# Capitolo 1

# Sistemi lineari

## 1.1 Equazioni lineari

Una *equazione lineare a n incognite*, è una equazione del tipo seguente:

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b\,, \tag{1.1}$$

dove $a_1, \cdots, a_n, b$ sono numeri reali. I simboli $x_1, \cdots, x_n$ sono detti *incognite* dell'equazione, i numeri reali $a_1, \cdots, a_n$ *coefficienti* e $b$ sarà chiamato *termine noto*.

Per semplificare la notazione, useremo spesso il simbolo di sommatoria, così da scrivere l'equazione lineare nel modo seguente:

$$\sum_{i=1}^{n} a_i x_i = b\,. \tag{1.2}$$

Consideriamo ora la seguente definizione:

**Definizione 1.1.1.** Una soluzione di una equazione lineare a $n$ incognite $\sum_{i=1}^{n} a_i x_i = b$, è una $n$-upla ordinata (leggasi ennupla) di numeri reali, ossìa un elemento $(x_1^0, \cdots, x_n^0)$ di $\mathbb{R}^n$ tale che si abbia:

$$\sum_{i=1}^{n} a_i x_i^0 = b\,.$$

Questo significa che se nell'equazione, al posto delle incognite mettiamo gli elementi della $n$-upla e svolgiamo i calcoli, otteniamo il termine noto.

Vediamo ora alcuni esempi di equazioni lineari e di soluzioni di tali equazioni:

**Esempio 1.1.2.** Consideriamo l'equazione in 4 incognite:

$$2x_1 - x_2 + \sqrt{2}x_4 = 0\,.$$

Si osservi innanzitutto che il coefficiente di $x_3$ è nullo e perciò tale incognita non appare nell'equazione ed inoltre il termine noto è anch'esso nullo. La quaterna di numeri reali $(0, 0, 0, 0)$ è una soluzione, infatti si ha:

$$2 \cdot 0 - 1 \cdot 0 + \sqrt{2} \cdot 0 = 0 + 0 + 0 = 0\,,$$

che è pari al termine noto. Allo stesso modo si verifica che le quaterne $(1, 2, 0, 0)$ e $(\sqrt{2}, 0, 5, -2)$ sono entrambe soluzioni.

**Esempio 1.1.3.** L'equazione a 5 incognite

$$0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 = 0\,,$$

ammette come soluzione qualunque 5-upla di numeri reali.

**Esempio 1.1.4.** L'equazione ad un'incognita (che denotiamo con $x$)

$$5x = -3\,,$$

ammette come unica soluzione il numero reale $-\frac{3}{5}$. Infatti si verifica sostituendo che $-\frac{3}{5}$ è una soluzione. Inoltre, se $x^0$ è una soluzione di tale equazione allora si ha

$$5x^0 = -3$$

e, dividendo entrambi i membri per 5, si ottiene $x^0 = -\frac{3}{5}$.

Prima di analizzare più dettagliatamente alcuni casi particolari, ci occorre ancora la seguente definizione:

**Definizione 1.1.5.** Una equazione lineare a $n$ incognite è detta *omogenea* se il suo termine noto è nullo.

In tal caso si verifica facilmente che la $n$-upla in cui ogni elemento è zero, è una soluzione dell'equazione. Tale $n$-upla verrà chiamata *soluzione nulla*.

Data un'equazione lineare a $n$ incognite, ci interessa determinare e descrivere l'insieme formato da tutte le soluzioni di tale equazione. Tale insieme sarà per definizione un sottoinsieme di $\mathbb{R}^n$ e verrà indicato con $\mathcal{S}$. Di seguito analizzeremo in dettaglio la generica equazione ad una e a due incognite per poi analizzare il caso generale.

### 1.1.1 Equazione lineare ad una incognita

In questo caso denoteremo con $x$ l'incognita, $a$ il coefficiente di $x$ e $b$ il termine noto. L'equazione sarà allora:

$$ax = b\,.$$

Nella ricerca delle soluzioni si presentano vari casi a seconda dei valori assunti dal coefficiente e dal termine noto:

Caso $a \neq 0$ In questo caso, dividendo entrambi i membri dell'equazione per $a$ (dato che $a \neq 0$) si ha che l'unica soluzione è data dal numero reale $\frac{b}{a}$ e quindi $\mathcal{S} = \{\frac{b}{a}\}$.

Caso $a = 0$ Se $a$ è nullo, l'equazione diventa

$$0x = b$$

e quindi abbiamo due possibilità. Se $b$ è non nullo, l'equazione non ha soluzioni poichè qualunque numero reale $x^0$ moltiplicato per 0 dà 0 e quindi non può essere uguale a $b$; in tal caso si ha $\mathcal{S} = \emptyset$. Se $b$ è nullo ogni numero reale è soluzione dell'equazione e allora avremo $\mathcal{S} = \mathbb{R}$.

### 1.1.2 Equazione lineare a due incognite

Indicheremo con $x$ e $y$ le due incognite, con $a$ e $b$ i rispettivi coefficienti e con $c$ il termine noto cosicché l'equazione diventa:

$$ax + by = c\,.$$

Come per l'equazione ad una incognita distingueremo due casi. Se entrambi i coefficienti sono nulli avremo $0x+0y = c$ e quindi se $c \neq 0$ non esistono soluzioni, mentre se $c = 0$ l'insieme delle soluzioni è $\mathbb{R}$.

Se invece almeno uno dei coefficienti è non nullo, supponiamo che sia $a$, allora, dividendo ambo i membri per $a$ si ottiene $x + \frac{b}{a}y = \frac{c}{a}$ e, portando al secondo membro $\frac{b}{a}y$, si arriva a

$$x = -\frac{b}{a}y + \frac{c}{a}\,.$$

Quindi una qualunque soluzione $(x^0, y^0) \in \mathbb{R}^2$ dell'equazione ha la forma

$$\left(-\frac{b}{a}y^0 + \frac{c}{a}, y^0\right)$$

e viceversa ogni coppia della forma $\left(-\frac{b}{a}y^0 + \frac{c}{a}, y^0\right)$ per un qualunque numero reale $y^0$ soddisfa l'equazione. Possiamo quindi scrivere l'insieme delle soluzioni come segue:

$$\mathcal{S} = \left\{\left(-\frac{b}{a}t + \frac{c}{a}, t\right) \in \mathbb{R}^2 \middle| t \in \mathbb{R}\right\}\,.$$

Diremo che le soluzioni sono *infinito a uno*, ossìa che dipendono da un parametro (in questo caso abbiamo usato $t$ come parametro).

Nel caso in cui $b$ sia non nullo, sarà possibile descrivere $\mathcal{S}$ in modo simile, in cui il parametro apparirà come primo elemento della coppia.

### 1.1.3 Equazioni lineari: caso generale

Il caso dell'equazione ad $n$ incognite

$$\sum_{i=1}^{n} a_i x_i = b\,.$$

è una semplice generalizzazione di quanto fatto per due incognite e quindi si avrà:

- Se i coefficienti sono tutti nulli, e il termine noto è nullo, avremo $\mathcal{S} = \mathbb{R}^n$, ciè quella che viene chiamata *identità*.
- Se i coefficienti sono tutti nulli, e il termine noto è non nullo, avremo $\mathcal{S} = \emptyset$ e diremo che l'equazione non ammette soluzioni oppure che è *incompatibile*.
- Se almeno uno dei coefficienti è diverso da zero sarà possibile descrivere l'insieme delle soluzioni utilizzando $n-1$ parametri (uno in meno del numero delle incognite). Ad esempio, se $a_1 \neq 0$, operando come per il caso a due incognite avremo

  $$x_1 = -\frac{1}{a_1}\sum_{i=2}^{n} a_i x_i + \frac{b}{a_1}$$

  e quindi $x_2, \cdots, x_n$ possono assumere qualunque valore reale mentre $x_1$ deve avere la forma suddetta. L'insieme $\mathcal{S}$ sarà allora descritto come segue:

  $$\mathcal{S} = \left\{\left(-\frac{1}{a_1}\sum_{i=2}^{n} a_i t_i + \frac{b}{a_1},\ t_2, \cdots, t_n\right) \in \mathbb{R}^n \middle| t_2, \cdots, t_n \in \mathbb{R}\right\}.$$

**Esempio 1.1.6.** Consideriamo l'equazione a 4 incognite

$$3x_1 + x_3 - x_4 = 5\,.$$

Poiché il coefficiente di $x_3$ è diverso da zero, possiamo isolare $x_3$ e portare i restanti termini al secondo membro e ottenere

$$x_3 = 5 - 3x_1 + x_4$$

e quindi l'insieme delle soluzioni sarà

$$\mathcal{S} = \left\{(t_1, t_2, 5 - 3t_1 + t_4, t_4) \in \mathbb{R}^4 \middle| t_1, t_2, t_4 \in \mathbb{R}\right\}.$$

Quindi si avranno $\infty^3$ (si legga infinito a tre) soluzioni.

## 1.2 Sistemi lineari

Ora ci occuperemo dei sistemi di equazioni e quindi daremo la seguente definizione:

**Definizione 1.2.1.** Un *sistema lineare a n incognite e m equazioni* è un insieme di $m$ equazioni lineari ad $n$ incognite.

Per avere un sistema a $n$ incognite e $m$ equazioni ci occorrono dunque $m \times n$ coefficienti e $n$ termini noti. Per poter scrivere più concisamente un sistema, siccome ci occorre identificare quali sono i coefficienti e i termini noti di ogni equazione, utilizzeremo un doppio indice per il coefficiente generico, in modo tale che il primo indice identifichi l'equazione e il secondo la corrispondente incognita.

Ad esempio, se il sistema ha 2 equazioni e 3 incognite, ci occorrono $2 \times 3$ coefficienti che indicheremo ad esempio con $a_{ij}$, dove l'indice $i$ assume i valori $1, 2$, mentre $j$ assume valori $1, 2, 3$. Allora avremo che $a_{11}$ è il coefficiente relativo a $x_1$ della prima equazione, $a_{12}$ è il coefficiente relativo a $x_2$ della prima equazione e $a_{13}$ è il coefficiente relativo a $x_3$ della prima equazione. Per la seconda equazione avremo invece che $a_{21}$ è il coefficiente relativo a $x_1$ , $a_{22}$ è il coefficiente relativo a $x_2$ e $a_{23}$ è il coefficiente relativo a $x_3$. Per i termini noti ci servirà un unico indice dato che per identificarli ci occorre solo il numero di equazione.

In generale dunque, se abbiamo $n$ incognite ed $m$ equazioni e indichiamo con $a_{ij}$ i coefficienti e $b_i$ i termini noti, per descrivere il sistema scriveremo:

$$\sum_{j=1}^{n} a_{ij}x_j = b_i\,,\ i = 1, \cdots, m. \tag{1.3}$$

Quindi per $i$ fissato avremo la $i$-esima equazione del sistema

$$\sum_{j=1}^{n} a_{ij}x_j = b_i\,.$$

La notazione classica per scrivere un sistema per esteso (che useremo in particolare quando il sistema è dato esplicitamente):

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = b_1 \\ \qquad\qquad\vdots & \ \ \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n & = b_i \\ \qquad\qquad\vdots & \ \ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = b_m \end{cases} \tag{1.4}$$

**Definizione 1.2.2.** Una soluzione del sistema (1.4) è una $n$-upla di numeri reali che sia soluzione di ogni equazione del sistema.

*Osservazione* 1.2.3. Se $\mathcal{S}_i$ indica l'insieme delle soluzioni della $i$-esima equazione del sistema, l'insieme delle soluzioni del sistema sarà $\mathcal{S} = \cap_{i=1}^{m} \mathcal{S}_i$.

**Definizione 1.2.4.** Un sistema si dice *compatibile* se ammette soluzioni (ossìa se l'insieme delle soluzioni è non vuoto), altrimenti si dice *incompatibile*.

**Esempio 1.2.5.** Il sistema a tre incognite e due equazioni

$$\begin{cases} x - 2y + z = 0 \\ \pi x + \sqrt{5} y - (\ln 5) z = 0 \end{cases}$$

è compatibile poiché entrambe le equazioni ammettono la soluzione nulla.

**Esempio 1.2.6.** l sistema a due incognite e due equazioni

$$\begin{cases} x - y = 0 \\ x - y = 1 \end{cases}$$

non è compatibile. Infatti se $(a, b)$ fosse una soluzione, si dovrebbe avere $a = b$ poiché la coppia $(a, b)$ è soluzione della prima equazione, ma allora la loro differenza sarebbe 0 e quindi la coppia $(a, b)$ non può essere soluzione della seconda equazione.

Di particolare importanza sono i sistemi in cui tutte le equazioni sono omogenee. Diamo quindi la seguente definizione:

**Definizione 1.2.7.** Un sistema è detto *omogeneo* se tutte le sue equazioni sono omogenee.

Poiché una equazione omogenea ammette sempre la soluzione nulla, un sistema omogeneo è sempre compatibile.

### 1.2.1 Risoluzione dei sistemi lineari

Per poter affrontare la risoluzione dei sistemi lineari, ci occorre il concetto di equivalenza di sistemi lineari e di combinazione lineare di equazioni di un sistema:

**Definizione 1.2.8.** Due sistemi sono detti *equivalenti* se ammettono le stesse soluzioni.

*Osservazione* 1.2.9. Dovrebbe esser chiaro che due sistemi equivalenti devono avere lo stesso numero di incognite!

**Definizione 1.2.10.** Dato il sistema lineare (1.4) e $m$ numeri reali $\lambda_1, \cdots, \lambda_m$, è detta *combinazione lineare di coefficienti* $\lambda_1, \cdots, \lambda_m$, l'equazione ottenuta come segue:

$$\lambda_1 \cdot \left( \sum_{j=1}^{n} a_{1j} x_j \right) + \cdots + \lambda_m \cdot \left( \sum_{j=1}^{n} a_{mj} x_j \right) = \lambda_1 \cdot b_1 + \cdots + \lambda_m \cdot b_m \, .$$

L'utilizzo delle combinazioni lineari di equazioni di un sistema permette di passare da un sistema lineare ad uno equivalente. In effetti si ha il seguente risultato, chiamato *mettodo di Gauss-Jordan* per la risoluzione dei sistemi lineari:

**Teorema 1.2.11.** *Se all'equazione di un sistema si somma una combinazione lineare delle altre equazioni, si ottiene un sistema equivalente. Se si moltiplicano entrambi i membri di una equazione per un numero non nullo, si ottiene un sistema equivalente.*

Il metodo di Gauss-Jordan si dimostra molto efficace nella risoluzione dei sistemi lineari e può essere utilizzato anche assieme al metodo di sostituzione.

Per semplificare la dimostrazione del Teorema 1.2.11, ci occorre introdurre alcune notazioni:

- Indicheremo con $x$ la $n$-upla di incognite $(x_1, \cdots, x_n)$ e con $x^0$ la $n$-upla di numeri reali $(x_1^0, \cdots, x_n^0)$;
- il primo membro della $i$-esima equazione è un polinomio omogeneo di grado 1 in $x$ e lo indicheremo con $P_i(x)$;
- scriveremo la $i$-esima equazione come $P_i(x) = b_i$;
- il polinomio $P_i(x)$ valutato in $x^0$ verrà indicato con $P_i(x^0)$ e quindi se $x^0$ è una soluzione del sistema avremo $P_i(x^0) = b_i$;
- dato un numero reale $\lambda$ e un polinomio omogeneo $P(x) = a_1 x_1 + \cdots + a_n x_n$, indicheremo con $(\lambda P)(x)$ il polinomio $(\lambda a_1) x_1 + \cdots + (\lambda a_n) x_n$;
- dati due polinomi omogenei $P(x) = a_1 x_1 + \cdots + a_n x_n$ e $Q(x) = b_1 x_1 + \cdots + b_n x_n$ indicheremo con $(P + Q)(x)$ il polinomio somma, cioè $(P + Q)(x) = (a_1 + b_1) x_1 + \cdots + (a_n + b_n) x_n$.

Si verifica facilmente ed è lasciato come esercizio per il lettore che si ha:

$$(P + Q)(x^0) = P(x^0) + Q(x^0) \, ; \; (\lambda P)(x^0) = \lambda \cdot P(x^0) \, .$$

*Dimostrazione del Teorema 1.2.11.* Consideriamo il sistema lineare

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = b_1 \\ \vdots & \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n & = b_i \\ \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = b_m \end{cases}$$

che riscritto utilizzando le notazioni precedenti diventerà

$$\begin{cases} P_1(x) & = b_1 \\ \vdots & \vdots \\ P_i(x) & = b_i \\ \vdots & \vdots \\ P_m(x) & = b_m \end{cases} . \tag{1.5}$$

Consideriamo ora il sistema ottenuto dal precedente sommando ad una equazione una combinazione lineare delle altre. Per l'Osservazione 1.2.3, scambiando l'ordine delle equazioni si ottiene un sistema equivalente, quindi possiamo supporre di sommare alla prima equazione una combinazione lineare delle altre. Perciò il nuovo sistema sarà

$$\begin{cases} P_1(x) + \lambda_2 P_2(x) + \cdots + \lambda_m P_m(x) & = b_1 + \lambda_2 b_2 + \cdots + \lambda_m b_m \\ \vdots & \vdots \\ P_i(x) & = b_i \\ \vdots & \vdots \\ P_m(x) & = b_m \end{cases} . \tag{1.6}$$

Siano ora $\mathcal{S}_1$ e $\mathcal{S}_2$ gli insiemi delle soluzioni di (1.5) e (1.6) rispettivamente. Vogliamo dimostrare che $\mathcal{S}_1 = \mathcal{S}_2$ ovvero che si ha $\mathcal{S}_1 \subset \mathcal{S}_2$ e $\mathcal{S}_2 \subset \mathcal{S}_1$. Sia dunque $x^0 \in \mathcal{S}_1$. Poiché $x^0$ è soluzione del sistema (1.5), esso è soluzione di ogni equazione del sistema e quindi si ha $P_i(x^0) = b_i$ per $i = 1, 2, \cdots, m$. Ma questo implica che $x^0$ è soluzione di tutte le equazioni del sistema 1.6, tranne eventualmente della prima. Verifichiamo quindi che $x^0$ è anche soluzione delle prima equazione, che può essere scritta anche come

$$(P_1 + \lambda_2 P_2 + \cdots + \lambda_m P_m)(x) = b_1 + \lambda_2 b_2 + \cdots + \lambda_m b_m .$$

Se valutiamo il polinomio $(P_1 + \lambda_2 P_2 + \cdots + \lambda_m P_m)(x)$ in $x^0$ otteniamo:

$$P_1(x^0) + \lambda_2 \cdot P_2(x^0) + \cdots + \lambda_m \cdot P_m(x^0) = b_1 + \lambda_2 b_2 + \cdots + \lambda_m b_m .$$

poiché $P_i(x^0) = b_i$. Quindi $x^0 \in \mathcal{S}_2$. L'altra inclusione si dimostra analogamente ed è lasciata per esercizio così come la seconda proprietà dell'enunciato del teorema. $\square$

**Esempio 1.2.12.** Consideriamo il sistema a tre incognite

$$\begin{cases} x - 2y + z &= 0 \\ -x + 2y - 5z &= 0 \\ x - y &= 0 \end{cases} .$$

Se alla prima riga sommiamo la seconda moltiplicata per 1, otteniamo il seguente sistema equivalente

$$\begin{cases} - 4z &= 0 \\ -x + 2y - 5z &= 0 \\ x - y &= 0 \end{cases} .$$

Sommando la terza equazione alla seconda si ottiene

$$\begin{cases} - 4z &= 0 \\ y - 5z &= 0 \\ x - y &= 0 \end{cases} .$$

Dividendo la prima equazione per $-4$ e riordinando le equazioni si ottiene

$$\begin{cases} x - y &= 0 \\ y - 5z &= 0 \\ z &= 0 \end{cases} .$$

A questo punto si può sommare alla seconda la terza moltiplicata per 5 per ottenere

$$\begin{cases} x - y &= 0 \\ y &= 0 \\ z &= 0 \end{cases} .$$

Infine sommando alla prima la seconda si ha

$$\begin{cases} x &= 0 \\ y &= 0 \\ z &= 0 \end{cases} .$$

e quindi l'unica soluzione del sistema è $(0, 0, 0)$.

## Esercizi

1. Si risolvano i seguenti sistemi lineari:

   - $$\begin{cases} x_1 + x_2 + x_3 = 1 \\ 2x_1 + 2x_2 + x_3 = 0 \\ x_1 + x_2 + 2x_3 = -1 \end{cases}$$

- $$\begin{cases} 2x_1 - x_2 - x_3 - 4x_4 = 9 \\ 4x_1 - 3x_3 - x_4 = 0 \\ 8x_1 - 2x_2 - 5x_3 - 9x_4 = 18 \end{cases}$$
- $$\begin{cases} 2x_1 + x_2 - x_3 = 1 \\ x_1 + 2x_2 - 2x_3 = 0 \\ 3x_1 - x_2 + 2x_3 = -1 \\ x_1 - x_2 + x_3 = 1 \end{cases}$$

2. Discutere e risolvere il seguente sistema lineare al variare dei parametri reali $a, b$:
$$\begin{cases} x_1 + x_2 + ax_3 = 1 \\ x_1 + 2x_2 + bx_3 = 3 \\ x_2 + x_3 = 2 \end{cases}$$

3. Si calcoli la combinazione lineare di coefficienti $(2, 2, -1)$ delle tre equazioni del primo sistema dell'esercizio 1. Si faccia lo stesso per il secondo sistema ma con coefficienti $(2, 1, -1)$.

4. Si dica quali dei sistemi del primo esercizio sono equivalenti.

# Capitolo 2

# Matrici

In questo capitolo ci occuperemo dello studio delle matrici ad entrate reali, benché gran parte delle definizioni e risultati siano validi per un *campo* qualunque e quindi in particolare quando si considera il campo dei numeri complessi.

**Definizione 2.0.13.** Una *matrice con n righe e m colonne ad entrate reali* è una tabella con $n$ righe ed $m$ colonne di numeri reali.

I numeri reali che si trovano nella tabella verranno chiamati *entrate* della matrice. Solitamente indicheremo le matrici con lettere maiuscole, mentre se vogliamo scrivere una matrice per esteso scriveremo le entrate tra due parentesi tonde, come nei seguenti esempi:

- esempio di matrice con 3 righe e 3 colonne

$$A = \begin{pmatrix} 0 & 1 & -2 \\ -5 & -\pi & 0 \\ 1 & -1 & -7 \end{pmatrix};$$

- esempio di matrice con 4 righe e 3 colonne

$$A = \begin{pmatrix} 0 & 1 \\ -\pi & 0 \\ -1 & -7 \\ 0 & 1 \\ \sqrt{5} & -\sqrt{\pi} \end{pmatrix};$$

- esempio di matrice con 2 righe e 3 colonne

$$A = \begin{pmatrix} 0 & 1 & 10 \\ -\pi & 0 & e \end{pmatrix}.$$

Indicheremo con $\mathcal{M}_{m\times n}(\mathbb{R})$ l'insieme delle matrici con $m$ righe e $n$ colonne. Talvolta, visto che considereremo quasi esclusivamente matrici reali, denoteremo questo insieme semplicemente con $\mathcal{M}_{m\times n}$

Se $A \in \mathcal{M}_{m\times n}$, le sue entrate saranno in totale $m \times n$ e le indicheremo con una lettera minuscola con due indici in basso, ad esempio $a_{ij}$, dove l'indice $i$ assume valori (nell'insieme dei numeri naturali) da 1 a $m$ e $j$ da 1 ad $n$. Scriveremo quindi

$$A = (a_{ij}).$$

L'indice $i$ (quello che sta al primo posto) sarà chiamato *indice di riga*, mentre $j$ (al secondo posto) *indice di colonna*.

Per i ragionamenti teorici talvolta sarà utile, data una matrice $A$, indicare la sua entrata che sta alla $i$-esima riga e $j$-esima colonna, con $[A]_{ij}$.

Un'altra notazione molto utile è quella che descrive la matrice come insieme di righe o di colonne. Scriveremo cioè $A = (A_1, A_2, \cdots, A_m)$ (usando gli indici in basso) dove $A_1$ è la prima riga, $A_2$ la seconda e così via, oppure $A = (A^1, A^2, \cdots, A^n)$ (usando gli indici in alto) dove $A^j$ è la $j$-esima colonna.

## 2.1 Matrici speciali

Prima di introdurre le operazioni fondamentali tra matrici, vedremo alcune matrici di tipo speciale, che sono fondamentali nella teoria.

**Definizione 2.1.1.** Due matrici $A, B \in \mathcal{M}_{m\times n}$ sono *uguali* se $[A]_{ij} = [B]_{ij}$ per ogni $i, j$.

**Definizione 2.1.2.** Data una matrice $A \in \mathcal{M}_{m\times n}$, sarà detta *trasposta* di $A$ la matrice indicata con $A^t$ tale che $A^t \in \mathcal{M}_{n\times m}$ e $[A^t]_{ij} = [A]_{ji}$ per ogni $i, j$. La matrice *opposta* di $A$, indicata con $-A$ è la matrice di $\mathcal{M}_{n\times m}$ tale che $[-A]_{ij} = -[A]_{ji}$.

**Esempio 2.1.3.** La trasposta della matrice

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

è la matrice

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

**Definizione 2.1.4.** Una matrice *quadrata di ordine n* è una matrice avente $n$ righe ed $n$ colonne.

Indicheremo con $\mathcal{M}_n$ l'insieme delle matrici quadrate di ordine $n$.

Data una matrice quadrata $A = (a_{ij})$ di ordine $n$, la $n$-upla di elementi $(a_{11}, \cdots, a_{nn})$ è detta *diagonale principale* di $A$, mentre la $n$-upla $(a_{n1}, a_{(n-1)2}, \cdots, a_{1n})$ è detta *diagonale secondaria*.

**Definizione 2.1.5.** La matrice $O_{m\times n} \in \mathcal{M}_{m\times n}$ aventi entrate tutte nulle è detta *matrice nulla* di $\mathcal{M}_{m\times n}$.

Nel caso delle matrici quadrate esiste un'altra matrice particolarmente importante chiamata *matrice identità*. Per poterne dare una definizione più concisa ci occorre la seguente definizione:

**Definizione 2.1.6.** Il *simbolo di Kronecker*[1] o *delta di Kronecker* indicato con $\delta_{ij}$, dove $i, j$ sono numeri naturali, è definito come segue:

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

**Definizione 2.1.7.** La *matrice identità di ordine n*, denotata con $I_n$, è la matrice quadrata di ordine $n$ tale che $[I_n]_{ij} = \delta_{ij}$.

Quindi la matrice identità ha tutti 1 nella diagonale principale e 0 altrove. Quando sarà chiaro l'ordine della matrice identità, la indicheremo semplicemente con $I$.

**Esempio 2.1.8.** La matrice identità di ordine 1 è la matrice (1), quella di ordine due è la matrice $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, mentre quella di ordine 3 è la matrice $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Più in generale si può parlare di matrici *diagonali*, cioè matrici quadrate le cui entrate, tranne eventualmente quelle della diagonale principale, sono nulle. Più precisamente abbiamo:

**Definizione 2.1.9.** Una matrice quadrata $A = (a_{ij})$ è detta *diagonale* se $a_{ij} = 0$ per ogni $i \neq j$.

In particolare la matrice identità e la matrice nulla di ordine $n$ sono diagonali.

Spesso è utile scrivere una matrice diagonale di ordine $n$ mettendo in evidenza solo le entrate della diagonale. Scriveremo quindi $diag(\lambda_1, \cdots, \lambda_n)$, intendendo la matrice diagonale i cui elementi della diagonale principale sono $(\lambda_1, \cdots, \lambda_n)$.

Possiamo ancora generalizzare le matrici diagonali, nel modo seguente:

[1]*Leopold Kronecker*: matematico tedesco (Liegnitz, 1823 - Berlino, 1981).

**Definizione 2.1.10.** Una matrice quadrata $A = (a_{ij})$ è detta *triangolare superiore* (rispettivamente *triangolare inferiore*) se $a_{ij} = 0$ quando $i > j$ (rispettivamente $i < j$).

**Esempio 2.1.11.** Le seguenti matrici sono triangolari superiori

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

mentre sono triangolari inferiori le matrici

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

Di particolare interesse sono le matrici definite come segue:

**Definizione 2.1.12.** Una matrice quadrata $A = (a_{ij})$ è detta *simmetrica* (rispettivamente *antisimmetrica*) se per ogni $i, j$ si ha $a_{ij} = a_{ji}$ o equivalentemente se $A^t = A$ (rispettivamente $a_{ij} = -a_{ji}$ o equivalentemente se $A^t = -A$).

Le matrici identità e le matrici nulle di ordine qualunque sono simmetriche.

**Esempio 2.1.13.** La matrice $\begin{pmatrix} 1 & -2 & 3 \\ -2 & 0 & 0 \\ 3 & 0 & 1 \end{pmatrix}$ è simmetrica.

**Esempio 2.1.14.** La matrice $\begin{pmatrix} 0 & -2 & 3 \\ 2 & 0 & 0 \\ -3 & 0 & 0 \end{pmatrix}$ è antisimmetrica, mentre non lo è la matrice $\begin{pmatrix} 0 & -2 & 3 \\ 2 & 0 & 0 \\ -3 & 0 & 1 \end{pmatrix}$.

Terminiamo questa sezione con una definizione che zarà utile più avanti:

**Definizione 2.1.15.** La *traccia* di una matrice quadrata $A = (a_{ij})$, di ordine $n$, è il numero reale definito come segue:

$$\mathrm{tr}A := \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \cdots + a_{nn}.$$

La traccia di una matrice è quindi la somma degli elementi delle diagonale principale.

