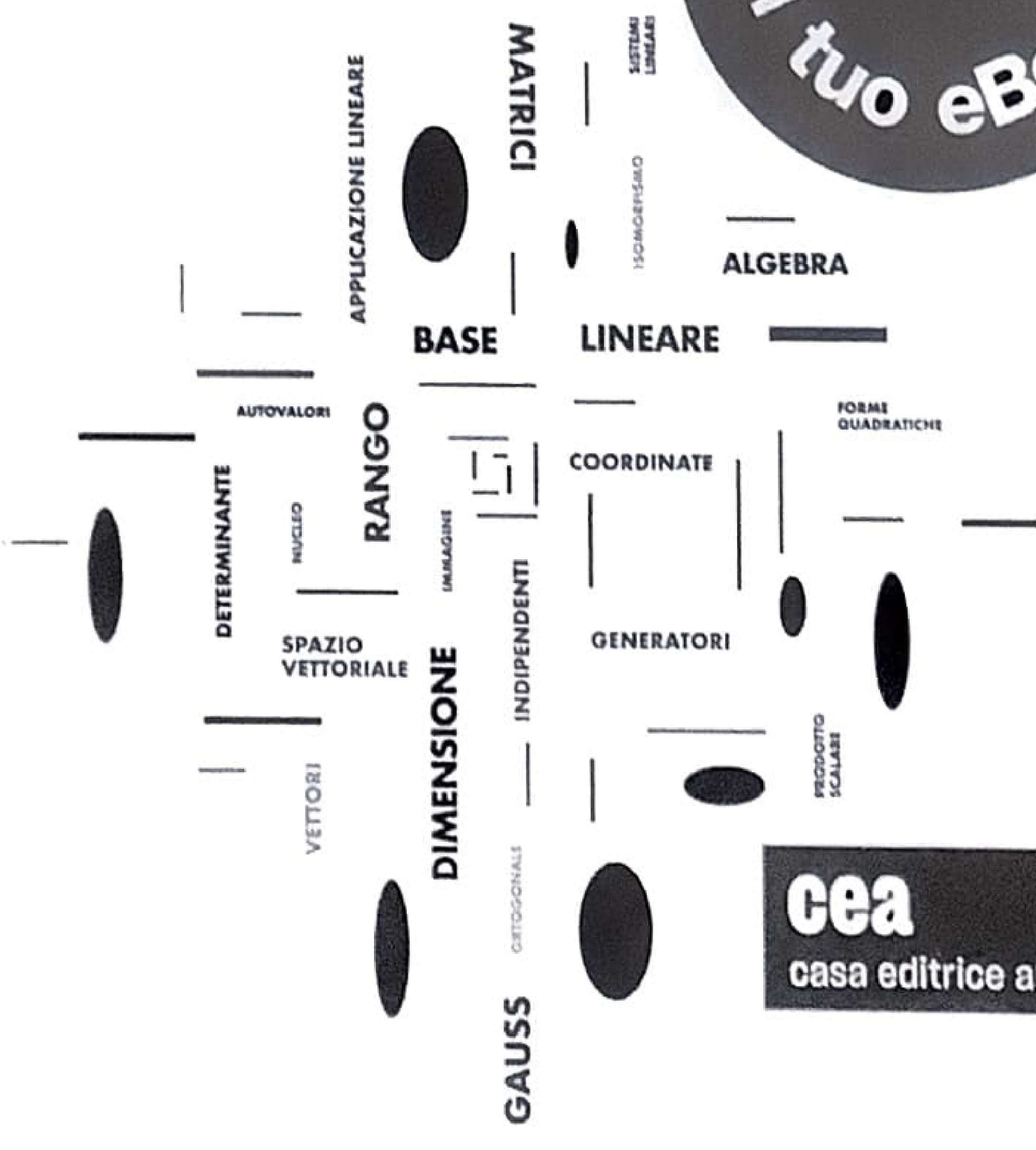
Rita Fioresi
Marta Morigi

# Introduzione all'algebra lineare

**Seconda edizione**






cea
casa editrice ambrosiana

Rita Fioresi
Marta Morigi

# Introduzione all'algebra lineare

## Seconda edizione



**Se vuoi accedere alle risorse online riservate**

1. Vai su **my.zanichelli.it**
2. Clicca su *Registrati*.
3. Scegli *Studente*.
4. Segui i passaggi richiesti per la registrazione.
5. Riceverai un'email: clicca sul link per completare la registrazione.
6. Cerca il tuo codice di attivazione stampato in verticale sul bollino argentato in questa pagina.
7. Inseriscilo nella tua area personale su **my.zanichelli.it**

Se sei già registrato, per accedere ai contenuti riservati di altri volumi ti serve solo il relativo codice di attivazione.

**cea**
casa editrice ambrosiana

*Impaginazione*: CompoMat, Configni (RI)
*Copertina*: © Anchora, Milano
*Immagine di copertina*: © Anchora, Milano

Prima edizione: febbraio 2015
Seconda edizione: giugno 2021

Ristampa: **prima tiratura**
5 4 3 2 1 2021 2022 2023 2024 2025

Realizzare un libro è un'operazione complessa, che richiede numerosi controlli: sul testo, sulle immagini e sulle relazioni che si stabiliscono tra loro. L'esperienza suggerisce che è praticamente impossibile pubblicare un libro privo di errori. Saremo quindi grati ai lettori che vorranno segnalarceli.

Per segnalazioni o suggerimenti relativi a questo libro rivolgersi a:
CEA – Casa Editrice Ambrosiana
viale Romagna 5, 20089 Rozzano (MI)
fax 02 52202260 e-mail: segreteria_cea@ceaedizioni.it

Sul sito **online.universita.zanichelli.it/fioresi2e** è possibile verificare se sono disponibili errata corrige o aggiornamenti per questo volume.

Stampa: Fotoincisa Bi-Co
Via della Fisica 33, San Lazzaro di Savena (BO)
per conto di Zanichelli editore S.p.A.
Via Irnerio 34, 40126 Bologna

# Indice generale

# Prefazione

Questo testo è rivolto principalmente, ma non solo, a chi studia Informatica e Ingegneria e vuole studiare l'algebra lineare attraverso un approccio rigoroso e allo stesso tempo concreto, che fornisca l'abilità di affrontare e risolvere varie tipologie di problemi, indispensabili per le applicazioni.

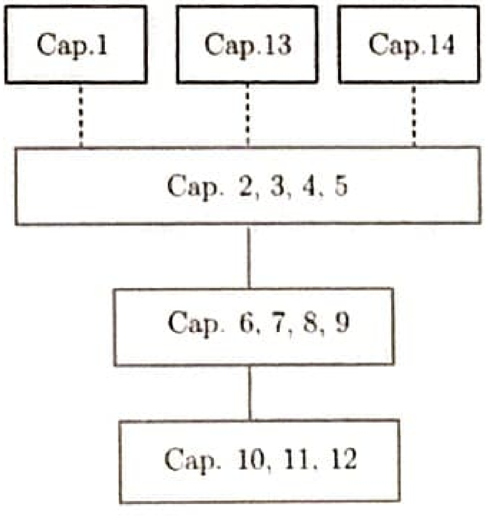
Per questo motivo, pur riportando integralmente le dimostrazioni, abbiamo dato grande spazio agli esempi e alle spiegazioni intuitive di tutti i risultati che proponiamo.

Vogliamo ringraziare in primo luogo tutti gli studenti e le studentesse di Informatica e di Fisica, che ci hanno insegnato in questi anni a rendere più accessibile questa bellissima materia proposta al primo anno di corso, che talvolta risulta ostica in quanto astratta. Invitiamo i nostri futuri studenti ad appassionarsi, come alcuni dei loro predecessori, e a cogliere la perfezione di questa teoria.

In questa seconda edizione abbiamo aggiunto una parte riguardante i prodotti scalari, con i relativi risultati più salienti, insieme a un'introduzione ad argomenti più geometrici, quali lo studio di rette e piani nello spazio tridimensionale e di alcune coniche nel piano.

Vogliamo infine ringraziare il professor Francesco Faglioni per l'aiuto tecnico e il Dipartimento di Matematica dell'Università di Bologna, che ci ha sostenuto fornendoci il supporto informatico e logistico in questi anni.

Il seguente diagramma indica alcune possibili modalità di lettura del libro: i Capitoli 1, 13, 14 possono, ciascuno indipendentemente dagli altri due, essere affrontati prima delle definizioni fondamentali dell'algebra lineare, oppure possono essere omessi e alcune nozioni eventualmente richiamate successivamente.




I Capitoli 2, 3, 4, 5 introducono le nozioni essenziali sugli spazi vettoriali e sulle applicazioni lineari, mentre nei Capitoli 6, 7, 8, 9 si sviluppa la teoria sino ad arrivare al problema della determinazione di autovalori e autovettori. Un corso essenziale di algebra lineare può terminare dopo il Capitolo 5 oppure, più compiutamente, dopo il Capitolo 9. Nei restanti Capitoli 10, 11, 12 si studiano i prodotti scalari e le forme quadratiche.

# 1 Introduzione ai sistemi lineari

Ci proponiamo di risolvere un qualsiasi sistema lineare a coefficienti reali attraverso un metodo noto come *algoritmo di Gauss*. In seguito useremo questo metodo anche per risolvere problemi diversi dalla risoluzione dei sistemi lineari e, nello stesso tempo, interpreteremo i sistemi lineari come casi particolari di una teoria molto più ampia.

## 1.1 SISTEMI LINEARI: PRIMI ESEMPI

Un'*equazione lineare* è un'equazione in cui le incognite compaiono con grado 1, cioè un'equazione della forma:

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

dove $a_1, a_2, \ldots, a_n$ e $b$ sono numeri assegnati e $x_1, x_2, \ldots, x_n$ sono le incognite. I numeri $a_1, \ldots, a_n$ si dicono *coefficienti* dell'equazione lineare, $b$ si chiama *termine noto*. Se $b = 0$ l'equazione si dice *omogenea*. Una *soluzione* della equazione (1.1) è una $n$-upla di numeri $(s_1, s_2, \ldots, s_n)$ che, sostituiti ordinatamente alle incognite, verificano l'uguaglianza, cioè tali che

$$a_1s_1 + a_2s_2 + \ldots + a_ns_n = b$$

Per esempio $(3, -1, 4)$ è una soluzione dell'equazione $2x_1 + 7x_2 - x_3 = -5$ perché $2 \cdot 3 + 7 \cdot (-1) - 4 = -5$.

Un *sistema lineare di $m$ equazioni nelle $n$ incognite* $x_1, x_2, \ldots, x_n$ è un insieme di $m$ equazioni lineari nelle $n$ incognite $x_1, x_2, \ldots, x_n$ che devono essere soddisfatte contemporaneamente:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \quad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{1.2}$$

I numeri $a_{11}, \ldots, a_{1n}, \ldots, a_{m1}, \ldots, a_{mn}$ si chiamano *coefficienti* del sistema, $b_1, \ldots, b_m$ *termini noti*. Se $b_i = 0$ per ogni $i = 1, \ldots, m$ il sistema si dice *omogeneo*. Una *soluzione* del sistema lineare (1.2) è una $n$-upla $(s_1, s_2, \ldots, s_n)$

di numeri che è soluzione di tutte le equazioni del sistema. Per esempio $(1, 2)$ è soluzione del sistema lineare

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

In questo libro ci occuperemo esclusivamente di sistemi lineari a *coefficienti reali* cioè di sistemi della forma (1.2), in cui tutti i coefficienti $a_{ij}$ delle incognite e tutti i termini noti $b_i$ sono numeri reali. Le soluzioni che cercheremo, dunque, saranno sempre $n$-uple (ordinate) di numeri reali.

Dato un sistema lineare, ci prefiggiamo di rispondere alle seguenti domande:

1) Il sistema ammette soluzioni?

2) In caso affermativo, quante e quali sono?

In certi casi rispondere a queste domande è particolarmente facile. Vediamo qualche esempio:

---

**Esempio 1.1.1**

Consideriamo il seguente sistema lineare nelle incognite $x_1, x_2$:

$$\begin{cases} x_1 + x_2 = 3 \\ x_1 + x_2 = 1 \end{cases}$$

È immediato osservare che la somma di due numeri reali non può essere contemporaneamente uguale a 3 e a 1. Le condizioni assegnate dalle due equazioni del sistema sono incompatibili, dunque il sistema non può avere soluzioni.

---

L'esempio appena fatto giustifica la seguente definizione:

**Definizione 1.1.2** Un sistema si dice *compatibile* se ammette soluzioni.

---

**Esempi 1.1.3**

Consideriamo il seguente sistema lineare nelle incognite $x_1, x_2$:

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

Sostituendo nella prima equazione il valore di $x_2$ ottenuto in quanto fissato dalla seconda equazione, otteniamo: $x_1 = 3 - x_2 = 3 + 1 = 4$. Il sistema è dunque compatibile e ammette un'unica soluzione: $(4, -1)$. In questo esempio vengono assegnate due variabili (le incognite $x_1$ e $x_2$) e due condizioni (le due equazioni del sistema). Tali condizioni sono compatibili, cioè non si contraddicono a vicenda, e sono "indipendenti" nel senso che non possono essere ottenute una dall'altra. In sintesi: *due variabili reali, insieme a due condizioni compatibili, danno luogo a una e una sola soluzione.*

**Esempio 1.1.4**

Consideriamo ora il sistema lineare

$$\begin{cases} x_1 + x_2 = 3 \\ 2x_1 + 2x_2 = 6 \end{cases}$$

nelle incognite $x_1, x_2$. Diversamente da quanto succedeva nell'esempio precedente, qui le condizioni assegnate dalle due equazioni non sono "indipendenti": la seconda equazione si ottiene moltiplicando la prima per 2. Le due equazioni individuano dunque la stessa relazione tra le variabili $x_1$ e $x_2$, quindi risolvere il sistema lineare assegnato equivale a risolvere semplicemente l'equazione $x_1 + x_2 = 3$. Tale equazione ha certamente soluzioni: abbiamo visto nell'esempio precedente che $(4, -1)$ è soluzione dell'equazione, ma anche $(1, 2)$ o $(0, 3)$ sono soluzioni dell'equazione. Quante sono allora esattamente le soluzioni di questa equazione? E come sono fatte? In questo caso abbiamo due variabili e una sola condizione su di esse. Questo significa che una variabile è libera e, siccome varia nell'insieme dei numeri reali, che sono infiniti, essa può assumere infiniti valori diversi. L'equazione assegnata ci permette di esprimere una variabile in funzione della variabile libera. Le soluzioni dell'equazione saranno solo quelle della forma: $(x_1, 3 - x_1)$. Con questa scrittura si intende che la variabile $x_1$ può assumere tutti gli infiniti valori reali e che, affinché l'equazione $x_1 + x_2 = 3$ sia soddisfatta, deve essere $x_2 = 3 - x_1$. Un modo più esplicito, ma ovviamente equivalente di descrivere le soluzioni, è dire che esse sono date dall'insieme $\{(t, 3 - t) | t \in \mathbb{R}\}$. Naturalmente avremmo potuto decidere di far variare liberamente la variabile $x_2$ e di esprimere $x_1$ in funzione di $x_2$. In tal caso avremmo descritto ogni soluzione del sistema nella forma $(3 - x_2, x_2)$, o equivalentemente avremmo affermato che l'insieme delle soluzioni è: $\{(3 - s, s) | s \in \mathbb{R}\}$. Il sistema assegnato ha dunque infinite soluzioni. In sintesi: *due variabili reali insieme a una sola condizione danno luogo a infinite soluzioni.*

**Definizione 1.1.5** Due sistemi lineari si dicono *equivalenti* se hanno le stesse soluzioni.

Nell'Esempio 1.1.4 abbiamo osservato che il sistema lineare

$$\begin{cases} x_1 + x_2 = 3 \\ 2x_1 + 2x_2 = 6 \end{cases}$$

è equivalente all'equazione $x_1 + x_2 = 3$. Riuscire a capire se due sistemi sono equivalenti può essere molto utile; per esempio potremmo tentare di risolvere un sistema lineare riducendolo a uno ad esso equivalente ma più semplice da risolvere.

Nel prossimo paragrafo introdurremo alcune nozioni utili per semplificare la scrittura di un sistema lineare.

## 1.2 MATRICI

Dati due numeri naturali $m, n$, si chiama *matrice* $m \times n$ a coefficienti reali una tabella di $mn$ numeri reali collocati su $m$ righe e $n$ colonne. Per esempio:

$$\begin{pmatrix} 5 & -6 & 0 \\ 4 & 3 & -1 \end{pmatrix}$$

è una matrice $2 \times 3$.

Se $m = n$ la matrice si dice *quadrata* di ordine $n$. Per esempio:

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

è una matrice quadrata di ordine 2.

Indicheremo con $M_{m,n}(\mathbb{R})$ l'insieme delle matrici $m \times n$ a coefficienti reali e con $M_n(\mathbb{R})$ l'insieme delle matrici quadrate di ordine $n$ a coefficienti reali.

Data una matrice $A$, il numero reale che compare nella $i$-esima riga e nella $j$-esima colonna di $A$ viene detto *elemento di posto* $(i, j)$ di $A$, o anche *coefficiente di posto* $(i, j)$. Per esempio nella matrice

$$A = \begin{pmatrix} 5 & -6 & 0 \\ 4 & 3 & -1 \end{pmatrix}$$

l'elemento di posto $(1, 3)$ è $0$ e l'elemento di posto $(2, 2)$ è $3$. Naturalmente due matrici $m \times n$, $A$ e $B$, sono uguali se coincidono coefficiente per coefficiente, cioè se l'elemento di posto $(i, j)$ di $A$ coincide con l'elemento di posto $(i, j)$ di $B$, per ogni $i = 1, \dots, m$ e per ogni $j = 1, \dots, n$.

Data una generica matrice $m \times n$:

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

essa può essere sinteticamente indicata nel modo seguente:

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

dove $a_{ij}$ è l'elemento di posto $(i, j)$, $i = 1, \dots, m$, $j = 1, \dots, n$.

Date due matrici $A$ e $B$ appartenenti entrambe a $M_{m,n}(\mathbb{R})$, è possibile affettuare la loro somma, coefficiente per coefficiente, cioè se

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

la matrice $C = A + B$ ha come coefficiente di posto $(i, j)$ l'elemento

$$c_{ij} = a_{ij} + b_{ij}$$

con $i = 1, \dots, m$, $j = 1, \dots, n$.

Per esempio:

$$\begin{pmatrix} 1 & 0 & -1 \\ 3 & -5 & 1 \end{pmatrix} + \begin{pmatrix} -1 & 2 & -1 \\ 1 & 7 & -2 \end{pmatrix} = \begin{pmatrix} 0 & 2 & -2 \\ 4 & 2 & -1 \end{pmatrix}$$

Vogliamo ora definire il prodotto righe per colonne, o semplicemente prodotto, tra due matrici $A$ e $B$, nel caso in cui le righe di $A$ abbiano la stessa lunghezza delle colonne di $B$.

Se $A$ è una matrice $m \times s$ e $B$ è una matrice $s \times n$, definiamo il prodotto $c_{ij}$ della riga $i$ di $A$ e della colonna $j$ di $B$ nel modo seguente:

$$c_{ij} = \begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{is} \end{pmatrix} \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{sj} \end{pmatrix} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{is}b_{sj}$$

che si indica in forma compatta con

$$c_{ij} = \sum_{h=1}^{s} a_{ih}b_{hj}$$

In pratica abbiamo moltiplicato, nell'ordine, i coefficienti della riga $i$-esima di $A$ per i coefficienti della colonna $j$-esima di $B$, dopodiché abbiamo sommato i numeri ottenuti.

Per esempio, se

$$A = \begin{pmatrix} 1 & 0 & 3 & -1 \\ 0 & -2 & 2 & 1 \\ 1 & 0 & -1 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 1 \\ -3 & 5 \\ 1 & 0 \\ 2 & -1 \end{pmatrix}$$

si ha

$$\begin{aligned} c_{12} &= 1 \cdot 1 + 0 \cdot 5 + 3 \cdot 0 + (-1) \cdot (-1) = 2 \\ c_{31} &= 1 \cdot 0 + 0 \cdot (-3) + (-1) \cdot 1 + 0 \cdot 2 = -1 \end{aligned}$$

A questo punto definiamo il prodotto di $A$ e $B$ come

$$C = AB = (c_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$$

La matrice $C$ prodotto di $A$ e $B$ è quindi una matrice $m \times n$.

Nell caso dell'esempio precedente abbiamo

$$C = \begin{pmatrix} 1 & 0 & 3 & -1 \\ 0 & -2 & 2 & 1 \\ 1 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -3 & 5 \\ 1 & 0 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 10 & -11 \\ -1 & 1 \end{pmatrix}$$

Notiamo che in generale il numero di righe di $AB$ è uguale al numero di righe di $A$ e il numero di colonne di $AB$ è uguale al numero di colonne di $B$.

Osserviamo anche che il prodotto di una matrice $m \times n$ e una matrice $n \times 1$ dà come risultato una matrice di $m \times 1$.

**Proposizione 1.2.1** *L'operazione di prodotto tra matrici gode della proprietà associativa, cioè $(AB)C = A(BC)$ ove $A$, $B$, $C$ sono matrici con un numero di righe e colonne tale che i prodotti che compaiono nella formula siano definiti. Inoltre vale la proprietà distributiva del prodotto rispetto alla somma, cioè $A(B+C) = AB + AC$, purché le operazioni di somma e prodotto che compaiono nella formula siano definite.*

**Dimostrazione** – Le dimostrazioni sono un puro calcolo di applicazione della definizione. Mostreremo solo l'associatività del prodotto. Per comodità, indicheremo l'elemento di posto $(i, j)$ di una generica matrice $M$ con $(M)_{ij}$. Siano $A \in \mathrm{M}_{m,s}(\mathbb{R})$, $B \in \mathrm{M}_{s,r}(\mathbb{R})$, $C \in \mathrm{M}_{r,n}(\mathbb{R})$. Osserviamo che:

$$(AB)_{iu} = \sum_{h=1}^{s} a_{ih}b_{hu}, \; (BC)_{hj} = \sum_{u=1}^{r} b_{hu}c_{uj}$$

quindi:

$$\begin{aligned}
((AB)C)_{ij} &= \sum_{u=1}^{r}(AB)_{iu}c_{uj} = \sum_{u=1}^{r}\left(\sum_{h=1}^{s} a_{ih}b_{hu}\right)c_{uj} \\
&= \sum_{u=1}^{r}\sum_{h=1}^{s} a_{ih}b_{hu}c_{uj} = \sum_{h=1}^{s}\sum_{u=1}^{r} a_{ih}b_{hu}c_{uj} \\
&= \sum_{h=1}^{s} a_{ih}\left(\sum_{u=1}^{r} b_{hu}c_{uj}\right) = \sum_{h=1}^{s} a_{ih}(BC)_{hj} = (A(BC))_{ij}
\end{aligned}$$

La dimostrazione della distributività è analoga. □

Si noti che l'operazione di prodotto tra matrici non gode della proprietà *commutativa*. Per esempio se

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

si ha che

$$AB = \begin{pmatrix} -1 & 5 \\ 0 & -6 \end{pmatrix} \quad BA = \begin{pmatrix} -1 & -5 \\ 0 & -6 \end{pmatrix}$$

Addirittura se è definito il prodotto $AB$ tra due matrici $A$ e $B$, il prodotto $BA$ potrebbe non essere neppure definito. Per esempio se

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

si ha

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

mentre $BA$ non è definito.

## 1.3 MATRICI E SISTEMI LINEARI

Vediamo ora come sia possibile utilizzare le matrici e il prodotto righe per colonne per descrivere un sistema lineare.

Consideriamo un sistema lineare della forma:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \qquad\qquad \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{cases}$$

Possiamo scrivere questo sistema in forma matriciale nel seguente modo:

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

quindi il sistema può essere riscritto utilizzando il prodotto righe per colonne, in questo modo:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

o, più sinteticamente, come

$$A\underline{x} = \underline{b}$$

dove $A = (a_{ij})$ è la matrice $m \times n$ dei coefficienti delle incognite, $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ è la colonna delle $n$ incognite, e $\underline{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$ è la colonna degli $m$ termini noti. La matrice $A = (a_{ij})$ si chiama matrice *incompleta* associata al sistema, la matrice

$$(A|\underline{b}) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right)$$

si chiama matrice *completa* associata al sistema.

---

**Esempio 1.3.1**

Consideriamo il sistema lineare

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

nelle incognite $x_1, x_2, x_3$. Allora la matrice incompleta e la matrice completa associate al sistema sono, rispettivamente:

$$A = \begin{pmatrix} 2 & \sqrt{2} & -1 \\ 1 & 0 & -1 \end{pmatrix} \text{ e } (A|\underline{b}) = \left(\begin{array}{ccc|c} 2 & \sqrt{2} & -1 & 2 \\ 1 & 0 & -1 & 1 \end{array}\right)$$

---

Quello di usare le matrici è semplicemente un modo più comodo di scrivere e trattare i sistemi lineari. Ogni riga della matrice completa associata a un sistema lineare equivale a un'equazione del sistema in cui vengono sottointese le incognite e i segni di somma e uguaglianza.

**Definizione 1.3.2** Una matrice si dice *in forma a scala* (per righe) o, semplicemente, *a scala* se sono soddisfatte le seguenti condizioni:

(a) eventuali righe nulle si trovano in fondo alla matrice;

**Dimostrazione** – Le dimostrazioni sono un puro calcolo di applicazione della definizione. Mostreremo solo l'associatività del prodotto. Per comodità, indicheremo l'elemento di posto $(i,j)$ di una generica matrice $M$ con $(M)_{ij}$. Siano $A \in \mathrm{M}_{m,s}(\mathbb{R})$, $B \in \mathrm{M}_{s,r}(\mathbb{R})$, $C \in \mathrm{M}_{r,n}(\mathbb{R})$. Osserviamo che:

$$(AB)_{iu} = \sum_{h=1}^{s} a_{ih}b_{hu}, \quad (BC)_{hj} = \sum_{u=1}^{r} b_{hu}c_{uj}$$

quindi:

$$\begin{aligned}((AB)C)_{ij} &= \sum_{u=1}^{r}(AB)_{iu}c_{uj} = \sum_{u=1}^{r}\left(\sum_{h=1}^{s} a_{ih}b_{hu}\right)c_{uj} \\ &= \sum_{u=1}^{r}\sum_{h=1}^{s} a_{ih}b_{hu}c_{uj} = \sum_{h=1}^{s}\sum_{u=1}^{r} a_{ih}b_{hu}c_{uj} \\ &= \sum_{h=1}^{s} a_{ih}\left(\sum_{u=1}^{r} b_{hu}c_{uj}\right) = \sum_{h=1}^{s} a_{ih}(BC)_{hj} = (A(BC))_{ij}\end{aligned}$$

La dimostrazione della distributività è analoga. □

Si noti che l'operazione di prodotto tra matrici non gode della proprietà *commutativa*. Per esempio se

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

si ha che

$$AB = \begin{pmatrix} -1 & 5 \\ 0 & -6 \end{pmatrix} \quad BA = \begin{pmatrix} -1 & -5 \\ 0 & -6 \end{pmatrix}$$

Addirittura se è definito il prodotto $AB$ tra due matrici $A$ e $B$, il prodotto $BA$ potrebbe non essere neppure definito. Per esempio se

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

si ha

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

mentre $BA$ non è definito.

## 1.3 MATRICI E SISTEMI LINEARI

Vediamo ora come sia possibile utilizzare le matrici e il prodotto righe per colonne per descrivere un sistema lineare.

Consideriamo un sistema lineare della forma:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{cases}$$

Possiamo scrivere questo sistema in forma matriciale nel seguente modo:

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

quindi il sistema può essere riscritto utilizzando il prodotto righe per colonne, in questo modo:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

o, più sinteticamente, come

$$A\underline{x} = \underline{b}$$

dove $A = (a_{ij})$ è la matrice $m \times n$ dei coefficienti delle incognite, $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ è la colonna delle $n$ incognite, e $\underline{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$ è la colonna degli $m$ termini noti. La matrice $A = (a_{ij})$ si chiama matrice *incompleta* associata al sistema, la matrice

$$(A|\underline{b}) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right)$$

si chiama matrice *completa* associata al sistema.

---

**Esempio 1.3.1**

Consideriamo il sistema lineare

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

nelle incognite $x_1, x_2, x_3$. Allora la matrice incompleta e la matrice completa associate al sistema sono, rispettivamente:

$$A = \begin{pmatrix} 2 & \sqrt{2} & -1 \\ 1 & 0 & -1 \end{pmatrix} \text{ e } (A|\underline{b}) = \left(\begin{array}{ccc|c} 2 & \sqrt{2} & -1 & 2 \\ 1 & 0 & -1 & 1 \end{array}\right)$$

---

Quello di usare le matrici è semplicemente un modo più comodo di scrivere e trattare i sistemi lineari. Ogni riga della matrice completa associata a un sistema lineare equivale a un'equazione del sistema in cui vengono sottointese le incognite e i segni di somma e uguaglianza.

**Definizione 1.3.2** Una matrice si dice *in forma a scala* (per righe) o, semplicemente, *a scala* se sono soddisfatte le seguenti condizioni:

(a) eventuali righe nulle si trovano in fondo alla matrice;

(b) il primo elemento non nullo di ogni riga (non nulla) si trova più a destra del primo elemento non nullo della riga precedente.

**Esempio 1.3.3**

La matrice

$$A = \begin{pmatrix} 1 & -1 & -1 & 2 & -4 \\ 0 & 0 & -1 & 3 & 5 \\ 0 & 0 & 0 & \frac{1}{3} & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

è una matrice in forma a scala (per righe) perché soddisfa le condizioni (a) e (b) della Definizione 1.3.2. Al contrario, la matrice

$$B = \begin{pmatrix} 2 & -1 & -1 & 2 & -4 \\ 0 & 1 & -1 & 3 & 5 \\ 0 & 2 & 0 & 1 & \frac{1}{5} \end{pmatrix}$$

non è in forma a scala perché il primo elemento non nullo della terza riga non si trova più a destra del primo elemento non nullo della seconda riga (ma sotto di esso).

**Definizione 1.3.4** Sia $A$ una matrice in forma a scala (per righe). Si chiama *pivot* di $A$ il primo elemento non nullo di ogni riga (non nulla) di $A$. Si chiama *rango righe* di $A$ e si indica con $\mathrm{rr}(A)$ il numero di righe non nulle di $A$ o, equivalentemente, il numero dei suoi pivot.

**Esempio 1.3.5**

Data

$$A = \begin{pmatrix} 1 & -1 & -1 & 2 & -4 \\ 0 & 0 & -1 & 3 & 5 \\ 0 & 0 & 0 & \frac{1}{3} & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

i pivot di $A$ sono $1, -1, \frac{1}{3}$, perciò $\mathrm{rr}(A) = 3$.

♦ **Osservazione 1.3.6** Sia $A \in \mathrm{M}_{m,n}(\mathbb{R})$ una matrice a scala.

Per definizione di rango si ha

$$\mathrm{rr}(A) \leq m \tag{1.3}$$

Vale però anche la disuguaglianza

$$\mathrm{rr}(A) \leq n \tag{1.4}$$

Se $m \leq n$, (1.4) segue ovviamente da (1.3). Se $m > n$ è facile rendersi conto, disegnando una matrice a scala con un numero $m$ di righe maggiore del numero $n$ di colonne, che la proprietà (b) della Definizione 1.3.2 implica che il massimo numero di pivot di $A$ è $n$.

**Definizione 1.3.7** Il sistema lineare $A\underline{x} = \underline{b}$ si dice *a scala* se la matrice $A$ è in forma a scala.

Mostreremo di seguito come risolvere velocemente un sistema lineare a scala.

**Esempi 1.3.8**

Consideriamo il seguente sistema lineare nelle incognite $x_1, x_2, x_3, x_4$:

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

La matrice completa associata al sistema è:

$$(A|\underline{b}) = \left(\begin{array}{cccc|c} 4 & 2 & 3 & 4 & 1 \\ 0 & 1 & -2 & 0 & 2 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right)$$

che è in forma a scala e ha rango 4. Ovviamente anche la matrice incompleta $A$ è in forma a scala e notiamo che anch'essa ha rango 4. Il fatto che la matrice $A$ sia in forma a scala indica che in ogni equazione del sistema compare un'incognita che non compare nelle equazioni successive. Il sistema lineare può dunque essere facilmente risolto per sostituzioni successive dal basso, cioè a partire dall'ultima equazione e risalendo verso la prima: dalla quarta equazione abbiamo $x_4 = 1$; sostituendo $x_4 = 1$ nella terza equazione otteniamo $x_3 = x_4 = 1$. Sostituendo $x_3 = 1$ nella seconda equazione otteniamo $x_2 = 2 + 2x_3 = 2 + 2 = 4$. Infine, sostituendo $x_2 = 4$ e $x_3 = x_4 = 1$ nella prima equazione, otteniamo $x_1 = \frac{1}{4}(1 - 2x_2 - 3x_3 - 4x_4) = \frac{1}{4}(1 - 8 - 3 - 4) = -\frac{7}{2}$. Il sistema assegnato ha dunque una sola soluzione: $(-\frac{7}{2}, 4, 1, 1)$.

**Esempio 1.3.9**

Consideriamo il sistema lineare nelle incognite $x_1, x_2, x_3, x_4$ ottenuto da quello dell'esempio precedente cancellando l'ultima equazione:

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

La matrice completa associata al sistema è:

$$(A|\underline{b}) = \left(\begin{array}{cccc|c} 4 & 2 & 3 & 4 & 1 \\ 0 & 1 & -2 & 0 & 2 \\ 0 & 0 & 1 & -1 & 0 \end{array}\right)$$

che è in forma a scala e ha rango 3. Anche la matrice incompleta $A$ è in forma a scala e anch'essa ha rango 3. Naturalmente la soluzione $(-\frac{7}{2}, 4, 1, 1)$, trovata nell'esempio precedente, è una soluzione anche di questo sistema, che quindi è senz'altro compatibile. Quante sono, tuttavia, in questo caso le soluzioni del sistema? Anche in questo caso possiamo procedere per sostituzioni successive dal basso perché, come prima, in ogni equazione compare un'incognita che non compare nelle equazioni successive. Dall'ultima equazione abbiamo $x_3 = x_4$. Sostituendo $x_3 = x_4$ nella seconda equazione, otteniamo $x_2 = 2 + 2x_3 = 2 + 2x_4$. Sostituendo $x_2$ e $x_3$ nella prima equazione otteniamo $x_1 = \frac{1}{4}(1 - 2x_2 - 3x_3 - 4x_4) = \frac{1}{4}(1 - 4 - 4x_4 - 3x_4 - 4x_4) = \frac{1}{4}(-3 - 11x_4)$. Il sistema ha dunque infinite soluzioni della forma $\frac{1}{4}(-3 - 11x_4), 2 + 2x_4, x_4, x_4)$ al variare di $x_4$ nell'insieme dei numeri reali.

Quanto illustrato negli Esempi 1.3.8, 1.3.9 è un fatto del tutto generale. Vale infatti la seguente proposizione:

**Proposizione 1.3.10** *Sia $A\underline{x} = \underline{b}$ un sistema lineare a scala nelle $n$ incognite $x_1, \dots, x_n$. Allora:*

(a) *il sistema ammette soluzioni se e solo se* $\mathrm{rr}(A) = \mathrm{rr}(A|\underline{b})$;

(b) *se* $\mathrm{rr}(A) = \mathrm{rr}(A|\underline{b}) = n$ *il sistema ammette una sola soluzione;*

(c) *se* $\mathrm{rr}(A) = \mathrm{rr}(A|\underline{b}) = k < n$ *il sistema ammette infinite soluzioni, che dipendono da $n-k$ variabili libere.*

**Dimostrazione** – Osserviamo innanzitutto che, cancellando la colonna $\underline{b}$ dalla matrice $(A|\underline{b})$, si ottiene ancora una matrice in forma a scala, quindi anche la matrice incompleta $A$ associata al sistema è una matrice a scala. Inoltre, cancellando la colonna $(\underline{b})$ dalla matrice $(A|\underline{b})$, il numero di pivot può diminuire al più di 1. Più precisamente questo succede se e soltanto se la matrice $A$ ha almeno una riga nulla, diciamo la $i$-esima, e l'elemento $b_i$ è diverso da 0. In termini di equazioni questo equivale alla condizione $0 = b_i \neq 0$ che, evidentemente, non può essere soddisfatta. Pertanto se $\mathrm{rr}(A) \neq \mathrm{rr}(A|\underline{b})$, il sistema non ammette soluzioni. Supponiamo ora che sia $\mathrm{rr}(A) = \mathrm{rr}(A|\underline{b}) = n$. Questo significa che il numero dei pivot, ossia il numero degli "scalini", coincide con il numero delle incognite, quindi il sistema è costituito esattamente da $n$ equazioni, l'incognita $x_1$ compare solo nella prima equazione, $x_2$ solo nelle prime due equazioni, $x_3$ solo nelle prime tre e così via. In particolare l'ultima equazione del sistema contiene solo l'incognita $x_n$ e quindi ne fissa il valore. Sostituendo tale valore nella penultima equazione si ottiene l'unico valore della variabile $x_{n-1}$ e così via, procedendo per sostituzioni successive dal basso come nell'Esempio 1.3.8, si ottiene l'unica soluzione del sistema. Se, invece, $\mathrm{rr}(A) = \mathrm{rr}(A|\underline{b}) = k < n$ è possibile, procedendo per sostituzioni successive dal basso, esprimere le $k$ variabili corrispondenti ai pivot delle righe non nulle in funzione delle altre $n-k$ che restano libere di variare nell'insieme dei numeri reali. Si ottengono così infinite soluzioni. □

---

**Esempio 1.3.11**

Risolviamo il seguente sistema lineare a scala nelle incognite $x_1, x_2, x_3, x_4$:

$$\begin{cases} x_1 - x_2 + x_3 - x_4 = 1 \\ x_3 + \frac{1}{2}x_4 = 0 \end{cases}$$

La matrice completa associata al sistema è:

$$(A|\underline{b}) = \left(\begin{array}{cccc|c} 1 & -1 & 1 & -1 & 1 \\ 0 & 0 & 1 & \frac{1}{2} & 0 \end{array}\right)$$

Notiamo che $\mathrm{rr}(A) = \mathrm{rr}(A|\underline{b}) = 2$ quindi, per la proprietà (a) della Proposizione 1.3.10, il sistema ammette soluzioni. Dal momento che il numero delle variabili è $4 > 2$, per la proprietà (c) della Proposizione 1.3.10, il sistema ammette infinite soluzioni. In sostanza abbiamo 4 variabili e due condizioni su di esse, perciò due variabili restano libere di variare nell'insieme dei numeri reali e potremo esprimere due variabili in funzione delle altre due. Procedendo per sostituzioni successive dal basso abbiamo:

$$x_3 = -\frac{1}{2}x_4$$

$$x_1 = x_2 - x_3 + x_4 + 1 = x_2 + \frac{1}{2}x_4 + x_4 + 1 = x_2 + \frac{3}{2}x_4 + 1$$

Le infinite soluzioni del sistema sono pertanto della forma $(x_2 + \frac{3}{2}x_4 + 1, x_2, -\frac{1}{2}x_4, x_4)$, con $x_2, x_4 \in \mathbb{R}$.

Osserviamo che avremmo potuto decidere di esprimere la variabile $x_4$ in funzione della variabile $x_3$ ($x_4 = -2x_3$) e, per esempio, la variabile $x_2$ in funzione delle variabili $x_1$ e $x_3$ ($x_2 = x_1 + 3x_3 - 1$). In altri termini, la scelta delle variabili "libere" non è obbligata. Tuttavia è sempre possibile scegliere come variabili libere quelle corrispondenti alle colonne della matrice $A$ non contenenti pivot ed esprimere in funzione di queste le incognite corrispondenti alle colonne contenenti i pivot. Per esempio, in questo caso i pivot, entrambi uguali a 1, si trovano sulla prima e sulla terza colonna di $A$ e nella nostra prima scelta, la più naturale, abbiamo lasciato libere le variabili $x_2$ e $x_4$ ed espresso $x_1$ e $x_3$ in funzione di $x_2$ e $x_4$.

## 1.4 ALGORITMO DI GAUSS

Abbiamo stabilito come risolvere un sistema lineare a scala. Cosa succede nel caso di un qualsiasi sistema lineare $A\underline{x} = \underline{b}$? Sarebbe comodo poter ottenere un nuovo sistema lineare $A'\underline{x} = \underline{b}'$, questa volta a scala, equivalente al sistema di partenza, cioè avente le stesse soluzioni, in modo tale da poter calcolare le soluzioni di $A\underline{x} = \underline{b}$ risolvendo il sistema a scala $A'\underline{x} = \underline{b}'$. Questo è esattamente quello che faremo.

**Esempio 1.4.1**

I seguenti sistemi lineari nelle incognite $x_1, x_2$, sono equivalenti:

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

Si verifica, infatti, facilmente che l'unica soluzione di ciascuno dei due sistemi è: $(\frac{3}{2}, \frac{1}{2})$. Notiamo che la prima equazione è la stessa nei due sistemi e che la seconda equazione del secondo sistema può essere ottenuta dalla differenza tra la seconda e la prima equazione del primo sistema:

$$2^a\text{equazione} \to 2^a\text{equazione} - 1^a\text{equazione}.$$

Come si può passare da un sistema a uno a esso equivalente? Per esempio eseguendo le seguenti operazioni:

(a) scambio di due equazioni;

(b) moltiplicazione di un'equazione per un numero reale diverso da 0;

(c) sostituzione dell'equazione $i$-esima con la somma dell'equazione $i$-esima e della $j$-esima moltiplicata per un numero reale $\alpha$ qualsiasi. In sintesi:

$$i\text{-esima equazione} \longrightarrow i\text{-esima equazione} + \alpha(j\text{-esima equazione}).$$

È immediato verificare che le operazioni (a) e (b) non alterano le soluzioni del sistema. L'operazione (c) coinvolge soltanto la $i$-esima e la $j$-esima equazione del sistema, quindi basta osservare che i sistemi

$$\begin{cases} j\text{-esima equazione} \\ i\text{-esima equazione} \end{cases} \qquad \begin{cases} j\text{-esima equazione} \\ i\text{-esima equazione} + \alpha(j\text{-esima equazione}) \end{cases}$$

sono equivalenti, cioè hanno le stesse soluzioni.

Traduciamo ora le operazioni (a), (b) e (c) in termini matriciali: scambiare due equazioni del sistema equivale a scambiare due righe della matrice completa associata al sistema; moltiplicare un'equazione per un numero reale diverso da 0 equivale a moltiplicare una riga della matrice completa associata al sistema per un numero reale diverso da 0, cioè moltiplicare per tale numero ogni elemento della riga; infine l'operazione (c) equivale a sostituire la riga $i$-esima della matrice completa associata al sistema con la somma della riga $i$-esima e della $j$-esima moltiplicata per un numero reale $\alpha$. Spieghiamo un po' meglio che cosa si intende con tale somma: siano $(a_{i1} \ldots a_{in}\, b_i)$ e $(a_{j1} \ldots a_{jn}\, b_j)$ rispettivamente la $i$-esima e la $j$-esima riga della matrice $(A|\underline{b})$. Sommare la $i$-esima riga con la $j$-esima moltiplicata per un numero $\alpha$, significa effettuare la somma coefficiente per coefficiente:

$$(a_{i1} \ldots a_{in}\; b_i) + \alpha(a_{j1} \ldots a_{jn}\; b_j) = (a_{i1} + \alpha a_{j1} \ldots a_{in} + \alpha a_{jn}\; b_i + \alpha b_j)$$

In virtù dell'importanza che tali operazioni avranno in seguito, diamo ad esse un nome:

**Definizione 1.4.2** Data una matrice $A$ si chiamano *operazioni elementari* sulle righe di $A$ le seguenti operazioni:

(a) scambio di due righe;

(b) moltiplicazione di una riga per un numero reale diverso da 0;

(c) sostituzione della riga $i$-esima con la somma della riga $i$-esima e della $j$-esima moltiplicata per un numero reale $\alpha$ qualsiasi.

♦ **Osservazione 1.4.3** Osserviamo che nell'operazione elementare (c) non richiediamo che il numero $\alpha$ sia non nullo. In effetti se $\alpha = 0$ l'operazione (c) equivale a lasciare la riga $i$-esima inalterata.

Data una qualsiasi matrice $A = (a_{ij})$ è possibile trasformare $A$ in una matrice a scala attraverso operazioni elementari sulle righe di $A$. Tale procedimento è noto come *riduzione di Gauss* e l'algoritmo che si utilizza si chiama *algoritmo di Gauss* e funziona nel modo seguente:

1) Se il primo elemento di ogni riga di $A$ è nullo, si considera la matrice ottenuta cancellando la prima colonna della matrice in esame e si ricomincia dal principio del punto 1). Altrimenti, se $a_{11} = 0$ si scambia la prima riga di $A$ con una riga in cui il primo elemento è non nullo. Indichiamo con $a$ tale elemento non nullo.

2) Si controllano una dopo l'altra tutte le righe tranne la prima. Se il primo elemento di una riga è nullo si lascia quella riga inalterata. Se il primo elemento di una riga, diciamo la $i$-esima ($i > 1$), è uguale a $b \neq 0$, si sostituisce la riga $i$-esima con la somma della riga $i$-esima e della prima riga moltiplicata per $-\frac{b}{a}$.

3) A questo punto tutti gli elementi della prima colonna, tranne eventualmente il primo, sono nulli. Si considera dunque la matrice che si ottiene cancellando la prima riga e la prima colonna della matrice ottenuta e si ricomincia dal punto 1.

---

**Esempio 1.4.4**

Sia

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

Utilizziamo l'algoritmo di Gauss per ridurre $A$ a scala.

Dal momento che l'elemento di posto $(1,1)$ è nullo, scambiamo la prima con la seconda riga, ottenendo la matrice:

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

Il primo elemento della seconda riga della matrice ottenuta è 0, perciò lasciamo questa riga inalterata. Il primo elemento della terza riga, invece, è 2, quindi sostituiamo la terza riga con la somma della terza riga e della prima moltiplicata per $-2$. Otteniamo così la matrice:

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

Ora ogni elemento della prima colonna tranne il primo è uguale a 0. Passiamo a considerare la matrice che otteniamo cancellando la prima riga e la prima colonna della matrice ottenuta:

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

Applichiamo un'altra volta l'algoritmo di Gauss: il primo elemento della prima riga della matrice rimasta, questa volta, è diverso da 0, perciò lasciamo inalterata la prima riga. Ora sostituiamo la seconda riga con la somma della seconda e della prima moltiplicata per 5, ottenendo:

$$\begin{pmatrix} \cancel{1} & \cancel{2} & \cancel{0} & \cancel{1} \\ \cancel{0} & 1 & -1 & 0 \\ \cancel{0} & -5 & 1 & 0 \end{pmatrix} \to \begin{pmatrix} \cancel{1} & \cancel{2} & \cancel{0} & \cancel{1} \\ \cancel{0} & 1 & -1 & 0 \\ \cancel{0} & 0 & -4 & 0 \end{pmatrix}$$

Abbiamo così ottenuto la matrice a scala:

$$B = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -4 & 0 \end{pmatrix}$$

---

A questo punto siamo in grado di risolvere qualsiasi sistema lineare $A\underline{x} = \underline{b}$. La matrice completa associata al sistema è $(A|\underline{b})$. Utilizzando l'algoritmo di Gauss possiamo "ridurre" $(A|\underline{b})$ a scala ottenendo una matrice $(A'|\underline{b}')$. Il sistema lineare $A'\underline{x} = \underline{b}'$ è equivalente al sistema $A\underline{x} = \underline{b}$ dal momento che ogni operazione elementare sulle righe di $(A|\underline{b})$ equivale a un'operazione sulle equazioni del sistema che ne preserva le soluzioni. Quindi, per trovare le soluzioni del sistema di partenza, risolveremo il sistema a scala $A'\underline{x} = \underline{b}'$ tenendo conto della Proposizione 1.3.10. Notiamo in particolare che, in conseguenza del ragionamento appena illustrato e del contenuto della Proposizione 1.3.10, dato un qualsiasi sistema lineare $A\underline{x} = \underline{b}$ a coefficienti reali *solo una* delle seguenti situazioni si può presentare:

1) il sistema *non* ha soluzioni;

2) il sistema ha *una sola* soluzione;

3) il sistema ha *infinite* soluzioni.

Questo significa che non esiste alcun sistema lineare a coefficienti reali con un numero finito di soluzioni strettamente più grande di 1. Nel momento in cui un sistema lineare a coefficienti reali ha 2 soluzioni allora ne ha infinite.

♦ **Osservazione 1.4.5** Le mosse dell'algoritmo di Gauss non sono necessariamente obbligate. Nell'Esempio 1.4.4, Per esempio, anziché scambiare la prima con la seconda riga, avremmo potuto scambiare la prima con la terza riga. In questo modo, portando a termine l'algoritmo, avremmo ottenuto una matrice a scala diversa dalla matrice $B$. Dal punto di vista dei sistemi lineari questo significa ottenere sistemi a scala diversi, ma tutti equivalenti al sistema di partenza (e quindi equivalenti tra loro).

---

**Esempio 1.4.6**

Risolvere il seguente sistema lineare di quattro equazioni nelle cinque incognite $u, v, w, x, y$:

$$\begin{cases} u + 2v + 3w + x + y = 4 \\ u + 2v + 3w + 2x + 3y = -2 \\ u + v + w + x + y = -2 \\ -3u - 5v - 7w - 4x - 5y = 0 \end{cases}$$

La matrice completa associata al sistema è:

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

Si tratta ora di ridurre la matrice $(A|\underline{b})$ in forma a scala utilizzando l'algoritmo di Gauss e, successivamente, di risolvere il sistema lineare associato alla matrice ridotta.

In questo primo esempio riportiamo i passi dell'algoritmo di Gauss descrivendo contemporaneamente le operazioni sulle equazioni del sistema che equivalgono a ogni passo. Tuttavia, poiché il vantaggio dell'algoritmo di Gauss consiste proprio nel poter dimenticare equazioni e incognite concentrandosi solo sulle matrici, questa descrizione è puramente esplicativa.

L'elemento di posto $(1,1)$ è non nullo, perciò lasciamo la prima riga inalterata. Dopodiché effettuiamo le seguenti operazioni elementari sulle righe di $(A|\underline{b})$:

- $2^a$ riga $\to$ $2^a$ riga $-$ $1^a$ riga;
- $3^a$ riga $\to$ $3^a$ riga $-$ $1^a$ riga;
- $4^a$ riga $\to$ $4^a$ riga $+ 3(1^a$ riga$)$.

Otteniamo così la seguente matrice (e l'equivalente sistema lineare):

$$\left(\begin{array}{ccccc|c} 1 & 2 & 3 & 1 & 1 & 4 \\ 0 & 0 & 0 & 1 & 2 & -6 \\ 0 & -1 & -2 & 0 & 0 & -6 \\ 0 & 1 & 2 & -1 & -2 & 12 \end{array}\right) \Leftrightarrow \begin{cases} u + 2v + 3w + x + y = 4 \\ x + 2y = -6 \\ -v - 2w = -6 \\ v + 2w - x - 2y = 12 \end{cases}$$

Ora scambiamo la seconda con la quarta riga:

$$\left(\begin{array}{ccccc|c} 1 & 2 & 3 & 1 & 1 & 4 \\ 0 & 1 & 2 & -1 & -2 & 12 \\ 0 & -1 & -2 & 0 & 0 & -6 \\ 0 & 0 & 0 & 1 & 2 & -6 \end{array}\right) \Leftrightarrow \begin{cases} u + 2v + 3w + x + y = 4 \\ v + 2w - x - 2y = 12 \\ -v - 2w = -6 \\ x + 2y = -6 \end{cases}$$

Ora sostituiamo alla terza riga la somma della terza riga e della seconda:

$$\left(\begin{array}{ccccc|c} 1 & 2 & 3 & 1 & 1 & 4 \\ 0 & 1 & 2 & -1 & -2 & 12 \\ 0 & 0 & 0 & -1 & -2 & 6 \\ 0 & 0 & 0 & 1 & 2 & -6 \end{array}\right) \Leftrightarrow \begin{cases} u + 2v + 3w + x + y = 4 \\ v + 2w - x - 2y = 12 \\ -x - 2y = 6 \\ x + 2y = -6 \end{cases}$$

Infine sostituiamo alla quarta riga la somma della quarta e della terza:

$$\left(\begin{array}{ccccc|c} 1 & 2 & 3 & 1 & 1 & 4 \\ 0 & 1 & 2 & -1 & -2 & 12 \\ 0 & 0 & 0 & -1 & -2 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \Leftrightarrow \begin{cases} u + 2v + 3w + x + y = 4 \\ v + 2w - x - 2y = 12 \\ -x - 2y = 6 \\ 0 = 0 \end{cases}$$

Il sistema di partenza è equivalente al sistema a scala che abbiamo ottenuto, in cui l'ultima equazione è diventata un'identità. Il rango della matrice incompleta e il rango della matrice completa della matrice a scala ottenuta coincidono e sono uguali a 3. Il numero delle incognite del sistema è 5, quindi il sistema ammette infinite soluzioni che dipenderanno da $5 - 3 = 2$ variabili libere. Risolviamo il sistema per sostituzioni successive dal basso: usando la terza equazione possiamo esprimere la variabile $x$ in funzione di $y$:

$$x = -2y - 6$$

Nella seconda equazione sostituiamo $x$ con la sua espressione in funzione di $y$ e ricaviamo $v$ in funzione di $w$ e di $y$:

$$v = -2w + 6$$

Infine nella prima equazione sostituiamo $x$ con la sua espressione in funzione di $y$, $v$ con la sua espressione in funzione di $w$ e ricaviamo $u$ in funzione di $w$ e di $y$:

$$u = -2v - 3w - x - y + 4 = -2(-2w + 6) - 3w - (-2y - 6) - y + 4 = w + y - 2$$

Dunque il sistema ha infinite soluzioni del tipo $(w + y - 2, -2w + 6, w, -2y - 6, y)$ che dipendono da due variabili libere $w, y \in \mathbb{R}$.

## 1.5 ESERCIZI SVOLTI

### Esercizio 1.5.1

*Si risolva il seguente sistema lineare nelle quattro incognite $x, y, z, t$:*

$$\begin{cases} x - 2y = 5 \\ -x + 2y - 3z = -2 \\ -2y + 3z - 4t = -11 \\ -3z + 4t = 15 \end{cases}$$

**Svolgimento**

La matrice completa associata al sistema è:

$$(A|\underline{b}) = \left(\begin{array}{rrrr|r} 1 & -2 & 0 & 0 & 5 \\ -1 & 2 & -3 & 0 & -2 \\ 0 & -2 & 3 & -4 & -11 \\ 0 & 0 & -3 & 4 & 15 \end{array}\right)$$

Riduciamo la matrice $(A|\underline{b})$ in forma a scala utilizzando l'algoritmo di Gauss:

$$\left(\begin{array}{rrrr|r} 1 & -2 & 0 & 0 & 5 \\ 0 & 0 & -3 & 0 & 3 \\ 0 & -2 & 3 & -4 & -11 \\ 0 & 0 & -3 & 4 & 15 \end{array}\right) \to \left(\begin{array}{rrrr|r} 1 & -2 & 0 & 0 & 5 \\ 0 & -2 & 3 & -4 & -11 \\ 0 & 0 & -3 & 0 & 3 \\ 0 & 0 & -3 & 4 & 15 \end{array}\right)$$

$$\to \left(\begin{array}{rrrr|r} 1 & -2 & 0 & 0 & 5 \\ 0 & -2 & 3 & -4 & -11 \\ 0 & 0 & -3 & 0 & 3 \\ 0 & 0 & 0 & 4 & 12 \end{array}\right) = (A'|\underline{b}')$$

La matrice a scala ottenuta è la matrice completa associata al sistema lineare:

$$\begin{cases} x - 2y = 5 \\ -2y + 3z - 4t = -11 \\ -3z = 3 \\ 4t = 12 \end{cases}$$

Notiamo che $\text{rr}(A') = \text{rr}(A'|\underline{b}') = 4 =$ numero incognite. Il sistema di partenza ammette dunque un'unica soluzione che possiamo calcolare procedendo per sostituzioni successive dal basso: dalla quarta equazione abbiamo

$$t = 3$$

e dalla terza equazione abbiamo

$$z = -1$$

sostituendo questi valori di $t$ e di $z$ nella seconda equazione otteniamo

$$y = -2$$

infine, sostituendo i valori di $t, z, y$ nella prima equazione otteniamo

$$x = 1$$

Dunque il sistema ha come unica soluzione la quaterna $(1, -2, -1, 3)$. ■

### Esercizio 1.5.2

*Si determinino le soluzioni del seguente sistema lineare nelle incognite $x$, $y$, $z$, $t$, al variare del parametro reale $\alpha$:*

$$\begin{cases} x + y + z + t = 0 \\ x - z - t = -1 \\ x + 2y + (2\alpha + 1)z + 3t = 2\alpha - 1 \\ 3x + 4y + (3\alpha + 2)z + (\alpha + 5)t = 3\alpha - 1 \end{cases}$$

**Svolgimento**

In questo esercizio si ha a che fare con un sistema lineare in cui compare un parametro reale $\alpha$. Questo significa che al variare di $\alpha \in \mathbb{R}$ si ottengono infiniti sistemi lineari diversi che noi risolveremo trattandoli il più possibile come un solo sistema lineare. Il modo di procedere resta sempre lo stesso, come se il parametro fosse un numero reale fissato. Per prima cosa, dunque, scriviamo la matrice completa associata al sistema:

$$(A|\underline{b}) = \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & -1 & -1 & -1 \\ 1 & 2 & 2\alpha+1 & 3 & 2\alpha-1 \\ 3 & 4 & 3\alpha+2 & \alpha+5 & 3\alpha-1 \end{array}\right)$$

e riduciamola in forma a scala mediante l'algoritmo di Gauss:

$$(A|\underline{b}) \to \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & -1 & -2 & -2 & -1 \\ 0 & 1 & 2\alpha & 2 & 2\alpha-1 \\ 0 & 1 & 3\alpha-1 & \alpha+2 & 3\alpha-1 \end{array}\right)$$

$$\to \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & -1 & -2 & -2 & -1 \\ 0 & 0 & 2\alpha-2 & 0 & 2\alpha-2 \\ 0 & 0 & 3\alpha-3 & \alpha & 3\alpha-2 \end{array}\right)$$

$$\to \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & -1 & -2 & -2 & -1 \\ 0 & 0 & \alpha-1 & 0 & \alpha-1 \\ 0 & 0 & 3\alpha-3 & \alpha & 3\alpha-2 \end{array}\right)$$

$$\to \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & -1 & -2 & -2 & -1 \\ 0 & 0 & \alpha-1 & 0 & \alpha-1 \\ 0 & 0 & 0 & \alpha & 1 \end{array}\right) = (A'|\underline{b}')$$

Dobbiamo ora stabilire cosa succede al variare del parametro $\alpha$ nell'insieme dei numeri reali. Dobbiamo cioè rispondere alle seguenti domande:

1) per quali valori di $\alpha$ il sistema è compatibile?

2) per i valori di $\alpha$ per cui il sistema è compatibile, quante soluzioni ammette e quali sono?

Come sappiamo la risposta viene fornita dalla proprietà (a) della Proposizione 1.3.10: dobbiamo confrontare il rango di $A'$ con il rango di $(A'|\underline{b}')$. Notiamo che questi ranghi dipendono dal valore di $\alpha$.

Più precisamente: $\operatorname{rr}(A') = \operatorname{rr}(A'|\underline{b}') = 4$ per $\alpha \neq 0, 1$. In questo caso il sistema ha un'unica soluzione che possiamo calcolare procedendo per sostituzioni successive dal basso: la soluzione è $(\frac{1}{\alpha}, -\frac{\alpha+2}{\alpha}, 1, \frac{1}{\alpha})$.

Per $\alpha = 0$ si ha

$$(A'|\underline{b}') = \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & -1 & -2 & -2 & -1 \\ 0 & 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

perciò $\operatorname{rr}(A') = 3$ e $\operatorname{rr}(A'|\underline{b}') = 4$, dunque il sistema non è risolubile.

Per $\alpha = 1$ si ha

$$(A'|\underline{b}') = \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & -1 & -2 & -2 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right)$$

perciò $\operatorname{rr}(A') = 3 = \operatorname{rr}(A'|\underline{b}')$, quindi il sistema ha infinite soluzioni dipendenti da una variabile libera. Procedendo come al solito per sostituzioni successive dal basso si ha: $(x_3, -1 - 2x_3, x_3, 1)$, $x_3 \in \mathbb{R}$. ■

### Esercizio 1.5.3

*Si consideri il sistema lineare $\Sigma_\alpha$ nelle incognite $x_1, x_2, x_3$ dipendente dal parametro reale $\alpha$:*

$$\Sigma_\alpha : \begin{cases} \alpha x_1 + (\alpha + 3)x_2 + 2\alpha x_3 = \alpha + 2 \\ \alpha x_1 + (2\alpha + 2)x_2 + 3\alpha x_3 = 2\alpha + 2 \\ 2\alpha x_1 + (\alpha + 7)x_2 + 4\alpha x_3 = 2\alpha + 4 \end{cases}$$

(a) *Si determinino le soluzioni di $\Sigma_\alpha$ al variare di $\alpha \in \mathbb{R}$.*

(b) *Si determinino le soluzioni di $\Sigma_\alpha$ interpretato ora come sistema lineare nelle 4 incognite $x_1, x_2, x_3, x_4$.*

### Svolgimento

(a) Consideriamo la matrice completa $(A|\underline{b})$ associata al sistema lineare $\Sigma_\alpha$:

$$\left(\begin{array}{ccc|c} \alpha & \alpha + 3 & 2\alpha & \alpha + 2 \\ \alpha & 2\alpha + 2 & 3\alpha & 2\alpha + 2 \\ 2\alpha & \alpha + 7 & 4\alpha & 2\alpha + 4 \end{array}\right)$$

e riduciamola a scala con l'algoritmo di Gauss:

$$\left(\begin{array}{ccc|c} \alpha & \alpha+3 & 2\alpha & \alpha+2 \\ \alpha & 2\alpha+2 & 3\alpha & 2\alpha+2 \\ 2\alpha & \alpha+7 & 4\alpha & 2\alpha+4 \end{array}\right) \to \left(\begin{array}{ccc|c} \alpha & \alpha+3 & 2\alpha & \alpha+2 \\ 0 & \alpha-1 & \alpha & \alpha \\ 0 & -\alpha+1 & 0 & 0 \end{array}\right)$$

$$\to \left(\begin{array}{ccc|c} \alpha & \alpha+3 & 2\alpha & \alpha+2 \\ 0 & \alpha-1 & \alpha & \alpha \\ 0 & 0 & \alpha & \alpha \end{array}\right) = (A'|\underline{b}')$$

Se $\alpha \neq 0$ e $\alpha - 1 \neq 0$, cioè per ogni $\alpha \in \mathbb{R}\setminus\{0,1\}$, si ha $\mathrm{rr}(A') = \mathrm{rr}(A'|b') = 3$ quindi, per la Proposizione 1.3.10, il sistema $\Sigma_\alpha$ ammette un'unica soluzione: $(\frac{2-\alpha}{\alpha}, 0, 1)$;
se $\alpha = 0$ otteniamo la matrice:

$$(A'|\underline{b}') = \left(\begin{array}{ccc|c} 0 & 3 & 0 & 2 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

che non è in forma a scala ma può essere ridotta a scala sostituendo la seconda riga con la somma della seconda riga e della prima moltiplicata per $\frac{1}{3}$:

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

Il sistema di partenza risulta dunque equivalente al sistema lineare

$$\begin{cases} 3x_2 = 2 \\ 0 = 2/3 \end{cases}$$

che, ovviamente, non ha soluzioni;
infine, se $\alpha = 1$ si ha:

$$(A'|\underline{b}') = \left(\begin{array}{ccc|c} 1 & 4 & 2 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right)$$

che possiamo ridurre in forma a scala ottenendo la matrice

$$(A''|\underline{b}'') = \left(\begin{array}{ccc|c} 1 & 4 & 2 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Si ha che $\mathrm{rr}(A'') = \mathrm{rr}(A''|\underline{b}'') = 2$ perciò il sistema $\Sigma_1$ è equivalente al sistema lineare di 2 equazioni in 3 incognite

$$\begin{cases} x_1 + 4x_2 + 2x_3 = 3 \\ x_3 = 1 \end{cases}$$

Tale sistema ha infinite soluzioni dipendenti da una variabile e l'insieme delle soluzioni risulta: $\{(1 - 4x_2, x_2, 1) \mid x_2 \in \mathbb{R}\}$.

(b) Aggiungere l'incognita $x_4$ significa aggiungere alla matrice completa $(A|\underline{b})$ associata al sistema una colonna di zeri corrispondenti ai coefficienti di $x_4$. Pertanto, riducendo $(A|\underline{b})$ in forma a scala, si otterrà la matrice:

$$(A'|\underline{b}') = \left(\begin{array}{cccc|c} \alpha & \alpha+3 & 2\alpha & 0 & \alpha+2 \\ 0 & \alpha-1 & \alpha & 0 & \alpha \\ 0 & 0 & \alpha & 0 & \alpha \end{array}\right)$$

Quindi, ragionando come sopra ma tenendo conto che in questo caso il numero di variabili è 4, si ottiene che:

per $\alpha \in \mathbb{R}\setminus\{0,1\}$ il sistema ha infinite soluzioni, tutte della forma $\left(\frac{2-\alpha}{\alpha}, 0, 1, x_4\right)$ con $x_4 \in \mathbb{R}$;

per $\alpha = 0$ il sistema non ha soluzioni;

per $\alpha = 1$ il sistema ha infinite soluzioni, tutte della forma $(1-4x_2, x_2, 1, x_4)$, con $x_2, x_4 \in \mathbb{R}$. ■

**Esercizio 1.5.4**

*Si stabilisca se esistono valori del parametro reale $k$ tali che il sistema lineare*

$$\Sigma : \begin{cases} 2x_1 + x_2 - x_3 = 0 \\ 4x_1 - x_2 = 0 \\ x_1 + \frac{1}{2}x_2 - x_3 = -\frac{3}{2} \end{cases}$$

*sia equivalente al sistema lineare*

$$\Pi_k : \begin{cases} x_1 + x_2 - \frac{1}{2}x_3 = 1 \\ 2x_1 - x_2 + x_3 = 2 \\ kx_1 - 4x_2 + 3x_3 = k \end{cases}$$

*(entrambi i sistemi si intendono nelle variabili $x_1$, $x_2$, $x_3$).*

**Svolgimento**

Due sistemi sono equivalenti se hanno le stesse soluzioni. Risolviamo innanzitutto il sistema lineare $\Sigma$. La matrice completa associata al sistema è:

$$(A|\underline{b}) = \left(\begin{array}{ccc|c} 2 & 1 & -1 & 0 \\ 4 & -1 & 0 & 0 \\ 1 & \frac{1}{2} & -1 & -\frac{3}{2} \end{array}\right)$$

Utilizzando l'algoritmo di Gauss possiamo ridurre $(A|\underline{b})$ in forma a scala, ottenendo la matrice

$$(A'|\underline{b}') = \left(\begin{array}{ccc|c} 2 & 1 & -1 & 0 \\ 0 & -3 & 2 & 0 \\ 0 & 0 & -\frac{1}{2} & -\frac{3}{2} \end{array}\right)$$

Abbiamo: $\text{rr}(A') = \text{rr}(A'|\underline{b}') = 3$, dunque il sistema $\Sigma$ ammette una sola soluzione che possiamo determinare procedendo per sostituzioni successive dal basso:

$$x_3 = 3; \quad x_2 = 2; \quad x_1 = \frac{1}{2}$$

L'unica soluzione del sistema $\Sigma$ è pertanto $\left(\frac{1}{2}, 2, 3\right)$.

Affinché $\Sigma$ sia equivalente a $\Pi_k$ è dunque necessario che $(\frac{1}{2}, 2, 3)$ soddisfi tutte le equazioni di $\Pi_k$. Sostituiamo dunque $x_1 = \frac{1}{2}$, $x_2 = 2$ e $x_3 = 3$ nelle equazioni di $\Pi_k$:

$$\begin{cases} \frac{1}{2} + 2 - \frac{3}{2} = 1 \\ 1 - 2 + 3 = 2 \\ \frac{k}{2} - 8 + 9 = k \end{cases}$$

Abbiamo così ottenuto due identità e la condizione necessaria $k = 2$. Pertanto $(\frac{1}{2}, 2, 3)$ è soluzione del sistema $\Pi_k$ solo se $k = 2$. Possiamo quindi affermare che per $k \neq 2$ i sistemi $\Pi_k$ e $\Sigma$ non sono equivalenti, ma non sappiamo ancora se i sistemi $\Sigma$ e $\Pi_2$ sono equivalenti. Infatti ciò accade se e solo se $(\frac{1}{2}, 2, 3)$ è l'unica soluzione anche del sistema $\Pi_2$. Consideriamo allora la matrice completa associata al sistema $\Pi_k$ per $k = 2$:

$$\left(\begin{array}{ccc|c} 1 & 1 & -\frac{1}{2} & 1 \\ 2 & -1 & 1 & 2 \\ 2 & -4 & 3 & 2 \end{array}\right)$$

Riducendo questa matrice in forma a scala otteniamo la matrice:

$$(A''|\underline{b}'') = \left(\begin{array}{ccc|c} 1 & 1 & -\frac{1}{2} & 1 \\ 0 & -3 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Si ha dunque $\operatorname{rr}(A'') = \operatorname{rr}(A''|\underline{b}'') = 2 < 3$, dunque il sistema $\Pi_2$ ha infinite soluzioni. Possiamo allora concludere che non esistono valori di $k$ tali che i sistemi $\Sigma$ e $\Pi_k$ siano tra loro equivalenti. ∎

## 1.6 ESERCIZI PROPOSTI

### Esercizio 1.6.1

*Si risolvano i seguenti sistemi lineari nelle incognite $x, y, z$:*

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

(b) $\begin{cases} x - y + 4z = 10 \\ 3x + y + 5z = 15 \\ x + 3y - 3z = 6 \end{cases}$

### Esercizio 1.6.2

*Si risolvano i seguenti sistemi lineari nelle incognite $x, y, z, w$:*

(a) $\begin{cases} x - y + 2z - 3w = 0 \\ 2x + y - w = 3 \\ 2y + z + w = -3 \\ 2x + z = 0 \end{cases}$

(b) $\begin{cases} x + y - z + w = 0 \\ 2x - z - w = 0 \\ x - y - 2w = 0 \\ 3x + y - 2z = 0 \end{cases}$

(c) $\begin{cases} x + z = 7 \\ x + y = 2 \\ 4x + 12y + z = 1 \\ 5x + 6y + 2z = -1 \end{cases}$

**Esercizio 1.6.3**

*Si discuta, al variare del parametro reale $k$, il seguente sistema lineare nelle incognite $x, y, z$:*

$$\begin{cases} x + 2y + kz = 0 \\ x + y = -1 \\ x + ky = -2 \end{cases}$$

*Si stabilisca per quali valori di $k$ il sistema ammette soluzioni e, quando possibile, si determinino.*

**Esercizio 1.6.4**

*Si dica per quali valori di $k \in \mathbb{R}$ il seguente sistema lineare ammette soluzioni:*

$$\begin{cases} x_1 + 2x_2 + 5x_4 = 0 \\ 2x_1 + 5x_2 + 11x_4 = 4 \\ x_1 + x_2 + (k - k^2)x_3 + (5 - k^2)x_4 = -k^2 - 3 \\ -x_2 + 2(k^2 - k)x_3 + (3k^2 - 4)x_4 = 2(k^2 + k - 4) \end{cases}$$

*Si risolva il sistema per i valori di $k$ per i quali la soluzione non è unica.*

**Esercizio 1.6.5**

*Si dica per quali valori di $a \in \mathbb{R}$ il seguente sistema lineare, nelle incognite $x$, $y$, $z$, $t$, ammette soluzioni e, quando possibile, si determinino tali soluzioni:*

$$\begin{cases} 2x + y - z = 1 \\ -2x + 3z + t = 1 \\ 2x + 3y + (a^2 + 2a + 3)z + (a^2 - 2)t = a + 6 \\ y + 2(a^2 + 2a + 1)z + (3a^2 - 2a - 7)t = 3a + 4 \end{cases}$$

**Esercizio 1.6.6**

*Dato il sistema lineare nelle incognite $x, y, z$:*

$$\Sigma_{a,b} : \begin{cases} x + (2 + a)y = b \\ (2 + 2a)x + 3y - (b + 1)z = 1 + b \\ bx + by - (b + 4)z = b^2 + 3b. \end{cases}$$

(a) *Si stabilisca per quali valori di $a, b \in \mathbb{R}$ il sistema omogeneo associato ammette la soluzione $(a, -a, 0)$. (Si chiama sistema omogeneo associato al sistema lineare $A\underline{x} = \underline{b}$ il sistema $A\underline{x} = \mathbf{0}$, ove $\mathbf{0}$ è la colonna nulla.)*

(b) *Si dica per quali tra i valori $a, b$ trovati al punto precedente il sistema $\Sigma_{a,b}$ è risolubile e se ne determinino le soluzioni.*

### Esercizio 1.6.7

*Si stabilisca per quali valori del parametro reale $k$ il seguente sistema lineare nelle incognite $x_1, x_2, x_3, x_4$ è compatibile. Si determinino, quando possibile, le soluzioni del sistema:*

$$\begin{cases} x_1 + 3x_2 + kx_3 + 2x_4 = k \\ x_1 + 6x_2 + kx_3 + 3x_4 = 2k + 1 \\ -x_1 - 3x_2 + (k-2)x_4 = 1 - k \\ kx_3 + (2-k)x_4 = 1 \end{cases}$$

# 2 Spazi vettoriali

In questo capitolo vogliamo introdurre il vero protagonista dell'algebra lineare: lo *spazio vettoriale*. Si tratta di una astrazione di concetti che già ben conosciamo. Il piano cartesiano, l'insieme delle funzioni studiate in analisi, l'insieme delle matrici $m \times n$ introdotte nel capitolo precedente, l'insieme dei polinomi, l'insieme dei numeri reali stessi sono tutti esempi di insiemi che hanno una struttura naturale di spazio vettoriale. Lo spazio vettoriale sarà anche l'ambiente giusto in cui leggere e interpretare i risultati ottenuti nel capitolo precedente. Prima di darne la definizione vera a propria vediamo qualche esempio concreto.

## 2.1 PREMESSA: L'INSIEME DEI NUMERI REALI

Ricordiamo brevemente le principali proprietà delle operazioni che siamo abituati a effettuare con i numeri, in particolare con i numeri reali.

La somma di due numeri reali è un'operazione che associa a ogni coppia di numeri reali $a$ e $b$ un altro numero reale, indicato con $a + b$. Quindi la somma è una funzione che ha come dominio $\mathbb{R} \times \mathbb{R}$ e come codominio $\mathbb{R}$:

$$\begin{aligned} + : \mathbb{R} \times \mathbb{R} &\to \mathbb{R} \\ (a, b) &\mapsto a + b \end{aligned}$$

La somma di numeri reali è:

(a) *commutativa*: $a + b = b + a$ per ogni $a, b \in \mathbb{R}$;

(b) *associativa*: $(a + b) + c = a + (b + c)$ per ogni $a, b, c \in \mathbb{R}$;

(c) ammette *elemento neutro*, cioè esiste un numero, lo 0, tale che $0 + a = a + 0 = a$ per ogni $a \in \mathbb{R}$;

(d) ogni numero reale $a$ ammette *opposto*, cioè esiste un altro numero, che indichiamo con $-a$, tale che $a + (-a) = 0$.

Il prodotto di due numeri reali è un'operazione che associa a ogni coppia di numeri reali $a$ e $b$ un altro numero reale, indicato con $ab$. Quindi anche il prodotto è una funzione che ha come dominio $\mathbb{R} \times \mathbb{R}$ e come codominio $\mathbb{R}$:

$$\begin{aligned} \cdot : \mathbb{R} \times \mathbb{R} &\to \mathbb{R} \\ (a, b) &\mapsto ab \end{aligned}$$

Il prodotto di numeri reali è:

(a) *commutativo*: $ab = ba$ per ogni $a, b \in \mathbb{R}$;

(b) *associativo*: $(ab)c = a(bc)$ per ogni $a, b, c \in \mathbb{R}$;

(c) ammette *elemento neutro*, cioè esiste un numero, 1, tale che $1a = a1 = a$ per ogni $a \in \mathbb{R}$;

(d) *distributivo rispetto alla somma*: $a(b + c) = ab + ac$ per ogni $a, b, c \in \mathbb{R}$.

Una delle proprietà più importanti dei numeri reali, che li distingue da altri insiemi di numeri, è la loro continuità. Geometricamente questo significa che pensiamo i numeri reali distribuiti lungo una retta. Più precisamente, data una retta e fissati su di essa un punto (origine) e un'unità di misura, esiste una corrispondenza biunivoca tra i punti sulla retta e l'insieme dei numeri reali. In altre parole ogni numero reale individua univocamente uno e un solo punto sulla retta.

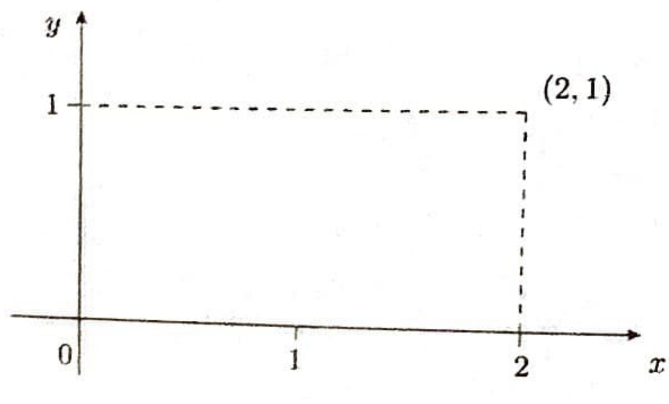

## 2.2 SPAZIO VETTORIALE $\mathbb{R}^N$ E SPAZIO DELLE MATRICI

Indichiamo con il simbolo $\mathbb{R}^2$ l'insieme delle coppie ordinate di numeri reali:

$$\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}$$

Il fatto che le coppie siano ordinate significa, per esempio, che l'elemento $(1, 2)$ è diverso dall'elemento $(2, 1)$.

Fissato nel piano un sistema di riferimento cartesiano, esiste una corrispondenza biunivoca tra $\mathbb{R}^2$ e l'insieme dei punti del piano. Fissare un riferimento cartesiano nel piano significa fissare due rette ortogonali orientate $r$ e $s$ e un'unità di misura. Il punto di intersezione tra le due rette si chiama origine del sistema di riferimento. Ogni punto del piano è allora univocamente individuato da una coppia di numeri reali, detti coordinate del punto, che indicano, rispettivamente, la distanza orientata del punto dalla retta $s$ e la sua distanza orientata dalla retta $r$. Lo studente che non avesse familiarità con il piano cartesiano, può pensare al piano di gioco della "battaglia navale".



È naturale cercare di estendere alle coppie di numeri reali le operazioni che sappiamo eseguire con i numeri. Definiamo allora:

- somma:

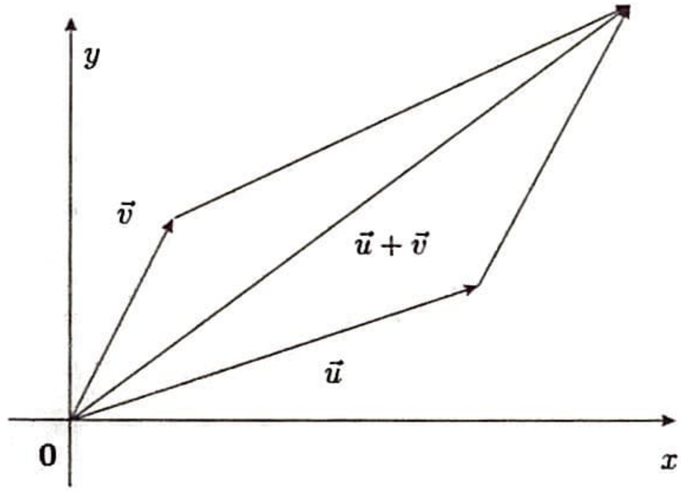
$$\begin{array}{rccc} + : & \mathbb{R}^2 \times \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & ((x,y),(x',y')) & \mapsto & (x,y)+(x',y') = (x+x', y+y') \end{array}$$

- prodotto per un numero reale:

$$\begin{array}{rccc} \cdot : & \mathbb{R} \times \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & (\lambda,(x,y)) & \mapsto & \lambda(x,y) = (\lambda x, \lambda y) \end{array}$$

Si noti che, pur avendo indicato l'operazione di prodotto per un numero reale con il simbolo $\cdot$, nel prodotto $\lambda(x,y)$ abbiamo omesso tale simbolo, esattamente come si fa solitamente quando si moltiplicano due numeri reali.

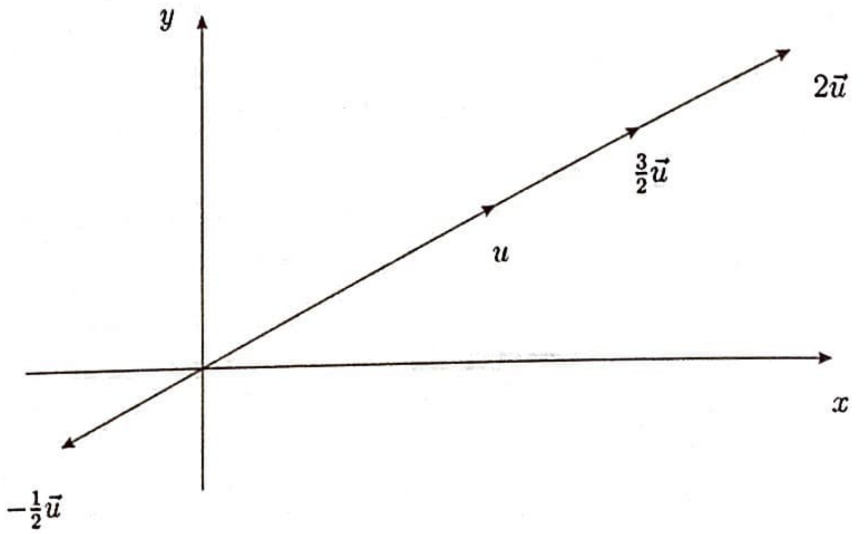
Cerchiamo di interpretare geometricamente le operazioni definite nel caso di $\mathbb{R}^2$. A tal fine pensiamo ogni elemento $(a,b)$ di $\mathbb{R}^2$ come il secondo estremo di un *vettore* applicato nell'origine, cioè di un segmento orientato uscente dall'origine con la freccia diretta verso il punto di coordinate $(a,b)$. In questo caso il modo di sommare due elementi di $\mathbb{R}^2$ coincide con la ben nota *regola del parallelogramma* con cui si sommano le forze in fisica. Secondo questa regola la somma di due vettori $\vec{u}$ e $\vec{v}$ applicati in un punto è un vettore applicato nello stesso punto avente direzione e lunghezza della diagonale del parallelogramma avente per lati $\vec{u}$ e $\vec{v}$, e verso uscente dal punto di applicazione.




Moltiplicando invece un vettore $\vec{v}$ per un numero reale $\alpha$ si ottiene un vettore avente la stessa direzione di $\vec{v}$, lunghezza moltiplicata per il valore assoluto di $\alpha$ e verso concorde o discorde da quello di $\vec{v}$ a seconda che $\alpha$ sia positivo o negativo.

È pressoché immediato verificare che la somma di elementi di $\mathbb{R}^2$ soddisfa le seguenti proprietà:

1. *Commutativa*: $(x,y)+(x',y') = (x',y')+(x,y)$
   per ogni $(x,y),(x',y') \in \mathbb{R}^2$;




2. *Associativa*: $((x,y)+(x',y'))+(x'',y'') = (x,y)+((x',y')+(x'',y''))$ per ogni $(x,y),(x',y'),(x'',y'') \in \mathbb{R}^2$;
3. *Esistenza dell'elemento neutro* $(0,0)$:

   $(x,y)+(0,0)=(0,0)+(x,y)=(x,y)$ per ogni $(x,y) \in \mathbb{R}^2$.
4. *Esistenza dell'opposto*: per ogni $(x,y) \in \mathbb{R}^2$ esiste un elemento $(a,b)$, detto *opposto* di $(x,y)$, tale che $(a,b)+(x,y)=(x,y)+(a,b)=(0,0)$; $(a,b)$ è ovviamente: $(a,b)=(-x,-y)$;
5. $\lambda((x,y)+(x',y')) = \lambda(x,y)+\lambda(x',y')$,

   per ogni $(x,y)$, $(x',y') \in \mathbb{R}^2$ e per ogni $\lambda \in \mathbb{R}$;
6. $(\lambda+\mu)(x,y) = \lambda(x,y)+\mu(x,y)$, per ogni $(x,y) \in \mathbb{R}^2$ e per ogni $\lambda, \mu \in \mathbb{R}$;
7. $(\lambda\mu)(x,y) = \lambda(\mu(x,y))$, per ogni $(x,y) \in \mathbb{R}^2$ e per ogni $\lambda, \mu \in \mathbb{R}$;
8. $1(x,y)=(x,y)$, per ogni $(x,y) \in \mathbb{R}^2$.

Naturalmente possiamo generalizzare quanto fatto in $\mathbb{R}^2$ all'insieme delle $n$-uple ordinate di numeri reali, per ogni $n \in \mathbb{N}$:

$$\mathbb{R}^n = \{(x_1,\dots,x_n) \mid x_1,\dots,x_n \in \mathbb{R}\}$$

Definiamo le seguenti operazioni di somma + e di prodotto $\cdot$ per i numeri reali:

$$+ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$$

$$(x_1,\dots,x_n)+(x_1',\dots,x_n') = (x_1+x_1',\dots,x_n+x_n')$$

$$\cdot : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$$

$$\lambda(x_1,\dots,x_n) = (\lambda x_1,\dots,\lambda x_n)$$

Con un po' di pazienza si possono verificare le proprietà 1-8 elencate sopra per la somma e il prodotto per i numeri reali in $\mathbb{R}^2$.

Esaminiamo un altro esempio. Consideriamo l'insieme delle matrici $2 \times 2$ a coefficienti reali:

$$\mathrm{M}_2(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{R} \right\}$$

introdotto nel Capitolo 1. Definiamo in $\mathrm{M}_2(\mathbb{R})$ le seguenti operazioni di somma $+$ e di prodotto $\cdot$ per i numeri reali:

$$+ : \mathrm{M}_2(\mathbb{R}) \times \mathrm{M}_2(\mathbb{R}) \to \mathrm{M}_2(\mathbb{R})$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} a + a' & b + b' \\ c + c' & d + d' \end{pmatrix}$$

$$\cdot : \mathbb{R} \times \mathrm{M}_2(\mathbb{R}) \to \mathrm{M}_2(\mathbb{R})$$

$$\lambda \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{pmatrix}$$

Anche in questo caso, armandosi di pazienza, è possibile verificare le proprietà 1-8. Gli studenti sono caldamente invitati a farlo. Dimostriamo per esempio la commutatività di $+$:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} a + a' & b + b' \\ c + c' & d + d' \end{pmatrix}$$

Poiché la somma di numeri reali è commutativa, possiamo scrivere:

$$\begin{pmatrix} a + a' & b + b' \\ c + c' & d + d' \end{pmatrix} = \begin{pmatrix} a' + a & b' + b \\ c' + c & d' + d \end{pmatrix} = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} + \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Ancora una volta tutto dipende dalle proprietà delle operazioni dei numeri reali. È questa la strategia per dimostrare le altre proprietà della somma e del prodotto per i reali in $\mathrm{M}_2(\mathbb{R})$ e più in generale, come vedremo, in un qualsiasi spazio vettoriale.

È chiaro che, analogamente a quanto fatto per le matrici $2 \times 2$, è possibile definire una somma e un prodotto per i numeri reali anche nell'insieme delle matrici $m \times n$ e verificare che tali operazioni godono di tutte le proprietà elencate sopra. Diamo dunque la definizione delle operazioni di somma e di prodotto per i numeri reali in $\mathrm{M}_{m,n}(\mathbb{R})$:

- somma

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} a'_{11} & a'_{12} & \dots & a'_{1n} \\ a'_{21} & a'_{22} & \dots & a'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a'_{m1} & a'_{m2} & \dots & a'_{mn} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} + a'_{11} & a_{12} + a'_{12} & \dots & a_{1n} + a'_{1n} \\ a_{21} + a'_{21} & a_{22} + a'_{22} & \dots & a_{2n} + a'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + a'_{m1} & a_{m2} + a'_{2m} & \dots & a_{mn} + a'_{mn} \end{pmatrix}$$

- prodotto per un numero reale:

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

Si noti che l'elemento neutro rispetto alla somma in $\mathrm{M}_{m,n}(\mathbb{R})$ è la *matrice nulla*

$$\begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

Gli insiemi fin qui descritti con le operazioni di somma e prodotto per i numeri reali sono esempi di spazi vettoriali.

## 2.3 SPAZI VETTORIALI

In questo paragrafo diamo la definizione di spazio vettoriale, cioè formalizziamo la struttura di cui, nel paragrafo precedente, abbiamo fatto alcuni esempi[1].

**Definizione 2.3.1** Si dice *spazio vettoriale* reale un insieme $V$ munito di due operazioni dette rispettivamente *somma e prodotto per scalari*:

$$\begin{aligned} + : V \times V &\longrightarrow V & \qquad \cdot : \mathbb{R} \times V &\longrightarrow V \\ (\mathbf{u}, \mathbf{v}) &\mapsto \mathbf{u} + \mathbf{v} & (\lambda, \mathbf{u}) &\mapsto \lambda \mathbf{u} \end{aligned}$$

che soddisfino le seguenti proprietà:
la somma + è:

1. *commutativa*, cioè $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$, per ogni $\mathbf{u}, \mathbf{v} \in V$;
2. *associativa*, cioè $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ per ogni $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$;
3. ammette un *elemento neutro*,
   cioè esiste $\mathbf{0} \in V$ tale che $\mathbf{0} + \mathbf{u} = \mathbf{u} + \mathbf{0} = \mathbf{u}$ per ogni $\mathbf{u}$ in $V$;
4. ogni elemento di $V$ ha un *opposto*,
   cioè per ogni $\mathbf{u} \in V$ esiste un vettore $\mathbf{a}$ tale che $\mathbf{a} + \mathbf{u} = \mathbf{u} + \mathbf{a} = \mathbf{0}$.

Inoltre valgono le seguenti proprietà:

5. $1\mathbf{u} = \mathbf{u}$;
6. $(\lambda\mu)\mathbf{u} = \lambda(\mu\mathbf{u})$, per ogni $\mathbf{u} \in V$ e per ogni $\lambda, \mu \in \mathbb{R}$;

---

[1] Definiremo il concetto di spazio vettoriale reale, in cui i vettori potranno essere moltiplicati per numeri reali che chiameremo *scalari*. È importante però sapere che è possibile sviluppare questa teoria anche sostituendo ai reali altri insiemi numerici, quali per esempio i numeri razionali, oppure i numeri complessi.

7. $\lambda(\mathbf{u}+\mathbf{v}) = \lambda\mathbf{u} + \lambda\mathbf{v}$, per ogni $\mathbf{u}, \mathbf{v} \in V$ e per ogni $\lambda \in \mathbb{R}$;
8. $(\lambda + \mu)\mathbf{u} = \lambda\mathbf{u} + \mu\mathbf{u}$, per ogni $\mathbf{u} \in V$ e per ogni $\lambda, \mu \in \mathbb{R}$.

Gli elementi di uno spazio vettoriale si dicono *vettori* mentre i numeri reali si dicono *scalari*. L'elemento neutro rispetto alla somma in $V$ si chiama *vettore nullo*. Per distinguere i vettori dai numeri indichiamo, come abbiamo fatto nella precedente definizione, i vettori in grassetto.

Nel paragrafo precedente abbiamo visto che $\mathbb{R}^n$ e $\mathrm{M}_{m,n}(\mathbb{R})$ sono spazi vettoriali reali. Vediamo ora altri esempi.

**Esempio tratto dalla fisica 2.3.2 I vettori applicati in un punto**

Consideriamo l'insieme $V$ dei vettori dello spazio applicati in uno stesso punto. Definiamo la somma di due siffatti vettori usando la regola del parallelogramma. L'operazione di somma gode chiaramente delle proprietà 1-4 e l'elemento neutro è il vettore di lunghezza zero.

Poi possiamo definire la moltiplicazione fra un numero reale e un vettore nel modo seguente: se $\alpha \in \mathbb{R}$ e $\vec{v} \in V$ allora $\alpha \cdot \vec{v}$ è il vettore applicato con direzione uguale a $\vec{v}$, lunghezza moltiplicata per il fattore $|\alpha|$ (ove $|\alpha|$ è il valore assoluto di $\alpha$) e verso uguale a quello di $\vec{v}$ o discorde da esso a seconda che il segno di $\alpha$ sia positivo o negativo.

In questo modo l'insieme dei vettori dello spazio applicati in un punto risulta essere uno spazio vettoriale reale.

**Esempio tratto dall'analisi matematica 2.3.3 Le funzioni**

Sia $F(\mathbb{R})$ l'insieme delle funzioni $f : \mathbb{R} \longrightarrow \mathbb{R}$ con le due operazioni:

$$(f+g)(x) = f(x) + g(x), \qquad (\lambda \cdot f)(x) = \lambda f(x)$$

Per esercizio si verifichi che $F(\mathbb{R})$ è uno spazio vettoriale reale.

**Esempio tratto dall'algebra 2.3.4 I polinomi**

Sia $\mathbb{R}[x]$ l'insieme dei polinomi a coefficienti reali nella variabile $x$. Con le usuali operazioni di somma e prodotto per un numero reale $\mathbb{R}[x]$ è uno spazio vettoriale. Di nuovo lasciamo la verifica delle proprietà 1-8 per esercizio. Osserviamo, per esempio, che il vettore nullo in $\mathbb{R}[x]$ è il polinomio identicamente nullo cioè 0; l'opposto del polinomio $5x^2 - x$ è il polinomio $-5x^2 + x$.

Mostriamo ora alcune utili proprietà che valgono in qualunque spazio vettoriale:

**Proposizione 2.3.5** *Sia $V$ uno spazio vettoriale. Allora valgono le seguenti proprietà:*

i) *Il vettore nullo è unico e verrà indicato con $\mathbf{0}_V$.*

ii) *Se $\mathbf{u}$ è un vettore di $V$ il suo opposto è unico e lo indicheremo con $-\mathbf{u}$.*

iii) $\lambda \mathbf{0}_V = \mathbf{0}_V$, *per ogni scalare* $\lambda \in \mathbb{R}$.

iv) $0\mathbf{u} = \mathbf{0}_V$, *per ogni* $\mathbf{u} \in V$ *(si noti il diverso significato dello zero al primo e al secondo membro!).*

v) *Se* $\lambda\mathbf{u} = \mathbf{0}_V$ *allora è o* $\lambda = 0$ *o* $\mathbf{u} = \mathbf{0}_V$.

vi) $(-\lambda)\mathbf{u} = \lambda(-\mathbf{u}) = -\lambda\mathbf{u}$.

**Dimostrazione** – È utile notare che mentre le proprietà 1-8 della Definizione 2.3.1 sono parte integrante della definizione e non vanno dimostrate, le proprietà i)-vi), che spesso allo studente appaiono ovvie, sono una conseguenza delle 1-8 e vanno invece dimostrate.

i) Se $\mathbf{0}'$ è un altro vettore che soddisfa la proprietà 3 della Definizione 2.3.1, si ha che $\mathbf{0}' + \mathbf{0}_V = \mathbf{0}_V$ (qui abbiamo preso $\mathbf{u} = \mathbf{0}_V$). Inoltre, usando il fatto che $\mathbf{0}_V$ soddisfa la proprietà 3 e prendendo $\mathbf{u} = \mathbf{0}'$ si ha anche $\mathbf{0}' + \mathbf{0}_V = \mathbf{0}'$. Segue che $\mathbf{0}_V = \mathbf{0}' + \mathbf{0}_V = \mathbf{0}'$.

ii) Se $\mathbf{a}$ e $\mathbf{a}'$ sono entrambi opposti di $\mathbf{u}$, per la proprietà 4 si ha in particolare che $\mathbf{a} + \mathbf{u} = \mathbf{0}_V$ e $\mathbf{u} + \mathbf{a}' = \mathbf{0}_V$. Quindi $\mathbf{a} = \mathbf{a} + \mathbf{0}_V = \mathbf{a} + (\mathbf{u} + \mathbf{a}') = (\mathbf{a} + \mathbf{u}) + \mathbf{a}' = \mathbf{0}_V + \mathbf{a}' = \mathbf{a}'$, ove abbiamo usato l'associatività della somma.

iii) Si ha che $\lambda\mathbf{0}_V = \lambda(\mathbf{0}_V + \mathbf{0}_V)$ per la proprietà 3. Inoltre $\lambda(\mathbf{0}_V + \mathbf{0}_V) = \lambda\mathbf{0}_V + \lambda\mathbf{0}_V$ per la proprietà 7, dunque $\lambda\mathbf{0}_V = \lambda\mathbf{0}_V + \lambda\mathbf{0}_V$. Sommiamo a entrambi i membri $-\lambda\mathbf{0}_V$ e otteniamo, per la proprietà 4, che $\mathbf{0}_V = \lambda\mathbf{0}_V$.

iv) Si ha che $0\mathbf{u} = (0 + 0)\mathbf{u} = 0\mathbf{u} + 0\mathbf{u}$ per la proprietà 8, quindi $0\mathbf{u} = 0\mathbf{u} + 0\mathbf{u}$. Sommiamo a entrambi i membri $-0\mathbf{u}$ e otteniamo, per la proprietà 4, che $\mathbf{0}_V = 0\mathbf{u}$.

v) Mostriamo che se $\lambda\mathbf{u} = \mathbf{0}_V$ e $\lambda \neq 0$ allora deve essere $\mathbf{u} = \mathbf{0}_V$. Sia $\frac{1}{\lambda}$ l'inverso di $\lambda$, che esiste perché $\lambda \neq 0$. Moltiplicando per $\frac{1}{\lambda}$ entrambi i membri dell'uguaglianza $\lambda\mathbf{u} = \mathbf{0}_V$, si ha: $\mathbf{u} = (\frac{1}{\lambda}\lambda)\mathbf{u} = \frac{1}{\lambda}\mathbf{0}_V = \mathbf{0}_V$.

vi) Si ha che $(-\lambda)\mathbf{u} + \lambda\mathbf{u} = (-\lambda + \lambda)\mathbf{u} = 0\mathbf{u} = \mathbf{0}_V$, per le proprietà 8 e iv). Analogamente $\lambda\mathbf{u} + (-\lambda)\mathbf{u} = \mathbf{0}_V$. Quindi $(-\lambda)\mathbf{u}$ è l'opposto di $\lambda\mathbf{u}$, cioè $(-\lambda)\mathbf{u} = -\lambda\mathbf{u}$.

Inoltre $\lambda(-\mathbf{u}) + \lambda\mathbf{u} = \lambda(-\mathbf{u} + \mathbf{u}) = \lambda\mathbf{0}_V = \mathbf{0}_V$ per le proprietà 7, 4 e iii). Analogamente $\lambda\mathbf{u} + (-\lambda)\mathbf{u} = \mathbf{0}_V$, quindi anche $\lambda(-\mathbf{u})$ è l'opposto di $\lambda\mathbf{u}$, cioè $\lambda(-\mathbf{u}) = -\lambda\mathbf{u}$. □

**Definizione 2.3.6** Si chiama *spazio vettoriale banale* e si indica con $\{\mathbf{0}_V\}$, uno spazio vettoriale costituito solo dal vettore nullo.

♦ **Osservazione 2.3.7** Per definizione di vettore nullo e per la proprietà iii) della Proposizione 2.3.5, le operazioni di somma e prodotto per scalari nello spazio banale sono ben definite e ovviamente banali:

$$\mathbf{0}_V + \mathbf{0}_V = \mathbf{0}_V$$
$$\lambda\mathbf{0}_V = \mathbf{0}_V \text{ per ogni } \lambda \in \mathbb{R}$$

♦ **Osservazione 2.3.8** Riflettiamo ancora un po' sulla definizione di $\mathbb{R}$-spazio vettoriale. Innanzitutto osserviamo che, per definizione, *uno spazio vettoriale reale non può mai essere vuoto.* Esso deve infatti contenere almeno il vettore nullo, cioè l'elemento neutro rispetto alla somma.

Supponiamo ora che $V$ sia uno spazio vettoriale *non* banale, ossia che esso contenga almeno un vettore $\mathbf{v} \neq \mathbf{0}_V$. Quanti elementi contiene $V$? Poiché disponiamo del prodotto per i numeri reali, che sono infiniti, $V$ conterrà tutti gli infiniti multipli di $\mathbf{v}$, cioè tutti i vettori della forma $\lambda\mathbf{v}$ al variare di $\lambda \in \mathbb{R}$.

Al variare di $\lambda$ nell'insieme dei numeri reali si ottengono così infiniti elementi diversi di $V$. Per dimostrare in modo rigoroso questa affermazione, occorre mostrare che, se $\lambda$ e $\mu$ sono numeri reali distinti, cioè $\lambda \neq \mu$, e $\mathbf{v} \in V$ è un vettore non nullo, allora $\lambda\mathbf{v} \neq \mu\mathbf{v}$. In effetti, se così non fosse, avremmo:

$$\lambda\mathbf{v} = \mu\mathbf{v} \Leftrightarrow (\lambda - \mu)\mathbf{v} = \mathbf{0}_V$$

con $\lambda - \mu \neq 0$ e $\mathbf{v} \neq \mathbf{0}_V$, il che contraddirebbe la proprietà $(v)$ della Proposizione 2.3.5.

## 2.4 SOTTOSPAZI VETTORIALI

Come possiamo riconoscere e descrivere uno spazio vettoriale dentro un altro? Come distinguiamo un sottoinsieme qualsiasi di uno spazio vettoriale da uno che ne possiede le stesse caratteristiche? Per rispondere a queste domande è necessario introdurre la definizione di sottospazio vettoriale.

**Definizione 2.4.1** Sia $W$ un sottoinsieme dello spazio vettoriale $V$. Diremo che $W$ è un *sottospazio vettoriale* di $V$ se soddisfa le seguenti proprietà:

1) $W$ è diverso dall'insieme vuoto;

$\forall u, v \in W. \ u \cdot v \in W$

2) $W$ è chiuso rispetto alla somma, cioè per ogni $\mathbf{u}, \mathbf{v} \in W$ si ha che $\mathbf{u} + \mathbf{v} \in W$;

$\forall u \in W \wedge \forall \lambda \in \mathbb{R}, \ \lambda u \in W$

3) $W$ è chiuso rispetto al prodotto per scalari, cioè per ogni $\mathbf{u} \in W$ e ogni $\lambda \in \mathbb{R}$ si ha che $\lambda\,\mathbf{u} \in W$.

Osserviamo che poiché $W$ non è vuoto, e $\lambda\mathbf{u} \in W$ per ogni $\lambda \in \mathbb{R}$, allora prendendo un vettore qualsiasi di $W$ e moltiplicandolo per $\lambda = 0$ si ottiene che $\mathbf{0}_V \in W$. In effetti la proprietà (1) puo essere efficacemente sostituita dalla proprietà:

1') $\mathbf{0}_V \in W$,

ottenendo una definizione equivalente.

È importante notare che un sottospazio vettoriale $W$ di $V$ è uno spazio vettoriale con le operazioni di $V$ ristrette a $W$. Infatti, la proprietà 2) della Definizione 2.4.1 assicura che la restrizione a $W$ della somma definita in $V$ dia come risultato un vettore di $W$:

$$+_V|_{W \times W} : W \times W \to W$$

Analogamente la proprietà 3) della Definizione 2.4.1 assicura che la restrizione a $\mathbb{R} \times W$ del prodotto per scalari definito su $\mathbb{R} \times V$ dia come risultato un vettore di $W$. Perciò le proprietà 1-8 della Definizione 2.3.1 continuano a valere perché valgono in $V$.

In particolare, dunque, ogni spazio vettoriale $V$ possiede sempre almeno due sottospazi: $V$ stesso e il sottospazio banale, costituito solo dal vettore nullo $\mathbf{0}_V$. In virtù dell'Osservazione 2.3.8, se $V$ stesso non è banale, ogni sottospazio vettoriale di $V$ non banale conterrà infiniti elementi.

Cerchiamo di chiarire il concetto di sottospazio vettoriale attraverso degli esempi e dei controesempi.

---

**Esempi 2.4.2**

L'insieme $X = \{(x, y) \in \mathbb{R}^2 \mid y = 0\}$ è un sottospazio vettoriale di $\mathbb{R}^2$. $X$ è infatti:

1) non vuoto: contiene le infinite coppie di numeri reali $(x, 0)$;

2) chiuso rispetto alla somma: presi due elementi $(x_1, 0)$, $(x_2, 0)$ in $X$, la loro somma $(x_1, 0) + (x_2, 0) = (x_1 + x_2, 0)$ appartiene ancora all'insieme $X$;

3) chiuso rispetto al prodotto per scalari: presi un qualsiasi numero reale $\alpha$ e un qualsiasi elemento $(x, 0) \in X$, il prodotto $\alpha(x, 0) = (\alpha x, 0)$ appartiene a $X$.

