considerare prodotti scalari non necessariamente definiti positivi non è masochismo da matematici. Per fare un esempio, nella relatività ristretta di Einstein è fondamentale considerare nello spazio-tempo  $\mathbb{R}^4$  il prodotto scalare

$$\left\langle \begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ t_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \\ z_2 \\ t_2 \end{pmatrix} \right\rangle = x_1 x_2 + y_1 y_2 + z_1 z_2 - t_1 t_2,$$

che è non degenere e indefinito.

**Esempio 11.6**

Il prodotto scalare probabilmente più importante in assoluto, oltre al prodotto scalare canonico, è il seguente

$$\langle p, q \rangle = \int_0^1 p(t)q(t) dt,$$

dove  $p$  e  $q$  sono due polinomi a coefficienti reali (o anche semplicemente due funzioni continue definite sull'intervallo  $[0, 1]$ ). Le proprietà dell'integrale permettono di dimostrare che si tratta di un prodotto scalare definito positivo.  $\square$

**Esempio 11.7**

Un prodotto scalare definito sullo spazio  $\mathbb{R}_2[t]$  è

$$\langle p, q \rangle = p(0)q(0) + p(1)q(1) + p(-1)q(-1).$$

Dimostriamo che è un prodotto scalare. Siano  $p_1, p_2, q \in \mathbb{R}_2[t]$ ; allora

$$\begin{aligned} \langle p_1 + p_2, q \rangle &= (p_1 + p_2)(0)q(0) + (p_1 + p_2)(1)q(1) + (p_1 + p_2)(-1)q(-1) \\ &= (p_1(0)q(0) + p_1(1)q(1) + p_1(-1)q(-1)) \\ &\quad + (p_2(0)q(0) + p_2(1)q(1) + p_2(-1)q(-1)) = \langle p_1, q \rangle + \langle p_2, q \rangle. \end{aligned}$$

Analogamente si dimostra l'omogeneità rispetto alla prima variabile. Siccome  $\langle \cdot, \cdot \rangle$  è evidentemente simmetrico, la linearità rispetto alla seconda variabile è conseguenza di quanto già visto (perché?), e quindi abbiamo verificato che è un prodotto scalare. Dimostriamo adesso che è anche definito positivo. Infatti, per ogni  $p \in \mathbb{R}_2[t]$  si ha  $\langle p, p \rangle = p(0)^2 + p(1)^2 + p(-1)^2 \geq 0$ , e inoltre  $\langle p, p \rangle = 0$  se e solo se il polinomio  $p$  ha almeno le tre radici 0, 1 e -1. Ma questo può accadere se e solo se  $p \equiv 0$ , in quanto  $p$  è di secondo grado.  $\square$

### Esempio 11.8

Sia  $\langle \cdot, \cdot \rangle: M_{m,n}(\mathbb{R}) \times M_{m,n}(\mathbb{R}) \rightarrow \mathbb{R}$  data da  $\langle A, B \rangle = \text{tr}(B^T A)$ , dove  $\text{tr}: M_{n,n}(\mathbb{R}) \rightarrow \mathbb{R}$  la traccia definita nell'Esercizio 5.10; nota che  $B^T A$  è una matrice quadrata di ordine  $n$ , per cui possiamo calcolarne la traccia. Vogliamo dimostrare che  $\langle \cdot, \cdot \rangle$  è un prodotto scalare definito positivo. Che sia bilineare segue subito dalla linearità della traccia. Poi, è subito evidente che  $\text{tr}(C^T) = \text{tr}(C)$  per qualunque matrice quadrata  $C$ ; quindi  $\langle B, A \rangle = \text{tr}(A^T B) = \text{tr}((A^T B)^T) = \text{tr}(B^T A) = \langle A, B \rangle$ . Dunque è un prodotto scalare. Vediamo ora che è definito positivo; infatti

$$\begin{aligned}\langle A, A \rangle &= \text{tr}(A^T A) = (a_{11}a_{11} + \cdots + a_{m1}a_{m1}) + \cdots + (a_{1n}a_{1n} + \cdots + a_{mn}a_{mn}) \\ &= \sum_{i=1}^m \sum_{j=1}^n a_{ij}^2 \geq 0,\end{aligned}$$

e che inoltre è uguale a zero se e solo se  $A$  è la matrice nulla.  $\square$

### Definizione 11.7

Uno spazio vettoriale metrico è uno spazio vettoriale  $V$  su  $\mathbb{R}$  provvisto di un prodotto scalare  $\langle \cdot, \cdot \rangle$  definito positivo. La norma  $\|\cdot\|: V \rightarrow \mathbb{R}^+$  in questo spazio vettoriale metrico è definita da  $\|v\| = \sqrt{\langle v, v \rangle}$ . La norma  $\|v\|$  di un vettore  $v$  si dice anche lunghezza di  $v$ .

Concludiamo il paragrafo introducendo il prodotto hermitiano canonico, un analogo del prodotto scalare ma definito su  $\mathbb{C}^n$ . In realtà, esiste una teoria sistematica dei prodotti hermitiani su spazi vettoriali complessi qualsiasi, ma per i nostri scopi è sufficiente introdurre un esempio e lasciare alla tua fantasia e intuizione capire come si procede nel caso generale (vedi gli Esercizi 11.27-11.29).

Un primo tentativo naturale per definire un “prodotto scalare canonico” su  $\mathbb{C}^n$  potrebbe essere porre  $\langle v, w \rangle = v_1w_1 + \cdots + v_nw_n$  per ogni  $v, w \in \mathbb{C}^n$ . Questa definizione ha però un problema fondamentale:  $\langle v, v \rangle$  si guarda bene dall'essere sempre positivo. Anzi, può benissimo essere negativo, o addirittura immaginario puro (verificalo su degli esempi). Quindi se volessimo definire una lunghezza come radice quadrata di  $\langle v, v \rangle$  ci troveremmo in grossi guai. Se ci pensi bene, è lo stesso tipo di problema che avevamo avuto nel tentare di generalizzare al caso complesso la Proposizione 5.11. Allora, il problema era stato risolto (nella Proposizione 5.12) ponendo dei coniugati in posti strategici: anche qui possiamo operare nello stesso modo.

### Definizione 11.8

Il prodotto hermitiano canonico  $\langle \cdot, \cdot \rangle: \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}$  su  $\mathbb{C}^n$  è l'applicazione data da

$$\langle v, w \rangle = v_1\overline{w_1} + \cdots + v_n\overline{w_n} = w^H \cdot v \quad (11.7)$$

per ogni  $v = (v_1, \dots, v_n)^T$ ,  $w = (w_1, \dots, w_n)^T \in \mathbb{C}^n$ , dove il prodotto all'ultimo membro è il solito prodotto righe per colonne, e  $w^H = \overline{w^T}$ . La norma associata al prodotto hermitiano canonico è la funzione  $\|\cdot\|: \mathbb{C}^n \rightarrow \mathbb{R}^+$  data da

$$\forall v \in \mathbb{C}^n \quad \|v\| = \sqrt{\langle v, v \rangle} = (\lvert v_1 \rvert^2 + \cdots + \lvert v_n \rvert^2)^{1/2}. \quad (11.8)$$

Nota che la norma adesso è ben definita in quanto  $\langle v, v \rangle = \lvert v_1 \rvert^2 + \cdots + \lvert v_n \rvert^2 \geq 0$  per tutti i vettori  $v \in \mathbb{C}^n$ .

Il prodotto hermitiano canonico gode di proprietà simili a quelle del prodotto scalare canonico, ma non proprio delle stesse: ogni tanto appare qualche coniugato. Per essere precisi, valgono (a<sub>1</sub>), (a<sub>2</sub>) e (b<sub>1</sub>), ma (b<sub>2</sub>) è sostituita da:

(b'<sub>2</sub>) per ogni  $v, w \in \mathbb{C}^n$  e  $\lambda \in \mathbb{C}$  si ha  $\langle v, \lambda w \rangle = \bar{\lambda} \langle v, w \rangle$  (*anti-omogeneità nella seconda variabile*).

Inoltre, valgono la (d) e la (e), mentre la proprietà (c) è sostituita dalla:

(c') per ogni  $v, w \in \mathbb{C}^n$  si ha  $\langle w, v \rangle = \overline{\langle v, w \rangle}$  (*hermitianità*).

## 11.2 Disuguaglianza di Cauchy-Schwarz

Possiamo finalmente iniziare a fare qualcosa di concreto con i prodotti scalari. Cominciamo con un piccolo risultato introduttivo.

### Proposizione 11.1

Sia  $\langle \cdot, \cdot \rangle$  un prodotto scalare su uno spazio vettoriale  $V$ . Allora:

- (1) se  $\langle v, v \rangle \neq 0$  per tutti i  $v \in V$  non nulli, allora  $\langle \cdot, \cdot \rangle$  è non degenere;
- (2) sia  $\langle \cdot, \cdot \rangle$  non degenere. Allora  $\langle v_1, w \rangle = \langle v_2, w \rangle$  per ogni  $w \in V$  se e solo se  $v_1 = v_2$ ;
- (3)  $\langle \cdot, \cdot \rangle$  è (semi)definito negativo se e solo se  $-\langle \cdot, \cdot \rangle$  è (semi)definito positivo;
- (4) se  $\langle \cdot, \cdot \rangle$  è semidefinito positivo o negativo, allora è degenere;
- (5) se  $\langle \cdot, \cdot \rangle$  è indefinito, allora esiste  $v \neq O$  tale che  $\langle v, v \rangle = 0$ .

### Dimostrazione

(1) Se  $\langle \cdot, \cdot \rangle$  fosse degenere, esisterebbe  $v_0 \in V^\perp \setminus \{O\}$ , e quindi in particolare avremmo  $\langle v_0, v_0 \rangle = 0$  con  $v_0 \neq O$ .

(2) Se  $\langle v_1, w \rangle = \langle v_2, w \rangle$ , allora  $\langle v_2 - v_1, w \rangle = 0$ ; siccome questo deve valere per tutti i  $w \in V^\perp = \{O\}$  per ipotesi, ne segue che  $v_2 - v_1 = O$ , cioè  $v_1 = v_2$ .

(3) Ovvio.

(4) Supponiamo  $\langle \cdot, \cdot \rangle$  semidefinito positivo; grazie a (3), il caso semidefinito negativo è identico. Siccome il nostro prodotto scalare è soltanto semidefinito, esiste un  $v_0 \neq O$  tale che  $\langle v_0, v_0 \rangle = 0$ ; vogliamo dimostrare che  $\langle v_0, w \rangle = 0$  per tutti i  $w \in V$ .

Supponiamo per assurdo che esista  $w_0 \in V$  tale che  $\langle v_0, w_0 \rangle \neq 0$ . Il nostro prodotto scalare è semidefinito positivo; quindi per ogni  $t \in \mathbb{R}$  abbiamo

$$\begin{aligned} 0 \leq \langle w_0 + tv_0, w_0 + tv_0 \rangle &= \langle w_0, w_0 \rangle + t[\langle w_0, v_0 \rangle + \langle v_0, w_0 \rangle] + t^2 \langle v_0, v_0 \rangle \\ &= \langle w_0, w_0 \rangle + 2t \langle v_0, w_0 \rangle, \end{aligned}$$

(in quanto  $\langle v_0, v_0 \rangle = 0$  per ipotesi), cioè  $-2 \langle v_0, w_0 \rangle t \leq \langle w_0, w_0 \rangle$  per ogni  $t \in \mathbb{R}$ , impossibile. Quindi un tale  $w_0$  non esiste e il nostro prodotto scalare è degenere.

(5) Siano  $v_0, v_1 \in V$  tali che  $\langle v_0, v_0 \rangle < 0$  e  $\langle v_1, v_1 \rangle > 0$ , e definiamo  $f : \mathbb{R} \rightarrow \mathbb{R}$  ponendo  $f(t) = \langle v_0 + tv_1, v_0 + tv_1 \rangle$ . Allora si ha  $f(0) = \langle v_0, v_0 \rangle < 0$  e

$$\begin{aligned} f(t) &= \langle v_0 + tv_1, v_0 + tv_1 \rangle = \langle v_0, v_0 + tv_1 \rangle + t \langle v_1, v_0 + tv_1 \rangle \\ &= \langle v_0, v_0 \rangle + 2t \langle v_0, v_1 \rangle + t^2 \langle v_1, v_1 \rangle. \end{aligned}$$

Dunque  $f$  è un polinomio di secondo grado con coefficiente direttore positivo e termine noto negativo; quindi esiste (perché?) un  $t_0 \in \mathbb{R}^*$  tale che  $f(t_0) = 0$ . Per concludere, basta dimostrare che  $v_0 + t_0 v_1 \neq O$ . Se fosse  $v_0 + t_0 v_1 = O$ , avremmo  $v_0 = -t_0 v_1$ , ma allora  $\langle v_0, v_0 \rangle = \langle -t_0 v_1, -t_0 v_1 \rangle = t_0^2 \langle v_1, v_1 \rangle > 0$ , mentre  $\langle v_0, v_0 \rangle < 0$ .  $\square$

L'idea è che su uno spazio vettoriale metrico è possibile, usando il prodotto scalare e la norma associata, studiare la geometria metrica come nel piano o nello spazio euclideo. Per far ciò abbiamo però bisogno di definire la distanza fra due punti, e l'angolo fra due vettori (o direzioni).

### Definizione 11.9

Sia  $V$  uno spazio vettoriale metrico. Allora la *distanza*  $d(v, w)$  fra due punti  $v, w \in V$  è data da  $d(v, w) = \|w - v\|$ .

### Esempio 11.9

Se  $\langle \cdot, \cdot \rangle$  è il prodotto scalare canonico su  $\mathbb{R}^2$ , la Definizione 11.9 dà la distanza euclidea come visto in (11.2). Considerando invece altri prodotti scalari definiti positivi otteniamo distanze diverse. Per esempio, con il prodotto scalare definito positivo dell'Esempio 11.2 otteniamo la distanza

$$\begin{aligned} d(v, w) &= \sqrt{2(w_1 - v_1)^2 + 2(w_1 - v_1)(w_2 - v_2) + (w_2 - v_2)^2 + 3(w_3 - v_3)^2} \\ &= \sqrt{(w_1 - v_1)^2 + (w_1 + w_2 - v_1 - v_2)^2 + 3(w_3 - v_3)^2}. \end{aligned}$$

Perché la funzione introdotta nella Definizione 11.9 meriti il nome di “distanza” deve verificare le proprietà principali che ci aspettiamo valgano per la distanza fra punti. Per esempio, un punto deve ovviamente essere a distanza zero da se stesso: e infatti  $d(v_0, v_0) = \|v_0 - v_0\| = \|O\| = 0$ . Inoltre, la distanza fra  $v_0$  e  $v_1$  dev'essere uguale alla distanza fra  $v_1$  e  $v_0$ : e infatti

$$d(v_0, v_1) = \|v_1 - v_0\| = \|-(v_0 - v_1)\| = \|v_0 - v_1\| = d(v_1, v_0),$$

(dove abbiamo usato il fatto che  $\|-\nu\| = \|\nu\|$  per ogni  $\nu \in V$ ; vedi la Proposizione 11.2.(3) qui sotto). Infine, se andiamo da  $v_0$  direttamente a  $v_2$  dobbiamo percorrere al più una distanza uguale a quella percorsa andando da  $v_0$  a  $v_1$  e poi da  $v_1$  a  $v_2$ , quale che sia  $v_1$ ; in simboli, si deve avere

$$d(v_0, v_2) \leq d(v_0, v_1) + d(v_1, v_2). \quad (11.9)$$

Anche questa proprietà è verificata dalla nostra distanza, ma non è ovvia come le altre due. La dimostrazione è contenuta nella seguente proposizione, che raccoglie tutte le proprietà fondamentali della norma in uno spazio vettoriale metrico.

### Proposizione 11.2

Sia  $V$  uno spazio vettoriale metrico con norma  $\|\cdot\|$ . Allora:

- (1)  $\|\nu\| = 0$  se e solo se  $\nu = O$ ;
- (2)  $\|\nu\| > 0$  per ogni  $\nu \neq O$ ;
- (3)  $\|\lambda\nu\| = |\lambda| \|\nu\|$  per ogni  $\lambda \in \mathbb{R}$  e  $\nu \in V$ , e  $\|\nu + w\|^2 = \|\nu\|^2 + 2\langle \nu, w \rangle + \|w\|^2$  per ogni  $\nu, w \in V$ ;
- (4) per ogni  $\nu, w \in V$  si ha

$$|\langle \nu, w \rangle| \leq \|\nu\| \|w\|,$$

(disuguaglianza di Cauchy-Schwarz), e l'uguaglianza vale se e solo se  $v$  e  $w$  sono linearmente dipendenti;

(5) per ogni  $v, w \in V$  si ha

$$|\|v\| - \|w\|| \leq \|v + w\| \leq \|v\| + \|w\|,$$

(disuguaglianza triangolare);

(6) per ogni  $v, w \in V$  si ha  $\langle v, w \rangle = \frac{1}{4} [\|v + w\|^2 - \|v - w\|^2]$ .

### Dimostrazione

La (1) e la (2) seguono dalle definizioni.

(3) Infatti  $\|\lambda v\| = \sqrt{\langle \lambda v, \lambda v \rangle} = \sqrt{|\lambda|^2 \langle v, v \rangle} = |\lambda| \|v\|$ . Inoltre, si ha

$$\|v + w\|^2 = \langle v + w, v + w \rangle = \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle = \|v\|^2 + 2\langle v, w \rangle + \|w\|^2.$$

(4) Se  $v = O$  oppure  $w = O$  è ovvio; quindi possiamo supporli entrambi non nulli. Siccome  $\langle \cdot, \cdot \rangle$  è definito positivo, si ha

$$\begin{aligned} 0 &\leq \|av + bw\|^2 = \langle av + bw, av + bw \rangle = a\langle v, av + bw \rangle + b\langle w, av + bw \rangle \\ &= a^2\langle v, v \rangle + ab\langle v, w \rangle + ab\langle w, v \rangle + b^2\langle w, w \rangle = a^2\|v\|^2 + 2ab\langle v, w \rangle + b^2\|w\|^2 \end{aligned}$$

per ogni  $a, b \in \mathbb{R}$ . In particolare, prendendo  $a = \|w\|^2$  e  $b = -\langle v, w \rangle$  otteniamo

$$0 \leq \|w\|^4\|v\|^2 - 2\|w\|^2|\langle v, w \rangle|^2 + \|w\|^2|\langle v, w \rangle|^2 = \|w\|^2(\|v\|^2\|w\|^2 - |\langle v, w \rangle|^2);$$

dividendo per  $\|w\|^2$  troviamo la disuguaglianza di Cauchy-Schwarz. L'uguaglianza vale se e solo se  $av + bw = O$ , cioè se e solo se  $v$  e  $w$  sono linearmente dipendenti.

(5) La disuguaglianza di Cauchy-Schwarz ci dà

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

e abbiamo dimostrato la disuguaglianza triangolare.

(6) Abbiamo

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2 + 2\langle v, w \rangle \quad \text{e} \quad \|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2\langle v, w \rangle,$$

per cui l'asserto segue sottraendo opportunamente.  $\square$

### Osservazione 11.5

La Proposizione 11.2.(6) ci dice una cosa importante: in uno spazio vettoriale metrico, la norma determina completamente il prodotto scalare. In altri termini, se due prodotti scalari hanno la stessa norma, allora coincidono.

### Osservazione 11.6

La disuguaglianza triangolare implica che per ogni  $v_0, v_1, v_2 \in V$  si ha

$$\|v_2 - v_0\| \leq \|v_1 - v_0\| + \|v_2 - v_1\|$$

(basta prendere  $v = v_1 - v_0$  e  $w = v_2 - v_1$ ), per cui (11.9) è verificata. Il nome "disuguaglianza triangolare" è dovuto al fatto che può essere interpretata come l'affermazione che

un lato del triangolo di vertici  $O$ ,  $v$  e  $v + w$  è più corto della somma e più lungo della differenza delle lunghezze degli altri due lati (vedi la Figura 11.3), un risultato classico della geometria euclidea dei triangoli.

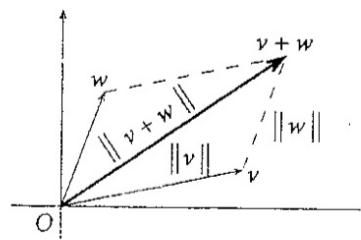


Figura 11.3 - La diseguaglianza triangolare.

Dunque la nostra distanza su uno spazio vettoriale metrico soddisfa effettivamente le proprietà basilari che si richiedono a una distanza. Infine, la diseguaglianza di Cauchy-Schwarz ci permette di definire anche l'angolo fra due vettori.

### Definizione 11.10

Sia  $V$  uno spazio vettoriale metrico, e  $v, w \in V$  due vettori non nulli. Diremo *angolo* fra  $v$  e  $w$  quel numero reale  $\widehat{vw} \in [0, \pi]$  tale che

$$\cos \widehat{vw} = \frac{\langle v, w \rangle}{\|v\| \|w\|}.$$

La diseguaglianza di Cauchy-Schwarz dice che il rapporto a secondo membro è in modulo minore o uguale a 1, per cui è il coseno di qualcosa; ed essendo il coseno una funzione iniettiva nell'intervallo  $[0, \pi]$ , abbiamo definito in modo univoco l'angolo fra  $v$  e  $w$ .

Dunque abbiamo recuperato anche la definizione di angolo in uno spazio vettoriale metrico. Nota che il caso di uguaglianza nella Proposizione 11.2.(4) ci dice che due vettori  $v$  e  $w$  sono linearmente dipendenti se e solo se il coseno del loro angolo è uguale a  $\pm 1$ , cioè se e solo se formano un angolo di  $0$  o  $\pi$  radianti, com'è giusto che sia. D'altro canto, formano un angolo di  $\pi/2$  se e solo se il coseno è zero, cioè se e solo se  $\langle v, w \rangle = 0$ . Questo ci porta alla seguente definizione.

### Definizione 11.11

Sia  $\langle \cdot, \cdot \rangle$  un prodotto scalare su uno spazio vettoriale  $V$ . Due vettori  $v, w \in V$  si dicono *ortogonali* (o *perpendicolari*) se  $\langle v, w \rangle = 0$ , e scriveremo  $v \perp w$ . Se poi  $S_1, S_2 \subset V$  sono due sottoinsiemi di  $V$  tali che tutti gli elementi di  $S_1$  sono ortogonali a tutti gli elementi di  $S_2$  diremo che  $S_1$  è *ortogonale* a  $S_2$  (vedi la Figura 11.4), e scriveremo  $S_1 \perp S_2$ . Se  $S_1 = \{v_1\}$ , scriveremo  $v_1 \perp S_2$  invece di  $\{v_1\} \perp S_2$ .

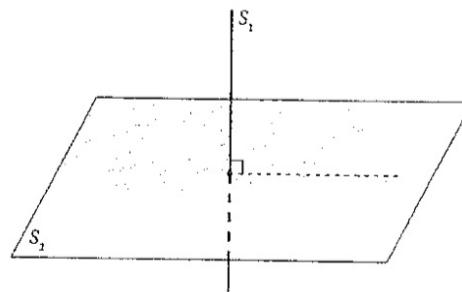


Figura 11.4 - Due sottoinsiemi ortogonali.

## 11.3 Basi ortogonali

Uno dei concetti più utili per lavorare con gli spazi vettoriali è stato senza dubbio quello di base; vediamo se riusciamo ad adattarlo in modo da renderlo utile anche nel caso degli spazi vettoriali metrici. Cominciamo col seguente lemma.

### Lemma 11.3

*Sia  $V$  uno spazio vettoriale metrico, e  $v_1, \dots, v_k \in V$  vettori non nulli ortogonali a due a due, cioè tali che  $v_i \perp v_j$  non appena  $i \neq j$ . Allora  $v_1, \dots, v_k$  sono linearmente indipendenti. In particolare, se  $\dim V = n$ , allora  $n$  vettori non nulli a due a due ortogonali sono automaticamente una base di  $V$ .*

### Dimostrazione

Sia  $\alpha_1 v_1 + \dots + \alpha_k v_k = O$  una relazione di dipendenza lineare fra i vettori  $v_1, \dots, v_k$ . Moltiplicando scalarmente entrambi i membri dell'uguaglianza per il vettore  $v_j$  si trova  $0 = \langle O, v_j \rangle = \alpha_1 \langle v_1, v_j \rangle + \dots + \alpha_j \langle v_j, v_j \rangle + \dots + \alpha_k \langle v_k, v_j \rangle = \alpha_j \|v_j\|^2$ , perché  $\langle v_i, v_j \rangle = 0$  se  $i \neq j$ ; quindi  $\alpha_j = 0$ . Siccome questo valeva per qualunque  $j$ , tutti gli  $\alpha_j$  sono zero, e dunque  $v_1, \dots, v_k$  sono linearmente indipendenti.  $\square$

Questo risultato ci porta a identificare il tipo di base che ci interessa.

### Definizione 11.12

Una *base ortogonale* di uno spazio vettoriale metrico  $V$  è una base  $\{v_1, \dots, v_n\}$  di  $V$  composta da vettori a due a due ortogonali, ovvero tali che valga  $\langle v_i, v_j \rangle = 0$  per  $1 \leq i \neq j \leq n$ .

### Esempio 11.10

La base canonica  $\{e_1, e_2\}$  è una base ortogonale di  $\mathbb{R}^2$  col prodotto scalare canonico. Un'altra base ortogonale è  $\left\{ \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} -3 \\ 1 \end{pmatrix} \right\}$ , e se ne possono trovare facilmente infinite altre.  $\square$

Vi è anche un altro tipo di base ancora più speciale.

### Definizione 11.13

Una *base ortonormale* di uno spazio vettoriale metrico  $V$  è una base ortogonale composta da vettori di lunghezza unitaria (ovvero, con una terminologia che useremo raramente, composta da *versori*).