### Esercizi

1. Si scrivano per esteso tutte le matrici diagonali di ordine 2 e 3 con entrate 1 e 0.

2. Si dimostri che le entrate della diagonale principale di una matrice antisimmetrica sono nulle.

3. Si scriva la forma generale delle matrici simmetriche di ordine $Si$ e delle matrici antisimmetriche di ordine 3.

4. Si dimostri che l'unica matrice antisimmetrica di ordine 1 è la matrice nulla. Si mostri con un controesempio che questo non vale per le matrici antisimmetriche di ordine 2.

5. Si dimostri che tutte le matrici quadrate di ordine uno sono simmetriche.

6. Si calcoli la traccia delle seguenti matrici:
$$\begin{pmatrix} 0 & -2 & 3 \\ 2 & 0 & 0 \\ -3 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -2 & 3 \\ 2 & 0 & 0 \\ -3 & 0 & 1 \end{pmatrix}, I_3.$$

7. Si dimostri che la traccia delle matrici nulle di qualunque ordine vale 0.

8. Si dimostri che la traccia di $I_n$ vale $n$.

9. Si calcoli la traccia di una matrice antisimmetrica.

10. Si dimostri che una matrice quadrata è diagonale se e solo se è sia triangolare inferiore che superiore.

11. Si dimostri che la trasposta di una matrice triangolare inferiore è triangolare superiore e viceversa.

## 2.2 Operazioni sulle matrici

In questa sezione definiremo alcune operazioni negli insieme di matrici. Inizieremo con la *somma* ad il *prodotto per uno scalare*:

**Definizione 2.2.1.** La somma su $\mathcal{M}_{m\times n}$ è una applicazione

$$\begin{array}{rcl} + : \mathcal{M}_{m\times n} \times \mathcal{M}_{m\times n} & \to & \mathcal{M}_{m\times n} \\ (A, B) & \mapsto & A + B \end{array}$$

definita come segue

$$[A + B]_{ij} := [A]_{ij} + [B]_{ij}.$$

**Definizione 2.2.2.** Il prodotto di una matrice in $\mathcal{M}_{m\times n}$ per una o scalare è una applicazione

$$\begin{array}{rcl} \cdot : \mathbb{R} \times \mathcal{M}_{m\times n} & \to & \mathcal{M}_{m\times n} \\ (\lambda, A) & \mapsto & \lambda \cdot A \end{array}$$

definita come segue

$$[\lambda \cdot A]_{ij} := \lambda([A]_{ij}).$$

Per abbreviare la notazione, denoteremo $\lambda \cdot A$ anche con $\lambda A$.

*Osservazione* 2.2.3. Per definizione si può fare la somma tra due matrici dello stesso tipo e si ottiene ancora una matrice dello stesso tipo. Allo stesso modo se si moltiplica uno numero reale per una matrice si ottiene ancora una matrice dello stesso tipo. Vedremo più avanti che il prodotto tra matrici è definito in modo molto diverso.

**Esempio 2.2.4.** Eseguiamo la somma tra le due matrici

$$A = \begin{pmatrix} 2 & -3 & 0 \\ -4 & 3 & 10 \end{pmatrix}, \ B = \begin{pmatrix} 0 & -2 & 1 \\ -\pi & 2\pi & 0 \end{pmatrix}.$$

Sommando le entrate corrispondenti avremo

$$A + B = \begin{pmatrix} 2+0 & -3+(-2) & 0+1 \\ -4+(-\pi) & 3+2\pi & 10+0 \end{pmatrix} = \begin{pmatrix} 2 & -5 & 1 \\ -4-\pi & 3+2\pi & 10 \end{pmatrix}.$$

**Esempio 2.2.5.** Facciamo il prodotto del numero reale $\pi$ per la matrice

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

Moltilicando $\pi$ per ogni entrata di $A$ otterremo:

$$\pi \cdot A = \begin{pmatrix} -\pi & 3\pi \\ 0 & \pi \end{pmatrix}.$$

Vediamo ora alcune proprietà fondamentali della somma di matrici e del prodotto per un scalare:

**Proposizione 2.2.6.** *La somma di matrici e il prodotto per uno scalare su $\mathcal{M}_{m\times n}$ soddisfano le seguenti proprietà:*

1. *esiste una matrice $O \in \mathcal{M}_{m\times n}$ tale che $A + O = O + A$ per ogni $A \in \mathcal{M}_{m\times n}$;*

2. *per ogni $A \in \mathcal{M}_{m\times n}$ esiste una matrice $A' \in \mathcal{M}_{m\times n}$ tale che $A + A' = A' + A = O$;*

3. *per ogni* $A, B, C \in \mathcal{M}_{m\times n}$ *si ha* $(A + B) + C = A + (B + C)$*;*

4. *per ogni* $A, B \in \mathcal{M}_{m\times n}$ *si ha* $A + B = B + A$*;*

5. *per ogni* $A \in \mathcal{M}_{m\times n}$ *si ha* $1A = A$*;*

6. *per ogni* $\lambda, \mu \in \mathbb{R}$ *e per ogni* $A \in \mathcal{M}_{m\times n}$ *si ha* $\lambda(\mu A) = (\lambda\mu)A$*;*

7. *per ogni* $\lambda \in \mathbb{R}$ *e per ogni* $A, B \in \mathcal{M}_{m\times n}$ *si ha* $\lambda(A + B) = \lambda A + \lambda B$*;*

8. *per ogni* $\lambda, \mu \in \mathbb{R}$ *e ogni* $A \in \mathcal{M}_{m\times n}$ *si ha* $(\lambda + \mu)A = \lambda A + \mu A$*.*

*Dimostrazione.* Per quanto riguarda la prima e seconda proprietà osserviamo che la matrice $O$ cercata è la matrice nulla di $\mathcal{M}_{m\times n}$, mentre la matrice $A'$ è la matrice opposta di $A$. Visto che le dimostrazioni della varie proprietà sono molto simili, dimostreremo solo la terza proprietà, cioè la proprietà associativa. Per dimostrare che le matrici $(A + B) + C$ e $A + (B + C)$ sono uguali, dobbiamo dimostrare che le entrate corrispondenti sono uguali. Se $i, j$ sono due indici generici abbiamo:

$$\begin{aligned}
[(A + B) + C]_{ij} &\overset{\text{per definizione di somma}}{=} [A + B]_{ij} + [C]_{ij} \\
&\overset{\text{per definizione di somma}}{=} ([A]_{ij} + [B]_{ij}) + [C]_{ij} \\
&\overset{\text{per l'associatività della somma in}\mathbb{R}}{=} [A]_{ij} + ([B]_{ij} + [C]_{ij}) \\
&\overset{\text{per definizione di somma}}{=} [A]_{ij} + [B + C]_{ij} \\
&\overset{\text{per definizione di somma}}{=} [A + (B + C)]_{ij}.
\end{aligned}$$

□

La seguente proposizione, la cui dimostrazione è lasciata per esercizio, può essere dimostrata direttamente utilizzando le definizioni di somma e prodotto per uno scalare oppure utilizzando unicamente le proprietà della Proposizione 2.2.6:

**Proposizione 2.2.7.** *Per ogni* $A \in \mathcal{M}_{m\times n}$ *e* $\lambda \in \mathbb{R}$ *si ha:*

1. $0A = O_{m\times n}$*;*

2. $(-1)A = -A$*;*

3. $(-\lambda)A = -(\lambda A) = \lambda(-A)$.

*Osservazione* 2.2.8. Nella proposizione precedente, si noti che $-A$ è l'opposta di $A$, mentre $(-1)A$ è il prodotto di $-1$ per $A$ come definito precedentemente.

**Definizione 2.2.9.** Date delle matrici $B_1, \cdots, B_k \in \mathcal{M}_{m\times n}$ e dei numeri reali $a_1, \cdots, a_k$, chiameremo *combinazione lineare* di $B_1, \cdots, B_k$ di coefficienti $a_1, \cdots, a_k$, la matrice $\sum_{i=1}^k a_i B_i \in \mathcal{M}_{m\times n}$.

Vedremo ora la definizione di *prodotto* di matrici:

**Definizione 2.2.10.** Il prodotto tra matrici è una applicazione

$$\begin{array}{ccc} \mathcal{M}_{m\times n} \times \mathcal{M}_{n\times k} & \to & \mathcal{M}_{m\times k} \\ (A, B) & \mapsto & AB \end{array}$$

definita come segue

$$[AB]_{ij} := \sum_{s=1}^n [A]_{is}[B]_{sj}.$$

Si osservi che per poter moltiplicare due matrici, il numero di righe della seconda deve essere uguale al numero di colonne della prima. La matrice che si ottiene ha il numero di righe delle prima e il numero di colonne della seconda.

Il prodotto fra matrici viene anche chiamato *prodotto riga per colonna*. Il motivo diventa chiaro se si analizza più in dettaglio come viene calcolato il prodotto. Vediamo innanzitutto come funziona il prodotto di una matrice riga $A \in \mathcal{M}_{1\times n}$ per una matrice colonna $B \in \mathcal{M}_{n\times 1}$. Il risultato sarà una matrice $1 \times 1$ la cui unica entrata è data da

$$[AB]_{11} = \sum_{s=1}^n [A]_{1s}[B]_{s1} = [A]_{11}[B]_{11} + [A]_{12}[B]_{21} + \cdots + [A]_{1j}[B]_{j1} + \cdots + [A]_{1n}[B]_{n1}.$$

In questo modo stiamo moltiplicando una riga per una colonna. Ora è facile convincersi che se $A \in \mathcal{M}_{m\times n}$ e $B \in \mathcal{M}_{n\times k}$, l'entrata $ij$ della matrice $AB$ è proprio il prodotto della $i$-esima riga di $A$ per la $j$-esima colonna di $B$.

Un esempio con due matrici generiche chiarirà ulteriormente questo fatto.

**Esempio 2.2.11.** Calcoliamo il prodotto della matrice $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ per la matrice $B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix}$. L'entrata $[AB]_{11}$ sarà il prodotto della prima riga di $A$ per la prima colonna di $B$ e dunque varrà $a_{11}b_{11} + a_{12}b_{21}$. Continuando in questo modo si avrà quindi

$$AB = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} & a_{11}b_{13} + a_{12}b_{23} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} & a_{21}b_{13} + a_{22}b_{23} \end{pmatrix}.$$

L'esempio precedente mostra inoltre che non è possibile fare il prodotto $BA$, per via di come è stato definito il prodotto. A volte è possibile fare entrambi i prodotti ma i due risultati non sono confrontabili in quanto si ottengono matrici di tipo diverso. Ad esempio se $A \in \mathcal{M}_{2\times 3}$ e $B \in \mathcal{M}_{3\times 2}$, allora è possibile fare $AB$ e $BA$, ma $AB \in \mathcal{M}_{2\times 2}$, mentre $BA \in \mathcal{M}_{3\times 3}$.

In effetti ci si rende facilmente conto (verificare!) che $AB$ e $BA$ sono dello stesso tipo solo se sono matrici quadrate (dello stesso ordine). Anche in questo caso comunque $AB$ e $BA$ non necessariamente coincidono, come mostra il seguente esempio:

**Esempio 2.2.12.** Consideriamo le matrici

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

Facendo i due prodotti si ottiene $AB = B$ e $BA = O$ e quindi i due prodotti sono diversi.

L'esempio precedente mostra che in generale il prodotto di matrici non è commutativo. Il prodotto di matrici gode comunque di altre interessanti proprietà:

**Proposizione 2.2.13.** *Il prodotto di matrici gode delle seguenti proprietà:*

1. $A(B + C) = AB + AC$, $\forall A \in \mathcal{M}_{m\times n}$, $\forall B, C \in \mathcal{M}_{n\times k}$;

2. $(B + C)A = BA + CA$, $\forall A \in \mathcal{M}_{m\times n}$, $\forall B, C \in \mathcal{M}_{k\times m}$;

3. $\lambda(AB) = (\lambda A)B = A(\lambda B)$, $\forall \lambda \in \mathbb{R}$, $\forall A \in \mathcal{M}_{m\times n}$, $\forall B \in \mathcal{M}_{n\times k}$;

4. $A(BC) = (AB)C$, $\forall A \in \mathcal{M}_{m\times n}$, $\forall B \in \mathcal{M}_{n\times k}$, $\forall C \in \mathcal{M}_{k\times r}$.

*Dimostrazione.* La dimostrazione delle quattro proprietà è molto simile e quindi dimostreremo solamente la quarta, lasciando le altre per esercizio.

Dobbiamo dimostrare che $A(BC) = (AB)C$ $\forall A \in \mathcal{M}_{m\times n}$, $\forall B \in \mathcal{M}_{n\times k}$, $\forall C \in \mathcal{M}_{k\times r}$. Per far ciò basta dimostrare che effettivamente le due matrici sono dello stesso tipo (verifica immediata) e che le entrate corrispondenti siano uguali.

Avremo quindi:

$$
\begin{aligned}
[A(BC)]_{ij} &\overset{\text{def. di prodotto}}{=} \sum_{s=1}^{n} [A]_{is}[BC]_{sj} \\
&\overset{\text{def. di prodotto}}{=} \sum_{s=1}^{n} [A]_{is} \left( \sum_{l=1}^{k} [B]_{sl}[C]_{lj} \right) \\
&\overset{\text{propr. distrib. in } \mathbb{R}}{=} \sum_{s=1}^{n} \left( \sum_{l=1}^{k} [A]_{is}[B]_{sl}[C]_{lj} \right) \\
&\overset{\text{propr. comm. e assoc. in } \mathbb{R}}{=} \sum_{l=1}^{k} \left( \sum_{s=1}^{n} [A]_{is}[B]_{sl}[C]_{lj} \right) \\
&\overset{\text{propr. distrib. in } \mathbb{R}}{=} \sum_{l=1}^{k} \left( \sum_{s=1}^{n} [A]_{is}[B]_{sl} \right) [C]_{lj} \\
&\overset{\text{def. prod.}}{=} \sum_{l=1}^{k} [AB]_{il}[C]_{lj} \\
&\overset{\text{def. prod.}}{=} [(AB)C]_{ij}.
\end{aligned}
$$

□

### 2.2.1 Sistemi in forma matriciale

Le matrici possono essere utilizzate per scrivere i sistemi in *forma matriciale*. Vedremo di seguito due forme particolarmente utili più avanti.

Abbiamo visto che avere un sistema lineare a $n$ incognite e $m$ equazioni equivale ad avere i coefficienti $a_{ij}$ e i termini noti $b_i$, per $i = 1, \cdots, m$ (dove $m$ è il numero di equazioni) e $j = 1, \cdots, n$ (dove $n$ è il numero di incognite). I coefficienti del sistema definiscono in modo naturale le entrate di una matrice $m \times n$, che sarà la matrice $A = (a_{ij})$. La matrice dei termini noti sarà invece la matrice colonna $B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$.

Sfruttando la definizione di prodotto tra matrici, se indichiamo con $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ la matrice colonna le cui entrate sono le incognite del sistema, il sistema potrà essere scritto come segue:

$$AX = B$$

e quindi una soluzione del sistema sarà un elemento $X_0 \in \mathcal{M}_{n\times 1}$ tale che

$$AX_0 = B.$$

Se invece indichiamo con $A^1, \cdots, A^n$ le colonne della matrice, stavolta utilizzando la definizione di somma e di prodotto per uno scalare, scriveremo il sistema come segue:

$$x_1 A^1 + \cdots + x_n A^n = B.$$

### 2.2.2 Matrici invertibili

Nelle sezioni precedenti abbiamo introdotto la somma e la moltiplicazione tra matrici. Nel caso di matrici $1 \times 1$, queste operazioni coincidono con la somma e la moltiplicazione in $\mathbb{R}$. Mentre la somma tra matrici ha formalmente proprietà molti simili a quelle della somma tra numeri reali, abbiamo notato che la moltiplicazione, benché associativa, non è in generale commutativa.

Una proprietà importante nel caso della somma e della moltiplicazione in $\mathbb{R}$ è quella dell'esistenza dell'opposto per la somma e del reciproco per la moltiplicazione (tranne nel caso in cui il numero reale sia 0). Abbiamo visto che per la somma di matrici esiste un elemento neutro e data una matrice esiste la sua opposta. Cosa possiamo dire nel caso del prodotto?

Si può verificare che, data $A \in \mathcal{M}_{m\times n}$, si ha $A \times I_n = A$ e $I_m \times A = A$. Quindi la matrice identità (di ordine tale da dare un senso alla moltiplicazione) moltiplicata per una matrice la lascia invariata. Il problema è che non sempreè possibile moltiplicare $A$ per l'identità dello stesso ordine sia a destra che a sinistra. Questo si può ovviare considerando solamente matrici quadrate. Avremo allora che per qualunque $A \in \mathcal{M}_n$ si ottiene $AI_n = I_n A = A$.

Ora possiamo dare la seguente definizione:

**Definizione 2.2.14.** Una matrice quadrata $A \in \mathcal{M}_n$ è detta *ivertibile* se esiste $A' \in \mathcal{M}_n$ tale che

$$A \cdot A' = A' \cdot A = I_n.$$

Si verifica facilmente che, data $A$, esiste un'unica $A'$ che soddisfa le condizioni di cui sopra. Infatti se esistesse $A''$ che soddisfa le stesse condizioni, moltiplicando per $A''$ i; primo e il secondo membro di $A \cdot A' = I_n$ (attenzione perché il prodotto non è commutativo ) si avrebbe:

$$A''(A \cdot A') = A'' I_n = A''.$$

Ma per l'associatività e per le relazioni che soddisfa $A''$ abbiamo

$$A''(A \cdot A') = (A''A)A' = I_n A' = A'$$

e quindi si ha $A' = A''$.

La matrice $A'$ verrà chiamata l'*inversa* di $A$ e indicata semplicemente con $A^{-1}$.

Ancora non abbiamo i mezzi necessari per poter dare una condizione necessaria e sufficiente affinché una matrice sia invertibile, però possiamo dimostrare le seguenti proprietà:

**Proposizione 2.2.15.** *Se $A \in \mathcal{M}_n$ è invertibile, allora $A^{-1}$ è invertibile e si la sua inversa è $A$. Se $A, B \in \mathcal{M}_n$ sono invertibili, allora $AB$ è invertibile e la sua inversa è $B^{-1} \cdot A^{-1}$.*

*Dimostrazione.* Dato che $A$ è invertibile si ha $A \cdot A' = A' \cdot A = I_n$ e quindi anche $A'$ è invertibile. Se $A$ e $B$ sono invertibili, allora esistono $A^{-1}$ e $B^{-1}$ e dunque possiamo costruire la matrice $B^{-1} \cdot A^{-1}$. Per dimostrare che $B^{-1} \cdot A^{-1}$ è l'inversa di $AB$ basta verificare che $(B^{-1} \cdot A^{-1})(AB) = (AB)B^{-1} \cdot A^{-1} = I_n$. Per la proprietà associativa e per le proprietà di $I_n$ si ha:

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

Allo stesso modo si verifica che $(AB)(B^{-1} \cdot A^{-1}) = I_n$. □

**Esempio 2.2.16.** La matrice nulla $O_n$ non è invertibile. Infatti, se lo fosse e A fosse la sua inversa, si avrebbe $A \cdot O_n = I_n$, ma $A \cdot O_n = O_n \neq I_n$ e questo è assurdo.

**Esempio 2.2.17.** La matrice $I_n$ è invertibile e la sua inversa è sè stessa.

**Esempio 2.2.18.** Utilizzando la definizione si verifica facilmente che la matrice $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ è invertibile e la sua inversa è $\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$.

L'insieme delle matrici invertibili di ordine $n$ vereà denotato con $GL(n, \mathbb{R})$ o anche $GL_n(\mathbb{R})$ ed è chiamato *gruppo lineare*. In effetti la moltiplicazione tra matrici induce una moltiplicazione in $GL(n, \mathbb{R})$ per la Proposizione 2.2.15 e rispetto a tale prodotto $GL(n, \mathbb{R})$ è un gruppo non abeliano (tranne nel caso in cui $n = 1$ dove si riduce al gruppo moltiplicativo dei reali non nulli).

### Matrici ortogonali

In geometria e non solo, hanno un ruolo importante le matrici definite come segue:

**Definizione 2.2.19.** Una matrice $A$ (necessariamente quadrata) è detta *matrice ortogonale* se si ha

$$A^t A = AA^t = I.$$

Per definizione una matrice ortogonale è invertibile ed ha come inversa la sua trasposta. L'insieme delle matrici ortogonali di ordine $n$ verrà indicato con $O(n, \mathbb{R})$ ed è chiamato *gruppo ortogonale*. Si dimostra che anche $O(n, \mathbb{R})$ è un gruppo (si veda l'Esercizio 8).

**Esempio 2.2.20.** La matrice identità di qualunque ordine è ortogonale.

**Esempio 2.2.21.** Per ogni numero reale $\theta$, le seguenti matrici quadrate di ordine 2 sono ortogonali:

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}; \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

**Esempio 2.2.22.** La seguente matrice quadrate di ordine 3 è ortogonale:

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

## Esercizi

1. Si completi la dimostrazione delle Proposizioni 2.2.6 e 2.2.7.

2. Si completi la dimostrazione della Proposizione 2.2.13.

3. Si scrivano i seguenti sistemi lineari nelle forme matricolai viste:

   - $\begin{cases} x_1 + x_2 + x_3 = 1 \\ 2x_1 + 2x_2 + x_3 = 0 \\ x_1 + x_2 + 2x_3 = -1 \end{cases}$

   - $\begin{cases} 2x_1 - x_2 - x_3 - 4x_4 = 9 \\ 4x_1 - 3x_3 - x_4 = 0 \\ 8x_1 - 2x_2 - 5x_3 - 9x_4 = 18 \end{cases}$

   - $\begin{cases} 2x_1 + x_2 - x_3 = 1 \\ x_1 + 2x_2 - 2x_3 = 0 \\ 3x_1 - x_2 + 2x_3 = -1 \\ x_1 - x_2 + x_3 = 1 \end{cases}$

4. Si dimostri che $\forall A \in \mathcal{M}_{m\times n}$, si ha $A \times I_n = A$ e $I_m \times A = A$.

5. Si dimostri che $\forall A \in \mathcal{M}_{m\times n}$, si ha $A \times O_{n\times k} = O_{m\times k}$ e $O_{k\times m} \times A = O_{k\times n}$.

6. Si dimostrino le seguenti proprietà della trasposta di una matrice:

- $(AB)^t = B^t \cdot A^t$, $(A+B)^t = A^t + B^t$, $(\lambda A)^t = \lambda(A^t)$ ($\lambda \in \mathbb{R}$), $(A^t)^t = A$, $I_n^t = I_n$;
- se $A$ è invertibile, allora $(A^{-1})^t = (A^t)^{-1}$ (sugg: si calcoli la traccia d $A^{-1}A = AA^{-1} = I_n$).

7. Si dimostrino le seguenti proprietà della traccia di una matrice quadrata:
   - $\mathrm{tr}(AB) = \mathrm{tr}(BA)$, $\mathrm{tr}(A+B) = \mathrm{tr}A + \mathrm{tr}B$, $\mathrm{tr}(\lambda A) = \lambda\,\mathrm{tr}A$ ($\lambda \in \mathbb{R}$), $\mathrm{tr}\, I_n = n$;
   - se $B$ è invertibile $\mathrm{tr}(B^{-1}AB) = \mathrm{tr}(A)$ (sugg: si utilizzi la proprietà precedente e la definizione di inversa).

8. Utilizzando l'Esercizio 6, si dimostri che l'inversa di una matrice ortogonale è ortogonale e che il prodotto di due matrici ortogonali sono ancora ortogonali.

9. Si dimostri che ogni matrice quadrata di ordine 2 si scrive come una delle matrici dell'Esempio 2.2.21.

10. Date le seguenti matrici,
$$A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \\ 1 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & -1 \\ 1 & 3 \\ 2 & 1 \end{pmatrix}, C = \begin{pmatrix} 0 & -1 \\ 1 & 3 \\ 2 & 1 \end{pmatrix},$$
    - si calcoli $2A + 3B - C$,
    - si calcoli ${}^tA - {}^tB + 2\ {}^tC$,
    - si determini la matrice $D = (d_{ij})$ tale che $D - A + B - C = O$, dove $O$ è la matrice nulla.

11. Date le matrici
$$A_1 = \begin{pmatrix} 1 & -1 & -1 \\ 2 & 3 & 1 \\ 1 & 0 & -2 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & -1 & 1 \\ 0 & 3 & 1 \\ 2 & 0 & -2 \end{pmatrix},$$
    Si dica se esistono $\lambda_1$ e $\lambda_2$ tali che $\sum_{i=1}^{2} \lambda_i A_i = I$, dove $I$ è la matrice identità di ordine 3.
    Posto $A_3 = {}^tA_1$ e $A_4 = {}^tA_2$, si trovino le soluzioni del sistema $\sum_{i=1}^{4} \lambda_i A_i = I$.

12. Si dimostri che se è possibile eseguire la somma $A +^t A$, allora la matrice $A$ deve essere quadrata.

13. Si descrivano tutte le matrici quadrate di ordine 2 e antisimmetriche.

14. Sia $I$ la matrice identità di ordine 3. Si calcoli $\sum_{k=1}^{4} kI$. Si calcoli più in generale $\sum_{k=1}^{n} kI$.

15. Data le matrice $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, si trovi una matrice quadrata $X$ tale che $AX = I_2$

# Capitolo 3

# Determinante

### 3.0.3 Permutazioni

Ricordiamo innanzitutto che, dato un insieme $X$, l'insieme $S(X)$ delle applicazioni biunivoche da $X$ in sè stesso, può essere munito di una operazione, indicata con $\cdot$, nel modo seguente:

$$\begin{array}{rcl} \cdot : S(X) \times S(X) & \rightarrow & S(X) \\ (f, g) & \mapsto & f \cdot g := f \circ g \,, \end{array}$$

dove $f \circ g$ è la composizione di $f$ con $g$, ossìa la funzione

$$\begin{array}{rrcl} f \circ g : & X & \rightarrow & X \\ & x & \mapsto & f(g(x)). \end{array}$$

Poiché l'applicazione identità è biunivoca, l'inversa di una applicazione biunivoca è anch'essa biunivoca e la composizione è associativa, l'insieme $S(X)$ munito della suddetta operazione è un gruppo. L'elemento neutro rispetto a tale operazione è l'applicazione identità che indicheremo con $\mathbb{1}_X$. In generale non vale la proprietà commutativa.

Noi siamo particolarmente interessati al caso in cui $X$ sia finito e sia un sottoinsieme di $\mathbb{N}$. Considereremo quindi $S_n = \{x \in \mathbb{N} | 1 \leq x \leq n\}$. Possiamo ora dare la seguente definizione:

**Definizione 3.0.23.** Una *permutazione* su $n$ elementi è una applicazione biunivoca da $S_n$ in sè.

Indicheremo con $\sigma_n$ l'insieme delle permutazioni su $n$ elementi e lo muniremo del prodotto definito precedentemente. La cardinalità di $\sigma_n$ è pari a $n!$. Utilizzeremo la seguente notazione per una permutazione $\tau : S_n \rightarrow S_n$:

$$\tau = \begin{pmatrix} 1 & 2 & \cdots & n \\ \tau(1) & \tau(2) & \cdots & \tau(n) \end{pmatrix}.$$

In particolare la premutazione identità di $\sigma_n$ sarà:

$$\mathbb{1}_n = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}.$$

Per ottenere la permutazione inversa basterà dunque capovolgere la tabella e riordinare le colonne in modo tale da ottenere nella prima riga $1, 2, \cdots, n$. Ad esempio, se consideriamo la permutazione

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix},$$

la sua inversa si ottiene capovolgendo la tabella

$$\begin{pmatrix} 4 & 1 & 2 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix},$$

e riordinando le colonne in modo da avere

$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}.$$

Date due permutazioni

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}, \tau = \begin{pmatrix} 1 & 2 & \cdots & n \\ \tau(1) & \tau(2) & \cdots & \tau(n) \end{pmatrix},$$

il prodotto $\sigma\tau$, sarà

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(\tau(1)) & \sigma(\tau(2)) & \cdots & \sigma(\tau(n)) \end{pmatrix}.$$

Vediamo su un esempio come svolgere il prodotto.

**Esempio 3.0.24.** Siano $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$ e $\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$.

Per calcolare il prodotto $\sigma\tau$, serve l'immagine di 1 e quindi occorre calcolare $\tau(1)$ che è 3 e poi calcolare $\sigma(3)$ che è 4 e quindi il prodotto $\sigma\tau$ manda 1 in 4. Procedendo in questo modo si ottiene:

$$\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}.$$

**Cicli, trasposizioni**

In questo paragrafo vedremo che una permutazione qualunque può essere scritta come prodotto di permutino più semplici. Vediamo innanzitutto cosa intendiamo per permutazioni più semplici:

**Definizione 3.0.25.** Un *ciclo* di $\sigma_n$, di lunghezza $k \leq n$ è una permutazione $\tau$ per cui esistono $k$ elementi distinti $i_1, \cdots, i_k \in S_n$ tali che:

- $\tau(i_1) = i_2$, $\tau(i_2) = i_3$, $\cdots$, $\tau(i_{k-1}) = i_k$, $\tau(i_k) = i_1$;
- $\tau(j) = j$ per qualunque $j \in S_n - \{i_1, \cdots, i_k\}$.

Quindi un ciclo permuta gli elementi $i_1, \cdots, i_k$ in modo ciclico, mentre lascia invariati tutti gli altri. Indicheremo un ciclo semplicemente scrivendo $(i_1, \cdots, i_k)$.