Geometricamente, fissato un sistema di riferimento cartesiano in $\mathbb{R}^2$, l'insieme $X$ può essere pensato come l'asse delle ascisse. Allora sommando due vettori che giacciono sull'asse delle $x$ o moltiplicando per uno scalare uno di essi, si ottiene ancora un vettore che giace sull'asse delle $x$.

Più in generale sia $a$ un numero reale fissato e sia $W_a = \{(x, y) \in \mathbb{R}^2 \mid y = ax\}$. Osserviamo innanzitutto che $(0, 0) \in W$. Inoltre, presi due elementi $(x_1, ax_1)$ e $(x_2, ax_2)$ in $W_a$, la loro somma

$$(x_1, ax_1) + (x_2, ax_2) = (x_1 + x_2, a(x_1 + x_2))$$

appartiene a $W_a$. Cioè $W_a$ è chiuso rispetto alla somma. Analogamente, presi $\lambda \in \mathbb{R}$ e $(x_1, ax_1) \in W_a$, si ha:

$$\lambda(x_1, ax_1) = (\lambda x_1, \lambda a x_1) = (\lambda x_1, a(\lambda x_1))$$

che appartiene a $W_a$, cioè $W_a$ è chiuso rispetto al prodotto per scalari. Dunque $W_a$ è un sottospazio vettoriale di $\mathbb{R}^2$.

Fissato ancora una volta un sistema di riferimento cartesiano nel piano, $W_a$ può essere pensato come l'insieme dei punti di una retta passante per l'origine del sistema di riferimento.

Tutte le rette del piano descrivono dei sottospazi vettoriali?

---

**Esempio 2.4.3**

Consideriamo l'insieme $S = \{(x, y) \in \mathbb{R}^2 \mid y = x + 1\}$. Possiamo immediatamente affermare che $S$ *non* è un sottospazio vettoriale di $\mathbb{R}^2$ poiché non contiene $\mathbf{0}_{\mathbb{R}^2} = (0, 0)$. Dunque non tutte le rette del piano individuano dei sottospazi vettoriali di $\mathbb{R}^2$, ma solo quelle passanti per $(0, 0)$.

---

**Esempio 2.4.4**

L'insieme $X = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{M}_2(\mathbb{R}) \mid b = 1 \right\} \subseteq \mathrm{M}_2(\mathbb{R})$ non è un sottospazio vettoriale di $\mathrm{M}_2(\mathbb{R})$, poiché la matrice nulla $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ non appartiene a $X$.

♦ **Osservazione 2.4.5** Se $S$ è un sottospazio vettoriale di uno spazio vettoriale $V$ la condizione $\mathbf{0}_V \in S$ è necessaria ma non sufficiente affinché $S$ sia un sottospazio di $V$. Facciamo subito un controesempio.

**Esempio 2.4.6**

Sia $S = \{(x, y, z) \in \mathbb{R}^3 \mid xy = z\}$. Nonostante l'insieme $S$ contenga il vettore nullo $(0, 0, 0)$ di $\mathbb{R}^3$, $S$ non è un sottospazio vettoriale di $\mathbb{R}^3$, perché non è chiuso rispetto alla somma. Infatti i vettori $\mathbf{v} = (1, 1, 1)$ e $\mathbf{w} = (-1, -1, 1)$ appartengono a $S$, dal momento che soddisfano l'equazione $xy = z$, tuttavia la loro somma $\mathbf{v} + \mathbf{w} = (1, 1, 1) + (-1, -1, 1) = (0, 0, 2)$ non appartiene a $S$ poiché $0 \cdot 0 \neq 2$.

♦ **Osservazione 2.4.7** Dire che un sottoinsieme $S$ di uno spazio vettoriale $V$ è chiuso rispetto al prodotto per scalari, significa dire che, se $S$ contiene un vettore non nullo, allora *deve contenere anche tutti i suoi multipli*. Se pensiamo all'interpretazione geometrica del prodotto per scalari che abbiamo dato nel caso di $\mathbb{R}^2$, questo significa che, se un sottospazio di $\mathbb{R}^2$ contiene un vettore non nullo $\mathbf{v}$, allora esso contiene tutta la retta del piano passante per l'origine e individuata da $\mathbf{v}$. Questo ragionamento geometrico ci permette di dire immediatamente che alcuni sottoinsiemi del piano che ben conosciamo *non* sono spazi vettoriali, per esempio

$$C = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$$

(una circonferenza),

$$P = \{(x, y) \in \mathbb{R}^2 \mid y = x^2\}$$

(una parabola) e così via.

In un certo senso uno spazio vettoriale non può essere *curvo*, da cui il nome appunto di algebra *lineare*.

♦ **Osservazione 2.4.8** Gli esempi fatti finora mettono in luce due diversi tipi di ragionamento. Per dimostrare che un sottoinsieme $S$ di uno spazio vettoriale $V$ è un sottospazio vettoriale di $V$, bisogna dimostrare che valgono le proprietà 1), 2) e 3) della Definizione 2.4.1. Tali proprietà devono valere *sempre* cioè per ogni coppia di vettori di $S$ (proprietà 2)) e per tutti i numeri reali (proprietà 3)).

Per dimostrare, al contrario, che un sottoinsieme di uno spazio vettoriale $V$ *non* è un sottospazio vettoriale di $V$, basta mostrare che una delle proprietà 1), 2) o 3) della Definizione 2.4.1 fallisce, ossia, se $S \neq \emptyset$, che esiste (anche solo) una coppia di vettori di $S$ la cui somma non sta in $S$ o che esistono (anche solo) un vettore di $S$ e uno scalare il cui prodotto non sta in $S$.

**Esempi 2.4.9**

Nello spazio vettoriale $\mathbb{R}[x]$ dei polinomi a coefficienti reali in una variabile $x$, consideriamo il sottoinsieme $\mathbb{R}_2[x]$ costituito dai polinomi di grado minore o uguale a 2:

$$\mathbb{R}_2[x] = \{p(x) = a + bx + cx^2 \mid a, b, c \in \mathbb{R}\}$$

Allora $\mathbb{R}_2[x]$ è un sottospazio vettoriale di $\mathbb{R}[x]$. Infatti, sommando due polinomi di grado minore o uguale a 2, si ottiene ancora un polinomio di grado minore o uguale a 2:

$$(a + bx + cx^2) + (a' + b'x + c'x^2) = (a + a') + (b + b')x + (c + c')x^2$$

Analogamente, moltiplicando un polinomio di grado minore o uguale a 2 per un numero reale $\lambda$, si ottiene ancora un polinomio di grado minore o uguale a 2:

$$\lambda(a + bx + cx^2) = (\lambda a) + (\lambda b)x + (\lambda c)x^2$$

In altre parole $\mathbb{R}_2[x]$ (che è certamente non vuoto) è chiuso rispetto alla somma e al prodotto per scalari, pertanto è un sottospazio vettoriale di $\mathbb{R}[x]$.

**Esempio 2.4.10**

Nello spazio vettoriale $\mathbb{R}[x]$ dei polinomi a coefficienti reali in una variabile $x$, consideriamo il sottoinsieme $S = \{p(x) = a + bx + cx^2 \in \mathbb{R}[x] \mid ac = 0\}$. Il sottoinsieme $S$ contiene, per esempio, il polinomio identicamente nullo e il monomio $x$, pertanto è diverso dall'insieme vuoto. Tuttavia esso non è chiuso rispetto alla somma definita in $\mathbb{R}[x]$. Infatti $S$ contiene i polinomi $p(x) = 1 + x$ e $q(x) = x + x^2$, ma non contiene la loro somma: $p(x) + q(x) = 1 + 2x + x^2$.

Dal momento che un sottospazio vettoriale di uno spazio vettoriale $V$ è in primo luogo un sottoinsieme di $V$, è naturale chiedersi che cosa succeda effettuando le operazioni insiemistiche di unione e intersezione su due (o più) sottospazi vettoriali di $V$.

**Esempio 2.4.11**

Consideriamo l'insieme $W = X \cup Y$ con

$$X = \{(x, y) \in \mathbb{R}^2 \mid y = 0\} \quad \text{e} \quad Y = \{(x, y) \in \mathbb{R}^2 \mid x = 0\}$$

Nell'Esempio 2.4.2 abbiamo mostrato che $X$ è un sottospazio vettoriale di $\mathbb{R}^2$. In modo del tutto analogo si può mostrare che anche $Y$ è un sottospazio vettoriale di $\mathbb{R}^2$. Tuttavia la loro unione $W$ non è un sottospazio vettoriale di $\mathbb{R}^2$, perché non è chiusa rispetto alla somma: infatti, il vettore $(1, 0)$ appartiene a $W$ perché è un elemento di $X$ e il vettore $(0, 1)$ appartiene a $W$ perché è un elemento di $Y$, ma la loro somma $(1, 0) + (0, 1) = (1, 1)$ non appartiene a $W$, poiché non appartiene né a $X$ né a $Y$.

Geometricamente il ragionamento è altrettanto semplice: possiamo pensare a $X$ e $Y$ come agli assi rispettivamente delle ascisse e delle ordinate in un riferimento cartesiano nel piano e a $W$ come all'unione dei due assi. È chiaro dalla regola del parallelogramma che la somma di un vettore che giace sull'asse delle ascisse e di uno che giace sull'asse delle ordinate starà *al di fuori* delle due rette, dunque $W$ non è un sottospazio vettoriale. Osserviamo che l'insieme $W$ può essere descritto nel modo seguente:

$$W = \{(x, y) \in \mathbb{R}^2 \mid xy = 0\}$$

Essendo, infatti, $x$ e $y$ numeri reali, il loro prodotto è nullo se e solo se almeno uno dei due fattori è nullo.

L'esempio appena fatto mostra che, in generale, l'unione di due sottospazi di uno spazio vettoriale $V$ *non* è un sottospazio di $V$. Più precisamente vale la seguente proposizione:

**Proposizione 2.4.12** *Siano $W_1, W_2$ due sottospazi di uno spazio vettoriale $V$. Allora $W_1 \cup W_2$ è un sottospazio di $V$ se e solo se $W_1 \subseteq W_2$ o $W_2 \subseteq W_1$.*

**Dimostrazione** – "$\Leftarrow$" Se $W_1 \subseteq W_2$ (rispettivamente se $W_2 \subseteq W_1$), allora $W_1 \cup W_2 = W_2$ (risp. $W_1 \cup W_2 = W_1$) che è sottospazio vettoriale per ipotesi.

"$\Rightarrow$" Per dimostrare questa implicazione mostriamo che se $W_1 \nsubseteq W_2$ e $W_2 \nsubseteq W_1$ allora $W_1 \cup W_2$ non è un sottospazio di $V$. Essendo $W_1 \nsubseteq W_2$ esiste un vettore $\mathbf{v}_1 \in W_1 \setminus W_2$, analogamente essendo $W_2 \nsubseteq W_1$ esiste un vettore $\mathbf{v}_2 \in W_2 \setminus W_1$. Se $W_1 \cup W_2$ fosse un sottospazio di $V$ allora $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$ dovrebbe essere un elemento di $W_1 \cup W_2$ in quanto somma di un elemento di $W_1$ e di uno di $W_2$. Se $\mathbf{v}$ fosse in $W_1$ allora anche $\mathbf{v}_2 = \mathbf{v} - \mathbf{v}_1$ apparterrebbe a $W_1$, ma avevamo scelto $\mathbf{v}_2 \in W_2 \setminus W_1$. Analogamente se $\mathbf{v}$ fosse in $W_2$ allora anche $\mathbf{v}_1 = \mathbf{v} - \mathbf{v}_2$ apparterrebbe a $W_2$, ma avevamo scelto $\mathbf{v}_1 \in W_1 \setminus W_2$.
Quindi $\mathbf{v} \notin W_1 \cup W_2$. □

Con l'intersezione di due sottospazi si hanno meno problemi.

**Proposizione 2.4.13** *L'intersezione $S_1 \cap S_2$ di due sottospazi $S_1$ e $S_2$ di uno spazio vettoriale $V$ è un sottospazio vettoriale di $V$.*

**Dimostrazione** – Dobbiamo mostrare che $S_1 \cap S_2$ è un sottospazio di $V$: osserviamo innanzitutto che tale intersezione è non vuota dal momento che $\mathbf{0}_V$ appartiene sia a $S_1$ sia a $S_2$, quindi appartiene a $S_1 \cap S_2$. Ora mostriamo che $S_1 \cap S_2$ è chiuso rispetto alla somma di $V$: siano quindi $\mathbf{v}_1, \mathbf{v}_2 \in S_1 \cap S_2$. Questo significa, in particolare, che $\mathbf{v}_1, \mathbf{v}_2 \in S_1$, che è un sottospazio di $V$, quindi $\mathbf{v}_1 + \mathbf{v}_2 \in S_1$. Analogamente $\mathbf{v}_1, \mathbf{v}_2 \in S_2$ che è un sottospazio di $V$, quindi $\mathbf{v}_1 + \mathbf{v}_2 \in S_2$, dunque $\mathbf{v}_1 + \mathbf{v}_2 \in S_1 \cap S_2$.

Analogamente mostriamo che $S_1 \cap S_2$ è chiuso rispetto al prodotto per scalari. Siano $\mathbf{v} \in S_1 \cap S_2$ e $\lambda \in \mathbb{R}$. In particolare $\mathbf{v} \in S_1$, che è un sottospazio, quindi $\lambda\mathbf{v} \in S_1$; analogamente $\mathbf{v} \in S_2$, che è un sottospazio, quindi $\lambda\mathbf{v} \in S_2$ da cui $\lambda\mathbf{v}$ appartiene sia a $S_1$ sia a $S_2$ quindi appartiene alla loro intersezione. □

**Esempio 2.4.14**

Consideriamo i sottospazi vettoriali:

$$S = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{M}_2(\mathbb{R}) \mid b = -c \right\} \quad \text{e}$$

$$T = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{M}_2(\mathbb{R}) \mid a + b + c + d = 0 \right\}$$

di $\mathrm{M}_2(\mathbb{R})$.

Come è fatto $S \cap T$? Il sottospazio $S \cap T$ è costituito dagli elementi di $M_2(\mathbb{R})$ che appartengono sia a $S$ sia a $T$, cioè:

$$S \cap T = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R}) \mid b = -c, a + b + c + d = 0 \right\}$$

Si ha dunque:

$$S \cap T = \left\{ \begin{pmatrix} a & -c \\ c & -a \end{pmatrix} \in M_2(\mathbb{R}) \right\}$$

È facile verificare che tale sottoinsieme di $M_2(\mathbb{R})$ è chiuso rispetto alla somma e al prodotto per scalari, come garantito dalla Proposizione 2.4.13.

## 2.5 ESERCIZI SVOLTI

### Esercizio 2.5.1

*Si stabilisca se l'insieme* $X = \{(r, s, r-s) \in \mathbb{R}^3\}$ *è un sottospazio vettoriale di* $\mathbb{R}^3$.

**Svolgimento**

Innanzitutto osserviamo che $X$ è diverso dall'insieme vuoto perché $(0,0,0) \in X$ (basta prendere $r = s = 0$).

Consideriamo ora due elementi generici di $X$:

$$(r_1, s_1, r_1 - s_1) \quad \text{e} \quad (r_2, s_2, r_2 - s_2).$$

La loro somma è:

$$\begin{aligned}(r_1, s_1, r_1 - s_1) + (r_2, s_2, r_2 - s_2) &= (r_1 + r_2, s_1 + s_2, r_1 - s_1 + r_2 - s_2) \\ &= (r_1 + r_2, s_1 + s_2, r_1 + r_2 - (s_1 + s_2))\end{aligned}$$

e appartiene ancora a $X$ in quanto è del tipo $(r, s, r-s)$, con

$$r = r_1 + r_2 \quad \text{e} \quad s = s_1 + s_2.$$

Siano poi $(r_1, s_1, r_1 - s_1) \in X$ e $\lambda \in \mathbb{R}$. Allora $\lambda(r_1, s_1, r_1 - s_1) = (\lambda r_1, \lambda s_1, \lambda(r_1 - s_1)) = (\lambda r_1, \lambda s_1, \lambda r_1 - \lambda s_1)$ appartiene ancora a $X$ in quanto è del tipo $(r, s, r-s)$, con $r = \lambda r_1$ e $s = \lambda s_1$. Quindi $X$ è un sottospazio vettoriale di $\mathbb{R}^3$. ■

### Esercizio 2.5.2

*Si stabilisca se l'insieme* $W = \{(x, y, z) \in \mathbb{R}^3 \mid 2x + z^2 = 0\} \subseteq \mathbb{R}^3$ *è un sottospazio vettoriale di* $\mathbb{R}^3$.

**Svolgimento**

Sicuramente $W$ è diverso dall'insieme vuoto perché $(0,0,0) \in W$.

Consideriamo ora due elementi generici di $W$, $(x_1, y_1, z_1)$ e $(x_2, y_2, z_2)$, con $2x_1 + z_1^2 = 0$ e $2x_2 + z_2^2 = 0$. Si ha che $(x_1, y_1, z_1) + (x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2)$ e tale somma appartiene a $W$ se e solo se $2(x_1 + x_2) + (z_1 + z_2)^2 = 0$. Ma $2(x_1 + x_2) + (z_1 + z_2)^2 = 2x_1 + 2x_2 + z_1^2 + z_2^2 + 2z_1 z_2 = (2x_1 + z_1^2) + (2x_2 + z_2^2) + 2z_1 z_2 = 0 + 0 + 2z_1 z_2 = 2z_1 z_2$, che, in generale, non è uguale a zero.

Per esempio gli elementi $(-2,1,2)$ e $(-8,3,4)$ appartengono a $W$, ma $(-2,1,2)+(-8,3,4)=(-10,4,6) \notin W$, perché $2\cdot(-10)+(6)^2 \neq 0$. (Si noti che tali elementi di $W$ non sono stati scelti a caso, ma in modo da soddisfare la richiesta $2z_1z_2 \neq 0$).

Quindi $W$ non è un sottospazio vettoriale di $\mathbb{R}^3$. ∎

### Esercizio 2.5.3

*Si determini un sottoinsieme non vuoto di $\mathbb{R}^3$ chiuso rispetto alla somma, ma non al prodotto per scalari.*

**Svolgimento**

L'insieme $X = \{(x,y,z) | x,y,z \in \mathbb{R}, x \geq 0\}$ ha questa proprietà. Infatti $X$ è non vuoto poiché, per esempio, $(0,0,0) \in X$. Verifichiamo ora che sia chiuso rispetto alla somma. Siano $(x_1,y_1,z_1)$, $(x_2,y_2,z_2) \in X$, con $x_1,x_2 \geq 0$. Allora $(x_1,y_1,z_1)+(x_2,y_2,z_2) = (x_1+x_2, y_1+y_2, z_1+z_2) \in X$, perché $x_1+x_2 \geq 0$ (la somma di due numeri reali non negativi è un numero reale non negativo). Siano ora $(x_1,y_1,z_1) \in X$ e $\lambda \in \mathbb{R}$. Abbiamo che $\lambda(x_1,y_1,z_1) = (\lambda x_1, \lambda y_1, \lambda z_1)$ appartiene a $X$ se e solo se $\lambda x_1 \geq 0$, ma se scegliamo $\lambda$ negativo e $x_1 > 0$, per esempio $\lambda = -1$ e $(x_1,y_1,z_1) = (3,-2,1)$, sicuramente tale condizione non è verificata. Quindi $X$ non è chiuso rispetto al prodotto per scalari. ∎

### Esercizio 2.5.4

*Si stabilisca per quali valori del parametro $k$ l'insieme*

$$X_k = \left\{ \begin{pmatrix} r & s \\ r+k & k^2-k \end{pmatrix} \middle| r,s \in \mathbb{R} \right\} \subseteq \mathrm{M}_2(\mathbb{R})$$

*è un sottospazio vettoriale di $\mathrm{M}_2(\mathbb{R})$.*

**Svolgimento**

Sappiamo che affinché $X_k$ sia un sottospazio vettoriale di $\mathrm{M}_2(\mathbb{R})$, la matrice nulla deve appartenere a $X_k$, cioè deve succedere che $\left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ sia del tipo $\left(\begin{smallmatrix} r & s \\ r+k & k^2-k \end{smallmatrix}\right)$ per qualche $r,s \in \mathbb{R}$. Questo succede se

$$\begin{cases} r = 0 \\ s = 0 \\ k = 0 \\ k^2 - k = 0 \end{cases}$$

cioè $k = 0$.

Vediamo ora se $X_0 = \left\{ \left(\begin{smallmatrix} r & s \\ r & 0 \end{smallmatrix}\right) \middle| r,s \in \mathbb{R} \right\}$ è un sottospazio di $\mathrm{M}_2(\mathbb{R})$. Sicuramente $X_0$ è non vuoto poiché contiene la matrice nulla.

Siano ora $\left(\begin{smallmatrix} r_1 & s_1 \\ r_1 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} r_2 & s_2 \\ r_2 & 0 \end{smallmatrix}\right) \in X_0$. Si ha:

$$\begin{pmatrix} r_1 & s_1 \\ r_1 & 0 \end{pmatrix} + \begin{pmatrix} r_2 & s_2 \\ r_2 & 0 \end{pmatrix} = \begin{pmatrix} r_1+r_2 & s_1+s_2 \\ r_1+r_2 & 0 \end{pmatrix}$$

La matrice ottenuta appartiene dunque a $X_0$. Analogamente, preso $\lambda \in \mathbb{R}$, anche $\lambda\left(\begin{smallmatrix} r_1 & s_1 \\ r_1 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} \lambda r_1 & \lambda s_1 \\ \lambda r_1 & 0 \end{smallmatrix}\right)$ appartiene a $X_0$, quindi $X_0$ è chiuso rispetto alla somma e al prodotto per scalari, dunque è un sottospazio vettoriale di $\mathrm{M}_2(\mathbb{R})$.

In conclusione, $X_k$ è un sottospazio di $\mathrm{M}_2(\mathbb{R})$ per $k = 0$. ∎

## 2.6 ESERCIZI PROPOSTI

### Esercizio 2.6.1

*Si dica quali dei seguenti sottoinsiemi di spazi vettoriali sono sottospazi:*

a) $S = \{(x, y, z) \in \mathbb{R}^3 \mid x = 0\}$;

b) $T = \{(x, y, z) \in \mathbb{R}^3 \mid x = y\}$;

c) $W_n = \{p(x) \in \mathbb{R}[x] \mid deg(p(x)) = n\}$, $n \in \mathbb{N}$. *(Qui $deg(p(x))$ indica il grado del polinomio $p(x)$;*

d) $D = \left\{\begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} \in \mathrm{M}_2(\mathbb{R})\right\}$;

e) $T = \left\{\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in \mathrm{M}_2(\mathbb{R})\right\}$;

f) $A = \{(a_{ij})_{\substack{i=1,\dots,3 \\ j=1,\dots,3}} \in \mathrm{M}_3(\mathbb{R}) \mid a_{11} + a_{22} + a_{33} = 0\}$;

g) $X = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y = 0\}$;

h) $X = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = -1\}$;

i) $X = \{p(x) = 3x^2 + rx + s \mid r, s \in \mathbb{R}\} \subseteq \mathbb{R}_2[x]$;

l) $X = \left\{\begin{pmatrix} 0 & r \\ 2r & s \end{pmatrix} \mid r, s \in \mathbb{R}\right\} \subseteq \mathrm{M}_2(\mathbb{R})$;

m) $X = \left\{\begin{pmatrix} r & 2r \\ r & r^2 \end{pmatrix} \mid r \in \mathbb{R}\right\} \subseteq \mathrm{M}_2(\mathbb{R})$.

### Esercizio 2.6.2

*Si dimostri che l'insieme delle soluzioni del sistema lineare omogeneo*

$$\begin{cases} x_1 + 3x_2 - x_3 = 0 \\ 2x_1 + 4x_2 - 4x_3 - x_4 = 0 \\ x_2 + x_3 + 2x_4 = 0 \end{cases}$$

*nelle incognite $x_1, x_2, x_3, x_4$ è un sottospazio vettoriale di $\mathbb{R}^4$.*

### Esercizio 2.6.3

*Si stabilisca se l'insieme $X = \{p(x) \in \mathbb{R}_3[x] \mid p(-1) = 0\}$ è un sottospazio vettoriale di $\mathbb{R}_3[x]$, ove per $p(-1)$ si intende il valore del polinomio calcolato in $-1$.*

### Esercizio 2.6.4

*Si stabilisca se l'insieme $X = \{p(x) \in \mathbb{R}_2[x] \mid p(1) = -1\}$ è un sottospazio vettoriale di $\mathbb{R}_2[x]$.*

### Esercizio 2.6.5

*Si stabilisca se l'insieme $X = \{g : \mathbb{R} \to \mathbb{R} \mid g$ è continua e derivabile in $x = 2\}$ è un sottospazio vettoriale dello spazio vettoriale delle funzioni continue $f : \mathbb{R} \to \mathbb{R}$.*

### Esercizio 2.6.6

*Si stabilisca se l'insieme $X = \{g : \mathbb{R} \to \mathbb{R} \mid g$ è continua ma non derivabile in $x = 0\}$ è un sottospazio vettoriale dello spazio vettoriale delle funzioni continue $f : \mathbb{R} \to \mathbb{R}$.*

### Esercizio 2.6.7

*Si scriva, se possibile, l'insieme $S = \{(x, y) \in \mathbb{R}^2 \mid x^2 + xy - 2y^2 = 0\}$ come unione di due sottospazi vettoriali di $\mathbb{R}^2$ e si dica se $S$ è un sottospazio vettoriale di $\mathbb{R}^2$.*

### Esercizio 2.6.8

*Si dice successione a elementi in $\mathbb{R}$ una qualunque funzione $s : \mathbb{N} \to \mathbb{R}$. Se $s(n) = a_n$ la successione si indica anche con $(a_n)$. Sia $\mathcal{S}_{\mathbb{R}}$ l'insieme di tutte le successioni a elementi in $\mathbb{R}$ e su di esso si definiscono le seguenti operazioni:*

$$(a_n) + (b_n) = (a_n + b_n) \qquad k(a_n) = (ka_n)$$

*per ogni $(a_n), (b_n) \in \mathcal{S}_{\mathbb{R}}$ e $k \in \mathbb{R}$. Si dimostri che con queste operazioni $\mathcal{S}_{\mathbb{R}}$ è uno spazio vettoriale su $\mathbb{R}$.*

### Esercizio 2.6.9

*Sia $C(\mathbb{R}; \mathbb{R})$ l'insieme delle funzioni continue da $\mathbb{R}$ in $\mathbb{R}$. Si considerino l'operazione di somma di funzioni e quella di prodotto fra una funzione qualunque e un numero reale, definite come nell'Esempio 2.3.3. Si dimostri che con queste operazioni $C(\mathbb{R}; \mathbb{R})$ è uno spazio vettoriale reale.*

# 3 Combinazioni lineari e lineare indipendenza

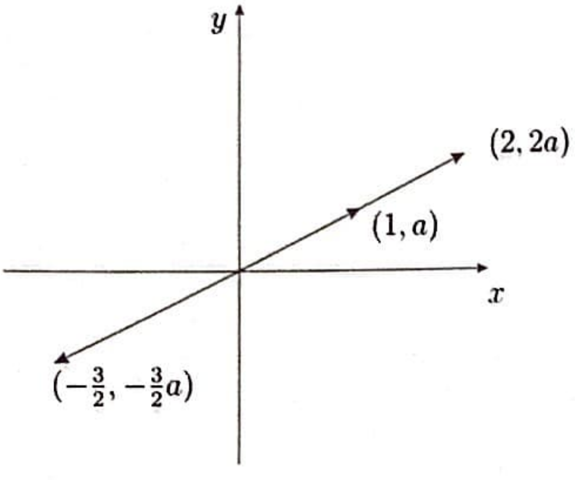
Nel capitolo precedente abbiamo visto la definizione di spazio vettoriale e di sottospazio. Ora vogliamo descrivere questi oggetti in modo più efficiente. Introdurremo a tale scopo il concetto di *combinazione lineare* di un insieme di vettori e il concetto di vettori *linearmente indipendenti*. Si tratta di due definizioni di importanza fondamentale all'interno della teoria degli spazi vettoriali, la cui comprensione è indispensabile per poter procedere oltre e arrivare ai concetti chiave di *base* e *applicazione lineare* che affronteremo nei capitoli successivi.

## 3.1 COMBINAZIONI LINEARI E GENERATORI

Ogni spazio vettoriale $V \neq \{\mathbf{0}\}$ contiene infiniti vettori; infatti basta che $V$ contenga un vettore $\mathbf{v}$ e immediatamente deve contenere anche tutti i suoi multipli, cioè $\lambda\mathbf{v} \in V$ per ogni $\lambda \in \mathbb{R}$. Vediamo un esempio per capire meglio.

Consideriamo il sottospazio vettoriale $W = \{(x, ax) | x \in \mathbb{R}\}$ in $\mathbb{R}^2$ esaminato nella lezione precedente. Esso è rappresentato dalla retta del piano cartesiano di equazione $y = ax$ e possiamo descriverlo in modo alternativo come l'insieme dei multipli del vettore $(1, a)$

$$W = \{x(1, a) \mid x \in \mathbb{R}\}$$



Diciamo allora che il vettore $(1, a)$ *genera* il sottospazio vettoriale $W$ rappresentato dalla retta $y = ax$. La parola "genera" non è casuale in quanto appunto tutti i vettori del sottospazio $W$ sono multipli di $(1, a)$. Notiamo inoltre che la scelta del vettore $(1, a)$ come generatore di $W$ è arbitraria, avremmo potuto anche scegliere un suo qualunque multiplo come $(2, 2a)$ oppure $(-\frac{3}{2}, -\frac{3}{2}a)$.

Graficamente è chiaro che, se conosciamo un punto di una retta (nel piano, ma anche nello spazio tridimensionale) diverso dall'origine, allora possiamo disegnare subito la retta passante per esso e per l'origine. Vedremo in seguito che il fatto di conoscere i generatori di uno spazio vettoriale permette di determinarlo univocamente.

Vediamo ora un altro esempio. In $\mathbb{R}^2$ consideriamo i due vettori $(1, 0)$ e $(0, 1)$. Ci chiediamo: qual è il più piccolo sottospazio vettoriale $W$ di $\mathbb{R}^2$ che contiene entrambi questi vettori? Per i ragionamenti precedenti sappiamo che questo sottospazio vettoriale dovrà contenere i due sottospazi $W_1$ e $W_2$ generati da $(1, 0)$ e $(0, 1)$:

$$W_1 = \{\lambda(1, 0) | \lambda \in \mathbb{R}\} \qquad \text{rappresentato dall'asse } x$$
$$W_2 = \{\mu(0, 1) | \mu \in \mathbb{R}\} \qquad \text{rappresentato dall'asse } y$$

Sappiamo inoltre che la somma di due vettori di $W$ deve ancora appartenere a $W$ (per definizione di sottospazio vettoriale). Per esempio $(1, 0) + (0, 1) = (1, 1) \in W$, ma anche $(1, 2) + (3, 4) = (4, 6) \in W$. Lo studente è invitato a disegnare somme di vettori in $\mathbb{R}^2$ considerando i punti del piano ad essi associati e utilizzando la regola del parallelogramma. In tal modo ci si può convincere che effettivamente $W = \mathbb{R}^2$. La costruzione grafica però non è sufficiente per dimostrare questo fatto, poiché non è possibile disegnare tutti i vettori del piano, vediamo quindi una dimostrazione algebrica. Prendiamo il vettore generico $(\lambda, 0)$ in $W_1$ e il vettore generico $(0, \mu)$ in $W_2$ e facciamone la somma: $(\lambda, 0) + (0, \mu) = (\lambda, \mu)$. È chiaro che *ogni* vettore $(x, y)$ in $\mathbb{R}^2$, si può scrivere in questo modo scegliendo $\lambda = x$ e $\mu = y$. Quindi abbiamo trovato che il più piccolo sottospazio vettoriale di $\mathbb{R}^2$ contenente i vettori $(1, 0)$ e $(0, 1)$ è tutto $\mathbb{R}^2$.

Ora formalizziamo il concetto di *generazione* di un sottospazio vettoriale che abbiamo descritto con gli esempi precedenti.

**Definizione 3.1.1** Siano $V$ uno spazio vettoriale, $\mathbf{v}_1, \ldots, \mathbf{v}_n$ vettori di $V$ e $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Il vettore $\mathbf{w} = \lambda_1\mathbf{v}_1 + \cdots + \lambda_n\mathbf{v}_n$ si dice *combinazione lineare* di $\mathbf{v}_1, \ldots, \mathbf{v}_n$ con scalari $\lambda_1, \ldots, \lambda_n$.

Per esempio $(1, 1)$ è combinazione lineare di $(1, 0)$ e $(0, 1)$ con scalari $\lambda_1 = 1$ e $\lambda_2 = 1$, ma anche combinazione lineare di $(2, 1)$ e $(1, 0)$ con scalari $\lambda_1 = 1$ e $\lambda_2 = -1$.

Veniamo adesso al concetto di spazio vettoriale generato da alcuni vettori, che è il protagonista di questo capitolo insieme al concetto di lineare indipendenza.

**Definizione 3.1.2** Siano $V$ uno spazio vettoriale e $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ un insieme di vettori di $V$. Il *sottospazio generato*[1] dai vettori $\mathbf{v}_1, \ldots, \mathbf{v}_n$ è l'insieme di tutte

[1] Il sottospazio generato dai vettori $\mathbf{v}_1, \ldots, \mathbf{v}_n$ si indica talvolta anche con la notazione $\mathrm{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$.

le loro combinazioni lineari, in simboli

$$\langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle = \{\lambda_1 \mathbf{v}_1 + \cdots + \lambda_n \mathbf{v}_n \mid \lambda_1, \dots, \lambda_n \in \mathbb{R}\}$$

Abbiamo visto che, per esempio, il sottospazio generato da un vettore non nullo in $\mathbb{R}^2$ corrisponde a una retta, mentre il sottospazio generato dai due vettori $(1,0)$ e $(0,1)$ di $\mathbb{R}^2$ è tutto $\mathbb{R}^2$.

♦ **Osservazione 3.1.3** Se $V$ è uno spazio vettoriale e $\mathbf{v} \in V$, allora il sottospazio generato da $\mathbf{v}$ è l'insieme dei multipli di $\mathbf{v}$, cioè $\langle \mathbf{v} \rangle = \{\lambda \mathbf{v} | \lambda \in \mathbb{R}\}$.

Inoltre, il sottospazio generato dal vettore nullo è il sottospazio banale, cioè contiene solo il vettore nullo: $\langle \mathbf{0} \rangle = \{\mathbf{0}\}$.

**Definizione 3.1.4** Siano $V$ uno spazio vettoriale e $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ un insieme di vettori di $V$. Si dice che $\mathbf{v}_1, \dots, \mathbf{v}_n$ *generano* $V$, o che $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ è un *insieme di generatori* di $V$ se $V = \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$.

Nell'esempio visto inizialmente abbiamo che i vettori $(1,0)$ e $(0,1)$ generano lo spazio vettoriale $\mathbb{R}^2$ in quanto ogni vettore $(a,b)$ di $\mathbb{R}^2$ si può scrivere come combinazione lineare di $(1,0)$ e $(0,1)$:

$$(a,b) = a(1,0) + b(0,1)$$

**Proposizione 3.1.5** *Siano $V$ uno spazio vettoriale e $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ un insieme di vettori di $V$. Allora abbiamo che $\langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$ è un sottospazio vettoriale di $V$. Inoltre se $Z$ è un sottospazio vettoriale di $V$ contenente $\mathbf{v}_1, \dots \mathbf{v}_n$, allora $\langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle \subseteq Z$, quindi $\langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$ è il più piccolo sottospazio vettoriale di $V$ contenente $\mathbf{v}_1, \dots, \mathbf{v}_n$.*

**Dimostrazione** – Per prima cosa notiamo che $\mathbf{0} \in \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$, infatti $\mathbf{0} = 0\mathbf{v}_1 + \cdots + 0\mathbf{v}_n$. Siano $\mathbf{v}, \mathbf{w} \in \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$. Allora per definizione esistono degli scalari $\alpha_1, \dots, \alpha_n$ e $\beta_1, \dots, \beta_n$ tali che:

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n, \quad \mathbf{w} = \beta_1 \mathbf{v}_1 + \cdots + \beta_n \mathbf{v}_n$$

e pertanto

$$\mathbf{v} + \mathbf{w} = (\alpha_1 + \beta_1)\mathbf{v}_1 + \cdots + (\alpha_n + \beta_n)\mathbf{v}_n \in \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$$

Inoltre se $k \in \mathbb{R}$

$$k\mathbf{v} = (k\alpha_1)\mathbf{v}_1 + \cdots + (k\alpha_n)\mathbf{v}_n \in \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$$

Questo dimostra che $\langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$ è un sottospazio vettoriale di $V$.

Vediamo ora che $\langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$ è il più piccolo sottospazio vettoriale contenente $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. Siano $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ e sia $\mathbf{v} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_n \mathbf{v}_n \in \langle \mathbf{v}_1 \dots \mathbf{v}_n \rangle$, sia inoltre $Z$ un sottospazio vettoriale di $V$ contenente $\mathbf{v}_1, \dots \mathbf{v}_n$. Allora $Z$ contiene anche $\lambda_1 \mathbf{v}_1, \dots, \lambda_n \mathbf{v}_n$, perché essendo uno spazio vettoriale se contiene un vettore, contiene anche tutti i suoi multipli. Inoltre, poiché è chiuso rispetto alla somma, contiene anche $\lambda_1\mathbf{v}_1 + \cdots + \lambda_n\mathbf{v}_n = \mathbf{v}$. Quindi $\langle \mathbf{v}_1, \ldots, \mathbf{v}_n \rangle \subseteq Z$. □

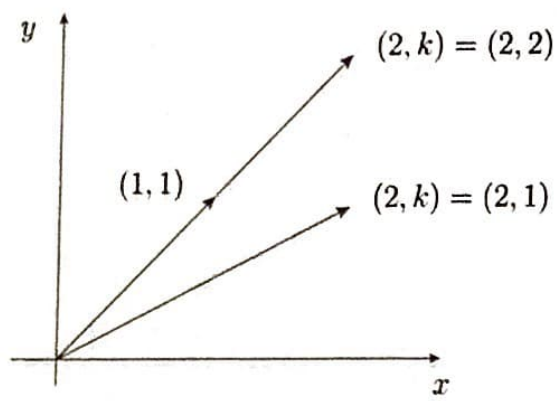
Consideriamo ora un esempio che si riallaccia a quanto abbiamo visto nel Capitolo 1 a proposito della soluzione di sistemi lineari dipendenti da un parametro.

**Esempi 3.1.6**

Vogliamo determinare il sottospazio generato dai vettori $(1,1)$, $(2,k)$ al variare del parametro $k$.

$$\begin{aligned}\langle (1,1),(2,k)\rangle &= \{\lambda_1(1,1) + \lambda_2(2,k) | \lambda_1, \lambda_2 \in \mathbb{R}\}\\ &= \{(\lambda_1 + 2\lambda_2, \lambda_1 + k\lambda_2) | \lambda_1, \lambda_2 \in \mathbb{R}\}\end{aligned}$$

Dato che stiamo considerando vettori di $\mathbb{R}^2$ possiamo rappresentare i vettori tramite punti del piano cartesiano. Il disegno seguente illustra i vettori $(1,1)$ e $(2,k)$ per i valori $k = 1$ e $k = 2$.




Vediamo subito che, se $k = 2$, allora i due punti giacciono sulla stessa retta per l'origine, perciò il più piccolo sottospazio che li contiene entrambi sarà appunto tale retta e cioè la retta di equazione $y = x$.

Se invece $k \neq 2$, i due punti giacciono su due rette distinte passanti per l'origine, quindi il più piccolo sottospazio che li contiene entrambi deve contenere tali rette, e anche la somma di due punti qualsiasi su tali rette, per cui, con un ragionamento analogo a quello fatto all'inizio di questo capitolo, si ha che il più piccolo sottospazio che contiene entrambi i punti di coordinate $(1,1)$, $(2,k)$ è tutto il piano, cioè i vettori $(1,1)$, $(2,k)$ generano $\mathbb{R}^2$.

Vediamo ora una dimostrazione algebrica di questo fatto. Sia $(a,b)$ un generico vettore di $\mathbb{R}^2$, ci chiediamo quando $(a,b)$ appartiene a $\langle (1,1),(2,k)\rangle$, cioè quando esistono $\lambda_1, \lambda_2 \in \mathbb{R}$ tali che

$$(\lambda_1 + 2\lambda_2, \lambda_1 + k\lambda_2) = (a,b)$$

In altre parole dobbiamo risolvere il sistema lineare:

$$\begin{cases} \lambda_1 + 2\lambda_2 = a \\ \lambda_1 + k\lambda_2 = b \end{cases}$$

Lasciamo per esercizio la verifica che questo sistema nelle incognite $\lambda_1, \lambda_2$ ammette sempre soluzione se $k \neq 2$. Se invece $k = 2$ la matrice completa associata al sistema è:

$$\left(\begin{array}{cc|c} 1 & 2 & a \\ 1 & 2 & b \end{array}\right)$$

che ridotta a scala diventa

$$\left(\begin{array}{cc|c} 1 & 2 & a \\ 0 & 0 & b-a \end{array}\right)$$

Dunque, se $a \neq b$ il sistema non ammette soluzioni, cioè si ha che $(a, b) \notin \langle(1,1),(2,2)\rangle$, se invece $a = b$ il sistema ammette soluzioni, cioè $(a, a) \in \langle(1,1),(2,2)\rangle$. Quindi $\langle(1,1),(2,2)\rangle$ è l'insieme dei vettori che hanno prima coordinata uguale alla seconda, cioè $\langle(1,1),(2,2)\rangle = \{(a,a) | a \in \mathbb{R}\}$.

**Esempio 3.1.7**

Modifichiamo ora leggermente l'esempio precedente. Vogliamo determinare il sottospazio generato dai vettori: $(1,1)$, $(2,k)$, $(-1,-1)$ al variare del parametro $k$.

$$\begin{aligned}\langle(1,1),(2,k),(-1,-1)\rangle &= \{\lambda_1(1,1) + \lambda_2(2,k) + \lambda_3(-1,-1) | \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}\} \\ &= \{(\lambda_1 + 2\lambda_2 - \lambda_3, \lambda_1 + k\lambda_2 - \lambda_3) | \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}\}\end{aligned}$$

Ci aspettiamo, dal ragionamento precedente, che questi vettori generino quasi sempre $\mathbb{R}^2$ (cioè per quasi ogni valore di $k$). Ma vediamo una dimostrazione rigorosa. Vogliamo mostrare che per ogni vettore $(a, b)$ fissato possiamo sempre scegliere $\lambda_1$, $\lambda_2$, $\lambda_3$ tali che

$$(\lambda_1 + 2\lambda_2 - \lambda_3, \lambda_1 + k\lambda_2 - \lambda_3) = (a, b)$$

Un facile calcolo di risoluzione dei sistemi lineari dipendenti da un parametro con l'algoritmo di Gauss mostra che questo sistema ammette sempre soluzione (per ogni $a$ e $b$ fissati) purché sia $k \neq 2$. Quando invece $k = 2$ abbiamo:

$$(\lambda_1 + 2\lambda_2 - \lambda_3, \lambda_1 + 2\lambda_2 - \lambda_3) = (a, b)$$

dunque, affinché il sistema ammetta soluzione, deve necessariamente essere $a = b$, perciò gli unici vettori che appartengono al sottospazio generato dai due vettori dati sono quelli del tipo $(a, a)$, e cioè quelli corrispondenti ai punti che giacciono sulla retta di equazione $y = x$. In realtà un'accurata analisi avrebbe potuto dare subito la risposta a questo problema senza alcun calcolo: infatti sarebbe bastato notare che il vettore $(-1,-1)$ era *superfluo* nel calcolo del sottospazio generato, in quanto appartenente al sottospazio generato da $(1,1)$. Graficamente, è chiaro che il punto $P$ di coordinate $(-1,-1)$ appartiene alla retta passante per l'origine e per il punto $Q$ di coordinate $(1,1)$, quindi ogni sottospazio del piano che contiene $Q$ contiene automaticamente anche $P$. Dunque avremmo potuto tranquillamente ignorare $(-1,-1)$ e dare subito come risposta la soluzione dell'Esempio 3.1.6.

Abbiamo dunque visto che, nel descrivere un sottospazio vettoriale usando un insieme di generatori, alcuni vettori sono *superflui*, cioè anche eliminandoli il sottospazio generato non cambia. Ciò accade per esempio quando abbiamo un vettore multiplo di un altro, ma anche quando un vettore è la somma di altri due. Per esempio abbiamo visto che $\langle(1,0),(0,1)\rangle = \mathbb{R}^2$, ma anche (come lo studente può direttamente verificare):

$$\langle(1,0),(0,1),(1,1)\rangle = \mathbb{R}^2$$

Questo fatto è formalizzato dalla seguente proposizione.

**Proposizione 3.1.8** *Siano $V$ uno spazio vettoriale, $\mathbf{v}_1, \ldots, \mathbf{v}_n$ vettori di $V$ e $\mathbf{w}$ una loro combinazione lineare, cioè: $\mathbf{w} = \lambda_1\mathbf{v}_1 + \cdots + \lambda_n\mathbf{v}_n$. Allora*

$$\langle\mathbf{v}_1, \ldots, \mathbf{v}_n\rangle = \langle\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{w}\rangle$$

*Viceversa se*

$$\langle \mathbf{v}_1, \ldots, \mathbf{v}_n \rangle = \langle \mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{w} \rangle$$

*allora* $\mathbf{w}$ *è combinazione lineare di* $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

**Dimostrazione** – Per mostrare la prima affermazione è sufficiente osservare che per ipotesi $\mathbf{w} \in \langle \mathbf{v}_1, \ldots, \mathbf{v}_n \rangle$, quindi dalla Proposizione 3.1.5 segue che $Z = \langle \mathbf{v}_1, \ldots, \mathbf{v}_n \rangle$ è un sottospazio che contiene $\{\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{w}\}$, allora $\langle \mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{w} \rangle \subseteq \langle \mathbf{v}_1, \ldots, \mathbf{v}_n \rangle$ sempre per la Proposizione 3.1.5. L'inclusione opposta è ovvia.

Per mostrare la seconda affermazione, basta notare che, siccome si ha che $\langle \mathbf{v}_1, \ldots, \mathbf{v}_n \rangle = \langle \mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{w} \rangle$, segue in particolare che $\mathbf{w} \in \langle \mathbf{v}_1, \ldots, \mathbf{v}_n \rangle$, cioè $\mathbf{w}$ è combinazione lineare di $\mathbf{v}_1, \ldots, \mathbf{v}_n$. □

## 3.2 INDIPENDENZA LINEARE

Il problema che ci poniamo ora è: come facciamo a stabilire quali sono i vettori "superflui" nella descrizione del sottospazio generato da un insieme di vettori? Se vogliamo essere efficienti nella descrizione di un sottospazio vettoriale, dobbiamo poterlo descrivere come il sottospazio generato dal numero minimo di vettori possibile. La risposta a questa domanda viene dal concetto di *indipendenza lineare*. Questo è di gran lunga il concetto più difficile da digerire ma è il pilastro su cui si basa l'intera teoria che noi svilupperemo. In pillole la storia è: se un insieme di generatori di un sottospazio è un insieme di vettori linearmente indipendenti, allora siamo sicuri che utilizzare questo insieme di vettori come generatori del sottospazio è il modo più efficiente di descrivere tale sottospazio, cioè questo insieme è formato dal minimo numero di vettori necessari per descrivere il sottospazio. Vediamo allora la definizione, poi, con una serie di piccoli passi e lemmi arriveremo a dimostrare quanto asserito sopra.

**Definizione 3.2.1** Sia $V$ uno spazio vettoriale. I vettori $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ si dicono *linearmente indipendenti* se per ogni combinazione lineare

$$\lambda_1 \mathbf{v}_1 + \cdots + \lambda_n \mathbf{v}_n = \mathbf{0}$$

abbiamo $\lambda_1 = \cdots = \lambda_n = 0$. In altre parole, l'unica combinazione lineare dei vettori $\mathbf{v}_1, \ldots, \mathbf{v}_n$ uguale al vettore nullo è quella con scalari tutti nulli. Diremo anche che l'insieme dei vettori $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ è linearmente indipendente[2].

I vettori $\mathbf{v}_1, \ldots, \mathbf{v}_n$ si dicono *linearmente dipendenti* se non sono indipendenti. In altre parole, i vettori dell'insieme $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ sono linearmente dipendenti se esistono scalari $\lambda_1, \ldots, \lambda_n$ non tutti nulli tali che $\lambda_1 \mathbf{v}_1 + \cdots + \lambda_n \mathbf{v}_n = \mathbf{0}$.

Rivediamo gli esempi precedenti. L'insieme di vettori $\{(1,0),(0,1)\}$ in $\mathbb{R}^2$ è un insieme di vettori linearmente indipendenti, infatti l'unica combinazione lineare

[2] Le parole "linearmente indipendente/i" possono essere usate per i vettori $\mathbf{v}_1, \ldots, \mathbf{v}_n$, ma anche per l'insieme dei vettori $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ indifferentemente, cioè le due terminologie hanno lo stesso significato.

che dà il vettore nullo è quella ottenuta con scalari tutti nulli:

$$\alpha(1,0) + \beta(0,1) = (\alpha,\beta) = (0,0)$$

solo se $\alpha = \beta = 0$.

Invece i vettori dell'insieme $\{(1,0),(0,1),(1,1)\}$ sono linearmente dipendenti poiché esiste una combinazione lineare dei vettori dati con scalari non tutti nulli che è uguale al vettore nullo.

$$1\cdot(1,0) + 1\cdot(0,1) - 1\cdot(1,1) = (0,0)$$

Vediamo un esempio più complesso.

**Esempio 3.2.2**

Consideriamo il seguente insieme di vettori in $\mathbb{R}_2[x]$: $\{x+1, x^2-1, 2, x-1\}$. Ci chiediamo: è un insieme di vettori linearmente indipendenti? Se conoscessimo qualche cosa in più di algebra lineare la risposta sarebbe immediata, per il momento dobbiamo fare i calcoli.

Scriviamo una loro combinazione lineare generica e poniamola uguale al vettore nullo:

$$\alpha_1(x+1) + \alpha_2(x^2-1) + 2\alpha_3 + \alpha_4(x-1) = 0$$

Da cui:

$$\alpha_2 x^2 + (\alpha_1 + \alpha_4)x + (\alpha_1 - \alpha_2 + 2\alpha_3 - \alpha_4) = 0$$

Poiché un polinomio è nullo se e solo se tutti i suoi coefficienti sono nulli, ricaviamo il sistema lineare:

$$\begin{cases} \alpha_2 = 0 \\ \alpha_1 + \alpha_4 = 0 \\ \alpha_1 - \alpha_2 + 2\alpha_3 - \alpha_4 = 0 \end{cases}$$

Lasciamo allo studente per esercizio la verifica del fatto che questo sistema ammette infinite soluzioni. Per esempio ammette la soluzione: $\alpha_1 = 1$, $\alpha_2 = 0$, $\alpha_3 = -1$, $\alpha_4 = -1$. Dunque possiamo scrivere esplicitamente una combinazione lineare dei vettori dati che dia il vettore nullo e non abbia scalari tutti nulli:

$$1\cdot(x+1) - 1\cdot 2 - 1\cdot(x-1) = 0$$

Quindi i vettori dell'insieme dato sono linearmente dipendenti.

Quando abbiamo una combinazione lineare di questo tipo possiamo sempre scrivere uno dei vettori come combinazione lineare degli altri, nel nostro caso si ha:

$$(x+1) = 2 + (x-1)$$

Naturalmente in un insieme di vettori linearmente dipendenti non tutti i vettori dati possono essere espressi in funzione degli altri, per esempio vediamo che non c'è modo di esprimere $x^2-1$ come combinazione lineare degli altri.

La cosa importante da notare è che, in un insieme di vettori linearmente dipendenti eliminandone uno che sia combinazione lineare degli altri, il sottospazio da essi generato non cambia (per la Proposizione 3.1.8) e i vettori del nuovo insieme ottenuto potrebbero essere diventati linearmente indipendenti. Attenzione però che ciò non è detto, per esempio nell'insieme $\{2x, 3x, 4x\}$, anche eliminando un vettore, i vettori rimanenti restano linearmente dipendenti, come lo studente può verificare. In qualche modo, la lineare indipendenza ci dice che abbiamo raggiunto il numero minimo di vettori necessari a descrivere il sottospazio generato. Questo concetto sarà esplorato con molta attenzione nel capitolo successivo sulle basi.

♦ **Osservazione 3.2.3** Notiamo che, se un insieme di vettori contiene il vettore nullo, allora è sempre un insieme di vettori linearmente dipendenti. Infatti, se consideriamo l'insieme $\{\mathbf{v}_1 = 0, \mathbf{v}_2 \ldots, \mathbf{v}_n\}$, si ha che:

$$1\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_n = 0$$

Dunque abbiamo ottenuto una combinazione lineare uguale al vettore nullo in cui il primo scalare è non nullo.

**Proposizione 3.2.4** *In uno spazio vettoriale $V$ i vettori $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sono linearmente dipendenti se e solo se almeno uno di essi è combinazione lineare degli altri.*

**Dimostrazione** – Supponiamo che $\mathbf{v}_1, \ldots, \mathbf{v}_n$ siano linearmente dipendenti. Allora esistono degli scalari $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$, non tutti nulli, tali che

$$\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n = 0$$

Dato che almeno uno degli scalari è non nullo, supponiamo $\alpha_k \neq 0$. Allora:

$$\mathbf{v}_k = -\frac{\alpha_1}{\alpha_k}\mathbf{v}_1 - \cdots - \frac{\alpha_{k-1}}{\alpha_k}\mathbf{v}_{k-1} - \frac{\alpha_{k+1}}{\alpha_k}\mathbf{v}_{k+1} - \cdots - \frac{\alpha_n}{\alpha_k}\mathbf{v}_n$$

e dunque $\mathbf{v}_k$ è combinazione lineare degli altri vettori.

Viceversa supponiamo che esistano degli scalari $\alpha_1, \ldots, \alpha_{k-1}, \alpha_{k+1}, \ldots, \alpha_n$ tali che

$$\mathbf{v}_k = \alpha_1\mathbf{v}_1 + \cdots + \alpha_{k-1}\mathbf{v}_{k-1} + \alpha_{k+1}\mathbf{v}_{k+1} + \cdots + \alpha_n\mathbf{v}_n$$

allora si ottiene che

$$\alpha_1\mathbf{v}_1 + \cdots + \alpha_{k-1}\mathbf{v}_{k-1} + (-1)\mathbf{v}_k + \alpha_{k+1}\mathbf{v}_{k+1} + \cdots + \alpha_n\mathbf{v}_n = 0$$

e almeno uno dei coefficienti è non nullo, quello di $\mathbf{v}_k$. Quindi i vettori $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sono linearmente dipendenti. □

Vediamo un caso particolare di questa proposizione, molto utile negli esercizi.

**Proposizione 3.2.5** *Due vettori sono linearmente indipendenti se e solo se non sono uno multiplo dell'altro.*

La dimostrazione di questa proposizione è immediata: basta utilizzare la proposizione precedente e ricordare che un vettore è combinazione lineare di un altro se e solo se è un suo multiplo.

♦ **Osservazione 3.2.6** Affinché i vettori di un insieme siano linearmente dipendenti è sufficiente trovare un vettore che sia combinazione lineare degli altri. Per esempio, se vediamo che un vettore è multiplo di un altro, allora sappiamo subito che i vettori dell'insieme sono linearmente dipendenti. I vettori dei seguenti insiemi sono linearmente dipendenti e possiamo verificarlo senza fare alcun calcolo (ma lo studente dovrebbe farli se non vede il perché e vuole convincersene!).

- In $\mathbb{R}^3$: $\{(1,0,0),(0,1,0),(0,0,1),(0,2,0)\}$;
- in $\mathbb{R}^3$: $\{(1,0,0),(0,1,0),(-1,2,0),(1,2,1)\}$;
- in $\mathbb{R}_3[x]$: $\{2, x+7, x^3-3x, 1\}$;
- in $\mathbb{R}_3[x]$: $\{0, x, 1-x, x^3\}$;
- in $\mathrm{M}_2(\mathbb{R})$:

$$\left\{\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}0 & 0\\ 0 & 3\end{pmatrix}, \begin{pmatrix}0 & 0\\ 0 & \sqrt{2}\end{pmatrix}, \begin{pmatrix}3 & 0\\ 0 & 3\end{pmatrix}\right\}$$

♦ **Osservazione 3.2.7** Anche se è vero che in un insieme di vettori basta trovare che uno è multiplo di un altro affinché i vettori siano linearmente dipendenti, *non è vero il contrario*! Come mostra l'Esempio 3.2.2, possiamo avere un insieme di vettori linearmente dipendenti in cui *nessuno è multiplo di un altro*.

La prossima proposizione ci mostra che invece togliendo alcuni vettori a un insieme di vettori linearmente indipendenti, si ottiene un insieme di vettori che sono ancora linearmente indipendenti.

**Proposizione 3.2.8** *Un sottoinsieme non vuoto di un insieme di vettori linearmente indipendenti è formato da vettori che sono ancora linearmente indipendenti.*

**Dimostrazione** – Supponiamo, per assurdo, che $I$ sia un insieme di vettori linearmente indipendenti e che $\varnothing \neq J \subseteq I$ sia un sottoinsieme di vettori dipendenti. Allora esiste un vettore in $J$ che si scrive come combinazione lineare degli altri. Ma allora si esprime anche come combinazione lineare dei vettori di $I$ e quindi $I$ è un insieme di vettori linearmente dipendenti, contraddicendo l'ipotesi. □

---

**Esempio 3.2.9**

Supponiamo di voler stabilire per quali valori di $k$ i seguenti vettori in $\mathrm{M}_2(\mathbb{R})$:

$$\begin{pmatrix}1 & 0\\ 1 & -3\end{pmatrix}, \begin{pmatrix}6 & 0\\ k & -18\end{pmatrix}, \begin{pmatrix}k & 0\\ 1 & 5\end{pmatrix}$$

sono linearmente indipendenti. Procediamo come suggerisce la definizione: scriviamo una loro combinazione lineare e vediamo se ci sono scalari non nulli che ci permettono di ottenere il vettore nullo. In corrispondenza dei valori di $k$ per cui ciò accade avremo che i vettori dati sono linearmente dipendenti.

Scriviamo dunque una combinazione lineare generica dei vettori dati e poniamola uguale al vettore nullo:

$$\lambda_1\begin{pmatrix}1 & 0\\ 1 & -3\end{pmatrix} + \lambda_2\begin{pmatrix}6 & 0\\ k & -18\end{pmatrix} + \lambda_3\begin{pmatrix}k & 0\\ 1 & 5\end{pmatrix} = \begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix} \tag{3.1}$$

da cui:

$$\begin{pmatrix} \lambda_1 + 6\lambda_2 + k\lambda_3 & 0 \\ \lambda_1 + k\lambda_2 + \lambda_3 & -3\lambda_1 - 18\lambda_2 + 5\lambda_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Quindi l'uguaglianza (3.1) è soddisfatta se e solo se $\lambda_1, \lambda_2, \lambda_3$ sono soluzioni del sistema lineare omogeneo:

$$\begin{cases} \lambda_1 + 6\lambda_2 + k\lambda_3 = 0 \\ \lambda_1 + k\lambda_2 + \lambda_3 = 0 \\ -3\lambda_1 - 18\lambda_2 + 5\lambda_3 = 0 \end{cases}$$

La matrice completa associata al sistema è:

$$(A|b) = \left(\begin{array}{ccc|c} 1 & 6 & k & 0 \\ 1 & k & 1 & 0 \\ -3 & -18 & 5 & 0 \end{array}\right)$$

che ridotta a scala con l'algoritmo di Gauss diventa:

$$(A'|\underline{b}') = \left(\begin{array}{ccc|c} 1 & 6 & k & 0 \\ 0 & k-6 & 1-k & 0 \\ 0 & 0 & 5+3k & 0 \end{array}\right)$$

Si noti che il sistema ammette sempre la soluzione nulla $\lambda_1 = \lambda_2 = \lambda_3 = 0$. Si tratta di capire se ci sono anche soluzioni non nulle oppure no. La forma della matrice scala per righe si presta molto bene a capire se abbiamo o meno soltanto la soluzione nulla. Possiamo subito osservare che se i vettori di partenza fossero stati 5, certamente almeno una delle 5 incognite (gli scalari che danno la combinazione lineare nulla) sarebbe stata indeterminata, in altre parole i vettori sarebbero stati sicuramente linearmente dipendenti, perché avremmo potuto attribuire arbitrariamente un valore non nullo a tale incognita. Nel prossimo capitolo, utilizzando il concetto di base, potremo formalizzare questo ragionamento, che tuttavia anche ora dovrebbe essere chiaro intuitivamente.

Tornando all'esempio in questione, se $k \neq 6$ e $k \neq -\frac{3}{5}$, $\mathrm{rr}(A') = \mathrm{rr}(A'|\underline{b}') = 3$ è uguale al numero delle incognite, il sistema ammette un'unica soluzione, $\lambda_1 = \lambda_2 = \lambda_3 = 0$ e i vettori dati sono dunque linearmente indipendenti.

Se $k = 6$, riducendo ulteriormente a scala la matrice, si ottiene:

$$(A''|\underline{b}'') = \left(\begin{array}{ccc|c} 1 & 6 & 6 & 0 \\ 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

quindi $\mathrm{rr}(A'') = \mathrm{rr}(A''|\underline{b}'') = 2$ e il sistema ammette infinite soluzioni dipendenti da un paramentro. In particolare ci sono soluzioni non nulle e i vettori dati sono linearmente dipendenti.

Infine, se $k = -\frac{3}{5}$ si ottiene:

$$(A'|\underline{b}') = \left(\begin{array}{ccc|c} 1 & 6 & -\frac{3}{5} & 0 \\ 0 & -\frac{33}{5} & -\frac{8}{5} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Quindi $\mathrm{rr}(A') = \mathrm{rr}(A'|\underline{b}') = 2$ e come prima i vettori dati sono linearmente dipendenti.

---

Concludiamo questo capitolo con alcuni esercizi svolti, che chiariscono la meccanica della verifica della dipendenza o indipendenza lineare e del concetto di generatore.

## 3.3 ESERCIZI SVOLTI

### Esercizio 3.3.1

*Si stabilisca per quali valori di $k$ i vettori $x^2+2x+k, 5x^2+2kx+k^2, kx^2+x+3$ generano $\mathbb{R}_2[x]$.*

**Svolgimento**

Prendiamo un vettore $ax^2+bx+c \in \mathbb{R}_2[x]$ e ci chiediamo quando $ax^2+bx+c \in \langle x^2+2x+k, 5x^2+2kx+k^2, kx^2+x+3\rangle$. Questo succede se si ha che

$$ax^2+bx+c = \lambda_1(x^2+2x+k)+\lambda_2(5x^2+2kx+k^2)+\lambda_3(kx^2+x+3)$$

per qualche $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$, cioè se

$$ax^2+bx+c = (\lambda_1+5\lambda_2+k\lambda_3)x^2+(2\lambda_1+2k\lambda_2+\lambda_3)x+(k\lambda_1+k^2\lambda_2+3\lambda_3)$$

Quindi $\lambda_1, \lambda_2, \lambda_3$ devono essere soluzioni del sistema:

$$\begin{cases} \lambda_1+5\lambda_2+k\lambda_3 = a \\ 2\lambda_1+2k\lambda_2+\lambda_3 = b \\ k\lambda_1+k^2\lambda_2+3\lambda_3 = c \end{cases}$$

La matrice completa associata al sistema è:

$$(A|\underline{b}) = \left(\begin{array}{ccc|c} 1 & 5 & k & a \\ 2 & 2k & 1 & b \\ k & k^2 & 3 & c \end{array}\right)$$

che, ridotta a scala con l'algoritmo di Gauss, diventa:

$$(A'|\underline{b}') = \left(\begin{array}{ccc|c} 1 & 5 & k & a \\ 0 & 2k-10 & 1-2k & b-2a \\ 0 & 0 & 3-\frac{1}{2}k & c-\frac{b}{2}k \end{array}\right)$$

Se $k \neq 5$ e $k \neq 6$ si ha che $\operatorname{rr}(A') = \operatorname{rr}(A'|\underline{b}') = 3$, quindi il sistema ammette soluzione, indipendentemente da quali siano $a, b, c$, perciò ogni vettore del tipo $ax^2+bx+c$ appartiene a $\langle x^2+2x+k, 5x^2+2kx+k^2, kx^2+x+3\rangle$ e i vettori dati generano $\mathbb{R}_2[x]$.

Se $k = 5$ riducendo ulteriormente a scala la matrice si ottiene:

$$(A''|\underline{b}'') = \left(\begin{array}{ccc|c} 1 & 5 & 5 & a \\ 0 & 0 & 1 & 2c-5b \\ 0 & 0 & 0 & -a-22b+9c \end{array}\right)$$

e, se $-a-22b+9c \neq 0$, il sistema non ammette soluzione, perché $3 = \operatorname{rr}(A'') \neq \operatorname{rr}(A''|\underline{b}'') = 2$. Questo significa che, se $-a-22b+9c \neq 0$, non è possibile scrivere $ax^2+bx+c$ come combinazione lineare di $x^2+2x+k, 5x^2+2kx+k^2, kx^2+x+3$, quindi i vettori dati non generano $\mathbb{R}_2[x]$. Per esempio il vettore $x^2-x+1 \notin \langle x^2+2x+k, 5x^2+2kx+k^2, kx^2+x+3\rangle$.

Se $k = 6$ si ottiene

$$(A'|\underline{b}') = \left(\begin{array}{ccc|c} 1 & 5 & 6 & a \\ 0 & 2 & -11 & b-2a \\ 0 & 0 & 0 & c-3b \end{array}\right)$$

e, se $c \neq 3b$, il sistema non ammette soluzione, quindi i vettori dati non generano $\mathbb{R}_2[x]$. Per esempio $3x^2 + 5x - 8 \notin \langle x^2 + 2x + k, 5x^2 + 2kx + k^2, kx^2 + x + 3 \rangle$. ∎

### Esercizio 3.3.2

*Sia $W$ il sottospazio vettoriale di $\mathbb{R}^4$ dato dall'insieme delle soluzioni del sistema lineare omogeneo:*

$$\begin{cases} x_1 + x_2 - x_4 = 0 \\ 2x_1 + x_2 - x_3 + 3x_4 = 0 \end{cases}$$

*nelle incognite $x_1, x_2, x_3, x_4$. Si determini, se possibile, un insieme finito di generatori di $W$.*

W è l'insieme delle sol. del sistema lin. om. $A\underline{x} = \underline{0}$, con matrice →

**Svolgimento**

La matrice completa associata al sistema è

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

l'insieme delle sol. di un qualunque sistema lineare omogeneo è sempre un sottospazio

che, ridotta a scala con l'algoritmo di Gauss, diventa:

$$R_2 - 2R_1 \qquad (A'|\underline{b}') = \left(\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 0 & -1 & -1 & 5 & 0 \end{array}\right)$$

$rr(A) = 2 = rr(A|b)$
$\Rightarrow \infty$ sol che dipendono da 4 - 2 = 2 param.

Le soluzioni del sistema sono: $(x_3 - 4x_4, -x_3 + 5x_4, x_3, x_4)$, con $x_3, x_4 \in \mathbb{R}$.

Per determinare i generatori di $W$ separiamo le variabili libere. Quindi

$$\begin{aligned} W &= \{(x_3 - 4x_4, -x_3 + 5x_4, x_3, x_4) | x_3, x_4 \in \mathbb{R}\} \\ &= \{(x_3, -x_3, x_3, 0) + (-4x_4, 5x_4, 0, x_4) | x_3, x_4 \in \mathbb{R}\} \\ &= \{x_3 \underbrace{(1, -1, 1, 0)}_{v_1} + x_4 \underbrace{(-4, 5, 0, 1)}_{v_2} | x_3, x_4 \in \mathbb{R}\} \end{aligned}$$

$\langle (1,-1,1,0), (-4,5,0,1) \rangle$

$x_4 = s \quad x_3 = t$
$x_2 = -x_3 + 5x_4 = -t + 5s$
$x_1 = x_3 - 4x_4 = t - 4s$

Ricordiamo che $x_3, x_4 \in \mathbb{R}$ possono assumere qualsiasi valore reale. A questo punto è chiaro che $W = \langle (1, -1, 1, 0), (-4, 5, 0, 1) \rangle$, cioè i vettori $(1, -1, 1, 0)$, $(-4, 5, 0, 1)$ generano $W$. ∎

cerchiamo una base → $v_1, v_2$ sono indipendenti (non sono uno multiplo dell'altro)
$\Rightarrow v_1, v_2$ è BAse di $W$ $\Rightarrow$ la dimensione di $W = 2$

### Esercizio 3.3.3

*Si stabilisca per quali valori di $k$ i seguenti vettori di $\mathbb{R}^4$ sono linearmente dipendenti:*

$$\mathbf{v}_1 = (2, 2k, k^2, 2k + 2), \quad \mathbf{v}_2 = (-1, -k, 2k + 2, -k - 1)$$

*Scelto a piacere uno di tali valori di $k$, si mostri che $\mathbf{v}_1 \in \langle \mathbf{v}_2 \rangle$.*

**Svolgimento**

Dobbiamo capire se esistono due scalari $\lambda_1, \lambda_2$ non entrambi nulli tali che $\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 = \mathbf{0}$. Deve succedere che:

$$\lambda_1(2, 2k, k^2, 2k + 2) + \lambda_2(-1, -k, 2k + 2, -k - 1) = (0, 0, 0, 0)$$

cioè

$$(2\lambda_1 - \lambda_2, 2k\lambda_1 - k\lambda_2, k^2\lambda_1 + (2k + 2)\lambda_2, (2k + 2)\lambda_1 - (k + 1)\lambda_2) = (0, 0, 0, 0)$$

ovverosia $\lambda_1, \lambda_2$ devono soddisfare il seguente sistema lineare omogeneo:

$$\begin{cases} 2\lambda_1 - \lambda_2 = 0 \\ 2k\lambda_1 - k\lambda_2 = 0 \\ k^2\lambda_1 + (2k+2)\lambda_2 = 0 \\ (2k+2)\lambda_1 - (k+1)\lambda_2 = 0 \end{cases}$$

La matrice completa associata al sistema è:

$$(A|b) = \left(\begin{array}{rr|r} 2 & -1 & 0 \\ 2k & -k & 0 \\ k^2 & 2k+2 & 0 \\ 2k+2 & -k-1 & 0 \end{array}\right)$$

che, ridotta a scala con l'algoritmo di Gauss, diventa:

$$(A'|\underline{b}') = \left(\begin{array}{lr|l} 2 & -1 & 0 \\ 0 & (k+2)^2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

Se $k \neq -2$ si ha che $\mathrm{rr}(A') = \mathrm{rr}(A'|\underline{b}') = 2$, quindi il sistema ha una sola soluzione, che è $\lambda_1 = \lambda_2 = 0$, i vettori dati sono perciò linearmente indipendenti.

Per $k = -2$ si ottiene:

$$(A'|\underline{b}') = \left(\begin{array}{rr|r} 2 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

quindi $\mathrm{rr}(A') = \mathrm{rr}(A'|\underline{b}') = 1$. Il sistema ha infinite soluzioni che dipendono da un parametro e i vettori dati sono linearmente dipendenti.

Per $k = -2$ in effetti si ha che:

$$\mathbf{v}_1 = (2, -4, 4, -2), \mathbf{v}_2 = (-1, 2, -2, -1)$$

e $\mathbf{v}_1 = -2\mathbf{v}_2$. ∎

## 3.4 ESERCIZI PROPOSTI

### Esercizio 3.4.1

*Si dica se i vettori dei seguenti insiemi sono linearmente dipendenti o indipendenti e, nel caso siano dipendenti, si scriva un vettore come combinazione lineare degli altri:*

(a) $\{(2,1,1), (3,2,1), (6,2,2)\} \subseteq \mathbb{R}^3$;

(b) $\{(1,0,1,4), (2,1,0,6), (1,-2,5,8)\} \subseteq \mathbb{R}^4$;

(c) *L'insieme dei polinomi* $\{2x^2+1, x^2+2x, 4x-1\} \subseteq \mathbb{R}_2[x]$;

(d) *L'insieme dei polinomi* $\{1, x, x^2, x^3\} \subseteq \mathbb{R}[x]$;

(c) *L'insieme delle seguenti matrici:*

$$\left\{\begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 0 & 3 \\ 1 & -3 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ -2 & -1 \end{pmatrix}\right\} \subseteq \mathrm{M}_2(\mathbb{R})$$

**Esercizio 3.4.2**

*Dati i vettori*

$$\mathbf{u} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \ \mathbf{v} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \ \mathbf{w} = \begin{pmatrix} 2 \\ -2 \end{pmatrix}$$

*si stabilisca se sono linearmente indipendenti e si determini il sottospazio da essi generato.*

**Esercizio 3.4.3**

*Si stabilisca se* $x^3 - x$ *appartiene* $\langle x^3 + x^2 + x, x^2 + 2x, x^2 \rangle$.

**Esercizio 3.4.4**

*Si stabilisca per quali valori di* $k$ *il polinomio* $k^2x^2 + x + 1$ *appartiene a* $\langle 2x^2 - x, -x^2 + 3x + 1 \rangle$.

**Esercizio 3.4.5**

*Si stabilisca per quali valori di* $k$ *si ha che*

$$\mathbf{w} = \begin{pmatrix} 2 \\ 5 \end{pmatrix} \in \left\langle \begin{pmatrix} k \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \end{pmatrix} \right\rangle$$

**Esercizio 3.4.6**

*Si stabilisca per quali valori di* $k$ *il polinomio* $x^2 + 2k$ *appartiene a* $\langle x^2 + kx, x^2 - (k+1)x - k \rangle$.

**Esercizio 3.4.7**

*In* $\mathbb{R}_2[x]$ *si diano esempi dei seguenti insiemi:*

a) *un insieme di generatori che non siano linearmente indipendenti;*

b) *un insieme di vettori linearmente indipendenti che non generino lo spazio.*

**Esercizio 3.4.8**

*In* $\mathrm{M}_{2,3}(\mathbb{R})$ *si considerino le seguenti matrici:*

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

*Si determinino una matrice* $C$ *tale che* $C \in \langle A, B \rangle$, $C \notin \langle A \rangle$ *e* $C \notin \langle B \rangle$ *e una matrice* $D$ *tale che* $D \notin \langle A, B \rangle$. *Si motivi la risposta.*

### Esercizio 3.4.9

*Siano dati in* $\mathbb{R}^3$ *i vettori:*

$$\mathbf{v}_1 = \begin{pmatrix}1\\0\\1\end{pmatrix} \quad \mathbf{v}_2 = \begin{pmatrix}2\\k\\1\end{pmatrix} \quad \mathbf{v}_3 = \begin{pmatrix}1\\-2\\k\end{pmatrix} \quad \mathbf{w} = \begin{pmatrix}-2\\3\\1\end{pmatrix}$$

a) *Si determinino i valori di* $k$ *per i quali i tre vettori* $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ *sono linearmente indipendenti;*

b) *si determinino i valori di* $k$ *per i quali* $\mathbf{w} \in \langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle$.

### Esercizio 3.4.10

a) *Si studino, al variare di* $k$, *le soluzioni del seguente sistema lineare:*

$$\begin{cases} x - z = 1 \\ kx - ky + 2z = 0 \\ 2x + 3ky - 11z = -1 \end{cases}$$

b) *Si stabilisca per quali valori del parametro* $k$ *il polinomio* $x^2 - 1$ *appartiene al sottospazio generato dai polinomi* $x^2 + kx + 2$, $kx - 3k^2$, $x^2 - 2x + 11$.

### Esercizio 3.4.11

*Sia* $V$ *uno spazio vettoriale e siano* $\mathbf{v}_1, \dots, \mathbf{v}_n, \mathbf{w}_1, \dots, \mathbf{w}_m \in V$. *Se* $\mathbf{v}_1, \dots, \mathbf{v}_n$ *generano* $V$, *è vero che anche* $\mathbf{v}_1, \dots, \mathbf{v}_n, \mathbf{w}_1, \dots, \mathbf{w}_m$ *generano* $V$? *Se sì, lo si dimostri, altrimenti si esibisca un controesempio.*

### Esercizio 3.4.12

*Si trovino i valori di* $h$ *e* $k$ *per i quali i vettori dell'insieme* $\{x^2 + h, kx - h, x^2 + 2kx - h\} \subseteq \mathbb{R}_2[x]$ *sono linearmente indipendenti. Se* $h = 1$ *e* $k = 2$ *tali vettori generano* $\mathbb{R}_2[x]$?

### Esercizio 3.4.13

a) *Si stabilisca per quali valori del parametro* $k$ *i vettori di* $\mathbb{R}_3[x]$: $\mathbf{v}_1 = x + 3$, $\mathbf{v}_2 = kx + 5$, $\mathbf{v}_3 = kx^2 + 5x$, $\mathbf{v}_4 = x^3$ *sono linearmente indipendenti;*

b) *scelto un valore di* $k$ *per cui* $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ *sono linearmente dipendenti, si scriva uno di essi come combinazione lineare degli altri.*

### Esercizio 3.4.14

*Si stabilisca per quali valori del parametro* $k$ *le matrici*

$$\begin{pmatrix}1 & 2\\0 & 0\end{pmatrix} \quad \begin{pmatrix}k & 0\\4 & 0\end{pmatrix} \quad \begin{pmatrix}-1 & k-2\\k & 0\end{pmatrix}$$

a) *sono linearmente indipendenti;*

b) *generano il sottospazio* $W = \left\{ \left( \begin{smallmatrix} r & s \\ t & 0 \end{smallmatrix} \right) \middle| r, s, t \in \mathbb{R} \right\}$ *di* $\mathrm{M}_2(\mathbb{R})$.

**Esercizio 3.4.15**

*Si stabilisca per quali valori del parametro $k$ si ha che:*

$$\begin{pmatrix} 3 & -2 \\ 2 & 2 \end{pmatrix} \in \left\langle \begin{pmatrix} 2 & 0 \\ 2 & 0 \end{pmatrix}, \begin{pmatrix} 1 & k \\ 0 & -k \end{pmatrix}, \begin{pmatrix} k & 6 \\ 1 & -6 \end{pmatrix} \right\rangle$$

**Esercizio 3.4.16**

*Si determini un insieme di generatori dei seguenti spazi vettoriali:*

a) $S = \{(x, y, z) \in \mathbb{R}^3 \mid x + 2y = 0\}$;

b) $T = \{hx^3 - kx^2 + 4hx + k \mid h, k \in \mathbb{R}\} \subseteq \mathbb{R}_3[x]$.

# 4 Basi e dimensione

Il concetto di base e quello di dimensione, ad esso strettamente correlato, sono centrali nella teoria degli spazi vettoriali.

Cominciamo prima con qualche esempio, principalmente, ma non solo, in $\mathbb{R}^2$ e $\mathbb{R}^3$. Grazie agli esempi svilupperemo un'intuizione geometrica che sarà preziosa e ci servirà per capire quanto accade in spazi vettoriali che non si possono visualizzare. Esporremo poi la parte teorica e dimostreremo il *Teorema del completamento*, il risultato più importante di questo capitolo, che, a partire dal concetto di base, ci permette di arrivare a quello di dimensione. Al termine ritroviamo l'algoritmo di Gauss visto nel Capitolo 1 e vediamo come possa essere utilizzato per rispondere, in modo efficace e abbastanza rapido, alle principali domande riguardanti una base o la dimensione di uno spazio vettoriale.

## 4.1 BASE: DEFINIZIONE ED ESEMPI

Come la parola stessa suggerisce, il concetto di base di uno spazio vettoriale racchiude tutte le informazioni necessarie a ricostruire lo spazio vettoriale stesso a partire da "pochi" vettori. Vediamo alcuni esempi per motivarci.

---

**Esempio 4.1.1**

Nel capitolo precedente abbiamo visto diversi esempi di insiemi di generatori dello spazio vettoriale $\mathbb{R}^2$. Ne ricordiamo un paio:

$$\mathbb{R}^2 = \langle (1,0),(0,1)\rangle = \langle (1,0),(0,1),(1,1)\rangle$$

Se a un insieme che genera $\mathbb{R}^2$ aggiungiamo un vettore, questo insieme continua a generare $\mathbb{R}^2$ per la Proposizione 3.1.8. La domanda che ci poniamo è: come possiamo trovare un insieme minimale, cioè il più piccolo possibile, di generatori per lo spazio $\mathbb{R}^2$?

Sempre la Proposizione 3.1.8 ci viene in aiuto: se eliminiamo vettori che sono combinazione lineare degli altri lo spazio vettoriale da essi generato non cambia. Nell'esempio che stiamo esaminando possiamo eliminare il vettore $(1,1)$, in quanto combinazione lineare di $(1,0)$ e $(0,1)$: $(1,1) = (1,0) + (0,1)$. Se ora proviamo però a ridurre ulteriormente l'insieme di generatori, lo spazio vettoriale da essi generato cambia. Infatti $\langle (1,0)\rangle$ è la sola retta delle ascisse, mentre $\langle (0,1)\rangle$ è la sola retta delle ordinate. Dunque, se eliminiamo uno tra i due vettori $(1,0)$ e $(0,1)$, lo spazio vettoriale generato cambia, non ci sono in altre parole "generatori superflui". Non ci può essere sfuggita l'importante differenza tra i due insiemi: $\{(1,0),(0,1)\}$ e $\{(1,0),(0,1),(1,1)\}$. Il primo insieme è costituito da vettori linearmente indipendenti, mentre i vettori del secondo insieme sono linearmente dipendenti. Quindi abbiamo visto in questo esempio che, a partire da un insieme di generatori, possiamo eliminare uno a uno i generatori che sono combinazione lineare degli altri, fino a giungere a un insieme di vettori linearmente indipendenti, in cui appunto nessun vettore è combinazione lineare degli altri (Proposizione 3.2.4). A questo punto non possiamo più eliminare vettori senza cambiare lo spazio vettoriale da essi generato.

---

La prossima proposizione formalizza le conclusioni dell'esempio precedente e ci fornisce un algoritmo per arrivare a un sistema minimale di generatori; come vedremo tale sistema sara denominato *base*.

**Proposizione 4.1.2** *Sia $V = \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle \neq \{0\}$. Allora esiste un sottoinsieme di vettori linearmente indipendenti di $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ che genera $V$.*

**Dimostrazione** – Procediamo per passi in modo algoritmico.

*Passo* 1. Si ha che $V = \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$ per ipotesi. Se $\mathbf{v}_1, \dots, \mathbf{v}_n$ sono linearmente indipendenti, allora abbiamo dimostrato l'enunciato. Altrimenti uno dei vettori, supponiamo $\mathbf{v}_n$, è combinazione lineare degli altri, per la Proposizione 3.2.4. Per la Proposizione 3.1.8 abbiamo:

$$V = \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle = \langle \mathbf{v}_1, \dots, \mathbf{v}_{n-1} \rangle$$

*Passo* 2. Nel passo 1 abbiamo eliminato dall'insieme dei generatori di $V$ il vettore $\mathbf{v}_n$, dunque $V = \langle \mathbf{v}_1, \dots, \mathbf{v}_{n-1} \rangle$. Se $\mathbf{v}_1, \dots, \mathbf{v}_{n-1}$ sono linearmente indipendenti, allora abbiamo terminato. Altrimenti torniamo al passo 1, cioè uno dei vettori, supponiamo $\mathbf{v}_{n-1}$, è combinazione lineare degli altri. Per la Proposizione 3.1.8 vista nel Capitolo 3:

$$V = \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle = \langle \mathbf{v}_1, \dots, \mathbf{v}_{n-1} \rangle = \langle \mathbf{v}_1, \dots, \mathbf{v}_{n-2} \rangle$$

È chiaro che, dopo un numero finito di passi, al massimo $n - 1$, si arriva a un insieme in cui nessun vettore è combinazione lineare degli altri, dunque, per la Proposizione 3.2.4, a un insieme di vettori linearmente indipendenti. □

Siamo pronti per la definizione di base.

**Definizione 4.1.3** Sia $V$ uno spazio vettoriale. L'insieme $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ si dice *base* se:

1. I vettori $\mathbf{v}_1, \dots, \mathbf{v}_n$ sono linearmente indipendenti;

2. I vettori $\mathbf{v}_1, \dots, \mathbf{v}_n$ generano $V$.

Diciamo inoltre che $V$ è *finitamente generato* se esiste un insieme di generatori di $V$ finito, cioè $V = \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$.

Se $V$ ammette una base $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ allora certamente è finitamente generato. Vedremo tra poco che è vero anche il contrario.

D'ora in poi, diremo che un insieme $X$ è *massimale* (*minimale*) rispetto a una certa proprietà se $X$ gode di quella proprietà, ma non appena aggiungiamo (togliamo) un elemento all'insieme $X$, allora $X$ non ne gode più.

**Teorema 4.1.4** *Siano $\mathbf{v}_1, \dots \mathbf{v}_n$ vettori in uno spazio vettoriale $V$.*

1. *$\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ è una base di $V$ se e solo se è un insieme minimale di generatori di $V$.*

2. *$\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ è una base di $V$ se e solo se è un insieme massimale di vettori linearmente indipendenti.*

**Dimostrazione** – 1. Se $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ è una base di uno spazio vettoriale $V$, allora per definizione è un insieme di generatori. Vediamo ora che è anche un insieme minimale rispetto a questa proprietà. Infatti se togliamo a $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ un qualunque vettore, allora lo spazio vettoriale generato cambia. Ciò accade in quanto altrimenti, per la Proposizione 3.1.8, un vettore tra i $\mathbf{v}_1, \dots, \mathbf{v}_n$ sarebbe combinazione lineare degli altri, mentre sappiamo che tali vettori sono linearmente indipendenti per ipotesi. Viceversa, se consideriamo un insieme minimale di generatori, allora è una base in quanto è costituito da vettori linearmente indipendenti. Infatti per la minimalità, abbiamo che, togliendo uno qualsiasi dei generatori, i rimanenti non generano più lo spazio vettoriale dato e dunque per le Proposizioni 3.1.8 e 3.2.4, ciò significa che nessuno è combinazione lineare degli altri.

2. Se $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ è una base di uno spazio vettoriale $V$, per definizione è un insieme di vettori linearmente indipendenti ed è anche massimale rispetto a questa proprietà. Infatti, poiché $\mathbf{v}_1, \dots, \mathbf{v}_n$ generano $V$, avremo che, se $\mathbf{w} \in V$, allora

$$\langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle = \langle \mathbf{v}_1, \dots, \mathbf{v}_n, \mathbf{w} \rangle = V$$

e dunque per la Proposizione 3.1.8 $\mathbf{w}$ è necessariamente una combinazione lineare di $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, e dunque i vettori $\mathbf{v}_1, \dots, \mathbf{v}_n, \mathbf{w}$ sono linearmente dipendenti per la Proposizione 3.2.4.

Viceversa, se $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ è un insieme massimale di vettori linearmente indipendenti, se aggiungiamo un qualunque altro vettore $\mathbf{v}$, ottieniamo un insieme di vettori linearmente dipendenti, cioè esistono scalari non tutti nulli $\alpha, \alpha_1, \dots, \alpha_n$ tali che

$$\alpha \mathbf{v} + \alpha_1 \mathbf{v}_1 + \dots + \alpha_n \mathbf{v}_n = \mathbf{0}$$

Notiamo che deve essere $\alpha \neq 0$, altrimenti i vettori $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ sarebbero linearmente dipendenti. Allora abbiamo che

$$\mathbf{v} = -\frac{\alpha_1}{\alpha}\mathbf{v}_1 - \dots - \frac{\alpha_n}{\alpha}\mathbf{v}_n$$

quindi $\mathbf{v} \in \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$. Poiché abbiamo scelto $\mathbf{v}$ arbitrariamente, abbiamo che $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ genera $V$. □

---

**Esempio 4.1.5**

Siano dati in $\mathbb{R}_3[x]$ i vettori:

$$x^3, x^2, 2, 5, x + 2, 3x, -7x, 2x^3$$

Vogliamo trovare una base per il sottospazio vettoriale da essi generato. La procedura che seguiamo non è quella standard, ma solo una esemplificazione della procedura descritta nella Proposizione 4.1.2. Innanzitutto vediamo subito che $5$ è combinazione lineare di $2$, in quanto è un suo multiplo: $5 = (5/2)2$. Dunque possiamo eliminare il vettore $5$ (in virtù della Proposizione 3.1.8). Allo stesso modo possiamo eliminare $-7x = (-7/3)3x$ e anche $2x^3 = 2(x^3)$. Dunque abbiamo:

$$W = \langle x^3, x^2, 2, 5, x + 2, 3x, -7x, 2x^3 \rangle = \langle x^3, x^2, 2, x + 2, 3x \rangle$$

Ora notiamo che $x + 2 = 1/3 \cdot 3x + 2$, dunque:

$$W = \langle x^3, x^2, 2, x + 2, 3x \rangle = \langle x^3, x^2, 2, 3x \rangle$$

Per verificare che questi vettori sono linearmente indipendenti e quindi formano una base per $W$ dobbiamo mostrare che l'equazione:

$$ax^3 + bx^2 + 2c + 3dx = 0$$

è soddisfatta soltanto per $a = b = c = d = 0$, ma cio è chiaro, poiché un polinomio è identicamente nullo se e solo se sono nulli tutti i suoi coefficienti. Perciò $\{x^3, x^2, 2, 3x\}$ è una base di $W$. Lasciamo dimostrare per esercizio allo studente che $W = \mathbb{R}_3[x]$. Vedremo più avanti che quest'ultima affermazione è ovvia utilizzando il concetto di dimensione.

---

Sorge spontanea la domanda: dato uno spazio vettoriale $V$, esiste sempre una base per $V$? La risposta è affermativa, anche se noi vedremo la dimostrazione soltanto per gli spazi vettoriali finitamente generati, cioè generati da un numero finito di vettori.

**Proposizione 4.1.6** *Se uno spazio vettoriale $V \neq \{0\}$ è generato da un numero finito di vettori $\mathbf{v}_1, \dots, \mathbf{v}_n$ allora esiste una base di $V$.*

**Dimostrazione** – Per ipotesi $V = \langle \mathbf{v}_1 \dots \mathbf{v}_n \rangle$. Allora per la Proposizione 4.1.2 abbiamo che esiste un sottoinsieme di $\{\mathbf{v}_1 \dots \mathbf{v}_n\}$ costituito da vettori linearmente indipendenti che generano $V$, cioè esiste una base di $V$. □

Per completezza, notiamo che l'insieme vuoto $\emptyset$ è un insieme di vettori linearmente indipendenti e che è una base dello spazio vettoriale banale $V = \{0\}$.

♦ **Osservazione 4.1.7** Non tutti gli spazi vettoriali sono finitamente generati. Un esempio di tali spazi è $\mathbb{R}[x]$, lo spazio vettoriale dei polinomi a coefficienti in $\mathbb{R}$. Infatti, se per assurdo esistesse una base di $\mathbb{R}[x]$ con un numero finito di elementi, detto $N$ il loro grado massimo, il polinomio $x^{N+1}$ non potrebbe essere espresso come combinazione lineare degli elementi della base e otterremmo una contraddizione.

Tuttavia anche per gli spazi vettoriali generati da un numero infinito di elementi esiste sempre una base: la dimostrazione è un solo un po' più difficile e non la vediamo qui. Nel caso particolare di $\mathbb{R}[x]$ una base deve necessariamente contenere un numero infinito di elementi; Per esempio lo studente è invitato a verificare che $\{1, x, x^2, x^3, \dots\}$ è una base di $\mathbb{R}[x]$.

♦ **Osservazione 4.1.8** La Proposizione 4.1.6 garantisce che ogni spazio vettoriale diverso dallo spazio nullo e generato da un numero finito di vettori abbia almeno una base. Tale base però non è unica. Per esempio è facile verificare che se $k \neq 0$ allora l'insieme

$\mathcal{B}_k = \{(k, 1), (0, 1)\}$ è una base di $\mathbb{R}^2$, quindi $\mathbb{R}^2$ ha infinite basi. Invitiamo lo studente a convincersi che se uno spazio vettoriale generico ammette una base allora ne ammette infinite.

## 4.2 IL CONCETTO DI DIMENSIONE

Uno spazio vettoriale, come già sappiamo, ammette diverse basi, tuttavia, come vedremo:

*tutte le basi hanno lo stesso numero di elementi* e tale numero si dice *dimensione* dello spazio vettoriale.

È difficile che lo studente comprenda subito l'importanza di questo numero associato univocamente a uno spazio vettoriale. Per motivarlo, basti dire che è veramente una chiave di volta per rispondere a molte domande riguardo la lineare indipendenza di certi insiemi di vettori, o per stabilire se un insieme di vettori di uno spazio vettoriale $V$ può o meno generare $V$. Per poter definire con precisione la dimensione è necessario prima il Teorema del completamento, che dimostreremo in appendice.

**Teorema del completamento 4.2.1** *Sia $\{\mathbf{v}_1, \dots \mathbf{v}_m\}$ un insieme di vettori linearmente indipendenti in uno spazio vettoriale $V$ finitamente generato. Se $\mathcal{B} = \{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ è una base di $V$ (sappiamo che ne esiste sempre almeno una), allora $m \leq n$ e possiamo sempre aggiungere a $\mathbf{v}_1, \dots, \mathbf{v}_m$ $n - m$ vettori di $\mathcal{B}$, in modo da ottenere una base di $V$.*

Vedremo al termine di questo capitolo che è possibile mettere in azione in modo molto esplicito il Teorema del completamento in $\mathbb{R}^n$ e completare a una base di $\mathbb{R}^n$ un qualunque insieme di vettori dati linearmente indipendenti.

**Proposizione 4.2.2** *Tutte le basi di uno stesso spazio vettoriale finitamente generato hanno lo stesso numero di elementi.*

**Dimostrazione** – Siano $\mathcal{B}_1 = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ e $\mathcal{B}_2 = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ due basi di $V$. Poiché i vettori di $\mathcal{B}_1$ sono linearmente indipendenti e $\mathcal{B}_2$ è una base, per il Teorema del completamento abbiamo che $n \leq m$. Scambiando i ruoli di $\mathcal{B}_1$ e $\mathcal{B}_2$ abbiamo $m \leq n$, quindi $n = m$. □

**Definizione 4.2.3** Il numero degli elementi di una base di uno spazio vettoriale si dice *dimensione* dello spazio vettoriale e si indica con $\dim(V)$. Quando tale numero è finito, cioè $V$ è generato da un numero finito di vettori, $V$ si dice di *dimensione finita*.

Vediamo ora delle basi particolarmente semplici, dette *basi canoniche*, per gli spazi vettoriali incontrati fino a qui e le dimensioni relative. Lasciamo allo studente la verifica che si tratta effettivamente di basi.

- $\mathbb{R}^n$. La base canonica è data da $\mathcal{C} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$, ove $\mathbf{e}_i$ è il vettore che ha 1 nella posizione $i$ e 0 nelle altre posizioni.

Per esempio una base di $\mathbb{R}^3$ è $\mathcal{C} = \{(1,0,0),(0,1,0),(0,0,1)\}$ e $\mathbb{R}^3$ ha dimensione 3, oppure possiamo anche dire che è uno spazio vettoriale tridimensionale.

- $\mathbb{R}_n[x]$. La base canonica è data da $\mathcal{C} = \{x^n, x^{n-1}, \dots, x, 1\}$.

- $\mathrm{M}_{m,n}(\mathbb{R})$. La base canonica è data da $\mathcal{C} = \{\mathbf{E}_{1,1}, \dots, \mathbf{E}_{m,n}\}$, ove $\mathbf{E}_{i,j}$ è la matrice che ha 1 nella posizione $(i,j)$ e 0 nelle altre posizioni. Per esempio la base canonica di $\mathrm{M}_{2,3}$ è

$$\mathcal{C} = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}$$

La conoscenza delle basi canoniche ci dice subito la dimensione degli spazi vettoriali considerati sopra:

$$\dim(\mathbb{R}^n) = n \qquad \dim(\mathbb{R}_n[x]) = n + 1 \qquad \dim(\mathrm{M}_{m,n}) = mn$$

Se si tiene bene a mente questo fatto molti esercizi diventano semplicissimi. Per esempio ci chiediamo se l'insieme

$$\{(1,2,1), \quad (1,1,5), \quad (2,3,1), \quad (0,1,0)\}$$

è una base di $\mathbb{R}^3$. La risposta è immediata: l'insieme non può essere una base, poiché sappiamo che tutte le basi in $\mathbb{R}^3$ hanno esattamente 3 elementi, mentre l'insieme dato ne ha 4.

La proposizione che enunciamo qui di seguito è particolarmente utile nelle applicazioni.

**Proposizione 4.2.4** *Sia $V$ uno spazio vettoriale di dimensione $n$ e sia $W$ un sottospazio vettoriale di $V$. Allora:*

a) $\dim(W) \leq \dim(V)$;

b) $\dim(W) = \dim(V)$ *se e solo se* $V = W$.

**Dimostrazione** – a) Ricordiamo che la dimensione di uno spazio vettoriale è il numero di elementi di una base, che è anche un insieme massimale di vettori linearmente indipendenti. Poiché $W$ è contenuto in $V$, non possiamo scegliere in $W$ più vettori linearmente indipendenti di quelli che troviamo in $V$ e dunque la dimensione di $W$ non può essere più grande di quella di $V$.

b) Poiché i vettori di una base di $W$ sono linearmente indipendenti, per il Teorema del completamento possiamo aggiungere ad essi $\dim(V) - \dim(W)$ vettori per ottenere una base di $V$. Se $\dim(W) = \dim(V)$ significa che una base di $W$ è già una base di $V$, quindi in particolare genera $V$, cioè $W = V$. □

Vediamo ora come questo teorema semplifichi notevolmente la risoluzione degli esercizi.

**Esempio 4.2.5**

Vogliamo dimostrare che $\mathcal{B} = \{(1,-1),(2,0)\}$ è una base di $\mathbb{R}^2$. In teoria dobbiamo verificare che i due vettori siano linearmente indipendenti e che generino tutto $\mathbb{R}^2$. Tuttavia mentre la lineare indipendenza è ovvia in quanto i vettori non sono uno multiplo dell'altro, per la generazione dovremmo fare dei calcoli. Ora vediamo che i calcoli sono superflui, infatti abbiamo trovato due vettori linearmente indipendenti in $\mathbb{R}^2$, pertanto il sottospazio che generano ha dimensione due. Dunque per il teorema precedente deve essere uguale a $\mathbb{R}^2$.

Analogamente, nell'Esempio 4.1.5 abbiamo dimostrato che i vettori $x^3, x^2, 2, 3x$ sono linearmente indipendenti. Allora dato che sono 4, possiamo subito concludere *senza fare calcoli* che sono una base di $\mathbb{R}_3[x]$, perciò $\langle x^3, x^2, 2, 3x\rangle = \mathbb{R}_3[x]$.

In realtà è vera una proprietà molto più forte. In generale per un insieme di vettori il fatto di essere linearmente indipendenti o di generare uno spazio vettoriale sono due proprietà che non hanno nulla a che vedere l'una con l'altra, però, se siamo in uno spazio vettoriale di dimensione $n$ e consideriamo un insieme con esattamente $n$ vettori, allora le due proprietà si equivalgono.

**Proposizione 4.2.6** *Sia $V$ uno spazio vettoriale di dimensione $n$ e sia $\{\mathbf{v}_1,\dots,\mathbf{v}_n\}$ un insieme di $n$ vettori di $V$. Le seguenti affermazioni sono equivalenti:*

a) $\{\mathbf{v}_1,\dots,\mathbf{v}_n\}$ *è una base di $V$.*

b) $\mathbf{v}_1,\dots,\mathbf{v}_n$ *sono linearmente indipendenti.*

c) $\mathbf{v}_1,\dots,\mathbf{v}_n$ *generano $V$.*

**Dimostrazione** – a) $\Longrightarrow$ b) per definizione di base.

Vediamo b) $\Longrightarrow$ c). Sia $W = \langle \mathbf{v}_1,\dots,\mathbf{v}_n\rangle$. Si ha che $\{\mathbf{v}_1,\dots,\mathbf{v}_n\}$ è una base di $W$, quindi $W$ ha dimensione $n$, allora per la Proposizione 4.2.4 b) si ha che $W = V$, dunque i vettori dati generano $V$.

Vediamo ora c) $\Longrightarrow$ a). Per la Proposizione 4.1.2 esiste un sottoinsieme di $\{\mathbf{v}_1,\dots,\mathbf{v}_n\}$ che è una base di $V$, ma, poiché $V$ ha dimensione $n$, tale sottoinsieme deve contenere esattamente $n$ elementi, quindi è proprio $\{\mathbf{v}_1,\dots,\mathbf{v}_n\}$. □

Vogliamo ora renderci conto del fatto che una base è un modo "efficiente" di rappresentare i vettori di uno spazio vettoriale. Vediamo un esempio.

**Esempio 4.2.7**

In $\mathbb{R}^2$ sappiamo che tutti i vettori sono combinazioni lineari dei due vettori della base canonica $\mathbf{e}_1 = (1,0)$, $\mathbf{e}_2 = (0,1)$. Possiamo inoltre verificare che ogni vettore in $\mathbb{R}^2$ non solo è combinazione lineare di $\mathbf{e}_1$ e $\mathbf{e}_2$ ma lo è in modo *unico*. Infatti, se prendiamo il vettore $(2,3)$, possiamo scrivere $(2,3) = 2(1,0) + 3(0,1)$ e i numeri 2 e 3 sono gli unici scalari che ci danno $(2,3)$ come combinazione lineare di $(1,0)$, $(0,1)$. Diverso è il discorso se invece prendiamo i tre vettori $(1,0)$, $(0,1)$ e $(1,1)$. Infatti abbiamo già capito che il vettore $(1,1)$ è in qualche modo "superfluo", cioè sappiamo che: $\langle (1,0),(0,1),(1,1)\rangle = \langle (1,0),(0,1)\rangle$

e questo perché $(1,1)$ è combinazione lineare di $(1,0)$ e $(0,1)$. Ciò si riflette nel fatto che un vettore in $\mathbb{R}^2$ *non è più combinazione lineare in modo unico* di questi tre vettori. Infatti

$$(2,3) = (1,0) + 2(0,1) + (1,1) = 0(1,0) + (0,1) + 2(1,1)$$

Il concetto di unicità di espressione visto nell'esempio precedente è il contenuto della seguente proposizione.

**Teorema 4.2.8** *Sia $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base ordinata per lo spazio vettoriale $V$ (cioè abbiamo fissato un ordine nell'insieme dei vettori numerandoli) e sia $\mathbf{v} \in V$. Allora esiste ed è unica la $n$-upla di scalari $(\alpha_1, \ldots, \alpha_n)$ tale che*

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n$$

**Dimostrazione** – Essendo $\mathcal{B}$ un sistema di generatori, ogni $\mathbf{v} \in V$ si scrive come combinazione lineare degli elementi di $\mathcal{B}$, cioè esistono scalari $\alpha_1, \ldots, \alpha_n$ tali che:

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n$$

Dimostriamo l'unicità degli $\alpha_i$. Supponiamo che sia anche

$$\mathbf{v} = \beta_1\mathbf{v}_1 + \cdots + \beta_n\mathbf{v}_n$$

Sottraendo membro a membro queste due equazioni si ottiene

$$\mathbf{0} = (\alpha_1 - \beta_1)\mathbf{v}_1 + \cdots + (\alpha_n - \beta_n)\mathbf{v}_n$$

da cui $\alpha_1 - \beta_1 = \ldots = \alpha_n - \beta_n = 0$ in virtù della lineare indipendenza dei vettori di $\mathcal{B}$, dunque $\alpha_1 = \beta_1, \ldots, \alpha_n = \beta_n$. □

**Definizione 4.2.9** Gli scalari $(\alpha_1, \ldots, \alpha_n)$ definiti nel Teorema 4.2.8 si dicono le *componenti* di $\mathbf{v} \in V$ nella base $\mathcal{B}$ o anche le *coordinate* di $\mathbf{v}$ rispetto alla base $\mathcal{B}$ e verranno indicate con la notazione $(\mathbf{v})_{\mathcal{B}} = (\alpha_1, \ldots, \alpha_n)$.

**Esempio 4.2.10**

A titolo esemplificativo, dimostriamo per esercizio che $\mathcal{B} = \{(1,-1),(2,0)\}$ è una base di $\mathbb{R}^2$ e determiniamo le coordinate di $\mathbf{v} = (-3,1)$ rispetto a questa base.