In altri termini, una base  $\{v_1, \dots, v_n\}$  di uno spazio vettoriale metrico è una base ortonormale se e solo se

$$\langle v_i, v_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{se } i = j, \\ 0 & \text{se } i \neq j. \end{cases} \quad (11.10)$$

### Esempio 11.11

La base canonica è una base ortonormale di  $\mathbb{R}^n$  rispetto al prodotto scalare canonico per ogni  $n$  naturale positivo.  $\square$

**Esempio 11.12**

Un'altra base ortonormale di  $\mathbb{R}^2$  rispetto al prodotto scalare canonico è data da

$$\left\{ \begin{vmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{vmatrix}, \begin{vmatrix} -3/\sqrt{10} \\ 1/\sqrt{10} \end{vmatrix} \right\}.$$

**Esempio 11.13**

La base canonica non è una base ortonormale per il prodotto scalare definito positivo dell'Esempio 11.2 su  $\mathbb{R}^3$  (non è neppure ortogonale, in quanto  $\langle e_1, e_2 \rangle = 1$ ). Per trovare una base ortonormale  $\{v_1, v_2, v_3\}$  per questo prodotto scalare cominciamo prendendo  $v_1 = e_1/\sqrt{2}$ . Siccome  $\langle e_1, e_1 \rangle = 2$ , allora il vettore  $v_1 = e_1/\sqrt{2}$  ha norma 1. Il secondo vettore  $v_2$  deve essere ortogonale a  $v_1$ ; quindi se scriviamo  $v_2 = (x, y, z)^T$  si deve avere

$$\langle v_1, v_2 \rangle = \frac{2}{\sqrt{2}} x + \frac{1}{\sqrt{2}} y = 0.$$

Ricordandosi che  $\|v_2\| = 1$ , una possibile scelta è  $v_2 = (1/\sqrt{2}, -2/\sqrt{2}, 0)^T$ . Le coordinate del terzo vettore  $v_3$  devono soddisfare il sistema

$$\begin{cases} \langle v_1, v_3 \rangle = \frac{2}{\sqrt{2}} x + \frac{1}{\sqrt{2}} y = 0, \\ \langle v_2, v_3 \rangle = -\frac{1}{\sqrt{2}} y = 0, \end{cases}$$

con inoltre  $\|v_3\| = 1$ ; una scelta possibile (l'unica a meno del segno) è  $v_3 = e_3/\sqrt{3}$ . Dunque una base ortonormale per questo prodotto scalare è data da

$$\left\{ \begin{vmatrix} 1/\sqrt{2} \\ 0 \\ 0 \end{vmatrix}, \begin{vmatrix} 1/\sqrt{2} \\ -2/\sqrt{2} \\ 0 \end{vmatrix}, \begin{vmatrix} 0 \\ 0 \\ 1/\sqrt{3} \end{vmatrix} \right\}.$$

Ogni vettore di uno spazio vettoriale si scrive in maniera unica come combinazione lineare degli elementi di una base, e i coefficienti sono le coordinate del vettore rispetto a quella base. Nel caso delle basi ortogonali, le coordinate si trovano immediatamente usando il prodotto scalare.

**Proposizione 11.4**

*Sia  $\{v_1, \dots, v_n\}$  una base ortogonale di uno spazio vettoriale metrico  $V$ . Allora*

$$v = \frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 + \cdots + \frac{\langle v, v_n \rangle}{\langle v_n, v_n \rangle} v_n \quad (11.1)$$

*per ogni  $v \in V$ . In particolare, se  $\{v_1, \dots, v_n\}$  è una base ortonormale si ha*

$$\forall v \in V \quad v = \langle v, v_1 \rangle v_1 + \cdots + \langle v, v_n \rangle v_n. \quad (11.12)$$

**Dimostrazione**

Siccome  $\{v_1, \dots, v_n\}$  è una base, esistono unici numeri reali  $\alpha_1, \dots, \alpha_n \in \mathbb{R}$  tali che  $v = \alpha_1 v_1 + \dots + \alpha_n v_n$ . Moltiplicando scalarmente per  $v_j$ , l'ortogonalità della base implica  $\langle v, v_j \rangle = \alpha_j \langle v_j, v_j \rangle$ , ed è fatta.  $\square$

Dunque il vettore  $v \in V$  è completamente individuato dai numeri  $\langle v, v_j \rangle / \langle v_j, v_j \rangle$ , che sono detti *coefficienti di Fourier* di  $v$  rispetto alla base  $\{v_1, \dots, v_n\}$ . In particolare, la sua norma si deve esprimere in termini di questi coefficienti.

**Corollario 11.5**

Sia  $\{v_1, \dots, v_n\}$  una base ortogonale di uno spazio vettoriale metrico  $V$ . Allora

$$\forall v, w \in V \quad \langle v, w \rangle = \frac{\langle v, v_1 \rangle \langle v_1, w \rangle}{\langle v_1, v_1 \rangle} + \dots + \frac{\langle v, v_n \rangle \langle v_n, w \rangle}{\langle v_n, v_n \rangle} \quad (11.13)$$

$$\forall v \in V \quad \|v\|^2 = \frac{|\langle v, v_1 \rangle|^2}{\|v_1\|^2} + \dots + \frac{|\langle v, v_n \rangle|^2}{\|v_n\|^2}. \quad (11.14)$$

In particolare, se  $\{v_1, \dots, v_n\}$  è una base ortonormale di  $V$  si ha

$$\forall v, w \in V \quad \langle v, w \rangle = \langle v, v_1 \rangle \langle v_1, w \rangle + \dots + \langle v, v_n \rangle \langle v_n, w \rangle \quad (11.15)$$

(formula di Parseval) e

$$\forall v \in V \quad \|v\|^2 = |\langle v, v_1 \rangle|^2 + \dots + |\langle v, v_n \rangle|^2 \quad (11.16)$$

(teorema di Pitagora).

**Dimostrazione**

Moltiplicando scalarmente (11.11) e (11.12) per  $w$  si ottengono le (11.13) e (11.15); infine (11.14) e (11.16) seguono prendendo  $w = v$ .  $\square$

Si pone ora il problema dell'esistenza di basi ortonormali. La risposta è che non solo ci sono sempre, ma si possono costruire a partire da una base qualunque con una procedura effettiva che conserva molte proprietà (per esempio gli span) della base di partenza. Questa procedura, che è descritta nella dimostrazione del prossimo teorema, è detta *procedimento di ortogonalizzazione di Gram-Schmidt*.

**Teorema 11.6**

Sia  $V$  uno spazio vettoriale metrico, e  $v_1, \dots, v_r \in V$  linearmente indipendenti. Allora esistono  $w_1, \dots, w_r \in V$  che soddisfano le seguenti proprietà:

- (1)  $\text{Span}(w_1, \dots, w_j) = \text{Span}(v_1, \dots, v_j)$  per ogni  $j = 1, \dots, r$ ;
- (2)  $w_j$  è ortogonale a ogni elemento di  $\text{Span}(w_1, \dots, w_{j-1})$  per ogni  $j = 2, \dots, r$ ;
- (3)  $\langle w_j, v_j \rangle > 0$  per  $j = 1, \dots, r$ .

Gli elementi  $w_1, \dots, w_r$  sono univocamente determinati a meno di fattori scalari positivi; in altri termini, se  $w'_1, \dots, w'_r$  sono altri elementi di  $V$  che soddisfano le proprietà (1)-(3) allora  $w'_j = \alpha_j w_j$  con  $\alpha_j > 0$  per  $j = 1, \dots, r$ . Infine, gli elementi  $w_1, \dots, w_r$  sono ottenuti ponendo  $w_1 = v_1$  e

$$w_j = v_j - \sum_{h=1}^{j-1} \frac{\langle v_j, w_h \rangle}{\langle w_h, w_h \rangle} w_h, \quad (11.17)$$

per  $j = 2, \dots, r$ .

**Dimostrazione**

Procediamo per induzione su  $r$ . Per  $r = 1$ , la (1) ci dice che dev'essere  $w_1 = cv_1$  per qualche  $c \neq 0$ . La (2) non fornisce alcuna condizione, e la (3) diventa  $c\|v_1\|^2 > 0$ , cioè  $c > 0$ , e quindi in questo caso ci siamo.

Supponiamo il teorema vero per  $r - 1$  vettori, e dimostriamolo per  $r$  vettori. Abbiamo dunque costruito  $w_1, \dots, w_{r-1}$  usando la (11.17); dobbiamo trovare  $w_r$ .

Prima di tutto notiamo che, essendo i vettori  $v_1, \dots, v_r$  linearmente indipendenti, dimensione di  $\text{Span}(v_1, \dots, v_r)$  è esattamente  $r$ . Siccome, per ipotesi induttiva,  $\{w_1, \dots, w_{r-1}, v_r\}$  è un sistema di generatori di  $\text{Span}(v_1, \dots, v_r)$ , questo implica che  $w_1, \dots, w_{r-1}$  e  $v_r$  sono linearmente indipendenti.

In particolare, vogliamo trovare  $w_r$  della forma  $w_r = \alpha v_r - \beta_1 w_1 - \dots - \beta_{r-1} w_{r-1}$ , dove  $\alpha, \beta_1, \dots, \beta_{r-1} \in \mathbb{R}$  devono essere determinati in modo che (1)-(3) valgano. Prima di tutto, la (1) implica  $\alpha \neq 0$ . La condizione (2) è equivalente (Esercizio 11.12) a dire che  $w_r$  dev'essere ortogonale a  $w_1, \dots, w_{r-1}$  (i quali a loro volta sono ortogonali a due a due per ipotesi induttiva); quindi si deve avere

$$0 = \langle w_r, w_j \rangle = \alpha \langle v_r, w_j \rangle - \beta_1 \langle w_1, w_j \rangle - \dots - \beta_{r-1} \langle w_{r-1}, w_j \rangle = \alpha \langle v_r, w_j \rangle - \beta_j \langle w_j, w_j \rangle$$

cioè  $\beta_j = \alpha \langle v_r, w_j \rangle / \langle w_j, w_j \rangle$  per  $j = 1, \dots, r-1$ . Dunque

$$w_r = \alpha \left[ v_r - \sum_{h=1}^{r-1} \frac{\langle v_r, w_h \rangle}{\langle w_h, w_h \rangle} w_h \right].$$

Dobbiamo scoprire per quali  $\alpha \in \mathbb{R}$  la (3) è verificata. Notiamo innanzitutto che

$$\langle w_r, v_r \rangle = \alpha \left[ \|v_r\|^2 - \sum_{h=1}^{r-1} \frac{|\langle v_r, w_h \rangle|^2}{\|w_h\|^2} \right]. \quad (11.18)$$

Vogliamo dimostrare che  $\langle w_r, v_r \rangle > 0$  se e solo se  $\alpha > 0$ ; per far ciò dobbiamo trovare segno della quantità fra parentesi quadre. Per costruzione, il vettore  $v_r$  appartiene a  $\text{Span}(w_1, \dots, w_r)$ , e  $\{w_1, \dots, w_r\}$  è una base ortogonale di  $\text{Span}(w_1, \dots, w_r)$ . Quindi (11.14) ci dice che

$$\|v_r\|^2 = \sum_{h=1}^r \frac{|\langle v_r, w_h \rangle|^2}{\|w_h\|^2} = \sum_{h=1}^{r-1} \frac{|\langle v_r, w_h \rangle|^2}{\|w_h\|^2} + \frac{|\langle v_r, w_r \rangle|^2}{\|w_r\|^2} > \sum_{h=1}^{r-1} \frac{|\langle v_r, w_h \rangle|^2}{\|w_h\|^2},$$

in quanto  $v_r \notin \text{Span}(v_1, \dots, v_{r-1}) = \text{Span}(w_1, \dots, w_{r-1})$ , e quindi  $\langle v_r, w_r \rangle \neq 0$ . Dunque la quantità fra parentesi quadre in (11.18) è positiva, e vediamo che (3) è verificata se e solo se  $\alpha > 0$ , come voluto.

Come conseguenza otteniamo l'esistenza delle basi ortonormali.

**Corollario 11.7**

Sia  $\mathcal{B} = \{v_1, \dots, v_n\}$  una base di uno spazio vettoriale metrico  $V$ . Allora esiste un'unica base ortonormale  $\mathcal{B}' = \{u_1, \dots, u_n\}$  di  $V$  tale che si abbia  $\langle u_j, v_j \rangle > 0$  e  $\text{Span}(v_1, \dots, v_j) = \text{Span}(u_1, \dots, u_j)$  per  $j = 1, \dots, n$ .

**Dimostrazione**

Il Teorema 11.6 fornisce una base ortogonale  $\{w_1, \dots, w_n\}$  soddisfacente le condizioni indicate; ponendo  $u_j = w_j / \|w_j\|$  per  $j = 1, \dots, n$  troviamo la base ortonormale cercata. Se  $\{u'_1, \dots, u'_n\}$  è un'altra base ortonormale che soddisfa le stesse condizioni, per il Teorema 11.6 si deve avere  $u'_j = \alpha_j u_j$  con opportuni  $\alpha_j > 0$ , per  $j = 1, \dots, n$ . Ma allora  $1 = \|u'_j\| = \|\alpha_j u_j\| = \alpha_j \|u_j\| = \alpha_j$ , e quindi  $u'_j = u_j$  per  $j = 1, \dots, n$ .  $\square$

C'è anche una versione metrica del teorema del completamento (Teorema 4.10).

**Corollario 11.8**

*Siano  $v_1, \dots, v_r$  vettori a due a due ortogonali di uno spazio vettoriale metrico  $V$  di dimensione  $n \geq r$ . Allora esistono  $n - r$  vettori  $w_{r+1}, \dots, w_n \in V$  tali che  $\mathcal{B} = \{v_1, \dots, v_r, w_{r+1}, \dots, w_n\}$  sia una base ortogonale di  $V$ . Inoltre, se  $\|v_j\| = 1$  per  $j = 1, \dots, r$ , allora  $w_{r+1}, \dots, w_n$  possono essere scelti in modo che  $\mathcal{B}$  sia una base ortonormale di  $V$ .*

**Dimostrazione**

Completiamo i vettori  $v_1, \dots, v_r$  a una base  $\{v_1, \dots, v_r, v'_{r+1}, \dots, v'_{n-1}\}$  di  $V$ ; il Teorema 11.6 ci fornisce allora una base ortogonale  $\{w_1, \dots, w_r, w_{r+1}, \dots, w_n\}$  di  $V$ , in cui, grazie a (11.17), possiamo supporre  $w_j = v_j$  per  $j = 1, \dots, r$ . La parte restante dell'enunciato è ora ovvia.  $\square$

Quindi non solo le basi ortonormali esistono sempre, ma possiamo anche fare tutte le operazioni di completamento che facevamo con le basi qualunque. Vediamo come procedere con un esempio numerico.

**Esempio 11.14**

Consideriamo lo spazio vettoriale

$$V = \text{Span} \left( v_1 = \begin{vmatrix} 0 \\ 1 \\ 0 \\ 1 \end{vmatrix}, v_2 = \begin{vmatrix} 1 \\ 1 \\ 1 \\ 1 \end{vmatrix}, v_3 = \begin{vmatrix} 1 \\ 1 \\ 1 \\ 0 \end{vmatrix} \right) \subset \mathbb{R}^4$$

equipaggiato col prodotto scalare canonico; ne vogliamo una base ortonormale. Applichiamo il procedimento di Gram-Schmidt, come descritto in (11.17), e il Corollario 11.7. Siccome  $\|v_1\|^2 = 2$ , il primo passo è

$$w_1 = v_1, \quad u_1 = \frac{w_1}{\|w_1\|} = \begin{vmatrix} 0 \\ 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{vmatrix}.$$

Per il secondo passo, calcoliamo  $\|w_1\|^2 = 2$  e  $\langle v_2, w_1 \rangle = 2$ . Quindi

$$w_2 = v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 = \begin{vmatrix} 1 \\ 1 \\ 1 \\ 1 \end{vmatrix} - \frac{2}{2} \begin{vmatrix} 0 \\ 1 \\ 0 \\ 1 \end{vmatrix} = \begin{vmatrix} 1 \\ 0 \\ 1 \\ 0 \end{vmatrix}, \quad u_2 = \frac{w_2}{\|w_2\|} = \begin{vmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \\ 0 \end{vmatrix}.$$

Infine per il terzo passo calcoliamo anche la quantità  $\|w_2\|^2 = 2$ ,  $\langle v_3, w_1 \rangle = 1$ ,  $\langle v_3, w_2 \rangle = 2$ . Quindi

$$w_3 = v_3 - \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2 = \begin{vmatrix} 1 \\ 1 \\ 1 \\ 0 \end{vmatrix} - \frac{1}{2} \begin{vmatrix} 0 \\ 1 \\ 0 \\ 1 \end{vmatrix} - \frac{2}{2} \begin{vmatrix} 1 \\ 0 \\ 1 \\ 0 \end{vmatrix} = \begin{vmatrix} 0 \\ 1/2 \\ 0 \\ -1/2 \end{vmatrix},$$

$$u_3 = \frac{w_3}{\|w_3\|} = \begin{vmatrix} 0 \\ 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{vmatrix},$$

e la base ortonormale cercata è

$$\left\{ \begin{vmatrix} 0 \\ 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{vmatrix}, \begin{vmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \\ 0 \end{vmatrix}, \begin{vmatrix} 0 \\ 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{vmatrix} \right\}.$$

## 11.4 Proiezioni ortogonali

Vogliamo esaminare meglio dal punto di vista geometrico il procedimento di Gram-Schmidt. Se  $r = 1$  non c'è nulla da dire, ma già se  $r = 2$  notiamo qualcosa di interessante. Per trovare  $w_2$  abbiamo sottratto da  $v_2$  un vettore  $w$  in modo da ottenere qualcosa ortogonale a  $v_1$ ; inoltre questo  $w$  era un multiplo di  $v_1$ . È un procedimento che ricorda la proiezione ortogonale di  $v_2$  su  $v_1$  (vedi la Figura 11.2), e la somiglianza diventa ancora più spiccata osservando che

$$w = \frac{\langle v_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 = \|v_2\| \cos \theta \frac{v_1}{\|v_1\|},$$

dove  $\theta$  è l'angolo fra  $v_1$  e  $v_2$ ; in altre parole abbiamo usato proprio la stessa formula della proiezione ortogonale nel piano (vedi l'Osservazione 11.3).

Per  $r > 2$  succede essenzialmente la stessa cosa: per trovare  $w_r$  prendiamo  $v_r$  e vi sottraiamo la sua proiezione ortogonale sul sottospazio generato da  $w_1, \dots, w_{r-1}$ . In altre parole, a ogni passo applichiamo il seguente risultato:

### Proposizione 11.9

Sia  $V$  uno spazio vettoriale metrico, e  $U$  un sottospazio di  $V$ . Allora per ogni  $v_0 \in V$  esiste un'unico  $u_0 \in U$  tale che  $v_0 - u_0$  sia ortogonale a tutti gli elementi di  $U$ .

### Dimostrazione

Fissiamo una base ortonormale  $\{u_1, \dots, u_r\}$  di  $U$  (costruita per esempio applicando il Corollario 11.7 a una qualunque base del sottospazio  $U$ ), e supponiamo che  $u_0 = \alpha_1 u_1 + \dots + \alpha_r u_r \in U$  sia come richiesto. Ma allora deduciamo che

$$0 = \langle v_0 - u_0, u_j \rangle = \langle v_0, u_j \rangle - \alpha_1 \langle u_1, u_j \rangle - \dots - \alpha_r \langle u_r, u_j \rangle = \langle v_0, u_j \rangle - \alpha_j$$

(in quanto  $\{u_1, \dots, u_r\}$  è una base ortonormale di  $U$ ), e quindi si ha  $\alpha_j = \langle v_0, u_j \rangle$  per  $j = 1, \dots, r$ . Dunque il vettore  $u_0$ , se esiste, è unico ed è dato da

$$u_0 = \langle v_0, u_1 \rangle u_1 + \cdots + \langle v_0, u_r \rangle u_r. \quad (11.19)$$

Definiamo allora  $u_0 \in U$  come in (11.19). Il conto già fatto mostra che  $v_0 - u_0$  è ortogonale a  $u_1, \dots, u_r$ , per cui (Esercizio 11.12) è ortogonale a tutto  $U$ .  $\square$

Questo risultato suggerisce di definire un'applicazione  $P_U: V \rightarrow V$  ponendo

$$P_U(v) = \langle v, u_1 \rangle u_1 + \cdots + \langle v, u_r \rangle u_r$$

per ogni  $v \in V$ , dove  $\{u_1, \dots, u_r\}$  è una base ortonormale del sottospazio  $U$ . L'applicazione  $P_U$  è chiaramente lineare, e non dipende dalla base scelta: infatti la dimostrazione della Proposizione 11.9 mostra che  $P_U(v)$  è l'unico elemento di  $U$  tale che  $v - P_U(v)$  sia ortogonale a tutti gli elementi di  $U$ .

### Definizione 11.14

Sia  $U$  un sottospazio vettoriale di uno spazio vettoriale metrico  $V$ . L'applicazione lineare  $P_U: V \rightarrow V$  appena definita si dice *proiezione ortogonale* di  $V$  su  $U$  (vedi la Figura 11.5). Nota che  $P_U(u) = u$  per ogni  $u \in U$  (grazie alla Proposizione 11.4), e che  $\text{Im } P_U = U$  (ovvio).

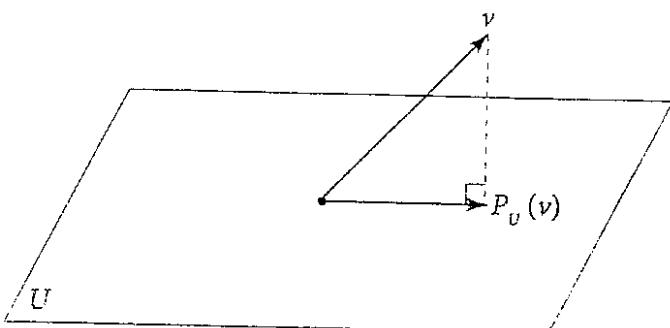


Figura 11.5 - Proiezione ortogonale.

Visto che abbiamo parlato di proiezione ortogonale, introduciamo anche i supplementi ortogonali (in modo da poter generalizzare al caso degli spazi vettoriali metrici concetti come quello di retta ortogonale a un piano in  $\mathbb{R}^3$ ).

### Definizione 11.15

Sia  $V$  uno spazio vettoriale equipaggiato con un prodotto scalare  $\langle \cdot, \cdot \rangle$ . L'*ortogonale* di un sottoinsieme  $S \subseteq V$  è l'insieme  $S^\perp$  di tutti gli elementi di  $V$  ortogonali a  $S$ . Nel caso in cui  $S$  sia un sottospazio vettoriale  $U$  di  $V$  (e  $\langle \cdot, \cdot \rangle$  sia definito positivo), allora  $U^\perp$  è detto *supplemento ortogonale* di  $U$ .

Questa definizione richiede dei chiarimenti: bisogna far vedere che  $U^\perp$  è un sottospazio, e che è un supplementare di  $U$  (nel senso della Definizione 4.11).

### Lemma 11.10

Sia  $V$  uno spazio vettoriale metrico. Allora:

- (1) se  $S$  è un sottoinsieme di  $V$ , allora  $S^\perp$  è un sottospazio vettoriale;
- (2) se  $U$  è un sottospazio di  $V$ , allora  $U^\perp = \text{Ker } P_U$ .

**Dimostrazione**

(1) Infatti se  $v_1, v_2 \in S^\perp$ ,  $w \in S$  e  $\lambda$  è uno scalare, allora abbiamo

$$\langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle = 0 + 0 = 0 \text{ e } \langle \lambda v_1, w \rangle = \lambda \langle v_1, w \rangle = \lambda 0 = 0,$$

per cui  $v_1 + v_2, \lambda v_1 \in S^\perp$ .

(2) Infatti  $v \in \text{Ker } P_U$  se e solo se (perché?)  $v$  è ortogonale a tutti gli elementi di  $U$ , se e solo se  $v \in U^\perp$ .

Allora possiamo dimostrare la seguente proposizione.

**Proposizione 11.11**

Sia  $U$  un sottospazio vettoriale di uno spazio vettoriale metrico  $V$ . Allora  $V = U \oplus U^\perp$ , in particolare,  $\dim U^\perp = \dim V - \dim U$ . Infine, se scriviamo  $v = u + u^\perp$  con  $v \in V$ ,  $u \in U$  e  $u^\perp \in U^\perp$ , si ha

$$\|v\|^2 = \|u\|^2 + \|u^\perp\|^2. \quad (11.20)$$

**Dimostrazione**

Prima di tutto,  $U \cap U^\perp = \{O\}$ ; infatti, se  $u \in U \cap U^\perp$ , allora il vettore  $u$  (essendo in  $U^\perp$ ) è ortogonale a se stesso (essendo in  $U$ ), per cui ha norma zero e quindi è il vettore nullo.

Per dimostrare che  $V = U + U^\perp$  basta notare che ogni  $v \in V$  può essere decomposto come  $v = P_U(v) + (v - P_U(v))$  con, per definizione,  $P_U(v) \in U$  e  $v - P_U(v) \in U^\perp$ .

Infine, se scriviamo  $v = u + u^\perp$  si ha  $\|v\|^2 = \langle u + u^\perp, u + u^\perp \rangle = \|u\|^2 + \|u^\perp\|^2$ , perché  $\langle u, u^\perp \rangle = 0$ .

Come corollario si ha la seguente caratterizzazione della proiezione ortogonale che giustificherà le definizioni di distanza fra un punto e una retta o un piano che vedremo nel prossimo capitolo. □

**Corollario 11.12**

Sia  $U$  un sottospazio di uno spazio vettoriale metrico  $V$ , e prendiamo  $v \in V$ . Allora  $P_U(v)$  è l'elemento di  $U$  più vicino a  $v$ . In altri termini,  $\|v - u\|^2 \geq \|v - P_U(v)\|^2$  per ogni  $u \in U$ , e l'uguaglianza vale se e solo se  $u = P_U(v)$ .

**Dimostrazione**

Scriviamo  $v = u_0 + u_0^\perp$ , dove  $u_0 = P_U(v)$  e  $u_0^\perp = v - P_U(v) \in U^\perp$ . Allora se  $u \in U$  l'uguaglianza (11.20) ci dice che (vedi la Figura 11.6)

$$\|v - u\|^2 = \|u_0 - u\|^2 + \|u_0^\perp\|^2 = \|u_0 - u\|^2 + \|v - P_U(v)\|^2 \geq \|v - P_U(v)\|^2,$$

e l'uguaglianza vale se e solo se  $\|u_0 - u\|^2 = 0$ , cioè se e solo se  $u = u_0 = P_U(v)$ . □

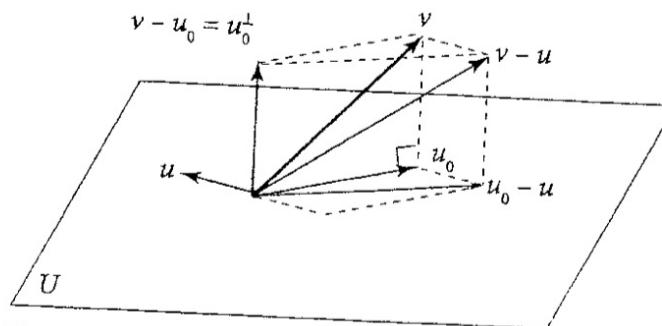


Figura 11.6 - Caratterizzazione della proiezione ortogonale.

