

Note di Geometria 2

**Laurea Triennale in Matematica
Università di Salerno - DipMat**

Luca Vitagliano

Copyright © 2021 Luca Vitagliano

Licensed under the Creative Commons Attribution-NonCommercial 3.0 Unported License (the “License”). You may not use this file except in compliance with the License. You may obtain a copy of the License at <http://creativecommons.org/licenses/by-nc/3.0>. Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an “AS IS” BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.

First printing, February 2021

Indice

	Introduction	7
1	Forme Bilineari	9
1.1	Applicazioni Multilineari	9
1.2	Rango di una Forma Bilineare	14
1.3	Forme Bilineari Simmetriche	17
1.4	Forme Bilineari Alternanti	32
2	Spazi Vettoriali Euclidei	37
2.1	Prodotti Scalari, Norme, Distanze e Angoli	37
2.2	Ortogonalità	41
2.3	Il Gruppo Ortogonale	52
2.4	Applicazioni Ortogonali e Isometrie	55
2.5	Isomorfismi Musicali	63
3	Diagonalizzazione di Endomorfismi	67
3.1	Autovalori e Autovettori	67
3.2	Autospazi	73
3.3	Il Criterio di Diagonalizzabilità	77
3.4	Diagonalizzazione di Matrici	84
3.5	Diagonalizzazione Ortogonale	87
4	Spazi Vettoriali Hermitiani	97
4.1	Richiami sui Numeri Complessi	97

4.2	Forme Hermitiane e Prodotti Hermitiani	100
4.3	Endomorfismi Unitari ed Hermitiani	107
5	Spazi Vettoriali Orientati	119
5.1	Orientazioni	119
5.2	Automorfismi che Conservano le Orientazioni	121
5.3	Spazi Vettoriali Euclidei Orientati	123
5.4	Classificazione degli Endomorfismi Ortogonali in Dimensione 2 e 3	127
6	Spazi Affini	135
6.1	Spazi Affini e Sottospazi Affini	135
6.2	Rappresentazioni di Sottospazi Affini	143
6.3	Geometria Affine in Dimensione 2 e 3	155
6.4	Applicazioni Affini	181
6.5	Affinità	190
7	Classificazione Affine delle Quadriche	197
7.1	Insiemi Algebrici	197
7.2	Equivalenza di Polinomi di Secondo Grado	199
7.3	Forma Normale delle Quadriche	207
8	Geometria Affine Euclidea	211
8.1	Spazi Affini Euclidei	211
8.2	Geometria Affine Euclidea in Dimensione 2 e 3	219
8.3	Isometrie di uno Spazio Affine Euclideo	232
	Indice Analitico	239

Introduzione

Notazioni e Convenzioni

Sia \mathbb{K} un campo e n un intero positivo. I vettori dello spazio vettoriale numerico n -dimensionale saranno sempre considerati come colonne di scalari:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad x_1, \dots, x_n \in \mathbb{K}.$$

Il \mathbb{K} -spazio vettoriale delle matrici di ordine $m \times n$ su \mathbb{K} sarà indicato con $M_{m,n}(\mathbb{K})$. Data una matrice $A \in M_{m,n}(\mathbb{K})$ con entrate a_{ij} , scriveremo anche

$$A = (a_{ij})_{\substack{j=1,\dots,n \\ i=1,\dots,m}} \quad (1)$$

per sottolineare che il numero di righe è m e il numero di colonne è n (si osservi che nella (1), e in formule simili, il primo indice, in questo caso i , è l'indice di riga e il secondo indice, in questo caso j , è l'indice di colonna). Data una matrice $A \in M_{m,n}(\mathbb{K})$, la i -esima riga sarà indicata $A_{(i)}$, mentre la j -esima colonna sarà indicata $A^{(j)}$. La matrice trasposta sarà indicata con A^T . Per esempio

$$(x_1, \dots, x_n)^T = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad x_1, \dots, x_n \in \mathbb{K}.$$

Il rango di una matrice A sarà indicato con $\text{rk} A$. Se A è una matrice quadrata, il suo determinante sarà indicato con $\det A$ oppure con $|A|$.

Le matrici quadrate invertibili (o, equivalentemente, le matrici con determinante non nullo) di ordine n sul campo \mathbb{K} , con il prodotto di matrici, formano un gruppo, detto il *gruppo lineare generale* di ordine n sul campo \mathbb{K} , che sarà indicato con $GL_n(\mathbb{K})$.

In questo corso, dato uno spazio vettoriale V , un insieme di vettori in V sarà anche detto un *sistema* di vettori. Sistemi e basi di V saranno di norma *ordinati* cioè si assumerà che è stato

dichiarato quale elemento del sistema si debba considerare il primo, quale il secondo, etc. Il sistema ordinato contenente i vettori v_1, \dots, v_p (v_1 il primo, v_2 il secondo, etc.) sarà indicato con il simbolo (v_1, \dots, v_p) (l'uso delle parentesi tonde al posto delle parentesi graffe dice appunto che un ordine è stato fissato). Una base ordinata prende il nome di *referimento*. Dunque, in questo corso, si parlerà di riferimenti piuttosto che di basi. Dato un referimento \mathcal{R} , il sistema di coordinate che associa ad un vettore di V la colonna delle coordinate di v nel referimento \mathcal{R} sarà indicato con $c_{\mathcal{R}} : V \rightarrow \mathbb{R}^n$. Si tratta di un isomorfismo di spazi vettoriali. Dato $v \in V$, la colonna $c_{\mathcal{R}}(v)$ è anche detta il *vettore coordinato* di v . Dato un altro referimento \mathcal{R}' di V , la matrice del cambio di coordinate dal referimento \mathcal{R} al referimento \mathcal{R}' sarà indicata con $M_{\mathcal{R}', \mathcal{R}}$. Se $f : V \rightarrow W$ è un'applicazione lineare tra spazi vettoriali, \mathcal{R} è un referimento di V ed \mathcal{S} è un referimento di W , la matrice rappresentativa di f nei riferimenti \mathcal{R}, \mathcal{S} , sarà indicata con $M_{\mathcal{S}, \mathcal{R}}^f$, o semplicemente $M_{\mathcal{R}}^f$ se $V = W$ e $\mathcal{R} = \mathcal{S}$.

Dato uno spazio vettoriale V di dimensione n sul campo \mathbb{K} , l'insieme degli automorfismi $f : V \rightarrow V$ (cioè degli isomorfismi da V a V stesso) con la composizione di funzioni, è un gruppo, detto il *gruppo lineare generale dello spazio vettoriale V* , che sarà indicato con $GL(V)$.

In queste note considereremo solo spazi vettoriali di dimensione finita senza ulteriori commenti.

1. Forme Bilineari

Il concetto di funzione lineare può essere generalizzato al caso di funzioni con più entrate, richiedendo la linearità in ciascuna entrata. Si ottiene in questo modo una nuova nozione di funzione (o applicazione) *multilineare*. In questo capitolo studieremo il caso di una funzione scalare a due entrate. Si parla in questo caso di *forma bilineare*. Un esempio standard di forma bilineare è il prodotto scalare. Questo esempio sarà studiato nel dettaglio più avanti. In effetti, lo scopo principale di questo capitolo è di raccogliere il materiale preliminare necessario ad affrontare la teoria dei prodotti scalari.

1.1 Applicazioni Multilineari

Sia \mathbb{K} un campo e V_1, \dots, V_k, V spazi vettoriali di dimensione finita sul campo \mathbb{K} .

Definizione 1.1.1 — Applicazione Multilineare. Un'applicazione k -lineare, o semplicemente *multilineare*, da V_1, \dots, V_k in V è un'applicazione

$$\mu : V_1 \times \dots \times V_k \rightarrow V$$

lineare in ciascuno dei suoi argomenti, cioè, per ogni $i = 1, \dots, k$, per ogni $v_i, u_i, w_i \in V_i$ e per ogni $a_i, b_i \in \mathbb{K}$, con $i = 1, \dots, k$, si ha

$$\mu(v_1, \dots, \underbrace{a_i u_i + b_i w_i}_{\text{posto } i}, \dots, v_k) = a_i \mu(v_1, \dots, \underbrace{u_i}_{\text{posto } i}, \dots, v_k) + b_i \mu(v_1, \dots, \underbrace{w_i}_{\text{posto } i}, \dots, v_k).$$

Se $k = 2, 3, 4, \dots$ si parla di *applicazione bilineare, trilineare, quadrilineare*, etc. Se $V = \mathbb{K}$ si parla di *forma multilineare* su V_1, \dots, V_k (o, per $k = 2, 3, 4, \dots$, *forma bilineare, trilineare, quadrilineare*, etc.).

Si noti che, nel caso $k = 1$, la Definizione 1.1.1 restituisce la nozione di applicazione lineare. Dunque le applicazioni multilineari generalizzano le applicazioni lineari al caso di più di una entrata.

■ **Esempio 1.1 — Il Prodotto di Matrici è Bilineare.** Siano n, m, p interi positivi. Il prodotto di matrici

$$M_{m,n}(\mathbb{K}) \times M_{n,p}(\mathbb{K}) \rightarrow M_{m,p}(\mathbb{K}), \quad (A, B) \mapsto AB$$

è un'applicazione bilineare (il lettore lo vede?). ■

■ **Esempio 1.2 — La Composizione di Applicazioni Lineari è Bilineare.** Questo Esempio generalizza l'Esempio 1.1. Siano U, V, W spazi vettoriali sul campo \mathbb{K} . Si ricordi che l'insieme $\text{Hom}_{\mathbb{K}}(U, V)$ delle applicazioni lineari $f : U \rightarrow V$ è esso stesso uno spazio vettoriale su \mathbb{K} in cui la somma e il prodotto per uno scalare sono definiti come segue. Per ogni $f, f' \in \text{Hom}_{\mathbb{K}}(U, V)$ e per ogni $a \in \mathbb{K}$, $f + f', af : U \rightarrow V$ sono le applicazioni lineari date da:

$$(f + f')(u) := f(u) + f'(u) \quad \text{e} \quad (af)(u) = af(u), \quad u \in U.$$

La composizione di applicazioni lineari

$$\circ : \text{Hom}_{\mathbb{K}}(V, W) \times \text{Hom}_{\mathbb{K}}(U, V) \rightarrow \text{Hom}_{\mathbb{K}}(U, W), \quad (g, f) \mapsto g \circ f \quad (1.1)$$

è un'applicazione bilineare. La semplice verifica è lasciata come Esercizio 1.1. ■

Esercizio 1.1 Siano U, V, W spazi vettoriali sul campo \mathbb{K} . Si dimostri che la composizione di applicazioni lineari (1.1) è un'applicazione bilineare. ■

■ **Esempio 1.3** Sia n un intero positivo. Il determinante di una matrice di ordine $n \times n$ può essere interpretato come una forma n -lineare

$$\det : \underbrace{\mathbb{K}^n \times \cdots \times \mathbb{K}^n}_{n \text{ volte}} \rightarrow \mathbb{K}$$

ponendo

$$\det(A^{(1)}, \dots, A^{(n)}) := \det A, \quad A^{(i)} \in \mathbb{K}^n \text{ per ogni } i = 1, \dots, n,$$

dove A è la matrice che ha per colonne i vettori numerici $A^{(1)}, \dots, A^{(n)}$:

$$A := (A^{(1)} \cdots A^{(n)}).$$

Siano V_1, \dots, V_k, V spazi vettoriali sul campo \mathbb{K} e sia k un numero naturale. L'insieme $\text{Mult}_k(V_1, \dots, V_k; V)$ delle applicazioni k -lineari $\mu : V_1 \times \cdots \times V_k \rightarrow V$ è munito di (almeno) due operazioni naturali: una *somma*

$$+ : \text{Mult}_k(V_1, \dots, V_k; V) \times \text{Mult}_k(V_1, \dots, V_k; V) \rightarrow \text{Mult}_k(V_1, \dots, V_k; V), \quad (\mu, \nu) \mapsto \mu + \nu$$

definita da

$$(\mu + \nu)(v_1, \dots, v_k) := \mu(v_1, \dots, v_k) + \nu(v_1, \dots, v_k),$$

e un *prodotto per uno scalare*

$$\cdot : \mathbb{K} \times \text{Mult}_k(V_1, \dots, V_k; V) \rightarrow \text{Mult}_k(V_1, \dots, V_k; V), \quad (a, \mu) \mapsto a \cdot \mu$$

definito da

$$(a \cdot \mu)(v_1, \dots, v_k) := a \cdot \mu(v_1, \dots, v_k),$$

per ogni $v_i \in V_i, i = 1, \dots, k$.

Proposizione 1.1.1 Con le due operazioni appena definite l'insieme $\text{Mult}_k(V_1, \dots, V_k; V)$ delle applicazioni k -lineari $\mu : V_1 \times \dots \times V_k \rightarrow V$ è un \mathbb{K} -spazio vettoriale. Il vettore nullo in $\text{Mult}_k(V_1, \dots, V_k; V)$ è l'applicazione multilineare nulla: $0 : V_1 \times \dots \times V_k \rightarrow V, (v_1, \dots, v_k) \mapsto 0$.

Dimostrazione. Si tratta di una semplice verifica degli assiomi che lasciamo al lettore come Esercizio 1.2. ■

Esercizio 1.2 Dimostrare la Proposizione 1.1.1. ■

Si noti che lo spazio vettoriale $\text{Mult}_1(V_1; V)$ delle applicazioni 1-lineari ($k = 1$) $\mu : V_1 \rightarrow V$ (incluse le sue operazioni) coincide con lo spazio vettoriale delle applicazioni lineari $V_1 \rightarrow V$. Per esempio se, in aggiunta, $V = \mathbb{K}$, si tratta dello spazio vettoriale duale di V_1 .

Per la convenienza del lettore esplicitiamo ora la definizione di *forma bilineare*, che giocherà il ruolo principale in questo capitolo. Dunque, in base alla Definizione 1.1.1, una *forma bilineare* sugli spazi vettoriali V, W è un'applicazione

$$\beta : V \times W \rightarrow \mathbb{K}$$

tale che, per ogni $v, v' \in V$, ogni $w, w' \in W$, e ogni $a, a', b, b' \in \mathbb{K}$,

$$\beta(av + a'v', w) = a\beta(v, w) + a'\beta(v', w),$$

e inoltre

$$\beta(v, bw + b'w') = b\beta(v, w) + b'\beta(v, w').$$

■ **Esempio 1.4** Siano n, m interi positivi, e sia $B = (b_{ij})_{i=1, \dots, m}^{j=1, \dots, n} \in M_{m,n}(\mathbb{K})$ una matrice di ordine $m \times n$ sul campo \mathbb{K} . È facile dimostrare, utilizzando le proprietà del prodotto di matrici, che l'applicazione

$$\beta_B : \mathbb{K}^m \times \mathbb{K}^n \rightarrow \mathbb{K},$$

definita da

$$\beta_B(x, y) := x^T B y = \sum_{i=1}^m \sum_{j=1}^n x_i b_{ij} y_j,$$

per ogni $x = (x_1, \dots, x_m)^T \in \mathbb{K}^m$ e ogni $y = (y_1, \dots, y_n)^T \in \mathbb{K}^n$, è una forma bilineare su $\mathbb{K}^m, \mathbb{K}^n$. I dettagli sono lasciati al lettore come Esercizio 1.3. ■

Esercizio 1.3 Dimostrare che l'applicazione β_B definita nell'Esempio 1.4 è una forma bilineare su $\mathbb{K}^m, \mathbb{K}^n$. ■

Il prossimo teorema mostra che, in coordinate, "ogni forma bilineare è del tipo β_B " discusso nell'Esempio 1.4, per qualche matrice B .

Teorema 1.1.2 Siano V, W spazi vettoriali sul campo \mathbb{K} e sia

$$\beta : V \times W \rightarrow \mathbb{K}$$

una forma bilineare su V, W . Siano inoltre $\mathcal{R} = (v_1, \dots, v_m)$ ed $\mathcal{S} = (w_1, \dots, w_n)$ un riferimento di V e un riferimento di W rispettivamente (sicché $m = \dim V$ e $n = \dim W$). Allora esiste un'unica matrice $B \in M_{m,n}(\mathbb{K})$ tale che, per ogni $v \in V$ e ogni $w \in W$,

$$\beta(v, w) = \beta_B(x, y) = x^T B y, \tag{1.2}$$

in cui $x \in \mathbb{K}^m$ e $y \in \mathbb{K}^n$ sono i vettori coordinati di v e w nei riferimenti \mathcal{R} ed \mathcal{S} rispettivamente.

Dimostrazione. Per dimostrare l'esistenza della matrice B come nell'enunciato, poniamo $B = (b_{ij})_{i=1, \dots, m}^{j=1, \dots, n}$ con

$$b_{ij} := \beta(v_i, w_j), \quad i = 1, \dots, m, \quad j = 1, \dots, n. \quad (1.3)$$

Consideriamo ora $v \in V$ e $w \in W$ e siano $x = (x_1, \dots, x_m)^T \in \mathbb{K}^m$ e $y = (y_1, \dots, y_n)^T \in \mathbb{K}^n$ i rispettivi vettori coordinati (nei riferimenti \mathcal{R} ed \mathcal{S}). Questo vuol dire che

$$v = \sum_{i=1}^m x_i v_i \quad \text{e} \quad w = \sum_{j=1}^n y_j w_j.$$

Dunque

$$\beta(v, w) = \beta\left(\sum_{i=1}^m x_i v_i, \sum_{j=1}^n y_j w_j\right) = \sum_{i=1}^m \sum_{j=1}^n x_i y_j \beta(v_i, w_j) = \sum_{i=1}^m \sum_{j=1}^n x_i y_j b_{ij} = x^T B y$$

come richiesto (si noti come è stata usata, nel secondo passaggio, la linearità di β in entrambi gli argomenti).

Per dimostrare l'unicità supponiamo che esista un'altra matrice $B' = (b'_{ij})_{i=1, \dots, m}^{j=1, \dots, n} \in M_{m,n}(\mathbb{K})$ tale che, per ogni $v \in V$ e $w \in W$,

$$\beta(v, w) = x^T B' y$$

dove $x = c_{\mathcal{R}}(v)$ (il vettore coordinato di v nel riferimento \mathcal{R}) e $y = c_{\mathcal{S}}(w)$ (il vettore coordinato di w nel riferimento \mathcal{S}). In particolare possiamo scegliere $v = v_i$ e $w = w_j$, da cui

$$c_{\mathcal{R}}(v) = c_{\mathcal{R}}(v_i) = E_i := (0, \dots, \underbrace{1}_{\text{posto } i}, \dots, 0)^T,$$

l' i -esimo vettore del riferimento canonico di \mathbb{K}^m e

$$c_{\mathcal{S}}(w) = c_{\mathcal{S}}(w_j) = E_j := (0, \dots, \underbrace{1}_{\text{posto } j}, \dots, 0)^T,$$

il j -esimo vettore del riferimento canonico di \mathbb{K}^n . Perciò

$$b_{ij} = \beta(v_i, w_j) = E_i^T B' E_j = b'_{ij}, \quad \text{per ogni } i = 1, \dots, m \text{ e } j = 1, \dots, n,$$

cioè $B = B'$. Questo conclude la dimostrazione. ■

Definizione 1.1.2 — Matrice Rappresentativa di una Forma Bilineare. La matrice B nel Teorema 1.1.2 è detta la *matrice rappresentativa* della forma bilineare β nei riferimenti \mathcal{R}, \mathcal{S} e qualche volta si scrive anche

$$B =: M_{\mathcal{R}, \mathcal{S}}^{\beta}.$$

■ **Esempio 1.5** Sia $B = (b_{ij})_{i=1, \dots, m}^{j=1, \dots, n} \in M_{m,n}(\mathbb{K})$ una matrice di ordine $m \times n$ sul campo \mathbb{K} e sia $\beta_B : \mathbb{K}^m \times \mathbb{K}^n \rightarrow \mathbb{K}$ la forma bilineare definita da B come nell'Esempio 1.4. La matrice rappresentativa di β_B nei riferimenti canonici di $\mathbb{K}^m, \mathbb{K}^n$ è esattamente B . Infatti,

$$\beta_B(E_i, E_j) = E_i^T B E_j = b_{ij}$$

per ogni $i = 1, \dots, m$ e per ogni $j = 1, \dots, n$ (se necessario, il lettore verifichi il conto nel dettaglio). ■

Se $V = W$ ed $\mathcal{R} = \mathcal{S}$, la matrice $M_{\mathcal{R},\mathcal{S}}^\beta$ è più spesso detta la matrice rappresentativa della forma bilineare β nel riferimento \mathcal{R} e si indica anche con $M_{\mathcal{R}}^\beta$ (invece che $M_{\mathcal{R},\mathcal{R}}^\beta$). Si osservi che, una volta fissati i riferimenti $\mathcal{R} = (v_1, \dots, v_m)$ ed $\mathcal{S} = (w_1, \dots, w_m)$, la matrice rappresentativa è completamente determinata dalla forma bilineare tramite la Formula (1.3) e, viceversa, la forma bilineare è completamente determinata dalla matrice rappresentativa tramite la Formula (1.2). Per questo motivo, una forma bilineare e la sua matrice rappresentativa dovrebbero essere interpretate come due versioni (equivalenti) dello stesso oggetto matematico.

- Come abbiamo appena discusso, una forma bilineare è completamente determinata dalla sua matrice rappresentativa (in una coppia di riferimenti prefissati). In vista della Formula (1.3) dunque, una forma bilineare è completamente determinata dai suoi valori sui vettori di base. D'altro canto non è difficile dimostrare che, data una matrice B (invece che una forma bilineare), la Formula (1.2) definisce una forma bilineare che ha proprio B come matrice rappresentativa (il lettore provi a verificarlo per esercizio!). Dunque, per definire una forma bilineare, è sufficiente assegnare i suoi valori sui vettori di base. In altre parole, vale una "versione bilineare" del Teorema dell'Estensione Lineare.

La matrice rappresentativa di una forma bilineare dipende dai riferimenti fissati e cambia quando cambiamo i riferimenti. Vogliamo ora vedere come. Sia dunque

$$\beta : V \times W \rightarrow \mathbb{K}$$

una forma bilineare sugli spazi vettoriali V, W , siano $\mathcal{R} = (v_1, \dots, v_m)$ ed $\mathcal{R}' = (v'_1, \dots, v'_m)$ riferimenti di V (sicché $\dim V = m$) e siano $\mathcal{S} = (w_1, \dots, w_n)$ ed $\mathcal{S}' = (w'_1, \dots, w'_n)$ riferimenti di W (sicché $\dim W = n$). Infine siano $M_{\mathcal{R}',\mathcal{R}} = (a_{ij})_{i=1,\dots,m}^{j=1,\dots,m}$ e $M_{\mathcal{S}',\mathcal{S}} = (c_{kl})_{k=1,\dots,n}^{l=1,\dots,n}$ le matrici del cambio di coordinate da \mathcal{R} a \mathcal{R}' e da \mathcal{S} a \mathcal{S}' rispettivamente.

- Ricordiamo che, per definizione di matrice del cambio di coordinate, $M_{\mathcal{R}',\mathcal{R}}$ è l'unica matrice invertibile con la proprietà che, per ogni $v \in V$, detti x, x' i vettori coordinati di v nei riferimenti $\mathcal{R}, \mathcal{R}'$ rispettivamente, si ha che

$$x' = M_{\mathcal{R}',\mathcal{R}}x.$$

In questo senso, la matrice $M_{\mathcal{R}',\mathcal{R}}$ consente di *passare* dalle coordinate nel riferimento \mathcal{R} alle coordinate nel riferimento \mathcal{R}' . La j -esima colonna della matrice $M_{\mathcal{R}',\mathcal{R}}$ è il vettore coordinato del j -esimo vettore v_j del riferimento \mathcal{R} nel riferimento \mathcal{R}' . In altre parole se $M_{\mathcal{R}',\mathcal{R}} = (a_{ij})_{i=1,\dots,m}^{j=1,\dots,m}$, allora

$$v_j = \sum_{i=1}^m a_{ij}v'_i, \quad j = 1, \dots, m.$$

Ricordiamo anche per futuri utilizzi che

$$M_{\mathcal{R},\mathcal{R}} = I_m \quad \text{e} \quad M_{\mathcal{R},\mathcal{R}'} = M_{\mathcal{R}',\mathcal{R}}^{-1},$$

per ogni due riferimenti $\mathcal{R}, \mathcal{R}'$. Inoltre, se \mathcal{R}'' è un terzo riferimento, si ha che

$$M_{\mathcal{R}'',\mathcal{R}} = M_{\mathcal{R}'',\mathcal{R}'}M_{\mathcal{R}',\mathcal{R}}.$$

Le matrici rappresentative di β nei riferimenti \mathcal{R}, \mathcal{S} e nei riferimenti $\mathcal{R}', \mathcal{S}'$ sono legate dalla seguente formula:

$$M_{\mathcal{R},\mathcal{S}}^\beta = M_{\mathcal{R}',\mathcal{R}}^T M_{\mathcal{R}',\mathcal{S}'}^\beta M_{\mathcal{S}',\mathcal{S}}. \quad (1.4)$$

Infatti, per ogni $j = 1, \dots, m$ e per ogni $l = 1, \dots, n$, si ha

$$\beta(v_j, w_l) = \beta\left(\sum_{i=1}^m a_{ij}v'_i, \sum_{k=1}^n c_{kl}w'_k\right) = \sum_{j=1}^m \sum_{k=1}^n a_{ij}\beta(v'_i, w'_k)c_{kl} \quad (1.5)$$

dove abbiamo usato la bilinearità di β . Ma al primo membro della (1.5) c'è appunto l'elemento di posto (j, l) della matrice $M_{\mathcal{R}, \mathcal{S}}^\beta$, mentre all'ultimo membro c'è l'elemento di ugual posto di $M_{\mathcal{R}', \mathcal{R}}^T M_{\mathcal{R}', \mathcal{S}}^\beta M_{\mathcal{S}, \mathcal{S}}$ (il lettore lo vede?). L'uguaglianza (1.4) è dunque verificata.

1.2 Rango di una Forma Bilineare

La Formula (1.4) mostra, fra l'altro, che il rango della matrice rappresentativa di una forma bilineare non dipende dalla scelta dei riferimenti (infatti le matrici del cambio di coordinate sono invertibili, e il rango di una matrice non cambia se la si moltiplica, a destra o a sinistra, per una matrice invertibile). Proporranno tra poco una dimostrazione alternativa di questo fatto. Prima di fare ciò è necessario introdurre la nozione di *sottospazio annullatore*. Sia dunque $\beta : V \times W \rightarrow \mathbb{K}$ una forma bilineare sui \mathbb{K} -spazi vettoriali V, W .

Definizione 1.2.1 — Sottospazi Annullatori di una Forma Bilineare. Il *sottospazio annullatore sinistro* di β è il sottoinsieme

$$\overleftarrow{W}^0 := \{v \in V : \beta(v, w) = 0 \text{ per ogni } w \in W\} \subseteq V.$$

Similmente, il *sottospazio annullatore destro* di β è il sottoinsieme

$$\overrightarrow{V}^0 := \{w \in W : \beta(v, w) = 0 \text{ per ogni } v \in V\} \subseteq W.$$

Proposizione 1.2.1 I sottospazi annullatori sono sottospazi vettoriali.

Dimostrazione. Discutiamo l'annullatore sinistro. Per l'annullatore destro la dimostrazione procede in modo analogo. Innanzitutto il vettore nullo $0 \in V$ appartiene a \overleftarrow{W}^0 infatti (analogamente a quanto accade per le applicazioni lineari), per ogni $w \in W$,

$$\beta(0, w) = \beta(0 + 0, w) = \beta(0, w) + \beta(0, w) \implies \beta(0, w) = 0.$$

Resta da dimostrare che, se $v, v' \in \overleftarrow{W}^0$ e $a, a' \in \mathbb{K}$, allora $av + a'v' \in \overleftarrow{W}^0$. È così, infatti sia $w \in W$. Allora, dalla bilinearità di β ,

$$\beta(av + a'v', w) = a\beta(v, w) + a'\beta(v', w) = 0,$$

perché $v, v' \in \overleftarrow{W}^0$. Questo conclude la dimostrazione. ■

In virtù della Proposizione 1.2.1 ha senso parlare di dimensioni dei sottospazi annullatori.

Proposizione 1.2.2 Sia $\beta : V \times W \rightarrow \mathbb{K}$ una forma bilineare sui \mathbb{K} -spazi vettoriali V, W , sia $\mathcal{R} = (v_1, \dots, v_m)$ un riferimento di V , e sia $\mathcal{S} = (w_1, \dots, w_n)$ un riferimento di W . Indichiamo con $B = M_{\mathcal{R}, \mathcal{S}}^\beta = (b_{ij})_{i=1, \dots, m}^{j=1, \dots, n}$ la matrice rappresentativa di β nei riferimenti \mathcal{R}, \mathcal{S} . Allora

$$\dim V - \dim \overleftarrow{W}^0 = \dim W - \dim \overrightarrow{V}^0 = \text{rk } B, \quad (1.6)$$

il rango di B . In particolare, il rango della matrice rappresentativa di β non dipende dalla scelta

dei riferimenti.

Dimostrazione. Dimostriamo innanzitutto che $\dim W - \dim \vec{V}^0 = \text{rk } B$. A questo scopo osserviamo preliminarmente che un vettore $w \in W$ appartiene al sottospazio annullatore destro \vec{V}^0 sse

$$\beta(v_i, w) = 0 \quad \text{per ogni } i = 1, \dots, m = \dim V. \quad (1.7)$$

Infatti, se $w \in \vec{V}^0$ allora, in particolare, vale la (1.7). Viceversa, valga la (1.7) e sia $v \in V$. Allora v si può scrivere come combinazione lineare dei vettori v_1, \dots, v_m del riferimento \mathcal{R} : $v = \sum_{i=1}^m a_i v_i$ per qualche $a_1, \dots, a_m \in \mathbb{K}$, e dunque

$$\beta(v, w) = \beta\left(\sum_{i=1}^m a_i v_i, w\right) = \sum_{i=1}^m a_i \beta(v_i, w) = 0.$$

Dall'arbitrarietà di v segue che $w \in \vec{V}^0$. Ricapitolando, un vettore $w \in W$ appartiene al sottospazio annullatore destro \vec{V}^0 sse, per ogni $i = 1, \dots, m$,

$$0 = \beta(v_i, w) = x_i^T B y,$$

dove x_i, y sono i vettori coordinati di v_i, w nei riferimenti \mathcal{R}, \mathcal{S} rispettivamente. In particolare

$$x_i = \left(0, \dots, \underbrace{1}_{\text{posto } i}, \dots, 0\right)^T = E_i,$$

l' i -esimo vettore del riferimento canonico di \mathbb{K}^m . Perciò $x_i^T B = E_i^T B = (b_{i1}, \dots, b_{in}) = B_{(i)}$, l' i -esima riga di B , $i = 1, \dots, m$ (se occorre, verificare i dettagli per esercizio). Riassumendo, w appartiene a \vec{V}^0 sse il suo vettore coordinato è soluzione del seguente sistema lineare omogeneo di m equazioni in n incognite y_1, \dots, y_n :

$$S: \begin{cases} B_{(1)}y = 0 \\ \vdots \\ B_{(m)}y = 0 \end{cases},$$

la cui matrice dei coefficienti è B . La dimensione di \vec{V}^0 coincide perciò con la dimensione dello spazio delle soluzioni di S che è $n - \text{rk } B$, da cui $\dim W - \dim \vec{V}^0 = n - (n - \text{rk } B) = \text{rk } B$ come desiderato.

Per dimostrare che si ha anche $\dim V - \dim \overleftarrow{W}^0 = \text{rk } B$ consideriamo la forma bilineare

$$\beta^T : W \times V \rightarrow \mathbb{K}, \quad (w, v) \mapsto \beta^T(w, v) := \beta(v, w)$$

(il lettore vede che si tratta effettivamente di una forma bilineare?). Ovviamente l'annullatore destro di β^T coincide con l'annullatore sinistro di β . Inoltre la matrice rappresentativa di β^T nei riferimenti \mathcal{S}, \mathcal{R} è B^T . Segue allora dalla prima parte della dimostrazione applicata alla nuova forma bilineare β^T che

$$\dim V - \dim \overleftarrow{W}^0 = \text{rk } B^T = \text{rk } B.$$

Questo conclude la dimostrazione. ■

Definizione 1.2.2 — Rango di una Forma Bilineare. L'intero non negativo

$$\text{rk}\beta := \dim V - \dim \overleftarrow{W}^0 = \dim W - \dim \overrightarrow{V}^0$$

è detto *rango* della forma bilineare β (e coincide con il rango della matrice rappresentativa di β in una qualunque coppia di riferimenti). Una forma bilineare $\beta : V \times W \rightarrow \mathbb{K}$ si dice *non-degenere* se il rango di β è il massimo possibile cioè

$$\text{rk}\beta = \min\{\dim V, \dim W\}.$$

Sia $\beta : V \times W \rightarrow \mathbb{K}$ una forma bilineare non-degenere. Si osservi che, in questo caso, se $\dim V \leq \dim W$, allora il sottospazio annullatore sinistro è il sottospazio nullo di V , cioè l'unico vettore $v \in V$ tale che $\beta(v, w) = 0$ per ogni $w \in W$ è il vettore nullo $0 \in V$. Se invece $\dim W \leq \dim V$, il sottospazio annullatore destro è il sottospazio nullo di W . Infine, se $\dim V = \dim W$, allora i sottospazi annullatori sono entrambi nulli.

■ **Esempio 1.6** Si consideri la matrice reale

$$B = \begin{pmatrix} 2 & 1 \\ -1 & -1 \\ 0 & 1 \end{pmatrix} \in M_{3,2}(\mathbb{R}).$$

Come sappiamo, B definisce una forma bilineare

$$\beta_B : \mathbb{R}^3 \times \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto \beta_B(x, y) = x^T B y.$$

In altre parole

$$\beta_B \left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right) = x_1(2y_1 + y_2) - x_2(y_1 + y_2) + x_3 y_2.$$

La matrice rappresentativa di β_B nei riferimenti canonici di $\mathbb{R}^3, \mathbb{R}^2$ è esattamente B (Esempio 1.5) e segue dalla dimostrazione della Proposizione 1.2.2 che il sottospazio annullatore destro di β_B è lo spazio delle soluzioni del sistema lineare omogeneo $S : \{B y = 0$ cioè

$$S : \begin{cases} 2y_1 + y_2 = 0 \\ -y_1 - y_2 = 0 \\ y_2 = 0 \end{cases}.$$

L'unica soluzione di S è quella banale, dunque il sottospazio annullatore destro è il sottospazio banale $\{0\} \subseteq \mathbb{R}^2$. Ne consegue che il rango di β_B è

$$\dim \mathbb{R}^2 - \dim(\text{sottospazio annullatore destro di } \beta_B) = 2 - 0 = 2,$$

che coincide anche con il rango di B , come atteso. La forma bilineare β_B è non-degenere.

Sempre dalla dimostrazione della Proposizione 1.2.2 segue anche che il sottospazio annullatore sinistro di β_B è (il sottospazio annullatore destro di $\beta_B^T = \beta_{B^T}$ e dunque) lo spazio delle soluzioni del sistema lineare omogeneo $T : \{B^T x = 0$ cioè

$$T : \begin{cases} 2x_1 - x_2 = 0 \\ x_1 - x_2 + x_3 = 0 \end{cases},$$

Dalla prima equazione $x_2 = 2x_1$ che, sostituita nella seconda equazione, dà $x_3 = x_1$. Concludiamo che il sottospazio annullatore sinistro di β_B è il sottospazio unidimensionale

$$\left\{ \begin{pmatrix} t \\ 2t \\ t \end{pmatrix} : t \in \mathbb{R} \right\} \subseteq \mathbb{R}^3$$

il che è dovutamente consistente con la (1.6), infatti

$$\dim \mathbb{R}^3 - \dim(\text{sottospazio annullatore sinistro di } \beta_B) = 3 - 1 = 2 = \text{rk} \beta_B.$$

Esercizio 1.4 Si consideri l'applicazione

$$\beta : \mathbb{R}^3 \times \mathbb{R}^4 \rightarrow \mathbb{R}$$

data da

$$\beta \left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} \right) = 2x_1(y_1 + y_2) - (x_2 + x_3)(y_3 - y_4).$$

Dopo aver dimostrato che β è una forma bilineare, se ne calcoli

- (1) l'annullatore destro,
- (2) l'annullatore sinistro,
- (3) il rango.

La forma bilineare β è non-degenere?

1.3 Forme Bilineari Simmetriche

Siano V, W spazi vettoriali sul campo \mathbb{K} e sia $\beta : V \times W \rightarrow \mathbb{K}$ una forma bilineare. Se $V = W$, e $v, w \in V = W$, allora nell'espressione $\beta(v, w)$ ha senso scambiare le entrate v, w (altrimenti no!). Questa osservazione suggerisce la seguente

Definizione 1.3.1 — Forma Bilineare Simmetrica. Una forma bilineare $\beta : V \times V \rightarrow \mathbb{K}$ sullo spazio vettoriale V si dice *simmetrica* se

$$\beta(v, w) = \beta(w, v), \quad \text{per ogni } v, w \in V.$$

È facile costruire esempi di forme bilineari simmetriche grazie alla

Proposizione 1.3.1 — Matrice rappresentativa di una forma bilineare simmetrica. Sia V uno spazio vettoriale di dimensione n sul campo \mathbb{K} , e sia $\beta : V \times V \rightarrow \mathbb{K}$ una forma bilineare. Le seguenti condizioni sono equivalenti:

- (1) β è una forma bilineare simmetrica;
- (2) per ogni riferimento $\mathcal{R} = (v_1, \dots, v_n)$ di V , la matrice rappresentativa $M_{\mathcal{R}}^{\beta}$ è una matrice simmetrica (cioè coincide con la sua trasposta);
- (3) esiste un riferimento $\mathcal{R} = (v_1, \dots, v_n)$ di V tale che la matrice rappresentativa $M_{\mathcal{R}}^{\beta}$ è una matrice simmetrica.

Dimostrazione. Dimosteremo che (1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (1). Dimostriamo innanzitutto che (1) \Rightarrow (2). È facile, infatti sia $\mathcal{R} = (v_1, \dots, v_n)$ un qualunque riferimento di V , allora l'elemento di posto (i, j)

della matrice rappresentativa $M_{\mathcal{R}}^{\beta}$ è

$$b_{ij} = \beta(v_i, v_j) = \beta(v_j, v_i) = b_{ji},$$

in cui, nel secondo passaggio, abbiamo usato la simmetria di β . Questo mostra che la matrice $M_{\mathcal{R}}^{\beta}$ coincide con la sua trasposta, come richiesto. Che (2) \Rightarrow (3) è ovvio e resta da dimostrare che (3) \Rightarrow (1). Sia dunque \mathcal{R} un riferimento di V tale che la matrice $M_{\mathcal{R}}^{\beta} =: B$ è una matrice simmetrica. Consideriamo due vettori $v, w \in V$ e indichiamo con $x, y \in \mathbb{K}^n$ i rispettivi vettori coordinati nel riferimento \mathcal{R} . Allora

$$\beta(v, w) = x^T B y = (x^T B y)^T = y^T B^T x^T = y^T B x = \beta(w, v),$$

dove, nel secondo passaggio, abbiamo usato che $x^T B y$ è uno scalare e dunque, interpretato come una matrice di ordine 1×1 , coincide necessariamente con la sua matrice trasposta. ■

Sia $\beta: V \times V \rightarrow \mathbb{K}$ una forma bilineare (simmetrica) sullo spazio vettoriale V . La Formula (1.4), specializzata al caso $W = V$, $\mathcal{R} = \mathcal{S}$ e $\mathcal{R}' = \mathcal{S}'$ dice che

$$M_{\mathcal{R}}^{\beta} = M_{\mathcal{R}', \mathcal{R}}^T M_{\mathcal{R}'}^{\beta} M_{\mathcal{R}, \mathcal{R}'}. \quad (1.8)$$

Definizione 1.3.2 — Congruenza di Matrici. Due matrici quadrate di ordine n sul campo \mathbb{K} , $A, B \in M_n(\mathbb{K})$, si dicono *congruenti* se esiste una matrice invertibile $M \in GL_n(\mathbb{K})$ (il gruppo lineare generale di ordine n) tale che

$$B = M^T A M.$$

Le Formula (1.8) dice allora che due matrici rappresentative della stessa forma bilineare (simmetrica) sono necessariamente congruenti.

Proposizione 1.3.2 La congruenza di matrici è una relazione di equivalenza.

Dimostrazione. Siano $A, B, C \in M_n(\mathbb{K})$. Allora A è congruente a se stessa infatti, detta $I_n \in GL_n(\mathbb{K})$ la matrice identica, si ha $A = I_n A I_n = I_n^T A I_n$. Quindi la congruenza di matrici è una relazione riflessiva. Inoltre, se A e B sono congruenti, esiste una matrice invertibile $M \in GL_n(\mathbb{K})$ tale che

$$B = M^T A M. \quad (1.9)$$

Consideriamo la matrice inversa $M^{-1} \in GL_n(\mathbb{K})$. Sappiamo che anche M^T è una matrice invertibile e la sua inversa è $(M^{-1})^T$: $(M^T)^{-1} = (M^{-1})^T$. Moltiplicando la (1.9) a sinistra per $(M^T)^{-1}$ e a destra per M^{-1} troviamo

$$A = (M^T)^{-1} B M^{-1} = (M^{-1})^T B M^{-1}.$$

Questo mostra che B ed A sono congruenti. Quindi la congruenza di matrici è una relazione simmetrica. Infine supponiamo che A e B siano congruenti e che anche B e C siano congruenti. Questo vuol dire che esistono matrici invertibili $M, N \in GL_n(\mathbb{K})$ tali che

$$B = M^T A M, \quad \text{e} \quad C = N^T B N. \quad (1.10)$$

Sappiamo che anche il prodotto MN è una matrice invertibile e la sua trasposta è $N^T M^T$: $(MN)^T = N^T M^T$. Sostituendo la prima delle (1.10) nella seconda troviamo

$$C = N^T M^T A M N = (MN)^T A (MN).$$

Questo mostra che A e C sono congruenti. Quindi la congruenza di matrici è una relazione transitiva e ciò conclude la dimostrazione. ■

Definizione 1.3.3 — Forma Quadratica Associata ad una Forma Bilineare Simmetrica. La funzione

$$Q : V \rightarrow \mathbb{K}, \quad v \mapsto Q(v) := \beta(v, v) \quad (1.11)$$

è detta *forma quadratica associata a β* .

La forma quadratica Q associata ad una forma bilineare simmetrica β gode chiaramente della seguente proprietà:

$$Q(av) = a^2 Q(v),$$

per ogni $v \in V$ e $a \in \mathbb{K}$, proprietà che segue immediatamente dalla bilinearità di β (il lettore lo vede?). Sotto appropriate ipotesi sul campo \mathbb{K} , la forma quadratica Q consente di ricostruire la forma bilineare β .

○ Sia \mathbb{K} un campo. La *caratteristica* di \mathbb{K} è il più piccolo numero naturale n tale che

$$\underbrace{1 + \cdots + 1}_{n \text{ volte}} = 0.$$

Se tale numero naturale non esiste, allora la caratteristica di \mathbb{K} è 0 per definizione. Si può dimostrare che la caratteristica di un campo è un numero primo. Se p è un numero primo allora, in un campo di caratteristica diversa da $p \in \mathbb{N}$, lo scalare

$$\underbrace{1 + \cdots + 1}_{p \text{ volte}}$$

(che talvolta, con abuso di notazione, indichiamo con lo stesso simbolo p) è diverso da 0 e perciò è invertibile. In tal caso, il suo inverso si indica con $1/p$.

Proposizione 1.3.3 — Formula di Polarizzazione. Sia $\beta : V \times V \rightarrow \mathbb{K}$ una forma bilineare simmetrica e sia $Q : V \rightarrow \mathbb{K}$ la forma quadratica associata. Se la caratteristica di \mathbb{K} è diversa da 2, allora vale la seguente *Formula di Polarizzazione*: per ogni $v, w \in V$

$$\beta(v, w) = \frac{Q(v+w) - Q(v) - Q(w)}{2}. \quad (1.12)$$

In particolare, la forma bilineare simmetrica β è completamente determinata dalla forma quadratica associata.

Dimostrazione. Siano $v, w \in V$. Calcoliamo

$$\begin{aligned} Q(v+w) &= \beta(v+w, v+w) = \beta(v, v) + \beta(v, w) + \beta(w, v) + \beta(w, w) \\ &= Q(v) + 2\beta(v, w) + Q(w), \end{aligned} \quad (1.13)$$

dove abbiamo usato sia la bilinearità che la simmetria di β . Ora, per ipotesi, la caratteristica del campo \mathbb{K} è diversa da 2, sicché $2 := 1 + 1 \in \mathbb{K}$ è uno scalare invertibile e, dalla (1.13) segue la (1.12). ■

Consideriamo di nuovo una forma bilineare simmetrica $\beta : V \times V \rightarrow \mathbb{K}$. Ovviamente, in questo caso, lo spazio annullatore destro e lo spazio annullatore sinistro di β coincidono e nel seguito saranno entrambi indicati semplicemente con V^0 .

La terminologia nella seguente definizione sarà chiarita nel prossimo capitolo.

Definizione 1.3.4 — Vettori Ortogonali rispetto ad una Forma Bilineare Simmetrica. Due vettori $v, w \in V$ sono *ortogonali rispetto a β* se $\beta(v, w) = 0$.

Nei termini della Definizione 1.3.4, il sottospazio annullatore V^0 consiste dei vettori v che sono ortogonali (rispetto a β) ad ogni altro vettore di V .

Definizione 1.3.5 — Riferimento Ortogonale rispetto ad una Forma Bilineare Simmetrica. Un riferimento $\mathcal{R} = (v_1, \dots, v_n)$ di V è *ortogonale rispetto a β* se i vettori di \mathcal{R} sono a due a due ortogonali, cioè $\beta(v_i, v_j) = 0$ per ogni $i \neq j$.

In altre parole, un riferimento ortogonale è un riferimento \mathcal{R} tale che la matrice rappresentativa $M_{\mathcal{R}}^{\beta}$ di β è una matrice diagonale. Si noti che, in questo caso, sulla diagonale principale di $M_{\mathcal{R}}^{\beta}$ compaiono gli scalari $Q_1 := Q(v_1), \dots, Q_n := Q(v_n)$:

$$M_{\mathcal{R}}^{\beta} = \begin{pmatrix} Q_1 & 0 & \cdots & 0 \\ 0 & Q_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & Q_n \end{pmatrix}.$$

Da qui l'utilità dei riferimenti ortogonali: in coordinate, rispetto ad un riferimento ortogonale, una forma bilineare simmetrica è particolarmente semplice: per ogni $v, w \in V$,

$$\beta(v, w) = Q_1 x_1 y_1 + \cdots + Q_n x_n y_n,$$

dove $x = (x_1, \dots, x_n)^T$ e $y = (y_1, \dots, y_n)^T$ sono i vettori coordinati di v e w nel riferimento ortogonale $\mathcal{R} = (v_1, \dots, v_n)$. Ne consegue anche che

$$Q(v) = Q_1 x_1^2 + \cdots + Q_n x_n^2.$$

Possiamo garantire l'esistenza di un riferimento ortogonale almeno nel caso in cui la caratteristica del campo \mathbb{K} sia diversa da 2, in base al seguente

Teorema 1.3.4 — Esistenza di Riferimenti Ortogonali. Sia V uno spazio vettoriale sul campo \mathbb{K} e sia $\beta : V \times V \rightarrow \mathbb{K}$ una forma bilineare simmetrica su V . Se la caratteristica di \mathbb{K} è diversa da 2, allora esiste un riferimento di V ortogonale rispetto a β .

Dimostrazione. La dimostrazione è per induzione sulla dimensione n di V . Se $n = 1$ ogni riferimento di V è ortogonale e non c'è nulla da dimostrare. Supponiamo ora che l'asserto sia vero per $n = k$ e dimostriamo che è vero anche per $n = k + 1$. Sia dunque V un \mathbb{K} -spazio vettoriale di dimensione $k + 1$ e sia $\beta : V \times V \rightarrow \mathbb{K}$ una forma bilineare simmetrica su V . È opportuno, a questo punto, introdurre una definizione: un vettore $u \in V$ si dice *isotropo* rispetto a β se $Q(u) = \beta(u, u) = 0$ (in altre parole, u è ortogonale a se stesso). Se $\beta = 0$, allora ogni riferimento di V è ortogonale. Se invece $\beta \neq 0$, vuol dire che esistono $v, w \in V$ non nulli e tali che $\beta(v, w) \neq 0$. Si considerino i tre vettori $v, w, v + w$. Dei tre, almeno uno non è isotropo. Infatti se, per assurdo, $v, w, v + w$ fossero tutti e tre isotropi, allora, dalla Formula di Polarizzazione (1.12) seguirebbe $\beta(v, w) = 0$ il che non è. Si noti che qui abbiamo usato l'ipotesi che la caratteristica del campo \mathbb{K} degli scalari sia diversa da 2. Questa breve discussione mostra che, se $\beta \neq 0$, allora esiste un vettore $v_0 \in V$ non isotropo. Consideriamo ora il sottoinsieme

$$W := \{w \in V : w \text{ è ortogonale a } v_0\} \subseteq V.$$

Si tratta di un sottospazio vettoriale, infatti, innanzitutto contiene chiaramente il vettore nullo, inoltre se $w_1, w_2 \in W$ e $a_1, a_2 \in \mathbb{K}$ allora

$$\beta(a_1 w_1 + a_2 w_2, v_0) = a_1 \beta(w_1, v_0) + a_2 \beta(w_2, v_0) = 0,$$

dove abbiamo usato la bilinearità di β insieme al fatto che w_1, w_2 sono ortogonali a v_0 . Dunque anche $a_1 w_1 + a_2 w_2$ è ortogonale a v_0 , cioè $a_1 w_1 + a_2 w_2 \in W$ come desiderato. Il prossimo passo della dimostrazione consiste nel dimostrare che W è un sottospazio di dimensione k . A questo scopo osserviamo che $v_0 \notin W$ altrimenti si avrebbe $Q(v_0) = \beta(v_0, v_0) = 0$ cioè v_0 sarebbe un vettore isotropo. Per lo stesso motivo $av_0 \notin W$ per alcuno scalare non nullo $a \in \mathbb{K}$. Dunque W e lo spazio $W_0 := \text{Span}(v_0)$ generato da v_0 sono in somma diretta:

$$W + W_0 = W \oplus W_0.$$

Dimostriamo ora che $W + W_0 = V$: il fatto che la dimensione di W è k seguirà dalla Relazione di Grassmann $\dim W + \dim W_0 = \dim V$. Consideriamo dunque un generico vettore $v \in V$. Ovviamente

$$v = \left(v - \frac{\beta(v_0, v)}{\beta(v_0, v_0)} v_0 \right) + \frac{\beta(v_0, v)}{\beta(v_0, v_0)} v_0. \quad (1.14)$$

Si noti che la combinazione lineare a secondo membro è ben definita perché $\beta(v_0, v_0) \neq 0$. Inoltre il primo addendo

$$w := v - \frac{\beta(v_0, v)}{\beta(v_0, v_0)} v_0$$

appartiene a W , infatti

$$\beta(v_0, w) = \beta \left(v_0, v - \frac{\beta(v_0, v)}{\beta(v_0, v_0)} v_0 \right) = \beta(v_0, v) - \frac{\beta(v_0, v)}{\beta(v_0, v_0)} \beta(v_0, v_0) = \beta(v_0, v) - \beta(v_0, v) = 0.$$

Il secondo addendo

$$\frac{\beta(v_0, v)}{\beta(v_0, v_0)} v_0$$

della (1.14) appartiene a W_0 . Dunque ogni vettore v di V si può scrivere come somma di un vettore in W e di un vettore in W_0 . Questo mostra che $V = W + W_0$ come desiderato. Ne consegue che $\dim W = k$ come annunciato.

Ora possiamo *restringere* la forma bilineare β a W definendo una nuova applicazione

$$\beta_W : W \times W \rightarrow \mathbb{K}, \quad (w_1, w_2) \mapsto \beta_W(w_1, w_2) := \beta(w_1, w_2).$$

È facile mostrare che β_W è una forma bilineare simmetrica su W (il lettore lo verifichi per esercizio). Inoltre, poiché β_W agisce come β (ma solo sui vettori di W), due vettori di $W \subseteq V$ sono ortogonali rispetto a β_W sse sono ortogonali rispetto a β (il lettore lo vede?). A β_W si applica l'ipotesi di induzione, dunque esiste un riferimento (v_1, \dots, v_k) di W ortogonale rispetto a β_W . Ma allora, il sistema $\mathcal{R} := (v_0, v_1, \dots, v_k)$ è un riferimento di V ortogonale rispetto a β . Infatti \mathcal{R} è un riferimento di V perché è un sistema di $k+1$ vettori linearmente indipendenti: (v_1, \dots, v_k) sono linearmente indipendenti e $v_0 \notin \text{Span}(v_1, \dots, v_k) = W$. Inoltre i vettori di \mathcal{R} sono a due a due ortogonali perché v_0 è ortogonale a v_i per ogni $i = 1, \dots, k$ giacché $v_1, \dots, v_k \in W$, e inoltre v_1, \dots, v_k sono ortogonali tra loro rispetto a β_W e dunque rispetto a β . ■

■ **Esempio 1.7** Sullo spazio vettoriale numerico tridimensionale \mathbb{R}^3 consideriamo la forma bilineare simmetrica

$$\beta : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$$

data da

$$\beta \left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \right) = 2(x_1 + x_2)(y_1 + y_2) - x_3 y_3$$

(il lettore verifichi per esercizio che β è effettivamente una forma bilineare simmetrica). La matrice rappresentativa di β nel riferimento canonico $\mathcal{R}_{can} = (E_1, E_2, E_3)$ è

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

e la forma quadratica $Q: \mathbb{R}^3 \rightarrow \mathbb{R}$ associata a β è data da

$$Q \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 2(x_1 + x_2)^2 - x_3^2.$$

Vogliamo trovare un riferimento di \mathbb{R}^3 ortogonale rispetto a β . Il riferimento canonico non è ortogonale infatti

$$\beta(E_1, E_2) = 2 \neq 0.$$

Per costruire un riferimento ortogonale (v_0, v_1, v_2) possiamo ragionare come nella dimostrazione del Teorema 1.3.4. Osserviamo innanzitutto che $v_0 := E_1$ non è un vettore isotropo. Infatti

$$\beta(E_1, E_1) = 2 \neq 0.$$

Un vettore $y = (y_1, y_2, y_3)^T$ è ortogonale a v_0 sse

$$0 = \beta(E_1, y) = 2(y_1 + y_2)$$

o, equivalentemente, $y_2 = -y_1$. Dunque il sottospazio vettoriale $W \subseteq \mathbb{R}^3$ dei vettori ortogonali a v_0 è il sottospazio bidimensionale dato da:

$$W = \left\{ \begin{pmatrix} s \\ -s \\ t \end{pmatrix} : s, t \in \mathbb{R} \right\}.$$

Un riferimento di W è costituito dai vettori

$$\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, E_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Ora dobbiamo trovare un riferimento di W ortogonale rispetto alla forma bilineare simmetrica ristretta β_W . Osserviamo che il vettore

$$\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \in W$$

è isotropo rispetto a β e quindi anche rispetto a β_W . Invece

$$\beta_W(E_3, E_3) = \beta(E_3, E_3) = 1 \neq 0,$$

cioè $v_1 := E_3$ è un vettore non isotropo di W . Un vettore $w = (s, -s, t)^T \in W$ è ortogonale a v_1 rispetto a β_W sse è ortogonale a v_1 rispetto a β , cioè

$$0 = \beta(v_1, w) = -t.$$

Dunque il sottospazio vettoriale $W' \subseteq W$ dei vettori ortogonali a v_1 è il sottospazio 1-dimensionale dato da

$$W' = \left\{ \begin{pmatrix} s \\ -s \\ 0 \end{pmatrix} : s \in \mathbb{R} \right\},$$

il quale è generato dal vettore

$$v_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

Dunque i vettori v_1, v_2 formano un riferimento di W ortogonale di rispetto a β_W (si noti che avremmo potuto verificare quest'ultimo fatto molto prima, semplicemente calcolando $\beta(v_1, v_2)$). Utilizzando la dimostrazione del Teorema 1.3.4 ne deduciamo che

$$\mathcal{R} = (v_0, v_1, v_2) = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right)$$

è un riferimento di \mathbb{R}^3 ortogonale rispetto a β . La matrice rappresentativa di β nel riferimento \mathcal{R} è la matrice diagonale che ha sulla diagonale principale gli scalari $Q(v_0) = 2$, $Q(v_1) = 1$ e $Q(v_2) = 0$, cioè

$$M_{\mathcal{R}}^{\beta} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

■

Esercizio 1.5 Si consideri l'applicazione

$$\beta : \mathbb{R}^4 \times \mathbb{R}^4 \rightarrow \mathbb{R}$$

data da

$$\beta \left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} \right) = (x_1 - x_2)(y_1 - y_2) + x_3y_2 + x_2y_3 - x_4y_3 - x_3y_4.$$

Dopo aver dimostrato che β è una forma bilineare simmetrica, si determini un riferimento $\mathcal{R} = (v_0, v_1, v_2, v_3)$ di \mathbb{R}^4 ortogonale rispetto a β . Qual'è la matrice rappresentativa di β nel riferimento \mathcal{R} ? ■

Corollario 1.3.5 Ogni matrice simmetrica su un campo \mathbb{K} di caratteristica diversa da 2 è congruente ad una matrice diagonale.

Dimostrazione. Sia $B \in M_n(\mathbb{K})$ una matrice simmetrica. Consideriamo la forma bilineare simmetrica

$$\beta_B : \mathbb{K}^n \times \mathbb{K}^n \rightarrow \mathbb{K}.$$

In virtù del Teorema 1.3.4 esiste un riferimento $\mathcal{R} = (M^{(1)}, \dots, M^{(n)})$ di \mathbb{K}^n ortogonale rispetto a β_B .

In altre parole la matrice rappresentativa $D = M_{\mathcal{R}}^{\beta_B}$ è una matrice diagonale. Ma, ricordando che

- (1) la matrice rappresentativa di β_B nel riferimento canonico \mathcal{R}_{can} è proprio B ,
- (2) la matrice del cambio di coordinate dal riferimento \mathcal{R} al riferimento canonico è proprio la matrice M che ha i vettori numerici $M^{(1)}, \dots, M^{(n)}$ come colonne,

in vista della Formula (1.8) troviamo

$$D = M_{\mathcal{R}}^{\beta_B} = M_{\mathcal{R}_{can}, \mathcal{R}}^T M_{\mathcal{R}_{can}}^{\beta_B} M_{\mathcal{R}_{can}, \mathcal{R}} = M^T B M,$$

cioè B e D sono matrici congruenti. ■

■ **Esempio 1.8** Consideriamo la matrice reale simmetrica

$$B = \begin{pmatrix} 0 & 1 & 0 & -\frac{1}{2} \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -\frac{1}{2} & 0 & 1 & 0 \end{pmatrix} \in M_4(\mathbb{R}).$$

Vogliamo trovare una matrice invertibile $M \in GL_4(\mathbb{R})$ tale che $M^T B M$ sia una matrice diagonale. A questo scopo procediamo come nella dimostrazione del Corollario 1.3.5. Cominciamo con il considerare la forma bilineare simmetrica

$$\beta_B : \mathbb{R}^4 \times \mathbb{R}^4 \rightarrow \mathbb{R}, \quad (x, y) \mapsto x^T B y.$$

Ricordiamo che la matrice rappresentativa di β_B nel riferimento canonico è esattamente B . La forma quadratica associata $Q : \mathbb{R}^4 \rightarrow \mathbb{R}$ è data da:

$$Q \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = 2(x_1 x_2 + x_2 x_3 + x_3 x_4) - x_1 x_4.$$

Costruiamo un riferimento di \mathbb{R}^4 ortogonale rispetto a β_B . I vettori E_1, E_2, E_3, E_4 del riferimento canonico di \mathbb{R}^4 sono tutti isotropi rispetto a β . D'altronde $\beta_B(E_1, E_2) = 1$ (l'elemento di posto (1, 2) di B), perciò il vettore

$$v_0 := E_1 + E_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

non è isotropo (dimostrazione del Teorema 1.3.4). Un vettore $y = (y_1, y_2, y_3, y_4)^T$ è ortogonale a v_0 sse

$$0 = \beta(v_0, y) = y_1 + y_2 + y_3 - \frac{1}{2}y_4$$

o, equivalentemente $y_4 = 2(y_1 + y_2 + y_3)$. Dunque il sottospazio vettoriale $W \subseteq \mathbb{R}^4$ dei vettori ortogonali a v_0 è il sottospazio tridimensionale dato da

$$W = \left\{ \begin{pmatrix} r \\ s \\ t \\ 2(r+s+t) \end{pmatrix} : r, s, t \in \mathbb{R} \right\}.$$

Un riferimento di W è costituito dai vettori

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \end{pmatrix}.$$

Il prossimo passo è trovare un riferimento di W ortogonale rispetto alla forma bilineare simmetrica ristretta, che indichiamo con β_W . Il vettore

$$v_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \end{pmatrix}$$

non è isotropo. Infatti

$$Q(v_1) = Q \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \end{pmatrix} = -2 \neq 0.$$

Un vettore $w = (r, s, t, 2(r+s+t))^T \in W$ è ortogonale a v_1 rispetto a β_W sse è ortogonale a v_1 rispetto a β_B cioè

$$0 = \beta_B(v_1, w) = \beta \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} r \\ s \\ t \\ 2(r+s+t) \end{pmatrix} \right) = -2r + t,$$

o, equivalentemente, $t = 2r$. Dunque il sottospazio $W' \subseteq W$ dei vettori ortogonali a v_1 è il sottospazio 2-dimensionale dato da

$$W' = \left\{ \begin{pmatrix} r \\ s \\ 2r \\ 6r+2s \end{pmatrix} : r, s \in \mathbb{R} \right\}.$$

Un riferimento di W' è costituito dai vettori

$$\begin{pmatrix} 1 \\ 0 \\ 2 \\ 6 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 2 \end{pmatrix}.$$

Dobbiamo trovare un riferimento di W' ortogonale rispetto alla forma bilineare ristretta. Il vettore

$$v_2 := \begin{pmatrix} 1 \\ 0 \\ 2 \\ 6 \end{pmatrix} \in W'$$

non è isotropo, infatti $Q(v_2) = 18 \neq 0$. Un vettore $w' = (r, s, 2r, 6r + 2s)^T \in W'$ è ortogonale a v_2 sse

$$0 = \beta_B(v_2, w') = 18r + 6s$$

o, equivalentemente, $s = -3r$. Dunque il sottospazio $W'' \subseteq W'$ dei vettori ortogonali a v_2 è il sottospazio 1-dimensionale dato da

$$W'' = \left\{ \begin{pmatrix} r \\ -3r \\ 2r \\ 0 \end{pmatrix} : r \in \mathbb{R} \right\}$$

il quale è generato dal vettore

$$v_3 = \begin{pmatrix} 1 \\ -3 \\ 2 \\ 0 \end{pmatrix}.$$

Concludendo,

$$\mathcal{R} = (v_0, v_1, v_2, v_3) = \left(\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \\ 6 \end{pmatrix}, \begin{pmatrix} 1 \\ -3 \\ 2 \\ 0 \end{pmatrix} \right) \quad (1.15)$$

è un riferimento di \mathbb{R}^4 ortogonale rispetto a β_B . La matrice rappresentativa di β_B nel riferimento \mathcal{R} è

$$D = \begin{pmatrix} Q(v_0) & 0 & 0 & 0 \\ 0 & Q(v_1) & 0 & 0 \\ 0 & 0 & Q(v_2) & 0 \\ 0 & 0 & 0 & Q(v_3) \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 18 & 0 \\ 0 & 0 & 0 & -18 \end{pmatrix}.$$

A questo punto segue dalla dimostrazione del Corollario 1.3.5 che la matrice M che stiamo cercando è esattamente la matrice che ha v_0, v_1, v_2, v_3 come colonne, cioè

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

Senza fare ulteriori conti, deduciamo che

$$M^T B M = D.$$

■

Esercizio 1.6 Si consideri la matrice reale simmetrica

$$B = \begin{pmatrix} 0 & -1 & -1 & -1 \\ -1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 1 \\ -1 & 0 & 1 & 1 \end{pmatrix} \in M_4(\mathbb{R}).$$

Si determini una matrice invertibile $M \in GL_4(\mathbb{R})$ tale che $M^T B M$ è una matrice diagonale. ■

Nel caso di una forma bilineare simmetrica β su uno spazio vettoriale *reale* un riferimento ortogonale può essere “modificato” in modo da rendere la matrice rappresentativa di β ancora più “semplice”. Sia dunque V uno spazio vettoriale di dimensione m sul campo \mathbb{R} dei numeri reali e sia $\beta : V \times V \rightarrow \mathbb{R}$ una forma bilineare simmetrica su V .

Definizione 1.3.6 — Riferimento Ortonormale. Un riferimento $\mathcal{R} = (v_1, \dots, v_m)$ di V si dice *ortonormale* rispetto a β se è ortogonale e inoltre

$$Q(v_i) = \pm 1, 0$$

per ogni $i = 1, \dots, m$.

Proposizione 1.3.6 Sia V uno spazio vettoriale sul campo \mathbb{R} dei numeri reali e sia $\beta : V \times V \rightarrow \mathbb{R}$ una forma bilineare simmetrica su V . Allora esiste un riferimento di V ortonormale rispetto a β .

Dimostrazione. Sia $m = \dim V$ e sia $\tilde{\mathcal{R}} = (\tilde{v}_1, \dots, \tilde{v}_m)$ un riferimento di V ortogonale rispetto a β ($\tilde{\mathcal{R}}$ esiste in virtù del Teorema 1.3.4). Per ogni $i = 1, \dots, m$ poniamo

$$\alpha_i = \begin{cases} \frac{1}{\sqrt{|Q(\tilde{v}_i)|}} & \text{se } Q(\tilde{v}_i) \neq 0 \\ 1 & \text{se } Q(\tilde{v}_i) = 0 \end{cases}.$$

Consideriamo il sistema $\mathcal{R} = (v_1, \dots, v_m)$ definito da

$$v_i = \alpha_i \tilde{v}_i, \quad i = 1, \dots, m.$$

È chiaro che \mathcal{R} è ancora un riferimento di V . Inoltre è un riferimento ortonormale, infatti, per ogni $i \neq j$,

$$\beta(v_i, v_j) = \beta(\alpha_i \tilde{v}_i, \alpha_j \tilde{v}_j) = \alpha_i \alpha_j \beta(\tilde{v}_i, \tilde{v}_j) = 0.$$

Inoltre, se $Q(\tilde{v}_i) \neq 0$, allora

$$Q(v_i) = Q(\alpha_i \tilde{v}_i) = \alpha_i^2 Q(\tilde{v}_i) = \frac{Q(\tilde{v}_i)}{|Q(\tilde{v}_i)|} = \pm 1,$$

mentre, se $Q(\tilde{v}_i) = 0$, allora

$$Q(v_i) = Q(\tilde{v}_i) = 0.$$

■

In altre parole, un riferimento ortonormale è un riferimento $\mathcal{R} = (v_1, \dots, v_m)$ nel quale la matrice rappresentativa N di β è una matrice diagonale in cui, sulla diagonale principale, compaiono solo i

numeri $\pm 1, 0$. Riordinando opportunamente i vettori di \mathcal{R} , se necessario, è sempre possibile fare in modo che N sia del tipo

$$N = \begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0_n \end{pmatrix}, \quad (1.16)$$

dove $0_n \in M_n(\mathbb{R})$ è la matrice nulla di ordine $n \times n$ e $p + q + n = m$ (è sufficiente mettere ai primi posti i p vettori v_i di \mathcal{R} tali che $Q(v_i) = +1$, ai successivi posti i q vettori v_i tali che $Q(v_i) = -1$ e agli ultimi posti gli n vettori v_i tali che $Q(v_i) = 0$; il lettore lo vede?). Nel seguito, dato un riferimento ortonormale \mathcal{R} , riordereremo (quasi) sempre i vettori di \mathcal{R} in questo modo.

Definizione 1.3.7 La terna (p, q, n) costituita dagli interi non negativi che definiscono l'ordine dei blocchi nella matrice rappresentativa (1.16) della forma bilineare simmetrica β prende il nome di *segnatura* di β . L'intero n è la *nullità* di β . La matrice (1.16) è detta la *forma normale* di β (e si dice anche che, nel riferimento \mathcal{R} in questione, β è *in forma normale*).

Teorema 1.3.7 — Legge di Inerzia di Sylvester. Sia V uno spazio vettoriale di dimensione m sul campo \mathbb{R} dei numeri reali e sia $\beta : V \times V \rightarrow \mathbb{R}$ una forma bilineare simmetrica su V . La segnatura di β non dipende dalla scelta del riferimento ortonormale (in altre parole, tutti i riferimenti ortonormali definiscono la stessa segnatura).

Dimostrazione. Il rango della matrice (1.16) è $p + q$. Dunque la nullità è

$$n = \dim V - \text{rk } B = \dim V^0,$$

la dimensione del sottospazio annullatore, e non dipende dalla scelta del riferimento. Ora siano

$$\begin{aligned} \mathcal{R} &= (v_1, \dots, v_p, v_{p+1}, \dots, v_{p+q}, v_{p+q+1}, \dots, v_m) \\ \bar{\mathcal{R}} &= (\bar{v}_1, \dots, \bar{v}_{\bar{p}}, \bar{v}_{\bar{p}+1}, \dots, \bar{v}_{\bar{p}+\bar{q}}, \bar{v}_{\bar{p}+\bar{q}+1}, \dots, \bar{v}_m) \end{aligned}$$

riferimenti ortonormali e, più precisamente, riferimenti ortogonali tali che

$$\begin{aligned} Q(v_i) &= 1, & \text{se } i &= 1, \dots, p; \\ Q(v_i) &= -1, & \text{se } i &= p+1, \dots, p+q; \\ Q(v_i) &= 0, & \text{se } i &= p+q+1, \dots, m; \end{aligned}$$

e analogamente

$$\begin{aligned} Q(\bar{v}_i) &= 1, & \text{se } i &= 1, \dots, \bar{p}; \\ Q(\bar{v}_i) &= -1, & \text{se } i &= \bar{p}+1, \dots, \bar{p}+\bar{q}; \\ Q(\bar{v}_i) &= 0, & \text{se } i &= \bar{p}+\bar{q}+1, \dots, m. \end{aligned}$$

Poiché le nullità definite da \mathcal{R} e $\bar{\mathcal{R}}$ coincidono, abbiamo

$$m - p - q = m - \bar{p} - \bar{q} \quad \Rightarrow \quad p + q = \bar{p} + \bar{q}. \quad (1.17)$$

Ora vogliamo dimostrare che $p = \bar{p}$ e $q = \bar{q}$. Ma, dalla (1.17), è sufficiente dimostrare che $p = \bar{p}$. A questo scopo consideriamo il sistema di vettori

$$\mathcal{S} = (\bar{v}_1, \dots, \bar{v}_{\bar{p}}, v_{p+1}, \dots, v_{p+q}, v_{p+q+1}, \dots, v_m)$$

e mostriamo che \mathcal{S} è linearmente indipendente. Sia dunque

$$\sum_{i=1}^{\bar{p}} a_i \bar{v}_i + \sum_{i=p+1}^{p+q} b_i v_i + \sum_{i=p+q+1}^m c_i v_i = 0 \quad (1.18)$$

una combinazione lineare nulla dei vettori di \mathcal{S} . Questo vuol dire che

$$v := \sum_{i=1}^{\bar{p}} a_i \bar{v}_i = - \sum_{i=p+1}^{p+q} b_i v_i - \sum_{i=p+q+1}^m c_i v_i.$$

Perciò $\alpha := \beta(v, v)$ può essere calcolato in due modi:

$$\alpha = \beta \left(\sum_{i=1}^{\bar{p}} a_i \bar{v}_i, \sum_{j=1}^{\bar{p}} a_j \bar{v}_j \right) = \sum_{i,j=1}^{\bar{p}} a_i a_j \beta(\bar{v}_i, \bar{v}_j) = \sum_{i,j=1}^{\bar{p}} a_i a_j \delta_{ij} = \sum_{i=1}^{\bar{p}} a_i^2 \geq 0,$$

ma anche

$$\begin{aligned} \alpha &= \beta \left(- \sum_{i=p+1}^{p+q} b_i v_i - \sum_{i=p+q+1}^m c_i v_i, - \sum_{j=p+1}^{p+q} b_j v_j - \sum_{j=p+q+1}^m c_j v_j \right) \\ &= \sum_{i,j=p+1}^{p+q} b_i b_j \beta(v_i, v_j) + \sum_{i=p+1}^{p+q} \sum_{j=p+q+1}^m b_i c_j \beta(v_i, v_j) \\ &\quad + \sum_{i=p+q+1}^m \sum_{j=p+1}^{p+q} c_i b_j \beta(v_i, v_j) + \sum_{i,j=p+q+1}^m c_i c_j \beta(v_i, v_j) \\ &= - \sum_{i,j=p+1}^{p+q} b_i b_j \delta_{ij} = - \sum_{i=p+1}^{p+q} b_i^2, \end{aligned}$$

dove abbiamo usato, fra l'altro, che vettori diversi di \mathcal{R} sono ortogonali e che i vettori v_{p+q+1}, \dots, v_m sono anche isotropi. Poiché $\alpha \geq 0$, ma anche $\alpha \leq 0$, allora $\alpha = 0$, cioè

$$\sum_{i=1}^{\bar{p}} a_i^2 = 0 \quad \Rightarrow \quad a_i = 0 \quad \text{per ogni } i,$$

ma anche

$$- \sum_{j=p+1}^{p+q} b_j^2 = 0 \quad \Rightarrow \quad b_j = 0 \quad \text{per ogni } j.$$

Infine, usando $a_i = b_j = 0$ nella (1.18) troviamo

$$\sum_{i=p+q+1}^m c_i v_i = 0.$$

Ma i vettori v_{p+q+1}, \dots, v_m sono linearmente indipendenti e, perciò, anche $c_i = 0$ per ogni i . Dunque \mathcal{S} è un sistema linearmente indipendente e, per il Lemma di Steinitz, ha un numero di elementi minore o uguale alla dimensione di V cioè

$$\bar{p} + q + (m - p - q) \leq m \quad \Rightarrow \quad \bar{p} \leq p.$$

Scambiando il ruolo di \mathcal{R} e $\bar{\mathcal{R}}$ si dimostra che è anche $p \leq \bar{p}$, perciò $p = \bar{p}$ e questo conclude la dimostrazione. ■

Sottolineamo, dalla dimostrazione del Teorema 1.3.7, che la nullità della forma bilineare simmetrica β , coincide con la dimensione del sottospazio annullatore V^0 . Osserviamo inoltre che, se $\mathcal{R} = (v_1, \dots, v_m)$ è un riferimento ortonormale ordinato come in (1.16), allora gli ultimi n vettori $v_{m-n+1} = v_{m-p-q+1}, \dots, v_m$ formano una base di V^0 . Infatti sono ortogonali a tutti i vettori di \mathcal{R} e perciò sono ortogonali ad ogni vettore di V e dunque appartengono a V_0 (si veda, a questo riguardo, la dimostrazione della Proposizione 1.2.2). Inoltre sono linearmente indipendenti e sono proprio tanti quanti è la dimensione di V_0 .

■ **Esempio 1.9** La forma bilineare simmetrica β dell'Esempio 1.7 ha segnatura $(1, 1, 1)$. Un riferimento di \mathbb{R}^3 ortonormale rispetto a β è

$$\left(\frac{1}{\sqrt{Q(v_0)}} v_0, \frac{1}{\sqrt{Q(v_1)}} v_1, v_2 \right) = \left(\left(\begin{array}{c} \frac{\sqrt{2}}{2} \\ 0 \\ 0 \end{array} \right), \left(\begin{array}{c} 0 \\ 0 \\ 1 \end{array} \right), \left(\begin{array}{c} 1 \\ -1 \\ 0 \end{array} \right) \right).$$

La forma normale di β è la matrice

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

Il vettore

$$v_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

genera il sottospazio annullatore. ■

○ Sia V uno spazio vettoriale reale di dimensione m , sia $\beta : V \times V \rightarrow \mathbb{R}$ una forma bilineare simmetrica non nulla su V , e sia (p, q, n) la segnatura di β . Si dice che β è

- ✓ *semidefinita* se $n \neq 0$, e $q = 0$ oppure $p = 0$; più precisamente,
- ✓ *semidefinita positiva* se $n \neq 0$ e $q = 0$ (questo si verifica sse $Q(v) \geq 0$ per ogni $v \in V$, ma esistono vettori isotropi non banali);
- ✓ *semidefinita negativa* se $n \neq 0$ e $p = 0$ (questo si verifica sse $Q(v) \leq 0$ per ogni $v \in V$, ma esistono vettori isotropi non banali);
- ✓ *definita* se $n = q = 0$ oppure $n = p = 0$ (e, in questo caso, β è non degenera); più precisamente,
- ✓ *definita positiva* se $n = q = 0$ (questo si verifica sse $Q(v) \geq 0$ per ogni $v \in V$ e $Q(v) = 0$ solo se $v = 0$ e, in questo caso, β è non degenera);
- ✓ *definita negativa* se $n = p = 0$ (questo si verifica sse $Q(v) \leq 0$ per ogni $v \in V$ e $Q(v) = 0$ solo se $v = 0$ e, in questo caso, β è non degenera);
- ✓ *indefinita* se p e q sono entrambi diversi da 0.

Esercizio 1.7 Si determini un riferimento di \mathbb{R}^4 ortonormale rispetto alla forma bilineare simmetrica β dell'Esercizio 1.5. Qual'è la segnatura di β ? ■

Esercizio 1.8 Si consideri l'applicazione

$$\beta : \mathbb{R}^4 \times \mathbb{R}^4 \rightarrow \mathbb{R}$$

data da

$$\beta \left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} \right) = 3(x_1 - x_3 + x_4)(y_1 - y_3 + y_4) + x_1 y_2 + x_2 y_1 - 2x_2 y_2.$$

Dopo aver dimostrato che β è una forma bilineare simmetrica, si determini un riferimento $\mathcal{R} = (v_0, v_1, v_2, v_3)$ di \mathbb{R}^4 ortonormale rispetto a β . Qual'è la segnatura di β ? ■

Sia infine $B \in M_m(\mathbb{R})$ una matrice reale simmetrica.

Corollario 1.3.8 Ogni matrice reale simmetrica è congruente ad un'unica matrice del tipo (1.16).

Dimostrazione. Sia $B \in M_n(\mathbb{R})$ una matrice reale simmetrica. Consideriamo la forma bilineare simmetrica $\beta_B: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$. In virtù della Proposizione 1.3.6 esiste un riferimento $\mathcal{R} = (M^{(1)}, \dots, M^{(n)})$ di \mathbb{R}^n ortonormale rispetto a β_B . In altre parole la matrice rappresentativa N di β nel riferimento \mathcal{R} è del tipo (1.16). Ora, esattamente come nella dimostrazione del Corollario 1.3.5 si ha $N = M^T B M$, dove M è la matrice che ha $M^{(1)}, \dots, M^{(n)}$ come colonne. I rimanenti dettagli, inclusa l'unicità, sono lasciati come Esercizio 1.9. ■

Esercizio 1.9 Si colmino le lacune nella dimostrazione del Corollario 1.3.8 (Suggerimento: si proceda come nella dimostrazione del Corollario 1.3.5). ■

■ **Esempio 1.10** Si consideri la matrice reale simmetrica $B \in M_4(\mathbb{R})$ dell'Esempio 1.8. Il riferimento (1.15) di \mathbb{R}^4 è ortogonale rispetto alla forma bilineare simmetrica β_B . Perciò il sistema

$$\begin{aligned} \bar{\mathcal{R}} &= \left(\frac{1}{\sqrt{|Q(v_0)|}} v_0, \frac{1}{\sqrt{|Q(v_1)|}} v_1, \frac{1}{\sqrt{|Q(v_2)|}} v_2, \frac{1}{\sqrt{|Q(v_3)|}} v_3, \right) \\ &= \left(\begin{pmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sqrt{2}/2 \\ 0 \\ 0 \\ \sqrt{2} \end{pmatrix}, \begin{pmatrix} \sqrt{2}/6 \\ 0 \\ \sqrt{2}/3 \\ \sqrt{2} \end{pmatrix}, \begin{pmatrix} \sqrt{2}/6 \\ -\sqrt{2}/2 \\ \sqrt{2}/3 \\ 0 \end{pmatrix} \right) \end{aligned}$$

è un riferimento ortonormale rispetto a β_B . Scambiando il secondo e il terzo vettore nel riferimento $\bar{\mathcal{R}}$ otteniamo il riferimento

$$\left(\begin{pmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sqrt{2}/6 \\ 0 \\ \sqrt{2}/3 \\ \sqrt{2} \end{pmatrix}, \begin{pmatrix} \sqrt{2}/2 \\ 0 \\ 0 \\ \sqrt{2} \end{pmatrix}, \begin{pmatrix} \sqrt{2}/6 \\ -\sqrt{2}/2 \\ \sqrt{2}/3 \\ 0 \end{pmatrix} \right)$$

nel quale la matrice rappresentativa di β_B è la forma normale

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

Ne deduciamo che, posto

$$M = \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/6 & \sqrt{2}/2 & \sqrt{2}/6 \\ \sqrt{2}/2 & 0 & 0 & -\sqrt{2}/2 \\ 0 & \sqrt{2}/3 & 0 & \sqrt{2}/3 \\ 0 & \sqrt{2} & \sqrt{2} & 0 \end{pmatrix},$$

si ha $M^T B M = N$. ■

Esercizio 1.10 Si consideri la matrice simmetrica reale $B \in M_4(\mathbb{R})$ dell'Esercizio 1.6. Si determini una matrice invertibile $M \in GL_4(\mathbb{R})$ tale che $M^T B M$ è una forma normale. ■

Esercizio 1.11 Si consideri la matrice simmetrica reale

$$B = \begin{pmatrix} 3 & 0 & -1 & 0 \\ 0 & 3 & 1 & 1 \\ -1 & 1 & 2 & -1 \\ 0 & 1 & -1 & 2 \end{pmatrix} \in M_4(\mathbb{R}).$$

Si determini una matrice invertibile $M \in GL_4(\mathbb{R})$ tale che $M^T B M$ è una forma normale. ■

1.4 Forme Bilineari Alternanti

Nell'ultima sezione di questo capitolo studiamo brevemente le forme bilineari alternanti e le loro forme normali. Sia V uno spazio vettoriale sul campo \mathbb{K} .

Definizione 1.4.1 — Forma Bilineare Alternante. Una forma bilineare $\omega : V \times V \rightarrow \mathbb{K}$ sullo spazio vettoriale V si dice *alternante* se

$$\omega(v, v) = 0, \quad \text{per ogni } v \in V.$$

Proposizione 1.4.1 Sia $\omega : V \times V \rightarrow \mathbb{K}$ una forma bilineare alternante sullo spazio vettoriale \mathbb{K} . Allora ω è *antisimmetrica*, cioè

$$\omega(v, w) = -\omega(w, v), \quad \text{per ogni } v, w \in V.$$

Se il campo \mathbb{K} ha caratteristica diversa da 2 allora è vero anche il viceversa.

Dimostrazione. Sia $\omega : V \times V \rightarrow \mathbb{K}$ una forma bilineare alternante. Allora, per ogni $v, w \in V$, si ha

$$0 = \omega(v+w, v+w) = \omega(v, v) + \omega(v, w) + \omega(w, v) + \omega(w, w) = \omega(v, w) + \omega(w, v)$$

da cui $\omega(v, w) = -\omega(w, v)$ come desiderato. Viceversa sia $\omega : V \times V \rightarrow \mathbb{K}$ una forma bilineare alternante e sia \mathbb{K} un campo di caratteristica diversa da due. Allora, per ogni $v \in V$,

$$\omega(v, v) = -\omega(v, v) \quad \Rightarrow \quad 2\omega(v, v) = 0 \quad \Rightarrow \quad \omega(v, v) = 0.$$

Questo conclude la dimostrazione. ■

Teorema 1.4.2 — Forma Normale di una Forma Bilineare Alternante. Sia V uno spazio vettoriale di dimensione m sul campo \mathbb{K} e sia $\omega : V \times V \rightarrow \mathbb{K}$ una forma bilineare alternante su V . Allora esiste un riferimento di V nel quale la matrice rappresentativa di ω è del tipo

$$U = \begin{pmatrix} 0 & I_r & 0 \\ -I_r & 0 & 0 \\ 0 & 0 & 0_n \end{pmatrix}, \quad (1.19)$$

con $2r + n = m$. In particolare, il rango di ω è sempre un numero pari.

Dimostrazione. La dimostrazione è per induzione sulla dimensione m di V ed è simile in spirito alla dimostrazione del Teorema 1.3.4. Se $n = 1$, la matrice rappresentativa di ω in ogni riferimento (v) è la matrice 1×1 la cui unica entrata è $\omega(v, v) = 0$, cioè la matrice nulla che è del tipo (1.19). Supponiamo ora che l'asserto sia vero per ogni $m \leq k$ e dimostriamo che è vero anche per $m = k + 1$. Sia dunque V uno spazio vettoriale di dimensione $k + 1$ e sia $\omega : V \times V \rightarrow \mathbb{K}$ una forma bilineare alternante su V . Se $\omega = 0$, allora in ogni riferimento la matrice rappresentativa di ω è la matrice

nulla che (come abbiamo già detto almeno nel caso 1×1) è del tipo (1.19). Se invece $\omega \neq 0$, vuol dire che esistono due vettori $v_0, v' \in V$ non nulli tali che $\alpha := \omega(v_0, v') \neq 0$. Poniamo

$$v_1 = \frac{1}{\alpha} v'$$

e osserviamo che $\omega(v_0, v_1) = 1$, infatti

$$\omega(v_0, v_1) = \omega\left(v_0, \frac{1}{\alpha} v'\right) = \frac{1}{\alpha} \omega(v_0, v') = \frac{\alpha}{\alpha} = 1.$$

Si noti che v_0, v_1 sono linearmente indipendenti. Infatti, se esistesse $a \in \mathbb{K}$ tale che $v_1 = av_0$, allora si avrebbe

$$1 = \omega(v_0, v_1) = \omega(v_0, av_0) = a\omega(v_0, v_0) = 0,$$

il che è assurdo. Dunque v_0, v_1 generano un sottospazio

$$W_0 = \text{Span}(v_0, v_1)$$

di dimensione 2. Consideriamo ora il sottoinsieme

$$W := \{w \in V : \omega(v_0, w) = \omega(v_1, w) = 0\} \subseteq V.$$

Si tratta di un sottospazio vettoriale (il lettore lo vede?). Il prossimo passo della dimostrazione consiste nel dimostrare che W è un sottospazio di dimensione $k - 1$. A questo scopo osserviamo che

$$W_0 \cap W = \{0\}.$$

Infatti sia $w \in W_0 \cap W$. Allora esistono $a_0, a_1 \in \mathbb{K}$ tali che $w = a_0 v_0 + a_1 v_1$. D'altro canto abbiamo anche

$$0 = \omega(v_0, w) = \omega(v_0, a_0 v_0 + a_1 v_1) = a_0 \omega(v_0, v_0) + a_1 \omega(v_0, v_1) = a_1.$$

Similmente $a_0 = 0$ e dunque $w = 0$. Perciò i sottospazi W_0 e W_1 sono in somma diretta:

$$W + W_0 = W \oplus W_0.$$

Dimostriamo ora che $W + W_0 = V$: il fatto che la dimensione di W è $k - 1$ seguirà dalla Relazione di Grassmann $\dim W + \dim W_0 = \dim V$. Consideriamo dunque un generico vettore $v \in V$. Ovviamente

$$v = \left(v - \omega(v_0, v)v_1 + \omega(v_1, v)v_0\right) + \left(\omega(v_0, v)v_1 - \omega(v_1, v)v_0\right). \quad (1.20)$$

Il primo addendo

$$w := v - \omega(v_0, v)v_1 + \omega(v_1, v)v_0$$

appartiene a W , infatti

$$\begin{aligned} \omega(v_0, w) &= \omega(v_0, v - \omega(v_0, v)v_1 + \omega(v_1, v)v_0) \\ &= \omega(v_0, v) - \omega(v_0, v)\omega(v_0, v_1) + \omega(v_1, v)\omega(v_0, v_0) \\ &= \omega(v_0, v) - \omega(v_0, v) = 0, \end{aligned}$$

e similamente

$$\begin{aligned}\omega(v_1, w) &= \omega(v_1, v - \omega(v_0, v)v_1 + \omega(v_1, v)v_0) \\ &= \omega(v_1, v) - \omega(v_0, v)\omega(v_1, v_1) + \omega(v_1, v)\omega(v_1, v_0) \\ &= \omega(v_1, v) - \omega(v_1, v) = 0.\end{aligned}$$

Il secondo addendo

$$\omega(v_0, v)v_1 - \omega(v_1, v)v_0$$

della (1.20) appartiene a W_0 . Dunque ogni vettore v di V si può scrivere come somma di un vettore di W e di un vettore di W_0 . Questo mostra che $V = W + W_0$ come desiderato. Ne consegue che $\dim W = k - 1$ come annunciato.

Ora possiamo restringere la forma bilineare ω a W definendo una nuova applicazione

$$\omega_W : W \times W \rightarrow \mathbb{K}, \quad (w_1, w_2) \mapsto \omega_W(w_1, w_2) := \omega(w_1, w_2).$$

È facile mostrare che ω_W è una forma bilineare alternante su W (il lettore lo vede?). A ω_W si applica l'ipotesi di induzione, dunque esiste un riferimento (v_2, \dots, v_k) di W nel quale la matrice rappresentativa di ω_W è

$$U_W = \begin{pmatrix} 0 & I_r & 0 \\ -I_r & 0 & 0 \\ 0 & 0 & 0_p \end{pmatrix},$$

con $2r + p = k - 1$. Consideriamo ora il sistema

$$\tilde{\mathcal{R}} = (\underbrace{v_0, v_1}_{\in W_0}, \underbrace{v_2, \dots, v_k}_{\in W}).$$

Si tratta d un riferimento di W perchè ottenuto unendo i riferimenti di due sottospazi in somma diretta che generano V . La matrice rappresentativa di ω nel riferimento $\tilde{\mathcal{R}}$ è

$$\left(\omega(v_i, v_j) \right)_{\substack{i=0, \dots, k \\ j=0, \dots, k}} = \begin{pmatrix} \omega(v_0, v_0) & \omega(v_0, v_1) & \omega(v_0, v_2) & \cdots & \omega(v_0, v_k) \\ \omega(v_1, v_0) & \omega(v_1, v_1) & \omega(v_1, v_2) & \cdots & \omega(v_1, v_k) \\ \omega(v_2, v_0) & \omega(v_2, v_1) & \omega(v_2, v_2) & \cdots & \omega(v_2, v_k) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \omega(v_k, v_0) & \omega(v_k, v_1) & \omega(v_k, v_2) & \cdots & \omega(v_k, v_k). \end{pmatrix}.$$

Il blocco 2×2 in alto a sinistra è

$$\begin{pmatrix} \omega(v_0, v_0) & \omega(v_0, v_1) \\ \omega(v_1, v_0) & \omega(v_1, v_1) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Il blocco $(k - 1) \times (k - 1)$ in basso a destra è la matrice rappresentativa di ω_W nel riferimento (v_2, \dots, v_k) (infatti, sui vettori di W , ω e ω_W agiscono nello stesso modo) e dunque è la matrice U_W . I rimanenti due blocchi sono nulli, perché $v_i \in W$, e dunque $\omega(v_0, v_i) = \omega(v_1, v_i) = 0$, per ogni $i = 2, \dots, k$. Riassumendo, la matrice rappresentativa di ω nel riferimento $\tilde{\mathcal{R}}$ è

$$\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & U_W \end{pmatrix}.$$

Riordinando opportunamente i vettori di $\tilde{\mathcal{R}}$ possiamo fare in modo che la matrice rappresentativa diventi del tipo (1.19). È sufficiente passare al riferimento

$$\mathcal{R} = (v_0, v_{r+1}, v_2, \dots, v_r, v_1, v_{r+2}, \dots, v_m)$$

scambiando i vettori v_1 e v_{r+1} . I dettagli sono lasciati al lettore. ■

- Una forma bilineare alternante non degenera è detta una *forma simplettica*. Uno spazio vettoriale V munito di una forma simplettica è detto uno *spazio vettoriale simplettico*. Segue dal Teorema 1.4.2 che uno spazio vettoriale simplettico (V, ω) ha necessariamente dimensione pari. Inoltre esiste un riferimento di V nel quale la matrice rappresentativa della forma simplettica ω è

$$U = \begin{pmatrix} 0 & I_r \\ -I_r & 0 \end{pmatrix}$$

con $2r = \dim V$.

2. Spazi Vettoriali Euclidei

Informalmente, uno spazio vettoriale euclideo è uno spazio vettoriale munito di una struttura aggiuntiva che dà senso a concetti di natura metrica come *lunghezza di un vettore*, *distanza tra due vettori*, *angolo tra due vettori*, in particolare *vettori ortogonali*. In questo capitolo definiremo gli spazi vettoriali euclidei e tutte le costruzioni metriche appena citate, discutendone le principali proprietà. Discuteremo anche di rappresentazione coordinata di uno spazio vettoriale euclideo e introdurremo una classe di applicazioni tra spazi vettoriali euclidei particolarmente adatte a confrontarli. Infine parleremo di *simmetrie* di uno spazio vettoriale euclideo.

2.1 Prodotti Scalari, Norme, Distanze e Angoli

In questo capitolo lavoreremo solo con *spazi vettoriali reali*, cioè spazi vettoriali sul campo \mathbb{R} dei numeri reali. Sia dunque V uno spazio vettoriale reale di dimensione finita.

Definizione 2.1.1 — Prodotto Scalare. Un *prodotto scalare* in V è una forma bilineare simmetrica

$$V \times V \rightarrow \mathbb{R}, \quad (v, w) \mapsto \langle v, w \rangle,$$

definita positiva, cioè:

- (1) **[bilinearità]** per ogni $v, v', w \in V$ e per ogni $a, a' \in \mathbb{R}$,

$$\langle av + a'v', w \rangle = a\langle v, w \rangle + a'\langle v', w \rangle,$$

e per ogni $v, w, w' \in V$ e ogni $b, b' \in \mathbb{R}$,

$$\langle v, bw + b'w' \rangle = b\langle v, w \rangle + b'\langle v, w' \rangle;$$

- (2) **[simmetria]** per ogni $v, w \in V$, $\langle v, w \rangle = \langle w, v \rangle$;

- (3) **[positività]** per ogni $v \in V$, $\langle v, v \rangle \geq 0$, e $\langle v, v \rangle = 0$ solo se $v = 0$.

Uno *spazio vettoriale euclideo* è uno spazio vettoriale reale munito di un prodotto scalare. In altre parole, uno spazio vettoriale euclideo è una coppia $(V, \langle -, - \rangle)$, in cui V è uno spazio vettoriale reale (di dimensione finita) e $\langle -, - \rangle$ è un prodotto scalare in V .

Segue immediatamente dalla definizione che un prodotto scalare è una forma bilineare (simmetrica e) non-degenere, infatti dalla positività segue che l'unico vettore v tale che $\langle v, w \rangle = 0$ per ogni $w \in V$ è il vettore nullo (se $v \neq 0$ allora si ha $\langle v, v \rangle \neq 0$).

■ **Esempio 2.1** Sia n un intero positivo. Lo spazio vettoriale numerico \mathbb{R}^n è munito di un prodotto scalare canonico, che indichiamo con il simbolo " $\langle -, - \rangle_{can}$ ", definito da

$$\langle x, y \rangle_{can} := x^T y = x_1 y_1 + \cdots + x_n y_n = \sum_{i=1}^n x_i y_i,$$

per ogni $x = (x_1, \dots, x_n)^T, y = (y_1, \dots, y_n)^T \in \mathbb{R}^n$. Il prodotto scalare $\langle -, - \rangle_{can}$ è detto *prodotto scalare standard* e lo spazio numerico \mathbb{R}^n , munito del prodotto scalare standard, è detto lo *spazio vettoriale euclideo numerico* o lo *spazio vettoriale euclideo standard*. Questa terminologia sarà chiara più avanti. ■

Il seguente esercizio fornisce una intera famiglia di esempi di prodotto scalare sullo spazio numerico.

Esercizio 2.1 Sia $A \in GL_n(\mathbb{R})$ una matrice reale invertibile. Si dimostri che l'applicazione

$$\langle -, - \rangle_A : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad (x, y) \mapsto \langle x, y \rangle_A := x^T A^T A y$$

è un prodotto scalare in \mathbb{R}^n (si osservi che, per ogni $x, y \in \mathbb{R}^n$, si ha che $\langle x, y \rangle_A = \langle Ax, Ay \rangle_{can}$; in particolare, se $A = I_n$ è la matrice identica, allora $\langle -, - \rangle_A = \langle -, - \rangle_{I_n}$ è esattamente il prodotto scalare standard). ■

Ora sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo, e sia $W \subseteq V$ un sottospazio vettoriale.

Proposizione 2.1.1 La restrizione $\langle -, - \rangle_W$ del prodotto scalare a $W \times W \subseteq V \times V$:

$$\langle -, - \rangle_W : W \times W \rightarrow \mathbb{R}, \quad (w_1, w_2) \mapsto \langle w_1, w_2 \rangle_W := \langle w_1, w_2 \rangle,$$

è un prodotto scalare in W .

Dimostrazione. La dimostrazione è lasciata per Esercizio 2.2. ■

Esercizio 2.2 Dimostrare la Proposizione 2.1.1. ■

La Proposizione 2.1.1 mostra che ogni sottospazio vettoriale di uno spazio vettoriale euclideo è esso stesso uno spazio vettoriale euclideo in modo naturale (con il prodotto scalare ristretto). Dato un sottospazio vettoriale $W \subseteq V$ di uno spazio vettoriale euclideo $(V, \langle -, - \rangle)$, lo spazio vettoriale euclideo $(W, \langle -, - \rangle_W)$ sarà talvolta detto un *sottospazio vettoriale euclideo*.

Introduciamo ora le prime costruzioni metriche in uno spazio vettoriale euclideo. Sia dunque $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo.

Definizione 2.1.2 — Norma e Distanza. La *norma quadra* è la forma quadratica $\|-\|^2$ associata al prodotto scalare $\langle -, - \rangle$, cioè la funzione:

$$V \rightarrow \mathbb{R}, \quad v \mapsto \|v\|^2 := \langle v, v \rangle.$$

La *norma* o *lunghezza* è la funzione

$$V \rightarrow \mathbb{R}, \quad v \mapsto \|v\| := \sqrt{\|v\|^2} = \sqrt{\langle v, v \rangle}.$$

Un vettore $v \in V$ di *norma unitaria*, cioè tale che $\|v\| = 1$, è detto un *versore*.

La *distanza* è la funzione

$$d: V \times V \rightarrow \mathbb{R}, \quad (v, w) \mapsto d(v, w) := \|v - w\|.$$

Osserviamo che la norma e la distanza sono ben definite giacché, dalla positività del prodotto scalare, la norma quadra è sempre non-negativa. Si osservi inoltre che la proprietà (3) nella Definizione 2.1.1 dice precisamente che la norma $\|v\|$ di un vettore $v \in V$ è nulla sse $v = 0$. Sarà bene, nel seguito, tenere a mente questa osservazione.