Chiaramente i vettori di $\mathcal{B}$ sono linearmente indipendenti in quanto $(2,0)$ non è multiplo di $(1,-1)$. A questo punto, poiché $\mathbb{R}^2$ ha dimensione 2, sappiamo già che $\mathcal{B}$ è una base. Però è istruttivo dimostrare direttamente che è un insieme di generatori. Sia $(x,y) \in \mathbb{R}^2$ e cerchiamo $a, b \in \mathbb{R}$ tali che

$$(x,y) = a(1,-1) + b(2,0) = (a+2b, -a)$$

dovrà dunque essere $a = -y$ e, pertanto $b = \frac{x+y}{2}$ e ciò è sempre possibile, dunque $\mathcal{B}$ genera $\mathbb{R}^2$. In particolare $(-3,1) = (-1)(1,-1) + (-1)(2,0)$, quindi le coordinate di $\mathbf{v}$ rispetto alla base $\mathcal{B}$ sono $(\mathbf{v})_{\mathcal{B}} = (-1,-1)$.

## 4.3 L'ALGORITMO DI GAUSS COME METODO PRATICO PER LA RISOLUZIONE DI PROBLEMI DI ALGEBRA LINEARE

Abbiamo già visto che, per determinare se alcuni vettori di uno spazio vettoriale $V$ sono linearmente indipendenti o se generano $V$, ci si riconduce a risolvere un sistema lineare. Dunque è chiaro che l'algoritmo di Gauss è di grande aiuto in problemi legati alla lineare dipendenza o indipendenza e al concetto di generazione o, più in generale, al concetto di base. Quello che vogliamo vedere ora è come sia possibile usare *direttamente* l'algoritmo di Gauss, senza dover impostare un sistema lineare.

Osserviamo che, se abbiamo una matrice $A \in \mathrm{M}_{m,n}(\mathbb{R})$, possiamo considerare le sue righe come tanti vettori di $\mathbb{R}^n$, tali vettori si dicono i *vettori riga* di $A$. Per esempio, se $A = \left(\begin{smallmatrix} 0 & 1 & -3 \\ 2 & -1 & 1 \end{smallmatrix}\right)$, i suoi vettori riga sono $R_1 = (0, 1, -3)$ e $R_2 = (2, -1, 1)$.

**Proposizione 4.3.1** *Data una matrice $A \in \mathrm{M}_{m,n}(\mathbb{R})$, le operazioni elementari di riga non cambiano il sottospazio di $\mathbb{R}^n$ generato dai vettori riga di $A$.*

**Dimostrazione** – Ricordiamo che le operazioni elementari di riga (Definizione 1.4.2) sono:

(a) scambio di due righe;

(b) moltiplicazione di una riga per un numero reale diverso da 0;

(c) sostituzione della riga $i$-esima con la somma della riga $i$-esima e della $j$-esima moltiplicata per un numero reale $\alpha$ qualsiasi.

È immediato verificare che l'enunciato è vero per le operazioni di tipo (a) e (b). Per le operazioni di tipo (c), è sufficiente mostrare che se $R_i$ e $R_j$ sono due vettori riga di $A$ e $\alpha \in \mathbb{R}$, si ha che $\langle R_i, R_j + \alpha R_i \rangle = \langle R_i, R_j \rangle$. Si ha ovviamente che $R_i, R_j + \alpha R_i \in \langle R_i, R_j \rangle$, quindi $\langle R_i, R_j \rangle$ è un sottospazio vettoriale di $\mathbb{R}^n$ contenente $R_i, R_j + \alpha R_i$. Allora per la Proposizione 3.1.5 si ha che $\langle R_i, R_j + \alpha R_i \rangle \subseteq \langle R_i, R_j \rangle$.

L'inclusione $\langle R_i, R_j \rangle \subseteq \langle R_i, R_j + \alpha R_i \rangle$ si dimostra in modo analogo tenendo conto del fatto che $R_i = (R_i + \alpha R_j) - \alpha R_j$, quindi $R_i \in \langle R_i, R_j + \alpha R_i \rangle$. □

♦ **Osservazione 4.3.2** Le operazioni elementari di riga non cambiano il sottospazio di $\mathbb{R}^n$ generato dai vettori riga di $A$, ma cambiano il sottospazio di $\mathbb{R}^m$ generato dai vettori colonna di $A$. Si invita il lettore a verificare questo fatto con un esempio per convincersene.

**Proposizione 4.3.3** *Se una matrice $A$ è a scala per riga, i suoi vettori riga non nulli sono linearmente indipendenti.*

**Dimostrazione** – Siano $R_1, \dots, R_k$ le righe non nulle di $A$ e siano $a_{1j_1}, \dots, a_{kj_k}$ i rispettivi pivot. Sia ora $\lambda_1 R_1 + \cdots + \lambda_k R_k = \mathbf{0}$, vogliamo dimostrare che $\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$. Nel vettore $\lambda_1 R_1 + \cdots + \lambda_k R_k$ l'elemento di posto $j_1$ è $\lambda_1 a_{1j_1}$, l'elemento di posto $j_2$ è $\lambda_1 a_{1j_2} + \lambda_2 a_{2j_2}$, e così via, sino all'elemento di posto $j_k$, che è $\lambda_1 a_{1j_k} + \lambda_2 a_{2j_k} + \cdots + \lambda_k a_{kj_k}$. Quindi dal fatto che $\lambda_1 R_1 + \cdots + \lambda_k R_k = \mathbf{0}$ segue che:

$$\begin{cases} \lambda_1 a_{1j_1} = 0 \\ \lambda_1 a_{1j_2} + \lambda_2 a_{2j_2} = 0 \\ \vdots \\ \lambda_1 a_{1j_k} + \lambda_2 a_{2j_k} + \cdots + \lambda_k a_{kj_k} = 0 \end{cases}$$

Poiché $a_{1j_1} \neq 0$, dalla prima equazione otteniamo $\lambda_1 = 0$. Sostituendo $\lambda_1 = 0$ nella seconda equazione e, poiché $a_{2j_2} \neq 0$, si ottiene che $\lambda_2 = 0$, e così via. Dopo $k$ passi abbiamo che anche $\lambda_k = 0$, dunque $\lambda_1 = \cdots = \lambda_k = 0$. Ciò dimostra che le righe $R_1, \ldots, R_k$ sono vettori di $\mathbb{R}^n$ linearmente indipendenti. □

Vediamo ora un esempio di come queste proposizioni si possano applicare agli esercizi.

---

**Esempio 4.3.4**

Dati i seguenti vettori di $\mathbb{R}^3$

$$\mathbf{v}_1 = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \qquad \mathbf{v}_2 = \begin{pmatrix} 0 \\ -1 \\ 3 \end{pmatrix} \qquad \mathbf{v}_3 = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$$

ci chiediamo se sono linearmente indipendenti e se sono una base di $\mathbb{R}^3$. Inoltre, vogliamo trovare una base del sottospazio $W$ da essi generato e calcolare la dimensione di $W$.

La matrice:

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

ha forma ridotta:

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

Per la Proposizione 4.3.1 abbiamo che

$$\begin{aligned} W &= \left\langle \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} \right\rangle \\ &= \left\langle \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 3 \end{pmatrix} \right\rangle \end{aligned}$$

Si ha quindi che il sottospazio $W$ è generato da due vettori $\mathbf{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix}$ e $\mathbf{u}_2 = \begin{pmatrix} 0 \\ -1 \\ 3 \end{pmatrix}$ linearmente indipendenti (per la Proposizione 4.3.3). Dunque $\{\mathbf{u}_1, \mathbf{u}_2\}$ è una base per $W$, che di conseguenza ha dimensione 2. Dato che, per il Teorema 4.1.4, il numero di vettori in una base è il massimo numero di vettori linearmente indipendenti nello spazio vettoriale, abbiamo che $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ sono linearmente dipendenti, quindi non possono essere una base di $\mathbb{R}^3$.

---

Vediamo in generale come si deve procedere per trovare una base del sottospazio $W$ generato dai vettori $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ e per decidere se $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$

sono linearmente indipendenti, cioè formalizziamo schematicamente quello che abbiamo imparato dall'esempio precedente.

- Scriviamo la matrice $A$ che ha come righe i vettori $\mathbf{v}_1, \ldots, \mathbf{v}_k$ (sarà una matrice $k \times n$).
- Utilizzando l'algoritmo di Gauss, otteniamo una matrice $A'$ in forma ridotta a scala per righe.
- I vettori che formano le righe non nulle di $A'$ generano $W$ e sono linearmente indipendenti, quindi sono una base di $W$.
- Sia $r$ il numero di righe non nulle di $A'$; si ha quindi che $\dim W = r$. Se $k = r$ i vettori $\mathbf{v}_1, \ldots, \mathbf{v}_k$ generano uno spazio vettoriale di dimensione $k$ e quindi sono linearmente indipendenti, se $k > r$ i vettori $\mathbf{v}_1, \ldots, \mathbf{v}_k$ sono linearmente dipendenti, perché un insieme massimale di vettori linearmente indipendenti di $W$ ha $r < k$ elementi.

Adesso invece vediamo un metodo per ottenere una base di un sottospazio $W \subset \mathbb{R}^n$ e completarla a una base di $\mathbb{R}^n$.

Sia $W \subset \mathbb{R}^n$ il sottospazio vettoriale dato generato da $\mathbf{w}_1, \ldots, \mathbf{w}_k$.

- Scriviamo la matrice che ha come righe i vettori $\mathbf{w}_1, \ldots, \mathbf{w}_k$ (sarà una matrice $k \times n$).
- Usando l'algoritmo di Gauss otteniamo una matrice $A'$ in forma ridotta a scala per righe.
- Gli $r$ vettori che formano le righe non nulle di $A'$ generano $W$ e sono linearmente indipendenti, quindi sono una base di $W$.
- Per completare tale base a una base di $\mathbb{R}^n$, basta aggiungere $n - r$ vettori riga nelle posizioni corrispondenti ai gradini mancanti nella forma a scala per righe di $A'$, in modo da ottenere una matrice $A''$ a scala per righe con $n$ pivot.
  Infatti, per la Proposizione 4.3.3, le righe di $A''$ sono linearmente indipendenti, e, poiché $\mathbb{R}^n$ ha dimensione $n$, per la Proposizione 4.2.6 gli $n$ vettori riga di $A''$ sono una base di $\mathbb{R}^n$.

**Esempio 4.3.5**

Nell'Esempio 4.3.4 per completare la base $\{\mathbf{u}_1, \mathbf{u}_2\}$ a una base di $\mathbb{R}^3$ basta aggiungere il vettore $(0, 0, 1)$, oppure qualunque vettore del tipo $(0, 0, h)$, con $h \neq 0$.

♦ **Osservazione 4.3.6** Osserviamo a questo punto che, se fissiamo una base ordinata $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ dello spazio vettoriale $V$, possiamo considerare la funzione $c : V \to \mathbb{R}^n$ che a ogni vettore associa le sue coordinate rispetto alla base $\mathcal{B}$. Se scriviamo $\mathbf{v} \in V$ come combinazione lineare degli elementi di $\mathcal{B}$, $\mathbf{v} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n$, abbiamo $c(v) = (\mathbf{v})_{\mathcal{B}} = (\alpha_1, \ldots, \alpha_n)$. Per il Teorema 4.2.8, $c$ è un'applicazione biunivoca. Se

$\mathbf{v}$ ha coordinate $(\alpha_1, \dots, \alpha_n)$ e $\lambda \in \mathbb{R}$ è uno scalare, allora le coordinate di $\lambda\mathbf{v}$ sono $(\lambda\alpha_1, \dots, \lambda\alpha_n)$, mentre, se $\mathbf{w}$ è un altro vettore, le cui coordinate sono $(\beta_1, \dots, \beta_n)$, allora le coordinate di $\mathbf{v} + \mathbf{w}$ sono $c(\mathbf{v} + \mathbf{w}) = (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n)$. Lo studente è invitato a verificare le due asserzioni fatte, che comunque riprenderemo anche in seguito, quando parleremo di isomorfismi di spazi vettoriali. Per ora ci limiteremo a osservare che grazie a queste proprietà, si ha che $c(\lambda_1\mathbf{v}_1 + \dots + \lambda_n\mathbf{v}_n) = \lambda_1 c(\mathbf{v}_1) + \dots + \lambda_n c(\mathbf{v}_n)$ per ogni $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ e per ogni $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Da questo segue che i vettori $\mathbf{v}_1, \dots, \mathbf{v}_n$ sono linearmente indipendenti se e solo se lo sono le loro coordinate $c(\mathbf{v}_1), \dots, c(\mathbf{v}_n)$, viste come vettori in $\mathbb{R}^n$. Analogamente, $\mathbf{w}$ è combinazione lineare di $\mathbf{v}_1, \dots, \mathbf{v}_n$ se e solo se $c(\mathbf{w})$ è combinazione lineare di $c(\mathbf{v}_1), \dots, c(\mathbf{v}_n)$, e $\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ è una base di $\langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$ se e solo se $\{c(\mathbf{w}_1), \dots, c(\mathbf{w}_k)\}$ è una base di $\langle c(\mathbf{v}_1), \dots, c(\mathbf{v}_n) \rangle$. Quindi, anziché operare sui vettori, possiamo operare sulle loro coordinate, e poi da esse risalire nuovamente ai vettori. Ciò ci dà modo di utilizzare tutte le tecniche che abbiamo visto per $\mathbb{R}^n$ anche per uno spazio vettoriale qualsiasi, purché ovviamente finitamente generato. Naturalmente, per poter fare questo, è sempre necessario *fissare una base*, altrimenti non è possibile parlare di coordinate come $n$-uple univocamente associate a ogni vettore. In altre parole uno stesso vettore può avere coordinate diverse ad esso associate rispetto a basi diverse. Approfondiremo questo discorso in un capitolo successivo, per il momento ci limitiamo a vedere un esempio di come possiamo fare calcoli in spazi vettoriali qualsiasi, ma di dimensione finita, utilizzando tecniche studiate per $\mathbb{R}^n$.

**Esempio 4.3.7**

Dati i polinomi $2x^2 + 3x - 1$, $-x + 3$, $2x^2 + 2x + 2 \in \mathbb{R}_2[x]$, vogliamo stabilire se sono linearmente indipendenti, e determinare una base del sottospazio $W$ da essi generato. Consideriamo la base canonica $C = \{x^2, x, 1\}$ di $\mathbb{R}_2[x]$. Rispetto ad essa, le coordinate dei polinomi sono

$$\mathbf{v}_1 = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ -1 \\ 3 \end{pmatrix} \quad \mathbf{v}_3 = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$$

Possiamo quindi procedere come nell'Esempio 4.3.4. Facendo esattamente gli stessi calcoli ottiamo una base di $W$ data dai polinomi le cui coordinate sono $\mathbf{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix}$, $\mathbf{u}_2 = \begin{pmatrix} 0 \\ -1 \\ 3 \end{pmatrix}$. Tornando ai polinomi, si ha che una base di $W$ è $\{x^2 + 4, -x + 3\}$.

## 4.4 ESERCIZI SVOLTI

### Esercizio 4.4.1

*Sia* $W = \langle (1-k, 1, 1, -k), (2, 2-k, 2, 0), (1, 1, 1-k, k) \rangle \subseteq \mathbb{R}^4$. *Si determini, al variare di* $k$, *la dimensione di* $W$ *e, scelto un valore di* $k$ *a piacere, si completi una base di* $W$ *a una base di* $\mathbb{R}^4$.

**Svolgimento**

Scriviamo la matrice $A$ che ha per righe i vettori dati e applichiamo l'algoritmo di Gauss per ridurla a scala. Si ha che:

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

e riducendo a scala si ottiene:

$$A' = \begin{pmatrix} 1 & 1 & 1-k & k \\ 0 & -k & 2k & -2k \\ 0 & 0 & 4k-k^2 & -4k+k^2 \end{pmatrix}$$

Se $k \neq 0$ e $4k - k^2 \neq 0$, cioè se $k \neq 0$ e $k \neq 4$, la matrice $A'$ ha 3 righe non nulle, tra loro linearmente indipendenti, quindi, poiché

$$\begin{aligned} W &= \langle (1-k, 1, 1, -k), (2, 2-k, 2, 0), (1, 1, 1-k, k) \rangle \\ &= \langle (1, 1, 1-k, k), (0, -k, 2k, -2k), (0, 0, 4k-k^2, -4k+k^2) \rangle \end{aligned}$$

si ha che $W$ ha dimensione 3.

Se $k = 0$ la matrice $A'$ ha una sola riga non nulla, quindi $W = \langle (1, 1, 1, 0) \rangle$ ha dimensione 1.

Se $k = 4$, si ottiene:

$$A' = \begin{pmatrix} 1 & 1 & -3 & 4 \\ 0 & -4 & 8 & -8 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

quindi $A'$ ha 2 righe non nulle e $W = \langle (1, 1, -3, 4), (0, -4, 8, -8) \rangle$ ha dimensione 2.

Scegliamo ora $k = 4$. Per completare $\{(1, 1, -3, 4), (0, -4, 8, -8)\}$ a una base di $\mathbb{R}^4$, bisogna aggiungere 2 vettori riga aventi i pivot nei "gradini mancanti", cioè al terzo e quarto posto. Per esempio possiamo aggiungere $(0, 0, -1, 2)$, $(0, 0, 0, 5)$.

Quindi $\{(1, 1, -3, 4), (0, -4, 8, -8), (0, 0, -1, 2), (0, 0, 0, 5)\}$ è una base di $\mathbb{R}^4$ ottenuta completando una base di $W$. ∎

### Esercizio 4.4.2

*Siano* $\mathbf{v}_1 = (1, 2, 1, 0), \mathbf{v}_2 = (4, 8, k, 5), \mathbf{v}_3 = (-1, -2, 3-k, -k)$. *Si determini per quali valori di* $k$ *i vettori* $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ *sono linearmente indipendenti. Posto* $k = 1$, *si determini, se possibile, un vettore* $\mathbf{w} \in \mathbb{R}^4$ *tale che* $\mathbf{w} \notin \langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle$.

**Svolgimento**

Scriviamo la matrice $A$ che ha per righe i vettori dati e applichiamo l'algoritmo di Gauss per ridurla a scala. Si ha che:

$$A = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 4 & 8 & k & 5 \\ -1 & -2 & 3-k & -k \end{pmatrix}$$

e, riducendo a scala, si ottiene:

$$A' = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & k-4 & 5 \\ 0 & 0 & 0 & 5-k \end{pmatrix}$$

Sia $W = \langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle$. Se $k \neq 4$ e $k \neq 5$, la matrice $A'$ ha tre righe non nulle, tra loro linearmente indipendenti, quindi $W = \langle (1, 2, 1, 0), (0, 0, k-4, 5), (0, 0, 0, k-5) \rangle$ ha dimensione 3. Poiché $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ generano $W$, per la Proposizione 4.2.6 sono linearmente indipendenti.

Se $k = 4$ otteniamo:

$$A' = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Ora $W = \langle(1,2,1,0),(0,0,0,5),(0,0,0,1)\rangle = \langle(1,2,1,0),(0,0,0,1)\rangle$ ha dimensione 2, quindi $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ sono linearmente dipendenti.

Se $k = 5$ otteniamo:

$$A' = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

che è una matrice a scala con due righe non nulle, quindi $W$ ha dimensione 2 e $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ sono linearmente dipendenti.

Sia ora $k = 1$. Sostituiamo tale valore in $A'$ per avere una base di $W$:

$$A' = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & -3 & 5 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

Si ha che $W$ ha dimensione 3 e, se scegliamo un vettore riga che abbia il secondo pivot non nullo, Per esempio $\mathbf{w} = (0,-2,3,-1)$, si ha che per la Proposizione 4.3.3 i vettori $(1,2,1,0),(0,-2,3,-1),(0,0,-3,5),(0,0,0,4)$ sono linearmente indipendenti, quindi per la Proposizione 3.2.4 si ha che $(0,-2,3,-1) \notin \langle(1,2,1,0),(0,0,-3,5),(0,0,0,4)\rangle$, cioé $\mathbf{w} \notin W$. ∎

### Esercizio 4.4.3

*Sia $W$ il sottospazio di $\mathbb{R}^5$ generato dall'insieme*

$$\mathcal{B} = \{(1,3,-1,1,2),(2,6,-2,4,4)\}$$

*Si completi la base $\mathcal{B}$ di $W$ a una base di $\mathbb{R}^5$.*

**Svolgimento**

Osserviamo che $\mathcal{B}$ è effettivamente una base di $W$, perché i 2 vettori di $\mathcal{B}$ sono linearmente indipendenti, non essendo multipli l'uno dell'altro.

Un modo per completare $\mathcal{B}$ a una base di $\mathbb{R}^5$ è procedere come indicato nella dimostrazione del Teorema del completamento. Un altro modo è il seguente.

Utilizzando l'algoritmo di Gauss, determiniamo una base $\mathcal{B}'$ di $W$, in modo che la matrice $A'$ avente come righe i vettori di $\mathcal{B}'$ sia a scala. Cioè consideriamo

$$A = \begin{pmatrix} 1 & 3 & -1 & 1 & 2 \\ 2 & 6 & -2 & 4 & 4 \end{pmatrix}$$

e, applicando l'algoritmo di Gauss, otteniamo:

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

A questo punto, per completare $\{(1,3,-1,1,2),(0,0,0,2,0)\}$ a una base di $\mathbb{R}^5$, bisogna aggiungere 3 vettori riga aventi i pivot nei "gradini mancanti", cioè al secondo, terzo e quinto posto. Per esempio possiamo aggiungere $(0,1,-1,0,1)$, $(0,0,2,1,-3)$, $(0,0,0,0,1)$. Poiché

$$W = \langle(1,3,-1,1,2),(2,6,-2,4,4)\rangle = \langle(1,3,-1,1,2),(0,0,0,2,0)\rangle$$

è facile vedere che

$$\begin{aligned}&\langle(1,3,-1,1,2),(2,6,-2,4,4),(0,1,-1,0,1),(0,0,2,1,-3),(0,0,0,0,1)\rangle\\ =&\langle(1,3,-1,1,2),(0,0,0,2,0),(0,1,-1,0,1),(0,0,2,1,-3),(0,0,0,0,1)\rangle\\ =&\mathbb{R}^5\end{aligned}$$

Quindi se $\tilde{\mathcal{B}}$ è l'insieme

$$\{(1,3,-1,1,2),(2,6,-2,4,4),(0,1,-1,0,1),(0,0,2,1,-3),(0,0,0,0,1)\}$$

si ha che i vettori di $\tilde{\mathcal{B}}$ generano $\mathbb{R}^5$, quindi, per la Proposizione 4.2.6, $\tilde{\mathcal{B}}$ è una base di $\mathbb{R}^5$ ed è stata ottenuta completando la base $\mathcal{B}$ di $W$. ■

**Esercizio 4.4.4**

*Si considerino i seguenti sottospazi vettoriali di* $\mathrm{M}_{2,3}(\mathbb{R})$*:*

$$U=\left\{\begin{pmatrix}a&0&b\\c&a&d\end{pmatrix}\middle| a,b,c,d\in\mathbb{R}\right\}$$

$$W=\left\{\begin{pmatrix}r&s&t\\u&x&y\end{pmatrix}\middle| r+s+t+u+x+y=0\right\}$$

*e si determini una base di* $U\cap W$*.*

**Svolgimento**

Si ha che:

$$U\cap W=\left\{\begin{pmatrix}a&0&b\\c&a&d\end{pmatrix}\middle| a,b,c,d\in\mathbb{R},\ \ a+b+c+a+d=0\right\}$$

cioè

$$\begin{aligned}U\cap W&=\left\{\begin{pmatrix}a&0&b\\c&a&d\end{pmatrix}\middle| a,b,c,d\in\mathbb{R},\ \ d=-2a-b-c\right\}\\&=\left\{\begin{pmatrix}a&0&b\\c&a&-2a-b-c\end{pmatrix}\middle| a,b,c\in\mathbb{R}\right\}\\&=\left\{a\begin{pmatrix}1&0&0\\0&1&-2\end{pmatrix}+b\begin{pmatrix}0&0&1\\0&0&-1\end{pmatrix}+c\begin{pmatrix}0&0&0\\1&0&-1\end{pmatrix}\middle| a,b,c\in\mathbb{R}\right\}\\&=\left\langle\begin{pmatrix}1&0&0\\0&1&-2\end{pmatrix},\begin{pmatrix}0&0&1\\0&0&-1\end{pmatrix},\begin{pmatrix}0&0&0\\1&0&-1\end{pmatrix}\right\rangle\end{aligned}$$

Quindi i vettori $\mathbf{v}_1=\left(\begin{smallmatrix}1&0&0\\0&1&-2\end{smallmatrix}\right)$, $\mathbf{v}_2=\left(\begin{smallmatrix}0&0&1\\0&0&-1\end{smallmatrix}\right)$, $\mathbf{v}_3=\left(\begin{smallmatrix}0&0&0\\1&0&-1\end{smallmatrix}\right)$ generano $U\cap V$. Per mostrare che sono linearmente indipendenti, consideriamo le loro coordinate rispetto alla base canonica, cioè: $(\mathbf{v}_1)_{\mathcal{C}}=(1,0,0,0,1,-2)$, $(\mathbf{v}_2)_{\mathcal{C}}=(0,0,1,0,0,-1)$, $(\mathbf{v}_3)_{\mathcal{C}}=(0,0,0,1,0,-1)$. Osserviamo che la matrice $A$ che ha per righe $(\mathbf{v}_1)_{\mathcal{C}},(\mathbf{v}_2)_{\mathcal{C}},(\mathbf{v}_3)_{\mathcal{C}}$ è a scala, quindi per la Proposizione 4.3.3 si ha che $(\mathbf{v}_1)_{\mathcal{C}},(\mathbf{v}_2)_{\mathcal{C}},(\mathbf{v}_3)_{\mathcal{C}}$ sono linearmente indipendenti. Allora anche $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ sono linearmente indipendenti, quindi sono una base di $U\cap W$. ■

## 4.5 ESERCIZI PROPOSTI

### Esercizio 4.5.1

*Si determini una base di $\mathbb{R}_4[x]$ diversa dalla base canonica, motivando la risposta.*

### Esercizio 4.5.2

*Si stabilisca per quali valori di $k$ i seguenti polinomi di $\mathbb{R}_2[x]$ generano $\mathbb{R}_2[x]$:* $x^2 + kx + 1, 3x^2 + 4x + 1, 5x^2 + kx - 3$.

### Esercizio 4.5.3

*Si trovi una base di $\langle(1,0,3),(2,3,0),(1,1,1)\rangle$ e la si completi a una base di $\mathbb{R}^3$.*

### Esercizio 4.5.4

*Si stabilisca quali dei seguenti insiemi di vettori generano $\mathbb{R}^3$ e quali sono una base di $\mathbb{R}^3$.*

a) $\mathcal{S}_1 = \left\{ \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -3 \\ 0 \end{pmatrix} \right\}$

b) $\mathcal{S}_2 = \left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 6 \\ -3 \end{pmatrix} \right\}$

c) $\mathcal{S}_3 = \left\{ \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ -9 \end{pmatrix}, \begin{pmatrix} 0 \\ 5 \\ -5 \end{pmatrix} \right\}$

### Esercizio 4.5.5

*Si stabilisca se i vettori $\mathbf{v}_1 = (1,1), \mathbf{v}_2 = (-1,1), \mathbf{v}_3 = (2,1)$ sono linearmente indipendenti. Essi generano $\mathbb{R}^2$?*

### Esercizio 4.5.6

*Si stabilisca per quali valori di $k$ il vettore $\mathbf{w} = (-1,k,1)$ appartiene a $\langle(1,-1,0), (k,-k,1),(-1,k^2,1)\rangle$.*

### Esercizio 4.5.7

*Si stabilisca per quali valori di $k$ i vettori $\mathbf{v}_1 = (1,2k,0), \mathbf{v}_2 = (1,0,-3), \mathbf{v}_3 = (-1,0,k+2)$ sono linearmente indipendenti.*

*Posto $k = 1$, si stabilisca se $(1,-2,-6) \in \langle \mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\rangle$.*

### Esercizio 4.5.8

*Siano $\mathbf{v}_1 = (1,1,1,0), \mathbf{v}_2 = (2,0,2,-1), \mathbf{v}_3 = (-1,1,-1,1)$. Si determini una base di $\langle \mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\rangle$ e la si completi a una base di $\mathbb{R}^4$. Si determini inoltre per quali valori di $k$ il vettore $\mathbf{w} = (k,-3k,k,-2k)$ appartiene a $\langle \mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\rangle$.*

### Esercizio 4.5.9

*Sia* $X = \{(x,y,z) \in \mathbb{R}^3 | x - y + z = 0 \text{ e } x - 2y = 0\}$. *Si dimostri che* $X$ *è un sottospazio di* $\mathbb{R}^3$ *e se ne determini una base.*

### Esercizio 4.5.10

*Si stabilisca per quali valori di* $k$ *i vettori* $\mathbf{v}_1 = (0,1,0)$, $\mathbf{v}_2 = (1,k,4), \mathbf{v}_3 = (k,2,3)$ *generano* $\mathbb{R}^3$. *Posto* $k = 0$, *si stabilisca se il vettore* $(4,-1,6)$ *appartiene a* $\langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle$.

### Esercizio 4.5.11

a) *Si stabilisca per quali valori di* $k$ *i vettori* $\mathbf{v}_1 = x^2 + 2x - 1, \mathbf{v}_2 = x^2 + kx + 1 - k, \mathbf{v}_3 = 5x + k$ *sono linearmente dipendenti.*
b) *Scelto a piacere un valore di* $k$ *trovato al punto a), si scriva uno dei 3 vettori come combinazione lineare degli altri e si trovi una base di* $\langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle$.

### Esercizio 4.5.12

a) *Si stabilisca per quali valori di* $k$ *i vettori* $\mathbf{v}_1 = 6x^2 - 6x - k, \mathbf{v}_2 = -kx^2 + kx + 6$ *sono linearmente indipendenti.*
b) *Posto* $k = 0$ *si determini la dimensione di* $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle$ *e si trovi, se possibile, un vettore* $\mathbf{w}$ *in modo che* $\mathbf{w}$ *non appartenga a* $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle$ *e* $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{w}\}$ *non generino* $\mathbb{R}_2[x]$.

### Esercizio 4.5.13

*Si stabilisca per quali valori di* $k$ *i vettori* $\mathbf{v}_1 = (1,2,0), \mathbf{v}_2 = (-k,3,0)$, $\mathbf{v}_3 = (2,0,1)$ *costituiscono una base* $\mathcal{B}$ *di* $\mathbb{R}^3$. *Posto* $k = 0$, *si determinino le coordinate del vettore* $\mathbf{w} = (-2,0,-1)$ *rispetto alla base* $\mathcal{B}$.

### Esercizio 4.5.14

*Si determinino, se possibile, 4 vettori non nulli di* $\mathbb{R}_2[x]$ *che non generino* $\mathbb{R}_2[x]$. *Si trovino, se possibile, due sottospazi distinti di* $\mathbb{R}_2[x]$ *di dimensione 2 che contengano entrambi il vettore* $x^2 + x$.

### Esercizio 4.5.15

*Sia* $X = \{\left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right) \in \mathrm{M}_2(\mathbb{R}) | a, b \in \mathbb{R}\}$. *Si dimostri che* $X$ *è un sottospazio di* $\mathrm{M}_2(\mathbb{R})$ *e se ne determini la dimensione.*

### Esercizio 4.5.16

*Si dimostri che* $\mathcal{B} = \{\left(\begin{smallmatrix} -1 \\ 2 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)\}$ *è una base di* $\mathbb{R}^2$ *e si determinino le coordinate dei vettori* $\left(\begin{smallmatrix} 3 \\ -1 \end{smallmatrix}\right)$ *e* $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$ *rispetto a tale base.*

### Esercizio 4.5.17

a) *Si stabilisca per quali valori di* $k$ *i vettori* $\mathbf{v}_1 = (-1,0,2k), \mathbf{v}_2 = (1,k,3k), \mathbf{v}_3 = (0,4k,2)$ *generano* $\mathbb{R}^3$.
b) *Scelto a piacere un valore di* $k$, *si stabilisca se il vettore* $(2,k,k)$ *appartiene a* $\langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle$.

**Esercizio 4.5.18**

a) *Si stabilisca per quali valori di $k$ i vettori* $\mathbf{v}_1 = x^2 + 2x + 2$, $\mathbf{v}_2 = -x^2 + 2kx + k - 1$, $\mathbf{v}_3 = kx^2 + (2k+4)x + 3k$ *sono linearmente indipendenti.*
b) *Posto* $k = 0$, *si determini, se possibile, un vettore* $\mathbf{w}$ *che non appartenga a* $\langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle$.

**Esercizio 4.5.19**

*Si determinino, se possibile, 4 vettori* $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ *di* $\mathbb{R}_2[x]$ *che abbiano contemporaneamente le seguenti proprietà:*

a) *nessuno di essi è multiplo di un altro;*

b) $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ *generano* $\mathbb{R}_2[x]$;

c) $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ *non generano* $\mathbb{R}_2[x]$.

**Esercizio 4.5.20**

*Si stabilisca se* $X = \{sx^3 + rx^2 + sx - 2r \mid r, s \in \mathbb{R}\} \subseteq \mathbb{R}^3[x]$ *è un sottospazio vettoriale di* $\mathbb{R}_3[x]$ *e, in caso affermativo, se ne determini una base.*

**Esercizio 4.5.21**

*Si determini una base del sottospazio di* $\mathbb{R}^4$ *dato da* $A \cap B$ *dove* $A = \langle (1,0,0,0), (0,1,0,0) \rangle$ *e* $B = \{(x,y,z,t) \in \mathbb{R}^4 \mid y = t = 0\}$.

**Esercizio 4.5.22**

*Si dimostri che uno spazio vettoriale non nullo di dimensione finita ha infinite basi.*

## 4.6 APPENDICE: IL TEOREMA DEL COMPLETAMENTO

Riportiamo qui di seguito, per il lettore interessato, la dimostrazione del Teorema del completamento 4.2.1. È necessario prima dimostrare un lemma tecnico.

**Lemma di sostituzione 4.6.1** *Se* $\mathcal{B} = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ *è una base di* $V$ *e* $\mathbf{v}$ *è un vettore del tipo* $\mathbf{v} = \lambda_1 \mathbf{w}_1 + \cdots + \lambda_k \mathbf{w}_k + \cdots + \lambda_n \mathbf{w}_n$ *con* $\lambda_k \neq 0$, *allora anche* $\mathcal{B}' = \{\mathbf{w}_1, \ldots, \mathbf{w}_{k-1}, \mathbf{v}, \mathbf{w}_{k+1}, \ldots, \mathbf{w}_n\}$ *è una base di* $V$.

**Dimostrazione** – Osserviamo che $\mathbf{w}_1, \ldots, \mathbf{w}_{k-1}, \mathbf{w}_{k+1}, \ldots, \mathbf{w}_n \in \langle \mathcal{B}' \rangle$, ma, poiché

$$\mathbf{v} = \lambda_1 \mathbf{w}_1 + \cdots + \lambda_k \mathbf{w}_k + \cdots + \lambda_n \mathbf{w}_n$$

con $\lambda_k \neq 0$, abbiamo anche che:

$$\mathbf{w}_k = -\frac{\lambda_1}{\lambda_k}\mathbf{w}_1 - \cdots + \frac{1}{\lambda_k}\mathbf{v} - \cdots - \frac{\lambda_n}{\lambda_k}\mathbf{w}_n \in \langle \mathcal{B}' \rangle$$

Quindi $\mathcal{B} \subseteq \langle \mathcal{B}' \rangle$. Poiché $\mathcal{B}$ genera $V$, abbiamo che $V = \langle \mathcal{B} \rangle \subseteq \langle \mathcal{B}' \rangle$. L'inclusione $\langle \mathcal{B}' \rangle \subseteq V$ è ovvia, quindi si ha che $\langle \mathcal{B}' \rangle = V$, cioè $\mathcal{B}'$ genera $V$.

Mostriamo ora che i vettori di $\mathcal{B}'$ sono linearmente indipendenti.

Siano $\beta_1, \ldots, \beta_n \in \mathbb{R}$ tali che:

$$\beta_1 \mathbf{w}_1 + \cdots + \beta_k \mathbf{v} + \cdots + \beta_n \mathbf{w}_n = \mathbf{0}$$

Sostituendo $\lambda_1 \mathbf{w}_1 + \cdots + \lambda_k \mathbf{w}_k + \cdots + \lambda_n \mathbf{w}_n$ al posto di $\mathbf{v}$ e riordinando si ottiene che:

$$(\beta_1 + \beta_k \lambda_1)\mathbf{w}_1 + \cdots + \beta_k \lambda_k \mathbf{w}_k + \cdots + (\beta_n + \beta_k \lambda_n)\mathbf{w}_n = \mathbf{0}$$

Poiché i $\mathbf{w}_i$ sono linearmente indipendenti, abbiamo che tutti i loro coefficienti devono essere nulli. In particolare deve succedere che $\beta_k \lambda_k = 0$ e $\beta_i + \beta_k \lambda_i = 0$ per ogni $i \neq k$. Dalla prima uguaglianza, essendo $\lambda_k \neq 0$, segue $\beta_k = 0$, e, per sostituzione, dalle altre si ottiene $\beta_i = 0$ per ogni $i \neq k$. Quindi tutti i $\beta_i$ sono nulli, il che dimostra che i vettori di $\mathcal{B}'$ sono linearmente indipendenti. □

Possiamo ora dimostrare il Teorema del completamento 4.2.1.

**Teorema del completamento** *Sia $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ un insieme di vettori linearmente indipendenti in uno spazio vettoriale $V$ finitamente generato. Se $\mathcal{B} = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ è una base di $V$ (sappiamo che ne esiste sempre almeno una), allora $m \leq n$ e possiamo sempre aggiungere a $\mathbf{v}_1, \ldots, \mathbf{v}_m$ $n - m$ vettori di $\mathcal{B}$, in modo da ottenere una base di $V$.*

**Dimostrazione** – Poiché $\mathcal{B}$ genera $V$, possiamo scrivere $\mathbf{v}_1$ nella forma

$$\mathbf{v}_1 = \alpha_1 \mathbf{w}_1 + \cdots + \alpha_n \mathbf{w}_n$$

dove non tutti i coefficienti sono nulli. Eventualmente, riordinando i $\mathbf{w}_k$, possiamo supporre che sia $\alpha_1 \neq 0$, e quindi per il Lemma di sostituzione 4.6.1 abbiamo che $\{\mathbf{v}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ è una base di $V$.

Ora si ripete il ragionamento considerando $\mathbf{v}_2$ e la base $\{\mathbf{v}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$. Possiamo scrivere $\mathbf{v}_2$ nella forma:

$$\mathbf{v}_2 = \beta_1 \mathbf{v}_1 + \beta_2 \mathbf{w}_2 + \cdots + \beta_n \mathbf{w}_n$$

ove almeno uno dei $\beta_j$ con $j \geq 2$ è non nullo, altrimenti $\mathbf{v}_2$ sarebbe un multiplo di $\mathbf{v}_1$, contraddicendo l'ipotesi che i vettori $\mathbf{v}_1, \ldots, \mathbf{v}_m$ siano linearmente indipendenti. Non è restrittivo supporre che sia $\beta_2 \neq 0$, e quindi per il Lemma di sostituzione 4.6.1 abbiamo che $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{w}_3, \ldots, \mathbf{w}_n\}$ è una base di $V$.

Si prosegue poi allo stesso modo. All'$i$-esimo passo possiamo supporre che $\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{w}_i, \ldots, \mathbf{w}_n\}$ sia una base di $V$. Possiamo scrivere $\mathbf{v}_i$ nella forma:

$$\mathbf{v}_i = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_{i-1} \mathbf{v}_{i-1} + \lambda_i \mathbf{w}_i + \cdots + \lambda_n \mathbf{w}_n$$

ove almeno uno dei $\lambda_j$ con $j \geq i$ è non nullo, altrimenti avremmo che $\mathbf{v}_i \in \langle \mathbf{v}_1, \ldots, \mathbf{v}_{i-1} \rangle$ contraddicendo l'ipotesi che i vettori $\mathbf{v}_1, \ldots, \mathbf{v}_m$ sono linearmente indipendenti. Non è restrittivo supporre che sia $\lambda_i \neq 0$, e quindi per il Lemma di sostituzione 4.6.1 abbiamo che $\{\mathbf{v}_1, \ldots, \mathbf{v}_i, \mathbf{w}_{i+1}, \ldots, \mathbf{w}_n\}$ è una base di $V$.

Se $m \leq n$, eventualmente riordinando opportunamente i vettori $\mathbf{w}_k$, dopo $m$ passi arriviamo a ottenere che $\{\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{w}_{m+1}, \ldots, \mathbf{w}_n\}$ è una base di $V$, come volevamo.

Se $m > n$ dopo $n$ passi otteniamo che $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ è una base di $V$, da cui segue che $\mathbf{v}_{n+1} \in \langle \mathbf{v}_1, \ldots, \mathbf{v}_n \rangle$, ma questo contraddice l'ipotesi che i vettori $\mathbf{v}_1, \ldots, \mathbf{v}_m$ siano linearmente indipendenti. Quindi deve essere $m \leq n$, e questo conclude la dimostrazione. $\square$

# 5 Applicazioni lineari

Le applicazioni lineari sono funzioni tra spazi vettoriali che ne rispettano la struttura, cioè sono compatibili con le operazioni di somma tra vettori e moltiplicazione di un vettore per uno scalare. Come vedremo le applicazioni lineari si rappresentano in modo molto efficace attraverso le matrici. Lo scopo di questo capitolo è quello di introdurre il concetto di applicazione lineare e capire come sia possibile associare univocamente una matrice a ogni applicazione lineare tra $\mathbb{R}^n$ e $\mathbb{R}^m$, una volta fissata in entrambi gli spazi la base canonica. Studieremo poi il nucleo e l'immagine di un'applicazione lineare fino ad arrivare al Teorema della dimensione, che rappresenta uno dei risultati più importanti della teoria sugli spazi vettoriali di dimensione finita.

## 5.1 DEFINIZIONE DI APPLICAZIONE LINEARE

In analisi matematica si studiano le funzioni reali a valori reali: sono leggi che associano a un numero reale in un certo insieme, detto *dominio* della funzione, un altro numero reale, appartenente al *codominio* della funzione.

Per esempio, $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = x^2$ è una funzione da $\mathbb{R}$ a $\mathbb{R}$ che ben conosciamo, il suo grafico è la parabola di equazione $y = x^2$. Come sappiamo $\mathbb{R}$ è anche uno spazio vettoriale, quindi $f$ è una funzione tra lo spazio vettoriale di partenza $\mathbb{R}$ e lo spazio vettoriale di arrivo $\mathbb{R}$. Però questa funzione non si comporta bene rispetto alla struttura di spazio vettoriale di $\mathbb{R}$. Infatti, se prendiamo due vettori $\mathbf{u}$ e $\mathbf{v}$ e ne facciamo la somma $\mathbf{u}+\mathbf{v}$, abbiamo subito che $f(\mathbf{u}+\mathbf{v}) \neq f(\mathbf{u})+f(\mathbf{v})$ (disegnando un grafico lo si constata immediatamente). Nello stesso modo possiamo vedere che anche la moltiplicazione per uno scalare non è rispettata. Per esempio, vediamo che $f(2 \cdot 3) = 6^2 \neq 2 \cdot f(3) = 2 \cdot 9$.

Ci sono invece altre funzioni $f : \mathbb{R} \longrightarrow \mathbb{R}$ che si comportano bene rispetto alla struttura di spazio vettoriale, cioè verificano le uguaglianze $f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$ e $f(\lambda\mathbf{v}) = \lambda f(\mathbf{v})$. Consideriamo, per esempio, la funzione $f(x) = 3x$. Osserviamo subito che $f(x_1 + x_2) = 3(x_1 + x_2) = 3x_1 + 3x_2 = f(x_1) + f(x_2)$ e che $f(\lambda x_1) = 3\lambda x_1 = \lambda(3x_1) = \lambda f(x_1)$.

Come vedremo tra poco, chiamiamo *applicazioni lineari* quelle funzioni tra spazi vettoriali che ne rispettano la struttura, cioè la somma di vettori ha per immagine, tramite la funzione, la somma delle immagini dei vettori e il prodotto di un vettore per uno scalare ha per immagine il prodotto dello scalare per l'immagine del vettore.

Prima della definizione formale di applicazione lineare richiamiamo la definizione di funzione e di immagine.

**Definizione 5.1.1** Definiamo *funzione* $f$ tra due insiemi $A$ e $B$ una legge che associa a un elemento di $A$ uno e un solo elemento di $B$ e denotiamo questa legge come:

$$\begin{array}{rcl} f : A & \longrightarrow & B \\ a & \mapsto & f(a) \end{array}$$

L'insieme $A$ si dice *dominio* della funzione, mentre l'insieme $B$ si dice *codominio* della funzione. Definiamo *immagine* di un elemento $a \in A$ l'elemento $f(a) \in B$. L'insieme delle immagini di tutti gli elementi di $A$ si dice *immagine* di $f$ e si indica con Im $(f)$ o talvolta con $f(A)$.

Non tutte le leggi che associano a elementi di un insieme elementi di un altro insieme sono funzioni. Per esempio, possiamo definire una legge che va dall'insieme $A$ degli esseri umani all'insieme $B$ degli esseri umani ($A$ e $B$ possono essere lo stesso insieme) che a ogni persona associa un fratello. Questa non è una funzione poiché qualcuno può avere più di un fratello.

Un altro esempio. Consideriamo la legge dall'insieme dei numeri naturali all'insieme dei numeri naturali che a ogni numero associa un suo divisore. Anche questa legge non è una funzione.

Andiamo ora a definire le applicazioni lineari.

**Definizione 5.1.2** Siano $V$ e $W$ due spazi vettoriali e sia $F : V \longrightarrow W$ una funzione; $F$ si dice *applicazione lineare* se:
1. $F(\mathbf{u} + \mathbf{v}) = F(\mathbf{u}) + F(\mathbf{v})$, per ogni $\mathbf{u}, \mathbf{v} \in V$,
2. $F(\lambda \mathbf{u}) = \lambda F(\mathbf{u})$ per ogni $\lambda \in \mathbb{R}$ e per ogni $\mathbf{u} \in V$.

♦ **Osservazione 5.1.3** È facile verificare che, se $V$ e $W$ sono due spazi vettoriali, una funzione $F : V \longrightarrow W$ è un'applicazione lineare se e solo se $F(\lambda\mathbf{u}+\mu\mathbf{v}) = \lambda F(\mathbf{u})+\mu F(\mathbf{v})$, per ogni $\mathbf{u}, \mathbf{v} \in V$ e per ogni $\lambda, \mu \in \mathbb{R}$.

**Proposizione 5.1.4** *Sia $F : V \to W$ un'applicazione lineare; allora $F(\mathbf{0}_V) = \mathbf{0}_W$.*

**Dimostrazione** – Sia $\mathbf{v}$ un qualsiasi vettore di $V$. Si ha:

$$F(\mathbf{0}_V) = F(0\mathbf{v}) = 0F(\mathbf{v}) = \mathbf{0}_W$$

in virtù delle proprietà in 5.1.2 e in 2.3.5. □

---

**Esempio 5.1.5**

i) Sia $F : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ la funzione definita da: $F(x, y) = (x+1, x-y)$. Vogliamo determinare se $F$ è lineare, cioè vedere se sono verificate le proprietà 1 e 2 della Definizione 5.1.2. Vediamo la proprietà 1.

$$F((x_1, y_1) + (x_2, y_2)) = (x_1 + x_2 + 1, x_1 + x_2 - y_1 - y_2)$$

D'altra parte:

$$F(x_1, y_1) + F(x_2, y_2) = (x_1 + x_2 + 2, x_1 + x_2 - y_1 - y_2)$$

Dunque la funzione data non è un'applicazione lineare.

Avremmo potuto concludere anche in modo più rapido, osservando che $F(0,0) = (1,0) \neq (0,0)$, dunque, per la proposizione precedente, l'applicazione non è lineare.

ii) Sia $D : \mathbb{R}[x] \longrightarrow \mathbb{R}[x]$ definita da: $D(p(x)) = p'(x)$, cioè $D$ è la funzione che a un polinomio associa la sua derivata. Dall'analisi sappiamo che $D(p(x) + q(x)) = D(p(x)) + D(q(x))$, cioè la derivata della somma di due polinomi è la somma delle derivate. Inoltre sappiamo anche che $D(kp(x)) = kD(p(x))$, per ogni costante $k \in \mathbb{R}$. Quindi abbiamo dimostrato che la derivata $D$ è un'applicazione lineare. Invitiamo lo studente a verificare in modo analogo, che anche l'integrale è un'applicazione lineare da $\mathbb{R}[x]$ a $\mathbb{R}[x]$.

iii) L'applicazione identità $\mathrm{id} : V \to V$ definita da: $\mathbf{v} \mapsto \mathbf{v}$ per ogni $\mathbf{v} \in V$ è un'applicazione lineare.

iv) L'applicazione nulla $T : V \to V$ definita da: $\mathbf{v} \mapsto \mathbf{0}_V$ per ogni $\mathbf{v} \in V$ è un'applicazione lineare.

v) Data una base $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ di $V$, l'applicazione che associa a ogni vettore $\mathbf{v} = \lambda_1\mathbf{v}_1 + \dots + \lambda_n\mathbf{v}_n$ le sue coordinate $(\lambda_1, \dots, \lambda_n)$ rispetto alla base $\mathcal{B}$ è un'applicazione lineare. Si lascia la facile verifica per esercizio.

$$\begin{array}{cccc} c: & V & \longrightarrow & \mathbb{R}^n \\ & \mathbf{v} = \lambda_1\mathbf{v}_1 + \dots + \lambda_n\mathbf{v}_n & \mapsto & (\lambda_1, \dots, \lambda_n) \end{array}$$

---

La seguente osservazione è di importanza capitale per comprendere la corrispondenza tra applicazioni lineari e matrici. Riprenderemo questo concetto più in dettaglio anche nelle sezioni successive.

♦ **Osservazione 5.1.6** A ogni matrice

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in \mathrm{M}_{m,n}(\mathbb{R})$$

possiamo associare la funzione $L_A : \mathbb{R}^n \to \mathbb{R}^m$ così definita:

$$\begin{array}{cccc} L_A: & \mathbb{R}^n & \to & \mathbb{R}^m \\ & \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} & \mapsto & \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix} \end{array}$$

$$= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

ove il prodotto di $A$ per il vettore $(x_1, \dots, x_n)$ è il *prodotto righe per colonne* definito nel Capitolo 1.

È facile verificare che $L_A$ è un'applicazione lineare. La proprietà 1 della Definizione 5.1.2 vale per la Proposizione 1.2.1 (distributività del prodotto righe per colonne rispetto alla somma) e la proprietà 2 è una semplice verifica.

Più concisamente possiamo scrivere:

$$L_A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Vediamo un esempio concreto. Se consideriamo la matrice:

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

si ha che l'applicazione lineare $L_A : \mathbb{R}^3 \to \mathbb{R}^2$ è definita da:

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

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

Si noti che si ha:

$$L_A \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \qquad L_A \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad L_A \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$$

cioè: *le immagini dei vettori della base canonica sono le colonne della matrice $A$.*

Questo fatto si rivelerà di importanza fondamentale per le applicazioni, quando dovremo determinare l'immagine di un'applicazione lineare.

Ci chiediamo ora come deve essere fatta una funzione $F : \mathbb{R} \longrightarrow \mathbb{R}$, affinché sia lineare. Sia $F(1) = a \in \mathbb{R}$. Osserviamo che, per la proprietà 1 di 5.1.2 si ha che $F(x) = xF(1) = ax$. Dunque le uniche applicazioni lineari da $\mathbb{R}$ a $\mathbb{R}$ hanno come grafico una retta passante per l'origine (da cui appunto il nome di *applicazione lineare*). Questo esempio è particolarmente istruttivo perché ci ha mostrato che, per conoscere completamente un'applicazione lineare da $\mathbb{R}$ a $\mathbb{R}$, è sufficiente conoscerne un solo valore; noi abbiamo scelto $F(1)$, ma lo studente può convincersi che il valore di $F$ in qualunque altro punto (purché diverso da zero) avrebbe determinato completamente $F$. Ciò è vero in generale, cioè un'applicazione lineare è completamente determinata conoscendone solo alcuni valori e precisamente i valori corrispondenti ai vettori di una base del dominio (e infatti nell'esempio considerato qualunque numero diverso da zero è base di $\mathbb{R}$). Questo fatto basilare è codificato dal seguente teorema.

**Teorema 5.1.7** *Siano $V$ e $W$ due spazi vettoriali. Se $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ è una base di $V$ e consideriamo $n$ vettori $\mathbf{w}_1, \dots, \mathbf{w}_n \in W$, non necessariamente distinti, allora esiste ed è unica un'applicazione lineare $L : V \to W$ tale che $L(\mathbf{v}_1) = \mathbf{w}_1, \dots, L(\mathbf{v}_n) = \mathbf{w}_n$.*

**Dimostrazione** – Poiché $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ è una base di $V$ esistono, e sono unici, gli scalari $\alpha_1, \dots, \alpha_n$ (coordinate di $\mathbf{v}$) tali che $\mathbf{v} = \alpha_1\mathbf{v}_1 + \dots + \alpha_n\mathbf{v}_n$. Definiamo $L(\mathbf{v})$ nel modo seguente:

$$L(\mathbf{v}) = \alpha_1\mathbf{w}_1 + \cdots + \alpha_n\mathbf{w}_n$$

Verifichiamo che tale $L$ è lineare e cioè che valgono le proprietà 1 e 2 della Definizione 5.1.2. Siano $\mathbf{v} = \alpha_1\mathbf{v}_1 + \dots + \alpha_n\mathbf{v}_n$, $\mathbf{u} = \beta_1\mathbf{v}_1 + \dots + \beta_n\mathbf{v}_n$. Abbiamo che $\mathbf{v} + \mathbf{u} = (\alpha_1 + \beta_1)\mathbf{v}_1 + \dots + (\alpha_n + \beta_n)\mathbf{v}_n$ e, se $\lambda \in \mathbb{R}$, abbiamo $\lambda\mathbf{v} = \lambda\alpha_1\mathbf{v}_1 + \dots + \lambda\alpha_n\mathbf{v}_n$. Inoltre:

$$\begin{aligned} L(\mathbf{v} + \mathbf{u}) &= L((\alpha_1 + \beta_1)\mathbf{v}_1 + \dots + (\alpha_n + \beta_n)\mathbf{v}_n) \\ &= (\alpha_1 + \beta_1)\mathbf{w}_1 + \dots + (\alpha_n + \beta_n)\mathbf{w}_n \\ &= \alpha_1\mathbf{w}_1 + \dots + \alpha_n\mathbf{w}_n + \beta_1\mathbf{w}_1 + \dots + \beta_n\mathbf{w}_n = L(\mathbf{v}) + L(\mathbf{u}) \end{aligned}$$

Ora vediamo
$$\begin{aligned} L(\lambda\mathbf{v}) &= L(\lambda\alpha_1\mathbf{v}_1 + \dots + \lambda\alpha_n\mathbf{v}_n) \\ &= \lambda\alpha_1\mathbf{w}_1 + \dots + \lambda\alpha_n\mathbf{w}_n \\ &= \lambda(\alpha_1\mathbf{w}_1 + \dots + \alpha_n\mathbf{w}_n) = \lambda L(\mathbf{v}) \end{aligned}$$

Quindi $L$ è un'applicazione lineare.

Vediamo ora l'unicità. Supponiamo che $G$ sia un'applicazione lineare $G : V \to W$ tale che $G(\mathbf{v}_1) = \mathbf{w}_1, \dots, G(\mathbf{v}_n) = \mathbf{w}_n$ e che $L$ sia l'applicazione lineare definita sopra. Si ha allora che:

$$\begin{aligned} G(\mathbf{v}) &= G(\alpha_1\mathbf{v}_1 + \dots + \alpha_n\mathbf{v}_n) = \alpha_1 G(\mathbf{v}_1) + \cdots + \alpha_n G(\mathbf{v}_n) \\ &= \alpha_1\mathbf{w}_1 + \cdots + \alpha_n\mathbf{w}_n = L(\mathbf{v}) \end{aligned}$$

Quindi $G = L$, come volevasi dimostrare. □

**Corollario 5.1.8** *Siano $V$ e $W$ due spazi vettoriali. Se due applicazioni lineari $T, S : V \to W$ coincidono su di una base di $V$, allora coincidono su tutto $V$.*

Come vedremo il Teorema 5.1.7 è fondamentale per stabilire una corrispondenza biunivoca tra applicazioni lineari tra spazi vettoriali con basi fissate e matrici. Prima di procedere vogliamo fermarci a considerare il fatto straordinario che un'applicazione lineare da $\mathbb{R}^n$ a $\mathbb{R}^m$ è nota se si conoscono i suoi valori su di un insieme di $n$ elementi! Se ricordiamo le funzioni studiate in analisi, ciò è ben lontano dal vero, cioè nello studio di una funzione $f : \mathbb{R} \longrightarrow \mathbb{R}$ non basta conoscere un valore di $f$, ma è necessario studiare faticosamente i suoi massimi e minimi, gli asintoti ecc. La differenza qui è che stiamo considerando applicazioni lineari, cioè funzioni con proprietà estremamente restrittive riguardo al loro comportamento.

## 5.2 APPLICAZIONI LINEARI E MATRICI

In questo paragrafo vogliamo esaminare diversi modi di scrivere e rappresentare le applicazioni lineari tra i due spazi vettoriali $\mathbb{R}^n$ e $\mathbb{R}^m$. Per il Teorema 5.1.7 sappiamo che un'applicazione lineare è determinata univocamente dai suoi valori su di una qualsiasi base. Vediamo in pratica come questo accade, tenendo ben presente l'Osservazione 5.1.6. Cominciamo con l'esaminare un esempio.

**Esempio 5.2.1**

Si consideri l'applicazione lineare $F : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ tale che $F(\mathbf{e}_1) = \mathbf{e}_2 + \mathbf{e}_3$, $F(\mathbf{e}_2) = 2\mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_3$, ove $\{\mathbf{e}_1, \mathbf{e}_2\}$ è la base canonica di $\mathbb{R}^2$, cioè $\mathbf{e}_1 = (1,0)$, $\mathbf{e}_2 = (0,1)$, e $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ è la base canonica di $\mathbb{R}^3$, cioè $\mathbf{e}_1 = (1,0,0)$, $\mathbf{e}_2 = (0,1,0)$, $\mathbf{e}_3 = (0,0,1)$. Ci chiediamo come determinare $F(x,y)$ per un generico vettore $(x,y) \in \mathbb{R}^2$. Per le proprietà 1 e 2 della Definizione 5.1.2 di applicazione lineare abbiamo che:

$$\begin{aligned} F(x,y) &= F(x(1,0) + y(0,1)) = F(x\mathbf{e}_1 + y\mathbf{e}_2) = xF(\mathbf{e}_1) + yF(\mathbf{e}_2) \\ &= x(\mathbf{e}_2 + \mathbf{e}_3) + y(2\mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_3) = x(0,1,1) + y(2,-1,1) \\ &= (2y, x-y, x+y) \end{aligned}$$

Vediamo ora ciò che accade in generale.

Sia $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ un'applicazione lineare definita da:

$$\begin{aligned} F(\mathbf{e}_1) &= a_{11}\mathbf{e}_1 + a_{21}\mathbf{e}_1 \cdots + a_{m1}\mathbf{e}_m \\ F(\mathbf{e}_2) &= a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_1 \cdots + a_{m2}\mathbf{e}_m \\ &\vdots \\ F(\mathbf{e}_n) &= a_{1n}\mathbf{e}_1 + a_{2n}\mathbf{e}_1 \cdots + a_{mn}\mathbf{e}_m \end{aligned}$$

Ci chiediamo come possiamo esprimere $F(x_1, \dots, x_n)$ cioè come possiamo scrivere l'immagine di un qualsiasi vettore $(x_1, \dots, x_n) \in \mathbb{R}^n$.

Procediamo esattamente come nell'esempio, il ragionamento è lo stesso, solo più complicato da scriversi in generale.

$$\begin{aligned} F(x_1, \dots, x_n) &= F(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n) \\ &= x_1F(\mathbf{e}_1) + x_2F(e_2) + \cdots + x_nF(\mathbf{e}_n) \\ &= x_1(a_{11}\mathbf{e}_1 + a_{12}\mathbf{e}_2 + \cdots + a_{m1}\mathbf{e}_m) \\ &\quad + x_2(a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_1 \cdots + a_{m2}\mathbf{e}_m) + \dots \\ &\quad + x_n(a_{1n}\mathbf{e}_1 + \cdots + a_{mn}\mathbf{e}_m) \\ &= (a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n)\mathbf{e}_1 \\ &\quad + (a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n)\mathbf{e}_2 + \dots \\ &\quad + (a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n)\mathbf{e}_m \\ &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{pmatrix} \end{aligned}$$

Facciamo un ulteriore passo, notando che $F(x_1, \dots, x_n)$ si può anche scrivere in modo più compatto usando la notazione di moltiplicazione di una matrice per un vettore:

$$F\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{ove} \quad A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

In pratica abbiamo appena dimostrato che $F$ è proprio l'applicazione $L_A$ associata ad $A$ descritta nell'Osservazione 5.1.6.

Raccogliamo tutte le nostre osservazioni nel seguente teorema, fondamentale per lo svolgimento degli esercizi. Indicheremo le basi canoniche di $\mathbb{R}^n$ e $\mathbb{R}^m$ rispettivamente con $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ e $\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$ (la notazione non è ambigua perché ogni vettore $\mathbf{e}_i$ è univocamente determinato non appena sappiamo a quale spazio vettoriale appartiene).

**Teorema 5.2.2** *Data un'applicazione lineare $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, fissiamo in $\mathbb{R}^n$ e $\mathbb{R}^m$ le rispettive basi canoniche. Allora possiamo equivalentemente rappresentare $F$ in uno dei seguenti tre modi:*

1. $F(\mathbf{e}_1) = a_{11}\mathbf{e}_1 + \dots + a_{m1}\mathbf{e}_m$
   $\vdots$
   $F(\mathbf{e}_n) = a_{1n}\mathbf{e}_1 + \dots + a_{mn}\mathbf{e}_m$

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

3. $F(\mathbf{x}) = A\mathbf{x}$, *ove*

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

**Corollario 5.2.3** *Esiste una corrispondenza biunivoca tra le matrici $m \times n$ e le applicazioni lineari tra gli spazi vettoriali $\mathbb{R}^n$ e $\mathbb{R}^m$, ove si siano fissate in entrambi gli spazi le basi canoniche per rappresentare i vettori. Più precisamente all'applicazione lineare $F(x_1, \dots, x_n) = (a_{11}x_1 + \dots + a_{1n}x_n, a_{21}x_1 + \dots + a_{2n}x_n, \dots, a_{m1}x_1 + \dots + a_{mn}x_n)$ è associata la matrice:*

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

*e viceversa.*

♦ **Osservazione 5.2.4** Osserviamo che, se all'applicazione $F : \mathbb{R}^n \to \mathbb{R}^m$ è associata la matrice $A$, rispetto alle basi canoniche di dominio e codominio, indicando come al solito con $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ la base canonica di $\mathbb{R}^n$, si ha che la $i$-esima colonna della matrice $A$ (che indicheremo con $A_i$) è $F(\mathbf{e}_i)$.

## 5.3 LA COMPOSIZIONE DI APPLICAZIONI LINEARI

In questo paragrafo ci occupiamo della composizione di applicazioni lineari. Se $F$ e $G$ sono applicazioni lineari, a cui sono associate rispettivamente due matrici $A$ e $B$ (rispetto alle basi canoniche) ed esiste la funzione composta $F \circ G$, si ha che $F \circ G$ è un'applicazione lineare e ci chiediamo quale sia la matrice ad essa associata.

Richiamiamo prima la definizione di funzione composta. Se $g : A \to B$ e $f : B \to C$ sono funzioni, e *il dominio di* $f$ *coincide con il codominio di* $g$ allora si definisce la *funzione composta* $f \circ g : A \to C$, che corrisponde al fare agire prima $g$ e poi $f$. Formalmente:

$$\begin{aligned} f \circ g : A &\longrightarrow C \\ a &\mapsto f(g(a)) \end{aligned}$$

Osserviamo che l'operazione di composizione è *associativa*, cioè, se abbiamo tre funzioni $h : A \to B$, $g : B \to C$, $f : C \to D$, si ha che $f \circ (g \circ h) = (f \circ g) \circ h$. Infatti, per ogni $a \in A$, si ha che

$$(f \circ (g \circ h))(a) = f((g \circ h)(a)) = f(g(h(a))) = (f \circ g)(h(a)) = ((f \circ g) \circ h)(a)$$

L'operazione di composizione non è commutativa, cioè in generale $f \circ g \neq g \circ f$, addirittura potrebbe succedere che $g \circ f$ non sia neppure definita.

**Esempio 5.3.1**

Siano

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ y &\mapsto y^2 - 1 \end{aligned} \qquad \begin{aligned} g : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto x + 2 \end{aligned}$$

Calcoliamo $f \circ g : \mathbb{R} \to \mathbb{R}$ e $g \circ f : \mathbb{R} \to \mathbb{R}$,

$$f \circ g : \; x \longmapsto g(x) = x + 2 \longmapsto f(x+2) = (x+2)^2 - 1 = x^2 + 2x + 3$$

$$g \circ f : y \longmapsto f(y) = y^2 - 1 \longmapsto g(y^2 - 1) = (y^2 - 1) + 2 = y^2 + 1$$

In questo caso $f \circ g \neq g \circ f$.

Vediamo ora un esempio di pertinenza dell'algebra lineare.

**Esempio 5.3.2**

Consideriamo le due applicazioni lineari $L_A : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$, $L_B : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ associate alle matrici:

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

rispetto alle basi canoniche in $\mathbb{R}^2$ e $\mathbb{R}^3$. Vediamo subito che $L_B \circ L_A$ è definita, mentre $L_A \circ L_B$ non è definita. Ciò accade perché $L_A$ deve avere come argomento un vettore in $\mathbb{R}^3$, mentre per ogni $\mathrm{v} \in \mathbb{R}^2$ abbiamo che $L_B(\mathrm{v}) \in \mathbb{R}^2$, dunque $L_A(L_B(\mathrm{v}))$ non ha senso.

Vediamo ora che la funzione composta di due applicazioni lineari, qualora tale composizione sia possibile, è ancora un'applicazione lineare.

**Proposizione 5.3.3** *Siano $U, V, W$ spazi vettoriali e siano $G : U \to V$, $F : V \to W$ due applicazioni lineari. Allora la funzione $F \circ G : U \to W$ è un'applicazione lineare.*

**Dimostrazione** – Dobbiamo verificare le due proprietà della Definizione 5.1.2.
1. Siano $\mathbf{u}_1, \mathbf{u}_2 \in U$. Allora $(F \circ G)(\mathbf{u}_1 + \mathbf{u}_2) = F(G(\mathbf{u}_1 + \mathbf{u}_2)) = F(G(\mathbf{u}_1) + G(\mathbf{u}_2)) = F(G(\mathbf{u}_1)) + F(G(\mathbf{u}_2)) = (F \circ G)(\mathbf{u}_1) + (F \circ G)(\mathbf{u}_2)$, dove abbiamo utilizzato prima la definizione di funzione composta, poi la linearità di $G$, la linearità di $F$ e infine di nuovo la definizione di funzione composta.
2. Siano $\mathbf{u} \in U$ e $\lambda \in \mathbb{R}$. Allora $(F \circ G)(\lambda \mathbf{u}) = F(G(\lambda \mathbf{u})) = F(\lambda(G(\mathbf{u})) = \lambda F(G(\mathbf{u})) = \lambda(F \circ G)(\mathbf{u})$, dove, come nel caso precedente, abbiamo utilizzato prima la definizione di funzione composta, poi la linearità di $G$, la linearità di $F$ e infine di nuovo la definizione di funzione composta. □

Da questa proposizione possiamo ricavare un facile corollario che è molto utile nelle applicazioni.

**Corollario 5.3.4** *Siano $L_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, $L_B : \mathbb{R}^m \longrightarrow \mathbb{R}^s$ due applicazioni lineari associate rispettivamente alle matrici $A$ e $B$. Allora l'applicazione lineare $L_B \circ L_A$ è associata alla matrice $BA$, cioè:*

$$L_B \circ L_A = L_{BA}$$

**Dimostrazione** – La dimostrazione è una semplice verifica:

$$(L_B \circ L_A)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = L_B\left(A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = B\left(A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right)$$

$$= (BA)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = L_{BA}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

In questa verifica abbiamo utilizzato la proprietà associativa del prodotto righe per colonne tra matrici.

□

## 5.4 NUCLEO E IMMAGINE

Vogliamo ora entrare nel vivo della teoria delle applicazioni lineari e introdurre i concetti di *nucleo* e *immagine*, che sono sottospazi vettoriali rispettivamente del dominio e del codominio di un'applicazione lineare.

**Definizione 5.4.1** Siano $V$ e $W$ due spazi vettoriali e sia $L : V \to W$ un'applicazione lineare. Si dice *nucleo* di $L$ l'insieme dei vettori di $V$ la cui immagine è il vettore nullo di $W$. Tale insieme si indica con $\operatorname{Ker} L$.

$$\operatorname{Ker}(L) = \{\mathbf{v} \in V \mid L(\mathbf{v}) = 0_W\}$$

Si dice *immagine* di $L$ l'insieme dei vettori di $W$ che sono immagini di qualche vettore nel codominio $W$, cioè

$$\mathrm{Im}\,(L) = \{\mathbf{w} \in W \mid \mathbf{w} = L(\mathbf{v}) \text{ per qualche } \mathbf{v} \in V\}$$

Vediamo subito qualche esempio.

**Esempio 5.4.2**

i) Consideriamo la funzione derivazione $D : \mathbb{R}[x] \longrightarrow \mathbb{R}[x]$, definita da $D(p(x)) = p'(x)$. Come abbiamo visto $D$ è un'applicazione lineare. Ci chiediamo quali siano i polinomi $p(x)$ tali che la loro immagine sia zero, cioè tali che $D(p(x)) = 0$. Dall'analisi sappiamo che sono tutti e soli i polinomi costanti. Dunque $\mathrm{Ker}\,(D) = \{c \mid c \in \mathbb{R}\}$. Vediamo ora l'immagine di $D$. Ci chiediamo quali siano i polinomi che sono derivata di altri polinomi. Dall'analisi sappiamo che sono tutti i polinomi (che infatti si possono integrare), quindi $\mathrm{Im}\,(D) = \mathbb{R}[x]$.

ii) Consideriamo ora l'applicazione lineare $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ definita da: $L(\mathbf{e}_1) = 2\mathbf{e}_1 - \mathbf{e}_2$, $L(\mathbf{e}_2) = \mathbf{e}_1$, $L(\mathbf{e}_3) = \mathbf{e}_1 + 2\mathbf{e}_2$.

Per quanto visto nel Teorema 5.2.2, sappiamo che $L(x,y,z) = (2x + y + z, -x + 2z)$. Vogliamo determinare nucleo e immagine di $L$. Il nucleo di $L$ è l'insieme dei vettori la cui immagine è il vettore nullo, cioè

$$\begin{aligned}
\mathrm{Ker}\,(L) &= \{(x,y,z) \mid (2x + y + z, -x + 2z) = (0,0)\} \\
&= \{(x,y,z) \mid 2x + y + z = 0, -x + z = 0\} \\
&= \left\{(x,y,z) \mid z = \frac{1}{2}x, y = -2x - z = -\frac{5}{2}x\right\} = \left\{\left(x, -\frac{5}{2}x, \frac{1}{2}x\right) \mid x \in \mathbb{R}\right\} \\
&= \langle(2,-5,1)\rangle
\end{aligned}$$

Vediamo ora l'immagine.

$$\begin{aligned}
\mathrm{Im}\,(L) &= \left\{\mathbf{w} \in \mathbb{R}^2 \mid \mathbf{w} = \begin{pmatrix} 2x + y + z \\ -x + 2z \end{pmatrix} \text{ con } x,y,z \in \mathbb{R}\right\} \\
&= \left\{\mathbf{w} \in \mathbb{R}^2 \mid \mathbf{w} = \begin{pmatrix} 2x \\ -x \end{pmatrix} + \begin{pmatrix} y \\ 0 \end{pmatrix} + \begin{pmatrix} z \\ 2z \end{pmatrix} \text{ con } x,y,z \in \mathbb{R}\right\} \\
&= \left\{\mathbf{w} \in \mathbb{R}^2 \mid \mathbf{w} = x\begin{pmatrix} 2 \\ -1 \end{pmatrix} + y\begin{pmatrix} 1 \\ 0 \end{pmatrix} + z\begin{pmatrix} 1 \\ 2 \end{pmatrix} \text{ con } x,y,z \in \mathbb{R}\right\} \\
&= \left\langle\begin{pmatrix} 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}\right\rangle
\end{aligned}$$

Il fatto che le applicazioni lineari siano definite in modo da conservare le due operazioni proprie degli spazi vettoriali, fa sì che sia nucleo sia immagine di una data applicazione lineare siano sottospazi vettoriali.

**Proposizione 5.4.3** *Sia $L : V \longrightarrow W$ un'applicazione lineare.*
1) *Il nucleo di $L$ è un sottospazio vettoriale del dominio $V$.*
2) *L'immagine di $L$ è un sottospazio vettoriale del codominio $W$.*

**Dimostrazione** – 1) Osserviamo innanzitutto che $\mathrm{Ker}\,(L)$ non è l'insieme vuoto, perché $\mathbf{0}_V \in \mathrm{Ker}\,(L)$ per la Proposizione 5.1.4. Dobbiamo poi verificare che $\mathrm{Ker}\,(L)$ sia chiuso rispetto alla somma di vettori e alla moltiplicazione di un vettore per uno scalare. Cominciamo con la somma. Siano $\mathbf{u}, \mathbf{v} \in \mathrm{Ker}\,(L)$. Allora $L(\mathbf{u}) = L(\mathbf{v}) = \mathbf{0}_W$ e quindi $L(\mathbf{u}+\mathbf{v}) = L(\mathbf{u}) + L(\mathbf{v}) = \mathbf{0}_W + \mathbf{0}_W = \mathbf{0}_W$, quindi $\mathbf{u}+\mathbf{v} \in \mathrm{Ker}\,(L)$. Verifichiamo ora la chiusura di $L$ rispetto al prodotto per uno scalare. Se $\alpha \in \mathbb{R}$ e $\mathbf{u} \in \mathrm{Ker}\,(L)$ si ha $L(\alpha\mathbf{u}) = \alpha L(\mathbf{u}) = \alpha\mathbf{0}_W = \mathbf{0}_W$, quindi $\alpha\mathbf{u} \in \mathrm{Ker}\,(L)$.
2) Vediamo ora le stesse due proprietà per $\mathrm{Im}\ (L)$. Abbiamo che $\mathbf{0}_W \in \mathrm{Im}\ (L)$ per la Proposizione 5.1.4. Siano ora $\mathbf{w}_1, \mathbf{w}_2 \in \mathrm{Im}L$. Allora esistono $\mathbf{v}_1, \mathbf{v}_2 \in V$ tali che $L(\mathbf{v}_1) = \mathbf{w}_1$ e $L(\mathbf{v}_2) = \mathbf{w}_2$. Dunque $\mathbf{w}_1 + \mathbf{w}_2 = L(\mathbf{v}_1) + L(\mathbf{v}_2) = L(\mathbf{v}_1 + \mathbf{v}_2) \in \mathrm{Im}\ L$ e $\alpha\mathbf{w}_1 = \alpha L(\mathbf{v}_1) = L(\alpha\mathbf{v}_1) \in \mathrm{Im}(L)$ per ogni $\alpha \in \mathbb{R}$. □

**Proposizione 5.4.4** *Sia $L : V \longrightarrow W$ un'applicazione lineare. Allora il sottospazio vettoriale* $\mathrm{Im}\ (L)$ *è generato dall'immagine di una qualsiasi base di $V$, cioè, se $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ è una base di $V$, allora:*

$$\mathrm{Im}\ (L) = \langle L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n)\rangle$$

**Dimostrazione** – Mostriamo prima che $\mathrm{Im}\ (L) \subseteq \langle L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n)\rangle$. $\mathrm{Im}\ (L)$ è formata da tutti i vettori del tipo $L(\mathbf{v})$ al variare di $\mathbf{v} \in V$. Sia $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base di $V$. Se $\mathbf{v} \in V$ allora $\mathbf{v} = \lambda_1\mathbf{v}_1 + \cdots + \lambda_n\mathbf{v}_n$, con $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Dunque:

$$L(\mathbf{v}) = L(\lambda_1\mathbf{v}_1 + \cdots + \lambda_n\mathbf{v}_n) = \lambda_1 L(\mathbf{v}_1) + \cdots + \lambda_n L(\mathbf{v}_n) \in \langle L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n)\rangle$$

Questo prova che $\mathrm{Im}\ (L) \subseteq \langle L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n)\rangle$.

L'inclusione $\langle L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n)\rangle \subseteq \mathrm{Im}\ (L)$ è vera per la Proposizione 3.1.5, perché $\mathrm{Im}\ (L)$ è un sottospazio di $V$. □

Riprendiamo l'esempio precedente.

**Esempio 5.4.5**

Vogliamo determinare una base per l'immagine dell'applicazione lineare $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ definita da: $L(\mathbf{e}_1) = 2\mathbf{e}_1 - \mathbf{e}_2$, $L(\mathbf{e}_2) = \mathbf{e}_1$, $L(\mathbf{e}_3) = \mathbf{e}_1 + 2\mathbf{e}_2$. Per la proposizione precedente sappiamo che

$$\mathrm{Im}\ (L) = \langle 2\mathbf{e}_1 - \mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_1 + 2\mathbf{e}_2\rangle$$

Osserviamo che $\mathrm{Im}\ L \subseteq \mathbb{R}^2$, quindi $\mathrm{Im}\ L$ ha dimensione al più 2. Vediamo facilmente che due vettori tra i generatori di $\mathrm{Im}\ L$ sono linearmente indipendenti. Possiamo dunque concludere che sono una base di $\mathbb{R}^2$ e quindi che $\mathrm{Im}\ (L) = \mathbb{R}^2$.

Vedremo meglio nel prossimo paragrafo un metodo generale per determinare il nucleo e l'immagine di un'applicazione lineare.

Nucleo e immagine sono legate all'iniettività e suriettività dell'applicazione lineare data. Richiamiamo brevemente questi concetti basilari.

**Definizione 5.4.6** Sia data una funzione tra due insiemi $A$ e $B$,

$$f : A \longrightarrow B$$

1. Diciamo che $f$ è *iniettiva* quando accade che: "se $f(x) = f(y)$ allora $x = y$", cioè due elementi distinti $x$ e $y$ non possono mai avere la stessa immagine.

2. Diciamo che $f$ è *suriettiva* se ogni elemento di $B$ è immagine di un elemento di $A$, cioè il codominio di $f$ coincide con l'immagine di $f$.

3. Diciamo che $f$ è *biettiva o biunivoca* se è sia iniettiva sia suriettiva.

4. Diciamo che $f$ è *invertibile* se esiste una funzione $g : B \longrightarrow A$ detta *inversa* di $f$, tale che $f \circ g = \mathrm{id}_B$ e $g \circ f = \mathrm{id}_A$, ove in generale $\mathrm{id}_X : X \longrightarrow X$ è la funzione identità che associa a ogni elemento se stesso, cioè $\mathrm{id}_X(x) = x$. Molto spesso si indica l'inversa di $f$ con $f^{-1}$.

Nella prossima proposizione enunciamo il fatto che una funzione è biunivoca se e solo se è invertibile e pertanto d'ora in poi useremo questi due termini in modo intercambiabile.

**Proposizione 5.4.7** *Sia $f : A \longrightarrow B$ una funzione tra due insiemi $A$ e $B$, allora $f$ è biettiva se e solo se è invertibile.*

**Dimostrazione** – Supponiamo che $f$ sia biettiva e vogliamo costruire l'inversa $g$ di $f$. Sia $b \in B$. Poiché $f$ è biettiva, è in particolare suriettiva, quindi esiste $a \in A$ tale che $f(a) = b$. Inoltre tale $a$ è unico, perché $f$ è anche iniettiva e quindi, se anche $a'$ è tale che $f(a') = b$, allora $a = a'$. Definiamo $g : B \longrightarrow A$ con la regola $g(b) = a$. Ora per costruzione $f(g(b)) = f(a) = b$, quindi $f \circ g = \mathrm{id}_B$. Inoltre $g(f(a)) = g(b) = a$, quindi $g \circ f = \mathrm{id}_A$. Questo dimostra che $g$ è l'inversa di $f$.

Supponiamo ora che $f$ sia invertibile e che $g : B \longrightarrow A$ sia l'inversa di $f$. Quindi $f(g(b)) = b$ per ogni $b \in B$ e $g(f(a)) = a$ per ogni $a \in A$. Mostriamo che $f$ è iniettiva. Siano $a_1, a_2 \in A$ tali che $f(a_1) = f(a_2)$. Allora $g(f(a_1)) = g(f(a_2))$, quindi $a_1 = g(f(a_1)) = g(f(a_2)) = a_2$. Per la suriettività, sia $b \in B$ e poniamo $a = g(b)$, allora $f(a) = f(g(b)) = b$ e quindi $b$ è immagine di $a$ tramite $f$. □

Ritorniamo ora alle applicazioni lineari.

**Proposizione 5.4.8** *Sia $L : V \longrightarrow W$ un'applicazione lineare.*

1) *$L$ è iniettiva se e solo se* $\mathrm{Ker}\,(L) = \{\mathbf{0}_V\}$, *cioè il suo nucleo è il sottospazio nullo del dominio $V$.*

2) *$L$ è suriettiva se e solo se* $\mathrm{Im}\,(L) = W$, *cioè l'immagine coincide con il codominio.*

**Dimostrazione** – 1) Mostriamo che se $L$ è iniettiva allora $\mathrm{Ker}\,(L) = \{\mathbf{0}_V\}$. Se $\mathbf{u} \in \mathrm{Ker}\,(L)$ allora $L(\mathbf{u}) = \mathbf{0}_W = L(\mathbf{0}_V)$ e poiché $L$ è iniettiva, ciò implica che $\mathbf{u} = \mathbf{0}_V$.

Viceversa sia $\text{Ker}(L) = \{\mathbf{0}_V\}$ e supponiamo che $L(\mathbf{u}) = L(\mathbf{v})$ per qualche $\mathbf{u}, \mathbf{v} \in V$. Allora $L(\mathbf{u}-\mathbf{v}) = L(\mathbf{u}) - L(\mathbf{v}) = \mathbf{0}_W$ e dunque poiché $\mathbf{u}-\mathbf{v} \in \text{Ker}(L) = \{\mathbf{0}_V\}$ si ha che $\mathbf{u}-\mathbf{v} = \mathbf{0}_V$, quindi $\mathbf{u} = \mathbf{v}$, pertanto $f$ è iniettiva.

2) È esattamente la definizione di suriettività. □

La prossima proposizione ci dice che le applicazioni lineari iniettive conservano la lineare indipendenza.

**Proposizione 5.4.9** *Siano $V$ e $W$ spazi vettoriali, $\mathbf{v}_1, \dots, \mathbf{v}_r$ vettori di $V$ linearmente indipendenti e sia $L : V \to W$ un'applicazione lineare iniettiva. Allora $L(\mathbf{v}_1), \dots, L(\mathbf{v}_r)$ sono vettori di $W$ linearmente indipendenti.*

**Dimostrazione** – Sia $\alpha_1 L(\mathbf{v}_1) + \dots + \alpha_r L(\mathbf{v}_r) = \mathbf{0}_W$, con $\alpha_1, \dots, \alpha_r \in \mathbb{R}$. Allora per la Proposizione 5.1.4 si ha che $L(\alpha_1\mathbf{v}_1 + \dots + \alpha_r\mathbf{v}_r) = \mathbf{0}_W$ e, poiché $L$ è iniettiva, ne segue $\alpha_1\mathbf{v}_1 + \dots + \alpha_r\mathbf{v}_r = \mathbf{0}_V$. Ma, poiché $\mathbf{v}_1, \dots, \mathbf{v}_r$ sono vettori linearmente indipendenti, deve essere $\alpha_1 = \ldots = \alpha_r = 0$, dunque anche $L(\mathbf{v}_1), \dots, L(\mathbf{v}_r)$ sono linearmente indipendenti. □

## 5.5 IL TEOREMA DELLA DIMENSIONE

Il Teorema della dimensione è forse il risultato più importante nella teoria delle applicazioni lineari e rappresenta uno strumento formidabile nella risoluzione degli esercizi.

**Teorema della dimensione 5.5.1** *Sia $L : V \longrightarrow W$ un'applicazione lineare. Allora*

$$\dim V = \dim(\text{Ker}(L)) + \dim(\text{Im}(L)) \tag{5.1}$$

**Dimostrazione** – Sia $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ una base per il sottospazio vettoriale $\text{Ker}\, L$. Per il Teorema del completamento (4.2.1) possiamo completare tale insieme di vettori linearmente indipendenti a una base $\mathcal{B}$ di $V$. Sia

$$\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_r, \mathbf{z}_{r+1}, \dots, \mathbf{z}_n\}$$

Se proviamo che $\mathcal{B}_1 = \{L(\mathbf{z}_{r+1}), \dots, L(\mathbf{z}_n)\}$ è una base per $\text{Im}(L)$ il teorema è dimostrato, in quanto $\dim(\text{Ker}(L)) = r$, $\dim(V) = n$ e $\dim(\text{Im}(L)) = n - r$ (la dimensione di $\text{Im}(L)$ è il numero di vettori in una base e $\mathcal{B}_1$ contiene $n - r$ vettori).

Utilizzando la Proposizione 5.4.4 e il fatto che $\mathbf{u}_1, \dots, \mathbf{u}_r$ appartengono al nucleo di $L$ possiamo subito dedurre che $\mathcal{B}_1$ è un sistema di generatori per $\text{Im}(L)$. Ora mostriamo che i vettori di $\mathcal{B}_1$ sono linearmente indipendenti. Sia

$$\alpha_{r+1} L(\mathbf{z}_{r+1}) + \dots + \alpha_n L(\mathbf{z}_n) = \mathbf{0}$$

con $\alpha_{r+1}, \dots, \alpha_n \in \mathbb{R}$. Vogliamo dimostrare che $\alpha_{r+1} = \dots = \alpha_n = 0$.

Abbiamo:

$$\mathbf{0} = \alpha_{r+1} L(\mathbf{z}_{r+1}) + \dots + \alpha_n L(\mathbf{z}_n) = L(\alpha_{r+1}\mathbf{z}_1 + \dots + \alpha_n\mathbf{z}_n)$$

e dunque $\mathbf{z} = \alpha_{r+1}\mathbf{z}_1 + \cdots + \alpha_n\mathbf{z}_n$ appartiene al nucleo di $L$. Poiché $\mathrm{Ker}(L) = \langle \mathbf{u}_1, \ldots, \mathbf{u}_r \rangle$, possiamo scrivere $\mathbf{z}$ nella forma $\mathbf{z} = \alpha_1\mathbf{u}_1 + \cdots + \alpha_r\mathbf{u}_r$, con $\alpha_1, \ldots, \alpha_r \in \mathbb{R}$. Quindi

$$\alpha_{r+1}\mathbf{z}_1 + \cdots + \alpha_n\mathbf{z}_n = \alpha_1\mathbf{u}_1 + \cdots + \alpha_r\mathbf{u}_r$$

da cui segue che

$$\alpha_{r+1}\mathbf{z}_1 + \cdots + \alpha_n\mathbf{z}_n - (\alpha_1\mathbf{u}_1 + \cdots + \alpha_r\mathbf{u}_r) = \mathbf{0}$$

e, essendo $\mathcal{B}$ una base per $V$, questo implica che $\alpha_1 = \ldots = \alpha_n = 0$, concludendo la dimostrazione del teorema.

□

La formula (5.1) pone delle restrizioni sulla tipologia e sull'esistenza di applicazioni lineari tra due spazi vettoriali dati.

**Proposizione 5.5.2** *Siano $V$ e $W$ due spazi vettoriali.*

1) *Se* $\dim V > \dim W$, *non esistono applicazioni lineari iniettive da $V$ in $W$.*

2) *Se* $\dim V < \dim W$, *non esistono applicazioni lineari suriettive da $V$ in $W$.*

**Dimostrazione** – È una semplice applicazione del Teorema 5.5.1.
1) Se $L : V \to W$ è un'applicazione lineare iniettiva, allora $\dim(\mathrm{Ker}\, L) = 0$ e dunque $\dim V = \dim(\mathrm{Im}\, L)$. Essendo $\mathrm{Im}\, L$ un sottospazio di $W$ risulta $\dim V \leq \dim W$.

2) Analogamente, se $L : V \to W$ è suriettiva, allora $\mathrm{Im}\, L = W$ e pertanto $\dim W = \dim V - \dim(\mathrm{Ker}\, L)$, quindi $\dim W \leq \dim V$.

□

**Esempio 5.5.3**

Si consideri l'applicazione lineare $F : \mathbb{R}^4 \longrightarrow \mathbb{R}^2$ definita da: $F(\mathbf{e}_1) = \mathbf{e}_1 - \mathbf{e}_2$, $F(\mathbf{e}_2) = 3\mathbf{e}_1 - 4\mathbf{e}_2$, $F(\mathbf{e}_3) = -\mathbf{e}_1 - 5\mathbf{e}_2$, $F(\mathbf{e}_4) = 3\mathbf{e}_1 + \mathbf{e}_2$. Ci chiediamo se l'applicazione sia iniettiva, suriettiva, biunivoca. Per il teorema precedente, senza bisogno di fare alcun calcolo, abbiamo che l'applicazione non può essere iniettiva in quanto la dimensione di $\mathbb{R}^4$ è maggiore della dimensione di $\mathbb{R}^2$. Veniamo ora alla suriettività. L'immagine di $F$ è generata dai vettori: $\mathbf{e}_1 - \mathbf{e}_2$, $3\mathbf{e}_1 - 4\mathbf{e}_2$, $-\mathbf{e}_1 - 5\mathbf{e}_2$, $3\mathbf{e}_1 + \mathbf{e}_2$. Poiché almeno due sono linearmente indipendenti, essi formano una base di $\mathbb{R}^2$ e dunque l'applicazione è suriettiva. Poiché non è iniettiva, $F$ non è biunivoca.

## 5.6 ISOMORFISMO DI SPAZI VETTORIALI

Il concetto di isomorfismo ci permette di identificare due spazi vettoriali e quindi di trattarli nello stesso modo quando dobbiamo risolvere problemi di algebra lineare, quali, per esempio, determinare se i vettori di un insieme siano linearmente indipendenti, oppure calcolare il nucleo e l'immagine di un'applicazione lineare.

**Definizione 5.6.1** Un'applicazione lineare $L : V \longrightarrow W$ si dice *isomorfismo* se è invertibile, o equivalentemente se è iniettiva e suriettiva.

Analogamente due spazi vettoriali $V$ e $W$ si dicono *isomorfi* se esiste un isomorfismo $L : V \longrightarrow W$; in tal caso e si scrive $V \cong W$.

**Esempio 5.6.2**

Consideriamo l'applicazione lineare $L : \mathbb{R}_2[x] \longrightarrow \mathbb{R}^3$ definita da: $L(x^2) = (1,0,0)$, $L(x) = (0,1,0)$, $L(1) = (0,0,1)$. Questa applicazione è invertibile. Per vederlo possiamo determinare il nucleo e verificare che è il sottospazio nullo, determinare l'immagine e verificare che coincide con $\mathbb{R}^3$. Lasciamo questa verifica per esercizio. In alternativa possiamo definire l'applicazione lineare $T : \mathbb{R}^3 \longrightarrow \mathbb{R}_2[x]$ tale che $T(e_1) = x^2$, $T(e_2) = x$, $T(e_3) = 1$ e verificare che è l'inversa di $L$. Lasciamo queste verifiche per esercizio. Dunque $\mathbb{R}_2[x]$ e $\mathbb{R}^3$ sono isomorfi. In qualche modo è come se fossero lo stesso spazio, in quanto abbiamo realizzato una *corrispondenza biunivoca* che a un vettore di $\mathbb{R}_2[x]$ fa corrispondere uno e un solo vettore di $\mathbb{R}^3$, e viceversa. Questa corrispondenza conserva inoltre le operazioni di somma di vettori e moltiplicazione di un vettore per uno scalare. Avevamo infatti già notato che, una volta fissata una base in $\mathbb{R}_2[x]$, ogni vettore si scrive usando 3 coordinate, proprio come un vettore in $\mathbb{R}^3$. Se fissiamo, per esempio, la base canonica $\{x^2, x, 1\}$, l'applicazione lineare che a ogni polinomio associa le sue coordinate è proprio l'isomorfismo $L : \mathbb{R}_2[x] \longrightarrow \mathbb{R}^3$ descritto sopra, e, una volta scritte le coordinate di un polinomio, a tutti gli effetti possiamo trattarlo come un elemento di $\mathbb{R}^3$. Per esempio, per stabilire se alcuni polinomi sono linearmente indipendenti o determinare una base del sottospazio vettoriale che essi generano, possiamo passare alle loro coordinate e utilizzare l'algoritmo di Gauss come descritto nel Capitolo 1.

Il prossimo teorema è particolarmente importante, in quanto ci dice che non solo $\mathbb{R}_2[x]$, ma ogni spazio vettoriale di dimensione finita è isomorfo a $\mathbb{R}^N$ per un certo $N$ (che naturalmente dipende dallo spazio vettoriale in questione). Quindi i metodi di calcolo che abbiamo descritto per risolvere vari problemi nello spazio vettoriale $\mathbb{R}^N$ si possono applicare a ogni spazio vettoriale $V$ anche molto strano (purché di dimensione finita), utilizzando al posto delle $N$-uple di numeri reali le coordinate dei vettori dello spazio $V$ rispetto a una base fissata. Vedremo nel corso di questo capitolo esempi di tutto ciò.

**Teorema 5.6.3** *Due spazi vettoriali $V$ e $W$ sono isomorfi se e solo se hanno la stessa dimensione.*

**Dimostrazione** – Supponiamo che $\dim V = \dim W = n$ e consideriamo due basi $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ e $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ di $V$ e $W$ rispettivamente. Allora l'applicazione lineare $L : V \to W$ definita da $L(\mathbf{v}_i) = \mathbf{w}_i$ per ogni $i = 1, \dots, n$ è un isomorfismo. Infatti per la Proposizione 5.4.4 si ha che $\operatorname{Im} L = \langle \mathbf{w}_1, \dots, \mathbf{w}_n \rangle$, e, poiché i vettori $\mathbf{w}_1, \dots, \mathbf{w}_n$ generano $W$, si ha che $L$ è suriettiva. Allora per il Teorema della dimensione (5.5.1) si ha che $\dim(\operatorname{Ker} L) = \dim V - \dim(\operatorname{Im} L) = \dim V - \dim W = 0$, quindi $\operatorname{Ker} L$ è il sottospazio nullo e per la Proposizione 5.4.8 si ha che $L$ è iniettiva.

Viceversa, se due spazi vettoriali sono isomorfi tra loro, allora esiste un'applicazione lineare $L : V \to W$ che è sia iniettiva sia suriettiva. Allora per la Proposizione 5.4.8 si ha che $\dim(\operatorname{Ker} L) = 0$ e $\dim(\operatorname{Im} L) = \dim W$; applicando il Teorema della dimensione (5.5.1), si ottiene che $\dim V = \dim(\operatorname{Ker} L) + \dim(\operatorname{Im} L) = \dim W$.

□

Poiché conosciamo la dimensione degli spazi vettoriali $\mathrm{M}_{m,n}(\mathbb{R})$ e $\mathbb{R}_d[x]$, abbiamo immediatamente il seguente corollario.

**Corollario 5.6.4** $\mathrm{M}_{m,n}(\mathbb{R}) \cong \mathbb{R}^{mn}$, $\mathbb{R}_d[x] \cong \mathbb{R}^{d+1}$.

## 5.7 CALCOLO DEL NUCLEO E DELL'IMMAGINE

Questo paragrafo è estremamente importante per le applicazioni in quanto ci fornisce dei metodi pratici per il calcolo delle basi di nucleo e immagine di un'applicazione lineare data.

Iniziamo con il *calcolo di una base* del nucleo di un'applicazione lineare.

Supponiamo di avere un'applicazione lineare $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ e di voler determinare una base per il nucleo. Fissiamo in $\mathbb{R}^n$ e $\mathbb{R}^m$ le basi canoniche, allora per la Proposizione 5.2.2 si ha che $F(\mathbf{x}) = A\mathbf{x}$ per un'opportuna matrice $A \in \mathrm{M}_{m,n}(\mathbb{R})$. Per definizione di nucleo abbiamo:

$$\operatorname{Ker} F = \{\mathbf{x} \in \mathbb{R}^n | A\mathbf{x} = \mathbf{0}\}$$

cioè *il nucleo di $F$ è l'insieme delle soluzioni del sistema lineare omogeneo associato ad $A$.*

Vediamo un esempio concreto.

---

**Esempio 5.7.1**

Si consideri l'applicazione lineare $F : \mathbb{R}^4 \longrightarrow \mathbb{R}^2$ definita da: $F(\mathbf{e}_1) = -\mathbf{e}_2$, $F(\mathbf{e}_2) = 3\mathbf{e}_1 - 4\mathbf{e}_2$, $F(\mathbf{e}_3) = -\mathbf{e}_1$, $F(\mathbf{e}_4) = 3\mathbf{e}_1 + \mathbf{e}_2$. Vogliamo determinare una base per il nucleo di $F$. Scriviamo la matrice $A$ associata a $F$ rispetto alle basi canoniche:

$$A = \begin{pmatrix} 0 & 3 & -1 & 3 \\ -1 & -4 & 0 & 1 \end{pmatrix}$$

Pertanto $F(x_1, x_2, x_3, x_4) = (3x_2 - x_3 + 3x_4, -x_1 - 4x_2 + x_4)$ e $\operatorname{Ker} F$ è l'insieme delle soluzioni del sistema lineare omogeneo:

$$\begin{cases} 3x_2 - x_3 + 3x_4 = 0 \\ -x_1 - 4x_2 + x_4 = 0 \end{cases}$$

che è appunto associato alla matrice $A$.

Per ridurre $A$ a scala è sufficiente scambiare le sue due righe e si ottiene:

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

Risolvendo il sistema si ha quindi che

$$\begin{aligned} \operatorname{Ker} F &= \left\{\left(-\frac{4}{3}x_3 + 5x_4, \frac{1}{3}x_3 - x_4, x_3, x_4\right) | x_3, x_4 \in \mathbb{R}\right\} \\ &= \left\{\left(-\frac{4}{3}x_3, \frac{1}{3}x_3, x_3, 0\right) + (5x_4, -x_4, 0, x_4) | x_3, x_4 \in \mathbb{R}\right\} \\ &= \left\{x_3\left(-\frac{4}{3}, \frac{1}{3}, 1, 0\right) + x_4\,(5, -1, 0, 1)\,| x_3, x_4 \in \mathbb{R}\right\} \\ &= \left\langle\left(-\frac{4}{3}, \frac{1}{3}, 1, 0\right), (5, -1, 0, 1)\right\rangle \end{aligned}$$

Osserviamo che i vettori $\left(-\frac{4}{3}, \frac{1}{3}, 1, 0\right), (5, -1, 0, 1)$ non solo generano $\operatorname{Ker} F$, ma sono anche linearmente indipendenti, perché non sono l'uno multiplo dell'altro, quindi sono una base di $\operatorname{Ker} F$.

Un altro modo di leggere le uguaglianze scritte sopra è il seguente: $\operatorname{Ker} F$ è l'insieme delle combinazioni lineari dei vettori $\left(-\frac{4}{3}, \frac{1}{3}, 1, 0\right)$, $(5, -1, 0, 1)$, ottenuti ponendo rispettivamente $x_3 = 1$, $x_4 = 0$ e poi $x_3 = 0, x_4 = 1$.

♦ **Osservazione 5.7.2** Il fenomeno descritto nell'esempio precedente si verifica in generale. Se l'insieme delle soluzioni di un sistema lineare omogeneo dipende da $k$ variabili libere, allora tale insieme è lo spazio vettoriale generato dai $k$ vettori $\mathbf{v}_1, \dots, \mathbf{v}_k$, ove $\mathbf{v}_i$ si ottiene ponendo la $i$-esima variabile libera uguale a 1 e le altre variabili libere uguali a 0, per ogni $i = 1, \dots, k$. Inoltre i $k$ vettori così ottenuti sono linearmente indipendenti, e sono quindi una base di tale spazio vettoriale.

Formalizziamo quanto appena osservato. È necessaria prima una definizione:

**Definizione 5.7.3** Si chiama *rango righe* di una matrice $M \in \mathrm{M}_{m,n}(\mathbb{R})$ il massimo numero di righe linearmente indipendenti di $M$, cioè la dimensione del sottospazio di $\mathbb{R}^n$ generato dalle righe di $M$. Il rango righe di $A$ si indica con $\operatorname{rr}(A)$.

Se $A \in \mathrm{M}_{m,n}(\mathbb{R})$ è una matrice in forma a scala, la definizione di rango righe appena data coincide con la Definizione 1.3.4. Infatti, per la Proposizione 4.3.3 le righe non nulle di una matrice a scala $A$ sono linearmente indipendenti, quindi la dimensione del sottospazio generato dalle righe di $A$ coincide con il numero di righe non nulle di $A$.

**Proposizione 5.7.4** *Sia $A\mathbf{x} = \mathbf{0}$ un sistema lineare omogeneo, ove $A \in \mathrm{M}_{m,n}(\mathbb{R})$ e sia $W$ l'insieme delle sue soluzioni. Allora $W$ è uno spazio vettoriale di dimensione $n - \operatorname{rr}(A)$.*

**Dimostrazione** – Sia $L_A$ l'applicazione lineare $L_A : \mathbb{R}^n \to \mathbb{R}^m$ associata alla matrice $A$, rispetto alle basi canoniche del dominio e del codominio. Allora $W$ è proprio il nucleo di $L_A$ e quindi è un sottospazio vettoriale di $\mathbb{R}^n$ per la Proposizione 5.4.3. Per determinare $W$, possiamo utilizzare l'algoritmo di Gauss e ridurre a scala la matrice $(A|\mathbf{0})$, ottenendo una matrice $(A'|\mathbf{0})$. Ora per la Proposizione 4.3.1 abbiamo che $\operatorname{rr}(A) = \operatorname{rr}(A') = r$ e, a sua volta, $r$ è uguale al numero di righe non nulle di $A'$, cioè al numero di pivot di $A'$. Allora, per quanto abbiamo visto nel Capitolo 1, possiamo assegnare un valore arbitario a $n - r$ variabili, e ricavare in funzione di esse le $r$ variabili corrispondenti ai pivot. Siano $x_{i_1}, \dots, x_{i_{n-r}}$ le variabili libere e sia $\mathbf{w}_j$ la soluzione del sistema ottenuta ponendo $x_{i_j} = 1$ e le altri variabili libere uguali a zero. Procedendo esattamente come nell'Esempio 5.7.1 si ottiene che i vettori $\mathbf{w}_1, \dots, \mathbf{w}_{n-r}$ generano $W$. Mostriamo ora che sono anche linearmente indipendenti, da cui segue che sono una base di $W$ e quindi $W$ ha dimensione $n - r$. Supponiamo che sia $\mathbf{w} = \lambda_1 \mathbf{w}_1 + \dots + \lambda_{n-r}\mathbf{w}_{n-r} = \mathbf{0}$, con $\lambda_1, \dots, \lambda_{n-r} \in \mathbb{R}$. Osserviamo che l'elemento di posto $i_j$ di $\mathbf{w}_h$ è 1 se $h = j$, altrimenti è zero, per cui l'elemento di posto $i_j$ di $\mathbf{w}$ è esattamente $\lambda_j$. L'ipotesi che sia $\mathbf{w} = \mathbf{0}$ significa che tutti gli elementi di $\mathbf{w}$ sono nulli, in particolare $\lambda_1 = \cdots = \lambda_{n-r} = 0$ e questo dimostra che i vettori $\mathbf{w}_1, \ldots, \mathbf{w}_{n-r}$ sono linearmente indipendenti.

□

Vogliamo ora procedere con il *calcolo di una base dell'immagine* di un'applicazione lineare.

Supponiamo di avere un'applicazione lineare $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ e di voler determinare una base per l'immagine. Fissiamo in $\mathbb{R}^n$ e $\mathbb{R}^m$ le basi canoniche, allora per la Proposizione 5.2.2 si ha che $F(\mathbf{x}) = A\mathbf{x}$ per un'opportuna matrice $A \in \mathrm{M}_{m,n}(\mathbb{R})$. Per la Proposizione 5.4.4 abbiamo:

$$\mathrm{Im}\,(F) = \langle F(\mathbf{e}_1), \ldots, F(\mathbf{e}_n) \rangle = \langle A_1, \ldots, A_n \rangle$$

ove $A_1, \ldots, A_n$ sono le colonne di $A$. A questo punto è sufficiente applicare l'algoritmo di Gauss ai vettori che formano le colonne di $A$. Ricordiamo che, per eseguire l'algoritmo di Gauss, è necessario scrivere i vettori come righe. Vediamo un esempio.