## 11.5 Prodotti scalari e matrici

Abbiamo visto che per fare i conti con le applicazioni lineari conviene fissare una base; in questo modo si associa all'applicazione lineare una matrice e poi si lavora con quest'ultima. Nel caso dei prodotti scalari, o più in generale delle forme bilineari, possiamo fare esattamente la stessa cosa.

Sia  $V$  uno spazio vettoriale su  $\mathbb{R}$ , e  $g$  una forma bilineare qualunque su  $V$ . Scelta una base  $\mathcal{B} = \{v_1, \dots, v_n\}$  di  $V$ , dati  $v, w \in V$  indichiamo con  $x = F_{\mathcal{B}}(v)$ , rispettivamente  $y = F_{\mathcal{B}}(w) \in \mathbb{R}^n$ , le coordinate di  $v$  e  $w$  rispetto a  $\mathcal{B}$ . Allora si ha

$$\begin{aligned} g(v, w) &= g(x_1 v_1 + \dots + x_n v_n, y_1 v_1 + \dots + y_n v_n) = g\left(\sum_{k=1}^n x_k v_k, \sum_{h=1}^n y_h v_h\right) \\ &= \sum_{k=1}^n \sum_{h=1}^n x_k y_h g(v_k, v_h). \end{aligned}$$

Quindi se introduciamo la matrice  $S = (s_{hk})$  data da (attenzione all'ordine degli indici!)  $s_{hk} = g(v_k, v_h)$ , abbiamo

$$g(v, w) = \sum_{h=1}^n \sum_{k=1}^n y_h s_{hk} x_k = y^T S x. \quad (11.21)$$

In particolare, *una forma bilineare è completamente determinata dai valori che assume su una base*.

### Definizione 11.16

Sia  $g: V \times V \rightarrow \mathbb{R}$  una forma bilineare su uno spazio vettoriale  $V$ , e  $\mathcal{B} = \{v_1, \dots, v_n\}$  una base di  $V$ . La matrice  $S \in M_{n,n}(\mathbb{R})$  che al posto  $(h, k)$  contiene il numero  $g(v_k, v_h)$  si dice *matrice associata* a (o che *rappresenta*) la forma bilineare  $g$  rispetto alla base  $\mathcal{B}$ .

### Esempio 11.15

Se  $S \in M_{n,n}(\mathbb{R})$  rappresenta una forma bilineare  $g$  su  $\mathbb{R}^n$  rispetto alla base canonica, allora  $s_{hk}$  è il coefficiente di  $x_k y_h$  (perché? Vedi anche gli Esempi 11.2-11.5).  $\square$

Dunque, fissata una base, a ogni forma bilineare possiamo associare una matrice in modo che (11.21) sia valida. Viceversa, se  $S \in M_{n,n}(\mathbb{R})$  è una matrice quadrata (e abbiamo scelto una base  $\mathcal{B}$ ) allora possiamo usare (11.21) per definire una forma bilineare su  $V$ . Potremmo formalizzare questo discorso dicendo che lo spazio delle forme bilineari su  $V$  è isomorfo a  $M_{n,n}(\mathbb{R})$ , dove  $n = \dim V$ ; ma siccome non useremo questo fatto lo lasciamo per esercizio (vedi gli Esercizi 11.15-11.16).

A questo punto si pongono due problemi fondamentali. Come cambia la matrice associata cambiando base? Come possiamo trovare le proprietà della forma bilineare semplicemente guardando la matrice associata?

Il primo problema ha una risposta immediata.

### Proposizione 11.13

Sia  $g: V \times V \rightarrow \mathbb{R}$  una forma bilineare su uno spazio vettoriale  $V$  su  $\mathbb{R}$ . Siano  $\mathcal{B} = \{v_1, \dots, v_n\}$  e  $\mathcal{B}' = \{v'_1, \dots, v'_n\}$  due basi di  $V$ , e  $B$  la matrice di cambiamento di base da  $\mathcal{B}$  a  $\mathcal{B}'$ . Indichiamo infine con  $S$  e  $S'$  le matrici associate a  $g$  rispetto a  $\mathcal{B}$  e  $\mathcal{B}'$ . Allora si ha  $S' = B^T S B$ .

**Dimostrazione**

Per definizione la  $i$ -esima colonna di  $B$  contiene le coordinate di  $v'_i$  rispetto alla base  $\mathcal{B}$  per  $i = 1, \dots, n$ . Quindi  $s'_{hk} = g(v'_k, v'_h) = (B^h)^T S B^k$  e pertanto si ha  $S' = B^T S B$ .

**Osservazione 11.7**

Nonostante le apparenze, le matrici associate agli endomorfismi lineari e quelle associate alle forme bilineari *non* sono la stessa cosa. Certo, in entrambi i casi le matrici sembrano essere semplicemente tabelle quadrate di numeri, ma in realtà si comportano in modo profondamente diverso. Per esempio, cambiando base le matrici associate a forme bilineari cambiano secondo la legge  $S \mapsto B^T S B$ , mentre abbiamo visto che le matrici associate a endomorfismi cambiano secondo la legge  $A \mapsto B^{-1} A B$ . Il punto è che in entrambi i casi le matrici sono soltanto dei *mezzi* per fare calcoli con oggetti che sono intrinsecamente diversi; quello che studiamo, a cui siamo interessati, sono questi oggetti (gli endomorfismi, le forme bilineari), e non le matrici in sé. Quindi quando si fanno conti con le matrici bisogna avere ben chiaro se si tratta di endomorfismi o di forme bilineari.

**Definizione 11.17**

Due matrici  $A, A' \in M_{n,n}(\mathbb{R})$  si dicono *congruenti* se esiste una matrice invertibile  $B \in GL_n(\mathbb{R})$  tale che  $A' = B^T A B$ . La classe di congruenza  $\mathcal{C}_A$  della matrice  $A$  è l'insieme delle matrici congruenti ad  $A$ .

Dunque se due matrici rappresentano la stessa forma bilineare rispetto a basi diverse, allora sono congruenti. Come nel caso degli endomorfismi (vedi la Proposizione 8.2) vale anche il viceversa (Esercizio 11.18).

**Osservazione 11.8**

Nota che la matrice associata a un prodotto scalare rispetto a una base ortogonale è necessariamente diagonale (perché?). Inoltre, se il prodotto scalare è definito positivo, allora la base  $\mathcal{B}$  è ortonormale se e solo se la matrice associata rispetto a questa base è la matrice identica  $I_n$ . Dunque se rispetto alla base ortonormale  $\mathcal{B}$  i vettori  $v, w \in V$  hanno come coordinate rispettivamente i vettori  $x$  e  $y$ , il loro prodotto scalare sarà allora dato da  $\langle v, w \rangle = y^T x$ . In altri termini, ogni prodotto scalare definito positivo si esprime in coordinate rispetto a una base *ortonormale* esattamente come il prodotto scalare canonico.

Questa osservazione ha due conseguenze rilevanti.

**Lemma 11.14**

*Sia  $\langle \cdot, \cdot \rangle$  un prodotto scalare definito positivo su uno spazio vettoriale  $V$ , ed  $S$  la matrice associata a  $\langle \cdot, \cdot \rangle$  rispetto a una base  $\mathcal{B}$ . Allora  $\det S > 0$ .*

**Dimostrazione**

Sia  $\mathcal{B}_0$  una base ortonormale di  $V$ , e  $B$  la matrice di cambiamento di base da  $\mathcal{B}_0$  a  $\mathcal{B}$ . Allora  $\langle \cdot, \cdot \rangle$  è rappresentato da  $I_n$  rispetto a  $\mathcal{B}_0$ , per cui  $S = B^T B$  e di conseguenza  $\det S = \det(B^T B) = (\det B)^2 > 0$ .

**Proposizione 11.15**

*Sia  $V$  uno spazio vettoriale su  $\mathbb{R}$ , e  $\mathcal{B} = \{v_1, \dots, v_n\}$  una base di  $V$ . Allora esiste unico prodotto scalare definito positivo su  $V$  rispetto a cui  $\mathcal{B}$  è una base ortonormale.*

**Dimostrazione**

Per l'Osservazione 11.8, questo prodotto scalare, se esiste, deve essere definito ponendo  $\langle v, w \rangle = y^T x$  per ogni  $v, w \in V$  dove  $x = F_B(v)$  e  $y = F_B(w)$  sono le coordinate di  $v$  e  $w$  rispetto a  $B$ . Dunque basta verificare che  $\langle \cdot, \cdot \rangle$  è davvero un prodotto scalare definito positivo su  $V$ . Per esempio, è omogeneo nella prima variabile: se  $x = F_B(v)$  allora

$$\lambda x = F_B(\lambda v) \text{ e } \langle \lambda v, w \rangle = y^T (\lambda x) = \lambda (y^T x) = \lambda \langle v, w \rangle.$$

In modo analogo (vedi l'Esercizio 11.19) si dimostrano la bilinearità e la simmetria. È definito positivo: infatti per ogni  $v \in V$  si ha  $\langle v, v \rangle = x^T x = x_1^2 + \cdots + x_n^2 \geq 0$ , e l'ugualianza vale se e solo se  $x = O$ , cioè se e solo se  $v = O$ . Infine,  $B$  è una base ortonormale, perché  $F_B(v_j) = e_j$  per  $j = 1, \dots, n$ , per cui  $\langle v_i, v_j \rangle = e_j^T e_i = \delta_{ij}$ .  $\square$

Torniamo al secondo problema, ancora aperto: come leggere dalle matrici associate le proprietà delle forme bilineari? Ci serve una definizione.

**Definizione 11.18**

Una matrice quadrata  $A \in M_{n,n}(\mathbb{R})$  a coefficienti reali si dice *simmetrica* se  $A^T = A$ ; *antisimmetrica* se  $A^T = -A$ .

**Lemma 11.16**

Sia  $V$  uno spazio vettoriale su  $\mathbb{R}$  di dimensione  $n$ , e  $g$  una forma bilineare su  $V$ . Fissiamo una base  $B = \{v_1, \dots, v_n\}$  di  $V$ , e sia  $S$  la matrice che rappresenta  $g$  rispetto a  $B$ . Allora:

- (1) la forma  $g$  è simmetrica se e solo se  $S$  è simmetrica;
- (2) supponiamo  $g$  simmetrica. Allora  $\text{Ker } S = F_B(V^\perp)$ . In particolare,  $g$  è non degenera se e solo se  $S$  è non singolare, cioè se e solo se  $\det S \neq 0$ .

**Dimostrazione**

- (1) Se  $g$  è simmetrica, allora  $s_{hk} = g(v_k, v_h) = g(v_h, v_k) = s_{kh}$ , e  $S$  è simmetrica. Viceversa, siano  $v, w \in V$  con coordinate  $x$  e  $y$  rispetto a  $B$ . Se la matrice  $S$  è simmetrica si vede facilmente che

$$g(v, w) = y^T S x = y^T S^T x = (x^T S y)^T = x^T S y = g(w, v),$$

dove  $(x^T S y)^T = x^T S y$  per il semplice motivo che  $x^T S y$  è un numero reale, e quindi  $g$  è simmetrica.

- (2) Un vettore  $v_0 \in V$  appartiene a  $V^\perp$  se e solo se  $g(v_0, w) = 0$  per ogni  $w \in V$ . Ponendo  $x_0 = F_B(v_0) \in \mathbb{R}^n$ , questo accade se e solo se  $y^T S x_0 = 0$  per ogni  $y \in \mathbb{R}^n$ , che a sua volta può succedere se e solo se  $S x_0 = O$  (perché?), cioè se e solo se  $x_0 \in \text{Ker } S$ . Dunque  $v_0 \in V^\perp$  se e solo se abbiamo  $F_B(v_0) \in \text{Ker } S$ , cioè  $\text{Ker } S = F_B(V^\perp)$ . In particolare,  $V^\perp = \{O\}$  se e solo se  $\text{Ker } S = \{O\}$ , per cui  $g$  è non degenera se e solo se  $S$  è non singolare.  $\square$

Dunque simmetria e degenericità si leggono bene sulle matrici rappresentative. Per sapere come stabilire se un certo prodotto scalare è definito positivo guardandone una matrice associata dobbiamo però aspettare il Paragrafo 14.3, in quanto servono strumenti teorici che non abbiamo ancora introdotto.

## 11.6 Endomorfismi simmetrici e isometrie

Il prodotto scalare ci permette di definire alcune classi importanti di endomorfismi.

### Definizione 11.19

Sia  $T: V \rightarrow V$  un endomorfismo di uno spazio vettoriale metrico  $V$ . Diremo che  $T$  è *simmetrico* (o *autoaggiunto*) se  $\langle T(v_1), v_2 \rangle = \langle v_1, T(v_2) \rangle$  per ogni  $v_1, v_2 \in V$ .

### Osservazione 11.9

Sia  $T: V \rightarrow V$  un endomorfismo di uno spazio vettoriale metrico, e indichiamo con  $A$  la matrice associata a  $T$  rispetto a una base  $B$ , e con  $S$  la matrice associata al prodotto scalare rispetto a  $B$ . Se  $x, y$  sono le coordinate dei vettori  $v, w \in V$  rispetto a  $B$ , allora si ha  $\langle T(v), w \rangle = y^T S A x$  e  $\langle v, T(w) \rangle = (Ay)^T S x = y^T A^T S x$ . Quindi l'endomorfismo  $T$  è simmetrico (perché?) se e solo se

$$S A = A^T S.$$

In particolare, se  $B$  è una base *ortonormale* di  $V$  allora l'endomorfismo  $T: V \rightarrow V$  è simmetrico se e solo se la matrice associata a  $T$  rispetto a  $B$  è simmetrica.

Se la base  $B$  non è ortonormale, non è più vero che  $T$  è simmetrico se e solo se la matrice associata a  $T$  rispetto a  $B$  è simmetrica.

### Esempio 11.16

Preso  $\mathbb{R}^2$  col prodotto scalare canonico, sia  $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  dato da

$$T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_1 - x_2 \end{pmatrix}.$$

L'endomorfismo  $T$  è simmetrico, per esempio perché rispetto alla base canonica è rappresentato da una matrice simmetrica. Scegliamo invece come base di  $\mathbb{R}^2$  la base

$B = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$ . Rispetto a questa base il prodotto scalare canonico è rappresentato dalla

matrice  $S = \begin{vmatrix} 2 & 3 \\ 3 & 5 \end{vmatrix}$  mentre  $T$  è rappresentato dalla matrice

$$A = \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix}^{-1} \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = \begin{vmatrix} 4 & 7 \\ -2 & -4 \end{vmatrix},$$

che si guarda bene dall'essere simmetrica. In questa base il fatto che  $T$  è un endomorfismo simmetrico è indicato dall'eguaglianza  $S^{-1} A^T S = A$ , eguaglianza che avrai cura di verificare.

Mentre l'importanza degli endomorfismi simmetrici diventerà evidente solo nel Capitolo 13, c'è un'altra classe di endomorfismi il cui interesse è subito chiaro.

### Definizione 11.20

Sia  $T: V \rightarrow V$  un endomorfismo di uno spazio vettoriale metrico  $V$ . Diremo che  $T$  è un'*isometria* (lineare), o *endomorfismo ortogonale*, se  $\langle T(v_1), T(v_2) \rangle = \langle v_1, v_2 \rangle$  per ogni  $v_1, v_2 \in V$ .

Dunque un'isometria conserva il prodotto scalare; in particolare, conserva distanze e angoli. Prima di discutere degli esempi, vediamo alcune proprietà.

**Proposizione 11.17**

Sia  $T: V \rightarrow V$  un endomorfismo di uno spazio vettoriale metrico  $V$ . Allora le seguenti affermazioni sono equivalenti:

- (1)  $T$  è un'isometria;
- (2) se  $\{v_1, \dots, v_n\}$  è una base ortonormale di  $V$ , anche  $\{T(v_1), \dots, T(v_n)\}$  lo è;
- (3)  $\|T(v)\| = \|v\|$  per ogni  $v \in V$  (e, in particolare,  $T$  è invertibile).

**Dimostrazione**

(1)  $\Rightarrow$  (2). Infatti  $\langle T(v_i), T(v_j) \rangle = \langle v_i, v_j \rangle = \delta_{ij} = \begin{cases} 0 & \text{se } i \neq j, \\ 1 & \text{se } i = j. \end{cases}$

(2)  $\Rightarrow$  (3). Utilizzando (11.12) e (11.16) si ha

$$\begin{aligned} \|T(v)\|^2 &= \left\langle \sum_{i=1}^n \langle v, v_i \rangle T(v_i), \sum_{j=1}^n \langle v, v_j \rangle T(v_j) \right\rangle = \sum_{i=1}^n \sum_{j=1}^n \langle v, v_i \rangle \langle v_j, v \rangle \langle T(v_i), T(v_j) \rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n \langle v, v_i \rangle \langle v_j, v \rangle \delta_{ij} = \sum_{h=1}^n |\langle v, v_h \rangle|^2 = \|v\|^2. \end{aligned}$$

(3)  $\Rightarrow$  (1). Utilizzando la Proposizione 11.2.(6) troviamo

$$\begin{aligned} \langle T(v), T(w) \rangle &= \frac{1}{4} [\|T(v) + T(w)\|^2 - \|T(v) - T(w)\|^2] \\ &= \frac{1}{4} [\|T(v+w)\|^2 - \|T(v-w)\|^2] = \frac{1}{4} [\|v+w\|^2 - \|v-w\|^2] = \langle v, w \rangle. \end{aligned}$$

□

**Corollario 11.18**

Sia  $T: V \rightarrow V$  un endomorfismo di uno spazio vettoriale metrico  $V$ . Fissiamo una base  $\mathcal{B}$  di  $V$ ; sia  $A$  la matrice che rappresenta  $T$  rispetto a  $\mathcal{B}$ , ed  $S$  la matrice che rappresenta il prodotto scalare rispetto a questa base. Allora  $T$  è ortogonale se e solo se  $A$  è invertibile e  $A^{-1} = S^{-1}A^T S$ .

**Dimostrazione**

Ragionando come nell'Osservazione 11.9 si vede che  $T$  è ortogonale se e solo se  $S = A^T S A$ , e ci siamo (perché?). □

**Definizione 11.21**

Una matrice  $A \in M_{n,n}(\mathbb{R})$  si dice *ortogonale* se  $A^T A = I_n$ . L'insieme delle matrici ortogonali di ordine  $n$  si indica con  $O(n)$ .

**Osservazione 11.10**

Nota che il Corollario 7.8 ci dice che ogni matrice ortogonale è automaticamente invertibile, e che  $A^{-1} = A^T$ .

**Osservazione 11.11**

Un endomorfismo  $T: V \rightarrow V$  è ortogonale se e solo se è rappresentato da una matrice ortogonale rispetto a una base *ortonormale*; confronta con l'Osservazione 11.9, che descrive quello che accade agli endomorfismi simmetrici.

**Corollario 11.19**

Se  $A \in M_{n,n}(\mathbb{R})$ , le seguenti affermazioni sono equivalenti:

- (1)  $A$  è ortogonale;
- (2)  $L_A$  è un'isometria rispetto al prodotto scalare canonico;
- (3) le colonne di  $A$  formano una base ortonormale di  $\mathbb{R}^n$  rispetto al prodotto scalare canonico.

**Dimostrazione**

(1)  $\Rightarrow$  (2). Poiché  $A$  è la matrice che rappresenta  $L_A$  rispetto alla base canonica, che è una base ortonormale rispetto al prodotto scalare canonico, il risultato segue subito da Corollario 11.18.

(2)  $\Rightarrow$  (3). Le colonne di  $A$  sono l'immagine tramite  $L_A$  della base canonica, e quindi sono una base ortonormale per la Proposizione 11.17.

(3)  $\Rightarrow$  (1). Infatti l'elemento di posto  $(i,j)$  del prodotto  $A^T A$  è dato da

$$(A^T A)_{ij} = (A^T)_i A^j = (A^i)^T A^j = \langle A^j, A^i \rangle = \delta_{ij},$$

e quindi  $A^T A = I_n$ . □

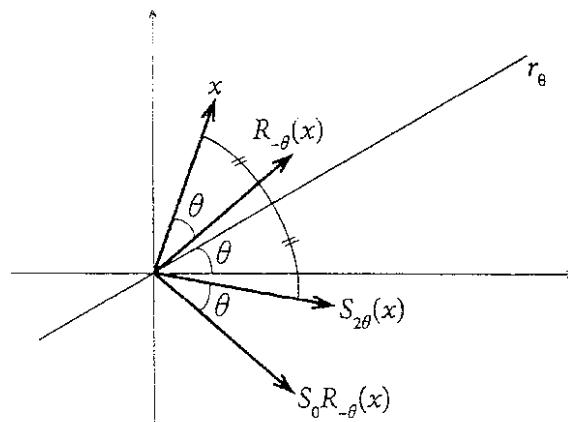
**Esempio 11.17**

Vogliamo trovare tutte le applicazioni di  $\mathbb{R}^2$  in sé ortogonali rispetto al prodotto scalare canonico – ovvero tutte le matrici ortogonali  $A \in O(2)$ . La relazione  $A^T A = I_2$  in coordinate diventa

$$\begin{vmatrix} a_{11}^2 + a_{21}^2 & a_{11}a_{12} + a_{21}a_{22} \\ a_{11}a_{12} + a_{21}a_{22} & a_{12}^2 + a_{22}^2 \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}.$$

Quindi devono esistere angoli  $\theta, \phi \in \mathbb{R}$  tali che  $a_{11} = \cos \theta$ ,  $a_{21} = \sin \theta$ ,  $a_{12} = \sin \phi$ ,  $a_{22} = \cos \phi$  e inoltre  $0 = \cos \theta \sin \phi + \sin \theta \cos \phi = \sin(\theta + \phi)$ . Dunque abbiamo due possibilità: o  $\phi = -\theta$  oppure  $\phi = \pi - \theta$ . Corrispondentemente, la matrice  $A$  è necessariamente di uno dei due tipi seguenti.

$$R_\theta = \begin{vmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{vmatrix} \quad \text{oppure} \quad S_\theta = \begin{vmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{vmatrix},$$



**Figura 11.7** - Matrici ortogonali e simmetrie.

ed è facile verificare che  $R_\theta, S_\theta \in O(2)$  per ogni  $\theta \in \mathbb{R}$ . Interpretiamo ora geometricamente le isometrie  $R_\theta$  e  $S_\theta$ . Prendiamo  $x \in \mathbb{R}^2$  e scriviamolo in coordinate polari: si ha  $x = r(\cos \psi, \sin \psi)$ , per qualche  $r \geq 0$  e  $\psi \in \mathbb{R}$ . Allora

$$R_\theta(x) = \begin{vmatrix} r(\cos \theta \cos \psi - \sin \theta \sin \psi) \\ r(\sin \theta \cos \psi + \cos \theta \sin \psi) \end{vmatrix} = \begin{vmatrix} r \cos(\theta + \psi) \\ r \sin(\theta + \psi) \end{vmatrix},$$

cioè  $R_\theta$  è la rotazione di angolo  $\theta$  in verso antiorario. Per capire chi è  $S_\theta$ , notiamo che (Figura 11.7)  $S_{2\theta} = R_\theta S_0 R_{-\theta}$ ; dunque  $S_{2\theta}$  si ottiene effettuando prima una rotazione di angolo  $-\theta$ , poi una simmetria  $S_0$  rispetto all'asse  $x$  e infine una rotazione di angolo  $\theta$ . Il risultato è una simmetria rispetto alla retta  $r_\theta$  che forma un angolo  $\theta$  con l'asse  $x$ . Pertanto, gli endomorfismi ortogonali di  $\mathbb{R}^2$  rispetto al prodotto scalare canonico sono esattamente le rotazioni di centro l'origine e le simmetrie rispetto a rette passanti per l'origine. Infine, si verifica subito che  $S_{\theta_1} S_{\theta_2} = R_{\theta_1 - \theta_2}$ , per cui *ogni rotazione si ottiene come composizione di due simmetrie* (e viceversa). □

- 11.1** Trova quali delle seguenti applicazioni  $\langle \cdot, \cdot \rangle: \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$  sono dei prodotti scalari, scrivili nella forma  $\langle v, w \rangle = w^T S v$  e stabilisci quali sono (semi)definiti:

$$\langle v, w \rangle_1 = v_1^2 + w_2^2 + v_3 w_3; \quad \langle v, w \rangle_2 = v_1 w_1 - v_1 w_2 - v_2 w_1 + 3v_2 w_2 + \pi v_3 w_3.$$

$$\langle v, w \rangle_3 = v_1 w_2 + v_2 w_1 + v_2 w_3 + v_3 w_2 + v_3 w_1 + v_1 w_3; \quad \langle v, w \rangle_4 = 3v_1 w_1 - v_1 w_2 + v_2 w_1 - v_2 w_2 - v_3 w_3.$$

$$\langle v, w \rangle_5 = -v_1 w_1 + v_1 w_2 + v_2 w_1 - v_2 w_2 - v_3 w_3; \quad \langle v, w \rangle_6 = v_1 w_1 - v_2 w_2.$$

$$\langle v, w \rangle_7 = 400v_1 w_1 + 3\sqrt{\pi}v_1 w_3 + 3\sqrt{\pi}v_3 w_1 + 227v_3 w_3.$$

- 11.2** Supponiamo che  $\mathcal{R}_0 = RA(O, \vec{i}, \vec{j}_1)$  sia un sistema di riferimento affine in  $A^2$  in cui i vettori  $\vec{i}$  e  $\vec{j}$  siano perpendicolari e di lunghezza unitaria. Considera il sistema di riferimento affine  $\mathcal{R} = RA(O, \vec{i}, \vec{j})$ , dove  $\vec{j}_1 = 2\vec{j}$ , e definisci la "distanza" fra i punti di  $A^2$  con la formula (11.2), ma usando le coordinate rispetto a  $\mathcal{R}$  invece di quelle rispetto a  $\mathcal{R}_0$ . L'esercizio consiste nel disegnare l'insieme dei punti di  $A^2$  che distano 1 dall'origine, rispetto a questa particolare "distanza".