**Esempio 3.0.26.** La permutazione $\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}$. è un ciclo, più precisamente è il ciclo $(1\,4\,3)$. Infatti si ha $\tau(1) = 4$, $\tau(4) = 3$, $\tau(3) = 1$, mentre $\tau(2) = 2$.

Si osservi che $(i_1, \cdots, i_k)$ e $(i_2, i_3, \cdots, i_k, i_1)$ rappresentano lo stesso ciclo.

**Definizione 3.0.27.** Un ciclo di lunghezza 2 è detto *trasposizione*.

Si osservi che una trasposizione $t$ verifica $t^2 = \mathbb{1}$ o, in modo equivalente, che l'inversa di $t$ è $t$ stessa.

**Esempio 3.0.28.** La permutazione $\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}$ è una trasposizione.

**Definizione 3.0.29.** Due cicli $(i_1, \cdots, i_k), (j_1, \cdots, j_s) \in \sigma_n$ sono detti *disgiunti* se gli insiemi $\{i_1, \cdots, i_k\}, \{j_1, \cdots, j_s\}$ sono disgiunti.

È facile dimostrare che il prodotto di due cicli disgiunti è commutativo.

Possiamo ora enunciare due teoremi fondamentali sulle permutazioni, le cui dimostrazioni possono essere trovate in [1].

**Teorema 3.0.30.** *Ogni permutazione si scrive in modo unico (a meno dell'ordine) come prodotto di cicli disgiunti.*

**Teorema 3.0.31.** *Ogni permutazione può essere scritta (non necessariamente in modo unico) come prodotto di trasposizioni.*

Vediamo ora come scrivere una permutazione come prodotto di cicli disgiunti. Consideriamo $\sigma \in \sigma_n$ e calcoliamo $\sigma(1)$, $\sigma(\sigma(1))$ e così via fino a quando il risultato non sarà nuovamente 1 e sia $k$ il primo numero naturale tale che $\sigma^k(1) = 1$. Questo si verifica poiché gli elementi di $S_n$ finiti. In questo modo determiniamo il ciclo $(1\,\sigma(1)\,\cdots\,\sigma^{k-1}(1))$. A questo punto sia $j \in S_n - \{1, \sigma(1), \cdots, \sigma^{k-1}(1)\}$. Con lo stesso procedimento di prima calcoliamo le iterate di $j$ tramite $\sigma$ e otteniamo un altro ciclo. Continuando in questo modo si ottengono cicli disgiunti il cui prodotto è esattamente $\sigma$.

Per ottenere una scomposizione in prodotto di trasposizioni sarà sufficiente dare un metodo per scrivere un ciclo come prodotto di trasposizioni. Se il ciclo è $(i_1, \cdots, i_k)$, ci si convince facilmente che esso è uguale al seguente prodotto di trasposizioni:

$$(i_k\, i_{k-1})(i_k\, i_{k-2})\cdots(i_k\, i_2)(i_k\, i_1).$$

**Esempio 3.0.32.** Vogliamo scrivere come prodotto di cicli disgiunti e come prodotto di trasposizioni la seguente permutazione:

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 9 & 1 & 3 & 7 & 2 & 5 & 6 & 8 \end{pmatrix}.$$

Iniziamo a calcolare l'iterata di 1 e otteniamo il ciclo $(1\,4\,3)$. Ora prendiamo un elemento di $S_9 - \{1, 4, 3\}$, ad esempio 2 e calcoliamone l'iterata. In tal modo otteniamo il ciclo $(2\,9\,8\,6)$. Prendiamo ora un elemento che non appare nei cicli precedenti, ciè in $S_9 - \{1, 4, 3, 2, 9, 8, 6\}$ e calcoliamone l'iterata. Possiamo prendere 5 e ottenere quindi il ciclo $(5\,7)$. Adesso abbiamo esaurito tutti gli elementi e quindi avremo:

$$\tau = (1\,4\,3)(2\,9\,8\,6)(5\,7).$$

Se decomponiamo ogni ciclo come prodotto di trasposizioni otteniamo

$$\tau = (3\,4)(3\,1)(6\,8)(6\,9)(6\,2)(5\,7).$$

### Segno di una permutazione

Per definire il segno di una permutazione, ci occorre la seguente definizione:

**Definizione 3.0.33.** Data una permutazione $\sigma \in \sigma_n$, una *inversione* per $\sigma$ è una coppia $(i, j) \in S_n \times S_n$ tale che $i < j$ e $\sigma(i) > \sigma(j)$.

Denoteremo con $i(\sigma)$ il numero totale di inversioni di $\sigma$ e porremo

$$\epsilon(\sigma) := (-1)^{i(\sigma)} = \prod_{1\le i<j\le n} \frac{\sigma_j - \sigma_i}{j - i}.$$

**Definizione 3.0.34.** una permutazione $\sigma$ è detta *pari* se $\epsilon(\sigma) = 1$, *dispari* se $\epsilon(\sigma) = -1$.

**Esempio 3.0.35.** Consideriamo la premutzaione

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 6 & 1 & 3 & 7 & 2 & 5 \end{pmatrix}.$$

Le coppie $(1,3), (1,4), (1,6), (2,3), (2,4), (2,6), (4,6), (5,6), (5,7)$ sono tutte le inversioni di $\tau$ e quindi $\tau$ è dispari.

Se scriviamo una permutazione $\sigma$ come prodotto di trasposizioni e supponiamo che compaia un numero di trasposizioni $\#t$, è possibile dimostrare che una qualunque altra scomposizione ha un numero di trasposizioni che differisce da $\#t$ per un numero pari e che si ha la seguente proprietà:

**Proposizione 3.0.36.** *Per una permutazione $\sigma$ si ha $\epsilon(\sigma) = (-1)^{\#t}$, dove $\#t$ è il numero di trasposizioni di una sua qualunque decomposizione.*

Vediamo ora alcune utili proprietà del segno di una permutazione:

**Proposizione 3.0.37.** *Il segno di una permutazione gode delle seguenti proprietà:*

1. $\epsilon(\mathbb{1}) = 1$;
2. *se $t$ è una trasposizione allora $\epsilon(t) = -1$;*
3. $\forall \sigma, \tau \in \sigma_n$ *si ha* $\epsilon(\sigma\tau) = \epsilon(\sigma)\epsilon(\tau)$;
4. $\forall \sigma \in \sigma_n$ *si ha* $\epsilon(\sigma^{-1}) = \epsilon(\sigma)$;
5. *se $t$ è una trasposizione, $\forall \sigma \in \sigma_n$ si ha $\epsilon(\sigma t) = -\epsilon(\sigma)$.*

Per le proprietà delle proposizione precedente, si ha che il prodotto di due permutazioni pari è pari e quello di due permutazioni dispari è pari. Inoltre l'inversa di una permutazione pari è pari. Indicheremo con $\mathcal{S}_n$ l'insieme delle permutazioni pari su $n$ elementi e con $\mathcal{A}_n$ l'insieme delle permutazioni dispari.

L'insieme $\mathcal{S}_n$ munito del prodotto di parmutazioni è un *sottogruppo* del gruppo $\sigma_n$.

Consideriamo ora una permutazione

$$\sigma = \begin{pmatrix} 1 & \cdots & i & \cdots & j & \cdots & n \\ \sigma(1) & \cdots & \sigma(i) & \cdots & \sigma(j) & \cdots & \sigma(n) \end{pmatrix}.$$

Se moltiplichiamo $\sigma$ per la trasposizione $(i\ j)$ otterremo

$$\sigma(i,j) = \begin{pmatrix} 1 & \cdots & i & \cdots & j & \cdots & n \\ \sigma(1) & \cdots & \sigma(j) & \cdots & \sigma(i) & \cdots & \sigma(n) \end{pmatrix}.$$

Moltiplicando successivamente per un certo numero trasposizioni, diciamo $t_1, \cdots t_k$, sarà possibile ottenere la permutazione identità. Si avrà ciè:

$$\sigma t_1 t_2 \cdots t_k = \mathbb{1}.$$

Moltiplicando a destra il primo e il secondo membro per l'inversa di $t_k$ (che è ancora $t_k$ per quanto detto precedentemente) si otterrà

$$\sigma t_1 t_2 \cdots t_{k-1} = t_k.$$

Continuando in questo modo si avrà:

$$\sigma = t_k t_{k-1} \cdots t_1,$$

e quindi abbiamo scritto $\sigma$ come prodotto di trasposizioni. Vediamo un esempio di quanto appena fatto:

**Esempio 3.0.38.** Consideriamo la permutazione

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 6 & 1 & 3 & 7 & 2 & 5 \end{pmatrix}.$$

Moltiplicando $\sigma$ per la trasposizione $(5\,7)$ si ottiene:

$$\tau(5\,7) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 6 & 1 & 3 & 5 & 2 & 7 \end{pmatrix}.$$

Moltiplicando $\tau(5\,7)$ per $(1\,3)$ si ottiene

$$\tau(5\,7)(1\,3) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 6 & 4 & 3 & 5 & 2 & 7 \end{pmatrix}.$$

Moltiplicando quanto ottenuto per $(6\,2)$ si ottiene

$$\tau(5\,7)(1\,3)(6\,2) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 2 & 4 & 3 & 5 & 6 & 7 \end{pmatrix}.$$

Infine, moltiplicando per $(3\,4)$ avremo

$$\tau(5\,7)(1\,3)(6\,2)(3\,4) = \mathbb{1}$$

e quindi

$$\tau = (3\,4)(6\,2)(1\,3)(5\,7)$$

è una permutazione pari.

### 3.0.4 Determinante di una matrice quadrata

In questa sezione ad ogni matrice quadrata assoceremo un numero reale che chiameremo determinante

**Definizione 3.0.39.** Data una matrice quadrata $A = (a_{ij})$ di ordine $n$, chiameremo *determinante* di $A$ il numero reale così definito:

$$\det A := \sum_{\tau \in \sigma_n} \epsilon(\tau) a_{1\,\tau(1)} a_{2\,\tau(2)} \cdots a_{n\,\tau(n)}.$$

Si osservi che nella definizione appare un addendo per ogni permutazione su $n$ elementi, per un totale di $n!$ addendi. Inoltre ogni addendo è il prodotto di $n$ entrate che si trovano tutte in righe diverse e colonne diverse.

Per indicare il determinante di una matrice utilizzeremo anche un'altra notazione. Data la matrice $A = (a_{ij})$, denoteremo il determinante di $A$ con $|a_{ij}|$.

Vediamo in dettaglio, utilizzando la definizione, come si calcola il determinante delle matrici di ordine 2. Ricordiamo innanzitutto che l'insieme delle permutazioni su due elementi ha esattamente 2 elementi: la permutazione identità $\mathbb{1}$ che è pari e la permutazione $\sigma = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ che è dispari.

Avremo quindi

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \epsilon(\mathbb{1}) a_{1\,\mathbb{1}(1)} a_{2\,\mathbb{1}(2)} + \epsilon(\sigma) a_{1\,\sigma(1)} a_{2\,\sigma(2)} = a_{11}a_{22} - a_{12}a_{21}.$$

**Esempio 3.0.40.** Calcoliamo i determinati delle seguenti matrici:

$$\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \begin{pmatrix} \cos a & -\sin a \\ \sin a & \cos a \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} \cos a & \sin a \\ \sin a & -\cos a \end{pmatrix}, I_2.$$

Avremo

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

$$\begin{vmatrix} a & 0 \\ 0 & b \end{vmatrix} = a \cdot b - 0 \cdot 0 = ab,$$

$$\begin{vmatrix} \cos a & -\sin a \\ \sin a & \cos a \end{vmatrix} = \cos a \cdot \cos a - (-\sin a \cdot \sin a) = \cos^2 a + \sin^2 a = 1,$$

$$\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = 1 \cdot (-1) - 1 \cdot 1 = -2,$$

$$\begin{vmatrix} \cos a & \sin a \\ \sin a & -\cos a \end{vmatrix} = -\cos a \cdot \cos a - (\sin a \cdot \sin a) = -\cos^2 a - \sin^2 a = -1.$$

Allo stesso modo si ottiene $\det I_2 = 1$.

Per le matrici di ordine 3, ricordando che vi sono 6 permutazioni du 3 elementi e calcolando di ognuna il segno si otterrà:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}-a_{11}a_{23}a_{32}.$$

Dalla definizione dovrebbe essere chiaro che il calcolo del determinante diventa tanto più lungo quanto più è grande l'ordine di una matrice. Quindi è importante trovare delle proprietà che ci permettano di semplificare questo calcolo. Vediamone alcune:

**Proposizione 3.0.41.** *Il determinante di una matrice quadrata gode delle seguenti proprietà:*

1. $\det(A^t) = \det A$ *(questa proprietà ci dice in particolare che le proprietà riguardanti le righe valgono anche per le colonne e viceversa);*

2. *se una riga (o una colonna) della matrice A è nulla, allora* $\det A = 0$*;*

3. *se* $A = (a_{ij})$ *e* $B = (b_{ij})$ *si ottiene da A permutando le righe con una permutazione* $\sigma$*, allora si ha* $\det B = \epsilon(\sigma)\det A$*;*

4. *se A ha due righe o due colonne uguali allora* $\det A = 0$*;*

5. *se* $A = (A_1, \cdots, A_n)$ *e* $B = (A_1, \cdots, A_{j-1}, \alpha A_j, A_{j+1}, \cdots, A_n)$ *allora* $\det B = \alpha \det A$*;*

6. *se* $B = \alpha A$ *allora* $\det B = \alpha^n \det A$*;*

7. *se* $A = (A_1, \cdots, A_{j-1}, A_j + B_j, A_{j+1}, \cdots, A_n)$ *allora si ha*

   $$\det A = \det(A_1, \cdots, A_{j-1}, A_j, A_{j+1}, \cdots, A_n) + \det(A_1, \cdots, A_{j-1}, B_j, A_{j+1}, \cdots, A_n).$$

8. *se* $A = (A_1, \cdots, A_{j-1}, \sum_s \alpha_s B_s, A_{j+1}, \cdots, A_n)$ *allora si ha*

   $$\det A = \sum_s \alpha_s \det(A_1, \cdots, A_{j-1}, B_s, A_{j+1}, \cdots, A_n).$$

9. *il determinante di una matrice non cambia se ad una sua riga (rispettivamente colonna) si somma una combinazione lineare delle altre righe (rispettivamente colonne).*

*Dimostrazione.* 1. Per definizione abbiamo

$$\det A := \sum_{\tau\in\sigma_n} \epsilon(\tau) a_{1\,\tau(1)} a_{2\,\tau(2)} \cdots a_{n\,\tau(n)}$$
$$\det A^t := \sum_{\sigma\in\sigma_n} \epsilon(\sigma) a_{\sigma(1)\,1} a_{\sigma(2)\,2} \cdots a_{\sigma(n)\,n}.$$

Si osservi innanzitutto che il determinante di $A$ può essere scritto anche come segue:

$$\sum_{\tau^{-1}\in\sigma_n} \epsilon(\tau) a_{1\,\tau(1)} a_{2\,\tau(2)} \cdots a_{n\,\tau(n)},$$

descrivendo gli addendi tramite $\tau^{-1}$. Ogni addendo del $\det A^t$ può essere scritto come

$$\epsilon(\sigma) a_{1\,\sigma^{-1}(1)} a_{2\,\sigma^{-1}(2)} \cdots a_{n\,\sigma^{-1}(n)},$$

per via dell'invertibilità di $\sigma$ e per la proprietà commutativa dei numeri reali. Dato che $\epsilon(\sigma) = \epsilon(\sigma^{-1})$, avremo

$$\det A^t = \sum_{\sigma\in\sigma_n} \epsilon(\sigma^{-1}) a_{1\,\sigma^{-1}(1)} a_{2\,\sigma^{-1}(2)} \cdots a_{n\,\sigma^{-1}(n)}.$$

Ponendo $\tau = \sigma$ si ottiene

$$\det A^t = \sum_{\tau^{-1}\in\sigma_n} \epsilon(\tau) a_{1\,\tau(1)} a_{2\,\tau(2)} \cdots a_{n\,\tau(n)} = \det A.$$

2. Poiché ogni addendo è della forma $\epsilon(\tau) a_{1\,\tau(1)} a_{2\,\tau(2)} \cdots a_{n\,\tau(n)}$ per qualche permutazione $\tau$, in esso appare un elemento di ogni riga e uno di ogni colonna. Quindi se una riga o una colonna ha elementi tutti nulli, ogni addendo del determinante è nullo e dunque il determinante è nullo.

3. Sia $\tau$ una permutazione su $n$ elementi, $A = (a_{ij}), C = (c_{ij})$ due matrici quadrate di ordine $n$ tali che $c_{ij} = a_{\tau(i)j}$. Avremo allora

$$\begin{aligned}
\det C &= \sum_{\sigma\in\sigma_n} \epsilon(\sigma) c_{1\,\sigma(1)} c_{2\,\sigma(2)} \cdots c_{n\,\sigma(n)} \\
&= \sum_{\sigma\in\sigma_n} \epsilon(\sigma) a_{\tau(1)\,\sigma(1)} a_{\tau(2)\,\sigma(2)} \cdots a_{\tau(n)\,\sigma(n)} \\
&= \sum_{\sigma\in\sigma_n} \epsilon(\sigma) a_{\tau(1)\,\sigma\tau^{-1}\tau(1)} a_{\tau(2)\,\sigma\tau^{-1}\tau(2)} \cdots a_{\tau(n)\,\sigma\tau^{-1}\tau(n)}.
\end{aligned}$$

Utilizzando la proprietà commutativa dei numeri reali, si possono riordinare i fattori che compongono ogni addendo in modo tale che si abbia

$$\det C = \sum_{\sigma\in\sigma_n} \epsilon(\sigma) a_{1\,\sigma\tau^{-1}(1)} a_{2\,\sigma\tau^{-1}(2)} \cdots a_{n\,\sigma\tau^{-1}(n)}.$$

Moltiplicando il secondo membro per $1 = \epsilon(\tau)\epsilon(\tau^{-1})$ e ricordando che $\epsilon(\sigma\tau^{-1}) = \epsilon(\sigma)\epsilon(\tau^{-1})$, si ottiene

$$\begin{aligned}\det C =&\epsilon(\tau)\sum_{\sigma\in\sigma_n}\epsilon(\sigma)\epsilon(\tau^{-1})a_{1\,\sigma\tau^{-1}(1)}a_{2\,\sigma\tau^{-1}(2)}\cdots a_{n\,\sigma\tau^{-1}(n)}\\ =&\epsilon(\tau)\sum_{\sigma\in\sigma_n}\epsilon(\sigma\tau^{-1})a_{1\,\sigma\tau^{-1}(1)}a_{2\,\sigma\tau^{-1}(2)}\cdots a_{n\,\sigma\tau^{-1}(n)}\\ =&\epsilon(\tau)\det A,\end{aligned}$$

dove nell'ultima riga si è sfruttato il fatto che ogni permutazione può essere scritta come $\sigma\tau^{-1}$ per una certa $\sigma$ e quindi al variare di $\sigma \in \sigma_n$ si ottengono tutte le permutazioni su $n$ elementi.

4. Supponiamo che due righe $A_i$ e $A_j$ di $A$ siano uguali. La matrice $B$ che si ottiene permutando $A_i$ e $A_j$ sarà uguale ad $A$. Per la proprietà precedente si ha $\det B = \epsilon((i\,j))\det A = -\det A$. Essendo $B = A$ di ottiene $\det A = -\det A$ e quindi $\det A = 0$.

5. Lasciata per esercizio.

6. Basta applicare la precedente proprietà ad ognuna delle righe.

7. Lasciata per esercizio.

8. Questa proprietà è una generalizzazione della precedente che può essere facilmente dimostrata per induzione su $s$.

9. Lasciata per esercizio.

□

Un altro risultato importante riguardante il determinante è il seguente teorema, che ha un corollario particolarmente utile nel seguito:

**Teorema 3.0.42** (Teorema di Binet[1]). *Date due matrici quadrate $A, B \in \mathcal{M}_n$ si ha:*

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

*Dimostrazione.* Per le matrici $A$ e $B$ utilizziamo la notazione per righe, cioé $A = (A_1, \cdots, A_n)$ e $B = (B_1, \cdots, B_n)$. Detta $C$ la matrice prodotto $AB$, per definizione si ha $[C]_{ij} = \sum_{s=1}^{n}[A]_{is}[B]_{sj}$. Osserviamo che le righe di $C$ possono essere espresse come $C_i = \sum_{s=1}^{n}[A]_{is}B_s$.

[1] *Jacques Philippe Marie Binet*: matematico e astronomo francese (Rennes, 1786 - Parigi, 1856).

Calcoliamo il determinante di $C$ utilizzando le proprietà del determinante e otteniamo:

$$\begin{aligned}
\det C &= \det\Big(\sum_{s_1=1}^{n}[A]_{1s_1}B_{s_1}, \sum_{s_2=1}^{n}[A]_{2s_2}B_{s_2}, \cdots, \sum_{s_n=1}^{n}[A]_{ns_n}B_{s_n}\Big) \\
&= \sum_{s_1=1}^{n}[A]_{1s_1}\det\Big(B_{s_1}, \sum_{s_2=1}^{n}[A]_{2s_2}B_{s_2}, \cdots, \sum_{s_n=1}^{n}[A]_{ns_n}B_{s_n}\Big) \\
&= \sum_{s_1=1}^{n}[A]_{1s_1}\sum_{s_2=1}^{n}[A]_{2s_2}\cdots\sum_{s_n=1}^{n}[A]_{ns_n}\det(B_{s_1}, B_{s_2}, \cdots, B_{s_n}) \\
&= \sum_{s_1=1}^{n}\sum_{s_2=1}^{n}\cdots\sum_{s_n=1}^{n}[A]_{1s_1}[A]_{2s_2}\cdots[A]_{ns_n}\det(B_{s_1}, B_{s_2}, \cdots, B_{s_n}).
\end{aligned} \tag{3.1}$$

Di tutte le sommatorie rimmarrà solamente quella in cui $s = (s_1, s_2, \cdot, s_n)$ è una permutazione, poiché in caso contrario avremmo uno degli $s_i$ uguale ad un certo $s_j$ e dunque si annullerebbe il determinante $\det(B_{s_1}, \cdots, B_{s_i}, \cdots, B_{s_j}, \cdots, B_{s_n})$. Posto $s_i = s(i)$, la sommatoria diventerà allora:

$$\begin{aligned}
\det C &= \sum_{s\in\sigma_n}[A]_{1s(1)}[A]_{2s(2)}\cdots[A]_{ns(n)}\det(B_{s(1)}, B_{s(2)}, \cdots, B_{s(n)}) \\
&= \sum_{s\in\sigma_n}[A]_{1s(1)}[A]_{2s(2)}\cdots[A]_{ns(n)}\epsilon(s)\det(B_1, B_2, \cdots, B_n) \\
&= \sum_{s\in\sigma_n}\epsilon(s)[A]_{1s(1)}[A]_{2s(2)}\cdots[A]_{ns(n)}\det(B_1, B_2, \cdots, B_n) \\
&= \sum_{s\in\sigma_n}\epsilon(s)[A]_{1s(1)}[A]_{2s(2)}\cdots[A]_{ns(n)}\det B \\
&= \det A\det B.
\end{aligned} \tag{3.2}$$

□

**Corollario 3.0.43.** *Se una matrice quadrata $A$ è invertibile allora il suo determinante è non nullo. Inoltre il determinante di $A^{-1}$ è il reciproco del determinante di $A$.*

*Dimostrazione.* Se $A \in \mathcal{M}_n$ è invertibile, allora esiste $A^{-1}$ che verifica:

$$A^{-1}A = AA^{-1} = I.$$

Se si considera la prima uguaglianza e si calcola il determinante di ambo i membri si ottiene

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

Applicando il Teorema di Binet si ha

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

Ciò implica che $\det A \neq 0$ e quindi si possono dividere entrambi i membri per $\det A$ ottenendo infine

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

□

### 3.0.5 Sottomatrici e minori

Data una matrice $A \in \mathcal{M}_{m\times n}$ e dei numeri naturali $1 \leq i_1 < i_2 \cdots < i_k \leq m$ e $1 \leq j_1 < j_2 \cdots < j_l \leq n$, la *sottomatrice* di $A$ individuata da $I = \{i_1, i_2, \cdots, i_k\}$ e $J = \{j_1, j_2, \cdots, j_l\}$ è la matrice $B$ con $k$ righe e $l$ colonne definita come segue:

$$[B]_{sr} := a_{I_s J_r}.$$

Quando $k = l$, la matrice $B$ è quadrata e sarà detta *minore* di $A$ di ordine $k$.

**Esempio 3.0.44.** Consideriamo la matrice

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix},$$

$I = \{2, 3\}$ e $J = \{2, 3, 4\}$. La sottamatrice individuata da $I$ e $J$ sarà allora:

$$\begin{pmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \end{pmatrix}.$$

**Definizione 3.0.45.** Data una matrice quadrata $A$ di ordine $n$ ed una sua entrata $a_{ij}$, diremo *minore complementare* di $a_{ij}$ il minore di ordine $n-1$ ottenuto da $A$ eliminando la $i$-esima riga e la $j$-esima colonna. Detto $M_{ij}$ tale minore, sarà detto *complemento algebrico* o *cofattore* di $a_{ij}$ il numero reale definito come segue:

$$A_{ij} := (-1)^{i+j} \det M_{ij}.$$

Possiamo ora enunciare un importante teorema che ci permette di ridurre il calcolo del determinante di una matrice di ordine $n$ a $n$ determinanti di matrici di ordine $n-1$. L'iterazione di questo procedimento ci permetterà, almeno in via teorica, il calcolo di qualunque determinante.

**Teorema 3.0.46** (Primo teorema di Laplace)**.** *Sia $A \in \mathcal{M}_n$, allora per ogni $h = 1, \cdots, n$ si ha:*

$$\det A = \sum_{j=1}^{n} a_{hj} A_{hj} = \sum_{i=1}^{n} a_{ih} A_{ih}.$$

*Dimostrazione.* Poniamo $A = (A_1, \cdots, A_n)$ e denotiamo con $R_j = (0, \cdots, 1, \cdots, 0)$ la matrice riga che ha 1 nella $j$-esima colonna e zero altrove. Sia inoltre $B_j$ la matrice che si ottiene da $A$ sostituendo la riga $A_h$ con la matrice $R_j$, ossìa $B_j = (A_1, \cdots, A_{h-1}, R_j, A_{h+1}, \cdots, A_n)$.

Si ha $A_h = (a_{h1}, \cdots, a_{hn}) = \sum_{j=1}^{n} a_{hj} R_j$ e quindi

$$\begin{aligned} \det A &= \sum_{j=1}^{n} a_{hj} \det(A_1, \cdots, A_{h-1}, R_j, A_{h+1}, \cdots, A_n) \\ &= \sum_{j=1}^{n} a_{hj} \det B_j. \end{aligned} \tag{3.3}$$

Resta quindi da dimostrare che $\det B_j = \det M_{hj}$, dove $M_{hj}$ è il minore complementare di $a_{hj}$. Facendo $n - j$ trasposizioni si può portare la $j$-esima colonna nell'ultima, ottenendo quindi una matrice del tipo seguente:

$$B'_j = \left(\begin{array}{c|c} & \vdots \\ M_{hj} & \\ & \vdots \\ \hline 0 \cdots 0 & 1 \end{array}\right).$$

Osserviamo innanzitutto che una permutazione $\sigma$ su $n$ elementi tale che $\sigma(n) = n$ può essere identificata con una permutazione $p$ su $n - 1$ elementi ponendo $p(i) = \sigma(i)$. Viceversa, data $p \in \sigma_{n-1}$ esiste un'unica $\sigma \in \sigma_n$ tale che $\sigma(n) = n$ e $\sigma(i) = i$ per $i \neq n$. Inoltre le permutazioni $p$ e $\sigma$ hanno lo stesso segno.

Poniamo $B'_j = (b_{sk})$ e osserviamo che se ne calcoliamo il determinante gli addendi per cui $\sigma(n) \neq n$ sono nulli poiché l'unica entrata non nulla della $n$-esima riga è quella nella $n$-esima colonna e inoltre si ha $b_{n\sigma(n)} = 1$. Calcolando il determinante si ha dunque:

$$\begin{aligned} \det B'_j &= \sum_{\sigma \in \sigma_n} \epsilon(\sigma) b_{1\sigma(1)} \cdots b_{n-1\sigma(n-1)} b_{n\sigma(n)} \\ &= \sum_{p \in \sigma_{n-1}} \epsilon(p) b_{1p(1)} \cdots b_{n-1p(n-1)} \cdot 1 \\ &= \sum_{p \in \sigma_{n-1}} \epsilon(p) b_{1p(1)} \cdots b_{n-1p(n-1)} \\ &= \det M_{ij}. \end{aligned} \tag{3.4}$$

□

Il calcolo del determinante utilizzando il primo teorema di Laplace è chiamato *sviluppo per riga* o *sviluppo per colonna*. Il seguente risultato generalizza il primo teorema di Laplace e ha importanti corollari:

**Teorema 3.0.47** (Secondo teorema di Laplace)**.** *Sia* $A \in \mathcal{M}_n$*, allora per ogni* $h, k = 1, \cdots, n$ *si ha:*

$$\sum_{j=1}^{n} a_{hj}A_{kj} = \sum_{i=1}^{n} a_{ih}A_{ik} = \det A\, \delta_{hk}.$$

*Dimostrazione.* Per $h = k$ si ha il primo teorema di Laplace. Per $h \neq k$ consideriamo $A = (A_1, \cdots, A_n)$ e sia $A' = (A_1, \cdots, A_{h-1}, A_k, A_{h+1}, \cdots A_k, \cdots, A_n)$ la matrice ottenuta da $A$ sostituendo la $h$-esima riga con $A_k$ e lasciando le altre invariate. Allora si ha $\det A' = 0$ poiché $A'$ ha due righe uguali. Sviluppando il determinante di $A'$ lungo la $h$-esima riga (che è uguale a $A_k$) si ottiene (ricordando che per $h \neq k$ si ha $\delta_{hk} = 0$)

$$\det A\, \delta_{hk} = 0 = \det A' = \sum_{j=1}^{n} a_{kj}A_{hj}.$$

Passando alla trasposta si ottiene l'altra uguaglianza. □

Il secondo teorema di Laplace può essere espresso in termini matriciali costruendo, a partire da una matrice quadrata $A$ , un'altra matrice quadrata, chiamata *aggiunta* di $A$, denotata con $Ad(A)$ e definita come segue:

$$Ad(A) := (A_{ij})^t,$$

cioé la trasposta della matrice che ha come entrate i complementi algebrici di $A$.