Esercizio 2.3 Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo e sia $\ell \subseteq V$ un sottospazio vettoriale 1-dimensionale. Dimostrare che, per ogni numero reale $a > 0$, esistono esattamente 2 vettori in ℓ di norma pari ad a e sono uno l'opposto dell'altro (in particolare, esistono esattamente 2 versori in ℓ , uno l'opposto dell'altro). ■

■ **Esempio 2.2** La norma quadra, la norma e la distanza nello spazio vettoriale euclideo standard $(\mathbb{R}^n, \langle -, - \rangle_{can})$ sono date da

$$\|x\|^2 = x^T x = x_1^2 + \cdots + x_n^2 = \sum_{i=1}^n x_i^2,$$

$$\|x\| = \sqrt{x^T x} = \sqrt{x_1^2 + \cdots + x_n^2} = \sqrt{\sum_{i=1}^n x_i^2},$$

$$d(x, y) = \sqrt{(x-y)^T(x-y)} = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2},$$

per ogni $x = (x_1, \dots, x_n)^T, y = (y_1, \dots, y_n)^T \in \mathbb{R}^n$. ■

Dalla bilinearità del prodotto scalare segue immediatamente che

$$\|av\|^2 = a^2 \|v\|^2, \quad \text{e} \quad \|av\| = |a| \|v\|$$

per ogni $v \in V$ e $a \in \mathbb{R}$. Inoltre vale il seguente importante

Teorema 2.1.2 — Disuguaglianza di Cauchy-Schwarz. Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo. Allora

$$|\langle v, w \rangle| \leq \|v\| \|w\| \tag{2.1}$$

per ogni $v, w \in V$. Inoltre, nella (2.1), l'uguaglianza vale sse v e w sono linearmente dipendenti.

Dimostrazione. Innanzitutto osserviamo che, se $v = 0$, la disuguaglianza (2.1) vale ed è, in effetti, l'uguaglianza $0 = 0$, in accordo anche con la seconda parte dell'enunciato. Perciò, nel seguito, potremo sempre assumere $v \neq 0$. Siano ora $v, w \in V$ e siano $a, b \in \mathbb{R}$. In virtù della positività del prodotto scalare si ha:

$$\begin{aligned} 0 &\leq \|av + bw\|^2 = \langle av + bw, av + bw \rangle = a^2 \langle v, v \rangle + ab \langle v, w \rangle + ab \langle w, v \rangle + b^2 \langle w, w \rangle \\ &= a^2 \|v\|^2 + 2ab \langle v, w \rangle + b^2 \|w\|^2, \end{aligned} \tag{2.2}$$

dove abbiamo anche usato la bilinearità e la simmetria (dove e come sono state usate queste proprietà?). In particolare, quando $a = -\langle v, w \rangle$ e $b = \|v\|^2$ otteniamo

$$0 \leq \langle v, w \rangle^2 \|v\|^2 - 2\|v\|^2 \langle v, w \rangle^2 + \|v\|^4 \|w\|^2 = -\langle v, w \rangle^2 \|v\|^2 + \|v\|^4 \|w\|^2. \quad (2.3)$$

A questo punto si ricordi che stiamo assumendo $v \neq 0$, perciò $\|v\|^2 > 0$ e, dividendo la (2.3) per $\|v\|^2$, si ottiene

$$0 \leq -\langle v, w \rangle^2 + \|v\|^2 \|w\|^2$$

che è equivalente alla (2.1). Per la seconda parte dell'enunciato, supponiamo innanzitutto che v e w siano linearmente dipendenti. Giacché $v \neq 0$ questo equivale a dire che esiste $\alpha \in \mathbb{R}$ tale che $w = \alpha v$ e, in tal caso,

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

D'altro canto

$$\|v\| \|w\| = \|v\| \|\alpha v\| = |\alpha| \|v\|^2.$$

Dunque nella (2.1) vale l'uguaglianza. Viceversa, se nella (2.1) vale l'uguaglianza, percorrendo a ritroso i conti (2.3) e (2.2), troviamo $\|av + bw\|^2 = 0$, e perciò $av + bw = 0$ con $a = -\langle v, w \rangle$ e $b = \|v\|^2$. Giacché $v \neq 0$, si ha $b \neq 0$, da cui $w = -b^{-1}av$, in particolare w dipende linearmente da v . Questo conclude la dimostrazione. ■

Veniamo ora alla distanza: essa gode delle seguenti proprietà:

- (1) $d(v, w) = d(w, v)$, per ogni $v, w \in V$,
- (2) $d(v, w) \geq 0$, per ogni $v, w \in V$, e $d(v, w) = 0$ sse $v = w$.

Il lettore lo verifichi per esercizio. Inoltre la disuguaglianza di Cauchy-Schwarz ha la seguente notevole conseguenza:

Corollario 2.1.3 — Disuguaglianza Triangolare. Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo. Per ogni terna di vettori $u, v, w \in V$ si ha che

$$d(v, w) \leq d(v, u) + d(u, w). \quad (2.4)$$

Dimostrazione. Dimostriamo innanzitutto che, per ogni $\tilde{v}, \tilde{u} \in V$,

$$\|\tilde{v} + \tilde{u}\| \leq \|\tilde{v}\| + \|\tilde{u}\|. \quad (2.5)$$

È così, infatti

$$\|\tilde{v} + \tilde{u}\|^2 = \langle \tilde{v} + \tilde{u}, \tilde{v} + \tilde{u} \rangle = \|\tilde{v}\|^2 + 2\langle \tilde{v}, \tilde{u} \rangle + \|\tilde{u}\|^2 \leq \|\tilde{v}\|^2 + 2|\langle \tilde{v}, \tilde{u} \rangle| + \|\tilde{u}\|^2,$$

dove abbiamo usato la bilinearità e la simmetria del prodotto scalare. Ora, dalla disuguaglianza di Cauchy-Schwarz, $|\langle \tilde{v}, \tilde{u} \rangle| \leq \|\tilde{v}\| \|\tilde{u}\|$, perciò

$$\|\tilde{v} + \tilde{u}\|^2 \leq \|\tilde{v}\|^2 + 2\|\tilde{v}\| \|\tilde{u}\| + \|\tilde{u}\|^2 \leq \|\tilde{v}\|^2 + 2\|\tilde{v}\| \|\tilde{u}\| + \|\tilde{u}\|^2 = (\|\tilde{v}\| + \|\tilde{u}\|)^2,$$

il che è equivalente alla (2.5). Infine, la (2.4) si ottiene dalla (2.5) ponendo $\tilde{v} = v - u$ e $\tilde{u} = u - w$ e usando che, da queste posizioni, segue $v - w = \tilde{v} + \tilde{u}$. ■

La Disuguaglianza Triangolare (2.4) mostra, in particolare, che in uno spazio vettoriale euclideo “in ogni triangolo, la lunghezza di ogni lato è minore della somma delle lunghezze degli altri due lati” (Figura 2.1). Questo giustifica la terminologia.

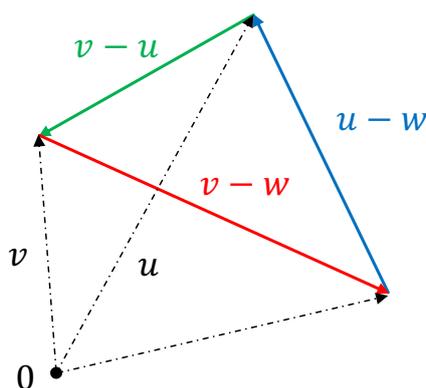


Figura 2.1: Disuguaglianza triangolare.

- Sia X un insieme. Una *distanza* su X è, per definizione, una funzione $d : X \times X \rightarrow \mathbb{R}$ tale che
- (1) $d(x, y) = d(y, x)$ per ogni $x, y \in X$,
 - (2) $d(x, y) \geq 0$ per ogni $x, y \in X$ e $d(x, y) = 0$ sse $x = y$,
 - (3) **[disuguaglianza triangolare]** $d(x, y) \leq d(x, z) + d(z, y)$ per ogni $x, y, z \in X$.

Uno *spazio metrico* è un insieme munito di una distanza. Gli spazi metrici hanno notevoli applicazioni in Geometria e Analisi. Il Corollario 2.1.3, insieme all'osservazione che lo precede, può essere riformulato dicendo che *uno spazio vettoriale euclideo, munito della distanza indotta dal prodotto scalare, è uno spazio metrico*.

In uno spazio vettoriale euclideo ha anche senso parlare di *angolo tra due vettori*.

Definizione 2.1.3 — Angolo tra Vettori. L'*angolo convesso* tra due vettori non nulli v e w dello spazio vettoriale euclideo $(V, \langle -, - \rangle)$ è l'angolo \widehat{vw} univocamente determinato dalle due condizioni:

- (1) $0 \leq \widehat{vw} \leq \pi$, e
- (2) $\cos \widehat{vw} = \langle v, w \rangle / \|v\| \|w\|$.

Per convenzione l'angolo tra il vettore nullo e un qualunque altro vettore è $\pi/2$.

Si osservi che la formula al punto (2) della Definizione 2.1.3 è ben posta in virtù della disuguaglianza di Cauchy-Schwarz. Inoltre, l'angolo tra i vettori v, w non dipende dall'ordine in cui consideriamo i due vettori: $\widehat{vw} = \widehat{wv}$. Infine, l'angolo tra v e w non cambia se si moltiplica uno dei due vettori per uno scalare positivo, cioè, per ogni $a \in \mathbb{R}$ con $a > 0$, si ha $\widehat{(av)w} = \widehat{vw}$. Se invece si moltiplica uno dei due vettori per uno scalare negativo, l'angolo cambia per il suo supplementare, cioè, per ogni $b \in \mathbb{R}$ con $b < 0$, si ha $\widehat{(bv)w} = \pi - \widehat{vw}$ (Figura 2.2). Il lettore è invitato a verificare nel dettaglio ciascuna di queste asserzioni (compreso il caso in cui uno tra v e w è il vettore nullo).

2.2 Ortogonalità

In questa sezione studieremo la struttura di uno spazio vettoriale euclideo. In un certo senso opportuno, *tutti gli spazi euclidei della stessa dimensione sono isomorfi*. Allo scopo di chiarire quest'ultima affermazione è opportuno discutere preliminarmente di 1) matrice rappresentativa di un prodotto scalare, e 2) ortogonalità tra vettori. Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo n -dimensionale. Ricordiamo che, dato un riferimento $\mathcal{R} = (v_1, \dots, v_n)$ di V , la *matrice rappresentativa* della forma bilineare $\langle -, - \rangle$ nel riferimento \mathcal{R} , è la matrice quadrata

$$G := \begin{pmatrix} \langle v_1, v_1 \rangle & \cdots & \langle v_1, v_n \rangle \\ \vdots & \ddots & \vdots \\ \langle v_n, v_1 \rangle & \cdots & \langle v_n, v_n \rangle \end{pmatrix} = \left(\langle v_i, v_j \rangle \right)_{i=1, \dots, n}^{j=1, \dots, n}.$$

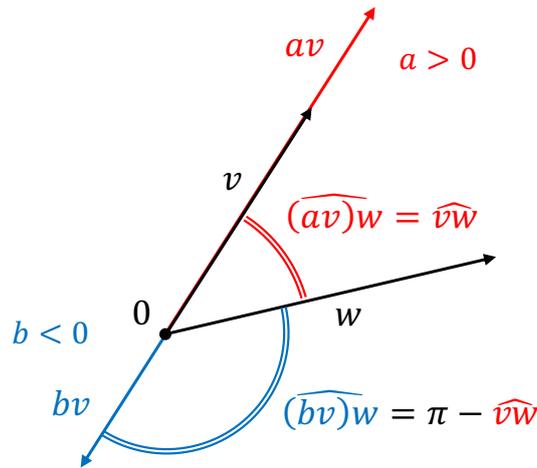


Figura 2.2: Angolo tra vettori non nulli.

La matrice G è una matrice simmetrica perché il prodotto scalare è una forma bilineare simmetrica. Inoltre G è una matrice invertibile, perché il prodotto scalare è una forma bilineare non-degenere. Infine G gode della seguente proprietà notevole: Per ogni $v, w \in V$

$$\langle v, w \rangle = x^T G y$$

in cui $x, y \in \mathbb{R}^n$ sono i vettori coordinati di v, w nel riferimento \mathcal{R} , rispettivamente. Questi richiami sulla matrice rappresentativa di un prodotto scalare saranno molto utili nel seguito.

Definizione 2.2.1 — Vettori Ortogonali. Due vettori $v, w \in V$ sono *ortogonali* se $\langle v, w \rangle = 0$. In tal caso si scrive $v \perp w$. Un vettore $v \in V$ è *ortogonale ad un sottoinsieme* $S \subseteq V$ se è ortogonale ad ogni vettore di S . In tal caso si scrive $v \perp S$. In altre parole

$$v \perp S \iff \langle v, w \rangle = 0 \text{ per ogni } w \in S.$$

Si osservi che i vettori $v, w \in V$ sono ortogonali sse l'angolo \widehat{vw} è un angolo retto, cioè $\widehat{vw} = \pi/2$.

■ **Esempio 2.3** Il vettore nullo è ortogonale ad ogni altro vettore. ■

Definizione 2.2.2 — Sistema Ortogonale. Un sistema di vettori $S = (v_1, \dots, v_p)$ in uno spazio vettoriale euclideo $(V, \langle -, - \rangle)$ si dice *ortogonale* se i vettori di S sono a due a due ortogonali, cioè $v_i \perp v_j$ per ogni $i \neq j, i, j = 1, \dots, p$.

Lemma 2.2.1 Sia $S = (v_1, \dots, v_p)$ un sistema ortogonale. Se $v_i \neq 0$ per ogni $i = 1, \dots, p$, allora S è un sistema linearmente indipendente.

Dimostrazione. Sia $S = (v_1, \dots, v_p)$ un sistema ortogonale di vettori non nulli. Consideriamo ora una combinazione lineare

$$v = a_1 v_1 + \dots + a_p v_p, \quad a_i \in \mathbb{R}, \quad i = 1, \dots, p.$$

Se $v = 0$, allora, per ogni $i = 1, \dots, p$ si ha

$$0 = \langle v_i, v \rangle = \left\langle v_i, \sum_{j=1}^p a_j v_j \right\rangle = \sum_{j=1}^p a_j \langle v_i, v_j \rangle.$$

Giacché i vettori di S sono a due a due ortogonali, l'unico termine che sopravvive nell'ultima sommatoria è l' i -esimo, e troviamo

$$0 = a_i \langle v_i, v_i \rangle = a_i \|v_i\|^2.$$

Ma $v_i \neq 0$, e dunque $\|v_i\|^2 \neq 0$ per ogni i . Perciò l'unica possibilità è che sia $a_i = 0$ per ogni i . Questo mostra che l'unica combinazione lineare nulla dei vettori v_1, \dots, v_p è quella banale, cioè S è un sistema linearmente indipendente come desiderato. ■

Definizione 2.2.3 — Complemento Ortogonale. Il *complemento ortogonale* di un sottoinsieme $S \subseteq V$ è il sottoinsieme

$$S^\perp := \{v \in V : v \perp S\} \subseteq V.$$

In altre parole S^\perp è l'insieme di tutti i vettori di V ortogonali ad ogni vettore di S .

Lemma 2.2.2 Sia $S \subseteq V$ un sottoinsieme e $W \subseteq V$ un sottospazio. Allora

- (1) se $T \subseteq S$ è un sottoinsieme di S , allora $T^\perp \supseteq S^\perp$;
- (2) $S^\perp \subseteq V$ è un sottospazio vettoriale;
- (3) $S^\perp = \text{Span}(S)^\perp$;
- (4) i sottospazi W, W^\perp sono in somma diretta e $W \oplus W^\perp = V$ (Figura 2.3);
- (5) $W^{\perp\perp} = W$.

Dal punto (4) segue, in particolare, che $\dim W^\perp = \dim V - \dim W$.

Dimostrazione. Il punto (1) è chiaro perché se un vettore è ortogonale a tutti i vettori di S in particolare è ortogonale a tutti i vettori di T .

Per il punto (2) dobbiamo verificare che S^\perp contiene il vettore nullo e che è stabile rispetto a combinazioni lineari. È chiaro che il vettore nullo appartiene ad S^\perp (Esempio 2.3). Resta da mostrare che, se $v, u \in S^\perp$ e $a, b \in \mathbb{R}$ allora $av + bu \in S^\perp$. È così infatti sia $w \in S$, allora, dalla bilinearità del prodotto scalare,

$$\langle av + bu, w \rangle = a \langle v, w \rangle + b \langle u, w \rangle = 0,$$

perché v e u sono ortogonali ai vettori di S . Dunque anche $av + bu$ è ortogonale ad ogni vettore di S .

Per il punto (3), l'inclusione $\text{Span}(S)^\perp \subseteq S^\perp$ segue banalmente dal punto (1). Verifichiamo anche l'altra inclusione $S^\perp \subseteq \text{Span}(S)^\perp$. Dobbiamo verificare che se un vettore $v \in V$ è ortogonale ai vettori di S , allora è anche ortogonale ad ogni vettore in $\text{Span}(S)$. Dunque, sia $v \in S^\perp$ e sia $w \in \text{Span}(S)$. Questo vuol dire che $w = a_1 w_1 + \dots + a_p w_p$ per qualche $w_1, \dots, w_p \in S$, e $a_1, \dots, a_p \in \mathbb{R}$. Perciò, dalla bilinearità del prodotto scalare,

$$\langle v, w \rangle = \langle v, a_1 w_1 + \dots + a_p w_p \rangle = a_1 \langle v, w_1 \rangle + \dots + a_p \langle v, w_p \rangle = 0,$$

dove abbiamo usato che $v \perp w_i$ per ogni $i = 1, \dots, p$. Questo conclude la dimostrazione del punto (3).

Per il punto (4), consideriamo un vettore $w \in W \cap W^\perp$. Dunque w è un vettore di W ortogonale a tutti i vettori di W stesso. In particolare $w \perp w$, cioè $\langle w, w \rangle = 0$, da cui $w = 0$. Questo prova che W e W^\perp hanno intersezione banale e, dunque, sono in somma diretta. Per dimostrare che $W + W^\perp = V$ è ora sufficiente dimostrare che $\dim W^\perp = \dim V - \dim W$. Infatti, in tal caso, segue dalla relazione di Grassmann che $W + W^\perp \subseteq V$ è un sottospazio della stessa dimensione di V e perciò coincide con V stesso. Per calcolare $\dim W^\perp$ poniamo $n = \dim V$, e $p = \dim W$. Inoltre, fissiamo un riferimento

(w_1, \dots, w_p) di W e completiamolo ad un riferimento $\mathcal{R} = (w_1, \dots, w_p, v_1, \dots, v_{n-p})$ di V . L'ultimo ingrediente di cui abbiamo bisogno è

$$G = (g_{ij})_{i=1, \dots, p}^{j=1, \dots, n},$$

la matrice rappresentativa del prodotto scalare nel riferimento \mathcal{R} . Giacché W è il sottospazio generato dai vettori w_1, \dots, w_p , per il punto (3), un vettore $v \in V$ è in W^\perp sse $\langle w_i, v \rangle = 0$ per ogni $i = 1, \dots, p$. Ora

$$\langle w_i, v \rangle = x_i^T G y,$$

dove x_i, y sono i vettori coordinati di w_i, v nel riferimento \mathcal{R} . In particolare

$$x_i = \left(0, \dots, \underbrace{1}_{\text{posto } i}, \dots, 0 \right)^T = E_i,$$

l' i -esimo vettore del riferimento canonico di \mathbb{R}^n . Perciò $x_i^T G = E_i^T G = (g_{i1}, \dots, g_{in}) = G_{(i)}$, l' i -esima riga di G , $i = 1, \dots, p$ (se occorre, verificare i dettagli per esercizio). Riassumendo, il vettore v appartiene a W^\perp sse il suo vettore coordinato y è soluzione del seguente sistema lineare omogeneo di p equazioni in n incognite y_1, \dots, y_n :

$$S: \begin{cases} G_{(1)}y = 0 \\ \vdots \\ G_{(p)}y = 0 \end{cases}.$$

Giacché G è una matrice invertibile, le righe $G_{(1)}, \dots, G_{(p)}$ sono linearmente indipendenti e il sistema S è una rappresentazione cartesiana di W^\perp . La dimensione di W^\perp coincide perciò con la dimensione dello spazio delle soluzioni di S che è $n - p$. Questo conclude la dimostrazione del punto (4).

Per concludere, dimostriamo il punto (5). È chiaro che $W \subseteq W^{\perp\perp}$ (perché i vettori di W sono ortogonali a tutti i vettori di W^\perp). D'altronde, per il punto (4), i sottospazi W e $W^{\perp\perp}$ hanno la stessa dimensione e, dunque, necessariamente coincidono. ■

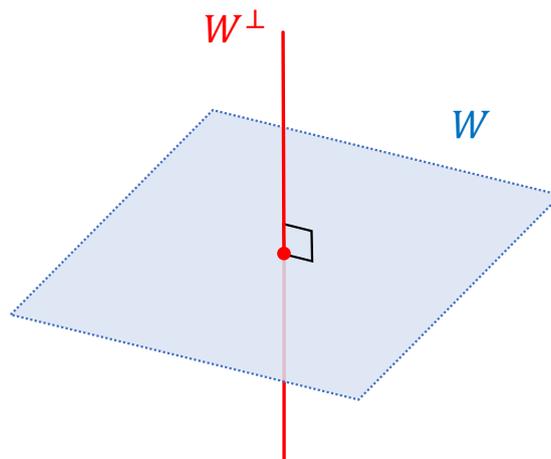


Figura 2.3: Il complemento ortogonale di un sottospazio 2D in uno spazio 3D.

Esercizio 2.4 Trovare la dimensione, una rappresentazione cartesiana e una parametrica per il complemento ortogonale del sottospazio W generato dai vettori

$$\begin{pmatrix} 1 \\ 0 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 1 \\ -1 \\ 1/2 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 1 \\ -2 \end{pmatrix}$$

nello spazio vettoriale Euclideo standard $(\mathbb{R}^4, \langle -, - \rangle_{can})$. ■

Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo, sia $W \subseteq V$ un sottospazio e sia $v \in V$. Dal Lemma 2.2.2 segue che esistono un unico vettore $w \in W$ e un unico vettore $v^\perp \in W^\perp$ tali che (Figura 2.4)

$$v = w + v^\perp \quad (2.6)$$

Definizione 2.2.4 — Proiezione Ortogonale. Il vettore w nella (2.6) è detto la *proiezione ortogonale* di v sul sottospazio W .

Proposizione 2.2.3 Sia W un sottospazio vettoriale k -dimensionale nello spazio vettoriale euclideo $(V, \langle -, - \rangle)$ e sia (u_1, \dots, u_k) un riferimento ortogonale di W (cioè i vettori u_1, \dots, u_k sono a due a due ortogonali rispetto al prodotto scalare ristretto $\langle -, - \rangle_W$ e quindi anche rispetto al prodotto scalare in V). Allora, la proiezione ortogonale w di un vettore $v \in V$ su W è data dalla seguente formula

$$w = \frac{\langle u_1, v \rangle}{\langle u_1, u_1 \rangle} u_1 + \dots + \frac{\langle u_k, v \rangle}{\langle u_k, u_k \rangle} u_k = \sum_{i=1}^k \frac{\langle u_i, v \rangle}{\langle u_i, u_i \rangle} u_i. \quad (2.7)$$

Dimostrazione. Indichiamo con \tilde{w} il membro di destra della (2.7) e consideriamo il vettore $v - \tilde{w}$. È chiaro che $\tilde{w} \in W$. Perciò è sufficiente dimostrare che $v - \tilde{w} \in W^\perp$. In tal caso, infatti, l'asserto segue immediatamente dall'unicità della decomposizione (2.6). Ora, dal Lemma 2.2.2.(3) $v - \tilde{w} \in W^\perp$ sse $v - \tilde{w} \perp u_j$ per ogni $j = 1, \dots, k$. Basta allora calcolare

$$\langle u_j, v - \tilde{w} \rangle = \left\langle u_j, v - \sum_{i=1}^k \frac{\langle u_i, v \rangle}{\langle u_i, u_i \rangle} u_i \right\rangle = \langle u_j, v \rangle - \sum_{i=1}^k \frac{\langle u_i, v \rangle}{\langle u_i, u_i \rangle} \langle u_j, u_i \rangle$$

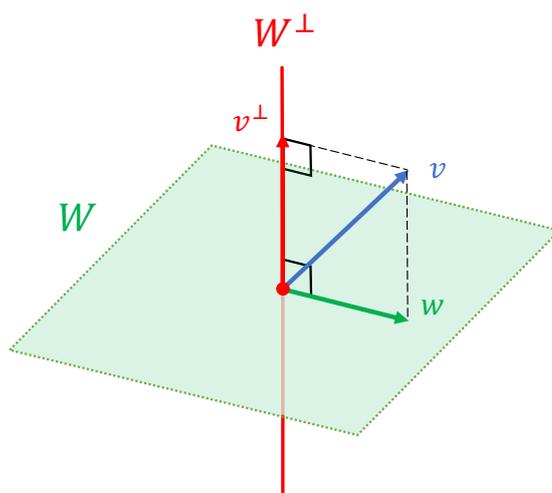
Ma $u_j \perp u_i$ per ogni $i \neq j$, perciò l'unico termine che sopravvive nell'ultima sommatoria è il j -esimo e si ha

$$\langle u_j, v - \tilde{w} \rangle = \langle u_j, v \rangle - \frac{\langle u_j, v \rangle}{\langle u_j, u_j \rangle} \langle u_j, u_j \rangle = 0.$$

Cioè $v \perp u_j$ per ogni j come desiderato. ■

In uno spazio vettoriale euclideo $(V, \langle -, - \rangle)$ è possibile distinguere una famiglia di riferimenti “adatti al prodotto scalare”, cioè una famiglia di riferimenti nei quali il prodotto scalare appaia particolarmente semplice.

Definizione 2.2.5 — Riferimento Ortonormale. Un *riferimento ortonormale* è un riferimento $\mathcal{R} = (e_1, \dots, e_n)$ di V tale che la matrice rappresentativa G del prodotto scalare nel riferimento \mathcal{R} è la matrice identica I_n ($n = \dim V$). In altre parole gli e_i sono versori a due a due ortogonali:

Figura 2.4: Proiezione ortogonale su W .

■ $\langle e_i, e_j \rangle = \delta_{ij}$, per ogni $i, j = 1, \dots, n$.

La Definizione 2.2.5 è compatibile con la Definizione 1.3.6 (il lettore lo vede?). In particolare, segue dalla Proposizione 1.3.6 che ogni spazio vettoriale euclideo ammette riferimenti ortonormali.

■ **Esempio 2.4** Il riferimento canonico $\mathcal{R}_{can} = (E_1, \dots, E_n)$ di \mathbb{R}^n è un riferimento ortonormale rispetto al prodotto scalare standard $\langle -, - \rangle_{can}$. ■

Sia $\mathcal{R} = (e_1, \dots, e_n)$ un riferimento ortonormale. Per definizione, la matrice rappresentativa del prodotto scalare nel riferimento \mathcal{R} è la matrice identica I_n . Dunque per ogni $v, w \in V$, detti $x = (x_1, \dots, x_n)^T, y = (y_1, \dots, y_n)^T \in \mathbb{R}^n$ i rispettivi vettori coordinati nel riferimento \mathcal{R} si ha

$$\langle v, w \rangle = x^T I_n y = x^T y = \langle x, y \rangle_{can}. \quad (2.8)$$

In altre parole *il prodotto scalare tra due vettori è uguale al prodotto scalare standard dei rispettivi vettori coordinati in un riferimento ortonormale*. Questa osservazione sarà molto utile nel seguito.

Le coordinate di un vettore in un riferimento ortonormale si possono calcolare in modo particolarmente semplice, in base alla seguente

Proposizione 2.2.4 Sia $\mathcal{R} = (e_1, \dots, e_n)$ un riferimento ortonormale, e sia $v \in V$. La i -esima coordinata di v nel riferimento \mathcal{R} è data dal prodotto scalare $\langle e_i, v \rangle$.

Dimostrazione. Sia $(x_1, \dots, x_n)^T \in \mathbb{R}^n$ il vettore coordinato di v nel riferimento \mathcal{R} , cioè

$$v = \sum_{j=1}^n x_j e_j.$$

Vogliamo dimostrare che $x_i = \langle e_i, v \rangle$ per ogni $i = 1, \dots, n$. A questo scopo calcoliamo

$$\langle e_i, v \rangle = \left\langle e_i, \sum_{j=1}^n x_j e_j \right\rangle = \sum_{j=1}^n x_j \langle e_i, e_j \rangle = \sum_{j=1}^n x_j \delta_{ij} = x_i.$$

■

Le coordinate di un vettore in un riferimento ortonormale si chiamano anche *coordinate cartesiane*. Nel prossimo esempio discutiamo di rappresentazioni cartesiane e parametriche di un sottospazio vettoriale e del suo complemento ortogonale in coordinate cartesiane.

■ **Esempio 2.5** Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo n -dimensionale, e sia $\mathcal{R} = (e_1, \dots, e_n)$ un riferimento ortonormale. Indichiamo con (x_1, \dots, x_n) le corrispondenti coordinate cartesiane. Consideriamo un sottospazio $W \subseteq V$. A partire da una rappresentazione cartesiana (risp. parametrica) di W (in coordinate cartesiane) possiamo facilmente trovare una rappresentazione parametrica (risp. cartesiana) del complemento ortogonale W^\perp , come segue. Sia

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n = 0 \end{cases} \quad (2.9)$$

una rappresentazione cartesiana di W nelle coordinate (x_1, \dots, x_n) . Siano $u_1, \dots, u_m \in V$ i vettori di coordinate

$$a_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{1n} \end{pmatrix}, \dots, a_m = \begin{pmatrix} a_{m1} \\ \vdots \\ a_{mn} \end{pmatrix} \in \mathbb{R}^n.$$

Vogliamo dimostrare che il sistema (u_1, \dots, u_m) è un riferimento di W^\perp . A questo scopo, osserviamo, innanzitutto, che, giacché i vettori numerici $a_1, \dots, a_m \in \mathbb{R}^n$ sono indipendenti per definizione di rappresentazione cartesiana, anche i vettori $u_1, \dots, u_m \in V$ sono indipendenti. Ora consideriamo un vettore $w \in W$. Il vettore coordinato $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n)^T$ di w nel riferimento \mathcal{R} è una soluzione del sistema (2.9). Cioè, per ogni $i = 1, \dots, m$,

$$0 = a_{i1}\bar{x}_1 + \dots + a_{in}\bar{x}_n = \langle a_i, \bar{x} \rangle_{can} = \langle u_i, w \rangle.$$

Giacché w è arbitrario, questo mostra che $u_i \in W^\perp$ per ogni $i = 1, \dots, m$. D'altronde dalla (2.9) leggiamo che la dimensione di W è $n - m$, perciò la dimensione di W^\perp è $n - (n - m) = m$. Ne deduciamo che (u_1, \dots, u_m) è un riferimento di W^\perp che avrà perciò la rappresentazione parametrica

$$\begin{cases} x_1 = t_1 a_{11} + \dots + t_m a_{m1} \\ \vdots \\ x_n = t_1 a_{1n} + \dots + t_m a_{mn} \end{cases}, \quad t_1, \dots, t_m \in \mathbb{R}. \quad (2.10)$$

Poiché $W = W^{\perp\perp}$, ripercorrendo il ragionamento all'inverso, ricaviamo anche che, se la (2.10) è una rappresentazione parametrica di W , allora la (2.9) è una rappresentazione cartesiana di W^\perp . Si osservi che le stesse conclusioni non valgono se \mathcal{R} non è un riferimento ortonormale. ■

Esercizio 2.5 Sia V uno spazio vettoriale reale n -dimensionale. Dimostrare che, per ogni riferimento $\mathcal{R} = (e_1, \dots, e_n)$ di V esiste un unico prodotto scalare $\langle -, - \rangle_{\mathcal{R}}$ in V tale che \mathcal{R} è un riferimento ortonormale rispetto a $\langle -, - \rangle_{\mathcal{R}}$ (Suggerimento: per ogni $v, w \in V$ si definisca $\langle v, w \rangle_{\mathcal{R}} := x^T y$, dove x, y sono i vettori coordinati di v, w nel riferimento \mathcal{R}). ■

○ L'Esercizio 2.5 mostra, fra l'altro, che un prodotto scalare è completamente determinato da un riferimento ortonormale. Cioè se un riferimento è ortonormale rispetto a due prodotti scalari, allora i due prodotti scalari necessariamente coincidono.

Come abbiamo già osservato, l'esistenza di basi ortogonali e il Teorema di Sylvester garantiscono l'esistenza di riferimenti ortonormali. La prossima proposizione illustra un algoritmo che permette, fra l'altro, di costruire un riferimento ortonormale a partire da un riferimento qualunque.

Teorema 2.2.5 — Ortogonalizzazione di Gram-Schmidt. Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo e sia $S = (v_1, \dots, v_p)$ un sistema di vettori di V linearmente indipendente. Allora esiste un unico sistema (u_1, \dots, u_p) linearmente indipendente e ortogonale che, in aggiunta, soddisfa le seguenti proprietà: per ogni $q = 1, \dots, p$

(1) i vettori u_1, \dots, u_q generano lo stesso sottospazio dei vettori v_1, \dots, v_q :

$$\text{Span}(u_1, \dots, u_q) = \text{Span}(v_1, \dots, v_q);$$

(2) Il prodotto scalare $\langle u_q, v_q \rangle$ coincide con la norma quadra di u_q :

$$\langle u_q, v_q \rangle = \|u_q\|^2.$$

I vettori u_1, \dots, u_p possono essere calcolati induttivamente mediante la formula

$$u_q = v_q - \sum_{i=1}^{q-1} \frac{\langle u_i, v_q \rangle}{\langle u_i, u_i \rangle} u_i, \quad q = 1, \dots, p. \quad (2.11)$$

Dimostrazione. Innanzitutto, per ogni $q = 1, \dots, p$, indichiamo con W_q il sottospazio (di dimensione q) generato da v_1, \dots, v_q :

$$W_q = \text{Span}(v_1, \dots, v_q).$$

La dimostrazione della prima parte dell'enunciato procede per induzione sul numero p di vettori del sistema S . Se $S = (v_1)$ è costituito da un unico vettore, poniamo $u_1 = v_1$, ed è chiaro che (u_1) è l'unico sistema che soddisfa tutte le proprietà richieste. Ora supponiamo che l'asserto sia vero quando $p = k$ e dimostriamo che l'asserto è vero anche quando $p = k + 1$. Dunque consideriamo un sistema (v_1, \dots, v_{k+1}) di vettori linearmente indipendenti. Anche il sistema $S_k = (v_1, \dots, v_k)$ è linearmente indipendente e ad S_k si applica l'ipotesi di induzione. Perciò esiste un unico sistema (u_1, \dots, u_k) indipendente e ortogonale che soddisfa le proprietà (1) e (2) dell'enunciato, per ogni $q = 1, \dots, k$. Consideriamo il sottospazio W_k generato dai vettori v_1, \dots, v_k o, equivalentemente, dai vettori u_1, \dots, u_k . Per il Lemma 2.2.2.(4) si ha $V = W_k \oplus W_k^\perp$. In particolare v_{k+1} si scrive in un unico modo come somma di un vettore w_{k+1} in W_k e di un vettore u_{k+1} in W_k^\perp :

$$v_{k+1} = w_{k+1} + u_{k+1}, \quad w_{k+1} \in W_k, \quad u_{k+1} \in W_k^\perp \quad (2.12)$$

(in altre parole w_{k+1} è la proiezione ortogonale di v_{k+1} sul sottospazio W_k , vedi Figura 2.5).

Dimostriamo che il sistema (u_1, \dots, u_{k+1}) soddisfa le proprietà (1) e (2) per ogni $q = 1, \dots, k + 1$. Già sappiamo che (u_1, \dots, u_{k+1}) soddisfa (1) e (2) per $q = 1, \dots, k$ e resta da dimostrare che

(1) _{$k+1$} $\text{Span}(u_1, \dots, u_{k+1}) = W_{k+1}$;

(2) _{$k+1$} $\langle u_{k+1}, v_{k+1} \rangle = \|u_{k+1}\|^2$.

Per il punto (1) _{$k+1$} osserviamo che, dalla (2.12), segue che $u_{k+1} = v_{k+1} - w_{k+1} \in \text{Span}(v_{k+1}) + W_k = W_{k+1}$. Perciò

$$\text{Span}(u_1, \dots, u_{k+1}) \subseteq W_{k+1},$$

e, per dimostrare l'uguaglianza, è sufficiente mostrare che i due sottospazi $\text{Span}(u_1, \dots, u_{k+1})$ e W_{k+1} hanno la stessa dimensione. Già sappiamo che $\dim W_k = k$ e resta da calcolare la dimensione di $\text{Span}(u_1, \dots, u_{k+1})$. Ma il vettore u_{k+1} è certamente diverso dal vettore nullo. Infatti, se fosse $u_{k+1} = 0$, si avrebbe $v_{k+1} = w_{k+1} \in W_k$ e il sistema (v_1, \dots, v_{k+1}) sarebbe linearmente dipendente contro le ipotesi. Inoltre, per costruzione, il vettore u_{k+1} è ortogonale a tutti i precedenti vettori $u_1, \dots, u_k \in W_k$. Perciò (u_1, \dots, u_{k+1}) è un sistema ortogonale di vettori non nulli. Dal Lemma

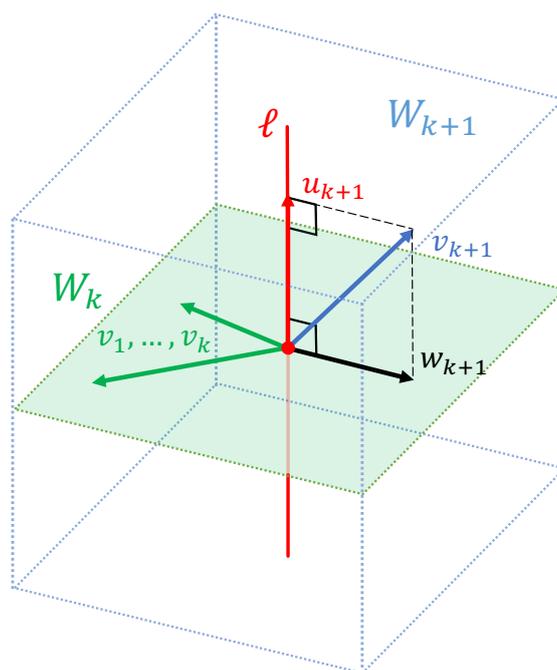


Figura 2.5: Ortogonalizzazione di Gram-Schmidt.

2.2.1 segue che (u_1, \dots, u_{k+1}) è linearmente indipendente e $\dim \text{Span}(u_1, \dots, u_{k+1}) = k + 1$ come desiderato. Questo dimostra il punto $(1)_{k+1}$. Per il punto $(2)_{k+1}$ calcoliamo

$$\langle u_{k+1}, v_{k+1} \rangle = \langle u_{k+1}, w_{k+1} + u_{k+1} \rangle = \langle u_{k+1}, w_{k+1} \rangle + \langle u_{k+1}, u_{k+1} \rangle = \|u_{k+1}\|^2,$$

in cui abbiamo usato la (2.12). Resta ancora da verificare che (u_1, \dots, u_{k+1}) è l'unico sistema ortogonale e indipendente che soddisfa (1) e (2) per ogni $q = 1, \dots, k + 1$. Per ipotesi di induzione, i vettori u_1, \dots, u_k sono univocamente determinati dalle condizioni (1) e (2) e dobbiamo solo dimostrare che u_{k+1} è l'unico vettore ortogonale a W_k che soddisfa $(1)_{k+1}$ e $(2)_{k+1}$. Sia dunque u'_{k+1} un altro vettore ortogonale a W_k tale che $\text{Span}(u_1, \dots, u_{k+1}) = W_{k+1}$ e $\langle u'_{k+1}, v_{k+1} \rangle = \|u'_{k+1}\|^2$. Osserviamo che u_{k+1} e u'_{k+1} sono entrambi nel complemento ortogonale ℓ del sottospazio W_k nello spazio vettoriale euclideo $(W_{k+1}, \langle -, - \rangle_{W_{k+1}})$ (verificarlo nei dettagli per esercizio). Giacché $\dim W_{k+1} = k + 1 = \dim W_k + 1$, segue che ℓ ha dimensione 1, e siccome $u_{k+1} \neq 0$ deve esistere $a \in \mathbb{R}$ tale che

$$u'_{k+1} = au_{k+1},$$

da cui, da un lato,

$$\|u'_{k+1}\|^2 = \langle u'_{k+1}, v_{k+1} \rangle = a \langle u_{k+1}, v_{k+1} \rangle = a \|u_{k+1}\|^2. \quad (2.13)$$

Dall'altro

$$\|u'_{k+1}\|^2 = \|au_{k+1}\|^2 = a^2 \|u_{k+1}\|^2 \quad (2.14)$$

Confrontando la (2.13) con la (2.14) si trova

$$a = a^2 \Rightarrow a = 1 \Rightarrow u'_{k+1} = u_{k+1}.$$

Questo conclude la dimostrazione della prima parte dell'enunciato.

Dobbiamo ora dimostrare la formula (2.11). Abbiamo già visto che $u_1 = v_1$ perciò la (2.11) è corretta per $q = 1$. Per $q > 1$, la formula (2.11) segue immediatamente dalla Proposizione 2.2.3, dalla (2.12) con $k = q - 1$, e dal fatto che w_q è esattamente la proiezione ortogonale di v_q sul sottospazio W_{q-1} . ■

Si osservi che la Formula (2.11) fornisce anche un algoritmo per calcolare il sistema (u_1, \dots, u_p) a partire dal sistema (v_1, \dots, v_p) . Tale algoritmo prende talvolta il nome di *procedura di ortogonalizzazione di Gram-Schmidt* o *algoritmo di Gram-Schmidt*.

■ **Esempio 2.6** In \mathbb{R}^3 con il prodotto scalare standard, consideriamo il sistema indipendente

$$S = (v_1, v_2, v_3) = \left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right).$$

Vogliamo applicare ad S la procedura di ortogonalizzazione di Gram-Schmidt. Al primo passo dell'algoritmo abbiamo

$$u_1 = v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Abbiamo anche

$$\langle u_1, u_1 \rangle_{can} = \langle v_1, v_1 \rangle_{can} = 3, \quad \langle u_1, v_2 \rangle_{can} = 2 \quad \text{e} \quad \langle u_1, v_3 \rangle_{can} = 0.$$

Perciò, al secondo passo dell'algoritmo troviamo

$$u_2 = v_2 - \frac{\langle u_1, v_2 \rangle_{can}}{\langle u_1, u_1 \rangle_{can}} u_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \frac{2}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1/3 \\ -2/3 \\ 1/3 \end{pmatrix}.$$

Inoltre

$$\langle u_2, u_2 \rangle_{can} = \frac{2}{3} \quad \text{e} \quad \langle u_2, u_3 \rangle_{can} = 0,$$

perciò, all'ultimo passo dell'algoritmo di Gram-Schmidt troviamo semplicemente

$$u_3 = v_3 - \frac{\langle u_1, v_3 \rangle_{can}}{\langle u_1, u_1 \rangle_{can}} u_1 - \frac{\langle u_2, v_3 \rangle_{can}}{\langle u_2, u_2 \rangle_{can}} u_2 = v_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Si noti che cambiando l'ordine dei vettori nel sistema iniziale, i vettori del sistema finale cambiano.

■

Esercizio 2.6 Si consideri la matrice

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

Dopo aver verificato che si tratta di una matrice invertibile, applicare la procedura di ortogonalizzazione di Gram-Schmidt al sistema (v_1, v_2, v_3) dell'Esempio 2.6 nello spazio vettoriale euclideo $(\mathbb{R}^3, \langle -, - \rangle_A)$ (si veda l'Esercizio 2.1 per la definizione del prodotto scalare $\langle -, - \rangle_A$). ■

Corollario 2.2.6 Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo n -dimensionale e sia $\mathcal{R} = (v_1, \dots, v_n)$ un riferimento di V . Allora esiste un unico riferimento ortonormale (e_1, \dots, e_n) tale che: per ogni $q = 1, \dots, n$

- (1) i vettori e_1, \dots, e_q generano lo stesso sottospazio dei vettori v_1, \dots, v_q ;
- (2) il prodotto scalare $\langle e_q, v_q \rangle$ è positivo.

I vettori e_1, \dots, e_n possono essere calcolati induttivamente mediante le formule

$$e_q = \frac{1}{\|u_q\|} u_q, \quad e \quad u_q = v_q - \sum_{i=1}^{q-1} \langle e_i, v_q \rangle e_i, \quad q = 1, \dots, n. \quad (2.15)$$

Dimostrazione. Sia (u_1, \dots, u_n) il sistema che si ottiene applicando la procedura di ortogonalizzazione di Gram-Schmidt ad \mathcal{R} . Ovviamente, (u_1, \dots, u_n) è un riferimento ortogonale. Poniamo

$$e_i = \frac{1}{\|u_i\|} u_i, \quad i = 1, \dots, n.$$

I vettori e_1, \dots, e_n sono, in effetti, versori, infatti

$$\|e_i\| = \left\| \frac{1}{\|u_i\|} u_i \right\| = \frac{\|u_i\|}{\|u_i\|} = 1, \quad i = 1, \dots, n$$

(si dice anche che e_i è il versore che si ottiene da u_i *normalizzando*). Inoltre i vettori e_i sono ancora a due a due ortogonali (il lettore lo vede?). Perciò (e_1, \dots, e_n) è un riferimento ortonormale. È anche chiaro che (e_1, \dots, e_n) soddisfa tutte le proprietà richieste (verificare i dettagli per esercizio). L'esistenza è così provata.

Per quanto riguarda l'unicità, osserviamo che, per ogni q , le condizioni nell'enunciato impongono che e_q appartenga al complemento ortogonale ℓ del sottospazio $W_{q-1} = \text{Span}(v_1, \dots, v_{q-1})$ nello spazio vettoriale euclideo $(W_q, \langle -, - \rangle_{W_q})$. Ma ℓ è un sottospazio 1-dimensionale perciò contiene solo due versori, uno l'opposto dell'altro, e solo uno dei due può soddisfare la (2). Ne deduciamo che e_q è completamente determinato da tutte le condizioni richieste. Ne consegue l'unicità.

Infine la Formula (2.15) segue facilmente dalla (2.11) (verificare i dettagli per esercizio). ■

- La dimostrazione del Corollario 2.2.6 non ricorre al teorema di esistenza delle basi ortogonali rispetto ad una forma bilineare simmetrica (né direttamente, né indirettamente) perciò, almeno nel caso di una forma bilineare simmetrica definita positiva (cioè un prodotto scalare), si può anche riguardare come una dimostrazione alternativa dell'esistenza di basi ortonormali.

Esercizio 2.7 Sia $S = (v_1, v_2, v_3)$ il sistema (in $(\mathbb{R}^3, \langle -, - \rangle_{can})$) considerato all'Esempio 2.6. Dopo aver osservato che S è, in effetti, un riferimento, utilizzare la Formula (2.15) per costruire un riferimento ortonormale a partire da S . ■

Esercizio 2.8 Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo di dimensione n e sia $S = (e_1, \dots, e_p)$ un sistema ortogonale di versori, $p \leq n$. Dimostrare che è possibile estendere S ad un riferimento ortonormale, cioè esistono vettori e_{p+1}, \dots, e_n tali che il sistema $(e_1, \dots, e_p, e_{p+1}, \dots, e_n)$ è un riferimento ortonormale. ■

2.3 Il Gruppo Ortogonale

In questa sezione studiamo le matrici del cambio di coordinate tra riferimenti ortonormali. Scopriremo che tali matrici formano un sottogruppo notevole nel gruppo lineare generale e impareremo a caratterizzarle.

Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo di dimensione n , siano $\mathcal{R} = (e_1, \dots, e_n)$ ed $\mathcal{R}' = (e'_1, \dots, e'_n)$ riferimenti ortonormali di V , e sia $A = M_{\mathcal{R}', \mathcal{R}} = (a_{ij})_{i=1, \dots, n}^{j=1, \dots, n}$ la matrice del cambio di coordinate da \mathcal{R} ad \mathcal{R}' . La j -esima colonna $A^{(j)}$ di A è il vettore coordinato di e_j nel riferimento \mathcal{R}' . Siccome \mathcal{R}' è un riferimento ortonormale, si ha

$$\langle e_i, e_j \rangle = \langle A^{(i)}, A^{(j)} \rangle_{can} = A_{(i)}^T A^{(j)},$$

l'elemento di posto (i, j) della matrice prodotto $A^T A$. D'altro canto, anche \mathcal{R} è un riferimento ortonormale, perciò $\langle e_i, e_j \rangle = \delta_{ij}$, l'elemento di posto (i, j) della matrice identica I_n . Abbiamo così scoperto che

$$A^T A = I_n. \quad (2.16)$$

Definizione 2.3.1 — Matrice Ortogonale. Una *matrice ortogonale* di ordine n è una matrice quadrata $A \in M_n(\mathbb{R})$ tale che $A^T A = I_n$.

La (2.16) dice perciò che *la matrice del cambio di coordinate tra due riferimenti ortonormali è una matrice ortogonale*. Vale anche il viceversa, nel senso che, dato un riferimento ortonormale \mathcal{R}' e un altro riferimento $\tilde{\mathcal{R}} = (\tilde{e}_1, \dots, \tilde{e}_n)$, se la matrice del cambio di coordinate $B = M_{\mathcal{R}', \tilde{\mathcal{R}}}$ è ortogonale, allora anche il riferimento $\tilde{\mathcal{R}}$ è ortonormale. Infatti, similmente a quanto visto sopra, in questa situazione abbiamo

$$\delta_{ij} = B_{(i)}^T B^{(j)} = \langle B^{(i)}, B^{(j)} \rangle_{can} = \langle \tilde{e}_i, \tilde{e}_j \rangle$$

dove, nell'ultimo passaggio, abbiamo utilizzato che \mathcal{R}' è un riferimento ortonormale. Riassumendo, abbiamo dimostrato il seguente

Teorema 2.3.1 Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo di dimensione n e sia \mathcal{R}' un riferimento ortonormale di V . Allora, un altro riferimento \mathcal{R} è ortonormale sse la matrice del cambio di coordinate $M_{\mathcal{R}', \mathcal{R}}$ è una matrice ortogonale.

Studiamo ora alcune proprietà elementari delle matrici ortogonali.

Proposizione 2.3.2 Sia $A \in M_n(\mathbb{R})$ una matrice reale quadrata di ordine n . Allora le seguenti asserzioni sono equivalenti:

- (1) A è una matrice ortogonale,
- (2) le colonne di A formano un riferimento ortonormale nello spazio vettoriale euclideo standard $(\mathbb{R}^n, \langle -, - \rangle_{can})$,
- (3) (le trasposte del)le righe di A formano un riferimento ortonormale nello spazio vettoriale euclideo standard $(\mathbb{R}^n, \langle -, - \rangle_{can})$.

Dimostrazione. Dimostriamo innanzitutto che (1) \Leftrightarrow (2). Questo segue da un conto identico a quelli fatti poco più sopra e che ripetiamo per la convenienza del lettore: la matrice A è ortogonale sse $A^T A = I_n$. In termini di entrate questa condizione si scrive:

$$\delta_{ij} = A_{(i)}^T A^{(j)} = \langle A^{(i)}, A^{(j)} \rangle_{can}.$$

Dunque A è una matrice ortogonale sse il sistema $S = (A^{(1)}, \dots, A^{(n)})$ è un sistema ortogonale di versori in $(\mathbb{R}^n, \langle -, - \rangle_{can})$. In questo caso, poiché ogni versore è un vettore non nullo, il sistema S è anche indipendente, perciò è un riferimento (ortonormale).

Ora una matrice A è ortogonale sse la matrice A^T è ortogonale (il lettore è d'accordo?). Da questo e dall'equivalenza (1) \Leftrightarrow (2) segue immediatamente anche l'equivalenza (1) \Leftrightarrow (3). ■

■ **Esempio 2.7** Per ogni numero reale θ , la matrice 2×2

$$R_\theta := \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

è una matrice ortogonale (il lettore lo dimostri per esercizio). La matrice R_θ è qualche volta detta *matrice di rotazione* (di angolo θ). Ritorniamo su questa terminologia nella prossima sezione. Anche la matrice

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

è ovviamente ortogonale. ■

Teorema / Definizione 2.3.3 — Gruppo Ortogonale. Le matrici ortogonali di ordine n costituiscono un sottogruppo del gruppo lineare generale $GL_n(\mathbb{R})$, detto *gruppo ortogonale di ordine n* , che si indica con O_n .

Dimostrazione. Chiaramente, la matrice identica I_n è ortogonale. Inoltre, segue immediatamente dalla definizione, che ogni matrice ortogonale A è invertibile e $A^{-1} = A^T$, sicché anche l'inversa di A è una matrice ortogonale. Resta da dimostrare che il prodotto AB di due matrici ortogonali A e B è ancora una matrice ortogonale. A questo scopo calcoliamo

$$(AB)^T AB = B^T A^T AB = B^T I_n B = B^T B = I_n.$$

Questo conclude la dimostrazione. ■

Proposizione 2.3.4 Sia $A \in M_n(\mathbb{R})$ una matrice ortogonale. Allora $\det A = \pm 1$.

Dimostrazione. Per il Teorema di Binet si ha

$$1 = \det I_n = \det(A^T A) = \det A^T \det A = (\det A)^2.$$

■

Si osservi che una matrice con determinante pari ad 1 non è necessariamente ortogonale. Una matrice ortogonale con determinante pari ad 1 si dice *matrice ortogonale speciale*. La matrice identica è una matrice ortogonale speciale. Inoltre l'inversa di una matrice ortogonale speciale è ancora una matrice ortogonale speciale. Infine, per il Teorema di Binet, il prodotto di matrici ortogonali speciali è sempre una matrice ortogonale speciale (il lettore lo vede?). Questo mostra che le matrici ortogonali speciali di ordine n formano un sottogruppo di O_n (e quindi di $GL_n(\mathbb{R})$). Tale sottogruppo si indica con SO_n e prende il nome di *gruppo ortogonale speciale di ordine n* .

Concludiamo questa sezione descrivendo in dettaglio tutte le matrici ortogonali di ordine 2×2 . Cominciamo con l'osservare che, per ogni intero positivo n , la matrice

$$U := \begin{pmatrix} -1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

è una matrice ortogonale. Inoltre $\det U = -1$. Ogni matrice ortogonale A di ordine n e determinante pari a -1 è il prodotto di U per una matrice ortogonale speciale. Infatti, poniamo $A' = U^{-1}A$. Ovviamente A' è una matrice ortogonale, inoltre

$$\det A' = \det(U^{-1}A) = \det U^{-1} \det A = (\det U)^{-1} \det A = -1 \cdot (-1) = 1.$$

Dunque A' è una matrice ortogonale speciale e

$$A = UA',$$

come annunciato. Perciò se vogliamo descrivere le matrici ortogonali (di ordine 2) è sufficiente descrivere le matrici ortogonali speciali (di ordine 2) e poi moltiplicare per la matrice U (con $n = 2$).

Proposizione 2.3.5 Ogni matrice ortogonale speciale di ordine 2 è una matrice di rotazione e viceversa, cioè, una matrice $A \in M_2(\mathbb{R})$ è ortogonale speciale sse esiste $\theta \in \mathbb{R}$ tale che

$$A = R_\theta := \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

Dimostrazione. Già sappiamo che ogni matrice di rotazione è ortogonale. D'altronde, per ogni $\theta \in \mathbb{R}$,

$$\det R_\theta = \det \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} = \cos^2 \theta + \sin^2 \theta = 1.$$

Quindi R_θ è una matrice ortogonale speciale. Viceversa, sia A una matrice ortogonale speciale. Vogliamo dimostrare che A è una matrice di rotazione. A questo scopo sono opportune due premesse. Innanzitutto sia $v = (a, b)^T$ un vettore nello spazio vettoriale euclideo standard $(\mathbb{R}^2, \langle -, - \rangle_{can})$. Se $v \neq 0$, il complemento ortogonale v^\perp è un sottospazio di dimensione 1, ed è generato dal vettore $w = (-b, a)$ (Figura 2.6). Infatti $w \neq 0$ e inoltre

$$\langle w, v \rangle_{can} = -ba + ab = 0,$$

cioè $w \perp v$.

La seconda premessa è la seguente: i versori in $(\mathbb{R}^2, \langle -, - \rangle_{can})$ sono i vettori sulla *circonferenza unitaria*

$$S^1 := \{(a, b)^T \in \mathbb{R}^2 : a^2 + b^2 = 1\} \subseteq \mathbb{R}^2.$$

Perciò, se $v \in \mathbb{R}^2$ è un versore, allora esiste un numero reale $\theta \in \mathbb{R}$ tale che $v = (\cos \theta, \sin \theta)^T$ (Figura 2.7).

Ora sia $A \in SO_2$. Le colonne $A^{(1)}, A^{(2)}$ di A sono versori in $(\mathbb{R}^2, \langle -, - \rangle_{can})$. In particolare, esiste $\theta \in \mathbb{R}$ tale che $A^{(1)} = (\cos \theta, \sin \theta)^T$. Inoltre $A^{(2)}$ è ortogonale ad $A^{(1)}$. Esistono esattamente 2 versori ortogonali ad $A^{(1)}$ e, in base alle considerazioni nella prima premessa, sono $(-\sin \theta, \cos \theta)^T$ e il suo opposto $(\sin \theta, -\cos \theta)^T$. Perciò A può essere solo una delle due matrici

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad \text{oppure} \quad \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}.$$

La prima matrice è la matrice di rotazione R_θ , mentre la seconda è una matrice ortogonale sì, ma ha determinante pari a -1 (il lettore lo vede?). Concludiamo che $A = R_\theta$ (Figura 2.8). ■

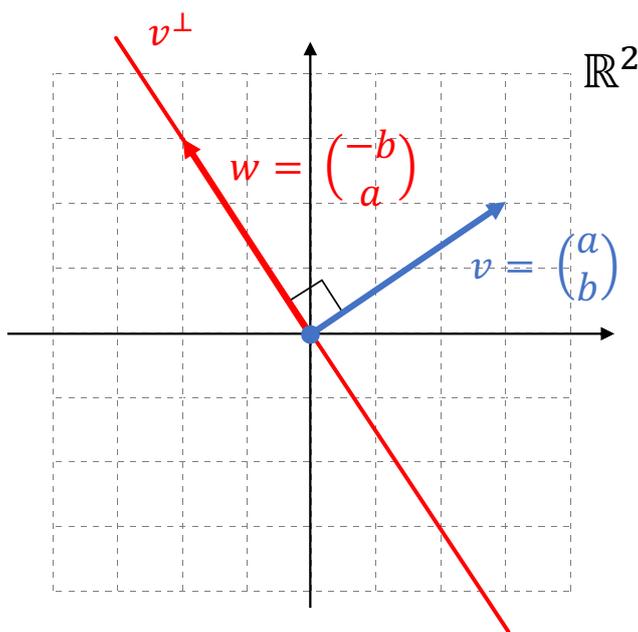


Figura 2.6: Complemento ortogonale di un vettore in $(\mathbb{R}^2, \langle -, - \rangle_{can})$.

2.4 Applicazioni Ortogonali e Isometrie

In questa sezione introduciamo una classe di applicazioni adatte a confrontare due spazi vettoriali euclidei. Dunque siano $(V, \langle -, - \rangle)$ e $(V', \langle -, - \rangle')$ spazi vettoriali euclidei di dimensione n ed n' rispettivamente. Indicheremo con $\|-\|, \|-\|'$ e d, d' rispettivamente le norme e le distanze in V, V' .

Definizione 2.4.1 — Applicazione Ortogonale. Un'applicazione ortogonale tra $(V, \langle -, - \rangle)$ e $(V', \langle -, - \rangle')$ è un'applicazione lineare

$$f : V \rightarrow V'$$

che, in aggiunta, *conserva i prodotti scalari*, cioè, per ogni $v, w \in V$,

$$\langle f(v), f(w) \rangle' = \langle v, w \rangle.$$

Un'applicazione ortogonale biettiva è anche detta *isomorfismo ortogonale* (se $(V', \langle -, - \rangle') = (V, \langle -, - \rangle)$ diciamo anche *automorfismo ortogonale*).

Nel seguito, dati due spazi vettoriali euclidei $(V, \langle -, - \rangle)$ e $(V', \langle -, - \rangle')$ e un'applicazione lineare $f : V \rightarrow V'$ (a priori non-necessariamente ortogonale) che vogliamo usare per confrontare $(V, \langle -, - \rangle)$ e $(V', \langle -, - \rangle')$, se vorremo insistere che su V e V' consideriamo proprio i prodotti scalari $\langle -, - \rangle$ e $\langle -, - \rangle'$ (e non altri due prodotti scalari), allora scriveremo anche $f : (V, \langle -, - \rangle) \rightarrow (V', \langle -, - \rangle')$.

■ **Esempio 2.8** Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo di dimensione n e sia $\mathcal{R} = (e_1, \dots, e_n)$ un riferimento ortonormale. L'isomorfismo coordinato

$$c_{\mathcal{R}} : (V, \langle -, - \rangle) \rightarrow (\mathbb{R}^n, \langle -, - \rangle_{can})$$

(che associa ad ogni vettore $v \in V$ il suo vettore coordinato $c_{\mathcal{R}}(v)$ nel riferimento \mathcal{R}) è un'applicazione ortogonale. Infatti, siano $v, v' \in V$ e siano $x = c_{\mathcal{R}}(v), x' = c_{\mathcal{R}}(v')$ i vettori coordinati di v, v' nel

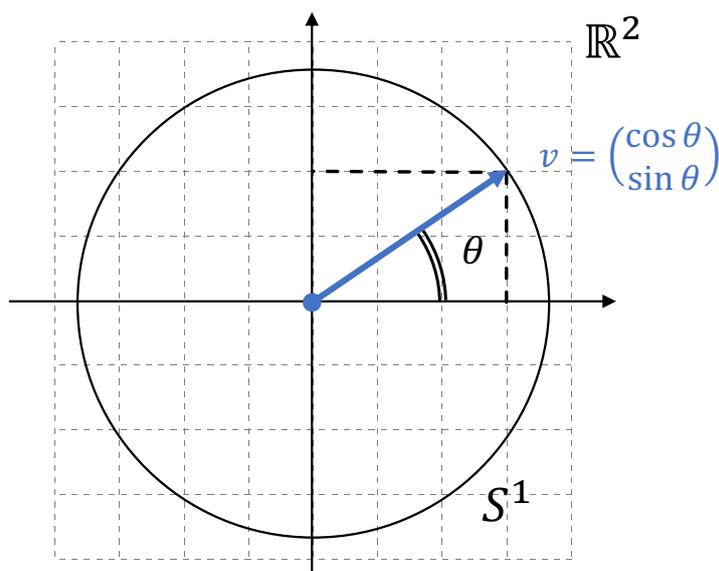


Figura 2.7: Vettori sulla circonferenza unitaria.

riferimento \mathcal{R} . Dalla (2.8) abbiamo che

$$\langle c_{\mathcal{R}}(v), c_{\mathcal{R}}(v') \rangle_{can} = \langle x, x' \rangle_{can} = \langle v, v' \rangle,$$

per ogni v, v' . ■

Prima di fare qualche altro esempio, studiamo le proprietà fondamentali delle applicazioni ortogonali.

Proposizione 2.4.1

- (1) L'applicazione identica $\text{id} : (V, \langle -, - \rangle) \rightarrow (V, \langle -, - \rangle)$ è un'applicazione ortogonale;
- (2) la composta di applicazioni ortogonali è un'applicazione ortogonale, cioè se $(V'', \langle -, - \rangle')$ è un terzo spazio vettoriale euclideo e $f : (V, \langle -, - \rangle) \rightarrow (V', \langle -, - \rangle')$ e $g : (V', \langle -, - \rangle') \rightarrow (V'', \langle -, - \rangle'')$ sono applicazioni ortogonali, allora anche

$$g \circ f : (V, \langle -, - \rangle) \rightarrow (V'', \langle -, - \rangle'')$$

è un'applicazione ortogonale;

- (3) l'inversa $f^{-1} : (V', \langle -, - \rangle') \rightarrow (V, \langle -, - \rangle)$ di un isomorfismo ortogonale $f : (V, \langle -, - \rangle) \rightarrow (V', \langle -, - \rangle')$ è un isomorfismo ortogonale;
- (4) l'inclusione $i_W : (W, \langle -, - \rangle_W) \hookrightarrow (V, \langle -, - \rangle)$ di un sottospazio vettoriale euclideo è un'applicazione ortogonale;
- (5) ogni applicazione ortogonale *conserva le norme*, cioè se $f : (V, \langle -, - \rangle) \rightarrow (V', \langle -, - \rangle')$ è un'applicazione ortogonale, allora

$$\|f(v)\|' = \|v\|$$

per ogni $v \in V$;

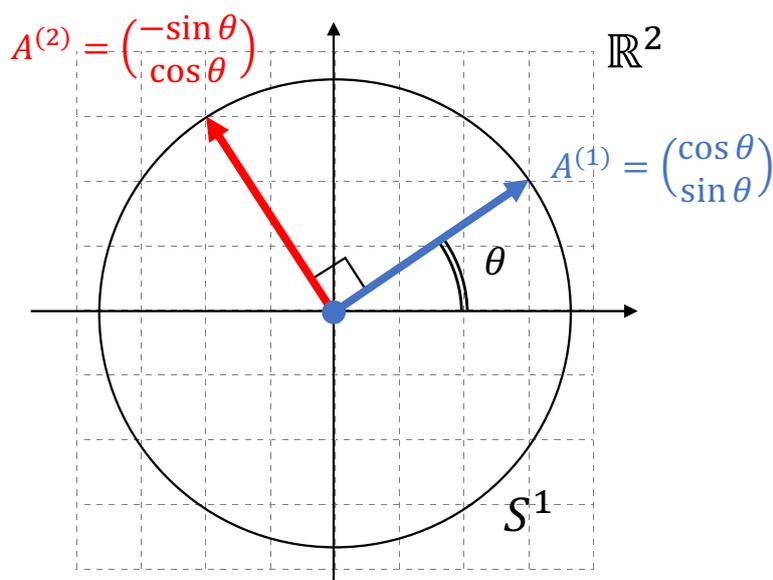


Figura 2.8: Matrice di rotazione.

- (6) ogni applicazione ortogonale *conserva le distanze*, cioè se $f : (V, \langle -, - \rangle) \rightarrow (V', \langle -, - \rangle')$ è un'applicazione ortogonale, allora

$$d'(f(v), f(w)) = d(v, w)$$

per ogni $v, w \in V$;

- (7) ogni applicazione ortogonale è iniettiva;
 (8) ogni applicazione ortogonale $f : (V, \langle -, - \rangle) \rightarrow (V', \langle -, - \rangle')$ è la composta di un isomorfismo ortogonale $f_W : (V, \langle -, - \rangle) \rightarrow (W, \langle -, - \rangle'_W)$ con l'inclusione $i_W : (W, \langle -, - \rangle'_W) \hookrightarrow (V', \langle -, - \rangle')$ di un sottospazio vettoriale euclideo.
 (9) ogni endomorfismo ortogonale $f : (V, \langle -, - \rangle) \rightarrow (V, \langle -, - \rangle)$ è un automorfismo ortogonale.

Dimostrazione. I punti (1)–(4) sono facili da verificare e vengono lasciati come Esercizio 2.9. Discutiamo la (5) e la (6). Siano $v, w \in V$, allora

$$\|f(v)\|' = \sqrt{\langle f(v), f(v) \rangle'} = \sqrt{\langle v, v \rangle} = \|v\|.$$

Similmente

$$d'(f(v), f(w)) = \|f(v) - f(w)\|' = \|f(v - w)\|' = \|v - w\| = d(v, w),$$

dove abbiamo usato anche il punto (5). Per la (7) dimostriamo che ogni applicazione ortogonale $f : (V, \langle -, - \rangle) \rightarrow (V', \langle -, - \rangle')$ ha nucleo $\ker f \subseteq V$ banale. Dunque sia $v \in \ker f$. Questo vuol dire che $f(v) = 0$. Allora

$$0 = \|f(v)\|' = \|v\| \quad \Rightarrow \quad v = 0,$$

come desiderato.

Per il punto (8), sia $f : (V, \langle -, - \rangle) \rightarrow (V', \langle -, - \rangle')$ un'applicazione ortogonale. Poniamo $W := \text{im } f$ e consideriamo il prodotto scalare ristretto $\langle -, - \rangle'_W$. Per costruzione, l'applicazione lineare $f : V \rightarrow V'$ si può restringere a W nel codominio. Diciamo $f_W : V \rightarrow W$ l'applicazione ristretta.

Ovviamente f_W è ancora un'applicazione lineare. Vogliamo ora dimostrare che $f_W : (V, \langle -, - \rangle) \rightarrow (W, \langle -, - \rangle'_W)$ è un'applicazione ortogonale. A questo scopo, per ogni $v, u \in V$ calcoliamo

$$\langle f_W(v), f_W(u) \rangle'_W = \langle f(v), f(u) \rangle = \langle v, u \rangle.$$

Ora f_W , per costruzione, è un'applicazione suriettiva. Giacché è un'applicazione ortogonale, per il punto (7), è anche iniettiva, perciò si tratta di un isomorfismo ortogonale. È anche chiaro che $f : V \rightarrow V'$ è la composta di f_W seguita dall'inclusione $W \hookrightarrow V'$. Questo conclude la dimostrazione del punto (8).

Il punto (9) è chiaro: dato un endomorfismo ortogonale $f : (V, \langle -, - \rangle) \rightarrow (V, \langle -, - \rangle)$, per il punto (7), f è un'applicazione iniettiva. Ma è un'applicazione lineare tra spazi vettoriali della stessa dimensione, perciò, dall'iniettività, segue la biiettività, cioè f è un isomorfismo (in effetti un automorfismo) ortogonale. ■

Esercizio 2.9 Dimostrare i punti (1)–(4) della Proposizione 2.4.1. ■

Corollario / Definizione 2.4.2 Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo. Gli automorfismi ortogonali $f : (V, \langle -, - \rangle) \rightarrow (V, \langle -, - \rangle)$ di V formano un sottogruppo del gruppo lineare generale $GL(V)$, che prende il nome di *gruppo ortogonale* di $(V, \langle -, - \rangle)$ e si indica con $O(V, \langle -, - \rangle)$.

Dimostrazione. Segue immediatamente dai punti (1)–(3) della Proposizione 2.4.1 (il lettore lo vede?). ■

In virtù della Proposizione 2.4.1.(7), nello studiare le applicazioni ortogonali, possiamo limitarci agli isomorfismi ortogonali, ed è quello che faremo per il resto di questa sezione. Gli isomorfismi ortogonali possono essere caratterizzati in vari modi come mostra la seguente

Proposizione 2.4.3 Siano $(V, \langle -, - \rangle)$ e $(V', \langle -, - \rangle')$ spazi vettoriali euclidei della stessa dimensione n , e sia $f : V \rightarrow V'$ un'applicazione lineare. Le seguenti condizioni sono equivalenti:

- (1) f è un isomorfismo ortogonale;
- (2) f conserva le norme;
- (3) f conserva le distanze;
- (4) per ogni riferimento ortormale $\mathcal{R} = (e_1, \dots, e_n)$ di $(V, \langle -, - \rangle)$, il sistema $f(\mathcal{R}) = (f(e_1), \dots, f(e_n))$ è un riferimento ortonormale di $(V', \langle -, - \rangle')$;
- (5) esiste un riferimento ortonormale $\mathcal{R} = (e_1, \dots, e_n)$ di $(V, \langle -, - \rangle)$ tale che il sistema $f(\mathcal{R}) = (f(e_1), \dots, f(e_n))$ è un riferimento ortonormale di $(V', \langle -, - \rangle')$;
- (6) per ogni riferimento ortonormale $\mathcal{R} = (e_1, \dots, e_n)$ di $(V, \langle -, - \rangle)$ e per ogni riferimento ortonormale $\mathcal{R}' = (e'_1, \dots, e'_n)$ di $(V', \langle -, - \rangle')$, la matrice rappresentativa $M_{\mathcal{R}', \mathcal{R}}^f$ è una matrice ortogonale;
- (7) esistono un riferimento ortonormale $\mathcal{R} = (e_1, \dots, e_n)$ di $(V, \langle -, - \rangle)$ e un riferimento ortonormale $\mathcal{R}' = (e'_1, \dots, e'_n)$ di $(V', \langle -, - \rangle')$, tali che la matrice rappresentativa $M_{\mathcal{R}', \mathcal{R}}^f$ è una matrice ortogonale.

○ Prima di dimostrare la Proposizione 2.4.3 ricordiamo la definizione e le proprietà principali della matrice rappresentativa di un'applicazione lineare. Innanzitutto, data un'applicazione lineare $f : V \rightarrow W$ e due riferimenti $\mathcal{R} = (v_1, \dots, v_n)$ e $\mathcal{S} = (w_1, \dots, w_m)$ di V e W rispettivamente, per definizione di matrice rappresentativa, $M_{\mathcal{S}, \mathcal{R}}^f = (a_{ij})_{i=1, \dots, m}^{j=1, \dots, n}$ è l'unica matrice di ordine $m \times n$ con la proprietà che, per ogni $v \in V$, detto x il vettore coordinato di v nel riferimento \mathcal{R} , e y il vettore coordinato di $f(v)$ nel riferimento \mathcal{S} , si ha che

$$y = M_{\mathcal{S}, \mathcal{R}}^f x.$$

In questo senso, la matrice $M_{S,\mathcal{R}}^f$ rappresenta l'applicazione lineare f in coordinate. La j -esima colonna di $M_{S,\mathcal{R}}^f$ è il vettore coordinato di $f(e_j)$ nel riferimento \mathcal{S} . In altre parole

$$f(e_j) = \sum_{i=1}^n a_{ij} e'_i, \quad j = 1, \dots, n,$$

(dove, come indicato sopra, le a_{ij} sono esattamente le entrate di $M_{S,\mathcal{R}}^f$). Ricordiamo anche per futuri utilizzi che, se \mathcal{R}' ed \mathcal{S}' sono altri due riferimenti, il primo di V e il secondo di W , si ha

$$M_{S',\mathcal{R}'}^f = M_{S',\mathcal{S}} M_{S,\mathcal{R}}^f M_{\mathcal{R},\mathcal{R}'}.$$

Inoltre, la matrice rappresentativa di un isomorfismo $\varphi : V \rightarrow V'$ rispetto ad un riferimento \mathcal{R} di V e al riferimento immagine $\varphi(\mathcal{R})$ di V' è la matrice identica:

$$M_{\varphi(\mathcal{R}),\mathcal{R}}^\varphi = I_n,$$

in particolare, la matrice rappresentativa dell'applicazione identica $\text{id}_V : V \rightarrow V$ in un qualunque riferimento è la matrice identica:

$$M_{\mathcal{R}}^{\text{id}_V} = I.$$

Se $g : W \rightarrow U$ è un'altra applicazione lineare e \mathcal{T} è un riferimento di U ,

$$M_{\mathcal{T},\mathcal{R}}^{g \circ f} = M_{\mathcal{T},\mathcal{S}}^g M_{S,\mathcal{R}}^f.$$

Infine, se $f : V \rightarrow W$ è un isomorfismo

$$M_{\mathcal{R},\mathcal{S}}^{f^{-1}} = \left(M_{S,\mathcal{R}}^f \right)^{-1}.$$

Da queste ultime osservazioni segue in particolare che, per ogni riferimento \mathcal{R} di V , l'applicazione $GL(V) \rightarrow GL_n(\mathbb{K})$, $f \mapsto M_{\mathcal{R}}^f$ che associa all'automorfismo f la sua matrice rappresentativa nel riferimento \mathcal{R} , è un isomorfismo di gruppi.

Dimostrazione della Proposizione 2.4.3. Concentriamoci innanzitutto sui primi tre punti. Già sappiamo che (1) \Rightarrow (2), (3) (Proposizione 2.4.1 Punti (5) e (6)). Per dimostrare che (2) \Rightarrow (1) osserviamo che il campo dei numeri reali \mathbb{R} ha caratteristica $0 \neq 2$. Perciò per un prodotto scalare e la sua norma quadra vale la Formula di Polarizzazione (Proposizione 1.3.3): per ogni $v, w \in V$

$$\langle v, w \rangle = \frac{\|v+w\|^2 - \|v\|^2 - \|w\|^2}{2}. \quad (2.17)$$

Ora supponiamo che l'applicazione lineare $f : V \rightarrow V'$ conservi la norma. In tal caso, per ogni $v, w \in V$

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

Dunque f conserva anche i prodotti scalari. Per dimostrare che (3) \Rightarrow (1) supponiamo che f conservi le distanze. Allora, per ogni $v \in V$

$$\|f(v)\|' = \|f(v) - 0\|' = d'(f(v), 0) = d'(f(v), f(0)) = d(v, 0) = \|v - 0\| = \|v\|.$$

Dunque f conserva le norme e quindi, giacché $(2) \Rightarrow (1)$, anche i prodotti scalari. Questo dimostra che $(1) \Leftrightarrow (2) \Leftrightarrow (3)$.

Occupiamoci ora dei successivi punti. Per i punti (4) e (5) dimostreremo che $(1) \Rightarrow (4) \Rightarrow (5) \Rightarrow (1)$. Dunque sia $f : (V, \langle -, - \rangle) \rightarrow (V', \langle -, - \rangle')$ un'applicazione ortogonale e sia $\mathcal{R} = (e_1, \dots, e_n)$ un qualunque riferimento ortonormale. Poiché f è un isomorfismo, anche il sistema $f(\mathcal{R}) = (f(e_1), \dots, f(e_n))$ è un riferimento. Inoltre, per ogni $i, j = 1, \dots, n$

$$\langle f(e_i), f(e_j) \rangle' = \langle e_i, e_j \rangle = \delta_{ij}$$

perché \mathcal{R} è un riferimento ortonormale. Perciò anche $f(\mathcal{R})$ è un riferimento ortonormale. Questo dimostra che $(1) \Rightarrow (4)$. Chiaramente $(4) \Rightarrow (5)$ e dobbiamo dimostrare ancora che $(5) \Rightarrow (1)$. Dunque sia $\mathcal{R} = (e_1, \dots, e_n)$ un riferimento ortonormale di $(V, \langle -, - \rangle)$ come al punto (5) dell'enunciato, e siano $v_1, v_2 \in V$. È chiaro che la matrice rappresentativa di f nei riferimenti $\mathcal{R}, f(\mathcal{R})$ è la matrice identica: $M_{f(\mathcal{R}), \mathcal{R}}^f$. Perciò, detti x_1, x_2 i vettori coordinati di v_1, v_2 nel riferimento \mathcal{R} , allora x_1, x_2 sono anche i vettori coordinati di $f(v_1), f(v_2)$ nel riferimento $f(\mathcal{R})$ e si ha

$$\langle f(v_1), f(v_2) \rangle' = \langle x_1, x_2 \rangle_{can} = \langle v_1, v_2 \rangle,$$

dove, nel primo passaggio, abbiamo usato che $f(\mathcal{R})$ è un riferimento ortonormale, mentre, nel secondo, abbiamo usato che \mathcal{R} è un riferimento ortonormale. Dunque f conserva i prodotti scalari. Questo dimostra che $(1) \Leftrightarrow (4) \Leftrightarrow (5)$.

Dobbiamo solo occuparci dei punti (6) e (7). Dimostreremo che $(1) \Rightarrow (6) \Rightarrow (7) \Rightarrow (1)$ e questo concluderà la dimostrazione della proposizione. Dunque sia $f : (V, \langle -, - \rangle) \rightarrow (V', \langle -, - \rangle')$ un isomorfismo ortogonale e siano $\mathcal{R} = (e_1, \dots, e_n)$ ed $\mathcal{R}' = (e'_1, \dots, e'_n)$ riferimenti ortonormali di $(V, \langle -, - \rangle)$ e $(V', \langle -, - \rangle')$ rispettivamente. Indichiamo con $A = M_{\mathcal{R}', \mathcal{R}}^f$ la matrice rappresentativa di f nei riferimenti $\mathcal{R}, \mathcal{R}'$. Per dimostrare che A è una matrice ortogonale, mostreremo che le sue colonne $A^{(1)}, \dots, A^{(n)}$ sono versori ortogonali nello spazio vettoriale euclideo standard $(\mathbb{R}^n, \langle -, - \rangle_{can})$. Ricordiamo che, per ogni $j = 1, \dots, n$, $A^{(j)}$ è il vettore coordinato di $f(e_j)$ nel riferimento \mathcal{R}' . Giacché \mathcal{R}' è un riferimento ortonormale, per ogni $i, j = 1, \dots, n$ si ha

$$\langle A^{(i)}, A^{(j)} \rangle_{can} = \langle f(e_i), f(e_j) \rangle' = \langle e_i, e_j \rangle = \delta_{ij},$$

come desiderato. È chiaro che $(6) \Rightarrow (7)$ e resta da dimostrare che $(7) \Rightarrow (1)$. Siano $\mathcal{R} = (e_1, \dots, e_n), \mathcal{R}' = (e'_1, \dots, e'_n)$ riferimenti ortonormali di $(V, \langle -, - \rangle), (V', \langle -, - \rangle')$ come al punto (7) dell'enunciato, e siano $v_1, v_2 \in V$. Indichiamo con $A = M_{\mathcal{R}', \mathcal{R}}^f$ la matrice rappresentativa di f nei riferimenti $\mathcal{R}, \mathcal{R}'$. Finalmente, detti x_1, x_2 i vettori coordinati di v_1, v_2 nel riferimento \mathcal{R} , e y_1, y_2 i vettori coordinati di $f(v_1), f(v_2)$ nel riferimento $f(\mathcal{R})$, calcoliamo

$$\begin{aligned} \langle f(v_1), f(v_2) \rangle' &= \langle y_1, y_2 \rangle_{can} && (\mathcal{R}' \text{ è un riferimento ortonormale}) \\ &= \langle Ax_1, Ax_2 \rangle_{can} && (A \text{ è la matrice rappresentativa di } f) \\ &= (Ax_1)^T Ax_2 && (\text{definizione di prodotto scalare standard}) \\ &= x_1^T A^T Ax_2 \\ &= x_1^T I_n x_2 && (A \text{ è una matrice ortogonale}) \\ &= x_1^T x_2 \\ &= \langle x_1, x_2 \rangle_{can} && (\text{definizione di prodotto scalare standard}) \\ &= \langle v_1, v_2 \rangle. && (\mathcal{R} \text{ è un riferimento ortonormale}) \end{aligned}$$

Dunque f conserva i prodotti scalari. ■

Corollario 2.4.4 Siano $(V, \langle -, - \rangle)$ e $(V', \langle -, - \rangle')$ spazi vettoriali euclidei della stessa dimensione n , e siano \mathcal{R} e \mathcal{R}' riferimenti ortonormali di $(V, \langle -, - \rangle)$ e $(V', \langle -, - \rangle')$ rispettivamente. Esiste un unico isomorfismo ortogonale $f : (V, \langle -, - \rangle) \rightarrow (V', \langle -, - \rangle')$ tale che $f(\mathcal{R}) = \mathcal{R}'$.

Dimostrazione. Per il *Teorema dell'Estensione Lineare* esiste un'unica applicazione lineare $f : V \rightarrow V'$ tale che $f(\mathcal{R}) = \mathcal{R}'$. Poiché f trasforma un riferimento ortonormale in un riferimento ortonormale, f è necessariamente un isomorfismo ortogonale. ■

○ Due spazi vettoriali euclidei $(V, \langle -, - \rangle)$ e $(V', \langle -, - \rangle')$ si dicono *isomorfi* se esiste un isomorfismo ortogonale $f : (V, \langle -, - \rangle) \rightarrow (V', \langle -, - \rangle')$ che li collega. Due spazi vettoriali euclidei isomorfi andrebbero identificati perché l'isomorfismo ortogonale ne conserva tutte le proprietà rilevanti. Il Corollario 2.4.4 (combinato con l'esistenza di riferimenti ortonormali) dimostra in particolare che due spazi vettoriali euclidei sono isomorfi sse hanno la stessa dimensione. Dunque *tutti gli spazi vettoriali euclidei della stessa dimensione possono essere identificati* (ma non in modo canonico: occorre fissare riferimenti ortonormali per ottenere un'identificazione).

Corollario 2.4.5 Siano $(V, \langle -, - \rangle)$ e $(V', \langle -, - \rangle')$ spazi vettoriali euclidei della stessa dimensione n , e siano \mathcal{R} e \mathcal{R}' riferimenti ortonormali di $(V, \langle -, - \rangle)$ e $(V', \langle -, - \rangle')$ rispettivamente. Per ogni matrice ortogonale $A \in O_n$ (di ordine n) esiste un unico isomorfismo ortogonale $f : (V, \langle -, - \rangle) \rightarrow (V', \langle -, - \rangle')$ tale che $A = M_{\mathcal{R}', \mathcal{R}}^f$. In altre parole la corrispondenza $f \mapsto M_{\mathcal{R}', \mathcal{R}}^f$ che associa ad un isomorfismo ortogonale la sua matrice rappresentativa nei riferimenti $\mathcal{R}, \mathcal{R}'$ è una biezione tra l'insieme degli isomorfismi ortogonali e l'insieme delle matrici ortogonali.

Dimostrazione. Si ricordi, innanzitutto, che, per ogni matrice $M_n(\mathbb{R})$ esiste un'unica applicazione lineare $f : V \rightarrow V'$ tale che $A = M_{\mathcal{R}', \mathcal{R}}^f$. Se $\mathcal{R}, \mathcal{R}'$ sono riferimenti ortonormali e A è una matrice ortogonale, necessariamente f è un isomorfismo ortogonale. ■

Corollario 2.4.6 Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo di dimensione n , e sia \mathcal{R} un riferimento ortonormale di $(V, \langle -, - \rangle)$. L'applicazione $O(V, \langle -, - \rangle) \rightarrow O_n, f \mapsto M_{\mathcal{R}}^f$ che associa ad un automorfismo ortogonale la sua matrice rappresentativa nel riferimento \mathcal{R} è un ben definito isomorfismo di gruppi.

Dimostrazione. Innanzitutto l'applicazione $O(V, \langle -, - \rangle) \rightarrow O_n, f \mapsto M_{\mathcal{R}}^f$ è ben definita, infatti la matrice rappresentativa di un automorfismo ortogonale in un riferimento ortonormale è dovutamente ortogonale. Inoltre è un omomorfismo di gruppi, il che segue facilmente dalle proprietà della matrice rappresentativa (riassunte poco sopra). Infine il Corollario 2.4.5 garantisce che si tratta di un'applicazione biunivoca (in caso di necessità, verificare tutti i dettagli per esercizio). ■

■ **Esempio 2.9** Nel caso in cui $(V, \langle -, - \rangle) = (\mathbb{R}^n, \langle -, - \rangle_{can})$ è lo spazio vettoriale euclideo standard, possiamo scegliere il riferimento canonico \mathcal{R}_{can} . In questo caso il Corollario 2.4.6 dice che gli automorfismi ortogonali di $(\mathbb{R}^n, \langle -, - \rangle_{can})$ sono esattamente le applicazioni lineari $L_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ associate alle matrici ortogonali $A \in O_n$. Esiste dunque un isomorfismo naturale di gruppi $O(\mathbb{R}^n, \langle -, - \rangle_{can}) \cong O_n$ il cui isomorfismo inverso è dato da $A \mapsto L_A$. ■

Esercizio 2.10 Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo di dimensione n e siano $W, W' \subseteq V$ sottospazi vettoriali di dimensione $p \leq n$. Dimostrare che esiste un automorfismo ortogonale $f :$

$V \rightarrow V$ tale che $f(W) = W'$ (Suggerimento: si considerino riferimenti ortonormali di $(W, \langle -, - \rangle_W)$ e $(W', \langle -, - \rangle_{W'})$ e poi si usino l'Esercizio 2.8 e il Corollario 2.4.4). ■

○ Siano (X, d) e (X', d') spazi metrici (cfr. Osservazione a pag. 41). Una biezione $f : X \rightarrow X'$ è un'isometria se conserva le distanze, cioè

$$d'(f(x), f(y)) = d(x, y)$$

per ogni $x, y \in X$. La Proposizione 2.4.1.(3) mostra che un isomorfismo ortogonale è, in particolare, un'isometria. Vogliamo mostrare che, in una certa misura, vale anche il viceversa. Siano dunque $(V, \langle -, - \rangle)$ e $(V', \langle -, - \rangle')$ spazi vettoriali euclidei, siano $\|-\|$ e $\|-\|'$ le rispettive norme, e siano d e d' le rispettive distanze. Vale la seguente

Proposizione 2.4.7 Sia $f : V \rightarrow V'$ una funzione (non necessariamente lineare). Allora le seguenti asserzioni sono equivalenti:

- (1) $f : (V, \langle -, - \rangle) \rightarrow (V', \langle -, - \rangle')$ è un isomorfismo ortogonale;
- (2) f è una isometria (rispetto alle distanze d, d') tale che $f(0) = 0$.

Dimostrazione. È chiaro che (1) \Rightarrow (2). Dimostriamo che (2) \Rightarrow (1). In virtù della Proposizione 2.4.3 è sufficiente dimostrare che se $f : V \rightarrow V'$ è un'isometria tale che $f(0) = 0$, allora f è un'applicazione lineare. A questo scopo cominciamo a dimostrare che f conserva i prodotti scalari. È così, infatti, per ogni $v, w \in V$, dalla (2.17) segue che

$$\begin{aligned} \langle f(v), f(w) \rangle &= \frac{\|f(v)\|^2 + \|f(w)\|^2 - \|f(v) - f(w)\|^2}{2} \\ &= \frac{\|f(v) - f(0)\|^2 + \|f(w) - f(0)\|^2 - \|f(v) - f(w)\|^2}{2} \\ &= \frac{d'(f(v), f(0))^2 + d'(f(w), f(0))^2 - d'(f(v), f(w))^2}{2} \\ &= \frac{d(v, 0)^2 + d(w, 0)^2 - d(v, w)^2}{2} \\ &= \frac{\|v\|^2 + \|w\|^2 - \|v - w\|^2}{2} \\ &= \langle v, w \rangle \end{aligned}$$

(si osservi che, in questo conto, abbiamo usato solo le ipotesi e, come dovuto, non abbiamo mai usato che f è un'applicazione lineare). Di conseguenza f trasforma ogni riferimento ortonormale $\mathcal{R} = (e_1, \dots, e_n)$ di $(V, \langle -, - \rangle)$ in un riferimento ortonormale $f(\mathcal{R}) = (f(e_1), \dots, f(e_n))$ di $(V', \langle -, - \rangle')$. Questo ci consente di scrivere un'utile formula per f : sia $\mathcal{R} = (e_1, \dots, e_n)$ un riferimento ortonormale di $(V, \langle -, - \rangle)$, sia $v \in V$ e sia $x = (x_1, \dots, x_n)^T$ il suo vettore coordinato nel riferimento \mathcal{R} . Allora

$$\begin{aligned} f(v) &= \sum_{i=1}^n \langle f(e_i), f(v) \rangle f(e_i) && (f(\mathcal{R}) \text{ è un riferimento ortonormale}) \\ &= \sum_{i=1}^n \langle e_i, v \rangle f(e_i) && (f \text{ conserva i prodotti scalari}) \\ &= \sum_{i=1}^n x_i f(e_i) && (\mathcal{R} \text{ è un riferimento ortonormale}). \end{aligned} \quad (2.18)$$

Con questa ultima formula a disposizione possiamo facilmente dimostrare che f è un'applicazione lineare e concludere la dimostrazione. Infatti siano $v, w \in V$, siano $a, b \in \mathbb{R}$ e siano $x = (x_1, \dots, x_n)^T, y = (y_1, \dots, y_n)^T$ i vettori coordinati di v, w nel riferimento \mathcal{R} . Il vettore

coordinato della combinazione lineare $av + bw$ è $ax + by$. Perciò

$$\begin{aligned} f(av + bw) &= \sum_{i=1}^n (ax_i + by_i) f(e_i) && \text{(formula (2.18))} \\ &= a \sum_{i=1}^n x_i f(e_i) + b \sum_{i=1}^n y_i f(e_i) \\ &= af(v) + bf(w) && \text{(di nuovo la formula (2.18)).} \end{aligned}$$

■

2.5 Isomorfismi Musicali

Sia V uno spazio vettoriale di dimensione n sul campo \mathbb{K} . Sappiamo che lo spazio duale V^* è isomorfo a V , ma non in modo canonico (un modo per costruire un isomorfismo è, per esempio, scegliere un riferimento in V). Ricordiamo che V^* consiste delle *forme lineari* su V , cioè le applicazioni lineari

$$\varphi : V \rightarrow \mathbb{K}.$$

Le operazioni in V^* sono definite “vettore a vettore”, cioè se $\varphi, \psi \in V^*$ sono due forme lineari e $a \in \mathbb{K}$ è uno scalare, allora la somma $\varphi + \psi$ è la forma lineare definita da

$$\varphi + \psi : V \rightarrow \mathbb{K}, \quad u \mapsto (\varphi + \psi)(u) := \varphi(u) + \psi(u),$$

e il prodotto $a\varphi$ è la forma lineare definita da

$$a\varphi : V \rightarrow \mathbb{K}, \quad u \mapsto (a\varphi)(u) := a \cdot \varphi(u).$$

Vedremo a brevissimo che se $(V, \langle -, - \rangle)$ è uno spazio vettoriale euclideo, allora il prodotto scalare $\langle -, - \rangle$ determina un isomorfismo tra V e il suo spazio duale V^* che indichiamo

$$b : V \rightarrow V^*.$$

Per ogni vettore $v \in V$, $b(v)$ è la forma lineare su V definita da

$$b(v) : V \rightarrow \mathbb{R}, \quad u \mapsto b(v)(u) := \langle v, u \rangle.$$

Poiché il prodotto scalare è lineare nel secondo argomento, $b(v)$ è effettivamente una forma lineare per ogni $v \in V$, infatti, per ogni $u, u' \in V$ e $a, a' \in \mathbb{R}$ si ha

$$b(v)(au + a'u') = \langle v, au + a'u' \rangle = a \langle v, u \rangle + a' \langle v, u' \rangle = ab(v)(u) + a' b(v)(u').$$

Inoltre, poiché il prodotto scalare è anche lineare nel primo argomento, l'applicazione $b : V \rightarrow V^*$ è un'applicazione lineare. Infatti, usando le definizioni delle operazioni in V^* ricordate poco fa, troviamo, per ogni $v, v', u \in V$ e per ogni $a, a' \in \mathbb{R}$,

$$\begin{aligned} b(av + a'v')(u) &= \langle av + a'v', u \rangle = a \langle v, u \rangle + a' \langle v', u \rangle \\ &= ab(v)(u) + a' b(v')(u) = (ab(v) + a' b(v'))(u). \end{aligned}$$

Questo mostra che le due forme lineari $b(av + a'v')$ e $ab(v) + a' b(v')$ agiscono nello stesso modo sullo stesso vettore u , per ogni $u \in V$. Dunque

$$b(av + a'v') = ab(v) + a' b(v')$$

come desiderato.

Lemma 2.5.1 Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo di dimensione n . L'applicazione lineare

$$b : V \rightarrow V^*$$

è un isomorfismo. Per ogni riferimento ortonormale $\mathcal{R} = (e_1, \dots, e_n)$ di $(V, \langle -, - \rangle)$, il sistema $b(\mathcal{R}) = (b(e_1), \dots, b(e_n))$ è esattamente il riferimento duale \mathcal{R}^* di V^* .

Dimostrazione. Per la prima parte dell'enunciato, poiché V e V^* sono spazi vettoriali della stessa dimensione, è sufficiente dimostrare che b è un'applicazione (lineare) iniettiva. A sua volta, dalla linearità, per dimostrare che b è un'applicazione iniettiva, basta verificare che il nucleo di b sia banale. Dunque, sia $v \in \ker b$. Questo vuol dire che v è un vettore in V tale che $b(v)$ è la forma lineare nulla, cioè la forma lineare tale che $b(v)(u) = 0$ per ogni $u \in V$. In particolare $0 = b(v)(v) = \langle v, v \rangle$, da cui $v = 0$, e la prima parte dell'enunciato è dimostrata.

Per la seconda parte dell'enunciato, consideriamo $\mathcal{R} = (e_1, \dots, e_n)$: un riferimento ortonormale di $(V, \langle -, - \rangle)$. Si ricordi che il riferimento duale $\mathcal{R}^* = (e_1^*, \dots, e_n^*)$ è univocamente determinato dalle condizioni

$$e_i^*(e_j) = \delta_{ij}, \quad \text{per ogni } i, j = 1, \dots, n$$

(equivalentemente e_i^* è la forma lineare su V che associa ad un vettore v la sua i -esima coordinata nel riferimento \mathcal{R}). Dunque, per dimostrare che $b(e_i) = e_i^*$ è sufficiente calcolare

$$b(e_i)(e_j) = \langle e_i, e_j \rangle = \delta_{ij},$$

in cui abbiamo usato che \mathcal{R} è un riferimento ortonormale. Questo conclude la dimostrazione. ■

Definizione 2.5.1 — Isomorfismi Musicali. L'isomorfismo $b : V \rightarrow V^*$ e il suo inverso $b^{-1} : V^* \rightarrow V$ prendono il nome di *isomorfismi musicali*. L'isomorfismo b^{-1} si indica anche $\sharp : V^* \rightarrow V$.

Teorema / Definizione 2.5.2 — Applicazione Aggiunta. Siano $(V, \langle -, - \rangle_V)$ e $(W, \langle -, - \rangle_W)$ spazi vettoriali euclidei di dimensione n ed m . Per ogni applicazione lineare $f : V \rightarrow W$, esiste un'unica applicazione lineare $f^\dagger : W \rightarrow V$, detta *applicazione aggiunta* o *coniugata*, tale che

$$\langle f^\dagger(w), v \rangle_V = \langle w, f(v) \rangle_W, \quad \text{per ogni } v \in V \text{ e } w \in W. \quad (2.19)$$

L'applicazione aggiunta f^\dagger è data da

$$f^\dagger = \sharp \circ f^* \circ b \quad (2.20)$$

in cui $f^* : W^* \rightarrow V^*$ è l'applicazione trasposta di f . In altre parole, il diagramma

$$\begin{array}{ccc} W & \xrightarrow{f^\dagger} & V \\ b \downarrow & & \uparrow \sharp \\ W^* & \xrightarrow{f^*} & V^* \end{array} \quad (2.21)$$

commuta (qui, abusando della notazione, abbiamo indicato con gli stessi simboli gli isomorfismi musicali di $(V, \langle -, - \rangle_V)$ e $(W, \langle -, - \rangle_W)$).

Se $\mathcal{R} = (e_1, \dots, e_n)$ è un riferimento ortonormale di $(V, \langle -, - \rangle_V)$, $\mathcal{S} = (u_1, \dots, u_m)$ è un riferimento ortonormale di $(W, \langle -, - \rangle_W)$ e $A = M_{\mathcal{S}, \mathcal{R}}^f$ è la matrice rappresentativa di f nei riferimenti

\mathcal{R}, \mathcal{S} , allora la matrice trasposta A^T è la matrice rappresentativa di f^\dagger nei riferimenti \mathcal{S}, \mathcal{R} :

$$A^T = M_{\mathcal{R}, \mathcal{S}}^{f^\dagger}.$$

- Prima di dimostrare il Teorema 2.5.2 è utile qualche richiamo sull'applicazione trasposta. Siano dunque V, W spazi vettoriali di dimensione n, m sul campo \mathbb{K} e siano V^*, W^* gli spazi duali. Per ogni applicazione lineare $f : V \rightarrow W$ è possibile costruire un'applicazione lineare $f^* : W^* \rightarrow V^*$, detta *applicazione trasposta*, ponendo, per ogni forma lineare $\varphi : W \rightarrow \mathbb{R}$,

$$f^*(\varphi) := \varphi \circ f : V \rightarrow \mathbb{R}.$$

Essendo composta di applicazioni lineari, $f^*(\varphi)$ è un'applicazione lineare. Inoltre è facile mostrare, utilizzando le definizioni delle operazioni nello spazio duale, che l'applicazione $f^* : W^* \rightarrow V^*$ così definita è lineare.

Se $\mathcal{R} = (v_1, \dots, v_n)$ ed $\mathcal{S} = (w_1, \dots, w_m)$ sono riferimenti di V e W rispettivamente, e $A = M_{\mathcal{S}, \mathcal{R}}^f$ è la matrice rappresentativa di f nei riferimenti \mathcal{R}, \mathcal{S} , allora la matrice trasposta A^T è la matrice rappresentativa di f^* nei riferimenti duali $\mathcal{S}^*, \mathcal{R}^*$:

$$A^T = M_{\mathcal{R}, \mathcal{S}}^{f^*}.$$

Dimostrazione del Teorema 2.5.2. Per l'esistenza, definiamo f^\dagger mediante la formula (2.20) e dimostriamo che, così definito, f^\dagger soddisfa la (2.19). Cominciamo innanzitutto ad osservare che, dalla (2.20) segue che

$$b \circ f^\dagger = b \circ \sharp \circ f^* \circ b = \text{id}_{W^*} \circ f^* \circ b = f^* \circ b.$$

Viceversa, se $\tilde{f}^\dagger : W \rightarrow V$ è un'applicazione lineare tale che $b \circ \tilde{f}^\dagger = f^* \circ b$, allora necessariamente $\tilde{f}^\dagger = \sharp \circ f^* \circ b$ (il lettore lo vede?).