- 11.3** Dimostra che  $\langle p, q \rangle = p(0)q(0) + p(1)q(1) + \cdots + p(n)q(n)$  è un prodotto scalare definito positivo su  $\mathbb{R}_n[t]$ .

- 11.4** Posto  $P = \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix}$ , dimostra che l'applicazione  $\langle \cdot, \cdot \rangle: M_{2,2}(\mathbb{R}) \times M_{2,2}(\mathbb{R}) \rightarrow \mathbb{R}$  data da  $\langle A, B \rangle = \text{tr}(B^T P A)$  è un prodotto scalare definito positivo.

- 11.5** Sia  $P \in M_{n,n}(\mathbb{R})$  una matrice simmetrica data e siano  $\langle \cdot, \cdot \rangle: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  e  $\langle \cdot, \cdot \rangle: M_{n,n}(\mathbb{R}) \times M_{n,n}(\mathbb{R}) \rightarrow \mathbb{R}$  le applicazioni definite da  $\langle x, y \rangle = y^T P x$  e  $\langle A, B \rangle = \text{tr}(B^T P A)$  per ogni  $x, y \in \mathbb{R}^n$  e  $A, B \in M_{n,n}(\mathbb{R})$ . Dimostra che sono entrambi prodotti scalari, e che  $\langle \cdot, \cdot \rangle$  è definito positivo se e solo se  $\langle \cdot, \cdot \rangle$  lo è.

- 11.6** Sia  $V$  uno spazio vettoriale su  $\mathbb{R}$ , e  $\langle \cdot, \cdot \rangle$  un prodotto scalare su  $V$  (anche degenere). Dimostra che  $2\langle v, w \rangle = \langle v + w, v + w \rangle - \langle v, v \rangle - \langle w, w \rangle$  per ogni  $v, w \in V$ .

- 11.7** Sia  $V$  uno spazio vettoriale metrico. Dimostra che due vettori  $v$  e  $w$  sono linearmente dipendenti se e solo se vale una delle uguaglianze nella diseguaglianza triangolare (Proposizione 11.2.(5)).

- 11.8** Sia  $V$  uno spazio vettoriale metrico con norma  $\|\cdot\|$ . Dimostra che per ogni  $v, w \in V$  si ha

$$\|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2)$$

(uguaglianza del parallelogramma). Interpreta geometricamente questa formula considerando il parallelogramma di vertici  $O, v, w$  e  $v + w$ .

- 11.9** Trova basi ortonormali di  $\mathbb{R}^3$  rispetto ai prodotti scalari dell'Esercizio 11.1 che sono definiti positivi.
- 11.10** Dimostra che se  $\{v_1, \dots, v_n\}$  è una base ortogonale di uno spazio vettoriale metrico  $V$ , allora  $\{v_1/\|v_1\|, \dots, v_n/\|v_n\|\}$  è una base ortonormale di  $V$ .
- 11.11** Trova una base ortonormale, rispetto al prodotto scalare canonico, del sottospazio  $V$  di  $\mathbb{R}^4$  generato dai vettori  $(1, -1, 1, -1)^T$ ,  $(5, 1, 1, 1)^T$  e  $(-3, -3, 1, -3)^T$ , e completa la base ortonormale di tutto  $\mathbb{R}^4$ .
- 11.12** Sia  $U$  un sottospazio vettoriale di uno spazio vettoriale  $V$ ,  $\{u_1, \dots, u_r\}$  una base di  $U$ , e sia  $\langle \cdot, \cdot \rangle$  un prodotto scalare su  $V$ . Dimostra che un vettore  $v \in V$  è ortogonale a  $U$  se e solo se  $v \perp u_j$  per  $j = 1, \dots, r$ .
- 11.13** Mostra che una proiezione ortogonale è una proiezione nel senso dell'Esercizio 7.7.
- 11.14** Siano  $U, U_1, U_2$  sottospazi vettoriali di uno spazio vettoriale metrico  $V$ . Dimostra che  $U_1 \subseteq U_2$  implica  $U_1^\perp \supseteq U_2^\perp$ ; che  $(U^\perp)^\perp = U$ ; che  $(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp$ ; e che  $(U_1 \cap U_2)^\perp = U_1^\perp + U_2^\perp$ .
- 11.15** Dimostra che tutte le forme bilineari su  $\mathbb{R}^n$  sono del tipo  $g(x, y) = y^T S x$  per qualche matrice  $S \in M_{n,n}(\mathbb{R})$ .
- 11.16** Sia  $\text{Bil}(V)$  l'insieme costituito dalle forme bilineari su uno spazio vettoriale reale  $V$  di dimensione  $n$ . Definisci una somma e un prodotto per scalari su  $\text{Bil}(V)$  in modo da renderlo uno spazio vettoriale. Fissata una base di  $V$ , dimostra che l'applicazione che associa a ogni elemento di  $\text{Bil}(V)$  la matrice che lo rappresenta rispetto a questa base è un isomorfismo di  $\text{Bil}(V)$  con  $M_{n,n}(\mathbb{R})$ , e da questo deduci quindi che  $\dim \text{Bil}(V) = n^2$ .
- 11.17** Trova la matrice associata al prodotto scalare dell'Esempio 11.7 rispetto alla solita base  $\mathcal{B} = \{1, t, t^2\}$ .
- 11.18** Sia  $g: V \times V \rightarrow \mathbb{R}$  una forma bilineare su uno spazio vettoriale  $V$ , sia  $\mathcal{B}$  una base di  $V$ , e  $S$  la matrice associata a  $g$  rispetto a  $\mathcal{B}$ . Sia  $S'$  una matrice congruente a  $S$ ; dimostra che esiste una base  $\mathcal{B}'$  di  $V$  tale che  $S'$  rappresenta  $g$  rispetto a  $\mathcal{B}'$ .
- 11.19** Completa i dettagli della dimostrazione della Proposizione 11.15.
- 11.20** Dimostra che ogni proiezione ortogonale è un endomorfismo simmetrico.
- 11.21** Sia  $U$  un sottospazio di uno spazio vettoriale metrico  $V$ , e siano  $\mathcal{B}_1$  una base di  $U$  e  $\mathcal{B}_2$  una base di  $U^\perp$ . Trova la matrice che rappresenta la proiezione ortogonale  $P_U: V \rightarrow V$  rispetto alla base  $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$  di  $V$ . Se  $\mathcal{B}_0$  è una base ortonormale di  $V$ , dimostra che la matrice  $P$  associata a  $P_U$  rispetto a  $\mathcal{B}_0$  è  $P = B(B^T B)^{-1} B^T$ , dove  $B$  è la matrice che ha per colonne le coordinate dei vettori di  $\mathcal{B}_1$  rispetto a  $\mathcal{B}_0$ .
- 11.22** Sia  $T: V \rightarrow V$  un endomorfismo simmetrico di uno spazio vettoriale metrico  $V$ . Dimostra che la forma  $\langle \cdot, \cdot \rangle_T$  definita da  $\langle v, w \rangle_T = \langle T(v), w \rangle$  è un prodotto scalare, non degenere se e solo se  $T$  è invertibile.
- 11.23** Se  $T_1, T_2: V \rightarrow V$  sono isometrie di uno spazio vettoriale metrico  $V$ , dimostra che anche  $T_1 \circ T_2$  e  $T_1^{-1}$  lo sono.

- 11.24** Dimostra che l'insieme  $O(n)$  delle matrici ortogonali di ordine  $n$  è un gruppo (non commutativo se  $n > 1$ ) rispetto al prodotto di matrici.
- 11.25** Dimostra che se  $A$  è una matrice ortogonale allora  $|\det A| = 1$ .
- 11.26** Siano  $\mathcal{B}_1 = \{v_1, \dots, v_n\}$  e  $\mathcal{B}_2 = \{w_1, \dots, w_n\}$  due basi ortonormali di  $\mathbb{R}^n$ . Dimostra che esiste un'unica matrice ortogonale  $B \in O(n)$  tale che  $Bv_j = w_j$  per  $j = 1, \dots, n$ . (*Suggerimento:* usa la Proposizione 11.17 e il Corollario 11.19).
- 11.27** Definisci il concetto di prodotto hermitiano su uno spazio vettoriale complesso, e generalizza a questa situazione tutti i risultati dei Paragrafi 11.2-11.4.
- 11.28** Definisci il concetto di matrice  $S$  associata a un prodotto hermitiano rispetto a una base, nota che è *hermitiana* (cioè tale che  $S^H = S$ ), e generalizza a questa situazione tutti i risultati del Paragrafo 11.5.
- 11.29** Sia  $\langle \cdot, \cdot \rangle$  un prodotto hermitiano definito positivo su uno spazio vettoriale complesso  $V$ , e  $T: V \rightarrow V$  un endomorfismo. Diremo che  $T$  è *hermitiano* (o *autoaggiunto*) se  $\langle T(v_1), v_2 \rangle = \langle v_1, T(v_2) \rangle$  per ogni  $v_1, v_2 \in V$ ; diremo che è *unitario* (o un'*isometria*) se  $\langle T(v_1), T(v_2) \rangle = \langle v_1, v_2 \rangle$  per ogni  $v_1, v_2 \in V$ . Infine, diremo che una matrice  $A \in M_{n,n}(\mathbb{C})$  è *unitaria* se  $A^H A = I_n$ . Usando queste definizioni, generalizza al caso complesso il contenuto del Paragrafo 11.6.
- 11.30** Data  $A \in M_{n,n}(\mathbb{R})$ , dimostra che si ha  $\langle L_A(x), y \rangle = \langle x, L_{A^T}(y) \rangle$  per ogni  $x, y \in \mathbb{R}^n$ , dove  $\langle \cdot, \cdot \rangle$  è il prodotto scalare canonico.
- 11.31** Data  $A \in M_{n,n}(\mathbb{C})$ , dimostra che si ha  $\langle L_A(x), y \rangle = \langle x, L_{A^H}(y) \rangle$  per ogni  $x, y \in \mathbb{C}^n$ , dove  $\langle \cdot, \cdot \rangle$  è il prodotto hermitiano canonico.

# Geometria euclidea

# 12

## Sommario

**12.1** Angoli e ortogonalità

**12.2** Distanze

**12.3** Il prodotto vettore

## Complementi

**12C.1** Isometrie

**12C.2** Isometrie lineari dello spazio

## Esercizi

In questo capitolo parleremo della geometria metrica del piano e dello spazio, utilizzando gli strumenti introdotti nei capitoli precedenti. Definiremo l'angolo fra rette e/o piani, la distanza fra punti, rette e/o piani, e vedremo come calcolarli. Parleremo anche del prodotto vettore, un'operazione definita esclusivamente su  $\mathbb{R}^3$  che permette di associare a una coppia di vettori un terzo vettore ortogonale a entrambi. Nei complementi a questo capitolo discuteremo le isometrie del piano e dello spazio, dedicando particolare attenzione alle simmetrie.

## 12.1 Angoli e ortogonalità

Cominciamo definendo l'ambiente in cui studieremo la geometria metrica.

### Definizione 12.1

Il piano euclideo  $\mathcal{E}^2$  (rispettivamente, lo spazio euclideo  $\mathcal{E}^3$ ) è il piano affine  $\mathcal{A}^2$  (lo spazio affine  $\mathcal{A}^3$ ) in cui sia stata fissata un'unità di misura per le lunghezze.

È utile ricordare che nel piano e nello spazio euclideo è definita in maniera univoca anche un'unità di misura per gli angoli, in quanto l'angolo di 1 radiante è quello che in una circonferenza di raggio 1 sottende un arco lungo 1.

### Definizione 12.2

Un sistema di riferimento cartesiano  $RC(O, A_1, A_2, A_3)$  di  $\mathcal{E}^3$  è un sistema di riferimento affine  $RA(O, A_1, A_2, A_3)$  tale che i vettori  $\vec{i} = \overrightarrow{OA_1}$ ,  $\vec{j} = \overrightarrow{OA_2}$  e  $\vec{k} = \overrightarrow{OA_3}$  siano di lunghezza unitaria e ortogonali a due a due.

In altri termini,  $\{\vec{i}, \vec{j}, \vec{k}\}$  è una base ortonormale di  $\mathcal{V}_O^3$ , e quindi l'usuale identificazione fra  $\mathcal{A}^3$  (e quindi  $\mathcal{E}^3$ ) ed  $\mathbb{R}^3$  data dalle coordinate rispetto a questa base ci permette di misurare lunghezze e angoli in  $\mathcal{E}^3$  usando il prodotto scalare canonico di  $\mathbb{R}^3$ . Inutile dire che analoghe definizioni e osservazioni si applicano a  $\mathcal{E}^2$  ed  $\mathbb{R}^2$ .

Innanzitutto vediamo velocemente cosa accade quando cambiamo sistema di riferimento cartesiano. Siano  $\mathcal{R} = RC(O, A_1, A_2, A_3)$  e  $\mathcal{R}' = RC(O', A'_1, A'_2, A'_3)$  due sistemi di riferimento cartesiani, e  $P \in \mathcal{E}^3$ . Indichiamo con  $x, x' \in \mathbb{R}^3$  le coordinate di  $P$  rispetto a  $\mathcal{R}$ , rispettivamente a  $\mathcal{R}'$ . Trattandosi di particolari sistemi di riferimento affini, quanto visto nel Paragrafo 10.7 ci dice che  $x$  e  $x'$  sono legati da

$$x = Bx' + c,$$

dove  $c \in \mathbb{R}^3$  e  $B \in GL_3(\mathbb{R})$  è la matrice di cambiamento di base. In particolare,  $B$  contiene per colonne le coordinate della base di  $\mathcal{V}_{O'}^3$ , relativa a  $\mathcal{R}'$ , che per ipotesi è una base ortonormale; quindi (Corollario 11.19) la matrice  $B$  è ortogonale. Viceversa, se  $B$  è ortogonale allora la base di  $\mathcal{V}_O^3$ , relativa a  $\mathcal{R}'$  è ortonormale, per cui  $\mathcal{R}'$  è un sistema di riferimento cartesiano. Abbiamo quindi dimostrato la seguente proposizione.

### Proposizione 12.1

Sia  $\mathcal{R}$  un sistema di riferimento cartesiano su  $\mathcal{E}^3$  (o  $\mathcal{E}^2$ ), e  $\mathcal{R}'$  un altro sistema di riferimento affine. Allora anche  $\mathcal{R}'$  è un sistema di riferimento cartesiano se e solo se la formula di cambiamento di coordinate da  $\mathcal{R}$  a  $\mathcal{R}'$  è data da  $x = Bx' + c$ , dove  $B$  è una matrice ortogonale.

### Definizione 12.3

Un cambiamento di coordinate di  $\mathcal{E}^2$  o  $\mathcal{E}^3$  che trasforma un sistema di riferimento cartesiano in un altro sistema di riferimento cartesiano verrà detto *cambiamento di coordinate metrico*.

Fissiamo allora una volta per tutte un riferimento cartesiano  $RC(O, A_1, A_2, A_3)$ , in modo da identificare  $\mathcal{E}^3$  con  $\mathbb{R}^3$  (ed  $\mathcal{E}^2$  con  $\mathbb{R}^2$ ); poniamo anche  $\vec{i} = \overrightarrow{OA_1}$ ,  $\vec{j} = \overrightarrow{OA_2}$  e  $\vec{k} = \overrightarrow{OA_3}$ .

Il nostro primo obiettivo è definire l'angolo fra rette e/o piani. Cominciamo con le rette. Ci troviamo subito di fronte a un piccolo ostacolo: due rette che si intersecano determinano *due* angoli convessi, e non uno (vedi la Figura 12.1); quale prendiamo?

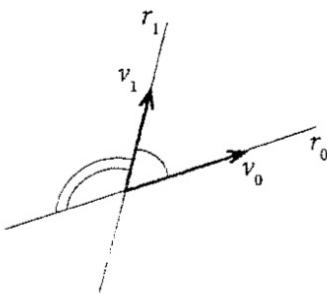


Figura 12.1 - L'angolo fra due rette.

La risposta è che la domanda è mal posta: entrambi vanno bene, per poter decidere occorrono altre informazioni. Per esempio, scegliendo un verso di percorrenza delle due rette (equivalentemente, fissando un vettore direttore per le due rette), allora quale angolo bisogna prendere risulta chiarissimo: l'unico determinato dalle due semirette positive. Questo suggerisce una definizione.

#### Definizione 12.4

Una *retta orientata* è una retta in cui si è fissato un vettore direttore (e quindi un verso di percorrenza). Due diversi vettori direttori determinano la stessa *orientazione* (lo stesso verso di percorrenza) se e solo se sono un multiplo positivo l'uno dell'altro (vedi anche il Paragrafo 10.8). In particolare, esiste un unico vettore di lunghezza unitaria che determina l'orientazione scelta: il *versore direttore* della retta orientata. Se  $v = \alpha\vec{i} + \beta\vec{j} + \gamma\vec{k}$  è l'espressione del versore direttore  $v$  come combinazione lineare della base  $\{\vec{i}, \vec{j}, \vec{k}\}$  di  $\mathcal{V}_O^3$ , allora  $\alpha, \beta, \gamma \in \mathbb{R}$  sono detti *coseni direttori* della retta.

I coseni direttori sono effettivamente dei coseni: infatti la Proposizione 11.4 implica che  $\alpha = \langle v, \vec{i} \rangle = \cos \widehat{v\vec{i}}$ , dato che  $\|v\| = 1 = \|\vec{i}\|$ . Analogamente,  $\beta = \cos \widehat{v\vec{j}}$  e  $\gamma = \cos \widehat{v\vec{k}}$ .

#### Definizione 12.5

L'angolo  $\widehat{r_0 r_1}$  fra due rette orientate  $r_0$  ed  $r_1$  è l'angolo fra i loro versori direttori.

In particolare, l'angolo fra due rette dipende soltanto dai loro sottospazi di giacitura<sup>(1)</sup>.

Dunque se le due rette  $r_0$  ed  $r_1$  hanno equazioni parametriche  $P = P_0 + t v_0$  e  $P = P_1 + s v_1$  rispettivamente, dove  $v_0 = (l_0, m_0, n_0)$  e  $v_1 = (l_1, m_1, n_1)$  sono vettori direttori che determinano le orientazioni scelte, allora l'angolo fra le rette  $r_0$  ed  $r_1$  è dato da

$$\cos \widehat{r_0 r_1} = \frac{\langle v_0, v_1 \rangle}{\|v_0\| \|v_1\|} = \frac{l_0 l_1 + m_0 m_1 + n_0 n_1}{\sqrt{l_0^2 + m_0^2 + n_0^2} \sqrt{l_1^2 + m_1^2 + n_1^2}}. \quad (12.1)$$

In particolare, due rette sono ortogonali se e solo se  $\cos \widehat{r_0 r_1} = 0$ , cioè se e solo se

$$l_0 l_1 + m_0 m_1 + n_0 n_1 = 0 \quad (12.2)$$

(condizione di ortogonalità).

#### Osservazione 12.1

In  $\mathcal{E}^2$  l'angolo fra due rette orientate si definisce e si calcola nello stesso modo; semplicemente nelle formule compare una coordinata in meno.

<sup>(1)</sup> E quindi due rette possono essere ortogonali anche senza intersecarsi.

Consideriamo adesso una retta orientata  $r_0$  passante per l'origine, di vettore direttore  $v_0 = (l_0, m_0, n_0)$ . Un altro vettore  $v = (x, y, z)$  è ortogonale a questa retta se e solo se  $l_0x + m_0y + n_0z = 0$ ; quindi l'ortogonale  $r_0^\perp$  di  $r_0$  è un piano (come si poteva anche dedurre dalla Proposizione 11.11). In particolare, l'equazione del piano ortogonale a  $r_0$  e passante per il punto  $P_0 = (x_0, y_0, z_0)$  è

$$l_0(x - x_0) + m_0(y - y_0) + n_0(z - z_0) = 0. \quad (12.3)$$

Nota che questa è l'equazione del piano ortogonale a  $r_0$  anche se  $r_0$  non passa per l'origine (perché?).

### Definizione 12.6

La *proiezione ortogonale* di un punto  $P_0$  su una retta  $r$  è il punto d'intersezione di  $r$  con l'unico piano passante per  $P_0$  ortogonale a  $r$ .

Per calcolare l'angolo fra due piani, ci serve la seguente definizione.

### Definizione 12.7

Un *piano orientato* è un piano  $\pi$  in cui è stata fissata una base nel sottospazio di giacitura. Due basi definiscono la stessa *orientazione* se la matrice di cambiamento di base ha determinante positivo (vedi il Paragrafo 10.8).

Sia  $\pi$  un piano orientato, con base  $\{v_1, v_2\}$  del sottospazio di giacitura. Se  $v_3$  e  $-v_3$  sono i due versori direttori della retta ortogonale  $\pi^\perp$ , allora solo una delle due basi  $\{v_1, v_2, v_3\}$  e  $\{v_1, v_2, -v_3\}$  ha la stessa orientazione della base  $\{\vec{i}, \vec{j}, \vec{k}\}$  (vedi l'Esercizio 10.29). Dunque siamo portati a dare la seguente definizione:

### Definizione 12.8

Sia  $\pi$  un piano orientato, e  $\{v_1, v_2\}$  una base del sottospazio di giacitura che determina l'orientazione scelta. Il *versore ortogonale* di  $\pi$  è l'unico versore direttore  $v_3$  della retta ortogonale al piano tale che  $\{v_1, v_2, v_3\}$  sia una base di  $\mathcal{V}_O^3$  con la stessa orientazione della base  $\{\vec{i}, \vec{j}, \vec{k}\}$ .

### Osservazione 12.2

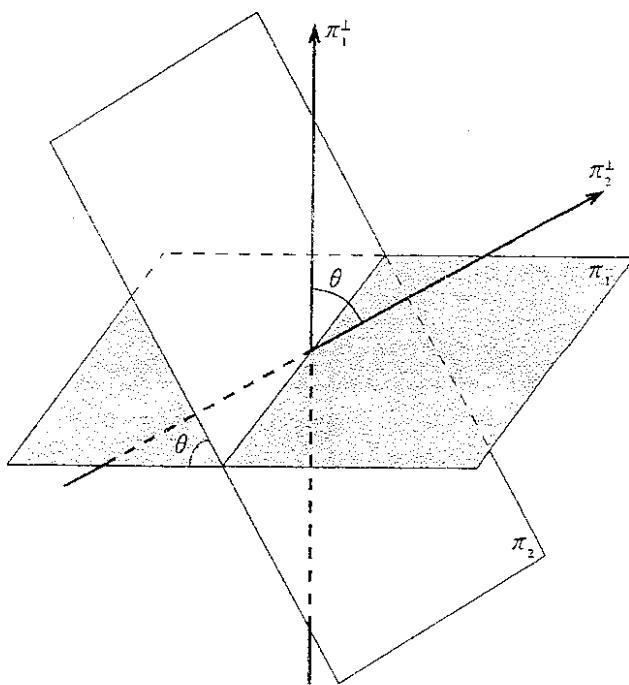
Prendiamo un piano  $\pi$  di equazione cartesiana  $ax + by + cz = d$ . Per quanto detto prima dell'equazione (12.3), il vettore  $v = (a, b, c)$  è un vettore direttore della retta  $\pi^\perp$  (passante per l'origine) ortogonale al piano. Dunque il versore ortogonale di  $\pi$  è  $v/\|v\|$  oppure  $-v/\|v\|$  a seconda dell'orientazione scelta per  $\pi$ .

### Definizione 12.9

L'*angolo*  $\widehat{\pi_0\pi_1}$  fra due piani orientati  $\pi_0$  e  $\pi_1$  è per definizione l'angolo fra i loro versori ortogonali.

### Osservazione 12.3

In geometria elementare l'“angolo diedro” formato da due piani non paralleli è l'angolo formato dalle due rette ottenute intersecando i due piani con un terzo piano ortogonale a entrambi. La nostra definizione sostituisce a queste due rette altre due, ruotate di  $\pi/2$ ; l'angolo quindi non è cambiato (vedi la Figura 12.2). Infine, di nuovo l'angolo fra due piani dipende solo dai sottospazi di giacitura.



**Figura 12.2** - L'angolo fra due piani.

Per calcolare l'angolo fra due piani  $\pi_0$  e  $\pi_1$ , supponiamo che abbiano equazione cartesiana  $a_0x + b_0y + c_0z = d_0$  e  $a_1x + b_1y + c_1z = d_1$  rispettivamente. Allora i vettori  $v_0 = (a_0, b_0, c_0)$  e  $v_1 = (a_1, b_1, c_1)$  sono vettori direttori delle rette ortogonali; possiamo supporre che i piani siano orientati in modo che  $v_0/\|v_0\|$  e  $v_1/\|v_1\|$  siano i loro versori ortogonali. Allora

$$\cos \widehat{\pi_0 \pi_1} = \cos \widehat{v_0 v_1} = \frac{a_0 a_1 + b_0 b_1 + c_0 c_1}{\sqrt{a_0^2 + b_0^2 + c_0^2} \sqrt{a_1^2 + b_1^2 + c_1^2}}. \quad (12.4)$$

In particolare, i due piani sono ortogonali se e solo se

$$a_0 a_1 + b_0 b_1 + c_0 c_1 = 0 \quad (12.5)$$

(condizione di ortogonalità). La stessa formula, con una coordinata in meno, ci fornisce la condizione di ortogonalità fra due rette in  $\mathcal{E}^2$  di equazione cartesiana  $a_0x + b_0y = c_0$  e  $a_1x + b_1y = c_1$ :

$$a_0 a_1 + b_0 b_1 = 0.$$

Di nuovo, dato un piano  $\pi$  di equazione cartesiana  $ax + by + cz = d$  e un punto  $P_0$ , esiste un'unica retta ortogonale a  $\pi$  passante per  $P_0$ , che ha equazione parametrica  $P = P_0 + tv$ , dove  $v = (a, b, c)$ .

### Definizione 12.10

La proiezione ortogonale di un punto  $P_0$  su un piano  $\pi$  è il punto d'intersezione di  $\pi$  con l'unica retta passante per  $P_0$  ortogonale al piano.

