

Geometria I

Contents

I	Geometria Affine	2
1	Vettori Geometrici	3
2	Algebra Lineare	5
2.1	Spazi Vettoriali	5
2.2	Matrici	9
2.3	Sistemi di Equazioni Lineari	15
2.4	Sottospazi Vettoriali	23
2.5	Rango	34
2.6	Determinanti	38
3	Spazi Affini	47
3.1	Spazi e Sottospazi Affini	47
3.2	Geometria in un Piano Affine	59
3.3	Geometria in uno Spazio Affine	65
4	Applicazioni Lineari	70
4.1	Applicazioni lineari	70
4.2	Applicazioni Lineari e matrici, cambiamenti di coordinate affini	83
4.3	Operatori Lineari	94
II	Geometria Euclidea	103
5	Forme bilineari e quadratiche	104
6	Diagonalizzazione delle forme quadratiche	116
7	Prodotti scalari	120
8	Prodotto vettoriale	131
9	Spazi euclidei	135
10	Diagonalizzazione di operatori simmetrici	145

Part I

Geometria Affine

1 Vettori Geometrici

Vettore applicato

Definizione 1.1. Un **vettore applicato** (o **segmento orientato**) dello spazio ordinario è individuato da un **punto iniziale** A e da un **punto finale** B e viene denotato con il simbolo (A, B) . Due vettori applicati (A, B) e (C, D) si dicono **equipollenti** se hanno la stessa **direzione**, la stessa **lunghezza**, e lo stesso **verso**, cioè se giacciono su rette parallele (eventualmente coincidenti) e se, muovendo una delle due rette parallelamente a se stessa, è possibile portare i due segmenti a sovrapporsi in modo che i loro punti iniziali e i loro punti finali coincidano.

Nota 1. Nell'insieme dei vettori applicati l'equipollenza è una relazione di equivalenza, perchè riflessiva, simmetrica e transitiva.

Vettore geometrico

Definizione 1.2. Un **vettore geometrico** (o **vettore**) è una classe di equipollenza di vettori applicati, cioè è l'insieme di tutti i segmenti orientati equipollenti a un segmento orientato assegnato. Ogni vettore applicato che individua il vettore \mathbf{a} si dirà **rappresentante di \mathbf{a}** . Il vettore avente rappresentante (A, B) verrà anche denotato con \overrightarrow{AB} . Dato un punto A , ogni vettore \mathbf{a} ha un rappresentante, e uno solo, applicato in A .

Vettore nullo

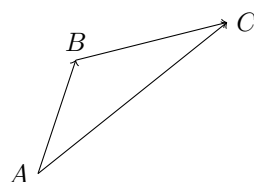
Definizione 1.3. Se $A = B$, il vettore individuato da uno qualunque dei segmenti orientati (A, A) si chiama **vettore nullo**: esso ha lunghezza nulla, direzione e verso indeterminati e si denota con $\mathbf{0}$.

Somma di due vettori Si può definire la **somma di due vettori** mediante loro rappresentanti nel modo seguente. Siano

$$\mathbf{a} = \overrightarrow{AB} \wedge \mathbf{b} = \overrightarrow{BC};$$

allora

$$\mathbf{a} + \mathbf{b} = \overrightarrow{AC}.$$



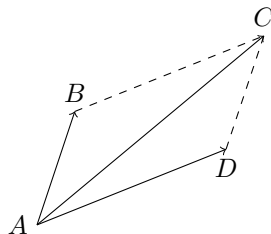
Se invece i vettori \mathbf{a} e \mathbf{b} sono dati mediante loro rappresentati applicati nello stesso punto, cioè

$$\mathbf{a} = \overrightarrow{AB} \wedge \mathbf{b} = \overrightarrow{AD}$$

rispettivamente, allora

$$\mathbf{a} + \mathbf{b} = \overrightarrow{AC},$$

dove il punto C è il quarto vertice del parallelogramma i cui altri vertici sono A, B, D . Questo modo di costruire $\mathbf{a} + \mathbf{b}$ è chiamato **regola del parallelogramma**.



Associatività L'operazione di somma di due vettori è **associativa**, cioè

$$\forall \mathbf{a}, \mathbf{b}, \mathbf{c}, \quad \mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}.$$

Ciò implica che sommando tre vettori si possono omettere le parentesi perchè la scrittura $\mathbf{a} + \mathbf{b} + \mathbf{c}$ ha un solo significato, e questo ragionamento si estende a un numero finito qualunque di vettori.

Commutatività L'operazione di somma è **commutativa**, cioè

$$\forall \mathbf{a}, \mathbf{b}, \quad \mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}.$$

Il vettore $\mathbf{0}$ soddisfa

$$\forall \mathbf{a}, \quad \mathbf{a} + \mathbf{0} = \mathbf{0} + \mathbf{a}.$$

Se $\mathbf{a} = \overrightarrow{AB}$ denoteremo con $-\mathbf{a}$ il vettore \overrightarrow{BA} . Esso verifica l'identità

$$\mathbf{a} + (-\mathbf{a}) = \mathbf{0}.$$

Prodotto di un vettore \mathbf{a} per un numero reale k Esso è, per definizione, il vettore $k\mathbf{a}$ che ha la stessa direzione di \mathbf{a} , lunghezza uguale a quella di \mathbf{a} moltiplicata per $|k|$, e verso discorde o concorde con \mathbf{a} a seconda che k abbia segno positivo o negativo; se $k = 0$ oppure $\mathbf{a} = \mathbf{0}$. L'operazione di moltiplicazione di un vettore per uno scalare è compatibile con quella di somma di vettori e con le operazioni di somma e di prodotto tra scalari. Ad esempio la seguente identità è di immediata verifica:

$$\forall \mathbf{a}, \forall n \in \mathbb{N}, \quad n\mathbf{a} = \underbrace{\mathbf{a} + \dots + \mathbf{a}}_{n \text{ volte}}.$$

In particolare

$$1\mathbf{a} = \mathbf{a} \quad \text{e} \quad (-1)\mathbf{a} = -\mathbf{a}.$$

Distributività Si può verificare facilmente che

$$\forall \mathbf{a}, \forall k, h, \quad (k + h)\mathbf{a} = k\mathbf{a} + h\mathbf{a} \quad \wedge \quad (kh)\mathbf{a} = k(h\mathbf{a})$$

Inoltre, abbiamo anche

$$\forall \mathbf{a}, \mathbf{b}, \quad \forall k, \quad k(\mathbf{a} + \mathbf{b}) = k\mathbf{a} + k\mathbf{b}.$$

Questa parte può essere allargata con un capitolo sulla geometria nel piano e nello spazio.

2 Algebra Lineare

2.1 Spazi Vettoriali

Spazio vettoriale

Definizione 2.1. Un insieme non vuoto \mathbf{V} si dice **spazio vettoriale** sopra un corpo commutativo \mathbb{K} (oppure spazio vettoriale su \mathbb{K} , o \mathbb{K} -spazio vettoriale) se sono soddisfatte le seguenti condizioni:

- L'insieme \mathbf{V} è dotato di una **legge di composizione interna**, indicata additivamente, definita al seguente modo:

$$+ : \mathbf{V} \times \mathbf{V} \rightarrow \mathbf{V} \quad (\mathbf{v}, \mathbf{w}) \mapsto \mathbf{v} + \mathbf{w}$$

l'elemento $\mathbf{v} + \mathbf{w}$ viene chiamato **somma** di \mathbf{v} e \mathbf{w} .

- L'insieme \mathbf{V} è dotato di una **legge di composizione esterna** indicata moltiplicativamente, definita al seguente modo:

$$\cdot : \mathbb{K} \times \mathbf{V} \rightarrow \mathbf{V} \quad (\lambda, \mathbf{v}) \mapsto \lambda \cdot \mathbf{v} = \lambda \mathbf{v}$$

l'elemento $\lambda \mathbf{v}$ viene chiamato **prodotto** di λ per \mathbf{v} .

- Le operazioni di addizione e di moltiplicazione per un elemento di \mathbb{K} soddisfano i seguenti assiomi:

SV1 Proprietà associativa:

$$\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{V}, \quad (\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$$

SV2 Esistenza dello zero:

$$\exists \mathbf{0} \in \mathbf{V} : \forall \mathbf{v} \in \mathbf{V}, \quad \mathbf{0} + \mathbf{v} = \mathbf{v} + \mathbf{0} = \mathbf{v}.$$

SV3 Esistenza dell'opposto:

Per ogni $\mathbf{v} \in \mathbf{V}$ l'elemento $(-1)\mathbf{v}$ soddisfa l'identità

$$\mathbf{v} + (-1)\mathbf{v} = \mathbf{0}.$$

SV4 Proprietà commutativa

$$\forall \mathbf{u}, \mathbf{v} \in \mathbf{V}, \quad \mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}.$$

SV5 Proprietà distributiva rispetto alla somma di vettori

$$\forall \mathbf{u}, \mathbf{v} \in \mathbf{V}, \forall \lambda \in \mathbb{K}, \quad \lambda(\mathbf{u} + \mathbf{v}) = \lambda \mathbf{u} + \lambda \mathbf{v}.$$

SV6 Proprietà distributiva rispetto alla somma di scalari

$$\forall \mathbf{v} \in \mathbf{V}, \forall \mu, \lambda \in \mathbb{K}, \quad (\mu + \lambda)\mathbf{v} = \mu \mathbf{v} + \lambda \mathbf{v}.$$

SV7

$$\forall \mathbf{v} \in \mathbf{V}, \forall \mu, \lambda \in \mathbb{K}, \quad (\mu \lambda)\mathbf{v} = \mu(\lambda \mathbf{v}).$$

SV8

$$\forall \mathbf{v} \in \mathbf{V}, \quad 1\mathbf{v} = \mathbf{v}.$$

Nota 2.

- In base ai campi e alle operazioni di somma e prodotto, si possono configurare diverse strutture di spazio vettoriale.
- Gli elementi dello spazio vettoriale \mathbf{V} si dicono **vettori**, e gli elementi di \mathbb{K} si dicono **scalari**.
- I vettori della forma $\lambda \mathbf{v}$, $\lambda \in \mathbb{K}$, si dicono **proporzionali** a (o **multipli** di) \mathbf{v} .
- Per ogni $\mathbf{v} \in \mathbf{V}$ il vettore $(-1)\mathbf{v}$, chiamato **l'opposto** di \mathbf{v} , si indica con $-\mathbf{v}$, e si scrive $\mathbf{u} - \mathbf{v}$ invece di $\mathbf{u} + (-1)\mathbf{v}$.

- Quando $\mathbb{K} = \mathbb{R}$ ($\mathbb{K} = \mathbb{C}$), si dice anche **spazio vettoriale reale** (**spazio vettoriale complesso**).
- Un insieme costituito da un solo elemento è in modo banale un \mathbb{K} -spazio vettoriale il cui unico elemento è il vettore nullo.

n -spazio numerico su \mathbb{K}

Proposizione 2.1. Siano $n \geq 1$ un intero e $\mathbf{V} = \mathbb{K}^n$, l'insieme delle n -uple ordinate di elementi di \mathbb{K} . Definiamo le seguenti operazioni:

$$+ : \mathbb{K}^n \times \mathbb{K}^n \rightarrow \mathbb{K}^n \quad ((x_1, \dots, x_n), (y_1, \dots, y_n)) \mapsto (x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$$

$$\cdot : \mathbb{K} \times \mathbb{K}^n \rightarrow \mathbb{K}^n \quad (\lambda, (x_1, \dots, x_n)) \mapsto \lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n)$$

Con queste operazioni, \mathbb{K}^n soddisfa gli assiomi SV1, ..., SV8 e quindi è un \mathbb{K} -spazio vettoriale. Esso viene solitamente chiamato l' **n -spazio numerico** su \mathbb{K} . L'1-spazio numerico è \mathbb{K} stesso, il quale, con le sue proprie operazioni di somma e di prodotto, è uno spazio vettoriale su sé stesso. Se $(x_1, \dots, x_n) \in \mathbb{K}^n$, gli scalari x_i , $i = 1, \dots, n$, si dicono **componenti**, e x_i l' **i -esima componente**, di (x_1, \dots, x_n) .

Proof. Già fatta. □

Spazio vettoriale di applicazioni

Proposizione 2.2. Sia I un insieme non vuoto qualunque e sia \mathbf{V} l'insieme i cui elementi sono le applicazioni da I a \mathbb{K} . Definiamo le seguenti operazioni:

$$+ : \mathbf{V} \times \mathbf{V} \rightarrow \mathbf{V} \quad (f, g) \mapsto f + g \quad \cdot : \mathbb{K} \times \mathbf{V} \rightarrow \mathbf{V} \quad (\lambda, f) \mapsto \lambda f$$

Con le operazioni che abbiamo introdotto, l'insieme \mathbf{V} è un \mathbb{K} -spazio vettoriale.

Proof. Da fare. □

Spazio vettoriale su un sottocampo

Proposizione 2.3. Sia \mathbf{F} un sottocampo di \mathbb{K} . Se \mathbf{V} è uno spazio vettoriale su \mathbb{K} , allora su \mathbf{V} resta indotta una struttura di \mathbf{F} -spazio vettoriale dalle stesse operazioni che definiscono la struttura di \mathbb{K} -spazio vettoriale. In altre parole, la somma di due vettori rimane la stessa, e la moltiplicazione di un vettore per uno scalare $\alpha \in \mathbf{F}$ si definisce considerando α come un elemento di \mathbb{K} e quindi utilizzando la definizione di moltiplicazione per elementi di \mathbb{K} .

Esempio 2.1. Il campo \mathbb{C} dei numeri complessi può essere considerato, oltre che uno spazio vettoriale complesso (l'1-spazio numerico su \mathbb{C}), anche come uno spazio vettoriale reale, perché \mathbb{R} è un sottocampo di \mathbb{C} .

Unicità del vettore nullo

Proposizione 2.4. Sia \mathbf{V} un \mathbb{K} -spazio vettoriale, allora

$$\exists! \mathbf{0} \in \mathbf{V} : \forall \mathbf{v} \in \mathbf{V}, \quad \mathbf{v} + \mathbf{0} = \mathbf{0} + \mathbf{v} = \mathbf{v}.$$

Alternativamente, siano $\mathbf{0}_1, \mathbf{0}_2 \in \mathbf{V}$ tali che

$$\forall \mathbf{v} \in \mathbf{V}, \quad \mathbf{0}_1 + \mathbf{v} = \mathbf{v} + \mathbf{0}_1 = \mathbf{v} \quad \wedge \quad \mathbf{0}_2 + \mathbf{v} = \mathbf{v} + \mathbf{0}_2 = \mathbf{v},$$

allora $\mathbf{0}_1 = \mathbf{0}_2$.

Proof. Infatti essendo entrambi appartenenti a \mathbf{V} , avremo che

$$\mathbf{0}_1 + \mathbf{0}_2 = \mathbf{0}_2 + \mathbf{0}_1 = \mathbf{0}_2 \quad \wedge \quad \mathbf{0}_2 + \mathbf{0}_1 = \mathbf{0}_1 + \mathbf{0}_2 = \mathbf{0}_1,$$

e dunque

$$\mathbf{0}_1 = \mathbf{0}_1 + \mathbf{0}_2 = \mathbf{0}_2 + \mathbf{0}_1 = \mathbf{0}_2.$$

□

Unicità del vettore opposto

Proposizione 2.5. *Sia \mathbf{V} un \mathbb{K} -spazio vettoriale. Allora*

$$\forall \mathbf{v} \in \mathbf{V}, \exists! (-1)\mathbf{v} \in \mathbf{V} : \mathbf{v} + (-1)\mathbf{v} = \mathbf{0}.$$

Alternativamente, siano $\mathbf{v}_1, \mathbf{v}_2 \in \mathbf{V}$ tali che

$$\forall \mathbf{v} \in \mathbf{V}, \quad \mathbf{v} + \mathbf{v}_1 = \mathbf{0} \quad \wedge \quad \mathbf{v} + \mathbf{v}_2 = \mathbf{0}$$

allora $\mathbf{v}_1 = \mathbf{v}_2$.

Proof. Infatti

$$\mathbf{v}_1 = \mathbf{v}_1 + \mathbf{0} = \mathbf{v}_1 + (\mathbf{v} + \mathbf{v}_2) = (\mathbf{v}_1 + \mathbf{v}) + \mathbf{v}_2 = \mathbf{0} + \mathbf{v}_2 = \mathbf{v}_2.$$

□

Proposizione 2.6. *Sia \mathbf{V} un \mathbb{K} -spazio vettoriale. Allora per ogni $\mathbf{a}, \mathbf{b} \in \mathbf{V}$, \mathbf{x} è l'unica soluzione dell'equazione $\mathbf{x} + \mathbf{a} = \mathbf{b}$ se e solo se $\mathbf{x} = \mathbf{b} - \mathbf{a}$.*

Proof. Dimostriamo le due implicazioni:

\Rightarrow Se $\mathbf{x} + \mathbf{a} = \mathbf{b}$, allora

$$\mathbf{x} = (\mathbf{x} + \mathbf{a}) - \mathbf{a} = \mathbf{b} - \mathbf{a}.$$

\Leftarrow Se $\mathbf{x} = \mathbf{b} - \mathbf{a}$, allora

$$\mathbf{x} + \mathbf{a} = (\mathbf{b} - \mathbf{a}) + \mathbf{a} = \mathbf{b}.$$

□

Proposizione 2.7. *Siano \mathbf{V} un \mathbb{K} -spazio vettoriale, $0 \in \mathbb{K}$ e $\mathbf{0} \in \mathbf{V}$. Allora*

$$\forall \mathbf{v} \in \mathbf{V}, \quad 0\mathbf{v} = \mathbf{0}.$$

Proof. Infatti $0\mathbf{v} = (0 + 0)\mathbf{v} = 0\mathbf{v} + 0\mathbf{v}$, e quindi $0\mathbf{v} = 0\mathbf{v} - 0\mathbf{v} = \mathbf{0}$. Analogamente si ha

$$\forall k \in \mathbb{K}, \quad k\mathbf{0} = \mathbf{0}.$$

Infatti $k\mathbf{0} = k(\mathbf{0} + \mathbf{0}) = k\mathbf{0} + k\mathbf{0}$, e quindi $k\mathbf{0} = k\mathbf{0} - k\mathbf{0} = \mathbf{0}$.

□

Associatività della somma di n vettori Sia \mathbf{V} un \mathbb{K} -spazio vettoriale. Dati tre vettori $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{V}$, scriveremo $\mathbf{u} + \mathbf{v} + \mathbf{w}$ per denotarne la somma effettuata in uno dei due modi possibili, i quali, per l'assioma SV1, danno lo stesso risultato. Supponiamo ora dati $n \geq 3$ vettori $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbf{V}$. Vogliamo dimostrare che la loro somma eseguita disponendo in un modo qualunque le parentesi:

$$(\mathbf{v}_1 + (\mathbf{v}_2 + (\dots + \mathbf{v}_n) \dots))$$

dà un risultato che dipende solo da $\mathbf{v}_1, \dots, \mathbf{v}_n$ e che quindi può essere denotato con $\mathbf{v}_1 + \dots + \mathbf{v}_n$, omettendo le parentesi. Procediamo per induzione su n : se $n = 3$ l'affermazione è vera per l'assioma SV1. Sia $n \geq 4$ e supponiamo dimostrato l'asserto per ogni intero k compreso tra 3 ed $n-1$. Possiamo scrivere due diverse somme come $\alpha + \beta$ e $\gamma + \delta$ rispettivamente, dove α, β sono somme in cui compaiono $\mathbf{v}_1, \dots, \mathbf{v}_k$, e $\mathbf{v}_{k+1}, \dots, \mathbf{v}_n$, rispettivamente, mentre in γ, δ compaiono $\mathbf{v}_1, \dots, \mathbf{v}_{j-1}, \mathbf{v}_j, \dots, \mathbf{v}_n$, rispettivamente, per opportuni interi positivi j, k minori di n . Se $k = j$, allora, per l'ipotesi induttiva, si ha $\alpha = \gamma$ e $\beta = \delta$, e quindi $\alpha + \beta = \gamma + \delta$. Supponiamo ora $k < j$. Per l'ipotesi induttiva si ha

$$\alpha + (\mathbf{v}_{k+1} + \dots + \mathbf{v}_j) = \gamma \quad \wedge \quad (\mathbf{v}_{k+1} + \dots + \mathbf{v}_j) + \gamma = \beta$$

e quindi $\alpha + \beta = \alpha + (\mathbf{v}_{k+1} + \dots + \mathbf{v}_j) + \delta = \gamma + \delta$.

2.2 Matrici

Matrice

Definizione 2.2. Siano $m, n \in \mathbb{N}$. Una **matrice** $m \times n$ a elementi in \mathbb{K} è una tabella rettangolare

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

di mn elementi di \mathbb{K} . Scriveremo $A = (a_{ij})$ con $1 \leq i \leq m, 1 \leq j \leq n$. L'**i-esima riga** della matrice A è la matrice $1 \times n$

$$A^{(i)} = (a_{i1} \ a_{i2} \ \cdots \ a_{in}),$$

e la **j-esima colonna** è la matrice $m \times 1$

$$A_{(j)} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

A possiede m righe ed n colonne. Ognuno degli elementi a_{ij} , della matrice è contrassegnato da due indici, di cui il primo denota la riga (**indice di riga**) e il secondo la colonna (**indice di colonna**) cui l'elemento appartiene. a_{ij} viene chiamato **l'elemento di posto i, j** . Se $m = n$ la matrice A si dice **quadrata di ordine n** . Una matrice $1 \times n$, cioè a una sola riga ed n colonne, viene anche chiamata **vettore riga** oppure **n -vettore riga**, mentre una matrice $n \times 1$, a n righe e una colonna, si dice **vettore colonna**, oppure **n -vettore colonna**.

Matrice trasposta

Definizione 2.3. La trasposta della matrice $m \times n$ $A = (a_{ij})$ è la matrice $n \times m$

$${}^t A = (a_{ji}) = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix},$$

ottenuta scambiando tra loro le righe e le colonne di A .

Insiemi di matrici a elementi in \mathbb{K}

Definizione 2.4. L'insieme di tutte le matrici $m \times n$ a elementi in \mathbb{K} si denota con $M_{m,n}(\mathbb{K})$ e l'insieme di tutte le matrici quadrate di ordine n con $M_n(\mathbb{K})$. Ovviamente $M_{n,n}(\mathbb{K}) = M_n(\mathbb{K})$, ed $M_1(\mathbb{K}) = \mathbb{K}$, perché una matrice 1×1 non è altro che un elemento di \mathbb{K} .

Nota 3. In seguito, identificheremo \mathbb{K}^n con $M_{n,1}(\mathbb{K})$, l'insieme degli n -vettori colonna. Pertanto un n -vettore colonna

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

verrà identificato con l'elemento $(x_1, x_2, \dots, x_n) \in \mathbb{K}^n$. Scriveremo anche

$$\mathbf{x} = {}^t(x_1 \ x_2 \ \dots \ x_n).$$

Insieme di matrici come spazio vettoriale

Proposizione 2.8. Ponendo

$$+ : M_{m,n}^2(\mathbb{K}) \rightarrow M_{m,n}(\mathbb{K}) \quad ((a_{ij}), (b_{ij})) \mapsto (a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$$

$$\cdot : \mathbb{K} \times M_{m,n}(\mathbb{K}) \rightarrow M_{m,n}(\mathbb{K}) \quad (\lambda, (a_{ij})) \mapsto \lambda(a_{ij}) = (\lambda a_{ij}),$$

si definisce su $M_{m,n}(\mathbb{K})$ una struttura di \mathbb{K} -spazio vettoriale. Lo zero è la matrice nulla $m \times n$, cioè la matrice $m \times n$ avente tutti i suoi elementi uguali a 0.

Proof. Da fare. □

Prodotto righe per colonne Se $A = (a_{ij})$, $B = (b_{ij})$, $k \in \mathbb{K}$, le matrici $(a_{ij} + b_{ij})$ e (ka_{ij}) si denoteranno con $A + B$ e kA rispettivamente. È possibile definire un'operazione di prodotto tra matrici. Dati un n -vettore riga

$$A = (a_1 \ a_2 \ \dots \ a_n)$$

e un n -vettore colonna

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix},$$

il loro **prodotto** è l'elemento di \mathbb{K} definito dalla seguente identità:

$$(a_1 \ a_2 \ \dots \ a_n) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + a_2 b_2 + \dots + a_n b_n.$$

Più in generale, date le seguenti matrici

$$A = (a_{ij}) \in M_{m,n}(\mathbb{K}) \quad B = (b_{jk}) \in M_{n,p}(\mathbb{K}),$$

il loro **prodotto righe per colonne** è la matrice $AB \in M_{m,p}(\mathbb{K})$ il cui elemento di posto i, k è il prodotto della i -esima riga di A per la k -esima colonna di B . In formule si ha

$$AB = (A^{(i)} B_{(k)}) = a_{i1} b_{1k} + a_{i2} b_{2k} + \dots + a_{in} b_{nk}.$$

Il prodotto di un vettore riga per un vettore colonna, è un caso particolare di prodotto righe per colonne di due matrici. Si noti che il prodotto AB è stato definito nell'ipotesi che il numero delle colonne di A sia uguale al numero delle righe di B : esprimeremo questo fatto dicendo che le matrici A e B possono essere moltiplicate. In particolare due qualsiasi matrici quadrate $A, B \in M_n(\mathbb{K})$ possono essere moltiplicate. Si noti però che in generale $AB \neq BA$.

Matrici diagonali, matrici identità

Definizione 2.5. Se $A = (a_{ij}) \in M_n(\mathbb{K})$, gli elementi $a_{11}, a_{22}, \dots, a_{nn}$ costituiscono la **diagonale principale** di A . Se tutti gli elementi a_{ij} con $i \neq j$ sono nulli, A si dice **matrice diagonale**. Una particolare matrice diagonale $n \times n$ è la **matrice unità** $\mathbf{I}_n = (\delta_{ij})$, dove

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

Il δ_{ij} è detto **simbolo di Kronecker**.

Esempio 2.2. Ad esempio

$$\mathbf{I}_1 = (1), \quad \mathbf{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Matrici triangolari

Definizione 2.6. Una matrice quadrata $A \in M_n(\mathbb{K})$ si dice **triangolare superiore** (**triangolare inferiore**) se $a_{ij} = 0$ per ogni $i > j$ (per ogni $i < j$). Si dice **strettamente triangolare superiore** (**strettamente triangolare inferiore**) se $a_{ij} = 0$ per ogni $i \geq j$ (per ogni $i \leq j$).

Esempio 2.3. Ad esempio, delle tre matrici quadrate a elementi reali:

$$\begin{pmatrix} 7 & 0 \\ 0 & 2 \end{pmatrix} \quad \begin{pmatrix} -1 & 2 & 4 \\ 0 & 0 & 1 \\ 0 & 0 & -3 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 \\ 2 & \pi & 0 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix}$$

la prima è diagonale, la seconda è triangolare superiore, la terza è strettamente triangolare inferiore.

Matrici simmetriche e antisimmetriche

Definizione 2.7. Una matrice quadrata $A \in M_n(\mathbb{K})$ tale che ${}^t A = A$ si dice **simmetrica**; se invece ${}^t A = -A$, allora A si dice **antisimmetrica**. Si noti che ogni matrice antisimmetrica ha necessariamente nulli gli elementi della diagonale principale.

Proposizioni fondamentali su somme e prodotti tra matrici

Proposizione 2.9.

1. Siano $A, B \in M_{m,n}(\mathbb{K})$, $C, D \in M_{n,p}(\mathbb{K})$, e $\lambda \in \mathbb{K}$. Allora:

- $(A + B)C = AC + BC$,
- $A(C + D) = AC + AD$,
- $A(\lambda C) = \lambda(AC) = (\lambda A)C$,
- $A\mathbf{I}_n = A$, $\mathbf{I}_n C = C$.

2. Siano $A \in M_{m,n}(\mathbb{K})$, $B \in M_{n,p}(\mathbb{K})$, $C \in M_{p,s}(\mathbb{K})$. Allora:

$$(AB)C = A(BC).$$

3. Siano A e B moltiplicabili. Allora ${}^t B$ e ${}^t A$ sono moltiplicabili e si ha

$${}^t(AB) = {}^t B {}^t A.$$

4. Siano $A, B \in M_{m,n}(\mathbb{K})$. Allora

$${}^t A + {}^t B = {}^t(A + B).$$

Proof.

1. Da fare.
2. Si osservi che $AB \in M_{m,p}(\mathbb{K})$ e $BC \in M_{n,s}(\mathbb{K})$, e pertanto sia AB e C che A e BC possono essere moltiplicate. L' i -esima riga di AB , $i = 1, \dots, m$, è

$$(AB)^{(i)} = (A^{(i)}B_{(1)} \ A^{(i)}B_{(2)} \ \dots \ A^{(i)}B_{(p)})$$

mentre la h -esima colonna di BC , $h = 1, \dots, s$, è

$$(BC)_{(h)} = \begin{pmatrix} B^{(1)}C_{(h)} \\ B^{(2)}C_{(h)} \\ \vdots \\ B^{(n)}C_{(h)} \end{pmatrix},$$

Pertanto l'elemento di posto i, h di $(AB)C$ è

$$\begin{aligned}(AB)^{(i)}C_{(h)} &= (A^{(i)}B_{(1)})c_{1h} + (A^{(i)}B_{(2)})c_{2h} + \dots + (A^{(i)}B_{(p)})c_{ph} \\ &= (a_{i1}b_{1i} + a_{i2}b_{2i} + \dots + a_{in}b_{ni})c_{1h} + \\ &\quad (a_{i1}b_{12} + a_{i2}b_{22} + \dots + a_{in}b_{n2})c_{2h} + \dots + \\ &\quad (a_{i1}b_{1p} + a_{i2}b_{2p} + \dots + a_{in}b_{np})c_{ph}.\end{aligned}$$

L'elemento di posto i, h di $A(BC)$ invece

$$\begin{aligned}A^{(i)}(BC)_{(h)} &= a_{i1}(B^{(1)}C_{(h)}) + a_{i2}(B^{(2)}C_{(h)}) + \dots + a_{in}(B^{(n)}C_{(h)}) \\ &= a_{i1}(b_{11}c_{1h} + b_{12}c_{2h} + \dots + b_{1p}c_{ph}) + \\ &\quad a_{i2}(b_{21}c_{1h} + b_{22}c_{2h} + \dots + b_{2p}c_{ph}) + \dots + \\ &\quad a_{in}(b_{n1}c_{1h} + b_{n2}c_{2h} + \dots + b_{np}c_{ph}).\end{aligned}$$

Confrontando vediamo che coincidono, perché entrambi sono la somma di tutti i prodotti della forma $a_{ij}b_{jk}c_{kh}$ al variare di $j = 1, \dots, n$ e di $k = 1, \dots, p$. Quindi le due matrici $(AB)C$ e $A(BC)$ coincidono elemento per elemento e l'asserzione è provata.

3. Supponiamo $A \in M_{m,n}(\mathbb{K})$, $B \in M_{n,p}(\mathbb{K})$. Allora ${}^tB \in M_{p,n}(\mathbb{K})$ e ${}^tA \in M_{n,m}(\mathbb{K})$, e quindi tB e tA possono essere moltiplicate. Si ha

$${}^t(AB)_{ji} = (AB)_{ij} = A^{(i)}B_{(j)} = ({}^tB)^{(j)}({}^tA)_{(i)} = ({}^tB {}^tA)_{ji}.$$

4. La dimostrazione è lasciata al lettore. □

Nota 4. Dalla (2.9.2) segue che possiamo scrivere ABC per denotare indifferentemente $(AB)C$ oppure $A(BC)$, perché queste due matrici coincidono.

Associatività del prodotto tra matrici

Proposizione 2.10. *Siano A_1, A_2, \dots, A_m matrici ad elementi in \mathbb{K} tali che A_k ed A_{k+1} possono essere moltiplicate per $k = 1, \dots, m-1$, il loro prodotto eseguito disponendo in un modo qualunque le parentesi:*

$$A_1(A_2(\dots A_m)\dots)$$

dà un risultato che dipende solo da A_1, A_2, \dots, A_m .

Matrice invertibile

Definizione 2.8. Sia $A \in M_n(\mathbb{K})$. Diciamo che A è **invertibile** se

$$\exists M \in M_n(\mathbb{K}) : AM = MA = \mathbf{I}_n.$$

M si dice **inversa** di A , e si denota con A^{-1} .

Unicità dell'inversa di una matrice

Proposizione 2.11. *Sia $A \in M_n(\mathbb{K})$ invertibile. Allora sua inversa è unica.*

Proof. Sia $N \in M_n(\mathbb{K})$ tale che $AN = NA = \mathbf{I}_n$, allora

$$M = M\mathbf{I}_n = M(AN) = (MA)N = \mathbf{I}_n N = N.$$

□

Se $A \in M_n(\mathbb{K})$ è invertibile, allora affinché una matrice $M \in M_n(\mathbb{K})$ sia l'inversa di A è sufficiente che sia verificata una sola delle due condizioni $AM = \mathbf{I}_n$ o $MA = \mathbf{I}_n$. Infatti, se per esempio $AM = \mathbf{I}_n$, si ha anche

$$MA = (A^{-1}A)MA = A^{-1}(AM)A = A^{-1}\mathbf{I}_nA = A^{-1}A = \mathbf{I}_n.$$

Similmente si dimostra che $MA = \mathbf{I}_n$ implica $AM = \mathbf{I}_n$.

$$AM = AM(AA^{-1}) = A(MA)A^{-1} = A\mathbf{I}_nA^{-1} = AA^{-1} = \mathbf{I}_n.$$

La matrice unità \mathbf{I}_n è invertibile e coincide con la sua inversa.

Proposizione 2.12. Per ogni $A \in M_n(\mathbb{K})$ invertibile, si ha che $(A^{-1})^{-1} = A$.

Se $A, B \in M_n(\mathbb{K})$ sono invertibili, allora anche AB lo è, e si ha

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

Infatti

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}\mathbf{I}_nB = B^{-1}B = \mathbf{I}_n.$$

Proposizione 2.13. Siano $A_1, A_2, \dots, A_{k-1}, A_k \in M_n(\mathbb{K})$ invertibili. Allora il loro prodotto $A_1A_2 \dots A_{k-1}A_k$ è invertibile, e si ha

$$(A_1A_2 \dots A_{k-1}A_k)^{-1} = A_k^{-1}A_{k-1}^{-1} \dots A_2^{-1}A_1^{-1}.$$

Proof. Procediamo per induzione su k . Per $k = 1$ l'affermazione è vera per definizione. Supponiamo l'affermazione vera per $k - 1$ e dimostriamola per k . Sia $B = A_1A_2 \dots A_{k-1}$. Per ipotesi induttiva:

$$B^{-1} = A_{k-1}^{-1} \dots A_2^{-1}A_1^{-1}$$

Allora:

$$\begin{aligned} (A_k^{-1}A_{k-1}^{-1} \dots A_2^{-1}A_1^{-1})(A_1A_2 \dots A_{k-1}A_k) &= \\ (A_k^{-1})(A_{k-1}^{-1} \dots A_2^{-1}A_1^{-1})(A_1A_2 \dots A_{k-1})A_k &= \\ (A_k^{-1})(B^{-1}B)A_k &= \\ A_k^{-1}\mathbf{I}_nA_k &= \\ A_k^{-1}A_k &= \\ \mathbf{I}_n \end{aligned}$$

Analogamente si dimostra che:

$$(A_1A_2 \dots A_{k-1}A_k)(A_k^{-1}A_{k-1}^{-1} \dots A_2^{-1}A_1^{-1}) = \mathbf{I}_n$$

Quindi per definizione:

$$(A_1A_2 \dots A_{k-1}A_k)^{-1} = A_k^{-1}A_{k-1}^{-1} \dots A_2^{-1}A_1^{-1}$$

□

Il sottoinsieme di $M_n(\mathbb{K})$ costituito dalle matrici invertibili viene denotato con $GL_n(\mathbb{K})$.

Se $k \geq 1$ è un intero, denoteremo con A^k il prodotto $AA \dots A$ di una matrice $A \in M_n(\mathbb{K})$ per sé stessa k volte; porremo $A^0 = \mathbf{I}$. Se $A \in GL_n(\mathbb{K})$ definiamo

$$A^{-k} = (A^{-1})^k.$$

Matrice ortogonale

Definizione 2.9. Una matrice quadrata reale $A \in M_n(\mathbb{R})$ si dice **ortogonale** se ${}^tAA = \mathbf{I}_n$, cioè se ${}^tA = A^{-1}$. L'insieme delle matrici ortogonali di ordine n si indica con $O(n)$. Per definizione $O(n) \subset GL_n(\mathbb{R})$.

Esempio 2.4. Le uniche matrici ortogonali 1×1 sono (1) e (-1) . Una matrice $A \in M_2(\mathbb{R})$ è ortogonale se e solo se è della forma

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad (1)$$

oppure

$$A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \quad (2)$$

con $a^2 + b^2 = 1$. Infatti, se $A = (a_{ij})$ si ha

$${}^tAA = \begin{pmatrix} a_{11}^2 + a_{21}^2 & a_{11}a_{12} + a_{21}a_{22} \\ a_{12}a_{11} + a_{22}a_{21} & a_{12}^2 + a_{22}^2 \end{pmatrix},$$

e quindi $A \in O(2)$ se e solo se

$$a_{11}^2 + a_{21}^2 = 1 = a_{12}^2 + a_{22}^2 \quad \wedge \quad a_{11}a_{12} + a_{21}a_{22} = 0.$$

Dall'ultima condizione si deduce che esiste $\rho \neq 0$ tale che

$$(a_{11}, a_{21}) = (-\rho a_{22}, \rho a_{12}).$$

Dalle condizioni precedenti discende che $\rho^2 = 1$, cioè $\rho = \pm 1$. Quindi $a_{12} = \pm a_{21}$, $a_{22} = \mp a_{11}$, e A è di una delle due forme dette.

Notazione a blocchi Per descrivere le matrici è spesso utile la cosiddetta **notazione a blocchi**. Tale notazione consiste nello scrivere una matrice $A \in M_{m,n}(\mathbb{K})$ nella forma seguente:

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1k} \\ A_{21} & A_{22} & \dots & A_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ A_{h1} & A_{h2} & \dots & A_{hk} \end{pmatrix},$$

dove le A_{ij} sono la loro volta matrici di ordini opportuni: precisamente $A_{ij} \in M_{m_i, n_j}(K)$, dove $m_1 + m_2 + \dots + m_h = m$, e $n_1 + n_2 + \dots + n_k = n$.

2.3 Sistemi di Equazioni Lineari

Equazione lineare

Definizione 2.10. Siano X_1, \dots, X_n , indeterminate. Un'equazione lineare (o di **primo grado**) nelle incognite X_1, \dots, X_n , a coefficienti in \mathbb{K} è un'equazione della forma

$$a_1X_1 + \dots + a_nX_n = b \quad \sum_{i=1}^n a_iX_i = b \quad (3)$$

oppure della forma equivalente

$$a_1X_1 + \dots + a_nX_n - b = 0 \quad \sum_{i=1}^n a_iX_i - b = 0$$

in cui $a_1, \dots, a_n, b \in \mathbb{K}$. Una **soluzione dell'equazione** (3) è un elemento

$$(x_1, \dots, x_n) \in \mathbb{K}^n,$$

sostituito nella (3) al posto della n -upla (X_1, \dots, X_n) , dà luogo a una identità. La (3) si dice **omogenea** (**non omogenea**) se $b = 0$ (se $b \neq 0$).

Sistema di equazioni lineari

Definizione 2.11. Se si considerano simultaneamente $m \geq 1$ equazioni lineari nelle incognite X_1, \dots, X_n :

$$\begin{array}{l} a_{11}X_1 + a_{12}X_2 + \dots + a_{1n}X_n = b_1 \\ a_{21}X_1 + a_{22}X_2 + \dots + a_{2n}X_n = b_2 \\ \vdots \\ a_{m1}X_1 + a_{m2}X_2 + \dots + a_{mn}X_n = b_m \end{array} \quad \sum_{j=1}^n a_{ij}X_j = b_j \quad i = 1, \dots, m \quad (4)$$

si ottiene un **sistema di m equazioni lineari nelle n incognite** X_1, \dots, X_n . Il sistema (4) si dice **omogeneo** (**non omogeneo**) se $b_1 = b_2 = \dots = b_m = 0$ (se $b_i \neq 0$ per qualche i). Una **soluzione del sistema** (4) è un elemento (x_1, \dots, x_n) di \mathbb{K}^n che è soluzione simultanea delle m equazioni (4). Il sistema si dice **compatibile** (**incompatibile**) se possiede almeno una soluzione (se non possiede soluzioni). Ogni sistema omogeneo ammette almeno la soluzione $(0, \dots, 0)$, che viene detta **soluzione banale**, e quindi è compatibile; ogni sua altra soluzione si dice **non banale**. Si noti che, viceversa, se il sistema (4) ammette la soluzione $(0, \dots, 0)$, allora è omogeneo.

Esempio 2.5. Il sistema di equazioni a coefficienti reali

$$\begin{array}{l} X_1 + 2X_2 = 1 \\ X_1 + 2X_2 = 0 \end{array}$$

è incompatibile, perché i primi membri delle due equazioni sono uguali, ma non lo sono i secondi membri e quindi non esiste alcun (x_1, x_2) in \mathbb{R}^2 che soddisfi entrambe le equazioni.

Esempio 2.6. Il sistema

$$\begin{array}{l} X_1 + X_2 = 1 \\ X_1 - X_2 = 3 \end{array}$$

è compatibile e ammette l'unica soluzione $(2, -1)$, che si ottiene nel modo seguente. Sommando membro a membro le due equazioni, si ottiene la nuova equazione $2X_1 = 4$, che è soddisfatta dall'unico valore $X_1 = 2$; sostituendo questo valore nella prima equazione si ottiene l'unico valore $X_2 = -1$ che la soddisfa. Inoltre la coppia $(2, -1)$ è soluzione anche della seconda equazione, e quindi è l'unica soluzione del sistema.

Esempio 2.7. Il sistema

$$\begin{aligned} X_1 + 3X_2 &= -1 \\ 2X_1 + 6X_2 &= -2 \end{aligned}$$

è compatibile ed ammette le infinite soluzioni $(-1 - 3t, t)$ al variare del parametro t in \mathbb{R} . Infatti le due equazioni sono proporzionali e quindi hanno le stesse soluzioni: risolvendo per esempio la prima si trovano le soluzioni dette.

Sistema omogeneo associato

Definizione 2.12. Il sistema

$$\begin{aligned} a_{11}X_1 + a_{12}X_2 + \dots + a_{1n}X_n &= 0 \\ a_{21}X_1 + a_{22}X_2 + \dots + a_{2n}X_n &= 0 \\ &\vdots \\ a_{m1}X_1 + a_{m2}X_2 + \dots + a_{mn}X_n &= 0 \end{aligned} \tag{5}$$

si dice il **sistema omogeneo associato** al sistema (4).

Proposizione 2.14. Se il sistema (4) è compatibile, le sue soluzioni sono tutte e sole le n -uple ottenute sommando a una qualsiasi di esse una soluzione del sistema omogeneo associato (5).

Proof. Denotiamo con Σ e Σ_0 i due sottoinsiemi di \mathbb{K}^n i cui elementi sono rispettivamente le soluzioni del sistema (4) e del sistema (5). Siano

$$(y_1, \dots, y_n) \in \Sigma \quad \text{e} \quad (z_1, \dots, z_n) \in \Sigma_0,$$

allora

$$(y_1, \dots, y_n) + (z_1, \dots, z_n) = (y_1 + z_1, \dots, y_n + z_n) \in \Sigma.$$

Infatti

$$\forall j = 1, \dots, m, \quad \sum_{i=1}^n a_{ji}(y_i + z_i) = \sum_{i=1}^n (a_{ji}y_i + a_{ji}z_i) = \sum_{i=1}^n a_{ji}y_i + \sum_{i=1}^n a_{ji}z_i = b_j + 0 = b_j.$$

Viceversa, fissata $(y_1, \dots, y_n) \in \Sigma$, per ogni altra $(z_1, \dots, z_n) \in \Sigma$ si ha

$$(z_1 - y_1, \dots, z_n - y_n) \in \Sigma_0.$$

Infatti

$$\forall j = 1, \dots, m, \quad \sum_{i=1}^n a_{ji}(z_i - y_i) = \sum_{i=1}^n (a_{ji}z_i - a_{ji}y_i) = \sum_{i=1}^n a_{ji}z_i - \sum_{i=1}^n a_{ji}y_i = b_j - b_j = 0.$$

Poiché

$$(z_1, \dots, z_n) = (y_1, \dots, y_n) + (z_1 - y_1, \dots, z_n - y_n)$$

abbiamo l'asserzione. □

Matrice dei coefficienti, matrice orlata associata

Definizione 2.13. Al sistema (4) possiamo associare la matrice $A = (a_{ij}) \in M_{m,n}(\mathbb{K})$ formata dai coefficienti delle incognite delle m equazioni del sistema, che si dice la **matrice dei coefficienti** del sistema (4). Aggiungendo ad A come $(n+1)$ -esima colonna la

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

formata dai termini costanti delle equazioni (4), si ottiene la matrice ad m righe e $n+1$ colonne:

$$(A\mathbf{b}) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix},$$

che diremo la **matrice orlata** del sistema (4).

Possiamo interpretare gli m primi membri di (4) come le componenti di un vettore colonna e riscrivere la (4) come un'uguaglianza di vettori colonna:

$$\begin{pmatrix} a_{11}X_1 + a_{12}X_2 + \dots + a_{1n}X_n \\ a_{21}X_1 + a_{22}X_2 + \dots + a_{2n}X_n \\ \vdots \\ a_{m1}X_1 + a_{m2}X_2 + \dots + a_{mn}X_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n a_{1i}X_i \\ \sum_{i=1}^n a_{2i}X_i \\ \vdots \\ \sum_{i=1}^n a_{mi}X_i \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \quad (6)$$

Ponendo

$$\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix},$$

e considerando \mathbf{X} come un vettore colonna, il primo membro della (6) è il prodotto righe per colonne $A\mathbf{X}$. Il sistema (4) si scrive quindi anche nella seguente forma più concisa:

$$A\mathbf{X} = \mathbf{b}. \quad (7)$$

Viceversa è evidente che per ogni matrice a m righe ed $n + 1$ colonne esiste un sistema di m equazioni lineari nelle incognite X_1, \dots, X_n di cui essa è la matrice orlata.

Sistema a gradini

Definizione 2.14. Un sistema di equazioni lineari nelle incognite X_1, \dots, X_n si dice a **gradini** se ha la forma seguente:

$$\begin{array}{rcl} a_{11}X_1 + a_{12}X_2 + \dots & + a_{1n}X_n = b_1 \\ & a_{22}X_2 + \dots & + a_{2n}X_n = b_2 \\ & & \vdots \\ & a_{mm}X_m + \dots + a_{mn}X_n = b_m \end{array} \quad (8)$$

con $a_{11}a_{22} \dots a_{mm} \neq 0$. La matrice dei coefficienti di (8) è

$$\begin{pmatrix} a_{11} & a_{12} & \dots & & a_{1n} \\ 0 & a_{22} & \dots & & a_{2n} \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & \dots & a_{mm} & \dots & a_{mn} \end{pmatrix}.$$

Ovviamente abbiamo che $m \leq n$. Analizziamo però i casi specifici separatamente.

$m = n$ Riscriviamo il sistema al seguente modo:

$$\begin{array}{rcl} a_{11}X_1 + a_{12}X_2 + & \dots + a_{1n}X_n = b_1 \\ & a_{22}X_2 + & \dots + a_{2n}X_n = b_2 \\ & & \vdots \\ & & a_{nn}X_n = b_n \end{array}$$

L'ultima equazione di è soddisfatta dal solo valore $x_n = b_n a_{nn}^{-1}$, il quale, sostituito nella penultima equazione, fornisce un unico valore x_{n-1} , che la soddisfa. I valori x_1, \dots, x_{n-1} , così ottenuti, sostituiti nella terz'ultima equazione, danno luogo a un unico valore x_{n-2} che la soddisfa. Procedendo in questo modo si arriva a ottenere un'unica soluzione. Quindi un sistema a gradini di n equazioni in n incognite possiede un'unica soluzione.

$m < n$ Riscriviamo il sistema al seguente modo:

$$\begin{aligned} a_{11}X_1 + a_{12}X_2 + \cdots + a_{1m}X_m &= b_1 - (a_{1m+1}X_{m+1} + \cdots + a_{1n}X_n) \\ a_{22}X_2 + \cdots + a_{2m}X_m &= b_2 - (a_{2m+1}X_{m+1} + \cdots + a_{2n}X_n) \\ &\vdots \\ a_{mm}X_m &= b_m - (a_{mm+1}X_{m+1} + \cdots + a_{mn}X_n) \end{aligned}$$

Dando valori arbitrari $t_{m+1}, \dots, t_n \in \mathbb{K}$ alle incognite X_{m+1}, \dots, X_n si ottiene un sistema a gradini di m equazioni nelle m incognite X_1, \dots, X_m :

$$\begin{aligned} a_{11}X_1 + a_{12}X_2 + \cdots + a_{1m}X_m &= b_1 - (a_{1m+1}t_{m+1} + \cdots + a_{1n}t_n) \\ a_{22}X_2 + \cdots + a_{2m}X_m &= b_2 - (a_{2m+1}t_{m+1} + \cdots + a_{2n}t_n) \\ &\vdots \\ a_{mm}X_m &= b_m - (a_{mm+1}t_{m+1} + \cdots + a_{mn}t_n) \end{aligned} \tag{9}$$

il quale ha un'unica soluzione. Ne deduciamo che il sistema (8) ammette le infinite soluzioni ottenute dalle (9) al variare dei parametri $t_{m+1}, \dots, t_n \in \mathbb{K}$. Dal modo in cui si calcolano le soluzioni si deduce che ogni soluzione di (9) si esprime come una n -upla

$$(S_1(t_{m+1}, \dots, t_n), S_2(t_{m+1}, \dots, t_n), \dots, S_n(t_{m+1}, \dots, t_n)) \tag{10}$$

in cui gli $S_i(t_{m+1}, \dots, t_n)$ sono polinomi di primo grado nei parametri t_{m+1}, \dots, t_n . La (10) è la soluzione generale del sistema (8). La n -upla dei termini costanti (c_1, \dots, c_n) degli S_i è una delle soluzioni, precisamente quella corrispondente ai valori $t_{m+1} = \dots = t_n = 0$. Da ciò e dalla proposizione 3.1 segue che la n -upla di polinomi omogenei in t_{m+1}, \dots, t_n ,

$$(S_1(t_{m+1}, \dots, t_n) - c_1, S_2(t_{m+1}, \dots, t_n) - c_2, \dots, S_n(t_{m+1}, \dots, t_n) - c_n)$$

è la soluzione generale del sistema omogeneo associato a (8).

In particolare vediamo che un sistema a gradini è sempre compatibile. Esprimiamo il fatto che le soluzioni di (8) si ottengono come funzioni di $n - m$ parametri liberi di variare arbitrariamente, dicendo che il sistema (8) possiede ∞^{n-m} **soluzioni**. Nel caso $n = m$ intenderemo con ciò dire che il sistema possiede una sola soluzione.

Caso 1. Un'equazione lineare (3) in cui $(a_1, \dots, a_n) \neq (0, \dots, 0)$ si può considerare come un particolare sistema a gradini, salvo cambiare l'ordine delle variabili se $a_1 = 0$; pertanto essa possiede ∞^{n-1} soluzioni.

Sistemi equivalenti

Definizione 2.15. Due sistemi di equazioni lineari nelle stesse incognite X_1, \dots, X_n si dicono **equivalenti** se possiedono le stesse soluzioni. Per essere equivalenti due sistemi non devono necessariamente avere lo stesso numero di equazioni.

Metodo di eliminazione di Gauss-Jordan Il metodo di eliminazione di Gauss-Jordan è un processo che permette di stabilire se un sistema è compatibile oppure no, e nel caso affermativo di trovare sistematicamente tutte le soluzioni. Tale procedimento consiste nel sostituire il sistema assegnato con un sistema a gradini, ad esso equivalente, mediante passaggi successivi detti **operazioni elementari sulle equazioni del sistema**. Esse corrispondono ad altrettante operazioni sulle righe della matrice orlata. Esistono tre tipi di operazioni elementari sulle righe di una matrice:

1. scambiare tra loro due righe della matrice;
2. moltiplicare una riga della matrice per uno scalare non nullo;
3. sostituire una riga della matrice con quella ottenuta sommando ad essa un multiplo di un'altra riga.

Le corrispondenti operazioni elementari sulle equazioni di un sistema sono le seguenti:

1. scambiare tra loro due equazioni del sistema;
2. moltiplicare (primo e secondo membro di) un'equazione per uno stesso scalare non nullo;

3. sostituire un'equazione con quella ottenuta sommando ad essa un multiplo di un'altra equazione.

Se si effettua su di un sistema un'operazione elementare del tipo (1), il nuovo sistema che si ottiene è equivalente al precedente, perché le soluzioni di un sistema non dipendono dall'ordine in cui si considerano le sue equazioni. Similmente un'operazione del tipo (2) non cambia l'insieme delle soluzioni del sistema perché due equazioni proporzionali hanno le stesse soluzioni. Anche un'operazione del tipo (3) non modifica l'insieme delle soluzioni del sistema: infatti se una n -upla (x_1, \dots, x_n) di \mathbb{K}^n soddisfa due equazioni del sistema,

$$\begin{aligned} a_{i1}X_1 + a_{i2}X_2 + \dots + a_{in}X_n &= b_i \\ a_{j1}X_1 + a_{j2}X_2 + \dots + a_{jn}X_n &= b_j \end{aligned} \quad (11)$$

allora per un qualsiasi $c \in \mathbb{K}$ essa è soluzione delle due equazioni

$$\begin{aligned} a_{i1}X_1 + a_{i2}X_2 + \dots + a_{in}X_n &= b_i \\ a_{j1}X_1 + a_{j2}X_2 + \dots + a_{jn}X_n + c(a_{i1}X_1 + a_{i2}X_2 + \dots + a_{in}X_n) &= b_j + cb_i \end{aligned} \quad (12)$$

Si verifica in modo simile che viceversa ogni soluzione delle (12) soddisfa le (11). Quindi, se si effettua una qualsiasi operazione elementare sulle equazioni di un sistema si ottiene un sistema ad esso equivalente.

Supponiamo dunque di avere assegnato un sistema (4). Osserviamo preliminarmente che se una delle sue equazioni, diciamo la i -esima, ha identicamente nullo il primo membro, cioè è della forma

$$0 = b_i,$$

allora essa è identicamente soddisfatta se $b_i = 0$, mentre è incompatibile se $b_i \neq 0$. Nel primo caso potremo cancellare l'equazione e ottenere un sistema equivalente al precedente, nel secondo caso il sistema (4) è incompatibile. Possiamo pertanto supporre che nessuno dei primi membri di (4) sia identicamente nullo. Possiamo inoltre supporre che sia $a_{i1} \neq 0$ per qualche $i = 1, \dots, m$: ciò può essere ottenuto scambiando eventualmente tra loro due delle incognite. Con un'operazione elementare (1) possiamo ottenere $a_{i1} \neq 0$, e moltiplicando per a_{i1}^{-1} la prima equazione (operazione elementare (2)) possiamo ridurci al caso $a_{i1} = 1$. Sommando alle successive equazioni la prima moltiplicata rispettivamente per

$$-a_{21} - a_{31}, \dots, -a_{m1} \quad \text{operazione elementare (3)}$$

si ottiene un nuovo sistema della forma seguente:

$$\begin{aligned} X_1 + a'_{12}X_2 + \dots + a'_{1n}X_n &= b'_1 \\ a'_{22}X_2 + \dots + a'_{2n}X_n &= b'_2 \\ &\vdots \\ a'_{m2}X_2 + \dots + a'_{mn}X_n &= b'_m \end{aligned} \quad (13)$$

Se qualcuna delle equazioni del sistema (13) è della forma $0 = 0$, possiamo ometterla senza modificare l'insieme delle soluzioni. Se invece compare un'equazione della forma $0 = b'_i$, con $b'_i \neq 0$, allora il sistema è incompatibile, e pertanto anche (4) è incompatibile e il procedimento si arresta. Possiamo pertanto supporre che nessuno dei primi membri del sistema (13) sia identicamente nullo.

Ora procediamo sul sistema (13) senza più occuparci della prima equazione e ragionando, sulle rimanenti equazioni, come nel caso precedente. Effettuando eventualmente un cambiamento dell'ordine delle variabili ed operazioni elementari (I) (II) possiamo supporre $a'_{22} = 1$. Sommando alle successive equazioni la prima moltiplicata rispettivamente per

$$-a'_{32} - a'_{42}, \dots, -a'_{m2} \quad \text{operazione elementare III,}$$

si ottiene un nuovo sistema della forma seguente:

$$\begin{aligned} X_1 + a'_{12}X_2 + a'_{13}X_3 + \dots + a'_{1n}X_n &= b'_1 \\ X_2 + a''_{23}X_3 + \dots + a''_{2n}X_n &= b''_2 \\ a''_{33}X_3 + \dots + a''_{3n}X_n &= b''_3 \\ &\vdots \\ a''_{s3}X_3 + \dots + a''_{sn}X_n &= b''_s \end{aligned} \quad (14)$$

Dopo aver eliminato dal sistema (14) tutte le equazioni della forma $0 = 0$, verifichiamo se vi compare un'equazione della forma $0 = b''_i$, $b''_i \neq 0$; in caso affermativo il sistema è incompatibile e pertanto anche (4) lo è, ed il procedimento ha termine. In caso contrario applichiamo di nuovo lo stesso procedimento al sistema (14) escludendo le prime due equazioni. Questo procedimento potrà essere iterato fin tanto che non si arriva un sistema incompatibile oppure a un

sistema a gradini equivalente al sistema (4) da cui eravamo partiti. Nel primo caso possiamo concludere che il sistema (4) è incompatibile. Nel secondo caso possiamo calcolare le soluzioni del sistema (4) e incominciamo. La soluzione generale del sistema a gradini che si è ottenuto è detta **soluzione generale del sistema** (4).

Nota 5. Ogni sistema omogeneo di m equazioni in n incognite, con $n \geq m$, possiede ∞^N soluzioni per qualche $N \geq n - m$. Infatti il sistema è compatibile perché omogeneo, e il procedimento di Gauss-Jordan lo trasforma in un sistema a gradini di p equazioni con $p \leq m$. Quindi il sistema originario possiede ∞^{n-p} soluzioni ed $n - p \geq n - m$.

Osservazione 1. Per i sistemi lineari a gradini, possiamo distinguere i seguenti casi:

- $n = m$
1. Sistema omogeneo:
 - È sempre compatibile
 - Ha un'unica soluzione: il vettore nullo
 - Non presenta parametri liberi ($\infty^0 = 1$ soluzione)
 2. Sistema non omogeneo:
 - È compatibile
 - Ha un'unica soluzione
 - Non presenta parametri liberi ($\infty^0 = 1$ soluzione)
- $n > m$
1. Sistema omogeneo:
 - È sempre compatibile
 - Ha infinite soluzioni (∞^{n-m} soluzioni)
 - Dipende da $(n - m)$ parametri liberi
 - Include la soluzione nulla quando tutti i parametri sono zero
 2. Sistema non omogeneo:
 - È compatibile
 - Ha infinite soluzioni (∞^{n-m} soluzioni)
 - Dipende da $(n - m)$ parametri liberi
 - Include una soluzione particolare quando tutti i parametri sono zero

Metodo dell'inversa

Proposizione 2.15. Siano $A \in M_n(\mathbb{K})$ e $\mathbf{b} \in \mathbb{K}^n$. Se il sistema

$$A\mathbf{X} = \mathbf{b}$$

di n equazioni in n incognite ha matrice dei coefficienti invertibile, allora esso è compatibile e ammette un'unica soluzione, data da

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

Proof. Sostituendo $\mathbf{x} = A^{-1}\mathbf{b}$ in $A\mathbf{X} = \mathbf{b}$ si ottiene

$$A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = \mathbf{I}_n \mathbf{b} = \mathbf{b},$$

che dimostra che \mathbf{x} è soluzione. Viceversa, se $\mathbf{y} \in \mathbb{K}^n$ è una soluzione del sistema, allora $A\mathbf{y} = \mathbf{b}$ e, moltiplicando entrambi i membri per A^{-1} , si ha

$$\mathbf{y} = A^{-1}(A\mathbf{y}) = A^{-1}\mathbf{b},$$

da cui segue che la soluzione è unica. □

Matrice elementare

Definizione 2.16. Una **matrice elementare** di ordine n è una matrice $R \in M_n(\mathbb{K})$ che può essere ottenuta dalla matrice identità per mezzo di un'operazione elementare sulle righe.

Ad esempio, ognuna delle seguenti matrici è elementare di ordine 4 a elementi reali:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Introduciamo le seguenti notazioni per le matrici elementari di ordine n :

- R_{ij}^n : matrice ottenuta scambiando tra loro l' i -esima e la j -esima riga di \mathbf{I}_n ;
- $R_i^n(c)$: matrice ottenuta moltiplicando per $c \in \mathbb{K}^*$ la i -esima riga di \mathbf{I}_n ;
- $R_{ij}^n(c)$: matrice ottenuta sommando alla i -esima riga di \mathbf{I}_n , la j -esima moltiplicata per $c \in \mathbb{K}$.

L'utilità di queste matrici sta nel fatto che, se $A \in M_{n,m}(\mathbb{K})$, allora ogni operazione elementare sulle righe di A si ottiene moltiplicando a sinistra A per la corrispondente matrice elementare. Le seguenti identità sono di verifica immediata:

1. $R_{ij}^{-1} = R_{ji}$,
2. $R_i(c)^{-1} = R_i(c^{-1})$,
3. $R_{ij}(c)^{-1} = R_{ij}(-c)$.

In particolare vediamo che le matrici elementari sono invertibili e hanno per inverse matrici elementari dello stesso tipo.

Collegamento tra invertibilità e prodotto di matrici elementari

Proposizione 2.16. *Sia $A \in M_n(\mathbb{K})$. Le seguenti condizioni sono equivalenti:*

1. A è invertibile;
2. A si può esprimere come prodotto di matrici elementari.

Proof.

$1 \Rightarrow 2$ Il sistema omogeneo di n equazioni in n incognite $A\mathbf{X} = \mathbf{0}$ ha l'unica soluzione $\mathbf{x} = \mathbf{0}$. Quindi, utilizzando il procedimento di Gauss-Jordan, con operazioni elementari esso può essere trasformato in un sistema a gradini di n equazioni omogenee in n incognite, cioè della forma

$$\begin{aligned} X_1 + a'_{12}X_2 + \dots + a'_{1n}X_n &= 0 \\ X_2 + \dots + a'_{2n}X_n &= 0 \\ &\vdots \\ X_{n-1} + a'_{n-1n}X_n &= 0 \\ X_n &= 0. \end{aligned}$$

Con ulteriori operazioni elementari del tipo (III) è possibile ridurre questo sistema nella forma $\mathbf{X} = \mathbf{0}$, cioè

$$\begin{aligned} X_1 &= 0 \\ X_2 &= 0 \\ &\vdots \\ X_n &= 0. \end{aligned}$$

Tale trasformazione corrisponde alla moltiplicazione a sinistra del primo membro del sistema $A\mathbf{X} = \mathbf{0}$ per il prodotto di un numero finito di matrici elementari, cioè si ha

$$R^{(1)} \dots R^{(s)} A\mathbf{X} = \mathbf{X} = \mathbf{I}_n \mathbf{X}$$

per opportune matrici elementari $R^{(1)}, \dots, R^{(s)}$. Pertanto $R^{(1)}, R^{(2)}, \dots, R^{(s)} A = \mathbf{I}_n$, per l'unicità di A^{-1} si ha

$$R^{(1)} \dots R^{(s)} = A^{-1}$$

e quindi

$$A = (R^{(1)} \dots R^{(s)})^{-1} = R^{(s)-1} \dots R^{(2)-1} R^{(1)-1}$$

è un prodotto di matrici elementari.

$2 \Rightarrow 1$ Se A è un prodotto di matrici elementari, allora è invertibile perché ognuno dei fattori lo è.

□

Calcolo dell'inversa Siano $A, B \in M_n(\mathbb{K})$. Se $M \in M_n(\mathbb{K})$, allora M e la matrice $(A \ B) \in M_{n, 2n}$, possono essere moltiplicate, e dalla definizione di prodotto righe per colonne segue che

$$M(A \ B) = (MA \ MB).$$

Supponiamo che la matrice $A \in M_n(\mathbb{K})$ sia invertibile e consideriamo la matrice

$$(A \mid \mathbf{I}_n) \in M_{n, 2n}.$$

Moltiplicandola a sinistra per A^{-1} otteniamo

$$A^{-1}(A \mid \mathbf{I}_n) = (\mathbf{I}_n \mid A^{-1}).$$

Siccome A^{-1} è esprimibile come prodotto di matrici elementari, vediamo che la matrice $(\mathbf{I}_n \mid A^{-1})$ può essere ottenuta a partire da $(A \mid \mathbf{I}_n)$ mediante operazioni elementari sulle righe.

Si ottiene così il seguente metodo pratico per stabilire se una matrice data $A \in M_n(\mathbb{K})$ è invertibile e, se lo è, per trovare A^{-1} . Si considera la matrice $(A \mid \mathbf{I}_n)$: effettuando operazioni elementari sulle sue righe, è possibile ottenere una matrice della forma $(\mathbf{I}_n \mid B)$, allora A è invertibile e la matrice B così ottenuta necessariamente coincide con A^{-1} . Se invece ciò non è possibile, allora A non è invertibile.

2.4 Sottospazi Vettoriali

Sottospazio vettoriale

Definizione 2.17. Sia \mathbf{V} un \mathbb{K} -spazio vettoriale. Un sottoinsieme non vuoto \mathbf{W} di \mathbf{V} si dice **sottospazio vettoriale** di \mathbf{V} se:

1. $\forall \mathbf{w}_1, \mathbf{w}_2 \in \mathbf{W}, \quad \mathbf{w}_1 + \mathbf{w}_2 \in \mathbf{W};$
2. $\forall \mathbf{w} \in \mathbf{W}, \forall c \in \mathbb{K}, \quad c\mathbf{w} \in \mathbf{W}.$

Le condizioni (1) e (2) implicano che le operazioni di somma e di prodotto per uno scalare definite in \mathbf{V} inducono altrettante operazioni in \mathbf{W} .

Per la proprietà (2) applicata agli scalari 0, -1, si ha rispettivamente

$$\mathbf{0} = 0\mathbf{w} \in \mathbf{W} \quad \text{e} \quad -\mathbf{w} = -1\mathbf{w} \in \mathbf{W}.$$

Quindi \mathbf{W} soddisfa gli assiomi SV2 ed SV3. Poiché gli altri assiomi SV1, SV4, ..., SV8 sono soddisfatti da \mathbf{V} , essi sono a maggior ragione soddisfatti da \mathbf{W} . Quindi \mathbf{W} è esso stesso uno spazio vettoriale. È inoltre evidente che se \mathbf{W} è un sottospazio di \mathbf{V} e \mathbf{U} è un sottospazio di \mathbf{W} , allora \mathbf{U} è un sottospazio di \mathbf{V} . Analogamente, se \mathbf{U} e \mathbf{W} sono sottospazi di \mathbf{V} ed $\mathbf{U} \subset \mathbf{W}$, allora \mathbf{U} è un sottospazio di \mathbf{W} .

Esempi di sottospazi vettoriali

Esempio 2.8. Per ogni spazio vettoriale \mathbf{V} , \mathbf{V} stesso e il sottoinsieme costituito dal solo $\mathbf{0}$ sono sottospazi vettoriali detti **sottospazi impropri** o **banali** di \mathbf{V} .

Copertura lineare di un vettore

Esempio 2.9. Sia $\mathbf{v} \in \mathbf{V}$ un vettore qualsiasi. L'insieme

$$\langle \mathbf{v} \rangle = \{c\mathbf{v} : c \in \mathbb{K}\}$$

costituito da tutti i multipli di \mathbf{v} costituisce un sottospazio di \mathbf{V} .

Esempio 2.10. Sia $\mathbf{V} = \mathbb{K}^n$, $n \geq 2$. Il sottoinsieme \mathbf{H}_1 di \mathbf{V} costituito dalle n -uple della forma $(0, x_2, \dots, x_n)$, al variare di $x_2, \dots, x_n \in \mathbb{K}$, è un sottospazio vettoriale di \mathbb{K}^n . Infatti \mathbf{H}_1 non è vuoto; inoltre, presi qualsiasi $x_2, \dots, x_n, y_2, \dots, y_n, c \in \mathbb{K}$, si ha

$$(0, x_2, \dots, x_n) + (0, y_2, \dots, y_n) = (0, x_2 + y_2, \dots, x_n + y_n) \in \mathbf{H}_1,$$

e

$$c(0, x_2, \dots, x_n) = (0, cx_2, \dots, cx_n) \in \mathbf{H}_1.$$

In modo analogo si verifica che per ogni indice i compreso tra 1 ed n il sottoinsieme \mathbf{H}_i di \mathbb{K}^n costituito dalle n -uple il cui i -esimo elemento è uguale a 0 è un sottospazio vettoriale. Dati $a_1, \dots, a_n \in \mathbb{K}$ non tutti uguali a 0, sia

$$\mathbf{H} = \left\{ (x_1, \dots, x_n) \in \mathbb{K}^n : \sum_{i=1}^n a_i x_i = 0 \right\}.$$

Se $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbf{H}$, $c \in \mathbb{K}$, allora

$$\sum_{i=1}^n a_i (x_i + y_i) = \sum_{i=1}^n (a_i x_i + a_i y_i) = \sum_{i=1}^n a_i x_i + \sum_{i=1}^n a_i y_i = 0 + 0 = 0.$$

$$\sum_{i=1}^n a_i (cx_i) = c \sum_{i=1}^n a_i x_i = c \cdot 0 = 0.$$

cioè $(x_1, \dots, x_n) + (y_1, \dots, y_n) \in \mathbf{H}$ e $c(x_1, \dots, x_n) \in \mathbf{H}$; pertanto \mathbf{H} è un sottospazio vettoriale di \mathbb{K}^n . I sottospazi \mathbf{H}_i , sono casi particolari di \mathbf{H} che si ottengono prendendo $a_i = 1$ e $a_j = 0$, se $j \neq i$.

Esempio 2.11. L'insieme $\Sigma \subset \mathbb{K}^n$ costituito dalle soluzioni di un assegnato sistema di equazioni lineari omogenee in n incognite è un sottospazio vettoriale di \mathbb{K}^n .

Intersezione di sottospazi

Proposizione 2.17. Siano \mathbf{U} e \mathbf{W} due sottospazi vettoriali di \mathbf{V} , l'intersezione $\mathbf{U} \cap \mathbf{W}$ è ancora un sottospazio vettoriale di \mathbf{V} . In generale, l'intersezione $\bigcap_{i \in I} \mathbf{W}_i$ di una famiglia qualsiasi $\{\mathbf{W}_i\}_{i \in I}$ di sottospazi vettoriali di \mathbf{V} è un sottospazio vettoriale di \mathbf{V} .

Unione di due sottospazi L'unione $\mathbf{U} \cup \mathbf{W}$ di due sottospazi non è in generale un sottospazio di \mathbf{V} . Ad esempio, se \mathbf{u} e \mathbf{w} sono due vettori non proporzionali di \mathbf{V} , allora $\langle \mathbf{u} \rangle \cup \langle \mathbf{w} \rangle$ contiene i vettori \mathbf{u} e \mathbf{w} ma non contiene $\mathbf{u} + \mathbf{w}$, perché questo non è multiplo né di \mathbf{u} né di \mathbf{w} .

Sottospazio somma

Definizione 2.18. Il sottoinsieme di \mathbf{V} costituito da tutti i vettori della forma $\mathbf{u} + \mathbf{w}$, al variare di \mathbf{u} in \mathbf{U} e di \mathbf{w} in \mathbf{W} , verrà denotato con $\mathbf{U} + \mathbf{W}$. Se $\mathbf{u}_1, \mathbf{u}_2 \in \mathbf{U}$, $\mathbf{w}_1, \mathbf{w}_2 \in \mathbf{W}$, e $c \in \mathbb{K}$ allora

$$(\mathbf{u}_1 + \mathbf{w}_1) + (\mathbf{u}_2 + \mathbf{w}_2) = (\mathbf{u}_1 + \mathbf{u}_2) + (\mathbf{w}_1 + \mathbf{w}_2) \in \mathbf{U} + \mathbf{W}$$

e

$$c(\mathbf{u}_1 + \mathbf{w}_1) = c\mathbf{u}_1 + c\mathbf{w}_1 \in \mathbf{U} + \mathbf{W}$$

e quindi $\mathbf{U} + \mathbf{W}$ è un sottospazio vettoriale di \mathbf{V} . Esso si chiama il **sottospazio somma dei sottospazi \mathbf{U} e \mathbf{W}** .

Nota 6. Si osservi che $\mathbf{U} + \mathbf{W}$ contiene $\mathbf{U} \cup \mathbf{W}$ perché contiene tutti i vettori $\mathbf{u} = \mathbf{u} + \mathbf{0}$ e $\mathbf{w} = \mathbf{0} + \mathbf{w}$ al variare di $\mathbf{u} \in \mathbf{U}$ e di $\mathbf{w} \in \mathbf{W}$.

Somma diretta di due sottospazi

Definizione 2.19. Se $\mathbf{U} \cap \mathbf{W} = \langle \mathbf{0} \rangle$, allora $\mathbf{U} + \mathbf{W}$ è detto **somma diretta** di \mathbf{U} e \mathbf{W} , e si denota con $\mathbf{U} \oplus \mathbf{W}$. Ogni vettore di $\mathbf{U} \oplus \mathbf{W}$ si esprime in modo unico nella forma $\mathbf{u} + \mathbf{w}$. Infatti, se

$$\mathbf{u} + \mathbf{w} = \mathbf{u}' + \mathbf{w}'$$

per qualche $\mathbf{u}, \mathbf{u}' \in \mathbf{U}$, $\mathbf{w}, \mathbf{w}' \in \mathbf{W}$, allora $\mathbf{u} - \mathbf{u}' = \mathbf{w}' - \mathbf{w} \in \mathbf{U} \cap \mathbf{W}$, e quindi

$$\mathbf{u} - \mathbf{u}' = \mathbf{w}' - \mathbf{w} = \mathbf{0},$$

cioè

$$\mathbf{u} = \mathbf{u}' \quad \text{e} \quad \mathbf{w} = \mathbf{w}'.$$

Sottospazio supplementare

Definizione 2.20. Se $\mathbf{V} = \mathbf{U} \oplus \mathbf{W}$ i sottospazi \mathbf{U} e \mathbf{W} si dicono **supplementari** in \mathbf{V} . Nel caso in cui $\mathbf{U} = \langle \mathbf{u} \rangle$, $\mathbf{W} = \langle \mathbf{w} \rangle$, con $\mathbf{u}, \mathbf{w} \in \mathbf{V}$ non proporzionali, si ha $\langle \mathbf{u} \rangle \cap \langle \mathbf{w} \rangle = \langle \mathbf{0} \rangle$. $\langle \mathbf{u} \rangle \oplus \langle \mathbf{w} \rangle$ consiste di tutti i vettori della forma $a\mathbf{u} + b\mathbf{w}$, al variare di $a, b \in \mathbb{K}$.