A questo punto il secondo teorema di Laplace può essere espresso come segue (la verifica è lasciata per esercizio):

$$A \cdot Ad(A) = Ad(A) \cdot A = (\det A)I_n.$$

Ecco un importante corollario del secondo teorema di Laplace:

**Corollario 3.0.48.** *Una matrice quadrata $A$ è invertibile se e solo se il suo determinante è non nullo.*

*Dimostrazione.* Se $A$ è invertibile, per il teorema di Binet $\det A$ è diverso da zero. Viceversa, se $\det A \neq 0$, per il secondo teorema di Laplace e per le proprieà del prodotto di uno scalare per una matrice si ha

$$A \cdot [(\det A)^{-1} Ad(A)] = [(\det A)^{-1} Ad(A)] \cdot A = I_n$$

e dunque l'inversa di $A$ è $(\det A)^{-1} Ad(A)$. □

Si osservi che il corollario permette di calcolare esplicitamente l'inversa di una matrice invertibile. Infatti, data la matrice $A$, sarà sufficiente calcolarne il determinante e il cofattore di ogni sua entrata per poter costruire l'aggiunta di $A$.

## Esercizi

1. Si scrivano tutti gli elementi di $\sigma_1$, di $\sigma_2$, di $\sigma_3$ e di $\sigma_4$.

2. Si consideri la permutazione
$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 9 & 1 & 3 & 7 & 2 & 5 & 6 & 8 \end{pmatrix}.$$
Si calcolino le potenze di $\tau$ fino all'esponente 9. Si dica qual'è il più piccolo esponente $k$ tale che $\tau^k = \mathbb{1}$.

3. Di ogni permutazione su 4 elementi si scriva la decomposizione in cicli disgiunti e una scomposizione come prodotto di trasposizioni. Di ognuna si dica se è pari o dispari.

4. Date le seguenti permutazioni,
$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 5 & 4 & 3 & 2 & 1 & 9 & 8 & 7 & 6 \end{pmatrix},$$
$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 5 & 7 & 9 & 1 & 2 & 4 & 6 & 8 \end{pmatrix},$$
$$\rho = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 9 & 5 & 2 & 8 & 6 & 1 & 4 & 3 & 7 \end{pmatrix},$$
   - si calcolino le permutazioni inverse di $\sigma$, $\tau$, $\rho$;
   - si calcolino i prodotti $\sigma\tau\rho$, $\sigma\rho$, $\rho\sigma$, $\tau^3$, $\rho^2\tau\sigma^{-1}$, $\sigma\tau^{-1}\rho^{-1}$;
   - si scrivano tutte le permutazioni dei punti precedenti come prodotto di cicli e come prodotto di trasposizioni;
   - si calcolino i segni di tutte le permutazioni precedenti.

5. Data una permutazione $\sigma \in S_n$, è possibile stabilire quanto vale il segno di $\sigma^2$ senza conoscere $\sigma$ esplicitamente?

6. Si verifichi che per una matrice $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ si ha:
$$\det A = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}.$$

7. Si dimostri che una matrice ortogonale ha determinante uguale $\pm 1$. Si mostrino degli esempi di matrici ortogonali con determinante pari a 1 e con determinante pari a $-1$.

8. Si diano degli esempi di matrici quadrate con determinante uguale a $\pm 1$, che non siano ortogonali.

9. Date due matrici quadrate $A, B$, si dimostri che se $AB$ è invertibile, allora $A$ e $B$ hanno determinante non nullo.

# Capitolo 4

# Spazi vettoriali

In questo capitolo considereremo la nozione di spazio vettoriale reale. In realtà, sostituendo $\mathbb{R}$ con l'insieme dei numeri complessi, si ottiene la struttura di spazio vettoriale complesso. Più in generale si possono considerare spazi vettoriali su un campo qualunque[1]. Denotiamo con $\mathbb{R}$ l'insieme dei numeri reali munito delle operazioni di somma e prodotto usuali.

**Definizione 4.0.49.** Un spazio vettoriale è una terna $(V, \boxplus, \boxdot)$, dove $V$ è un insieme munito di due operazioni $\boxplus : V \times V \to V$ e $\boxdot : \mathbb{R} \times V \to V$ (una interna ed una esterna) che verificano le seguenti proprietà:

1. $\forall u, v, w \in V$ si ha $(u \boxplus v) \boxplus w = u \boxplus (v \boxplus w)$ (proprietà associativa);

2. esiste $0 \in V$ tale che $\forall u \in V$ si ha $u \boxplus 0 = 0 \boxplus u = u$ (esistenza dell'elemento neutro);

3. $\forall u \in V$ esiste $u'$ tale che $u \boxplus u' = u' \boxplus u = 0$ (esistenza dell'opposto);

4. $\forall u, v \in V$ si ha $u \boxplus v = v \boxplus u$ (proprietà commutativa);

5. $\forall u \in V$ e $\forall \alpha, \beta \in \mathbb{R}$ si ha $(\alpha + \beta) \boxdot u = (\alpha \boxdot u) \boxplus (\beta \boxdot u)$;

6. $\forall u, v \in V$ e $\forall \alpha, \in \mathbb{R}$ si ha $\alpha \boxdot (u \boxplus v) = (\alpha \boxdot u) \boxplus (\alpha \boxdot v)$;

7. $\forall u \in V$ e $\forall \alpha, \beta \in \mathbb{R}$ si ha $(\alpha\beta) \boxdot u = \alpha \boxdot (\beta \boxdot u)$;

8. $\forall u \in V$ si ha $1 \boxdot u = u$.

[1] Un campo è un insieme munito di due operazioni che hanno le stesse proprietà della somma e del prodotto tra numeri reali o complessi. Un esempio è dato dall'insieme degli interi modulo $p$, dove $p$ è un numero primo.

Le prime tre proprietà si possono riassumere dicendo che $V$ è un gruppo rispetto all'operazione $\boxplus$ mentre la quarta ci dice che il gruppo $(V, \boxplus)$ è abeliano (o commutativo). D'ora in poi per snellire la notazione indicheremo le due operazioni con $+$ e $\cdot$ come nella somma e prodotto dei numeri reali.

Gli elementi di uno spazio vettoriale saranno chiamati *vettori*.

L'elemento neutro 0 del secondo assioma è unico. Se infatti supponiamo per assurdo che ne esista un altro e lo indichiamo con $0'$, per il secondo assioma si avrebbe

$$0 = 0 + 0' = 0'.$$

Allo stesso modo, dato un vettore $u \in V$ esiste un unico $u'$ che verifica il terzo assioma. Difatti se per assurdo $u''$ verificasse la stessa proprietà, avremmo:

$$u' = u' + 0 = u' + (u + u'') = (u' + u) + u'' = 0 + u'' = u''.$$

Tale elemento verrà detto opposto di $u$ e denotato con $-u$.

Vale inoltre la legge di annullamento del prodotto:

$$\alpha \cdot u = 0 \iff \alpha = 0 \, o \, u = 0.$$

### 4.0.6 Esempi di spazi vettoriali

Di seguito daremo alcuni esempi di spazi vettoriali e di insiemi che non sono spazi vettoriali. Si verifichino le proprietà per ognuno degli esempi:

**Esempio 4.0.50.** L'insieme dei vettori liberi dello spazio $V_3$ è uno spazio vettoriale rispetto alla somma e al prodotto per uno scalare, così come definiti per i vettori liberi.

**Esempio 4.0.51.** Fissati due numeri naturali non nulli $m, n$, l'insieme delle matrici con $m$ righe ed $n$ colonne, munito delle operazioni di somma e prodotto di uno scalare per una matrice, è uno spazio vettoriale.

**Esempio 4.0.52.** Dato un numero naturale non nullo $n$, l'insieme $\mathbb{R}^n$ delle $n$-uple di numeri reali munito delle seguenti operazioni:

$$(x_1, x_2, \cdots, x_n) + (y_1, y_2, \cdots, y_n) = (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n)$$

$$\lambda(x_1, x_2, \cdots, x_n) = (\lambda x_1, \lambda x_2, \cdots, \lambda x_n) \;\; \lambda \in \mathbb{R} \,,$$

è uno spazio vettoriale.

**Esempio 4.0.53.** L'insieme $\mathbb{R}^2$ munito della somma come nell'esempio precedente e del prodotto per uno scalare come segue

$$\lambda(x_1, x_2) = (\lambda x_1, 0) \quad , \lambda \in \mathbb{R} \, ,$$

non è uno spazio vettoriale. In effetti sono verificati tutti gli assiomi tranne l'ultimo.

**Esempio 4.0.54.** L'insieme $\mathbb{R}_n[x]$ dei polinomi nella variabile $x$ a coefficienti reali e di grado minore o uguale a $n$ è uno spazio vettoriale rispetto alla somma usuale di polinomi e al prodotto di un numero reale per un polinomio.

**Esempio 4.0.55.** L'insieme $\mathbb{R}[x]$ dei polinomi nella variabile $x$ a coefficienti reali di grado qualunque è uno spazio vettoriale rispetto alla somma usuale di polinomi e al prodotto di un numero reale per un polinomio. Vedremo più avanti che questo spazio vettoriale, a differenza di quello dell'esempio precedente, non è *finitamente generato*.

**Esempio 4.0.56.** L'insieme $\mathbb{R}^2$ munito della somma e del prodotto per uno scalare definiti come segue

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, 1) \quad , \lambda(x_1, x_2) = (\lambda x_1, \lambda x_2) \quad , \lambda \in \mathbb{R} \, ,$$

non è uno spazio vettoriale. (Si veda in dettaglio quali proprietà non sono verificate)

**Esempio 4.0.57.** L'insieme $\mathbb{R}^2$ munito della somma e del prodotto per uno scalare definiti come segue

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_2, x_2) \quad , \lambda(x_1, x_2) = (\lambda x_1, \lambda x_2) \quad , \lambda \in \mathbb{R} \, ,$$

non è uno spazio vettoriale. (Si veda in dettaglio quali proprietà non sono verificate)

**Esempio 4.0.58.** L'insieme $C^r[(a, b)]$ delle funzioni reali sull'intervallo aperto $(a, b)$ e derivabili $r$ volte (per $r = 0$ si intendono le funzioni continue), è uno spazio vettoriale rispetto alla usuale somma di funzioni e prodotto di un numero reale per una funzione.

## 4.1 Sottospazi vettoriali

Dato uno spazio vettoriale $(V, +, \cdot)$ e un suo sottoinsieme $U$, è naturale chiedersi se tale $U$ sia uno spazio vettoriale rispetto alle stesse operazioni di $V$:

**Definizione 4.1.1.** Un sottoinsieme $U$ di uno spazio vettoriale $(V, +, \cdot)$ è detto sottospazio vettoriale se $(U, +, \cdot)$ è uno spazio vettoriale.

Poiché le operazioni di $V$ verificano tutte le proprietà di spazio vettoriale, l'unica cosa di cui ci si deve accertare perché $U$ sia un sottospazio vettoriale è che dati due qualunque vettori di $U$, la loro somma appartenga ancora a $U$ e che ogni volta che si moltiplica uno scalare per un vettore di $U$, si ottiene ancora un vettore di $U$. Si ha dunque la seguente proposizione:

**Proposizione 4.1.2.** *Un sottoinsieme $U$ di uno spazio vettoriale $(V, +, \cdot)$ è un sottospazio vettoriale se e solo se $\forall u, v \in U$ si ha $u + v \in U$ e $\forall \lambda \in \mathbb{R}$ $\forall u \in U$ si ha $\lambda \cdot u \in U$.*

*Dimostrazione.* La dimostrazione è lasciata per esercizio. □

Un modo equivalente per verificare se un sottoinsieme è un sottospazio vettoriale è dato dalla seguente

**Proposizione 4.1.3.** *Un sottoinsieme $U$ di uno spazio vettoriale $(V, +, \cdot)$ è un sottospazio vettoriale se e solo se $\forall u, v \in U$ e $\forall \lambda, \mu \in \mathbb{R}$ si ha $\lambda \cdot u + \mu \cdot v \in U$.*

*Dimostrazione.* Se $U$ è un sottospazio vale la Proposizione 4.1.2 e quindi $\forall u, v \in U$ e $\forall \lambda, \mu \in \mathbb{R}$ si ha $\lambda \cdot u, \mu \cdot v \in U$ e sempre per la stessa proposizione si ha $\lambda \cdot u + \mu \cdot v \in U$.

Viceversa se $\forall u, v \in U$ e $\forall \lambda, \mu \in \mathbb{R}$ si ha $\lambda \cdot u + \mu \cdot v \in U$ allora presi $\lambda = \mu = 1$ si ottiene che $\forall u, v \in U$ $u + v \in U$. Se si prende $\mu = 0$ si ottiene la seconda parte della Proposizione 4.1.2. □

## Esempi di sottospazi vettoriali

**Esempio 4.1.4.** Se si considera lo spazio vettoriale $\mathbf{M}_n(\mathbb{R})$ delle matrici quadrate di ordine $n$, l'insieme delle matrici triangolari superiori e quello delle matrici triangolari inferiori sono sottospazi vettoriali.

**Esempio 4.1.5.** Se si considera lo spazio vettoriale $\mathbf{M}_{n\times 1}(\mathbb{R})$ delle matrici colonna con $n$ righe, l'insieme delle soluzioni di un dato sistema lineare omogeneo con $n$ incognite è un suo sottospazio vettoriale.

### 4.1.1 Somma e intersezione di sottospazi vettoriali

In questo paragrafo vedremo alcuni importanti sottospazi vettoriali costruiti a partire da sottospazi dati. Innanzitutto vediamo la seguente

**Proposizione 4.1.6.** *Dati due sottospazi $X, Y$ di uno spazio vettoriale $V$, la loro intersezione è un sottospazio vettoriale.*

*Dimostrazione.* Siano $u, v \in X \cap Y$ e $a, b \in \mathbb{R}$, allora $a{\cdot}u + b{\cdot}v$ appartiene sia a $X$ che a $Y$ poiché sono sottospazi vettoriali e quindi $a{\cdot}u + b{\cdot}v \in X \cap Y$. □

In genere l'unione di due sottospazi vettoriali non è un sottospazio vettoriale. Per rendersi conto di questo fatto si consideri il seguente esempio:

**Esempio 4.1.7.** Sia $\mathbf{M}_2(\mathbb{R})$ lo spazio vettoriale delle matrici quadrate e siano $U = \left\{ \begin{pmatrix} \alpha & 0 \\ 0 & 0 \end{pmatrix} | \alpha \in \mathbb{R} \right\}$ e $V = \left\{ \begin{pmatrix} 0 & 0 \\ 0 & \beta \end{pmatrix} | \beta \in \mathbb{R} \right\}$. Gli insiemi $U, V$ sono sottospazi vettoriali di $\mathbf{M}_2(\mathbb{R})$ (da verificare!), ma non la loro unione. Infatti le matrici $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ appartengono a $U \cup V$, mentre la loro somma no.

La definizione seguente ci permette di costruire un sottospazio vettoriale a partire da due sottospazi dati:

**Definizione 4.1.8.** Dati due sottospazi $X, Y$ di uno spazio vettoriale $V$, la somma di $X$ e $Y$ è il seguente insieme:

$$X + Y = \{z \in V | z = x + y \ , x \in X \ , y \in Y\}$$

**Proposizione 4.1.9.** *La somma di due sottospazi vettoriali è un sottospazio vettoriale*

*Dimostrazione.* Siano $X, Y$ due sottospazi vettoriali di $V$ e $u, v \in X + Y$. Avremo quindi $u = x_1 + y_1$ e $v = x_2 + y_2$, dove $x_1, x_2 \in X$ e $y_1, y_2 \in Y$. Dati due numeri reali qualunque $a$, $b$ si ha:

$$a{\cdot}u + b{\cdot}v = a{\cdot}(x_1 + y_1) + b{\cdot}(x_2 + y_2) = (a{\cdot}x_1 + b{\cdot}x_2) + (a{\cdot}y_1 + b{\cdot}y_2) \in X + Y.$$

□

**Definizione 4.1.10.** La somma di due sottospazi vettoriali $X, Y$ sarà detta diretta e si scriverà $X \oplus Y$, se $X \cap Y = \{0\}$.

Vale la seguente proprietà:

**Proposizione 4.1.11.** *La somma $X + Y$ di due sottospazi vettoriali $X, Y$ è diretta se e solo se ogni vettore $u \in X + Y$ si scrive in modo unico come $u_X + u_Y$, dove $u_X \in X$ e $u_Y \in Y$.*

*Dimostrazione.* Se la somma $X + Y$ è diretta e $u = u_X + u_Y = u'_X + u'_Y$ allora si ha:

$$u_X - u'_X = u'_Y - u_Y \in X \cap Y = \{0\}$$

e quindi $u_X - u'_X = u'_Y - u_Y = 0$, che implica $u_X = u'_X$ e $u_Y = u'_Y$ e quindi la scrittura è unica.

Viceversa supponiamo che ogni vettore di $X + Y$ si scriva in modo unico. Sia $v \in X \cap Y$ e dimostriamo che $v$ è nullo. Ma $v = 0 + v = v + 0$ e $0, v \in X \cap Y$ e quindi per l'unicità della scrittura si deve avere $v = 0$. □

## 4.2 Sistemi di generatori

Sia $(V, +, \cdot)$ uno spazio vettoriale. Diamo la seguente

**Definizione 4.2.1.** Dati $n$ vettori ($n$ numero naturale non nullo) $v_1, \cdots, v_n \in V$ e $n$ scalari $\lambda_1, \cdots, \lambda_n \in \mathbb{R}$, chiameremo combinazione lineare di $v_1, \cdots, v_n \in V$ di coefficienti $\lambda_1, \cdots, \lambda_n \in \mathbb{R}$, il vettore

$$\sum_{i=1}^{n} \lambda_i \cdot v_i.$$

Utilizzando le combinazioni lineari, è possibile costruire un sottospazio vettoriale a partire da qualunque sottoinsieme di $V$. Per far questo occorre la seguente definizione:

**Definizione 4.2.2.** Dato un sottoinsieme $X$ dello spazio vettoriale $V$, si chiama *chiusura lineare* di $X$ il seguente insieme

$$L(X) = \{\sum_{i=1}^{n} \lambda_i \cdot v_i | n \in \mathbb{N}, v_i \in X, \lambda_i \in \mathbb{R}\}$$

La chiusura lineare di un sottoinsieme $X$ non è altro che l'insieme di tutte le possibili combinazioni lineari di elementi di $X$.

*Osservazione* 4.2.3. Per convenzione porremo $L(\emptyset) = \{0\}$

La seguente proposizione è lasciata come esercizio:

**Proposizione 4.2.4.** *Siano $X, Y$ due sottoinsiemi di uno spazio vettoriale $V$, tali che $X \subset Y$. Allora si ha $L(X) \subset L(Y)$.*

La chiusura lineare è un sottospazio vettoriale e tra tutti gli spazi vettoriali che contengono $X$ è il più piccolo possibile, ossìa tale che ogni sottospazio che contiene $X$ deve contenere anche $L(X)$:

**Proposizione 4.2.5.** *La chiusura lineare di un sottoinsieme $X$ di uno spazio vettoriale $V$ è il più piccolo sottospazio vettoriale di $V$ contenente $X$.*

*Dimostrazione.* Se $X = \emptyset$, per convenzione abbiamo posto $L(X) = \{0\}$ e quindi in questo caso $L(X)$ è un sottospazio vettoriale ed è contenuto in qualunque sottospazio vettoriale.

Possiamo allora supporre che $X$ sia non vuoto. Mostriamo innanzitutto che $L(X)$ è un sottospazio vettoriale. Siano $u, v \in L(X)$ e $a, b \in \mathbb{R}$. Poiché $u$ e $v$ appartengono alla chiusura lineare di $X$, si potranno scrivere come combinazione lineare di elementi di $X$ e quindi avremo $u = \sum_{i=1}^{k} \alpha_i{\cdot}u_i$ e $v = \sum_{j=1}^{h} \beta_j{\cdot}v_j$ con $u_i, v_j \in X$. Dobbiamo dimostrare che il vettore $a{\cdot}u + b{\cdot}v$ appartiene anch'esso a $L(X)$. Abbiamo allora

$$a{\cdot}u + b{\cdot}v = a{\cdot}\sum_{i=1}^{k} \alpha_i{\cdot}u_i + b{\cdot}\sum_{j=1}^{h} \beta_j{\cdot}v_j = \sum_{i=1}^{k}(a\alpha_i){\cdot}u_i + \sum_{j=1}^{h}(b\beta_j){\cdot}v_j\ ,$$

che è ancora una combinazione lineare di elementi di $X$.

Per dimostrare che $L(X)$ è il più piccolo sottospazio contenente $X$, consideriamo un sottospazio vettoriale $W$ tale che $X \subset W$ e sia $u \in L(X)$. Vogliamo dimostrare che $u \in W$. Dato che $u \in L(X)$, esistono dei vettori $u_1, \cdots, u_s \in X$ e degli scalari $a_1, \cdots, a_s \in \mathbb{R}$ tali che $u = \sum_{i=1}^{s} a_i{\cdot}u_i$. Ma $X \subset W$ e quindi i vettori $u_i$ appartengono a $W$. Poiché $W$ è un sottospazio vettoriale, una combinazione lineare di vettori di $W$ appartiene ancora a $W$ e quindi $\sum_{i=1}^{s} a_i{\cdot}u_i = u \in W$. Questo significa che $L(X) \subset W$. □

Un sottoinsime $X$ di uno spazio vettoriale $V$ è un sottospazio vettoriale se e solo se $X = L(X)$. (da verificare per esercizio !)

**Definizione 4.2.6.** Un sottoinsieme $X$ di uno spazio vettoriale è detto *sistema di generatori* se $L(X) = V$, cioé se ogni vettore di $V$ è combinazione lineare di elementi di $X$.

*Osservazione* 4.2.7. Ogni spazio vettoriale $V$ ammette un sistema di generatori, infatti si ha $L(V) = V$.

**Definizione 4.2.8.** Uno spazio vettoriale $V$ è detto *finitamente generato* se esiste un numero finito di vettori $v_1, \cdots, v_n$ tali che $V = L(v_1, \cdots, v_n)$.

Vediamo ora alcune proprietà della chiusura lineare che saranno molto utili in seguito:

**Proposizione 4.2.9.** *Sia $X$ un sottoinsieme di uno spazio vettoriale $V$ e $Y \subset L(X)$, allora si ha $L(X) = L(X \cup Y)$.*

*Dimostrazione.* Il caso in cui $X$ è vuoto è banale, quindi supponiamo $X$ non vuoto. Poiché per la Proposizione 4.2.4 si ha $L(X) \subset L(X \cup Y)$, ci rimane da dimostrare l'atra inclusione. Sia quindi $u \in L(X \cup Y)$. Allora $u$ si potrà scrivere nella seguente forma:

$$u = \sum_i a_i \cdot u_i + \sum_j b_j \cdot v_j \quad , a_i, b_j \in \mathbb{R} \quad , u_i \in X, v_j \in Y.$$

Dato che $Y \subset L(X)$ i vettori $v_j$ sono combinazioni lineari di vettori di $X$ e quindi $u$ è combinazione linere di vettori di $X$, ossìa $L(X \cup Y) \subset L(X)$. □

Il seguente corollario è particolarmente utile perché ci fornisce un metodo per eliminare elementi superflui da un sistema di generatori e ci dice inoltre che aggiungendo elementi ad un sistema di generatori, si ottiene ancora un sistema di generatori.

**Corollario 4.2.10.** *Sia $V$ uno spazio vettoriale:*

1. *se $X$ è un sistema di generatori per $V$ e $X \subset Y \subset V$ allora anche $Y$ è un sistema di generatori;*

2. *se $Z$ è un sistema di generatori per $V$ e $Y \subset Z$ è tale che $Y \subset L(Z - Y)$ allora anche $Z - Y$ è un sistema di generatori.*

*Dimostrazione.*

1. Si ha $Y \subset L(X) = V$ e quindi dato che $X$ è un sistema di generatori ed è contenuto in $Y$, per la Proposizione 4.2.9 $V = L(X) = L(X \cup Y) = L(Y)$ e dunque $Y$ genera $V$.

2. Basta porre $X = Z - Y$ ed applicare la Proposizione 4.2.9 a $X, Y$.

□

## 4.3 Dipendenza lineare

In questo paragrafo studieremo il concetto di base di uno spazio vettoriale. Per far questo ci occorre la seguente

**Definizione 4.3.1.** Sia $V$ uno spazio vettoriale. I vettori $v_1, \cdots, v_n \in V$ $(n \geq 1)$ saranno detti *linearmente dipendenti* se esistono $\lambda_1, \cdots, \lambda_n \in \mathbb{R}$ non tutti nulli tali che

$$\sum_{i=1}^{n} \lambda_i \cdot v_i = 0.$$

In caso contrario saranno detti *linearmente indipendenti*.

**Proposizione 4.3.2.** *Una $n$-pla di vettori di uno spazio vettoriale è linearmente dipendente se e solo se uno dei vettori è combinazione lineare degli altri.*

*Dimostrazione.* Siano $v_1, \cdots, v_n \in V$ linearmente dipendenti. Allora esistono degli scalari $a_1, \cdots, a_n \in \mathbb{R}$ non tutti nulli tali che:

$$\sum_{i=1}^{n} a_i \cdot v_i = 0.$$

Possiamo supporre che il primo scalare sia non nullo, altrimenti basta riordinare i vettori. Allora avremo:

$$a_1 \cdot v_1 = -\sum_{i=2}^{n} a_i \cdot v_i$$

e poiché $a_1$ è non nullo, applicando le proprietà del prodotto per uno scalare, si avrà:

$$v_1 = -\sum_{i=2}^{n} (a_1^{-1} a_i) \cdot v_i$$

e quindi un vettore è combinazione lineare degli altri.

Viceversa supponiamo che un vettore sia combinazione lineare degli altri. Riordinandoli possiamo supporre che questo vettore sia $v_1$, trovando

$$v_1 = \sum_{i=2}^{n} b_i \cdot v_i \ , \ b_i \in \mathbb{R} \ ,$$

da cui otteniamo la seguente combinazione lineare

$$\sum_{i=2}^{n} b_i \cdot v_i - v_1 = 0$$

in cui il coefficiente di $v_1$ è non nullo. □

Ora possiamo definire un insieme linearmente dipendente:

**Definizione 4.3.3.** Un sottoinsieme $X$ di uno spazio vettoriale $V$ è detto linearmente dipendente se esistono dei vettori $v_1, \cdots, v_n \in X$ che sono linearmente dipendenti. Altrimenti sarà detto linearmente indipendente.

Se $X$ è linearmente indipendente ogni volta che si prendono dei vettori di $X$ questi sono linearmente indipendenti.

*Osservazione* 4.3.4. L'insieme vuoto è banalmente linearmente indipendente.

**Proposizione 4.3.5.** *Se $X \subset V$ è linearmente dipendente e $Y \subset V$ allora anche $X \cup Y$ è linearmente dipendente.*

*Dimostrazione.* Dato che $X$ è linearmente dipendente, allora esistono $v_1, \cdots, v_n \in X$ che sono linearmente dipendenti. Ma $v_1, \cdots, v_n \in X \cup Y$ e sono linearmente dipendenti e dunque $X \cup Y$ è linearmente dipendente. □

## 4.4 Basi di uno spazio vettoriale

Il concetto di base di uno spazio vettoriale generalizza quello di base dell'insieme dei vettori liberi dello spazio:

**Definizione 4.4.1.** Un sottoinsieme $B$ di uno spazio vettoriale $V$ è detto *base* se sono verificate le seguenti condizioni:

1. $L(B) = V$, cioé $B$ è un sistema di generatori;

2. $B$ è linearmente indipendente.

Ecco alcuni esempi di basi di alcuni spazi vettoriali:

**Esempio 4.4.2.** Una base dello spazio vettoriale $\mathbb{R}^n$ è data dalla $n$-upla di vettori $e_1, \cdots, e_n \in \mathbb{R}^n$ tali che la $i$-esima componente della $n$-upla $e_i$ valga 1, mentre tutte le altre sono nulle. Questa base viene anche chiamata *base canonica* di $\mathbb{R}^n$.

**Esempio 4.4.3.** Una base dello spazio vettoriale $\mathbb{R}[x]$ dei polinomi nella variabile $x$ è data dall'insieme di tutti i monomi. Questo spazio vettoriale non è finitamente generato e ammette una base infinita.