Resta da definire l'angolo fra una retta e un piano. Prendiamo una retta  $r$  orientata con vettore direttore  $v = (l, m, n)$ , e un piano  $\pi$  di equazione  $ax + by + cz = d$  non ortogonale alla retta  $r$  (se sono ortogonali l'angolo è chiaramente  $\pi/2$ ). Consideriamo allora la retta  $r'$  proiezione ortogonale di  $r$  su  $\pi$  (quando  $r$  e  $\pi$  sono ortogonali la proiezione ortogonale di  $r$  su  $\pi$  si riduce a un punto); la retta  $r'$  ha come vettore direttore la proiezione ortogonale di  $v$  sul sottospazio di giacitura di  $\pi$ .

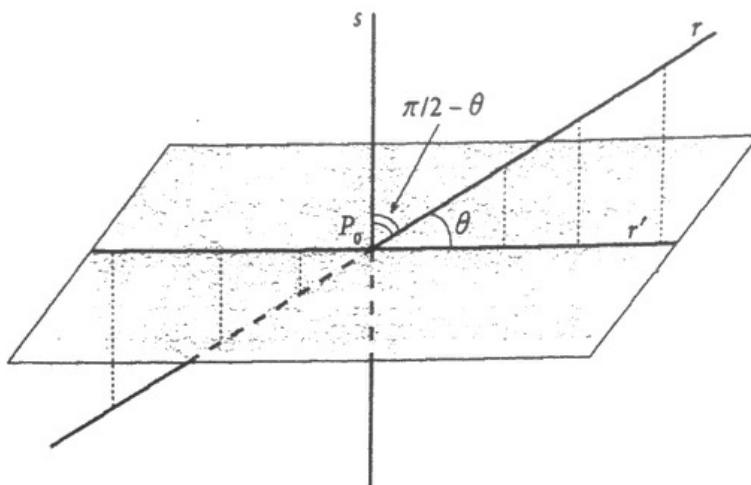


Figura 12.3 - L'angolo fra una retta e un piano.

**Definizione 12.11**

L'angolo  $\hat{r}\pi$  fra una retta  $r$  e un piano  $\pi$  è l'angolo fra  $r$  e la sua proiezione ortogonale su  $\pi$  se  $r$  e  $\pi$  non sono ortogonali; è pari a  $\pi/2$  se sono ortogonali. Cambiando l'orientazione di  $r$  cambia anche quella di  $r'$ , per cui l'angolo non dipende dall'orientazione di  $r$  scelta.

Per definizione, l'angolo fra una retta e un piano è sempre compreso fra 0 e  $\pi/2$ ; quindi per trovarlo esplicitamente ci basta calcolarne il seno. Se  $r$  e  $\pi$  sono paralleli, allora il loro angolo è zero; altrimenti, sia  $P_0$  il loro punto d'intersezione, ed  $s$  la retta ortogonale a  $\pi$  passante per  $P_0$ .

Se indichiamo con  $\theta \in [0, \pi/2]$  l'angolo fra  $r$  e  $\pi$ , a meno di moltiplicare per  $-1$  la direzione del piano possiamo supporre che la retta orientata  $s$  di vettore direttore  $(a, b, c)$  abbia un angolo di  $\pi/2 - \theta$  radianti con  $r$  (confronta l'Esercizio 12.3).

Allora (vedi la Figura 12.3)

$$\sin \hat{r}\pi = \sin \hat{rr'} = \cos \hat{rs} = \frac{al + bm + cn}{\sqrt{a^2 + b^2 + c^2} \sqrt{l^2 + m^2 + n^2}}.$$

In particolare, retta e piano sono ortogonali se e solo se  $\sin \hat{r}\pi = 1$  se e solo se (per la diseguaglianza di Cauchy-Schwarz) i vettori  $(a, b, c)$  e  $(l, m, n)$  sono linearmente indipendenti (*condizione di ortogonalità*).

Vediamo ora un esempio di un tipico problema di geometria metrica.

**Esempio 12.1**

Vogliamo trovare la retta  $r$  passante per  $P_0 = (1, 0, -1)$  ortogonale e incidente alla retta di equazioni cartesiane

$$\begin{cases} x - 2z = 1, \\ y + 2z = 0. \end{cases}$$

Prima di tutto, le equazioni parametriche di  $s$  sono (esercizio)

$$P = P_1 + tv_1 = \begin{vmatrix} 1 \\ 0 \\ 0 \end{vmatrix} + t \begin{vmatrix} 2 \\ -2 \\ 1 \end{vmatrix};$$

Ora diamo trovare le equazioni parametriche di  $r$ , che devono essere del tipo

$$P = P_0 + sv_0 = \begin{vmatrix} 1 \\ 0 \\ -1 \end{vmatrix} + s \begin{vmatrix} l \\ m \\ n \end{vmatrix}.$$

Ora, la condizione di ortogonalità ci dice che

$$2l - 2m + n = 0. \quad (12.7)$$

In particolare,  $\text{rg } [v_0 \ v_1] = 2$ . La condizione d'incidenza (vedi il Paragrafo 10.4) ci dice invece che

$$0 = \det |v_0 \ v_1 \ P_0 - P_1| = \det \begin{vmatrix} l & 2 & 0 \\ m & -2 & 0 \\ m & 1 & -1 \end{vmatrix} = 2l + 2m. \quad (12.8)$$

Risolvendo il sistema composto da (12.7) e (12.8) otteniamo quindi la soluzione: la retta  $r$  ha equazioni parametriche

$$P = P_0 + sv_0 = \begin{vmatrix} 1 \\ 0 \\ -1 \end{vmatrix} + s \begin{vmatrix} -1 \\ 1 \\ 4 \end{vmatrix}.$$

Ma vi è anche un altro modo per risolvere questo problema. Prima di tutto, la retta  $r$  deve appartenere al piano  $\pi_1$  passante per  $P_0$  e ortogonale alla retta  $s$ , che ha equazione cartesiana

$$2(x - 1) - 2y + (z + 1) = 0. \quad (12.9)$$

Poi la retta  $r$  deve anche appartenere al piano  $\pi_2$  passante per  $P_0$  e contenente  $s$  (infatti questo piano contiene sia  $P_0$  che il punto d'intersezione fra  $r$  ed  $s$ ). Per quanto visto nel Paragrafo 10.5.(I) il piano  $\pi_2$  ha equazione

$$-2(x - 2z - 1) - 2(y + 2z) = 0. \quad (12.10)$$

Mettendo insieme (12.9) e (12.10) abbiamo trovato che  $r$  ha equazioni cartesiane

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

□

## 12.2 Distanze

Passiamo ora a calcolare le distanze fra punti, rette e piani. La distanza fra due punti  $P_0 = (x_0, y_0, z_0)$  e  $P_1 = (x_1, y_1, z_1)$  è ovviamente data da

$$d(P_0, P_1) = \|\overrightarrow{OP_1} - \overrightarrow{OP_0}\| = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2}.$$

Più interessante è invece il calcolo della distanza fra un punto e una retta.

### Definizione 12.12

Siano  $r$  una retta e  $P_0$  un punto. La distanza  $d(P_0, r)$  fra  $P_0$  ed  $r$  è per definizione la distanza fra  $P_0$  e la sua proiezione ortogonale su  $r$ .

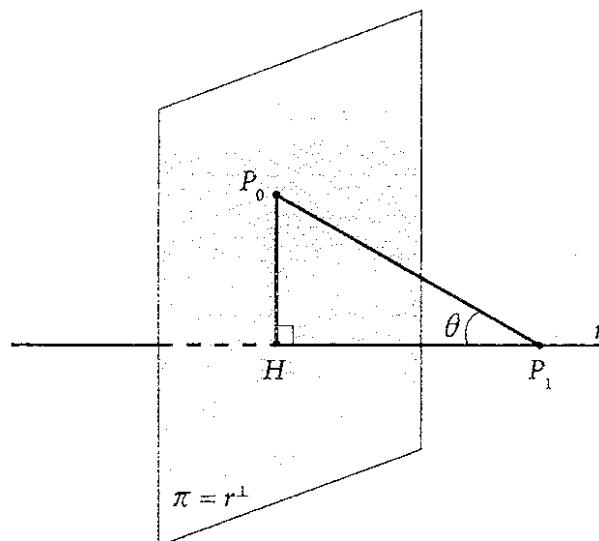


Figura 12.4 - La distanza fra un punto e una retta.

**Osservazione 12.4**

La scelta di questa definizione per la distanza è dovuta al Corollario 11.12.

Supponiamo che la retta  $r$  abbia equazione parametrica  $P = P_1 + tv_0$ , e indichiamo con  $\theta$  l'angolo fra  $v_0$  e il vettore  $v_1 = \overrightarrow{OP_1} - \overrightarrow{OP_0}$  (vedi la Figura 12.4, dove abbiamo supposto  $r$  orientata in modo che  $\theta \in [0, \pi/2]$ ).

Sia  $H$  la proiezione ortogonale di  $P_0$  su  $r$ . Allora, considerando il triangolo rettangolo di vertici  $P_0$ ,  $P_1$  e  $H$ , vediamo che

$$\begin{aligned}
 d(P_0, r) &= d(P_0, H) = d(P_0, P_1) |\sin \theta| = d(P_0, P_1) \sqrt{1 - (\cos \theta)^2} \\
 &= \|v_1\| \sqrt{1 - \frac{|\langle v_1, v_0 \rangle|^2}{\|v_0\|^2 \|v_1\|^2}} = \sqrt{\|v_1\|^2 - \frac{|\langle v_1, v_0 \rangle|^2}{\|v_0\|^2}} \\
 &= \sqrt{d(P_0, P_1)^2 - \frac{|\langle \overrightarrow{OP_1} - \overrightarrow{OP_0}, v_0 \rangle|^2}{\|v_0\|^2}}.
 \end{aligned} \tag{12.11}$$

**Osservazione 12.5**

Ovviamente, questa stessa formula letta in  $\mathcal{E}^2$  dà la distanza fra un punto e una retta del piano espressa in equazioni parametriche. Un'altra formula utilizzabile in  $\mathcal{E}^3$  è invece descritta nel prossimo paragrafo (Osservazione 12.9).

**Esempio 12.2**

Vogliamo la distanza fra il punto  $P_0 = (1, 0, 1)$  e la retta  $r$  di equazione parametrica

$$P = \begin{vmatrix} 0 \\ 0 \\ 1 \end{vmatrix} + t \begin{vmatrix} -1 \\ 1 \\ 0 \end{vmatrix}.$$

Si ha

$$\overrightarrow{OP_1} - \overrightarrow{OP_0} = \begin{vmatrix} -1 \\ 0 \\ 0 \end{vmatrix},$$

per cui

$$d(P_0, r) = \sqrt{1 - \frac{1}{2}} = \frac{1}{\sqrt{2}}.$$

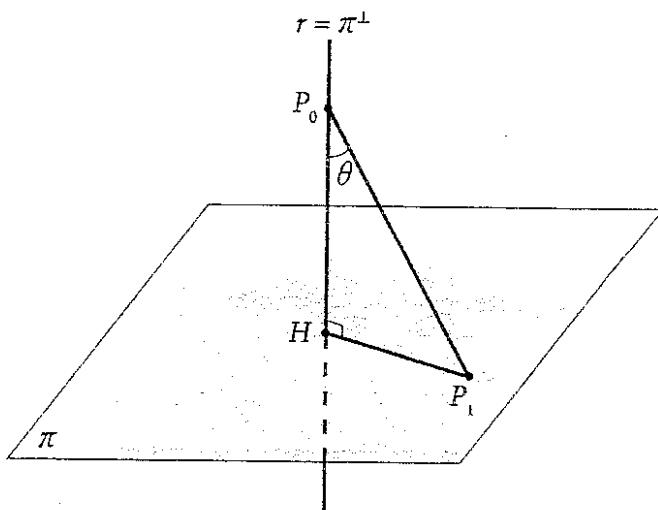
□

Passiamo ora alla distanza fra un punto e un piano.

### Definizione 12.13 |

Sia  $\pi$  un piano e  $P_0$  un punto. La *distanza*  $d(P_0, \pi)$  fra il punto e il piano è la distanza fra  $P_0$  e la sua proiezione ortogonale su  $\pi$ .

Considera ora un punto  $P_0$  di coordinate  $(x_0, y_0, z_0)$ , e  $\pi$  un piano di equazione cartesiana  $ax + by + cz = d$ . Sia  $H$  la proiezione ortogonale di  $P_0$  su  $\pi$ ; indichiamo con  $r$  la retta ortogonale a  $\pi$  passante per  $P_0$  (e  $H$ ; vedi la Figura 12.5).



**Figura 12.5** - La distanza fra un punto e un piano.

Sia  $P_1 = (x_1, y_1, z_1) \in \pi$  un punto del piano, e indichiamo con  $\theta$  l'angolo fra il vettore  $v_1 = \overrightarrow{OP_0} - \overrightarrow{OP_1}$  e il vettore  $v_0 = (a, b, c)$ , che è un vettore direttore della retta  $r$ ; a meno di moltiplicare per  $-1$  l'equazione del piano possiamo supporre che  $\theta \in [0, \pi/2]$ . Allora considerando il triangolo rettangolo di vertici  $P_0$ ,  $P_1$  e  $H$  vediamo che

$$\begin{aligned} d(P_0, \pi) &= d(P_0, H) = d(P_0, P_1) |\cos \theta| = \frac{|\langle v_0, v_1 \rangle|}{\|v_0\|} \\ &= \frac{|a(x_0 - x_1) + b(y_0 - y_1) + c(z_0 - z_1)|}{\sqrt{a^2 + b^2 + c^2}}. \end{aligned}$$

Ma ora il fatto che  $P_1 \in \pi$  ci dice che  $ax_1 + by_1 + cz_1 = d$ ; quindi

$$d(P_0, \pi) = \frac{|ax_0 + by_0 + cz_0 - d|}{\sqrt{a^2 + b^2 + c^2}}. \quad (12.12)$$

### Osservazione 12.6

Una formula assolutamente analoga (in cui non compare  $c$ ) fornisce la distanza in  $\mathbb{E}^2$  fra un punto e una retta del piano in equazioni cartesiane.

La distanza fra una retta e un piano, o fra due piani, è ottenuta molto facilmente: se la retta e il piano (o i due piani) sono incidenti, la distanza è zero; se invece sono paralleli, è

uguale alla distanza di un qualunque punto della retta (o del primo piano) dal (secondo) piano, per cui si calcola come abbiamo appena visto. Dunque l'unica distanza rimasta da determinare è quella fra due rette.

Siano  $r_0$  ed  $r_1$  due rette, di equazione parametrica rispettivamente  $P = P_0 + tv_0$  e  $P = P_1 + sv_1$ . Se le due rette non sono parallele, allora esiste un unico piano parallelo a  $r_0$  che contiene  $r_1$ : è il piano  $\pi$  di equazione parametrica  $P = P_1 + sv_1 + tv_0$ .

### Definizione 12.14

Siano  $r_0$  ed  $r_1$  due rette. Se sono parallele, definiamo loro *distanza*  $d(r_0, r_1)$  il numero  $d(P_0, r_1)$ , dove  $P_0$  è un qualunque punto di  $r_0$ ; se non sono parallele, la loro *distanza*  $d(r_0, r_1)$  è invece data da  $d(P_0, \pi)$ , dove  $P_0$  è ancora un punto qualunque di  $r_0$  e  $\pi$  è l'unico piano contenente  $r_1$  e parallelo a  $r_0$ .

Vediamo allora di trovare la formula della distanza di  $r_0$  ed  $r_1$  quando non sono parallele (vedi la Figura 12.6). Scriviamo  $v_0 = (l_0, m_0, n_0)$  e  $v_1 = (l_1, m_1, n_1)$ . Per l'Esercizio 10.7, il piano  $\pi$  ha equazione cartesiana  $ax + by + cz = ax_1 + by_1 + cz_1$ , dove

$$a = \det \begin{vmatrix} m_0 & m_1 \\ n_0 & n_1 \end{vmatrix} = m_0n_1 - m_1n_0, \quad b = -\det \begin{vmatrix} l_0 & l_1 \\ n_0 & n_1 \end{vmatrix} = n_0l_1 - n_1l_0,$$

$$c = \det \begin{vmatrix} l_0 & l_1 \\ m_0 & m_1 \end{vmatrix} = l_0m_1 - l_1m_0.$$

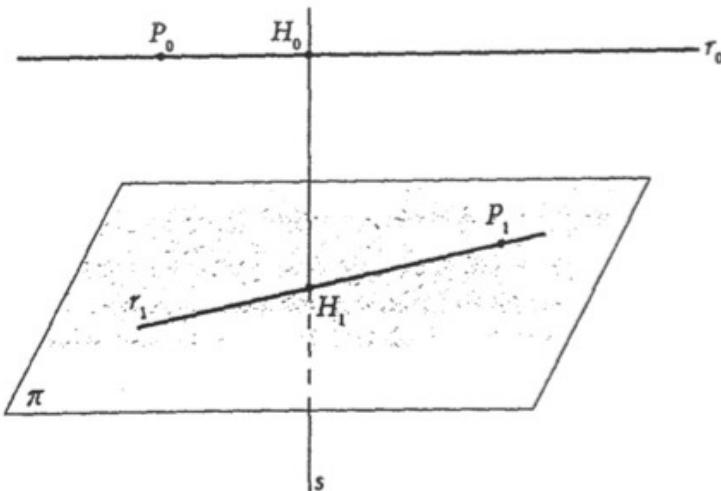


Figura 12.6 - La distanza fra due rette.

In particolare si verifica facilmente (esercizio) che

$$ax_0 + by_0 + cz_0 - (ax_1 + by_1 + cz_1) = \det |v_0 \ v_1 \ P_0 - P_1|,$$

e che

$$\begin{aligned} a^2 + b^2 + c^2 &= (m_0n_1 - m_1n_0)^2 + (n_0l_1 - n_1l_0)^2 + (l_0m_1 - l_1m_0)^2 \\ &= (l_0^2 + m_0^2 + n_0^2)(l_1^2 + m_1^2 + n_1^2) - (l_0l_1 + m_0m_1 + n_0n_1)^2 \\ &= \|v_0\|^2\|v_1\|^2 - (\langle v_0, v_1 \rangle)^2 = \|v_0\|^2\|v_1\|^2(1 - (\cos \widehat{v_0 v_1})^2) \\ &= \|v_0\|^2\|v_1\|^2(\sin \widehat{v_0 v_1})^2. \end{aligned}$$

Quindi applicando (12.12) troviamo che la distanza fra  $r_0$  ed  $r_1$  è data da

$$d(r_0, r_1) = d(P_0, \pi) = \left| \frac{\det |v_0 \ v_1 \ P_1 - P_0|}{\|v_0\| \|v_1\| \sin \widehat{v_0 v_1}} \right|. \quad (12.13)$$

Nota che le due rette (non parallele) sono incidenti se e solo se  $\det |v_0 \ v_1 \ P_1 - P_0| = 0$ , cioè se e solo se si ha  $d(r_0, r_1) = 0$ , come dev'essere. Nel prossimo paragrafo (Osservazione 12.10) vedremo un'ulteriore interpretazione di questa formula.

### Esempio 12.3

Vogliamo calcolare la distanza fra le rette  $r_0$  ed  $r_1$  di equazione parametrica rispettivamente

$$r_0: \quad P = \begin{vmatrix} 1 \\ 0 \\ 1 \end{vmatrix} + s \begin{vmatrix} 1 \\ 1 \\ 1 \end{vmatrix} \quad \text{e} \quad r_1: \quad P = \begin{vmatrix} -1 \\ 1 \\ 0 \end{vmatrix} + t \begin{vmatrix} -2 \\ 1 \\ 2 \end{vmatrix}.$$

Utilizzando le notazioni precedenti troviamo  $a = 1$ ,  $b = -4$  e  $c = 3$ ; dunque

$$\|v_0\| \|v_1\| |\sin \widehat{v_0 v_1}| = (a^2 + b^2 + c^2)^{1/2} = \sqrt{26}, \quad \det |v_0 \ v_1 \ P_1 - P_0| = -5,$$

e quindi  $d(r_0, r_1) = 5/\sqrt{26}$ . ◻

## 12.3 Il prodotto vettore

Questo paragrafo è dedicato alla descrizione di una particolare operazione definita unicamente su  $\mathbb{R}^3$ . Due vettori  $v, w \in \mathbb{R}^3$  linearmente indipendenti determinano un unico piano  $\pi$  che li contiene, di equazioni parametriche  $P = sv + tw$ . L'equazione cartesiana di  $\pi$  è  $ax + by + cz = 0$ , dove (vedi l'Esercizio 10.7)

$$a = \det \begin{vmatrix} v_2 & w_2 \\ v_3 & w_3 \end{vmatrix}, \quad b = -\det \begin{vmatrix} v_1 & w_1 \\ v_3 & w_3 \end{vmatrix}, \quad c = \det \begin{vmatrix} v_1 & w_1 \\ v_2 & w_2 \end{vmatrix}.$$

Inoltre, come abbiamo visto nel Paragrafo 12.1, il vettore  $(a, b, c)$  è ortogonale al piano  $\pi$ ; in particolare, è ortogonale sia a  $v$  che a  $w$ . Dunque a ogni coppia di vettori (linearmente indipendenti) di  $\mathbb{R}^3$  possiamo associare un terzo vettore ortogonale a entrambi.

### Definizione 12.15

Siano  $v, w \in \mathbb{R}^3$ . Il *prodotto vettore* di  $v$  e  $w$  è il vettore

$$v \wedge w = \det \begin{vmatrix} v_2 & w_2 \\ v_3 & w_3 \end{vmatrix} e_1 - \det \begin{vmatrix} v_1 & w_1 \\ v_3 & w_3 \end{vmatrix} e_2 + \det \begin{vmatrix} v_1 & w_1 \\ v_2 & w_2 \end{vmatrix} e_3 \in \mathbb{R}^3. \quad (12.14)$$

A volte si scrive  $v \times w$  invece di  $v \wedge w$ .

### Osservazione 12.7

La formula (12.14) viene spesso scritta nel seguente modo

$$v \wedge w = \det \begin{vmatrix} e_1 & v_1 & w_1 \\ e_2 & v_2 & w_2 \\ e_3 & v_3 & w_3 \end{vmatrix},$$

dove si intende che questo "determinante" viene sviluppato solo lungo la prima colonna.

Vediamo che proprietà possiede il prodotto vettore.

### Proposizione 12.2

Sia  $v \wedge w$  il prodotto vettore di due vettori  $v, w \in \mathbb{R}^3$ . Allora:

- (1)  $v \wedge w$  è ortogonale sia a  $v$  che a  $w$ ;
- (2)  $\|v \wedge w\|$  è l'area del parallelogramma di vertici  $O, v, w$  e  $v + w$ , cioè

$$\|v \wedge w\| = \|v\| \|w\| |\sin \theta|,$$

dove  $\theta$  è l'angolo fra  $v$  e  $w$ ;

- (3)  $v \wedge w = 0$  se e solo se  $v$  e  $w$  sono linearmente dipendenti;
- (4) se  $v$  e  $w$  sono linearmente indipendenti, allora la base  $\{v, w, v \wedge w\}$  determina la stessa orientazione della base canonica;
- (5)  $v \wedge w$  è l'unico vettore di  $\mathbb{R}^3$  che soddisfa (1), (2) e (4);
- (6)  $w \wedge v = -v \wedge w$ , e l'applicazione  $\wedge : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$  è bilineare.

### Dimostrazione

- (1) Infatti

$$\langle v, v \wedge w \rangle = \det \begin{vmatrix} v_1 & v_1 & w_1 \\ v_2 & v_2 & w_2 \\ v_3 & v_3 & w_3 \end{vmatrix} = 0 = \det \begin{vmatrix} w_1 & v_1 & w_1 \\ w_2 & v_2 & w_2 \\ w_3 & v_3 & w_3 \end{vmatrix} = \langle w, v \wedge w \rangle.$$

- (2) Prima di tutto, la Figura 12.7 mostra che l'area del parallelogramma di vertici  $O, v, w$  e  $v + w$  è effettivamente data da  $\|v\| \|w\| |\sin \theta|$ . Poi,

$$\begin{aligned} \|v \wedge w\|^2 &= (v_2 w_3 - w_2 v_3)^2 + (v_1 w_3 - w_1 v_3)^2 + (v_1 w_2 - w_1 v_2)^2 \\ &= (v_1^2 + v_2^2 + v_3^2)(w_1^2 + w_2^2 + w_3^2) - (v_1 w_1 + v_2 w_2 + v_3 w_3)^2 \\ &= \|v\|^2 \|w\|^2 (1 - \cos^2 \theta), \end{aligned}$$

e quindi (2) è soddisfatta.

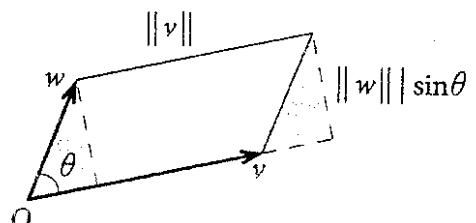


Figura 12.7 - La norma del prodotto vettore.

- (3) I vettori  $v$  e  $w$  sono linearmente dipendenti se e solo se  $\sin \theta = 0$ , per cui l'asserto segue da (2).

- (4) Infatti

$$\begin{aligned} \det |v \ w \ v \wedge w| &= \det \begin{vmatrix} v_1 & w_1 & \det \begin{vmatrix} v_2 & w_2 \\ v_3 & w_3 \end{vmatrix} \\ v_2 & w_2 & -\det \begin{vmatrix} v_1 & w_1 \\ v_3 & w_3 \end{vmatrix} \\ v_3 & w_3 & \det \begin{vmatrix} v_1 & w_1 \\ v_2 & w_2 \end{vmatrix} \end{vmatrix} \\ &= \left( \det \begin{vmatrix} v_2 & w_2 \\ v_3 & w_3 \end{vmatrix} \right)^2 + \left( \det \begin{vmatrix} v_1 & w_1 \\ v_3 & w_3 \end{vmatrix} \right)^2 + \left( \det \begin{vmatrix} v_1 & w_1 \\ v_2 & w_2 \end{vmatrix} \right)^2 \geq 0, \end{aligned}$$

ed è uguale a zero se e solo se  $v \wedge w = O$ , cioè se e solo se  $v$  e  $w$  sono linearmente dipendenti, e anche (4) vale.