Dunque, siano $v \in V$ e $w \in W$ e calcoliamo

$$\langle f^\dagger(w), v \rangle_V = b(f^\dagger(w))(v) = f^*(b(w))(v) = (b(w) \circ f)(v) = b(w)(f(v)) = \langle w, f(v) \rangle_W.$$

Questo dimostra l'esistenza. Per l'unicità, supponiamo che $\tilde{f}^\dagger : W \rightarrow V$ sia un'altra applicazione lineare tale che $\langle \tilde{f}^\dagger(w), v \rangle_V = \langle w, f(v) \rangle_W$ per ogni $v \in V$ e $w \in W$. In tal caso abbiamo

$$0 = \langle \tilde{f}^\dagger(w), v \rangle_V - \langle f^\dagger(w), v \rangle_V = \langle \tilde{f}^\dagger(w) - f^\dagger(w), v \rangle_V.$$

Dunque il vettore $\tilde{f}^\dagger(w) - f^\dagger(w)$ è ortogonale a v per ogni $v \in V$, e dunque è il vettore nullo, per ogni $w \in W$. In altre parole $\tilde{f}^\dagger(w) = f^\dagger(w)$ per ogni $w \in W$, cioè $\tilde{f}^\dagger = f^\dagger$. Questo dimostra l'unicità.

La Formula (2.20) è soddisfatta per costruzione e abbiamo anche già dimostrato che essa è equivalente alla commutatività del diagramma (2.21). Resta dunque da dimostrare solo l'ultima parte dell'enunciato. Siano dunque $\mathcal{R} = (e_1, \dots, e_n)$ ed $\mathcal{S} = (u_1, \dots, u_m)$ riferimenti come nell'enunciato e siano \mathcal{R}^* ed \mathcal{S}^* i riferimenti duali di V^* e W^* . Infine diciamo A la matrice rappresentativa di f nei riferimenti \mathcal{R}, \mathcal{S} . Allora, in virtù dell'osservazione che precede la presente dimostrazione, A^T è la matrice rappresentativa di f^* nei riferimenti $\mathcal{S}^*, \mathcal{R}^*$. Poiché $f^\dagger = \sharp \circ f^* \circ b$, la matrice rappresentativa di f^\dagger nei riferimenti \mathcal{S}, \mathcal{R} è (il prodotto delle matrici rappresentative cioè)

$$M_{\mathcal{R}, \mathcal{R}^*}^\sharp A^T M_{\mathcal{S}^*, \mathcal{S}}^b.$$

Ma l'ultima parte del Lemma 2.5.1 equivale a dire che $M_{\mathcal{R}^*, \mathcal{R}}^b = I_n$. Analogamente $M_{\mathcal{R}, \mathcal{R}^*}^\sharp = I_n$. Ne consegue che

$$M_{\mathcal{R}, \mathcal{S}}^{f^\dagger} = M_{\mathcal{R}, \mathcal{R}^*}^\sharp A^T M_{\mathcal{S}^*, \mathcal{S}}^b = A^T$$

come desiderato. ■

Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo di dimensione n e sia $f : V \rightarrow V$ un endomorfismo. In questo caso, l'applicazione aggiunta $f^\dagger : V \rightarrow V$ ha stesso dominio e codominio e ha senso confrontare f e f^\dagger .

Definizione 2.5.2 — Endomorfismo Simmetrico. Un *endomorfismo simmetrico* di $(V, \langle -, - \rangle)$ è un endomorfismo $f : V \rightarrow V$ tale che $f = f^\dagger$. In altre parole, f soddisfa la seguente identità: per ogni $v, w \in V$

$$\langle f(v), w \rangle = \langle v, f(w) \rangle.$$

La caratterizzazione nella prossima proposizione giustifica la terminologia.

Proposizione 2.5.3 Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo di dimensione n , e sia $f : V \rightarrow V$ un endomorfismo. Le seguenti condizioni sono equivalenti:

- (1) f è un endomorfismo simmetrico;
- (2) per ogni riferimento ortonormale $\mathcal{R} = (e_1, \dots, e_n)$ di $(V, \langle -, - \rangle)$, la matrice rappresentativa $M_{\mathcal{R}}^f$ è una matrice simmetrica;
- (3) esiste un riferimento ortonormale $\mathcal{R} = (e_1, \dots, e_n)$ di $(V, \langle -, - \rangle)$ tale che la matrice rappresentativa $M_{\mathcal{R}}^f$ è una matrice simmetrica.

Dimostrazione. Dimostriamo che $(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (1)$. Che $(1) \Rightarrow (2)$ è una conseguenza immediata dell'ultima parte del Teorema 2.5.2. Che $(2) \Rightarrow (3)$ è ovvio e resta da dimostrare che $(3) \Rightarrow (1)$. Sia dunque $\mathcal{R} = (e_1, \dots, e_n)$ un riferimento ortonormale nel quale la matrice rappresentativa $A = M_{\mathcal{R}}^f$ di f è simmetrica. Inoltre, siano $v, w \in V$ e siano $x, y \in \mathbb{R}^n$ i loro vettori coordinati nel riferimento \mathcal{R} . I vettori coordinati di $f(v), f(w)$ nel riferimento \mathcal{R} sono Ax, Ay rispettivamente. Perciò:

$$\langle f(v), w \rangle = \langle Ax, y \rangle_{can} = (Ax)^T y = x^T A^T y = x^T A y = \langle x, Ay \rangle_{can} = \langle v, f(w) \rangle,$$

come desiderato. Questo conclude la dimostrazione. ■

Corollario 2.5.4 Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo di dimensione n , e sia \mathcal{R} un riferimento ortonormale di $(V, \langle -, - \rangle)$. Per ogni matrice quadrata simmetrica $A \in M_n(\mathbb{R})$ (di ordine n) esiste un unico endomorfismo simmetrico $f : (V, \langle -, - \rangle) \rightarrow (V, \langle -, - \rangle)$ tale che $A = M_{\mathcal{R}}^f$. In altre parole la corrispondenza $f \mapsto M_{\mathcal{R}}^f$ che associa ad un endomorfismo simmetrico la sua matrice rappresentativa nel riferimento \mathcal{R} è una biezione tra l'insieme degli endomorfismi simmetrici e l'insieme delle matrici simmetriche.

■ **Esempio 2.10** Nel caso in cui $(V, \langle -, - \rangle) = (\mathbb{R}^n, \langle -, - \rangle_{can})$ è lo spazio vettoriale euclideo standard, possiamo scegliere il riferimento canonico \mathcal{R}_{can} . In questo caso il Corollario 2.5.4 dice che gli endomorfismi simmetrici di $(\mathbb{R}^n, \langle -, - \rangle_{can})$ sono esattamente le applicazioni lineari $L_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ associate alle matrici simmetriche A . ■

Esercizio 2.11 Dimostrare il Corollario 2.5.4. ■

3. Diagonalizzazione di Endomorfismi

In questo capitolo ci interesseremo alla struttura degli endomorfismi di uno spazio vettoriale V . Lavoreremo di nuovo su un campo arbitrario \mathbb{K} e affronteremo il problema di capire sotto quali condizioni la matrice rappresentativa di un endomorfismo (e quindi l'endomorfismo stesso) risulta particolarmente semplice, nella fattispecie una matrice diagonale.

3.1 Autovalori e Autovettori

Sia \mathbb{K} un campo e V uno spazio vettoriale di dimensione n su \mathbb{K} . Consideriamo un endomorfismo di V , cioè un'applicazione lineare $f : V \rightarrow V$. In questo contesto è naturale rappresentare f scegliendo nel dominio e nel codominio lo stesso riferimento $\mathcal{R} = (v_1, \dots, v_n)$. Ci chiediamo: è possibile scegliere \mathcal{R} in modo che la matrice rappresentativa $M_{\mathcal{R}}^f$ sia "particolarmente semplice"? Si noti che $M_{\mathcal{R}}^f$ è, in ogni caso, una matrice quadrata di ordine n : $M_{\mathcal{R}}^f \in M_n(\mathbb{K})$. Una classe di matrici sufficientemente semplici per i nostri scopi è la classe delle *matrici diagonali*, cioè le matrici quadrate del tipo:

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \quad (3.1)$$

con $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{K}$. In questo capitolo ci poniamo la seguente domanda: *dato un endomorfismo $f : V \rightarrow V$, sotto quali condizioni è possibile trovare un riferimento \mathcal{R} tale che la matrice rappresentativa $M_{\mathcal{R}}^f$ sia una matrice diagonale?*

Definizione 3.1.1 — Endomorfismo Diagonalizzabile. Un endomorfismo $f : V \rightarrow V$ si dice *diagonalizzabile* se esiste un riferimento $\mathcal{R} = (v_1, \dots, v_n)$ di V tale che la matrice rappresentativa $M_{\mathcal{R}}^f$ di f nel riferimento \mathcal{R} è una matrice diagonale. In tal caso si dice anche che il riferimento \mathcal{R} *diagonalizza* l'endomorfismo f .

■ **Esempio 3.1** L'endomorfismo nullo $0 : V \rightarrow V$, che associa ad ogni vettore il vettore nullo, è diagonalizzabile e ogni riferimento \mathcal{R} di V lo diagonalizza, infatti, per ogni riferimento \mathcal{R} , la matrice rappresentativa $M_{\mathcal{R}}^f$ è la matrice nulla. Anche l'endomorfismo identico $\text{id} : V \rightarrow V$ è diagonalizzabile e ogni riferimento lo diagonalizza, infatti, per ogni riferimento \mathcal{R} , la matrice rappresentativa $M_{\mathcal{R}}^{\text{id}}$ è la matrice identica I_n .

Più in generale, sia $\lambda \in \mathbb{K}$ uno scalare. Consideriamo l'applicazione $m_\lambda : V \rightarrow V$, $v \mapsto m_\lambda(v) := \lambda v$. Chiaramente m_λ è un'applicazione lineare e, perciò, un endomorfismo. Inoltre m_λ è un endomorfismo diagonalizzabile e ogni riferimento lo diagonalizza, infatti, per ogni riferimento \mathcal{R} si ha

$$M_{\mathcal{R}}^{m_\lambda} = \lambda I_n = \begin{pmatrix} \lambda & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda \end{pmatrix}$$

che è una matrice diagonale. Si osservi che m_0 e m_1 sono rispettivamente l'endomorfismo nullo e l'endomorfismo identico. Saremo in grado di fornire più esempi (e più controesempi) più avanti in questo capitolo. ■

Sia $f : V \rightarrow V$ un endomorfismo diagonalizzabile e sia $\mathcal{R} = (v_1, \dots, v_n)$ un riferimento che diagonalizza f . Allora la matrice rappresentativa $M_{\mathcal{R}}^f$ è una matrice diagonale. Questo vuol dire che, per ogni $i = 1, \dots, n$, le coordinate di $f(v_i)$ nel riferimento \mathcal{R} sono tutte nulle eccetto, al più, proprio la i -esima. In altre parole, per ogni $i = 1, \dots, n$, esiste $\lambda \in \mathbb{K}$ tale che

$$f(v_i) = \lambda v_i.$$

Questa considerazione suggerisce la seguente

Definizione 3.1.2 — Autovalori e Autovettori. Un *autovalore* dell'endomorfismo $f : V \rightarrow V$ è uno scalare $\lambda \in \mathbb{K}$ tale che esiste un vettore non nullo $v \in V$ che soddisfa:

$$f(v) = \lambda v. \tag{3.2}$$

In tal caso il vettore $v \in V \setminus \{0\}$ viene detto *autovettore* relativo all'autovalore λ e, viceversa, λ viene detto *autovalore relativo all'autovettore* v . Dunque un autovettore dell'endomorfismo $f : V \rightarrow V$ è un vettore non nullo $v \in V$ per cui esiste uno scalare λ (il relativo autovalore) tale che valga la (3.2).

■ **Esempio 3.2** Sia $\lambda \in \mathbb{K}$ e sia $m_\lambda : V \rightarrow V$ l'endomorfismo dell'esempio 3.1. Allora λ è un autovalore di m_λ e ogni vettore non nullo $v \in V$ è un autovettore (relativo all'autovalore λ). ■

■ **Esempio 3.3** Sia $f : V \rightarrow V$ un endomorfismo non iniettivo. Dunque ci sono vettori non nulli nel nucleo di f : $\ker f \neq 0$. In questo caso, 0 è un autovalore di f , e ogni vettore non nullo $v \in \ker f$ è un autovettore relativo all'autovalore 0 . ■

Lemma 3.1.1 Sia $f : V \rightarrow V$ un endomorfismo. Allora ogni autovettore di f è relativo ad un unico autovalore, cioè se $v \in V$ è un autovettore, allora esiste un unico scalare λ tale che $f(v) = \lambda v$.

Dimostrazione. Siano $\lambda, \lambda' \in \mathbb{K}$ entrambi autovalori relativi all'autovettore v . Allora si ha

$$f(v) = \lambda v = \lambda' v \quad \Rightarrow \quad 0 = \lambda v - \lambda' v = (\lambda - \lambda')v.$$

Poiché $v \neq 0$ per definizione di autovettore, necessariamente $\lambda - \lambda' = 0$, cioè $\lambda = \lambda'$. ■

La definizione di autovettore (e autovalore) ci consente di riformulare la Definizione 3.1.2. Infatti è chiaro che un endomorfismo $f : V \rightarrow V$ è diagonalizzabile sse esiste un riferimento $\mathcal{R} = (v_1, \dots, v_n)$ formato da autovettori (cioè v_i è un autovettore di f per ogni $i = 1, \dots, n$). In tal caso la matrice rappresentativa di f nel riferimento \mathcal{R} è la matrice diagonale (3.1) dove gli elementi di diagonale $\lambda_1, \dots, \lambda_n$ sono esattamente gli autovalori relativi agli autovettori v_1, \dots, v_n . Sarà bene, nel prosieguo, tenere a mente questa riformulazione.

In questo capitolo affrontiamo il *problema della diagonalizzazione* di un endomorfismo, cioè il problema di capire se un dato endomorfismo è diagonalizzabile oppure no ed eventualmente di trovare un riferimento che lo diagonalizza. Per risolvere questo problema è opportuno, innanzitutto, imparare ad individuare gli autovalori di un endomorfismo. A questo scopo, consideriamo un endomorfismo $f : V \rightarrow V$, fissiamo un riferimento $\mathcal{R} = (v_1, \dots, v_n)$ di V e osserviamo che uno scalare $\lambda \in \mathbb{K}$ è un autovalore di f sse l'equazione

$$f(v) = \lambda v \quad (3.3)$$

nell'incognita v ammette soluzioni non banali, cioè ammette altre soluzioni oltre alla ovvia soluzione $v = 0$. Detta $A = M_{\mathcal{R}}^f$ la matrice rappresentativa di f nel riferimento \mathcal{R} , è chiaro che l'equazione (3.3) è rappresentata, in coordinate, dall'equazione

$$Ax = \lambda x \quad (3.4)$$

nell'incognita $x = (x_1, \dots, x_n)^T \in \mathbb{K}^n$. Perciò l'equazione (3.3) ammette soluzioni non banali sse l'equazione (3.4) ammette soluzioni non banali. A sua volta l'equazione (3.4) si può riscrivere come un sistema lineare nelle incognite x_1, \dots, x_n . Infatti, per ogni $x \in \mathbb{K}^n$, si ha $\lambda x = \lambda I_n x$, da cui

$$Ax = \lambda x \Leftrightarrow Ax - \lambda x = 0 \Leftrightarrow Ax - \lambda I_n x = 0 \Leftrightarrow (A - \lambda I_n)x = 0.$$

Dunque l'equazione (3.4) è equivalente a

$$(A - \lambda I_n)x = 0, \quad (3.5)$$

che è un sistema lineare omogeneo di n equazioni nelle n incognite x_1, \dots, x_n sul campo \mathbb{K} . Riassumendo, lo scalare $\lambda \in \mathbb{K}$ è un autovalore dell'endomorfismo f sse il sistema lineare omogeneo (3.5) ammette soluzioni non banali. Per il Teorema di Rouché-Capelli questo avviene sse il determinante della matrice dei coefficienti del sistema (3.5) è nullo, cioè

$$\det(A - \lambda I_n) = 0.$$

In altre parole gli autovalori dell'endomorfismo f sono le soluzioni dell'equazione

$$\det(A - tI_n) = 0 \quad (3.6)$$

nell'incognita $t \in \mathbb{K}$. Vogliamo ora osservare da vicino l'Equazione (3.6). Il membro di sinistra dell'equazione è un *polinomio nell'indeterminata t* . La prossima osservazione contiene alcuni richiami sui polinomi che saranno utili da qui in avanti. Questo materiale è presentato senza dimostrazione e i dettagli saranno oggetto del corso di Algebra del secondo anno.

 **Polinomi su un Campo.** Sia \mathbb{K} un campo e sia n un intero non negativo. Un *polinomio di grado n sul campo \mathbb{K} nell'indeterminata t* è un'espressione formale $P(t)$ del tipo:

$$P(t) = a_0 + a_1 t + \dots + a_n t^n = \sum_{i=0}^n a_i t^i, \quad a_0, \dots, a_n \in \mathbb{K}, \quad a_n \neq 0. \quad (3.7)$$

L'indeterminata t si deve riguardare come una *variabile* che all'occorrenza può assumere specifici valori nel campo \mathbb{K} ma, a priori, non ha alcun valore predefinito. L'insieme dei polinomi nell'indeterminata t si indica con il simbolo $\mathbb{K}[t]$ ed è munito di due operazioni naturali: una *somma* $+$ e un *prodotto* \cdot definiti come segue. Se

$$P(t) = a_0 + a_1t + \cdots + a_nt^n = \sum_{i=1}^n a_i t^i, \quad e \quad Q(t) = b_0 + b_1t + \cdots + b_mt^m = \sum_{j=1}^m b_j t^j$$

sono polinomi di grado n, m rispettivamente, allora

$$P(t) + Q(t) := (a_0 + b_0) + (a_1 + b_1)t + \cdots + (a_m + b_m)t^m + a_{m+1}t^{m+1} + \cdots + a_nt^n, \quad \text{se } m \leq n$$

(analogamente se $m > n$) e

$$P(t)Q(t) = \sum_{j=1}^{m+n} \left(\sum_{i+j=k} a_i b_j \right) t^k.$$

Con queste due operazioni i polinomi formano un *anello commutativo* (perciò valgono tutte le usuali regole di calcolo: proprietà associativa di somma e prodotto, proprietà distributiva, etc.).

Ogni polinomio

$$P(t) = \sum_{i=0}^n a_i t^i \in \mathbb{K}[t]$$

determina una funzione:

$$P: \mathbb{K} \rightarrow \mathbb{K}, \quad \lambda \mapsto P(\lambda) := \sum_{i=0}^n a_i \lambda^i,$$

detta talvolta *funzione polinomiale* (si noti che esistono campi per i quali un polinomio non nullo può dar luogo ad una funzione polinomiale nulla, quindi un polinomio *non* va identificato con la corrispondente funzione polinomiale; se $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$, il campo dei numeri razionali, reali, complessi, allora è possibile identificare un polinomio con la corrispondente funzione polinomiale). Dato un polinomio $P(t) \in \mathbb{K}[t]$, uno scalare λ è *radice* di $P(t)$ se, per definizione, $P(\lambda) = 0$. Vale il seguente

Teorema 3.1.2 Sia $P(t)$ un polinomio. Uno scalare λ è radice di $P(t)$ sse il polinomio $t - \lambda$ divide $P(t)$. Cioè esiste un polinomio $Q(t)$ tale che

$$P(t) = (t - \lambda) Q(t)$$

Corollario 3.1.3 Ogni polinomio di grado n ha al più n radici distinte.

Corollario 3.1.4 Sia $P(t)$ un polinomio e siano $\lambda_1, \dots, \lambda_k$ tutte le sue radici distinte. Allora esistono interi positivi m_1, \dots, m_k tali che

$$P(t) = (t - \lambda_1)^{m_1} \cdots (t - \lambda_k)^{m_k} R(t) \tag{3.8}$$

dove $R(t)$ è un polinomio privo di radici. Viceversa, se $P(t)$ è della forma (3.8) con $\lambda_1, \dots, \lambda_k$ scalari distinti ed $R(t)$ polinomio privo di radici, allora $\lambda_1, \dots, \lambda_k$ sono tutte le radici distinte di $P(t)$.

I numeri naturali m_1, \dots, m_k nella (3.8) prendono il nome di *molteplicità* delle radici $\lambda_1, \dots, \lambda_k$. Ovviamente la somma $m_1 + \cdots + m_k$ delle molteplicità delle radici è minore o uguale del grado del polinomio. Abusando un po' la terminologia corrente (che vorrebbe anche $Q(t)$ fattorizzato nelle sue componenti irriducibili), chiameremo la presentazione del polinomio $P(t)$ nella forma data dalla (3.8) una *fattorizzazione completa* di $P(t)$.

Torniamo ora all'Equazione (3.6). Se interpretiamo t come un'indeterminata, allora il membro di destra $\det(A - tI_n)$ è un polinomio di grado n . Questo segue dal fatto che il determinante di una matrice è una somma di prodotti di elementi di matrice, ed ogni addendo in questa somma contiene esattamente n fattori. Si osservi che il termine nel quale l'indeterminata t è elevata al massimo esponente n , proviene dal prodotto degli elementi sulla diagonale principale: se $A = (a_{ij})_{i=1, \dots, n}^{j=1, \dots, n}$, allora

$$\begin{aligned} \det(A - tI_n) &= \det \begin{pmatrix} a_{11} - t & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - t & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - t \end{pmatrix} \\ &= (a_{11} - t)(a_{22} - t) \cdots (a_{nn} - t) + \cdots = (-1)^n t^n + \cdots \end{aligned}$$

Definizione 3.1.3 — Polinomio Caratteristico di un Endomorfismo. Il polinomio $P_f(t) := \det(A - tI_n)$ è detto il *polinomio caratteristico* dell'endomorfismo f . L'equazione $P_f(t) = 0$ è detta *equazione caratteristica*.

La lunga discussione fatta fin qui si può riassumere nel seguente

Teorema 3.1.5 Sia $f : V \rightarrow V$ un endomorfismo. Gli autovalori di f sono esattamente le radici del polinomio caratteristico $P_f(t)$ o, equivalentemente, le soluzioni dell'equazione caratteristica $P_f(t) = 0$.

Quindi, per trovare gli autovalori di un endomorfismo, è necessario risolvere l'equazione caratteristica associata.

Definizione 3.1.4 — Molteplicità Algebrica di un Autovalore. La molteplicità di un autovalore λ in quanto radice del polinomio caratteristico è detta *molteplicità algebrica* dell'autovalore λ e si indica con $m_a(\lambda)$.

Osserviamo, per utilizzo futuro, che se $\lambda_1, \dots, \lambda_k$ sono tutti gli autovalori distinti di un endomorfismo $f : V \rightarrow V$ allora la somma delle loro molteplicità algebriche non supera $n = \dim V$: $m_a(\lambda_1) + \cdots + m_a(\lambda_k) \leq n$.

Nella definizione di polinomio caratteristico interviene la matrice rappresentativa dell'endomorfismo f in un fissato riferimento \mathcal{R} . Sarebbe dunque che la definizione dipenda dalla scelta del riferimento. In realtà vale la seguente

Proposizione 3.1.6 Sia $f : V \rightarrow V$ un endomorfismo. Il polinomio caratteristico $P_f(t)$ non dipende dalla scelta del riferimento.

Dimostrazione. Siano $\mathcal{R}, \mathcal{R}'$ due riferimenti di V e siano $A = M_{\mathcal{R}}^f, A' = M_{\mathcal{R}'}^f$ le matrici rappresentative di f nei riferimenti $\mathcal{R}, \mathcal{R}'$ rispettivamente. Se chiamiamo $M = M_{\mathcal{R}', \mathcal{R}}$ la matrice del cambio di coordinate, allora dalle proprietà della matrice rappresentativa (e delle matrici del cambio di coordinate) segue che

$$A' = MAM^{-1}.$$

Perciò

$$\begin{aligned} \det(A' - tI_n) &= \det(MAM^{-1} - tMI_nM^{-1}) = \det(M(A - tI_n)M^{-1}) \\ &= \det M \det(A - tI_n) \det(M^{-1}) = \det M \det(A - tI_n) (\det M)^{-1} \\ &= \det(A - tI_n), \end{aligned}$$

dove abbiamo usato varie proprietà del prodotto di matrici (il lettore è in grado di individuarle tutte?) oltre che il Teorema di Binet. Questo mostra che i polinomi caratteristici calcolati a partire dai riferimenti $\mathcal{R}, \mathcal{R}'$ in effetti coincidono. ■

Sia $f : V \rightarrow V$ un endomorfismo, sia $\mathcal{R} = (v_1, \dots, v_n)$ un riferimento di V e sia A la matrice rappresentativa di f nel riferimento \mathcal{R} . Supponiamo di aver risolto l'equazione caratteristica $P_f(t) = 0$ e di aver così trovato gli autovalori di f . Sia $\lambda \in \mathbb{K}$ un tale autovalore. Per trovare gli autovettori relativi all'autovalore λ dobbiamo risolvere l'equazione (3.3) (e selezionare le soluzioni non nulle). Equivalentemente possiamo risolvere il sistema (3.5) e poi sostituire le soluzioni non nulle $(x_1, \dots, x_n)^T$ in $v = x_1 v_1 + \dots + x_n v_n$. Quindi, se siamo in grado di trovare tutti gli autovalori di f , possiamo anche trovare tutti gli autovettori.

■ **Esempio 3.4** Consideriamo l'endomorfismo

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x & y & -z \\ -x & +2y & -z \\ & & -2z \end{pmatrix}.$$

Vogliamo trovare autovalori e autovettori di f . Cominciamo dagli autovalori. Dobbiamo calcolare il polinomio caratteristico $P_f(t)$. La matrice rappresentativa di f nel riferimento canonico è

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

Il polinomio caratteristico è

$$P_f(t) = \det(A - tI_3) = \det \begin{pmatrix} -t & 1 & -1 \\ -1 & 2-t & -1 \\ 0 & 0 & -2-t \end{pmatrix}.$$

Questo determinante si può calcolare, per esempio, sviluppando mediante la regola di Laplace rispetto alla terza riga. Troviamo

$$P_f(t) = (-2-t)(-t(2-t) + 1) = (-2-t)(t^2 - 2t + 1) = -(t+2)(t-1)^2. \quad (3.9)$$

La (3.9) già fornisce una fattorizzazione completa di $P_f(t)$ e da qui leggiamo che $P_f(t)$ ha due radici, -2 e 1 , con molteplicità rispettivamente 1 e 2 . Dunque l'endomorfismo f ha due autovalori -2 e 1 con molteplicità algebriche 1 e 2 .

Determiniamo ora gli autovettori. Gli autovettori relativi all'autovalore $\lambda = -2$ sono le soluzioni non nulle del sistema (3.5). In questo caso, la matrice $A - \lambda I_n$ è

$$A + 2I_3 = \begin{pmatrix} 2 & 1 & -1 \\ -1 & 4 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

e il sistema (3.5) diventa

$$S : \begin{cases} 2x + y - z = 0 \\ -x + 4y - z = 0 \\ 0 = 0 \end{cases}.$$

Le prime due equazioni del sistema sono indipendenti, perciò lo spazio $\text{Sol}(S)$ delle soluzioni è un sottospazio di dimensione 1 . Un semplice conto rivela che

$$\text{Sol}(S) = \left\{ \begin{pmatrix} s/3 \\ s/3 \\ s \end{pmatrix} : s \in \mathbb{R} \right\}.$$

Dunque gli autovettori relativi all'autovalore -2 sono i vettori del tipo

$$\begin{pmatrix} s/3 \\ s/3 \\ s \end{pmatrix}, \quad s \neq 0.$$

Resta da calcolare gli autovettori relativi all'autovalore $\lambda = 1$. Procediamo come prima. In questo caso, la matrice $A - \lambda I_n$ è

$$A - I_3 = \begin{pmatrix} -1 & 1 & -1 \\ -1 & 1 & -1 \\ 0 & 0 & -3 \end{pmatrix}$$

e il sistema (3.5) diventa

$$S' : \begin{cases} -x + y - z = 0 \\ -x + y - z = 0 \\ -3z = 0 \end{cases}.$$

Solo le ultime due equazioni sono indipendenti e lo spazio delle soluzioni è

$$\text{Sol}(S') = \left\{ \begin{pmatrix} s' \\ s' \\ 0 \end{pmatrix} : s' \in \mathbb{R} \right\}.$$

Perciò gli autovettori relativi all'autovalore 1 sono i vettori del tipo

$$\begin{pmatrix} s' \\ s' \\ 0 \end{pmatrix}, \quad s' \neq 0.$$

Esercizio 3.1 Calcolare autovalori e autovettori dell'endomorfismo

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x + y \\ x + y \\ x - y + z \end{pmatrix}.$$

3.2 Autospazi

Sia V uno spazio vettoriale di dimensione n sul campo \mathbb{K} , e sia $f : V \rightarrow V$ un endomorfismo di V . Gli autovettori di f si organizzano in sottospazi vettoriali che studiamo in dettaglio in questa sezione. Sia dunque λ un autovalore di f .

Definizione 3.2.1 — Autospazio Relativo ad un Autovalore. L'autospazio relativo all'autovalore λ è l'insieme V_λ dei vettori $v \in V$ che soddisfano l'equazione (3.2):

$$V_\lambda := \{v \in V : f(v) = \lambda v\} \subseteq V.$$

In altre parole

$$V_\lambda = \{\text{autovettori relativi all'autovalore } \lambda\} \cup \{0\}.$$

Chiaramente, gli autospazi sono sottospazi vettoriali. Infatti $0 \in V_\lambda$, inoltre se $v, w \in V_\lambda$ e $a, b \in \mathbb{K}$, allora

$$f(av + bw) = af(v) + bf(w) = a\lambda v + b\lambda w = \lambda(av + bw),$$

il che mostra che anche $av + bw \in V_\lambda$. In particolare, ha senso parlare di *dimensione dell'autospazio* V_λ .

Definizione 3.2.2 — Molteplicità Geometrica di un Autovalore. La dimensione dell'autospazio V_λ è detta *molteplicità geometrica* dell'autovalore λ e si indica con $m_g(\lambda)$:

$$m_g(\lambda) := \dim V_\lambda.$$

Esercizio 3.2 Sia $f : V \rightarrow V$ un endomorfismo, sia $\lambda \in \mathbb{K}$ un autovalore di f e sia V_λ il relativo autospazio. Dimostrare che $V_\lambda \subseteq V$ è un sottospazio f -invariante, cioè $f(V_\lambda) \subseteq V_\lambda$. ■

■ **Esempio 3.5** Sia f l'endomorfismo dell'Esempio 3.4. Le molteplicità geometriche dei due autovalori -2 e 1 sono entrambe 1. Infatti i rispettivi autospazi sono $V_{-2} = \text{Sol}(S)$ e $V_1 = \text{Sol}(S')$ che hanno entrambi dimensione 1. ■

L'Esempio 3.5 suggerisce il modo di calcolare la molteplicità geometrica. Sia $f : V \rightarrow V$ un endomorfismo e sia $\lambda \in \mathbb{K}$ un autovalore di f . Fissiamo un riferimento \mathcal{R} di V e indichiamo con A la matrice rappresentativa di f nel riferimento \mathcal{R} . Da ora in avanti, in questa situazione, indicheremo sempre

$$A_\lambda := A - \lambda I_n,$$

$n = \dim V$ (anche se t è un'indeterminata, indicheremo $A_t := A - tI_n$).

Proposizione 3.2.1 La molteplicità geometrica dell'autovalore λ è data da

$$m_g(\lambda) = n - \text{rk} A_\lambda.$$

Dimostrazione. I vettori coordinati x degli elementi dell'autospazio V_λ , sono le soluzioni del sistema lineare omogeneo (3.5) che indichiamo con S e, nella notazione appena introdotta, si scrive

$$S : \{A_\lambda x = 0. \tag{3.10}$$

In altre parole, detto $c_{\mathcal{R}} : V \rightarrow \mathbb{K}^n$ l'isomorfismo coordinato (che associa ad un vettore di V il suo vettore coordinato nel riferimento \mathcal{R}) si ha $c_{\mathcal{R}}(V_\lambda) = \text{Sol}(S)$ (lo spazio delle soluzioni di S). Siccome $c_{\mathcal{R}}$ è appunto un isomorfismo, esso conserva la dimensione dei sottospazi e si ha:

$$m_g(\lambda) = \dim V_\lambda = \dim \text{Sol}(S) = n - \text{rk} A_\lambda.$$

■

Se λ è un autovalore di f , allora esiste almeno un autovettore, cioè una soluzione non nulla dell'equazione $f(v) = \lambda v$. In altre parole l'autospazio V_λ contiene almeno un vettore non nullo e quindi

$$m_g(\lambda) = \dim V_\lambda > 0.$$

In effetti, si può dire di più sulla molteplicità geometrica di un autovalore, in base alla seguente

Proposizione 3.2.2 Sia λ un autovalore dell'endomorfismo $f : V \rightarrow V$. Allora la molteplicità geometrica di λ è minore o uguale alla molteplicità algebrica:

$$m_g(\lambda) \leq m_a(\lambda).$$

Dimostrazione. Poniamo $m = m_g(\lambda)$ e consideriamo l'autospazio V_λ . Scegliamo un riferimento $\mathcal{R}_\lambda = (v_1, \dots, v_m)$ di V_λ . I vettori v_1, \dots, v_m sono linearmente indipendenti, perciò il sistema \mathcal{R}_λ si può completare ad un riferimento $\mathcal{R} = (v_1, \dots, v_m, u_{m+1}, \dots, u_n)$ di V ($n = \dim V$). La matrice rappresentativa A dell'endomorfismo f nel riferimento \mathcal{R} assume la seguente forma

$$A = \begin{pmatrix} \lambda I_m & B \\ 0 & C \end{pmatrix},$$

dove $0 \in M_{n-m, m}(\mathbb{K})$ è la matrice nulla, $B \in M_{m, n-m}(\mathbb{K})$ e $C \in M_{n-m, n-m}(\mathbb{K})$. Infatti sia $i = 1, \dots, m$. La i -esima colonna $A^{(i)}$ di A è il vettore coordinato di $f(v_i)$ nel riferimento \mathcal{R} . Ma $f(v_i) = \lambda v_i$ perché $v_i \in V_\lambda$, perciò il suo vettore coordinato è

$$(0, \dots, 0, \underbrace{\lambda}_{\text{posto } i}, 0, \dots, 0)^T.$$

Il polinomio caratteristico è

$$P_f(t) = \det(A - tI_n) = \det \begin{pmatrix} (\lambda - t)I_m & B \\ 0 & C - tI_{n-m} \end{pmatrix}.$$

Sviluppando questo determinante con la regola di Laplace rispetta alla prima colonna, si può dimostrare per induzione su m che

$$P_f(t) = (\lambda - t)^m Q(t), \tag{3.11}$$

in cui $Q(t)$ è il polinomio dato da

$$Q(t) = \det(C - tI_{n-m})$$

(il lettore tenti per esercizio una dimostrazione dettagliata). La (3.11) già mostra che $m = m_g(\lambda)$ non può superare la molteplicità algebrica di λ (infatti $(t - \lambda)$ divide $P_f(t)$ almeno m volte). ■

Dalla Proposizione 3.2.2 (e dal commento subito dopo la Definizione 3.1.4) segue anche che se $\lambda_1, \dots, \lambda_k$ sono tutti gli autovalori distinti di un endomorfismo $f : V \rightarrow V$ allora la somma delle loro molteplicità geometriche non supera $n = \dim V$: $m_g(\lambda_1) + \dots + m_g(\lambda_k) \leq m_a(\lambda_1) + \dots + m_a(\lambda_k) \leq n$.

Vogliamo ora determinare la posizione reciproca degli autospazi relativi ad autovalori distinti.

Proposizione 3.2.3 Siano $\lambda_1, \dots, \lambda_p$ autovalori distinti di un endomorfismo $f : V \rightarrow V$. Allora gli autospazi $V_{\lambda_1}, \dots, V_{\lambda_p}$ sono in somma diretta.

🔴 L'enunciato della Proposizione 3.2.3 richiede forse qualche spiegazione. Sia V uno spazio vettoriale e siano $V_1, \dots, V_p \subseteq V$ sottospazi vettoriali. Il *sottospazio somma*

$$V_1 + \dots + V_p = \sum_{i=1}^p V_i \subseteq V$$

è, per definizione, il sottospazio definito da:

$$V_1 + \dots + V_p = \{v_1 + \dots + v_p : v_i \in V_i \text{ per ogni } i = 1, \dots, p\}.$$

Si dice che V_1, \dots, V_p sono in *somma diretta* se, per ogni $v \in V_1 + \dots + V_p$ esiste un'unica p -upla di vettori (v_1, \dots, v_p) con $v_i \in V_i$ per ogni $i = 1, \dots, p$ tale che

$$v = v_1 + \dots + v_p,$$

e, in questo caso, si scrive $V_1 + \dots + V_p = V_1 \oplus \dots \oplus V_p$. Similmente al caso $p = 2$ le somme dirette si caratterizzano facilmente. Per la precisione, i sottospazi V_1, \dots, V_p sono in somma diretta sse per ogni $j = 1, \dots, p$ si ha che

$$V_j \cap \sum_{i \neq j} V_i = 0, \quad (3.12)$$

in altre parole sse, per ogni j , V_j interseca la somma dei rimanenti sottospazi solo nel vettore nullo (in altre parole ancora sse per ogni j , il sottospazio V_j è in somma diretta con la somma dei rimanenti sottospazi). Per completezza, dimostriamo questa equivalenza. Supponiamo prima che V_1, \dots, V_p siano in somma diretta, fissiamo $j = 1, \dots, p$ e consideriamo un vettore

$$v \in V_j \cap \sum_{i \neq j} V_i.$$

In particolare $v \in \sum_{i \neq j} V_i$, quindi esistono vettori $v_i \in V_i$ con $i \neq j$ tali che $v = \sum_{i \neq j} v_i$. Da cui

$$v_1 + \dots + v_{j-1} - v + v_{j+1} + \dots + v_p = 0.$$

In altre parole, posto $v_j = -v$, (v_1, \dots, v_p) è una p -upla di vettori con $v_i \in V_i$ per ogni i e tale che $\sum_{i=1}^p v_i = 0$. Ma $(0, \dots, 0)$ è un'altra tale p -upla. Per l'unicità, deve essere, in particolare, $v_j = 0$ e quindi $v = 0$. Viceversa supponiamo che, per ogni $j = 1, \dots, p$, valga la (3.12). Consideriamo $v \in \sum_{i=1}^p V_i$ e supponiamo che

$$v = v_1 + \dots + v_p = v'_1 + \dots + v'_p \quad (3.13)$$

con $v_i, v'_i \in V_i$ per ogni i . Vogliamo dimostrare che $v_j = v'_j$ per ogni j . A questo scopo osserviamo che, dalla (3.13), segue che

$$v_j - v'_j = \sum_{i \neq j} (v'_i - v_i)$$

per ogni j . Ma il primo membro appartiene a V_j mentre il secondo appartiene a $\sum_{i \neq j} V_i$, perciò devono essere entrambi nulli, in particolare $v_j = v'_j$.

Prima di dimostrare la Proposizione 3.2.3 è necessario il seguente

Lemma 3.2.4 Siano v_1, \dots, v_p autovettori dell'endomorfismo $f : V \rightarrow V$ relativi ad autovalori distinti $\lambda_1, \dots, \lambda_p$. Allora v_1, \dots, v_p sono linearmente indipendenti.

Dimostrazione. La dimostrazione è per induzione su p . Se $p = 1$ abbiamo un unico autovettore v_1 . Per definizione di autovettore, $v_1 \neq 0$ perciò v_1 è anche linearmente indipendente. Supponiamo ora che l'asserto sia vero per $p = k$ e dimostriamolo per $p = k + 1$. Consideriamo dunque $k + 1$ autovettori v_1, \dots, v_{k+1} relativi ad autovalori distinti $\lambda_1, \dots, \lambda_{k+1}$ e una loro combinazione lineare nulla:

$$a_1 v_1 + \dots + a_k v_k + a_{k+1} v_{k+1} = 0, \quad (3.14)$$

$a_i \in \mathbb{K}$. Dalla (3.14) ricaviamo

$$a_{k+1} v_{k+1} = -a_1 v_1 - \dots - a_k v_k. \quad (3.15)$$

Vogliamo dimostrare che $a_i = 0$ per ogni $i = 1, \dots, k + 1$. A questo scopo, manipoliamo la (3.15) in due modi diversi. Moltiplicando ambo i membri per λ_{k+1} troviamo

$$a_{k+1} \lambda_{k+1} v_{k+1} = -a_1 \lambda_{k+1} v_1 - \dots - a_k \lambda_{k+1} v_k. \quad (3.16)$$

D'altro canto applicando f ad ambo i membri della (3.15) troviamo anche

$$\begin{aligned} f(a_{k+1}v_{k+1}) &= f(-a_1v_1 - \dots - a_kv_k) \\ \Leftrightarrow a_{k+1}f(v_{k+1}) &= -a_1f(v_1) - \dots - a_kf(v_k) \\ \Leftrightarrow a_{k+1}\lambda_{k+1}v_{k+1} &= -a_1\lambda_1v_1 - \dots - a_k\lambda_kv_k \end{aligned} \quad (3.17)$$

dove abbiamo usato, al primo passo, che f è un'applicazione lineare e, al secondo, che v_i è autovettore relativo all'autovalore λ_i per ogni $i = 1, \dots, k+1$. Infine, sottraendo la (3.17) dalla (3.16) membro a membro, troviamo

$$\begin{aligned} 0 &= -a_1\lambda_{k+1}v_1 - \dots - a_k\lambda_{k+1}v_k + a_1\lambda_1v_1 + \dots + a_k\lambda_kv_k \\ &= a_1(\lambda_1 - \lambda_{k+1})v_1 + \dots + a_k(\lambda_k - \lambda_{k+1})v_k. \end{aligned}$$

Ma per ipotesi di induzione gli autovettori v_1, \dots, v_k sono linearmente indipendenti, dunque

$$a_1(\lambda_1 - \lambda_{k+1}) = \dots = a_k(\lambda_k - \lambda_{k+1}) = 0.$$

Inoltre gli autovalori $\lambda_1, \dots, \lambda_k$ sono tutti distinti da λ_{k+1} , cioè $\lambda_i - \lambda_{k+1} \neq 0$, $i = 1, \dots, k$, perciò deve necessariamente essere $a_1 = \dots = a_k = 0$. Sostituendo nella (3.15) troviamo $a_{k+1}v_{k+1} = 0$, e siccome $v_{k+1} \neq 0$, anche $a_{k+1} = 0$. Questo conclude il passo di induzione e la dimostrazione. ■

Siamo ora pronti a dimostrare la Proposizione 3.2.3.

Dimostrazione della Proposizione 3.2.3. Consideriamo un vettore $v \in V_{\lambda_1} + \dots + V_{\lambda_p}$. Allora esistono vettori v_1, \dots, v_p in $V_{\lambda_1}, \dots, V_{\lambda_p}$ rispettivamente, tali che

$$v = v_1 + \dots + v_p. \quad (3.18)$$

Vogliamo dimostrare che la decomposizione (3.18) è unica. A questo scopo siano v'_1, \dots, v'_p (altri) vettori in $V_{\lambda_1}, \dots, V_{\lambda_p}$ rispettivamente, tali che

$$v = v_1 + \dots + v_p = v'_1 + \dots + v'_p. \quad (3.19)$$

Dobbiamo dimostrare che $v_i = v'_i$ per ogni $i = 1, \dots, p$. Dalla (3.19) segue che

$$0 = v_1 + \dots + v_p - v'_1 - \dots - v'_p = (v_1 - v'_1) + \dots + (v_p - v'_p).$$

Supponiamo per assurdo che, tra i vettori $v_i - v'_i$, ce ne siano alcuni diversi da zero. Senza ledere la generalità della dimostrazione, assumiamo che siano i primi q , con $q \leq p$, mentre gli ultimi $p - q$ sono uguali a zero (altrimenti, possiamo sempre rinominare gli autovalori, gli autospazi e i vettori in questione in modo che siano non nulli proprio i primi q tra i $v_i - v'_i$). In questo caso si ha

$$(v_1 - v'_1) + \dots + (v_q - v'_q) = 0.$$

Ma i vettori $v_1 - v'_1, \dots, v_q - v'_q$ sono non nulli, quindi sono autovettori relativi agli autovalori distinti $\lambda_1, \dots, \lambda_q$. Perciò, per il Lemma 3.2.4, sono indipendenti e la loro somma non può annullarsi (sarebbe altrimenti una combinazione lineare nulla a coefficienti non tutti nulli). Siamo così giunti ad una contraddizione e dobbiamo concludere che $v_i - v'_i = 0$, cioè $v_i = v'_i$, per ogni $i = 1, \dots, p$, come desiderato. ■

3.3 Il Criterio di Diagonalizzabilità

I risultati presentati nella Sezione 3.2 consentono di mettere a punto facilmente un criterio per stabilire se un dato endomorfismo $f : V \rightarrow V$ è diagonalizzabile oppure no. Vale infatti il seguente

Teorema 3.3.1 — Criterio di Diagonalizzabilità. Sia V uno spazio vettoriale di dimensione n e sia $f : V \rightarrow V$ un endomorfismo di f . Detti $\lambda_1, \dots, \lambda_k$ tutti gli autovalori distinti di f , le seguenti affermazioni sono equivalenti:

(I) f è diagonalizzabile;

(II) La somma delle molteplicità algebriche degli autovalori $\lambda_1, \dots, \lambda_k$ è pari ad n :

$$m_a(\lambda_1) + \dots + m_a(\lambda_k) = n, \quad (3.20)$$

e inoltre, per ogni $i = 1, \dots, k$, molteplicità algebrica e geometrica dell'autovalore λ_i coincidono:

$$m_a(\lambda_i) = m_g(\lambda_i), \quad i = 1, \dots, k. \quad (3.21)$$

(III) La somma delle molteplicità geometriche degli autovalori $\lambda_1, \dots, \lambda_k$ è pari ad n :

$$m_g(\lambda_1) + \dots + m_g(\lambda_k) = n;$$

Dimostrazione. Dimosteremo che $(I) \Rightarrow (II) \Rightarrow (III) \Rightarrow (I)$. Assumiamo innanzitutto che l'endomorfismo f sia diagonalizzabile e sia $\mathcal{R} = (v_1, \dots, v_n)$ un riferimento che diagonalizza f . Cambiando l'ordine dei vettori v_1, \dots, v_n possiamo sempre fare in modo che

(1) i primi m_1 vettori di \mathcal{R} siano autovettori relativi allo stesso autovalore λ_1 ;

(2) i successivi m_2 vettori di \mathcal{R} siano autovettori relativi ad un altro autovalore λ_2 ;

\vdots $\quad \quad \quad \vdots$

(k) gli ultimi m_k vettori di \mathcal{R} siano autovettori relativi ad un altro autovalore ancora λ_k ,

con $\lambda_1, \dots, \lambda_k$ autovalori tutti distinti e $m_1 + \dots + m_k = n$. La situazione è illustrata dal seguente schema:

$$\mathcal{R} = \left(\underbrace{v_1, \dots, v_{m_1}}_{\substack{m_1 \text{ autovettori} \\ \text{relativi a } \lambda_1}}, \underbrace{v_{m_1+1}, \dots, v_{m_1+m_2}}_{\substack{m_2 \text{ autovettori} \\ \text{relativi a } \lambda_2}}, \dots, \underbrace{v_{n-m_k+1}, \dots, v_n}_{\substack{m_k \text{ autovettori} \\ \text{relativi a } \lambda_k}} \right).$$

In altre parole, la matrice rappresentativa di f nel riferimento \mathcal{R} è

$$A = \begin{pmatrix} \lambda_1 I_{m_1} & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{m_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k I_{m_k} \end{pmatrix}.$$

Perciò

$$A_t = A - tI_n = \begin{pmatrix} (\lambda_1 - t)I_{m_1} & 0 & \cdots & 0 \\ 0 & (\lambda_2 - t)I_{m_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (\lambda_k - t)I_{m_k} \end{pmatrix}$$

e il polinomio caratteristico è

$$P_f(t) = \det A_t = (\lambda_1 - t)^{m_1} (\lambda_2 - t)^{m_2} \cdots (\lambda_k - t)^{m_k},$$

da cui leggiamo che $\lambda_1, \dots, \lambda_k$ sono *tutti* gli autovalori distinti di f e che le loro molteplicità algebriche sono proprio m_1, \dots, m_k : in altre parole, la molteplicità algebrica $m_a(\lambda_i)$ (dell'autovalore

λ_i di un endomorfismo diagonalizzabile) è anche il numero di autovettori relativi all'autovalore λ_i in un riferimento diagonalizzante. Poiché $m_1 + \dots + m_k = n$, la (3.20) è provata. Ora dobbiamo dimostrare che $m_a(\lambda_i) = m_g(\lambda_i)$ per ogni $i = 1, \dots, k$. A questo scopo possiamo usare la Proposizione 3.2.1 per calcolare le molteplicità geometriche: per esempio A_{λ_1} è la matrice

$$A_{\lambda_1} = A - \lambda_1 I_n = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & (\lambda_2 - \lambda_1)I_{m_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (\lambda_k - \lambda_1)I_{m_k} \end{pmatrix}.$$

Giacché $\lambda_1, \dots, \lambda_k$ sono tutti distinti, $\text{rk} A_{\lambda_1} = m_2 + \dots + m_k = n - m_1$ e

$$m_g(\lambda_1) = n - \text{rk} A_{\lambda_1} = m_1 = m_a(\lambda_1).$$

Esattamente nello stesso modo $m_g(\lambda_i) = m_a(\lambda_i)$ per ogni $i = 2, \dots, k$. Questo prova anche la (3.21). Abbiamo così dimostrato che (I) \Rightarrow (II).

Che (II) \Rightarrow (III) è ovvio e resta da dimostrare che (III) \Rightarrow (I). Questo segue dalla Proposizione 3.2.3. Infatti, supponiamo vera la (III). Poniamo $m_i := m_g(\lambda_i) = \dim V_{\lambda_i}$ e scegliamo riferimenti

$$\mathcal{R}_{\lambda_1} = (v_1, \dots, v_{m_1}), \quad \mathcal{R}_{\lambda_2} = (v_{m_1+1}, \dots, v_{m_1+m_2}), \quad \dots, \quad \mathcal{R}_{\lambda_k} = (v_{n-m_k+1}, \dots, v_n)$$

di $V_{\lambda_1}, \dots, V_{\lambda_k}$ rispettivamente. Consideriamo anche il sistema

$$\mathcal{R} = \underbrace{(v_1, \dots, v_{m_1})}_{\mathcal{R}_{\lambda_1}} \underbrace{(v_{m_1+1}, \dots, v_{m_1+m_2})}_{\mathcal{R}_{\lambda_2}} \dots \underbrace{(v_{n-m_k+1}, \dots, v_n)}_{\mathcal{R}_{\lambda_k}}. \quad (3.22)$$

Se dimostriamo che \mathcal{R} è un sistema indipendente, allora, poiché ha esattamente n elementi, è anche un riferimento di V e, per costruzione, è un riferimento di autovettori di f , perciò è un riferimento diagonalizzante e la dimostrazione si conclude. Dobbiamo dunque solo dimostrare che \mathcal{R} è un sistema indipendente. Consideriamo una combinazione lineare nulla dei vettori di \mathcal{R} :

$$\underbrace{\sum_{i=1}^{m_1} a_i v_i}_{:= u_1} + \underbrace{\sum_{i=m_1+1}^{m_1+m_2} a_i v_i}_{:= u_2} + \dots + \underbrace{\sum_{i=n-m_k+1}^n a_i v_i}_{:= u_k} = 0. \quad (3.23)$$

Se adottiamo la notazione indicata nella (3.23), si ha $u_i \in V_{\lambda_i}$ per ogni $i = 1, \dots, k$ (perché u_i è combinazione lineare dei vettori di \mathcal{R}_{λ_i} che è un riferimento di V_{λ_i}). Ma, in virtù della Proposizione 3.2.3, il vettore nullo si può scrivere in un unico modo come somma di vettori u_1, \dots, u_k in $V_{\lambda_1}, \dots, V_{\lambda_k}$, perciò deve essere $u_1 = \dots = u_k = 0$. In particolare

$$u_1 = \sum_{i=1}^{m_1} a_i v_i = 0.$$

Ma i vettori v_1, \dots, v_{m_1} sono linearmente indipendenti (formano il riferimento \mathcal{R}_{λ_1} di V_{λ_1}), quindi deve essere $a_i = 0$ per ogni $i = 1, \dots, m_1$. Esattamente nello stesso modo, $a_i = 0$ per ogni $i = m_1 + 1, \dots, n$. Questo mostra che la combinazione lineare (3.23) ha coefficienti tutti nulli e il sistema \mathcal{R} è indipendente, il che conclude la dimostrazione. ■

Corollario 3.3.2 Sia V uno spazio vettoriale di dimensione n e sia $f : V \rightarrow V$ un endomorfismo di f . Se f possiede esattamente n autovalori distinti, allora è diagonalizzabile.

Dimostrazione. Siano $\lambda_1, \dots, \lambda_n$ gli n autovalori distinti di f . Poiché $m_g(\lambda_i) \geq 1$ per ogni $i = 1, \dots, n$ abbiamo

$$n \leq m_g(\lambda_1) + \dots + m_g(\lambda_n) \leq n,$$

da cui $m_g(\lambda_1) + \dots + m_g(\lambda_n) = n$. Dunque vale la condizione (III) nell'enunciato del Teorema 3.3.1 ed f è diagonalizzabile (si osservi che, in questo caso, abbiamo anche, necessariamente $m_g(\lambda_i) = 1$ per ogni $i = 1, \dots, n$). ■

■ **Esempio 3.6** L'endomorfismo $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dell'Esempio 3.4 non è diagonalizzabile. Infatti, sebbene la somma delle molteplicità algebriche dei due autovalori sia $3 = \dim \mathbb{R}^3$, l'autovalore 1 ha molteplicità algebrica pari a 2 ma, come visto nell'Esempio 3.5, molteplicità geometrica pari ad 1. ■

Oltre a fornire, nell'enunciato, un criterio di diagonalizzabilità piuttosto efficiente, il Teorema 3.3.1 indica anche, nella dimostrazione, un metodo per determinare un riferimento diagonalizzante. Riassumendo, per stabilire se un endomorfismo f è diagonalizzabile ed eventualmente determinare un riferimento diagonalizzante si procede così:

- (1) si calcola il polinomio caratteristico $P_f(t)$;
- (2) se possibile, si risolve l'equazione caratteristica $P_f(t) = 0$ e si determinano così tutti gli autovalori distinti $\lambda_1, \dots, \lambda_k$ e le relative molteplicità algebriche $m_a(\lambda_1), \dots, m_a(\lambda_k)$. Se $m_a(\lambda_1) + \dots + m_a(\lambda_k) < n$ allora f non è diagonalizzabile, e la procedura si interrompe. Se invece $m_a(\lambda_1) + \dots + m_a(\lambda_k) = n$ allora la procedura continua con il passo successivo;
- (3) si calcolano le molteplicità geometriche $m_g(\lambda_1), \dots, m_g(\lambda_k)$, per esempio usando la Proposizione 3.2.1, oppure risolvendo direttamente i sistemi (3.10) con $\lambda = \lambda_1, \dots, \lambda_k$. Se $m_g(\lambda_j) < m_a(\lambda_j)$ per qualche j , allora f non è diagonalizzabile e la procedura si interrompe. Se invece $m_g(\lambda_i) = m_a(\lambda_i)$ per ogni $i = 1, \dots, k$, allora f è diagonalizzabile e la procedura continua con il passo successivo;
- (4) si determina un riferimento \mathcal{R}_{λ_i} per ciascun autospazio V_{λ_i} risolvendo i sistemi (3.10) con $\lambda = \lambda_1, \dots, \lambda_k$.
- (5) si costruisce un riferimento \mathcal{R} di autovettori di f mettendo insieme tutti i vettori dei sistemi \mathcal{R}_{λ_i} al variare di $i = 1, \dots, k$ come nella (3.22).

■ **Esempio 3.7** Consideriamo l'endomorfismo

$$f : \mathbb{R}^4 \rightarrow \mathbb{R}^4, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto \begin{pmatrix} -x_1 & & & \\ & x_2 & - & x_3 & + & 2x_4 \\ & & & & & - & x_4 \end{pmatrix}.$$

Vogliamo dimostrare che f è diagonalizzabile e determinare una base di autovettori di f . Calcoliamo innanzitutto il polinomio caratteristico $P_f(t)$. La matrice rappresentativa di f nel riferimento canonico è

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

Il polinomio caratteristico è

$$P_f(t) = \det A_t = \det \begin{pmatrix} -1-t & 0 & 0 & 0 \\ 0 & -t & 0 & 2 \\ 0 & 1 & -1-t & 2 \\ 0 & 0 & 0 & -1-t \end{pmatrix},$$

che si può calcolare, per esempio, sviluppando mediante la regola di Laplace rispetto alla prima riga (e poi rispetto all'ultima). Troviamo

$$P_f(t) = t(1+t)^3. \quad (3.24)$$

Dobbiamo ora risolvere l'equazione caratteristica $P_f(t) = 0$. La (3.24) fornisce una fattorizzazione completa di $P_f(t)$ da cui leggiamo che $P_f(t)$ ha due radici, 0 e -1 , con molteplicità rispettivamente 1 e 3. Dunque f ha due autovalori 0 e -1 con molteplicità algebriche 1 e 3. Si noti che la somma delle molteplicità algebriche è $1 + 3 = 4 = \dim \mathbb{R}^4$.

Ora calcoliamo le molteplicità geometriche. La matrice A_0 è

$$A_0 = A = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

e ha rango 3, infatti il minore individuato dalle prime tre righe e dalla prima, dalla seconda e dalla quarta colonna è

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

Perciò le prime tre righe di A_0 sono indipendenti. L'ultima dipende palesemente dalla seconda e $\text{rk} A_0 = 3$ come annunciato. Di conseguenza

$$m_g(0) = 4 - \text{rk} A_0 = 4 - 3 = 1 = m_a(0).$$

Si osservi che, giacché per ogni autovalore λ si ha $1 \leq m_g(\lambda) \leq m_a(\lambda)$, se $m_a(\lambda) = 1$, come in questo caso, si può concludere direttamente che $m_g(\lambda) = 1$ senza necessariamente calcolare il rango della matrice A_λ .

La matrice A_{-1} è

$$A_{-1} = A + I_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

ed ha chiaramente rango 1, sicché

$$m_g(-1) = 4 - \text{rk} A_{-1} = 4 - 1 = 3 = m_a(-1).$$

La condizione (II) nell'enunciato del Teorema 3.3.1 è dunque soddisfatta e l'endomorfismo f è diagonalizzabile.

Per determinare un riferimento di autovettori, calcoliamo prima un riferimento per ciascun autospazio. L'autospazio V_0 è il nucleo $\ker f$ di f cioè lo spazio delle soluzioni del sistema lineare

omogeneo $Ax = 0$. Si osservi che già sappiamo, in base ai conti fatti sopra che, delle 4 equazioni del sistema $Ax = 0$, solo 3 sono linearmente indipendenti (infatti $\text{rk} A = 3$). Il sistema $Ax = 0$ è

$$\begin{cases} -x_1 & & & = 0 \\ & & & 2x_4 = 0 \\ & x_2 - x_3 + 2x_4 & = 0 \\ & & -x_4 & = 0 \end{cases}$$

L'ultima equazione è equivalente alla seconda e può essere eliminata. Dalla prima e dalla seconda ricaviamo $x_1 = x_4 = 0$ e dalla terza $x_2 = x_3$. Dunque V_0 è il sottospazio 1-dimensionale dato da

$$V_0 = \left\{ \begin{pmatrix} 0 \\ s \\ s \\ 0 \end{pmatrix} : s \in \mathbb{R} \right\}. \quad (3.25)$$

Un riferimento per V_0 si trova ponendo, per esempio, $s = 1$ nella (3.25), ed è

$$\mathcal{R}_0 = \left(\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right).$$

L'autospazio V_{-1} è lo spazio delle soluzioni del sistema $A_{-1}x = 0$. Già sappiamo che delle 4 equazioni di questo sistema solo una è indipendente, per esempio la seconda che è

$$x_2 + 2x_4 = 0.$$

Perciò

$$V_{-1} = \left\{ \begin{pmatrix} s \\ -2u \\ t \\ u \end{pmatrix} : s, t, u \in \mathbb{R} \right\},$$

Ed un riferimento di V_{-1} si trova ponendo $(s, t, u) = E_1, E_2, E_3$ i vettori del riferimento canonico di \mathbb{R}^3 , ed è

$$\mathcal{R}_{-1} = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \end{pmatrix} \right).$$

Mettendo insieme i vettori di $\mathcal{R}_0, \mathcal{R}_{-1}$ in un unico sistema, troviamo un sistema \mathcal{R} di autovettori di f :

$$\mathcal{R} = \left(\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \end{pmatrix} \right).$$

Si osservi che la matrice rappresentativa di f nel riferimento \mathcal{R} è la matrice diagonale che si trova scrivendo gli autovalori $0, -1$ nell'ordine e nel numero opportuno sulla diagonale:

$$M_{\mathcal{R}}^f = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

■

■ **Esempio 3.8** Consideriamo l'endomorfismo

$$f: \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} -\frac{y}{4} + \frac{z}{2} \\ \frac{x}{2} - \frac{y}{4} \\ 0 \end{pmatrix}.$$

Vogliamo dimostrare che f è diagonalizzabile. La matrice rappresentativa di f nel riferimento canonico è

$$A = \begin{pmatrix} 0 & -1/4 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & -1/4 & 0 \end{pmatrix},$$

e il polinomio caratteristico è

$$P_f(t) = \det A_t = \det \begin{pmatrix} -t & -1/4 & 1/2 \\ 0 & 1-t & 0 \\ 1/2 & -1/4 & -t \end{pmatrix} = -(t-1)\left(t-\frac{1}{2}\right)\left(t+\frac{1}{2}\right),$$

Da cui leggiamo che f ha 3 autovalori distinti: $1, 1/2, -1/2$. Per il Corollario 3.3.2 f è dunque diagonalizzabile. Lasciamo al lettore il compito di determinare un riferimento di autovettori di f . ■

Esercizio 3.3 Dimostrare che l'endomorfismo

$$f: \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x - y \\ -x + y \\ 2z \end{pmatrix}$$

è diagonalizzabile. Determinare un riferimento di autovettori di f . ■

■ **Esempio 3.9** Si consideri ancora l'endomorfismo $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dell'Esempio 3.4. Abbiamo visto che la somma delle molteplicità algebriche degli autovalori di f è $3 = \dim \mathbb{R}^3$. Nonostante ciò, f non è diagonalizzabile perché, per uno dei due autovalori, la molteplicità geometrica è strettamente minore della molteplicità algebrica (Esempio 3.5). Dato un endomorfismo $g: V \rightarrow V$ può anche capitare che la somma delle molteplicità algebriche degli autovalori di g sia strettamente minore di $n = \dim V$. È questo il caso dell'endomorfismo

$$g: \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} y \\ -x \\ 0 \end{pmatrix}.$$

Infatti, la matrice rappresentativa di g nel riferimento canonico è

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

e il polinomio caratteristico è

$$P_f(t) = \det A_t = \det \begin{pmatrix} -t & 1 & 0 \\ -1 & -t & 0 \\ 0 & 0 & -t \end{pmatrix} = -t(t^2 + 1),$$

Da cui vediamo che g ha un unico autovalore, 0, con molteplicità algebrica pari ad 1. ■

Esercizio 3.4 Si consideri l'endomorfismo

$$f: \mathbb{Q}^4 \rightarrow \mathbb{Q}^4, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto \begin{pmatrix} -\frac{x_1}{2} + \frac{x_2}{2} + \frac{x_3}{12} + \frac{x_4}{12} \\ x_2 + \frac{x_3}{12} + \frac{x_4}{12} \\ -\frac{x_3}{4} + \frac{x_4}{4} \\ \frac{x_3}{4} - \frac{x_4}{4} \end{pmatrix}.$$

Si stabilisca se f è diagonalizzabile oppure no e, in ogni caso, si determini un riferimento per ogni autospazio. ■

■ **Esempio 3.10** Esistono addirittura endomorfismi privi di autovalori. Consideriamo, per esempio, l'endomorfismo

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x & y \\ -x & 0 \end{pmatrix}.$$

La matrice rappresentativa di f nel riferimento canonico è

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

e il polinomio caratteristico è

$$P_f(t) = \det A_t = \det \begin{pmatrix} -t & 1 \\ -1 & -t \end{pmatrix} = t^2 + 1,$$

che non ammette radici (reali). ■

Esercizio 3.5 Dimostrare che l'endomorfismo

$$f: \mathbb{Q}^2 \rightarrow \mathbb{Q}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x & 2y \\ x & 0 \end{pmatrix}$$

è privo di autovalori. ■

3.4 Diagonalizzazione di Matrici

Sia $A \in M_n(\mathbb{K})$ una matrice quadrata.

Definizione 3.4.1 — Autovalori e Autovettori di una Matrice. Gli *autovalori*, gli *autovettori* e gli *autospazi* di A sono, per definizione, gli autovalori, gli autovettori e gli autospazi dell'endomorfismo associato

$$L_A: \mathbb{K}^n \rightarrow \mathbb{K}^n, \quad x \mapsto L_A(x) := Ax.$$

Il *polinomio caratteristico* di A è il polinomio caratteristico di L_A :

$$P_A(t) := P_{L_A}(t).$$

La matrice A si dice *diagonalizzabile*, se l'endomorfismo L_A è diagonalizzabile.

Si osservi che il polinomio caratteristico della matrice A è semplicemente

$$P_A(t) = \det A_t = \det(A - tI_n).$$

Infatti A è anche la matrice rappresentativa di L_A nel riferimento canonico. Similmente, se λ è un autovalore di A , l'autospazio $V_\lambda \subseteq \mathbb{R}^n$ è semplicemente lo spazio delle soluzioni del sistema lineare omogeneo:

$$A_\lambda x = 0.$$

Proposizione / Definizione 3.4.1 Una matrice quadrata $A \in M_n(\mathbb{K})$ è diagonalizzabile sse esiste una matrice invertibile $M \in GL_n(\mathbb{K})$ tale che la matrice

$$D := M^{-1}AM$$

è diagonale. In tal caso si dice che M è una *matrice diagonalizzante* per A .

Dimostrazione. Supponiamo che A sia diagonalizzabile. Allora L_A è diagonalizzabile ed esiste un riferimento $\mathcal{R} = (M^{(1)}, \dots, M^{(n)})$ che diagonalizza L_A . In altre parole la matrice $D = M_{\mathcal{R}}^{L_A}$ è una matrice diagonale. Ma, ricordando che

- (1) la matrice rappresentativa di L_A nel riferimento canonico \mathcal{R}_{can} è proprio A ,
- (2) la matrice del cambio di coordinate dal riferimento \mathcal{R} al riferimento canonico è proprio la matrice M che ha i vettori numerici $M^{(1)}, \dots, M^{(n)}$ come colonne,

per le proprietà della matrice rappresentativa troviamo

$$D = M_{\mathcal{R}}^{L_A} = M_{\mathcal{R}, \mathcal{R}_{can}} M_{\mathcal{R}_{can}}^{L_A} M_{\mathcal{R}_{can}, \mathcal{R}} = M^{-1}AM, \quad (3.26)$$

come desiderato. Viceversa, supponiamo che esista una matrice invertibile M tale che $M^{-1}AM$ è una matrice diagonale. Allora le colonne $(M^{(1)}, \dots, M^{(n)})$ di M formano un riferimento \mathcal{R} di \mathbb{K}^n e lo stesso conto (3.26) di prima mostra che \mathcal{R} è un riferimento diagonalizzante per L_A , dunque A è diagonalizzabile. ■

Esercizio 3.6 Si dimostri che una matrice quadrata $A \in M_n(\mathbb{K})$ è diagonalizzabile sse esiste una matrice invertibile $P \in GL_n(\mathbb{K})$ tale che la matrice PAP^{-1} è diagonale. ■

Sia $A \in M_n(\mathbb{K})$ una matrice diagonalizzabile. Sappiamo che esiste una matrice diagonalizzante M per A . Supponiamo di voler determinare M . La dimostrazione della Proposizione 3.4.1 mostra come fare: prima determiniamo un riferimento \mathcal{R} diagonalizzante per L_A . I vettori di \mathcal{R} sono allora le colonne della matrice M cercata.

■ **Esempio 3.11** Si consideri la matrice

$$A = \begin{pmatrix} 1/2 & 1/2 & 1/2 \\ 1 & 2/3 & 1 \\ -1/2 & -1/6 & -1/2 \end{pmatrix} \in M_3(\mathbb{R}),$$

Dimostriamo che esiste una matrice invertibile $M \in GL_3(\mathbb{R})$ tale che la matrice $D := M^{-1}AM$ è una matrice diagonale. Dobbiamo dimostrare che A è diagonalizzabile. Procediamo nel solito modo: il polinomio caratteristico di A è

$$\begin{aligned} P_A(t) &= \det A_t = \det \begin{pmatrix} 1/2-t & 1/2 & 1/2 \\ 1 & 2/3-t & 1 \\ -1/2 & -1/6 & -1/2-t \end{pmatrix} \\ &= -t^3 + \frac{2}{3}t^2 + \frac{1}{3}t = -t(t-1)\left(t + \frac{1}{3}\right). \end{aligned}$$

Dunque A ha 3 autovalori distinti: $0, 1, -1/3$, ed è perciò diagonalizzabile. Determiniamo un riferimento diagonalizzante per L_A . L'autospazio V_0 relativo all'autovalore 0 è il nucleo di L_A cioè lo spazio delle soluzioni del sistema $Ax = 0$: esplicitamente

$$\begin{cases} \frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 = 0 \\ x_1 + \frac{2}{3}x_2 + x_3 = 0 \\ -\frac{1}{2}x_1 - \frac{1}{6}x_2 - \frac{1}{2}x_3 = 0 \end{cases} \quad (3.27)$$

Già sappiamo che, di queste 3 equazioni, solo due sono indipendenti, per esempio le prime due. In altre parole, il sistema (3.27) è equivalente a

$$\begin{cases} \frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 = 0 \\ x_1 + \frac{2}{3}x_2 + x_3 = 0 \end{cases} \Leftrightarrow \begin{cases} x_1 + x_2 + x_3 = 0 \\ x_1 + \frac{2}{3}x_2 + x_3 = 0 \end{cases}.$$

Sottraendo membro a membro troviamo $x_2 = 0$, da cui $x_3 = -x_1$. Ponendo $x_1 = 1$ troviamo che un riferimento per V_0 è costituito dal vettore

$$\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

L'autospazio V_1 relativo all'autovalore 1 è lo spazio delle soluzioni del sistema $A_1x = 0$:

$$\begin{cases} -\frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 = 0 \\ x_1 - \frac{1}{3}x_2 + x_3 = 0 \\ -\frac{1}{2}x_1 - \frac{1}{6}x_2 - \frac{3}{2}x_3 = 0 \end{cases} \quad (3.28)$$

Solo due di queste equazioni sono indipendenti, per esempio le prime due. Perciò V_1 è lo spazio delle soluzioni del sistema

$$\begin{cases} -\frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 = 0 \\ x_1 - \frac{1}{3}x_2 + x_3 = 0 \\ -\frac{1}{2}x_1 - \frac{1}{6}x_2 - \frac{3}{2}x_3 = 0 \end{cases} \Leftrightarrow \begin{cases} -x_1 + x_2 + x_3 = 0 \\ -x_1 - \frac{1}{3}x_2 - 3x_3 = 0 \end{cases}.$$

Dalla prima equazione ricaviamo $x_1 = x_2 + x_3$ e, sostituendo nella seconda, $x_2 = -3x_3$, da cui anche $x_1 = -2x_3$. Ponendo $x_3 = 1$, troviamo che un riferimento per V_1 è costituito dal vettore

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

Infine, l'autospazio $V_{-1/3}$ relativo all'autovalore $-1/3$ è lo spazio delle soluzioni del sistema $A_{-1/3}x = 0$: esplicitamente

$$\begin{cases} \frac{5}{6}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 = 0 \\ x_1 + x_2 + x_3 = 0 \\ -\frac{1}{2}x_1 - \frac{1}{6}x_2 - \frac{1}{6}x_3 = 0 \end{cases} \quad (3.29)$$

Di nuovo, dalla teoria, di queste 3 equazioni, solo due sono indipendenti, per esempio le prime due. Quindi $V_{-1/3}$ è lo spazio delle soluzioni del sistema

$$\begin{cases} \frac{5}{6}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 = 0 \\ x_1 + x_2 + x_3 = 0 \end{cases} \Leftrightarrow \begin{cases} \frac{5}{3}x_1 + x_2 + x_3 = 0 \\ x_1 + x_2 + x_3 = 0 \end{cases}.$$

Sottraendo membro a membro troviamo $x_1 = 0$ da cui anche $x_2 = -x_3$. Un riferimento per $V_{1/3}$ è perciò costituito dal vettore

$$\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

e una matrice diagonalizzante per A è la matrice

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

Per scrivere la matrice diagonale $D = M^{-1}AM$ non occorre invertire M né tantomeno moltiplicare le matrici M^{-1} , A , M : la teoria dice che sulla diagonale di D ci sono gli autovalori di A nello stesso ordine in cui i relativi autovettori compaiono come colonne della matrice M e dunque

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

Esercizio 3.7 Si consideri la matrice

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & -1/2 & 0 \\ 0 & 1/2 & 1 \end{pmatrix} \in M_3(\mathbb{Q}).$$

Dopo aver dimostrato che A è diagonalizzabile, si determini una matrice diagonalizzante $M \in GL_3(\mathbb{Q})$. Si scriva anche esplicitamente la matrice diagonale $M^{-1}AM$.

3.5 Diagonalizzazione Ortogonale

Affrontiamo ora il problema della diagonalizzazione nel contesto degli spazi vettoriali euclidei. Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo finito-dimensionale e sia $f : V \rightarrow V$ un endomorfismo. In questa situazione è naturale cercare un riferimento \mathcal{R} che sia simultaneamente diagonalizzante e ortonormale, in altre parole un *riferimento ortonormale di autovettori*. Questo suggerisce la seguente.

Definizione 3.5.1 — Endomorfismo Ortogonalmente Diagonalizzabile. Un endomorfismo $f : V \rightarrow V$ di uno spazio vettoriale euclideo $(V, \langle -, - \rangle)$ si dice *ortogonalmente diagonalizzabile* se esiste un riferimento ortonormale $\mathcal{R} = (e_1, \dots, e_n)$ tale che la matrice rappresentativa $M_{\mathcal{R}}^f$ di f nel riferimento \mathcal{R} è una matrice diagonale.

Si osservi che un endomorfismo ortogonalmente diagonalizzabile è, in particolare, diagonalizzabile. Il prossimo teorema fornisce una caratterizzazione particolarmente semplice degli endomorfismi ortogonalmente diagonalizzabili, ed è anche il principale risultato discusso in questa sezione.

Teorema 3.5.1 — Teorema Spettrale. Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo di dimensione n . Un endomorfismo $f : V \rightarrow V$ è ortogonalmente diagonalizzabile sse è simmetrico.

Prima di dimostrare il Teorema Spettrale occorre discutere un lemma che ha anche un'importanza indipendente.

Lemma 3.5.2 Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo di dimensione n e sia $f : V \rightarrow V$ un endomorfismo simmetrico. Allora f ammette almeno un autovalore. Se $\lambda_1, \dots, \lambda_k$ sono tutti gli autovalori distinti di f , allora la somma delle molteplicità algebriche $m_a(\lambda_1) + \dots + m_a(\lambda_k)$ è pari ad n .

Dimostrazione. La dimostrazione di questo lemma sfrutta alcune costruzioni con i numeri complessi che verranno discusse nel prossimo capitolo ed è perciò rinviata (si veda il Lemma 4.3.13). ■

Siamo ora pronti a dimostrare il Teorema Spettrale.

Dimostrazione del Teorema 3.5.1. Supponiamo innanzitutto che f sia ortogonalmente diagonalizzabile e sia $\mathcal{R} = (e_1, \dots, e_n)$ un riferimento ortonormale di autovettori di f . La matrice rappresentativa di f nel riferimento \mathcal{R} è diagonale, in particolare è simmetrica, perciò, in virtù della Proposizione 2.5.3, f è un endomorfismo simmetrico.

Dimostrare che gli endomorfismi simmetrici sono ortogonalmente diagonalizzabili è più laborioso. Si procede per induzione sulla dimensione n di V . Se $n = 1$, è chiaro che ogni endomorfismo di V è ortogonalmente diagonalizzabile, perché ogni versore e di V costituisce un riferimento ortonormale nel quale la matrice rappresentativa di f è (una matrice di ordine 1×1 e perciò è in particolare) diagonale. Supponiamo ora che l'enunciato sia vero quando $n = k$ e dimostriamo che l'enunciato è vero anche quando $n = k + 1$. Sia dunque $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo di dimensione $k + 1$ e sia $f : (V, \langle -, - \rangle) \rightarrow (V, \langle -, - \rangle)$ un endomorfismo simmetrico. In virtù del Lemma 3.5.2, f possiede un autovalore $\lambda \in \mathbb{R}$ e dunque possiede anche un autovettore v . Giacché $v \neq 0$, si ha $\|v\| \neq 0$ e possiamo considerare il vettore

$$e_0 := \frac{1}{\|v\|} v.$$

Si tratta di un autovettore relativo all'autovalore λ . In più $\|e_0\| = 1$, cioè e_0 è un versore (il lettore lo vede?). Consideriamo il complemento ortogonale $W := e_0^\perp$. Il sottospazio W è f -invariante, cioè per ogni $w \in W$, anche $f(w) \in W$, infatti

$$\langle f(w), e_0 \rangle = \langle w, f(e_0) \rangle = \langle w, \lambda e_0 \rangle = \lambda \langle w, e_0 \rangle = 0,$$

dove abbiamo usato che f è un endomorfismo simmetrico, e, nell'ultimo passaggio, che $w \perp e_0$. Questo mostra che anche $f(w) \perp e_0$ come annunciato. Possiamo perciò restringere l'endomorfismo f a W simultaneamente nel dominio e nel codominio ottenendo un'applicazione

$$f_W : W \rightarrow W, \quad w \mapsto f_W(w) := f(w).$$

Chiaramente f_W è un'applicazione lineare. Inoltre si tratta di un endomorfismo simmetrico del sottospazio vettoriale euclideo $(W, \langle -, - \rangle_W)$. Infatti, per ogni $w_1, w_2 \in W$

$$\langle f_W(w_1), w_2 \rangle_W = \langle f(w_1), w_2 \rangle = \langle w_1, f(w_2) \rangle = \langle w_1, f_W(w_2) \rangle_W,$$

dove abbiamo usato le definizioni di f_W e $\langle -, - \rangle_W$, oltre al fatto che f è un endomorfismo simmetrico di $(V, \langle -, - \rangle)$ per ipotesi (il lettore è in grado di giustificare in dettaglio ogni passaggio?). Poiché $\dim W = k + 1 - \dim \text{Span}(e_0) = k + 1 - 1 = k$, all'endomorfismo simmetrico f_W di $(W, \langle -, - \rangle_W)$ si applica l'ipotesi di induzione. Dunque esiste un riferimento ortonormale (e_1, \dots, e_k) di W formato da autovettori di f_W . Consideriamo il sistema (e_0, e_1, \dots, e_k) di V . Poiché $e_1, \dots, e_k \in W = e_0^\perp$, si ha

$$\langle e_0, e_j \rangle = 0, \quad \text{per ogni } j = 1, \dots, k.$$

Inoltre per ogni $i, j = 1, \dots, k$

$$\langle e_i, e_j \rangle = \langle e_i, e_j \rangle_W = \delta_{ij}.$$

Quindi e_0, e_1, \dots, e_k sono versori ortogonali di V , cioè (e_0, e_1, \dots, e_k) è un riferimento ortonormale di V . D'altro canto, per ogni $i = 1, \dots, k$,

$$f(e_i) = f_W(e_i) = \lambda_i e_i$$

per qualche $\lambda_i \in \mathbb{R}$. Quindi gli e_i , oltre ad essere autovettori di f_W , sono anche autovettori di f . Concludiamo che (e_0, e_1, \dots, e_k) è un riferimento ortonormale di $(V, \langle -, - \rangle)$ formato da autovettori di f ed f è ortogonalmente diagonalizzabile come desiderato. ■

La dimostrazione del Teorema Spettrale suggerisce un procedimento induttivo per determinare un riferimento ortonormale di autovettori di un endomorfismo simmetrico. Esiste tuttavia un procedimento più diretto, basato sul seguente

Lemma 3.5.3 Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo di dimensione n , e sia $f : V \rightarrow V$ un endomorfismo simmetrico. Siano λ, μ autovalori distinti di f . Allora gli autospazi V_λ, V_μ sono ortogonali, cioè $v \perp u$, per ogni $v \in V_\lambda$ e ogni $u \in V_\mu$.

Dimostrazione. Siano $v \in V_\lambda$ e $u \in V_\mu$. Calcoliamo

$$\lambda \langle v, u \rangle = \langle \lambda v, u \rangle = \langle f(v), u \rangle = \langle v, f(u) \rangle = \langle v, \mu u \rangle = \mu \langle v, u \rangle$$

(il lettore è in grado di giustificare tutti i passaggi?). Questo mostra che $\lambda \langle v, u \rangle = \mu \langle v, u \rangle$ cioè

$$0 = \lambda \langle v, u \rangle - \mu \langle v, u \rangle = (\lambda - \mu) \langle v, u \rangle. \quad (3.30)$$

Poiché $\lambda \neq \mu$, si ha $\lambda - \mu \neq 0$, perciò, dalla (3.30), necessariamente $\langle v, u \rangle = 0$ come desiderato. ■

Corollario 3.5.4 Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo di dimensione n , sia $f : V \rightarrow V$ un endomorfismo simmetrico di V e siano $\lambda_1, \dots, \lambda_k$ tutti gli autovalori distinti di f . Se

$$\mathcal{R}_{\lambda_1} = (e_1, \dots, e_{m_1}), \quad \mathcal{R}_{\lambda_2} = (e_{m_1+1}, \dots, e_{m_1+m_2}), \quad \dots, \quad \mathcal{R}_{\lambda_k} = (e_{n-m_k+1}, \dots, e_n)$$

sono riferimenti ortonormali per gli autospazi $V_{\lambda_1}, V_{\lambda_2}, \dots, V_{\lambda_k}$ rispettivamente, allora

$$\underbrace{(e_1, \dots, e_{m_1})}_{V_{\lambda_1}}, \underbrace{(e_{m_1+1}, \dots, e_{m_1+m_2})}_{V_{\lambda_2}}, \dots, \underbrace{(e_{n-m_k+1}, \dots, e_n)}_{V_{\lambda_k}}$$

è un riferimento ortonormale di V formato da autovettori di f .

Dimostrazione. La dimostrazione è molto semplice ed è lasciata come Esercizio 3.8. ■

Esercizio 3.8 Dimostrare il Corollario 3.5.4. ■

Il Corollario 3.5.4 suggerisce la seguente procedura per determinare un riferimento ortonormale di autovettori di un endomorfismo simmetrico f :

- (1) si determinano gli autovalori $\lambda_1, \dots, \lambda_k$ di f ;
- (2) si determina un riferimento per ciascun autospazio V_{λ_i} , $i = 1, \dots, k$;
- (3) utilizzando per esempio il procedimento di ortogonalizzazione di Gram-Schmidt si ortonormalizzano i riferimenti trovati al passo (2) e si trova così un riferimento ortonormale \mathcal{R}_{λ_i} per ciascun autospazio V_{λ_i} , $i = 1, \dots, k$;
- (4) si uniscono i riferimenti \mathcal{R}_{λ_i} in un unico riferimento ortonormale di V di autovettori di f (come nel Corollario 3.5.4 appunto).

■ **Esempio 3.12** Consideriamo l'endomorfismo

$$f: \mathbb{R}^4 \rightarrow \mathbb{R}^4, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto \begin{pmatrix} \frac{x_1}{2} - \frac{x_2}{2} - \frac{x_3}{2} \\ -\frac{x_1}{2} + \frac{x_2}{2} - \frac{x_3}{2} \\ -\frac{x_1}{2} - \frac{x_2}{2} + \frac{x_3}{2} \\ x_4 \end{pmatrix}.$$

La matrice rappresentativa di f nel riferimento canonico è

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

che è una matrice simmetrica. Poiché il riferimento canonico è ortonormale rispetto al prodotto scalare standard, f è un endomorfismo simmetrico dello spazio vettoriale euclideo standard $(\mathbb{R}^4, \langle -, - \rangle_{can})$. Cerchiamo ora un riferimento ortonormale di \mathbb{R}^4 formato da autovettori di f .

Il polinomio caratteristico di f è

$$P_f(t) = \det A_t = \det \begin{pmatrix} 1/2-t & -1/2 & -1/2 & 0 \\ -1/2 & 1/2-t & -1/2 & 0 \\ -1/2 & -1/2 & 1/2-t & 0 \\ 0 & 0 & 0 & 1-t \end{pmatrix} = \left(t + \frac{1}{2}\right)(t-1)^3.$$

Dunque f ha 2 autovalori, $-1/2$ e 1 , con molteplicità algebriche 1 e 3 rispettivamente. Poiché f è ortogonalmente diagonalizzabile, in particolare è diagonalizzabile, conosciamo già anche le molteplicità geometriche dei due autovalori: esse coincidono con le molteplicità algebriche.

Determiniamo ora un riferimento in ciascun autospazio. L'autospazio $V_{-1/2}$ è lo spazio delle soluzioni del sistema $A_{-1/2}x = 0$. Delle 4 equazioni del sistema $A_{-1/2}x = 0$, solo 3 sono linearmente indipendenti, per esempio le ultime 3 che sono:

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

Dall'ultima equazione $x_4 = 0$. Nelle altre due, si può scegliere per esempio x_3 come parametro e si trova

$$\begin{cases} x_1 = x_3 \\ x_2 = x_3 \end{cases}$$

Ponendo $x_3 = 1$, vediamo che un riferimento di $V_{-1/2}$ è

$$\tilde{\mathcal{R}}_{-1/2} = \left(\begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right). \quad (3.31)$$

Per ottenere un riferimento ortonormale dobbiamo normalizzare. La norma dell'unico vettore di $\tilde{\mathcal{R}}_{-1/2}$ è

$$\sqrt{1^2 + 1^2 + 1^2} = \sqrt{3}.$$

Un riferimento ortonormale si ottiene dividendo per tale norma ed è dunque

$$\mathcal{R}_{-1/2} = \left(\left(\begin{array}{c} \sqrt{3}/3 \\ \sqrt{3}/3 \\ \sqrt{3}/3 \\ 0 \end{array} \right) \right).$$

L'autospazio V_1 è lo spazio delle soluzioni del sistema $A_1 x = 0$. Delle 4 equazioni di tale sistema solo una è indipendente, per esempio la prima che è

$$-\frac{1}{2}x_1 - \frac{1}{2}x_2 - \frac{1}{2}x_3 = 0,$$

da cui $x_1 = -x_2 - x_3$ e un riferimento per V_1 è

$$\tilde{\mathcal{R}}_1 = \left(\left(\begin{array}{c} -1 \\ 1 \\ 0 \\ 0 \end{array} \right), \left(\begin{array}{c} -1 \\ 0 \\ 1 \\ 0 \end{array} \right), \left(\begin{array}{c} 0 \\ 0 \\ 0 \\ 1 \end{array} \right) \right).$$

Per trovare un riferimento ortonormale di V_1 applichiamo ad $\tilde{\mathcal{R}}_1$ il procedimento di ortogonalizzazione di Gram-Schmidt (prima, e poi normalizziamo). Poniamo

$$v_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Dunque $u_1 = v_1$. Abbiamo

$$\langle u_1, u_1 \rangle_{can} = 2, \quad \text{e} \quad \langle u_1, v_2 \rangle_{can} = 1,$$

perciò

$$u_2 = v_2 - \frac{\langle u_1, v_2 \rangle_{can}}{\langle u_1, u_1 \rangle_{can}} u_1 = \begin{pmatrix} -1/2 \\ -1/2 \\ 1 \\ 0 \end{pmatrix}.$$

Abbiamo anche

$$\langle u_2, u_2 \rangle_{can} = \frac{3}{2}, \quad \langle u_1, v_3 \rangle_{can} = 0 \quad \text{e} \quad \langle u_2, v_3 \rangle_{can} = 0,$$

perciò

$$u_3 = v_3 - \frac{\langle u_1, v_3 \rangle_{can}}{\langle u_1, u_1 \rangle_{can}} u_1 - \frac{\langle u_2, v_3 \rangle_{can}}{\langle u_2, u_2 \rangle_{can}} u_2 = v_3.$$

Calcoliamo anche

$$\frac{1}{\|u_1\|} u_1 = \begin{pmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \\ 0 \end{pmatrix}, \quad \frac{1}{\|u_2\|} u_2 = \begin{pmatrix} -\sqrt{6}/6 \\ -\sqrt{6}/6 \\ \sqrt{6}/3 \\ 0 \end{pmatrix}, \quad \frac{1}{\|u_3\|} u_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Concludiamo che un riferimento ortonormale per V_1 è

$$\mathcal{R}_1 = \left(\frac{1}{\|u_1\|}u_1, \frac{1}{\|u_2\|}u_2, \frac{1}{\|u_3\|}u_3 \right) = \left(\begin{pmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -\sqrt{6}/6 \\ -\sqrt{6}/6 \\ \sqrt{6}/3 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right)$$

e un riferimento ortonormale di $(\mathbb{R}^4, \langle -, - \rangle_{can})$ di autovettori di f è

$$\left(\begin{pmatrix} \sqrt{3}/3 \\ \sqrt{3}/3 \\ \sqrt{3}/3 \\ 0 \end{pmatrix}, \begin{pmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -\sqrt{6}/6 \\ -\sqrt{6}/6 \\ \sqrt{6}/3 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right)$$

Esercizio 3.9 Si consideri l'endomorfismo

$$f : \mathbb{R}^4 \rightarrow \mathbb{R}^4, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ x_2 \\ \frac{x_3}{4} + \frac{\sqrt{3}x_4}{4} \\ \frac{\sqrt{3}x_3}{4} + \frac{3x_4}{4} \end{pmatrix}.$$

dello spazio vettoriale euclideo standard $(\mathbb{R}^4, \langle -, - \rangle_{can})$. Dopo aver dimostrato che f è ortogonalmente diagonalizzabile, si determini un riferimento ortonormale diagonalizzante. ■

Sia $A \in M_n(\mathbb{R})$ una matrice quadrata (reale).

Definizione 3.5.2 — Matrice Ortogonalmente Diagonalizzabile. La matrice A si dice *ortogonalmente diagonalizzabile* se l'endomorfismo associato $L_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ è un endomorfismo ortogonalmente diagonalizzabile dello spazio vettoriale euclideo standard n -dimensionale $(\mathbb{R}^n, \langle -, - \rangle_{can})$.

Proposizione 3.5.5 Una matrice quadrata $A \in M_n(\mathbb{R})$ è ortogonalmente diagonalizzabile sse è simmetrica, cioè $A = A^T$.

Dimostrazione. Per definizione, A è ortogonalmente diagonalizzabile sse l'endomorfismo $L_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ è ortogonalmente diagonalizzabile rispetto al prodotto scalare standard. In virtù del Teorema Spettrale questo avviene sse L_A è un endomorfismo simmetrico (rispetto a $\langle -, - \rangle_{can}$) e questo avviene, a sua volta, sse la matrice rappresentativa di L_A in un riferimento ortonormale di $(\mathbb{R}^n, \langle -, - \rangle_{can})$ è una matrice simmetrica (Proposizione 2.5.3). Ma A è proprio la matrice rappresentativa di L_A nel riferimento canonico, che è un riferimento ortonormale rispetto a $\langle -, - \rangle_{can}$. Questo conclude la dimostrazione. ■

Proposizione 3.5.6 Una matrice quadrata $A \in M_n(\mathbb{R})$ è ortogonalmente diagonalizzabile (equivalentemente simmetrica) sse esiste una matrice ortogonale $O \in O_n$ tale che la matrice

$$D := O^{-1}AO = O^T AO$$

è diagonale, cioè esiste una matrice ortogonale diagonalizzante.

Dimostrazione. La dimostrazione è simile alla dimostrazione della Proposizione 3.4.1 ed è lasciata per Esercizio 3.10. ■

Esercizio 3.10 Dimostrare la Proposizione 3.5.6. Dimostrare anche che una matrice quadrata $A \in M_n(\mathbb{R})$ è ortogonalmente diagonalizzabile (equivalentemente simmetrica) sse esiste una matrice ortogonale $Q \in O_n$ tale che la matrice $QAQ^{-1} = QAQ^T$ è diagonale. ■

Sia $A \in M_n(\mathbb{R})$ una matrice simmetrica. Sappiamo che A è ortogonalmente diagonalizzabile, perciò esiste una matrice ortogonale diagonalizzante O per A . Per determinare O possiamo procedere così: prima determiniamo un riferimento \mathcal{R} ortonormale e diagonalizzante per L_A . I vettori di \mathcal{R} sono allora le colonne della matrice O cercata. Il lettore dimostri per esercizio che questa procedura funziona effettivamente discutendone tutti i dettagli.

■ **Esempio 3.13** Si consideri la matrice

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \in M_3(\mathbb{R}).$$

Vogliamo determinare una matrice ortogonale $O \in O_3$ tale che $O^{-1}AO = O^T A O$ è una matrice diagonale. A questo scopo determiniamo innanzitutto un riferimento ortonormale di $(\mathbb{R}^3, \langle -, - \rangle_{can})$ di autovettori di A . Il polinomio caratteristico di A è

$$P_A(t) = \det \begin{pmatrix} 1-t & 1 & 1 \\ 1 & 1-t & 1 \\ 1 & 1 & 1-t \end{pmatrix} = -t^2(t-3).$$

Dunque A ha 2 autovalori, 0 e 3, con molteplicità algebriche 2 e 1. Poiché A è una matrice simmetrica è anche ortogonalmente diagonalizzabile, in particolare diagonalizzabile, dunque le molteplicità geometriche di entrambi gli autovalori coincidono con le molteplicità algebriche. L'autospazio V_0 è lo spazio delle soluzioni del sistema $A_0 x = Ax = 0$. Una sola delle equazioni del sistema è indipendente, per esempio la prima che è

$$x_1 + x_2 + x_3 = 0.$$

Perciò V_0 è generato, per esempio, dal sistema

$$\tilde{\mathcal{R}} = \left(\left(\begin{array}{c} -1 \\ 1 \\ 0 \end{array} \right), \left(\begin{array}{c} -1 \\ 0 \\ 1 \end{array} \right) \right).$$

Un riferimento ortonormale di V_0 si trova ortonormalizzando $\tilde{\mathcal{R}}$. Poniamo

$$v_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad \text{e} \quad v_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix},$$

e applichiamo il procedimento di ortogonalizzazione di Gram-Schmidt:

$$u_1 = v_1 \quad \Rightarrow \quad \langle u_1, u_1 \rangle_{can} = 2, \quad \text{e} \quad \langle u_1, v_2 \rangle_{can} = 1,$$

da cui

$$u_2 = v_2 - \frac{\langle u_1, v_2 \rangle_{can}}{\langle u_1, u_1 \rangle_{can}} u_1 = \begin{pmatrix} -1/2 \\ -1/2 \\ 1 \end{pmatrix} \quad \Rightarrow \quad \langle u_2, u_2 \rangle_{can} = 3/2,$$

e un riferimento ortonormale per V_0 è

$$\mathcal{R} = \left(\frac{1}{\|u_1\|} u_1, \frac{1}{\|u_2\|} u_2 \right) = \left(\left(\begin{array}{c} -\sqrt{2}/2 \\ +\sqrt{2}/2 \\ 0 \end{array} \right), \left(\begin{array}{c} -\sqrt{6}/6 \\ -\sqrt{6}/6 \\ \sqrt{6}/3 \end{array} \right) \right).$$

L'autospazio V_3 è lo spazio delle soluzioni del sistema $A_3 x = 0$. Solo due equazioni del sistema sono indipendenti, per esempio le prime due che sono

$$\begin{cases} -2x_1 + x_2 + x_3 = 0 \\ x_1 - 2x_2 + x_3 = 0 \end{cases}.$$

Risolvendo si trova che V_3 è generato dal sistema

$$\left(\left(\begin{array}{c} 1 \\ 1 \\ 1 \end{array} \right) \right).$$

Posto

$$u_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

abbiamo

$$\langle u_3, u_3 \rangle_{can} = 3,$$

perciò un riferimento ortonormale di V_2 è

$$\left(\frac{1}{\|u_3\|} u_3 \right) = \left(\left(\begin{array}{c} \sqrt{3}/3 \\ \sqrt{3}/3 \\ \sqrt{3}/3 \end{array} \right) \right).$$

Concludendo, la matrice O cercata è

$$O = \begin{pmatrix} -\sqrt{2}/2 & -\sqrt{6}/6 & \sqrt{3}/3 \\ \sqrt{2}/2 & -\sqrt{6}/6 & \sqrt{3}/3 \\ 0 & \sqrt{6}/3 & \sqrt{3}/3 \end{pmatrix}$$

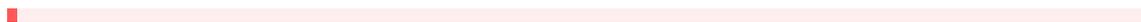
e

$$O^{-1} A O = O^T A O = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Esercizio 3.11 Si consideri la matrice

$$A = \begin{pmatrix} -1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \in M_4(\mathbb{R})$$

e si determini una matrice ortogonale $O \in O_4$ tale che $O^{-1} A O = O^T A O$ è una matrice diagonale. Dopo aver trovato O si calcoli la matrice diagonale $O^T A O$. ■



4. Spazi Vettoriali Hermitiani

Nella definizione di prodotto scalare su uno spazio vettoriale, è cruciale che il campo degli scalari \mathbb{K} sia il campo dei numeri reali \mathbb{R} . È naturale chiedersi se i prodotti scalari hanno un analogo se \mathbb{K} non è \mathbb{R} . In questo capitolo studieremo con qualche dettaglio il caso $\mathbb{K} = \mathbb{C}$: il campo dei numeri complessi. In questo caso esistono analoghi dei prodotti scalari, detti *prodotti hermitiani*. I prodotti hermitiani sono oggetti importanti sia in Matematica che in Fisica (particolarmente in Meccanica Quantistica) e in questo capitolo ne studieremo le principali proprietà. Molte di queste ultime assomigliano alle analoghe per i prodotti scalari. In alcuni aspetti, però, i prodotti hermitiani si differenziano significativamente dai loro *cugini reali*. Come vedremo la teoria dei prodotti hermitiani trova applicazioni anche all'algebra lineare degli spazi vettoriali euclidei.

4.1 Richiami sui Numeri Complessi

In questa sezione introduciamo il campo \mathbb{C} dei numeri complessi e ne discutiamo le principali proprietà. Il campo \mathbb{C} ha la notevole proprietà di essere *algebricamente chiuso*, cioè ogni polinomio su \mathbb{C} ammette radici e questo ha importanti conseguenze anche per i \mathbb{C} -spazi vettoriali.

Si consideri lo spazio vettoriale numerico bidimensionale sul campo \mathbb{R} dei numeri reali: \mathbb{R}^2 . Oltre alla somma $+$, \mathbb{R}^2 può essere dotato di un'altra operazione interna, detta *prodotto* (di numeri complessi), e indicata con \cdot , definita come segue

$$\cdot : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad ((a,b), (a',b')) \mapsto (a,b) \cdot (a',b') := (aa' - bb', ab' + ba'). \quad (4.1)$$

Con la somma usuale e il prodotto \cdot così definito \mathbb{R}^2 è un anello commutativo che si indica con il simbolo \mathbb{C} (la verifica dei dettagli è lasciata al lettore come Esercizio 4.1).