**Esempio 5.7.5**

Sia $F : \mathbb{R}^3 \longrightarrow \mathbb{R}^4$ definita da $F(x, y, z) = (x, 2x, x + y + z, y)$.

La matrice associata a $F$ rispetto alle basi canoniche è:

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

Si ha che l'immagine di $F$ è generata dalle colonne di $A$, cioè

$$\mathrm{Im}\, F = \langle (1, 2, 1, 0), (0, 0, 1, 1), (0, 0, 1, 0) \rangle.$$

Quindi applichiamo l'algoritmo di Gauss alla matrice

$$A^{\mathrm{T}} = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

ove $A^{\mathrm{T}}$ denota la *trasposta* della matrice $A$, cioè la matrice che ha per righe le colonne della matrice $A$. Riducendo $A^T$ a scala, si ottiene

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

Abbiamo quindi che una base per l'immagine di $F$ è:

$$\{(1, 2, 1, 0), (0, 0, 1, 1), (0, 0, 0, -1)\}$$

## 5.8 ESERCIZI SVOLTI

**Esercizio 5.8.1**

*Sia $F : \mathbb{R}^4 \to \mathbb{R}^3$ l'applicazione lineare definita da:* $F(\mathbf{e}_1) = \mathbf{e}_1 + 2\mathbf{e}_2 + k\mathbf{e}_3$, $F(\mathbf{e}_2) = k\mathbf{e}_2 + k\mathbf{e}_3$, $F(\mathbf{e}_3) = k\mathbf{e}_1 + k\mathbf{e}_2 + 6\mathbf{e}_3$, $F(\mathbf{e}_4) = k\mathbf{e}_1 + (6 - k)\mathbf{e}_3$.

a) *Si determini per quali valori di $k$ si ha che $F$ è iniettiva e per quali valori di $k$ si ha che $F$ è suriettiva.*

b) *Scelto un valore di $k$ per cui $F$ non è suriettiva, si determini un vettore $\mathbf{v} \in \mathbb{R}^3$ tale che $\mathbf{v} \notin \operatorname{Im} F$.*

**Svolgimento**

Per la Proposizione 5.5.2 $F$ non è mai iniettiva.

Studiamo ora la suriettività. La matrice associata a $F$ rispetto alle basi canoniche nel dominio e nel codominio è:

$$A = \begin{pmatrix} 1 & 0 & k & k \\ 2 & k & k & 0 \\ k & k & 6 & 6-k \end{pmatrix}$$

(vedi l'Osservazione 5.2.4). L'immagine di $F$ è il sottospazio generato dalle colonne di $A$, per cui scriviamo tali colonne in riga (cioè consideriamo la matrice trasposta $A^T$ di $A$) e applichiamo l'algoritmo di Gauss. Si ha che:

$$A^T = \begin{pmatrix} 1 & 2 & k \\ 0 & k & k \\ k & k & 6 \\ k & 0 & 6-k \end{pmatrix}$$

e riducendo a scala si ottiene la matrice:

$$\begin{pmatrix} 1 & 2 & k \\ 0 & k & k \\ 0 & 0 & k^2-k-6 \\ 0 & 0 & 0 \end{pmatrix}$$

Se $k \neq 0, k \neq -2$ e $k \neq 3$ tale matrice ha 3 righe non nulle, quindi $\operatorname{Im} F$ ha dimensione 3 e $F$ è suriettiva.

Se $k = 0$, dopo aver scambiato la seconda e la terza riga otteniamo:

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & -6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

quindi $\operatorname{Im} F$ ha dimensione 2 e $F$ non è suriettiva.

Se $k = -2$ o $k = 3$ si ottiene in entrambi i casi una matrice a scala con 2 righe non nulle, quindi $\operatorname{Im} F$ ha dimensione 2 e $F$ non è suriettiva.

Scelto $k = 0$, si ha che $\operatorname{Im} F = \langle (1,2,0), (0,0,6) \rangle$ e $\mathbf{v} = (0,-1,3) \notin \operatorname{Im} F$, in quanto i 3 vettori $(1,2,0), (0,-1,3), (0,0,6)$ sono linearmente indipendenti perché sono le righe non nulle di una matrice a scala. ∎

### Esercizio 5.8.2

*Si determini, se è possibile, un'applicazione lineare $G : \mathbb{R}^4 \to \mathbb{R}^3$ tale che* $\operatorname{Im} G = \langle (1,1,0), (0,3,-1), (3,0,1) \rangle$ *e* $\operatorname{Ker} G = \langle (-1,0,1,-3) \rangle$.

**Svolgimento**

Vediamo innanzitutto se le richieste fatte sono compatibili con il Teorema della dimensione. Determiniamo una base di $\langle (1,1,0), (0,3,-1), (3,0,1) \rangle$. Per farlo, riduciamo a scala la matrice:

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

e otteniamo
$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 3 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Quindi avremmo che $\operatorname{Im} G$ ha dimensione 2 e $\operatorname{Ker} G$ ha dimensione 1, ma $\dim \mathbb{R}^4 = 4 \neq 1 + 2 = \dim(\operatorname{Ker} G) + \dim(\operatorname{Im} G)$, di conseguenza un'applicazione lineare con le proprietà richieste non può esistere. ∎

### Esercizio 5.8.3

*Si determini, se possibile, un'applicazione lineare $F : \mathbb{R}^3 \to \mathbb{R}^3$ tale che* $\operatorname{Ker} F = \langle \mathbf{e}_1 - 2\mathbf{e}_3 \rangle$ *e* $\operatorname{Im} F = \langle 2\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3, \mathbf{e}_1 - \mathbf{e}_2 - \mathbf{e}_3 \rangle$. *Tale applicazione è unica?*

**Svolgimento**

Osserviamo che le richieste fatte sono compatibili con il Teorema della dimensione ($s \leq 1$). Infatti $\langle \mathbf{e}_1 - 2\mathbf{e}_3 \rangle$ ha dimensione 1, $\langle 2\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3, \mathbf{e}_1 - \mathbf{e}_2 - \mathbf{e}_3 \rangle$ ha dimensione 2 (perché i due vettori non sono l'uno multiplo dell'altro) e $\dim \mathbb{R}^3 = 3 = 1 + 2 = \dim(\operatorname{Ker} F) + \dim(\operatorname{Im} F)$.

Cerchiamo ora di determinare la matrice $A$ associata a $F$ rispetto alla base canonica (nel dominio e nel codominio). Deve succedere che $F(\mathbf{e}_1 - 2\mathbf{e}_3) = \mathbf{0}$, cioè $F(\mathbf{e}_1) - 2F(\mathbf{e}_3) = \mathbf{0}$ (perché $F$ è lineare), quindi $F(\mathbf{e}_1) = 2F(\mathbf{e}_3)$. Poiché $F(\mathbf{e}_1)$ è rappresentata dalla prima colonna di $A$, e $F(\mathbf{e}_3)$ dalla terza colonna, abbiamo che la prima colonna di $A$ deve essere il doppio della terza, inoltre il sottospazio generato dalle colonne di $A$, e cioè $\operatorname{Im} F$, deve essere uguale a $\langle (2, 1, -1), (1, -1, -1) \rangle$. Per esempio, la matrice:

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

soddisfa le richieste date. Infatti per costruzione $\mathbf{e}_1 - 2\mathbf{e}_3 \in \operatorname{Ker} F$ e $\operatorname{Im} F = \langle 2\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3, \mathbf{e}_1 - \mathbf{e}_2 - \mathbf{e}_3 \rangle$. Inoltre dal Teorema della dimensione (5.5.1) otteniamo che $\dim(\operatorname{Ker} F) = \dim \mathbb{R}^3 - \dim(\operatorname{Im} F) = 3 - 2 = 1$, quindi $\operatorname{Ker} F = \langle \mathbf{e}_1 - 2\mathbf{e}_3 \rangle$.

Tale $F$ non è unica, infatti si può verificare che, per esempio, anche la matrice

$$A = \begin{pmatrix} 6 & 2 & 3 \\ -6 & 1 & -3 \\ -6 & -1 & -3 \end{pmatrix}$$

soddisfa le richieste. ∎

### Esercizio 5.8.4

*Sia $G : \mathbb{R}^2 \to \mathbb{R}^3$ l'applicazione lineare definita da:*
$G(x, y) = (kx + 5y, 2x + (k + 3)y, (2k - 2)x + (7 - k)y)$. *Si stabilisca per quali valori di $k$ si ha che $G$ è iniettiva.*

**Svolgimento**

Per stabilire se $G$ è iniettiva dobbiamo capire se il nucleo di $G$ contiene solo il vettore nullo oppure no. Per il Corollario 5.2.3 la matrice associata a $G$ rispetto alle basi canoniche nel dominio e nel codominio è:

$$A = \begin{pmatrix} k & 5 \\ 2 & k + 3 \\ 2k - 2 & 7 - k \end{pmatrix}$$

Per trovare il nucleo di $G$ dobbiamo risolvere il sistema lineare omogeneo associato ad $A$. Riducendo a scala matrice $(A|\underline{0})$ si ottiene:

$$(A'|\underline{0}) = \left(\begin{array}{cc|c} 1 & \frac{k+3}{2} & 0 \\ 0 & -k^2-3k+10 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

Se $k^2+3k-10 \neq 0$, cioè se $k \neq -5$ e $k \neq 2$, abbiamo che $\mathrm{rr}(A'|\underline{0}) = \mathrm{rr}(A') = 2$. Poiché le incognite sono 2, il sistema ha un'unica soluzione, quella nulla, quindi $\operatorname{Ker} G = \{(0,0)\}$ e $G$ è iniettiva. Se $k=-5$ o $k=2$ si vede facilmente che $\mathrm{rr}(A'|\underline{0}) = \mathrm{rr}(A') = 1$, quindi il sistema ammette infinite soluzioni che dipendono da un parametro e $G$ non è iniettiva.

Un modo alternativo di procedere è il seguente. Per il Teorema della dimensione (5.5.1) abbiamo che $2 = \dim \mathbb{R}^2 = \dim(\operatorname{Ker} G) + \dim(\operatorname{Im} G)$, $G$ è iniettiva se e solo se l'immagine di $G$ ha dimensione 2. Calcoliamo quindi una base per l'immagine di $G$. Dobbiamo considerare la matrice

$$\begin{pmatrix} k & 2 & 2k-2 \\ 5 & k+3 & 7-k \end{pmatrix}$$

che ha per righe le colonne di $A$ e ridurla a scala con l'algoritmo di Gauss. Si ottiene:

$$\begin{pmatrix} 1 & \frac{k+3}{5} & \frac{7-k}{5} \\ 0 & -k^2-3k+10 & k^2+3k-10 \end{pmatrix}$$

Se $k \neq -5$ e $k \neq 2$, ci sono due righe non nulle, quindi l'immagine di $G$ ha dimensione 2 e $G$ è iniettiva. Se $k=-5$ o $k=2$, c'è solo una riga non nulla, quindi l'immagine di $G$ ha dimensione 1 e $G$ non è iniettiva. Abbiamo così ritrovato il risultato ottenuto con il metodo precedente. ■

## 5.9 ESERCIZI PROPOSTI

### Esercizio 5.9.1

*Sia data l'applicazione $F : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definita da: $F(x,y) = (x+2ky, x-y)$. Si determinino i valori di $k$ per cui tale applicazione è lineare.*

### Esercizio 5.9.2

*Sia data l'applicazione $F : \mathbb{R}_1[x] \longrightarrow \mathbb{R}_2[x]$ definita da: $F(ax+b) = (a-b)x^2 + kb^2x + 2a$. Si determinino i valori di $k$ per cui tale applicazione è lineare.*

### Esercizio 5.9.3

*Date le applicazioni lineari $F : \mathbb{R}^3 \to \mathbb{R}^2$ e $G : \mathbb{R}^2 \to \mathbb{R}^3$ definite da: $F(x,y,z) = (x-y, 2x+y+z)$ e $G(x,y) = (3y, -x, 4x+2y)$, si determinino, se possibile, $F \circ G$ e $G \circ F$.*

### Esercizio 5.9.4

*Date le applicazioni lineari $F : \mathbb{R}^2 \to \mathbb{R}^2$ e $G : \mathbb{R}^2 \to \mathbb{R}$ definite da: $F(\mathbf{e}_1) = -\mathbf{e}_1 - \mathbf{e}_2$, $F(\mathbf{e}_2) = \mathbf{e}_1 + \mathbf{e}_2$, $G(\mathbf{e}_1) = 2$, $G(\mathbf{e}_2) = -1$; si determinino, se possibile, $F \circ G$ e $G \circ F$.*

### Esercizio 5.9.5

*Sia data l'applicazione lineare* $F : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *definita da* $F(\mathbf{e}_1) = 3\mathbf{e}_1 - 3\mathbf{e}_2$, $F(\mathbf{e}_2) = 2\mathbf{e}_1 - 2\mathbf{e}_2$. *Si calcolino una base del nucleo e una base dell'immagine di* $F$.

### Esercizio 5.9.6

*Si stabilisca quali delle seguenti applicazioni lineari sono isomorfismi.*

a) $F : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *definita da* $F(x, y, z) = (x + 2z, y + z, z)$

b) $F : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ *definita da* $F(x, y, z) = (2x - z, x - y + z)$

c) $F : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *definita da* $F(\mathbf{e}_1) = 2\mathbf{e}_1 + \mathbf{e}_2$, $F(\mathbf{e}_2) = 3\mathbf{e}_1 - \mathbf{e}_3$, $F(\mathbf{e}_3) = \mathbf{e}_1 - \mathbf{e}_2 - \mathbf{e}_3$

### Esercizio 5.9.7

*Si trovino una base per il nucleo e una per l'immagine di ciascuna delle seguenti applicazioni lineari. Si stabilisca inoltre se sono iniettive, suriettive e/o biunivoche.*

a) $F : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *definita da* $F(x, y, z) = (x - z, x + 2y - z, x - 4y - z)$

b) $F : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ *definita da* $F(\mathbf{e}_1) = \mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3$, $F(\mathbf{e}_2) = 2\mathbf{e}_1 - 2\mathbf{e}_2 - \mathbf{e}_3$

c) $F : \mathbb{R}^4 \longrightarrow \mathbb{R}^2$ *definita da* $F(x, y, z, t) = (2x - t, 3y - x + 2z - t)$

d) $F : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *associata alla matrice* $A$ *rispetto alla base canonica, ove*

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

e) $F : \mathbb{R}^3 \longrightarrow \mathbb{R}^4$ *definita da* $F(x, y, z) = (x + y - z, z, x + y, z)$

### Esercizio 5.9.8

*Data l'applicazione lineare* $F : \mathbb{R}^4 \to \mathrm{M}_{2,2}(\mathbb{R})$ *definita da:*

$$F(x_1, x_2, x_3, x_4) = \begin{pmatrix} x_1 - x_3 - x_4 & x_1 + 2x_3 \\ x_1 + x_4 & 0 \end{pmatrix}$$

*si trovi una base per* $\operatorname{Ker} F$ *e si determini la dimensione di* $\operatorname{Im} F$.

### Esercizio 5.9.9

*Data l'applicazione lineare* $F : \mathbb{R}^3 \to \mathbb{R}^3$ *definita da:* $F(\mathbf{e}_1) = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$, $F(\mathbf{e}_2) = 2\mathbf{e}_1 + 2\mathbf{e}_2 + 2\mathbf{e}_3$, $F(\mathbf{e}_3) = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$; *si trovi una base per* $\operatorname{Ker} F$ *e si stabilisca se il vettore* $\mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_3$ *appartiene a* $\operatorname{Im} F$.

### Esercizio 5.9.10

*Si determinino, se possibile, un'applicazione lineare suriettiva* $T : \mathbb{R}^3 \to \mathbb{R}^2$ *e un'applicazione lineare iniettiva* $F : \mathbb{R}^4 \to \mathbb{R}^3$.

### Esercizio 5.9.11

*Esistono applicazioni iniettive* $T : \mathbb{R}^3 \to \mathbb{R}^4$? *Se sì, se ne determini una, se no, si motivi la risposta.*

### Esercizio 5.9.12

*Sia* $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *l'applicazione lineare associata alla matrice* $A$ *rispetto alla base canonica, ove*

$$A = \begin{pmatrix} 1 & 4 \\ k & 0 \end{pmatrix}$$

*Si determinino* $\operatorname{Ker}(T)$ *e* $\operatorname{Im}(T)$ *al variare del parametro* $k$.

### Esercizio 5.9.13

*Sia* $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *l'applicazione lineare associata alla matrice* $A$ *rispetto alla base canonica, ove*

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

*Si dica per quali valori del parametro* $k$ *si ha che* $T$ *è un isomorfismo.*

### Esercizio 5.9.14

*Sia* $: \mathbb{R}^3 \to \mathbb{R}^4$ *l'applicazione lineare definita da:* $T(x, y, z) = (x + 2z, y, 2x + 3y + 4z, 3x - y + 6z)$. *Si trovi una base per* $\operatorname{Ker}(T)$ *e una base di* $\operatorname{Im}(T)$ *e si verifichi il Teorema della dimensione (5.5.1).* $T$ *è iniettiva?*

### Esercizio 5.9.15

*Si determini, se possibile, un'applicazione lineare* $F : \mathbb{R}^3 \to \mathbb{R}^2$ *tale che* $\operatorname{Ker} F = \langle \mathbf{e}_1 \rangle$ *e* $\operatorname{Im} F = \langle \mathbf{e}_1 - \mathbf{e}_2 \rangle$.

### Esercizio 5.9.16

*Si determini, se possibile, un'applicazione lineare* $F : \mathbb{R}^2 \to \mathbb{R}^3$ *tale che* $\operatorname{Ker} F = \langle \mathbf{e}_2 \rangle$ *e* $\operatorname{Im} F = \langle \mathbf{e}_1 - \mathbf{e}_2 + 2\mathbf{e}_3 \rangle$.

### Esercizio 5.9.17

*Si determini, se possibile, un'applicazione lineare* $F : \mathbb{R}^2 \to \mathbb{R}^4$ *tale che* $\operatorname{Im} F$ *abbia dimensione* 1.

### Esercizio 5.9.18

*Si determini, se possibile, un'applicazione lineare non nulla* $F : \mathbb{R}^3 \to \mathbb{R}^2$ *tale che* $(1, 1) \notin \operatorname{Im} F$.

### Esercizio 5.9.19

*Sia* $F : \mathbb{R}^3 \to \mathbb{R}^2$ *l'applicazione lineare definita da:* $F(\mathbf{e}_1) = 2\mathbf{e}_1 + k\mathbf{e}_2$, $F(\mathbf{e}_2) = k\mathbf{e}_1 + 2\mathbf{e}_2$, $F(\mathbf{e}_3) = -2\mathbf{e}_1 - k\mathbf{e}_2$. *Si stabilisca per quali valori di* $k$ $F$ *non è suriettiva. Posto* $k = -1$, *si determinino un vettore* $\mathbf{v}_1$ *che appartiene a* $\operatorname{Ker} F$ *e un vettore* $\mathbf{v}_2$ *che non appartiene a* $\operatorname{Ker} F$.

### Esercizio 5.9.20

*Data l'applicazione lineare* $F : \mathbb{R}^3 \to \mathbb{R}^2$ *definita da:* $F(x, y, z) = (kx - 3y + kz, kx + ky - 3z)$, *si determini per quali valori di* $k$ $F$ *è iniettiva e per quali valori di* $k$ $F$ *è suriettiva.*

**Esercizio 5.9.21**

*Data l'applicazione lineare* $F : \mathbb{R}^2 \to \mathbb{R}^3$ *definita da:* $F(\mathbf{e}_1) = k\mathbf{e}_1 - 4\mathbf{e}_2 + k\mathbf{e}_3$, $F(\mathbf{e}_2) = -3\mathbf{e}_1 + k\mathbf{e}_2 + 3\mathbf{e}_3$, *si determini per quali valori di* $k$ $F$ *è iniettiva e per quali valori di* $k$ $F$ *è suriettiva.*

# 6 Sistemi lineari

In questo capitolo vogliamo rivisitare i sistemi lineari e interpretare i risultati già illustrati nel Capitolo 1 in termini di applicazioni lineari, sfruttando le conoscenze che abbiamo acquisito nel Capitolo 5. Naturalmente utilizzeremo le notazioni e la terminologia introdotte nel Capitolo 1.

## 6.1 CONTROIMMAGINE

La controimmagine o preimmagine di un vettore $\mathbf{w} \in W$ mediante un'applicazione lineare $f : V \longrightarrow W$ è costituita dai vettori dello spazio vettoriale $V$ che hanno per immagine, mediante $f$, il vettore $\mathbf{w}$. Si tratta di un concetto elementare in matematica, per noi sarà uno strumento utile per esprimere in modo conciso le soluzioni di un sistema lineare.

Conosciamo già un esempio di controimmagine, per noi molto importante: il nucleo di un'applicazione lineare $F$. Infatti $\mathrm{Ker}\,(F)$ è la controimmagine del vettore nullo di $W$, cioè consiste dei vettori di $V$ la cui immagine mediante $F$ è $\mathbf{0}_W$.

Vediamo ora la definizione precisa.

**Definizione 6.1.1** Sia $F : V \to W$ un'applicazione lineare e sia $\mathbf{w} \in W$. Si chiama *controimmagine* o *preimmagine* di $\mathbf{w}$ mediante $F$ l'insieme

$$F^{-1}(\mathbf{w}) = \{\mathbf{v} \in V \mid F(\mathbf{v}) = \mathbf{w}\}$$

Si osservi che la notazione $F^{-1}(\mathbf{w})$ introdotta nella precedente definizione *non* ha nulla a che fare con l'invertibilità della funzione. Parlando di controimmagine di un vettore mediante una funzione $F$ *non* stiamo affermando che $F$ è una funzione invertibile: la notazione $F^{-1}(\mathbf{w})$ indica semplicemente un sottoinsieme del dominio della funzione.

---

**Esempio 6.1.2**

Sia $F : \mathbb{R}^2 \to \mathbb{R}^3$ l'applicazione lineare definita da:

$$F(x, y) = (x + y, x + y, x)$$

Qual è la controimmagine del vettore $(1,1,3)$ mediante $F$? Per definizione abbiamo:

$$\begin{aligned} F^{-1}(1,1,3) &= \{(x,y)\in\mathbb{R}^2 \mid f(x,y)=(1,1,3)\} \\ &= \{(x,y)\in\mathbb{R}^2 \mid (x+y,x+y,x)=(1,1,3)\} \\ &= \left\{(x,y)\in\mathbb{R}^2 \mid \begin{cases} x+y=1 \\ x+y=1 \\ x=3 \end{cases}\right\} = \{(3,-2)\} \end{aligned}$$

Questo significa che $F(3,-2)=(1,1,3)$, e non ci sono altri elementi di $\mathbb{R}^2$ che hanno $(1,1,3)$ come immagine. In particolare si ha che $(1,1,3)\in \operatorname{Im} F$.

Calcoliamo ora la controimmagine del vettore $(1,0,0)$ mediante $F$. Per definizione abbiamo:

$$\begin{aligned} F^{-1}(1,0,0) &= \{(x,y)\in\mathbb{R}^2 | F(x,y)=(1,0,0)\} \\ &= \{(x,y)\in\mathbb{R}^2 | (x+y,x+y,x)=(1,0,0)\} \\ &= \left\{(x,y)\in\mathbb{R}^2 \mid \begin{cases} x+y=1 \\ x+y=0 \\ x=0 \end{cases}\right\} = \varnothing \end{aligned}$$

Il vettore $(1,0,0)$ non è immagine di alcun vettore di $\mathbb{R}^2$, cioè $(1,0,0)\notin \operatorname{Im} F$.

---

Gli esempi appena illustrati mostrano che calcolare la controimmagine di un vettore mediante un'applicazione lineare equivale a risolvere un sistema lineare. Approfondiremo tra poco questo punto di vista.

◆ **Osservazione 6.1.3** L'insieme $F^{-1}(\mathbf{w})$ è un sottospazio vettoriale di $V$ se e solo se $F^{-1}(\mathbf{w}) = \operatorname{Ker} f$ e in tal caso $\mathbf{w} = \mathbf{0}_W$.

Infatti se $F^{-1}(\mathbf{w}) = \operatorname{Ker} f$ abbiamo che la controimmagine è un sottospazio vettoriale. Viceversa, ragionando per assurdo, se $\mathbf{w} \neq \mathbf{0}_W$ allora l'insieme $F^{-1}(\mathbf{w})$ non contiene $\mathbf{0}_V$. Infatti, essendo $F$ un'applicazione lineare, si ha $F(\mathbf{0}_V) = \mathbf{0}_W$, pertanto $\mathbf{0}_V$ appartiene soltanto alla controimmagine di $\mathbf{0}_W$ (uno stesso vettore in $V$ non può appartenere alla controimmagine di vettori diversi di $W$). Quindi $F^{-1}(\mathbf{w})$ non può essere un un sottospazio vettoriale di $V$.

**Proposizione 6.1.4** *Sia $F : V \to W$ un'applicazione lineare e sia $\mathbf{w} \in W$. Allora $F^{-1}(\mathbf{w})$ è non vuoto se e solo se $\mathbf{w} \in \operatorname{Im} F$. In tal caso si ha:*

$$F^{-1}(\mathbf{w}) = \{\mathbf{v} + \mathbf{z} \mid \mathbf{z} \in \operatorname{Ker} F\} \tag{6.1}$$

*ove $\mathbf{v}$ è un qualsiasi elemento di $V$ tale che $F(\mathbf{v}) = \mathbf{w}$.*

**Dimostrazione** – Se $\mathbf{w} \notin \operatorname{Im} F$, allora, per definizione, $F^{-1}(\mathbf{w}) = \varnothing$. Vediamo ora il caso $\mathbf{w} \in \operatorname{Im} F$. Per definizione di immagine, esiste un elemento $\mathbf{v} \in V$ tale che $F(\mathbf{v}) = \mathbf{w}$. Non è detto, tuttavia, che $\mathbf{v}$ sia l'unico elemento in $F^{-1}(\mathbf{w})$. Supponiamo che $\mathbf{v}'$ sia un altro elemento in $F^{-1}(\mathbf{w})$. Allora

$$F(\mathbf{v}') = F(\mathbf{v}) = \mathbf{w}$$

quindi

$$F(\mathbf{v}' - \mathbf{v}) = \mathbf{0}_W$$

cioè:

$$\mathbf{v}' - \mathbf{v} \in \operatorname{Ker} F$$

Quindi un qualsiasi elemento $\mathbf{v}'$ di $F^{-1}(\mathbf{w})$ si puó scrivere come $\mathbf{v}' = \mathbf{v} + (\mathbf{v}' - \mathbf{v}) = \mathbf{v} + \mathbf{z}$ con $\mathbf{z} \in \operatorname{Ker} F$. Abbiamo così dimostrato un'inclusione.

Consideriamo ora $\mathbf{v} + \mathbf{z}$, con $\mathbf{z} \in \operatorname{Ker} F$. Allora

$$F(\mathbf{v} + \mathbf{z}) = F(\mathbf{v}) + F(\mathbf{z}) = \mathbf{w} + \mathbf{0}_W = \mathbf{w}$$

Quindi vale anche l'altra inclusione e abbiamo dimostrato il risultato. □

## 6.2 SISTEMI LINEARI: LA TEORIA

Nel Capitolo 5 abbiamo definito il rango righe di una matrice ( *vedi* Definizione 5.7.3). Vogliamo ora approfondire questa nozione, per avere una visione più chiara del legame tra matrici, applicazioni lineari e sistemi lineari.

Data una matrice $A \in \mathrm{M}_{m,n}(\mathbb{R})$, possiamo leggere le sue righe come vettori di $\mathbb{R}^n$ e le sue colonne come vettori di $\mathbb{R}^m$. Appare dunque naturale introdurre la seguente definizione.

**Definizione 6.2.1** Si chiama *rango colonne* di una matrice $A \in \mathrm{M}_{m,n}(\mathbb{R})$ il massimo numero di colonne linearmente indipendenti di $A$, cioè la dimensione del sottospazio di $\mathbb{R}^m$ generato dalle colonne di $A$.

La seguente osservazione ci è già nota e tuttavia, data la sua importanza nel contesto che stiamo esaminando, vogliamo puntualizzarla.

**Osservazione 6.2.2** Se leggiamo $A$ come la matrice associata all'applicazione lineare $L_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ rispetto alle basi canoniche, allora il rango colonne di $A$ è la dimensione dell'immagine di $L_A$. Infatti tale immagine è generata dalle colonne della matrice $A$.

Benché in generale i vettori riga e i vettori colonna di una matrice $A \in \mathrm{M}_{m,n}(\mathbb{R})$ siano elementi di spazi vettoriali diversi, il rango righe e il rango colonne di $A$ *coincidono sempre*. Tale numero si chiama semplicemente *rango* di $A$ e si indica con $\mathrm{rk}(A)$.

**Proposizione 6.2.3** *Se $A \in \mathrm{M}_{m,n}(\mathbb{R})$, allora il rango righe di $A$ è uguale al rango colonne di $A$.*

**Dimostrazione** – Sia $L_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ l'applicazione lineare associata alla matrice $A$ rispetto alle basi canoniche. Il nucleo di $L_A$ è l'insieme delle soluzioni del sistema lineare omogeneo associato alla matrice $A$, e per la Proposizione 5.7.4 ha dimensione $n - \mathrm{rr}(A)$, ove $\mathrm{rr}(A)$ è il rango righe di $A$. Per il Teorema della dimensione (5.5.1) sappiamo anche che la dimensione di $\operatorname{Ker} L_A$ è uguale a $n - \dim(\mathrm{Im} L_A)$. Segue che $\mathrm{rr}(A) = \dim(\mathrm{Im} L_A)$, cioè, $\mathrm{rr}(A)$ è uguale al rango colonne di $A$. □

Vediamo un esempio per chiarire.

**Esempio 6.2.4**

Consideriamo la matrice

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

Il rango righe di $A$ è 1 dal momento che $A$ ha due righe uguali non nulle. Anche il rango colonne è 1 dal momento che i vettori $(1,1)$, $(2,2)$, $(-1,-1)$ di $\mathbb{R}^2$ sono linearmente dipendenti. Dunque $\mathrm{rk}(A) = 1$.

◆ **Osservazione 6.2.5** Se $A \in \mathrm{M}_{m,n}(\mathbb{R})$ è una matrice in forma a scala, la definizione di rango appena data coincide con la Definizione 1.3.4. Infatti, per la Proposizione 4.3.3 le righe non nulle di una matrice a scala $A$ sono linearmente indipendenti, quindi la dimensione del sottospazio generato dalle righe di $A$ coincide con il numero di righe non nulle di $A$. Calcolare il rango di una matrice in forma a scala è dunque immediato mentre calcolare il rango di una matrice qualsiasi richiede in generale piú tempo.

La Proposizione 4.3.1 fornisce un metodo efficace per calcolare il rango di una matrice qualsiasi: poiché le operazioni elementari sulle righe di una matrice ne preservano il rango, e poiché il sottospazio generato dalle righe resta invariato, per calcolare il rango di una matrice $A$, si riduce prima $A$ in forma a scala con l'algoritmo di Gauss e poi si calcola il rango della matrice ridotta, il che si riduce semplicemente a contare il numero di righe non nulle.

**Esempio 6.2.6**

Calcolare il rango della matrice

$$A = \begin{pmatrix} 0 & 1 & 3 \\ 1 & -1 & 5 \\ -1 & 1 & 0 \end{pmatrix}$$

Abbiamo:

$$\mathrm{rk}(A) = \mathrm{rk}\begin{pmatrix} 1 & -1 & 5 \\ 0 & 1 & 3 \\ 0 & 0 & 5 \end{pmatrix} = 3$$

Diamo ora una definizione utile, già anticipata nel primo capitolo.

**Definizione 6.2.7** Sia $A\mathbf{x} = \mathbf{b}$ un sistema lineare. Il sistema lineare $A\mathbf{x} = \mathbf{0}$ è detto il *sistema omogeneo associato* al sistema $A\mathbf{x} = \mathbf{b}$.

**Teorema di struttura dei sistemi lineari 6.2.8** *Sia* $A\mathbf{x} = \mathbf{b}$ *un sistema di* $m$ *equazioni in* $n$ *incognite che ammette almeno una soluzione. Allora l'insieme delle soluzioni del sistema è*

$$S = \{\mathbf{v} + \mathbf{z} \mid \mathbf{z} \in \operatorname{Ker} A\}$$

*ove* $\mathbf{v}$ *è una soluzione particolare del sistema e* $\operatorname{Ker} A$ *rappresenta l'insieme delle soluzioni del sistema lineare omogeneo associato* $A\mathbf{x} = \mathbf{0}$.

**Dimostrazione** – Questo teorema è semplicemente una riscrittura della Proposizione 6.1.4 nel caso in cui $F = L_A$. Infatti, se leggiamo la matrice $A$ come la matrice dell'applicazione lineare $L_A : \mathbb{R}^n \to \mathbb{R}^m$ associata alla matrice $A$ rispetto alle basi canoniche, allora determinare le soluzioni del sistema lineare $A\mathbf{x} = \mathbf{b}$ equivale a determinare i vettori $\mathbf{x} \in \mathbb{R}^n$ tali che $L_A(\mathbf{x}) = \mathbf{b}$, cioè determinare la controimmagine $L_A^{-1}(\mathbf{b})$ del vettore $\mathbf{b} \in \mathbb{R}^m$. Poiché per ipotesi il sistema ammette soluzione, $\mathbf{b}$ appartiene all'immagine di $L_A$, cioè $L_A(\mathbf{v}) = A\mathbf{v} = \mathbf{b}$ per un opportuno vettore $\mathbf{v}$ di $\mathbb{R}^n$. L'insieme $S$ delle soluzioni del sistema è $L_A^{-1}(\mathbf{b})$ che, per la Proposizione 6.1.4, è esattamente $\{\mathbf{v} + \mathbf{z} \mid \mathbf{z} \in \operatorname{Ker} L_A\}$. □

Il teorema che segue è il risultato piú importante nella teoria dei sistemi lineari.

**Teorema di Rouché-Capelli 6.2.9** *Un sistema lineare* $A\mathbf{x} = \mathbf{b}$ *di* $m$ *equazioni in* $n$ *incognite ammette soluzioni se e solo se* $\operatorname{rk}(A) = \operatorname{rk}(A|\mathbf{b})$. *Se tale uguaglianza è soddisfatta, allora il sistema ha:*

1. *una sola soluzione se e solo se* $\operatorname{rk}(A) = \operatorname{rk}(A|\mathbf{b}) = n$;
2. *infinite soluzioni se e solo se* $\operatorname{rk}(A) = \operatorname{rk}(A|\mathbf{b}) < n$. *In tal caso le soluzioni del sistema dipendono da* $n - \operatorname{rk}(A)$ *parametri.*

**Dimostrazione** – 1. Leggiamo la matrice $A$ come la matrice dell'applicazione lineare $L_A : \mathbb{R}^n \to \mathbb{R}^m$ associata ad $A$ rispetto alle basi canoniche. Come prima, determinare le soluzioni del sistema lineare $A\mathbf{x} = \mathbf{b}$, equivale a determinare i vettori $\mathbf{x} \in \mathbb{R}^n$ tali che $L_A(\mathbf{x}) = \mathbf{b}$, cioè determinare la controimmagine $L_A^{-1}(\mathbf{b})$ del vettore $\mathbf{b} \in \mathbb{R}^m$. Per definizione tale controimmagine è non vuota se e solo se $\mathbf{b} \in \operatorname{Im} L_A$. In altre parole il sistema ha soluzioni se e solo se $\mathbf{b} \in \operatorname{Im} L_A$. Poiché $\operatorname{Im} L_A$ è generata dai vettori colonna della matrice $A$, $\mathbf{b} \in \operatorname{Im} L_A$ se e solo se il sottospazio generato dai vettori colonna di $A$ coincide con il sottospazio generato dai vettori colonna di $A$ e dal vettore $\mathbf{b}$ (scritto in colonna), cioè se e solo se

$$\dim\langle \text{colonne di } A\rangle = \dim\langle \text{colonne di } (A|\mathbf{b})\rangle$$

ossia se e solo se $\operatorname{rk}(A) = \operatorname{rk}(A|\mathbf{b})$.

2. Se $A\mathbf{x} = \mathbf{b}$ ammette soluzioni, allora per il Teorema 6.2.8 abbiamo che l'insieme delle sue soluzioni è della forma

$$S = \{\mathbf{v} + \mathbf{z} \mid \mathbf{z} \in \operatorname{Ker} L_A\}$$

essendo $\mathbf{v}$ una soluzione particolare del sistema. Allora si verifica solo uno dei due seguenti casi:

1. $\dim(\operatorname{Ker} A) = 0$, cioè $\operatorname{Ker} A = \{\mathbf{0}_{\mathbb{R}^n}\}$, perciò $S = \{\mathbf{v}\}$, cioè il sistema ammette una sola soluzione;
2. $\dim(\operatorname{Ker} A) > 0$ quindi $\operatorname{Ker} A$ contiene infiniti elementi, essendo un sottospazio vettoriale reale.

Per il Teorema della dimensione, abbiamo $\dim(\operatorname{Ker} A) = n - \dim \operatorname{Im} L_A = n - \operatorname{rk}(A)$ e dunque le soluzioni dipendono da $n - \operatorname{rk}(A)$ parametri. □

**Definizione 6.2.10** Dato un sistema lineare compatibile $A\mathbf{x} = \mathbf{b}$, cioè un sistema che ammette soluzione, il numero $\operatorname{rk}(A) = \operatorname{rk}(A|\mathbf{b})$ si chiama anche *rango del sistema*.

Nel Capitolo 1 abbiamo imparato a risolvere un qualsiasi sistema lineare $A\mathbf{x} = \mathbf{b}$ nel modo seguente:

1. si scrive la matrice completa $(A|\mathbf{b})$ associata al sistema;
2. si utilizza l'algoritmo di Gauss per ridurre $(A|\mathbf{b})$ a una matrice in forma a scala $(A'|\mathbf{b}')$;
3. il sistema lineare di partenza $A\mathbf{x} = \mathbf{b}$ è equivalente al sistema lineare a scala $A'\mathbf{x} = \mathbf{b}'$;
4. il sistema lineare $A'\mathbf{x} = \mathbf{b}'$ ha soluzioni se e solo se $\operatorname{rk}(A') = \operatorname{rk}(A'|\mathbf{b}')$. In tal caso, utilizzando sostituzioni successive dal basso si ottengono tutte le soluzioni del sistema.

Ora il Teorema di Rouché-Capelli (6.2.9) afferma a priori che il sistema $A\mathbf{x} = \mathbf{b}$, che in generale non sarà un sistema a scala, ammette soluzioni se e solo se $\operatorname{rk}(A) = \operatorname{rk}(A|\mathbf{b})$. Come si concilia il Teorema di Rouché-Capelli con il metodo di risoluzione dei sistemi lineari appena ricordato? Tutto funziona poiché, come dimostrato nella Proposizione 4.3.1, l'algoritmo di Gauss preserva il rango di una matrice, quindi $\operatorname{rk}(A) = \operatorname{rk}(A')$ e $\operatorname{rk}(A|\mathbf{b}) = \operatorname{rk}(A'|\mathbf{b}')$, perciò $\operatorname{rk}(A) = \operatorname{rk}(A|\mathbf{b})$ se e solo se $\operatorname{rk}(A') = \operatorname{rk}(A'|\mathbf{b}')$.

Il Teorema di Rouché-Capelli fornisce tutte le informazioni sui sistemi lineari che abbiamo già illustrato nel Capitolo 1. In particolare afferma che: *un sistema lineare a coefficienti reali compatibile possiede una soluzione o infinite.*

Questa è esattamente la situazione che avevamo descritto nel Capitolo 1, riletta in termini di applicazioni lineari. Nella sostanza, un sistema lineare compatibile in $n$ incognite reali è un insieme di condizioni compatibili che vengono assegnate sulle $n$ variabili reali prese in considerazione. Tali condizioni abbassano quindi il numero di gradi di libertà del sistema: se il rango del sistema è $k$, l'insieme delle soluzioni non dipende più da $n$ variabili libere ma da $n-k$. Quello che conta è il rango del sistema e non il numero delle equazioni perché il rango del sistema quantifica le condizioni "indipendenti" ed elimina le condizioni che possono essere dedotte dalle altre e che quindi sono superflue.

Vediamo qualche esempio per illustrare i risultati dimostrati sopra.

---

**Esempio 6.2.11**

Consideriamo l'applicazione lineare $L_A : \mathbb{R}^3 \to \mathbb{R}^3$ la cui matrice rispetto alla base canonica (sia nel dominio sia nel codominio) è:

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

Vogliamo stabilire se il vettore $\mathbf{b} = (3, 0, 3)$ appartiene a $\mathrm{Im}\,(L_A)$ e, in caso affermativo, calcolare la controimmagine di $\mathbf{b}$ mediante $L_A$.

Osserviamo che $\mathrm{Im}\,(L_A)$ è generata dalle colonne della matrice $A$, perciò $\mathbf{b}$ appartiene a $\mathrm{Im}\,(L_A)$ se e solo se $\mathrm{rk}(A) = \mathrm{rk}(A|\mathbf{b})$. Inoltre la controimmagine di $\mathbf{b}$ mediante $L_A$ è costituita dai vettori $(x, y, z) \in \mathbb{R}^3$ tali che $L_A(x, y, z) = (3, 0, 3)$, cioè dai vettori $(x, y, z)$ tali che

$$A\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 3 \end{pmatrix}$$

ossia tali che:

$$\begin{pmatrix} x + 2z \\ 2x + y + z \\ 3x + y + 3z \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 3 \end{pmatrix}$$

In altre parole calcolare la controimmagine di $\mathbf{b}$ mediante $L_A$ significa risolvere il sistema lineare

$$\begin{cases} x + 2z = 3 \\ 2x + y + z = 0 \\ 3x + y + 3z = 3 \end{cases}$$

Per calcolare il rango della matrice $(A|\mathbf{b})$ e confrontarlo con il rango di $A$, ridurremo la matrice $(A|\mathbf{b})$ e, simultaneamente, la matrice $A$ in forma a scala utilizzando l'algoritmo di Gauss e calcoleremo il rango delle matrici ridotte. Abbiamo:

$$(A|\mathbf{b}) \to \left(\begin{array}{ccc|c} 1 & 0 & 2 & 3 \\ 2 & 1 & 1 & 0 \\ 3 & 1 & 3 & 3 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & 2 & 3 \\ 0 & 1 & -3 & -6 \\ 0 & 1 & -3 & -6 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & 2 & 3 \\ 0 & 1 & -3 & -6 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Dunque $\mathrm{rk}(A) = \mathrm{rk}(A|\mathbf{b}) = 2$. Questo significa che il vettore $\mathbf{b}$ appartiene all'immagine di $L_A$ e che $\dim(\mathrm{Ker}\, L_A) = 3 - 2 = 1$. La controimmagine di $\mathbf{b}$ è data dagli elementi della forma: $\mathbf{v}+\mathbf{z}$, essendo $\mathbf{v}$ una soluzione particolare del sistema $A\mathbf{x} = \mathbf{b}$ ed essendo $\mathbf{z} \in \mathrm{Ker}\, A$, ove $A$ è il nucleo della matrice $A$, cioè l'insieme delle soluzioni del sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$. Per calcolare $\mathbf{v}$ osserviamo che il sistema di partenza è equivalente al sistema ridotto a scala:

$$\begin{cases} x + 2z = 3 \\ y - 3z = -6 \end{cases}$$

pertanto una soluzione particolare può essere calcolata sostituendo, per esempio, $z = 0$ nelle equazioni trovate: se $z = 0$, si ha $y = -6$, $x = 3$, quindi possiamo scegliere $\mathbf{v} = (3, -6, 0)$.

Il sistema omogeneo $A\mathbf{x} = \mathbf{0}$ è equivalente al sistema omogeneo ridotto a scala:

$$\begin{cases} x + 2z = 0 \\ y - 3z = 0 \end{cases}$$

che possiamo risolvere per sostituzioni successive dal basso, ottenendo: $y = 3z$, $x = -2z$. La controimmagine di $\mathbf{b}$ è dunque $S = \{(3, -6, 0) + (-2z, 3z, z) \mid z \in \mathbb{R}\}$.

## 6.3 ESERCIZI SVOLTI

### Esercizio 6.3.1

*Si calcolino le soluzioni del seguente sistema lineare nelle incognite $x_1$, $x_2$, $x_3$, $x_4$, $x_5$:*

$$\begin{cases} x_1 - x_2 + 3x_3 + x_5 = 2 \\ 2x_1 + x_2 + 8x_3 - 4x_4 + 2x_5 = 3 \\ x_1 + 2x_2 + 5x_3 - 3x_4 + 4x_5 = 1 \end{cases}$$

**Svolgimento**
La matrice completa associata al sistema è:

$$(A|\mathbf{b}) = \left(\begin{array}{ccccc|c} 1 & -1 & 3 & 0 & 1 & 2 \\ 2 & 1 & 8 & -4 & 2 & 3 \\ 1 & 2 & 5 & -3 & 4 & 1 \end{array}\right)$$

che, ridotta a scala, diventa:

$$(A'|\mathbf{b}') = \left(\begin{array}{ccccc|c} 1 & -1 & 3 & 0 & 1 & 2 \\ 0 & 3 & 2 & -4 & 0 & -1 \\ 0 & 0 & 0 & 1 & 3 & 0 \end{array}\right)$$

Si ha $\mathrm{rk}(A') = \mathrm{rk}(A'|\mathbf{b}') = 3$ quindi il sistema è risolubile, e le soluzioni dipendono da $5 - 3 = 2$ parametri.

I pivot di $A'$ si trovano sulla prima, sulla seconda e sulla quarta colonna, quindi possiamo ricavare le incognite $x_1, x_2$ e $x_4$ in funzione delle rimanenti $x_3$ e $x_5$.

Il sistema associato alla matrice a scala $(A'|\mathbf{b}')$ è:

$$\begin{cases} x_1 - x_2 + 3x_3 + x_5 = 2 \\ 3x_2 + 2x_3 - 4x_4 = -1 \\ x_4 + 3x_5 = 0 \end{cases}$$

Abbiamo dunque $x_4 = -3x_5$, $x_2 = -\frac{1}{3} - 4x_5 - \frac{2}{3}x_3$, $x_1 = \frac{5}{3} - 5x_5 - \frac{11}{3}x_3$. Le soluzioni sono quindi:

$$\left\{\left(-\frac{11}{3}x_3 - 5x_5 + \frac{5}{3}, -\frac{2}{3}x_3 - 4x_5 - \frac{1}{3}, x_3, -3x_5, x_5\right) | x_3, x_5 \in \mathbb{R}\right\}$$
$$= \left\{\left(\frac{5}{3}, -\frac{1}{3}, 0, 0, 0\right) + \mathbf{z} \,|\mathbf{z} \in \left\langle\left(-\frac{11}{3}, -\frac{2}{3}, 1, 0, 0\right), (-5, -4, 0, -3, 1)\right\rangle\right\}$$

■

**Esercizio 6.3.2**

*Si determinino le soluzioni del seguente sistema nelle incognite $x_1$, $x_2$, $x_3$:*

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

**Svolgimento**
La matrice completa associata al sistema è:

$$(A|\mathbf{b}) = \left(\begin{array}{ccc|c} 1 & -1 & 1 & 2 \\ 2 & -1 & 3 & -1 \\ 1 & 0 & 2 & 1 \end{array}\right)$$

che ridotta a scala diventa:

$$(A'|\mathbf{b}') = \left(\begin{array}{ccc|c} 1 & -1 & 1 & 2 \\ 0 & 1 & 1 & -5 \\ 0 & 0 & 0 & 4 \end{array}\right)$$

Si ha dunque $\mathrm{rk}(A') = 2 \neq \mathrm{rk}(A'|\mathbf{b}') = 3$. Il sistema pertanto non ammette soluzioni per il Teorema di Rouché-Capelli (6.2.9). Notiamo che il sistema lineare associato alla matrice $(A'|\mathbf{b}')$ è

$$\begin{cases} x_1 - x_2 + x_3 = 2 \\ x_2 + x_3 = -5 \\ 0 = 4 \end{cases}$$

che chiaramente non è compatibile.

■

**Esercizio 6.3.3**

a) *Si determini, se possibile, un sistema lineare avente come insieme di soluzioni*

$$S = (1, 2, 1) + \langle(1, 1, 1)\rangle$$

b) *Si determini, se possibile, un sistema lineare di tre equazioni avente come insieme di soluzioni*

$$S = (1, 2, 1) + \langle(1, 1, 1)\rangle$$

c) *Si determini, se possibile, un sistema lineare di rango 1 avente come insieme di soluzioni*

$$S = (1, 2, 1) + \langle(1, 1, 1)\rangle$$

**Svolgimento**

Poiché $S$ è un sottoinsieme di $\mathbb{R}^3$, ogni sistema lineare avente $S$ come insieme di soluzioni è un sistema lineare in 3 incognite. Indichiamo queste incognite con $x$, $y$, $z$. L'insieme delle soluzioni di un sistema lineare della forma $A\mathbf{x} = \mathbf{b}$ è $S = \{\mathbf{v} + \mathbf{z} | \mathbf{z} \in +\mathrm{Ker}\, A\}$, essendo $\mathbf{v}$ una soluzione particolare del sistema. Perciò, se $A\mathbf{x} = \mathbf{b}$ è un sistema lineare avente $S$ come insieme di soluzioni, $(1, 2, 1)$ è una soluzione del sistema e $\mathrm{Ker}\, A = \langle(1, 1, 1)\rangle$. In particolare notiamo che $(1, 2, 1) \notin \langle(1, 1, 1)\rangle$, perciò $S$ non è un sottospazio vettoriale di $\mathbb{R}^3$, dunque $\mathbf{b} \neq \mathbf{0}$. Inoltre $\dim(\mathrm{Ker}\, A) = 1 = 3 - \mathrm{rk}(A)$. Dunque il sistema che cerchiamo ha rango 2 e quindi deve essere necessariamente costituito da almeno 2 equazioni. Questo ci permette immediatamente di rispondere alla domanda c): non esiste alcun sistema lineare di rango 1 avente $S$ come insieme di soluzioni.

Per determinare un sistema lineare avente $S$ come insieme di soluzioni e quindi rispondere alla domanda a), potremmo dunque scrivere un sistema lineare generico costituito da due equazioni e imporre che:

1. $(1, 2, 1)$ sia soluzione del sistema;
2. $(1, 1, 1)$ sia soluzione del sistema omogeneo associato $A\mathbf{x} = \mathbf{0}$.

Questo metodo, che certamente funziona, non è tuttavia il più efficace. Scegliamo allora una strada più astuta.

Quello che vogliamo fare è descrivere mediante delle equazioni l'insieme degli elementi $(x, y, z) \in S$ cioè l'insieme degli elementi $(x, y, z)$ di $\mathbb{R}^3$ tali che:

$$(x, y, z) = (1, 2, 1) + z, \text{ con } z \in \langle(1, 1, 1)\rangle$$

o, equivalentemente,

$$(x, y, z) - (1, 2, 1) = z, \text{ con } z \in \langle(1, 1, 1)\rangle$$

Notiamo che il vettore $(x, y, z) - (1, 2, 1) = (x - 1, y - 2, z - 1)$ appartiene al sottospazio $\langle(1, 1, 1)\rangle$ se e solo se esso è un multiplo del vettore $(1, 1, 1)$, cioè se e solo se

$$\operatorname{rk}\begin{pmatrix} 1 & 1 & 1 \\ x-1 & y-2 & z-1 \end{pmatrix} = 1$$

Utilizzando l'algoritmo di Gauss, abbiamo che:

$$\operatorname{rk}\begin{pmatrix} 1 & 1 & 1 \\ x-1 & y-2 & z-1 \end{pmatrix} = \operatorname{rk}\begin{pmatrix} 1 & 1 & 1 \\ 0 & y-1-x & z-x \end{pmatrix}$$

Il rango della matrice trovata è uguale a 1 se e solo se la seconda riga della matrice trovata è nulla, cioè se e solo se

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

Abbiamo quindi trovato un sistema lineare avente $S$ come insieme di soluzioni. Naturalmente ogni sistema equivalente a quello trovato ha $S$ come insieme di soluzioni. In particolare, per rispondere alla domanda b) dovremo determinare un sistema lineare di 3 equazioni equivalente a quello appena scritto. Basterà aggiungere un'equazione che sia una combinazione lineare delle due equazioni trovate. Per esempio:

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

■

## ■ 6.4 ESERCIZI PROPOSTI

### Esercizio 6.4.1

*Sia* $F : \mathbb{R}^3 \to \mathbb{R}^3$ *l'applicazione lineare definita da:*

$$F(x, y, z) = (x - y, x + 2y, x + y + 3z)$$

*Si calcoli la controimmagine mediante* $F$ *del vettore* $(1, 0, -2)$.

### Esercizio 6.4.2

*Data l'applicazione lineare* $T_k : \mathbb{R}^4 \to \mathbb{R}^3$ *definita da:*

$$\begin{aligned} &T_k(x_1, x_2, x_3, x_4) \\ &= (x_1 - 5x_2 + kx_3 - kx_4, x_1 + kx_2 + kx_3 + 5x_4, 2x_1 - 10x_2 + (k+1)x_3 - 3kx_4) \end{aligned}$$

*si stabilisca per quali valori di* $k$ *il vettore* $\mathbf{w}_k = (1, k, -1)$ *appartiene a* $\operatorname{Im}(T_k)$. *Posto* $k = 0$, *si determini la controimmagine di* $\mathbf{w}_0$ *mediante* $T_0$.

### Esercizio 6.4.3

*Data l'applicazione lineare* $T : \mathbb{R}^3 \to \mathbb{R}^3$ *associata alla matrice*

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

*si stabilisca per quali valori di* $k$ *il vettore* $(k, 3, 3)$ *appartiene a* $\mathrm{Im}\,(T)$.
*Posto* $k = 2$ *si trovino tutti i vettori* $(x, y, z)$ *tali che* $T(x, y, z) = (2, 3, 3)$.
*Posto* $k = 3$ *si trovino tutti i vettori* $(x, y, z)$ *tali che* $T(x, y, z) = (3, 3, 3)$.

### Esercizio 6.4.4

*Sia* $S = \{(1, 2, 0, 3) + \mathbf{z} | \mathbf{z} \in \langle(1, -1, 2, 1), (1, 5, -2, 5)\rangle\}$. *Si stabilisca se* $S$ *è un sottospazio vettoriale di* $\mathbb{R}^4$ *e si determini, se possibile, un sistema lineare omogeneo avente* $S$ *come insieme di soluzioni.*

### Esercizio 6.4.5

*Si costruisca, se possibile, un'applicazione lineare* $F : \mathbb{R}^3 \to \mathbb{R}^3$ *tale che* $F^{-1}(1, 0, 0) = \{(1, 0, 0) + \mathbf{v} \,|\, \mathbf{v} \in \langle(1, 1, 1), (0, 1, -1)\rangle\}$. *Si stabilisca se una siffatta applicazione è unica.*

### Esercizio 6.4.6

*Si determini, se possibile:*

a) *un'equazione lineare avente* $S = (2, 1, 0, 1) + \langle(2, 1, 2, 2), (1, -1, 2, 1)\rangle$ *come insieme di soluzioni;*

b) *un sistema lineare di due equazioni avente* $S = (2, 1, 0, 1) + \langle(2, 1, 2, 2), (1, -1, 2, 1)\rangle$ *come insieme di soluzioni;*

c) *un sistema lineare di tre equazioni avente* $S = (2, 1, 0, 1) + \langle(2, 1, 2, 2), (1, -1, 2, 1)\rangle$ *come insieme di soluzioni.*

# 7 Determinante e inversa di una matrice

In questo capitolo vogliamo introdurre due concetti fondamentali: il *determinante* e l' *inversa* di una matrice. L'importanza di questi due concetti sarà sintetizzata dal Teorema 7.6.1 che contiene essenzialmente tutto ciò che avremo imparato sulle applicazioni lineari da $\mathbb{R}^n$ a $\mathbb{R}^n$.

## 7.1 DEFINIZIONE DI DETERMINANTE

Le nozioni di algebra introdotte fino a qui non sono sufficienti a permetterci di dare una definizione diretta di determinante; per completezza abbiamo riportato tale definizione nell'appendice al termine di questo capitolo. Vogliamo invece introdurre il determinante con una definizione "indiretta"; lo calcoleremo poi in alcuni casi particolari, che si riveleranno i più significativi per le applicazioni, e arriveremo infine a un metodo algoritmico per calcolarlo in generale.

Sia $A$ una matrice quadrata, cioè una matrice $n \times n$. Ad essa vogliamo associare un numero reale, detto il *determinante* di $A$, che si calcola a partire dagli elementi della matrice $A$. La definizione che diamo non è apparentemente una definizione costruttiva, tuttavia vedremo che è possibile ricavare da essa semplici regole con cui calcolare il determinante di una matrice.

Prima di iniziare, è necessario introdurre la definizione di matrice identità.

**Definizione 7.1.1** La *matrice identità* o *matrice identica* di ordine $n$, è la matrice $n \times n$ avente tutti gli elementi della diagonale principale uguali al numero 1, mentre i restanti elementi sono uguali a 0. Solitamente viene indicata con $I_n$, o con $I$ se non ci sono ambiguità.

Per esempio, la matrice identità di ordine 3 è

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

**Definizione 7.1.2** Il *determinante* è una funzione che a ogni matrice quadrata $A$ di ordine $n$ associa un numero reale, indicato con $\det(A)$, in modo che valgano le seguenti proprietà:

1. Se la riga $j$-esima di $A$ è somma di due elementi $\mathbf{u}$ e $\mathbf{v}$ di $\mathbb{R}^n$, allora il determinante di A è la somma dei determinanti delle due matrici ottenute sostituendo alla riga $j$-esima di $A$ rispettivamente $\mathbf{u}$ e $\mathbf{v}$.

2. Se la riga $j$-esima di $A$ è il prodotto $\lambda\mathbf{u}$, ove $\mathbf{u}$ è un elemento di $\mathbb{R}^n$ e $\lambda$ è uno scalare, allora il determinante di $A$ è il prodotto di $\lambda$ e del determinante della matrice ottenuta sostituendo la riga $j$-esima di $A$ con $\mathbf{u}$.

3. Se due righe di $A$ sono uguali, allora il determinante di $A$ è nullo.

4. Se $I$ è la matrice identità, allora $\det(I) = 1$.

Grazie a queste proprietà è possibile calcolare immediatamente il determinante di alcune matrici. Per esempio grazie alla proprietà (2), se consideriamo la matrice

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

allora

$$\det(A) = 2\det\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} = 2 \cdot 3\det\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 2 \cdot 3\det(I) = 6$$

Vedremo in seguito come sfruttare queste proprietà in modo opportuno per ricavare il determinante di una matrice qualsiasi.

Per il momento abbiamo definito il determinante come una funzione che possiede alcune proprietà, tuttavia ciò non garantisce che tale funzione esista o che, se esiste, sia unica. La prossima proposizione stabilisce tali fatti. Per la dimostrazione si rimanda all'appendice a questo capitolo.

**Proposizione 7.1.3** *Esiste una funzione che soddisfa le proprietà della Definizione* 7.1.2 *e tale funzione è unica.*

Grazie alla definizione di determinante possiamo subito dimostrare le seguenti ulteriori proprietà, che si riveleranno molto utili.

**Proposizione 7.1.4** *Siano $A$ e $B$ due matrici quadrate di ordine $n$.*

(a) *Se $B$ è ottenuta da $A$ scambiando due righe, allora:*

$$\det(A) = -\det(B)$$

(b) *Se $B$ è ottenuta da $A$ sommando a una riga di $A$ una qualunque combinazione lineare delle altre righe, allora:*

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

(c) *Se $A$ è una matrice triangolare superiore (o inferiore), cioè i coefficienti al di sotto (rispettivamente al di sopra) della diagonale principale sono tutti uguali a zero, allora il determinante di $A$ è il prodotto degli elementi che si trovano sulla sua diagonale principale.*

**Dimostrazione** – (a). Utilizziamo per una matrice $A$ la notazione in forma compatta, $A = (R_1, R_2, \dots, R_n)$ ove $R_1, \dots, R_n$ sono gli elementi di $\mathbb{R}^n$ che costituiscono le righe della matrice. Consideriamo ora la matrice $(R_1 + R_2, R_1 + R_2, R_3, \dots, R_n)$. Per la proprietà (3) della Definizione 7.1.2 abbiamo che:

$$\det(R_1 + R_2, R_1 + R_2, R_3, \dots, R_n) = 0$$

Per la proprietà (1) invece abbiamo che:

$$\begin{aligned}
&\det(R_1 + R_2, R_1 + R_2, R_3, \dots, R_n)\\
&= \det(R_1 + R_2, R_1, R_3, \dots R_n)\\
&\quad + \det(R_1 + R_2, R_2, R_3, \dots, R_n)\\
&= \det(R_1, R_1, R_3, \dots R_n) + \det(R_1, R_2, R_3, \dots, R_n)\\
&\quad + \det(R_2, R_1, R_3, \dots, R_n) + \det(R_2, R_2, R_3, \dots, R_n)\\
&= \det(R_1, R_2, R_3, \dots, R_n) + \det(R_2, R_1, R_3, \dots, R_n)
\end{aligned}$$

ove nell'ultimo passaggio abbiamo utilizzato nuovamente la proprietà (3). Quindi

$$\det(R_1, R_2, R_3, \dots, R_n) + \det(R_2, R_1, R_3, \dots, R_n) = 0$$

da cui segue (a) relativamente alle prime due righe. È chiaro che si può ripetere il discorso in modo identico per due righe generiche.

(b). Siano
$A = (R_1, R_2, \dots, R_n)$ e $B = (R_1 + \lambda_2 R_2 + \dots + \lambda_n R_n, R_2, \dots, R_n)$. Allora per le proprietà (1) e (2) della Definizione 7.1.2 abbiamo:

$$\begin{aligned}
\det(B) &= \det(R_1, R_2, \dots, R_n) + \lambda_2 \det(R_2, R_2, \dots, R_n) + \dots\\
&\quad + \lambda_n \det(R_n, R_2, \dots, R_n) = \det(A)
\end{aligned}$$

perché tutti i determinanti, a parte il primo, sono uguali a zero in quanto le matrici corrispondenti hanno due righe uguali.

(c). Sia $A$ una matrice triangolare inferiore. Supponiamo per il momento che tutti i coefficienti sulla diagonale siano diversi da zero. Utilizzando l'algoritmo di Gauss come descritto nel primo capitolo, possiamo sommare alla seconda riga un multiplo della prima in modo da azzerare il coefficiente di posto $(2, 1)$ e poi procedere analogamente azzerando tutti i coefficienti di posti $(3, 1), \dots, (n, 1)$. Il determinante della matrice così ottenuta non è cambiato, per la proprietà (b). Ripetiamo poi lo stesso procedimento per la seconda colonna, poi per la terza, sino all'$n$-esima. In tal modo otteniamo una matrice diagonale. Per la proprietà (2) della Definizione 7.1.2 abbiamo:

$$\det \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & \vdots & 0 & d_n \end{pmatrix} = d_1 d_2 \cdots d_n$$

Nel caso che uno o più coefficienti sulla diagonale siano uguali a zero non possiamo ottenere una matrice diagonale, tuttavia è facile convincersi che in questo caso, applicando l'algoritmo di Gauss, si ottiene una matrice in cui una riga è costituita da tutti zeri e, conseguentemente, il determinante è uguale a zero. Lasciamo al lettore i dettagli di questo caso. Il ragionamento per una matrice triangolare superiore è del tutto analogo, ma si deve iniziare ad azzerare i coefficienti della colonna $n$-esima, e procedere in ordine decrescente di colonna sino alla seconda colonna. □

Data una matrice quadrata $A$, tramite operazioni elementari di riga è sempre possibile ridurre $A$ in forma triangolare e quindi calcolarne il determinante tramite la proprietà appena viste. Si faccia attenzione al fatto che le operazioni elementari di riga possono cambiare il determinante: se si scambiano tra loro due righe si deve ricordare che il determinante cambia di segno; se si moltiplica una riga per uno scalare il determinante viene moltiplicato per lo stesso scalare; mentre se si somma a una riga una combinazione lineare delle altre, il determinante non cambia.

Vediamo un esempio esplicito, anche se il metodo che descriviamo non è il più efficiente per il calcolo del determinante in generale.

---

**Esempio 7.1.5**

Calcoliamo il determinante della matrice $A$:

$$A = \begin{pmatrix} 0 & 2 & 4 & 6 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 2 & -1 \\ 1 & 1 & 1 & 2 \end{pmatrix}$$

Vogliamo portare tale matrice in forma triangolare, utilizzando l'algoritmo di Gauss. Dobbiamo però tener conto di tutti gli scambi e degli eventuali fattori per cui moltiplichiamo le righe.

Scambiamo la prima riga con la seconda, in tal caso per la Proposizione 7.1.4 (a) il determinante cambia segno e la matrice diventa:

$$A = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 2 & 4 & 6 \\ 1 & 1 & 2 & -1 \\ 1 & 1 & 1 & 2 \end{pmatrix}$$

Ora effettuiamo le seguenti operazioni elementari, in modo da azzerare tutti i coefficienti nella prima colonna:

- $3^a$ riga $\to$ $3^a$ riga $-$ $1^a$ riga;
- $4^a$ riga $\to$ $4^a$ riga $-$ $1^a$ riga.

In tal modo il determinante non cambia, per la Proposizione 7.1.4 (b), e si ottiene:

$$\begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 2 & 4 & 6 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix}$$

Poi scambiamo la seconda con la terza riga; per la Proposizione 7.1.4 (a) il determinante cambia segno e la matrice diventa:

$$\begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 2 & 4 & 6 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Ora possiamo usare la proprietà (c) della proposizione precedente che ci dice che il determinante di una matrice triangolare è il prodotto dei coefficienti sulla diagonale.

$$\det\begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 2 & 4 & 6 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix} = 1 \cdot 2 \cdot (-1) \cdot (-1) = 2$$

Per ottenere il determinante di $A$ dobbiamo poi moltiplicare per $(-1)$ tante volte quanti sono stati gli scambi di riga (in questo caso due), dunque:

$$\det(A) = 2 \cdot (-1) \cdot (-1) = 2$$

## 7.2 CALCOLO DEL DETERMINANTE: CASI 2 × 2 E 3 × 3

Per le matrici $2 \times 2$ e $3 \times 3$ esistono formule semplici per il calcolo del determinante.

Cominciamo con l'esaminare il caso delle matrici $2 \times 2$. Sia $A$ la matrice:

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

Procediamo con l'algoritmo esemplificato qui di sopra, considerando due casi.

**Caso 1.** Supponiamo dapprima che sia $a_{11} \neq 0$. In tal caso effettuiamo la seguente operazione elementare:

- $2^a$ riga $\rightarrow$ $2^a$ riga $- \frac{a_{21}}{a_{11}} \cdot 1^a$ riga.

In tal modo per la Proposizione 7.1.4 (b) il determinante non cambia e otteniamo la matrice triangolare:

$$\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} - \frac{a_{21}}{a_{11}} \cdot a_{12} \end{pmatrix}$$

Ora è sufficiente fare il prodotto dei coefficienti della diagonale e si ha che:

$$\det(A) = a_{11} \cdot \left(a_{22} - \frac{a_{21}}{a_{11}} \cdot a_{12}\right) = a_{11}a_{22} - a_{12}a_{21}$$

**Caso 2.** Supponiamo $a_{11} = 0$. Scambiamo la prima e la seconda riga; il determinante cambia segno e otteniamo:

$$\begin{pmatrix} a_{21} & a_{22} \\ 0 & a_{12} \end{pmatrix}$$

Quindi, tenendo conto del cambio di segno,

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

essendo $a_{11} = 0$.

In entrambi i casi quindi vale la formula:

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

Dunque possiamo calcolare immediatamente il determinante di una qualsiasi matrice $2 \times 2$. Per esempio:

$$\det\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = 1 \cdot 4 - 2 \cdot 3 = -2$$

Consideriamo ora il caso delle matrici $3 \times 3$. Sia $A$ la matrice:

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

Procedendo in modo analogo al caso $2 \times 2$ (ovviamente l'algoritmo di Gauss richiede un maggior numero di passaggi), è possibile mostrare che il suo determinante è:

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

Vediamo un aiuto mnemonico per ricordare la formula. Riscriviamo le prime due colonne di $A$ di fianco ad $A$, a destra.

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

Per ottenere il determinante di $A$ è necessario fare la somma dei prodotti dei coefficienti sulle diagonali principali $a_{11}a_{22}a_{33}$, $a_{12}a_{23}a_{31}$, $a_{13}a_{21}a_{32}$ e poi sottrarre la somma dei prodotti di coefficienti sulle diagonali "opposte": $a_{13}a_{22}a_{31}$, $a_{11}a_{23}a_{32}$, $a_{12}a_{21}a_{33}$.

---

**Esempio 7.2.1**

Sia

$$A = \begin{pmatrix} 1 & 3 & 0 \\ 2 & 1 & -1 \\ 0 & -3 & 5 \end{pmatrix}$$

Consideriamo

$$\begin{pmatrix} 1 & 3 & 0 & 1 & 3 \\ 2 & 1 & -1 & 2 & 1 \\ 0 & -3 & 5 & 0 & -3 \end{pmatrix}$$

Abbiamo quindi $\det(A) = 5 + 0 - 0 - (0 + 3 + 30) = -28$.

---

## 7.3 CALCOLO DEL DETERMINANTE: METODO RICORSIVO

Per matrici di ordine più grande di due o tre, ma anche come metodo alternativo rispetto a quanto studiato nel paragrafo precedente, diamo un metodo basato su di una definizione *ricorsiva* di determinante del tutto equivalente a quella già vista. Per la dimostrazione di tutte le affermazioni, rimandiamo all'appendice di questo capitolo.

Cominciamo con il definire il concetto di *minore* di una matrice.

**Definizione 7.3.1** Sia $A \in M_{n\times n}(\mathbb{R})$ una matrice quadrata di ordine $n$ (cioè con $n$ righe e $n$ colonne). Indichiamo con $A_{ij}$ la sottomatrice quadrata di $A$ ottenuta cancellando la $i$-esima riga e la $j$-esima colonna di $A$. Allora $A_{ij}$ si dice un *minore* di $A$ di ordine $n-1$.

Vediamo degli esempi.

---

**Esempio 7.3.2**

Se

$$A = \begin{pmatrix} -1 & 0 & 4 & 7 \\ 3 & -1 & 2 & 1 \\ 3 & 6 & -1 & 0 \\ -2 & -1 & 0 & -1 \end{pmatrix}$$

abbiamo

$$A_{11} = \begin{pmatrix} -1 & 2 & 1 \\ 6 & -1 & 0 \\ -1 & 0 & -1 \end{pmatrix} \quad A_{23} = \begin{pmatrix} -1 & 0 & 7 \\ 3 & 6 & 0 \\ -2 & -1 & -1 \end{pmatrix}$$

$$A_{42} = \begin{pmatrix} -1 & 4 & 7 \\ 3 & 2 & 1 \\ 3 & -1 & 0 \end{pmatrix}$$

---

Vogliamo ora enunciare una proposizione che permette di calcolare il determinante di una matrice $A$ procedendo ricorsivamente sull'ordine di $A$, cioè sul numero delle sue righe (o colonne). Rimandiamo all'appendice per una dimostrazione di tutte le nostre affermazioni.

**Teorema di Laplace 7.3.3** *Sia $A$ una matrice quadrata.*

1. *Se $A$ ha ordine 1, cioè $A = (a_{11})$ ha una riga e una colonna, poniamo*

$$\det(A) = a_{11}$$

2. *Supponiamo ora di saper calcolare il determinante delle matrici di ordine $n-1$. Sia*

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

*allora*

$$\det(A) = a_{11}\Gamma_{11} + a_{12}\Gamma_{12} + \ldots + a_{1n}\Gamma_{1n} = \sum_{k=1}^{n} a_{1k}\Gamma_{1k}$$

Questo è il metodo per calcolo del determinante sviluppando secondo la prima riga. Vediamo come si traduce in pratica, per le matrici $2\times 2$ e $3\times 3$ ritroviamo i risultati visti in precedenza.

Vediamo subito infatti che se $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ è una matrice $2 \times 2$ abbiamo

$$\det(A) = a\Gamma_{11} + b\Gamma_{12} = ad - bc$$

Per esempio: $\det \left(\begin{smallmatrix} 2 & 3 \\ 1 & 4 \end{smallmatrix}\right) = 2 \cdot 4 - 3 \cdot 1 = 8 - 3 = 5.$

Vediamo ora il caso delle matrici $3 \times 3$. Se

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

abbiamo per definizione:

$$\Gamma_{11} = (-1)^{1+1} \det A_{11} = \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}$$

$$\Gamma_{12} = (-1)^{1+2} \det A_{12} = -\det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix}$$

$$\Gamma_{13} = (-1)^{1+3} \det A_{13} = \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

Quindi

$$\begin{aligned}
\det(A) &= a_{11}\Gamma_{11} + a_{12}\Gamma_{12} + a_{13}\Gamma_{13} \\
&= a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13} \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \\
&= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) \\
&\quad + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\
&= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} \\
&\quad - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}
\end{aligned}$$

come abbiamo visto in precedenza.

È possibile sviluppare il determinante secondo una qualsiasi riga o colonna. Lo sviluppo di $\det(A)$ secondo la $r$-esima riga è dato da:

$$\det(A) = a_{r1}\Gamma_{r1} + a_{r2}\Gamma_{r2} + \ldots + a_{rn}\Gamma_{rn} = \sum_{k=1}^{n} a_{rk}\Gamma_{rk}$$

Lo sviluppo di $\det(A)$ secondo la $s$-esima colonna è dato da:

$$\det(A) = a_{1s}\Gamma_{1s} + a_{1s}\Gamma_{2s} + \ldots + a_{ns}\Gamma_{ns} = \sum_{k=1}^{n} a_{ks}\Gamma_{ks}$$

Si può dimostrare (si veda l'appendice a questo capitolo) che in generale, se si sviluppa il determinante di una matrice $A \in \mathrm{M}_n(\mathbb{R})$ secondo una qualsiasi riga o colonna, si ottiene sempre lo stesso numero, quindi conviene sviluppare secondo la riga o la colonna con il maggior numero di zeri.

Vediamo ora un esempio, rimandando all'appendice per i dettagli e la trattazione più generale.

---

**Esempio 7.3.4**

Sia

$$A = \begin{pmatrix} 1 & 3 & 0 & 0 \\ 2 & 0 & 0 & -1 \\ 0 & -3 & 5 & 0 \\ 1 & 0 & -1 & 0 \end{pmatrix}$$

Lo sviluppo di $\det(A)$ secondo la terza riga è:

$$\det(A) = 0\Gamma_{31} - 3\Gamma_{32} + 5\Gamma_{33} + 0\Gamma_{34} = -3\Gamma_{32} + 5\Gamma_{33}$$

Ora

$$\Gamma_{32} = (-1)^{3+2} \det \begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix} = -(-1) = 1$$

$$\Gamma_{33} = (-1)^{3+3} \det \begin{pmatrix} 1 & 3 & 0 \\ 2 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix} = +(-3) = -3$$

quindi $\det(A) = -3 \cdot 1 + 5(-3) = -18$.

Vediamo ora invece lo sviluppo di $\det(A)$ secondo la seconda colonna:

$$\det(A) = 3\Gamma_{12} + 0\Gamma_{22} - 3\Gamma_{32} + 0\Gamma_{42} = 3\Gamma_{12} - 3\Gamma_{32}$$

Ora

$$\Gamma_{12} = (-1)^{1+2} \det \begin{pmatrix} 2 & 0 & -1 \\ 0 & 5 & 0 \\ 1 & -1 & 0 \end{pmatrix} = -(+5) = -5$$

$$\Gamma_{32} = (-1)^{3+2} \det \begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix} = -(-1) = 1$$

quindi $\det(A) = 3(-5) - 3 \cdot 1 = -18$.

Questo esempio particolare conferma che se sviluppiamo il determinante secondo la terza riga o la seconda colonna il risultato non cambia.

---

Un'importanza particolare è riservata al Teorema di Binet, che afferma che la funzione determinante gode della proprietà moltiplicativa, cioè il determinante di un prodotto di matrici è il prodotto dei determinanti. Si noti che per la somma ciò è molto lontano dall'essere vero. Per la dimostrazione del Teorema di Binet rimandiamo all'appendice.

**Teorema di Binet 7.3.5** *Siano $A$ e $B$ due matrici quadrate $n \times n$, allora*

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

## 7.4 INVERSA DI UNA MATRICE

Nell'insieme dei numeri reali, l'inverso di un numero $t \in \mathbb{R} \setminus \{0\}$ è il numero $s$ definito attraverso la proprietà: $st = ts = 1$ (dunque $s$ si indica con $1/t$, come tutti sappiamo). Poiché è possibile moltiplicare le matrici quadrate tra loro, possiamo chiederci se, data una matrice quadrata $A$, esista una matrice quadrata $B$ tale che $AB = BA = I$, ove $I$ è la matrice identità, che qui gioca lo stesso ruolo dell'unità nei numeri reali. Tale matrice $B$ viene detta l'inversa di $A$. Per il Teorema di Binet, che abbiamo appena visto, è chiaro che se $\det(A) = 0$ non è possibile trovare una matrice $B$ con tale proprietà, perché deve succedere che $\det(AB) = \det(A)\det(B) = \det(I) = 1$. Vedremo tra breve che tale condizione è anche sufficiente.

Cominciamo la nostra trattazione con la definizione di inversa di una matrice quadrata per poi passare ai vari metodi di calcolo.

**Definizione 7.4.1** Una matrice quadrata $A$ si dice *invertibile* se esiste una matrice $B$ (denotata con $A^{-1}$) tale che:

$$AB = BA = I$$

Vediamo ora come calcolare l'inversa. Un primo metodo diretto ci viene dalla dimostrazione del seguente teorema, che caratterizza le matrici invertibili.

**Teorema 7.4.2** *La matrice $A$ è invertibile se e solo se il suo determinante è diverso da zero.*

**Dimostrazione** – Dimostriamo prima che, se $A$ è invertibile, allora il suo determinante è diverso da zero. Per definizione abbiamo $AA^{-1} = I$, quindi, per il Teorema di Binet (7.3.5), $\det(A)\det(A^{-1}) = \det(AA^{-1}) = \det(I) = 1$, pertanto $\det(A) \neq 0$.

Viceversa, se il determinante di $A$ è diverso da zero, si può costruire l'inversa di $A$ nel modo seguente. Sia $\det(A_{ij})$ il determinante della matrice ottenuta da $A$ rimuovendo la $i$-esima riga e la $j$-esima colonna. Abbiamo allora che

$$(A^{-1})_{ij} = \frac{1}{\det(A)}(-1)^{i+j}\det(A_{ji}) \tag{7.1}$$

dove $(A^{-1})_{ij}$ indica l'elemento di posto $i, j$ della matrice $A^{-1}$. Dimostreremo in appendice a questo capitolo che si tratta proprio dell'inversa di $A$. □

♦ **Osservazione 7.4.3** Supponiamo di avere due matrici quadrate $A, B$ tali che $AB = I$. Allora per il Teorema di Binet (7.3.5) abbiamo che $1 = \det(I) = \det(AB) = \det(A)\det(B)$, quindi $\det(A)$ e $\det(B)$ sono diversi da zero. Per il Teorema 7.4.2 abbiamo che $A$ e $B$ sono entrambe invertibili, quindi moltiplicando a sinistra entrambi i membri dell'uguaglianza $AB = I$ per $A^{-1}$ si ottiene che $B = A^{-1}$ è l'inversa di $A$.

Grazie a questa osservazione possiamo dire che, date due matrici quadrate $A$ e $B$, si ha che $AB = I$ se e solo se $BA = I$ e che l'inversa di una matrice è unica.

Vediamo ora la formula esplicita per calcolare l'inversa di una matrice $2 \times 2$. Sia $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Se il determinante $\det(A) = ad - bc$ di $A$ è diverso da zero, l'inversa di $A$ si può calcolare grazie alla formula (7.1) della proposizione precedente. Si ha che:

$$\begin{aligned}
(A^{-1})_{11} &= \frac{1}{ad-bc}d & (A^{-1})_{12} &= \frac{1}{ad-bc}(-b) \\
(A^{-1})_{21} &= \frac{1}{ad-bc}(-c) & (A^{-1})_{22} &= \frac{1}{ad-bc}a
\end{aligned}$$

quindi

$$A^{-1} = \begin{pmatrix} \dfrac{d}{ad-bc} & \dfrac{-b}{ad-bc} \\ \dfrac{-c}{ad-bc} & \dfrac{a}{ad-bc} \end{pmatrix}$$

Lasciamo per esercizio la facile verifica che si tratta proprio dell'inversa di $A$, cioè che:

$$\begin{pmatrix} \dfrac{d}{ad-bc} & \dfrac{-b}{ad-bc} \\ \dfrac{-c}{ad-bc} & \dfrac{a}{ad-bc} \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \dfrac{d}{ad-bc} & \dfrac{-b}{ad-bc} \\ \dfrac{-c}{ad-bc} & \dfrac{a}{ad-bc} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Ora illustreremo un metodo alternativo per il calcolo della matrice inversa. È difficile dire quale dei due metodi sia più veloce, in quanto dipende molto dal tipo di matrice che si considera.

## 7.5 CALCOLO DELL'INVERSA CON IL METODO DI GAUSS

Supponiamo di avere una matrice $A$ invertibile e di volerne calcolare l'inversa. Consideriamo la matrice:

$$M = (A \mid I)$$

ottenuta mettendo accanto ad $A$ la matrice identità. Poi, tramite operazioni elementari sulle righe della matrice $M$, è possibile fare in modo di ottenere a sinistra la matrice identità. Indichiamo brevemente la procedura, che poi chiariremo negli esempi successivi. Sappiamo già come ottenere a sinistra una matrice a scala $C$ e, poiché le operazioni elementari sulle righe preservano il rango e la matrice $A$ è invertibile, la matrice $C$ avrà esattamente $n$ pivot. Dividendo ciascuna riga per il pivot corrispondente possiamo inoltre supporre che tutti i pivot di $C$ siano uguali a 1. Si procede poi "dal basso" utilizzando l'ultimo pivot e opportune operazioni elementari sulle righe per far comparire tutti zeri nella colonna tale pivot, poi si passa al penultimo pivot e così via, ottenendo al termine la matrice identità sulla sinistra. A questo punto, la matrice che compare a destra è la matrice inversa di $A$. Per il momento applichiamo questo metodo senza motivarlo, per una trattazione accurata, rimandiamo all'appendice di questo capitolo.

Vediamo un esempio.

**Esempio 7.5.1**

Consideriamo la matrice

$$A = \begin{pmatrix} 1 & 3 \\ 1 & 4 \end{pmatrix}$$

Per calcolarne l'inversa, dobbiamo applicare il metodo di Gauss alla matrice:

$$\left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 1 & 4 & 0 & 1 \end{array}\right)$$

Effettuiamo la seguente operazione elementare: $2^a$ riga $\to$ $2^a$ riga $-$ $1^a$ riga, e otteniamo:

$$\left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & 1 & -1 & 1 \end{array}\right)$$

Poi facciamo: $1^a$ riga $\to$ $1^a$ riga $- 3 \cdot 2^a$, ottenendo:

$$\left(\begin{array}{cc|cc} 1 & 0 & 4 & -3 \\ 0 & 1 & -1 & 1 \end{array}\right)$$

Abbiamo dunque:

$$A^{-1} = \begin{pmatrix} 4 & -3 \\ -1 & 1 \end{pmatrix}$$

---

Vediamo un altro esempio, questa volta parametrico.

---

**Esempio 7.5.2**

Vogliamo stabilire per quali valori di $k$ la matrice

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

sia invertibile e, per tali valori, vogliamo calcolarne l'inversa.

Calcoliamo prima il determinante di $A$, per esempio sviluppando secondo la terza colonna.

$$\det(A) = 0 \cdot \Gamma_{13} + 0 \cdot \Gamma_{23} + k \cdot \Gamma_{33} = k(-1)^6 \det\begin{pmatrix} 1 & k \\ 1 & 2k-2 \end{pmatrix} = k(k-2)$$

$A$ è invertibile per $k \neq 0$ e $k \neq 2$. Consideriamo ora:

$$(A \mid I) = \left(\begin{array}{ccc|ccc} 1 & k & 0 & 1 & 0 & 0 \\ 1 & 2k-2 & 0 & 0 & 1 & 0 \\ 2k & 0 & k & 0 & 0 & 1 \end{array}\right)$$

e applichiamo il metodo di Gauss:

$$\begin{array}{l} 2^a \text{ riga} \to \quad 2^a \text{ riga} - 1^a \text{ riga} \\ 3^a \text{ riga} \to 3^a \text{ riga} - 2k \cdot 1^a \text{ riga} \end{array} \to \left(\begin{array}{ccc|ccc} 1 & k & 0 & 1 & 0 & 0 \\ 0 & k-2 & 0 & -1 & 1 & 0 \\ 0 & -2k^2 & k & -2k & 0 & 1 \end{array}\right)$$

$$2^a \text{ riga} \to \frac{1}{k-2} \cdot 2^a \text{ riga} \to \left(\begin{array}{ccc|ccc} 1 & k & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -\dfrac{1}{k-2} & \dfrac{1}{k-2} & 0 \\ 0 & -2k^2 & k & -2k & 0 & 1 \end{array}\right)$$

$$3^a \text{ riga} \to 3^a \text{ riga} + 2k^2 \cdot 2^a \text{ riga} \to \left(\begin{array}{ccc|ccc} 1 & k & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -\dfrac{1}{k-2} & \dfrac{1}{k-2} & 0 \\ 0 & 0 & k & \dfrac{4k(1-k)}{k-2} & \dfrac{2k^2}{k-2} & 1 \end{array}\right)$$

$$\begin{array}{l} 1^a \text{ riga} \to 1^a \text{ riga} - k \cdot 2^a \text{ riga} \\ 3^a \text{ riga} \to \frac{1}{k} \cdot 3^a \text{ riga} \end{array} \quad \to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{2k-2}{k-2} & -\frac{k}{k-2} & 0 \\ 0 & 1 & 0 & -\frac{1}{k-2} & \frac{1}{k-2} & 0 \\ 0 & 0 & 1 & \frac{4(1-k)}{k-2} & \frac{2k}{k-2} & \frac{1}{k} \end{array}\right)$$

L'inversa è quindi:

$$\begin{pmatrix} \frac{2k-2}{k-2} & -\frac{k}{k-2} & 0 \\ -\frac{1}{k-2} & \frac{1}{k-2} & 0 \\ \frac{4(1-k)}{k-2} & \frac{2k}{k-2} & \frac{1}{k} \end{pmatrix}.$$

---

Nel prossimo paragrafo metteremo in relazione il concetto di determinante e inversa di una matrice, con le proprietà dell'applicazione lineare ad essa associata una volta fissate le basi canoniche nel dominio e nel codominio.

## 7.6 LE APPLICAZIONI LINEARI DA $\mathbb{R}^N$ A $\mathbb{R}^N$

Grazie al concetto di determinante e di inversa di una matrice, possiamo ora dare un risultato importantissimo che ci permette di caratterizzare le applicazioni lineari invertibili da $\mathbb{R}^n$ a $\mathbb{R}^n$.

**Teorema 7.6.1** *Sia $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ un'applicazione lineare, e sia $A$ la matrice associata a $F$ rispetto alla base canonica (nel dominio e nel codominio). Le seguenti affermazioni sono equivalenti:*

1. *$F$ è un isomorfismo;*
2. *$F$ è iniettiva;*
3. *$F$ è suriettiva;*
4. $\dim(\text{Im}\,(F)) = n$;
5. $\text{rk}(A) = n$;
6. *le colonne di $A$ sono linearmente indipendenti;*
7. *le righe di $A$ sono linearmente indipendenti;*
8. *il sistema $A\mathbf{x} = \mathbf{0}$ ha un'unica soluzione;*
9. *per ogni $\mathbf{b} \in \mathbb{R}^n$ il sistema $A\mathbf{x} = \mathbf{b}$ ha un'unica soluzione;*
10. *$A$ è invertibile;*
11. *il determinante di $A$ è diverso da zero.*

**Dimostrazione** – Per la Proposizione 5.5.2 abbiamo subito l'equivalenza tra (1), (2), (3). Mostriamo ora che le affermazioni (3)-(9) sono equivalenti, dimostrando che ciascuna di esse implica la successiva e poi che (9) implica (2). Mostreremo infine che (1), (10), (11) sono equivalenti.

(3) implica (4) perché, se $F$ è suriettiva, allora Im $F = \mathbb{R}^n$ ha dimensione $n$.

(4) implica (5) perché $\mathrm{rk}(A) = \dim(\mathrm{Im}\ (F))$ per l'Osservazione 6.2.2.

(5) implica (6) per definizione di rango di una matrice (che in particolare è il rango colonne).

(6) implica (7) perché il rango righe di una matrice è uguale al rango colonne (Proposizione 6.2.3).

(7) implica (8) perché, se le righe di $A$ sono linearmente indipendenti, quando applichiamo il metodo di Gauss per risolvere il sistema $A\mathbf{x} = \mathbf{0}$ troviamo una matrice a scala con esattamente $n$ pivot, quindi c'è un unica soluzione.

Mostriamo ora che (8) implica (9). Se il sistema $A\mathbf{x} = \mathbf{0}$ ha un'unica soluzione, allora riducendo a scala la matrice $A$, si ottiene una matrice $A'$ con esattamente $n$ pivot. Allora riducendo a scala la matrice $A|\mathbf{b}$ si ottiene una matrice del tipo $A'|\mathbf{b}'$, che ha anch'essa esattamente $n$ pivot (quelli di $A'$). Quindi il sistema $A\mathbf{x} = \mathbf{b}$ ammette un'unica soluzione.

Mostriamo che (9) implica (2). Per la Proposizione 5.4.8 basta mostrare che $\ker(F) = \mathbf{0}$. Ma $\ker(F)$ è costituito dalle soluzioni del sistema lineare omogeneo $A \cdot \mathbf{x} = \mathbf{0}$, e grazie a (9), prendendo $\mathbf{x} = \mathbf{0}$, esso ha solo una soluzione, che deve essere per forza quella nulla, quindi $F$ è iniettiva.

Abbiamo mostrato che le condizioni (2) – (9) sono tra loro equivalenti.

Mostriamo ora che (1) è implica (10). Sia $G$ l'inversa di $F$, quindi $F \circ G = G \circ F = \mathrm{id}_{\mathbb{R}^n}$. Sia $B$ la matrice associata a $G$ rispetto alla base canonica. Allora $AB = BA = I$, quindi $B$ è l'inversa di $A$.

(10) implica (1) perché, se $B$ è l'inversa di $A$ e $L_B : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ è l'applicazione lineare ad essa associata, allora $L_B$ è l'inversa di $F$.

(10) è equivalente a (11) per il Teorema 7.4.2. □

## 7.7 ESERCIZI SVOLTI

### Esercizio 7.7.1

*Si stabilisca per quali valori di $k$ la matrice*

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

*è invertibile e per tali valori se ne calcoli l'inversa.*

**Svolgimento**

Calcoliamo prima il determinante: $\det(A) = -(k-3) - 4k = -5k + 3$, quindi $A$ è invertibile per $k \neq \frac{3}{5}$.

Usiamo la formula (7.1) per l'inversa:

$$(A^{-1})_{11} = \frac{1}{\det(A)}(-1)^{1+1}\det(A_{11}) = \frac{-1}{-5k+3}$$

$$(A^{-1})_{12} = \frac{1}{\det(A)}(-1)^{1+2}\det(A_{21}) = -\frac{2k}{-5k+3}$$
$$(A^{-1})_{21} = \frac{1}{\det(A)}(-1)^{2+1}\det(A_{12}) = -\frac{2}{-5k+3}$$
$$(A^{-1})_{22} = \frac{1}{\det(A)}(-1)^{2+2}\det(A_{22}) = \frac{k-3}{-5k+3}$$

L'inversa è dunque:

$$A^{-1} = \begin{pmatrix} \frac{1}{5k-3} & \frac{2k}{5k-3} \\ \frac{2}{5k-3} & \frac{k-3}{-5k+3} \end{pmatrix}$$

■

### Esercizio 7.7.2

*Data l'applicazione lineare $F : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definita da: $F(\mathbf{e}_1) = 2\mathbf{e}_1 - \mathbf{e}_2$, $F(\mathbf{e}_2) = \mathbf{e}_2 + k\mathbf{e}_3$, $F(\mathbf{e}_3) = \mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_3$, si dica per quali valori di $k$ $F$ è un isomorfismo. Scelto un opportuno valore di $k$, si calcoli $F^{-1}$.*

**Svolgimento**

La matrice associata a $F$ rispetto alla base canonica di dominio e codominio è

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

Se calcoliamo il determinante di $A$ con uno qualunque dei metodi che abbiamo visto, per esempio sviluppando secondo la prima riga, otteniamo:

$$\det(A) = (-1)^{1+1}2(1+k) + 0 + (-1)^{1+3}(-k) = k+2$$

Per il Teorema 7.6.1 sappiamo che $F$ è isomorfismo se e solo se il determinante di $A$ è diverso da zero, pertanto $F$ è isomorfismo se e solo se $k \neq -2$. Possiamo dunque scegliere un qualsiasi valore di $k$ diverso da 2 per calcolare $F^{-1}$. Scegliamo $k = 0$, in quanto ciò ci semplificherà i calcoli. La matrice associata all'inversa di $F$ rispetto alla base canonica nel dominio e nel codominio è l'inversa della matrice $A$. Calcoliamo pertanto tale inversa utilizzando la formula diretta.

$$A^{-1} = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1/2 & 0 & -1/2 \\ 1/2 & 1 & 1/2 \\ 0 & 0 & 1 \end{pmatrix}$$

Anche se non necessario per l'esercizio è sempre una buona norma accertarsi che $A^{-1}$ sia effettivamente l'inversa di $A$. A tale scopo è sufficiente fare il prodotto righe per colonne di $A$ e $A^{-1}$ e verificare che il risultato sia l'identità.

$$AA^{-1} = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1/2 & 0 & -1/2 \\ 1/2 & 1 & 1/2 \\ 0 & 0 & 1 \end{pmatrix} = I$$

Pertanto l'inversa di $F$ è

$$F^{-1}(x,y,z) = \left(\frac{1}{2}x - \frac{1}{2}z, \frac{1}{2}x + y + \frac{1}{2}z, z\right)$$

■

## 7.8 ESERCIZI PROPOSTI

### Esercizio 7.8.1

*Sia data l'applicazione lineare* $F : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *definita da:* $F(x,y,z) = (x + 2z, y + z, z)$.
a) *Si stabilisca se* $F$ *è un isomorfismo, motivando la risposta.*
b) *In caso di risposta affermativa al punto* (a) *si calcoli l'inversa di* $F$.

### Esercizio 7.8.2

*Si consideri l'applicazione* $F : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *tale che:* $F(x,y,z) = (x + y - 2z, -y + z, -z)$.
*Si determini un'applicazione* $G : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *tale che* $F \circ G = id$.

### Esercizio 7.8.3

*Si determini per quali valori di* $a$ *la matrice*

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

*è invertibile.*
*Scelto uno dei valori per cui è invertibile se ne calcoli l'inversa.*

### Esercizio 7.8.4

*Si determini per quali valori di* $a$ *la matrice*

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

*è invertibile.*
*Scelto uno dei valori per cui è invertibile, se ne calcoli l'inversa.*

### Esercizio 7.8.5

*Sia* $\mathbf{e}_1 = (1,0,0)$, $\mathbf{e}_2 = (0,1,0)$, $\mathbf{e}_3 = (0,0,1)$ *la base canonica dello spazio vettoriale reale* $\mathbb{R}^3$ *e sia* $a \in \mathbb{R}$. *Sia* $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *l'applicazione lineare tale che* $T(\mathbf{e}_1) = \mathbf{e}_1 - a\mathbf{e}_2$, $T(\mathbf{e}_2) = \mathbf{e}_2 + \mathbf{e}_3$, $T(\mathbf{e}_3) = a\mathbf{e}_3$.
a) *Si trovino i valori di* $a$ *per i quali l'applicazione è invertibile.*
b) *Scelto uno dei valori di* $a$ *per i quali* $T$ *è invertibile, se ne calcoli l'inversa.*

### Esercizio 7.8.6

*Si calcoli, se è possibile, l'inversa della matrice* $A$.

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

### Esercizio 7.8.7

a) *Si calcoli l'inversa della matrice:*

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

*con un qualunque metodo.*

b) *Sia $L_A$ l'applicazione lineare associata ad $A$ e sia $T$ l'applicazione lineare:*

$$\begin{array}{rccc} T: & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & (x,y,z) & \longrightarrow & (-x+y-z, z, x-y) \end{array}$$

*Si determini $T \circ L_A$.*

## 7.9 APPENDICE: APPROFONDIMENTI

In questa appendice vogliamo dare una definizione alternativa, ma equivalente, di determinante di una matrice quadrata $n \times n$. Anziché definirlo indirettamente tramite le sue proprietà come abbiamo fatto nel testo, vedremo una definizione diretta, tramite il concetto di *permutazione* che è di per sé estremamente importante, anche se marginale nella nostra scelta espositiva della teoria.

Questa appendice non è necessaria per poter continuare la lettura, ma rappresenta un approfondimento dei concetti esposti in questo capitolo. Per la natura e profondità degli argomenti di questa appendice ci sarà impossibile dare una trattazione completa, rimandiamo il lettore interessato al testo fondamentale di S. Lang, *Algebra Lineare*, per ulteriori dettagli.

**Definizione 7.9.1** Sia $\{1,\dots,n\}$ l'insieme dei primi $n$ numeri naturali. Si definisce *permutazione* una funzione biettiva $\sigma : \{1,\dots,n\} \longrightarrow \{1,\dots,n\}$.

Una permutazione dunque associa a un numero compreso tra 1 e $n$ un numero compreso tra 1 ed $n$. Per esempio la funzione $\sigma : \{1,2\} \longrightarrow \{1,2\}$, definita da: $\sigma(1) = 2$, $\sigma(2) = 1$ è la permutazione di $\{1,2\}$ che *scambia* il numero 1 con il numero 2. Una permutazione che scambia soltanto due numeri $i$, $j$ e lascia gli altri invariati si dice *trasposizione* e si indica brevemente con $(i,j)$. Per esempio la trasposizione $\sigma$ descritta sopra si indica con $\sigma = (1,2)$.

Un altro esempio di permutazioni è costituito dai cicli. Un *ciclo* si indica con $(i_1, i_2, \dots, i_s)$, ove $\{i_1, i_2, \dots, i_s\}$ è un sottoinsieme di $\{1,\dots,n\}$, ed è la permutazione che manda ogni $i_j$ in $i_{j+1}$ per $j = 1,\dots,n-1$, infine manda $i_n$ in $i_1$, lasciando invariati tutti gli altri interi. Per esempio il ciclo $(3,1,5)$ è la permutazione di $\{1,\dots,5\}$ che manda 3 in 1, 1 in 5, 5 in 3 e fissa 2 e 4.

Si dice che $r$ cicli $s_1,\dots,s_r$ sono *disgiunti* se ciascun elemento di $\{1,\dots,n\}$ viene fissato da tutti i cicli tranne al più uno di essi.

Indichiamo le permutazioni anche in questo modo:

$$\sigma = \begin{pmatrix} 1 & \dots & n \\ \sigma(1) & \dots & \sigma(n) \end{pmatrix}$$

e il loro insieme con $S_n$.

Vediamo qualche esempio.

**Esempio 7.9.2**

i) Consideriamo la permutazione

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}$$

Questa permutazione scambia (oppure diciamo anche *permuta*) gli elementi 2 e 4 e lascia inalterati 1 e 3. Si indica anche per semplicità $\sigma_1 = (2,4)$ e, come abbiamo visto, si dice *trasposizione*.
ii) Consideriamo la permutazione:

$$\sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}$$

Questa permutazione è la funzione che si ottiene come composizione delle permutazioni $(2,3)$ e $(1,2)$, cioè $\sigma_2 = (1,2) \circ (2,3)$. Infatti prima scambiamo 2 con 3 e poi 1 con 2 ottenendo il ciclo $(1,2,3)$. Si noti che il fatto che 4 non compaia nella scrittura del ciclo significa che il numero 4 è lasciato invariato dalla permutazione.

---

Definiamo ora la *parità* di una permutazione $\sigma$.

**Definizione 7.9.3** Data una permutazione $\sigma$, consideriamo il numero di *inversioni* e cioè di coppie $(i,j)$ tali che $i < j$ ma $\sigma(i) > \sigma(j)$. Se tale numero è pari diremo che la permutazione $\sigma$ è *pari*, o che la sua parità è $p(\sigma) = 1$, se tale numero è dispari diremo allora che la permutazione è *dispari*, cioè che la sua parità è $p(\sigma) = -1$. In altre parole $p(\sigma) = (-1)^i$ ove $i$ è il numero di inversioni.

Nell'esempio considerato sopra $p(\sigma_1) = -1$, mentre $p(\sigma_2) = 1$.

Ogni permutazione può essere scritta come composizione di trasposizioni, infatti si può dimostrare che ogni funzione da $\{1,\ldots, n\}$ a $\{1, \ldots, n\}$ può essere realizzata effettuando scambi successivi. Si faccia tuttavia attenzione al fatto che, in generale, non c'è unicità di scrittura: una stessa permutazione può apparire in due modi diversi come composizione di trasposizioni. Per esempio, l'identità in $S_3$ si ottiene come la composizione $(1,2) \circ (1,2)$, ma anche come $(1,3) \circ (1,3)$.

La parità di una permutazione $\sigma$ può anche essere definita equivalentemente come $(-1)^m$, dove $m$ è il numero di trasposizioni la cui composizione è $\sigma$.

Nonostante una stessa permutazione possa apparire come composizione di trasposizioni diverse, anche in numero diverso, una volta fissata una permutazione il numero di tali trasposizioni è sempre della stessa parità (cioè sempre pari o sempre dispari), dunque la definizione alternativa di parità che abbiamo dato è ben posta. Le dimostrazioni di questo fatto e dell'equivalenza tra le due definizioni proposte di parità non sono per nulla ovvie, ma ci portebbero troppo lontano dai nostri scopi.

Enunciamo ora due risultati fondamentali nella teoria delle permutazioni, omettendone anche in questo caso la dimostrazione per le ragioni già esposte.

**Proposizione 7.9.4** *Ogni permutazione si scrive in modo unico come composizione di cicli disgiunti.*

**Proposizione 7.9.5** *La parità gode della seguente proprietà:*

$$p(s_1 \circ s_2) = p(s_1)p(s_2)$$

*per ogni $s_1, s_2 \in S_n$.*

Ora che abbiamo introdotto il concetto di permutazione possiamo dare una definizione alternativa di determinante.

**Definizione 7.9.6** Sia $A = (a_{ij})$ una matrice quadrata $n \times n$. Definiamo *determinante* di $A$ il numero:

$$\det(A) = \sum_{\sigma \in S_n} (-1)^{p(\sigma)} a_{1,\sigma(1)} \dots a_{n,\sigma(n)}$$

dove con $\sum_{\sigma \in S_n}$ si indica il fatto che stiamo facendo la somma degli elementi $a_{1,\sigma(1)} \dots a_{n,\sigma(n)}$ al variare di $\sigma$ tra tutte le permutazioni di $S_n$.

Verifichiamo che, nel caso di matrici $2 \times 2$ e $3 \times 3$, la nuova definizione di determinante corrisponda a quella vista precedentemente in 7.1.2.

Sia $A$ una matrice $2 \times 2$:

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

Secondo la nuova definizione abbiamo che il suo determinante è:

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

in quanto $S_2$, l'insieme delle permutazioni di due elementi, consiste soltanto dell'identità e della trasposizione $(1, 2)$. Possiamo subito constatare che tale espressione coincide con la formula per i determinanti di matrici $2 \times 2$ ricavata nel Paragrafo 7.2.

Sia $A$ una matrice $3 \times 3$:

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

L'insieme delle permutazioni di tre elementi è:

$$S_3 = \{\mathrm{id}, (1, 2), (2, 3), (1, 3), (3, 2, 1), (2, 3, 1)\}$$

con rispettive parità: $p(1, 2) = p(2, 3) = p(1, 3) = -1$, $p(\mathrm{id}) = p(3, 2, 1) = p(2, 3, 1) = 1$. Il determinante di $A$ è dunque dato da:

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

Introduciamo una notazione che ci sarà utile in seguito.