Prodotto cartesiano di due \mathbb{K} -spazi vettoriali Se \mathbf{U} e \mathbf{W} sono due \mathbb{K} -spazi vettoriali, il loro prodotto cartesiano $\mathbf{U} \times \mathbf{W}$ è un \mathbb{K} -spazio vettoriale se si definiscono in esso le operazioni di somma di due vettori e di prodotto di un vettore per uno scalare nel modo seguente:

$$(\mathbf{u}, \mathbf{w}) + (\mathbf{u}', \mathbf{w}') = (\mathbf{u} + \mathbf{u}', \mathbf{w} + \mathbf{w}') \quad \wedge \quad \lambda(\mathbf{u}, \mathbf{w}) = (\lambda\mathbf{u}, \lambda\mathbf{w}) \quad \forall (\mathbf{u}, \mathbf{w}), (\mathbf{u}', \mathbf{w}') \in \mathbf{U} \times \mathbf{W}, \lambda \in \mathbb{K}.$$

Il vettore nullo di $\mathbf{U} \times \mathbf{W}$ è $(\mathbf{0}, \mathbf{0})$. I sottoinsiemi

$$\mathbf{U}' = \{(\mathbf{u}, \mathbf{0}) : \mathbf{u} \in \mathbf{U}\}, \quad \mathbf{W}' = \{(\mathbf{0}, \mathbf{w}) : \mathbf{w} \in \mathbf{W}\}$$

sono due sottospazi di $\mathbf{U} \times \mathbf{W}$. Si ha evidentemente

$$\mathbf{U}' \cap \mathbf{W}' = \{(\mathbf{0}, \mathbf{0})\};$$

inoltre

$$\forall (\mathbf{u}, \mathbf{w}) \in \mathbf{U} \times \mathbf{W}, \quad (\mathbf{u}, \mathbf{w}) = (\mathbf{u}, \mathbf{0}) + (\mathbf{0}, \mathbf{w}).$$

Pertanto si ha

$$\mathbf{U} \times \mathbf{W} = \mathbf{U}' \oplus \mathbf{W}'.$$

Ad esempio lo spazio vettoriale \mathbb{K}^{m+n} si identifica con $\mathbb{K}^n \times \mathbb{K}^m$, ed è somma diretta dei due sottospazi

$$\mathbb{K}^{n'} = \{(x_1, x_2, \dots, x_n, 0, \dots, 0) : x_1, \dots, x_n \in \mathbb{K}\}$$

$$\mathbb{K}^{m'} = \{(0, \dots, 0, y_1, y_2, \dots, y_m) : y_1, \dots, y_m \in \mathbb{K}\}$$

Un procedimento per costruire sottospazi vettoriali è fornito dalla nozione di **combinazione lineare**.

Combinazione lineare

Definizione 2.21. Siano $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbf{V}$ e $a_1, \dots, a_n \in \mathbb{K}$. Il vettore

$$\sum_{i=1}^n a_i \mathbf{v}_i = a_1 \mathbf{v}_1 + \dots + a_n \mathbf{v}_n$$

si dice **combinazione lineare dei vettori** $\mathbf{v}_1, \dots, \mathbf{v}_n$ e a_1, \dots, a_n si dicono **coefficienti della combinazione lineare**. Se i coefficienti sono tutti uguali a 0, allora abbiamo una **combinazione lineare banale** di $\mathbf{v}_1, \dots, \mathbf{v}_n$ che sarà uguale al vettore $\mathbf{0}$. Ogni combinazione lineare di $\mathbf{v}_1, \dots, \mathbf{v}_n$ in cui i coefficienti non siano tutti nulli si dice **non banale**.

Nota 7.

- Una combinazione lineare può essere non banale e tuttavia essere uguale al vettore nullo. Questo è il caso ad esempio di $0\mathbf{v} + a\mathbf{0}$, per ogni $\mathbf{v} \in \mathbf{V}$, $a \in \mathbb{K}^*$.
- Le combinazioni lineari di un vettore $\mathbf{v} \in \mathbf{V}$ sono i suoi multipli.
- Si noti che

$$\mathbf{v}_i = 0\mathbf{v}_1 + \dots + 0\mathbf{v}_{i-1} + 1\mathbf{v}_i + 0\mathbf{v}_{i+1} + \dots + 0\mathbf{v}_n$$

è combinazione lineare di $\mathbf{v}_1, \dots, \mathbf{v}_n$, $i = 1, \dots, n$.

- Se \mathbf{W} è un sottospazio vettoriale di \mathbf{V} e $\mathbf{v}_1, \dots, \mathbf{v}_n$ sono elementi di \mathbf{W} , allora ogni combinazione lineare di $\mathbf{v}_1, \dots, \mathbf{v}_n$ appartiene a \mathbf{W} .

Sottospazio generato e sistema di generatori

Proposizione 2.18. Sia $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ un sottoinsieme finito di vettori di \mathbf{V} . L'insieme $\langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$ costituito da tutte le combinazioni lineari di $\mathbf{v}_1, \dots, \mathbf{v}_n$ è un sottospazio vettoriale di \mathbf{V} . Esso è uguale all'intersezione di tutti i sottospazi di \mathbf{V} che contengono $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$.

Proof. Siano $\sum_{i=1}^n a_i \mathbf{v}_i, \sum_{i=1}^n b_i \mathbf{v}_i \in \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$ e $\lambda \in \mathbb{K}$, allora

$$\sum_{i=1}^n a_i \mathbf{v}_i + \sum_{i=1}^n b_i \mathbf{v}_i = \sum_{i=1}^n (a_i \mathbf{v}_i + b_i \mathbf{v}_i) = \sum_{i=1}^n (a_i + b_i) \mathbf{v}_i$$

e

$$\lambda \sum_{i=1}^n a_i \mathbf{v}_i = \sum_{i=1}^n \lambda a_i \mathbf{v}_i$$

sono elementi di $\langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$, e quindi $\langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$ è un sottospazio vettoriale di \mathbf{V} . □

Denotiamo con \mathbf{W} l'intersezione di tutti i sottospazi di \mathbf{V} che contengono $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. Poiché $\langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$ è un sottospazio vettoriale contenente $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, si ha $\mathbf{W} \subset \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$. D'altra parte \mathbf{W} , essendo un sottospazio, contiene tutte le combinazioni lineari di suoi elementi, e quindi contiene quelle di $\mathbf{v}_1, \dots, \mathbf{v}_n$, cioè $\mathbf{W} \supset \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$. In conclusione $\mathbf{W} = \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$. Chiameremo $\langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$ il **sottospazio generato** da $\mathbf{v}_1, \dots, \mathbf{v}_n$. Osserviamo che, se $1 \leq m \leq n$, il sottospazio $\langle \mathbf{v}_1, \dots, \mathbf{v}_m \rangle$ è contenuto in $\langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$, perché ogni combinazione lineare di $\mathbf{v}_1, \dots, \mathbf{v}_m$ è anche una combinazione lineare di $\mathbf{v}_1, \dots, \mathbf{v}_n$:

$$a_1 \mathbf{v}_1 + \dots + a_m \mathbf{v}_m = a_1 \mathbf{v}_1 + \dots + a_m \mathbf{v}_m + 0 \mathbf{v}_{m+1} + \dots + 0 \mathbf{v}_n.$$

Sistema di generatori

Definizione 2.22. Diremo che $\mathbf{v}_1, \dots, \mathbf{v}_n$ **generano** \mathbf{V} , oppure che $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ è un **sistema di generatori** di \mathbf{V} , se $\langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle = \mathbf{V}$. Quindi $\mathbf{v}_1, \dots, \mathbf{v}_n$ generano \mathbf{V} se e solo se

$$\forall \mathbf{v} \in \mathbf{V}, \exists a_1, \dots, a_n \in \mathbb{K} : \mathbf{v} = \sum_{i=1}^n a_i \mathbf{v}_i.$$

Vettori linearmente dipendenti e linearmente indipendenti

Definizione 2.23. I vettori $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbf{V}$ si dicono **linearmente dipendenti** se esistono scalari $a_1, \dots, a_n \in \mathbb{K}$ non tutti nulli tali che

$$\sum_{i=1}^n a_i \mathbf{v}_i = \mathbf{0},$$

o, equivalentemente, se il vettore $\mathbf{0}$ si può esprimere come loro combinazione lineare non banale. Altrimenti $\mathbf{v}_1, \dots, \mathbf{v}_n$ si dicono **linearmente indipendenti**.

Proprietà sui vettori linearmente dipendenti e indipendenti

Proposizione 2.19. *Un vettore \mathbf{v} è linearmente dipendente se e solo se $\mathbf{v} = \mathbf{0}$.*

Proposizione 2.20. *Se \mathbf{v}_1 e \mathbf{v}_2 sono due vettori tali che \mathbf{v}_2 sia proporzionale a \mathbf{v}_1 , cioè tali che $\mathbf{v}_2 = a \mathbf{v}_1$ per qualche $a \in \mathbb{K}$, allora \mathbf{v}_1 e \mathbf{v}_2 sono linearmente dipendenti.*

Proof. Infatti $a \mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0}$ è una loro combinazione lineare con coefficienti $a \neq -1$, e quindi non entrambi nulli. Viceversa, se \mathbf{v}_1 e \mathbf{v}_2 sono due vettori linearmente dipendenti, allora uno di essi è multiplo dell'altro. Infatti

$$a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 = \mathbf{0},$$

cioè $a_2 \mathbf{v}_2 = -a_1 \mathbf{v}_1$, con a_1, a_2 non entrambi uguali a 0, implica, supponendo ad esempio $a_2 \neq 0$, che $\mathbf{v}_2 = a \mathbf{v}_1$ dove $a = -\frac{a_1}{a_2}$. □

Proposizione 2.21. $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbf{V}$, $n \geq 2$, sono linearmente dipendenti se e solo se uno almeno di essi si può esprimere come combinazione lineare dei rimanenti.

Proof. Infatti, se $\mathbf{v}_1, \dots, \mathbf{v}_n$ sono linearmente dipendenti, allora

$$\mathbf{0} = a_1 \mathbf{v}_1 + \dots + a_n \mathbf{v}_n$$

con $a_i \neq 0$ per qualche i ; quindi

$$a_i \mathbf{v}_i = -(a_1 \mathbf{v}_1 + \dots + a_{i-1} \mathbf{v}_{i-1} + a_{i+1} \mathbf{v}_{i+1} + \dots + a_n \mathbf{v}_n)$$

cioè

$$\begin{aligned} \mathbf{v}_i &= -a_i^{-1} (a_1 \mathbf{v}_1 + \dots + a_{i-1} \mathbf{v}_{i-1} + a_{i+1} \mathbf{v}_{i+1} + \dots + a_n \mathbf{v}_n) \\ &= -a_i^{-1} a_1 \mathbf{v}_1 - \dots - a_i^{-1} a_{i-1} \mathbf{v}_{i-1} - a_i^{-1} a_{i+1} \mathbf{v}_{i+1} - \dots - a_i^{-1} a_n \mathbf{v}_n. \end{aligned}$$

Viceversa, se per qualche i

$$\mathbf{v}_i = b_1 \mathbf{v}_1 + \dots + b_{i-1} \mathbf{v}_{i-1} + b_{i+1} \mathbf{v}_{i+1} + \dots + b_n \mathbf{v}_n$$

allora

$$\mathbf{0} = b_1 \mathbf{v}_1 + \dots + b_{i-1} \mathbf{v}_{i-1} - \mathbf{v}_i + b_{i+1} \mathbf{v}_{i+1} + \dots + b_n \mathbf{v}_n$$

e quindi $\mathbf{v}_1, \dots, \mathbf{v}_n$ sono linearmente dipendenti. \square

Proposizione 2.22. Se l'insieme $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ contiene il vettore $\mathbf{0}$, allora $\mathbf{v}_1, \dots, \mathbf{v}_n$ sono linearmente dipendenti.

Proof. Supponiamo infatti che si abbia $\mathbf{v}_i = \mathbf{0}$ per qualche i compreso tra 1 ed n . Allora si ha

$$0\mathbf{v}_1 + \dots + 0\mathbf{v}_{i-1} + 1\mathbf{v}_i + 0\mathbf{v}_{i+1} + \dots + 0\mathbf{v}_n = \mathbf{v}_i = \mathbf{0},$$

e $\mathbf{0}$ è una combinazione lineare non banale di $\mathbf{v}_1, \dots, \mathbf{v}_n$. Si può anche osservare che $\mathbf{v}_i = \mathbf{0}$ può essere espresso come combinazione lineare di $\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_n$ la loro combinazione lineare banale, e quindi $\mathbf{v}_1, \dots, \mathbf{v}_n$ sono linearmente dipendenti per la proposizione (2.21). \square

Proposizione 2.23. Se $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ sono linearmente indipendenti e $1 \leq m < n$, allora $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ sono linearmente indipendenti. Equivalentemente, se $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ sono linearmente dipendenti, allora anche $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ sono linearmente dipendenti.

Proof. Dimostriamo la seconda affermazione. Se $\mathbf{v}_1, \dots, \mathbf{v}_m$ sono linearmente dipendenti, allora

$$\mathbf{0} = a_1 \mathbf{v}_1 + \dots + a_m \mathbf{v}_m = a_1 \mathbf{v}_1 + \dots + a_m \mathbf{v}_m + 0\mathbf{v}_{m+1} + \dots + 0\mathbf{v}_n$$

per qualche scelta di $a_1, \dots, a_m \in \mathbb{K}$ non tutti nulli, e quindi il terzo membro è una combinazione lineare non banale di $\mathbf{v}_1, \dots, \mathbf{v}_n$, che è uguale a $\mathbf{0}$, cioè $\mathbf{v}_1, \dots, \mathbf{v}_n$ sono linearmente dipendenti. \square

Proposizione 2.24. Se $\mathbf{v}_1, \dots, \mathbf{v}_n$ sono linearmente indipendenti, e $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{K}$ sono tali che

$$\sum_{i=1}^n a_i \mathbf{v}_i = \sum_{i=1}^n b_i \mathbf{v}_i$$

allora $a_1 = b_1, a_2 = b_2, \dots, a_n = b_n$.

Proof. Infatti, essendo

$$\mathbf{0} = \sum_{i=1}^n a_i \mathbf{v}_i - \sum_{i=1}^n b_i \mathbf{v}_i = \sum_{i=1}^n (a_i \mathbf{v}_i - b_i \mathbf{v}_i) = \sum_{i=1}^n (a_i - b_i) \mathbf{v}_i$$

si deve avere $a_1 - b_1 = a_2 - b_2 = \dots = a_n - b_n = 0$. \square

Base

Definizione 2.24. Un sottoinsieme finito $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ di \mathbf{V} si dice **base finita** (o **base**) di \mathbf{V} se $\mathbf{v}_1, \dots, \mathbf{v}_n$ sono linearmente indipendenti e generano \mathbf{V} .

Se $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ è una base, allora poiché $\mathbf{v}_1, \dots, \mathbf{v}_n$ generano \mathbf{V} , avremo che

$$\forall \mathbf{v} \in \mathbf{V}, \exists a_1, \dots, a_n \in \mathbb{K} : \mathbf{v} = \sum_{i=1}^n a_i \mathbf{v}_i \quad (15)$$

Per la proposizione (2.24), a_1, \dots, a_n , sono univocamente determinati da \mathbf{v} . I coefficienti a_1, \dots, a_n , della combinazione lineare (15) si dicono le **coordinate di \mathbf{v} rispetto alla base** $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, e a_1, \dots, a_n si dice la **n -upla delle coordinate di \mathbf{v}** . Quindi, una volta assegnata una base $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, ad ogni vettore \mathbf{v} viene univocamente associata una n -upla di coordinate; viceversa ogni n -upla $(a_1, \dots, a_n) \in \mathbb{K}^n$ individua univocamente il vettore (15) di cui essa è la n -upla delle coordinate. Un vettore \mathbf{v} di coordinate a_1, \dots, a_n verrà spesso denotato con $\mathbf{v}(a_1, \dots, a_n)$. Lo spazio vettoriale $\{\mathbf{0}\}$ costituito dal solo vettore $\mathbf{0}$ non possiede una base finita, perché il suo unico elemento è linearmente dipendente. Non tutti gli spazi vettoriali diversi da $\{\mathbf{0}\}$ possiedono una base finita.

Teorema 2.25. Sia $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ un sistema di generatori di \mathbf{V} e siano $\mathbf{w}_1, \dots, \mathbf{w}_m$ elementi di \mathbf{V} . Se $m > n$, allora $\mathbf{w}_1, \dots, \mathbf{w}_m$ sono linearmente dipendenti.

Corollario 1. Siano $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ e $\{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ due basi dello spazio vettoriale \mathbf{V} . Allora $m = n$.

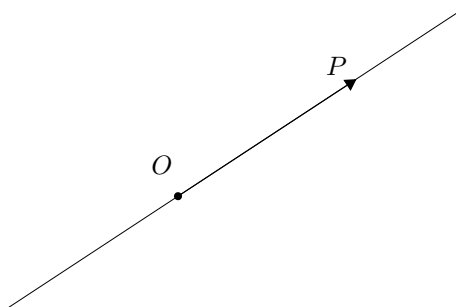
Proof. Poiché $\mathbf{v}_1, \dots, \mathbf{v}_n$ generano \mathbf{V} e $\mathbf{w}_1, \dots, \mathbf{w}_m$ sono linearmente indipendenti, dal teorema (2.25) segue che $m \leq n$. D'altra parte, poiché $\mathbf{w}_1, \dots, \mathbf{w}_m$ generano \mathbf{V} e $\mathbf{v}_1, \dots, \mathbf{v}_n$ sono linearmente indipendenti, dallo stesso teorema segue che anche $n \leq m$. Quindi $m = n$. \square

Dimensione

Definizione 2.25. Se un \mathbb{K} -spazio vettoriale \mathbf{V} possiede una base finita $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, il numero n si dice la **dimensione** di \mathbf{V} , e si denota con $\dim_{\mathbb{K}}(\mathbf{V})$, o semplicemente $\dim(\mathbf{V})$. Se $\mathbf{V} = \{\mathbf{0}\}$ consiste del solo vettore nullo, si pone $\dim(\mathbf{V}) = 0$. Se $\mathbf{V} = \{\mathbf{0}\}$, oppure \mathbf{V} possiede una base finita, diremo che \mathbf{V} ha **dimensione finita**.

Esempio 2.12.

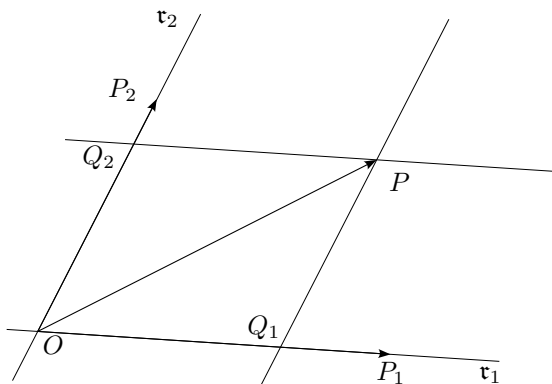
1. Sia $\mathbf{v} = OP$ un vettore geometrico non nullo della retta ordinaria. Poiché ogni altro vettore geometrico è multiplo di \mathbf{v} , $\{\mathbf{v}\}$ è una base dello spazio dei vettori geometrici della retta ordinaria, il quale quindi ha dimensione 1.



2. Sia \mathbf{V} lo spazio vettoriale reale dei vettori geometrici della retta ordinaria, e siano \mathbf{v}_1 e \mathbf{v}_2 due vettori non proporzionali di \mathbf{V} , che penseremo applicati in uno stesso punto O , rappresentati nella forma

$$\mathbf{v}_1 = \overrightarrow{OP_1} \quad \mathbf{v}_2 = \overrightarrow{OP_2}$$

per opportuni punti P_1 e P_2 . Dalla proposizione (2.20) segue che \mathbf{v}_1 e \mathbf{v}_2 sono linearmente indipendenti.



Denotiamo con \mathbf{r}_1 (con \mathbf{r}_2) la retta contenente i punti O e P_1 (i punti O e P_2). Sia $\mathbf{v} = \overrightarrow{OP} \in \mathbf{V}$ un vettore qualsiasi. Detto Q_1 (rispettivamente Q_2) il punto di intersezione con \mathbf{r}_1 (con \mathbf{r}_2) della retta parallela a \mathbf{r}_2 (a \mathbf{r}_1) e passante per P , sia

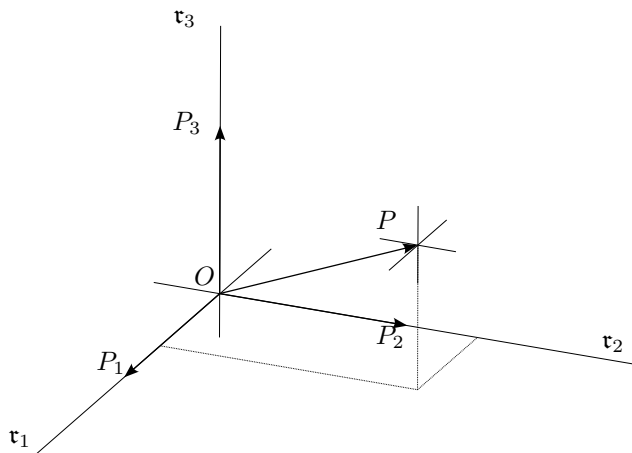
$$\overrightarrow{OQ_1} = a_1 \overrightarrow{OP_1} \quad \text{e} \quad \overrightarrow{OQ_2} = a_2 \overrightarrow{OP_2}$$

per opportuni $a_1, a_2 \in \mathbb{R}$. Poiché

$$\mathbf{v} = \overrightarrow{OQ_1} + \overrightarrow{OQ_2} = a_1 \overrightarrow{OP_1} + a_2 \overrightarrow{OP_2} = a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2$$

vediamo che \mathbf{v} è combinazione lineare di \mathbf{v}_1 e \mathbf{v}_2 , e che quindi $\mathbf{v}_1, \mathbf{v}_2$ generano \mathbf{V} . Dunque $\{\mathbf{v}_1, \mathbf{v}_2\}$ è una base di \mathbf{V} . Pertanto \mathbf{V} ha dimensione 2.

3. Nello spazio ordinario tre vettori geometrici $\mathbf{v}_1 = \overrightarrow{OP_1}$, $\mathbf{v}_2 = \overrightarrow{OP_2}$, $\mathbf{v}_3 = \overrightarrow{OP_3}$ sono linearmente indipendenti se e solo se i punti O, P_1, P_2, P_3 non giacciono in un stesso piano: ciò segue subito dal fatto che la dipendenza lineare di $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ equivale all'essere uno di essi combinazione lineare dei rimanenti. Se $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ sono linearmente indipendenti, allora si verifica che ogni vettore geometrico si può esprimere come loro combinazione lineare. Quindi $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ è una base, e i vettori geometrici dello spazio ordinario costituiscono uno spazio vettoriale reale di dimensione 3.



Base canonica Consideriamo lo spazio vettoriale numerico \mathbb{K}^n con $n \geq 1$, e siano

$$\mathbf{E}_1 = (1, 0, \dots, 0), \quad \mathbf{E}_2 = (0, 1, \dots, 0), \quad \dots, \quad \mathbf{E}_n = (0, \dots, 0, 1).$$

I vettori $\mathbf{E}_1, \dots, \mathbf{E}_n$ generano \mathbb{K}^n . Infatti per ogni $(x_1, \dots, x_n) \in \mathbb{K}^n$ si ha

$$(x_1, \dots, x_n) = \sum_{i=1}^n x_i \mathbf{E}_i. \quad (16)$$

Inoltre, essendo $(x_1, \dots, x_n) = \mathbf{0}$ se e solo se $x_1 = x_2 = \dots = x_n = 0$, segue dalla (16) che $\mathbf{E}_1, \dots, \mathbf{E}_n$ sono linearmente indipendenti. Quindi $\{\mathbf{E}_1, \dots, \mathbf{E}_n\}$ è una base di \mathbb{K}^n e \mathbb{K}^n ha dimensione n .

$$\{\mathbf{E}_1, \dots, \mathbf{E}_n\}$$

si chiama la **base canonica** di \mathbb{K}^n . Per $n = 1$ otteniamo il campo \mathbb{K} , considerato come uno spazio vettoriale su sé stesso, ha dimensione 1, la sua base canonica essendo costituita dal solo $1 \in \mathbb{K}$. Dalla (16) segue che le coordinate di un vettore (x, \dots, x_n) , rispetto alla base canonica coincidono con le sue componenti x_1, \dots, x_n .

Equazioni cartesiane di un sottospazio Siano $A \in M_{m,n}(\mathbb{K})$, $\mathbf{X} = {}^t(X_1, X_2, \dots, X_n)$ indeterminate, e

$$A\mathbf{X} = \mathbf{0} \quad (17)$$

il sistema omogeneo, di m equazioni nelle incognite \mathbf{X} , la cui matrice dei coefficienti è A . Denotiamo con Σ_0 l'insieme delle soluzioni del sistema. Come abbiamo già osservato nell'esempio (2.10), Σ_0 è un sottospazio vettoriale di \mathbb{K}^n . Le m equazioni (17) si dicono **equazioni cartesiane del sottospazio Σ_0 di \mathbb{K}^n** . Quindi, per definizione, due diversi sistemi di equazioni lineari omogenee nelle incognite \mathbf{X} sono equivalenti se e solo se sono equazioni cartesiane dello stesso sottospazio di \mathbb{K}^n . Gli elementi di Σ_0 possono interpretarsi come relazioni di dipendenza lineare tra le colonne $A_{(1)}, A_{(2)}, \dots, A_{(n)}$ di A , considerate come vettori di \mathbb{K}^m . Infatti $\mathbf{x} = {}^t(x_1, x_2, \dots, x_n) \in \mathbb{K}^n$ è soluzione di (17) se e solo se $A\mathbf{x} = \mathbf{0}$, equivalentemente, se e solo se

$$A_{(1)}x_1 + A_{(2)}x_2 + \dots + A_{(n)}x_n = \sum_{i=1}^n A_{(i)}x_i = \mathbf{0}.$$

Consideriamo un sistema di m equazioni lineari omogenee a gradini:

$$\begin{aligned} a_{11}X_1 + a_{12}X_2 + \dots &+ a_{1n}X_n = 0 \\ a_{22}X_2 + \dots &+ a_{2n}X_n = 0 \\ &\vdots \\ a_{mm}X_m + \dots + a_{mn}X_n &= 0, \end{aligned} \quad (18)$$

dove $a_{ii} \neq 0$ per $i = 1, \dots, m$. Se $n = m$, il sistema (18) ammette solo la soluzione banale. Se $n > m$, le ∞^{n-m} soluzioni sono ottenute dando valori arbitrari alle incognite X_{m+1}, \dots, X_n e risolvendo il corrispondente sistema di m equazioni nelle X_1, \dots, X_m . Consideriamo in particolare le $n - m$ soluzioni

$$S_{m+1} = \begin{pmatrix} S_{m+1,1} \\ S_{m+1,2} \\ \vdots \\ S_{m+1,m} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad S_{m+2} = \begin{pmatrix} S_{m+2,1} \\ S_{m+2,2} \\ \vdots \\ S_{m+2,m} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} \quad \dots \quad S_n = \begin{pmatrix} S_{n,1} \\ S_{n,2} \\ \vdots \\ S_{n,m} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

ottenute in corrispondenza alle $(n - m)$ -uple

$$(t_{m+1}, t_{m+2}, \dots, t_n) = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

rispettivamente. Per ogni $t_{m+1}, t_{m+2}, \dots, t_n \in \mathbb{K}$, si ha

$$\begin{aligned}
t_{m+1}S_{m+1} + t_{m+2}S_{m+2} + \dots + t_n S_n &= t_{m+1} \begin{pmatrix} S_{m+11} \\ S_{m+12} \\ \vdots \\ S_{m+1m} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + t_{m+2} \begin{pmatrix} S_{m+21} \\ S_{m+22} \\ \vdots \\ S_{m+2m} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + t_n \begin{pmatrix} S_{n1} \\ S_{n2} \\ \vdots \\ S_{nm} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \\
&= \begin{pmatrix} t_{m+1}S_{m+11} \\ t_{m+1}S_{m+12} \\ \vdots \\ t_{m+1}S_{m+1m} \\ t_{m+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} t_{m+2}S_{m+21} \\ t_{m+2}S_{m+22} \\ \vdots \\ t_{m+2}S_{m+2m} \\ 0 \\ t_{m+2} \\ \vdots \\ 0 \end{pmatrix} + \dots + \begin{pmatrix} t_n S_{n1} \\ t_n S_{n2} \\ \vdots \\ t_n S_{nm} \\ 0 \\ 0 \\ \vdots \\ t_n \end{pmatrix} = \begin{pmatrix} t_{m+1}S_{m+11} + \dots + t_n S_{n1} \\ t_{m+1}S_{m+12} + \dots + t_n S_{n2} \\ \vdots \\ t_{m+1}S_{m+1m} + \dots + t_n S_{nm} \\ t_{m+1} \\ t_{m+2} \\ \vdots \\ t_n \end{pmatrix}
\end{aligned} \tag{19}$$

Il secondo membro è $(0, \dots, 0)$ se e solo se

$$t_{m+1} = t_{m+2} = \dots = t_n = 0.$$

Quindi $S_{m+1}, S_{m+2}, \dots, S_n$ sono linearmente indipendenti. Il secondo membro della (19) è la soluzione del sistema (18) che corrisponde alla scelta dei valori $t_{m+1}, t_{m+2}, \dots, t_n$ per le incognite X_{m+1}, \dots, X_n ; la (19) mostra che tale soluzione è combinazione lineare di $S_{m+1}, S_{m+2}, \dots, S_n$. Quindi s_{m+1}, \dots, s_n generano lo spazio Σ_0 delle soluzioni del sistema (18). Segue che

$$\{s_{m+1}, s_{m+2}, \dots, s_n\}$$

è una base di Σ_0 . In particolare, $\dim(\Sigma_0) = n - m$. Pertanto l'infinità delle soluzioni del sistema omogeneo a gradini (18), coincide con la dimensione dello spazio delle soluzioni. Più in generale, poiché ogni sistema di equazioni lineari omogenee (17) è equivalente a un sistema a gradini, abbiamo che l'infinità delle soluzioni di un sistema di equazioni lineari omogenee è uguale alla dimensione dello spazio delle sue soluzioni.

Proposizione 2.26. *Lo spazio vettoriale $M_{m,n}(\mathbb{K})$ delle matrici $m \times n$ a elementi in \mathbb{K} ha dimensione mn .*

Proof. Consideriamo infatti, per ogni $1 \leq i \leq m$, $1 \leq j \leq n$, la matrice 1_{ij} che ha 1 nel posto i, j e 0 altrove; Otteniamo così un insieme di nm matrici $\{1_{11}, 1_{12}, \dots, 1_{mn}\}$, che costituisce una base di $M_{m,n}(\mathbb{K})$. Infatti, se $A = (a_{ij}) \in M_{m,n}(\mathbb{K})$, allora si ha

$$A = a_{11}1_{11} + a_{12}1_{12} + \dots + a_{mn}1_{mn}.$$

□

Esempio 2.13. Consideriamo una matrice 2×3 A :

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

Le matrici 1_{ij} per una matrice 2×3 sono:

$$\begin{aligned}
1_{11} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & 1_{12} &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & 1_{13} &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \\
1_{21} &= \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, & 1_{22} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, & 1_{23} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{aligned}$$

La matrice A può essere espressa come combinazione lineare delle matrici 1_{ij} :

$$A = a_{11}1_{11} + a_{12}1_{12} + a_{13}1_{13} + a_{21}1_{21} + a_{22}1_{22} + a_{23}1_{23} = \sum_{i=1}^m \sum_{j=1}^n a_{ij}1_{ij}.$$

Esplicitamente:

$$A = a_{11} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + a_{12} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + a_{13} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} + a_{21} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + a_{22} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} + a_{23} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Questa combinazione lineare produce esattamente la matrice A :

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

Proposizione 2.27. *Supponiamo che $\dim(\mathbf{V}) = n$.*

1. *Se $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbf{V}$ sono linearmente indipendenti, $\mathbf{v}_1, \dots, \mathbf{v}_n$ è una base di \mathbf{V} .*
2. *Se $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbf{V}$ sono linearmente indipendenti, esistono $\mathbf{v}_{n+1}, \dots, \mathbf{v}_n$ tali che $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ sia una base di \mathbf{V} .*

Proof.

1. È sufficiente mostrare che $\langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle = \mathbf{V}$. Per ipotesi esiste una base $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ di \mathbf{V} , e ciò implica, per il teorema (2.25), che per ogni $\mathbf{v} \in \mathbf{V}$ i vettori $\mathbf{v}_1, \dots, \mathbf{v}_n, \mathbf{v}$ sono linearmente dipendenti. Pertanto esistono $a_1, \dots, a_n \in \mathbb{K}$ non tutti nulli tali che

$$a_1 \mathbf{v}_1 + \dots + a_n \mathbf{v}_n + a \mathbf{v} = \mathbf{0}.$$

Poiché $\mathbf{v}_1, \dots, \mathbf{v}_n$ sono linearmente indipendenti, deve essere $a \neq 0$. Quindi si ha

$$\mathbf{v} = -a^{-1}a_1 \mathbf{v}_1 - \dots - a^{-1}a_n \mathbf{v}_n,$$

cioè $\mathbf{v} \in \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$. Essendo $\mathbf{v} \in \mathbf{V}$ arbitrario, si ha l'asserto.

2. Per il teorema (2.25) deve essere $k \leq n$. Se $k = n$ la conclusione segue dalla parte (1); supponiamo dunque $k < n$. $\mathbf{v}_1, \dots, \mathbf{v}_k$ non possono generare \mathbf{V} , altrimenti si avrebbe una base costituita da $k \neq n$ elementi, il che contraddirebbe il corollario (1). Pertanto possiamo trovare $\mathbf{v}_{k+1} \in \mathbf{V} \setminus \langle \mathbf{v}_1, \dots, \mathbf{v}_k \rangle$. Supponiamo che $a_1, \dots, a_k, a_{k+1} \in \mathbb{K}$ siano tali che

$$a_1 \mathbf{v}_1 + \dots + a_k \mathbf{v}_k + a_{k+1} \mathbf{v}_{k+1} = \mathbf{0}.$$

Deve essere $a_{k+1} = 0$, perché se fosse $a_{k+1} \neq 0$ si avrebbe

$$\mathbf{v}_{k+1} = -a_{k+1}^{-1}a_1 \mathbf{v}_1 - \dots - a_{k+1}^{-1}a_k \mathbf{v}_k,$$

cioè $\mathbf{v}_{k+1} \in \langle \mathbf{v}_1, \dots, \mathbf{v}_k \rangle$, il che è contro l'ipotesi; pertanto si deve anche avere $a_1 = \dots = a_k = 0$, perché $\mathbf{v}_1, \dots, \mathbf{v}_k$ sono linearmente indipendenti. Dunque i vettori $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}$ sono linearmente indipendenti.

Se $k+1 = n$ si conclude, per la parte (1), che $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ è una base, e l'asserito è provato; se invece $k+1 < n$ possiamo ripetere il ragionamento precedente e trovare $\mathbf{v}_{k+2} \in \mathbf{V}$ tale che $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \mathbf{v}_{k+2}\}$ siano linearmente indipendenti. Iterando questo procedimento $n - k$ volte è possibile trovare $\mathbf{v}_{k+1}, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n \in \mathbf{V}$ tali che $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ siano linearmente indipendenti. Per la (1), $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ è una base, e l'asserto è provato. □

Corollario 2. *Se $\dim(\mathbf{V}) = n$, ogni sottospazio vettoriale \mathbf{W} di \mathbf{V} ha dimensione finita non superiore a n .*

Proof. Per il teorema (2.25), per ogni sottoinsieme finito $\{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ di \mathbf{W} costituito da vettori linearmente indipendenti deve essere $m \leq n$. Quindi, se \mathbf{W} ha dimensione finita, ogni sua base deve essere costituita da non più di n elementi, cioè $\dim(\mathbf{W}) \leq n$. D'altra parte, se \mathbf{W} non avesse dimensione finita, per ogni sottoinsieme finito $\{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ di vettori linearmente indipendenti di \mathbf{W} si avrebbe $\langle \mathbf{w}_1, \dots, \mathbf{w}_m \rangle \neq \mathbf{W}$ e quindi esisterebbe $\mathbf{w}_{m+1} \in \mathbf{W}$ tale che $\{\mathbf{w}_1, \dots, \mathbf{w}_m, \mathbf{w}_{m+1}\}$ siano linearmente indipendenti. Ciò implica l'esistenza di insiemi finiti di vettori linearmente indipendenti di \mathbf{V} con un numero arbitrariamente alto di elementi, il che è impossibile per il teorema (2.25). □

Nota 8. Se in particolare $\dim(\mathbf{W}) = 0$ oppure n , allora $\mathbf{W} = \langle \mathbf{0} \rangle$ oppure $\mathbf{W} = \mathbf{V}$ rispettivamente.

Codimensione

Definizione 2.26. Per ogni sottospazio vettoriale \mathbf{W} di \mathbf{V} , il numero

$$\dim(\mathbf{V}) - \dim(\mathbf{W})$$

si dice **codimensione** di \mathbf{W} in \mathbf{V} .

Teorema 2.28. Siano \mathbf{U} e \mathbf{W} due sottospazi di dimensione finita dello spazio vettoriale \mathbf{V} . Allora $\mathbf{U} \cap \mathbf{W}$ e $\mathbf{U} + \mathbf{W}$ hanno dimensione finita e

$$\dim(\mathbf{U} + \mathbf{W}) = \dim(\mathbf{U}) + \dim(\mathbf{W}) - \dim(\mathbf{U} \cap \mathbf{W}). \quad (20)$$

In particolare, $\mathbf{U} + \mathbf{W}$ è somma diretta di \mathbf{U} e \mathbf{W} se e solo se

$$\dim(\mathbf{U} + \mathbf{W}) = \dim(\mathbf{U}) + \dim(\mathbf{W}).$$

Proof. $\mathbf{U} \cap \mathbf{W}$ è un sottospazio di \mathbf{U} , che ha dimensione finita, e quindi anche $\mathbf{U} \cap \mathbf{W}$ ha dimensione finita. Sia dunque $\{\mathbf{z}_1, \dots, \mathbf{z}_q\}$ una base di $\mathbf{U} \cap \mathbf{W}$. Per la proposizione (2.27.2) esistono $\mathbf{u}_1, \dots, \mathbf{u}_t \in \mathbf{U}$ e $\mathbf{w}_1, \dots, \mathbf{w}_s \in \mathbf{W}$ tali che $\{\mathbf{z}_1, \dots, \mathbf{z}_q, \mathbf{u}_1, \dots, \mathbf{u}_t\}$ sia una base di \mathbf{U} e $\{\mathbf{z}_1, \dots, \mathbf{z}_q, \mathbf{w}_1, \dots, \mathbf{w}_s\}$ sia una base di \mathbf{W} . Poiché

$$\dim(\mathbf{U}) + \dim(\mathbf{W}) - \dim(\mathbf{U} \cap \mathbf{W}) = q + t + s,$$

per dimostrare la prima parte del teorema sarà sufficiente dimostrare che

$$\{\mathbf{z}_1, \dots, \mathbf{z}_q, \mathbf{u}_1, \dots, \mathbf{u}_t, \mathbf{w}_1, \dots, \mathbf{w}_s\}$$

è una base di $\mathbf{U} + \mathbf{W}$. Sia $\mathbf{u} + \mathbf{w} \in \mathbf{U} + \mathbf{W}$. Esistono scalari $a_1, \dots, a_q, a'_1, \dots, a'_q, b_1, \dots, b_t, c_1, \dots, c_s$, tali che

$$\begin{aligned} \mathbf{u} &= a_1 \mathbf{z}_1 + \dots + a_q \mathbf{z}_q + b_1 \mathbf{u}_1 + \dots + b_t \mathbf{u}_t \\ \mathbf{w} &= a'_1 \mathbf{z}_1 + \dots + a'_q \mathbf{z}_q + c_1 \mathbf{w}_1 + \dots + c_s \mathbf{w}_s \end{aligned}$$

Si ha quindi

$$\mathbf{u} + \mathbf{w} = (a_1 + a'_1) \mathbf{z}_1 + \dots + (a_q + a'_q) \mathbf{z}_q + b_1 \mathbf{u}_1 + \dots + b_t \mathbf{u}_t + c_1 \mathbf{w}_1 + \dots + c_s \mathbf{w}_s,$$

Dunque $\mathbf{z}_1, \dots, \mathbf{z}_q, \mathbf{u}_1, \dots, \mathbf{u}_t, \mathbf{w}_1, \dots, \mathbf{w}_s$, generano $\mathbf{U} + \mathbf{W}$. Supponiamo ora che $a_1, \dots, a_q, b_1, \dots, b_t, c_1, \dots, c_s \in \mathbb{K}$ siano tali che

$$a_1 \mathbf{z}_1 + \dots + a_q \mathbf{z}_q + b_1 \mathbf{u}_1 + \dots + b_t \mathbf{u}_t + c_1 \mathbf{w}_1 + \dots + c_s \mathbf{w}_s = \mathbf{0}. \quad (21)$$

Dalla (21) segue che

$$c_1 \mathbf{w}_1 + \dots + c_s \mathbf{w}_s = -(a_1 \mathbf{z}_1 + \dots + a_q \mathbf{z}_q + b_1 \mathbf{u}_1 + \dots + b_t \mathbf{u}_t). \quad (22)$$

Poiché il primo membro della (22) appartiene a \mathbf{W} , il secondo membro sta in $\mathbf{U} \cap \mathbf{W}$. Essendo $\{\mathbf{z}_1, \dots, \mathbf{z}_q\}$ una base di $\mathbf{U} \cap \mathbf{W}$, si ha

$$a_1 \mathbf{z}_1 + \dots + a_q \mathbf{z}_q + b_1 \mathbf{u}_1 + \dots + b_t \mathbf{u}_t = e_1 \mathbf{z}_1 + \dots + e_q \mathbf{z}_q,$$

per opportuni $e_1, \dots, e_q \in \mathbb{K}$, o equivalentemente,

$$(a_1 - e_1) \mathbf{z}_1 + \dots + (a_q - e_q) \mathbf{z}_q + b_1 \mathbf{u}_1 + \dots + b_t \mathbf{u}_t = \mathbf{0}.$$

Per l'indipendenza lineare di $\mathbf{z}_1, \dots, \mathbf{z}_q, \mathbf{u}_1, \dots, \mathbf{u}_t$, tutti i coefficienti sono 0, e in particolare $b_1 = \dots = b_t = 0$. Per la (21) si ha dunque

$$a_1 \mathbf{z}_1 + \dots + a_q \mathbf{z}_q + c_1 \mathbf{w}_1 + \dots + c_s \mathbf{w}_s = \mathbf{0}.$$

Dall'indipendenza lineare di $\mathbf{z}_1, \dots, \mathbf{z}_q, \mathbf{w}_1, \dots, \mathbf{w}_s$, segue che anche

$$a_1 = \dots = a_q = c_1 = \dots = c_s = 0.$$

Quindi $\mathbf{z}_1, \dots, \mathbf{z}_q, \mathbf{u}_1, \dots, \mathbf{u}_t, \mathbf{w}_1, \dots, \mathbf{w}_s$ sono linearmente indipendenti. L'ultima asserzione del teorema segue immediatamente dalla (20) e dalla definizione di somma diretta. \square

La (20) è detta **formula di Grassmann vettoriale**.

2.5 Rango

Rango

Definizione 2.27. Siano \mathbf{V} un \mathbb{K} -spazio vettoriale e $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ un sottoinsieme finito di \mathbf{V} , il **rango** di $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ è il massimo numero di vettori linearmente indipendenti appartenenti a $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$. Equivalentemente, il rango di $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ è la dimensione del sottospazio vettoriale $\langle \mathbf{v}_1, \dots, \mathbf{v}_m \rangle$.

Rango per le matrici

Definizione 2.28. Sia $A \in M_{m,n}(\mathbb{K})$. Il **rango per righe** di A è il rango dell'insieme delle sue righe, cioè il massimo numero di righe linearmente indipendenti di A , considerate come vettori di \mathbb{K}^n . Analogamente si definisce il **rango per colonne** di A .

Teorema 2.29. Il rango per righe e il rango per colonne di una matrice $A \in M_{m,n}(\mathbb{K})$ coincidono.

Proof. Siano r il rango per righe e c il rango per colonne di A . Se $r = 0$ allora tutti gli elementi di A sono nulli e quindi anche $c = 0$. Pertanto possiamo supporre $r > 0$. Una relazione di dipendenza lineare tra le colonne di A è una soluzione non banale del sistema omogeneo

$$A\mathbf{X} = \mathbf{0} \quad (23)$$

dove $\mathbf{X} = {}^t(X_1, \dots, X_n)$. Pertanto il rango per colonne di A è individuato dall'insieme delle soluzioni del sistema (23). Se le righe di A si dispongono in ordine diverso, c non cambia perché un cambiamento nell'ordine delle equazioni del sistema, cioè una successione di operazioni elementari (I), non influisce sull'insieme delle sue soluzioni. Neanche r cambia se si riordinano le righe di A , perché il rango di un insieme di vettori non dipende dal loro ordine. Pertanto non è restrittivo supporre che le prime r righe di A siano linearmente indipendenti. La matrice

$$A^* = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rn} \end{pmatrix}$$

ha rango per righe uguale a r . Consideriamo il sistema

$$A^*\mathbf{X} = \mathbf{0} \quad (24)$$

e sia (x, \dots, x_n) una sua soluzione. Se $r < m$, allora per ogni $i = r + 1, \dots, m$ l' i -esima riga di A è combinazione lineare delle righe di A^* , e quindi si ha

$$a_{i1}x_1 + \dots + a_{in}x_n = \alpha_1(a_{11}x_1 + \dots + a_{1n}x_n) + \dots + \alpha_r(a_{r1}x_1 + \dots + a_{rn}x_n) = 0 + \dots + 0 = 0$$

per opportuni $\alpha_1, \dots, \alpha_r \in \mathbb{K}$. Pertanto (x_1, \dots, x_n) è anche soluzione di (23). Viceversa è ovvio che ogni soluzione di (23) è soluzione di (24), che è un suo sotto-sistema. In conclusione (23) e (24) sono equivalenti, e quindi il rango per colonne di A^* è uguale a c . Poiché le colonne di A^* sono vettori di \mathbb{K}^r , si ha $c \leq r$. Ragionando allo stesso modo su tA si deduce che anche $r \leq c$, e l'asserto è provato. \square

Il valore comune del rango per righe e del rango per colonne di A si dice rango di A e si denota con $r(A)$. È evidente che

$$\forall A \in M_{m,n}(\mathbb{K}), r(A) \leq \min\{m, n\}.$$

Proposizione 2.30. Sia $A \in M_{m,n}(\mathbb{K})$. Se $B \in M_{m,n}(\mathbb{K})$ è ottenuta da A mediante una successione di operazioni elementari sulle righe, allora

$$r(A) = r(B).$$

Proof. È sufficiente dimostrare l'asserto nel caso in cui B è ottenuta da A mediante una sola operazione elementare sulle righe.

1. Se B è ottenuta da A per mezzo di un'operazione elementare (I), le sue righe sono

$$A^{(1)}, \dots, A^{(t)}, \dots, A^{(s)}, \dots, A^{(m)}$$

per qualche $1 \leq s < t \leq m$, e quindi si ha

$$\langle A^{(1)}, \dots, A^{(t)}, \dots, A^{(s)}, \dots, A^{(m)} \rangle = \langle A^{(1)}, \dots, A^{(m)} \rangle,$$

cioè $r(A) = r(B)$, perché il rango di un insieme di vettori non dipende dall'ordine in cui si considerano.

2. Se si è effettuata un'operazione elementare (II), le righe di B sono

$$A^{(1)}, \dots, cA^{(t)}, \dots, A^{(m)}$$

per qualche $c \in \mathbb{K}^*$ e $1 \leq t \leq m$, e si ha evidentemente

$$\langle A^{(1)}, \dots, cA^{(t)}, \dots, A^{(m)} \rangle = \langle A^{(1)}, \dots, A^{(m)} \rangle,$$

cioè ancora $r(A) = r(B)$.

3. Se l'operazione elementare è del tipo (III), le righe di B sono

$$A^{(1)}, \dots, A^{(s-1)}, A^{(s)} + cA^{(t)}, A^{(s+1)}, \dots, A^{(m)},$$

per qualche $c \in \mathbb{K}$ e $1 \leq s, t \leq m$, $s \neq t$; poiché tali righe sono combinazioni lineari di $A^{(1)}, \dots, A^{(m)}$, si ha

$$\langle A^{(1)}, \dots, A^{(s-1)}, A^{(s)} + cA^{(t)}, A^{(s+1)}, \dots, A^{(m)} \rangle \subset \langle A^{(1)}, \dots, A^{(m)} \rangle.$$

D'altra parte

$$A^{(s)} = (A^{(s)} + cA^{(t)}) - cA^{(t)}$$

e quindi è vera anche l'inclusione opposta:

$$\langle A^{(1)}, \dots, A^{(s-1)}, A^{(s)} + cA^{(t)}, A^{(s+1)}, \dots, A^{(m)} \rangle \supset \langle A^{(1)}, \dots, A^{(m)} \rangle,$$

e di nuovo $r(A) = r(B)$.

□

Proposizione 2.31.

1. Se A e B sono due matrici che possono essere moltiplicate, allora

$$r(AB) \leq \min\{r(A), r(B)\}.$$

2. Se $A \in \text{GL}_m(\mathbb{K})$, $B \in M_{m,n}(\mathbb{K})$, $C \in \text{GL}_n(\mathbb{K})$ allora

$$r(AB) = r(B) = r(BC).$$

Proof.

1. Siano $A = (a_{ij}) \in M_{m,n}(\mathbb{K})$ e $B = (b_{jh}) \in M_{n,s}(\mathbb{K})$. Per ogni $i = 1, \dots, m$ la i -esima riga di AB è

$$\begin{aligned} A^{(i)}B &= (A^{(i)}B_{(1)} \ A^{(i)}B_{(2)} \ \dots \ A^{(i)}B_{(s)}) = \left(\sum_{j=1}^n a_{ij}b_{j1} \ \sum_{j=1}^n a_{ij}b_{j2} \ \dots \ \sum_{j=1}^n a_{ij}b_{js} \right) = \sum_{j=1}^n (a_{ij}b_{j1} \ a_{ij}b_{j2} \ \dots \ a_{ij}b_{js}) \\ &= \sum_{j=1}^n a_{ij}(b_{j1} \ b_{j2} \ \dots \ b_{js}) = \sum_{j=1}^n a_{ij}B^{(j)}. \end{aligned}$$

Quindi, essendo ogni $A^{(i)}B$ combinazione lineare delle righe di B , si ha

$$\langle A^{(1)}B, A^{(2)}B, \dots, A^{(m)}B \rangle \subset \langle B^{(1)}, B^{(2)}, \dots, B^{(n)} \rangle$$

e pertanto $r(AB) \leq r(B)$.

D'altra parte si ha anche

$$r(AB) = r({}^t(AB)) = r({}^tB {}^tA) \leq r({}^tA) = r(A)$$

e la (1) è dimostrata.

2. Per la (1) si ha

$$r(AB) \leq r(B) = r(A^{-1}(AB)) \leq r(AB)$$

e quindi $r(AB) = r(B)$. La seconda uguaglianza si dimostra in modo simile.

$$r(BC) \leq r(B) = r((BC)C^{-1}) \leq r(BC).$$

□

Teorema 2.32. Una matrice quadrata di ordine n è invertibile se e solo se ha rango n .

Proof. Per la proposizione (2.31.2) una matrice invertibile A ha lo stesso rango di $\mathbf{I}_n = A^{-1}A$. Il rango di \mathbf{I}_n è n , perché le sue righe costituiscono la base canonica $\{\mathbf{E}_1, \dots, \mathbf{E}_n\}$ di \mathbb{K}^n . Viceversa, se A ha rango n , le sue righe $A^{(1)}, \dots, A^{(n)}$ costituiscono una base di \mathbb{K}^n . Quindi, per ogni $i = 1, \dots, n$ esistono $b_{i1}, \dots, b_{in} \in \mathbb{K}$ tali che

$$\mathbf{E}_i = b_{i1}A^{(1)} + b_{i2}A^{(2)} + \dots + b_{in}A^{(n)}. \quad (25)$$

Consideriamo la matrice $B = (b_{ij}) \in M_n(\mathbb{K})$. La (25) equivale all'identità

$$\mathbf{I}_n = BA$$

e quindi A è invertibile. □

Sottomatrice

Definizione 2.29. Una sottomatrice $p \times q$ di una matrice $A \in M_{m,n}(\mathbb{K})$ è una matrice costituita dagli elementi di A comuni a p righe e q colonne fissate in A . Scelti indici

$$1 \leq i_1 < i_2 < \dots < i_p \leq m \quad \text{e} \quad 1 \leq j_1 < j_2 < \dots < j_q \leq n,$$

la sottomatrice di A corrispondente alle righe i_1 -esima, i_2 -esima, ..., i_p -esima, e alle colonne j_1 -esima, j_2 -esima, ..., j_q -esima, si denota con $A(i_1 \ i_2 \ \dots \ i_p \mid j_1 \ j_2 \ \dots \ j_q)$.

Proposizione 2.33. Se B è una sottomatrice della matrice A , allora $r(B) \leq r(A)$.

Proof. Sia $A \in M_{m,n}(\mathbb{K})$ e $B = A(i_1 \ i_2 \ \dots \ i_p \mid j_1 \ j_2 \ \dots \ j_q)$. Consideriamo la sottomatrice di A

$$C = A(i_1 \ i_2 \ \dots \ i_p \mid 1 \ 2 \ \dots \ n) \in M_{p,n}(\mathbb{K})$$

Costituita dalla i_1 -esima, i_2 -esima, ..., i_p -esima riga. La disuguaglianza

$$r(C) \leq r(A) \quad (26)$$

è ovvia se interpretata come una relazione tra i ranghi per righe delle due matrici. D'altra parte B è una sottomatrice di C , e precisamente

$$B = C(1 \ 2 \ \dots \ p \mid j_1 \ j_2 \ \dots \ j_q),$$

la sottomatrice costituita dalla j_1 -esima, j_2 -esima, ..., j_q -esima colonna di C . Anche in questo caso la disuguaglianza

$$r(B) \leq r(C) \quad (27)$$

è ovvia se interpretata come relazione tra i ranghi per colonne delle due matrici. La (26) e la (27) insieme implicano $r(B) \leq r(A)$. □

Teorema 2.34. Il rango di una matrice A è uguale al massimo degli ordini delle sue sottomatrici quadrate invertibili.

Proof. Sia p il massimo degli ordini delle sottomatrici quadrate invertibili di A . Dalle proposizioni (2.32) e (2.33) segue che $p \leq r(A)$. D'altra parte, posto $r = r(A)$, se le righe $A^{(i_1)}, A^{(i_2)}, \dots, A^{(i_r)}$ di A sono linearmente indipendenti, la sottomatrice

$$B = A(i_1 \dots i_r \mid 1 \dots n)$$

ha rango r , e quindi possiede r colonne linearmente indipendenti $B_{(j_1)}, B_{(j_2)}, \dots, B_{(j_r)}$. La sottomatrice quadrata

$$B(1 \dots r \mid j_1 \dots j_r)$$

di B ha rango r , cioè è invertibile. Poiché

$$B(1 \dots r \mid j_1 \dots j_r) = A(i_1 \dots i_r \mid j_1 \dots j_r)$$

è una sottomatrice di A , si ha anche $p \geq r(A)$. □

Dalla dimostrazione del teorema (2.34) segue che se $B = A(i_1, \dots, i_n, j_1, \dots, j_m)$ è una sottomatrice quadrata invertibile di A , allora le righe $A^{(i_1)}, A^{(i_2)}, \dots, A^{(i_n)}$ di A sono linearmente indipendenti. Similmente, le colonne $A_{(j_1)}, A_{(j_2)}, \dots, A_{(j_m)}$ sono linearmente indipendenti.

Kronecker-Rouché-Capelli

Teorema 2.35. *Un sistema di m equazioni in n incognite*

$$A\mathbf{X} = \mathbf{b}, \quad \text{con } A \in M_{m,n}(\mathbb{K}), \mathbf{b} \in M_{m,1}(\mathbb{K}), \mathbf{X} = {}^t(X_1 \dots X_n), \quad (28)$$

è compatibile se e solo se $r(A) = r(A\mathbf{b})$. In tal caso il sistema (28) possiede ∞^{n-r} soluzioni, dove $r = r(A)$.

Proof. Sia

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Una n -upla $(x_1, \dots, x_n) \in \mathbb{K}^n$ è soluzione di (28) se e solo se si ha

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}. \quad (29)$$

La (29) esprime la condizione che il vettore \mathbf{b} sia combinazione lineare delle colonne di A . Questa condizione è verificata se e solo se la matrice $(A\mathbf{b})$ ha lo stesso rango per colonne di A , cioè se e solo se $r(A) = r(A\mathbf{b})$. La prima parte del teorema è dimostrata.

Se il sistema (28) è compatibile, e $r = r(A)$, possiamo supporre che le sue prime r equazioni siano linearmente indipendenti e sostituire (28) con il sistema equivalente

$$\begin{aligned} a_{11}X_1 + a_{12}X_2 + \dots + a_{1n}X_n &= b_1 \\ a_{21}X_1 + a_{22}X_2 + \dots + a_{2n}X_n &= b_2 \\ &\vdots \\ a_{m1}X_1 + a_{m2}X_2 + \dots + a_{mn}X_n &= b_m \end{aligned} \quad (30)$$

Applicando il procedimento di Gauss-Jordan al sistema (30), nessuna delle equazioni si riduce $0 = 0$, perchè ciò implicherebbe che essa è linearmente dipendente da quelle che la precedono. Quindi il sistema (30) si può trasformare in un sistema a gradini di r equazioni; ciò significa che (30), e quindi (28), possiede ∞^{n-r} soluzioni. □

2.6 Determinanti

Determinante

Definizione 2.30. Sia $n \geq 1$ e

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \in M_n(\mathbb{K}).$$

Il **determinante** di A è l'elemento di \mathbb{K}

$$\det(A) = \sum_{p \in \sigma_n} \varepsilon(p) \prod_{i=1}^n a_{ip(i)}, \quad (31)$$

dove $\sigma(n)$ denota l'insieme di tutte le permutazioni di $\{1, 2, \dots, n\}$ e dove $\varepsilon(p)$ è il segno della permutazione $p \in \sigma_n$; $\det(A)$ verrà anche indicato con $|A|$.

La (31) è una somma di $n!$ termini, che, a meno del segno, sono tutti i possibili prodotti di n elementi di A appartenenti a righe ed a colonne diverse.

$n = 1$ $A = (a)$ e si ha $\det(A) = a$;

$n = 2$

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

allora

$$\det(A) = a_{11}a_{22} - a_{12}a_{21};$$

$n = 3$

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

allora

$$\det(A) = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}.$$

Se $A \in M_n(\mathbb{K})$, denoteremo con $A^{(1)}, \dots, A^{(n)}$ le sue righe e con $A_{(1)}, \dots, A_{(n)}$ le sue colonne; scriveremo, usando la notazione a blocchi,

$$A = (A_{(1)} \dots A_{(n)})$$

oppure

$$A = \begin{pmatrix} A^{(1)} \\ \vdots \\ A^{(n)} \end{pmatrix}.$$

Scriveremo quindi

$$\det(A) = \det(A_{(1)} \dots A_{(n)}) = \det \begin{pmatrix} A^{(1)} \\ \vdots \\ A^{(n)} \end{pmatrix}.$$

Teorema 2.36. Sia $n \geq 1$ un intero e sia

$$A = (a_{ij}) = (A_{(1)} \dots A_{(n)}) = \begin{pmatrix} A^{(1)} \\ \vdots \\ A^{(n)} \end{pmatrix} \in M_n(\mathbb{K}).$$

Allora:

1. $\det({}^t A) = \det(A)$

2. Se $A^{(i)} = cV + c'V'$, per qualche $1 \leq i \leq n$, $c, c' \in \mathbb{K}$ cioè se la i -esima riga di A è combinazione lineare di due vettori n -riga, allora

$$\det(A) = c \det \begin{pmatrix} A^{(1)} \\ \vdots \\ V \\ \vdots \\ A^{(n)} \end{pmatrix} + c' \det \begin{pmatrix} A^{(1)} \\ \vdots \\ V' \\ \vdots \\ A^{(n)} \end{pmatrix}$$

Analogamente, se $A_{(i)} = cW + c'W'$ per qualche $1 \leq i \leq n$, $c, c' \in \mathbb{K}$, dove W e W' sono due n -vettori colonna, si ha

$$\det(A) = c \det(A_{(1)} \dots W \dots A_{(n)}) + c' \det(A_{(1)} \dots W' \dots A_{(n)}).$$

3. Se la matrice $B \in M_n(\mathbb{K})$ è ottenuta da A scambiando tra loro due righe oppure due colonne, si ha $\det(B) = -\det(A)$.
4. Se A ha due righe oppure due colonne uguali, allora $\det(A) = 0$.
5. $\det(\mathbf{I}_n) = 1$.

Proof.

1. Si ha

$$\det({}^t A) = \sum_{p \in \sigma_n} \varepsilon(p) \prod_{i=1}^n a_{p(i) i}. \quad (32)$$

A meno del segno, gli addendi di (32) sono gli stessi di (31); infatti

$$\prod_{i=1}^n a_{p(i) i} = \prod_{i=1}^n a_{i q(i)}$$

con $q = p^{-1} \in \sigma_n$. Osservando poi che $\varepsilon(\sigma^{-1}) = \varepsilon(p)$, si conclude che gli addendi di $\det(A)$ e di $\det({}^t A)$ sono gli stessi, cioè $\det(A) = \det({}^t A)$.

2. Le due affermazioni sono equivalenti per la (1), e quindi è sufficiente dimostrare la prima. Siano

$$V = (v_1 \dots v_n), \quad V' = (v'_1 \dots v'_n).$$

Si ha

$$(a_{i1} \dots a_{in}) = (cv_1 + c'v'_1 \quad \dots \quad cv_n + c'v'_n),$$

e

$$\begin{aligned} \det(A) &= \sum_{p \in \sigma_n} \varepsilon(p) a_{1p(1)} \dots a_{np(n)} = \sum_{p \in \sigma_n} \varepsilon(p) a_{1p(1)} \dots (cv_{p(i)} + c'v'_{p(i)}) \dots a_{np(n)} \\ &= c \sum_{p \in \sigma_n} \varepsilon(p) a_{1p(1)} \dots v_{p(i)} \dots a_{np(n)} + c' \sum_{p \in \sigma_n} \varepsilon(p) a_{1p(1)} \dots v'_{p(i)} \dots a_{np(n)} \\ &= c \det \begin{pmatrix} A^{(1)} \\ \vdots \\ V \\ \vdots \\ A^{(n)} \end{pmatrix} + c' \det \begin{pmatrix} A^{(1)} \\ \vdots \\ V' \\ \vdots \\ A^{(n)} \end{pmatrix}. \end{aligned}$$

3. Per la (1) è sufficiente dimostrare la (3) nel caso in cui B sia ottenuta da A scambiando due righe, siano esse la i -esima e la j -esima, dove $1 \leq i < j \leq n$. Posto $B = (b_{hk})$, si ha

$$\begin{aligned} \det(B) &= \sum_{p \in \sigma_n} \varepsilon(p) b_{1p(1)} \dots b_{ip(i)} \dots b_{jp(j)} \dots b_{np(n)}, \\ &= \sum_{p \in \sigma_n} \varepsilon(p) a_{1p(1)} \dots a_{jp(i)} \dots a_{ip(j)} \dots a_{np(n)}, \\ &= \sum_{p \in \sigma_n} \varepsilon(p) a_{1(tp)(1)} \dots a_{i(tp)(i)} \dots a_{j(tp)(j)} \dots a_{n(tp)(n)}, \end{aligned}$$

dove abbiamo denotato con t la trasposizione che scambia i con j . Poichè $\varepsilon(p) = -\varepsilon(t \circ p)$ e poichè al variare di $p \in \sigma_n$, $t \circ p$ descrive tutto σ_n , si deduce:

$$\det(B) = \sum_{q \in \sigma_n} -\varepsilon(q) a_{1q(1)} \dots a_{iq(i)} \dots a_{jq(j)} \dots a_{nq(n)} = -\det(A).$$

4. Supponiamo che A abbia due righe uguali. Scambiando tra loro tali righe si ottiene ancora la matrice A . Per la (3) si ha quindi $\det(A) = -\det(A)$, sicchè $\det(A) = 0$.
5. L'unico addendo di $\det(\mathbf{I}_n)$ diverso da 0 è $a_{11} a_{22} \dots a_{nn} = 1 \cdot 1 \cdot \dots \cdot 1 = 1$.

□

Corollario 3. Se $A, B \in M_n(\mathbb{K})$, allora $\det(AB) = \det(A) \det(B)$. In particolare, se A è invertibile, allora $\det(A^{-1}) = \det(A)^{-1}$.

Proof. Siano $A = (a_{ij})$ e $B = \begin{pmatrix} B^{(1)} \\ \vdots \\ B^{(n)} \end{pmatrix}$. Si ha, utilizzando la (2) e la (3) del teorema (2.36):

$$\begin{aligned} \det(AB) &= \det \begin{pmatrix} \sum_{i=1}^n a_{1i} B^{(i)} \\ \vdots \\ \sum_{i=1}^n a_{ni} B^{(i)} \end{pmatrix} = \sum_{p \in \sigma_n} \det \begin{pmatrix} a_{1p(1)} B^{(p(1))} \\ \vdots \\ a_{np(n)} B^{(p(n))} \end{pmatrix} = \sum_{p \in \sigma_n} a_{1p(1)} \dots a_{np(n)} \det \begin{pmatrix} B^{(p(1))} \\ \vdots \\ B^{(p(n))} \end{pmatrix} = \\ &= \sum_{p \in \sigma_n} \varepsilon(p) a_{1p(1)} \dots a_{np(n)} \det \begin{pmatrix} B^{(1)} \\ \vdots \\ B^{(n)} \end{pmatrix} = \det(A) \det(B) \end{aligned}$$

L'ultima affermazione segue immediatamente dalla prima applicata al prodotto $\mathbf{I}_n = AA^{-1}$, tenuto conto del teorema (2.36.5). □

Teorema 2.37. Sia $A \in M_n(\mathbb{K})$. Allora $\det(A) \neq 0 \iff r(A) = n$.

Proof. Se $r(A) = n$, allora A è invertibile per il teorema (2.32). Dal corollario (3) segue che dev'essere $\det(A^{-1}) = \det(A)^{-1}$, e pertanto $\det(A) \neq 0$. Supponiamo $r(A) < n$, cioè che le righe di A siano linearmente dipendenti. Salvo scambiare tra loro due righe di A , il che, per il teorema (2.36.3), non influisce sulla conclusione, possiamo supporre che

$$A^{(1)} = c_2 A^{(2)} + \dots + c_n A^{(n)}, \quad c_2, \dots, c_n \in \mathbb{K}$$

cioè che la prima riga sia combinazione lineare delle rimanenti righe di A . Per il teorema (2.36.2) si ha

$$\det(A) = c_2 \det \begin{pmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(n)} \end{pmatrix} + c_3 \det \begin{pmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(n)} \end{pmatrix} + \dots + c_n \det \begin{pmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(n)} \end{pmatrix}.$$

Poiché ognuna delle matrici che compaiono a secondo membro ha due righe uguali, per il teorema (2.36.4) il suo determinante è 0. Pertanto

$$\det(A) = c_2 0 + c_3 0 + \dots + c_n 0 = 0.$$

□

Minore

Definizione 2.31. Sia $M \in M_{m,n}(\mathbb{K})$. Un **minore** di M è il determinante di una sua sottomatrice quadrata. L'**ordine** del minore è l'ordine della sottomatrice quadrata corrispondente.

Corollario 4. Sia $M \in M_{m,n}(\mathbb{K})$. Il rango di M è uguale al massimo ordine dei minori non nulli di M .

Proof. Il corollario è immediata conseguenza dei teoremi (2.34) e (2.37). \square

Proposizione 2.38. Sia $A = (a_{ij}) \in M_n(\mathbb{K})$:

1. Se A ha una riga oppure una colonna nulla, $\det(A) = 0$
2. Se $B \in M_n(\mathbb{K})$ è ottenuta da A sommando a una sua riga(colonna) un multiplo scalare di un'altra riga (colonna), allora $\det(B) = \det(A)$.

Proof.

1. Se A ha una riga, o una colonna, nulla, $r(A) < n$, e quindi $\det(A) = 0$ per il teorema (2.37).
2. Supponiamo che si abbia $B = (B_{(1)} \dots B_{(n)})$ con

$$B_{(i)} = A_{(i)} + cA_{(j)},$$

per qualche $i \neq j$, e

$$\forall k \neq i, B_{(k)} = A_{(k)}.$$

Allora

$$\det(B) = \det(A) + c \det(A_{(1)} \dots A_{(j)} \dots A_{(n)}) = \det(A) + c0 = \det(A)$$

perchè $(A_{(1)} \dots A_{(j)} \dots A_{(n)})$ ha uguali la i -esima e la j -esima colonna, e quindi ha determinante uguale a 0. In modo simile si procede nel caso in cui

$$B^{(i)} = A^{(i)} + cA^{(j)}.$$

\square

Definizione 2.32. Data $A = (a_{ij}) \in M_n(\mathbb{K})$, per ogni $1 \leq i, j \leq n$ sia

$$A(1 \dots \hat{i} \dots n | 1 \dots \hat{j} \dots n)$$

la sottomatrice quadrata di ordine $n - 1$ di A ottenuta cancellando la i -esima riga e la j -esima colonna. Il **complemento algebrico** (o **cofattore**) dell'elemento a_{ij} di A è

$$A_{ij} = (-1)^{i+j} \det(A(1 \dots \hat{i} \dots n | 1 \dots \hat{j} \dots n)).$$

La **matrice cofattore** di A è

$$\text{cof}(A) = (A_{ij}) \in M_n(\mathbb{K}).$$

Proposizione 2.39. Sia $A \in M_n(\mathbb{K})$. Allora

$$\forall i = 1, \dots, n \quad \det(A) = a_{i1}A_{i1} + a_{i2}A_{i2} + \dots + a_{in}A_{in} = \sum_{j=1}^n a_{ij}A_{ij} \quad (33)$$

e

$$\forall j = 1, \dots, n \quad \det(A) = a_{1j}A_{1j} + a_{2j}A_{2j} + \dots + a_{nj}A_{nj} = \sum_{i=1}^n a_{ij}A_{ij}. \quad (34)$$

La 33 e la 34 sono rispettivamente lo sviluppo di $\det(A)$ secondo la i -esima riga e secondo la j -esima colonna.

Proof. Sostituendo A con la sua trasposta ci si riduce a dimostrare solo la 33. Mediante $i - 1$ scambi fra righe contigue di A si ottiene una matrice

$$A' = \begin{pmatrix} A'^{(1)} \\ A'^{(2)} \\ \vdots \\ A'^{(n)} \end{pmatrix}$$

la cui prima riga è $A'^{(1)} = A^{(i)}$

e le rimanenti righe sono nella stessa posizione relativa delle righe di A , cioè

$$\begin{pmatrix} A'^{(1)} \\ A'^{(2)} \\ \vdots \\ A'^{(n)} \end{pmatrix} = \begin{pmatrix} A^{(1)} \\ \vdots \\ A^{(i-1)} \\ A^{(i+1)} \\ \vdots \\ A^{(n)} \end{pmatrix}.$$

Si ha pertanto

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

Inoltre, per ogni $1 \leq j \leq n$:

$$\begin{aligned} A'_{1j} &= (-1)^{j+1} \det[A'(2 \dots n \mid 1 \dots \hat{j} \dots n)] = \\ &= (-1)^{j+1} \det[A(1 \dots \hat{i} \dots n \mid 1 \dots \hat{j} \dots n)] = (-1)^{i+1} A_{ij}. \end{aligned}$$

Quindi, se la 33 è vera per A' , lo è anche per A . Quest'osservazione ci consente di limitarci a considerare il caso $i = 1$, cioè a dimostrare lo sviluppo ?? di A secondo la prima riga.

A questo scopo consideriamo i termini della somma 31 in cui compare a_{1j} che sono della forma

$$\varepsilon(p) a_{1j} a_{2p(2)} \dots a_{np(n)} \quad (35)$$

dove $p \in \sigma_n$ è una permutazione tale che $p(1) = j$. Ad ogni tale p possiamo far corrispondere la permutazione $q \in \sigma_{n-1}$ definita, per $k = 1, \dots, n-1$, da

$$q(k) = \begin{cases} p(k+1) & p(k+1) < j \\ p(k+1) - 1 & p(k+1) > j \end{cases} \quad (36)$$

Si ha

$$(-1)^{j-1} \varepsilon(q) = \varepsilon(p).$$

Infatti la permutazione $r \in \sigma_n$ definita da

$$\begin{aligned} r(1) &= 1, \\ r(k) &= q(k-1), \quad k = 2, \dots, n \end{aligned}$$

è ottenuta componendo p con $j-1$ trasposizioni di elementi contingui, e quindi soddisfa $\varepsilon(r) = (-1)^{j-1} \varepsilon(p)$. D'altra parte, per definizione, si ha evidentemente $\varepsilon(q) = \varepsilon(r)$. Poniamo

$$B = (b_{hk}) = A(2 \dots n \mid 1 \dots \hat{j} \dots n).$$

Poichè al variare di $p \in \sigma_n$ tale che $p(1) = j$ la permutazione q definita dalla (36) descrive tutto σ_{n-1} , la somma dei termini (35) è uguale a

$$a_{1j} (-1)^{j-1} \sum_{q \in \sigma_{n-1}} \varepsilon(q) b_{1q(1)} \dots b_{n-1, q(n-1)} = a_{1j} (-1)^{j-1} \det(B) = a_{1j} A_{1j}.$$

In conclusione, la somma di tutti i termini della (31) è uguale a $\sum_j a_{1j} A_{1j}$, che è la (33) per $i = 1$. □

Nota 9. Nell'applicare la proposizione (2.39) è conveniente, quando possibile, scegliere una riga o una colonna nella quale compaiano degli zeri, allo scopo di abbreviare i calcoli.

