Appunti di Algebra Lineare

Francesco Mazzocca

anno accademico 2006/07

Indice

Notazioni

Riportiamo l'elenco di alcuni simboli e notazioni di cui faremo uso nel corso di queste note.

- $''\forall''$, (*quantificatore universale*) che si legge "*per ogni*", o "*qualunque sia*".
- $''\exists''$, (*quantificatore esistenziale*) che si legge "*esiste almeno un*".
- $''\Rightarrow''$, (*implicazione logica*) che si legge "*implica*".
- $''\Leftrightarrow''$, (*equivalenza logica*) che si legge "*equivale*".
- $''\,|\,''$ e $''\,:\,''$ che si leggono "*tale che*".
- $a \in A$ indica che a è un elemento dell'insieme A;
- $a \notin A$ indica che a non è un elemento dell'insieme A;
- $\emptyset :=$ insieme vuoto [1].

Se A e B sono insiemi, poniamo

- $A \subseteq B$, $B \supseteq A \Leftrightarrow A$ è *sottoinsieme* di B;
- $A \subset B$, $B \supset A \Leftrightarrow A$ è sottoinsieme di B e non è uguale a B (*sottoinsieme proprio*);
- $A \cup B := \{x \ : \ x \in A \ \ o \ \ x \in B\}$ (*unione di* A e B);
- $A \cap B := \{x \ : \ x \in A \ \ e \ \ x \in B\}$ (*intersezione di* A e B);
- $A \setminus B := \{x \ : \ x \in A \ \ e \ \ x \notin B\}$ (*differenza fra* A e B);
- $A \times B := \{(a, b) \ : \ a \in A, \, b \in B\}$ (*prodotto cartesiano di* A e B).
- $A^n := \{(a_1, a_2, \ldots, a_n) \ : \ a_1, a_2, \ldots, a_n \in A\}$ (*prodotto cartesiano di* A *per se stesso* n *volte*).

Poniamo, inoltre,

[1] La notazione := indica che il simbolo scritto alla sua sinistra è definito da ciò che è scritto alla sua destra.

- $\mathbb{N} :=$ insieme dei numeri naturali (compreso lo zero);
- $\mathbb{N}^+ :=$ insieme dei numeri naturali diversi da zero;
- $\mathbb{Z} :=$ insieme dei numeri interi (relativi);
- $\mathbb{Q} :=$ insieme dei numeri razionali;
- $\mathbb{R} :=$ insieme dei numeri reali;
- $\mathbb{C} :=$ insieme dei numeri complessi;
- $A^* := A \setminus \{0\}$, $A = \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$;
- $A[x] :=$ insieme dei polinomi nell'indeterminata x a coefficienti in A, $A = \mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$;
- $P(S) :=$ insieme delle parti di un insieme non vuoto S;
- $Perm(S) :=$ insieme delle permutazioni[2] su un insieme non vuoto S.

Siano A, B, C insiemi. La scrittura

$$A \to B$$

indica che è assegnata una *funzione,* o *applicazione,* tra A e B. In queste ipotesi, A si dice *insieme di definizione,* o *dominio,* e B *codominio* della funzione. Per indicare che tra A e B è assegnata una funzione f usiamo una delle notazioni

$$f \ : \ A \to B \qquad , \qquad A \xrightarrow{f} B \qquad , \qquad x \in A \to f(x) \in B,$$

Il simbolo $f(x)$ denota il *corrispondente,* o *immagine,* dell'elemento x nella funzione f.

Assegnati una funzione $f : A \to B$, un sottoinsieme X di A ed uno Y di B, il sottoinsieme $f(X)$ di B definito da

$$f(X) := \{y \in B \ : \ y = f(x), \text{ per qualche } x \in X\}$$

si chiama *immagine di X in f*, mentre il sottoinsieme $f^{-1}(Y)$ di A definito da

$$f^{-1}(Y) := \{x \in A \ : \ f(x) \in Y\}$$

si chiama *controimmagine di Y in f.*

Assegnate le funzioni $f : A \to B$ e $g : B \to C$, il simbolo $g \circ f$ indica la *funzione composta,* o *composizione,* di f e g, cioè

$$g \circ f \ : x \in A \ \to \ g(f(x)) \in C.$$

Nel seguito porremo quasi sempre $fg := g \circ f$, cioè

$$(fg)(x) = g(f(x)) \,, \quad \text{per ogni} \ \ x \in A \,.$$

[2]Una *permutazione* su S è una funzione biunivoca di S su S.

Capitolo 1

Strutture Algebriche

La lettura di questo capitolo, dedicato ad una breve introduzione al linguaggio dell'algebra, è consigliata ma non è essenziale per la comprensione degli argomenti successivi. Il Lettore che voglia affrontare subito lo studio dell'algebra lineare può iniziare dal secondo capitolo.

1.1 Operazioni su un insieme

Alcuni esempi di *operazioni* note al Lettore sono i seguenti:

- Addizione e moltiplicazione in $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$.
- Addizione e moltiplicazione nell'insieme dei polinomi a coefficienti in $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$.
- Massimo comune divisore e minimo comune multiplo in $\mathbb{N}$.
- Unione e intersezione in $P(S)$.
- Moltiplicazione di un numero reale per un segmento orientato dello spazio ordinario.

Osservazione 1.1.1. Tutte le operazioni elencate si presentano come delle leggi che permettono di individuare un elemento di un insieme a partire da una coppia di elementi assegnati. È pertanto ragionevole pensare che la nozione di *operazione* possa generalizzarsi ed esprimersi in termini precisi nel linguaggio della teoria degli insiemi. È appunto ciò che ci proponiamo di fare in questo paragrafo. □

Nel seguito del capitolo S denoterà sempre un insieme non vuoto.

Un'applicazione di $S \times S$ in S prende il nome di *operazione interna ad* S. Assegnata un'operazione $\star : S \times S \to S$, si pone

$$\star(a, b) = a \star b\,;$$

per ogni $a, b \in S$. L'elemento $a \star b$ di S si chiama *composto* di a e b mediante l'operazione "$\star$". Le notazioni più usate per le operazioni sono

- la notazione *additiva:* si usa il segno $+$, il composto di due elementi a, b si denota con $a + b$ e si chiama *somma* di a e b;

- la notazione *moltiplicativa:* si usa il segno $\cdot$ o $\times$, il composto di due elementi a, b si denota con $a \cdot b$ o $a \times b$, o anche con ab, e si chiama *prodotto* di a e b;

- la notazione *esponenziale:* il composto di due elementi a, b si denota con a^b.

Sia A un insieme non vuoto, $A \neq S$. Un'applicazione di $A \times S$ in S prende il nome di *operazione esterna ad S con dominio di operatori A.*

Per le operazioni esterne si usano la simbologia e la terminologia introdotte per quelle interne. Un esempio di operazione esterna familiare al Lettore dovrebbe essere, per esempio, la *moltiplicazione di un numero reale per un segmento orientato.*

Si chiama *struttura algebrica ad n operazioni su S* ogni $(n+1)-pla$ del tipo

$$(S, \star_1, \star_2, \dots, \star_n),$$

ove ogni $\star_i$ è una operazione su S (interna o esterna). L'insieme S si chiama *sostegno* della struttura. Una struttura algebrica si dice *finita* (risp. *infinita*) se il suo sostegno è un insieme finito (risp. infinito[1]). Se una struttura algebrica è finita, il numero di elementi del suo sostegno si chiama *ordine* della struttura; nel caso contrario si dice che la struttura ha *ordine infinito.*

Esempi 1.1.2. Di seguito riportiamo alcuni esempi di strutture algebriche. Questi, tra l'altro, mostrano che uno stesso insieme S può essere sostegno di strutture algebriche diverse.

- $(\mathbb{N}, +), (\mathbb{Z}, +), (\mathbb{Q}, +), (\mathbb{R}, +), (\mathbb{C}, +)$.
- $(\mathbb{N}, \cdot)$, $(\mathbb{Z}, \cdot)$, $(\mathbb{Q}, \cdot)$, $(\mathbb{R}, \cdot)$, $(\mathbb{C}, \cdot)$.
- $(\mathbb{N}, +, \cdot)$, $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$, $(\mathbb{C}, +, \cdot)$.
- $(\mathbb{Z}[x], +, \cdot)$, $(\mathbb{Q}[x], +, \cdot)$, $(\mathbb{R}[x], +, \cdot)$, $(\mathbb{C}[x], +, \cdot)$.
- $(P(S), \cap)$, $(P(S), \cup)$, $(P(S), \cup, \cap)$. $\square$

Indicate con $\circ$ e $\star$ operazioni interne ad S, si definiscono le seguenti proprietà:

[1]Un insieme è infinito se può essere messo in corrispondenza biunivoca con un suo sottoinsieme proprio.

- Proprietà *commutativa*:

$$a \circ b = b \circ a\,, \quad \textit{per ogni } a, b \in S.$$

- Proprietà *associativa*:

$$a \circ (b \circ c) = (a \circ b) \circ c\,, \quad \textit{per ogni} \quad a, b, c \in S.$$

- Proprietà *distributiva (a destra)* di $\circ$ rispetto a $\star$:

$$(a \star b) \circ c = (a \circ c) \star (b \circ c)\,, \quad \textit{per ogni } a, b, c \in S.$$

Esercizio 1.1.3. *Verificare che su un insieme con un solo elemento si può definire un'unica operazione e che questa risulta commutativa e associativa.*

Sia $(S, \circ)$ una struttura algebrica con operazione interna. Un elemento $u \in S$ si dice *neutro* se

$$u \circ a = a \circ u = a \ \ , \text{ per ogni } a \in S.$$

Se la struttura $(S, \circ)$ possiede un elemento neutro si dice *unitaria*.

Proposizione 1.1.4. *Se* $(S, \circ)$ *possiede un elemento neutro, questo è unico.*

Dimostrazione. Se u e v sono elementi neutri, dalla definizione si ha

$$u = u \circ v = v\,,$$

cioè l'asserto. □

In notazione moltiplicativa l'elemento neutro si chiama *unità* e si denota con 1. In notazione additiva l'elemento neutro si chiama *zero* e si denota con 0.

Esempi 1.1.5.
- $(A, +)$, $A = \mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$, ha 0 come elemento neutro.

- $(A^*, \cdot)$, $A = \mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$, ha 1 come elemento neutro.

- Sia / l'operazione di divisione in $\mathbb{Q}^*$. $(\mathbb{Q}^*, /)$ (struttura non associativa e non commutativa) non possiede elemento neutro.

- $(P(S), \cap)$ ha S come l'elemento neutro.

- $(P(S), \cup)$ ha $\emptyset$ come l'elemento neutro. □

Sia $(S, \circ)$ una struttura algebrica dotata di elemento neutro u. Un elemento $a \in S$ si dice *simmetrizzabile* se esiste $a' \in S$ tale che

$$a' \circ a = a \circ a' = u \ \ .$$

L'elemento a' si chiama *simmetrico* di a.

Osservazione 1.1.6. In una una struttura algebrica $(S, \circ)$ dotata di elemento neutro u si ha:

- a' simmetrico di $a \Leftrightarrow a$ simmetrico di a'.
- L'elemento neutro u è simmetrizzabile e si ha $u' = u$. □

Nella notazione moltiplicativa il simmetrico di un elemento a si chiama *inverso* di a e si denota con a^{-1}. Nella notazione additiva il simmetrico di a si chiama *opposto* di a e si denota con $-a$.

Esempi 1.1.7.

- In $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$ tutti gli elementi sono simmetrizzabili.
- In $(\mathbb{Z}, \cdot)$, 1 e -1 sono gli unici elementi simmetrizzabili.
- In $(\mathbb{Q}^*, \cdot)$, $(\mathbb{R}^*, \cdot)$, $(\mathbb{C}^*, \cdot)$ tutti gli elementi sono simmetrizzabili.
- In $(\mathbb{Z}[x], +)$, $(\mathbb{Q}[x], +)$, $(\mathbb{R}[x], +)$, $(\mathbb{C}[x], +)$ tutti gli elementi sono simmetrizzabili.
- In $(\mathbb{Z}[x], \cdot)$, 1 e -1 sono gli unici elementi simmetrizzabili.
- In $(\mathbb{Q}[x], \cdot)$, $(\mathbb{R}[x], \cdot)$, $(\mathbb{C}[x], \cdot)$, gli elementi simmetrizzabili sono tutte e sole le costanti diverse da zero. □

1.2 Semigruppi

Una struttura algebrica $(S, \circ)$, con operazione interna, si chiama *semigruppo* se l'operazione $\circ$ è associativa. Se $\circ$ è anche commutativa, il semigruppo si dice *commutativo* o *abeliano*.

Esempi 1.2.1. Le strutture sottoelencate sono esempi di semigruppi.

- $(\mathbb{N}, +), (\mathbb{Z}, +), (\mathbb{Q}, +), (\mathbb{R}, +), (\mathbb{C}, +)$.
- $(\mathbb{N}, \cdot), (\mathbb{Z}, \cdot), (\mathbb{Q}, \cdot), (\mathbb{R}, \cdot), (\mathbb{C}, \cdot)$.
- $(\mathbb{Z}[x], +)$, $(\mathbb{Q}[x], +)$, $(\mathbb{R}[x], +)$, $(\mathbb{C}[x], +)$.
- $(\mathbb{Z}[x], \cdot)$, $(\mathbb{Q}[x], \cdot)$, $(\mathbb{R}[x], \cdot)$, $(\mathbb{C}[x], \cdot)$.
- $(P(S), \cap)$, $(P(S), \cup)$. □

Proposizione 1.2.2. *In un semigruppo $(S, \circ)$ con elemento neutro valgono le seguenti proprietà:*

- *a simmetrizzabile $\Rightarrow$ a ha un unico simmetrico a';*
- *a simmetrizzabile e tale che $a \circ b = b \circ a \Rightarrow a' \circ b = b \circ a'$.*

Dimostrazione. Siano u l'elemento neutro di $(S, \circ)$ e a', a'' due simmetrici dell'elemento a, allora

$$a' = a' \circ u = a' \circ (a \circ a'') = (a \circ a') \circ a'' = u \circ a'' = a''.$$

Per la seconda proprietà abbiamo

$$a' \circ b = a' \circ b \circ u = a' \circ b \circ (a \circ a') = a' \circ (b \circ a) \circ a' =$$

$$a' \circ (a \circ b) \circ a' = (a' \circ a) \circ b \circ a' = u \circ b \circ a' = b \circ a'$$

e l'asserto è completamente provato. □

Proposizione 1.2.3. *Siano a, b elementi simmetrizzabili di un semigruppo unitario $(S, \circ)$. Allora $a \circ b$ è simmetrizzabile e*

$$(a \circ b)' = b' \circ a'; \tag{1.1}$$

inoltre, $(a \circ b)' = a' \circ b'$ se, e solo se, risulta $a \circ b = b \circ a$.

Dimostrazione. Detto u l'elemento neutro di $(S, \circ)$, risulta

$$(a \circ b) \circ (b' \circ a') = a \circ (b \circ b') \circ a' = a \circ u \circ a' = a \circ a' = u.$$

La seconda parte della proposizione è ovvia. □

Esercizio 1.2.4. *Sull'insieme $\mathbb{R}^3 = \{(a, b, c) \ : \ a, b, c \in \mathbb{R}\}$ si consideri l'operazione "$\circ$" definita da*

$$(a, b, c) \circ (a', b', c') = (aa', (a + b + c)b' + ba', (a + b + c)c' + ca').$$

Provare che $(\mathbb{R}^3, \circ)$ è un semigruppo unitario e che un elemento (a, b, c) è invertibile se, e solo se, a e $a + b + c$ sono entrambi diversi da zero.

Sia $(S, \circ)$ una struttura algebrica con operazione interna. Un elemento a si dice *cancellabile a sinistra* (*a destra*) se:

$$a \circ b = a \circ c \Rightarrow b = c \quad (b \circ a = c \circ a \Rightarrow b = c).$$

L'elemento a si dice *cancellabile* o *regolare* se è cancellabile a sinistra e a destra. Se tutti gli elementi di S sono regolari, si dice che $(S, \circ)$ è *regolare* o anche che in $(S, \circ)$ vale la *legge di cancellazione* .

Proposizione 1.2.5. *Sia $(S, \circ)$ un semigruppo unitario. Allora ogni suo elemento simmetrizzabile è regolare.*

Dimostrazione. Siano u l'unità del semigruppo, a un elemento simmetrizzabile e a' il suo simmetrico. Se, per due elementi $b, c \in S$, risulta

$$a \circ b = a \circ c\,,$$

abbiamo

$$a' \circ (a \circ b) = a' \circ (a \circ c) \ \Rightarrow \ (a' \circ a) \circ b = (a' \circ a) \circ c \ \Rightarrow \ u \circ b = u \circ c \ \Rightarrow \ b = c.$$

Abbiamo così che a è cancellabile a sinistra. Allo stesso modo si vede che a è cancellabile a destra e l'asserto è provato. □

Sia $(S, \circ)$ una struttura algebrica con operazione interna. Due elementi a, b di S si dicono *permutabili* se

$$a \circ b = b \circ a.$$

Un elemento permutabile con tutti gli elementi di S si dice *centrale*. L'insieme di tutti gli elementi centrali si chiama *centro* di $(S, \circ)$ e si denota con $Z(S)$.

Osservazione 1.2.6. Valgono le seguenti proprietà:

– ogni elemento è permutabile con se stesso;

– $S = Z(S)$ se, e solo se, l'operazione $\circ$ è commutativa;

– Se $(S, \circ)$ è un semigruppo si ha:
  (1) a, b permutabili con $c \Rightarrow a \circ b$ permutabile con c;
  (2) Se $Z(S)$ è non vuoto, allora: $a, b \in Z(S) \ \Rightarrow \ a \circ b \in Z(S)$. □

1.3 Isomorfismi

Consideriamo le operazioni $\cdot, +, *$ rispettivamente su

$$S_1 = \{1, r_1, r_2, r_3\}, \quad S_2 = \{0, 1, 2, 3\}, \quad S_3 = \{e, a, b, c\},$$

definite dalle seguenti tabelle[2]:

$\cdot$	1	r_1	r_2	r_3
1	1	r_1	r_2	r_3
r_1	r_1	r_2	r_3	1
r_2	r_2	r_3	1	r_1
r_3	r_3	1	r_1	r_2

$+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$*$	e	a	b	c
e	e	a	b	c
a	a	b	c	e
b	b	c	e	a
c	c	e	a	b

E' evidente che, a meno dei nomi dati alle tre operazioni e agli elementi dei tre insiemi, le strutture considerate sono la stessa struttura; esse sono cioè *algebricamente equivalenti* o, come usualmente si dice, *isomorfe*. Per esempio possiamo identificare la seconda e la terza ponendo:

$$+ = *, \quad 0 = e, \quad 1 = a, \quad 2 = b, \quad 3 = c.$$

[2]Le tabelle che definiscono operazioni su un insieme prendono il nome di *tabelle di Cayley*.

Nasce, così , l'esigenza di definire rigorosamente le situazioni che permettono di ritenere algebricamente equivalenti due strutture algebriche. Ciò si può fare introducendo il concetto di *isomorfismo*.

Siano $(S, \circ_1)$ e $(S', \circ_2)$ strutture algebriche con operazioni interne. Una applicazione biunivoca $f\colon S \to S'$ si dice *isomorfismo* se

$$f(a \circ_1 b) = f(a) \circ_2 f(b), \quad \text{per ogni } a, b \in S.$$

Siano $(S, \circ_1)$ e $(S', \circ_2)$ strutture algebriche con operazioni esterne aventi lo stesso dominio di operatori A. Un'applicazione biunivoca $f\colon S \to S'$ si dice *isomorfismo* se

$$f(\alpha \circ_1 a) = \alpha \circ_2 f(a), \quad \text{per ogni } \alpha \in A \text{ e } a \in S.$$

Un *isomorfismo* fra due strutture algebriche ad n operazioni

$$(S, \star_1, \star_2, \dots, \star_n) \quad \text{e} \quad (S', \circ_1, \circ_2, \dots, \circ_n)$$

è un'applicazione biunivoca $f\colon S \to S'$ per cui esiste una permutazione σ degli indici $1, 2, ..., n$, tale che f è un isomorfismo fra $(S, \star_j)$ e $(S', \circ_{\sigma(j)})$, per ogni $j = 1, 2, \dots, n$.

Riportiamo di seguito alcune proprietà degli isomorfismi di facile dimostrazione:

- L'identità è un isomorfismo di ogni struttura algebrica in se stessa.
- f isomorfismo $\Rightarrow f^{-1}$ isomorfismo (*l'inverso di* f).
- $S_1 \xrightarrow{f} S_2 \xrightarrow{g} S_3$, f, g isomorfismi $\Rightarrow S_1 \xrightarrow{fg} S_3$ isomorfismo.
- La relazione di isomorfismo fra strutture algebriche è riflessiva, simmetrica e transitiva.

Osservazione 1.3.1. Dal punto di vista algebrico due strutture isomorfe possono considerarsi equivalenti. Questo fatto si esprime dicendo che *lo studio delle strutture algebriche si fa a meno di isomorfismi.* □

Esempio 1.3.2. Si considerino le strutture algebriche $(\mathbb{R}^+, \cdot)$, ove $\mathbb{R}^+$ denota l'insieme dei numeri reali positivi, e $(\mathbb{R}, +)$. E' noto che la funzione logaritmo

$$\log \ : \ a \in R^+ \ \to \ log\, a \in \mathbb{R}$$

è biunivoca. Essa è un isomorfismo di $(\mathbb{R}^+, \cdot)$ in $(\mathbb{R}, +)$ perché risulta

$$\log(ab) = \log(a) \, + \, \log(b)\,, \quad \text{per ogni } a, b \in \mathbb{R}^+.$$

L'isomorfismo inverso della funzione logaritmo è la funzione esponenziale. □

1.4 Morfismi

Le definizioni che seguono introducono un'importante generalizzazione del concetto di isomorfismo.

Siano $(S_1, \circ)$ e $(S_2, *)$ due strutture algebriche con operazioni interne. Un'applicazione $f\colon S_1 \to S_2$ si chiama *morfismo* o *omomorfismo* se:

$$f(a \circ b) = f(a) * f(b) \ \ \textit{per ogni } a, b \in S_1.$$

Siano $(S_1, \circ)$ e $(S_2, *)$ strutture algebriche con operazioni esterne aventi lo stesso dominio di operatori A. Una applicazione $f\colon S_1 \to S_2$ si chiama *morfismo* se:

$$f(\alpha \circ a) = \alpha * f(a), \ \ \textit{per ogni } \alpha \in A \textit{ e } a \in S_1.$$

Un *morfismo* fra due strutture

$$(S_1, \circ_1, \circ_2, \cdots, \circ_n) \ \ , \ \ (S_2, *_1, *_2, \cdots, *_n)$$

ad n operazioni è una applicazione $f\colon S_1 \to S_2$ se è un morfismo fra $(S_1, \circ_j)$ e $(S_2, *_j)$, per ogni $j = 1, 2, \ldots n$.

Osservazione 1.4.1. Un morfismo biettivo è un isomorfismo. La nozione di morfismo è dunque una generalizzazione di quella di isomorfismo. Essa è molto utile perché due strutture non isomorfe possono essere legate tra loro da un morfismo e, come vedremo, vi sono proprietà delle strutture algebriche che sono conservate dai morfismi. □

Un morfismo f si dice *monomorfismo* se è iniettivo, *epimorfismo* se è suriettivo.

Esercizio 1.4.2. *Provare le seguenti implicazioni:*

- $S_1 \xrightarrow{f} S_2 \xrightarrow{g} S_3$, f, g *morfismi* $\Rightarrow S_1 \xrightarrow{fg} S_3$ *morfismo.*
- $S_1 \xrightarrow{f} S_2$, f *morfismo,* X *parte stabile di* $S_1 \Rightarrow f(X)$ *parte stabile di* S_2.

Esempio 1.4.3. Le applicazioni $a \in \mathbb{N} \to a \in \mathbb{Z}$, $a \in \mathbb{Z} \to a \in \mathbb{Q}$, $a \in \mathbb{Q} \to a \in \mathbb{R}$, $a \in \mathbb{R} \to a \in \mathbb{C}$ sono morfismi rispetto alle operazioni $+$ e $\cdot$. □

Un morfismo di una struttura algebrica S in se stessa si chiama *endomorfismo* di S. Un isomorfismo di una struttura algebrica in se stessa si chiama *automorfismo* di S.

1.5 Gruppi

Una struttura algebrica $(S, \circ)$ si chiama *gruppo* se sono verificate le seguenti proprietà:

1. l'operazione $\circ$ è associativa;

2. esiste l'elemento neutro u;

3. ogni elemento di S è simmetrizzabile.

Se l'operazione $\circ$ è commutativa il gruppo si dice *commutativo* o *abeliano.* Un gruppo si dice *moltiplicativo* (risp. *additivo*) se per la sua operazione si usa la notazione moltiplicativa (risp. additiva). Richiamiamo esplicitamente l'attenzione del Lettore sul fatto che la scelta della notazione per l'operazione di un gruppo è ininfluente sulle proprietà algebriche del gruppo stesso.

Esempi 1.5.1. Le strutture sottoelencate sono esempi di gruppi.

- $(\mathbb{Z},+),(\mathbb{Q},+),(\mathbb{R},+),(\mathbb{C},+)$ (*gruppi additivi* di $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$).
- $(\mathbb{Q}^*,\cdot),(\mathbb{R}^*,\cdot),(\mathbb{C}^*,\cdot)$ (*gruppi moltiplicativi* di $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$).
- $(\mathbb{Z}[x],+)$, $(\mathbb{Q}[x],+)$, $(\mathbb{R}[x],+)$, $(\mathbb{C}[x],+)$ (*gruppi additivi* di $\mathbb{Z}[x]$, $\mathbb{Q}[x]$, $\mathbb{R}[x]$, $\mathbb{C}[x]$). □

Nel seguito del presente paragrafo, tranne esplicito avviso, $G=(G,\cdot\,)$ denoterà sempre un gruppo moltiplicativo. Sarà un utile esercizio per il Lettore trovare l'analogo nella notazione additiva di tutti i concetti che si daranno in notazione moltiplicativa.

Esercizio 1.5.2. *Siano a,b elementi di un gruppo G. Provare che ciascuna delle equazioni $ax=b$ e $xa=b$ ammette un'unica soluzione in G.*

Si chiama *potenza* $n-$esima di un elemento $a\in G$, e si denota con a^n, l'elemento di G definito per ricorrenza da

- *per* $n\geq 0$:
$$a^0=1,\;\; a^n=a^{n-1}a,$$
- *per* $n<0$:
$$a^n=(a^{-1})^{-n}.$$

Esercizio 1.5.3. *Provare che sono verificate le seguenti proprietà, per ogni $a,b\in G$:*

- $a^m a^n = a^{m+n} = a^n a^m$.
- $ab=ba \Rightarrow (ab)^n = a^n b^n$. □

Sia a un elemento di G. Se esiste un intero $m\neq 0$ tale che $a^m=1$ si dice che a è *periodico* o che ha *ordine finito. In questo caso, il più piccolo intero positivo n tale che $a^n=1$ si chiama ordine* o *periodo* di a e si denota con $|a|$ o con $o(a)$. Se per ogni intero positivo m risulta $a^m\neq 1$ si dice che a è *aperiodico* o che ha *ordine infinito.*

Esercizio 1.5.4. *Provare che un gruppo finito G d'ordine pari contiene almeno un elemento di periodo due.*

Soluzione. Osserviamo che un elemento $a \in G \setminus \{1\}$ ha periodo due se, e solo se, $a = a^{-1}$ e che gli elementi b di $G \setminus \{1\}$ per cui è $b \neq b^{-1}$ sono in numero pari; deve, dunque, esistere in G almeno un elemento di periodo due. □

Proposizione 1.5.5. *Siano a un elemento di un gruppo G di periodo n ed m un intero positivo. Allora risulta $a^m = 1$ se, e solo se, m è un multiplo di n.*

Dimostrazione. Se è $m = hn$, risulta

$$a^m = a^{hn} = (a^n)^h = 1^h = 1,$$

cioè la prima parte dell'asserto. Se è $a^m = 1$, detti q ed r il quoziente ed il resto della divisione tra m ed n, risulta

$$1 = a^m = a^{nq+r} = (a^n)^q a^r = a^r.$$

Allora, essendo r un intero non negativo minore di n, deve essere $r = 0$ e l'asserto è provato. □

Osservazione 1.5.6. L'unità è l'unico elemento di un gruppo di periodo 1. □

Esercizio 1.5.7. *L'analogo additivo del concetto di potenza $n-$sima di un elemento a si chiama multiplo di a secondo l'intero n e si denota con na. Definire i multipli in un gruppo additivo e studiarne le prime proprietà. Definire l'ordine di un elemento di un gruppo additivo.*

Esempi 1.5.8. In $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$ ogni elemento diverso da zero ha ordine infinito.

In $(\mathbb{Q}^*, \cdot)$, $(\mathbb{R}^*, \cdot)$ ogni elemento diverso da 1 e -1 ha ordine infinito e risulta $o(-1) = 2$. □

Proposizione 1.5.9. *In un gruppo ogni elemento è regolare, cioè in esso vale la legge di cancellazione.*

Dimostrazione. È una conseguenza immediata della prop.1.2.5. □

1.6 Anelli e campi

Una struttura algebrica con due operazioni interne (*addizione e moltiplicazione*) $A = (A, +, \cdot)$ si chiama *anello* se sono verificate le seguenti proprietà:

1. $(A, +)$ è un gruppo abeliano,

2. $(A, \cdot\,)$ è un semigruppo,

3. la moltiplicazione è distributiva rispetto all'addizione.

L'anello A si dice *commutativo* se la moltiplicazione è commutativa, si dice *unitario* se la moltiplicazione ammette elemento neutro 1.

ESEMPI 1.6.1. $(\mathbb{Z},+,\cdot)$, $(\mathbb{Q},+,\cdot)$, $(\mathbb{R},+,\cdot)$, $(\mathbb{C},+,\cdot,)$ $(\mathbb{Z}[x],+,\cdot)$, $(\mathbb{Q}[x],+,\cdot)$, $(\mathbb{R}[x],+,\cdot)$, $(\mathbb{C}[x],+,\cdot)$ sono anelli commutativi unitari. □

ESEMPIO 1.6.2. (ANELLO DEI RESTI MODULO UN INTERO) Sia $n > 1$ un intero e consideriamo l'insieme $\mathbb{Z}_n := \{0, 1, 2, \ldots, n-1\}$. Si osservi che gli interi da 0 ad $n-1$ costituiscono tutti i possibili resti della divisione di un intero relativo per n. Per ogni intero relativo a, denotiamo con

$$a \mod n$$

il resto della divisione tra a e n e definiamo in $\mathbb{Z}_n$ le seguenti operazioni:

$$a+b = (a+b) \mod n,$$

$$ab = (ab) \mod n,$$

per ogni $a, b \in \mathbb{Z}_n$. Non è difficile provare che la struttura algebrica $(\mathbb{Z}_n, +, \cdot)$ è un anello commutativo unitario. Tale anello si chiama *anello dei resti modulo* n. □

Non è difficile provare che in un anello $(A,+,\cdot)$ valgono le seguenti proprietà, per ogni $a,b,c \in A$ e $n \in \mathbb{Z}$.