Sia $A = (a_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$ una matrice generica. La riga $i$-esima di $A$ è data da:

$$(a_{i1}, a_{i2}, \dots, a_{in}) = a_{i1}\mathbf{e}_1 + a_{i2}\mathbf{e}_2 + \dots + a_{in}\mathbf{e}_n$$

dove $\mathbf{e}_i = (0, \dots, 1, \dots, 0)$ è l'$i$-esimo vettore della base canonica di $\mathbb{R}^n$. Scriviamo la matrice $A$ come successione delle sue righe:

$$A = \left( \sum_{k_1=1}^{n} a_{1k_1}\mathbf{e}_{k_1}, \sum_{k_2=1}^{n} a_{2k_2}\mathbf{e}_{k_2}, \dots, \sum_{k_n=1}^{n} a_{nk_n}\mathbf{e}_{k_n} \right)$$

Per chiarire la nuova notazione appena introdotta, se $A = \left(\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right)$, scriviamo:

$$A = (a_{11}\mathbf{e}_1 + a_{12}\mathbf{e}_2,\ a_{21}\mathbf{e}_1 + a_{22}\mathbf{e}_2)$$

Vogliamo adesso mostrare che nel caso di una matrice $2 \times 2$ le proprietà che definiscono il determinante (vedi la Definizione 7.1.2) lo determinano in modo unico. Sia $A = (a_{11}\mathbf{e}_1 + a_{12}\mathbf{e}_2, a_{21}\mathbf{e}_1 + a_{22}\mathbf{e}_2)$. Grazie alla proprietà (1) della Definizione 7.1.2 abbiamo:

$$\begin{aligned}\det(A) &= \det(a_{11}\mathbf{e}_1 + a_{12}\mathbf{e}_2, a_{21}\mathbf{e}_1 + a_{22}\mathbf{e}_2) \\ &= \det(a_{11}\mathbf{e}_1, a_{21}\mathbf{e}_1 + a_{22}\mathbf{e}_2) + \det(a_{12}\mathbf{e}_2, a_{21}\mathbf{e}_1 + a_{22}\mathbf{e}_2) \\ &= \det(a_{11}\mathbf{e}_1, a_{21}\mathbf{e}_1) + \det(a_{11}\mathbf{e}_1, a_{22}\mathbf{e}_2) + \det(a_{12}\mathbf{e}_2, a_{21}\mathbf{e}_1) \\ &\quad + \det(a_{12}\mathbf{e}_2, a_{22}\mathbf{e}_2)\end{aligned}$$

Per la proprietà (2) della Definizione 7.1.2 possiamo poi scrivere:

$$\begin{aligned}\det(A) &= a_{11}a_{21}\det(\mathbf{e}_1, \mathbf{e}_1) + a_{11}a_{22}\det(\mathbf{e}_1, \mathbf{e}_2) \\ &\quad + a_{12}a_{21}\det(\mathbf{e}_2, \mathbf{e}_1) + a_{12}a_{22}\det(\mathbf{e}_2, \mathbf{e}_2)\end{aligned}$$

Per la Proposizione 7.1.4 abbiamo che: $\det(\mathbf{e}_1, \mathbf{e}_1) = \det(\mathbf{e}_2, \mathbf{e}_2) = 0$ (in quanto sono matrici con due righe uguali) e $\det(\mathbf{e}_1, \mathbf{e}_2) = -\det(\mathbf{e}_2, \mathbf{e}_1)$. Infine per la proprietà (3) della Definizione 7.1.2, $\det(\mathbf{e}_1, \mathbf{e}_2) = 1$. Da ciò consegue che

$$\det(A) = (a_{11}a_{22} - a_{12}a_{21})\det(\mathbf{e}_1, \mathbf{e}_2) = a_{11}a_{22} - a_{12}a_{21}$$

Questo dimostra che la funzione determinante definita in 7.1.2 deve necessariamente essere espressa dalla formula nella Definizione 7.9.6 e quindi tale funzione è unica.

Il procedimento che abbiamo descritto per le matrici $2 \times 2$ può essere replicato in modo identico nel caso di matrici $n \times n$, permettendoci di ottenere l'equivalenza delle due definizioni date di determinante. Vediamo più in dettaglio tutto ciò nella dimostrazione del seguente teorema che rappresenta il risultato più significativo di questa appendice.

**Teorema 7.9.7** *La funzione definita in* 7.1.2 *esiste, è unica ed è espressa dalla formula della Definizione* 7.9.6. *Le due Definizioni* 7.9.6 *e* 7.1.2 *di determinante sono pertanto equivalenti.*

**Dimostrazione** – Sia

$$A = \left(\sum_{k_1=1}^{n} a_{1k_1}\mathbf{e}_{k_1}, \sum_{k_2=1}^{n} a_{2k_2}\mathbf{e}_{k_2}, \dots, \sum_{k_n=1}^{n} a_{nk_n}\mathbf{e}_{k_n}\right)$$

una matrice $n \times n$ scritta secondo la notazione introdotta qui sopra (cioè come successione delle sue righe). Sia $\det(A)$ il numero definito in 7.1.2. In modo del tutto analogo a quanto abbiamo visto per il caso $2 \times 2$, per le proprietà (1) e (2) della Definizione 7.1.2 possiamo scrivere immediatamente

$$\det(A) = \sum_{1 \le k_1 \dots k_n \le n} a_{1,k_1} \dots a_{n,k_n} \det(\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \dots, \mathbf{e}_{k_n})$$

Osserviamo che la matrice $(\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \ldots, \mathbf{e}_{k_n})$ (scritta come successione di righe, secondo la nostra convenzione) ha, nella $i$-esima riga, 1 nella posizione $k_i$ e zero altrove. È dunque chiaro che, se ci sono valori ripetuti tra i $k_1, \ldots, k_n$, la matrice $(\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \ldots, \mathbf{e}_{k_n})$ ha due righe uguali e quindi per la proprietà (3) della Definizione 7.1.2 il suo determinante è zero. Dunque $\det(\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \ldots, \mathbf{e}_{k_n}) \neq 0$ solo se $k_1, \ldots, k_n$ sono tutti distinti e cioè la funzione $s$ definita da $s(i) = k_i$ è una permutazione di $\{1, \ldots, n\}$. A questo punto abbiamo

$$\det(\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \ldots, \mathbf{e}_{k_n}) = \begin{cases} 1 & \text{se} \quad p(s) = 1 \\ -1 & \text{se} \quad p(s) = -1 \end{cases}$$

perché è possibile riordinare le righe di $(\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \ldots, \mathbf{e}_{k_n})$ in modo da ottenere la matrice identità (che ha determinante 1) e per fare ciò si effettua un numero di scambi corrispondente alla parità di $s$. Abbiamo pertanto ottenuto che il determinante di $A$, come definito in 7.1.2, deve essere espresso dalla formula:

$$\det(A) = \sum_{s \in S_n} (-1)^{p(s)} a_{1,s(1)} \ldots a_{n,s(n)}$$

Ciò dimostra l'equivalenza tra le due definizioni, ma anche che il numero definito dalle quattro proprietà in 7.1.2 esiste ed è unico, cioè come tali proprietà lo determinino univocamente. □

**Corollario 7.9.8** *Il determinante è dato anche dalla formula:*

$$\det(A) = \sum_{\sigma \in S_n} (-1)^{p(\sigma)} a_{\sigma(1),1} \ldots a_{\sigma(n),n}$$

*(si noti che rispetto alla Definizione 7.9.6 le permutazioni sono effettuate sugli indici di riga e non di colonna). In particolare abbiamo che*

$$\det(A) = \det(A^T)$$

**Dimostrazione** – Per il teorema precedente abbiamo:

$$\det(A) = \sum_{\sigma \in S_n} (-1)^{p(\sigma)} a_{1,\sigma(1)} \ldots a_{n,\sigma(n)}$$

Osserviamo ora che

$$a_{\sigma(r),r} = a_{s,\sigma^{-1}(s)}$$

dove $s = \sigma(r)$, poiché $\sigma^{-1}(s) = \sigma^{-1}(\sigma(r)) = r$. Inoltre, poiché $s$ è una biezione di $\{1, \ldots, n\}$ in sé, abbiamo che

$$a_{1,\sigma(1)} \ldots a_{n,\sigma(n)} = a_{\sigma^{-1}(1),1} \ldots a_{\sigma^{-1}(n),n}$$

Pertanto:

$$\begin{aligned} \det(A) &= \sum_{\sigma \in S_n} (-1)^{p(\sigma)} a_{1,\sigma(1)} \ldots a_{n,\sigma(n)} \\ &= \sum_{\sigma \in S_n} (-1)^{p(\sigma)} a_{\sigma^{-1}(1),1} \ldots a_{\sigma^{-1}(n),n} \\ &= \sum_{\tau \in S_n} (-1)^{p(\tau)} a_{\tau(1),1} \ldots a_{\tau(n),n} \end{aligned}$$

in quanto, se $\sigma$ varia tra tutte le permutazioni in $S_n$, anche $\tau = \sigma^{-1}$ varia tra tutte le permutazioni in $S_n$ e inoltre $p(\tau) = p(\sigma)$. □

Grazie a questa nuova definizione e al teorema visto siamo in grado di dimostrare quanto abbiamo solamente enunciato nel testo relativamente alle procedure di calcolo del determinante.

Vogliamo ora ottenere una dimostrazione del Teorema 7.3.3, che ci fornisce un valido strumento per il calcolo del determinante. Prima di procedere alla dimostrazione della formula che appare nel Teorema 7.3.3, nota anche come formula per l' *espansione di Laplace* del determinante, abbiamo bisogno di un lemma tecnico.

**Lemma 7.9.9** *Se $B$ è una matrice quadrata del tipo:*

$$B = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ b_{21} & b_{22} & b_{23} & \dots & b_{2n} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ b_{n1} & b_{n2} & b_{n3} & \dots & b_{nn} \end{pmatrix}$$

*allora*

$$\det(B) = \det \begin{pmatrix} b_{22} & b_{23} & \dots & b_{2n} \\ \vdots & \vdots & \dots & \vdots \\ b_{n2} & b_{n3} & \dots & b_{nn} \end{pmatrix}$$

**Dimostrazione** – Si ha che $\det(B) = \sum_{\sigma \in S_n} (-1)^{p(\sigma)} b_{1,\sigma(1)} \dots b_{n,\sigma(n)}$. Poiché $b_{1,j} = 0$ per ogni $j \neq 1$, nella sommatoria le permutazioni $\sigma \in S_n$ tali che $\sigma(1) \neq 1$ danno un contributo nullo, per cui si ha che

$$\det(B) = \sum_{\sigma \in S_n} (-1)^{p(\sigma)} b_{1,1} b_{2,\sigma(2)} \dots b_{n,\sigma(n)}$$

L'insieme delle permutazioni $\sigma$ di $\{1, \dots, n\}$ che fissano 1 può essere visto come l'insieme di tutte le permutazioni $\tau$ dell'insieme $\{2, \dots, n\}$. Tenendo conto che $b_{1,1} = 1$, si ha:

$$\det(B) = \sum_{\tau \in S_{n-1}} (-1)^{p(\tau)} b_{2,\tau(2)} \dots b_{n,\tau(n)}$$

che è esattamente il determinante della matrice:

$$\begin{pmatrix} b_{22} & b_{23} & \dots & b_{2n} \\ \vdots & \vdots & \dots & \vdots \\ b_{n2} & b_{n3} & \dots & b_{nn} \end{pmatrix}$$

□

**Teorema di Laplace** *Sia $A$ una matrice quadrata $n \times n$ e sia $i \in \{1, \dots, n\}$ un indice fissato. Abbiamo allora che:*

$$\det(A) = a_{i,1}\Gamma_{i,1} + \dots a_{i,n}\Gamma_{i,n} \quad \text{e}$$

$$\det(A) = a_{1,i}\Gamma_{1,i} + \dots a_{n,i}\Gamma_{n,i}$$

*dove $\Gamma_{kl}$ denota il determinante della matrice ottenuta da $A$ sopprimendo la riga $k$-esima e la colonna $l$-esima, moltiplicato per $(-1)^{k+l}$.*

**Dimostrazione** – Diamo un cenno della dimostrazione della prima di queste proprietà, essendo la seconda del tutto analoga. Supponiamo $i = 1$. Il caso generale è più complicato da scrivere, ma non offre alcuna difficoltà concettuale aggiuntiva. Scriviamo $A$ come $A = (\sum_{j=1}^{n} a_{1j}\mathbf{e}_j, A_2, \dots, A_n)$. Per le proprietà (1) e (2) della Definizione (7.1.2) abbiamo che:

$$\det(A) = \sum_{j=1}^{n} a_{1j} \det(\mathbf{e}_j, A_2, \dots, A_n)$$

È sufficiente quindi calcolare il determinante di matrici del tipo:

$$M_j = \det(\mathbf{e}_j, A_2, \dots, A_n) = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2(j-1)} & a_{2j} & a_{2(j+1)} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{n(j-1)} & a_{nj} & a_{n(j+1)} & \dots & a_{2n} \end{pmatrix}$$

Osserviamo ora che, come conseguenza della proprietà (1) della Proposizione 7.1.4 e del Corollario 7.9.8, se si scambiano due colonne di una matrice il determinante cambia segno, quindi

$$\det(M_j) = (-1)^{j-1} \det \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ a_{2j} & a_{21} & a_{22} & \dots & a_{2(j-1)} & a_{2(j+1)} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \dots & \vdots & & & \\ a_{nj} & a_{n1} & a_{n2} & \dots & a_{n(j-1)} & a_{n(j+1)} & \dots & a_{nn} \end{pmatrix}$$

Osservando che $(-1)^{j-1} = (-1)^{1+j}$ e utilizzando il Lemma 7.9.9 si ha ora che:

$$\det(M_j) = (-1)^{1+j} \det \begin{pmatrix} a_{21} & a_{22} & \dots & a_{2(j-1)} & a_{2(j+1)} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{n(j-1)} & a_{n(j+1)} & \dots & a_{nn} \end{pmatrix}$$

dove la matrice è quella ottenuta da $A$ cancellando la prima riga e la $j$-esima colonna.

Si ottiene quindi:

$$\det(A) = \sum_{j=1}^{n} a_{1j}(-1)^{1+j} \det \begin{pmatrix} a_{21} & a_{22} & \dots & a_{2(j-1)} & a_{2(j+1)} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{n(j-1)} & a_{n(j+1)} & \dots & a_{nn} \end{pmatrix}$$

che è la formula voluta. □

Dimostriamo ora il Teorema di Binet (7.3.5). Poiché le due definizioni date di determinante sono tra di loro equivalenti, possiamo utilizzare quella che meglio si adatta alla dimostrazione che vogliamo fare.

**Teorema di Binet** *Siano $A = (a_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$ e $B = (b_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$ due matrici quadrate di ordine $n$. Allora:*

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

**Dimostrazione** – Sappiamo che $AB$ è la matrice il cui coefficiente di posto $i,j$ è $\sum_{k=1}^n a_{ik}b_{kj}$, dunque applichiamo la Definizione 7.9.6:

$$\begin{aligned}\det(AB) &= \sum_{\sigma\in S_n}(-1)^{p(\sigma)}\left(\sum_{k_1=1}^n a_{1k_1}b_{k_1,\sigma(1)}\right)\cdots\left(\sum_{k_n=1}^n a_{nk_n}b_{k_n,\sigma(n)}\right)\\ &= \sum_{1\le k_1,\dots,k_n\le n}\sum_{\sigma\in S_n}(-1)^{p(\sigma)}a_{1,k_1}b_{k_1,\sigma(1)}\cdots a_{n,k_n}b_{k_n,\sigma(n)}\\ &= \left(\sum_{1\le k_1,\dots,k_n\le n}a_{1,k_1}\cdots a_{n,k_n}\right)\left(\sum_{\sigma\in S_n}(-1)^{p(\sigma)}b_{k_1,\sigma(1)}\cdots b_{k_n,\sigma(n)}\right)\end{aligned}$$

riordinando i termini. Fissiamo $k_1,\dots,k_n$ e consideriamo il numero:

$$\sum_{\sigma\in S_n}(-1)^{p(\sigma)}b_{k_1,\sigma(1)}\cdots b_{k_n,\sigma(n)}$$

Si tratta del determinante della matrice avente, come righe, le righe $k_1,\dots,k_n$ della matrice $B$ cioè:

$$\sum_{\sigma\in S_n}(-1)^{p(\sigma)}b_{k_1,\sigma(1)}\cdots b_{k_n,\sigma(n)} = \det\left(\sum_{i_1}b_{k_1,i_1}\mathbf{e}_{i_1},\dots,\sum_{i_n}b_{k_1,i_n}\mathbf{e}_{i_n}\right)$$

utilizzando la notazione introdotta precedentemente. Per la proprietà (3) della Definizione 7.1.2 abbiamo che, se $k_1,\dots,k_n$ non sono tutti distinti, tale determinante è zero. Quindi, d'ora in avanti, sommiamo nell'espressione $\det(AB)$ soltanto i termini ove $k_1,\dots,k_n$ sono distinti. In questo caso, notiamo che la matrice $(\sum_{i_1}b_{k_1,i_1}\mathbf{e}_{i_1},\dots,\sum_{i_n}b_{k_1,i_n}\mathbf{e}_{i_n})$ è ottenuta a partire da $B$ con un numero di scambi di righe corrispondente alla parità della permutazione $\tau$ definita da $\tau(i)=k_i$. Pertanto:

$$\det\left(\sum_{i_1}b_{k_1,i_1}\mathbf{e}_{i_1},\dots,\sum_{i_n}b_{k_1,i_n}\mathbf{e}_{i_n}\right) = (-1)^{p(\tau)}\det(B)$$

Poiché $(k_1,\dots,k_n)=(\tau(1),\dots,\tau(n))$, sommare al variare di tutti i termini con $k_1,\dots,k_n$ distinti equivale a sommare al variare di $\tau$ tra tutte le permutazioni di $S_n$.

Abbiamo pertanto ottenuto:

$$\begin{aligned}\det(AB) &= \textstyle\sum_{1\le k_1,\dots,k_n\le n}(-1)^{p(\tau)}a_{1,k_1}\cdots a_{n,k_n}\det(B)\\ &= \textstyle\sum_{\tau\in S_n}(-1)^{p(\tau)}a_{1,\tau(1)}\cdots a_{n,\tau(n)}\det(B) = \det(A)\det(B) \qquad \square\end{aligned}$$

Se $A$ è una matrice quadrata indichiamo con $A_1,\dots,A_n$ le righe di $A$ e con $\tilde{\Gamma}_i$ il vettore colonna $(\Gamma_{i1},\dots,\Gamma_{kn})^T$, ove i numeri $\Gamma_{ij}$ sono definiti come nel Teorema di Laplace, e si dicono i *complementi algebrici* della matrice $A$. La prima formula del Teorema di Laplace si può allora scrivere come:

$$A_i\tilde{\Gamma}_i = \det(A), \qquad \text{per ogni } i=1,\dots,n \tag{7.2}$$

dove il prodotto è l'usuale prodotto righe per colonne.

Ha senso chiedersi quindi se il prodotto $A_i\widetilde{\Gamma}_j$ abbia qualche significato anche quando $i \neq j$. La risposta è data dalla seguente proposizione.

**Proposizione 7.9.12** *Sia la notazione come sopra. Allora:*

$$A_i\widetilde{\Gamma}_j = 0 \quad \textit{per ogni } i,j = 1,\dots,n, \;\; i \neq j \tag{7.3}$$

**Dimostrazione** – Per dimostrare la (7.3) consideriamo la matrice $A' = (a'_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}}$ ottenuta da $A$ sostituendo la $j$-esima riga con la $i$-esima riga. Si ha quindi che $a'_{rs} = a_{rs}$ per ogni $r,s = 1,\dots,n$, con $r \neq j$, e $a'_{js} = a_{is}$ per ogni $s = 1,\dots,n$.

Siano $\Gamma'_{rs}$ i complementi algebrici di $A'$. Poiché nel calcolo di $\Gamma'_{js}$ non interviene la $j$-esima riga di $A'$, si ha che $\Gamma'_{js} = \Gamma_{js}$ per ogni $s = 1,\dots,n$. Calcoliamo ora il determinante di $A'$ utilizzando il Teorema di Laplace e sviluppando secondo la $j$-esima riga::

$$\det(A') = a'_{j,1}\Gamma'_{j,1} + \dots + a'_{j,n}\Gamma'_{j,n} = a_{i,1}\Gamma_{j,1} + \dots + a_{i,n}\Gamma_{j,n}$$

D'altra parte la matrice $A'$ ha due righe uguali (la $i$-esima e la $j$-esima), quindi per la proprietà (3) della Definizione 7.1.2 il determinante di $A'$ è uguale a zero. Questo dimostra proprio la formula (7.3), cioè si ha che:

$$a_{i,1}\Gamma_{j,1} + \dots + a_{i,n}\Gamma_{j,n} = 0 \quad \text{per ogni } i,j = 1,\dots,n, \;\; i \neq j$$ □

Dimostriamo ora la formula (7.1) per il calcolo della matrice inversa di una matrice quadrata $A$ con determinante diverso da zero. Sia $\det(A_{ij})$ il determinante della matrice ottenuta da $A$ rimuovendo la $i$-esima riga e la $j$-esima colonna e consideriamo la matrice $B$ i cui coefficienti sono definiti da:

$$(B)_{ij} = \frac{1}{\det(A)}(-1)^{i+j}\det(A_{ji}) \tag{7.4}$$

Osserviamo che la $j$-esima colonna della matrice $B$ è $\frac{1}{\det(A)}\widetilde{\Gamma}_j$. Effettuiamo ora il prodotto righe per colonne $AB$. L'elemento di posto $i,j$ è $A_iB_j = \frac{1}{\det(A)}A_i\widetilde{\Gamma}_j$ e, per le formule (7.2) e (7.3), tale valore è $\frac{1}{\det(A)}\det(A) = 1$ se $i = j$ e $0$ se $i \neq j$.

Quindi $AB = I$. Analogamente, partendo dal Teorema di Laplace e sviluppando il determinante secondo le colonne, si dimostra che $BA = I$, quindi $B$ è l'inversa di $A$.

Dimostriamo infine la correttezza del metodo descritto nel Paragrafo 7.5 per calcolare l'inversa di una matrice invertibile utilizzando l'algoritmo di Gauss.

Sia $A = (a_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}}$ una matrice quadrata invertibile e sia $B = (b_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}}$ la sua inversa. Si ha che $AB = I$, ove $I$ è la matrice identità di ordine $n$. Siano $A_1,\dots,A_n$ i vettori riga di $A$ e siano $\widetilde{B}_1,\dots,\widetilde{B}_n$ i vettori colonna di $B$. I coefficienti $b_{11},\dots,b_{n1}$ di $\widetilde{B}_1$ soddisfano le relazioni: $A_1\widetilde{B}_1 = (AB)_{11} = 1$, $A_2\widetilde{B}_1 = (AB)_{21} = 0,\dots,A_n\widetilde{B}_1 = (AB)_{n1} = 0$ (ove i prodotti si intendono righe per colonne), cioè costituiscono una soluzione del sistema lineare associato alla matrice:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & 1 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 \end{array}\right) \tag{7.5}$$

Poiché $A$ è invertibile, per il Teorema 7.6.1 il sistema ammette un'unica soluzione, e quindi, risolvendo tale sistema, determiniamo in modo univoco gli elementi della colonna $\widetilde{B}_1$. Come descritto nel Paragrafo 7.5, tramite operazioni elementari sulle righe, è possibile ottenere a sinistra la matrice identità, cioè si ottiene una matrice del tipo

$$\left(\begin{array}{cccc|c} 1 & 0 & \dots & 0 & c_{11} \\ 0 & 1 & \dots & 0 & c_{21} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 1 & c_{n1} \end{array}\right) \tag{7.6}$$

e, poiché i due sistemi associati alle matrici in (7.5) e (7.6) hanno le stesse soluzioni, deve essere $b_{j1} = c_{j1}$ per ogni $j = 1, \dots, n$, cioè in (7.6) la colonna a sinistra è proprio $\widetilde{B}_1$.

Procediamo in modo analogo per la generica colonna $\widetilde{B}_i$, i cui coefficienti sono soluzioni del sistema lineare:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 \end{array}\right) \tag{7.7}$$

Per risolvere questo sistema, possiamo effettuare sulle righe di $A$ *esattamente le stesse operazioni elementari* che nel caso di $\widetilde{B}_1$, e si ottiene così una matrice del tipo:

$$\left(\begin{array}{cccc|c} 1 & 0 & \dots & 0 & c_{1i} \\ 0 & 1 & \dots & 0 & c_{2i} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 1 & c_{ni} \end{array}\right) \tag{7.8}$$

Di nuovo, poiché i due sistemi associati alle matrici in (7.7) e (7.8) hanno le stesse soluzioni, deve essere $b_{ji} = c_{ji}$ per ogni $j = 1, \dots, n$, cioè in (7.8) la colonna a sinistra è proprio $\widetilde{B}_i$. Visto che dobbiamo risolvere $n$ sistemi lineari aventi tutti la stessa matrice dei coefficienti, possiamo risolverli in contemporanea considerando la matrice

$$A|I$$

Per quanto detto, dopo aver effettuato le operazioni elementari sulle righe necessarie per ridurre $A$ alla matrice identità, si ottiene una matrice del tipo:

$$\left(\begin{array}{cccc|cccc} 1 & 0 & \dots & 0 & c_{11} & c_{12} & \dots & c_{1n} \\ 0 & 1 & \dots & 0 & c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & & \vdots & \vdots & & & \\ 0 & 0 & \dots & 1 & c_{n1} & c_{n2} & \dots & c_{nn} \end{array}\right)$$

ove $c_{ji} = b_{ji}$ per ogni $i, j = 1, \ldots, n$, cioè la matrice a sinistra è proprio l'inversa $B$ di $A$.

Facciamo un sunto delle proprietà salienti del determinante di una matrice che provengono dalla definizione o abbiamo dimostrato nel testo.

### Proprietà del determinante

Il determinante gode delle seguenti proprietà:

- Il determinante di una matrice $A$ è nullo se $A$ ha una riga o una colonna nulla.
- Il determinante di una matrice $A$ è nullo se e solo se la matrice $A$ ha una riga (o una colonna) che è combinazione lineare delle altre.
- Se la matrice $A'$ si ottiene dalla matrice $A$ scambiando due righe (o due colonne), il determinante di $A'$ è l'opposto del determinante di $A$.
- Se la matrice $A'$ si ottiene dalla matrice $A$ moltiplicando una riga (o una colonna) per lo scalare $\lambda$, il determinante di $A'$ è il prodotto di $\lambda$ per il determinante di $A$.
- Se la matrice $A'$ si ottiene dalla matrice $A$ aggiungendo a una riga (o una colonna) una combinazione lineare delle altre, il determinante di $A'$ è uguale al determinante di $A$.
- Il determinante della matrice identità di ordine $n$ è 1.
- Il determinante di una matrice diagonale è il prodotto degli elementi sulla diagonale.
- $\det(A) = \det(A^T)$, per ogni matrice quadrata $A$.
- $\det(AB) = \det(BA) = \det(A)\det(B)$ per ogni coppia di matrici quadrate $A$ e $B$ dello stesso ordine (ma attenzione, in generale $AB \neq BA$!)
- $\det(A^{-1}) = \det(A)^{-1}$, per ogni matrice quadrata invertibile $A$.

# 8 Cambio di base

In questo capitolo vogliamo affrontare uno degli argomenti più tecnici della teoria, cioè il cambio di base all'interno di uno spazio vettoriale. Cercheremo anche di capire come cambia la matrice associata a un'applicazione lineare se cambiamo le basi nel dominio e nel codominio.

## 8.1 APPLICAZIONI LINEARI E MATRICI

Come abbiamo avuto modo di vedere nel Capitolo 5, fissate le basi canoniche in $\mathbb{R}^n$ e $\mathbb{R}^m$, possiamo identificare le applicazioni lineari $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ e le matrici $m \times n$ in $\mathrm{M}_{m,n}(\mathbb{R})$

$$\begin{array}{ccc}\{\text{applicazioni lineari}\, \mathbb{R}^n \longrightarrow \mathbb{R}^m\} & \leftrightarrow & \mathrm{M}_{m,n}(\mathbb{R}) \\ F : \mathbb{R}^n \longrightarrow \mathbb{R}^m & \mapsto & (F(\mathbf{e}_1), \dots, F(\mathbf{e}_n))\end{array}$$

La matrice $(F(\mathbf{e}_1), \dots, F(\mathbf{e}_n))$ associata all'applicazione lineare $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ in tale corrispondenza biunivoca ha per colonne le immagini tramite $F$ dei vettori della base canonica di $\mathbb{R}^n$. Vediamo un esempio.

Consideriamo l'applicazione $F : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, $F(\mathbf{e}_1) = \mathbf{e}_1 - \mathbf{e}_2$, $F(\mathbf{e}_2) = 3\mathbf{e}_2$. Questa applicazione è associata, fissando la base canonica nel dominio e codominio, alla matrice:

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

Sappiamo bene che la scelta della base canonica per rappresentare i vettori in $\mathbb{R}^n$ è arbitraria, pur essendo estremamente conveniente. Per esempio abbiamo già visto che un vettore espresso in due basi diverse ha ovviamente coordinate diverse. Per il momento l'utilizzo di una base diversa da quella canonica per rappresentare i vettori sembra una fatica inutile, tuttavia, come vedremo nel capitolo successivo, ci fornisce la chiave per capire i concetti di autovalori e autovettori, che sono di importanza fondamentale non solo nella teoria dell'algebra lineare, ma anche e soprattutto nelle applicazioni.

Vogliamo ora generalizzare il concetto di matrice associata a un'applicazione lineare. Cominciamo con alcune osservazioni che richiamano la teoria svolta nel Capitolo 5.

Siano $V$ e $W$ due spazi vettoriali di dimensione finita e siano $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ e $\mathcal{B}' = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ due basi ordinate fissate di $V$ e $W$ rispettivamente. Se $F : V \to W$ è un'applicazione lineare, per il Teorema 5.1.7, sappiamo che $F$ è univocamente determinata da $F(\mathbf{v}_1), \dots, F(\mathbf{v}_n)$.

Siano:

$$\begin{aligned} F(\mathbf{v}_1) &= a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + \cdots + a_{m1}\mathbf{w}_m \\ F(\mathbf{v}_2) &= a_{12}\mathbf{w}_1 + a_{22}\mathbf{w}_2 + \cdots + a_{m2}\mathbf{w}_m \\ &\vdots \\ F(\mathbf{v}_n) &= a_{1n}\mathbf{w}_1 + a_{2n}\mathbf{w}_2 + \cdots + a_{mn}\mathbf{w}_m \end{aligned} \tag{8.1}$$

e sia $\mathbf{v} = x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n$ un generico vettore di $V$. Analogamente a quanto fatto nel Capitolo 5, vogliamo determinare $F(\mathbf{v})$.

Si ha che:

$$\begin{aligned} F(\mathbf{v}) &= F(x_1\mathbf{v}_1 + x_2\mathbf{v}_2 \cdots + x_n\mathbf{v}_n) \\ &= x_1F(\mathbf{v}_1) + x_2F(\mathbf{v}_2) + \cdots + x_nF(\mathbf{v}_n) \\ &= x_1(a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + \cdots + a_{m1}\mathbf{w}_m) \\ &\quad + x_2(a_{12}\mathbf{w}_1 + a_{22}\mathbf{w}_2 + \cdots + a_{m2}\mathbf{w}_m) + \dots \\ &\quad + x_n(a_{1n}\mathbf{w}_1 + a_{2n}\mathbf{w}_2 + \cdots + a_{mn}\mathbf{w}_m) \\ &= (a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n)\mathbf{w}_1 \\ &\quad + (a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n)\mathbf{w}_2 + \dots \\ &\quad + (a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n)\mathbf{w}_m \end{aligned}$$

Quindi, se le coordinate di $\mathbf{v}$ rispetto alla base ordinata $\mathcal{B}$ sono $(\mathbf{v})_{\mathcal{B}} = (x_1, \dots, x_n)$, abbiamo che le coordinate di $F(\mathbf{v})$ rispetto alla base ordinata $\mathcal{B}'$ sono:

$$(F(\mathbf{v}))_{\mathcal{B}'} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix} = A \cdot (\mathbf{v})_{\mathcal{B}}$$

dove $A$ è la matrice definita in (8.1), che ha per colonne le coordinate dei vettori $F(\mathbf{v}_1), \dots, F(\mathbf{v}_n)$ rispetto alla base ordinata $\mathcal{B}' = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ e $A \cdot (\mathbf{v})_{\mathcal{B}}$ denota il prodotto righe per colonne della matrice $A$ per il vettore $(\mathbf{v})_{\mathcal{B}}$ delle coordinate di $\mathbf{v}$ rispetto alla base $\mathcal{B}$.

Siamo dunque in grado di associare a $F$ una matrice $A$ $m \times n$ una volta fissate *basi ordinate arbitrarie* $\mathcal{B}$ e $\mathcal{B}'$ nel dominio e nel codominio. Se $V = \mathbb{R}^n$, $W = \mathbb{R}^m$ e fissiamo come $\mathcal{B}$ e $\mathcal{B}'$ le basi canoniche del dominio e del codominio si verifica facilmente che la matrice $A$ è proprio la matrice associata a $F$ definita nei capitoli precedenti e richiamata all'inizio di questo capitolo. Tale matrice ha per colonne $F(\mathbf{e}_1), \dots, F(\mathbf{e}_n)$ e cioè le coordinate delle immagini dei vettori della base canonica rispetto alla base canonica stessa (in $\mathbb{R}^n$ qualora non specifichiamo rispetto quale base consideriamo le coordinate dei vettori, supponiamo di considerare la base canonica).

Vogliamo ora formalizzare nella seguente definizione quanto abbiamo osservato.

Facciamo una precisazione importante. Mentre fino a questo momento abbiamo utilizzato indifferentemente righe o colonne per indicare le coordinate, in questo capitolo sarà necessaria una maggiore precisione, in quanto le coordinate di un vettore rispetto a una base data formeranno un vettore colonna che dovrà poi essere moltiplicato (righe per colonne) per una matrice. D'ora in avanti denoteremo quindi le coordinate di un vettore rispetto a una base data tramite un vettore *colonna*.

**Definizione 8.1.1** Sia $F : V \longrightarrow W$ un'applicazione lineare, dove $V$ e $W$ sono spazi vettoriali e siano $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, $\mathcal{B}' = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ rispettivamente una base ordinata del dominio e una base ordinata del codominio.

Siano

$$(\mathbf{v})_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad (F(\mathbf{v}))_{\mathcal{B}'} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

le coordinate di un vettore $\mathbf{v}$ di $V$ e le coordinate della sua immagine tramite $F$, rispettivamente.

La matrice $A_{\mathcal{B},\mathcal{B}'}$ associata a $F$ rispetto alle basi ordinate $\mathcal{B}$ e $\mathcal{B}'$ è per definizione la matrice $m \times n$ tale che:

$$(F(\mathbf{v}))_{\mathcal{B}'} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = A_{\mathcal{B},\mathcal{B}'} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Dalle osservazioni precedenti abbiamo che la $i$-esima colonna di $A_{\mathcal{B},\mathcal{B}'}$ è costituita dalle coordinate di $F(\mathbf{v}_i)$ rispetto alla base $\mathcal{B}'$.

Nel caso in cui l'applicazione lineare sia del tipo $F : V \longrightarrow V$ e si scelga la stessa base ordinata $\mathcal{B}$ nel dominio e nel codominio, la matrice associata a $F$ rispetto alla base $\mathcal{B}$ verrà indicata semplicemente con $A_{\mathcal{B}}$.

Torniamo all'esempio visto sopra. Come base ordinata di $\mathbb{R}^2$ scegliamo $\mathcal{B} = \{\mathbf{v}_1 = \mathbf{e}_1 - \mathbf{e}_2, \mathbf{v}_2 = 3\mathbf{e}_2\}$ e ci poniamo le seguenti domande:

1. quali sono le coordinate dei vettori $\mathbf{v}_1$ e $\mathbf{v}_2$ rispetto alla nuova base $\mathcal{B}$?

2. Come possiamo scrivere la matrice associata a $F$ rispetto alla base canonica del dominio e alla base $\mathcal{B}$ del codominio?

La risposta alla prima domanda è ovvia. I vettori $\mathbf{v}_1$ e $\mathbf{v}_2$ hanno rispettivamente coordinate $(1,0)^T$ e $(0,1)^T$ ($T$ indica la trasposta e cioè il fatto che le coordinate rappresentano un vettore colonna). Infatti $\mathbf{v}_1 = 1 \cdot \mathbf{v}_1 + 0 \cdot \mathbf{v}_2$, $\mathbf{v}_2 = 0 \cdot \mathbf{v}_1 + 1 \cdot \mathbf{v}_2$.

Anche la risposta alla seconda domanda è abbastanza semplice. Abbiamo infatti già visto che modificare la base che noi scegliamo per rappresentare i vettori all'interno di uno spazio vettoriale, non cambia i vettori, ma soltanto il modo di scriverli, cioè le loro coordinate. Nell'esempio precedente i vettori $\mathbf{v}_1$

e $\mathbf{v}_2$ sono sempre gli stessi, quello che cambia, passando dalla base canonica alla base ordinata $\mathcal{B}$, sono le loro coordinate che passano da $(\mathbf{v}_1)_\mathcal{C} = (1,-1)^T$, $(\mathbf{v}_2)_\mathcal{C} = (0,3)^T$ a $(\mathbf{v}_1)_\mathcal{B} = (1,0)^T$, $(\mathbf{v}_2)_\mathcal{B} = (0,1)^T$. La stessa cosa avviene per le applicazioni lineari, dove il concetto di coordinate è sostituito dal concetto di matrice associata all'applicazione lineare rispetto a due basi ordinate date nel dominio e nel codominio. Vediamo un esempio concreto.

L'applicazione lineare che vogliamo considerare è $F : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, tale che $F(\mathbf{e}_1) = \mathbf{v}_1$, $F(\mathbf{e}_2) = \mathbf{v}_2$. Vediamo ora di rappresentare $F$ usando le coordinate rispetto alla base canonica $\mathcal{C} = \{\mathbf{e}_1, \mathbf{e}_2\}$ nel dominio e le coordinate rispetto alla base ordinata $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ nel codominio. Si ha che $(F(\mathbf{e}_1))_\mathcal{B} = (1,0)^T$, $(F(\mathbf{e}_2))_\mathcal{B} = (0,1)^T$, dove mettiamo un indice per ricordarci che si tratta di coordinate rispetto a una base, che non è necessariamente quella canonica. Quindi abbiamo che la matrice associata a $F$ rispetto alle basi $\mathcal{C}$ del dominio e $\mathcal{B}$ del codominio è

$$A_{\mathcal{C},\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Infatti facendo il prodotto righe per colonne vediamo che:

$$A_{\mathcal{C},\mathcal{B}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = (F(\mathbf{e}_1))_\mathcal{B}$$

$$A_{\mathcal{C},\mathcal{B}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = (F(\mathbf{e}_2))_\mathcal{B}$$

Dunque la matrice associata a $F$ rispetto alle basi $\mathcal{C}$ nel dominio e $\mathcal{B}$ nel codominio è proprio $A_{\mathcal{C},\mathcal{B}}$, cioè la matrice identità.

Quanto abbiamo detto si può facilmente generalizzare.

**Proposizione 8.1.2** *Sia $F : V \longrightarrow V$ un'applicazione lineare tale che $F(\mathbf{v}_1) = \mathbf{w}_1, \dots, F(\mathbf{v}_n) = \mathbf{w}_n$ dove $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, $\mathcal{B}' = \{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ sono due basi ordinate di $V$. Allora la matrice associata a $F$ rispetto alla base ordinata $\mathcal{B}$ nel dominio e alla base ordinata $\mathcal{B}'$ nel codominio è la matrice identità.*

La dimostrazione è un facile esercizio e ripercorre il ragionamento precedente.

## 8.2 L'IDENTITÀ

Da questo paragrafo in poi la nostra attenzione sarà esclusivamente volta a esaminare il caso di $V = \mathbb{R}^n$ per acquistare concretezza, anche se tutto quello che diremo può facilmente essere generalizzato al caso di uno spazio vettoriale generico, purché di dimensione finita [1]. La nostra scelta di non trattare il cambio di base in questa generalità è dettata soltanto da questioni espositive e ha il solo scopo di aumentare la chiarezza.

---

[1] Uno spazio vettoriale $V$ di dimensione finita può sempre essere identificato con $\mathbb{R}^n$ *purché si fissi una base*. Vogliamo tuttavia sottolineare che *non* è questo il caso, cioè tutto quello che diremo d'ora in poi può essere generalizzato a uno spazio vettoriale $V$ qualsiasi di dimensione finita $n$, senza identificarlo con $\mathbb{R}^n$.

Ci poniamo ora una domanda in un qualche modo inversa a quanto abbiamo stabilito nell'ultima proposizione del paragrafo precedente.

*Qual è la matrice associata all'applicazione identità* id : $\mathbb{R}^n \longrightarrow \mathbb{R}^n$ *rispetto alle basi* $\mathcal{B} = (\mathbf{v}_1 \dots \mathbf{v}_n)$, $\mathcal{B}' = (\mathbf{w}_1 \dots \mathbf{w}_n)$ *rispettivamente nel dominio e nel codominio?*

Sappiamo che la matrice identità è associata all'applicazione identità qualora fissiamo le basi canoniche nel dominio e nel codominio, tuttavia abbiamo già capito dall'esempio precedente che, cambiando la base, può cambiare radicalmente l'aspetto della matrice associata alla stessa applicazione lineare.

Vediamo un esempio semplice per aiutarci nella comprensione.

---

**Esempio 8.2.1**

Consideriamo l'applicazione identità id : $\mathbb{R}^2 \longrightarrow \mathbb{R}^2$ e supponiamo di fissare la base ordinata $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ nel dominio e la base canonica $\mathcal{C}$ nel codominio, con $\mathbf{v}_1 = 2\mathbf{e}_1, \mathbf{v}_2 = \mathbf{e}_1 + \mathbf{e}_2$.

L'applicazione identità si comporta sempre allo stesso modo anche se cambiamo il modo di rappresentarla: id manderà comunque un vettore in se stesso. Vediamo che cosa accade:

$$\mathrm{id}(\mathbf{v}_1) = \mathbf{v}_1 = 2\mathbf{e}_1 \qquad \mathrm{id}(\mathbf{v}_2) = \mathbf{v}_2 = \mathbf{e}_1 + \mathbf{e}_2$$

Scriviamo le coordinate dei vettori rispetto alla base canonica:

$$(\mathrm{id}(\mathbf{v}_1))_{\mathcal{C}} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \qquad (\mathrm{id}(\mathbf{v}_2))_{\mathcal{C}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Dunque la matrice associata all'identità rispetto alla base $\mathcal{B}$ nel dominio e alla base canonica $\mathcal{C}$ nel codominio è:

$$I_{\mathcal{B},\mathcal{C}} = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$$

Ora chiediamoci invece che cosa accade se vogliamo rappresentare l'applicazione identità utilizzando la base canonica $\mathcal{C}$ nel dominio e la base $\mathcal{B}$ nel codominio. L'identità associa sempre a ogni vettore se stesso, il problema è capire quali sono le coordinate giuste.

$$\mathrm{id}(\mathbf{e}_1) = \mathbf{e}_1 = (1/2)\mathbf{v}_1 \qquad \mathrm{id}(\mathbf{e}_2) = \mathbf{e}_2 = -(1/2)\mathbf{v}_1 + \mathbf{v}_2$$

Quindi:

$$(\mathrm{id}(\mathbf{e}_1))_{\mathcal{B}} = \begin{pmatrix} 1/2 \\ 0 \end{pmatrix} \qquad (\mathrm{id}(\mathbf{e}_2))_{\mathcal{B}} = \begin{pmatrix} -1/2 \\ 1 \end{pmatrix}$$

Dunque la matrice associata all'identità rispetto alla base canonica $\mathcal{C}$ del dominio e alla base $\mathcal{B}$ del codominio è:

$$I_{\mathcal{C},\mathcal{B}} = \begin{pmatrix} 1/2 & -1/2 \\ 0 & 1 \end{pmatrix}$$

In questo esempio molto semplice è stato possibile calcolare facilmente le coordinate di $\mathbf{e}_1$ ed $\mathbf{e}_2$ rispetto alla base ordinata $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$, in generale ciò non è sempre così facile. Tuttavia nell'esempio si nota che $I_{\mathcal{C},\mathcal{B}} = I^{-1}_{\mathcal{B},\mathcal{C}}$. Dunque le coordinate dei vettori $\mathbf{e}_1$ ed $\mathbf{e}_2$ rispetto alla base ordinata $\mathcal{B}$ si possono leggere come le colonne della matrice $I^{-1}_{\mathcal{B},\mathcal{C}}$. Si ricordi che la matrice $I_{\mathcal{B},\mathcal{C}}$ è facilmente calcolabile, in quanto ha per colonne le coordinate dei vettori $\mathbf{v}_1$, $\mathbf{v}_2$ rispetto alla base canonica.

---

Questo fenomeno è vero in generale: otteniamo le coordinate di un vettore $\mathbf{v}$ rispetto a una base ordinata $\mathcal{B}$ attraverso la moltiplicazione $(\mathbf{v})_{\mathcal{B}} = I^{-1}_{\mathcal{B},\mathcal{C}}(\mathbf{v})_{\mathcal{C}}$, ma è necessario ancora un po' di lavoro per dimostrarlo. Ci occupiamo prima di un altro caso particolare.

**Proposizione 8.2.2** *Sia $\mathcal{B}$ una base ordinata di $\mathbb{R}^n$; allora la matrice $I_\mathcal{B}$ associata all'applicazione identità* id : $\mathbb{R}^n \to \mathbb{R}^n$ *rispetto alla base $\mathcal{B}$ nel dominio e nel codominio è la matrice identità.*

**Dimostrazione** – La dimostrazione è immediata, poiché è una conseguenza della Proposizione 8.1.2. Tuttavia si vede anche direttamente, perché se $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, allora nella colonna $i$-esima di $I_\mathcal{B}$ ci sono le coordinate di $\mathrm{id}(\mathbf{v}_i) = \mathbf{v}_i$ rispetto alla base $\mathcal{B}$, cioè il vettore che ha zero ovunque e 1 al posto $i$-esimo. □

♦ **Osservazione 8.2.3** In generale, se abbiamo la composizione di applicazioni lineari $V \xrightarrow{F} W \xrightarrow{G} Z$, e indichiamo con $A_F$ la matrice associata a $F$ e con $A_G$ la matrice associata a $G$ (rispetto a una qualunque terna di basi fissate per $V$, $W$ e $Z$), allora alla composizione $G \circ F$ è associata la matrice $A_G \cdot A_F$, dove il prodotto si intende come il prodotto usuale righe per colonne. La dimostrazione di questo fatto è del tutto analoga alla dimostrazione del Corollario 5.3.4.

Possiamo ora formalizzare l'uguaglianza $I_{\mathcal{C},\mathcal{B}} = I_{\mathcal{B},\mathcal{C}}^{-1}$ che abbiamo verificato nel caso particolare dell'Esempio 8.2.1 insieme ad altri fatti essenziali per il calcolo delle coordinate di un vettore rispetto a una base data.

**Teorema 8.2.4** *Sia* id : $\mathbb{R}^n \longrightarrow \mathbb{R}^n$ *l'applicazione lineare identità che associa a ogni vettore se stesso. Sia $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ una base ordinata per $\mathbb{R}^n$. Allora abbiamo che:*

1. *la matrice associata a* id *rispetto alla base $\mathcal{B}$ nel dominio e alla base canonica $\mathcal{C}$ nel codominio è:*
$$I_{\mathcal{B},\mathcal{C}} = ((\mathbf{v}_1)_\mathcal{C}, \dots, (\mathbf{v}_n)_\mathcal{C})$$
*dove $(\mathbf{v}_i)_\mathcal{C}$ sono le coordinate del vettore $\mathbf{v}_i$ rispetto alla base canonica;*

2. *la matrice associata a* id *rispetto alla base canonica $\mathcal{C}$ nel dominio e $\mathcal{B}$ nel codominio è: $I_{\mathcal{B},\mathcal{C}}^{-1}$.*

**Dimostrazione** – 1. Si tratta di quanto visto nell'Esempio 8.2.1. Infatti $(\mathrm{id}(\mathbf{v}_1))_\mathcal{C} = (\mathbf{v}_1)_\mathcal{C}$ sono proprio le coordinate del vettore $\mathbf{v}_1$ rispetto alla base canonica. Lo stesso per $(\mathrm{id}(\mathbf{v}_2))_\mathcal{C}, \dots, (\mathrm{id}(\mathbf{v}_n))_\mathcal{C}$.

2 Per dimostrare il secondo punto, dimostriamo che $I_{\mathcal{C},\mathcal{B}}$ è la matrice inversa di $I_{\mathcal{B},\mathcal{C}}$, cioè che $I_{\mathcal{B},\mathcal{C}} I_{\mathcal{C},\mathcal{B}} = I_{\mathcal{C},\mathcal{B}} I_{\mathcal{B},\mathcal{C}} = I$, dove $I$ è la matrice identità.

Consideriamo la composizione dell'identità con se stessa, fissando però in $\mathbb{R}^n$ basi diverse, come indicato nel seguente diagramma:

$$\begin{array}{ccccc} \mathbb{R}^n & \xrightarrow{\mathrm{id}} & \mathbb{R}^n & \xrightarrow{\mathrm{id}} & \mathbb{R}^n \\ \mathcal{B} & & \mathcal{C} & & \mathcal{B} \end{array}$$

La funzione composta id ∘ id = id è ancora l'identità e, se consideriamo le matrici associate alle applicazioni in gioco, per l'Osservazione 8.2.3 otteniamo:

$I_{\mathcal{C},\mathcal{B}} I_{\mathcal{B},\mathcal{C}} = I_{\mathcal{B}}$. Ora, per la Proposizione 8.2.2 si ha che $I_{\mathcal{B}} = I$, quindi $I_{\mathcal{C},\mathcal{B}} I_{\mathcal{B},\mathcal{C}} = I$. Analogamente, considerando il diagramma:

$$\begin{array}{ccccc} \mathbb{R}^n & \xrightarrow{\text{id}} & \mathbb{R}^n & \xrightarrow{\text{id}} & \mathbb{R}^n \\ \mathcal{C} & & \mathcal{B} & & \mathcal{C} \end{array}$$

si ottiene che $I_{\mathcal{B},\mathcal{C}} I_{\mathcal{C},\mathcal{B}} = I$. Questo conclude la dimostrazione. □

Questo teorema risponde anche a un'altra domanda che ci eravamo posti in precedenza e cioè: come possiamo scrivere le coordinate di un vettore $\mathbf{v}$ rispetto a una data base ordinata $\mathcal{B}$?

Fino a questo momento abbiamo risposto con un calcolo molto esplicito caso per caso. Ora però possiamo enunciare un corollario che contiene la risposta in generale.

**Corollario 8.2.5** *Sia $\mathcal{B}$ una base ordinata di $\mathbb{R}^n$ e sia $\mathbf{v}$ un vettore fissato di $\mathbb{R}^n$. Allora le coordinate di $\mathbf{v}$ rispetto alla base $\mathcal{B}$ sono date da:*

$$(\mathbf{v})_{\mathcal{B}} = I_{\mathcal{B},\mathcal{C}}^{-1} (\mathbf{v})_{\mathcal{C}}$$

**Dimostrazione** – Si ha che $\mathbf{v} = \text{id}(\mathbf{v})$, quindi basta scegliere la base canonica $\mathcal{C}$ nel dominio e la base $\mathcal{B}$ nel codominio. A questo punto per il Teorema 8.2.4 sappiamo che, rispetto a queste basi, l'applicazione lineare identità è rappresentata dalla matrice $I_{\mathcal{C},\mathcal{B}} = I_{\mathcal{B},\mathcal{C}}^{-1}$. □

Vediamo un esempio per illustrare questi concetti.

---

**Esempio 8.2.6**

Consideriamo la base ordinata $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ di $\mathbb{R}^2$, dove $\mathbf{v}_1 = -\mathbf{e}_1 + 2\mathbf{e}_2$, $\mathbf{v}_2 = -\mathbf{e}_2$, e vogliamo trovare le coordinate del vettore $\mathbf{v} = -\mathbf{e}_1 + 3\mathbf{e}_2$ rispetto alla base $\mathcal{B}$. Sappiamo che $(\mathbf{v})_{\mathcal{C}} = (-1, 3)$ e $I_{\mathcal{B},\mathcal{C}} = \left(\begin{smallmatrix} -1 & 0 \\ 2 & -1 \end{smallmatrix}\right)$. Con un facile calcolo, si ottiene che $I_{\mathcal{B},\mathcal{C}}^{-1} = \left(\begin{smallmatrix} -1 & 0 \\ -2 & -1 \end{smallmatrix}\right)$, quindi le coordinate sono:

$$(\mathbf{v})_{\mathcal{B}} = I_{\mathcal{B},\mathcal{C}}^{-1} \cdot (\mathbf{v})_{\mathcal{C}} = \begin{pmatrix} -1 & 0 \\ -2 & -1 \end{pmatrix} \begin{pmatrix} -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Infatti:

$$\mathbf{v} = 1 \cdot \mathbf{v}_1 - 1 \cdot \mathbf{v}_2 = -\mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_2 = -\mathbf{e}_1 + 3\mathbf{e}_2$$

---

## 8.3 CAMBIO DI BASE PER UN'APPLICAZIONE LINEARE

In questo paragrafo vogliamo affrontare una generalizzazione del problema studiato nel paragrafo precedente, vediamo più precisamente.

Sia $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ un'applicazione lineare a cui è associata la matrice $A_{\mathcal{C},\mathcal{C}'}$ rispetto alle basi canoniche $\mathcal{C}$ e $\mathcal{C}'$ nel dominio e nel codominio rispettivamente. Vogliamo stabilire qual è la matrice associata a $F$ se fissiamo le basi ordinate $\mathcal{B}$ per $\mathbb{R}^n$ e $\mathcal{B}'$ per $\mathbb{R}^m$.

Sia l'applicazione lineare $F$, sia i vettori dello spazio vettoriale si comportano in modo indipendente dalla base che noi arbitrariamente scegliamo per rappresentarli. L'esempio dell'identità è particolarmente utile, perché l'identità resta sempre quell'applicazione che a ogni vettore associa se stesso, ma, se cambiamo la base in cui rappresentarla, la matrice associata subisce variazioni drastiche.

Consideriamo le seguenti applicazioni lineari, dove abbiamo posto accanto a ogni spazio vettoriale la base che scegliamo per rappresentare i vettori; nella riga superiore, scegliamo le basi canoniche, in quella inferiore, due basi ordinate $\mathcal{B}$ e $\mathcal{B}'$ scelte a piacere.

$$\begin{array}{ccc} \mathcal{C}\ \mathbb{R}^n & \xrightarrow{F} & \mathbb{R}^m\ \mathcal{C}' \\ \mathrm{id}\uparrow & & \downarrow \mathrm{id} \\ \mathcal{B}\ \mathbb{R}^n & \xrightarrow{F} & \mathbb{R}^m\ \mathcal{B}' \end{array}$$

Questo è quello che si chiama un *diagramma commutativo* in quanto il percorso che scegliamo nel diagramma non influenza il risultato:

$$F = \mathrm{id} \circ F \circ \mathrm{id}$$

L'uguaglianza che abbiamo scritto appare come una tautologia e non particolarmente interessante, tuttavia, nel momento in cui associamo a ciascuna applicazione lineare la rispettiva matrice rispetto alle basi fissate nel dominio e nel codominio, questa stessa uguaglianza ci fornirà una risposta completa alla domanda che ci siamo posti all'inizio di questo paragrafo e che rappresenta forse il punto più tecnico del nostro testo.

Associamo dunque a ogni applicazione lineare la matrice corrispondente e riscriviamo lo stesso diagramma utilizzando le coordinate, cioè usando le matrici per rappresentare le applicazioni lineari. Per quanto sappiamo dal paragrafo precedente si ottiene:

$$\begin{array}{cccccc} & \mathbb{R}^n & \xrightarrow{F} & \mathbb{R}^m & \\ I_{\mathcal{B},\mathcal{C}}(\mathbf{v})_{\mathcal{B}} & (\mathbf{v})_{\mathcal{C}} & \mapsto & A_{\mathcal{C},\mathcal{C}'}(\mathbf{v})_{\mathcal{C}} & (\mathbf{v})_{\mathcal{C}'} \\ \uparrow & \mathrm{id}\uparrow & & \downarrow \mathrm{id} & \downarrow \\ (\mathbf{v})_{\mathcal{B}} & \mathbb{R}^n & \xrightarrow{F} & \mathbb{R}^m & I_{\mathcal{C}',\mathcal{B}'}(\mathbf{v})_{\mathcal{C}'} \\ & (\mathbf{v})_{\mathcal{B}} & \mapsto & A_{\mathcal{B},\mathcal{B}'}(\mathbf{v})_{\mathcal{B}} & \end{array}$$

Vogliamo sfruttare questo diagramma per determinare $A_{\mathcal{B},\mathcal{B}'}$, cioè la matrice associata a $F$ rispetto alle basi $\mathcal{B}$ nel dominio e $\mathcal{B}'$ nel codominio.

Per l'Osservazione 8.2.3 abbiamo che l'uguaglianza:

$$F = \mathrm{id} \circ F \circ \mathrm{id}$$

corrisponde a:

$$A_{\mathcal{B},\mathcal{B}'} = I_{\mathcal{C}',\mathcal{B}'} A_{\mathcal{C},\mathcal{C}'} I_{\mathcal{B},\mathcal{C}} = I^{-1}_{\mathcal{B}',\mathcal{C}'} A_{\mathcal{C},\mathcal{C}'} I_{\mathcal{B},\mathcal{C}}$$

dove nell'ultima uguaglianza abbiamo utilizzato il Teorema 8.2.4.

Abbiamo pertanto dimostrato il seguente teorema.

**Teorema 8.3.1** *Sia $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ un'applicazione lineare, $A_{\mathcal{C},\mathcal{C}'}$, la matrice ad essa associata rispetto alle basi canoniche $\mathcal{C}$ e $\mathcal{C}'$ nel dominio e nel codominio rispettivamente. Allora la matrice associata a $F$ rispetto alle basi ordinate $\mathcal{B}$ nel dominio e $\mathcal{B}'$ nel codominio è data da:*

$$A_{\mathcal{B},\mathcal{B}'} = I^{-1}_{\mathcal{B}',\mathcal{C}'} A_{\mathcal{C},\mathcal{C}'} I_{\mathcal{B},\mathcal{C}}$$

*dove:*

1. *$A_{\mathcal{C},\mathcal{C}'}$ ha per colonne le coordinate delle immagini dei vettori della base canonica, cioè di $F(\mathbf{e}_1), \dots, F(\mathbf{e}_n)$, rispetto alla base canonica di $\mathbb{R}^m$;*

2. *$I_{\mathcal{B}',\mathcal{C}'}$ ha per colonne le coordinate dei vettori della base $\mathcal{B}'$, rispetto alla base canonica di $\mathbb{R}^m$;*

3. *$I_{\mathcal{B},\mathcal{C}}$ ha per colonne le coordinate dei vettori della base $\mathcal{B}$, rispetto alla base canonica di $\mathbb{R}^n$.*

Vediamo un esempio per chiarire come sia possibile calcolare esplicitamente $A_{\mathcal{B},\mathcal{B}'}$.

---

**Esempio 8.3.2**

Sia data l'applicazione $F : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ la cui matrice associata rispetto alle basi canoniche è:

$$A_{\mathcal{C},\mathcal{C}'} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Vogliamo trovare la matrice $A_{\mathcal{B},\mathcal{B}'}$ associata a $F$ rispetto alle basi ordinate: $\mathcal{B} = \{\mathbf{e}_1 + \mathbf{e}_2, -\mathbf{e}_1 - 2\mathbf{e}_2\}$ nel dominio e $\mathcal{B}' = \{2\mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_2\}$ nel codominio.

Troviamo le matrici $I_{\mathcal{B},\mathcal{C}}$ e $I_{\mathcal{B}',\mathcal{C}'}$:

$$I_{\mathcal{B},\mathcal{C}} = \begin{pmatrix} 1 & -1 \\ 1 & -2 \end{pmatrix} \qquad I_{\mathcal{B}',\mathcal{C}'} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}$$

Calcoliamo con un qualunque metodo $I^{-1}_{\mathcal{B}',\mathcal{C}'}$:

$$I^{-1}_{\mathcal{B}',\mathcal{C}'} = \begin{pmatrix} -1/2 & 1/2 & 1/2 \\ 1 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

La matrice $A_{\mathcal{B},\mathcal{B}'}$ è data dalla formula:

$$A_{\mathcal{B},\mathcal{B}'} = I^{-1}_{\mathcal{B}',\mathcal{C}'} A_{\mathcal{C},\mathcal{C}'} I_{\mathcal{B},\mathcal{C}} = \begin{pmatrix} -1/2 & 1/2 & 1/2 \\ 1 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & -2 \end{pmatrix} = \begin{pmatrix} -3/2 & 2 \\ 2 & -3 \\ 1 & -2 \end{pmatrix}$$

---

Concludiamo con un'osservazione generale.

Sia l'applicazione lineare $F$, sia i vettori dello spazio vettoriale si comportano in modo indipendente dalla base che noi arbitrariamente scegliamo per rappresentarli. L'esempio dell'identità è particolarmente utile, perché l'identità resta sempre quell'applicazione che a ogni vettore associa se stesso, ma, se cambiamo la base in cui rappresentarla, la matrice associata subisce variazioni drastiche.

Consideriamo le seguenti applicazioni lineari, dove abbiamo posto accanto a ogni spazio vettoriale la base che scegliamo per rappresentare i vettori; nella riga superiore, scegliamo le basi canoniche, in quella inferiore, due basi ordinate $\mathcal{B}$ e $\mathcal{B}'$ scelte a piacere.

$$\begin{array}{ccccc}
\mathcal{C} & \mathbb{R}^n & \xrightarrow{F} & \mathbb{R}^m & \mathcal{C}' \\
& \mathrm{id} \uparrow & & \downarrow \mathrm{id} & \\
\mathcal{B} & \mathbb{R}^n & \xrightarrow{F} & \mathbb{R}^m & \mathcal{B}'
\end{array}$$

Questo è quello che si chiama un *diagramma commutativo* in quanto il percorso che scegliamo nel diagramma non influenza il risultato:

$$F = \mathrm{id} \circ F \circ \mathrm{id}$$

L'uguaglianza che abbiamo scritto appare come una tautologia e non particolarmente interessante, tuttavia, nel momento in cui associamo a ciascuna applicazione lineare la rispettiva matrice rispetto alle basi fissate nel dominio e nel codominio, questa stessa uguaglianza ci fornirà una risposta completa alla domanda che ci siamo posti all'inizio di questo paragrafo e che rappresenta forse il punto più tecnico del nostro testo.

Associamo dunque a ogni applicazione lineare la matrice corrispondente e riscriviamo lo stesso diagramma utilizzando le coordinate, cioè usando le matrici per rappresentare le applicazioni lineari. Per quanto sappiamo dal paragrafo precedente si ottiene:

$$\begin{array}{cccccc}
& \mathbb{R}^n & \xrightarrow{F} & \mathbb{R}^m & \\
I_{\mathcal{B},\mathcal{C}}(\mathbf{v})_{\mathcal{B}} & (\mathbf{v})_{\mathcal{C}} & \mapsto & A_{\mathcal{C},\mathcal{C}'}(\mathbf{v})_{\mathcal{C}} & (\mathbf{v})_{\mathcal{C}'} \\
\uparrow & \mathrm{id} \uparrow & & \downarrow \mathrm{id} & \downarrow \\
(\mathbf{v})_{\mathcal{B}} & \mathbb{R}^n & \xrightarrow{F} & \mathbb{R}^m & I_{\mathcal{C}',\mathcal{B}'}(\mathbf{v})_{\mathcal{C}'} \\
& (\mathbf{v})_{\mathcal{B}} & \mapsto & A_{\mathcal{B},\mathcal{B}'}(\mathbf{v})_{\mathcal{B}} &
\end{array}$$

Vogliamo sfruttare questo diagramma per determinare $A_{\mathcal{B},\mathcal{B}'}$, cioè la matrice associata a $F$ rispetto alle basi $\mathcal{B}$ nel dominio e $\mathcal{B}'$ nel codominio.

Per l'Osservazione 8.2.3 abbiamo che l'uguaglianza:

$$F = \mathrm{id} \circ F \circ \mathrm{id}$$

corrisponde a:

$$A_{\mathcal{B},\mathcal{B}'} = I_{\mathcal{C}',\mathcal{B}'} A_{\mathcal{C},\mathcal{C}'} I_{\mathcal{B},\mathcal{C}} = I^{-1}_{\mathcal{B}',\mathcal{C}'} A_{\mathcal{C},\mathcal{C}'} I_{\mathcal{B},\mathcal{C}}$$

dove nell'ultima uguaglianza abbiamo utilizzato il Teorema 8.2.4.

Abbiamo pertanto dimostrato il seguente teorema.

**Teorema 8.3.1** *Sia $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ un'applicazione lineare, $A_{\mathcal{C},\mathcal{C}'}$ la matrice ad essa associata rispetto alle basi canoniche $\mathcal{C}$ e $\mathcal{C}'$ nel dominio e nel codominio rispettivamente. Allora la matrice associata a $F$ rispetto alle basi ordinate $\mathcal{B}$ nel dominio e $\mathcal{B}'$ nel codominio è data da:*

$$A_{\mathcal{B},\mathcal{B}'} = I^{-1}_{\mathcal{B}',\mathcal{C}'} A_{\mathcal{C},\mathcal{C}'} I_{\mathcal{B},\mathcal{C}}$$

*dove:*

1. *$A_{\mathcal{C},\mathcal{C}'}$ ha per colonne le coordinate delle immagini dei vettori della base canonica, cioè di $F(\mathbf{e}_1), \dots, F(\mathbf{e}_n)$, rispetto alla base canonica di $\mathbb{R}^m$;*

2. *$I_{\mathcal{B}',\mathcal{C}'}$ ha per colonne le coordinate dei vettori della base $\mathcal{B}'$, rispetto alla base canonica di $\mathbb{R}^m$;*

3. *$I_{\mathcal{B},\mathcal{C}}$ ha per colonne le coordinate dei vettori della base $\mathcal{B}$, rispetto alla base canonica di $\mathbb{R}^n$.*

Vediamo un esempio per chiarire come sia possibile calcolare esplicitamente $A_{\mathcal{B},\mathcal{B}'}$.

**Esempio 8.3.2**

Sia data l'applicazione $F : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ la cui matrice associata rispetto alle basi canoniche è:

$$A_{\mathcal{C},\mathcal{C}'} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Vogliamo trovare la matrice $A_{\mathcal{B},\mathcal{B}'}$ associata a $F$ rispetto alle basi ordinate: $\mathcal{B} = \{\mathbf{e}_1 + \mathbf{e}_2, -\mathbf{e}_1 - 2\mathbf{e}_2\}$ nel dominio e $\mathcal{B}' = \{2\mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_2\}$ nel codominio.

Troviamo le matrici $I_{\mathcal{B},\mathcal{C}}$ e $I_{\mathcal{B}',\mathcal{C}'}$:

$$I_{\mathcal{B},\mathcal{C}} = \begin{pmatrix} 1 & -1 \\ 1 & -2 \end{pmatrix} \qquad I_{\mathcal{B}',\mathcal{C}'} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}$$

Calcoliamo con un qualunque metodo $I^{-1}_{\mathcal{B}',\mathcal{C}'}$:

$$I^{-1}_{\mathcal{B}',\mathcal{C}'} = \begin{pmatrix} -1/2 & 1/2 & 1/2 \\ 1 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

La matrice $A_{\mathcal{B},\mathcal{B}'}$ è data dalla formula:

$$A_{\mathcal{B},\mathcal{B}'} = I^{-1}_{\mathcal{B}',\mathcal{C}'} A_{\mathcal{C},\mathcal{C}'} I_{\mathcal{B},\mathcal{C}} = \begin{pmatrix} -1/2 & 1/2 & 1/2 \\ 1 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & -2 \end{pmatrix} = \begin{pmatrix} -3/2 & 2 \\ 2 & -3 \\ 1 & -2 \end{pmatrix}$$

Concludiamo con un'osservazione generale.

♦ **Osservazione 8.3.3** Il Teorema 5.2.2 stabilisce una corrispondenza biunivoca tra applicazioni lineari da $\mathbb{R}^n$ a $\mathbb{R}^m$ e matrici in $M_{m,n}(\mathbb{R})$, una volta fissate le basi canoniche in $\mathbb{R}^n$ e $\mathbb{R}^m$. Tuttavia il lettore attento non faticherà a convincersi che la scelta delle basi canoniche è totalmente arbitraria e dunque tale corrispondenza biunivoca si ha anche nel caso che si fissino due basi ordinate qualsiasi di $\mathbb{R}^n$ e $\mathbb{R}^m$ rispettivamente. Ovviamente la matrice associata a una stessa applicazione lineare sarà in generale diversa a seconda delle basi che decidiamo di fissare. Possiamo calcolare tale matrice in modo diretto, semplicemente conoscendo le coordinate dei vettori delle due basi fissate rispetto alle basi canoniche, grazie al Teorema 8.3.1.

Possiamo inoltre stabilire in modo del tutto analogo una corrispondenza biunivoca tra applicazioni lineari tra gli spazi vettoriali $V$ a $W$ e matrici in $M_{m,n}(\mathbb{R})$, con $\dim(V) = n$ e $\dim(W) = m$. Invitiamo il lettore a ripercorrere i calcoli effettuati prima di enunciare il Teorema 5.2.2, sostituendo gli spazi vettoriali $V$ e $W$ a $\mathbb{R}^n$ e $\mathbb{R}^m$ rispettivamente, e rimpiazzando le basi canoniche con due basi ordinate arbitrarie di $V$ e $W$ rispettivamente.

## 8.4 ESERCIZI SVOLTI

### Esercizio 8.4.1

*Sia $F : M_2(\mathbb{R}) \longrightarrow \mathbb{R}$ l'applicazione lineare che associa a ogni matrice la sua traccia, cioè:*

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

*Si determini la matrice $A$ associata a $F$ rispetto alle basi canoniche di $M_2(\mathbb{R})$ e $\mathbb{R}$.*

**Svolgimento**

Osserviamo che $M_2(\mathbb{R})$ ha dimensione 4 e $\mathbb{R}$ ha dimensione 1, quindi $A$ è una matrice 4×1. La prima "colonna" di $A$ è costituita da $F(\mathbf{e}_{11}) = F\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right) = 1+0 = 1$, la seconda da $F(\mathbf{e}_{12}) = F\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) = 0+0 = 0$, la terza da $F(\mathbf{e}_{21}) = F\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right) = 0$ e la quarta da $F(\mathbf{e}_{22}) = F\left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right) = 1$, quindi $A = (1, 0, 0, 1)$. ∎

### Esercizio 8.4.2

*Sia $F : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ l'applicazione lineare definita da: $F(\mathbf{e}_1) = 2\mathbf{e}_1 - \mathbf{e}_2$, $F(\mathbf{e}_2) = \mathbf{e}_1$, $F(\mathbf{e}_3) = \mathbf{e}_1 + \mathbf{e}_2$. Sia $\mathcal{B} = \{2\mathbf{e}_1 - \mathbf{e}_2, \mathbf{e}_1 - \mathbf{e}_2\}$ una base ordinata di $\mathbb{R}^2$. Si determini la matrice $A_{\mathcal{C},\mathcal{B}}$ associata a $F$ rispetto alla base canonica nel dominio e alla base $\mathcal{B}$ nel codominio.*

**Svolgimento**

La matrice associata a $F$ rispetto alle basi canoniche del dominio e del codominio è:

$$A_{\mathcal{C},\mathcal{C}'} = \begin{pmatrix} 2 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}$$

Vogliamo cambiare base nel codominio, quindi consideriamo la seguente composizione di funzioni:

$$\begin{array}{ccccc} \mathbb{R}^3 & \xrightarrow{F} & \mathbb{R}^2 & \xrightarrow{\text{id}} & \mathbb{R}^2 \\ \mathcal{C} & & \mathcal{C}' & & \mathcal{B} \end{array}$$

L'applicazione composta è id $\circ$ $F = F$ e la matrice ad essa associata è $A_{\mathcal{C},\mathcal{B}} = I_{\mathcal{C}',\mathcal{B}}A_{\mathcal{C},\mathcal{C}'}$. Inoltre $I_{\mathcal{C}',\mathcal{B}} = I^{-1}_{\mathcal{B},\mathcal{C}'}$, dove $I_{\mathcal{B},\mathcal{C}'}$ è la matrice che ha per colonne le coordinate dei vettori di $\mathcal{B}$ rispetto alla base canonica $\mathcal{C}'$, quindi: $I_{\mathcal{B},\mathcal{C}'} = \left(\begin{smallmatrix} 2 & 1 \\ -1 & -1 \end{smallmatrix}\right)$. Abbiamo che $I^{-1}_{\mathcal{B},\mathcal{C}'} = \left(\begin{smallmatrix} 1 & 1 \\ -1 & -2 \end{smallmatrix}\right)$, dunque

$$A_{\mathcal{C},\mathcal{B}} = I_{\mathcal{C}',\mathcal{B}}A_{\mathcal{C},\mathcal{C}'} = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix}\begin{pmatrix} 2 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 2 \\ 0 & -1 & -3 \end{pmatrix}$$

■

### Esercizio 8.4.3

*Sia $\mathcal{B} = \{(2,-1),(1,1)\}$ una base ordinata di $\mathbb{R}^2$ e sia $G : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ l'applicazione lineare definita da: $G(2,-1) = (1,-1,0)$, $G(1,1) = (2,1,-2)$. Si determini la matrice $A_{\mathcal{C},\mathcal{C}'}$ associata a $G$ rispetto alle basi canoniche $\mathcal{C}$ del dominio e $\mathcal{C}'$ del codominio.*

**Svolgimento**

La matrice associata a $G$ rispetto alla base $\mathcal{B}$ del dominio e alla base $\mathcal{C}'$ del codominio è:

$$A_{\mathcal{B},\mathcal{C}'} = \begin{pmatrix} 1 & 2 \\ -1 & 1 \\ 0 & -2 \end{pmatrix}$$

Avendo cambiato base nel dominio, consideriamo la seguente composizione di funzioni:

$$\begin{array}{ccccc} \mathbb{R}^2 & \xrightarrow{\text{id}} & \mathbb{R}^2 & \xrightarrow{F} & \mathbb{R}^3 \\ \mathcal{B} & & \mathcal{C} & & \mathcal{C}' \end{array}$$

L'applicazione composta è $G \circ \text{id} = G$ e la matrice ad essa associata è $A_{\mathcal{B},\mathcal{C}'} = A_{\mathcal{C},\mathcal{C}'}I_{\mathcal{B},\mathcal{C}}$, quindi, moltiplicando entrambi i membri a destra per $I^{-1}_{\mathcal{B},\mathcal{C}}$, si ottiene: $A_{\mathcal{C},\mathcal{C}'} = A_{\mathcal{B},\mathcal{C}'}I^{-1}_{\mathcal{B},\mathcal{C}}$.

Abbiamo che $I_{\mathcal{B},\mathcal{C}} = \left(\begin{smallmatrix} 2 & 1 \\ -1 & 1 \end{smallmatrix}\right)$ e $I^{-1}_{\mathcal{B},\mathcal{C}'} = \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}$, quindi

$$A_{\mathcal{C},\mathcal{C}'} = \begin{pmatrix} 1 & 2 \\ -1 & 1 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ -\frac{2}{3} & -\frac{4}{3} \end{pmatrix}$$

■

## 8.5 ESERCIZI PROPOSTI

### Esercizio 8.5.1

*Si scrivano le coordinate del vettore $\mathbf{v} = (1,2,3)$ rispetto alla base ordinata $\mathcal{B} = \{\mathbf{e}_1 - \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3, -\mathbf{e}_1 + \mathbf{e}_3\}$ di $\mathbb{R}^3$.*

### Esercizio 8.5.2

*Sia data l'applicazione lineare $F : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ definita da: $F(\mathbf{e}_1) = -\mathbf{e}_2 + \mathbf{e}_3$, $F(\mathbf{e}_2) = 3\mathbf{e}_1 - 2\mathbf{e}_3$.*

a) *Si determinino la matrice $A$ associata a $F$ rispetto alla base canonica, e la matrice $A_{\mathcal{B},\mathcal{B}'}$ associata a $F$ rispetto alle basi ordinate $\mathcal{B} = \{(1,-2),(-2,1)\}$ di $\mathbb{R}^2$ e $\mathcal{B}' = \{(2,0,5),(0,-1,0),(1,1,3)\}$ di $\mathbb{R}^3$.*
b) *Si determinino le coordinate del vettore $\mathbf{v} = (1,-1,1)$ rispetto alla base $\mathcal{B}'$ di $\mathbb{R}^3$.*

### Esercizio 8.5.3

*Sia data l'applicazione lineare $F : \mathbb{R}^4 \longrightarrow \mathbb{R}^2$ definita da: $F(\mathbf{e}_1) = 2\mathbf{e}_2$, $F(\mathbf{e}_2) = \mathbf{e}_1 - 3\mathbf{e}_2$, $F(\mathbf{e}_3) = -2\mathbf{e}_1$, $F(\mathbf{e}_4) = \mathbf{e}_1 + \mathbf{e}_2$.*
a) *Si determinino la matrice $A$ associata a $F$ rispetto alla base canonica, e la matrice $A_{\mathcal{B},\mathcal{C}}$ associata a $F$ rispetto alla basi ordinate $\mathcal{B} = \{\mathbf{e}_2, \mathbf{e}_4, \mathbf{e}_1, \mathbf{e}_3\}$ di $\mathbb{R}^4$ e la base canonica $\mathcal{C}$ di $\mathbb{R}^2$.*
b) *Si determinino le coordinate del vettore $\mathbf{v} = (6,-1,3,-2)$ rispetto alla base $\mathcal{B}$ di $\mathbb{R}^4$.*

### Esercizio 8.5.4

*Sia data l'applicazione lineare $F : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definita da:*

$$F(x,y,z) = (8x - 9y, 6x - 7y, 2x - 3y - z)$$

a) *Si determinino la matrice $A$ associata a $F$ rispetto alla base canonica, e la matrice $A_{\mathcal{B}}$ assciata a $F$ rispetto alla base ordinata $\mathcal{B} = \{-\mathbf{e}_1 + \mathbf{e}_2, -2\mathbf{e}_1 + 3\mathbf{e}_2, -\mathbf{e}_3\}$.*
b) *Si dica se $F$ è un isomorfismo. Si motivi la risposta.*
c) *In caso di risposta affermativa al punto (b), si calcoli l'inversa di $F$.*

### Esercizio 8.5.5

*Sia data l'applicazione lineare $T : \mathbb{R}^3 \to \mathbb{R}^2$ definita da:*

$$T(x,y,z) = (3kx + y - 2kz, 3x + ky - 2z)$$

a) *Si determini per quali valori di $k$ si ha che $T$ è suriettiva, motivando il procedimento seguito. Posto $k = 0$, si determini $\ker T$.*
b) *Sia $\mathcal{B} = \{4\mathbf{e}_1 - 2\mathbf{e}_2, -\mathbf{e}_1 + \mathbf{e}_2\}$ un'altra base ordinata di $\mathbb{R}^2$. Posto $k = 0$ si determini la matrice $A_{\mathcal{C},\mathcal{B}}$ associata a $T$ rispetto alla base canonica $\mathcal{C}$ di $\mathbb{R}^3$ nel dominio e alla base $\mathcal{B}$ nel codominio.*

### Esercizio 8.5.6

a) *Sia $F : \mathbb{R}^3 \to \mathbb{R}^3$ l'applicazione lineare definita da:*

$$F(x,y,z) = (x - 4y - 2z, -x + ky + kz, kx - 4ky + z)$$

*Si stabilisca per quali valori di $k$ $F$ è suriettiva.*
b)*Posto $k = 0$ si determini, se possibile, un'applicazione lineare $G : \mathbb{R}^3 \to \mathbb{R}^3$ tale che $G \circ F$ sia l'identità.*
c) *Sia $\mathcal{B} = \{\mathbf{e}_1 + \mathbf{e}_2, -\mathbf{e}_1 + \mathbf{e}_3, 2\mathbf{e}_2\}$ un'altra base ordinata di $\mathbb{R}^3$. Posto $k = 0$ si determini la matrice $A_{\mathcal{C},\mathcal{B}}$ associata a $F$ rispetto alla base $\mathcal{B}$ nel dominio e alla base canonica $\mathcal{C}$ di $\mathbb{R}^3$ nel codominio.*

## Esercizio 8.5.7

*Sia* $D : \mathbb{R}_3[x] \longrightarrow \mathbb{R}_3[x]$ *l'applicazione lineare che a ogni polinomio associa la sua derivata. Si determini la matrice associata a* $D$ *rispetto alla base ordinata* $\{x^3, x^2, x, 1\}$ *di* $\mathbb{R}_3[x]$.

## Esercizio 8.5.8

*Si determini un'applicazione lineare* $F : \mathbb{R}^3 \to \mathbb{R}^3$ *tale che* $\ker F = \langle \mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_2 - \mathbf{e}_3 \rangle$ *e* $\operatorname{Im} F = \langle 2\mathbf{e}_1 + \mathbf{e}_3 \rangle$. (Suggerimento: conviene scegliere opportune basi nel dominio e nel codominio).
*Si determini la matrice associata a* $F$ *rispetto alla base canonica.*

# 9 Autovalori e autovettori

In questo capitolo vogliamo affrontare uno degli argomenti più profondi e interessanti dell'algebra lineare e cioè il problema della diagonalizzazione di un'applicazione lineare e i concetti di autovalore e autovettore.

## 9.1 DIAGONALIZZABILITÀ

L'idea dietro al problema della diagonalizzabilità è molto semplice: data un'applicazione lineare $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ ci chiediamo se esista una base ordinata, la stessa sia per il dominio sia per il codominio, rispetto alla quale la matrice associata a $F$ abbia la forma più semplice possibile e cioè la forma diagonale. Vediamo subito un esempio.

---

**Esempio 9.1.1**

Sia $\phi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ la trasformazione lineare definita da: $\phi(\mathbf{e}_1) = \mathbf{e}_2$, $\phi(\mathbf{e}_2) = \mathbf{e}_1$. Rispetto alla base canonica di dominio e codominio, $\phi$ è rappresentata dalla matrice:

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

Osserviamo che:

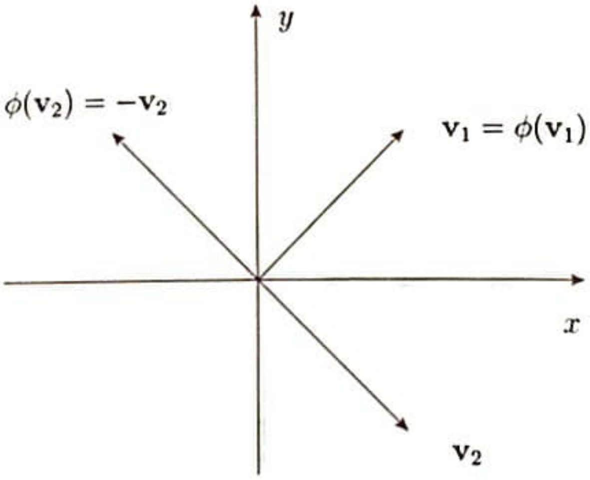
$$\phi(\mathbf{e}_1 + \mathbf{e}_2) = \mathbf{e}_1 + \mathbf{e}_2 \qquad \phi(\mathbf{e}_1 - \mathbf{e}_2) = -(\mathbf{e}_1 - \mathbf{e}_2)$$

Quindi, se scegliamo come base ordinata $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$, con $\mathbf{v}_1 = \mathbf{e}_1 + \mathbf{e}_2$ e $\mathbf{v}_2 = \mathbf{e}_1 - \mathbf{e}_2$, abbiamo che $\phi(\mathbf{v}_1) = \mathbf{v}_1$ e $\phi(\mathbf{v}_2) = -\mathbf{v}_2$. Perciò la matrice associata a $\phi$, rispetto alla base $\mathcal{B}$ (nel dominio e nel codominio), è una matrice diagonale:

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

Possiamo vederlo subito senza calcoli, tuttavia per convincersene basta utilizzare la Definizione 8.1.1 o la formula del Paragrafo 8.3.

Geometricamente in questo caso è molto semplice vedere che cosa accade. La trasformazione $\phi$ consiste in una riflessione del piano cartesiano rispetto alla retta $x = y$. Infatti $\phi(\mathbf{e}_1) = \mathbf{e}_2$, $\phi(\mathbf{e}_2) = \mathbf{e}_1$. È perciò immediato geometricamente stabilire che il vettore $\mathbf{v}_1$, che giace sulla retta $x = y$, è fissato dalla trasformazione mentre il vettore $\mathbf{v}_2$, che è perpendicolare alla retta $x = y$, viene mandato in $-\mathbf{v}_2$. In base a queste osservazioni geometriche si può concludere senza alcun calcolo che, rispetto alla base ordinata $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$, la matrice associata a $\phi$ ha la forma diagonale specificata.




**Definizione 9.1.2** Un'applicazione lineare $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ si dice *diagonalizzabile* se esiste una base ordinata $\mathcal{B}$ per $\mathbb{R}^n$ tale che la matrice $A_\mathcal{B}$ associata a $T$ rispetto a $\mathcal{B}$ (nel dominio e nel codominio) sia una matrice diagonale.

Nell'esempio precedente l'applicazione $\phi$ è diagonalizzabile e $\mathcal{B} = \{\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 - \mathbf{e}_2\}$ è la base ordinata rispetto alla quale la matrice associata a $\phi$ è diagonale.

Così come abbiamo dato la definizione di applicazione lineare diagonalizzabile, è possibile dare anche quella di matrice diagonalizzabile: essenzialmente si tratta di una matrice associata a un'applicazione lineare diagonalizzabile, ma vediamo la definizione precisa.

**Definizione 9.1.3** Una matrice quadrata $A$ si dice *diagonalizzabile* se esiste una matrice invertibile $P$ tale che $P^{-1}AP$ sia diagonale.

**Proposizione 9.1.4** *Sia $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ un'applicazione lineare associata alla matrice $A$ rispetto alla base canonica (in dominio e codominio). Allora $T$ è diagonalizzabile se e solo se $A$ è diagonalizzabile. Inoltre, se $T$ e $A$ sono diagonalizzabili, la base $\mathcal{B}$ che diagonalizza $T$ corrisponde alle colonne della matrice $P$ tale che $P^{-1}AP$ sia diagonale.*

**Dimostrazione** – L'enunciato risulta abbastanza immediato se ci ricordiamo come fare i cambi di base dal capitolo precedente. Supponiamo che $T$ sia diagonalizzabile. Allora esiste una base ordinata $\mathcal{B}$ rispetto a cui la matrice $A_\mathcal{B}$ associata a $T$ è diagonale. Detta $P = I_{\mathcal{B},\mathcal{C}}$ abbiamo per la formula del Paragrafo 8.3:

$$A_\mathcal{B} = P^{-1}AP$$

dunque per definizione $A$ è diagonalizzabile. Viceversa, se $A$ è diagonalizzabile, significa che esiste $P$ tale che $P^{-1}AP$ sia diagonale e, se consideriamo la base ordinata $\mathcal{B}$ che ha per elementi le colonne di $P$, si ha che $P$ è precisamente la matrice di cambio di base $I_{\mathcal{B},\mathcal{C}}$ per cui la matrice associata a $T$ rispetto alla base $\mathcal{B}$ è proprio $I_{\mathcal{B},\mathcal{C}}^{-1} A I_{\mathcal{B},\mathcal{C}} = P^{-1}AP$ ed è diagonale. Pertanto $T$ è diagonalizzabile □

♦ **Osservazione 9.1.5** A questo punto è chiaro che un'applicazione lineare è diagonalizzabile se e solo se la matrice ad essa associata rispetto a una qualunque base ordinata è diagonalizzabile. Lasciamo per esercizio la dimostrazione di questa affermazione.

Sorgono ora spontanee due domande:
1) in generale, un'applicazione lineare $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ è sempre diagonalizzabile? Equivalentemente, data una matrice quadrata $A$, esiste sempre una matrice invertibile $P$ tale che $P^{-1}AP$ sia diagonale?

2) Esiste una procedura per "diagonalizzare" un'applicazione lineare o una matrice diagonalizzabile? Cioè, esiste una procedura per determinare la base $\mathcal{B}$ o, equivalentemente, la matrice $P$?

La risposta alla prima domanda in generale è no, mentre alla seconda domanda è "sì"; vedremo in dettaglio la risposta nel prossimo paragrafo. Geometricamente è facile convincersi che esistono molte matrici non diagonalizzabili. Vediamo un esempio.

**Esempio 9.1.6**

Consideriamo l'applicazione lineare $\psi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ defnita da: $\psi(\mathbf{e}_1) = -\mathbf{e}_2$, $\psi(\mathbf{e}_2) = \mathbf{e}_1$. La matrice associata a $\psi$ rispetto alla base canonica è data da:

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

Nonostante sia molto simile a quella dell'esempio precedente, c'è un'importante differenza. Geometricamente $\psi$ corrisponde a una rotazione oraria di $90^\circ$ del piano attorno all'origine. Affinché $A$ o equivalentemente $\psi$ sia diagonalizzabile, devono esistere due vettori $\mathbf{v}_1$, $\mathbf{v}_2$ linearmente indipendenti tali che $\psi(\mathbf{v}_1) = \lambda_1\mathbf{v}_1$, $\psi(\mathbf{v}_2) = \lambda_2\mathbf{v}_2$ per $\lambda_1, \lambda_2 \in \mathbb{R}$, cioè due vettori che *conservino la propria direzione*. In tal caso infatti la matrice di $\psi$ rispetto alla base ordinata $\mathcal{B} = (\mathbf{v}_1, \mathbf{v}_2)$ sarebbe:

$$A_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

Ma è immediato vedere che una rotazione non fissa la direzione di nessun vettore, quindi $\mathbf{v}_1$ e $\mathbf{v}_2$ non esistono, cioè $A$ non è diagonalizzabile.

Si noti che il discorso sarebbe diverso se permettessimo agli scalari di assumere valori complessi. Infatti, in tal caso esisterebbero due vettori $\mathbf{v}_1 = (1, -i)$ e $\mathbf{v}_2 = (i, -1)$ tali che $\psi(\mathbf{v}_1) = -i\mathbf{v}_1$, $\psi(\mathbf{v}_2) = i\mathbf{v}_2$.

Questo esempio suggerisce una terza domanda:

3) Se permettiamo agli scalari di assumere valori complessi allora possiamo sempre diagonalizzare una data matrice (o applicazione lineare)?

La risposta rimane "no", tuttavia è possibile sempre ricondurre una matrice a una forma "quasi" diagonale, detta forma di Jordan, che non discuteremo, poiché ci richiederebbe molto tempo.

## 9.2 AUTOVALORI E AUTOVETTORI

Abbiamo visto nel paragrafo precedente che, per quanto riguarda la diagonalizzabilità, giocano un ruolo fondamentale i vettori la cui direzione non viene cambiata dall'applicazione lineare data, cioè quei vettori che vengono trasformati in multipli di se stessi. Qualora l'applicazione lineare data sia diagonalizzabile, tali vettori infatti costituiscono una base $\mathcal{B}$ rispetto alla quale l'applicazione lineare è associata a una matrice diagonale. Equivalentemente, le coordinate di tali vettori rispetto alla base canonica costituiscono le colonne della matrice $P$ del cambio di base, tale che $P^{-1}AP$ sia diagonale (dove $A$ è la matrice associata all'applicazione lineare data rispetto alla base canonica). Questi vettori, cruciali per il nostro problema, sono gli *autovettori*. Vediamone la definizione precisa.

**Definizione 9.2.1** Data un'applicazione lineare $T : V \longrightarrow V$ diciamo che un vettore non nullo $\mathbf{v} \in V$ è un *autovettore* di $T$ di *autovalore* $\lambda$ se $T(\mathbf{v}) = \lambda\mathbf{v}$ con $\lambda \in \mathbb{R}$.

Si osservi che un autovalore $\lambda$ può essere nullo, mentre un autovettore, per definizione, deve essere diverso dal vettore nullo.

Data una matrice $A$ si dicono *autovalori* e *autovettori* di $A$ gli autovalori e autovettori dell'applicazione lineare $L_A$ associata ad $A$ rispetto alla base canonica (nel dominio e nel codominio).

È importante capire che il fatto che un vettore sia autovettore e che un numero sia autovalore di un'applicazione lineare $T$ *non dipende dalla base scelta per rappresentare* $T$!

Nell'Esempio 9.1.1 abbiamo che i vettori $\mathbf{v}_1 = \mathbf{e}_1 + \mathbf{e}_2$ e $\mathbf{v}_2 = \mathbf{e}_1 - \mathbf{e}_2$ sono autovettori di autovalori rispettivamente 1 e $-1$ dell'applicazione $\phi$. Se prendiamo la base ordinata formata dagli autovettori $\mathcal{B} = (\mathbf{v}_1, \mathbf{v}_2)$, abbiamo che $\mathbf{v}_1 = (1,1)_\mathcal{C} = (1,0)_\mathcal{B}$, $\mathbf{v}_2 = (1,-1)_\mathcal{C} = (0,1)_\mathcal{B}$. Pur cambiando le coordinate, cioè il modo di scrivere questi vettori, i vettori $\mathbf{v}_1$ e $\mathbf{v}_2$ non cambiano e allo stesso modo non cambiano gli autovalori ad essi associati. Infatti nell'esempio abbiamo visto che la matrice associata a $\phi$ è

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

dunque gli autovalori restano $-1$, 1. Sintetizziamo il nostro discorso nella seguente osservazione, di validità generale.

**Osservazione 9.2.2** Uno scalare $\lambda$ è autovalore di una trasformazione $T$ se e solo se è autovalore della matrice $A_\mathcal{B}$ associata a $T$ rispetto a una qualunque base $\mathcal{B}$.

Veniamo ora a caratterizzare la proprietà di diagonalizzabilità di un'applicazione lineare.

**Proposizione 9.2.3** *Un'applicazione lineare $T : V \longrightarrow V$ è diagonalizzabile se e solo se esiste una base $\mathcal{B}$ di $V$ costituita da autovettori di $T$.*

La dimostrazione di questo enunciato è particolarmente istruttiva in quanto contiene fatti fondamentali per la comprensione dei concetti di autovalore e autovettore.

**Dimostrazione** – Se esiste una base $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ di autovettori, allora la matrice associata a $T$ rispetto alla base ordinata $\mathcal{B}$ è diagonale. Infatti:

$$T(\mathbf{v}_1) = \lambda_1\mathbf{v}_1 = \lambda_1\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_n$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$T(\mathbf{v}_n) = \lambda_n\mathbf{v}_n = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + \lambda_n\mathbf{v}_n$$

quindi la matrice associata a $T$ rispetto alla base ordinata $\mathcal{B}$ è:

$$A_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}$$

Dunque la trasformazione $T$ è diagonalizzabile per definizione.

Viceversa, se $T$ è diagonalizzabile, significa che esiste una base ordinata $\mathcal{B}$ rispetto a cui la matrice associata a $T$ è diagonale, cioè

$$A_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}$$

Ma allora tale matrice ha proprio gli autovalori sulla diagonale, in quanto

$$(T(\mathbf{v}_1))_{\mathcal{B}} = A(\mathbf{v}_1)_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}\begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \vdots \\ 0 \end{pmatrix}$$

$$\vdots$$

$$(T(\mathbf{v}_n))_{\mathcal{B}} = A(\mathbf{v}_n)_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}\begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ \lambda_n \end{pmatrix}$$

Dunque $T(\mathbf{v}_1) = \lambda_1\mathbf{v}_1, \dots, T(\mathbf{v}_n) = \lambda_n\mathbf{v}_n$ e la base $\mathcal{B}$ è costituita da autovettori. □

Vogliamo adesso dare un metodo operativo per calcolare autovalori e autovettori di una matrice o applicazione lineare data.

**Definizione 9.2.4** Data una matrice quadrata $A$ definiamo *polinomio caratteristico* $p_A$ di $A$ il seguente polinomio in $x$:

$$p_A(x) = \det(A - xI)$$

dove det denota il determinante e $I$ la matrice identità.

Il fatto che $p_A(x)$ sia effettivamente un polinomio in $x$ segue per esempio dal calcolo ricorsivo del determinante, sviluppando secondo una riga (o una colonna) di $A - xI$.

Nel seguito, se $A$ è una matrice, per brevità indicheremo con $\ker A$ il nucleo dell'applicazione lineare $L_A$ associata ad $A$ rispetto alle basi canoniche (nel dominio e nel codominio). Equivalentemente, $\ker A$ è l'insieme delle soluzioni del sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$.

**Teorema 9.2.5** *Uno scalare $\lambda$ è un autovalore di $A$ se e solo se è uno zero del polinomio caratteristico cioè* $\det(A - \lambda I) = 0$.

**Dimostrazione** – Se $\lambda$ è autovalore di $A$ allora esiste $\mathbf{v} \neq \mathbf{0}$ tale che $A\mathbf{v} = \lambda\mathbf{v}$, cioè $A\mathbf{v} - \lambda\mathbf{v} = \mathbf{0}$, cioè $(A - \lambda I)\mathbf{v} = \mathbf{0}$ e quindi $\mathbf{v} \in \ker(A - \lambda I)$. Quindi, per il Teorema 7.6.1, il determinante della matrice $A - \lambda I$ è zero.

Viceversa, se $\det(A - \lambda I) = 0$, per il Teorema 7.6.1 esiste un vettore non nullo $\mathbf{v} \in \ker(A - \lambda I)$. Allora $(A - \lambda I)\mathbf{v} = \mathbf{0}$ e quindi $A\mathbf{v} = \lambda\mathbf{v}$ e $\mathbf{v}$ è autovettore di autovalore $\lambda$.

□

**Definizione 9.2.6** Siano $A$ e $B$ due matrici $n \times n$. $A$ e $B$ si dicono *simili* se esiste una matrice $n \times n$ invertibile $P$ tale che:

$$B = P^{-1}AP$$

◆ **Osservazione 9.2.7**
i) Se $A$ e $B$ sono simili, allora $A$ e $B$ corrispondono alla *stessa* applicazione lineare rispetto a basi diverse. Ciò segue immediatamente dalla nostra trattazione sul cambio di base.
ii) $A$ è diagonalizzabile se e solo se è simile ad una matrice diagonale.

**Teorema 9.2.8** *Matrici simili hanno lo stesso polinomio caratteristico. (Attenzione, l'inverso è falso!).*

**Dimostrazione** – Siano $A$ e $B$ matrici simili, con $B = P^{-1}AP$. Utilizzando il Teorema di Binet (7.3.5) e le proprietà di associatività e distributività del prodotto di matrici, si ha che

$$\begin{aligned} p_B(x) &= \det(P^{-1}AP - xI) = \det(P^{-1}AP - P^{-1}PxI) \\ &= \det(P^{-1}AP - P^{-1}xIP) = \det(P^{-1}(A - xI)P) \\ &= \det P^{-1}\det(A - xI)\det P = \det(A - xI) = p_A(x) \end{aligned}$$

(In questa dimostrazione abbiamo anche utilizzato il fatto, di facile verifica, che, se $P$ è una matrice qualsiasi e $x$ uno scalare, allora $P(xI) = (xI)P$). □

◆ **Osservazione 9.2.9** Osserviamo che il viceversa del Teorema 9.2.8 è falso. Infatti, le due matrici

$$A = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 0 \\ 1 & 3 \end{pmatrix}$$

hanno lo stesso polinomio caratteristico $p_A(x) = (3 - x)^2 = p_B(x)$, ma, con le tecniche che acquisiremo alla fine del capitolo, è facile dedurre che $A$ e $B$ non sono simili, perché $B$ non è diagonalizzabile mentre $A$, essendo diagonale, lo è.

♦ **Osservazione 9.2.10** Dalla dimostrazione del Teorema 9.2.5 segue che un vettore $\mathbf{v} \neq \mathbf{0}$ è autovettore di una trasformazione lineare $T$ associato all'autovalore $\lambda$ se e solo se appartiene a $\ker(A - \lambda I)$, dove $A$ è la matrice associata a $T$ rispetto alla base canonica.

**Definizione 9.2.11** Se $\lambda$ è un autovalore di $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ allora $V_\lambda = \{\mathbf{v} \in \mathbb{R}^n \mid T(\mathbf{v}) = \lambda \mathbf{v}\}$ si dice *autospazio* relativo all'autovalore $\lambda$.

♦ **Osservazione 9.2.12** Osserviamo che gli elementi di $V_\lambda$ sono esattamente gli autovettori di autovalore $\lambda$ e il vettore nullo, quindi $V_\lambda = \ker(A - \lambda I)$, dove $A$ è la matrice associata a $T$ rispetto alla base canonica. Ovviamente $V_\lambda$ è un sottospazio vettoriale, essendo il nucleo di un'applicazione lineare.

Mettiamo ora in pratica quanto abbiamo appreso teoricamente e calcoliamo autovalori e autovettori di una matrice arbitraria.

---

**Esempio 9.2.13**

Vogliamo trovare autovalori e autovettori della matrice

$$A = \begin{pmatrix} 5 & -4 \\ 3 & -2 \end{pmatrix}$$

e stabilire se la matrice $A$ sia diagonalizzabile.

- Il polinomio caratteristico di $A$ è:

$$p_A(x) = \det \begin{pmatrix} 5 - x & -4 \\ 3 & -2 - x \end{pmatrix} = x^2 - 3x + 2$$

  L'equazione associata ha per soluzioni 1 e 2, quindi ci sono due autovalori per $A$, $\lambda_1 = 1$ e $\lambda_2 = 2$.

- Calcoliamo l'autospazio corrispondente all'autovalore $\lambda_1 = 1$. È costituito dai vettori $(x, y) \in \mathbb{R}^2$ tali che:

$$A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

  e quindi dalle soluzioni del sistema

$$\begin{cases} 5x - 4y = x \\ 3x - 2y = y \end{cases}$$

  Si tratta di un sistema lineare omogeneo, e la matrice dei coefficienti ad esso associata è:

$$\begin{pmatrix} 5 - 1 & -4 \\ 3 & -2 - 1 \end{pmatrix}$$

  che è la matrice usata per calcolare il polinomio caratteristico, in cui abbiamo sostituito l'autovalore 1 al posto di $x$, cioè è la matrice di $A - I$.
  Abbiamo quindi che

$$V_1 = \ker(A - I) = \ker \begin{pmatrix} 4 & -4 \\ 3 & -3 \end{pmatrix} = \langle (1, 1) \rangle$$

- Per calcolare l'autospazio relativo all'autovalore 2 si procede in modo analogo e si ottiene:

$$V_2 = \ker(A - 2I) = \ker \begin{pmatrix} 3 & -4 \\ 3 & -4 \end{pmatrix} = \langle (4/3, 1) \rangle$$

Il fatto che $p_A(x)$ sia effettivamente un polinomio in $x$ segue per esempio dal calcolo ricorsivo del determinante, sviluppando secondo una riga (o una colonna) di $A - xI$.

Nel seguito, se $A$ è una matrice, per brevità indicheremo con $\ker A$ il nucleo dell'applicazione lineare $L_A$ associata ad $A$ rispetto alle basi canoniche (nel dominio e nel codominio). Equivalentemente, $\ker A$ è l'insieme delle soluzioni del sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$.

**Teorema 9.2.5** *Uno scalare $\lambda$ è un autovalore di $A$ se e solo se è uno zero del polinomio caratteristico cioè* $\det(A - \lambda I) = 0$.

**Dimostrazione** – Se $\lambda$ è autovalore di $A$ allora esiste $\mathbf{v} \neq \mathbf{0}$ tale che $A\mathbf{v} = \lambda\mathbf{v}$, cioè $A\mathbf{v} - \lambda\mathbf{v} = \mathbf{0}$, cioè $(A - \lambda I)\mathbf{v} = \mathbf{0}$ e quindi $\mathbf{v} \in \ker(A - \lambda I)$. Quindi, per il Teorema 7.6.1, il determinante della matrice $A - \lambda I$ è zero.

Viceversa, se $\det(A - \lambda I) = 0$, per il Teorema 7.6.1 esiste un vettore non nullo $\mathbf{v} \in \ker(A - \lambda I)$. Allora $(A - \lambda I)\mathbf{v} = \mathbf{0}$ e quindi $A\mathbf{v} = \lambda\mathbf{v}$ e $\mathbf{v}$ è autovettore di autovalore $\lambda$.

□

**Definizione 9.2.6** Siano $A$ e $B$ due matrici $n \times n$. $A$ e $B$ si dicono *simili* se esiste una matrice $n \times n$ invertibile $P$ tale che:

$$B = P^{-1}AP$$

♦ **Osservazione 9.2.7**
i) Se $A$ e $B$ sono simili, allora $A$ e $B$ corrispondono alla *stessa* applicazione lineare rispetto a basi diverse. Ciò segue immediatamente dalla nostra trattazione sul cambio di base.
ii) $A$ è diagonalizzabile se e solo se è simile ad una matrice diagonale.