Formula di Laplace per l'inversa

Corollario 5. Per ogni $A = (a_{ij}) \in M_n(\mathbb{K})$ sussiste l'identità

$$A {}^t[\text{cof}(A)] = \det(A) \mathbf{I}_n. \quad (37)$$

In particolare, se A è invertibile, si ha

$$A^{-1} = \frac{{}^t[\text{cof}(A)]}{\det(A)} \quad (38)$$

Proof. La (37) è equivalente alle n^2 identità seguenti:

$$\sum_{k=1}^n a_{ik} A_{jk} = \det(A) \delta_{ij} \quad (39)$$

dove δ_{ij} è il simbolo di Kronecker. Infatti il primo membro è l'elemento di posto i, j della matrice $A {}^t[\text{cof}(A)]$. Nel caso $i = j$ la (39) coincide con la (??), che è già stata dimostrata. Se $i \neq j$ la (39) è

$$\sum_{k=1}^n a_{ik} A_{jk} = 0 \quad (40)$$

Per la proposizione (2.39), il primo membro della (40) è lo sviluppo secondo la i -esima riga del determinante della matrice B ottenuta da A sostituendo la sua riga i -esima alla j -esima. Poichè B ha due righe uguali il suo determinante è 0, e pertanto la (40) è vera. La (38) segue dalla (37) moltiplicando ambo i membri a sinistra per $\det(A)^{-1} A^{-1}$. \square

Regola di Cramer

Corollario 6. Siano $A = (a_{ij}) \in \text{GL}_n(\mathbb{K})$, $\mathbf{b} = {}^t(b_1 \dots b_n)$ un n -vettore colonna e

$$A\mathbf{x} = \mathbf{b} \quad (41)$$

il corrispondente sistema di n equazioni in n incognite. L'unica soluzione di (41), $\mathbf{x} = {}^t(x_1, \dots, x_n)$ è data dalla formula:

$$x_i = \frac{\sum_{k=1}^n b_k A_{ki}}{\det(A)}, \quad i = 1, \dots, n. \quad (42)$$

Proof. Basta ricordare la regola dell'inversa che da $\mathbf{x} = A^{-1}\mathbf{b}$, e sostituire la (38) al posto di A^{-1} . \square

Complemento algebrico o cofattore

Definizione 2.33. Sia $A \in M_n(\mathbb{K})$, e sia $M = A(i_1 \dots i_k | j_1 \dots j_k)$ una sottomatrice quadrata di ordine k di A . Il **complemento algebrico** (o **cofattore**) di M è

$$(-1)^{i_1 + \dots + i_k + j_1 + \dots + j_k} \det(A(\{1, \dots, n\} \setminus \{i_1, \dots, i_k\} | \{1, \dots, n\} \setminus \{j_1, \dots, j_k\})),$$

cioè è il determinante della sottomatrice quadrata di ordine $n - k$ di A ottenuta cancellando le righe e le colonne di M , preso con il segno $+$ o $-$ a seconda che $i_1 + \dots + i_k + j_1 + \dots + j_k$ sia pari o dispari.

Nota 10. Nel caso particolare di $k = 1$ si riottiene la definizione (36) di complemento algebrico di un elemento di A .

Metodo di Laplace o Sviluppo di Laplace

Teorema 2.40. Sia $A \in M_n(\mathbb{K})$, e siano assegnate $k \leq n$ righe (colonne) di A . Allora $\det(A)$ è uguale alla somma dei prodotti dei minori di ordine k di A estratti dalle righe (colonne) assegnate per i corrispondenti cofattori.

Proof. La daremo nel caso delle righe, lasciando al lettore l'immediata estensione al caso delle colonne. Denotiamo con $\mathcal{D}(A)$ la somma indicata nell'enunciato, che vogliamo dimostrare essere uguale a $\det(A)$. Il prodotto di un minore di ordine k estratto dalle k righe assegnate per il corrispondente complemento algebrico è una somma di $k!(n - k)!$ termini, perché il minore è un determinante di ordine k , e quindi ottenuto sommando $k!$ termini, mentre il cofattore è un determinante di ordine $n - k$, quindi somma di $(n - k)!$ termini. Inoltre il numero dei minori di ordine k distinti che si possono estrarre dalle k righe assegnate è

$$\binom{n}{k} = \frac{n!}{k!(n - k)!}$$

Pertanto $\mathcal{D}(A)$ è somma di

$$\binom{n}{k} k!(n-k)! = n!$$

termini. Poiché anche $\det(A)$ è somma di $n!$ termini, sarà sufficiente dimostrare che ogni termine di $\mathcal{D}(A)$ appare almeno una volta in $\det(A)$. Supponiamo dapprima che le righe assegnate siano $A^{(1)}, A^{(2)}, \dots, A^{(k)}$. I termini di $\mathcal{D}(A)$ sono prodotti di un termine di

$$\det(A(k+1 \dots n \mid \{1, \dots, n\} \setminus \{j_1, \dots, j_k\})),$$

per qualche scelta di j_1, \dots, j_k e l , moltiplicati per $(-1)^{1+\dots+k+j_1+\dots+j_k}$. Consideriamo il caso particolare in cui $\{j_1, \dots, j_k\} = \{1, \dots, k\}$. Un termine di $\mathcal{D}(A)$ corrispondente è della forma

$$\begin{aligned} & (-1)^{1+\dots+k+\dots+k} [\varepsilon(s) a_{1s(1)} \dots a_{ks(k)}] [\varepsilon(t) a_{k+1, k+t(1)} \dots a_{n, k+t(n-k)} \dots a_{ks(k)}] \\ & = e(s)e(t) [a_{s_1 t_1} \dots a_{s_k t_k} a_{k+1 i_1} \dots a_{n k-t}] \end{aligned} \quad (43)$$

dove $s \in \sigma_k$, $t \in \sigma_{n-k}$. Per riconoscere che (43) è un termine di $\det(A)$ osserviamo che la permutazione

$$u = \begin{pmatrix} 1 & \dots & k & k+1 & \dots & n \\ s(1) & \dots & s(k) & k+t(1) & \dots & k+t(n-k) \end{pmatrix}$$

ha segno uguale a $\varepsilon(s)\varepsilon(t)$, perché ogni indice $s(h)$ è minore di ogni indice $t(l)$, e quindi il numero di trasposizioni di cui u è prodotto è la somma del numero di trasposizioni di s e di quelle di t . Da ciò segue che (43) è un termine di $\det(A)$. Per dimostrare che il prodotto di un termine di $\det(A(1 \dots k, j_1 \dots j_k))$ per uno di $\det(A(k+1 \dots n \mid \{1, \dots, n\} \setminus \{j_1, \dots, j_k\}))$ moltiplicati per $(-1)^{1+\dots+k+j_1+\dots+j_k}$, per scelta arbitraria di $1 \leq j_1 < j_2 < \dots < j_k \leq n$, è un termine di $\det(A)$, scambiamo tra loro opportunamente le colonne di A in modo da ottenere una matrice B in cui la j_1 -esima colonna, la j_2 -esima colonna ecc.

di A si trovino rispettivamente come prima, seconda, \dots , k -esima colonna. Così facendo si operano $(j_1 - 1) + (j_2 - 1) + \dots + (j_k - 1) = j_1 + \dots + j_k - (1 + \dots + k)$ scambi di colonne, e quindi

$$\det(B) = (-1)^{j_1+\dots+j_k-(1+\dots+k)} \det(A) = (-1)^{1+\dots+k+j_1+\dots+j_k} \det(A).$$

Il termine di $\mathcal{D}(A)$ che stiamo considerando è uguale a $(-1)^{1+\dots+k+j_1+\dots+j_k}$, per un termine di $\mathcal{D}(B)$ del tipo che abbiamo considerato nella prima parte della dimostrazione, e quindi è uguale a $(-1)^{1+\dots+k+j_1+\dots+j_k}$ per un termine di $\det(B)$, cioè è uguale a un termine di $\det(A)$.

Se le righe assegnate sono $A^{(i_1)}, A^{(i_2)}, \dots, A^{(i_k)}$, consideriamo la matrice C ottenuta da A con opportune trasposizioni delle righe in modo che la i_1 -esima riga, la i_2 -esima riga ecc. di A siano rispettivamente prima, seconda, \dots , k -esima riga di C . Il numero di inversioni effettuate è

$$(i_1 - 1) + (i_2 - 1) + \dots + (i_k - 1) = i_1 + \dots + i_k - (1 + \dots + k),$$

e quindi $\det(C) = (-1)^{i_1+\dots+i_k-(1+\dots+k)} \det(A)$. D'altra parte ogni termine di $\mathcal{D}(C)$ è uguale a $(-1)^{i_1+\dots+i_k-(1+\dots+k)}$ per un termine di $\mathcal{D}(A)$, e quindi

$$\mathcal{D}(C) = (-1)^{i_1+\dots+i_k-(1+\dots+k)} \mathcal{D}(A).$$

Confrontando otteniamo $\mathcal{D}(A) = \det(A)$, come si voleva. □

Nota 11. Esso è particolarmente utile quando la matrice A di cui si vuole calcolare il determinante ha molti elementi uguali a 0.

Esempio 2.14. Sia A una matrice triangolare superiore (inferiore). Sviluppandone il determinante secondo la prima colonna (l prima riga), si trova subito, procedendo per induzione su n , che $\det(A)$ è uguale al prodotto degli elementi della diagonale principale.

Principio dei minori orlati Sia data una matrice $M \in M_{m,n}(\mathbb{K})$ non nulla (cioè tale che $r(M) \geq 1$). Per determinare il rango di M si procede nel seguente modo:

1. Si cercano, per $k \geq 2$, sottomatrici quadrate di ordine k e si calcola il rispettivo determinante. Definiamo r come il massimo intero tale che esista almeno una sottomatrice di ordine r con determinante non nullo, e, contestualmente, ogni sottomatrice quadrata di ordine $r+1$ abbia determinante nullo (oppure $r = \min\{m, n\}$). In tal modo si conclude formalmente che

$$r(M) = r.$$

Infatti, se per ogni $s > r$ ogni sottomatrice quadrata di ordine s ha determinante zero, è immediato (per induzione su s) che non si possiede nessuna famiglia di s colonne linearmente indipendenti, dunque il rango è al massimo r .

2. Il procedimento può essere reso più efficiente mediante il seguente **principio dei minori orlati**. Sia

$$B = M(i_1, i_2, \dots, i_r \mid j_1, j_2, \dots, j_r)$$

una sottomatrice quadrata di ordine r tale che

$$\det(B) \neq 0.$$

Supponiamo inoltre che ogni sottomatrice quadrata di ordine $r + 1$ ottenuta “estendendo” B (cioè, aggiungendo una riga e la corrispondente colonna di M) abbia determinante nullo. In altre parole, i cosiddetti **minori orlati** relativi a B sono tutti nulli. Da ciò ne consegue che le colonne di M aventi indice j_1, \dots, j_r sono linearmente indipendenti e, essendo ogni altra colonna (cioè, ogni colonna non compresa in $\{j_1, \dots, j_r\}$) genera un minore di ordine $r + 1$ nullo, deve essere espressa come combinazione lineare di quelle in B . Quindi l'intero insieme delle colonne di M è generato dalle colonne con indici j_1, \dots, j_r e, per definizione, il rango della matrice è esattamente r .

Esempio 2.15. Supponiamo assegnato un sistema di m equazioni in n incognite, in cui i coefficienti delle incognite e i termini noti siano funzioni di uno o più parametri variabili in \mathbb{K} . Per ogni valore assunto dai parametri si ottiene un diverso sistema a coefficienti in \mathbb{K} di cui si vuole accertare la compatibilità e ricercare le eventuali soluzioni: lo studio dei casi che si presentano e la ricerca delle rispettive soluzioni si dice la **discussione del sistema assegnato**. Il modo più efficace e naturale di procedere in questo caso è quello di utilizzare il teorema (2.35) analizzando i valori possibili del rango della matrice dei coefficienti e della matrice orlata in funzione dei parametri. Una volta stabiliti i valori dei parametri per cui il sistema è compatibile, e in ogni caso l'infinità delle soluzioni, si procederà a risolverlo in ciascun caso.

Esempio 2.16. Siano $x_1, x_2, \dots, x_n \in \mathbb{K}$. Il determinante $\mathcal{V}(x_1, x_2, \dots, x_n)$ della **matrice di Vandermonde**

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

è detto **determinante di Vandermonde** relativo a x_1, x_2, \dots, x_n . Si ha

$$\mathcal{V}(x_1, x_2, \dots, x_n) = (x_2 - x_1) \dots (x_n - x_1)(x_3 - x_2) \dots (x_n - x_n)(x_4 - x_3) \dots (x_n - x_{n-1}) \quad (44)$$

il secondo membro essendo il prodotto di tutti i termini della forma $(x_i - x_j)$, con $1 \leq j < i \leq n$. La (44) si dimostra per induzione su n , il caso $n = 2$ essendo ovvio. Supponiamo $n \geq 3$ e che la (44) sia stata dimostrata per $n - 1$. Operiamo sulle righe della matrice di Vandermonde, sottraendo dalla n -esima riga la $(n - 1)$ -esima moltiplicata per x_1 , dalla $(n - 1)$ -esima la $(n - 2)$ -esima moltiplicata per x_1 ecc. Si ottiene

$$\begin{aligned} \mathcal{V}(x_1, x_2, \dots, x_n) &= \begin{vmatrix} 1 & 1 & \dots & 1 \\ 0 & x_2 - x_1 & \dots & x_n - x_1 \\ 0 & x_2^2(x_2 - x_1) & \dots & x_n^2(x_n - x_1) \\ \vdots & \vdots & & \vdots \\ 0 & x_2^{n-2}(x_2 - x_1) & \dots & x_n^{n-2}(x_n - x_1) \end{vmatrix} \\ &= (x_2 - x_1)(x_3 - x_1) \dots (x_n - x_1) \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_2 & x_3 & \dots & x_n \\ x_2^2 & x_3^2 & \dots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_2^{n-2} & x_3^{n-2} & \dots & x_n^{n-2} \end{pmatrix} \\ &= (x_2 - x_1)(x_3 - x_1) \dots (x_n - x_1) \mathcal{V}(x_2, \dots, x_n) = \mathcal{V}(x_1, x_2, \dots, x_n). \end{aligned}$$

Esempio 2.17. Non è difficile dimostrare che le proprietà (2), (3) e (5) del teorema (2.36) sono sufficienti a caratterizzare univocamente il determinante. Precisamente, si dimostra che, se

$$D : M_n(\mathbb{K}) \rightarrow \mathbb{K}$$

$$\mathbf{A} = (\mathbf{A}_{(1)} \dots \mathbf{A}_{(n)}) \mapsto \begin{cases} D(\mathbf{A}_{(1)} \dots cV + c'V' \dots \mathbf{A}_{(n)}) = cD(\mathbf{A}_{(1)} \dots V \dots \mathbf{A}_{(n)}) + \\ \quad c'D(\mathbf{A}_{(1)} \dots V' \dots \mathbf{A}_{(n)}) \quad \text{se } \mathbf{A}_{(j)} = cV + c'V' \\ D(\mathbf{A}_{(1)} \dots \mathbf{A}_{(i)} \dots \mathbf{A}_{(j)} \dots \mathbf{A}_{(n)}) = \\ \quad -D(\mathbf{A}_{(1)} \dots \mathbf{A}_{(j)} \dots \mathbf{A}_{(i)} \dots \mathbf{A}_{(n)}) \quad 1 \leq i < j \leq n, \\ D(\mathbf{I}_n) = 1 \end{cases}$$

Allora

$$\forall \mathbf{A} \in M_n(\mathbb{K}), \quad D(\mathbf{A}) = \det(\mathbf{A}).$$

3 Spazi Affini

3.1 Spazi e Sottospazi Affini

Spazio affine

Definizione 3.1. Sia \mathbf{V} un \mathbb{K} -spazio vettoriale. Uno **spazio affine** su \mathbf{V} (ovvero uno **spazio affine con spazio vettoriale associato \mathbf{V}**) è un insieme non vuoto \mathbf{A} , i cui elementi si dicono **punti** di \mathbf{A} , tale che sia data un'applicazione definita al seguente modo

$$(P, Q) \in \mathbf{A} \times \mathbf{A} \rightarrow \overrightarrow{PQ} \in \mathbf{V}. \quad (45)$$

tale che soddisfi i seguenti due assiomi:

- **SA1:**

$$\forall P \in \mathbf{A}, \forall \mathbf{v} \in \mathbf{V}, \exists! Q \in \mathbf{A} : \overrightarrow{PQ} = \mathbf{v}.$$

- **SA2:**

$$\forall P, Q, R \in \mathbf{A}, \overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}.$$

Per ogni $(P, Q) \in \mathbf{A} \times \mathbf{A}$, diremo P **punto di applicazione** o **punto iniziale** e Q **punto finale** del vettore \overrightarrow{PQ} . Se $\mathbb{K} = \mathbb{R}$ ($\mathbb{K} = \mathbb{C}$), \mathbf{A} si dice **spazio affine reale** (**spazio affine complesso**). L'applicazione (45) definisce una **struttura di spazio affine** sull'insieme \mathbf{A} .

Dal secondo assioma notiamo che

$$P = Q = R \implies \overrightarrow{PP} + \overrightarrow{PP} = \overrightarrow{PP} \implies \overrightarrow{PP} = \vec{0}.$$

Prendendo invece $R = P$ troviamo che

$$\forall P, Q \in \mathbf{A}, \overrightarrow{QP} = -\overrightarrow{PQ}.$$

Su un insieme non vuoto \mathbf{A} possono esistere diverse strutture di spazio affine, cioè diversi modi di assegnare uno spazio vettoriale \mathbf{V} e una applicazione $\mathbf{A} \times \mathbf{A} \rightarrow \mathbf{V}$ che soddisfa SA1 e SA2. Nel seguito considereremo esclusivamente spazi affini tali che lo spazio vettoriale associato \mathbf{V} abbia dimensione finita.

Dimensione di uno spazio affine

Definizione 3.2. Sia \mathbf{A} uno spazio affine su \mathbf{V} . Si dice **dimensione dello spazio affine \mathbf{A}** , e scriveremo $\dim(\mathbf{A})$, la dimensione dello spazio vettoriale associato \mathbf{V} .

Nota 12. Uno spazio affine di dimensione 1 (2) viene comunemente chiamato **retta affine** (**piano affine**).

Esempio 3.1. La retta, il piano e lo spazio ordinari sono rispettivamente una retta affine reale, un piano affine reale e uno spazio affine reale di dimensione 3. Gli spazi vettoriali associati sono quelli dei vettori geometrici dei rispettivi spazi, e l'operazione che associa un vettore a una coppia ordinata di punti è quella con cui abbiamo definito inizialmente i vettori geometrici. Quindi gli spazi affini sono generalizzazioni della retta, del piano e dello spazio ordinari.

Spazio affine su sè stesso

Proposizione 3.1. Sia \mathbf{V} uno spazio vettoriale di dimensione finita su \mathbb{K} . Mediante la seguente applicazione

$$\mathbf{V}^2 \rightarrow \mathbf{V}, \quad (\mathbf{a}, \mathbf{b}) \mapsto \overrightarrow{\mathbf{ab}} = \mathbf{b} - \mathbf{a}$$

si definisce su \mathbf{V} una struttura di spazio affine su sè stesso. Con questa struttura di spazio affine, \mathbf{V} si denoterà \mathbf{V}_a .

Proof. Infatti

• SA1:

$$\forall \mathbf{p}, \mathbf{v} \in \mathbf{V}, \quad \exists! \mathbf{q} \in \mathbf{V} : \overrightarrow{\mathbf{q}\mathbf{p}} = \mathbf{p} - \mathbf{q} = \mathbf{v} \quad \text{con} \quad \mathbf{q} = \mathbf{p} + \mathbf{v};$$

• SA2:

$$\forall \mathbf{p}, \mathbf{q}, \mathbf{r} \in \mathbf{V}, \quad \overrightarrow{\mathbf{p}\mathbf{q}} + \overrightarrow{\mathbf{q}\mathbf{r}} = \overrightarrow{\mathbf{p}\mathbf{r}} \quad \because \quad \mathbf{q} - \mathbf{p} + \mathbf{r} - \mathbf{q} = \mathbf{r} - \mathbf{p} = \mathbf{pr}.$$

□

Nota 13. Quindi ogni spazio vettoriale \mathbf{V} può considerarsi come uno spazio affine su sè stesso.

***n*-spazio affine numerico**

Caso 2. Un caso particolare dell'esempio precedente si ha prendendo $\mathbf{V} = \mathbb{K}^n$. Lo spazio affine $(\mathbb{K}^n)_a$ si chiama ***n*-spazio affine numerico su \mathbb{K}** . esso si denota con $\mathbf{A}^n(\mathbb{K})$, o semplicemente \mathbf{A}^n quando il campo è chiaro dal contesto.

Nota 14. L'assioma SA1 implica che se in uno spazio affine \mathbf{A} si fissa un punto $O \in \mathbf{A}$ e si associa ad ogni punto $P \in \mathbf{A}$ il vettore $\overrightarrow{OP} \in \mathbf{V}$, si ottiene una corrispondenza biunivoca di \mathbf{A} su \mathbf{V} . Tale corrispondenza è la generalizzazione di quella che, nello spazio ordinario in cui si sia fissato un punto O , associa ad ogni punto P il vettore geometrico rappresentato dal segmento orientato di punto iniziale O e punto finale P .

Sistema di coordinate affini

Definizione 3.3. Siano \mathbf{V} un \mathbb{K} -spazio vettoriale e \mathbf{A} uno spazio affine su \mathbf{V} . Un **sistema di coordinate affini** (ovvero un **referimento affine**) nello spazio \mathbf{A} è assegnato una volta fissati un punto $O \in \mathbf{A}$ e una base $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ di \mathbf{V} ; esso viene denotato con $O\mathbf{e}_1 \dots \mathbf{e}_n$.

Coordinate affini, *n*-upla delle coordinate, e origine

Definizione 3.4. Per ogni punto $P \in \mathbf{A}$ si ha

$$\overrightarrow{OP} = a_1\mathbf{e}_1 + \dots + a_n\mathbf{e}_n = \sum_{i=1}^n a_i\mathbf{e}_i$$

per opportuni $a_1, \dots, a_n \in \mathbb{K}$. Gli scalari a_1, \dots, a_n si dicono **coordinate affini** (o semplicemente **coordinate**) e (a_1, \dots, a_n) si dice ***n*-upla delle coordinate**, di P rispetto al riferimento $O\mathbf{e}_1 \dots \mathbf{e}_n$. Il punto O si dice **origine** del sistema di coordinate, esso ha $(0, \dots, 0)$ per *n*-upla delle coordinate. Dato un riferimento affine $O\mathbf{e}_1 \dots \mathbf{e}_n$ in \mathbf{A} , scriveremo $P(x_1, \dots, x_n)$ per denotare un punto $P \in \mathbf{A}$ di coordinate x_1, \dots, x_n .

Proposizione 3.2. Siano $A(a_1, \dots, a_n), B(b_1, \dots, b_n) \in \mathbf{A}$. Il vettore \overrightarrow{AB} ha $(b_1 - a_1, \dots, b_n - a_n)$ come *n*-upla di coordinate rispetto alla base $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$.

Proof. Per l'identità $\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA}$, le coordinate di \overrightarrow{AB} si ottengono sottraendo alle coordinate di B quelle di A .

$$\sum_{i=1}^n b_i\mathbf{e}_i - \sum_{i=1}^n a_i\mathbf{e}_i = \sum_{i=1}^n (b_i - a_i)\mathbf{e}_i.$$

□

Riferimento affine standard

Definizione 3.5. Se $\mathbf{A} = \mathbf{A}^n$, il riferimento affine $O\mathbf{E}_1 \dots \mathbf{E}_n$ in cui $O = (0, \dots, 0)$ ed $\{\mathbf{E}_1, \dots, \mathbf{E}_n\}$ è la base canonica di \mathbb{K}^n si dice **riferimento affine standard**. In questo riferimento ogni punto $(x_1, \dots, x_n) \in \mathbf{A}^n$ ha sè stesso come n -upla di coordinate.

Sottospazio affine

Definizione 3.6. Siano \mathbf{V} un \mathbb{K} -spazio vettoriale e \mathbf{A} uno spazio affine su \mathbf{V} . Siano assegnati un punto $Q \in \mathbf{A}$ e un sottospazio vettoriale \mathbf{W} di \mathbf{V} . Il **sottospazio affine passante per Q e parallelo a \mathbf{W}** è il sottoinsieme \mathbf{S} di \mathbf{A} costituito da tutti i punti $P \in \mathbf{A}$ tali che $\overrightarrow{QP} \in \mathbf{W}$. In simboli:

$$\mathbf{S} = \{P \in \mathbf{A} : \overrightarrow{QP} \in \mathbf{W}\}.$$

Nota 15. Si noti che $Q \in \mathbf{S}$, perchè il sottospazio vettoriale \mathbf{W} contiene il vettore nullo $\mathbf{0} = \overrightarrow{QQ}$; in particolare $\mathbf{S} \neq \emptyset$.

Giacitura e dimensione di un sottospazio affine

Definizione 3.7. Il sottospazio $\mathbf{W} \subset \mathbf{V}$ è chiamato **giacitura** di \mathbf{S} ; il numero $\dim(\mathbf{W})$ è detto **dimensione** di \mathbf{S} e si indica con $\dim(\mathbf{S})$.

Considerazioni sulle dimensioni dei sottospazi affini

- Se $\dim(\mathbf{S}) = 0$, allora $\mathbf{S} = \{Q\}$ è un punto; viceversa, ogni sottoinsieme di \mathbf{A} costituito da un solo punto è un sottospazio affine di dimensione 0.
- Se $\dim(\mathbf{S}) = 1$, \mathbf{S} si dice **retta di \mathbf{A}** e \mathbf{W} **direzione di \mathbf{S}** ; un qualunque vettore non nullo $\mathbf{a} \in \mathbf{W}$ è un **vettore di direzione** della retta. Segue dalla definizione che la retta \mathbf{S} consiste di tutti i punti $P \in \mathbf{A}$ tali che $\overrightarrow{QP} = t\mathbf{a}$ per qualche $t \in \mathbb{K}$.
- Se $\dim(\mathbf{S}) = 2$, \mathbf{S} si dice **piano di \mathbf{A}** .

La dimensione di un sottospazio affine non può superare $\dim(\mathbf{A})$. Se $\dim(\mathbf{S}) = \dim(\mathbf{A})$, allora $\mathbf{S} = \mathbf{A}$, poichè in tal caso $\mathbf{W} = \mathbf{V}$ e quindi

$$\forall P \in \mathbf{A}, \overrightarrow{QP} \in \mathbf{W}.$$

Se $\dim(\mathbf{S}) = \dim(\mathbf{A}) - 1$, \mathbf{S} è detto **iperpiano**. Ad esempio una retta in un piano affine è un iperpiano, così come un piano in uno spazio affine di dimensione 3.

Sia \mathbf{A} lo spazio ordinario. I sottospazi affini di \mathbf{A} sono i punti, le rette, i piani e \mathbf{A} stesso.

Traslato di uno spazio vettoriale

Caso 3. Sia \mathbf{V} un \mathbb{K} -spazio vettoriale non nullo di dimensione finita. Consideriamo un sottospazio vettoriale $\mathbf{W} \subset \mathbf{V}$ e un punto $\mathbf{q} \in \mathbf{V}_a$. Il sottospazio affine di \mathbf{V}_a passante per \mathbf{q} e parallelo a \mathbf{W} è l'insieme

$$\mathbf{q} + \mathbf{W} = \{\mathbf{q} + \mathbf{w} \mid \mathbf{w} \in \mathbf{W}\}.$$

Infatti $\mathbf{q} + \mathbf{W}$ consiste di tutti i $\mathbf{v} \in \mathbf{V}_a$ tali che $\mathbf{v} - \mathbf{q} \in \mathbf{W}$. Se in particolare $\mathbf{q} \in \mathbf{W}$, allora $\mathbf{q} + \mathbf{W} = \mathbf{W}$. Vediamo quindi che i sottospazi vettoriali di \mathbf{V} sono particolari sottospazi affini di \mathbf{V}_a e che ogni sottospazio affine è della forma $\mathbf{q} + \mathbf{W}$, per qualche $\mathbf{q} \in \mathbf{V}_a$ e per qualche sottospazio vettoriale $\mathbf{W} \subset \mathbf{V}$, cioè è un **traslato** di un sottospazio vettoriale di \mathbf{V} .

Sottospazio affine generato da punti

Definizione 3.8. Dati $N + 1 \geq 2$ punti P_0, \dots, P_N di uno spazio affine \mathbf{A} , il sottospazio passante per P_0 e avente giacitura

$$\langle \overrightarrow{P_0P_1}, \overrightarrow{P_0P_2}, \dots, \overrightarrow{P_0P_N} \rangle$$

viene indicato con

$$\overline{P_0P_1 \dots P_N}$$

e si chiama **sottospazio affine generato da** P_0, \dots, P_N .

Proposizione 3.3. Il sottospazio affine generato $\overline{P_0P_1 \dots P_N}$ non dipende dall'ordine in cui vengono presi i punti P_0, \dots, P_N .

Proof. Infatti il sottospazio vettoriale $\langle \overrightarrow{P_0P_1}, \overrightarrow{P_0P_2}, \dots, \overrightarrow{P_0P_N} \rangle$ contiene tutti i vettori

$$\overrightarrow{P_iP_j} = \overrightarrow{P_iP_0} + \overrightarrow{P_0P_j} = \overrightarrow{P_0P_j} - \overrightarrow{P_0P_i},$$

e quindi

$$\forall i = 1, \dots, N, \quad \langle \overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_N} \rangle \supset \langle \overrightarrow{P_iP_0}, \dots, \overrightarrow{P_iP_N} \rangle. \quad (46)$$

Viceversa, per ogni $i = 1, \dots, N$, ogni vettore $\overrightarrow{P_0P_j}$ si può esprimere come

$$\overrightarrow{P_0P_j} = \overrightarrow{P_0P_i} + \overrightarrow{P_iP_j} = \overrightarrow{P_iP_j} - \overrightarrow{P_iP_0},$$

e quindi anche

$$\langle \overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_N} \rangle \subset \langle \overrightarrow{P_iP_0}, \dots, \overrightarrow{P_iP_N} \rangle. \quad (47)$$

Pertanto, se $P \in \mathbf{A}$ soddisfa

$$\overrightarrow{P_0P} \in \langle \overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_N} \rangle$$

allora

$$\overrightarrow{P_iP} = \overrightarrow{P_0P} - \overrightarrow{P_0P_i} \in \langle \overrightarrow{P_iP_0}, \dots, \overrightarrow{P_iP_N} \rangle$$

per la (47). D'altra parte, se $\overrightarrow{P_iP} \in \langle \overrightarrow{P_iP_0}, \dots, \overrightarrow{P_iP_N} \rangle$ allora

$$\overrightarrow{P_0P} = \overrightarrow{P_iP} - \overrightarrow{P_iP_0} \in \langle \overrightarrow{P_0P_i}, \dots, \overrightarrow{P_0P_N} \rangle$$

per la (46). □

Punti dipendenti e indipendenti

Definizione 3.9. Dalla definizione di $\overline{P_0P_1 \dots P_N}$ segue che $\dim(\overline{P_0P_1 \dots P_N}) \leq N$.

- Se $\dim(\overline{P_0P_1 \dots P_N}) = N$, allora i punti P_0, \dots, P_N si dicono **indipendenti**;
- Se $\dim(\overline{P_0P_1 \dots P_N}) < N$, allora i punti P_0, \dots, P_N si dicono **dipendenti**.

Per definizione i punti P_0, \dots, P_N sono indipendenti se e solo se i vettori $\overrightarrow{P_0P_1}, \overrightarrow{P_0P_2}, \dots, \overrightarrow{P_0P_N}$ sono linearmente indipendenti. Se P_0, \dots, P_N sono indipendenti, allora $N \leq \dim(\mathbf{A})$. Due punti $P_0, P_1 \in \mathbf{A}$ sono indipendenti se e solo se sono distinti; in questo caso $\overline{P_0P_1}$ è una retta. Tre punti P_0, P_1, P_2 sono indipendenti se e solo se non appartengono a una retta; in questo caso $\overline{P_0P_1P_2}$ è un piano.

Punti allineati, non allineati, e complanari

Definizione 3.10. I punti $P_0, \dots, P_N \in \mathbf{A}$ si dicono **allineati** (o **collineari**) se esista una retta che li contiene, o, equivalentemente, se

$$\dim(\overline{P_0P_1 \dots P_N}) \leq 1;$$

altrimenti P_0, \dots, P_N si dicono **non allineati** (o **non collineari**). I punti $P_0, \dots, P_N \in \mathbf{A}$ si dicono **complanari**

se esiste un piano che li contiene, o, equivalentemente, se

$$\dim(\overline{P_0 P_1 \dots P_N}) \leq 2.$$

Punto simmetrico di P rispetto a C

Definizione 3.11. Siano \mathbf{A} uno spazio affine su \mathbf{V} , e $C \in \mathbf{A}$. Per ogni $P \in \mathbf{A}$, il **punto simmetrico di P rispetto a C** è il punto $\sigma_C(P)$ che soddisfa l'identità vettoriale

$$\overrightarrow{C\sigma_C(P)} = -\overrightarrow{CP}.$$

Si osservi che si ha

$$\forall P \in \mathbf{A} \quad \sigma_C(\sigma_C(P)) = P,$$

perchè

$$\overrightarrow{C\sigma_C(\sigma_C(P))} = -\overrightarrow{C\sigma_C(P)} = \overrightarrow{CP}.$$

Se in \mathbf{A} è fissato un riferimento affine $O\mathbf{e}_1, \dots, \mathbf{e}_n$, e i punti C e P hanno rispettivamente coordinate c_1, \dots, c_n e x_1, \dots, x_n , il punto $\sigma_C(P)$ ha coordinate $2c_1 - x_1, \dots, 2c_n - x_n$. Infatti

$$(2c_1 - x_1, \dots, 2c_n - x_n) - (c_1, \dots, c_n) = (c_1 - x_1, \dots, c_n - x_n)$$

è la n -upla delle coordinate di $\overrightarrow{PC} = -\overrightarrow{CP}$. Nel caso particolare $C = O$, le coordinate di $\sigma_C(P)$ sono $-x_1, \dots, -x_n$.

L'applicazione $\sigma_C : \mathbf{A} \rightarrow \mathbf{A}$ manda sottospazi affini in sottospazi affini. Più precisamente, sia $\mathbf{S} \subset \mathbf{A}$ il sottospazio affine avente giacitura $\mathbf{W} \subset \mathbf{V}$ e passante per il punto $Q \in \mathbf{A}$. Allora

$$\sigma_C(\mathbf{S}) = \{\sigma_C(P) : P \in \mathbf{S}\}$$

è il sottospazio affine \mathbf{S}' avente la stessa giacitura \mathbf{W} e passante per $\sigma_C(Q)$. Infatti abbiamo che

$$\forall P \in \mathbf{S}, \quad \overrightarrow{\sigma_C(Q)\sigma_C(P)} = \overrightarrow{\sigma_C(Q)C} + \overrightarrow{C\sigma_C(P)} = \overrightarrow{CQ} - \overrightarrow{CP} = -\overrightarrow{QP} \in \mathbf{W}.$$

e quindi $\sigma_C(P) \in \mathbf{S}'$. Viceversa, se $P' \in \mathbf{S}'$, allora, posto $P = \sigma_C(P')$, si ha

$$\overrightarrow{QP} = \overrightarrow{Q\sigma_C(P')} = -\overrightarrow{\sigma_C(Q)P'} = -\overrightarrow{\sigma_C(Q)P'} \in \mathbf{W}$$

e quindi $P \in \mathbf{S}$. Dunque

$$P' = \sigma_C(P) \in \sigma_C(\mathbf{S}).$$

Proposizione 3.4.

1. Un sottospazio affine è individuato dalla sua giacitura e da uno qualsiasi dei suoi punti.
2. Sia \mathbf{S} un sottospazio affine di \mathbf{A} avente giacitura \mathbf{W} . Associando ad ogni coppia di punti P, Q di \mathbf{S} il vettore \overrightarrow{PQ} si definisce su \mathbf{S} una struttura di spazio affine su \mathbf{W}

Proof.

1. Sia S il sottospazio affine di \mathbf{A} passante per Q ed avente giacitura \mathbf{W} . Sia $M \in S$ e sia T il sottospazio affine passante per M ed avente giacitura \mathbf{W} . Se $P \in S$, allora si ha

$$\overrightarrow{MP} = \overrightarrow{MQ} + \overrightarrow{QP} = -\overrightarrow{QM} + \overrightarrow{QP}$$

che è un vettore di \mathbf{W} perchè entrambi gli addendi vi appartengono; quindi $P \in T$. Se viceversa $P \in T$, allora

$$\overrightarrow{QP} = \overrightarrow{QM} + \overrightarrow{MP} = -\overrightarrow{MQ} + \overrightarrow{MP} \in \mathbf{W},$$

e quindi $P \in S$. In conclusione $S = T$.

2. Se $P, Q \in S$, allora $\overrightarrow{PQ} \in \mathbf{W}$ perchè per la (1), S coincide con il sottospazio affine passante per P e parallelo a \mathbf{W} . Otteniamo quindi un'applicazione

$$S \times S \rightarrow \mathbf{W}, \quad (P, Q) \mapsto \overrightarrow{PQ}$$

la quale soddisfa le proprietà SA1 e SA2, perchè esse sono verificate in \mathbf{A} .

□

Semirette

Definizione 3.12. Sia \mathbf{A} un sottospazio affine reale con spazio vettoriale associato \mathbf{V} . Fissati un punto $Q \in \mathbf{A}$ e un vettore non nullo $\mathbf{a} \in \mathbf{V}$, l'insieme dei punti $P \in \mathbf{A}$ tali che

$$\overrightarrow{QP} = t\mathbf{a}$$

per qualche $t \geq 0$ è la **semiretta di origine Q e direzione \mathbf{a}** .

Segmenti

Definizione 3.13. Se A e B sono due punti distinti di \mathbf{A} , il **segmento di estremi A e B** è l'insieme dei punti $P \in \mathbf{A}$ tali che

$$\overrightarrow{AP} = t\overrightarrow{AB} \quad t \in [0, 1] \subset \mathbb{R},$$

e si denota con AB .

I punti $P_1, \dots, P_{t-1} \in AB$ che dividono il segmento in t parti uguali, cioè tali che $\overrightarrow{AP_1} = \overrightarrow{P_1P_2} = \dots = \overrightarrow{P_{t-1}B}$ sono definiti dalle condizioni

$$\overrightarrow{AP_i} = \frac{i}{t}\overrightarrow{AB}, \quad i = 1, \dots, t-1.$$

Se in \mathbf{A} è fissato un riferimento affine $O\mathbf{e}_1, \dots, \mathbf{e}_n$ e $A = A(a_1, \dots, a_n)$, $B = B(b_1, \dots, b_n)$, $P_i = P_i(x_1, \dots, x_n)$, la condizione è equivalente alla seguente:

$$(x_1 - a_1, \dots, x_n - a_n) = \frac{i}{t}(b_1 - a_1, \dots, b_n - a_n)$$

e pertanto

$$(x_1, \dots, x_n) = \frac{1}{t}(ib_1 + (t-i)a_1, \dots, ib_n + (t-i)a_n).$$

In particolare, il **punto medio del segmento AB** ha la n -upla di coordinate

$$(x_1, \dots, x_n) = \frac{1}{t} \left(\frac{a_1 + b_1}{2}, \frac{a_2 + b_2}{2}, \dots, \frac{a_n + b_n}{2} \right).$$

Triangolo

Definizione 3.14. Sia \mathbf{A} uno spazio affine reale, e siano $A, B, C \in \mathbf{A}$ tre punti non allineati. Il **triangolo di vertici A, B, C** è l'insieme dei punti $P \in \mathbf{A}$ tali che

$$\overrightarrow{AP} = t\overrightarrow{AB} + u\overrightarrow{AC}$$

per qualche $t, u \in \mathbb{R}$, con $t, u \geq 0$ e $t + u \leq 1$.

Parallelogramma

Definizione 3.15. Il **parallelogramma** individuato da A, B, C è l'insieme dei punti $P \in \mathbf{A}$ tali che

$$\overrightarrow{AP} = t\overrightarrow{AB} + u\overrightarrow{AC}$$

per qualche $t, u \in \mathbb{R}$, con $0 \leq t, u \leq 1$.

Tetraedro

Definizione 3.16. Sia \mathbf{A} uno spazio affine reale, e siano $A, B, C, D \in \mathbf{A}$ punti non complanari. Il **tetraedro** di vertici A, B, C, D è l'insieme dei punti $P \in \mathbf{A}$ tali che

$$\overrightarrow{AP} = t\overrightarrow{AB} + u\overrightarrow{AC} + v\overrightarrow{AD}$$

per qualche $t, u, v \in \mathbb{R}$, con $t, u, v \geq 0$ e $t + u + v \leq 1$.

Parallelepipedo

Definizione 3.17. Il **parallelepipedo** individuato da A, B, C, D è l'insieme dei punti $P \in \mathbf{A}$ tali che

$$\overrightarrow{AP} = t\overrightarrow{AB} + u\overrightarrow{AC} + v\overrightarrow{AD}$$

per qualche $t, u, v \in \mathbb{R}$, con $0 \leq t, u, v \leq 1$.

Simplessi

Definizione 3.18. Siano \mathbf{A} uno spazio affine reale, e $P_0, P_1, \dots, P_k \in \mathbf{A}$ punti indipendenti. Il **k-simpleso di vertici** P_0, P_1, \dots, P_k è l'insieme dei punti $P \in \mathbf{A}$ tali che

$$\overrightarrow{P_0P} = t_1\overrightarrow{P_0P_1} + t_2\overrightarrow{P_0P_2} + \dots + t_k\overrightarrow{P_0P_k}$$

per qualche $t_1, t_2, \dots, t_k \in \mathbb{R}$ con $t_1, t_2, \dots, t_k \geq 0$ e $\sum_{i=1}^k t_i \leq 1$.

Esempio 3.2. I segmenti, i triangoli e i tetraedri sono casi particolari di simpleso; in particolare, per $k = 1, 2, 3$, un k -simpleso corrisponde rispettivamente a un segmento, a un triangolo e a un tetraedro.

Sottoinsiemi e involuppi convessi

Definizione 3.19. Sia \mathbf{A} uno spazio affine reale. Un sottoinsieme \mathbf{S} di \mathbf{A} si dice **convesso** se per ogni $A, B \in \mathbf{S}$ il segmento AB è contenuto in \mathbf{S} . Se γ è un sottoinsieme di \mathbf{A} , l'**inviluppo convesso** di γ è il più piccolo sottoinsieme convesso di \mathbf{A} contenente γ , cioè è l'intersezione di tutti i sottoinsiemi convessi di \mathbf{A} che contengono γ .

Sia \mathbf{A} uno spazio affine reale con spazio vettoriale associato \mathbf{V} . L'assioma SA1 definisce la seguente applicazione

$$t : \mathbf{A} \times \mathbf{V} \rightarrow \mathbf{A}, \quad (A, \mathbf{a}) \mapsto B : \overrightarrow{AB} = \mathbf{a}.$$

L'applicazione t gode delle seguenti proprietà:

1. $\forall A \in \mathbf{A}, \forall \mathbf{a}, \mathbf{b} \in \mathbf{V}, \quad t(t(A, \mathbf{a}), \mathbf{b}) = t(A, \mathbf{a} + \mathbf{b});$
2. $\forall A, B \in \mathbf{A}, \exists! \mathbf{a} \in \mathbf{V} : B = t(A, \mathbf{a})$

Ponendo infatti $B = t(A, \mathbf{a}), C = t(t(A, \mathbf{a}), \mathbf{b})$, si ha $\overrightarrow{AB} = \mathbf{a}, \overrightarrow{BC} = \mathbf{b}$, e la (a) afferma che $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$, che è vero per SA2. La (b) è una riformulazione di SA1. Viceversa, dati uno spazio vettoriale \mathbf{V} , un insieme \mathbf{A} e un'applicazione $t : \mathbf{A} \times \mathbf{V} \rightarrow \mathbf{A}$ che soddisfa le condizioni (a),(b), resta definita su \mathbf{A} una struttura di spazio affine con spazio vettoriale associato \mathbf{V} . Infatti, per la (b), t individua un'applicazione $\mathbf{A} \times \mathbf{A} \rightarrow \mathbf{V}$ che soddisfa SA1, e che, per la (a), soddisfa anche SA2.

Sia \mathbf{A} uno spazio affine e $A, B, C, D \in \mathbf{A}$. Se $\overrightarrow{AB} = \overrightarrow{CD}$, allora $\overrightarrow{AC} = \overrightarrow{BD}$. Infatti si ha $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$ e $\overrightarrow{BD} = \overrightarrow{BC} + \overrightarrow{CD}$, e quindi $\overrightarrow{AC} - \overrightarrow{BD} = \overrightarrow{AB} - \overrightarrow{CD} = \mathbf{0}$, cioè $\overrightarrow{AC} = \overrightarrow{BD}$.

Equazioni parametriche di un sottospazio affine Sia \mathbf{A} uno spazio affine su \mathbf{V} in cui supponiamo fissato un riferimento affine $O\mathbf{e}_1, \dots, \mathbf{e}_n$. Sia \mathbf{S} il sottospazio affine passante per il punto $Q = (q_1, q_2, \dots, q_n)$ e parallelo al sottospazio vettoriale \mathbf{W} di \mathbf{V} . Sia

$$\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_s\} = \left\{ \begin{pmatrix} w_{11} \\ w_{21} \\ \vdots \\ w_{n1} \end{pmatrix}, \begin{pmatrix} w_{12} \\ w_{22} \\ \vdots \\ w_{n2} \end{pmatrix}, \dots, \begin{pmatrix} w_{1s} \\ w_{2s} \\ \vdots \\ w_{ns} \end{pmatrix} \right\}$$

una base di \mathbf{W} . Allora $\forall P = (x_1, x_2, \dots, x_n) \in \mathbf{S}$ e per opportuni $t_1, \dots, t_s \in \mathbb{K}$, si ha

$$\overrightarrow{QP} = t_1 \mathbf{w}_1 + t_2 \mathbf{w}_2 + \dots + t_s \mathbf{w}_s \implies \begin{pmatrix} x_1 - q_1 \\ x_2 - q_2 \\ \vdots \\ x_n - q_n \end{pmatrix} = t_1 \begin{pmatrix} w_{11} \\ w_{21} \\ \vdots \\ w_{n1} \end{pmatrix} + t_2 \begin{pmatrix} w_{12} \\ w_{22} \\ \vdots \\ w_{n2} \end{pmatrix} + \dots + t_s \begin{pmatrix} w_{1s} \\ w_{2s} \\ \vdots \\ w_{ns} \end{pmatrix}$$

da cui

$$\begin{array}{ccc} x_1 - q_1 = t_1 w_{11} + t_2 w_{12} + \dots + t_s w_{1s} & & x_1 = q_1 + t_1 w_{11} + t_2 w_{12} + \dots + t_s w_{1s} \\ x_2 - q_2 = t_1 w_{21} + t_2 w_{22} + \dots + t_s w_{2s} & \implies & x_2 = q_2 + t_1 w_{21} + t_2 w_{22} + \dots + t_s w_{2s} \\ \vdots & & \vdots \\ x_n - q_n = t_1 w_{n1} + t_2 w_{n2} + \dots + t_s w_{ns} & & x_n = q_n + t_1 w_{n1} + t_2 w_{n2} + \dots + t_s w_{ns} \end{array}$$

Queste equazioni, generalizzabili al seguente modo

$$\forall i = 1, \dots, n, \quad x_i = q_i + \sum_{k=1}^s t_k w_{ik} \quad (48)$$

sono chiamate **equazioni parametriche** di \mathbf{S} . Al variare dei parametri $t_1, \dots, t_s \in \mathbb{K}$ le (48) danno le coordinate di tutti i punti di \mathbf{S} . Si noti che le equazioni parametriche non sono univocamente determinate da \mathbf{S} , ma dipendono dalla scelta di Q e di $\mathbf{w}_1, \dots, \mathbf{w}_s$.

Equazioni parametriche di una retta

Caso 4. Nel caso in cui il sottospazio affine è una retta r , passante per il punto $Q = (q_1, q_2, \dots, q_n)$ e avente vettore di direzione $\mathbf{a} = (a_1, a_2, \dots, a_n)$, avremo

$$\overrightarrow{QP} = t\mathbf{a} \implies \begin{pmatrix} x_1 - q_1 \\ x_2 - q_2 \\ \vdots \\ x_n - q_n \end{pmatrix} = t \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \implies \begin{array}{l} x_1 = q_1 + ta_1 \\ x_2 = q_2 + ta_2 \\ \vdots \\ x_n = q_n + ta_n \end{array}$$

cioè

$$\forall i = 1, \dots, n \quad x_i = q_i + ta_i \quad (49)$$

le (49) sono **equazioni parametriche** della retta r . Quando la retta r è assegnata mediante due suoi punti distinti $Q = (q_1, q_2, \dots, q_n)$ e $Q' = (q'_1, q'_2, \dots, q'_n)$, $\overrightarrow{QQ'}$ è un vettore di direzione di r e nelle (49) si può prendere $\mathbf{a} = \overrightarrow{QQ'}$ cioè $(a_1, a_2, \dots, a_n) = (q'_1 - q_1, q'_2 - q_2, \dots, q'_n - q_n)$.

- Se $n = 2$, cioè se \mathbf{A} è un piano affine, le (49) si riducono a due equazioni. Le coordinate di \mathbf{a} si denotano con l, m , le coordinate del punto variabile con x, y e quelle del punto Q con a, b ; quindi avremo

$$\begin{array}{ccc} x_1 = q_1 + ta_1 & \implies & x = a + lt \\ x_2 = q_2 + ta_2 & & y = b + mt \end{array};$$

- Se $n = 3$, cioè se \mathbf{A} è uno spazio affine, le (49) si riducono a tre equazioni. Le coordinate di \mathbf{a} si denotano con l, m, n , le coordinate del punto variabile con x, y, z e quelle del punto Q con a, b, c ; quindi avremo

$$\begin{array}{ccc} x_1 = q_1 + ta_1 & \implies & x = a + lt \\ x_2 = q_2 + ta_2 & & y = b + mt \\ x_3 = q_3 + ta_3 & & z = c + nt \end{array}.$$

Un altro modo di rappresentare un sottospazio affine mediante equazioni è dato dal seguente teorema.

Equazioni cartesiane di un sottospazio affine

Teorema 3.5. 1. Sia

$$\begin{aligned} a_{11}X_1 + a_{12}X_2 + \dots + a_{1n}X_n &= b_1 \\ a_{21}X_1 + a_{22}X_2 + \dots + a_{2n}X_n &= b_2 \\ &\vdots \\ a_{t1}X_1 + a_{t2}X_2 + \dots + a_{tn}X_n &= b_t \end{aligned} \quad (50)$$

un sistema di equazioni lineari nelle incognite X_1, \dots, X_n . L'insieme \mathbf{S} dei punti di \mathbf{A} le cui coordinate sono soluzioni del sistema, se non è vuoto, è un sottospazio affine di dimensione $n - r$, dove r è il rango della matrice dei coefficienti del sistema. La giacitura di \mathbf{S} è il sottospazio vettoriale \mathbf{W} di \mathbf{V} avente per equazioni cartesiane il sistema omogeneo associato

$$\begin{aligned} a_{11}X_1 + a_{12}X_2 + \dots + a_{1n}X_n &= 0 \\ a_{21}X_1 + a_{22}X_2 + \dots + a_{2n}X_n &= 0 \\ &\vdots \\ a_{t1}X_1 + a_{t2}X_2 + \dots + a_{tn}X_n &= 0 \end{aligned} \quad ; \quad (51)$$

2. Per ogni sottospazio affine \mathbf{S} di \mathbf{A} di dimensione s esiste un sistema di $n - s$ equazioni lineari in n incognite le cui soluzioni sono le coordinate di tutti e soli i punti di \mathbf{S} .

Proof.

1. Siano $A \in M_{t,n}(\mathbb{K})$, $\mathbf{b} \in M_{t,1}(\mathbb{K})$. Per ipotesi $\exists Q = (q_1, q_2, \dots, q_n) \in \mathbf{S}$. Avremo che $\forall P = (x_1, x_2, \dots, x_n) \in \mathbf{S}$,

$$A\overrightarrow{QP} = A(P - Q) = AP - AQ = \mathbf{b} - \mathbf{b} = \mathbf{0},$$

cioè \overrightarrow{QP} è soluzione di (51), ossia $\overrightarrow{QP} \in \mathbf{W}$, quindi \mathbf{S} è contenuto nel sottospazio affine Σ passante per Q e parallelo a \mathbf{W} . Viceversa, se $P = (x_1, x_2, \dots, x_n) \in \Sigma$, allora $\overrightarrow{QP} \in \mathbf{W}$ e quindi le coordinate di \overrightarrow{QP} sono soluzione delle (51). Si avrà dunque

$$\mathbf{0} = A\overrightarrow{QP} = AP - AQ \implies AP = AQ,$$

da cui $P \in \mathbf{S}$. Quindi $S = \Sigma$ ed S è un sottospazio affine.

2. Supponiamo che $S \subset A$ sia il sottospazio affine passante per il punto $Q = (q_1, q_2, \dots, q_n)$ e avente giacitura \mathbf{W} . Siano

$$\begin{aligned} a_{11}X_1 + a_{12}X_2 + \dots + a_{1n}X_n &= 0 \\ a_{21}X_1 + a_{22}X_2 + \dots + a_{2n}X_n &= 0 \\ &\vdots \\ a_{n-s+1}X_1 + a_{n-s+2}X_2 + \dots + a_{n-s+n}X_n &= 0 \end{aligned}$$

equazioni cartesiane di \mathbf{W} . I punti $P = (x_1, x_2, \dots, x_n) \in \mathbf{S}$ sono caratterizzati dalla condizione $\overrightarrow{QP} \in \mathbf{W}$, cioè $A\overrightarrow{QP} = \mathbf{0} \in M_{n-s+1}(\mathbb{K})$ ovvero $AP = \mathbf{b}$ con $\mathbf{b} = AQ$. Dunque i punti $P \in \mathbf{S}$ sono precisamente quei punti di \mathbf{A} le cui coordinate sono soluzioni del sistema

$$\begin{aligned} a_{11}X_1 + a_{12}X_2 + \dots + a_{1n}X_n &= b_1 \\ a_{21}X_1 + a_{22}X_2 + \dots + a_{2n}X_n &= b_2 \\ &\vdots \\ a_{n-s+1}X_1 + a_{n-s+2}X_2 + \dots + a_{n-s+n}X_n &= b_{n-s} \end{aligned}$$

□

Le (50) si dicono **equazioni cartesiane del sottospazio \mathbf{S} rispetto al riferimento $O\mathbf{e}_1 \dots \mathbf{e}_n$** . È evidente che le (50) non sono univocamente determinate da \mathbf{S} : due sistemi di equazioni lineari definiscono lo stesso sottospazio affine di \mathbf{A} se e solo se sono equivalenti.

Il sottospazio \mathbf{S} di equazioni cartesiane (50) contiene l'origine se e solo se il sistema è omogeneo. Quindi ogni sistema di equazioni lineari omogenee (51), oltre a definire un sottospazio vettoriale \mathbf{W} di \mathbf{V} , definisce anche un sottospazio affine \mathbf{S} di \mathbf{A} passante per l'origine e avente per giacitura \mathbf{W} . Per ogni punto $P \in \mathbf{S}$ il vettore $\mathbf{w} \in \mathbf{W}$ avente le stesse coordinate di P appartiene alla giacitura \mathbf{W} di \mathbf{S} ; si ottiene così una corrispondenza biunivoca tra \mathbf{S} e la sua giacitura \mathbf{W} .

Dal teorema (3.5.2) segue che ogni iperpiano \mathbf{H} di \mathbf{A} si rappresenta con un'equazione cartesiana $a_1X_1 + \dots + a_nX_n = b$,

- Se $\dim(\mathbf{A}) = 2$, ogni retta di \mathbf{A} ha un'equazione cartesiana della forma

$$a_1X_1 + a_2X_2 = b;$$

- Se $\dim(\mathbf{A}) = 3$, ogni piano di \mathbf{A} ha un'equazione cartesiana della forma

$$a_1X_1 + a_2X_2 + a_3X_3 = b.$$

L'iperpiano \mathbf{H} dello spazio affine \mathbf{A} contiene l'origine se e solo se $b = 0$. In particolare $\forall i = 1, \dots, n$, l'iperpiano di equazione $X_i = 0$ passa per l'origine; esso si dice ***i*-esimo iperpiano coordinato**.

Sottospazi affini paralleli

Definizione 3.20. Siano \mathbf{S} e \mathbf{T} due sottospazi affini di \mathbf{A} , aventi giaciture \mathbf{W} ed \mathbf{U} rispettivamente, di dimensione maggiore di zero. \mathbf{S} e \mathbf{T} si dicono **paralleli** se $\mathbf{W} \subset \mathbf{U}$ oppure $\mathbf{U} \subset \mathbf{W}$. Scriveremo $\mathbf{S} \parallel \mathbf{T}$.

Se $\dim(\mathbf{S}) = \dim(\mathbf{T})$, i due sottospazi affini \mathbf{S} e \mathbf{T} sono paralleli se e solo se le loro $\mathbf{W} = \mathbf{U}$; in particolare due sottospazi affini uguali sono paralleli.

Nel caso in cui \mathbf{S} e \mathbf{T} sono due rette il loro parallelismo significa che hanno la stessa direzione, cioè che hanno vettori di direzione proporzionali.

Se il sottospazio affine \mathbf{S} è contenuto in un altro sottospazio affine \mathbf{T} , allora $\mathbf{S} \parallel \mathbf{T}$: infatti al variare $P, Q \in \mathbf{S}$, il vettore \overrightarrow{QP} descrive la giacitura \mathbf{W} di \mathbf{S} , ma è anche contenuto in quella \mathbf{U} di \mathbf{T} , perchè $P, Q \in \mathbf{T}$. Quindi $\mathbf{W} \subset \mathbf{U}$.

Proposizione 3.6. Siano $\mathbf{S}, \mathbf{T} \subset \mathbf{A}$ due sottospazi affini paralleli, tali che $\dim(\mathbf{S}) \leq \dim(\mathbf{T})$.

1. $\mathbf{S} \cap \mathbf{T} \neq \emptyset \implies \mathbf{S} \subset \mathbf{T}$;
2. $(\dim(\mathbf{S}) = \dim(\mathbf{T}) \wedge \mathbf{S} \cap \mathbf{T} \neq \emptyset) \implies \mathbf{S} = \mathbf{T}$.

Proof.

1. Sia $Q \in \mathbf{S} \cap \mathbf{T}$, $\forall P \in \mathbf{S}$, $\overrightarrow{QP} \in \mathbf{W} \subset \mathbf{U}$, quindi $P \in \mathbf{T}$; dunque $\mathbf{S} \subset \mathbf{T}$.
2. $\dim(\mathbf{S}) = \dim(\mathbf{T})$, allora $\mathbf{W} = \mathbf{U}$, si deduce che $\mathbf{S} \subset \mathbf{T}$ e $\mathbf{T} \subset \mathbf{S}$, cioè $\mathbf{S} = \mathbf{T}$.

□

Corollario 7. Se \mathbf{S} è un sottospazio affine di \mathbf{A} e $P \in \mathbf{A}$, esiste un unico sottospazio affine \mathbf{T} contenente P , parallelo a \mathbf{S} e tale che $\dim(\mathbf{T}) = \dim(\mathbf{S})$.

Nel caso in cui \mathbf{A} è il piano affine ordinario ed \mathbf{S} una retta, il corollario è equivalente al quinto postulato della geometria euclidea piana, il cosiddetto postulato delle parallele, che afferma che per un punto esterno ad una retta passa una ed una sola retta parallela alla retta data. Gli assiomi di piano affine implicano quindi la validità di tale postulato in un piano affine qualunque.

Spazio vettoriale quoziente di \mathbf{V} di modulo \mathbf{W} Siano \mathbf{V} un \mathbb{K} -spazio vettoriale e \mathbf{W} un sottospazio vettoriale di \mathbf{V} . Da 3.6(2) segue che, se $\mathbf{v}, \mathbf{v}' \in \mathbf{V}$, i due sottospazi affini $\mathbf{v} + \mathbf{W}$ e $\mathbf{v}' + \mathbf{W}$ di $\mathbf{V}_{\mathbf{a}}$ o coincidono oppure sono disgiunti. Da ciò segue che la famiglia dei sottospazi affini di $\mathbf{V}_{\mathbf{a}}$ aventi giacitura \mathbf{W} è una partizione di \mathbf{V} . L'insieme quoziente di \mathbf{V} rispetto a tale partizione, cioè l'insieme i cui elementi sono i sottospazi affini di $\mathbf{V}_{\mathbf{a}}$ aventi giacitura \mathbf{W} , si denoterà al seguente modo:

$$\mathbf{V}/\mathbf{W} = \{\mathbf{v} + \mathbf{W} : \mathbf{v} \in \mathbf{V}\}.$$

Si noti che $\mathbf{v} + \mathbf{W} = \mathbf{v}' + \mathbf{W}$ se e solo se $\mathbf{v}' \in \mathbf{v} + \mathbf{W}$, perchè ciò equivale alla condizione

$$(\mathbf{v} + \mathbf{W}) \cap (\mathbf{v}' + \mathbf{W}) \neq \emptyset$$

cioè all'esistenza di $\mathbf{w}, \mathbf{w}' \in \mathbf{W}$ tali che $\mathbf{v} + \mathbf{w} = \mathbf{v}' + \mathbf{w}'$, cioè alla condizione

$$\mathbf{v}' = \mathbf{v} + (\mathbf{w} - \mathbf{w}') \in \mathbf{v} + \mathbf{W}.$$

Definiamo un'operazione di somma in \mathbf{V}/\mathbf{W} ponendo

$$(\mathbf{v}_1 + \mathbf{W}) + (\mathbf{v}_2 + \mathbf{W}) = (\mathbf{v}_1 + \mathbf{v}_2) + \mathbf{W}.$$

Quest'operazione è ben definita, perchè non dipende dai vettori \mathbf{v}_1 e \mathbf{v}_2 che sono stati scelti per rappresentare i due sottospazi. Infatti, se

$$\mathbf{v}_1 + \mathbf{W} = \mathbf{v}'_1 + \mathbf{W}, \quad \mathbf{v}_2 + \mathbf{W} = \mathbf{v}'_2 + \mathbf{W},$$

allora

$$\mathbf{v}'_1 = \mathbf{v}_1 + \mathbf{w}_1, \mathbf{v}'_2 = \mathbf{v}_2 + \mathbf{w}_2$$

per qualche $\mathbf{w}_1, \mathbf{w}_2 \in \mathbf{W}$ e quindi

$$(\mathbf{v}'_1 + \mathbf{v}'_2) + \mathbf{W} = (\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{w}_1 + \mathbf{w}_2) + \mathbf{W} = (\mathbf{v}_1 + \mathbf{v}_2) + \mathbf{W}.$$

Definiamo \mathbf{V}/\mathbf{W} un'operazione di moltiplicazione per elementi di \mathbb{K} ponendo

$$c(\mathbf{v} + \mathbf{W}) = c\mathbf{v} + \mathbf{W}$$

Anche quest'operazione è ben definita, perchè $\mathbf{v} + \mathbf{W} = \mathbf{v}' + \mathbf{W}$ significa $\mathbf{v}' = \mathbf{v} + \mathbf{w}, \mathbf{w} \in \mathbf{W}$, e quindi

$$c\mathbf{v}' + \mathbf{W} = c(\mathbf{v} + \mathbf{w}) + \mathbf{W} = (c\mathbf{v} + c\mathbf{w}) + \mathbf{W} = c\mathbf{v} + \mathbf{W}.$$

Lo zero rispetto alla somma in \mathbf{V}/\mathbf{W} è il sottospazio $\mathbf{W} = \mathbf{0} + \mathbf{W}$, come si vede subito. È immediato verificare che \mathbf{V}/\mathbf{W} è un \mathbb{K} -spazio vettoriale, che si chiama **spazio vettoriale quoziente di \mathbf{V} modulo \mathbf{W}** .

Si noti che se $\mathbf{V} = \mathbf{W}$, lo spazio \mathbf{V}/\mathbf{W} possiede un unico elemento, \mathbf{V} stesso, e quindi si identifica con lo spazio costituito dal solo vettore nullo. Se invece $\mathbf{W} = \langle \mathbf{0} \rangle$, i sottospazi affini di giacitura \mathbf{W} sono i punti di \mathbf{V}_a , cioè lo spazio \mathbf{V}/\mathbf{W} coincide con \mathbf{V} .

Sottospazi Sghembi o Incidenti

Definizione 3.21. Due sottospazi affini \mathbf{S} e \mathbf{T} di \mathbf{A} si dicono **sghembi (incidenti)** se sono privi di punti in comune (se hanno almeno un punto in comune).

Consideriamo due sottospazi affini \mathbf{S} e \mathbf{T} , con $\dim(\mathbf{S}) = s, \dim(\mathbf{T}) = t$. Supponiamo che essi abbiano equazioni cartesiane:

$$\mathbf{S}: \sum_{j=1}^n m_{ij} X_j = b_i, \quad i = 1, \dots, n-s, \quad (52)$$

$$\mathbf{T}: \sum_{j=1}^n n_{kj} X_j = c_k, \quad k = 1, \dots, n-t. \quad (53)$$

L'intersezione $\mathbf{S} \cap \mathbf{T}$ è il luogo dei punti di \mathbf{A} le cui coordinate sono simultaneamente soluzioni delle equazioni cartesiane di \mathbf{S} e di \mathbf{T} ; quindi i punti di $\mathbf{S} \cap \mathbf{T}$ si ottengono in corrispondenza delle soluzioni del sistema

$$\begin{aligned} \sum_{j=1}^n m_{ij} X_j &= b_i, \quad i = 1, \dots, n-s, \\ \sum_{j=1}^n n_{kj} X_j &= c_k, \quad k = 1, \dots, n-t. \end{aligned} \quad (54)$$

Per il teorema 3.5 il sistema (54), se ha soluzioni, rappresenta un sottospazio affine. Quindi, se non è vuoto, $\mathbf{S} \cap \mathbf{T}$ è un sottospazio affine di \mathbf{A} . La dimensione di $\mathbf{S} \cap \mathbf{T}$ è uguale a $n-r$, dove r è il rango della matrice dei coefficienti del sistema (54).

In particolare si ha

$$r \leq (n-s) + (n-t) = 2n - (s+t), \quad (55)$$

da cui, se $\mathbf{S} \cap \mathbf{T} \neq \emptyset$,

$$\dim(\mathbf{S} \cap \mathbf{T}) = n-r \geq n - [2n - (s+t)] = s+t-n. \quad (56)$$

D'altra parte, essendo $\mathbf{S} \cap \mathbf{T}$ contenuto sia in \mathbf{S} che in \mathbf{T} , la sua dimensione non può superare il minimo tra s e t . Riassumendo quanto detto, possiamo enunciare la seguente proposizione.

Proposizione 3.7. *L'intersezione $\mathbf{S} \cap \mathbf{T}$ di due sottospazi affini \mathbf{S} e \mathbf{T} di \mathbf{A} , se non è vuota, è un sottospazio affine tale che*

$$\dim(\mathbf{S}) + \dim(\mathbf{T}) - \dim(\mathbf{A}) \leq \dim(\mathbf{S} \cap \mathbf{T}) \leq \min\{\dim(\mathbf{S}), \dim(\mathbf{T})\}. \quad (57)$$

Dalla dimostrazione segue che nella (57) sussiste la seconda uguaglianza se e solo se $\mathbf{S} \subset \mathbf{T}$ oppure $\mathbf{T} \subset \mathbf{S}$. Se invece \mathbf{S} e \mathbf{T} sono paralleli e disgiunti, allora $\mathbf{S} \cap \mathbf{T} = \emptyset$.

Il caso in cui vale la prima uguaglianza nella (57) è considerato nella seguente proposizione.

Proposizione 3.8. *Siano \mathbf{S} e \mathbf{T} sottospazi affini di \mathbf{A} di giaciture \mathbf{W} e \mathbf{U} rispettivamente. Allora $\mathbf{S} \cap \mathbf{T} \neq \emptyset$ e*

$$\dim(\mathbf{S} \cap \mathbf{T}) = \dim(\mathbf{S}) + \dim(\mathbf{T}) - \dim(\mathbf{A}) \quad (58)$$

se e solo se $\mathbf{V} = \mathbf{W} + \mathbf{U}$.

Proof. Supponiamo che \mathbf{S} e \mathbf{T} abbiano equazioni cartesiane (52) e (53) rispettivamente. La (58) sussiste precisamente quando $\mathbf{S} \cap \mathbf{T} \neq \emptyset$ e si ha l'uguaglianza nella (55). La (55) si può leggere come una relazione tra $s = \dim(\mathbf{W})$, $t = \dim(\mathbf{U})$ e $\dim(\mathbf{W} \cap \mathbf{U}) = \dim(\mathbf{S} \cap \mathbf{T})$, ed è una conseguenza diretta della formula di Grassmann. Più precisamente, la formula di Grassmann afferma che la (55) è un'uguaglianza se e solo se $n = \dim(\mathbf{W} + \mathbf{U})$, cioè se solo se $\mathbf{V} = \mathbf{W} + \mathbf{U}$. Per concludere è quindi sufficiente mostrare che in questo caso il sistema (54) è compatibile. Ma la matrice dei coefficienti del sistema (54) ha rango $r = (n - s) + (n - t) = 2n - (s + t)$, che è il numero delle sue righe, e la matrice omogenea ha lo stesso numero di righe e quindi lo stesso rango: il sistema è compatibile per il teorema di Kronecker-Rouché-Capelli. \square

Applicate ai casi $\dim(\mathbf{A}) = 2, 3$, le proposizioni 3.7 e 3.8 permettono di ritrovare i risultati sulla posizione reciproca di rette e piani che sappiamo essere veri nel piano e nello spazio ordinari.

Un caso particolare importante della proposizione 3.8 è quello in cui le giaciture \mathbf{W} e \mathbf{U} dei sottospazi affini \mathbf{S} e \mathbf{T} sono supplementari in \mathbf{V} , ovvero tali che $\mathbf{V} = \mathbf{W} \oplus \mathbf{U}$: poichè in questo caso $\dim(\mathbf{S}) + \dim(\mathbf{T}) = \dim(\mathbf{A})$, da (58) afferma che \mathbf{S} e \mathbf{T} hanno un solo punto in comune. Ciò è quanto avviene ad esempio per due rette non parallele in un piano affine, oppure per un piano e una retta non paralleli in uno spazio di dimensione 3.

Da quest'osservazione si deduce la seguente costruzione geometrica. Sia \mathbf{S} un sottospazio affine di dimensione s di \mathbf{A} , e sia \mathbf{W} la sua giacitura. Fissiamo un sottospazio vettoriale \mathbf{U} di \mathbf{V} tale che $\mathbf{W} \oplus \mathbf{U} = \mathbf{V}$. Per il corollario 7, per ogni punto $P \in \mathbf{A}$ esiste un unico sottospazio affine $\mathbf{T}_{P,\mathbf{U}}$ contenente P avente giacitura \mathbf{U} . Dalla (58) segue che $\mathbf{S} \cap \mathbf{T}_{P,\mathbf{U}} = \{Q\}$. L'applicazione $p_{\mathbf{S},\mathbf{U}} : \mathbf{A} \rightarrow \mathbf{S}$, definita ponendo $p_{\mathbf{S},\mathbf{U}}(P) = Q$, è chiamata **proiezione di \mathbf{A} su \mathbf{S} parallela a \mathbf{U}** . Se $\dim(\mathbf{U}) = 1$, allora \mathbf{S} è un iperpiano e $p_{\mathbf{S},\mathbf{U}}$ è la proiezione di \mathbf{A} su \mathbf{S} nella direzione \mathbf{U} .

3.2 Geometria in un Piano Affine

Siano \mathbf{V} un \mathbb{K} -spazio vettoriale tale che $\dim(\mathbf{V}) = 2$ e \mathbf{A} un piano affine su \mathbf{V} . Fissiamo un riferimento affine $O\mathbf{e}_1\mathbf{e}_2$ in \mathbf{A} e denotiamo con x, y le coordinate di un punto variabile $P \in \mathbf{A}$ e con X, Y indeterminate. I sottospazi affini di \mathbf{A} sono le rette, oltre ad \mathbf{A} stesso e ai punti di \mathbf{A} . Ogni retta \mathbf{r} di \mathbf{A} ha equazioni parametriche

$$\begin{aligned} x &= a + lt \\ y &= b + mt \end{aligned} \quad (59)$$

in cui $Q(a, b)$ è un punto qualsiasi di \mathbf{r} e $\mathbf{v} = (l, m)$ è un vettore di direzione di \mathbf{r} . Poiché ha codimensione 1 in \mathbf{A} , la retta \mathbf{r} si può anche rappresentare per mezzo di un'equazione cartesiana

$$AX + BY + C = 0, \quad (60)$$

per opportuni $A, B, C \in \mathbb{K}$ tali che $(A, B) \neq (0, 0)$. Le costanti A, B, C sono individuate da \mathbf{r} solo a meno di un comune fattore di proporzionalità non nullo, e quindi una retta \mathbf{r} possiede infinite equazioni 60, tutte tra loro proporzionali. La retta \mathbf{r} di equazione 60 contiene l'origine O se e solo se $C = 0$. Le rette di equazioni $X = 0$ e $Y = 0$ si dicono **assi coordinati**, rispettivamente **asse delle Y** e **asse delle X**.

- $A, B \neq 0$
 - $C \neq 0$ $\mathbf{r} = \{(x, y) \in \mathbb{R}^2 : Ax + By + C = 0\}$
 - $C = 0$ $\mathbf{r} = \{(x, y) \in \mathbb{R}^2 : Ax + By = 0\}$
- $A \neq 0, B = 0$
 - $C \neq 0$ $\mathbf{r} = \{(-C/A, y) : y \in \mathbb{R}\}$
 - $C = 0$ $\mathbf{r} = \{(0, y) : y \in \mathbb{R}\}$
- $A = 0, B \neq 0$
 - $C \neq 0$ $\mathbf{r} = \{(x, -C/B) : x \in \mathbb{R}\}$
 - $C = 0$ $\mathbf{r} = \{(x, 0) : x \in \mathbb{R}\}$
- $A, B = 0$
 - $C \neq 0$ $\mathbf{r} = \{(x, y) \in \mathbb{R}^2 : C = 0\}$ Viene dunque descritto un insieme vuoto, che non è una retta.
 - $C = 0$ $\mathbf{r} = \{(x, y) \in \mathbb{R}^2 : 0 = 0\}$ In questo caso viene descritto tutto il piano, che ancora una volta non è una retta.