Esercizio 4.1 Dimostrare che, con la somma usuale e il prodotto dato dalla (4.1), lo spazio \mathbb{R}^2 è un anello commutativo in cui lo *zero* (cioè l'elemento neutro rispetto alla somma) è la coppia nulla $(0,0)$. ■

In effetti \mathbb{C} è più che un anello: è un campo, come vedremo adesso. Innanzitutto, il prodotto (4.1) ammette elemento neutro: si tratta della coppia $(1,0)$ (il lettore lo vede?). Inoltre ogni

elemento (a, b) diverso da $(0, 0)$ ammette inverso rispetto al prodotto, cioè esiste un elemento (a', b') , necessariamente unico, tale che $(a, b) \cdot (a', b') = (1, 0)$. Per vedere ciò, basta porre

$$(a', b') = \left(\frac{a}{a^2 + b^2}, -\frac{b}{a^2 + b^2} \right).$$

Tale elemento, che dipende naturalmente da (a, b) , verrà indicato anche con $(a, b)^{-1}$ (il lettore verifichi per esercizio che effettivamente $(a, b) \cdot (a, b)^{-1} = (1, 0)$). Dunque \mathbb{C} è un campo. Gli elementi di \mathbb{C} vengono anche detti *numeri complessi*.

L'applicazione

$$\mathbb{R} \rightarrow \mathbb{C}, \quad a \mapsto (a, 0) \tag{4.2}$$

è un'applicazione iniettiva che conserva sia la somma che il prodotto, oltre che l'elemento neutro rispetto al prodotto, cioè, per ogni $a, b \in \mathbb{R}$

$$\begin{aligned} a + b &\mapsto (a + b, 0) = (a, 0) + (b, 0), \\ ab &\mapsto (ab, 0) = (a, 0) \cdot (b, 0) \\ 1 &\mapsto (1, 0). \end{aligned}$$

Possiamo perciò identificare il campo dei numeri reali \mathbb{R} con la sua immagine mediante la (4.2) in \mathbb{C} e scrivere $\mathbb{R} \subseteq \mathbb{C}$. Per questo si dice anche che *ogni numero reale è anche un numero complesso*.

Introduciamo ora una notazione più standard (e più utile) per i numeri complessi. Il numero complesso (a, b) viene più frequentemente indicato con il simbolo

$$a + ib.$$

Dato un numero complesso $z = a + ib$, il numero reale a viene anche detto la *parte reale* di z e si indica con $\operatorname{Re} z \in \mathbb{R}$. Il numero reale b viene detto la *parte immaginaria* di z e si indica con $\operatorname{Im} z \in \mathbb{R}$. Dato $z = a + ib$, se $b = \operatorname{Im} z = 0$, allora z si indica anche semplicemente con a (invece che con $a + i0$). Per esempio l'elemento neutro rispetto al prodotto è $1 + i0$, ma si indica semplicemente 1 . Si osservi che, così facendo, stiamo usando lo stesso simbolo a per

- (1) un numero reale,
- (2) il numero complesso $(a, 0)$,
- (3) l'immagine di a mediante la (4.2).

Ma questo è consistente perché l'immagine di a mediante la (4.2) è esattamente $(a, 0)$. Se invece $a = \operatorname{Re} z = 0$, allora z si indica anche semplicemente ib (invece che con $0 + ib$). Il numero complesso $i1 = 0 + i1$ si indica anche semplicemente con i e si chiama l'*unità immaginaria*. È facile verificare che $i \cdot i = -1$ (cioè $(0, 1) \cdot (0, 1) = (1, 0)$). Un numero complesso del tipo ib si dice *puramente immaginario*. Queste notazioni sono consistenti, nel senso che, per ogni $a, b \in \mathbb{R}$, il numero complesso $a + ib$ è esattamente la somma del numero complesso $a = (a, 0)$ e del numero complesso $i \cdot b = (0, 1) \cdot (b, 0)$:

$$a + ib = (a, b) = (a, 0) + (0, b) = (a, 0) + (0, 1) \cdot (b, 0).$$

Per questo, lavorare con i numeri complessi è facile: tutte le usuali *regole di calcolo* con somme e prodotti (di elementi in un campo, per esempio il campo dei numeri reali) si applicano ai simboli $a + ib$, con l'accortezza che $i^2 = i \cdot i = -1$. Per esempio, per calcolare il prodotto $(a + ib) \cdot (a' + ib')$, invece di ricordarsi la formula data dalla (4.1), si possono usare tali regole di calcolo. In questo modo si trova

$$\begin{aligned} (a + ib) \cdot (a' + ib') &= a \cdot a' + a \cdot ib' + ib \cdot a' + ib' \cdot ib' \\ &= aa' + i(ab' + ba') + i^2 bb' = (aa' - bb') + i(ab' + ba'), \end{aligned}$$

consistentemente con la (4.1). Da adesso in poi, come è uso fare, ometteremo spesso il puntino “.” nella notazione per il prodotto di numeri complessi. Utilizzeremo anche la notazione $1/z$ per l'inverso z^{-1} del numero complesso z . Si noti che $i^{-1} = -i$ (il lettore lo vede?).

Dato un numero complesso $z := a + ib \in \mathbb{C}$, il numero complesso $\bar{z} = a - ib$ è detto *coniugato* di z . L'applicazione

$$\mathbb{C} \rightarrow \mathbb{C}, \quad z \mapsto \bar{z},$$

è detta *coniugio*.

Proposizione 4.1.1 Il coniugio gode delle seguenti proprietà: per ogni $z, z_1, z_2, z' \in \mathbb{C}$

- (1) $\overline{\bar{z}} = z$;
- (2) $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$;
- (3) $\operatorname{Re} z = (z + \bar{z})/2$;
- (4) $\operatorname{Im} z = (z - \bar{z})/2i$;
- (5) $\overline{zz'} = \bar{z}\bar{z}'$.

In particolare, un numero complesso z è un numero reale (cioè $\operatorname{Im} z = 0$) sse $\bar{z} = z$. Inoltre, se $z \neq 0$, allora

- (6) $\overline{z^{-1}} = \bar{z}^{-1}$.

Esercizio 4.2 Dimostrare la Proposizione 4.1.1. ■

Sia $z = a + ib \in \mathbb{C}$ un numero complesso. Calcoliamo

$$z\bar{z} = (a + ib)(a - ib) = a^2 + b^2 \in \mathbb{R}.$$

Cioè $z\bar{z}$ è sempre un numero reale non negativo, detto il *modulo quadro* di z , e si indica con $|z|^2$. La radice quadrata $\sqrt{|z|^2}$ è detta *modulo* di z e si indica con $|z|$. Si osservi che il modulo non è altro che la norma rispetto al prodotto scalare standard in \mathbb{R}^2 . Si noti anche che, se z è un numero reale, allora il suo modulo coincide con il suo valore assoluto (dunque la notazione “ $|-|$ ” è consistente).

Proposizione 4.1.2 Il modulo gode delle seguenti proprietà: per ogni $z, z_1, z_2 \in \mathbb{C}$

- (1) $|z| = 0$ sse $z = 0$;
- (2) $|z_1 z_2| = |z_1| |z_2|$;
- (3) $|\bar{z}| = |z|$;
- (4) se $z \neq 0$, allora $|z^{-1}| = |z|^{-1}$, e inoltre $z^{-1} = \bar{z}/|z|^2$.

Esercizio 4.3 Dimostrare la Proposizione 4.1.2. ■

Discutiamo ora brevemente di *polinomi* sul campo dei numeri complessi. Vale il seguente

Teorema 4.1.3 — Teorema Fondamentale dell'Algebra. Ogni polinomio di grado positivo sul campo \mathbb{C} dei numeri complessi ammette radici. Perciò ogni polinomio $P(t) \in \mathbb{C}[t]$ di grado $n > 0$ si scrive nella forma

$$P(t) = a(t - \lambda_1)^{m_1} \cdots (t - \lambda_k)^{m_k}, \quad \text{dove } m_1 + \cdots + m_k = n,$$

inoltre $a \in \mathbb{C} \setminus \{0\}$ (un polinomio di grado 0), $\lambda_1, \dots, \lambda_k$ sono tutte le radici distinte di $P(t)$ e m_1, \dots, m_k sono le loro molteplicità.

Ogni polinomio $P(t)$ sul campo dei numeri reali si può riguardare come un polinomio sul campo dei numeri complessi usando l'inclusione $\mathbb{R} \subseteq \mathbb{C}$. Per esempio si può parlare di *radici complesse*

di un polinomio reale. In virtù del Teorema Fondamentale dell'Algebra, ogni polinomio reale ammette radici complesse. Si può anche dire qualcosa in più a riguardo: sia

$$P(t) = a_0 + a_1 t + \dots + a_n t^n \in \mathbb{R}[t]$$

un polinomio reale, $a_0, a_1, \dots, a_n \in \mathbb{R}$. Se $z \in \mathbb{C}$ è una radice complessa di $P(t)$, allora anche \bar{z} è una radice (complessa) di $P(t)$, infatti, in virtù della Proposizione 4.1.1,

$$P(\bar{z}) = \overline{P(z)} = \overline{\sum_{i=1}^n a_i z^i} = \sum_{i=1}^n \overline{a_i z^i} = \sum_{i=1}^n \overline{a_i} \overline{z^i} = \sum_{i=1}^n a_i z^i = P(z) = \bar{0} = 0,$$

dove abbiamo usato che $\overline{\bar{a}_i} = a_i$ per ogni $i = 1, \dots, n$. Ne consegue il

Corollario 4.1.4 Ogni polinomio reale $P(t) \in \mathbb{R}[t]$ di grado $n > 0$ si scrive nella forma

$$P(t) = a(t - \lambda_1)^{m_1} \dots (t - \lambda_k)^{m_p} (t - \mu_1)^{n_1} (t - \bar{\mu}_1)^{n_1} \dots (t - \mu_q)^{n_q} (t - \bar{\mu}_q)^{n_q}$$

dove

$$\sum_{i=1}^p m_i + 2 \sum_{j=1}^q n_j = n,$$

inoltre $a \in \mathbb{R} \setminus \{0\}$, $\lambda_1, \dots, \lambda_k$ sono le radici reali di $P(t)$ e $\mu_1, \bar{\mu}_1, \dots, \mu_q, \bar{\mu}_q$ sono le radici complesse di $P(t)$ con parte immaginaria non nulla.

4.2 Forme Hermitiane e Prodotti Hermitiani

Cominciamo discutendo brevemente di *matrici complesse*, cioè matrici sul campo dei numeri complessi. Innanzitutto, ogni matrice reale si può riguardare come una matrice complessa, usando $\mathbb{R} \subseteq \mathbb{C}$. Inoltre, data una matrice complessa $Z = (z_{ij})_{i=1, \dots, m}^{j=1, \dots, n} \in M_{m,n}(\mathbb{C})$, essa si può sempre scrivere in un unico modo nella forma

$$Z = A + iB$$

con $A, B \in M_{m,n}(\mathbb{R})$ matrici reali. A questo scopo è sufficiente porre $A = (\operatorname{Re} z_{ij})_{i=1, \dots, m}^{j=1, \dots, n}$ e $B = (\operatorname{Im} z_{ij})_{i=1, \dots, m}^{j=1, \dots, n}$. Per questo si scrive anche $A = \operatorname{Re} Z$ e $B = \operatorname{Im} Z$. Sia di nuovo $Z = (z_{ij})_{i=1, \dots, m}^{j=1, \dots, n} \in M_{m,n}(\mathbb{C})$. La matrice $\bar{Z} = (\bar{z}_{ij})_{i=1, \dots, m}^{j=1, \dots, n} \in M_{m,n}(\mathbb{C})$ si dice la matrice *coniugata*. L'applicazione

$$M_{m,n}(\mathbb{C}) \rightarrow M_{m,n}(\mathbb{C}), \quad A \mapsto \bar{A}$$

è detta *coniugio di matrici*.

Proposizione 4.2.1 Il coniugio di matrici gode delle seguenti proprietà (ereditate direttamente dal coniugio di numeri complessi): per ogni $Z, Z_1, Z_2 \in M_{m,n}(\mathbb{C})$

- (0) $\overline{\bar{Z}} = Z$;
- (1) $\overline{\bar{Z}} = Z$;
- (2) $\overline{Z_1 + Z_2} = \bar{Z}_1 + \bar{Z}_2$;
- (3) $\operatorname{Re} Z = \frac{1}{2}(Z + \bar{Z})$;
- (4) $\operatorname{Im} Z = \frac{1}{2i}(Z - \bar{Z})$.

In particolare, una matrice complessa Z è una matrice reale (cioè $\operatorname{Im} Z = 0$) sse $\bar{Z} = Z$. Inoltre se Z, Z' sono matrici moltiplicabili, si ha anche

- (5) $\overline{ZZ'} = \bar{Z}'\bar{Z}$.

e, se Z è una matrice quadrata, allora $\det \bar{Z} = \overline{\det Z}$ e, se Z è invertibile, anche \bar{Z} è invertibile e
 (6) $\bar{Z}^{-1} = \overline{Z^{-1}}$.

Esercizio 4.4 Dimostrare la Proposizione 4.2.1. ■

Data una matrice complessa $Z \in M_{m,n}(\mathbb{C})$, la matrice $\bar{Z}^T \in M_{n,m}(\mathbb{C})$ è detta *matrice hermitiana coniugata* di Z .

Passiamo ora agli spazi vettoriali complessi. Uno *spazio vettoriale complesso* è uno spazio vettoriale sul campo dei numeri complessi. Per gli spazi vettoriali complessi, oltre alla nozione di applicazione lineare è utile la nozione di *applicazione antilineare*. Siano V, W spazi vettoriali complessi.

Definizione 4.2.1 — Applicazione Antilineare. Un'applicazione $f : V \rightarrow W$ è *antilineare* se, per ogni $v, u \in V$ e per ogni $z \in \mathbb{C}$ si ha

- (1) $f(v + u) = f(v) + f(u)$,
- (2) $f(zv) = \bar{z}f(v)$.

■ **Esempio 4.1** Il coniugio di matrici:

$$M_{m,n}(\mathbb{C}) \rightarrow M_{m,n}(\mathbb{C}), \quad Z \mapsto \bar{Z}$$

è un'applicazione antilineare. Infatti, per ogni $Z_1, Z_2 \in M_{m,n}(\mathbb{C})$, come abbiamo già visto nella Proposizione 4.2.1.(2), $\overline{Z_1 + Z_2} = \bar{Z}_1 + \bar{Z}_2$. Inoltre se $Z = (z_{ij})_{i=1, \dots, m}^{j=1, \dots, n} \in M_{m,n}(\mathbb{C})$ e $z \in \mathbb{C}$, allora

$$\overline{zZ} = \overline{z(z_{ij})} = \overline{(z z_{ij})} = (\bar{z} \bar{z}_{ij}) = (\bar{z} \bar{z}_{ij}) = \bar{z}(\bar{z}_{ij}) = \bar{z}\bar{Z}.$$

La prossima definizione introduce l'analogo per uno spazio complesso di una forma bilineare simmetrica per uno spazio reale. Sia V uno spazio vettoriale complesso di dimensione finita.

Definizione 4.2.2 — Forma Hermitiana. Una *forma hermitiana* su V è un'applicazione

$$h : V \times V \rightarrow \mathbb{C},$$

tale che

- (1) h è lineare nel secondo argomento, cioè

$$h(v, b_1 w_1 + b_2 w_2) = b_1 h(v, w_1) + b_2 h(v, w_2),$$

per ogni $v, w_1, w_2 \in V$ e $b_1, b_2 \in \mathbb{C}$;

- (2) h è **antilineare** nel primo argomento, cioè

$$h(a_1 v_1 + a_2 v_2, w) = \bar{a}_1 h(v_1, w) + \bar{a}_2 h(v_2, w),$$

per ogni $v_1, v_2, w \in V$ e $a_1, a_2 \in \mathbb{C}$;

- (3) h gode della seguente proprietà di “*simmetria*”:

$$h(v, w) = \overline{h(w, v)},$$

per ogni $v, w \in V$.

In effetti, la Definizione 4.2.2 è ridondante: la proprietà (2) segue dalla (1) e dalla (3) (oppure la (1) segue dalla (2) e dalla (3)). Infatti supponiamo che $h : V \times V \rightarrow \mathbb{C}$ sia un'applicazione che

soddisfa le proprietà (1) e (3) della Definizione 4.2.2, allora, per ogni $v_1, v_2, w \in V$ e $a_1, a_2 \in \mathbb{C}$ si ha

$$\begin{aligned} h(a_1v_1 + a_2v_2, w) &= \overline{h(w, a_1v_1 + a_2v_2)} = \overline{a_1h(w, v_1) + a_2h(w, v_2)} \\ &= \overline{a_1} \overline{h(w, v_1)} + \overline{a_2} \overline{h(w, v_2)} = \overline{a_1} h(v_1, w) + \overline{a_2} h(v_2, w), \end{aligned}$$

come desiderato.

■ **Esempio 4.2** Sia $V = \mathbb{C}^n$ lo spazio complesso numerico. Consideriamo l'applicazione

$$\langle -, - \rangle_{can} : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}, \quad (x, y) \mapsto \langle x, y \rangle_{can} := \overline{x}^T y.$$

È facile dimostrare che $\langle -, - \rangle_{can}$ è una forma hermitiana (il lettore lo dimostri per esercizio!). Tale forma hermitiana è detta il *prodotto hermitiano standard* e, nel prosieguo, giocherà un ruolo importante. Più in generale sia $H \in M_n(\mathbb{C})$ una matrice quadrata tale che $H = \overline{H}^T$ cioè H coincide con la sua matrice hermitiana coniugata, e consideriamo l'applicazione:

$$h_H : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}, \quad (x, y) \mapsto \langle x, y \rangle_{can} := \overline{x}^T H y.$$

Allora h_H è una forma hermitiana, infatti

(1) h_H è lineare nel secondo argomento: per ogni $x, y_1, y_2 \in \mathbb{C}^n$, e per ogni $a_1, a_2 \in \mathbb{C}$,

$$h_H(x, a_1y_1 + a_2y_2) = \overline{x}^T H(a_1y_1 + a_2y_2) = a_1 \overline{x}^T H y_1 + a_2 \overline{x}^T H y_2 = a_1 h_H(x, y_1) + a_2 h_H(x, y_2).$$

(2) h_H è "simmetrica": per ogni $x, y \in \mathbb{C}^n$,

$$h_H(x, y) = \overline{x}^T H y = y^T \overline{\overline{H}^T \overline{x}} = \overline{y^T \overline{H}^T \overline{x}} = \overline{y^T H x} = \overline{h_H(y, x)}.$$

Si osservi che, quando $H = I_n$, allora h_H è il prodotto hermitiano standard. ■

In modo simile a quanto avviene per le forme bilineari, ogni forma hermitiana è del tipo descritto nell'Esempio 4.2 a meno della scelta di un sistema di coordinate, in base alla seguente

Proposizione / Definizione 4.2.2 — Matrice Rappresentativa di una Forma Hermitiana. Sia $h : V \times V \rightarrow \mathbb{C}$ una forma hermitiana su uno spazio vettoriale complesso V di dimensione n e sia $\mathcal{R} = (v_1, \dots, v_n)$ un riferimento di V . Allora esiste un'unica matrice quadrata complessa $H \in M_n(\mathbb{C})$ tale che, per ogni $v, w \in V$,

$$h(v, w) = \overline{x}^T H y,$$

dove x e y sono i vettori coordinati di v e w nel riferimento \mathcal{R} . La matrice H prende il nome di *matrice rappresentativa* della forma hermitiana h nel riferimento \mathcal{R} e coincide con la sua matrice hermitiana coniugata:

$$H^T = \overline{H}.$$

Dimostrazione. La dimostrazione è simile alla analoga dimostrazione per le forme bilineari. Sia $H = (h_{ij})_{i=1, \dots, n}^{j=1, \dots, n} \in M_n(\mathbb{C})$ la matrice definita da

$$h_{ij} := h(v_i, v_j), \quad i, j = 1, \dots, n.$$

È chiaro che

$$\overline{h_{ji}} = \overline{h(v_j, v_i)} = h(v_i, v_j) = h_{ij}, \quad i, j = 1, \dots, n.$$

Questo mostra che $\overline{H}^T = H$. Ora siano $v, w \in V$ e siano $x = (x_1, \dots, x_n)^T, y = (y_1, \dots, y_n)^T$ i rispettivi vettori coordinati nel riferimento \mathcal{R} . Allora

$$h(v, w) = h\left(\sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j\right) = \sum_{i=1}^n \sum_{j=1}^n \overline{x_i} y_j h(v_i, v_j) = \sum_{i=1}^n \sum_{j=1}^n \overline{x_i} y_j h_{ij}, \quad (4.3)$$

dove abbiamo usato simultaneamente la linearità nel secondo argomento e l'antilinearità nel primo argomento. L'ultimo termine della (4.3) è esattamente $\overline{x}^T H y$ e così è dimostrata l'esistenza. Per l'unicità, sia $H' = (h'_{ij})_{i=1, \dots, n}^{j=1, \dots, n} \in M_n(\mathbb{C})$ un'altra matrice tale che

$$h(v, w) = \overline{x}^T H' y$$

per ogni $v, w \in V$ (dove, come al solito, x, y sono i vettori coordinati di v, w nel riferimento \mathcal{R}). In particolare si ha

$$h(v_i, v_j) = \overline{E_i}^T H' E_j, \quad i, j = 1, \dots, n$$

dove E_i è l' i -esimo vettore del riferimento canonico di \mathbb{C}^n . Poiché

$$E_i = \left(0, \dots, 0, \underbrace{1}_{\text{posto } i}, 0, \dots, 0\right)^T,$$

E_i è, in effetti, un vettore reale, perciò $\overline{E_i} = E_i$ e

$$h_{ij} = h(v_i, v_j) = \overline{E_i}^T H' E_j = E_i^T H' E_j = h'_{ij}, \quad i, j = 1, \dots, n,$$

cioè $H' = H$, il che dimostra l'unicità. ■

■ **Esempio 4.3** Sia $H \in M_n(\mathbb{C})$ una matrice complessa tale che $H = \overline{H}^T$ e sia h_H la forma hermitiana nell'Esempio 4.2. Allora H è anche la matrice rappresentativa di h_H nel riferimento canonico (il lettore lo vede?). ■

Definizione 4.2.3 — Matrice Hermitiana. Una *matrice hermitiana* è una matrice quadrata complessa $H \in M_n(\mathbb{C})$ tale che

$$H = \overline{H}^T.$$

In altre parole H coincide con la sua matrice hermitiana coniugata.

■ **Esempio 4.4** Sia $H = (h_{ij})_{i=1, \dots, n}^{j=1, \dots, n} \in M_n(\mathbb{C})$ una matrice hermitiana. In termini degli elementi di matrice la condizione $\overline{H}^T = H$ si scrive

$$\overline{h_{ji}} = h_{ij}, \quad \text{per ogni } i, j = 1, \dots, n. \quad (4.4)$$

In particolare

$$\overline{h_{ii}} = h_{ii} \Leftrightarrow h_{ii} \in \mathbb{R}, \quad \text{per ogni } i = 1, \dots, n.$$

Cioè gli elementi sulla diagonale di una matrice hermitiana sono necessariamente numeri reali. Inoltre, dalla (4.4), gli elementi sotto la diagonale sono i coniugati degli elementi sopra la diagonale (e viceversa). Per esempio la matrice

$$\begin{pmatrix} 1 & 2i & 1-i \\ -2i & 0 & -2 \\ 1+i & -2 & -1/2 \end{pmatrix}$$

è una matrice hermitiana. Si osservi infine che una matrice reale, rivista come matrice complessa, è hermitiana sse è una matrice simmetrica (il lettore lo vede?). ■

Introduciamo ora l'analogo complesso di un prodotto scalare. Sia V uno spazio vettoriale complesso (di dimensione finita). Osserviamo preliminarmente che, se h è una forma hermitiana in V , allora, per ogni $v \in V$, il numero complesso $h(v, v)$ è, in effetti, un numero reale, infatti, dalla "simmetria" di h , abbiamo $h(v, v) = \overline{h(v, v)}$.

Definizione 4.2.4 — Prodotto Hermitiano. Un *prodotto hermitiano* in V è una forma hermitiana

$$\langle -, - \rangle : V \times V \rightarrow \mathbb{C}, \quad (v, w) \mapsto \langle v, w \rangle$$

definita positiva, cioè $\langle v, v \rangle \geq 0$ per ogni $v \in V$, e $\langle v, v \rangle = 0$ solo se $v = 0$. Uno *spazio vettoriale hermitiano* (o, semplicemente, uno *spazio hermitiano*) è uno spazio vettoriale complesso munito di prodotto hermitiano, cioè una coppia $(V, \langle -, - \rangle)$, in cui V è uno spazio vettoriale complesso (di dimensione finita) e $\langle -, - \rangle$ è un prodotto hermitiano in V .

○ Stiamo indicando con lo stesso simbolo $\langle -, - \rangle$ tanto un prodotto scalare quanto un prodotto hermitiano. Dovrebbe essere sempre chiaro dal contesto (e, in particolare, da quale sia il campo degli scalari: \mathbb{R} oppure \mathbb{C}) se intendiamo l'uno o l'altro.

■ **Esempio 4.5** Sia n un intero positivo. Il prodotto hermitiano standard $\langle -, - \rangle_{can}$ sullo spazio vettoriale numerico \mathbb{C}^n (vedi Esempio 4.2) è, in effetti, un prodotto hermitiano. Si osservi che $\langle -, - \rangle_{can}$ è la forma hermitiana in \mathbb{C}^n la cui matrice rappresentativa nel riferimento canonico è la matrice identica I_n . Lo spazio hermitiano $(\mathbb{C}^n, \langle -, - \rangle_{can})$ è detto *spazio hermitiano numerico* o *spazio hermitiano standard*. ■

Anche i prodotti hermitiani consentono di introdurre alcune costruzioni metriche. Sia $(V, \langle -, - \rangle)$ uno spazio hermitiano.

Definizione 4.2.5 — Norma e Distanza. La *norma quadra* è la funzione

$$V \rightarrow \mathbb{R}, \quad v \mapsto \|v\|^2 := \langle v, v \rangle.$$

La *norma* è la funzione

$$V \rightarrow \mathbb{R}, \quad v \mapsto \|v\| := \sqrt{\|v\|^2} = \sqrt{\langle v, v \rangle}.$$

I vettori di norma unitaria sono detti *versori*.

La *distanza* è la funzione

$$d : V \times V \rightarrow \mathbb{R}, \quad (v, w) \mapsto d(v, w) := \|v - w\|.$$

È chiaro che norma quadra, norma e distanza sono ben definite (il lettore lo vede?). Inoltre la norma di un vettore v è nulla sse $v = 0$. Consideriamo $v \in V$ e $a \in \mathbb{C}$. Allora

$$\|av\|^2 = \langle av, av \rangle = \bar{a}a \langle v, v \rangle = |a|^2 \|v\|^2 \quad \text{da cui anche} \quad \|av\| = |a| \|v\|.$$

Anche negli spazi hermitiani valgono la *disuguaglianza di Cauchy-Schwarz* e la *disuguaglianza triangolare* con qualche piccola necessaria variazione nella dimostrazione che ora discutiamo.

Teorema 4.2.3 — Disuguaglianza di Cauchy-Schwarz. Sia $(V, \langle -, - \rangle)$ uno spazio hermitiano. Allora

$$|\langle v, w \rangle| \leq \|v\| \|w\| \tag{4.5}$$

per ogni $v, w \in V$. Inoltre, nella (4.5), l'uguaglianza vale sse v e w sono linearmente dipendenti.

Dimostrazione. Se $v = 0$ la (4.5) vale ed è un'uguaglianza, in accordo con la seconda parte dell'enunciato. Quindi possiamo assumere $v \neq 0$. Siano $a, b \in \mathbb{C}$. Allora

$$\begin{aligned} 0 \leq \|av + bw\|^2 &= \langle av + bw, av + bw \rangle = \bar{a}a \langle v, v \rangle + \bar{a}b \langle v, w \rangle + \bar{b}a \langle w, v \rangle + \bar{b}b \langle w, w \rangle \\ &= |a|^2 \|v\|^2 + |b|^2 \|w\|^2 + \bar{a}b \langle v, w \rangle + \overline{\bar{a}b \langle v, w \rangle} \\ &= |a|^2 \|v\|^2 + |b|^2 \|w\|^2 + 2\operatorname{Re}(\bar{a}b \langle v, w \rangle). \end{aligned}$$

Quando $a = -\langle v, w \rangle$ e $b = \|v\|^2$ otteniamo

$$\begin{aligned} 0 \leq |\langle v, w \rangle|^2 \|v\|^2 + \|v\|^4 \|w\|^2 - 2\operatorname{Re}(|\langle v, w \rangle|^2 \|v\|^2) \\ = |\langle v, w \rangle|^2 \|v\|^2 + \|v\|^4 \|w\|^2 - 2|\langle v, w \rangle|^2 \|v\|^2 \\ = -|\langle v, w \rangle|^2 \|v\|^2 + \|v\|^4 \|w\|^2. \end{aligned}$$

Per il resto la dimostrazione prosegue e si conclude esattamente come nel caso degli spazi vettoriali euclidei e lasciamo i dettagli al lettore. ■

Corollario 4.2.4 — Disuguaglianza Triangolare. Sia $(V, \langle -, - \rangle)$ uno spazio hermitiano. Per ogni terna di vettori $u, v, w \in V$ si ha che

$$d(v, w) \leq d(v, u) + d(u, w).$$

Dimostrazione. Innanzitutto, per ogni $\tilde{v}, \tilde{u} \in V$,

$$\|\tilde{v} + \tilde{u}\| \leq \|\tilde{v}\| + \|\tilde{u}\|.$$

Infatti

$$\begin{aligned} \|\tilde{v} + \tilde{u}\|^2 &= \langle \tilde{v} + \tilde{u}, \tilde{v} + \tilde{u} \rangle = \|\tilde{v}\|^2 + 2\operatorname{Re} \langle \tilde{v}, \tilde{u} \rangle + \|\tilde{u}\|^2 \\ &\leq \|\tilde{v}\|^2 + 2|\langle \tilde{v}, \tilde{u} \rangle| + \|\tilde{u}\|^2 \\ &\leq \|\tilde{v}\|^2 + 2\|\tilde{v}\|\|\tilde{u}\| + \|\tilde{u}\|^2 = (\|\tilde{v}\| + \|\tilde{u}\|)^2, \end{aligned}$$

dove abbiamo usato, fra l'altro, la disuguaglianza di Cauchy-Schwarz. La dimostrazione si conclude come nel caso degli spazi vettoriali euclidei. ■

○ Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale hermitiano e sia d la relativa distanza. È chiaro che

- (1) $d(v, w) = d(w, v)$ per ogni $v, w \in V$;
- (2) $d(v, w) \geq 0$ per ogni $v, w \in V$ e $d(v, w) = 0$ sse $v = w$.

Queste due proprietà, insieme alla disuguaglianza triangolare, mostrano che (V, d) è uno spazio metrico (vedi Osservazione a pag. 41).

Esiste un'ovvia nozione di *ortogonalità* per i vettori in uno spazio hermitiano $(V, \langle -, - \rangle)$. Due vettori $v, w \in V$ sono *ortogonali*, e si scrive $v \perp w$, se $\langle v, w \rangle = 0$, e un vettore $v \in V$ è *ortogonale ad un sottoinsieme* $S \subseteq V$ se $v \perp w$, per ogni $w \in S$. La nozione di *complemento ortogonale* si introduce esattamente come per gli spazi vettoriali euclidei. Il Lemma 2.2.2 vale invariato anche per gli spazi hermitiani, e la dimostrazione è essenzialmente la stessa (il lettore provi ad individuare le minime necessarie modifiche per esercizio). Anche la *proiezione ortogonale di un vettore su un sottospazio* è definita nello stesso modo e la Proposizione 2.2.3 continua a valere e si dimostra nello

stesso modo. Un *riferimento ortonormale* di uno spazio hermitiano $(V, \langle -, - \rangle)$ di dimensione n è un riferimento $\mathcal{R} = (e_1, \dots, e_n)$ tale che la matrice rappresentativa di $\langle -, - \rangle$ nel riferimento \mathcal{R} è la matrice identica, cioè $\langle e_i, e_j \rangle = \delta_{ij}$ per ogni $i, j = 1, \dots, n$. Per esempio, il riferimento canonico di \mathbb{C}^n è un riferimento ortonormale rispetto al prodotto hermitiano standard. Si osservi che, se \mathcal{R} è un riferimento ortonormale, per ogni coppia di vettori $v, w \in V$, il prodotto hermitiano $\langle v, w \rangle$ coincide con il prodotto hermitiano standard dei rispettivi vettori coordinati $x, y \in \mathbb{C}^n$ nel riferimento \mathcal{R} :

$$\langle v, w \rangle = \bar{x}^T y = \langle x, y \rangle_{can}.$$

La Proposizione 2.2.4 vale anche per gli spazi vettoriali hermitiani $(V, \langle -, - \rangle)$: se $\mathcal{R} = (e_1, \dots, e_n)$ è un riferimento ortonormale di $(V, \langle -, - \rangle)$ e $v \in V$, allora la i -esima componente di v nel riferimento \mathcal{R} è $\langle e_i, v \rangle$. Si presti attenzione, però, al fatto che l'ordine dei fattori nel prodotto hermitiano $\langle e_i, v \rangle$ è importante! La dimostrazione è uguale al caso reale. Il Teorema di ortogonalizzazione di Gram-Schmidt (inclusa la formula (2.11)) vale invariato per gli spazi hermitiani (attenzione però all'ordine dei fattori nei prodotti hermitiani che compaiono nella Formula (2.11)) e anche la dimostrazione è la stessa (fatta eccezione per una minima variazione nella Formula (2.14), in cui, nel caso hermitiano, a è, a priori, un numero complesso e a^2 nell'ultimo termine, va sostituito con $|a|^2$; da $a = |a|^2$ segue di nuovo $a = 1$). In particolare, *in uno spazio hermitiano esistono sempre riferimenti ortogonali*, e quindi ortonormali (a partire da un riferimento ortogonale è sempre possibile costruirne uno ortonormale mediante la solita *normalizzazione*).

■ **Esempio 4.6** In \mathbb{C}^3 con il prodotto hermitiano standard, consideriamo il sistema

$$S = (v_1, v_2, v_3) = \left(\begin{pmatrix} i \\ -i \\ i \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ i \end{pmatrix} \right).$$

Vogliamo applicare ad S la procedura di ortogonalizzazione di Gram-Schmidt. Al primo passo dell'algoritmo abbiamo

$$u_1 = v_1 = \begin{pmatrix} i \\ -i \\ i \end{pmatrix}.$$

Abbiamo anche

$$\langle u_1, u_1 \rangle_{can} = \langle v_1, v_1 \rangle_{can} = 3, \quad \langle u_1, v_2 \rangle_{can} = -2i \quad \text{e} \quad \langle u_1, v_3 \rangle_{can} = 1 + i.$$

Perciò, al secondo passo dell'algoritmo troviamo

$$u_2 = v_2 - \frac{\langle u_1, v_2 \rangle_{can}}{\langle u_1, u_1 \rangle_{can}} u_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \frac{2i}{3} \begin{pmatrix} i \\ -i \\ i \end{pmatrix} = \begin{pmatrix} 1/3 \\ 2/3 \\ 1/3 \end{pmatrix}.$$

Inoltre

$$\langle u_2, u_2 \rangle_{can} = \frac{2}{3} \quad \text{e} \quad \langle u_2, v_3 \rangle_{can} = \frac{2+i}{3},$$

perciò, all'ultimo passo dell'algoritmo di Gram-Schmidt troviamo

$$\begin{aligned} u_3 &= v_3 - \frac{\langle u_1, v_3 \rangle_{can}}{\langle u_1, u_1 \rangle_{can}} u_1 - \frac{\langle u_2, v_3 \rangle_{can}}{\langle u_2, u_2 \rangle_{can}} u_2 = \begin{pmatrix} 0 \\ 1 \\ i \end{pmatrix} - \frac{1+i}{3} \begin{pmatrix} i \\ -i \\ i \end{pmatrix} - \frac{2+i}{2} \begin{pmatrix} 1/3 \\ 2/3 \\ 1/3 \end{pmatrix} \\ &= \begin{pmatrix} -i/2 \\ 0 \\ i/2 \end{pmatrix}. \end{aligned}$$

Si noti che $S' = (u_1, u_2, u_3)$ è un sistema indipendente. È perciò un riferimento ortogonale di $(\mathbb{C}^3, \langle -, - \rangle_{can})$. Possiamo costruire anche un riferimento ortonormale $\mathcal{R} = (e_1, e_2, e_3)$ dividendo ciascun vettore di S' per la sua norma. Poiché

$$\langle u_3, u_3 \rangle = 1/2,$$

otteniamo

$$e_1 = \frac{1}{\|u_1\|} u_1 = \begin{pmatrix} \sqrt{3}i/3 \\ -\sqrt{3}i/3 \\ \sqrt{3}i/3 \end{pmatrix}, \quad e_2 = \frac{1}{\|u_2\|} u_2 = \begin{pmatrix} \sqrt{6}/6 \\ \sqrt{6}/3 \\ \sqrt{6}/6 \end{pmatrix}, \quad \frac{1}{\|u_3\|} u_3 = \begin{pmatrix} -\sqrt{2}i/2 \\ 0 \\ \sqrt{2}i/2 \end{pmatrix}.$$

Dunque il sistema

$$\mathcal{R} = \left(\begin{pmatrix} \sqrt{3}i/3 \\ -\sqrt{3}i/3 \\ \sqrt{3}i/3 \end{pmatrix}, \begin{pmatrix} \sqrt{6}/6 \\ \sqrt{6}/3 \\ \sqrt{6}/6 \end{pmatrix}, \begin{pmatrix} -\sqrt{2}i/2 \\ 0 \\ \sqrt{2}i/2 \end{pmatrix} \right)$$

è un riferimento ortonormale di $(\mathbb{C}^3, \langle -, - \rangle_{can})$. ■

4.3 Endomorfismi Unitari ed Hermitiani

In questa sezione introduciamo gli analoghi degli endomorfismi ortogonali e simmetrici per gli spazi hermitiani (non considereremo invece, l'analogo delle applicazioni ortogonali tra spazi vettoriali euclidei di dimensione diversa, ma invitiamo il lettore ad indovinare la definizione e le prime proprietà per esercizio, incluse le dimostrazioni). Sia $(V, \langle -, - \rangle)$ uno spazio hermitiano di dimensione n .

Definizione 4.3.1 — Endomorfismo Unitario. Un'endomorfismo unitario di $(V, \langle -, - \rangle)$ è un endomorfismo

$$f : V \rightarrow V$$

che conserva il prodotto hermitiano, cioè tale che, per ogni $v, w \in V$,

$$\langle f(v), f(w) \rangle = \langle v, w \rangle.$$

Proposizione / Definizione 4.3.1 — Gruppo Unitario di uno Spazio Hermitiano. Gli endomorfismi unitari di uno spazio hermitiano $(V, \langle -, - \rangle)$ formano un sottogruppo nel gruppo lineare generale $GL(V)$ degli automorfismi di V , detto il *gruppo unitario* di $(V, \langle -, - \rangle)$ ed indicato con $U(V, \langle -, - \rangle)$.

Dimostrazione. Innanzitutto ogni endomorfismo unitario è iniettivo. La dimostrazione è identica alla analoga nel caso delle applicazioni ortogonali. Dunque ogni endomorfismo unitario è un automorfismo. In altre parole gli endomorfismi unitari formano un sottoinsieme di $GL(V)$. Che l'endomorfismo identico, la composta di due endomorfismi unitari, e l'inversa di un endomorfismo unitario sono endomorfismi unitari è molto facile da verificare e viene lasciato come Esercizio 4.5. ■

Esercizio 4.5 Completare la dimostrazione della Proposizione 4.3.1 dimostrando che l'endomorfismo identico, la composta di due endomorfismi unitari, e l'inversa di un endomorfismo unitario sono endomorfismi unitari. ■

Come vedremo a breve, la matrice rappresentativa di un endomorfismo unitario è una *matrice unitaria*. Le matrici unitarie sono l'analogo complesso delle matrici ortogonali in base alla seguente

Definizione 4.3.2 — Matrice Unitaria. Una *matrice unitaria* di ordine n è una matrice quadrata $A \in M_n(\mathbb{C})$ tale che $\overline{A}^T A = I_n$.

Esercizio 4.6 Dimostrare che la matrice del cambio di coordinate tra due riferimenti ortonormali di uno spazio hermitiano è una matrice unitaria e viceversa. Dato un riferimento ortonormale \mathcal{R}' e un altro riferimento \mathcal{R} , se la matrice del cambio di coordinate $M_{\mathcal{R}', \mathcal{R}}$ è unitaria, allora anche il riferimento \mathcal{R} è ortonormale. ■

Proposizione 4.3.2 Sia $A \in M_n(\mathbb{C})$ una matrice complessa quadrata di ordine n . Allora le seguenti asserzioni sono equivalenti:

- (1) A è una matrice unitaria,
- (2) le colonne di A formano un riferimento ortonormale nello spazio hermitiano standard $(\mathbb{C}^n, \langle -, - \rangle_{can})$,
- (3) le colonne hermitiane coniugate delle righe di A formano un riferimento ortonormale nello spazio hermitiano standard $(\mathbb{C}^n, \langle -, - \rangle_{can})$.

Dimostrazione. La dimostrazione è molto simile a quella della Proposizione 2.3.2 ed è lasciata al lettore come Esercizio 4.7. ■

Esercizio 4.7 Dimostrare la Proposizione 4.3.2. ■

■ **Esempio 4.7** Ogni matrice ortogonale è una matrice reale e può, perciò, essere rivista come una matrice complessa. In quanto matrice complessa, chiaramente, è anche una matrice unitaria. Anzi, una matrice reale è unitaria sse è una matrice ortogonale. Esistono, però, matrici unitarie che non sono ortogonali. Per esempio la matrice

$$\begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}$$

è una matrice unitaria (il lettore lo vede?), ma, non essendo reale, non è una matrice ortogonale. ■

Teorema / Definizione 4.3.3 — Gruppo Unitario. Le matrici unitarie di ordine n costituiscono un sottogruppo del gruppo lineare generale $GL_n(\mathbb{C})$, detto *gruppo unitario di ordine n* che si indica con U_n .

Dimostrazione. Chiaramente la matrice identica è unitaria. Inoltre ogni matrice unitaria A è invertibile e la sua inversa è \overline{A}^T , la matrice hermitiana coniugata, sicché anche A^{-1} è unitaria (le colonne di \overline{A}^T sono le colonne hermitiane coniugate delle righe di A). Infine, siano A, B matrici unitarie (di ordine n). Allora

$$(\overline{AB})^T AB = (\overline{A} \overline{B})^T AB = \overline{B}^T \overline{A}^T AB = I_n.$$

■

Esercizio 4.8 Dimostrare che il determinante di una matrice unitaria ha modulo pari ad 1. ■

Torniamo ora agli endomorfismi unitari

Proposizione 4.3.4 Sia $(V, \langle -, - \rangle)$ uno spazio hermitiano di dimensione n e sia $f : V \rightarrow V$ un endomorfismo. Le seguenti condizioni sono equivalenti:

- (1) f è un endomorfismo unitario;
- (2) f conserva la norma;

- (3) f conserva la distanza;
- (4) per ogni riferimento ortonormale \mathcal{R} di $(V, \langle -, - \rangle)$, anche il sistema $f(\mathcal{R})$ è un riferimento ortonormale;
- (5) esiste un riferimento ortonormale \mathcal{R} di $(V, \langle -, - \rangle)$, tale che il sistema $f(\mathcal{R})$ è un riferimento ortonormale;
- (6) per ogni riferimento ortonormale \mathcal{R} di $(V, \langle -, - \rangle)$, la matrice rappresentativa $M_{\mathcal{R}}^f$ è una matrice unitaria;
- (7) esiste un riferimento ortonormale \mathcal{R} di $(V, \langle -, - \rangle)$ tale che la matrice rappresentativa $M_{\mathcal{R}}^f$ è una matrice unitaria.

Dimostrazione. Già sappiamo che (1) \Rightarrow (2), (3). Dimostriamo che (2) \Rightarrow (1). Per ogni $v, w \in V$ abbiamo

$$\|v+w\|^2 = \langle v+w, v+w \rangle = \|v\|^2 + 2\operatorname{Re}\langle v, w \rangle + \|w\|^2.$$

Da cui

$$\operatorname{Re}\langle v, w \rangle = \frac{\|v+w\|^2 - \|v\|^2 - \|w\|^2}{2}.$$

Perciò se f è un'applicazione lineare che conserva la norma, si ha

$$\begin{aligned} \operatorname{Re}\langle f(v), f(w) \rangle &= \frac{\|f(v)+f(w)\|^2 - \|f(v)\|^2 - \|f(w)\|^2}{2} \\ &= \frac{\|f(v+w)\|^2 - \|f(v)\|^2 - \|f(w)\|^2}{2} \\ &= \frac{\|v+w\|^2 - \|v\|^2 - \|w\|^2}{2} \\ &= \operatorname{Re}\langle v, w \rangle. \end{aligned}$$

Ora osserviamo che, per ogni numero complesso $z = a + ib \in \mathbb{C}$, con $a, b \in \mathbb{R}$,

$$\operatorname{Im}z = b = -\operatorname{Re}(-b + ia) = -\operatorname{Re}(iz).$$

Perciò

$$\begin{aligned} \operatorname{Im}\langle f(v), f(w) \rangle &= -\operatorname{Re}(i\langle f(v), f(w) \rangle) = -\operatorname{Re}\langle f(v), if(w) \rangle \\ &= -\operatorname{Re}\langle f(v), f(iw) \rangle = -\operatorname{Re}\langle v, iw \rangle = -\operatorname{Re}(i\langle v, w \rangle) \\ &= \operatorname{Im}\langle v, w \rangle. \end{aligned}$$

Dunque f conserva sia la parte reale che la parte immaginaria del prodotto hermitiano e perciò conserva il prodotto hermitiano stesso, infatti

$$\langle f(v), f(w) \rangle = \operatorname{Re}\langle f(v), f(w) \rangle + i \cdot \operatorname{Im}\langle f(v), f(w) \rangle = \operatorname{Re}\langle v, w \rangle + i \cdot \operatorname{Im}\langle v, w \rangle = \langle v, w \rangle.$$

Il resto è come nella dimostrazione della Proposizione 2.4.3 (con qualche minima modifica nei conti finali che il lettore è invitato ad individuare). ■

Corollario 4.3.5 Sia $(V, \langle -, - \rangle)$ uno spazio hermitiano di dimensione n , e sia \mathcal{R} un riferimento ortonormale di $(V, \langle -, - \rangle)$. L'applicazione $U(V, \langle -, - \rangle) \rightarrow U_n$, $f \mapsto M_{\mathcal{R}}^f$ è un ben definito isomorfismo di gruppi.

■ **Esempio 4.8** Nel caso in cui $(V, \langle -, - \rangle) = (\mathbb{C}^n, \langle -, - \rangle_{can})$ è lo spazio hermitiano standard, possiamo scegliere il riferimento canonico \mathcal{R}_{can} . In questo caso il Corollario 4.3.5 dice che gli endomorfismi unitari di $(\mathbb{C}^n, \langle -, - \rangle_{can})$ sono esattamente le applicazioni lineari $L_A : \mathbb{C}^n \rightarrow \mathbb{C}^n$ associate alle matrici unitarie $A \in U_n$. Esiste dunque un isomorfismo naturale di gruppi $U(\mathbb{C}^n, \langle -, - \rangle_{can}) \cong U_n$ il cui isomorfismo inverso è dato da $A \mapsto L_A$. ■

Esercizio 4.9 Dimostrare il Corollario 4.3.5. ■

Concludiamo questa breve presentazione degli endomorfismi unitari discutendone la diagonalizzabilità, a partire dagli autovalori.

Lemma 4.3.6 Sia $(V, \langle -, - \rangle)$ uno spazio hermitiano di dimensione n , sia $f : (V, \langle -, - \rangle) \rightarrow (V, \langle -, - \rangle)$ un endomorfismo unitario, e sia $\lambda \in \mathbb{C}$ un autovalore di f . Allora $|\lambda| = 1$. Inoltre autovettori associati ad autovalori distinti sono ortogonali.

Dimostrazione. Per la prima parte dell'enunciato, sia v un autovettore relativo all'autovalore λ . Allora

$$\langle v, v \rangle = \langle f(v), f(v) \rangle = \langle \lambda v, \lambda v \rangle = |\lambda|^2 \langle v, v \rangle. \quad (4.6)$$

Ma $v \neq 0$ per definizione di autovettore, perciò $\langle v, v \rangle = \|v\|^2 \neq 0$ e dunque, dalla (4.6), dividendo tutto per $\|v\|^2$, si ottiene $|\lambda|^2 = 1$, come desiderato.

Ora sia μ un altro autovalore distinto da λ , e sia w un autovettore relativo all'autovalore μ . Calcoliamo

$$\langle v, w \rangle = \langle f(v), f(w) \rangle = \langle \lambda v, \mu w \rangle = \bar{\lambda} \mu \langle v, w \rangle.$$

Moltiplicando tutto per λ e usando che $\lambda \bar{\lambda} = |\lambda|^2 = 1$ otteniamo

$$\lambda \langle v, w \rangle = \mu \langle v, w \rangle \quad \Rightarrow \quad (\lambda - \mu) \langle v, w \rangle.$$

Poiché $\lambda \neq \mu$, deve essere $\langle v, w \rangle = 0$, come desiderato. ■

Il Lemma 4.3.6 spiega la terminologia “endomorfismo unitario”.

Proposizione 4.3.7 Sia $(V, \langle -, - \rangle)$ uno spazio hermitiano di dimensione n , e sia

$$f : (V, \langle -, - \rangle) \rightarrow (V, \langle -, - \rangle)$$

un endomorfismo unitario. Allora f è *ortogonalmente diagonalizzabile*, cioè esiste un riferimento ortonormale $\mathcal{R} = (e_1, \dots, e_n)$ di $(V, \langle -, - \rangle)$ di autovettori di f .

Dimostrazione. La dimostrazione è simile a quella del Teorema Spettrale e procede per induzione sulla dimensione n di V . Se $n = 1$, è chiaro che ogni endomorfismo di V è ortogonalmente diagonalizzabile, e quindi non c'è nulla da dimostrare. Supponiamo ora che l'enunciato sia vero quando $n = k$ e dimostriamo che l'enunciato è vero anche quando $n = k + 1$. Sia dunque $(V, \langle -, - \rangle)$ uno spazio hermitiano di dimensione $k + 1$ e sia $f : (V, \langle -, - \rangle) \rightarrow (V, \langle -, - \rangle)$ un endomorfismo unitario. Innanzitutto f (come ogni endomorfismo su uno spazio vettoriale complesso!) ammette autovalori. Infatti il suo polinomio caratteristico $P_f(t)$ è un polinomio complesso e, per il Teorema Fondamentale dell'Algebra, ammette radici. Sia dunque λ un autovalore di f e sia e_0 un autovettore relativo all'autovalore λ . Dopo aver normalizzato (se necessario), possiamo assumere che e_0 sia un versore. Consideriamo il complemento ortogonale $W := e_0^\perp$. Il sottospazio W è f -invariante, cioè per ogni $w \in W$, anche $f(w) \in W$. Per vederlo calcoliamo

$$\lambda \langle f(w), e_0 \rangle = \langle f(w), \lambda e_0 \rangle = \langle f(w), f(e_0) \rangle = \langle w, e_0 \rangle = 0,$$

dove abbiamo usato che f è un endomorfismo unitario e, nell'ultimo passaggio, che $w \perp e_0$. Dunque $\lambda \langle f(w), e_0 \rangle = 0$. Ma $|\lambda| = 1$, perciò $\lambda \neq 0$ e (moltiplicando per λ^{-1}) $\langle f(w), e_0 \rangle = 0$, cioè $f(w) \perp e_0$ come annunciato. Possiamo perciò restringere l'endomorfismo f a W simultaneamente nel dominio e nel codominio ottenendo un'applicazione

$$f_W : W \rightarrow W, \quad w \mapsto f_W(w) := f(w).$$

A questo punto la dimostrazione prosegue esattamente come nel caso del Teorema Spettrale e i dettagli sono lasciati per esercizio al lettore (il lettore presti particolare attenzione alla necessità di restringere il prodotto hermitiano al sottospazio W e al fatto che tale restrizione è ancora un prodotto hermitiano). ■

■ **Esempio 4.9** Consideriamo l'endomorfismo

$$f: \mathbb{C}^3 \rightarrow \mathbb{C}^3, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} ix_1 & & \\ & & -x_3 \\ & x_2 & \end{pmatrix}.$$

La matrice rappresentativa di f nel riferimento canonico è

$$A = \begin{pmatrix} i & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix},$$

che è una matrice unitaria (il lettore lo vede?). Poiché il riferimento canonico è ortonormale rispetto al prodotto hermitiano standard, f è un endomorfismo unitario dello spazio hermitiano standard $(\mathbb{C}^3, \langle -, - \rangle_{can})$. Vogliamo determinare un riferimento ortonormale di \mathbb{C}^3 di autovettori di f .

Il polinomio caratteristico di f è

$$P_f(t) = \det A_t = \det \begin{pmatrix} i-t & 0 & 0 \\ 0 & -t & -1 \\ 0 & 1 & -t \end{pmatrix} = -(t+i)(t-i)^2.$$

Dunque f ha 2 autovalori, $-i$ e i , con molteplicità algebriche 1 e 2 rispettivamente. Poiché f è ortogonalmente diagonalizzabile, in particolare è diagonalizzabile, conosciamo già anche le molteplicità geometriche dei due autovalori: esse coincidono con le molteplicità algebriche.

Determiniamo ora un riferimento in ciascun autospazio. L'autospazio V_{-i} è lo spazio delle soluzioni del sistema $A_{-i}x = 0$. Delle 3 equazioni del sistema $A_{-i}x = 0$, solo 2 sono linearmente indipendenti, per esempio le prime 2 che sono:

$$\begin{cases} 2ix_1 & = 0 \\ ix_2 - x_3 & = 0 \end{cases},$$

Da cui $x_1 = 0$ e $x_3 = ix_2$. Ponendo $x_2 = 1$, vediamo che un riferimento di V_{-i} è

$$\tilde{\mathcal{R}}_{-i} = \left(\begin{pmatrix} 0 \\ 1 \\ i \end{pmatrix} \right). \quad (4.7)$$

Per ottenere un riferimento ortonormale dobbiamo normalizzare. La norma dell'unico vettore di $\tilde{\mathcal{R}}_{-i}$ è

$$\sqrt{1^2 - i \cdot i} = \sqrt{2}.$$

Un riferimento ortonormale si ottiene dividendo per tale norma ed è dunque

$$\mathcal{R}_{-i} = \left(\begin{pmatrix} 0 \\ \sqrt{2}/2 \\ \sqrt{2}i/2 \end{pmatrix} \right).$$

L'autospazio V_i è lo spazio delle soluzioni del sistema $A_i x = 0$. Delle 3 equazioni di tale sistema solo una è indipendente, per esempio la seconda che è:

$$-ix_2 - x_3 = 0,$$

da cui $x_3 = -ix_2$ e un riferimento per V_i è

$$\tilde{\mathcal{R}}_i = \left(\left(\begin{array}{c} 1 \\ 0 \\ 0 \end{array} \right), \left(\begin{array}{c} 0 \\ 1 \\ -i \end{array} \right) \right).$$

I vettori di $\tilde{\mathcal{R}}_i$ sono già ortogonali. Per trovare un riferimento ortonormale di V_i è sufficiente normalizzare. Detti (v_1, v_2) i due vettori di $\tilde{\mathcal{R}}_i$ (nel loro ordine), un riferimento ortonormale per V_i è

$$\mathcal{R}_1 = \left(\frac{1}{\|v_1\|} v_1, \frac{1}{\|v_2\|} v_2 \right) = \left(\left(\begin{array}{c} 1 \\ 0 \\ 0 \end{array} \right), \left(\begin{array}{c} 0 \\ \sqrt{2}/2 \\ -\sqrt{2}i/2 \end{array} \right) \right)$$

e, in virtù (dell'ultima parte) del Lemma 4.3.6, un riferimento ortonormale di $(\mathbb{C}^3, \langle -, - \rangle_{can})$ di autovettori di f è

$$\left(\left(\begin{array}{c} 0 \\ \sqrt{2}/2 \\ \sqrt{2}i/2 \end{array} \right), \left(\begin{array}{c} 1 \\ 0 \\ 0 \end{array} \right), \left(\begin{array}{c} 0 \\ \sqrt{2}/2 \\ -\sqrt{2}i/2 \end{array} \right) \right).$$

■

Corollario 4.3.8 Sia $A \in U_n$ una matrice unitaria (per esempio una matrice ortogonale) di ordine n . Allora esiste un'altra matrice unitaria $U \in U_n$ tale che la matrice $U^{-1}AU = \bar{U}^T AU$ è una matrice diagonale.

Dimostrazione. Vedi l'Esercizio 4.10. ■

Esercizio 4.10 Dimostrare il Corollario 4.3.8 (*Suggerimento: si osservi che l'endomorfismo $L_A : \mathbb{C}^n \rightarrow \mathbb{C}^n$ associato ad A è unitario rispetto al prodotto hermitiano standard. In particolare, dalla Proposizione 4.3.7, L_A è ortogonalmente diagonalizzabile. Ora si ragioni come nella dimostrazione della Proposizione 3.4.1 (si veda anche la Proposizione 3.5.6).*) ■

Esercizio 4.11 Si consideri la matrice

$$A = \left(\begin{array}{ccc} \sqrt{2}/2 & 0 & -\sqrt{2}/2 \\ 0 & -1 & 0 \\ \sqrt{2}/2 & 0 & \sqrt{2}/2 \end{array} \right) \in M_3(\mathbb{C}).$$

Dopo aver dimostrato che si tratta di una matrice unitaria, si determini un'altra matrice unitaria $U \in U_3$ tale che $U^{-1}AU = \bar{U}^T AU$ è una matrice diagonale. Si calcoli $U^{-1}AU$. ■

Anche un prodotto hermitiano definisce una coppia di isomorfismi musicali, ma questi ultimi hanno proprietà leggermente diverse rispetto al loro analogo per i prodotti scalari (per esempio si tratta di applicazioni antilineari, invece che lineari). Per evitare questi tecnicismi, introdurremo gli analoghi degli endomorfismi simmetrici senza fare riferimento agli isomorfismi musicali. Cominciamo da una versione hermitiana del Teorema / Definizione 2.5.2.

Teorema / Definizione 4.3.9 — Endomorfismo Hermitiano Coniugato. Sia $(V, \langle -, - \rangle)$ uno spazio hermitiano di dimensione n . Per ogni endomorfismo $f : V \rightarrow V$, esiste un unico endomorfismo $f^\dagger : V \rightarrow V$, detto *endomorfismo aggiunto*, o *endomorfismo hermitiano coniugato*, tale che

$$\langle f^\dagger(w), v \rangle = \langle w, f(v) \rangle, \quad \text{per ogni } v, w \in V. \quad (4.8)$$

Se $\mathcal{R} = (e_1, \dots, e_n)$ è un riferimento ortonormale di $(V, \langle -, - \rangle)$ e $A = M_{\mathcal{R}}^f$ è la matrice rappresentativa di f nel riferimento \mathcal{R} , allora la matrice hermitiana coniugata \overline{A}^T è la matrice rappresentativa di f^\dagger nello stesso riferimento:

$$\overline{A}^T = M_{\mathcal{R}}^{f^\dagger}.$$

Dimostrazione. Avendo scelto di non introdurre gli isomorfismi musicali determinati da un prodotto hermitiano, non possiamo usare una formula del tipo (2.20) per definire f^\dagger . Perciò, per dimostrare l'esistenza, dobbiamo adottare un'altra strategia. Più precisamente, fissiamo fin dal principio un riferimento ortonormale $\mathcal{R} = (e_1, \dots, e_n)$ di $(V, \langle -, - \rangle)$. Sia $A = M_{\mathcal{R}}^f$ la matrice rappresentativa di f nel riferimento \mathcal{R} . Esiste un unico endomorfismo $f^\dagger : V \rightarrow V$ la cui matrice rappresentativa nel riferimento \mathcal{R} è \overline{A}^T . Al vettore w di coordinate $y = (y_1, \dots, y_n)^T$, l'endomorfismo f^\dagger associa il vettore $f^\dagger(w)$ di coordinate $\overline{A}^T y$. Verifichiamo che l'endomorfismo f^\dagger così definito verifica la proprietà (4.8). Dunque, sia $v \in V$ un altro vettore e sia x il vettore coordinato di v nel riferimento \mathcal{R} , sicché il vettore coordinato di $f(v)$ è Ax . Poiché \mathcal{R} è un riferimento ortonormale, si ha

$$\begin{aligned} \langle f^\dagger(w), v \rangle &= \langle \overline{A}^T y, x \rangle_{can} = (\overline{\overline{A}^T y})^T x = (\overline{A}^T \overline{y})^T x \\ &= \overline{y}^T \overline{A}^T x = \overline{y}^T Ax = \langle y, Ax \rangle_{can} = \langle w, f(v) \rangle, \end{aligned}$$

come desiderato. L'unicità si dimostra come nella dimostrazione del Teorema 2.5.2 e l'ultima parte dell'enunciato è vera per costruzione. La dimostrazione è conclusa. ■

○ Nella dimostrazione del Teorema 4.3.9 abbiamo definito l'endomorfismo hermitiano coniugato ricorrendo ad un riferimento ortonormale \mathcal{R} . A priori questa definizione potrebbe dipendere dalla scelta di \mathcal{R} . Tuttavia l'unicità garantisce l'indipendenza da \mathcal{R} . Infatti, se anche definiamo l'endomorfismo hermitiano coniugato utilizzando un altro riferimento ortonormale \mathcal{R}' , in base alla dimostrazione, il nuovo f^\dagger soddisfa la stessa proprietà (4.8) e, per l'unicità, deve necessariamente coincidere con il vecchio.

Sia $(V, \langle -, - \rangle)$ uno spazio hermitiano di dimensione n .

Definizione 4.3.3 — Endomorfismo Hermitiano. Un *endomorfismo hermitiano*, o *autoaggiunto*, di $(V, \langle -, - \rangle)$ è un endomorfismo $f : V \rightarrow V$ tale che $f = f^\dagger$. In altre parole, f soddisfa la seguente identità: per ogni $v, w \in V$

$$\langle f(v), w \rangle = \langle v, f(w) \rangle.$$

Proposizione 4.3.10 Sia $(V, \langle -, - \rangle)$ uno spazio hermitiano di dimensione n , e sia $f : V \rightarrow V$ un endomorfismo. Le seguenti condizioni sono equivalenti:

- (1) f è un endomorfismo hermitiano;
- (2) per ogni riferimento ortonormale \mathcal{R} di $(V, \langle -, - \rangle)$, la matrice rappresentativa $M_{\mathcal{R}}^f$ è una matrice hermitiana;

(3) esiste un riferimento ortonormale \mathcal{R} di $(V, \langle -, - \rangle)$ tale che la matrice rappresentativa $M_{\mathcal{R}}^f$ è una matrice hermitiana.

Dimostrazione. La dimostrazione è molto simile a quella della Proposizione 2.5.3 e non la ripeteremo. Il lettore, tuttavia, è invitato a verificare tutti i dettagli. ■

Concludiamo questa sezione discutendo la diagonalizzabilità degli endomorfismi hermitiani. Cominciamo con un

Lemma 4.3.11 Sia $(V, \langle -, - \rangle)$ uno spazio hermitiano di dimensione n e sia $f : (V, \langle -, - \rangle) \rightarrow (V, \langle -, - \rangle)$ un endomorfismo hermitiano. Gli autovalori di f sono tutti reali, inoltre autovettori associati ad autovalori distinti sono ortogonali.

Dimostrazione. Osserviamo innanzitutto che, per il Teorema Fondamentale dell'Algebra, f ammette sempre autovalori. Ora sia λ un autovalore di f e sia v un autovettore relativo all'autovalore λ . Calcoliamo

$$\bar{\lambda}\|v\|^2 = \bar{\lambda}\langle v, v \rangle = \langle \lambda v, v \rangle = \langle f(v), v \rangle = \langle v, f(v) \rangle = \langle v, \lambda v \rangle = \lambda\langle v, v \rangle = \lambda\|v\|^2.$$

Poiché $v \neq 0$, si ha $\|v\| \neq 0$ e, dividendo tutto per $\|v\|^2$, troviamo $\bar{\lambda} = \lambda$, cioè $\lambda \in \mathbb{R}$, come desiderato. La seconda parte dell'enunciato si dimostra esattamente come il Lemma 3.5.3, usando che gli autovalori di f sono tutti reali. ■

Corollario 4.3.12 Sia $A \in M_n(\mathbb{R})$ una matrice simmetrica. Allora A ammette almeno un autovalore. Se $\lambda_1, \dots, \lambda_k$ sono tutti gli autovalori distinti di A , allora la somma delle molteplicità algebriche $m_a(\lambda_1) + \dots + m_a(\lambda_k)$ è pari ad n .

Dimostrazione. Possiamo riguardare A come matrice complessa $A \in M_n(\mathbb{C})$. In tal caso, si tratta di una matrice hermitiana infatti $\bar{A}^T = A^T = A$. Allora l'endomorfismo associato $L_A : \mathbb{C}^n \rightarrow \mathbb{C}^n$ è hermitiano rispetto al prodotto hermitiano standard (infatti A è proprio la matrice rappresentativa di L_A nel riferimento canonico che è un riferimento ortonormale rispetto al prodotto hermitiano standard). Il polinomio caratteristico di A coincide con il polinomio caratteristico di L_A e, per il Teorema Fondamentale dell'Algebra, si scrive nella forma

$$P_A = (-1)^n(t - \lambda_1)^{m_1} \dots (t - \lambda_k)^{m_k}, \quad (4.9)$$

dove $\lambda_1, \dots, \lambda_k$ sono tutti gli autovalori distinti di A (o, equivalentemente, di L_A) e m_1, \dots, m_k sono le rispettive molteplicità algebriche. Ma per il Lemma 4.3.11, $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ cosicché la (4.9) è anche una fattorizzazione completa di P_A nell'anello dei polinomi reali $\mathbb{R}[t]$. Questo conclude la dimostrazione. ■

Possiamo finalmente dimostrare il Lemma 3.5.2, che enunciamo nuovamente sotto per la convenienza del lettore. In questo modo concludiamo anche la dimostrazione del Teorema Spettrale.

Lemma 4.3.13 Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo di dimensione n e sia $f : V \rightarrow V$ un endomorfismo simmetrico. Allora f ammette almeno un autovalore. Se $\lambda_1, \dots, \lambda_k$ sono tutti gli autovalori distinti di f , allora la somma delle molteplicità algebriche $m_a(\lambda_1) + \dots + m_a(\lambda_k)$ è pari ad n .

Dimostrazione. Sia \mathcal{R} un riferimento ortonormale di $(V, \langle -, - \rangle)$ e sia A la matrice rappresentativa di f nel riferimento \mathcal{R} . Allora A è una matrice simmetrica. Inoltre gli autovalori di f e le relative molteplicità algebriche coincidono con gli autovalori di A e le relative molteplicità algebriche. Adesso è sufficiente applicare il Corollario 4.3.12. ■

Proposizione 4.3.14 Sia $(V, \langle -, - \rangle)$ uno spazio hermitiano di dimensione n , e sia

$$f : (V, \langle -, - \rangle) \rightarrow (V, \langle -, - \rangle)$$

un endomorfismo hermitiano. Allora f è ortogonalmente diagonalizzabile, cioè esiste un riferimento ortonormale \mathcal{R} di $(V, \langle -, - \rangle)$ di autovettori di f .

Dimostrazione. La dimostrazione è identica alla dimostrazione della seconda parte del Teorema Spettrale e la omettiamo. ■

■ **Esempio 4.10** Consideriamo l'endomorfismo

$$f : \mathbb{C}^3 \rightarrow \mathbb{C}^3, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} \frac{x_1}{4} & & + \frac{ix_3}{4} \\ & \frac{x_2}{2} & \\ -\frac{ix_1}{4} & & + \frac{x_3}{4} \end{pmatrix}.$$

La matrice rappresentativa di f nel riferimento canonico è

$$A = \begin{pmatrix} 1/4 & 0 & i/4 \\ 0 & 1/2 & 0 \\ -i/4 & 0 & 1/4 \end{pmatrix}.$$

Si tratta di una matrice hermitiana, perciò f è un endomorfismo hermitiano di $(\mathbb{C}^3, \langle -, - \rangle_{can})$. Il polinomio caratteristico di f è

$$P_f(t) = \det A_t = \det \begin{pmatrix} 1/4 - t & 0 & i/4 \\ 0 & 1/2 - t & 0 \\ -i/4 & 0 & 1/4 - t \end{pmatrix} = -t \left(t - \frac{1}{2} \right)^2$$

Dunque f ha due autovalori, 0 e 1/2, con molteplicità algebriche, e quindi geometriche, pari a 1 e 2 rispettivamente. L'autospazio V_0 è lo spazio delle soluzioni del sistema $A_0 x = 0$. Solo due equazioni del sistema sono indipendenti, per esempio le prime due che sono

$$\begin{cases} \frac{1}{4}x_1 & + \frac{i}{4}x_3 = 0 \\ \frac{1}{2}x_2 & = 0 \end{cases}.$$

Dalla seconda equazione $x_2 = 0$ e dalla prima $x_1 = -ix_3$. Perciò V_0 è generato dal vettore

$$\begin{pmatrix} -i \\ 0 \\ 1 \end{pmatrix}.$$

Normalizzando, troviamo un riferimento ortonormale per V_0 :

$$\mathcal{R}_0 = \left(\begin{pmatrix} -\sqrt{2}i/2 \\ 0 \\ \sqrt{2}/2 \end{pmatrix} \right).$$

L'autospazio $V_{1/2}$ è lo spazio delle soluzioni del sistema $A_{1/2} x = 0$. Solo una equazione del sistema è indipendente, per esempio la prima che è

$$-\frac{1}{4}x_1 + \frac{i}{4}x_3 = 0,$$

da cui $x_1 = ix_3$. Perciò $V_{1/2}$ è generato dal sistema

$$S = \left(\left(\begin{array}{c} 1 \\ 0 \\ i \end{array} \right), \left(\begin{array}{c} 0 \\ 1 \\ 0 \end{array} \right) \right).$$

Il sistema S è già ortogonale. Per trovare un riferimento ortonormale per $V_{1/2}$ è sufficiente normalizzare. Dunque un riferimento ortonormale di $V_{1/2}$ è

$$\mathcal{R}_{1/2} = \left(\left(\begin{array}{c} \sqrt{2}/2 \\ 0 \\ \sqrt{2}i/2 \end{array} \right), \left(\begin{array}{c} 0 \\ 1 \\ 0 \end{array} \right) \right),$$

e un riferimento ortonormale di autovettori di f è

$$\mathcal{R} = \left(\left(\begin{array}{c} -\sqrt{2}i/2 \\ 0 \\ \sqrt{2}/2 \end{array} \right), \left(\begin{array}{c} \sqrt{2}/2 \\ 0 \\ \sqrt{2}i/2 \end{array} \right), \left(\begin{array}{c} 0 \\ 1 \\ 0 \end{array} \right) \right).$$

■

Corollario 4.3.15 Sia $A \in M_n(\mathbb{C})$ una matrice hermitiana di ordine n . Allora esiste una matrice unitaria $U \in U_n$ tale che la matrice $U^{-1}AU = \overline{U}^T AU$ è una matrice diagonale.

Dimostrazione. Vedi l'Esercizio 4.12. ■

Esercizio 4.12 Dimostrare il Corollario 4.3.15 (*Suggerimento: si osservi che l'endomorfismo $L_A : \mathbb{C}^n \rightarrow \mathbb{C}^n$ associato ad A è hermitiano rispetto al prodotto hermitiano standard. In particolare, dalla Proposizione 4.3.14, L_A è ortogonalmente diagonalizzabile. Ora si ragioni come nella dimostrazione della Proposizione 3.4.1 (si veda anche la Proposizione 3.5.6).*) ■

Esercizio 4.13 Si consideri la matrice

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

Dopo aver osservato che si tratta di una matrice hermitiana, si determini una matrice unitaria $U \in U_3$ tale che $U^{-1}AU = \overline{U}^T AU$ è una matrice diagonale. Si calcoli $U^{-1}AU$. ■

La Proposizione 4.3.14 non è altrettanto potente quanto il Teorema Spettrale. Mentre un endomorfismo di uno spazio euclideo è ortogonalmente diagonalizzabile sse è simmetrico, esistono endomorfismi di uno spazio hermitiano che sono ortogonalmente diagonalizzabili ma non sono hermitiani. Per esempio esistono endomorfismi unitari (che sono ortogonalmente diagonalizzabili in base alla Proposizione 4.3.7) che non sono hermitiani. Per convincersi di ciò è sufficiente esibire una matrice unitaria che non sia hermitiana. È il caso, per esempio, della matrice nell'Esercizio 4.11 (il lettore lo vede?). Un altro esempio di endomorfismo unitario che non è hermitiano è l'endomorfismo nell'Esempio 4.9 (il lettore lo vede?). Resta dunque aperta la domanda: *è possibile caratterizzare gli endomorfismi ortogonalmente diagonalizzabili di uno spazio hermitiano in modo efficiente quanto nel Teorema Spettrale?* La risposta è affermativa, vale infatti il seguente

Teorema 4.3.16 Sia $(V, \langle -, - \rangle)$ uno spazio hermitiano di dimensione n . Un endomorfismo $f : V \rightarrow V$ è ortogonalmente diagonalizzabile (cioè esiste un riferimento ortonormale \mathcal{R} di $(V, \langle -, - \rangle)$ di autovettori di f) sse f commuta con l'endomorfismo hermitiano coniugato f^\dagger , cioè $f \circ f^\dagger = f^\dagger \circ f$.

La dimostrazione del Teorema 4.3.16 va oltre gli scopi di queste note ed è omessa. Ci limitiamo ad osservare che la condizione $f \circ f^\dagger = f^\dagger \circ f$ è chiaramente soddisfatta dagli endomorfismi hermitiani. In effetti, è anche soddisfatta dagli endomorfismi unitari f per i quali si ha $f^\dagger = f^{-1}$ (a tal proposito si veda l'Esercizio 4.14).

Esercizio 4.14 Sia $(V, \langle -, - \rangle)$ uno spazio hermitiano di dimensione n . Dimostrare che un endomorfismo $f : V \rightarrow V$ è unitario sse è invertibile e $f^\dagger = f^{-1}$. ■

5. Spazi Vettoriali Orientati

In questo capitolo introduciamo il nuovo concetto di *orientazione* di uno spazio vettoriale reale. Un'orientazione in uno spazio vettoriale euclideo 3-dimensionale consente di introdurre una nuova operazione di *prodotto vettoriale* che tornerà utile nel Capitolo 6.

5.1 Orientazioni

Sia V uno spazio vettoriale reale di dimensione n e siano $\mathcal{R}, \mathcal{R}'$ riferimenti di V .

Definizione 5.1.1 — Riferimenti Concordemente Orientati. I riferimenti \mathcal{R} ed \mathcal{R}' si dicono *concordemente orientati* se la matrice del cambio di coordinate $M_{\mathcal{R}', \mathcal{R}}$ ha determinante positivo:

$$\det M_{\mathcal{R}', \mathcal{R}} > 0.$$

In tal caso scriviamo $\mathcal{R} \sim \mathcal{R}'$. In caso contrario \mathcal{R} e \mathcal{R}' si dicono *discordemente orientati*.

Si osservi che la definizione di *riferimenti concordemente orientati* non ha senso in uno spazio vettoriale su un campo \mathbb{K} generico, perché il concetto di “scalare positivo” non ha senso in un campo generico.

■ **Esempio 5.1** Sia $V = \mathbb{R}^3$ lo spazio vettoriale standard 3-dimensionale, sia $\mathcal{R}' = \mathcal{R}_{can}$ il riferimento canonico di \mathbb{R}^3 e sia \mathcal{R} il riferimento

$$\mathcal{R} = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right).$$

I riferimenti \mathcal{R} ed \mathcal{R}' sono concordemente orientati, infatti la matrice del cambio di coordinate $M_{\mathcal{R}', \mathcal{R}}$ è la matrice

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

che ha determinante pari ad $1 > 0$. Invece, il riferimento \mathcal{R}'' , con

$$\mathcal{R}'' = \left(\left(\begin{array}{c} -1 \\ 0 \\ 0 \end{array} \right), \left(\begin{array}{c} 1 \\ 1 \\ 0 \end{array} \right), \left(\begin{array}{c} 1 \\ 1 \\ 1 \end{array} \right) \right).$$

e il riferimento \mathcal{R}' sono discordemente orientati, infatti la matrice del cambio di coordinate $M_{\mathcal{R}', \mathcal{R}''}$ è la matrice

$$\begin{pmatrix} -1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

che ha determinante pari a $-1 < 0$. ■

Proposizione 5.1.1 La relazione \sim (“essere concordemente orientati”) è una ben definita relazione di equivalenza sull’insieme $\text{Fr}(V)$ dei riferimenti di V . Rispetto a tale relazione di equivalenza esistono esattamente 2 classi di equivalenza.

Dimostrazione. Per la prima parte della dimostrazione dobbiamo verificare che la relazione \sim è riflessiva, simmetrica e transitiva. Perciò siano $\mathcal{R}, \mathcal{R}', \mathcal{R}''$ tre riferimenti di V . Chiaramente $\mathcal{R} \sim \mathcal{R}$, infatti

$$\det M_{\mathcal{R}, \mathcal{R}} = \det I_n = 1 > 0.$$

Quindi \sim è una relazione riflessiva. Ora, sia $\mathcal{R} \sim \mathcal{R}'$. Allora $\det(M_{\mathcal{R}', \mathcal{R}}) > 0$. Per verificare che $\mathcal{R}' \sim \mathcal{R}$ dobbiamo calcolare

$$\det M_{\mathcal{R}, \mathcal{R}'} = \det(M_{\mathcal{R}', \mathcal{R}}^{-1}) = \det(M_{\mathcal{R}', \mathcal{R}})^{-1} > 0.$$

Quindi \sim è una relazione simmetrica. Infine sia $\mathcal{R} \sim \mathcal{R}'$ e $\mathcal{R}' \sim \mathcal{R}''$. Allora $\det M_{\mathcal{R}', \mathcal{R}}, \det M_{\mathcal{R}'', \mathcal{R}'} > 0$. Per verificare che $\mathcal{R} \sim \mathcal{R}''$ dobbiamo calcolare

$$\det M_{\mathcal{R}'', \mathcal{R}} = \det(M_{\mathcal{R}'', \mathcal{R}'} M_{\mathcal{R}', \mathcal{R}}) = \det M_{\mathcal{R}'', \mathcal{R}'} \det M_{\mathcal{R}', \mathcal{R}} > 0,$$

dove abbiamo usato il Teorema di Binet. Quindi \sim è anche una relazione transitiva.

Per dimostrare che ci sono solo due classi di equivalenza consideriamo due riferimenti $\mathcal{R}, \mathcal{R}'$ discordemente orientati. Allora $\det M_{\mathcal{R}', \mathcal{R}} < 0$. È sufficiente dimostrare che se un terzo riferimento \mathcal{R}'' è discordemente orientato con \mathcal{R} , allora è concordemente orientato con \mathcal{R}' (in altre parole \mathcal{R} e \mathcal{R}' sono in due classi di equivalenza distinte e ogni altro riferimento appartiene all’una o all’altra). Consideriamo dunque \mathcal{R}'' un tale riferimento. Si ha $\det M_{\mathcal{R}'', \mathcal{R}} < 0$, ma

$$\det M_{\mathcal{R}'', \mathcal{R}'} = \det(M_{\mathcal{R}'', \mathcal{R}} M_{\mathcal{R}, \mathcal{R}'}) = \det M_{\mathcal{R}'', \mathcal{R}} \det M_{\mathcal{R}, \mathcal{R}'} > 0.$$

■

Definizione 5.1.2 — Orientazione. Un’*orientazione* sullo spazio vettoriale reale finito-dimensionale V è ciascuna delle due classi di equivalenza in cui l’insieme $\text{Fr}(V)$ dei riferimenti di V è diviso dalla relazione di equivalenza \sim . Uno *spazio vettoriale orientato* è una coppia (V, \mathfrak{o}_+) in cui V è uno spazio reale finito-dimensionale e \mathfrak{o}_+ è un’orientazione su V . Dato uno spazio vettoriale orientato (V, \mathfrak{o}_+) , l’orientazione \mathfrak{o}_+ è detta *positiva*, ed ogni riferimento \mathcal{R} che appartiene alla classe di equivalenza \mathfrak{o}_+ si dice *positivamente orientato* o, semplicemente, *positivo* (si dice anche che \mathcal{R} *determina l’orientazione* \mathfrak{o}_+). L’altra orientazione \mathfrak{o}_- è detta *negativa*, ed ogni riferimento \mathcal{R}' che appartiene alla classe di equivalenza \mathfrak{o}_- si dice *negativamente orientato* o,

■ semplicemente, *negativo*.

Si osservi che i concetti di *orientazione positiva* e *orientazione negativa* hanno senso solo *dopo* che un'orientazione sia stata fissata e, di norma, su uno spazio vettoriale reale, non esiste una orientazione privilegiata.

■ **Esempio 5.2** Sia \mathbb{R}^n lo spazio vettoriale reale standard n -dimensionale. Il riferimento canonico \mathcal{R}_{can} determina un'orientazione ν_{can} su \mathbb{R}^n che si chiama *orientazione standard*. ■

Proposizione 5.1.2 — Orientazione e Trasformazioni Elementari. Sia V uno spazio vettoriale reale di dimensione n e sia $\mathcal{R} = (v_1, \dots, v_n)$ un riferimento di V . Allora

- (1) Fissati $i, j = 1, \dots, n$, con $i \neq j$, ed $a \in \mathbb{R}$, il riferimento \mathcal{R}' che si ottiene da \mathcal{R} mediante la trasformazione elementare

$$v_i \rightarrow v_i + av_j$$

è concordemente orientato con \mathcal{R} per ogni i, j , e per ogni a ;

- (2) Fissati $i, j = 1, \dots, n$, $i \neq j$, il riferimento \mathcal{R}'' che si ottiene da \mathcal{R} mediante la trasformazione elementare

$$v_i \leftrightarrow v_j$$

è discordemente orientato con \mathcal{R} per ogni i, j ;

- (3) Fissati $i = 1, \dots, n$ e $a \in \mathbb{R}$, con $a \neq 0$, il riferimento \mathcal{R}''' che si ottiene da \mathcal{R} mediante la trasformazione elementare

$$v_i \rightarrow av_i$$

è concordemente orientato con \mathcal{R} se $a > 0$, ed è discordemente orientato con \mathcal{R} se $a < 0$.

Dimostrazione. La matrice del cambio di coordinate $M_{\mathcal{R}', \mathcal{R}}$ si ottiene dalla matrice identica addizionando alla i -esima colonna la j -esima colonna moltiplicata per a . Poiché il determinante di una matrice non cambia sotto trasformazioni elementari di prima specie sulle colonne, si ha $\det M_{\mathcal{R}', \mathcal{R}} = \det I_n = 1 > 0$. Questo dimostra il punto (1). I punti (2) e (3) si dimostrano in modo simile e i dettagli sono lasciati al lettore. ■

5.2 Automorfismi che Conservano le Orientazioni

Dato uno spazio vettoriale reale V di dimensione n è possibile individuare una classe speciale di automorfismi $f : V \rightarrow V$ che, in un certo senso opportuno, *conservano le orientazioni*. Cominciamo con una

Proposizione 5.2.1 Sia $f : V \rightarrow V$ un automorfismo. Le seguenti condizioni sono equivalenti:

- (1) per ogni riferimento \mathcal{R} di V , il riferimento $f(\mathcal{R})$ è un riferimento concordemente orientato con \mathcal{R} ;
- (2) esiste un riferimento \mathcal{R} di V tale che $f(\mathcal{R})$ è un riferimento concordemente orientato con \mathcal{R} .
- (3) per ogni riferimento $\tilde{\mathcal{R}}$ di V , la matrice $M_{\tilde{\mathcal{R}}}^f$ che rappresenta f nel riferimento $\tilde{\mathcal{R}}$ ha determinante positivo;
- (4) esiste un riferimento $\tilde{\mathcal{R}}$ di V tale che la matrice rappresentativa $M_{\tilde{\mathcal{R}}}^f$ ha determinante positivo.

Dimostrazione. Dimostriamo che $(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4) \Rightarrow (1)$. L'implicazione $(1) \Rightarrow (2)$ è ovvia. Ora assumiamo (2). Indichiamo con \mathcal{R}' il riferimento $f(\mathcal{R})$. Allora la matrice rappresentativa $M_{\mathcal{R}',\mathcal{R}}^f$ è semplicemente la matrice identica I_n . Sia $\tilde{\mathcal{R}}$ un qualunque altro riferimento di V . Abbiamo

$$\begin{aligned} M_{\tilde{\mathcal{R}}}^f &= M_{\tilde{\mathcal{R}},\mathcal{R}'} M_{\mathcal{R}',\mathcal{R}}^f M_{\mathcal{R},\tilde{\mathcal{R}}} = M_{\tilde{\mathcal{R}},\mathcal{R}'} I_n M_{\mathcal{R},\tilde{\mathcal{R}}} = M_{\tilde{\mathcal{R}},\mathcal{R}'} M_{\mathcal{R},\tilde{\mathcal{R}}} = M_{\tilde{\mathcal{R}},\mathcal{R}'} M_{\mathcal{R},\mathcal{R}'} M_{\mathcal{R}',\tilde{\mathcal{R}}} \\ &= M_{\tilde{\mathcal{R}},\mathcal{R}'} M_{\mathcal{R},\mathcal{R}'} M_{\tilde{\mathcal{R}},\mathcal{R}'}^{-1}. \end{aligned}$$

Perciò, dal Teorema di Binet,

$$\det M_{\tilde{\mathcal{R}}}^f = \det M_{\tilde{\mathcal{R}},\mathcal{R}'} \det M_{\mathcal{R},\mathcal{R}'} \det M_{\tilde{\mathcal{R}},\mathcal{R}'}^{-1} = \det M_{\mathcal{R},\mathcal{R}'} > 0,$$

come desiderato. Ora è chiaro che $(3) \Rightarrow (4)$ e resta da dimostrare che $(4) \Rightarrow (1)$. Dunque sia $\tilde{\mathcal{R}}$ un riferimento di V come in (4) e sia \mathcal{R} un qualunque altro riferimento di V . Poniamo $\mathcal{R}' = f(\mathcal{R})$. Lo stesso conto che abbiamo fatto sopra mostra che

$$0 < \det M_{\tilde{\mathcal{R}}}^f = \det M_{\mathcal{R},\mathcal{R}'},$$

come desiderato. ■

Definizione 5.2.1 — Automorfismo che Conserva le Orientazioni. Un automorfismo $f : V \rightarrow V$ conserva le orientazioni se f soddisfa una delle, e quindi tutte le, condizioni nella Proposizione 5.2.1. In caso contrario si dice che f inverte le orientazioni.

Proposizione 5.2.2 Gli automorfismi che conservano le orientazioni formano un sottogruppo nel gruppo $GL(V)$ degli automorfismi di V .

Dimostrazione. La proposizione viene lasciata come Esercizio 5.1 ■

Esercizio 5.1 Dimostrare la Proposizione 5.2.2. ■

Il sottogruppo di $GL(V)$ formato dagli automorfismi che conservano le orientazioni si indica con $GL^+(V)$.

Corollario 5.2.3 Sia V uno spazio vettoriale reale di dimensione n e siano $\mathcal{R}, \mathcal{R}'$ riferimenti concordemente orientati di V . Esiste un unico automorfismo $f : V \rightarrow V$ che preserva le orientazioni tale che $f(\mathcal{R}) = \mathcal{R}'$.

Prima di procedere osserviamo che, essenzialmente dal Teorema di Binet, le matrici quadrate di ordine n con determinante positivo formano un sottogruppo del gruppo lineare generale $GL_n(\mathbb{R})$. Tale sottogruppo si indica con GL_n^+ .

Corollario 5.2.4 Sia V uno spazio vettoriale reale di dimensione n e sia \mathcal{R} un riferimento di V . L'applicazione $GL^+(V) \rightarrow GL_n^+$, $f \mapsto M_{\mathcal{R}}^f$ che associa ad un automorfismo $f : V \rightarrow V$ che conserva le orientazioni la sua matrice rappresentativa nel riferimento \mathcal{R} è un ben definito isomorfismo di gruppi.

■ **Esempio 5.3** Nel caso in cui $V = \mathbb{R}^n$ è lo spazio vettoriale numerico, possiamo scegliere il riferimento canonico \mathcal{R}_{can} . In questo caso il Corollario 5.2.4 dice che gli automorfismi di \mathbb{R}^n che preservano le orientazioni sono esattamente le applicazioni lineari $L_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ associate alle matrici di determinante positivo $A \in GL_n^+$. Esiste dunque un isomorfismo naturale di gruppi $GL^+(\mathbb{R}^n) \cong GL_n^+$ il cui isomorfismo inverso è dato da $A \mapsto L_A$ (cfr. la presente situazione con quella nell'Esempio 2.9). ■

5.3 Spazi Vettoriali Euclidei Orientati

Dato uno spazio vettoriale euclideo $(V, \langle -, - \rangle)$ ha senso considerare gli automorfismi ortogonali $f : (V, \langle -, - \rangle) \rightarrow (V, \langle -, - \rangle)$ che conservano le orientazioni.

In questo contesto valgono proposizioni analoghe a quelle già viste per gli automorfismi ortogonali e per gli automorfismi che conservano le orientazioni. Prima di elencare i fatti più rilevanti a tale riguardo, osserviamo che, per ogni $n \in \mathbb{N}$, l'intersezione $O_n \cap GL_n^+$ del gruppo ortogonale e del gruppo delle matrici con determinante positivo, è esattamente il gruppo ortogonale speciale: $O_n \cap GL_n^+ = SO_n$. Ora valgono le seguenti osservazioni:

- ✓ dato un riferimento ortonormale \mathcal{R} , un altro riferimento \mathcal{R}' è ortonormale e concordemente orientato con \mathcal{R} sse la matrice del cambio di coordinate $M_{\mathcal{R}', \mathcal{R}}$ è una matrice ortogonale speciale;
- ✓ un automorfismo è ortogonale e, simultaneamente, conserva le orientazioni sse la sua matrice rappresentativa in un riferimento ortonormale (equivalentemente in *tutti* i riferimenti ortonormali) è una matrice ortogonale speciale;
- ✓ gli automorfismi ortogonali $f : (V, \langle -, - \rangle) \rightarrow (V, \langle -, - \rangle)$ di uno spazio vettoriale euclideo $(V, \langle -, - \rangle)$ che preservano le orientazioni formano un sottogruppo nel gruppo lineare generale (che si indica $SO(V, \langle -, - \rangle) = O(V, \langle -, - \rangle) \cap GL^+(V)$);
- ✓ L'applicazione $SO(V, \langle -, - \rangle) \rightarrow SO_n, f \mapsto M_{\mathcal{R}}^f$ che associa ad un automorfismo ortogonale f che preserva le orientazioni la sua matrice rappresentativa in un riferimento ortonormale fissato \mathcal{R} , è un isomorfismo di gruppi.

I dettagli sono lasciati per esercizio al lettore.

Ci concentriamo ora su uno spazio vettoriale euclideo orientato 3-dimensionale $(V, \langle -, - \rangle, \nu_+)$. In questo caso, è possibile definire su V una nuova operazione interna

$$\wedge : V \times V \rightarrow V,$$

detta il *prodotto vettoriale*, che è uno strumento particolarmente utile in Geometria Affine Euclidea 3D (si veda la Sezione 8.1). Sia dunque $(V, \langle -, - \rangle, \nu_+)$ uno spazio vettoriale euclideo orientato 3-dimensionale e sia $\mathcal{R} = (e_1, e_2, e_3)$ un riferimento ortonormale positivo. Dati due vettori $v, w \in V$, indichiamo con $x = (x_1, x_2, x_3)^T, y = (y_1, y_2, y_3)^T$ i vettori coordinati di v, w rispettivamente. Consideriamo la matrice

$$\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{pmatrix} \quad (5.1)$$

e indichiamo anche con z_1, z_2, z_3 i relativi minori di ordine 2, presi con segni alterni. In altre parole

$$z_1 = \det \begin{pmatrix} x_2 & y_2 \\ x_3 & y_3 \end{pmatrix}, \quad z_2 = -\det \begin{pmatrix} x_1 & y_1 \\ x_3 & y_3 \end{pmatrix}, \quad \text{e} \quad z_3 = \det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}. \quad (5.2)$$

Il vettore $v \wedge w$ è, per definizione, il vettore di coordinate z_1, z_2, z_3 nel riferimento \mathcal{R} . Cioè

$$v \wedge w = z_1 e_1 + z_2 e_2 + z_3 e_3.$$

L'operazione

$$\wedge : V \times V \rightarrow V$$

così introdotta prende il nome di *prodotto vettoriale* o *prodotto wedge* (dall'inglese *wedge* = cuneo, per il simbolo usato per indicarla). Apparentemente il prodotto vettoriale dipende dalla scelta del riferimento \mathcal{R} . Ma in effetti possiamo dimostrare che esso è indipendente da tale scelta, purché \mathcal{R} sia un riferimento *ortonormale e positivo*. Prima di fare ciò è conveniente dimostrare alcune proprietà dell'operazione \wedge appena introdotta.

Proposizione 5.3.1 Sia $(V, \langle -, - \rangle, \nu_+)$ uno spazio vettoriale euclideo orientato 3-dimensionale e sia \mathcal{R} un riferimento ortonormale e positivo di V . Il prodotto vettoriale \wedge (associato ad \mathcal{R}) è bilineare e antisimmetrico (cioè $w \wedge v = -v \wedge w$, per ogni $v, w \in V$). Inoltre, per ogni $v, w \in V$

- (1) $v \wedge w \perp v$ e $v \wedge w \perp w$;
- (2) $v \wedge w = 0$ sse v e w sono linearmente dipendenti;
- (3) $\|v \wedge w\|^2 = \|v\|^2 \|w\|^2 - \langle v, w \rangle^2$.

Dimostrazione. La linearità del prodotto vettoriale nel primo argomento segue dal fatto che i minori z_1, z_2, z_3 nella (5.2) sono funzioni lineari del vettore v . Nel dettaglio, siano $v, v', w \in V$ e siano $a, a' \in \mathbb{R}$. Calcoliamo $(av + a'v') \wedge w$. A questo scopo indichiamo con $x = (x_1, x_2, x_3)^T, x' = (x'_1, x'_2, x'_3)^T, y = (y_1, y_2, y_3)^T$ i vettori coordinati di v, v', w nel riferimento \mathcal{R} rispettivamente. Indichiamo infine con z_1, z_2, z_3 le coordinate di $v \wedge w$ e con z'_1, z'_2, z'_3 le coordinate di $v' \wedge w$ (nel riferimento \mathcal{R}). Allora il vettore coordinato di $av + a'v'$ è

$$ax + a'x' = \begin{pmatrix} ax_1 + a'x'_1 \\ ax_2 + a'x'_2 \\ ax_3 + a'x'_3 \end{pmatrix}.$$

Perciò

$$(av + a'v') \wedge w = Z_1 e_1 + Z_2 e_2 + Z_3 e_3$$

dove Z_1, Z_2, Z_3 sono i minori di ordine 2 della matrice

$$\begin{pmatrix} ax_1 + a'x'_1 & y_1 \\ ax_2 + a'x'_2 & y_2 \\ ax_3 + a'x'_3 & y_3 \end{pmatrix}$$

presi a segni alterni come sopra. Ma è chiaro, dalle proprietà del determinante, che

$$Z_i = az_i + a'z'_i, \quad i = 1, 2, 3.$$

Ne consegue immediatamente che

$$(av + a'v') \wedge w = a(v \wedge w) + a'(v' \wedge w),$$

cioè il prodotto wedge è lineare nel primo argomento (il lettore che non abbia seguito tutti i dettagli della discussione è invitato a fare tutti i dovuti conti per esercizio). La linearità nel secondo argomento si dimostra nello stesso modo.

L'antisimmetria segue dal fatto che i minori z_1, z_2, z_3 nella (5.2) cambiano di segno quando i fattori v, w nel prodotto vettoriale $v \wedge w$ si scambiano di posto. Il lettore è invitato a verificare i dettagli per esercizio.

Per il punto (1), come al solito indichiamo con (x_1, x_2, x_3) le coordinate di v , con (y_1, y_2, y_3) le coordinate di w e con (z_1, z_2, z_3) le coordinate di $v \wedge w$ nel riferimento \mathcal{R} . Giacché \mathcal{R} è un riferimento ortonormale si ha

$$\langle v, v \wedge w \rangle = x_1 z_1 + x_2 z_2 + x_3 z_3.$$

Ma il membro di destra è lo sviluppo di Laplace (rispetto alla prima colonna) del determinante

$$\det \begin{pmatrix} x_1 & x_1 & y_1 \\ x_2 & x_2 & y_2 \\ x_3 & x_3 & y_3 \end{pmatrix}.$$

Poiché la matrice in questione ha due colonne uguali, tale determinante è 0. Analogamente $\langle w, v \wedge w \rangle = 0$.

Per il punto (2), osserviamo che $v \wedge w = 0$ sse sono nulle le sue coordinate (z_1, z_2, z_3) , ma questo vuol dire esattamente che i minori di ordine 2 della matrice (5.1) sono nulli, e ciò avviene sse le colonne $x = (x_1, x_2, x_3)^T$ e $y = (y_1, y_2, y_3)^T$ sono linearmente dipendenti. Infine, poiché x, y sono proprio i vettori coordinati di v, w si ha che x, y sono linearmente dipendenti sse v, w sono linearmente dipendenti, come voluto.

Per il punto (3) dobbiamo calcolare la norma quadra $\|v \wedge w\|^2$. Poiché \mathcal{R} è un riferimento ortonormale

$$\|v \wedge w\|^2 = z_1^2 + z_2^2 + z_3^2.$$

La formula (3) ora segue da un conto diretto che sfrutta la definizione (5.2) e che lasciamo al lettore per esercizio. Qui ci limitiamo ad osservare che la (3) si può anche scrivere nella forma

$$\|v \wedge w\|^2 = \det \begin{pmatrix} \langle v, v \rangle & \langle v, w \rangle \\ \langle w, v \rangle & \langle w, w \rangle \end{pmatrix}.$$

Il determinante al membro di destra è talvolta detto il *determinante di Gram*. ■

Prima di dimostrare l'indipendenza del prodotto vettoriale dalla scelta del riferimento \mathcal{R} (tra i riferimenti ortonormali e positivi) abbiamo bisogno di un ulteriore

Lemma 5.3.2 Siano $v, w \in V$ vettori linearmente indipendenti. Allora il sistema $S = (v, w, v \wedge w)$ è un riferimento positivo di V (vedi Figura 5.1).

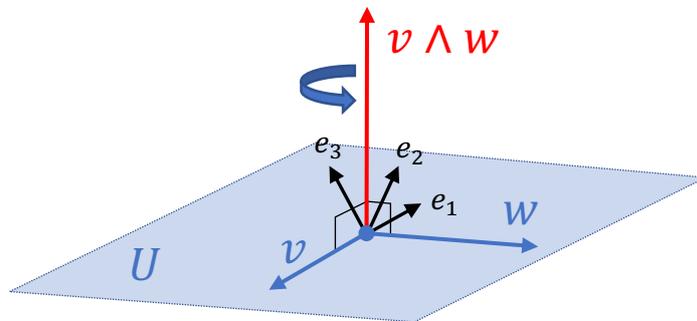


Figura 5.1: Prodotto vettoriale di due vettori indipendenti.