**Teorema 9.2.8** *Matrici simili hanno lo stesso polinomio caratteristico. (Attenzione, l'inverso è falso!).*

**Dimostrazione** – Siano $A$ e $B$ matrici simili, con $B = P^{-1}AP$. Utilizzando il Teorema di Binet (7.3.5) e le proprietà di associatività e distributività del prodotto di matrici, si ha che

$$\begin{aligned} p_B(x) &= \det(P^{-1}AP - xI) = \det(P^{-1}AP - P^{-1}PxI) \\ &= \det(P^{-1}AP - P^{-1}xIP) = \det(P^{-1}(A - xI)P) \\ &= \det P^{-1} \det(A - xI) \det P = \det(A - xI) = p_A(x) \end{aligned}$$

(In questa dimostrazione abbiamo anche utilizzato il fatto, di facile verifica, che, se $P$ è una matrice qualsiasi e $x$ uno scalare, allora $P(xI) = (xI)P$. □

♦ **Osservazione 9.2.9** Osserviamo che il viceversa del Teorema 9.2.8 è falso. Infatti, le due matrici

$$A = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 0 \\ 1 & 3 \end{pmatrix}$$

hanno lo stesso polinomio caratteristico $p_A(x) = (3 - x)^2 = p_B(x)$, ma, con le tecniche che acquisiremo alla fine del capitolo, è facile dedurre che $A$ e $B$ non sono simili, perché $B$ non è diagonalizzabile mentre $A$, essendo diagonale, lo è.

♦ **Osservazione 9.2.10** Dalla dimostrazione del Teorema 9.2.5 segue che un vettore $\mathbf{v} \neq 0$ è autovettore di una trasformazione lineare $T$ associato all'autovalore $\lambda$ se e solo se appartiene a $\ker(A - \lambda I)$, dove $A$ è la matrice associata a $T$ rispetto alla base canonica.

**Definizione 9.2.11** Se $\lambda$ è un autovalore di $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ allora $V_\lambda = \{\mathbf{v} \in \mathbb{R}^n \mid T(\mathbf{v}) = \lambda \mathbf{v}\}$ si dice *autospazio* relativo all'autovalore $\lambda$.

♦ **Osservazione 9.2.12** Osserviamo che gli elementi di $V_\lambda$ sono esattamente gli autovettori di autovalore $\lambda$ e il vettore nullo, quindi $V_\lambda = \ker(A - \lambda I)$, dove $A$ è la matrice associata a $T$ rispetto alla base canonica. Ovviamente $V_\lambda$ è un sottospazio vettoriale, essendo il nucleo di un'applicazione lineare.

Mettiamo ora in pratica quanto abbiamo appreso teoricamente e calcoliamo autovalori e autovettori di una matrice arbitraria.

---

**Esempio 9.2.13**

Vogliamo trovare autovalori e autovettori della matrice

$$A = \begin{pmatrix} 5 & -4 \\ 3 & -2 \end{pmatrix}$$

e stabilire se la matrice $A$ sia diagonalizzabile.

- Il polinomio caratteristico di $A$ è:

$$p_A(x) = \det \begin{pmatrix} 5 - x & -4 \\ 3 & -2 - x \end{pmatrix} = x^2 - 3x + 2$$

  L'equazione associata ha per soluzioni 1 e 2, quindi ci sono due autovalori per $A$, $\lambda_1 = 1$ e $\lambda_2 = 2$.

- Calcoliamo l'autospazio corrispondente all'autovalore $\lambda_1 = 1$. È costituito dai vettori $(x, y) \in \mathbb{R}^2$ tali che:

$$A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

  e quindi dalle soluzioni del sistema

$$\begin{cases} 5x - 4y = x \\ 3x - 2y = y \end{cases}$$

  Si tratta di un sistema lineare omogeneo, e la matrice dei coefficienti ad esso associata è:

$$\begin{pmatrix} 5 - 1 & -4 \\ 3 & -2 - 1 \end{pmatrix}$$

  che è la matrice usata per calcolare il polinomio caratteristico, in cui abbiamo sostituito l'autovalore 1 al posto di $x$, cioè è la matrice di $A - I$.
  Abbiamo quindi che

$$V_1 = \ker(A - I) = \ker \begin{pmatrix} 4 & -4 \\ 3 & -3 \end{pmatrix} = \langle (1, 1) \rangle$$

- Per calcolare l'autospazio relativo all'autovalore 2 si procede in modo analogo e si ottiene:

$$V_2 = \ker(A - 2I) = \ker \begin{pmatrix} 3 & -4 \\ 3 & -4 \end{pmatrix} = \langle (4/3, 1) \rangle$$

- I vettori $(1,1), (4/3,1)$ sono linearmente indipendenti, perciò abbiamo che $A$ è diagonalizzabile ed è simile alla matrice diagonale:

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = P^{-1}AP$$

$$\text{dove } P = \begin{pmatrix} 1 & 4/3 \\ 1 & 1 \end{pmatrix}$$

$P$ è la matrice del cambio di base, che ci permette di passare dalla base canonica alla base costituita dagli autovettori di $A$. Per la teoria sul cambio di base (*vedi* il Capitolo 8) le colonne di $P$ sono costituite dalle coordinate degli autovettori rispetto alla base canonica.

Ci sarà inoltre molto utile il seguente teorema.

**Teorema 9.2.14** *Sia $T$ una trasformazione lineare e siano $\mathbf{v}_1, \ldots, \mathbf{v}_n$ autovettori di $T$ relativi agli autovalori $\lambda_1, \ldots, \lambda_n$ rispettivamente. Se $\lambda_1, \ldots, \lambda_n$ sono tutti distinti tra loro, allora $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sono linearmente indipendenti.*

**Dimostrazione** – La dimostrazione si articola in $n$ passi, tutti simili tra loro.

*Passo* 1. Abbiamo che $\mathbf{v}_1$ è un vettore linearmente indipendente, perché $\mathbf{v}_1 \neq \mathbf{0}$.

*Passo* 2. Mostriamo che $\mathbf{v}_1, \mathbf{v}_2$ sono linearmente indipendenti. Siano $\beta_1, \beta_2 \in \mathbb{R}$ tali che:

$$\beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 = \mathbf{0} \tag{9.1}$$

Applicando $T$ ad entrambi i membri e tenendo conto che $T(\mathbf{v}_1) = \lambda_1\mathbf{v}_1$, $T(\mathbf{v}_2) = \lambda_2\mathbf{v}_2$, si ottiene:

$$\beta_1\lambda_1\mathbf{v}_1 + \beta_2\lambda_2\mathbf{v}_2 = \mathbf{0} \tag{9.2}$$

Ora sottraiamo dalla seconda uguaglianza la prima moltiplicata per $\lambda_2$ e otteniamo:

$$\beta_1(\lambda_1 - \lambda_2)\mathbf{v}_1 = \mathbf{0}$$

Poiché $\mathbf{v}_1 \neq \mathbf{0}$ si ottiene che $\beta_1(\lambda_1 - \lambda_2) = 0$, quindi $\beta_1 = 0$, essendo $\lambda_1 \neq \lambda_2$. Sostituendo $\beta_1 = 0$ nella (9.1), si ottiene che $\beta_2\mathbf{v}_2 = \mathbf{0}$. Ma $\mathbf{v}_2$ è un autovettore, quindi $\mathbf{v}_2 \neq \mathbf{0}$. Segue che $\beta_2 = 0$, come volevasi dimostrare.

Dopo il *passo* $k-1$ avremo dimostrato che i vettori $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$ sono linearmente indipendenti.

*Passo* $k$. Mostriamo che anche $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}, \mathbf{v}_k$ sono linearmente indipendenti. Siano $\beta_1, \ldots, \beta_{k-1}, \beta_k \in \mathbb{R}$ tali che:

$$\beta_1\mathbf{v}_1 + \cdots + \beta_{k-1}\mathbf{v}_{k-1} + \beta_k\mathbf{v}_k = \mathbf{0} \tag{9.3}$$

Applicando $T$ ad entrambi i membri e tenendo conto che $T(\mathbf{v}_i) = \lambda_i\mathbf{v}_i$ si ottiene:

$$\beta_1\lambda_1\mathbf{v}_1 + \cdots + \beta_{k-1}\lambda_{k-1}\mathbf{v}_{k-1} + \beta_k\lambda_k\mathbf{v}_k = \mathbf{0} \tag{9.4}$$

Ora sottraiamo dall'uguaglianza (9.4) l'uguaglianza (9.3) moltiplicata per $\lambda_k$ e otteniamo:

$$\beta_1(\lambda_1 - \lambda_k)\mathbf{v}_1 + \cdots + \beta_k(\lambda_k - \lambda_k)\mathbf{v}_k = \mathbf{0}$$

Poiché $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$ sono linearmente indipendenti e $\lambda_i - \lambda_k \neq 0$ per ogni $i = 1, \ldots, k-1$ (ricordiamo che gli autovalori sono tutti distinti), segue che $\beta_i = 0$ per ogni $i = 1, \ldots, k-1$. A questo punto l'uguaglianza (9.3) diventa:

$$\beta_k \mathbf{v}_k = \mathbf{0}$$

quindi anche $\beta_k = 0$, essendo $\mathbf{v}_k \neq \mathbf{0}$. Quindi i $\beta_i$ sono tutti nulli e $\mathbf{v}_1, \ldots, \mathbf{v}_k$ sono linearmente indipendenti, come volevasi dimostrare.

Dopo $n$ passi si ottiene quanto voluto, e cioè che $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sono linearmente indipendenti. □

Una conseguenza immediata del teorema precedente è il seguente corollario.

**Corollario 9.2.15** *Una matrice $A \in \mathrm{M}_n(\mathbb{R})$ con $n$ autovalori distinti è diagonalizzabile.*

**Dimostrazione** – Gli $n$ autovettori relativi agli $n$ autovalori distinti di $A$, per il Teorema 9.2.14 sono linearmente indipendenti, quindi formano una base di $\mathbb{R}^n$. □

Vediamo un esempio concreto di applicazione di quanto abbiamo appena visto.

---

**Esempio 9.2.16**

Vogliamo trovare gli autovalori e gli autovettori della matrice

$$A = \begin{pmatrix} -1 & 2 & 0 \\ 1 & 1 & 0 \\ -1 & 1 & 4 \end{pmatrix}$$

e stabilire se è diagonalizzabile.

Vediamo schematicamente come procedere.

- Calcoliamo il polinomio caratteristico di $A$:

$$p_A(x) = \det \begin{pmatrix} -1-x & 2 & 0 \\ 1 & 1-x & 0 \\ -1 & 1 & 4-x \end{pmatrix} = (x^2 - 3)(4 - x)$$

- Troviamo le radici del polinomio caratteristico, cioè le soluzioni di $(x^2-3)(4-x)=0$:

$$x_1 = \sqrt{3} \qquad x_2 = -\sqrt{3} \qquad x_3 = 4$$

  Tali radici sono gli autovalori di $A$. Poiché $A$ ha tre autovalori distinti, per il teorema precedente esiste una base di $\mathbb{R}^3$ costituita da autovettori di $A$. Dunque possiamo immediatamente rispondere ad una delle domande: la matrice $A$ è diagonalizzabile.

- Calcoliamo l'autospazio $V_4$ corrispondente all'autovalore 4; che è costituito dai vettori $(x, y, z) \in \mathbb{R}^3$ tali che:

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 4 \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

e quindi dalle soluzioni del sistema

$$\begin{cases} -x + 2y = 4x \\ x + y = 4y \\ -x + y + 4z = 4z \end{cases}$$

Notiamo che la matrice associata a questo sistema lineare è:

$$A = \begin{pmatrix} -5 & 2 & 0 \\ 1 & -3 & 0 \\ -1 & 1 & 0 \end{pmatrix}$$

che è la matrice $A - 4I$.
Riducendo la matrice con il metodo di Gauss, si ottiene:

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

Quindi le soluzioni del sistema dipendono da $3 - 2 = 1$ parametri. Si ricavano $x$ e $y$, mentre $z$ ha un valore arbitrario $s$. Si ottiene che $V_4 = \{(0, 0, s) \mid s \in \mathbb{R}\} = \langle(0, 0, 1)\rangle$ e ha dimensione 1.

- Calcoliamo l'autospazio $V_{\sqrt{3}}$ corrispondente all'autovalore $\sqrt{3}$; che è costituito dai vettori $(x, y, z) \in \mathbb{R}^3$ tali che:

$$A\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \sqrt{3}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

e quindi, analogamente a quanto visto prima, occorre trovare le soluzioni del sistema lineare omogeneo associato alla matrice $A - \sqrt{3}I$:

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

Riducendo la matrice con il metodo di Gauss, si ottiene

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

quindi $V_{\sqrt{3}} = \{((-1 - 3\sqrt{3})z, -(5 + 2\sqrt{3})z, z) \mid z \in \mathbb{R}\} = \langle(-1 - 3\sqrt{3}, -5 - 2\sqrt{3}, 1)\rangle$ e ha dimensione 1.

- Calcoliamo l'autospazio corrispondente all'autovalore $-\sqrt{3}$; che è costituito dai vettori $(x, y, z) \in \mathbb{R}^3$ tali che:

$$A\begin{pmatrix} x \\ y \\ z \end{pmatrix} = -\sqrt{3}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

quindi dalle soluzioni del sistema lineare omogeneo associato alla matrice $A + \sqrt{3}I$:

$$A = \begin{pmatrix} -1 + \sqrt{3} & 2 & 0 \\ 1 & 1 + \sqrt{3} & 0 \\ -1 & 1 & 4 + \sqrt{3} \end{pmatrix}$$

Riducendo la matrice con il metodo di Gauss, si ottiene

$$\begin{pmatrix} 1 & 1+\sqrt{3} & 0 \\ 0 & 2+\sqrt{3} & 4+\sqrt{3} \\ 0 & 0 & 0 \end{pmatrix}$$

quindi $V_{-\sqrt{3}} = \{((3\sqrt{3}-1)z, (2\sqrt{3}-5)z, z) \mid z \in \mathbb{R}\} = \langle(3\sqrt{3}-1, 2\sqrt{3}-5, 1)\rangle$ e ha dimensione 1.

- Una matrice diagonale $D$ simile ad $A$ è

$$D = \begin{pmatrix} 4 & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & -\sqrt{3} \end{pmatrix}$$

e si ha che $D = P^{-1}AP$, dove

$$P = \begin{pmatrix} 0 & -1-3\sqrt{3} & 3\sqrt{3}-1 \\ 0 & -5-2\sqrt{3} & 2\sqrt{3}-5 \\ 1 & 1 & 1 \end{pmatrix}$$

---

Torniamo ora alla trattazione teorica. Sappiamo che gli autovalori di una matrice $A$ sono le radici del suo polinomio caratteristico, e sappiamo che dire $p_A(\lambda) = 0$ equivale a dire $x - \lambda$ divide $p_A(x)$, ossia possiamo scrivere $p_A(x) = (x-\lambda)f(x)$, dove $f(x)$ è un polinomio in $x$.

Vogliamo ora essere ancora più precisi.

**Definizione 9.2.17** Sia $A$ una matrice quadrata. Uno scalare $\lambda$ si dice autovalore di $A$ di *molteplicità algebrica* $m$ se $p_A(x) = (x-\lambda)^m f(x)$, dove $f(x)$ è un polinomio tale che $f(\lambda) \neq 0$. In pratica $(x-\lambda)^m$ è la massima potenza di $x - \lambda$ che divide $p_A(x)$.

Se $\lambda$ è un autovalore di $A$, la dimensione di $\ker(A - \lambda I)$ si dice *molteplicità geometrica* di $\lambda$.

♦ **Osservazione 9.2.18** Osserviamo che, se $\lambda$ è un autovalore di $A$, per definizione c'è almeno un autovettore non nullo, quindi $\ker(A - \lambda I)$ ha dimensione almeno 1. Questo ci dice che la molteplicità geometrica di un autovalore è sempre maggiore o uguale a 1.

**Proposizione 9.2.19** *Se $A$ è una matrice quadrata e $\lambda$ è un autovalore di $A$, allora la molteplicità geometrica di $\lambda$ è minore o uguale alla sua molteplicità algebrica.*

**Dimostrazione** – Supponiamo che $A$ sia una matrice quadrata di ordine $n$ e sia $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ l'applicazione lineare associata ad $A$ rispetto alla base canonica. La molteplicità geometrica $s$ di $\lambda$ è la dimensione di $\ker(A - \lambda I) = \ker(T - \lambda \mathrm{id}) = V_\lambda$. Sia $\{\mathbf{v}_1, \dots, \mathbf{v}_s\}$ una base di $\ker(T - \lambda \mathrm{id})$ e completiamola a una base $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_s, \dots, \mathbf{v}_n\}$ di $\mathbb{R}^n$ (si noti che è sempre possibile farlo per il Teorema del completamento (4.2.1)). Poiché $T(\mathbf{v}_i) = \lambda \mathbf{v}_i$ per ogni $i = 1, \dots, s$

la matrice $A_{\mathcal{B}}$ associata a $T$ rispetto alla base ordinata $\mathcal{B}$ è del tipo:

$$A_{\mathcal{B}} = \begin{pmatrix} \lambda & 0 & \cdots & 0 & b_{1\,s+1} & \cdots & b_{1\,n} \\ 0 & \lambda & \cdots & 0 & b_{2\,s+1} & \cdots & b_{2\,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda & b_{s\,s+1} & \cdots & b_{s\,n} \\ 0 & 0 & \cdots & 0 & b_{s+1\,s+1} & \cdots & b_{s+1\,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & b_{n\,s+1} & \cdots & b_{n\,n} \end{pmatrix}$$

Osserviamo ora che per il Teorema 9.2.8 si ha che $p_A(x) = p_{A_{\mathcal{B}}}(x)$, perché le due matrici $A$ e $A_{\mathcal{B}}$ sono simili. Calcoliamo quindi $\det(A - xI) = \det(A_{\mathcal{B}} - xI)$ sviluppando secondo la prima colonna, e poi ancora secondo la prima colonna dell'unico minore di ordine $n-1$ a determinante non nullo che compare nella formula, e così via. Otteniamo:

$$\det(A_{\mathcal{B}} - xI)$$

$$= \det \begin{pmatrix} \lambda - x & 0 & \cdots & 0 & b_{1\,s+1} & \cdots & b_{1\,n} \\ 0 & \lambda - x & \cdots & 0 & b_{2\,s+1} & \cdots & b_{2\,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda - x & b_{s\,s+1} & \cdots & b_{s\,n} \\ 0 & 0 & \cdots & 0 & b_{s+1\,s+1} - x & \cdots & b_{s+1\,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & b_{n\,s+1} & \cdots & b_{n\,n} - x \end{pmatrix}$$

$$= (\lambda - x) \det \begin{pmatrix} \lambda - x & \cdots & 0 & b_{2\,s+1} & \cdots & b_{2\,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda - x & b_{s\,s+1} & \cdots & b_{s\,n} \\ 0 & \cdots & 0 & b_{s+1\,s+1} - x & \cdots & b_{s+1\,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & b_{n\,s+1} & \cdots & b_{n\,n} - x \end{pmatrix}$$

$$= (\lambda - x)^{s-1} \det \begin{pmatrix} \lambda - x & b_{s\,s+1} & \cdots & b_{s\,n} \\ 0 & b_{s+1\,s+1} - x & \cdots & b_{s+1\,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & b_{n\,s+1} & \cdots & b_{n\,n} - x \end{pmatrix}$$

$$= (\lambda - x)^{s} \det \begin{pmatrix} b_{s+1\,s+1} - x & \cdots & b_{s+1\,n} \\ \vdots & \ddots & \vdots \\ b_{n\,s+1} & \cdots & b_{n\,n} - x \end{pmatrix}$$

Posto

$$B = \begin{pmatrix} b_{s+1\,s+1} & \cdots & b_{s+1\,n} \\ \vdots & \ddots & \vdots \\ b_{n\,s+1} & \cdots & b_{n\,n} \end{pmatrix}$$

abbiamo ottenuto che $p_A(x) = p_{A_{\mathcal{B}}}(x) = (x - \lambda)^s p_B(x)$, cioè che $(x-\lambda)^s$ divide $p_A(x)$. Poiché la molteplicità algebrica $m$ di $\lambda$ è la massima potenza di $x - \lambda$

che divide $p_A(x)$, abbiamo che $s \leq m$, cioè la molteplicità geometrica di $\lambda$ è minore o uguale alla sua molteplicità algebrica. □

♦ **Osservazione 9.2.20** Combinando la precedente proposizione e l'Osservazione 9.2.18 si ottiene che, se $\lambda$ è un autovalore di una matrice e $\lambda$ ha molteplicità algebrica 1, allora la molteplicità geometrica di $\lambda$ è esattamente 1. Vedremo che ciò ci sarà di grande aiuto negli esercizi.

La seguente proposizione ci permette di avere una strategia precisa per capire se una certa matrice $n \times n$ sia diagonalizzabile.

**Proposizione 9.2.21** *Sia $A$ una matrice quadrata di ordine $n$ e siano $\lambda_1, \dots, \lambda_r$ i suoi autovalori distinti, con molteplicità geometriche rispettivamente $n_1, \dots, n_r$. Allora $A$ è diagonalizzabile se e solo se $n_1 + \cdots + n_r = n$.*

**Dimostrazione** – Sia $T$ l'applicazione lineare associata ad $A$ rispetto alla base canonica.

Supponiamo che $n_1 + \cdots + n_r = n$ e per ogni autovalore $\lambda_i$, fissiamo una base $\mathcal{B}_i = \{\mathbf{v}_{i1}, \dots, \mathbf{v}_{in_i}\}$ dell'autospazio $V_{\lambda_i}$. Mostriamo che allora l'unione $\mathcal{B} = \{\mathbf{v}_{11}, \dots, \mathbf{v}_{1n_1}, \dots, \mathbf{v}_{r1}, \dots, \mathbf{v}_{rn_r}\}$ di tali basi è una base di $\mathbb{R}^n$. Una volta mostrato questo, per la Proposizione 9.2.3 segue subito che $T$ è diagonalizzabile, perché $\mathcal{B}$ è una base di autovettori di $T$.

Poiché $n_1 + \cdots + n_r = n$ l'insieme $\mathcal{B}$ contiene $n$ vettori, quindi per la Proposizione 4.2.6 per dimostrare che $\mathcal{B}$ è una base di $\mathbb{R}^n$ è sufficiente dimostrare che i vettori di $\mathcal{B}$ sono linearmente indipendenti. Supponiamo che sia

$$\lambda_{11}\mathbf{v}_{11} + \cdots + \lambda_{1n_1}\mathbf{v}_{1n_1} + \cdots + \lambda_{r1}\mathbf{v}_{r1} + \cdots + \lambda_{rn_r}\mathbf{v}_{rn_r} = \mathbf{0} \tag{9.5}$$

Osserviamo che gli $r$ vettori $\mathbf{w}_1 = \lambda_{11}\mathbf{v}_{11} + \cdots + \lambda_{1n_1}\mathbf{v}_{1n_1}, \dots, \mathbf{w}_r = \lambda_{r1}\mathbf{v}_{r1} + \cdots + \lambda_{rn_r}\mathbf{v}_{rn_r}$ appartengono rispettivamente agli autospazi $V_{\lambda_1}, \dots, V_{\lambda_r}$. Osserviamo inoltre che l'uguaglianza (9.5) si può leggere come:

$$1 \cdot \mathbf{w}_1 + 1 \cdot \mathbf{w}_2 + \cdots + 1 \cdot \mathbf{w}_r = \mathbf{0} \tag{9.6}$$

Se alcuni dei $\mathbf{w}_i$ fossero non nulli, dall'uguaglianza (9.6) tali $\mathbf{w}_i$ sarebbero linearmente dipendenti, ma questo contraddice il fatto che autovettori relativi ad autovalori distinti sono linearmente indipendenti (Teorema 9.2.14). Quindi $\mathbf{w}_i = \mathbf{0}$ per ogni $i = 1, \dots, r$, cioè $\lambda_{i1}\mathbf{v}_{i1} + \cdots + \lambda_{in_1}\mathbf{v}_{in_1} = \mathbf{0}$, per ogni $i = 1, \dots, r$. Sfruttiamo ora il fatto che i vettori $\mathbf{v}_{i1}, \dots, \mathbf{v}_{in_i}$ sono linearmente indipendenti, perché costituiscono una base di $V_{\lambda_i}$, e otteniamo che: $\lambda_{i1} = \cdots = \lambda_{in_i} = 0$. Dunque nella combinazione lineare al primo membro dell'uguaglianza (9.5) i coefficienti devono essere tutti nulli. Questo dimostra che i vettori di $\mathcal{B}$ sono linearmente indipendenti, quindi $\mathcal{B}$ è una base di autovettori. Questo mostra che $A$ è diagonalizzabile.

Viceversa, supponiamo che $A$ sia diagonalizzabile e sia $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ una base di $\mathbb{R}^n$ costituita da autovettori di $T$. La matrice $D$ associata a $T$ rispetto alla base $\mathcal{B}$ è una matrice diagonale $D$ simile ad $A$, e il suo polinomio caratteristico è il prodotto di $n$ fattori lineari del tipo $x - \lambda_i$, ove $\lambda_i$ è un autovalore di $A$. Per la Proposizione 9.2.8, matrici simili hanno lo stesso polinomio caratteristico, quindi $p_A(x) = p_D(x) = (x - \lambda_1)^{m_1} \cdots (x - \lambda_r)^{m_r}$ dove $m_i$ è la molteplicità algebrica dell'autovalore $\lambda_i$ ma è anche il numero di autovettori distinti di $\mathcal{B}$ che hanno autovalore $\lambda_i$, quindi $m_i \leq n_i = \dim(V_{\lambda_i})$ per ogni $i = 1, \dots, r$. Si osservi inoltre che $m_1 + \dots + m_r = n$. Poiché per la Proposizione 9.2.19 abbiamo che $n_i \leq m_i$, segue che $n_i = m_i$ per ogni $i = 1, \dots, r$. Quindi $n_1 + \dots + n_r = n$, come volevasi dimostrare. □

Possiamo ora descrivere come si può procedere per stabilire se una matrice quadrata $A$ di ordine $n$ sia diagonalizzabile.

- Calcoliamo le radici del polinomio caratteristico. Se otteniamo $n$ radici distinte, allora abbiamo $n$ autovalori distinti, corrispondenti a $n$ autovettori linearmente indipendenti, quindi $A$ è diagonalizzabile.
- Se gli autovalori non sono distinti, per ciascun autovalore $\lambda$ calcoliamo la molteplicità geometrica. Se la somma di tutte le molteplicità geometriche è proprio $n$, ciò ci permetterà di trovare $n$ autovettori linearmente indipendenti e quindi una base di $V$, di conseguenza $A$ è diagonalizzabile. Se la somma di tutte le molteplicità geometriche è minore di $n$, allora $A$ non è diagonalizzabile.

Concludiamo con un'ultima osservazione, che mette in relazione il calcolo del nucleo di un'applicazione lineare e il fatto che abbia un autovalore nullo.

♦ **Osservazione 9.2.22** Supponiamo che l'applicazione lineare $T : \mathbb{R}^3 \to \mathbb{R}^3$ abbia un autovalore uguale a zero. Cosa possiamo dire di $T$? Si ha che l'autospazio $V_0$ ha dimensione almeno 1, per l'Osservazione 9.2.18. Inoltre, se $A$ è la matrice associata a $T$ rispetto alla base canonica, si ha che $V_0$ è l'insieme delle soluzioni del sistema lineare omogeneo associato alla matrice $A - 0I = A$, cioè $V_0$ è proprio il nucleo di $T$. Quindi, se l'applicazione lineare $T$ ha un autovalore uguale a zero, possiamo dire che $V_0 = \ker T$ e $T$ non è iniettiva.

## 9.3 ESERCIZI SVOLTI

### Esercizio 9.3.1

*Si stabilisca se l'applicazione lineare $L : \mathbb{R}^2 \to \mathbb{R}^2$ definita da $L(x, y) = (x - y, x + 3y)$ è diagonalizzabile. Si determinino inoltre, se possibile, una matrice diagonale $D$ che sia simile ad $A$ e una matrice $B$ che non sia simile ad $A$.*

**Svolgimento**

Determiniamo prima la matrice associata a $L$ rispetto alla base canonica:

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

Ora determiniamo il polinomio caratteristico di questa matrice:

$$p_A(x) = \det\begin{pmatrix} 1-x & -1 \\ 1 & 3-x \end{pmatrix} = x^2 - 4x + 4 = (x-2)^2$$

Il polinomio caratteristico ha come unica radice $x = 2$, con molteplicità algebrica 2. Quindi $L$ ha un solo autovalore.

Per trovare l'autospazio di autovalore 2, risolviamo il sistema lineare omogeneo associato alla matrice $A - 2I$:

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

Si ottiene che $V_2 = \langle(-1,1)\rangle$ ha dimensione 1. Quindi non è possibile trovare una base di $\mathbb{R}^2$ costituita da autovettori e la matrice $A$ non è diagonalizzabile.

Poiché $A$ non è diagonalizzabile, non è possibile trovare nessuna matrice diagonale $D$ che sia simile ad $A$. In particolare, se prendiamo per esempio $B = \begin{pmatrix} -1 & 0 \\ 0 & 5 \end{pmatrix}$, sicuramente $B$ non è simile ad $A$. ∎

### Esercizio 9.3.2

*Sia $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ l'applicazione lineare definita da $L(\mathbf{e}_1) = 8\mathbf{e}_1 + 3\mathbf{e}_2$, $L(\mathbf{e}_2) = -18\mathbf{e}_1 - 7\mathbf{e}_2$, $L(\mathbf{e}_3) = 9\mathbf{e}_1 + 3\mathbf{e}_2 - \mathbf{e}_3$. Si stabilisca se $A$ è diagonalizzabile e, in caso affermativo, si determini una matrice $D$ simile ad $A$, dove $A$ è la matrice associata ad $L$ rispetto alla base canonica. Si determinino inoltre, se possibile, due matrici distinte $P_1$ e $P_2$ tali che $P_1^{-1}AP_1 = P_2^{-1}AP_2 = D$.*

**Svolgimento**

La matrice $A$ associata a $L$ rispetto alla base canonica è:

$$A = \begin{pmatrix} 8 & -18 & 9 \\ 3 & -7 & 3 \\ 0 & 0 & -1 \end{pmatrix}$$

Calcoliamo ora il polinomio caratteristico di $A$:

$$p_A(x) = \det\begin{pmatrix} 8-x & -18 & 9 \\ 3 & -7-x & 3 \\ 0 & 0 & -1-x \end{pmatrix} = -(1+x)(x^2 - x - 2)$$

Poi troviamo le radici del polinomio caratteristico, cioè le soluzioni di $(1 + x)(x^2 - x - 2) = 0$, ossia $(1+x)^2(x-2) = 0$:

$$x_1 = -1 \qquad x_2 = 2$$

Tali radici sono gli autovalori di $A$. Notiamo che l'autovalore $-1$ ha molteplicità algebrica 2 e l'autovalore 2 ha molteplicità algebrica 1.

Per l'Osservazione 9.2.20 sappiamo che l'autospazio $V_2$ ha dimensione 1, mentre per la Proposione 9.2.19 di $V_{-1}$ possiamo dire solo che ha dimensione al massimo 2. Se $V_{-1}$ ha dimensione 2, allora per la Proposizione 9.2.21 si ha che $A$ è diagonalizzabile, se invece $V_{-1}$ ha dimensione minore di 2, allora non è possibile trovare una base di $\mathbb{R}^3$ costituita da autovettori e quindi $A$ non è diagonalizzabile.

Calcoliamo quindi l'autospazio corrispondente all'autovalore $x = -1$; che è costituito dalle soluzioni del sistema lineare omogeneo associato alla matrice $A + I$:

$$\begin{pmatrix} 8+1 & -18 & 9 \\ 3 & -7+1 & 3 \\ 0 & 0 & -1+1 \end{pmatrix}$$

Riducendo la matrice con il metodo di Gauss, si ottiene

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

quindi le soluzioni del sistema dipendono da $3 - 1 = 2$ parametri. Si ricava $x$, mentre $y, z$ possono assumere valori arbitrari $s$ e $t$. Si ottiene che $V_{-1} = \{(2s - t, s, t) \mid s \in \mathbb{R}\} = \langle(2,1,0),(-1,0,1)\rangle$ e ha dimensione 2.

Quindi $A$ è diagonalizzabile.

Le matrici $P_i$ richieste sono matrici le cui colonne devono essere costituite dalle coordinate di una base di autovettori (rispetto alla base canonica), perciò è necessario determinare anche l'autospazio $V_2$.

Dobbiamo risolvere il sistema lineare omogeneo associato alla matrice $A-2I$:

$$\begin{pmatrix} 8-2 & -18 & 9 \\ 3 & -7-2 & 3 \\ 0 & 0 & -1-2 \end{pmatrix}$$

Riducendo la matrice con il metodo di Gauss, si ottiene

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

quindi le soluzioni del sistema dipendono da $3 - 2 = 1$ parametro. Si ricavano $x, z$, mentre $y$ ha un valore arbitrario $s$. Si ottiene che $V_2 = \{(3s, s, 0) \mid s \in \mathbb{R}\} = \langle\{(3,1,0)\rangle$ ed ha dimensione 1, come ci aspettavamo.

Ora una base ordinata di autovettori è per esempio $\mathcal{B}_1 = \{(2,1,0),(-1,0,1),(3,1,0)\}$ e la matrice associata a $L$ rispetto a tale base ordinata ha sulla diagonale gli autovalori corrispondenti ai vettori di base, cioè:

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

Una matrice $P_1$ tale che $P_1^{-1}AP_1 = D$ è per esempio la matrice di cambio di base $I_{\mathcal{B}_1,\mathcal{C}}$:

$$P_1 = \begin{pmatrix} 2 & -1 & 3 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

Se vogliamo un'altra matrice $P_2$ tale che $P_2^{-1}AP_2 = D$, dobbiamo scegliere un'altra base ordinata di $\mathbb{R}^3$ costituita da autovettori, facendo in modo che ai primi due posti ci siano sempre due autovettori di autovalore $-1$ e al terzo posto ci sia un autovettore di autovalore 2, per esempio $\mathcal{B}_1 = \{(-3,0,3),(2,1,0),(-3,-1,0)\}$. In tal caso la matrice di cambio di base $P_2 = I_{\mathcal{B}_2,\mathcal{C}}$ è:

$$P_2 = \begin{pmatrix} -3 & 2 & -3 \\ 0 & 1 & -1 \\ 3 & 0 & 0 \end{pmatrix}$$

■

**Esercizio 9.3.3**

*Si stabilisca se la matrice*

$$A = \begin{pmatrix} -8 & 18 & 2 \\ -3 & 7 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

*è diagonalizzabile.*

**Svolgimento**
Determiniamo il polinomio caratteristico di $A$:

$$p_A(x) = \det\begin{pmatrix} -8-x & 18 & 2 \\ -3 & 7-x & 1 \\ 0 & 0 & 1-x \end{pmatrix} = (1-x)(x^2+x-2)$$

Troviamo le radici del polinomio caratteristico, cioè le soluzioni di $(1-x)(x^2+x-2)=0$, ossia $(1-x)^2(2+x)=0$:

$$x_1 = 1 \qquad x_1 = -2$$

e tali radici sono gli autovalori di $A$. Notiamo che l'autovalore 1 ha molteplicità algebrica 2 e l'autovalore $-2$ ha molteplicità algebrica 1.

Come nell'esercizio precedente, sappiamo che l'autospazio $V_{-2}$ ha dimensione 1, mentre per determinare la dimensione di $V_1$ bisogna calcolare esplicitamente tale autospazio.

$V_1$ è costituito dalle soluzioni del sistema lineare omogeneo associato alla matrice $A - I$:

$$\begin{pmatrix} -8-1 & 18 & 2 \\ -3 & 7-1 & 1 \\ 0 & 0 & 1-1 \end{pmatrix}$$

Riducendo la matrice con il metodo di Gauss, si ottiene

$$\begin{pmatrix} 1 & -2 & -1/9 \\ 0 & -12 & 1/3 \\ 0 & 0 & 0 \end{pmatrix}$$

quindi le soluzioni del sistema dipendono da $3-2=1$ parametro. Dunque $V_1$ ha dimensione 1.

In questo caso $A$ non è diagonalizzabile, perché non abbiamo una base di autovettori. Questo dipende dal fatto che l'autovalore 1 ha molteplicità algebrica 2, ma $V_1$ ha dimensione 1 e non 2. ■

**Esercizio 9.3.4**

*Si stabilisca per quali valori di $k$ la seguente matrice:*

$$A = \begin{pmatrix} 1 & 1 & 1+k \\ 2 & 2 & 2 \\ 0 & 0 & -k \end{pmatrix}$$

*è diagonalizzabile. Si stabilisca inoltre, per quali valori di $k$, 5 è un autovalore di $A$.*

**Svolgimento**
Determiniamo il polinomio caratteristico di $A$:

$$p_A(x) = \det\begin{pmatrix} 1-x & 1 & 1+k \\ 2 & 2-x & 2 \\ 0 & 0 & -k-x \end{pmatrix} = (x^2-3x)(-k-x)$$

Le radici del polinomio caratteristico sono:

$$x_1 = 0 \qquad x_2 = 3 \qquad x_3 = -k$$

e tali radici sono gli autovalori di $A$. Se $k \neq 0$ e $k \neq -3$ otteniamo che $A$ ha 3 autovalori distinti, quindi è diagonalizzabile.

Se $k = 0$ si ottiene che

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

e, per quanto appena visto, sappiamo che $A$ ha autovalori 0 e $-3$ con molteplicità algebrica 2 e 1, rispettivamente. Per l'Osservazione 9.2.20 sappiamo che l'autospazio $V_{-3}$ ha dimensione 1, mentre per determinare la dimensione di $V_0$ bisogna calcolare esplicitamente tale autospazio.

$V_0$ è costituito dalle soluzioni del sistema lineare omogeneo associato alla matrice $A - 0I = A$, cioè $V_0 = \ker A$. Riducendo la matrice con il metodo di Gauss, si ottiene

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

quindi le soluzioni del sistema dipendono da $3 - 1 = 2$ parametri. Dunque $V_0$ ha dimensione 2 e $A$ è diagonalizzabile.

Se $k = -3$ si ottiene che

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

e, per quanto appena visto, sappiamo che $A$ ha autovalori 0 e 3 con molteplicità algebrica 1 e 2, rispettivamente. Come prima, per stabilire se $A$ è diagonalizzabile, dobbiamo determinare la dimensione dell'autospazio relativo all'autovalore di molteplicità algebrica 2, cioè $V_3$.

Dobbiamo risolvere il sistema lineare omogeneo associato alla matrice:

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

Riducendo la matrice con il metodo di Gauss, si ottiene

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

quindi le soluzioni del sistema dipendono da $3 - 1 = 2$ parametri. Dunque $V_3$ ha dimensione 2 e $A$ è diagonalizzabile.

Riassumendo, $A$ è diagonalizzabile per ogni valore di $k$.

Poiché gli autovalori di $A$ sono $0, 3$ e $-k$, si ha che 5 è un autovalore di $A$ se se solo se $k = -5$. ∎

## 9.4 ESERCIZI PROPOSTI

### Esercizio 9.4.1

*Si trovino autovalori e autovettori delle seguenti matrici o applicazioni lineari:*

a) *la matrice:*

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

b) *l'applicazione lineare* $L : \mathbb{R}^2 \to \mathbb{R}^2$ *definita da:*

$$L(x, y) = (2x + y, 2x + 3y)$$

c) *l'applicazione lineare* $L : \mathbb{R}^3 \to \mathbb{R}^3$ *definita da:*

$$L(x, y, z) = (x + y, x + z, y + z)$$

d) *l'applicazione lineare* $L : \mathbb{R}^2 \to \mathbb{R}^2$ *definita da:*

$$L(x, y) = (x - 3y, -2x + 6y)$$

e) *L'applicazione lineare* $L : \mathbb{R}^2 \to \mathbb{R}^2$ *definita da:*

$$L(\mathbf{e}_1) = \mathbf{e}_1 - \mathbf{e}_2, \ \ L(\mathbf{e}_2) = 2\mathbf{e}_1$$

**Esercizio 9.4.2**

*Sia data la matrice:*

$$A = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 7 & -1 \\ 0 & 14 & -2 \end{pmatrix}$$

a) *Si calcolino autovalori e autovettori.*
*$A$ è diagonalizzabile? In caso affermativo, si determini una matrice diagonale $A'$ simile ad $A$.*
b) *È possibile trovare una matrice $B$ tale che $AB = I$ (dove $I$ è la matrice identità)? Si motivi accuratamente la risposta.*

**Esercizio 9.4.3**

*Si determini un'applicazione lineare* $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *che abbia* $\mathbf{e}_1 - \mathbf{e}_2$ *come autovettore di autovalore 2.*

**Esercizio 9.4.4**

*Sia data la matrice:*

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

*Si calcolino autovalori e autovettori.*
*$A$ è diagonalizzabile? In caso affermativo, si determinino tutte le matrici diagonali simili ad $A$. Si determini inoltre una matrice diagonale $D$ che non sia simile ad $A$.*

**Esercizio 9.4.5**

*Sia data la matrice:*

$$A = \begin{pmatrix} -4 & -6 & 1 \\ 3 & 5 & -2 \\ 0 & 0 & -1 \end{pmatrix}$$

a) *Si calcolino autovalori e autovettori.*
b) *$A$ è diagonalizzabile? In caso affermativo, si scriva una matrice diagonale $A'$ simile ad $A$.*
c) *$A$ è invertibile?*

### Esercizio 9.4.6

*Sia data la matrice:*

$$A = \begin{pmatrix} 9 & 0 & 0 \\ 6 & 3 & 6 \\ 0 & 0 & 9 \end{pmatrix}$$

a) *Si calcolino autovalori e autovettori.*
b) *Si stabilisca se $A$ è diagonalizzabile e in caso affermativo si determinino una matrice diagonale $D$ simile ad $A$ e una matrice $P$ tale che $P^{-1}AP = D$. Una volta fissata $D$, $P$ è unica?*

### Esercizio 9.4.7

*Sia data la matrice:*

$$A = \begin{pmatrix} 6 & -9 & -3 \\ 0 & -2 & 2 \\ 0 & -1 & 1 \end{pmatrix}$$

a) *Si calcolino gli autovalori e autovettori.*
b) *Si determinino, se possibile, due matrici $P_1$ e $P_2$ tale che $P_1^{-1}AP_1$ e $P_2^{-1}AP_2$ sia diagonale.*

### Esercizio 9.4.8

*Sia data l'applicazione lineare $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definita da: $T(\mathbf{e}_1) = \mathbf{e}_1 - 2\mathbf{e}_2$, $T(\mathbf{e}_2) = 2\mathbf{e}_1 + 6\mathbf{e}_2$, $T(\mathbf{e}_3) = 3\mathbf{e}_1 - \mathbf{e}_2 + 5\mathbf{e}_3$.*
a) *Si calcolino gli autovalori e un autospazio di $T$.*
b) *Si stabilisca se $T$ è diagonalizzabile.*
c) *Sia $\mathcal{B} = \{3\mathbf{e}_2+\mathbf{e}_3, 5\mathbf{e}_2+\mathbf{e}_3, -\mathbf{e}_1+\mathbf{e}_3\}$ un'altra base ordinata di $\mathbb{R}^3$. Si determini la matrice $A_{\mathcal{B}}$ associata a $T$ rispetto alla base $\mathcal{B}$ (nel dominio e nel codominio).*

### Esercizio 9.4.9

*Sia data l'applicazione lineare $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definita da: $T(x, y, z) = (-3x + 6y, -x + 4y, x - 6y - 2z)$.*
a) *Si calcolino gli autovalori e un autospazio di $T$.*
b) *Si stabilisca se $T$ è diagonalizzabile e si determini, se possibile, una base ordinata $\mathcal{B}$ di $\mathbb{R}^3$ tale che la matrice $A_{\mathcal{B}}$ associata a $T$ 2351182 rispetto alla base $\mathcal{B}$ sia diagonale.*

### Esercizio 9.4.10

*Sia $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ l'applicazione lineare definita da: $T(\mathbf{e}_1) = (k+1)\mathbf{e}_1 - \mathbf{e}_3$, $T(\mathbf{e}_2) = k\mathbf{e}_2 + (k+1)\mathbf{e}_3$, $T(\mathbf{e}_3) = k\mathbf{e}_3$ e sia $A$ la matrice associata a $T$ rispetto alla base canonica.*
a) *Si stabilisca per quali valori di $k$ $T$ è diagonalizzabile.*
b) *Si stabilisca per quali valori di $k$ $2\mathbf{e}_1 - 2\mathbf{e}_3$ è un autovettore di $T$.*

c) *Per i valori di $k$ trovati al punto a) si determinino, se possibile, due matrici diagonali distinte $D_1$ e $D_2$ simili ad $A$. Si determini inoltre una matrice $P$ tale che $P^{-1}AP = D_1$.*

### Esercizio 9.4.11

*Sia data la matrice:*

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

a) *Si determini per quali valori di $k$, $A$ è invertibile.*
b) *Si stabilisca per quali valori di $k$, $A$ è diagonalizzabile.*

### Esercizio 9.4.12

*Sia $F : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ l'applicazione lineare definita da: $F(x,y,z) = (2x + 2y + z, 2x - y - 2z, kz)$ e sia $A$ la matrice associata a $F$ rispetto alla base canonica.*
a) *Si stabilisca per quali valori di $k$ $F$ è diagonalizzabile.*
b) *Scelto a piacere un valore di $k$ per cui $F$ è diagonalizzabile, si determinino tutte le matrici diagonali $D$ simili ad $A$.*

### Esercizio 9.4.13

*Sia data la matrice:*

$$A = \begin{pmatrix} k & 0 & 0 \\ k & 2 & 3 \\ k & 3 & 2 \end{pmatrix}$$

a) *Si stabilisca per quali valori di $k$ $A$ è diagonalizzabile.*
b) *Si stabilisca per quali valori di $k$ $-2\mathbf{e}_1$ è un autovettore di $A$.*
c) *Posto $k = 3$ si determini, se possibile, una matrice $B \in \mathrm{M}_3(\mathbb{R})$ avente gli stessi autovalori di $A$, tale che $B$ non sia simile ad $A$.*

# 10 Prodotti scalari

Nella definizione di spazio vettoriale, compaiono le due operazioni di somma di vettori e moltiplicazione di un vettore per uno scalare (si veda il Capitolo 2). In questo capitolo vogliamo introdurre una nuova operazione: il *prodotto scalare* di due vettori. Tale operazione ha per risultato uno *scalare* e cioè un numero reale. Oltre ad avere una enorme importanza nelle applicazioni alla fisica, vedremo come il prodotto scalare sia indispensabile in algebra lineare nella soluzione del problema della diagonalizzazione delle matrici simmetriche, che tratteremo più avanti.

## 10.1 FORME BILINEARI

In questa sezione vogliamo introdurre il concetto di forma o applicazione bilineare e studiare la corrispondenza biunivoca che esiste tra l'insieme delle forme bilineari su di uno spazio vettoriale $V$ di dimensione $n$ e l'insieme delle matrici di ordine $n$, una volta che sia stata fissata una base ordinata per $V$.

**Definizione 10.1.1** Sia $V$ uno spazio vettoriale. La funzione $g : V \times V \longrightarrow \mathbb{R}$ si dice una *forma bilineare* o una *applicazione bilineare* se:

1. $g(\mathbf{u}+\mathbf{u}',\mathbf{v}) = g(\mathbf{u},\mathbf{v}) + g(\mathbf{u}',\mathbf{v})$,
   $g(\mathbf{u},\mathbf{v}+\mathbf{v}') = g(\mathbf{u},\mathbf{v}) + g(\mathbf{u},\mathbf{v}')$ per ogni $\mathbf{u}$, $\mathbf{u}'$, $\mathbf{v}$, $\mathbf{v}' \in V$.

2. $g(\lambda\mathbf{u},\mathbf{v}) = \lambda g(\mathbf{u},\mathbf{v})$,
   $g(\mathbf{u},\mu\mathbf{v}) = \mu g(\mathbf{u},\mathbf{v})$ per ogni $\mathbf{u}$, $\mathbf{v} \in V$ e per ogni $\lambda$, $\mu \in \mathbb{R}$.

In altre parole, per ogni vettore fissato $\mathbf{u} \in V$ le funzioni $g(\mathbf{u},\cdot) : V \longrightarrow \mathbb{R}$ e $g(\cdot,\mathbf{u}) : V \longrightarrow \mathbb{R}$ sono applicazioni lineari, da cui il termine *bilineare*.

$g$ si dice *simmetrica* se $g(\mathbf{u},\mathbf{v}) = g(\mathbf{v},\mathbf{u})$ per ogni $\mathbf{u}, \mathbf{v} \in V$. Una forma bilineare simmetrica si dice un *prodotto scalare* su $V$ e sarà denotata con $<\,,\,>$.

Ritorneremo, con maggiori dettagli, sulla definizione di prodotto scalare nel Paragrafo 10.4.

**Esempio 10.1.2**

Definiamo la funzione

$$g : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}, \qquad g((x_1, x_2), (y_1, y_2)) = x_1 y_1 + 2x_1 y_2.$$

Verifichiamo la proprietà (1) della definizione precedente.

$$\begin{aligned} g((x_1, x_2) + (x_1', x_2'), (y_1, y_2)) &= (x_1 + x_1')y_1 + 2(x_1 + x_1')y_2 \\ &= g((x_1, x_2), (y_1, y_2)) + g((x_1', x_2'), (y_1, y_2)) \end{aligned}$$

$$g(\lambda(x_1, x_2), (y_1, y_2)) = \lambda x_1 y_1 + 2\lambda x_1 y_2 = \lambda(g((x_1, x_2), (y_1, y_2)))$$

In modo del tutto analogo possiamo verificare anche la proprietà (2). Si tratta dunque di una forma bilineare. Si osservi tuttavia che $g(\mathbf{e}_1, \mathbf{e}_2) = 2$ mentre $g(\mathbf{e}_2, \mathbf{e}_1) = 0$, dunque $g$ non è un prodotto scalare.

Vediamo ora un altro esempio, importante nelle applicazioni alla fisica.

**Esempio 10.1.3**

Definiamo in $\mathbb{R}^n$ la funzione:

$$g((x_1, \dots, x_n), (y_1, \dots, y_n)) = x_1 y_1 + \dots + x_n y_n.$$

per ogni $\mathbf{u} = (x_1, \dots, x_n)$ e $\mathbf{v} = (y_1, \dots, y_n)$ in $\mathbb{R}^n$.

Lasciamo al lettore la facile verifica delle proprietà (1) e (2) della Definizione 10.1.1. Si tratta quindi di una forma bilineare. Poiché $g(\mathbf{u}, \mathbf{v}) = g(\mathbf{v}, \mathbf{u})$, $g$ è un prodotto scalare. Tale prodotto scalare su $\mathbb{R}^n$ si dice *prodotto euclideo* o *prodotto standard*. Indicheremo questo prodotto scalare tra due vettori $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ come $< \mathbf{u}, \mathbf{v} >_e$ o anche come $\mathbf{u} \cdot \mathbf{v}$.

In modo analogo a quanto avviene per le applicazioni lineari, un'applicazione bilineare è completamente determinata dai valori che assume sulle coppie degli elementi di una base fissata.

**Proposizione 10.1.4** *Sia $V$ uno spazio vettoriale di dimensione $n$ e fissiamo in $V$ una base $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. Siano $c_{ij} \in \mathbb{R}$, $i, j = 1, \dots, n$, scalari arbitrari fissati. Allora esiste ed è unica l'applicazione bilineare $g$ tale che*

$$g(\mathbf{v}_i, \mathbf{v}_j) = c_{ij}$$

**Dimostrazione** – Dimostriamo prima l'esistenza di $g$. Poiché ogni vettore di $V$ si esprime in modo unico come combinazione lineare degli elementi della base fissata $\mathcal{B}$, dati $\mathbf{u}$ e $\mathbf{w}$ in $V$ possiamo scrivere:

$$\mathbf{u} = \alpha_1 \mathbf{v}_1 + \dots + \alpha_n \mathbf{v}_n, \qquad \mathbf{w} = \beta_1 \mathbf{v}_1 + \dots + \beta_n \mathbf{v}_n. \tag{10.1}$$

Definiamo pertanto la funzione $g : V \times V \longrightarrow \mathbb{R}$ nel modo seguente:

$$g(\mathbf{u}, \mathbf{w}) = \sum_{i,j=1}^{n} \alpha_i \beta_j c_{ij} \tag{10.2}$$

ove abbiamo utilizzato il simbolo $\sum_{i,j=1}^{n}$ per indicare la somma per tutti i possibili $i, j = 1, \dots n$. Per esteso:

$$\begin{aligned} g(\mathbf{u}, \mathbf{w}) &= \alpha_1\beta_1 c_{11} + \alpha_1\beta_2 c_{12} + \dots \\ &\quad + \alpha_1\beta_n c_{1n} + \alpha_2\beta_1 c_{21} + \dots + \alpha_n\beta_n c_{nn}. \end{aligned}$$

Ciò implica che $g(\mathbf{v}_i, \mathbf{v}_j) = c_{ij}$. Dobbiamo ora verificare che si tratti di una applicazione bilineare e cioè che verifichi le proprietà della Definizione 10.1.1. Verifichiamo la prima delle condizioni in (1) lasciando le altre per esercizio.

Si considerino i tre vettori in $V$:

$$\mathbf{u} = \alpha_1\mathbf{v}_1 + \dots + \alpha_n\mathbf{v}_n, \ \mathbf{u}' = \alpha_1'\mathbf{v}_1 + \dots + \alpha_n'\mathbf{v}_n, \ \mathbf{w} = \beta_1\mathbf{v}_1 + \dots + \beta_n\mathbf{v}_n.$$

Per definizione di $g$ abbiamo che:

$$\begin{aligned} g(\mathbf{u} + \mathbf{u}', \mathbf{w}) &= \sum_{i,j=1}^{n} (\alpha_i + \alpha_i')\beta_j c_{ij} \\ &= \sum_{i,j=1}^{n} \alpha_i\beta_j c_{ij} + \sum_{i,j=1}^{n} \alpha_i'\beta_j c_{ij} \\ &= g(\mathbf{u}, \mathbf{w}) + g(\mathbf{u}', \mathbf{w}) \end{aligned}$$

Analogamente possiamo dimostrare anche le altre tre proprietà.

Veniamo ora all'unicità. Supponiamo che esista un'altra forma bilineare $\tilde{g}$ tale che $g(\mathbf{v}_i, \mathbf{v}_j) = \tilde{g}(\mathbf{v}_i, \mathbf{v}_j)$. Vogliamo dimostrare che $g = \tilde{g}$ cioè $g(\mathbf{u}, \mathbf{w}) = \tilde{g}(\mathbf{u}, \mathbf{w})$ per ogni $\mathbf{u}$, $\mathbf{w}$ in $V$. Se esprimiamo $\mathbf{u}$ e $\mathbf{w}$ come combinazioni lineari degli elementi della base $\mathcal{B}$ come in (10.1), grazie alle proprietà (1) e (2) della Definizione 10.1.1 possiamo scrivere:

$$\begin{aligned} \tilde{g}(\mathbf{u}, \mathbf{w}) &= \tilde{g}(\alpha_1\mathbf{v}_1 + \dots + \alpha_n\mathbf{v}_n, \beta_1\mathbf{v}_1 + \dots + \beta_n\mathbf{v}_n) \\ &= \alpha_1\beta_1\tilde{g}(\mathbf{v}_1, \mathbf{v}_1) + \alpha_1\beta_2\tilde{g}(\mathbf{v}_1, \mathbf{v}_2) + \dots + \alpha_1\beta_n\tilde{g}(\mathbf{v}_1, \mathbf{v}_n) \\ &\quad + \alpha_2\beta_1\tilde{g}(\mathbf{v}_2, \mathbf{v}_1) + \dots + \alpha_n\beta_n\tilde{g}(\mathbf{v}_n, \mathbf{v}_n) \\ &= \sum_{i,j=1}^{n} \alpha_i\beta_j\tilde{g}(\mathbf{v}_i, \mathbf{v}_j) \end{aligned}$$

Poiché $\tilde{g}(\mathbf{v}_i, \mathbf{v}_j) = g(\mathbf{v}_i, \mathbf{v}_j)$, per la definizione stessa (vedi (10.2)) di $\tilde{g}$ otteniamo $\tilde{g}(\mathbf{u}, \mathbf{w}) = g(\mathbf{u}, \mathbf{w})$. □

## 10.2 FORME BILINEARI E MATRICI

In questa sezione, vogliamo dimostrare, in modo del tutto analogo a quanto avviene per le applicazioni lineari, che abbiamo una corrispondenza biunivoca tra l'insieme delle applicazioni bilineari su uno spazio vettoriale $V$ di dimensione $n$ e l'insieme delle matrici quadrate di ordine $n$, una volta che sia stata fissata una base ordinata di $V$. Questa corrispondenza è fondamentale per poter fare qualunque calcolo riguardante forme bilineari e prodotti scalari.

Prima di procedere osserviamo una proprietà delle matrici che ci sarà particolarmente utile, la cui verifica è un facile calcolo. La trasposta di un prodotto (righe per colonne) di matrici è il prodotto delle matrici trasposte, invertendo però l'ordine dei fattori. In formule: se $A \in \mathrm{M}_{m,r}(\mathbb{R})$, $B \in \mathrm{M}_{r,n}(\mathbb{R})$, allora:

$$(AB)^T = B^T A^T \tag{10.3}$$

Continuiamo ora la nostra trattazione sulla corrispondenza biunivoca tra forme bilineari e matrici, una volta fissata una base dello spazio vettoriale dato.

**Proposizione 10.2.1** *Sia $V$ uno spazio vettoriale reale di dimensione $n$ e sia $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ una base ordinata fissata. Esiste una corrispondenza biunivoca tra l'insieme delle forme bilineari $g : V \times V \longrightarrow \mathbb{R}$ e l'insieme $\mathrm{M}_n(\mathbb{R})$. In tale corrispondenza:*

- *Alla forma bilineare $g : V \times V \longrightarrow \mathbb{R}$ è associata la matrice:*

$$C = \begin{pmatrix} g(\mathbf{v}_1, \mathbf{v}_1) & \dots & g(\mathbf{v}_1, \mathbf{v}_n) \\ \vdots & \ddots & \vdots \\ g(\mathbf{v}_n, \mathbf{v}_1) & \dots & g(\mathbf{v}_n, \mathbf{v}_n) \end{pmatrix}$$

- *Alla matrice $C \in \mathrm{M}_n(\mathbb{R})$ è associata la forma bilineare*

$$g(\mathbf{u}, \mathbf{v}) = (\mathbf{u})_{\mathcal{B}}^T C(\mathbf{v})_{\mathcal{B}}, \qquad \forall \mathbf{u}, \mathbf{v} \in V$$

*ove $(\mathbf{u})_{\mathcal{B}}$ denota la colonna delle coordinate del vettore $\mathbf{u}$ rispetto alla base $\mathcal{B}$.*

*I prodotti scalari, cioè le forme bilineari simmetriche, corrispondono alle matrici simmetriche in $\mathrm{M}_n(\mathbb{R})$.*

**Dimostrazione –** Il primo punto della corrispondenza è una conseguenza diretta della proposizione precedente: a ogni applicazione bilineare possiamo associare gli $n^2$ scalari $g(\mathbf{v}_i, \mathbf{v}_j)$.

Vediamo ora il secondo punto e cioè come associare direttamente a una matrice una applicazione bilineare.

Definiamo

$$g(\mathbf{u}, \mathbf{v}) = (\mathbf{u})_{\mathcal{B}}^T C (\mathbf{v})_{\mathcal{B}} = (x_1 \dots x_n) \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \dots & c_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

ove

$$(\mathbf{u})_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad (\mathbf{v})_{\mathcal{B}} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

sono le coordinate di $\mathbf{u}$ e $\mathbf{v}$ rispetto alla base $\mathcal{B}$. È immediato verificare che

$$g(\mathbf{v}_i, \mathbf{v}_j) = c_{ij}$$

e dunque $C$ è proprio la matrice associata a $g$ come al punto (1).

La bilinearità di $g$ è pressoché immediata. Verifichiamo la prima parte della condizione (1) della Definizione 10.1.1 lasciando le altre verifiche per esercizio.

$$g(\mathbf{u}+\mathbf{u}',\mathbf{v}) = (\mathbf{u}+\mathbf{u}')_{\mathcal{B}}^T C(\mathbf{v})_{\mathcal{B}} = (\mathbf{u})_{\mathcal{B}}^T C(\mathbf{v})_{\mathcal{B}} + (\mathbf{u}')_{\mathcal{B}}^T C(\mathbf{v})_{\mathcal{B}}$$
$$= g(\mathbf{u},\mathbf{v}) + g(\mathbf{u}',\mathbf{v})$$

Notiamo ora che $g$ è simmetrica se e solo se la matrice corrispondente $C = (c_{ij})$ è simmetrica cioè $C = C^T$.

Se $g$ è simmetrica allora $c_{ij} = g(\mathbf{v}_i,\mathbf{v}_j) = g(\mathbf{v}_j,\mathbf{v}_i) = c_{ji}$ per ogni $i,j = 1,\dots,n$, quindi $C$ è simmetrica.

Viceversa, supponiamo che $C$ sia simmetrica, cioè $C = C^T$. Osserviamo che per ogni $\mathbf{u},\mathbf{v} \in V$ si ha che $g(\mathbf{u},\mathbf{v})^T = g(\mathbf{u},\mathbf{v})$, poiché $g(\mathbf{u},\mathbf{v}) \in \mathbb{R}$.

Per ogni $\mathbf{u},\mathbf{v} \in V$ per la (10.3) abbiamo dunque che:

$$g(\mathbf{u},\mathbf{v}) = (\mathbf{u})_{\mathcal{B}}^T\, C\,(\mathbf{v})_{\mathcal{B}} = ((\mathbf{u})_{\mathcal{B}}^T\, C\,(\mathbf{v})_{\mathcal{B}})^T$$
$$= (\mathbf{v})_{\mathcal{B}}^T\, C^T\,(\mathbf{u})_{\mathcal{B}} = (\mathbf{v})_{\mathcal{B}}^T\, C\,(\mathbf{u})_{\mathcal{B}} = g(\mathbf{v},\mathbf{u})$$

Questo dimostra che $g$ è simmetrica. □

Vediamo un esempio concreto della corrispondenza descritta qui sopra.

**Esempio 10.2.2**

Osserviamo che la matrice:

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

è associata, rispetto alla base canonica di $\mathbb{R}^2$, al prodotto scalare $<(x_1,x_2),(y_1,y_2)>= 3x_1y_1 + x_1y_2 + x_2y_1 + 2x_2y_2$. Infatti:

$$\begin{pmatrix} x_1 & x_2 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = 3x_1y_1 + x_1y_2 + x_2y_1 + 2x_2y_2$$

## 10.3 CAMBIO DI BASE

Ci chiediamo ora come cambi la matrice associata a una data forma bilineare su $\mathbb{R}^n$ fissata la base canonica, se decidiamo di cambiare base[1].

Ricordiamo, in analogia a quanto visto per le applicazioni lineari, che al variare della base, la matrice associata a un'applicazione lineare può assumere varie forme (per esempio se riusciamo a trovare una base di autovettori, la matrice associata è diagonale), tuttavia *l'applicazione lineare non cambia!* La situazione qui è del tutto simile: la matrice associata alla forma bilineare data può assumere aspetti molto diversi tra loro e tuttavia la forma bilineare non cambia.

[1] Il fatto che prendiamo $V = \mathbb{R}^n$ è solo per semplicità: tutti i nostri ragionamenti valgono in modo identico per uno spazio generico $V$ di dimensione finita, in cui fissiamo una base.

Sia $I_{\mathcal{B},\mathcal{C}}$ la matrice del cambio di base; si tratta della matrice associata all'applicazione identica ove abbiamo fissato una data base ordinata $\mathcal{B}$ nel dominio e la base canonica $\mathcal{C}$ nel codominio. Abbiamo allora che per ogni vettore $\mathbf{u} \in \mathbb{R}^n$:

$$(\mathbf{u})_{\mathcal{C}} = I_{\mathcal{B},\mathcal{C}}(\mathbf{u})_{\mathcal{B}}$$

ove $(\mathbf{u})_{\mathcal{B}}$ denota la colonna delle coordinate del vettore $\mathbf{u}$ rispetto alla base ordinata $\mathcal{B}$.

Supponiamo che $C$ sia la matrice associata a una forma bilineare data, fissata la base canonica di $\mathbb{R}^n$. Vogliamo determinare qual è la matrice $C'$ associata allo stessa forma bilineare rispetto alla base $\mathcal{B}$. Sostituiamo pertanto i vettori generici nell'espressione della forma bilineare data:

$$< \mathbf{u}, \mathbf{v} >= (\mathbf{u})_{\mathcal{C}}^T C(\mathbf{v})_{\mathcal{C}} = (I_{\mathcal{B}\mathcal{C}}(\mathbf{u})_{\mathcal{B}})^T C(I_{\mathcal{B}\mathcal{C}}(\mathbf{v})_{\mathcal{B}}) = (\mathbf{u})_{\mathcal{B}}^T I_{\mathcal{B}\mathcal{C}}^T C I_{\mathcal{B}\mathcal{C}}(\mathbf{v})_{\mathcal{B}}$$

Abbiamo pertanto ottenuto che, fissata una base ordinata arbitraria $\mathcal{B}$ la matrice $C'$ associata a una forma bilineare $g$ è legata alla matrice $C$ associata a $g$ rispetto alla base canonica dalla formula:

$$C' = I_{\mathcal{B},\mathcal{C}}^T C I_{\mathcal{B},\mathcal{C}} \tag{10.4}$$

Più in generale, come già notato nel Capitolo 8, tale formula vale anche nel caso in cui $\mathcal{C}$ venga sostituita da una base arbitraria $\mathcal{B}'$; la dimostrazione di questa affermazione resta la stessa del caso visto. Dunque la formula (10.4) diventa:

$$C' = I_{\mathcal{B},\mathcal{B}'}^T C I_{\mathcal{B},\mathcal{B}'}, \qquad \text{con} \quad I_{\mathcal{B},\mathcal{B}'} = I_{\mathcal{B}',\mathcal{C}}^{-1} I_{\mathcal{B},\mathcal{C}} \tag{10.5}$$

ove per l'ultima uguaglianza abbiamo usato il Teorema 8.3.1, considerando come $F$ l'applicazione identità di $\mathbb{R}^n$.

Vediamo l'esempio precedente.

---

**Esempio 10.3.1**

Consideriamo il prodotto scalare $<,>$ (dunque in particolare una forma bilineare) associato, rispetto alla base canonica, alla matrice

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

Supponiamo di scegliere come base $\mathcal{B} = \{\mathbf{v}_1 = \mathbf{e}_1, \mathbf{v}_2 = \mathbf{e}_1 + \mathbf{e}_2\}$, dunque:

$$I_{\mathcal{B}\mathcal{C}} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

La matrice $C'$ associata al medesimo prodotto scalare rispetto alla base $\mathcal{B}$ è:

$$C' = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 3 & 3 \end{pmatrix}$$

Consideriamo il prodotto scalare di $\mathbf{v}_1$ e $\mathbf{v}_2$ cioè $< \mathbf{e}_1, \mathbf{e}_1 + \mathbf{e}_2 >$ utilizzando prima la base canonica e poi la base $\mathcal{B}$ e verifichiamo che il risultato è lo stesso.

$$< \mathbf{e}_1, \mathbf{e}_1 + \mathbf{e}_2 > = (1 \ 0)\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3$$

$$< \mathbf{e}_1 + \mathbf{e}_1 + \mathbf{e}_2 > = (1 \ 0)\begin{pmatrix} 2 & 3 \\ 3 & 3 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 3$$

---

♦ **Osservazione 10.3.2** È utile confrontare la formula del cambio di base per una applicazione lineare e quella del cambio di base per una forma bilineare.

- Due matrici $A$ e $B$ rappresentano la stessa applicazione lineare (rispetto a basi diverse) se e solo se sono simili, cioè esiste una matrice invertibile $P$ tale che $B = P^{-1}AP$.
- Due matrici $A$ e $B$ rappresentano la stessa forma bilineare (rispetto a basi diverse) se e solo se esiste una matrice $P$ invertibile tale che $B = P^T AP$.

È chiaro, guardando queste due formule, che matrici con la proprietà $P^T = P^{-1}$, cioè tali che la loro trasposta coincide con l'inversa, rivestono una importanza particolare. Più avanti faremo uno studio più approfondito di queste matrici e delle loro proprietà.

## 10.4 PRODOTTI SCALARI

D'ora in avanti considereremo forme bilineari simmetriche e cioè prodotti scalari. Richiamiamo la definizione, aggiungendo alcune precisazioni.

**Definizione 10.4.1** Sia $V$ uno spazio vettoriale. La funzione $<,>: V \times V \longrightarrow \mathbb{R}$ si dice un *prodotto scalare* se:

1. $< \mathbf{u} + \mathbf{u}', \mathbf{v} >=< \mathbf{u}, \mathbf{v} > + < \mathbf{u}', \mathbf{v} >$,
   $< \mathbf{u}, \mathbf{v} + \mathbf{v}' >=< \mathbf{u}, \mathbf{v} > + < \mathbf{u}, \mathbf{v}' >$ per ogni $\mathbf{u}$, $\mathbf{u}'$, $\mathbf{v}$, $\mathbf{v}' \in V$.
2. $< \lambda\mathbf{u}, \mathbf{v} >= \lambda < \mathbf{u}, \mathbf{v} >$,
   $< \mathbf{u}, \mu\mathbf{v} >= \mu < \mathbf{u}, \mathbf{v} >$ per ogni $\mathbf{u}$, $\mathbf{v} \in V$ e per ogni $\lambda$, $\mu \in \mathbb{R}$.
3. $< \mathbf{u}, \mathbf{v} >=< \mathbf{v}, \mathbf{u} >$ per ogni $\mathbf{u}$, $\mathbf{v} \in V$.

In altre parole, per ogni vettore fissato $\mathbf{u} \in V$ le funzioni $< \mathbf{u}, \cdot >: V \longrightarrow \mathbb{R}$ e $< \cdot, \mathbf{u} >: V \longrightarrow \mathbb{R}$ sono applicazioni lineari.

$<,>$ è *non degenere* quando, se $< \mathbf{u}, \mathbf{v} >= 0$ per ogni $\mathbf{v} \in V$, allora $\mathbf{u} = \mathbf{0}$.

$<,>$ è *definito positivo* se $< \mathbf{u}, \mathbf{u} > \geq 0$ per ogni $\mathbf{u} \in V$ e $< \mathbf{u}, \mathbf{u} >= 0$ se e solo se $\mathbf{u} = \mathbf{0}$. Se un prodotto scalare è definito positivo, diciamo che $\sqrt{< \mathbf{u}, \mathbf{u} >}$ è la *norma* del vettore $\mathbf{u}$ e la indichiamo con $\|\mathbf{u}\|$.

Come avremo modo di vedere, i prodotti scalari definiti positivi sono particolarmente importanti e tuttavia esistono esempi di prodotti non definiti positivi che rivestono un'importanza fondamentale in fisica.

---

**Esempio 10.4.2**

È immediato verificare che il *prodotto scalare standard* o *euclideo* in $\mathbb{R}^n$, definito nell'Esempio 10.1.3 e dato da:

$$< (x_1, \ldots, x_n), (x_1', \ldots, x_n') >_e = x_1x_1' + \cdots + x_nx_n'$$

è definito positivo, infatti:

$$< (x_1, \ldots, x_n), (x_1, \ldots, x_n) >_e = x_1^2 + \cdots + x_n^2 \geq 0$$

Inoltre $x_1^2 + \cdots + x_n^2 = 0$ se e solo se $x_1 = x_2 = \cdots = x_n = 0$.

Si noti che la matrice associata al prodotto euclideo rispetto alla base canonica di $\mathbb{R}^n$ è la matrice identica.

---

**Definizione 10.4.3** Sia $V$ uno spazio vettoriale con un prodotto scalare. Diciamo che $\mathbf{u}, \mathbf{v} \in V$ sono *perpendicolari (ortogonali) tra loro* se $< \mathbf{u}, \mathbf{v} >= 0$. Useremo anche la notazione $\mathbf{u} \perp \mathbf{v}$ per indicare due vettori perpendicolari tra loro.

Dunque possiamo riformulare le nozioni definite sopra nel seguente modo:
*Un prodotto scalare è non degenere se e solo se non esiste un vettore non nullo perpendicolare a tutti gli altri. Inoltre un prodotto scalare definito positivo è automaticamente non degenere*: infatti l'essere definito positivo implica che non esista un vettore ortogonale a se stesso, mentre l'essere degenere richiede che tale vettore esista.

♦ **Osservazione 10.4.4** 1. La nozione di ortogonalità dipende dal prodotto scalare scelto. Per esempio in $\mathbb{R}^2$ consideriamo i due prodotti scalari:

$$< (x_1, x_2), (y_1, y_2) >_e = x_1y_1 + x_2y_2,$$

$$< (x_1, x_2), (y_1, y_2) >' = x_1y_2 + x_2y_1.$$

Vediamo che i vettori $\mathbf{e}_1$ ed $\mathbf{e}_2$ sono perpendicolari rispetto al prodotto euclideo $<,>_e$ ma non rispetto all'altro prodotto scalare, infatti:

$$< (1,0), (0,1) >' = 1$$

2. Il fatto che esista un vettore ortogonale a se stesso è una condizione necessaria, ma non sufficiente, per avere un prodotto scalare degenere. Infatti se consideriamo in $\mathbb{R}^2$ il prodotto scalare:

$$< (x_1, x_2), (y_1, y_2) >= x_1y_1$$

vediamo subito che è degenere, in quanto il vettore $\mathbf{e}_2$ è ortogonale a ogni altro vettore.

D'altra parte, se consideriamo in $\mathbb{R}^2$ il prodotto scalare:

$$< (x_1, x_2), (y_1, y_2) >_m = x_1y_1 - x_2y_2$$

abbiamo che il vettore $(1,1)$ è ortogonale a se stesso, infatti:

$$< (1,1), (1,1) >_m = 1 - 1 = 0$$

tuttavia, tale prodotto è non degenere, infatti non esiste alcun vettore non nullo che sia ortogonale a tutti gli altri. Se, per assurdo, tale vettore $(a, b)$ esistesse, sarebbe ortogonale a $\mathbf{e}_1$ ed $\mathbf{e}_2$:

$$< (a,b), (1,0) >_m = a = 0 \qquad < (a,b), (0,1) >_m = -b = 0$$

Dunque $(a, b) = (0, 0)$.

Quest'ultimo prodotto scalare è particolarmente importante per la fisica e si dice *prodotto scalare di Minkowski*.

## 10.5 SOTTOSPAZI ORTOGONALI

In questa sezione introduciamo la nozione di *sottospazio ortogonale* $W^\perp$ a un sottospazio $W$ di uno spazio vettoriale $V$.

**Proposizione 10.5.1** *Siano $W \subseteq V$ un sottospazio vettoriale di $V$ e $<,>$ un prodotto scalare su $V$. Allora l'insieme*

$$W^\perp = \{\mathbf{u} \in V \mid <\mathbf{u}, \mathbf{w}>= 0, \text{ per ogni } \mathbf{w} \in W\}$$

*è un sottospazio vettoriale di $V$.*

**Dimostrazione** – Verifichiamo le tre proprietà della Definizione 10.4.1. Abbiamo subito che $\mathbf{0}_V \in W^\perp$. Infatti, per la proprietà (2) della Definizione 10.4.1:

$$<\mathbf{0}_V, \mathbf{w}>=<0 \cdot \mathbf{0}_V, \mathbf{w}>= 0 \cdot <\mathbf{0}_V, \mathbf{w}>= 0$$

Verifichiamo ora che per ogni $\mathbf{u}_1$, $\mathbf{u}_2 \in W^\perp$ abbiamo che $\mathbf{u}_1 + \mathbf{u}_2 \in W^\perp$. Per la proprietà (1) della Definizione 10.4.1:

$$<\mathbf{u}_1 + \mathbf{u}_2, \mathbf{w}>=<\mathbf{u}_1, \mathbf{w}> + <\mathbf{u}_2, \mathbf{w}>= 0$$

Infine, verifichiamo che per ogni scalare $\lambda$ e ogni $\mathbf{u} \in W^\perp$ abbiamo che $\lambda\mathbf{u} \in W^\perp$. Per le proprietà (2) della Definizione 10.4.1:

$$<\lambda\mathbf{u}, \mathbf{w}>= \lambda <\mathbf{u}, \mathbf{w}>= 0$$ □

**Definizione 10.5.2** Dati un sottospazio vettoriale $W$ di $V$ e un prodotto scalare $<,>$ su $V$, il *sottospazio ortogonale* a $W$ è:

$$W^\perp = \{\mathbf{u} \in V \mid <\mathbf{u}, \mathbf{w}>= 0, \text{ per ogni } \mathbf{w} \in W\}$$

L'osservazione che segue è fondamentale per lo svolgimento degli esercizi.

♦ **Osservazione 10.5.3** Osserviamo che se $W = \langle \mathbf{w}_1, \dots, \mathbf{w}_n \rangle \subseteq V$, allora

$$W^\perp = \{\mathbf{u} \in V \mid <\mathbf{u}, \mathbf{w}_i>= 0, \text{ per ogni } i = 1, \dots, n\} \tag{10.6}$$

Infatti se $\mathbf{u} \in W^\perp$ certamente $<\mathbf{u}, \mathbf{w}_i>= 0$, ma vale anche il viceversa. Infatti sia $\mathbf{u} \in V$ tale che $<\mathbf{u}, \mathbf{w}_i>= 0$, Se $\mathbf{w} = \lambda_1\mathbf{w}_1 + \dots + \lambda_n\mathbf{w}_n \in W$ allora

$$<\mathbf{u}, \mathbf{w}>=<\mathbf{u}, \lambda_1\mathbf{w}_1 + \dots + \lambda_n\mathbf{w}_n>= \lambda_1 <\mathbf{u}, \mathbf{w}_1> + \dots + \lambda_n <\mathbf{u}, \mathbf{w}_n>= 0$$

Dunque, fissata una base di $V$, possiamo esprimere le condizioni in (10.6) mediante equazioni lineari, e pertanto il calcolo del sottospazio ortogonale a un sottospazio dato si riduce alla soluzione di un sistema lineare omogeneo.

**Esempio 10.5.4**

Calcoliamo il sottospazio ortogonale a $W = \langle (1, 1) \rangle$ rispetto al prodotto scalare euclideo e al prodotto scalare di Minkowski introdotti nell'Osservazione 10.4.4.

Rispetto al prodotto scalare euclideo abbiamo:

$$(W^\perp)_e = \{(x,y) \mid <(x,y),(1,1)>_e = 0\} = \{(x,y) \mid x+y=0\} = \langle(1,-1)\rangle$$

Rispetto al prodotto scalare di Minkowski abbiamo:

$$(W^\perp)_m = \{(x,y) \mid <(x,y),(1,1)>_m = 0\} = \{(x,y) \mid x-y=0\} = \langle(1,1)\rangle$$

---

Grazie all'Osservazione 10.5.3 siamo in grado di determinare la dimensione del sottospazio ortogonale a $W \subseteq \mathbb{R}^n$ rispetto al prodotto scalare euclideo.

**Proposizione 10.5.5** *Sia $W$ un sottospazio vettoriale di $\mathbb{R}^n$ e sia $W^\perp$ il sottospazio ortogonale a $W$ rispetto al prodotto scalare euclideo. Allora:*

$$\dim(W) + \dim(W^\perp) = n$$

**Dimostrazione** – Sia $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ una base ordinata di $W$ e sia $A = (a_{ij}) \in M_{m,n}(\mathbb{R})$ la matrice avente come righe le coordinate dei vettori $\mathbf{v}_1, \dots, \mathbf{v}_m$. Per l'Osservazione 10.5.3 abbiamo che $(x_1, \dots, x_n) \in W^\perp$ se e solo se:

$$(a_{i1}, \dots, a_{in}) \cdot (x_1, \dots, x_n) = 0 \text{ per ogni } i = 1, \dots, n$$

In altre parole:

$$W^\perp = \{\mathbf{x} = (x_1, \dots, x_n) | A\mathbf{x} = \mathbf{0}\} = \ker(L_A),$$

ove $L_A$ è l'applicazione lineare associata alla matrice $A$ rispetto alla base canonica.

Per il Teorema della dimensione (5.5.1):

$$\dim(W^\perp) = n - \dim \mathrm{Im}\,(A)$$

Sappiamo che la dimensione dell'immagine di $L_A$ è il rango di $A$, cioè la dimensione del sottospazio generato dalle sue colonne o equivalentemente dalle sue righe. Poiché le righe di $A$ sono date dalle coordinate dei vettori di una base di $W$, tale dimensione è proprio $\dim(W)$. Pertanto:

$$\dim(W^\perp) = n - \dim(W)$$ □

## 10.6 ALGORITMO DI GRAM-SCHMIDT

Sia $V$ uno spazio vettoriale munito di un prodotto scalare definito positivo. L'algoritmo di Gram-Schmidt ci permette di costruire una base che consiste di vettori mutualmente ortogonali.

**Definizione 10.6.1** Siano $V$ uno spazio vettoriale con un prodotto scalare $< , >$ definito positivo e $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ una sua base. Diciamo che $\mathcal{B}$ è una *base ortogonale* se $\mathbf{u}_i \perp \mathbf{u}_j$ per ogni $i, j = 1, \ldots, n$ e che $\mathcal{B}$ è *base ortonormale* se:

$$< \mathbf{u}_i, \mathbf{u}_j >= \begin{cases} 1 & \text{per } i = j \\ 0 & \text{per } i \neq j \end{cases}$$

cioè se è una base ortogonale e ogni vettore della base ha norma uguale a 1.

In questa situazione, e nel seguito, ci sarà utile utilizzare la notazione compatta:

$$< \mathbf{u}_i, \mathbf{u}_j >= \delta_{ij},$$

ove la funzione $\delta_{ij}$ si dice *delta di Kronecker* ed è definita da:

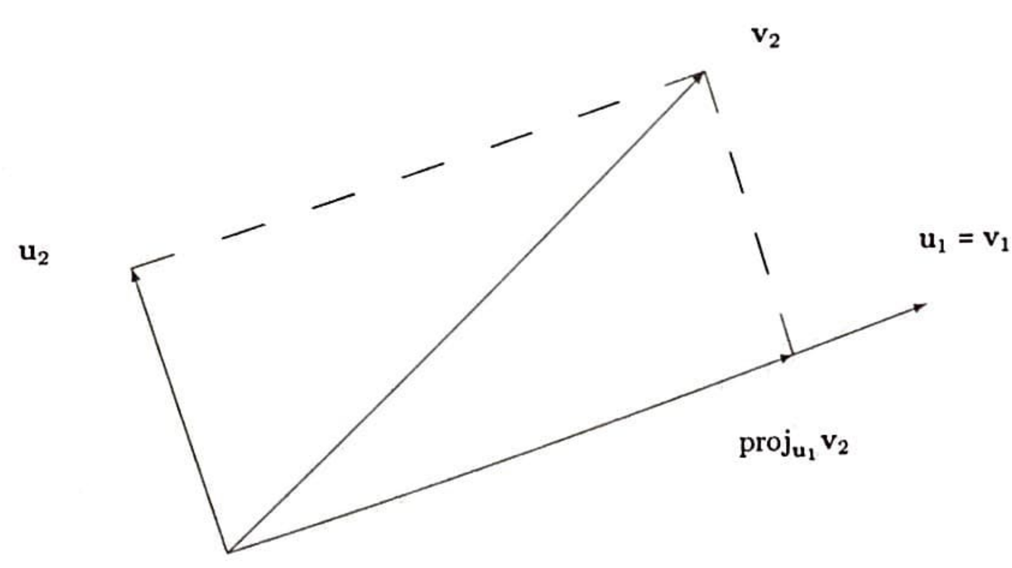
$$\delta_{ij} = \begin{cases} 1 & \text{per } i = j \\ 0 & \text{per } i \neq j \end{cases}$$

Osserviamo che data una base ortogonale possiamo immediatamente ottenere una base ortonormale moltiplicando ogni vettore per l'inverso della sua norma.

Per ottenere una base ortogonale a partire da una base data, è necessario introdurre la nozione di *proiezione ortogonale*.

**Definizione 10.6.2** Sia $V$ uno spazio vettoriale con un prodotto scalare definito positivo e siano $\mathbf{u}, \mathbf{v} \in V$. La *proiezione ortogonale* del vettore $\mathbf{v}$ sul vettore $\mathbf{u}$ è data da:

$$\text{proj}_{\mathbf{u}}(\mathbf{v}) = \frac{< \mathbf{v}, \mathbf{u} >}{< \mathbf{u}, \mathbf{u} >}\mathbf{u}$$




Come si vede dalla figura e si può verificare con un facile calcolo, se $\{\mathbf{u}, \mathbf{v}\}$ è una base di $\mathbb{R}^2$, $\{\mathbf{u}_1 = \mathbf{v}_1, \mathbf{u}_2 = \mathbf{v}_2 - \text{proj}_{\mathbf{u}_1}(\mathbf{v}_2)\}$ è una base ortogonale di $\mathbb{R}^2$. Questo procedimento può essere iterato e permette di costruire una base ortogonale $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ a partire da una base generica $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ di $V$.

Si definisce allora, in modo simile a quanto abbiamo visto per $\mathbb{R}^2$:

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1, & \mathbf{f}_1 &= \frac{1}{\|\mathbf{u}_1\|}\mathbf{u}_1 \\
\mathbf{u}_2 &= \mathbf{v}_2 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_2), & \mathbf{f}_2 &= \frac{1}{\|\mathbf{u}_2\|}\mathbf{u}_2 \\
\mathbf{u}_3 &= \mathbf{v}_3 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_3) - \mathrm{proj}_{\mathbf{u}_2}(\mathbf{v}_3), & \mathbf{f}_3 &= \frac{1}{\|\mathbf{u}_3\|}\mathbf{u}_3 \qquad (10.7) \\
&\vdots & &\vdots \\
\mathbf{u}_k &= \mathbf{v}_k - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_k) \cdots - \mathrm{proj}_{\mathbf{u}_{k-1}}(\mathbf{v}_k), & \mathbf{f}_k &= \frac{1}{\|\mathbf{u}_k\|}\mathbf{u}_k
\end{aligned}$$

ove $\|\mathbf{u}_i\| = \sqrt{<\mathbf{u}_i, \mathbf{u}_i>}$. Si verifica immediatamente che con il procedimento descritto sopra, denominato *Algoritmo di Gram-Schmidt*, si ottiene un insieme di vettori mutuamente ortogonali.

**Teorema 10.6.3** **(Algoritmo di Gram-Schmidt)** Sia $V$ uno spazio vettoriale con un prodotto scalare $<,>$ definito positivo. Sia $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ una base di $V$. Allora l'insieme di vettori $\mathbf{f}_1, \dots, \mathbf{f}_n$ ricavati attraverso (10.7) è una base ortonormale di $V$.

**Dimostrazione** – Il fatto che i vettori $\mathbf{f}_1, \dots, \mathbf{f}_n$ siano mutuamente ortogonali e di norma uguale a 1 è una facile verifica. Poiché la dimensione di $V$ è uguale a $n$, per mostrare che formano una base è sufficiente verificare che sono linearmente indipendenti. Supponiamo che

$$\lambda_1\mathbf{f}_1 + \cdots + \lambda_n\mathbf{f}_n = \mathbf{0}$$

Allora:

$$<\mathbf{f}_i, \lambda_1\mathbf{f}_1 + \cdots + \lambda_n\mathbf{f}_n >= \lambda_i = 0 \ \text{ per ogni } i = 1, \dots, n$$

Dunque $\mathbf{f}_1, \dots, \mathbf{f}_n$ sono linearmente indipendenti e base ortonormale di $V$. □

Concludiamo con un'osservazione che stabilisce l'importanza della scelta delle coordinate rispetto a una base ortonormale per un prodotto scalare definito positivo, nella descrizione esplicita di tale prodotto.

♦ **Osservazione 10.6.4** Sia $V$ uno spazio vettoriale di dimensione finita con un prodotto scalare arbitrario $<,>$ definito positivo e sia $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ una base ortonormale di $V$. Una tale base esiste sempre grazie all'Algoritmo di Gram-Schmidt.

La matrice associata al prodotto scalare $<>$, fissata la base $\mathcal{B}$ è data, per definizione, da $C = (<\mathbf{v}_i, \mathbf{v}_j>)_{i,j=1,\cdots,n}$. Quindi $C$ coincide con la matrice identità $I$. Inoltre se esprimiamo due vettori $\mathbf{u}$ e $\mathbf{w}$ come combinazione lineare dei vettori della base $\mathcal{B}$, $\mathbf{u} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n$, $\mathbf{w} = \beta_1\mathbf{v}_1 + \cdots + \beta_n\mathbf{v}_n$, abbiamo che:

$$\begin{aligned}
<\mathbf{u}, \mathbf{w}> &=< \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n, \beta_1\mathbf{v}_1 + \cdots + \beta_n\mathbf{v}_n > \\
&= \alpha_1\beta_1 <\mathbf{v}_1, \mathbf{v}_1> + \alpha_1\beta_2 <\mathbf{v}_1, \mathbf{v}_2> + \cdots + \alpha_n\beta_n <\mathbf{v}_n, \mathbf{v}_n> \\
&= \alpha_1\beta_1 + \cdots + \alpha_n\beta_n
\end{aligned}$$

ove abbiamo utilizzato le proprietà di bilinearità del prodotto scalare e il fatto che $<\mathbf{v}_i, \mathbf{v}_j >= \delta_{ij}$. Quindi a un prodotto scalare arbitrario, definito positivo, se scegliamo una base ortonormale $\mathcal{B}$, è associata la matrice identità, proprio come accade per il prodotto scalare standard in $\mathbb{R}^n$ e il prodotto di $\mathbf{u}$ e $\mathbf{v}$ è dato dal prodotto scalare euclideo delle loro componenti rispetto alla base $\mathcal{B}$.

## 10.7 ESERCIZI SVOLTI

**Esercizio 10.7.1**

*Si consideri la forma bilineare* $<(x_1,x_2),(y_1,y_2)>= x_1y_1 - 2x_1y_2 - 2x_1y_2$. *Si determini la matrice ad essa associata rispetto alla base canonica, e si mostri che si tratta di un prodotto scalare. Si scriva la matrice associata allo stesso prodotto scalare rispetto alla base* $\mathcal{B} = \{\mathbf{e}_1 + \mathbf{e}_2, 2\mathbf{e}_1 - \mathbf{e}_2\}$.

**Svolgimento**

Possiamo subito scrivere la matrice associata alla forma bilineare data rispetto alla base canonica:

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

Si tratta di un prodotto scalare in quanto la matrice è simmetrica. La matrice $C'$ associata al medesimo prodotto scalare rispetto alla base $\mathcal{B}$ è:

$$C' = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}^T \begin{pmatrix} 1 & -2 \\ -2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} -3 & 0 \\ 0 & 12 \end{pmatrix}$$

■

**Esercizio 10.7.2**

*Si consideri il sottospazio vettoriale* $W$ *generato in* $\mathbb{R}^3$ *dai vettori* $\mathbf{w}_1 = \mathbf{e}_1 + \mathbf{e}_2 - 3\mathbf{e}_3$ *e* $\mathbf{w}_2 = -\mathbf{e}_1 + 2\mathbf{e}_2 - 3\mathbf{e}_3$. *Si determini una base per* $W^\perp$, *calcolato rispetto al prodotto scalare euclideo in* $\mathbb{R}^3$. *Si determinino inoltre una base ortonormale per* $W$ *e una per* $W^\perp$.

**Svolgimento**

$W^\perp$ consiste dei vettori $(x,y,z) \in \mathbb{R}^3$ tali che:

$$(1,1,-3)\cdot(x,y,z) = 0, \qquad (-1,2,-3)\cdot(x,y,z) = 0$$

Dobbiamo pertanto risolvere il sistema lineare:

$$\begin{cases} x + y - 3z = 0 \\ -x + 2y - 3z = 0 \end{cases}$$

Otteniamo subito che $W^\perp = \{(z,2z,z) | z \in \mathbb{R}\}$ pertanto una base per $W^\perp$ è data dal vettore $(1,2,1)$.

Per determinare una base ortonormale di $W$ è necessario utilizzare l'algoritmo di Gram-Schmidt descritto in (10.7). Abbiamo pertanto:

$$\mathbf{u}_1 = \mathbf{w}_1 = (1,1,-3), \quad \mathbf{f}_1 = \frac{1}{\|\mathbf{u}_1\|}\mathbf{u}_1 = (1/\sqrt{11}, 1/\sqrt{11}, -3/\sqrt{11})$$

$$\mathbf{u}_2 = \mathbf{w}_2 - \frac{<\mathbf{w}_2,\mathbf{u}_1>}{<\mathbf{u}_1,\mathbf{u}_1>}\mathbf{u}_1,$$

$$\mathbf{f}_2 = \frac{1}{\|\mathbf{u}_2\|}\mathbf{u}_2 = (-7/\sqrt{66}, 2\sqrt{2/33}, -1/\sqrt{66})$$

Una base ortonormale per $W$ è pertanto data da:

$$\left\{(1/\sqrt{11}, 1/\sqrt{11}, -3/\sqrt{11}), (-7/\sqrt{66}, 2\sqrt{2/33}, -1/\sqrt{66})\right\}$$

Una base ortonormale per $W^{\perp}$ si ottiene prendendo un generatore, per esempio $(1, 2, 1)$, e dividendolo per la sua norma:

$$\left\{(1/\sqrt{6}, 2/\sqrt{6}, 1/\sqrt{6})\right\}$$

■

**Esercizio 10.7.3**

*Si consideri il sottospazio vettoriale $W$ di $\mathbb{R}^4$ costituito dalle soluzioni del sistema lineare:*

$$\begin{cases} x + y + z - t = 0 \\ x + 2y - z + t = 0 \end{cases}$$

*Si determini una base per $W^{\perp}$, calcolato rispetto al prodotto scalare euclideo in $\mathbb{R}^3$.*

**Svolgimento**

Osserviamo che possiamo scrivere le equazioni che definiscono $W$ nel seguente modo:

$$(1, 1, 1, -1) \cdot (x, y, z, t) = 0, \qquad (1, 2, -1, 1) \cdot (x, y, z, t) = 0$$

Pertanto i vettori $(1, 1, 1, -1)$ e $(1, 2, -1, 1)$ appartengono a $W^{\perp}$ in quanto sono perpendicolari a tutti i vettori di $W$. Poiché per la Proposizione 10.5.1 $\dim(W^{\perp}) = 4 - \dim(W) = 2$, segue che essi formano una base di $W^{\perp}$. ■

## 10.8 ESERCIZI PROPOSTI

**Esercizio 10.8.1**

*Si consideri la forma bilineare in $\mathbb{R}^2$ data da $g((x_1, x_2), (y_1, y_2)) = 2x_1y_1 - x_1y_2 + 3x_2y_2$.*

a) *Si scriva la matrice ad essa associata rispetto alla base canonica.*

b) *Si scriva la matrice ad essa associata rispetto alla base $\mathcal{B} = \{\mathbf{v}_1 = \mathbf{e}_1 + \mathbf{e}_2, \mathbf{v}_2 = -2\mathbf{e}_2\}$.*

**Esercizio 10.8.2**

*Sia $W$ il sottospazio vettoriale di $\mathbb{R}^4$ definito dall'equazione $x + y + 2z - t = 0$ e si consideri il prodotto scalare euclideo di $\mathbb{R}^4$.*

a) *Si determini una base per $W^{\perp}$.*

b) *Si determini una base ortogonale di $W$.*

**Esercizio 10.8.3**

*Sia $W$ il sottospazio vettoriale di $\mathbb{R}^4$ generato dai vettori* $\mathbf{e}_1 + \mathbf{e}_4$, $\mathbf{e}_2 - 2\mathbf{e}_3 + \mathbf{e}_4$.

a) *Si determini una base per $W^\perp$.*

b) *Si determini una base ortogonale di $W$.*

**Esercizio 10.8.4**

*Sia $W$ il sottospazio vettoriale di $\mathbb{R}^4$ definito dalle seguenti equazioni:*

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

*Si calcoli una base ortonormale per $W$ rispetto al prodotto scalare euclideo di $\mathbb{R}^4$.*

**Esercizio 10.8.5**

*Sia $W$ il sottospazio vettoriale di $\mathbb{R}^3$ generato dai vettori* (1,2,−1), (−1/2,−1,−1), (1,2,1).

a) *Si trovi una base di $W$ ortonormale rispetto al prodotto scalare euclideo,*

b) *Si trovi una base ortonormale per $W^\perp$.*

**Esercizio 10.8.6**

*Si consideri il sottospazio vettoriale $W$ di $\mathbb{R}^5$ descritto dalle soluzioni del sistema lineare omogeneo:*

$$\begin{cases} x_1 + 2x_2 + x_5 = 0 \\ x_1 - x_2 - x_3 = 0 \\ 2x_1 - x_3 - x_4 = 0 \end{cases}$$

a) *Si determini una base ortogonale di $W$ rispetto al prodotto scalare euclideo.*

b) *Si determini $W^\perp$. Che relazione sussiste tra $W^\perp$ e la matrice del sistema?*

**Esercizio 10.8.7**

*Sia dato il sottospazio vettoriale $W$ di $\mathbb{R}^4$:*

$$W = \langle (1,-1,0,1), (2,-1,0,1), (1,1,1,0) \rangle$$

a) *Si determini $W^\perp$.*

b) *Si determinino una base ortogonale per $W$ e una base ortogonale per $W^\perp$ rispetto al prodotto scalare standard in $\mathbb{R}^4$.*

c) *Si stabilisca se la base trovata al punto* (b) *resta una base ortogonale anche rispetto al prodotto scalare:*

$$< (x_1, x_2, x_3, x_4), (y_1, y_2, y_3, y_4) > = x_1y_1 + 2x_2y_2 + 3x_3y_3 + 4x_4y_4$$

d) *Si scriva la matrice associata al prodotto scalare del punto* (c) *rispetto alla base ordinata*

$$\mathcal{B} = \{(1,1,1,-1), (0,1,0,1), (0,0,1,0), (1,0,1,2)\}.$$

**Esercizio 10.8.8**

a) *Si dimostri che un prodotto scalare in uno spazio vettoriale $V$ di dimensione finita, associato a una matrice diagonale $D$, rispetto a una base data, è non degenere se e solo se tutti gli elementi sulla diagonale di $D$ sono non nulli.*

b) *Si dimostri che un prodotto scalare in uno spazio vettoriale $V$ di dimensione finita, associato a una matrice diagonale $D$, rispetto a una base data, è definito positivo se e solo se tutti gli elementi sulla diagonale di $D$ sono positivi.*

**Esercizio 10.8.9**

*Sia $V$ uno spazio vettoriale con un prodotto scalare $<,>$.*

a) *Se $<,>$ è un prodotto scalare reale definito positivo è vero che allora è non degenere?*

b) *Se $<,>$ è un prodotto scalare reale non degenere è vero che allora è definito positivo?*

# 11 Teorema spettrale

Il Teorema spettrale rappresenta uno dei risultati più importanti dell'algebra lineare elementare. Nel Capitolo 9 abbiamo esaminato il problema del calcolo degli autovalori e autovettori e della diagonalizzabilità di una matrice quadrata $A$ a coefficienti reali. Abbiamo visto che non sempre è possibile trovare una matrice diagonale simile alla matrice data, perché talvolta non esiste una base di autovettori di $A$ per lo spazio $\mathbb{R}^n$. Tuttavia, se la matrice $A$ è simmetrica, cioè coincide con la sua trasposta, il Teorema spettrale ci garantisce che sia diagonalizzabile. Inoltre, non solo $A$ è simile a una matrice diagonale, ma, se indichiamo con $<,>$ il prodotto scalare ad essa associato, esiste una base costituita da autovettori di $A$ tale che la matrice associata a $<,>$ rispetto a tale base sia diagonale.

Per dimostrare tutti questi risultati è necessario introdurre il concetto di applicazione lineare ortogonale e di applicazione lineare simmetrica.

## 11.1 APPLICAZIONI LINEARI ORTOGONALI

In questa sezione vogliamo definire le applicazioni lineari ortogonali e un insieme di matrici estremamente importanti per una comprensione più approfondita del concetto di prodotto scalare: le *matrici ortogonali*. Cominciamo con il primo di questi concetti.

**Definizione 11.1.1** Sia $V$ uno spazio vettoriale reale con un prodotto scalare definito positivo. Diciamo che un'applicazione lineare $U : V \longrightarrow V$ è *ortogonale* se

$$< U(\mathbf{u}), U(\mathbf{v}) >=< \mathbf{u}, \mathbf{v} >, \qquad \text{per ogni} \quad \mathbf{u}, \mathbf{v} \in V$$

In altre parole, l'applicazione lineare $U$ conserva il prodotto scalare dato in $V$.

**Proposizione 11.1.2** *Sia $V$ uno spazio vettoriale reale di dimensione $n$ con un prodotto scalare $<,>$ definito positivo. Sia $U : V \longrightarrow V$ un'applicazione lineare e siano $\mathbf{u}, \mathbf{v} \in V$. Le seguenti affermazioni sono equivalenti;*

1. *$U$ è ortogonale, cioè $< U(\mathbf{u}), U(\mathbf{v}) >=< \mathbf{u}, \mathbf{v} >$ per ogni $\mathbf{u}, \mathbf{v} \in V$;*

2. *$U$ conserva la norma dei vettori, equivalentemente:* $< U(\mathbf{u}), U(\mathbf{u}) >=< \mathbf{u}, \mathbf{u} >$ *per ogni* $\mathbf{u} \in V$;

3. *Se* $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ *è una base ortonormale di $V$ (rispetto al prodotto scalare $<,>$) allora anche* $\{U(\mathbf{v}_1), \dots, U(\mathbf{v}_n)\}$ *è una base ortonormale di $V$.*

**Dimostrazione** – Vediamo (2) $\implies$ (1). Siano $\mathbf{u}, \mathbf{v} \in V$ e si consideri:

$$< U(\mathbf{u}-\mathbf{v}), U(\mathbf{u}-\mathbf{v}) >=< (\mathbf{u}-\mathbf{v}), (\mathbf{u}-\mathbf{v}) >$$

Da cui:

$$\begin{aligned} &< U(\mathbf{u}), U(\mathbf{u}) > -2 < U(\mathbf{u}), U(\mathbf{v}) > + < U(\mathbf{v}), U(\mathbf{v}) > \\ &=< \mathbf{u}, \mathbf{u} > -2 < \mathbf{u}, \mathbf{v} > + < \mathbf{v}, \mathbf{v} > \end{aligned}$$

Utilizzando la (2) vediamo subito che $< U(\mathbf{u}), U(\mathbf{v}) >=< \mathbf{u}, \mathbf{v} >$.

(1) $\implies$ (3). Se $\mathcal{B}$ è una base ortonormale allora $< \mathbf{v}_i, \mathbf{v}_j >= \delta_{ij}$, ove $\delta_{ij}$ è la delta di Kronecker, cioè $\delta_{ij} = 1$ se $i = j$ e $\delta_{ij} = 0$ se $i \neq j$. Per (1) si ha che $< \mathbf{v}_i, \mathbf{v}_j >=< U(\mathbf{v}_i), U(\mathbf{v}_j) >= \delta_{ij}$ dunque anche $\{U(\mathbf{v}_1), \dots, U(\mathbf{v}_n)\}$ è una base ortonormale di $V$.

(3) $\implies$ (2). Sia $\mathbf{u} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n$ un vettore di $V$ espresso come combinazione lineare degli elementi di $\mathcal{B}$. Allora

$$\begin{aligned} < \mathbf{u}, \mathbf{u} > &=< \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n, \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n > \\ &= \alpha_1^2 < \mathbf{v}_1, \mathbf{v}_1 > + \alpha_1\alpha_2 < \mathbf{v}_1, \mathbf{v}_2 > + \cdots + \alpha_n^2 < \mathbf{v}_n, \mathbf{v}_n > \end{aligned}$$

dove stiamo utilizzando la bilinearità del prodotto scalare. Poiché $\mathcal{B}$ è una base ortonormale

$$< \mathbf{u}, \mathbf{u} >= \alpha_1^2 + \cdots + \alpha_n^2 \tag{11.1}$$

Calcoliamo ora $< U(\mathbf{u}), U(\mathbf{u}) >$:

$$\begin{aligned} &< U(\mathbf{u}), U(\mathbf{u}) > =< \alpha_1 U(\mathbf{v}_1) + \cdots + \alpha_n U(\mathbf{v}_n), \\ &\alpha_1 U(\mathbf{v}_1) + \cdots + \alpha_n U(\mathbf{v}_n) > = \alpha_1^2 < U(\mathbf{v}_1), U(\mathbf{v}_1) > + \alpha_1\alpha_2 < U(\mathbf{v}_1), U(\mathbf{v}_2) > + \\ &\cdots + \alpha_n^2 < U(\mathbf{v}_n), U(\mathbf{v}_n) > \end{aligned}$$

dove stiamo utilizzando la bilinearità del prodotto scalare e la linearità di $U$. Poiché per ipotesi anche $\{U(\mathbf{v}_1), \dots, U(\mathbf{v}_n)\}$ è base ortonormale e cioè $< U(\mathbf{v}_i), U(\mathbf{v}_j) >= \delta_{ij}$, otteniamo:

$$< U(\mathbf{u}), U(\mathbf{u}) >= \alpha_1^2 + \cdots + \alpha_n^2$$

Dunque per (11.1) abbiamo $< U(\mathbf{u}), U(\mathbf{u}) >=< \mathbf{u}, \mathbf{u} >$.