- $a0 = 0a = 0$,
- $a(-b) = (-a)b = -ab$,
- $(-a)(-b) = ab$,
- $(na)b = a(nb) = n(ab)$,
- $a(b-c) = ab - ac$ e $(b-a)c = bc - ac$.

OSSERVAZIONE 1.6.3. Se in un anello unitario A risulta $1 = 0$, abbiamo:

$$a = a1 = a0 = 0\,, \quad \text{per ogni } \ a \in A \ \Rightarrow \ A = \{0\}.$$

L'anello $A = \{0\}$ si chiama *anello nullo*. □

Nel seguito A denoterà un anello che, tranne esplicito avviso, supporremo sempre non nullo.

Sia $a \in A$ con $a \neq 0$. L'elemento a si dice *divisore sinistro* (*destro*) *dello zero* se esiste in A un elemento $b \neq 0$ tale che $ab = 0$ ($ba = 0$.) L'elemento a si dice *divisore dello zero* se è divisore sia sinistro che destro dello zero.

ESERCIZIO 1.6.4. *Provare che in un anello esiste un divisore sinistro dello zero se, e solo se, esiste un divisore destro dello zero.*

Un elemento $a \in A$ si dice *nilpotente* se esiste un intero positivo n tale che $a^n = 0$.

Un anello non nullo privo di divisori sinistri dello zero si chiama *anello integro*. Un anello integro commutativo si chiama *dominio di integrità*.

Sia A unitario. Un elemento $a \in A$ si dice *invertibile* se è tale nel semigruppo $(A, \cdot)$. Ovviamente l'unità di A è invertibile e coincide col proprio inverso. L'insieme degli elementi invertibili di A, che è non vuoto perché 1 è invertibile, si denota con $U(A)$ e si prova che, rispetto alla moltiplicazione, é un gruppo.

Un anello commutativo unitario A si chiama *campo* se ogni suo elemento non nullo è invertibile, cioè se $U(A) = A \setminus \{0\}$.

Esempi 1.6.5. Le strutture algebriche $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ e $(\mathbb{C}, +, \cdot)$ sono esempi di campi. $\square$

Esempio 1.6.6. L'anello commutativo $(\mathbb{Z}_2 = \{0, 1\}, +, \cdot)$ é campo perchè il suo unico elemento non nullo è 1, che è invertibile. $\square$

Esempio 1.6.7. L'anello commutativo $(\mathbb{Z}_3 = \{0, 1, 2\}, +, \cdot)$ é campo perchè i suoi elementi non nulli sono tutti invertibili. In questo anello, infatti, risulta $2^{-1} = 2$ perché $2 \cdot 2 = 4 \mod 3 = 1$. $\square$

Esempio 1.6.8. L'anello commutativo $(\mathbb{Z}_4 = \{0, 1, 2, 3\}, +, \cdot)$ non é campo. L'elemento 2, infatti, non è invertibile, risultando $2 \cdot 1 = 2$, $2 \cdot 2 = 4 \mod 4 = 0$, $2 \cdot 3 = 6 \mod 4 = 2$. $\square$

Gli ultimi tre esempi sono casi particolari del seguente teorema.

Proposizione 1.6.9. *L'anello $(\mathbb{Z}_n, +, \cdot)$ dei resti modulo un intero n è un campo se, e solo se, n è un numero primo.*

Capitolo 2

Vettori numerici

2.1 $n-$ple di elementi di un campo

Sia $\mathbb{F}$ un campo[1], i cui elementi si diranno *scalari*[2], e consideriamo l'insieme $\mathbb{F}^n$ delle $n-$ple ordinate degli elementi di $\mathbb{F}$:

$$\mathbb{F}^n = \mathbb{F} \times \mathbb{F} \times \cdots \times \mathbb{F} = \{(a_1, a_2, \ldots, a_n) \ : \ a_1, a_2, \ldots, a_n \in \mathbb{F}\}.$$

Gli scalari $a_1, a_2, \ldots, a_n$ che compaiono nell'$n-$pla $(a_1, a_2, \ldots, a_n)$ si chiamano *componenti* dell'$n-$pla. Spesso, se le componenti di una $n-$pla sono assegnate usando una lettera con indici da 1 ad n, useremo la corrispondente lettera in grassetto per denotare l'$n-$pla in questione; questo significa che porremo:

$$\boldsymbol{a} = (a_1, a_2, \ldots, a_n) \ , \ \ \boldsymbol{b} = (b_1, b_2, \ldots, b_n) \ , \ldots \text{ e così via.}$$

Quando l'intero n è chiaro dal contesto, si usa anche la notazione

$$\boldsymbol{a} = (a_j), \tag{2.1}$$

invece di $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$. Per ogni $n-$pla $\boldsymbol{a} \in \mathbb{F}^n$, diciamo *$n-$pla opposta di $\boldsymbol{a}$*, e la denotiamo con $-\boldsymbol{a}$, quella definita da

$$-\boldsymbol{a} = (-a_1, -a_2, \ldots, -a_n);$$

inoltre, diciamo *$n-$pla nulla*, e la denotiamo con $\mathbf{0}$, quella che ha tutte le componenti uguali a zero, cioè

$$\mathbf{0} = (0, 0, \ldots, 0).$$

[1] Il Lettore che non cocnosca, o che non abbia ancora assimilato bene, il concetto di campo può continuare la lettura di queste note pensando che $\mathbb{F}$ sia il campo $\mathbb{R}$ dei numeri reali, o il campo $\mathbb{Q}$ dei numeri razionali.

[2] Il termine *scalare* è preso in prestito dalla fisica, dove per grandezza scalare si intende una grandezza fisica caratterizzata da un numero (in generale reale) che rappresenta la sua misura rispetto ad una fissata "scala" (ad sempio, densità, tempo, temperatura, pressione).

Esempio 2.1.1. Per i seguenti elementi

$$\boldsymbol{a} = (1, -2, -5)\,,\ \boldsymbol{b} = (\frac{3}{5}, -7, 2, 0, 0, 1)\,,\ \boldsymbol{c} = (\frac{4}{3}, -1, 2, -9),$$

rispettivamente di $\mathbb{R}^3$,$\mathbb{R}^6$,$\mathbb{R}^4$, risulta

$$-\boldsymbol{a} = (-1, 2, 5)\,,\ -\boldsymbol{b} = (-\frac{3}{5}, 7, -2, 0, 0, -1)\,,\ -\boldsymbol{c} = (-\frac{4}{3}, 1, -2, 9) \qquad \square$$

Nell'insieme $\mathbb{F}^n$ si definiscono un'operazione interna di *addizione* fra $n-$ple ed una esterna di *moltiplicazione* di uno scalare per un'$n-$pla, nel seguente modo:

- addizione di due n-ple: per ogni $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{F}^n$,

$$\boldsymbol{a} + \boldsymbol{b} = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n); \tag{2.2}$$

- moltiplicazione di uno scalare per una $n-$pla: per ogni $\boldsymbol{a} \in \mathbb{F}^n$ e $\alpha \in \mathbb{F}$,

$$\alpha\boldsymbol{a} = (\alpha a_1, \alpha a_2, \dots, \alpha a_n). \tag{2.3}$$

Se si usa la notazione (2.1), le (2.2) e (2.3) si scrivono rispettivamente

$$\boldsymbol{a} + \boldsymbol{b} = (a_j + b_j) \ \text{ e } \ \alpha\boldsymbol{a} = (\alpha a_j).$$

Esempio 2.1.2. Considerati gli elementi

$$\boldsymbol{a} = (1, 1, -1, -1)\,,\ \boldsymbol{b} = (2, -5, 4, 3)\,,\ \boldsymbol{c} = (0, 6, 0, -2)$$

di $\mathbb{R}^4$, risulta

$$\boldsymbol{a} + \boldsymbol{b} = (3, -4, 3, 2)\,,\ \boldsymbol{b} + \boldsymbol{c} = (2, 1, 4, 1)\,,\ \boldsymbol{c} + \boldsymbol{a} = (1, 7, -1, -3)\,,$$

$$\boldsymbol{a} + (-\boldsymbol{a}) = \boldsymbol{b} + (-\boldsymbol{b}) = \boldsymbol{c} + (-\boldsymbol{c}) = (0, 0, 0, 0).$$

e

$$3\boldsymbol{a} = (3, 3, -3, -3)\,,\ 2\boldsymbol{b} = (4, -10, 8, 6)\,,\ \frac{1}{2}\boldsymbol{c} = (0, 3, 0, -1). \qquad \square$$

Le operazioni definite dalle (2.2) e (2.3) verificano le seguenti proprietà:

$SV1$. *proprietà associativa dell'addizione*:
per ogni $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{F}^n$,

$$(\boldsymbol{a} + \boldsymbol{b}) + \boldsymbol{c} = \boldsymbol{a} + (\boldsymbol{b} + \boldsymbol{c}); \tag{2.4}$$

$SV2$. *esistenza dell'elemento neutro per l'addizione*:
per ogni $\boldsymbol{a} \in \mathbb{F}^n$,

$$\boldsymbol{a} + \boldsymbol{0} = \boldsymbol{0} + \boldsymbol{a} = \boldsymbol{a}; \tag{2.5}$$

$SV3$. *simmetrizzabilità rispetto all'addizione*:

per ogni $\boldsymbol{a} \in \mathbb{F}^n$,
$$\boldsymbol{a} + (-\boldsymbol{a}) = -\boldsymbol{a} + \boldsymbol{a} = \boldsymbol{0}; \tag{2.6}$$

$SV4$. *proprietà commutativa dell'addizione*:

per ogni $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{F}^n$,
$$\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{b} + \boldsymbol{a}; \tag{2.7}$$

$SV5$. *proprietà distributiva della moltiplicazione di una $n-$pla per uno scalare rispetto all'addizione fra $n-$ple*:

per ogni $\alpha \in \mathbb{F}$ e per ogni $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{F}^n$,
$$\alpha(\boldsymbol{a} + \boldsymbol{b}) = \alpha\boldsymbol{a} + \alpha\boldsymbol{b}; \tag{2.8}$$

$SV6$. *proprietà distributiva della moltiplicazione di una $n-$pla per uno scalare rispetto all'addizione fra scalari*:

per ogni $\alpha, \beta \in \mathbb{F}$ e per ogni $\boldsymbol{a} \in \mathbb{F}^n$,
$$(\alpha + \beta)\boldsymbol{a} = \alpha\boldsymbol{a} + \beta\boldsymbol{a}; \tag{2.9}$$

$SV7$. *proprietà distributiva della moltiplicazione di una $n-$pla per uno scalare con la moltiplicazione fra scalari*:

per ogni $\alpha, \beta \in \mathbb{F}$ e per ogni $\boldsymbol{a} \in \mathbb{F}^n$,
$$(\alpha\beta)\boldsymbol{a} = \alpha(\beta\boldsymbol{a}); \tag{2.10}$$

$SV8$. per ogni $\boldsymbol{a} \in \mathbb{F}^n$,
$$1\,\boldsymbol{a} = \boldsymbol{a}. \tag{2.11}$$

ESERCIZIO 2.1.3. *Dimostrare le proprietà* $SV1$, $SV2$, $SV3$, $SV4$, $SV5$, $SV6$, $SV7$, $SV8$.

A titolo di esempio, dimostriamo la propietà $SV5$:

$$\begin{aligned}
\alpha(\boldsymbol{a} + \boldsymbol{b}) =& \alpha(a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n) \\
=& (\alpha(a_1 + b_1), \alpha(a_2 + b_2), \ldots, \alpha(a_n + b_n)) \\
=& (\alpha a_1 + \alpha b_1, \alpha a_2 + \alpha b_2, \ldots, \alpha a_n + \alpha b_n) \\
=& (\alpha a_1, \alpha a_2, \ldots, \alpha a_n) + (\alpha b_1, \alpha b_2, \ldots, \alpha b_n) \\
=& \alpha(a_1, a_2, \ldots, a_n) + \alpha(b_1, b_2, \ldots, b_n) \\
=& \alpha\boldsymbol{a} + \alpha\boldsymbol{b}
\end{aligned}$$

Il fatto che in $\mathbb{F}^n$ le operazioni di addizione e moltiplicazione appena definite verificano le precedenti otto proprietà si esprime dicendo che $\mathbb{F}^n$ è uno *spazio vettoriale sul campo* $\mathbb{F}$. Più precisamente, $\mathbb{F}^n$ si dice *spazio vettoriale numerico* su $\mathbb{F}$ e, per tale motivo, i suoi elementi si chiamano anche *vettori numerici di lunghezza n*, o più semplicemente *vettori*.

Osservazione 2.1.4. In $\mathbb{F}^n$, per la somma di tre vettori $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$, si può scrivere

$$\boldsymbol{a} + \boldsymbol{b} + \boldsymbol{c},$$

senza indicare con le parentesi l'ordine delle operazioni. Infatti, le due possibili interpretazioni $(\boldsymbol{a} + \boldsymbol{b}) + \boldsymbol{c}$ e $\boldsymbol{a} + (\boldsymbol{b} + \boldsymbol{c})$, essendo l'addizione fra vettori numerici associativa (*cfr.* proprietà $SV1$), danno luogo allo stesso risultato. Più in generale, è ben definita la scrittura

$$\boldsymbol{a}_1 + \boldsymbol{a}_2 + \cdots + \boldsymbol{a}_h$$

per la somma di h vettori. □

2.2 Dipendenza lineare

Nello spazio vettoriale numerico $\mathbb{F}^n$ si considerino h vettori $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h$ e h scalari $\alpha_1, \alpha_2, \ldots, \alpha_h$. Il vettore

$$\boldsymbol{a} = \alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2 + \cdots + \alpha_h\boldsymbol{a}_h \tag{2.12}$$

prende il nome di *combinazione lineare* dei vettori $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h$ mediante gli scalari $\alpha_1, \alpha_2, \ldots, \alpha_h$. Quando vale la (2.12) si dice anche che il vettore $\boldsymbol{a}$ *dipende linearmente* dai vettori $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h$.

Esempio 2.2.1. In $\mathbb{R}^3$ la combinazione lineare dei vettori

$$(1, -1, 1),\ (\frac{1}{2}, -1, 0),\ (2, 0, -1),\ (0, -\frac{2}{3}, 1)$$

mediante gli scalari $2, -2, 1, 3$ è

$$\begin{aligned} 2(1,-1,1) - 2(\tfrac{1}{2},-1,0) + (2,0,-1) + 3(0,-\tfrac{2}{3},1) &= \\ (2,-2,2) + (-1,2,0) + (2,0,-1) + (0,-2,3) &= \\ (2-1+2+0, -2+2+0-2, 2+0-1+3) &= (3,-2,4). \end{aligned}$$

□

Osservazione 2.2.2. Supponiamo che nella (2.12) sia

$$\boldsymbol{a} = (a_1, a_2, \ldots, a_n) \ \text{ e } \ \boldsymbol{a}_i = (a_{i1}, a_{i2}, \ldots, a_{in}),$$

per ogni $i = 1, 2, \ldots, n$. Allora la relazione vettoriale (2.12) equivale alle h relazioni scalari

$$\begin{cases} a_1 = \alpha_1 a_{11} + \alpha_2 a_{12} + \cdots + \alpha_n a_{1n} \\ a_2 = \alpha_1 a_{21} + \alpha_2 a_{22} + \cdots + \alpha_n a_{2n} \\ \qquad \vdots \\ a_h = \alpha_1 a_{h1} + \alpha_2 a_{h2} + \cdots + \alpha_n a_{hn} \end{cases} \tag{2.13}$$

e queste equivalgono al fatto che il sistema di equazioni

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = a_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = a_2 \\ \qquad \vdots \\ a_{h1}x_1 + a_{h2}x_2 + \cdots + a_{hn}x_n = a_h \end{cases}, \tag{2.14}$$

nelle incognite $x_1, x_2, \ldots, x_n$, ha la soluzione $x_1 = \alpha_1$, $x_2 = \alpha_2, \ldots, x_n = \alpha_n$. Di conseguenza, se un vettore $\boldsymbol{b} = (b_1, b_2, \ldots, b_n)$ non è combinazione lineare dei vettori $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h$, allora il sistema di equazioni

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad \vdots \\ a_{h1}x_1 + a_{h2}x_2 + \cdots + a_{hn}x_n = b_h \end{cases}, \tag{2.15}$$

nelle incognite $x_1, x_2, \ldots, x_n$, non ammette soluzioni; non esiste, cioè, alcuna $n-$pla di scalari $\beta_1, \beta_2, \ldots, \beta_n$, tali che

$$\begin{cases} a_{11}\beta_1 + a_{12}\beta_2 + \cdots + a_{1n}\beta_n = b_1 \\ a_{21}\beta_1 + a_{22}\beta_2 + \cdots + a_{2n}\beta_n = b_2 \\ \qquad \vdots \\ a_{h1}\beta_1 + a_{h2}\beta_2 + \cdots + a_{hn}\beta_n = b_h \end{cases}. \tag{2.16}$$

I sistemi di equazioni del tipo (2.14) si dicono *sistemi lineari* e saranno studiati nei prossimi capitoli. □

Osservazione 2.2.3. La combinazione lineare di h vettori numerici $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\ldots$, $\boldsymbol{a}_h$ con coefficienti tutti nulli è uguale al vettore nullo, cioè

$$0\boldsymbol{a}_1 + 0\boldsymbol{a}_2 + \cdots + 0\boldsymbol{a}_h = \boldsymbol{0}\,,$$

ma, in generale, non è vero il viceversa. Esistono, cioè, vettori che ammettono combinazioni lineari a coefficienti non tutti nulli uguali al vettore nullo; ad esempio, per i vettori

$$(3,3,3),\ (1,0,0),\ (0,0,1),\ (0,2,2) \in \mathbb{R}^3 \tag{2.17}$$

risulta

$$\begin{aligned} -2(3,3,3) + 6(1,0,0) + 0(0,0,1) + 3(0,2,2) &= \\ (-6+6+0+0, -6+0+0+6, -6+0+0+6) &= (0,0,0). \end{aligned}$$

È importante osservare che esistono anche vettori la cui unica combinazione lineare uguale al vettore nullo è quella che ha tutti i coefficienti nulli. Questo, ad esempio, accade per i vettori

$$(1,0,0),\ (0,1,0),\ (0,0,1) \in \mathbb{R}^3; \tag{2.18}$$

infatti, per $\alpha, \beta, \gamma \in \mathbb{R}$, si ha

$$\begin{aligned} \alpha(1,0,0)+\beta(0,1,0)+\gamma(0,0,1) &= (0,0,0) \quad \Leftrightarrow \\ (\alpha,0,0)+(0,\beta,0)+(0,0,\gamma)=(\alpha,\beta,\gamma) &= (0,0,0) \quad \Leftrightarrow \\ \alpha=\beta=\gamma=0. & \end{aligned}$$

□

I vettori $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h$ si dicono *linearmente dipendenti,* o semplicemente *dipendenti,* se uno di essi dipende linearmente dai rimanenti, cioè se

$$\boldsymbol{a}_i = \alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2 + \cdots + \alpha_{i-1}\boldsymbol{a}_{i-1} + \alpha_{i+1}\boldsymbol{a}_{i+1} + \cdots + \alpha_h\boldsymbol{a}_h, \tag{2.19}$$

per qualche indice $i = 1, 2, \ldots, h$. In quest'ultimo caso si dice anche che l'insieme $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h\}$ è *linearmente dipendente,* o *dipendente.* Vettori numerici, o insiemi finiti di tali vettori, che non sono dipendenti si dicono *linearmente indipendenti,* o *indipendenti.*

Proposizione 2.2.4. *I vettori numerici $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h$ di $\mathbb{F}^n$ sono linearmente dipendenti se, e solo se, esiste una loro combinazione lineare a coefficienti non tutti nulli uguale al vettore nullo, cioè*

$$\alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2 + \cdots + \alpha_h\boldsymbol{a}_h = \mathbf{0} \tag{2.20}$$

con qualche scalare $\alpha_i \in \mathbb{F}$ diverso da zero.

Dimostrazione. Se $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h$ sono linearmente dipendenti, vale la relazione (2.19), per qualche indice i. Allora risulta

$$\alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2 + \cdots + \alpha_{i-1}\boldsymbol{a}_{i-1} - \boldsymbol{a}_i + \alpha_{i+1}\boldsymbol{a}_{i+1} + \cdots + \alpha_h\boldsymbol{a}_h = \mathbf{0},$$

con il coefficiente di $\boldsymbol{a}_i$ uguale a $-1 \neq 0$. Viceversa, se vale la (2.20), con $\alpha_i \neq 0$ per qualche indice i, risulta

$$\boldsymbol{a}_i = -\frac{\alpha_1}{\alpha_i}\boldsymbol{a}_1 - \frac{\alpha_2}{\alpha_i}\boldsymbol{a}_2 - \cdots - \frac{\alpha_{i-1}}{\alpha_i}\boldsymbol{a}_{i-1} - \frac{\alpha_{i+1}}{\alpha_i}\boldsymbol{a}_{i+1} - \cdots - \frac{\alpha_h}{\alpha_i}\boldsymbol{a}_h$$

e i nostri vettori sono linearmente dipendenti. □

Corollario 2.2.5. *I vettori numerici $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h$ di $\mathbb{F}^n$ sono linearmente indipendenti se, e solo se, la loro combinazione lineare con coefficienti tutti nulli è l'unica che risulta uguale al vettore nullo, cioè*

$$\alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2 + \cdots + \alpha_h\boldsymbol{a}_h = \mathbf{0} \;\Rightarrow\; \alpha_1 = \alpha_2 = \cdots = \alpha_h = 0. \tag{2.21}$$

Esempio 2.2.6. I quattro vettori di $\mathbb{R}^3$ di cui alla (2.17) sono linearmente dipendenti. I tre vettori di $\mathbb{R}^3$ di cui alla (2.18) sono linearmente indipendenti. I vettori $(1,1,0)$, $(0,1,1)$, $(1,0,1) \in \mathbb{R}^3$ sono linearmente indipendenti; infatti, per $\alpha, \beta, \gamma \in \mathbb{R}$, si ha

$$\begin{aligned} \alpha(1,1,0)+\beta(0,1,1)+\gamma(1,0,1) = (0,0,0) &\quad \Leftrightarrow \\ (\alpha,\alpha,0)+(0,\beta,\beta)+(\gamma,0,\gamma) = (\alpha+\gamma,\alpha+\beta,\beta+\gamma) = (0,0,0) &\quad \Leftrightarrow \\ \alpha+\gamma=\alpha+\beta=\beta+\gamma=0 &\quad \Leftrightarrow \\ \alpha=\beta=\gamma=0 &\quad . \end{aligned}$$

□

Le nozioni di dipendenza e indipendenza lineare per gli insiemi finiti si estendono in modo naturale alle $h-$ple ordinate di vettori. Si dice, infatti, che una $h-$pla $(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h)$ è *dipendente* o *indipendente*, se valgono rispettivamente la (2.19) o la (2.21). Si noti che, a differenza degli elementi di un insieme, le componenti di una $h-$pla non sono necessariamente a due a due distinte. Nel seguito un'$h-$pla di vettori sarà anche detta *sistema* di vettori e le componenti dell'$h-$pla si diranno *vettori* o *elementi* del sistema. È chiaro che ogni proprietà relativa alla dipendenza o indipendenza di insiemi finiti si traduce in una corrispondente proprietà dei sistemi, e viceversa.

2.3 Sottospazi vettoriali

Un sottoinsieme non vuoto U di $\mathbb{F}^n$ si dice *sottospazio vettoriale*, o semplicemente *sottospazio*, di $\mathbb{F}^n$ se sono verificate le seguenti due proprietà:

1. *Proprietà di chiusura rispetto all'addizione*:

$$\text{per ogni } \boldsymbol{a}, \boldsymbol{b} \in U \;\Rightarrow\; \boldsymbol{a} + \boldsymbol{b} \in U.$$

2. *Proprietà di chiusura rispetto alla moltiplicazione per scalari*:

$$\text{per ogni } \boldsymbol{a} \in U \text{ e per ogni } \alpha \in \mathbb{F} \;\Rightarrow\; \alpha\boldsymbol{a} \in U.$$

Per indicare che U è sottospazio di $\mathbb{F}^n$ si usa la notazione $U \leq \mathbb{F}^n$. Tra i sottospazi di $\mathbb{F}^n$ ve ne sono due detti *banali*: $\mathbb{F}^n$ stesso e il singleton[3] del vettore nullo $\{\mathbf{0}\}$. Quest'ultimo sottospazio si chiama *sottospazio nullo* e, con abuso di notazione, si denota con $\mathbf{0}$.

Esempio 2.3.1. Il sottoinsieme U di $\mathbb{F}^3$ definito da

$$U = \{(a, b, 0) \;:\; a, b \in \mathbb{F}\}$$

è un sottospazio. Si ha, infatti che:

1. $(a, b, 0), (a', b', 0) \in U \;\Rightarrow\; (a, b, 0) + (a', b', 0) = (a + a', b + b', 0) \in U$;
2. $(a, b, 0) \in U\,, \alpha \in \mathbb{F} \;\Rightarrow\; \alpha(a, b, 0) = (\alpha a, \alpha b, 0) \in U$.

Si noti che in $\mathbb{F}^3$ il sottospazio U può anche essere definito da

$$U = \{(x_1, x_2, x_3) \;:\; x_3 = 0\}$$

e, per tale motivo, l'equazione $x_3 = 0$ si dice *equazione di U*. □

Proposizione 2.3.2. *Ogni sottospazio U di $\mathbb{F}^n$ contiene l'opposto di ogni suo elemento e il vettore nullo.*

[3] Il *singleton* di un elemento a è, per definizione, l'insieme $\{a\}$ contenente il solo elemento a.

Dimostrazione. Se $\boldsymbol{a}$ è un elemento di U, allora $-\boldsymbol{a} = (-1)\boldsymbol{a}$ e $\boldsymbol{0} = \boldsymbol{a}+(-\boldsymbol{a})$ appartengono ad U in forza delle proprietà 2 e 1, rispettivamente. □

Osservazione 2.3.3. Sia U un sottospazio di $\mathbb{F}^n$. Allora $\mathbb{F}^n$ induce su U un'operazione di addizione fra vettori e una di moltiplicazione per uno scalare. Rispetto a tali operazioni, U verifica le stesse otto proprietà $SV1$, $SV2, \dots,$ $SV8$ di $\mathbb{F}^n$. Per tale motivo si dice che U è a sua volta uno spazio vettoriale. □

Proposizione 2.3.4. *Un sottoinsieme non vuoto U di $\mathbb{F}^n$ è un sottospazio se, e solo se, è verificata la seguente proprietà:*

$$\text{per ogni } \boldsymbol{a},\, \boldsymbol{b} \in U \text{ e per ogni } \alpha,\, \beta \in \mathbb{F} \;\Rightarrow\; \alpha\boldsymbol{a} + \beta\boldsymbol{b} \in U. \tag{2.22}$$

Dimostrazione. Se vale la (2.22), le due proprietà che definiscono un sottospazio si ottengono per $(\alpha, \beta) = (1, 1)$ e $(\alpha, \beta) = (1, 0)$. La seconda parte della dimostrazione è immediata. □

2.4 Prodotto scalare

Nello spazio vettoriale $\mathbb{F}^n$ si definisce un'operazione di moltiplicazione tra $n-$ple a valori nel campo $\mathbb{F}$ nel seguente modo:

$$(\boldsymbol{a}, \boldsymbol{b}) \in \mathbb{F}^n \times \mathbb{F}^n \quad \to \quad a_1b_1 + a_2b_2 + \cdots + a_nb_n \in \mathbb{F}\,. \tag{2.23}$$

L'operazione definita dalla (2.23) si chiama *moltiplicazione scalare* (*standard*) in $\mathbb{F}^n$ e il risultato della moltiplicazione di due $n-$ple $\boldsymbol{a}, \boldsymbol{b}$ si chiama *prodotto scalare* (*standard*) di $\boldsymbol{a}$ e $\boldsymbol{b}$ e si denota con $\boldsymbol{ab}$, o con $\boldsymbol{a} \cdot \boldsymbol{b}$; si pone cioè

$$\boldsymbol{ab} := a_1b_1 + a_2b_2 + \cdots + a_nb_n\,. \tag{2.24}$$

Il prodotto scalare verifica le seguenti proprietà:

1. $\boldsymbol{ab} = \boldsymbol{ba}$ (*proprietà commutativa*),
2. $(\boldsymbol{a} + \boldsymbol{b})\boldsymbol{c} = \boldsymbol{ac} + \boldsymbol{bc}$ (*proprietà distributiva*),
3. $(\alpha\boldsymbol{a})\boldsymbol{b} = \alpha(\boldsymbol{ab})$,

per ogni $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{F}^n$ e per ogni scalare α. In particolare, per $\boldsymbol{a} = \boldsymbol{b}$, si ha

$$\boldsymbol{aa} = a_1^2 + a_2^2 + \cdots + a_n^2\,. \tag{2.25}$$

Capitolo 3

Matrici

3.1 Generalità

Sia $\mathbb{F}$ un campo e, per ogni intero $h > 0$, denotiamo con $\mathbb{N}_h$ l'insieme dei primi h interi positivi, cioè

$$\mathbb{N}_h := \{1, 2, \ldots, h\}.$$

Per esempio, è $\mathbb{N}_1 = \{1\}$, $\mathbb{N}_2 = \{1, 2\}$, $\mathbb{N}_5 = \{1, 2, 3, 4, 5\}$.