**Esempio 4.4.4.** Una base dello spazio vettoriale $\mathbf{M}_{m\times n}$ delle matrici con $m$ righe e $n$ colonne è data dall'insieme delle matrici $\{E_{hk}\}$ definite nel modo seguente

$$[E_{hk}]_{ij} = \delta_{ih}\delta_{jk}.$$

Quindi la matrice $\{E_{hk}\}$ ha entrate tutte nulle tranne quella della $h$-esima riga, $k$-esima colonna che vale 1. Questa base è composta da $m{\cdot}n$ vettori.

### 4.4.1 Dimensione e esistenza di una base

In questo paragrafo ci proponiamo di dimostrare che ogni spazio vettoriale finitamente generato ammette una base e che due basi dello stesso spazio hanno lo stesso numero di elementi. Questo vale anche per gli spazi non finitamente generati ma richiede l'uso dell'*assioma di scelta*.

D'ora in poi $V$ sarà uno spazio vettoriale finitamente generato, se non specificato diversamente. Iniziamo con alcune proposizioni che ci serviranno per il seguito:

**Proposizione 4.4.5.** *Se $X \subset V$ è linearmente indipendente e $u \in V - L(X)$ allora anche $X \cup \{u\}$ è linearmente indipendente.*

*Dimostrazione.* Siano $u_1, \cdots, u_k \in X \cup \{u\}$ e $a_1, \cdots a_k \in \mathbb{R}$ e sia

$$\sum_{i=1}^{k} a_i \cdot u_i = 0.$$

Se nella combinazione lineare appaiono solo vettori di $X$ allora i coefficienti sono tutti nulli perchè $X$ è linearmente indipendente. Quindi uno dei vettori deve essere $u$ e il suo coefficiente non deve essere nullo. Ma allora $u$ si può esprimere come combinazione lineare di vettori di $X$ il che è assurdo perché abbiamo supposto che $u \notin L(X)$. □

**Proposizione 4.4.6.** *Un sottoinsieme $X$ di $V$ è linearmente dipendente se e solo se esiste $u \in X$ tale che $u \in L(X - \{u\})$.*

*Dimostrazione.* Se $X$ è linearmente dipendente allora esiste $u \in X$ che è combinazione lineare di altri vettori di $X$ cioé $u \in L(X - \{u\})$.

Viceversa se esiste $u \in X$ tale che $u \in L(X - \{u\})$, allora $u$ è combinazione lineare di certi vettori $v_1 \cdots, v_n \in X$ diversi da $u$ e quindi i vettori $v_1 \cdots, v_n, u$ sono linearmente dipendenti e per definizione anche $X$. □

**Lemma 4.4.7.** *Sia $X$ un sottoinsieme finito di $V$. Allora esiste $X' \subset X$ (necessariamente finito) tale che $L(X) = L(X')$ e $X'$ è linearmente indipendente.*

*Dimostrazione.* Dimostriamo il Lemma per induzione sul numero $n$ di elementi di $X = \{v_1 \cdots, v_n\}$.

Per $n = 1$ si ha $X = \{v_1\}$. Se $v_1$ è non nullo allora è linearmente indipendente e quindi basta porre $X' = X$. Se $v_1 = 0$ allora poniamo $X' = \emptyset$ e abbiamo che $X'$ è linearmente indipendente e inoltre $L(X') = L(\emptyset) = \{0\} = L(X)$. Quindi la proprietà è vera per $n = 1$.

Supponiamo che la proprietà sia vera per un sottoinsieme di $n$ vettori e dimostriamo che questo implica che è vera per $n + 1$ vettori.

Sia dunque $X = \{v_1 \cdots, v_{n+1}\}$. Se $X$ è linearmente indipendente basta porre $X' = X$ e la proprietà è soddisfatta. Se $X$ è linearmente dipendente allora per la Proposizione 4.4.6 esiste $u \in X$ tale che $u \in L(X - \{u\})$. Posto $X' = X - \{u\}$, abbiamo $L(X') = L(X)$ per la Proposizione 4.2.9. Ora possiamo applicare l'ipotesi induttiva a $X'$ che ha $n$ elementi. Esisterà allora $X'' \subset X' \subset X$ tale che $X''$ è linearmente indipendente e $L(X'') = L(X')$. Poiché $L(X') = L(X)$, abbiamo terminato. □

**Corollario 4.4.8.** *Uno spazio vettoriale finitamente generato ammette sempre una base.*

*Dimostrazione.* Essendo $V$ finitamente generato si ha $V = L(X)$ per un certo insieme finito $X$. Applicando il Lemma 4.4.7 a $X$, si ottiene $X'$ linearmente indipendente tale che $L(X') = L(X) = V$ e quindi $X'$ è una base. Inoltre la base ottenuta è finita in quanto sottoinsieme di $X$. □

Ci occorrono ora i seguenti risultati:

**Teorema 4.4.9** (Teorema dello scambio)**.** *Sia $V$ uno spazio vettoriale finitamente generato e $B$ una sua base finita di cardinalità $n$. Sia $X \subset V$ linearmente indipendente e di cardinalità $h \leq n$. Allora esiste $Y \subset B$ di cardinalità $h$ e tale che $(B - Y) \cup X$ sia una base di $V$. Inoltre se $h = n$ allora anche $X$ è una base.*

**Corollario 4.4.10.** *Sia $V$ uno spazio vettoriale finitamente generato e $B$ una sua base finita di cardinalità $n$. Allora ogni $X \subset V$ linearmente indipendente ha cardinalità $h \leq n$.*

*Dimostrazione.* Supponiamo per assurdo che $h > n$ e sia $S \subset X$ tale che $\mathbf{card}(S) = n$. Se $S$ è linearmente dipendente allora lo è anche $X$. Dunque $S$ è linearmente indipendente e, per il Teorema dello scambio, è una base di $V$ e quindi $L(S)$=V. Siccome $S$ è strettamente contenuto in $X$, allora esiste $v \in X - S$. Poichè $S$ genera $V$ allora $v \in L(S)$. Quindi $S \cup \{v\}$ è linearmente dipendente e inoltre $S \cup \{v\} \subset X$. Questo implica che anche $X$ è linearmente dipendente il che è assurdo. □

Possiamo ora enunciare il seguente teorema:

**Teorema 4.4.11.** *Sia $V$ uno spazio vettoriale finitamente generato. Allora abbiamo:*

1. *esiste una base finita $B$ di $V$;*
2. *due basi di $V$ hanno la stessa cardinalità.*

*Dimostrazione.*
1. L'esistenza di una base è garantita dal Corollario 4.4.8;
2. siano $B, B'$ due basi di cardinalità $h$ e $k$ rispettivamente. Per il Corollario 4.4.10, poiché $B$ è una base e $B'$ un sottoinsieme linearmente indipendente si ha $k \leq h$. Per lo stesso corollario, poiché $B'$ è una base e $B$ linearmente indipendente si ha $h \leq k$ e quindi $h = k$. Questo mostra che due basi hanno la stessa cardinalità.

□

Il teorema precedente ci dice che il numero di elementi di due basi qualunque è lo stesso. Questo permette di dare la seguente definizione:

**Definizione 4.4.12.** Dato uno spazio vettoriale finitamente generato si dice *dimensione* di $V$ e si indica con $\dim V$ la cardinalità di una qualunque sua base.

Se uno spazio vettoriale non ammette una base finita si pone $\dim V = \infty$.

Il seguente teorema ci permette di caratterizzare il numero di elementi di una base come *numero minimo* di generatori o anche come *numero massimo* di vettori linearmente indipendenti.

**Teorema 4.4.13.** *Sia $V$ uno spazio vettoriale finitamente generato e* $\dim V = n$. *Allora si ha:*

1. *per ogni sottoinsieme $X$ linearmente indipendente si ha* $\mathbf{card}(X) \leq n$ *e inoltre* $\mathbf{card}(X) = n$ *se e solo se $X$ è una base;*

2. *per ogni sottoinsieme di generatori $X$ si ha* $\mathbf{card}(X) \geq n$ *e inoltre* $\mathbf{card}(X) = n$ *se e solo se $X$ è una base.*

*Dimostrazione.*

1. Se $X$ è linearmente indipendente, per il Corollario 4.4.10 si ha $\mathbf{card}(X) \leq n$.

   Se $\mathbf{card}(X) = n$, mostriamo che si tratta di una base. Dobbiamo dimostrare che $X$ genera $V$. Se per assurdo questo non fosse vero, esisterebbe $u \in V - L(X)$ e quindi $X \cup \{u\}$ sarebbe linearmente indipendente. Ma $\mathbf{card}(X \cup \{u\}) = \mathbf{card}(X) + 1 = n + 1 > n$ e ciò sarebbe assurdo per il Corollario 4.4.10.

   Viceversa se $X$ è una base allora ha cardinalità $n$ per il Teorema 4.4.11.

2. Se $X$ è sistema di generatori e per assurdo $\mathbf{card}(X) < n$ allora per il Lemma 4.4.7 esiste $X' \subset X$ linearmente indipendente che genera $V = L(X)$. Ma allora si avrebbe $\mathbf{card}(X') \leq \mathbf{card}(X) < n$ e quindi $X'$ sarebbe una base con cardinalità minore di $n$, il che è assurdo.

   Se $\mathbf{card}(X) = n$ e $X$ non fosse una base, allora $X$ sarebbe linearmente dipendente e per il Lemma 4.4.7 esisterebbe $X' \subset X$ linearmente indipendente che genera $V = L(X)$. Inoltre $\mathbf{card}(X') < \mathbf{card}(X) = n$ perché $X'$ è strettamente contenuto in $X$ in quanto linearmente indipendente mentre $X$ non lo è. Ma allora $X'$ sarebbe una base di $V$ con cardinalità minore di $n$ il che è assurdo.

   Viceversa se $X$ è una base, ovviamente $\mathbf{card}(X) = n$.

□

Terminiamo con un teorema utile per la costruzione di una base a partire da un sistema di vettori linearmente indipendenti:

**Teorema 4.4.14** (Teorema del completamento di una base)**.** *Sia $V$ uno spazio vettoriale di dimensione finita $n$ e $X$ un sottoinsieme linearmente indipendente con* $\mathbf{card}(X) = h < n$. *Allora esiste $Y \subset V$ tale che* $\mathbf{card}(Y) = n - h$ *e $X \cup Y$ è una base.*

*Dimostrazione.* Sia $B$ una base di $V$. Per il Teorema dello scambio esiste $Z$ con $\mathbf{card}(Z) = h$ tale che $(B - Z) \cup X$ è una base. Allora $Y = B - Z$ soddisfa le proprietà volute. □

Si dice che la base $X \cup Y$ è ottenuta completando $X$.

### 4.4.2 Componenti rispetto a una base

Dato uno spazio vettoriale $V$ di dimensione finita $n$, una sua base ordinata $B = (e_1, \cdots, e_n)$ è una $n$-upla ordinata tale che $\{e_1, \cdots, e_n\}$ sia una base.

Dato un qualunque vettore $v \in V$, esiste una $n$-upla ordinata $(v_1, \cdots, v_n) \in \mathbb{R}^n$ tale che $v = \sum_{i=1}^n v_i \cdot e_i$. Poichè gli elementi della base sono linearmente indipendenti la $n$-upla è univocamente determinata. infatti se $(v'_1, \cdots, v'_n)$ fosse un'altra $n$-upla tale che $v = \sum_{i=1}^n v'_i \cdot e_i$, si avrebbe

$$\sum_{i=1}^{n}(v_i - v'_i) \cdot e_i = 0$$

e quindi, per la lineare indipendenza di $B$, $v_i = v'_i \ \forall i = 1, \cdots, n$.

**Definizione 4.4.15.** Data una base ordinata $B = (e_1, \cdots, e_n)$, gli unici $v_i$ tali che l'$n$-upla ordinata $(v_1, \cdots, v_n) \in \mathbb{R}^n$ sia tale che $v = \sum_{i=1}^n v_i \cdot e_i$, sono detti *componenti* di $v$ rispetto alla base $B$.

Data una base di $V$, l'introduzione delle componenti ci permette di costruire una applicazione biunivoca $f : V^n \to \mathbb{R}^n$ che ad ogni vettore $v \in V$ associa la sua $n$-upla di componenti $(v_1, \cdots, v_n)$ rispetto a $B$.

Si osservi che, se si considera in $\mathbb{R}^n$ l'usuale struttura di spazio vettoriale, alla somma di due vettori di $V$, tramite $f$ è associata la $n$-upla data dalla somma (in $\mathbb{R}^n$) delle loro componenti. Allo stesso modo al prodotto di uno scalare $\alpha$ per un vettore $v \in V$, tramite $f$ è associato il prodotto (in $\mathbb{R}^n$) di $\alpha$ per la $n$-upla delle componenti di $v$. Questo ci permette di identificare lo spazio vettoriale $V$ con $\mathbb{R}^n$. Più in là vedremo che questa applicazione $f$ è un esempio di applicazione vettoriale tra spazi vettoriali.

### 4.4.3 Dimensione di sottospazi vettoriali

Un sottospazio $W$ di uno spazio vettoriale $V$ di dimensione $n$, deve necessariamente avere dimensione finita. Infatti se $u_1, \cdots, u_k \in W$ sono linearmente indipendenti allora lo sono anche come vettori di $V$ e quindi il numero massimo di vettori linearmente indipendenti di $W$ è minore o uguale a $\dim V$.

Si osservi che $\dim W = 0$ se e solo se $W = \{0\}$ e $\dim W = n$ se e solo se $V = W$. (Verificare!)

**Esercizio 4.4.16.** Si dimostri che se $W_1, W_2$ sono sottospazi vettoriali di $V$ allora $W_1 + W_2 = L(W_1 \cup W_2)$.

Diamo inoltre la seguente definizione:

**Definizione 4.4.17.** Dato un numero finito di vettori $v_1, \cdots, v_k \in V$, si dice *rango* dei vettori $v_1, \cdots, v_k$ il numero massimo di vettori linearmente indipendenti, cioè la dimensione di $L(v_1, \cdots, v_k)$.

*Osservazione* 4.4.18. Se $A$ è una matrice con $m$ righe ed $n$ colonne, le sue righe $A_1, \cdots, A_m$(equiv. le colonne $A^1, \cdots, A^n$) possono essere considerate come dei vettori dello spazio vettoriale delle matrici riga (equiv. colonna) con $n$ colonne (equiv. $m$ righe). Si definisce quindi il rango per riga di A la dimensione di $L(A_1, \cdots, A_m)$ e il rango per colonna come la dimensione di $L(A^1, \cdots, A^n)$. Si può dimostrare che queste due quantità sono uguali al rango della matrice $A$.

Il seguente risultato permette di calcolare la dimensione della somma di due sottospazi finitamente generati:

**Teorema 4.4.19** (Relazione di Grassmann)**.** *Dati due sottospazi finitamente generati $W_1, W_2$ di uno spazio vettoriale $V$ (non necessariamente finitamente generato), allora anche $W_1 + W_2$ e $W_1 \cap W_2$ sono finitamente generati e si ha:*

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2) .$$

*Dimostrazione.* Poiché $W_1 \cap W_2$ è un sottospazio di $W_1$ e $W_2$ allora è necessariamente finitamente generato. Il sottospazio $W_1 + W_2$ è generato dall'unione di generatori di $W_1$ e di $W_2$ i quali sono finitamente generati e quindi anche $W_1 + W_2$ lo è.

Siano ora $r = \dim W_1$ e $s = \dim W_2$. Se $W_1 \subset W_2$ o $W_2 \subset W_1$ la relazione è banalmente soddisfatta. Possiamo quindi supporre che non valga nessuna delle due precedenti inclusioni. Distinguiamo due casi:

Caso 1) Supponiamo $W_1 \cap W_2 = \{0\}$. Consideriamo una base $f_1, \cdots, f_r$ di $W_1$ e una base $g_1, \cdots, g_s$ di $W_2$. Consideriamo una loro combinazione lineare che dia il vettore nullo:

$$\sum_i a_i \cdot f_i + \sum_j b_j \cdot g_j = 0 .$$

Se i $b_j$ fossero tutti nulli si avrebbe

$$\sum_i a_i \cdot f_i = 0 ,$$

e allora sarebbero nulli anche gli $a_i$ perché $f_1, \cdots, f_r$ sono linearmente indipendenti. Quindi almeno uno dei $b_j$ è non nullo e quindi il vettore $u = \sum_j b_j \cdot g_j = -\sum_i a_i \cdot f_i$ non è nullo e appartiene a $W_1 \cap W_2$ che è assurdo perché abbiamo supposto $W_1 \cap W_2 = \{0\}$.

Quindi l'insieme $\{f_1, \cdots, f_r\} \cup \{g_1, \cdots, g_s\}$ è linearmente indipendente e anche un sistema di generatori per $W_1 + W_2$ e dunque una sua base. In tal caso la relazione di Grassman è verificata poiché $\dim(W_1 \cap W_2) = 0$.

Caso 2) Supponiamo che $W_1 \cap W_2 \neq \{0\}$. Supponiamo che $\dim(W_1 \cap W_2) = h$ e sia $\{e_1, \cdots, e_h\}$ una base di $W_1 \cap W_2$. Poiché $W_1 \cap W_2$ è un sottospazio di $W_1$ possiamo completare la base di $W_1 \cap W_2$ per ottenere una base $B_1 = \{e_1, \cdots, e_h\} \cup \{f_1, \cdots, f_{r-h}\}$ di $W_1$. Allo stesso modo troviamo una base $B_2 = \{e_1, \cdots, e_h\} \cup \{g_1, \cdots, g_{s-h}\}$ di $W_2$. Vogliamo ora dimostrare che $B = \{e_1, \cdots, e_h\} \cup \{f_1, \cdots, f_{r-h}\} \cup \{g_1, \cdots, g_{s-h}\}$ è una base di $W_1 + W_2$. Poichè per costruzione questo insieme genera $W_1 + W_2$, ci basta dimostrare l'indipendenza lineare di $B$. Si noti che una volta dimostrato ciò, la relazione di Grassmann è verificata in quanto si ha:

$$\begin{aligned} \dim(W_1 + W_2) = & \quad \mathbf{card}(B) = (r - h) + (s - h) + h = r + s - h \\ = & \quad \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2) \; . \end{aligned}$$

Ora dimostriamo che $B$ è linearmente indipendente. Supponiamo che non lo sia e consideriamo una combinazione lineare dei suoi elementi a coefficienti non tutti nulli che dia il vettore nullo:

$$\sum_k c_k \cdot e_k + \sum_i a_i \cdot f_i + \sum_j b_j \cdot g_j = 0 \; .$$

Se i $b_j$ fossero tutti nulli si avrebbe

$$\sum_k c_k \cdot e_k + \sum_i a_i \cdot f_i = 0 \; ,$$

e quindi anche i $c_k$ e gli $a_i$ sarebbero nulli poiché $\{e_1, \cdots, e_h\} \cup \{f_1, \cdots, f_{r-h}\}$ è una base di $W_1$. Quindi il vettore

$$u = \sum_j b_j \cdot g_j = -\sum_k c_k \cdot e_k - \sum_i a_i \cdot f_i$$

sarebbe non nullo. Inoltre si ha $u \in W_2$ poiché combinazione lineare degli elementi di $B_2$ e $u \in W_1$ in quanto combinazione lineare di elementi di $B_1$. Ma allora avremmo $u \in W_1 \cap W_2$. Quindi si avrebbe $u = \sum_j b_j \cdot g_j = \sum_l d_l \cdot e_l$ da cui si ha

$$\sum_j b_j \cdot g_j - \sum_l d_l \cdot e_l = 0 \; .$$

in cui non tutti i coefficienti sono nulli (poiché i $b_j$ non sono tutti nulli), il che è assurdo perchè $B_2$ è una base di $W_2$ e quindi linearmente indipendente.

□

*Osservazione* 4.4.20. Si noti che, per la relazione di Grassmann, la somma $W_1 + W_2$ è diretta se e solo se $\dim(W_1 + W_2) = \dim W_1 + \dim W_2$.

### 4.4.4 Somma di sottospazi

La definizione di somma di due sottospazi vettoriali di uno spazio vettoriale $V$, si può generalizzare al caso di un numero finito di sottospazi $W_1, \cdots, W_k$. La definizione è analoga a quella di somma di due sottospazi:

$$+_{i=1}^{k} W_i = W_1 + \cdots + W_k = \{z \in V | z = \sum_{i=1}^{k} u_i \ , u_i \in W_i\}.$$

Si verifica facilmente che il sottospazio somma non cambia se si cambia l'ordine dei sottospazi $W_1, \cdots, W_k$. Anche per un numero finito di sottospazi si definisce la somma diretta, in modo leggermente diverso:

**Definizione 4.4.21.** Dato un numero finito di sottospazi $W_1, \cdots, W_k$ di uno spazio vettoriale $V$, si dice che la loro somma è diretta e verrà indicata $\oplus_{i=1}^{k} W_i$, se l'intersezione di ogni $W_i$ con $+_{j=1, j\neq i}^{k} W_j$ è vuota.

Nel caso di più di due sottospazi, perchè la somma sia diretta, deve essere vuota l'intersezione di ognuno dei sottospazi con la somma dei restanti. Questo ci assicura l'unicità della decomposizione come nel caso della somma diretta di due sottospazi:

**Proposizione 4.4.22.** *La somma di un numero finito di sottospazi $W_1, \cdots, W_k$ è diretta se e solo se ogni per ogni vettore $u$ di $+_{i=1}^{k} W_i$ esiste un'unica k-upla $(u_1, \cdots, u_k) \in W_1 \times \cdots \times W_k$ tale che $u = \sum_{i=1}^{k} u_i$.*

*Dimostrazione.* La dimostrazione è simile al caso della somma di due sottospazi ed è lasciata per esercizio. □

### 4.4.5 Cambiamento di base e matrice di passaggio

Consideriamo due basi $B_1 = \{e_1, \cdots, e_n\}$ e $B_2 = \{e'_1, \cdots, e'_n\}$ di uno spazio vettoriale $V^n$. Dato un vettore $v$ di $V^n$, questo avrà delle componenti rispetto ad ognuna delle due basi e quindi vogliamo capire in che modo sono legate queste componenti.

Sia dunque $v = \sum_{j=1}^{n} v_j \cdot e_j = \sum_{i=1}^{n} v'_i \cdot e'_i$. Poichè $B_2$ è una base, ogni elemento di $B_1$ avrà delle componenti rispetto a $B_2$. Scriviamo queste componenti come segue:

$$e_j = \sum_{i=1}^{n} a_{ij} \cdot e'_i \ , j = 1, \cdots, n.$$

In tal modo costruiamo una matrice $A = (a_{ij})$ in cui la $j$-esima colonna è formata dalle componenti di $e_j$ rispetto alla base $B_2$.

Ora possiamo sostituire $e_j$ nell'espressione di $v$ e ottenere

$$v \quad = \sum_{j=1}^{n} v_j \cdot (\sum_{i=1}^{n} a_{ij} \cdot e'_i) = \sum_{i=1}^{n} (\sum_{j=1}^{n} v_j a_{ij}) \cdot e'_i.$$

Confrontando le due espressioni di $v$, si ha

$$\sum_{i=1}^{n} (\sum_{j=1}^{n} v_j a_{ij}) \cdot e'_i = \sum_{i=1}^{n} v'_i \cdot e'_i$$

e quindi, per l'unicità delle componenti otteniamo $v'_i = \sum_{j=1}^{n} a_{ij} v_j$, che è il legame tra le componenti rispetto a $B_2$ e quelle rispetto a $B_1$. Questo legame si può esprimere in forma matriciale. In fatti, detta $X$ la matrice colonna delle componenti di $v$ rispetto a $B_1$ e $X'$ quella delle componenti di $v$ rispetto a $B_2$ si ha:

$$X' = AX.$$

La matrice $A$ è detta *matrice di passaggio* dalla base $B_1$ alla base $B_2$. Vediamo un esempio:

**Esempio 4.4.23.** Consideriamo in $\mathbb{R}^3$ la base canonica $B_1 = (e_1, e_2, e_3)$ e la base $B_2 = (e'_1 = (1, 1, 1), e'_2 = (1, 1, 0), e'_3 = (1, 0, 0))$ e sia $v = (2, 0, -1)$. La matrice colonna delle componenti di $v$ rispetto a $B_1$ è data da $X = \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}$. Troviamo ora le componenti di ogni elemento della base canonica rispetto alla base $B_2$.

Si ha $e_1 = e'_3$ quindi le sue componenti sono $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ che ci da la prima colonna della matrice di passaggio; $e_2 = e'_2 - e'_3$ e dunque le componenti sono $\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$; $e_3 = e'_1 - e'_2$ e quindi le sue componenti sono $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$. La matrice di passaggio da $B_1$ a $B_2$ sarà:

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

Quindi la colonna $X'$ delle componenti di $v$ rispetto a $B_2$ è data da:

$$X' = AX = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}.$$

Si può verificare che in effetti si ha $v = -e'_1 + e'_2 + 2e'_3$.

Supponiamo ora di avere tre basi $B_1 = \{e_1, \cdots, e_n\}$, $B_2 = \{e'_1, \cdots, e'_n\}$ e $B_3 = \{e''_1, \cdots, e''_n\}$ e siano $A = (a_{ij}), B = (b_{ij}), C = (c_{ij})$ le matrici di passaggio rispettivamente da $B_1$ a $B_2$, da $B_2$ a $B_3$ e da $B_1$ a $B_3$. Avremo quindi

$$e_j = \sum_{i=1}^{n} a_{ij} \cdot e'_i \ , e'_i = \sum_{k=1}^{n} b_{ki} \cdot e''_k \ , e_j = \sum_{k=1}^{n} c_{kj} \cdot e''_k \ .$$

Si ottiene

$$e_j = \sum_{i=1}^{n} a_{ij} \cdot e'_i = \sum_{i=1}^{n} a_{ij} \cdot (\sum_{k=1}^{n} b_{ki} \cdot e''_k) = \sum_{k=1}^{n} \cdot (\sum_{i=1}^{n} b_{ki} a_{ij}) \cdot e''_k .$$

Confrontando quanto ottenuto con $e_j = \sum_{k=1}^{n} c_{kj} \cdot e''_k$, per l'unicità delle componenti di ottiene $c_{kj} = \sum_{i=1}^{n} b_{ki} a_{ij}$ che matricialmente significa $C = BA$.

Si osservi che se la base $B_3$ coincide con la base $B_1$, allora $C = I$ e $B$ è la matrice di passaggio da $B_2$ a $B_1$ e quindi si ha $BA = I$ e questo significa che $A$ e $B$ sono una l'inversa dell'altra.

# Capitolo 5

# Applicazioni lineari

Dati due spazi vettoriali $V, W$, studieremo le applicazioni tra di essi. Non ci interessano delle applicazioni qualunque tra $V$ e $W$, bensì delle applicazioni che in qualche modo tengano conto della struttura di spazio vettoriale e che la conservino. Diamo quindi la seguente definizione:

**Definizione 5.0.24.** Siano $(V, +, \cdot), (W, \boxplus, \boxdot)$ due spazi vettoriali. Una applicazione $f : V \to W$ è detta *applicazione lineare* se per ogni $u, v \in V$ e per ogni $a \in \mathbb{R}$ si ha

$$f(u+v) = f(u) \boxplus f(v) \tag{5.1}$$
$$f(a \cdot u) = a \boxdot f(u). \tag{5.2}$$

Diamo ora alcune definizioni:

**Definizione 5.0.25.** Una applicazione lineare $f : V \to W$ è detta

- epimorfismo se è suriettiva;
- isomorfismo se è bigettiva;
- endomorfismo (o operatore lineare) se $V = W$;
- automorfismo se è un isomorfismo e $V = W$;
- forma lineare se $W = \mathbb{R}$.

Denoteremo con $L(V, W)$ l'insieme delle applicazioni lineari da $V$ in $W$, con $End(V)$ l'insieme degli endomorfismi di $V$ e con $Aut(V)$ l'insieme degli automorfismi di $V$.

*Osservazione* 5.0.26. Le condizioni che deve verificare $f$ sono equivalenti alla condizione seguente (verificare!):

$$f(\sum_{i=1}^{k} a_i \cdot u_i) = \sum_{i=1}^{k} a_i \boxdot f(u_i)$$

per ogni $k \in \mathbb{N}^*$, $\forall u_1, \cdots, u_k \in V$ e $\forall a_1, \cdots, a_k \in \mathbb{R}$.

Prima di studiare le proprietà delle applicazioni lineari vediamo alcuni esempi:

**Esempio 5.0.27.** Dati gli spazi vettoriali $\mathbb{R}^3, \mathbb{R}^2$, le applicazioni $f, g : \mathbb{R}^3 \to \mathbb{R}^2$ definite da $f(x, y, z) = (2x + y - 3z, 0)$ e $g(x, y, z) = (x + y - 10z, -x - z)$ sono applicazioni lineari.

**Esempio 5.0.28.** La funzione $f : \mathbb{R} \to \mathbb{R}$ definita da $f(x) = -2x$ è lineare mentre non lo è la funzione $g(x) = 2x + 1$.

**Esempio 5.0.29.** Siano $C^1[(a, b)], C^0[(a, b)]$ rispettivamente lo spazio vettoriale delle funzioni derivabili con derivata continua sull'intervallo $(a, b)$ e quello delle funzioni continue su $(a, b)$, muniti dell'usuale somma e prodotto per uno scalare. L'applicazione $D : C^1[(a, b)] \to C^0[(a, b)]$ definita da $D(f) = f'$, dove $f'$ è la derivata della funzione $f$, è una applicazione lineare. Questo è vero per via delle proprietà della derivata della somma di due funzioni e del prodotto di una costante per una funzione.