Passaggio da equazioni parametriche a cartesiane Per ottenere un'equazione cartesiana della retta \mathbf{r} a partire da sue equazioni parametriche 59, cioè una volta noti un suo punto $Q(a, b)$ e un suo vettore di direzione $\mathbf{v} = (l, m)$, si può osservare che le 59 esprimono la condizione che il vettore \overrightarrow{QP} sia parallelo a \mathbf{v} , e che la stessa condizione può esprimersi imponendo che (x, y) sia soluzione dell'equazione in X, Y :

$$\begin{vmatrix} X - a & Y - b \\ l & m \end{vmatrix} = 0 \quad (61)$$

cioè

$$m(X - a) - l(Y - b) = 0, \quad (62)$$

o, equivalentemente,

$$mX - lY + lb - ma = 0. \quad (63)$$

L'equazione 61 è soddisfatta da tutti e soli i punti $P(x, y) \in \mathbf{r}$. Quindi la 61, oppure la 62 o la 63, è un'equazione cartesiana di \mathbf{r} .

Se la retta \mathbf{r} è assegnata mediante due suoi punti distinti $Q(a, b)$ e $Q'(a', b')$, allora come vettore di direzione si può prendere $\mathbf{v} = (a' - a, b' - b) = \overrightarrow{QQ'}$ e la 61 prende la forma

$$\begin{vmatrix} X - a & Y - b \\ a' - a & b' - b \end{vmatrix} = 0. \quad (64)$$

Proposizione 3.9. Siano \mathbf{r} ed \mathbf{r}' due rette di \mathbf{A} di equazioni cartesiane

$$AX + BY + C = 0 \quad A'X + B'Y + C' = 0,$$

rispettivamente. Allora:

1. \mathbf{r} ed \mathbf{r}' sono parallele se e solo se la matrice

$$\begin{pmatrix} A & B \\ A' & B' \end{pmatrix} \quad (65)$$

ha rango 1, cioè se $AB' - A'B = 0$.

2. Se \mathbf{r} ed \mathbf{r}' sono parallele, allora sono disgiunte oppure coincidono a seconda che la matrice

$$\begin{pmatrix} A & B & C \\ A' & B' & C' \end{pmatrix} \quad (66)$$

abbia rango 2 oppure 1.

3. \mathbf{r} ed \mathbf{r}' hanno uno e un solo punto in comune se e solo se la matrice 65 ha rango 2. In questo caso il punto $\mathbf{r} \cap \mathbf{r}'$ ha coordinate

$$x_0 = \frac{CB' - C'B}{AB' - A'B}, \quad y_0 = \frac{AC' - A'C}{AB' - A'B}. \quad (67)$$

Proof.

1. Le direzioni di \mathbf{r} e di \mathbf{r}' sono determinate dalle equazioni omogenee

$$AX + BY = 0 \quad A'X + B'Y = 0,$$

rispettivamente. Quindi \mathbf{r} ed \mathbf{r}' sono parallele se e solo se

$$\begin{cases} A' = \rho A \\ B' = \rho B \end{cases} \quad \rho \in \mathbb{K} \setminus \{0\},$$

cioè se e solo se la 65 ha rango 1.

2. La matrice 66 ha rango 1 se e solo se le equazioni di \mathbf{r} e di \mathbf{r}' sono proporzionali, cioè se e solo se $\mathbf{r} \equiv \mathbf{r}'$. Quindi, se la 66 ha rango 2 e la 65 ha rango 1, \mathbf{r} ed \mathbf{r}' sono distinte e, per la (1), parallele: pertanto non hanno punti in comune.
3. Infine, la 65 ha rango 2 se e solo se \mathbf{r} ed \mathbf{r}' non sono parallele. In questo caso, per il teorema di Cramer, il sistema costituito dalle equazioni di \mathbf{r} e di \mathbf{r}' ha un'unica soluzione (x_0, y_0) data dalla 67.

□

I casi che vengono considerati nell'enunciato della proposizione 3.9 sono tutti quelli che si possono presentare per la posizione reciproca di due rette di \mathbf{A} , perché corrispondono a tutte le possibilità per un sistema di due equazioni lineari in due incognite. In particolare vediamo che l'unica possibilità perché due rette \mathbf{r} ed \mathbf{r}' non abbiano punti in comune è che la 65 abbia rango 1 e la 66 abbia rango 2, e questo caso corrisponde al parallelismo di \mathbf{r} ed \mathbf{r}' . Quindi due rette di un piano affine non possono essere sghembe.

Fascio proprio di rette

Definizione 3.22. Sia $Q(x_0, y_0) \in \mathbf{A}$. L'insieme Φ i cui elementi sono le rette di \mathbf{A} passanti per Q si dice **fascio proprio di rette** e Q si dice il suo **centro**.

Siano

$$AX + BY + C = 0 \quad A'X + B'Y + C' = 0$$

equazioni cartesiane di due rette distinte, aventi in comune il centro Q di Φ . Siano $\lambda, \mu \in \mathbb{K}$ due scalari non entrambi nulli; la retta di equazione cartesiana

$$\lambda(AX + BY + C) + \mu(A'X + B'Y + C') = 0 \quad (68)$$

passa per Q e quindi appartiene al fascio Φ .

Viceversa, se $\mathbf{r} \in \Phi$ e $P(x, y) \in \mathbf{r}$ è un punto diverso da Q , la condizione

$$\lambda(AX + BY + C) + \mu(A'X + B'Y + C') = 0 \quad (69)$$

è un'equazione lineare omogenea in λ e μ che ha un'unica soluzione, determinata a meno di un fattore di proporzionalità: la retta 68 in cui λ e μ soddisfano la 69 appartiene a Φ e passa per P , quindi coincide con \mathbf{r} .

Pertanto tutte le rette di Φ si rappresentano nella forma 68; una retta di Φ determina λ, μ solo a meno di un fattore di proporzionalità.

Un caso particolare della 68 si ha prendendo \mathbf{r} ed \mathbf{r}' parallele agli assi: si ottiene che ogni retta del fascio Φ si può scrivere nella forma

$$\lambda(X - x_0) + \mu(Y - y_0) = 0$$

per opportuni λ, μ .

I fasci di rette sono utili in pratica soprattutto quando il punto Q è individuato da due rette che lo contengono ma le sue coordinate non sono note, e si vuole individuare una retta contenente Q soddisfacente a condizioni assegnate, ad esempio la condizione di passare per un punto P diverso da Q , oppure di essere parallela ad una retta assegnata.

È talvolta più conveniente utilizzare un solo parametro non omogeneo t , invece della coppia di parametri omogenei λ, μ . In altre parole, anziché nella forma 68, si può scrivere la retta variabile nel fascio nella forma

$$AX + BY + C + t(A'X + B'Y + C') = 0, \quad (70)$$

e la condizione di passaggio 69 diventa un'equazione lineare in t :

$$Ax + By + C + t(A'x + B'y + C') = 0, \quad (71)$$

la cui soluzione, se esiste, determina la retta cercata.

Si faccia però attenzione: nella forma 71 si rappresentano tutte le rette del fascio Φ ad eccezione della retta

$$A'X + B'Y + C' = 0. \quad (72)$$

Ciò corrisponde al fatto che l'equazione 71 è incompatibile se (e solo se)

$$A'x + B'y + C' = 0$$

e

$$Ax + By + C \neq 0,$$

cioè se $P \notin Q$ e P appartiene alla retta 72.

Con questa avvertenza, si può sempre utilizzare un parametro non omogeneo per determinare una retta del fascio Φ , purché si interpreti nel modo detto l'equazione 71 quando essa risulti incompatibile.

Considerazioni simili alle precedenti si possono fare nel caso dei "fasci impropri".

Fascio improprio di rette

Definizione 3.23. Dato un vettore non nullo $\mathbf{v} \in \mathbf{V}$, l'insieme i cui elementi sono le rette di \mathbf{A} aventi direzione $\langle \mathbf{v} \rangle$ è un **fascio improprio di rette**, e $\langle \mathbf{v} \rangle$ è la sua direzione.

Data una retta del fascio improprio Φ , di equazione

$$AX + BY + C = 0,$$

ogni altra retta di Φ è della forma

$$AX + BY + t = 0$$

al variare del parametro $t \in \mathbb{K}$. Ciò segue immediatamente dalla proposizione 3.9.

Per ogni punto $P(x, y) \in \mathbb{A}$ esiste un'unica retta di Φ contenente P , che è quella corrispondente al valore

$$t = -(Ax + By).$$

Si noti che, mentre per individuare un fascio proprio sono necessarie due sue rette, un fascio improprio è individuato una volta assegnata una sola sua retta.

Esempio 3.3. Sia $C(x_0, y_0) \in \mathbf{A}$ e sia $P(x, y)$ un altro punto qualsiasi. Il punto simmetrico di P rispetto a C è il punto $\sigma_C(P)$ tale che $\overrightarrow{C\sigma_C(P)} = \overrightarrow{PC}$. Le coordinate x', y' di $\sigma_C(P)$ sono individuate dalla condizione

$$(x' - x_0, y' - y_0) = (x_0 - x, y_0 - y)$$

e quindi sono

$$(x', y') = (2x_0 - x, 2y_0 - y).$$

Sia \mathbf{r} la retta di equazioni parametriche

$$x = a + lt \quad y = b + mt.$$

Se $P(a + lt, b + mt) \in \mathbf{r}$, il punto $\sigma_C(P)$ è

$$(2x_0 - a - lt, 2y_0 - b - mt).$$

Da ciò discende che al variare di $P \in \mathbf{r}$ il punto $\sigma_C(P)$ descrive la retta di equazioni parametriche

$$x = 2x_0 - a - lt \quad y = 2y_0 - b - mt.$$

che denotiamo con $\sigma_C(\mathbf{r})$. Si noti che $\sigma_C(\mathbf{r})$ passa per $\sigma_C(Q)$, dove $Q(a, b) \in \mathbf{r}$, ed è parallela a \mathbf{r} , avendo un vettore di direzione di coordinate $(-l, -m)$.

Teorema di Talete

Teorema 3.10. Siano $\mathbf{H}, \mathbf{H}', \mathbf{H}''$ rette parallele e distinte del piano affine \mathbf{A} , ed $\mathbf{r}_1, \mathbf{r}_2$ due rette non parallele ad $\mathbf{H}, \mathbf{H}', \mathbf{H}''$. Siano inoltre

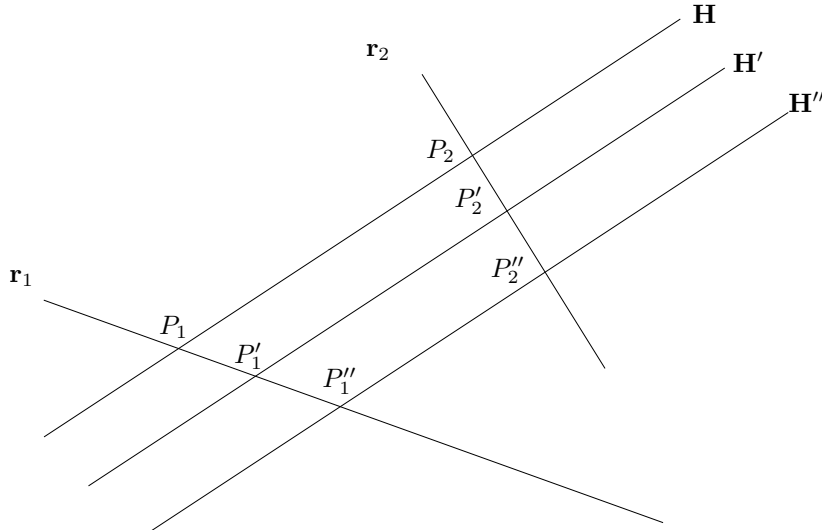
$$P_i = \mathbf{r}_i \cap \mathbf{H}, \quad P'_i = \mathbf{r}_i \cap \mathbf{H}', \quad P''_i = \mathbf{r}_i \cap \mathbf{H}'', \quad i = 1, 2,$$

e $k_1, k_2 \in \mathbb{K}$ tali che

$$\overrightarrow{P_i P''_i} = k_i \overrightarrow{P_i P'_i}, \quad i = 1, 2.$$

Allora $k_1 = k_2$.

Proof. Supponiamo $\mathbf{r}_1 \neq \mathbf{r}_2$, perché se $\mathbf{r}_1 = \mathbf{r}_2$ il teorema è banalmente vero. Salvo scambiare tra loro due delle rette $\mathbf{H}, \mathbf{H}', \mathbf{H}''$, possiamo supporre $P_1 \neq P_2$.



Poniamo $\mathbf{v} = \overrightarrow{P_1 P_2}$. Si ha

$$\overrightarrow{P_2 P'_2} - \overrightarrow{P_1 P'_1} = \overrightarrow{P'_1 P'_2} - \overrightarrow{P_1 P'_2} = \alpha \mathbf{v}, \quad \overrightarrow{P_2 P''_2} - \overrightarrow{P_1 P''_1} = \overrightarrow{P''_1 P''_2} - \overrightarrow{P_1 P''_2} = \beta \mathbf{v}$$

per opportuni $\alpha, \beta \in \mathbb{K}$. Se $\alpha = 0$, \mathbf{r}_1 ed \mathbf{r}_2 sono parallele; allora anche $\beta = 0$, perché se fosse $\beta \neq 0$ il vettore \mathbf{v} sarebbe parallelo a \mathbf{r}_1 e a \mathbf{r}_2 , contro l'ipotesi. In questo caso:

$$\overrightarrow{P_2 P_2''} = k_2 \overrightarrow{P_2 P_2'} = k_2 \overrightarrow{P_1 P_1'} \quad \overrightarrow{P_1 P_1''} = k_1 \overrightarrow{P_1 P_1'}$$

e poiché $\overrightarrow{P_1 P_1''} = k_1 \overrightarrow{P_1 P_1'}$ otteniamo $k_1 = k_2$. Se $\alpha \neq 0$ allora

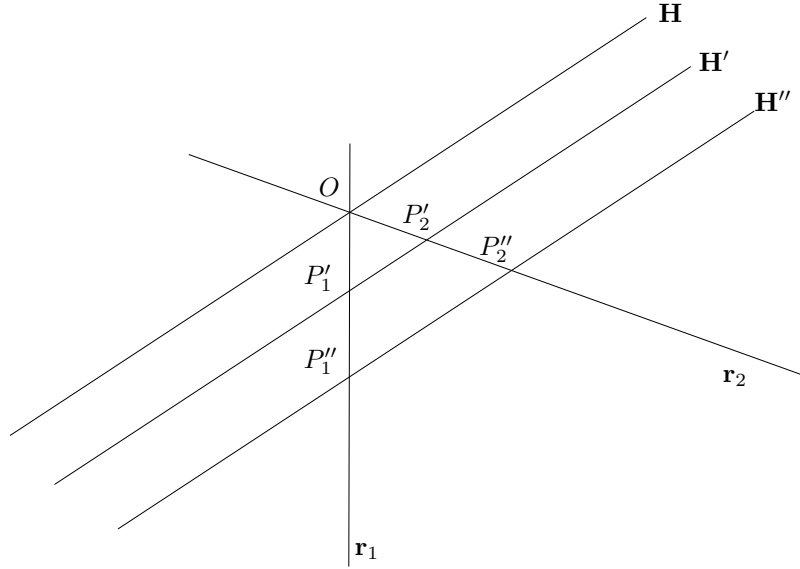
$$\overrightarrow{P_2 P_2''} - \overrightarrow{P_1 P_1''} = \beta \mathbf{v} = \alpha^{-1} \beta (\alpha \mathbf{v}) = \alpha^{-1} \beta \overrightarrow{P_2 P_2'} - \alpha^{-1} \beta \overrightarrow{P_1 P_1'}. \quad (73)$$

D'altra parte si ha anche

$$\overrightarrow{P_2 P_2''} - \overrightarrow{P_1 P_1''} = k_2 \overrightarrow{P_2 P_2'} - k_1 \overrightarrow{P_1 P_1'}. \quad (74)$$

Poiché $\alpha \neq 0$, $\overrightarrow{P_1 P_1'}$ e $\overrightarrow{P_2 P_2'}$ non sono paralleli, e quindi sono linearmente indipendenti. Confrontando 73 e 74 deduciamo che $k_2 = \alpha^{-1} \beta = k_1$. \square

Il teorema di Talete afferma essenzialmente che lo scalare $k = k_1 = k_2$ dipende solo da $\mathbf{H}, \mathbf{H}', \mathbf{H}''$ e non dalle rette trasversali $\mathbf{r}_1, \mathbf{r}_2$. Un caso particolare si ottiene prendendo \mathbf{r}_1 e \mathbf{r}_2 incidenti in un punto $O \in \mathbf{H}$. In questo caso si ha $P_1 = O = P_2$ e il teorema afferma che $\overrightarrow{OP_1''} = k \overrightarrow{OP_1'}$, $\overrightarrow{OP_2''} = k \overrightarrow{OP_2'}$.

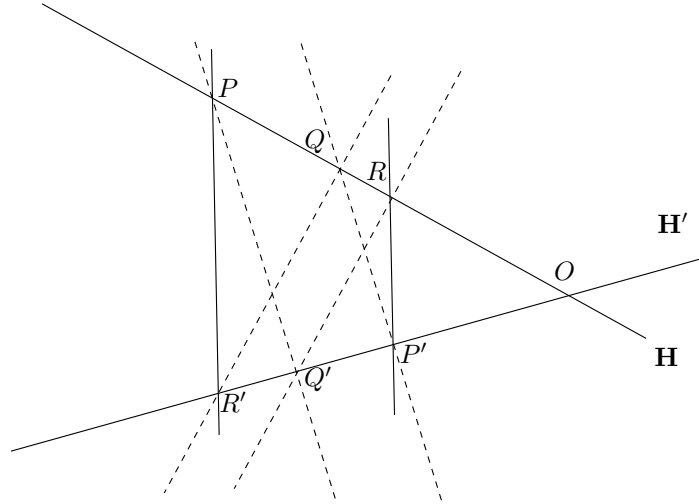


Teorema di Pappo

Teorema 3.11. Siano \mathbf{H} e \mathbf{H}' due rette distinte, del piano affine \mathbf{A} . Siano $P, Q, R \in \mathbf{H}$, $P', Q', R' \in \mathbf{H}'$ punti distinti, nessuno dei quali comune ad \mathbf{H} e ad \mathbf{H}' . Allora

$$\overrightarrow{PQ'} \parallel \overrightarrow{P'Q} \wedge \overrightarrow{QR'} \parallel \overrightarrow{Q'R} \implies \overrightarrow{PR} \parallel \overrightarrow{P'R}.$$

Proof. Supponiamo che \mathbf{H} e \mathbf{H}' non siano parallele, e sia $\{O\} = \mathbf{H} \cap \mathbf{H}'$.



Per il teorema di Talete si ha, per opportuni $h, k \in \mathbb{K}^*$:

$$\begin{aligned} \overrightarrow{OP'} &= k\overrightarrow{OQ'} & \overrightarrow{OQ'} &= h\overrightarrow{OR'} \\ \overrightarrow{OQ} &= k\overrightarrow{OP} & \overrightarrow{OR} &= h\overrightarrow{OQ} \end{aligned} \quad \because \quad \overrightarrow{PQ'} \parallel \overrightarrow{P'Q} \quad \wedge \quad \because \quad \overrightarrow{QR'} \parallel \overrightarrow{Q'R}.$$

Ma allora:

$$\begin{aligned} \overrightarrow{PR'} &= \overrightarrow{OR'} - \overrightarrow{OP} = h^{-1}\overrightarrow{OQ'} - k^{-1}\overrightarrow{OQ} \\ \overrightarrow{RP'} &= \overrightarrow{OP'} - \overrightarrow{OR} = k\overrightarrow{OQ'} - h\overrightarrow{OQ}, \end{aligned}$$

e quindi $\overrightarrow{RP'} = hk\overrightarrow{PR'}$, cioè $\overrightarrow{RP'} \parallel \overrightarrow{PR'}$.

Se $\mathbf{H} \parallel \mathbf{H'}$, allora si ha

$$\overrightarrow{PQ} = \overrightarrow{Q'P'} \quad \because \quad \overrightarrow{PQ'} \parallel \overrightarrow{P'Q}, \quad \overrightarrow{QR} = \overrightarrow{R'Q'} \quad \because \quad \overrightarrow{QR'} \parallel \overrightarrow{Q'R},$$

e quindi $\overrightarrow{PR} = \overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{Q'P'} + \overrightarrow{R'Q'} = \overrightarrow{R'P'}$.

Pertanto $\overrightarrow{PR'} \parallel \overrightarrow{P'R}$. □

Teorema di Desargues

Teorema 3.12. Siano $A, B, C, A', B', C' \in \mathbf{A}$ punti a tre a tre non allineati, tali che $\overline{AB} \parallel \overline{A'B'}$, $\overline{BC} \parallel \overline{B'C'}$, $\overline{AC} \parallel \overline{A'C'}$. Allora le tre rette $\overline{AA'}, \overline{BB'}, \overline{CC'}$ sono parallele oppure hanno un punto in comune.

Proof. Supponiamo che $\overline{AA'}, \overline{BB'}, \overline{CC'}$ non siano parallele. Allora due di esse si incontrano, e possiamo quindi supporre che $\overline{AA'} \cap \overline{BB'} = \{O\}$.

Per il teorema di Talete applicato ad $\overline{AB}, \overline{A'B'}$ si ha

$$\overrightarrow{OA'} = k\overrightarrow{OA}, \quad \overrightarrow{OB'} = k\overrightarrow{OB}, \quad k \in \mathbb{K}.$$

Sia $\{C''\} = \overline{OC} \cap \overline{A'C'}$. Per il teorema di Talete applicato ad $\overline{AC}, \overline{A'C'}$ si ha

$$\overrightarrow{OC''} = k\overrightarrow{OC} \tag{75}$$

perché $\overrightarrow{OA'} = k\overrightarrow{OA}$. D'altra parte, posto $\{C'''\} = \overline{OC} \cap \overline{B'C'}$, il teorema di Talete applicato alle rette $\overline{BC}, \overline{B'C'}$ implica che

$$\overrightarrow{OC'''} = k\overrightarrow{OC} \tag{76}$$

perché $\overrightarrow{OB'} = k\overrightarrow{OB}$. Confrontando le 75 e 76 vediamo che $C'' = C''' = C'$, e quindi O, C, C' sono allineati. □

3.3 Geometria in uno Spazio Affine

Considereremo ora il caso degli spazi affini di dimensione 3. Tra questi rientra come caso particolare lo spazio ordinario.

Supponiamo assegnato un \mathbb{K} -spazio vettoriale \mathbf{V} di dimensione 3 e uno spazio affine \mathbf{A} su \mathbf{V} . Fissiamo un riferimento affine $O\mathbf{e}_1\mathbf{e}_2\mathbf{e}_3$ in \mathbf{A} ; denotiamo con x, y, z le coordinate di un punto variabile $P \in \mathbf{A}$ e con X, Y, Z indeterminate.

I sottospazi affini di \mathbf{A} sono i piani e le rette, oltre ad \mathbf{A} stesso e ai suoi punti.

Un piano \mathbf{p} di \mathbf{A} passante per un punto $Q(q_1, q_2, q_3)$ ha equazioni parametriche della forma

$$\begin{aligned} x &= q_1 + a_1u + b_1v \\ y &= q_2 + a_2u + b_2v \\ z &= q_3 + a_3u + b_3v \end{aligned} \quad (77)$$

dove $\{\mathbf{a}(a_1, a_2, a_3), \mathbf{b}(b_1, b_2, b_3)\}$ è una base della giacitura \mathbf{W} del piano.

Il piano \mathbf{p} , avendo codimensione 1, si rappresenta anche con un'equazione cartesiana

$$AX + BY + CZ + D = 0 \quad (78)$$

per opportuni $A, B, C, D \in \mathbb{K}$ tali che $(A, B, C) \neq (0, 0, 0)$. Le costanti A, B, C, D sono individuate da \mathbf{p} solo a meno di un comune fattore di proporzionalità non nullo. Il piano \mathbf{p} contiene l'origine O se e solo se $D = 0$. I piani di equazione $X = 0, Y = 0, Z = 0$ si dicono **piani coordinati**, e rispettivamente *piano* YZ , *piano* XZ , *piano* XY . Possiamo ottenere un'equazione cartesiana 78 di \mathbf{p} a partire dalle 77 osservando che esse esprimono la dipendenza lineare dei vettori $\overrightarrow{QP}, \mathbf{a}, \mathbf{b}$; le 77 sono pertanto equivalenti all'annullarsi del determinante della matrice le cui righe sono le coordinate di questi vettori. Quindi il punto $P(x, y, z)$ appartiene al piano passante per $Q(q_1, q_2, q_3)$ e parallelo a $\langle \mathbf{a}(a_1, a_2, a_3), \mathbf{b}(b_1, b_2, b_3) \rangle$ se e solo se le sue coordinate soddisfano l'equazione in X, Y, Z

$$\begin{vmatrix} X - q_1 & Y - q_2 & Z - q_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = 0. \quad (79)$$

Espandendo il primo membro della 79 secondo la prima riga, e ponendo

$$A = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}, \quad B = - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}, \quad C = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}, \quad (80)$$

si ottiene

$$A(X - q_1) + B(Y - q_2) + C(Z - q_3) = 0. \quad (81)$$

Si noti che A, B, C non sono tutti nulli: ciò segue dalla loro definizione e dal fatto che \mathbf{a} e \mathbf{b} sono linearmente indipendenti. Se poniamo

$$D = -Aq_1 - Bq_2 - Cq_3, \quad (82)$$

possiamo scrivere la 81 nella forma

$$AX + BY + CZ + D = 0. \quad (83)$$

La 83 è soddisfatta da tutti e soli i punti $P(x, y, z) \in \mathbf{p}$ e quindi è un'equazione cartesiana del piano \mathbf{p} .

Se \mathbf{p} è assegnato mediante tre suoi punti non allineati $P_0(x_0, y_0, z_0), P_1(x_1, y_1, z_1), P_2(x_2, y_2, z_2)$, la sua giacitura è generata dai vettori $\overrightarrow{P_0P_1}$ e $\overrightarrow{P_0P_2}$, e la 79 prende la forma seguente:

$$\begin{vmatrix} X - x_0 & Y - y_0 & Z - z_0 \\ x_1 - x_0 & y_1 - y_0 & z_1 - z_0 \\ x_2 - x_0 & y_2 - y_0 & z_2 - z_0 \end{vmatrix} = 0. \quad (84)$$

La 84 è un'equazione cartesiana del piano passante per i punti P_0, P_1, P_2 .

Una retta $\mathbf{r} \subset \mathbf{A}$ ha equazioni parametriche della forma

$$\begin{aligned} x &= a + lt \\ y &= b + mt \\ z &= c + nt \end{aligned} \quad (85)$$

dove $Q(a, b, c) \in \mathbf{r}$, e $\mathbf{v}(l, m, n)$ è un suo vettore di direzione.

La retta \mathbf{r} può anche essere definita da due equazioni cartesiane, cioè come intersezione di due piani, perché ha codimensione 2:

$$\begin{aligned} AX + BY + CZ + D &= 0 \\ A'X + B'Y + C'Z + D' &= 0 \end{aligned} \quad (86)$$

La direzione di \mathbf{r} è il sottospazio vettoriale di dimensione 1 di \mathbf{V} definito dal sistema omogeneo associato

$$\begin{aligned} AX + BY + CZ &= 0 \\ A'X + B'Y + C'Z &= 0 \end{aligned} \quad (87)$$

Da ciò segue che il vettore $\mathbf{v}(l, m, n)$ di coordinate

$$l = \begin{vmatrix} B & C \\ B' & C' \end{vmatrix}, \quad m = - \begin{vmatrix} A & C \\ A' & C' \end{vmatrix}, \quad n = \begin{vmatrix} A & B \\ A' & B' \end{vmatrix}$$

è un vettore di direzione di \mathbf{r} , perché (l, m, n) è una soluzione non nulla del sistema 87.

Quest'osservazione fornisce un metodo pratico per calcolare un vettore di direzione di una retta \mathbf{r} assegnata mediante equazioni cartesiane 77.

Se viceversa la retta \mathbf{r} è assegnata mediante un suo punto $Q(a, b, c)$ e un suo vettore di direzione $\mathbf{v}(l, m, n)$, ovvero per mezzo della 85, se ne possono ottenere equazioni cartesiane imponendo la condizione che si annullino i minori di ordine 2 della matrice

$$\begin{pmatrix} X - a & Y - b & Z - c \\ l & m & n \end{pmatrix} \quad (88)$$

Infatti questa condizione nelle indeterminate X, Y, Z è soddisfatta dalle coordinate dei punti $P(x, y, z) \in \mathbf{A}$ tali che il vettore \overrightarrow{QP} sia proporzionale a \mathbf{v} , cioè dai punti di \mathbf{r} . Supponendo ad esempio $l \neq 0$, la condizione detta è equivalente alle due equazioni seguenti:

$$m(X - a) - l(Y - b) = 0 \quad (89)$$

$$n(X - a) - l(Z - c) = 0 \quad (90)$$

cioè alle

$$mX - lY - (ma - lb) = 0, \quad nX - lZ - (na - lc) = 0 \quad (10.14)$$

che sono equazioni cartesiane di due piani **distinti** contenenti \mathbf{r} (che i due piani siano distinti segue dal fatto che, essendo $l \neq 0$, la matrice dei coefficienti delle incognite delle due equazioni 10.14 ha rango due), i quali quindi definiscono \mathbf{r} .

Nota 16. Si noti che le 8990 implicano l'annullamento del rimanente minore della 88:

$$n(Y - b) - m(Z - c) = l^{-1} [m[n(X - a) - l(Z - c)] - n[m(X - a) - l(Y - b)]] = 0.$$

Proposizione 3.13. *Siano \mathbf{p}, \mathbf{p}' due piani di \mathbf{A} , di equazioni cartesiane*

$$\begin{aligned} AX + BY + CZ + D &= 0 \\ A'X + B'Y + C'Z + D' &= 0 \end{aligned} \quad (91)$$

rispettivamente. Allora:

1. $\mathbf{p} \parallel \mathbf{p}'$ se e solo se la matrice

$$\begin{pmatrix} A & B & C \\ A' & B' & C' \end{pmatrix} \quad (92)$$

ha rango 1.

2. *Se la matrice 92 ha rango 1, allora $\mathbf{p} \equiv \mathbf{p}'$ sono paralleli e disgiunti oppure coincidono a seconda che la matrice*

$$\begin{pmatrix} A & B & C & D \\ A' & B' & C' & D' \end{pmatrix} \quad (93)$$

abbia rango 2 oppure 1.

3. *Se \mathbf{p} e \mathbf{p}' non sono paralleli, allora sono incidenti e $\mathbf{p} \cap \mathbf{p}'$ è una retta; ciò avviene se e solo se la matrice 92 ha rango 2.*

Proof. I casi considerati nell'enunciato corrispondono alle possibilità che si hanno per il sistema 91, di cui la 92 e la 93 sono rispettivamente la matrice dei coefficienti e la matrice orlata. Poiché le giaciture di \mathbf{p} e di \mathbf{p}' sono determinate dalle equazioni omogenee associate

$$\begin{aligned} AX + BY + CZ &= 0 \\ A'X + B'Y + C'Z &= 0 \end{aligned} \quad (94)$$

il parallelismo di \mathbf{p} e \mathbf{p}' è equivalente alla condizione che le 94 siano proporzionali, cioè che la 92 abbia rango 1. In questo caso la matrice 93 ha rango 1 o 2 rispettivamente, a seconda che il sistema 91 sia compatibile oppure no, cioè che $\mathbf{p} \cap \mathbf{p}' \neq \emptyset$ oppure $\mathbf{p} \cap \mathbf{p}' = \emptyset$. Nel primo caso dev'essere $\mathbf{p} \equiv \mathbf{p}'$ perché i due piani sono paralleli. Ciò dimostra (1) e (2). Per la (1) la 92 ha rango 2 se e solo se \mathbf{p} e \mathbf{p}' non sono paralleli. In tal caso il sistema 91 è compatibile, perché anche la 93 ha rango 2. Il sistema 91 possiede ∞^1 soluzioni e $\mathbf{p} \cap \mathbf{p}'$ è una retta. Anche la (3) è dimostrata. \square

Passiamo ora a considerare il caso di una retta e di un piano. La proposizione seguente descrive tutti i casi che si possono presentare nella loro posizione relativa.

Proposizione 3.14. *Sia \mathbf{r} una retta di equazioni parametriche 85 e cartesiane 86, e sia \mathbf{p}'' un piano di equazione*

$$A''X + B''Y + C''Z + D'' = 0. \quad (95)$$

1. $\mathbf{r} \parallel \mathbf{p}''$ se e solo se

$$\begin{vmatrix} A & B & C \\ A' & B' & C' \\ A'' & B'' & C'' \end{vmatrix} = 0 \quad (96)$$

o, equivalentemente, se e solo se

$$A''l + B''m + C''n = 0 \quad (97)$$

2. Se la 96 è soddisfatta, allora \mathbf{r} e \mathbf{p}'' sono paralleli e disgiunti oppure coincidono a seconda che la matrice

$$\begin{pmatrix} A & B & C & D \\ A' & B' & C' & D' \\ A'' & B'' & C'' & D'' \end{pmatrix} \quad (98)$$

abbia rango 3 oppure 2.

3. Se \mathbf{r} e \mathbf{p}'' non sono paralleli, allora sono incidenti ed $\mathbf{r} \cap \mathbf{p}''$ consiste di un solo punto; ciò avviene se e solo se

$$\begin{vmatrix} A & B & C \\ A' & B' & C' \\ A'' & B'' & C'' \end{vmatrix} \neq 0 \quad (99)$$

o, equivalentemente, se e solo se

$$A''l + B''m + C''n \neq 0. \quad (100)$$

Proof. 1. Supponiamo soddisfatta la 96. Allora il sistema omogeneo

$$\begin{aligned} AX + BY + CZ &= 0 \\ A'X + B'Y + C'Z &= 0 \\ A''X + B''Y + C''Z &= 0 \end{aligned} \quad (101)$$

è equivalente al sistema costituito dalle sue due prime equazioni, che definiscono la direzione di \mathbf{r} . Pertanto ogni vettore di direzione di \mathbf{r} soddisfa la terza equazione 101, che è un'equazione della giacitura di \mathbf{p}'' ; cioè \mathbf{p}'' ed \mathbf{r} sono paralleli. Se viceversa \mathbf{p}'' ed \mathbf{r} sono paralleli, necessariamente la terza equazione 101 è dipendente dalle prime due, e la 96 è soddisfatta. L'equivalenza delle condizioni 96 e 97 si vede sviluppando il determinante 96 secondo la terza riga.

2. Se \mathbf{r} ed \mathbf{p}'' sono paralleli, allora coincidono oppure sono disgiunti a seconda che il sistema

$$\begin{aligned} AX + BY + CZ + D &= 0 \\ A'X + B'Y + C'Z + D' &= 0 \\ A''X + B''Y + C''Z + D'' &= 0 \end{aligned} \quad (102)$$

sia compatibile oppure no, applichiamo dunque il teorema di Kronecker-Rouchè-Capelli.

\square

Per la (1), \mathbf{r}'' ed ν non sono paralleli se e solo se la [10.23] o la [10.24], che come già osservato sono equivalenti, sono soddisfatte. In questo caso il sistema [10.26] è compatibile e ammette un'unica soluzione, per il teorema 5.7. Quindi \mathbf{r}'' e ρ'' hanno un unico punto in comune.

La [10.20] e la [10.21] danno la **condizione di parallelismo di una retta e di un piano**.

Passiamo ora a considerare due rette e le loro posizioni relative.

Due rette \mathbf{r} e \mathbf{r}_1 di Δ si dicono **complanari** se esiste un piano che le contiene entrambe.

10.3 Proposizione

Due rette \mathfrak{r} ed \mathfrak{r}_1 di Δ sono complanari se e solo se una delle condizioni seguenti è verificata:

1. \mathfrak{r} ed \mathfrak{r}_1 sono parallele;
2. \mathfrak{r} ed \mathfrak{r}_1 sono incidenti.

In particolare \mathfrak{r} ed \mathfrak{r}_1 sono complanari se e solo se non sono sghembe.

Dimostrazione. Poiché due rette di un piano affine che non sono incidenti sono necessariamente parallele, se \mathfrak{r} ed \mathfrak{r}_1 sono complanari, allora o sono parallele o sono incidenti.

Viceversa, se $\mathfrak{r} = \mathfrak{r}_1$, allora ovviamente \mathfrak{r} ed \mathfrak{r}_1 sono complanari. Se \mathfrak{r} ed \mathfrak{r}_1 sono parallele e distinte sono contenute nel piano che passa per un punto qualsiasi $Q \in \mathfrak{r}$ e di giacitura $\langle \vec{v}, \overrightarrow{OQ_1} \rangle$, dove \vec{v} è un vettore di direzione di \mathfrak{r} e di \mathfrak{r}_1 , e $Q_1 \in \mathfrak{r}_1 \setminus \mathfrak{r}$. Se \mathfrak{r} ed \mathfrak{r}_1 sono incidenti e distinte sono contenute nel piano passante per il punto $\mathfrak{r} \cap \mathfrak{r}_1$ e di giacitura $\langle \vec{v}, \vec{v}_1 \rangle$, dove \vec{v} e \vec{v}_1 sono vettori di direzione di \mathfrak{r} e di \mathfrak{r}_1 , rispettivamente.

La proposizione che segue dà dei criteri per riconoscere la complanarità di due rette assegnate.

10.4 Proposizione

Siano \mathfrak{r} ed \mathfrak{r}_1 due rette di Δ ; supponiamo che \mathfrak{r} abbia equazioni cartesiane [10.10] e che \mathfrak{r}_1 abbia equazioni cartesiane

$$\begin{aligned} A_1 X + B_1 Y + C_1 Z + D_1 &= 0 \\ A'_1 X + B'_1 Y + C'_1 Z + D'_1 &= 0. \end{aligned} \quad (10.27)$$

Siano $Q(a, b, c) \in \mathfrak{r}$, $Q_1(a_1, b_1, c_1) \in \mathfrak{r}_1$; siano inoltre $\vec{v}(m, n, v) \in \nu(\mathfrak{r})$, $\vec{v}_1(m_1, n_1, v_1) \in \nu(\mathfrak{r}_1)$ vettori di direzione di \mathfrak{r} e di \mathfrak{r}_1 rispettivamente. Le seguenti condizioni sono equivalenti:

1. \mathfrak{r} ed \mathfrak{r}_1 sono complanari;

$$2. \begin{vmatrix} a - a_1 & b - b_1 & c - c_1 \\ l & m & n \\ l_1 & m_1 & n_1 \end{vmatrix} = 0;$$

$$3. \begin{vmatrix} A & B & C & D \\ A' & B' & C' & D' \\ A_1 & B_1 & C_1 & D_1 \\ A'_1 & B'_1 & C'_1 & D'_1 \end{vmatrix} = 0.$$

Dimostrazione

(1) \Rightarrow (2) Se ν ed ν_1 sono parallele, allora \vec{v} e \vec{v}_1 sono proporzionali, e quindi la (2) è soddisfatta. Se ν ed ν_1 sono incidenti, allora sono contenute in uno stesso piano \mathcal{G} : il vettore $\overrightarrow{QQ_1}$ appartiene alla giacitura di \mathcal{G} , che è generata da \vec{v} e da \vec{v}_1 , e di nuovo la (2) è soddisfatta.

(2) \Rightarrow (1) Se la (2) è soddisfatta, allora o le ultime due righe della matrice

$$\begin{pmatrix} a - a_1 & b - b_1 & c - c_1 \\ l & m & n \\ l_1 & m_1 & n_1 \end{pmatrix} \quad (10.28)$$

sono proporzionali, e le due rette sono parallele, oppure la prima riga della (10.28) è combinazione lineare delle ultime due, e il piano \mathcal{G} di giacitura $\langle \vec{v}, \vec{v}_1 \rangle$ passante per Q , che contiene ν , contiene anche Q_1 , e quindi contiene ν_1 . In ogni caso le due rette sono complanari.

(3) \Rightarrow (1) Se il sistema costituito dalle [10.10] e dalle [10.27] è compatibile, allora ν ed ν_1 sono complanari perché hanno punti in comune.

Supponiamo che il sistema non sia compatibile. Allora, poiché il rango R della matrice

$$\begin{pmatrix} A & B & C & D \\ A' & B' & C' & D' \\ A_1 & B_1 & C_1 & D_1 \\ A'_1 & B'_1 & C'_1 & D'_1 \end{pmatrix} \quad (10.29)$$

è al più 3, e quello r della matrice

$$\begin{pmatrix} A & B & C \\ A' & B' & C' \\ A_1 & B_1 & C_1 \\ A'_1 & B'_1 & C'_1 \end{pmatrix} \quad (10.30)$$

è almeno 2, si deve avere $R = 3$ ed $r = 2$. La condizione $r = 2$ implica che le due ultime righe della [10.30] sono combinazione lineare delle prime due; e quindi ν , ν_1 hanno pertanto la stessa direzione, cioè sono parallele, e in particolare complanari.

(3) \Rightarrow (3') Se ν ed ν_1 sono incidenti il sistema costituito dalle [10.10] e dalle [10.27] è compatibile, e quindi la [10.29] ha rango minore di 4. Se invece ν ed ν_1 sono parallele, la [10.30] ha rango 2 e conseguentemente la [10.29] ha rango al più 3. In entrambi i casi la (3) è verificata.

Sia ν una retta di \mathbb{A} . L'insieme Φ dei piani di \mathbb{A} che contengono ν si dice *fascio proprio di piani*, ed ν è detta *asse del fascio* (fig. 10.2).

Consideriamo due piani distinti $\rho, \rho_1 \in \Phi$, di equazioni rispettive

$$AX + BY + CZ + D = 0$$

$$A_1X + B_1Y + C_1Z + D_1 = 0.$$

Come nel caso dei fasci di rette, si verifica che ogni piano appartenente a Φ ha equazione

$$\lambda(AX + BY + CZ + D) + \mu(A_1X + B_1Y + C_1Z + D_1) = 0$$

per opportuni $\lambda, \mu \in \mathbb{K}$ non entrambi nulli.

Anche in questo caso si può utilizzare un parametro non omogeneo t per rappresentare i piani del fascio nella forma

$$AX + BY + CZ + D + t(A_1X + B_1Y + C_1Z + D_1) = 0,$$

tenendo presente che il piano ρ_1 è l'unico elemento di Φ che non si può ottenere in questo modo.

Se \mathcal{G} è un piano non parallelo a ν , i piani appartenenti al fascio Φ di asse ν intersecano \mathcal{G} nelle rette appartenenti al fascio di rette di centro il punto $\nu \cap \mathcal{G}$ (fig. 10.3).

Sia W un sottospazio di dimensione 2 di V . L'insieme Ψ dei piani di \mathbb{A} che hanno giacitura W si dice *fascio improprio di piani* e W si dice *giacitura del fascio*.

Se \mathcal{G} è un piano del fascio Ψ , di equazione

Se φ è un piano del fascio Ψ , di equazione

$$AX + BY + CZ + D = 0,$$

allora ogni altro elemento di Ψ ha equazione

$$AX + BY + CZ + t = 0$$

al variare di $t \in K$.

10.4 Esempio

Siano τ ed τ' due rette sghembe di \mathbf{A} , e sia $P \in \mathbf{A}$ un punto non appartenente a $\tau \cup \tau'$ (fig. 10.4).

Esiste una e una sola retta contenente P e complanare sia con τ che con τ' . Per vederlo si osservi che esiste un unico piano φ contenente sia τ che P , e un unico piano φ' contenente sia τ' che P . Poiché τ e τ' hanno in comune il punto P e sono distinti (altrimenti τ e τ' sarebbero complanari), $\varphi \cap \varphi'$ è una retta che contiene P e, per costruzione, è complanare sia con τ che con τ' . L'unicità di φ segue da quella di φ e di φ' .

Si noti che certamente τ non è parallela sia a τ che a τ' , perché in tal caso τ ed τ' sarebbero parallele tra loro. Quindi τ interseca almeno una delle due rette.

4 Applicazioni Lineari

4.1 Applicazioni lineari

Applicazione lineare o omomorfismo

Definizione 4.1. Siano \mathbf{V} e \mathbf{W} due \mathbb{K} -spazi vettoriali. Un'applicazione

$$F : \mathbf{V} \rightarrow \mathbf{W}.$$

si dice **lineare** (o **omomorfismo**) se per ogni $\mathbf{v}, \mathbf{v}' \in \mathbf{V}$ e $c \in \mathbb{K}$ si ha

$$F(\mathbf{v} + \mathbf{v}') = F(\mathbf{v}) + F(\mathbf{v}') \quad (103)$$

$$F(c\mathbf{v}) = cF(\mathbf{v}) \quad (104)$$

Proposizione 4.1. Le due condizioni (103) e (104) sono equivalenti all'unica condizione

$$\forall \mathbf{v}, \mathbf{v}' \in \mathbf{V}, \forall c, c' \in \mathbb{K}, \quad F(c\mathbf{v} + c'\mathbf{v}') = cF(\mathbf{v}) + c'F(\mathbf{v}'). \quad (105)$$

Proof. Infatti la (105) si traduce nella (103) prendendo $c = c' = 1$ e nella (104) prendendo $c' = 0$; viceversa, se F soddisfa le (103) e (104), allora

$$F(c\mathbf{v} + c'\mathbf{v}') = F(c\mathbf{v}) + F(c'\mathbf{v}') = cF(\mathbf{v}) + c'F(\mathbf{v}')$$

dove la prima uguaglianza segue dalla (103) e la seconda dalla (104). □

Applicando ripetutamente la (105) si deduce che, se F è lineare, avremo

$$\forall \mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbf{V}, \forall c_1, \dots, c_n \in \mathbb{K}, \quad F\left(\sum_{i=1}^n c_i \mathbf{v}_i\right) = \sum_{i=1}^n c_i F(\mathbf{v}_i).$$

Nota 17. Si osservi che la (104), applicata a $c = 0$ e $\mathbf{v} \in \mathbf{V}$ qualsiasi, implica che se F è lineare,

$$F(\mathbf{0}) = \mathbf{0}.$$

Operatore lineare o endomorfismo

Definizione 4.2. Un'applicazione lineare $F : \mathbf{V} \rightarrow \mathbf{V}$ è detta **operatore lineare** su \mathbf{V} , oppure **endomorfismo** di \mathbf{V} .

Funzionale lineare

Definizione 4.3. Un'applicazione lineare $F : \mathbf{V} \rightarrow \mathbb{K}$ è detta **funzionale lineare** su \mathbf{V} .

Composizione di applicazioni lineari

Proposizione 4.2. Siano $G : \mathbf{U} \rightarrow \mathbf{V}$ ed $F : \mathbf{V} \rightarrow \mathbf{W}$ applicazioni lineari. Allora la composizione $F \circ G : \mathbf{U} \rightarrow \mathbf{W}$ è un'applicazione lineare.

Proof. Siano $\mathbf{u}, \mathbf{u}' \in \mathbf{U}$ e $c, c' \in \mathbb{K}$. Poniamo $G(\mathbf{u}) = \mathbf{v}$ e $G(\mathbf{u}') = \mathbf{v}'$, da cui segue che $F(\mathbf{v}) = \mathbf{w}$ e $F(\mathbf{v}') = \mathbf{w}'$. Consideriamo ora:

$$\begin{aligned}(F \circ G)(c\mathbf{u} + c'\mathbf{u}') &= F(G(c\mathbf{u} + c'\mathbf{u}')) = F(cG(\mathbf{u}) + c'G(\mathbf{u}')) \\ &= F(c\mathbf{v} + c'\mathbf{v}') = cF(\mathbf{v}) + c'F(\mathbf{v}') \\ &= c\mathbf{w} + c'\mathbf{w}' = c(F \circ G)(\mathbf{u}) + c'(F \circ G)(\mathbf{u}').\end{aligned}$$

Dunque, $F \circ G$ soddisfa le proprietà di linearità, e quindi è anch'essa un'applicazione lineare. \square

Isomorfismo

Definizione 4.4. Diremo che $F : \mathbf{V} \rightarrow \mathbf{W}$ è un **isomorfismo** se è un'applicazione lineare biunivoca.

Proposizione 4.3. L'applicazione inversa $F^{-1} : \mathbf{W} \rightarrow \mathbf{V}$ è anch'essa lineare e quindi è un isomorfismo.

Proof. Siano infatti $\mathbf{w}, \mathbf{w}' \in \mathbf{W}$, e $\mathbf{v} = F^{-1}(\mathbf{w})$, $\mathbf{v}' = F^{-1}(\mathbf{w}')$. Si ha

$$F^{-1}(\mathbf{w} + \mathbf{w}') = F^{-1}(F(\mathbf{v}) + F(\mathbf{v}')) = F^{-1}(F(\mathbf{v} + \mathbf{v}')) = \mathbf{v} + \mathbf{v}' = F^{-1}(\mathbf{w}) + F^{-1}(\mathbf{w}').$$

Analogamente, se $c \in \mathbb{K}$,

$$F^{-1}(c\mathbf{w}) = F^{-1}(cF(\mathbf{v})) = F^{-1}(F(c\mathbf{v})) = c\mathbf{v} = cF^{-1}(\mathbf{w}).$$

\square

Automorfismo

Definizione 4.5. Un isomorfismo di uno spazio \mathbf{V} in sè stesso è detto **automorfismo** di \mathbf{V} .

Notazioni

- $\text{hom}(\mathbf{V}, \mathbf{W})$: Insieme di tutte le applicazioni lineari di \mathbf{V} in \mathbf{W} ;
- $\text{end}(\mathbf{V})$: Insieme di tutti gli operatori lineari su \mathbf{V} ;
- \mathbf{V}^\sim : Insieme di tutti i funzionali lineari su \mathbf{V} ;
- $\text{GL}(\mathbf{V})$: Sottoinsieme di $\text{end}(\mathbf{V})$ costituito dagli automorfismi.

Proposizione 4.4. Per ogni coppia di \mathbb{K} -spazi vettoriali \mathbf{V} e \mathbf{W} , l'applicazione nulla

$$0 : \mathbf{V} \rightarrow \mathbf{W} \quad \mathbf{v} \mapsto \mathbf{0}$$

è un'applicazione lineare.

Proof. Siano $\mathbf{v}, \mathbf{v}' \in \mathbf{V}$ e $c, c' \in \mathbb{K}$. Allora:

$$0(c\mathbf{v} + c'\mathbf{v}') = \mathbf{0} = \mathbf{0} + \mathbf{0} = c\mathbf{0} + c'\mathbf{0} = c0(\mathbf{v}) + c'0(\mathbf{v}'),$$

Quindi 0 è lineare. \square

Proposizione 4.5. Per ogni \mathbb{K} -spazio vettoriale \mathbf{V} , l'applicazione

$$\mathbf{1}_{\mathbf{V}} : \mathbf{V} \rightarrow \mathbf{V} \quad \mathbf{v} \mapsto \mathbf{v}$$

è un'automorfismo.

Proof. Siano $\mathbf{v}, \mathbf{v}' \in \mathbf{V}$ e $c, c' \in \mathbb{K}$. Allora si ha

$$\mathbf{1}_{\mathbf{V}}(c\mathbf{v} + c'\mathbf{v}') = c\mathbf{v} + c'\mathbf{v}' = c\mathbf{1}_{\mathbf{V}}(\mathbf{v}) + c'\mathbf{1}_{\mathbf{V}}(\mathbf{v}').$$

Inoltre, $\mathbf{1}_{\mathbf{V}}$ è invertibile e la sua inversa è ancora $\mathbf{1}_{\mathbf{V}}$ stessa, che è lineare per quanto appena dimostrato. Quindi $\mathbf{1}_{\mathbf{V}}$ è un automorfismo. \square

Proposizione 4.6. Siano $F, G \in \text{GL}(\mathbf{V})$. Allora F^{-1} e $F \circ G$ sono automorfismi di \mathbf{V} .

Proof. Poiché $F \in \text{GL}(\mathbf{V})$, per definizione F è un isomorfismo, cioè un'applicazione lineare invertibile. La sua inversa F^{-1} è quindi anch'essa lineare, e per costruzione è biiettiva. Dunque $F^{-1} \in \text{GL}(\mathbf{V})$, cioè è un automorfismo. Analogamente, poiché $F, G \in \text{GL}(\mathbf{V})$, allora sono entrambi isomorfismi. La composizione di due isomorfismi è ancora un isomorfismo: infatti la composizione $F \circ G$ è lineare (poiché composizione di applicazioni lineari) e invertibile, con inversa data da $G^{-1} \circ F^{-1}$ (che è ancora lineare). Pertanto $F \circ G \in \text{GL}(\mathbf{V})$, cioè è un automorfismo. \square

Proposizione 4.7.

1. Se $c \in \mathbb{K}$, l'applicazione

$$c\mathbf{1}_{\mathbf{V}} : \mathbf{V} \rightarrow \mathbf{V} \quad \mathbf{v} \mapsto c\mathbf{v}$$

è lineare.

2. Se $c = 0$, allora

$$0\mathbf{1}_{\mathbf{V}} = \mathbf{0}.$$

3. Se $c \neq 0$, allora $c\mathbf{1}_{\mathbf{V}}$ è un automorfismo il cui inverso è $c^{-1}\mathbf{1}_{\mathbf{V}}$.

Proof.

1. Siano $\mathbf{v}, \mathbf{v}' \in \mathbf{V}$ e $\lambda \in \mathbb{K}$. Allora:

$$(c\mathbf{1}_{\mathbf{V}})(\mathbf{v} + \mathbf{v}') = c(\mathbf{v} + \mathbf{v}') = c\mathbf{v} + c\mathbf{v}' = (c\mathbf{1}_{\mathbf{V}})(\mathbf{v}) + (c\mathbf{1}_{\mathbf{V}})(\mathbf{v}'),$$

$$(c\mathbf{1}_{\mathbf{V}})(\lambda\mathbf{v}) = c(\lambda\mathbf{v}) = \lambda(c\mathbf{v}) = \lambda(c\mathbf{1}_{\mathbf{V}})(\mathbf{v}).$$

Dunque $c\mathbf{1}_{\mathbf{V}}$ è un'applicazione lineare.

2. Se $c = 0$, allora per ogni $\mathbf{v} \in \mathbf{V}$ si ha $0\mathbf{1}_{\mathbf{V}}(\mathbf{v}) = 0 \cdot \mathbf{v} = \mathbf{0}$. Quindi l'applicazione è costantemente nulla, ovvero $0\mathbf{1}_{\mathbf{V}} = \mathbf{0}$.

3. Se $c \neq 0$, definiamo $f = c\mathbf{1}_{\mathbf{V}}$ e consideriamo $g = c^{-1}\mathbf{1}_{\mathbf{V}}$. Allora per ogni $\mathbf{v} \in \mathbf{V}$:

$$(g \circ f)(\mathbf{v}) = g(c\mathbf{v}) = c^{-1}(c\mathbf{v}) = \mathbf{v}, \quad (f \circ g)(\mathbf{v}) = f(c^{-1}\mathbf{v}) = c(c^{-1}\mathbf{v}) = \mathbf{v}.$$

Quindi g è l'inversa di f e, poiché è lineare, f è un automorfismo. \square

Proposizione 4.8. Sia \mathbf{V} un \mathbb{K} -spazio vettoriale tale che $\dim(\mathbf{V}) = 1$. Allora gli unici operatori lineari su \mathbf{V} sono quelli della forma $c\mathbf{1}_{\mathbf{V}}$.

Proof. Infatti $F \in \text{end}(\mathbf{V})$ ed $\mathbf{e} \in \mathbf{V}$, $\mathbf{e} \neq \mathbf{0}$; poichè $\{\mathbf{e}\}$ è una base di \mathbf{V} , si ha $F(\mathbf{e}) = c\mathbf{e}$ per qualche $c \in \mathbb{K}$, e quindi per ogni $\mathbf{v} = x\mathbf{e} \in \mathbf{V}$:

$$F(\mathbf{v}) = F(x\mathbf{e}) = xF(\mathbf{e}) = x(c\mathbf{e}) = c(x\mathbf{e}) = c\mathbf{v}.$$

\square

Isomorfismo definito da una base

Proposizione 4.9. Sia $\mathbf{e} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una base del \mathbb{K} -spazio vettoriale \mathbf{V} , e

$$\varphi_{\mathbf{e}} : \mathbf{V} \rightarrow \mathbb{K}^n, \quad \mathbf{v} = \sum_{i=1}^n x_i \mathbf{e}_i \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

l'applicazione che associa ad ogni vettore $\mathbf{v} \in \mathbf{V}$ la n -upla (x_1, \dots, x_n) delle sue coordinate rispetto alla base \mathbf{e} . Allora $\varphi_{\mathbf{e}}$ è un isomorfismo.

Proof. Infatti, $\forall c \in \mathbb{K}$ e $\forall \mathbf{v} = \sum_{i=1}^n x_i \mathbf{e}_i$, $\mathbf{v}' = \sum_{i=1}^n y_i \mathbf{e}_i \in \mathbf{V}$ si ha

$$\varphi_{\mathbf{e}}(\mathbf{v} + \mathbf{v}') = \varphi_{\mathbf{e}}\left(\sum_{i=1}^n (x_i + y_i) \mathbf{e}_i\right) = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \varphi_{\mathbf{e}}(\mathbf{v}) + \varphi_{\mathbf{e}}(\mathbf{v}'),$$

e

$$\varphi_{\mathbf{e}}(c\mathbf{v}) = \begin{pmatrix} cx_1 \\ \vdots \\ cx_n \end{pmatrix} = c \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = c\varphi_{\mathbf{e}}(\mathbf{v}).$$

e quindi $\varphi_{\mathbf{e}}$ è lineare. Inoltre, per le proprietà delle coordinate di un vettore rispetto a una base, $\varphi_{\mathbf{e}}$ è biunivoca. \square

Definizione 4.6. L'isomorfismo $\varphi_{\mathbf{e}}$ è detto **isomorfismo definito dalla base \mathbf{e}** .

Caso 5. Se ad esempio $\mathbf{V} = \mathbb{K}^n$, ed \mathbf{e} è la sua base canonica, allora $\varphi_{\mathbf{e}}$ è l'identità di \mathbb{K}^n in sè stesso.

Proiezione definita dalla decomposizione in sottospazi supplementari

Definizione 4.7. Sia \mathbf{V} un \mathbb{K} -spazio vettoriale e siano \mathbf{U} e \mathbf{W} due sottospazi vettoriali tali che

$$\mathbf{V} = \mathbf{U} \oplus \mathbf{W}. \quad (106)$$

Allora, per ogni $\mathbf{v} \in \mathbf{V}$, esiste una unica coppia $(\mathbf{u}, \mathbf{w}) \in \mathbf{U} \times \mathbf{W}$ tale che $\mathbf{v} = \mathbf{u} + \mathbf{w}$. Possiamo dunque definire le seguenti applicazioni:

$$p_{\mathbf{U}} : \mathbf{V} \rightarrow \mathbf{U}, \quad \mathbf{v} = \mathbf{u} + \mathbf{w} \mapsto \mathbf{u}, \quad p_{\mathbf{W}} : \mathbf{V} \rightarrow \mathbf{W}, \quad \mathbf{v} = \mathbf{u} + \mathbf{w} \mapsto \mathbf{w}.$$

Chiamiamo $p_{\mathbf{U}}$ e $p_{\mathbf{W}}$ le **proiezioni di \mathbf{V} su \mathbf{U} e \mathbf{W} definite dalla decomposizione (106)**.

Proposizione 4.10. La proiezione $p_{\mathbf{W}} : \mathbf{V} \rightarrow \mathbf{W}$ definita dalla decomposizione $\mathbf{V} = \mathbf{U} \oplus \mathbf{W}$ è un'applicazione lineare.

Proof. Siano $\mathbf{v}, \mathbf{v}' \in \mathbf{V}$ con $\mathbf{v} = \mathbf{u} + \mathbf{w}$ e $\mathbf{v}' = \mathbf{u}' + \mathbf{w}'$ per $\mathbf{u}, \mathbf{u}' \in \mathbf{U}$ e $\mathbf{w}, \mathbf{w}' \in \mathbf{W}$, e siano $c, c' \in \mathbb{K}$. Allora:

$$c\mathbf{v} + c'\mathbf{v}' = c(\mathbf{u} + \mathbf{w}) + c'(\mathbf{u}' + \mathbf{w}') = (c\mathbf{u} + c'\mathbf{u}') + (c\mathbf{w} + c'\mathbf{w}').$$

Applicando $p_{\mathbf{W}}$:

$$p_{\mathbf{W}}(c\mathbf{v} + c'\mathbf{v}') = c\mathbf{w} + c'\mathbf{w}' = c p_{\mathbf{W}}(\mathbf{v}) + c' p_{\mathbf{W}}(\mathbf{v}'),$$

cioè $p_{\mathbf{W}}$ è lineare. Un ragionamento analogo mostra che anche $p_{\mathbf{U}}$ è un'applicazione lineare. \square

Caso 6. Se in particolare \mathbf{U} è un iperpiano, allora $\mathbf{W} = \langle \mathbf{w} \rangle$ per qualche $\mathbf{w} \in \mathbf{V} \setminus \mathbf{U}$, e dunque

$$\mathbf{V} = \mathbf{U} \oplus \langle \mathbf{w} \rangle$$

e in questo caso $p_{\mathbf{U}}$ è detta **proiezione di \mathbf{V} su \mathbf{U} nella direzione $\langle \mathbf{w} \rangle$** .

Se $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ è una base di \mathbf{V} e $1 \leq k < n$, la proiezione di \mathbf{V} su $\langle \mathbf{e}_{k+1}, \dots, \mathbf{e}_n \rangle$ definita dalla decomposizione

$$\mathbf{V} = \langle \mathbf{e}_1, \dots, \mathbf{e}_k \rangle \oplus \langle \mathbf{e}_{k+1}, \dots, \mathbf{e}_n \rangle$$

è:

$$p \left(\sum_{i=1}^n c_i \mathbf{e}_i \right) = \sum_{i=k+1}^n c_i \mathbf{e}_i.$$

Proiezione rispetto a una decomposizione indotta da una base Sia $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una base di \mathbf{V} e sia $1 \leq k < n$. Allora i sottospazi

$$\mathbf{U} := \langle \mathbf{e}_1, \dots, \mathbf{e}_k \rangle \quad \text{e} \quad \mathbf{W} := \langle \mathbf{e}_{k+1}, \dots, \mathbf{e}_n \rangle$$

sono supplementari in \mathbf{V} , cioè $\mathbf{V} = \mathbf{U} \oplus \mathbf{W}$. Possiamo quindi definire la proiezione su \mathbf{W} secondo \mathbf{U} come:

$$p \left(\sum_{i=1}^n c_i \mathbf{e}_i \right) = \sum_{i=k+1}^n c_i \mathbf{e}_i,$$

che restituisce la componente di un vettore lungo \mathbf{W} , annullando quella lungo \mathbf{U} .

Esempio 4.1. Siano π il piano ordinario, r una sua retta, \mathbf{V} e \mathbf{W} gli \mathbb{R} -spazi vettoriali dei vettori geometrici di π e di r rispettivamente, e $\mathbf{u} \in \mathbf{V} \setminus \mathbf{W}$ un vettore non parallelo a r . Scrivendo $\mathbf{U} = \langle \mathbf{u} \rangle$, arriviamo alla forma $\mathbf{V} = \mathbf{U} \oplus \mathbf{W}$. Sia $\mathbf{v} \in \mathbf{V}$ tale che \mathbf{v} non sia parallelo nè a \mathbf{U} nè a \mathbf{W} . In questo caso, entrambi i sottospazi supplementari sono iperpiani di \mathbf{V} , dunque dato un $\mathbf{v} \in \mathbf{V}$, possiamo analizzare due casi distinti, quello di $p_{\mathbf{U}}(\mathbf{v})$ e quello di $p_{\mathbf{W}}(\mathbf{v})$ che sono rispettivamente la proiezione di \mathbf{V} su \mathbf{U} nella direzione \mathbf{W} e la proiezione di \mathbf{V} su \mathbf{W} nella direzione \mathbf{U} .

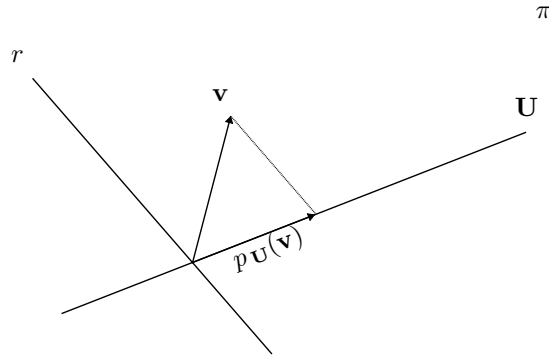


Figure 1: Proiezione di \mathbf{V} su \mathbf{U} nella direzione \mathbf{W}

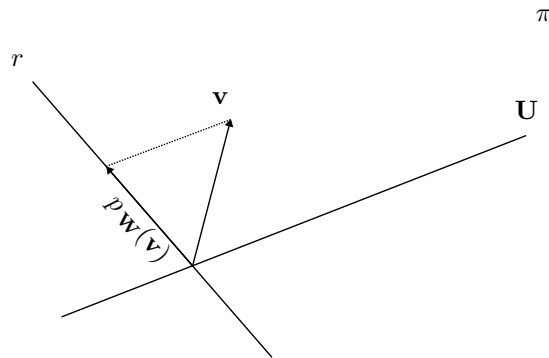


Figure 2: Proiezione di \mathbf{V} su \mathbf{W} nella direzione \mathbf{U}

Notiamo che possiamo riscrivere \mathbf{v} al seguente modo

$$\mathbf{v} = \begin{pmatrix} p_{\mathbf{U}}(\mathbf{v}) \\ p_{\mathbf{W}}(\mathbf{v}) \end{pmatrix}$$

Esempio 4.2. In modo simile, fissato un piano π nello spazio ordinario Σ , detti \mathbf{V} e \mathbf{W} gli spazi dei vettori geometrici di Σ e di π rispettivamente, e fissato $\mathbf{u} \in \mathbf{V} \setminus \mathbf{W}$, si descrive la proiezione p di \mathbf{V} su \mathbf{W} nella direzione $\langle \mathbf{u} \rangle$.

Siano \mathbf{V} un \mathbb{K} -spazio vettoriale ed $\mathbf{e} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una sua base. Per ogni $i = 1, \dots, n$ definiamo

$$\eta_i : \mathbf{V} \rightarrow \mathbb{K}, \quad c_1 \mathbf{e}_1 + \dots + c_n \mathbf{e}_n \mapsto c_i.$$

η_i è un funzionale lineare su \mathbf{V} . Infatti presi $\mathbf{v} = \sum_{k=1}^n c_k \mathbf{e}_k$, $\mathbf{v}' = \sum_{k=1}^n d_k \mathbf{e}_k \in \mathbf{V}$, $\lambda, \lambda' \in \mathbb{K}$, si ha

$$\eta_i(\lambda \mathbf{v} + \lambda' \mathbf{v}') = \eta_i \left(\sum_{k=1}^n (\lambda c_k + \lambda' d_k) \mathbf{e}_k \right) = \lambda c_i + \lambda' d_i = \lambda \eta_i(\mathbf{v}) + \lambda' \eta_i(\mathbf{v}')$$

Siano \mathbf{V} un \mathbb{K} -spazio vettoriale e \mathbf{U} , \mathbf{W} sottospazi supplementari in \mathbf{V} . Definiamo

$$\rho : \mathbf{V} \rightarrow \mathbf{V}, \quad \mathbf{u} + \mathbf{w} \mapsto \mathbf{u} - \mathbf{w} \quad \forall \mathbf{v} = \mathbf{u} + \mathbf{w}, \mathbf{u} \in \mathbf{U}, \mathbf{w} \in \mathbf{W}$$

ρ è un operatore lineare. Infatti presi $\mathbf{v} = \mathbf{u} + \mathbf{w}$, $\mathbf{v}' = \mathbf{u}' + \mathbf{w}'$, $\lambda \in \mathbb{K}$:

$$\rho(\mathbf{v} + \mathbf{v}') = \rho(\mathbf{u} + \mathbf{w} + \mathbf{u}' + \mathbf{w}') = \rho(\mathbf{u} + \mathbf{u}' + \mathbf{w} + \mathbf{w}') = \mathbf{u} + \mathbf{u}' - \mathbf{w} - \mathbf{w}' = \mathbf{u} - \mathbf{w} + \mathbf{u}' - \mathbf{w}' = \rho(\mathbf{v}) + \rho(\mathbf{v}')$$

$$\rho(\lambda \mathbf{v}) = \rho(\lambda \mathbf{u} + \lambda \mathbf{w}) = \lambda \mathbf{u} - \lambda \mathbf{w} = \lambda(\mathbf{u} - \mathbf{w}) = \lambda \rho(\mathbf{v}).$$

L'operatore ρ gode delle seguenti proprietà

1. $\rho(\mathbf{u}) = \mathbf{u} \quad \forall \mathbf{u} \in \mathbf{U}$;
2. $\rho \circ \rho = \mathbf{1}_{\mathbf{V}}$.

La seconda proprietà implica che ρ è invertibile, e quindi $\rho \in \text{GL}(\mathbf{V})$. Se $\mathbf{e} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ è una base di \mathbf{V} e $1 \leq k < n$,

$$\mathbf{U} = \langle \mathbf{e}_1, \dots, \mathbf{e}_k \rangle, \quad \mathbf{W} = \langle \mathbf{e}_{k+1}, \dots, \mathbf{e}_n \rangle,$$

allora

$$\rho \left(\sum_{i=1}^n c_i \mathbf{e}_i \right) = \rho \left(\sum_{i=1}^k c_i \mathbf{e}_i + \sum_{i=k+1}^n c_i \mathbf{e}_i \right) = \sum_{i=1}^k c_i \mathbf{e}_i - \sum_{i=k+1}^n c_i \mathbf{e}_i$$

Siano \mathbf{V} un \mathbb{K} -spazio vettoriale, \mathbf{U} un suo sottospazio e consideriamo lo spazio vettoriale quoziente \mathbf{V}/\mathbf{U} . L'applicazione

$$p : \mathbf{V} \rightarrow \mathbf{V}/\mathbf{U}, \quad \mathbf{v} \mapsto \mathbf{v} + \mathbf{U}$$

è lineare. Essa è chiamata **proiezione naturale** di \mathbf{V} su \mathbf{V}/\mathbf{U}

Sia \mathbf{U} un sottospazio del \mathbb{K} -spazio vettoriale \mathbf{V} . L'inclusione i di \mathbf{U} in \mathbf{V} è un'applicazione lineare.

Proposizione 4.11. Siano $F : \mathbf{V} \rightarrow \mathbf{W}$ un'applicazione lineare di \mathbb{K} -spazi vettoriali, $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbf{V}$ e $\mathbf{w}_i = F(\mathbf{v}_i)$, $i = 1, \dots, n$. Allora

1. se $\mathbf{v}_1, \dots, \mathbf{v}_n$ sono linearmente dipendenti, anche $\mathbf{w}_1, \dots, \mathbf{w}_n$ sono linearmente dipendenti;
2. se $\mathbf{w}_1, \dots, \mathbf{w}_n$ sono linearmente indipendenti, anche $\mathbf{v}_1, \dots, \mathbf{v}_n$ sono linearmente indipendenti.

Proof.

1. Siano $c_1, \dots, c_n \in \mathbb{K}$ non tutti nulli tali che $\sum_{i=1}^n c_i \mathbf{v}_i = \mathbf{0}$. Allora si ha

$$\sum_{i=1}^n c_i \mathbf{w}_i = \sum_{i=1}^n c_i F(\mathbf{v}_i) = F \left(\sum_{i=1}^n c_i \mathbf{v}_i \right) = F(\mathbf{0}) = \mathbf{0},$$

e quindi $\mathbf{w}_1, \dots, \mathbf{w}_n$ sono linearmente dipendenti.

2. Da fare tu.

□

Teorema 4.12. Siano \mathbf{V} e \mathbf{W} due \mathbb{K} -spazi vettoriali, $\mathbf{e} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una base di \mathbf{V} , e $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$ vettori arbitrari di \mathbf{W} . Allora

$$\exists! F \in \text{hom}(\mathbf{V}, \mathbf{W}) : \forall i = 1, \dots, n, \quad F(\mathbf{e}_i) = \mathbf{w}_i.$$

Proof. Se F esiste, è unica, perchè

$$\forall \mathbf{v} = \sum_{i=1}^n x_i \mathbf{e}_i \in \mathbf{V}$$

avremo

$$F(\mathbf{v}) = F\left(\sum_{i=1}^n x_i \mathbf{e}_i\right) = \sum_{i=1}^n x_i F(\mathbf{e}_i) = \sum_{i=1}^n x_i \mathbf{w}_i. \quad (107)$$

e i coefficienti x_1, \dots, x_n saranno univocamente determinati poichè \mathbf{e} è una base. Sarà pertanto sufficiente dimostrare che l'applicazione F definita dalla 107 è lineare. Verifichiamo la condizione 103. Se

$$\mathbf{v} = \sum_{i=1}^n x_i \mathbf{e}_i, \quad \mathbf{v}' = \sum_{i=1}^n y_i \mathbf{e}_i$$

sono elementi di \mathbf{V} , si ha

$$F(\mathbf{v} + \mathbf{v}') = \sum_{i=1}^n (x_i + y_i) \mathbf{w}_i = \sum_{i=1}^n x_i \mathbf{w}_i + \sum_{i=1}^n y_i \mathbf{w}_i = F(\mathbf{v}) + F(\mathbf{v}').$$

Verifichiamo la 104. Se $c \in \mathbb{K}$ e $\mathbf{v} = \sum_{i=1}^n x_i \mathbf{e}_i$ si ha

$$F(c\mathbf{v}) = \sum_{i=1}^n cx_i \mathbf{e}_i = c \left(\sum_{i=1}^n x_i \mathbf{e}_i \right) = cF(\mathbf{v}).$$

Quindi F è lineare. □

Nucleo e Immagine di un omomorfismo

Definizione 4.8. Sia $F : \mathbf{V} \rightarrow \mathbf{W}$ un'applicazione lineare di \mathbb{K} -spazi vettoriali. Il **nucleo** di F è

$$\ker(F) = \{\mathbf{v} \in \mathbf{V} : F(\mathbf{v}) = \mathbf{0}\}.$$

L'**immagine** di F è il sottoinsieme di \mathbf{W}

$$\text{Im}(F) = \{\mathbf{w} = F(\mathbf{v}) : \mathbf{v} \in \mathbf{V}\}.$$

Proposizione 4.13. $\ker(F)$ e $\text{Im}(F)$ sono sottospazi vettoriali di \mathbf{V} e \mathbf{W} rispettivamente.