Assegnati due interi positivi m, n, una funzione

$$A \;:\; (i, j) \in \mathbb{N}_m \times \mathbb{N}_n \;\rightarrow\; a_{ij} \in \mathbb{F} \tag{3.1}$$

prende il nome di *matrice di tipo* $m \times n$ *su* $\mathbb{F}$. Nella pratica è comodo identificare la matrice A, definita dalla (3.1), con la tabella di m righe e n colonne avente l'elemento a_{ij} nel posto (i, j) corrispondente alla $i-$esima riga e $j-$ma colonna, per ogni $i = 1, 2, \ldots, m$ e $j = 1, 2, \ldots, n$; si pone cioè

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}. \tag{3.2}$$

Per ogni $i = 1, 2, \ldots, m$ e $j = 1, 2, \ldots, n$, gli elementi a_{ij} di $\mathbb{F}$ si chiamano *elementi*, o *coefficienti*, o *entrate*, della matrice A *di posto* (i, j). Per tale motivo, A si dice anche matrice *ad elementi*, o *a coefficienti*, o *ad entrate*, in $\mathbb{F}$. Gli indici i, j, che individuano la posizione dell'elemento a_{ij} nella (3.2), si chiamano rispettivamente *indice di riga* e *indice di colonna* di a_{ij}.

Esempio 3.1.1. Quelle che seguono sono due matrici a coefficienti reali:

$$A = \begin{pmatrix} 1 & 5 & 0 & 2 \\ 3 & -1 & 9 & 12 \\ 5 & 11 & 3 & 6 \end{pmatrix} \quad , \quad B = \begin{pmatrix} 9 & 0 & 5 \\ 5 & 1 & -4 \\ 0 & -9 & 1 \\ 2 & 22 & 6 \end{pmatrix}.$$

La prima è di tipo 3×4, la seconda ti tipo 4×3 e, per esempio, risulta

$$a_{12} = 5\ ,\ a_{34} = 6\ ,\ b_{22} = 1\ ,\ b_{41} = 2\,.$$

Nella matrice A, l'indice di riga dell'elemento 9 è 2 e l'indice di colonna è 3. Gli indici di riga e colonna dell'elemento 9 nella matrice B sono entrambi uguali ad 1. □

Quando il numero di righe m e il numero di colonne n di una matrice A sono chiari dal contesto, si usa anche la notazione

$$A = (a_{ij}) \tag{3.3}$$

invece della (3.2). In analogia a quanto fatto per le $n-$ple, se una matrice è assegnata mediante una lettera maiuscola, useremo la stessa lettera con due indici per denotare i suoi elementi; questo significa che porremo

$$B = (b_{ij})\ ,\ \ C = (c_{ij})\ ,\ ...\ \text{e così via.}$$

Una matrice di tipo $1 \times n$ si dice *vettore riga di lunghezza* n e una matrice di tipo $m \times 1$ si dice *matrice colonna di lunghezza* m; è chiaro che tali matrici si possono pensare rispettivamente come elementi di $\mathbb{F}^n$ e $\mathbb{F}^m$. Spesso per un vettore riga del tipo

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \end{pmatrix}$$

si userà la notazione

$$(a_{11}, a_{12}, \ldots, a_{1n}).$$

Esempio 3.1.2. Le matrici

$$\begin{pmatrix} 2 & 5 & 22 & 7 \end{pmatrix} = (2, 5, 22, 7)\ ,\ \ \begin{pmatrix} 12 \\ 8 \\ 2 \\ 2 \\ 1 \end{pmatrix}$$

sono rispettivamente esempi di un vettore riga di lunghezza 4 e di un vettore colonna di lunghezza 5. □

Le righe e le colonne di una matrice A si dicono anche *vettori riga* e *vettori colonna* di A, rispettivamente. A volte, la riga $i-$esima e la colonna $j-$esima di A si denotano rispettivamente con $\boldsymbol{a}_{(i)}$ e $\boldsymbol{a}^{(j)}$; si pone, cioè,

$$\boldsymbol{a}_{(i)} := (a_{i1}, a_{i2}, \ldots, a_{1n})\ ,\ \ \boldsymbol{a}^{(j)} := \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

ed è chiaro che $\boldsymbol{a}_i$ si può pensare come un elemento di $\mathbb{F}^n$ e $\boldsymbol{a}^j$ come uno di $\mathbb{F}^m$. In particolare, una matrica $A = (a)$ di tipo 1×1 sul campo $\mathbb{F}$ può identificarsi con l'elemento a di $\mathbb{F}$.

Nel seguito, useremo il termine di *linea* di A per indicare indifferentemente una riga o una colonna della matrice A. Due linee dello stesso tipo, cioè due righe o due colonne, si dicono anche *parallele*.

Esempio 3.1.3. Con riferimento alle due matrici dell'esempio 3.1.1 si ha:

$$\boldsymbol{a}_{(2)} = \begin{pmatrix} 3 & -1 & 9 & 12 \end{pmatrix} \ , \ \boldsymbol{b}_{(4)} = \begin{pmatrix} 2 & 22 & 6 \end{pmatrix}$$

e

$$\boldsymbol{a}^{(4)} = \begin{pmatrix} 2 \\ 12 \\ 6 \end{pmatrix} \ , \ \boldsymbol{b}^{(1)} = \begin{pmatrix} 9 \\ 5 \\ 0 \\ 2 \end{pmatrix}.$$

□

La matrice di tipo $n \times m$ che si ottiene dalla (3.2) scambiando le righe con le colonne si chiama *matrice trasposta* della matrice $A = (a_{ij})$ e si denota con A^t; si pone cioè

$$A^t = (a_{ji}) = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}. \tag{3.4}$$

Per la trasposta della matrice A si usano anche le notazioni ${}^tA$ e A^T. È immediato verificare che la trasposta di A^t è A, cioè

$$(A^t)^t = A. \tag{3.5}$$

La funzione $A \rightarrow A^t$, che ad ogni matrice associa la propria trasposta, si chiama *trasposizione*.

Esempio 3.1.4. La trasposta della matrice

$$A = \begin{pmatrix} 1 & 5 & 7 & 2 \\ 3 & 1 & 9 & 12 \\ 5 & 11 & 3 & 6 \end{pmatrix}$$

di tipo 3×4 è la matrice

$$A^t = \begin{pmatrix} 1 & 3 & 5 \\ 5 & 1 & 11 \\ 7 & 9 & 3 \\ 2 & 12 & 6 \end{pmatrix}$$

di tipo 4×3. □

Una matrice ottenuta eliminando qualche linea di una fissata matrice A si dice *sottomatrice di A*.

Esempio 3.1.5. Le matrici

$$A' = \begin{pmatrix} 1 & 5 & 2 \\ 3 & -1 & 12 \\ 5 & 11 & 6 \end{pmatrix} \quad , \quad B' = \begin{pmatrix} 1 & -4 \\ 22 & 6 \end{pmatrix}.$$

sono sottomatrici rispettivamente delle matrici A e B dell'esempio 3.1.1. A' si ottiene eliminando la terza colonna di A. B' si ottiene eliminando la prima colonna e la prima e la terza riga di B. $\square$

Una matrice con numero di righe pari a quello delle colonne, cioè $m = n$, si dice *quadrata di ordine n*. In una tale matrice, l'$n-$pla degli elementi che si trovano nei posti corrispondenti a coppie (i, i) di indici uguali si chiama *diagonale principale* o, più semplicemente, *diagonale*. Si chiama, invece, *diagonale secondaria*, l'$n-$pla degli elementi che si trovano nei posti corrispondenti alle coppie di indici (i, j) con $i + j = n + 1$. Una matrice quadrata che abbia uguali a zero tutti gli elementi non appartenenti alla diagonale principale si dice *matrice diagonale*.

Esempio 3.1.6. Quelle che seguono sono esempi di matrici quadrate d'ordine rispettivamente 2, 3 e 4:

$$A = \begin{pmatrix} -1 & 3 \\ 8 & 11 \end{pmatrix} \quad , \quad B = \begin{pmatrix} 1 & 3 & 5 \\ 5 & 1 & 11 \\ 7 & 9 & 3 \end{pmatrix} \quad , \quad C = \begin{pmatrix} -6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 14 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Le diagonali di A e B sono rispettivamente $(-1, 11)$ e $(1, 1, 3)$. La diagonale secondaria di B è $(5, 1, 7)$. La matrice C è una matrice diagonale. $\square$

3.2 Addizione e moltiplicazione per scalari

Denotiamo con $\mathbb{F}^{m,n}$ l'insieme di tutte le matrici d'ordine $m \times n$ sul campo $\mathbb{F}$.

Per ogni matrice $A = (a_{ij}) \in \mathbb{F}^{m,n}$, diciamo *matrice opposta di A*, e la denotiamo con $-A$, quella definita da

$$-A = (-a_{ij});$$

inoltre, diciamo *matrice nulla*, e la denotiamo con $\mathbf{0}$, quella che ha tutti gli elementi uguali a zero, cioè

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Nell'insieme $\mathbb{F}^{m,n}$ si definiscono un'operazione interna di *addizione* fra matrici ed una esterna di *moltiplicazione* di uno scalare per una matrice, nel seguente modo:

- addizione di due matrici: per ogni $A, B \in \mathbb{F}^{m,n}$,

$$A + B = (a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij}); \tag{3.6}$$

- moltiplicazione di uno scalare per una matrice: per ogni $A \in \mathbb{F}^{m,n}$ e $\alpha \in \mathbb{F}$,

$$\alpha A = \alpha(a_{ij}) = (\alpha a_{ij}). \tag{3.7}$$

Esempio 3.2.1. Assegnate le matrici

$$A = \begin{pmatrix} 1 & 5 & 0 \\ 3 & -1 & 9 \end{pmatrix} \ , \ B = \begin{pmatrix} 9 & 0 & 5 \\ 5 & 1 & -4 \end{pmatrix},$$

risulta

$$A + B = \begin{pmatrix} 1+9 & 5+0 & 0+5 \\ 3+5 & -1+1 & 9-4 \end{pmatrix} = \begin{pmatrix} 10 & 5 & 5 \\ 8 & 0 & 5 \end{pmatrix}$$

e

$$2A = \begin{pmatrix} 2\cdot 1 & 2\cdot 5 & 2\cdot 0 \\ 2\cdot 3 & 2\cdot(-1) & 2\cdot 9 \end{pmatrix} = \begin{pmatrix} 2 & 10 & 0 \\ 6 & -2 & 18 \end{pmatrix} \ ,$$

$$-3B = \begin{pmatrix} -27 & 0 & -15 \\ -15 & -3 & 12 \end{pmatrix}.$$

□

Esempio 3.2.2. Assegnate le matrici

$$A = \begin{pmatrix} -1 & 3 & 0 \\ 8 & 11 & -1 \\ 2 & -4 & 8 \end{pmatrix} \ , \ B = \begin{pmatrix} 1 & 3 & 5 \\ 5 & 1 & 11 \\ 7 & 9 & 3 \end{pmatrix},$$

risulta

$$A + B = \begin{pmatrix} 0 & 6 & 5 \\ 13 & 12 & 10 \\ 9 & 5 & 11 \end{pmatrix}$$

e

$$0,5\,A = \begin{pmatrix} -0,5 & 1,5 & 0 \\ 4 & 5,5 & -0,5 \\ 1 & -2 & 4 \end{pmatrix} \ , \ -2\,B = \begin{pmatrix} -2 & -6 & -10 \\ -10 & -2 & -22 \\ -14 & -18 & -6 \end{pmatrix}.$$

□

Le operazioni definite dalle (3.6) e (3.7) verificano le seguenti proprietà:

$SV1$. *proprietà associativa dell'addizione*:
per ogni $A, B, C \in \mathbb{F}^{m,n}$,

$$(A + B) + C = A + (B + C); \tag{3.8}$$

$SV2$ *esistenza dell'elemento neutro per l'addizione*:
per ogni $A \in \mathbb{F}^{m,n}$,

$$A + \mathbf{0} = \mathbf{0} + A = A; \tag{3.9}$$

$SV3$. *simmetrizzabilità rispetto all'addizione*:

per ogni $A \in \mathbb{F}^{m,n}$,
$$A + (-A) = -A + A = \mathbf{0}; \tag{3.10}$$

$SV4$. *Proprietà commutativa dell'addizione*:

per ogni $A, B \in \mathbb{F}^{m,n}$,
$$A + B = B + A; \tag{3.11}$$

$SV5$. *proprietà distributiva della moltiplicazione di una matrice per uno scalare rispetto all'ddizione fra matrici*:

per ogni $\alpha \in \mathbb{F}$ e per ogni $A, B \in \mathbb{F}^{m,n}$,
$$\alpha(A + B) = \alpha A + \alpha B; \tag{3.12}$$

$SV6$. *proprietà distributiva della moltiplicazione di una matrice per uno scalare rispetto all'ddizione fra scalari*:

per ogni $\alpha, \beta \in \mathbb{F}$ e per ogni $A \in \mathbb{F}^{m,n}$,
$$(\alpha + \beta)A = \alpha A + \beta A; \tag{3.13}$$

$SV7$. *proprietà associativa della moltiplicazione di una matrice per uno scalare con la moltiplicazione fra scalari*:

per ogni $\alpha, \beta \in \mathbb{F}$ e per ogni $A \in \mathbb{F}^{n}$,
$$(\alpha\beta)A = \alpha(\beta A); \tag{3.14}$$

$SV8$. per ogni $A \in \mathbb{F}^{m,n}$,
$$1\,A = A. \tag{3.15}$$

Il fatto che in $\mathbb{F}^{m,n}$ le operazioni di addizione e moltiplicazione appena definite verificano le precedenti otto proprietà si esprime dicendo che $\mathbb{F}^{m,n}$ è uno *spazio vettoriale sul campo* $\mathbb{F}$.

Se $A_1, A_2, \ldots, A_k$ sono elementi di $\mathbb{F}^{m,n}$ e $\alpha_1, \alpha_2, \ldots, \alpha_k$ scalari, la matrice definita da
$$\alpha_1 A_1 + \alpha_2 A_2 + \cdots + \alpha_k A_k \tag{3.16}$$
si chiama *combinazione lineare* di $A_1, A_2, \ldots, A_k$ mediante gli scalari α_1, α_2, $\ldots$, α_k.

Esempio 3.2.3. Assegnate le matrici
$$A = \begin{pmatrix} 1 & 5 & 0 \\ 3 & -1 & 9 \end{pmatrix} \ , \ B = \begin{pmatrix} 9 & 0 & 5 \\ 5 & 1 & -4 \end{pmatrix} \ , \ C = \begin{pmatrix} -25 & 10 & -15 \\ -9 & -6 & 6 \end{pmatrix} ,$$
si ha che C è combinazione lineare di A e B mediante gli scalari 2 e -3; risulta, infatti, $C = 2A - 3B$. □

Esercizio 3.2.4. *Siano $A = (a_{ij}) \in \mathbb{F}^{m,n}$ e $\alpha \in \mathbb{F}$. Provare che vale la seguente proprietà:*

$$\alpha A = \mathbf{0} \;\Rightarrow\; \alpha = 0 \;\; o \;\; A = \mathbf{0}\,. \tag{3.17}$$

Esercizio 3.2.5. *Siano $A = (a_{ij}), B = (b_{ij}) \in \mathbb{F}^{m,n}$ e $\alpha \in \mathbb{F}$. Provare che valgono le seguenti proprietà:*

$$(A+B)^t = A^t + B^t \quad , \quad (\alpha A)^t = \alpha A^t\,. \tag{3.18}$$

Osservazione 3.2.6. Una matrice $A \in \mathbb{F}^{m,n}$ può identificarsi col vettore numerico $\boldsymbol{a} \in \mathbb{F}^{nm}$ ottenuto disponendo le righe di A l'una di seguito alla successiva, partendo dalla prima. Per esempio, la matrice

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

viene identificata col vettore numerico

$$\boldsymbol{a} = (a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}).$$

In questa identificazione, alla somma fra matrici e al prodotto di uno scalare per una matrice corrispondono le analoghe operazioni tra i vettori associati. Abbiamo così che $F^{m,n}$ e F^{mn} possono considerarsi come uno stesso spazio vettoriale[1]. □

L'osservazione precedente permette di estendere in modo ovvio agli spazi vettoriali di matrici le nozioni e i risultati dei paragrafi 2.2 e 2.3 relativi alla dipendenza lineare e ai sottospazi in $\mathbb{F}^n$.

3.3 Moltiplicazione righe per colonne

Se $A = (a_{ij}), B = (b_{ij})$ sono due matrici su $\mathbb{F}$, rispettivamente di tipo $m \times n$ e $n \times p$, la matrice

$$AB = (c_{ij})$$

di tipo $m \times p$ definita da

$$c_{ij} = \sum_{h=1}^{n} a_{ih} b_{hj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} \tag{3.19}$$

prende il nome di *prodotto* (*righe per colonne*) *di A e B*. Si noti che l'elemento di posto (i,j) della matrice AB è il prodotto scalare della $i-$esima riga $\boldsymbol{a}_{(i)}$ di A e della $j-$ma colonna $\boldsymbol{b}^{(j)}$ di B, cioè

$$AB = \left(\boldsymbol{a}_{(i)}\boldsymbol{b}^{(j)}\right) = \begin{pmatrix} \boldsymbol{a}_{(1)}\boldsymbol{b}^{(1)} & \boldsymbol{a}_{(1)}\boldsymbol{b}^{(2)} & \cdots & \boldsymbol{a}_{(1)}\boldsymbol{b}^{(p)} \\ \boldsymbol{a}_{(2)}\boldsymbol{b}^{(1)} & \boldsymbol{a}_{(2)}\boldsymbol{b}^{(2)} & \cdots & \boldsymbol{a}_{(2)}\boldsymbol{b}^{(p)} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{a}_{(m)}\boldsymbol{b}^{(1)} & \boldsymbol{a}_{(m)}\boldsymbol{b}^{(1)} & \cdots & \boldsymbol{a}_{(m)}\boldsymbol{b}^{(p)} \end{pmatrix}. \tag{3.20}$$

[1]Questo concetto sarà chiarito e reso preciso nel seguito con l'introduzione della nozione di *isomorfismo*.

ESEMPIO 3.3.1. Assegnate le matrici

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

risulta

$$AB = \begin{pmatrix} 5 & 0 & 3 \\ 0 & 5 & -2 \end{pmatrix} \ , \ AC = \begin{pmatrix} 3 & -7 \\ -2 & 3 \end{pmatrix} .$$

Osserviamo che non è definito il prodotto BA, mentre lo è il prodotto CA e risulta

$$CA = \begin{pmatrix} 5 & -5 \\ -2 & 1 \end{pmatrix} .$$

Osserviamo ancora che è $AC \neq CA$. □

ESERCIZIO 3.3.2. *Provare che:*

- *se una matrice A ha la $i-$esima riga nulla, allora ogni matrice prodotto del tipo AB ha ancora la $i-$esima riga nulla;*
- *se una matrice A ha la $j-$esima colonna nulla, allora ogni matrice prodotto del tipo CA ha ancora la $j-$esima colonna nulla.*

ESEMPIO 3.3.3. Assegnate le matrici

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

risulta

$$AB = \begin{pmatrix} -1 & 11 & 3 \\ 0 & 0 & 0 \end{pmatrix} \ , \ BC = \begin{pmatrix} 0 & -1 \\ 0 & -4 \end{pmatrix} .$$

□

La moltiplicazione fra matrici definita dalla (3.19) verifica le seguenti proprietà:

- *prima proprietà associativa*: per ogni A, B, C di tipo rispettivamente $m \times n$, $n \times p$, $p \times q$,
$$(AB)C = A(BC); \tag{3.21}$$
- *seconda proprietà associativa*: per ogni scalare α e per ogni A, B di tipo rispettivamente $m \times n$, $n \times p$,
$$\alpha(AB) = (\alpha A)B = A(\alpha B); \tag{3.22}$$
- *proprietà distributive (destra e sinistra) rispetto all'ddizione fra matrici*: per ogni A, B, C, D di tipo rispettivamente $m \times n$, $m \times n$, $n \times p$, $p \times m$,
$$(A + B)C = AC + BC \ , \ D(A + B) = DA + DB; \tag{3.23}$$

- per ogni A, B di tipo rispettivamente $m \times n$ e $n \times p$:

$$(AB)^t = B^t A^t. \tag{3.24}$$

Per ogni intero positivo n, si chiama *matrice identica*, o *unitaria* o *identità*, la matrice diagonale d'ordine n avente tutti gli elementi della diagonale principale uguali ad 1. Tale matrice si denota con I_n e possiamo scrivere

$$I_n = (\delta_{ij}) \ , \tag{3.25}$$

ove δ_{ij}, detto *simbolo di Kronecker*, è definito da

$$\delta_{ij} = \begin{cases} 1 & \text{se } i = j \,, \\ 0 & \text{altrimenti.} \end{cases} \tag{3.26}$$

Quando n è chiaro dal contesto, si scrive anche I invece di I_n. È immediato verificare che, per ogni matrice A di tipo $m \times n$, risulta

$$AI_n = I_m A = A \,. \tag{3.27}$$

Esempio 3.3.4. Per $n = 1, 2, 3, 4$ le matrici identità sono le seguenti:

$$I_1 = (1) \ , \ I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \ , \ I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \ , \ I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

□

Osservazione 3.3.5. La matrice diagonale avente uno stesso elemento λ sulla diagonale principale si ottiene moltiplicando lo scalare λ per la matrice identità, cioè

$$\begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{pmatrix} = \lambda \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} .$$

Questo implica che moltiplicare per λ una matrice A di tipo $m \times n$ equivale a moltiplicare A a sinistra per λI_m, o a destra per λI_n, cioè

$$\lambda A = (\lambda I_m) A = A(\lambda I_n).$$

□

Se A, B sono due matrici quadrate d'ordine n, le matrici prodotto AB e BC sono ancora quadrate d'ordine n. Il prodotto fra matrici definisce, dunque, un'operazione interna all'insieme $\mathbb{F}^{n,n}$ che, come abbiamo notato alla fine dell'esempio 3.3.1, non è commutativa. Il fatto che nell'insieme $\mathbb{F}^{n,n}$ le operazioni di addizione e di moltiplicazione righe per colonne verificano le proprietà $SV1$, $SV2$, $SV3$, $SV4$, (3.21), (3.23) e (3.27), si esprime dicendo che, rispetto a queste due operazioni, $\mathbb{F}^{n,n}$ é un *anello unitario*, l'unità essendo la matrice I_n.

Osservazione 3.3.6. È opportuno notare esplicitamente che quando consideriamo $\mathbb{F}^{n,n}$ come spazio vettoriale su $\mathbb{F}$ ci riferiamo alle operazioni di somma fra matrici e di moltiplicazione di uno scalare per una matrice. Quando, invece, consideriamo $\mathbb{F}^{n,n}$ come anello ci riferiamo alle operazioni di addizione e di moltiplicazione righe per colonne fra matrici. □

L'anello $\mathbb{F}^{n,n}$, per quanto precedentemente osservato, non è commutativo. Due matrici $A, B \in \mathbb{F}^{n,n}$ tali che $AB = BA$ si dicono *permutabili*, o che *commutano*.

Esempio 3.3.7. Le due matrici

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ e } B = \begin{pmatrix} 3 & 2 \\ 0 & 3 \end{pmatrix}$$

sono permutabili. Risulta, infatti,

$$AB = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 5 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = BA\,.$$

□

Se k è un intero non negativo e A una matrice quadrata d'ordine n, definiamo *potenza $k-$esima* di A la matrice definita per ricorrenza nel seguente modo

$$A^k = \begin{cases} I_n & \text{se } k = 0\,, \\ AA^{k-1} & \text{se } k > 0. \end{cases} \tag{3.28}$$

Per definizione, la potenza $k-$esima di A è quadrata e dello stesso ordine n di A.

Esempio 3.3.8. Si consideri la matrice

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

Risulta:

$$A^2 = \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}\,,\; A^3 = \begin{pmatrix} -2 & 2 \\ -2 & -2 \end{pmatrix}\,,\; A^4 = \begin{pmatrix} -4 & 0 \\ 0 & -4 \end{pmatrix}.$$

□

Esercizio 3.3.9. *Provare che, per una matrice diagonale con elementi diagonali $a_1, a_2, \dots, a_n$, risulta*

$$\begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}^k = \begin{pmatrix} a_1^k & 0 & \cdots & 0 \\ 0 & a_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n^k \end{pmatrix}, \tag{3.29}$$

per ogni intero $k \geq 0$.

3.4 Matrici invertibili

Una matrice quadrata non nulla $A \in \mathbb{F}^{n,n}$ si dice *invertibile*, o *non singolare*, se esiste una matrice $B \in \mathbb{F}^{n,n}$ tale che

$$AB = BA = I_n; \tag{3.30}$$

in tale ipotesi la matrice B si dice *inversa* di A.

Proposizione 3.4.1. *Una matrice invertibile $A \in \mathbb{F}^{n,n}$ ha un'unica matrice inversa.*

Dimostrazione. Se B, C sono due matrici inverse di A, risulta

$$B = BI_n = B(AC) = (BA)C = I_nC = C\,,$$

così è $B = C$ e l'asserto è provato. □

La matrice inversa di una matrice invertibile $A \in \mathbb{F}^{n,n}$ si denota co A^{-1}; essa è per definizione a sua volta invertibile e risulta

$$(A^{-1})^{-1} = A\,. \tag{3.31}$$

Esempio 3.4.2. Assegnate le matrici

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} \quad , \quad B = \begin{pmatrix} -0,6 & 0,8 \\ 0,4 & -0,2 \end{pmatrix},$$

si ha che $AB = BA = I$ e, quindi, A è invertibile e $A^{-1} = B$. □

Esempio 3.4.3. La matrice identità I_n è invertibile e risulta

$$I_n^{-1} = I_n. \tag{3.32}$$

Più in generale, una matrice diagonale con gli elementi $a_1, a_2, \ldots, a_n$ tutti diversi da zero è invertibile e risulta

$$\begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}^{-1} = \begin{pmatrix} a_1^{-1} & 0 & \cdots & 0 \\ 0 & a_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n^{-1} \end{pmatrix}. \tag{3.33}$$

Per esempio:

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -7 \end{pmatrix}^{-1} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & -\frac{1}{7} \end{pmatrix}.$$

□

Non tutte le matrici quadrate sono invertibili, come prova il seguente esempio.

Esempio 3.4.4. Una matrice quadrata A avente una linea nulla non è invertibile (*cfr.eserc.*3.3.2). Per esempio, se A ha una riga nulla, moltiplicando A per una qualsiasi altra matrice quadrata B dello stesso ordine, si ottiene una matrice che ha ancora una riga nulla e, quindi, non potrà mai aversi $AB = I$. □

Proposizione 3.4.5. *Siano A, B due matrici quadrate invertibili dello stesso ordine. Allora le matrici prodotto AB e BA sono invertibili e risulta*

$$(AB)^{-1} = B^{-1}A^{-1} \ , \ (BA)^{-1} = A^{-1}B^{-1}. \tag{3.34}$$

Dimostrazione. Risulta

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

e, allo stesso modo, si prova che

$$(B^{-1}A^{-1})(AB) = I \, , \ (BA)(A^{-1}B^{-1}) = (A^{-1}B^{-1})(BA) = I;$$

cioè l'asserto. □

Proposizione 3.4.6. *L'insieme delle matrici quadrate invertibili d'ordine n sul campo $\mathbb{F}$, rispetto alla moltiplicazione righe per colonne, costituisce un gruppo.*

Il gruppo delle matrici quadrate invertibili d'ordine n su $\mathbb{F}$, definito dal teorema precedente, si chiama *gruppo (generale) lineare di grado n su* $\mathbb{F}$ e si denota con $GL(n, F)$.

3.5 Operazioni e matrici elementari

Sia $A = (a_{ij})$ una matrice su $\mathbb{F}$ di tipo $m \times n$. Le operazioni di seguito definite si dicono *operazioni elementari sulle righe di A* :

- *primo tipo:* lo scambio di due righe, che consiste nel sostituire la matrice A con quella ottenuta scambiando tra loro due fissate righe di A (lo scambio della riga $i-$esima con la $j-$esima si denota con $R_i \leftrightarrow R_j$);

- *secondo tipo:* la moltiplicazione di una riga per uno scalare, che consiste nel sostituire la matrice A con quella ottenuta moltiplicando per uno stesso scalare gli elementi di una fissata riga di A (la moltiplicazione della riga $i-$esima per lo scalare λ si denota con $R_i \to \lambda R_i$);

- *terzo tipo:* La somma ad una riga di un'altra riga moltiplicata per uno scalare che consiste nel sostituire la matrice A con quella ottenuta sommando ad una fissata riga di A un'altra riga moltiplicata per uno scalare (l'addizione alla riga $i-$esima della riga $j-$esima moltiplicata per lo scalare λ si denota con $R_i \to R_i + \lambda R_j$).

In modo analogo si definiscono le *operazioni elementari sulle colonne di* A. È chiaro che queste ultime operazioni corrispondono ad operazioni elementari sulle righe della matrice A^t trasposta di A. Per questo motivo, nel seguito, prenderemo in considerazione quasi esclusivamente operazioni elementari sulle righe e, per brevità, parleremo semplicemente di operazioni elementari, omettendo il riferimento alle righe.

Esempio 3.5.1. Se alla matrice

$$A = \begin{pmatrix} 0 & 6 & 5 & 5 \\ 3 & -2 & 1 & -1 \\ 2 & 5 & 4 & 0 \end{pmatrix},$$

applichiamo le operazioni elementari $R_1 \leftrightarrow R_3$ (scambio della prima riga con la terza), $R_2 \to 3R_2$ (moltiplicazione per 3 della seconda riga), $R_2 \to R_2 - 2R_1$ (sottrazione della prima riga moltiplicata per 2 alla seconda riga) otteniamo rispettivamente

$$B = \begin{pmatrix} 2 & 5 & 4 & 0 \\ 3 & -2 & 1 & -1 \\ 0 & 6 & 5 & 5 \end{pmatrix} \ , \ C = \begin{pmatrix} 0 & 6 & 5 & 5 \\ 9 & -6 & 3 & -3 \\ 2 & 5 & 4 & 0 \end{pmatrix} \ ,$$

$$D = \begin{pmatrix} 0 & 6 & 5 & 5 \\ 3 & -14 & -9 & -11 \\ 2 & 5 & 4 & 0 \end{pmatrix}.$$

□

La matrice identità e le matrici che si ottengono dalle operazioni elementari applicate alla matrice identità (d'ordine n) si chiamano *matrici elementari* (d'ordine n); più precisamente queste si definiscono nel seguente modo:

- E_{ij} è la matrice che si ottiene dalla matrice identità mediante l'operazione $R_i \leftrightarrow R_j$.
- $E_i(\lambda)$ è la matrice che si ottiene dalla matrice identità mediante l'operazione $R_i \to \lambda R_i$.
- $E_{ij}(\lambda)$ è la matrice che si ottiene dalla matrice identità mediante l'operazione $R_i \to R_i + \lambda R_j$.