**Proposizione 5.0.30.** *Se $f : V \to W$ è lineare allora si ha:*

1. $f(0_V) = 0_W$;
2. $f(-v) = -f(v)$ *per ogni* $v \in V$.

*Dimostrazione.* Si ha

$$f(0_V) = f(0_V + 0_V) = f(0_V) + f(0_V) \ ,$$

per cui

$$f(0_V) = f(0_V) + f(0_V)$$

e quindi sommando l'opposto di $f(0_V)$ ad ambo i membri si ottiene $f(0_V) = 0_W$.

Sia ora $v \in V$. Per quanto appena dimostrato e per la linearità, abbiamo

$$0_W = f(0_V) = f(v + (-v)) = f(v) + f(-v)$$

e quindi $f(-v)$ è l'opposto di $f(v)$. □

**Proposizione 5.0.31.** *L'insieme $L(V, W)$, munito delle seguenti operazioni*

$$\begin{array}{rl} f, g \in L(V, W) & f \boxplus g : V \to W \ , (f \boxplus g)(u) = f(u) + g(u) \ , \\ g \in L(V, W) \ , \alpha \in \mathbb{R} & \alpha \boxdot g : V \to W \ , (\alpha \boxdot g)(u) = \alpha g(u) \ , \end{array}$$

*è uno spazio vettoriale.*

*Dimostrazione.* Dobbiamo innanzitutto dimostrare che, date $f, g \in L(V, W)$ e $\alpha \in \mathbb{R}$, le applicazioni $f \boxplus g$ e $\alpha \boxdot g$ sono lineari. Le proprietà si verificano tutte allo stesso modo, quindi ne dimostriamo una sola per $f \boxplus g$. Mostriamo che per ogni $u.v \in V$ si ha $(f \boxplus g)(u + v) = (f \boxplus g)(u) + (f \boxplus g)(v)$.

Infatti per come è definita $f \boxplus g$ e per la linearità di $f$ e $g$, si ha:

$$\begin{aligned} (f \boxplus g)(u + v) &= f(u + v) + g(u + v) = f(u) + f(v) + g(u) + g(v) \\ &= f(u) + g(u) + f(v) + g(v) = (f \boxplus g)(u) + (f \boxplus g)(v). \end{aligned}$$

Quindi le due operazioni sono ben definite. A questo punto occorre verificare che $(L(V, W), \boxplus, \boxdot)$ è uno spazio vettoriale. Questa verifica è lasciata come esercizio. Si osservi che il vettore nullo di $L(V, W)$ è l'applicazione $0 : V \to W$ che ad ogni vettore di $V$ associa il vettore nullo di $W$ □

## 5.1 Proprietà delle applicazioni lineari

Vediamo ora alcune proprietà delle applicazioni lineari che saranno utili in seguito. Diamo innanzitutto la seguente definizione:

**Definizione 5.1.1.** Data una applicazione lineare $f : V \to W$, è detto *nucleo* di $f$ l'insieme $\ker f = \{v \in V | f(v) = 0\}$ e *immagine* di $f$ l'insieme $\mathrm{Im} f = f(V)$.

**Proposizione 5.1.2.** *Una applicazione lineare $f : V \to W$ è suriettiva se e solo se* $\mathrm{Im} f = W$ *ed è iniettiva se e solo se* $\ker f = \{0\}$.

*Dimostrazione.* La $f$ è suriettiva se e solo se $f(V) = W$, cioè se e solo se $\mathrm{Im} f = W$.

Supponiamo ora che $f$ sia iniettiva e dimostriamo che $\ker f = \{0\}$. Sia $v \in \ker f$. Poichè si ha $f(v) = 0 = f(0)$, per l'iniettività di $f$ si ha $v = 0$ e quindi $\ker f = \{0\}$.

Viceversa, sia $\ker f = \{0\}$ e siano $u, v \in V$ tali che $f(u) = f(v)$. Per la linearità di $f$ si ha $f(u - v) = 0$ e dunque $u - v \in \ker f = \{0\}$ e quindi $u - v = 0$, cioè $u = v$ e allora $f$ è iniettiva. □

Vediamo ora come vengono trasformati i sottospazi tramite una applicazione lineare:

**Proposizione 5.1.3.** *Sia $f : V \to W$ una applicazione lineare. Allora si ha:*

1. *Se $X \subset V$ è un sottospazio vettoriale allora $f(X)$ è un sottospazio vettoriale di $W$.*

2. *Se $Y \subset W$ è un sottospazio vettoriale allora $f^{-1}(Y)$ è un sottospazio vettoriale di $V$.*

*In particolare per $X = V$ si ha che* $\mathrm{Im} f$ *è un sottospazio di $W$ e per $Y = \{0\}$ si ha che* $\ker f = f^{-1}(0)$ *è un sottospazio di $V$.*

*Dimostrazione.* 1. Siano $a, b \in \mathbb{R}$ e $u, v \in f(X)$. Allora esistono $x, y \in V$ tali che $u = f(x)$ e $v = f(y)$. Vogliamo dimostrare che la combinazione lineare $a \cdot u + b \cdot v$ appartiene a $f(X)$. Per la linearità di $f$ si ha

$$a \cdot u + b \cdot v = a \cdot f(x) + b \cdot f(y) = f(a \cdot x + b \cdot y)$$

e poiché $X$ è un sottospazio, $a \cdot x + b \cdot y \in X$, e quindi $a \cdot u + b \cdot v = f(a \cdot x + b \cdot y)$ appartiene a $f(X)$.

2. Siano $a, b \in \mathbb{R}$ e $u, v \in f^{-1}(Y)$. Allora $f(u), f(v) \in Y$. Poichè $Y$ è un sottospazio allora $a \cdot f(u) + b \cdot f(v) \in Y$. Per la linearità di $f$, si ha $a \cdot f(u) + b \cdot f(v) = f(a \cdot u + b \cdot v)$ e quindi $a \cdot u + b \cdot v \in f^{-1}(Y)$ e dunque $f^{-1}(Y)$ è un sottospazio.

$\square$

La proposizione seguente ci dice in particolare come trovare un sistema di generatori di $\mathrm{Im} f$:

**Proposizione 5.1.4.** *Sia $f : V \to W$ una applicazione lineare. Allora si ha:*

1. *se $\{v_1, \cdots, v_s\}$ è un sistema di generatori per $V$, $\{f(v_1), \cdots, f(v_s)\}$ è un sistema di generatori per* $\mathrm{Im} f$*;*

2. *se $f$ è iniettiva e i vettori $\{v_1, \cdots, v_s\}$ sono linearmente indipendenti, allora i vettori $\{f(v_1), \cdots, f(v_s)\}$ sono linearmente indipendenti.*

*Dimostrazione.* 1. sia $u \in \mathrm{Im} f$. Allora esiste $x \in V$ tale che $f(x) = u$. Poiché $V$ è generato da $\{v_1, \cdots, v_s\}$ abbiamo $x = \sum_i x_i \cdot v_i$, $x_i \in \mathbb{R}$. Per la linearità di $f$ si ha

$$u = f(x) = f(\sum_i x_i \cdot v_i) = \sum_i x_i \cdot f(v_i)$$

e quindi $\mathrm{Im} V$ è generato da $\{f(v_1), \cdots, f(v_s)\}$.

2. Consideriamo una combinazione lineare di $\{f(v_1), \cdots, f(v_s)\}$ che dia il vettore nullo:
$$\sum_i a_i \cdot f(v_i) = 0.$$
Per la linearità di $f$ si ha
$$f(\sum_i a_i \cdot v_i) = 0$$
e quindi $\sum_i a_i \cdot v_i$ appartiene a $\ker f$. Essendo $f$ iniettiva si ha $\ker f = \{0\}$ e quindi si ottiene
$$\sum_i a_i \cdot v_i = 0$$
e dunque gli $a_i$ sono nulli perchè per ipotesi i vettori $\{v_1, \cdots, v_s\}$ sono linearmente indipendenti. Questo implica che i vettori $\{f(v_1), \cdots, f(v_s)\}$ sono linearmente indipendenti.

□

Come conseguenza immediata si ha il seguente

**Corollario 5.1.5.** *Sia $f : V^n \to W^m$ una applicazione lineare iniettiva. Se $\{v_1, \cdots, v_n\}$ è una base di $V$ allora $\{f(v_1), \cdots, f(v_n)\}$ è una base di* Im$f$.

*Dimostrazione.* Per la proposizione precedente, i vettori $\{f(v_1), \cdots, f(v_n)\}$ sono linearmente indipendenti e generano Im$f$ e quindi sono una base di Im$f$. □

Si osservi che data una applicazione lineare $f : V^n \to W^m$, se $\{e_1, \cdots, e_n\}$ è una base di $V$ e $v \in V$, per calcolare $f(v)$ è sufficiente conoscere le componenti di $v$ rispetto alla base. Infatti, posto $v = \sum_{i=1}^n v_i \cdot e_i$, per la linearità di $f$ si ha $f(v) = f(\sum_{i=1}^n v_i \cdot e_i) = \sum_{i=1}^n v_i \cdot f(e_i)$.

Questo suggerisce il fatto che l'applicazione sia completamente determinata da $\{f(e_1), \cdots, f(e_n)\}$. Il seguente risultato esprime quanto detto.

**Teorema 5.1.6** (Teorema fondamentale delle applicazioni lineari)**.** *Sia $B = \{e_1, \cdots, e_n\}$ una base di uno spazio vettoriale $V$. Dato un spazio vettoriale $W$ e una applicazione $F : B \to W$, esiste un'unica applicazione lineare $f : V \to W$ tale che la sua restrizione a $B$ coincida con $F$ ($f_{|B} = F$).*

*Dimostrazione.* Definiamo $f$ come segue
$$f(v) := \sum_{i=1}^n v_i \cdot F(e_i) \ ,$$
dove i $v_i$ sono le componenti di $v$ rispetto a $B$. Dalla definizione risulta evidente che $f(e_i) = F(e_i)$.

Verifichiamo che $f$ è lineare. Siano $a, b \in \mathbb{R}$, $u = \sum_{i=1}^{n} u_i \cdot e_i$ e $v = \sum_{i=1}^{n} v_i \cdot e_i$. Abbiamo

$$\begin{aligned} f(a \cdot u + b \cdot v) &= f(a \cdot \sum_{i=1}^{n} u_i \cdot e_i + b \cdot \sum_{i=1}^{n} v_i \cdot e_i) = f(\sum_{i=1}^{n} (au_i + bv_i) \cdot e_i) \\ &\overset{def}{=} \sum_{i=1}^{n} (au_i + bv_i) \cdot F(e_i) = a \cdot \sum_{i=1}^{n} u_i \cdot F(e_i) + b \cdot \sum_{i=1}^{n} v_i \cdot F(e_i) \\ &= a \cdot f(u) + b \cdot f(v). \end{aligned}$$

Ora dimostriamo l'unicità di $f$. Per questo supponiamo che esista $g$ lineare che coincide con $F$ su $B$ e mostriamo che $f, g$ coincidono su ogni vettore. Infatti dato $v = \sum_{i=1}^{n} v_i \cdot e_i$ si ha

$$g(v) = g(\sum_{i=1}^{n} v_i \cdot e_i) \overset{\text{linearità di } g}{=} \sum_{i=1}^{n} v_i \cdot g(e_i) \overset{g(e_i)=F(e_i)}{=} \sum_{i=1}^{n} v_i \cdot F(e_i) = f(v).$$

□

Il prossimo teorema fornisce una utile relazione tra la dimensione del nucleo, dell'immagine e del dominio di una applicazione lineare:

**Teorema 5.1.7** (Equazione dimensionale per le applicazioni lineari)**.** *Sia* $f : V \to W$ *una applicazione lineare e supponiamo che* $\dim V = n$*. Allora si ha*

$$\dim V = \dim \ker f + \dim \mathrm{Im} f \ .$$

*Dimostrazione.* Essendo $V$ di dimensione finita, lo è anche il nucleo di $f$ in quanto sottospazio di $V$. Sia quindi $h$ (eventualmente nullo) la dimensione di $\ker f$ e $\{e_1, \cdots, e_h\}$ una sua base. Completiamo la base del nucleo per ottenere una base $B = \{e_1, \cdots, e_h, g_1, \cdots, g_{n-h}\}$ di $V$. Vogliamo dimostrare che $B' = \{f(g_1), \cdots, f(g_{n-h})\}$ è una base di $\mathrm{Im} f$. Fatto ciò è chiaro che si ha la tesi del teorema. Per la Proposizione 5.1.4, $B' \cup \{0\}$ è un sistema di generatori in quanto $B' \cup \{0\} = f(B)$ e $B$ è un sistema di generatori per $V$. Inoltre $0 \in L(B')$ e quindi anche $B'$ è un sistema di generatori. Basta quindi dimostrare che l'insieme $B'$ è linearmente indipendente. Siano quindi $a_1, \cdots, a_{n-h} \in \mathbb{R}$ tali che

$$0 = \sum_{i=1}^{n-h} a_i \cdot f(g_i) = f(\sum_{i=1}^{n-h} a_i \cdot g_i).$$

Questo implica che $\sum_{i=1}^{n-h} a_i \cdot g_i \in \ker f$ e quindi gli $a_i$ sono tutti nulli per la lineare indipendenza di $B$. □

**Definizione 5.1.8.** La dimensione di $\mathrm{Im}f$ è detta rango di $f$.

Vediamo alcuni corollari importanti del Teorema 5.1.7:

**Corollario 5.1.9.** *Sia $f : V \to W$ una applicazione lineare e supponiamo che $\dim V = n$, allora si ha:*

1. *$f$ è iniettiva se e solo $\dim \ker f = 0$ ovvero se e solo $\dim \mathrm{Im}f = n$;*
2. *$f$ è suriettiva se e solo se $\dim \mathrm{Im}f = \dim W$;*
3. *$f$ è un isomorfismo se e solo se $\dim \mathrm{Im}f = \dim W = \dim V$;*
4. *se $\dim V = \dim W$ allora $f$ è un isomorfismo se e solo se è iniettiva.*

*Dimostrazione.* La dimostrazione è immediata utilizzando l'equazione dimensionale ed è quindi lasciata per esercizio. □

**Corollario 5.1.10.** *Due spazi vettoriali $V, W$ di dimensione finita sono isomorfi se e solo se hanno la stessa dimensione.*

*Dimostrazione.* Per il Corollario 5.1.9, se $V$ e $W$ sono isomorfi allora hanno la stessa dimensione. Viceversa, se hanno la stessa dimensione $n$, siano $B = \{e_1, \cdots, e_n\}$, $B' = \{f_1, \cdots, f_n\}$ basi di $V$ e $W$ rispettivamente. Per il Teorema 5.1.6 esiste un'unica applicazione lineare $g : V \to W$ tale che $g(e_i) = f_i$. La $f$ così costruita è iniettiva in quanto iniettiva su $B$ (verificare!) e quindi un isomorfismo per il Corollario 5.1.9. □

## 5.2 Matrice associata ad una applicazione lineare

Data una applicazione lineare $f : V \to W$ tra due spazi vettoriali di dimensione finita $n$ ed $m$ rispettivamente, scelta una base per $V$ e una per $W$, è possibile associare una matrice ad $f$ come spiegato di seguito.

Siano $B = (e_1, \cdots, e_n)$ e $B' = (b_1, \cdots, b_m)$ basi ordinate di $V$ e $W$ rispettivamente. Il vettore $f(e_j) \in W$ è combinazione lineare degli elementi della base $B'$ e quindi esistono e sono uniche le componenti di $f(e_j)$ rispetto a $B'$. Indichiamo con $(a_{1j}, a_{2j}, \cdots, a_{mj})$ tali componenti. Si noti che l'indice $j$ nelle componenti sta ad indicare che si tratta delle componenti del vettore $f(e_j)$.

Si ottengono in questo modo le entrate di una matrice $A = (a_{ij})$. La $j$-esima colonna di $A$ è formata dalle componenti di $f(e_j)$ rispetto alla base ordinata $B'$.

Possiamo schematizzare la costruzione della matrice nel modo seguente:

$$A_{B,B'} = \begin{array}{c} \begin{array}{ccccc} f(e_1) & \cdots & f(e_j) & \cdots & f(e_n) \end{array} \\ \left\| \begin{array}{ccccc} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{array} \right\| \begin{array}{c} b_1 \\ b_2 \\ \vdots \\ b_m \end{array} \end{array} \tag{5.3}$$

**Definizione 5.2.1.** La matrice (5.3) è detta matrice di $f$ rispetto alle basi $B$ e $B'$.

La matrice di $f$ rispetto alle basi $B$ e $B'$ permette di trovare le componenti di un qualsiasi vettore $f(v)$ note quelle di $v$:

**Proposizione 5.2.2.** *Sia $f : V \to W$ una applicazione lineare e siano $B = (e_1, \cdots, e_n)$ e $B' = (b_1, \cdots, b_m)$ basi ordinate di $V$ e $W$ rispettivamente. Siano $v = \sum_{j=1}^{n} v_j{\cdot}e_j$ e $X$ la matrice colonna delle componenti di $v$ rispetto a $B$. Sia $A$ la matrice di $f$ rispetto alle basi $B, B'$ e sia $Y$ la matrice colonna delle componenti di $f(v)$ rispetto alla base $B'$. Allora si ha:*

$$Y = AX \,. \tag{5.4}$$

*Dimostrazione.* Sia $A = (a_{ij})$. Si ha quindi $f(e_j) = \sum_{i=1}^{m} a_{ij}{\cdot}b_i$. Sostituendo questa espressione in quella del vettore $f(v)$ e sfruttando la linearità di $f$, si ottiene

$$f(v) = \sum_{j=1}^{n} v_j{\cdot}f(e_j) = \sum_{j=1}^{n} v_j{\cdot}(\sum_{i=1}^{m} a_{ij}{\cdot}b_i) = \sum_{i=1}^{m}(\sum_{j=1}^{n} a_{ij}v_j)b_i \,,$$

da cui si ricava che la $i$-esima componente di $f(v)$ rispetto a $B'$ vale $\sum_{j=1}^{n} a_{ij}v_j$ e questo equivale alla (5.4). □

*Osservazione* 5.2.3. Quando $f \in End(V)$, di solito si considera la stessa base nel dominio e nel codominio.

**Esempio 5.2.4.** Consideriamo l'applicazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^2$ data da $f(x, y, z) = (x + y - z, 2x + z)$. Siano $B = (e_1, e_2, e_3)$, $B' = (b_1, b_2)$ le basi canoniche di $\mathbb{R}^3$ e $\mathbb{R}^2$ rispettivamente e troviamo la matrice $A$ di $f$ rispetto a tali basi.

Abbiamo $f(e_1) = (1, 2) = b_1 + 2b_2$ e quindi nella prima colonna di $A$ dobbiamo mettere tali componenti. Facciamo lo stesso per gli altri vettori di $B$ e otteniamo $f(e_2) = (1, 0) = b_1, f(e_3) = (-1, 1) = -b_1 + b_2$. Quindi la matrice di $f$ rispetto a $B, B'$ sarà

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

Consideriamo ora la base $B'' = (c_1 = b_1 + b_2, c_2 = b_1 - b_2)$ di $\mathbb{R}^2$ e calcoliamo la matrice di $f$ rispetto a $B, B''$. Dobbiamo calcolare le componenti di $(f(e_1), f(e_2), f(e_3))$ rispetto a $B''$. Abbiamo quindi $f(e_1) = (1,2) = 3/2c_1 - 1/2c_2$, $f(e_2) = (1,0) = 1/2c_1 + 1/2c_2$, $f(e_3) = (-1,1) = -c_2$ e quindi la matrice di $f$ rispetto a $B, B''$ sarà

$$\begin{Vmatrix} 3/2 & 1/2 & 0 \\ -1/2 & 1/2 & -1 \end{Vmatrix}.$$

L'esempio precedente mostra che se si cambiano le basi, la matrice associata alla stessa applicazione lineare cambia. Vedremo più avanti in che modo avviene questo cambiamento.

Abbiamo visto che l'insieme delle applicazioni lineari tra due spazi vettoriali $V, W$ è anch'esso uno spazio vettoriale. Il seguente risultato stabilisce un isomorfismo tra $L(V^n, W^m)$ e $\mathbf{M}_{m\times n}(\mathbb{R})$:

**Teorema 5.2.5.** *Siano $V, W$ due spazi vettoriali di dimensione finita. Siano inoltre $B = (e_1, \cdots, e_n)$ e $B' = (h_1, \cdots, h_m)$ basi ordinate di $V$ e $W$ rispettivamente. Allora l'applicazione $F_{B,B'} : L(V, W) \to \mathbf{M}_{m\times n}(\mathbb{R})$ che ad una applicazione lineare associa la sua matrice rispetto a $B, B'$, è un isomorfismo tra spazi vettoriali. In particolare si ha* $\dim L(V, W) = \dim \mathbf{M}_{m\times n}(\mathbb{R}) = mn$.

*Dimostrazione.* Siano $f, g \in L(V, W)$ e $\lambda \in \mathbb{R}$. Sia $A = (a_{ij})$ la matrice associata a $f$, $B = (b_{ij})$ quella associata a $g$ e $C = (c_{ij})$ quella associata a $f \boxplus g$. Dobbiamo dimostrare che $C = A + B$. Le entrate di $C$ sono definite come segue

$$(f \boxplus g)(e_j) = \sum_{i=1}^{m} c_{ij}\cdot h_i.$$

Sfruttiamo ora la definizione di somma di due applicazioni lineari

$$\sum_{i=1}^{m} c_{ij}\cdot h_i = (f \boxplus g)(e_j) \overset{\text{def di } \boxplus}{=} f(e_j) + g(e_j) = \sum_{i=1}^{m} a_{ij}\cdot h_i + \sum_{i=1}^{m} b_{ij}\cdot h_i = \sum_{i=1}^{m} (a_{ij} + b_{ij})\cdot h_i,$$

da cui si ricava $c_{ij} = a_{ij} + b_{ij}$ che equivale a $C = A + B$.

Ora dimostriamo che la matrice di $\lambda \boxdot f$ è $\lambda A$. Come prima si ha

$$\lambda \boxdot f(e_j) = \lambda\cdot f(e_j) = \lambda\cdot \sum_{i=1}^{m} a_{ij}\cdot h_i = \sum_{i=1}^{m} (\lambda a_{ij})\cdot h_i$$

e quindi la matrice di $\lambda \boxdot f$ è $\lambda A$. Questo mostra la linearità di $F_{B,B'}$.

Dimostriamo che $F_{B,B'}$ è iniettiva. Per far questo basta dimostrare che il suo nucleo è formato dalla applicazione nulla. Sia dunque $f \in \ker F_{B,B'}$. Allora la sua matrice associata è la matrice nulla. Questo significa che $f(e_j) = 0$ per ogni $j$ e quindi, per il Teorema 5.1.6 si deve avere $f = 0$.

Mostriamo la suriettività di $F_{B,B'}$. Data $A = (a_{ij}) \in \mathbf{M}_{m\times n}(\mathbb{R})$ basta definire $f$ come l'unica applicazione lineare tale che per ogni $j$ si abbia $f(e_j) = \sum_{i=1}^{m} a_{ij} \cdot h_i$.

Quindi $F_{B,B'}$ è un isomorfismo e si ha $\dim L(V, W) = \dim \mathbf{M}_{m\times n}(\mathbb{R}) = mn$. □

Il teorema seguente descrive in modo preciso come cambia la matrice associata ad una applicazione lineare se si cambiano le basi:

**Teorema 5.2.6.** *Sia $f : V \to W$ una applicazione lineare e siano $B, B'$ basi ordinate di $V$ e $C, C'$ basi ordinate di $W$. Sia $A_{B,C}$ la matrice di $f$ rispetto alle basi $B, C$ e $A_{B',C'}$ la matrice di $f$ rispetto a $B', C'$. Siano $X, X'$ le matrici delle componenti di un vettore $v$ rispetto a $B$ e $B'$ e $Y, Y'$ le matrici delle componenti di $f(v)$ rispetto a $C$ e $C'$ rispettivamente. Inoltre sia $P$ la matrice di passaggio dalla base $B$ alla base $B'$ e $Q$ la matrice di passaggio dalla base $C$ alla base $C'$. Allora si ha*

$$A_{B',C'} = Q\, A_{B,C}\, P^{-1}.$$

*Dimostrazione.* Per ipotesi si ha:

$$Y' = QY\ , X' = PX\ , Y = A_{B,C}X\ , Y' = A_{B',C'}X'\ ,$$

da cui, ricavando $Y$ e $X$, si ottiene

$$Q^{-1}Y' = A_{B,C}P^{-1}X'$$

e quindi ancora

$$Y' = QA_{B,C}P^{-1}X'\ .$$

Per l'unicità della matrice associata rispetto a due basi si deve avere

$$A_{B',C'} = Q\, A_{B,C}\, P^{-1}.$$

□

Nel caso particolare di un endomorfismo, si avrebbe $B = C$ e $B' = C'$, e quindi $A_{B',B'} = P\, A_{B,B}\, P^{-1}$. In questo caso le due matrici $A_{B',C'}$ e $A_{B,C}$ sono simili.

Si può quindi definire la traccia e il determinante di $f$ come la traccia e il determinante di una sua qualunque matrice associata. Infatti per i risultati sulle matrici visti precedentemente, due matrici simili hanno stessa traccia e stesso determinante.

Il prossimo Teorema ci permette di calcolare la matrice associata ad una composizione di applicazioni lineari.

**Teorema 5.2.7.** *Siano $f : U \to V$ e $g : V \to W$ applicazioni lineari tra spazi vettoriali finitamente generati e $B, B', B''$ basi ordinate di $U, V, W$ rispettivamente. Sia $A$ la matrice di $f$ rispetto a $B, B'$ e $C$ la matrice di $g$ rispetto a $B', B''$. Allora la matrice di $g \circ f : U \to W$ rispetto a $B, B''$ è $CA$.*

*Dimostrazione.* Siano $B = (e_1, \cdots, e_n)$, $B' = (b_1, \cdots, b_m)$ e $B'' = (h_1, \cdots, h_s)$ le basi di $U, V, W$ rispettivamente. Si ha

$$f(e_j) = \sum_{i=1}^{m} a_{ij} \cdot b_i \ , \ g(b_i) = \sum_{k=1}^{s} c_{ki} \cdot h_k.$$

Sostituendo si ottiene

$$\begin{aligned}(g \circ f)(e_j) &= g(f(e_j)) = \sum_{i=1}^{m} a_{ij} \cdot g(b_i) = \sum_{i=1}^{m} a_{ij} \cdot (\sum_{k=1}^{s} c_{ki} \cdot h_k) \\ &= \sum_{k=1}^{s} (\sum_{i=1}^{m} c_{ki} a_{ij}) \cdot h_k\end{aligned}$$

e quindi la matrice di $g \circ f$ rispetto a $B, B''$ è proprio la matrice $CA$. □

**Corollario 5.2.8.** *Se $f : U \to V$ è un isomorfismo, $B, B'$ basi ordinate di $U, V$ rispettivamente e $A$ la matrice di $f$ rispetto a $B, B'$. Allora la matrice di $f^{-1}$ rispetto a $B', B$ è $A^{-1}$.*

*Dimostrazione.* Si applichi il teorema precedente a $f \circ f^{-1} = id_V$ e si osservi che l'applicazione identità ha come matrice l'identità. □

*Osservazione* 5.2.9. Una matrice è invertibile se e solo se è la matrice associata a un isomorfismo. Infatti per il corollario precedente se è associata ad un isomorfismo è invertibile. Viceversa se è invertibile e ha ordine $n$, la si può pensare come associata all'applicazione lineare da $\mathbf{M}_{n\times 1}(\mathbb{R})$ in se, definita nel modo seguente:

$$f_A(X) = AX \ , X \in \mathbf{M}_{n\times 1}(\mathbb{R}).$$

Si verifica immediatamente che $f_A$ è lineare, invertibile e ha come matrice associata rispetto alla base canonica la matrice $A$.

### 5.2.1 Sistemi lineari e teorema di Rouché - Capelli

Utilizzando la teoria degli spazi vettoriali dimostreremo il teorema di Rouché-Capelli. Consideriamo un sistema lineare con $m$ equzioni ed $n$ incognite

$$AX = B,$$

dove $A \in \mathbf{M}_{m\times n}(\mathbb{R})$ è la matrice dei coefficienti del sistema, $B \in \mathbf{M}_{m\times 1}(\mathbb{R})$ è la matrice dei termini noti e $X =^t (x_1 \cdots x_n)$ è la matrice delle incognite. Indichiamo con $(A|B)$ la matrice completa. Indichiamo con $A^i$ la $i$-esima colonna di $A$.

**Teorema 5.2.10** (di Rouché - Capelli)**.** *Il sistema lineare*

$$AX = B \tag{5.5}$$

*è compatibile se e solo se* $\mathrm{rg}A = \mathrm{rg}(A|B)$. *Se è compatibile, il numero di soluzioni è* $\infty^{n-\mathrm{rg}A}$.

*Dimostrazione.* Il sistema è compatibile se e solo se esiste $X_0 =^t (x_1^0 \cdots x_n^0)$ tale che

$$x_1^0 A_1 + \cdots + x_n^0 A_n = B.$$

Questo è vero se e solo se $L(A^1, \cdots, A^n)$ coincide con $L(A^1, \cdots, A^n, B)$, poichè $L(A^1, \cdots, A^n)$ è un sottospazio di $L(A^1, \cdots, A^n, B)$. Questo equivale a dire che $\dim L(A^1, \cdots, A^n) = \dim L(A^1, \cdots, A^n, B)$ e quindi il sistema è compatibile se e solo se $\mathrm{rg}A = \mathrm{rg}(A|B)$.