Proof. 1. Siano $\mathbf{v}, \mathbf{v}' \in \ker(F)$ e $c, c' \in \mathbb{K}$. Si ha:

$$F(c\mathbf{v} + c'\mathbf{v}') = cF(\mathbf{v}) + c'F(\mathbf{v}') = c\mathbf{0} + c'\mathbf{0} = \mathbf{0}.$$

Quindi $c\mathbf{v} + c'\mathbf{v}' \in \ker(F)$, e dunque $\ker(F)$ è un sottospazio vettoriale di \mathbf{V} .

2. Siano $\mathbf{w}, \mathbf{w}' \in \text{Im}(F)$ e $c, c' \in \mathbb{K}$. Allora esistono $\mathbf{v}, \mathbf{v}' \in \mathbf{V}$ tali che:

$$\mathbf{w} = F(\mathbf{v}), \quad \mathbf{w}' = F(\mathbf{v}').$$

Quindi:

$$c\mathbf{w} + c'\mathbf{w}' = cF(\mathbf{v}) + c'F(\mathbf{v}') = F(c\mathbf{v} + c'\mathbf{v}') \in \text{Im}(F).$$

Pertanto $\text{Im}(F)$ è un sottospazio vettoriale di \mathbf{W} . □

Rango e Nullità di un omomorfismo

Definizione 4.9. Se $\ker(F)$ e $\text{Im}(F)$ hanno dimensione finita, la dimensione di $\text{Im}(F)$ si dice **rank** di F , e si denota con $r(F)$, la dimensione di $\ker(F)$ si dice **nullità** di F .

Nota 18. Dalla 4.9 segue che se $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ è una base di \mathbf{V} , allora

$$\text{Im}(F) = \langle F(\mathbf{e}_1), \dots, F(\mathbf{e}_n) \rangle.$$

Proposizione 4.14. Un'applicazione lineare $F: \mathbf{V} \rightarrow \mathbf{W}$ di \mathbb{K} -spazi vettoriali è iniettiva se e solo se $\ker(F) = \langle \mathbf{0} \rangle$.

Proof. Se F è iniettiva, allora $\mathbf{0}$ è l'unico elemento di $\ker(F)$, e quindi $\ker(F) = \langle \mathbf{0} \rangle$. Viceversa, supponiamo che $\ker(F) = \langle \mathbf{0} \rangle$. Se $\mathbf{v}, \mathbf{v}' \in \mathbf{V}$ sono tali che

$$F(\mathbf{v}) = F(\mathbf{v}'),$$

allora

$$F(\mathbf{v} - \mathbf{v}') = F(\mathbf{v}) - F(\mathbf{v}') = \mathbf{0}$$

e quindi $\mathbf{v} - \mathbf{v}' \in \ker(F)$. Poiché $\mathbf{0}$ è l'unico elemento di $\ker(F)$, si ha $\mathbf{v} - \mathbf{v}' = \mathbf{0}$, cioè $\mathbf{v} = \mathbf{v}'$. Quindi F è iniettiva. \square

Abbiamo il seguente importante teorema.

Teorema 4.15. Sia $F: \mathbf{V} \rightarrow \mathbf{W}$ un'applicazione lineare di \mathbb{K} -spazi vettoriali, con $\dim(\mathbf{V}) = n$. Allora $\ker(F)$ e $\text{Im}(F)$ hanno dimensione finita, e

$$\dim(\ker(F)) + r(F) = n.$$

Proof. Poiché \mathbf{V} ha dimensione finita, anche $\ker(F)$, che è un sottospazio di \mathbf{V} , ha dimensione finita; sia $s = \dim(\ker(F))$. Fissiamo una base $\{\mathbf{n}_1, \dots, \mathbf{n}_s\}$ di $\ker(F)$, e siano $\mathbf{v}_{s+1}, \dots, \mathbf{v}_n \in \mathbf{V}$ tali che $\{\mathbf{n}_1, \dots, \mathbf{n}_s, \mathbf{v}_{s+1}, \dots, \mathbf{v}_n\}$ sia una base di \mathbf{V} . Per completare la dimostrazione del teorema sarà sufficiente dimostrare che $\{F(\mathbf{v}_{s+1}), \dots, F(\mathbf{v}_n)\}$ è una base di $\text{Im}(F)$.

Ogni vettore $\mathbf{w} \in \text{Im}(F)$ è della forma

$$\mathbf{w} = F\left(\sum_{i=1}^s a_i \mathbf{n}_i + \sum_{j=s+1}^n b_j \mathbf{v}_j\right) = \sum_{i=1}^s a_i F(\mathbf{n}_i) + \sum_{j=s+1}^n b_j F(\mathbf{v}_j) = \sum_{j=s+1}^n b_j F(\mathbf{v}_j).$$

per opportuni scalari $a_1, \dots, a_s, b_{s+1}, \dots, b_n$. Quindi $F(\mathbf{v}_{s+1}), \dots, F(\mathbf{v}_n)$ generano $\text{Im}(F)$. Supponiamo che $c_{s+1}, \dots, c_n \in \mathbb{K}$ siano tali che

$$\sum_{i=s+1}^n c_i F(\mathbf{v}_i) = \mathbf{0}.$$

Allora

$$F\left(\sum_{i=s+1}^n c_i \mathbf{v}_i\right) = \mathbf{0} \implies \sum_{i=s+1}^n c_i \mathbf{v}_i \in \ker(F).$$

Poiché $\{\mathbf{n}_1, \dots, \mathbf{n}_s\}$ è una base di $\ker(F)$, esistono $d_1, \dots, d_s \in \mathbb{K}$ tali che

$$\sum_{i=s+1}^n c_i \mathbf{v}_i = \sum_{i=1}^s d_i \mathbf{n}_i,$$

cioè tali che

$$\sum_{i=1}^s d_i \mathbf{n}_i - \sum_{i=s+1}^n c_i \mathbf{v}_i = \mathbf{0}.$$

Ma $\mathbf{n}_1, \dots, \mathbf{n}_s, \mathbf{v}_{s+1}, \dots, \mathbf{v}_n$ sono linearmente indipendenti, e quindi tutti i coefficienti sono nulli; in particolare

$$c_{s+1} = \dots = c_n = 0,$$

ed $F(\mathbf{v}_{s+1}), \dots, F(\mathbf{v}_n)$ sono linearmente indipendenti. \square

Si noti che nell'enunciato del teorema precedente non si è supposto che \mathbf{W} abbia dimensione finita. Si ha il seguente:

Corollario 8. Se \mathbf{U} è un sottospazio del \mathbb{K} -spazio vettoriale \mathbf{V} di dimensione finita, allora

$$\dim(\mathbf{V}/\mathbf{U}) = \dim(\mathbf{V}) - \dim(\mathbf{U}).$$

Proof. La proiezione naturale $p: \mathbf{V} \rightarrow \mathbf{V}/\mathbf{U}$ è suriettiva e ha nucleo uguale a \mathbf{U} . \square

Corollario 9. Se \mathbf{V} e \mathbf{W} sono \mathbb{K} -spazi vettoriali di dimensione finita tali che $\dim(\mathbf{V}) = \dim(\mathbf{W})$, ed $F: \mathbf{V} \rightarrow \mathbf{W}$ è un'applicazione lineare, allora le seguenti condizioni sono equivalenti:

1. $\ker(F) = \langle \mathbf{0} \rangle$;
2. $\text{Im}(F) = \mathbf{W}$;
3. F è un isomorfismo.

Teorema di omomorfismo per gli spazi vettoriali

Teorema 4.16. Sia $F: \mathbf{V} \rightarrow \mathbf{W}$ un'applicazione lineare di \mathbb{K} -spazi vettoriali. F definisce un isomorfismo

$$F': \mathbf{V}/\ker(F) \rightarrow \text{Im}(F)$$

tale che

$$F = i \circ F' \circ p, \tag{108}$$

dove $p: \mathbf{V} \rightarrow \mathbf{V}/\ker(F)$ è la proiezione naturale e $i: \text{Im}(F) \rightarrow \mathbf{W}$ è l'inclusione.

Proof. Siano $\mathbf{v}_1, \mathbf{v}_2 \in \mathbf{V}$. Si ha $F(\mathbf{v}_1) = F(\mathbf{v}_2)$ se e solo se $F(\mathbf{v}_2 - \mathbf{v}_1) = \mathbf{0}$, cioè se e solo se $\mathbf{v}_2 - \mathbf{v}_1 \in \ker(F)$, ovvero

$$F(\mathbf{v}_1) = F(\mathbf{v}_2) \iff \mathbf{v}_2 \in [\mathbf{v}_1 + \ker(F)] \tag{109}$$

Possiamo quindi definire $F': \mathbf{V}/\ker(F) \rightarrow \text{Im}(F)$ ponendo

$$F'(\mathbf{v} + \ker(F)) = F(\mathbf{v}).$$

F' è biunivoca perché, oltre ad essere ovviamente suriettiva, è anche iniettiva per 109. La verifica della linearità di F' e della 4.16 è lasciata al lettore. \square

Spazi Vettoriali Isomorfi

Definizione 4.10. Siano \mathbf{V} e \mathbf{W} due \mathbb{K} -spazi vettoriali. \mathbf{V} e \mathbf{W} si dicono **isomorfi**, oppure si dice che \mathbf{V} è **isomorfo** a \mathbf{W} , se esiste un isomorfismo $\mathbf{V} \rightarrow \mathbf{W}$.

Ogni spazio vettoriale è isomorfo a se stesso, perché $1_{\mathbf{V}}$ è un isomorfismo. Se \mathbf{V} è isomorfo a \mathbf{W} , anche \mathbf{W} è isomorfo a \mathbf{V} , perché l'inverso di un isomorfismo è un isomorfismo. Infine, poiché la composizione di isomorfismi è un isomorfismo, se \mathbf{V} è isomorfo a \mathbf{W} e \mathbf{W} è isomorfo a \mathbf{U} , allora \mathbf{V} è isomorfo a \mathbf{U} . Quindi l'isomorfismo è una relazione di equivalenza tra spazi vettoriali.

Teorema 4.17. Due \mathbb{K} -spazi vettoriali di dimensione finita sono isomorfi se e solo se hanno la stessa dimensione.

Proof. Supponiamo \mathbf{V} e \mathbf{W} isomorfi, e sia $F: \mathbf{V} \rightarrow \mathbf{W}$ un isomorfismo. Poiché $\ker(F) = \langle \mathbf{0} \rangle$, si ha

$$\dim(\mathbf{W}) = r(F) = \dim(\mathbf{V}).$$

Viceversa supponiamo $\dim(\mathbf{V}) = n = \dim(\mathbf{W})$, e siano $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ e $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ basi di \mathbf{V} e di \mathbf{W} rispettivamente. L'applicazione lineare $F: \mathbf{V} \rightarrow \mathbf{W}$ definita da

$$F(\mathbf{v}_i) = \mathbf{w}_i, \quad i = 1, \dots, n,$$

è suriettiva perché $\mathbf{w}_1, \dots, \mathbf{w}_n$ generano \mathbf{W} . Segue $\dim(\ker(F)) = 0$, e quindi F è anche iniettiva. Dunque F è un isomorfismo. \square

Spazio Vettoriale Duale

Proposizione 4.18. Sia \mathbf{V} un \mathbb{K} -spazio vettoriale. È possibile introdurre in $\mathbf{V}^\sim (\text{hom}(\mathbf{V}, \mathbb{K}))$, l'insieme dei funzionali lineari su \mathbf{V} , una struttura di \mathbb{K} -spazio vettoriale definendo le seguenti operazioni.

$$+ : \mathbf{V}^\sim \times \mathbf{V}^\sim \rightarrow \mathbf{V}^\sim, \quad (L_1, L_2) \mapsto L_1 + L_2 \quad \cdot : \mathbb{K} \times \mathbf{V}^\sim \rightarrow \mathbf{V}^\sim, \quad (c, L) \mapsto cL.$$

Ponendo rispettivamente

$$L_1 + L_2 : \mathbf{V} \rightarrow \mathbb{K}, \quad \mathbf{v} \mapsto (L_1 + L_2)(\mathbf{v}) = L_1(\mathbf{v}) + L_2(\mathbf{v}),$$

$$cL : \mathbf{V} \rightarrow \mathbb{K}, \quad \mathbf{v} \mapsto (cL)(\mathbf{v}) = cL(\mathbf{v}).$$

Proof. Da ultimare. □

Supponiamo che \mathbf{V} abbia dimensione finita, e sia $\mathbf{e} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una sua base. Sia $1 \leq i \leq n$. Per la (4.12) abbiamo che

$$\exists! \eta_i \in \mathbf{V}^\sim : \eta_i(\mathbf{e}_j) = \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}. \quad (110)$$

Otteniamo così n funzionali lineari η_1, \dots, η_n .

Teorema 4.19. Sia \mathbf{V} un \mathbb{K} -spazio vettoriale di dimensione finita, e sia $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una sua base. L'insieme $\{\eta_1, \dots, \eta_n\}$ dei funzionali lineari definiti dalla (110) è una base di \mathbf{V}^\sim ; in particolare $\dim(\mathbf{V}^\sim) = \dim(\mathbf{V})$ e quindi \mathbf{V} e \mathbf{V}^\sim sono isomorfi.

Proof. Siano $L \in \mathbf{V}^\sim$ e $a_k = L(\mathbf{e}_k) \quad \forall k = 1, \dots, n$. Il funzionale $\sum_{k=1}^n a_k \eta_k$ soddisfa l'identità

$$\left(\sum_{k=1}^n a_k \eta_k \right) (\mathbf{e}_i) = \sum_{k=1}^n (a_k \eta_k)(\mathbf{e}_i) = \sum_{k=1}^n a_k \eta_k(\mathbf{e}_i) = a_i \quad \forall i = 1, \dots, n.$$

e quindi $\sum_{k=1}^n a_k \eta_k$ ha gli stessi valori di L sulla base $\{\eta_1, \dots, \eta_n\}$. Per la (4.12) abbiamo che $\sum_{k=1}^n a_k \eta_k = L$, quindi η_1, \dots, η_n generano \mathbf{V}^\sim . Supponiamo ora che $a_1, \dots, a_n \in \mathbb{K}$ tali che $\sum_{k=1}^n a_k \eta_k = \mathbf{0}$. Allora

$$\forall i = 1, \dots, n, \quad 0 = \mathbf{0}(\mathbf{e}_i) = \left(\sum_{k=1}^n a_k \eta_k \right) (\mathbf{e}_i) = a_i.$$

Quindi $a_1 = \dots = a_n = 0$ e pertanto η_1, \dots, η_n sono anche linearmente indipendenti. □

L'insieme dei funzionali lineari $\{\eta_1, \dots, \eta_n\}$ è la **base duale** della base $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ di \mathbf{V} . Se $L = \sum_{k=1}^n a_k \eta_k \in \mathbf{V}^\sim$, si ha

$$L \left(\sum_{i=1}^n x_i \mathbf{e}_i \right) = \sum_{k=1}^n a_k \eta_k \sum_{i=1}^n x_i \mathbf{e}_i = \sum_{k=1}^n \sum_{i=1}^n a_k \eta_k x_i \mathbf{e}_i = a_1 x_1 + a_2 x_2 + \dots + a_n x_n,$$

perchè

$$\eta_i \left(\sum_{k=1}^n x_k \mathbf{e}_k \right) = x_i \quad \forall i = 1, \dots, n.$$

Quindi ogni funzionale lineare su \mathbf{V} si esprime in modo unico come un polinomio omogeneo di primo grado nelle coordinate dei vettori rispetto alla base $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$. I coefficienti del polinomio sono le coordinate del funzionale rispetto alla base duale $\{\eta_1, \dots, \eta_n\}$.

Caso 7. In particolare, se $\mathbf{V} = \mathbb{K}^n$, e $\{\mathbf{E}_1, \dots, \mathbf{E}_n\}$ è la base canonica, ogni funzionale lineare L su \mathbb{K}^n si esprime in modo unico come un polinomio omogeneo di primo grado:

$$L \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^n a_i x_i. \quad (111)$$

Si osservi che se il funzionale $L : \mathbf{V} \rightarrow \mathbb{K}$ non è nullo, la sua immagine è un sottospazio vettoriale di \mathbb{K} diverso da $\langle 0 \rangle$, e quindi $\text{Im}(L) = \mathbb{K}$; da ciò e dal teorema (ancora da fare) segue che $\dim(\text{N}(L)) = n - 1$.

Esempio 4.3. Siano

$$\mathbf{e} = \left\{ \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 4 \end{pmatrix}, \begin{pmatrix} 5 \\ 10 \\ 0 \end{pmatrix} \right\}, \quad \mathbf{E} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

due basi di $\mathbf{V} = \mathbb{K}^3$. Siano $L_1, L_2 \in \mathbf{V}^\sim$ definite al seguente modo

$$L_1 : \mathbf{V} \rightarrow \mathbb{K}, \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto x + y + z, \quad L_2 : \mathbf{V} \rightarrow \mathbb{K}, \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto 2x + 3y + 4z.$$

Siano ora

$$\mathbf{v} = \begin{pmatrix} 19 \\ 31 \\ 2 \end{pmatrix} = 2 \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix} - \begin{pmatrix} 0 \\ -1 \\ 4 \end{pmatrix} + 3 \begin{pmatrix} 5 \\ 10 \\ 0 \end{pmatrix}$$

- base \mathbf{e} , funzionale L_1 :

$$L_1(\mathbf{e}_1) = 5, L_1(\mathbf{e}_2) = 3, L_1(\mathbf{e}_3) = 15 \implies L_1 = 5\eta_1 + 3\eta_2 + 15\eta_3$$

$$L_1(\mathbf{v}) = 19 + 31 + 2 = 52 \quad (5\eta_1 + 3\eta_2 + 15\eta_3)(2\mathbf{e}_1 - \mathbf{e}_2 + 3\mathbf{e}_3) = 5 \cdot 2 + 3 \cdot (-1) + 15 \cdot 3 = 52$$

- base \mathbf{e} , funzionale L_2 :

$$L_2(\mathbf{e}_1) = 16, L_2(\mathbf{e}_2) = 13, L_2(\mathbf{e}_3) = 40 \implies L_2 = 16\eta_1 + 13\eta_2 + 40\eta_3$$

- base \mathbf{E} , funzionale L_1 :

$$L_1(\mathbf{E}_1) = 1, L_1(\mathbf{E}_2) = 1, L_1(\mathbf{E}_3) = 1 \implies L_1 = \eta_1 + \eta_2 + \eta_3$$

- base \mathbf{E} , funzionale L_2 :

$$L_2(\mathbf{E}_1) = 2, L_2(\mathbf{E}_2) = 3, L_2(\mathbf{E}_3) = 4 \implies L_2 = 2\eta_1 + 3\eta_2 + 4\eta_3$$

Proposizione 4.20. Siano \mathbf{V} un \mathbb{K} -spazio vettoriale ed $f, g \in \mathbf{V}^\sim$ tali che si abbia $f(\mathbf{v}) = 0$ per ogni $\mathbf{v} \in \text{N}(g)$, cioè $\ker(f) \supset \text{N}(g)$. Allora

$$\exists x \in \mathbb{K} : f = cg.$$

Proof. Se $g = 0$, allora $\text{N}(g) = \mathbf{V}$ e quindi $\ker(f) = \mathbf{V}$, cioè $f = 0 = g$. Supponiamo viceversa $g \neq 0$. In tal caso $\text{N}(g)$ è un iperpiano. Fissiamo $\mathbf{e} \in \mathbf{V} \setminus \text{N}(g)$, e sia $c \in \mathbb{K}$ tale che $f(\mathbf{e}) = cg(\mathbf{e})$. Si ha $\mathbf{V} = \langle \mathbf{e} \rangle \oplus \text{N}(g)$ e quindi ogni $\mathbf{v} \in \mathbf{V}$ è della forma $\mathbf{v} = x\mathbf{e} + \mathbf{w}$, per qualche $\mathbf{w} \in \text{N}(g)$ e $x \in \mathbb{K}$. Pertanto

$$\begin{aligned} f(\mathbf{v}) &= f(x\mathbf{e} + \mathbf{w}) = f(x\mathbf{e}) + f(\mathbf{w}) = xf(\mathbf{e}) = xcg(\mathbf{e}) = cg(x\mathbf{e}) = \\ &= cg(x\mathbf{e}) + c\mathbf{0} = cg(x\mathbf{e}) + cg(\mathbf{w}) = cg(x\mathbf{e} + \mathbf{w}) = cg(\mathbf{v}). \end{aligned}$$

cioè $f = cg$. □

Dati due \mathbb{K} -spazi vettoriali \mathbf{V} e \mathbf{W} , è possibile introdurre in $\text{hom}(\mathbf{V}, \mathbf{W})$ una struttura di \mathbb{K} -spazio vettoriale definendo le seguenti operazioni.

$$+ : \text{hom}(\mathbf{V}, \mathbf{W}) \times \text{hom}(\mathbf{V}, \mathbf{W}) \rightarrow \text{hom}(\mathbf{V}, \mathbf{W}), (F, G) \mapsto F + G \quad \cdot : \mathbb{K} \times \text{hom}(\mathbf{V}, \mathbf{W}) \rightarrow \text{hom}(\mathbf{V}, \mathbf{W}), (\lambda, F) \mapsto cF.$$

Ponendo rispettivamente

$$F + G : \mathbf{V} \rightarrow \mathbf{W}, \quad \mathbf{v} \mapsto (F + G)(\mathbf{v}) = F(\mathbf{v}) + G(\mathbf{v}),$$

$$\lambda F : \mathbf{V} \rightarrow \mathbf{W}, \quad \mathbf{v} \mapsto (\lambda F)(\mathbf{v}) = \lambda F(\mathbf{v}).$$

Verifichiamo la linearità di entrambi i funzionali, $\forall \mathbf{v}, \mathbf{v}' \in \mathbf{V}, \lambda \in \mathbb{K}$,

$F + G$

$$\begin{aligned}(F + G)(\mathbf{v} + \mathbf{v}') &= F(\mathbf{v} + \mathbf{v}') + G(\mathbf{v} + \mathbf{v}') = F(\mathbf{v}) + F(\mathbf{v}') + G(\mathbf{v}) + G(\mathbf{v}') \\ &= F(\mathbf{v}) + G(\mathbf{v}) + F(\mathbf{v}') + G(\mathbf{v}') = (F + G)(\mathbf{v}) + (F + G)(\mathbf{v}')\end{aligned}$$

$$(F + G)(\lambda \mathbf{v}) = F(\lambda \mathbf{v}) + G(\lambda \mathbf{v}) = \lambda F(\mathbf{v}) + \lambda G(\mathbf{v}) = \lambda(F(\mathbf{v}) + G(\mathbf{v})) = \lambda(F + G)(\mathbf{v}).$$

λF

$$(\lambda F)(\mathbf{v} + \mathbf{v}') = \lambda F(\mathbf{v} + \mathbf{v}') = \lambda(F(\mathbf{v}) + F(\mathbf{v}')) = \lambda F(\mathbf{v}) + \lambda F(\mathbf{v}') = (\lambda F)(\mathbf{v}) + (\lambda F)(\mathbf{v}')$$

$$(\lambda F)(\mu \mathbf{v}) = \lambda F(\mu \mathbf{v}) = \mu(\lambda F(\mathbf{v})) = \mu(\lambda F)(\mathbf{v}).$$

L'applicazione nulla $\mathbf{0} : \mathbf{V} \rightarrow \mathbf{W}$, $\mathbf{v} \mapsto \mathbf{0}$ soddisfa $F + \mathbf{0} = F \quad \forall F \in \text{hom}(\mathbf{V}, \mathbf{W})$. Con queste operazioni resta definita su $\text{hom}(\mathbf{V}, \mathbf{W})$ una struttura di \mathbb{K} -spazio vettoriale.

Nel caso particolare $\mathbf{V} = \mathbb{K}$, lo spazio $\text{hom}(\mathbb{K}, \mathbf{W})$ è isomorfo a \mathbf{W} : un isomorfismo $\lambda : \mathbf{W} \rightarrow \text{hom}(\mathbb{K}, \mathbf{W})$ è definito associando a $\mathbf{w} \in \mathbf{W}$ l'applicazione lineare $\lambda(\mathbf{w}) : \mathbb{K} \rightarrow \mathbf{W}$ tale che $\lambda(\mathbf{w})(1) = \mathbf{w}$; verifica che λ sia un isomorfismo. λ si chiama **isomorfismo canonico** di \mathbf{W} su $\text{hom}(\mathbb{K}, \mathbf{W})$, perchè è indicato univocamente da \mathbf{W} .

Sia \mathbf{V} un \mathbb{K} -spazio vettoriale. Il duale di \mathbf{V}^\sim , cioè lo spazio $\text{hom}(\mathbf{V}^\sim, \mathbb{K})$, si chiama **spazio biduale** di \mathbf{V} , e si denota con $\mathbf{V}^{\sim\sim}$. Se $\dim(\mathbf{V}) = n$, allora segue dal teorema 11.12 che

$$\dim(\mathbf{V}^{\sim\sim}) = \dim(\mathbf{V}^\sim) = \dim(\mathbf{V}) = n,$$

e i tre spazi \mathbf{V} , \mathbf{V}^\sim , $\mathbf{V}^{\sim\sim}$ sono tra loro isomorfi.

Se si sceglie una base $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ di \mathbf{V} , resta definita la base duale $\{\eta_1, \dots, \eta_n\}$ di \mathbf{V}^\sim , e un isomorfismo $\varphi : \mathbf{V} \rightarrow \mathbf{V}^{\sim\sim}$ è definito ponendo $\varphi(\mathbf{e}_i) = \eta_i$, $i = 1, \dots, n$.

L'isomorfismo φ dipende dalla base assegnata, cioè se si sceglie un'altra base di \mathbf{V} , si ottiene un isomorfismo diverso da φ .

Per lo spazio $\mathbf{V}^{\sim\sim}$ si ha una situazione diversa: è infatti possibile definire un isomorfismo

$$\beta : \mathbf{V} \rightarrow \mathbf{V}^{\sim\sim}$$

in modo del tutto intrinseco, cioè indipendente da qualunque scelta ulteriore, il quale per questo motivo si dice **l'isomorfismo canonico** di \mathbf{V} su $\mathbf{V}^{\sim\sim}$.

β è l'applicazione che ad ogni $\mathbf{v} \in \mathbf{V}$ associa il funzionale $\beta(\mathbf{v}) : \mathbf{V}^{\sim\sim} \rightarrow \mathbb{K}$ definito nel modo seguente:

$$\beta(\mathbf{v})(L) = L(\mathbf{v})$$

per ogni $L \in \mathbf{V}^\sim$.

Verifichiamo che $\beta(\mathbf{v})$ è lineare:

$$\begin{aligned}\beta(\mathbf{v})(cL + c'L') &= (cL + c'L')(\mathbf{v}) = cL(\mathbf{v}) + c'L'(\mathbf{v}) \\ &= c(L(\mathbf{v})) + c'(L'(\mathbf{v})) = c\beta(\mathbf{v})(L) + c'\beta(\mathbf{v})(L'),\end{aligned}$$

per ogni $L, L' \in \mathbf{V}^\sim$, $c, c' \in \mathbb{K}$, e quindi $\beta(\mathbf{v})$ è un funzionale lineare su $\mathbf{V}^{\sim\sim}$.

Verifichiamo ora che $\beta : \mathbf{V} \rightarrow \mathbf{V}^{\sim\sim}$ è lineare. Per ogni $\mathbf{v}, \mathbf{v}' \in \mathbf{V}$, $c, c' \in \mathbb{K}$, si ha:

$$\begin{aligned}\beta(c\mathbf{v} + c'\mathbf{v}')(L) &= L(c\mathbf{v} + c'\mathbf{v}') = cL(\mathbf{v}) + c'L(\mathbf{v}') \\ &= c\beta(\mathbf{v})(L) + c'\beta(\mathbf{v}')(L) = [c\beta(\mathbf{v}) + c'\beta(\mathbf{v}')](L)\end{aligned}$$

per ogni $L \in \mathbf{V}^\sim$, e quindi $\beta(c\mathbf{v} + c'\mathbf{v}') = c\beta(\mathbf{v}) + c'\beta(\mathbf{v}')$, cioè β è lineare.

Per dimostrare che β è un isomorfismo è sufficiente far vedere che

$$\mathbf{N}(\beta) = \{0\},$$

perché $\dim(\mathbf{V}) = \dim(\mathbf{V}^{\sim\sim})$. Supponiamo per assurdo che esista $\mathbf{v} \in \mathbf{V}$, $\mathbf{v} \neq 0$, tale che $\beta(\mathbf{v}) = 0 \in \mathbf{V}^{\sim\sim}$. Allora si ha:

$$\beta(\mathbf{v})(L) = L(\mathbf{v}) = 0$$

per ogni $L \in \mathbf{V}^\sim$. Siano $\mathbf{v}_2, \dots, \mathbf{v}_n \in \mathbf{V}$ tali che $\{\mathbf{v}, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ sia una base di \mathbf{V} . Il funzionale $L \in \mathbf{V}^\sim$ definito da

$$L(\mathbf{v}) = 1, \quad L(\mathbf{v}_i) = 0, \quad i = 2, \dots, n,$$

dà una contraddizione.

Sia \mathbf{V} un \mathbb{K} -spazio vettoriale, $\dim(\mathbf{V}) = n$. Se $L \in \mathbf{V}^\sim$, i vettori del nucleo di L , cioè i $\mathbf{v} \in \mathbf{V}$ tali che $L(\mathbf{v}) = 0$, si dicono talvolta **ortogonali** a L . Se Φ è un sottoinsieme di \mathbf{V}^\sim , un vettore $\mathbf{v} \in \mathbf{V}$ si dirà **ortogonale a Φ** se \mathbf{v} è ortogonale ad ogni $L \in \Phi$, cioè se $L(\mathbf{v}) = 0$ per ogni $L \in \Phi$. L'insieme

$$\Phi^\perp = \{\mathbf{v} \in \mathbf{V} : \mathbf{v} \text{ è ortogonale a } \Phi\} = \bigcap_{L \in \Phi} N(L) \subset \mathbf{V}$$

è un sottospazio vettoriale perché è l'intersezione di una famiglia di sottospazi vettoriali di \mathbf{V} . Chiameremo Φ^\perp il **sottospazio di \mathbf{V} ortogonale a Φ** .

Supponiamo in particolare che Φ sia un sottospazio vettoriale di \mathbf{V}^\sim . Sia $\dim(\Phi) = t$ e $\{L_1, \dots, L_t\}$ una base di Φ . Allora si ha

$$\Phi^\perp = N(L_1) \cap \dots \cap N(L_t) \quad (112)$$

e

$$\dim(\Phi^\perp) = n - t. \quad (113)$$

Infatti, se $\mathbf{v} \in N(L_1) \cap \dots \cap N(L_t)$, allora per ogni $L = a_1 L_1 + a_2 L_2 + \dots + a_t L_t$ si ha

$$L(\mathbf{v}) = a_1 L_1(\mathbf{v}) + a_2 L_2(\mathbf{v}) + \dots + a_t L_t(\mathbf{v}) = 0,$$

cioè $\mathbf{v} \in N(L)$. Pertanto $N(L_1) \cap \dots \cap N(L_t) \subset \Phi^\perp$. L'inclusione opposta è ovviamente vera, e da 112 segue.

Sia $m = \dim(\Phi^\perp)$. Osserviamo che Φ^\perp coincide con il nucleo dell'applicazione lineare

$$\varphi : \mathbf{V} \rightarrow \mathbb{K}^t$$

definita da $\varphi(\mathbf{v}) = (L_1(\mathbf{v}), L_2(\mathbf{v}), \dots, L_t(\mathbf{v}))$. Dal teorema 11.6 segue che $m \geq n - t$.

Supponiamo per assurdo che sia $m > n - t$. Scegliamo una base $\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$ di \mathbf{V} tale che $\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$ sia una base di Φ^\perp .

La matrice $t \times n$:

$$\begin{pmatrix} 0 & \dots & 0 & L_1(\mathbf{e}_{m+1}) & \dots & L_1(\mathbf{e}_n) \\ 0 & \dots & 0 & L_2(\mathbf{e}_{m+1}) & \dots & L_2(\mathbf{e}_n) \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & L_t(\mathbf{e}_{m+1}) & \dots & L_t(\mathbf{e}_n) \end{pmatrix}$$

ha rango non superiore a $t - 1$, perché ha $m > n - t$ colonne nulle. D'altra parte le sue t righe sono linearmente indipendenti, perché sono i vettori delle coordinate di L_1, \dots, L_t rispetto alla base $\{\eta_1, \dots, \eta_n\}$, duale di $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$. Abbiamo una contraddizione, e quindi dev'essere $m = n - t$.

Siano dati due spazi vettoriali complessi \mathbf{V}, \mathbf{W} . Un'applicazione

$$F : \mathbf{V} \rightarrow \mathbf{W}$$

si dice **antilineare** se soddisfa le seguenti condizioni:

$$F(\mathbf{v} + \mathbf{v}') = F(\mathbf{v}) + F(\mathbf{v}'), \quad (114)$$

$$F(c\mathbf{v}) = \bar{c}F(\mathbf{v}), \quad (115)$$

per ogni $\mathbf{v}, \mathbf{v}' \in \mathbf{V}$, $c \in \mathbb{C}$. Una condizione equivalente alle 114, 115 è la seguente:

$$F(c\mathbf{v} + c'\mathbf{v}') = \bar{c}F(\mathbf{v}) + \bar{c}'F(\mathbf{v}')$$

per ogni $\mathbf{v}, \mathbf{v}' \in \mathbf{V}$, $c, c' \in \mathbb{C}$. La verifica dell'equivalenza è lasciata al lettore.

4.2 Applicazioni Lineari e matrici, cambiamenti di coordinate affini

Matrice associata a F rispetto alle basi ψ e ϕ

Definizione 4.11. Siano \mathbf{V} e \mathbf{W} due \mathbb{K} -spazi vettoriali di dimensione finita, $\psi = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \in b(\mathbf{V})$ e $\phi = \{\mathbf{w}_1, \dots, \mathbf{w}_m\} \in b(\mathbf{W})$, e $F \in \text{hom}(\mathbf{V}, \mathbf{W})$. La matrice $m \times n$ la cui j -esima colonna, $j \in \{1, \dots, n\}$, è costituita dalle coordinate del vettore $F(\mathbf{v}_j) \in \mathbf{W}$ rispetto alla base ϕ è detta **matrice associata a F rispetto alle basi ψ e ϕ** , e si denota con $M_{\phi, \psi}(F)$. Esplicitamente:

$$M_{\phi, \psi}(F) = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{m1} & m_{m2} & \cdots & m_{mn} \end{pmatrix}, \quad F(\mathbf{v}_j) = \sum_{i=1}^m m_{ij} \mathbf{w}_i \quad j = 1, \dots, n.$$

Nota 19. $M_{\phi, \psi}(F)$ dipende, oltre che da F , anche dalle basi ψ e ϕ .

L'utilità di $M_{\phi, \psi}(F)$ sta nel fatto che una volta assegnate ψ e ϕ , da essa si può risalire a F , come illustrato nella seguente proposizione.

Proposizione 4.21. Siano \mathbf{V} e \mathbf{W} due \mathbb{K} -spazi vettoriali di dimensione finita, $\psi = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \in b(\mathbf{V})$ e $\phi = \{\mathbf{w}_1, \dots, \mathbf{w}_m\} \in b(\mathbf{W})$, e $F \in \text{hom}(\mathbf{V}, \mathbf{W})$. Allora

$$\forall \mathbf{v} = \sum_{j=1}^n x_j \mathbf{v}_j \in \mathbf{V}, \quad F(\mathbf{v}) = \sum_{i=1}^m y_i \mathbf{w}_i$$

dove

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = M_{\phi, \psi}(F) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Proof. Si ha:

$$F(\mathbf{v}) = F\left(\sum_{j=1}^n x_j \mathbf{v}_j\right) = \sum_{j=1}^n x_j F(\mathbf{v}_j) = \sum_{j=1}^n x_j \left(\sum_{i=1}^m m_{ij} \mathbf{w}_i\right) = \sum_{i=1}^m \left(\sum_{j=1}^n m_{ij} x_j\right) \mathbf{w}_i$$

Quindi il vettore colonna delle coordinate di $F(\mathbf{v})$ è

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n m_{1j} x_j \\ \sum_{j=1}^n m_{2j} x_j \\ \vdots \\ \sum_{j=1}^n m_{mj} x_j \end{pmatrix} = M_{\phi, \psi}(F) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

□

Esempio 4.4. Siano

$$F: \mathbb{R}^2 \rightarrow \mathbb{R}^3, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x - y \\ x + 2y \\ 3x \end{pmatrix}$$

con

$$\psi = \left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \end{pmatrix} \right\}, \quad \phi = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} \right\}.$$

basi di \mathbb{R}^2 e \mathbb{R}^3 rispettivamente.

$$F \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 6 \end{pmatrix} \Rightarrow \begin{cases} x - z = 1 \\ x + 2y + z = 4 \\ x - y - z = 6 \end{cases} = \begin{pmatrix} \frac{15}{2} \\ -5 \\ \frac{13}{2} \end{pmatrix}$$

$$F \begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \\ 9 \end{pmatrix} \Rightarrow \begin{cases} x - z = 4 \\ x + 2y + z = 1 \\ x - y - z = 9 \end{cases} = \begin{pmatrix} \frac{15}{2} \\ -5 \\ \frac{7}{2} \end{pmatrix}$$

Otterremo dunque

$$M_{\phi, \psi}(F) = \begin{pmatrix} \frac{15}{2} & \frac{15}{2} \\ -5 & -5 \\ \frac{13}{2} & \frac{7}{2} \end{pmatrix}$$

Sia dunque ora

$$\begin{pmatrix} -5 \\ 5 \end{pmatrix} = 2 \begin{pmatrix} 2 \\ 1 \end{pmatrix} - 3 \begin{pmatrix} 3 \\ -1 \end{pmatrix} \Rightarrow \mathbf{x} = \begin{pmatrix} 2 \\ -3 \end{pmatrix}$$

Mediante applicazione il vettore iniziale si trasforma al seguente modo

$$F \begin{pmatrix} -5 \\ 5 \end{pmatrix} = \begin{pmatrix} -10 \\ 5 \\ -15 \end{pmatrix}$$

verifichiamo dunque prendendo la matrice associata all'applicazione e moltiplichiamola per il vettore delle coordinate del vettore rispetto alla base ψ :

$$M_{\phi, \psi}(F) \cdot \mathbf{x} = \begin{pmatrix} -\frac{15}{2} \\ 5 \\ \frac{5}{2} \end{pmatrix} \Rightarrow -\frac{15}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 5 \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} + \frac{5}{2} \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -10 \\ 5 \\ -15 \end{pmatrix}$$

Teorema 4.22. Siano \mathbf{V} e \mathbf{W} due \mathbb{K} -spazi vettoriali, e siano $\psi = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ e $\phi = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ basi di \mathbf{V} e di \mathbf{W} rispettivamente. L'applicazione

$$M_{\phi, \psi} : \text{hom}(\mathbf{V}, \mathbf{W}) \rightarrow M_{m, n}(\mathbb{K}), \quad F \mapsto M_{\phi, \psi}(F)$$

è un isomorfismo di \mathbb{K} -spazi vettoriali. In particolare

$$\text{hom}(\mathbf{V}, \mathbf{W}) = mn.$$

Proof. Siano $F, G \in \text{hom}(\mathbf{V}, \mathbf{W})$, $M = M_{\phi, \psi}(F)$, $N = M_{\phi, \psi}(G)$, $c \in \mathbb{K}$. Se $\mathbf{v} = \sum_{j=1}^n x_j \mathbf{v}_j$, denotiamo con $\mathbf{x} = (x_1, \dots, x_n)$ l' n -vettore colonna delle coordinate di \mathbf{v} rispetto alla base ψ . Gli m -vettori colonna delle coordinate di $F(\mathbf{v})$ e di $G(\mathbf{v})$ rispetto a ϕ sono rispettivamente $M\mathbf{x}$ ed $N\mathbf{x}$, mentre il vettore colonna delle coordinate di $(F+G)(\mathbf{v}) = F(\mathbf{v}) + G(\mathbf{v})$ è $M\mathbf{x} + N\mathbf{x} = (M+N)\mathbf{x}$; quindi

$$M_{\phi, \psi}(F+G) = M_{\phi, \psi}(F) + M_{\phi, \psi}(G).$$

Inoltre l' m -vettore colonna delle coordinate di $(cF)(\mathbf{v}) = cF(\mathbf{v})$ è $c(M\mathbf{x}) = (cM)\mathbf{x}$, e quindi

$$M_{\phi, \psi}(cF) = cM_{\phi, \psi}(F).$$

Pertanto $M_{\phi, \psi}$ è un'applicazione lineare. Sia $A \in M_{m, n}(\mathbb{K})$ e definiamo

$$F_A : \mathbf{V} \rightarrow \mathbf{W}, \quad \sum_{j=1}^n x_j \mathbf{v}_j \mapsto \sum_{i=1}^m (A^{(i)} \mathbf{x}) \mathbf{w}_i.$$

dove abbiamo denotato con $\mathbf{x} = (x_1, \dots, x_n)$ e $A^{(1)}, A^{(2)}, \dots, A^{(m)}$ le righe di A . In altre parole, il vettore colonna delle coordinate di $F_A(\mathbf{v})$ è $A\mathbf{x}$. Verifichiamo che F_A è un'applicazione lineare. Consideriamo due vettori qualsiasi $\mathbf{v} = \sum_{j=1}^n x_j \mathbf{v}_j$, $\mathbf{v}' = \sum_{j=1}^n x'_j \mathbf{v}_j \in \mathbf{V}$. Si ha

$$F_A(\mathbf{v} + \mathbf{v}') = \sum_{i=1}^m A^{(i)}(\mathbf{x} + \mathbf{x}') \mathbf{w}_i = \sum_{i=1}^m (A^{(i)} \mathbf{x} + A^{(i)} \mathbf{x}') \mathbf{w}_i = \sum_{i=1}^m (A^{(i)} \mathbf{x}) \mathbf{w}_i + \sum_{i=1}^m (A^{(i)} \mathbf{x}') \mathbf{w}_i = F_A(\mathbf{v}) + F_A(\mathbf{v}').$$

Se $c \in \mathbb{K}$ si ha inoltre

$$F_A(c\mathbf{v}) = \sum_{i=1}^m (A^{(i)} c\mathbf{x}) \mathbf{w}_i = \sum_{i=1}^m c(A^{(i)} \mathbf{x}) \mathbf{w}_i = c \left(\sum_{i=1}^m (A^{(i)} \mathbf{x}) \mathbf{w}_i \right) = cF_A(\mathbf{v}).$$

Quindi F_A è lineare. Per definizione si ha $M_{\phi, \psi}(F_A) = A$. D'altra parte, se $A = M_{\phi, \psi}(F)$ si ha evidentemente $F_A = F$. In conclusione, l'applicazione

$$M_{m,n}(\mathbb{K}) \rightarrow \text{hom}(\mathbf{V}, \mathbf{W}), \quad A \mapsto F_A.$$

è l'inversa di $M_{\phi, \psi}$. L'ultima affermazione del teorema segue da 2.26. \square

Applicazione lineare associata alla matrice A rispetto alle basi ψ e ϕ .

Definizione 4.12. L'applicazione F_A introdotta nel corso della dimostrazione del teorema 4.22 si dice **applicazione lineare associata alla matrice A rispetto alle basi ψ e ϕ** . Essa è definita, per ogni $A \in M_{m,n}(\mathbb{K})$, ponendo

$$F_A : M_{m,n}(\mathbb{K}) \rightarrow \text{hom}(\mathbf{V}, \mathbf{W}), \quad A \mapsto F_A \left(\sum_{j=1}^n x_j \mathbf{v}_j \right) = \sum_{i=1}^m y_i \mathbf{w}_i,$$

dove $\mathbf{y} = A\mathbf{x}$, essendo $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{y} = (y_1, \dots, y_m)$.

$$\begin{array}{ccc} M_{m,n}(\mathbb{K}) & \longrightarrow & \text{hom}(\mathbf{V}, \mathbf{W}) \\ A & \longmapsto & F_A \end{array} \quad \text{con} \quad \begin{array}{ccc} F_A : \mathbf{V} & \longrightarrow & \mathbf{W} \\ \mathbf{v} & \longmapsto & F_A(\mathbf{v}) = A\mathbf{x} = \mathbf{y} \end{array}$$

dove \mathbf{x} è il vettore delle coordinate di \mathbf{v} rispetto alla base ψ , e \mathbf{y} è il vettore delle coordinate di $F_A(\mathbf{v})$ rispetto alla base ϕ .

Nota 20. Si noti che $M_{\phi, \psi}$, fa corrispondere all'applicazione lineare $\mathbf{0}$ la matrice nulla $m \times n$. Dalla definizione segue inoltre che per ogni $\mathbf{v} \in \mathbf{V}$ il vettore $F_A(\mathbf{v})$ ha per coordinate polinomi omogenei di primo grado nelle coordinate di \mathbf{v} .

Proposizione 4.23. Siano $\mathbf{U}, \mathbf{V}, \mathbf{W}$ spazi vettoriali su \mathbb{K} di dimensioni s, n ed m , e siano $\alpha = \{\mathbf{u}_1, \dots, \mathbf{u}_s\}$, $\beta = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, $\gamma = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ le loro rispettive basi. Siano $G : \mathbf{U} \rightarrow \mathbf{V}$ ed $F : \mathbf{V} \rightarrow \mathbf{W}$ applicazioni lineari. Si ha

$$M_{\gamma, \alpha}(F \circ G) = M_{\gamma, \beta}(F) M_{\beta, \alpha}(G).$$

Proof. Se $\mathbf{u} = \sum_{k=1}^s z_k \mathbf{u}_k \in \mathbf{U}$, siano

$$G(\mathbf{u}) = \sum_{j=1}^n x_j \mathbf{v}_j, \quad F(G(\mathbf{u})) = (F \circ G)(\mathbf{u}) = \sum_{i=1}^m y_i \mathbf{w}_i.$$

Ponendo $\mathbf{z} = {}^t(z_1 \ z_2 \ \dots \ z_s)$, $\mathbf{x} = {}^t(x_1 \ x_2 \ \dots \ x_n)$, $\mathbf{y} = {}^t(y_1 \ y_2 \ \dots \ y_m)$, si ha

$$\mathbf{x} = M_{\beta, \alpha}(G) \mathbf{z}, \quad \mathbf{y} = M_{\gamma, \beta}(F) \mathbf{x} = M_{\gamma, \beta}(F) M_{\beta, \alpha}(G) \mathbf{z} = [M_{\gamma, \beta}(F) M_{\beta, \alpha}(G)] \mathbf{z}.$$

\square

Se $\mathbf{V} = \mathbf{W}$ e se $\psi = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ e $\phi = \{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ sono due basi di \mathbf{V} , ad ogni operatore lineare F su \mathbf{V} è associata una matrice quadrata $M_{\phi, \psi}(F) \in M_n(\mathbb{K})$.

Se $\psi = \phi$, dalla proposizione 4.23 segue che

$$F \in \text{GL}(\mathbf{V}) \iff M_{\psi, \psi}(F) \in \text{GL}_n(\mathbb{K}),$$

e in questo caso si ha

$$M_{\psi, \psi}(F^{-1}) = M_{\psi, \psi}(F)^{-1}.$$

Inoltre

$$M_{\psi, \psi}(F) = \mathbf{I}_n \iff F = \mathbf{1}_{\mathbf{V}}.$$

Proposizione 4.24. Sia \mathbf{V} un \mathbb{K} -spazio vettoriale con $\dim(\mathbf{V}) = n$. Sia

$$\psi = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$$

una base di \mathbf{V} . Per ogni $F \in \text{end}(\mathbf{V})$ indichiamo con $M_{\psi, \psi}(F) \in M_n(\mathbb{K})$ la matrice di F rispetto alla base ψ sia in dominio sia in codominio. Allora valgono:

1.

$$F \in \text{GL}(\mathbf{V}) \iff M_{\psi, \psi}(F) \in \text{GL}_n(\mathbb{K});$$

2.

$$F \in \text{GL}(\mathbf{V}) \implies M_{\psi, \psi}(F^{-1}) = [M_{\psi, \psi}(F)]^{-1};$$

3.

$$M_{\psi, \psi}(F) = \mathbf{I}_n \iff F = \mathbf{1}_{\mathbf{V}}.$$

Proof.

1. \implies Dall'ipotesi $F \in \text{GL}(\mathbf{V})$ segue che

$$\exists F^{-1} \in \text{GL}(\mathbf{V}) : F^{-1} \circ F = \mathbf{1}_{\mathbf{V}} \quad \wedge \quad F \circ F^{-1} = \mathbf{1}_{\mathbf{V}}.$$

Applicando la proposizione di composizione (con la stessa base ψ in ogni spazio), otteniamo

$$M_{\psi, \psi}(F^{-1} \circ F) = M_{\psi, \psi}(F^{-1}) M_{\psi, \psi}(F) = M_{\psi, \psi}(\mathbf{1}_{\mathbf{V}}) = \mathbf{I}_n.$$

Analogamente,

$$M_{\psi, \psi}(F \circ F^{-1}) = M_{\psi, \psi}(F) M_{\psi, \psi}(F^{-1}) = \mathbf{I}_n.$$

Ne segue che $M_{\psi, \psi}(F)$ ammette per inversa $M_{\psi, \psi}(F^{-1})$, quindi

$$M_{\psi, \psi}(F) \in \text{GL}_n(\mathbb{K}).$$

\Leftarrow Viceversa, supponiamo $M_{\psi, \psi}(F) \in \text{GL}_n(\mathbb{K})$. Denotiamo $A = M_{\psi, \psi}(F)$ e sia A^{-1} la matrice inversa di A . Definiamo un operatore lineare $G \in \text{GL}(\mathbf{V})$ in questo modo:

- Ad ogni vettore $\mathbf{v} \in \mathbf{V}$ facciamo corrispondere il vettore colonna $\mathbf{z} \in \mathbb{K}^n$ delle sue coordinate in base ψ .
- Poniamo $G(\mathbf{v})$ uguale al vettore di \mathbf{V} le cui coordinate in base ψ sono $A^{-1} \mathbf{z}$.

Per costruzione, $M_{\psi, \psi}(G) = A^{-1}$. Ora consideriamo $F \circ G$. Poiché

$$M_{\psi, \psi}(F \circ G) = M_{\psi, \psi}(F) \cdot M_{\psi, \psi}(G) = A A^{-1} = \mathbf{I}_n,$$

si ottiene $F \circ G = \mathbf{1}_{\mathbf{V}}$. Allo stesso modo $G \circ F = \mathbf{1}_{\mathbf{V}}$. Quindi $G = F^{-1}$ e F è invertibile.

Pertanto

$$F \in \text{GL}(\mathbf{V}) \iff M_{\psi, \psi}(F) \in \text{GL}_n(\mathbb{K}).$$

2. Se $F \in \text{GL}(\mathbf{V})$, poniamo $A = M_{\psi, \psi}(F)$. Dal punto precedente sappiamo che A è invertibile e che l'inversa A^{-1} corrisponde alla matrice di F^{-1} . Più precisamente:

$$F^{-1} \circ F = \mathbf{1}_{\mathbf{V}} \implies M_{\psi, \psi}(F^{-1} \circ F) = M_{\psi, \psi}(F^{-1}) M_{\psi, \psi}(F) = \mathbf{I}_n.$$

Analogamente $F \circ F^{-1} = \mathbf{1}_{\mathbf{V}}$ implica $M_{\psi, \psi}(F) M_{\psi, \psi}(F^{-1}) = \mathbf{I}_n$. Poiché l'inversa di A in $M_n(\mathbb{K})$ è unica, si conclude:

$$M_{\psi, \psi}(F^{-1}) = A^{-1} = [M_{\psi, \psi}(F)]^{-1}.$$

In simboli,

$$M_{\psi, \psi}(F^{-1}) = [M_{\psi, \psi}(F)]^{-1}.$$

3. \implies Se $F = \mathbf{1}_{\mathbf{V}}$, allora per ciascun $\mathbf{v}_j \in \psi$ vale $\mathbf{1}_{\mathbf{V}}(\mathbf{v}_j) = \mathbf{v}_j$. Le colonne della matrice $M_{\psi, \psi}(F)$ sono le coordinate di $\mathbf{1}_{\mathbf{V}}(\mathbf{v}_j)$ nella base ψ . Poiché $\mathbf{1}_{\mathbf{V}}(\mathbf{v}_j) = \mathbf{v}_j$, la colonna j -esima è il vettore colonna con uno in posizione j e zeri altrove. Pertanto la matrice è \mathbf{I}_n .

\Leftarrow Viceversa, se $M_{\psi, \psi}(F) = \mathbf{I}_n$, allora per ogni vettore con coordinate \mathbf{z} si ha

$$\text{coordinate di } F(\mathbf{v}) = \mathbf{I}_n \mathbf{z} = \mathbf{z}.$$

Ne segue $F(\mathbf{v}) = \mathbf{v}$ per ogni $\mathbf{v} \in \mathbf{V}$, cioè $F = \mathbf{1}_{\mathbf{V}}$.

Così dimostriamo che

$$M_{\psi, \psi}(F) = \mathbf{I}_n \iff F = \mathbf{1}_{\mathbf{V}}.$$

□

Matrice del cambiamento di coordinate

Definizione 4.13. Siano alternativamente $\psi \neq \phi$ basi distinte di \mathbf{V} e $F = \mathbf{1}_V$. In questo caso $M_{\phi,\psi}(\mathbf{1}_V)$ è detta **matrice del cambiamento di coordinate dalla base ψ alla base ϕ** . Per definizione la j -esima colonna di $M_{\phi,\psi}(\mathbf{1}_V)$ è costituita dalle coordinate di \mathbf{v}_j rispetto alla base ϕ , per ogni $j = 1, \dots, n$. Per ogni vettore $\mathbf{v} \in \mathbf{V}$ si ha

$$\mathbf{v} = \sum_{j=1}^n x_j \mathbf{v}_j = \sum_{j=1}^n y_j \mathbf{w}_j,$$

e, posto $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{y} = (y_1, \dots, y_n)$, si ha

$$\mathbf{y} = M_{\phi,\psi}(\mathbf{1}_V) \mathbf{x}.$$

Quindi la matrice $M_{\phi,\psi}(\mathbf{1}_V)$ permette di ottenere le coordinate \mathbf{y} di un vettore \mathbf{v} rispetto alla base ϕ una volta note le sue coordinate \mathbf{x} rispetto alla base ψ .

Esempio 4.5. Sia $\mathbf{V} = \mathbb{R}^2$. Consideriamo le due basi

$$\psi = \{\mathbf{v}_1, \mathbf{v}_2\} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}, \quad \phi = \{\mathbf{w}_1, \mathbf{w}_2\} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}.$$

Vogliamo costruire la matrice del cambiamento di coordinate

$$M_{\phi,\psi}(\mathbf{1}_V),$$

cioè la matrice 2×2 che esprime le coordinate dei vettori \mathbf{v}_j (della base ψ) rispetto alla base ϕ .

$$\begin{cases} x + y = 1 \\ x - y = 0 \end{cases} \implies \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

$$\begin{cases} x + y = 0 \\ x - y = 1 \end{cases} \implies \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$$

Da quanto ottenuto, la matrice del cambiamento di coordinate

$$M_{\phi,\psi}(\mathbf{1}_V) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

Ciò significa che, se un vettore $\mathbf{v} \in \mathbb{R}^2$ ha coordinate $\mathbf{x} = (x_1, x_2)$ rispetto a ψ , allora le sue coordinate $\mathbf{y} = (y_1, y_2)$ rispetto a ϕ si ottengono come

$$\mathbf{y} = M_{\phi,\psi}(\mathbf{1}_V) \mathbf{x} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Ad esempio, se $\mathbf{v} = (3, 2)$ ha $\mathbf{x} = (3, 2)$ in ψ , allora

$$\mathbf{y} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{5}{2} \\ \frac{1}{2} \end{pmatrix},$$

il che significa

$$(3, 2) = \frac{5}{2}(1, 1) + \frac{1}{2}(1, -1)$$

rispetto alla base ϕ .

Nota 21. Si noti che, per la proposizione 4.23:

$$M_{\psi,\phi}(\mathbf{1}_V) M_{\phi,\psi}(\mathbf{1}_V) = M_{\psi,\psi}(\mathbf{1}_V) = \mathbf{I}_n,$$

e quindi

$$M_{\psi,\phi}(\mathbf{1}_V) = M_{\phi,\psi}(\mathbf{1}_V)^{-1}. \quad (116)$$

Basi orientate

Definizione 4.14. Sia \mathbf{V} un \mathbb{R} -spazio vettoriale. Due basi $\psi = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ ed $\phi = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ di \mathbf{V} si dicono **orientate concordemente** se $\det(M_{\psi, \phi}(\mathbf{1}_V)) > 0$, e si scrive $e \sim f$. Altrimenti le due basi si dicono **orientate discordemente**.

Ogni base è orientata concordemente a sé stessa perché

$$M_{\psi, \psi}(\mathbf{1}_V) = \mathbf{I}_n \quad \therefore \quad \det(M_{\psi, \psi}(\mathbf{1}_V)) = 1 > 0.$$

Inoltre, poiché $M_{\psi, \phi}(\mathbf{1}_V) = M_{\phi, \psi}(\mathbf{1}_V)^{-1}$, il fatto che due basi siano orientate concordemente o discordemente non dipende dall'ordine in cui esse sono state prese. Infine, se ψ ed ϕ sono orientate concordemente, e così pure ϕ e $\varphi = \{\mathbf{g}_1, \dots, \mathbf{g}_n\}$, allora anche ψ e φ sono orientate concordemente. Infatti si ha:

$$\det(M_{\psi, \varphi}(\mathbf{1}_V)) = \det(M_{\psi, \phi}(\mathbf{1}_V)M_{\phi, \varphi}(\mathbf{1}_V)) = \det(M_{\psi, \phi}(\mathbf{1}_V))\det(M_{\phi, \varphi}(\mathbf{1}_V)) > 0.$$

Deduciamo che \sim è una relazione di equivalenza nell'insieme \mathcal{B} di tutte le basi di \mathbf{V} . Ogni classe di equivalenza si chiama **orientazione** di \mathbf{V} .

Quante sono le orientazioni di \mathbf{V} ? Certamente almeno due, perché se $\psi = \{\mathbf{e}_1, \dots, \mathbf{e}_n\} \in \mathcal{B}$, esiste qualche base orientata discordemente da ψ , ad esempio la base $\phi = \{-\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$. Infatti

$$M_{\psi, \phi}(\mathbf{1}_V) = \begin{pmatrix} -1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

D'altra parte, le orientazioni non sono più di due: infatti se esistessero tre basi ψ, ϕ, φ orientate discordemente a due a due, si avrebbe l'assurdo

$$0 > \det(M_{\psi, \varphi}(\mathbf{1}_V)) = \det(M_{\psi, \phi}(\mathbf{1}_V))\det(M_{\phi, \varphi}(\mathbf{1}_V)) > 0.$$

Quindi \mathbf{V} possiede due orientazioni, cioè \mathcal{B} è ripartito in due classi di equivalenza rispetto a \sim . L'orientazione cui una data base $\psi \in \mathcal{B}$ appartiene è l'**orientazione di \mathbf{V} definita da ψ** .

Se $\mathbf{V} = \mathbb{K}^n$, $\mathbf{W} = \mathbb{K}^m$ e ψ e ϕ sono le basi canoniche di \mathbb{K}^n e di \mathbb{K}^m rispettivamente, l'applicazione F_A associata a una matrice $A \in M_{m,n}(\mathbb{K})$ è data per ogni $\mathbf{x} \in \mathbb{K}^n$ da $F_A(\mathbf{x}) = A\mathbf{x}$, se gli elementi di \mathbb{K}^n e di \mathbb{K}^m sono visti come vettori colonna. Poiché le colonne di A sono i vettori $F_A(\mathbf{E}_1), \dots, F_A(\mathbf{E}_n)$, si ha

$$\text{Im}(F_A) = \langle F_A(\mathbf{E}_1), \dots, F_A(\mathbf{E}_n) \rangle.$$

In particolare $r(F_A) = r(A)$.

Si noti che le coordinate del vettore $A\mathbf{x}$ sono m polinomi omogenei di primo grado in x_1, x_2, \dots, x_n . Viceversa, una qualunque applicazione lineare $F: \mathbb{K}^n \rightarrow \mathbb{K}^m$ è della forma

$$F \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} F_1(x_1, x_2, \dots, x_n) \\ F_2(x_1, x_2, \dots, x_n) \\ \vdots \\ F_m(x_1, x_2, \dots, x_n) \end{pmatrix}$$

in cui ognuno degli $F_j(x_1, x_2, \dots, x_n)$ è un funzionale lineare su \mathbb{K}^n : infatti F_j è uguale alla composizione

$$\eta_j \circ F : \mathbb{K}^n \rightarrow \mathbb{K}^m \rightarrow \mathbb{K},$$

che è lineare. Ricordando la 111, vediamo che ognuno degli $F_j(x_1, x_2, \dots, x_n)$ è un polinomio omogeneo del primo grado in x_1, x_2, \dots, x_n .

Segue che ogni applicazione lineare $F \in \text{hom}(\mathbb{K}^n, \mathbb{K}^m)$ è definita da m polinomi omogenei di primo grado in x_1, \dots, x_n , e la corrispondente matrice ha per prima, seconda, ..., m -esima riga i coefficienti di

$$F_1(x_1, x_2, \dots, x_n), F_2(x_1, x_2, \dots, x_n), \dots, F_m(x_1, x_2, \dots, x_n).$$

Esempio 4.6. Alla matrice

$$A = \begin{pmatrix} 1 & 2 & \sqrt{2} \\ 3 & 1 & -\frac{1}{2} \end{pmatrix} \in M_{2,3}(\mathbb{R})$$

è associata l'applicazione lineare così definita:

$$F_A: \mathbb{R}^3 \rightarrow \mathbb{R}^2, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} x_1 + 2x_2 + \sqrt{2}x_3 \\ 3x_1 + x_2 - \frac{x_3}{2} \end{pmatrix}.$$

Esempio 4.7. All'applicazione lineare

$$F: \mathbb{R}^4 \rightarrow \mathbb{R}^3, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto \begin{pmatrix} 2x_1 - x_3 + x_4 \\ x_2 - \sqrt{3}x_3 + \frac{3}{4}x_4 \\ x_1 - x_2 + x_3 + 5x_4 \end{pmatrix}$$

è associata la matrice

$$\begin{pmatrix} 2 & 0 & -1 & 1 \\ 0 & 1 & -\sqrt{3} & \frac{3}{4} \\ 1 & -1 & 1 & 5 \end{pmatrix}.$$

Siano $A \in M_{m,n}(\mathbb{K})$, $\mathbf{b} = (b_1, \dots, b_m) \in \mathbb{K}^m$, e consideriamo il sistema di m equazioni lineari in n incognite

$$A\mathbf{X} = \mathbf{b} \quad (117)$$

dove $\mathbf{X} = (x_1, \dots, x_n)$. Un vettore $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{K}^n$ è una soluzione del sistema 117 se e solo se $F_A(\mathbf{x}) = \mathbf{b}$, dove $F_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$ è l'applicazione lineare associata ad A . Affinché un $\mathbf{x} \in \mathbb{K}^n$ siffatto esista è necessario e sufficiente che $\mathbf{b} \in \text{Im}(F_A)$. D'altra parte, poiché $\text{Im}(F_A)$ è generata dalle colonne di A , affinché $\mathbf{b} \in \text{Im}(F_A)$, cioè affinché il sistema 117 sia compatibile, è necessario e sufficiente che

$$r(A) = r(A\mathbf{b}).$$

In questo modo abbiamo ottenuto una nuova dimostrazione del teorema di Kronecker-Rouché-Capelli. Sappiamo che, se è compatibile, il sistema 117 possiede ∞^{n-r} soluzioni, dove $r = r(A)$. Ciò può essere dimostrato anche osservando che il sistema 117 possiede un'infinità di soluzioni uguali alla dimensione dello spazio delle soluzioni del sistema omogeneo associato

$$A\mathbf{x} = \mathbf{0},$$

e che tale spazio coincide con $N(F_A)$. Per il teorema 4.1 abbiamo

$$\dim[N(F_A)] = n - r(F) = n - r.$$

Sia \mathbf{V} uno spazio vettoriale reale di dimensione 3, e sia $\psi = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ una base di \mathbf{V} . Consideriamo le seguenti basi, i cui vettori assegnamo in coordinate rispetto a ψ :

$$\alpha = \left\{ \mathbf{a}_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{a}_2 \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \mathbf{a}_3 \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\}, \quad \beta = \left\{ \mathbf{b}_1 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \mathbf{b}_2 \begin{pmatrix} 1 \\ -2 \\ -3 \end{pmatrix}, \mathbf{b}_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

Si ha

$$M_{\psi,\alpha}(\mathbf{1}_V) = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & -2 \\ 0 & 1 & 1 \end{pmatrix} \quad M_{\alpha,\psi}(\mathbf{1}_V) = M_{\psi,\alpha}(\mathbf{1}_V)^{-1} = \begin{pmatrix} 3 & -2 & -4 \\ -1 & 1 & 2 \\ 1 & -1 & -1 \end{pmatrix}.$$

Similmente

$$M_{\psi,\beta}(\mathbf{1}_V) = \begin{pmatrix} -1 & 1 & 1 \\ 0 & -2 & 1 \\ 1 & -3 & 1 \end{pmatrix}, \quad M_{\beta,\psi}(\mathbf{1}_V) = M_{\psi,\beta}(\mathbf{1}_V)^{-1} = \begin{pmatrix} \frac{1}{2} & -2 & \frac{3}{2} \\ \frac{1}{2} & -1 & \frac{1}{2} \\ 1 & -1 & 1 \end{pmatrix}$$

Per calcolare $M_{\beta,\alpha}(\mathbf{1}_V)$ conviene utilizzare la proposizione 4.23. Si ha

$$M_{\beta,\alpha}(\mathbf{1}_V) = M_{\beta,\psi}(\mathbf{1}_V)M_{\psi,\alpha}(\mathbf{1}_V) = \begin{pmatrix} \frac{1}{2} & -2 & \frac{3}{2} \\ \frac{1}{2} & -1 & \frac{1}{2} \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & -2 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} -\frac{3}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{5}{2} \\ 0 & 2 & 3 \end{pmatrix}.$$

Infine, per calcolare $M_{\alpha,\beta}(\mathbf{1}_V)$ possiamo utilizzare l'identità

$$M_{\alpha,\beta}(\mathbf{1}_V) = M_{\beta,\alpha}(\mathbf{1}_V)^{-1}$$

e calcolare l'inversa di $M_{\beta,\alpha}(\mathbf{1}_V)$, oppure scrivere

$$M_{\alpha,\beta}(\mathbf{1}_V) = M_{\alpha,\psi}(\mathbf{1}_V)M_{\psi,\beta}(\mathbf{1}_V) = \begin{pmatrix} 3 & -2 & -4 \\ -1 & 1 & 2 \\ 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} -1 & 1 & 1 \\ 0 & -2 & 1 \\ 1 & -3 & 1 \end{pmatrix} = \begin{pmatrix} -7 & 19 & -3 \\ 3 & -9 & 2 \\ -2 & 6 & -1 \end{pmatrix}.$$

Siano V e W \mathbb{R} -spazi vettoriali con $\dim(V) = 4$, $\dim(W) = 2$,

$$\psi = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}, \text{ e } \phi = \{\mathbf{w}_1, \mathbf{w}_2\}$$

basi di V e W rispettivamente, e $F: V \rightarrow W$ l'applicazione lineare tale che

$$M_{\phi,\psi}(F) = \begin{pmatrix} 1 & 3 & -2 & \frac{1}{2} \\ 2 & 0 & 1 & 0 \end{pmatrix}.$$

Siano

$$\alpha = \left\{ \mathbf{a}_1 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2 \end{pmatrix}, \mathbf{a}_2 \begin{pmatrix} 2 \\ -1 \\ 3 \\ 0 \end{pmatrix}, \mathbf{a}_3 \begin{pmatrix} \sqrt{2} \\ 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{a}_4 \begin{pmatrix} 1 \\ -\frac{1}{2} \\ 1 \\ 5 \end{pmatrix} \right\}, \quad \beta = \left\{ \mathbf{b}_1 \begin{pmatrix} 2 \\ -3 \end{pmatrix}, \mathbf{b}_2 \begin{pmatrix} -1 \\ 3 \end{pmatrix} \right\}$$

nuove basi di V e W rispettivamente, assegnate mediante le loro coordinate nelle basi ψ e ϕ . Per la proposizione 4.23 si ha

$$M_{\beta,\alpha}(F) = M_{\beta,\phi}(\mathbf{1}_W)M_{\phi,\psi}(F)M_{\psi,\alpha}(\mathbf{1}_V).$$

Poiché

$$M_{\phi,\beta}(\mathbf{1}_W) = \begin{pmatrix} 2 & -1 \\ -3 & 3 \end{pmatrix},$$

avremo

$$M_{\beta,\phi}(\mathbf{1}_W) = M_{\phi,\beta}(\mathbf{1}_W)^{-1} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

e dunque

$$M_{\psi,\alpha}(\mathbf{1}_V) = \begin{pmatrix} 1 & 2 & \sqrt{2} & 1 \\ 1 & -1 & 1 & -\frac{1}{2} \\ 1 & 3 & 0 & 1 \\ 2 & 0 & 0 & 5 \end{pmatrix}$$

Si ottiene:

$$\begin{aligned} M_{\beta,\alpha}(F) &= \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} 1 & 3 & -2 & 1 \\ 2 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & \sqrt{2} & 1 \\ 1 & -1 & 1 & -\frac{1}{2} \\ 1 & 3 & 0 & 1 \\ 2 & 0 & 0 & 5 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 & -\frac{1}{3} & \frac{1}{6} \\ 1 & -1 & \frac{4}{3} & -\frac{1}{6} \end{pmatrix} \begin{pmatrix} 1 & 2 & \sqrt{2} & 1 \\ 1 & -1 & 1 & -\frac{1}{2} \\ 1 & 3 & 0 & 1 \\ 2 & 0 & 0 & 5 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 0 & \sqrt{2} + 1 & 1 \\ 1 & 7 & \sqrt{2} - 1 & 2 \end{pmatrix} \end{aligned}$$

Nello spazio vettoriale $M_2(\mathbb{C})$, consideriamo le basi:

$$\psi = \left\{ \mathbf{1}_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{1}_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{1}_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{1}_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

$$\phi = \left\{ \mathbf{I}_2, \quad \Sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \Sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \Sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\}$$

dove Σ_1 , Σ_2 e Σ_3 sono le matrici di Pauli . Poiché

$$\mathbf{I}_2 = \mathbf{1}_{11} + \mathbf{1}_{22}, \quad \Sigma_1 = \mathbf{1}_{12} + \mathbf{1}_{21}, \quad \Sigma_2 = -i\mathbf{1}_{12} + i\mathbf{1}_{21}, \quad \Sigma_3 = \mathbf{1}_{11} - \mathbf{1}_{22},$$

si ha:

$$M_{\psi,\phi}(\mathbf{1}) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & -i & 0 \\ 0 & 1 & i & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}, \quad M_{\phi,\psi}(\mathbf{1}) = M_{\psi,\phi}(\mathbf{1})^{-1} = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & -\frac{i}{2} & \frac{i}{2} & 0 \\ \frac{1}{2} & 0 & 0 & -\frac{1}{2} \end{pmatrix}.$$

In uno spazio affine \mathbf{A} con spazio vettoriale associato \mathbf{V} consideriamo due riferimenti affini,

$$\psi_{\mathbf{A}} = E\mathbf{e}_1, \dots, \mathbf{e}_n, \quad \phi_{\mathbf{A}} = F\mathbf{f}_1, \dots, \mathbf{f}_n$$

e un punto qualsiasi $P \in \mathbf{A}$. Sia $\mathbf{x} = (x_1, \dots, x_n)$ il vettore colonna delle coordinate di P nel riferimento $\psi_{\mathbf{A}}$ e sia $\mathbf{y} = (y_1, \dots, y_n)$ quello delle coordinate di $P \in \mathbf{A}$ rispetto a $\phi_{\mathbf{A}}$. Si ha

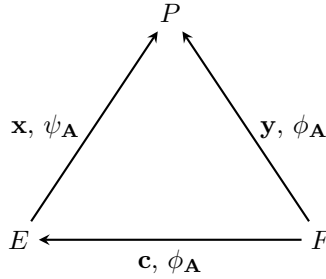
$$\overrightarrow{EP} = \sum_{j=1}^n x_j \mathbf{e}_j \quad \overrightarrow{FP} = \sum_{j=1}^n y_j \mathbf{f}_j.$$

Supponiamo di conoscere \mathbf{x} e di voler trovare \mathbf{y} . Denotiamo con $\psi = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ e $\phi = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ le due basi di \mathbf{V} , con $A = (a_{ij}) = M_{\phi,\psi}(\mathbf{1}_{\mathbf{V}})$ e con $\mathbf{c} = (c_1, \dots, c_n)$ il vettore delle coordinate di E rispetto a $F\mathbf{f}_1, \dots, \mathbf{f}_n$. L'identità vettoriale

$$\overrightarrow{FP} = \overrightarrow{FE} + \overrightarrow{EP},$$

espressa rispetto alla base ϕ , ci dà

$$\mathbf{y} = A\mathbf{x} + \mathbf{c}. \quad (118)$$



La 118 è la formula del **cambiamento di coordinate affini dal riferimento $E\mathbf{e}_1, \dots, \mathbf{e}_n$ al riferimento $F\mathbf{f}_1, \dots, \mathbf{f}_n$** .

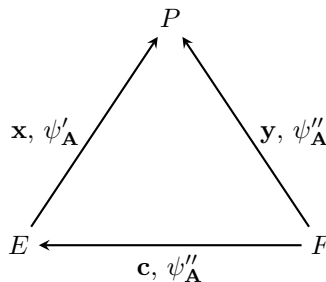
La 118 dipende solo dalle basi ψ ed ϕ e dai punti E ed F , le origini nei due riferimenti affini assegnati.

Nel caso in cui lo spazio \mathbf{A} è uno spazio affine reale, due riferimenti affini $E\mathbf{e}_1, \dots, \mathbf{e}_n$ e $F\mathbf{f}_1, \dots, \mathbf{f}_n$ si diranno **orientati concordemente** (**orientati discordemente**) se le basi ϕ ed ψ sono orientate concordemente (orientate discordemente).

Consideriamo l'insieme \mathcal{R} di tutti i riferimenti affini di \mathbf{A} , e in \mathcal{R} diciamo equivalenti due riferimenti se sono orientati concordemente. Procedendo in modo simile al caso vettoriale, possiamo verificare che quella che abbiamo definito è effettivamente una relazione di equivalenza, e che le classi di equivalenza sono due. Esse si dicono le **orientazioni di \mathbf{A}** . L'orientazione cui un dato riferimento $E\mathbf{e}_1, \dots, \mathbf{e}_n \in \mathcal{R}$ appartiene è l'**orientazione di \mathbf{A} definita da $E\mathbf{e}_1, \dots, \mathbf{e}_n$** .