Esempio 3.5.2. Quelle che seguono sono le matrici elementari d'ordine tre che si ottengono rispettivamente dalle operazioni elementari $R_1 \leftrightarrow R_3$, $R_2 \to 3R_2$, $R_2 \to R_2 - 2R_1$:

$$E_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \ , \ E_2(3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \ , \ E_{21}(-2) = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Con riferimento all'esempio 3.5.1, notiamo che si ha

$$A \xrightarrow{R_1 \to R_3} E_{13}A = B\,, \ \ A \xrightarrow{R_2 \to 3R_2} E_2(3)A = C\,, \ \ A \xrightarrow{R_2 \to R_2 - 2R_1} E_{21}(-2)A = D\,.$$

Queste relazioni, come fra poco vedremo, possono generalizzarsi a tutte le trasformazioni elementari. □

Per le matrici elementari risulta

$$E_{ij}E_{ij} = I\,,$$
$$E_i(\lambda)E_i(\lambda^{-1}) = E_i(\lambda^{-1})E_i(\lambda) = I\,,$$
$$E_{ij}(\lambda)E_{ij}(\lambda^{-1}) = E_{ij}(\lambda^{-1})E_{ij}(\lambda) = I.$$

Ne segue che tutte le matrici elementari sono invertibili e si ha:

$$E_{ij}^{-1} = E_{ij}\,,\ \ E_i(\lambda)^{-1} = E_i(\lambda^{-1})\,,\ \ E_{ij}(\lambda)^{-1} = E_{ij}(\lambda^{-1}). \tag{3.35}$$

Le operazioni elementari sulla matrice A possono effettuarsi mediante moltiplicazione delle matrici elementari per A. È, infatti, un esercizio provare il seguente teorema.

Proposizione 3.5.3. *Per ogni matrice A si ha che:*

- *$E_{ij}A$ è la matrice che si ottiene da A mediante l'operazione $R_i \leftrightarrow R_j$, cioè scambiando tra loro la riga $i-$esima e quella $j-$esima.*
- *$E_i(\lambda)A$ è la matrice che si ottiene da A mediante l'operazione $R_i \to \lambda R_i$, cioè moltiplicando per lo scalare λ la riga $i - esima$.*
- *$E_{ij}(\lambda)A$ è la matrice che si ottiene da A mediante l'operazione $R_i \to R_i + \lambda R_j$, cioè addizionando alla $i-$esima riga la $j-$esima moltiplicata per lo scalare λ.* □

3.6 Matrici equivalenti

Due matrici A e B dello stesso ordine $m \times n$ si dicono *equivalenti*, e si scrive $A \sim B$, se B si può ottenere da A mediante un numero finito di operazioni elementari e pertanto, in forza della prop.3.5.3, si ha che $A \sim B$ se, e solo se, esistono h matrici elementari $E_1, E_2, \dots, E_h$ tali che

$$B = E_1E_2\cdots E_hA. \tag{3.36}$$

La relazione appena definita tra le matrici di tipo $m\times n$ è riflessiva, simmetrica e transitiva e, quindi, è un'effettiva relazione d'equivalenza nell'insieme $\mathbb{F}^{m,n}$.

Esempio 3.6.1. Le matrici

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

sono equivalenti. Si ha, infatti

$$\begin{pmatrix} 1 & -2 & -3 \\ 5 & 1 & 0 \\ -4 & -3 & -3 \end{pmatrix} \xrightarrow{R_3 \to -R_3} \begin{pmatrix} 1 & -2 & -3 \\ 5 & 1 & 0 \\ 4 & 3 & 3 \end{pmatrix} \xrightarrow{R_3 \to R_3 + R_1} \begin{pmatrix} 1 & -2 & -3 \\ 5 & 1 & 0 \\ 5 & 1 & 0 \end{pmatrix}$$

$$\xrightarrow{R_3 \to R_3 - R_2} \begin{pmatrix} 1 & -2 & -3 \\ 5 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

o, equivalentemente,

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

$$= E_{32}(-1)E_{31}(1)E_3(-1)A$$

□

Nel seguito denoteremo col simbolo $R_i \to \lambda R_i + \mu R_j$ la trasformazione della matrice A ottenuta sostituendo la riga $i-$esima con la somma della stessa riga moltiplicata per uno scalare λ e della riga $j-$esima moltiplicata per uno scalare μ. È chiaro che $R_i \to \lambda R_i + \mu R_j$ si ottiene applicando consecutivamente le trasformazioni elementari λR_i e $R_i \to \mu R_j$.

Esempio 3.6.2. Con riferimento all'esempio 3.6.1, per provare che A e B sono equivalenti, si può usare anche la seguente successione di trasformazioni:

$$\begin{pmatrix} 1 & -2 & -3 \\ 5 & 1 & 0 \\ -4 & -3 & -3 \end{pmatrix} \xrightarrow{R_3 \to -R_3 + R_1} \begin{pmatrix} 1 & -2 & -3 \\ 5 & 1 & 0 \\ 5 & 1 & 0 \end{pmatrix}$$

$$\xrightarrow{R_3 \to R_3 - R_2} \begin{pmatrix} 1 & -2 & -3 \\ 5 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

□

3.7 Matrici ridotte e algoritmo di Gauss

In una matrice di tipo $m \times n$ si chiama *pivot* della riga $i-$esima, o *$i-$esimo pivot*, il primo elemento non nullo di questa riga. I pivot di una matrice si chiamano anche *elementi principali*. Una matrice

$$R = (r_{ij}) = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{pmatrix}$$

si dice *ridotta* (per righe) se valgono le seguenti due proprietà:

- la riga successiva ad una riga nulla di R è nulla (questo significa che le eventuali righe nulle della matrice occupano le ultime posizioni);

- per $i = 1, 2, \dots m$, se l'$i-$esima riga è non nulla, il suo pivot r_{ij_i} risulta l'unico elemento diverso da zero della matrice

$$\begin{pmatrix} r_{i1} & r_{i2} & \cdots & r_{ij_i} \\ r_{i+1,1} & r_{i+1,2} & \cdots & r_{i+1,j_i} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mj_i} \end{pmatrix},$$

cioè

$$\begin{pmatrix} r_{i1} & r_{i2} & \cdots & r_{ij_i} \\ r_{i+1,1} & r_{i+1,2} & \cdots & r_{i+1,j_i} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mj_i} \end{pmatrix} = \begin{pmatrix} 0 & 0 & \cdots & 0 & r_{ij_1} \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

(questo significa che r_{ij_i} si trova alla destra del pivot della riga precedente).

Esempio 3.7.1. Quelle che seguono sono esempi di matrici ridotte:

$$A = \begin{pmatrix} 3 & -2 & 3,7 & 0 & 9 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 7 & -2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \ B = \begin{pmatrix} 1 & -1 & 7 & 22 & 19 \\ 0 & 2 & 1 & -1 & 0 \\ 0 & 0 & 3 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Non sono, invece, ridotte le matrici:

$$C = \begin{pmatrix} 3 & -2 & 3,7 & 0 & 9 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 7 & -2 \end{pmatrix}, \ D = \begin{pmatrix} 1 & -1 & 7 & 22 & 19 \\ 0 & 2 & 1 & -1 & 0 \\ 0 & 0 & 3 & 0 & -2 \\ 0 & -5 & 0 & 0 & 0 \end{pmatrix}.$$

□

Osservazione 3.7.2. Se a_{hk} è l'$h-$esimo pivot di una matrice $A = (a_{ij})$ e a_{tk} è un elemento diverso da zero al di sotto di a_{hk}, cioè $t > h$, allora la matrice $A' = (a'_{ij})$ ottenuta da A mediante la trasformazione $R_t \to a_{hk}R_t - a_{tk}R_h$ differisce da A solo sulla $t-$esima riga e risulta

$$a'_{tk} = a_{hk}a_{tk} - a_{tk}a_{hk} = 0.$$

Per esempio, nella matrice

$$A = \begin{pmatrix} 3 & -2 & 3,7 & 0 & 9 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 5 & -3 & 7 & -2 \\ 0 & -1 & 0 & 2 & 1 \end{pmatrix},$$

il pivot della terza riga è $a_{32} = 5$ e $a_{42} = -1 \neq 0$. Allora la matrice

$$A' = \begin{pmatrix} 3 & -2 & 3,7 & 0 & 9 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 5 & -3 & 7 & -2 \\ 0 & 0 & -3 & 3 & 7 \end{pmatrix},$$

ottenuta da A mediante la trasformazione $R_4 \to a_{32}R_4 - a_{42}R_3$, cioè $R_4 \to 5R_4 + R_3$, differisce da A solo sulla quarta riga e risulta $a'_{42} = 0$. □

Osservazione 3.7.3. Si considerino una matrice A del tipo

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

e le sue sottomatrici A', A'' definite da:

$$A' = \begin{pmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix}, \quad A'' = \begin{pmatrix} a_{12} & a_{13} & \cdots & a_{1n} \\ a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix}.$$

Se è $a_{11} \neq 0$ e R' è una matrice ridotta equivalente ad A', allora la matrice

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ \vdots & & R' & \\ 0 & & & \end{pmatrix}$$

è ridotta ed equivalente ad A.
Se è $a_{11} = 0$ e R'' è una matrice ridotta equivalente ad A'', allora la matrice

$$\begin{pmatrix} 0 & & \\ 0 & R'' & \\ \vdots & & \\ 0 & & \end{pmatrix}$$

è ridotta ed equivalente ad A. □

Esempio 3.7.4. Assegnata la matrice

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

proviamo a costruire una matrice ridotta equivalente ad A, tenendo conto delle osservazioni 3.7.2 e 3.7.3. Dobbiamo, dunque, trovare una successione di trasformazioni elementari che, applicate una di seguito all'altra, trasformino A in una matrice ridotta. Operiamo nel modo seguente:

- PASSO 1: Usando l'osservazione 3.7.2, riduciamo a zero gli elementi al di sotto del primo pivot a_{11} :

$$A = \begin{pmatrix} 2 & 3 & 3 & 5 \\ 1 & -1 & -3 & -4 \\ 3 & 1 & -1 & 0 \end{pmatrix} \xrightarrow{R_2 \to R_1 - 2R_2} \begin{pmatrix} 2 & 3 & 3 & 5 \\ 0 & 5 & 9 & 13 \\ 3 & 1 & -1 & 0 \end{pmatrix}$$

$$\xrightarrow{R_3 \to 3R_1 - 2R_3} \begin{pmatrix} 2 & 3 & 3 & 5 \\ 0 & 5 & 9 & 13 \\ 0 & 7 & 11 & 15 \end{pmatrix} = B.$$

- PASSO 2: Adesso, sempre usando l'osservazione 3.7.2, riduciamo a zero gli elementi di B al di sotto del secondo pivot $b_{22} = 5$:

$$B = \begin{pmatrix} 2 & 3 & 3 & 5 \\ 0 & 5 & 9 & 13 \\ 0 & 7 & 11 & 15 \end{pmatrix} \xrightarrow{R_3 \to 7R_2 - 5R_3} \begin{pmatrix} 2 & 3 & 3 & 5 \\ 0 & 5 & 9 & 13 \\ 0 & 0 & 8 & 16 \end{pmatrix} = R.$$

- RISULTATO FINALE: La matrice R è ridotta ed equivalente ad A.

Osserviamo che R non è l'unica matrice ridotta equivalente ad A. Se, infatti, applichiamo ad R trasformazioni elementari che conservano la sua proprietà di essere ridotta, otteniamo ancora matrici ridotte equivalenti ad A. È, per esempio, equivalente ad A la matrice

$$R' = \begin{pmatrix} 1 & 1,5 & 1,5 & 2,5 \\ 0 & 1 & 1,8 & 2,6 \\ 0 & 0 & 1 & 2 \end{pmatrix},$$

che si ottiene da R moltiplicando le sue righe rispettivamente per $1/2, 1/5, 1/8$. □

Esempio 3.7.5. Proviamo, in un altro caso particolare, a trasformare una matrice A in una ridotta ad essa equivalente e sia

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

Rispetto all'esempio precedente abbiamo due notevoli differenze:

1. la prima colonna di A è nulla;
2. il pivot della prima riga di A si trova alla destra di altri pivot.

Cerchiamo di eliminare queste differenze. A tale scopo, possiamo ignorare la prima colonna di A (*cfr.* osservazione 3.7.3) e, prima di iniziare la procedura dell'esempio precedente, scambiamo tra loro le prime due righe di A in modo che il primo pivot si trovi il più a sinistra possibile. Operiamo, dunque, nel modo seguente:

- PASSO 1: Scambiamo tra loro la prima e la seconda riga di A :

$$\begin{pmatrix} 0 & 0 & 1 & -1 & 1 & 1 \\ 0 & 5 & 0 & 0 & -3 & 0 \\ 0 & 3 & 1 & 0 & -1 & 2 \\ 0 & 2 & 1 & -2 & 3 & 1 \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{pmatrix} 0 & 5 & 0 & 0 & -3 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \\ 0 & 3 & 1 & 0 & -1 & 2 \\ 0 & 2 & 1 & -2 & 3 & 1 \end{pmatrix} = B.$$

  In B, ora, non c'è alcun pivot a destra del primo.

- PASSO 2: Riduciamo a zero gli elementi di B al di sotto del primo pivot $b_{12} = 5$:

$$B \xrightarrow{R_3 \to 3R_1 - 5R_3} \begin{pmatrix} 0 & 5 & 0 & 0 & -3 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \\ 0 & 0 & -5 & 0 & -4 & -10 \\ 0 & 2 & 1 & -2 & 3 & 1 \end{pmatrix}$$

$$\xrightarrow{R_4 \to 2R_1 - 5R_4} \begin{pmatrix} 0 & 5 & 0 & 0 & -3 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \\ 0 & 0 & -5 & 0 & -4 & -10 \\ 0 & 0 & -5 & 10 & -21 & -5 \end{pmatrix} = C$$

- PASSO 3: Riduciamo a zero gli elementi di C al di sotto del secondo pivot $c_{23} = 1$:

$$C \xrightarrow{R_3 \to 5R_2 + R_3} \begin{pmatrix} 0 & 5 & 0 & 0 & -3 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \\ 0 & 0 & 0 & -5 & 1 & -5 \\ 0 & 0 & -5 & 10 & -21 & -5 \end{pmatrix}$$

$$\xrightarrow{R_4 \to 5R_2 + R_4} \begin{pmatrix} 0 & 5 & 0 & 0 & -3 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \\ 0 & 0 & 0 & -5 & 1 & -5 \\ 0 & 0 & 0 & 5 & -16 & 0 \end{pmatrix} = D.$$

- PASSO 4: Riduciamo a zero gli elementi di D al di sotto del terzo pivot $d_{34} = 1$:

$$C \xrightarrow{R_4 \to -R_- R_4} \begin{pmatrix} 0 & 5 & 0 & 0 & -3 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \\ 0 & 0 & 0 & -5 & 1 & -5 \\ 0 & 0 & 0 & 0 & 15 & 5 \end{pmatrix} = R.$$

- RISULTATO FINALE: La matrice R è ridotta ed equivalente ad A.

Anche questa volta siamo riusciti nel nostro intento. □

A questo punto dovrebbe essere chiaro che il procedimento utilizzato negli ultimi due esempi è del tutto generale e, quindi, se applicato ad una qualsiasi matrice non nulla A, restituisce una matrice ridotta equivalente ad A. Esso è noto come *algoritmo di Gauss* e può descriversi nel modo seguente.

Algoritmo 3.7.6. (Gauss)

- INPUT: Una matrice $A = (a_{ij})$ di tipo $m \times n$ con la prima colonna non nulla[2].

[2] Questa assunzione non è restrittiva perché nel caso contrario, in forza dell'osservazione 3.7.3, possiamo applicare l'algoritmo alla matrice che si ottiene eliminando da A le colonne nulle che occupano le prime posizioni.

- PASSO 1: Se $a_{11} = 0$, si consideri il più piccolo indice i tale che $a_{i1} \neq 0$ e si applichi ad A la trasformazione $R_i \longleftrightarrow R_j$; si scambi cioè la prima riga di A con la $i-$esima. Si ottiene così una matrice $A' = (a'_{ij})$ con $a'_{11} \neq 0$.
- PASSO 2: Per ogni $i = 2, 3, \dots, m$, se $a'_{i1} \neq 0$, si applichi ad A' la trasformazione $R_i \longrightarrow a'_{i1}R_1 - a'_{11}R_i$. Si ottiene così una matrice $A'' = (a''_{ij})$ del tipo
$$A'' = \begin{pmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ 0 & a''_{22} & \cdots & a''_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a''_{m2} & \cdots & a''_{mn} \end{pmatrix}.$$
- PASSO 3: Si consideri il più piccolo indice di colonna j tale che
$$a''_{2j}\,,\, a''_{3j}\,, \dots,\, a''_{mj}$$
non siano tutti nulli e si ritorni al PASSO 1, applicando la procedura alla seguente sottomatrice di A''
$$\overline{A} = \begin{pmatrix} a''_{2j} & a''_{2,j+1} & \cdots & a''_{2n} \\ a''_{3j} & a''_{3,j+1} & \cdots & a''_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ a''_{mj} & a''_{m,j+1} & \cdots & a''_{mn} \end{pmatrix}.$$
Si osservi che la procedura termina dopo al più $m-1$ di queste iterazioni.
- OUTPUT: Una matrice J ridotta equivalente ad A (usando l'osservazione 3.7.3). □

Il teorema che segue è una conseguenza immediata della prop.3.5.3 e dell'algoritmo di Gauss.

Proposizione 3.7.7. *Ogni matrice $A \in \mathbb{F}^{m,n}$ è equivalente ad una matrice ridotta. Esistono, quindi, matrici elementari $E_1, E_2, \dots, E_h$ ed una matrice ridotta J tali che*
$$E_h E_{h-1} \cdots E_1 A = J. \tag{3.37}$$

Dimostrazione. L'algoritmo di Gauss permette di costruire una matrice ridotta J equivalente ad A mediante l'applicazione successiva di operazioni elementari sulle righe di A. Ciascuna di tali operazioni, in forza della prop.3.5.3, ha su A lo stesso effetto della moltiplicazione a sinistra per una matrice elementare e da ciò segue l'asserto. □

3.8 Matrici completamente ridotte

Una matrice ridotta R si dice *completamente ridotta* se ogni suo pivot è uguale ad 1 e se in ogni colonna contenente un pivot tutti gli elementi diversi da questo sono uguali a zero.

Esempio 3.8.1. Quelle che seguono sono esempi di matrici completamente ridotte:

$$A = \begin{pmatrix} 1 & -2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 & 22 & 19 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Non sono, invece, completamente ridotte le matrici:

$$C = \begin{pmatrix} 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 & 0 & 22 & 19 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 2 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Nella matrice C, infatti, il pivot della seconda riga $c_{23} = 1$ occupa la terza colonna e questa contiene anche $c_{13} = 1 \neq 0$. Nella matrice D, invece, il pivot della terza riga è 2 e non 1. □

Una lieve modifica dell'algoritmo di Gauss permette di costruire, a partire da una qualsiasi matrice non nulla A, una matrice completamente ridotta equivalente ad A. Cominciamo a renderci conto di ciò con un paio di esempi.

Esempio 3.8.2. Assegnata la matrice

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

costruiamo una matrice completamente ridotta equivalente ad A. Operiamo nel modo seguente:

- PASSO 1: Usando l'algoritmo 3.7.6 di Gauss, costruiamo una matrice ridotta
$$R = \begin{pmatrix} 2 & 3 & 3 & 5 \\ 0 & 5 & 9 & 13 \\ 0 & 0 & 8 & 16 \end{pmatrix}$$
equivalente ad A (*cfr.* esempio 3.7.4).
- PASSO 2: Riduciamo a zero gli elementi di R che si trovano al di sopra dei pivot:
$$R \xrightarrow{R_1 \to 5R_1 - 3R_2} \begin{pmatrix} 10 & 0 & -12 & -14 \\ 0 & 5 & 9 & 13 \\ 0 & 0 & 8 & 16 \end{pmatrix}$$
$$\xrightarrow{R_1 \to 8R_1 + 12R_3} \begin{pmatrix} 80 & 0 & 0 & 80 \\ 0 & 5 & 9 & 13 \\ 0 & 0 & 8 & 16 \end{pmatrix}$$
$$\xrightarrow{R_2 \to 8R_2 - 9R_3} \begin{pmatrix} 80 & 0 & 0 & 80 \\ 0 & 40 & 0 & -40 \\ 0 & 0 & 8 & 16 \end{pmatrix} = R'.$$

- PASSO 3: Moltiplichiamo ciascuna riga di R' per l'inverso del suo pivot; cioè la prima per 1/80, la seconda per 1/40, e la terza per 1/8. Otteniamo, così , la matrice
$$R'' = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{pmatrix}.$$
- RISULTATO FINALE: La matrice R'', che è completamente ridotta ed equivalente ad A. □

Esempio 3.8.3. Assegnata la matrice
$$A = \begin{pmatrix} 0 & 0 & 1 & -1 & 1 & 1 \\ 0 & 5 & 0 & 0 & -3 & 0 \\ 0 & 3 & 1 & 0 & -1 & 2 \\ 0 & 2 & 1 & -2 & 3 & 1 \end{pmatrix},$$
costruiamo una matrice completamente ridotta equivalente ad A. Operiamo nel modo seguente:

- PASSO 1: Usando l'algoritmo 3.7.6 di Gauss, costruiamo la matrice ridotta
$$R = \begin{pmatrix} 0 & 5 & 0 & 0 & -3 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \\ 0 & 0 & 0 & -5 & 1 & -5 \\ 0 & 0 & 0 & 0 & 15 & 5 \end{pmatrix}$$
equivalente ad A (*cfr.* esempio 3.7.5).
- PASSO 2: Riduciamo a zero gli elementi di R che si trovano al di sopra dei pivot:
$$R \xrightarrow{R_2 \to 5R_2 - R_3} \begin{pmatrix} 0 & 5 & 0 & 0 & -3 & 0 \\ 0 & 0 & 1 & 0 & 4 & 10 \\ 0 & 0 & 0 & -5 & 1 & -5 \\ 0 & 0 & 0 & 0 & 15 & 5 \end{pmatrix}$$
$$\xrightarrow{R_1 \to 15R_1 + 3R_4} \begin{pmatrix} 0 & 75 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 4 & 10 \\ 0 & 0 & 0 & -5 & 1 & -5 \\ 0 & 0 & 0 & 0 & 15 & 5 \end{pmatrix}$$
$$\xrightarrow{R_2 \to 15R_2 - 4R_4} \begin{pmatrix} 0 & 75 & 0 & 0 & 0 & 0 \\ 0 & 0 & 15 & 0 & 0 & 130 \\ 0 & 0 & 0 & -5 & 1 & -5 \\ 0 & 0 & 0 & 0 & 15 & 5 \end{pmatrix}$$
$$\xrightarrow{R_3 \to 15R_3 - R_4} \begin{pmatrix} 0 & 75 & 0 & 0 & 0 & 0 \\ 0 & 0 & 15 & 0 & 0 & 130 \\ 0 & 0 & 0 & -75 & 0 & -80 \\ 0 & 0 & 0 & 0 & 15 & 5 \end{pmatrix} = R'.$$

- PASSO 3: Moltiplichiamo ciascuna riga di R' per l'inverso del suo pivot, se diverso da 1; cioè la prima per 1/75, la seconda per 1/15, la terza per −1/75, e la quarta per 1/15. Otteniamo, così , la matrice
$$R'' = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & \frac{26}{3} \\ 0 & 0 & 0 & 1 & 0 & \frac{16}{15} \\ 0 & 0 & 0 & 0 & 1 & \frac{1}{3} \end{pmatrix}.$$

- RISULTATO FINALE: La matrice R'', che è completamente ridotta ed equivalente ad A. □

I due esempi precedenti suggeriscono quali integrazioni apportare all'algoritmo di Gauss per ottenerne uno che trasformi una matrice A in una equivalente e completamente ridotta. Descriviamo di seguito questa variazione dell'algoritmo di Gauss, nota come *algoritmo di Gauss-Jordan.*

ALGORITMO 3.8.4. (GAUSS-JORDAN)

- INPUT: Una matrice $A = (a_{ij})$ di tipo $m \times n$ con la prima colonna non nulla[3].

- PASSO 1: Si trasformi la matrice A nella matrice ridotta $R = (r_{ij})$ mediante l'algoritmo 3.7.6 di Gauss.

- PASSO 2: Per ogni pivot r_{ij}, con $i > 1$, e per ogni $s = 1, 2, \dots, i-1$, si applichi ad R la trasformazione $R_s \longrightarrow r_{ij}R_s - r_{sj}R_i$. Si ottiene così una matrice ridotta $C = (c_{ij})$ che, al di sopra di ogni pivot, la tutti gli elementi uguali a zero.

- PASSO 3: Per ogni pivot c_{ij} della matrice C, si applichi a C la trasformazione $R_i \longrightarrow c_{ij}^{-1}R_i$.

- OUTPUT: Una matrice J completamente ridotta equivalente ad A. □

La prop.3.5.3 e l'algoritmo 3.8.4 di Gauss-Jordan possono sintetizzarsi nel seguente teorema.

PROPOSIZIONE 3.8.5. *Ogni matrice $A \in \mathbb{F}^{m,n}$ è equivalente ad una matrice completamente ridotta. Esistono, quindi, matrici elementari $E_1, E_2, \dots, E_h$ ed una matrice completamente ridotta J tali che*
$$E_h E_{h-1} \cdots E_1 A = J. \tag{3.38}$$

DIMOSTRAZIONE. L'algoritmo di Gauss-Jordan permette di costruire una matrice completamente ridotta J equivalente ad A mediante l'applicazione successiva di operazioni elementari sulle righe di A. Ciascuna di tali operazioni, in forza della prop.3.5.3, ha su A lo stesso effetto della moltiplicazione a sinistra per una matrice elementare e da ciò segue l'asserto. □

[3]Si osservi che, anche qui, questa assunzione non è restrittiva, in forza dell'osservazione 3.7.3.

3.9 Riduzione di matrici quadrate e calcolo dell'inversa

Una matrice quadrata $T = (t_{ij})$ si dice *triangolare* se sono nulli tutti i suoi elementi che si trovano al di sotto, o al di sopra, della diagonale principale, cioè se T è del tipo

$$T = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1,n-1} & t_{1n} \\ & t_{22} & \cdots & t_{2,n-1} & t_{2n} \\ & & \ddots & \vdots & \vdots \\ & 0 & & t_{n-1,n-1} & t_{n-1,n} \\ & & & & t_{nn} \end{pmatrix}$$

o del tipo

$$T = \begin{pmatrix} t_{11} & & & & \\ t_{21} & t_{22} & & 0 & \\ \vdots & \vdots & \ddots & & \\ t_{n-1,1} & t_{n-1,2} & \cdots & t_{n-1,n-1} & \\ t_{n1} & t_{n2} & \cdots & t_{n,n-1} & t_{nn} \end{pmatrix}$$

Nel primo caso T si dice *triangolare superiore*, o *sopratriangolare*, e nel secondo *triangolare inferiore*, o *sottotriangolare*. Notiamo esplicitamente che un elemento della diagonale principale di una matrice triangolare può essere nullo.

Esempio 3.9.1. Le matrici elementari $E_{ij}(\lambda)$ sono triangolari inferiori per $i > j$ e triangolari superiori per $i < j$. Le matrici diagonali sono triangolari sia superiori che inferiori. □

Esempio 3.9.2. Quelle che seguono sono esempi di matrici triangolari.

$$A = \begin{pmatrix} 1 & 5 & 2 & -8 & 11 \\ 0 & \sqrt{2} & 0 & 4 & 0 \\ 0 & 0 & 12 & 0 & 22 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix},$$

$$B = \begin{pmatrix} 1 & 5 & 2 & -8 \\ 0 & \sqrt{2} & 0 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, C = \begin{pmatrix} 1 & 5 & 2 & -8 \\ 0 & \sqrt{2} & 0 & 4 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$G = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ \sqrt{2} & -1 & 0 & 0 & 0 \\ 3 & -5 & 32 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 8 & 0 & -9 & 0 \end{pmatrix}, H = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ \sqrt{2} & -1 & 0 & 0 & 0 \\ 3 & -5 & 32 & 0 & 0 \\ 3 & 3 & 1 & 4 & 0 \\ 0 & 1 & 0 & -9 & 7 \end{pmatrix}$$

Le matrici A, B, C sono sopratriangolari e di queste A e C sono anche ridotte. La matrice B non è ridotta perché ha la terza riga nulla e la quarta non nulla. $\square$

Osservazione 3.9.3. Nelle matrici sopratriangolari la seconda proprietà delle matrici ridotte è sempre verificata. Ne segue che una matrice sopratriangolare la cui diagonale contiene solo elementi non nulli è ridotta. $\square$

Esercizio 3.9.4. *Provare che:*

- *il prodotto di due matrici sopratriangolari d'ordine n è una matrice sopratriangolare d'ordine n;*
- *il prodotto di due matrici sottotriangolari d'ordine n è una matrice sottotriangolare d'ordine n;*
- *il prodotto di due matrici diagonali d'ordine n è una matrice diagonale d'ordine n.*

Sia $A \in \mathbb{F}^{n,n}$ una matrice quadrata d'ordine n. In queste ipotesi le matrici che intervengono nelle relazioni (3.37) e (3.38) sono tutte quadrate d'ordine n. La matrice J, inoltre, risulta di forma triangolare superiore.

Le (3.37) e (3.38) permettono di controllare se A è invertibile e, se questo è il caso, di costruire un algoritmo per il calcolo dell'inversa di A.

Proposizione 3.9.5. *Per una matrice $A \in \mathbb{F}^{n,n}$ sono equivalenti le seguenti proprietà:*

(i) *A è prodotto di matrici elementari;*

(ii) *A è invertibile;*

(iii) *A è equivalente ad una matrice ridotta senza righe nulle;*

(iv) *A è equivalente alla matrice identità; esistono, cioè, matrici elementari $E_1, E_2, \ldots, E_h$ tali che*
$$E_1 E_2 \cdots E_h A = I. \tag{3.39}$$

Proposizione 3.9.6. *Sia $A \in \mathbb{F}^{n,n}$ una matrice quadrata invertibile d'ordine n e*
$$E_1 E_2 \cdots E_h A = I$$
la relazione di cui alla (3.39)*. Allora risulta*
$$A^{-1} = E_1 E_2 \cdots E_h \tag{3.40}$$
e, quindi,
$$E_1 E_2 \cdots E_h (A|I) = (I|A^{-1}). \tag{3.41}$$

La relazione (3.41) sostanzialmente dice che, se si eseguono sulla matrice identità, le operazioni elementari che trasformano una matrice invertibile A in I, si ottiene A^{-1}. Questa osservazione suggerisce il seguente algoritmo per calcolare l'inversa di una matrice invertibile.