Supponiamo ora che il sistema sia compatibile. Allora esiste una soluzione $X_0$. Ogni altra soluzione $Y$ di (5.5) è della forma $X + X_0$, dove stavolta $X$ è soluzione del sistema omogeneo associato

$$AX = O. \tag{5.6}$$

L'insieme delle soluzioni di (5.6) forma uno spazio vettoriale che può essere visto come il nucleo della applicazione lineare $f_A : \mathbf{M}_{n\times 1}(\mathbb{R}) \to \mathbf{M}_{m\times 1}(\mathbb{R})$

$$f_A(X) = AX,$$

che ha come matrice associata rispetto alle basi canoniche proprio la matrice $A$. La dimensione di tale nucleo ci è fornita dall'equazione dimensionale relativa a $f_A$ e quindi si ottiene $\dim\ker f_A = \dim \mathbf{M}_{n\times 1} - \dim \mathrm{Im} f_A = n - \mathrm{rg}A$. Questo significa che date $(n - \mathrm{rg}A)$ soluzioni di (5.6) linearmente indipendenti, ogni soluzione di (5.5) è data da una combinazione lineare di esse sommata alla soluzione particolare $X_0$. In questo senso ci sono $\infty^{n-\mathrm{rg}A}$ soluzioni.

□

## 5.3 Spazio duale e biduale di uno spazio vettoriale

Abbiamo visto precedentemente che dati due spazi vettoriali $V, W$ di dimensione $n, m$ rispettivamente, lo spazio vettoriale $L(V, W)$ ha dimensione pari a $mn$. Nel caso particolare in cui $W = \mathbb{R}$ indicheremo tale spazio con $V^*$ e lo chiameremo *spazio duale* di $V$. Chiameremo *biduale* di $V$ il duale di $V^*$ e lo indicheremo con $V^{**}$. Per quanto detto sopra si ha

$$\dim V = \dim V^* = \dim V^{**}$$

e quindi i tre spazi sono isomorfi.

Ora costruiremo degli isomorfismi espliciti tra tali spazi. Sia $B = (e_1, \cdots, e_n)$ una base ordinata di $V$ e per $j$ fissato $j = 1, \cdots, n$, definiamo il seguente elemento di $V^*$ (che è una applicazione lineare da $V$ in $\mathbb{R}$)

$$e^j : V \to \mathbb{R} \ \ , e^j(e_i) = \delta_{ij}.$$

Per il teorema fondamentale sulle applicazioni lineari, l'applicazione lineare $e^j$ è completamente determinata dal suo valore sulla base $B$ di V. Si ottengono in questo modo $n$ elementi di $V^*$, $B^* = (e^1, \cdots, e^n)$. Vogliamo dimostrare che $B^*$ è una base di $V^*$ e per far questo è sufficiente dimostrare che i vettori di $B^*$ sono linearmente indipendenti, visto che $\dim V^* = n$.

Siano quindi $a_1, \cdots, a_n \in \mathbb{R}$ tali che $f = \sum_{i=1}^n a_i e^i = 0$, dove 0 è l'applicazione nulla da $V$ in $\mathbb{R}$. Se $f$ è l'applicazione nulla, allora in particolare vale zero sui vettori di $B$. Quindi si ottiene per ogni $j$:

$$0 = f(e_j) = (\sum_{i=1}^n a_i e^i)(e_j) = \sum_{i=1}^n a_i[e^i(e_j)] = \sum_{i=1}^n a_i \delta_{ij} = a_j.$$

Quindi tutti gli $a_j$ devono essere nulli e dunque $B^*$ è una base di $V^*$.

La base $B^*$ è detta *base duale* di $B$. A questo punto è facile costruire un isomorfismo tra $V$ e $V^*$. Infatti basta definire $F : V \to V^*$ come l'unica applicazione lineare tale che $F(e_i) = e^i$, per $i = 1, \cdots, n$.

Bisogna osservare che per costruire questo isomorfismo è necessario utilizzare una base di $V$ e quindi in questo senso si dice che tale isomorfismo non è *canonico*, ma dipende dalla base scelta. Ora vedremo invece come sia possibile costruire un isomorfismo canonico tra $V$ e il suo biduale.

Innanzitutto, dato un vettore $v \in V$, sia $f_v : V^* \to \mathbb{R}$ definita da

$$f_v(\omega) = \omega(v).$$

Quanto fatto ha senso poichè $\omega$ è un elemento di $V^*$, cioè una forma lineare, e quindi si può applicare ai vettori di $V$. Inoltre si verifica facilmente che

$$f_v(\omega \boxplus \eta) = f_v(\omega) + f_v(\eta) \ , \ f_v(a \cdot \omega) = a f_v(\omega)$$

e quindi $f_v$ è lineare, cioè appartiene a $V^{**}$.

Sia ora $F : V \to V^{**}$ definita come segue:

$$F(v) = f_v.$$

Dobbiamo dimostrare che $F$ è lineare. Abbiamo $F(u + v) = f_{u+v}$. Calcoliamo $f_{u+v}$ su un vettore generico $\omega \in V^*$:

$$f_{u+v}(\omega) = \omega(u + v) = \omega(u) + \omega(v) = f_u(\omega) + f_v(\omega) = (f_u \boxplus f_v)(\omega)$$

e quindi si ha $f_{u+v} = f_u \boxplus f_v$, cioè $F(u+v) = F(u)+F(v)$. Analogamente si dimostra che $F(a\cdot u) = a\cdot F(u)$ e quindi $F$ è lineare.

Per dimostrare che $F$ è un isomorfismo è sufficiente dimostrare che è iniettiva, considerato che $\dim V = \dim V^{**}$. Poiché $F$ è lineare, basta dimostrare che il suo nucleo è ridotto al vettore nullo.

Sia quindi $v_1 \in \ker F$. Si ha allora $F(v_1) = f_{v_1} = 0$ e quindi l'applicazione $f_{v_1}$ vale zero su ogni vettore. Se per assurdo $v_1$ non fosse nullo, potremmo completarlo per ottenere una base $(v_1, v_2, \cdots, v_n)$ di $V$. Consideriamo allora la sua base duale $(v^1, \cdots, v^n)$. Ricordiamo che i $v^i$ appartengono a $V^*$ e applichiamo ambo i membri di $f_{v_1} = 0$ a $v^1$. Si ha quindi

$$0 = f_{v_1}(v^1) \overset{\text{definizione di } f_{v_1}}{=} v^1(v_1) = 1$$

e questo è assurdo, quindi si deve avere $v_1 = 0$.

Vedremo più avanti come sia possibile definire un isomorfismo canonico tra $V$ e $V^*$ quando lo spazio vettoriale $V$ è munito di un prodotto scalare.

**Esempio 5.3.1.** Sia $V = \mathbb{R}^3$ munito della base canonica $B = (e_1, e_2, e_3)$. Definiamo la base duale di $B$. Per esempio si ha che $e^1 : \mathbb{R}^3 \to \mathbb{R}$ è definita da $e^1(e_j) = \delta_{1j}$ e quindi

$$e^1(e_1) = 1\ , e^1(e_2) = e^1(e_3) = 0.$$

Per calcolare $e^1$ su un vettore qualunque di $\mathbb{R}^3$ ci occorrono le componenti del vettore rispetto a $B$. Sia dunque $v = (x, y, z)$, cioè $v = x\cdot e_1 + y\cdot e_2 + z\cdot e_3$, dato che $B$ è la base canonica. Per la linearità di $e^1$ otteniamo

$$e^1(v) = xe^1(e_1) + ye^1(e_2) + ze^1(e_3) = x.$$

Allo stesso modo si ottiene $e^2(x, y, z) = y$ e $e^3(x, y, z) = z$.

Sia ora $f$ una forma lineare su $V$, cioè $f \in V^*$. Allora si avrà $f(x, y, z) = ax + by + cz$. Poiché $f \in V^*$, allora è combinazione lineare di $(e^1, e^2, e^3)$ e si avrà $f = \alpha\cdot e^1 \boxplus \beta\cdot e^2 \boxplus \gamma\cdot e^3$. Per calcolare i coefficienti della combinazione lineare basterà applicare $f$ ai vettori di $B$. Applicando $f$ a $e_1$ si ha

$$f(e_1) = (\alpha\cdot e^1 \boxplus \beta\cdot e^2 \boxplus \gamma\cdot e^3)(e_1) = \alpha.$$

Poichè conosciamo l'espressione di $f$ e di $e_1 = (1, 0, 0)$, otteniamo $f(e_1) = a$ e quindi $\alpha = a$. Continuando allo stesso modo si ottiene $\beta = b$ e $\gamma = c$. Ovviamente se si cambia base di $\mathbb{R}^3$ e quindi anche base duale, le componenti rispetto alla nuova base duale non saranno più $a, b, c$.

# Capitolo 6

# Diagonalizzazione di un endomorfismo

Consideriamo uno spazio vettoriale $V$ di dimensione $n$ e $f \in End(V)$. Precedentemente abbiamo osservato che la matrice associata ad $f$ rispetto ad una base cambia in modo ben preciso se si cambia la base. Detta $A$ la matrice di $f$ rispetto ad una base $B$ e $P$ la matrice di passaggio dalla base $B$ ad una base $B'$, la matrice di $f$ rispetto a $B'$ sarà $PAP^{-1}$. È quindi naturale porsi il problema di trovare una base tale che la matrice rispetto ad essa sia la più semplice possibile, cioè una matrice diagonale.

Se $B = (e_1, \cdots, e_n)$ è una base rispetto alla quale $f$ ha matrice diagonale le cui entrate nella diagonale principale sono date dagli scalari $a_1, \cdots, a_n$ (non necessariamente tutti distinti), allora si ha

$$f(e_i) = a_i \cdot e_i.$$

Per poter cercare una base siffatta, occorre quindi trovare dei vettori non nulli $v$ tali che $f(v) = \lambda \cdot v$ per un certo $\lambda \in \mathbb{R}$. Questo giustifica la seguente definizione:

**Definizione 6.0.2.** Si dice che $\lambda$ è un *autovalore* di $f \in End(V)$ se esiste un vettore non nullo $v$ tale che $f(v) = \lambda \cdot v$. Il vettore $v$ sarà detto *autovettore* relativo a $\lambda$.

*Osservazione* 6.0.3. Gli autovettori relativi all'autovalore nullo non sono altro che i vettori non nulli di $\ker f$ (se esistono).

Vediamo alcuni esempi:

**Esempio 6.0.4.** Consideriamo l'insieme $V_3$ dei vettori liberi dello spazio e sia $u$ un vettore non nullo di modulo 1. Sia $f : V_3 \to V_3$ l'applicazione che a ogni vettore $v$ associa il vettore proiezione ortogonale su $u$, cioè $f(v) = < v, u > \cdot u$. La linearità di $f$ discende dalle proprietà del prodotto scalare.

Si osservi che $f(u) = u$ e quindi $u$ è un autovettore relativo all'autovalore 1. Se $v$ è un vettore non nullo perpendicolare ad $u$, allora si ha $f(v) = 0 = 0 \cdot v$ e quindi $v$ è un autovettore relativo a 0.

**Definizione 6.0.5.** Si chiama autospazio relativo all'autovalore $\lambda$ l'insieme

$$V(\lambda) = \{v \in V | f(v) = \lambda \cdot v\} \cup \{0\}.$$

In effetti ogni autospazio risulta essere un sottospazio vettoriale di $V$ in quanto si verifica facilmente che

$$V(\lambda) = \ker(f - \lambda \cdot Id) ,$$

dove $Id$ è l'applicazione identità di $V$. Vediamo ora alcune proprietà degli autospazi:

**Proposizione 6.0.6.** *Siano $\lambda_1, \cdots, \lambda_k$ autovalori distinti di $f$ e $v_1, \cdots, v_k$ autovettori corrispondenti. Allora i vettori $v_1, \cdots, v_k$ sono linearmente indipendenti.*

*Dimostrazione.* Dimostriamo la proprietà per induzione sul numero $k$ di autovalori.

Per $k = 1$ la proprietà è vera perché $v_1$ è un autovettore e quindi non può essere nullo, il che significa che è linearmente indipendente.

Supponiamo vera la proprietà per $k - 1$ autovalori distinti e consideriamo $k$ autovalori distinti $\lambda_1, \cdots, \lambda_k$ con corrispondenti autovettori $v_1, \cdots, v_k$. Per ipotesi induttiva, gli autovettori $v_1, \cdots, v_{k-1}$ sono linearmente indipendenti. Se per assurdo $v_1, \cdots, v_k$ fossero dipendenti allora si avrebbe

$$v_k = \sum_{i=1}^{k-1} a_i \cdot v_i,$$

per certi scalari $a_i$. Applicando $f$ ad ambo i membri, si ottiene

$$\lambda_k \cdot v_k = f(v_k) = \sum_{i=1}^{k-1} a_i \cdot f(v_i) = \sum_{i=1}^{k-1} (a_i \lambda_i) \cdot v_i$$

e sostituendo l'espressione di $v_k = \sum_{i=1}^{k-1} a_i \cdot v_i$

$$\lambda_k (\sum_{i=1}^{k-1} a_i \cdot v_i) = \sum_{i=1}^{k-1} (a_i \lambda_i) \cdot v_i.$$

Portando tutto al primo membro si ha

$$\sum_{i=1}^{k-1} a_i (\lambda_k - \lambda_i) \cdot v_i = 0,$$

da cui si ottiene $a_i(\lambda_k - \lambda_i) = 0$ per ogni $i$ poiché per ipotesi induttiva $v_1, \cdots, v_{k-1}$ sono linearmente indipendenti. Ma questo implica che $a_i = 0$ per ogni $i$ in quanto gli autovalori sono tutti distinti. □

**Proposizione 6.0.7.** *Siano $\lambda_1, \cdots, \lambda_k$ autovalori distinti di $f$ e $V(\lambda_1), \cdots, V(\lambda_k)$ gli autospazi corrispondenti. Allora la somma $+_{i=1}^{k} V(\lambda_i)$ è diretta.*

*Dimostrazione.* Dobbiamo dimostrare che l'intersezione di ogni $V(\lambda_j)$ con la somma dei restanti autospazi (che indicheremocon $+_{i\neq j}V(\lambda_i)$) contiene solo il vettore nullo. Sia allora $v \in V(\lambda_j) \cap +_{i\neq j}V(\lambda_i)$. Allora si ha

$$v = v_1 + \cdots + v_{j-1} + v_{j+1} + \cdots + v_k \ ,$$

dove $v_i \in V(\lambda_i)$. Inoltre $v \in V(\lambda_j)$ poiché appartiene a $V(\lambda_j) \cap +_{i\neq j}V(\lambda_i)$. Quindi ricaviamo la seguente combinazione lineare di vettori appartenenti ad autospazi relativi ad autovalori distinti

$$v_1 + \cdots + v_{j-1} - v + v_{j+1} + \cdots + v_k = 0 \ .$$

Se uno dei vettori della combinazione lineare fosse non nullo, allora sarebbe un autovettore e quindi si avrebbe una combinazione lineare, a coefficienti non tutti nulli, di autovettori relativi ad autovalori distinti, che contraddice la Proposizione 6.0.6. Quindi tutti i vettori devono essere nulli ed in particolare $v$. □

### 6.0.1 Polinomio caratteristico

Per cercare gli autovalori di un endomorfismo $f$, consideriamo una base $B$ di $V$ e sia $A$ la matrice di $f$ rispetto a tale base. La matrice $X$ delle componenti di un autovettore $v$ relativo a $\lambda$ verifica la seguente equazione

$$AX = \lambda X$$

che si può scrivere nella forma seguente

$$(A - \lambda I)X = O, \tag{6.1}$$

dove $I$ è la matrice identità.

Si tratta quindi di capire per quali $\lambda$ il sistema omogeneo (6.1) ammette soluzioni non nulle $X$, in quanto gli autovettori devono essere non nulli. Per il teorema di Rouché-Capelli, questo può avvenire se e solo se $\det(A - \lambda I) = 0$. Quindi un autovalore deve essere soluzione della equazione

$$\det(A - \lambda I) = 0,$$

che è una equazione polinomiale di grado $n = \dim V$ in $\lambda$.

Il polinomio $P(\lambda) = \det(A - \lambda I)$ è detto *polinomio caratteristico* di $f$ e non dipende dalla matrice scelta per rappresentare $f$. Infatti se $A'$ è la matrice di $f$ rispetto alla base $B'$, e $P$ la matrice di passaggio da $B$ a $B'$, si ha $A' = PAP^{-1}$ e il suo polinomio caratteristico è dato da

$$\begin{aligned}
\det(A' - \lambda I) &= \det(PAP^{-1} - \lambda I) = \det(PAP^{-1} - \lambda PIP^{-1}) \\
&= \det(P(A - \lambda I)P^{-1}) \overset{\text{per Binet}}{=} \det P \det(A - \lambda I) \det(P^{-1}) \\
&= \det(A - \lambda I).
\end{aligned}$$

Quindi gli autovalori sono le soluzioni del polinomio caratteristico.

Per poter enunciare il teorema che ci permette di caratterizzare gli endomorfismi che ammettono una base rispetto alla quale la matrice associata è diagonale, ci occorrono le seguenti definizioni:

**Definizione 6.0.8.** Un endomorfismo $f \in End(V)$ si dice *diagonalizzabile* se esite una base di $V$ formata da autovettori di $f$.

Chiaramente la matrice di $f$ rispetto ad una tale base è diagonale.

**Definizione 6.0.9.** Una matrice quadrata $A$ sarà detta diagonalizzabile se è la matrice associata a qualche endomorfismo diagonalizzabile.

In tal caso la matrice è simile ad una matrice diagonale. Viceversa se $A$ è simile ad una matrice diagonale, allora è diagonalizzabile (verificare!).

**Definizione 6.0.10.** Sia $\lambda_0$ un autovalore di $f \in End(V)$ e $P(\lambda)$ il polinomio caratteristico di $f$. Chiameremo molteplicità algebrica $ma(\lambda_0)$ e molteplicità geometrica $mg(\lambda_0)$ di $\lambda$ i numeri naturali definiti come segue:

- $ma(\lambda_0)$ è l'unico numero naturale tale che $(\lambda - \lambda_0)^{ma(\lambda_0)}$ divide il polinomio $P(\lambda)$ mentre $(\lambda - \lambda_0)^{ma(\lambda_0)+1}$ non divide $P(\lambda)$.

- $mg(\lambda_0) = \dim V(\lambda_0)$, dove $V(\lambda_0)$ è l'autospazio relativo a $\lambda_0$.

Quindi $ma(\lambda_0)$ è la molteplicità di $\lambda_0$ come soluzione di $P(\lambda)$ ed è chiaramente maggiore o uguale a 1 poiché $\lambda_0$ è soluzione di $P(\lambda)$. Abbiamo la seguente proprietà:

**Proposizione 6.0.11.** *Siano $V$ uno spazio vettoriale di dimensione $n$, $\lambda_0$ un autovalore di $f \in End(V)$ e $P(\lambda)$ il polinomio caratteristico di $f$. Allora si ha*

$$1 \leq mg(\lambda_0) \leq ma(\lambda_0).$$

*Dimostrazione.* Poichè $\lambda_0$ è un autovalore di $f$ allora esiste almeno un autovettore, cioè un vettore non nullo che appartiene a $V(\lambda_0)$ e quindi si deve avere $mg(\lambda_0) = \dim V(\lambda_0) \geq 1$.

Siano ora $h = ma(\lambda_0)$, $k = mg(\lambda_0)$ e supponiamo per assurdo che $k > h$. Sia $(v_1, \cdots, v_k)$ una base ordinata di $V(\lambda_0)$ e completiamola per ottenere una base di $V$. Sia $B = (v_1, \cdots, v_k, g_1, \cdots, g_{n-k})$ la base ottenuta e scriviamo la matrice di $f$ rispetto a tale base:

$$A = \begin{array}{c} \begin{array}{cccccc} f(v_1) & f(v_2) & \cdots f(v_k) f(g_1) & \cdots f(g_{n-k}) \end{array} \\ \left\| \begin{array}{cccc|ccc} \lambda_0 & 0 & 0 & 0 & ???? & ???? & ???? \\ 0 & \lambda_0 & 0 & 0 & ???? & ???? & ???? \\ \vdots & \vdots & \vdots & \vdots & ???? & ???? & ???? \\ 0 & 0 & 0 & \lambda_0 & ???? & ???? & ???? \\ 0 & 0 & 0 & 0 & ???? & ???? & ???? \\ 0 & 0 & 0 & 0 & ???? & ???? & ???? \\ 0 & 0 & 0 & 0 & ???? & ???? & ???? \end{array} \right\| \end{array} \begin{array}{c} \\ v_1 \\ v_2 \\ \vdots \\ v_k \\ g_1 \\ \vdots \\ g_{n-k} \end{array} \tag{6.2}$$

Quindi se calcoliamo $\det(A - \lambda I)$, otteniamo il polinomio caratteristico della forma

$$P(\lambda) = (\lambda - \lambda_0)^k Q(\lambda),$$

il che implica che la molteplicità algebrica di $\lambda_0$ è almeno $k$. Ma ciò è assurdo perchè $k > h = ma(\lambda_0)$. Quindi si deve avere $mg(\lambda_0) \leq ma(\lambda_0)$. □

Possiamo ora enunciare il teorema che va sotto il nome di **Teorema spettrale**:

**Teorema 6.0.12.** *Siano $V$ uno spazio vettoriale di dimensione $n$, $f \in End(V)$ e $P(\lambda)$ il polinomio caratteristico di $f$. Allora $f$ è diagonalizzabile se e solo se sono verificate le due seguenti proprietà*

1. *tutti gli autovalori di $f$ sono reali,*

2. *per ogni autovalore $\lambda_0$ si ha $mg(\lambda_0) = ma(\lambda_0)$.*

*Dimostrazione.* Se $f$ è diagonalizzabile le due condizioni sono verificate (dimostrare in dettaglio per esercizio).

Viceversa se le due condizioni del teorema sono verificate, siano $\lambda_1, \cdots, \lambda_k$ gli autovalori di $f$ e denotiamo con $V_i$ l'autospazio relativo a $\lambda_i$. Abbiamo precedentemente dimostrato che la somma degli autospazi è diretta. La somma diretta $\oplus_{i=1}^k V_i$ è un sottospazio di $V$ e vogliamo dimostrare che coincide con $V$. Per far questo è sufficiente dimostrare che hanno la stessa dimensione. Calcoliamo allora la dimensione di $\oplus_{i=1}^{k} V_i$:

$$\begin{aligned}\dim \oplus_{i=1}^{k} V_i &\overset{\text{somma diretta}}{=} \sum_{i=1}^{k} \dim V_i \overset{\text{def. di } mg}{=} \sum_{i=1}^{k} mg(\lambda_i) \\ &\overset{\text{ipotesi Teor.}}{=} \sum_{i=1}^{k} ma(\lambda_i) \overset{\text{autov. tutti reali}}{=} \deg P = \dim V\end{aligned}$$

Quindi abbiamo $\oplus_{i=1}^{k} V_i = V$. Per avere una base di $V$ formata da autovettori basta prendere una base $B_i$ di ogni $V_i$ e l'unione di tali base sarà la base cercata. □

Il seguente corollario è molto utile e può essere dimostrato indipendentemente dal Teorema spettrale:

**Corollario 6.0.13.** *Se un endomorfismo $f$ ha autovalori tutti reali e distinti, allora è diagonalizzabile.*

*Dimostrazione.* Se gli autovalori sono tutti distinti, allora ognuno ha molteplicità 1 e quindi le ipotesi del Teorema 6.0.12 sono verificate, se si considera che $1 \leq mg(\lambda_i) \leq ma(\lambda_i)$. □

# Capitolo 7

# Forme bilineari

In questo capitolo ci occuperemo di forme bilineari. Il prodotto scalare nell'insieme dei vettori liberi dello spazio ne è un esempio. La definizione precisa è la seguente:

**Definizione 7.0.14.** Dicesi *forma bilineare* su uno spazio vettoriale $V$, una applicazione

$$\phi : V \times V \to \mathbb{R}$$

che è lineare in entrambi gli argomenti, ossìa tale che $\forall u, v, w \in V$ e $\forall a, b \in \mathbb{R}$ si abbia:

$$\phi(a\cdot u + b\cdot v, w) = a\phi(u, w) + b\phi(v, w) \tag{7.1}$$

$$\phi(w, a\cdot u + b\cdot v) = a\phi(w, u) + b\phi(w, v) \tag{7.2}$$

Indicheremo con $B^2(V)$ l'insieme delle forme bilineari su $V$.

*Esercizio* 7.0.15. Sull'insieme $B^2(V)$ si considerino le seguenti operazioni di somma e prodotto per uno scalare:

$$(\phi \boxplus \psi)(u, v) = \phi(u, v) + \psi(u, v) \quad , \quad (\lambda\cdot\phi)(u, v) = \lambda\phi(u, v).$$

Si dimostri che rispetto a tali operazioni l'insieme $B^2(V)$ è uno spazio vettoriale.

**Esempio 7.0.16.**

- la forma $\phi(x, y) = x_1y_1 - x_2y_1 + x_3y_1$ è una forma bilineare su $\mathbb{R}^3$, dove si è posto $x = (x_1, x_2, x_3)$ e $y = (y_1, y_2, y_3)$.
- la forma $\phi(x, y) = x_1y_1 - x_2y_2 + x_3y_3$ è una forma bilineare su $\mathbb{R}^3$.
- la forma $\phi(x, y) = x_1y_2 - x_2y_1$ è una forma bilineare su $\mathbb{R}^3$.
- il prodotto scalare definito sull'insieme dei vettori liberi dello spazio è una forma bilineare.

Tra le forme bilineari ce ne sono alcune particolari:

**Definizione 7.0.17.** Una forma bilineare $\phi \in B^2(V)$ sarà detta

- simmetrica se per ogni $u, v \in V$ si ha $\phi(u, v) = \phi(v, u)$;
- antisimmetrica se per ogni $u, v \in V$ si ha $\phi(u, v) = -\phi(v, u)$.

Si osservi che se $\phi$ è antisimmetrica allora si ha $\phi(u, u) = 0$.

**Esempio 7.0.18.**

- la forma bilineare su $\mathbb{R}^3$ data da $\phi(x, y) = x_1y_1 - x_2y_2 + x_3y_3$ è simmetrica.
- la forma bilineare su $\mathbb{R}^3$ data da $\phi(x, y) = x_1y_2 - x_2y_1$ è antisimmetrica.
- la forma bilineare su $\mathbb{R}^3$ data da $\phi(x, y) = x_1y_1 - x_2y_3$ non è ne simmetrica ne antisimmetrica.

Denoteremo con $S^2(V)$ e $A^2(V)$ i sottoinsiemi di $B^2(V)$ formati dalle forme bilineari simmetriche e antisimmetriche rispettivamente.

*Esercizio* 7.0.19. Si verifichi che $S^2(V)$ e $A^2(V)$ sono sottospazi di $B^2(V)$ (rispetto alle operazioni precedentemente definite), che la loro somma è diretta ed è uguale a tutto lo spazio $B^2(V)$.

## 7.1 Matrice associata ad una forma bilineare

Se $V$ è uno spazio vettoriale di dimensione finita $n$, $B = (e_1, \cdots, e_n)$ una sua base ordinata e $\phi \in B^2(V)$, è possibile associare a $\phi$ una matrice $A$ nel modo seguente

$$[A]_{ij} = \phi(e_i, e_j).$$

Tale matrice è detta *matrice della forma bilineare $\phi$* rispetto alla base $B$. La matrice $A$ ci permette di calcolare $\phi$ su una coppia qualunque di vettori una volta che si conoscono le loro componenti rispetto alla base $B$.

Infatti, dette $u_i$ e $v_i$ le componenti rispetto a $B$ di $u, v$ rispettivamente, e $X, Y$ le matrici colonna di tali componenti, si ha:

$$\begin{aligned}
\phi(u, v) &= \phi(\sum_{i=1}^{n} u_i \cdot e_i, \sum_{j=1}^{n} v_j \cdot e_j) \overset{\text{bilinearità}}{=} \sum_{i=1}^{n} u_i (\sum_{j=1}^{n} v_j \phi(e_i, e_j)) \\
&= \sum_{i=1}^{n} u_i [AY]_i = {}^t XAY.
\end{aligned}$$

In particolare si osservi che la forma bilineare è simmetrica (risp. antisimmetrica) se e solo se lo è la matrice rispetto ad una base qualsiasi.

A questo punto ci si può chiedere come cambia la matrice associata se si cambia base:

**Proposizione 7.1.1.** *Siano $B = (e_1, \cdots, e_n)$, $B' = (e'_1, \cdots, e'_n)$ due basi ordinate di $V^n$ e $P$ la matrice di passaggio da $B'$ a $B$. Le matrici $A, A'$ di una forma bilineare $\phi$ su $V$, rispetto a $B$ e $B'$ rispettivamente, sono legate dalla seguente relazione*

$$A' =^t PAP.$$

*Dimostrazione.* Posto $P = (p_{ij})$, si ha

$$\begin{aligned}
[A']_{ij} = \phi(e'_i, e'_j) &= \phi(\sum_{k=1}^{n} p_{ki}\cdot e_k, \sum_{s=1}^{n} p_{sj}\cdot e_s) = \sum_{k=1}^{n} p_{ki} \sum_{s=1}^{n} p_{sj}\phi(e_k, e_s) \\
&= \sum_{k=1}^{n} p_{ki} \sum_{s=1}^{n} p_{sj}[A]_{ks} = \sum_{k=1}^{n} p_{ki}[AP]_{kj} = \sum_{k=1}^{n} [{}^tP]_{ik}[AP]_{kj} \\
&= [{}^tPAP]_{ij}.
\end{aligned}$$

Questo significa che $A' =^t PAP$. □

Il fatto che due matrici della stessa forma bilineare $\phi$ rispetto a due basi siano legate come visto sopra, permette di definire il *rango* di $\phi$ come rango di una sua matrice associata. Questa definizione non dipende dalla matrice associata poichè si ha $\mathrm{rg}A' = \mathrm{rg}({}^tPAP) = \mathrm{rg}A$, dato che la matrice $P$ e la sua trasposta sono invertibili.