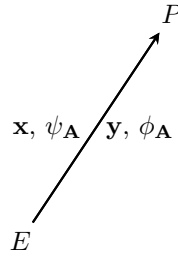
Se $\psi'_{\mathbf{A}} = E\mathbf{e}_1, \dots, \mathbf{e}_n$ è un riferimento affine dello spazio affine \mathbf{A} e se un secondo riferimento affine assegnato è $\psi''_{\mathbf{A}} = F\mathbf{e}_1, \dots, \mathbf{e}_n$, cioè è ottenuto solo cambiando la posizione dell'origine e lasciando invariata la base dei vettori, la formula 118 si riduce alla seguente:

$$\mathbf{y} = \mathbf{x} + \mathbf{c}. \quad (119)$$



Se invece il secondo riferimento è $E\mathbf{f}_1, \dots, \mathbf{f}_n$, cioè è ottenuto cambiando la base dei vettori, ma non l'origine, la formula 118 diventa

$$\mathbf{y} = A\mathbf{x} \quad (120)$$



Proposizione 4.25. *Ogni cambiamento di riferimento affine si può ottenere come la composizione di uno del tipo 119 seguito da uno del tipo 120, o viceversa.*

Proof. Da fare. □

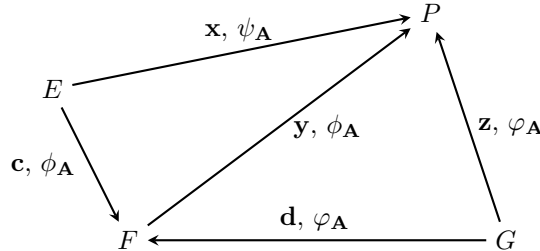
Siano

$$\psi_{\mathbf{A}} = E\mathbf{e}_1, \dots, \mathbf{e}_n, \quad \phi_{\mathbf{A}} = F\mathbf{f}_1, \dots, \mathbf{f}_n, \quad \varphi_{\mathbf{A}} = G\mathbf{g}_1, \dots, \mathbf{g}_n$$

tre riferimenti affini in \mathbf{A} , e siano

$$\mathbf{x} = (x_1, \dots, x_n), \quad \mathbf{y} = (y_1, \dots, y_n), \quad \mathbf{z} = (z_1, \dots, z_n)$$

i vettori delle coordinate di un punto $P \in \mathbf{A}$ rispetto ad ognuno dei riferimenti dati.



Siano $A = M_{\phi, \psi}(\mathbf{1}_V)$, $B = M_{\varphi, \phi}(\mathbf{1}_V)$, e

$$\overrightarrow{EP} = \sum_{j=1}^n x_j \mathbf{e}_j, \quad \overrightarrow{FP} = \sum_{j=1}^n y_j \mathbf{f}_j, \quad \overrightarrow{GP} = \sum_{j=1}^n z_j \mathbf{g}_j$$

Supponiamo che il passaggio dalle coordinate \mathbf{x} alle \mathbf{y} sia dato da

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

e che quello dalle \mathbf{y} alle \mathbf{z} sia dato da

$$\mathbf{z} = B\mathbf{y} + \mathbf{d}.$$

Allora la formula che esprime il passaggio dalla \mathbf{x} alle \mathbf{z} si ottiene al seguente modo mediante sostituzione:

$$\mathbf{z} = B\mathbf{y} + \mathbf{d} = B(A\mathbf{x} + \mathbf{c}) + \mathbf{d} = B A \mathbf{x} + (B\mathbf{c} + \mathbf{d}).$$

Con le notazioni della 118 la formula che esprime il passaggio dalle coordinate \mathbf{y} alle \mathbf{x} , cioè il passaggio inverso di quello dato dalla 118, è

$$\mathbf{x} = A^{-1}\mathbf{y} - A^{-1}\mathbf{c}. \quad (121)$$

Infatti, sostituendo la 118 nel secondo membro della 121 si ottiene l'identità $\mathbf{x} = \mathbf{x}$.

Cambiamento di coordinate affini Sia \mathbf{A} un \mathbb{R} -spazio affine di dimensione 3, e sia \mathbf{V} lo spazio vettoriale associato. Siano

$$\psi_{\mathbf{A}} = E\mathbf{e}_1\mathbf{e}_2\mathbf{e}_3, \quad \text{e} \quad \phi_{\mathbf{A}} = F\mathbf{f}_1\mathbf{f}_2\mathbf{f}_3$$

due riferimenti affini in \mathbf{A} . Supponiamo che

$$\mathbf{f}_1 = \mathbf{e}_1 + \mathbf{e}_2, \quad \mathbf{f}_2 = -\mathbf{e}_1 + \mathbf{e}_3, \quad \mathbf{f}_3 = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3,$$

e che F abbia coordinate $(5, -\frac{3}{2}, \frac{1}{2})$ nel riferimento $\psi_{\mathbf{A}}$. Per conoscere la formula di cambiamento di coordinate dal riferimento $\psi_{\mathbf{A}}$ al riferimento $\phi_{\mathbf{A}}$ occorre conoscere la matrice $M_{\phi,\psi}(\mathbf{1}_{\mathbf{V}})$, nonché le coordinate c_1, c_2, c_3 di E rispetto a $\phi_{\mathbf{A}}$. Si ha

$$M_{\psi,\phi}(\mathbf{1}_{\mathbf{V}}) = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

e quindi

$$M_{\phi,\psi}(\mathbf{1}_{\mathbf{V}}) = M_{\psi,\phi}(\mathbf{1}_{\mathbf{V}})^{-1} = \begin{pmatrix} -1 & 2 & -1 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}.$$

Per determinare (c_1, c_2, c_3) utilizziamo la matrice $M_{\phi,\psi}(\mathbf{1}_{\mathbf{V}})$ che abbiamo appena calcolato. Si ha

$$c_1\mathbf{f}_1 + c_2\mathbf{f}_2 + c_3\mathbf{f}_3 = \overrightarrow{FE} = -\overrightarrow{EF} = -\left(5\mathbf{e}_1 - \frac{3}{2}\mathbf{e}_2 + \frac{1}{2}\mathbf{e}_3\right).$$

Pertanto

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = M_{\phi,\psi}(\mathbf{1}_{\mathbf{V}}) \begin{pmatrix} -5 \\ \frac{3}{2} \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} -1 & 2 & -1 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} -5 \\ \frac{3}{2} \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{17}{2} \\ \frac{13}{2} \\ -7 \end{pmatrix}.$$

In conclusione, la formula del cambiamento di coordinate dal riferimento $\psi_{\mathbf{A}}$ al riferimento $\phi_{\mathbf{A}}$ è

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -1 & 2 & -1 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} \frac{17}{2} \\ \frac{13}{2} \\ -7 \end{pmatrix}$$

ovvero:

$$y_1 = -x_1 + 2x_2 - x_3 + \frac{17}{2} \quad y_2 = -x_1 + x_2 + \frac{13}{2} \quad y_3 = x_1 - x_2 + x_3 - 7.$$

Proposizione 4.26. Siano $\psi = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$, $\phi = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ due basi del \mathbb{K} -spazio vettoriale \mathbf{V} , e siano

$$\eta = \{\eta_1, \dots, \eta_n\}, \quad \beta = \{\beta_1, \dots, \beta_n\}$$

le basi di \mathbf{V}^* duali di ψ e di ϕ rispettivamente. Allora

$$M_{\beta,\eta}(\mathbf{1}_{\mathbf{V}^*}) = [M_{\psi,\phi}(\mathbf{1}_{\mathbf{V}})]^t.$$

Proof. Da fare. □

4.3 Operatori Lineari

Matrice di un endomorfismo rispetto a una base

Definizione 4.15. Siano \mathbf{V} un \mathbb{K} -spazio vettoriale di dimensione finita, e $\psi = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una base di \mathbf{V} . Per ogni operatore $F \in \text{end}(\mathbf{V})$, scriveremo $M_\psi(F)$ invece di $M_{\psi,\psi}(F)$, e chiameremo $M_\psi(F)$ la **matrice di F rispetto alla base ψ** .

Dal teorema 117 segue che l'applicazione

$$M_\psi : \text{end}(\mathbf{V}) \rightarrow M_n(\mathbb{K}), \quad F \mapsto M_\psi(F)$$

è un isomorfismo di \mathbb{K} -spazi vettoriali. Si ha $M_\psi(\mathbf{1}_\mathbf{V}) = \mathbf{I}_n$ e

$$M_\psi(F) \in \text{GL}_n(\mathbb{K}) \iff F \in \text{GL}(\mathbf{V}),$$

cioè un operatore F è un automorfismo se e solo se $M_\psi(F)$ è invertibile. Quindi l'applicazione M_ψ induce una biezione che denotiamo con lo stesso simbolo:

$$M_\psi : \text{GL}(\mathbf{V}) \rightarrow \text{GL}_n(\mathbb{K}).$$

Sia $\phi = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ un'altra base di \mathbf{V} . Dalla proposizione 4.23 deduciamo che

$$M_\phi(F) = M_{\phi,\psi}(\mathbf{1}_\mathbf{V}) M_\psi(F) M_{\psi,\phi}(\mathbf{1}_\mathbf{V}).$$

Poiché $M_{\phi,\psi}(\mathbf{1}_\mathbf{V}) = M_{\psi,\phi}(\mathbf{1}_\mathbf{V})^{-1}$, otteniamo

$$M_\phi(F) = M_{\psi,\phi}(\mathbf{1}_\mathbf{V})^{-1} M_\psi(F) M_{\psi,\phi}(\mathbf{1}_\mathbf{V}) \quad (122)$$

da cui segue immediatamente che

$$\det(M_\phi(F)) = \det(M_\psi(F)),$$

cioè $\det(M_\psi(F))$ non dipende dalla base ψ , ma solo da F .

Determinante dell'operatore

Definizione 4.16. Chiameremo pertanto $\det(M_\psi(F))$ il **determinante dell'operatore F** e lo denoteremo con $\det(F)$, senza dover specificare la matrice $M_\psi(F)$ attraverso la quale è stato calcolato.

Similitudine tra matrici

Definizione 4.17. Due matrici $A, B \in M_n(\mathbb{K})$ si dicono **simili** se:

$$\exists M \in \text{GL}_n(\mathbb{K}) : B = M^{-1}AM.$$

Proposizione 4.27. La similitudine è una relazione di equivalenza in $M_n(\mathbb{K})$.

Proof.

1. *Riflessività:*

$$\forall A \in M_n(\mathbb{K}), \quad A = \mathbf{I}_n^{-1}A\mathbf{I}_n;$$

2. *Simmetria*

$$\forall A, B \in M_n(\mathbb{K}), \quad B = M^{-1}AM \implies A = (MM^{-1})A(MM^{-1}) = M(M^{-1}AM)M^{-1} = MBM^{-1};$$

3. *Transitività*

$$B = M^{-1}AM \wedge C = N^{-1}BN \implies C = N^{-1}(M^{-1}AM)N = (MN)^{-1}A(MN).$$

□

Proposizione 4.28. *Siabo \mathbf{V} un \mathbb{K} -spazio vettoriale, $\dim(\mathbf{V}) = n$, e $A, B \in M_n(\mathbb{K})$. A e B sono simili se e solo se esistono un operatore lineare $F \in \text{end}(\mathbf{V})$ e basi ψ ed ϕ di \mathbf{V} tali che $M_\psi(F) = A$ ed $M_\phi(F) = B$.*

Proof. Se F, ψ, ϕ esistono, allora dalla 122 segue che A e B sono simili. Supponiamo viceversa che

$$B = M^{-1}AM. \quad (123)$$

Sia ψ una base arbitraria di \mathbf{V} e sia $F = F_A$ l'operatore associato alla matrice A rispetto alla base ψ . Per ogni $j \in \{1, \dots, n\}$ sia \mathbf{f}_j il vettore le cui coordinate rispetto a ψ sono gli elementi della j -esima colonna di M , cioè

$$\mathbf{f}_j = \sum_{i=1}^n m_{ij} \mathbf{e}_i.$$

Poiché M ha rango n , i vettori $\mathbf{f}_1, \dots, \mathbf{f}_n$ sono linearmente indipendenti e quindi $\{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ è una base di \mathbf{V} . Si ha inoltre

$$M = M_{\psi, \phi}(\mathbf{1}_{\mathbf{V}}).$$

Dalla 123 discende che $B = M_\phi(F)$. □

Operatore diagonalizzabile e base diagonalizzante

Definizione 4.18. Sia \mathbf{V} un \mathbb{K} -spazio vettoriale, $\dim(\mathbf{V}) = n$. Un operatore $F \in \text{end}(\mathbf{V})$ si dice **diagonalizzabile** se esiste una base ψ di \mathbf{V} tale che $M_\psi(F)$ sia una matrice diagonale, cioè della forma

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

per opportuni $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{K}$. Se ciò avviene ψ è una **base diagonalizzante** per F .

Matrice diagonalizzabile

Definizione 4.19. Una matrice $A \in M_n(\mathbb{K})$ si dice **diagonalizzabile** se è simile a una matrice diagonale.

Ovviamente, se $F \in \text{end}(\mathbf{V})$ e ψ è una base di \mathbf{V} , F è diagonalizzabile se e solo se $M_\psi(F)$ è una matrice diagonalizzabile. In particolare $A \in M_n(\mathbb{K})$ è diagonalizzabile se e solo se l'operatore $F_A \in \text{end}(\mathbb{K}^n)$ definito da A è diagonalizzabile. Se $F \in \text{end}(\mathbf{V})$ è diagonalizzabile e ψ è una base diagonalizzante per F , si ha

$$F(\mathbf{e}_i) = \lambda_i \mathbf{e}_i, \quad i \in \{1, \dots, n\}. \quad (124)$$

Viceversa, se una base ψ soddisfacente la 124 esiste, allora la matrice $M_\psi(F)$ è diagonale, e quindi F è diagonalizzabile e ψ è una base diagonalizzante. Si noti che, se $\dim(\mathbf{V}) = 1$, allora ogni $F \in \text{end}(\mathbf{V})$ è diagonalizzabile e ogni base di \mathbf{V} è diagonalizzante per F . Se $\dim(\mathbf{V}) \geq 2$ non tutti gli operatori $F \in \text{end}(\mathbf{V})$ sono diagonalizzabili. Similmente non tutte le matrici $A \in M_n(\mathbb{K})$ sono diagonalizzabili se $n \geq 2$.

In relazione al problema di stabilire l'esistenza di basi diagonalizzanti si presentano in modo naturale le nozioni di **autovettore** e di **autovalore**.

Autovettore, autovalore, e spettro di un operatore

Definizione 4.20. Siano \mathbf{V} un \mathbb{K} -spazio vettoriale, e $F \in \text{end}(\mathbf{V})$. Un vettore $\mathbf{v} \in \mathbf{V}$ si dice **autovettore** di F se

$$\mathbf{v} \neq \mathbf{0} \wedge \exists \lambda \in \mathbb{K} : F(\mathbf{v}) = \lambda \mathbf{v}.$$

λ è l'**autovalore** di F relativo all'autovettore \mathbf{v} . Il sottoinsieme di \mathbb{K} costituito dagli autovalori di F è lo **spettro** di F .

Se $A \in M_n(\mathbb{K})$, un autovettore di A è un autovettore $\mathbf{x} \in \mathbb{K}^n$ dell'operatore $F_A \in \text{end}(\mathbb{K}^n)$ definito da A , e un autovalore di A è un autovalore di F_A .

Esempio 4.8. Se $F = \mathbf{1}_V$, ogni $\mathbf{v} \neq \mathbf{0}$ è un autovettore di F con autovalore $\lambda = 1$.

Esempio 4.9. Se F è un operatore tale che $\ker(F) \neq \langle \mathbf{0} \rangle$, ogni $\mathbf{v} \in \ker(F) \setminus \langle \mathbf{0} \rangle$ è un autovettore di F con autovalore $\lambda = 0$.

Diamo qui di seguito alcune semplici proprietà degli autovettori e degli autovalori di un operatore $F \in \text{end}(V)$. Supponiamo $\dim(V) = n \geq 1$.

Proposizione 4.29. *L'autovalore relativo a un autovettore \mathbf{v} è univocamente determinato.*

Proof. Se $\lambda \mathbf{v} = F(\mathbf{v}) = \mu \mathbf{v}$ per qualche $\lambda, \mu \in \mathbb{K}$, allora $(\lambda - \mu)\mathbf{v} = \mathbf{0}$, poiché $\mathbf{v} \neq \mathbf{0}$, dev'essere $\lambda - \mu = 0$, cioè $\lambda = \mu$. \square

Proposizione 4.30. *Se $\mathbf{v}_1, \mathbf{v}_2 \in V$ sono autovettori relativi allo stesso autovalore λ , allora per ogni $c_1, c_2 \in \mathbb{K}$ il vettore $c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2$, se non è il vettore nullo, è ancora un autovettore con autovalore λ .*

Proof. Si ha

$$F(c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2) = c_1 F(\mathbf{v}_1) + c_2 F(\mathbf{v}_2) = c_1 \lambda \mathbf{v}_1 + c_2 \lambda \mathbf{v}_2 = \lambda(c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2).$$

\square

Autospazio relativo a un'autovalore

Definizione 4.21. Dalla 4.30 segue che l'insieme

$$V_\lambda(F) = \{\mathbf{v} \in V : \mathbf{v} \text{ è un autovettore di } F \text{ con autovalore } \lambda\} \cup \{\mathbf{0}\}$$

è un sottospazio vettoriale di V , detto **l'autospazio relativo all'autovalore λ** . Per una matrice $A \in M_n(\mathbb{K})$ si definisce **l'autospazio $V_\lambda(A)$ relativo all'autovalore λ** come il sottospazio di \mathbb{K}^n

$$V_\lambda(A) = V_\lambda(F_A).$$

Proposizione 4.31. *Se $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ sono autovettori relativi agli autovalori $\lambda_1, \dots, \lambda_k$ e se questi λ_i con $i \in \{1, \dots, k\}$ sono a due a due distinti, allora $\mathbf{v}_1, \dots, \mathbf{v}_k$ sono linearmente indipendenti.*

Proof. L'asserzione è banalmente vera se $k = 1$, perché $\mathbf{v}_1 \neq \mathbf{0}$. Procediamo per induzione su k , e supponiamo $k \geq 2$. Se

$$\sum_{i=1}^k c_i \mathbf{v}_i = \mathbf{0}, \tag{125}$$

allora, applicando F a entrambi i membri, si ha anche

$$\sum_{i=1}^k c_i F(\mathbf{v}_i) = \mathbf{0},$$

cioè

$$\sum_{i=1}^k c_i \lambda_i \mathbf{v}_i = \mathbf{0}. \tag{126}$$

D'altra parte, moltiplicando ambo i membri della 125 per λ_1 , si ottiene

$$\lambda_1 \left(\sum_{i=1}^k c_i \mathbf{v}_i \right) = \sum_{i=1}^k \lambda_1 c_i \mathbf{v}_i = 0 \quad (127)$$

e sottraendo la 127 dalla 126:

$$\sum_{i=1}^k c_i \lambda_i \mathbf{v}_i - \sum_{i=1}^k \lambda_1 c_i \mathbf{v}_i = \sum_{i=1}^k (c_i \lambda_i \mathbf{v}_i - \lambda_1 c_i \mathbf{v}_i) = \sum_{i=1}^k c_i (\lambda_i \mathbf{v}_i - \lambda_1 \mathbf{v}_i) = \sum_{i=1}^k c_i (\lambda_i - \lambda_1) \mathbf{v}_i = 0. \quad (128)$$

Per l'ipotesi induttiva $\mathbf{v}_2, \dots, \mathbf{v}_k$ sono linearmente indipendenti, e quindi i coefficienti del primo membro della 128 sono tutti uguali a 0. Poiché $\lambda_j - \lambda_1 \neq 0$ per ogni $j \in \{2, \dots, k\}$, si deduce che $c_2 = \dots = c_k = 0$. Quindi la 125 si riduce a $c_1 \mathbf{v}_1 = 0$, e quest'identità implica che anche $c_1 = 0$, perché $\mathbf{v}_1 \neq 0$. \square

Proposizione 4.32. *Se ogni $\mathbf{v} \in \mathbf{V} \setminus \{\mathbf{0}\}$ è un autovettore di F , allora*

$$\exists \lambda \in \mathbb{K} : F = \lambda \mathbf{1}_{\mathbf{V}}.$$

Proof. Se $\dim(\mathbf{V}) = 1$ l'asserzione è ovvia. Possiamo quindi supporre $\dim(\mathbf{V}) \geq 2$. Sia $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una base di \mathbf{V} . Dall'ipotesi segue che esistono $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{K}$ tali che $F(\mathbf{e}_i) = \lambda_i \mathbf{e}_i$, per $i \in \{1, \dots, n\}$. Siano $1 \leq i, j \leq n$ due indici distinti, e sia $\mathbf{v}_{ij} = \mathbf{e}_i + \mathbf{e}_j$. Per l'ipotesi esiste $\lambda_{ij} \in \mathbb{K}$ tale che

$$F(\mathbf{v}_{ij}) = \lambda_{ij} \mathbf{v}_{ij} = \lambda_{ij} \mathbf{e}_i + \lambda_{ij} \mathbf{e}_j.$$

D'altra parte si ha

$$F(\mathbf{v}_{ij}) = F(\mathbf{e}_i + \mathbf{e}_j) = F(\mathbf{e}_i) + F(\mathbf{e}_j) = \lambda_i \mathbf{e}_i + \lambda_j \mathbf{e}_j,$$

e, per l'indipendenza lineare di \mathbf{e}_i ed \mathbf{e}_j , deduciamo che $\lambda_i = \lambda_j = \lambda_{ij}$. In conclusione $\lambda_1 = \lambda_2 = \dots = \lambda_n$, e l'asserto è provato. \square

In pratica nella ricerca degli autovalori di un operatore o di una matrice si utilizza il cosiddetto **polinomio caratteristico**. La sua definizione si avvale del seguente semplice risultato.

Proposizione 4.33. *Sia \mathbf{V} uno spazio vettoriale di dimensione finita e sia $F \in \text{end}(\mathbf{V})$. Uno scalare $\lambda \in \mathbb{K}$ è un autovalore di F se e solo se l'operatore*

$$F - \lambda \mathbf{1}_{\mathbf{V}} : \mathbf{V} \rightarrow \mathbf{V}, \quad \mathbf{v} \mapsto F(\mathbf{v}) - \lambda \mathbf{v},$$

non è un isomorfismo, o, equivalentemente, se e solo se

$$\det(F - \lambda \mathbf{1}_{\mathbf{V}}) = 0.$$

Proof. $F - \lambda \mathbf{1}_{\mathbf{V}}$ non è un isomorfismo se e solo se $N(F - \lambda \mathbf{1}_{\mathbf{V}}) \neq \langle \mathbf{0} \rangle$, cioè se

$$\exists \mathbf{v} \in \mathbf{V} \setminus \{\mathbf{0}\} : (F - \lambda \mathbf{1}_{\mathbf{V}})(\mathbf{v}) = \mathbf{0},$$

ovvero tale che

$$F(\mathbf{v}) = \lambda \mathbf{v}. \quad (129)$$

La 129 afferma che F possiede l'autovettore \mathbf{v} con autovalore λ . Sia $\psi = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una base di \mathbf{V} . La matrice associata all'operatore $\lambda \mathbf{1}_{\mathbf{V}}$ è

$$\begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{pmatrix}$$

e, se $A = (a_{ij}) = M_{\psi}(F)$, allora

$$M_{\psi}(F - \lambda \mathbf{1}_{\mathbf{V}}) = \begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{pmatrix}$$

\square

Polinomio caratteristico

Definizione 4.22. Sia $A \in M_n(\mathbb{K})$ e sia T un'indeterminata. Il determinante

$$P_A(T) = |A - T\mathbf{I}_n| = \begin{vmatrix} a_{11} - T & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - T & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - T \end{vmatrix}$$

è un polinomio di grado n in T , detto **polinomio caratteristico** di A . Se $F \in \text{end}(\mathbf{V})$, $\psi = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ è una base di \mathbf{V} e $A = M_\psi(F)$, allora $P_A(T)$ è il **polinomio caratteristico** di F , e si denota con $P_F(T)$

La definizione di $P_F(T)$ è indipendente dalla base ψ , come vediamo dalla seguente proposizione:

Proposizione 4.34. Due matrici simili, come sono quelle che rappresentano F in due basi diverse, hanno lo stesso polinomio caratteristico.

Proof. Siano A e B due matrici simili, cioè si abbia $B = M^{-1}AM$, per qualche $M \in \text{GL}_n(\mathbb{K})$. Allora

$$B - T\mathbf{I}_n = M^{-1}AM - T(M^{-1}M) = M^{-1}(AM - TM) = M^{-1}(A - T\mathbf{I}_n)M.$$

Pertanto

$$|B - T\mathbf{I}_n| = |M^{-1}| \cdot |A - T\mathbf{I}_n| \cdot |M| = |A - T\mathbf{I}_n|.$$

□

Nota 22. Il coefficiente di T^n in $P_A(T)$ è $(-1)^n$, e quindi $(-1)^n P_A(T)$ è un polinomio monico.

Dalla proposizione 4.33 deduciamo il seguente corollario.

Corollario 10. Sia \mathbf{V} uno spazio vettoriale di dimensione finita n , e sia $F \in \text{end}(\mathbf{V})$. Allora $\lambda \in \mathbb{K}$ è un autovalore di F se e solo se λ è radice di $P_F(T)$. In particolare F possiede al più n autovalori distinti.

Proof. La prima asserzione è una riformulazione della proposizione 4.33. Poiché $P_A(T)$ ha grado uguale a $\dim(\mathbf{V})$, l'ulteriore asserzione segue dal fatto che un polinomio di grado n a coefficienti in \mathbb{K} possiede al più n radici in \mathbb{K} . □

Il corollario 10 fornisce un metodo pratico per calcolare gli autovalori e autovettori di un operatore.

Il problema di stabilire se un operatore è diagonalizzabile è un problema di ricerca di autovalori e dei relativi autovettori.

Proposizione 4.35. Sia \mathbf{V} uno spazio vettoriale di dimensione finita. Un operatore $F \in \text{end}(\mathbf{V})$ è diagonalizzabile se e solo se \mathbf{V} possiede una base costituita da autovettori di F .

Proof. Segue immediatamente dalla 124. □

Il risultato seguente dà una condizione necessaria e sufficiente affinché un operatore sia diagonalizzabile.

Teorema 4.36. Sia \mathbf{V} un \mathbb{K} -spazio vettoriale, $\dim(\mathbf{V}) = n$, e sia $F \in \text{end}(\mathbf{V})$. Se $\{\lambda_1, \dots, \lambda_k\} \subset \mathbb{K}$ è lo spettro di F , allora si ha:

$$\sum_{i=1}^k \dim(\mathbf{V}_{\lambda_i}(F)) \leq n \quad (130)$$

e l'uguaglianza sussiste se e solo se F è diagonalizzabile.

Proof. Per ogni $i \in \{1, \dots, k\}$ poniamo $d(i) = \dim(V_{\lambda_i}(F))$, e sia $\{e_{i1}, \dots, e_{id(i)}\}$ una base di $V_{\lambda_i}(F)$. In virtù della proposizione 4.35 sarà sufficiente dimostrare che i vettori

$$\mathbf{e}_{11}, \dots, \mathbf{e}_{1d(1)}, \mathbf{e}_{21}, \dots, \mathbf{e}_{2d(2)}, \dots, \mathbf{e}_{k1}, \dots, \mathbf{e}_{kd(k)}$$

sono linearmente indipendenti. Supponiamo che si abbia

$$\mathbf{0} = \sum_{j=1}^{d(1)} c_{1j} \mathbf{e}_{1j} + \sum_{j=1}^{d(2)} c_{2j} \mathbf{e}_{2j} + \dots + \sum_{j=1}^{d(k)} c_{kj} \mathbf{e}_{kj}$$

cioè

$$\mathbf{0} = \sum_{i=1}^k \sum_{j=1}^{d(i)} c_{ij} \mathbf{e}_{ij} \quad (131)$$

per opportuni scalari c_{ij} . Ponendo

$$\mathbf{v}_i = \sum_{j=1}^{d(i)} c_{ij} \mathbf{e}_{ij},$$

si ha $\mathbf{v}_i \in V_{\lambda_i}(F)$ e la 131 può essere riscritta nella forma seguente:

$$\mathbf{0} = \sum_{i=1}^k \mathbf{v}_i \quad (132)$$

Poiché

$$\mathbf{v}_i = \mathbf{0} \iff c_{i1} = c_{i2} = \dots = c_{id(i)} = 0,$$

sarà sufficiente dimostrare che

$$\mathbf{v}_1 = \mathbf{v}_2 = \dots = \mathbf{v}_k = \mathbf{0}.$$

Se $\mathbf{v}_i \neq \mathbf{0}$, allora \mathbf{v}_i è un autovettore relativo a λ_i e dalla 4.31 segue che $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\} \setminus \{\mathbf{0}\}$ è un insieme di vettori linearmente indipendenti. Quindi il secondo membro della 132 può essere uguale a $\mathbf{0}$ se e solo se tutti gli addendi sono $\mathbf{0}$. \square

Un caso particolare importante del teorema 4.36 è il seguente immediato corollario.

Corollario 11. *Se $\dim(\mathbf{V}) = n$ ed $F \in \text{end}(\mathbf{V})$ possiede n autovalori distinti, allora F è diagonalizzabile.*

Si noti che la condizione sufficiente di diagonalizzabilità espressa dal corollario 11 non è necessaria. Infatti l'operatore $\mathbf{1}_{\mathbf{V}}$ è diagonalizzabile qualunque sia $n = \dim(\mathbf{V})$, ma possiede l'unico autovalore $\lambda = 1$.

Il corollario 10 fornisce un metodo pratico per calcolare autovettori e autovalori di un operatore o di una matrice. Scegliendo una base di \mathbf{V} ci si riduce a considerare il solo caso delle matrici. Sia dunque assegnata $A \in M_n(\mathbb{K})$. Si cominci con il calcolare gli eventuali autovalori, che si trovano calcolando il polinomio caratteristico $P_A(T)$ e le sue radici in \mathbb{K} . Per ogni autovalore $\lambda \in \mathbb{K}$, il sistema omogeneo di n equazioni nelle n incognite $\mathbf{X} = {}^t(X_1, \dots, X_n)$:

$$(A - \lambda I_n) \mathbf{X} = \mathbf{0}$$

ha rango $r < n$ e possiede quindi soluzioni non banali. Lo spazio delle soluzioni è l'autospazio $V_{\lambda}(A)$. Se la somma delle dimensioni degli autospazi così trovati al variare di λ è tra tutte le radici di $P_A(T)$ uguale a n , allora A è diagonalizzabile, per il teorema 4.36. Una base diagonalizzante si ottiene scegliendo una base di ciascun autospazio e prendendo l'unione. Ciò fornisce in linea di principio un metodo di calcolo di tutti i vettori di una base diagonalizzante e della matrice del corrispondente cambiamento di base. Si osservi che un operatore può non avere autovalori, e quindi neanche autovettori, perché il polinomio caratteristico $P_A(T)$ può non avere radici in \mathbb{K} . Se però $\mathbb{K} = \mathbb{C}$, allora dal teorema fondamentale dell'algebra segue che $P_A(T)$ possiede radici in \mathbb{C} . Pertanto: *ogni operatore di uno spazio vettoriale complesso di dimensione finita possiede necessariamente almeno un autovalore, e quindi possiede autovettori.* Ciò non significa necessariamente che l'operatore sia diagonalizzabile. Se $\mathbb{K} = \mathbb{R}$ può accadere che un operatore di uno spazio \mathbf{V} non possieda autovalore, allora il polinomio caratteristico non ha grado dispari, e quindi possiede almeno una radice reale. In conclusione, *ogni operatore di uno spazio vettoriale reale di dimensione dispari possiede almeno un autovalore e quindi possiede autovettori.*

Il polinomio caratteristico della matrice identità \mathbf{I}_n è

$$P_{\mathbf{I}_n}(T) = (1 - T)^n.$$

Il polinomio caratteristico della matrice nulla $n \times n$ è

$$P_0(T) = (-1)^n T^n.$$

In entrambi i casi l'unico autospazio della matrice è \mathbb{K}^n .

Se $A = (a_{ij}) \in M_n(\mathbb{K})$ è una matrice triangolare (superiore o inferiore), si ha

$$P_A(T) = (a_{11} - T)(a_{22} - T) \cdots (a_{nn} - T).$$

Se $a_{11}, a_{22}, \dots, a_{nn}$ sono distinti, allora, per il corollario 11, A è diagonalizzabile perché possiede n autovalori distinti altrimenti in generale non è vero.

Esempio 4.10. La matrice $n \times n$, $n \geq 2$,

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \in M_n(\mathbb{K})$$

ha polinomio caratteristico

$$P_A(T) = (-1)^n T^n$$

e quindi possiede l'unico autovalore $\lambda = 0$. Da ciò segue che se A fosse diagonalizzabile sarebbe simile alla matrice $\mathbf{0}$. Ma $\mathbf{0}$ è simile soltanto a sé stessa. Infatti per ogni $M \in GL_n(\mathbb{K})$ si ha

$$M^{-1}\mathbf{0}M = \mathbf{0}.$$

Quindi A non è diagonalizzabile.

Esempio 4.11. La matrice

$$B = A + \mathbf{I}_n = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \in M_n(\mathbb{K})$$

ha polinomio caratteristico $P_B(T) = (1 - T)^n$, uguale a quello della matrice \mathbf{I}_n . B non è diagonalizzabile, perché se lo fosse sarebbe simile a \mathbf{I}_n , la quale invece è simile solo a sé stessa: infatti $M^{-1}\mathbf{I}_n M = \mathbf{I}_n$ per ogni $M \in GL_n(\mathbb{K})$.

Esempio 4.12. La matrice

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in M_2(\mathbb{R})$$

ha polinomio caratteristico $1 + T^2$. Poiché questo polinomio non ha radici reali, A non possiede autovalori né autovettori in \mathbb{R}^2 . Se però A viene considerata come una matrice ad elementi complessi, allora essa possiede i due autovalori distinti $\lambda = \pm i$. Per il corollario 11, A è diagonalizzabile in $M_2(\mathbb{C})$, ed è simile alla matrice diagonale

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

Per trovare la matrice M tale che $B = M^{-1}AM$, dobbiamo trovare una base di \mathbb{C}^2 costituita da autovettori di A , il che equivale a trovare un autovettore per ciascuno dei due autovalori. Possiamo procedere nel seguente modo.

Per la proposizione 4.33 gli autovettori relativi a $\lambda = i$ sono gli elementi del nucleo di $A - i\mathbf{I}_2$, cioè sono le soluzioni del sistema omogeneo

$$\begin{aligned} -iX + Y &= 0 \\ -X - iY &= 0 \end{aligned}$$

che ha rango 1, ed è quindi equivalente alla prima equazione. Soluzioni sono i vettori della forma (t, it) , i quali, al variare di $t \in \mathbb{C}$, descrivono l'autospazio $\mathbb{C}_i^2(A)$. Prendendo ad esempio $t = 1$ si ottiene l'autovettore $(1, i)$ relativo a $\lambda = i$. Analogamente, considerando il sistema omogeneo corrispondente a $A + i\mathbf{I}_2$, otteniamo i vettori della forma $(t, -it)$, e quindi $(i, 1)$ è un autovettore relativo a $\lambda = -i$. La base $\phi = \{(1, i), (i, 1)\}$ è diagonalizzante. Detta ψ la base canonica, si ha

$$M = M_{\psi, \phi}(\mathbf{1}) = \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$$

e

$$M^{-1}AM = \begin{pmatrix} \frac{1}{2} & -\frac{i}{2} \\ -\frac{i}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

Molteplicità geometrica e algebrica

Definizione 4.23. Sia \mathbf{V} un \mathbb{K} -spazio vettoriale di dimensione finita. Sia $F \in \text{end}(\mathbf{V})$ e sia $\lambda \in \mathbb{K}$ un autovalore di F . La $\dim(\mathbf{V}_\lambda(F))$ si dice **molteplicità geometrica** di λ per F . La **molteplicità algebrica** di λ per F è la molteplicità $h(\lambda)$ di λ come radice del polinomio caratteristico di F .

In generale la molteplicità geometrica e quella algebrica sono diverse. Ciò accade ad esempio per le matrici A e B dell'esempio 4.10. Per A si ha infatti, evidentemente, $h(0) = n$, mentre per B si ha $h(1) = n$. D'altra parte A non è diagonalizzabile e quindi $\dim(\mathbf{V}_0(A)) < n$. Similmente per B .

Proposizione 4.37. Per ogni $F \in \text{end}(\mathbf{V})$ e $\lambda \in \mathbb{K}$ autovalore di F sussiste la disuguaglianza

$$\dim(\mathbf{V}_\lambda(F)) \leq h(\lambda), \quad (133)$$

cioè la molteplicità geometrica non supera la molteplicità algebrica.

Proof. Supponiamo che $d = \dim(\mathbf{V}_\lambda(F)) \geq 1$, e sia $\psi = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una base di \mathbf{V} tale che $\{\mathbf{e}_1, \dots, \mathbf{e}_d\}$ sia una base di $\mathbf{V}_\lambda(F)$. Si ha

$$A = M_\psi(F) = \begin{pmatrix} \lambda \mathbf{I}_d & B \\ \mathbf{0} & C \end{pmatrix}$$

dove $B \in M_{d, n-d}(\mathbb{K})$, $\mathbf{0} \in M_{n-d, d}(\mathbb{K})$ è la matrice nulla, e $C \in M_{n-d, n-d}(\mathbb{K})$. Sviluppando $|A - T\mathbf{I}_n|$ con la regola di Laplace rispetto alle prime d righe si trova

$$P_F(T) = (\lambda - T)^d p(T),$$

dove $p(T) = P_C(T)$ è un polinomio di grado $n - d$. Pertanto $h(\lambda) \geq d$. □

Se il campo \mathbb{K} è algebricamente chiuso e $\lambda_1, \dots, \lambda_k$ sono gli autovalori di F , si ha

$$h(\lambda_1) + \dots + h(\lambda_k) = n.$$

Pertanto, dalla 133 e dal teorema 4.36 segue che:

Corollario 12. Se il campo \mathbb{K} è algebricamente chiuso l'operatore F è diagonalizzabile se e solo se per ogni autovalore λ di F si ha

$$\dim(\mathbf{V}_\lambda(F)) = h(\lambda),$$

cioè se e solo se la molteplicità geometrica e la molteplicità algebrica di ogni autovalore λ coincidono.

Proposizione 4.38. Se $A = (a_{ij}) \in M_2(\mathbb{K})$, allora si ha

$$P_A(T) = T^2 - \text{tr}(A)T + \det(A),$$

dove $\text{tr}(A) = a_{11} + a_{22}$ è la **traccia** di A .

Proof. Da fare. □

Proposizione 4.39. Sia

$$P(T) = g_0 + g_1T + \cdots + g_{n-1}T^{n-1} + T^n$$

un polinomio monico di grado n a coefficienti in \mathbb{K} , e sia

$$M_P = \begin{pmatrix} 0 & 0 & \cdots & 0 & -g_0 \\ 1 & 0 & \cdots & 0 & -g_1 \\ 0 & 1 & \cdots & 0 & -g_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -g_{n-1} \end{pmatrix}$$

Il polinomio caratteristico di M_P è uguale a $(-1)^n P(T)$.

Proof. Se $n = 1$, l'affermazione è evidente: $M_P - T = -g_0 - T$. Procediamo per induzione su n , e supponiamo $n \geq 2$. Si ha

$$\begin{aligned} |M_P - T\mathbf{I}_n| &= \begin{vmatrix} -T & 0 & \cdots & 0 & -g_0 \\ 1 & -T & \cdots & 0 & -g_1 \\ 0 & 1 & \cdots & 0 & -g_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -g_{n-1} - T \end{vmatrix} = \\ &= -T \begin{vmatrix} -T & \cdots & 0 & -g_1 \\ 1 & \cdots & 0 & -g_2 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & \cdots & -g_{n-1} - T \end{vmatrix} + (-1)^n g_0 \begin{vmatrix} 1 & -T & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} \\ &= -T(-1)^{n-1}(g_1 + g_2T + \cdots + g_{n-1}T^{n-2} + T^{n-1}) + (-1)^n g_0 = (-1)^n P(T), \end{aligned}$$

dove il valore del primo determinante è dedotto dall'ipotesi induttiva. □

Questa proposizione dimostra che per ogni polinomio monico di grado $n \geq 1$ in $\mathbb{K}[T]$ esistono matrici quadrate di ordine n di cui esso, o il suo opposto, a seconda che n sia pari o dispari, è il polinomio caratteristico.

Part II

Geometria Euclidea

5 Forme bilineari e quadratiche

Forme bilineari

Definizione 5.1. Sia \mathbf{V} un \mathbb{K} -spazio vettoriale. Un'applicazione

$$b : \mathbf{V}^2 \rightarrow \mathbb{K}$$

si dice **forma bilineare su \mathbf{V}** se è lineare in ognuno dei due argomenti, cioè se soddisfa le seguenti condizioni:

$$\text{FB1 } b(\mathbf{v} + \mathbf{v}', \mathbf{w}) = b(\mathbf{v}, \mathbf{w}) + b(\mathbf{v}', \mathbf{w})$$

$$\text{FB2 } b(\mathbf{v}, \mathbf{w} + \mathbf{w}') = b(\mathbf{v}, \mathbf{w}) + b(\mathbf{v}, \mathbf{w}')$$

$$\text{FB3 } b(k\mathbf{v}, \mathbf{w}) = b(\mathbf{v}, k\mathbf{w}) = kb(\mathbf{v}, \mathbf{w})$$

per ogni $\mathbf{v}, \mathbf{v}', \mathbf{w}, \mathbf{w}' \in \mathbf{V}$, $k \in \mathbb{K}$. La forma bilineare b si dice **simmetrica** se

$$\forall \mathbf{v}, \mathbf{w} \in \mathbf{V}, \quad b(\mathbf{v}, \mathbf{w}) = b(\mathbf{w}, \mathbf{v});$$

b si dice **antisimmetrica**, o **alterna**, se

$$\forall \mathbf{v}, \mathbf{w} \in \mathbf{V}, \quad b(\mathbf{v}, \mathbf{w}) = -b(\mathbf{w}, \mathbf{v}).$$

Insieme delle forme bilineari

Definizione 5.2. Sia \mathbf{V} un \mathbb{K} -spazio vettoriale. Allora definiremo l'insieme delle forme bilineari al seguente modo:

$$\mathbf{V}^{\mathbb{K}} = \{b \in \text{hom}(\mathbf{V}^2, \mathbb{K}) : b \text{ è una forma bilineare}\}.$$

Siano definiti ulteriormente gli insiemi delle forme bilineari simmetriche e antisimmetriche:

$$\mathbf{V}_s^{\mathbb{K}} = \{b \in \mathbf{V}^{\mathbb{K}} : b \text{ è simmetrica}\},$$

$$\mathbf{V}_a^{\mathbb{K}} = \{b \in \mathbf{V}^{\mathbb{K}} : b \text{ è antisimmetrica}\}.$$

Proposizione 5.1. Sia $b \in \mathbf{V}^{\mathbb{K}}$. Allora

$$b \text{ è antisimmetrica} \iff \forall \mathbf{v} \in \mathbf{V}, \quad b(\mathbf{v}, \mathbf{v}) = 0.$$

Proof. Infatti dalle FB1 e FB2, segue

$$\forall \mathbf{v}, \mathbf{w} \in \mathbf{V}, \quad 0 = b(\mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w}) = b(\mathbf{v}, \mathbf{v}) + b(\mathbf{v}, \mathbf{w}) + b(\mathbf{w}, \mathbf{v}) + b(\mathbf{w}, \mathbf{w}) = b(\mathbf{v}, \mathbf{w}) + b(\mathbf{w}, \mathbf{v}).$$

□

Forma bilineare nulla

Definizione 5.3. L'applicazione

$$0 : \mathbf{V}^2 \rightarrow \mathbb{K}, \quad (\mathbf{v}, \mathbf{w}) \mapsto 0$$

si dice **forma bilineare nulla**.

Proposizione 5.2. La forma bilineare nulla è sia simmetrica che antisimmetrica.

Proof. Simmetrica perchè $\forall \mathbf{v}, \mathbf{w} \in \mathbf{V}$, $b(\mathbf{v}, \mathbf{w}) = 0 = b(\mathbf{w}, \mathbf{v})$, e antisimmetrica per la 5.1. \square

Sia $A = (a_{ij}) \in M_n(\mathbb{K})$; considerando i vettori di \mathbb{K}^n come degli n -vettori colonna, otteniamo una forma bilineare su \mathbb{K}^n ponendo

$$b : (\mathbb{K}^n)^2 \rightarrow \mathbb{K}, \quad (\mathbf{x}, \mathbf{y}) \mapsto {}^t\mathbf{x}A\mathbf{y} = \sum_{i,j} a_{ij}x_iy_j.$$

Dalle proprietà del prodotto di matrici segue che in questo modo si è definita una forma bilineare. Se $\mathbf{E}_1, \dots, \mathbf{E}_n$ è la base canonica di \mathbb{K}^n , si ha

$$\forall 1 \leq i, j \leq n, \quad b(\mathbf{E}_i, \mathbf{E}_j) = a_{ij}.$$

Esempio 5.1. Prendiamo $n = 3$ e

$$A = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 0 & \frac{1}{2} \\ -1 & \pi & 0 \end{pmatrix}$$

allora la corrispondente forma bilineare su \mathbb{K}^3 è

$$b(\mathbf{x}, \mathbf{y}) = x_1y_1 - 2x_1y_3 + \frac{x_2y_3}{2} - x_3y_1 + \pi x_2y_3.$$

Questa forma bilineare non è simmetrica perchè

$$b(\mathbf{E}_1, \mathbf{E}_3) = -2 \neq -1 = b(\mathbf{E}_3, \mathbf{E}_1).$$

Forma simmetrica standard

Definizione 5.4. Prendendo $A = \mathbf{I}_n$, otteniamo

$$b(\mathbf{x}, \mathbf{y}) = {}^t\mathbf{x}\mathbf{y} = \sum_{i=1}^n x_iy_i. \quad (134)$$

La b definita dalla 134 è una forma bilineare simmetrica, chiamata **forma simmetrica standard** su \mathbb{K}^n .

Forma alterna standard

Definizione 5.5. Se $n = 2k$, cioè se n è pari, e se si prende $A = \mathbf{J}_k$, dove

$$\mathbf{J}_k = \begin{pmatrix} \mathbf{0}_k & \mathbf{I}_k \\ -\mathbf{I}_k & \mathbf{0}_k \end{pmatrix}$$

si ottiene

$$b(\mathbf{x}, \mathbf{y}) = x_1y_{k+1} + \dots + x_ky_{2k} - \dots - x_{2k}y_k$$

che è una forma bilineare alterna, chiamata **forma alterna standard** su \mathbb{K}^n .

Sia $b : \mathbf{V}^2 \rightarrow \mathbb{K}$ una forma bilineare. La bilinearità di b permette di definire due applicazioni lineari in $\text{hom}(\mathbf{V}, \mathbf{V}^\vee)$ nel modo seguente.

Proposizione 5.3. Per ogni $\mathbf{v} \in \mathbf{V}$ l'applicazione

$$b_{\mathbf{v}} : \mathbf{V} \rightarrow \mathbb{K}, \quad \mathbf{w} \mapsto b(\mathbf{v}, \mathbf{w})$$

è un funzionale lineare.

Proof. Infatti dalla definizione segue che per ogni $\mathbf{w}_1, \mathbf{w}_2 \in \mathbf{V}$, $c_1, c_2 \in \mathbb{K}$,

$$\begin{aligned} b_{\mathbf{v}}(c_1 \mathbf{w}_1 + c_2 \mathbf{w}_2) &= b(\mathbf{v}, c_1 \mathbf{w}_1 + c_2 \mathbf{w}_2) = c_1 b(\mathbf{v}, \mathbf{w}_1) + c_2 b(\mathbf{v}, \mathbf{w}_2) \\ &= c_1 b_{\mathbf{v}}(\mathbf{w}_1) + c_2 b_{\mathbf{v}}(\mathbf{w}_2). \end{aligned}$$

□

Quindi, ponendo $\delta_b(\mathbf{v}) = b_{\mathbf{v}}$ per ogni $\mathbf{v} \in \mathbf{V}$, si ottiene un'applicazione

$$\delta_b : \mathbf{V} \rightarrow \mathbf{V}^\sim.$$

Proposizione 5.4. *La δ_b è lineare.*

Proof. Infatti per ogni $\mathbf{v}_1, \mathbf{v}_2 \in \mathbf{V}$, $c_1, c_2 \in \mathbb{K}$ si ha

$$\begin{aligned} [\delta_b(c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2)](\mathbf{w}) &= b_{c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2}(\mathbf{w}) = b(c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2, \mathbf{w}) \\ &= c_1 b(\mathbf{v}_1, \mathbf{w}) + c_2 b(\mathbf{v}_2, \mathbf{w}) = c_1 b_{\mathbf{v}_1}(\mathbf{w}) + c_2 b_{\mathbf{v}_2}(\mathbf{w}) \\ &= [c_1 \delta_b(\mathbf{v}_1) + c_2 \delta_b(\mathbf{v}_2)](\mathbf{w}) \end{aligned}$$

per ogni $\mathbf{w} \in \mathbf{V}$, cioè $\delta_b(c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2) = c_1 \delta_b(\mathbf{v}_1) + c_2 \delta_b(\mathbf{v}_2)$.

□

In modo simile si verifica che ponendo $\delta'_b(\mathbf{w}) = b'_{\mathbf{w}}$, dove

$$b'_{\mathbf{w}} : \mathbf{V} \rightarrow \mathbb{K}, \quad b'_{\mathbf{w}}(\mathbf{v}) = b(\mathbf{v}, \mathbf{w}),$$

si definisce un'applicazione lineare

$$\delta'_b : \mathbf{V} \rightarrow \mathbf{V}^\sim.$$

Proposizione 5.5. *$\delta_b = \delta'_b$ se e solo se b è simmetrica.*

Proof.

$$\delta_b(\mathbf{v}) = b_{\mathbf{v}}(\mathbf{w}) = b(\mathbf{v}, \mathbf{w}) = b(\mathbf{w}, \mathbf{v}) = b'_{\mathbf{v}}(\mathbf{w}) = \delta'_b(\mathbf{v}).$$

□

Matrice di una forma bilineare rispetto a una base

Definizione 5.6. Siano \mathbf{V} un \mathbb{K} -spazio vettoriale di dimensione n , $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una sua base, e $b \in \mathbf{V}^{\mathbb{K}}$. La **matrice di b (o che rappresenta b) rispetto alla base $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$** è la matrice $A = (a_{ij}) \in M_n(\mathbb{K})$ così definita:

$$a_{ij} = b(\mathbf{e}_i, \mathbf{e}_j), \quad 1 \leq i, j \leq n.$$

Proposizione 5.6. *La matrice A individua la forma bilineare b .*

Proof. Infatti per ogni $\mathbf{v} = \sum_{i=1}^n x_i \mathbf{e}_i$, $\mathbf{w} = \sum_{j=1}^n y_j \mathbf{e}_j \in \mathbf{V}$, si ha

$$\begin{aligned} b(\mathbf{v}, \mathbf{w}) &= b\left(\sum_{i=1}^n x_i \mathbf{e}_i, \sum_{j=1}^n y_j \mathbf{e}_j\right) = \sum_{i=1}^n b\left(x_i \mathbf{e}_i, \sum_{j=1}^n y_j \mathbf{e}_j\right) = \sum_{i=1}^n x_i b\left(\mathbf{e}_i, \sum_{j=1}^n y_j \mathbf{e}_j\right) \\ &= \sum_{i=1}^n x_i \left(\sum_{j=1}^n b(\mathbf{e}_i, y_j \mathbf{e}_j)\right) = \sum_{i=1}^n x_i \left(\sum_{j=1}^n y_j b(\mathbf{e}_i, \mathbf{e}_j)\right) = \sum_{i,j} x_i y_j b(\mathbf{e}_i, \mathbf{e}_j) = \mathbf{x}^t A \mathbf{y}. \end{aligned}$$

dove abbiamo denotato con \mathbf{x} e \mathbf{y} i vettori colonna delle coordinate di \mathbf{v} e \mathbf{w} rispettivamente.

□

Proposizione 5.7. Viceversa, se $A = (a_{ij}) \in M_n(\mathbb{K})$ è una qualunque matrice quadrata di ordine n ed $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ è una base di \mathbf{V} , ponendo

$$b(\mathbf{v}, \mathbf{w}) = \sum_{i,j} a_{ij} x_i y_j = {}^t \mathbf{x} A \mathbf{y}$$

per ogni $\mathbf{v}(x_1, \dots, x_n), \mathbf{w}(y_1, \dots, y_n) \in \mathbf{V}$, si definisce una forma bilineare su \mathbf{V} .

Proof. Infatti, per ogni $\mathbf{v}, \mathbf{w}, \mathbf{v}'(x'_1, \dots, x'_n), \mathbf{w}'(y'_1, \dots, y'_n) \in \mathbf{V}, k \in \mathbb{K}$ si ha

$$\begin{aligned} b(\mathbf{v} + \mathbf{v}', \mathbf{w}) &= {}^t(\mathbf{x} + \mathbf{x}') A \mathbf{y} = {}^t \mathbf{x} A \mathbf{y} + {}^t \mathbf{x}' A \mathbf{y} = b(\mathbf{v}, \mathbf{w}) + b(\mathbf{v}', \mathbf{w}), \\ b(\mathbf{v}, \mathbf{w} + \mathbf{w}') &= {}^t \mathbf{x} A (\mathbf{y} + \mathbf{y}') = {}^t \mathbf{x} A \mathbf{y} + {}^t \mathbf{x} A \mathbf{y}' = b(\mathbf{v}, \mathbf{w}) + b(\mathbf{v}, \mathbf{w}'), \\ b(k\mathbf{v}, \mathbf{w}) &= {}^t(k\mathbf{x}) A \mathbf{y} = k {}^t \mathbf{x} A \mathbf{y} = kb(\mathbf{v}, \mathbf{w}), \\ b(\mathbf{v}, k\mathbf{w}) &= {}^t \mathbf{x} A (k\mathbf{y}) = k {}^t \mathbf{x} A \mathbf{y} = kb(\mathbf{v}, \mathbf{w}). \end{aligned}$$

È evidente che la forma bilineare b così definita ha proprio A come matrice rispetto alla base $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$. \square

Si osservi inoltre che si ha

$$b(\mathbf{w}, \mathbf{v}) = {}^t \mathbf{y} A \mathbf{x} = {}^t \mathbf{x}^t A \mathbf{y}$$

e quindi

$$b(\mathbf{w}, \mathbf{v}) = b(\mathbf{v}, \mathbf{w})$$

per ogni $\mathbf{v}(x_1, \dots, x_n), \mathbf{w}(y_1, \dots, y_n) \in \mathbf{V}$ se e solo se $A = {}^t A$. In altre parole, la forma bilineare b è simmetrica se e solo se la matrice A è simmetrica. Analogamente, b è antisimmetrica se e solo se $A = -{}^t A$, cioè se e solo se A è antisimmetrica.

Riassumendo, possiamo enunciare la seguente proposizione.

Proposizione 5.8. Siano \mathbf{V} un \mathbb{K} -spazio vettoriale di dimensione finita, e $\psi = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una sua base. Associando ad ogni forma bilineare la sua matrice rispetto a ψ si ottiene una corrispondenza biunivoca tra l'insieme $\mathbf{V}^{\mathbb{K}}$ ed $M_n(\mathbb{K})$. Tale corrispondenza induce un isomorfismo bilineare dell'insieme delle forme bilineari simmetriche (forme bilineari antisimmetriche) sull'insieme delle matrici simmetriche (matrici antisimmetriche).

Ovviamente la corrispondenza biunivoca descritta dalla proposizione 5.8 dipende dalla base che si è scelta, cioè le matrici che rappresentano una data forma bilineare rispetto a due diverse basi sono in generale diverse.

Siano $b \in \mathbf{V}^{\mathbb{K}}$, e $\psi = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ e $\phi = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ due basi di \mathbf{V} . Siano

$$A = (a_{ij}) = (b(\mathbf{e}_i, \mathbf{e}_j)), \quad B = (b_{ij}) = (b(\mathbf{f}_i, \mathbf{f}_j))$$

le matrici che rappresentano b rispetto a ψ e ϕ rispettivamente. Se $\mathbf{v}, \mathbf{w} \in \mathbf{V}$ sono due vettori qualunque, con coordinate rispetto alle due basi

$$\mathbf{v} = \sum_{i=1}^n x_i \mathbf{e}_i = \sum_{i=1}^n x'_i \mathbf{f}_i, \quad \mathbf{w} = \sum_{j=1}^n y_j \mathbf{e}_j = \sum_{j=1}^n y'_j \mathbf{f}_j,$$

si ha

$$b(\mathbf{v}, \mathbf{w}) = {}^t \mathbf{x} A \mathbf{y} = {}^t \mathbf{x}' B \mathbf{y}' \quad (135)$$

Posto $M = M_{\psi, \phi}(\mathbf{1}_{\mathbf{V}})$, si ha $\mathbf{x} = M \mathbf{x}'$, $\mathbf{y} = M \mathbf{y}'$ e, sostituendo nella 135, si ottiene

$${}^t \mathbf{x}' B \mathbf{y}' = {}^t (M \mathbf{x}') A (M \mathbf{y}') = {}^t \mathbf{x}' ({}^t M A M) \mathbf{y}' \quad (136)$$

Poiché la 136 è vera per ogni $\mathbf{x}', \mathbf{y}' \in \mathbb{K}^n$, deduciamo che

$$B = {}^t M A M \quad (137)$$

La 137 esprime la relazione esistente tra le matrici A e B che rappresentano la forma bilineare b rispetto alle due basi ψ e ϕ rispettivamente. Viceversa, se A è la matrice che rappresenta la forma bilineare b rispetto alla base ψ , e se $M \in \text{GL}_n(\mathbb{K})$ è una qualunque matrice invertibile, e poniamo ϕ la base di \mathbf{V} tale che $M = M_{\phi, \psi}(\mathbf{1}_{\mathbf{V}})$. Pertanto $B = {}^t M A M$ è la matrice che rappresenta b rispetto a ϕ .

Matrici congruenti

Diremo che due matrici $A, B \in M_n(\mathbb{K})$ sono **congruenti** se

$$\exists M \in GL_n(\mathbb{K}) : B = {}^t M A M. \quad (138)$$

Proposizione 5.9. *La congruenza di matrici è una relazione di equivalenza in $M_n(\mathbb{K})$.*

Proposizione 5.10. *Sia \mathbf{V} un \mathbb{K} -spazio vettoriale di dimensione n . Due matrici $A, B \in M_n(\mathbb{K})$ rappresentano la stessa forma bilineare b su \mathbf{V} rispetto a due diverse basi se e solo se sono congruenti.*

Rango di una forma bilineare Dalla proposizione 4.3(2) (?) segue che due matrici congruenti hanno lo stesso rango. Pertanto, per la proposizione 5.10, il rango della matrice A che rappresenta una data forma bilineare b rispetto a una base qualsiasi non dipende dalla base, ma solo da b : chiameremo r **rango della forma bilineare b** .

Forma bilineare non degenera e degenera

Definizione 5.7. Se b ha rango $r = \dim(\mathbf{V})$ ($r < \dim(\mathbf{V})$) la forma bilineare si dice **non degenera (degenera)**.

La seguente proposizione dà diverse caratterizzazioni delle forme bilineari non degeneri.

Proposizione 5.11. *Siano \mathbf{V} un \mathbb{K} -spazio vettoriale di dimensione finita e $b \in \mathbf{V}^{\mathbb{K}}$. Sono equivalenti le seguenti condizioni:*

1. b è non degenera;
2. $\forall \mathbf{v} \in \mathbf{V} \setminus \{\mathbf{0}\}, \exists \mathbf{w} \in \mathbf{V} : b(\mathbf{v}, \mathbf{w}) \neq 0$;
3. $\forall \mathbf{w} \in \mathbf{V} \setminus \{\mathbf{0}\}, \exists \mathbf{v} \in \mathbf{V} \setminus \{\mathbf{0}\} : b(\mathbf{v}, \mathbf{w}) \neq 0$;
4. L'applicazione $\delta_b : \mathbf{V} \rightarrow \mathbf{V}^\sim$ è un isomorfismo;
5. L'applicazione $\delta'_b : \mathbf{V} \rightarrow \mathbf{V}^\sim$ è un isomorfismo.

Proof. Scelta una base $\psi = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ di \mathbf{V} , sia $A \in M_n(\mathbb{K})$ la matrice di b rispetto a ψ .

- (1) \rightarrow (2) Se A ha rango n e $\mathbf{x} \neq \mathbf{0}$ è il vettore delle coordinate di \mathbf{v} , allora ${}^t \mathbf{x} A \neq (0, \dots, 0)$, e quindi esiste $\mathbf{y} \in \mathbb{K}^n$ tale che ${}^t \mathbf{x} A \mathbf{y} \neq 0$; il vettore \mathbf{w} di coordinate \mathbf{y} è tale che $b(\mathbf{v}, \mathbf{w}) \neq 0$;
- (2) \rightarrow (1) Per ipotesi per ogni $\mathbf{x} \neq \mathbf{0}$ esiste \mathbf{y} tale che ${}^t \mathbf{x} A \mathbf{y} \neq 0$; ciò implica ${}^t \mathbf{x} A \neq (0, \dots, 0)$, e quindi ogni riga di A ha rango n ;
- (1) \leftrightarrow (3) Si dimostra in modo simile.
- (2) \rightarrow (4) Poiché $\dim(\mathbf{V}) = \dim(\mathbf{V}^\sim)$, è sufficiente far vedere che $N(\delta_b) = \{\mathbf{0}\}$. Sia dunque $\mathbf{v} \in \mathbf{V}$ tale che $b_{\mathbf{v}}$ sia il funzionale nullo. Allora

$$\mathbf{0} = b_{\mathbf{v}}(\mathbf{w}) = b(\mathbf{v}, \mathbf{w})$$

per ogni $\mathbf{w} \in \mathbf{V}$. Ciò contraddice la (2) a meno che $\mathbf{v} = \mathbf{0}$;

- (4) \rightarrow (2) Per ogni $\mathbf{v} \in \mathbf{V} \setminus \{\mathbf{0}\}$, si ha $\delta_b(\mathbf{v}) = b_{\mathbf{v}} \neq 0$, e quindi esiste $\mathbf{w} \in \mathbf{V}$ tale che $0 \neq b_{\mathbf{v}}(\mathbf{w}) = b(\mathbf{v}, \mathbf{w})$;
- (4) \rightarrow (5) Si dimostra in modo simile.

□

D'ora in poi ci limiteremo a considerare le forme bilineari simmetriche, le quali hanno una particolare importanza per gli argomenti geometrici che svilupperemo.

Vettori ortogonali

Definizione 5.8. Siano $b \in \mathbf{V}_s^{\mathbb{K}}$, e $\mathbf{v} \in \mathbf{V}$. Un vettore $\mathbf{w} \in \mathbf{V}$ si dice **ortogonale** (o **perpendicolare**) a \mathbf{v} , se $b(\mathbf{v}, \mathbf{w}) = 0$. In tal caso i due vettori \mathbf{v} e \mathbf{w} si dicono **ortogonali** (o **perpendicolari**).

Siano $b \in \mathbf{V}_s^{\mathbb{K}}$ e \mathbf{S} un sottinsieme di \mathbf{V} . L'insieme dei vettori ortogonali a tutti gli elementi di \mathbf{S} si denota con \mathbf{S}^\perp . In simboli:

$$\mathbf{S}^\perp = \{\mathbf{w} \in \mathbf{V} : \forall \mathbf{v} \in \mathbf{S}, b(\mathbf{v}, \mathbf{w}) = 0\}.$$

Proposizione 5.12. \mathbf{S}^\perp è un sottospazio vettoriale di \mathbf{V} .

Proof. Infatti, se $\mathbf{w}, \mathbf{w}' \in \mathbf{S}^\perp$, $k \in \mathbb{K}$, allora per ogni $\mathbf{v} \in \mathbf{S}$

$$\begin{aligned} b(\mathbf{v}, \mathbf{w} + \mathbf{w}') &= b(\mathbf{v}, \mathbf{w}) + b(\mathbf{v}, \mathbf{w}') = 0 + 0 = 0 \\ b(\mathbf{v}, k\mathbf{w}) &= kb(\mathbf{v}, \mathbf{w}) = k \cdot 0 = 0, \end{aligned}$$

□

Sottospazio ortogonale a un insieme di vettori

Definizione 5.9. Chiamiamo \mathbf{S}^\perp il **sottospazio ortogonale** ad \mathbf{S} .

Se $\mathbf{S} = \{\mathbf{v}\}$, scriveremo \mathbf{v}^\perp anziché $\{\mathbf{v}\}^\perp$.

Sottospazi ortogonali, Radicale di uno spazio vettoriale

Definizione 5.10. Due sottospazi \mathbf{U} e \mathbf{W} di \mathbf{V} si dicono **ortogonali** se $\mathbf{U} \subset \mathbf{W}^\perp$; dalla simmetria di b segue immediatamente che questa condizione è equivalente a $\mathbf{W} \subset \mathbf{U}^\perp$. Il sottospazio \mathbf{V}^\perp è detto **radicale** di \mathbf{V} .

Dalla proposizione 5.11, segue che b è non degenere se e solo se $\mathbf{V}^\perp = \langle \mathbf{0} \rangle$.

Vettore isotropo rispetto a una forma bilineare

Definizione 5.11. Un vettore $\mathbf{v} \in \mathbf{V}$ è detto **isotropo** rispetto alla forma bilineare simmetrica b se $\mathbf{v} \in \mathbf{v}^\perp$, cioè se $b(\mathbf{v}, \mathbf{v}) = 0$.

Ovviamente $\mathbf{0}$ è isotropo. Se $\mathbf{v} \in \mathbf{V}$ è isotropo e $k \in \mathbb{K}$, allora si ha

$$b(k\mathbf{v}, k\mathbf{v}) = k^2 b(\mathbf{v}, \mathbf{v}) = k^2 \cdot 0 = 0,$$

e quindi il sottospazio $\langle \mathbf{v} \rangle$ consiste di vettori isotropi.

Se \mathbf{v} non è un vettore isotropo posto, per ogni $\mathbf{w} \in \mathbf{V}$,

$$a_{\mathbf{v}}(\mathbf{w}) = \frac{b(\mathbf{v}, \mathbf{w})}{b(\mathbf{v}, \mathbf{v})}, \quad (139)$$

si ha

$$b(\mathbf{v}, \mathbf{w} - a_{\mathbf{v}}(\mathbf{w})\mathbf{v}) = 0$$

cioè $\mathbf{w} - a_{\mathbf{v}}(\mathbf{w})\mathbf{v} \in \mathbf{v}^\perp$. Poiché

$$\mathbf{w} = a_{\mathbf{v}}(\mathbf{w})\mathbf{v} + (\mathbf{w} - a_{\mathbf{v}}(\mathbf{w})\mathbf{v})$$

deduciamo che

$$\mathbf{V} = \langle \mathbf{v} \rangle + \mathbf{v}^\perp.$$

D'altra parte $\langle \mathbf{v} \rangle \cap \mathbf{v}^\perp = \langle \mathbf{0} \rangle$ perché \mathbf{v} non è isotropo e pertanto $\langle \mathbf{v} \rangle \cap \mathbf{v}^\perp$ è un sottospazio proprio di $\langle \mathbf{v} \rangle$. Quindi per ogni vettore non isotropo $\mathbf{v} \in \mathbf{V}$ si ha

$$\mathbf{V} = \langle \mathbf{v} \rangle \oplus \mathbf{v}^\perp. \quad (140)$$

Lo scalare $a_{\mathbf{v}}(\mathbf{w})$ definito dalla 139 è detto **coefficiente di Fourier** di \mathbf{w} rispetto a \mathbf{v} .

Nota 23. $a_{\mathbf{v}}(\mathbf{w})$ è definito solo se \mathbf{v} non è un vettore isotropo.

Base ortogonale (diagonalizzante)

Definizione 5.12. Se \mathbf{V} ha dimensione finita e $\psi = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ è una base di \mathbf{V} tale che i vettori che ne fanno parte siano a due a due ortogonali, cioè soddisfino la condizione

$$\forall i \neq j, \quad b(\mathbf{e}_i, \mathbf{e}_j) = 0$$

allora ψ è una **base ortogonale**, o **diagonalizzante**, per b .

Se ψ è una base ortogonale, allora la matrice $A = (a_{ij})$ che rappresenta b rispetto a ψ è una matrice diagonale, perché $a_{ij} = b(\mathbf{e}_i, \mathbf{e}_j) = 0$ per ogni $i \neq j$. In tale base la forma bilineare si esprime pertanto nel modo seguente:

$$b(\mathbf{v}, \mathbf{w}) = \sum_{i=1}^n a_i x_i y_i. \quad (141)$$

Nota 24. Se una base ortogonale esiste, essa non è unica: ad esempio ogni base della forma

$$\{\lambda_1 \mathbf{e}_1, \lambda_2 \mathbf{e}_2, \dots, \lambda_n \mathbf{e}_n\}, \quad \lambda_1, \dots, \lambda_n \in \mathbb{K}^*,$$

è ancora ortogonale.

Forma quadratica determinata da una forma bilineare

Definizione 5.13. L'applicazione definita al seguente modo

$$q : \mathbf{V} \rightarrow \mathbb{K}, \mathbf{v} \mapsto b(\mathbf{v}, \mathbf{v})$$

è detta **forma quadratica determinata da** (o **associata a**) b .

Se ad esempio b è la forma bilineare (simmetrica) standard su \mathbb{K}^n , la forma quadratica ad essa associata è

$$q(\mathbf{x}) = \sum_{i=1}^n x_i^2,$$

chiamata **forma quadratica standard su \mathbb{K}^n** .

Proposizione 5.13. Sia \mathbf{V} uno spazio vettoriale su cui sia assegnata una forma bilineare simmetrica $b : \mathbf{V}^2 \rightarrow \mathbb{K}$. La forma quadratica q associata a b soddisfa le seguenti condizioni per ogni $k \in \mathbb{K}$, $\mathbf{v}, \mathbf{w} \in \mathbf{V}$:

1. $q(k\mathbf{v}) = k^2 q(\mathbf{v})$;
2. $2b(\mathbf{v}, \mathbf{w}) = q(\mathbf{v} + \mathbf{w}) - q(\mathbf{v}) - q(\mathbf{w})$.

Proof. La prima proprietà è immediata conseguenza della FB3. Si ha poi

$$q(\mathbf{v} + \mathbf{w}) - q(\mathbf{v}) - q(\mathbf{w}) = b(\mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w}) - b(\mathbf{v}, \mathbf{v}) - b(\mathbf{w}, \mathbf{w}) = b(\mathbf{v}, \mathbf{w}) + b(\mathbf{w}, \mathbf{v}) = 2b(\mathbf{v}, \mathbf{w}).$$

□

Dalla proposizione 5.13 discende in particolare che una forma quadratica q individua univocamente la forma bilineare simmetrica b cui è associata, perché b si esprime per mezzo di q ; segue da ciò che è equivalente assegnare su \mathbf{V} una forma bilineare simmetrica oppure la forma quadratica ad essa associata.

Forma bilineare polare

Definizione 5.14. La forma bilineare simmetrica b si dice **forma bilineare polare** della forma quadratica q .

Nel caso in cui \mathbf{V} ha dimensione finita diremo che q **ha rango** r se è il rango di b . Se $\psi = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ è una base dello spazio \mathbf{V} , e se $A = (a_{ij})$ è la matrice (simmetrica) che rappresenta la forma bilineare simmetrica b , si ha, per ogni $\mathbf{v}(x_1, \dots, x_n) \in \mathbf{V}$:

$$q(\mathbf{v}) = {}^t\mathbf{x}A\mathbf{x} = \sum_{i,j} a_{ij}x_ix_j.$$

Quindi $q(\mathbf{v}) = Q(\mathbf{x})$, dove

$$Q(\mathbf{X}) = {}^t\mathbf{X}A\mathbf{X} = \sum_{i,j} a_{ij}X_iX_j \quad (142)$$

è un polinomio omogeneo di grado 2 nelle indeterminate X_1, \dots, X_n che sono state rappresentate simbolicamente come un vettore colonna $\mathbf{X} = {}^t(X_1, \dots, X_n)$. Diremo che $Q(\mathbf{X})$ **rappresenta la forma quadratica q nella base ψ** .

Se ad esempio $\dim(\mathbf{V}) = 3$ e la matrice di b rispetto alla base $\psi = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ è

$$\begin{pmatrix} 2 & -1 & 5 \\ -1 & 0 & \frac{1}{3} \\ 5 & \frac{1}{3} & -3 \end{pmatrix}, \quad (143)$$

si ha

$$Q(\mathbf{X}) = 2X_1^2 - 2X_1X_2 + 10X_1X_3 + \frac{2}{3}X_2X_3 - 3X_3^2. \quad (144)$$

Si noti che ogni polinomio omogeneo di secondo grado in n indeterminate

$$Q(\mathbf{X}) = \sum_{1 \leq i \leq j \leq n} q_{ij}X_iX_j$$

può rappresentarsi nella forma 142 per un'opportuna matrice simmetrica A , e quindi, in una data base ψ di \mathbf{V} , Q rappresenta una forma quadratica q la cui forma bilineare polare è rappresentata dalla matrice A . La matrice $A = (a_{ij})$ è data dalla seguente formula:

$$\begin{aligned} a_{ii} &= q_{ii} & i &\in \{1, \dots, n\} \\ a_{ij} &= \frac{q_{ij}}{2}, & i &\neq j. \end{aligned}$$

Un polinomio omogeneo di secondo grado $Q(\mathbf{X})$ può sempre essere considerato come un'applicazione $Q : \mathbb{K}^n \rightarrow \mathbb{K}$. Ovviamente Q è una forma quadratica il cui polinomio associato rispetto alla base canonica è $Q(\mathbf{X})$ stesso. Pertanto spesso identificheremo il polinomio omogeneo di secondo grado con la forma quadratica. Un polinomio siffatto viene anche chiamato **forma quadratica n -aria**. Se A è una matrice diagonale il polinomio 142 è privo dei termini misti $q_{ij}X_iX_j$, $i \neq j$, e quindi è della forma

$$Q(\mathbf{X}) = \sum_{i=1}^n a_{ii}X_i^2. \quad (145)$$

Pertanto una base ψ di \mathbf{V} è diagonalizzante per la forma bilineare b se e solo se il polinomio che rappresenta la forma quadratica q è della forma 145. Diremo in tal caso che ψ è una **base diagonalizzante** per q .