(5) Supponiamo che  $u \in \mathbb{R}^3$  soddisfi (1), (2) e (4). Se  $v$  e  $w$  sono linearmente dipendenti, allora (2) implica  $u = O = v \wedge w$ . Se sono linearmente indipendenti, l'ortogonale di  $\text{Span}(v, w)$  ha dimensione 1, per cui si ha necessariamente  $u = \lambda(v \wedge w)$  per qualche  $\lambda \in \mathbb{R}$ . La (2) allora implica  $\lambda = \pm 1$ , e la (4) ci dà  $\lambda = 1$ .

(6) È una conseguenza immediata della definizione.  $\square$

### Osservazione 12.8

Il prodotto vettore non è un “prodotto” come lo intendiamo di solito: *non* è associativo. Per esempio,

$$e_1 \wedge (e_2 \wedge e_2) = e_1 \wedge O = O \neq -e_1 = e_3 \wedge e_2 = (e_1 \wedge e_2) \wedge e_2.$$

### Osservazione 12.9

Il prodotto vettore può venire usato per esprimere la distanza di un punto da una retta. Consideriamo il punto  $P_0 \in \mathbb{R}^3$  e la retta  $r$  di equazione parametrica  $P = P_1 + t\vec{v}_1$ . Ponendo  $\vec{v}_0 = \overrightarrow{OP_1} - \overrightarrow{OP_0}$ , la formula (12.11) ci dice che

$$d(P_0, r) = d(P_0, P_1)|\sin \theta| = \|\vec{v}_0\| |\sin \theta| = \frac{\|\vec{v}_0 \wedge \vec{v}_1\|}{\|\vec{v}_1\|}. \quad (12.15)$$

Questa uguaglianza ha una interessante interpretazione geometrica: consideriamo infatti il parallelogrammo avente come vertici i punti  $P_0, P_1, P_1 + \vec{v}_1$  e  $P_0 + \vec{v}_1$ . La Proposizione 12.2. ci dice che l'area di tale parallelogrammo è data da  $\|\vec{v}_0 \wedge \vec{v}_1\|$ ; ma, d'altra parte, l'area di un parallelogrammo si ottiene moltiplicando la lunghezza della base, che in questo caso vale  $\|\vec{v}_1\|$ , per l'altezza, che è proprio  $d(P_0, r)$ .

### Osservazione 12.10

Possiamo usare il prodotto vettore anche per esprimere la distanza fra due rette. Date  $r_0$  e  $r_1$  rette di equazione parametrica  $P = P_j + t\vec{v}_j$  con  $\vec{v}_j = (x_j, y_j, z_j)$  per  $j = 0, 1$ , poniamo  $\vec{v}_2 = \overrightarrow{OP_1} - \overrightarrow{OP_0}$ . Si verifica subito che

$$\langle \vec{v}_0 \wedge \vec{v}_1, \vec{v}_2 \rangle = \det \begin{vmatrix} \vec{v}_0 & \vec{v}_1 & \vec{v}_2 \end{vmatrix}, \quad (12.16)$$

e di conseguenza la (12.13) diventa

$$d(r_0, r_1) = \frac{|\langle \vec{v}_0 \wedge \vec{v}_1, \overrightarrow{OP_1} - \overrightarrow{OP_0} \rangle|}{\|\vec{v}_0 \wedge \vec{v}_1\|} \quad (12.17)$$

Quest'ultima osservazione suggerisce la seguente definizione.

### Definizione 12.16

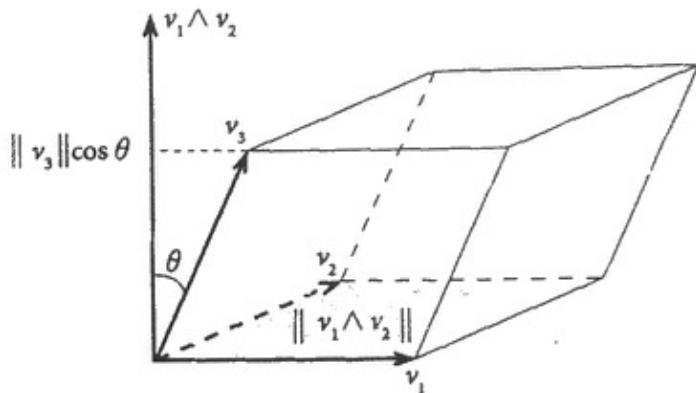
Se  $\vec{v}_1, \vec{v}_2, \vec{v}_3 \in \mathbb{R}^3$  sono tre vettori, il loro *prodotto misto* è

$$[\vec{v}_1, \vec{v}_2, \vec{v}_3] = \langle \vec{v}_1 \wedge \vec{v}_2, \vec{v}_3 \rangle \in \mathbb{R}.$$

Se  $\theta$  è l'angolo fra  $\vec{v}_1 \wedge \vec{v}_2$  e  $\vec{v}_3$ , si ha

$$[\vec{v}_1, \vec{v}_2, \vec{v}_3] = \|\vec{v}_1 \wedge \vec{v}_2\| \|\vec{v}_3\| \cos \theta.$$

Ora, il parallelepipedo inclinato generato da  $v_1$ ,  $v_2$  e  $v_3$  (vedi la Figura 12.8) ha area di base esattamente  $\|v_1 \wedge v_2\|$  e altezza  $\|v_3\| \cos \theta$ ; quindi il prodotto misto è il volume di questo parallelepipedo. Ricordando la (12.16) abbiamo quindi fatto vedere come il determinante dia il volume (orientato) del parallelepipedo generato dai tre vettori; confronta con le Osservazioni 9.1 e 9.8.



**Figura 12.8** - Prodotto misto e volume del parallelepipedo.

# Complementi

# 12

## 12C.1 Isometrie

L'obiettivo di questo paragrafo è scoprire quali sono le affinità (vedi i complementi al Capitolo 10) che conservano le distanze; nel prossimo paragrafo ne daremo una descrizione particolareggiata. Cominciamo con due definizioni.

### Definizione 12C.1

Uno spazio (piano, retta) euclideo è uno spazio (piano, retta) affine su cui sia stata fissata un'unità di misura.

### Definizione 12C.2

Sia  $\varphi$  un'affinità fra due spazi (piani, rette) euclidei. Diremo che  $\varphi$  è un'isometria (non necessariamente lineare) se conserva le distanze, cioè se

$$d(\varphi(P_1), \varphi(P_2)) = d(P_1, P_2)$$

per ogni coppia di punti  $P_1, P_2$  dello spazio (piano, retta) affine di partenza.

Consideriamo un'affinità  $\varphi: \Sigma_1 \rightarrow \Sigma_2$  fra due spazi euclidei; il caso di affinità fra piani o rette sarà analogo. Fissato un sistema di riferimento cartesiano  $\mathcal{R}_1$  in  $\Sigma_1$  e uno  $\mathcal{R}_2$  in  $\Sigma_2$ , sappiamo che  $\varphi$  si esprime in coordinate come  $y = Bx + c$  per qualche  $B \in GL_3(\mathbb{R})$  e  $c \in \mathbb{R}^3$ , dove  $x \in \mathbb{R}^3$  sono le coordinate di un punto  $P \in \Sigma_1$  rispetto a  $\mathcal{R}_1$  e  $y \in \mathbb{R}^3$  sono le coordinate di  $\varphi(P) \in \Sigma_2$  rispetto a  $\mathcal{R}_2$ . Dunque  $\varphi$  è un'isometria se e solo se per ogni coppia di punti  $P, P' \in \Sigma_1$  si ha

$$\|x' - x\| = d(P, P') = d(\varphi(P), \varphi(P')) = \|(Bx' + c) - (Bx + c)\| = \|B(x' - x)\|.$$

Dunque la matrice  $B$  conserva la norma; la Proposizione 11.17 e il Corollario 11.19 implicano che  $B$  è ortogonale. Viceversa, se  $B$  è una matrice ortogonale, è chiaro che un'affinità data in coordinate da  $y = Bx + c$  conserva le distanze (per non parlare degli angoli); quindi abbiamo dimostrato la seguente proposizione.

### Proposizione 12C.1

Sia  $\varphi$  un'affinità fra due spazi (piani, rette) euclidei. Allora  $\varphi$  è un'isometria se e solo se rispetto a qualunque sistema di riferimento cartesiano scelto nei due spazi (piani, rette) rappresentata in coordinate da  $y = Bx + c$ , dove  $B$  è una matrice ortogonale. In particolare, un'isometria conserva anche gli angoli.

In realtà possiamo dimostrare un risultato molto più forte: una qualunque applicazione fra spazi (rette, piani) euclidei che conserva le distanze è necessariamente un'affinità – e quindi un'isometria nel senso della Definizione 12C.2.

Considereremo solo il caso degli spazi; per rette e piani la dimostrazione è analoga.

### **Teorema 12C.2**

Sia  $\varphi: \Sigma_1 \rightarrow \Sigma_2$  un'applicazione fra spazi euclidei tale che

$$\forall P, Q \in \Sigma_1 \quad d_2(\varphi(P), \varphi(Q)) = d_1(P, Q),$$

dove  $d_j$  è la distanza in  $\Sigma_j$  per  $j = 1, 2$ . Allora  $\varphi$  è necessariamente un'affinità – e quindi un'isometria.

### **Dimostrazione**

Fissiamo dei sistemi di riferimento cartesiani in  $\Sigma_1$  e  $\Sigma_2$  in modo da identificare sia  $\Sigma_1$  sia  $\Sigma_2$  con  $\mathbb{R}^3$ . Allora  $\varphi$  diventa un'applicazione  $\tilde{\varphi}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  che conserva la distanza indotta dal prodotto scalare canonico; dobbiamo dimostrare che  $\tilde{\varphi}(x) = Bx + c$  per qualche  $c \in \mathbb{R}^3$  e  $B \in GL_3(\mathbb{R})$ .

Prima di tutto, definiamo  $\tilde{\varphi}_1: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  ponendo

$$\tilde{\varphi}_1(x) = \tilde{\varphi}(x) - \tilde{\varphi}(O);$$

chiaramente  $\tilde{\varphi}_1$  conserva ancora le distanze, e inoltre  $\tilde{\varphi}_1(O) = O$ . Vogliamo dimostrare che  $\tilde{\varphi}_1$  è lineare.

Siccome  $\tilde{\varphi}_1$  conserva le distanze abbiamo

$$\forall v, w \in \mathbb{R}^3 \quad \|\tilde{\varphi}_1(v) - \tilde{\varphi}_1(w)\|^2 = \|v - w\|^2;$$

quindi

$$\|\tilde{\varphi}_1(v)\|^2 - 2\langle \tilde{\varphi}_1(v), \tilde{\varphi}_1(w) \rangle + \|\tilde{\varphi}_1(w)\|^2 = \|v\|^2 - 2\langle v, w \rangle + \|w\|^2.$$

In particolare prendendo  $w = O$  otteniamo  $\|\tilde{\varphi}_1(v)\| = \|v\|$ . Ne segue che

$$\forall v, w \in \mathbb{R}^3 \quad \langle \tilde{\varphi}_1(v), \tilde{\varphi}_1(w) \rangle = \langle v, w \rangle. \tag{12C.1}$$

Sia  $\{e_1, e_2, e_3\}$  la base canonica di  $\mathbb{R}^3$ . La (12C.1) dice che  $\{\tilde{\varphi}_1(e_1), \tilde{\varphi}_1(e_2), \tilde{\varphi}_1(e_3)\}$  è una base ortonormale di  $\mathbb{R}^3$ . Allora se  $v = (x_1, x_2, x_3) \in \mathbb{R}^3$  la Proposizione 11.4 dà

$$\tilde{\varphi}_1(v) = \sum_{j=1}^3 \langle \tilde{\varphi}_1(v), \tilde{\varphi}_1(e_j) \rangle \tilde{\varphi}_1(e_j) = \sum_{j=1}^3 \langle v, e_j \rangle \tilde{\varphi}_1(e_j) = \sum_{j=1}^3 x_j \tilde{\varphi}_1(e_j),$$

per cui  $\tilde{\varphi}_1$  è lineare (vedi la dimostrazione della Proposizione 5.2). Essendo necessariamente iniettiva (perché?) esiste allora una matrice invertibile  $B \in GL_3(\mathbb{R})$  tale che  $\tilde{\varphi}_1(x) = Bx$ , e quindi  $\tilde{\varphi}(x) = Bx + c$  dove  $c = \tilde{\varphi}(O)$ . □

### **Osservazione 12C.1**

La stessa dimostrazione può venire adattata per far vedere che un'applicazione  $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$  che conserva la distanza indotta dal prodotto scalare canonico è necessariamente della forma  $\varphi(x) = Bx + c$  per qualche  $B \in O(n)$  e  $c \in \mathbb{R}^n$ .

## 12C.2 Isometrie lineari dello spazio

Concentriamo adesso la nostra attenzione sulle isometrie che fissano l'origine degli assi; per quanto visto nel paragrafo precedente, sono tutte applicazioni lineari, e ogni altra isometria si ottiene componendo una isometria lineare con una traslazione.

Nell'Esempio 11.17 abbiamo studiato le isometrie lineari del piano, dimostrando che sono o simmetrie rispetto a una retta per l'origine (quando hanno determinante uguale a  $-1$ ), oppure sono rotazioni attorno all'origine (quando hanno determinante uguale a  $1$ ). Inoltre, abbiamo osservato che ogni rotazione si ottiene come composizione di due simmetrie. Scopo di questo paragrafo è dimostrare risultati simili per le isometrie dello spazio. Cominciamo col definire gli analoghi spaziali di simmetrie e rotazioni.

### Definizione 12C.3

La simmetria (o riflessione) rispetto a un piano  $\pi \subset \mathbb{R}^3$  per l'origine è l'unica (perché?) applicazione lineare  $S_\pi: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  tale che  $S(v) = v$  per ogni  $v \in \pi$  e  $S(w) = -w$  per ogni  $w \in \mathbb{R}^3$  ortogonale a  $\pi$ .

### Definizione 12C.4

La rotazione di angolo  $\theta \in (0, 2\pi)$  attorno a una retta  $r$  per l'origine è l'unica (perché?) applicazione lineare  $R_{r,\theta}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  tale che  $R_{r,\theta}(v) = v$  per ogni  $v \in r$  e che agisce come una rotazione di angolo  $\theta$  sul piano  $r^\perp$  ortogonale a  $r$  passante per l'origine.

### Definizione 12C.5

L'antirotazione di angolo  $\theta \in (0, 2\pi)$  attorno a una retta  $r$  per l'origine è l'unica (perché?) applicazione lineare  $A_{r,\theta}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  tale che  $A_{r,\theta}(v) = -v$  per ogni  $v \in r$  e che agisce come una rotazione di angolo  $\theta$  sul piano  $r^\perp$  ortogonale a  $r$  passante per l'origine.

### Osservazione 12C.2

Le rotazioni attorno a una retta  $r$  sono isometrie lineari di determinante  $+1$ . Infatti, sia  $w_1 \in r$  un vettore di lunghezza unitaria, e  $\{w_2, w_3\}$  una base ortonormale di  $\pi = r^\perp$ . Allora  $\{w_1, w_2, w_3\}$  è una base ortonormale di  $\mathbb{R}^3$  rispetto a cui la rotazione è rappresentata da una matrice della forma

$$\begin{vmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{vmatrix},$$

che è chiaramente ortogonale e con determinante  $+1$ . In maniera analoga è facile verificare (Esercizio 12C.1) che le simmetrie e le antirotazioni sono isometrie lineari di determinante  $-1$ . Inoltre, ogni simmetria  $S$  è una *involtuzione*, cioè  $S^2 = \text{id}$ .

### Osservazione 12C.3

La simmetria rispetto all'origine  $S_O(v) = -v$  è una antirotazione di angolo  $\pi$  attorno a una qualsiasi retta per l'origine.

Un modo per classificare le isometrie lineari consiste nel guardarne i punti fissi.

### Definizione 12C.6

Se  $B: V \rightarrow V$  è un'applicazione lineare, l'insieme

$$\text{Fix}(B) = \{v \in V \mid B(v) = v\}$$

dei punti fissi di  $B$  è un sottospazio vettoriale di  $V$  (verificare, prego).

**Esempio 12C.1**

L'insieme dei punti fissi di una simmetria  $S_\pi$  è esattamente il piano  $\pi$ , e quindi ha dimensione 2. L'insieme dei punti fissi di una rotazione  $R_{r,\theta}$  (con  $\theta \neq 0$ ) è esattamente la retta  $r$  per cui ha dimensione 1. Infine, l'unico punto fisso di una antirotazione  $A_{r,\theta}$  (con  $\theta \neq 0$ ) è l'origine.

Il principale risultato di questo paragrafo è che in questo modo abbiamo descritto tutte le possibili isometrie lineari dello spazio.

**Teorema 12C.3**

Ogni isometria lineare di  $\mathbb{R}^3$  si può ottenere come composizione di al più 3 simmetrie rispetto a opportuni piani per l'origine. Più precisamente:

- (1) le simmetrie sono tutte e sole le isometrie lineari il cui insieme dei punti fissi abbia dimensione 2;
- (2) le rotazioni sono tutte e sole le isometrie lineari che si ottengono come composizione di due simmetrie, e sono tutte e sole le isometrie lineari il cui insieme di punti fissi abbia dimensione 1;
- (3) le antirotazioni si ottengono come composizione di tre simmetrie, e sono tutte e sole le isometrie lineari il cui insieme di punti fissi abbia dimensione 0.

**Dimostrazione**

Sia  $B \in O(3)$  un'isometria lineare (ovvero una matrice ortogonale) diversa dall'identità. Se  $\dim \text{Fix}(B) = 2$ , cioè  $\pi = \text{Fix}(B)$  è un piano, prendiamo un versore  $v_1$  ortogonale a  $\pi$ . La (12C.1) ci dice che  $B(v_1)$  dev'essere un versore ancora ortogonale a  $\pi$ , per cui necessariamente  $B(v_1) = \pm v_1$ . Se fosse  $B(v_1) = v_1$  allora  $\text{Fix}(B)$  avrebbe dimensione 3, e  $B$  sarebbe l'identità, impossibile. Quindi  $B(v_1) = -v_1$ , e  $B$  è la simmetria rispetto al piano  $\pi = v_1^\perp$ .

Supponiamo ora che  $\dim \text{Fix}(B) = 1$ , e prendiamo un versore  $v_1 \in r = \text{Fix}(B)$ . La (12C.1) ci dice che se  $w$  è ortogonale a  $v_1$  allora  $Bw$  è ortogonale a  $Bv_1 = v_1$ ; quindi il piano  $\pi = v_1^\perp$  viene mandato in sé da  $B$ , e, inoltre, la restrizione di  $B$  a  $\pi$  è ancora una isometria lineare. Siccome  $B$  non può avere punti fissi in  $\pi$  (a parte l'origine), l'Esempio 11.17 ci dice che  $B|_\pi$  è una rotazione, e quindi  $B$  è una rotazione attorno alla retta  $r$ . Inoltre, sempre nell'Esempio 11.17 abbiamo visto che una rotazione nel piano è composizione di due simmetrie; estendendo queste simmetrie a  $\mathbb{R}^3$  richiedendo che siano l'identità su  $r$  otteniamo  $B$  come composizione di due simmetrie dello spazio.

Supponiamo infine che  $\dim \text{Fix}(B) = 0$ . Dato un qualsiasi versore  $v_0 \in \mathbb{R}^3$ , poniamo  $w = B(v_0) - v_0 \neq 0$ . Allora  $B(v_0) + v_0$  è ortogonale a  $w$ , in quanto

$$\langle B(v_0) + v_0, w \rangle = \langle B(v_0) + v_0, B(v_0) - v_0 \rangle = \|B(v_0)\|^2 - \|v_0\|^2 = 0.$$

Questo implica che  $S_\pi(B(v_0) + v_0) = B(v_0) + v_0$  e  $S_\pi(B(v_0) - v_0) = -B(v_0) + v_0$ , dove  $\pi$  è il piano per l'origine ortogonale a  $w$ ; sommando otteniamo  $S_\pi \circ B(v_0) = v_0$ . Dunque  $\dim \text{Fix}(S_w \circ B) \geq 1$ , per cui  $S_\pi \circ B$  è una rotazione o una simmetria.

Se  $S_\pi \circ B$  fosse una simmetria rispetto a un piano  $\pi_1$ , potremmo trovare un vettore non nullo  $v_2 \in \pi \cap \pi_1$ ; ma allora  $B(v_2) = S_\pi(S_\pi \circ B(v_2)) = S_\pi(v_2) = v_2$  (dove abbiamo sfruttato il fatto che  $S_\pi$  è un'involuzione), contro l'ipotesi che  $B$  non avesse punti fissi non nulli. Quindi  $S_\pi \circ B$  è una rotazione; in particolare, si ottiene come composizione di due simmetrie, e (componendo ancora con  $S_\pi$ ) la nostra  $B$  si ottiene come composizione di tre simmetrie.

Siccome il determinante della composizione è uguale al prodotto dei determinanti, e le simmetrie hanno determinante  $-1$ , il determinante di  $B$  è  $-1$ . Ma allora il Teorema 9.9 e il Corollario 9.7 implicano che

$$\det(B + I_3) = \det(B + B^T B) = \det(I_3 + B^T) \det B = -\det(B + I_3);$$

Quindi  $\det(B + I_3) = 0$ , cioè  $B + I_3$  è una matrice singolare. Quindi esiste un versore  $v_1$  tale che  $B(v_1) = -v_1$ , e  $B$  manda il piano  $\pi_1$  per l'origine ortogonale a  $v_1$  in se stesso. Non potendo avere punti fissi diversi dall'origine, l'azione di  $B$  su questo piano dev'essere una rotazione, e quindi  $B$  è un'antirotazione.

Infine, la composizione di due simmetrie ha determinante  $+1$ , per cui non può essere né una simmetria né un'antirotazione, e quindi è una rotazione.  $\square$

Un'importante conseguenza di questo teorema è la seguente.

#### **Corollario 12C.4**

*Un'isometria lineare ha determinante  $+1$  se e solo se è una rotazione attorno a una retta per l'origine. In particolare, la composizione di due rotazioni attorno a rette per l'origine è ancora una rotazione.*

#### **Dimostrazione**

La prima affermazione segue dal teorema precedente e dall'Osservazione 12C.2. Siccome la composizione di due isometrie lineari di determinante  $+1$  è ancora un'isometria lineare di determinante  $+1$ , abbiamo finito.  $\square$

# Esercizi 12

**[12.1]** Presa una retta  $r \subset \mathcal{E}^3$  passante per l'origine, mostra che la Definizione 12.6 di proiezione ortogonale su  $r$  coincide con quella data nel Paragrafo 11.4. Se  $r$  non passa per l'origine, dimostra la stessa cosa lavorando in  $\mathcal{V}_Q^3$  con  $Q \in r$ .

**[12.2]** Preso un piano  $\pi$  passante per l'origine, dimostra che la Definizione 12.10 di proiezione ortogonale su  $\pi$  coincide con quella data nel Paragrafo 11.4.

**[12.3]** Nel ricavare la formula (12.6) dell'angolo fra una retta e un piano, abbiamo implicitamente assunto che le tre rette  $r$ ,  $r'$  e  $s$  siano complanari (altrimenti l'angolo fra  $s$  ed  $r$  non sarebbe stato  $\pi/2 - \theta$ ). Dimostra che sono davvero complanari.

**[12.4]** Calcola il coseno dell'angolo fra le rette  $r_0$  e  $r_1$  (orientate a tuo piacimento) di equazioni parametriche rispettivamente  $r_0 : x = 2t - 9$ ,  $y = 3t + 1$ ,  $z = t - 6$  e  $r_1 : x = 4t - 1$ ,  $y = t - 1$ ,  $z = 5t - 7$ .

**[12.5]** Calcola il coseno dell'angolo fra le rette (orientate a tuo piacimento)  $r_0$  di equazioni parametriche  $r_0 : x = t + 2$ ,  $y = 2t - 3$ ,  $z = -2t + 5$  e  $r_1$  di equazioni cartesiane  $r_1 : \begin{cases} 3x - y - 2z = 4, \\ x - 3y + 2z = 5. \end{cases}$

**[12.6]** Calcola il coseno dell'angolo fra le rette  $r_0$  e  $r_1$  (orientate a tuo piacimento) di equazioni cartesiane

$$r_0 : \begin{cases} 2x + y - z = 0, \\ x - y + z = 1; \end{cases} \quad r_1 : \begin{cases} x - y = 1, \\ y - z = 0. \end{cases}$$

**[12.7]** Trova equazioni parametriche e cartesiane di tutte le rette e del piano passanti per  $P_0 = (0, 1, 0)$  e ortogonali alla retta  $r_1$  dell'esercizio precedente.

**[12.8]** Scrivi equazioni parametriche e cartesiane del piano passante per il punto  $(2, -2, 1)$  e perpendicolare alla retta di equazioni cartesiane  $\begin{cases} 2x + y + z = 1, \\ 3x + y - 2z = 7. \end{cases}$

**[12.9]** Determina equazioni parametriche e cartesiane della retta passante per il punto  $(3, 6, -11)$  e perpendicolare al piano di equazione  $x - 5y + 2z = \sqrt{3}$ .

**[12.10]** Determina equazioni parametriche e cartesiane della retta passante per il punto  $(0, -4, 2)$  e perpendicolare al piano passante per l'origine e avente vettori di giacitura  $w_1 = \begin{vmatrix} 1 \\ 3 \\ -4 \end{vmatrix}$  e  $w_2 = \begin{vmatrix} 8 \\ 1 \\ -3 \end{vmatrix}$ .

**12.11** Calcola il coseno dell'angolo fra i piani  $\pi_0$  e  $\pi_1$  (orientati a tuo piacimento) di equazioni cartesiane rispettivamente  $x + y + z = 1$  e  $x - y - z = 2$ .

**12.12** Calcola il coseno dell'angolo fra i due piani (orientati a tuo piacimento)  $\pi_0$  di equazione  $2x - y + 4z = 5$  e  $\pi_1$  passante per l'origine e avente vettori di giacitura

$$u_1 = \begin{vmatrix} 0 \\ -2 \\ 1 \end{vmatrix} \text{ e } u_2 = \begin{vmatrix} 3 \\ 4 \\ -5 \end{vmatrix}.$$

**12.13** Calcola il coseno dell'angolo fra i piani (orientati a tuo piacimento)  $\pi_0$  e  $\pi_1$  di equazioni parametriche rispettivamente

$$\pi_0 : \begin{cases} x = t - s + 2, \\ y = t + s + 3, \\ z = 2t - 4s - 8, \end{cases} \quad \text{e} \quad \pi_1 : \begin{cases} x = 2t + 3s - 1, \\ y = 3t - 9s - 5, \\ z = 7t + 2s + 3. \end{cases}$$

**12.14** Calcola il seno dell'angolo fra la retta  $r_0$  dell'Esercizio 12.6 e il piano  $\pi_1$  dell'Esercizio 12.11.