Questo conclude la dimostrazione. □

## 11.2 MATRICI ORTOGONALI

Vogliamo ora definire l'insieme delle *matrici ortogonali* e stabilire la relazione che sussiste tra matrici ortogonali e applicazioni lineari ortogonali.

Consideriamo lo spazio vettoriale $\mathbb{R}^n$ dotato del prodotto scalare standard $<,>_e$ e cioè:

$$< \mathbf{u}, \mathbf{v} >_e = (\alpha_1 \dots \alpha_n) \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \alpha_1\beta_1 + \cdots + \alpha_n\beta_n \tag{11.2}$$

ove $\mathbf{u} = (\alpha_1 \dots \alpha_n)^T$, $\mathbf{v} = (\beta_1 \dots \beta_n)^T$ sono vettori (colonna) in $\mathbb{R}^n$.

**Definizione 11.2.1** Una matrice $A \in \mathrm{M}_n(\mathbb{R})$ si dice *ortogonale* se conserva il prodotto scalare standard, cioè se:

$$< A\mathbf{u}, A\mathbf{v} >_e = < \mathbf{u}, \mathbf{v} >_e \qquad \text{per ogni} \quad \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$$

ove con $A\mathbf{u}$ intendiamo il prodotto righe per colonne della matrice $A$ per il vettore colonna $\mathbf{u}$.

Ricordiamo ora che, data una matrice $A \in \mathrm{M}_n(\mathbb{R})$, denotiamo con $L_A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ l'applicazione lineare ad essa associata una volta fissate le basi canoniche in dominio e codominio; se $\mathbf{u} \in \mathbb{R}^n$, $L_A(\mathbf{u}) = A\mathbf{u}$.

La relazione tra le matrici ortogonali e le applicazioni lineari ortogonali è espressa dalla seguente proposizione la cui dimostrazione è immediata se confrontiamo le Definizioni 11.1.1 e 11.2.1.

**Proposizione 11.2.2** *Una matrice $A \in \mathrm{M}_n(\mathbb{R})$ è ortogonale se e solo se l'applicazione lineare $L_A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ è ortogonale rispetto al prodotto scalare standard in $\mathbb{R}^n$* (si veda la formula (11.2)).

Le matrici ortogonali godono di molte proprietà, che riassumiamo in una proposizione.

**Proposizione 11.2.3** *Sia $A \in \mathrm{M}_n(\mathbb{R})$ una matrice quadrata. Le seguenti affermazioni sono equivalenti.*

a) *$A$ è ortogonale, cioè $< A\mathbf{u}, A\mathbf{v} >_e = < \mathbf{u}, \mathbf{v} >_e$ per ogni $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.*

b) *$A$ conserva la norma dei vettori, cioè: $< A\mathbf{u}, A\mathbf{u} >_e = < \mathbf{u}, \mathbf{u} >_e$ per ogni $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.*

c) *$A^T A = I = A^T A$ ove $I$ è la matrice identità. In particolare $A^{-1} = A^T$.*

d) *Le colonne e le righe di $A$ formano, rispettivamente, due basi ortonormali.*

**Dimostrazione** – L'equivalenza tra (a) e (b) è immediata per la Proposizione 11.1.2. Prima di procedere con la dimostrazione, ricordiamo che il prodotto scalare standard di due vettori colonna $\mathbf{u} = (\alpha_1 \dots \alpha_n)^T$, $\mathbf{v} = (\beta_1 \dots \beta_n)^T$, può essere espresso come il prodotto righe per colonne:

$$\mathbf{u}^T \mathbf{v} = (\alpha_1 \dots \alpha_n) \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$$

Dunque abbiamo che:

$$< A\mathbf{u}, A\mathbf{v} >_e = (A\mathbf{u})^T (A\mathbf{v}) = \mathbf{u}^T (A^T A) \mathbf{v} \tag{11.3}$$

ricordando che se $X$ e $Y$ sono matrici abbiamo $(XY)^T = Y^T X^T$.

(a) $\Longrightarrow$ (c). Per la formula (11.3) dalle ipotesi segue che

$$\mathbf{u}^T (A^T A) \mathbf{v} = \mathbf{u}^T \mathbf{v}$$

per ogni $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. Sia $A^T A = C = (c_{ij})_{i,j=1,\dots,n}$ e scegliamo $\mathbf{u} = \mathbf{e}_i$ e $\mathbf{v} = \mathbf{e}_j$ nella formula precedente. Notiamo che

$$\mathbf{e}_i^T (A^T A) \mathbf{e}_j = c_{ij}$$

mentre

$$\mathbf{e}_i^T \mathbf{e}_j = \delta_{ij}$$

ove $\delta_{ij}$ è il delta di Kronecker. Dunque abbiamo

$$A^T A = I$$

in quanto abbiamo mostrato che le matrici $C = A^T A$ e $I$ hanno (ordinatamente) gli stessi coefficienti.

In particolare segue che il determinante di $A$ è diverso da zero, quindi $A$ è invertibile e moltiplicando a sinistra entrambi i membri dell'uguaglianza per la matrice $A^{-1}$ otteniamo che $A^T = A^{-1}$, quindi $AA^T = I$.

(c) $\Longrightarrow$ (a). Supponiamo che sia $AA^T = I$. Allora per la formula (11.3) abbiamo che

$$< A\mathbf{u}, A\mathbf{v} >_e = (A\mathbf{u})^T (A\mathbf{v}) = \mathbf{u}^T (A^T A) \mathbf{v} = \mathbf{u}^T \mathbf{v} =< \mathbf{u}, \mathbf{v} >_e,$$

come volevasi dimostrare.

(c) $\Longleftrightarrow$ (d). Se $\mathbf{v}_1, \dots, \mathbf{v}_n$ sono i vettori colonna di $A$ l'equazione $A^T A = I$ equivale a:

$$\mathbf{v}_i \cdot \mathbf{v}_j = 0, \quad i \neq j, \qquad \mathbf{v}_i \cdot \mathbf{v}_i = 1.$$

cioè equivale al fatto che tali colonne siano vettori ortonormali.

Analogamente, ragionando sui vettori riga, l'equazione $A A^T = I$ equivale al fatto che le righe di $A$ siano vettori ortonormali. Concludiamo ricordando che $A$ è una matrice invertibile se e solo se le sue righe (colonne) formano una base di $\mathbb{R}^n$.

□

◆ **Osservazione 11.2.4** Osserviamo che nella proposizione precedente la condizione (c) $AA^T = I = A^T A$ è equivalente a una sola delle due uguaglianze. Infatti, in generale, date $X, Y \in \mathrm{M}_n(\mathbb{R})$, se $XY = I$ allora per il Teorema di Binet, $\det(XY) = \det(X)\det(Y) = 1$ dunque entrambe le matrici sono invertibili e sono una l'inversa dell'altra (si veda l'Osservazione 7.4.3 del Capitolo 7).

Allo stesso modo nella parte (d) le due condizioni "le colonne di $A$ formano una base" e "le righe di $A$ formano una base" sono tra di loro equivalenti. Ciò è una conseguenza del fatto generale che il rango righe di una matrice è uguale al suo rango colonne.

Grazie all'Osservazione 10.6.4 e per quanto detto qui sopra abbiamo subito la seguente proposizione.

**Proposizione 11.2.5** *Sia $V$ uno spazio vettoriale reale di dimensione finita con un prodotto scalare definito positivo. Sia $U : V \longrightarrow V$ un'applicazione lineare e sia $A$ la matrice ad essa associata rispetto a una base $\mathcal{B}$ ortonormale. Allora $U$ è un'applicazione ortogonale se e solo se la matrice $A$ è ortogonale.*

Concludiamo questa sezione con alcune proprietà delle matrici ortogonali.

◆ **Osservazione 11.2.6**

1. Il determinante di una matrice ortogonale può essere solo $\pm 1$. Infatti
$$\det(A^T A) = \det(A^T)\det(A) = \det(A)^2 = \det(I) = 1$$
perché il determinante di una matrice è uguale al determinante della sua trasposta, per il Corollario 7.9.8.
2. Si noti inoltre che, se $A$ è una matrice ortogonale, abbiamo che $A^{-1} = A^T$ è ortogonale. Questo segue immediatamente dalla proprietà (c) della Proposizione 11.2.3. Per verificare che $(A^T)^T A^T = I$ basta ricordare che $(A^T)^T = A$.
3. Se $A$ e $B$ sono matrici ortogonali allora $AB$ è una matrice ortogonale. Infatti possiamo verificare direttamente che:
$$(AB)^T(AB) = B^T A^T AB = B^T IB = B^T B = I$$

Vediamo alcuni esempi salienti di matrici ortogonali, lasciamo per esercizio la facile verifica che $A^{-1} = A^T$.

---

**Esempio 11.2.7**

1. Rotazioni in $\mathbb{R}^2$. Sia
$$A = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}$$

Come il lettore potrà facilmente verificare, tale matrice è associata a una rotazione del piano di un angolo $t$, avente per centro l'origine degli assi cartesiani.

2. Riflessione rispetto alla retta $x = y$ in $\mathbb{R}^2$. Sia
$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Come il lettore potrà facilmente verificare, tale matrice è associata a una riflessione, cioè i punti che appartengono alla retta di equazione $x = y$ sono fissati, e l'immagine di ogni altro punto del piano è il suo simmetrico rispetto a tale retta.

3. Rotazioni in $\mathbb{R}^3$. Consideriamo una applicazione lineare che ruota ogni vettore applicato nell'origine intorno a un asse dato anch'esso passante per l'origine, di un certo angolo fissato $t$. Tale applicazione lineare conserva la norma euclidea e dunque è una trasformazione ortogonale rispetto al prodotto scalare euclideo. Un celebre Teorema di Eulero stabilisce che ogni trasformazione ortogonale di $\mathbb{R}^3$ la cui matrice abbia determinante uguale a 1 è di questo tipo.

## 11.3 APPLICAZIONI LINEARI SIMMETRICHE

In questa sezione vogliamo definire le applicazioni lineari simmetriche e vedere quali sono le matrici ad esse associate fissata un'opportuna base in dominio e codominio.

**Definizione 11.3.1** Sia $V$ uno spazio vettoriale reale con un prodotto scalare $<,>$ definito positivo. Diciamo che l'applicazione lineare $T : V \longrightarrow V$ è *simmetrica* se

$$< T(\mathbf{u}), \mathbf{v} >=< \mathbf{u}, T(\mathbf{v}) > \qquad \text{per ogni} \quad \mathbf{u}, \mathbf{v} \in V \tag{11.4}$$

Vediamo ora il caso particolare di una applicazione lineare $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ e del prodotto scalare euclideo. Fissata la base canonica $\mathcal{C}$, possiamo scrivere:

$$< \mathbf{u}, \mathbf{v} >_e = \mathbf{u}^T I \mathbf{v} = \mathbf{u}^T \mathbf{v} \qquad \text{per ogni} \quad \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$$

ove $<,>_e$ denota il prodotto scalare standard in $\mathbb{R}^n$ e $\mathbf{u}$ denota un vettore *colonna* in $\mathbb{R}^n$. Questa notazione corrisponde all'identificazione di un vettore in $\mathbb{R}^n$ con la colonna delle sue coordinate rispetto alla base canonica.

In altre parole, il prodotto euclideo è associato alla matrice identità, fissata la base canonica. Consideriamo ora la matrice $A$ associata all'applicazione lineare $T$, fissata la base canonica, cioè $T = L_A$.

Possiamo dunque scrivere:

$$< T(\mathbf{u}), \mathbf{v} >_e = (A\mathbf{u})^T \mathbf{v} = \mathbf{u}^T A^T \mathbf{v} = \mathbf{u}^T A^T \mathbf{v} \qquad \text{per ogni} \quad \mathbf{u}, \mathbf{v} \in \mathbb{R}^n \tag{11.5}$$

Se $A$ è una matrice simmetrica, cioè $A = A^T$, abbiamo che $T$ è un'applicazione lineare simmetrica. Infatti da (11.5):

$$< T(\mathbf{u}), \mathbf{v} >_e =< A\mathbf{u}, \mathbf{v} >_e = \mathbf{u}^T A^T \mathbf{v} = \mathbf{u}^T A \mathbf{v} =< \mathbf{u}, A\mathbf{v} >_e =< \mathbf{u}, T(\mathbf{v}) >_e$$

Viceversa, se $T$ è simmetrica, vediamo che la matrice $A$ è simmetrica. Infatti:

$$< T(\mathbf{e}_i), \mathbf{e}_j >_e = (A\mathbf{e}_i)^T \mathbf{e}_j = \begin{pmatrix} a_{1i} & \dots & a_{ni} \end{pmatrix} \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & & \dots & 1 \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = a_{ji}$$

$$< \mathbf{e}_i, T(\mathbf{e}_j) >_e = \mathbf{e}_i^T A \mathbf{e}_j = \begin{pmatrix} 0 \dots 1 \dots 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & & \dots & 1 \end{pmatrix} \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix} = a_{ij}$$

e dunque per la condizione (11.4) abbiamo $a_{ij} = a_{ji}$, cioè la matrice $A$ è simmetrica.

Lo stesso accade per il caso generale: un'applicazione lineare simmetrica è associata, fissata una base ortonormale in dominio e codominio, a una matrice simmetrica.

**Proposizione 11.3.2** *Sia $V$ uno spazio vettoriale reale, con un prodotto scalare $<,>$ definito positivo e sia $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ una base ortonormale. Sia $T : V \longrightarrow V$ un'applicazione lineare e sia $A = (a_{ij})$ la matrice associata a $T$ rispetto alla base $\mathcal{B}$. Allora $T$ è simmetrica se e solo se la matrice $A$ è simmetrica.*

**Dimostrazione** – Dimostriamo prima che la matrice associata all'applicazione lineare simmetrica $T$ rispetto alla base $\mathcal{B}$ è simmetrica. Abbiamo:

$$T(\mathbf{v}_j) = a_{1j}\mathbf{v}_1 + \dots + a_{nj}\mathbf{v}_n$$

da cui:

$$< T(\mathbf{v}_j), \mathbf{v}_i > \; = a_{1j} < \mathbf{v}_1, \mathbf{v}_i > + \dots + a_{ij} < \mathbf{v}_i, \mathbf{v}_i > + \dots$$

$$+ a_{nj} < \mathbf{v}_n, \mathbf{v}_j >= a_{ij}$$

e analogamente

$$< \mathbf{v}_j, T(\mathbf{v}_i) > \; = a_{j1} < \mathbf{v}_j, \mathbf{v}_1 > + \dots + a_{ji} < \mathbf{v}_j, \mathbf{v}_i > + \dots$$

$$+ a_{nj} < \mathbf{v}_n, \mathbf{v}_j >= a_{ji}$$

Quindi abbiamo che:

$$a_{ij} =< T(\mathbf{v}_j), \mathbf{v}_i >=< \mathbf{v}_j, T(\mathbf{v}_i) >= a_{ji} \tag{11.6}$$

D'altra parte, l'equazione (11.6) dimostra che se la matrice $A$ associata a $T$ rispetto a una base ortonormale è simmetrica allora $< T(\mathbf{v}_j), \mathbf{v}_i >=< \mathbf{v}_j, T(\mathbf{v}_i) >$. Poiché ogni $\mathbf{u}, \mathbf{v} \in V$ è combinazione lineare di $\mathbf{v}_1, \dots, \mathbf{v}_n$ abbiamo che l'uguaglianza precedente dà immediatamente $< T(\mathbf{u}), \mathbf{v} >=< \mathbf{u}, T(\mathbf{v}) >$ per ogni $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. Ciò conclude la dimostrazione. □

♦ **Osservazione 11.3.3** Si noti che, se $A$ è una matrice $n \times n$ a coefficienti reali, nelle osservazioni preliminari alla proposizione precedente abbiamo dimostrato che:

$$< A\mathbf{u}, \mathbf{v} >_e =< \mathbf{u}, A^T\mathbf{v} >_e$$

ove $<,>_e$ è il prodotto scalare euclideo e $\mathbf{u}$, $\mathbf{v} \in \mathbb{R}^n$ sono vettori colonna.

## 11.4 IL TEOREMA SPETTRALE

In questa sezione vogliamo enunciare e dimostrare uno dei risultati più importanti dell'algebra lineare: il Teorema spettrale.

Iniziamo con due lemmi seguiti da alcune osservazioni praticamente immediate. Il primo lemma ci dice che ogni matrice simmetrica ammette almeno un autovalore reale, il secondo lemma, che in realtà è una conseguenza del primo, ci dice che autovettori relativi ad autovalori distinti di una matrice simmetrica sono sempre perpendicolari tra loro. Questi sono i passi chiave per la dimostrazione del Teorema spettrale.

**Lemma 11.4.1** *Sia $A \in \mathrm{M}_n(\mathbb{R})$ una matrice simmetrica. Allora $A$ ammette un autovalore reale.*

**Dimostrazione** – Si veda l'Appendice 11.7. □

Osserviamo che, una volta dimostrato che una matrice $A$ a coefficienti reali ammette un autovalore $\lambda \in \mathbb{R}$, abbiamo subito che l'autospazio $V_\lambda = \ker(A - \lambda I) \subseteq \mathbb{R}^n$ non contiene solo il vettore nullo e quindi esiste anche un autovettore *reale* relativo all'autovalore $\lambda$.

Andiamo ora a stabilire un altro risultato che si rivelerà fondamentale nella dimostrazione del Teorema spettrale.

**Lemma 11.4.2** *Siano $A \in \mathrm{M}_n(\mathbb{R})$ una matrice simmetrica, $\lambda$ un suo autovalore reale e $\mathbf{u} \in \mathbb{R}^n$ un autovettore di autovalore $\lambda$. Sia $\mathbf{w}$ un vettore perpendicolare a $\mathbf{u}$ rispetto al prodotto scalare euclideo. Allora $\mathbf{u}$ è perpendicolare a $A\mathbf{w}$.*

**Dimostrazione** – Poiché $\mathbf{u}$ è perpendicolare a $\mathbf{w}$ e $A = A^T$ abbiamo

$$0 = \lambda < \mathbf{u}, \mathbf{w} >_e = < \lambda\mathbf{u}, \mathbf{w} >_e = < A\mathbf{u}, \mathbf{w} >_e = < \mathbf{u}, A\mathbf{w} >_e$$

e dunque anche $A\mathbf{w}$ è perpendicolare a $\mathbf{u}$. □

Abbiamo quasi immediatamente un corollario particolarmente importante.

**Corollario 11.4.3** *Siano $\lambda$ e $\mu$ autovalori distinti di una matrice simmetrica $A$ e $\mathbf{u}$, $\mathbf{w}$ due autovettori corrispondenti. Allora $\mathbf{u}$ è perpendicolare a $\mathbf{w}$.*

**Dimostrazione** – Dobbiamo mostrare che $< \mathbf{u}, \mathbf{w} >_e = 0$. Abbiamo

$$\lambda < \mathbf{u}, \mathbf{w} >_e = < A\mathbf{u}, \mathbf{w} >_e = < \mathbf{u}, A\mathbf{w} >_e = \mu < \mathbf{u}, \mathbf{w} >_e$$

Dunque $(\lambda - \mu) < \mathbf{u}, \mathbf{w} >_e = 0$ e poiché $\lambda \neq \mu$ otteniamo $< \mathbf{u}, \mathbf{w} >_e = 0$. □

Riassumiamo con un corollario quanto abbiamo dimostrato per le matrici simmetriche nel linguaggio delle applicazioni lineari simmetriche.

**Corollario 11.4.4** *Sia $V$ uno spazio vettoriale reale di dimensione finita con un prodotto scalare definito positivo e sia $T : V \longrightarrow V$ un'applicazione lineare simmetrica. Allora:*

1. *$T$ ammette almeno un autovalore reale.*
2. *Se $\mathbf{w} \in V$ è perpendicolare a un autovettore $\mathbf{u}$ di $T$ allora anche $T(\mathbf{w})$ è perpendicolare a $\mathbf{u}$.*
3. *Autovettori di $T$ relativi ad autovalori distinti sono perpendicolari tra loro.*

**Dimostrazione** – Sia $\mathcal{B}$ una base ortonormale per il prodotto scalare definito positivo. Una tale base esiste sempre grazie all'algoritmo di Gram-Schmidt. Per la Proposizione 11.3.2, la matrice $A$ associata a $T$ rispetto alla base $\mathcal{B}$ è simmetrica. Dunque le affermazioni del corollario seguono immediatamente dai Lemmi 11.4.1, 11.4.2 e dal Corollario 11.4.3. □

Possiamo finalmente enunciare il Teorema spettrale simultaneamente per le matrici simmetriche reali e le applicazioni lineari simmetriche.

**Teorema 11.4.5** *Sia $V$ uno spazio vettoriale reale di dimensione $n$ con un prodotto scalare definito positivo. Sia $T : V \longrightarrow V$ un'applicazione lineare simmetrica e sia $A \in \mathrm{M}_n(\mathbb{R})$ la matrice simmetrica associata a $T$ rispetto a una base ortonormale $\mathcal{B}$.*

*Allora:*

- *$T$ è diagonalizzabile e inoltre esiste una base ortonormale $\mathcal{N}$ costituita da autovettori di $T$.*
- *$A$ è diagonalizzabile mediante una matrice ortogonale, cioè esiste una matrice $P$ ortogonale tale che $D = P^{-1}AP$ sia diagonale.*

Prima della dimostrazione osserviamo che le due affermazioni dell'enunciato sono completamente equivalenti. La base ortonormale $\mathcal{N}$ è una base di autovettori $T$ mutuamente perpendicolari con norma 1. L'esistenza di tale base di autovettori equivale all'esistenza di una matrice $P$ ortogonale tale che $P^{-1}AP$ sia diagonale. Tale matrice ha per colonne le coordinate degli autovettori rispetto alla base $\mathcal{B}$.

**Dimostrazione** – Sia $\lambda_1$ un autovalore reale di $T$ e sia $\mathbf{u}_1 \in V$ un autovettore di norma 1 relativo all'autovalore $\lambda_1$. Sappiamo che tali $\lambda_1$ e $\mathbf{u}_1$ esistono per il Lemma 11.4.1. Sia $W_1 = \langle \mathbf{u}_1 \rangle^{\perp}$. Allora abbiamo $\dim(W_1) = n-1$. Consideriamo ora l'applicazione lineare $T_1 = T|_{W_1}$, cioè guardiamo la restrizione di $T_1$ al solo sottospazio $W_1$, quindi $T_1 : W_1 \longrightarrow V$. Per il Lemma 11.4.2, poiché $\mathbf{u}$ è perpendicolare anche a $T_1(\mathbf{w})$ per ogni $\mathbf{w} \in W_1$ abbiamo che $\mathrm{Im}\,(T_1) \subseteq W_1 = \langle \mathbf{u}_1 \rangle^{\perp}$ e pertanto possiamo scrivere $T_1 : W_1 \longrightarrow W_1$. Ripetiamo ora tutti i ragionamenti fatti fino a questo momento per $T : V \longrightarrow V$ anche per l'applicazione simmetrica $T_1 : W_1 \longrightarrow W_1$. Tale applicazione ha un autovalore reale $\lambda_2$ e un autovettore reale di norma 1, $\mathbf{u}_2 \in W_1 \subseteq \mathbb{R}^n$. Chiaramente poiché $T_1$ non è altro che $T$ ristretta a $W_1$, l'autovalore $\lambda_2$ e il corrispondente autovettore $\mathbf{u}_2$ saranno anche rispettivamente un autovalore e un autovettore di $T$. Dunque ragionando come sopra, definiamo $W_2 = \langle \mathbf{u}_2 \rangle^{\perp} \subseteq W_1$, $\dim(W_2) = n-2$. Notiamo che ogni vettore di $W_2$ è anche perpendicolare a $\mathbf{u}_1$ in quanto $W_2 \subseteq W_1$. È chiaro che procedendo in questo modo dopo $n$ passi avremo trovato $n$ autovettori $\mathbf{u}_1, \ldots, \mathbf{u}_n$ di $T$ di norma 1 e perpendicolari tra loro, dunque essi formano la base ortonormale di autovettori richiesta e pertanto $T$ è diagonalizzabile. □

**Esempio 11.4.6**

Consideriamo la matrice reale

$$A = \begin{pmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{pmatrix}$$

Gli autovalori sono $\lambda = 7$ con molteplicità algebrica 2 e $\lambda = -2$ con molteplicità algebrica 1. Gli autospazi sono:

$$V_7 = \langle \mathbf{v}_1 = (1,0,1), \mathbf{v}_2 = (-1/2,1,0) \rangle,$$
$$V_{-2} = \langle \mathbf{v}_3 = (-1,-1/2,1) \rangle$$

La matrice $Q$ avente per colonne gli autovettori della base di autovettori $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ diagonalizza la matrice $A$, tuttavia non è ortogonale:

$$D = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{pmatrix} = Q^{-1}AQ, \qquad Q = \begin{pmatrix} 1 & -1/2 & -1 \\ 0 & 1 & -1/2 \\ 1 & 0 & 1 \end{pmatrix}$$

Se vogliamo diagonalizzare $A$ tramite una cambio di base ortogonale, dobbiamo ortogonalizzare con l'algoritmo di Gram Schmidt la base di autovettori di ciascun autospazio. La matrice ortogonale che stiamo cercando è:

$$P = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{5} & -2/3 \\ 0 & 2/\sqrt{5} & -1/3 \\ 1/\sqrt{2} & 0 & 2/3 \end{pmatrix}$$

e abbiamo $D = P^{-1}AP$.

## 11.5 ESERCIZI SVOLTI

**Esercizio 11.5.1**

*Si consideri l'applicazione lineare* $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *definita da:* $T(\mathbf{e}_1) = \mathbf{e}_1 + 2\mathbf{e}_2$, $T(\mathbf{e}_2) = 2\mathbf{e}_1 + \mathbf{e}_2$. *Si verifichi che si tratta di un'applicazione lineare simmetrica e si determini una base* $\mathcal{B}$ *ortonormale rispetto alla quale* $T$ *sia associata a una matrice diagonale.*

**Svolgimento**

La matrice associata a $T$ rispetto alla base canonica (in dominio e codominio) è:

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

Poiché $A$ è una matrice simmetrica, $T$ è una applicazione lineare simmetrica.

Gli autovalori di $A$ (e di $T$) sono: $-1$, $3$. Calcoliamo gli autospazi:

$$V_{-1} = \langle (1,-1) \rangle, \qquad V_3 = \langle (1,1) \rangle$$

Per ottenere una base ortonormale dobbiamo scegliere autovettori di norma uguale a 1. Pertanto:

$$\mathcal{B} = \{(1/\sqrt{2}, -1/\sqrt{2}), (1/\sqrt{2}, 1/\sqrt{2})\}$$

è una base che soddisfa i requisiti richiesti.

■

## Esercizio 11.5.2

*Sia* $\mathbf{u} = \mathbf{e}_1 - 2\mathbf{e}_3 \in \mathbb{R}^3$ *e si consideri l'applicazione lineare* $\mathrm{proj}_{\mathbf{u}} : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *che a ogni vettore* $\mathbf{v}$ *di* $\mathbb{R}^3$ *associa la sua proiezione ortogonale sul vettore* $\mathbf{u}$ *rispetto al prodotto scalare euclideo, cioè*

$$\mathrm{proj}_{\mathbf{u}}(\mathbf{v}) = \frac{<\mathbf{v},\mathbf{u}>_e}{<\mathbf{u},\mathbf{u}>_e}\mathbf{u}$$

(si veda la Definizione 10.6.2). *Si verifichi che si tratta di un'applicazione lineare simmetrica e si determini una base $\mathcal{B}$ ortonormale rispetto alla quale* $\mathrm{proj}_{\mathbf{u}}$ *sia associata a una matrice diagonale $D$. Si scriva poi esplicitamente $D$.*

**Svolgimento**

Determiniamo la matrice $A$ associata a $\mathrm{proj}_{\mathbf{u}}$ rispetto alla base canonica (in dominio e codominio). Abbiamo:

$\mathrm{proj}_{\mathbf{u}}(\mathbf{e}_1) = \frac{<\mathbf{e}_1,\mathbf{u}>_e}{<\mathbf{u},\mathbf{u}>_e}(\mathbf{e}_1 - 2\mathbf{e}_3) = \frac{1}{5}(\mathbf{e}_1 - 2\mathbf{e}_3) = \frac{1}{5}\mathbf{e}_1 - \frac{2}{5}\mathbf{e}_3$

$\mathrm{proj}_{\mathbf{u}}(\mathbf{e}_2) = \frac{<\mathbf{e}_2,\mathbf{u}>_e}{<\mathbf{u},\mathbf{u}>_e}(\mathbf{e}_1 - 2\mathbf{e}_3) = 0(\mathbf{e}_1 - 2\mathbf{e}_3) = \mathbf{0}$

$\mathrm{proj}_{\mathbf{u}}(\mathbf{e}_3) = \frac{<\mathbf{e}_3,\mathbf{u}>_e}{<\mathbf{u},\mathbf{u}>_e}(\mathbf{e}_1 - 2\mathbf{e}_3) = -\frac{2}{5}(\mathbf{e}_1 - 2\mathbf{e}_3) = -\frac{2}{5}\mathbf{e}_1 + \frac{4}{5}\mathbf{e}_3$

Quindi:

$$A = \begin{pmatrix} \frac{1}{5} & 0 & -\frac{2}{5} \\ 0 & 0 & 0 \\ -\frac{2}{5} & 0 & \frac{4}{5} \end{pmatrix}$$

Poiché $A$ è una matrice simmetrica, $\mathrm{proj}_{\mathbf{u}}$ è una applicazione lineare simmetrica.

Per determinare una base $\mathcal{B}$ ortonormale rispetto alla quale $\mathrm{proj}_{\mathbf{u}}$ sia associata a una matrice diagonale si può procedere come nell'esercizio precedente oppure osservare che $\mathbf{u}$ è un autovettore di $\mathrm{proj}_{\mathbf{u}}$ relativo all'autovalore 1, infatti: $\mathrm{proj}_{\mathbf{u}}(\mathbf{u}) = \frac{<\mathbf{u},\mathbf{u}>_e}{<\mathbf{u},\mathbf{u}>_e}\mathbf{u} = \mathbf{u}$. Inoltre se $\mathbf{v}$ è un qualsiasi vettore non nullo ortogonale a $\mathbf{u}$, cioè $<\mathbf{v},\mathbf{u}>_e = 0$, si ha che $\mathrm{proj}_{\mathbf{u}}(\mathbf{u}) = 0\mathbf{u}$, quindi $\mathbf{v}$ è un autovettore di $\mathrm{proj}_{\mathbf{u}}$ relativo all'autovalore 0.

Consideriamo ora una base ortonormale $\mathcal{B}$ ottenuta applicando l'algoritmo di Gram-Schmidt alla base $\{\mathbf{u}, \mathbf{e}_1, \mathbf{e}_2\}$.

Usando le notazioni del Teorema 10.6.3, abbiamo:

$$\mathbf{u}_1 = \mathbf{u}, \qquad \mathbf{f}_1 = \frac{1}{\|\mathbf{u}\|}\mathbf{u} = \frac{1}{\sqrt{5}}\mathbf{e}_1 - \frac{2}{\sqrt{5}}\mathbf{e}_3$$

$$\mathbf{u}_2 = \mathbf{e}_1 - \mathrm{proj}_{\mathbf{u}}(\mathbf{e}_1) = \frac{4}{5}\mathbf{e}_1 + \frac{2}{5}\mathbf{e}_3, \qquad \mathbf{f}_2 = \frac{1}{\|\mathbf{u}_2\|}\mathbf{u}_2 = \frac{2}{\sqrt{5}}\mathbf{e}_1 + \frac{1}{\sqrt{5}}\mathbf{e}_3$$

$$\mathbf{u}_3 = \mathbf{e}_2 - \mathrm{proj}_{\mathbf{u}}(\mathbf{e}_2) - \mathrm{proj}_{\mathbf{u}_2}(\mathbf{e}_2) = \mathbf{e}_2, \qquad \mathbf{f}_3 = \frac{1}{\|\mathbf{u}_3\|}\mathbf{u}_3 = \mathbf{e}_2.$$

Il vettore $\mathbf{f}_1$ è un multiplo di $\mathbf{u}$ quindi è autovettore di $\mathrm{proj}_{\mathbf{u}}$ di autovalore 1, i vettori $\mathbf{f}_2$, $\mathbf{f}_3$ per costruzione sono perpendicolari a $\mathbf{u}$ quindi sono autovettori di $\mathrm{proj}_{\mathbf{u}}$ di autovalore 0. Una base $\mathcal{B}$ con le proprietà richieste è quindi: $\mathcal{B} = \{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ e la matrice $D$ è:

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

## 11.6 ESERCIZI PROPOSTI

### Esercizio 11.6.1

*Si dica quali tra le seguenti matrici sono ortogonali, motivando la risposta.*

$$\text{a)}\begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix}, \quad b)\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \text{c)}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \text{d)}\begin{pmatrix} -1/\sqrt{5} & 4/\sqrt{45} & -2/3 \\ 2/\sqrt{5} & 2/\sqrt{5} & -1/3 \\ 0 & 5/\sqrt{5} & 2/3 \end{pmatrix}$$

### Esercizio 11.6.2

*Si consideri la matrice:*

$$A = \begin{pmatrix} 1/2 & 0 & \sqrt{3}/2 \\ 0 & 1 & 0 \\ a & b & c \end{pmatrix}$$

*Si determinino, se possibile, valori per $a$, $b$, $c$ che la rendano una matrice ortogonale.*

### Esercizio 11.6.3

a) *Si determini se esiste un valore dell'angolo $t$ per il quale la rotazione in $\mathbb{R}^2$ descritta nell'Esempio* 11.2.7 *è un'applicazione lineare simmetrica.*

b) *Si determini l'applicazione lineare che a ogni punto del piano, identificato con $\mathbb{R}^2$, associa il suo simmetrico rispetto all'asse delle ascisse, e si stabilisca se è un'applicazione lineare ortogonale.*

c) *Si determini se la proiezione ortogonale descritta nell'Esercizio* 11.5.2 *è un'applicazione lineare ortogonale.*

### Esercizio 11.6.4

*Diagonalizzare le seguenti matrici utilizzando una matrice ortogonale:*

$$\begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix}, \quad \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 3 & 2 & 0 \\ 2 & 6 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & -2 & 4 \end{pmatrix}$$

### Esercizio 11.6.5

*Si determini l'applicazione lineare* $\mathrm{proj}_{\mathbf{u}} : \mathbb{R}^2 \to \mathbb{R}^2$ *che a ogni vettore associa la sua proiezione sul vettore* $\mathbf{u} = \mathbf{e}_1 - \mathbf{e}_2$, *scrivendo esplicitamente la matrice $T$ ad essa associata rispetto alla base canonica. Si dica se si tratta di un'applicazione lineare simmetrica e/o ortogonale. Si determini, se possibile, una matrice diagonale simile a $T$.*

### Esercizio 11.6.6

*Sia* $\mathrm{proj}_{\mathbf{u}} : \mathbb{R}^2 \to \mathbb{R}^2$ *l'applicazione lineare che a ogni vettore associa la sua proiezione sul vettore* $\mathbf{u} = 3\mathbf{e}_1 - 4\mathbf{e}_2$ *e sia $A$ la matrice ad essa associata rispetto alla base canonica. Si stabilisca se esiste una matrice ortogonale $P$ tale che* $P^{-1}AP = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$.

## 11.7 APPENDICE: IL CASO COMPLESSO

In questa appendice vogliamo dare la dimostrazione del Lemma 11.4.1 e pertanto dobbiamo introdurre i *prodotti hermitiani*, che rappresentano una generalizzazione al caso complesso dei prodotti scalari. La lettura di questa appendice richiede familiarità con i numeri complessi. Invitiamo il lettore a consultare l'Appendice A.

Definiamo $\mathbb{C}^n = \{(x_1, \dots, x_n) \mid x_i \in \mathbb{C}\}$ l'insieme delle $n$-ple di numeri complessi. In completa analogia con quanto avviene per $\mathbb{R}^n$, possiamo definire una somma ed una moltiplicazione per uno scalare $\lambda \in \mathbb{C}$ in modo che le proprietà (1)-(8) della Definizione 2.3.1 siano verificate. Otteniamo pertanto che $\mathbb{C}^n$ è uno *spazio vettoriale complesso* (si veda l'Appendice A). Tutta la teoria che abbiamo svolto nei Capitoli (1)-(9) per gli spazi vettoriali reali vale anche per gli spazi vettoriali complessi, cioè per gli insiemi che soddisfino le proprietà (1)-(8) della Definizione 2.3.1, ove si sostituisca $\mathbb{C}$ a $\mathbb{R}$ come insieme degli scalari.

Vediamo ora un'importante generalizzazione della nozione di prodotto scalare.

**Definizione 11.7.1** Sia $V$ uno spazio vettoriale complesso. La funzione $<,>$: $V \times V \longrightarrow \mathbb{C}$ si dice un *prodotto hermitiano* se:

1. $<\mathbf{u}+\mathbf{u}', \mathbf{v}> = <\mathbf{u}, \mathbf{v}> + <\mathbf{u}', \mathbf{v}>$,
   $<\mathbf{u}, \mathbf{v}+\mathbf{v}'> = <\mathbf{u}, \mathbf{v}> + <\mathbf{u}, \mathbf{v}'>$ per ogni $\mathbf{u}$, $\mathbf{u}'$, $\mathbf{v}$, $\mathbf{v}' \in V$.

2. $<\lambda\mathbf{u}, \mathbf{v}> = \lambda <\mathbf{u}, \mathbf{v}>$,
   $<\mathbf{u}, \mu\mathbf{v}> = \overline{\mu} <\mathbf{u}, \mathbf{v}>$ per ogni $\mathbf{u}$, $\mathbf{v} \in V$ e per ogni $\lambda$, $\mu \in \mathbb{C}$.

3. $<\mathbf{u}, \mathbf{v}> = <\overline{\mathbf{v}, \mathbf{u}}>$ per ogni $\mathbf{u}$, $\mathbf{v} \in V$.

$<,>$ è *non degenere*, se $<\mathbf{u}, \mathbf{v}> = 0$ per ogni $\mathbf{v} \in V$, allora $\mathbf{u} = 0$.
$<,>$ è *definito positivo* se $<\mathbf{u}, \mathbf{u}> \geq 0$ per ogni $\mathbf{u} \in V$ e $<\mathbf{u}, \mathbf{u}> = 0$ se solo se $\mathbf{u} = 0$.

La differenza tra un prodotto hermitiano e un prodotto scalare è che per un prodotto hermitiano richiediamo la linearità della funzione $<\mathbf{u}, \cdot>: V \longrightarrow \mathbb{C}$, ma *l'antilinearità* della funzione $<\cdot, \mathbf{u}>: V \longrightarrow \mathbb{C}$, cioè per $\mathbf{u} \in V$ fissato abbiamo $<\mathbf{u}, \mu\mathbf{v}> = \overline{\mu} <\mathbf{u}, \mathbf{v}>$.

Ci interessa in modo particolare il seguente esempio.

---

**Esempio 11.7.2**

Nello spazio vettoriale complesso $\mathbb{C}^n$ definiamo:

$$<(x_1, \dots, x_n), (x_1', \dots, x_n')>_h = x_1\bar{x}_1' + \dots + x_n\bar{x}_n'$$

Lasciamo al lettore la facile verifica che si tratta di un prodotto hermitiano. Tale prodotto si dice il *prodotto hermitiano standard* in $\mathbb{C}^n$. È inoltre immediato verificare che tale prodotto è definito positivo, infatti:

$$<(x_1, \dots, x_n), (x_1, \dots, x_n)>_h = x_1\bar{x}_1 + \dots + x_n\bar{x}_n = |x_1|^2 + \dots + |x_n|^2 \geq 0$$

e $|x_1|^2 + \cdots + |x_n|^2 = 0$ se e solo se $x_1 = x_2 = \cdots = x_n = 0$. Dunque il prodotto hermitiano standard è anche non degenere in quanto definito positivo.
Consideriamo la base canonica $C = \{e_1, \dots, e_n\}$ di $\mathbb{C}^n$, ove $e_i$ è il vettore avente 1 nella posizione $i$-esima e tutti 0 altrove. Osserviamo che, come abbiamo fatto per i prodotti scalari, a ogni prodotto hermitiano $<,>$ in $\mathbb{C}^n$ possiamo associare una matrice $C$, con $c_{ij} =< e_i, e_j >$, tale che

$$< (x_1, \dots, x_n), (x_1', \dots, x_n') >= (x_1, \dots, x_n) C \begin{pmatrix} \bar{x}_1' \\ \vdots \\ \bar{x}_n' \end{pmatrix}$$

Nel caso del prodotto hermitiano standard, la matrice ad esso associata è la matrice identità $I$ poiché $< e_i, e_j >_h = \delta_{ij}$. Non è difficile dimostrare, in completa analogia con il caso del prodotti scalari, che una matrice $C$ è associata a un prodotto hermitiano se e solo se $C = \overline{C}^T$, cioè coincide con la sua trasposta complessa coniugata.

---

L'osservazione che segue è cruciale nella dimostrazione del Lemma 11.4.1.

♦ **Osservazione 11.7.3** Consideriamo lo spazio vettoriale complesso $\mathbb{C}^n$ con il prodotto hermitiano standard, descritto nell'esempio precedente.

Se $(x_1, \dots, x_n), (x_1', \dots, x_n') \in \mathbb{R}^n \subseteq \mathbb{C}^n$, poiché $\overline{x_i'} = x_i'$ abbiamo che:

$$< (x_1, \dots, x_n), (x_1', \dots, x_n') >_h = x_1 x_1' + \cdots + x_n x_n'$$

$$=< (x_1, \dots, x_n), (x_1', \dots, x_n') >_e$$

In altre parole, il prodotto hermitiano di vettori in $\mathbb{R}^n$ coincide con l'usuale prodotto euclideo in $\mathbb{R}^n$.

Se $A$ è una matrice a coefficienti reali, abbiamo che:

$$< Au, v >_h =< u, A^T v >_h, \qquad \text{per ogni } u, v \in \mathbb{C}^n \tag{11.7}$$

Infatti, (11.7) è vera per $u = e_i$, $v = e_j$ (ove gli $e_i$ sono i vettori della base canonica), dunque, per la linearità e l'antilinearità del prodotto hermitiano, non è difficile verificare che sia vera anche per vettori generici $u, v \in \mathbb{C}^n$.

Siamo dunque pronti per enunciare e dimostrare il Lemma 11.4.1.

**Lemma 11.7.4** *Sia $A \in M_n(\mathbb{R})$ una matrice simmetrica. Allora $A$ ammette un autovalore reale.*

**Dimostrazione** – $A$ è una matrice a coefficienti reali, tuttavia poiché i numeri reali sono contenuti nel campo complesso abbiamo anche che $A \in M_n(\mathbb{C})$. Per il Teorema fondamentale dell'algebra (vedi Appendice A), il polinomio caratteristico di $A$, $\det(A - \lambda I)$, si annulla in almeno un numero complesso, $\lambda_0 \in \mathbb{C}$. Vogliamo dimostrare che $\lambda_0$ è reale, cioè $\lambda_0 = \overline{\lambda}_0$. Sia $u \in \mathbb{C}^n$ un autovettore di autovalore $\lambda_0$ e sia $<,>_h$ è il prodotto hermitiano standard in $\mathbb{C}^n$, cioè:

$$< u, v >_h = u^T \overline{v}, \qquad \text{per ogni} \quad u, v \in \mathbb{C}^n$$

ove $u = (\alpha_1, \dots, \alpha_n)^T, v = (\beta_1, \dots, \beta_n)^T \in \mathbb{C}^n$ sono vettori colonna e $\overline{v} = (\overline{\beta}_1, \dots, \overline{\beta}_n)^T$. Abbiamo:

$$< Au, u >_h = (Au)^T \overline{u} = u^T A^T \overline{u} = u^T A \overline{u} = u^T (\overline{Au}) =< u, Au >_h$$

in quanto $A = A^T$ e $\overline{A} = A$ poiché $A$ è una matrice a coefficienti reali. Dunque

$$\lambda_0 < \mathbf{u}, \mathbf{u} >_h = < A\mathbf{u}, \mathbf{u} >_h = < \mathbf{u}, A\mathbf{u} >_h = \overline{\lambda}_0 < \mathbf{u}, \mathbf{u} >_h$$

Pertanto:

$$(\lambda_0 - \overline{\lambda}_0) < \mathbf{u}, \mathbf{u} >_h = 0$$

Dal fatto che $\mathbf{u} \neq \mathbf{0}$, in quanto è un autovettore, segue che $< \mathbf{u}, \mathbf{u} >_h \neq 0$, quindi $\lambda_0 = \overline{\lambda}_0$. □

Nella restante parte di questa appendice vogliamo rivedere i risultati che abbiamo enunciato in questo capitolo per uno spazio vettoriale reale con un prodotto scalare definito positivo, per il caso di uno spazio vettoriale complesso con un prodotto hermitiano definito positivo. Come vedremo tutti i teoremi principali, incluso il Teorema spettrale, hanno enunciati e dimostrazioni analoghe a quelle viste, che pertanto lasceremo per esercizio.

La lettura di questa parte non è necessaria per la comprensione del caso reale, la includiamo per completezza.

In completa analogia con il caso reale possiamo dare le seguenti definizioni, ove nel caso complesso le nozioni di applicazioni lineari o matrici *unitarie* e *hermitiane* sostituiscono le corrispettive nozioni di applicazioni o matrici reali ortogonali e simmetriche.

**Definizione 11.7.5** Sia $V$ uno spazio vettoriale complesso con un prodotto hermitiano definito positivo $< , >_h$. Diciamo che un'applicazione lineare $T : V \longrightarrow V$ è *unitaria* se

$$< T\mathbf{u}, T\mathbf{v} >_h = < \mathbf{u}, \mathbf{v} >_h$$

Diciamo che un'applicazione lineare $T : V \longrightarrow V$ è *hermitiana* se

$$< T\mathbf{u}, \mathbf{v} >_h = < \mathbf{u}, T\mathbf{v} >_h$$

Diciamo che una matrice a coefficienti complessi $A$ è *unitaria* se $A^{-1} = A^*$, ove $A^* = \overline{A^T}$; invece diciamo che è *hermitiana* se $A = A^*$, cioè $A$ coincide con la sua trasposta complessa coniugata. Notiamo che queste due operazioni, cioè trasposizione di $A$ e coniugazione di ogni coefficiente di $A$, commutano tra loro, cioè il risultato è indipendente da quale scegliamo di fare per prima.

È facile verificare che la condizione $A$ hermitiana, corrisponde al fatto che rispetto al prodotto scalare hermitiano in $\mathbb{C}^n$ si abbia:

$$< A\mathbf{u}, \mathbf{v} >_h = < \mathbf{u}, A\mathbf{v} >_h$$

Si noti inoltre che se $A$ è una matrice simmetrica reale abbiamo immediatamente che è anche una matrice hermitiana, infatti è banale che soddisfi la condizione $A = A^*$, in quanto il complesso coniugato di un numero reale è il numero reale stesso.

Dato uno spazio vettoriale $V$ complesso con un prodotto hermitiano $< , >_h$, diciamo che $\mathbf{u}, \mathbf{v} \in V$ sono *perpendicolari (ortogonali) tra loro* se $< \mathbf{u}, \mathbf{v} >_h = 0$.

e $|x_1|^2 + \cdots + |x_n|^2 = 0$ se e solo se $x_1 = x_2 = \cdots = x_n = 0$. Dunque il prodotto hermitiano standard è anche non degenere in quanto definito positivo.
Consideriamo la base canonica $\mathcal{C} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ di $\mathbb{C}^n$, ove $\mathbf{e}_i$ è il vettore avente 1 nella posizione $i$-esima e tutti 0 altrove. Osserviamo che, come abbiamo fatto per i prodotti scalari, a ogni prodotto hermitiano $<,>$ in $\mathbb{C}^n$ possiamo associare una matrice $C$, con $c_{ij} =< \mathbf{e}_i, \mathbf{e}_j >$, tale che

$$< (x_1, \dots, x_n), (x_1', \dots, x_n') >= (x_1, \dots, x_n) C \begin{pmatrix} \overline{x}_1' \\ \vdots \\ \overline{x}_n' \end{pmatrix}$$

Nel caso del prodotto hermitiano standard, la matrice ad esso associata è la matrice identità $I$ poiché $< \mathbf{e}_i, \mathbf{e}_j >_h = \delta_{ij}$. Non è difficile dimostrare, in completa analogia con il caso del prodotti scalari, che una matrice $C$ è associata a un prodotto hermitiano se e solo se $C = \overline{C}^T$, cioè coincide con la sua trasposta complessa coniugata.

---

L'osservazione che segue è cruciale nella dimostrazione del Lemma 11.4.1.

♦ **Osservazione 11.7.3** Consideriamo lo spazio vettoriale complesso $\mathbb{C}^n$ con il prodotto hermitiano standard, descritto nell'esempio precedente.

Se $(x_1, \dots, x_n), (x_1', \dots, x_n') \in \mathbb{R}^n \subseteq \mathbb{C}^n$, poiché $\overline{x_i'} = x_i'$ abbiamo che:

$$\begin{aligned} < (x_1, \dots, x_n), (x_1', \dots, x_n') >_h &= x_1 x_1' + \cdots + x_n x_n' \\ &=< (x_1, \dots, x_n), (x_1', \dots, x_n') >_e \end{aligned}$$

In altre parole, il prodotto hermitiano di vettori in $\mathbb{R}^n$ coincide con l'usuale prodotto euclideo in $\mathbb{R}^n$.

Se $A$ è una matrice a coefficienti reali, abbiamo che:

$$< A\mathbf{u}, \mathbf{v} >_h =< \mathbf{u}, A^T \mathbf{v} >_h, \qquad \text{per ogni } \mathbf{u}, \mathbf{v} \in \mathbb{C}^n \tag{11.7}$$

Infatti, (11.7) è vera per $\mathbf{u} = \mathbf{e}_i$, $\mathbf{v} = \mathbf{e}_j$ (ove gli $\mathbf{e}_i$ sono i vettori della base canonica), dunque, per la linearità e l'antilinearità del prodotto hermitiano, non è difficile verificare che sia vera anche per vettori generici $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$.

Siamo dunque pronti per enunciare e dimostrare il Lemma 11.4.1.

**Lemma 11.7.4** *Sia $A \in \mathrm{M}_n(\mathbb{R})$ una matrice simmetrica. Allora $A$ ammette un autovalore reale.*

**Dimostrazione** – $A$ è una matrice a coefficienti reali, tuttavia poiché i numeri reali sono contenuti nel campo complesso abbiamo anche che $A \in \mathrm{M}_n(\mathbb{C})$. Per il Teorema fondamentale dell'algebra (vedi Appendice A), il polinomio caratteristico di $A$, $\det(A - \lambda I)$, si annulla in almeno un numero complesso, $\lambda_0 \in \mathbb{C}$. Vogliamo dimostrare che $\lambda_0$ è reale, cioè $\lambda_0 = \overline{\lambda}_0$. Sia $\mathbf{u} \in \mathbb{C}^n$ un autovettore di autovalore $\lambda_0$ e sia $<,>_h$ è il prodotto hermitiano standard in $\mathbb{C}^n$, cioè:

$$< \mathbf{u}, \mathbf{v} >_h = \mathbf{u}^T \overline{\mathbf{v}}, \qquad \text{per ogni} \quad \mathbf{u}, \mathbf{v} \in \mathbb{C}^n$$

ove $\mathbf{u} = (\alpha_1, \dots, \alpha_n)^T, \mathbf{v} = (\beta_1, \dots, \beta_n)^T \in \mathbb{C}^n$ sono vettori colonna e $\overline{\mathbf{v}} = (\overline{\beta}_1, \dots, \overline{\beta}_n)^T$. Abbiamo:

$$< A\mathbf{u}, \mathbf{u} >_h = (A\mathbf{u})^T \overline{\mathbf{u}} = \mathbf{u}^T A^T \overline{\mathbf{u}} = \mathbf{u}^T A \overline{\mathbf{u}} = \mathbf{u}^T (\overline{A\mathbf{u}}) =< \mathbf{u}, A\mathbf{u} >_h$$

in quanto $A = A^T$ e $\overline{A} = A$ poiché $A$ è una matrice a coefficienti reali. Dunque

$$\lambda_0 < \mathbf{u}, \mathbf{u} >_h = < A\mathbf{u}, \mathbf{u} >_h = < \mathbf{u}, A\mathbf{u} >_h = \overline{\lambda}_0 < \mathbf{u}, \mathbf{u} >_h$$

Pertanto:

$$(\lambda_0 - \overline{\lambda}_0) < \mathbf{u}, \mathbf{u} >_h = 0$$

Dal fatto che $\mathbf{u} \neq \mathbf{0}$, in quanto è un autovettore, segue che $< \mathbf{u}, \mathbf{u} >_h \neq 0$, quindi $\lambda_0 = \overline{\lambda}_0$. □

Nella restante parte di questa appendice vogliamo rivedere i risultati che abbiamo enunciato in questo capitolo per uno spazio vettoriale reale con un prodotto scalare definito positivo, per il caso di uno spazio vettoriale complesso con un prodotto hermitiano definito positivo. Come vedremo tutti i teoremi principali, incluso il Teorema spettrale, hanno enunciati e dimostrazioni analoghe a quelle viste, che pertanto lasceremo per esercizio.

La lettura di questa parte non è necessaria per la comprensione del caso reale, la includiamo per completezza.

In completa analogia con il caso reale possiamo dare le seguenti definizioni, ove nel caso complesso le nozioni di applicazioni lineari o matrici *unitarie* e *hermitiane* sostituiscono le corrispettive nozioni di applicazioni o matrici reali ortogonali e simmetriche.

**Definizione 11.7.5** Sia $V$ uno spazio vettoriale complesso con un prodotto hermitiano definito positivo $< , >_h$. Diciamo che un'applicazione lineare $T : V \longrightarrow V$ è *unitaria* se

$$< T\mathbf{u}, T\mathbf{v} >_h = < \mathbf{u}, \mathbf{v} >_h$$

Diciamo che un'applicazione lineare $T : V \longrightarrow V$ è *hermitiana* se

$$< T\mathbf{u}, \mathbf{v} >_h = < \mathbf{u}, T\mathbf{v} >_h$$

Diciamo che una matrice a coefficienti complessi $A$ è *unitaria* se $A^{-1} = A^*$, ove $A^* = \overline{A^T}$; invece diciamo che è *hermitiana* se $A = A^*$, cioè $A$ coincide con la sua trasposta complessa coniugata. Notiamo che queste due operazioni, cioè trasposizione di $A$ e coniugazione di ogni coefficiente di $A$, commutano tra loro, cioè il risultato è indipendente da quale scegliamo di fare per prima.

È facile verificare che la condizione $A$ hermitiana, corrisponde al fatto che rispetto al prodotto scalare hermitiano in $\mathbb{C}^n$ si abbia:

$$< A\mathbf{u}, \mathbf{v} >_h = < \mathbf{u}, A\mathbf{v} >_h$$

Si noti inoltre che se $A$ è una matrice simmetrica reale abbiamo immediatamente che è anche una matrice hermitiana, infatti è banale che soddisfi la condizione $A = A^*$, in quanto il complesso coniugato di un numero reale è il numero reale stesso.

Dato uno spazio vettoriale $V$ complesso con un prodotto hermitiano $< , >_h$, diciamo che $\mathbf{u}, \mathbf{v} \in V$ sono *perpendicolari (ortogonali) tra loro* se $< \mathbf{u}, \mathbf{v} >_h = 0$.

Se $V$ ha dimensione finita e il prodotto hermitiano è definito positivo, con gli stessi calcoli descritti nel Paragrafo 10.6 si dimostra che $V$ ha una base $\mathcal{B}$ costituita da vettori di norma 1 a due a due ortogonali tra loro, ove la norma di un vettore $\mathbf{u}$ è definita come $\|\mathbf{u}\| = \sqrt{<\mathbf{u}, \mathbf{u}>_{\mathrm{h}}}$.

Possiamo enunciare l'analogo della Proposizione 11.7.6, la cui dimostrazione è uguale a quanto visto per il caso reale.

**Proposizione 11.7.6** *Sia $V$ uno spazio vettoriale complesso di dimensione finita, con un prodotto hermitiano definito positivo $<\,,\,>_{\mathrm{h}}$. Sia $T : V \longrightarrow V$ un'applicazione lineare e sia $\mathcal{B}$ una base ortonormale. Allora:*

1. *$T$ è hermitiana se e solo se la matrice associata a $T$ rispetto alla base $\mathcal{B}$ è una matrice hermitiana.*
2. *$T$ è unitaria se e solo se la matrice associata a $T$ rispetto alla base $\mathcal{B}$ è una matrice unitaria.*

Possiamo inoltre enunciare e dimostrare in modo del tutto analogo al caso reale i seguenti risultati.

**Lemma 11.7.7** *Sia $A$ una matrice hermitiana. Allora:*

1. *$A$ ammette almeno un autovalore reale.*
2. *Se $\mathbf{u}$ è autovettore di $A$ e $\mathbf{w}$ è perpendicolare a $\mathbf{u}$ allora $A\mathbf{w}$ è perpendicolare a $\mathbf{u}$.*
3. *autovettori di $A$ relativi ad autovalori distinti sono perpendicolari tra loro.*

Possiamo infine enunciare il Teorema spettrale per applicazioni lineari hermitiane ed equivalentemente per matrici hermitiane. La dimostrazione è identica a quella per il caso reale.

**Teorema 11.7.8** *Sia $V$ uno spazio vettoriale complesso di dimensione $n$ con un prodotto hermitiano definito positivo. Sia $T : V \longrightarrow V$ un'applicazione lineare hermitiana e sia $A \in \mathrm{M}_n(\mathbb{C})$ la matrice hermitiana associata a $T$ rispetto a una base ortonormale $\mathcal{B}$ di $V$.*

*Allora:*

- *$T$ è diagonalizzabile ed inoltre esiste una base ortonormale $\mathcal{N}$ di $V$ costituita da autovettori di $T$.*
- *$A$ è diagonalizzabile mediante una matrice unitaria, cioè esiste una matrice $P$ unitaria tale che $D = P^{-1}AP$ sia diagonale.*

# 12 Applicazioni: Teorema spettrale e forme quadratiche

In questo capitolo vogliamo studiare alcune conseguenze del Teorema spettrale per quanto riguarda i prodotti scalari e le forme quadratiche ad essi associate.

## 12.1 DIAGONALIZZAZIONE DI PRODOTTI SCALARI

Vogliamo ora riprendere il problema del cambio di base per i prodotti scalari; ci chiediamo se e come sia possibile determinare una base per lo spazio vettoriale rispetto alla quale la matrice associata a un dato prodotto scalare sia il più semplice possibile, e cioè diagonale. Abbiamo già risolto questo problema, grazie al Teorema di Gram-Schmidt (10.6.3), per il caso di un prodotto scalare definito positivo. Ci interessa adesso il caso più generale e inoltre vogliamo imporre alcuni vincoli sulla base da determinare.

Dato uno spazio vettoriale $V$ ed un prodotto scalare $<,>$, fissata una base ordinata $\mathcal{A} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$, possiamo associare a $<,>$ in modo univoco una matrice $C$ i cui coefficienti sono $c_{ij} =< \mathbf{u}_i, \mathbf{u}_j >$, come abbiamo visto nel Capitolo 10.
Se scegliamo una base diversa, $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ avremo che la matrice associata al prodotto scalare cambia secondo la seguente formula (si veda 10.5):

$$C' = I_{\mathcal{B},\mathcal{A}}^T C I_{\mathcal{B},\mathcal{A}} \tag{12.1}$$

ove $I_{\mathcal{B},\mathcal{A}}$ è la matrice del cambio di base tra le basi $\mathcal{B}$ e $\mathcal{A}$.

Supponiamo ora di avere uno spazio vettoriale $V$ con un prodotto scalare $<,>_V$ definito positivo. Come abbiamo visto nell'Osservazione 10.6.4, fissata una base ortonormale, $<,>_V$ è associato alla matrice identità. Dunque, se scriviamo i vettori di $V$ utilizzando le coordinate rispetto alla base ortonormale scelta, possiamo identificare $V$ con $\mathbb{R}^n$ e $<,>_V$ con il prodotto scalare standard.

Ora, consideriamo in $V$ un prodotto scalare arbitrario $<,>$ (non necessariamente definito positivo o non degenere). Vedremo tra poco che, grazie al Teorema Spettrale, è possibile scegliere una base $\mathcal{N}$, ortonormale rispetto a $<,>_V$, tale che la matrice associata al prodotto scalare $<,>$ rispetto alla base

$\mathcal{N}$ sia diagonale. Dunque, $\mathcal{N}$ sarà una base ortogonale (non necessariamente ortonormale) anche rispetto a $<,>$. Ciò ci permetterà di determinare immediatamente alcune proprietà fondamentali del prodotto scalare $<,>$, come per esempio il fatto che sia non degenere oppure definito positivo, semplicemente guardando i segni degli elementi sulla diagonale (che di fatto sono gli autovalori) della matrice associata a $<,>$ rispetto a $\mathcal{N}$.

Iniziamo con un enunciato equivalente del Teorema Spettrale.

**Teorema 12.1.1** *Sia $V$ uno spazio vettoriale di dimensione $n$ con un prodotto scalare $<,>_V$ definito positivo. Sia $<,>$ un'altro prodotto scalare in $V$. Allora esiste una base $\mathcal{N}$ ortonormale per $<,>_V$ e ortogonale per $<,>$.*

**Dimostrazione** – Per il Teorema di Gram-Schmidt (10.6.3) esiste una base $\mathcal{A}$ ortonormale per $<,>_V$. Sia $C$ la matrice associata a $<,>$ rispetto alla base $\mathcal{A}$ e sia $T : V \longrightarrow V$, l'applicazione lineare associata a $C$ rispetto alla base $\mathcal{A}$ in dominio e codominio. Per il Teorema Spettrale, esiste una base $\mathcal{N}$, ortonormale per il prodotto scalare definito positivo $<,>_V$, costituita da autovettori di $T$. Se $P = I_{\mathcal{N},\mathcal{A}}$ è la matrice del cambio di base tra $\mathcal{N}$ e $\mathcal{A}$, abbiamo, sempre per il Teorema Spettrale, che $P$ è ortogonale, cioè $P^{-1} = P^T$. Dunque, essendo $\mathcal{N}$ una base di autovettori, possiamo scrivere:

$$D = P^{-1}CP = P^T CP \tag{12.2}$$

ove $D$ è una matrice diagonale, con gli autovalori di $C$ sulla diagonale.

Dalla formula (12.2), abbiamo che la matrice diagonale $D$ è la matrice associata al prodotto scalare $<,>$ rispetto alla base $\mathcal{N}$, che dunque risulta una base ortogonale per $<,>$. Ciò conclude la dimostrazione. □

♦ **Osservazione 12.1.2** La formula (12.2) ci dice un fatto sorprendente. Dato uno spazio vettoriale $V$ con un prodotto scalare definito positivo e fissata una base ortonormale $\mathcal{A}$, esiste una base ortonormale $\mathcal{N}$ grazie alla quale possiamo scrivere la formula del cambio di base per un'applicazione lineare simmetrica $T$ e la formula del cambio di base per un prodotto scalare $<,>$ nello stesso modo!

Dunque possiamo utilizzare la teoria della diagonalizzazione delle applicazioni lineari che abbiamo studiato nel Capitolo 9, per risolvere il problema della diagonalizzabilità per i prodotti scalari. Due cose dobbiamo tenere presenti:

1. i prodotti scalari sono sempre diagonalizzabili, a differenza di quanto accade per le applicazioni lineari. Ciò avviene perché, fissata una base ordinata, un prodotto scalare è associato a una matrice simmetrica e il Teorema Spettrale ci garantisce la diagonalizzabilità di tali matrici tramite una matrice $P$ ortogonale.
2. La formula (12.2) risolve il problema della diagonalizzabilità di un prodotto scalare *attraverso trasformazioni (cambi di base) ortogonali*. Se rilassiamo questa richiesta è possibile dimostrare in modo più generale il risultato della diagonalizzabilità dei prodotti scalari, attraverso il Teorema di Sylvester, che tuttavia non tratteremo. Vogliamo però sottolineare che cambi di base ortogonali sono particolarmente utili nelle applicazioni, specialmente quelle fisiche.

Dal teorema precedente segue un corollario immediato e tuttavia molto importante per le applicazioni.

**Corollario 12.1.3** *Sia $<,>$ un prodotto scalare in $\mathbb{R}^n$. Allora esiste una base $\mathcal{N}$, ortonormale per il prodotto scalare euclideo, rispetto a cui la matrice associata a $<,>$ è diagonale.*

Vediamo ora come sia possibile applicare i risultati visti sopra per determinare se un prodotto scalare arbitrario è definito positivo e non degenere.

**Proposizione 12.1.4** *Sia $V$ uno spazio vettoriale di dimensione $n$ con un prodotto scalare $<,>$ associato a una matrice diagonale $D$ rispetto a una base data $\mathcal{N}$. Allora:*

1. *$<,>$ è non degenere se e solo se tutti gli elementi sulla diagonale di $D$ sono non nulli;*
2. *$<,>$ è definito positivo se e solo se tutti gli elementi sulla diagonale di $D$ sono positivi.*

**Dimostrazione** – Dati due vettori $\mathbf{u}, \mathbf{v} \in V$, siano $(\mathbf{u})_{\mathcal{N}} = (\alpha_1, \ldots, \alpha_n)^T$, $(\mathbf{v})_{\mathcal{N}} = (\beta_1, \ldots, \beta_n)^T$ le loro coordinate rispetto alla base ordinata $\mathcal{N} = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$. Abbiamo allora:

$$< \mathbf{u}, \mathbf{v} >= (\mathbf{u})_{\mathcal{N}}^T D (\mathbf{v})_{\mathcal{N}} = \lambda_1 \beta_1 \alpha_1 + \cdots + \lambda_n \beta_n \alpha_n \tag{12.3}$$

ove $\lambda_1, \ldots, \lambda_n$ sono, ordinatamente, gli elementi della diagonale principale della matrice diagonale $D$, dunque:

$$< \mathbf{w}_i, \mathbf{w}_i >= \lambda_i \tag{12.4}$$

Vediamo le affermazioni.

(1). Supponiamo che $<,>$ sia non degenere e supponiamo per assurdo che sia $\lambda_i = 0$ per qualche $i$. Allora dalla formula (12.4) segue che il vettore $\mathbf{w}_i$ risulta ortogonale a ogni altro vettore dello spazio $V$, e questo ci dà una contraddizione.

Vediamo ora l'altra implicazione. Supponiamo per assurdo che $<,>$ sia degenere; allora esiste un vettore $\mathbf{u}$ non nullo ortogonale a tutti i vettori dello spazio $V$, in particolare ai vettori di $\mathcal{N}$. Poniamo $(\mathbf{u})_{\mathcal{N}} = (\alpha_1, \ldots, \alpha_n)$; pertanto:

$$< \mathbf{u}, \mathbf{w}_i >= (\mathbf{u})_{\mathcal{N}}^T D (\mathbf{w}_i)_{\mathcal{N}} = \lambda_i \alpha_i = 0$$

per ogni $i$. Poiché per ipotesi $\lambda_i \neq 0$ abbiamo $\alpha_i = 0$ per ogni $i$, che è assurdo.

(2). Se $<,>$ è definito positivo da (12.4) segue subito che $\lambda_i > 0$ per ogni $i$.

Vediamo l'altra implicazione. Per assurdo supponiamo che esista un vettore non nullo $\mathbf{u}$ tale che $< \mathbf{u}, \mathbf{u} > \leq 0$. Allora dalla (12.3) segue:

$$< \mathbf{u}, \mathbf{u} >= \lambda_1 \alpha_1^2 + \cdots + \lambda_n \alpha_n^2 \leq 0$$

Ciò implica che almeno uno dei $\lambda_i$ sia negativo o nullo, ottenendo l'assurdo. □

Vediamo un esempio di applicazione di quanto visto sopra.

**Esempio 12.1.5**

Consideriamo in $\mathbb{R}^2$ il prodotto scalare definito come:

$$< (x_1, y_1), (x_2, y_2) >= -4x_1y_1 + 2x_1y_2 + 2x_2y_1 - 4x_2y_2$$

La matrice ad esso associata rispetto alla base canonica è:

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

Il Teorema Spettrale ci garantisce che tale matrice può essere diagonalizzata attraverso un cambio di base ortogonale. Con un facile calcolo, otteniamo che gli autovalori di $C$ sono $\lambda_1 = -2, \lambda_2 = -6$, e i relativi autospazi sono: $V_{-2} = \langle (1,1) >$, $V_{-6} = \langle (1,-1) >$.

Per il Teorema spettrale abbiamo subito che

$$\begin{pmatrix} -2 & 0 \\ 0 & -6 \end{pmatrix} = P^{-1}CP = P^{t}CP =$$

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

Per la proposizione precedente, poiché al prodotto scalare $<,>$, fissata la base ordinata $\mathcal{N} = ((1/\sqrt{2}, 1/\sqrt{2}), (1/\sqrt{2}, -1/\sqrt{2}))$, è associata la matrice diagonale

$$D = \begin{pmatrix} -2 & 0 \\ 0 & -6 \end{pmatrix}$$

possiamo concludere che $<,>$ sia non degenere, ma non sia definito positivo.

---

Grazie all'esempio precedente possiamo fare un'osservazione immediata, ma molto importante per gli esercizi.

♦ **Osservazione 12.1.6** Sia $<,>$ un prodotto scalare in uno spazio vettoriale $V$ di dimensione $n$, e sia $C$ la matrice associata $C$ rispetto a una qualsiasi base di $V$. Allora:

1. $<,>$ è non degenere se e solo se $C$ non ha autovalori nulli.
2. $<,>$ è definito positivo se e solo se $C$ non ha autovalori negativi o nulli.

Infatti, per il Corollario 12.1.3 esiste una base $\mathcal{N}$ rispetto a cui la matrice associata a $<,>$ sia diagonale. Analizzando la dimostrazione del Teorema 12.1.1 vediamo che tale matrice ha sulla diagonale gli autovalori di $C$. Dunque le affermazioni risultano ovvie per la Proposizione 12.1.4.

Vediamo un esempio.

---

**Esempio 12.1.7**

In $\mathbb{R}^3$ definiamo il prodotto scalare:

$$< (x_1, y_1), (x_2, y_2) >= -2x_1y_1 + 2x_1y_2 + 2x_2y_1 - x_2y_3 - x_3y_2 + x_3y_3$$

Ci chiediamo se tale prodotto sia non degenere e definito positivo.

La matrice associata a $<,>$ rispetto alla base canonica è:

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

Gli autovalori sono:

$$\lambda_1 = 1/2(-3 - \sqrt{13}), \qquad \lambda_2 = 2, \qquad \lambda_3 = 1/2(\sqrt{13} - 3)$$

Per l'osservazione precedente possiamo concludere *senza fare altri calcoli* che il prodotto dato è non degenere, ma non è definito positivo, in quanto tutti gli autovalori sono diversi da zero, ma l'autovalore $\lambda_1$ è negativo.

## 12.2 FORME QUADRATICHE

In questa sezione vogliamo discutere le forme quadratiche e come sia possibile utilizzare le informazioni che conosciamo sul prodotto scalare per diagonalizzare una forma quadratica e, nel caso di due variabili, utilizzare le informazioni per disegnare curve nel piano.

**Definizione 12.2.1** Sia $V$ uno spazio vettoriale reale e sia $<,>$ un prodotto scalare in $V$. Definiamo la *forma quadratica reale* $q$ associata a $<,>$ come la funzione $q : V \longrightarrow \mathbb{R}$ tale che $q(\mathbf{v}) = <\mathbf{v},\mathbf{v}>$.

Per esempio in $\mathbb{R}^n$, dato il prodotto euclideo, la forma quadratica ad esso associata è la funzione che associa a un vettore la sua norma al quadrato: $q(\mathbf{v}) = \|\mathbf{v}\|^2 = \alpha_1^2 + \cdots + \alpha_n^2$, per ogni $\mathbf{v} = (\alpha_1, \ldots, \alpha_n)$.

♦ **Osservazione 12.2.2** Se $q$ è una forma quadratica, allora $q$ determina univocamente il prodotto scalare che la definisce. Infatti tale prodotto su due vettori arbitrari è dato da:

$$<\mathbf{u},\mathbf{v}> = (1/2)[q(\mathbf{u}+\mathbf{v}) - q(\mathbf{u}) - q(\mathbf{v})] \qquad (12.5)$$

Infatti:

$$\begin{aligned}(1/2)[q(\mathbf{u}+\mathbf{v}) - q(\mathbf{u}) - q(\mathbf{v})] &= (1/2)[<\mathbf{u}+\mathbf{v},\mathbf{u}+\mathbf{v}> - <\mathbf{u},\mathbf{u}> - <\mathbf{v},\mathbf{v}>] \\ &= (1/2)[2<\mathbf{u},\mathbf{v}>]\end{aligned}$$

**Definizione 12.2.3** Data una forma quadratica $q$ su uno spazio vettoriale $V$, diremo che $q$ è *non degenere* o *definita positiva* se lo è il prodotto scalare ad essa associato.

Data una forma quadratica $q$, fissata una base $\mathcal{B}$ per $V$, possiamo dunque scrivere:

$$q(\mathbf{v}) = (\mathbf{v})_{\mathcal{B}}^T C (\mathbf{v})_{\mathcal{B}}$$

ove $C$ è la matrice associata al prodotto scalare corrispondente a $q$, rispetto alla base $\mathcal{B}$, e $(\mathbf{v})_{\mathcal{B}}$ la colonna delle coordinate di $\mathbf{v}$, sempre rispetto alla base $\mathcal{B}$.

♦ **Osservazione 12.2.4** In $\mathbb{R}^n$ consideriamo la seguente funzione $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ data da:

$$\begin{aligned}q(x_1,\ldots,x_n) &= a_{11}x_1^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + \ldots a_{1n}x_1x_n + \\ &\quad + a_{22}x_2^2 + a_{23}x_2x_3 + \cdots + a_{2n}x_2x_n + \cdots + a_{nn}x_n^2\end{aligned}$$

Costruiamo la matrice simmetrica $C$ nel seguente modo:

$$C = \begin{pmatrix} a_{11} & \frac{a_{12}}{2} & \frac{a_{13}}{2} & \ldots & \frac{a_{1n}}{2} \\ \frac{a_{12}}{2} & a_{22} & \frac{a_{23}}{2} & \ldots & \frac{a_{2n}}{2} \\ \ldots \\ \frac{a_{1n}}{2} & \frac{a_{2n}}{2} & \frac{a_{3n}}{2} & \ldots & a_{nn} \end{pmatrix} \qquad (12.6)$$

È facile verificare che

$$q(x_1, \dots, x_n) = (x_1, \dots, x_n) C \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Dunque la funzione $q$ rappresenta una forma quadratica in $\mathbb{R}^n$ ed è immediato verificare usando (12.6) che ogni forma quadratica in $\mathbb{R}^n$ è di questo tipo.

Vediamo un esempio.

---

**Esempio 12.2.5**

Consideriamo $q : \mathbb{R}^3 \longrightarrow \mathbb{R}$, definita da $q(x, y, z) = x^2 + 2xy + 3zy - 2z^2$. Abbiamo che $q$ è una forma quadratica, e la matrice ad essa associata è:

$$q(x, y, z) = (x \; y \; z) \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 3/2 \\ 0 & 3/2 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Lasciamo al lettore la facile verifica dell'uguaglianza che abbiamo scritto.

---

Il Teorema Spettrale e la sua versione per i prodotti scalari data dal Teorema 12.1.1 hanno una conseguenza immediata per quanto riguarda le forme quadratiche.

**Corollario 12.2.6** **(Teorema degli Assi Principali).** *Sia $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ una forma quadratica associata alla matrice $C$ rispetto alla base canonica di $\mathbb{R}^n$. Allora esiste sempre una base ortonormale $\mathcal{N}$ di $\mathbb{R}^n$, costituita da autovettori di $C$, rispetto a cui la matrice associata a $q$ assume la forma diagonale. Possiamo pertanto scrivere:*

$$q(x_1, \dots, x_n) = \lambda_1 x_1^2 + \cdots + \lambda_n x_n^2$$

*ove $(x_1, \dots, x_n)$ sono le coordinate rispetto alla base $\mathcal{N}$ e $\lambda_1, \dots, \lambda_n$ sono gli autovalori della matrice $C$.*

Concludiamo la sezione con una definizione, fondamentale per la classificazione delle forme quadratiche reali, argomento che tuttavia non svolgeremo in questo testo[1].

**Definizione 12.2.7** Sia $q$ una forma quadratica su $\mathbb{R}^n$. Definiamo *segnatura* di $q$ la coppia $(r, s)$, ove $r$ ed $s$ sono il numero di autovalori positivi, rispettivamente negativi, della matrice $C$ associata a $q$ rispetto alla base canonica, ciascuno contato con la propria molteplicità.

Notiamo che in questa definizione, anziché la base canonica, possiamo scegliere qualunque base ortonormale per determinare una matrice associata a $q$, in quanto sappiamo che tutte le matrici associate a $q$ in questo modo avranno gli stessi autovalori.

---

[1] Il Teorema di Sylvester, che non trattiamo, afferma infatti che due matrici simmetriche rappresentano la stessa forma quadratica, o equivalentemente lo stesso prodotto scalare, rispetto a basi diverse se e solo se hanno la stessa segnatura.

## 12.3 FORME QUADRATICHE E CURVE NEL PIANO

Come applicazione dei risultati sulle forme quadratiche vogliamo descrivere il luogo dei punti del piano che soddisfano l'equazione

$$q(x,y) = c \tag{12.7}$$

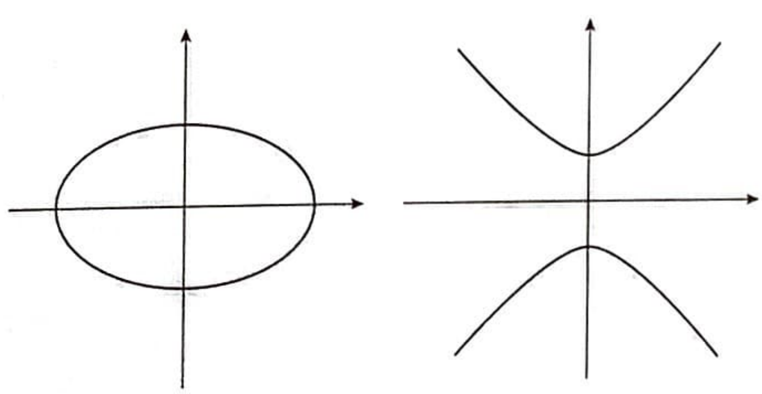
ove $q$ è una forma quadratica non degenere e $c \in \mathbb{R}$ è una costante positiva.

Cominciamo con il considerare il caso particolare

$$q(x,y) = \lambda_1 x^2 + \lambda_2 y^2 \tag{12.8}$$

con $\lambda_1$, $\lambda_2 \neq 0$. Vediamo subito che se $\lambda_1$, $\lambda_2 < 0$ nessun punto del piano ha coordinate che soddisfino (12.7). Suddividiamo gli altri casi a seconda del segno di $\lambda_1$ e $\lambda_2$ ottenendo la seguente classificazione:

- $\lambda_1, \lambda_2 > 0$ ellisse;
- $\lambda_1 < 0$, $\lambda_2 > 0$ (oppure $\lambda_1 > 0$, $\lambda_2 < 0$) iperbole;



Nel caso in cui la matrice associata a $q$ rispetto alla base canonica abbia due autovalori non nulli e non entrambi negativi, è sempre possibile ricondurci a queste due figure geometriche grazie al Teorema degli Assi Principali (12.2.6). Diremo pertanto che una forma quadratica è in *forma canonica*, se assume l'espressione (12.8). Vediamo un esempio.

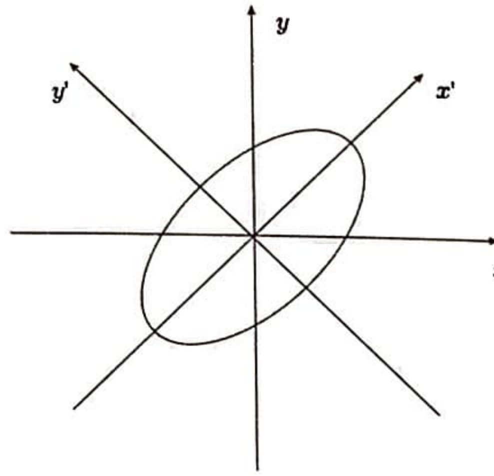
**Esempio 12.3.1**

Vogliamo disegnare il luogo dei punti del piano descritti dall'equazione: $5x^2 - 4xy + 5y^2 = 48$.

La matrice associata alla forma quadratica $q(x,y) = 5x^2 - 4xy + 5y^2$ è:

$$A = \begin{pmatrix} 5 & -2 \\ -2 & 5 \end{pmatrix}$$

Gli autovalori di $A$ sono 3 e 7, i corrispondenti autovettori di lunghezza unitaria $\mathbf{u}_1 = (1/\sqrt{2}, 1/\sqrt{2})$, $\mathbf{u}_2 = (-1/\sqrt{2}, 1/\sqrt{2})$. Utilizzando le coordinate rispetto alla base $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2\}$ abbiamo $q(x', y') = 3(x')^2 + 7(y')^2$, pertanto $q(x,y) = 48$ è un'ellisse che possiamo subito disegnare:




Vediamo ora un altro esempio relativo all'iperbole.

**Esempio 12.3.2**

Vogliamo disegnare il luogo dei punti del piano le cui coordinate soddisfano l'equazione: $x^2 - 8xy - 5y^2 = 16$. La matrice associata alla forma quadratica $q(x, y) = x^2 - 8xy - 5y^2$ è:

$$A = \begin{pmatrix} 1 & -4 \\ -4 & -5 \end{pmatrix}$$

Gli autovalori di $A$ sono $-7$ e $3$, i corrispondenti autovettori di lunghezza unitaria $\mathbf{u}_1 = (1/\sqrt{5}, 2/\sqrt{5})$, $\mathbf{u}_2 = (-2/\sqrt{5}, 1/\sqrt{5})$. Utilizzando le coordinate rispetto alla base $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2\}$ abbiamo $q(x', y') = -7(x')^2 + 3(y')^2$, pertanto $q(x, y) = 16$ è un'iperbole, che possiamo subito disegnare, analogamente a quanto abbiamo fatto precedentemente.

## 12.4 ESERCIZI SVOLTI

### Esercizio 12.4.1

*Si consideri la forma quadratica in $\mathbb{R}^3$ data da:*

$$q(x, y, z) = x^2 + 2xz - y^2 - 2yz$$

*Si dica se è non degenere, definita positiva e se ne dia la segnatura. Si scriva il prodotto scalare ad essa associato rispetto alla base canonica.*

*Si determini inoltre una base rispetto a cui tale prodotto scalare è associato a una matrice diagonale.* ■

**Svolgimento**

Per prima cosa scriviamo la matrice associata alla forma quadratica data, fissata la base canonica:

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

Gli autovalori sono:

$$\lambda_1 = -\sqrt{3}, \qquad \lambda_2 = \sqrt{3}, \qquad \lambda_3 = 0$$

La forma quadratica data è degenere in quanto uno degli autovalori della matrice associata è uguale a zero. Non è definita positiva in quanto è degenere. La segnatura è $(1,1)$.

Il prodotto scalare associato a $q$ rispetto alla base canonica è dato da:

$$< (x,y,z),(x',y',z') >= (x,y,z)\begin{pmatrix}1 & 0 & 1\\ 0 & -1 & -1\\ 1 & -1 & 0\end{pmatrix}\begin{pmatrix}x'\\ y'\\ z'\end{pmatrix}$$

Una base ortonormale $\mathcal{N}$ rispetto a cui la matrice $C$ è diagonale è costituita dagli autovettori di $C$ opportunamente normalizzati. Infatti calcolati gli autovettori:

$$\mathbf{v}_1 = (-(2-\sqrt{3})/(-1+\sqrt{3}), 1/(-1+\sqrt{3}), 1)$$

$$\mathbf{v}_2 = (-(-2-\sqrt{3})/(1+\sqrt{3}), -1/(1+\sqrt{3}), 1)$$

$$\mathbf{v}_3 = (-1,-1,1)$$

abbiamo $\mathcal{N} = (\mathbf{v}_1/\|\mathbf{v}_1\|, \mathbf{v}_2/\|\mathbf{v}_2\|, \mathbf{v}_3/\|\mathbf{v}_3\|)$.

### Esercizio 12.4.2

*Si consideri nel piano la curva costituita da punti che soddisfano l'equazione:*

$$3x^2 - 2xy - y^2 = 1$$

*Si determini di che curva si tratta e se ne dia un disegno di massima.*

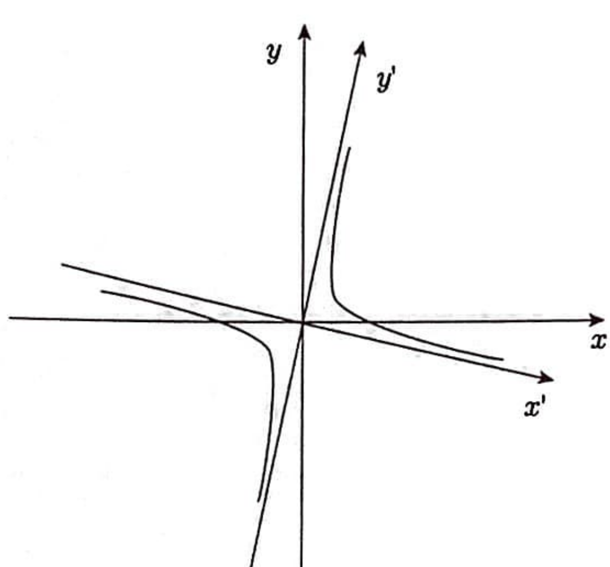
**Svolgimento**

La matrice associata alla forma quadratica $q(x,y) = 3x^2 - 2xy - y^2$ rispetto alla base canonica è:

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

Gli autovalori sono:

$$\lambda_1 = 1+\sqrt{5}, \qquad \lambda_2 = 1-\sqrt{5}$$



Vediamo subito, senza bisogno di ulteriori calcoli che si tratta di un'iperbole, la cui forma canonica è data da:

$$q(x', y') = (1 + \sqrt{5})(x')^2 + (1 - \sqrt{5})(y')^2$$

Tuttavia, se vogliamo fare un disegno, anche di massima è necessario calcolare gli autovettori, in quanto le loro direzioni sono quelle degli assi $x'$ e $y'$ del nuovo sistema di coordinate. Notiamo che, a differenza del caso in cui dobbiamo calcolare il cambio di base per una diagonalizzazione di un prodotto scalare, o equivalentemente di una matrice simmetrica, in questo caso non è necessario normalizzare gli autovettori, cioè dividerli per la norma, in quanto ci interessa soltanto la direzione degli assi e non il cambio di base.

Calcoliamo gli autovalori:

$$\mathbf{v}_1 = (-2 - \sqrt{5}, 1), \qquad \mathbf{v}_2 = (-2 + \sqrt{5}, 1)$$

Possiamo pertanto disegnare l'iperbole. È utile calcolare le intersezioni con gli assi $x'$ e $y'$: per $y' = 0$, $x' = \pm 1/(1 + \sqrt{5})$. ∎

## 12.5 ESERCIZI PROPOSTI

### Esercizio 12.5.1

*Data la forma quadratica* $q(x_1, x_2) = 5x_1^2 - 4x_1x_2 + 5x_2^2$.

a) *Si scriva la matrice ad essa associata rispetto alla base canonica.*

b) *Si scriva la forma quadratica in forma canonica* $q(x_1, x_2) = ax_1^2 + bx_2^2$ *per opportuni* $a$ *e* $b$.

c) *Si disegni nel piano la curva* $q(x_1, x_2) = 48$.

### Esercizio 12.5.2

*Data la forma quadratica* $q(x_1, x_2) = 3x_1^2 + 2x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3$.

a) *Si scriva la matrice* $A$ *ad essa associata rispetto alla base canonica.*

b) *Si dica se* $q$ *è definita positiva.*

c) *Si trovi (se possibile) una matrice* $P$ *tale che* $D = P^{-1}AP$ *sia diagonale.*

d) *Si scriva la forma quadratica* $q_1$ *associata a* $D$ *e si stabilisca la relazione sussiste tra* $q_1$ *e* $q$.

### Esercizio 12.5.3

*Data la matrice:*

$$C = \begin{pmatrix} 4 & 2 & 3 \\ 2 & 0 & 2 \\ 3 & 2 & 4 \end{pmatrix}$$

a) *Si scrivano il prodotto scalare e la forma quadratica ad essa associati.*

b) *Si determini se il prodotto scalare del punto* (a) *è definito positivo e/o non degenere.*

c) *Si calcoli la segnatura della forma quadratica del punto* (a).

d) *Si determini una base rispetto alla quale il prodotto scalare del punto* (a) *è associato a una matrice diagonale.*

[Suggerimento: 8 e $-1$ sono autovalori di $C$.]

### Esercizio 12.5.4

*Si consideri la forma quadratica:* $q(x, y) = x^2 + 5xy$.

a) *Si scriva il prodotto scalare ad essa associato e si determini se è non degenere, definito positivo. Si calcoli inoltre la segnatura di* $q$.

b) *Dato il luogo dei punti del piano descritto dall'equazione* $x^2 + 5xy = 1$, *di che curva si tratta? Si motivi la risposta.*

### Esercizio 12.5.5

*Si disegni la curva di equazione* $-7x_1^2 - 12x_1x_2 + 2x_2^2 = 16$.

### Esercizio 12.5.6

*Data la curva descritta dall'equazione*

$$5x^2 + 5y^2 - 6xy - 16 = 0$$

*se ne trovi la forma canonica e se ne dia un disegno di massima.*

### Esercizio 12.5.7

*Data l'equazione:*

$$x^2 - 4xy + y^2 = 4$$

a) *Si dica che luogo dei punti del piano individua.*

b) *Si determini la sua forma canonica.*

c) *Se ne tracci un disegno di massima.*

### Esercizio 12.5.8

*Data la curva descritta dall'equazione:* $x^2 + y^2 - 16xy = 1$

a) *Se ne trovi la forma canonica e se ne dia un disegno di massima.*

b) *Si scriva il prodotto scalare ad essa associato e si dica se è non degenere o definito positivo.*

# 13 Rette e piani in $\mathbb{R}^3$

In questo capitolo vogliamo introdurre in modo elementare la geometria tridimensionale e ci concentreremo particolarmente sullo studio delle rette e dei piani nello spazio $\mathbb{R}^3$.

Avremo quindi bisogno di introdurre il concetto di *prodotto scalare* e *prodotto vettoriale* in $\mathbb{R}^3$. Poiché questa appendice è indipendente dal resto del nostro testo, tali nozioni saranno introdotte senza alcun riferimento alle definizioni presenti nei Capitoli 10, 11, 12. Il lettore è avvisato che il prodotto scalare definito qui di sotto rappresenta soltanto un esempio della nozione più generale di prodotto scalare così come è definita nel Capitolo 11.

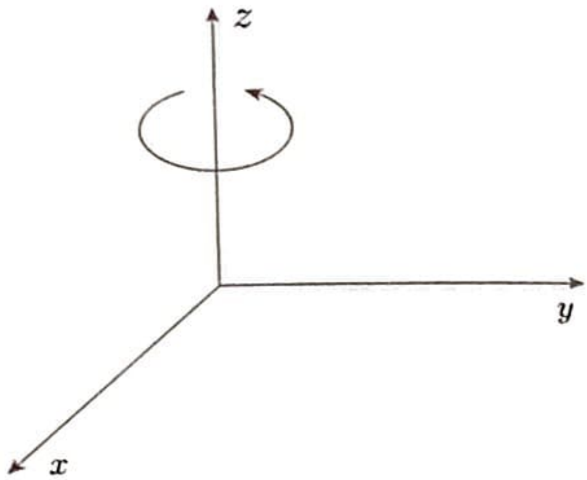
## 13.1 PUNTI E VETTORI IN $\mathbb{R}^3$

Consideriamo $\mathbb{R}^3$, l'insieme delle terne ordinate di numeri reali:

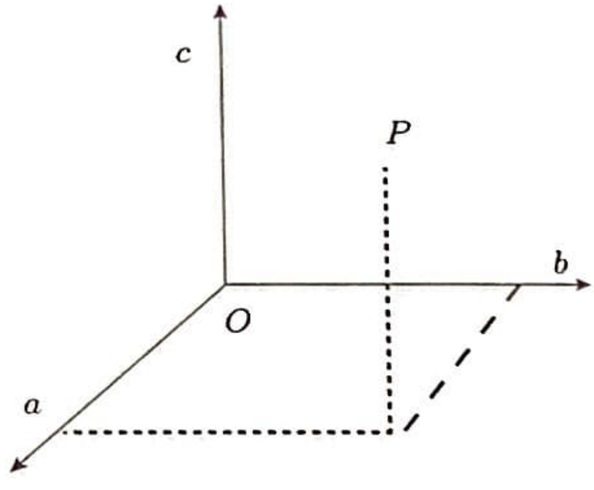
$$\mathbb{R}^3 = \{(x, y, z) \mid x, y, z \in \mathbb{R}\}$$

Possiamo rappresentare e visualizzare gli elementi di $\mathbb{R}^3$ come punti dello spazio in cui:

- Fissiamo prima un punto $O$ detto *origine* corrispondente all'elemento $(0, 0, 0)$.
- Scegliamo tre rette attraverso $O$ perpendicolari tra loro, dette *assi coordinati*, e denominate rispettivamente *asse* $x$, *asse* $y$ e *asse* $z$. Di solito pensiamo agli assi $x$ e $y$ come *orizzontali* e all'asse $z$ come *verticale* (si veda la figura successiva).
- Fissiamo una orientazione degli assi secondo la *regola della mano destra* (si veda la figura successiva). Scegliamo prima una direzione, denominata *positiva*, per l'asse $x$ e una per l'asse $y$. Per l'asse $z$ la direzione positiva è individuata nel seguente modo: avvolgendo le dita della mano destra (indice, medio, anulare, mignolo) attorno all'asse $z$ dall'asse $x$ in direzione positiva, all'asse $y$ in direzione positiva, il pollice punta nella direzione che definiamo positiva per l'asse $z$.




Possiamo dunque associare a un punto $P$ nello spazio la terna ordinata di numeri reali $(a, b, c)$ che chiamiamo *coordinate* di $P$ (si veda la figura sottostante). D'ora in avanti identificheremo $\mathbb{R}^3$ con i punti dello spazio attraverso tale rappresentazione.




Dati due punti $P = (a, b, c)$ e $Q = (a', b', c') \in \mathbb{R}^3$, possiamo definire la distanza tra loro come:

$$d(P, Q) = \sqrt{(a-a')^2 + (b-b')^2 + (c-c')^2}$$

Lasciamo al lettore il facile esercizio di dimostrare che effettivamente $d(P, Q)$ rappresenta la lunghezza del segmento che congiunge $P$ con $Q$.

Veniamo ora a definire la nozione di *vettore*, estremamente importante per la nostra trattazione. In questa appendice, a differenza del resto del testo, useremo le parentesi tonde per i punti di $\mathbb{R}^3$ e le parentesi quadre per i vettori, in quanto è necessario rimarcare la differenza.

Dato un punto $P = (a, b, c) \in \mathbb{R}^3$ definiamo il *vettore posizione* di $P$ come:

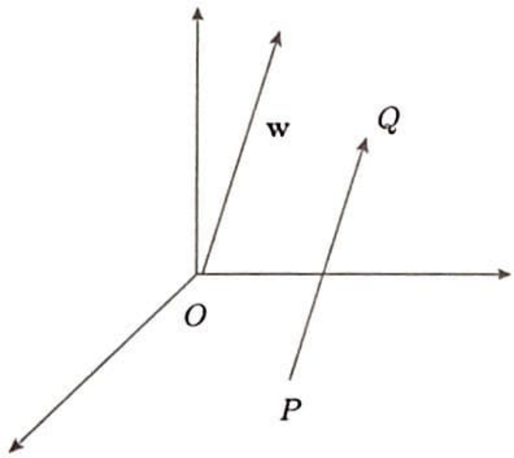
$$\mathbf{v} = \overrightarrow{OP} = [a, b, c]$$

Rappresentiamo il vettore posizione come una freccia uscente dall'origine e con la punta in $P$.

Dati due punti $P = (a, b, c), Q = (a', b', c') \in \mathbb{R}^3$ definiamo il vettore $\overrightarrow{PQ}$ come:

$$\mathbf{w} = \overrightarrow{PQ} = [(a - a'), (b - b'), (c - c')]$$

Se pensiamo a una freccia uscente da $P$ e con la punta in $Q$, come si vede nella figura seguente, possiamo rappresentare il vettore $\mathbf{w}$ come tale freccia traslata e uscente dall'origine.




Dati due vettori $\mathbf{u} = [u_1, u_2, u_3]$, $\mathbf{v} = [v_1, v_2, v_3]$ possiamo definire la loro somma nel seguente modo:

$$\mathbf{u} + \mathbf{v} = [u_1 + v_1, u_2 + v_2, u_3 + v_3]$$

Analogamente possiamo definire la moltiplicazione per uno scalare e cioè per un numero reale $\lambda$:

$$\lambda\mathbf{u} = [\lambda u_1, \lambda u_2, \lambda u_3]$$

Le operazioni di somma di vettori e moltiplicazione per uno scalare rendono l'insieme dei vettori dello spazio uno spazio vettoriale, cioè valgono le 8 proprietà della Definizione 2.3.1 del Capitolo 2. Lasciamo la facile verifica al lettore.

## 13.2 PRODOTTO SCALARE E PRODOTTO VETTORIALE

Vogliamo ora definire il prodotto scalare tra due vettori. Siano $\mathbf{u} = [u_1, u_2, u_3]$, $\mathbf{v} = [v_1, v_2, v_3]$ due vettori. Definiamo il loro *prodotto scalare* come:

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3$$

Il prodotto scalare gode delle seguenti proprietà, che lasciamo come facile esercizio al lettore. Per ogni vettore $\mathbf{u}, \mathbf{v}, \mathbf{w}$ e per ogni scalare $\lambda$ valgono:

- commutatività
$$\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$$
- compatibilità con la moltiplicazione per uno scalare
$$(\lambda\mathbf{u}) \cdot \mathbf{v} = \lambda(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (\lambda\mathbf{v})$$
- distributività
$$\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$$

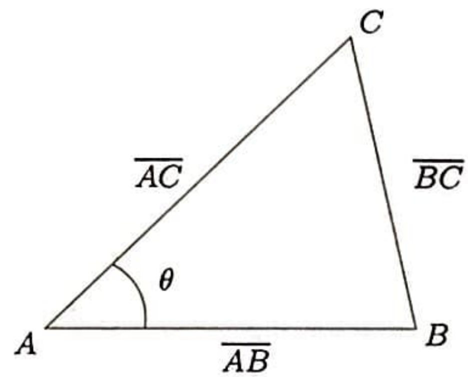
Si noti che $\sqrt{\mathbf{u}\cdot\mathbf{u}} = \sqrt{u_1^2+u_2^2+u_3^2}$ rappresenta la distanza del punto $P = (u_1, u_2, u_3)$ dall'origine e pertanto rappresenta la lunghezza del segmento $OP$. Definiamo pertanto come *lunghezza* del vettore $\mathbf{u}$ la quantità $\sqrt{\mathbf{u}\cdot\mathbf{u}}$ che denotiamo anche come $\|\mathbf{u}\|$.

Prima di proseguire, ricordiamo al lettore un risultato di geometria euclidea elementare, il *Teorema del coseno*, che permette di trovare la lunghezza di un lato di un triangolo di vertici $A$, $B$, $C$, conoscendo le lunghezze degli altri due lati e l'ampiezza dell'angolo $\theta$ tra essi (si veda la figura seguente). Tale teorema afferma che:

$$\overline{BC}^2 = \overline{AC}^2 + \overline{AB}^2 - \overline{AC}\cdot\overline{AB}\cos\theta,$$

ove $\theta$ è l'angolo di vertice $C$.

Grazie a questo teorema possiamo dimostrare un risultato che si rivelerà fondamentale per la nostra trattazione.




**Teorema 13.2.1** *Siano* $\mathbf{u}$ *e* $\mathbf{v}$ *due vettori non nulli e sia* $\theta$ *l'angolo avente per lati le semirette individuate dai due vettori. Allora*

$$\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\ \|\mathbf{v}\|\cos\theta$$

**Dimostrazione** – Con un facile calcolo, abbiamo che:

$$\|\mathbf{u}-\mathbf{v}\|^2 = (\mathbf{u}-\mathbf{v})\cdot(\mathbf{u}-\mathbf{v}) = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\mathbf{u}\cdot\mathbf{v} \tag{13.1}$$

Per il Teorema del coseno, se consideriamo il triangolo con un vertice nell'origine e lati di lunghezza $\|\mathbf{u}\|$, $\|\mathbf{v}\|$ e $\|\mathbf{u}-\mathbf{v}\|$, abbiamo subito che:

$$\|\mathbf{u}-\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - \|\mathbf{u}\|\ \|\mathbf{v}\|^2\cos(\theta) \tag{13.2}$$

Le uguaglianze (13.1) e (13.2) ci danno subito quanto cercato. □

Dal teorema precedente possiamo dedurre immediatamente il seguente corollario, che stabilisce quando due vettori sono perpendicolari e cioè quando l'angolo compreso tra di essi è di $\pi/2$. Tale risultato si rivelerà molto utile per gli esercizi.

**Corollario 13.2.2** *Due vettori* $\mathbf{u}$ *e* $\mathbf{v}$ *sono perpendicolari tra loro se e solo se* $\mathbf{u}\cdot\mathbf{v} = 0$.

Passiamo ora a esaminare un altro prodotto estremamente importante per la nostra trattazione: il *prodotto vettoriale*. Siano $\mathbf{u} = [u_1, u_2, u_3]$, $\mathbf{v} = [v_1, v_2, v_3]$ due vettori. Definiamo il loro prodotto vettoriale come:

$$\mathbf{u} \times \mathbf{v} = (u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1)$$

Il prodotto vettoriale gode delle seguenti proprietà che lasciamo per esercizio. Per ogni vettore $\mathbf{u}, \mathbf{v}, \mathbf{w}$ e per ogni scalare $\lambda$ valgono:

- il prodotto di un vettore per se stesso è il vettore nullo
$$\mathbf{u} \times \mathbf{u} = \mathbf{0}$$
- anticommutatività
$$\mathbf{u} \times \mathbf{v} = -(\mathbf{v} \times \mathbf{u})$$
- distributività rispetto all'addizione
$$\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) + (\mathbf{u} \times \mathbf{w})$$
- compatibilità con la moltiplicazione per uno scalare
$$(\lambda \mathbf{u}) \times \mathbf{v} = \mathbf{u} \times (\lambda \mathbf{v}) = \lambda (\mathbf{u} \times \mathbf{v})$$

Concludiamo la sezione con un risultato di grande importanza per gli esercizi, che discende immediatamente dal Corollario 13.2.2

**Proposizione 13.2.3** *Il prodotto vettoriale tra due vettori* $\mathbf{u}$ *e* $\mathbf{v}$ *è perpendicolare sia a* $\mathbf{u}$ *sia a* $\mathbf{v}$.

## 13.3 RETTE IN $\mathbb{R}^3$

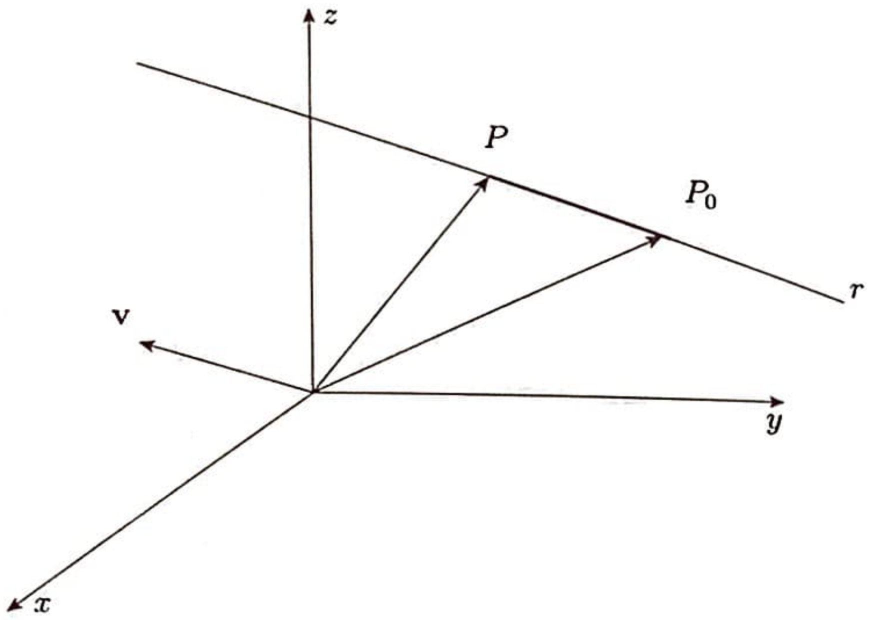
Vogliamo descrivere, tramite equazioni, il luogo geometrico dei punti corrispondenti a una retta $r$ passante per un punto $P_0 = (x_0, y_0, z_0)$ e con direzione individuata dal vettore $\mathbf{v} = [v_1, v_2, v_3]$ nello spazio $\mathbb{R}^3$. In forma concisa, aiutandoci con un disegno, possiamo subito scrivere l'equazione che lega i vettori posizione di $P = (x, y, z)$, punto generico della retta $r$, e $P_0$:

$$\overrightarrow{OP} = \overrightarrow{OP_0} + t[v_1, v_2, v_3] \tag{13.3}$$

Pertanto:

$$[x, y, z] = [x_0, y_0, z_0] + t[v_1, v_2, v_3] \tag{13.4}$$

L'equazione (13.4) si dice *equazione vettoriale* della retta $r$.