Dimostrazione. Sia $U \subseteq V$ il sottospazio vettoriale generato da v, w . Se v, w sono linearmente indipendenti, allora $\dim U = 2$. Inoltre, $v \wedge w \neq 0$ (Proposizione 5.3.1.(2)) e $v \wedge w \in U^\perp$ (Proposizione 5.3.1.(1)). Poiché U e U^\perp sono in somma diretta $v \wedge w$ non può appartenere ad U . In altre parole il sistema S è linearmente indipendente. Essendo formato da 3 vettori, è anche un riferimento di V e resta da dimostrare che tale riferimento è positivo. In effetti è sufficiente mostrare che il riferimento S è concordemente orientato con \mathcal{R} . A questo scopo dobbiamo calcolare il determinante della matrice del cambio di coordinate $M_{\mathcal{R},S}$. Usando le solite notazioni per le coordinate di $v, w, v \wedge w$ nel riferimento \mathcal{R} abbiamo

$$M_{\mathcal{R},S} = \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix},$$

il cui determinante può essere calcolato, per esempio, con la regola di Laplace, sviluppando rispetto all'ultima colonna. In tal caso troviamo

$$\det M_{\mathcal{R},S} = z_1^2 + z_2^2 + z_3^2 = \|v \wedge w\|^2,$$

dove abbiamo anche usato che \mathcal{R} è un riferimento ortonormale. Giacché $v \wedge w \neq 0$, abbiamo $\det M_{\mathcal{R},S} > 0$ come desiderato. ■

Siamo ora pronti a dimostrare il principale risultato di questa sezione.

Teorema 5.3.3 Sia $(V, \langle -, - \rangle, \mathfrak{o}_+)$ uno spazio vettoriale euclideo orientato 3-dimensionale. Il prodotto vettoriale \wedge non dipende dalla scelta del riferimento ortonormale positivo $\mathcal{R} = (e_1, e_2, e_3)$, cioè se $\mathcal{R}' = (e'_1, e'_2, e'_3)$ è un altro riferimento ortonormale positivo e \wedge' è il prodotto vettoriale definito a partire da \mathcal{R}' , allora

$$v \wedge w = v \wedge' w$$

per ogni $v, w \in V$.

Dimostrazione. Siano $\mathcal{R}, \mathcal{R}'$ riferimenti ortonormali positivi di V , e siano \wedge, \wedge' i prodotti vettoriali definiti a partire da $\mathcal{R}, \mathcal{R}'$ rispettivamente. Ovviamente, anche il prodotto vettoriale \wedge' soddisfa tutte le proprietà elencate nella Proposizione 5.3.1. In particolare se $v, w \in V$ sono vettori linearmente dipendenti, allora

$$v \wedge' w = 0 = v \wedge w.$$

Questo mostra che i prodotti vettoriali \wedge, \wedge' coincidono quando agiscono tra gli stessi due vettori dipendenti. Consideriamo ora il caso in cui $v, w \in V$ siano vettori linearmente indipendenti. In questo caso, v, w generano un sottospazio bidimensionale $U \subseteq V$. Di nuovo dalla Proposizione 5.3.1, i vettori $v \wedge w$ e $v \wedge' w$

- (1) sono entrambi non nulli;
- (2) appartengono entrambi al sottospazio 1-dimensionale U^\perp ;
- (3) hanno la stessa norma $\alpha = \|v \wedge w\| = \|v \wedge' w\|$.

Inoltre i sistemi $S = (v, w, v \wedge w)$ ed $S' = (v, w, v \wedge' w)$ sono riferimenti concordemente orientati con \mathcal{R} , e quindi sono concordemente orientati tra loro. Siccome in U^\perp ci sono esattamente due vettori di norma pari ad α e sono uno l'opposto dell'altro, ne consegue che $v \wedge w$ e $v \wedge' w$ possono differire al più per il segno. Supponiamo per assurdo che $v \wedge' w = -v \wedge w$, allora i riferimenti S ed S' sarebbero discordemente orientati, che è un assurdo. Perciò deve necessariamente essere

$$v \wedge' w = v \wedge w.$$

■

○ Sia $(V, \langle -, - \rangle, \mathfrak{o}_+)$ uno spazio vettoriale euclideo orientato 3-dimensionale e sia

$$\wedge : V \times V \rightarrow V$$

il prodotto vettoriale. Si può usare un riferimento negativamente orientato per definire un altro prodotto vettoriale

$$\wedge_- : V \times V \rightarrow V$$

esattamente nello stesso modo. Chiaramente, il prodotto \wedge_- godrà di tutte le proprietà elencate dalla Proposizione 5.3.1. Inoltre, dati due vettori indipendenti $v, w \in V$, il sistema $(v, w, v \wedge_- w)$

sarà un riferimento negativo. Ragionando come nella dimostrazione del Teorema 5.3.3 si vede allora facilmente che

$$v \wedge_{-} w = -v \wedge w,$$

per ogni v, w . In altre parole, dato uno spazio vettoriale euclideo 3-dimensionale V , ciascuna delle due orientazioni di V definisce un diverso prodotto vettoriale. I due prodotti vettoriali differiscono solo per il segno.

Proposizione 5.3.4 Sia $(V, \langle -, - \rangle, v_+)$ uno spazio vettoriale euclideo orientato 3-dimensionale, e siano $v, w \in V$ con $v \neq 0$. Diciamo w' la proiezione ortogonale di w sul sottospazio $\text{Span}(v)$, di modo che $w = w' + w^\perp$ con $w^\perp \in v^\perp$. Allora

$$\|v \wedge w\| = \|v\| \|w^\perp\|$$

Dimostrazione. Innanzitutto

$$v \wedge w = v \wedge (w' + w^\perp) = v \wedge w' + v \wedge w^\perp = v \perp w^\perp, \quad (5.3)$$

dove abbiamo usato la bilinearità del prodotto scalare e il fatto che v e w' sono linearmente dipendenti. Dalla (5.3) segue che

$$\|v \wedge w\|^2 = \|v \wedge w^\perp\|^2 = \|v\|^2 \|w^\perp\|^2 - \langle v, w \rangle^2 = \|v\|^2 \|w^\perp\|^2,$$

dove abbiamo usato che $v \perp w^\perp$. Questo conclude la dimostrazione. ■

○ Per la norma del prodotto vettoriale vale anche

$$\|v \wedge w\| = \|v\| \|w\| \sin \widehat{vw} \quad v, w \in V. \quad (5.4)$$

Infatti, se $v = 0$ o $w = 0$, la (5.4) è chiaramente vera. Se entrambi v e w sono diversi da 0, allora

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

da cui

$$\begin{aligned} \|v \wedge w\|^2 &= \|v\|^2 \|w\|^2 - \langle v, w \rangle^2 = \|v\|^2 \|w\|^2 - \|v\|^2 \|w\|^2 \cos^2 \widehat{vw} \\ &= \|v\|^2 \|w\|^2 (1 - \cos^2 \widehat{vw}) = \|v\|^2 \|w\|^2 \sin^2 \widehat{vw}. \end{aligned}$$

Poiché il seno dell'angolo convesso \widehat{vw} è non-negativo per ogni v, w , la (5.4) segue.

5.4 Classificazione degli Endomorfismi Ortogonali in Dimensione 2 e 3

In questa sezione, come applicazione del Teorema Fondamentale dell'Algebra, discutiamo più da vicino gli endomorfismi ortogonali di uno spazio vettoriale euclideo in dimensione 2 e 3. Il nostro scopo è dimostrare un *Teorema di Classificazione*. In altre parole, mostreremo che è possibile suddividere gli endomorfismi ortogonali in un certo numero di classi. Ciascuna fissata classe C è caratterizzata dal fatto che esiste una matrice A_C , particolarmente semplice, tale che, per ogni endomorfismo f in C , la matrice rappresentativa di f in un opportuno riferimento è proprio A_C . La matrice A_C è talvolta detta una *forma normale* per gli endomorfismi nella classe C . Come risultato collaterale forniremo una descrizione geometrica degli endomorfismi ortogonali in dimensione 2 e 3. La sezione è collocata in questo capitolo perché faremo uso esplicito della nozione di *endomorfismo che preserva le orientazioni*.

Cominciamo con un

Lemma 5.4.1 Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo di dimensione n , sia

$$f : (V, \langle -, - \rangle) \rightarrow (V, \langle -, - \rangle)$$

un endomorfismo ortogonale, sia λ un autovalore di f . Allora $\lambda = \pm 1$. Inoltre, se $v \in V$ è un autovettore relativo all'autovalore λ , allora il sottospazio v^\perp è f -invariante, cioè per ogni $w \in v^\perp$, si ha $f(w) \in v^\perp$.

Dimostrazione. Siano λ e v come nell'enunciato. Allora, poiché f conserva la norma

$$\|v\| = \|f(v)\| = \|\lambda v\| = |\lambda| \|v\|. \quad (5.5)$$

Ma v è un autovettore, perciò $v \neq 0$ e dunque $\|v\| \neq 0$. Allora, nella (5.5), possiamo dividere tutto per $\|v\|$, trovando $|\lambda| = 1$, come desiderato.

Passiamo ora alla seconda parte dell'enunciato. Sia $w \in v^\perp$. Allora

$$0 = \langle w, v \rangle = \langle f(w), f(v) \rangle = \langle f(w), \pm v \rangle = \pm \langle f(w), v \rangle.$$

In ogni caso $\langle f(w), v \rangle = 0$, cioè $f(w) \in v^\perp$ come desiderato. ■

Si osservi che, qualunque sia la dimensione dello spazio vettoriale euclideo ambiente $(V, \langle -, - \rangle)$, esistono sia endomorfismi ortogonali con un autovalore pari a $+1$ sia endomorfismi ortogonali con un autovalore pari a -1 . Per esempio l'endomorfismo identico $\text{id}_V : (V, \langle -, - \rangle) \rightarrow (V, \langle -, - \rangle)$ è un endomorfismo ortogonale con autovalore $+1$. Analogamente l'endomorfismo $-\text{id}_V : (V, \langle -, - \rangle) \rightarrow (V, \langle -, - \rangle)$ è un endomorfismo ortogonale con autovalore -1 .

Discutiamo ora gli endomorfismi ortogonali in dimensione 2.

Teorema 5.4.2 Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo di dimensione 2 e sia

$$f : (V, \langle -, - \rangle) \rightarrow (V, \langle -, - \rangle)$$

un endomorfismo ortogonale. Allora

[D2-1] Se f preserva le orientazioni, allora solo uno dei seguenti tre casi si può presentare:

- $f = \text{id}_V$, in particolare f è ortogonalmente diagonalizzabile con un unico autovalore pari ad 1;
- $f = -\text{id}_V$, in particolare f è ortogonalmente diagonalizzabile con un unico autovalore pari a -1 . In altre parole f è la *riflessione rispetto all'origine* (Figura 5.2);
- f non possiede autovalori e, in tal caso, la matrice rappresentativa di f in ogni riferimento ortonormale è una matrice di rotazione

$$R_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix},$$

con $\theta \in [0, 2\pi)$, ma $\theta \neq 0, \pi$. In altre parole f è una *rotazione di angolo* $\theta \neq 0, \pi$ intorno all'origine (Figura 5.3).

[D2-2] Se f inverte le orientazioni, allora f è ortogonalmente diagonalizzabile ed esiste un riferimento ortonormale nel quale la matrice rappresentativa di f è la matrice

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (5.6)$$

In altre parole, f è una *riflessione rispetto ad una retta* (Figura 5.4).

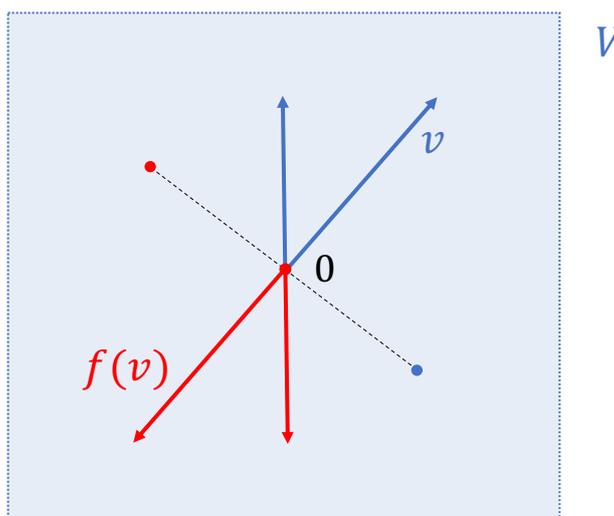


Figura 5.2: Riflessione rispetto all'origine nel piano 2D.

Dimostrazione. Sia \mathcal{R} un riferimento ortonormale di V e sia A la matrice rappresentativa di f nel riferimento \mathcal{R} . Supponiamo che f preservi le orientazioni. In tal caso, A è una matrice ortogonale speciale di ordine 2, perciò è una matrice di rotazione

$$R_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad \theta \in [0, 2\pi).$$

Si osservi che θ può assumere valore arbitrario, perché ogni matrice ortogonale è la matrice rappresentativa di un endomorfismo ortogonale in un fissato riferimento. Se $\theta = 0$, allora $R_\theta = I_2$ e $f = \text{id}_V$. In altre parole siamo nel caso [D2-1].(a). Se $\theta = \pi$, allora $R_\theta = -I_2$, $f = -\text{id}_V$ e siamo nel caso [D2-1].(b). Resta da dimostrare che, se $\theta \neq 0, \pi$, allora f non possiede autovalori. È così, infatti il polinomio caratteristico di f è

$$\begin{aligned} P_f(t) &= \det \begin{pmatrix} \cos \theta - t & -\sin \theta \\ \sin \theta & \cos \theta - t \end{pmatrix} = (\cos \theta - t)^2 + \sin^2 \theta \\ &= t^2 - 2 \cos \theta \cdot t + \cos^2 \theta + \sin^2 \theta = t^2 - 2 \cos \theta \cdot t + 1, \end{aligned}$$

il cui discriminante è

$$\Delta = 4(\cos^2 \theta - 1)$$

che è sempre negativo (eccetto i casi già esclusi $\theta = 0, \pi$).

Supponiamo ora che f inverta le orientazioni. Allora A è una matrice ortogonale con determinante pari a -1 . Abbiamo già osservato che ogni tale matrice è il prodotto di una matrice ortogonale speciale per la matrice

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

Dunque esiste $\theta \in \mathbb{R}$ tale che

$$A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} = \begin{pmatrix} -\cos \theta & \sin \theta \\ \sin \theta & \cos \theta \end{pmatrix},$$

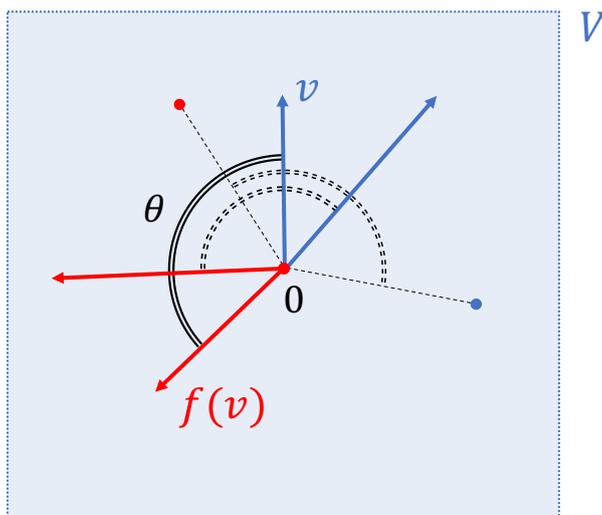


Figura 5.3: Rotazione di angolo θ intorno all'origine nel piano 2D.

che è una matrice simmetrica, dunque f è anche un endomorfismo simmetrico e perciò è ortogonalmente diagonalizzabile. Calcoliamo gli autovalori di f . In questo caso, il polinomio caratteristico di f è

$$\begin{aligned} P_f(t) &= \det \begin{pmatrix} -\cos\theta - t & \sin\theta \\ \sin\theta & \cos\theta - t \end{pmatrix} = (-\cos\theta - t)(\cos\theta - t) - \sin^2\theta \\ &= t^2 - \cos^2\theta - \sin^2\theta = t^2 - 1 = (t+1)(t-1). \end{aligned}$$

Perciò f ha due autovalori distinti $-1, 1$ ed esiste un (altro) riferimento ortonormale nel quale la sua matrice rappresentativa è la (5.6). ■

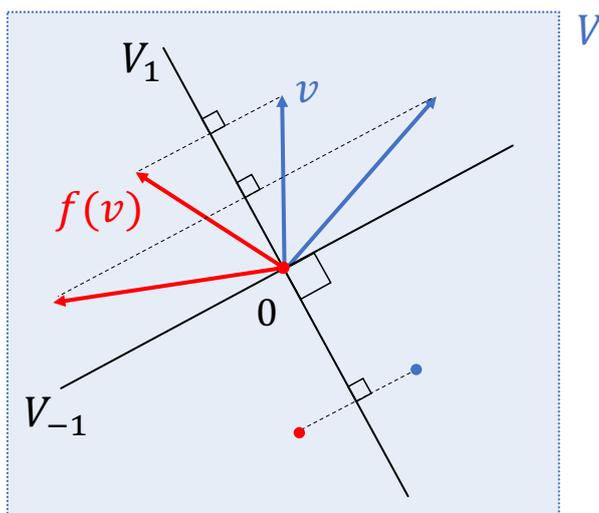


Figura 5.4: Riflessione rispetto ad una retta nel piano 2D.

Consideriamo ora il caso in cui $(V, \langle -, - \rangle)$ è uno spazio vettoriale euclideo 3-dimensionale. Sia $f : (V, \langle -, - \rangle) \rightarrow (V, \langle -, - \rangle)$ un endomorfismo ortogonale e cominciamo ad osservare che, in questo

caso, f possiede almeno un autovalore ± 1 . Infatti, il polinomio caratteristico di f è un polinomio reale di grado 3. Se lo interpretiamo come polinomio complesso, in base al Corollario 4.1.4, le sue radici non reali vengono a coppie di numeri complessi coniugati (con parte immaginaria non nulla). Poiché la somma delle molteplicità delle radici complesse di $P_f(t)$ deve essere comunque pari a 3, $P_f(t)$ deve necessariamente avere almeno una radice reale. Dunque f ha almeno un autovalore. Ciò premesso, abbiamo il seguente

Teorema 5.4.3 Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo di dimensione 3 e sia

$$f : (V, \langle -, - \rangle) \rightarrow (V, \langle -, - \rangle)$$

un endomorfismo ortogonale. Allora

[D3-1] Se f preserva le orientazioni, allora solo uno dei seguenti casi si può presentare:

- (a) $f = \text{id}_V$, in particolare f è ortogonalmente diagonalizzabile con un unico autovalore pari ad 1;
- (b) f è ortogonalmente diagonalizzabile e possiede entrambi gli autovalori $-1, +1$ con molteplicità (algebriche e geometriche) rispettivamente 2 ed 1. In altre parole f è una *riflessione rispetto ad una retta*, l'autospazio V_1 (Figura 5.5);
- (c) f possiede solo un autovalore pari ad 1 con molteplicità algebrica e geometrica pari ad 1, inoltre esiste un riferimento ortonormale nel quale la matrice rappresentativa di f è

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & R_\theta & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}, \quad (5.7)$$

con $\theta \in [0, 2\pi)$ ma $\theta \neq 0, \pi$. In altre parole f è una *rotazione di angolo $\theta \neq 0, \pi$ intorno ad una retta*, l'autospazio V_1 (Figura 5.6).

[D3-2] Se f inverte le orientazioni, allora solo uno dei seguenti casi si può presentare:

- (a) $f = -\text{id}_V$, in particolare f è ortogonalmente diagonalizzabile con un unico autovalore pari a -1 . In altre parole f è la *riflessione rispetto all'origine*;
- (b) f è ortogonalmente diagonalizzabile e possiede entrambi gli autovalori $-1, +1$ con molteplicità (algebriche e geometriche) rispettivamente 1 ed 2. In altre parole f è la *riflessione rispetto ad un piano*, l'autospazio V_1 (Figura 5.7);
- (c) f possiede solo un autovalore pari a -1 con molteplicità algebrica e geometrica pari a -1 , inoltre esiste un riferimento ortonormale nel quale la matrice rappresentativa di f è

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & R_\theta & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}, \quad (5.8)$$

con $\theta \in [0, 2\pi)$ ma $\theta \neq 0, \pi$. In altre parole f è una *rotazione di angolo $\theta \neq 0, \pi$ intorno ad una retta*, l'autospazio V_{-1} , composta con una *riflessione rispetto al piano ortogonale*, l'autospazio V_1 .

Dimostrazione. Sappiamo già che f possiede almeno un autovalore λ . Sia e_1 un autovettore di norma unitaria relativo all'autovalore λ . Consideriamo anche il complemento ortogonale $W := e_1^\perp$. In base al Lemma 5.4.1, W è f -invariante e dunque f si può restringere a W sia nel dominio che nel codominio ottenendo un endomorfismo

$$f_W : W \rightarrow W.$$

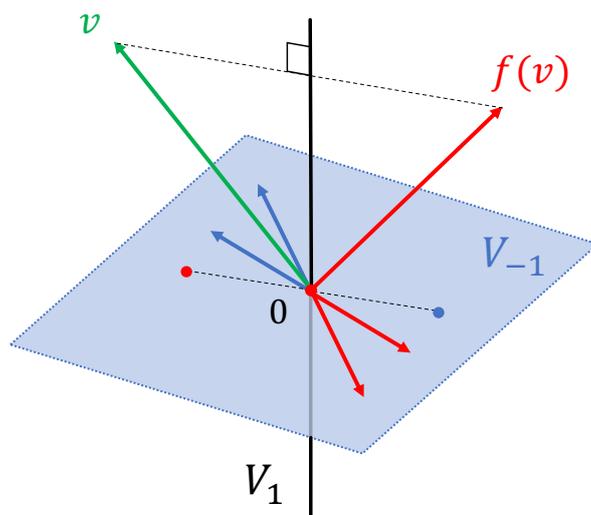


Figura 5.5: Riflessione rispetto ad una retta nello spazio 3D.

È facile dimostrare che f_W è un endomorfismo ortogonale rispetto al prodotto scalare ristretto $\langle -, - \rangle_W$. Siccome $\dim W = 3 - 1 = 2$, per f_W si può presentare solo uno dei 4 casi nel Teorema 5.4.2. Prima di elencare questi casi è opportuna un'osservazione: se (e_2, e_3) è un riferimento ortonormale di W nel quale la matrice rappresentativa di f_W è $B \in O_2$, allora il sistema (e_1, e_2, e_3) è un riferimento ortonormale di $(V, \langle -, - \rangle)$ nel quale la matrice rappresentativa di f è

$$A = \begin{pmatrix} \lambda & 0 \\ 0 & B \end{pmatrix}$$

(il lettore lo vede?). Con questa osservazione in mente elenchiamo tutti i possibili casi per λ e per f_W (quest'ultimo in base al Teorema 5.4.2):

$$\lambda = 1$$

[D2-1]

- (a) In questo caso la matrice B è la matrice identica I_2 , perciò la matrice A è la matrice identica I_3 , cioè $f = \text{id}_V$ e siamo nel caso [D3-1](a).
- (b) In questo caso, la matrice B è $-I_2$, e la matrice A è la matrice

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

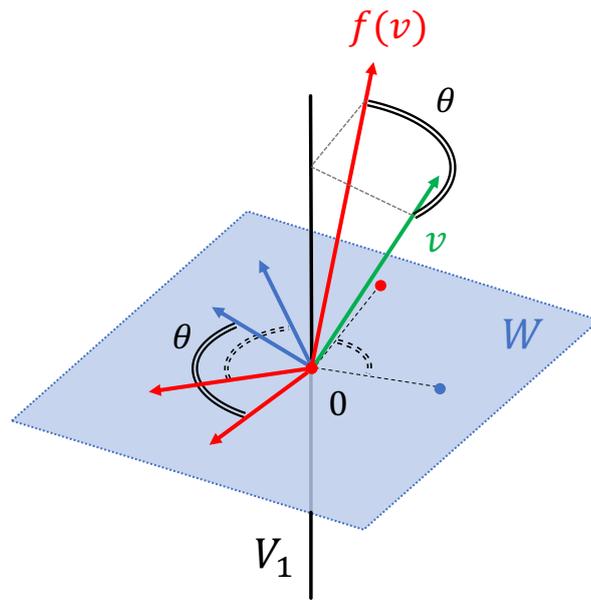
e siamo nel caso [D3-1](b).

- (c) la matrice B è una matrice di rotazione R_θ con $\theta \neq 0, \pi$, perciò la matrice A è la matrice (5.7) e siamo nel caso [D3-1](c).

[D2-2] Possiamo scegliere il riferimento (e_2, e_3) di W in modo che la matrice B sia la matrice (5.6), oppure, scambiando l'ordine dei vettori e_2, e_3 , la matrice opposta. In questo modo la matrice A diventa

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

Siamo nel caso [D3-2](b).

Figura 5.6: Rotazione di angolo θ intorno ad una retta nello spazio 3D.

$$\lambda = -1$$

[D2-1]

(a) La matrice B è la matrice identica I_2 , e la matrice A è la matrice

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

Siamo di nuovo nel caso [D3-2](b).

(b) La matrice B è $-I_2$, e la matrice A è la matrice $-I_3$, dunque $f = -\text{id}_V$ e siamo nel caso [D3-2](a).(c) la matrice B è una matrice di rotazione R_θ con $\theta \neq 0, \pi$, e la matrice A è la matrice (5.8). Siamo nel caso [D3-2](c).[D2-2] Possiamo scegliere il riferimento (e_2, e_3) di W in modo che la matrice B sia la matrice (5.6). In questo modo la matrice A diventa

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

Siamo di nuovo nel caso [D3-1](b).

Questo conclude la dimostrazione. ■

Per completezza enunciamo anche, senza dimostrazione, un *Teorema di Classificazione degli Endomorfismi Ortogonali* di uno spazio vettoriale euclideo di dimensione arbitraria.**Teorema 5.4.4** Sia $(V, \langle -, - \rangle)$ uno spazio vettoriale euclideo di dimensione n e sia

$$f : (V, \langle -, - \rangle) \rightarrow (V, \langle -, - \rangle)$$

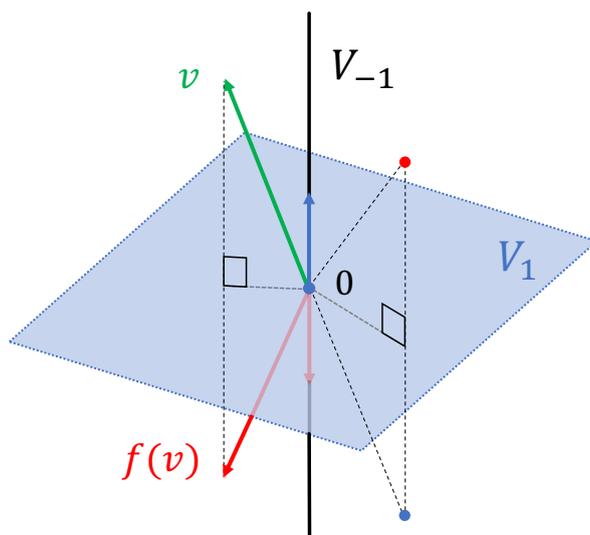


Figura 5.7: Riflessione rispetto ad un piano nello spazio 3D.

un endomorfismo ortogonale. Allora esistono interi non negativi p, q, r (eventualmente nulli) con

$$p + q + 2r = n,$$

ed esiste un riferimento ortonormale \mathcal{R} di $(V, \langle -, - \rangle)$ tale che la matrice rappresentativa di f nel riferimento \mathcal{R} è

$$\begin{pmatrix} I_p & 0 & 0 & \cdots & 0 \\ 0 & -I_q & 0 & \cdots & 0 \\ 0 & 0 & R_{\theta_1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & R_{\theta_r} \end{pmatrix}$$

per qualche $\theta_1, \dots, \theta_r \in [0, 2\pi)$, con $\theta_i \neq 0, \pi$, $i = 1, \dots, r$. In altre parole f è la composizione di una riflessione rispetto all'origine in un sottospazio U di dimensione q e r rotazioni in piani mutuamente ortogonali e ortogonali ad U .

Dimostrazione. La dimostrazione non è complicata e procede per induzione sulla dimensione n di V in modo simile alla dimostrazione del Teorema 5.4.3. In ogni caso, discutere i dettagli va al di là degli scopi di queste note. ■

6. Spazi Affini

Questo Capitolo 6 è il capitolo dal contenuto più geometrico. Gli *spazi affini* sono modelli per il piano e lo spazio 3-dimensionale della Geometria Euclidea. Sono anche modelli per lo spazio fisico nel quale noi ci muoviamo (almeno su piccola scala). Uno spazio affine contiene punti, rette e piani e in uno spazio affine ha senso la nozione di *parallelismo*. Tuttavia, a differenza degli spazi vettoriali, gli spazi affini non posseggono un'origine privilegiata e tutti i punti sono *equivalenti*. Un esempio di spazio affine è lo spazio delle soluzioni di un sistema di equazioni lineari non omogeneo.

6.1 Spazi Affini e Sottospazi Affini

Sia \mathbb{K} un campo.

Definizione 6.1.1 — Spazio Affine. Uno *spazio affine* sul campo \mathbb{K} è una terna (A, V, α) in cui

- (1) A è un insieme non vuoto;
- (2) V è uno spazio vettoriale sul campo \mathbb{K} ;
- (3) α è una funzione

$$\alpha : A \times A \rightarrow V, \quad (P, Q) \mapsto \alpha(P, Q)$$

che gode delle seguenti proprietà:

- (a) Per ogni $P \in A$ la funzione

$$\alpha_P : A \rightarrow V, \quad Q \mapsto \alpha_P(Q) := \alpha(P, Q)$$

è una funzione biettiva. In altre parole, per ogni $P \in A$ e per ogni vettore $v \in V$, esiste un unico $Q \in A$ tale che $v = \alpha(P, Q)$;

- (b) [**proprietà triangolare**] Per ogni $P, Q, R \in A$ si ha che

$$\alpha(P, Q) + \alpha(Q, R) = \alpha(P, R).$$

Gli elementi di A vengono anche detti *punti*. Per ogni coppia di punti P, Q , il vettore $\alpha(P, Q)$ viene anche detto *vettore di separazione* (tra P e Q), oppure *vettore congiungente* (P e Q) e verrà spesso indicato, più semplicemente, con il simbolo \overrightarrow{PQ} . Se $P \in A$ è un punto e $v \in V$ è un vettore, il punto $\alpha_P^{-1}(v)$, cioè l'unico punto Q tale che $\overrightarrow{PQ} = v$ verrà talvolta indicato con il simbolo $P + v$:

$$P + v := \alpha_P^{-1}(v).$$

Lo spazio vettoriale V è detto lo *spazio vettoriale modello*, o anche la *giacitura*, dello spazio affine e si dice anche che A è uno spazio affine *modellato* sullo spazio vettoriale V . La dimensione di V è anche detta *dimensione dello spazio affine* (A, V, α) e si indica anche con $\dim A$. Uno spazio affine di dimensione 1 è detto *retta* (affine) e uno spazio affine di dimensione 2 è detto *piano* (affine).

Nel seguito uno spazio affine (A, V, α) sarà spesso indicato semplicemente con A (per evitare l'eccessivo proliferare di simboli).

Prima di presentare qualche esempio, vogliamo discutere le primissime proprietà elementari degli spazi affini. Osserviamo innanzitutto che, se A è uno spazio affine, applicando più volte la proprietà triangolare, è facile verificare che, per ogni k -upla di punti $P_1, \dots, P_k \in A$ si ha

$$\overrightarrow{P_1 P_k} = \overrightarrow{P_1 P_2} + \overrightarrow{P_2 P_3} + \dots + \overrightarrow{P_{k-1} P_k}.$$

Nel prosieguo, occasionalmente, faremo uso di quest'ultima versione della proprietà triangolare.

Lemma 6.1.1 Sia A uno spazio affine sul campo \mathbb{K} . Allora, per ogni $P, Q \in A$ si ha che

- (1) $\overrightarrow{PP} = 0$,
- (2) $\overrightarrow{QP} = -\overrightarrow{PQ}$,

Dimostrazione. Per la (1), dalla proprietà triangolare,

$$\overrightarrow{PP} = \overrightarrow{PP} + \overrightarrow{PP} \implies \overrightarrow{PP} = 0.$$

Per la (2) calcoliamo

$$\overrightarrow{PQ} + \overrightarrow{QP} = \overrightarrow{PP} = 0,$$

dove abbiamo usato la proprietà triangolare, oltre che la (1). ■

Esercizio 6.1 Dimostrare che, se A è uno spazio affine modellato sullo spazio vettoriale V , allora, per ogni punto $P \in A$ e per ogni $v_1, v_2 \in V$ si ha che

- (1) $P + 0 = P$,
- (2) $P + (v_1 + v_2) = (P + v_1) + v_2$.

(Suggerimento: per la (2), posto $P' = (P + v_1) + v_2$ usare la proprietà triangolare per dimostrare che $\overrightarrow{PP'} = v_1 + v_2$ e dedurre che $P + (v_1 + v_2) = P' = (P + v_1) + v_2$ come richiesto). ■

Ora presentiamo due esempi notevoli.

■ **Esempio 6.1 — Spazio Affine Numerico.** Ogni spazio vettoriale V sul campo \mathbb{K} si può anche riguardare come uno spazio affine modellato sullo spazio vettoriale V stesso, ponendo $A = V$ e $\alpha =$ *funzione differenza*, cioè

$$\alpha : V \times V \rightarrow V, \quad (v, w) \mapsto \overrightarrow{vw} := w - v.$$

Con questa definizione (A, V, α) è effettivamente uno spazio affine sul campo \mathbb{K} . Infatti, per ogni $v \in A = V$, la funzione $\alpha_v : A = V \rightarrow V$ è la funzione data da

$$\alpha_v(w) = \overrightarrow{vw} = w - v, \quad \text{per ogni } w \in A = V,$$

che è invertibile con inversa $\alpha_v^{-1} : V \rightarrow A = V$ data da

$$\alpha_v^{-1}(u) = u + v, \quad \text{per ogni } u \in V \tag{6.1}$$

(il lettore lo vede?). Questo dimostra la proprietà (a). Riguardo alla proprietà triangolare, consideriamo 3 punti $v, w, u \in A = V$. Allora

$$\overrightarrow{vw} = w - v = w - u + u - v = \overrightarrow{vu} + \overrightarrow{uw},$$

come desiderato. Si osservi che la (6.1) mostra che, in questo caso, la notazione $\alpha_v^{-1}(v) = P + v$ è consistente con la notazione per la somma di vettori e non dà luogo ad ambiguità.

Nel caso particolare in cui $V = \mathbb{K}^n$ otteniamo uno spazio affine modellato su \mathbb{K}^n , detto *spazio affine numerico n-dimensionale*, che indicheremo anche con $(A_{\mathbb{K}}^n, \mathbb{K}^n, \alpha_{can})$, o semplicemente $A_{\mathbb{K}}^n$. Come vedremo, ogni spazio affine si identifica con uno spazio affine numerico tramite la scelta di un *riferimento affine*. ■

■ **Esempio 6.2 — Lo Spazio 3D della Geometria Euclidea.** Sia E^3 lo spazio 3-dimensionale della Geometria Euclidea. Questo spazio può essere introdotto in modo rigoroso per via assiomatica mediante gli *assiomi di Hilbert* (si veda, ad esempio, la pagina Wikipedia a riguardo per un brevissimo resoconto: https://it.wikipedia.org/wiki/Assiomi_di_Hilbert). Per motivi di tempo, rinunciamo a presentare tali assiomi e ci accontentiamo dell'intuizione che il lettore ha di E^3 e delle usuali nozioni di punto, retta, piano, etc., incluse le loro principali proprietà, così come vengono presentate, per esempio, nelle lezioni di Matematica a scuola. Vogliamo brevemente mostrare che E^3 è uno spazio affine. In questo esempio ometteremo tutte le dimostrazioni (ma invitiamo il lettore a tentarle). A questo scopo introduciamo, innanzitutto, lo spazio vettoriale su cui lo spazio affine E^3 è modellato. Un *segmento* è ogni porzione di retta *compresa* tra due punti $P, Q \in E^3$. I punti P, Q sono gli *estremi* del segmento e possono anche coincidere. Un *segmento orientato* è un segmento insieme ad una scelta di quale dei due estremi P, Q sia il primo. Per ogni scelta di $P, Q \in E^3$, esiste un unico segmento di estremi P, Q e il segmento orientato di estremi P, Q in cui P è il primo estremo e Q è il secondo si indica con PQ e si rappresenta graficamente mediante una freccia di origine P e di termine Q (Figura 6.1).

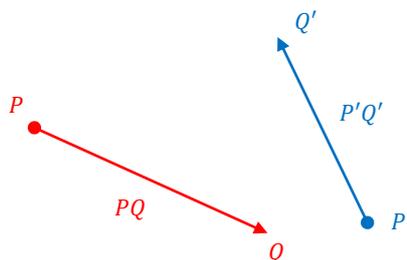


Figura 6.1: Vettori applicati.

Si osservi che la coppia $(P, Q) \in E^3 \times E^3$ contiene un'informazione completa sul segmento orientato PQ . Fissato $O \in E^3$, un segmento orientato con primo estremo O è anche detto *vettore applicato* in O , e O è detto il *punto di applicazione*. Lo spazio V_O^3 dei vettori applicati in O è uno spazio vettoriale reale di dimensione 3:

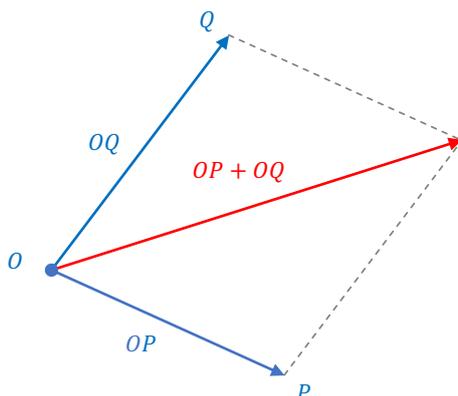


Figura 6.2: Regola del parallelogramma.

- ✓ la somma di due vettori OP, OQ applicati in O è data dalla *regola del parallelogramma* (Figura 6.2);
- ✓ il prodotto di un numero reale $a \in \mathbb{R}$, per il vettore applicato OP , è il vettore aOP applicato in O , tale che
 - (1) aOP è contenuto nella stessa retta di OP ,
 - (2) aOP ha verso uguale od opposto ad OP a seconda che $a > 0$ oppure $a < 0$, mentre $aOP = OO$ se $a = 0$,
 - (3) aOP ha lunghezza pari ad $|a|$ -volte la lunghezza di OP
 (Figura 6.3). Si osservi che questo ha senso anche nel caso in cui $P = O$ (e in questo caso $aOP = aOO = OO$), nonostante siano più di una le rette che passano per un punto O . Il vettore nullo nello spazio vettoriale V^3 è il vettore OO di lunghezza nulla.

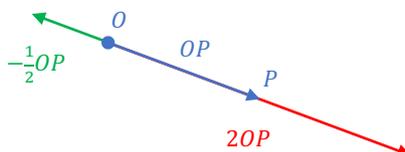


Figura 6.3: Prodotto di uno scalare per un vettore applicato.

Sullo spazio dei segmenti orientati è possibile definire una relazione di equivalenza, detta *equipollenza*, come segue. Due segmenti orientati P_1Q_1 e P_2Q_2 sono *equipollenti* se hanno stessa direzione, stesso verso e stessa lunghezza (Figura 6.4).

La *classe di equipollenza* (cioè la classe di equivalenza rispetto alla relazione di equipollenza) del segmento orientato PQ si indica con \overrightarrow{PQ} . Ogni tale classe di equipollenza è detta *vettore libero* in E^3 o anche *vettore geometrico* in E^3 . Dunque il vettore libero \overrightarrow{PQ} è l'insieme di tutti i segmenti orientati che hanno stessa direzione, stesso verso e stessa lunghezza di PQ , in altre parole che differiscono da PQ al più per il punto di applicazione. Quindi possiamo pensare ad un vettore geometrico come ad un segmento orientato da cui sia stata "rimossa l'informazione sul punto di applicazione". L'insieme V^3 dei vettori liberi è uno spazio vettoriale reale di dimensione 3. Per descrivere le operazioni in V^3 è opportuno innanzitutto osservare che, per ogni segmento orientato PQ e per ogni punto $O \in E^3$, esiste un unico vettore OQ' applicato in O ed equipollente a PQ .

- ✓ Per definire la somma $\overrightarrow{P_1Q_1} + \overrightarrow{P_2Q_2}$ di due vettori liberi $\overrightarrow{P_1Q_1}, \overrightarrow{P_2Q_2}$ si scelga un punto $O \in E^3$ e si considerino i vettori OQ'_1, OQ'_2 applicati in O ed equipollenti a P_1Q_1, P_2Q_2 ; allora

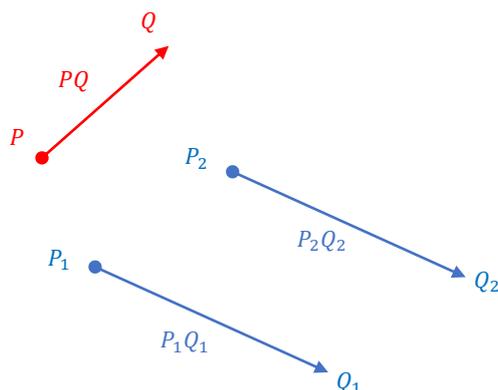


Figura 6.4: I segmenti orientati P_1Q_1, P_2Q_2 sono equipollenti, il segmento orientato PQ non è equipollente agli altri due.

$\overrightarrow{P_1Q_1} + \overrightarrow{P_2Q_2}$ è, per definizione, la classe di equipollenza del vettore applicato $OO'_1 + OO'_2$ (Figura 6.5).

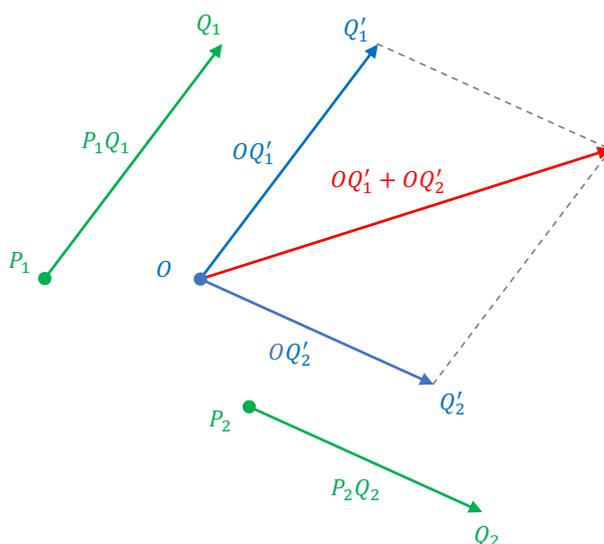


Figura 6.5: Somma di vettori liberi.

Questa definizione è ben posta, nel senso che non dipende dalla scelta del punto O né dai rappresentanti P_1Q_1, P_2Q_2 selezionati nelle classi di equipollenza $\overrightarrow{P_1Q_1}, \overrightarrow{P_2Q_2}$.

- ✓ Per definire il prodotto di un numero reale $a \in \mathbb{R}$, per il vettore libero \overrightarrow{PQ} , procediamo in modo analogo. Scegliamo un punto $O \in E^3$ e consideriamo il vettore OO' applicato in O ed equipollente a PQ ; allora $a\overrightarrow{PQ}$ è, per definizione, la classe di equipollenza del vettore applicato aOO' e questa definizione è ben posta.

Infine consideriamo la funzione

$$\alpha : E^3 \times E^3 \rightarrow V^3, \quad (P, Q) \mapsto \alpha(P, Q) := \overrightarrow{PQ}.$$

Allora la terna (E^3, V^3, α) è uno spazio affine di dimensione 3, detto lo *spazio affine della Geometria Euclidea*. Per esempio, vale la proprietà triangolare (Figura 6.6). In effetti si può considerare questo esempio come prototipo di spazio affine per maturare un'intuizione a riguardo. Si osservi infine che

si possono anche considerare solo segmenti orientati nel piano, nel qual caso la stessa costruzione produce uno spazio affine bidimensionale: il *piano affine della Geometria Euclidea*. ■

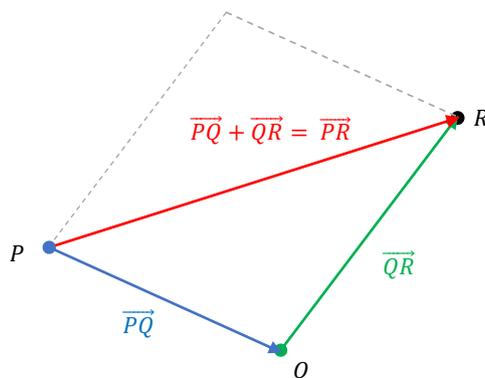


Figura 6.6: Proprietà triangolare.

In analogia con i sottospazi vettoriali, introduciamo ora una classe di sottoinsiemi di uno spazio affine che ereditano la struttura di spazio affine dall'ambiente, i cosiddetti *sottospazi affini*. Sia (A, V, α) uno spazio affine di dimensione n sul campo \mathbb{K} .

Definizione 6.1.2 — Sottospazio Affine. Un sottoinsieme non vuoto $S \subseteq A$ è un *sottospazio affine* di dimensione $m \leq n$ se esiste un sottospazio vettoriale $W \subseteq V$ di dimensione m tale che per ogni $P \in S$ si ha che

$$S = \alpha_P^{-1}(W) = \{Q \in A : \overrightarrow{PQ} \in W\},$$

o, equivalentemente,

$$W = \alpha_P(S) = \{\overrightarrow{PQ} \in V : Q \in S\}. \quad (6.2)$$

In altre parole, per ogni $P \in S$,

- (1) il vettore congiungente \overrightarrow{PQ} appartiene a W per ogni punto $Q \in S$, e
- (2) il punto $P + w$ appartiene ad S per ogni vettore $w \in W$.

In questo caso il sottospazio vettoriale W si dice la *giacitura* di S .

Dalla (6.2) segue immediatamente che un sottospazio affine $S \subseteq A$ ha un'unica giacitura W .

Se $S \subseteq A$ è un sottospazio affine, chiaramente la funzione $\alpha : A \times A \rightarrow V$ si restringe (nel dominio e nel codominio) ad una funzione

$$\alpha^S : S \times S \rightarrow W, \quad (P, Q) \mapsto \alpha^S(P, Q) := \alpha(P, Q) = \overrightarrow{PQ}.$$

Lemma 6.1.2 Sia (A, V, α) uno spazio affine e sia $S \subseteq A$ un sottospazio affine di giacitura W . Allora (S, W, α^S) è uno spazio affine.

Dimostrazione. Poiché α^S agisce come α (sui punti di S) chiaramente α^S soddisfa la proprietà triangolare e resta da dimostrare che, per ogni punto $P \in S$, la funzione

$$\alpha_P^S : S \rightarrow W, \quad Q \mapsto \alpha^S(P, Q) = \overrightarrow{PQ}$$

è biettiva. Poiché α_P^S è la restrizione di α_P ad S nel dominio e a W nel codominio e α_P è iniettiva, segue che anche α_P^S è iniettiva (punti diversi hanno immagini diverse). Infine, dalla (6.2), α_P^S è anche suriettiva: $\alpha_P^S(S) = \alpha_P(S) = W$. ■

Prima di presentare qualche esempio, vogliamo descrivere i sottospazi affini in modo più efficiente. Lo facciamo attraverso il seguente

Lemma 6.1.3 Sia A uno spazio affine di giacitura V , sia $W \subseteq V$ un sottospazio vettoriale, e sia $S \subseteq A$ un sottoinsieme non vuoto. Allora le seguenti affermazioni sono equivalenti:

- (1) S è un sottospazio affine di giacitura W ;
- (2) esiste un punto $P_0 \in A$ tale che $S = \alpha_{P_0}^{-1}(W)$.

Dimostrazione. Poiché S è non vuoto, è chiaro che (1) \Rightarrow (2) (basta scegliere $P_0 \in S$). Ora supponiamo che S sia della forma $S = \alpha_{P_0}^{-1}(W)$ per qualche $P_0 \in A$. Vogliamo dimostrare che, per ogni $P \in S$, abbiamo $S = \alpha_P^{-1}(W)$. Potremmo dimostrare direttamente l'uguaglianza. Per maggiore chiarezza, dimostriamo la doppia inclusione. Innanzitutto, poiché $P \in S = \alpha_{P_0}^{-1}(W)$, allora $\overrightarrow{P_0P} \in W$ e dunque, per ogni altro punto $Q \in S$ abbiamo

$$\overrightarrow{PQ} = \overrightarrow{PP_0} + \overrightarrow{P_0Q} = -\overrightarrow{P_0P} + \overrightarrow{P_0Q} \in W,$$

in cui abbiamo usato la proprietà triangolare e il fatto che $W \subseteq V$ è un sottospazio vettoriale. Questo mostra che $Q \in \alpha_P^{-1}(W)$ e dunque $S \subseteq \alpha_P^{-1}(W)$. Viceversa, supponiamo che $Q \in \alpha_P^{-1}(W)$, cioè $\overrightarrow{PQ} \in W$, allora

$$\overrightarrow{P_0Q} = \overrightarrow{P_0P} + \overrightarrow{PQ} \in W,$$

il che mostra che $Q \in \alpha_{P_0}^{-1}(W) = S$. Dunque $\alpha_P^{-1}(W) \subseteq S$ e questo conclude la dimostrazione. ■

Sia A uno spazio affine di giacitura V , sia $P \in A$ un punto e sia $W \subseteq V$ un sottospazio vettoriale. In base al Lemma 6.1.3.(2), il sottoinsieme $\alpha_P^{-1}(W) \subseteq A$ dei punti $Q \in A$ tali che $\overrightarrow{PQ} \in W$ è un sottospazio affine, anche detto il *sottospazio affine che passa per P ed è parallelo a W* (oppure il *sottospazio affine di giacitura W che passa per P*). Veniamo ora agli esempi.

■ **Esempio 6.3** Sia A uno spazio affine di dimensione n modellato sullo spazio vettoriale V . Ogni sottoinsieme $\{P\} \subseteq A$ che contiene un unico punto $P \in A$ è un sottospazio affine di dimensione 0. Ogni tale sottospazio affine viene anche detto, semplicemente, un *punto*. Il sottoinsieme banale $A \subseteq A$ è un sottospazio affine di dimensione n , detto il *sottospazio affine banale*. I sottospazi affini di dimensione 1 vengono detti *rette* in A , e i sottospazi affini di dimensione 2 vengono detti *piani* in A . ■

■ **Esempio 6.4** Consideriamo lo spazio affine numerico $(A_{\mathbb{K}}^n, \mathbb{K}^n, \alpha_{can})$ di dimensione n sul campo \mathbb{K} . Ricordiamo che $A_{\mathbb{K}}^n = \mathbb{K}^n$ e α_{can} è la funzione differenza. Consideriamo un sistema lineare

$$S : Ax = b$$

di m equazioni in n incognite $x = (x_1, \dots, x_n)^T$ sul campo \mathbb{K} : $A \in M_{m,n}(\mathbb{K})$, $b \in \mathbb{K}^m$. Consideriamo anche il sistema omogeneo S_0 associato ad S ;

$$S_0 : Ax = 0.$$

Vogliamo dimostrare che, se S è un sistema compatibile, allora lo spazio $\text{Sol}(S) \subseteq A_{\mathbb{K}}^n = \mathbb{K}^n$ delle soluzioni di S è un sottospazio affine la cui giacitura è lo spazio $\text{Sol}(S_0) \subseteq \mathbb{K}^n$ delle soluzioni di

S_0 . In altre parole vogliamo dimostrare che, data una soluzione x_0 del sistema S , un altro vettore numerico $x \in A_{\mathbb{K}}^n = \mathbb{K}^n$ è soluzione di S sse $\overrightarrow{x_0x} = x - x_0$ è soluzione di S_0 (in altre parole ancora cioè lo soluzioni di S si possono trovare a partire da una soluzione data x_0 aggiungendo tutte le soluzioni di S_0). Calcoliamo dunque

$$A(x - x_0) = Ax - Ax_0 = Ax - b,$$

da cui

$$x - x_0 \in \text{Sol}(S_0) \Leftrightarrow A(x - x_0) = 0 \Leftrightarrow Ax - b = 0 \Leftrightarrow Ax = b \Leftrightarrow x \in \text{Sol}(S).$$

Il che conferma quanto annunciato. Viceversa, ogni sottospazio affine S di $A_{\mathbb{K}}^n$ è lo spazio delle soluzioni di un sistema lineare S in n incognite sul campo \mathbb{K} . Infatti, sia $S \subseteq A_{\mathbb{K}}^n = \mathbb{K}^n$ un sottospazio affine e sia $W \subseteq \mathbb{K}^n$ la sua giacitura. Sappiamo che W può essere presentato come lo spazio delle soluzioni di un sistema lineare omogeneo

$$S_0 : Ax = 0$$

(per esempio una sua rappresentazione cartesiana nel riferimento canonico): $W = \text{Sol}(S_0)$. Ora, sia x_0 un punto di S . Sappiamo che $S = (\alpha_{can})_{x_0}^{-1}(W)$, cioè un punto $x \in A_{\mathbb{K}}^n = \mathbb{K}^n$ appartiene ad S sse $\overrightarrow{x_0x} = x - x_0 \in W = \text{Sol}(S_0)$ sse

$$A(x - x_0) = 0 \Leftrightarrow Ax - Ax_0 = 0 \Leftrightarrow Ax = Ax_0.$$

Ciò vuol dire che S è lo spazio delle soluzioni del sistema lineare

$$S : Ax = b,$$

dove si è posto $b = Ax_0$. ■

■ **Esempio 6.5** I sottospazi affini nello spazio affine E^3 della Geometria Euclidea sono esattamente i punti, le rette e i piani oltre che tutto E^3 . ■

Sia A uno spazio affine di dimensione n modellato sullo spazio vettoriale V .

Definizione 6.1.3 — Posizione Reciproca di Due Sottospazi. Due sottospazi affini $S_1, S_2 \subseteq A$ di dimensione $m_1, m_2 \leq n$ e giaciture $W_1, W_2 \subseteq V$ si dicono

- ✓ *paralleli* se $W_1 \subseteq W_2$ oppure $W_2 \subseteq W_1$ (a seconda che $m_1 \leq m_2$ oppure $m_2 \leq m_1$);
- ✓ *incidenti* se $S_1 \cap S_2 \neq \emptyset$;
- ✓ *sghebbi* se non sono né paralleli né incidenti.

Il problema di stabilire la *posizione reciproca* di due sottospazi affini S_1, S_2 consiste nello stabilire se S_1, S_2 sono paralleli, incidenti o sghembi (e verrà affrontato in dettaglio nella prossima sezione). Si osservi che due sottospazi affini $S_1, S_2 \subseteq A$ della stessa dimensione sono paralleli sse hanno la stessa giacitura. Si osservi anche che due sottospazi affini possono essere simultaneamente paralleli e incidenti. Questo caso è discusso nella prossima semplice

Proposizione 6.1.4 Sia A uno spazio affine di dimensione n e siano $S_1, S_2 \subseteq A$ sottospazi affini di dimensione $m_1, m_2 \leq n$ e giaciture W_1, W_2 . Le seguenti condizioni sono equivalenti:

- (1) S_1 ed S_2 sono sia paralleli che incidenti;
- (2) $S_1 \subseteq S_2$ oppure $S_2 \subseteq S_1$ (a seconda che $m_1 \leq m_2$ oppure $m_2 \leq m_1$).

Dimostrazione. Supponiamo che S_1 ed S_2 siano sia paralleli che incidenti e sia $P_0 \in S_1 \cap S_2$. Allora $S_1 = \alpha_{P_0}^{-1}(W_1)$ e $S_2 = \alpha_{P_0}^{-1}(W_2)$. Se $m_1 \leq m_2$, allora $W_1 \subseteq W_2$ e $S_1 = \alpha_{P_0}^{-1}(W_1) \subseteq \alpha_{P_0}^{-1}(W_2) = S_2$. Nello stesso modo, se $m_2 \leq m_1$, allora $S_2 \subseteq S_1$.

Viceversa, supponiamo che $S_1 \subseteq S_2$, sicché certamente S_1 ed S_2 sono incidenti. Per dimostrare che sono anche paralleli, scegliamo $P_0 \in S_1$ (dunque anche $P_0 \in S_2$). Allora $W_1 = \alpha_{P_0}(S_1) \subseteq \alpha_{P_0}(S_2) = W_2$. Nello stesso modo, se $S_2 \subseteq S_1$, allora $W_2 \subseteq W_1$. ■

Ora ci concentriamo su sottospazi affini incidenti.

Proposizione 6.1.5 Sia A uno spazio affine di dimensione n e siano $S_1, S_2 \subseteq A$ sottospazi affini di giaciture W_1, W_2 . Se S_1 ed S_2 sono incidenti, allora la loro intersezione $S_1 \cap S_2$ è un sottospazio affine di giacitura $W_1 \cap W_2$.

Dimostrazione. Sia $P_0 \in S_1 \cap S_2$, allora $S_1 = \alpha_{P_0}^{-1}(W_1)$ e $S_2 = \alpha_{P_0}^{-1}(W_2)$ perciò

$$S_1 \cap S_2 = \alpha_{P_0}^{-1}(W_1) \cap \alpha_{P_0}^{-1}(W_2) = \alpha_{P_0}^{-1}(W_1 \cap W_2).$$

Questo conclude la dimostrazione. ■

○ Esiste anche il concetto di “sottospazio affine generato da un sottoinsieme (non vuoto)”: sia A uno spazio affine di giacitura V , e sia $X \subseteq A$ un sottoinsieme non vuoto. Sia

$$\{S_i\}_{i \in I}$$

la famiglia dei sottospazi affini di A che contengono X , e sia

$$\{W_i\}_{i \in I}$$

la famiglia delle loro giaciture. Si osservi che la famiglia $\{S_i\}_{i \in I}$ è certamente non vuota, perché contiene almeno A . Consideriamo

$$S(X) := \bigcap_{i \in I} S_i \subseteq A, \quad \text{e} \quad W(X) := \bigcap_{i \in I} W_i \subseteq V.$$

Allora $S(X)$ è un sottospazio affine di giacitura $W(X)$. Infatti $S(X)$ è non vuoto perché contiene almeno i punti di X . Inoltre, siano $P, Q \in S(X)$, in particolare $P, Q \in S_i$ per ogni $i \in I$. Ne consegue che $\overrightarrow{PQ} \in W_i$ per ogni $i \in I$ e dunque $\overrightarrow{PQ} \in W(X)$. Viceversa se $P \in S(X)$ e $v \in W(X)$, allora $P + v \in S_i$ per ogni $i \in I$ e dunque $P + v \in S(X)$. Il sottospazio affine $S(X)$ prende il nome di *sottospazio affine generato dal sottoinsieme X* ed è il “più piccolo” sottospazio affine che contiene X .

Si osservi che la giacitura $W(X)$ di $S(X)$ si può anche presentare come segue: dato un punto qualsiasi $P_0 \in X$, allora

$$W(X) = \text{Span} \left\{ \overrightarrow{P_0 P} : P \in X \right\} \subseteq V, \quad (6.3)$$

cioè $W(X)$ è il sottospazio generato dai vettori congiungenti P_0 e ogni altro punto $P \in X$. Per dimostrarlo, diciamo W il sottospazio a destra nella (6.3). Consideriamo il sottospazio affine $S = \alpha_{P_0}^{-1}(W)$ di giacitura W che passa per P_0 . Per costruzione $X \subseteq S$ (il lettore lo vede?). Dunque $S(X) \subseteq S$ da cui $W(X) \subseteq W$. Viceversa ogni vettore del tipo $\overrightarrow{P_0 P}$, con $P \in X$, appartiene a W_i per ogni $i \in I$ (perché $P_0, P \in X$ e quindi $P_0, P \in S_i$ per ogni $i \in I$), perciò $\overrightarrow{P_0 P} \in W(X)$ e anche $W \subseteq W(X)$. Concludendo, $W(X) = W$.

6.2 Rappresentazioni di Sottospazi Affini

Similmente a quanto avviene per gli spazi vettoriali, è possibile introdurre coordinate su uno spazio affine n -dimensionale in modo da “rappresentare” i punti mediante n -uple di scalari. Per fare ciò, occorre fissare un *riferimento affine*. Sia A uno spazio affine di dimensione n sul campo \mathbb{K} modellato sullo spazio vettoriale V .

Definizione 6.2.1 — Riferimento Affine. Un *riferimento affine* di A è una coppia (O, \mathcal{R}) , in cui O è un punto di A , e $\mathcal{R} = (v_1, \dots, v_n)$ è un riferimento di V . Il punto O viene anche detto *origine* del riferimento affine (O, \mathcal{R}) .

■ **Esempio 6.6** Lo spazio affine numerico n -dimensionale $A_{\mathbb{K}}^n$ è dotato di un riferimento affine naturale $(O_{can}, \mathcal{R}_{can})$, detto il *riferimento canonico*, in cui l'origine O_{can} è semplicemente il vettore nullo $0 \in A_{\mathbb{K}}^n = \mathbb{K}^n$ e il riferimento \mathcal{R}_{can} è il riferimento canonico (E_1, \dots, E_n) di \mathbb{K}^n . ■

○ C'è un modo alternativo di introdurre i riferimenti affini. Per discuterlo, introduciamo innanzitutto la nozione di *punti indipendenti*. Sia (P_0, \dots, P_k) una $(k+1)$ -upla di punti di A . Le seguenti condizioni su (P_0, \dots, P_k) sono equivalenti:

- (1) per ogni $i = 0, \dots, k$, i k vettori $\overrightarrow{P_i P_0}, \dots, \overrightarrow{P_i P_i}, \dots, \overrightarrow{P_i P_k}$ sono linearmente indipendenti (qui il cappello “ $\widehat{}$ ” indica che il termine sottostante è omissso dalla sequenza);
- (2) esiste $i = 0, \dots, k$ tale che i k vettori $\overrightarrow{P_i P_0}, \dots, \overrightarrow{P_i P_i}, \dots, \overrightarrow{P_i P_k}$ sono linearmente indipendenti.

Infatti, chiaramente (1) \Rightarrow (2), per vedere che (2) \Rightarrow (1), senza ledere la generalità della discussione, assumiamo che i vettori

$$u_1 = \overrightarrow{P_0 P_1}, u_2 = \overrightarrow{P_0 P_2}, \dots, u_k = \overrightarrow{P_0 P_k}$$

siano linearmente indipendenti, e dimostriamo che anche i vettori

$$v_1 = \overrightarrow{P_1 P_0}, v_2 = \overrightarrow{P_1 P_2}, \dots, v_k = \overrightarrow{P_1 P_k}$$

sono linearmente indipendenti. È così, infatti $v_1 = -u_1$ mentre, per ogni $j = 2, \dots, k$

$$v_j = \overrightarrow{P_1 P_j} = \overrightarrow{P_1 P_0} + \overrightarrow{P_0 P_j} = \overrightarrow{P_0 P_j} - \overrightarrow{P_0 P_1} = u_j - u_1.$$

Ora consideriamo una combinazione lineare nulla

$$a_1 v_1 + a_2 v_2 + \dots + a_k v_k = 0$$

dei vettori v_1, v_2, \dots, v_k e dimostriamo che i coefficienti a_1, a_2, \dots, a_k sono necessariamente tutti nulli. Abbiamo

$$\begin{aligned} 0 &= a_1 v_1 + a_2 v_2 + \dots + a_k v_k = -a_1 u_1 + a_2 (u_2 - u_1) + \dots + a_k (u_k - u_1) \\ &= -\left(\sum_{j=1}^k a_j \right) u_1 + a_2 u_2 + \dots + a_k u_k. \end{aligned}$$

Ma i vettori u_1, u_2, \dots, u_k sono linearmente indipendenti, perciò, in particolare $a_2 = \dots = a_k = 0$, e poiché inoltre

$$\sum_{j=1}^k a_j = a_1 + a_2 + \dots + a_k = 0,$$

allora anche $a_1 = 0$. I punti (P_0, \dots, P_k) di A si dicono *indipendenti* se una delle (e quindi entrambe le) condizioni (1) e (2) è verificata. Un *riferimento affine di punti* di A è una k -upla di punti (P_0, \dots, P_k) indipendenti che inoltre *generano* il sottospazio affine banale A , cioè

$$S(\{P_0, \dots, P_k\}) = A$$

(si veda la definizione nell'Osservazione a pag. 143). Sia dunque (P_0, \dots, P_k) un riferimento affine di punti di A , allora, in particolare,

$$W(\{P_0, \dots, P_k\}) = V.$$

Ma, in base all'Osservazione a pag. 143, i vettori $\overrightarrow{P_0 P_1}, \dots, \overrightarrow{P_0 P_k}$ generano $W(\{P_0, \dots, P_k\}) = V$. Siccome sono anche indipendenti, per definizione di punti indipendenti, essi formano un

riferimento di V , in particolare $k = n$. Questo mostra che, se (P_0, \dots, P_n) è un riferimento affine di punti, allora $(P_0, \mathcal{R} = (\overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_n}))$ è un riferimento affine (secondo la Definizione 6.2.1). Viceversa, sia $(O, \mathcal{R} = (v_1, \dots, v_n))$ un riferimento affine di A (secondo la Definizione 6.2.1). Allora la $(n+1)$ -upla di punti $(O, O+v_1, \dots, O+v_n)$ è un riferimento affine di punti, infatti posto $P_i = O+v_i$, per ogni $i = 1, \dots, n$, si ha $\overrightarrow{OP_i} = v_i$, ma (v_1, \dots, v_n) è un riferimento di V per definizione di riferimento affine. Questa discussione mostra che le nozioni di *riferimento affine* (secondo la Definizione 6.2.1) e *riferimento affine di punti* sono in effetti equivalenti, perché si può passare dall'una all'altra (e tornare indietro). Nel seguito adotteremo sempre la nozione di *riferimento affine* (secondo la Definizione 6.2.1) come nozione *primaria*.

Sia (A, V, α) uno spazio affine di dimensione n sul campo \mathbb{K} modellato sullo spazio vettoriale V . Un riferimento affine $(O, \mathcal{R} = (v_1, \dots, v_n))$ di A consente di introdurre coordinate su A come segue: consideriamo l'applicazione

$$c_{(O, \mathcal{R})} : A \rightarrow \mathbb{K}^n, \quad P \mapsto c_{(O, \mathcal{R})}(P) := c_{\mathcal{R}}(\overrightarrow{OP}).$$

In altre parole $c_{(O, \mathcal{R})}$ associa ad un punto P le coordinate del vettore \overrightarrow{OP} nel riferimento \mathcal{R} di V .

Proposizione 6.2.1 L'applicazione $c_{(O, \mathcal{R})} : A \rightarrow \mathbb{K}^n$ è una biezione. Inoltre, per ogni $P, Q \in A$, si ha che

$$c_{(O, \mathcal{R})}(Q) - c_{(O, \mathcal{R})}(P) = c_{\mathcal{R}}(\overrightarrow{PQ}).$$

Dimostrazione. Per definizione di $c_{(O, \mathcal{R})}$ il diagramma

$$\begin{array}{ccc} A & \xrightarrow{c_{(O, \mathcal{R})}} & \mathbb{K}^n \\ & \searrow \alpha_O & \nearrow c_{\mathcal{R}} \\ & & V \end{array}$$

commuta, cioè $c_{(O, \mathcal{R})} = c_{\mathcal{R}} \circ \alpha_O$ (il lettore lo vede?). Poiché sia $c_{\mathcal{R}}$ che α_O sono funzioni biettive, anche la loro composta è una funzione biettiva. Questo dimostra la prima parte dell'asserto. Per la seconda parte calcoliamo

$$c_{\mathcal{R}}(\overrightarrow{PQ}) = c_{\mathcal{R}}(\overrightarrow{PO} + \overrightarrow{OQ}) = -c_{\mathcal{R}}(\overrightarrow{OP}) + c_{\mathcal{R}}(\overrightarrow{OQ}) = c_{(O, \mathcal{R})}(Q) - c_{(O, \mathcal{R})}(P),$$

come desiderato. ■

Definizione 6.2.2 — Coordinate Affini. L'applicazione $c_{(O, \mathcal{R})} : A \rightarrow \mathbb{K}^n$ è detta *applicazione coordinata* associata al riferimento affine (O, \mathcal{R}) . Dato un punto $P \in A$, le entrate $x_1, \dots, x_n \in \mathbb{K}$ del vettore numerico $c_{(O, \mathcal{R})}(P) = (x_1, \dots, x_n)^T \in \mathbb{K}^n$ sono le *coordinate affini* di P nel riferimento affine (O, \mathcal{R}) .

Esplicitando la definizione, le coordinate affini $x_1, \dots, x_n \in \mathbb{K}$ di un punto $P \in A$ nel riferimento affine $(O, \mathcal{R} = (v_1, \dots, v_n))$ sono gli scalari definiti implicitamente da

$$\overrightarrow{OP} = \sum_{i=1}^n x_i v_i$$

(il lettore lo vede?).

■ **Esempio 6.7** Consideriamo lo spazio affine numerico $A_{\mathbb{K}}^n$ e il suo riferimento affine canonico $(O_{can}, \mathcal{R}_{can})$. È chiaro che, in questo caso, l'applicazione coordinata

$$c_{(O_{can}, \mathcal{R}_{can})} : A_{\mathbb{K}}^n = \mathbb{K}^n \rightarrow \mathbb{K}^n$$

è semplicemente la funzione identica, infatti, per ogni punto $P = (P_1, \dots, P_n)^T \in A_{\mathbb{K}}^n = \mathbb{K}^n$,

$$c_{(O_{can}, \mathcal{R}_{can})}(P) = c_{\mathcal{R}_{can}}(\overrightarrow{O_{can}P}) = c_{\mathcal{R}_{can}}(P - 0) = c_{\mathcal{R}_{can}}(P) = (P_1, \dots, P_n)^T = P.$$

In altre parole, le coordinate affini di un punto $P = (P_1, \dots, P_n)$ di $A_{\mathbb{K}}^n = \mathbb{K}^n$ nel riferimento affine canonico sono semplicemente le entrate di P .

Consideriamo ora un altro riferimento affine di $A_{\mathbb{K}}^n$. Più precisamente consideriamo un riferimento affine del tipo (O, \mathcal{R}_{can}) in cui $O = (O_1, \dots, O_n)^T$ è un altro punto di $A_{\mathbb{K}}^n = \mathbb{K}^n$ (non necessariamente l'origine), ma il riferimento \mathcal{R}_{can} di V è ancora il riferimento canonico (E_1, \dots, E_n) . Calcoliamo le coordinate affini di un punto $P = (P_1, \dots, P_n)^T \in A_{\mathbb{K}}^n = \mathbb{K}^n$ nel nuovo riferimento affine (O, \mathcal{R}_{can}) . Abbiamo

$$c_{(O, \mathcal{R}_{can})}(P) = c_{\mathcal{R}_{can}}(\overrightarrow{OP}) = c_{\mathcal{R}_{can}}(P - O) = P - O = (P_1 - O_1, \dots, P_n - O_n)^T.$$

Esercizio 6.2 Sia $(O, \mathcal{R} = (v_1, \dots, v_n))$ un generico riferimento affine dello spazio affine numerico $A_{\mathbb{K}}^n$. I vettori $v_1, \dots, v_n \in \mathbb{K}^n$ sono vettori colonna:

$$v_1 = \begin{pmatrix} v_{11} \\ \vdots \\ v_{n1} \end{pmatrix}, \quad \dots, \quad v_n = \begin{pmatrix} v_{1n} \\ \vdots \\ v_{nn} \end{pmatrix}.$$

Si esprimano le coordinate affini di un punto $P = (P_1, \dots, P_n)^T \in A_{\mathbb{K}}^n = \mathbb{K}^n$ nel riferimento $(O, \mathcal{R} = (v_1, \dots, v_n))$ in termini de

- (1) (le entrate de) il punto $O = (O_1, \dots, O_n)^T \in A_{\mathbb{K}}^n = \mathbb{K}^n$,
- (2) la matrice

$$M = \begin{pmatrix} v_{11} & \cdots & v_{n1} \\ \vdots & \ddots & \vdots \\ v_{n1} & \cdots & v_{nn} \end{pmatrix},$$

- (3) oltre che (delle entrate) di P .

In altre parole si trovi la formula che lega le coordinate di P nel riferimento affine (O, \mathcal{R}) alle coordinate di P nel riferimento affine canonico. ■

Il resto di questa sezione è dedicato ad una descrizione in coordinate dei sottospazi di uno spazio affine. Similmente al caso dei sottospazi vettoriali in uno spazio vettoriale, esistono due modi di rappresentare i sottospazi affini di uno spazio affine: uno piuttosto esplicito mediante le cosiddette *rappresentazioni parametriche*, ed uno più implicito mediante le *rappresentazioni cartesiane*. Risulta più conveniente utilizzare una rappresentazione parametrica o una rappresentazione cartesiana di un sottospazio affine a seconda della situazione.

Teorema 6.2.2 Sia A uno spazio affine di dimensione n sul campo \mathbb{K} modellato sullo spazio vettoriale V , e sia (O, \mathcal{R}) un riferimento affine di A . Inoltre, consideriamo un sottoinsieme $S \subseteq A$ ed un sottospazio vettoriale $W \subseteq V$ di dimensione m . Allora, le seguenti affermazioni sono equivalenti:

- (1) S è un sottospazio affine di giacitura W ;
- (2) Esiste un punto $x_0 \in A_{\mathbb{K}}^n = \mathbb{K}^n$ ed esistono m vettori numerici indipendenti $d_1, \dots, d_m \in \mathbb{K}^n$ tali che

(a) la seguente

$$x = t_1 d_1 + \cdots + t_m d_m, \quad t_1, \dots, t_m \in \mathbb{K} \quad (6.4)$$

è una rappresentazione parametrica del sottospazio W , e inoltre

(b) un punto P di A appartiene ad S sse esistono $t_1, \dots, t_m \in \mathbb{K}$ tali che

$$x = x_0 + t_1 d_1 + \cdots + t_m d_m; \quad (6.5)$$

dove x è la colonna delle coordinate affini di P nel riferimento affine (O, \mathcal{R}) .

(3) Esiste un sistema lineare ridotto di $n - m$ equazioni in n incognite sul campo \mathbb{K} :

$$Ax = b \quad (6.6)$$

tale che

(a) il sistema omogeneo associato $Ax = 0$ è una rappresentazione cartesiana di W , e

(b) un punto $P \in A$ appartiene ad S sse la colonna $x = (x_1, \dots, x_n)^T$ delle coordinate di P nel riferimento affine (O, \mathcal{R}) è soluzione di $Ax = b$.

Dimostrazione. Cominciamo con il dimostrare che (1) \Rightarrow (2). Fissiamo un punto $P_0 \in S$ e un riferimento (w_1, \dots, w_m) in W . Poiché $S = \alpha_{P_0}^{-1}(W)$, un punto P di A appartiene ad S sse il vettore congiungente $\overrightarrow{P_0 P}$ appartiene a W , e questo si verifica sse $\overrightarrow{P_0 P}$ è combinazione lineare dei vettori w_1, \dots, w_m cioè sse esistono $t_1, \dots, t_m \in \mathbb{K}$ tali che

$$\overrightarrow{P_0 P} = t_1 w_1 + \cdots + t_m w_m. \quad (6.7)$$

Ora la (6.7) è verificata sse i due vettori a destra e a sinistra dell'uguale hanno le stesse coordinate nel riferimento \mathcal{R} . Cioè, detto $\tilde{x} \in \mathbb{K}^n$ il vettore coordinato di $\overrightarrow{P_0 P}$ e detti $d_i \in \mathbb{K}^n$ i vettori coordinati di w_i , $i = 1, \dots, m$, la (6.7) è verificata sse

$$\tilde{x} = t_1 d_1 + \cdots + t_m d_m,$$

dove abbiamo anche usato la linearità dell'isomorfismo $c_{\mathcal{R}} : V \rightarrow \mathbb{K}^n$ (il lettore lo vede?). Detta x_0 la colonna delle coordinate affini di P_0 e x la colonna delle coordinate affini di P nel riferimento affine (O, \mathcal{R}) , in virtù della Proposizione 6.2.1, abbiamo $\tilde{x} = x - x_0$. Riassumendo, il punto P di A appartiene ad S sse esistono $t_1, \dots, t_m \in \mathbb{K}$ tali che la colonna x delle coordinate affini di P soddisfa la condizione

$$x - x_0 = t_1 d_1 + \cdots + t_m d_m \quad \Leftrightarrow \quad x = x_0 + t_1 d_1 + \cdots + t_m d_m.$$

Poiché i vettori $w_1, \dots, w_m \in V$ sono linearmente indipendenti, anche i vettori numerici $d_1, \dots, d_m \in \mathbb{K}^n$ sono indipendenti, inoltre è anche chiaro che

$$x = t_1 d_1 + \cdots + t_m d_m, \quad t_1, \dots, t_m \in \mathbb{K}$$

è una rappresentazione parametrica di W . Questo conclude la dimostrazione del fatto che (1) \Rightarrow (2).

Ora supponiamo che valga (2). Diciamo $P_0 \in A$ il punto di coordinate affini $x_0 \in \mathbb{K}^n$ nel riferimento affine (O, \mathcal{R}) e diciamo $w_1, \dots, w_m \in V$ i vettori di coordinate rispettivamente $d_1, \dots, d_m \in \mathbb{K}^n$ nel riferimento \mathcal{R} . Poiché la (6.4) è una rappresentazione parametrica di W , allora il sistema (w_1, \dots, w_m) è un riferimento per W . Se x sono le coordinate di un generico punto $P \in A$, ragionando come sopra (ma a ritroso), vediamo che la (6.5) è verificata sse

$$\overrightarrow{P_0 P} = t_1 w_1 + \cdots + t_m w_m \quad \Leftrightarrow \quad \overrightarrow{P_0 P} \in \text{Span}(w_1, \dots, w_m) = W.$$

L'ultima condizione equivale a dire che $P \in \alpha_{P_0}^{-1}(W)$. Riassumendo $P \in S$ sse $P \in \alpha_{P_0}^{-1}(W)$ e dunque $S = \alpha_{P_0}^{-1}(W)$ è un sottospazio di giacitura W . Questo conclude la dimostrazione che (1) \Leftrightarrow (2).

Ora dimostriamo che (2) \Leftrightarrow (3). Innanzitutto esplicitiamo

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad x_0 = \begin{pmatrix} x_{01} \\ \vdots \\ x_{0n} \end{pmatrix}, \quad d_1 = \begin{pmatrix} d_{11} \\ \vdots \\ d_{n1} \end{pmatrix}, \quad \dots, \quad d_m = \begin{pmatrix} d_{1m} \\ \vdots \\ d_{nm} \end{pmatrix}. \quad (6.8)$$

Consideriamo la condizione (6.5) su un vettore numerico $x \in \mathbb{K}^n$. Essa equivale a dire che il vettore $x - x_0$ è combinazione lineare dei vettori d_1, \dots, d_m . Poiché i vettori colonna d_1, \dots, d_m sono linearmente indipendenti, questo equivale a dire che la matrice

$$\begin{pmatrix} d_{11} & \cdots & d_{1m} & x_1 - x_{01} \\ \vdots & \ddots & \vdots & \vdots \\ d_{n1} & \cdots & d_{nm} & x_n - x_{0n} \end{pmatrix} \quad (6.9)$$

ha rango pari ad m . Ora, nel primo blocco

$$\begin{pmatrix} d_{11} & \cdots & d_{1m} \\ \vdots & \ddots & \vdots \\ d_{n1} & \cdots & d_{nm} \end{pmatrix}$$

esiste un minore M di ordine m diverso da zero. Assumiamo per semplicità che M sia il minore individuato dalle prime m righe (e da tutte le colonne). In altre parole M è il determinante della sottomatrice quadrata

$$\begin{pmatrix} d_{11} & \cdots & d_{1m} \\ \vdots & \ddots & \vdots \\ d_{m1} & \cdots & d_{mm} \end{pmatrix}.$$

Se si tratta di un altro minore si ragiona in modo simile. Per il Teorema degli Orlati, la matrice (6.9) ha rango m sse tutti gli orlati del minore M si annullano cioè sse

$$\begin{vmatrix} d_{11} & \cdots & d_{1m} & x_1 - x_{01} \\ \vdots & \ddots & \vdots & \vdots \\ d_{m1} & \cdots & d_{mm} & x_m - x_{0m} \\ d_{m+1,1} & \cdots & d_{m+1,m} & x_{m+1} - x_{0,m+1} \end{vmatrix} = \dots = \begin{vmatrix} d_{11} & \cdots & d_{1m} & x_1 - x_{01} \\ \vdots & \ddots & \vdots & \vdots \\ d_{m1} & \cdots & d_{mm} & x_m - x_{0m} \\ d_{n,1} & \cdots & d_{nm} & x_n - x_{0,n} \end{vmatrix} = 0. \quad (6.10)$$

Sviluppando i determinanti mediante la regola di Laplace rispetto all'ultima colonna si vede che le equazioni (6.10) formano un sistema di $n - m$ equazioni nelle incognite x_1, \dots, x_n del tipo

$$\begin{cases} A_{11}(x_1 - x_{01}) + \cdots + A_{1n}(x_n - x_{0n}) = 0 \\ \vdots \\ A_{n-m,1}(x_1 - x_{01}) + \cdots + A_{n-m,n}(x_n - x_{0n}) = 0 \end{cases}$$

Che si può riscrivere nella forma

$$\begin{cases} A_{11}x_1 + \cdots + A_{1n}x_n = b_1 \\ \vdots \\ A_{n-m,1}x_1 + \cdots + A_{n-m,n}x_n = b_{n-m} \end{cases}, \quad (6.11)$$

dove si è posto

$$\begin{aligned} b_1 &= A_{11}x_{01} + \cdots + A_{1n}x_{0n} \\ &\vdots \\ b_{n-m} &= A_{n-m,1}x_{01} + \cdots + A_{n-m,n}x_{0n} \end{aligned}$$

Ora la (6.11) è un sistema lineare di $n - m$ equazioni in n incognite le cui soluzioni sono date, per costruzione, in forma parametrica, dalla (6.5). Poiché la generica soluzione dipende da m parametri, le $n - m$ equazioni devono essere indipendenti. Cioè il sistema (6.11) è un sistema ridotto. Riassumendo, le coordinate affini x di un punto $P \in A$ soddisfano la (6.5) sse sono soluzioni del sistema ridotto (6.11). Ragionando nello stesso modo sulla rappresentazione parametrica di W , vediamo anche che il sistema omogeneo associato a (6.11) è una rappresentazione cartesiana per W . Questo mostra che (2) \Rightarrow (3). Viceversa, supponiamo che il sottoinsieme $S \subseteq A$ sia descritto, in coordinate, da un sistema come nel punto (3) dell'enunciato. Risolvendo il sistema, troviamo la soluzione generale nella forma parametrica (6.5), in cui x_0 è una soluzione della (6.6) (quella che si ottiene quando $t_1 = \cdots = t_m = 0$) e d_1, \dots, d_m sono soluzioni indipendenti del sistema omogeneo associato a (6.6) (il lettore lo vede?). Dunque si ha anche che (3) \Leftrightarrow (2), e questo conclude la dimostrazione. ■

Sia A uno spazio affine di dimensione n e giacitura V , sia $(O, \mathcal{R} = (v_1, \dots, v_n))$ un riferimento affine di A , e sia $S \subseteq A$ un sottospazio affine di dimensione m e di giacitura $W \subseteq V$.

Definizione 6.2.3 — Rappresentazioni Parametrica e Cartesiana di un Sottospazio Affine.

(1) Ogni descrizione (6.5) di S come nel Teorema 6.2.2.(2) è detta una *rappresentazione parametrica* di S nel riferimento affine (O, \mathcal{R}) , e si scrive

$$S : x = x_0 + t_1 d_1 + \cdots + t_m d_m, \quad t_1, \dots, t_m \in \mathbb{K}.$$

(2) Ogni sistema lineare (6.6) come nel Teorema 6.2.2.(3) è detto una *rappresentazione cartesiana* di S nel riferimento affine (O, \mathcal{R}) , e si scrive

$$S : Ax = b.$$

Sia

$$S : x = x_0 + t_1 d_1 + \cdots + t_m d_m, \quad t_1, \dots, t_m \in \mathbb{K}, \quad (6.12)$$

una rappresentazione parametrica del sottospazio affine $S \subseteq A$ nel riferimento (O, \mathcal{R}) . Si osservi che, in base al Teorema 6.2.2, il vettore numerico $x_0 \in \mathbb{K}^n$ è la colonna delle coordinate affini di un punto $P_0 \in S$ (nel riferimento affine (O, \mathcal{R})). Inoltre i vettori numerici $d_1, \dots, d_m \in \mathbb{K}^n$ sono i vettori coordinati dei vettori w_1, \dots, w_m di un riferimento della giacitura W di S (nel riferimento \mathcal{R} di V). I vettori w_1, \dots, w_m vengono talvolta detti *vettori direttori* di S , mentre il punto P_0 è detto un *punto base*. Dunque, una rappresentazione parametrica di S è completamente determinata da un punto base e da un sistema di vettori direttori (tramite le loro coordinate). L'uguaglianza vettoriale (6.12) si può anche scrivere "componente a componente", come segue. Supponiamo che x, x_0, d_1, \dots, d_m siano esplicitamente dati dalla (6.8) in termini delle loro coordinate. Allora la (6.12) è equivalente a

$$\begin{cases} x_1 = x_{01} + t_1 d_{11} + \cdots + t_m d_{1m} \\ \vdots \\ x_n = x_{0n} + t_1 d_{n1} + \cdots + t_m d_{nm} \end{cases}, \quad t_1, \dots, t_m \in \mathbb{K}.$$

La matrice

$$D = \begin{pmatrix} d_{11} & \cdots & d_{1m} \\ \vdots & \ddots & \vdots \\ d_{n1} & \cdots & d_{nm} \end{pmatrix}$$

sarà talvolta chiamata la *matrice dei vettori direttori*.

■ **Esempio 6.8** Sia $P_0 \in A$ un punto di coordinate $x_0 \in \mathbb{K}^n$. Un sottospazio affine $S \subseteq A$ passa per il punto P_0 sse ha una rappresentazione cartesiana del tipo

$$S : A(x - x_0) = 0. \quad (6.13)$$

Infatti, sia $S : Ax = b$ una qualunque rappresentazione cartesiana di S . Allora P appartiene ad S sse la colonna x_0 delle coordinate di P è soluzione del sistema $Ax = b$, cioè $Ax_0 = b$. Perciò $S : Ax = Ax_0$ che è equivalente alla (6.13). ■

■ **Esempio 6.9 — Rappresentazione Parametrica di una Retta per Due Punti.** Sia A uno spazio affine di dimensione n sul campo \mathbb{K} , e sia (O, \mathcal{R}) un riferimento affine. Una retta $r \subseteq A$, cioè un sottospazio affine di dimensione 1, ha rappresentazione parametrica del tipo:

$$\begin{cases} x_1 = x_{01} + td_1 \\ \vdots \\ x_n = x_{0n} + td_n \end{cases}, \quad t \in \mathbb{K}.$$

in cui (x_{01}, \dots, x_{0n}) sono le coordinate affini di un punto di r e (d_1, \dots, d_n) sono le coordinate di un vettore direttore.

Consideriamo ora due punti distinti $P_0, P'_0 \in A$ di coordinate affini rispettivamente (x_{01}, \dots, x_{0n}) e $(x'_{01}, \dots, x'_{0n})$. Per i punti P_0, P'_0 passa un'unica retta r : la retta di punto base P_0 e vettore direttore $\overrightarrow{P_0P'_0}$. Poiché il vettore $\overrightarrow{P_0P'_0}$ ha coordinate $(x'_{01} - x_{01}, \dots, x'_{0n} - x_{0n})$, la retta r ha rappresentazione parametrica

$$r : \begin{cases} x_1 = x_{01} + t(x'_{01} - x_{01}) \\ \vdots \\ x_n = x_{0n} + t(x'_{0n} - x_{0n}) \end{cases}, \quad t \in \mathbb{K}.$$

■ **Esempio 6.10 — Rappresentazione Parametrica di un Piano per Tre Punti.** Sia A uno spazio affine di dimensione n sul campo \mathbb{K} , e sia (O, \mathcal{R}) un riferimento affine. Una piano $\Pi \subseteq A$, cioè un sottospazio affine di dimensione 2, ha rappresentazione parametrica del tipo:

$$\begin{cases} x_1 = x_{01} + t'd'_1 + t''d''_1 \\ \vdots \\ x_n = x_{0n} + t'd'_n + t''d''_n \end{cases}, \quad t', t'' \in \mathbb{K}.$$

in cui (x_{01}, \dots, x_{0n}) sono le coordinate di un punto di Π e (d'_1, \dots, d'_n) e (d''_1, \dots, d''_n) sono le coordinate di due vettori direttori di Π .

Consideriamo ora tre punti $P_0, P'_0, P''_0 \in A$ di coordinate affini rispettivamente (x_{01}, \dots, x_{0n}) , $(x'_{01}, \dots, x'_{0n})$ e $(x''_{01}, \dots, x''_{0n})$. Diciamo che P_0, P'_0, P''_0 sono *allineati* se esiste una retta che li contiene tutti e tre. Questo avviene sse i vettori $\overrightarrow{P_0P'_0}$ e $\overrightarrow{P_0P''_0}$ sono linearmente dipendenti. Infatti, se $r \subseteq A$ è una retta che contiene P_0, P'_0, P''_0 e ℓ è la giacitura di r , allora $\overrightarrow{P_0P'_0}, \overrightarrow{P_0P''_0}$ appartengono entrambi

ad ℓ che è un sottospazio di dimensione 1, perciò $\overrightarrow{P_0P'_0}, \overrightarrow{P_0P''_0}$ sono necessariamente linearmente dipendenti. Viceversa, supponiamo che $\overrightarrow{P_0P'_0}, \overrightarrow{P_0P''_0}$ siano linearmente dipendenti. Distinguiamo due casi:

- (1) I vettori $\overrightarrow{P_0P'_0}, \overrightarrow{P_0P''_0}$ sono entrambi nulli. In questo caso $P_0 = P'_0 = P''_0$ ed esiste certamente una retta che li contiene tutti e 3 perché per ogni punto passa almeno una retta (il lettore provi a dimostrare per esercizio quest'ultima asserzione).
- (2) Uno dei due vettori $\overrightarrow{P_0P'_0}, \overrightarrow{P_0P''_0}$, per esempio il primo $\overrightarrow{P_0P'_0}$, è diverso dal vettore nullo. In questo caso, $P'_0 \neq P_0$ e possiamo considerare la retta r che passa per i punti P_0, P'_0 (si veda l'Esempio 6.9). Sia ℓ la giacitura di r . Poiché $\overrightarrow{P_0P''_0}$ dipende linearmente da $\overrightarrow{P_0P'_0}$, allora $\overrightarrow{P_0P''_0} \in \ell$, cioè anche il punto P''_0 appartiene alla retta r .

Supponiamo ora che i punti P_0, P'_0, P''_0 non siano allineati. Dunque i vettori $\overrightarrow{P_0P'_0}, \overrightarrow{P_0P''_0}$ sono linearmente indipendenti e per i punti P_0, P'_0, P''_0 passa un unico piano Π : il piano di punto base P_0 e vettori direttori $\overrightarrow{P_0P'_0}, \overrightarrow{P_0P''_0}$. Poiché i vettori $\overrightarrow{P_0P'_0}$ hanno coordinate rispettivamente $(x'_{01} - x_{01}, \dots, x'_{0n} - x_{0n})$ e $(x''_{01} - x_{01}, \dots, x''_{0n} - x_{0n})$, il piano Π ha rappresentazione parametrica

$$\Pi : \begin{cases} x_1 = x_{01} + t'(x'_{01} - x_{01}) + t''(x''_{01} - x_{01}) \\ \vdots \\ x_n = x_{0n} + t'(x'_{0n} - x_{0n}) + t''(x''_{0n} - x_{0n}) \end{cases}, \quad t', t'' \in \mathbb{K}.$$

■

Nella prossima sezione faremo numerosi esempi di sottospazi affini e delle loro rappresentazioni parametriche e cartesiane. Qui discutiamo le principali proprietà di tali rappresentazioni, con particolare riferimento alle proprietà di incidenza e parallelismo di sottospazi affini.

Cominciamo ad osservare che è possibile passare da una rappresentazione parametrica ad una cartesiana di un sottospazio affine $S \subseteq A$ come indica la dimostrazione del Teorema 6.2.2. Data una rappresentazione parametrica (6.5), possiamo trovare una rappresentazione cartesiana “eliminando i parametri” t_1, \dots, t_m . Questo si può fare in pratica, per esempio, imponendo l'annullarsi degli $n - m$ orlati di un minore non nullo di ordine m nella matrice (6.9). Viceversa, data una rappresentazione cartesiana (6.6), è sufficiente risolvere il sistema trovando così la soluzione generale in forma parametrica.

Adesso discutiamo l'intersezione di due sottospazi affini in termini delle rappresentazioni. Siano $S, T \subseteq A$ sottospazi affini. È facile descrivere l'intersezione $S \cap T$ quando di S e T siano note rappresentazioni cartesiane. Dunque siano

$$S : Ax = b \quad \text{e} \quad T : Bx = c$$

rappresentazioni cartesiane di S e di T . Un punto P di A appartiene all'intersezione $S \cap T$ sse le sue coordinate x risolvono sia il sistema $Ax = b$ che il sistema $Bx = c$. Dunque, S e T sono incidenti sse il sistema

$$\begin{cases} Ax = b \\ Bx = c \end{cases} \quad (6.14)$$

è compatibile. In tal caso, un punto $P \in A$ appartiene ad $S \cap T$ sse x è soluzione del sistema (6.14). Ora, sebbene il sistema (6.14) “rappresenti” in buona misura l'intersezione $S \cap T$, non è detto che sia una rappresentazione cartesiana, infatti non è detto che sia un sistema ridotto. Non è detto, cioè, che le equazioni di (6.14) siano tutte linearmente indipendenti. Eliminando le equazioni dipendenti otteniamo un nuovo sistema

$$Cx = d$$

che è 1) ridotto e 2) equivalente a (6.14) (cioè ha le stesse soluzioni di (6.14)). Dunque il sistema $Cx = b$ è (finalmente) una rappresentazione cartesiana di $S \cap T$.

■ **Esempio 6.11** Sia A uno spazio affine reale 4-dimensionale, e sia (O, \mathcal{R}) un riferimento affine di A . Indichiamo con (x_1, x_2, x_3, x_4) le coordinate affini determinate su A da (O, \mathcal{R}) . Ora consideriamo $S, T \subseteq A$ i sottospazi affini di rappresentazioni coordinate

$$S : \begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ -x_1 - 2x_2 + 2x_3 - x_4 = 1 \end{cases}$$

e

$$T : \begin{cases} -x_2 + 3x_3 = 1 \\ x_1 - x_2 - x_3 + 2x_4 = -1 \end{cases}.$$

Dunque S e T sono sottospazi affini bidimensionali. Studiamo la loro intersezione $S \cap T$. Un punto di A appartiene ad $S \cap T$ sse le sue coordinate (x_1, x_2, x_3, x_4) sono soluzione del sistema

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ -x_1 - 2x_2 + 2x_3 - x_4 = 1 \\ -x_2 + 3x_3 = 1 \\ x_1 - x_2 - x_3 + 2x_4 = -1 \end{cases}. \quad (6.15)$$

La matrice completa di tale sistema è

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ -1 & -2 & 2 & -1 & 1 \\ 0 & -1 & 3 & 0 & 1 \\ 1 & -1 & -1 & 2 & -1 \end{pmatrix}. \quad (6.16)$$

Mentre la matrice incompleta è

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

Non è difficile verificare che queste matrici hanno entrambe rango 3 (il lettore lo verifichi per esercizio). Per il teorema di Rouché-Capelli, quindi, il sistema (6.15) è compatibile e ammette $\infty^{4-3} = \infty^1$ soluzioni. L'intersezione $S \cap T$ è perciò una retta $r \subseteq A$. Una rappresentazione cartesiana di r si trova osservando che la terza riga della matrice (6.16) dipende linearmente dalle prime due (ne è la somma) e quindi la terza equazione del sistema (6.15) dipende dalle prime due. Dunque il sistema (6.15) è equivalente al sistema che si ottiene eliminando la terza equazione cioè:

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ -x_1 - 2x_2 + 2x_3 - x_4 = 1 \\ x_1 - x_2 - x_3 + 2x_4 = -1 \end{cases}. \quad (6.17)$$

Per costruzione (6.17) è un sistema ridotto, ed è anche una rappresentazione cartesiana di r . Per trovare una rappresentazione parametrica di r basta risolvere il sistema (6.17). Per esempio, riducendo a scala mediante l'algoritmo di Gauss, troviamo

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ -x_2 + 3x_3 = 1 \\ -8x_3 + x_4 = -3 \end{cases}.$$

Dunque x_4 può assumere valore arbitrario $t \in \mathbb{R}$ e, in tal caso, dall'ultima equazione

$$x_3 = \frac{3}{8} + \frac{1}{8}t,$$

dalla seconda

$$x_2 = \frac{1}{8} + \frac{3}{8}t,$$

e dalla prima

$$x_1 = -\frac{1}{2} - \frac{3}{2}t,$$

e una rappresentazione parametrica di r è

$$\begin{cases} x_1 = -\frac{1}{2} - \frac{3}{2}t \\ x_2 = \frac{1}{8} + \frac{3}{8}t \\ x_3 = \frac{3}{8} + \frac{1}{8}t \\ x_4 = t \end{cases}, \quad t \in \mathbb{R}.$$

Da cui leggiamo anche che un punto (base) di r è il punto di coordinate affini $(-1/2, 1/8, 3/8, 0)$ e un vettore direttore di r è il vettore di coordinate $(-3/2, 3/8, 3/8, 1)$. ■

Discutiamo adesso in che modo stabilire la posizione reciproca di due sottospazi affini guardando le rispettive rappresentazioni. Sia A uno spazio affine di dimensione n sul campo \mathbb{K} , e sia (O, \mathcal{R}) un riferimento affine di A . Consideriamo due sottospazi affini $S, T \subseteq A$ di dimensione m, p , con $m \leq p$, e giaciture W, U , due loro rappresentazioni parametriche

$$S : x = x_0 + t_1 d_1 + \cdots + t_m d_m, \quad t_1, \dots, t_m \in \mathbb{K}, \quad (6.18)$$

e

$$T : x = y_0 + s_1 h_1 + \cdots + s_p h_p, \quad s_1, \dots, s_p \in \mathbb{K}, \quad (6.19)$$

e anche due loro rappresentazioni cartesiane

$$S : Ax = b, \quad e \quad T : Bx = c.$$

Indichiamo con

$$D = \begin{pmatrix} d_{11} & \cdots & d_{1m} \\ \vdots & \ddots & \vdots \\ d_{n1} & \cdots & d_{nm} \end{pmatrix}, \quad e \quad H = \begin{pmatrix} h_{11} & \cdots & h_{1p} \\ \vdots & \ddots & \vdots \\ h_{n1} & \cdots & h_{np} \end{pmatrix}$$

le matrici dei vettori direttori di S e T , rispettivamente.

Teorema 6.2.3 Sono equivalenti le seguenti affermazioni:

- (1) S e T sono sottospazi affini paralleli;
- (2)

$$\text{rk}(D, H) = p; \quad (6.20)$$

- (3)

$$\text{rk} \begin{pmatrix} A \\ B \end{pmatrix} = n - m; \quad (6.21)$$

Sono equivalenti anche le seguenti affermazioni:

- (4) S e T sono incidenti;

(5)

$$\text{rk}(x_0 - y_0, D, H) = \text{rk}(D, H); \quad (6.22)$$

(6)

$$\text{rk} \begin{pmatrix} A & b \\ B & c \end{pmatrix} = \text{rk} \begin{pmatrix} A \\ B \end{pmatrix}; \quad (6.23)$$

Dimostrazione. La dimostrazione è una semplice applicazione della teoria dei sistemi lineari e del rango. Prima di proporla osserviamo che, in pratica, è possibile stabilire la posizione relativa di due sottospazi guardando solo le rappresentazioni parametriche o cartesiane dei sottospazi, e, più precisamente, le matrici coinvolte, con particolare riguardo al loro rango.