Algoritmo 3.9.7. (Calcolo della matrice inversa)

- INPUT: Una matrice $A = (a_{ij})$ quadrata d'ordine n.
- PASSO 1: Si applichino in successione alla matrice (A, I) le trasformazioni elementari che, mediante l'algoritmo 3.8.4 di Gauss-Jordan, trasformano A in una matrice completamente ridotta ottenendo volta per volta una matrice del tipo $(C|D)$ e terminando l'algoritmo se su C si forma una riga nulla (in questo caso A non è invertibile).
- OUTPUT: La matrice $A^{-1} = C$.

Esempio 3.9.8. Controlliamo se la matrice

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

è invertibile e, in caso di risposta affermativa, calcoliamone l'inversa. Applicando l'algoritmo 3.9.7 abbiamo:

$$(A|I) = \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ -1 & -1 & 0 & 1 \end{array}\right) \xrightarrow{R_2 \to R_1 + R_2} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{array}\right)$$

$$\xrightarrow{R_1 \to R_1 - 2R_2} \left(\begin{array}{cc|cc} 1 & 0 & -1 & -2 \\ 0 & 1 & 1 & 1 \end{array}\right) = (I|A^{-1}).$$

Concludiamo che A è invertibile e risulta:

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

□

Esempio 3.9.9. Controlliamo se la matrice

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

è invertibile e, in caso di risposta affermativa, calcoliamone l'inversa. Applicando l'algoritmo 3.9.7 abbiamo (il Lettore individui per esercizio quali trasformazioni si utilizzano):

$$(A|I) = \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 2 & -1 & 3 & 0 & 1 & 0 \\ 4 & 2 & 2 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & -2 & 1 & 0 \\ 4 & 2 & 2 & 0 & 0 & 1 \end{array}\right)$$

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

$$\to \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & -8 & 2 & 1 \end{array}\right) = (C|D).$$

Abbiamo ottenuto una matrice C con una riga nulla e, quindi, A non è invertibile. □

Capitolo 4

Sistemi lineari

4.1 Equazioni lineari

Un'*equazione lineare* nelle n incognite $x_1, x_2, ..., x_n$ a coefficienti in un campo $\mathbb{F}$ è un'equazione del tipo

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b, \tag{4.1}$$

ove $a_1, a_2, \ldots, a_n, b$ sono elementi di $\mathbb{F}$. Gli elementi $a_1, a_2, \ldots, a_n$, si chiamano *coefficienti* e b *termine noto* dell'equazione. L'incognita corrispondente al primo coefficiente diverso da zero, se esiste, prende il nome di *incognita principale* dell'equazione, le altre si dicono *libere* o *secondarie*[1].

Esempio 4.1.1. L'equazione

$$x_1 + 3x_2 + \frac{2}{3}x_3 - 12x_4 + 2x_5 = 7$$

è un'equazione lineare a coefficienti reali nelle cinque incognite x_1, x_2, x_3, x_4, x_5. L'incognita principale è x_1, le incognite libere sono x_2, x_3, x_4,x_5, i coefficienti sono $1, 3, \frac{2}{3}, -12, 2$, il termine noto è 7. □

Una *soluzione dell'equazione* (4.1) è un'$n-$pla $\boldsymbol{c} = (c_1, c_2, \ldots, c_n)$ di elementi di $\mathbb{F}$, cioè un vettore numerico elemento di $\mathbb{F}^n$, tale che

$$a_1c_1 + a_2c_2 + \cdots + a_nc_n = b. \tag{4.2}$$

Esempio 4.1.2. Si consideri l'equazione a coefficienti reali

$$x_1 + x_2 - 2x_3 = 4.$$

Le terne $(4, 2, 1)$ e $(0, 0, -2)$ sono soluzioni dell'equazione; infatti, si ha

$$4 - 2 - 2 = 4 \quad \text{e} \quad 0 + 0 - 2(-2) = 4.$$

[1] È chiaro che queste definizioni dipendono dall'ordine con cui si assegnano le incognite.

Da notare che la nostra equazione ammette un numero infinito di soluzioni; infatti, per ogni terna del tipo $(4-\lambda+2\mu, \lambda, \mu)$, con λ, μ numeri reali arbitrari, risulta:

$$(4-\lambda+2\mu)+\lambda-2\mu=4.$$

Le terne $(1,1,-3)$ e $(0,0,0)$, invece, non sono soluzioni dell'equazione perchè

$$1+1-2(-3)\neq 4 \quad \text{e} \quad 0+0+0\neq 4\,.$$

□

OSSERVAZIONE 4.1.3. Quando nella (4.1) tutti i coefficienti ed il termine noto sono nulli, cioè

$$a_1=a_2=\cdots=a_n=b=0,$$

l'equazione si dice *identità* e, in questo caso, ogni elemento di $\mathbb{F}^n$ è una sua soluzione. □

Due equazioni lineari si dicono *equivalenti* se hanno lo stesso insieme di soluzioni.

PROPOSIZIONE 4.1.4. *L'equazione*

$$(\lambda a_1)x_1+(\lambda a_2)x_2+\cdots+(\lambda a_n)x_n=\lambda b,$$

ottenuta dalla (4.1) *moltiplicando tutti i coefficienti ed il termine noto per uno stesso scalare* $\lambda\in\mathbb{F}$ *diverso da zero, è equivalente alla* (4.1).

DIMOSTRAZIONE. Se $\boldsymbol{c}=(c_1,c_2,\ldots,c_n)$ è un'$n-$pla di elementi di $\mathbb{F}$, si ha

$$\begin{array}{rcl} a_1c_1+a_2c_2+\cdots+a_nc_n &=& b \;\Leftrightarrow \\ \lambda(a_1c_1+a_2c_2+\cdots+a_nc_n) &=& \lambda b \;\Leftrightarrow \\ (\lambda a_1)c_1+(\lambda a_2)c_2+\cdots+(\lambda a_n)c_n &=& \lambda b \end{array}$$

cioè l'asserto. □

Due equazioni lineari che si ottengono l'una dall'altra moltiplicando tutti i coefficienti ed il termine noto per uno stesso scalare non nullo si dicono *proporzionali.*

ESEMPIO 4.1.5. Le equazioni a coefficienti reali

$$x_1+2x_2-x_3=3 \quad \text{e} \quad -2x_1-4x_2+2x_3=-6$$

sono equivalenti in quanto la seconda si ottiene moltiplicando i coefficienti ed il termine noto della prima per -2. Non sono, invece, equivalenti le due equazioni

$$x_1+x_2=0 \quad \text{e} \quad x_1-x_2=0\,;$$

la coppia $(1,-1)$, infatti, è una soluzione della prima e non della seconda. □

ESEMPIO 4.1.6. (EQUAZIONE LINEARE IN UNA INCOGNITA) È facile verificare che per l'equazione (caso $n = 1$ della (4.1))

$$a_1 x_1 = b$$

valgono le seguenti proprietà:

- se $a_1 \neq 0$, l'equazione ha l'unica soluzione

$$x_1 = \frac{b}{a_1} \ ; \tag{4.3}$$

- se $a_1 = b = 0$, l'equazione è un'identità e ha come soluzioni tutti gli elementi di $\mathbb{F}$;

- se $a_1 = 0$ e $b \neq 0$, l'equazione non ha soluzioni. □

PROPOSIZIONE 4.1.7. (SOLUZIONI DI UN'EQUAZIONE LINEARE) *Sia assegnata l'equazione lineare*

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b.$$

Se tutti i coefficienti sono nulli e il termine noto è diverso da zero, allora l'equazione non ha soluzioni. Se almeno uno dei coefficienti è diverso da zero, l'equazione ha soluzioni e ciascuna di queste corrisponde biunivocamente ad un assegnato valore delle incognite libere[2]*; in particolare, quando $n = 1$ la soluzione è unica.*

DIMOSTRAZIONE. Se tutti i coefficienti sono nulli e il termine noto è diverso da zero, l'equazione è

$$0x_1 + 0x_2 + \cdots + 0x_n = b,$$

con $b \neq 0$, e evidentemente è priva di soluzioni. Se $n = 1$ e $a_1 \neq 0$, abbiamo l'unica soluzione $x_1 = \frac{b}{a_1}$.

Supponiamo $n > 1$ e almeno uno dei coefficienti diverso da zero, per esempio $a_1 \neq 0$, cosa che non lede le generalità del problema. Allora ogni specializzazione delle incognite libere

$$x_2 = c_2, \ x_3 = c_3, \ \ldots, \ x_n = c_n$$

dà luogo all'equazione lineare nella sola incognita principale x_1

$$a_1 x_1 = b - a_2 c_2 - \cdots - a_n c_n$$

che, per quanto detto prima, ammette un'unica soluzione

$$c_1 = \frac{b - a_2 c_2 - \cdots - a_n c_n}{a_1}$$

[2]In altre parole, le incognite libere devono essere considerate come dei parametri che variano nel campo $\mathbb{F}$.

e, in questo modo, otteniamo univocamente la soluzione $(c_1, c_2, \dots, c_n)$. Viceversa, ogni soluzione $(d_1, d_2, \dots, d_n)$ della nostra equazione corrisponde univocamente alla specializzazione

$$x_2 = d_2,\ x_3 = d_3, \dots, x_n = d_n$$

delle incognite libere. □

Corollario 4.1.8. *Un'equazione lineare ha per soluzioni tutti gli elementi di $\mathbb{F}^n$ se, e solo se, è un'identità.*

Quando nell'equazione (4.1) è $b = 0$, cioè

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0\,, \tag{4.4}$$

l'equazione si dice *omogenea*. In questo caso abbiamo almeno la soluzione

$$x_1 = x_2 = \cdots = x_n = 0,$$

che è detta *soluzione nulla* o *banale*. Un'equazione del tipo (4.1) individua l'equazione omogenea (4.4), che si dice *associata* a (4.1).

Osservazione 4.1.9. Quando tutte le soluzioni di un'equazione lineare si ottengono in corrispondenza dei valori assegnati ad m (> 0) incognite libere e il campo $\mathbb{F}$ è infinito[3], allora l'insieme delle soluzioni è infinito e si dice che l'equazione ammette "*infinito ad m*" soluzioni (in simboli ∞^m). □

Esempio 4.1.10. L'equazione omogenea a coefficienti reali

$$x_1 + 2x_2 = 0$$

ammette ∞^1 soluzioni date da

$$\{\,(-2\lambda,\, \lambda) \in \mathbb{R}^2 \ :\ \lambda \in \mathbb{R}\,\}.$$

Queste si ottengono assegnando valori arbitrari all'incognita libera, $x_2 = \lambda$, e ricavando, in corrispondenza di questa, l'unico possibile valore dell'incognita principale, $x_1 = -2\lambda$. Osserviamo che, nella nostra equazione, potremmo assumere x_2 come incognita principale e x_1 come incognita libera. In questo caso le soluzioni si esprimono nella forma

$$\{\,(\lambda,\, -\frac{\lambda}{2}) \in \mathbb{R}^2 \ :\ \lambda \in \mathbb{R}\,\}.$$

Naturalmente, qualunque sia la nostra ipotesi sulla scelta dell'incognita principale, l'insieme delle soluzioni dell'equazione rimane sempre lo stesso. □

[3] Come nel caso del campo reale $\mathbb{R}$, del campo razionale Q e del campo complesso $\mathbb{C}$.

Esempio 4.1.11. L'equazione a coefficienti reali

$$x_1 - x_2 + 3x_3 = 3$$

ammette ∞^2 soluzioni date da

$$\{\, (3 + \lambda - 3\mu,\ \lambda,\ \mu) \in \mathbb{R}^3 \ :\ \lambda, \mu \in \mathbb{R}\,\}.$$

Queste si ottengono assegnando valori arbitrari alle due incognite libere, $x_3 = \mu\,, x_2 = \lambda$, e ricavando, in corrispondenza di questi, l'unico possibile valore dell'incognita principale, $x_1 = 3 - \lambda + 3\mu$. Anche qui è possibile scegliere x_2 o x_3 come incognita principale; se, per esempio, scegliamo x_2, le soluzioni dell'equazione si scrivono

$$\{\, (\lambda,\ -3 - \lambda + 3\mu,\ \mu) \in \mathbb{R}^3 \ :\ \lambda, \mu \in \mathbb{R}\,\}.$$

□

È il caso di osservare che l'equazione (4.1) dipende soltanto dai suoi coefficienti e dal termine noto, cioè dalla $(n+1)-$pla $(a_1, a_2, \ldots, a_n, b)$, e non dalla notazione usata per le incognite; per esempio, l'equazione

$$a_1y_1 + a_2y_2 + \cdots + a_ny_n = b$$

nelle incognite $y_1, y_2, ..., y_n$ ha lo stesso insieme di soluzioni della (4.1) e, quindi, è ad essa equivalente. Spesso, quando il numero delle incognite è relativamente piccolo, può convenire denotare le incognite stesse mediante lettere diverse dell'alfabeto; per esempio, invece di

$$a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 = b,$$

si scrive

$$a_1x + a_2y + a_3z + a_4t = b,$$

oppure

$$a_1X + a_2Y + a_3Z + a_4T = b,$$

ove le incognite sono x, y, z, t nel primo caso e X, Y, Z, T nel secondo.

Esempio 4.1.12. Le due equazioni a coefficienti reali

$$3X = -9 \quad \text{e} \quad 2X - Y + \frac{2}{5}Z = 8$$

ammettono entrambe soluzioni. La prima ha l'unica soluzione $X = -3$. La seconda ha ∞^2 soluzioni date da

$$\{\, (\frac{1}{2}\lambda - \frac{1}{5}\mu,\ \lambda,\ \mu) \in \mathbb{R}^3 \ :\ \lambda, \mu \in \mathbb{R}\,\}.$$

4.2 Generalità sui sistemi lineari

Un *sistema lineare* S di m equazioni in n incognite è un insieme di m equazioni del tipo (4.1), cioè

$$S = \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}, \tag{4.5}$$

ove gli a_{ij} e i b_i sono elementi di un campo $\mathbb{F}$, per ogni $i = 1, 2, \ldots, n$ e $j = 1, 2, \ldots, m$. Gli elementi a_{ij} si chiamano *coefficienti* e i b_i *termini noti* del sistema. Il numero delle incognite n si chiama anche *ordine* del sistema.

ESEMPIO 4.2.1. Quello che segue è un sistema a coefficienti reali di tre equazioni nelle quattro incognite x_1, x_2, x_3, x_4.

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 3x_4 = 1 \\ \quad\ - 4x_2 + x_3 + x_4 = 1 \\ 9x_1 - 5x_2 + x_3 \qquad\quad = \frac{1}{2} \end{cases}.$$

□

Assegnato il sistema (4.5), restano definite le due matrici

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \ A' = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix} \tag{4.6}$$

che si chiamano, rispettivamente, *matrice dei coefficienti* (o *incompleta*) e *matrice completa* del sistema. Per $i = 1, 2, \ldots, n$, l'$i-$esimo pivot della matrice A si dice anche $i-$*esimo pivot,* o $i-$*esimo elemento principale,* del sistema. Se si pone

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \ \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \tag{4.7}$$

il sistema (4.5) si può scrivere nella forma

$$\boldsymbol{a}^{(1)}x_1 + \boldsymbol{a}^{(2)}x_2 + \cdots + \boldsymbol{a}^{(2)}x_n = \boldsymbol{b}, \tag{4.8}$$

ove $\boldsymbol{a}^{(1)}, \boldsymbol{a}^{(2)}, \ldots, \boldsymbol{a}^{(n)}$ denotano le colonne della matrice A, o nella forma

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \tag{4.9}$$

o, equivalentemente,

$$A\boldsymbol{x} = \boldsymbol{b}\,. \tag{4.10}$$

Si osservi che la scrittura (4.10) per un sistema lineare è la più generale possibile; in essa non compaiono neanche il numero m delle equazioni e il numero n delle incognite.

Esempio 4.2.2. Consideriamo il seguente sistema a coefficienti reali di tre equazioni nelle quattro incognite x_1, x_2, x_3, x_4.

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

Il pivot della prima equazione è 1(=coefficiente di x_1), quello della seconda è -4(=coefficiente di x_2), quello della terza è 9(=coefficiente di x_1). Le due matrici del sistema, quella incompleta e quella completa, sono rispettivamente

$$A = \begin{pmatrix} 1 & 2 & 3 & 3 \\ 0 & -4 & 1 & 1 \\ 9 & 0 & 1 & -5 \end{pmatrix}, \; A' = \begin{pmatrix} 1 & 2 & 3 & 3 & 1 \\ 0 & -4 & 1 & 1 & 1 \\ 9 & 0 & 1 & -5 & 0 \end{pmatrix}.$$

Il sistema può, dunque, scriversi nella forma

$$\begin{pmatrix} 1 \\ 0 \\ 9 \end{pmatrix} x_1 + \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} x_2 + \begin{pmatrix} 3 \\ 1 \\ 3 \end{pmatrix} x_3 + \begin{pmatrix} 3 \\ 1 \\ -5 \end{pmatrix} x_4 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

o nella forma

$$\begin{pmatrix} 1 & 1 & 3 & 3 \\ 0 & -2 & 1 & 1 \\ 9 & 0 & 3 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

$\square$

Una *soluzione del sistema* (4.5) è una soluzione comune a tutte le equazioni del sistema, cioè un'$n-$pla $(c_1, c_2, \ldots, c_n)$ di elementi di $\mathbb{F}$ tale che

$$\begin{cases} a_{11}c_1 + a_{12}c_2 + \cdots + a_{1n}c_n = b_1 \\ a_{21}c_1 + a_{22}c_2 + \cdots + a_{2n}c_n = b_2 \\ \qquad\vdots \\ a_{m1}c_1 + a_{m2}c_2 + \cdots + a_{mn}c_n = b_m \end{cases}. \tag{4.11}$$

Nel seguito, se $(c_1, c_2, \ldots, c_n)$ è una soluzione del sistema (4.5), porremo

$$\boldsymbol{c} = (c_1, c_2, \ldots, c_n)^t = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

e, quindi, avremo

$$A\boldsymbol{c} = \boldsymbol{b}\,. \tag{4.12}$$

Denoteremo, inoltre, con $Sol(S)$ l'insieme di tutte le soluzioni del sistema S; porremo, cioè

$$Sol(S) := \{\boldsymbol{c} \in \mathbb{F}^n \ : \ A\boldsymbol{c} = \boldsymbol{b}\}. \tag{4.13}$$

L'insieme delle soluzioni del sistema S è, dunque, un sottoinsieme dello spazio vettoriale numerico $\mathbb{F}^n$.

Osservazione 4.2.3. Se consideriamo il sistema (4.5), dalla (4.8) abbiamo che un vettore $\boldsymbol{c} = (c_1, c_2, \dots, c_n)^t$ è una sua soluzione se, e solo se,

$$\boldsymbol{a}^{(1)}c_1 + \boldsymbol{a}^{(2)}c_2 + \cdots + \boldsymbol{a}^{(2)}c_n = \boldsymbol{b}\,, \tag{4.14}$$

cioè se, e solo se, il vettore $\boldsymbol{b}$ è combinazione lineare delle colonne della matrice dei coefficienti del sistema mediante gli scalari $c_1, c_2, \dots, c_n$. □

Osservazione 4.2.4. Se, per ogni $i = 1, 2, \dots, m$, denotiamo con $\mathcal{S}_i$ l'insieme delle soluzioni della $i-$esima equazione del sistema (4.5), risulta

$$Sol(S) = \mathcal{S}_1 \cap \mathcal{S}_2 \cap \cdots \cap \mathcal{S}_m\,; \tag{4.15}$$

cioè, l'insieme $Sol(S)$ delle soluzioni del sistema è uguale all'intersezione di $\mathcal{S}_1$, $\mathcal{S}_2, \dots, \mathcal{S}_m$. □

Il sistema (4.5) si dice *compatibile* se ammette almeno una soluzione, *incompatibile* nel caso contrario. Quando tutte le equazioni del sistema (4.5) sono omogenee ($b_1 = b_2 = \cdots = b_m = 0$), cioè

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases} , \tag{4.16}$$

il sistema si dice *omogeneo*. In questo caso abbiamo almeno la soluzione

$$x_1 = x_2 = \cdots = x_n = 0,$$

che è detta *soluzione nulla* o *banale*. Un sistema omogeneo, dunque, è sempre compatibile. Un sistema del tipo (4.5) individua il sistema omogeneo (4.16), che si dice *associato* a (4.5).

Due sistemi lineari si dicono *equivalenti* se hanno lo stesso insieme di soluzioni. Anche qui, naturalmente, il sistema (4.5) dipende soltanto dalla sua matrice completa A' e non dalla notazione usata per le incognite; pertanto, quando l'ordine del sistema è relativamente piccolo, può convenire denotare le incognite stesse mediante lettere diverse dell'alfabeto, come nel caso di un'unica equazione lineare.

Proposizione 4.2.5. *Ogni sistema lineare è equivalente a quello che si ottiene eliminando le identità dalle equazioni che lo rappresentano. In altre parole, ogni sistema lineare è equivalente a quello che si ottiene eliminando le equazioni corrispondenti alle righe nulle della sua matrice completa.*

Dimostrazione. Un'equazione identica di un sistema lineare in n incognite ha per soluzioni tutti gli elementi di $\mathbb{F}^n$, così dalla (4.15) segue subito l'asserto. $\square$

Nei paragrafi che seguono ci occuperemo del problema di determinare degli algoritmi che permettano di stabilire se un fissato sistema lineare è compatibile e, in caso di risposta affermativa, di calcolarne le soluzioni.

4.3 Sistemi lineari ridotti

Sia S un sistema lineare di m equazioni in n incognite rappresentato dalle equazioni (4.5)

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

e supponiamo che la sua matrice completa A' sia ridotta. In queste ipotesi il sistema si dice *ridotto* e risulta ridotta anche la sua matrice incompleta A. Le incognite del sistema ridotto S corrispondenti ai pivot di A si dicono *principali*, e le rimanenti *libere*. Se la matrice A' è completamente ridotta, il sistema si dice *completamente ridotto*.

Esempio 4.3.1. Consideriamo il seguente sistema

$$\begin{cases} 2x_1 + x_2 - x_3 + 6x_4 - 2x_5 + x_6 & = & 3 \\ -2x_3 + x_4 - x_5 - x_6 & = & -1 \\ 2x_4 + 7x_5 + 3x_6 & = & 1 \end{cases}.$$

La sua matrice completa è

$$\begin{pmatrix} 2 & 1 & -1 & 6 & -2 & 1 & 3 \\ 0 & 0 & -2 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 2 & 7 & 3 & 1 \end{pmatrix}$$

ed è ridotta. Il sistema è, dunque, ridotto. Le incognite principali sono x_1, x_3, x_4, le libere sono x_2, x_5, x_6. $\square$

Esempio 4.3.2. Consideriamo il seguente sistema

$$\begin{cases} x_1 + 6x_4 - 2x_5 + x_6 & = & 3 \\ x_2 + x_4 - x_5 - x_6 & = & -1 \\ x_3 + 7x_5 + 3x_6 & = & 1 \end{cases}.$$

La sua matrice completa è

$$\begin{pmatrix} 1 & 0 & 0 & 6 & -2 & 1 & 3 \\ 0 & 1 & 0 & 1 & -1 & -1 & -1 \\ 0 & 0 & 1 & 0 & 7 & 3 & 1 \end{pmatrix}$$

ed è completamente ridotta. Il sistema è, dunque, completamente ridotto. Le incognite principali sono x_1, x_2, x_3, le libere sono x_4, x_5, x_6. □

Nel seguito del paragrafo, in forza della prop.4.2.5, supporremo che tra le equazioni del sistema S non vi siano identità o, equivalentemente, che le righe della matrice completa A' di S siano tutte non nulle. In queste ipotesi, le righe della matrice incompleta A diverse dall'ultima sono necessariamente non nulle.

Proposizione 4.3.3. *Se l'ultima riga della matrice incompleta A del sistema ridotto S è nulla, cioè*

$$(a_{m1}, a_{m2}, \ldots, a_{mn}) = (0, 0, \ldots, 0),$$

ed è $b_m \neq 0$, allora il sistema è incompatibile.

Dimostrazione. Nelle nostre ipotesi l'ultima equazione del sistema è

$$0x_1 + 0x_2 + \cdots + 0x_n = b_m,$$

con $b_m \neq 0$. Tale equazione non ha soluzioni e, quindi, il nostro sistema è incompatibile. □

Proposizione 4.3.4. *Supponiamo che nel sistema ridotto S sia $m = n$ (cioè il numero delle equazioni uguale al numero delle incognite) e che tutte le incognite siano principali (cioè la matrice incompleta A è triangolare con ciascun elemento della diagonale principale non nullo). Allora, la matrice completa A' è della forma*

$$A' = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} & a_{1n} & b_1 \\ & a_{22} & \cdots & a_{2,n-1} & a_{2n} & b_2 \\ & & \ddots & \vdots & \vdots & \\ & 0 & & a_{n-1,n-1} & a_{n-1,n} & b_{n-1} \\ & & & & & \\ & & & & a_{nn} & b_n \end{pmatrix},$$

e il sistema S ha un'unica soluzione $(\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_n)$ data da

$$\begin{cases} \overline{x}_n = \frac{b_n}{a_{nn}} \\ \overline{x}_{n-1} = \frac{b_{n-1}}{a_{n-1,n-1}} - \frac{a_{n-1,n}}{a_{n-1,n-1}}\overline{x}_n \\ \vdots \\ \overline{x}_2 = \frac{b_2}{a_{22}} - \frac{a_{23}}{a_{22}}\overline{x}_3 - \cdots - \frac{a_{2n}}{a_{22}}\overline{x}_n \\ \overline{x}_1 = \frac{b_1}{a_{11}} - \frac{a_{12}}{a_{11}}\overline{x}_2 - \frac{a_{13}}{a_{11}}\overline{x}_3 - \cdots - \frac{a_{1n}}{a_{11}}\overline{x}_n \end{cases}. \qquad (4.17)$$

Dimostrazione. Il sistema si presenta nella forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1,n-1}x_{n-1} + a_{1n}x_n &= b_1 \\ \qquad a_{22}x_2 + a_{23}x_3 + \cdots + a_{2,n-1}x_{n-1} + a_{2n}x_n &= b_2 \\ & \vdots \\ \qquad\qquad a_{n-1,n-1}x_{n-1} + a_{n-1,n}x_n &= b_{n-1} \\ \qquad\qquad\qquad a_{nn}x_n &= b_n \end{cases}$$

e allora, risolvendo a partire dall'ultima equazione, otteniamo la (4.17). □

Osservazione 4.3.5. L'unica soluzione di un sistema S completamente ridotto con $m = n$ si calcola immediatamente in quanto il sistema si presenta nella forma

$$\begin{cases} x_1 &= b_1 \\ x_2 &= b_2 \\ & \vdots \\ x_{n-1} &= b_{n-1} \\ x_n &= b_n \end{cases}.$$

□

Proposizione 4.3.6. *Se l'ultima riga della matrice incompleta A del sistema ridotto S è non nulla, allora il sistema è compatibile. In queste ipotesi vi sono m incognite principali, $n - m$ incognite libere e ciascuna soluzione del sistema corrisponde biunivocamente ad un assegnato valore delle incognite libere.*

Dimostrazione. Supponiamo, cosa che non lede le generalità del problema, che le incognite principali siano $x_1, x_2, \ldots, x_m$ e le libere $x_{m+1}, x_{m+2}, \ldots, x_n$. Allora ogni specializzazione delle incognite libere

$$x_{m+1} = c_{m+1},\ x_{m+2} = c_{m+2}, \ldots, x_n = c_n$$

dà luogo ad un sistema di m equazioni in m incognite principali che, in forza della prop.4.3.4, ammette un'unica soluzione

$$x_1 = c_1, x_2 = c_2, \ldots, x_m = c_m$$

e, in questo modo, otteniamo univocamente la soluzione $(c_1, c_2, \ldots, c_n)$ di S. Viceversa, ogni soluzione $(d_1, d_2, \ldots, d_n)$ di S corrisponde univocamente alla specializzazione

$$x_{m+1} = d_{m+1},\ x_{m+2} = d_{m+2}, \ldots, x_n = d_n$$

delle incognite libere. □

4.4 Risoluzione dei sistemi lineari

Le seguenti operazioni sulle equazioni di un sistema lineare S prendono il nome di *operazioni* elementari:

- lo scambio di posto di due equazioni (*lo scambio dell'equazione* $i-$*esima con la* $j-$*esima si denota con* $R_i \leftrightarrow R_j$);
- *la* moltiplicazione di un'equazione per uno scalare non nullo (*la moltiplicazione dell'equazione* $i-$*esima per lo scalare* λ *si denota con* $R_i \rightarrow \lambda R_i$);
- *l'*addizione ad un'equazione di un'altra equazione moltiplicata per uno scalare non nullo (*l'addizione all'equazione* $i-$*esima dell'equazione* $j-$*esima moltiplicata per lo scalare* λ *si denota con* $R_i \rightarrow R_i + \lambda R_j$).

Proposizione 4.4.1. *Sia* S *un sistema di equazioni lineari. Ogni sistema* S' *ottenuto applicando ad* S *una successione finita di trasformazioni elementari è equivalente ad* S.