**12.15** Date due rette sghembe, esiste sempre un piano parallelo a entrambe? E un piano ortogonale a entrambe?

**12.16** Trova equazioni cartesiane e parametriche della retta  $r$  passante per il punto  $P_0 = (1, 1, 1)$  e ortogonale e incidente alla retta  $s$  di equazioni cartesiane  $x - y - z = 1$  e  $2x - y + z = 1$ .

**12.17** Sia  $P_0 = (1, 0, 1)$ . Calcola la distanza di  $P_0$  dalle rette dell'Esercizio 12.6 e dai piani dell'Esercizio 12.11.

**12.18** Calcola la distanza del punto  $P_0 = (3, 1, -2)$  dalla retta  $r$  di equazioni cartesiane

$$r : \begin{cases} x + z = 4, \\ x - 4y + z = 5. \end{cases}$$

**12.19** La distanza del punto  $P_0 = (5, 0, 2)$  dal piano di equazione  $3x - 2y + 7z = 9$  è maggiore della distanza tra  $P_0$  e l'origine?

**12.20** Calcola la distanza fra le due rette dell'Esercizio 12.6.

**12.21** Calcola la distanza fra la retta  $r_0$  di equazioni parametriche

$$r_0 : x = t - 2, y = 2t - 1, z = 2t - 3$$

e la retta  $r_1$  di equazioni cartesiane  $r_1 : \begin{cases} x - y - z = 2, \\ 2x - y + 2z = 5. \end{cases}$

**12.22** Trova equazioni cartesiane e parametriche dei piani che distano  $2/\sqrt{6}$  dal punto  $P_0 = (1, 1, 1)$  e contenenti la retta  $r$  di equazioni cartesiane  $x - y = 1$  e  $y + z = 1$ .

**12.23** Siano  $r_0$  e  $r_1$  due rette non parallele. Dimostra che esiste un'unica retta  $s$  ortogonale e incidente sia a  $r_0$  che a  $r_1$ . Dimostra che  $d(r_0, r_1)$  è uguale a  $d(H_0, H_1)$ , dove  $H_j$  è il punto d'intersezione fra  $s$  ed  $r_j$ , per  $j = 0, 1$  (vedi la Figura 12.6).

**12.24** Date le rette  $r_1$  ed  $r_2$  di equazioni rispettivamente

$$r_1 : \begin{cases} x = 2 + s, \\ y = -2s, \\ z = -1 + 3s, \end{cases} \quad r_2 : \begin{cases} x + y + 2z = 0, \\ x + y + z = 0, \end{cases}$$

calcola la distanza fra  $r_1$  e la retta  $s$  passante per  $P_0 = (2, 1, 3)$ , ortogonale a  $r_1$  incidente a  $r_2$ .

**12.25** Verifica che la retta  $r_0$  di equazioni parametriche

$$r_0 : x = -2t, y = -1 - t, z = 1 - t$$

e la retta  $r_1$  di equazioni cartesiane

$$r_1 : \begin{cases} y - z = -1, \\ 2x - y - z = 0, \end{cases}$$

sono sghembe. Scrivi equazioni parametriche e cartesiane della retta  $\rho$  perpendicolare a entrambe e incidente sia a  $r_0$  che a  $r_1$ . Trova  $H_0$  e  $H_1$ , punti di intersezione di  $\rho$  rispettivamente con  $r_0$  e  $r_1$ , e verifica che  $d(r_0, r_1) = d(H_0, H_1)$ .

**12.26** Trova il prodotto vettore di  $v = (1, 2, 1)$  e  $w = (a, 2a, 1)$  al variare di  $a \in \mathbb{R}$

**12.27** Fissato un vettore  $v_0 \in \mathbb{R}^3$ , siano  $S, T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  le applicazioni definite da  $S(v) = v \wedge v_0$  e  $T(v) = v \wedge v_0 + \langle v, v_0 \rangle v_0$ , dove  $\langle \cdot, \cdot \rangle$  è il prodotto scalare canonico. Dimostra che  $S$  e  $T$  sono lineari, e determina le loro immagini e i loro nuclei.

### Esercizi relativi ai Complementi del Capitolo 12

**12C.1** Dimostra che le simmetrie e le antirotazioni sono isometrie lineari di determinante  $-1$ , e che le simmetrie sono delle involuzioni.

**12C.2** Sia  $v_1 \in \mathbb{R}^3$  un vettore unitario ortogonale a un piano  $\pi \subset \mathbb{R}^3$  per l'origine. Dimostra che la simmetria  $S_\pi$  rispetto a  $\pi$  è data da  $S_\pi(v) = v - 2\langle v, v_1 \rangle v_1$ .

**12C.3** Dimostra che la traccia (vedi gli Esercizi 5.10 e 7.25) di una rotazione di angolo  $\theta$  è  $2\cos\theta + 1$ , mentre la traccia di un'antirotazione di angolo  $\theta$  è  $2\cos\theta - 1$ .

**12C.4** Definisci il concetto di simmetria rispetto a un iperpiano per l'origine in  $\mathbb{R}^n$ , e dimostra che ogni isometria lineare di  $\mathbb{R}^n$  si può ottenere come composizione di al più  $n$  simmetrie rispetto a opportuni iperpiani. (Suggerimento: Procedi per induzione su  $n$ .)

# Autovalori e autovettori

13

## Sommario

- 13.1 Definizioni ed esempi
- 13.2 Il polinomio caratteristico
- 13.3 Molteplicità

## Esercizi

Questo capitolo è dedicato a uno degli argomenti più importanti dell'Algebra Lineare, sia dal punto di vista teorico sia per le applicazioni: autovettori e autovalori. Un autovettore di un endomorfismo  $T$  di uno spazio vettoriale  $V$  è un vettore  $v \in V$  non nullo tale che  $T(v)$  sia un multiplo di  $v$ ; il coefficiente di proporzionalità si chiama autovалore. Nelle applicazioni, autovettori e autovalori compaiono come assi preferenziali di rotazione, frequenze di risonanza, direzioni di maggior sforzo e così via; nella matematica pura sono uno dei mezzi più utili per lo studio degli endomorfismi degli spazi vettoriali. Dopo aver esaminato molteplici esempi, definiremo il polinomio caratteristico di un endomorfismo, che è lo strumento principale per la determinazione di autovalori e autovettori. Poi studieremo il problema (che ritroveremo anche nel prossimo capitolo) dell'esistenza di una base di autovettori per un dato endomorfismo, dimostrando un criterio necessario e sufficiente in termini della molteplicità algebrica e geometrica degli autovalori.

## 13.1 Definizioni ed esempi

Per giustificare l'importanza dell'argomento di questo capitolo vediamo un modello di situazione fisica (e ingegneristica) in cui compare naturalmente.

### Esempio 13.1

Consideriamo un corpo statico  $\Sigma$  sottoposto a delle forze; può essere un ponte, una macchina, una colonna o qualsiasi altra cosa. Vogliamo capire se c'è il rischio che il corpo, a causa dello stress a cui è sottoposto, si rompa.

Prendiamo un punto  $P_0 \in \Sigma$ ; cominciamo col rappresentare le forze a cui è soggetto il corpo in  $P_0$  tenendo presente che, siccome abbiamo a che fare con un corpo solido, ciò che accade in  $P_0$  è influenzato da quanto succede nelle immediate vicinanze. Un possibile modo di procedere consiste nel considerare le forze che agiscono sulle facce di un cubetto infinitesimo centrato in  $P_0$ . Chiamiamo  $T(e_j)$  la somma di tutte le forze applicate sulla faccia ortogonale alla direzione  $e_j$ , per  $j = 1, 2, 3$ . Estendendo per linearità (cioè applicando la Proposizione 5.2) otteniamo un'applicazione lineare  $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  che descrive che tipo di forze agiscono sul corpo in  $P_0$ : se  $v \neq O$  il vettore  $T(v)$  è la somma delle forze che agiscono su una superficie piana infinitesima passante per  $P_0$  e ortogonale a  $v$ . Possiamo quindi definire lo sforzo  $F$  a cui il corpo è soggetto in  $P_0$  nella direzione  $v \neq O$  ponendo

$$F(v) = \frac{\langle T(v), v \rangle}{\langle v, v \rangle}.$$

Infatti, la componente di  $T(v)$  parallela alla superficie ortogonale alla direzione  $v$  non crea alcuno stress su questa superficie (si limita a traslarla su se stessa, non la tira verso l'esterno); nello sforzo dunque conta solo la componente di  $T(v)$  ortogonale a questa superficie, cioè la proiezione ortogonale di  $T(v)$  lungo  $v$ , proiezione il cui modulo è proprio  $|F(v)|$ . La direzione in cui il nostro materiale si potrebbe rompere è ovviamente quella di massimo sforzo; vogliamo vedere se possiamo trovarla usando l'endomorfismo  $T$ . Prima di tutto, notiamo che  $F(\lambda v) = F(v)$  per ogni  $\lambda \in \mathbb{R}^*$  (esercizio). Come conseguenza, questa direzione di massimo sforzo esiste: infatti, per trovarla possiamo limitarci a considerare i vettori  $v$  di lunghezza uno, che è un insieme chiuso e limitato (un insieme compatto). Ma  $F$  è una funzione continua (è un rapporto di polinomi), e quindi ammette massimo su un insieme chiuso e limitato (è una generalizzazione del risultato che avrai certamente incontrato nel corso di Analisi Matematica a proposito delle funzioni continue definite su un intervallo chiuso e limitato). Sia allora  $v_0$  questa direzione di massimo sforzo, con  $\|v_0\| = 1$ . Presa un'altra direzione  $v \neq O$ , consideriamo la funzione  $f(t) = F(v_0 + tv)$ . Chiaramente, la  $f$  ha un massimo in 0; quindi  $f'(0) = 0$ . Calcoliamo questa derivata. Prima di tutto,

$$f(t) = \frac{\langle T(v_0), v_0 \rangle + t[\langle T(v), v_0 \rangle + \langle T(v_0), v \rangle] + t^2 \langle T(v), v \rangle}{\langle v_0, v_0 \rangle + 2t \langle v_0, v \rangle + t^2 \langle v, v \rangle}.$$

Ora si può dimostrare che, siccome il nostro corpo è fermo, l'endomorfismo  $T$  è simmetrico<sup>(1)</sup>; dunque abbiamo

$$f(t) = \frac{\langle T(v_0), v_0 \rangle + 2t \langle T(v_0), v \rangle + t^2 \langle T(v), v \rangle}{\langle v_0, v_0 \rangle + 2t \langle v_0, v \rangle + t^2 \langle v, v \rangle}.$$

<sup>(1)</sup> L'idea è che se non lo fosse allora il corpo dovrebbe cominciare a ruotare intorno a  $P_0$ .

Derivando e ponendo  $t = 0$  otteniamo

$$0 = f'(0) = 2 \frac{\langle T(v_0), v \rangle \langle v_0, v_0 \rangle - \langle v_0, v \rangle \langle T(v_0), v_0 \rangle}{\langle v_0, v_0 \rangle^2}.$$

Dunque, ricordando che  $\langle v_0, v_0 \rangle = 1$  e indicando con  $\lambda = \langle T(v_0), v_0 \rangle = F(v_0)$  lo sforzo in  $v_0$ , troviamo che

$$\langle T(v_0), v \rangle = \langle \lambda v_0, v \rangle.$$

Siccome questo dev'essere vero per ogni  $v \in \mathbb{R}^3$  ne segue che

$$T(v_0) = \lambda v_0. \quad (13.1)$$

Riassumendo, abbiamo trovato che la direzione di massimo sforzo in  $P_0$  soddisfa la (13.1), dove  $\lambda$  è proprio lo sforzo che il corpo subisce nella direzione  $v_0$ . Dunque per scoprire se il nostro materiale sottoposto alle sollecitazioni riassunte dall'endomorfismo  $T$  si romperà dobbiamo trovare tutti i vettori che soddisfano (13.1), e vedere se la resistenza del materiale in  $P_0$  è maggiore del valore dello sforzo lungo quei vettori. Se lo è, non ci sono problemi; se non lo è, meglio cambiare materiale.  $\square$

A questo punto le seguenti definizioni divengono piuttosto naturali.

### Definizione 13.1

Sia  $T: V \rightarrow V$  un endomorfismo di uno spazio vettoriale  $V$ . Un vettore  $v_0 \neq O$  di  $V$  è un *autovettore* di  $T$  relativo all'*autovalore*  $\lambda$  se si ha

$$T(v_0) = \lambda v_0. \quad (13.2)$$

L'insieme degli autovalori di  $T$  – cioè degli scalari  $\lambda$  per cui esiste un vettore non nullo  $v_0$  per cui la formula (13.2) è verificata – è lo *spettro* di  $T$ , ed è indicato con  $\text{sp}(T)$ . Se  $\lambda \in \text{sp}(T)$ , l'insieme

$$V_\lambda = \{v \in V \mid T(v) = \lambda v\}$$

è detto *autospazio* di  $T$  relativo all'autovalore  $\lambda$ , ed è un sottospazio vettoriale di  $V$  (Esercizio 13.1). Ovviamente,  $\lambda$  è un autovalore se e solo se  $V_\lambda \neq \{O\}$ .

### Osservazione 13.1

Dalla definizione è chiaro che  $V_0 = \text{Ker } T$ , cioè lo zero è un autovalore di  $T$  se e solo se  $T$  è singolare. Più in generale,  $V_\lambda = \text{Ker}(T - \lambda \text{id}_V)$ , per cui  $\lambda$  è un autovalore se e solo se l'endomorfismo  $T - \lambda \text{id}_V$  è singolare.

Con questa terminologia l'Esempio 13.1 può venire riassunto dicendo che la direzione di massimo sforzo è esattamente un autovettore dell'endomorfismo  $T$ , e lo sforzo è il relativo autovalore. Dunque dovrebbe essere chiara l'importanza del seguente problema: *dato un endomorfismo  $T: V \rightarrow V$ , trovare i suoi autovettori e relativi autovalori*.

Prima di affrontare il problema in generale vediamo alcuni esempi.

### Esempio 13.2

Sia  $V = \mathbb{R}^3$  e  $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  la rotazione di angolo  $\pi$  attorno all'asse  $z$ . Allora  $T(e_1) = -e_1$ ,  $T(e_2) = -e_2$  e  $T(e_3) = e_3$ ; quindi  $T$  è rappresentato dalla matrice

$$\begin{vmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

In particolare,  $T$  ha almeno due autovalori, 1 e  $-1$ , e i relativi autospazi contengono per lo meno i vettori della base canonica; per l'esattezza abbiamo  $V_1 \supseteq \text{Span}(e_3)$  e  $V_{-1} \supseteq \text{Span}(e_1, e_2)$ . Per scoprire se ve ne sono altri o abbiamo finito, studiamo il sistema  $T(x) = \lambda x$ , dipendente dal parametro  $\lambda$ ; vogliamo vedere se esistono dei valori del parametro per cui questo sistema ammette una soluzione non nulla. In coordinate dobbiamo studiare il sistema

$$\begin{cases} -x_1 = \lambda x_1, \\ -x_2 = \lambda x_2, \\ x_3 = \lambda x_3, \end{cases} \quad \text{cioè} \quad \begin{cases} (\lambda + 1)x_1 = 0, \\ (\lambda + 1)x_2 = 0, \\ (\lambda - 1)x_3 = 0. \end{cases} \quad (13.3)$$

Quindi se  $\lambda \neq \pm 1$  l'unica soluzione possibile è  $x = O$ , per cui un tale  $\lambda$  non è un autovalore. Di conseguenza,  $\text{sp}(T) = \{1, -1\}$ ; inoltre risolvendo (13.3) per  $\lambda = 1$  o  $\lambda = -1$  troviamo esattamente che  $V_1 = \text{Span}(e_3)$  e  $V_{-1} = \text{Span}(e_1, e_2)$ , in quanto  $V_\lambda$  è proprio lo spazio delle soluzioni del sistema  $T(x) = \lambda x$ . Abbiamo dunque recuperato tutti gli autovettori e autovalori di  $T$ ; notiamo in particolare che la base canonica è una base di autovettori per il nostro endomorfismo  $T$ .  $\square$

### Esempio 13.3

Sia ancora  $V = \mathbb{R}^3$ , ma stavolta  $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  è la rotazione di angolo  $\pi/2$  attorno all'asse  $z$ . Questa volta  $T(e_1) = e_2$ ,  $T(e_2) = -e_1$  e  $T(e_3) = e_3$ , per cui  $T$  è rappresentato dalla matrice

$$\begin{vmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix}. \quad (13.4)$$

Stavolta vi è un solo autovalore evidente,  $\lambda = 1$  con autospazio  $V_1 \supseteq \text{Span}(e_3)$ . Vediamo se ce ne sono altri, studiando il sistema  $T(x) = \lambda x$ . In coordinate abbiamo

$$\begin{cases} -x_2 = \lambda x_1, \\ x_1 = \lambda x_2, \\ x_3 = \lambda x_3, \end{cases} \quad \text{cioè} \quad \begin{cases} x_1 = \lambda x_2, \\ (1 + \lambda^2)x_2 = 0, \\ (1 - \lambda)x_3 = 0. \end{cases} \quad (13.5)$$

Siccome siamo su  $\mathbb{R}$ , il numero  $1 + \lambda^2$  è positivo; quindi (13.5) forza  $x_1 = x_2 = 0$  per qualunque valore di  $\lambda$ . Se poi  $\lambda \neq 1$ , la terza equazione implica anche  $x_3 = 0$ , e  $\lambda \notin \text{sp}(T)$ ; se invece  $\lambda = 1$ , la terza equazione è soddisfatta per qualunque valore di  $x_3$ . Quindi  $\text{sp}(T) = \{1\}$  e  $V_1 = \text{Span}(e_3)$ . Dunque  $T$  ha un solo autovalore, e l'autospazio relativo ha dimensione 1; in particolare, non esiste una base di  $\mathbb{R}^3$  composta da autovettori di  $T$ .  $\square$

### Esempio 13.4

Ora consideriamo lo stesso endomorfismo  $T$ , ma sui complessi. In altri termini, prendiamo  $V = \mathbb{C}^3$ , e sia  $T: \mathbb{C}^3 \rightarrow \mathbb{C}^3$  l'endomorfismo rappresentato ancora dalla matrice (13.4). Ovviamente,  $1 \in \text{sp}(T)$  e  $V_1 \supseteq \text{Span}(e_3)$  esattamente come prima; ma nel momento in cui andiamo a risolvere il sistema (13.5) le cose cambiano. Infatti, se  $\lambda$  è un numero complesso non è più detto che  $1 + \lambda^2$  sia sempre diverso da zero; per  $\lambda = \pm i$  abbiamo

$1 + \lambda^2 = 0$ . Dunque, se  $\lambda \neq 1, \pm i$  il sistema (13.5) continua ad ammettere soltanto la soluzione  $x = O$ , e se  $\lambda = 1$  otteniamo ancora semplicemente  $V_1 = \text{Span}(e_3)$ . Ma se  $\lambda = \pm i$  la seconda equazione è automaticamente soddisfatta, la terza fornisce  $x_3 = 0$  e la prima diventa  $x_1 = \pm ix_2$ . Dunque esistono soluzioni non nulle: per l'esattezza,

$$V_i = \text{Span} \left( \begin{vmatrix} i \\ 1 \\ 0 \end{vmatrix} \right), \quad V_{-i} = \text{Span} \left( \begin{vmatrix} -i \\ 1 \\ 0 \end{vmatrix} \right).$$

Di conseguenza,  $\text{sp}(T) = \{1, i, -i\}$ , e  $\mathbb{C}^3$  ammette una base composta da autovettori di  $T$ : la base

$$\mathcal{B} = \left\{ \begin{vmatrix} i \\ 1 \\ 0 \end{vmatrix}, \begin{vmatrix} -i \\ 1 \\ 0 \end{vmatrix}, \begin{vmatrix} 0 \\ 0 \\ 1 \end{vmatrix} \right\}.$$

Rispetto a  $\mathcal{B}$ , l'endomorfismo  $T$  è rappresentato dalla matrice diagonale

$$\begin{vmatrix} i & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & 1 \end{vmatrix}.$$

□

### Osservazione 13.2

La differenza fra questi due esempi è uno dei motivi che ci costringono a lavorare anche con spazi vettoriali complessi. Mentre un endomorfismo su uno spazio vettoriale complesso di dimensione  $n$  ammette sempre esattamente  $n$  autovalori, contati con le relative molteplicità<sup>(2)</sup>, in spazi vettoriali reali questo non è sempre vero: possono mancare degli autovalori. Il problema è che, come vedremo fra poco, gli autovalori si ottengono come radici di un polinomio di grado  $n$ ; sui complessi, grazie al Teorema fondamentale dell'Algebra (Teorema 4C.2), tutti i polinomi di grado  $n$  hanno, contando le molteplicità, esattamente  $n$  radici, mentre sui reali può non essere così (pensa al polinomio  $p(t) = t^2 + 1$ ). Questo rende la teoria nel caso reale sensibilmente più complicata: bisogna iniziare a distinguere fra endomorfismi con tutti gli autovalori reali ed endomorfismi con anche autovalori complessi. Allora tanto vale lavorare direttamente sui complessi, dove tale distinzione svanisce, e poi applicare i risultati al caso particolare in cui tutti i numeri complessi coinvolti sono reali – e questo è quanto faremo.

### Osservazione 13.3

Un'altra differenza evidente negli ultimi due esempi è che nel secondo caso esisteva una base di autovettori, mentre nel primo no. Si potrebbe pensare che sui complessi basi di autovettori esistano sempre: questo sarebbe particolarmente piacevole, in quanto rispetto a una base di autovettori un endomorfismo è rappresentato da una matrice molto semplice, addirittura diagonale. Sfortunatamente, questo non è vero: come mostrato nel prossimo esempio, anche sui complessi ci sono endomorfismi che non ammettono basi di autovettori. Rimane comunque vero che sui complessi è sempre possibile trovare basi rispetto a cui l'endomorfismo si esprime con matrici semplici (nella fattispecie, triangolari superiori, vedi il Paragrafo 14.2), mentre sui reali la situazione è più complicata.

<sup>(2)</sup> Nell'Esempio 13.2 l'autovalore  $-1$  appare con molteplicità 2, nel senso che l'autospazio relativo ha dimensione 2. Invece, nell'Esempio 13.3 l'autovalore 1 ha molteplicità 1. Daremo fra breve la definizione precisa di molteplicità di un autovalore.

**Esempio 13.5**

Prendiamo  $V = \mathbb{R}^2$  e sia  $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  rappresentato dalla matrice

$$\begin{vmatrix} -5 & -9 \\ 4 & 7 \end{vmatrix}.$$

Questa volta non ci sono autovettori o autovalori evidenti, per cui consideriamo subito il sistema  $T(x) = \lambda x$ . Questo sistema è equivalente al sistema  $(T - \lambda I_2)x = O$ , che a sua volta ammette soluzioni diverse dal vettore nullo se e solo se la matrice  $T - \lambda I_2$  è singolare, cioè se e solo se  $\det(T - \lambda I_2) = 0$ . Ora,

$$\det(T - \lambda I_2) = \det \begin{vmatrix} -5 - \lambda & -9 \\ 4 & 7 - \lambda \end{vmatrix} = (-5 - \lambda)(7 - \lambda) - 4(-9) = (\lambda - 1)^2.$$

Dunque  $\det(T - \lambda I_2) = 0$  se e solo se  $\lambda = 1$ , per cui  $\text{sp}(T) = \{1\}$ . Per trovare l'autospazio dobbiamo risolvere il sistema  $T(x) = x$ , cioè

$$\begin{cases} -5x_1 - 9x_2 = x_1, \\ 4x_1 + 7x_2 = x_2, \end{cases} \quad \text{ovvero} \quad \begin{cases} 6x_1 = -9x_2, \\ 4x_1 = -6x_2, \end{cases}$$

che ha come soluzione

$$V_1 = \text{Span} \left( \begin{vmatrix} -3 \\ 2 \end{vmatrix} \right).$$

Dunque non esiste una base di autovettori. E questo accade anche considerando  $T$  come endomorfismo su  $\mathbb{C}^2$ . Infatti, in nessun momento abbiamo utilizzato il fatto che  $\lambda$  possa essere reale: la stessa dimostrazione funziona anche sui complessi. Infine notiamo che

rispetto alla base  $\left\{ \begin{vmatrix} -3 \\ 2 \end{vmatrix}, e_2 \right\}$  l'endomorfismo  $T$  è rappresentato dalla matrice

$$\begin{vmatrix} 1 & 3 \\ 0 & 1 \end{vmatrix},$$

che è triangolare superiore. □

Dunque un'altra questione che questi esempi invitano a considerare è: *quando esiste una base di autovettori per un dato endomorfismo  $T$ ?* Questo problema è legato a quello, già accennato alla fine del Paragrafo 8.2, di trovare una base in cui l'endomorfismo sia rappresentato da una matrice semplice, come mostra la seguente proposizione.

**Proposizione 13.1**

*Sia  $T: V \rightarrow V$  un endomorfismo di uno spazio vettoriale  $V$ , e  $\mathcal{B}$  una base di  $V$ . Allora  $\mathcal{B}$  è composta da autovettori di  $T$  se e solo se la matrice che rappresenta  $T$  rispetto a  $\mathcal{B}$  è diagonale.*

**Dimostrazione**

Sia  $\mathcal{B} = \{v_1, \dots, v_n\}$  una base composta da autovettori di  $T$ . Allora  $T(v_j) = \lambda_j v_j$  per  $j = 1, \dots, n$ , e quindi la matrice che rappresenta  $T$  rispetto a questa base è la matrice diagonale di elementi  $\lambda_1, \dots, \lambda_n$ .

Viceversa,  $T$  è rappresentato da una matrice diagonale di elementi  $\lambda_1, \dots, \lambda_n$  rispetto alla base  $\mathcal{B} = \{v_1, \dots, v_n\}$  proprio quando  $T(v_j) = \lambda_j v_j$  per  $j = 1, \dots, n$ , e quindi quando  $\mathcal{B}$  è composta da autovettori di  $T$ .