Proposizione 5.14. Due polinomi omogenei di secondo grado $Q(\mathbf{X}) = {}^t\mathbf{X}A\mathbf{X}$ e $R(\mathbf{X}) = {}^t\mathbf{X}B\mathbf{X}$ nelle indeterminate $\mathbf{X} = {}^t(X_1, \dots, X_n)$ rappresentano lo stesso forma quadratica su \mathbf{V} in due basi diverse se e solo se le matrici simmetriche A e B sono congruenti.

Proof. Infatti, per la proposizione 5.10, l'essere congruenti è condizione necessaria e sufficiente affinché le matrici A e B rappresentino la stessa forma bilineare in due basi diverse. \square

Supponiamo che sul \mathbb{K} -spazio vettoriale \mathbf{V} sia definita una forma bilineare simmetrica $b : \mathbf{V}^2 \rightarrow \mathbf{K}$, con forma quadratica associata, q , e sia \mathbf{W} un sottospazio di \mathbf{V} . Allora b induce un'applicazione

$$b' : \mathbf{W}^2 \rightarrow \mathbf{K} \quad (146)$$

la quale evidentemente soddisfa ancora le condizioni della definizione, ed è pertanto una forma bilineare su \mathbf{W} . Inoltre b' è simmetrica perché b lo è. Nello stesso modo vediamo che la forma quadratica $q' : \mathbf{W} \rightarrow \mathbf{K}$ associata a b' coincide con la restrizione di q a \mathbf{W} , e pertanto la restrizione di q a \mathbf{W} è ancora una forma quadratica.

Proposizione 5.15. *Siano \mathbf{U} e \mathbf{W} sottospazi vettoriali di uno spazio \mathbf{V} tali che $\mathbf{V} = \mathbf{U} \oplus \mathbf{W}$. Supponiamo assegnate forme bilineari $h : \mathbf{U}^2 \rightarrow \mathbf{K}$, $k : \mathbf{W}^2 \rightarrow \mathbf{K}$. Definiamo la seguente applicazione:*

$$h \oplus k : \mathbf{V}^2 \rightarrow \mathbf{K}, \quad ((\mathbf{u}, \mathbf{w}), (\mathbf{u}', \mathbf{w}')) \mapsto h(\mathbf{u}, \mathbf{u}') + k(\mathbf{w}, \mathbf{w}').$$

*L'applicazione $h \oplus k$ è una forma bilineare, detta **somma diretta** di h e k . Se h e k sono entrambe simmetriche (entrambe alterne) $h \oplus k$ è simmetrica (alterna).*

Proof. Mostriamo che $h \oplus k$ soddisfa i tre assiomi di bilinearità e conserva simmetria/alternanza.

FB1 Siano $(\mathbf{u}_1, \mathbf{w}_1), (\mathbf{u}_2, \mathbf{w}_2), (\mathbf{u}_3, \mathbf{w}_3) \in \mathbf{V}^2$. Allora

$$\begin{aligned} (h \oplus k)((\mathbf{u}_1, \mathbf{w}_1) + (\mathbf{u}_2, \mathbf{w}_2), (\mathbf{u}_3, \mathbf{w}_3)) &= (h \oplus k)((\mathbf{u}_1 + \mathbf{u}_2, \mathbf{w}_1 + \mathbf{w}_2), (\mathbf{u}_3, \mathbf{w}_3)) \\ &= h(\mathbf{u}_1 + \mathbf{u}_2, \mathbf{u}_3) + k(\mathbf{w}_1 + \mathbf{w}_2, \mathbf{w}_3) \\ &= [h(\mathbf{u}_1, \mathbf{u}_3) + h(\mathbf{u}_2, \mathbf{u}_3)] + [k(\mathbf{w}_1, \mathbf{w}_3) + k(\mathbf{w}_2, \mathbf{w}_3)] \\ &= (h \oplus k)((\mathbf{u}_1, \mathbf{w}_1), (\mathbf{u}_3, \mathbf{w}_3)) + (h \oplus k)((\mathbf{u}_2, \mathbf{w}_2), (\mathbf{u}_3, \mathbf{w}_3)). \end{aligned}$$

FB2 Con la stessa procedura, per ogni $(\mathbf{u}_1, \mathbf{w}_1)$ e $(\mathbf{u}_2, \mathbf{w}_2), (\mathbf{u}_3, \mathbf{w}_3)$ si ottiene

$$\begin{aligned} (h \oplus k)((\mathbf{u}_1, \mathbf{w}_1), (\mathbf{u}_2, \mathbf{w}_2) + (\mathbf{u}_3, \mathbf{w}_3)) &= h(\mathbf{u}_1, \mathbf{u}_2 + \mathbf{u}_3) + k(\mathbf{w}_1, \mathbf{w}_2 + \mathbf{w}_3) \\ &= h(\mathbf{u}_1, \mathbf{u}_2) + h(\mathbf{u}_1, \mathbf{u}_3) + k(\mathbf{w}_1, \mathbf{w}_2) + k(\mathbf{w}_1, \mathbf{w}_3) \\ &= (h \oplus k)((\mathbf{u}_1, \mathbf{w}_1), (\mathbf{u}_2, \mathbf{w}_2)) + (h \oplus k)((\mathbf{u}_1, \mathbf{w}_1), (\mathbf{u}_3, \mathbf{w}_3)). \end{aligned}$$

FB3 Per $\lambda \in \mathbf{K}$ e $(\mathbf{u}, \mathbf{w}), (\mathbf{u}', \mathbf{w}') \in \mathbf{V}^2$,

$$\begin{aligned} (h \oplus k)(\lambda(\mathbf{u}, \mathbf{w}), (\mathbf{u}', \mathbf{w}')) &= (h \oplus k)((\lambda\mathbf{u}, \lambda\mathbf{w}), (\mathbf{u}', \mathbf{w}')) \\ &= h(\lambda\mathbf{u}, \mathbf{u}') + k(\lambda\mathbf{w}, \mathbf{w}') \\ &= \lambda h(\mathbf{u}, \mathbf{u}') + \lambda k(\mathbf{w}, \mathbf{w}') = \lambda (h \oplus k)((\mathbf{u}, \mathbf{w}), (\mathbf{u}', \mathbf{w}')). \end{aligned}$$

Analogamente si mostra l'omogeneità nel secondo argomento.

Simmetria Se h e k sono simmetriche, cioè

$$h(\mathbf{u}, \mathbf{u}') = h(\mathbf{u}', \mathbf{u}), \quad k(\mathbf{w}, \mathbf{w}') = k(\mathbf{w}', \mathbf{w}),$$

allora

$$\begin{aligned} (h \oplus k)((\mathbf{u}, \mathbf{w}), (\mathbf{u}', \mathbf{w}')) &= h(\mathbf{u}, \mathbf{u}') + k(\mathbf{w}, \mathbf{w}') \\ &= h(\mathbf{u}', \mathbf{u}) + k(\mathbf{w}', \mathbf{w}) = (h \oplus k)((\mathbf{u}', \mathbf{w}'), (\mathbf{u}, \mathbf{w})). \end{aligned}$$

Alternanza Se h e k sono entrambe alterne, ovvero $h(\mathbf{u}, \mathbf{u}) = 0$ per ogni $\mathbf{u} \in \mathbf{U}$ e analogamente per k , allora

$$(h \oplus k)((\mathbf{u}, \mathbf{w}), (\mathbf{u}, \mathbf{w})) = h(\mathbf{u}, \mathbf{u}) + k(\mathbf{w}, \mathbf{w}) = 0 + 0 = 0,$$

quindi $h \oplus k$ è alterna.

□

Proposizione 5.16. *Siano $b \in \text{Bil}(\mathbf{V})$, $\psi = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una base di \mathbf{V} e $A = (a_{ij})$ la matrice di b rispetto ad ψ . Se $\psi^\sim = \{\eta_1, \dots, \eta_n\}$ è la base di \mathbf{V}^\sim duale di ψ . Allora ${}^t A$ è la matrice che rappresenta l'applicazione lineare*

$$\delta_b : \mathbf{V} \rightarrow \mathbf{V}^\sim$$

rispetto alle basi ψ ed ψ^\sim .

Proof. Per dimostrare l'affermazione precedente è necessario dimostrare che $\forall i \in \{1, \dots, n\}$,

$$\delta_b(\mathbf{e}_i) = \sum_{t=1}^n a_{it} \eta_t. \quad (147)$$

Ma

$$\forall i \in \{1, \dots, n\}, \quad [\delta_b(\mathbf{e}_i)](\mathbf{e}_j) = b(\mathbf{e}_i, \mathbf{e}_j) = a_{ij},$$

mentre

$$\left[\sum_{t=1}^n a_{it} \eta_t \right](\mathbf{e}_j) = \sum_{t=1}^n a_{it} \delta_{tj} = a_{ij}.$$

Poiché assumono gli stessi valori sulla base ψ , il primo e il secondo membro di 147 sono uguali, e l'asserto è dimostrato. \square

Si verifica in modo simile che A è la matrice che rappresenta l'applicazione $\delta'_b : \mathbf{V} \rightarrow \mathbf{V}^\sim$ rispetto alle basi ψ e ψ^\sim .

Forme e piani iperbolici

Definizione 5.15. Sia \mathbf{U} un \mathbb{K} -spazio vettoriale tale che $\dim(\mathbf{U}) = 2$. Supponiamo assegnata su \mathbf{U} una forma bilineare simmetrica h non degenera tale che esista un vettore isotropo rispetto ad h e non nullo. Allora h si dice **forma iperbolica** su \mathbf{U} e la coppia (\mathbf{U}, h) si dice **piano iperbolico**.

Proposizione 5.17. Se (\mathbf{U}, h) è un piano iperbolico, \mathbf{U} possiede una base $\{\mathbf{u}_1, \mathbf{u}_2\}$ formata da due vettori isotropi tali che $h(\mathbf{u}_1, \mathbf{u}_2) = 1$.

Proof. Infatti, per definizione esiste $\mathbf{u}_1 \neq \mathbf{0}$ isotropo. Poiché h è non degenera,

$$\exists \mathbf{v} \in \mathbf{U} : h(\mathbf{u}_1, \mathbf{v}) = 1$$

I vettori \mathbf{u}_1 e \mathbf{v} non sono proporzionali, perché si ha

$$\forall k \in \mathbb{K}, \quad h(\mathbf{u}_1, k\mathbf{u}_1) = kh(\mathbf{u}_1, \mathbf{u}_1) = 0;$$

pertanto \mathbf{u}_1 e \mathbf{v} sono linearmente indipendenti. Prendendo $\mathbf{u}_2 = \mathbf{v} - \frac{h(\mathbf{v}, \mathbf{v})\mathbf{u}_1}{2}$ si ottiene la base richiesta. \square

Base iperbolica

Definizione 5.16. Una base $\{\mathbf{u}_1, \mathbf{u}_2\}$ con tali proprietà è detta **iperbolica**. La matrice che rappresenta h rispetto a una base iperbolica è

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (148)$$

Nota 25. È evidente, viceversa, che se lo spazio vettoriale \mathbf{U} ha una base $\{\mathbf{u}_1, \mathbf{u}_2\}$ tale che la matrice della forma bilineare sia la 148, allora (\mathbf{U}, h) è un piano iperbolico.

Esempio 5.2. La forma bilineare

$$h(\mathbf{x}, \mathbf{y}) = x_1 y_2 + x_2 y_1$$

su \mathbb{K}^2 è iperbolica, e la base canonica è iperbolica.

Esempio 5.3. La forma bilineare

$$k(\mathbf{x}, \mathbf{y}) = \frac{x_1 y_1}{2} - \frac{x_2 y_2}{2}$$

è iperbolica, e una base iperbolica è $\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$. Per la forma k la base canonica è diagonalizzante.

Proposizione 5.18. Siano \mathbf{V} un \mathbb{K} -spazio vettoriale e $b \in \text{Bil}_s(\mathbf{V})$ non degenerare tale che \mathbf{V} contenga un vettore isotropo $\mathbf{u} \neq \mathbf{0}$. Allora esiste un sottospazio \mathbf{U} di \mathbf{V} contenente \mathbf{u} e tale che la coppia $(\mathbf{U}, b_{\mathbf{U}})$ sia un piano iperbolico.

Proof. Poiché b è non degenerare,

$$\exists \mathbf{v} \in \mathbf{V} : b(\mathbf{u}, \mathbf{v}) = 1.$$

Il vettore

$$\mathbf{w} = \mathbf{v} - \frac{b(\mathbf{v}, \mathbf{v})}{2} \mathbf{u}$$

è isotropo e tale che $b(\mathbf{u}, \mathbf{w}) = 1$. Quindi il sottospazio $\mathbf{U} = \langle \mathbf{u}, \mathbf{w} \rangle$ ha le proprietà volute. \square

Scalare rappresentabile mediante forma quadratica

Definizione 5.17. Siano \mathbf{V} uno spazio vettoriale e $q : \mathbf{V} \rightarrow \mathbb{K}$ una forma quadratica. Uno scalare $\alpha \in \mathbb{K}$ si dice **rappresentabile mediante q** se

$$\exists \mathbf{v} \in \mathbf{V} \setminus \{\mathbf{0}\} : q(\mathbf{v}) = \alpha.$$

Proposizione 5.19. Se \mathbf{V} è uno spazio vettoriale complesso e q è non degenerare, allora ogni $\alpha \in \mathbb{C}$ è rappresentabile mediante q .

Proof. Infatti, fissata una base $\psi = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ tale che $b(\mathbf{e}_i) \neq 0$, e detta $\beta \in \mathbb{C}$ una radice quadrata di $\alpha q(\mathbf{e}_1)^{-1}$, si ha

$$q(\beta \mathbf{e}_1) = \beta^2 q(\mathbf{e}_1) = \alpha.$$

\square

Proposizione 5.20. Siano $\mathbb{K} = \mathbb{R}$, $\mathbf{V} = \mathbb{R}^n$, $n \geq 1$, e $q(\mathbf{x}) = \sum_{i=1}^n x_i^2$ la forma quadratica standard. Allora, nessun numero $\alpha \leq 0$ è rappresentabile mediante q . Invece ogni numero reale $\alpha > 0$ lo è.

Proof. La dimostrazione è simile al caso complesso, tenendo presente che α possiede una radice quadrata reale. \square

Nota 26. La nozione di rappresentabilità di uno scalare mediante q è importante se $\mathbb{K} = \mathbb{Q}$. L'insieme degli scalari rappresentabili mediante una data forma quadratica q su uno \mathbb{Q} -spazio vettoriale ha una notevole importanza aritmetica.

Proposizione 5.21. Se (\mathbf{U}, h) è un piano iperbolico e $q : \mathbf{U} \rightarrow \mathbb{K}$ è la forma quadratica associata ad h , allora $q(\mathbf{U} \setminus \{\mathbf{0}\}) = \mathbb{K}$, e quindi ogni elemento di \mathbb{K} è rappresentabile mediante q .

Proof. Infatti, se $(\mathbf{u}_1, \mathbf{u}_2)$ è una base iperbolica di \mathbf{U} e $\alpha \in \mathbb{K}$, si ha

$$q\left(\mathbf{u}_1 + \frac{\alpha}{2} \mathbf{u}_2\right) = \alpha.$$

\square

Dall'osservazione precedente e dalla 5.21 segue subito che se q è una forma quadratica non degenerare su uno spazio vettoriale \mathbf{V} , tale che esista in \mathbf{V} un vettore isotropo non nullo, allora ogni $\alpha \in \mathbb{K}$ è rappresentabile mediante q . In altre parole, se 0 è rappresentabile mediante q , lo è ogni $\alpha \in \mathbb{K}$.

Esempio 5.4. Prendendo $\mathbb{K} = \mathbb{Q}$ e $\mathbf{V} = \mathbb{Q}^2$, otteniamo che ogni numero razionale α può essere espresso nella forma

$$\alpha = \frac{x^2 - y^2}{2}, \quad x, y \in \mathbb{Q}.$$

Sia $b \in \text{Bil}_s(\mathbf{V})$ non degenerare, e sia \mathbf{U} un sottospazio di \mathbf{V} . Il sottospazio ortogonale \mathbf{U}^\perp coincide con l'ortogonale di $\delta_b(\mathbf{U}) \subset \mathbf{V}^\sim$, così com'è stato definito in 114. Supponiamo che \mathbf{V} abbia dimensione finita. Poiché b è non degenerare, δ_b è un isomorfismo, e quindi $\dim(\mathbf{U}) = \dim[\delta_b(\mathbf{U})]$. Dalla 113 segue pertanto che

$$\dim(\mathbf{U}^\perp) = n - \dim(\mathbf{U}). \quad (149)$$

Se \mathbf{U} non contiene vettori isotropi non nulli, si ha $\mathbf{U} \cap \mathbf{U}^\perp = \langle \mathbf{0} \rangle$, e dalla 149 segue allora che $\mathbf{V} = \mathbf{U} \oplus \mathbf{U}^\perp$. Prendendo in particolare $\mathbf{U} = \langle \mathbf{v} \rangle$, dove \mathbf{v} è un vettore non isotropo, si ottiene l'identità dei vettori 149 dimostrata in precedenza.

Cono isotropo

Definizione 5.18. Sia $b : \mathbf{V}^2 \rightarrow \mathbb{K}$ una forma bilineare simmetrica sullo spazio \mathbf{V} , e denotiamo con $\mathbf{I}_b(\mathbf{V}) \subset \mathbf{V}$ l'insieme dei vettori isotropi rispetto a b . $\mathbf{I}_b(\mathbf{V})$ è detto **cono isotropo** di \mathbf{V} (rispetto a b).

Sottospazi isotropi

Definizione 5.19. Un sottospazio \mathbf{U} di \mathbf{V} si dice **isotropo** se $\mathbf{U} \subset \mathbf{I}_b(\mathbf{V})$.

Ovviamente $\langle \mathbf{0} \rangle$ è un sottospazio isotropo di \mathbf{V} , che si dice **banale**.

Se b è degenerare, allora il suo radicale \mathbf{V}^\perp è un sottospazio isotropo non banale.

Se \mathbf{U} è un sottospazio isotropo a $\mathbf{u}_1, \mathbf{u}_2 \in \mathbf{U}$, allora, poichè $\mathbf{u}_1 + \mathbf{u}_2 \in \mathbf{U}$, si ha

$$0 = b(\mathbf{u}_1 + \mathbf{u}_2, \mathbf{u}_1 + \mathbf{u}_2) = 2b(\mathbf{u}_1, \mathbf{u}_2), \quad (150)$$

cioè \mathbf{u}_1 e \mathbf{u}_2 sono ortogonali. Da ciò discende che $\mathbf{U} \subset \mathbf{U}^\perp$. Viceversa, se $\mathbf{U} \subset \mathbf{U}^\perp$, allora è evidente che \mathbf{U} è isotropo.

La forma b si dice **anisotropia** se \mathbf{V} non possiede vettori isotropi non nulli, cioè se $\mathbf{I}_b(\mathbf{V}) = \{\mathbf{0}\}$. Ad esempio la forma bilineare standard su \mathbb{R}^n è anisotropia. Supponiamo b non degenerare, e sia \mathbf{U} un sottospazio isotropo. Allora si ha

$$\dim(\mathbf{U}) \leq \frac{\dim(\mathbf{V})}{2}. \quad (151)$$

Infatti \mathbf{U} isotropo significa che $\mathbf{U} \subset \mathbf{U}^\perp$, e quindi, per la 149, abbiamo

$$\dim(\mathbf{U}) \leq \dim(\mathbf{U}^\perp) = \dim(\mathbf{V}) - \dim(\mathbf{U}),$$

cioè la 151.

6 Diagonalizzazione delle forme quadratiche

Teorema 6.1. *Siano \mathbf{V} un \mathbb{K} -spazio vettoriale con $\dim(\mathbf{V}) = n \geq 1$, e $b \in \text{Bil}_s(\mathbf{V})$. In \mathbf{V} esiste una base diagonalizzante per b . Equivalentemente, ogni matrice simmetrica $A \in M_n(\mathbb{K})$, $n \geq 1$, è congruente a una matrice diagonale.*

Data l'importanza del teorema, ne diamo due dimostrazioni.

Prima dimostrazione

Proof. Procediamo per induzione su $n = \dim(\mathbf{V})$. Se $n = 1$ non c'è niente da dimostrare. Supponiamo dunque $n \geq 2$, e che ogni forma bilineare simmetrica su uno spazio di dimensione minore di n possieda una base diagonalizzante. Se b è la forma bilineare nulla non c'è niente da dimostrare, perché in una qualsiasi base la matrice di b è la matrice nulla, che è diagonale, e quindi ogni base di \mathbf{V} è diagonalizzante. Possiamo dunque supporre che b non sia la forma bilineare nulla, e che quindi

$$\exists \mathbf{v}, \mathbf{w} \in \mathbf{V} : b(\mathbf{v}, \mathbf{w}) \neq 0.$$

Da ciò segue che almeno uno dei tre vettori $\mathbf{v}, \mathbf{w}, \mathbf{v} + \mathbf{w}$ non è isotropo. Infatti se \mathbf{v} e \mathbf{w} sono entrambi isotropi, allora

$$b(\mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w}) = b(\mathbf{v}, \mathbf{v}) + b(\mathbf{w}, \mathbf{w}) + 2b(\mathbf{v}, \mathbf{w}) = 2b(\mathbf{v}, \mathbf{w}) \neq 0.$$

Pertanto

$$\exists \mathbf{e}_1 \in \mathbf{V} : b(\mathbf{e}_1, \mathbf{e}_1) \neq 0.$$

Dalla 140 segue che

$$\mathbf{V} = \langle \mathbf{e}_1 \rangle \oplus \mathbf{e}_1^\perp,$$

in particolare $\dim(\mathbf{e}_1^\perp) = n - 1$. Per l'ipotesi induttiva la forma bilineare b' indotta da b su \mathbf{e}_1^\perp possiede una base diagonalizzante: sia essa $\{\mathbf{e}_2, \dots, \mathbf{e}_n\}$. Allora $\psi = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ è una base di \mathbf{V} : infatti $\mathbf{e}_2, \dots, \mathbf{e}_n$ sono linearmente indipendenti e d'altra parte $\mathbf{e}_1 \notin \langle \mathbf{e}_2, \dots, \mathbf{e}_n \rangle = \mathbf{e}_1^\perp$, sicché $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ sono linearmente indipendenti. Inoltre $b(\mathbf{e}_1, \mathbf{e}_j) = 0$ per ogni $j \in \{2, \dots, n\}$, perché $\mathbf{e}_j \notin \mathbf{e}_1^\perp$. Infine $b(\mathbf{e}_i, \mathbf{e}_j) = b'(\mathbf{e}_i, \mathbf{e}_j) = 0$ per ogni $i \neq j$, $2 \leq i, j \leq n$, perché $\{\mathbf{e}_2, \dots, \mathbf{e}_n\}$ è una base di \mathbf{e}_1^\perp diagonalizzante per b' . Quindi ψ è una base diagonalizzante per b . \square

Seconda dimostrazione (Lagrange)

Proof. Per induzione su $n = \dim(\mathbf{V})$. Se $n = 1$ non c'è niente da dimostrare. Supponiamo $n \geq 2$, e che ogni forma bilineare simmetrica su uno spazio di dimensione minore di n possieda una base diagonalizzante. Scegliamo una base $\phi = [\mathbf{b}_1, \dots, \mathbf{b}_n]$ di \mathbf{V} . Se b è la forma bilineare nulla, allora b è diagonalizzante, e non c'è niente da dimostrare. Se b non è la forma bilineare nulla, allora possiamo ottenere da ϕ una nuova base $\varphi = [\mathbf{c}_1, \dots, \mathbf{c}_n]$ tale che $b(\mathbf{c}_1, \mathbf{c}_1) \neq 0$. Infatti se $b(\mathbf{b}_i, \mathbf{b}_i) \neq 0$ per qualche i , sarà sufficiente scambiare \mathbf{b}_1 con \mathbf{b}_i . Se viceversa $b(\mathbf{b}_i, \mathbf{b}_i) = 0$ per ogni i , allora dev'essere $b(\mathbf{b}_i, \mathbf{b}_j) \neq 0$ per qualche $i \neq j$, scambiando ancora possiamo supporre $b(\mathbf{b}_1, \mathbf{b}_2) \neq 0$, e la nuova base

$$\varphi = \{\mathbf{b}_1 + \mathbf{b}_2, \mathbf{b}_2, \dots, \mathbf{b}_n\}$$

ha la proprietà voluta. Nella base φ la forma quadratica q associata a b ha l'espressione

$$q(\mathbf{v}(y_1, \dots, y_n)) = h_{11}y_1^2 + 2 \sum_{i=2}^n h_{1i}y_1y_i + \sum_{i,j=2}^n h_{ij}y_iy_j, \quad (152)$$

dove $h_{ij} = b(\mathbf{c}_i, \mathbf{c}_j)$. Poiché $h_{11} = b(\mathbf{c}_1, \mathbf{c}_1) \neq 0$, possiamo riscrivere la 152 nel modo seguente:

$$q(\mathbf{v}(y_1, \dots, y_n)) = h_{11} \left(y_1 + \sum_{i=2}^n h_{11}^{-1} h_{1i} y_i \right)^2 + (\text{termini in cui non compare } y_1).$$

Eseguiamo il cambiamento di coordinate

$$z_1 = y_1 + \sum_{i=2}^n h_{11}^{-1} h_{1i} y_i, \quad z_2 = y_2, \quad \dots, \quad z_n = y_n$$

che corrisponde al passaggio dalla base φ alla nuova base

$$\gamma = \{\mathbf{d}_1, \dots, \mathbf{d}_n\} = \{\mathbf{c}_1, -h_{11}^{-1} h_{12} \mathbf{c}_1 + \mathbf{c}_2, -h_{11}^{-1} h_{13} \mathbf{c}_1 + \mathbf{c}_3, \dots, -h_{11}^{-1} h_{1n} \mathbf{c}_1 + \mathbf{c}_n\}.$$

In queste coordinate q ha la forma

$$q(\mathbf{v}(z_1, \dots, z_n)) = h_{11}z_1^2 + q'(z_2, \dots, z_n),$$

dove $q'(z_2, \dots, z_n)$ è un polinomio omogeneo di secondo grado in z_2, \dots, z_n e quindi è una forma quadratica sullo spazio $\langle \mathbf{d}_2, \dots, \mathbf{d}_n \rangle$. Per l'ipotesi induttiva $\langle \mathbf{d}_2, \dots, \mathbf{d}_n \rangle$ possiede una base $\{\mathbf{e}_2, \dots, \mathbf{e}_n\}$ diagonalizzante per q' . Pertanto la base $\{\mathbf{d}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ di \mathbf{V} è una base diagonalizzante per q . \square

Il teorema precedente afferma l'esistenza di una base ψ rispetto alla quale b ha la forma 141, con $a_{11}, a_{22}, \dots, a_{nn} \in \mathbb{K}$ non tutti nulli. Dimosteremo che nei casi $\mathbb{K} = \mathbb{R}, \mathbb{C}$ si possono ottenere risultati più precisi. Iniziamo dal caso $\mathbb{K} = \mathbb{C}$ (considereremo più in generale il caso in cui \mathbb{K} è algebricamente chiuso).

Teorema 6.2. *Supponiamo che \mathbb{K} sia algebricamente chiuso, e siano \mathbf{V} un \mathbb{K} -spazio vettoriale con $\dim(\mathbf{V}) = n \geq 1$ e $b \in \text{Bil}_s(\mathbf{V})$. Esiste una base $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ diagonalizzante per b tale che la matrice di b sia della forma*

$$D = \begin{pmatrix} \mathbf{I}_r & \mathbf{0}_1 \\ \mathbf{0}_2 & \mathbf{0}_3 \end{pmatrix} \quad (153)$$

dove r è il rango di b e $\mathbf{0}_1 \in M_{r,n-r}(\mathbb{K})$, $\mathbf{0}_2 \in M_{n-r,r}(\mathbb{K})$, $\mathbf{0}_3 \in M_{n-r,n-r}(\mathbb{K})$ sono le matrici nulle. Equivalentemente, ogni matrice simmetrica $A \in M_n(\mathbb{K})$ di rango r è congruente alla matrice 153.

Proof. L'equivalenza delle due affermazioni è evidente. Dimosteremo la prima. Per il teorema 6.1 esiste una base $\phi = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ tale che la matrice di b rispetto a ϕ sia diagonale della forma

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Salvo scambiare tra loro $\mathbf{f}_1, \dots, \mathbf{f}_n$, possiamo supporre $a_{11}, a_{22}, \dots, a_{nn}$ diversi da 0, e $a_{r+1,r+1} = \dots = a_{nn} = 0$. Siano $\alpha_1, \alpha_2, \dots, \alpha_r \in \mathbb{K}$ tali che $\alpha_i^2 = a_{ii}$, $i = 1, \dots, r$ (gli α_i esistono perché \mathbb{K} è algebricamente chiuso) e consideriamo i vettori

$$\mathbf{e}_1 = \alpha_1^{-1}\mathbf{f}_1, \dots, \mathbf{e}_r = \alpha_r^{-1}\mathbf{f}_r, \mathbf{e}_{r+1} = \mathbf{f}_{r+1}, \dots, \mathbf{e}_n = \mathbf{f}_n.$$

Ovviamente $\mathbf{e} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ è una base ortogonale. Inoltre

$$\begin{aligned} b(\mathbf{e}_i, \mathbf{e}_i) &= b(\alpha_i^{-1}\mathbf{f}_i, \alpha_i^{-1}\mathbf{f}_i) = \alpha_i^{-2}b(\mathbf{f}_i, \mathbf{f}_i) = \alpha_i^{-2}a_{ii} = 1, \quad i = 1, \dots, r \\ b(\mathbf{e}_i, \mathbf{e}_i) &= b(\mathbf{f}_i, \mathbf{f}_i) = 0 \quad i = r+1, \dots, n. \end{aligned}$$

Quindi \mathbf{e} ha le proprietà volute. \square

Nella dimostrazione del teorema precedente si è utilizzato il fatto che \mathbb{K} è algebricamente chiuso per ottenere l'esistenza degli scalari α_i . Nel caso $\mathbb{K} = \mathbb{R}$ si ottiene un risultato un po' più debole.

Teorema di Sylvester

Teorema 6.3. *Siano \mathbf{V} un \mathbb{R} -spazio vettoriale, $\dim(\mathbf{V}) = n \geq 1$, e $b \in \text{Bil}_s(\mathbf{V})$. Esistono un intero non negativo $p \leq r$, dove r è il rango di b , dipendente solo da b , e una base $\psi = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ di \mathbf{V} tali che rispetto a ψ la forma b abbia la seguente matrice:*

$$\begin{pmatrix} \mathbf{I}_p & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I}_{r-p} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad (154)$$

dove il simbolo $\mathbf{0}$ denota matrici nulle di ordini opportuni. Equivalentemente, ogni matrice simmetrica $A \in M_n(\mathbb{R})$ è congruente a una matrice diagonale della forma 154 in cui $r = r(A)$ e p dipende solo da A .

Proof. L'equivalenza delle due affermazioni è evidente. Dimostriamo quindi la prima. Per il teorema 6.1 esiste una base $\phi = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ di \mathbf{V} tale che

$$\forall \mathbf{v} = \sum_{i=1}^n y_i \mathbf{f}_i \in \mathbf{V}, \quad q(\mathbf{v}) = \sum_{i=1}^n a_{ii} y_i^2$$

Il numero di coefficienti a_{ii} che sono diversi da 0 è uguale al rango r della forma q , e quindi dipende solo da q . Salvo cambiare la numerazione possiamo supporre che i primi r coefficienti siano diversi da zero e che tra essi tutti quelli positivi figurino per primi. Si avrà dunque

$$a_{11} = \alpha_1^2, \dots, a_{pp} = \alpha_p^2, a_{p+1, p+1} = -\alpha_{p+1}^2, \dots, a_{rr} = -\alpha_r^2$$

per un opportuno intero $p \leq r$ e opportuni numeri reali positivi $\alpha_1, \alpha_2, \dots, \alpha_r$. Esattamente come nella dimostrazione del teorema precedente si verifica che rispetto alla base

$$\mathbf{e}_1 = \frac{\mathbf{f}_1}{\alpha_1}, \mathbf{e}_2 = \frac{\mathbf{f}_2}{\alpha_2}, \dots, \mathbf{e}_r = \frac{\mathbf{f}_r}{\alpha_r}, \mathbf{e}_{r+1} = \mathbf{f}_{r+1}, \dots, \mathbf{e}_n = \mathbf{f}_n$$

la matrice di b è la 154, e quindi la forma quadratica q associata a b è

$$\forall \mathbf{v} = \sum_{i=1}^n x_i \mathbf{e}_i \in \mathbf{V}, \quad q(\mathbf{v}) = \sum_{i=1}^p x_i^2 - \sum_{j=p+1}^r x_j^2. \quad (155)$$

Resta da dimostrare che p dipende solo da b , e non dalla base $\psi = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$. Supponiamo dunque che in un'altra base $\varphi = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ la forma b si esprima come

$$\forall \mathbf{v} = \sum_{i=1}^n z_i \mathbf{b}_i \in \mathbf{V}, \quad q(\mathbf{v}) = \sum_{i=1}^t z_i^2 - \sum_{j=t+1}^r z_j^2 \quad (156)$$

per un opportuno intero $t \leq r$. Dobbiamo far vedere che $t = p$. Se $p \neq t$ possiamo supporre che sia $t < p$. Consideriamo i sottospazi

$$\mathbf{S} = \langle \mathbf{e}_1, \dots, \mathbf{e}_p \rangle \quad \mathbf{T} = \langle \mathbf{b}_{t+1}, \dots, \mathbf{b}_n \rangle.$$

Poiché

$$\dim(\mathbf{S}) + \dim(\mathbf{T}) = p + n - t > n,$$

dalla formula di Grassmann segue che $\mathbf{S} \cap \mathbf{T} \neq \langle \mathbf{0} \rangle$, e quindi esiste $\mathbf{v} \in \mathbf{S} \cap \mathbf{T} \setminus \{\mathbf{0}\}$. Si ha quindi

$$\mathbf{v} = \sum_{i=1}^p x_i \mathbf{e}_i = \sum_{j=t+1}^n z_j \mathbf{b}_j. \quad (157)$$

Poiché $\mathbf{v} \neq \mathbf{0}$, dalla 155 segue che

$$q(\mathbf{v}) = \sum_{i=1}^p x_i^2 > 0,$$

mentre dalla 156 si deduce che

$$q(\mathbf{v}) = - \sum_{j=t+1}^r z_j^2 < 0, \quad (158)$$

e cioè una contraddizione. Quindi $p = t$. □

Forma canonica, indici di positività e negatività, segnatura

Definizione 6.1. L'espressione 155 si dice **la forma canonica della forma quadratica q** . Gli interi p ed $r - p$ si dicono rispettivamente **indice di positività** e **indice di negatività**, e la coppia $(p, r - p)$ è detta **segnatura**, di b e di q .

Forme definite e semidefinite positive e negative

Definizione 6.2. Una forma quadratica q sullo spazio vettoriale reale \mathbf{V} si dice:

- **definita positiva** se $q(\mathbf{v}) \geq 0$ per ogni $\mathbf{v} \in \mathbf{V}$;
- **definita negativa** se $q(\mathbf{v}) > 0$ per ogni $\mathbf{v} \neq \mathbf{0}$;
- **semidefinita positiva** se $q(\mathbf{v}) \leq 0$ per ogni $\mathbf{v} \neq \mathbf{0}$;
- **semidefinita negativa** se $q(\mathbf{v}) \leq 0$ per ogni $\mathbf{v} \in \mathbf{V}$.

Evidentemente, se q è definita positiva (negativa) allora è anche semidefinita positiva (negativa).

Forme indefinite

Definizione 6.3. Se non è semidefinita positiva né semidefinita negativa, q si dice **indefinita**.

Analogamente, la terminologia si usa per la forma bilineare polare di q .

Le forme canoniche corrispondenti ai diversi casi, e le relative segnature, sono le seguenti:

	Forma quadratica	Segnatura
definita positiva	$\sum_{i=1}^n x_i^2$	$(n, 0)$
semidefinita positiva	$\sum_{i=1}^r x_i^2 \quad r \leq n$	$(r, 0)$
definita negativa	$-\sum_{i=1}^n x_i^2$	$(0, n)$
semidefinita negativa	$-\sum_{i=1}^r x_i^2 \quad r \leq n$	$(0, r)$
indefinita	$\sum_{i=1}^p x_i^2 - \sum_{i=p+1}^r x_i^2 \quad 0 < p < r \leq n$	$(p, r - p)$

Definizione 6.4. Una matrice simmetrica $A \in M_n(\mathbb{R})$ si dirà **definita positiva** o, rispettivamente **semidefinita positiva**, **definita negativa**, **semidefinita negativa**, **indefinita** se la corrispondente forma quadratica è definita positiva, semidefinita positiva, definita negativa, semidefinita negativa, indefinita.

Dal teorema di Sylvester discende anche che ogni classe di congruenza di matrici simmetriche reali di ordine n contiene una matrice diagonale della forma 154. Si ha dunque una corrispondenza biunivoca tra tali classi di congruenza e le matrici 154, cioè le matrici 154 costituiscono un insieme completo di rappresentanti delle classi di congruenza di matrici simmetriche reali di ordine n .

Proposizione 6.4. Una matrice simmetrica reale di ordine n è definita positiva se e solo se è congruente alla matrice unità \mathbf{I}_n , cioè se e solo se

$$\exists M \in \text{GL}_n(\mathbb{R}) : A = {}^t M \mathbf{I}_n M = {}^t M M.$$

Corollario 13. Una matrice simmetrica $A \in M_n(\mathbb{R})$ è definita positiva se e solo se

$$\exists M \in \text{GL}_n(\mathbb{R}) : A = {}^t M M.$$

Le forme bilineari simmetriche definite positive sugli spazi vettoriali reali sono molto importanti in geometria euclidea, perché permettono di introdurre tutte le nozioni di natura metrica. Anche altri tipi di forme bilineari su spazi vettoriali reali hanno notevole importanza in geometria e in fisica. Un esempio particolare è lo spazio vettoriale reale \mathbb{R}^4 dotato dalla forma quadratica

$$\lambda(x_1, x_2, x_3, x_4) = x_1^2 + x_2^2 + x_3^2 - x_4^2,$$

che si chiama **forma di Minkowski**. λ è non degenera indefinita, di segnatura $(3, 1)$. La coppia (\mathbb{R}^4, λ) è detta **spazio di Minkowski** e ha importanza fondamentale in relatività ristretta.

7 Prodotti scalari

Prodotto scalare

Definizione 7.1. Sia \mathbf{V} un \mathbb{R} -spazio vettoriale. Una forma bilineare simmetrica definita positiva su \mathbf{V} si dice **prodotto scalare**.

Spazio vettoriale euclideo

Definizione 7.2. Se su \mathbf{V} è assegnato un prodotto scalare, \mathbf{V} si dice **spazio vettoriale euclideo**.

Si noti che la definizione non richiede che \mathbf{V} abbia dimensione finita, e infatti esistono esempi di spazi vettoriali euclidei che non hanno dimensione finita. La nostra trattazione sarà però finalizzata al caso finito dimensionale.

Prodotto scalare standard

Definizione 7.3. La forma bilineare standard 134 su \mathbb{R}^n , $n \geq 1$, è un prodotto scalare che chiameremo d'ora in poi **prodotto scalare standard** su \mathbb{R}^n ; se $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ denoteremo il loro prodotto scalare standard con il simbolo $\mathbf{x} \cdot \mathbf{y}$; si ha quindi

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^n x_i y_i = {}^t \mathbf{x} \mathbf{y}.$$

n -spazio vettoriale euclideo

Definizione 7.4. Dotato del prodotto scalare standard, \mathbb{R}^n è uno spazio vettoriale euclideo, denominato n -spazio vettoriale euclideo.

Sia \mathbf{V} uno spazio vettoriale euclideo qualsiasi; denoteremo il prodotto scalare di due vettori $\mathbf{v}, \mathbf{w} \in \mathbf{V}$ con il simbolo

$$\langle \mathbf{v}, \mathbf{w} \rangle,$$

che si legge \mathbf{v} scalare \mathbf{w} .

Teorema 7.1. Se $\mathbf{v}, \mathbf{w} \in \mathbf{V}$, allora

$$\langle \mathbf{v}, \mathbf{w} \rangle^2 \leq \langle \mathbf{v}, \mathbf{v} \rangle \langle \mathbf{w}, \mathbf{w} \rangle \quad (159)$$

e l'uguaglianza vale se e solo se \mathbf{v} e \mathbf{w} sono paralleli.

Proof. Se $\mathbf{w} = \mathbf{0}$ la 159 è ovvia, essendo uguali a 0 entrambi i membri. Possiamo dunque supporre $\mathbf{w} \neq \mathbf{0}$. Per ogni $a, b \in \mathbb{R}$ si ha

$$0 \leq \langle a\mathbf{v} + b\mathbf{w}, a\mathbf{v} + b\mathbf{w} \rangle = a^2 \langle \mathbf{v}, \mathbf{v} \rangle + 2ab \langle \mathbf{v}, \mathbf{w} \rangle + b^2 \langle \mathbf{w}, \mathbf{w} \rangle$$

e l'uguaglianza vale se e solo se $a\mathbf{v} + b\mathbf{w} = \mathbf{0}$, cioè se e solo se \mathbf{v} e \mathbf{w} sono paralleli. Prendendo in particolare

$$a = \langle \mathbf{w}, \mathbf{w} \rangle, \quad b = -\langle \mathbf{v}, \mathbf{w} \rangle,$$

si ottiene

$$0 \leq \langle \mathbf{w}, \mathbf{w} \rangle^2 \langle \mathbf{v}, \mathbf{v} \rangle - 2 \langle \mathbf{w}, \mathbf{w} \rangle \langle \mathbf{v}, \mathbf{w} \rangle^2 + \langle \mathbf{v}, \mathbf{w} \rangle^2 \langle \mathbf{w}, \mathbf{w} \rangle.$$

Poiché $\mathbf{w} \neq \mathbf{0}$, si ha $\langle \mathbf{w}, \mathbf{w} \rangle > 0$, e dividendo il secondo membro per $\langle \mathbf{w}, \mathbf{w} \rangle$ otteniamo

$$0 \leq \langle \mathbf{w}, \mathbf{w} \rangle \langle \mathbf{v}, \mathbf{v} \rangle - \langle \mathbf{v}, \mathbf{w} \rangle^2,$$

cioè la 159. □

Disuguaglianza di Schwartz

Definizione 7.5. La 159 è chiamata **disuguaglianza di Schwarz**.

Norma o lunghezza di un vettore

Definizione 7.6. Se $\mathbf{v} \in \mathbf{V}$, la **norma**, o **lunghezza**, di \mathbf{v} si definisce come

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

Utilizzando la norma, la disuguaglianza di Schwarz si può esprimere nella forma equivalente

$$|\langle \mathbf{v}, \mathbf{w} \rangle| \leq \|\mathbf{v}\| \|\mathbf{w}\|, \quad (160)$$

ottenuta estraendo la radice quadrata dal primo e secondo membro della 159.

Proposizione 7.2. La lunghezza di un vettore gode delle seguenti proprietà:

$$N1 \quad \forall \mathbf{v} \in \mathbf{V}, \quad \|\mathbf{v}\| \geq 0 \quad \wedge \quad (\|\mathbf{v}\| = 0 \iff \mathbf{v} = \mathbf{0});$$

$$N2 \quad \forall r \in \mathbb{R}, \forall \mathbf{v} \in \mathbf{V}, \quad \|r\mathbf{v}\| = |r| \|\mathbf{v}\|;$$

$$N3 \quad \forall \mathbf{v}, \mathbf{w} \in \mathbf{V}, \quad \|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\| \quad \wedge \quad (\|\mathbf{v} + \mathbf{w}\| = \|\mathbf{v}\| + \|\mathbf{w}\| \iff \mathbf{v}, \mathbf{w} \text{ paralleli}).$$

La N1 è una formulazione della proprietà del prodotto scalare di essere definito positivo.

La N2 segue estraendo le radici quadrate dal primo e secondo membro dell'uguaglianza seguente:

$$\|r\mathbf{v}\| = \sqrt{\langle r\mathbf{v}, r\mathbf{v} \rangle} = \sqrt{r^2 \langle \mathbf{v}, \mathbf{v} \rangle} = |r| \|\mathbf{v}\|.$$

La N3 è detta **disuguaglianza triangolare** e si dimostra nel modo seguente.

$$\|\mathbf{v} + \mathbf{w}\|^2 = \langle \mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w} \rangle = \|\mathbf{v}\|^2 + 2\langle \mathbf{v}, \mathbf{w} \rangle + \|\mathbf{w}\|^2$$

Per la 160 si ha

$$2\langle \mathbf{v}, \mathbf{w} \rangle \leq 2|\langle \mathbf{v}, \mathbf{w} \rangle| \leq 2\|\mathbf{v}\| \|\mathbf{w}\|,$$

e dunque

$$\|\mathbf{v}\|^2 + 2\langle \mathbf{v}, \mathbf{w} \rangle + \|\mathbf{w}\|^2 \leq \|\mathbf{v}\|^2 + 2\|\mathbf{v}\| \|\mathbf{w}\| + \|\mathbf{w}\|^2 = (\|\mathbf{v}\| + \|\mathbf{w}\|)^2$$

che è equivalente alla N3.

Per la N1 un vettore \mathbf{v} ha lunghezza 0 se e solo se $\mathbf{v} = \mathbf{0}$.

Versore e normalizzazione di un vettore

Definizione 7.7. Un vettore \mathbf{v} di lunghezza 1 si dice **versore**. Per la N2, se $\mathbf{v} \neq \mathbf{0}$,

$$\frac{\mathbf{v}}{\|\mathbf{v}\|}$$

è un versore parallelo a \mathbf{v} , che si dice ottenuto **normalizzando** \mathbf{v} .

Ricordiamo che due vettori $\mathbf{v}, \mathbf{w} \in \mathbf{V}$ sono detti **ortogonali**, o **perpendicolari**, se

$$\langle \mathbf{v}, \mathbf{w} \rangle = 0.$$

Dalla definizione segue che il vettore $\mathbf{0}$ è perpendicolare ad ogni altro vettore di \mathbf{V} . Poiché il prodotto scalare è definito positivo, $\mathbf{0}$ è l'unico vettore perpendicolare a sé stesso.

Insieme ortogonale di vettori

Definizione 7.8. Un insieme finito $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_t\}$ di vettori non nulli di \mathbf{V} è un **insieme ortogonale di vettori** se i suoi elementi sono a due a due ortogonali, cioè se

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0 \quad \forall i \neq j, 1 \leq i, j \leq t.$$

Insieme ortonormale di vettori

Definizione 7.9. $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_t\}$ si dice **insieme ortonormale di vettori** se è un insieme ortogonale e se i suoi elementi sono versori. Se $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_t\}$ è un insieme ortogonale, normalizzando i suoi elementi si ottiene un insieme ortonormale:

$$\left\{ \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|}, \dots, \frac{\mathbf{v}_t}{\|\mathbf{v}_t\|} \right\}.$$

Base ortogonale (ortonormale)

Definizione 7.10. Se \mathbf{V} ha dimensione finita, una **base ortogonale** (una **base ortonormale**) di \mathbf{V} è una base $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ che è un insieme ortogonale (un insieme ortonormale) di vettori.

Proposizione 7.3. Se $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_t\}$ è un insieme ortogonale di vettori, allora $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_t$ sono linearmente indipendenti. In particolare, se \mathbf{V} ha dimensione finita, un insieme ortogonale di $n = \dim(\mathbf{V})$ vettori è una base ortogonale di \mathbf{V} .

Proof. Se $\sum_{i=1}^t a_i \mathbf{v}_i = 0$, allora, per ogni $i \in \{1, \dots, t\}$,

$$0 = \langle \mathbf{v}_i, \sum_{j=1}^t a_j \mathbf{v}_j \rangle = \sum_{j=1}^t a_j \langle \mathbf{v}_i, \mathbf{v}_j \rangle = a_i \langle \mathbf{v}_i, \mathbf{v}_i \rangle.$$

Poiché $\langle \mathbf{v}_i, \mathbf{v}_i \rangle > 0$, dev'essere $a_i = 0$. Quindi $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_t$ sono linearmente indipendenti. □

Dal teorema 6.1 discende che \mathbf{V} possiede basi ortogonali se ha dimensione finita positiva; normalizzando (gli elementi di) una base ortogonale si ottiene una base ortonormale, e quindi \mathbf{V} possiede anche basi ortonormali. Questo fatto segue anche direttamente dal teorema di Sylvester, visto che il prodotto scalare ha segnatura $(n, 0)$.

Se $\psi = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ è una base ortonormale di \mathbf{V} , allora la matrice che rappresenta \langle, \rangle rispetto alla base ψ è \mathbf{I}_n . Pertanto

$$\forall \mathbf{v} = \sum_{i=1}^n x_i \mathbf{e}_i, \mathbf{w} = \sum_{i=1}^n y_i \mathbf{e}_i \in \mathbf{V}, \quad \langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^n x_i y_i = \mathbf{x} \cdot \mathbf{y}.$$

In altre parole in una base ortonormale il prodotto scalare di due vettori si esprime come il prodotto scalare standard dei vettori colonna delle loro coordinate. Nella pratica è conveniente utilizzare basi ortonormali anziché basi qualunque perché i calcoli con le coordinate dei vettori sono notevolmente più semplici.

Il risultato seguente descrive la matrice di un cambiamento di base ortonormale.

Proposizione 7.4. Siano $\psi = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}, \phi = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ due basi dello spazio vettoriale euclideo \mathbf{V} , e supponiamo che ψ sia ortonormale. La base ϕ è ortonormale se e solo se la matrice $M_{\psi, \phi}(\mathbf{1}_{\mathbf{V}})$ del cambiamento di coordinate da ϕ a ψ è una matrice ortogonale.

Proof. Le colonne della matrice $M = M_{\psi, \phi}(\mathbf{1}_V)$ sono le coordinate dei vettori $\mathbf{f}_1, \dots, \mathbf{f}_n$ rispetto ad ψ . Pertanto ϕ è ortonormale se e solo se

$$M_{(i)} \cdot M_{(j)} = \langle \mathbf{f}_i, \mathbf{f}_j \rangle = \delta_{ij}. \quad (161)$$

Le 161 esprimono precisamente la condizione ${}^tMM = \mathbf{I}_n$, cioè $M \in O(n)$. \square

Descriviamo ora un metodo molto semplice per costruire insiemi ortogonali e trovare esplicitamente basi ortogonali (e quindi basi ortonormali), il cosiddetto **procedimento di ortogonalizzazione di Gram-Schmidt**. Tale procedimento è basato sulla nozione di coefficienti di Fourier. Ricordiamo che, se \mathbf{v} è un vettore non nullo di V , per ogni $\mathbf{w} \in V$ il coefficiente di Fourier di \mathbf{w} rispetto a \mathbf{v} è lo scalare

$$a_{\mathbf{v}}(\mathbf{w}) = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}.$$

Proiezione ortogonale

Definizione 7.11. Definiamo la **proiezione ortogonale di \mathbf{w} nella direzione di \mathbf{v}** come il vettore

$$a_{\mathbf{v}}(\mathbf{w}) \mathbf{v}.$$

La terminologia è giustificata dal fatto che

$$\langle \mathbf{w} - a_{\mathbf{v}}(\mathbf{w}) \mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle - a_{\mathbf{v}}(\mathbf{w}) \langle \mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle - \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \langle \mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle - \langle \mathbf{w}, \mathbf{v} \rangle = 0,$$

cioè $\mathbf{w} - a_{\mathbf{v}}(\mathbf{w}) \mathbf{v}$ è perpendicolare a \mathbf{v} .

Caso 8. Nel caso particolare in cui \mathbf{v} è un versore, per ogni $\mathbf{w} \in V$ il coefficiente di Fourier di \mathbf{w} rispetto a \mathbf{v} è:

$$a_{\mathbf{v}}(\mathbf{w}) = \langle \mathbf{v}, \mathbf{w} \rangle;$$

la proiezione ortogonale di \mathbf{w} nella direzione di \mathbf{v} è $\langle \mathbf{v}, \mathbf{w} \rangle \mathbf{v}$ e $\mathbf{w} - \langle \mathbf{v}, \mathbf{w} \rangle \mathbf{v}$ è ortogonale a \mathbf{v} .

Teorema di ortogonalizzazione di Gram-Schmidt

Teorema 7.5. Sia $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots$ una successione (finita o infinita) di vettori dello spazio euclideo V . Allora:

1. Esiste una successione (corrispondentemente finita dello stesso numero di elementi o infinita) $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \dots$ di vettori di V tale che per ogni intero $k \geq 1$ si abbia:

$$(a) \quad \langle \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \rangle = \langle \mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k \rangle;$$

(b) i vettori $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k$ sono a due a due ortogonali.

2. Se $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \dots$ è un'altra successione che soddisfa le condizioni (a) e (b) per ogni $k \geq 1$, allora esistono scalari non nulli c_1, c_2, \dots tali che $\mathbf{u}_k = c_k \mathbf{w}_k$, $k = 1, 2, \dots$

Proof.

1. Costruiremo gli elementi $\mathbf{w}_1, \mathbf{w}_2, \dots$ per induzione su k . Prendiamo $\mathbf{w}_1 = \mathbf{v}_1$, che evidentemente soddisfa (a) e (b) per $k = 1$. Sia $t \geq 1$ e supponiamo di aver costruito $\mathbf{w}_1, \dots, \mathbf{w}_t$ soddisfacenti alle condizioni (a) e (b) per $k = t$. Definiamo

$$\mathbf{w}_{t+1} = \mathbf{v}_{t+1} - \sum_{i=1}^t a_{\mathbf{w}_i}(\mathbf{v}_{t+1}) \mathbf{w}_i,$$

dove \sum' denota la somma estesa ai soli indici i tali che $\mathbf{w}_i \neq 0$. Dalla definizione segue che \mathbf{v}_{t+1} è combinazione lineare di $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{t+1}$ e, per l'ipotesi induttiva i vettori $\mathbf{v}_1, \dots, \mathbf{v}_t$ sono combinazioni lineari di $\mathbf{w}_1, \dots, \mathbf{w}_t$. Quindi

$$\langle \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{t+1} \rangle \subset \langle \mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{t+1} \rangle.$$

D'altra parte, sempre per la definizione di \mathbf{w}_{t+1} , abbiamo che \mathbf{w}_{t+1} è combinazione lineare di $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_t, \mathbf{v}_{t+1}$. Per l'ipotesi induttiva i vettori $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_t$ sono combinazioni lineari di $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_t$ e quindi \mathbf{w}_{t+1} è combinazione lineare di $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{t+1}$; dunque si ha anche

$$\langle \mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{t+1} \rangle \subset \langle \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{t+1} \rangle.$$

Pertanto la (a) è soddisfatta per $k = t + 1$. Si ha poi, per ogni $i \in \{1, \dots, t\}$ tale che $\mathbf{w}_i \neq \mathbf{0}$:

$$\begin{aligned} \langle \mathbf{w}_{t+1}, \mathbf{w}_i \rangle &= \langle \mathbf{v}_{t+1} - \sum_{j=1}^t a_{\mathbf{w}_j}(\mathbf{v}_{t+1}) \mathbf{w}_j, \mathbf{w}_i \rangle = \langle \mathbf{v}_{t+1}, \mathbf{w}_i \rangle - a_{\mathbf{w}_i}(\mathbf{v}_{t+1}) \langle \mathbf{w}_i, \mathbf{w}_i \rangle \\ &= \langle \mathbf{v}_{t+1}, \mathbf{w}_i \rangle - \langle \mathbf{v}_{t+1}, \mathbf{w}_i \rangle = 0. \end{aligned}$$

Essendo per l'ipotesi induttiva anche

$$\langle \mathbf{w}_j, \mathbf{w}_i \rangle = 0$$

per ogni $i \neq j$, $1 \leq i, j \leq t$, la (b) è soddisfatta per $k = t + 1$. Ciò dimostra la parte (1) del teorema.

2. Procediamo per induzione su k . Per $k = 1$ la conclusione è ovvia. Supponiamo $t \geq 1$ e che esistano c_1, c_2, \dots, c_t tali che $\mathbf{u}_k = c_k \mathbf{w}_k$ per ogni $k \leq t$. Per la (a) si ha

$$\mathbf{u}_{t+1} = \mathbf{z} + c_{t+1} \mathbf{w}_{t+1},$$

per qualche $\mathbf{z} \in \langle \mathbf{w}_1, \dots, \mathbf{w}_t \rangle$, $c_{t+1} \in \mathbb{R}$. Poiché, per la (b), sia \mathbf{u}_{t+1} che \mathbf{w}_{t+1} sono ortogonali a \mathbf{z} , anche $\mathbf{u}_{t+1} - c_{t+1} \mathbf{w}_{t+1} = \mathbf{z}$ è ortogonale a \mathbf{z} , cioè \mathbf{z} è ortogonale a sé stesso; quindi $\mathbf{z} = \mathbf{0}$.

□

Nota 27. Si noti che la (b) del teorema 162 non afferma che $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k\}$ è un insieme ortogonale di vettori, perché può accadere che qualcuno dei vettori $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k$ sia uguale a $\mathbf{0}$. Se $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k$ sono tutti non nulli, allora $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k\}$ è un insieme ortogonale di vettori.

Se \mathbf{V} ha dimensione finita e $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ è una base di \mathbf{V} , il procedimento di Gram-Schmidt applicato ai vettori $\mathbf{v}_1, \dots, \mathbf{v}_n$ fornisce vettori $\mathbf{w}_1, \dots, \mathbf{w}_n$ a due a due ortogonali, che generano \mathbf{V} , e quindi tutti diversi da $\mathbf{0}$. Dalla proposizione 7.3 segue che $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ è una base ortogonale di \mathbf{V} . Per ottenere una base ortonormale non resta che normalizzare $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$.

Esempio 7.1. Sia \mathbf{V} uno spazio vettoriale euclideo, $\dim(\mathbf{V}) = 4$, e sia $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ una base ortonormale di \mathbf{V} . Consideriamo i vettori

$$\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_4 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Il procedimento di Gram-Schmidt applicato a $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ dà

$$\mathbf{w}_1 = \mathbf{v}_1$$

$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{\langle \mathbf{w}_1, \mathbf{v}_2 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 = \mathbf{v}_2 - \frac{2}{2} \mathbf{w}_1 = 2\mathbf{e}_1$$

$$\mathbf{w}_3 = \mathbf{v}_3 - \frac{\langle \mathbf{w}_1, \mathbf{v}_3 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 - \frac{\langle \mathbf{w}_2, \mathbf{v}_3 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} \mathbf{w}_2 = -\frac{\mathbf{e}_2}{2} + \frac{\mathbf{e}_4}{2}$$

$$\mathbf{w}_4 = \mathbf{v}_4 - \frac{\langle \mathbf{w}_1, \mathbf{v}_4 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 - \frac{\langle \mathbf{w}_2, \mathbf{v}_4 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} \mathbf{w}_2 - \frac{\langle \mathbf{w}_3, \mathbf{v}_4 \rangle}{\langle \mathbf{w}_3, \mathbf{w}_3 \rangle} \mathbf{w}_3 = \mathbf{v}_4$$

I vettori $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4$ costituiscono una base ortogonale di \mathbf{V} . Normalizzando questi vettori si ottiene la base ortonormale

$$\left\{ \frac{\mathbf{e}_1}{\sqrt{2}} + \frac{\mathbf{e}_2}{\sqrt{2}}, \mathbf{e}_1, -\frac{\mathbf{e}_1}{\sqrt{2}} + \frac{\mathbf{e}_2}{\sqrt{2}}, \mathbf{e}_3 \right\}.$$

Esempio 7.2. Consideriamo \mathbb{R}^4 con il prodotto scalare:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \frac{x_1 y_1}{2} + \frac{x_2 y_2}{2} + x_3 y_3 + x_4 y_4.$$

Applicando il procedimento di Gram-Schmidt ai vettori

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} -1 \\ -1 \\ -1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

otteniamo

$$\mathbf{w}_1 = \mathbf{v}_1$$

$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{\langle \mathbf{w}_1, \mathbf{v}_2 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 = \dots = \begin{pmatrix} 4/3 \\ 4/3 \\ 2/3 \\ 2/3 \end{pmatrix}$$

$$\mathbf{w}_3 = \mathbf{v}_3 - \frac{\langle \mathbf{w}_1, \mathbf{v}_3 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 - \frac{\langle \mathbf{w}_2, \mathbf{v}_3 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} \mathbf{w}_2 = \dots = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}$$

$$\mathbf{w}_4 = \mathbf{v}_4 - \frac{\langle \mathbf{w}_1, \mathbf{v}_4 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} \mathbf{w}_1 - \frac{\langle \mathbf{w}_2, \mathbf{v}_4 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} \mathbf{w}_2 - \frac{\langle \mathbf{w}_3, \mathbf{v}_4 \rangle}{\langle \mathbf{w}_3, \mathbf{w}_3 \rangle} \mathbf{w}_3 = \dots = \begin{pmatrix} 1/2 \\ -1/2 \\ 0 \\ 0 \end{pmatrix}$$

Quindi

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 4/3 \\ 4/3 \\ 2/3 \\ 2/3 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1/2 \\ -1/2 \\ 0 \\ 0 \end{pmatrix} \right\};$$

è una base di \mathbb{R}^4 ortogonale rispetto a \langle, \rangle .

Proposizione 7.6. Sia \mathbf{V} uno spazio vettoriale euclideo e $\mathbf{W} \neq \langle \mathbf{0} \rangle$ un suo sottospazio vettoriale di dimensione finita. Allora ogni elemento $\mathbf{v} \in \mathbf{V}$ si può esprimere in modo unico come

$$\mathbf{v} = \mathbf{w} + \mathbf{w}', \quad (162)$$

dove $\mathbf{w} \in \mathbf{W}$ e $\mathbf{w}' \in \mathbf{W}^\perp$. Quindi si ha

$$\mathbf{V} = \mathbf{W} \oplus \mathbf{W}^\perp.$$

Proof. Sia $\{\mathbf{e}_1, \dots, \mathbf{e}_t\}$ una base ortonormale di \mathbf{W} . Se $\mathbf{v} \in \mathbf{V}$, poniamo

$$\mathbf{w} = \sum_{i=1}^t \langle \mathbf{v}, \mathbf{e}_i \rangle \mathbf{e}_i, \quad \mathbf{w}' = \mathbf{v} - \mathbf{w}.$$

Evidentemente $\mathbf{w} \in \mathbf{W}$. Inoltre:

$$\forall i \in \{1, \dots, t\}, \quad \langle \mathbf{w}', \mathbf{e}_i \rangle = \langle \mathbf{v}, \mathbf{e}_i \rangle - \langle \mathbf{w}, \mathbf{e}_i \rangle = \langle \mathbf{v}, \mathbf{e}_i \rangle - \langle \mathbf{v}, \mathbf{e}_i \rangle \langle \mathbf{e}_i, \mathbf{e}_i \rangle = 0,$$

e quindi \mathbf{w}' , essendo perpendicolare a $\mathbf{e}_1, \dots, \mathbf{e}_t$, è perpendicolare ad ogni loro combinazione lineare, cioè $\mathbf{w}' \in \mathbf{W}^\perp$. Quindi $\mathbf{v} = \mathbf{w} + \mathbf{w}'$ si esprime nella forma voluta. Se $\mathbf{v} = \mathbf{u} + \mathbf{u}'$ per qualche $\mathbf{u} \in \mathbf{W}$, $\mathbf{u}' \in \mathbf{W}^\perp$, allora

$$\mathbf{w} - \mathbf{u} = \mathbf{u}' - \mathbf{w}' \in \mathbf{W} \cap \mathbf{W}^\perp$$

e quindi $\mathbf{w} - \mathbf{u}$, essendo perpendicolare ad ogni elemento di \mathbf{W} , è anche ortogonale a sé stesso. Ma allora

$$\mathbf{w} - \mathbf{u} = \mathbf{0} = \mathbf{u}' - \mathbf{w}',$$

cioè $\mathbf{w} = \mathbf{u}$ e $\mathbf{w}' = \mathbf{u}'$. □

Proiezione ortogonale di un vettore su un sottospazio

Definizione 7.12. L'elemento \mathbf{w} che compare al secondo membro della 162 si dice **proiezione ortogonale di \mathbf{v} sul sottospazio \mathbf{W}** .

Identità Pitagorica

Definizione 7.13. Esplicitando l'identità $\|\mathbf{v}\|^2 = \langle \mathbf{w} + \mathbf{w}', \mathbf{w} + \mathbf{w}' \rangle$ si ottiene immediatamente la seguente:

$$\|\mathbf{v}\|^2 = \|\mathbf{w}\|^2 + \|\mathbf{w}'\|^2,$$

che è detta **identità pitagorica**.

La disuguaglianza di Schwarz 159 permette di introdurre in uno spazio vettoriale euclideo \mathbf{V} la nozione di *angolo convesso di due vettori non nulli* nel modo seguente.

Siano $\mathbf{v}, \mathbf{w} \in \mathbf{V} \setminus \{\mathbf{0}\}$. La 160 si può anche scrivere nella forma equivalente:

$$-\|\mathbf{v}\|\|\mathbf{w}\| \leq \langle \mathbf{v}, \mathbf{w} \rangle \leq \|\mathbf{v}\|\|\mathbf{w}\|$$

dalla quale, dividendo per $\|\mathbf{v}\|\|\mathbf{w}\|$, si ottiene

$$-1 \leq \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{v}\|\|\mathbf{w}\|} \leq 1. \quad (163)$$

Dalla 163 e dalle proprietà della funzione coseno discende che

$$\exists! \theta \in [0, \pi] \subseteq \mathbb{R} : \cos \theta = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{v}\|\|\mathbf{w}\|}.$$

Potremo quindi scrivere

$$\langle \mathbf{v}, \mathbf{w} \rangle = \|\mathbf{v}\|\|\mathbf{w}\| \cos \theta. \quad (164)$$

Chiameremo

$$\theta = \arccos \left(\frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{v}\|\|\mathbf{w}\|} \right) \quad (165)$$

l'angolo convesso tra i vettori \mathbf{v} e \mathbf{w} oppure l'angolo non orientato formato dai vettori \mathbf{v} e \mathbf{w} .

Nota 28. L'angolo convesso formato dai vettori \mathbf{v} e \mathbf{w} non dipende dall'ordine in cui essi sono presi. Segue inoltre dalla definizione che due vettori non nulli \mathbf{v}, \mathbf{w} sono perpendicolari se e solo se il loro angolo convesso è $\frac{\pi}{2}$ e che avremo $\cos \theta = \langle \mathbf{v}, \mathbf{w} \rangle$ se \mathbf{v} e \mathbf{w} sono versori.

Le funzioni trigonometriche e le loro inverse si possono definire in modo puramente analitico, cosicché la definizione di angolo convesso di due vettori di uno spazio vettoriale euclideo qualunque data dalla 165 è indipendente dalla geometria euclidea elementare. L'angolo convesso così definito è un numero reale compreso tra 0 e π , e non ha il significato geometrico che tale nozione possiede nel piano euclideo ordinario. La relazione esistente tra questa definizione e il concetto di angolo che si dà nella geometria euclidea elementare è la seguente.

Sia \mathbf{V} lo spazio dei vettori geometrici del piano ordinario, in cui supponiamo introdotta un'unità di misura per le lunghezze dei segmenti. Per ogni $\mathbf{v} \in \mathbf{V}$ la **lunghezza di \mathbf{v}** , denotata con $|\mathbf{v}|$, si definisce come la lunghezza di uno qualsiasi dei rappresentanti di \mathbf{v} . Se $\mathbf{v}, \mathbf{w} \in \mathbf{V}$ sono entrambi non nulli, **l'angolo convesso (o non orientato) tra \mathbf{v} e \mathbf{w}** si definisce come l'angolo convesso θ ($0 \leq \theta \leq \pi$ se espresso in radianti) formato da rappresentanti di \mathbf{v} e \mathbf{w} applicati in uno stesso punto (qui stiamo utilizzando la definizione di angolo della geometria euclidea elementare). Queste definizioni sono state date senza utilizzare alcun prodotto scalare su \mathbf{V} . È possibile però introdurre in \mathbf{V} un prodotto scalare in modo che le nozioni di lunghezza di un vettore e di angolo convesso tra due vettori non nulli, che si ottengono utilizzando il prodotto scalare, coincidano con quelle che abbiamo appena introdotto geometricamente.

Proposizione 7.7. Per ogni $\mathbf{v}, \mathbf{w} \in \mathbf{V}$ poniamo

$$\mathbf{v} \times \mathbf{w} = \begin{cases} |\mathbf{v}|\|\mathbf{w}\| \cos \theta & \mathbf{v} \neq \mathbf{0} \neq \mathbf{w}, \\ 0 & \text{altrimenti.} \end{cases} \quad (166)$$

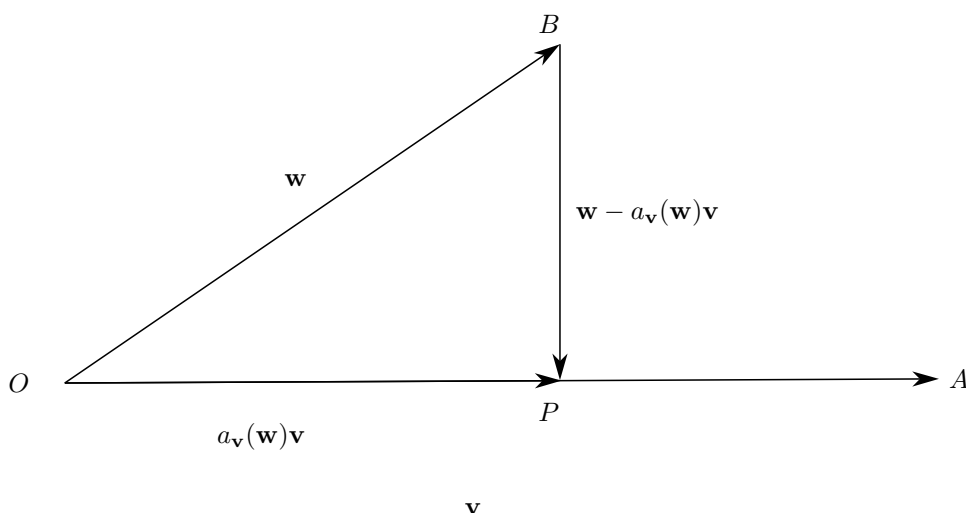
Allora \times definito dalla 166 è un prodotto scalare in \mathbf{V} .

Proof.

□

Si osservi che se \mathbf{v} e \mathbf{w} sono versori allora il loro prodotto scalare coincide con il coseno dell'angolo da essi formato. Inoltre la norma $\|\mathbf{v}\|$ di un vettore \mathbf{v} coincide con la sua lunghezza $|\mathbf{v}|$. Se due vettori \mathbf{v}, \mathbf{w} soddisfano $\mathbf{v} \times \mathbf{w} = 0$ e nessuno dei due è $\mathbf{0}$, allora $\cos \theta = 0$, cioè il loro angolo è $\frac{\pi}{2}$; quindi il concetto di perpendicolarità di vettori così come viene definito dal prodotto scalare 166 coincide con quello usuale.

Siano $\mathbf{v} = \overrightarrow{OA}$, $\mathbf{w} = \overrightarrow{OB} \in \mathbf{V}$, $\mathbf{v} \neq \mathbf{0}$. Sia P il piede della perpendicolare condotta da B sulla retta contenente i punti O e A . È immediato verificare che il coefficiente di Fourier $a_{\mathbf{v}}(\mathbf{w})$ è il rapporto tra la lunghezza di \overrightarrow{OP} e quella di \overrightarrow{OA} , preso con il segno $+$ o $-$ a seconda che \overrightarrow{OP} e \overrightarrow{OA} siano orientati concordemente o discordemente. La proiezione ortogonale di \mathbf{w} nella direzione di \mathbf{v} è il vettore $\overrightarrow{OP} = a_{\mathbf{v}}(\mathbf{w})\mathbf{v}$, e $\mathbf{w} - a_{\mathbf{v}}(\mathbf{w})\mathbf{v} = \overrightarrow{PB}$, che è perpendicolare a \mathbf{v} .



Consideriamo una base ortonormale $\{\mathbf{i}, \mathbf{j}\}$ di \mathbf{V} . Se $\mathbf{v} = x_1\mathbf{i} + x_2\mathbf{j}$, l'identità

$$|\mathbf{v}|^2 = x_1^2 + x_2^2$$

è nient'altro che il teorema di Pitagora.

Notiamo anche che la disuguaglianza di Schwarz

$$|\mathbf{v} \times \mathbf{w}| \leq |\mathbf{v}||\mathbf{w}|$$

segue immediatamente dal fatto che $|\cos \theta| \leq 1$.

Se si considera lo spazio vettoriale \mathbf{V} dei vettori geometrici dello spazio ordinario in cui sia stata introdotta un'unità di misura dei segmenti, la 166 definisce anche in questo caso un prodotto scalare in \mathbf{V} , per il quale valgono considerazioni del tutto simili a quelle fatte nel caso precedente. **Da fare con annesse visualizzazioni**

Per le esigenze dell'algebra lineare e della geometria euclidea non è sufficiente disporre del concetto di angolo non orientato di due vettori, dato dalla 165, è spesso necessaria l'introduzione di una nozione di **angolo orientato**. La definizione di angolo orientato, come nel caso precedente, sarà data ricorrendo alle proprietà dei numeri reali nel modo seguente.

Congruenza modulo 2π

Definizione 7.14. Definiamo in \mathbb{R} la seguente relazione: diciamo che $\theta, \varphi \in \mathbb{R}$ **sono congrui modulo 2π** , oppure che θ è congruo a φ modulo 2π , e scriveremo

$$\theta \equiv \varphi \pmod{2\pi}$$

se e solo se esiste $k \in \mathbb{Z}$ tale che

$$\theta - \varphi = 2k\pi.$$

Proposizione 7.8. *La congruenza modulo 2π è una relazione di equivalenza su \mathbb{R} .*

Proof.

- *Riflessività:* per ogni $\theta \in \mathbb{R}$ si ha

$$\theta - \theta = 0 = 2 \cdot 0 \cdot \pi,$$

quindi $\theta \equiv \theta \pmod{2\pi}$.

- *Simmetria:* se $\theta, \varphi \in \mathbb{R}$ e

$$\theta - \varphi = 2k\pi$$

per qualche $k \in \mathbb{Z}$, allora

$$\varphi - \theta = -2k\pi = 2(-k)\pi,$$

e poiché $-k \in \mathbb{Z}$, segue che $\varphi \equiv \theta \pmod{2\pi}$.