Cominciamo a dimostrare che (1) \Leftrightarrow (2). Poiché $\dim S \leq \dim T$, e quindi $\dim W \leq \dim U$ per ipotesi, allora S e T sono paralleli sse $W \subseteq U$, il che avviene sse i vettori direttori di S appartengono ad U , cioè dipendono linearmente dai vettori direttori di T . Ma i vettori numerici d_1, \dots, d_m (risp. h_1, \dots, h_p) sono appunto vettori coordinati dei vettori direttori di W (risp. U), perciò S e T sono paralleli sse i vettori numerici d_1, \dots, d_m dipendono linearmente da h_1, \dots, h_p . Poiché i vettori h_1, \dots, h_p sono linearmente indipendenti, quest'ultima condizione si verifica sse è verificata la (2).

Ora dimostriamo che (1) \Leftrightarrow (3). Abbiamo già osservato che S e T sono paralleli sse $W \subseteq U$. Ora il sistema $Ax = 0$ (risp. $Bx = 0$) è una rappresentazione cartesiana di W (risp. U) perciò $W \subseteq U$ sse ogni soluzione di $Ax = 0$ è anche soluzione di $Bx = 0$, il che si verifica sse i sistemi

$$\begin{cases} Ax = 0 \\ Bx = 0 \end{cases} \quad \text{e} \quad Ax = 0$$

sono equivalenti, cioè sse ogni equazione di $Bx = 0$ dipende linearmente dalle equazioni di $Ax = 0$. Poiché il sistema $Ax = 0$ è ridotto (dunque $\text{rk} A = n - m$), quest'ultima condizione equivale alla (3).

Ora dimostriamo che (4) \Leftrightarrow (5). Un punto $P \in A$ appartiene ad $S \cap T$ sse le sue coordinate x si possono esprimere contemporaneamente nella forma (6.18) e nella forma (6.19). Dunque esiste un punto in $S \cap T$ sse esistono scalari $s_1, \dots, s_m, t_1, \dots, t_p$ tali che

$$x_0 + t_1 d_1 + \dots + t_m d_m = y_0 + s_1 h_1 + \dots + s_p h_p \quad \Leftrightarrow \quad x_0 - y_0 = -t_1 d_1 - \dots - t_m d_m + s_1 h_1 + \dots + s_p h_p.$$

In altre parole, esiste un punto in $S \cap T$ sse il vettore numerico $x_0 - y_0$ dipende linearmente dai vettori numerici $d_1, \dots, d_m, h_1, \dots, h_p$, il che equivale alla (5) (il lettore lo vede?).

Dimostriamo che (4) \Leftrightarrow (6). Un punto P di A appartiene ad $S \cap T$ sse le sue coordinate affini x sono soluzione del sistema

$$\begin{cases} Ax = b \\ Bx = c \end{cases}. \quad (6.24)$$

Dunque esiste un punto in $S \cap T$ sse il sistema (6.24) è compatibile. Per il Teorema di Rouché-Capelli, quest'ultima condizione è equivalente alla (6).

■

Il Teorema 6.2.3 fornisce un metodo per studiare la posizione reciproca di due sottospazi affini S, T quando di S e T siano note due rappresentazioni parametriche o due rappresentazioni cartesiane. Quando di S è nota una rappresentazione parametrica e di T una rappresentazione cartesiana, si può passare dall'una all'altra, oppure si può ragionare come segue. Utilizzando le stesse notazioni del Teorema 6.2.3 (e di nuovo nell'ipotesi $m \leq p$), S, T sono paralleli sse $W \subseteq U$ il che avviene sse i vettori direttori di S appartengono ad U , il che, a sua volta, avviene sse i vettori numerici d_1, \dots, d_m risolvono il sistema $Bx = 0$ che rappresenta U . È anche possibile trovare condizioni affinché S e T siano incidenti, ma non ne avremo bisogno in questa generalità, e le cercheremo caso per caso.

Corollario 6.2.4 Sia A uno spazio affine di giacitura V e siano $S, T \subseteq A$ due sottospazi affini di giaciture $W, U \subseteq V$. Se $W + U = V$, allora S e T sono necessariamente incidenti e la loro intersezione è un sottospazio affine di dimensione pari a $\dim S + \dim T - \dim A$.

Dimostrazione. Sia $n = \dim A$ e inoltre $m = \dim S = \dim W$ e $p = \dim T = \dim U$. Poiché $W + U = V$, allora $m + p \geq n$. Senza ledere la generalità della dimostrazione, assumiamo $m \leq p$. Consideriamo rappresentazioni parametriche per S e T come nel Teorema 6.2.3. I vettori d_1, \dots, d_m sono vettori coordinati dei vettori direttori di S (che generano W), mentre i vettori h_1, \dots, h_p sono i vettori coordinati dei vettori direttori di T (che generano U). Da $W + U = V$ segue che i vettori $d_1, \dots, d_m, h_1, \dots, h_p$ generano tutto \mathbb{K}^n . Perciò tra di essi ce ne sono almeno n linearmente indipendenti e il rango della matrice (D, H) in (6.20) è esattamente n . Per lo stesso motivo, anche il rango della matrice $(x_0 - y_0, D, H)$ in (6.22) è n . Siccome i due ranghi coincidono, dal Teorema 6.2.3 (punto (5)) segue che S e T sono incidenti. In questo caso, dalla Proposizione 6.1.5, $S \cap T$ è un sottospazio affine di giacitura $W \cap U$ e quindi, per la formula di Grassmann,

$$\begin{aligned} \dim(S \cap T) &= \dim(W \cap U) = \dim W + \dim U - \dim(W + U) = \dim W + \dim U - \dim V \\ &= \dim S + \dim T - \dim A. \end{aligned}$$

■

Faremo altri esempi di applicazione del Teorema 6.2.3 nella prossima sezione.

6.3 Geometria Affine in Dimensione 2 e 3

In questa sezione trattiamo con qualche dettaglio in più gli spazi affini di dimensione 2 e 3 e i loro sottospazi. Cominciamo con il caso bidimensionale. Sia dunque A un piano affine, cioè uno spazio affine di dimensione 2, sul campo \mathbb{K} . Fissiamo una volta per tutte un riferimento affine (O, \mathcal{R}) di A e indichiamo con (x, y) le corrispondenti coordinate affini su A . Le possibili dimensioni per un sottospazio affine $S \subseteq A$ sono 0, 1, 2. Dunque, esclusi i casi banali in cui S è un punto oppure è tutto lo spazio affine A , gli unici sottospazi affini di A sono le rette. Data una retta $r \subseteq A$, una sua rappresentazione parametrica nel riferimento (O, \mathcal{R}) è del tipo

$$r: \begin{cases} x = x_0 + tl \\ y = y_0 + tm \end{cases}, \quad t \in \mathbb{K}, \quad (6.25)$$

dove (x_0, y_0) sono le coordinate affini di un punto di r (punto base), mentre (l, m) sono le coordinate di un vettore direttore di r . In particolare $(l, m) \neq (0, 0)$. Una rappresentazione cartesiana di r è data da un'unica equazione lineare nelle due incognite x, y :

$$r: ax + by = c, \quad (6.26)$$

con $(a, b) \neq (0, 0)$. Per passare dalla rappresentazione parametrica (6.25) ad una rappresentazione cartesiana, possiamo procedere come suggerito dalla dimostrazione del Teorema 6.2.2. In questo caso, la (6.25) dice che il vettore numerico $(x - x_0, y - y_0)^T$ dipende linearmente dal vettore $(l, m)^T$, il che equivale alla condizione

$$\det \begin{pmatrix} x - x_0 & l \\ y - y_0 & m \end{pmatrix} = 0$$

che è già una rappresentazione cartesiana di r (il lettore lo vede?). Viceversa per passare dalla rappresentazione cartesiana (6.26) ad una rappresentazione parametrica, bisogna risolvere l'equazione

(6.26). Per esempio, se $a \neq 0$, allora possiamo attribuire valore arbitrario $t \in \mathbb{K}$ all'incognita y . In tal caso $x = c/a - tb/a$ e una rappresentazione parametrica di r è:

$$\begin{cases} x = \frac{c}{a} - t\frac{b}{a} \\ y = t \end{cases}, \quad t \in \mathbb{K}. \quad (6.27)$$

Dalla (6.27) leggiamo che il punto di coordinate $(c/a, 0)$ appartiene ad r e che il vettore di coordinate $(-b/a, 1)$ è un vettore direttore per r . Si procede in modo simile se $a = 0$, nel qual caso necessariamente $b \neq 0$. Si osservi che, in entrambi i casi, gli scalari $(-b, a)$ sono le coordinate di un vettore direttore di r , infatti il vettore $(-b, a)^T$ è una soluzione della rappresentazione cartesiana $ax + by = 0$ della giacitura di r (il lettore lo vede?).

■ **Esempio 6.12 — Rappresentazione Cartesiana di una Retta per Due Punti nel Piano.** Consideriamo due punti distinti $P_0, P'_0 \in A$ e siano $(x_0, y_0), (x'_0, y'_0)$ le loro coordinate affini. Allora la retta r che passa per P_0 e P'_0 (Figura 6.7) ha rappresentazione parametrica (Esempio 6.9):

$$r: \begin{cases} x = x_0 + t(x'_0 - x_0) \\ y = y_0 + t(y'_0 - y_0) \end{cases}, \quad t \in \mathbb{K},$$

e quindi rappresentazione cartesiana

$$\det \begin{pmatrix} x - x_0 & x'_0 - x_0 \\ y - y_0 & y'_0 - y_0 \end{pmatrix} = 0.$$

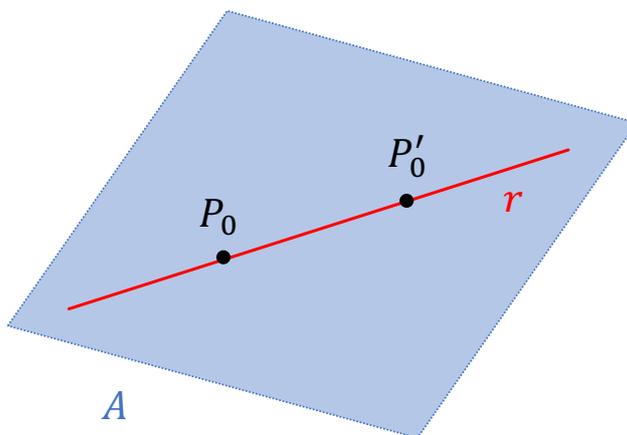


Figura 6.7: Retta r per due punti distinti P_0, P'_0 nel piano.

Consideriamo ora due rette $r, r' \subseteq A$. Siano

$$r: \begin{cases} x = x_0 + tl \\ y = y_0 + tm \end{cases}, \quad t \in \mathbb{K},$$

e

$$r': \begin{cases} x = x'_0 + t'l' \\ y = y'_0 + t'm' \end{cases}, \quad t' \in \mathbb{K}$$

due loro rappresentazioni parametriche, e siano

$$r: ax + by = c \quad \text{e} \quad r': a'x + b'y = c'$$

due loro rappresentazioni cartesiane. In questo caso, il Teorema 6.2.3 si specializza al

Teorema 6.3.1 Le seguenti affermazioni sono equivalenti

- (1) r ed r' sono parallele;
 (2)

$$\det \begin{pmatrix} l & l' \\ m & m' \end{pmatrix} = 0; \quad (6.28)$$

- (3)

$$\det \begin{pmatrix} a & b \\ a' & b' \end{pmatrix} = 0; \quad (6.29)$$

Se le rette r ed r' sono parallele, allora le seguenti affermazioni sono equivalenti

- (4) r ed r' sono anche incidenti, e quindi $r = r'$;
 (5) oltre alla (6.28) vale anche che

$$\det \begin{pmatrix} x_0 - x'_0 & l \\ y_0 - y'_0 & m \end{pmatrix} = 0 \quad \text{o, equivalentemente,} \quad \det \begin{pmatrix} x_0 - x'_0 & l' \\ y_0 - y'_0 & m' \end{pmatrix} = 0;$$

- (6) oltre alla (6.29) vale anche che

$$\text{rk} \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} = 1;$$

Se r ed r' sono parallele ma non coincidono, allora sono parallele e disgiunte. Se r ed r' non sono parallele, allora sono incidenti e si intersecano esattamente in un punto \bar{P} le cui coordinate (\bar{x}, \bar{y}) sono date da:

$$\bar{x} = \frac{\det \begin{pmatrix} c & b \\ c' & b' \end{pmatrix}}{\det \begin{pmatrix} a & b \\ a' & b' \end{pmatrix}}, \quad \bar{y} = \frac{\det \begin{pmatrix} a & c \\ a' & c' \end{pmatrix}}{\det \begin{pmatrix} a & b \\ a' & b' \end{pmatrix}}. \quad (6.30)$$

In particolare, due rette nel piano non sono mai sghembe.

Dimostrazione. Le equivalenze (1) \Leftrightarrow (2) \Leftrightarrow (3) si ottengono semplicemente specializzando la prima parte del Teorema 6.2.3 al caso $n = 2, m = p = 1$. Per esempio, la (6.28) e la (6.29) sono solo un modo alternativo di scrivere le condizioni (2) e (3) nel Teorema 6.2.3 in questo caso particolare (il lettore lo vede?). Analogamente, le equivalenze (4) \Leftrightarrow (5) \Leftrightarrow (6) si ottengono specializzando l'ultima parte del Teorema 6.2.3.

Supponiamo adesso che r ed r' non siano parallele. Allora

$$\det \begin{pmatrix} a & b \\ a' & b' \end{pmatrix} \neq 0 \quad \Rightarrow \quad \text{rk} \begin{pmatrix} a & b \\ a' & b' \end{pmatrix} = 2 \quad \Rightarrow \quad \text{rk} \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} = 2.$$

In particolare, le matrici

$$\begin{pmatrix} a & b \\ a' & b' \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix}$$

hanno lo stesso rango. Ma questa è esattamente la condizione di incidenza proposta al punto (7) del Teorema 6.2.3. Dunque r ed r' si intersecano. La loro intersezione è rappresentata dal sistema

$$r \cap r' : \begin{cases} ax + by = c \\ a'x + b'y = c' \end{cases}$$

che (è compatibile, ammette un'unica soluzione $(\bar{x}, \bar{y})^T$ e) può essere risolto con il metodo di Cramer, il che fornisce esattamente l'espressione (6.30) per la soluzione. ■

Dalle osservazioni che seguono il Teorema 6.2.3, sappiamo che, se di r è nota la rappresentazione parametrica mentre di r' è nota la rappresentazione cartesiana, allora r, r' sono parallele sse $a'l + b'm = 0$ (dove abbiamo utilizzato le stesse notazioni del Teorema 6.3.1). In questo caso, per verificare se $r = r'$ oppure $r \cap r' = \emptyset$ è sufficiente verificare se il punto base di r appartiene oppure no ad r' , cioè se $a'x_0 + b'y_0 = c'$ oppure no (il lettore lo vede?). A questo riguardo si veda l'Esempio 6.14 per un caso studiato in tutti i dettagli.

Le condizioni (6.28) e (6.29) sono anche dette *condizioni di parallelismo* di due rette nel piano. Se r ed r' non sono parallele, allora le coordinate (\bar{x}, \bar{y}) del punto di intersezione si possono anche esprimere in termini delle rappresentazioni parametriche di r ed r' ed invitiamo il lettore a farlo nel prossimo

Esercizio 6.3 Siano r ed r' rette non parallele in un piano affine. Esprimere le coordinate del punto di intersezione di r ed r' in termini delle rispettive rappresentazioni parametriche. ■

Riassumendo, date due rette r, r' in un piano affine, per stabilire la loro posizione reciproca è opportuno, nell'ordine (vedi Figura 6.8),

- (1) stabilire se r ed r' sono parallele (utilizzando una delle condizioni di parallelismo fornite dal Teorema 6.3.1);
- (2) se r ed r' sono parallele, stabilire se coincidono oppure sono parallele e disgiunte (utilizzando una delle condizioni di coincidenza fornite dal Teorema 6.3.1);
- (3) se r ed r' non sono parallele, concludere che allora si intersecano esattamente in un punto (ed eventualmente trovare questo punto, per esempio con la Formula di Cramer).

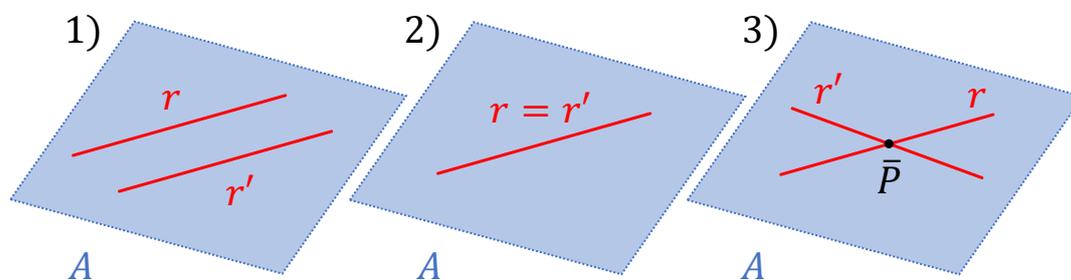


Figura 6.8: Posizioni reciproche di due rette r ed r' in un piano A : 1) r ed r' sono parallele e disgiunte, 2) r ed r' (sono parallele e) coincidono, 3) r ed r' (non sono parallele, ma sono incidenti e) si intersecano esattamente in un punto \bar{P} .

■ **Esempio 6.13** Nel piano affine numerico reale $A_{\mathbb{R}}^2$ consideriamo le rette r, r' di rappresentazioni cartesiane

$$r : 2x + y = 0, \quad r' : x - y = 3,$$

nel riferimento affine canonico. Vogliamo studiare la posizione reciproca di r ed r' . A questo scopo consideriamo il sistema

$$\begin{cases} 2x + y = 0 \\ x - y = 3 \end{cases} \quad (6.31)$$

che descrive l'intersezione $r \cap r'$. Per stabilire se r ed r' sono parallele oppure no, calcoliamo il determinante della matrice incompleta:

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

Dunque r ed r' non sono parallele e si intersecano in un unico punto $\bar{P} \in A_{\mathbb{R}}^2 = \mathbb{R}^2$ che può essere calcolato risolvendo il sistema (6.31), per esempio con la regola di Cramer. Si trova

$$\bar{P} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

■ **Esempio 6.14** Nel piano affine numerico reale $A_{\mathbb{R}}^2$ consideriamo la retta r di rappresentazione parametrica

$$r: \begin{cases} x = -1 + 2t \\ y = 2 - t \end{cases},$$

e la retta r' di rappresentazione cartesiana

$$r': x + 2y = 7$$

nel riferimento affine canonico. Vogliamo studiare la posizione reciproca di r ed r' . Per stabilire se r ed r' sono parallele oppure no osserviamo che

(1) il vettore $v = (2, -1)^T$ è un vettore direttore di r ,

(2) l'equazione $x + 2y = 0$ è una rappresentazione cartesiana per la giacitura W di r' .

Ora r ed r' sono parallele sse hanno la stessa giacitura, e questo si verifica sse un vettore direttore di r appartiene alla giacitura di r' . Per verificare se $v \in W$, dobbiamo verificare se $(2, -1)^T$ è soluzione dell'equazione $x + 2y = 0$. Calcoliamo dunque

$$2 + 2(-1) = 0.$$

Ne ricaviamo che r ed r' sono parallele. Per verificare se coincidono o no è sufficiente verificare se il punto base $(-1, 2)^T$ di r appartiene alla retta r' oppure no. In altre parole dobbiamo verificare se $(-1, 2)^T$ è soluzione dell'equazione $x + 2y = 7$ oppure no. Calcoliamo

$$-1 + 2 \cdot 2 = 3 \neq 7.$$

Dunque r ed r' sono parallele e disgiunte. ■

Esercizio 6.4 Nel piano affine numerico reale $A_{\mathbb{R}}^2$ si considerino le rette r, r' di rappresentazioni parametriche

$$r: \begin{cases} x = -t \\ y = 3 + t \end{cases}, \quad t \in \mathbb{R}, \quad r': \begin{cases} x = 1 + 2t \\ y = 2 - 2t \end{cases}, \quad t \in \mathbb{R}.$$

Stabilire la posizione reciproca di r ed r' . ■

Sia A un piano affine e sia \bar{P} un punto di A . È utile qualche volta considerare l'insieme di tutte le rette che passano per il punto \bar{P} . Tale insieme è detto il *fascio proprio* di rette di centro \bar{P} . Se \bar{P} è presentato come intersezione di due rette non parallele r ed r' , e di r ed r' sono note rappresentazioni cartesiane (vedi Figura 6.9), allora è facile scrivere una rappresentazione cartesiana della generica retta del fascio proprio di centro \bar{P} in base alla seguente

Proposizione 6.3.2 Sia A un piano affine sul campo \mathbb{K} , sia (O, \mathcal{R}) un riferimento affine di A , e siano r ed r' rette non parallele in A di rappresentazioni cartesiane

$$r : ax + by = c, \quad r' : a'x + b'y = c'$$

nel riferimento affine (O, \mathcal{R}) . Detto \bar{P} il punto di intersezione di r ed r' , una retta $r'' \subseteq A$ appartiene al fascio proprio di rette di centro \bar{P} sse r'' ha rappresentazione cartesiana data da

$$r'' : \lambda(ax + by) + \lambda'(a'x + b'y) = \lambda c + \lambda'c' \quad (6.32)$$

per qualche coppia di scalari $\lambda, \lambda' \in \mathbb{K}$, con $(\lambda, \lambda') \neq (0, 0)$.

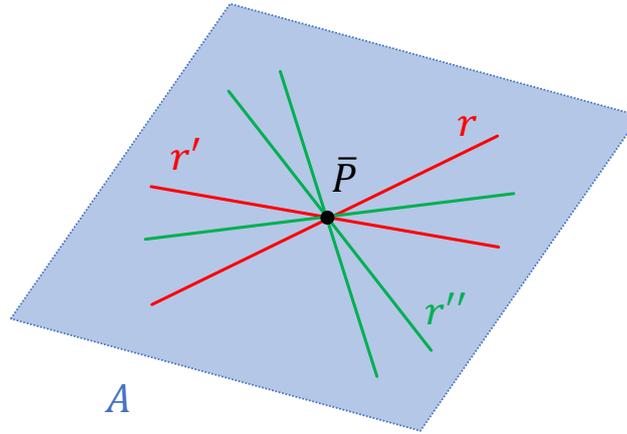


Figura 6.9: Fascio proprio di rette di centro \bar{P} nel piano.

Dimostrazione. Indichiamo con (\bar{x}, \bar{y}) le coordinate affini del punto \bar{P} nel riferimento affine (O, \mathcal{R}) . Poiché \bar{P} appartiene sia ad r che ad r' si ha

$$a\bar{x} + b\bar{y} = c, \quad \text{e} \quad a'\bar{x} + b'\bar{y} = c'.$$

Perciò, per ogni $\lambda, \lambda' \in \mathbb{K}$, si ha che

$$\lambda(a\bar{x} + b\bar{y}) + \lambda'(a'\bar{x} + b'\bar{y}) = \lambda c + \lambda'c',$$

cioè \bar{P} appartiene anche alla retta r'' di rappresentazione cartesiana

$$r'' : \lambda(ax + by) + \lambda'(a'x + b'y) = \lambda c + \lambda'c'.$$

In altre parole, r'' appartiene al fascio proprio di rette di centro \bar{P} . Viceversa, consideriamo una retta r'' appartenente al fascio proprio di rette di centro \bar{P} . Sia

$$r'' : a''x + b''y = c''$$

una rappresentazione cartesiana di r'' nel riferimento (O, \mathcal{R}) . Vogliamo dimostrare che esistono $\lambda, \lambda' \in \mathbb{K}$ non entrambi nulli tali che

$$(a'', b'', c'') = \lambda(a, b, c) + \lambda'(a', b', c'). \quad (6.33)$$

A questo scopo consideriamo il sistema

$$S : \begin{cases} ax + by = c \\ a'x + b'y = c' \\ a''x + b''y = c'' \end{cases}.$$

È chiaro che $r \cap r' \cap r'' = r \cap r' = \{\bar{P}\}$ perciò il sistema S ammette un'unica soluzione $(\bar{x}, \bar{y})^T$ ed è equivalente al sistema

$$r \cap r' : \begin{cases} ax + by = c \\ a'x + b'y = c' \end{cases}$$

che rappresenta $r \cap r'$. Ma allora, la terza equazione in S deve dipendere linearmente dalle altre due, il che è equivalente alla (6.33). Questo conclude la dimostrazione. ■

■ **Esempio 6.15** In un piano affine A sul campo \mathbb{K} , con un riferimento affine (O, \mathcal{R}) consideriamo un punto P di coordinate affini (\bar{x}, \bar{y}) . È chiaro che \bar{P} è l'intersezione delle rette r, r' di rappresentazioni cartesiane

$$r : x = \bar{x}, \quad r' : y = \bar{y}.$$

Perciò una retta r'' passa per \bar{P} (cioè appartiene al fascio proprio di rette di centro \bar{P}) sse esistono $\lambda, \lambda' \in \mathbb{K}$ (non entrambi nulli) tali che

$$r'' : \lambda x + \lambda' y = \lambda \bar{x} + \lambda' \bar{y},$$

che si può anche riscrivere nella forma

$$r'' : \lambda(x - \bar{x}) + \lambda'(y - \bar{y}) = 0.$$

■

■ **Esempio 6.16** L'equazione del fascio (Equazione (6.32)) torna utile quando si voglia determinare una rappresentazione cartesiana di una retta che passa per un punto \bar{P} (assegnato come intersezione di due rette non parallele) e soddisfa condizioni aggiuntive. Per esempio, nel piano affine numerico reale $A_{\mathbb{R}}^2$ (con il riferimento affine canonico) consideriamo le rette

$$r : -x + 3y = 5, \quad e \quad r' : 2x + 4y = -1.$$

È chiaro che r ed r' non sono parallele (il lettore lo vede?). Perciò si intersecano in un unico punto \bar{P} (il lettore trovi \bar{P} per esercizio). Consideriamo anche il punto $Q = (1, 0)^T$. Poiché $Q \notin r$, allora $\bar{P} \neq Q$, perciò esiste un'unica retta r'' che passa per \bar{P} e Q . Anche se non conosciamo le coordinate di \bar{P} , possiamo determinare una rappresentazione cartesiana di r'' come segue. La retta r'' appartiene al fascio proprio di rette di centro \bar{P} , perciò ogni sua rappresentazione cartesiana è del tipo:

$$r'' : \lambda(-x + 3y) + \lambda'(2x + 4y) = 5\lambda - \lambda', \tag{6.34}$$

con $\lambda, \lambda' \in \mathbb{R}$. Vogliamo determinare λ, λ' in modo tale che $Q \in r''$. Chiaramente $(1, 0)$ deve essere soluzione dell'Equazione (6.34), cioè

$$-\lambda + 2\lambda' = 5\lambda - \lambda' \quad \Rightarrow \quad \lambda' = 2\lambda.$$

Per esempio, se poniamo $\lambda = 1$, deve essere $\lambda' = 2$. Da cui

$$r'' : 3x + 11y = 3.$$

Si osservi che possiamo scegliere per λ un qualunque valore diverso da 0 (purchè poi si ponga $\lambda' = 2\lambda$). Al variare di λ otteniamo tutte le (diverse) rappresentazioni cartesiane della retta r'' cercata. ■

Dato un piano affine A di giacitura V ed un sottospazio vettoriale $\ell \subseteq V$ di dimensione 1, qualche volta è anche utile considerare l'insieme di tutte le rette affini in A che hanno giacitura pari ad ℓ (e quindi sono tutte parallele tra loro). Tale insieme è detto *fascio improprio* di rette di giacitura ℓ . Data una rappresentazione parametrica (risp. cartesiana) di ℓ , è facile scrivere una rappresentazione parametrica (risp. cartesiana) della generica retta del fascio improprio di giacitura ℓ (vedi Figura 6.10) in base alla

Proposizione 6.3.3 Sia A un piano affine sul campo \mathbb{K} modellato sullo spazio vettoriale V , sia (O, \mathcal{R}) un riferimento affine di A , e sia $\ell \subseteq V$ un sottospazio vettoriale di dimensione 1. Inoltre siano

$$\ell : \begin{cases} x = tl \\ y = tm \end{cases}, \quad t \in \mathbb{K},$$

e

$$\ell : ax + by = 0$$

una rappresentazione parametrica e una cartesiana di ℓ nel riferimento \mathcal{R} . Allora, data una retta $r \subseteq A$, le seguenti affermazioni sono equivalenti:

- (1) r appartiene al fascio improprio di rette di giacitura ℓ ;
- (2) r ha la seguente rappresentazione parametrica nel riferimento affine (O, \mathcal{R})

$$r : \begin{cases} x = x_0 + tl \\ y = y_0 + tm \end{cases}, \quad t \in \mathbb{K},$$

per qualche coppia di scalari $x_0, y_0 \in \mathbb{K}$;

- (3) r ha la seguente rappresentazione cartesiana nel riferimento affine (O, \mathcal{R})

$$r : ax + by = c$$

per qualche scalare $c \in \mathbb{K}$.

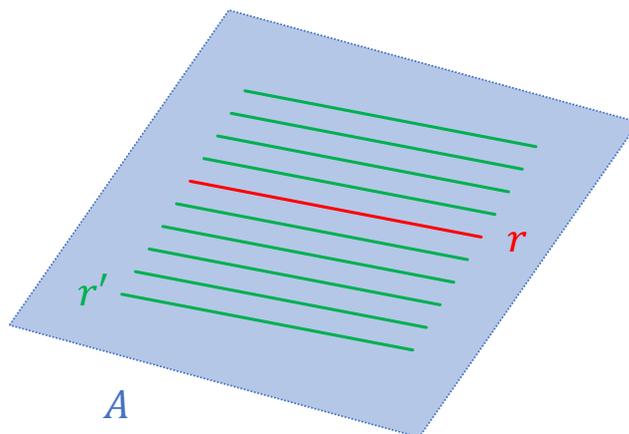


Figura 6.10: Fascio improprio di rette nel piano.

Dimostrazione. La dimostrazione è ovvia e i dettagli sono lasciati al lettore per esercizio. ■

■ **Esempio 6.17** Nel piano affine A consideriamo una retta $r_0 \subseteq A$ di rappresentazione parametrica

$$r_0 : \begin{cases} x = x_0 + tl \\ y = y_0 + tm \end{cases}, \quad t \in \mathbb{K},$$

e rappresentazione cartesiana

$$r_0 : ax + by = c.$$

Ogni retta r parallela ad r_0 ha una rappresentazione parametrica del tipo

$$r : \begin{cases} x = \bar{x}_0 + tl \\ y = \bar{y}_0 + tm \end{cases}, \quad t \in \mathbb{K},$$

e una rappresentazione cartesiana del tipo

$$r : ax + by = \bar{c}.$$

■ **Esempio 6.18** Nel piano affine numerico reale $A_{\mathbb{R}}^2$ consideriamo la retta

$$r_0 : 6x - 7y = 3,$$

e il punto $P = (1, 1)^T$. Determiniamo la retta r parallela ad r_0 che passa per il punto P . Affinchè sia parallela ad r_0 , la retta r deve avere una rappresentazione cartesiana del tipo

$$r : 6x - 7y = \bar{c}, \tag{6.35}$$

e si tratta solo di determinare $\bar{c} \in \mathbb{R}$ in modo tale che $P \in r$, cioè $(1, 1)^T$ sia soluzione della (6.35). Deve essere

$$6 - 7 = \bar{c} \quad \Rightarrow \quad \bar{c} = -1.$$

Perciò

$$r : 6x - 7y = -1.$$

Esercizio 6.5 Nel piano affine numerico reale $A_{\mathbb{R}}^2$ si considerino le rette

$$r : 2x - 3y = -2, \quad r' : x + 5y = 1, \quad r'' : x - y = 0.$$

Dopo aver verificato che r ed r' non sono parallele, determinare una rappresentazione cartesiana e una parametrica per la retta \bar{r} passante per il punto P di intersezione tra r ed r' e parallela alla retta r'' .

Passiamo ora a studiare i sottospazi affini in uno spazio affine tridimensionale. Sia dunque A uno spazio affine tridimensionale sul campo \mathbb{K} . Fissiamo una volta per tutte un riferimento affine (O, \mathcal{R}) di A e indichiamo con x, y, z le corrispondenti coordinate affini su A . Le possibili dimensioni per un sottospazio affine $S \subseteq A$ sono 0, 1, 2, 3. Dunque, esclusi i casi banali in cui S è un punto oppure è tutto lo spazio affine A , gli unici sottospazi affini di A sono i piani e le rette. Dato un piano $\Pi \subseteq A$, una sua rappresentazione parametrica è del tipo:

$$\Pi : \begin{cases} x = x_0 + sl + tl' \\ y = y_0 + sm + tm' \\ z = z_0 + sn + tn' \end{cases}, \quad s, t \in \mathbb{K}, \tag{6.36}$$

dove (x_0, y_0, z_0) sono le coordinate affini di un punto di Π , mentre $(l, m, n)^T, (l', m', n')^T$ sono i vettori coordinati di due vettori direttori di Π . In particolare $(l, m, n)^T, (l', m', n')^T$ sono linearmente indipendenti. Una rappresentazione cartesiana di Π è data da un'unica equazione lineare nelle tre incognite x, y, z :

$$\Pi : ax + by + cz = d, \quad (6.37)$$

con $(a, b, c) \neq (0, 0, 0)$. Per passare dalla rappresentazione parametrica (6.36) ad una rappresentazione cartesiana osserviamo che la (6.36) dice che il vettore numerico $(x - x_0, y - y_0, z - z_0)$ dipende linearmente dai vettori $(l, m, n)^T, (l', m', n')^T$, il che si può esprimere nella forma

$$\det \begin{pmatrix} x - x_0 & l & l' \\ y - y_0 & m & m' \\ z - z_0 & n & n' \end{pmatrix} = 0, \quad (6.38)$$

che è già una rappresentazione cartesiana di Π .

○ Si osservi che, in virtù del Primo Teorema di Laplace, i coefficienti delle incognite x, y, z nella rappresentazione cartesiana (6.38) sono rispettivamente i minori

$$\begin{vmatrix} m & m' \\ n & n' \end{vmatrix}, - \begin{vmatrix} l & l' \\ n & n' \end{vmatrix}, \begin{vmatrix} l & l' \\ m & m' \end{vmatrix}$$

della matrice

$$\begin{pmatrix} l & l' \\ m & m' \\ n & n' \end{pmatrix}$$

(presi a segni alterni). Questo dice che, nel caso particolare in cui $\mathbb{K} = \mathbb{R}$, il vettore $(a, b, c)^T$ dei coefficienti delle incognite nella (6.38) è esattamente il prodotto vettoriale dei vettori numerici $(l, m, n)^T, (l', m', n')^T$ fatto rispetto al prodotto scalare standard e all'orientazione standard nello spazio vettoriale euclideo standard $(\mathbb{R}^3, \langle -, - \rangle_{can})$:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} l \\ m \\ n \end{pmatrix} \wedge \begin{pmatrix} l' \\ m' \\ n' \end{pmatrix}.$$

Viceversa, per passare dalla rappresentazione cartesiana (6.37) ad una rappresentazione parametrica, bisogna risolvere l'equazione (6.37). Per esempio, se $a \neq 0$, allora possiamo attribuire valori arbitrari $s, t \in \mathbb{K}$ alle incognite y, z . In tal caso $x = d/a - sb/a - tc/a$ e una rappresentazione parametrica di Π è:

$$\begin{cases} x = \frac{d}{a} - s\frac{b}{a} - t\frac{c}{a} \\ y = s \\ z = t \end{cases}, \quad s, t \in \mathbb{K}. \quad (6.39)$$

Dalla (6.39) leggiamo che, in questo caso, il punto di coordinate $(d/a, 0, 0)$ appartiene a Π e che i vettori di coordinate $(-b/a, 1, 0)$ e $(-c/a, 0, 1)$ sono vettori direttori di Π . Si può procedere in modo analogo se $b \neq 0$ oppure $c \neq 0$.

■ **Esempio 6.19 — Rappresentazione Cartesiana di un Piano per Tre Punti nello Spazio 3D.** Consideriamo tre punti non allineati $P_0, P'_0, P''_0 \in A$ e siano $(x_0, y_0, z_0), (x'_0, y'_0, z'_0), (x''_0, y''_0, z''_0)$ le loro

coordinate affini. Allora il piano Π che passa per P_0, P'_0, P''_0 (Figura 6.11) ha rappresentazione parametrica (Esempio 6.10):

$$\pi : \begin{cases} x = x_0 + s(x'_0 - x_0) + t(x''_0 - x_0) \\ y = y_0 + s(y'_0 - y_0) + t(y''_0 - y_0) \\ z = z_0 + s(z'_0 - z_0) + t(z''_0 - z_0) \end{cases}$$

e quindi rappresentazione cartesiana

$$\det \begin{pmatrix} x - x_0 & x'_0 - x_0 & x''_0 - x_0 \\ y - y_0 & y'_0 - y_0 & y''_0 - y_0 \\ z - z_0 & z'_0 - z_0 & z''_0 - z_0 \end{pmatrix} = 0.$$

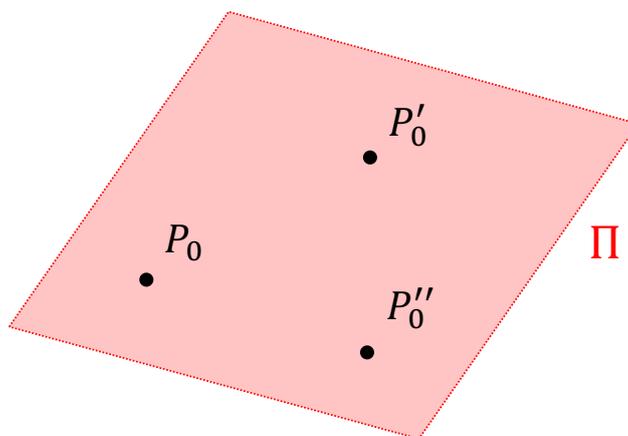


Figura 6.11: Piano per 3 punti non allineati in uno spazio affine 3D.

Ora consideriamo una retta $r \subseteq A$. Una rappresentazione parametrica di r è del tipo

$$r : \begin{cases} x = x_0 + tl \\ y = y_0 + tm \\ z = z_0 + tn \end{cases}, \quad t \in \mathbb{K} \quad (6.40)$$

dove (x_0, y_0, z_0) sono le coordinate affini di un punto di r e (l, m, n) sono le coordinate di un vettore direttore di r . In particolare $(l, m, n) \neq (0, 0, 0)$. Una rappresentazione cartesiana di r è data da un sistema di due equazioni lineari nelle 3 incognite x, y, z :

$$r : \begin{cases} ax + by + cz = d \\ a'x + b'y + c'z = d' \end{cases}, \quad (6.41)$$

in cui la matrice incompleta

$$\begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix}$$

(e quindi anche la matrice completa) ha rango 2, cioè i vettori riga (a, b, c) e (a', b', c') sono linearmente indipendenti. Si osservi che ciascuna delle due equazioni nel sistema (6.41) è la rappresentazione cartesiana di un piano e quindi la (6.41) presenta anche la retta r come intersezione di due piani. Per passare dalla rappresentazione parametrica (6.40) ad una rappresentazione

cartesiana ragioniamo di nuovo come nella dimostrazione del Teorema 6.2.2: la (6.41) dice che il vettore numerico $(x-x_0, y-y_0, z-z_0)$ dipende linearmente dal vettore $(l, m, n)^T$. Poiché $(l, m, n) \neq (0, 0, 0)$, questo equivale a dire che il rango della matrice

$$\begin{pmatrix} x-x_0 & l \\ y-y_0 & m \\ z-z_0 & n \end{pmatrix} \quad (6.42)$$

è 1. In altre parole i tre minori di ordine 2 della matrice (6.42) si annullano. Otteniamo così tre equazioni lineari nelle 3 incognite x, y, z . Di queste tre equazioni, solo 2 sono indipendenti e costituiscono la rappresentazione cartesiana di r che stiamo cercando. Per esempio, se $l \neq 0$, in virtù del Teorema degli Orlati, il rango della matrice (6.42) è pari a 2 sse si annullano i due orlati dell'entrata l . In questo caso, perciò, una rappresentazione cartesiana di r è data da

$$r: \begin{cases} \begin{vmatrix} x-x_0 & l \\ y-y_0 & m \end{vmatrix} = 0 \\ \begin{vmatrix} x-x_0 & l \\ z-z_0 & n \end{vmatrix} = 0 \end{cases}.$$

Se $m \neq 0$, o $n \neq 0$, si ragiona in modo simile. Viceversa, per passare dalla rappresentazione cartesiana (6.41) ad una rappresentazione parametrica, bisogna risolvere il sistema (6.41). Come risolvere questo sistema dipende dai valori dei coefficienti. In ogni caso possiamo dire che il vettore di coordinate

$$(X, Y, Z) := \left(\begin{vmatrix} b & c \\ b' & c' \end{vmatrix}, - \begin{vmatrix} a & c \\ a' & c' \end{vmatrix}, \begin{vmatrix} a & b \\ a' & b' \end{vmatrix} \right)$$

è un vettore direttore di r . Per vedere ciò, si osservi che $(X, Y, Z)^T$ è una soluzione del sistema

$$\begin{cases} ax + by + cz = 0 \\ a'x + b'y + c'z = 0 \end{cases}$$

omogeneo associato a (6.41) (che è una rappresentazione cartesiana per la giacitura di r). Per esempio $(X, Y, Z)^T$ risolve la prima equazione. Infatti, per il Primo Teorema di Laplace

$$aX + bY + cZ = \det \begin{pmatrix} a & b & c \\ a & b & c \\ a' & b' & c' \end{pmatrix} = 0,$$

dove abbiamo usato che (X, Y, Z) sono esattamente i complementi algebrici degli elementi nella prima riga della matrice a secondo membro o, equivalentemente, i minori di ordine 2 della matrice

$$\begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix}$$

(presi a segni alterni). Analogamente si vede che $(X, Y, Z)^T$ risolve la seconda equazione.

Studiamo ora in dettaglio le posizioni reciproche rispettivamente di *due piani, una retta e un piano, due rette*, in uno spazio affine tridimensionale A .

Consideriamo dunque, innanzitutto, due piani $\Pi, \Pi' \subseteq A$. Siano

$$\Pi: \begin{cases} x = x_0 + sl + tl' \\ y = y_0 + sm + tm' \\ z = z_0 + sn + tn' \end{cases}, \quad s, t \in \mathbb{K}, \quad (6.43)$$

e

$$\Pi' : \begin{cases} x = x'_0 + s'l'' + t'l''' \\ y = y'_0 + s'm'' + t'm''' \\ z = z'_0 + s'n'' + t'n''' \end{cases}, \quad s', t' \in \mathbb{K}$$

due loro rappresentazioni parametriche, e siano

$$\Pi : ax + by + cz = d \quad (6.44)$$

e

$$\Pi' : a'x + b'y + c'z = d'$$

due loro rappresentazioni cartesiane. In questo caso, il Teorema 6.2.3 si specializza al

Teorema 6.3.4 Le seguenti affermazioni sono equivalenti:

- (1) Π e Π' sono paralleli;
- (2)

$$\text{rk} \begin{pmatrix} l & l' & l'' & l''' \\ m & m' & m'' & m''' \\ n & n' & n'' & n''' \end{pmatrix} = 2; \quad (6.45)$$

- (3)

$$\text{rk} \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} = 1. \quad (6.46)$$

Se i piani Π e Π' sono paralleli, allora le seguenti affermazioni sono equivalenti:

- (6) Π e Π' sono anche incidenti, e quindi $\Pi = \Pi'$;
- (7) oltre alla (6.45) vale anche che

$$\det \begin{pmatrix} x_0 - x'_0 & l & l' \\ y_0 - y'_0 & m & m' \\ z_0 - z'_0 & n & n' \end{pmatrix} = 0, \quad \text{o, equivalentemente,} \quad \det \begin{pmatrix} x_0 - x'_0 & l'' & l''' \\ y_0 - y'_0 & m'' & m''' \\ z_0 - z'_0 & n'' & n''' \end{pmatrix} = 0;$$

- (8) oltre alla (6.46) vale anche che

$$\text{rk} \begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \end{pmatrix} = 1.$$

Se Π e Π' sono paralleli ma non coincidono, allora sono paralleli e disgiunti. Se Π e Π' non sono paralleli, allora sono incidenti e la loro intersezione è una retta \bar{r} . In particolare, due piani in uno spazio affine tridimensionale non sono mai sghembi.

Dimostrazione. La dimostrazione è similissima alla dimostrazione del Teorema 6.3.1 ed è lasciata al lettore per esercizio. Il lettore è anche invitato a studiare il caso in cui di Π sia assegnata una rappresentazione parametrica e di Π' una rappresentazione cartesiana (a quest'ultimo riguardo si veda anche l'Esempio 6.21). ■

Riassumendo, dati due piani Π, Π' in uno spazio affine tridimensionale, per stabilire la loro posizione reciproca è opportuno, nell'ordine (vedi Figura 6.12),

- (1) stabilire se Π e Π' sono paralleli (utilizzando una delle condizioni di parallelismo fornite dal Teorema 6.3.4);

- (2) se Π e Π' sono paralleli, stabilire se coincidono oppure sono paralleli e disgiunti (utilizzando una delle condizioni di coincidenza fornite dal Teorema 6.3.4);
- (3) se Π e Π' non sono paralleli, concludere che sono incidenti e la loro intersezione è una retta (eventualmente trovare questa retta).

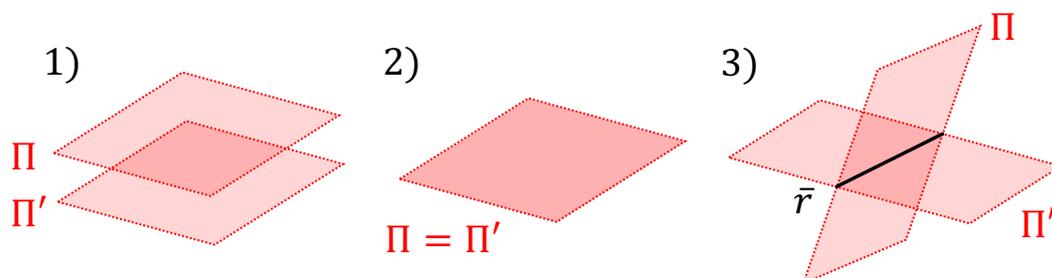


Figura 6.12: Posizioni reciproche di due piani Π e Π' in uno spazio affine tridimensionale: 1) Π e Π' sono paralleli e disgiunti, 2) Π e Π' (sono paralleli e) coincidono, 3) Π e Π' (non sono paralleli, ma sono incidenti e) si intersecano in una retta \bar{r} .

■ **Esempio 6.20** Nello spazio affine numerico reale tridimensionale $A_{\mathbb{R}}^3$ consideriamo i piani Π, Π' di rappresentazioni parametriche

$$\Pi : \begin{cases} x = 1 + s + t \\ y = -2s + t \\ z = -1 - s - 3t \end{cases}, \quad s, t \in \mathbb{R}$$

e

$$\Pi' : \begin{cases} x = + 2s' \\ y = 3 - s' - 3t' \\ z = -2 - 4s' + 2t' \end{cases}, \quad s', t' \in \mathbb{R}.$$

Vogliamo studiare la posizione reciproca di Π e Π' . Per stabilire se Π e Π' sono paralleli, studiamo le relative giaciture. La giacitura di Π è generata dai vettori

$$\begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -3 \end{pmatrix},$$

mentre la giacitura di Π' è generata dai vettori

$$\begin{pmatrix} 2 \\ -1 \\ -4 \end{pmatrix}, \begin{pmatrix} 0 \\ -3 \\ 2 \end{pmatrix}.$$

È facile verificare che la matrice

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

ha rango 2 (il lettore lo dimostri nel dettaglio). Perciò Π e Π' hanno la stessa giacitura e dobbiamo solo verificare se coincidono oppure no. A questo scopo basta verificare, per esempio, se il punto base $P_0 = (1, 0, -1)^T$ di Π appartiene o no a Π' . Detto $P'_0 = (0, 3, -2)^T$ il punto base di Π' , questo si

verifica a seconda che il vettore $\overrightarrow{P'_0 P_0} = (1, -3, 1)^T$ appartenga oppure no alla giacitura di Π' , in altre parole a seconda se il determinante della matrice

$$\begin{pmatrix} 1 & 2 & 0 \\ -3 & -1 & -3 \\ 1 & -4 & 2 \end{pmatrix} \quad (6.47)$$

sia 0 oppure no. Ma il determinante della matrice (6.47) è -8 (il lettore lo calcoli per verificare), perciò Π e Π' sono paralleli e disgiunti. ■

■ **Esempio 6.21** Nello spazio affine numerico reale tridimensionale $A_{\mathbb{R}}^3$ consideriamo il piano Π di rappresentazione parametrica

$$\Pi : \begin{cases} x = -1 & + & t \\ y = 1 & - & s \\ z = & s & + & \frac{1}{2}t \end{cases}, \quad s, t \in \mathbb{R} \quad (6.48)$$

e il piano Π' di rappresentazione cartesiana

$$\Pi' : x + \frac{3}{2}y - z = -1. \quad (6.49)$$

I vettori direttori di Π sono $d_1 = (0, -1, 1)^T$ e $d_2 = (1, 0, 1/2)^T$, mentre la giacitura W di Π' ha rappresentazione cartesiana

$$W : x + \frac{3}{2}y - z = 0. \quad (6.50)$$

Per verificare se Π e Π' sono paralleli oppure no è sufficiente verificare se d_1, d_2 appartengono a W oppure no, cioè se d_1, d_2 sono soluzione della (6.50) oppure no. Ora d_1 non è soluzione della (6.50) e tanto basta a concludere che Π e Π' non sono paralleli e perciò la loro intersezione è una retta \bar{r} . Per determinare \bar{r} è necessario individuare i punti (x, y, z) della forma (6.48) che, contemporaneamente, risolvono la (6.49). Sostituendo la (6.48) nella (6.49) troviamo la seguente equazione lineare per i parametri s, t ,

$$t - 5s = -3 \quad (6.51)$$

che, una volta risolta, fornisce i valori dei parametri s, t nella (6.48) corrispondenti ai punti di \bar{r} . Ora, dalla (6.51), troviamo $t = 5s - 3$. Risostituendo nella (6.48) troviamo la seguente rappresentazione parametrica della retta \bar{r} :

$$\Pi : \begin{cases} x = -4 & + & 5s \\ y = 1 & - & s \\ z = -\frac{3}{2} & + & \frac{7}{2}s \end{cases}, \quad s \in \mathbb{R}. \quad (6.52)$$

Per determinare una rappresentazione cartesiana di \bar{r} possiamo, eliminare il parametro dalla (6.52), oppure ragionare come segue. Dalla (6.48) ricaviamo una rappresentazione cartesiana di Π :

$$\Pi : \det \begin{pmatrix} x+1 & 0 & 1 \\ y-1 & -1 & 0 \\ z & 1 & 1/2 \end{pmatrix} = -\frac{1}{2}x + y + z - \frac{3}{2} = 0. \quad (6.53)$$

Una rappresentazione cartesiana di \bar{r} si ottiene allora unendo le equazioni (6.49) e (6.53):

$$\bar{r} : \begin{cases} -\frac{1}{2}x + y + z = \frac{3}{2} \\ x + \frac{3}{2}y - z = -1 \end{cases}.$$

Esercizio 6.6 Nello spazio affine numerico reale tridimensionale $A_{\mathbb{R}}^3$ studiare la posizione reciproca dei piani:

$$\Pi : 2x + 3y - z = -1, \quad \Pi' : -x + 5y + z = 2.$$

Più precisamente, dimostrare che Π e Π' non sono paralleli e determinare una rappresentazione parametrica e una cartesiana della retta $\bar{r} = \Pi \cap \Pi'$. ■

Consideriamo adesso un piano $\Pi \subseteq A$ di rappresentazione parametrica (6.43) e rappresentazione cartesiana (6.44) e una retta $r \subseteq A$ di rappresentazione parametrica

$$r : \begin{cases} x = x'_0 + tl'' \\ y = y'_0 + tm'' \\ z = z'_0 + tn'' \end{cases}, \quad t \in \mathbb{K}.$$

e rappresentazione cartesiana

$$r : \begin{cases} a'x + b'y + c'z = d' \\ a''x + b''y + c''z = d'' \end{cases}.$$

Il Teorema 6.2.3 si specializza al

Teorema 6.3.5 Le seguenti affermazioni sono equivalenti:

- (1) Π ed r sono sottospazi paralleli;
- (2)

$$\det \begin{pmatrix} l & l' & l'' \\ m & m' & m'' \\ n & n' & n'' \end{pmatrix} = 0; \quad (6.54)$$

- (3)

$$\det \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} = 0. \quad (6.55)$$

Se il piano Π e la retta r sono paralleli, allora le seguenti affermazioni sono equivalenti:

- (6) Π ed r sono anche incidenti, e quindi $r \subseteq \Pi$;
- (7) oltre alla (6.54) vale anche che

$$\det \begin{pmatrix} x_0 - x'_0 & l & l' \\ y_0 - y'_0 & m & m' \\ z_0 - z'_0 & n & n' \end{pmatrix} = 0;$$

- (8) oltre alla (6.55) vale anche che

$$\text{rk} \begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{pmatrix} = 2.$$

Se Π ed r sono paralleli ma r non è inclusa in Π , allora sono paralleli e disgiunti. Se Π ed r non sono paralleli, allora sono incidenti e la loro intersezione è un punto \bar{P} le cui coordinate $(\bar{x}, \bar{y}, \bar{z})$

sono date da

$$\bar{x} = \frac{\det \begin{pmatrix} d & b & c \\ d' & b' & c' \\ d'' & b'' & c'' \end{pmatrix}}{\det \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix}}, \quad \bar{y} = \frac{\det \begin{pmatrix} a & d & c \\ a' & d' & c' \\ a'' & d'' & c'' \end{pmatrix}}{\det \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix}}, \quad \bar{z} = \frac{\det \begin{pmatrix} a & b & d \\ a' & b' & d' \\ a'' & b'' & d'' \end{pmatrix}}{\det \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix}}.$$

In particolare, due piani in uno spazio affine tridimensionale non sono mai sghembi.

Dimostrazione. Lasciamo la dimostrazione come esercizio per il lettore. ■

Riassumendo, dati un piano Π e una retta r in uno spazio affine tridimensionale, per stabilire la loro posizione reciproca è opportuno, nell'ordine (vedi Figura 6.13),

- (1) stabilire se Π ed r sono paralleli (utilizzando una delle condizioni di parallelismo fornite dal Teorema 6.3.5);
- (2) se Π ed r sono paralleli, stabilire se $r \subseteq \Pi$ oppure $\Pi \cap r = \emptyset$ (utilizzando una delle condizioni fornite dal Teorema 6.3.5);
- (3) se Π ed r non sono paralleli, concludere che sono incidenti e la loro intersezione è un punto (eventualmente trovare questo punto).

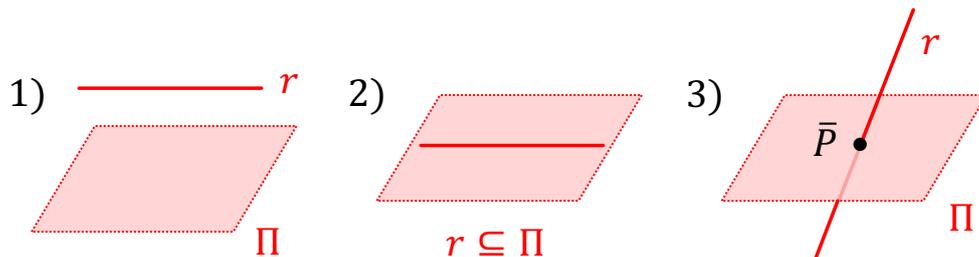


Figura 6.13: Posizioni reciproche di un piano Π e di una retta r in uno spazio affine tridimensionale: 1) Π ed r sono paralleli e disgiunti, 2) (Π ed r sono paralleli e) $r \subseteq \Pi$, 3) Π ed r (non sono paralleli, ma sono incidenti e) si intersecano in un punto \bar{P} .

■ **Esempio 6.22** Nello spazio affine numerico reale tridimensionale $A_{\mathbb{R}}^3$ consideriamo il piano Π e la retta r di rappresentazioni parametriche

$$\Pi : \begin{cases} x = 3 + \frac{1}{3}s - 2t \\ y = -\frac{1}{3} + s - 3t \\ z = \frac{2}{3} + t \end{cases}, \quad s, t \in \mathbb{R}, \quad r : \begin{cases} x = \frac{2}{3}t' \\ y = -\frac{1}{3} - t' \\ z = -1 - \frac{1}{3}t' \end{cases}, \quad t' \in \mathbb{R}.$$

Vogliamo studiare la posizione reciproca di Π ed r . Per stabilire se Π ed r sono paralleli dobbiamo verificare se il vettore direttore $(2/3, -1, -1/3)^T$ di r appartiene alla giacitura di Π , o, equivalentemente, se è combinazione lineare dei vettori direttori $(1/3, 1, 0)^T, (-2, -3, 1)^T$ di Π . Ma questo avviene sse il determinante

$$\det \begin{pmatrix} 1/3 & -2 & 2/3 \\ 1 & -3 & -1 \\ 0 & 1 & -1/3 \end{pmatrix}$$

si annulla. Invece

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

Dunque Π ed r non sono paralleli e quindi sono incidenti e si intersecano in un punto. Per determinare il punto di intersezione \bar{P} , possiamo procedere in vari modi. Per esempio trovando le rappresentazioni cartesiane di Π ed r e mettendole a sistema (al lettore è chiaro?). Oppure come segue. Sia $\bar{P} = (\bar{x}, \bar{y}, \bar{z})^T$. Poichè $\bar{P} \in \Pi$, esistono $\bar{s}, \bar{t} \in \mathbb{R}$ tali che

$$\begin{cases} \bar{x} = 3 + \frac{1}{3}\bar{s} - 2\bar{t} \\ \bar{y} = -\frac{1}{3} + \bar{s} - 3\bar{t} \\ \bar{z} = \frac{2}{3} + \bar{t} \end{cases} \quad (6.56)$$

D'altronde, poichè $\bar{P} \in r$, esiste anche $\bar{t}' \in \mathbb{R}$ tale che

$$\begin{cases} \bar{x} = \frac{2}{3}\bar{t}' \\ \bar{y} = -\frac{1}{3} - \bar{t}' \\ \bar{z} = -1 - \frac{1}{3}\bar{t}' \end{cases}. \quad (6.57)$$

Questo implica che la terna $(\bar{s}, \bar{t}, \bar{t}')$ soddisfa il sistema lineare

$$\begin{cases} 3 + \frac{1}{3}\bar{s} - 2\bar{t} = \frac{2}{3}\bar{t}' \\ -\frac{1}{3} + \bar{s} - 3\bar{t} = -\frac{1}{3} - \bar{t}' \\ \frac{2}{3} + \bar{t} = -1 - \frac{1}{3}\bar{t}' \end{cases},$$

che riscritto nella forma usuale è

$$\begin{cases} \frac{1}{3}\bar{s} - 2\bar{t} - \frac{2}{3}\bar{t}' = -3 \\ \bar{s} - 3\bar{t} + \bar{t}' = 0 \\ \bar{t} + \frac{1}{3}\bar{t}' = -\frac{5}{3} \end{cases}.$$

Per costruzione, il determinante della matrice incompleta è diverso da zero, perciò questo sistema ammette un'unica soluzione che si può trovare, per esempio, mediante la Formula di Cramer. Si trova

$$\bar{s} = -19, \quad \bar{t} = -4, \quad \bar{t}' = 7.$$

Sostituendo \bar{t}' nella (6.57) (o, equivalentemente, sostituendo \bar{s}, \bar{t} nella (6.56)) troviamo

$$\bar{P} = \begin{pmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{pmatrix} = \begin{pmatrix} 14/3 \\ -22/3 \\ -10/3 \end{pmatrix}.$$

■

■ **Esempio 6.23** Nello spazio affine numerico reale tridimensionale $A_{\mathbb{R}}^3$ consideriamo il piano Π e la retta r di rappresentazioni cartesiane

$$\Pi : -\frac{1}{2}x + y - \frac{3}{2}z = 2, \quad r : \begin{cases} x + \frac{5}{2}y + z = \frac{1}{2} \\ \frac{5}{4}y - z = \frac{1}{4} \end{cases}.$$

L'intersezione $\Pi \cap r$ è rappresentata dal sistema

$$\Pi \cap r : \begin{cases} -\frac{1}{2}x + y - \frac{3}{2}z = 2 \\ x + \frac{5}{2}y + z = \frac{1}{2} \\ \frac{5}{4}y - z = \frac{1}{4} \end{cases},$$

la cui matrice incompleta

$$\begin{pmatrix} -1/2 & 1 & -3/2 \\ 1 & 5/2 & 1 \\ 0 & 9/4 & -1 \end{pmatrix}$$

ha determinante pari a zero e quindi rango pari a 2 (il lettore lo verifichi per esercizio). Dunque Π ed r sono paralleli. La matrice completa

$$\begin{pmatrix} -1/2 & 1 & -3/2 & 2 \\ 1 & 5/2 & 1 & 1/2 \\ 0 & 9/4 & -1 & 9/4 \end{pmatrix}$$

ha rango pari a 2 (il lettore verifichi in dettaglio anche questo). Concludiamo che $r \subseteq \Pi$. ■

Esercizio 6.7 Nello spazio affine numerico reale tridimensionale $A_{\mathbb{R}}^3$ studiare la posizione reciproca del piano Π di rappresentazione parametrica

$$\Pi: \begin{cases} x = & 3s \\ y = 2 + 3s - 2t \\ z = -1 + 3s + 2t \end{cases}, s, t \in \mathbb{R}$$

e della retta r di rappresentazione cartesiana

$$r: \begin{cases} 2x - y - z = 3 \\ x + z = -5 \end{cases}.$$

Infine studiamo la posizione reciproca di due rette $r, r' \subseteq A$.

Definizione 6.3.1 — Rette Complanari. Le rette $r, r' \subseteq A$ si dicono *complanari* se esiste un piano $\Pi \subseteq A$ che le contiene entrambe: $\Pi \supseteq r \cup r'$.

Proposizione 6.3.6 Le rette $r, r' \subseteq A$ sono complanari sse sono parallele oppure incidenti. In altre parole r, r' sono complanari sse non sono sghembe.

Dimostrazione. Supponiamo innanzitutto che r, r' siano complanari e sia Π il piano che le contiene. Allora possiamo riguardare r, r' come rette nello spazio affine bidimensionale Π e sappiamo che, in quanto tali, sono necessariamente parallele o incidenti. Viceversa, supponiamo che r, r' siano parallele o incidenti. Supponiamo innanzitutto che siano parallele. Allora o coincidono, nel qual caso esiste certamente un piano che le contiene entrambe, oppure sono parallele e disgiunte. Se sono parallele e disgiunte ragioniamo come segue. Sia $v \in V$ un vettore direttore di r , quindi v è anche un vettore direttore di r' (che ha la stessa giacitura di r), sia P_0 un punto di r e sia P'_0 un punto di r' . Il vettore congiungente $\overrightarrow{P_0P'_0}$ e il vettore v sono linearmente indipendenti, altrimenti $\overrightarrow{P_0P'_0}$ apparterebbe alla giacitura di r e si avrebbe $P'_0 \in r$, il che non può essere perché r ed r' non hanno punti in comune. Dunque v e $\overrightarrow{P_0P'_0}$ generano un sottospazio vettoriale bidimensionale $W \subseteq V$. Consideriamo il piano $\Pi \subseteq A$ di giacitura W che passa per P_0 , cioè $\Pi = \alpha_{P_0}^{-1}(W)$ (Figura 6.14). Il piano Π contiene r , infatti contiene un punto di r , il punto P_0 , e la sua giacitura W contiene la giacitura di r (perché contiene un vettore direttore di r). Ma il piano Π contiene anche r' , infatti, per costruzione contiene un punto di r' , il punto P'_0 ($P_0 \in \Pi$ e $\overrightarrow{P_0P'_0}$ appartiene alla giacitura di Π), e la sua giacitura contiene la giacitura di r' (che coincide con la giacitura di r). Dunque Π è un piano che contiene sia r che r' che, perciò, sono rette complanari.

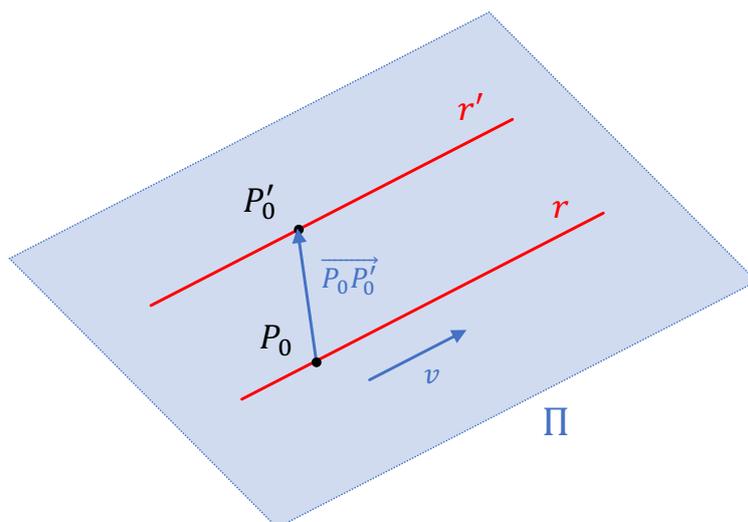


Figura 6.14: Due rette parallele e disgiunte r, r' in uno spazio affine 3D sono complanari.

Se invece r ed r' sono incidenti, o coincidono (nel qual caso abbiamo già osservato che sono complanari) oppure si intersecano in un solo punto \bar{P} (il lettore lo vede?). Consideriamo questo secondo caso e siano v, v' vettori direttori di r, r' rispettivamente. Poiché r ed r' non sono parallele (altrimenti, essendo anche incidenti, coinciderebbero), allora v, v' sono linearmente indipendenti e generano un sottospazio vettoriale bidimensionale $W \subseteq V$. Consideriamo il piano $\Pi \subseteq A$ di giacitura W che passa per \bar{P} , cioè $\Pi = \alpha_{\bar{P}}^{-1}(W)$ (Figura 6.15). Il piano Π contiene r perché contiene il punto \bar{P} di r e la sua giacitura W contiene la giacitura di r (che è generata da v). Per lo stesso motivo Π contiene anche r' . Questo conclude la dimostrazione. ■

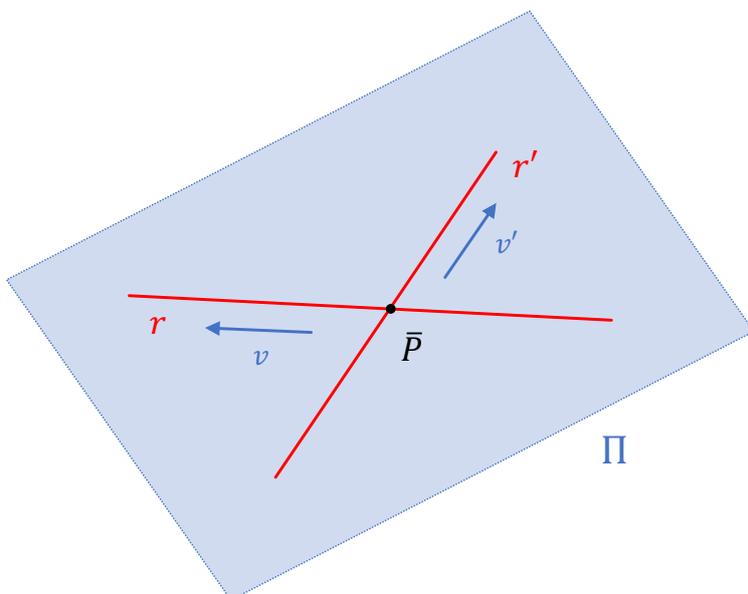


Figura 6.15: Due rette incidenti e non parallele r, r' in uno spazio affine 3D sono complanari.



Si osservi che la stessa identica dimostrazione funziona qualunque sia la dimensione dello

spazio affine ambiente A . Dunque due rette sono complanari sse non sono sghembe in ogni spazio affine.

Esercizio 6.8 Dimostrare che, se le rette $r, r' \subseteq A$ sono parallele e disgiunte oppure sono incidenti ma non parallele (e quindi hanno un solo punto in comune) allora, non solo (sono complanari cioè) esiste un piano Π che le contiene, ma tale piano Π è unico. ■

Ora siano

$$r: \begin{cases} x = x_0 + tl \\ y = y_0 + tm \\ z = z_0 + tn \end{cases}, \quad t \in \mathbb{K}, \quad r': \begin{cases} x = x'_0 + t'l' \\ y = y'_0 + t'm' \\ z = z'_0 + t'n' \end{cases}, \quad t' \in \mathbb{K} \quad (6.58)$$

le rappresentazioni parametriche, e

$$r: \begin{cases} ax + by + cz = d \\ a'x + b'y + c'z = d' \end{cases}, \quad r': \begin{cases} a''x + b''y + c''z = d'' \\ a'''x + b'''y + c'''z = d''' \end{cases} \quad (6.59)$$

le rappresentazioni cartesiane di r, r' . Per studiare la posizione reciproca di r, r' è opportuno, innanzitutto, stabilire se esse sono complanari oppure no. Questo si può fare guardando le loro rappresentazioni parametriche o cartesiane in base alla seguente

Proposizione 6.3.7 — Criteri di Complanarità. Le seguenti affermazioni sono equivalenti:

- (1) r, r' sono rette complanari;
- (2)

$$\det \begin{pmatrix} x_0 - x'_0 & l & l' \\ y_0 - y'_0 & m & m' \\ z_0 - z'_0 & n & n' \end{pmatrix} = 0;$$

- (3)

$$\det \begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \\ a''' & b''' & c''' & d''' \end{pmatrix} = 0.$$

Dimostrazione. La presente proposizione è un semplice corollario del Teorema 6.2.3. Per maggiore chiarezza ripetiamo in questo caso specifico alcuni dettagli della dimostrazione del Teorema 6.2.3.

Indichiamo preliminarmente

$$B = \begin{pmatrix} l & l' \\ m & m' \\ n & n' \end{pmatrix}, \quad \mathbb{B} = \begin{pmatrix} x_0 - x'_0 & l & l' \\ y_0 - y'_0 & m & m' \\ z_0 - z'_0 & n & n' \end{pmatrix}.$$

Indichiamo anche

$$A = \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \\ a''' & b''' & c''' \end{pmatrix}, \quad \mathbb{A} = \begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \\ a''' & b''' & c''' & d''' \end{pmatrix}.$$

Ora dimostriamo che (1) \Leftrightarrow (2). Se le rette r, r' sono complanari, allora sono parallele oppure sono incidenti. Se sono parallele, i loro vettori direttori, e quindi i relativi vettori coordinati, sono

linearmente dipendenti cioè la matrice B ha rango 1 e la matrice \mathbb{B} non può avere rango maggiore di 2, da cui $\det \mathbb{B} = 0$. Se r, r' sono incidenti, allora, dal Teorema 6.2.3, $\text{rk } B = \text{rk } \mathbb{B}$. Poiché il rango di B non può superare 2, anche il rango di \mathbb{B} non supera 2 e, anche in questo caso, $\det \mathbb{B} = 0$. Viceversa, sia $\det \mathbb{B} = 0$, da cui $\text{rk } \mathbb{B} < 3$. Se $\text{rk } B = 1$, allora i vettori direttori di r, r' sono linearmente dipendenti e quindi r, r' sono parallele e dunque complanari (in questa situazione il rango di \mathbb{B} può essere 1 oppure 2, nel primo caso si ha $r = r'$, altrimenti r, r' sono parallele e disgiunte). Se $\text{rk } B = 2$, allora $3 > \text{rk } \mathbb{B} \geq \text{rk } B = 2$, perciò $\text{rk } \mathbb{B} = 2 = \text{rk } B$ e (di nuovo dal Teorema 6.2.3) r, r' sono incidenti (più precisamente sono incidenti, ma non parallele, perciò si intersecano esattamente in un punto).

Infine dimostriamo che (1) \Leftrightarrow (3). A questo scopo ricordiamo, innanzitutto, che le giaciture di r, r' hanno rappresentazioni cartesiane

$$\begin{cases} ax + by + cz = 0 \\ a'x + b'y + c'z = 0 \end{cases}, \quad \begin{cases} a''x + b''y + c''z = 0 \\ a'''x + b'''y + c'''z = 0 \end{cases}, \quad (6.60)$$

rispettivamente, mentre l'intersezione $r \cap r'$ è rappresentata dal sistema lineare

$$r \cap r' : \begin{cases} ax + by + cz = d \\ a'x + b'y + c'z = d' \\ a''x + b''y + c''z = d'' \\ a'''x + b'''y + c'''z = d''' \end{cases} \quad (6.61)$$

le cui matrici incompleta e completa sono A e \mathbb{A} rispettivamente. Se r, r' sono complanari, come al solito, sono parallele oppure incidenti. Se sono parallele, hanno la stessa giacitura e i sistemi lineari (6.60) sono equivalenti. Ne consegue che le ultime due righe della matrice A dipendono linearmente dalle prime due e il rango di A è pari a 2. Perciò il rango di \mathbb{A} non può superare 3 e $\det \mathbb{A} = 0$. Se r, r' sono incidenti, il sistema (6.61) è compatibile e, di nuovo per il Teorema di Rouché-Capelli, $\text{rk } \mathbb{A} = \text{rk } A$. In particolare il rango di \mathbb{A} non può superare 3 e $\det \mathbb{A} = 0$. Viceversa, sia $\det \mathbb{A} = 0$, da cui $\text{rk } \mathbb{A} < 4$. Se $\text{rk } A = 2$, allora i sistemi (6.60), sono equivalenti, ed r, r' hanno la stessa giacitura, cioè sono parallele. In questo caso, il rango di \mathbb{A} può essere 2, e allora il sistema (6.61) è compatibile, r, r' sono anche incidenti e dunque $r = r'$. Oppure il rango di \mathbb{A} può essere 3, e allora il sistema (6.61) è incompatibile ed r, r' sono parallele e disgiunte. In ogni caso r, r' sono complanari. Se $\text{rk } A = 3$, allora $4 > \text{rk } \mathbb{A} \geq \text{rk } A = 3$, da cui $\text{rk } \mathbb{A} = 3 = \text{rk } A$ e il sistema (6.61) è compatibile. Dunque r, r' non sono parallele ma sono incidenti (e si intersecano in un unico punto), in particolare sono complanari. ■

Se le rette r, r' sono complanari, allora già sappiamo che le loro possibili posizioni reciproche sono le seguenti: o r, r' sono parallele e disgiunte, o $r = r'$, oppure $r \cap r'$ è esattamente un punto. La dimostrazione della Proposizione 6.3.7 indica come distinguere questi tre casi (nel caso in cui r, r' siano complanari): se abbiamo a disposizione le rappresentazioni parametriche (6.58) e le matrici B, \mathbb{B} sono come nella dimostrazione del Teorema 6.3.7, allora

- (1) r, r' sono parallele e disgiunte sse $\text{rk } B = 1$ e $\text{rk } \mathbb{B} = 2$;
- (2) $r = r'$ sse $\text{rk } B = \text{rk } \mathbb{B} = 1$;
- (3) r ed r' sono incidenti ma non parallele sse $\text{rk } B = 2$.

Se invece abbiamo a disposizione le rappresentazioni cartesiane (6.59) (e le matrici A, \mathbb{A} sono come nella dimostrazione del Teorema 6.3.7) allora

- (1) r, r' sono parallele e disgiunte sse $\text{rk } A = 2$ e $\text{rk } \mathbb{A} = 3$;
- (2) $r = r'$ sse $\text{rk } A = \text{rk } \mathbb{A} = 2$;
- (3) r ed r' sono incidenti ma non parallele sse $\text{rk } A = 3$.

È anche possibile trovare condizioni di complanarità in termini di una rappresentazione parametrica per r e una rappresentazione cartesiana per r' e invitiamo vivamente il lettore a farlo. Se r, r' non sono complanari, allora sono sghembe (Figura 6.16).

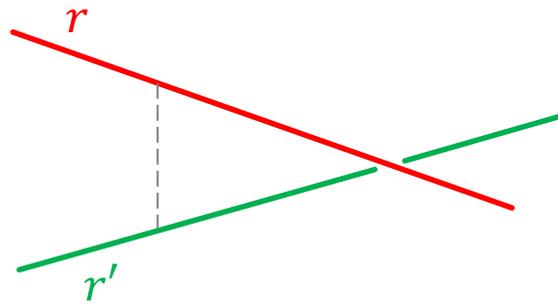


Figura 6.16: Due rette sghembe r, r' in uno spazio affine tridimensionale.

■ **Esempio 6.24** Nello spazio affine numerico reale tridimensionale $A_{\mathbb{R}}^3$ consideriamo le rette

$$r: \begin{cases} -y - z = 2 \\ 2x + y + 2z = 0 \end{cases}, \quad (6.62)$$

e

$$r': \begin{cases} x + y = 0 \\ -x + y - z = -2 \end{cases},$$

e siano A, \mathbb{A} matrici come nella dimostrazione della Proposizione 6.3.7. Un conto diretto mostra che $\det \mathbb{A} = 0$ (il lettore lo dimostri nel dettaglio). Dunque r, r' sono complanari. Inoltre $\text{rk } A = 3$ (il lettore dimostri anche questo). Concludiamo che r, r' sono incidenti ma non parallele, perciò si intersecano esattamente in un punto $\bar{P} = (\bar{x}, \bar{y}, \bar{z})^T$ che può essere determinato risolvendo il sistema

$$r \cap r': \begin{cases} -y - z = 2 \\ 2x + y + 2z = 0 \\ x + y = 0 \\ -x + y - z = -2. \end{cases}$$

Tale sistema è compatibile e solo tre delle sue equazioni sono linearmente indipendenti, per esempio le prime 3, dunque

$$r \cap r': \begin{cases} -y - z = 2 \\ 2x + y + 2z = 0 \\ x + y = 0. \end{cases}$$

Quest'ultimo sistema può essere risolto, per esempio, con il metodo di Cramer (il lettore lo vede senza fare ulteriori conti?), trovando

$$\bar{x} = 4/3, \quad \bar{y} = -4/3, \quad \bar{z} = -2/3.$$

Possiamo anche determinare il piano Π che contiene le rette r ed r' , per esempio come segue. Il piano Π passa per il punto \bar{P} e la sua giacitura è generata dai vettori direttori di r ed r' . Un vettore direttore v di r è il vettore le cui entrate sono i minori di ordine 2 della matrice dei coefficienti della (6.62) presi a segni alterni:

$$v = \left(\begin{vmatrix} -1 & -1 \\ 1 & 2 \end{vmatrix}, - \begin{vmatrix} 0 & -1 \\ 2 & 2 \end{vmatrix}, \begin{vmatrix} 0 & -1 \\ 2 & 1 \end{vmatrix} \right)^T = \begin{pmatrix} -1 \\ -2 \\ 2 \end{pmatrix}.$$

Un vettore direttore v' di r' si può trovare in modo analogo ed è:

$$v' = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}.$$

Dunque il piano Π ha rappresentazione parametrica

$$\Pi : \begin{cases} x = \frac{4}{3} - s - t \\ y = -\frac{4}{3} - 2s + t \\ z = -\frac{2}{3} + 2s + 2t \end{cases}, \quad s, t \in \mathbb{R}. \quad (6.63)$$

Per trovare una rappresentazione cartesiana di Π possiamo trasformare la rappresentazione parametrica (6.63) nel solito modo:

$$\Pi : \det \begin{pmatrix} x-4/3 & -1 & -1 \\ y+4/3 & -2 & 1 \\ z+2/3 & 2 & 2 \end{pmatrix} = -6x - 3z + 6 = 0.$$

Esercizio 6.9 Determinare la posizione reciproca delle rette r, r' nello spazio affine numerico reale tridimensionale $A_{\mathbb{R}}^3$ in ciascuno dei seguenti due casi:

(1)

$$r : \begin{cases} x = 1 + t \\ y = -7 - 2t \\ z = + 4t \end{cases}, \quad t \in \mathbb{R}, \quad r' : \begin{cases} x = 5 + 3t' \\ y = 1 - 6t' \\ z = -2 + 12t' \end{cases}, \quad t' \in \mathbb{R}.$$

(2)

$$r : \begin{cases} x = t \\ y = 1 \\ z = -1 - t \end{cases}, \quad t \in \mathbb{R}, \quad r' : \begin{cases} x - y = 3 \\ -x - y + \frac{1}{2}z = 2 \end{cases}.$$

Sia A uno spazio affine tridimensionale e sia $\bar{r} \subseteq A$ una retta. È utile qualche volta considerare l'insieme di tutti i piani che contengono la retta \bar{r} . Tale insieme è detto il *fascio proprio* di piani di centro \bar{r} . Se \bar{r} è presentata come intersezione di due piani non paralleli Π e Π' , e di Π e Π' sono note rappresentazioni cartesiane (è questo il caso, per esempio, se di \bar{r} è assegnata una rappresentazione cartesiana, Figura 6.17), allora è facile scrivere una rappresentazione cartesiana della generica retta del fascio proprio di centro \bar{r} in base alla seguente

Proposizione 6.3.8 Sia A uno spazio affine tridimensionale sul campo \mathbb{K} , sia (O, \mathcal{R}) un riferimento affine di A , e siano Π e Π' piani non paralleli in A di rappresentazioni cartesiane

$$\Pi : ax + by + cz = d, \quad \Pi' : a'x + b'y + c'z = d'$$

nel riferimento affine (O, \mathcal{R}) . Detta $\bar{r} = \Pi \cap \Pi'$ la retta intersezione di Π e Π' , un piano $\Pi'' \subseteq A$ appartiene al fascio proprio di piani di centro \bar{r} sse Π'' ha rappresentazione cartesiana data da

$$\Pi'' : \lambda(ax + by + cz) + \lambda'(a'x + b'y + c'z) = \lambda d + \lambda' d' \quad (6.64)$$

per qualche coppia di scalari $\lambda, \lambda' \in \mathbb{K}$, con $(\lambda, \lambda') \neq (0, 0)$.

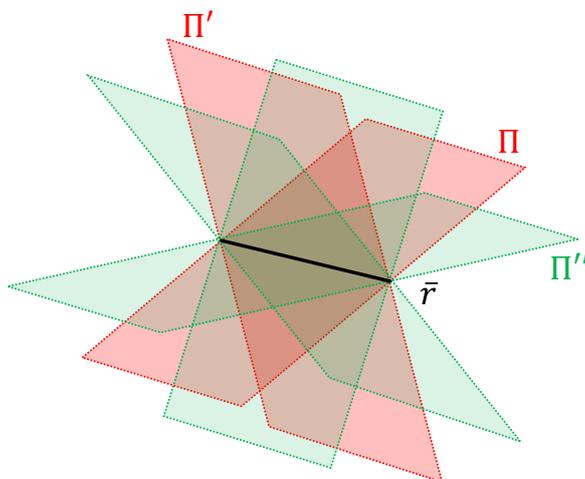


Figura 6.17: Fascio proprio di piani di centro la retta \bar{r} in uno spazio affine 3D.

Dimostrazione. La dimostrazione è similissima a quella della Proposizione 6.3.2 ed è lasciata per esercizio. ■

■ **Esempio 6.25** Similmente al caso del fascio di rette nel piano, l'equazione del fascio di piani (Equazione (6.64)) torna utile quando si voglia determinare una rappresentazione cartesiana per un piano che contiene una retta \bar{r} (di cui è nota una rappresentazione cartesiana) e soddisfa condizioni aggiuntive. Per esempio, nello spazio affine numerico reale tridimensionale $A_{\mathbb{R}}^3$ (con il riferimento affine canonico) consideriamo la retta

$$\bar{r} : \begin{cases} x - 2y + 4z = -6 \\ 3x - y + 5z = 0 \end{cases} \quad (6.65)$$

Consideriamo anche il punto $Q = (0, 0, 1)^T$. Poiché Q non risolve la prima equazione del sistema (6.65), allora $Q \notin \bar{r}$ ed esiste un unico piano Π'' che contiene \bar{r} e passa per Q (il lettore lo vede?). Possiamo determinare una rappresentazione cartesiana di Π'' come segue. Il piano Π'' appartiene al fascio proprio di piani di centro \bar{r} , perciò ogni sua rappresentazione cartesiana è del tipo:

$$\Pi'' : \lambda(x - 2y + 4z) + \lambda'(3x - y + 5z) = -6\lambda, \quad (6.66)$$

con $\lambda, \lambda' \in \mathbb{R}$ (in altre parole l'equazione di Π'' dipende linearmente dalle equazioni di \bar{r}). Vogliamo determinare λ, λ' in modo tale che $Q \in \Pi''$. Chiaramente $(0, 0, 1)^T$ deve essere soluzione dell'Equazione (6.66), cioè

$$10\lambda + 5\lambda' = 0.$$

Per esempio, se poniamo $\lambda = 1$, deve essere $\lambda' = -2$. Da cui

$$r : -5x - 6z = -6.$$

Si osservi che possiamo scegliere per λ un qualunque valore diverso da 0 (purchè poi si ponga $\lambda' = -2\lambda$). Al variare di λ otteniamo tutte le (diverse) rappresentazioni cartesiane del piano Π'' cercato. ■

Esercizio 6.10 Nello spazio affine numerico reale tridimensionale $A_{\mathbb{R}}^3$ si considerino le rette

$$\bar{r}: \begin{cases} x - 2y - 3z = 1 \\ 5x - 2y + 3z = 0 \end{cases} \quad \text{e} \quad r: \begin{cases} x + y = 1/2 \\ -y + 2z = -1 \end{cases}.$$

Dimostrare che esiste un unico piano $\bar{\Pi}$ che contiene la retta \bar{r} ed è parallelo alla retta r . Determinare una rappresentazione parametrica e una cartesiana del piano $\bar{\Pi}$. ■

Dato uno spazio affine tridimensionale A di giacitura V ed un sottospazio vettoriale $W \subseteq V$ di dimensione 2, qualche volta è anche utile considerare l'insieme di tutti i piani in A che hanno giacitura pari ad W (e quindi sono tutti paralleli tra loro). Tale insieme è detto *fascio improprio* di piani di giacitura W . Data una rappresentazione parametrica (risp. cartesiana) di W , è facile scrivere una rappresentazione parametrica (risp. cartesiana) del generico piano del fascio improprio di giacitura W (vedi Figura 6.18) in base alla

Proposizione 6.3.9 Sia A uno spazio affine tridimensionale sul campo \mathbb{K} modellato sullo spazio vettoriale V , sia (O, \mathcal{R}) un riferimento affine di A , e sia $W \subseteq V$ un sottospazio vettoriale di dimensione 2. Inoltre siano

$$W: \begin{cases} x = sl + t'l' \\ y = sm + tm' \\ z = sn + tn' \end{cases}, \quad s, t \in \mathbb{K},$$

e

$$W: ax + by + cz = 0$$

una rappresentazione parametrica e una cartesiana di W nel riferimento \mathcal{R} . Allora, dato un piano $\Pi \subseteq A$, le seguenti affermazioni sono equivalenti:

- (1) Π appartiene al fascio improprio di piani di giacitura W ;
- (2) Π ha la seguente rappresentazione parametrica nel riferimento affine (O, \mathcal{R})

$$\Pi: \begin{cases} x = x_0 + sl + t'l' \\ y = y_0 + sm + tm' \\ z = z_0 + sn + tn' \end{cases}, \quad s, t \in \mathbb{K},$$

per qualche terna di scalari $x_0, y_0, z_0 \in \mathbb{K}$;

- (3) Π ha la seguente rappresentazione cartesiana nel riferimento affine (O, \mathcal{R})

$$\Pi: ax + by + cz = d$$

per qualche scalare $d \in \mathbb{K}$.

Dimostrazione. La dimostrazione è ovvia e i dettagli sono lasciati al lettore per esercizio. ■

■ **Esempio 6.26** Nello spazio affine tridimensionale A consideriamo un piano $\Pi_0 \subseteq A$ di rappresentazione parametrica

$$\Pi_0: \begin{cases} x = x_0 + sl + t'l' \\ y = y_0 + sm + tm' \\ z = z_0 + sn + tn' \end{cases}, \quad s, t \in \mathbb{K},$$

e rappresentazione cartesiana

$$\Pi_0: ax + by + cz = d.$$

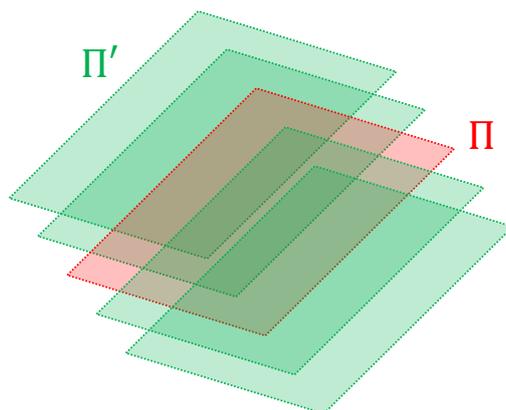


Figura 6.18: Fascio improprio di piani in uno spazio affine tridimensionale.

Ogni piano Π parallelo a Π_0 ha una rappresentazione parametrica del tipo

$$\Pi : \begin{cases} x = \bar{x}_0 + sl + tl' \\ y = \bar{y}_0 + sm + tm' \\ z = \bar{z}_0 + sn + tn' \end{cases}, \quad s, t \in \mathbb{K},$$

e una rappresentazione cartesiana del tipo

$$\Pi : ax + by + cz = \bar{d}.$$

■ **Esempio 6.27** Nello spazio affine numerico reale tridimensionale $A_{\mathbb{R}}^3$ consideriamo il piano

$$\Pi_0 : \frac{2}{3}x + \frac{5}{6}y - z = -\frac{1}{3},$$

e il punto $P = (0, 3, 1)^T$. Determiniamo il piano Π parallelo a Π_0 che passa per il punto P . Affinchè sia parallelo a Π_0 , il piano Π deve avere una rappresentazione cartesiana del tipo

$$\Pi : \frac{2}{3}x + \frac{5}{6}y - z = \bar{c}, \tag{6.67}$$

e si tratta solo di determinare $\bar{c} \in \mathbb{R}$ in modo tale che $P \in \Pi$, cioè $(0, 3, 1)^T$ sia soluzione della (6.67). Un facile conto mostra che

$$\bar{c} = 3/2.$$

Perciò

$$\Pi : \frac{2}{3}x + \frac{5}{6}y - z = \frac{3}{2}.$$

6.4 Applicazioni Affini

In questa sezione studiamo una classe di applicazioni adatte a confrontare due spazi affini. Siano dunque (A, V, α) e (A', V', α') spazi affini di dimensione n ed n' , sullo stesso campo \mathbb{K} .

Definizione 6.4.1 — Applicazione Affine. Un'applicazione affine tra gli spazi affini A e A' è una funzione

$$F : A \rightarrow A'$$

per la quale esista un'applicazione lineare $f_F : V \rightarrow V'$ tale che, per ogni $P, Q \in A$

$$\overrightarrow{F(P)F(Q)} = f_F(\overrightarrow{PQ}).$$

In tal caso, l'applicazione lineare $f_F : V \rightarrow V'$ è detta *applicazione lineare associata* all'applicazione affine F , o anche *parte lineare* di F .

Presentiamo ora una carrellata di esempi di applicazioni affini.

■ **Esempio 6.28 — Funzioni Costanti.** Ogni funzione costante tra spazi affini è un'applicazione affine. Più precisamente, sia $P'_0 \in A'$ e sia $c_{P'_0} : A \rightarrow A'$ la funzione che vale costantemente P'_0 , cioè

$$c_{P'_0} : A \rightarrow A', \quad P \mapsto c_{P'_0}(P) := P'_0.$$

Ora siano $P, Q \in A$. Calcoliamo

$$\overrightarrow{c_{P'_0}(P)c_{P'_0}(Q)} = \overrightarrow{P'_0P'_0} = 0.$$

Questo mostra che $c_{P'_0} : A \rightarrow A'$ è un'applicazione affine la cui parte lineare è l'applicazione lineare nulla $0 : V \rightarrow V'$, $v \mapsto 0(v) = 0$. Viceversa, consideriamo un'applicazione affine $F : A \rightarrow A'$ e supponiamo che la parte lineare $f_F : V \rightarrow V'$ di F sia l'applicazione lineare nulla. Allora, per ogni $P, Q \in A$ si ha che

$$\overrightarrow{F(P)F(Q)} = 0(\overrightarrow{PQ}) = 0 \quad \Rightarrow \quad F(P) = F(Q).$$

In altre parole F è una funzione costante. Questo mostra che le applicazioni affini con parte lineare nulla sono esattamente le funzioni costanti. ■

■ **Esempio 6.29 — Traslazioni.** Fissiamo un vettore $v \in V$ e consideriamo la funzione

$$\tau_v : A \rightarrow A, \quad P \mapsto \tau_v(P) := P + v = \alpha_P^{-1}(v).$$

In altre parole $\tau_v(P)$ è l'unico punto P' di A tale che $\overrightarrow{PP'} = v$. La funzione τ_v è detta *traslazione di vettore v* . Per ogni $P, Q \in A$, calcoliamo

$$\overrightarrow{\tau_v(P)\tau_v(Q)} = \overrightarrow{(P+v)(Q+v)} = \overrightarrow{(P+v)P} + \overrightarrow{PQ} + \overrightarrow{Q(Q+v)} = -v + \overrightarrow{PQ} + v = \overrightarrow{PQ},$$

dove abbiamo usato la proprietà triangolare. Questo mostra che la traslazione $\tau_v : A \rightarrow A$ è un'applicazione affine la cui parte lineare è la funzione identica $\text{id}_V : V \rightarrow V$. Viceversa, consideriamo un'applicazione affine $F : A \rightarrow A$ e supponiamo che la parte lineare $f_F : V \rightarrow V$ di F sia la funzione identica: $f_F = \text{id}_V$. Questo vuol dire che, per ogni $P, Q \in A$,

$$\overrightarrow{F(P)F(Q)} = \text{id}(\overrightarrow{PQ}) = \overrightarrow{PQ}. \quad (6.68)$$

Dalla (6.68) segue che il vettore $\overrightarrow{PF(P)}$ non dipende dalla scelta del punto P . In altre parole, per ogni coppia di punti $P, Q \in A$, si ha che $\overrightarrow{PF(P)} = \overrightarrow{QF(Q)}$, infatti

$$\overrightarrow{PF(P)} = \overrightarrow{PQ} + \overrightarrow{QF(Q)} + \overrightarrow{F(Q)F(P)} = \overrightarrow{PQ} + \overrightarrow{QF(Q)} + \overrightarrow{QP} = \overrightarrow{QF(Q)}.$$

Ora poniamo $v := \overrightarrow{PF(P)}$ (posso scegliere un qualunque punto $P \in A$ per definire il vettore v , e otterrò sempre lo stesso vettore, come abbiamo appena visto). Allora, per ogni punto $P \in A$,

$$F(P) = P + \overrightarrow{PF(P)} = P + v = \tau_v(P).$$

Poiché v non dipende da P , questo mostra che F è la traslazione di vettore v . Quindi le applicazioni affini $F : A \rightarrow A$ con parte lineare pari alla funzione identica sono esattamente le traslazioni. Si osservi che la traslazione di vettore nullo è la funzione identica $\text{id}_A : A \rightarrow A$. Perciò la funzione identica id_A è un'applicazione affine con parte lineare pari alla funzione identica $\text{id}_V : V \rightarrow V$. Invitiamo il lettore a dimostrare quest'ultima affermazione direttamente a partire dalla definizione di applicazione affine. ■

■ **Esempio 6.30** Fissiamo un punto $P_0 \in A$. Come sappiamo dall'Esempio (6.1), lo spazio vettoriale V può essere visto esso stesso come spazio affine di giacitura V con la differenza di vettori come funzione α . Se interpretiamo V come spazio affine in questo modo, allora la funzione

$$\alpha_{P_0} : A \rightarrow V$$

è un'applicazione affine con parte lineare pari alla funzione identica $\text{id} : V \rightarrow V$. Infatti, per ogni $P, Q \in A$,

$$\overrightarrow{\alpha_{P_0}(P)\alpha_{P_0}(Q)} = \alpha_{P_0}(Q) - \alpha_{P_0}(P) = \overrightarrow{P_0Q} - \overrightarrow{P_0P} = \overrightarrow{P_0Q} + \overrightarrow{PP_0} = \overrightarrow{PQ}.$$

■ **Esempio 6.31 — Applicazione Coordinata.** Fissiamo un riferimento affine (O, \mathcal{R}) di A . Come sappiamo, (O, \mathcal{R}) determina un'applicazione coordinata

$$c_{(O, \mathcal{R})} : A \rightarrow \mathbb{K}^n.$$

La Proposizione 6.2.1 mostra che $c_{(O, \mathcal{R})}$ è un'applicazione affine tra lo spazio affine A e lo spazio affine numerico $A_{\mathbb{K}}^n$, la cui parte lineare è l'isomorfismo coordinato

$$c_{\mathcal{R}} : V \rightarrow \mathbb{K}^n$$

(il lettore lo vede?). ■

■ **Esempio 6.32 — Applicazioni Affini tra Spazi Affini Numerici.** In questo esempio descriviamo tutte le applicazioni affini tra due spazi affini numerici $A_{\mathbb{K}}^n$ e $A_{\mathbb{K}}^m$. Specificamente, dimostriamo il seguente

Teorema 6.4.1 Per ogni matrice $A \in M_{m,n}(\mathbb{K})$ e per ogni vettore numerico $b \in \mathbb{K}^m$, la funzione $F_{(A,b)} : A_{\mathbb{K}}^n \rightarrow A_{\mathbb{K}}^m$ definita da

$$F_{(A,b)} : A_{\mathbb{K}}^n \rightarrow A_{\mathbb{K}}^m, \quad x \mapsto F_{(A,b)}(x) := Ax + b, \quad (6.69)$$

è un'applicazione affine la cui parte lineare è l'applicazione lineare

$$L_A : \mathbb{K}^n \rightarrow \mathbb{K}^m, \quad x \mapsto L_A(x) = Ax,$$

associata alla matrice A . Per ogni applicazione affine $F : A_{\mathbb{K}}^n \rightarrow A_{\mathbb{K}}^m$ esistono un'unica matrice $A \in M_{m,n}(\mathbb{K})$ e un unico vettore numerico $b \in \mathbb{K}^m$, tali che $F = F_{(A,b)}$.

Dimostrazione. Siano $A, b, F_{(A,b)}$ come nell'enunciato. Consideriamo $x, y \in A_{\mathbb{K}}^n = \mathbb{K}^n$ e calcoliamo

$$\overrightarrow{F_{(A,b)}(x)F_{(A,b)}(y)} = F_{(A,b)}(y) - F_{(A,b)}(x) = Ay + b - Ax - b = A(y - x) = L_A(\overrightarrow{xy}).$$

Questo dimostra la prima parte dell'enunciato. Per la seconda parte, consideriamo un'applicazione affine

$$F : A_{\mathbb{K}}^n \rightarrow A_{\mathbb{K}}^m$$

e sia $f : \mathbb{K}^n \rightarrow \mathbb{K}^m$ la parte lineare di F . Poiché f è un'applicazione lineare (tra spazi vettoriali numerici) esiste una matrice $A \in M_{m,n}(\mathbb{K})$ tale che $f = L_A$, cioè f è la funzione lineare associata ad A . Poniamo inoltre $b := F(0)$. Quindi, per ogni $x \in A_{\mathbb{K}}^n = \mathbb{K}^n$,

$$\begin{aligned} F(x) &= F(x) - b + b = F(x) - F(0) + b = \overrightarrow{F(0)F(x)} + b = L_A(\overrightarrow{0x}) + b = A(x - 0) + b \\ &= Ax + b. \end{aligned}$$

cioè $F = F_{(A,b)}$. Infine supponiamo di avere due matrici $A, A' \in M_{m,n}(\mathbb{K})$ e due vettori numerici $b, b' \in \mathbb{K}^m$ tali che $F = F_{(A,b)} = F_{(A',b')}$ e dimostriamo che $A = A'$ e $b = b'$. È così, infatti

$$b = F_{(A,b)}(0) = F_{(A',b')}(0) = b'.$$

Inoltre, da $F_{(A,b)} = F_{(A',b')}$ segue anche che le rispettive parti lineari coincidono. Ma la parte lineare di $F_{(A,b)}$ (risp. $F_{(A',b')}$) è L_A (risp. $L_{A'}$). Perciò $L_A = L_{A'}$ da cui anche $A = A'$. Questo conclude la dimostrazione del teorema. ■

Si osservi che il Teorema 6.4.1 stabilisce una biezione $(A, b) \mapsto F_{(A,b)}$ tra l'insieme $M_{m,n}(\mathbb{K}) \times \mathbb{K}^m$ delle coppie (matrice, vettore numerico) e l'insieme delle applicazioni affini da $A_{\mathbb{K}}^n$ a $A_{\mathbb{K}}^m$. ■

○ Più in generale, dati due spazi vettoriali V, W sul campo \mathbb{K} , si può dimostrare che, per ogni applicazione lineare $f : V \rightarrow W$ e per ogni vettore $w \in W$, l'applicazione $F_{(f,w)} : V \rightarrow W$ definita da $F_{(f,w)}(v) := f(v) + w$ è un'applicazione affine la cui parte lineare è f (qui stiamo interpretando gli spazi vettoriali V, W come spazi affini nel solito modo). Viceversa, per ogni applicazione affine $F : V \rightarrow W$ esistono un'unica applicazione lineare $f : V \rightarrow W$ ed un unico vettore $w \in W$ tali che $F = F_{(f,w)}$. Invitiamo il lettore a tentare la dimostrazione di queste ultime due affermazioni ispirandosi alla dimostrazione del Teorema 6.4.1.