Dimostrazione. Lo scambio di posto di due equazioni non altera l'insieme delle soluzioni del sistema e alla stessa conclusione si arriva se si moltiplica un'equazione per uno scalare non nullo (*cfr.* prop.4.1.4). Proviamo dunque che, se cambiamo l'equazione $i-$esima di S, addizionandole l'equazione $j-$esima moltiplicata per uno scalare non nullo λ, si ottiene un sistema S' equivalente ad S. Supponiamo che S sia il sistema (4.5) e, senza perdita di generalità, sommiamo alla seconda equazione la prima moltiplicata per λ, ottenendo, in questo modo, il sistema

$$S' = \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ (a_{21} + \lambda a_{11})x_1 + (a_{22} + \lambda a_{12})x_2 \\ \qquad\qquad + \cdots + (a_{2n} + \lambda a_{1n})x_n = b_2 + \lambda b_1 \\ a_{31}x_1 + a_{32}x_2 + \cdots + a_{3n}x_n = b_3 \\ \qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}, \tag{4.18}$$

Se $\boldsymbol{c} = (c_1, c_2, \dots, c_n)$ è una soluzione di S, risulta

$$a_{11}c_1 + a_{12}c_2 + \cdots + a_{1n}c_n = b_1, \tag{4.19}$$

$$a_{21}c_1 + a_{22}c_2 + \cdots + a_{2n}c_n = b_2, \tag{4.20}$$

da cui abbiamo

$$\lambda(a_{11}c_1 + \cdots + a_{1n}c_n) + (a_{21}c_1 + \cdots + a_{2n}c_n) = b_2 + \lambda b_1, \tag{4.21}$$

e di conseguenza

$$(a_{21} + \lambda a_{11})c_1 + \cdots + (a_{2n} + \lambda a_{1n})c_n = b_2 + \lambda b_1, \tag{4.22}$$

cioè $\boldsymbol{c}$ è soluzione di S'. Viceversa, se $\boldsymbol{c}$ è una soluzione di S', valgono le (4.22) e (4.19) e da queste si ricava subito la (4.20), cioè $\boldsymbol{c}$ è soluzione di S. □

Osservazione 4.4.2. Sia assegnato un sistema lineare S con matrice completa A'. Allora:

- lo scambio di due equazioni di S corrisponde allo scambio delle corrispondenti righe di A';
- la moltiplicazione dell'equazione $i-$esima di S per uno scalare non nullo λ corrisponde a moltiplicare per λ tutti gli elementi della riga $i-$esima di A';
- l'addizione all'equazione $i-$esima dell'equazione $j-$esima moltiplicata per uno scalare λ corrisponde nella matrice A' a sommare alla riga $i-$esima la riga $j-$esima moltiplicata per λ. □

Corollario 4.4.3. *Due sistemi lineari le cui matrici complete sono equivalenti sono a loro volta equivalenti.*

Esempio 4.4.4. Sia assegnato il seguente sistema S a coefficienti reali

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

con matrice completa

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

Di seguito riportiamo una successione di operazioni elementari su S e, a fianco, le corrispondenti sulla matrice completa:

$$R_1 \leftrightarrow R_2$$

$$\begin{cases} x_1 + 4x_2 + x_3 = 1 \\ \phantom{x_1 + {}} 2x_2 + 3x_3 = 1 \\ x_1 - x_2 + x_3 = 1 \end{cases} \qquad \begin{pmatrix} 1 & 4 & 1 & 1 \\ 0 & 2 & 3 & 1 \\ 1 & -1 & 1 & 1 \end{pmatrix}$$

$$R_3 \to R_3 - R_1$$

$$\begin{cases} x_1 + 4x_2 + x_3 = 1 \\ \phantom{x_1 + {}} 2x_2 + 3x_3 = 1 \\ - 5x_2 = 0 \end{cases} \qquad \begin{pmatrix} 1 & 4 & 1 & 1 \\ 0 & 2 & 3 & 1 \\ 1 & -1 & 1 & 0 \end{pmatrix}$$

$$R_2 \to \frac{5}{2} R_2$$

$$\begin{cases} x_1 + 4x_2 + x_3 = 1 \\ \phantom{x_1 + {}} 5x_2 + \frac{15}{2}x_3 = \frac{5}{2} \\ - 5x_2 = 0 \end{cases} \qquad \begin{pmatrix} 1 & 4 & 1 & 1 \\ 0 & 5 & \frac{15}{2} & \frac{5}{2} \\ 1 & -1 & 1 & 0 \end{pmatrix}$$

$$R_2 \to R_2 + R_3$$

$$\begin{cases} x_1 + 4x_2 + x_3 = 1 \\ \phantom{x_1 + 4x_2 + {}} \frac{15}{2}x_3 = \frac{5}{2} \\ - 5x_2 = 0 \end{cases} \qquad \begin{pmatrix} 1 & 4 & 1 & 1 \\ 0 & & \frac{15}{2} & \frac{5}{2} \\ 1 & -1 & 1 & 0 \end{pmatrix}$$

□

La prop.4.4.1, l'osservazione 4.4.2 e il corollario 4.4.3, visto che sappiamo trasformare una matrice in una equivalente ridotta (algoritmi di Gauss e di Gauss-Jordan), suggeriscono di studiare un sistema lineare riducendolo ad un sistema equivalente completamente ridotto del quale sappiamo calcolare le soluzioni (*cfr.* paragrafo 4.3). Proviamo ad applicare questa idea a qualche esempio.

Esempio 4.4.5. Consideriamo il sistema lineare S a coefficienti reali

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

e, utilizzando il risultato del corollario 4.4.3, proviamo a studiarlo riducendolo ad un sistema equivalente completamente ridotto. A tale scopo, trasformiamo la matrice completa di S in una equivalente e completamente ridotta, usando gli algoritmi 3.7.6 di Gauss e 3.8.4 di Gauss-Jordan. Possiamo operare nel modo seguente:

- PASSO 1: Scriviamo la matrice completa A' di S,
$$A' = \begin{pmatrix} 1 & 2 & 3 & 3 & 1 \\ 1 & 4 & 1 & 1 & 1 \end{pmatrix}.$$
- PASSO 2: Trasformiamo la matrice A' in una matrice equivalente ridotta, mediante l'algoritmo 3.7.6 di Gauss,
$$A'' = \begin{pmatrix} 1 & 2 & 3 & 3 & 1 \\ 0 & 2 & -2 & -2 & 0 \end{pmatrix}.$$
- PASSO 3: L'ultima riga non nulla di A'' non ha tutti i primi 4 elementi uguali a zero, quindi il sistema è compatibile (*cfr.* prop.4.3.6).
- PASSO 4: Trasformiamo la matrice A'' in una equivalente completamente ridotta, mediante l'algoritmo 3.8.4 di Gauss-Jordan,
$$\begin{pmatrix} 1 & 0 & 5 & 5 & 1 \\ 0 & 1 & -1 & -1 & 0 \end{pmatrix}.$$
- PASSO 5: Calcoliamo l'insieme delle soluzioni del sistema S' (equivalente ad S) la cui matrice completa è quella ottenuta al passo 4,
$$\begin{cases} x_1 + 5x_3 + 5x_4 = 1 \\ x_2 - x_3 - x_4 = 0 \end{cases}$$
assegnando dei valori arbitrari alle incognite libere x_3, x_4 e ricavando i corrispondenti valori delle incognite principali x_1, x_2 (*cfr.* prop.4.3.6). Otteniamo, così
$$\begin{cases} x_4 = \mu \\ x_3 = \lambda \\ x_2 = \lambda + \mu \\ x_1 = 1 - 5\lambda - 5\mu \end{cases}$$

e, quindi, il nostro sistema ha ∞^2 soluzioni, le quali sono tutte e sole le 4−ple di elementi di $\mathbb{R}^4$ del tipo

$$(1 - 5\lambda - 5\mu, \lambda + 5\mu, \lambda, \mu)$$

al variare di λ, μ e ν nel campo dei numeri reali. □

Esempio 4.4.6. Consideriamo il sistema lineare S a coefficienti reali

$$\begin{cases} 2x_1 + x_2 - x_3 + 6x_4 - 2x_5 + x_6 = 0 \\ x_1 - x_2 - 2x_3 + x_4 - x_5 - x_6 = 0 \\ 3x_1 + 2x_2 - x_3 + 2x_4 + 7x_5 + 3x_6 = 0 \end{cases}$$

e proviamo a studiare anche questo riducendolo ad un sistema equivalente completamente ridotto. A tale scopo, trasformiamo la matrice completa di S in una equivalente e completamente ridotta, usando gli algoritmi 3.7.6 di Gauss e 3.8.4 di Gauss-Jordan. Operiamo come nell'esempio precedente:

- PASSO 1: Scriviamo la matrice completa A' di S,

$$A' = \begin{pmatrix} 2 & 1 & -1 & 6 & -2 & 1 & 0 \\ 1 & -1 & -2 & 1 & -1 & -1 & 0 \\ 3 & 2 & -1 & 2 & 7 & 3 & 0 \end{pmatrix}.$$

- PASSO 2: Trasformiamo la matrice A' in una matrice equivalente ridotta, mediante l'algoritmo 3.7.6 di Gauss,

$$A'' = \begin{pmatrix} 2 & 1 & -1 & 6 & -2 & 1 & 0 \\ 0 & 3 & 3 & 4 & 0 & 3 & 0 \\ 0 & 0 & 0 & 23 & -30 & -3 & 0 \end{pmatrix}.$$

- PASSO 3: L'ultima riga non nulla di A'' non ha tutti i primi 6 elementi uguali a zero, quindi il sistema è compatibile.

- PASSO 4: Trasformiamo la matrice A'' in una equivalente completamente ridotta, mediante l'algoritmo 3.8.4 di Gauss-Jordan,

$$\begin{pmatrix} 1 & 0 & -1 & 0 & \frac{47}{23} & \frac{7}{23} & 0 \\ 0 & 1 & 1 & 0 & \frac{40}{23} & \frac{27}{23} & 0 \\ 0 & 0 & 0 & 1 & -\frac{30}{23} & -\frac{3}{23} & 0 \end{pmatrix}.$$

- PASSO 5: Calcoliamo l'insieme delle soluzioni del sistema S' (equivalente ad S) la cui matrice completa è quella ottenuta al passo 4,

$$\begin{cases} x_1 - x_3 + \frac{47}{23}x_5 + \frac{7}{23}x_6 = 0 \\ x_2 + x_3 + \frac{40}{23}x_5 + \frac{27}{23}x_6 = 0 \\ - x_4 - \frac{30}{23}x_5 - \frac{3}{23}x_6 = 0 \end{cases},$$

assegnando dei valori arbitrari alle incognite libere x_3, x_5, x_6 e ricavando i corrispondenti valori delle incognite principali x_1, x_2, x_4. Otteniamo, così

$$\begin{cases} x_6 = & \nu \\ x_5 = & \mu \\ x_3 = & \lambda \\ x_4 = & -\frac{30}{23}\mu - \frac{3}{23}\nu \\ x_2 = & -\lambda - \frac{40}{23}\mu - \frac{27}{23}\nu \\ x_1 = & \lambda - \frac{47}{23}\mu - \frac{7}{23}\nu \end{cases}$$

e, quindi, le soluzioni del nostro sistema sono tutte e sole le $6-$ple di elementi di $\mathbb{R}^6$ del tipo

$$\left(\lambda - \frac{47}{23}\mu - \frac{7}{23}\nu,\ -\lambda - \frac{40}{23}\mu - \frac{27}{23}\nu, \lambda,\ \frac{30}{23}\mu - \frac{3}{23}\nu,\ \mu,\ \nu\right)$$

al variare di λ, μ e ν nel campo dei numeri reali. Si osservi che, come accade per ogni sistema omogeneo, tra le soluzioni di S c'è quella nulla $(0, 0, 0, 0, 0, 0)$; questa ovviamente si ottiene in corrispondenza dei valori $\lambda = \mu = \nu = 0$ dei parametri. □

Esempio 4.4.7. Con lo stesso metodo dei due esempi precedenti, proviamo a studiare il seguente sistema lineare S a coefficienti reali

$$\begin{cases} 2x_1 + x_2 - x_3 + 6x_4 - 2x_5 + x_6 = 3 \\ x_1 - x_2 - 2x_3 + x_4 - x_5 - x_6 = -1 \\ 4x_1 - x_2 - 5x_3 + 8x_4 - 4x_5 - x_6 = 0 \end{cases}$$

Operariamo nel modo seguente:

- PASSO 1: Scriviamo la matrice completa A' di S,

$$A' = \begin{pmatrix} 2 & 1 & -1 & 6 & -2 & 1 & 3 \\ 1 & -1 & -2 & 1 & -1 & -1 & -1 \\ 4 & -1 & -5 & 8 & -4 & -1 & 0 \end{pmatrix}.$$

- PASSO 2: Trasformiamo la matrice A' in una matrice equivalente ridotta, mediante l'algoritmo 3.7.6 di Gauss,

$$A'' = \begin{pmatrix} 2 & 1 & -1 & 6 & -2 & 1 & 3 \\ 0 & 3 & 3 & 4 & 0 & 3 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}.$$

- PASSO 3: L'ultima riga non nulla di A'' ha i primi 6 elementi uguali a zero e l'ultimo diverso da zero, quindi il sistema è incompatibile (*cfr.* prop.4.3.3). □

I risultati e le osservazioni degli ultimi due paragrafi suggeriscono e giustificano il seguente algoritmo per lo studio di un sistema di equazioni lineari.

ALGORITMO 4.4.8. (STUDIO DI UN SISTEMA LINEARE)

- INPUT: Un sistema lineare S.
- PASSO 1: Scrivere la matrice completa A' di S.
- PASSO 2: Trasformare la matrice completa del sistema in una matrice equivalente ridotta, mediante l'algoritmo 3.7.6 di Gauss.
- PASSO 3: Se l'ultima riga non nulla della matrice ottenuta al passo 2 ha solo l'ultimo elemento diverso da zero, il sistema non è compatibile e l'algoritmo termina (*cfr.* prop.4.3.3).
- PASSO 4: Trasformare la matrice ottenuta al passo 2 in una equivalente completamente ridotta, mediante l'algoritmo 3.8.4 di Gauss-Jordan.
- PASSO 5: Calcolare l'insieme delle soluzioni del sistema S' (equivalente ad S) la cui matrice completa è quella ottenuta al passo 4, assegnando dei valori arbitrari alle incognite libere e ricavando i corrispondenti valori delle incognite principali (*cfr.* prop.4.3.6).
- OUTPUT: L'insieme $Sol(S') = Sol(S)$ delle soluzioni del sistema S. □

4.5 Alcune proprietà di $Sol(S)$

In questo paragrafo studieremo alcune proprietà algebrico-geometriche dell'insieme delle soluzioni di un sistema lineare S.

PROPOSIZIONE 4.5.1. *Sia assegnato un sistema lineare omogeneo. Allora, una combinazione lineare di due sue soluzioni è ancora una soluzione del sistema. In particolare, se* $\boldsymbol{c}$ *e* $\boldsymbol{d}$ *sono soluzioni e* λ *è uno scalare,* $\boldsymbol{c}+\boldsymbol{d}$ *e* $\alpha\boldsymbol{c}$ *sono ancora soluzioni del sistema.*

DIMOSTRAZIONE. Siano $A\boldsymbol{x} = 0$ un sistema omogeneo e $\boldsymbol{c}, \boldsymbol{d}$ due sue soluzioni, cioè $A\boldsymbol{c} = \boldsymbol{0}$, $A\boldsymbol{d} = \boldsymbol{0}$. Allora, se α, β sono scalari, risulta

$$A(\alpha\boldsymbol{c} + \beta\boldsymbol{d}) = A(\alpha\boldsymbol{c}) + A(\beta\boldsymbol{d}) = \alpha(A\boldsymbol{c}) + \beta(A\boldsymbol{d}) = \alpha\boldsymbol{0} + \beta\boldsymbol{0} = \boldsymbol{0}.$$

Per $(\alpha, \beta) = (1, 1)$ e $(\alpha, \beta) = (1, 0)$ si ottengono rispettivamente $\boldsymbol{c} + \boldsymbol{d}$ e $\alpha\boldsymbol{c}$ e l'asserto è provato. □

COROLLARIO 4.5.2. *L'insieme delle soluzioni del sistema lineare omogeneo* (4.16) *è un sottospazio vettoriale di* $\mathbb{F}^n$.

Supponiamo che il sistema S sia omogeneo e ridotto. Supponiamo, inoltre, che abbia m incognite principali e, di conseguenza, $n-m$, incognite libere. Se si assegna il valore 1 ad una delle incognite libere e zero alle rimanenti, si ottiene, in corrispondenza di tale assegnazione, un'unica soluzione di S, (*cfr.* prop.4.3.6) che diciamo *principale*. Ovviamente esistono $n - m$ soluzioni principali di S.

Proposizione 4.5.3. *Sia S un sistema lineare omogeneo e ridotto con m incognite principali. Allora ogni soluzione di S è combinazione lineare delle sue soluzioni principali.*

Dimostrazione. Supponiamo, cosa che non lede le generalità del problema, che le incognite principali siano $x_1, x_2, \ldots, x_m$ e le libere $x_{m+1}, x_{m+2}, \ldots, x_n$. Denotiamo con

$$\boldsymbol{c}_{m+1},\ \boldsymbol{c}_{m+2}, \ldots, \boldsymbol{c}_n$$

le soluzioni principali di S corrispondenti rispettivamente ai valori

$$(1, 0, \ldots, 0), (0, 1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)$$

dell'$(n-m)-$pla $(x_{m+1}, x_{m+2}, \ldots, x_n)$ delle incognite libere. Allora, in forza della prop. 4.5.1, ogni combinazione lineare di $\boldsymbol{c}_{m+1},\ \boldsymbol{c}_{m+2}, \ldots, \boldsymbol{c}_n$ è una soluzione di S. Ciò premesso, consideriamo una soluzione

$$\boldsymbol{b} = (b_1, \ldots, b_m, b_{m+1}, \ldots, b_n)$$

di S e osserviamo che questa è l'unica che corrisponde ai valori

$$(x_{m+1}, x_{m+2}, \ldots, x_n) = (b_{m+1}, b_{m+2}, \ldots, b_n) \tag{4.23}$$

delle incognite libere. D'altra parte, anche il vettore

$$b_{m+1}\boldsymbol{c}_{m+1} + b_{m+2}\boldsymbol{c}_{m+2} + \cdots + b_n\boldsymbol{c}_n$$

è una soluzione di S corrispondente ai valori (4.23) delle incognite libere e, di conseguenza, deve essere

$$\boldsymbol{b} = b_{m+1}\boldsymbol{c}_{m+1} + b_{m+2}\boldsymbol{c}_{m+2} + \cdots + b_n\boldsymbol{c}_n\ ,$$

cioè l'asserto. □

Proposizione 4.5.4. *La differenza di due soluzioni del sistema lineare* (4.5) *è una soluzione del sistema omogeneo associato* (4.16)*. La somma di una soluzione di* (4.5) *e di una di* (4.16) *è una soluzione di* (4.16)*.*

Dimostrazione. Siano $\boldsymbol{c}, \boldsymbol{d}$ due sue soluzioni del sistema (4.5) e $\boldsymbol{u}$ una soluzione del sistema omogeneo associato (4.16), cioè $A\boldsymbol{c} = \boldsymbol{b}$, $A\boldsymbol{d} = \boldsymbol{b}$, $A\boldsymbol{u} = \boldsymbol{0}$. Allora, risulta

$$A(\boldsymbol{d} - \boldsymbol{c}) = A\boldsymbol{d} - A\boldsymbol{c} = \boldsymbol{b} - \boldsymbol{b} = \boldsymbol{0}$$

e

$$A(\boldsymbol{c} + \boldsymbol{u}) = A\boldsymbol{c} + A\boldsymbol{u} = \boldsymbol{b} + \boldsymbol{0} = \boldsymbol{b},$$

cioè l'asserto. □

Proposizione 4.5.5. (Soluzioni di un sistema lineare) *Sia $\boldsymbol{c}$ una soluzione del sistema lineare S di equazioni* (4.5) *e sia U l'insieme delle soluzioni del sistema omogeneo associato* (4.16)*. Allora l'insieme*

$$\boldsymbol{c} + U = \{\boldsymbol{c} + \boldsymbol{u}\ :\ \boldsymbol{u} \in U\} \tag{4.24}$$

coincide con l'insieme $Sol(S)$ delle soluzioni del sistema (4.5)*.*

DIMOSTRAZIONE. La seconda parte della proposizione precedente assicura che è

$$\boldsymbol{c} + U \subseteq Sol(S).$$

Sia dunque $\boldsymbol{d}$ una soluzione del sistema S e osserviamo che, in forza della prima parte della precedente proposizione, il vettore $\boldsymbol{u} = \boldsymbol{d} - \boldsymbol{c}$ appartiene ad U e, quindi, risulta $\boldsymbol{d} = \boldsymbol{c} + \boldsymbol{u} \in \boldsymbol{c} + U$, cioè

$$Sol(S) \subseteq \boldsymbol{c} + U.$$

Dalle due inclusioni provate segue l'asserto. □

OSSERVAZIONE 4.5.6. La precedente proposizione assicura che, per determinare l'insieme $Sol(S)$ di un sistema lineare S, basta conoscere una soluzione del sistema stesso e l'insieme di tutte le soluzioni del sistema omogeneo associato. □

COROLLARIO 4.5.7. *Sia S il sistema lineare di equazioni* (4.5) *e sia U il sottospazio delle soluzioni* (cfr. corollario 4.5.2) *del sistema omogeneo associato* (4.16). *Allora $Sol(S)$ coincide con il laterale di U relativo ad un qualinque elemento di $Sol(S)$; cioè, se $\boldsymbol{c}, \boldsymbol{d}$ sono soluzioni del sistema S, risulta:*

$$Sol(S) = \boldsymbol{c} + U = \boldsymbol{b} + U. \tag{4.25}$$

Capitolo 5

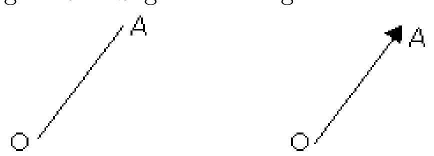
Spazi Vettoriali

Nei capitoli precedenti abbiamo studiato due classi di spazi vettoriali: gli spazi numerici $\mathbb{F}^n$ e gli spazi di matrici $\mathbb{F}^{m,n}$ su un campo $\mathbb{F}$. Vogliamo ora introdurre, nel modo piú generale possibile, il concetto di *spazio vettoriale su un campo* concentrando la nostra attenzione soltanto sulle proprietà delle operazioni tra vettori prescindendo da quelle relative alla natura dei vettori stessi. Prima di iniziare questo studio, però, riteniamo opportuno illustrare un altro esempio importante di spazio vettoriale che, almeno apparentemente, non coinvolge $n-$ple di elementi e matrici.

5.1 Segmenti orientati

Sia $\mathbb{E}_3$ l'insieme dei punti dello spazio della geometria euclidea. Se O e A sono due punti di $\mathbb{E}_3$, denotiamo con OA, o con AO, il segmento di estremi O e A[1]. Denotiamo, invece, con $\overrightarrow{OA}$, o con $A - O$, il segmento di estremi O (*primo estremo*) e A (*secondo estremo*) orientato da O verso A[2]. Un tale *segmento*

Figura 5.1: Segmento e segmento orientato



orientato si dice anche *uscente* da O e, nell'ipotesi $O \neq A$, si definiscono i seguenti parametri:

[1]Quando si assegna un segmento OA non ha alcuna importanza l'ordine con cui si elencano i punti O e A; un tale segmento è, dunque, individuato dall'insieme $\{O, A\}$.

[2]Nella considerazione del segmento orientato $\overrightarrow{OA}$ è essenziale precisare l'ordine con cui si assegnano i punti O e A : un tale segmento è, dunque, individuato dalla coppia (O, A).

- la *lunghezza* (rispetto ad una prefissata unità di misura), o *modulo*, cioè la misura del segmento OA, che denotiamo con $\mid OA \mid$;

- la *direzione*, che è quella individuata dalla retta per O e A;

- il *verso di percorrenza*, che è quello da O verso A.

Osserviamo esplicitamente che non escludiamo l'ipotesi che i punti O ed A coincidano. In questo caso otteniamo il segmento orientato $\overrightarrow{OO}$, che è l'unico di lunghezza zero uscente da O e che chiamiamo *segmento orientato nullo*.

Due segmenti orientati $\overrightarrow{OA}$ e $\overrightarrow{O'A'}$ si dicono *paralleli* se hanno la stessa direzione, cioè se la retta per O e A è parallela a quella per O' e A'. In questo caso, $\overrightarrow{OA}$ e $\overrightarrow{O'A'}$ si dicono *concordi* se inducono lo stesso verso su una retta ad essi parallela, *discordi* in caso contrario. Ovviamente, due segmenti orientati con la stessa origine $\overrightarrow{OA}$ e $\overrightarrow{OB}$ sono paralleli se i punti O, A, B sono allineati, cioè appartengono ad una stessa retta. Analogamente, un segmento orientato non nullo $\overrightarrow{OA}$ e una retta ℓ si dicono *paralleli*, se hanno la stessa direzione. Assumiamo, per convenzione, che il segmento orientato nullo sia parallelo ad ogni altro segmento orientato e ad ogni retta.

Se B è il punto di $\mathbb{E}_3$ simmetrico di A rispetto ad O, il segmento orientato $\overrightarrow{OB}$ si chiama *opposto* di $\overrightarrow{OA}$ e si denota con $-\overrightarrow{OA}$. Chiaramente $\overrightarrow{OA}$ e $\overrightarrow{OB}$ hanno lo stesso modulo, sono paralleli e discordi e si ha:

$$\overrightarrow{OB} = -\overrightarrow{OA} \iff \overrightarrow{OA} = -\overrightarrow{OB}\,.$$

Nell'insieme di tutti i segmenti orientati uscenti da O, che denotiamo con $\mathcal{V}_O$, si definiscono un'operazione interna di *addizione* ed una esterna di *moltiplicazione*, nel seguente modo:

- addizione di due segmenti orientati uscenti da O (Fig.5.2): per ogni $\overrightarrow{OA}, \overrightarrow{OB} \in \mathcal{V}_O$,

$$\overrightarrow{OA} + \overrightarrow{OB} = \overrightarrow{OT} \tag{5.1}$$

ove

- se O, A, B non sono allineati, T è il quarto vertice del parallelogramma individuato da $\overrightarrow{OA}$ e $\overrightarrow{OB}$ (regola del *parallelogramma*) ;

- se $\overrightarrow{OA}$ e $\overrightarrow{OB}$ sono paralleli e concordi, $\overrightarrow{OT}$ è l'unico segmento orientato parallelo e concorde ad $\overrightarrow{OA}$ e $\overrightarrow{OB}$ di modulo pari a $\mid OA \mid + \mid OB \mid$;

- se $\overrightarrow{OA}$ e $\overrightarrow{OB}$ sono paralleli e discordi, $\overrightarrow{OT}$ è l'unico segmento orientato parallelo ad $\overrightarrow{OA}$ e $\overrightarrow{OB}$, di modulo pari al valore assoluto[3] di $\mid OA \mid - \mid OB \mid$ e concorde con $\overrightarrow{OA}$ o $\overrightarrow{OB}$ a seconda che è $\mid OA \mid \geq \mid OB \mid$ o $\mid OA \mid \leq \mid OB \mid$.

[3] Il valore assoluto di un numero reale α, che si denota con $\mid \alpha \mid$, è, per definizione, uguale ad α se è $\alpha \geq 0$ e uguale a $-\alpha$ se è $\alpha < 0$. In ogni caso è $\mid \alpha \mid \geq 0$ e il segno di uguaglianza vale se, e solo se, $\alpha = 0$.

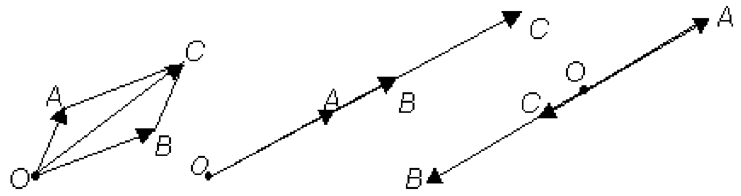
- moltiplicazione di un numero reale per un segmento orientato: per ogni $\overrightarrow{OA} \in \mathcal{V}_O$ e $\alpha \in \mathbb{R}$,

$$\alpha\, \overrightarrow{OA} = \overrightarrow{OT} \tag{5.2}$$

ove

- se $\alpha \neq 0$, $\overrightarrow{OT}$ è l'unico segmento orientato parallelo ad $\overrightarrow{OA}$, di modulo pari a $\mid \alpha \mid\mid OA \mid$ e concorde o discorde con $\overrightarrow{OA}$ a seconda che α sia positivo o negativo;
- se $\alpha = 0$, $\overrightarrow{OT}$ è il segmento orientato nullo .

Figura 5.2: Somma di segmenti orientati



Le operazioni ora definite verificano le seguenti proprietà:

$SV1$. *proprietà associativa dell'addizione*:

per ogni $\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC} \in \mathcal{V}_O$,

$$(\overrightarrow{OA} + \overrightarrow{OB}) + \overrightarrow{OC} = \overrightarrow{OA} + (\overrightarrow{OB} + \overrightarrow{OC});$$

$SV2$. *esistenza dell'elemento neutro per l'addizione*:

per ogni $\overrightarrow{OA} \in \mathcal{V}_O$,

$$\overrightarrow{OA} + \overrightarrow{OO} = \overrightarrow{OO} + \overrightarrow{OA} = \overrightarrow{OA};$$

$SV3$. *simmetrizzabilità rispetto all'addizione*:

per ogni $\overrightarrow{OA} \in \mathcal{V}_O$,

$$\overrightarrow{OA} + (-\overrightarrow{OA}) = -\overrightarrow{OA} + \overrightarrow{OA} = \overrightarrow{OO};$$

$SV4$. *proprietà commutativa dell'addizione*:

per ogni $\overrightarrow{OA}, \overrightarrow{OB} \in \mathcal{V}_O$,

$$\overrightarrow{OA} + \overrightarrow{OB} = \overrightarrow{OB} + \overrightarrow{OA};$$

$SV5$. *proprietà distributiva della moltiplicazione di un numero reale per un segmento orientato rispetto all'ddizione fra segmenti orientati*:

per ogni $\alpha \in \mathbb{R}$ e per ogni $\overrightarrow{OA}, \overrightarrow{OB} \in \mathcal{V}_O$,

$$\alpha(\overrightarrow{OA} + \overrightarrow{OB}) = \alpha\overrightarrow{OA} + \alpha\overrightarrow{OB};$$

$SV6$. *proprietà distributiva della moltiplicazione di un numero reale per un segmento orientato rispetto all'addizione fra numeri reali*:

per ogni $\alpha\beta \in \mathbb{R}$ e per ogni $\overrightarrow{OA} \in \mathcal{V}_O$,

$$(\alpha + \beta)\overrightarrow{OA} = \alpha\overrightarrow{OA} + \beta\overrightarrow{OA};$$

$SV7$. *proprietà associativa della moltiplicazione di un numero reale per un segmento orientato con la moltiplicazione fra numeri reali*:

per ogni $\alpha, \beta \in \mathbb{R}$ e per ogni $\overrightarrow{OA} \in \mathcal{V}_O$,

$$(\alpha\beta)\overrightarrow{OA} = \alpha(\beta\overrightarrow{OA}).$$

$SV8$. per ogni $\overrightarrow{OA} \in \mathcal{V}_O$,

$$1\,\overrightarrow{OA} = \overrightarrow{OA}.$$

Come di solito, il fatto che in $\mathcal{V}_O$ le operazioni di addizione e moltiplicazione verificano le precedenti otto proprietà si esprime dicendo che $\mathcal{V}_O$ è uno *spazio vettoriale sul campo* $\mathbb{R}$ *dei numeri reali*.