Non è invariante il determinante di una forma bilineare poichè si ha $\det A' = \det({}^tPAP) = (\det A)(\det P)^2$, mentre rimane invariato il suo segno, dato che si ha $(\det P)^2 > 0$.

**Definizione 7.1.2.** Una forma bilineare è detta *non degenere* se ha rango massimo, altrimenti è detta degenere.

Si osservi che una forma bilineare è non degenere se e solo se il determinante di una sua matrice associata (e quindi di una qualunque) è non nullo.

**Esempio 7.1.3.**
- la matrice rispetto alla base canonica della forma bilineare su $\mathbb{R}^3$ data da $\phi(x, y) = x_1y_1 - x_2y_2 + x_3y_3$ è la matrice diagonale $diag(1, -1, 1)$. (Qui si intende la matrice diagonale i cui elementi nella diagonale sono specificati nelle parentesi)

- la matrice rispetto alla base canonica della forma bilineare su $\mathbb{R}^3$ data da $\phi(x, y) = x_1y_2 - x_2y_1$ è $\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Si vede quindi che $\phi$ ha rango 2.

- la matrice rispetto alla base canonica della forma bilineare su $\mathbb{R}^3$ data da $\phi(x, y) = x_1y_1 - 3x_3y_2 - x_2y_3$ è $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -3 & 0 \end{pmatrix}$. In questo caso la forma bilineare è non degenere.

### 7.1.1 Dualità e rango di una forma bilineare

Data una forma bilineare $\phi \in B^2(V)$, questa determina due applicazioni lineari $s, d : V \to V^*$, che sono definite come segue:

$$s(x) = \phi(x, \cdot) \quad d(x) = \phi(\cdot, x).$$

Questo significa che $s(x) \in V^*$ è l'applicazione da $V$ in $\mathbb{R}$ tale che $s(x)(y) = \phi(x, y)$ mentre $d(x)(y) = \phi(y, x)$ .

La bilinearità di $\phi$ implica la linearità di $s(x)$ e $d(x)$ per ogni $x$ e anche la linearità delle applicazioni $s, t$ (verificare tutti i dettagli).

Vogliamo ora capire quando le applicazioni $s, d$ sono degli isomorfismi, nel caso in cui $V$ abbia dimensione finita. In tal caso è sufficiente verficare che $s, d$ siano iniettive. Vediamo il caso di $s$, in quanto quello di $d$ è del tutto simile.

Si ha $\ker s = \{x \in V | s(x) = 0\}$. Il nucleo è ridotto al solo vettore nullo se e solo se $s(x) = 0$ implica $x = 0$. Scriviamo quanto detto

$$s(x)(y) = 0 \ \forall y \in V \implies x = 0,$$

cioè $\phi(x, y) = 0 \ \forall y \in V \implies x = 0$. Allo stesso modo si ha che $d$ è iniettiva se e solo se

$$\phi(x, y) = 0 \ \forall x \in V \implies y = 0.$$

Vogliamo ora dimostrare che queste due proprietà sono equivalenti al fatto che $\phi$ sia non degenere. Per far questo consideriamo una base ordinata $B = (e_1, \cdots, e_n)$ di $V$ e la sua base duale $B^* = (e^1, \cdots, e^n)$. Scriviamo la matrice di $s$ rispetto a tali basi:

$$s(e_j) = \sum_i s_{ij} e^i.$$

Per calcolare $s_{ij}$ basta applicare $s(e_j)$ a $e^i$ e si ottiene $s_{ij} = s(e_j)(e_i) = \phi(e_j, e_i)$. Quindi la matrice di $s$ rispetto a tali basi è la trasposta della matrice di $\phi$ rispetto alla base $B$. In modo del tutto simile si ottiene che la matrice di $d$ rispetto a $B, B^*$ è la matrice di $\phi$ rispetto a $B$.

Questo implica che il rango delle applicazioni $s, d$ è uguale, poichè hanno come matrici associate due matrici che sono una la trasposta dell'altra, e inoltre tale rango è pari a quello della forma bilineare $\phi$. Riassumendo si ha:

**Proposizione 7.1.4.** *Una forma bilineare $\phi$ è non degenere se e solo se una delle due condizioni equivalenti seguenti è verificata*

- $\phi(x, y) = 0$ *per ogni* $x \in V$ *implica* $y = 0$*;*
- $\phi(x, y) = 0$ *per ogni* $y \in V$ *implica* $x = 0$*.*

## 7.2 Forme quadratiche

Ad ogni forma bilineare simmetrica $\phi \in B^2(V^n)$ è possibile associare una funzione $\Phi : V^n \to \mathbb{R}$ definita come segue:

$$\Phi(v) = \phi(v, v).$$

La funzione $\Phi$ è detta *forma quadratica associata* a $\phi$ ed è una funzione omogenea di grado due, ovvero $\Phi(\lambda \cdot v) = \lambda^2 \Phi(v)$ per ogni $v \in V$ e $\lambda$ reale.

Il nome forma quadratica è giustificato dal fatto che se $B$ è una base di $V^n$, $A = (a_{ij})$ la matrice (simmetrica) di $\phi$ rispetto a $B$ e i $v_i$ sono le componenti del vettore $v$ rispetto a $B$, allora $\Phi(v) = \sum_{i,j=1}^{n} a_{ij} v_i v_j$ è un polinomio omogeneo di secondo grado nei $v_i$.

Si osservi inoltre che il polinomio dato dalla $\Phi(v)$ si può anche scrivere come segue

$$\sum_{i=1}^{n} a_{ii} v_i^2 + \sum_{i<j} 2a_{ij} v_i v_j,$$

considerato che la matrice $A$ è simmetrica.

**Esempio 7.2.1.**

- la forma quadratica associata alla forma bilineare su $\mathbb{R}^3$ data da $\phi(x, y) = x_1 y_1 - x_2 y_2 + x_3 y_3$ è $\Phi(x) = x_1^2 - x_2^2 + x_3^2$
- la forma quadratica associata alla forma bilineare su $\mathbb{R}^3$ data da $\phi(x, y) = x_1 y_2 + x_2 y_1$ è $\Phi(x) = 2x_1 x_2$.
- la forma quadratica associata alla forma bilineare su $\mathbb{R}^3$ data da $\phi(x, y) = x_1 y_1 - 3x_3 y_3 - x_2 y_3 - x_3 y_2$ è $\Phi(x) = x_1^2 - 3x_3^2 - 2x_2 x_3$.

Nota la forma quadratica $\Phi$ associata ad una forma bilineare simmetrica $\phi$, è possibile risalire alla $\phi$ nel modo seguente

$$\begin{aligned} \Phi(u + v) &= \phi(u + v, u + v) \overset{\text{bilinearità e simmetria}}{=} \phi(u, u) + \phi(v, v) + 2\phi(u, v) \\ &= \Phi(u) + \Phi(v) + 2\phi(u, v), \end{aligned}$$

da cui si ottiene
$$\phi(u, v) = \frac{1}{2}(\Phi(u + v) - \Phi(u) - \Phi(v)),$$
che viene detta *forma polare* di $\phi$. In particolare questo dimostra la seguente proposizione

**Proposizione 7.2.2.** *Data una forma quadratica $\Phi$ associata ad una forma bilineare simmetrica su $V$, esiste un'unica $\phi \in S^2(V)$ tale che $\phi(v, v) = \Phi(v)$.*

Vediamo ora alcuni tipi particolari di forme quadratiche:

**Definizione 7.2.3.** Una forma quadratica $\Phi$ su uno spazio vettoriale $V$ è detta

1. semidefinita positiva se $\Phi(u) \geq 0$ per ogni $u \in V$;
2. definita positiva se $\Phi(u) > 0$ per ogni $u \neq 0$;
3. semidefinita negativa se $\Phi(u) \leq 0$ per ogni $u \in V$;
4. definita negativa se $\Phi(u) < 0$ per ogni $u \neq 0$;
5. indefinita se non è semidefinita positiva o negativa.

Diremo che una forma quadratica è non degenere se lo è la forma bilineare a cui è associata. Per le forme quadratiche definite positive si ha il risultato seguente:

**Proposizione 7.2.4.** *Una forma quadratica $\Phi$ definita (positiva o negativa) è non degenere.*

*Dimostrazione.* Se per assurdo $\Phi$ fosse degenere, esisterebbe $u \neq 0$ tale che $\phi(u, v) = 0$ per ogni $v \in V$. Questo varrebbe in particolare per $v = u$. Allora si avrebbe $\Phi(u) = 0$ per $u \neq 0$, contro l'ipotesi che $\Phi$ sia definita. □

**Esempio 7.2.5.** Consideriamo i seguenti esempi di forme quadratiche su $\mathbb{R}^3$:

- le forme quadratiche $x_1^2$, $x_1^2 + x_2^2$, $x_1^2 + x_2^2 + x_3^2$ sono semidefinite positive. In particolare la terza è anche definita positiva.
- le forme quadratiche $-x_1^2$, $-x_1^2 - x_2^2$, $-x_1^2 - x_2^2 - x_3^2$ sono semidefinite negative. In particolare la terza è anche definita negativa.
- le forme quadratiche $x_1^2 - x_2^2 - x_3^2$, $x_1^2 - x_2^2$, $x_1^2 - x_2^2 + x_3^2$ sono indefinite.

Un esempio importante di forma quadratica definita positiva è dato dal prodotto scalare dei vettori liberi dello spazio. Ora è possibile generalizzare questa nozione ad uno spazio vettoriale qualunque:

**Definizione 7.2.6.** Un *prodotto scalare* su uno spazio vettoriale $V$, è una forma bilineare simmetrica definita positiva (i.e. la forma quadratica associata è definita positiva). Lo spazio $V$ munito di un prodotto scalare è detto *spazio vettoriale Euclideo*.

Di solito si usa indicare un prodotto scalare col simbolo $<,>$.

### 7.2.1 Forma canonica di una forma quadratica

Abbiamo già osservato che una forma quadratica, in componenti rispetto ad una base, non è altro che un polinomio omogeneo di secondo grado. Poiché la matrice della forma quadratica dipende dalla base scelta, è naturale chiedersi se è possibile trovare una base rispetto alla quale la forma quadratica abbia un'espressione più semplice, cioè un polinomio in cui appaiono solamente termini quadratici.

**Definizione 7.2.7.** Si dice che una forma quadratica $\Phi$ su uno spazio vettoriale $V^n$ è in *forma canonica* se è espressa nel modo seguente

$$\Phi(x) = \sum_{i=1}^{n} a_i x_i^2,$$

dove le $x_i$ sono le componenti di $x$ rispetto ad una certa base e gli scalari $a_i$ possono anche essere tutti nulli.

In modo equivalente si può dire che $\Phi$ è in forma canonica se è scritta rispetto ad una base $(e_1, \cdots, e_n)$ tale che la forma bilineare $\phi$ a cui è associata $\Phi$ verifica $\phi(e_i, e_j) = 0$ per ogni $i \neq j$. La matrice di $\phi$ rispetto a tale base è una matrice diagonale.

Il teorema seguente afferma che una forma quadratica si può scrivere sempre in forma canonica:

**Teorema 7.2.8.** *Sia $\Phi$ una forma quadratica su $V^n$. Allora esiste una base (non unica) rispetto alla quale $\Phi$ si scrive in forma canonica.*

*Dimostrazione.* Dimostreremo il teorema per induzione sulla dimensione $n$ di $V$.

Per $n = 1$ la proprietà è ovvia. Supponiamo vera la proprietà per $n-1$ e dimostriamola per $n$. Siano $\Phi(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$ e $A = (a_{ij})$ la matrice di $\Phi$ rispetto alla base che si sta considerando. Distinguiamo due casi:

Caso 1 Nell'espressione di $\Phi$ almeno uno dei termini $a_{ii}$ è non nullo. Supponiamo che sia $a_{11}$ e consideriamo la forma quadratica

$$\Phi'(x) = \Phi(x) - \frac{1}{a_{11}}(a_{11}x_1 + \cdots + a_{1n}x_n)^2 = \Phi(x) - \frac{1}{a_{11}}(\sum_{j=1}^{n} a_{1j}x_j)^2.$$

Nella forma quadratica $\Phi'(x)$ scompaiono tutti i termini dove appare $x_1$, quindi dipende solo da $x_2, \cdots, x_n$. In questo modo abbiamo scritto $\Phi(x) = \Phi'(x) + \frac{1}{a_{11}}(\sum_{j=1}^n a_{1j}x_j)^2$.

A questo punto possiamo fare un cambiamento di variabile e porre

$$y_1 = \sum_{j=1}^n a_{1j}x_j \ , \ y_i = x_i \, \forall i \geq 2,$$

ottenendo in questo modo $\Phi(y_1, \cdots, y_n) = \Phi'(y_2, \cdots, y_n) + \frac{1}{a_{11}}y_1^2$. Ora è possibile applicare l'ipotesi induttiva alla forma quadratica $\Phi'(y_2, \cdots, y_n)$ che dipende da $n-1$ variabili.

Perché questa operazione abbia senso bisogna assicurarsi che il cambiamento di variabili effettuato derivi effettivamente da un cambiamento di base, cioè che $(y_1, \cdots, y_n)$ siano le componenti rispetto ad una nuova base del vettore di componenti $(x_1, \cdots, x_n)$ rispetto alla base attuale. Per renderci conto di questo osserviamo che si ha

$$Y = CX,$$

dove $Y$ è la matrice colonna di entrate $(y_1, \cdots, y_n)$, $X$ quella di entrate $(x_1, \cdots, x_n)$ e $C$ è la matrice seguente:

$$C = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Quindi $Y$ rappresenta le componenti del vettore rispetto alla nuova base in funzione delle componenti $X$ rispetto alla vecchia, mentre la matrice $C$ è la matrice di passaggio dalla vecchia base alla nuova. La matrice di $\Phi$ rispetto alla nuova base sarà allora ${}^tC^{-1}AC^{-1}$.

Il cambiamento di base che si è effettuato (che ha come matrice di passaggio $C^{-1}$) è il seguente:

$$\begin{aligned} e_1' &= \frac{1}{a_{11}}e_1 \\ e_i' &= e_i - \frac{a_{1i}}{a_{11}}e_1 \text{ per } i \geq 2. \end{aligned}$$

Caso 2 Tutti gli $a_{ii}$ sono nulli, ma almeno uno degli $a_{ij}$, con $i \neq j$, è non nullo. Se anche tutti gli $a_{ij}$ sono nulli allora $\Phi$ è la forma quadratica nulla e quindi è già in forma canonica.

Supponiamo che non sia nullo il coefficiente $a_{12}$ e operiamo il seguente cambiamento di variabile:

$$x_1 = z_1 + z_2 \quad , x_2 = z_1 - z_2 \quad , x_i = z_i \ \forall i \geq 3.$$

La forma quadratica diventa allora

$$\Phi(z_1, \cdots, z_n) = z_1^2 - z_2^2 + \sum_{ij} b_{ij} z_i z_j,$$

per certi scalari $b_{ij}$. Inoltre $z_1^2$ e $z_2^2$ appaiono effettivamente nell'espressione di $\Phi(z_1, \cdots, z_n)$ e quindi ci si può ricondurre al Caso 1.

Anche per questo tipo di operazione dobbiamo accertarci che si tratti di un cambiamento di base. In effetti, dette $X$ e $Z$ le matrici colonna di entrate $(x_1, \cdots, x_n)$ e $(z_1, \cdots, z_n)$ rispettivamente, si ha

$$X = PZ,$$

dove la matrice $P$ è

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

Quindi $P$ rappresenta le componenti della nuova base rispetto alla vecchia. Il cambiamento di base che si è effettuato è il seguente:

$$e_1' = e_1 + e_2 \quad , e_2' = e_1 - e_2 \quad e_i' = e_i \ \forall i \geq 3.$$

Se in $n$ passi si hanno $n$ cambiamenti di base dati dalle matrici $A_1, \cdots, A_n$, la matrice del cambiamento di base finale sarà data da $A_1 A_2 \cdot A_n$. □

*Esercizio* 7.2.9. Scrivere in forma canonica la forma quadratica su $\mathbb{R}^3$ data da $\Phi(x) = x_1x_2 + x_1x_3 + x_2x_3$.

Il teorema sulla forma canonica di una forma quadratica può essere migliorato ottenendo il seguente teorema:

**Teorema 7.2.10** (Teorema di Sylvester)**.** *Sia $\Phi$ una forma quadratica su uno spazio vettoriale $V$ di dimensione $n$ e $\phi$ la forma bilineare simmetrica associata. Sia*

*$p$ il rango di $\Phi$. Allora esiste un numero $s \leq p$ univocamente determinato da $\Phi$ e una base $(e_1, \cdots, e_n)$ tale che*

$$\begin{aligned} \phi(e_i, e_j) &= 0 \ \ i \neq j \ \ , \\ \phi(e_i, e_i) &= 1 \ per \ i \leq s \\ \phi(e_i, e_i) &= -1 \ per \ s+1 \leq i \leq p \\ \phi(e_i, e_i) &= 0 \ per \ i \geq p+1. \end{aligned}$$

*La forma quadratica rispetto a tale base ha la seguente espressione:*

$$\Phi(x) = \sum_{i=1}^{s} x_i^2 - \sum_{i=s+1}^{p} x_i^2.$$

I numeri $s$ e $p - s$ sono detti rispettivamente indice di positività e di negatività di $\Phi$ e la coppia $(s, p - s)$ è detta segnatura.

*Dimostrazione.* L'esistenza di una tale forma canonica, si ottiene a partire da una forma canonica di $\Phi$, la cui esistenza è assicurata dal Teorema 7.2.8. Supponiamo che $\Phi$ sia scritta nel modo seguente:

$$\sum_{i=1}^{p} a_i x_i^2,$$

dove $p$ è il rango di $\Phi$ e dove possiamo supporre che i primi $s$ coefficienti siano positivi e a partire dal $(s+1)$-esimo e fino al $p$-esimo negativi. Per ottenere questo, basta infatti operare un semplice cambiamento di variabile (e quindi di base).

A questo punto possiamo fare il seguente cambiamento di variabili

$$\begin{aligned} y_i &= \sqrt{a_i} x_i \ \ i \leq s \ \ , \\ y_i &= \sqrt{-a_i} x_i \ \ s+1 \leq i \leq p, \\ y_i &= x_i \ \ p+1 \leq i \leq n, \end{aligned}$$

che porta $\Phi$ nella forma voluta.

Dobbiamo ora dimostrare l'unicità di $s$, quella di $p$ essendo ovvia poichè pari al rango di $\phi$. Supponiamo per assurdo che esista $t \neq s$ tale che si abbia rispetto a due basi:

$$\Phi(v) = \sum_{i=1}^{s} x_i^2 - \sum_{i=s+1}^{p} x_i^2 = \sum_{i=1}^{t} y_i^2 - \sum_{i=s+1}^{p} y_i^2,$$

dove le $x_i$ e le $y_i$ sono le componenti di $v$ rispetto alle basi $(e_1, \cdots, e_n)$ ed $(e'_1, \cdots, e'_n)$ rispettivamente. Possiamo supporre $s > t$ (la dimostrazione è identica per $s < t$ ) e considerare i seguenti sottospazi:

$$V_1 = L(e_1, \cdots e_s) \ \ , V_2 = L(e'_{t+1}, \cdots e'_n).$$

Poichè si ha $\dim V_1 + \dim V_2 = s + n - t > n = \dim V$, deve esistere un vettore non nullo $u \in V_1 \cap V_2$.

Allora avremo $u = \sum_{i=1}^{s} x_i \cdot e_i = \sum_{i=t+1}^{n} y_i e'_i$ e quindi

$$\Phi(u) = \sum_{i=1}^{s} x_i^2 \geq 0 \ , \ \Phi(u) = -\sum_{i=t+1}^{n} y_i^2 \leq 0,$$

che implica $u = 0$ e quindi una contraddizione. □

Una conseguenza immediata è il seguente

**Corollario 7.2.11.** *Se $\Phi$ è definita positiva (risp. negativa), allora ha segnatura $(n, 0)$ (risp. $(0, n)$).*

# Capitolo 8

# Spazi euclidei

Abbiamo precedentemente dato la definizione di prodotto scalare in uno spazio vettoriale. Ora vogliamo studiare alcune proprietà relative a questa struttura.

**Definizione 8.0.12.** Uno spazio vettoriale $V$ è detto *spazio vettoriale euclideo* se è munito di un prodotto scalare $<, >$.

**Esempio 8.0.13.** Lo spazio vettoriale $\mathbb{R}^n$ munito del prodotto scalare standard $< x, y >= \sum_{i=1}^{n} x_i y_i$ è uno spazio euclideo. (Si verifichino tutti i dettagli!)

In analogia col prodotto scalare nell'insieme dei vettori liberi dello spazio, diamo le seguenti definizioni:

**Definizione 8.0.14.** Due vettori $u, v$ di uno spazio vettoriale euclideo sono detti ortogonali se $< u, v >= 0$.

Poichè il prodotto scalare è definito positivo ha senso dare la seguente

**Definizione 8.0.15.** Il *modulo* o *norma* di un vettore $u$ è il numero reale $||u|| = \sqrt{< u, u >}$.

Si osservi che il modulo di un vettore $v$ è zero se e solo se $v$ è il vettore nullo.

**Definizione 8.0.16.** Una base $(e_1, \cdots, e_n)$ di uno spazio vettoriale euclideo è detta *ortogonale* se $< e_i, e_j >= 0$ per ogni $i \neq j$, mentre è detta *ortonormale* se si ha $< e_i, e_j >= \delta_{ij}$.

Se $V$ ha dimensione finita $n$, il Teorema di Sylvester assicura l'esistenza di una base ortonormale.

**Definizione 8.0.17.** Siano $v \in V$ un vettore non nullo e $u \in V$. Il vettore proiezione ortogonale di $u$ su $v$, è il vettore $< u, \frac{1}{||v||}v > \frac{1}{||v||}v$. Il numero reale $< u, \frac{1}{||v||}v >$ è detto componente ortogonale di $u$ su $v$.

La seguente proposizione sarà utile per il seguito:

**Proposizione 8.0.18.** *Se $v_1, \cdots, v_k \in V$ sono vettori non nulli e a due a due ortogonali, allora sono linearmente indipendenti.*

*Dimostrazione.* Consideriamo $a_1, \cdots, a_k \in \mathbb{R}$ tali che

$$\sum_i a_i \cdot v_i = 0$$

e calcoliamo il prodotto scalare del primo e del secondo membro con $v_j$ (per $j = 1, \cdots, k$). Utilizzando la bilinearità del prodotto scalare si ottiene quanto segue:

$$0 =< \sum_i a_i \cdot v_i, v_j >= \sum_i a_i < v_i, v_j >= a_j ||v_j||,$$

da cui si ricava $a_j = 0$ per ogni $j$, dato che i $v_j$ sono tutti non nulli. Questo implica che i vettori $v_1, \cdots, v_k \in V$ sono linearmente indipendenti.

□

### 8.0.2 Metodo di ortonormalizzazione di Gram-Schmidt

In questo paragrafo mostriamo un metodo per ottenere una base ortonormale a partire da una base qualunque di $V$. Tale costruzione è nota come *metodo di ortonormalizzazione di Gram-Schmidt*.

Sia $(v_1, \cdots, v_n)$ una base di $V$. Poniamo innanzitutto $e_1 = \frac{1}{||v_1||} v_1$. In questo modo otteniamo un vettore $e_1$ che ha modulo 1. Per costruire $e_2$, partiamo da $v_2$ che però non è ortogonale a $e_1$. Per aggiustare le cose possiamo considerare il vettore

$$v_2 - < v_2, e_1 > e_1,$$

che è ortogonale ad $e_1$ e non nullo. Possiamo allora definire

$$e_2 = \frac{1}{||v_2 - < v_2, e_1 > e_1||}(v_2 - < v_2, e_1 > e_1).$$

Si verifica facilmente (verificare tutti i dettagli) che $e_1$ ed $e_2$ sono a due a due ortogonali e non nulli, e quindi linearmente indipendenti. A questo punto è chiaro come costruire la base ortonormale per induzione. Infatti, una volta costruita la $k$-upla ortonormale $(e_1, \cdots, e_k)$, si considera il vettore

$$v_{k+1} - \sum_{i=1}^{k} < v_{k+1}, e_i > e_i,$$

che è non nullo e ortogonale a $e_1, \cdots, e_k$, lo si divide per il suo modulo e si ottiene $e_{k+1}$.

### 8.0.3 Diseguaglianze fondamentali

In un spazio vettoriale euclideo sono verificate alcune diseguaglianze molto importanti.

**Proposizione 8.0.19** (Diseguaglianza di Schwarz)**.** *Dato uno spazio vettoriale euclideo $V$, per ogni $u, v \in V$ si ha:*

$$| < u, v > | \leq \|u\| {\cdot} \|v\|. \tag{8.1}$$

*l'uguaglianza vale se e solo se $u, v$ sono linearmente dipendenti.*

*Dimostrazione.* Se $v = 0$ la diseguaglianza è banalmente verificata. Se $v \neq 0$ allora per ogni numero reale $a$ e per ogni $u \in V$, si ha

$$\|u + a{\cdot}v\|^2 \geq 0 \tag{8.2}$$

ossìa

$$\|u + a{\cdot}v\|^2 = \|v\|^2 a^2 + 2 < u, v > a + \|u\|^2 \geq 0. \tag{8.3}$$

Poichè $\|v\| > 0$ la disequazione è verificata per ogni $a$ reale se e solo se $\Delta \leq 0$, cioè se e solo se

$$< u, v >^2 - \|u\|^2 {\cdot} \|v\|^2 \leq 0, \tag{8.4}$$

che equivale alla diseguaglianza di Schwarz.

Se i due vettori sono linearmente dipendenti si verifica facilmente che vale l'uguaglianza in (8.1). Viceversa se vale l'uguaglianza in (8.1), allora il $\Delta$ di (8.3) è nullo e l'equazione

$$\|u + a{\cdot}v\|^2 = \|v\|^2 a^2 + 2 < u, v > a + \|u\|^2 = 0. \tag{8.5}$$

ammette la soluzione doppia $a_0 = \frac{<u,v>}{\|v\|^2}$ e quindi si ha $\|u + a_0{\cdot}v\| = 0$ cioè $u = -a_0 v$. □

Si osservi che la diseguaglianza di Schwarz si può scrivere come segue:

$$- \|u\| {\cdot} \|v\| \leq < u, v > \leq \|u\| {\cdot} \|v\|. \tag{8.6}$$

Se $u$ e $v$ sono vettori non nulli, la (8.6) diventa

$$- 1 \leq \frac{< u, v >}{\|u\| {\cdot} \|v\|} \leq 1. \tag{8.7}$$

Questo è molto importante in quanto permette di dare la seguente definizione che generalizza ad uno spazio euclideo qualunque il concetto di angolo tra due vettori:

**Definizione 8.0.20.** Se $u$ e $v$ sono vettori non nulli di uno spazio euclideo $V$, si definisce l'*angolo* tra $u$ e $v$ l'unico numero reale $\theta$ tale che $0 \leq \theta \leq \pi$ e

$$\frac{< u, v >}{||u|| \cdot ||v||} = \cos\theta.$$

Si osservi che, con questa definizione di angolo, due vettori non nulli sono ortogonali se e solo se il loro angolo vale $\frac{\pi}{2}$. Inoltre, per quanto mostrato precedentemente, due vettori non nulli sono linearmente dipendenti (cioé paralleli) se e solo se il loro angolo vale $0$ o $\pi$.

Una conseguenza diretta della (8.1) è la seguente

**Proposizione 8.0.21** (Disuguaglianza di Minkowski o triangolare)**.** *Dato uno spazio vettoriale euclideo $V$, per ogni $u, v \in V$ si ha:*

$$||u + v|| \leq ||u|| + ||v||, \tag{8.8}$$

$$||u + v|| \geq |\, ||u|| - ||v||\, |. \tag{8.9}$$

*Dimostrazione.* Se si calcola $||u + v||^2$ e si applica la diseguaglianza di Schwarz (8.6) si ottiene

$$||u+v||^2 =< u+v, u+v >= ||u||^2+2 < u, v > +||v||^2 \leq ||u||^2+2||u|| \cdot ||v||+||v||^2 = (||u||+||v||)^2$$

da cui si ottiene immediatamente la (8.8) estraendo la radice. Utilizzando l'altra disuguaglianza di (8.6) si ottiene la (8.9).

□

Se si pensa ai vettori liberi dello spazio, la disuguaglianza di Minkowski ci dice che la somma delle lunghezze di due lati di un triangolo è maggiore o uguale alla lunghezza del terzo.

Si osservi infine che vale la seguente uguaglianza quando due vettori $u, v$ sono perpendicolari:

$$||u + v||^2 = ||u||^2 + 2 < u, v > +||v||^2 = ||u||^2 + ||v||^2.$$

Questa uguaglianza non è altro che il teorema di Pitagora.

# Bibliografia

[1] Herstein, Algebra