Ora discutiamo le principali proprietà delle applicazioni affini. Cominciamo discutendo la composizione di applicazioni affini.

Proposizione 6.4.2 Siano A, A', A'' spazi affini sul campo \mathbb{K} di giaciture V, V', V'' rispettivamente, e siano

$$A \xrightarrow{F} A' \xrightarrow{G} A''$$

applicazioni affini con parti lineari

$$V \xrightarrow{f_F} V' \xrightarrow{f_G} V''.$$

Allora la funzione composta

$$A \xrightarrow{G \circ F} A''$$

è un'applicazione affine con parte lineare

$$V \xrightarrow{f_{G \circ F}} V''.$$

Dimostrazione. Per ogni $P, Q \in A$ calcoliamo

$$\begin{aligned} \overrightarrow{(G \circ F)(P)(G \circ F)(Q)} &= \overrightarrow{G(F(P))G(F(Q))} = f_G(\overrightarrow{F(P)F(Q)}) \\ &= f_G\left(f_F(\overrightarrow{PQ})\right) = (f_G \circ f_F)(\overrightarrow{PQ}). \end{aligned}$$

Poiché la composta di applicazioni lineari è un'applicazione lineare, questo conclude la dimostrazione. ■

■ **Esempio 6.33** Consideriamo uno spazio affine A sul campo \mathbb{K} modellato sullo spazio vettoriale V e siano $v, w \in V$ due vettori. Le traslazioni

$$A \xrightarrow{\tau_v} A \xrightarrow{\tau_w} A$$

sono applicazioni affini con parti lineari

$$V \xrightarrow{\text{id}_V} V \xrightarrow{\text{id}_V} V,$$

perciò anche la loro composizione

$$A \xrightarrow{\tau_w \circ \tau_v} A$$

è un'applicazione affine con parte lineare

$$V \xrightarrow{\text{id}_V} V,$$

cioè una traslazione. L'Esercizio 6.1 mostra ora che $\tau_w \circ \tau_v$ è la traslazione di vettore $v + w$ (il lettore lo vede?): $\tau_w \circ \tau_v = \tau_{v+w}$. ■

■ **Esempio 6.34** Consideriamo gli spazi affini numerici $A_{\mathbb{K}}^n, A_{\mathbb{K}}^m$ e $A_{\mathbb{K}}^p$ e due applicazioni affini:

$$A_{\mathbb{K}}^n \xrightarrow{F} A_{\mathbb{K}}^m \xrightarrow{G} A_{\mathbb{K}}^p.$$

La composta:

$$A_{\mathbb{K}}^n \xrightarrow{G \circ F} A_{\mathbb{K}}^p$$

è un'applicazione affine. Ora, in virtù del Teorema 6.4.1, esistono matrici $A \in M_{m,n}(\mathbb{K}), B \in M_{p,m}(\mathbb{K}), C \in M_{p,n}(\mathbb{C})$ e vettori numerici $b \in \mathbb{K}^m, c, d \in \mathbb{K}^p$ tali che $F = F_{(A,b)}, G = F_{(B,c)}$ e $G \circ F = F_{(C,d)}$. Vogliamo vedere in che modo la coppia (C, d) è legata alle coppie $(A, b), (B, c)$. A questo scopo, per ogni $x \in A_{\mathbb{K}}^n = \mathbb{K}^n$ calcoliamo

$$F_{(C,d)}(x) = G(F(x)) = G(Ax + b) = B(Ax + b) + c = (BA)x + Bb + c.$$

Questo conto mostra che $F_{(C,d)} = F_{(BA, Bb+d)}$. Per l'unicità della coppia (C, d) , abbiamo allora $C = BA, d = Bb + d$. Riassumendo

$$F_{(B,c)} \circ F_{(A,b)} = F_{(BA, Bb+c)}.$$

Similmente a quanto avviene per le applicazioni lineari tra spazi vettoriali, è possibile *rappresentare* un'applicazione affine tra spazi affini in coordinate affini, in base al seguente

Teorema 6.4.3 Siano A, A' spazi affini di dimensioni n, n' sul campo \mathbb{K} di giaciture V, V' rispettivamente. Sia (O, \mathcal{R}) un riferimento affine di A e sia (O', \mathcal{R}') un riferimento affine di A' . Data un'applicazione affine

$$F : A \rightarrow A'$$

esistono un'unica matrice $A = M_{n',n}(\mathbb{K})$ ed un unico vettore numerico $b \in \mathbb{K}^{n'}$ tali che, per ogni punto $P \in A$, dette $x = (x_1, \dots, x_n)^T \in \mathbb{K}^n$ le coordinate affini di P nel riferimento affine (O, \mathcal{R}) , e dette $y = (y_1, \dots, y_{n'})^T \in \mathbb{K}^{n'}$ le coordinate affini di $F(P)$ nel riferimento (O', \mathcal{R}') , si ha che

$$y = Ax + b. \quad (6.70)$$

In altre parole, l'immagine di un punto di coordinate affini x ha coordinate $Ax + b = F_{(A,b)}(x)$.

Dimostrazione. Per dimostrare l'esistenza di (A, b) ragioniamo come segue. Sia $f_F : V \rightarrow V'$ la parte lineare di F e sia $A = M_{\mathcal{R}', \mathcal{R}}^{f_F}$ la sua matrice rappresentativa nei riferimenti $\mathcal{R}, \mathcal{R}'$ di V, V' . Inoltre indichiamo con $b = (b_1, \dots, b_{n'})^T \in \mathbb{K}^{n'}$ le coordinate affini del punto $F(O) \in A'$ nel riferimento affine (O', \mathcal{R}') . Dobbiamo dimostrare la (6.70). A questo scopo consideriamo un punto $P \in A$ e siano $x = (x_1, \dots, x_n)^T \in \mathbb{K}^n$ le sue coordinate affini nel riferimento affine (O, \mathcal{R}) . Come nell'enunciato indichiamo con $y = (y_1, \dots, y_{n'}) \in \mathbb{K}^{n'}$ le coordinate affini di $F(P)$. Calcoliamo innanzitutto

$$\overrightarrow{O'F(P)} = \overrightarrow{O'F(O)} + \overrightarrow{F(O)F(P)} = \overrightarrow{O'F(O)} + f_F(\overrightarrow{OP}).$$

Poiché i vettori a primo ed ultimo membro coincidono, essi hanno le stesse coordinate (nel riferimento \mathcal{R}'). Ma, per definizione di coordinate affini, le coordinate del vettore $\overrightarrow{O'F(P)}$ nel riferimento \mathcal{R}' sono proprio le coordinate affini y del punto $F(P)$ nel riferimento affine (O', \mathcal{R}') . Similmente, le coordinate del vettore $\overrightarrow{O'F(O)}$ sono le coordinate affini b di $F(O)$. Infine, per definizione di matrice rappresentativa di un'applicazione lineare, le coordinate di $f_F(\overrightarrow{OP})$ sono esattamente Ax (infatti x sono le coordinate affini di P , cioè le coordinate di \overrightarrow{OP} e A è la matrice rappresentativa di f_F). Riassumendo, deve essere

$$y = b + Ax = Ax + b,$$

come desiderato. L'unicità della coppia (A, b) si può verificare come segue. Supponiamo che $A' \in M_{n',n}(\mathbb{K})$ e $b' \in \mathbb{K}^{n'}$ siano un'altra matrice e un altro vettore numerico tali che le coordinate affini di $F(P)$ nel riferimento affine (O', \mathcal{R}') sono date da $A'x + b'$ per ogni $P \in A$ (dove, come al solito, x sono le coordinate affini di P nel riferimento affine (O, \mathcal{R})). Dall'arbitrarietà di P segue che

$$Ax + b = A'x + b', \quad \text{per ogni } x \in \mathbb{K}^n.$$

Ma questo vuol dire che le applicazioni affini $F_{(A,b)} : A_{\mathbb{K}}^n \rightarrow A_{\mathbb{K}}^{n'}$ e $F_{(A',b')} : A_{\mathbb{K}}^n \rightarrow A_{\mathbb{K}}^{n'}$ coincidono. Dunque, dal Teorema 6.4.1, deve essere $A = A'$ e $b = b'$. ■

La coppia (A, b) associata all'applicazione affine F dal Teorema 6.4.3 è anche detta, per ovvi motivi, la *rappresentazione coordinata* di F nei riferimenti $(O, \mathcal{R}), (O', \mathcal{R}')$. Si osservi che il Teorema 6.4.3 dice che la rappresentazione coordinata (A, b) dell'applicazione affine $F : A \rightarrow A'$ è l'unica coppia (matrice, vettore numerico) per cui il diagramma

$$\begin{array}{ccc} A & \xrightarrow{F} & A' \\ c_{(O,\mathcal{R})} \downarrow & & \downarrow c_{(O',\mathcal{R}')} \\ A_{\mathbb{K}}^n & \xrightarrow{F_{(A,b)}} & A_{\mathbb{K}}^{n'} \end{array}$$

commuta, cioè $c_{(O',\mathcal{R}')} \circ F = F_{(A,b)} \circ c_{(O,\mathcal{R})}$ (il lettore lo vede?).

Esercizio 6.11 Siano A e A' spazi affini di giaciture V, V' , siano $S \subseteq A$ e $S' \subseteq A'$ sottospazi affini di giaciture $W \subseteq V$ e $W' \subseteq V'$, e sia $F : A \rightarrow A'$ un'applicazione affine tale che $F(S) \subseteq S'$. Dimostrare che $f_F(W) \subseteq W'$ e la restrizione $F|_S^S : S \rightarrow S'$ di F ad S nel dominio e ad S' nel codominio è ancora un'applicazione affine con parte lineare data dalla restrizione $f_F|_W^W : W \rightarrow W'$ di f_F a W nel dominio e a W' nel codominio. ■

La seguente caratterizzazione delle applicazioni affini, torna spesso utile.

Proposizione 6.4.4 Siano A e A' spazi affini di giaciture V, V' , sia $F : A \rightarrow A'$ una funzione e sia $f : V \rightarrow V'$ un'applicazione lineare. Allora le seguenti affermazioni sono equivalenti:

- (1) $F : A \rightarrow A'$ è un'applicazione affine con parte lineare pari ad f ;
- (2) esiste un punto $\bar{P} \in A$ tale che, per ogni $Q \in A$ si ha che $\overrightarrow{F(\bar{P})F(Q)} = f(\overrightarrow{\bar{P}Q})$.

Dimostrazione. Che (1) \Rightarrow (2) è ovvio. Dimostriamo che (2) \Rightarrow (1). A questo scopo calcoliamo

$$\begin{aligned}
 \overrightarrow{F(P)F(Q)} &= \overrightarrow{F(P)F(\bar{P})} + \overrightarrow{F(\bar{P})F(Q)} && \text{(proprietà triangolare)} \\
 &= \overrightarrow{-F(\bar{P})F(P)} + \overrightarrow{F(\bar{P})F(Q)} \\
 &= -f(\overrightarrow{\bar{P}P}) + f(\overrightarrow{\bar{P}Q}) && \text{(ipotesi (2))} \\
 &= f(-\overrightarrow{\bar{P}P} + \overrightarrow{\bar{P}Q}) && \text{(} f \text{ è un'applicazione lineare)} \\
 &= f(\overrightarrow{\bar{P}P} + \overrightarrow{\bar{P}Q}) \\
 &= f(\overrightarrow{PQ}) && \text{(proprietà triangolare).}
 \end{aligned}$$

Questo conclude la dimostrazione. ■

Osserviamo, per futuro utilizzo, che, fissato il punto $P \in A$, la condizione

$$\overrightarrow{F(P)F(Q)} = f(\overrightarrow{PQ}) \quad \text{per ogni } Q \in A$$

può essere scritta in molti modi equivalenti. Innanzitutto equivale a dire che il diagramma

$$\begin{array}{ccc}
 A & \xrightarrow{F} & A' \\
 \alpha_P \downarrow & & \downarrow \alpha_{F(P)} \\
 V & \xrightarrow{f} & V'
 \end{array} \tag{6.71}$$

commuta, cioè

$$\alpha_{F(P)} \circ F = f \circ \alpha_P. \tag{6.72}$$

Infatti, per ogni $Q \in A$

$$(\alpha_{F(P)} \circ F)(Q) = \alpha_{F(P)}(F(Q)) = \overrightarrow{F(P)F(Q)},$$

il che coincide con

$$(f \circ \alpha_P)(Q) = f(\alpha_P(Q)) = f(\overrightarrow{PQ})$$

esse $\overrightarrow{F(P)F(Q)} = f(\overrightarrow{PQ})$ come desiderato. A sua volta l'identità $\alpha_{F(P)} \circ F = f \circ \alpha_P$ può essere riscritta nelle forme

$$F = \alpha_{F(P)}^{-1} \circ f \circ \alpha_P, \quad (6.73)$$

$$F \circ \alpha_P^{-1} = \alpha_{F(P)}^{-1} \circ f, \quad (6.74)$$

$$\alpha_{F(P)} \circ F \circ \alpha_P^{-1} = f. \quad (6.75)$$

Per esempio, la (6.73) si ottiene dalla (6.72) componendo ambo i membri a sinistra con $\alpha_{F(P)}^{-1}$ (il lettore lo vede?). La (6.74) e la (6.75) si ottengono in modo simile. Sottolineiamo che le identità (6.72)–(6.75) sono tutte soddisfatte per ogni P da un'applicazione affine F e dalla sua parte lineare f . Viceversa se un'applicazione $F : A \rightarrow A'$ soddisfa una qualsiasi delle (6.72)–(6.75) (e quindi tutte le altre) per qualche punto $P \in A$ (ne basta uno solo) e qualche applicazione lineare $f : V \rightarrow V'$, allora F è un'applicazione affine con parte lineare pari ad f .

Applichiamo ora la Proposizione 6.4.4, discutendo come un'applicazione affine interagisce con i sottospazi affini del suo dominio e del suo codominio. Vale la

Proposizione 6.4.5 Sia $F : A \rightarrow A'$ un'applicazione affine tra spazi affini con parte lineare $f : V \rightarrow V'$, e siano $S \subseteq A$ ed $S' \subseteq A'$ sottospazi affini di giacitura $W \subseteq V$ e $W' \subseteq V'$. Allora

- (1) $F(S) \subseteq A'$ è un sottospazio affine di giacitura $f(W) \subseteq V'$;
- (2) $F^{-1}(S') \subseteq A$ o è vuoto, oppure è un sottospazio affine di giacitura $f^{-1}(W')$.

Dimostrazione. Sia $\bar{P} \in S$. Allora $S = \alpha_{\bar{P}}^{-1}(W)$ e, poiché F è un'applicazione affine, possiamo usare la (6.74) per trovare

$$F(S) = F(\alpha_{\bar{P}}^{-1}(W)) = \alpha_{F(\bar{P})}^{-1}(f(W))$$

che è appunto un sottospazio affine di giacitura $f(W)$ come desiderato. Questo dimostra il punto (1).

Il punto (2) richiede un'argomentazione leggermente più articolata. Innanzitutto, assumiamo che $F^{-1}(S') \neq \emptyset$. Questo significa che esiste $P \in A$ tale che $F(P) \in S'$. Allora $S' = \alpha_{F(P)}^{-1}(W')$. Per calcolare la controimmagine $F^{-1}(S')$ osserviamo che un punto Q di A appartiene a $F^{-1}(S')$ sse

$$\begin{aligned} F(Q) &\in S' = \alpha_{F(P)}^{-1}(W') \\ \Leftrightarrow \alpha_{F(P)}(F(Q)) &\in W' \\ \Leftrightarrow f(\alpha_P(Q)) &\in W' && \text{(Identità (6.72))} \\ \Leftrightarrow \alpha_P(Q) &\in f^{-1}(W') \\ \Leftrightarrow Q &\in \alpha_P^{-1}(f^{-1}(W')), \end{aligned}$$

da cui $F^{-1}(S') = \alpha_P^{-1}(f^{-1}(W'))$ che è un sottospazio affine di giacitura $f^{-1}(W')$ (in pratica il conto appena effettuato mostra che la controimmagine $(g \circ h)^{-1}(B)$ di un sottoinsieme B mediante la composta $g \circ h$ di due funzioni h e g è la "composta" $h^{-1}(g^{-1}(B))$ delle controimmagini: $(g \circ h)^{-1}(B) = h^{-1}(g^{-1}(B))$ - il lettore è invitato ad identificare il sottoinsieme B e le due funzioni h e g nel presente caso). ■

■ **Esempio 6.35** Sia $f : V \rightarrow V'$ un'applicazione lineare tra spazi vettoriali. Sia V che V' possono essere riguardati come spazi affini nel solito modo. Se $v'_0 \in V'$ è un punto fissato, la sua controimmagine $f^{-1}(v'_0) \subseteq V$ non è un sottospazio vettoriale di V a meno che non sia $v'_0 = 0$ (nel qual caso $f^{-1}(v'_0) = \ker f$), tuttavia, se non è vuota, è sempre un sottospazio affine. Infatti, innanzitutto f è un'applicazione affine con parte lineare f stessa: per ogni $v, w \in V$

$$\overrightarrow{f(v)f(w)} = f(w) - f(v) = f(w - v) = f(\overrightarrow{vw}).$$

Inoltre il singleton $\{v'_0\} \subseteq V'$ è un sottospazio affine di dimensione 0. ■

Esercizio 6.12 Dimostrare che immagine e controimmagine mediante applicazioni affini conservano il parallelismo di sottospazi nel senso che segue. Se $F : A \rightarrow A'$ è un'applicazione affine tra spazi affini, ed $S, T \subseteq A$ ed $S', T' \subseteq A'$ sono sottospazi affini tali che S, T sono paralleli ed S', T' sono paralleli, allora

- (1) $F(S), F(T)$ sono sottospazi affini paralleli di A' ;
- (2) $F^{-1}(S'), F^{-1}(T')$, se entrambi non-vuoti, sono sottospazi affini paralleli di A .

Un'altra applicazione della Proposizione 6.4.4 è il seguente importante

Teorema 6.4.6 — Teorema dell'Estensione Affine. Siano A e A' spazi affini modellati sugli spazi vettoriali V, V' . Per ogni coppia di punti $O \in A, O' \in A'$ e per ogni applicazione lineare $f : V \rightarrow V'$ esiste un'unica applicazione affine $F : A \rightarrow A'$ tale che $F(O) = O'$ e inoltre $f_F = f$, cioè la parte lineare di F è proprio f .

Dimostrazione. Siano O, O', f come nell'enunciato. Definiamo $F : A \rightarrow A'$ come la composta delle seguenti funzioni

$$A \xrightarrow{\alpha_O} V \xrightarrow{f} V' \xrightarrow{\alpha_{O'}^{-1}} A',$$

cioè

$$F = \alpha_{O'}^{-1} \circ f \circ \alpha_O. \quad (6.76)$$

In altre parole, per ogni $P \in A$,

$$F(P) = \alpha_{O'}^{-1}(f(\alpha_O(P))) = O' + f(\alpha_O(P)) = O' + f(\overrightarrow{OP}).$$

In particolare

$$F(O) = O' + f(\overrightarrow{OO}) = O' + f(0) = O' + 0 = O'$$

come richiesto. Inoltre, poiché $O' = F(O)$, la F soddisfa l'Identità (6.73) con $P = O$, perciò è anche un'applicazione affine con parte lineare pari ad f , dunque è proprio l'applicazione affine che stiamo cercando.

Per dimostrare l'unicità, supponiamo che $F' : A \rightarrow A'$ sia un'altra applicazione affine tale che $F'(O) = O'$ e $f_{F'} = f$. Allora F' soddisfa la (6.73) con $P = O$, cioè

$$F' = \alpha_{F'(O)}^{-1} \circ f \circ \alpha_O = \alpha_{O'}^{-1} \circ f \circ \alpha_O = F,$$

come desiderato. ■

- Il Teorema dell'Estensione Affine ha un'altra interessante formulazione, più simile al Teorema dell'Estensione Lineare, in termini di riferimenti affini di punti (Osservazione a pag. 144). Innanzitutto si ricordi l'enunciato del Teorema dell'Estensione Lineare: *Siano V e V' spazi vettoriali, e sia $\mathcal{R} = (v_1, \dots, v_n)$ un riferimento di V , $n = \dim V$. Per ogni n -upla (v'_1, \dots, v'_n) di vettori di V' esiste un'unica applicazione lineare $f : V \rightarrow V'$ tale che $f(v_i) = v'_i$, per ogni $i = 1, \dots, n$. Similmente vale il seguente corollario del Teorema dell'Estensione Affine:*

Teorema 6.4.7 Siano A, A' spazi affini modellati sugli spazi vettoriali V, V' , e sia

$$(O, P_1, \dots, P_n)$$

un riferimento affine di punti di A . Per ogni $(n+1)$ -upla (O', P'_1, \dots, P'_n) di punti di A' esiste un'unica applicazione affine $F : A \rightarrow A'$ tale che $F(O) = O'$ e $F(P_i) = P'_i$ per ogni

$$i = 1, \dots, n.$$

Dimostrazione. Per definizione di riferimento affine di punti, il sistema $(\overrightarrow{OP_1}, \dots, \overrightarrow{OP_n})$ è un riferimento di V . Per il Teorema dell'Estensione Lineare esiste un'unica applicazione lineare $f : V \rightarrow V'$ tale che $f(\overrightarrow{OP_i}) = \overrightarrow{O'P'_i}$ per ogni $i = 1, \dots, n$ e, per il Teorema dell'Estensione Affine, esiste un'unica applicazione affine $F : A \rightarrow A'$ tale che $F(O) = O'$ e inoltre $f_F = f$. L'applicazione affine F è proprio quella che stiamo cercando, infatti, dalla dimostrazione del Teorema dell'Estensione Affine, per ogni $P \in A$, $F(P) = O' + f(\overrightarrow{OP})$. In particolare

$$F(P_i) = O' + f(\overrightarrow{OP_i}) = O' + \overrightarrow{O'P'_i} = P'_i,$$

per ogni $i = 1, \dots, n$, come desiderato. Infine, se $F' : A \rightarrow A'$ è un'altra applicazione affine tale che $F'(O) = O'$ e $F'(P_i) = P'_i$, per ogni $i = 1, \dots, n$, allora si ha

$$f_{F'}(\overrightarrow{OP_i}) = \overrightarrow{F'(O)F'(P_i)} = \overrightarrow{O'P'_i},$$

da cui, per l'unicità nel Teorema dell'Estensione Lineare, $f_{F'} = f$ e, per l'unicità nel Teorema dell'Estensione Affine, $F' = F$. ■

A titolo di esempio menzioniamo che l'unica applicazione affine che manda tutti i punti di un riferimento affine di punti di A nello stesso punto $P'_0 \in A'$, è l'applicazione affine costante $c_{P'_0}$ (il lettore lo vede?).

Come ultima applicazione della Proposizione 6.4.4 caratterizziamo le applicazioni affini iniettive, suriettive e biettive, mostrando come l'iniettività (risp. la suriettività, la biettività) di un'applicazione affine dipenda solo dall'iniettività (risp. la suriettività, la biettività) della sua parte lineare.

Proposizione 6.4.8 Sia $F : A \rightarrow A'$ un'applicazione affine tra spazi affini e sia $f : V \rightarrow V'$ la sua parte lineare. Allora F è iniettiva, risp. suriettiva, biettiva, sse f è iniettiva, risp. suriettiva, biettiva.

Dimostrazione. Fissiamo un punto $P \in A$. Vale la (6.73), da cui segue immediatamente che, se f è iniettiva (risp. suriettiva, biettiva) allora anche F è iniettiva (risp. suriettiva, biettiva), ricordando che la composta di funzioni iniettive (risp. suriettive, biettive) è una funzione iniettiva (risp. suriettiva, biettiva). Infine, dalla (6.75), analogamente a sopra, se F è iniettiva (risp. suriettiva, biettiva), anche f lo è. Questo conclude la dimostrazione. ■

Corollario 6.4.9 Sia $F : A \rightarrow A'$ un'applicazione affine iniettiva (risp. suriettiva, biettiva). Allora $\dim A \leq \dim A'$ (risp. $\dim A \geq \dim A'$, $\dim A = \dim A'$).

6.5 Affinità

Siano A, A' spazi affini sul campo \mathbb{K} modellati sugli spazi vettoriali V, V' .

Definizione 6.5.1 — Affinità. Un'affinità (tra A e A') è un'applicazione affine biettiva $\Phi : A \rightarrow A'$. Due spazi affini A e A' collegati da un'affinità $\Phi : A \rightarrow A'$ sono detti *isomorfi*.

■ **Esempio 6.36** Sia A uno spazio affine modellato sullo spazio vettoriale V , e sia $v \in V$ un vettore. La traslazione $\tau_v : A \rightarrow A$ è un'affinità, infatti è invertibile e la sua inversa è la traslazione τ_{-v} (il lettore lo vede?). In alternativa possiamo dedurre che τ_v è biettiva dalla Proposizione 6.4.8. Infatti la sua parte lineare è $\text{id}_V : V \rightarrow V$ che è biettiva. ■

■ **Esempio 6.37** Sia $P_0 \in A$ un punto. La funzione $\alpha_{P_0} : A \rightarrow V$ è un'affinità (quando si interpreta anche V come spazio affine nel solito modo) infatti è un'applicazione affine (Esempio 6.30) ed è biettiva. ■

■ **Esempio 6.38** Sia (O, \mathcal{R}) un riferimento affine di A . L'applicazione coordinata $c_{(O, \mathcal{R})} : A \rightarrow \mathbb{K}^n = A_{\mathbb{K}}^n$   un'affinit . Infatti   un'applicazione affine ed   biettiva. Quindi ogni spazio affine n -dimensionale sul campo \mathbb{K}   isomorfo allo spazio affine numerico n -dimensionale $A_{\mathbb{K}}^n$.

Si osservi anche che ogni affinit  $A \rightarrow A_{\mathbb{K}}^n$   un'applicazione coordinata cio , per ogni affinit  $\Phi : A \rightarrow A_{\mathbb{K}}^n$, esiste un riferimento affine (O, \mathcal{R}) di A tale che $\Phi = c_{(O, \mathcal{R})}$. Infatti, data l'affinit  Φ poniamo

$$O := \Phi^{-1}(0) \quad \text{e} \quad \mathcal{R} = f_{\Phi}^{-1}(\mathcal{R}_{can}).$$

Poich  $f_{\Phi} : V \rightarrow \mathbb{K}^n$   un isomorfismo di spazi vettoriali, \mathcal{R}   un riferimento di V e dunque (O, \mathcal{R})   un riferimento affine di A . Inoltre, le coordinate dell' i -esimo vettore di \mathcal{R} nel riferimento \mathcal{R} stesso sono

$$(0, \dots, \underbrace{1}_{\text{posto } i}, \dots, 0),$$

e quindi

$$c_{\mathcal{R}}(\mathcal{R}) = \mathcal{R}_{can} = f_{\Phi}(\mathcal{R}).$$

Segue allora dal Teorema dell'Estensione Lineare che $c_{\mathcal{R}} = f_{\Phi}$ (coincidono sugli elementi di base). Infine

$$c_{(O, \mathcal{R})}(O) = \overrightarrow{OO} = 0 = \Phi(O).$$

Dunque le applicazioni affini $c_{(O, \mathcal{R})}$ e Φ hanno la stessa parte lineare e coincidono sul punto O . Per il Teorema dell'Estensione affine coincidono, come annunciato. ■

■ **Esempio 6.39** Consideriamo lo spazio affine numerico $A_{\mathbb{K}}^n$ e un'applicazione affine

$$\Phi : A_{\mathbb{K}}^n \rightarrow A_{\mathbb{K}}^n.$$

Come sappiamo $\Phi = F_{(A, b)}$ per qualche matrice $A \in M_n(\mathbb{K})$ e qualche vettore numerico $b \in \mathbb{K}^n$. Allora Φ   biettiva sse A   una matrice invertibile. Infatti, dalla Proposizione (6.4.8), Φ   biettiva sse f_{Φ}   biettiva, ma $f_{\Phi} : \mathbb{K}^n \rightarrow \mathbb{K}^n$   l'applicazione lineare associata alla matrice A la quale   biettiva sse A   una matrice invertibile. ■

Proposizione 6.5.1 Due spazi affini A e A' sullo stesso campo \mathbb{K} sono isomorfi sse hanno la stessa dimensione.

Dimostrazione. Se A e A' sono isomorfi, allora hanno anche la stessa dimensione in virt  del Corollario 6.4.9. Viceversa, se A e A' hanno la stessa dimensione, possiamo costruire un'affinit  $\Phi : A \rightarrow A'$ per esempio come segue: consideriamo un punto $O \in A$ un punto $O' \in A'$ e un isomorfismo lineare $f : V \rightarrow V'$ tra le giaciture (quest'ultimo esiste perch  V e V' sono spazi vettoriali della stessa dimensione). Il Teorema dell'Estensione affine garantisce ora che esiste un'applicazione affine $\Phi : A \rightarrow A'$ tale che $\Phi(O) = O'$ e $f_{\Phi} = f$. Poich  f   biettiva, anche Φ lo  . Dunque Φ   un'affinit . ■

Proposizione 6.5.2 La composta $\Psi \circ \Phi$ di due affinit 

$$A \xrightarrow{\Phi} A' \xrightarrow{\Psi} A''$$

  un'affinit  con parte lineare data da $f_{\Psi} \circ f_{\Phi}$. L'inversa di un'affinit 

$$A \xrightarrow{\Phi} A'$$

è un'affinità con parte lineare data da f_{Φ}^{-1} .

Dimostrazione. La prima parte dell'enunciato segue dalla Proposizione 6.4.2 e dal fatto che la composta di funzioni biettive è una funzione biettiva. Per la seconda parte, innanzitutto ricordiamo che, poiché Φ è biettiva, anche la sua parte lineare f_{Φ} lo è. Dobbiamo fare vedere che Φ^{-1} è un'applicazione affine con parte lineare f_{Φ}^{-1} . A questo scopo, consideriamo $P', Q' \in A'$ e calcoliamo

$$\begin{aligned} \overrightarrow{\Phi^{-1}(P')\Phi^{-1}(Q')} &= f_{\Phi}^{-1}\left(\overrightarrow{f_{\Phi}(\Phi^{-1}(P')\Phi^{-1}(Q'))}\right) \\ &= f_{\Phi}^{-1}\left(\overrightarrow{\Phi(\Phi^{-1}(P'))\Phi(\Phi^{-1}(Q'))}\right) \\ &= f_{\Phi}^{-1}\left(\overrightarrow{P'Q'}\right). \end{aligned}$$

■

■ **Esempio 6.40** Sia A uno spazio affine di dimensione n e giacitura V sul campo \mathbb{K} , e siano $(O, \mathcal{R}), (O', \mathcal{R}')$ riferimenti affini di (A, V, α) . Le applicazioni coordinate

$$A_{\mathbb{K}}^n \xleftarrow{c_{(O, \mathcal{R})}} A \xrightarrow{c_{(O', \mathcal{R}')}} A_{\mathbb{K}}^n$$

sono affinità. Perciò anche la composta

$$\Phi := c_{(O', \mathcal{R}')} \circ c_{(O, \mathcal{R})}^{-1} : A_{\mathbb{K}}^n \rightarrow A_{\mathbb{K}}^n$$

è un'affinità (detta *cambio di coordinate affini*, o anche *mappa di transizione* dal riferimento (O, \mathcal{R}) al riferimento (O', \mathcal{R}')) che trasforma le coordinate affini di un punto P di A nel riferimento affine (O, \mathcal{R}) nelle coordinate affini dello stesso punto nel riferimento (O', \mathcal{R}') . La parte lineare di Φ è chiaramente l'applicazione di cambio di coordinate (nello spazio vettoriale V)

$$c_{\mathcal{R}'} \circ c_{\mathcal{R}}^{-1} : \mathbb{K}^n \rightarrow \mathbb{K}^n.$$

Dal Teorema 6.4.1 segue che $\Phi = F_{(A, b)}$, dove A è esattamente la matrice del cambio di coordinate dal riferimento \mathcal{R} al riferimento \mathcal{R}' nello spazio vettoriale V (il lettore lo vede?), in particolare si tratta di una matrice invertibile. ■

Un'affinità $\Phi : A \rightarrow A'$ consente di identificare gli spazi affini A e A' ed ogni costruzione nei due che dipenda solo dalla struttura di spazio affine. Per esempio un'affinità trasforma sottospazi affini in sottospazi affini della stessa dimensione e ne conserva anche le posizioni reciproche (il lettore lo vede?). In buona sostanza, due spazi affini collegati da una affinità andrebbero identificati e interpretati come "lo stesso spazio affine" (a meno di una mera rietichettatura dei punti). In questo spirito, ogni spazio affine può essere identificato con uno spazio affine numerico tramite l'applicazione coordinata determinata da un fissato riferimento affine. Discutiamo ora brevemente di affinità da uno spazio affine in sé stesso.

Proposizione / Definizione 6.5.3 Sia A uno spazio affine. Le affinità $\Phi : A \rightarrow A$ (da A in sé stesso), con la composizione di funzioni, formano un gruppo, detto *gruppo delle affinità*, o *gruppo affine*, dello spazio affine A e indicato con $\text{Aff}(A)$.

Dimostrazione. L'enunciato segue immediatamente dal fatto che la funzione identica $\text{id}_A : A \rightarrow A$ è un'affinità e dalla Proposizione 6.5.2. ■

Proposizione / Definizione 6.5.4 Sia A uno spazio affine. Le traslazioni di A formano un sottogruppo abeliano del gruppo $\text{Aff}(A)$ delle affinità, detto *gruppo delle traslazioni* di A .

Dimostrazione. Innanzitutto, ogni traslazione è un'affinità. Inoltre sappiamo anche che

- ✓ la funzione identica $\text{id}_A : A \rightarrow A$ è una traslazione;
- ✓ la composta di due traslazioni è una traslazione;
- ✓ l'inversa di una traslazione è una traslazione.

Questo mostra che le traslazioni formano un sottogruppo di $\text{Aff}(A)$. Per dimostrare che si tratta di un sottogruppo abeliano, ricordiamo che per ogni $v, w \in V$ si ha

$$\tau_w \circ \tau_v = \tau_{v+w} = \tau_{w+v} = \tau_v \circ \tau_w.$$

Questo conclude la dimostrazione. ■

- Sia A uno spazio affine modellato sullo spazio vettoriale V . In particolare l'insieme V , munito della sola somma di vettori, è un gruppo abeliano. La funzione

$$\tau : V \rightarrow \text{Aff}(A), \quad v \mapsto \tau_v$$

è un omomorfismo di gruppi la cui immagine $\tau(V)$ è esattamente il gruppo delle traslazioni di A . Questo segue dalle seguenti tre formule

- ✓ $\tau_0 = \text{id}_A$;
- ✓ $\tau_{v+w} = \tau_v \circ \tau_w$;
- ✓ $\tau_{-v} = \tau_v^{-1}$.

per ogni $v, w \in V$. Inoltre si tratta di un omomorfismo iniettivo. Per verificarlo è sufficiente verificare che il suo nucleo consiste del solo elemento neutro che, in questo caso, è il vettore nullo $0 \in V$. È così, infatti l'unica traslazione pari alla funzione identica è la traslazione di vettore nullo (il lettore provi a dimostrarlo in tutti i dettagli). In particolare la restrizione $\tau|_{\tau(V)} : V \rightarrow \tau(V)$ di τ alla sua immagine $\tau(V)$ nel codominio è un isomorfismo di gruppi che possiamo usare per identificare il gruppo abeliano $(V, +)$ con il gruppo delle traslazioni $\tau(V)$. Adesso consideriamo la funzione

$$f : \text{Aff}(A) \rightarrow GL(V), \quad \Phi \mapsto f_\Phi.$$

È ben definita, perché la parte lineare di un'affinità è un isomorfismo di spazi vettoriali. Inoltre è un omomorfismo di gruppi. Questo segue dalle formule

- ✓ $f_{\text{id}_A} = \text{id}_V$;
- ✓ $f_{\Psi \circ \Phi} = f_\Psi \circ f_\Phi$;
- ✓ $f_{\Phi^{-1}} = f_\Phi^{-1}$;

per ogni $\Phi, \Psi \in \text{Aff}(A)$. Inoltre, in virtù del Teorema di Estensione Affine, si tratta di un omomorfismo suriettivo. Il nucleo $\ker f \subseteq \text{Aff}(A)$ è esattamente il gruppo delle traslazioni. Infatti un'affinità $\Phi : A \rightarrow A$ è una traslazione sse $f_\Phi = \text{id}_V$. Riassumendo, nella sequenza

$$1 \longrightarrow V \xrightarrow{\tau} \text{Aff}(A, V, \alpha) \xrightarrow{f} GL(V) \longrightarrow 1 \quad (6.77)$$

l'omomorfismo $\tau : V \rightarrow \text{Aff}(A)$ è iniettivo, l'omomorfismo $f : \text{Aff}(A) \rightarrow GL(V)$ è suriettivo, e inoltre $\text{Im } \tau = \ker f$. Si dice anche, a questo proposito, che la sequenza (6.77) è *esatta*. In particolare il gruppo delle traslazioni $\tau(V)$ è un sottogruppo normale di $\text{Aff}(A)$ e, dal Primo Teorema di Omomorfismo per i gruppi, il gruppo lineare generale $GL(V)$ è isomorfo al gruppo quoziente $\text{Aff}(A)/\tau(V)$. Più precisamente, l'epimorfismo $f : \text{Aff}(A) \rightarrow GL(V)$ induce un isomorfismo $\text{Aff}(A)/\tau(V) \rightarrow GL(V)$, $\Phi \circ \tau(V) \mapsto f_\Phi$.

- Consideriamo lo spazio affine numerico n -dimensionale $A_{\mathbb{K}}^n$ sul campo \mathbb{K} . La funzione

$$F : GL_n(\mathbb{K}) \times \mathbb{K}^n \rightarrow \text{Aff}(A_{\mathbb{K}}^n), \quad (A, b) \mapsto F_{(A,b)}$$

è una ben definita biezione. È possibile dotare l'insieme $GL_n(\mathbb{K}) \times \mathbb{K}^n$ di un'operazione interna \cdot in modo che

- (1) $(GL_n(\mathbb{K}) \times \mathbb{K}^n, \cdot)$ diventi un gruppo,
 (2) $F : GL_n(\mathbb{K}) \times \mathbb{K}^n \rightarrow \text{Aff}(A_{\mathbb{K}}^n)$ diventi un isomorfismo di gruppi,

come segue. Per ogni $(A, b), (B, c) \in GL_n(\mathbb{K}) \times \mathbb{K}^n$ poniamo

$$(B, c) \cdot (A, b) = (BA, Bb + c).$$

Invitiamo il lettore a dimostrare che questa operazione è associativa e ammette elemento neutro, ed è la coppia $(I_n, 0)$. Inoltre ogni elemento $(A, b) \in GL_n(\mathbb{K}) \times \mathbb{K}^n$ è simmetrizzabile rispetto a \cdot e il suo simmetrico è l'elemento

$$(A, b)^{-1} := (A^{-1}, -A^{-1}b).$$

Dunque, $(GL_n(\mathbb{K}) \times \mathbb{K}^n, \cdot)$ è un gruppo che si indica anche $GL_n(\mathbb{K}) \ltimes \mathbb{K}^n$. Inoltre

- ✓ $F_{(I_n, 0)} = \text{id}_{A_{\mathbb{K}}^n}$;
- ✓ $F_{(B, c) \cdot (A, b)} = F_{(B, c)} \circ F_{(A, b)}$;
- ✓ $F_{(A, b)^{-1}} = F_{(A, b)}^{-1}$;

per ogni $(A, b), (B, c) \in GL_n(\mathbb{K}) \ltimes \mathbb{K}^n$ (il lettore dimostri per esercizio queste proprietà quando non già discusse in precedenza). Questo mostra che

$$F : GL_n(\mathbb{K}) \ltimes \mathbb{K}^n \rightarrow \text{Aff}(A_{\mathbb{K}}^n)$$

è un isomorfismo di gruppi come annunciato. Si può anche identificare il gruppo $GL_n(\mathbb{K}) \ltimes \mathbb{K}^n$ con un sottogruppo del gruppo lineare generale $GL_{n+1}(\mathbb{K})$ come segue. Consideriamo la funzione

$$H : GL_n(\mathbb{K}) \ltimes \mathbb{K}^n \rightarrow GL_{n+1}(\mathbb{K}), \quad (A, b) \mapsto H(A, b) := \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix}.$$

Si tratta chiaramente di una ben definita funzione iniettiva (in particolare la matrice $H(A, b)$ è invertibile per ogni $(A, b) \in GL_n(\mathbb{K}) \ltimes \mathbb{K}^n$). Inoltre è facile verificare che

- ✓ $H(I_n, 0) = I_{n+1}$;
- ✓ $H((B, c) \cdot (A, b)) = H(B, c)H(A, b)$;
- ✓ $H((A, b)^{-1}) = H(A, b)^{-1}$,

per ogni $(A, b), (B, c) \in GL_n(\mathbb{K}) \ltimes \mathbb{K}^n$, cioè H è un omomorfismo iniettivo di gruppi. Quindi l'immagine

$$H(GL_n(\mathbb{K}) \ltimes \mathbb{K}^n) \subseteq GL_{n+1}(\mathbb{K})$$

è un sottogruppo del gruppo lineare generale $GL_{n+1}(\mathbb{K})$, e precisamente il sottogruppo che consiste delle matrici invertibili $(n+1) \times (n+1)$ del tipo

$$\begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix},$$

equivalentemente, le matrici invertibili in cui l'ultima riga è $(0, \dots, 0, 1)$. Tale sottogruppo prende il nome di *gruppo affine* di ordine n sul campo \mathbb{K} e si indica con il simbolo $\text{Aff}_n(\mathbb{K})$. Inoltre la restrizione $\bar{H} := H|_{\text{Aff}_n(\mathbb{K})} : GL_n(\mathbb{K}) \ltimes \mathbb{K}^n \rightarrow \text{Aff}_n(\mathbb{K})$ del monomorfismo H alla sua immagine $\text{Aff}_n(\mathbb{K})$ nel codominio è un isomorfismo di gruppi. Componendo gli isomorfismi di gruppi

$$\text{Aff}_n(\mathbb{K}) \xrightarrow{\bar{H}^{-1}} GL_n(\mathbb{K}) \ltimes \mathbb{K}^n \xrightarrow{F} \text{Aff}(A_{\mathbb{K}}^n)$$

otteniamo anche un isomorfismo di gruppi

$$\text{Aff}_n(\mathbb{K}) \cong \text{Aff}(A_{\mathbb{K}}^n).$$

Infine, più in generale, consideriamo uno spazio affine n -dimensionale A sul campo \mathbb{K} . Fissiamo anche un riferimento affine (O, \mathcal{R}) . Sappiamo dal Teorema 6.4.3 che ad ogni applicazione affine $F : A \rightarrow A$, in particolare ad ogni affinità, è associata una *rappresentazione coordinata* (A, b) nel riferimento affine (O, \mathcal{R}) (qui consideriamo lo stesso riferimento affine nel dominio e nel codominio). Non è difficile dimostrare, e invitiamo il lettore a farlo, che

- (1) Se $F : A \rightarrow A$   un'affinit , allora la sua rappresentazione coordinata (A, b) appartiene a $GL_n(\mathbb{K}) \ltimes \mathbb{K}^n$ (cio  A   una matrice invertibile);
- (2) la funzione $R : \text{Aff}(A) \rightarrow GL_n(\mathbb{K}) \ltimes \mathbb{K}^n$ che associa a un'affinit  la sua rappresentazione coordinata   un isomorfismo di gruppi.

Componendo opportunamente con gli isomorfismi di gruppi gi  discussi, otteniamo una catena di isomorfismi di gruppi:

$$\text{Aff}(A) \cong \text{Aff}(A_{\mathbb{K}}^n) \cong GL_n(\mathbb{K}) \ltimes \mathbb{K}^n \cong \text{Aff}_n(\mathbb{K}).$$

Sia (A) uno spazio affine sul campo \mathbb{K} .

Definizione 6.5.2 — Centroaffinit . Una *centroaffinit * di (A)   un'affinit  $\Psi : A \rightarrow A$ per la quale esiste un punto $Q \in A$ tale che $\Psi(Q) = Q$. In tal caso si dice anche che Ψ *fixa* il punto Q , o che Q   un *punto fisso* di Ψ .

■ **Esempio 6.41** Sia V uno spazio vettoriale e $f : V \rightarrow V$ un automorfismo. Se interpretiamo V come uno spazio affine e f come un'affinit  (vedi Esempio 6.35), allora f   una centroaffinit  che fixa il vettore nullo $0 \in V$. ■

Esercizio 6.13 Sia A uno spazio affine e sia $Q \in A$. Dimostrare che le centroaffinit  di A che fissano Q formano un sottogruppo nel gruppo delle affinit  $\text{Aff}(A)$. ■

Proposizione 6.5.5 Sia $\Phi : A \rightarrow A$ un'affinit . Per ogni punto $Q \in A$ esistono

- (1) un'unica traslazione τ ed un'unica centroaffinit  Ψ che fixa Q , tali che $\Phi = \tau \circ \Psi$;
 - (2) un'unica traslazione τ' ed un'unica centroaffinit  Ψ' che fixa Q , tali che $\Phi = \Psi' \circ \tau'$.
- Inoltre $\Psi = \Psi'$.

Dimostrazione. Siano Φ e Q come nell'enunciato. Siano inoltre, $v := \overrightarrow{Q\Phi(Q)}$, $\tau = \tau_v$ la traslazione di vettore v e $\Psi := \tau^{-1} \circ \Phi = \tau_{-v} \circ \Phi$. Allora componendo ambo i membri a sinistra con τ troviamo

$$\tau \circ \Psi = \tau \circ \tau^{-1} \circ \Phi = \text{id}_A \circ \Phi = \Phi.$$

Ma Ψ   una centroaffinit  che fixa Q , infatti   un'affinit  perch    composta di affinit , e inoltre

$$\Psi(Q) = \tau_{-v}(\Phi(Q)) = \Phi(Q) - v = \Phi(Q) - \overrightarrow{Q\Phi(Q)} = \Phi(Q) + \overrightarrow{\Phi(Q)Q} = Q.$$

Per completare con il punto (1)   sufficiente dimostrare l'unicit . Siano dunque $\tilde{v} \in V$ e $\tilde{\Psi}$ un altro vettore e un'altra centroaffinit  che fixa Q tali che $\Phi = \tau_{\tilde{v}} \circ \tilde{\Psi}$. Osserviamo che la applicazioni lineari associate a Φ, Ψ e $\tilde{\Psi}$ coincidono. Infatti

$$f_{\Phi} = f_{\tau \circ \Psi} = f_{\tau} \circ f_{\Psi} = \text{id}_V \circ f_{\Psi} = f_{\Psi}. \quad (6.78)$$

Analogamente $f_{\Phi} = f_{\tilde{\Psi}}$. Perci  Ψ e $\tilde{\Psi}$ sono due applicazioni affini con la stessa parte lineare che associano la stessa immagine al punto Q : $\Psi(Q) = Q = \tilde{\Psi}(Q)$. Per il Teorema dell'Estensione Affine, allora, deve necessariamente essere $\Psi = \tilde{\Psi}$. Ne consegue anche che

$$\tilde{\tau} = \tilde{\tau} \circ \tilde{\Psi} \circ \tilde{\Psi}^{-1} = \Phi \circ \tilde{\Psi}^{-1} = \Phi \circ \Psi^{-1} = \tau.$$

e questo conclude la dimostrazione del punto (1).

Per il punto (2) si pu  ragionare in modo simile. Innanzitutto sia $v' = \overrightarrow{\Phi^{-1}(Q)Q}$ e sia $\tau' := \tau_{v'}$ la traslazione di vettore v' . Poniamo anche $\Psi' := \Phi \circ \tau'^{-1} = \Phi \circ \tau_{-v'}$.   facile dimostrare che $\Phi = \Psi' \circ \tau'$ (il lettore lo vede). Inoltre Ψ'   una centroaffinit  che fixa Q :

$$\Psi'(Q) = \Phi(\tau_{-v'}(Q)) = \Phi(Q - v') = \Phi(Q - \overrightarrow{\Phi^{-1}(Q)Q}) = \Phi(Q + \overrightarrow{Q\Phi^{-1}(Q)}) = \Phi(\Phi^{-1}(Q)) = Q.$$

Per l'unicità possiamo ragionare esattamente come nel caso (1) e lasciamo i dettagli al lettore per esercizio. Infine $\Psi = \Psi'$ (di nuovo per il Teorema dell'Estensione Affine) infatti sono affinità con la stessa parte lineare $f_\Psi = f_{\Psi'}$ che associano lo stesso punto $Q = \Psi(Q) = \Psi'(Q)$ al punto Q . Questo conclude la dimostrazione. ■

■ **Esempio 6.42** Data un'affinità $\Phi : A \rightarrow A$ e un punto $Q \in A$, la dimostrazione della Proposizione 6.5.5 fornisce la ricetta per costruire le traslazioni τ, τ' e le centroaffinità Ψ, Ψ' . A titolo di esempio, implementiamo questa ricetta nel caso in cui A sia lo spazio affine numerico $A_{\mathbb{K}}^n$. Sia dunque $\Phi : A_{\mathbb{K}}^n \rightarrow A_{\mathbb{K}}^n$ un'affinità. Allora $\Phi = F_{(A,b)}$ per una qualche matrice invertibile $A \in GL_n(\mathbb{K})$ e un qualche vettore numerico $b \in \mathbb{K}^n$. Il punto $Q \in A_{\mathbb{K}}^n$ è esso stesso un vettore numerico che preferiamo indicare con x_0 . In questa situazione il vettore v nella dimostrazione della Proposizione 6.5.5 è

$$v = \overrightarrow{Q\Phi(Q)} = \Phi(x_0) - x_0 = Ax_0 + b - x_0$$

Perciò, per ogni $x \in A_{\mathbb{K}}^n = \mathbb{K}^n$,

$$\Psi(x) = \tau_{-v}(\Phi(x)) = Ax + b - Ax_0 - b + x_0 = A(x - x_0) + x_0.$$

Per quanto riguarda τ' , osserviamo innanzitutto che l'affinità inversa di Φ è $\Phi^{-1} = F_{(A^{-1}, -A^{-1}b)}$ (se il lettore non ha guardato l'osservazione a pag. 193 lo dimostri ora per esercizio). Infine τ' è la traslazione di vettore

$$v' = \overrightarrow{\Phi^{-1}(Q)Q} = x_0 - \Phi^{-1}(x_0) = x_0 - A^{-1}x_0 + A^{-1}b.$$

Il lettore verifichi per esercizio che, effettivamente, $\tau_v \circ \Psi = \Psi \circ \tau_{v'} = \Phi$. ■

7. Classificazione Affine delle Quadriche

Nel Capitolo 6 abbiamo visto che i sottospazi affini di uno spazio affine sono, in coordinate, spazi di soluzioni di sistemi di equazioni lineari. Una naturale generalizzazione di questi ultimi sono gli spazi di soluzioni di sistemi di equazioni polinomiali. Tale generalizzazione è oggetto di un capitolo vasto e importante della Matematica contemporanea che prende il nome di *Geometria Algebrica*. In questo capitolo ci limiteremo a studiare le *quadriche*. Una quadrica è un sottoinsieme di uno spazio affine che, in coordinate, è lo spazio delle soluzioni di un'unica equazione polinomiale di secondo grado. Ci concentreremo sugli aspetti della teoria che non dipendono dalla scelta delle coordinate affini. Le classiche *coniche* sono quadriche in un piano affine.

7.1 Insiemi Algebrici

Sia \mathbb{K} un campo.

- Siano n, d interi non negativi. Un *monomio (monico) di grado d in n indeterminate x_1, \dots, x_n* è un'espressione formale del tipo

$$x_1^{i_1} \cdots x_n^{i_n}, \quad i_1, \dots, i_n \in \mathbb{N}_0, \quad i_1 + \cdots + i_n = d.$$

Un *polinomio in n indeterminate x_1, \dots, x_n sul campo \mathbb{K}* è una combinazione lineare formale $P(x)$ di (un numero finito di) monomi (distinti):

$$P(x) = \sum_{i_1, \dots, i_n} a_{i_1 \dots i_n} x_1^{i_1} \cdots x_n^{i_n}, \quad a_{i_1 \dots i_n} \in \mathbb{K}, \quad (7.1)$$

in cui abbiamo indicato anche con $x = (x_1, \dots, x_n)$ la n -upla delle indeterminate. Il *grado* del polinomio $P(x)$ è il massimo grado di un monomio (con coefficiente non nullo) nella (7.1). Due polinomi si possono sommare sommando i coefficienti dei monomi uguali e si possono moltiplicare ponendo

$$\begin{aligned} & \left(\sum_{i_1, \dots, i_n} a_{i_1 \dots i_n} x_1^{i_1} \cdots x_n^{i_n} \right) \cdot \left(\sum_{j_1, \dots, j_n} b_{j_1 \dots j_n} x_1^{j_1} \cdots x_n^{j_n} \right) \\ &= \sum_{i_1, \dots, i_n, j_1, \dots, j_n} a_{i_1 \dots i_n} b_{j_1 \dots j_n} x_1^{i_1+j_1} \cdots x_n^{i_n+j_n}. \end{aligned}$$

Con queste due operazioni, l'insieme dei polinomi in n -indeterminate sul campo \mathbb{K} è un anello commutativo con unità che indichiamo con il simbolo $\mathbb{K}[x_1, \dots, x_n]$.

Ogni polinomio (7.1) determina una *funzione polinomiale*

$$P : A_{\mathbb{K}}^n \rightarrow \mathbb{K}, \quad z := (z_1, \dots, z_n)^T \mapsto P(z) = \sum_{i_1, \dots, i_n} a_{i_1 \dots i_n} z_1^{i_1} \cdots z_n^{i_n},$$

sullo spazio affine numerico n -dimensionale. Uno *zero* (o una *radice*) del polinomio $P(x)$ è un punto $z \in A_{\mathbb{K}}^n$ tale che $P(z) = 0$.

Siano $P_1(x), \dots, P_p(x) \in \mathbb{K}[x_1, \dots, x_n]$ polinomi in n indeterminate $x = (x_1, \dots, x_n)$ sul campo \mathbb{K} . Il *luogo degli zeri simultanei* di $P_1(x), \dots, P_p(x)$ è il sottoinsieme $Z(P_1, \dots, P_p) \subseteq A_{\mathbb{K}}^n$ dello spazio affine numerico n -dimensionale formato da tutte le soluzioni del sistema (di equazioni polinomiali)

$$S : \begin{cases} P_1(x) = 0 \\ \vdots \\ P_p(x) = 0 \end{cases}.$$

In altre parole

$$Z(P_1, \dots, P_p) := \{z \in A_{\mathbb{K}}^n : P_i(z) = 0 \text{ per ogni } i = 1, \dots, p\}.$$

Consideriamo ora uno spazio affine A di dimensione n sul campo \mathbb{K} .

Proposizione 7.1.1 Sia $X \subseteq A$. Le seguenti condizioni sono equivalenti:

- (1) per ogni riferimento affine (O, \mathcal{R}) di A , esistono polinomi $Q_1(x), \dots, Q_q(x) \in \mathbb{K}[x_1, \dots, x_n]$, tali che $c_{(O, \mathcal{R})}(X) = Z(Q_1, \dots, Q_q)$.
- (2) esiste un riferimento affine (O, \mathcal{R}) di A ed esistono polinomi $P_1(x), \dots, P_p(x) \in \mathbb{K}[x_1, \dots, x_n]$, tali che $c_{(O, \mathcal{R})}(X) = Z(P_1, \dots, P_p)$.

Dimostrazione. È ovvio che (1) \Rightarrow (2). Sia ora (O, \mathcal{R}) un riferimento affine di A e siano

$$P_1(x), \dots, P_p(x)$$

polinomi come nella (2). Consideriamo un altro riferimento affine (O', \mathcal{R}') . Allora

$$c_{(O', \mathcal{R}')} (X) = (c_{(O', \mathcal{R}')} \circ c_{(O, \mathcal{R})}^{-1}) (c_{(O, \mathcal{R})} (X)) = (c_{(O', \mathcal{R}')} \circ c_{(O, \mathcal{R})}^{-1}) (Z(P_1, \dots, P_p)).$$

La biezione $c_{(O', \mathcal{R}')} \circ c_{(O, \mathcal{R})}^{-1} : A_{\mathbb{K}}^n \rightarrow A_{\mathbb{K}}^n$ è un'affinità. Indichiamo con F l'affinità inversa. Allora esistono una matrice invertibile $A \in \text{GL}_n(\mathbb{K})$ e un vettore numerico $b \in \mathbb{K}^n$, tali che $F = F_{(A, b)}$. Dobbiamo dimostrare che, dati polinomi $P_1(x), \dots, P_p(x)$, e data un'affinità $F = F_{(A, b)} : A_{\mathbb{K}}^n \rightarrow A_{\mathbb{K}}^n$ esistono polinomi $Q_1(x), \dots, Q_q(x)$ tali che

$$F^{-1}(Z(P_1, \dots, P_p)) = Z(Q_1, \dots, Q_q).$$

A questo scopo osserviamo che

$$\begin{aligned} F^{-1}(Z(P_1, \dots, P_p)) &= \{y \in A_{\mathbb{K}}^n : F(y) \in Z(P_1, \dots, P_p)\} \\ &= \{y \in A_{\mathbb{K}}^n : P_i(F(y)) = 0 \text{ per ogni } i = 1, \dots, p\}. \end{aligned} \tag{7.2}$$

Ora è facile verificare che, per ogni polinomio $P(x) \in \mathbb{K}[x_1, \dots, x_n]$, e per ogni applicazione affine $F = F_{(A, b)} : A_{\mathbb{K}}^n \rightarrow A_{\mathbb{K}}^n$, l'espressione formale $P(F(x)) = P(Ax + b)$, ottenuta da $P(x)$ mediante la *sostituzione lineare* $x \rightarrow Ax + b$ (può essere riorganizzata come combinazione lineare di monomi e

perciò) è di nuovo un polinomio. Posto allora $Q_i(x) = P_i(F(x))$, $i = 1, \dots, p$, deduciamo dalla (7.2) che

$$c_{(O', \mathcal{R}')} (X) = F^{-1}(Z(P_1, \dots, P_p)) = Z(Q_1, \dots, Q_p),$$

come desiderato. ■

Osserviamo, per futuri utilizzi, che se $P(x) \in \mathbb{K}[x_1, \dots, x_n]$ è un polinomio di grado d e $F = F_{(A,b)} : A_{\mathbb{K}}^n \rightarrow A_{\mathbb{K}}^n$ è un'affinità, allora anche $P(F(x))$ è un polinomio di grado d .

Definizione 7.1.1 — Insieme Algebrico. Un sottoinsieme $X \subseteq A$ è un *insieme algebrico* se soddisfa una delle, e quindi entrambe le, condizioni nella Proposizione 7.1.1. L'insieme algebrico X è una *quadrica* se, in coordinate affini rispetto ad un riferimento affine, e quindi rispetto ad ogni riferimento affine, è il luogo degli zeri di un singolo polinomio di secondo grado. Una *conica* è una quadrica in uno spazio affine di dimensione 2.

7.2 Equivalenza di Polinomi di Secondo Grado

Concentriamoci adesso sul caso di un insieme algebrico $X \subseteq A$ che, in coordinate, in un riferimento affine (O, \mathcal{R}) , sia il luogo $Z(P)$ degli zeri di un unico polinomio $P(x) \in \mathbb{K}[x_1, \dots, x_n]$. Vorremmo scegliere il riferimento (O, \mathcal{R}) in modo tale che il polinomio $P(x)$ sia il più semplice possibile. La dimostrazione della Proposizione 7.1.1 mostra che, se (O', \mathcal{R}') è un altro riferimento affine allora, in coordinate, nel riferimento (O', \mathcal{R}') , l'insieme algebrico X è il luogo $Z(Q)$ degli zeri di un altro polinomio $Q(x) \in \mathbb{K}[x_1, \dots, x_n]$ collegato a P da una opportuna affinità $F : A_{\mathbb{K}}^n \rightarrow A_{\mathbb{K}}^n$ mediante la formula $Q(x) = P(F(x))$. Inoltre, se $k \in \mathbb{K}$ è uno scalare non nullo è chiaro che $Z(Q) = Z(kQ)$ (il lettore lo vede?). Queste semplici considerazioni suggeriscono di introdurre la seguente relazione nell'insieme $\mathbb{K}[x_1, \dots, x_n]$ dei polinomi.

Definizione 7.2.1 Due polinomi $P, Q \in \mathbb{K}[x_1, \dots, x_n]$ si dicono *equivalenti* se esistono

- (1) uno scalare non nullo $k \in \mathbb{K}$, e
 - (2) un'affinità $F : A_{\mathbb{K}}^n \rightarrow A_{\mathbb{K}}^n$,
- tali che $Q(x) = kP(F(x))$.

Ovviamente polinomi equivalenti hanno lo stesso grado.

Esercizio 7.1 Dimostrare che l'equivalenza di polinomi è una relazione di equivalenza. ■

Se dunque, in un riferimento affine (O, \mathcal{R}) , un insieme algebrico X è il luogo degli zeri $Z(P)$ di un polinomio $P(x) \in \mathbb{K}[x_1, \dots, x_n]$, allora, in ogni altro riferimento affine (O', \mathcal{R}') , X è il luogo degli zeri $Z(Q)$ di un polinomio $Q(x)$ equivalente a $P(x)$. In vista di questa osservazione procediamo come segue: per trovare la più semplice rappresentazione di X , troviamo innanzitutto il più semplice polinomio equivalente a $P(x)$. Nel seguito ci limiteremo al caso delle quadriche in cui il polinomio $P(x)$ può essere scelto di secondo grado. In questo contesto i risultati più semplici si ottengono nel caso in cui la caratteristica del campo \mathbb{K} sia diversa da 2, cosa che assumeremo sempre di qui in avanti, spesso senza ulteriori commenti.

Cominciamo con l'osservare che un polinomio di secondo grado $P(x) \in \mathbb{K}[x_1, \dots, x_n]$ può essere scritto nella forma

$$P(x) = a_0 + \sum_{i=1}^n b_i x_i + \sum_{i,j=1}^n b_{ij} x_i x_j \quad (7.3)$$

dove $a_0, b_i, b_{ij} \in \mathbb{K}$ sono scalari (il lettore lo vede?). Consideriamo la riga $b := (b_1, \dots, b_n) \in M_{1,n}(\mathbb{K})$ e la matrice $B := (b_{ij}) \in M_n(\mathbb{K})$. Se interpretiamo la n -upla x delle indeterminate x_1, \dots, x_n come

una colonna $x = (x_1, \dots, x_n)^T$, la (7.3) può essere presentata nella forma:

$$P(x) = a_0 + bx + x^T Bx,$$

dove abbiamo usato il prodotto righe per colonne. Se la caratteristica del campo \mathbb{K} è diversa da 2, come stiamo assumendo, allora esistono una riga $a = (a_1, \dots, a_n) \in M_{1,n}(\mathbb{K})$ e una matrice simmetrica e non nulla $A \in M_n(\mathbb{K})$ tali che si abbia anche

$$P(x) = a_0 + 2ax + x^T Ax. \quad (7.4)$$

Infatti è sufficiente porre

$$a := \frac{1}{2}b, \quad e \quad A := \frac{1}{2}(B + B^T)$$

e si ha

$$A^T = \frac{1}{2}(B + B^T)^T = \frac{1}{2}(B^T + B^{TT}) = \frac{1}{2}(B^T + B) = A,$$

cioè A è una matrice simmetrica. Inoltre

$$x^T Ax = \frac{1}{2}x^T (B + B^T)x = \frac{1}{2}(x^T Bx + x^T B^T x), \quad (7.5)$$

ma $x^T B^T x$ è uno scalare (equivalentemente, una matrice di ordine 1×1) da cui

$$x^T B^T x = (x^T B^T x)^T = x^T B^{TT} x^{TT} = x^T Bx,$$

che, sostituita nella (7.5), dà

$$x^T Ax = \frac{1}{2}(x^T Bx + x^T B^T x) = \frac{1}{2}2x^T Bx = x^T Bx.$$

Riassumendo

$$P(x) = a_0 + bx + x^T Bx = a_0 + 2ax + x^T Ax.$$

Infine, come desiderato, la matrice A non può essere nulla, altrimenti $P(x)$ non sarebbe un polinomio di grado 2. D'ora in avanti, dato un polinomio di secondo grado $P(x) \in \mathbb{K}[x_1, \dots, x_n]$, lo presenteremo sempre nella forma (7.4) (assumendo che la caratteristica del campo \mathbb{K} sia diversa da 2).

Accanto alla matrice A considereremo anche la matrice

$$M := \begin{pmatrix} A & a^T \\ a & a_0 \end{pmatrix} \in M_{n+1}(\mathbb{K}).$$

Chiamiamo A ed M le *matrici associate al polinomio di secondo grado* $P(x)$. Indicheremo anche con $\mathfrak{a}(P)$ (o semplicemente \mathfrak{a} se questo non dà luogo a confusione) il rango della matrice A e con $\mathfrak{m}(P)$ (o semplicemente \mathfrak{m}) il rango della matrice M . Si osservi che, essendo A una sottomatrice non nulla di M , necessariamente $1 \leq \mathfrak{a}(P) \leq \mathfrak{m}(P)$, in particolare $\mathfrak{a}(P), \mathfrak{m}(P)$ sono numeri naturali.

Proposizione 7.2.1 Lo scalare a_0 , la riga a e la matrice A (e quindi la matrice M) sono univocamente determinati dal polinomio $P(x)$. In altre parole, se a'_0, a', A' sono un altro scalare, un'altra riga, e un'altra matrice simmetrica tali che $P(x)$ è anche dato da $a'_0 + 2a'x + x^T A'x$, allora necessariamente $a'_0 = a_0$, $a' = a$, e $A' = A$.

Dimostrazione. La dimostrazione è lasciata come Esercizio 7.2. ■

Esercizio 7.2 Dimostrare la Proposizione 7.2.1 (*Suggerimento: si ricordi che una forma bilineare simmetrica su un campo di caratteristica diversa da 2 è univocamente determinata dalla sua forma quadratica associata tramite la Formula di Polarizzazione*). ■

Proposizione 7.2.2 I numeri naturali $\alpha(P), m(P)$ sono *invarianti per equivalenze di polinomi di secondo grado*, cioè se $P'(x)$ è un polinomio (di secondo grado) equivalente a $P(x)$, allora $\alpha(P') = \alpha(P)$ e $m(P') = m(P)$.

Dimostrazione. Siano $P(x), P'(x) \in \mathbb{K}[x_1, \dots, x_n]$ polinomi di secondo grado equivalenti. Questo vuol dire che esistono uno scalare $k \in \mathbb{K} \setminus \{0\}$ e un'affinità $F: \mathbb{A}_{\mathbb{K}}^n \rightarrow \mathbb{A}_{\mathbb{K}}^n$ tali che $P'(x) = kP(F(x))$. Come al solito, poiché F è un'affinità tra spazi affini numerici, si ha anche $F = F_{(C,c)}$ per qualche matrice invertibile $C \in \text{GL}_n(\mathbb{K})$ e qualche vettore numerico $c \in \mathbb{K}^n$. Dette A, M le matrici associate a $P(x)$ e A', M' le matrici associate a $P'(x)$ vogliamo innanzitutto esprimere A', M' in termini di A, M, C, c . Se

$$M = \begin{pmatrix} A & a^T \\ a & a_0 \end{pmatrix} \in M_{n+1}(\mathbb{K}) \quad \text{e} \quad M' = \begin{pmatrix} A' & a'^T \\ a' & a'_0 \end{pmatrix} \in M_{n+1}(\mathbb{K}),$$

allora, da un lato,

$$P'(x) = a'_0 + 2a'x + x^T A' x.$$

Dall'altro

$$P'(x) = kP(F(x)) = kP(Cx + c) = k(a_0 + 2a(Cx + c) + (Cx + c)^T A(Cx + c))$$

Raccogliendo con un po' di pazienza i termini omogenei, utilizzando che $A = A^T$, e osservando che per ogni riga $d \in M_{1,n}(\mathbb{K})$ si ha $dx = x^T d^T$, otteniamo

$$P'(x) = k(a_0 + 2ac + c^T Ac) + 2k(aC + c^T AC)x + x^T (kC^T AC)x. \quad (7.6)$$

Ma $kC^T AC$ è di nuovo una matrice simmetrica. Segue ora dalla Proposizione 7.2.1 che

$$a'_0 = k(a_0 + 2ac + c^T Ac),$$

$$a' = k(aC + c^T AC),$$

$$A' = kC^T AC.$$

Inoltre, un conto diretto, che lasciamo al lettore, mostra che

$$M' = kN^T M N,$$

dove $N \in M_{n+1}(\mathbb{K})$ è la matrice data da

$$N = \begin{pmatrix} C & c \\ 0 & 1 \end{pmatrix}. \quad (7.7)$$

Si ricordi che la matrice C è invertibile, perciò le matrici A e $C^T AC$ sono congruenti. Poiché il prodotto per una matrice invertibile non cambia il rango, matrici congruenti hanno lo stesso

rango. D'altronde, poichè $k \neq 0$, si ha anche $\text{rk}(C^T AC) = \text{rk}(kC^T AC)$. Mettendo assieme le ultime osservazioni:

$$a(P) = \text{rk} A = \text{rk}(C^T AC) = \text{rk}(kC^T AC) = \text{rk} A' = a(P').$$

Infine, anche N è una matrice invertibile (il suo determinante è dato dal prodotto dei determinanti dei blocchi diagonali ed è perciò uguale a $\det C \neq 0$) da cui, per gli stessi motivi di sopra, $m(P) = m(P')$, come desiderato. Questo conclude la dimostrazione. ■

Sia $P(x) \in \mathbb{K}[x_1, \dots, x_n]$ un polinomio di secondo grado. Abbiamo già osservato che $1 \leq a(P) \leq m(P)$. In effetti, possiamo essere più precisi sulla relazione tra i numeri naturali $a(P), m(P)$ in base al seguente

Lemma 7.2.3 Sia $P(x) \in \mathbb{K}[x_1, \dots, x_n]$ un polinomio di secondo grado. Allora $0 < a(P) \leq n$. Inoltre, $a(P) \leq m(P) \leq a(P) + 2$.

Dimostrazione. La prima parte dell'enunciato è ovvia. Per la seconda parte si ricordi che, se $P(x) = a_0 + 2ax + x^T Ax$, allora $a(P)$ è il rango della matrice A , mentre $m(P)$ è il rango della matrice

$$M = \begin{pmatrix} A & a^T \\ a & a_0 \end{pmatrix}.$$

In particolare la matrice A ha esattamente $a(P)$ colonne indipendenti. Ne consegue che la matrice

$$\begin{pmatrix} A & a^T \end{pmatrix}$$

ha al massimo $a(P) + 1$ colonne, e quindi righe, indipendenti. Concludiamo che la matrice M ha al massimo $a(P) + 2$ righe indipendenti, come desiderato. ■

Il Lemma 7.2.3 dice in pratica che il numero naturale $m(P)$ può assumere solo i valori $a(P)$, $a(P) + 1$, oppure $a(P) + 2$. Ovviamente l'ultima circostanza è possibile solo se $a(P) < n$. In effetti tutti e tre i casi possono essere realizzati scegliendo opportunamente il polinomio $P(x)$. A titolo di esempio, sia a un intero compreso tra 1 ed n . Nel caso

$$P(x) = x_1^2 + \dots + x_a^2$$

si ha

$$M = \begin{pmatrix} I_a & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

da cui $m(P) = a(P) = a$. Nel caso invece

$$P(x) = 1 + x_1^2 + \dots + x_a^2$$

si ha

$$M = \begin{pmatrix} I_a & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

da cui $a(P) = a$ e $m(P) = a(P) + 1 = a + 1$. Infine, nel caso $a < n$, e

$$P(x) = 2x_{a+1} + x_1^2 + \dots + x_a^2$$

si ha

$$M = \begin{pmatrix} I_a & 0 & 0 \\ 0 & 0 & E_1 \\ 0 & E_1^T & 0 \end{pmatrix}$$

da cui $a(P) = a$ e $m(P) = a(P) + 2 = a + 2$.

Siamo ora pronti ad enunciare il risultato principale di questa sezione.

Teorema 7.2.4 — Classificazione dei Polinomi di Secondo Grado. Sia \mathbb{K} un campo di caratteristica diversa da 2 e sia $P(x) \in \mathbb{K}[x_1, \dots, x_n]$ un polinomio di secondo grado in n indeterminate $x = (x_1, \dots, x_n)^T$ sul campo \mathbb{K} , e indichiamo $a := a(P)$. Solo uno dei seguenti casi si può presentare:

- (1) $m(P) = a(P) = a$. In questo caso esistono scalari non nulli q_2, \dots, q_a tali che $P(x)$ è equivalente al polinomio

$$x_1^2 + q_2 x_2^2 + \dots + q_a x_a^2. \quad (7.8)$$

- (2) $m(P) = a(P) + 1 = a + 1$. In questo caso esistono scalari non nulli q_1, \dots, q_a tali che $P(x)$ è equivalente al polinomio

$$1 + q_1 x_1^2 + q_2 x_2^2 + \dots + q_a x_a^2. \quad (7.9)$$

- (3) $a(P) = a < n$ e $m(P) = a(P) + 2 = a + 2$. In questo caso esistono scalari non nulli q_1, \dots, q_a tali che $P(x)$ è equivalente al polinomio

$$2x_{a+1} + q_1 x_1^2 + q_2 x_2^2 + \dots + q_a x_a^2. \quad (7.10)$$

Dimostrazione. Sia $P(x) \in \mathbb{K}[x_1, \dots, x_n]$ un polinomio di secondo grado. Siano A ed

$$M = \begin{pmatrix} A & a^T \\ a & a_0 \end{pmatrix}$$

le matrici associate a $P(x)$. Cioè

$$P(x) = a_0 + 2ax + x^T Ax.$$

Consideriamo una traslazione $\tau : A_{\mathbb{K}}^n \rightarrow A_{\mathbb{K}}^n$ di vettore $c \in \mathbb{K}^n$. Posto $P'(x) := P(\tau(x))$, la Formula (7.6) mostra che

$$P'(x) := a_0 + 2ac + c^T Ac + 2(a + c^T A)x + x^T Ax.$$

[Caso Centrale]. Esiste un vettore $c \in \mathbb{K}^n$ tale che $a + c^T A = 0$. In questo caso, posto $\tau = \tau_c$ e

$P'(x) = P(\tau(x))$, otteniamo

$$P'(x) = a'_0 + x^T Ax$$

dove

$$a'_0 = a_0 + 2ac + c^T Ac = a_0 + 2ac - ac = a_0 + ac.$$

Distinguiamo due sottocasi:

[Caso Centrale–1]. $a'_0 = 0$. In questo (sotto)caso $P(x)$ è equivalente al polinomio

$$P'(x) = x^T A x,$$

da cui $m(P) = m(P') = a(P') = a(P) = a$. Inoltre, la matrice A è congruente ad una matrice diagonale, cioè esiste una matrice invertibile $C \in GL_n(\mathbb{K})$ tale che $C^T A C$ è una matrice diagonale (dello stesso rango a di A). Riordinando opportunamente le colonne di C si vede addirittura che esistono scalari non nulli q'_1, \dots, q'_a tali che

$$C^T A C = \begin{pmatrix} q'_1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & q'_a & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Se ora consideriamo l'affinità $F = F_{(C,0)} : A_{\mathbb{K}}^n \rightarrow A_{\mathbb{K}}^n$, di nuovo dalla Formula (7.6) otteniamo

$$\frac{1}{q'_1} P'(F(x)) = \frac{1}{q'_1} x^T C^T A C x = \frac{1}{q'_1} (q'_1 x_1^2 + q'_2 x_2^2 + \cdots + q'_a x_a^2) = x_1^2 + \frac{q'_2}{q'_1} x_2^2 + \cdots + \frac{q'_a}{q'_1} x_a^2.$$

In altre parole, posto $q_i = q'_i/q'_1$, $i = 2, \dots, a$, il polinomio $P'(x)$, e quindi, dall'Esercizio 7.1, anche il polinomio $P(x)$, è equivalente al polinomio (7.8).

[Caso Centrale–2] $a'_0 \neq 0$. In questo (sotto)caso la matrice M' associata al polinomio $P'(x)$ è

$$M' = \begin{pmatrix} A & 0 \\ 0 & a'_0 \end{pmatrix}$$

in cui l'ultima entrata di diagonale è diversa da 0. Perciò $m(P) = m(P') = a(P') + 1 = a(P) + 1 = a + 1$. Di nuovo esiste $C \in GL_n(\mathbb{K})$ come nel sottocaso precedente e, considerata di nuovo l'affinità $F := F_{(C,0)}$, otteniamo

$$\begin{aligned} \frac{1}{a'_0} P'(F(x)) &= \frac{1}{a'_0} (a'_0 + x^T C^T A C x) = \frac{1}{a'_0} (a'_0 + q'_1 x_1^2 + q'_2 x_2^2 + \cdots + q'_a x_a^2) \\ &= 1 + \frac{q'_1}{a'_0} x_1^2 + \frac{q'_2}{a'_0} x_2^2 + \cdots + \frac{q'_a}{a'_0} x_a^2, \end{aligned}$$

Posto $q_i = q'_i/a'_0$, $i = 1, \dots, a$, il polinomio $P(x)$ è equivalente al polinomio (7.9).

[Caso Parabolico]. Non esiste un vettore $c \in \mathbb{K}^n$ tale che $a + c^T A = 0$. Poiché A è una matrice simmetrica, questo equivale a dire che il sistema

$$S_p : Ax = -a^T$$

non è compatibile. In tal caso, non effettuiamo la sostituzione lineare $x \rightarrow \tau(x)$ discussa all'inizio della dimostrazione. Piuttosto, procediamo così: come nei casi precedenti, esiste una matrice invertibile $C \in GL_n(\mathbb{K})$ tale che $C^T A C$ è una matrice diagonale. Dunque, posto di nuovo $F = F_{(C,0)}$, abbiamo, per opportuni scalari non nulli $q_1, \dots, q_a \in \mathbb{K}$,

$$P''(x) := P(F(x)) = a''_0 + 2a''_1 x + q_1 x_1^2 + \cdots + q_a x_a^2,$$

dove abbiamo posto $a'' = aC$. I polinomi $P(x)$ e $P''(x)$ sono equivalenti, perciò, in virtù del Lemma 7.2.5 più sotto, neanche il sistema $S_{P''}$ è compatibile. Ma $S_{P''}$ è dato da

$$S_{P''} : \begin{cases} q_1 x_1 & = -a''_1 \\ \vdots & \\ q_a x_a & = -a''_a \\ 0 & = -a''_{a+1} \\ \vdots & \\ 0 & = -a''_n \end{cases} .$$

Questo implica che esiste $i = a + 1, \dots, n$ tale che $a''_i \neq 0$, sicché l'applicazione affine

$$G : A_{\mathbb{K}}^n \rightarrow A_{\mathbb{K}}^n, \quad \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ \vdots \\ a''_1 x_1 + \dots + a''_n x_n + a''_0/2 \\ \vdots \\ x_n \end{pmatrix} \leftarrow \text{posto } i$$

è un'affinità (il lettore lo vede?). Consideriamo ora il polinomio

$$\bar{P}(x) = 2x_i + q_1 x_1^2 + \dots + q_a x_a^2,$$

e calcoliamo

$$\bar{P}(G(x)) = 2 \left(a''_1 x_1 + \dots + a''_n x_n + \frac{a''_0}{2} \right) + q_1 x_1^2 + \dots + q_a x_a^2 = P''(x).$$

Ne deduciamo che $P(x)$ è equivalente a $\bar{P}(x)$. Applicando l'affinità che scambia x_i con x_{a+1} vediamo che $P(x)$ è anche equivalente al polinomio (7.10) come desiderato. Questo conclude la dimostrazione. ■

Lemma 7.2.5 La compatibilità del sistema S_P è *invariante per equivalenze di polinomi di secondo grado*, cioè se $P(x), P'(x) \in \mathbb{K}[x_1, \dots, x_n]$ sono polinomi di secondo grado equivalenti, allora il sistema S_P è compatibile sse lo è il sistema $S_{P'}$.

Dimostrazione. Sia $P(x) = a_0 + ax + x^T Ax$. Si osservi che, per il Teorema di Rouché-Capelli, il sistema S_P è compatibile sse

$$\text{rk } A = \text{rk} \begin{pmatrix} A & a^T \end{pmatrix}.$$

Siano ora $k \in \mathbb{K} \setminus \{0\}$ uno scalare non nullo e $F = F_{(C,c)} : A_{\mathbb{K}}^n \rightarrow A_{\mathbb{K}}^n$ un'affinità tali che $P'(x) = kP(F(x))$. Se $P'(x) = a'_0 + a'x + x^T A'x$, allora il sistema $S_{P'}$ è $A'x = -a'^T$ che è compatibile sse

$$\text{rk } A' = \text{rk} \begin{pmatrix} A' & a'^T \end{pmatrix}.$$

Ma il conto nella dimostrazione della Proposizione 7.2.2 mostra che

$$\begin{pmatrix} A' & a'^T \end{pmatrix} = kC^T \begin{pmatrix} A & a^T \end{pmatrix} N,$$

dove la matrice $N \in \text{GL}_{n+1}(\mathbb{K})$ è data dalla (7.7). Ne consegue che

$$\text{rk} \begin{pmatrix} A' & a'^T \end{pmatrix} = \text{rk} \begin{pmatrix} A & a^T \end{pmatrix}$$

(il prodotto per una matrice invertibile non cambia il rango). Poiché abbiamo anche $\text{rk} A = \text{rk} A'$, concludiamo che

$$\text{rk} A = \text{rk} \begin{pmatrix} A & a^T \end{pmatrix} \quad \text{sse} \quad \text{rk} A' = \text{rk} \begin{pmatrix} A' & a'^T \end{pmatrix}$$

come desiderato. ■

- La terminologia “caso centrale” che abbiamo adottato nella dimostrazione del Teorema 7.2.4 è motivata dalle seguenti osservazioni. Sia $P(x) = a_0 + ax + x^T Ax$ un polinomio di secondo grado in n indeterminate x_1, \dots, x_n . Supponiamo di essere nel *caso centrale*, cioè esiste $c \in \mathbb{K}^n$ tale che $a + c^T A = 0$. Allora il punto $c \in A_{\mathbb{K}}^n$ è un *centro di simmetria* per la quadrica $Z(P)$, cioè l'affinità

$$F : A_{\mathbb{K}}^n \rightarrow A_{\mathbb{K}}^n, \quad x \mapsto F(x) := 2c - x,$$

che *riflette rispetto a c*, trasforma $Z(P)$ in sé stessa: $F(Z(P)) = Z(P)$. Per dimostrarlo dobbiamo verificare che

$$P(x) = 0 \quad \Leftrightarrow \quad P(F(x)) = 0.$$

È così, infatti un conto diretto (che invitiamo il lettore a fare) mostra che

$$P(F(x)) = P(2c - x) = 4(a + c^T A)(c - x) + P(x) = P(x).$$

Per esempio, se $c = 0$, allora la quadrica $Z(P)$ è *simmetrica rispetto all'origine*, cioè l'affinità $x \mapsto -x$ trasforma $Z(P)$ in sé stessa.

Se $\mathbb{K} = \mathbb{R}$, in virtù della Legge di Inerzia di Sylvester, si possono ottenere valori specifici per gli scalari q_1, \dots, q_a nell'enunciato del Teorema 7.2.4. Più precisamente vale il seguente

Corollario 7.2.6 Sia $P(x) \in \mathbb{R}[x_1, \dots, x_n]$ un polinomio reale di secondo grado in n indeterminate $x = (x_1, \dots, x_n)^T$, e indichiamo $a := a(P)$. Solo uno dei seguenti casi si può presentare:

- (1) $m(P) = a(P) = a$ e in questo caso esiste un intero p , con $1 \leq p \leq a$, tale che $P(x)$ è equivalente al polinomio

$$Q_{p,a}(x) := x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_a^2; \quad (7.11)$$

- (2) $m(P) = a(P) + 1 = a + 1$ e in questo caso esiste un intero p , con $0 \leq p \leq a$, tale che $P(x)$ è equivalente al polinomio

$$Q_{p,a}(x) := 1 + x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_a^2; \quad (7.12)$$

- (3) $a(P) = a < n$ e $m(P) = a(P) + 2 = a + 2$ e in questo caso esiste un intero p , con $0 \leq p \leq a$, tale che $P(x)$ è equivalente al polinomio

$$Q_{p,a}(x) := 2x_{a+1} + x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_a^2. \quad (7.13)$$

Dimostrazione. La dimostrazione è lasciata come Esercizio 7.3. ■

Osserviamo che l'intero p nell'enunciato del Corollario 7.2.6 è determinato dal polinomio $P(x)$ come segue: la *segnatura della matrice A* è $(p, a - p, n - a)$ oppure $(a - p, p, n - a)$. Queste ultime due circostanze sono equivalenti nei casi (1) e (3) (perciò non occorre distinguerle). Infatti, in tali casi, i polinomi $Q_{p,a}(x)$ e $Q_{a-p,a}(x)$ sono equivalenti (il lettore lo vede?). Tuttavia, nel caso (2), i polinomi $Q_{p,a}(x)$ e $Q_{a-p,a}(x)$ *non* sono equivalenti ed è importante determinare con maggiore precisione l'intero p . In effetti, in tal caso, p può essere determinato osservando che la segnatura della matrice M è $(p + 1, a - p, n - a)$ oppure $(a - p, p + 1, n - a)$.

Esercizio 7.3 Dimostrare il Corollario 7.2.6. ■

Definizione 7.2.2 — Forma Normale dei Polinomi Reali di Secondo Grado. Ogni polinomio reale di secondo grado del tipo (7.11), (7.12) o (7.13) è detto in *forma normale*. Se un polinomio reale di secondo grado $P(x)$ è equivalente ad un polinomio $Q(x)$ in forma normale, allora $Q(x)$ è detta la *forma normale* di $P(x)$.

7.3 Forma Normale delle Quadriche

La più importante conseguenza geometrica del Teorema 7.2.4 è che una quadrica ha una descrizione particolarmente semplice in opportune coordinate affini.

Corollario 7.3.1 — Forma Normale delle Quadriche. Sia A uno spazio affine di dimensione n su un campo \mathbb{K} di caratteristica diversa da 2 e sia $X \subseteq A$ una quadrica. Allora esiste un riferimento affine (O, \mathcal{R}) di A tale che $c_{(O, \mathcal{R})}(X)$ è il luogo degli zeri di un polinomio del tipo

$$x_1^2 + q_2 x_2^2 + \cdots + q_a x_a^2,$$

oppure

$$1 + q_1 x_1^2 + q_2 x_2^2 + \cdots + q_a x_a^2,$$

oppure

$$2x_{a+1} + q_1 x_1^2 + q_2 x_2^2 + \cdots + q_a x_a^2,$$

per qualche $a = 0, \dots, n$ e qualche a -upla di scalari non nulli $q_1, \dots, q_a \in \mathbb{K}$.

Dimostrazione. Sia (O', \mathcal{R}') un qualunque riferimento affine di A . Allora esiste un polinomio di secondo grado $P(x) \in \mathbb{K}[x_1, \dots, x_n]$ tale che $c_{(O', \mathcal{R}')} (X) = Z(P)$. Per il Teorema 7.2.4 di Classificazione dei Polinomi di Secondo Grado, $P(x)$ è equivalente ad un polinomio di secondo grado $Q(x)$ del tipo (7.8), (7.9) o (7.10), cioè esistono un'affinità $F: A_{\mathbb{K}}^n \rightarrow A_{\mathbb{K}}^n$ e uno scalare non nullo $k \in \mathbb{K}$, tali che $Q(x) = kP(F(x))$. Segue ora dalla (7.2) che

$$F^{-1}(Z(P)) = Z(P(F)) = Z(kP(F)) = Z(Q)$$

e perciò

$$(F^{-1} \circ c_{(O', \mathcal{R}')})(X) = Z(Q).$$

Resta solo da dimostrare che esiste un riferimento affine (O, \mathcal{R}) tale che $F^{-1} \circ c_{(O', \mathcal{R}')} = c_{(O, \mathcal{R})}$. Ma questo segue dalla discussione nell'Esempio 6.38. ■

Il Corollario 7.3.1 si specializza come segue al caso reale in cui $\mathbb{K} = \mathbb{R}$.

Corollario 7.3.2 — Forma Normale delle Quadriche Reali. Sia A uno spazio affine reale di dimensione n e sia $X \subseteq A$ una quadrica. Allora esiste un riferimento affine (O, \mathcal{R}) di A tale che $c_{(O, \mathcal{R})}(X)$ è il luogo degli zeri di un polinomio di secondo grado in forma normale (cfr. Definizione 7.2.2).

Esercizio 7.4 Sia A uno spazio affine reale e siano $X, Y \subseteq A$ due quadriche con la stessa forma normale. Dimostrare che esiste un'affinità $F : A \rightarrow A$ che *trasforma* X in Y , cioè tale che $F(X) = Y$. ■

Concludiamo questa sezione e questo capitolo discutendo nel dettaglio il caso delle *coniche* reali. Sia dunque A un piano affine reale e sia $X \subseteq A$ una conica. In coordinate affini (x, y) , X è il luogo degli zeri di un polinomio reale di secondo grado $P(x, y)$ nelle indeterminate x, y . Dette $A \in M_2(\mathbb{R})$ ed $M \in M_3(\mathbb{R})$ le matrici associate a $P(x, y)$, il rango a di A e il rango m di M possono assumere solo i seguenti valori:

- ✓ $a = 1$, nel qual caso m può essere 1, 2, oppure 3,
- ✓ $a = 2$, nel qual caso m può essere solo 2 oppure 3.

Riassumendo, in opportune coordinate affini, X è il luogo degli zeri di un polinomio $P(x, y)$ in una delle seguenti forme normali:

- (1) $a = 1, m = 1, p = 1$ (il caso $p = 0$ è equivalente):

$$P(x, y) = x^2$$

in questo caso ($Z(P)$ è la retta di equazione $x = 0$ e) X è una *retta*;

- (2) $a = 1, m = 2, p = 1$:

$$P(x, y) = 1 + x^2,$$

in questo caso X è vuota e si dice che è una *coppia di rette immaginarie coniugate*;

- (3) $a = 1, m = 2, p = 0$:

$$P(x, y) = 1 - y^2,$$

in questo caso ($Z(P)$ è l'unione delle rette di equazione $y = \pm 1$ in $A_{\mathbb{R}}^2$ e) X è una *coppia di rette parallele e disgiunte*;

- (4) $a = 1, m = 3, p = 1$ (il caso $p = 0$ è equivalente):

$$P(x, y) = 2y + x^2,$$

in questo caso, si dice che X è una *parabola* (il che motiva anche la terminologia “caso parabolico”);

- (5) $a = 2, m = 2, p = 2$ (il caso $p = 0$ è equivalente):

$$P(x, y) = x^2 + y^2,$$

in questo caso X è un *singolo punto*;

- (6) $a = 2, m = 2, p = 1$:

$$P(x, y) = x^2 - y^2,$$

in questo caso ($Z(P)$ è l'unione delle rette di equazione $y = \pm x$ in $A_{\mathbb{R}}^2$ e) X è una *coppia di rette che si intersecano in un punto*;

- (7) $a = 2, m = 3, p = 2$:

$$P(x, y) = 1 + x^2 + y^2,$$

in questo caso X è vuota e si dice che è un' *ellisse immaginaria*;

- (8) $a = 2, m = 3, p = 1$:

$$P(x, y) = 1 + x^2 - y^2,$$

in questo caso ($Z(P)$ è l'iperbole unitaria in $A_{\mathbb{R}}^2$ e) si dice che X è un' *iperbole*;

(9) $a = 2, m = 3, p = 0$:

$$P(x, y) = 1 - x^2 - y^2,$$

in questo caso ($Z(P)$ è la circonferenza unitaria in $A_{\mathbb{R}}^2$ e) si dice che X è un'ellisse.

Le coniche dei casi $a = 1$ vengono dette *degeneri*, perchè corrispondono al caso in cui la matrice A è, appunto, degenera. Le rimanenti coniche vengono dette, *non degeneri*.

○ Le coniche dei casi centrali $m = a, a + 1$ vengono dette *coniche a centro* e, in effetti, posseggono (almeno un) *centro di simmetria* nell'origine del riferimento affine (il lettore lo vede?).

■ **Esempio 7.1** Nel piano affine numerico reale $A_{\mathbb{R}}^2$ con coordinate standard x, y si consideri la conica X di equazione

$$X: -1 + 5x - y + x^2 - 6xy + 2y^2 = 0.$$

In altre parole $X = Z(P)$, con $P(x, y) = -1 + 5x - y + x^2 - 6xy + 2y^2$. Le matrici associate al polinomio di secondo grado $P(x, y)$ sono

$$A = \begin{pmatrix} 1 & -3 \\ -3 & 2 \end{pmatrix} \quad \text{e} \quad M = \begin{pmatrix} 1 & -3 & 5/2 \\ -3 & 2 & -1/2 \\ 5/2 & -1/2 & -1 \end{pmatrix}.$$

Studiamo innanzitutto il rango e la segnatura della forma bilineare simmetrica $\beta_A: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$. Il primo vettore E_1 del riferimento canonico di \mathbb{R}^2 è non isotropo rispetto a β_A , infatti $\beta_A(E_1, E_1) = 1$ (l'elemento di posto (1, 1) della matrice A) che è diverso da 0. Un vettore $v = (x, y)^T$ è ortogonale a E_1 sse

$$0 = \beta_A(E_1, v) = x - 3y,$$

cioè $x = 3y$. Dunque il sottospazio vettoriale $W \subseteq \mathbb{R}^2$ dei vettori ortogonali ad E_1 è il sottospazio unidimensionale dato da:

$$W = \left\{ \begin{pmatrix} 3t \\ t \end{pmatrix} : t \in \mathbb{R} \right\}.$$

Un riferimento di W è costituito dal vettore

$$E'_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

I vettori E_1, E'_1 formano un riferimento di \mathbb{R}^2 ortogonale rispetto a β_A . In questo riferimento, la matrice rappresentativa di β_A è la matrice

$$A' = \begin{pmatrix} \beta_A(E_1, E_1) & 0 \\ 0 & \beta_A(E'_1, E'_1) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -7 \end{pmatrix},$$

da cui leggiamo che $a = 2$ (X è una conica non degenera) e la segnatura di A è $(1, 1, 0)$. Dunque $p = 1$. Resta da studiare il rango di M . Può essere solo 2 o 3. Per distinguere i due casi è sufficiente calcolare

$$\det M = \det \begin{pmatrix} 1 & -3 & 5/2 \\ -3 & 2 & -1/2 \\ 5/2 & -1/2 & -1 \end{pmatrix} = \frac{7}{4} \neq 0.$$

Dunque $m = 3$ e siamo nel caso (8): X è un'iperbole. ■

Esercizio 7.5 Studiare la quadrica $Z(P) \subseteq A_{\mathbb{R}}^2$ in ciascuno dei seguenti casi:

$$P(x, y) = -3y + 2x^2 + xy + y^2,$$

$$P(x, y) = -5 + 4x + x^2 - 2xy + y^2,$$

$$P(x, y) = -x - y + 3x^2 + 5xy + 2y^2.$$

Esercizio 7.6 Studiare la quadrica $Z(P_t) \subseteq A_{\mathbb{R}}^2$ al variare del parametro $t \in \mathbb{R}$ in ciascuno dei seguenti casi:

$$P_t(x, y) = 2y + tx^2 + 2xy + \frac{t}{4}y^2,$$

$$P_t(x, y) = -t + 6tx^2 - 2xy + 2y^2.$$

8. Geometria Affine Euclidea

In questo capitolo introduciamo il linguaggio degli *spazi affini euclidei* che consente di descrivere, per esempio, il piano e lo spazio tridimensionale della Geometria Euclidea (in presenza di una unità di misura per le lunghezze). In uno spazio affine euclideo sono disponibili le usuali nozioni metriche: distanze, angoli, ortogonalità, etc.

8.1 Spazi Affini Euclidei

Definizione 8.1.1 — Spazio Affine Euclideo. Uno *spazio affine euclideo* è uno spazio affine reale munito di un prodotto scalare sulla sua giacitura. In altre parole, uno spazio affine euclideo è una coppia $(A, \langle -, - \rangle)$ in cui A è uno spazio affine reale, e $\langle -, - \rangle$ è un prodotto scalare sulla giacitura di A .

Sia $(A, \langle -, - \rangle)$ uno spazio affine euclideo e sia V la giacitura di A . Munito del prodotto scalare $\langle -, - \rangle$, lo spazio vettoriale reale V è uno spazio vettoriale euclideo e si dice che $(A, \langle -, - \rangle)$ è *modellato* su $(V, \langle -, - \rangle)$, o anche che $(V, \langle -, - \rangle)$ è la *giacitura* di $(A, \langle -, - \rangle)$.

■ **Esempio 8.1** Lo spazio affine numerico reale n -dimensionale $A_{\mathbb{R}}^n$ è in modo naturale anche uno spazio affine euclideo quando viene munito del prodotto scalare standard $\langle -, - \rangle_{can}$ sulla sua giacitura \mathbb{R}^n . Lo spazio affine euclideo $(A_{\mathbb{R}}^n, \langle -, - \rangle_{can})$ è anche detto lo *spazio affine euclideo numerico* (o *standard*) n -dimensionale. ■

■ **Esempio 8.2** Lo spazio vettoriale tridimensionale V dei vettori liberi nello spazio 3D della Geometria Euclidea è munito di un prodotto scalare $\langle -, - \rangle_u$ per ogni scelta di un'unità di misura u per le lunghezze. Il prodotto scalare $\langle -, - \rangle_u$ è definito come segue. Siano $v, w \in V$. Indichiamo con $|v|_u$ e $|w|_u$ le lunghezze di v e w espresse nell'unità di misura u , rispettivamente. Inoltre sia θ l'angolo convesso formato da v e w . Si osservi che θ è indipendente dalla scelta dell'unità di misura u . Poniamo

$$\langle v, w \rangle_u := |v|_u |w|_u \cos \theta.$$

Questa definizione ha senso anche nel caso in cui uno o entrambi i vettori v, w siano nulli. Non è difficile dimostrare che $\langle -, - \rangle_u$ così definito è effettivamente un prodotto scalare e invitiamo il

lettore a farlo in tutti i dettagli. Lasciamo aperta la seguente domanda: *come dipende il prodotto scalare $\langle -, - \rangle_u$ dalla scelta dell'unità di misura u ?* In ogni caso, da queste osservazioni segue che la scelta di un'unità di misura per le lunghezze dà automaticamente una struttura di spazio affine euclideo (tridimensionale) allo spazio 3D della Geometria Euclidea. ■

Sia $(A, \langle -, - \rangle)$ uno spazio affine euclideo. Come accennato, in questo contesto ha senso parlare di *distanza* tra punti in base alla

Definizione 8.1.2 — Distanza. La *distanza* in $(A, \langle -, - \rangle)$ è la funzione

$$d_A : A \times A \rightarrow \mathbb{R}, \quad (P, Q) \mapsto d_A(P, Q) := \|\overrightarrow{PQ}\|.$$

Per distinguere la distanza tra punti dalla distanza tra vettori, indicheremo quest'ultima con il simbolo $d_V : V \times V \rightarrow \mathbb{R}$.