- *Transitività:* se $\theta, \varphi, \psi \in \mathbb{R}$ e

$$\theta - \varphi = 2k_1\pi, \quad \varphi - \psi = 2k_2\pi$$

con $k_1, k_2 \in \mathbb{Z}$, allora sommando membro a membro otteniamo

$$(\theta - \varphi) + (\varphi - \psi) = \theta - \psi = 2k_1\pi + 2k_2\pi = 2(k_1 + k_2)\pi,$$

e quindi $\theta \equiv \psi \pmod{2\pi}$.

□

Angoli orientati

Definizione 7.15. Le classi di congruenza modulo 2π di un angolo θ verranno chiamate **angoli orientati**, o semplicemente **angoli**, e si indicano con

$$[\theta]_{2\pi} = \{\theta + 2k\pi : k \in \mathbb{Z}\}. \quad (167)$$

Poiché le classi di congruenza costituiscono una partizione di \mathbb{R} , ogni numero reale θ individua un angolo 167 cui appartiene e uno solo. Gli elementi di un angolo sono le sue **determinazioni**; per definizione, due determinazioni di uno stesso angolo differiscono per un multiplo intero di 2π . Ogni angolo possiede un'unica determinazione θ_0 , tale che $0 \leq \theta_0 < 2\pi$, la cosiddetta **determinazione principale**. Con abuso di linguaggio identificheremo spesso un angolo con una sua determinazione, così che un numero reale, sottintendendo che due numeri reali rappresentano lo stesso angolo se e solo se la loro differenza è un multiplo intero di 2π .

Poiché ogni angolo possiede un'unica determinazione principale, si ha una corrispondenza biunivoca tra l'insieme \mathcal{D} di tutti gli angoli e l'intervallo chiuso a sinistra $[0, 2\pi)$.

In precedenza abbiamo definito un angolo convesso come un numero reale nell'intervallo $[0, \pi]$. Ad ogni angolo $\{\theta + 2k\pi : k \in \mathbb{Z}\}$ si può associare il numero reale $\min\{|\theta + 2k\pi| : k \in \mathbb{Z}\}$, che, è immediato verificarlo, appartiene a $[0, \pi]$, e si dice **l'angolo convesso associato**.

Gli angoli si possono sommare tra loro sommandone le determinazioni: se $[\theta]_{2\pi}, [\eta]_{2\pi}$ sono due angoli, la loro somma è l'angolo

$$[\theta + \eta]_{2\pi} = \{\theta + \eta + 2k\pi : k \in \mathbb{Z}\}$$

Rispetto all'operazione di somma gli angoli costituiscono un gruppo \mathcal{D} , il cui elemento neutro è l'angolo $[0]_{2\pi}$, e in cui l'opposto dell'angolo $[\theta]_{2\pi}$ è $[-\theta]_{2\pi}$.

Per ogni $\theta \in \mathbb{R}$ consideriamo la matrice

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

Dall'identità $(\cos \theta)^2 + (\sin \theta)^2 = 1$ segue che $R_\theta \in \text{SO}(2)$. Dalle proprietà elementari delle funzioni trigonometriche si deduce che

$$\forall (a, b) \in \mathbb{R}^2 : a^2 + b^2 = 1, \exists! \theta \in \mathbb{R} : \begin{cases} a = \cos \theta \\ b = \sin \theta \end{cases},$$

e quindi, poichè ogni matrice in $\text{SO}(2)$ è della forma 1, al variare di $\theta \in \mathbb{R}$ la matrice R_θ descrive tutto $\text{SO}(2)$. Infine, essendo le funzioni coseno e seno periodiche di periodo 2π , si ha $R_\theta = R_\varphi \iff \theta, \varphi$ sono due determinazioni dello stesso angolo. Da ciò segue che associando ad ogni angolo la matrice R_θ dove θ è una qualunque determinazione dell'angolo, si ottiene una corrispondenza biunivoca di \mathcal{D} su $\text{SO}(2)$. È facile verificare che questa corrispondenza è un'isomorfismo di gruppi. Allo scopo è sufficiente osservare che, se $\varphi, \theta \in \mathbb{R}$, allora

$$R_\varphi R_\theta = R_{\varphi+\theta},$$

cioè:

$$\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} = \begin{pmatrix} \cos(\varphi + \theta) & -\sin(\varphi + \theta) \\ \sin(\varphi + \theta) & \cos(\varphi + \theta) \end{pmatrix}. \quad (168)$$

Consideriamo ora uno spazio vettoriale euclideo \mathbf{V} di dimensione 2, in cui sia fissata una base ortonormale $\{\mathbf{i}, \mathbf{j}\}$. Siano $\mathbf{u}(u_1, u_2)$, $\mathbf{v}(v_1, v_2)$ due vettori, e siano $\varphi, \psi \in \mathbb{R}$ tali che

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \quad \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} \cos \psi \\ \sin \psi \end{pmatrix}.$$

Definiamo l'**angolo orientato** formato dai vettori \mathbf{u} e \mathbf{v} come l'angolo $\widehat{\mathbf{u}\mathbf{v}} = [\psi - \varphi]_{2\pi}$.

Siano $\mathbf{a}, \mathbf{b} \in \mathbf{V} \setminus \{\mathbf{0}\}$, il loro angolo orientato è definito come

$$\widehat{\mathbf{a}\mathbf{b}} = \frac{\widehat{\mathbf{a}} \cdot \widehat{\mathbf{b}}}{\|\widehat{\mathbf{a}}\| \|\widehat{\mathbf{b}}\|}.$$

Proposizione 7.9. *L'angolo orientato gode delle seguenti proprietà: $\forall \mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbf{V} \setminus \{\mathbf{0}\}, \forall \lambda, \mu > 0$.*

- a) $\widehat{\mathbf{a}\mathbf{a}} = 0$,
- b) $\widehat{\mathbf{a}\mathbf{b}} = -\widehat{\mathbf{b}\mathbf{a}}$,
- c) $\widehat{\mathbf{a}\mathbf{b}} + \widehat{\mathbf{b}\mathbf{c}} = \widehat{\mathbf{a}\mathbf{c}}$,
- d) $(\widehat{\lambda\mathbf{a}})(\widehat{\mu\mathbf{b}}) = \widehat{\mathbf{a}\mathbf{b}}$

Proof. Da fare. □

Si osservi che per la definizione di angolo orientato è stata fissata una base ortonormale di \mathbf{V} . In realtà la definizione dipende dall'orientazione di \mathbf{V} definita dalla base fissata.

Infatti, sia $\{\mathbf{i}', \mathbf{j}'\}$ un'altra base ortonormale, concordemente orientata con $\{\mathbf{i}, \mathbf{j}\}$, e sia

$$A = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

la matrice del cambiamento di coordinate da $\{\mathbf{i}, \mathbf{j}\}$ a $\{\mathbf{i}', \mathbf{j}'\}$. Se

$$\mathbf{u} = \cos \varphi \mathbf{i} + \sin \varphi \mathbf{j} = \cos \varphi' \mathbf{i}' + \sin \varphi' \mathbf{j}' \quad \mathbf{v} = \cos \psi \mathbf{i} + \sin \psi \mathbf{j} = \cos \psi' \mathbf{i}' + \sin \psi' \mathbf{j}'$$

allora si ha

$$\begin{pmatrix} \cos \varphi' \\ \sin \varphi' \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} = \begin{pmatrix} \cos(\varphi + \alpha) \\ \sin(\varphi + \alpha) \end{pmatrix}$$

e similmente

$$\begin{pmatrix} \cos \psi' \\ \sin \psi' \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} \cos \psi \\ \sin \psi \end{pmatrix} = \begin{pmatrix} \cos(\psi + \alpha) \\ \sin(\psi + \alpha) \end{pmatrix}$$

Quindi si ha

$$\psi' - \varphi' \equiv (\psi + \alpha) - (\varphi + \alpha) \equiv \psi - \varphi \pmod{2\pi},$$

e pertanto le due basi definiscono lo stesso angolo orientato tra \mathbf{u} e \mathbf{v} .

L'insieme \mathbb{C} dei numeri complessi è uno spazio vettoriale reale di dimensione 2. La base $\{1, i\}$ identifica \mathbb{C} con \mathbb{R}^2 associando ad un numero complesso $a + ib$ la coppia (a, b) . Se $z = a + ib \in \mathbb{C}$, il **modulo** di z , che è definito come

$$|z| = \sqrt{z\bar{z}} = \sqrt{a^2 + b^2}$$

è uguale a $\|(a, b)\|$. Ad ogni $z \in \mathbb{C}^*$ possiamo associare il numero complesso

$$\frac{z}{|z|} = \frac{a}{|z|} + i \frac{b}{|z|}$$

che ha modulo 1, e quindi è della forma

$$\frac{z}{|z|} = \cos \theta + i \sin \theta$$

L'angolo individuato da θ si dice l'**argomento** di z e si denota con $\arg(z)$; la determinazione principale di $\arg(z)$ è talvolta chiamata **argomento principale** di z . Poiché

$$z = |z| \frac{z}{|z|}$$

deduciamo da quanto detto che ogni numero complesso $z \neq 0$ si scrive nella forma seguente:

$$z = |z|(\cos \theta + i \sin \theta) \tag{169}$$

dove θ è una determinazione di $\arg(z)$. La 169 è una **rappresentazione trigonometrica** di z . Se $z, w \in \mathbb{C}^*$ allora si ha

$$|zw| = |z||w| \quad \text{e} \quad \arg(zw) = \arg(z) + \arg(w).$$

Ciò può essere verificato per mezzo delle rappresentazioni trigonometriche nel modo seguente. Siano $\theta \in \arg(z)$, $\varphi \in \arg(w)$. Allora

$$\begin{aligned} zw &= |z|(\cos \theta + i \sin \theta)|w|(\cos \varphi + i \sin \varphi) = \\ &= |z||w|(\cos \theta \cos \varphi - \sin \theta \sin \varphi + i(\cos \theta \sin \varphi + \sin \theta \cos \varphi)) = \\ &= |z||w|[\cos(\theta + \varphi) + i \sin(\theta + \varphi)] \end{aligned}$$

Consideriamo lo spazio vettoriale $\mathbb{R}[X]$ dei polinomi a coefficienti reali in una indeterminata X . Per ogni $f(X), g(X) \in \mathbb{R}[X]$ poniamo

$$\langle f, g \rangle = \int_{-1}^1 f(x)g(x) dx.$$

Dalle proprietà dell'integrale definito segue subito che \langle, \rangle è un prodotto scalare. Dotato di questo prodotto scalare, $\mathbb{R}[X]$ è uno spazio vettoriale euclideo che non ha dimensione finita.

8 Prodotto vettoriale

Denotiamo con \mathbf{V} uno spazio vettoriale euclideo di dimensione 3, con prodotto scalare \langle, \rangle . Fissiamo una base ortonormale $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ di \mathbf{V} .

Prodotto vettoriale

Definizione 8.1. Siano $\mathbf{v}_1(x_1, y_1, z_1)$ e $\mathbf{v}_2(x_2, y_2, z_2)$ due vettori di \mathbf{V} . Il **prodotto vettoriale** di \mathbf{v}_1 per \mathbf{v}_2 è il vettore $\mathbf{v}_1 \wedge \mathbf{v}_2$ (\mathbf{v}_1 vettore \mathbf{v}_2) le cui coordinate sono

$$\begin{pmatrix} y_1 z_2 - z_1 y_2 \\ z_1 x_2 - x_1 z_2 \\ x_1 y_2 - y_1 x_2 \end{pmatrix},$$

cioè i minori di ordine 2, presi a segni alterni $+, -, +$, della matrice

$$\begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{pmatrix}$$

L'operazione di prodotto vettoriale che abbiamo introdotto associa a una coppia ordinata di vettori un terzo vettore, cioè è un'applicazione

$$\wedge : \mathbf{V}^2 \rightarrow \mathbf{V}.$$

Poiché è stata definita utilizzando le coordinate dei vettori, dobbiamo aspettarci che quest'operazione dipenda dalla scelta della base ortonormale $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$. Studiandone le proprietà vedremo che in effetti il prodotto vettoriale dipende solo dall'orientazione di \mathbf{V} definita dalla base.

Il teorema seguente descrive le principali proprietà del prodotto vettoriale.

Teorema 8.1. $\forall \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbf{V}, \forall c \in \mathbb{R},$

1. $\mathbf{v}_1 \wedge \mathbf{v}_2 = -\mathbf{v}_2 \wedge \mathbf{v}_1$
2. $\mathbf{v}_1 \wedge (\mathbf{v}_2 + \mathbf{v}_3) = \mathbf{v}_1 \wedge \mathbf{v}_2 + \mathbf{v}_1 \wedge \mathbf{v}_3$
3. $c(\mathbf{v}_1 \wedge \mathbf{v}_2) = (c\mathbf{v}_1) \wedge \mathbf{v}_2 = \mathbf{v}_1 \wedge (c\mathbf{v}_2)$
4. $\langle \mathbf{v}_1, \mathbf{v}_1 \wedge \mathbf{v}_2 \rangle = 0$
5. $\langle \mathbf{v}_2, \mathbf{v}_1 \wedge \mathbf{v}_2 \rangle = 0$
6. $\|\mathbf{v}_1 \wedge \mathbf{v}_2\|^2 = \|\mathbf{v}_1\|^2 \|\mathbf{v}_2\|^2 - \langle \mathbf{v}_1, \mathbf{v}_2 \rangle^2$
7. $\mathbf{v}_1 \wedge \mathbf{v}_2 = \mathbf{0}$ se e solo se \mathbf{v}_1 e \mathbf{v}_2 sono paralleli.

Proof.

1.

$$\mathbf{v}_1 \wedge \mathbf{v}_2 = \begin{pmatrix} y_1 z_2 - z_1 y_2 \\ z_1 x_2 - x_1 z_2 \\ x_1 y_2 - y_1 x_2 \end{pmatrix} = \begin{pmatrix} -y_2 z_1 + z_2 y_1 \\ -z_2 x_1 + x_2 z_1 \\ -x_2 y_1 + y_2 x_1 \end{pmatrix} = - \begin{pmatrix} y_2 z_1 - z_2 y_1 \\ z_2 x_1 - x_2 z_1 \\ x_2 y_1 - y_2 x_1 \end{pmatrix} = -\mathbf{v}_2 \wedge \mathbf{v}_1$$

2.

$$\begin{aligned} \mathbf{v}_1 \wedge (\mathbf{v}_2 + \mathbf{v}_3) &= \begin{pmatrix} y_1(z_2 + z_3) - z_1(y_2 + y_3) \\ z_1(x_2 + x_3) - x_1(z_2 + z_3) \\ x_1(y_2 + y_3) - y_1(x_2 + x_3) \end{pmatrix} = \begin{pmatrix} y_1 z_2 + y_1 z_3 - z_1 y_2 - z_1 y_3 \\ z_1 x_2 + z_1 x_3 - x_1 z_2 - x_1 z_3 \\ x_1 y_2 + x_1 y_3 - y_1 x_2 - y_1 x_3 \end{pmatrix} \\ &= \begin{pmatrix} y_1 z_2 - z_1 y_2 + y_1 z_3 - z_1 y_3 \\ z_1 x_2 - x_1 z_2 + z_1 x_3 - x_1 z_3 \\ x_1 y_2 - y_1 x_2 + x_1 y_3 - y_1 x_3 \end{pmatrix} = \begin{pmatrix} y_1 z_2 - z_1 y_2 \\ z_1 x_2 - x_1 z_2 \\ x_1 y_2 - y_1 x_2 \end{pmatrix} + \begin{pmatrix} y_1 z_3 - z_1 y_3 \\ z_1 x_3 - x_1 z_3 \\ x_1 y_3 - y_1 x_3 \end{pmatrix} \\ &= \mathbf{v}_1 \wedge \mathbf{v}_2 + \mathbf{v}_1 \wedge \mathbf{v}_3 \end{aligned}$$

3.

$$c(\mathbf{v}_1 \wedge \mathbf{v}_2) = c \begin{pmatrix} y_1 z_2 - z_1 y_2 \\ z_1 x_2 - x_1 z_2 \\ x_1 y_2 - y_1 x_2 \end{pmatrix} = \begin{pmatrix} c(y_1 z_2) - c(z_1 y_2) \\ c(z_1 x_2) - c(x_1 z_2) \\ c(x_1 y_2) - c(y_1 x_2) \end{pmatrix}$$

Da qui possiamo avere

(a)

$$\begin{pmatrix} (cy_1)z_2 - (cz_1)y_2 \\ (cz_1)x_2 - (cx_1)z_2 \\ (cx_1)y_2 - (cy_1)x_2 \end{pmatrix}$$

che corrisponde ai minori della seguente matrice

$$\begin{pmatrix} cx_1 & cy_1 & cz_1 \\ x_2 & y_2 & z_2 \end{pmatrix}$$

e dunque al prodotto vettoriale $(c\mathbf{v}_1) \wedge \mathbf{v}_2$

(b)

$$\begin{pmatrix} y_1(cz_2) - z_1(cy_2) \\ z_1(cx_2) - x_1(cz_2) \\ x_1(cy_2) - y_1(cx_2) \end{pmatrix}$$

che corrisponde ai minori della seguente matrice

$$\begin{pmatrix} x_1 & y_1 & z_1 \\ cx_2 & cy_2 & cz_2 \end{pmatrix}$$

e dunque al prodotto vettoriale $\mathbf{v}_1 \wedge (c\mathbf{v}_2)$

4. $\langle \mathbf{v}_1, \mathbf{v}_1 \wedge \mathbf{v}_2 \rangle$ uguaglia il determinante della matrice

$$\begin{pmatrix} x_1 & y_1 & z_1 \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{pmatrix}$$

che ha due righe uguali e quindi ha determinante nullo.

5. Nello stesso modo della (4) si dimostra la (5).

6. Per dimostrare la (6) scriviamo esplicitamente

$$\|\mathbf{v}_1 \wedge \mathbf{v}_2\|^2 = (y_1 z_2 - z_1 y_2)^2 + (z_1 x_2 - x_1 z_2)^2 + (x_1 y_2 - y_1 x_2)^2$$

e

$$\|\mathbf{v}_1\|^2 \|\mathbf{v}_2\|^2 - \langle \mathbf{v}_1, \mathbf{v}_2 \rangle^2 = (x_1^2 + y_1^2 + z_1^2)(x_2^2 + y_2^2 + z_2^2) - (x_1 x_2 + y_1 y_2 + z_1 z_2)^2.$$

Svolgendo i calcoli si verifica che i secondi membri sono uguali.

7. È ovvia, perchè il prodotto vettoriale uguale a 0 indica che i due vettori sono linearmente dipendenti e dunque paralleli.

□

Si osservi che la (6) può anche esprimersi nella forma seguente:

$$\|\mathbf{v}_1 \wedge \mathbf{v}_2\|^2 = \begin{vmatrix} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle & \langle \mathbf{v}_1, \mathbf{v}_2 \rangle \\ \langle \mathbf{v}_2, \mathbf{v}_1 \rangle & \langle \mathbf{v}_2, \mathbf{v}_2 \rangle \end{vmatrix}$$

Corollario 14.

8. Se \mathbf{v}_1 e \mathbf{v}_2 sono linearmente indipendenti, ogni vettore ortogonale sia a \mathbf{v}_1 che a \mathbf{v}_2 è un multiplo scalare di $\mathbf{v}_1 \wedge \mathbf{v}_2$.
9. Se \mathbf{v}_1 e \mathbf{v}_2 sono linearmente indipendenti, $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_1 \wedge \mathbf{v}_2\}$ è una base di \mathbf{V} concordemente orientata con $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$.

Proof.

Poiché \mathbf{v}_1 e \mathbf{v}_2 sono linearmente indipendenti e $\dim(\mathbf{V}) = 3$, si ha

$$\dim(\{\mathbf{v}_1, \mathbf{v}_2\}^\perp) = \dim(\langle \mathbf{v}_1, \mathbf{v}_2 \rangle^\perp) = 1.$$

Per la (4) e la (5), $\mathbf{v}_1 \wedge \mathbf{v}_2 \in \{\mathbf{v}_1, \mathbf{v}_2\}^\perp$, e per la (7), $\mathbf{v}_1 \wedge \mathbf{v}_2 \neq 0$. Quindi $\mathbf{v}_1 \wedge \mathbf{v}_2$ genera $\{\mathbf{v}_1, \mathbf{v}_2\}^\perp$.

9. L'indipendenza lineare di $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_1 \wedge \mathbf{v}_2$ segue dal fatto che $\mathbf{v}_1, \mathbf{v}_2$ lo sono e che $\mathbf{v}_1 \wedge \mathbf{v}_2 \in \{\mathbf{v}_1, \mathbf{v}_2\}^\perp$. Poiché $\dim(\mathbf{V}) = 3$ si conclude che $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_1 \wedge \mathbf{v}_2\}$ è una base. Per dimostrare che $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_1 \wedge \mathbf{v}_2\}$ è concordemente orientata con $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ basta osservare che la matrice del cambiamento di coordinate da $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_1 \wedge \mathbf{v}_2\}$ a $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ ha determinante uguale a $\|\mathbf{v}_1 \wedge \mathbf{v}_2\|^2$. □

Ora non è difficile accertare in che modo $\mathbf{v}_1 \wedge \mathbf{v}_2$ dipenda dalla scelta della base $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$. Se \mathbf{v}_1 e \mathbf{v}_2 sono linearmente dipendenti, il loro prodotto vettoriale è $\mathbf{v}_1 \wedge \mathbf{v}_2 = 0$ e quindi è univocamente determinato da \mathbf{v}_1 e da \mathbf{v}_2 . Se viceversa \mathbf{v}_1 e \mathbf{v}_2 sono linearmente indipendenti, $\|\mathbf{v}_1 \wedge \mathbf{v}_2\|$ è univocamente individuato dalla proprietà (6), e $\mathbf{v}_1 \wedge \mathbf{v}_2$ appartiene a $\{\mathbf{v}_1, \mathbf{v}_2\}^\perp$, che ha dimensione 1. Vediamo dunque che una diversa scelta della base può cambiare $\mathbf{v}_1 \wedge \mathbf{v}_2$ nel suo opposto, e per la proprietà (9), questo avviene se e solo se le due basi non sono concordemente orientate. In conclusione $\mathbf{v}_1 \wedge \mathbf{v}_2$ dipende solo dall'orientazione dello spazio \mathbf{V} definita dalla base $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$.

La proprietà (6) ha una interessante interpretazione, data dalla seguente proposizione.

Proposizione 8.2. *Supponiamo che \mathbf{v} e \mathbf{w} siano linearmente indipendenti, e scriviamo $\mathbf{v} = \mathbf{a} + \mathbf{b}$, dove \mathbf{a} è parallelo a \mathbf{w} e \mathbf{b} è ortogonale a \mathbf{w} . Allora*

$$\|\mathbf{v} \wedge \mathbf{w}\| = \|\mathbf{b}\| \|\mathbf{w}\|.$$

Proof. Si ha, utilizzando le proprietà (2), (7) e (6):

$$\|\mathbf{v} \wedge \mathbf{w}\| = \|(\mathbf{a} + \mathbf{b}) \wedge \mathbf{w}\| = \|(\mathbf{a} \wedge \mathbf{w}) + (\mathbf{b} \wedge \mathbf{w})\| = \|\mathbf{b} \wedge \mathbf{w}\| = \|\mathbf{b}\| \|\mathbf{w}\|.$$

□

Si noti che nel caso in cui \mathbf{V} è lo spazio dei vettori dello spazio ordinario, la proposizione 8.2 afferma che $\|\mathbf{v} \wedge \mathbf{w}\|$ è uguale all'area del parallelogramma costruito sui vettori \mathbf{v} e \mathbf{w} .

Prodotto misto

Definizione 8.2. Dati $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbf{V}$, il prodotto scalare $\langle \mathbf{v}_1, \mathbf{v}_2 \wedge \mathbf{v}_3 \rangle$ è detto **prodotto misto** di $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

Se le coordinate dei vettori dati sono rispettivamente (x_1, y_1, z_1) , (x_2, y_2, z_2) , (x_3, y_3, z_3) , si ha

$$\langle \mathbf{v}_1, \mathbf{v}_2 \wedge \mathbf{v}_3 \rangle = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}.$$

Da quest'espressione di $\langle \mathbf{v}_1, \mathbf{v}_2 \wedge \mathbf{v}_3 \rangle$ segue immediatamente che scambiando fra loro due dei fattori il prodotto misto cambia di segno e che $\langle \mathbf{v}_1, \mathbf{v}_2 \wedge \mathbf{v}_3 \rangle = 0$ se e solo se $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ sono linearmente dipendenti.

Identità di Lagrange

Proposizione 8.3. *Per ogni $\mathbf{v}, \mathbf{w}, \mathbf{a}, \mathbf{b} \in \mathbf{V}$ si ha*

$$\langle \mathbf{v} \wedge \mathbf{w}, \mathbf{a} \wedge \mathbf{b} \rangle = \begin{vmatrix} \langle \mathbf{v}, \mathbf{a} \rangle & \langle \mathbf{v}, \mathbf{b} \rangle \\ \langle \mathbf{w}, \mathbf{a} \rangle & \langle \mathbf{w}, \mathbf{b} \rangle \end{vmatrix}. \quad (170)$$

Proof. Se le coordinate dei vettori dati sono (v_1, v_2, v_3) , (w_1, w_2, w_3) , (a_1, a_2, a_3) , (b_1, b_2, b_3) , il primo membro della 170 è

$$\begin{vmatrix} v_2 & v_3 \\ w_2 & w_3 \end{vmatrix} \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} + \begin{vmatrix} v_1 & v_3 \\ w_1 & w_3 \end{vmatrix} \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} + \begin{vmatrix} v_1 & v_2 \\ w_1 & w_2 \end{vmatrix} \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$$

il secondo membro è

$$(v_1 a_1 + v_2 a_2 + v_3 a_3)(w_1 b_1 + w_2 b_2 + w_3 b_3) - (v_1 b_1 + v_2 b_2 + v_3 b_3)(w_1 a_1 + w_2 a_2 + w_3 a_3)$$

L'identità 170 si verifica svolgendo i calcoli e confrontando le due espressioni. □

9 Spazi euclidei

Spazio euclideo

Definizione 9.1. Sia \mathbf{E} uno spazio affine reale con associato spazio vettoriale \mathbf{V} . Diremo che \mathbf{E} è uno **spazio euclideo** se in \mathbf{V} è assegnato un prodotto scalare definito positivo, cioè se \mathbf{V} è uno spazio vettoriale euclideo. Useremo la notazione $\langle \mathbf{v}, \mathbf{w} \rangle$ per il prodotto scalare di due vettori \mathbf{v} e \mathbf{w} in \mathbf{V} .

n -spazio euclideo numerico

Definizione 9.2. Lo spazio affine numerico $\mathbf{A}^n(\mathbb{R})$ diventa uno spazio euclideo se in \mathbb{R}^n si assegna il prodotto scalare standard. Questo spazio euclideo, denotato con \mathbf{E}^n , viene chiamato l' **n -spazio euclideo numerico**. Il suo spazio vettoriale associato è l' n -spazio vettoriale euclideo.

Sistema di coordinate cartesiane

Definizione 9.3. Un sistema di coordinate $O\mathbf{e}_1 \dots \mathbf{e}_n$ nello spazio euclideo \mathbf{E} è tale che $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ sia una base ortonormale di \mathbf{V} si chiama **sistema di coordinate cartesiane** oppure **riferimento cartesiano**.

Nota 29. Nello studio degli spazi euclidei si utilizzano sistemi di coordinate cartesiane: ciò facilita notevolmente i calcoli.

La formula del cambiamento di coordinate da un riferimento cartesiano a un altro ha forma particolare, dovuta al fatto che le corrispondenti basi dei vettori sono ortonormali. Siano infatti $O\mathbf{e}_1 \dots \mathbf{e}_n$ e $O'\mathbf{e}'_1 \dots \mathbf{e}'_n$ due riferimenti cartesiani in \mathbf{E} , e sia $P \in \mathbf{E}$ un punto di coordinate $\mathbf{x} = {}^t(x_1 \dots x_n)$ rispetto al primo riferimento e $\mathbf{x}' = {}^t(x'_1 \dots x'_n)$ rispetto al secondo. Allora si ha

$$\mathbf{x}' = A\mathbf{x} + \mathbf{c}, \quad (171)$$

dove $A = M_{\mathbf{e}', \mathbf{e}}(\mathbf{1}_V)$ è una matrice ortogonale e $\mathbf{c} = {}^t(c_1, \dots, c_n)$ è individuato dall'identità

$$\overrightarrow{O'O} = c_1\mathbf{e}'_1 + \dots + c_n\mathbf{e}'_n. \quad (172)$$

Nel caso $n = 2$, cioè in cui \mathbf{E} è un piano euclideo, la matrice A è di una delle due forme 1 o 2 a seconda che i due riferimenti siano concordemente o discordemente orientati.

Utilizzando il prodotto scalare definito nello spazio vettoriale associato \mathbf{V} , è possibile introdurre nello spazio euclideo \mathbf{E} nozioni di natura metrica come distanze, angoli, aree.

Distanza euclidea

Definizione 9.4. Siano P e Q punti dello spazio euclideo \mathbf{E} . La distanza tra P e Q , indicata con $d(P, Q)$, è definita da

$$d(P, Q) = \|\overrightarrow{PQ}\|.$$

Se in \mathbf{E} è fissato un sistema di coordinate cartesiane $O\mathbf{e}_1 \dots \mathbf{e}_n$, e se i punti dati sono $P(x_1, \dots, x_n)$ e $Q(y_1, \dots, y_n)$ rispettivamente, allora si ha

$$d(P, Q) = \sqrt{\sum_{i=1}^n (y_i - x_i)^2}.$$

Nel caso in cui $\mathbf{E} = \mathbf{V}_a$, dove \mathbf{V} è uno spazio vettoriale euclideo, si ha

$$d(\mathbf{v}, \mathbf{w}) = \|\mathbf{w} - \mathbf{v}\|.$$

Proposizione 9.1. La distanza gode delle seguenti proprietà:

SM1 $\forall P, Q \in \mathbf{E}, d(P, Q) \geq 0$ e $d(P, Q) = 0 \iff P = Q$.

SM2 $\forall P, Q \in \mathbf{E}, d(P, Q) = d(Q, P)$.

SM3 $\forall P, Q, R \in \mathbf{E}, d(P, Q) + d(Q, R) \geq d(P, R)$.

La 174 può essere presa come sistema assiomatico per definire una classe più generale di spazi.

Spazio metrico

Definizione 9.5. Uno **spazio metrico** è un insieme \mathbf{X} su cui sia definita una **distanza**, cioè un'applicazione $d : \mathbf{X}^2 \rightarrow \mathbb{R}$ che soddisfi alle tre condizioni SM1, SM2, SM3.

La 174 garantisce che, con la distanza definita dalla 9.4, ogni spazio euclideo è uno spazio metrico. Non è vero il viceversa: in generale uno spazio metrico non è uno spazio euclideo, e infatti gli spazi metrici sono oggetti molto più generali. Per rendersi conto della loro maggiore generalità si pensi che ogni sottoinsieme di uno spazio metrico è ancora uno spazio metrico; in particolare tutti i sottoinsiemi di uno spazio euclideo E sono spazi metrici.

È anche possibile definire la nozione di angolo convesso di due rette utilizzando quella di angolo convesso di due vettori, definito in \mathbf{V} grazie al prodotto scalare.

Versore di una retta

Definizione 9.6. Data una retta \mathbf{r} in \mathbf{E} , un vettore di direzione di \mathbf{r} di lunghezza 1 si dirà **versore** di \mathbf{r} .

Nota 30. Ci sono esattamente due versori di \mathbf{r} , l'uno opposto dell'altro.

Angolo convesso tra due rette

Definizione 9.7. Date \mathbf{r} ed \mathbf{r}' , fissiamo rispettivamente due vettori di direzione \mathbf{a} e \mathbf{a}' . Definiamo quindi l'**angolo convesso** φ tra \mathbf{r} ed \mathbf{r}' come l'angolo convesso tra \mathbf{a} e \mathbf{a}' , cioè mediante le condizioni seguenti:

$$0 \leq \varphi \leq \pi, \quad \cos \varphi = \frac{\langle \mathbf{a}, \mathbf{a}' \rangle}{\|\mathbf{a}\| \|\mathbf{a}'\|}$$

Questa definizione dipende dalla scelta di \mathbf{a} e di \mathbf{a}' : l'angolo φ così definito viene sostituito da $\pi - \varphi$ se uno dei due vettori viene moltiplicato per uno scalare negativo. Si noti che anche per definire l'angolo convesso di due rette nello spazio è necessario scegliere un verso di percorrenza su ognuna di esse, che è quanto fissare vettori di direzione.

Rette ortogonali

Definizione 9.8. Due rette in \mathbf{E} si dicono **ortogonali** (o **perpendicolari**) se il loro angolo convesso è $\frac{\pi}{2}$, cioè se un (e quindi ad ogni) vettore di direzione di una delle rette è ortogonale a un (e quindi ad ogni) vettore di direzione dell'altra.

Sia \mathbf{E} un piano euclideo in cui sia fissato un riferimento cartesiano Oij . Consideriamo una retta \mathbf{r} , avente equazione cartesiana

$$AX + BY + C = 0. \quad (173)$$

Ricordando che $\mathbf{a}(-B, A)$ è un vettore di direzione di \mathbf{r} , si deduce che i vettori

$$\pm \mathbf{u} = \pm \frac{\mathbf{a}}{\|\mathbf{a}\|}$$

di coordinate

$$\pm \left(\frac{-B}{\sqrt{A^2 + B^2}}, \frac{A}{\sqrt{A^2 + B^2}} \right)$$

sono i versori di \mathbf{r} .

Vettore normale a una retta

Definizione 9.9. Un vettore non nullo \mathbf{m} si dice **normale** (o **perpendicolare**, o **ortogonale**) a una retta \mathbf{r} se \mathbf{m} è ortogonale a un (e quindi ad ogni) vettore di direzione di \mathbf{r} .

È evidente che, essendo $\dim(\mathbf{V}) = 2$, due vettori qualunque normali a \mathbf{r} sono tra loro paralleli. Inoltre esistono esattamente due versori, l'uno opposto all'altro, normali a \mathbf{r} e si dicono **versori normali** a \mathbf{r} .

Poiché $\mathbf{a}(-B, A)$ è un vettore di direzione di \mathbf{r} , il vettore $\mathbf{n}(A, B)$ è ortogonale a \mathbf{r} , e quindi i versori normali a \mathbf{r} sono

$$\pm \mathbf{v} = \pm \frac{\mathbf{n}}{\|\mathbf{n}\|}$$

Dato un punto $Q(a, b) \in \mathbf{r}$, un'equazione cartesiana di \mathbf{r} si scrive nella forma

$$A(X - a) + B(Y - b) = 0. \quad (174)$$

Poiché la 174 è individuata da a, b, A, B , vediamo che \mathbf{r} è individuata da un suo punto $Q(a, b)$ e da un vettore normale $\mathbf{n}(A, B)$: una retta in \mathbf{E} può quindi essere individuata assegnando un suo punto e un vettore ad essa perpendicolare.

Siano \mathbf{r} ed \mathbf{r}' due rette di equazioni cartesiane

$$AX + BY + C = 0 \quad \text{e} \quad A_1X + B_1Y + C_1 = 0$$

rispettivamente, il loro **angolo convesso** φ , riferito ai vettori di direzione $\mathbf{a}(-B, A)$ e $\mathbf{a}_1(-B', A')$, è definito dalle condizioni $0 \leq \varphi \leq \pi$ e

$$\cos \varphi = \frac{\langle \mathbf{a}, \mathbf{a}_1 \rangle}{\|\mathbf{a}\| \|\mathbf{a}_1\|} = \frac{AA' + BB'}{\sqrt{A^2 + B^2} \sqrt{(A')^2 + (B')^2}}.$$

In particolare \mathbf{r} ed \mathbf{r}' sono **perpendicolari** se e solo se

$$AA' + BB' = 0.$$

Supponiamo data una retta \mathbf{r} di equazione cartesiana 173, e un punto $P_0(x_0, y_0) \in \mathbf{E}$. Consideriamo la retta \mathbf{r}' che passa per P_0 ed è ortogonale a \mathbf{r} ; evidentemente questa retta esiste ed è unica. Poiché \mathbf{r}' non è parallela a \mathbf{r} , essa ha uno e un solo punto N in comune con \mathbf{r} : N è il **pie' della perpendicolare condotta da P_0 a \mathbf{r}** .

La distanza $d(P_0, N)$ si chiama **distanza** di P_0 da \mathbf{r} , e si denota con $d(P_0, \mathbf{r})$.

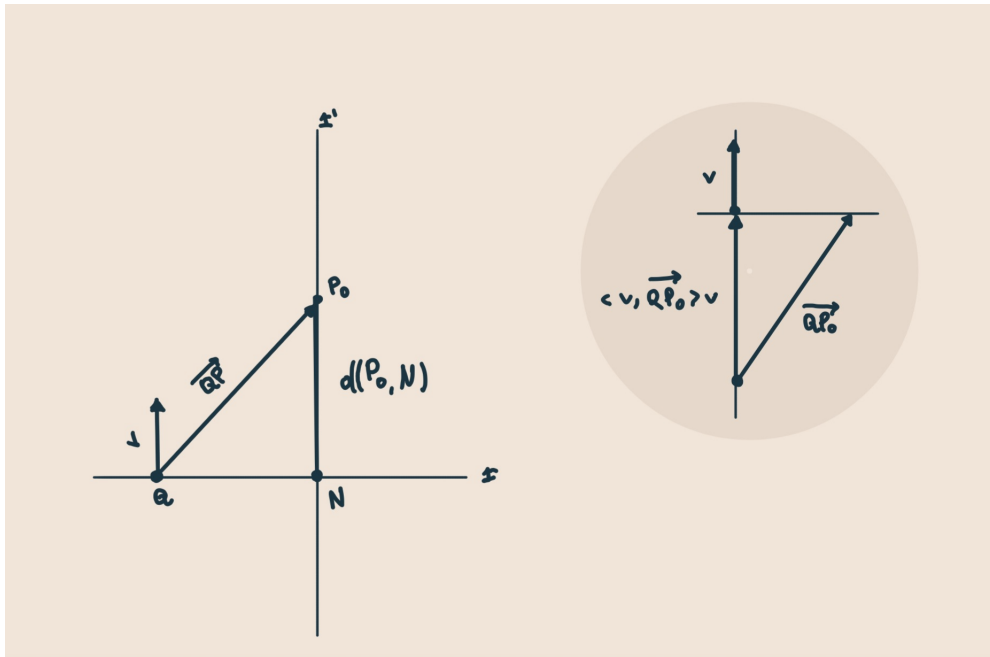
È possibile calcolare $d(P_0, \mathbf{r})$ trovando un'equazione di \mathbf{r}' , e poi le coordinate di N , e infine calcolando $d(P_0, N)$. Un metodo più efficiente è il seguente.

Proposizione 9.2. La distanza $d(P_0, \mathbf{r})$ di un punto $P_0(x_0, y_0) \in \mathbf{E}$ dalla retta \mathbf{r} di equazione 173 è data dalla seguente formula:

$$d(P_0, \mathbf{r}) = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}}.$$

Proof. Si osservi che se \mathbf{v} è un versore normale a \mathbf{r} e se $Q(a, b)$ è un punto qualsiasi di \mathbf{r} , allora

$$d(P_0, \mathbf{r}) = |\langle \mathbf{v}, \overrightarrow{QP_0} \rangle|,$$



Scegliamo

$$v \left(\frac{A}{\sqrt{A^2 + B^2}}, \frac{B}{\sqrt{A^2 + B^2}} \right)$$

come versore normale; poiché $\overrightarrow{QP_0}$ ha coordinate $(x_0 - a, y_0 - b)$, tenendo presente che $C = -(Aa + Bb)$, otteniamo

$$d(P_0, r) = \frac{|A(x_0 - a) + B(y_0 - b)|}{\sqrt{A^2 + B^2}} = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}}.$$

□

Se r ed r' sono due rette parallele di \mathbf{E} , la loro distanza $d(r, r')$ è per definizione uguale a $d(P, r')$, dove P è un punto qualunque di r . Per calcolare $d(r, r')$ è sufficiente trovare un punto $P \in r$ e applicare la formula della proposizione 9.2.

Sia \mathbf{E} uno spazio euclideo di dimensione 3 con spazio vettoriale associato \mathbf{V} . Supponiamo fissato in \mathbf{E} un sistema di coordinate cartesiane $Oijk$. Consideriamo un piano \mathcal{P} in \mathbf{E} , di equazione cartesiana

$$AX + BY + CZ + D = 0. \quad (175)$$

Un vettore \mathbf{m} si dice **normale** (o **perpendicolare**, o **ortogonale**) a \mathcal{P} se è ortogonale ad ogni vettore della giacitura di \mathcal{P} . Se inoltre \mathbf{m} è un versore, esso viene detto **versore normale** a \mathcal{P} .

Poiché $\dim(\mathbf{E}) = 3$, due qualsiasi vettori normali a \mathcal{P} sono tra loro paralleli; quindi \mathcal{P} possiede esattamente due versori normali.

Fissiamo un punto $P_0(x_0, y_0, z_0) \in \mathcal{P}$ e consideriamo l'equazione cartesiana di \mathcal{P}

$$A(X - x_0) + B(Y - y_0) + C(Z - z_0) = 0. \quad (176)$$

Ogni vettore appartenente alla giacitura di \mathcal{P} è della forma $\overrightarrow{P_0P}$, $P \in \mathcal{P}$; quindi dalla 176 vediamo che il vettore $\mathbf{n}(A, B, C)$ è normale a \mathcal{P} . Ne deduciamo che i due vettori $\pm \mathbf{v} = \pm \frac{\mathbf{n}}{\|\mathbf{n}\|}$ di coordinate

$$\pm \left(\frac{A}{\sqrt{A^2 + B^2 + C^2}}, \frac{B}{\sqrt{A^2 + B^2 + C^2}}, \frac{C}{\sqrt{A^2 + B^2 + C^2}} \right)$$

sono i versori normali a \mathcal{P} .

La 176 mostra inoltre che un qualsiasi piano è individuato da un suo punto e da un vettore ad esso perpendicolare.

L'angolo convesso fra due piani di \mathbf{E} può essere definito utilizzando i loro vettori normali.

Angolo convesso tra due piani

Definizione 9.10. Siano \mathcal{P} e \mathcal{P}' due piani, di equazioni cartesiane

$$AX + BY + CZ + D = 0 \quad A'X + B'Y + C'Z + D' = 0.$$

L'angolo convesso di \mathcal{P} e \mathcal{P}' è l'angolo convesso tra i vettori $\mathbf{n}(A, B, C)$ ed $\mathbf{n}'(A', B', C')$, cioè è l'angolo φ definito da $0 \leq \varphi \leq \pi$ e da

$$\cos \varphi = \frac{\langle \mathbf{n}, \mathbf{n}' \rangle}{\|\mathbf{n}\| \|\mathbf{n}'\|} = \frac{AA' + BB' + CC'}{\sqrt{A^2 + B^2 + C^2} \sqrt{A'^2 + B'^2 + C'^2}}. \quad (177)$$

Si osservi che la definizione di φ dipende dai vettori \mathbf{n} ed \mathbf{n}' , e quindi dalla scelta delle equazioni dei piani. Moltiplicando una o l'altra per un fattore di proporzionalità negativo, φ si cambia in $\pi - \varphi$.

Se $\varphi = \frac{\pi}{2}$ i due piani si dicono **perpendicolari** od **ortogonali**. Ciò avviene se e solo se

$$AA' + BB' + CC' = 0.$$

Angolo tra una retta e un piano

Definizione 9.11. Sia \mathcal{P} un piano di equazione

$$AX + BY + CZ + D = 0,$$

e sia \mathbf{r} una retta con vettore di direzione $\mathbf{a}(l, m, n)$. **L'angolo tra \mathcal{P} ed \mathbf{r}** è l'angolo la cui determinazione è

$$\varphi = \psi - \frac{\pi}{2},$$

dove ψ è l'angolo convesso tra i vettori $\mathbf{n}(A, B, C)$ e \mathbf{a} . Quindi φ è definito dalle condizioni

$$-\frac{\pi}{2} \leq \varphi \leq \frac{\pi}{2}$$

e

$$\sin \varphi = \frac{Al + Bm + Cn}{\sqrt{A^2 + B^2 + C^2} \sqrt{l^2 + m^2 + n^2}}.$$

Se $\varphi = \pm \frac{\pi}{2}$, allora \mathcal{P} ed \mathbf{r} si dicono **perpendicolari** (od **ortogonali**).

Si noti che ponendo $\sin \varphi = 0$ si ritrova la condizione di parallelismo tra \mathcal{P} ed \mathbf{r}

$$Al + Bm + Cn = 0$$

già data nella proposizione 3.14.

La distanza di un punto da un piano si definisce in modo simile al caso della distanza punto-retta in un piano euclideo.

Distanza di un punto da un piano

Definizione 9.12. Siano \mathcal{P} un piano di equazione

$$AX + BY + CZ + D = 0,$$

e $P_0(x_0, y_0, z_0) \in \mathbf{E}$ un punto. Consideriamo la retta \mathbf{r} passante per P_0 e perpendicolare a \mathcal{P} e denotiamo con $N = \mathcal{P} \cap \mathbf{r}$ il piede della perpendicolare condotta da P_0 a \mathcal{P} . Definiamo la distanza di P_0 da \mathcal{P} come

$$d(P_0, \mathcal{P}) = d(P_0, N).$$

Si dimostra che

$$d(P_0, \mathcal{P}) = \frac{|Ax_0 + By_0 + Cz_0 + D|}{\sqrt{A^2 + B^2 + C^2}}.$$

La dimostrazione è simile a quella della proposizione 19.3 ed è lasciata al lettore.

Distanza tra una retta e un piano paralleli Se \mathfrak{r} e \mathcal{P} sono rispettivamente una retta e un piano paralleli, la loro distanza $d(\mathfrak{r}, \mathcal{P})$ è per definizione $d(P, \mathcal{P})$, dove P è un qualsiasi punto di \mathfrak{r} . Per calcolare $d(\mathfrak{r}, \mathcal{P})$ è sufficiente trovare le coordinate di un punto $P \in \mathfrak{r}$ e applicare la formula che dà $d(P, \mathcal{P})$.

Distanza di un punto da una retta

Definizione 9.13. Sia \mathfrak{r} una retta e sia $P_0 \in \mathbf{E}$. Consideriamo il piano \mathcal{P} passante per P_0 e perpendicolare a \mathfrak{r} e sia $N = \mathfrak{r} \cap \mathcal{P}$. La distanza di P_0 da \mathfrak{r} è definita come

$$d(P_0, \mathfrak{r}) = d(P_0, N).$$

Se \mathfrak{r} passa per il punto $Q(a, b, c)$ e ha vettore di direzione $\mathbf{a}(l, m, n)$, e $P_0(x_0, y_0, z_0) \in \mathbf{E}$, allora si ha la seguente formula:

$$d(P_0, \mathfrak{r}) = \frac{\sqrt{\begin{vmatrix} y_0 - b & z_0 - c \\ m & n \end{vmatrix}^2 + \begin{vmatrix} x_0 - a & z_0 - c \\ l & n \end{vmatrix}^2 + \begin{vmatrix} x_0 - a & y_0 - b \\ l & m \end{vmatrix}^2}}{\sqrt{l^2 + m^2 + n^2}}$$

Infatti

$$\overrightarrow{QP_0} = \overrightarrow{QN} + \overrightarrow{NP_0},$$

e \overrightarrow{QN} è parallelo ad \mathbf{a} , mentre $\overrightarrow{NP_0}$ è perpendicolare ad \mathbf{a} .

Dalla proposizione 8.2 si deduce che

$$d(P_0, \mathfrak{r}) = \|\overrightarrow{NP_0}\| = \frac{\|\mathbf{a} \wedge \overrightarrow{QP_0}\|}{\|\mathbf{a}\|}.$$

Esplicitando l'uguaglianza si ottiene l'asserto.

Distanza tra due rette Date due rette \mathfrak{r} ed \mathfrak{r}' non parallele in \mathbf{E} , definiamo la loro distanza, che denoteremo con $d(\mathfrak{r}, \mathfrak{r}')$, nel modo seguente. Supponiamo che \mathfrak{r} passi per il punto $Q(a, b, c)$ e abbia $\mathbf{a}(l, m, n)$ come vettore di direzione, e che \mathfrak{r}' passi per $Q'(a', b', c')$ e abbia vettore di direzione $\mathbf{a}'(l', m', n')$.

Osserviamo che esiste una e una sola coppia di punti, $N \in \mathfrak{r}$ ed $N' \in \mathfrak{r}'$, tali che la retta che li contiene sia perpendicolare sia a \mathfrak{r} che a \mathfrak{r}' . Infatti le due condizioni di perpendicolarità sono

$$\langle \overrightarrow{NN'}, \mathbf{a} \rangle = 0 \quad \langle \overrightarrow{NN'}, \mathbf{a}' \rangle = 0 \quad (178)$$

Scrivendo

$$\overrightarrow{ON} = \overrightarrow{OQ} + t\mathbf{a} \quad \overrightarrow{ON'} = \overrightarrow{OQ'} + t'\mathbf{a}',$$

si ottiene

$$\overrightarrow{NN'} = \overrightarrow{QQ'} + t'\mathbf{a}' - t\mathbf{a},$$

e quindi le 178 diventano

$$\langle \overrightarrow{QQ'} + t'\mathbf{a}' - t\mathbf{a}, \mathbf{a} \rangle = 0 \quad \langle \overrightarrow{QQ'} + t'\mathbf{a}' - t\mathbf{a}, \mathbf{a}' \rangle = 0,$$

cioè

$$\langle \overrightarrow{QQ'}, \mathbf{a} \rangle + t'\langle \mathbf{a}', \mathbf{a} \rangle - t\langle \mathbf{a}, \mathbf{a} \rangle = 0 \quad \langle \overrightarrow{QQ'}, \mathbf{a}' \rangle + t'\langle \mathbf{a}', \mathbf{a}' \rangle - t\langle \mathbf{a}, \mathbf{a}' \rangle = 0. \quad (179)$$

Poiché \mathbf{a} ed \mathbf{a}' sono linearmente indipendenti, per il teorema 8.1(6) si ha

$$\begin{vmatrix} \langle \mathbf{a}', \mathbf{a} \rangle & \langle \mathbf{a}, \mathbf{a} \rangle \\ \langle \mathbf{a}', \mathbf{a}' \rangle & \langle \mathbf{a}, \mathbf{a}' \rangle \end{vmatrix} = \|\mathbf{a} \wedge \mathbf{a}'\|^2 \neq 0.$$

Quindi il sistema (179) ammette una e una sola soluzione (t, t') , e ciò significa appunto che esiste un'unica coppia (N, N') che soddisfa alla condizione voluta.

La retta \mathfrak{s} passante per N e per N' si chiama **perpendicolare comune** a \mathfrak{r} e a \mathfrak{r}' .

Ciò posto, definiamo la **distanza di \mathfrak{r} da \mathfrak{r}'** come

$$d(\mathfrak{r}, \mathfrak{r}') = d(N, N').$$

Si ha la seguente formula che permette di calcolare agevolmente la distanza di due rette assegnate:

$$d(\mathfrak{r}, \mathfrak{r}') = \frac{\begin{vmatrix} a - a' & b - b' & c - c' \\ l & m & n \\ l' & m' & n' \end{vmatrix}}{\sqrt{\begin{vmatrix} m & n \\ m' & n' \end{vmatrix}^2 + \begin{vmatrix} l & n \\ l' & n' \end{vmatrix}^2 + \begin{vmatrix} l & m \\ l' & m' \end{vmatrix}^2}} \quad (180)$$

La (180) si dimostra nel modo seguente. Sia \mathbf{b} un versore della perpendicolare comune a \mathfrak{r} e a \mathfrak{r}' . Si ha

$$\langle \mathbf{b}, \overrightarrow{QQ'} \rangle = \langle \mathbf{b}, \overrightarrow{QN} \rangle + \langle \mathbf{b}, \overrightarrow{NQ'} \rangle = \langle \mathbf{b}, \overrightarrow{NQ'} \rangle$$

e quindi

$$d(\mathfrak{r}, \mathfrak{r}') = |\langle \mathbf{b}, \overrightarrow{NQ'} \rangle| = |\langle \mathbf{b}, \overrightarrow{QQ'} \rangle|.$$

D'altra parte, poiché \mathbf{b} è ortogonale sia ad \mathbf{a} che ad \mathbf{a}' , si ha

$$\mathbf{b} = \pm \frac{\mathbf{a} \wedge \mathbf{a}'}{\|\mathbf{a} \wedge \mathbf{a}'\|}.$$

Quindi

$$d(\mathfrak{r}, \mathfrak{r}') = \frac{|\langle \mathbf{a} \wedge \mathbf{a}', \overrightarrow{QQ'} \rangle|}{\|\mathbf{a} \wedge \mathbf{a}'\|}.$$

Esplicitando quest'uguaglianza si ottiene la formula cercata.

Per trovare equazioni cartesiane della retta f perpendicolare comune a \mathfrak{r} e a \mathfrak{r}' si procede nel modo seguente. Conosciamo $\mathbf{a} \wedge \mathbf{a}'$, vettore di direzione \mathfrak{s} , le cui coordinate denotiamo con $(\beta_1, \beta_2, \beta_3)$. Denotando con $P(X, Y, Z)$ un punto di coordinate indeterminate, e imponendo le condizioni di complanarità di \mathfrak{r} e di \mathfrak{r}' con la retta passante per P e avente vettore di direzione $\mathbf{a} \wedge \mathbf{a}'$, otteniamo le due seguenti equazioni:

$$\begin{vmatrix} X - a & Y - b & Z - c \\ \beta_1 & \beta_2 & \beta_3 \\ l & m & n \end{vmatrix} = 0 \quad \begin{vmatrix} X - a' & Y - b' & Z - c' \\ \beta_1 & \beta_2 & \beta_3 \\ l' & m' & n' \end{vmatrix} = 0.$$

Poiché queste equazioni rappresentano due piani distinti e sono soddisfatte dai punti di \mathfrak{s} , esse sono equazioni cartesiane di \mathfrak{s} .

Supponiamo fissato in uno spazio euclideo \mathbf{E} un riferimento cartesiano $O\mathbf{e}_1 \dots \mathbf{e}_n$ ed un punto $C(c_1, \dots, c_n)$. Sia $r > 0$.

- La **sfera di centro C e raggio r** è l'insieme

$$\mathbf{S}(C, r) = \{P(x_1, \dots, x_n) \in \mathbf{E} : d(C, P) = r\},$$

cioè tali che

$$\sum_{i=1}^n (x_i - c_i)^2 = r^2. \quad (181)$$

- Il **disco di centro C e raggio r** è l'insieme

$$\mathbf{D}(C, r) = \{P(x_1, \dots, x_n) \in \mathbf{E} : d(C, P) \leq r\},$$

cioè tali che

$$\sum_{i=1}^n (x_i - c_i)^2 \leq r^2. \quad (182)$$

Nel caso in cui \mathbf{E} sia un piano, $\mathbf{S}(C, r)$ e $\mathbf{D}(C, r)$ sono la **circonferenza** e il **cerchio** rispettivamente, di centro C e raggio r . Se $\dim(\mathbf{E}) = 1$, $\mathbf{D}(C, r)$ è il segmento di lunghezza $2r$ e di centro C , ed $\mathbf{S}(C, r)$ è l'insieme costituito dai suoi estremi.

Quando $\mathbf{E} = \mathbf{E}^n$, si usano i simboli \mathbf{S}^{n-1} e \mathbf{D}^n per denotare $\mathbf{S}(\mathbf{0}, 1)$ e $\mathbf{D}(\mathbf{0}, 1)$ rispettivamente.

Dalla 181 deduciamo che $\mathbf{S}(C, r)$ è l'insieme dei punti di \mathbf{E} le cui coordinate sono soluzioni dell'equazione

$$\sum_{i=1}^n (X_i - c_i)^2 = r^2 \quad (183)$$

nelle indeterminate X_1, X_2, \dots, X_n . La 183 è detta **equazione cartesiana della sfera** $\mathbf{S}(C, r)$. Portando r^2 a primo membro e svolgendo i calcoli troviamo che la 183 è equivalente all'equazione

$$\sum_{i=1}^n X_i(X_i + d_i) + d = 0 \quad d_i = -2c_i, \quad d = \sum_{i=1}^n c_i^2 - r^2. \quad (184)$$

A causa della loro definizione e del fatto che $r^2 > 0$, i coefficienti d_1, \dots, d_n, d della 184 soddisfano la condizione

$$\sum_{i=1}^n \frac{d_i^2}{4} - d > 0. \quad (185)$$

Viceversa, un'equazione della forma 184 i cui coefficienti soddisfano la condizione 185 è equazione cartesiana di una sfera \mathbf{S} , il cui centro $C(c_1, \dots, c_n)$ e raggio ρ sono

$$c_i = -\frac{d_i}{2}, \quad r = \sqrt{\sum_{i=1}^n \frac{d_i^2}{4} - d}.$$

Nel caso particolare $n = 2$, la 184 è l'equazione di una circonferenza e ha la forma, nelle indeterminate X, Y ,

$$X^2 + Y^2 + d_1 X + d_2 Y + d = 0 \quad \frac{d_1^2}{4} + \frac{d_2^2}{4} - d > 0. \quad (186)$$

Il centro della circonferenza è

$$C\left(-\frac{d_1}{2}, -\frac{d_2}{2}\right),$$

mentre il raggio è

$$r = \sqrt{\frac{d_1^2}{4} + \frac{d_2^2}{4} - d}.$$

Sia \mathbf{E} un piano euclideo in cui sia assegnata un'orientazione (un piano euclideo orientato) e siano \mathfrak{s} ed \mathfrak{s}' due semirette aventi origine nello stesso punto O . Detti \mathbf{a} e \mathbf{a}' vettori di direzione di \mathfrak{s} ed \mathfrak{s}' rispettivamente, l'angolo $\widehat{\mathbf{a}\mathbf{a}'}$ si dice **angolo orientato tra \mathfrak{s} ed \mathfrak{s}'** e si denota con $\widehat{\mathfrak{s}\mathfrak{s}'}$.

In un piano euclideo orientato \mathbf{E} fissiamo una semiretta \mathfrak{s} di origine O . Per ogni punto $P \neq O$ lo associamo al piano E consideriamo la semiretta \mathfrak{s}_P di origine O e vettore di direzione \overrightarrow{OP} . Lo scalare positivo $\rho = \|\overrightarrow{OP}\|$, e l'angolo $\theta = \widehat{\mathfrak{s}\mathfrak{s}_P}$ si dicono **coordinate polari** di P , e rispettivamente il **modulo** e l'**anomalia** di P rispetto alla semiretta \mathfrak{s} . Diremo che \mathfrak{s} definisce in $\text{text}E$ un **sistema di coordinate polari**. Le coordinate polari (ρ, θ) di un punto $P \neq O$ lo individuano univocamente. Infatti θ individua una semiretta \mathfrak{s}_θ di origine O , e P è il punto di intersezione di \mathfrak{s}_θ con la circonferenza di centro O e raggio ρ . Quindi per ogni coppia di numeri reali (ρ, θ) , con $\rho > 0$, esiste un unico punto P le cui coordinate polari sono ρ e l'angolo definito da θ . Si conviene di estendere le coordinate polari anche al punto O assegnandogli modulo $\rho = 0$ ed anomalia indeterminata. Consideriamo il riferimento cartesiano $O\mathbf{i}\mathbf{j}$ appartenente all'orientazione assegnata in \mathbf{E} , avente origine in O e tale che \mathbf{i} sia il versore di direzione di \mathfrak{s} . ($O\mathbf{i}\mathbf{j}$ è univocamente individuato da queste condizioni). Sia $P \neq O$ un punto avente coordinate polari (ρ, θ) e coordinate cartesiane (x, y) . Allora si ha

$$x = \rho \cos \theta \quad y = \rho \sin \theta. \quad (187)$$

Viceversa:

$$\rho = \sqrt{x^2 + y^2} \quad \theta = \varepsilon(y) \arccos \left(\frac{x}{\sqrt{x^2 + y^2}} \right), \quad (188)$$

in cui si è denotato con $\varepsilon(y) = \pm 1$ a seconda che $y \geq 0$ oppure $y < 0$ (con questa formula si individua una determinazione dell'angolo θ compresa tra $-\pi$ e π). Le 187 e 188 sono le **formule di passaggio da coordinate polari a coordinate cartesiane e viceversa**. La validità delle suddette formule è lasciata al lettore.

Sia \mathbf{E} uno spazio euclideo di dimensione n , $O\mathbf{e}_1 \dots \mathbf{e}_n$ un riferimento cartesiano, e siano

$$A_i(a_1^i, \dots, a_n^i) \in \mathbf{E}, \quad i \in \{0, 1, \dots, n\},$$

punti indipendenti. Il **volume dell' n -parallelepipedo** rispettivamente determinato da A_0, A_1, \dots, A_n è definito come $|\det(M)|$, dove

$$M = \begin{pmatrix} a_1^1 - a_1^0 & a_2^1 - a_2^0 & \cdots & a_n^1 - a_n^0 \\ a_1^2 - a_1^0 & a_2^2 - a_2^0 & \cdots & a_n^2 - a_n^0 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^n - a_1^0 & a_2^n - a_2^0 & \cdots & a_n^n - a_n^0 \end{pmatrix}.$$

In particolare, per $n = 1, 2, 3$ parleremo di **lunghezza**, **area**, **volume** di un segmento, di un parallelogramma, di un parallelepipedo rispettivamente. La lunghezza di un segmento $A(a)B(b)$ di una retta euclidea è $|b - a|$, quindi coincide con $d(A, B)$. L'area del parallelogramma individuato da $A(a_1, a_2)$, $B(b_1, b_2)$, $C(c_1, c_2)$ in un piano euclideo è

$$\begin{vmatrix} b_1 - a_1 & b_2 - a_2 \\ c_1 - a_1 & c_2 - a_2 \end{vmatrix}.$$

Il volume del parallelepipedo individuato da $A(a_1, a_2, a_3)$, $B(b_1, b_2, b_3)$, $C(c_1, c_2, c_3)$, $D(d_1, d_2, d_3)$ in uno spazio euclideo di dimensione 3 è

$$\begin{vmatrix} b_1 - a_1 & b_2 - a_2 & b_3 - a_3 \\ c_1 - a_1 & c_2 - a_2 & c_3 - a_3 \\ d_1 - a_1 & d_2 - a_2 & d_3 - a_3 \end{vmatrix}.$$

Osserviamo che la definizione di volume di un n -parallelepipedo non dipende dalla scelta del riferimento cartesiano $O\mathbf{e}_1 \dots \mathbf{e}_n$. Infatti le righe della matrice M sono le coordinate dei vettori $\overrightarrow{A_0A_1}, \overrightarrow{A_0A_2}, \dots, \overrightarrow{A_0A_n}$ rispetto alla base $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$; se $O\mathbf{f}_1 \dots \mathbf{f}_n$ è un altro riferimento cartesiano, le coordinate di $\overrightarrow{A_0A_1}, \overrightarrow{A_0A_2}, \dots, \overrightarrow{A_0A_n}$ rispetto alla base $\{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ si ottengono dalle precedenti moltiplicando per una matrice ortogonale. Quindi la matrice N analoga di M nel riferimento $O\mathbf{f}_1 \dots \mathbf{f}_n$ si ottiene da M moltiplicandola a destra per una matrice ortogonale, che ha determinante uguale a ± 1 ; dunque

$$|\det(M)| = |\det(N)|.$$

Sottoinsieme limitato

Definizione 9.14. Sia \mathbf{E} uno spazio euclideo. Un sottoinsieme $\mathbf{S} \subset \mathbf{E}$ si dice **limitato** se è contenuto in un disco, cioè

$$\exists C \in \mathbf{E} \wedge r > 0 : \mathbf{S} \subset D(C, r).$$

Un **poliedro convesso** è un sottoinsieme limitato di \mathbf{E} che non è contenuto in un sottospazio affine proprio di \mathbf{E} e che è l'intersezione di un numero finito di semispazi. Un poliedro convesso è un insieme convesso perché lo è ogni semispazio. La dimensione di \mathbf{E} è detta **dimensione del poliedro**. Se $\dim(\mathbf{E}) = 1$ si ottiene un segmento. Se $\dim(\mathbf{E}) = 2, 3$ un poliedro convesso si dice **poligono convesso** e **solido convesso** rispettivamente.

Lasciamo al lettore di introdurre in modo appropriato tutte le nozioni elementari relative ai poligoni convessi, come quelle di *vertice*, *lato* ed *angolo*, di *n -agono regolare*, di *lati adiacenti* o *consecutivi* ecc. in analogia con quanto viene fatto in geometria euclidea elementare.

Supponiamo $\dim(\mathbf{E}) = 3$ e sia $\Pi \subset \mathbf{E}$ un solido convesso. Se \mathcal{S} è un piano di \mathbf{E} tale che Π sia contenuto in uno dei due semispazi definiti da \mathcal{S} , allora abbiamo le seguenti possibilità:

- $\mathcal{S} \cap \Pi = \emptyset$;
- $\mathcal{S} \cap \Pi$ è un punto, che si dice **vertice** di Π ;
- $\mathcal{S} \cap \Pi$ è un segmento, che si dice **spigolo** (o **lato**) di Π ;
- $\mathcal{S} \cap \Pi$ è un poligono, che si dice **faccia** di Π .

Dal fatto che Π è intersezione di un numero finito di semispazi segue che esso possiede un numero finito v di vertici, s di spigoli, ed f di facce. È facile vedere che ogni spigolo è lato di due facce e ogni vertice è vertice di almeno tre facce e di altrettanti spigoli.

Per ogni solido convesso Π sussiste la seguente relazione tra v , s ed f :

$$v - s + f = 2. \quad (189)$$

Daremo una dimostrazione della 189 che è simile a quella originale di Eulero. Per semplificare l'argomentazione supporremo che sia possibile costruire il solido Π partendo da una faccia e aggiungendone poi via via altre, in modo che ad ogni nuova faccia che si aggiunge abbia solo lati **adiacenti** in comune con quelle precedentemente iscritte.

Si osservi che necessariamente $f \geq 3$. Ad ogni stadio del procedimento denotiamo con Φ il numero $v - s + f - 1$. Per una sola faccia si ha $v = 0$. Procediamo per induzione sul numero di facce iscritte, dimostrando che finché il poliedro non è stato completato, si ha $\Phi = 0$. Supponiamo che ciò sia vero un dato stadio della costruzione in cui restano da iscrivere almeno due facce ancora. Aggiungendo una nuova faccia F avente p lati, di cui q consecutivi siano in comune con le precedenti; pertanto $q + 1$ vertici di F appartengono alle precedenti facce. Abbiamo quindi aggiunto 1 nuova faccia, $p - q$ nuovi vertici, $p - q$ nuovi spigoli. Denotando con Φ' la quantità corrispondente di Φ relativa alla nuova configurazione, si ha

$$\Phi' = \Phi + (p - q) - (p - q) + 1 = \Phi + 1 = 0 + 1 = 1 = \Phi, \quad (190)$$

come asserito. Osserviamo che quando si aggiunge l'ultima faccia non si modifica né il numero dei vertici né quello degli spigoli, mentre il numero delle facce aumenta di 1. Quindi per Π si ha $\Phi = 1$, cioè la (19.18).

Gli studi sulla estensione 3-dimensionale dei poligoni regolari detti “solidi regolari”. Un *solido regolare* è un solido convesso avente facce poligonali tutti uguali tra loro. È un fatto notevole che, diversamente da quello che avviene per i poligoni regolari, esistono solo un numero finito, precisamente 5, di classi di similitudine di solidi regolari (per la definizione di similitudine cfr. 20.10(2)): il *tetraedro*, l'*ottaedro*, il *cubo*, il *dodecaedro*, l'*icosaedro* (fig. 19.4).

Questi solidi erano noti fin dall'antichità. Poiché la scuola platonica se ne interessò, in particolare Platone ne parla nel *Timeo*, essi vengono anche chiamati *solidi platonici*. Teeteto li studiò sistematicamente attorno al 380 a.C. e di essi tratta il XIII libro degli *Elementi* di Euclide. (Per maggiori dettagli rimandiamo il lettore a [7]).

10 Diagonalizzazione di operatori simmetrici

Nelle pagine precedenti abbiamo introdotto due diverse relazioni di equivalenza tra matrici quadrate: la similitudine e la congruenza. Due matrici $A, B \in M_n(\mathbb{K})$, $n \geq 1$, sono dette simili (rispettivamente congruenti) se

$$\exists M \in GL_n(\mathbb{K}) : B = M^{-1}AM \quad (B = {}^tMAM).$$

La similitudine è stata introdotta allo scopo di studiare le matrici che rappresentano un operatore su di uno spazio vettoriale rispetto a due diverse basi; la congruenza è stata invece definita per descrivere le matrici di una forma bilineare rispetto a basi diverse.

In corrispondenza alle due nozioni si hanno due diversi problemi di diagonalizzazione, che possono così enunciarsi: data $A \in M_n(\mathbb{K})$, trovare una matrice diagonale $B \in M_n(\mathbb{K})$ simile (oppure congruente) ad A .

Il secondo problema, quello dell'esistenza di matrici diagonali in una data classe di congruenza, equivalente al problema della diagonalizzazione delle forme bilineari, è risolubile se ci si limita a considerare forme bilineari simmetriche, e cioè matrici A simmetriche, per le quali si afferma il teorema 6.1.

Come sappiamo, facili esempi mostrano che il primo dei due problemi non ammette soluzione in generale, cioè non tutte le classi di similitudine contengono una matrice diagonale.

In questo paragrafo considereremo un'altra questione, più particolare ma molto importante in geometria euclidea, vale a dire il problema di diagonalizzare matrici simmetriche reali per mezzo di matrici ortogonali.

Se $A \in M_n(\mathbb{R})$ e $M \in O$, si ha

$$M^{-1}AM = {}^tMAM \quad (191)$$

e quindi la matrice 191 è simultaneamente simile e congruente ad A . Parlando di diagonalizzazione di una matrice per mezzo di matrici ortogonali, non è dunque necessario specificare se ci si riferisce alla similitudine o alla congruenza perché le due nozioni sono equivalenti. Il limitarsi a considerare le matrici $M \in O(n)$ è equivalente a considerare, in uno spazio vettoriale euclideo di dimensione finita \mathbf{V} , solo basi ortonormali; quindi la diagonalizzabilità di una matrice simmetrica A per mezzo di matrici ortogonali significa che sia la forma quadratica definita da A che l'operatore di matrice A rispetto a una base ortonormale di \mathbf{V} sono diagonalizzabili in una base ortonormale.

Una semplice, ma fondamentale, proprietà delle matrici simmetriche reali è descritta dal seguente lemma.

Lemma 10.1. *Il polinomio caratteristico di una matrice simmetrica $A \in M_n(\mathbb{R})$ possiede solo radici reali.*

Proof. Possiamo considerare A come una matrice di numeri complessi e quindi come un operatore

$$T_A : \mathbb{C}^n \rightarrow \mathbb{C}^n.$$

Sia $\lambda \in \mathbb{C}$ una radice del polinomio caratteristico di A , e sia $\mathbf{x} \in \mathbb{C}^n$ un corrispondente autovettore. Si ha

$$A\mathbf{x} = \lambda\mathbf{x}. \quad (192)$$

Prendendo i complessi coniugati di primo e secondo membro, si ha anche

$$A\bar{\mathbf{x}} = \bar{\lambda}\bar{\mathbf{x}}. \quad (193)$$

Consideriamo lo scalare ${}^t\bar{\mathbf{x}}A\mathbf{x}$, e scriviamolo in due modi diversi utilizzando la 192 e la 193:

$${}^t\bar{\mathbf{x}}A\mathbf{x} = {}^t\bar{\mathbf{x}}(A\mathbf{x}) = {}^t\bar{\mathbf{x}}\lambda\mathbf{x} = \lambda{}^t\bar{\mathbf{x}}\mathbf{x} \quad (194)$$

$${}^t\bar{\mathbf{x}}A\mathbf{x} = ({}^t\bar{\mathbf{x}}A)\mathbf{x} = {}^t(A\bar{\mathbf{x}})\mathbf{x} = {}^t(\bar{\lambda}\bar{\mathbf{x}})\mathbf{x} = \bar{\lambda}{}^t\bar{\mathbf{x}}\mathbf{x}. \quad (195)$$

Osservando che

$${}^t\bar{\mathbf{x}}\mathbf{x} = \sum_{i=1}^n \bar{x}_i x_i$$

è un numero reale positivo, perché $\mathbf{x} \neq \mathbf{0}$, dalle 194 e 195 deduciamo che $\lambda = \bar{\lambda}$, cioè che λ è reale. □

Teorema Spettrale

Teorema 10.2. *Siano \mathbf{V} uno spazio vettoriale euclideo di dimensione finita e $T : \mathbf{V} \rightarrow \mathbf{V}$ un operatore simmetrico. Esiste una base ortonormale di \mathbf{V} rispetto alla quale la matrice che rappresenta T è diagonale.*

Proof. Procediamo per induzione su $n = \dim(\mathbf{V})$. Se $n = 1$ non c'è niente da dimostrare; supponiamo quindi $n \geq 2$ e che il teorema sia vero per spazi di dimensione $n - 1$. Poiché l'operatore T è simmetrico, il polinomio caratteristico di T possiede radici reali, per il lemma 191. Quindi T possiede un autovalore λ ; sia \mathbf{e}_1 un corrispondente autovettore, che possiamo supporre di norma 1, e sia $\mathbf{U} = \{\mathbf{e}_1\}^\perp$ il complemento ortogonale di \mathbf{e}_1 . Per ogni $\mathbf{u} \in \mathbf{U}$ si ha

$$\langle T(\mathbf{u}), \mathbf{e}_1 \rangle = \langle \mathbf{u}, T(\mathbf{e}_1) \rangle = \langle \mathbf{u}, \lambda \mathbf{e}_1 \rangle = \lambda \langle \mathbf{u}, \mathbf{e}_1 \rangle = \lambda \cdot 0 = 0,$$

e quindi $T(\mathbf{u}) \in \mathbf{U}$, cioè T induce un operatore $T_{\mathbf{U}} : \mathbf{U} \rightarrow \mathbf{U}$. Poiché $T_{\mathbf{U}}(\mathbf{u}) = T(\mathbf{u})$ per ogni $\mathbf{u} \in \mathbf{U}$, l'operatore $T_{\mathbf{U}}$ è simmetrico. Per l'ipotesi induttiva, \mathbf{U} possiede una base ortonormale $\{\mathbf{e}_2, \dots, \mathbf{e}_n\}$ che diagonalizza $T_{\mathbf{U}}$. Allora $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ è una base ortonormale di \mathbf{V} che diagonalizza T . □