Assegnati un piano Π ed un suo punto O, possiamo considerare l'insieme $\mathcal{V}_O^\Pi$ dei segmenti orientati di origine O e contenuti in Π. È immediato verificare le seguenti proprietà:

- (*proprietà di chiusura dell'addizione*) per ogni $\overrightarrow{OA}, \overrightarrow{OB} \in \mathcal{V}_O^\Pi$, il segmento orientato $\overrightarrow{OA} + \overrightarrow{OB}$ appartiene ancora a $\mathcal{V}_O^\Pi$;

- (*proprietà di chiusura della moltiplicazione di un numero reale per un segmento orientato*) per ogni $\alpha \in \mathbb{R}$ e per ogni $\overrightarrow{OA} \in \mathcal{V}_O^\Pi$, il segmento orientato $\alpha\overrightarrow{OA}$ appartiene ancora a $\mathcal{V}_O^\Pi$.

Ne segue che in $\mathcal{V}_O^\Pi$ si possono considerare le operazioni di addizione e moltiplicazione indotte da quelle definite in V_O. Tali operazioni verificano a loro volta le proprietà 5.2, $SV1$,...,$SV8$ e, così , $\mathcal{V}_O^\Pi$ è uno spazio vettoriale su $\mathbb{R}$, che si dice anche *sottospazio vettoriale* di $\mathcal{V}_O$ (in notazione: $\mathcal{V}_O^\Pi \leq \mathcal{V}_O$).

Analogamente a quanto fatto per un piano, assegnata una retta ℓ ed un suo punto O, si vede che l'insieme $\mathcal{V}_O^\ell$ dei segmenti orientati di origine O e contenuti in ℓ costituiscono un sottospazio vettoriale di $\mathcal{V}_O$. Chiaramente, se la retta ℓ è contenuta nel piano Π, risulta $\mathcal{V}_O^\ell \leq \mathcal{V}_O^\Pi \leq \mathcal{V}_O$.

5.2 Definizione di spazio vettoriale

Sia $\mathbb{V}$ un insieme non vuoto ed $\mathbb{F}$ un campo[4]. Diremo *vettori* gli elementi di $\mathbb{V}$, *scalari* quelli di $\mathbb{F}$ e, quasi sempre, denoteremo con lettere latine in grassetto i vettori e con lettere greche gli scalari. Denoteremo, inoltre, con 0 e 1 rispettivamente lo *zero* e l'*unità* del campo $\mathbb{F}$.

Siano "+" un'operazione binaria definita in $\mathbb{V}$ (*addizione fra vettori*) e "$\cdot$ " un'operazione esterna tra $\mathbb{F}$ e $\mathbb{V}$ a valori in $\mathbb{V}$ (*moltiplicazione di uno scalare per un vettore*). Se $\alpha \in \mathbb{F}$ è uno scalare e $\boldsymbol{b} \in \mathbb{V}$ un vettore, il vettore $\alpha \cdot \boldsymbol{b}$ è denotato semplicemente con $\alpha\boldsymbol{b}$, si pone cioè

$$\alpha \cdot \boldsymbol{b} := \alpha\boldsymbol{b}\,.$$

Così, per esempio, la scrittura $(\alpha\beta)\boldsymbol{a}$ indica il vettore che si ottiene moltiplicando lo scalare $\alpha\beta$ per il vettore $\boldsymbol{a}$, ove $\alpha\beta$ indica il prodotto in $\mathbb{F}$ degli scalari α e β. Notiamo esplicitamente che, con abuso di notazione, useremo lo stesso simbolo $+$ per la somma di vettori e per quella di scalari.

La quaterna $(\mathbb{V}, \mathbb{F}, +, \cdot)$ prende il nome di *spazio vettoriale sul campo* $\mathbb{F}$ se sono verificati i seguenti otto assiomi:

$SV1$. *proprietà associativa dell'addizione*:

per ogni $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{V}$,

$$(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c} = \boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c});$$

$SV2$. *esistenza dell'elemento neutro*[5] *per l'addizione*:

esiste un elemento $\mathbf{0} \in \mathbb{V}$, detto *vettore nullo*, tale che per ogni $\boldsymbol{a} \in \mathbb{V}$,

$$\boldsymbol{a}+\mathbf{0} = \mathbf{0}+\boldsymbol{a} = \boldsymbol{a};$$

$SV3$. *simmetrizzabilità rispetto all'addizione*[6]:

per ogni $\boldsymbol{a} \in \mathbb{V}$, esiste un vettore $-\boldsymbol{a} \in \mathbb{V}$, detto l'*opposto* di $\boldsymbol{a}$, tale che

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

$SV4$. *proprietà commutativa dell'addizione*:

per ogni $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{V}$,

$$\boldsymbol{a}+\boldsymbol{b} = \boldsymbol{b}+\boldsymbol{a}.$$

$SV5$. *proprietà distributiva della moltiplicazione per uno scalare rispetto all'addizione fra vettori*:

per ogni $\alpha \in \mathbb{F}$ e per ogni $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{V}$,

$$\alpha(\boldsymbol{a}+\boldsymbol{b}) = \alpha\boldsymbol{a}+\alpha\boldsymbol{b};$$

[4]Anche qui, il Lettore che non abbia ancora assimilato bene il concetto di campo può continuare la lettura delle note pensando che $\mathbb{F}$ sia il campo reale $\mathbb{R}$, o quello razionale $\mathbb{Q}$.

[5]Un elemento $\boldsymbol{e} \in \mathbb{V}$ si dice neutro rispetto all'addizione se: $\boldsymbol{a}+\boldsymbol{e} = \boldsymbol{e}+\boldsymbol{a} = \boldsymbol{a}$.

[6]Un elemento $\boldsymbol{b} \in \mathbb{V}$ si dice *simmetrico* di $\boldsymbol{a} \in \mathbb{V}$ rispetto all'addizione se: $\boldsymbol{a}+\boldsymbol{b} = \boldsymbol{b}+\boldsymbol{a} = \mathbf{0}$.

$SV6$. *proprietà distributiva della moltiplicazione per uno scalare rispetto all'addizione fra scalari*:

per ogni $\alpha\beta \in \mathbb{F}$ e per ogni $\boldsymbol{a} \in \mathbb{V}$,

$$(\alpha + \beta)\boldsymbol{a} = \alpha\boldsymbol{a} + \beta\boldsymbol{a};$$

$SV7$. *proprietà associativa della moltiplicazione per uno scalare con la moltiplicazione fra scalari*:

per ogni $\alpha\beta \in \mathbb{F}$ e per ogni $\boldsymbol{a} \in \mathbb{V}$,

$$(\alpha\beta)\boldsymbol{a} = \alpha(\beta\boldsymbol{a});$$

$SV8$. per ogni $\boldsymbol{a} \in \mathbb{V}$,

$$1\,\boldsymbol{a} = \boldsymbol{a}.$$

L'insieme dei vettori $\mathbb{V}$ si dice *sostegno* dello spazio vettoriale ed è un gruppo abeliano rispetto all'operazione di addizione (proprietà $SV1$, $SV2$, $SV3$, $SV4$). Di solito, per semplificare le notazioni, lo spazio vettoriale $(\mathbb{V}, \mathbb{F}, +, \cdot)$ si denota soltanto col simbolo del suo sostegno $\mathbb{V}$.

Esempio 5.2.1. (Spazi vettoriali numerici) L'insieme $\mathbb{F}^n$ delle $n-$ple ordinate di elementi del campo $\mathbb{F}$, come già osservato nel paragrafo 2.1, è uno spazio vettoriale su $\mathbb{F}$ rispetto alle operazioni di addizione fra $n-$ple e di moltiplicazione di uno scalare per una $n-$pla (*spazio vettoriale numerico su* $\mathbb{F}$). In particolare, per $n = 1$, si ha che ogni campo può essere riguardato come spazio vettoriale su se stesso. □

Esempio 5.2.2. (Spazi di matrici) L'insieme $\mathbb{F}^{m,n}$ delle matrici di tipo $m \times n$ su $\mathbb{F}$, come già osservato nel paragrafo 3.1, è uno spazio vettoriale su $\mathbb{F}$ rispetto alle operazioni di addizione fra matrici e di moltiplicazione di uno scalare per una matrice (*spazio vettoriale delle matrici di tipo* $m \times n$ *su* $\mathbb{F}$). □

Esempio 5.2.3. (Spazi di segmenti orientati) L'insieme $\mathcal{V}_O$ dei segmenti orientati di origine un punto O in $\mathbb{E}_3$ (o in $\mathbb{E}_1$, o in $\mathbb{E}_2$), come già osservato nel paragrafo 5.1, forma uno spazio vettoriale sul campo $\mathbb{R}$ dei numeri reali rispetto alle operazioni di addizione fra segmenti orientati e di moltiplicazione di uno scalare per un segmento orientato. □

Esempio 5.2.4. (Campi e sottocampi) Sia $\mathbb{K}$ un campo ed $\mathbb{F}$ un suo sottocampo[7]. Allora $\mathbb{K}$ risulta uno spazio vettoriale su $\mathbb{F}$ rispetto all'addizione fra elementi di $\mathbb{K}$ e alla moltiplicazione in $\mathbb{K}$ fra un elemento di $\mathbb{F}$ e uno di $\mathbb{K}$. In particolare, il campo reale $\mathbb{R}$ è uno spazio vettoriale sul campo razionale $\mathbb{Q}$ e il campo complesso $\mathbb{C}$ è uno spazio vettoriale sia su $\mathbb{R}$ che su $\mathbb{Q}$. □

[7] Un sottoinsieme non vuoto $\mathbb{F}$ di $\mathbb{K}$ è un *sottocampo* di $\mathbb{K}$ se è chiuso rispetto alle operazioni di addizione e moltiplicazione di $\mathbb{K}$ e se queste inducono su $\mathbb{F}$ una struttura di campo. Per esempio, $\mathbb{Q}$ è sottocampo di $\mathbb{R}$ e di $\mathbb{C}$; $\mathbb{R}$ è sottocampo di $\mathbb{C}$.

Esempio 5.2.5. (Spazio dei polinomi) Consideriamo l'insieme $\mathbb{F}[x]$ dei polinomi nell'indeterminata x a coefficienti in un campo $\mathbb{F}$, cioè

$$\mathbb{F}[x] = \{a_o + a_1x + a_2x^2 + \cdots + a_nx^n \; : \; a_o, a_1, \ldots, a_n \in \mathbb{F}\,,\; n \in \mathbb{N}\,\}.$$

Per $n = 0$ si ottengono gli elementi di $\mathbb{F}$, che si dicono *polinomi costanti* o, piú semplicemente, *costanti.* Per il polinomio $a_o + a_1x + a_2x^2 + \cdots + a_nx^n$ usiamo anche la notazione compatta $a(x)$; in altre parole poniamo

$$a(x) := a_o + a_1x + a_2x^2 + \cdots + a_nx^n\,, \tag{5.3}$$

utilizzando per la notazione compatta di un polinomio la stessa lettera adoperata per i suoi coefficienti, e viceversa. A volte per il polinomio $a(x)$ scriveremo anche

$$a(x) = \sum_{j=1}^{n} a_jx^j$$

o, sottointendendo l'insieme di variazione degli indici,

$$a(x) = \sum a_jx^j.$$

Gli scalari $a_o, a_1, \ldots, a_n$ che compaiono nell'espressione (5.3) si chiamano *coefficienti* di $a(x)$; per un tale polinomio si assume per convenzione che sia $a_j = 0$, per ogni intero $j > n$.

In $\mathbb{F}[\boldsymbol{x}]$ consideriamo le seguenti operazioni:

$a(x) + b(x) = \sum_j (a_j + b_j)x^j$ (*addizione fra polinomi*),

$\alpha a(x) = \sum_j (\alpha a_j)x^j$ (*moltiplicazione di uno scalare per un polinomio*),

per ogni $a(x), b(x) \in F[x]$ e $\alpha \in \mathbb{F}$. È un esercizio provare che, rispetto a tali operazioni, $\mathbb{F}[x]$ ha una struttura di spazio vettoriale su $\mathbb{F}$. Il vettore nullo di questo spazio è il *polinomio nullo* 0, cioè il polinomio con tutti i coeficienti uguali a zero. L'opposto $-a(x)$ del polinomio $a(x)$ è l'unico polinomio avente per coefficienti gli opposti dei coefficienti di $a(x)$, cioè

$$\begin{aligned} -a(x) &= (-a_o) + (-a_1)x + (-a_2)x^2 + \cdots + (-a_n)x^n \\ &= -a_o - a_1x - a_2x^2 - \cdots - a_nx^n. \end{aligned}$$

Ricordiamo che, il *grado* di un polinomio non nullo $a(x) \in \mathbb{F}[x]$ è il piú grande intero n per cui $a_n \neq 0$ e si denota con $\deg a$; al polinomio nullo non si attribuisce grado. I polinomi di grado zero, pertanto, sono tutti e soli gli elementi non nulli del campo $\mathbb{F}$. È facile dimostrare le seguenti relazioni

$$\begin{cases} \deg(a+b) \leq max\{\deg a, \deg b\}\,, \;\; a(x), b(x) \in \mathbb{F}[x], \;\; a(x) \neq -b(x) \\ \deg(\alpha a) = \deg a\,, \;\; a(x) \in \mathbb{F}[x]\,, \;\; \alpha \in \mathbb{F}\,, \;\; a(x) \neq 0\,, \;\; \alpha \neq 0 \end{cases} \tag{5.4}$$

ove $max\{\deg a, \deg b\}$ denota il più grande fra gli interi $\deg a$ e $\deg b$. □

Esempio 5.2.6. (Spazio dei polinomi lineari) Consideriamo l'insieme

$$\mathbb{F}_1[\boldsymbol{x}] = \mathbb{F}_1[x_1, x_2, \dots, x_n]$$

dei polinomi lineari nelle indeterminate $x_1, x_2, \dots, x_n$ a coefficienti in un campo $\mathbb{F}_1$, cioè

$$\mathbb{F}_1[\boldsymbol{x}] = \{a_o + a_1x_1 + a_2x_2 + \cdots + a_nx_n \ : \ a_o, a_1, \dots, a_n \in \mathbb{F}\}.$$

Per $a_1 = a_2 = \cdots = a_n = 0$ si ottengono gli elementi di $\mathbb{F}$, che si dicono *polinomi costanti* o, piú semplicemente, *costanti*. Per il polinomio lineare $a_o + a_1x_1 + a_2x_2 + \cdots + a_nx_n$ usiamo anche le notazioni compatte $a(x_1, x_2, \dots, x_n)$ o $a(x_j)$; in altre parole poniamo

$$a(x_j) := a_o + a_1x_1 + a_2x_2 + \cdots + a_nx_n \,, \tag{5.5}$$

utilizzando per la notazione compatta di un polinomio la stessa lettera adoperata per i suoi coefficienti, e viceversa. A volte per il polinomio $a(x_j)$ scriveremo anche

$$a(x_j) = a_0 + \sum_{j=1}^{n} a_jx_j$$

o, sottointendendo l'insieme di variazione degli indici,

$$a(x_j) = a_0 + \sum a_jx_j.$$

Gli scalari $a_o, a_1, \dots, a_n$ che compaiono nell'espressione (5.5) si chiamano *coefficienti* di $a(x_j)$. In $\mathbb{F}_1[\boldsymbol{x}]$ consideriamo le seguenti operazioni:

$a(x_j) + b(x_j) = \sum_j (a_j + b_j)x_j$ (*addizione fra polinomi*),

$\alpha a(x_j) = \alpha a_0 + \sum_j (\alpha a_j)x_j$ (*moltiplicazione di uno scalare per un polinomio*),

per ogni $a(x_j), b(x_j) \in \mathbb{F}_1[\boldsymbol{x}]$ e $\alpha \in \mathbb{F}$. È un esercizio provare che, rispetto a tali operazioni, $\mathbb{F}_1[\boldsymbol{x}]$ ha una struttura di spazio vettoriale su $\mathbb{F}$. Il vettore nullo di questo spazio è il *polinomio nullo* 0, cioè il polinomio con tutti i coeficienti uguali a zero. L'opposto $-a(x_j)$ del polinomio $a(x_j)$ è l'unico polinomio avente per coefficienti gli opposti dei coefficienti di $a(x_j)$, cioè

$$\begin{aligned} -a(x_j) &= (-a_o) + (-a_1)x_1 + (-a_2)x_2 + \cdots + (-a_n)x_n \\ &= -a_o - a_1x_1 - a_2x_2 - \cdots - a_nx_n. \end{aligned}$$

□

Esempio 5.2.7. (Spazi di funzioni) Sia A un insieme non vuoto e denotiamo con $\mathbb{F}^A$ l'insieme di tutte le funzioni di A in $\mathbb{F}$. Se f, g sono funzioni appartenenti a $\mathbb{F}^A$, definiamo la *somma* $f + g$ di f e g nel seguente modo:

$$(f + g)(a) := f(a) + g(a)\,, \ \text{ per ogni } \ 4a \in A. \tag{5.6}$$

Definiamo, inoltre, il *prodotto* αf di uno scalare $\alpha \in \mathbb{F}$ per una funzione $f \in F^A$ nel seguente modo:

$$(\alpha f)(a) := \alpha f(a)\,, \quad \text{per ogni } \; a \in A. \tag{5.7}$$

È facile provare che, rispetto alle operazioni appena definite, F^A risulta uno spazio vettoriale su $\mathbb{F}$. Il vettore nullo di F^A è la *funzione nulla*, cioè la funzione che ad ogni elemento di A associa lo zero di $\mathbb{F}$. L'opposto di un elemento $f \in \mathbb{F}^A$ è la funzione $-f$ definita da

$$-f \;:\; a \in A \;\to\; -f(a) \in \mathbb{F}\,. \qquad \square$$

Nel seguito riterremo sempre assegnato uno spazio vettoriale $\mathbb{V}$ su un campo $\mathbb{F}$. Porremo, inoltre,

$$\boldsymbol{b} - \boldsymbol{a} := \boldsymbol{b} + (-\boldsymbol{a})$$

per ogni due vettori $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{V}$.

Proposizione 5.2.8. *Per ogni due vettori $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{V}$, l'equazione $\boldsymbol{x} + \boldsymbol{a} = \boldsymbol{b}$, nell'incognita $\boldsymbol{x}$, ammette un'unica soluzione data da $\boldsymbol{c} = \boldsymbol{b} - \boldsymbol{a}$.*

Dimostrazione. Se nell'equazione si pone $\boldsymbol{x} = \boldsymbol{b} - \boldsymbol{a}$, abbiamo

$$(\boldsymbol{b} - \boldsymbol{a}) + \boldsymbol{a} = \boldsymbol{b} + (-\boldsymbol{a} + \boldsymbol{a}) = \boldsymbol{b} + \boldsymbol{0} = \boldsymbol{b}\,.$$

D'altra parte, per una soluzione $\boldsymbol{c}$ della nostra equazione, si ha

$$\boldsymbol{c} + \boldsymbol{a} = \boldsymbol{b}$$

da cui, sommando $-\boldsymbol{a}$ ai due membri dell'uguaglianza,

$$\begin{aligned}(\boldsymbol{c} + \boldsymbol{a}) + (-\boldsymbol{a}) = \boldsymbol{b} + (-\boldsymbol{a}) &\Rightarrow \boldsymbol{c} + (\boldsymbol{a} - \boldsymbol{a}) = \boldsymbol{b} - \boldsymbol{a} \\ &\Rightarrow \boldsymbol{c} + \boldsymbol{0} = \boldsymbol{b} - \boldsymbol{a} \Rightarrow \boldsymbol{c} = \boldsymbol{b} - \boldsymbol{a}\,.\end{aligned}$$

Si ha così l'asserto. $\square$

Proposizione 5.2.9. (Unicità dell'elemento neutro) *L'elemento neutro di $\mathbb{V}$ rispetto all'addizione è unico.*

Dimostrazione. Se $\boldsymbol{0}$ e $\boldsymbol{0}'$ sono due elementi neutri di $\mathbb{V}$ rispetto all'addizione si ha

$$\boldsymbol{0} = \boldsymbol{0} + \boldsymbol{0}' = \boldsymbol{0}'\,. \qquad \square$$

Proposizione 5.2.10. (Unicità dell'elemento opposto) *L'opposto di un qualunque vettore $\boldsymbol{a} \in \mathbb{V}$ è unico.*

Dimostrazione. Se $\boldsymbol{b}$ e $\boldsymbol{c}$ sono due vettori opposti di $\boldsymbol{a}$, si ha:

$$\boldsymbol{b} = \boldsymbol{b} + \boldsymbol{0} = \boldsymbol{b} + (\boldsymbol{a} + \boldsymbol{c}) = (\boldsymbol{b} + \boldsymbol{a}) + \boldsymbol{c} = \boldsymbol{0} + \boldsymbol{c} = \boldsymbol{c}\,. \qquad \square$$

La proposizione che segue si prova facilmente facendo uso degli assiomi $SV1$, $SV2, \ldots, SV8$.

PROPOSIZIONE 5.2.11. *Per ogni vettore $\boldsymbol{a} \in \mathbb{V}$ e per ogni scalare $\alpha \in \mathbb{F}$, si ha:*

(1) $\alpha \mathbf{0} = \mathbf{0}$;

(2) $0\,\boldsymbol{a} = \mathbf{0}$;

(3) $\alpha \boldsymbol{a} = \mathbf{0} \;\Rightarrow\; \alpha = 0$ *o* $\boldsymbol{a} = \mathbf{0}$;

(4) $\alpha(-\boldsymbol{a}) = -\alpha \boldsymbol{a}$;

(5) $(-\alpha)\boldsymbol{a} = -\alpha \boldsymbol{a}$;

(6) $-\alpha(\boldsymbol{a} - \boldsymbol{b}) = \alpha \boldsymbol{a} - \alpha \boldsymbol{b}$;

(7) $(\alpha - \beta)\boldsymbol{a} = \alpha \boldsymbol{a} - \beta \boldsymbol{a}$.

DIMOSTRAZIONE.

(1) $\alpha \boldsymbol{a} = \alpha(\boldsymbol{a} + \mathbf{0}) = \alpha \boldsymbol{a} + \alpha \mathbf{0} \;\Rightarrow\; \alpha \mathbf{0} = \alpha \boldsymbol{a} - \alpha \boldsymbol{a} = \mathbf{0}$;

(2) $\alpha \boldsymbol{a} = (\alpha + 0)\boldsymbol{a} = \alpha \boldsymbol{a} + 0\boldsymbol{a} \;\Rightarrow\; 0\boldsymbol{a} = \alpha \boldsymbol{a} - \alpha \boldsymbol{a} = \mathbf{0}$;

(3) $\alpha \boldsymbol{a} = \mathbf{0}$ e $\alpha \neq \mathbf{0} \;\Rightarrow\; \boldsymbol{a} = 1\boldsymbol{a} = (\alpha^{-1}\alpha)\boldsymbol{a} = \alpha^{-1}(\alpha \boldsymbol{a}) = \alpha^{-1}\mathbf{0} = \mathbf{0}$;

(4) $\alpha \boldsymbol{a} + \alpha(-\boldsymbol{a}) = \alpha(\boldsymbol{a} - \boldsymbol{a}) = \alpha \mathbf{0} = \mathbf{0} \;\Rightarrow\; \alpha(-\boldsymbol{a}) = -\alpha \boldsymbol{a}$;

(5) $\alpha \boldsymbol{a} + (-\alpha)\boldsymbol{a} = (\alpha - \alpha)\boldsymbol{a} = 0\boldsymbol{a} = \mathbf{0} \;\Rightarrow\; (-\alpha)\boldsymbol{a} = -\alpha \boldsymbol{a}$;

(6) $\alpha(\boldsymbol{a} - \boldsymbol{b}) = \alpha[\boldsymbol{a} + (-\boldsymbol{b})] = \alpha \boldsymbol{a} + \alpha(-\boldsymbol{b}) = \alpha \boldsymbol{a} - \alpha \boldsymbol{b}$;

(7) $(\alpha - \beta)\boldsymbol{a} = [\alpha + (-\beta)]\boldsymbol{a} = \alpha \boldsymbol{a} + (-\beta)\boldsymbol{a} = \alpha \boldsymbol{a} - \beta \boldsymbol{a}$. □

5.3 Combinazioni lineari

Sia $\mathbb{V}$ uno spazio vettoriale sul campo $\mathbb{F}$ e si considerino h vettori $\boldsymbol{a}_1$, $\boldsymbol{a}_2, \ldots,$ $\boldsymbol{a}_h$ e h scalari $\alpha_1, \alpha_2, \ldots, \alpha_h$. Il vettore

$$\boldsymbol{a} = \alpha_1 \boldsymbol{a}_1 + \alpha_2 \boldsymbol{a}_2 + \cdots + \alpha_h \boldsymbol{a}_h \tag{5.8}$$

prende il nome di *combinazione lineare* dei vettori $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h$, mediante gli scalari $\alpha_1, \alpha_2, \ldots, \alpha_h$. Quando vale la (5.8) si dice anche che il vettore $\boldsymbol{a}$ *dipende linearmente* dai vettori $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h$.

L'insieme di tutte le possibili combinazioni lineari dei vettori $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, …, $\boldsymbol{a}_h$ si chiama *chiusura lineare* dell'insieme $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h\}$ e si denota con

$$\mathcal{L}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h);$$

si pone, cioè,

$$\mathcal{L}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h) = \{\lambda_1 \boldsymbol{a}_1 + \lambda_2 \boldsymbol{a}_2 + \cdots + \lambda_h \boldsymbol{a}_h \;:\; \lambda_1, \lambda_2, \ldots, \lambda_h \in \mathbb{F}\}. \tag{5.9}$$

Esempio 5.3.1. La chiusura lineare dell'insieme

$$X = \{\boldsymbol{a} = (1,2), \boldsymbol{b} = (2,1)\}$$

in $\mathbb{R}^2$ è

$$\mathcal{L}(X) = \{\alpha(1,2) + \beta(2,1) \ : \ \alpha, \beta \in \mathbb{F}\} = \{(\alpha + 2\beta, 2\alpha + \beta) \ : \ \alpha, \beta \in \mathbb{F}\}.$$

Nelle due tabelle che seguono riportiamo alcuni vettori di $\mathcal{L}(X)$ ottenuti in corrispondenza di particolari valori di α e β :

α	β	$\alpha\boldsymbol{a} + \beta\boldsymbol{b}$
0	0	$(0,0)$
1	0	$(1,2)$
0	1	$(2,1)$

α	β	$\alpha\boldsymbol{a} + \beta\boldsymbol{b}$
1	1	$(3,3)$
1	-1	$(-1,1)$
2	1	$(4,5)$

Il Lettore, per esercizio, provi a controllare se risulta $\mathcal{L}(X) = \mathbb{R}^2$. □

Esempio 5.3.2. In $\mathbb{F}^3$, la chiusura lineare dell'insieme

$$X = \{(1,0,0),\ (0,1,0),\ (0,0,1)\}$$

è uguale ad $\mathbb{F}^3$; infatti

$$\begin{aligned}\mathcal{L}(X) =& \{\alpha(1,0,0) + \beta(0,1,0) + \gamma(0,0,1) \ : \ \alpha, \beta, \gamma \in \mathbb{F}\} \\ =& \{(\alpha, \beta, \gamma) \ : \ \alpha, \beta, \gamma \in \mathbb{F}\} = \mathbb{F}^3.\end{aligned}$$ □

Esempio 5.3.3. La chiusura lineare dell'insieme

$$X = \{(1,0,0,0),\ (0,1,1,0),\ (0,0,1,1)\}$$

in $\mathbb{F}^4$ è

$$\begin{aligned}\mathcal{L}(X) =& \{\alpha(1,0,0,0) + \beta(0,1,1,0) + \gamma(0,0,1,1) \ : \ \alpha, \beta, \gamma \in \mathbb{F}\} \\ =& \{(\alpha, \beta, \beta + \gamma, \gamma) \ : \ \alpha, \beta, \gamma \in \mathbb{F}\}.\end{aligned}$$

Il Lettore verifichi che i vettori $(1,-1,0,1), (-1,-1,-1,0)$ appartengono a $\mathcal{L}(X)$, mentre non vi appartengono i vettori $(1,1,1,1), (-1,-1,1,-1)$. □

Esempio 5.3.4. In un campo $\mathbb{F}$, considerato come spazio vettoriale su se stesso, risulta $\mathcal{L}(1) = \{\alpha 1 \ : \ \alpha \in \mathbb{F}\} = \mathbb{F}$. Inoltre, poiché per ogni scalare non nullo α è $\alpha^{-1}\alpha = 1$, risulta anche $\mathcal{L}(\alpha) = \mathbb{F}$, per ogni $\alpha \in \mathbb{F}$ diverso da zero. □

Esempio 5.3.5. La chiusura lineare dell'insieme

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

in $\mathbb{F}^{3,3}$ è l'insieme della matrici diagonali d'ordine 3. Si ha, infatti, per ogni $\alpha, \beta, \gamma \in \mathbb{F}$

$$\alpha \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \beta \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \gamma \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix}.$$ □

ESEMPIO 5.3.6. Sia n un intero non negativo. La chiusura lineare dell'insieme

$$X = \{1, x, x^2, \dots, x^n\}$$

in $\mathbb{F}[x]$ è l'insieme dei polinomi di grado non superiore ad n. Si ha, infatti,

$$\mathcal{L}(X) = \{a_o + a_1 x + a_2 x^2 + \cdots + a_n x^n : \ a_o, a_1, \dots, a_n \in \mathbb{F}\}.$$

Allo stesso modo si vede che la chiusura lineare di $\{x, x^2, \dots, x^n\}$ è l'insieme dei polinomi di grado non superiore ad n con termine noto nullo. □

ESEMPIO 5.3.7. Sia n un intero non negativo. La chiusura lineare dell'insieme

$$X = \{1, x_1, x_2, \dots, x_n\}$$

nello spazio dei polinomi lineari $\mathbb{F}_1[\boldsymbol{x}]$ è l'insieme di tutti i polinomi lineari nelle variabili $x_1, x_2, \dots .x_n$. Si ha, infatti,

$$\mathcal{L}(X) = \{a_o + a_1 x_1 + a_2 x_2 + \cdots + a_n x_n : \ a_o, a_1, \dots, a_n \in \mathbb{F}\}.$$ □

ESEMPIO 5.3.8. Sia S un sistema lineare omogeneo e ridotto con m incognite principali e denotiamo con $\boldsymbol{c}_{m+1}, \boldsymbol{c}_{m+2}, \dots, \boldsymbol{c}_n$ le sue soluzioni principali. Allora, in forza delle prop.4.5.1 e 4.5.3, risulta

$$Sol(S) = \mathcal{L}(\boldsymbol{c}_{m+1}, \boldsymbol{c}_{m+2}, \dots, \boldsymbol{c}_n).$$ □

La nozione di chiusura lineare si estende in modo naturale al caso di insiemi infiniti di vettori. Se Y è un insieme infinito di vettori di $\mathbb{V}$, si dice *chiusura lineare*, o semplicemente *chiusura*, di Y, e si denota con $\mathcal{L}(Y)$, l'unione degli insiemi $\mathcal{L}(X)$ ottenuti al variare di X tra tutti i sottoinsiemi finiti di Y, cioè

$$\mathcal{L}(Y) = \bigcup \{\mathcal{L}(X) \ : \ X \text{ sottoinsieme finito di } Y\}. \tag{5.10}$$

ESEMPIO 5.3.9. Nello spazio vettoriale dei polinomi $\mathbb{F}[x]$ risulta

$$\mathcal{L}(\{1, x, x^2, \dots, x^n, \dots\}) = \mathbb{F}[x].$$

Ogni polinomio, infatti, è una combinazione lineare di un numero finito di polinomi del tipo x^j. □

PROPOSIZIONE 5.3.10. *Per ogni insieme X di vettori di $\mathbb{V}$, risulta*

$$X \subseteq \mathcal{L}(X) \ \ e \ \ \mathcal{L}(\mathcal{L}(X)) = \mathcal{L}(X). \tag{5.11}$$

DIMOSTRAZIONE. È lasciata per esercizio al Lettore. □

ESERCIZIO 5.3.11. *Provare che, se A, B sono due insiemi di vettori con $A \subseteq B$, risulta*

$$\mathcal{L}(A) \subseteq \mathcal{L}(B).$$

5.4 Dipendenza lineare

Le nozioni di dipendenza e indipendenza lineare, già introdotte per i vettori numerici e le matrici, sono del tutto generali e possono considerarsi in un qualsiasi spazio vettoriale. Sia, dunque, assegnato uno spazio vettoriale $\mathbb{V}$ su un campo $\mathbb{F}$. I vettori $\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_h$ di $\mathbb{V}$ si dicono *linearmente dipendenti*, o semplicemente *dipendenti*, se uno di essi dipende linearmente dai rimanenti, cioè se

$$\boldsymbol{a}_i = \alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2 + \cdots + \alpha_{i-1}\boldsymbol{a}_{i-1} + \alpha_{i+1}\boldsymbol{a}_{i+1} + \cdots + \alpha_h\boldsymbol{a}_h, \tag{5.12}$$

per qualche indice $i = 1, 2, \ldots, h$. In quest'ultimo caso si dice anche che l'insieme $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h\}$ è *linearmente dipendente*, o *dipendente*. Vettori (in numero finito) o insiemi finiti di vettori che non siano dipendenti si dicono *linearmente indipendenti*, o *indipendenti*.