**Corollario 13.2**

Sia  $A$  una matrice quadrata di ordine  $n$ . Allora le seguenti affermazioni sono equivalenti:

- (1) esiste una base di autovettori per  $L_A$ ;
- (2)  $A$  è simile a una matrice diagonale;
- (3) la classe di similitudine di  $A$  contiene una matrice diagonale.

**Dimostrazione**

Segue dalle Proposizioni 13.1 e 8.2. □

**Definizione 13.2**

Un endomorfismo  $T: V \rightarrow V$  è *diagonalizzabile* se esiste una base di  $V$  composta da autovettori di  $T$ .

Dunque un endomorfismo è diagonalizzabile se e solo se esiste una base rispetto a cui è rappresentato da una matrice diagonale. Questo e il prossimo capitolo saranno dedicati a identificare classi di endomorfismi diagonalizzabili (gli endomorfismi simmetrici, per esempio), anche se non daremo una soluzione completa del problema di quali endomorfismi sono diagonalizzabili (vedi però l'Osservazione 13.8, e anche il Teorema 14.3 e relativi corollari).

Se l'endomorfismo  $T$  non è diagonalizzabile, possiamo provare a triangolarlo.

**Definizione 13.3**

Un endomorfismo  $T: V \rightarrow V$  è *triangolabile* se esiste una base  $\mathcal{B}$  di  $V$  rispetto a cui  $T$  è rappresentato da una matrice triangolare superiore; in tal caso diremo che la base  $\mathcal{B}$  *triangolarizza*  $T$ .

E chiaro che per proseguire abbiamo bisogno di maggiori informazioni su autovalori e autovettori, e soprattutto su come trovarli. Quindi concludiamo questa introduzione e mettiamoci al lavoro.

## 13.2 Il polinomio caratteristico

Nell'Esempio 13.5 abbiamo trovato gli autovalori di un endomorfismo come radici dell'equazione  $\det(T - \lambda I_2) = 0$ . Questo è il modo con cui si recuperano in generale.

**Teorema 13.3**

Sia  $T: V \rightarrow V$  un endomorfismo di uno spazio vettoriale  $V$  sul campo  $\mathbb{K}$ . Fissiamo una base  $\mathcal{B} = \{v_1, \dots, v_n\}$  di  $V$ , e sia  $A \in M_{n,n}(\mathbb{K})$  la matrice che rappresenta  $T$  rispetto a  $\mathcal{B}$ . Allora si ha:

- (1) la funzione  $p_T: \mathbb{K} \rightarrow \mathbb{K}$  data da

$$p_T(\lambda) = \det(A - \lambda I_n)$$

non dipende dalla base  $\mathcal{B}$  scelta;

- (2)  $p_T$  è un polinomio di grado  $n$  in cui il coefficiente direttivo è  $(-1)^n$ , il termine noto è  $\det A$  e il coefficiente di  $\lambda^{n-1}$  è  $(-1)^{n-1} \operatorname{tr} A$ , dove la traccia  $\operatorname{tr} A$  della matrice  $A$  è la somma degli elementi sulla diagonale;

- (3)  $\lambda_0 \in \mathbb{K}$  è un autovalore di  $T$  se e solo se  $p_T(\lambda_0) = 0$ .

**Osservazione 13.4**

Come sempre, buona parte di quanto faremo vale non solo su  $\mathbb{R}$  o su  $\mathbb{C}$  ma anche su altri campi, quali per esempio  $\mathbb{Q}$ . In pratica, però, in questo e nel prossimo capitolo frasi del tipo “sul campo  $\mathbb{K}$ ” saranno da intendersi come abbreviazioni di “su  $\mathbb{R}$  o su  $\mathbb{C}$ ”.

**Dimostrazione**

(1) Se la matrice  $A'$  rappresenta  $T$  rispetto a un'altra base, per la Proposizione 8.2 esiste una matrice invertibile  $B$  tale che  $A' = B^{-1}AB$  e dunque

$$\begin{aligned}\det(A' - \lambda I_n) &= \det(B^{-1}AB - \lambda I_n) = \det(B^{-1}AB - \lambda B^{-1}B) \\ &= \det(B^{-1}(A - \lambda I_n)B) = \det(B^{-1})\det(A - \lambda I_n)\det(B) \\ &= (\det B)^{-1}(\det B)\det(A - \lambda I_n) = \det(A - \lambda I_n);\end{aligned}$$

pertanto  $p_T$  non dipende dalla base scelta.

(2) Procediamo per induzione su  $n$ . Per  $n = 2$  si ha

$$\begin{aligned}p_T(\lambda) &= \det \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} \\ &= \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21},\end{aligned}$$

e ci siamo. Supponiamolo vero per  $n - 1$ , e dimostriamolo per  $n$ . Prima di tutto, il determinante è ottenuto facendo somme e prodotti degli elementi della matrice, quindi  $p_T(\lambda)$  contiene solo somme e prodotti di costanti e di  $\lambda$ , per cui è un polinomio in  $\lambda$ . Inoltre, siccome nel calcolo del determinante ogni elemento della matrice interviene al più una volta sola in ogni singolo addendo di uno sviluppo di Laplace, e nella matrice  $A - \lambda I_n$  il parametro  $\lambda$  compare solo in  $n$  elementi, ne segue che  $p_T(\lambda)$  ha al più grado  $n$ . Ora, possiamo scrivere

$$\begin{aligned}p_T(\lambda) &= \det \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} \\ &= (a_{11} - \lambda) \det \begin{vmatrix} a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} + \text{termini di grado al più } n - 2,\end{aligned}$$

dove abbiamo sviluppato il determinante lungo la prima colonna, e il resto sono determinanti di matrici in cui  $\lambda$  appare soltanto  $n - 2$  volte, per cui forniscono polinomi di grado al più  $n - 2$ . Per ipotesi induttiva abbiamo allora

$$\begin{aligned}p_T(\lambda) &= (a_{11} - \lambda)[(-1)^{n-1}\lambda^{n-1} + (-1)^{n-2}(a_{22} + \cdots + a_{nn})\lambda^{n-2} + \cdots] + \cdots \\ &= (-1)^n\lambda^n + (-1)^{n-1}(\text{tr } A)\lambda^{n-2} + \cdots,\end{aligned}$$

dove i puntini rappresentano termini di grado minore o uguale a  $n - 2$ . Dunque abbiamo dimostrato che  $p_T$  è un polinomio di grado esattamente  $n$  con coefficiente direttivo  $(-1)^n$ , e con  $(-1)^{n-1}\text{tr } A$  come coefficiente di  $\lambda^{n-1}$ . Infine il termine noto di  $p_T(\lambda)$  è dato da  $p_T(0) = \det(A - 0I_n) = \det A$ .

(3) Lo scalare  $\lambda_0 \in \mathbb{K}$  è un autovalore di  $T$  se e solo se il sistema  $Ax = \lambda_0 x$  ammette una soluzione  $x \neq 0$ , se e solo se il sistema  $(A - \lambda_0 I_n)x = 0$  ammette una soluzione non nulla, se e solo se  $0 = \det(A - \lambda_0 I_n) = p_T(\lambda_0)$ , grazie al Teorema 7.7 e al Corollario 9.10.  $\square$

### Definizione 13.4

Sia  $T: V \rightarrow V$  un endomorfismo di uno spazio vettoriale  $V$  sul campo  $\mathbb{K}$ . Il polinomio  $p_T \in \mathbb{K}[\lambda]$  definito nel Teorema 13.3 si chiama *polinomio caratteristico* dell'endomorfismo  $T$ . Ricordando la Definizione 9.3, si vede facilmente che  $p_T(\lambda) = \det(T - \lambda \text{id}_V)$ .

### Definizione 13.5

Sia  $T: V \rightarrow V$  un endomorfismo di uno spazio vettoriale  $V$ . La *traccia*  $\text{tr } T$  di  $T$  è la traccia della matrice associata a  $T$  rispetto a una qualunque base di  $V$ . Grazie al Teorema 13.3, la traccia di un endomorfismo è ben definita, in quanto a meno del segno è uguale al coefficiente di  $\lambda^{n-1}$  del polinomio caratteristico di  $T$  (ma vedi anche l'Esercizio 7.25).

Dunque gli autovalori di un endomorfismo  $T$  sono esattamente le radici del suo polinomio caratteristico. Questo è il motivo per cui su  $\mathbb{R}$  l'endomorfismo  $T$  potrebbe non avere tutti gli autovalori necessari: le radici del polinomio caratteristico potrebbero non essere tutte reali. Sui complessi questo non può succedere; infatti, il Teorema fondamentale dell'Algebra (Teorema 4C.2) ci assicura che un polinomio a coefficienti complessi di grado  $n$  ha esattamente  $n$  radici, contate con la relativa molteplicità.

### Definizione 13.6

Sia  $T: V \rightarrow V$  un endomorfismo di uno spazio vettoriale di dimensione  $n$  sul campo  $\mathbb{K}$ . Allora diremo che  $T$  ha tutti gli autovalori in  $\mathbb{K}$  se il polinomio caratteristico  $p_T$  ha esattamente  $n$  radici nel campo  $\mathbb{K}$ , contate con la relativa molteplicità (vedi la Definizione 4C.2), cioè se esistono  $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ , non necessariamente distinti, tali che

$$p_T(\lambda) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda).$$

Chiaramente, ogni endomorfismo di uno spazio vettoriale complesso ha tutti gli autovalori in  $\mathbb{C}$ , mentre sui reali questo non capita sempre (vedi l'Esempio 13.6). Tra parentesi, l'avere tutti gli autovalori nel campo è una condizione necessaria per la triangolabilità.

### Lemma 13.4

*Sia  $T: V \rightarrow V$  un endomorfismo triangolabile di uno spazio vettoriale sul campo  $\mathbb{K}$ . Allora  $T$  ha tutti gli autovalori in  $\mathbb{K}$ .*

### Dimostrazione

Sia  $B$  una base di  $V$  che triangolarizza  $T$ , e sia

$$A = \begin{vmatrix} \lambda_1 & a_{12} & \cdots & a_{1n} \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_n \end{vmatrix}$$

la matrice triangolare superiore che rappresenta  $T$  rispetto a  $\mathcal{B}$ . Allora

$$p_T(\lambda) = \det(A - \lambda I_n) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda),$$

e  $T$  ha tutti gli autovalori in  $\mathbb{K}$ .

Una conseguenza immediata del Teorema 13.3 è il seguente corollario.

### Corollario 13.5

*Sia  $T: V \rightarrow V$  un endomorfismo di uno spazio vettoriale sul campo  $\mathbb{K}$ . Supponiamo che  $T$  abbia tutti gli autovalori in  $\mathbb{K}$ , e siano  $\lambda_1, \dots, \lambda_n$  le radici del polinomio caratteristico di  $T$ . Allora*

$$\operatorname{tr} T = \lambda_1 + \cdots + \lambda_n \quad \text{e} \quad \det T = \lambda_1 \cdots \lambda_n.$$

### Dimostrazione

Per ipotesi possiamo scrivere

$$p_T(\lambda) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda); \quad (13.6)$$

in particolare,  $\det T = p_T(0) = \lambda_1 \cdots \lambda_n$ . Inoltre, è facile dimostrare per induzione su (esercizio) che (13.6) implica

$$p_T(\lambda) = (-1)^n \lambda^n + (-1)^{n-1} (\lambda_1 + \cdots + \lambda_n) \lambda^{n-1} + \cdots,$$

e il Teorema 13.3.(2) ci fornisce  $\operatorname{tr} T = \lambda_1 + \cdots + \lambda_n$ .  $\square$

### Osservazione 13.5

Dunque per trovare gli autovalori di un endomorfismo basta calcolare le radici di un polinomio. Questo risultato è molto gradevole dal punto di vista teorico, ma può creare dei problemi sul piano pratico. Infatti, mentre si sa perfettamente come trovare le radici di un polinomio di secondo grado, per gradi più alti la situazione non è così semplice.<sup>(3)</sup> Comunque, ci sono metodi molto efficienti per trovare delle approssimazioni delle radici di un polinomio qualunque, e in molte applicazioni queste approssimazioni sono altrettanto utili del valore esatto.

## 13.3 Molteplicità

Ora che sappiamo come trovare gli autovalori (basta calcolare le radici del polinomio caratteristico), e quindi i relativi autovettori (basta risolvere il sistema  $T(v) = \lambda v$  con  $\lambda$  autovalore), possiamo dedicarci a trovare un criterio di diagonalizzabilità. Cominciamo con la seguente proposizione.

### Proposizione 13.6

*Sia  $T: V \rightarrow V$  un endomorfismo di uno spazio vettoriale  $V$  sul campo  $\mathbb{K}$ , e siano  $v_1, \dots, v_k \neq 0$  autovettori di  $T$  corrispondenti ad autovalori distinti  $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ . Allora  $v_1, \dots, v_k$  sono linearmente indipendenti.*

<sup>(3)</sup> Esistono delle formule che esprimono le radici di polinomi di terzo e quarto grado in funzione dei coefficienti, ma si può dimostrare che non esistono formule del genere per i polinomi di grado dal quinto in su.

**Dimostrazione**

Procediamo per induzione su  $k$ . Per  $k = 1$  non c'è nulla da dimostrare; sia allora vero per  $k - 1$ , cioè supponiamo che  $v_1, \dots, v_{k-1}$  siano linearmente indipendenti, e siano  $\alpha_1, \dots, \alpha_k \in \mathbb{K}$  tali che

$$\alpha_1 v_1 + \dots + \alpha_k v_k = O. \quad (13.7)$$

Applicando  $T$  a questa relazione troviamo

$$O = \alpha_1 T(v_1) + \dots + \alpha_k T(v_k) = \alpha_1 \lambda_1 v_1 + \dots + \alpha_k \lambda_k v_k. \quad (13.8)$$

L'uguaglianza (13.7) ci dice che  $\alpha_k v_k = -\alpha_1 v_1 - \dots - \alpha_{k-1} v_{k-1}$ ; inserendolo in (13.8) otteniamo  $\alpha_1(\lambda_1 - \lambda_k)v_1 + \dots + \alpha_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1} = O$ . Essendo  $v_1, \dots, v_{k-1}$  linearmente indipendenti, questo implica

$$\alpha_1(\lambda_1 - \lambda_k) = \dots = \alpha_{k-1}(\lambda_{k-1} - \lambda_k) = 0.$$

Ma i  $\lambda_j$  sono tutti distinti per ipotesi; quindi  $\alpha_1 = \dots = \alpha_{k-1} = 0$ . Mettendo questo in (13.7) troviamo anche  $\alpha_k = 0$ , e  $v_1, \dots, v_k$  sono linearmente indipendenti.  $\square$

**Osservazione 13.6**

In particolare, se  $v_1, \dots, v_k \in V$  sono autovettori relativi ad autovalori distinti  $\lambda_1, \dots, \lambda_k$  di un endomorfismo  $T: V \rightarrow V$ , e si ha  $v_1 + \dots + v_k = O$ , allora  $v_1 = \dots = v_k = O$ . Infatti, se anche uno solo dei  $v_j$  fosse diverso dal vettore nullo avremmo trovato una relazione di dipendenza lineare fra autovettori relativi ad autovalori distinti, contraddicendo la Proposizione 13.6.

Una prima conseguenza è il seguente corollario.

**Corollario 13.7**

*Sia  $T: V \rightarrow V$  un endomorfismo di uno spazio vettoriale di dimensione  $n$  sul campo  $\mathbb{K}$ , e supponiamo che  $T$  abbia esattamente  $n$  autovalori distinti in  $\mathbb{K}$ . Allora  $T$  è diagonalizzabile.*

**Dimostrazione**

Siano  $v_1, \dots, v_n \in V$  degli autovettori relativi agli  $n$  autovalori distinti  $\lambda_1, \dots, \lambda_n$ . Dalla Proposizione 13.6, segue che  $v_1, \dots, v_n$  sono linearmente indipendenti, e quindi formano una base di  $V$ .  $\square$

**Esempio 13.6**

Consideriamo l'endomorfismo  $T$  rappresentato dalla matrice

$$T = \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}.$$

Allora  $p_T(\lambda) = \lambda^2 + 1$ , per cui  $T$  non ha autovalori reali e non è neppure triangolabile su  $\mathbb{R}$ . Sui complessi, invece, ha esattamente due autovalori distinti,  $i$  e  $-i$ , per cui è diagonalizzabile su  $\mathbb{C}$ .  $\square$

Ma cosa accade quando le radici del polinomio caratteristico non sono tutte distinte? Cominciamo con una definizione.

**Definizione 13.7**

Sia  $\lambda_0 \in \text{sp}(T)$  un autovalore di un endomorfismo  $T: V \rightarrow V$ . Diremo *molteplicità algebrica* di  $\lambda_0$  la sua molteplicità come radice del polinomio caratteristico; *molteplicità geometrica* la dimensione del relativo autospazio  $V_{\lambda_0}$ .

**Osservazione 13.7**

Per calcolare la molteplicità algebrica dell'autovalore  $\lambda_0$  di un endomorfismo  $T$  bisogna scoprire quante volte  $\lambda_0$  è radice del polinomio caratteristico. Per calcolare invece la molteplicità geometrica, basta trovare il rango di  $T - \lambda_0 \text{id}_V$ ; infatti l'autospazio  $V_{\lambda_0}$  è il nucleo di  $T - \lambda_0 \text{id}_V$ , per cui

$$\dim V_{\lambda_0} = \dim V - \text{rg}(T - \lambda_0 \text{id}_V).$$

**Esempio 13.7**

Sia  $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  l'endomorfismo rappresentato dalla matrice

$$\begin{vmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ 3 & 3 & 5 \end{vmatrix};$$

vogliamo calcolare molteplicità algebrica e geometrica dei suoi autovalori. Prima di tutto,  $p_T(\lambda) = \det(T - \lambda I_3) = -\lambda^3 + 12\lambda^2 - 36\lambda + 32$ . Si verifica subito che  $p_T(2) = 0$ , quindi  $p_T(\lambda) = (2 - \lambda)(\lambda^2 - 10\lambda + 16) = (2 - \lambda)^2(8 - \lambda)$ , e  $\text{sp}(T) = \{2, 8\}$ , dove 2 ha molteplicità algebrica 2 e 8 ha molteplicità algebrica 1. Per calcolare le molteplicità geometriche, troviamo

$$\text{rg}(T - 2I_3) = \text{rg} \begin{vmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{vmatrix} = 1,$$

e

$$\text{rg}(T - 8I_3) = \text{rg} \begin{vmatrix} -5 & 1 & 1 \\ 2 & -4 & 2 \\ 3 & 3 & -3 \end{vmatrix} = 2,$$

per cui 2 ha molteplicità geometrica 2 e 8 ha molteplicità geometrica 1.  $\square$

Le molteplicità geometriche e algebriche di un autovalore non sono completamente scollegate.

**Proposizione 13.8**

*La molteplicità algebrica di un autovalore è sempre maggiore o uguale alla sua molteplicità geometrica.*

**Dimostrazione**

Sia  $\lambda_0 \in \text{sp}(T)$  un autovalore di un endomorfismo  $T$  di uno spazio vettoriale  $V$  sul campo  $\mathbb{K}$ , e sia  $h$  la molteplicità geometrica di  $\lambda_0$ . Sia  $\{v_1, \dots, v_h\}$  una base dell'autospazio  $V_{\lambda_0}$ , e completiamola a una base  $\{v_1, \dots, v_h, v'_{h+1}, \dots, v'_n\}$  di  $V$ . Rispetto a questa base, l'endomorfismo  $T$  è rappresentato da una matrice della forma

$$A = \left| \begin{array}{c|c} \lambda_0 I_h & B \\ \hline O & C \end{array} \right|$$

dove  $B \in M_{h,n-h}(\mathbb{K})$  e  $C \in M_{n-h,n-h}(\mathbb{K})$ . Dunque

$$A - \lambda I_n = \left| \begin{array}{c|c} (\lambda_0 - \lambda)I_h & B \\ \hline O & C - \lambda J_{n-h} \end{array} \right|$$

per cui sviluppando il determinante rispetto alle prime  $h$  colonne troviamo

$$p_T(\lambda) = \det(A - \lambda I_n) = (\lambda_0 - \lambda)^h p_C(\lambda).$$

Siccome la molteplicità algebrica è il più grande  $m$  per cui  $(\lambda_0 - \lambda)^m$  può dividere  $p_T(\lambda)$  ne segue che la molteplicità algebrica è sempre maggiore o uguale alla molteplicità geometrica.  $\square$

E ora possiamo enunciare (e dimostrare) un criterio di diagonalizzabilità.

### Teorema 13.9

Sia  $T: V \rightarrow V$  un endomorfismo di uno spazio vettoriale di dimensione finita sul campo  $\mathbb{K}$ . Supponiamo che lo spettro di  $T$  sia  $\text{sp}(T) = \{\lambda_1, \dots, \lambda_h\}$ , e indichiamo con  $\mu_j$  (rispettivamente,  $\nu_j$ ) la molteplicità algebrica (rispettivamente, geometrica) di  $\lambda_j$ , per  $j = 1, \dots, h$ . Allora le seguenti affermazioni sono equivalenti:

- (1)  $T$  è diagonalizzabile;
- (2)  $T$  ha tutti gli autovalori in  $\mathbb{K}$  e la molteplicità geometrica di ciascun autovalore coincide con la sua molteplicità algebrica, cioè  $\mu_j = \nu_j$  per  $j = 1, \dots, h$ ;
- (3)  $\nu_1 + \dots + \nu_h = \dim V$ .

### Dimostrazione

(1)  $\Rightarrow$  (2) La prima affermazione è il contenuto del Lemma 13.4, mentre la seconda si verifica immediatamente su una matrice diagonale che rappresenta  $T$  rispetto a una base di autovettori.

(2)  $\Rightarrow$  (3) Siccome  $T$  ha tutti gli autovalori in  $\mathbb{K}$  si ha  $\mu_1 + \dots + \mu_h = \dim V$ , e ci siamo.

(3)  $\Rightarrow$  (1) Sia  $\mathcal{B}^j = \{v_1^j, \dots, v_{\nu_j}^j\}$  una base di  $V_{\lambda_j}$  per  $j = 1, \dots, h$ . Per ipotesi, l'insieme  $\mathcal{B} = \mathcal{B}^1 \cup \dots \cup \mathcal{B}^h$  contiene esattamente  $\dim V$  elementi; quindi se dimostriamo che sono linearmente indipendenti abbiamo trovato una base di  $V$  composta da autovettori di  $T$ . Scriviamo una relazione di dipendenza lineare fra gli elementi di  $\mathcal{B}$ , cioè scriviamo

$$\alpha_1^1 v_1^1 + \dots + \alpha_{\nu_1}^1 v_{\nu_1}^1 + \dots + \alpha_1^h v_1^h + \dots + \alpha_{\nu_h}^h v_{\nu_h}^h = 0. \quad (13.9)$$

Ponendo  $w_j = \alpha_1^j v_1^j + \dots + \alpha_{\nu_j}^j v_{\nu_j}^j \in V_{\lambda_j}$ , possiamo riscrivere la (13.9) nella forma  $w_1 + \dots + w_h = 0$ . Ma allora l'Osservazione 13.6 ci dice che  $w_1 = \dots = w_h = 0$ , e quindi tutti gli  $\alpha_i^j$  sono zero.  $\square$

### Esempio 13.8

Dunque l'endomorfismo dell'Esempio 13.7 è diagonalizzabile. Una base di autovettori, trovata risolvendo i sistemi  $T(v) = 2v$  e  $T(v) = 8v$ , è

$$\left\{ \left| \begin{array}{c} 1 \\ -1 \\ 0 \end{array} \right|, \left| \begin{array}{c} 1 \\ 0 \\ -1 \end{array} \right|, \left| \begin{array}{c} 1 \\ 2 \\ 3 \end{array} \right| \right\}.$$

### Osservazione 13.8

Il criterio dato nel Teorema 13.9 per decidere quando un endomorfismo è diagonalizzabile è abbastanza efficace ma non è perfetto, in quanto richiede il calcolo esplicito degli autovalori<sup>(4)</sup>. Vi è un altro criterio che non ha questo difetto, ma è di dimostrazione al quanto più complicata. Sia  $T: V \rightarrow V$  un endomorfismo. Indichiamo con  $T^k$  la composizione di  $T$  con se stesso  $k$  volte; per esempio,  $T^3 = T \circ T \circ T$  e così via. Per ogni polinomio  $p \in \mathbb{K}[t]$  possiamo definire un endomorfismo  $p(T): V \rightarrow V$  in questo modo, se  $p(t) = a_d t^d + a_{d-1} t^{d-1} + \dots + a_1 t + a_0$ , allora  $p(T) = a_d T^d + a_{d-1} T^{d-1} + \dots + a_1 T + a_0 \text{id}_V$ .

Ovviamente, può succedere che  $p(T)$  sia l'applicazione nulla (vedi per esempio l'Esercizio 13.17). Si può dimostrare che, a meno di costanti moltiplicative, esiste un unico polinomio  $p_0$  (che è denominato *polinomio minimo* di  $T$ ) di grado minimo tale che  $p_0(T) = O$ ; è esplicitamente calcolabile, e risulta essere un divisore del polinomio caratteristico di  $T$ . Il criterio summenzionato è il seguente:  $T$  è diagonalizzabile su  $\mathbb{K}$  se e solo se tutte le radici del polinomio minimo sono in  $\mathbb{K}$  e hanno molteplicità 1. Siccome per  $\mathbb{K} = \mathbb{R}$  e  $\mathbb{C}$  ci sono tecniche per stabilire se un polinomio ha tutte le radici in  $\mathbb{K}$  di molteplicità 1 senza bisogno di calcolarle, questo criterio permette di stabilire quando un endomorfismo è diagonalizzabile senza doverne calcolare gli autovalori. Ulteriori dettagli si possono trovare nei complementi al Capitolo 14 del testo *Geometria* di M. Abate, edito dalla McGraw-Hill Italia. Inoltre, nel prossimo capitolo vedremo un criterio molto semplice per stabilire quando un endomorfismo di uno spazio vettoriale metrico reale ammette una base *ortonormale* di autovettori.

<sup>(4)</sup> Nota che conoscere semplicemente il valore approssimato  $\tilde{\lambda}$  di un autovalore  $\lambda_0$  non basta: anche se  $\tilde{\lambda}$  è vicinissimo a  $\lambda_0$ , l'endomorfismo  $T - \lambda \text{id}_V$  rimane invertibile, e quindi non ci dice nulla sulla molteplicità geometrica di  $\lambda_0$ .