Più estesamente, la retta $r$ è l'insieme dei punti con coordinate $(x, y, z) \in \mathbb{R}^3$ espresse come:

$$\begin{cases} x = x_0 + tv_1 \\ y = y_0 + tv_2 \\ z = z_0 + tv_3 \end{cases} \tag{13.5}$$

al variare del parametro $t \in \mathbb{R}$. Le equazioni in (13.5) si dicono *equazioni parametriche* della retta $r$. Il vettore $\mathbf{v}$ si dice *vettore direttore o direzione*. Notiamo che il vettore $\mathbf{v}$ individua la *direzione* di $r$, ma in realtà tutti i vettori multipli non nulli di $\mathbf{v}$ possono essere utilizzati equivalentemente per definire la stessa retta.

Vediamo un esempio concreto.

**Esempio 13.3.1**

Si scrivano equazioni parametriche della retta $r$ per il punto $P_0 = (1, 0 - 1)$ con vettore direzione $\mathbf{v} = [2, 1, -1]$ e si dica se $r$ è parallela alla retta $r'$ data da equazioni parametriche $[x, y, z] = [2, 0, -3] + t'[-4, -2, 2]$. Sostituendo nella formula (13.5) abbiamo subito che equazioni parametriche della retta sono:

$$\begin{cases} x = 1 + 2t \\ y = t \\ z = -1 - t \end{cases} \tag{13.6}$$

Le due rette $r$ ed $r'$ hanno la stessa direzione, in quanto il vettore direttore di $r$, $\mathbf{v} = [2, 1, -1]$ è un multiplo del vettore direttore di $r'$, $\mathbf{v}' = [-4, -2, 2]$. Quindi le due rette sono parallele o coincidenti. Per stabilire la loro posizione reciproca è sufficiente verificare se il punto $P_0$ appartiene o meno alla retta $r'$, cioè se esiste un valore del parametro $t'$ tale che $[1, 0, -1] = [2, 0, -3] + t'[-4, -2, 2]$. Lasciamo per esercizio la facile verifica che tale valore non esiste. Dunque le due rette date sono parallele e non coincidenti.

Vediamo un altro esempio.

**Esempio 13.3.2**

Si scrivano equazioni parametriche della retta $\hat{r}$ per i punti $P = (3, 1, -2)$ e $Q = (5, 2, -3)$. Un vettore direttore di $\hat{r}$ si ottiene facendo ordinatamente la differenza tra le coordinate di $Q$ e quelle di $P$: $\mathbf{v} = [2, 1, -1]$. Dunque $\hat{r}$ è data da:

$$\begin{cases} x = 3 + 2\hat{t} \\ y = 1 + \hat{t} \\ z = -2 - \hat{t} \end{cases} \tag{13.7}$$

ove abbiamo scelto $P_0 = P$ da sostituire nella formula (13.5) (avremmo benissimo potuto scegliere $Q$).

Ci chiediamo ora se la retta $\hat{r}$ sia parallela o coincidente con la retta $r$ dell'esempio precedente, dato che hanno lo stesso vettore direttore. Un rapido calcolo mostra che $Q$ appartiene a $r$ e dunque le due rette sono coincidenti.

Questi esempi mostrano che una forma parametrica di una retta data non è unica: possiamo infatti cambiare il punto $P_0 = (x_0, y_0, z_0)$ usato nella rappresentazione (13.5), scegliendolo arbitrariamente tra tutti gli infiniti punti della retta, oppure possiamo moltiplicare il parametro $t$ per una costante arbitraria non nulla: in entrambi i casi la retta descritta non cambia, anche se le sue equazioni parametriche possono assumere un aspetto diverso.

Vediamo ora un modo equivalente di descrivere i punti di una retta in $\mathbb{R}^3$ senza fare ricorso a un parametro. Date equazioni parametriche di una retta $r$, è sempre possibile ricavare il parametro $t$ da una delle equazioni e, sostituendolo nelle altre due, ottenere un sistema lineare di due equazioni nelle sole incognite $x$, $y$, $z$. Tali equazioni si dicono *equazioni cartesiane* della retta $r$ data. È ovvio che se le coordinate di un punto verificano equazioni parametriche della retta $r$ allora esse verificano anche le equazioni cartesiane, il viceversa sarà chiaro alla fine della prossima sezione. Infatti, vedremo che le due equazioni lineari del sistema che otteniamo rappresentano due piani contenenti entrambi la retta $r$, quindi le soluzioni del sistema lineare sono i punti che giacciono nell'intersezione di due piani, cioè i punti di una retta, e tale retta deve essere proprio la retta $r$. Vediamo un esempio.

**Esempio 13.3.3**

Vogliamo scrivere in forma cartesiana la retta $r$ dell'Esempio 13.3.1. In questo caso è molto semplice; poiché $t = y$, basta sostituire $t$ al posto di $y$ direttamente nelle altre equazioni:

$$\begin{cases} x = 1 + 2y \\ z = -1 - y \end{cases} \tag{13.8}$$

Dunque $x - 2y - 1 = 0$ e $y + z + 1 = 0$ sono le equazioni cartesiane della retta $r$. Possiamo quindi pensare ai punti della retta anche come l'insieme delle soluzioni del sistema lineare (13.8).

Nella prossima sezione vedremo come l'esempio precedente possa essere reinterpretato in modo da vedere una retta come intersezione di due piani in $\mathbb{R}^3$.

## 13.4 PIANI IN $\mathbb{R}^3$

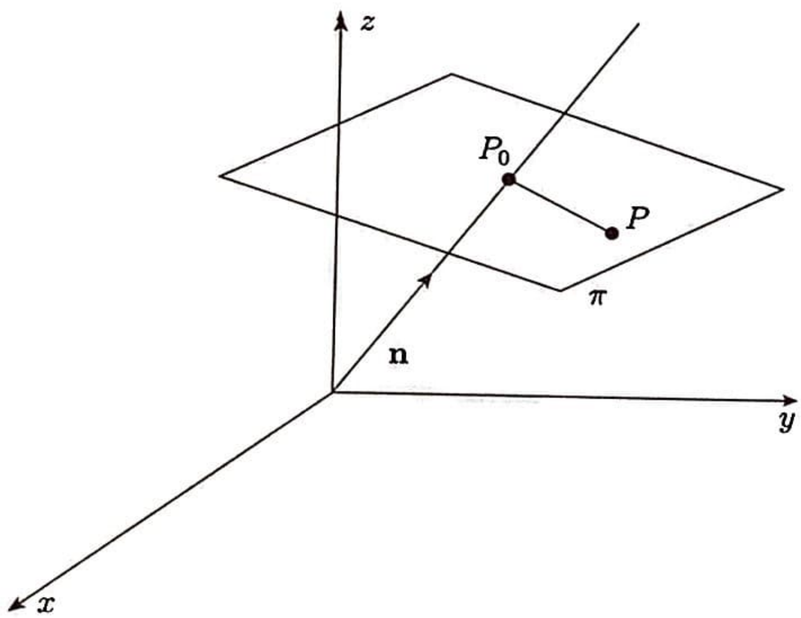
Vogliamo ora determinare l'equazione che descrive i punti di un piano perpendicolare a una retta data con direzione $\mathbf{n} = [a, b, c]$ e passante per un punto fissato $P_0 = (x_0, y_0, z_0)$. Per il Corollario 13.2.2 abbiamo che due vettori sono perpendicolari se e solo se il loro prodotto scalare è nullo. Dunque l'insieme dei punti del piano per $P_0 = (x_0, y_0, z_0)$, perpendicolare al vettore $\mathbf{n}$ si ottiene imponendo che il punto generico $P = (x, y, z)$ sul piano soddisfi l'equazione:

$$\mathbf{n} \cdot \overrightarrow{P_0P} = 0$$

Più estesamente scriviamo tale equazione come:

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \tag{13.9}$$

L'equazione (13.9) si dice *equazione cartesiana* del piano dato e il vettore $\mathbf{n}$ si dice *vettore normale* al piano.




### Esempio 13.4.1

Vogliamo determinare il piano per il punto $P_0 = (1, -1, 2)$ e con vettore normale $\mathbf{n} = [2, -3, -1]$. Sostituendo nell'equazione (13.9) abbiamo subito:

$$2(x - 1) - 3(y + 1) - (z - 2) = 0$$

Dunque il piano consiste di tutti i punti $(x, y, z)$ che soddisfano l'equazione $2x - 3y - z = 3$.

È facile verificare che, d'altra parte, ogni equazione lineare del tipo:

$$ax + by + cz = d$$

rappresenta un piano normale al vettore $\mathbf{n} = [a, b, c]$.

In modo analogo a quanto abbiamo visto nella sezione precedente possiamo anche scrivere equazioni parametriche che rappresentano il luogo geometrico dei punti appartenenti a un piano individuato da un punto $P_0 = (x_0, y_0, z_0)$ e con due vettori direttori $\mathbf{u} = [u_1, u_2, u_3]$ e $\mathbf{v} = [v_1, v_2, v_3]$ che non siano uno multiplo dell'altro.

Sinteticamente scriviamo:

$$[x, y, z] = [x_0, y_0, z_0] + t[u_1, u_2, u_3] + s[v_1, v_2, v_3]$$

o più estesamente:

$$\begin{cases} x = x_0 + tu_1 + sv_1 \\ y = y_0 + tu_2 + sv_2 \\ z = z_0 + tu_3 + sv_3 \end{cases} \tag{13.10}$$

Le equazioni in (13.10) si dicono *equazioni parametriche* del piano dato. Si noti che mentre le equazioni parametriche di una retta (si veda (13.5)) dipendono da un solo parametro $t$, le equazioni parametriche di un piano dipendono da due parametri $t$ ed $s$. Intuitivamente ciò corrisponde al fatto che mentre lungo una retta possiamo muoverci in una sola direzione (suggestivamente $\mathbf{v}$ si può pensare come la velocità e $t$ il tempo), nel piano abbiamo due gradi di libertà e quindi possiamo muoverci secondo tutte le combinazioni lineari possibili dei due vettori direzione $\mathbf{u}$ e $\mathbf{v}$ attraverso i due parametri $t$ ed $s$.

In modo analogo a quanto abbiamo visto per le rette, possiamo trasformare delle equazioni parametriche di un piano ricavando $t$ ed $s$ da due equazioni e sostituendoli nella terza, ottenendo così una sola equazione lineare, che è appunto un'equazione cartesiana del piano. Essa è determinata univocamente a meno di multipli.

Vediamo qualche esempio.

**Esempio 13.4.2**

Vogliamo determinare in forma parametrica e in forma cartesiana il piano passante per il punto $P = (3, 1, 0)$ con vettore normale $\mathbf{n} = [1, 2, -5]$. Immediatamente dalla formula (13.9) otteniamo l'equazione del piano in forma cartesiana:

$$1(x - 3) + 2(y - 1) - 5z = 0 \qquad \Longrightarrow \qquad x + 2y - 5z = 11$$

Per ottenere le equazioni del piano in forma parametrica procediamo in modo simile a quanto abbiamo visto per le rette: assegniamo a due variabili (in modo arbitrario) i valori dei parametri $t$ ed $s$ e sostituiamo nell'equazione cartesiana:

$$\begin{cases} x = 11 - 2t + 5s \\ y = t \\ z = s \end{cases}$$

Vediamo un altro esempio più complicato.

**Esempio 13.4.3**

Vogliamo determinare in forma parametrica e in forma cartesiana il piano passante per i tre punti: $P = (1, 0, -1)$, $Q = (2, 2, 1)$, $R = (4, 1, 2)$. Innanzitutto determiniamo due vettori direttori:

$$\overrightarrow{PQ} = Q - P = [1, 2, 2], \qquad \overrightarrow{PR} = R - P = [3, 1, 3]$$

Si noti che avremmo potuto benissimo scegliere per esempio $\overrightarrow{PR}$, $\overrightarrow{QR}$: il piano così ottenuto sarebbe lo stesso, invitiamo lo studente a verificare facendo i calcoli. A questo punto una forma parametrica è immediata:

$$\begin{cases} x = 1 + t + 3s \\ y = 0 + 2t + s \\ z = -1 + 2t + 3s \end{cases} \tag{13.11}$$

Per calcolare la forma cartesiana potremmo certamente ricavare $t$ ed $s$ da due equazioni e sostituirli nella terza, tuttavia dalla formula (13.9) sappiamo che per determinare il piano è sufficiente conoscere un punto e un vettore normale al piano. Per determinare un vettore normale al piano, possiamo fare il *prodotto vettoriale* dei due vettori direzione. Tale prodotto è infatti sempre perpendicolare a entrambi i vettori.

Vediamo il calcolo:

$$\mathbf{n} = \overrightarrow{PQ} \times \overrightarrow{PR} = [4, 3, -5]$$

Dunque il piano è dato immediatamente dall'equazione cartesiana:

$$4(x - 1) + 3y - 5(z + 1) = 0$$

cioè $4x + 3y - 5z = 9$.

Terminiamo la sezione dando la definizione di rette sghembe.

**Definizione 13.4.4** Due rette distinte $r$ ed $r'$ in $\mathbb{R}^3$ si dicono *sghembe* se non sono né parallele né incidenti.

## 13.5 ESERCIZI SVOLTI

In questa sezione vogliamo applicare quanto abbiamo introdotto nelle due sezioni precedenti per risolvere semplici problemi geometrici.

**Esercizio 13.5.1**

*Si consideri il piano $\pi$ passante per il punto $P = (1, 0, 1)$ con vettore normale $\mathbf{n} = [1, -2, 4]$. Tale piano è dato dall'equazione cartesiana:*

$$x - 2y + 4z = 5$$

*Si determini la retta $r$ perpendicolare a $\pi$ e passante per il punto $P = (1, -2, 3)$.*

**Svolgimento**

Tale retta ha per vettore direzione il vettore normale al piano e pertanto possiamo subito scrivere le equazioni della retta $r$ in forma parametrica:

$$\begin{cases} x = 1 + t \\ y = -2 - 2t \\ z = 3 + 4t \end{cases} \tag{13.12}$$

che corrispondono alle equazioni in forma cartesiana:

$$\begin{cases} 4x - z = -7 \\ 2x + y = 0 \end{cases} \tag{13.13}$$

■

Prima di procedere con altri esempi, ricordiamo la formula della distanza $d$ tra due punti in $\mathbb{R}^3$ di coordinate $(x_0, y_0, z_0)$ e $(x_1, y_1, z_1)$:

$$d = \sqrt{(x_0 - x_1)^2 + (y_0 - y_1)^2 + (z_0 - z_1)^2}. \tag{13.14}$$

**Esercizio 13.5.2**

*Determinare la distanza del punto $P = (2,-1,3)$ dal piano passante per $Q = (1,-1,-8)$ con vettore normale $\mathbf{n} = (2,-2,-1)$.*

**Svolgimento**

Scriviamo subito l'equazione cartesiana del piano:

$$2x - 2y - z = 12$$

e poi equazioni parametriche della retta passante per $P$ e con vettore direzione $\mathbf{n}$:

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

Calcoliamo poi il punto di intersezione $R$ tra tale retta e il piano dato sostituendo il punto generico ottenuto dalle equazioni parametriche della retta nell'equazione del piano:

$$2(2 + 2t) - 2(-1 - 2t) - (3 - t) = 12$$

Otteniamo $t = 1$ quindi utilizzando la formula della distanza (13.14) otteniamo che $R = (4, -3, 2)$ e la distanza tra $P$ ed $R$ è $\sqrt{17}$. ■

**Esercizio 13.5.3**

*Si considerino la retta $r$ per i due punti $P = (1,0,-2)$ e $Q = (-1,1,-1)$ e la retta $r'$ per il punto $R = (3,1,-5)$ con vettore direzione $\mathbf{v}' = (0,-1,1)$. Stabilire se la retta $r'$ sia perpendicolare a $r$ e calcolare la distanza tra $Q$ e la retta $r'$.*

**Svolgimento**

Calcoliamo un vettore direzione della retta $r$: $\mathbf{v} = [-2, 1, 1]$ e dunque tale retta ha equazioni parametriche:

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

Vediamo subito che $\mathbf{v} \cdot \mathbf{v}' = 0$, tuttavia due rette sono perpendicolari se sono incidenti e se hanno vettori direzione perpendicolari. Dunque dobbiamo verificare se sono incidenti. Scriviamo equazioni parametriche di $r'$:

$$\begin{cases} x = 3 \\ y = 1 - t' \\ z = -5 + t' \end{cases}$$

A questo punto, per verificare se siano o meno incidenti risolviamo il sistema:

$$\begin{cases} 3 = 1 - 2t \\ 1 - t' = 0 + t \\ -5 + t' = -2 + t \end{cases}$$

Vediamo che tale sistema ammette soluzione per $t = -1$ e $t' = 2$, cioè il punto $S = (3, -1, -3)$ appartiene a entrambe le rette. Quindi $r$ ed $r'$ sono perpendicolari.

Vogliamo ora calcolare la distanza tra $Q$ ed $r'$. Poiché le rette sono perpendicolari e si intersecano nel punto $S$ tale distanza sarà data dalla distanza tra i punti $Q$ ed $S$ che possiamo subito calcolare attraverso la formula (13.14):

$$\sqrt{16 + 4 + 4} = 2\sqrt{6}$$

■

### Esercizio 13.5.4

*Determinare la retta intersezione dei due piani:*

$$x + y - z = 0, \qquad y + 2z = 6 \tag{13.15}$$

**Svolgimento**

Prima di procedere notiamo un fatto molto importante: l'espressione di una retta in forma cartesiana e cioè attraverso un sistema di due equazioni nelle incognite $x$, $y$, $z$, non è altro che la sua realizzazione come intersezione di due piani individuati ciascuno dalla propria equazione cartesiana. La direzione della retta è individuata univocamente in quanto perpendicolare a entrambe le direzioni normali dei due piani. Quindi per trovare un vettore perpendicolare a entrambi i vettori normali, ciascuno perpendicolare a uno dei due piani, calcoliamo il prodotto vettoriale dei due vettori normali: $\mathbf{n}_1 = [1, 1, -1]$ e $\mathbf{n}_2 = [0, 1, 2]$:

$$\mathbf{n}_1 \times \mathbf{n}_2 = [3, -2, 1]$$

Scegliamo ora un punto arbitrario sulla retta, cioè un punto che soddisfi entrambe le equazioni (13.15). Possiamo per esempio porre $z = 0$, ottenendo così dalle equazioni $y = 6$ e $x = -6$. Dunque equazioni parametriche della retta intersezione dei due piani dati sono:

$$\begin{cases} x = -6 + 3t \\ y = 6 - 2t \\ z = t \end{cases}$$

■

## 13.6 ESERCIZI PROPOSTI

### Esercizio 13.6.1

*Si considerino le seguenti rette a coppie e si stabilisca qual è la loro posizione reciproca (rette parallele, incidenti o sghembe).*

$$\begin{aligned} r_1 &: \quad x = 1 + t, \quad y = t, \quad z = 2 - 5t \\ r_2 &: \quad x + 1 = y - 2 = 1 - z, \\ r_3 &: \quad x = 1 + t, \quad y = 4 + t, \quad z = 1 - t \\ r_4 &: \quad x = 2 + 2t, \ y = 1 + 2t, \quad z = -3 - 10t \end{aligned}$$

*Si calcoli inoltre la distanza tra le coppie di rette parallele.*

### Esercizio 13.6.2

a) *Sia dato il piano $\pi$:*

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

*Si determini la retta $r$ perpendicolare a $\pi$ e passante per il punto $P = (2, 1, 0)$.*

b) *Si determini il piano $\pi'$ per $Q = (1, 0, -1)$ e $r$, sia in forma parametrica sia in forma cartesiana.*

### Esercizio 13.6.3

*Si determinino equazioni parametriche per la retta intersezione dei piani $x + y - z = 2$, $3x - 4y + 5z = 6$. Si calcoli inoltre l'angolo formato dai due piani, cioè dalle due rette normali ai piani.*

### Esercizio 13.6.4

*In $\mathbb{R}^3$, dati il piano $\pi$, $x + y - z = 0$ e la retta $r$ $x = 2t$, $y = t$, $z = 2t + 1$ si calcoli (se esiste) la retta passante per $(0, 0, 0)$, e per $\pi \cap r$.*

### Esercizio 13.6.5

*Date in $\mathbb{R}^3$ le rette $r, s$ di equazioni cartesiane*

$$r : \begin{cases} x + y - 1 = 0 \\ z - 1 = 0 \end{cases} \qquad s : \begin{cases} x + y - 2z - 2 = 0 \\ z + 1 = 0; \end{cases}$$

a) *Si mostri che* $r, s$ *sono complanari, cioè appartengono allo stesso piano.*

b) *Si determini un'equazione del piano che contiene* $r, s$.

c) *Si determini la distanza tra* $r$ *e* $s$.

### Esercizio 13.6.6

*Dati in* $\mathbb{R}^3$ *i piani* $\pi_1 : x - y + 1 = 0$, $\pi_2 : x + y + 3z = 0$

a) *Si determinino equazioni parametriche per la retta* $r$ *intersezione di* $\pi_1$ *e* $\pi_2$.

b) *Si determini l'equazione cartesiana del piano* $\pi$ *contenente* $r$ *e il punto* $P = (1, 0, 1)$.

c) *Si determinino equazioni parametriche per la retta* $s$ *passante per* $P$ *e ortogonale a* $\pi$.

### Esercizio 13.6.7

*Dati in* $\mathbb{R}^3$ *il piano* $\pi$ *di equazione cartesiana* $\pi : x + y - 2z + 4 = 0$ *e la retta* $r$ *di equazioni cartesiane*

$$\begin{cases} x - 2z + 12 = 0 \\ y - 4 = 0 \end{cases}$$

a) *Si determinino equazioni parametriche per* $r$.

b) *Si determini la posizione relativa di* $\pi, r$.

c) *Si determini la distanza tra* $\pi$ *e* $r$.

### Esercizio 13.6.8

*Dati in* $\mathbb{R}^3$ *il piano* $\pi$ *di equazione cartesiana*

$$\pi : x - y + z + 1 = 0$$

*e il punto* $Q = (1, 1, 0)$

a) *Si determini un'equazione cartesiana del piano passante per* $Q$ *e parallelo a* $\pi$.

b) *Si determini la distanza tra questi due piani.*

c) *Si determinino equazioni cartesiane per la retta* $s$ *passante per* $Q$ *e parallela al vettore* $v = (1, 0, 1)$.

### Esercizio 13.6.9

*Si considerino due vettori generici* $\mathbf{u}$, $\mathbf{v}$ *giacenti sul piano* $xy$. *Si dimostri che la norma di* $\|\mathbf{u} \times \mathbf{v}\|$ *è l'area del parallelogramma di lati* $\mathbf{u}$ *e* $\mathbf{v}$.

### Esercizio 13.6.10

*Si considerino tre vettori generici* $\mathbf{u}$, $\mathbf{v}$ $\mathbf{w}$ *in* $\mathbb{R}^3$. *Si dimostri che* $\|\mathbf{u} \times \mathbf{v} \cdot \mathbf{w}\|$ *è il volume del parallelepipedo di lati* $\mathbf{u}$, $\mathbf{v}$ *e* $\mathbf{w}$.

# 14 Elementi di matematica discreta

In questo capitolo vogliamo studiare l'aritmetica degli interi. Inizieremo con il principio di induzione, un risultato di fondamentale rilievo che ha numerose applicazioni nei vari ambiti della matematica. Procederemo poi con l'algoritmo della divisione e l'algoritmo di Euclide, fino a giungere alle congruenze, l'argomento più importante della matematica discreta elementare.

## 14.1 IL PRINCIPIO DI INDUZIONE

Il principio di induzione è la tecnica principe per dimostrare enunciati riguardanti i numeri naturali cioè l'insieme $\mathbb{N} = \{0, 1, 2, \dots\}$. Facciamo innanzitutto un esempio.

Supponiamo di voler dimostrare che la somma dei numeri interi da 0 a $n$ è $n(n+1)/2$:

$$\sum_{k=0}^{n} k = \frac{n(n+1)}{2} \tag{14.1}$$

Indicheremo tale enunciato con $P(n)$. Verifichiamo innanzitutto la sua validità per $n = 0$: si tratta di mostrare che la somma dei numeri naturali da 0 a 0 è 0. In effetti, $0 = 0(0+1)/2 = 0$. Verifichiamo ora che vale $P(1)$: la somma dei numeri interi da 0 a 1 è $0 + 1 = 1 = 1(1+1)/2$. Analogamente vale $P(2)$: $0 + 1 + 2 = 2(2+1)/2 = 3$. È chiaro che, con un po' di pazienza, potremmo andare avanti così verificando la formula (14.1) per $n$ anche molto grande, ma noi vogliamo dimostrare che essa vale *per tutti i numeri naturali $n$*. Il principio di induzione ci viene in aiuto permettendoci di dimostrare la validità di un enunciato per tutti i numeri naturali.

Enunciamo in primo luogo l'assioma del buon ordinamento. Vedremo in seguito che il principio di induzione e l'assioma del buon ordinamento sono equivalenti tra loro, tuttavia la teoria che illustreremo si basa sulla necessità di prenderne uno come assioma e dimostrare l'altro a partire da questo.

**Assioma del buon ordinamento.** *Ogni sottoinsieme non vuoto dell'insieme dei numeri naturali contiene un elemento minore degli altri.*

**Teorema 14.1.1** **(Principio di induzione)** *Supponiamo che un enunciato $P(n)$ sia dato per ogni numero naturale. Se si verifica che:*

1) *$P(0)$ è vero;*

2) *se $P(k)$ è vero allora $P(k+1)$ è vero;*

*allora $P(n)$ è vero per ogni $n \in \mathbb{N}$.*

**Dimostrazione** – Sia $S$ il sottoinsieme di $\mathbb{N}$ per cui l'enunciato $P(n)$ non è vero. Vogliamo mostrare che $S = \varnothing$. Supponiamo per assurdo che $S$ sia non vuoto. Allora, per l'assioma del buon ordinamento, $S$ contiene un numero naturale $m$ più piccolo di tutti gli altri elementi di $S$. In particolare $P(m)$ non è vero e, dal momento che $P(0)$ è vero per ipotesi, $m \neq 0$. D'altra parte, essendo $m$ il più piccolo elemento di $S$, $m-1$ non appartiene a $S$, cioè $P(m-1)$ è vero. Per l'ipotesi 2), però, $P(m-1+1) = P(m)$ è vero e dunque abbiamo ottenuto una contraddizione.

□

---

**Esempio 14.1.2**

Vediamo come utilizzare il principio di induzione per dimostrare la formula (14.1), ossia il seguente enunciato $P(n)$:

$$\textit{la somma dei primi } n \textit{ numeri naturali è } \frac{n(n+1)}{2}$$

Abbiamo già visto che tale enunciato è vero per $n = 0$, cioè vale l'ipotesi 1) del principio di induzione. Ora supponiamo che $P(k)$ sia vero e cioè che:

$$0 + 1 + 2 + 3 + \cdots + k = \frac{k(k+1)}{2}$$

Dobbiamo mostrare che $P(k) \Longrightarrow P(k+1)$ (ipotesi 2) del principio di induzione), cioè che:

$$0 + 1 + 2 + 3 + \cdots + k + (k+1) = \frac{(k+1)(k+2)}{2}$$

Dal momento che vale $P(k)$, abbiamo:

$$\begin{aligned} 0 + 1 + \cdots + k + (k+1) &= \frac{k(k+1)}{2} + k + 1 = \frac{k^2 + k + 2k + 2}{2} \\ &= \frac{(k+1)(k+2)}{2} \end{aligned}$$

quindi $P(k+1)$ è vero. A questo punto il principio di induzione garantisce che l'enunciato $P(n)$ valga per ogni $n$.

---

L'utilizzo dell'enunciato $P(k)$ per dimostrare $P(k+1)$ prende il nome di *ipotesi induttiva*.

Enunciamo ora una variante del principio di induzione, nota come principio di induzione completa, che si rivela molto utile nelle applicazioni. Nonostante l'enunciato appaia più debole di quello del principio di induzione, al termine di questa sezione vedremo che in effetti il principio di induzione e il principio di induzione completa sono equivalenti.

**Teorema 14.1.3** **(Principio di induzione completa)** *Supponiamo che un enunciato $P(n)$ sia dato per ogni numero naturale. Se si verifica che:*

1. *$P(0)$ è vero;*
2. *se $P(j)$ è vero per ogni $j < k$, allora $P(k)$ è vero;*

*allora $P(n)$ è vero per ogni $n \in \mathbb{N}$.*

**Dimostrazione** – È una conseguenza immediata del principio di induzione. □

**Teorema 14.1.4** *Sono equivalenti:*

1. *l'assioma del buon ordinamento;*
2. *il principio di induzione;*
3. *il principio di induzione completa.*

**Dimostrazione** – Abbiamo già visto che 1. $\Longrightarrow$ 2. $\Longrightarrow$ 3. È dunque sufficiente mostrare che 3. $\Longrightarrow$ 1. L'assioma del buon ordinamento può essere formulato nel modo seguente: "se un insieme $S \subset \mathbb{N}$ non ha elemento minimo, allora $S = \varnothing$" o, equivalentemente, "se un insieme $S \subset \mathbb{N}$ non ha elemento minimo allora $n \notin S$ per ogni $n \in \mathbb{N}$". Indichiamo con $P(n)$ quest'ultima affermazione e mostriamo che essa è vera utilizzando il principio di induzione completa. Sia quindi $S$ un sottoinsieme di $\mathbb{N}$ che non ha elemento minimo. $P(0)$ è vera, cioè $0 \notin S$, altrimenti 0 sarebbe l'elemento minimo di $S$ (essendo 0 l'elemento minimo di $\mathbb{N}$). Supponiamo che $P(j)$ sia vero per ogni $j \le k-1$, cioè supponiamo che $0, 1, 2, \dots, k-1 \notin S$. Vogliamo mostrare che $k \notin S$. Ma, se $k \in S$, allora tale $k$ sarebbe il minimo di $S$ (perché tutti i numeri naturali più piccoli di $k$ non stanno in $S$) e avremmo un assurdo, perché stiamo supponendo che $S$ non abbia elemento minimo. Dunque, per il principio di induzione completa, $P(n)$ vale per ogni $n \in \mathbb{N}$, cioè vale l'assioma del buon ordinamento. □

Questa dimostrazione, apparentemente involuta, è tuttavia molto istruttiva, perché dimostra l'equivalenza di tre enunciati apparentemente diversi tra loro.

## 14.2 ALGORITMO DI DIVISIONE E ALGORITMO DI EUCLIDE

L'algoritmo di divisione formalizza un procedimento che ben conosciamo fin dalle scuole elementari. È necessario tuttavia capire che una procedura, per quanto sia ben definita, deve essere dimostrata, in quanto la sua validità diventa in questo modo assoluta e non confinata ai possibili esempi che riusciamo a costruire.

**Teorema 14.2.1** **(Algoritmo di divisione)** *Siano $n, b \in \mathbb{N}$, $b \neq 0$. Allora esistono e sono unici due numeri interi $q$ e $r$, detti rispettivamente quoziente e resto, tali che:*

$$n = qb + r \qquad \text{con} \quad 0 \le r < b$$

**Dimostrazione** – Dimostriamo dapprima l'esistenza di $q$ e $r$ con il principio di induzione completa. $P(0)$ è vero in quanto $n = 0 = 0b + 0$, con $q = r = 0$. Supponiamo che $P(j)$ sia vero per ogni $j$ tale che $0 \le j < k$ e vogliamo mostrare $P(k)$ (vogliamo, cioè verificare l'ipotesi 2. del principio di induzione completa). Se $k < b$ allora:

$$k = 0b + k \qquad \text{con resto } r = k \quad 0 \le k < b$$

dunque $P(k)$ è vero per $k < b$. Sia ora $k \ge b$. Poiché $0 \le k - b < k$, applicando l'ipotesi induttiva a $k - b$, abbiamo:

$$k - b = q_1 b + r_1 \qquad \text{con} \quad 0 \le r_1 < b$$

da cui ricaviamo:

$$k = (q_1 + 1)b + r_1 \qquad \text{con} \quad 0 \le r_1 < b$$

che dimostra quanto ci eravamo prefissati.

Mostriamo ora che $q$ e $r$ sono unici, e supponiamo che sia $n = q_1 b + r_1 = q_2 b + r_2$, con $0 \le r_1 \le r_2 < b$. Segue allora che $r_2 - r_1 = (q_1 - q_2)b$. Ora al primo membro dell'uguaglianza c'è un intero non negativo minore di $b$, al secondo membro c'è un multiplo di $b$, per cui devono essere entambi nulli, cioè $r_1 = r_2$ e $q_1 = q_2$. □

♦ **Osservazione 14.2.2** Con una dimostrazione analoga, che utilizza l'assioma del buon ordinamento, si ottiene che l'agoritmo di divisione vale per due interi qualsiasi $a$ e $b$, non necessiariamente appartenenti ai numeri naturali. Più precisamente si ha che:

**Teorema 14.2.3** *Dati due interi $a$ e $b$ con $b \ne 0$, esiste un'unica coppia di interi p e $q$ detti rispettivamente quoziente e resto tali che*

$$n = qb + r \qquad \text{con} \quad 0 \le r < |b|$$

*ove $|b|$ indica il valore assoluto di $b$.*

Vogliamo ora introdurre il concetto di *divisibilità* e di *massimo comun divisore.*

**Definizione 14.2.4** Siano $a$ e $b$ due numeri interi. Diciamo che $b$ *divide* $a$ e scriviamo $b|a$, se esiste un numero intero $c$ tale che $a = bc$. Diciamo che $d$ è il *massimo comun divisore* tra due numeri $a$ e $b$ se li divide entrambi ed è il più grande numero intero con questa proprietà. Indicheremo il massimo comun divisore tra $a$ e $b$ con $\gcd(a, b)$.

♦ **Osservazione 14.2.5** Osserviamo che, se $a, b$ e $c$ sono interi, e $a$ divide sia $b$ che $c$, allora $a$ divide anche $b + c$ e $b - c$. Lasciamo allo studente la facile verifica di questa proprietà, che in seguito useremo più volte.

Vogliamo ora trovare un algoritmo efficiente per determinare il massimo comun divisore tra due interi.

**Teorema 14.2.6** **(Algoritmo di Euclide)** *Siano $a$ e $b$ due numeri interi positivi tali che $b \le a$ e $b$ non divide $a$. Abbiamo allora:*

$$\begin{aligned} a &= q_0 b + r_0 \quad \text{ove} \quad 0 \le r_0 < b \\ b &= q_1 r_0 + r_1 \quad \text{ove} \quad 0 \le r_1 < r_0 \\ r_0 &= q_2 r_1 + r_2 \quad \text{ove} \quad 0 \le r_2 < r_1 \\ &\vdots \\ r_{t-2} &= q_t r_{t-1} + r_t \quad \text{ove} \quad 0 < r_t < r_{t-1} \\ r_{t-1} &= q_{t+1} r_t \end{aligned}$$

*e l'ultimo resto non nullo $r_t$ è il massimo comun divisore tra $a$ e $b$.*

**Dimostrazione** – Per il Teorema 14.2.1 possiamo scrivere: $a = q_0 b + r_0$. Ora vogliamo mostrare che $\gcd(a, b) = \gcd(b, r_0)$. Infatti, se $c|a$ e $c|b$, allora $c|r_0$, poiché $r_0 = a - q_0 b$. Analogamente, se $c|b$ e $c|r_0$, allora $c|a = q_0 b + r_0$. Dunque l'insieme degli interi che dividono sia $a$ che $b$ coincide con l'insieme degli interi che dividono sia $b$ che $r_0$ e dunque il massimo comune divisore delle due coppie $(a, b)$ e $(b, r_0)$ è lo stesso. Detto questo, il risultato segue immediatamente dalla catena di uguaglianze:

$$\gcd(a, b) = \gcd(b, r_0) = \gcd(r_0, r_1) = \cdots = \gcd(r_{t-1}, r_t) = r_t \qquad \square$$

Vediamo concretamente come utilizzare questo algoritmo per determinare il massimo comun divisore di due numeri dati.

---

**Esempio 14.2.7**

Vogliamo calcolare $\gcd(603, 270)$. Utilizziamo l'algoritmo di Euclide (Teorema 14.2.6).

$$\begin{aligned} 603 &= 2 \cdot 270 + 63 \\ 270 &= 4 \cdot 63 + 18 \\ 63 &= 3 \cdot 18 + 9 \\ 18 &= 2 \cdot 9 \end{aligned}$$

Dunque $\gcd(603, 270) = 9$.

---

Il teorema che segue è una conseguenza dell'algoritmo di Euclide e sarà lo strumento fondamentale per la risoluzione delle congruenze che studieremo nel Paragrafo 14.4.

**Teorema 14.2.8** **(Identità di Bézout)** *Siano $a, b$ interi positivi e sia $d = \gcd(a, b)$. Allora esistono due numeri interi $u$, $v$ (non unici) tali che:*

$$d = ua + vb$$

**Dimostrazione** – La dimostrazione di questo risultato ripercorre l'algoritmo di Euclide 14.2.6. Dimostriamo che a ogni passo esistono $u_i, v_i \in \mathbb{Z}$ tali che $r_i = u_i a + v_i b$.

Per $r_0$ il risultato è vero, con $u_0 = 1$ e $v_0 = -q_0$, infatti

$$r_0 = a - q_0 b$$

Si ha poi che:

$$r_1 = b - r_0 q_1 = b - (u_0 a + v_0 b) q_1 = -u_0 a q_1 + (1 - v_o q_1) b$$

quindi il risultato è vero anche per $r_1$, basta prendere $u_1 = -u_0 q_1$ e $v_1 = 1 - v_0 q_1$. In generale, dopo il passo $i - 1$, conosciamo $u_{i-2}, u_{i-1}, v_{i-2}, v_{i-1} \in \mathbb{Z}$ tali che:

$$r_{i-2} = u_{i-2} a + v_{i-2} b, \quad r_{i-1} = u_{i-1} a + v_{i-1} b$$

Allora abbiamo che:

$$\begin{aligned} r_i &= r_{i-2} - r_{i-1} q_i = u_{i-2} a + v_{i-2} b - (u_{i-1} a + v_{i-1} b) q_i \\ &= (u_{i-2} - u_{i-1} q_i) a + (v_{i-2} - v_{i-1} q_i) b \end{aligned}$$

quindi il risultato è vero per $r_i$, con $u_i = u_{i-2} - u_{i-1} q_i$ e $v_i = v_{i-2} - v_{i-1} q_i$. Poiché $\gcd(a, b)$ è l'ultimo resto non nullo $r_t$, dopo il passo $t$ conosciamo $u_t$ e $v_t$ tali che $r_t = \gcd(a, b) = u_t a + v_t b$ e $u = u_t$, $v = v_t$ sono gli interi cercati. □

♦ **Osservazione 14.2.9** Nel teorema precedente l'esistenza di due numeri $u$, $v$ tali che $d = ua + vb$ non garantisce che $d = \gcd(a, b)$. Per esempio $10 = 15 \cdot 14 - 25 \cdot 8$ ma $\gcd(14, 8) = 2$.

---

**Esempio 14.2.10**

Nell'Esempio 14.2.7 abbiamo calcolato $\gcd(603, 270) = 9$. Vogliamo ora ricavare due numeri $u$ e $v$ tali che $u\,603 + v\,270 = 9$. Procediamo a ritroso, sostituendo attentamene i resti nella sequenza di equazioni ricavate nell'Esempio 14.2.7.

$$\begin{aligned} 9 &= 63 - 3 \cdot 18 \\ &= 63 - 3 \cdot [270 - 4 \cdot 63] = 63 - 3 \cdot 270 + 12 \cdot 63 \\ &= (-3) \cdot 270 + 13 \cdot 63 = (-3) \cdot 270 + 13 \cdot [603 - 2 \cdot 270] \\ &= (-3) \cdot 270 + 13 \cdot 603 + (-26) \cdot 270 \\ &= 13 \cdot 603 + (-29) \cdot 270 \end{aligned}$$

Dunque $9 = 13 \cdot 603 + (-29) \cdot 270$ cioè vale la relazione $u\,603 + v\,270 = 9$ con $u = 13$ e $v = -29$.

---

Concludiamo questo paragrafo con un risultato che non dimostriamo e che non utilizzeremo in seguito, ma che riveste un ruolo fondamentale per la teoria degli interi e le cui generalizzazioni sono estremamente importanti in teoria dei numeri.

**Definizione 14.2.11** Diciamo che un intero $p$ maggiore di 1 è *primo* se ha come unici divisori $\pm p$ e $\pm 1$.

**Teorema 14.2.12** **(Teorema fondamentale dell'aritmetica)** *Ogni intero maggiore di 1 è prodotto di numeri primi, in modo unico a meno dell'ordine:*

$$n = p_1 p_2 \dots p_r$$

*ove* $p_1, \dots, p_r$ *sono numeri primi (non necessariamente distinti).*

## 14.3 CLASSI DI CONGRUENZA

Le classi di congruenza rappresentano un modo di contare, di sommare e di moltiplicare i numeri diverso da quello che conosciamo dall'aritmetica elementare e che, tuttavia, utilizziamo in ambiti particolari ogni giorno. Per esempio, se chiediamo a qualcuno di incontrarci precisamente tra 10 ore e sono le 23, la persona con cui stiamo parlando sa benissimo che l'incontro avverrà alle 9 e non alle 33, come la somma di numeri interi suggerirebbe. Analogamente, se le lancette del nostro orologio segnano le 8 e abbiamo un appuntamento tra 6 ore, sappiamo che all'ora dell'appuntamento l'orologio segnerà le 2.

Vogliamo formalizzare questo modo di effettuare operazioni, che nei casi esaminati consiste semplicemente nel fare la somma dei numeri, dividere il risultato per un certo intero (che nel primo esempio è 24, nel secondo 12) e poi prendere il resto della divisione.

**Definizione 14.3.1** Siano $a$, $b$ e $n$ tre interi, con $n > 0$. Diciamo che $a$ è *congruo* (o *congruente*) a $b$ modulo $n$ e scriviamo $a \equiv_n b$ se $n | a - b$.

La relazione di congruenza tra numeri interi verifica le seguenti proprietà:

(i) Proprietà riflessiva: $a \equiv_n a$, infatti $n | a - a = 0$.

(ii) Proprietà simmetrica: se $a \equiv_n b$ allora $b \equiv_n a$. Infatti, se $n | a - b$, allora $n | b - a$.

(iii) Proprietà transitiva: se $a \equiv_n b$ e $b \equiv_n c$ allora $a \equiv_n c$. Infatti, se $n | a - b$ e $n | b - c$, allora $n | a - b + b - c = a - c$.

Una relazione soddisfacente la proprietà riflessiva, la proprietà simmetrica e la proprietà transitiva si chiama *relazione di equivalenza*. Non ci soffermeremo sulle relazioni di equivalenza in generale, tuttavia utilizzeremo le tre proprietà delle congruenze elencate sopra.

Enunciamo ora un risultato che ci sarà utile in seguito, la cui dimostrazione lasciamo per esercizio.

**Proposizione 14.3.2** *Siano* $a, b, c, d, n \in \mathbb{Z}$, *con* $n > 0$. *Se* $a \equiv_n b$ *e* $c \equiv_n d$, *allora:*

1. $a + c \equiv_n b + d$;

2. $ac \equiv_n bd$.

**Dimostrazione** – La dimostrazione di questo risultato ripercorre l'algoritmo di Euclide 14.2.6. Dimostriamo che a ogni passo esistono $u_i, v_i \in \mathbb{Z}$ tali che $r_i = u_i a + v_i b$.

Per $r_0$ il risultato è vero, con $u_0 = 1$ e $v_0 = -q_0$, infatti

$$r_0 = a - q_0 b$$

Si ha poi che:

$$r_1 = b - r_0 q_1 = b - (u_0 a + v_0 b) q_1 = -u_0 a q_1 + (1 - v_o q_1) b$$

quindi il risultato è vero anche per $r_1$, basta prendere $u_1 = -u_0 q_1$ e $v_1 = 1 - v_0 q_1$. In generale, dopo il passo $i - 1$, conosciamo $u_{i-2}, u_{i-1}, v_{i-2}, v_{i-1} \in \mathbb{Z}$ tali che:

$$r_{i-2} = u_{i-2} a + v_{i-2} b, \quad r_{i-1} = u_{i-1} a + v_{i-1} b$$

Allora abbiamo che:

$$\begin{aligned} r_i = r_{i-2} - r_{i-1} q_i &= u_{i-2} a + v_{i-2} b - (u_{i-1} a + v_{i-1} b) q_i \\ &= (u_{i-2} - u_{i-1} q_i) a + (v_{i-2} - v_{i-1} q_i) b \end{aligned}$$

quindi il risultato è vero per $r_i$, con $u_i = u_{i-2} - u_{i-1} q_i$ e $v_i = v_{i-2} - v_{i-1} q_i$. Poiché $\gcd(a, b)$ è l'ultimo resto non nullo $r_t$, dopo il passo $t$ conosciamo $u_t$ e $v_t$ tali che $r_t = \gcd(a, b) = u_t a + v_t b$ e $u = u_t$, $v = v_t$ sono gli interi cercati. □

♦ **Osservazione 14.2.9** Nel teorema precedente l'esistenza di due numeri $u$, $v$ tali che $d = ua + vb$ non garantisce che $d = \gcd(a, b)$. Per esempio $10 = 15 \cdot 14 - 25 \cdot 8$ ma $\gcd(14, 8) = 2$.

---

**Esempio 14.2.10**

Nell'Esempio 14.2.7 abbiamo calcolato $\gcd(603, 270) = 9$. Vogliamo ora ricavare due numeri $u$ e $v$ tali che $u\,603 + v\,270 = 9$. Procediamo a ritroso, sostituendo attentamene i resti nella sequenza di equazioni ricavate nell'Esempio 14.2.7.

$$\begin{aligned} 9 &= 63 - 3 \cdot 18 \\ &= 63 - 3 \cdot [270 - 4 \cdot 63] = 63 - 3 \cdot 270 + 12 \cdot 63 \\ &= (-3) \cdot 270 + 13 \cdot 63 = (-3) \cdot 270 + 13 \cdot [603 - 2 \cdot 270] \\ &= (-3) \cdot 270 + 13 \cdot 603 + (-26) \cdot 270 \\ &= 13 \cdot 603 + (-29) \cdot 270 \end{aligned}$$

Dunque $9 = 13 \cdot 603 + (-29) \cdot 270$ cioè vale la relazione $u\,603 + v\,270 = 9$ con $u = 13$ e $v = -29$.

---

Concludiamo questo paragrafo con un risultato che non dimostriamo e che non utilizzeremo in seguito, ma che riveste un ruolo fondamentale per la teoria degli interi e le cui generalizzazioni sono estremamente importanti in teoria dei numeri.

**Definizione 14.2.11** Diciamo che un intero $p$ maggiore di 1 è *primo* se ha come unici divisori $\pm p$ e $\pm 1$.

**Teorema 14.2.12** **(Teorema fondamentale dell'aritmetica)** *Ogni intero maggiore di 1 è prodotto di numeri primi, in modo unico a meno dell'ordine:*

$$n = p_1 p_2 \dots p_r$$

*ove $p_1, \dots, p_r$ sono numeri primi (non necessariamente distinti).*

## 14.3 CLASSI DI CONGRUENZA

Le classi di congruenza rappresentano un modo di contare, di sommare e di moltiplicare i numeri diverso da quello che conosciamo dall'aritmetica elementare e che, tuttavia, utilizziamo in ambiti particolari ogni giorno. Per esempio, se chiediamo a qualcuno di incontrarci precisamente tra 10 ore e sono le 23, la persona con cui stiamo parlando sa benissimo che l'incontro avverrà alle 9 e non alle 33, come la somma di numeri interi suggerirebbe. Analogamente, se le lancette del nostro orologio segnano le 8 e abbiamo un appuntamento tra 6 ore, sappiamo che all'ora dell'appuntamento l'orologio segnerà le 2.

Vogliamo formalizzare questo modo di effettuare operazioni, che nei casi esaminati consiste semplicemente nel fare la somma dei numeri, dividere il risultato per un certo intero (che nel primo esempio è 24, nel secondo 12) e poi prendere il resto della divisione.

**Definizione 14.3.1** Siano $a$, $b$ e $n$ tre interi, con $n > 0$. Diciamo che $a$ è *congruo* (o *congruente*) a $b$ modulo $n$ e scriviamo $a \equiv_n b$ se $n | a - b$.

La relazione di congruenza tra numeri interi verifica le seguenti proprietà:

(i) Proprietà riflessiva: $a \equiv_n a$, infatti $n | a - a = 0$.

(ii) Proprietà simmetrica: se $a \equiv_n b$ allora $b \equiv_n a$. Infatti, se $n | a - b$, allora $n | b - a$.

(iii) Proprietà transitiva: se $a \equiv_n b$ e $b \equiv_n c$ allora $a \equiv_n c$. Infatti, se $n | a - b$ e $n | b - c$, allora $n | a - b + b - c = a - c$.

Una relazione soddisfacente la proprietà riflessiva, la proprietà simmetrica e la proprietà transitiva si chiama *relazione di equivalenza*. Non ci soffermeremo sulle relazioni di equivalenza in generale, tuttavia utilizzeremo le tre proprietà delle congruenze elencate sopra.

Enunciamo ora un risultato che ci sarà utile in seguito, la cui dimostrazione lasciamo per esercizio.

**Proposizione 14.3.2** *Siano $a, b, c, d, n \in \mathbb{Z}$, con $n > 0$. Se $a \equiv_n b$ e $c \equiv_n d$, allora:*

1. $a + c \equiv_n b + d$;

2. $ac \equiv_n bd$.

Definiamo ora le classi di congruenza, cioè gli insiemi che contengono numeri interi congrui tra loro.

**Definizione 14.3.3** Siano $a, n \in \mathbb{Z}$. Si chiama *classe di congruenza di a modulo n* e si indica con $[a]_n$ l'insieme di numeri interi congrui ad $a$ modulo $n$:

$$[a]_n = \{b \in \mathbb{Z} \mid b \equiv_n a\} = \{a + kn \mid k \in \mathbb{Z}\}$$

---

**Esempio 14.3.4**

Per esempio, se prendiamo $n = 4$, abbiamo che:

$$\begin{aligned}
[0]_4 &= \{0, 4, -4, 8, -8 \dots\} \\
[1]_4 &= \{1, 5, -3, 9, -7 \dots\} \\
[2]_4 &= \{2, 6, -2, 10, -6 \dots\} \\
[3]_4 &= \{3, 7, -1, 11, -5 \dots\}
\end{aligned}$$

Si noti che:

$$\begin{aligned}
[4]_4 &= [0]_4 = [-4]_4 = \dots \\
[5]_4 &= [1]_4 = [-3]_4 = \dots \\
[6]_4 &= [2]_4 = [-2]_4 = \dots \\
[7]_4 &= [3]_4 = [-1]_4 = \dots
\end{aligned}$$

Notiamo che in questo esempio:

- non ci sono elementi comuni a due classi di congruenza diverse;
- l'unione delle classi di congruenza è l'insieme di tutti i numeri interi;
- c'è un numero finito di classi di congruenza.

---

Questi fatti, come vedremo, valgono in generale.

**Proposizione 14.3.5** *Siano $[a]_n$ e $[b]_n$ due classi di congruenza modulo $n$. Allora $[a]_n = [b]_n$, oppure $[a]_n$ e $[b]_n$ sono disgiunte, cioè non hanno elementi in comune.*

**Dimostrazione** – Supponiamo che ci sia un elemento in comune $c$ tra $[a]_n$ e $[b]_n$, cioè $c \equiv_n a$ e $c \equiv_n a$. Allora, per la proprietà transitiva delle congruenze $a \equiv_n b$, quindi, ancora per la proprietà transitiva, $[a]_n = [b]_n$. □

La dimostrazione della seguente proposizione è immediata.

**Proposizione 14.3.6**

1. *Sia $r$ il resto della divisione di $a$ per $n$. Allora $[a]_n = [r]_n$;*

2. *$[0]_n, [1]_n, \dots, [n-1]_n$ sono tutte e sole le classi di congruenza distinte modulo $n$;*

3. *$[0]_n \cup [1]_n \cup \dots \cup [n-1]_n = \mathbb{Z}$.*

Siamo giunti alla definizione più importante di questo capitolo: l'insieme $\mathbb{Z}_n$.

**Definizione 14.3.7** Si chiama insieme degli *interi modulo n*, e si indica con $\mathbb{Z}_n$, l'insieme delle classi di congruenza modulo $n$:

$$\mathbb{Z}_n = \{[0]_n, [1]_n, \ldots, [n-1]_n\}$$

**Definizione 14.3.8** Definiamo sull'insieme $\mathbb{Z}_n$ le seguenti operazioni di somma e prodotto:

$$[a]_n + [b]_n = [a+b]_n \qquad [a]_n [b]_n = [ab]_n$$

♦ **Osservazione 14.3.9** Le operazioni appena definite non dipendono dai numeri $a$ e $b$ che scegliamo per rappresentare le classi di congruenza che sommiamo o moltiplichiamo, ma soltanto dalla loro classe di congruenza. Si dice in tal caso che le operazioni sono *ben definite*. Per esempio, in $\mathbb{Z}_4$ si ha: $[1]_4 = [5]_4$ e $[2]_4 = [6]_4$. Per definizione $[1]_4 + [2]_4 = [3]_4 = [11]_4 = [5]_4 + [6]_4$.

**Esempio 14.3.10**

Calcoliamo le tavole dell'addizione e della moltiplicazione per $\mathbb{Z}_3$ e $\mathbb{Z}_4$ invitando lo studente a esercitarsi a costruire le tavole analoghe per $\mathbb{Z}_5$ e $\mathbb{Z}_6$:

$\mathbb{Z}_3$

| + | $[0]_3$ | $[1]_3$ | $[2]_3$ |
|---|---|---|---|
| $[0]_3$ | $[0]_3$ | $[1]_3$ | $[2]_3$ |
| $[1]_3$ | $[1]_3$ | $[2]_3$ | $[0]_3$ |
| $[2]_3$ | $[2]_3$ | $[0]_3$ | $[1]_3$ |

| · | $[0]_3$ | $[1]_3$ | $[2]_3$ |
|---|---|---|---|
| $[0]_3$ | $[0]_3$ | $[0]_3$ | $[0]_3$ |
| $[1]_3$ | $[0]_3$ | $[1]_3$ | $[2]_3$ |
| $[2]_3$ | $[0]_3$ | $[2]_3$ | $[1]_3$ |

$\mathbb{Z}_4$

| + | $[0]_4$ | $[1]_4$ | $[2]_4$ | $[3]_4$ |
|---|---|---|---|---|
| $[0]_4$ | $[0]_4$ | $[1]_4$ | $[2]_4$ | $[3]_4$ |
| $[1]_4$ | $[1]_4$ | $[2]_4$ | $[3]_4$ | $[0]_4$ |
| $[2]_4$ | $[2]_4$ | $[3]_4$ | $[0]_4$ | $[1]_4$ |
| $[3]_4$ | $[3]_4$ | $[0]_4$ | $[1]_4$ | $[2]_4$ |

| · | $[0]_4$ | $[1]_4$ | $[2]_4$ | $[3]_4$ |
|---|---|---|---|---|
| $[0]_4$ | $[0]_4$ | $[0]_4$ | $[0]_4$ | $[0]_4$ |
| $[1]_4$ | $[0]_4$ | $[1]_4$ | $[2]_4$ | $[3]_4$ |
| $[2]_4$ | $[0]_4$ | $[2]_4$ | $[0]_4$ | $[2]_4$ |
| $[3]_4$ | $[0]_4$ | $[3]_4$ | $[2]_4$ | $[1]_4$ |

Notiamo alcuni fatti molto importanti: in $\mathbb{Z}_3$ ogni elemento diverso da $[0]_3$ ammette un *inverso*, cioè per ogni $[a]_3 \neq [0]_3$ esiste un elemento $[b]_3$ tale che $[a]_3[b]_3 = [1]_3$. Tale inverso si indica con $[a]_3^{-1}$. Abbiamo dunque: $[1]_3^{-1} = [1]_3$, $[2]_3^{-1} = [2]_3$. Tale proprietà invece non vale nel caso di $\mathbb{Z}_4$. Infatti la tavola moltiplicativa mostra che non esiste alcun inverso della classe $[2]_4$. Come vedremo dettagliatamente nel prossimo paragrafo, questa diversità è legata al fatto che 3 è un numero primo mentre 4 non lo è.

## 14.4 CONGRUENZE

In questo paragrafo ci proponiamo di risolvere equazioni lineari in cui l'incognita appartiene all'insieme $\mathbb{Z}_n$ introdotto nel paragrafo precedente.

Incominciamo esaminando la struttura di $\mathbb{Z}_p$ con $p$ numero primo.

**Proposizione 14.4.1** *Le seguenti affermazioni sono equivalenti:*

1) *$p$ è un numero primo.*

2) *L'equazione $[a]_p\, x = [1]_p$, con $[a]_p \neq [0]_p$, ha soluzione in $\mathbb{Z}_p$, cioè ogni elemento $[a]_p \neq [0]_p$ in $\mathbb{Z}_p$ ammette inverso.*

3) *Se $[a]_p[b]_p = [0]_p$ in $\mathbb{Z}_p$ allora $[a]_p = [0]_p$ oppure $[b]_p = [0]_p$.*

**Dimostrazione** – 1) $\Longrightarrow$ 2): poiché $[a]_p \neq [0]_p$, $p$ non divide $a$, quindi $\gcd(a, p) = 1$. Allora per il Teorema 14.2.8 esistono $u, v \in \mathbb{Z}$ tali che $1 = au+pv$. Prendendo le classi di congruenza modulo $p$, abbiamo: $[1]_p = [a]_p[u]_p + [p]_p[v]_p = [a]_p[u]_p$, dunque $x = [u]_p \in \mathbb{Z}_p$ è una soluzione di $[a]_p\, x = [1]_p$.

2) $\Longrightarrow$ 3): sia $[a]_p[b]_p = [0]_p$ con $[a]_p \neq [0]_p$. Per ipotesi esiste un inverso di $[a]_p$, cioè esiste un elemento $[u]_p$ tale che $[u]_p[a]_p = [1]_p$. Moltiplicando per $[u]_p$ entrambi i membri dell'uguaglianza $[a]_p[b]_p = [0]_p$ otteniamo: $[u]_p[a]_p[b]_p = [u]_p[0]_p = [0]_p$, cioè $[b]_p = [0]_p$.

3) $\Longrightarrow$ 1): supponiamo che sia $p = ab$ e mostriamo che necessariamente $a$ e $b$ sono uguali a $\pm 1$ o $\pm p$, cioè che gli unici divisori di $p$ sono, a meno del segno, $p$ stesso e 1. Osserviamo, considerando i valori assoluti, che $|a||b| = |ab| = |p| = p$ quindi $|a| \leq p$ e $|b| \leq p$. L'uguaglianza $p = ab$ si traduce in $\mathbb{Z}_p$ nell'uguaglianza $[p]_p = [a]_p[b]_p$, cioè $[a]_p[b]_p = [0]_p$. Per ipotesi sappiamo che $[a]_p = [0]_p$ oppure $[b]_p = [0]_p$. Se $[a]_p = [0]_p$, allora $a = \pm p$ e $b = \pm 1$. Se $[b]_p = [0]_p$, allora $b = \pm p$ e $a = \pm 1$. □

**Corollario 14.4.2** *Se $p$ è un numero primo, l'equazione $[a]_p\, x = [b]_p$, con $[a]_p \neq [0]_p$, ha un'unica soluzione in $\mathbb{Z}_p$.*

**Dimostrazione** – Per la proprietà 2) della proposizione precedente sappiamo che $[a]_p$ è invertibile. Moltiplicando l'equazione data per $[a]_p^{-1}$ otteniamo $x = [a]_p^{-1}[b]_p$, il che mostra allo stesso tempo che la soluzione esiste ed è unica. ora l'unicità. Se $[c_1]_p, [c_2]_p$ sono due soluzioni, abbiamo che: $[a]_p\, [c_1]_p = [b]_p$ e $[a]_p\, [c_2]_p = [b]_p$ , quindi $[a]_p\, ([c_1]_p - [c_2]_p) = [0]_p$ e moltiplicando l'equazione data per $[a]_p^{-1}$ otteniamo che $[c_1]_p - [c_2]_p = [0]_p$, cioè $[c_1]_p = [c_2]_p$. □

Il risultato che segue è di notevole utilità nella risoluzione degli esercizi.

**Proposizione 14.4.3** *Se $\gcd(a, n) = 1$ l'equazione $[a]_n\, x = [b]_n$ ha un'unica soluzione in $\mathbb{Z}_n$.*

**Dimostrazione** – Poiché $\gcd(a, n) = 1$, per il Teorema 14.2.8 esistono $u, v \in \mathbb{Z}$ tali che $au + nv = 1$, dunque $[a]_n$ è invertibile in $\mathbb{Z}_n$, con inverso $[u]_n$. Ragionando come nella dimostrazione del corollario precedente, otteniamo quanto voluto. □

A questo punto è facile caratterizzare quali elementi di $\mathbb{Z}_n$ hanno un inverso.

**Proposizione 14.4.4** *La classe $[a]_n$ ammette un inverso in $\mathbb{Z}_n$ se e solo se $\gcd(a, n) = 1$.*

**Dimostrazione** – Supponiamo che $\gcd(a,n) = 1$. Allora per la Proposizione 14.4.3 l'equazione $[a]_n\ x = [1]_n$ ha un'unica soluzione $[c]_n$ in $\mathbb{Z}_n$. Allora $[c]_n$ è proprio l'inverso di $[a]_n$.

Viceversa, supponiamo che esista una classe $[c]_n \in Z_n$ tale che $[a]_n\ [c]_n = [1]_n$, quindi $n|1 - ac$, cioè $1 - ac = nr$ per qualche $r \in Z$. Sia $d = \gcd(a,n)$; abbiamo che $d|a$, $d|n$, quindi $d|ac + nr = 1$, cioè $d = 1$, come volevasi. □

Esaminiamo ora il caso generale della Proposizione 14.4.3.

**Teorema 14.4.5** *Siano $a, n \in \mathbb{Z}$, $n > 0$, e sia $d = \gcd(a,n)$. Allora:*

1) *l'equazione $[a]_n\ x = [b]_n$ ha soluzione in $\mathbb{Z}_n$ se e solo se $d|b$;*
2) *se $d|b$ allora l'equazione $[a]_n\ x = [b]_n$ ha esattamente $d$ soluzioni distinte in $\mathbb{Z}_n$.*

**Dimostrazione** – Supponiamo che $[c]_n$ sia una soluzione dell'equazione data, quindi: $[a]_n[c]_n = [ac]_n = [b]_n$, cioè $ac \equiv_n b$ o, equivalentemente, $n|ac - b$. Di conseguenza $d|b$ dal momento che $d|n$ e $d|a$.

Supponiamo ora che $d$ divida $b$ e siano $a = a'd$, $n = n'd$, $b = b'd$. Osserviamo che $\gcd(a', n') = 1$, altrimenti $d$ non sarebbe il massimo comun divisore tra $a$ e $n$. Allora per la Proposizione 14.4.3 l'equazione $[a']_{n'}\ x = [b']_{n'}$ ha un'unica soluzione in $\mathbb{Z}_{n'}$, sia essa $[c]_{n'}$. Si ha dunque: $[a']_{n'}[c]_{n'} = [a'c]_{n'} = [b']_{n'}$, cioè $n'|a'c - b'$, quindi $n = n'd|a'dc - b'd = ac - b$, cioè $[c]_n$ è soluzione dell'equazione $[a]_n\ x = [b]_n$. Questo dimostra *1)*.

Resta da dimostrare che, se $d|b$, ci sono $d$ soluzioni distinte in $\mathbb{Z}_n$. Sia $[c]_n$ la soluzione trovata al punto 1). Se $[e]_n$ è un'altra soluzione dell'equazione data, allora si ha che $[a]_n[c]_n = [b]_n = [a]_n[e]_n$, quindi $n|ac - ae$, cioè $n'd|a'dc - a'de$ e dunque $n'|a'c - a'e$. Segue che $[a']_{n'}[e]_{n'} = [a']_{n'}[c]_{n'}$ e quindi $[e]_{n'}$ è soluzione dell'equazione $[a']_{n'}\ x = [b']_{n'}$. Per la Proposizione 14.4.3 si ha che $[e]_{n'} = [c]_{n'}$. Cioè $e = c + kn'$, con $k \in \mathbb{Z}$.

Si verifica poi facilmente che $[e]_n \in \{[c]_n, [c + n']_n, [c + 2n']_n, \dots, [c + (d-1)n']_n\} = X$ e che gli elementi di $X$ sono tutti distinti e sono tutti soluzioni dell'equazione $[a]_n\ x = [b]_n$. Questo dimostra quanto voluto. □

---

**Esempio 14.4.6**

Vogliamo determinare in $\mathbb{Z}_{74}$ tutte le soluzioni dell'equazione $[33]_{74}x = [5]_{74}$.

Applichiamo l'algoritmo di Euclide:

$$74 = 2 \cdot 33 + 8$$

$$33 = 4 \cdot 8 + 1$$

$$8 = 8 \cdot 1$$

Poiché $(33, 74) = 1$ sappiamo che la soluzione esiste ed è unica, e i conti appena fatti ci permettono di calcolare l'inverso di $[33]_{74}$. Abbiamo infatti che:

$$1 = 33 - 4 \cdot 8 = 33 - 4 \cdot (74 - 2 \cdot 33) = (-4) \cdot 74 + 9 \cdot 33.$$

Abbiamo perciò trovato: $1 = (-4) \cdot 74 + 9 \cdot 33$. Scriviamo questa equazione in $\mathbb{Z}_{74}$:

$$[1]_{74} = [9]_{74}[33]_{74}$$

Quindi la soluzione cercata è $x = [9]_{74} \cdot [5]_{74} = [45]_{74}$.

---

Vediamo ora come il Teorema 14.4.5 ci permetta anche di risolvere le congruenze lineari.

**Definizione 14.4.7** Si dice *congruenza lineare* nell'incognita $x$ (modulo $n$) una congruenza del tipo:

$$ax \equiv_n b$$

con $a, b, n \in \mathbb{Z}$, $n > 0$.

♦ **Osservazione 14.4.8** È chiaro che un intero $c$ è soluzione della congruenza lineare $ax \equiv_n b$ se e solo se $[c]_n$ è soluzione dell'equazione $[a]_n\, x = [b]_n$.

Riprendiamo un attimo in mano la dimostrazione del Teorema 14.4.5, utilizzando le stesse notazioni. Abbiamo che l'equazione $[a]_n\, x = [b]_n$ ha soluzione se e solo se $d$ divide $n$, ove $d = \gcd(a, n)$. In tal caso, posto $n = n'd$, se $[c]_n$ è una soluzione dell'equazione $[a]_n\, x = [b]_n$, per esempio quella trovata con il metodo del punto 1), gli interi $e$ tali che $[e]_n$ è soluzione dell'equazione $[a]_n\, x = [b]_n$ sono tutti e soli quelli del tipo $e = c + kn'$, con $k \in \mathbb{Z}$. Perciò le soluzioni della congruenza lineare $ax \equiv_n b$ sono precisamente tutte e sole quelle del tipo $e = c + kn'$, con $k \in \mathbb{Z}$.

## 14.5 ESERCIZI SVOLTI

### Esercizio 14.5.1

*Si dimostri per induzione che un insieme con $n$ elementi ha $2^n$ sottoinsiemi.*

**Svolgimento**

Vediamo di dimostrare $P(0)$. Se un insieme è vuoto allora ha zero elementi e dunque un solo sottoinsieme, se stesso; pertanto ha $1 = 2^0$ sottoinsieme. Supponiamo che l'enunciato $P(k-1)$ sia vero e vogliamo dimostrare $P(k)$, con $k \geq 1$. L'enunciato $P(k)$ consiste nell'affermazione che un insieme $S$ con $k$ elementi ha $2^k$ sottoinsiemi. Poiché $S$ contiene almeno un elemento $x$, possiamo pensare $S$ come l'unione di un suo sottoinsieme $S'$ e il sottoinsieme $\{x\}$. I sottoinsiemi di $S$ che non contengono $x$ sono anche sottoinsiemi di $S'$ e ce ne sono $2^{k-1}$ per l'ipotesi induttiva $P(k-1)$ in quanto $S'$ contiene $k-1$ elementi. D'altra parte un sottoinsieme di $S$ che contiene $x$ è dato da $T \cup \{x\}$, ove $T$ è un sottoinsieme di $S'$. È facile convincersi che ce ne sono esattamente $2^{k-1}$. Pertanto i sottoinsiemi di $S$ sono $2^{k-1}+2^{k-1} = 2^k$ e ciò conclude la dimostrazione per induzione. ∎

### Esercizio 14.5.2

*Vogliamo determinare tutte le soluzioni della congruenza $63x \equiv_{375} 24$.*

**Svolgimento**

Poiché $3 = \gcd(63, 375)$ divide 24, la congruenza ammette soluzioni in $\mathbb{Z}$. Abbiamo che $63 = 3 \cdot 21$, $375 = 3 \cdot 125$, $24 = 3 \cdot 8$, quindi risolviamo l'equazione

$[21]_{125}\, x = [8]_{125}$. Vogliamo trovare l'inverso di $[21]_{125}$. Per farlo, come prima cosa utilizziamo l'algoritmo di Euclide per calcolare $\gcd(125, 21)$.

$$\begin{aligned} 125 &= 5 \cdot 21 + 20 \\ 21 &= 1 \cdot 20 + 1 \\ 20 &= 20 \cdot 1 \end{aligned}$$

Procedendo a ritroso:

$$1 = 21 - 1 \cdot 20 = 21 - (125 - 5 \cdot 21) = -125 + 6 \cdot 21$$

Otteniamo quindi che:

$$[1]_{125} = [6]_{125}[21]_{125}$$

Quindi la soluzione dell'equazione $[21]_{125}\, x = [8]_{125}$ è $x = [6]_{125} \cdot [8]_{125} = [48]_{125}$. A questo punto le soluzioni della congruenza data sono: $x = 48 + k125$, con $k \in \mathbb{Z}$. ∎

## 14.6 ESERCIZI PROPOSTI

### Esercizio 14.6.1

*Si dimostri che* $1 + 2^2 + 3^2 + \ldots n^2 = n(n+1)(2n+1)/6$.

### Esercizio 14.6.2

*Si dimostri il Teorema fondamentale dell'aritmetica* (14.2.12) *utilizzando il principio di induzione completa.*

### Esercizio 14.6.3

*Si dimostri (per induzione) che, se $n$ è un intero non negativo, vale $2^n > n$.*

### Esercizio 14.6.4

*Si dica se esistono in $\mathbb{Z}_{37}$ due classi $[a]_{37}$, $[b]_{37}$, entrambe diverse da zero tali che $[a]_{37}[b]_{37} = [0]$. Se esistono si calcolino, se non esistono si motivi questo fatto. Si risponda alla stessa domanda sostituendo $\mathbb{Z}_{37}$ con $\mathbb{Z}_{36}$.*

### Esercizio 14.6.5

*Si considerino le due classi di congruenza: $[0]_6$ e $[3]_{12}$. Si dica se sono uguali o diverse tra loro, oppure se una è contenuta nell'altra.*

### Esercizio 14.6.6

*Si risolvano (se possibile) le seguenti equazioni:*

a) $[23]x = [7]$ *in* $\mathbb{Z}_{40}$

b) $[3]x = [13]$ *in* $\mathbb{Z}_{17}$

c) $[15]x = [9]$ *in* $\mathbb{Z}_{53}$

d) $[4]x = [1]$ *in* $\mathbb{Z}_{400}$

e) $[18]x = [30]$ *in* $\mathbb{Z}_{42}$

f) $[16]x = [36]$ *in* $\mathbb{Z}_{56}$

g) $[20]x = [56]$ *in* $\mathbb{Z}_{178}$

## 14.7 APPENDICE: NOZIONI ELEMENTARI DI INSIEMISTICA

Richiamiamo in questa appendice alcune nozioni elementari di insiemistica utilizzate nel testo.

Il concetto di *insieme* e di *appartenenza* sono *primitivi*, quindi non ne diamo una definizione rigorosa. Informalmente, un insieme è una collezione di oggetti e, per assegnare un insieme si possono elencare i suoi elementi, per esempio:

$$X = \{3, 4, 5, 6, 7\}$$

oppure si può assegnare una proprietà dei suoi elementi, per esempio, nel caso precedente, $X$ è l'insieme dei numeri naturali maggiori di 2 e minori di 8, e si può indicare come:

$$X = \{x \mid x \text{ è un numero naturale e } 2 < x < 8\}$$

Per indicare che un elemento appartiene a un insieme si utilizza il simbolo $\in$, la cui negazione è $\notin$. Per esempio nel caso precedente abbiamo $5 \in X$, $9 \notin X$.

Alcuni insiemi utilizzati spesso nel testo sono:

| | |
|---|---|
| $\mathbb{N} = \{0, 1, 2, \dots, n, n+1, \dots\}$ | l'insieme dei numeri naturali, |
| $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots, \pm n, \dots\}$ | l'insieme dei numeri interi, |
| $\mathbb{R}$ | l'insieme dei numeri reali. |

Due insiemi $X$ e $Y$ sono uguali se hanno gli stessi elementi e in tal caso scriviamo $X = Y$.

**Definizione 14.7.1** Se $X$ e $Y$ sono insiemi, diciamo che $X$ è un *sottoinsieme* di $Y$ se ogni elemento di $X$ è anche un elemento di $Y$, e si scrive con $X \subseteq Y$.

C'è poi un insieme speciale, l'*insieme vuoto*, che è l'insieme privo di elementi e viene indicato con $\varnothing$. Si noti che l'insieme vuoto è sottoinsieme di qualsiasi insieme $X$. Abbiamo per esempio:

$$\{x \mid x \in \mathbb{R},\ x^2 = -1\} = \varnothing$$

perché nessun numero reale elevato al quadrato dà come risultato $-1$. È necessario prestare attenzione, per esempio $X = \{0\}$ è il sottoinsieme dei numeri reali che contiene il solo elemento zero, tuttavia non è l'insieme vuoto in quanto contiene un elemento.

Richiamiamo ora le due operazioni fondamentali che si possono effettuare tra gli insiemi.

- L'*unione* di due insiemi $X$ e $Y$ è l'insieme di tutti gli elementi che appartengono a $X$ oppure a $Y$ e si indica con $X \cup Y$.
- L'*intersezione* di due insiemi $X$ e $Y$ è l'insieme di tutti gli elementi che appartengono sia a $X$ sia a $Y$ e si indica con $X \cap Y$.

# A Numeri complessi

In questa appendice vogliamo introdurre in modo elementare l'insieme dei numeri complessi, indispensabile per una comprensione più profonda delle questioni legate alle soluzioni di equazioni algebriche. Tutti i risultati di algebra lineare, riguardanti spazi vettoriali reali descritti in questo libro, restano validi se si sostituisce l'insieme dei numeri complessi a quello dei numeri reali, senza alcuna modifica della teoria. Tuttavia, poiché questo argomento, in generale, comporta una difficoltà aggiuntiva, abbiamo preferito esporre la nostra trattazione dell'algebra lineare limitandoci al caso di scalari reali e lasciando i numeri complessi in questa appendice.

## A.1 I NUMERI COMPLESSI

L'insieme $\mathbb{C}$ dei numeri complessi rappresenta un ampliamento dell'insieme dei reali, nato principalmente dall'esigenza di poter trovare la radice quadrata di un numero reale negativo. Vogliamo dunque introdurre un nuovo simbolo, denotato con $i$, detto *unità immaginaria*, con la proprietà che $i^2 = -1$. Definiamo pertanto l'insieme dei numeri complessi come

$$\mathbb{C} = \{a + bi | a, b \in \mathbb{R}\}$$

con due operazioni di somma e prodotto definite nel modo seguente:

- $(a + bi) + (c + di) = (a + c) + (b + d)i$,
- $(a + ib)(c + id) = (ac - bd) + (ad + bc)i$,

ove i numeri reali $a+c, b+d, ac-bd, ad+bc$ sono ottenuti con le usuali operazioni di somma e prodotto tra numeri reali. Per definizione, dati $a + bi, c + di \in \mathbb{C}$, abbiamo che $a + bi = c + di$ se e solo se $a = c$ e $b = d$.

Un modo semplice per ricordare l'operazione di prodotto tra due numeri complessi è quello di moltiplicarli come se fossero due polinomi in $i$ e di ricordare che $i^2 = -1$, quindi:

$$(a + ib)(c + id) = ac + adi + bci + bdi^2 = (ac - bd) + (ad + bc)i.$$

**Esempio A.1.1**

Dati i due numeri complessi $1 + 2i$, $3 - i$ ne vogliamo calcolare la somma ed il prodotto.

$$(1 + 2i) + (3 - i) = (1 + 3) + (2 - 1)i = 4 + i$$
$$(1 + 2i)(3 - i) = 3 + 6i - i - 2i^2 = 3 + (6 - 1)i - 2(-1) = 5 + 5i$$

L'insieme dei numeri reali è un sottoinsieme dei numeri complessi, perché possiamo scrivere ogni numero reale $a \in \mathbb{R}$ nella forma $a = a + 0i$. I numeri complessi del tipo $bi = 0 + bi$ si dicono *puramente immaginari* e se $z = a + bi$ è un numero complesso, i numeri reali $a$ e $b$ si dicono rispettivamente la *parte reale* e la *parte immaginaria* di $z$.

Possiamo rappresentare i numeri complessi utilizzando il piano cartesiano nel seguente modo: al numero complesso $a + bi$ associamo la coppia di reali $(a, b)$. In tale piano l'asse delle $x$ rappresenta i numeri reali ed è pertanto detto *asse reale*, mentre l'asse delle $y$ rappresenta gli immaginari puri e si dice *asse immaginario*.

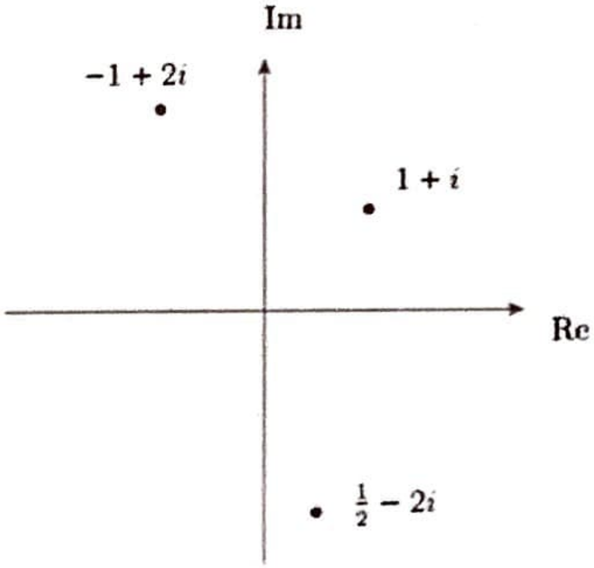



Figura A.1.

Dato un numero complesso $\alpha = a + bi$, definiamo il suo *complesso coniugato* (*o coniugato*) $\overline{\alpha}$ come $\overline{\alpha} = a - bi$. Definiamo inoltre il *modulo* di un numero complesso $\alpha$ come

$$|\alpha| = \sqrt{\alpha\overline{\alpha}} = \sqrt{a^2 + b^2}$$

È immediato vedere che, nella rappresentazione sul piano cartesiano, il modulo di $\alpha$ è la distanza del punto $P = (a, b)$, che rappresenta $\alpha$, dall'origine.

L'operazione di coniugio, che ad ogni numero complesso associa il suo complesso coniugato, soddisfa alcune proprietà, che lasciamo come facile esercizio:

- $\overline{(\overline{\alpha})} = \alpha$ per ogni $\alpha \in \mathbb{C}$.
- $\overline{\alpha} = \alpha$ se e solo se $\alpha \in \mathbb{R}$.

- $\overline{\alpha+\beta} = \overline{\alpha} + \overline{\beta}$ per ogni $\alpha, \beta \in \mathbb{C}$,
- $\overline{\alpha\beta} = \overline{\alpha}\overline{\beta}$ per ogni $\alpha, \beta \in \mathbb{C}$.

Una delle proprietà più importanti dei numeri complessi è che l'inverso $\alpha^{-1}$ di un numero complesso $\alpha \neq 0$ esiste sempre ed è ancora un numero complesso. Tale inverso si ottiene come $\alpha^{-1} = \overline{\alpha}/|\alpha|^2$. Più esplicitamente:

$$\alpha^{-1} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}i$$

Ciò ci permette di calcolare immediatamente il rapporto tra due numeri complessi. Anziché ricordare la formula, invitiamo lo studente a comprendere la procedura descritta nel seguente esempio.

**Esempio A.1.2**

Consideriamo il rapporto tra numeri complessi: $\frac{3-2i}{1-i}$. Vogliamo esprimere tale rapporto come $a + bi$ per opportuni $a$ e $b$. Procediamo moltiplicando numeratore e denominatore per il complesso coniugato del denominatore. Lo studente riconoscerà l'analogia con la procedura di razionalizzazione del denominatore di una frazione.

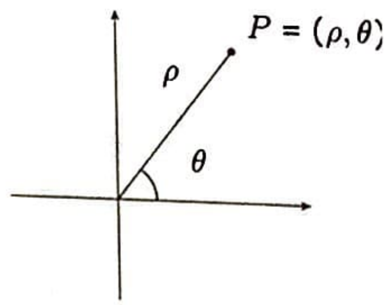
$$\frac{3-2i}{1-i} = \frac{3-2i}{1-i}\cdot\frac{1+i}{1+i} = \frac{(3-2i)(1+i)}{|1-i|^2} = \frac{3-2i+3i-2i^2}{2} = \frac{5}{2}+\frac{i}{2}$$

Concludiamo questa sezione con una lista delle proprietà delle operazioni nei numeri complessi, la cui verifica è lasciata al lettore come facile esercizio.

- *commutatività della somma*:
  $\alpha + \beta = \beta + \alpha$ per ogni $\alpha, \beta \in \mathbb{C}$
- *associatività della somma*:
  $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ per ogni $\alpha, \beta, \gamma \in \mathbb{C}$
- *esistenza dell'elemento neutro* 0 *per la somma*:
  $\alpha + 0 = \alpha$ per ogni $\alpha \in \mathbb{C}$, con $0 = 0 + 0i$
- *esistenza dell'opposto* $-\alpha$ di un numero complesso $\alpha$: $\alpha + (-\alpha) = 0$ per ogni $\alpha = a + bi \in \mathbb{C}$, con $-\alpha = -a - bi$
- *commutatività del prodotto*:
  $\alpha\beta = \beta\alpha$ per ogni $\alpha, \beta \in \mathbb{C}$
- *associatività del prodotto*:
  $(\alpha\beta)\gamma = \alpha(\beta\gamma)$ per ogni $\alpha, \beta, \gamma \in \mathbb{C}$
- *esistenza dell'elemento neutro* 1 *per il prodotto*:
  $1\alpha = \alpha$ per ogni $\alpha \in \mathbb{C}$, con $1 = 1 + 0i$
- *esistenza dell'inverso*: $\alpha\alpha^{-1} = 1$ per ogni $\alpha \neq 0$, $\alpha \in \mathbb{C}$
- *distributività della somma rispetto al prodotto*:
  $(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma$ per ogni $\alpha, \beta, \gamma \in \mathbb{C}$

## A.2 RAPPRESENTAZIONE POLARE

Vogliamo ora rappresentare un numero complesso nel piano cartesiano usando il sistema di coordinate *polari*. In tale sistema ogni punto del piano $P$ è individuato da due coordinate $(\rho, \theta)$, dette rispettivamente *coordinata radiale* e *coordinata angolare*. La coordinata $\rho$ rappresenta la distanza del punto $P$ dall'origine $O$, mentre la coordinata $\theta$ rappresenta l'angolo percorso in senso antiorario compreso tra l'asse $x$ e la semiretta $OP$:




Per definizione delle funzioni trigonometriche seno e coseno possiamo subito esprimere le coordinate cartesiane $(x, y)$ di un punto $P$ in termini delle sue coordinate polari $(\rho, \theta)$:

$$x = \rho \cos\theta \qquad y = \rho \sin\theta$$

Viceversa, date le coordinate cartesiane $(x, y)$ di $P$, grazie al Teorema di Pitagora e alla definizione di arcotangente, possiamo subito scrivere:

$$\rho = \sqrt{x^2 + y^2}, \qquad \theta = \arctan(y/x)$$

ove $\rho \geq 0$ e $\theta$ è determinato a meno di multipli di $2\pi$ (se $\rho \neq 0$).

Se ora consideriamo un numero complesso $\alpha = a + ib$ e la sua rappresentazione nel piano cartesiano, vediamo subito che in coordinate polari abbiamo $a = \rho\cos\theta$ e $b = \rho\sin\theta$.

Quindi possiamo anche scrivere

$$\alpha = \rho(\cos\theta + \sin\theta i)$$

Tale espressione viene detta *rappresentazione trigonometrica o polare* del numero complesso $\alpha = a + bi$. Come abbiamo visto $\rho$ è il modulo di $\alpha$, mentre l'angolo $\theta$ viene detto *argomento* di $\alpha$.

Se utilizziamo la rappresentazione trigonometrica, la formula per il prodotto di due numeri complessi $\alpha = \rho_1(\cos\theta_1 + \sin\theta_1\, i)$ e $\beta = \rho_2(\cos\theta_2 + \sin\theta_2\, i)$ diventa particolarmente elegante. Abbiamo infatti:

$$\begin{aligned}\alpha\beta &= \rho_1\rho_2(\cos\theta_1 + \sin\theta_1\, i)(\cos\theta_2 + \sin\theta_2\, i) \\ &= \rho_1\rho_2\big((\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + (\cos\theta_1\sin\theta_2 + \sin\theta_1\cos\theta_2)i\big) \qquad \text{(A.1)} \\ &= \rho_1\rho_2\big(\cos(\theta_1 + \theta_2) + \sin(\theta_1 + \theta_2)i\big)\end{aligned}$$

Quindi, per due numeri complessi abbiamo che: il modulo del prodotto è il prodotto dei moduli e l'argomento del prodotto è la somma degli argomenti.

La forma trigonometrica di un numero complesso ci permette di calcolare abbastanza rapidamente le sue radici $n$-esime, attraverso la *formula di De Moivre*. Grazie alla formula (A.1) possiamo calcolare le potenze di un numero complesso:

$$\begin{aligned} \alpha &= \rho(\cos\theta + i\sin\theta) \\ \alpha^2 &= \rho^2(\cos 2\theta + \sin 2\theta\, i) \\ \alpha^3 &= \rho^3(\cos 3\theta + \sin 3\theta\, i) \\ \vdots &= \\ \alpha^n &= \rho^n(\cos n\theta + \sin n\theta\, i) \end{aligned} \tag{A.2}$$

L'ultima uguaglianza prende il nome di *formula di De Moivre.*

Pertanto possiamo subito determinare le radici $n$-esime di un numero complesso $\alpha = \rho(\cos\theta + \sin\theta\, i)$:

$$\alpha^{1/n} = \rho\Big(\cos[(\theta + 2k\pi)/n] + \sin[(\theta + 2k\pi)/n]\, i\Big) \tag{A.3}$$

con $k = 0, \ldots, n-1$, in altre parole abbiamo $n$ radici $n$-esime complesse di un dato numero complesso non nullo.

Vediamo un esempio.

**Esempio A.2.1**

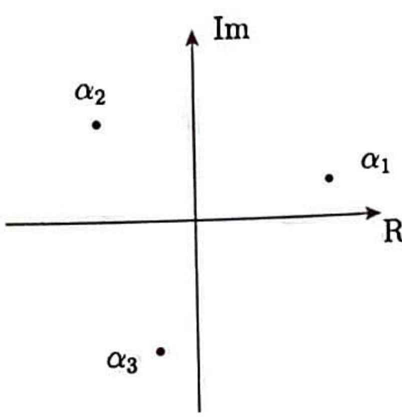
Vogliamo determinare tutte le radici cubiche di $1 + i$. Secondo la formula (A.3) sono date da:

$$\sqrt{2}\Big\{\cos[(\pi/4 + 2k\pi)/3] + \sin[(\pi/4 + 2k\pi)/3]\, i\Big\}$$

con $k = 0, 1, 2$. Più precisamente:

$$\begin{aligned} \alpha_1 &= \sqrt{2}\Big\{\cos[(\pi/4)/3] + \sin[(\pi/4)/3] i\Big\} = \sqrt{2}(\cos\tfrac{\pi}{12} + \sin\tfrac{\pi}{12}\, i\Big) \\ \alpha_2 &= \sqrt{2}\Big\{\cos[(\pi/4 + 2\pi)/3] + \sin[(\pi/4 + 2\pi)/3] i\Big\} = \\ &= \sqrt{2}\Big\{\cos\tfrac{3\pi}{4} + \sin\tfrac{3\pi}{4}\, i\Big\} \\ \alpha_3 &= \sqrt{2}\Big\{\cos[(\pi/4 + 4\pi)/3] + \sin[(\pi/4 + 4\pi)/3] i\Big\} = \\ &= \sqrt{2}\Big\{\cos\tfrac{17\pi}{12} + \sin\tfrac{17\pi}{12}\, i\Big\} \end{aligned}$$

Possiamo rappresentare tali radici nel piano complesso come segue:



Concludiamo questa sezione enunciando un risultato molto importante: il *Teorema fondamentale dell'algebra*, la cui dimostrazione è particolarmente difficile e poiché esula dagli scopi di questo libro rimandiamo il lettore ad uno dei numerosi testi specifici (ad esempio *S. Lang "Algebra"*).

**Teorema A.2.2** *Ogni polinomio di grado n a coefficienti complessi*

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0, \quad a_n, \ldots, a_o \in \mathbb{C}$$

*si fattorizza in un prodotto di n fattori lineari (non necessariamente distinti):*

$$p(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n),$$

*con* $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$.

Questo implica in particolare che una equazione polinomiale di grado $n$ a coefficienti in $\mathbb{C}$ ha sempre $n$ soluzioni complesse, anche se non necessariamente distinte.

Vediamo un esempio.

---

**Esempio A.2.3**

Vogliamo determinare tutte le soluzioni dell'equazione $x^4 - 16 = 0$. Possiamo subito fattorizzare il polinomio come:

$$x^4 - 16 = (x^2 - 4)(x^2 + 4) = (x - 2)(x + 2)(x + 2i)(x - 2i)$$

Pertanto gli zeri del polinomio, corrispondenti alle soluzioni dell'equazione data, sono: $\pm 2$, $\pm 2i$. Possiamo giungere allo stesso risultato anche applicando la formula (A.3):

$$\begin{aligned} 2 &= 2[\cos(0) + \sin(0)\, i], \\ 2i &= 2[\cos(2\pi/4) + \sin(2\pi/4)\, i], \\ -2 &= 2[\cos(4\pi/4) + \sin(4\pi/4)\, i], \\ -2i &= 2[\cos(6\pi/4) + \sin(6\pi/4)\, i] \end{aligned}$$

ricordando che in coordinate polari: $-1 = \cos(\pi) + \sin(\pi) i$ e che $-i = \cos(6\pi/4) + \sin(6\pi/4) i$.

---

# B Soluzioni di alcuni esercizi proposti

## Capitolo 1. Introduzione ai sistemi lineari

**1.6.1** a) $x = y = 0,\ z = 1$.
b) Il sistema non ammette soluzione.

**1.6.3** Il sistema ammette un'unica soluzione per $k \neq 0, 1$. Per $k = 0$ il sistema ammette infinite soluzioni dipendenti da un parametro, per $k = 1$ il sistema non ammette soluzioni.

## Capitolo 2. Spazi vettoriali

**2.6.1** a), b), d), e), f), l) sono sottospazi. c), g), h), i), m) non sono sottospazi.

**2.6.4** $X$ non è sottospazio vettoriale.

## Capitolo 3. Combinazioni lineari e lineare indipendenza

**3.4.1** a), d), e) sono insiemi di vettori linearmente indipendenti. b), c) sono insiemi di vettori dipendenti.

**3.4.3** Sì.

**3.4.4** $k = \pm\sqrt{3}$.

**3.4.5** $k \neq -1/2$.

**3.4.6** $k = -2/5$ e $k = 0$.

**3.4.9** a) $k \neq 2, -1$. b) $k \neq 2$.

**3.4.10** a) Il sistema ammette un'unica soluzione per $k \neq 0, 1$, non ha soluzioni per $k = 0, 1$. b) $k \neq 0, 1$.

**3.4.12** I vettori dati sono sempre linearmente dipendenti.

**3.4.13** a) $k \neq 0, 5/3$. Per $k = 0$ si ha che $\mathbf{v}_2 = \mathbf{v}_3$.

**3.4.14** a) $k \neq 0, -2$. b) $k \neq 0, -2$.

**3.4.15** $k \neq 3$.

**3.4.16** a) $\{(-2, 1, 0), (0, 0, 1)\}$.

**Capitolo 4. Basi e dimensione**

**4.5.2** $k \neq 2$.

**4.5.4** a) è base, b) è un insieme di generatori, c) non è un insieme di generatori.

**4.5.6** $k \neq -1$

**4.5.7** I vettori sono linearmente indipendenti per $k \neq 0$ e $k \neq 1$. Per $k = 1$, $(1, -2, -6) \in \langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle$.

**4.5.8** Per ogni $k \in \mathbb{R}$.

**4.5.10** $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ generano $\mathbb{R}^3$ per $k \neq \frac{3}{4}$; per $k = 0$ il vettore $(4, -1, 6)$ appartiene a $\langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle$.

**4.5.11** $k = -5, k = 2$.

**4.5.12** +- $k \neq \pm 6$.

**4.5.13** $k \neq -\frac{3}{2}$.

**4.5.17** a) $k \neq 0, k \neq 1/10$.

**Capitolo 5. Applicazioni lineari**

**5.9.1** Per ogni $k \in \mathbb{R}$.

**5.9.2** $k = 0$.

**5.9.6** a) $F$ è isomorfismo. b), c) le applicazioni non sono isomorfismi.

**5.9.13** $k \neq -2$.

**5.9.14** Una base di $\operatorname{Ker}(T)$ è $\{(-2, 0, 1)\}$ e una base di $\operatorname{Im}(T)$ è $\{(1, 0, 2, 3), (0, 1, 3, -1)\}$, $\dim(\operatorname{Ker} T) = 1$, $\dim(\operatorname{Im} T) = 2$ e $\dim(\mathbb{R}^3) = 3 = 1 + 2 = \dim(\operatorname{Ker} T) + \dim(\operatorname{Im} T)$, pertanto $T$ non è iniettiva.

**5.9.19** $k = \pm 2$.

**5.9.20** $k \neq -3$.

**5.9.21** $F$ è iniettiva per ogni $k \in \mathbb{R}$.

**Capitolo 6. Sistemi lineari**

**6.4.2** $k \neq -5$.

**6.4.4** $S$ è sottospazio.

**6.4.5** $F$ esiste ed è unica.

## Capitolo 7. Determinante e inversa

**7.8.1** a) $F$ è isomorfismo, perché il determinante della matrice associata ad $F$ nella base canonica è uguale a 1.

**7.8.2** $G(x, y, z) = (x + y - z, -y - z, -z)$.

**7.8.3** La matrice è invertibile per $a \neq 3/2$.

**7.8.7** a) L'inversa è:

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

## Capitolo 8. Cambio di base

**8.5.1** $(4, 6, 9)$.

**8.5.5** a) $k \neq \pm 1$, $\ker(T) = \langle (2/3, 0, 1) \rangle$. b) $\begin{pmatrix} 3/2 & 1/2 & -1 \\ 6 & 1 & -4 \end{pmatrix}$

**8.5.6** a) $k \neq 4$, $k \neq -1/2$.
b) $G$ corrisponde, fissate le basi canoniche in dominio e codominio, alla matrice:

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

## Capitolo 9. Autovalori e autovettori

**9.4.1** a) Autovalori 2, 3, la matrice non è diagonalizzabile. b) Autovalori 1, 4, $L$ è diagonalizzabile. c) Autovalori $\pm 1$, 2, $L$ è diagonalizzabile. d) Autovalori 0, 7, $L$ è diagonalizzabile. e) $L$ non ha autovalori reali.

**9.4.2** a) Autovalori 0, 5, 7, $A$ è diagonalizzabile.

**9.4.4** Autovalori 2, 4, $A$ è diagonalizzabile.

**9.4.5** Autovalori $-1$, 2, $A$ non è diagonalizzabile.

**9.4.6** Autovalori 9, 3, $A$ è diagonalizzabile.

**9.4.7** Autovalori 0, 6, $-1$, $A$ è diagonalizzabile.

**9.4.8** Autovalori 2, 5, $T$ non è diagonalizzabile.

**9.4.9** Autovalori $-2$, 3, $T$ è diagonalizzabile.

**9.4.10** a) $k = -1$. b) Per ogni $k$.

**9.4.11** $A$ è diagonalizzabile per $k = 0$.

**9.4.12** $F$ è diagonalizzabile per $k \neq 2$.

**9.4.13** a) $A$ è diagonalizzabile per $k \neq 5$. b) Per $k = 0$.

## Capitolo 10. Prodotti scalari

**10.8.2** a) $(1,1,2,-1)$. b) Data la base $\{(2,-1,0,1),\ (2,0,-1,0),\ (1,0,0,1)\}$ di $W$, per l'algoritmo di Gram-Schmidt otteniamo la base ortonormale: $\{(\sqrt{2/3}, -1/\sqrt{6}, 0, 1/\sqrt{6}), (2/\sqrt{21}, 2/\sqrt{21}, -\sqrt{3/7}, -2/\sqrt{21}), (0, 1/\sqrt{2}, 0, 1/\sqrt{2})\}$.

**10.8.3** a) Una base per $W^\perp$ è: $\{2\mathbf{e}_2 + \mathbf{e}_3, -\mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_4\}$.

b) Una base ortogonale per $W$ è: $\{\mathbf{e}_1+\mathbf{e}_4, -(1/2)\mathbf{e}_1+\mathbf{e}_2-2\mathbf{e}_3+(1/2)\mathbf{e}_4\}$.

**10.8.4** Data la base: $\left\{(0,-1,1,0),\ (\frac{1}{2},-\frac{1}{2},0,1)\right\}$ di $W$, per l'algoritmo di Gram-Schmidt otteniamo la base ortonormale:

$$\left(0,-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},0\right) \qquad \left(\frac{2}{\sqrt{22}},-\frac{1}{\sqrt{22}},-\frac{1}{\sqrt{22}},\frac{4}{\sqrt{22}}\right)$$

**10.8.7** a) $W^\perp = \langle(0,1,-1,1)\rangle$. b) Una base ortogonale (ortonormale) per $W$ è: $\{(1/\sqrt{3}, -1/\sqrt{3}, 0, 1/\sqrt{3}), (\sqrt{2/3}\}, 1/\sqrt{6}, 0, -1/\sqrt{6}), (0, 1/\sqrt{6}, \sqrt{2/3}, 1/\sqrt{6})$. Una base ortogonale (ortonormale) per $W^\perp$ è: $\{(0, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})\}$.
c) No.

## Capitolo 11. Il Teorema spettrale

**11.6.1** a) No, b) no, c) sì, d) sì.

**11.6.4** Gli autovalori di

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

sono $1/2(3 \pm \sqrt{5})$, mentre gli autovettori sono $\mathbf{v}_1 = (1/2(-1+\sqrt{5}),1)$ di norma $c_1 = \sqrt{(1/2)(5-\sqrt{5})}$ e $\mathbf{v}_2 = (1/2(-1-\sqrt{5}),1)$ di norma $c_2 = \sqrt{(1/2)(5+\sqrt{5})}$. Dunque la matrice $A$ viene diagonalizzata dalla trasformazione ortogonale $P$:

$$\begin{pmatrix} \frac{1}{2c_1}(-1+\sqrt{5}) & \frac{1}{2c_2}(-1-\sqrt{5}) \\ \frac{1}{c_1} & \frac{1}{c_2} \end{pmatrix}$$

Gli autovalori e autovettori di

$$A' = \begin{pmatrix} 3 & 2 & 0 \\ 2 & 6 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

sono $\lambda_1 = 7$, $\lambda_2 = 3$, $\lambda_3 = 2$, $\mathbf{v}_1 = (1,2,0)$, $\mathbf{v}_2 = (0,0,1)$, $\mathbf{v}_3 = (-2,1,0)$. Dunque la matrice $A'$ viene diagonalizzata dalla trasformazione ortogonale $P'$:

$$\begin{pmatrix} 7 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix} = (P')^{-1}AP', \qquad P' = \begin{pmatrix} 1/\sqrt{5} & 0 & -2\sqrt{5} \\ 2/\sqrt{5} & 0 & 1/\sqrt{5} \\ 0 & 1 & 0 \end{pmatrix}$$

## Capitolo 12. Applicazioni: Teorema spettrale e forme quadratiche

**12.5.1** Gli autovalori sono 7 e 3 e pertanto la curva considerata è un'ellisse. Gli autovettori sono $\mathbf{v}_1 = (-1,1)$, $\mathbf{v}_2 = (1,1)$, che rappresentano le direzioni dei nuovi assi cartesiani.

**12.5.3** b) Gli autovalori sono 8, $-1$, 1, dunque il prodotto scalare è non degenere, ma non definito positivo. c) La segnatura della forma quadratica è $(2,1)$. d) Una base rispetto alla quale il prodotto è diagonale è data da:

$$\mathbf{v}_1 = (2,1,2), \quad \mathbf{v}_2 = (1,-4,1), \quad \mathbf{v}_3 = (-1,0,1)$$

**12.5.4** Gli autovalori sono $\lambda_1 = 1/2(1+\sqrt{26})$, $\lambda_2 = 1/2(1-\sqrt{26})$ e dunque il prodotto scalare è non degenere, ma non definito positivo. La curva è un'iperbole. Gli assi hanno direzione data dagli autovettori:

$$\mathbf{v}_1 = (1/5(1+\sqrt{26}),1), \qquad \mathbf{v}_2 = (1/5(1-\sqrt{26}),1)$$

**12.5.5** Gli autovalori sono $-10,5$, pertanto si tratta di un'iperbole. La forma canonica è $-10(x')^2 + 5(y')^2 = 16$ ove gli assi cartesiani $x'$ e $y'$ hanno come direzione gli autovettori $(2,1)$, $(-1,2)$.

**12.5.8** Gli autovalori sono $\lambda_1 = 9$ e $\lambda_2 = -7$ e gli autovettori: $\mathbf{v}_1 = (-1,1)$, $\mathbf{v}_2 = (1,1)$. Il prodotto scalare è non degenere ma non definito positivo. La forma canonica è $-7(x')^2 + 9(y')^2$.

## Capitolo 13. Rette e piani in $\mathbb{R}^3$

**13.6.1** $r_1$, $r_2$: sono sghembe. $r_1$, $r_3$: sono sghembe. $r_1$, $r_4$: coincidono. $r_2$, $r_3$: sono parallele. $r_2$, $r_4$: sono sghembe. $r_3$, $r_4$: sono sghembe. La distanza tra $r_2$ e $r_3$ è $2\sqrt{2}$.

**13.6.6** Equazioni parametriche della retta $r$: $x = t$, $y = t+1$, $z = -2t/3 - 1/3$. Equazione cartesiana del piano $\pi$: $x - 3y - 3z = -2$. Equazioni parametriche della retta $s$: $x = 1+t$, $y = -3t$, $z = 1-3t$.

**13.6.7** Equazioni parametriche per $r$: $x = t$, $y = 4$, $z = t/2 + 6$. Retta e piano sono paralleli. La distanza è $2\sqrt{6}$.

## Capitolo 14. Elementi di matematica discreta

**14.6.4** Tali classi non esistono in $\mathbb{Z}_{37}$ perché 37 è primo. In $\mathbb{Z}_{36}$ invece abbiamo $[4]_{36}[9]_{36} = [0]_{36}$.

**14.6.5** Sono diverse e nessuna delle due è contenuta nell'altra.

**14.6.6** b) [10], d) Impossibile.

# Indice analitico

## M



## N



## O



## P



## R



## S



## T



## V



## Z