La distanza d_A gode chiaramente delle seguenti proprietà:

- (1) $d_A(P, Q) = d_A(Q, P)$, per ogni $P, Q \in A$,
- (2) $d_A(P, Q) \geq 0$, per ogni $P, Q \in A$, e $d_A(P, Q) = 0$ sse $P = Q$.

■ **Esempio 8.3** La distanza nello spazio affine euclideo standard $(A_{\mathbb{R}}^n, \langle -, - \rangle_{can})$ coincide chiaramente con la distanza tra vettori numerici indotta dal prodotto scalare standard. ■

Proposizione 8.1.1 — Disuguaglianza Triangolare. Sia $(A, \langle -, - \rangle)$ uno spazio affine euclideo. Per ogni terna di punti $P, Q, R \in A$ si ha che

$$d_A(Q, R) \leq d_A(Q, P) + d_A(P, R). \quad (8.1)$$

Dimostrazione.

$$d_A(Q, R) = \|\overrightarrow{QR}\| = \|\overrightarrow{QP} + \overrightarrow{PR}\| \leq \|\overrightarrow{QP}\| + \|\overrightarrow{PR}\| = d_A(Q, P) + d_A(P, R),$$

dove abbiamo usato la disuguaglianza triangolare per la distanza euclidea d_V nella versione data dalla (2.5). ■

Come già negli spazi vettoriali euclidei, la disuguaglianza triangolare mostra che, anche in uno spazio affine euclideo, “*in ogni triangolo, la lunghezza di ogni lato è minore della somma delle lunghezze degli altri due lati*” (Figura 8.1).

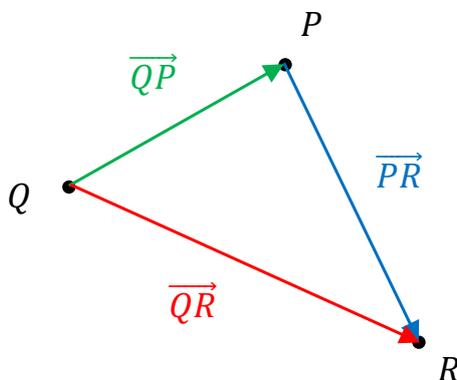


Figura 8.1: Disuguaglianza triangolare in uno spazio affine euclideo.

- La Disuguaglianza Triangolare (8.1) insieme all'osservazione che precede l'Esempio 8.3, mostra che (come già avviene per gli spazi vettoriali euclidei) *uno spazio affine euclideo, munito della distanza, è uno spazio metrico* (si veda l'osservazione a pag. 41 per la definizione di spazio metrico).

Proposizione 8.1.2 — Teorema di Pitagora. Sia $(A, \langle -, - \rangle)$ uno spazio affine euclideo e siano $P, Q, R \in A$ punti di A . Se i vettori \overrightarrow{PQ} e \overrightarrow{PR} sono ortogonali, allora

$$d_A(Q, R)^2 = d_A(P, Q)^2 + d_A(P, R)^2.$$

(Figura 8.2).

Dimostrazione. Calcoliamo

$$\begin{aligned} d_A(Q, R)^2 &= \|\overrightarrow{QR}\|^2 = \|\overrightarrow{QP} + \overrightarrow{PR}\|^2 = \langle \overrightarrow{QP} + \overrightarrow{PR}, \overrightarrow{QP} + \overrightarrow{PR} \rangle \\ &= \langle \overrightarrow{QP}, \overrightarrow{QP} \rangle + \langle \overrightarrow{PR}, \overrightarrow{PR} \rangle + 2\langle \overrightarrow{QP}, \overrightarrow{PR} \rangle = \|\overrightarrow{QP}\|^2 + \|\overrightarrow{PR}\|^2 \\ &= d_A(P, Q)^2 + d_A(P, R)^2, \end{aligned}$$

dove abbiamo usato che $\langle \overrightarrow{QP}, \overrightarrow{PR} \rangle = 0$. ■

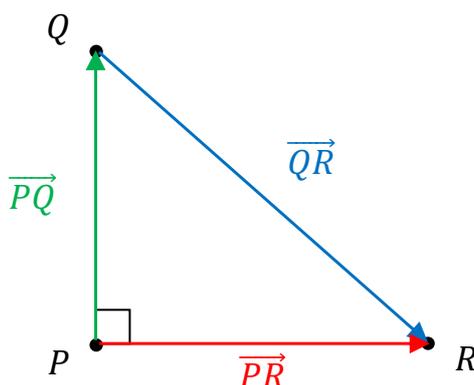


Figura 8.2: Teorema di Pitagora: $d_A(Q, R)^2 = d_A(P, Q)^2 + d_A(P, R)^2$.

In uno spazio affine euclideo ha anche senso parlare di *angolo tra due rette*.

Definizione 8.1.3 — Angolo tra Rette. L'angolo $\widehat{rr'}$ tra due rette $r, r' \subseteq A$ nello spazio affine euclideo $(A, \langle -, - \rangle)$ è l'angolo convesso $\widehat{vv'}$ tra un vettore direttore v di r e un vettore direttore v' di r' .

Sottolineiamo che l'angolo tra due rette r, r' è definito anche se r, r' non si intersecano. Si osservi però che esso dipende, almeno parzialmente, dalla scelta dei vettori direttori v, v' . Precisamente, moltiplicando uno dei due vettori direttori per uno scalare positivo (ottenendo un altro vettore direttore), l'angolo $\widehat{rr'}$ non cambia, ma moltiplicando uno dei vettori direttori per uno scalare negativo l'angolo $\widehat{rr'}$ cambia per il suo supplementare. Questo fenomeno esprime il fatto che, date due rette r, r' , esse formano, di fatto, non uno, ma due angoli supplementari (Figura 8.3). Per eliminare l'ambiguità nella Definizione 8.1.3 si parla talvolta di *angolo tra due rette riferito ad una coppia di vettori direttori*.

Infine, in uno spazio affine euclideo ha senso parlare di *ortogonalità di sottospazi affini*. Cominciamo con il dire che, dato un sottospazio affine $S \subseteq A$ di giacitura $W \subseteq V$ in uno spazio

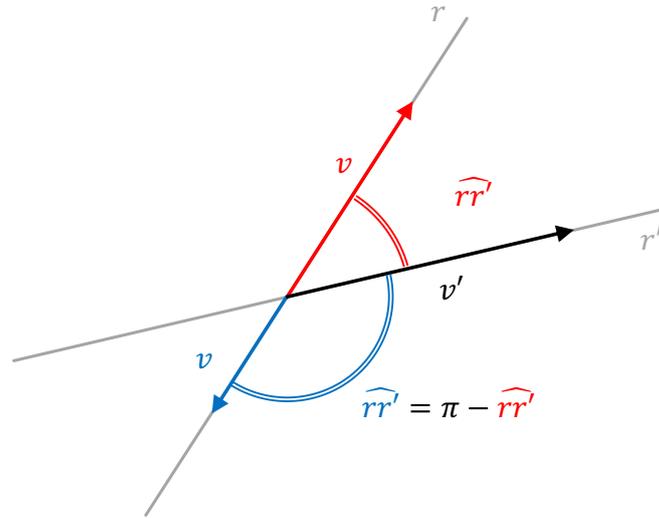


Figura 8.3: Angolo tra due rette in uno spazio affine euclideo.

affine euclideo $(A, \langle -, - \rangle)$ modellato sullo spazio vettoriale euclideo $(V, \langle -, - \rangle)$, e dato un vettore $v \in V$, si dice che v è *ortogonale ad S* se v appartiene al complemento ortogonale di W : $v \in W^\perp$.

Definizione 8.1.4 — Sottospazi Affini Ortogonali. Due sottospazi affini $S_1, S_2 \subseteq A$ di giaciture W_1, W_2 nello spazio affine euclideo $(A, \langle -, - \rangle)$ si dicono *ortogonali* se $W_1 \subseteq W_2^\perp$ oppure $W_1 \supseteq W_2^\perp$.

La definizione di sottospazi ortogonali non dipende dall'ordine in cui prendiamo i due sottospazi. Infatti se $W_1 \subseteq W_2^\perp$ (risp. $W_1 \supseteq W_2^\perp$), allora $W_1^\perp \supseteq W_2^{\perp\perp} = W_2$ (risp. $W_1^\perp \subseteq W_2^{\perp\perp} = W_2$).

■ **Esempio 8.4** Due rette $r_1, r_2 \subseteq A$ sono ortogonali sse formano un angolo $\widehat{r_1 r_2}$ pari a $\pi/2$. ■

Si osservi che, se S_1, S_2 sono sottospazi ortogonali, non è vero, in generale, che ogni vettore della giacitura W_1 di S_1 è ortogonale ad S_2 . Se, per esempio, $W_1 \supseteq W_2^\perp$ ma $W_1 \neq W_2^\perp$, allora W_1 contiene anche vettori che *non* sono ortogonali ad S_2 (Figura 8.4).

Proposizione / Definizione 8.1.3 — Proiezione Ortogonale su un Sottospazio Affine. Sia $(A, \langle -, - \rangle)$ uno spazio affine euclideo di dimensione n modellato sullo spazio vettoriale euclideo $(V, \langle -, - \rangle)$, e sia $S \subseteq A$ un sottospazio affine di dimensione m e giacitura $W \subseteq V$. Per ogni punto $P \in A$ esiste un unico sottospazio affine $T \subseteq A$ di dimensione $n - m$ tale che $P \in T$ e T è ortogonale ad S . Inoltre T ed S si intersecano esattamente in un punto \bar{P} detto *proiezione ortogonale di P sul sottospazio S* (vedi Figura 8.5).

Dimostrazione. Sia T il sottospazio che passa per P di giacitura W^\perp . In altre parole $T = \alpha_P^{-1}(W^\perp)$. Poiché $\dim W^\perp = \dim V - \dim W = n - m$, T è un sottospazio affine di dimensione $n - m$. Inoltre è chiaro che T ed S sono ortogonali. Ogni altro sottospazio affine di dimensione $n - m$ ortogonale ad S , ha giacitura necessariamente pari a W^\perp , perciò T è anche l'unico tale sottospazio che passa per P . Poiché $W^\perp \oplus W = V$, dal Corollario 6.2.4, T ed S sono incidenti. Ma allora la loro intersezione è un sottospazio affine di giacitura pari a $W^\perp \cap W = \{0\}$, cioè è un punto. Questo conclude la dimostrazione. ■

Nel seguito indicheremo con

$$\pi_W^\perp : V = W \oplus W^\perp \rightarrow W, \quad v \mapsto \pi_W^\perp(v) := v - v^\perp$$

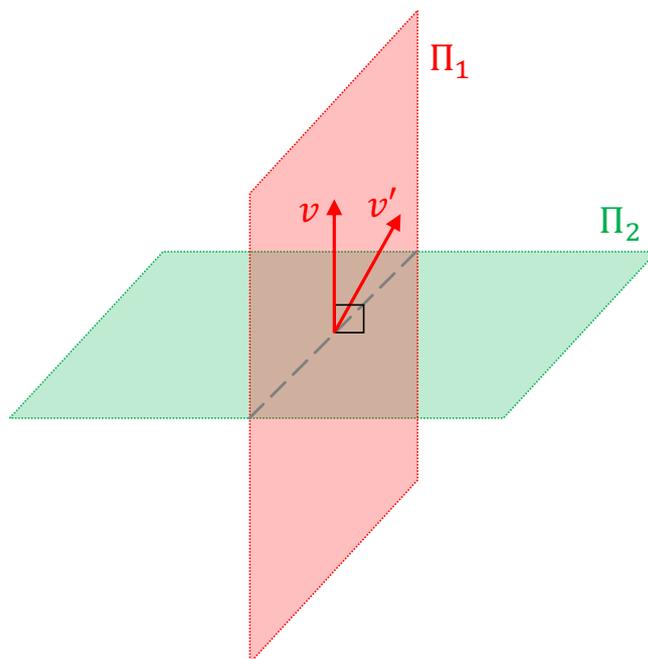


Figura 8.4: Due piani ortogonali Π_1, Π_2 in uno spazio affine euclideo. I vettori v, v' appartengono alla giacitura di Π_1 , ma solo v è ortogonale a Π_2 .

l'applicazione che associa ad un vettore $v \in V$ la sua proiezione ortogonale su W (vedi Formula (2.6), Definizione 2.2.4 e Figura 2.4). Chiaramente π_W^\perp è un'applicazione lineare. Consideriamo anche la funzione

$$\pi_S^\perp : A \rightarrow S, \quad P \mapsto \pi_S^\perp(P) := \bar{P}$$

che associa ad un punto $P \in A$ la sua proiezione ortogonale \bar{P} sul sottospazio affine S .

Proposizione 8.1.4 La funzione $\pi_S^\perp : A \rightarrow S$ è un'applicazione affine con parte lineare data da $\pi_W^\perp : V \rightarrow W$.

Dimostrazione. Siano $P, Q \in A$. Abbiamo (Figura 8.6)

$$\overrightarrow{PQ} = \overrightarrow{PP} + \overrightarrow{PQ} + \overrightarrow{QQ}.$$

Poiché $\bar{P}, \bar{Q} \in S$, abbiamo $\overrightarrow{\bar{P}\bar{Q}} \in W$. Inoltre P e \bar{P} appartengono entrambi al sottospazio affine $T \subseteq A$ di dimensione $\dim A - \dim S$ che passa per P ed è ortogonale ad S (come nella Proposizione 8.1.3). Ma la giacitura di T è precisamente W^\perp , e quindi $\overrightarrow{PP} \in W^\perp$. Similmente $\overrightarrow{QQ} \in W^\perp$. Ma il vettore \overrightarrow{PQ} si scrive in un unico modo come somma di un vettore $\overrightarrow{PQ}^\perp \in W^\perp$ e di un vettore $\overrightarrow{PQ} - \overrightarrow{PQ}^\perp = \pi_W^\perp(\overrightarrow{PQ}) \in W$, perciò deve essere

$$\pi_W^\perp(\overrightarrow{PQ}) = \overrightarrow{\bar{P}\bar{Q}} = \pi_S^\perp(P)\pi_S^\perp(Q),$$

come desiderato. ■

La proiezione ortogonale su un sottospazio affine $S \subseteq A$ consente di definire anche la *distanza di un punto da S* .

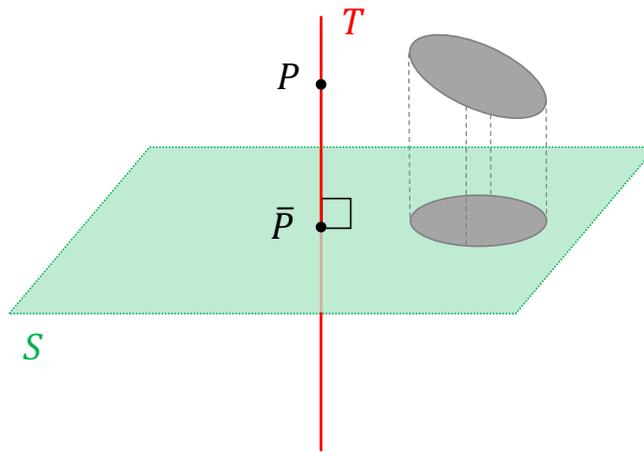


Figura 8.5: Proiezione ortogonale su un sottospazio affine di uno spazio affine euclideo.

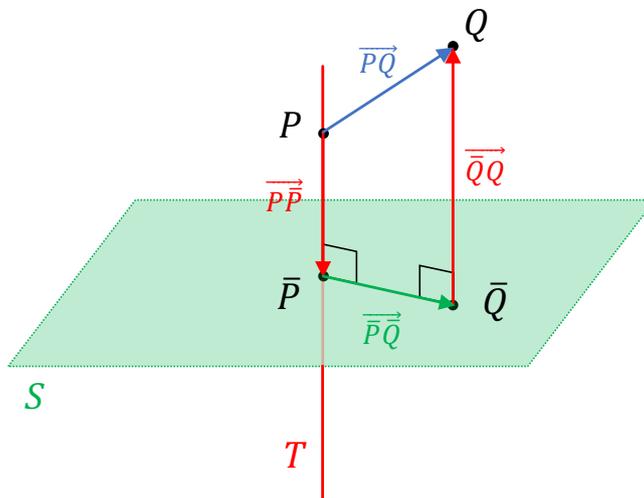


Figura 8.6: La proiezione ortogonale su un sottospazio affine è un'applicazione affine.

Definizione 8.1.5 — Distanza di un Punto da un Sottospazio Affine. La *distanza* di un punto $P \in A$ da un sottospazio affine $S \subseteq A$ in uno spazio affine euclideo $(A, \langle -, - \rangle)$ è

$$d_A(P, S) := d_A(P, \bar{P}),$$

dove $\bar{P} = \pi_S^\perp(P)$ è la proiezione ortogonale di P su S .

Nella prossima sezione troveremo formule per la distanza di un punto da un sottospazio affine in alcuni casi speciali. Ora vogliamo osservare che se $S', S \subseteq A$ sono sottospazi affini paralleli, con $\dim S' \leq \dim S$, allora *tutti i punti di S' hanno la stessa distanza da S* (Figura 8.7). Per dimostrarlo, diciamo W' e W le giaciture di S' ed S rispettivamente. Poichè S', S sono paralleli, abbiamo $W' \subseteq W$. Consideriamo $P, P' \in S'$, dimodochè $\overrightarrow{PP'} \in W' \subseteq W$ e quindi

$$\overrightarrow{P\bar{P}'} = \overrightarrow{\pi_S^\perp(P)\pi_S^\perp(P')} = \pi_W^\perp(\overrightarrow{PP'}) = \overrightarrow{PP'}.$$

Ne consegue che

$$\overrightarrow{P\bar{P}} = \overrightarrow{PP'} + \overrightarrow{P'\bar{P}'} + \overrightarrow{\bar{P}'\bar{P}} = \overrightarrow{PP'} + \overrightarrow{P'\bar{P}'} + \overrightarrow{\bar{P}'\bar{P}} = \overrightarrow{P'\bar{P}'},$$

e dunque

$$d_A(P, S) = d_A(P, \bar{P}) = \|\vec{P\bar{P}}\| = \|\vec{P'\bar{P}'}\| = d_A(P', \bar{P}') = d_A(P', S).$$

In particolare *due sottospazi affini paralleli della stessa dimensione sono equidistanti*.

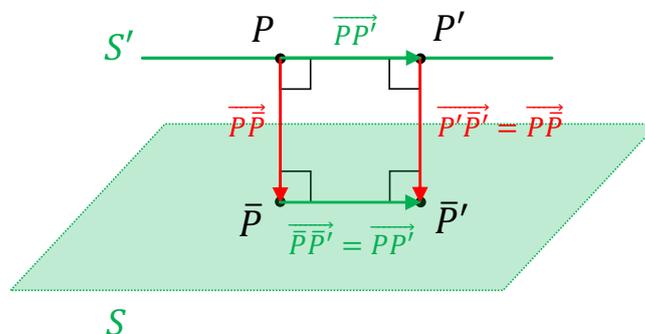


Figura 8.7: Due punti $P, P' \in S'$ hanno la stessa distanza da un sottospazio affine S parallelo ad S' .

Definizione 8.1.6 — Distanza tra Sottospazi Affini Paralleli. La *distanza* tra due sottospazi affini paralleli $S', S \subseteq A$ in uno spazio affine euclideo $(A, \langle -, - \rangle)$, con $\dim S' \leq \dim S$, è

$$d_A(S', S) = d_A(P, S)$$

dove P è un punto qualsiasi di S' .

Esercizio 8.1 Siano $S, S' \subseteq A$ sottospazi affini paralleli della stessa dimensione in uno spazio affine euclideo $(A, \langle -, - \rangle)$. Dimostrare che $d_A(S', S) = d_A(S, S')$. ■

Sia $(A, \langle -, - \rangle)$ uno spazio affine euclideo di dimensione n . È possibile distinguere una famiglia di riferimenti affini di A particolarmente adatta alla struttura aggiuntiva in base alla seguente

Definizione 8.1.7 — Riferimento Cartesiano. Un *riferimento cartesiano* di $(A, \langle -, - \rangle)$ è un riferimento affine (O, \mathcal{R}) tale che \mathcal{R} è un riferimento ortonormale rispetto al prodotto scalare $\langle -, - \rangle$. Le coordinate $c_{(O, \mathcal{R})} : A \rightarrow \mathbb{R}^n$ determinate da un riferimento cartesiano sono dette *coordinate cartesiane*.

■ **Esempio 8.5** Il riferimento affine canonico $(O_{can}, \mathcal{R}_{can})$ dello spazio affine euclideo standard $(A_{\mathbb{R}}^n, \langle -, - \rangle_{can})$ è un riferimento cartesiano. ■

A titolo di esempio, illustriamo come utilizzare un riferimento cartesiano per determinare un sottospazio affine T passante per un punto P , ortogonale ad un sottospazio affine dato S , e di dimensione complementare: $\dim A = \dim S + \dim T$. Sia dunque (O, \mathcal{R}) un riferimento cartesiano di $(A, \langle -, - \rangle)$ e supponiamo di conoscere una rappresentazione parametrica

$$S : \begin{cases} x_1 = x_{01} + d_{11}t_1 + \cdots + d_{1m}t_m \\ \vdots \\ x_n = x_{0n} + d_{n1}t_1 + \cdots + d_{nm}t_m \end{cases}, \quad t_1, \dots, t_m \in \mathbb{R},$$

e una cartesiana

$$S : \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{p1}x_1 + \cdots + a_{pn}x_n = b_p \end{cases}.$$

di S (qui $n = \dim A$, $m = \dim S$ e $p = n - m$). Allora conosciamo anche una rappresentazione parametrica

$$W : \begin{cases} x_1 = d_{11}t_1 + \cdots + d_{1m}t_m \\ \vdots \\ x_n = d_{n1}t_1 + \cdots + d_{nm}t_m \end{cases}, \quad t_1, \dots, t_m \in \mathbb{R},$$

e una cartesiana

$$W : \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0 \\ \vdots \\ a_{p1}x_1 + \cdots + a_{pn}x_n = 0 \end{cases}$$

della giacitura W di S . Dall'Esempio 2.5 sappiamo allora che il complemento ortogonale W^\perp ha rappresentazione cartesiana

$$W^\perp : \begin{cases} d_{11}x_1 + \cdots + d_{n1}x_n = 0 \\ \vdots \\ d_{1m}x_1 + \cdots + d_{nm}x_n = 0 \end{cases}$$

e rappresentazione parametrica

$$W^\perp : \begin{cases} x_1 = a_{11}s_1 + \cdots + a_{p1}s_p \\ \vdots \\ x_n = a_{1n}s_1 + \cdots + a_{pn}s_p \end{cases}, \quad s_1, \dots, s_p \in \mathbb{R}.$$

Di conseguenza, se le coordinate cartesiane di P sono (X_1, \dots, X_n) , allora T ha rappresentazione cartesiana

$$T : \begin{cases} d_{11}(x_1 - X_1) + \cdots + d_{n1}(x_n - X_n) = 0 \\ \vdots \\ d_{1m}(x_1 - X_1) + \cdots + d_{nm}(x_n - X_n) = 0 \end{cases}$$

e rappresentazione parametrica

$$T : \begin{cases} x_1 = X_1 + a_{11}s_1 + \cdots + a_{p1}s_p \\ \vdots \\ x_n = X_n + a_{1n}s_1 + \cdots + a_{pn}s_p \end{cases}, \quad s_1, \dots, s_p \in \mathbb{R}.$$

■ **Esempio 8.6** Nello spazio affine euclideo standard $(A_{\mathbb{R}}^4, \langle -, - \rangle_{can})$ consideriamo il piano

$$\Pi : \begin{cases} x_1 + x_2 + x_3 - x_4 = -1 \\ x_1 - 2x_2 + x_4 = 0 \end{cases}.$$

Per ogni punto $P \in A_{\mathbb{R}}^4$ esiste un unico piano per P ortogonale a Π . Per esempio, in base alle osservazioni appena fatte, il piano per il punto $P = (1, -1, 0, 1/2)^T$ ortogonale a Π è

$$\Pi' : \begin{cases} x_1 = 1 + s + t \\ x_2 = -1 + s - 2t \\ x_3 = s \\ x_4 = \frac{1}{2} - s + t \end{cases}, \quad s, t \in \mathbb{R}. \quad (8.2)$$

Per trovare una rappresentazione cartesiana di Π' dobbiamo eliminare i parametri dalla (8.2) per esempio annullando i due orlati di un minore di ordine 2 non nullo nella matrice

$$\begin{pmatrix} x_1 - 1 & 1 & 1 \\ x_2 + 1 & 1 & -2 \\ x_3 & 1 & 0 \\ x_4 - 1/2 & -1 & 1 \end{pmatrix}.$$

Il minore individuato dalle prime due righe e dalle ultime due colonne è

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

I suoi orlati sono

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

e

$$\det \begin{pmatrix} x_1 - 1 & 1 & 1 \\ x_2 + 1 & 1 & -2 \\ x_4 - 1/2 & -1 & 1 \end{pmatrix} = -x_1 - 2x_2 - 3x_4 + \frac{1}{2}.$$

Concludiamo che

$$\Pi' : \begin{cases} 2x_1 + x_2 - 3x_3 & = 1 \\ -x_1 - 2x_2 - 3x_4 & = -\frac{1}{2} \end{cases}.$$

■

Esercizio 8.2 Nello spazio affine euclideo standard $(A_{\mathbb{R}}^4, \langle -, - \rangle_{can})$ si consideri il piano

$$\Pi : \begin{cases} x_1 = -3 + 2s - \frac{1}{3}t \\ x_2 = 2 - \frac{4}{3}s - t \\ x_3 = s - t \\ x_4 = \frac{1}{3} + \frac{1}{3}s + \frac{1}{3}t \end{cases}, \quad s, t \in \mathbb{R}.$$

Determinare una rappresentazione parametrica e una cartesiana del piano Π' che passa per il punto $(0, 1, 0, 1)^T \in A_{\mathbb{R}}^4$ ed è ortogonale a Π . ■

8.2 Geometria Affine Euclidea in Dimensione 2 e 3

In questa sezione studiamo con particolare riguardo gli spazi affini euclidei di dimensione 2 e 3. Cominciamo dal caso bidimensionale.

Sia $(A, \langle -, - \rangle)$ un piano euclideo (cioè uno spazio affine euclideo di dimensione 2). Fissiamo una volta e per tutte un riferimento cartesiano (O, \mathcal{R}) di $(A, \langle -, - \rangle)$. Indichiamo con (x, y) le corrispondenti coordinate cartesiane. Consideriamo una retta r insieme con una rappresentazione parametrica

$$r : \begin{cases} x = x_0 + tl \\ y = y_0 + tm \end{cases}, \quad t \in \mathbb{R},$$

e una cartesiana

$$r : ax + by = c$$

nel riferimento (O, \mathcal{R}) . I vettori di coordinate (a, b) e $(-m, l)$ sono linearmente dipendenti e sono ortogonali ad r (il lettore lo vede?). Perciò le rette r' ortogonali ad r hanno rappresentazione parametrica del tipo

$$r' : \begin{cases} x = X + sa \\ y = Y + sb \end{cases}, \quad t \in \mathbb{R},$$

e rappresentazione cartesiana del tipo

$$r' : lx + my = C.$$

Data un'altra retta \tilde{r} di rappresentazione parametrica

$$\tilde{r} : \begin{cases} x = \tilde{x}_0 + t\tilde{l} \\ y = \tilde{y}_0 + t\tilde{m} \end{cases}, \quad t \in \mathbb{R},$$

e rappresentazione cartesiana

$$\tilde{r} : \tilde{a}x + \tilde{b}y = c.$$

L'angolo $\widehat{r\tilde{r}}$ coincide per definizione con l'angolo $\widehat{v\tilde{v}}$ tra due vettori direttori, per esempio quelli di coordinate (l, m) e (\tilde{l}, \tilde{m}) . Poiché \mathcal{R} è un riferimento ortonormale, abbiamo

$$\begin{aligned} \cos \widehat{r\tilde{r}} &= \frac{\langle v, \tilde{v} \rangle}{\|v\| \|\tilde{v}\|} = \frac{\left\langle \begin{pmatrix} l \\ m \end{pmatrix}, \begin{pmatrix} \tilde{l} \\ \tilde{m} \end{pmatrix} \right\rangle_{can}}{\sqrt{\left\langle \begin{pmatrix} l \\ m \end{pmatrix}, \begin{pmatrix} l \\ m \end{pmatrix} \right\rangle_{can} \left\langle \begin{pmatrix} \tilde{l} \\ \tilde{m} \end{pmatrix}, \begin{pmatrix} \tilde{l} \\ \tilde{m} \end{pmatrix} \right\rangle_{can}}} \\ &= \frac{\tilde{l}l + m\tilde{m}}{\sqrt{(l^2 + m^2)(\tilde{l}^2 + \tilde{m}^2)}}. \end{aligned} \quad (8.3)$$

Giacché possiamo scegliere $(a, b) = (-m, l)$ e $(\tilde{a}, \tilde{b}) = (-\tilde{m}, \tilde{l})$, abbiamo anche

$$\cos \widehat{r\tilde{r}} = \frac{a\tilde{a} + b\tilde{b}}{\sqrt{(a^2 + b^2)(\tilde{a}^2 + \tilde{b}^2)}} = \widehat{w\tilde{w}},$$

dove w, \tilde{w} sono i due vettori ortogonali a r, \tilde{r} di coordinate rispettivamente $(a, b), (\tilde{a}, \tilde{b})$. Concludiamo che *l'angolo tra due rette nel piano coincide con l'angolo tra due vettori ad esse ortogonali* (Figura 8.8). La Formula (8.3) conferma, tra l'altro, che *due rette nel piano sono parallele sse formano un'angolo pari a 0 o π e sono ortogonali sse formano un'angolo pari a $\pi/2$* . Sottolineiamo anche che la *condizione di ortogonalità* tra le rette r, \tilde{r} si può scrivere tanto nella forma

$$\tilde{l}l + m\tilde{m} = 0$$

quanto nella forma

$$a\tilde{a} + b\tilde{b} = 0.$$

Discutiamo ora la distanza di un punto dalla retta r .

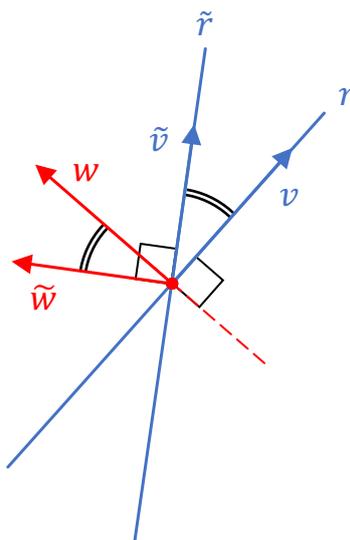


Figura 8.8: Angolo tra due rette nel piano (e angolo tra i vettori ad esse ortogonali).

Proposizione 8.2.1 Sia $P \in A$ e siano (X, Y) le coordinate di P nel riferimento cartesiano (O, \mathcal{R}) . La distanza di P da r è data dalla seguente formula:

$$d_A(P, r) = \frac{|aX + bY - c|}{\sqrt{a^2 + b^2}}.$$

Dimostrazione. Consideriamo un punto $Q \in r$. Siano (x_0, y_0) le sue coordinate cartesiane. Consideriamo anche

- (1) la proiezione ortogonale \overline{P} di P su r ;
- (2) un versore $u \in V$ ortogonale ad r .

Chiaramente u genera il complemento ortogonale della giacitura di r . Perciò esiste $\alpha \in \mathbb{R}$ tale che $\overrightarrow{P\overline{P}} = \alpha u$. Inoltre, da un lato

$$\langle \overrightarrow{P\overline{P}}, u \rangle = \langle \alpha u, u \rangle = \alpha \langle u, u \rangle = \alpha,$$

dall'altro

$$d_A(P, r) = \|\overrightarrow{P\overline{P}}\| = \|\alpha u\| = |\alpha| \|u\| = |\alpha|.$$

Mettendo assieme le ultime due uguaglianze troviamo

$$d_A(P, r) = |\alpha| = \left| \langle \overrightarrow{P\overline{P}}, u \rangle \right| = \left| \langle \overrightarrow{P\overline{Q}} + \overrightarrow{\overline{Q}\overline{P}}, u \rangle \right| = \left| \langle \overrightarrow{P\overline{Q}}, u \rangle + \langle \overrightarrow{\overline{Q}\overline{P}}, u \rangle \right| = \left| \langle \overrightarrow{P\overline{Q}}, u \rangle \right|$$

dove, nell'ultimo passaggio, abbiamo usato che u è ortogonale ad r (e che $\overrightarrow{P\overline{Q}}$ appartiene alla giacitura di r - vedi Figura 8.9). Ora il versore u può essere scelto normalizzando un vettore w ortogonale ad r , cioè possiamo scegliere

$$u = \frac{1}{\|w\|} w.$$

Se per w scegliamo il vettore di coordinate cartesiane (a, b) allora $\|w\| = \sqrt{a^2 + b^2}$ e u ha coordinate cartesiane

$$\left(\frac{a}{\sqrt{a^2 + b^2}}, \frac{b}{\sqrt{a^2 + b^2}} \right).$$

Il vettore \overrightarrow{PQ} ha coordinate cartesiane

$$(X - x_0, Y - y_0)$$

perciò

$$\begin{aligned} d_A(P, r) &= \left| \langle \overrightarrow{PQ}, u \rangle \right| = \left| \left\langle \begin{pmatrix} X - x_0 \\ Y - y_0 \end{pmatrix}, \begin{pmatrix} \frac{a}{\sqrt{a^2 + b^2}} \\ \frac{b}{\sqrt{a^2 + b^2}} \end{pmatrix} \right\rangle_{can} \right| = \frac{|a(X - x_0) + b(Y - y_0)|}{\sqrt{a^2 + b^2}} \\ &= \frac{|aX + bY - ax_0 - by_0|}{\sqrt{a^2 + b^2}}. \end{aligned}$$

Poiché $Q \in r$ le sue coordinate soddisfano

$$ax_0 + by_0 = c$$

e concludiamo che

$$d_A(P, r) = \frac{|aX + bY - c|}{\sqrt{a^2 + b^2}},$$

come desiderato. ■

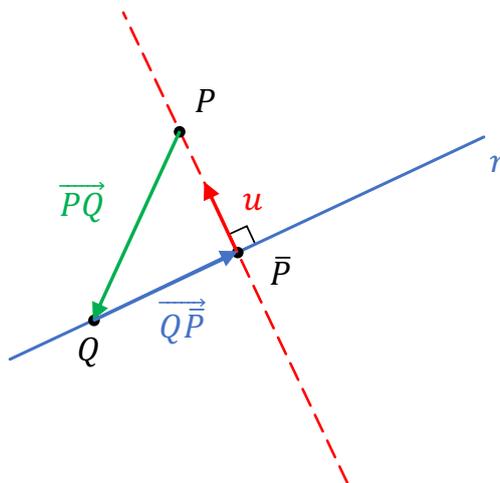


Figura 8.9: Distanza di un punto P da una retta r nel piano euclideo.

Esercizio 8.3 Trovare una formula per la distanza di un punto P di coordinate cartesiane (X, Y) dalla retta r in termini della rappresentazione parametrica di r . ■

Esercizio 8.4 Sia r' una retta parallela ad r . In particolare r' avrà una rappresentazione cartesiana del tipo

$$r' : ax + by = c'.$$

Dimostrare che la distanza di r' da r è data da

$$d_A(r', r) = \frac{|c - c'|}{\sqrt{a^2 + b^2}}.$$

Esercizio 8.5 Nel piano euclideo standard $(A_{\mathbb{R}}^2, \langle -, - \rangle_{can})$ si considerino la retta

$$r : x - 2y = 3$$

e il punto $P = (1, 1)^T$. Determinare

- (1) una rappresentazione parametrica e una cartesiana della retta per P ortogonale ad r ;
- (2) la proiezione ortogonale di P su r ;
- (3) la distanza di P da r .

Passiamo ora a studiare uno spazio affine euclideo tridimensionale $(A, \langle -, - \rangle)$. Fissiamo di nuovo un riferimento cartesiano (O, \mathcal{R}) e indichiamo con (x, y, z) le corrispondenti coordinate cartesiane. Utilizzeremo anche il prodotto vettoriale \wedge in V determinato dal riferimento \mathcal{R} (o, per meglio dire, dalla orientazione associata ad \mathcal{R}). Consideriamo un piano $\Pi \subseteq A$, una rappresentazione parametrica

$$\Pi : \begin{cases} x = X_0 + sL + tL' \\ y = Y_0 + sM + tM' \\ z = Z_0 + sN + tN' \end{cases}, \quad s, t \in \mathbb{R}, \quad (8.4)$$

e una cartesiana

$$\Pi : Ax + By + Cz = D. \quad (8.5)$$

Il vettore di coordinate (A, B, C) è ortogonale a Π . Per costruire un vettore ortogonale a Π dalla sua rappresentazione parametrica possiamo ragionare come segue: siano v, v' i vettori di coordinate $(L, M, N), (L', M', N')$ essi sono due vettori direttori di Π . Il vettore $v \wedge v'$ è diverso da 0 ed è ortogonale sia a v che a v' , perciò è ortogonale a Π . Ricordiamo che esso ha coordinate

$$\left(\begin{vmatrix} M & M' \\ N & N' \end{vmatrix}, - \begin{vmatrix} L & L' \\ N & N' \end{vmatrix}, \begin{vmatrix} L & L' \\ M & M' \end{vmatrix} \right)$$

Le rette r' ortogonali a Π hanno rappresentazione parametrica del tipo

$$r'' : \begin{cases} x = x'_0 + tA \\ y = y'_0 + tB \\ z = z'_0 + tC \end{cases}, \quad t \in \mathbb{R}, \quad (8.6)$$

e rappresentazione cartesiana del tipo

$$r' : \begin{cases} Lx + My + Nz = d \\ L'x + M'y + N'z = d' \end{cases}. \quad (8.7)$$

Viceversa, data una retta r' di rappresentazione parametrica (8.6) e rappresentazione cartesiana (8.7), i piani Π ortogonali ad r' hanno rappresentazione parametrica (8.4) e rappresentazione cartesiana (8.5). Date due rette

$$r : \begin{cases} x = x_0 + tl \\ y = y_0 + tm \\ z = z_0 + tn \end{cases}, \quad t \in \mathbb{R}, \quad (8.8)$$

ed

$$\tilde{r} : \begin{cases} x = \tilde{x}_0 + t\tilde{l} \\ y = \tilde{y}_0 + t\tilde{m} \\ z = \tilde{z}_0 + t\tilde{n} \end{cases}, \quad t \in \mathbb{R}$$

l'angolo $\widehat{r\tilde{r}}$ è dato da

$$\cos \widehat{r\tilde{r}} = \frac{\tilde{l} + m\tilde{m} + n\tilde{n}}{\sqrt{(l^2 + m^2 + n^2)(\tilde{l}^2 + \tilde{m}^2 + \tilde{n}^2)}}.$$

In questo contesto ha anche senso parlare dell'*angolo* $\widehat{\Pi r}$ tra il piano Π e la retta r : esso è definito come il complementare dell'angolo tra r ed una qualsiasi retta ortogonale a Π (Figura 8.10). Perciò

$$\widehat{\Pi r} := \frac{\pi}{2} - \widehat{r r'} \Rightarrow \sin \widehat{\Pi r} = \frac{lA + mB + nC}{\sqrt{(l^2 + m^2 + n^2)(A^2 + B^2 + C^2)}}, \quad \widehat{\Pi r} \in [-\pi/2, \pi/2].$$

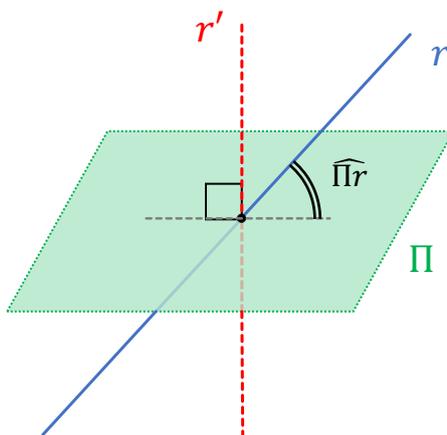


Figura 8.10: Angolo tra un piano Π e una retta r in uno spazio affine euclideo tridimensionale.

Si osservi che, in base alla definizione di angolo tra piano e retta, una retta e un piano sono paralleli sse formano un angolo pari a 0 o π , mentre sono ortogonali sse formano un angolo di $\pm\pi/2$ (il lettore lo vede?). Dato un altro piano

$$\tilde{\Pi} : \tilde{A}x + \tilde{B}y + \tilde{C}z = \tilde{D},$$

si definisce *angolo* $\widehat{\Pi\tilde{\Pi}}$ tra Π e $\tilde{\Pi}$, l'angolo formato da due rette ortogonali a Π e $\tilde{\Pi}$ rispettivamente (Figura 8.11). Dunque

$$\cos \widehat{\Pi\tilde{\Pi}} = \frac{A\tilde{A} + B\tilde{B} + C\tilde{C}}{\sqrt{(A^2 + B^2 + C^2)(\tilde{A}^2 + \tilde{B}^2 + \tilde{C}^2)}}, \quad \widehat{\Pi\tilde{\Pi}} \in [0, \pi].$$

In base a questa definizione due piani sono paralleli sse formano un angolo pari a 0 o π , e sono ortogonali sse formano un angolo pari a $\pi/2$. In particolare, la *condizione di ortogonalità* dei piani $\Pi, \tilde{\Pi}$ si scrive

$$A\tilde{A} + B\tilde{B} + C\tilde{C} = 0.$$

■ **Esempio 8.7** Nello spazio affine euclideo standard $(A_{\mathbb{R}}^3, \langle -, - \rangle_{can})$ consideriamo il piano

$$\Pi : x - 2y + 2z = -1$$

e la retta

$$r : \begin{cases} x + 2y - z = 2 \\ -\frac{1}{2}x + y + \frac{1}{2}z = 3 \end{cases} .$$

Il vettore v di coordinate $(1, -2, 2)$ è ortogonale a Π , mentre il vettore w di coordinate

$$\left(\begin{vmatrix} 2 & -1 \\ 1 & 1/2 \end{vmatrix}, -\begin{vmatrix} 1 & -1 \\ -1/2 & 1/2 \end{vmatrix}, \begin{vmatrix} 1 & 2 \\ -1/2 & 1 \end{vmatrix} \right) = (2, 0, 2)$$

è un vettore direttore per r . Poiché v e w non sono né linearmente dipendenti, né ortogonali, Π ed r non sono né paralleli, né ortogonali. In effetti, essi formano un angolo $\theta \in [-\pi/2, \pi/2]$ con

$$\sin \theta = \frac{\langle v, w \rangle_{can}}{\|v\| \|w\|} = \frac{1 \cdot 2 - 2 \cdot 0 + 2 \cdot 2}{\sqrt{(1^2 + 2^2 + 2^2)(2^2 + 0^2 + 2^2)}} = \frac{\sqrt{2}}{2},$$

cioè $\theta = \pi/4$. Esiste un unico piano Π' che contiene r ed è ortogonale a Π . Per dimostrarlo, osserviamo che, per contenere r , Π' deve avere rappresentazione cartesiana del tipo (Equazione del Fascio Proprio)

$$\Pi' : \lambda(x + 2y - z) + \lambda' \left(-\frac{1}{2}x + y + \frac{1}{2}z \right) = 2\lambda + 3\lambda',$$

che riscritta opportunamente è

$$\Pi' : \left(\lambda - \frac{1}{2}\lambda' \right) x + (2\lambda + \lambda') y + \left(-\lambda + \frac{1}{2}\lambda' \right) z = 2\lambda + 3\lambda'. \quad (8.9)$$

Dunque il vettore v' di coordinate

$$\left(\lambda - \frac{1}{2}\lambda', 2\lambda + \lambda', -\lambda + \frac{1}{2}\lambda' \right)$$

è ortogonale a Π' . La condizione di ortogonalità tra Π e Π' si scrive ora $\langle v, v' \rangle_{can} = 0$:

$$\lambda - \frac{1}{2}\lambda' - 2(2\lambda + \lambda') + 2 \left(-\lambda + \frac{1}{2}\lambda' \right) = 0$$

che dà $\lambda' = -10\lambda/3$. Dunque, a meno di un fattore di proporzionalità, c'è un'unica rappresentazione cartesiana per Π' che soddisfa tutte le richieste. Dunque il piano Π' esiste ed è unico. La sua rappresentazione cartesiana si ottiene ponendo, per esempio, $\lambda = 3$, e quindi necessariamente $\lambda' = -10$, nella (8.9) ed è:

$$\Pi' : 8x - 4y - 8z = -24,$$

o, dopo un'ulteriore ovvia semplificazione,

$$\Pi' : 2x - y - 2z = -6.$$

■

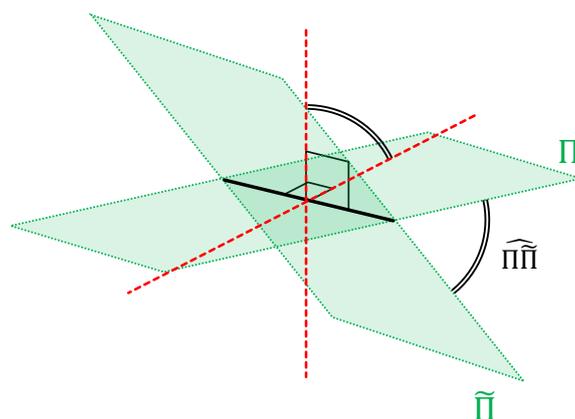


Figura 8.11: Angolo tra due piani $\Pi, \tilde{\Pi}$ in uno spazio affine euclideo tridimensionale.

Proposizione 8.2.2 Sia $P \in A$ e siano (X, Y, Z) le coordinate di P nel riferimento cartesiano (O, \mathcal{R}) . La distanza di P dal piano Π di rappresentazione cartesiana (8.5) è data dalla seguente formula:

$$d_A(P, \Pi) = \frac{|AX + BY + CZ - D|}{\sqrt{A^2 + B^2 + C^2}}.$$

Dimostrazione. La dimostrazione è molto simile a quella della Proposizione 8.2.1 ed è lasciata come Esercizio 8.6. ■

Esercizio 8.6 Dimostrare la Proposizione 8.2.2. Inoltre, sia Π' un piano parallelo a Π . Dunque Π' ha rappresentazione cartesiana del tipo

$$\Pi' : Ax + By + Cz = D'.$$

Dimostrare che la distanza di Π' da Π è data da

$$d_A(\Pi', \Pi) = \frac{|D - D'|}{\sqrt{A^2 + B^2 + C^2}}.$$

Esercizio 8.7 Nello spazio affine euclideo standard $(A_{\mathbb{R}}^3, \langle -, - \rangle_{can})$ si considerino il piano

$$\Pi : 2x - y - \frac{2}{3}z = -2,$$

la retta

$$r : \begin{cases} 3x - y - 4z = -8 \\ x - z = -\frac{8}{3} \end{cases},$$

e il punto $P = (-3, 1, 2/3)^T$. Determinare

- (1) una rappresentazione parametrica e una cartesiana della retta per P ortogonale a Π ;
- (2) la proiezione ortogonale di P su Π ;

- (3) la distanza di P da Π ;
 (4) una rappresentazione parametrica e una cartesiana della retta r' che passa per P , è parallela a Π e ortogonale ad r .

Argomentare sul perché la retta r' esiste ed è unica e, dopo aver osservato che r, r' non sono parallele, stabilire se r, r' sono incidenti o sono sghembe. ■

Proposizione 8.2.3 Sia $P \in A$ e sia $r \subseteq A$ una retta. Se Q è un punto di r e v è un vettore direttore, allora la distanza di P da r è data dalla formula

$$d_A(P, r) = \frac{\|\overrightarrow{PQ} \wedge v\|}{\|v\|}.$$

Se inoltre P ha coordinate (X, Y, Z) e r ha rappresentazione parametrica (8.8) nel riferimento (O, \mathcal{R}) , allora si ha anche

$$d_A(P, r) = \sqrt{\frac{\begin{vmatrix} y_0 - Y & m \\ z_0 - Z & n \end{vmatrix}^2 + \begin{vmatrix} x_0 - X & l \\ z_0 - Z & n \end{vmatrix}^2 + \begin{vmatrix} x_0 - X & l \\ y_0 - Y & m \end{vmatrix}^2}{l^2 + m^2 + n^2}}. \quad (8.10)$$

Dimostrazione. Indichiamo con \bar{P} la proiezione ortogonale di P su r . Poiché i vettori $\overrightarrow{P\bar{P}}$ e v sono ortogonali, dalla Proposizione 5.3.1.(3) abbiamo.

$$\|\overrightarrow{P\bar{P}} \wedge v\| = \|\overrightarrow{P\bar{P}}\| \|v\|.$$

Ma $v \neq 0$, perciò

$$d_A(P, r) = \|\overrightarrow{P\bar{P}}\| = \frac{\|\overrightarrow{P\bar{P}} \wedge v\|}{\|v\|} = \frac{\|(\overrightarrow{PQ} + \overrightarrow{Q\bar{P}}) \wedge v\|}{\|v\|} = \frac{\|\overrightarrow{PQ} \wedge v + \overrightarrow{Q\bar{P}} \wedge v\|}{\|v\|} = \frac{\|\overrightarrow{PQ} \wedge v\|}{\|v\|},$$

dove abbiamo usato che $\overrightarrow{Q\bar{P}}$ dipende linearmente da v (oltre alla bilinearità del prodotto vettoriale). Infine, se v è il vettore di coordinate (l, m, n) e Q è il punto di coordinate cartesiane (x_0, y_0, z_0) , allora l'uguaglianza

$$d_A(P, r) = \frac{\|\overrightarrow{PQ} \wedge v\|}{\|v\|},$$

appena ottenuta dà esattamente la (8.10) (il lettore lo vede?). ■

Esercizio 8.8 Nello spazio affine euclideo standard $(A_{\mathbb{R}}^3, \langle -, - \rangle_{can})$ si considerino la retta

$$r: \begin{cases} 3x - y = 6 \\ -4x - z = -7 \end{cases}$$

e il punto $P = (1, 0, 1)^T$. Determinare

- (1) una rappresentazione parametrica e una cartesiana del piano per P ortogonale ad r ;
- (2) la proiezione ortogonale di P su r ;
- (3) la distanza di P da r ;
- (4) una rappresentazione parametrica e una cartesiana del piano Π' che contiene r e P ;
- (5) una rappresentazione parametrica e una cartesiana della retta r' per P ortogonale a Π' ;
- (6) la posizione reciproca di r ed r' .

In uno spazio affine euclideo tridimensionale ha anche senso parlare di *distanza tra due rette non parallele*. Prima di dare una definizione occorre dimostrare la seguente

Proposizione 8.2.4 Sia $(A, \langle -, - \rangle)$ uno spazio affine euclideo di dimensione 3 e siano $r_1, r_2 \subseteq A$ rette non parallele. Allora esiste un unico punto $N_1 \in r_1$ ed un unico punto $N_2 \in r_2$ tali che il vettore $\overrightarrow{N_1 N_2}$ è ortogonale sia ad r_1 che ad r_2 . Inoltre esiste un'unica retta \tilde{r} passante per N_1 ed N_2 e ortogonale sia ad r_1 che ad r_2 (Figura 8.12).

Dimostrazione. Fissiamo punti $Q_i \in r_i$ e vettori direttori v_i di r_i , $i = 1, 2$. Un punto $N_i \in A$ appartiene ad r_i sse esiste $t_i \in \mathbb{R}$ tale che

$$\overrightarrow{Q_i N_i} = t_i v_i, \quad i = 1, 2. \quad (8.11)$$

Se fissiamo anche un altro punto $O \in A$ (per esempio l'origine di un riferimento affine (O, \mathcal{R})), la (8.11) si riscrive

$$\overrightarrow{O N_i} = \overrightarrow{O Q_i} + t_i v_i$$

(che poi, riscritta in coordinate rispetto al riferimento \mathcal{R} , non è altro che una rappresentazione parametrica di r_i), $i = 1, 2$. Noi stiamo cercando $t_1, t_2 \in \mathbb{R}$ tali che, in aggiunta, $\overrightarrow{N_1 N_2} \perp v_1, v_2$, cioè $\langle \overrightarrow{N_1 N_2}, v_1 \rangle = \langle \overrightarrow{N_1 N_2}, v_2 \rangle = 0$. Ma

$$\overrightarrow{N_1 N_2} = \overrightarrow{O N_2} - \overrightarrow{O N_1} = \overrightarrow{O Q_2} + t_2 v_2 - \overrightarrow{O Q_1} - t_1 v_1 = \overrightarrow{Q_1 Q_2} + t_2 v_2 - t_1 v_1,$$

e dunque stiamo cercando $t_1, t_2 \in \mathbb{R}$ tali che

$$\begin{aligned} 0 &= \langle \overrightarrow{Q_1 Q_2} + t_2 v_2 - t_1 v_1, v_1 \rangle = -\langle v_1, v_1 \rangle t_1 + \langle v_2, v_1 \rangle t_2 + \langle \overrightarrow{Q_1 Q_2}, v_1 \rangle, \\ 0 &= \langle \overrightarrow{Q_1 Q_2} + t_2 v_2 - t_1 v_1, v_2 \rangle = -\langle v_1, v_2 \rangle t_1 + \langle v_2, v_2 \rangle t_2 + \langle \overrightarrow{Q_1 Q_2}, v_2 \rangle. \end{aligned}$$

Riassumendo, N_1, N_2 come nell'enunciato esistono sse esistono $t_1, t_2 \in \mathbb{R}$ tali che

$$\begin{cases} \langle v_1, v_1 \rangle t_1 - \langle v_2, v_1 \rangle t_2 = \langle \overrightarrow{Q_1 Q_2}, v_1 \rangle \\ \langle v_1, v_2 \rangle t_1 - \langle v_2, v_2 \rangle t_2 = \langle \overrightarrow{Q_1 Q_2}, v_2 \rangle \end{cases} \quad (8.12)$$

In altre parole sse il sistema lineare di due equazioni in due incognite (8.12) è compatibile. La matrice incompleta del sistema (8.12) è

$$\begin{pmatrix} \langle v_1, v_1 \rangle & -\langle v_2, v_1 \rangle \\ \langle v_1, v_2 \rangle & -\langle v_2, v_2 \rangle \end{pmatrix}$$

il cui determinante è

$$\det \begin{pmatrix} \langle v_1, v_1 \rangle & -\langle v_2, v_1 \rangle \\ \langle v_1, v_2 \rangle & -\langle v_2, v_2 \rangle \end{pmatrix} = -\det \begin{pmatrix} \langle v_1, v_1 \rangle & \langle v_2, v_1 \rangle \\ \langle v_1, v_2 \rangle & \langle v_2, v_2 \rangle \end{pmatrix} = -\|v_1 \wedge v_2\|^2 \neq 0,$$

dove l'ultima disuguaglianza è dovuta all'ipotesi che r_1 e r_2 non sono parallele (e dunque i loro vettori direttori non sono linearmente dipendenti). In virtù del Teorema di Rouché-Capelli, il sistema (8.12) ammette perciò una ed una sola soluzione e dunque esiste solo una coppia di punti N_1, N_2 come nell'enunciato.

Per la seconda parte dell'enunciato distinguiamo due casi:

- (1) r_1 ed r_2 sono incidenti. In questo caso, non essendo parallele, si intersecano in un solo punto N . Poiché $\overrightarrow{NN} = 0$ e dunque \overrightarrow{NN} è ortogonale sia ad r_1 che ad r_2 , dall'unicità, deve necessariamente essere $N_1 = N_2 = N$. Poiché le giaciture di r_1 ed r_2 generano un sottospazio bidimensionale il cui complemento ortogonale ha dunque dimensione 1, esiste esattamente una retta \tilde{r} per N ortogonale sia ad r_1 che ad r_2 .
- (2) r_1 ed r_2 sono sghembe. In questo caso $N_1 \neq N_2$ e la retta \tilde{r} è necessariamente l'unica retta che contiene sia N_1 che N_2 .

■

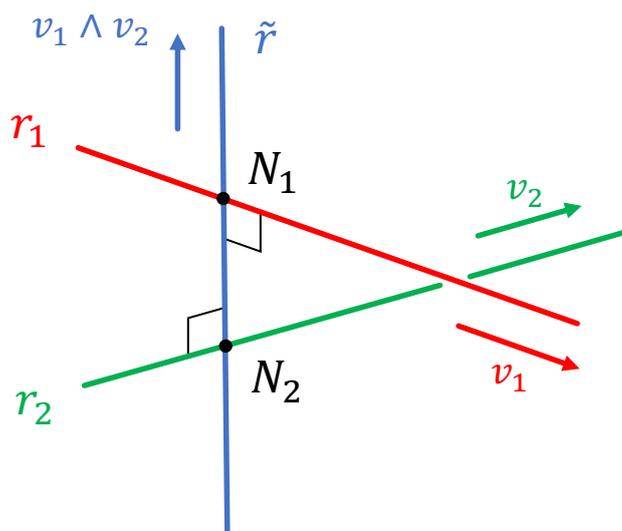


Figura 8.12: Perpendicolare comune \tilde{r} a due rette non parallele r_1, r_2 in uno spazio affine euclideo tridimensionale.

Definizione 8.2.1 — Distanza tra Due Rette Non Parallele. La *distanza* tra due rette non parallele r_1, r_2 in uno spazio affine euclideo tridimensionale $(A, \langle -, - \rangle)$ è

$$d_A(r_1, r_2) := d_A(N_1, N_2).$$

La retta \tilde{r} nella Proposizione 8.2.4 è detta la *perpendicolare comune* ad r_1 ed r_2 .

Proposizione 8.2.5 Sia $(A, \langle -, - \rangle)$ uno spazio affine euclideo tridimensionale e siano $r_1, r_2 \subseteq A$ rette non parallele. Se Q_i è un punto e v_i è un vettore direttore di r_i , $i = 1, 2$, allora la distanza di r_1 da r_2 è data dalla formula

$$d_A(r_1, r_2) = \frac{|\langle \overrightarrow{Q_1 Q_2}, v_1 \wedge v_2 \rangle|}{\|v_1 \wedge v_2\|}.$$

Se inoltre r_1, r_2 hanno rappresentazioni parametriche

$$r_1 : \begin{cases} x = x_1 + t_1 l_1 \\ y = y_1 + t_1 m_1 \\ z = z_1 + t_1 n_1 \end{cases}, \quad t_1 \in \mathbb{R}, \quad r_2 : \begin{cases} x = x_2 + t_2 l_2 \\ y = y_2 + t_2 m_2 \\ z = z_2 + t_2 n_2 \end{cases}, \quad t_2 \in \mathbb{R},$$

allora si ha anche

$$d_A(r_1, r_2) = \frac{\left| \det \begin{pmatrix} x_2 - x_1 & l_1 & l_2 \\ y_2 - y_1 & m_1 & m_2 \\ z_2 - z_1 & n_1 & n_2 \end{pmatrix} \right|}{\sqrt{\begin{vmatrix} m_1 & m_2 \\ n_1 & n_2 \end{vmatrix}^2 + \begin{vmatrix} l_1 & l_2 \\ n_1 & n_2 \end{vmatrix}^2 + \begin{vmatrix} l_1 & l_2 \\ m_1 & m_2 \end{vmatrix}^2}}. \quad (8.13)$$

Dimostrazione. Consideriamo il vettore

$$u := \frac{1}{\|v_1 \wedge v_2\|} v_1 \wedge v_2.$$

Si tratta di un versore ortogonale sia ad r_1 che ad r_2 , perciò esiste $a \in \mathbb{R}$ tale che $\overrightarrow{N_1 N_2} = au$. Inoltre

$$d_A(r_1, r_2) = d_A(N_1, N_2) = \|\overrightarrow{N_1 N_2}\| = |a| \|u\| = |a|.$$

D'altro canto

$$\langle \overrightarrow{N_1 N_2}, u \rangle = \langle au, u \rangle = a \langle u, u \rangle = a,$$

perciò

$$\begin{aligned} d_A(r_1, r_2) &= |a| = \left| \langle \overrightarrow{N_1 N_2}, u \rangle \right| = \left| \langle \overrightarrow{N_1 Q_1} + \overrightarrow{Q_1 Q_2} + \overrightarrow{Q_2 N_2}, u \rangle \right| \\ &= \left| \langle \overrightarrow{N_1 Q_1}, u \rangle + \langle \overrightarrow{Q_1 Q_2}, u \rangle + \langle \overrightarrow{Q_2 N_2}, u \rangle \right| = \left| \langle \overrightarrow{Q_1 Q_2}, u \rangle \right| \\ &= \frac{\left| \langle \overrightarrow{Q_1 Q_2}, v_1 \wedge v_2 \rangle \right|}{\|v_1 \wedge v_2\|}, \end{aligned}$$

dove abbiamo usato che u è ortogonale sia ad r_1 (e quindi a $\overrightarrow{N_1 Q_1}$) che ad r_2 (e quindi a $\overrightarrow{Q_2 N_2}$). Se Q_i è il punto di coordinate (x_i, y_i, z_i) e v_i è il vettore di coordinate (l_i, m_i, n_i) , $i = 1, 2$, allora l'uguaglianza

$$d_A(r_1, r_2) = \frac{\left| \langle \overrightarrow{Q_1 Q_2}, v_1 \wedge v_2 \rangle \right|}{\|v_1 \wedge v_2\|}$$

appena ottenuta dà esattamente la (8.13) (il lettore lo vede?). ■

Per determinare la perpendicolare comune \tilde{r} a due rette non parallele $r_1, r_2 \subseteq A$ si può procedere in due modi: si determinano i punti N_1, N_2 della Proposizione 8.2.4 risolvendo il sistema (8.12). Se $N_1 \neq N_2$ allora la retta per N_1 ed N_2 è già \tilde{r} . Se invece $N_1 = N_2$, allora $N_1 = N_2$ è un punto di \tilde{r} mentre il vettore $v_1 \wedge v_2$ è un vettore direttore, dunque scriviamo subito una rappresentazione parametrica. Oppure possiamo trovare direttamente una rappresentazione cartesiana di \tilde{r} come segue. Osserviamo che la retta \tilde{r} è completamente determinata dalle seguenti tre condizioni

- (1) $v_1 \wedge v_2$ è un vettore direttore di \tilde{r} ;
- (2) \tilde{r} è complanare ad r_1 ;
- (3) \tilde{r} è complanare ad r_2 ;

Diciamo (x, y, z) le coordinate di un generico punto (incognito) di \tilde{r} e diciamo (L, M, N) le coordinate del vettore $v_1 \wedge v_2$. La condizione di complanarità tra la retta per (x, y, z) di vettore direttore $v_1 \wedge v_2$ e r_1 è

$$\begin{vmatrix} x - x_1 & l_1 & L \\ y - y_1 & m_1 & M \\ z - z_1 & n_1 & N \end{vmatrix} = 0 \quad (8.14)$$

e la condizione di complanarità tra la stessa retta e r_2 è

$$\begin{vmatrix} x-x_2 & l_2 & L \\ y-y_2 & m_2 & M \\ z-z_2 & n_2 & N \end{vmatrix} = 0. \quad (8.15)$$

Riassumendo, il punto di coordinate (x, y, z) appartiene ad \tilde{r} sse (x, y, z) è una soluzione delle due equazioni lineari (8.14) e (8.15). Esse sono pertanto equazioni in una rappresentazione cartesiana di \tilde{r} .

■ **Esempio 8.8** Nello spazio affine euclideo standard tridimensionale $(A_{\mathbb{R}}^3, \langle -, - \rangle_{can})$ consideriamo le rette

$$r_1 : \begin{cases} x = -t_1 \\ y = 2 + t_1 \\ z = 2 + t_1 \end{cases}, \quad t_1 \in \mathbb{R}, \quad r_2 : \begin{cases} x = 3 - \frac{3}{2}t_2 \\ y = -\frac{7}{2} + 2t_2 \\ z = -\frac{3}{2} + t_2 \end{cases}, \quad t_2 \in \mathbb{R}, \quad (8.16)$$

Il vettore $v_1 = (-1, 1, 1)^T$ è un vettore direttore di r_1 , mentre il vettore $v_2 = (-3/2, 2, 1)^T$ è un vettore direttore di r_2 . Poiché v_1 e v_2 sono linearmente indipendenti, le rette r_1, r_2 non sono parallele. In questo caso, la Formula (8.13) dà

$$d_A(r_1, r_2) = \sqrt{6}/2$$

(il lettore faccia il conto in dettaglio). Vogliamo anche calcolare la perpendicolare comune ad r_1, r_2 nonché i suoi punti di intersezione N_1, N_2 con r_1, r_2 rispettivamente. Per trovare una rappresentazione parametrica di \tilde{r} determiniamo prima i punti N_1, N_2 risolvendo il sistema (8.12). Posto $Q_1 = (0, 2, 2)^T$ e $Q_2 = (3, -7/2, -3/2)^T$, il sistema (8.12) diventa

$$\begin{cases} 3t_1 - \frac{9}{2}t_2 = -12 \\ -\frac{9}{2}t_1 + \frac{29}{4}t_2 = 19 \end{cases}$$

(il lettore lo verifichi nel dettaglio), la cui unica soluzione è $t_1 = -1, t_2 = 2$. Sostituendo nelle rappresentazioni parametriche (8.16) troviamo $N_1 = (1, 1, 1)^T$ e $N_2 = (0, 1/2, 1/2)^T$. La perpendicolare comune \tilde{r} è la retta per N_1, N_2 cioè:

$$\tilde{r} : \begin{cases} x = 1 - t \\ y = 1 - \frac{1}{2}t \\ z = 1 - \frac{1}{2}t \end{cases}, \quad t \in \mathbb{R}.$$

Per trovare una rappresentazione cartesiana di \tilde{r} possiamo eliminare il parametro t dalla rappresentazione parametrica, oppure ragionare come nella discussione che segue la dimostrazione della Proposizione 8.2.5: il vettore $v_1 \wedge v_2$ è

$$v_1 \wedge v_2 = \left(\begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix}, - \begin{vmatrix} -1 & -3/2 \\ 1 & 1 \end{vmatrix}, \begin{vmatrix} -1 & -3/2 \\ 1 & 2 \end{vmatrix} \right)^T = \begin{pmatrix} -1 \\ -1/2 \\ -1/2 \end{pmatrix}$$

e le Equazioni (8.14), (8.15) diventano

$$\begin{vmatrix} x & -2 & -1 \\ y-2 & 1 & -1/2 \\ z-2 & 1 & -1/2 \end{vmatrix} = 0$$

e

$$\begin{vmatrix} x-3 & -3/2 & -1 \\ y-7/2 & 2 & -1/2 \\ z-3/2 & 1 & -1/2 \end{vmatrix} = 0,$$

che, riscritte opportunamente, danno la seguente rappresentazione cartesiana

$$\tilde{r} : \begin{cases} -2y + 2z = 0 \\ -\frac{1}{2}x - \frac{7}{4}y + \frac{11}{4}z = -\frac{7}{2} \end{cases}.$$

Esercizio 8.9 Nello spazio affine euclideo standard tridimensionale $(A_{\mathbb{R}}^3, \langle -, - \rangle_{can})$ si considerino le rette

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

Dopo aver osservato che r_1, r_2 non sono parallele, si calcoli la loro distanza e si determini una rappresentazione parametrica e una cartesiana per la loro perpendicolare comune.

8.3 Isometrie di uno Spazio Affine Euclideo

Nell'ultima sezione di queste note, studiamo brevemente le *simmetrie* di uno spazio affine euclideo, cioè le trasformazioni che ne preservano tutte le proprietà. Sia $(A, \langle -, - \rangle)$ uno spazio affine euclideo di dimensione n modellato sullo spazio vettoriale euclideo $(V, \langle -, - \rangle)$.

Definizione 8.3.1 — Affinità Isometrica. Una *affinità isometrica*, o, più semplicemente, una *isometria*, di $(A, \langle -, - \rangle)$ è un'affinità $\Phi : A \rightarrow A$ tale che l'applicazione lineare associata $f_{\Phi} : V \rightarrow V$ è un automorfismo ortogonale. Una isometria Φ si dice *diretta* se la sua parte lineare f_{Φ} conserva le orientazioni, altrimenti si dice *inversa*.

■ **Esempio 8.9** Ogni traslazione τ è una isometria diretta. Infatti la parte lineare di τ è semplicemente $\text{id}_V : V \rightarrow V$ che è un automorfismo ortogonale che preserva le orientazioni.

■ **Esempio 8.10** Consideriamo lo spazio affine euclideo standard $(A_{\mathbb{R}}^n, \langle -, - \rangle_{can})$ e sia $\Phi : A_{\mathbb{R}}^n \rightarrow A_{\mathbb{R}}^n$ un'affinità. Già sappiamo che $\Phi = F_{(A,b)}$ per qualche matrice invertibile $A \in GL_n(\mathbb{R})$ e qualche vettore numerico $b \in \mathbb{R}^n$. È chiaro che Φ è un'isometria sse A è una matrice ortogonale. Infatti la parte lineare di Φ è l'applicazione lineare $L_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ associata alla matrice A , la quale è un automorfismo ortogonale sse $A \in O_n$. Inoltre Φ è un'isometria diretta sse A è una matrice ortogonale speciale (il lettore lo vede?).

Proposizione / Definizione 8.3.1 — Gruppo delle Isometrie. Le isometrie di $(A, \langle -, - \rangle)$ formano un sottogruppo $\text{Isom}(A, \langle -, - \rangle) \subseteq \text{Aff}(A)$ nel gruppo delle affinità, detto *gruppo delle isometrie*.

Dimostrazione. L'applicazione identica $\text{id}_A : A \rightarrow A$ è un'isometria. Infatti è un'affinità e la sua parte lineare è $\text{id}_V : V \rightarrow V$ che è un automorfismo ortogonale. Ora siano $\Phi, \Psi : A \rightarrow A$ isometrie. La loro composta $\Psi \circ \Phi : A \rightarrow A$ è un'isometria, infatti (è un'affinità e) la sua parte lineare è $f_{\Psi} \circ f_{\Phi}$. Ma la composta di automorfismi ortogonali è un automorfismo ortogonale, perciò $f_{\Psi} \circ f_{\Phi}$ è un automorfismo ortogonale come annunciato. Infine l'inversa $\Phi^{-1} : A \rightarrow A$ di Φ è un'isometria, infatti (è un'affinità e) la sua parte lineare è f_{Φ}^{-1} . Ma l'inversa di un automorfismo ortogonale è un automorfismo ortogonale. Questo conclude la dimostrazione. ■

Esercizio 8.10 Dimostrare che le isometrie dirette formano un sottogruppo nel gruppo delle isometrie. ■

La seguente proposizione giustifica la terminologia “isometria”.

Proposizione 8.3.2 Le isometrie *conservano le distanze*, cioè se $\Phi : A \rightarrow A$ è una isometria di $(A, \langle -, - \rangle)$ allora, per ogni $P, Q \in A$, si ha che

$$d_A(\Phi(P), \Phi(Q)) = d_A(P, Q).$$

Dimostrazione. Basta calcolare

$$d_A(\Phi(P), \Phi(Q)) = \|\overrightarrow{\Phi(Q)\Phi(P)}\| = \|f_\Phi(\overrightarrow{PQ})\| = \|\overrightarrow{PQ}\| = d_A(P, Q),$$

dove abbiamo usato che f_Φ è un automorfismo ortogonale, e quindi conserva la norma. ■

Esercizio 8.11 Siano (O, \mathcal{R}) e (O', \mathcal{R}') riferimenti cartesiani di $(A, \langle -, - \rangle)$. Dimostrare che esiste un'unica isometria $\Phi : A \rightarrow A$ tale che $\Phi(O) = O'$ e $f_\Phi(\mathcal{R}) = \mathcal{R}'$. ■

○ Ricordiamo dall'Osservazione a pag. 213 che, dato uno spazio affine euclideo $(A, \langle -, - \rangle)$, l'insieme A munito della distanza d_A è uno spazio metrico. Ogni affinità isometrica di $(A, \langle -, - \rangle)$ è anche un'isometria dello spazio metrico (A, d_A) . Vale anche il viceversa, cioè ogni biezione $\Phi : A \rightarrow A$ che conserva la distanza d_A è necessariamente un'affinità isometrica in base alla seguente

Proposizione 8.3.3 Sia $\Phi : A \rightarrow A$ una funzione (non necessariamente un'applicazione affine). Allora le seguenti asserzioni sono equivalenti:

- (1) Φ è un'affinità isometrica;
- (2) Φ è un'isometria dello spazio metrico (A, d_A) , cioè è una biezione tale che

$$d_A(\Phi(P), \Phi(Q)) = d_A(P, Q),$$

per ogni $P, Q \in A$.

Dimostrazione. Abbiamo già osservato che (1) \Rightarrow (2). L'implicazione (2) \Rightarrow (1) è un corollario della Proposizione 2.4.7. Infatti sia Φ come in (2). Dimostriamo innanzitutto che Φ è un'affinità. Poiché Φ è biettiva per ipotesi, basta dimostrare che Φ è un'applicazione affine. A questo scopo utilizziamo la caratterizzazione delle applicazioni affini data dalla Proposizione 6.4.4. Fissiamo un punto $\bar{P} \in A$ e consideriamo la funzione

$$f := \alpha_{\Phi(\bar{P})} \circ \Phi \circ \alpha_{\bar{P}}^{-1} : V \rightarrow V.$$

(per maggiore chiarezza osserviamo che, a priori, f dipende dalla scelta del punto \bar{P}). Detta

$d : V \times V \rightarrow \mathbb{R}$ la distanza in V , per ogni $v, w \in V$ abbiamo

$$\begin{aligned}
 d(f(v), f(w)) &= \|f(w) - f(v)\| \\
 &= \|\alpha_{\Phi(\bar{P})}(\Phi(\alpha_{\bar{P}}^{-1}(w))) - \alpha_{\Phi(\bar{P})}(\Phi(\alpha_{\bar{P}}^{-1}(v)))\| \\
 &= \|\overrightarrow{\Phi(\bar{P})\Phi(\alpha_{\bar{P}}^{-1}(w))} - \overrightarrow{\Phi(\bar{P})\Phi(\alpha_{\bar{P}}^{-1}(v))}\| \\
 &= \|\overrightarrow{\Phi(\alpha_{\bar{P}}^{-1}(v))\Phi(\alpha_{\bar{P}}^{-1}(w))}\| \\
 &= d_A(\Phi(\alpha_{\bar{P}}^{-1}(v)), \Phi(\alpha_{\bar{P}}^{-1}(w))) \\
 &= d_A(\alpha_{\bar{P}}^{-1}(v), \alpha_{\bar{P}}^{-1}(w)) \\
 &= \|\overrightarrow{\alpha_{\bar{P}}^{-1}(v)\alpha_{\bar{P}}^{-1}(w)}\| \\
 &= \|w - v\| \\
 &= d(v, w).
 \end{aligned}$$

Cioè f preserva la distanza in V . Inoltre

$$f(0) = \alpha_{\Phi(\bar{P})}(\Phi(\alpha_{\bar{P}}^{-1}(0))) = \alpha_{\Phi(\bar{P})}(\Phi(\bar{P})) = 0.$$

Perciò, dalla Proposizione 2.4.7, f è un automorfismo ortogonale di $(V, \langle -, - \rangle)$. Siccome il diagramma

$$\begin{array}{ccc}
 A & \xrightarrow{\Phi} & A \\
 \alpha_{\bar{P}} \downarrow & & \downarrow \alpha_{\Phi(\bar{P})} \\
 V & \xrightarrow{f} & V
 \end{array}$$

commuta per costruzione (il lettore lo vede?), segue dalla Proposizione 6.4.4 che Φ è un'applicazione affine e f è la sua parte lineare (in particolare f , di fatto, non dipende dalla scelta di \bar{P}). Dunque Φ è un'affinità isometrica come annunciato. ■

- Sia $(A, \langle -, - \rangle)$ uno spazio affine euclideo di giacitura $(V, \langle -, - \rangle)$. Due sottoinsiemi $\mathcal{F}, \mathcal{F}' \subseteq A$ si dicono *congruenti* se esiste un'isometria $\Phi : A \rightarrow A$ tale che $\Phi(\mathcal{F}) = \mathcal{F}'$. Per esempio due sottospazi affini $S, S' \subseteq A$ sono congruenti sse hanno la stessa dimensione. Infatti, sia $\Phi : A \rightarrow A$ un'isometria tale che $\Phi(S) = S'$. Poiché le affinità conservano la dimensione dei sottospazi affini allora $\dim S = \dim S'$. Viceversa, se $\dim S = \dim S'$ possiamo costruire un'isometria $\Phi : A \rightarrow A$ tale che $\Phi(S) = S'$ come segue. Fissiamo punti $O \in S$ e $O' \in S'$. Fissiamo anche due riferimenti ortonormali \mathcal{R}_W e $\mathcal{R}_{W'}$, il primo nella giacitura $W \subseteq V$ di S e il secondo nella giacitura $W' \subseteq V$ di S' . I riferimenti $\mathcal{R}_W, \mathcal{R}_{W'}$ possono essere completati a riferimenti ortonormali $\mathcal{R}, \mathcal{R}'$ di V (il lettore discuta questo passaggio in tutti i dettagli). Per il Teorema dell'Estensione Lineare, esiste un unico endomorfismo $f : V \rightarrow V$ tale che $f(\mathcal{R}) = \mathcal{R}'$. Poiché f trasforma un riferimento ortonormale in un riferimento ortonormale, f è necessariamente un automorfismo ortogonale. Si osservi inoltre che $f(W) = W'$ (il lettore lo vede?). Per il Teorema dell'Estensione Affine, esiste un'unica applicazione affine $\Phi : A \rightarrow A$ tale che $\Phi(O) = O'$ e $f_{\Phi} = f$. Poiché f è un automorfismo ortogonale, Φ è necessariamente un'isometria. Inoltre è facile vedere che $\Phi(S) = S'$. Il lettore lo verifichi nei dettagli.

Facciamo un ultimo esempio di sottoinsiemi congruenti. Sia $n = \dim A$, sia $C \in A$ e sia $r \in \mathbb{R}$ un numero reale positivo. Il disco n -dimensionale di centro C e raggio r è il sottoinsieme

$$D_r(C) := \{P \in A : d_A(P, C) \leq r\}.$$

Se $m \leq n$, un disco m -dimensionale di centro C e raggio r è, per definizione, l'intersezione tra $D_r(C)$ e un sottospazio affine di dimensione m passante per C .

Esercizio 8.12 Si dimostri che due dischi sono congruenti sse hanno la stessa dimensione e lo stesso raggio (Suggerimento: si discuta prima il caso di due dischi di dimensione n).

Concludiamo questa breve sezione (e queste note) discutendo in maggiore dettaglio le isometrie di uno spazio affine euclideo bidimensionale. Sia dunque $(A, \langle -, - \rangle)$ uno spazio affine euclideo e sia $\dim A = 2$. In questo contesto, oltre alle traslazioni, distinguiamo altre 3 classi speciali di isometrie, in base alla seguente

Definizione 8.3.2 — Rotazioni, Riflessioni e Glissoriflessioni.

- (1) Una *rotazione* di centro $C \in A$ è una isometria diretta diversa dall'identità che fissa il punto C ;
- (2) una *riflessione* è una isometria inversa che fissa tutti i punti di una retta $r \subseteq A$ detta *asse* della riflessione;
- (3) una *glissoriflessione* è la composizione di una riflessione ρ seguita da una traslazione di vettore v non nullo, parallelo all'asse di ρ (cioè v è un vettore direttore dell'asse di ρ); l'asse di ρ è detto *asse* della glissoriflessione.

■ **Esempio 8.11** Consideriamo lo spazio affine euclideo standard bidimensionale

$$(A_{\mathbb{R}}^2, \langle -, - \rangle_{can}).$$

Sia $R = R_{\theta} \in O_2$ una matrice di rotazione. L'applicazione lineare associata $f_R : A_{\mathbb{R}}^2 = \mathbb{R}^2 \rightarrow A_{\mathbb{R}}^2 = \mathbb{R}^2$ è una rotazione di centro l'origine $O_{can} = (0, 0)^T \in A_{\mathbb{R}}^2 = \mathbb{R}^2$. Similmente l'applicazione lineare associata alla matrice ortogonale

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

è una riflessione di asse r pari all'asse delle y , cioè la retta $r : x = 0$ (il lettore lo vede?). ■

Esercizio 8.13 Nello spazio affine euclideo standard bidimensionale

- (1) si costruisca una rotazione di centro $(1, 1)^T \in A_{\mathbb{R}}^2 = \mathbb{R}^2$;
- (2) si costruisca una riflessione di asse $r : x - y = 0$.

La terminologia nella Definizione 8.3.2 sarà chiarita dalla dimostrazione del seguente

Teorema 8.3.4 — Teorema di Chasles. Ogni isometria di uno spazio affine euclideo bidimensionale è di uno dei quattro tipi: una traslazione (eventualmente di vettore nullo), una rotazione, una riflessione o una glissoriflessione.

Dimostrazione. Sia $(A, \langle -, - \rangle)$ uno spazio affine euclideo bidimensionale di giacitura $(V, \langle -, - \rangle)$, e sia $\Phi : A \rightarrow A$ un'isometria. Cominciamo con qualche osservazione preliminare. Sia (A, b) la rappresentazione coordinata di Φ in un riferimento cartesiano (O, \mathcal{R}) e siano $x = (x_1, x_2)^T$ le coordinate cartesiane nel riferimento (O, \mathcal{R}) . Poiché A è la matrice rappresentativa di f_{Φ} nel riferimento ortonormale \mathcal{R} di V , allora si tratta di una matrice ortogonale: $A \in O_2$. Ora distinguiamo 2 casi ciascuno suddiviso in due sottocasi:

- (1) Φ è una *isometria diretta*. In questo caso f_{Φ} fissa le orientazioni e $A \in SO_2$, cioè A è una matrice di rotazione: $A = R_{\theta}$ per qualche $\theta \in [0, 2\pi)$, e dunque

$$F_{(A,b)}(x) = R_{\theta}x + b.$$

Consideriamo i due sottocasi:

- (1).(i) Φ ha un punto fisso C . In questo caso, possiamo scegliere $O = C$. Poiché $\Phi(O) = \Phi(C) = C = O$ si deve anche avere $b = F_{(A,b)}0 = 0$ da cui

$$F_{(A,b)}x = R_\theta x. \quad (8.17)$$

Se $\theta = 0$, allora $R_\theta = I_2$ e $\Phi = \text{id}_A$, altrimenti Φ è una rotazione di centro C (e la rappresentazione coordinata (8.17) giustifica la terminologia, Figura 8.13).

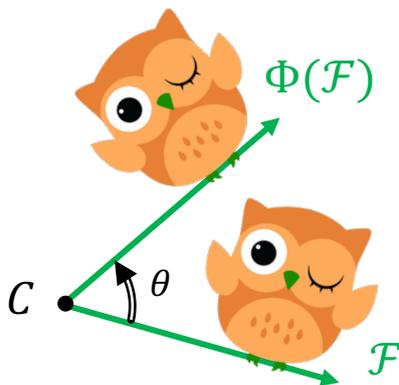


Figura 8.13: Una rotazione di centro C nel piano affine euclideo.

- (1).(ii) Φ non ha punti fissi. In questo caso il sistema lineare

$$F_{(A,b)}(x) = R_\theta x + b = x$$

non ha soluzioni. Tale sistema si riscrive

$$(R_\theta - I_2)x = -b. \quad (8.18)$$

Ora un semplice calcolo mostra che

$$\det(R_\theta - I_2) = 2(1 - \cos \theta),$$

dunque $\text{rk}(R_\theta - I_2) = 2$ per ogni $\theta \neq 0$. Perciò l'unica possibilità per la (8.18) di essere incompatibile è che sia $\theta = 0$, cioè $R_\theta = I_2$, e $b \neq 0$, da cui

$$F_{(A,b)}(x) = x + b,$$

cioè Φ è la traslazione di vettore v , dove v è il vettore di coordinate b nel riferimento \mathcal{R} (Figura 8.14).

- (2) Φ è una isometria inversa. In questo caso, in virtù del Teorema 5.4.2, esiste un riferimento ortonormale $\overline{\mathcal{R}}$ nel quale la matrice rappresentativa di f_Φ è la matrice

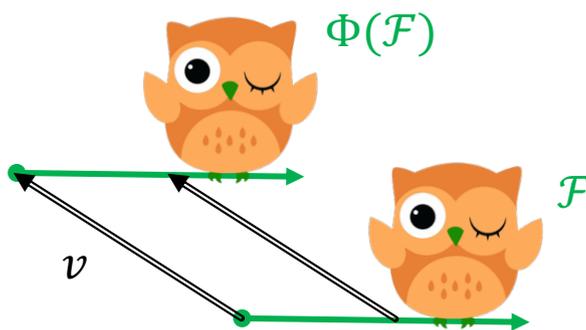
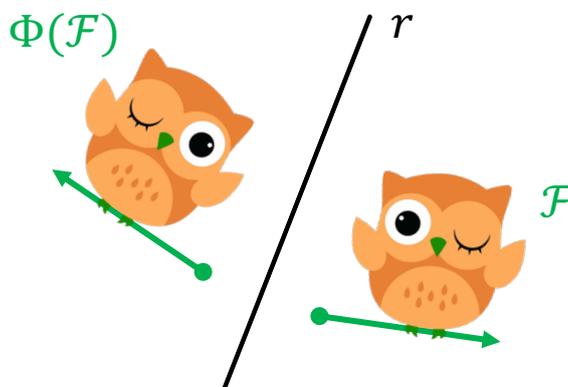
$$U = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Se scegliamo $\mathcal{R} = \overline{\mathcal{R}}$, allora $A = U$. Consideriamo i due sottocasi:

- (2).(i) Φ ha un punto fisso C . In questo caso possiamo scegliere $O = C$ e, come sopra, otteniamo $b = 0$, cioè

$$F_{(A,b)}x = Ux. \quad (8.19)$$

Allora Φ fissa la retta r di rappresentazione cartesiana $r : x_1 = 0$. Dunque Φ è una riflessione (Figura 8.15).

Figura 8.14: Una traslazione di vettore v nel piano affine euclideo.Figura 8.15: Una riflessione di asse r nel piano affine euclideo.

(2).(ii) Φ non ha punti fissi. In questo caso il sistema

$$Ux + b = x$$

non ha soluzioni. Se poniamo $b = (b_1, b_2)^T$, tale sistema è semplicemente

$$\begin{cases} -2x_1 = b_1 \\ 0 = b_2 \end{cases}$$

che è incompatibile esattamente quando $b_2 \neq 0$. La formula $F_{(A,b)}x = Ux + b$ mostra che Φ è la composta di una riflessione ρ' di asse $r : x_1 = 0$ (e rappresentazione coordinata $(U, 0)$) con la traslazione di vettore v' dove v' ha coordinate (b_1, b_2) , con $b_2 \neq 0$. Si osservi che la giacitura di r è l'autospazio di $f_{\rho'}$ relativo all'autovalore 1, mentre il suo complemento ortogonale è l'autospazio di $f_{\rho'}$ relativo all'autovalore -1 . Ora possiamo scrivere v' in un unico modo come somma di un vettore v parallelo ad r (e dunque tale che $\rho'(v) = v$) e un vettore w ortogonale ad r (e dunque tale che $\rho'(w) = -w$) e la condizione $b_2 \neq 0$ dice che $v \neq 0$. Riassumendo $\Phi = \tau_{v+w} \circ \rho' = \tau_v \circ \tau_w \circ \rho'$. Ora basta mostrare che $\rho := \tau_w \circ \rho'$ è una riflessione di asse parallelo a v . È così, infatti innanzitutto ρ è un'isometria inversa (il lettore lo vede?). Inoltre fissa i punti della retta $r := \tau_{w/2}(r')$, cioè, per ogni punto $P \in r$ si ha $\rho(P) = P$. Per dimostrare questo

consideriamo $P \in r$. Allora esiste $P' \in r'$ tale che $P = P' + w/2$ e

$$\begin{aligned}\rho(P) &= \tau_w(\rho'(P' + w/2)) = \tau_w(\rho'(P') + f_{\rho'}(w/2)) \\ &= \tau_w(P' - w/2) = P' - w/2 + w = P' + w/2 = P.\end{aligned}$$

Infine, poichè una traslazione preserva la giacitura di ogni sottospazio (infatti la sua parte lineare è la funzione identica) v è anche un vettore direttore di r . Questo conclude la dimostrazione (vedi anche la Figura 8.16 per un esempio di glissoriflessione).

■

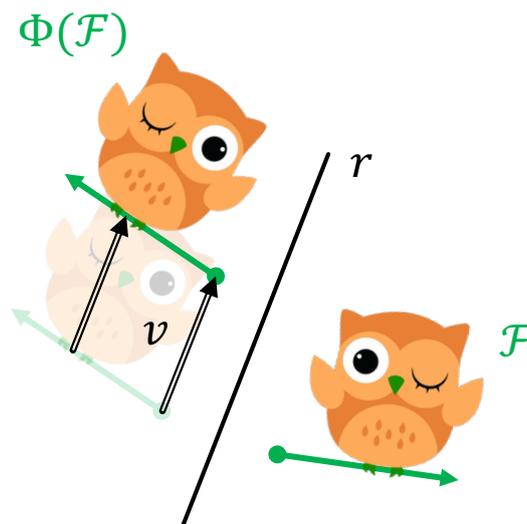


Figura 8.16: Una glissoriflessione di asse r nel piano affine euclideo.

La dimostrazione del Teorema di Chasles mostra, tra l'altro, che una riflessione è completamente determinata dal suo asse r .

Infine osserviamo che le quattro classi di isometrie di uno spazio affine euclideo bidimensionale, e cioè traslazioni, rotazioni, riflessioni e glissoriflessioni sono classi disgiunte, cioè una isometria può appartenere solo ad una di queste classi. E precisamente un'isometria è

- (1) una rotazione sse è un'isometria diretta che fissa un punto diversa dall'identità (per definizione);
- (2) una traslazione sse è l'identità oppure è un'isometria diretta senza punti fissi;
- (3) una riflessione sse è un'isometria inversa con un punto fisso;
- (4) una glissoriflessione sse è un'isometria inversa senza punti fissi.

Esercizio 8.14 Si consideri lo spazio affine euclideo bidimensionale standard $(A_{\mathbb{R}}^2, \langle -, - \rangle_{can})$ e la funzione:

$$\Phi : A_{\mathbb{R}}^2 \rightarrow A_{\mathbb{R}}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} y+1 \\ x \end{pmatrix}.$$

Dopo aver verificato che Φ è un'affinità, si dimostri che si tratta, in effetti, di una glissoriflessione. Si decomponga inoltre Φ come la composizione di una riflessione rispetto ad un asse r seguita da una traslazione di vettore non nullo nella giacitura di r . ■

Indice analitico

A

Affinità	190
isometrica	232
Angolo	
tra due piani	224
tra due rette	213
nel piano, 220	
nello spazio 3D, 224	
tra due vettori	41
tra un piano e una retta	224
Applicazione	
affine	182
applicazione lineare associata, 182	
e sottospazi affini, 188	
parte lineare, 182	
tra spazi affini numerici, 183	
aggiunta	64
antilineare	101
bilineare	9
coniugata	64
multilineare	9
ortogonale	55
Automorfismo	
che conserva le orientazioni	122
che inverte le orientazioni	122
ortogonale	55
che conserva le orientazioni, 123	
Autospazio	73

di una matrice	84
Autovalore	68
di un endomorfismo hermitiano	114
di un endomorfismo ortogonale	128
di un endomorfismo simmetrico	88
di un endomorfismo unitario	110
di una matrice	84
molteplicità algebrica	71
molteplicità geometrica	74
Autovettore	68
di una matrice	84

C

Centroaffinità	195
Complemento ortogonale	43, 105
proprietà	43
Congruenza	234
Congruenza di matrici	18
Conica	199
a centro	208
degenere	208
non degenere	208
Coordinate	
affini	145, 183, 190
cambio di, 191	
cartesiane	46, 217
Criterio	
di complanarità	175
di diagonalizzabilità	78

D

- Disco 234
 dischi congruenti 234
 Distanza 41
 di un punto da un piano 226
 di un punto da un sottospazio affine .. 216
 di un punto da una retta
 nel piano, 221
 nello spazio 3D, 227
 in uno spazio affine euclideo 212
 in uno spazio vettoriale euclideo 39
 in uno spazio vettoriale hermitiano .. 104
 nello spazio affine euclideo standard . 212
 tra due piani paralleli 226
 tra due rette non parallele 229
 tra due rette parallele 222
 tra sottospazi affini paralleli 217
 Disuguaglianza
 di Cauchy-Schwarz 39, 104
 triangolare
 in uno spazio affine euclideo, 212
 in uno spazio hermitiano, 105
 in uno spazio metrico, 41
 in uno spazio vettoriale euclideo, 40

E

- Endomorfismo
 aggiunto 113
 autoaggiunto 113
 diagonalizzabile 67
 hermitiano 113
 hermitiano coniugato 113
 ortogonale
 classificazione, 133
 classificazione in 2D, 128
 classificazione in 3D, 131
 ortogonalmente diagonalizzabile . 87, 110
 caratterizzazione, 117
 simmetrico 66
 autovalori, 88
 caratterizzazione, 66
 unitario 107
 è ortogonalmente diagonalizzabile, 110
 Equazione
 caratteristica 71
 del fascio proprio
 di piani, 178
 di rette, 160

Equipollenza 138

F

- Fascio
 improprio
 di piani, 180
 di rette nel piano, 162
 proprio
 di piani, 178
 di rette nel piano, 159
 Forma
 bilineare 9
 alternante, 32
 antisimmetrica, 32
 non-degenere, 16
 simmetrica, 17
 hermitiana 101
 multilineare 9
 quadratica associata ad una forma bilineare
 simmetrica 19
 simplettica 35
 Forma normale
 di un polinomio di secondo grado
 caso reale, 206
 di una conica
 caso reale, 207
 di una forma bilineare
 alternante, 32
 simmetrica reale, 28
 di una quadrica 206
 caso reale, 207
 Formula
 di polarizzazione 19
 Funzione
 differenza 136
 polinomiale 70

G

- Giacitura
 di un sottospazio affine 140
 di uno spazio affine 136
 euclideo, 211
 Glissoriflessione 235
 asse di una 235
 Grado
 di un polinomio 69
 Gruppo
 affine 194

di uno spazio affine, 192
 delle isometrie 232
 delle traslazioni 192
 ortogonale 53
 di uno spazio vettoriale euclideo, 58
 speciale, 53
 unitario 108
 di uno spazio hermitiano, 107

I

Insieme algebrico 199
 Isometria 232
 diretta 232
 inversa 232
 tra spazi metrici 62
 Isomorfismo
 musicale 63
 ortogonale
 proprietà, 58

L

Legge di Inerzia di Sylvester 28
 Lunghezza
 di un vettore 39

M

Matrice
 associata ad un polinomio di secondo grado 200
 complessa 100
 coniugio, 100
 hermitiana coniugata, 101
 parte immaginaria, 100
 parte reale, 100
 dei vettori direttori 150
 del cambio di coordinate 13
 tra riferimenti ortonormali, 52
 di rotazione 53
 diagonalizzabile 84
 diagonalizzante 85
 hermitiana 103
 ortogonale 52
 caratterizzazioni, 52
 speciale, 53
 ortogonalmente diagonalizzabile 92
 rappresentativa
 di un'applicazione lineare, 58

di una forma bilineare, 12
 di una forma bilineare simmetrica, 17
 di una forma hermitiana, 102
 unitaria 108
 è ortogonalmente diagonalizzabile, 112

Matrici

congruenti 18

N

Norma 39, 104
 quadra 38, 104
 Numeri complessi 97
 coniugio 99
 inverso 98
 modulo 99
 parte immaginaria 98
 parte reale 98
 prodotto 97
 puramente immaginari 98

O

Orientazione 120
 e trasformazioni elementari 121
 standard 121
 Ortogonalità
 tra sottospazi affini 214
 tra un vettore e un sottoinsieme .. 42, 105
 tra un vettore e un sottospazio affine . 214
 tra vettori 42, 105
 rispetto ad una forma bilineare simmetrica, 20
 Ortogonalizzazione di Gram-Schmidt. 48, 106
 Ortonormalizzazione 51

P

Perpendicolare comune
 a due rette non parallele 229
 Piano 141
 affine 136
 della Geometria Euclidea 140
 euclideo 219
 Polinomi
 equivalenti 199
 Polinomio 69
 caratteristico 71
 di una matrice, 84
 complesso 99

- di secondo grado
 forma normale, 206
 radice 70
 molteplicità, 70
 Posizione reciproca
 di due piani nello spazio 3D 167
 di due rette nel piano 157
 di sottospazi affini 142
 piano/retta nello spazio 3D 170
 Prodotto
 di matrici 9
 hermitiano 104
 standard, 102
 scalare 37
 standard, 38
 vettoriale 119, 123
 Proiezione ortogonale
 su un sottospazio affine 214
 su un sottospazio vettoriale 45, 105
 Proprietà triangolare 135
 Punto 141
 base 149
 fisso 195
 punti indipendenti 144
- Q**
- Quadrica 199
 forma normale 207
 caso reale, 207
- R**
- Rango
 di una forma bilineare 16
 Rappresentazione
 cartesiana
 di un piano nello spazio 3D, 164
 di un piano per 3 punti, 150, 164
 di un sottospazio affine, 146
 di una retta nel piano, 155
 di una retta nello spazio 3D, 165
 di una retta per due punti, 156
 coordinata di un'applicazione affine . 186
 parametrica
 di un piano nello spazio 3D, 163
 di un sottospazio affine, 146
 di una retta nel piano, 155
 di una retta nello spazio 3D, 165
 di una retta per 2 punti, 150
 Regola del parallelogramma 138
 Retta 141
 affine 136
 nel piano 155
 rette complanari 173, 175
 Riferimento
 affine 144
 canonico, 144, 217
 di punti, 144
 cartesiano 217
 concordemente orientato 119
 di autovettori 69, 79
 ortonormale, 87
 diagonalizzante 67
 discordemente orientato 119
 ortogonale
 esistenza, 20
 rispetto ad una forma bilineare simmetrica, 20
 ortonormale 27, 45, 106
 coordinate in un, 46
 di autovettori, 87, 89
 esistenza, 27
 Riflessione 235
 asse di una 235
 Rotazione 235
 centro di una 235
- S**
- Segmento 137
 orientato 137
 Segnatura
 di una forma bilineare simmetrica reale 28
 Sistema ortogonale 42
 Somma diretta 76
 Sottospazio
 affine 140
 banale, 141
 delle soluzioni di un sistema lineare, 141
 generato da un sottoinsieme, 143
 posizione reciproca, 142, 153
 sottospazi affini congruenti, 234
 sottospazi incidenti, 142
 sottospazi paralleli, 142
 sottospazi sghembi, 142
 annullatore di una forma bilineare 14
 somma 75
 vettoriale euclideo 38
 Spazio

affine	135
euclideo, 211	
numerico, 136	
spazi affini isomorfi, 190	
della Geometria Euclidea	137
hermitiano	104
numerico, 104	
standard, 104	
metrico	41
vettoriale	
dei vettori applicati, 137	
vettoriale complesso	101
vettoriale euclideo	38
numerico, 38	
standard, 38	
vettoriale modello	136
vettoriale orientato	120
Spazio vettoriale	
delle applicazioni multilineari	11

T

Teorema	
dell'Estensione Affine	189
di Chasles	235
di classificazione dei polinomi di secondo	
grado	203
caso reale, 206	
di Pitagora	213
di Sylvester	28
Fondamentale dell'Algebra	99
Spettrale	87
Traslazione	182

U

Unità immaginaria	98
-----------------------------	----

V

Versore	39
Vettore	
applicato	137
congiungente	136
di separazione	136
direttore	149
geometrico	138
isotropo	20
libero	138