Proposizione 5.4.1. *I vettori $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h$ di $\mathbb{V}$ sono linearmente dipendenti se, e solo se, esiste una loro combinazione lineare a coefficienti non tutti nulli uguale al vettore nullo, cioè*

$$\alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2 + \cdots + \alpha_h\boldsymbol{a}_h = \mathbf{0} \tag{5.13}$$

con qualche scalare $\alpha_i \in \mathbb{F}$ diverso da zero.

Dimostrazione. Se $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h$ sono linearmente dipendenti, vale la relazione (5.12), per almeno un indice i. Allora, sommando $-\boldsymbol{a}_i$ a primo e secondo membro della (5.12), risulta

$$\alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2 + \cdots + \alpha_{i-1}\boldsymbol{a}_{i-1} - \boldsymbol{a}_i + \alpha_{i+1}\boldsymbol{a}_{i+1} + \cdots + \alpha_h\boldsymbol{a}_h = \mathbf{0},$$

con il coefficiente di $\boldsymbol{a}_i$ uguale a $-1 \neq 0$. Viceversa, se vale la (5.13), con $\alpha_i \neq 0$ per qualche indice i, sommando $-\alpha_i\boldsymbol{a}_i$ a primo e secondo membro della (5.13) e dividendo per $-\alpha_i$ la successiva uguaglianza, risulta

$$\boldsymbol{a}_i = -\frac{\alpha_1}{\alpha_i}\boldsymbol{a}_1 - \frac{\alpha_2}{\alpha_i}\boldsymbol{a}_2 - \cdots - \frac{\alpha_{i-1}}{\alpha_i}\boldsymbol{a}_{i-1} - \frac{\alpha_{i+1}}{\alpha_i}\boldsymbol{a}_{i+1} - \cdots - \frac{\alpha_h}{\alpha_i}\boldsymbol{a}_h$$

e i nostri vettori sono linearmente dipendenti. □

Corollario 5.4.2. *I vettori $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h$ di $\mathbb{V}$ sono linearmente indipendenti se, e solo se, la loro combinazione lineare con coefficienti tutti nulli è l'unica che risulta uguale al vettore nullo, cioè*

$$\alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2 + \cdots + \alpha_h\boldsymbol{a}_h = \mathbf{0} \ \Leftrightarrow\ \alpha_1 = \alpha_2 = \cdots = \alpha_h = 0. \tag{5.14}$$

Per convenzione, assumiamo che l'insieme vuoto sia indipendente. Ne segue che un insieme dipendente è sempre non vuoto.

Esempio 5.4.3. Un sottoinsieme di $\mathbb{V}$ contenente un unico elemento $\boldsymbol{a}$ è indipendente se, e solo se, il vettore $\boldsymbol{a}$ è non nullo. □

Esempio 5.4.4. (Dipendenza lineare di due vettori) Due vettori $\boldsymbol{a}$ e $\boldsymbol{b}$ si dicono *proporzionali*, o *paralleli*, se esiste uno scalare α tale che $\boldsymbol{b} = \alpha\boldsymbol{a}$. Allora, segue subito dalla definizione che due vettori distinti sono dipendenti se, e solo se, sono proporzionali. In particolare, è dipendente ogni insieme del tipo $\{\boldsymbol{a}, -\boldsymbol{a}\}$, con $\boldsymbol{a}$ vettore non nullo. □

Esempio 5.4.5. I vettori $(0,1), (1,0) \in \mathbb{R}^2$, non essendo proporzionali, sono indipendenti. Verifichiamolo anche direttamente: se $\lambda, \mu \in \mathbb{R}$, si ha

$$(0,0) = \lambda(1,0) + \mu(0,1) = (\lambda, 0) + (0, \mu) = (\lambda, \mu)\,,$$

da cui segue $\lambda = \mu = 0$. □

Esempio 5.4.6. I vettori $(0,1), (1,0), (2,3) \in \mathbb{R}^2$ sono dipendenti. Infatti, se $\lambda, \mu, \nu \in \mathbb{R}$, si ha

$$(0,0) = \lambda(0,1) + \mu(1,0) + \nu(2,3) = (\mu + 2\nu, \lambda + 3\nu)$$

ed esistono soluzioni non banali del sistema

$$\mu + 2\nu = 0 \;\;,\;\; \lambda + 3\nu = 0\,.$$

Una di tali soluzioni è, per esempio, $(\lambda, \mu, \nu) = (-3, -2, 1)$ e in corrispondenza di questa si ha

$$-3(0,1) - 2(1,0) + (2,3) = (0,0)\,.$$ □

Esempio 5.4.7. Siano $n_1, n_2, \ldots, n_h$ interi non negativi distinti. In $\mathbb{F}[x]$, l'insieme di polinomi

$$X = \{x^{n_1}, x^{n_2}, \ldots, x^{n_h}\}$$

è indipendente. Infatti, una combinazione lineare degli elementi di X mediante scalari $a_1, a_2, \ldots, a_h$ è il polinomio $a_1x^{n_1} + a_2x^{n_2} + \cdots + a_hx^{n_h}$ e questo è il polinomio nullo se, e solo se, $a_1 = a_2 = \cdots = a_h = 0$. In particolare, per ogni intero non negativo n, risulta indipendente l'insieme

$$X = \{1, x, x^2, \ldots, x^n\}.$$ □

Esempio 5.4.8. Sia $A = (a_{ij}) \in \mathbb{F}^{m,n}$ una matrice di tipo $m \times n$ sul campo $\mathbb{F}$ e sia A' la matrice ottenuta ta A scambiando tra loro le colonne di posto h e k. È facile verificare che le righe di A sono indipendenti (risp. dipendenti) se, e solo se, lo sono quelle di A'.

Esempio 5.4.9. Sia $A = (a_{ij}) \in \mathbb{F}^{m,n}$ una matrice di tipo $m \times n$ sul campo $\mathbb{F}$ e si supponga che A contenga come sottomatrice la matrice identica I_m d'ordine m. Allora le righe di A sono linearmente indipendenti. Infatti, nell'ipotesi che I_m sia individuata dalle prime m colonne di A (cosa che non lede le generalità del problema (*cfr.* esempio 5.4.8)), cioè $A = (I_m|A')$, una combinazione lineare delle sue righe con gli scalari $\lambda_1, \lambda_2, \ldots, \lambda_m$ è un vettore di $\mathbb{F}^n$ del tipo

$$(\lambda_1, \lambda_2, \ldots, \lambda_m\,, \ldots)$$

e, quindi, è nullo se, e solo se,

$$\lambda_1 = \lambda_2 = \cdots = \lambda_m = 0\,.$$ □

Le nozioni di dipendenza e indipendenza lineare per gli insiemi finiti si estendono in modo naturale alle $h-$ple ordinate di vettori. Si dice, infatti, che una $h-$pla $(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h)$ è *dipendente* o *indipendente*, se valgono rispettivamente la (5.12) o la (5.14). Si noti che, a differenza degli elementi di un insieme, le componenti di una $h-$pla non sono necessariamente a due a due distinte. Nel seguito un'$h-$pla di vettori sarà anche detta *sistema* di vettori e le componenti dell'$h-$pla si diranno *vettori* o *elementi* del sistema. È chiaro che ogni proprietà relativa alla dipendenza o indipendenza di insiemi finiti si traduce in una corrispondente proprietà dei sistemi, e viceversa.

OSSERVAZIONE 5.4.10. La considerazione di sistemi, piuttosto che di insiemi, di vettori è particolarmente utile quando si lavora con le righe e le colonne di una matrice. Le righe (risp. colonne) di una matrice, infatti, non sono vettori numerici necessariamente distinti e, quindi, si deve parlare di *sistema delle righe* (*risp. colonne*) *di una matrice* e non di insieme delle righe (risp. colonne), anche se a volte, con abuso di linguaggio, si usa anche questa seconda dicitura. □

ESEMPIO 5.4.11. Se un sistema lineare di m equazioni in n incognite $AX = \boldsymbol{b}$ a coefficienti in un campo $\mathbb{F}$ è compatibile, allora il sistema di vettori

$$(\boldsymbol{a}^{(1)}, \boldsymbol{a}^{(2)}, \ldots, \boldsymbol{a}^{(n)}, \boldsymbol{b})$$

formato dalle colonne di A e da $\boldsymbol{b}$ è linearmente dipendente (*cfr.* esempio 4.2.3). □

PROPOSIZIONE 5.4.12. *Siano X, Y due sistemi di vettori tali che ogni vettore di X dipende linearmente da Y. Allora ogni vettore dipendente linearmente da X dipende anche da Y.*

DIMOSTRAZIONE. È un facile esercizio lasciato al Lettore. □

PROPOSIZIONE 5.4.13. *Sia $\mathcal{I} = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h\}$ un insieme indipendente di vettori. Allora ogni sottoinsieme $\mathcal{J}$ di $\mathcal{I}$ è indipendente.*

DIMOSTRAZIONE. Poiché l'insieme vuoto è indipendente, non è restrittivo supporre $\mathcal{J} \neq \varnothing$ e

$$\mathcal{J} = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k\} \ , \ \text{con } k < h \, .$$

Se consideriamo una combinazione lineare dei vettori di $\mathcal{J}$ uguale al vettore nullo, essendo $\mathcal{I}$ indipendente, abbiamo

$$\begin{aligned}
&\alpha_1 \boldsymbol{a}_1 + \alpha_2 \boldsymbol{a}_2 \cdots + \alpha_k \boldsymbol{a}_k = \mathbf{0} \\
\Rightarrow &\alpha_1 \boldsymbol{a}_1 + \alpha_2 \boldsymbol{a}_2 \cdots + \alpha_k \boldsymbol{a}_k + 0\boldsymbol{a}_{k+1} + \cdots + 0\boldsymbol{a}_h = \mathbf{0} \\
\Rightarrow &\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0 \, ,
\end{aligned}$$

cioè l'asserto. □

COROLLARIO 5.4.14. *Sia $\mathcal{D}$ un insieme dipendente. Allora ogni insieme finito di vettori contenente $\mathcal{D}$ è dipendente. In particolare, un insieme finito di vettori contenente il vettore nullo è dipendente.*

Dimostrazione. È una immediata conseguenza della proposizione precedente. □

Proposizione 5.4.15. *Un sistema* $(\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_h)$ *contenente due vettori uguali è dipendente.*

Dimostrazione. Supponiamo, per esempio, $\boldsymbol{a}_1 = \boldsymbol{a}_2$. Allora risulta

$$1\boldsymbol{a}_1 + (-1)\boldsymbol{a}_2 + 0\boldsymbol{a}_3 + \cdots + 0\boldsymbol{a}_h = \boldsymbol{a}_1 - \boldsymbol{a}_1 = \boldsymbol{0}$$

e, quindi, il nostro sistema è dipendente. □

Proposizione 5.4.16. *Siano* $\mathcal{I} = \{\boldsymbol{a}_1,\, \boldsymbol{a}_2, \dots, \boldsymbol{a}_h\}$ *un insieme indipendente di vettori e* $\boldsymbol{a} \in \mathbb{V}$. *Se i vettori* $\boldsymbol{a}, \boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_h$ *sono dipendenti, allora* $\boldsymbol{a}$ *si scrive in modo unico come combinazione lineare di* $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_h$. *Esistono, cioè, e sono univocamente determinati, h scalari* $\alpha_1, \alpha_2, \dots, \alpha_h$ *tali che*

$$\boldsymbol{a} = \alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2 + \cdots + \alpha_h\boldsymbol{a}_h.$$

Dimostrazione. Supponiamo

$$\boldsymbol{a} = \alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2 + \cdots + \alpha_h\boldsymbol{a}_h = \beta_1\boldsymbol{a}_1 + \beta_2\boldsymbol{a}_2 + \cdots + \beta_h\boldsymbol{a}_h\,.$$

Allora

$$\begin{aligned}
&\alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2 + \cdots + \alpha_h\boldsymbol{a}_h - (\beta_1\boldsymbol{a}_1 + \beta_2\boldsymbol{a}_2 + \cdots + \beta_h\boldsymbol{a}_h) = \boldsymbol{0}\\
\Rightarrow\ &\alpha_1\boldsymbol{a}_1 - \beta_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a} - \beta_2\boldsymbol{a}_2 + \cdots + \alpha_h\boldsymbol{a} - \beta_h\boldsymbol{a}_h = \boldsymbol{0}\\
\Rightarrow\ &(\alpha_1 - \beta_1)\boldsymbol{a}_1 + (\alpha_2 - \beta_2)\boldsymbol{a}_2 + \cdots + (\alpha_h - \beta_h)\boldsymbol{a}_h = \boldsymbol{0}\\
\Rightarrow\ &\alpha_1 - \beta_1 = 0,\ \alpha_2 - \beta_2 = 0, \cdots, \alpha_h - \beta_h = 0\\
\Rightarrow\ &\alpha_1 = \beta_1,\ \alpha_2 = \beta_2, \cdots, \alpha_h = \beta_h\,,
\end{aligned}$$

cioè l'asserto. □

Proposizione 5.4.17. *Siano* $\mathcal{I} = \{\boldsymbol{a}_1,\, \boldsymbol{a}_2, \dots, \boldsymbol{a}_h\}$ *un insieme di vettori e* $\boldsymbol{a} \in \mathbb{V}$. *Se* $\boldsymbol{a}$ *si scrive in un unico modo come combinazione lineare di* $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_h$, *allora* $\{\boldsymbol{a}_1,\, \boldsymbol{a}_2, \dots, \boldsymbol{a}_h\}$ *è linearmente indipendente.*

Dimostrazione. Sia

$$\boldsymbol{a} = \alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2 + \cdots + \alpha_h\boldsymbol{a}_h$$

l'unico modo di scrivere $\boldsymbol{a}$ come combinazione lineare di $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_h$ e sia

$$\lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2 + \cdots \lambda_h\boldsymbol{a}_h = \boldsymbol{0}.$$

Allora risulta

$$\begin{aligned}
\boldsymbol{a} &= \alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2 + \cdots + \alpha_h\boldsymbol{a}_h\\
&= (\alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2 + \cdots + \alpha_h\boldsymbol{a}_h) + (\lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2 + \cdots + \lambda_h\boldsymbol{a}_h)\\
&= (\alpha_1 - \lambda_1)\boldsymbol{a}_1 + (\alpha_2 - \lambda_2)\boldsymbol{a}_2 + \cdots + (\alpha_h - \lambda_h)\boldsymbol{a}_h\\
&\Rightarrow \lambda_1 = \lambda_2 = \cdots = \lambda_h = 0
\end{aligned}$$

e $\{\boldsymbol{a}_1,\, \boldsymbol{a}_2, \dots, \boldsymbol{a}_h\}$ è un insieme indipendente. □

5.5 Teoremi di unicità per i sistemi lineari

In tutto il paragrafo supporremo fissato un sistema lineare

$$S = \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} , \tag{5.15}$$

di m equazioni lineari in n incognite a coefficienti in un campo $\mathbb{F}$ che, come al solito, scriveremo anche in forma compatta

$$A\boldsymbol{x} = \boldsymbol{b}\,, \text{ con } A = (a_{ij}) \text{ e } \boldsymbol{b} = (b_1, b_2, \ldots, \boldsymbol{b}_n)^t.$$

Denoteremo, inoltre, con $\boldsymbol{a}^{(1)}, \boldsymbol{a}^{(2)}, \ldots, \boldsymbol{a}^{(n)}$ le colonne della matrice A.

Nell'ipotesi che tale sistema sia compatibile, vogliamo studiare le condizioni su A in corrispondenza delle quali il sistema ammette una soluzione unica verificante le condizioni stesse.

Proposizione 5.5.1. (Primo teorema di unicità per i sistemi lineari) *Un sistema lineare compatibile a coefficienti in un campo $\mathbb{F}$ possiede un'unica soluzione se, e solo se, il sistema delle colonne della matrice dei coefficienti è linearmente indipendente.*

Dimostrazione. Consideriamo il sistema lineare (5.15). Abbiamo già osservato che un vettore $\boldsymbol{c}$ è soluzione del il sistema se, e solo se, verifica la relazione

$$\boldsymbol{b} = A\boldsymbol{c} = c_1\boldsymbol{a}^{(1)} + c_2\boldsymbol{a}^{(2)} + \cdots + c_n\boldsymbol{a}^{(n)}\,,$$

cioè se $\boldsymbol{b}$ è combinazione lineare delle colonne di A mediante gli scalari c_1, $c_2, \ldots, c_n$. L'asserto è, allora, una conseguenza delle prop. 5.4.16 e 5.4.17. □

Proposizione 5.5.2. (Secondo teorema di unicità per i sistemi lineari) *Assegnato il sistema lineare* (5.15), *sia $C = (\boldsymbol{a}^{(1)}, \boldsymbol{a}^{(2)}, \ldots, \boldsymbol{a}^{(n)})$ il sistema delle colonne di A e*

$$C' = (\boldsymbol{a}^{(j_1)}, \boldsymbol{a}^{(j_2)}, \ldots, \boldsymbol{a}^{(j_h)})^8$$

un sistema massimo di colonne linearmente indipendenti di A. Allora, se il sistema è compatibile, ciascuna soluzione corrisponde biunivocamente ad un prefissato valore delle $n-h$ incognite $x_{t_1}, x_{t_2}, \ldots, x_{t_{n-h}}$ diverse dalle $x_{j_1}, x_{j_2}, \ldots, x_{j_h}$; in altre parole, fissati liberamente $n-h$ scalari $\alpha_{t_1}, \alpha_{t_2}, \ldots, \alpha_{t_{n-h}}$, esiste un'unica soluzione del sistema $\boldsymbol{c} = (c_1, c_2, \ldots, c_n)$ tale che

$$c_{t_1} = \alpha_{t_1}, c_{t_2} = \alpha_{t_2}, \ldots, c_{t_{n-h}} = \alpha_{t_{n-h}}\,.$$

[8] Questo significa che se a C' aggiungiamo una qualsiasi colonna di A otteniamo un sistema dipendente.

Dimostrazione. Per semplificare le notazioni, senza ledere le generalià delle nostre ipotesi, supponiamo che C' sia il sistema delle prime h colonne di A. Se nella (5.15) poniamo

$$x_{h+1} = \alpha_{h+1}\,,\ x_{h+2} = \alpha_{h+2}\,, \dots, x_n = \alpha_n$$

otteniamo il sistema

$$S' = \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1h}x_h = d_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2h}x_h = d_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mh}x_h = d_h \end{cases} \quad , \tag{5.16}$$

con

$$d_i = b_i - a_{i(h+1)}\alpha_{h+1} - a_{i(h+2)}\alpha_{h+2} - \cdots - a_{in}\alpha_n\,,$$

per ogni $i = h+1, h+2, \dots, n$, cioè

$$\boldsymbol{d} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_h \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_h \end{pmatrix} - \begin{pmatrix} a_{1(h+1)} \\ a_{2(h+1)} \\ \vdots \\ a_{h(h+1)} \end{pmatrix} \alpha_{h+1} - \cdots - \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{hn} \end{pmatrix} \alpha_n$$

Ora , il vettore $\boldsymbol{b}$ dipende linearmente da C perchè S è compatibile (*cfr.* osservazione (5.4.11) e, quindi, dipende linearmente da C' (*cfr.* prop.5.4.12). Ne segue che

$$(d_1, d_2, \dots, d_h)^t$$

dipende linearmente dalle colonne dei coefficienti di S'. D'altra parte, anche i vettori $(a_{1s}, a_{2s}, \dots, a_{hs})^t$, $s = h+1, h+2, \dots, n$, dipendono linearmente da C'. Ne segue che S' verifica le ipotesi del primo teorema di unicità e, quindi, ammette un'unica soluzione $(c_1, c_2, \dots, c_h)$. Allora S ha la soluzione

$$\boldsymbol{c} = (c_1, c_2, \dots, c_h, \alpha_{h+1}, \alpha_{h+2}, \dots, \alpha_n).$$

Ora , se $\boldsymbol{c}'$ è una soluzione di S del tipo

$$\boldsymbol{c}' = (c_1', c_2', \dots, c_h', \alpha_{h+1}, \alpha_{h+2}, \dots, \alpha_n)\,,$$

$(c_1', c_2', \dots, c_h')$ è una soluzione di S' e, ammettendo S' l'unica soluzione $\boldsymbol{c}$, risulta

$$c_1' = c_1\,,\ c_2' = c_2\,, \dots, c_h' = c_h.$$

□

5.6 Sottospazi vettoriali

Sia $\mathbb{V}$ uno spazio vettoriale sul campo $\mathbb{F}$. Un sottoinsieme non vuoto U di $\mathbb{V}$ si dice *sottospazio vettoriale*, o semplicemente *sottospazio*, di $\mathbb{V}$ se sono verificate le seguenti due proprietà:

1. *proprietà di chiusura rispetto all'addizione*:
$$\text{per ogni } \boldsymbol{a}, \boldsymbol{b} \in U \ \Rightarrow \ \boldsymbol{a} + \boldsymbol{b} \in U; \tag{5.17}$$
2. *proprietà di chiusura rispetto alla moltiplicazione per scalari*:
$$\text{per ogni } \boldsymbol{a} \in U \text{ e per ogni } \alpha \in \mathbb{F} \ \Rightarrow \ \alpha\boldsymbol{a} \in U. \tag{5.18}$$

Per indicare che U è sottospazio di $\mathbb{V}$ si usa la notazione $U \leq \mathbb{V}$. Tra i sottospazi di $\mathbb{V}$ ve ne sono due detti *banali*: $\mathbb{V}$ stesso e il singleton del vettore nullo $\{\mathbf{0}\}$. Quest'ultimo sottospazio si chiama *sottospazio nullo* e, con abuso di notazione, si denota anche con $\mathbf{0}$.

Osservazione 5.6.1. Sia U un sottospazio di $\mathbb{V}$. Allora $\mathbb{V}$ induce su U una operazione di addizione fra vettori e una di moltiplicazione per uno scalare. Rispetto a tali operazioni, U risulta a sua volta uno spazio vettoriale su $\mathbb{F}$. È chiaro, quindi, cosa si debba intendere quando nel seguito si dirà che un sottoinsieme W di U è un sottospazio di U. □

Esempio 5.6.2. Consideriamo lo spazio vettoriale $\mathcal{V}_O$ dei segmenti orientati di origine un punto O in $\mathbb{E}_3$. È facile provare che i segmenti orientati di origine O e contenuti in un piano o in una retta per O formano un sottospazio vettoriale di $\mathcal{V}_O$. □

Esempio 5.6.3. (Soluzioni di un sistema lineare omogeneo) L'insieme delle soluzioni di un sistema lineare omogeneo in n incognite a coefficienti in un campo $\mathbb{F}$ è un sottospazio vettoriale di $\mathbb{F}^n$ (*cfr.* corollario 4.5.2). □

Esempio 5.6.4. (Matrici triangolari e diagonali) Gli insiemi delle matrici triangolari (superiori o inferiori) e di quelle diagonali d'ordine un fissato intero positivo n su un campo $\mathbb{F}$ sono sottospazi vettoriali di $\mathbb{F}^{n,n}$. □

Esempio 5.6.5. (Matrici simmetriche) Una matrice quadrata $A = (a_{ij})$ d'ordine n su un campo $\mathbb{F}$ si dice *simmetrica* se risulta $a_{ij} = a_{ji}$, per ogni $i, j = 1, 2, \dots, n$. Per esempio, le matrici simmetriche d'ordine due e tre sono rispettivamente del tipo
$$\begin{pmatrix} a & c \\ c & b \end{pmatrix} \text{ e } \begin{pmatrix} a & d & e \\ d & b & f \\ e & f & c \end{pmatrix},$$
con $a, b, c, d, e, f \in \mathbb{F}$. Le matrici diagonali sono particolari matrici simmetriche. É un facile esercizio verificare che l'insieme $\mathbb{F}^{n,n}_s$ delle matrici simmetriche d'ordine n su $\mathbb{F}$ verifica le (5.17) e (5.18) e, quindi, è un sottospazio vettoriale di $\mathbb{F}^{n,n}$. □

Esempio 5.6.6. (Matrici antisimmetriche) Una matrice quadrata $A = (a_{ij})$ d'ordine n su un campo $\mathbb{F}$ si dice *antisimmetrica* se risulta $a_{ij} = -a_{ji}$, per ogni $i, j = 1, 2, \dots, n$. In particolare deve essere $a_{ii} = -a_{ii}$ e di conseguenza è

$a_{ii} = 0$; cioè una matrice antisimmetrica ha la diagonale principale nulla. Per esempio, le matrici antisimmetriche d'ordine due e tre sono rispettivamente del tipo

$$\begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} \text{ e } \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix},$$

con $a, b, c \in \mathbb{F}$. É un facile esercizio verificare che l'insieme $\mathbb{F}_a^{n,n}$ delle matrici antisimmetriche d'ordine n su $\mathbb{F}$ verifica le (5.17) e (5.18) e, quindi, è un sottospazio vettoriale di $\mathbb{F}^{n,n}$. □

Esempio 5.6.7. (Polinomi di grado non superiore ad n) L'insieme, che denotiamo con $\mathbb{F}_n[x]$, dei polinomi a coefficienti in un campo $\mathbb{F}$ e di grado non superiore ad un fissato intero non negativo n, in forza della (5.4), è un sottospazio vettoriale di $\mathbb{F}[x]$. Nei casi $n = 0, 1, 2$ si ha rispettivamente

$$\mathbb{F}_0[x] = \mathbb{F}\,, \ \ \mathbb{F}_1[x] = \{a + bx \ : \ a, b \in \mathbb{F}\}\,,$$

$$\mathbb{F}_2[x] = \{a + bx + cx^2 \ : \ a, b, c \in \mathbb{F}\}\,.$$

È facile rendersi conto che risulta $\mathbb{F}_0[x] \leq \mathbb{F}_1[x] \leq \mathbb{F}_2[x]$ e, più in generale, $\mathbb{F}_m[x] \leq \mathbb{F}_n[x]$ per ogni $m \leq n$. È da osservare che l'insieme dei polinomi di fissato grado n non è un sottospazio di $\mathbb{F}[x]$ perché la somma di due polinomi di grado n può avere grado minore di n. □

Esempio 5.6.8. (Forme lineari) Un polinomio lineare $a(x_j) \in \mathbb{F}_1[\boldsymbol{x}]$ si dice *forma lineare* se è $a_0 = 0$, cioè se

$$a(x_j) = a_1x_1 + a_2x_2 + \cdots + a_nx_n\,.$$

È facile verificare che le forme lineari costituiscono un sottospazio vettoriale di $\mathbb{F}_1[\boldsymbol{x}]$. □

Esempio 5.6.9. (Funzioni continue) Nello spazio vettoriale $\mathbb{R}^{\mathbb{R}}$ delle funzioni di $\mathbb{R}$ in $\mathbb{R}$ (*cfr.* esempio 5.2.7) consideriamo il sottoinsieme $C(\mathbb{R})$ di tutte le funzioni continue. Si prova che la somma di due funzioni continue e il prodotto di uno scalare per una funzione continua sono ancora funzioni continue. Ne segue che $C(\mathbb{R})$ è un sottospazio vettoriale di $\mathbb{R}^{\mathbb{R}}$. □

Proposizione 5.6.10. *Ogni sottospazio U di $\mathbb{V}$ contiene l'opposto di ogni suo elemento e il vettore nullo.*

Dimostrazione. Se $\boldsymbol{a}$ è un elemento di U, allora $-\boldsymbol{a} = (-1)\boldsymbol{a}$ e $\boldsymbol{0} = \boldsymbol{a}+(-\boldsymbol{a})$ appartengono ad U in forza delle (5.17) e (5.18), rispettivamente. □

Esercizio 5.6.11. *Provare che, se U è sottospazio di $\mathbb{V}$ e W è sottospazio di U, allora W è sottospazio di $\mathbb{V}$.*

Proposizione 5.6.12. *Un sottoinsieme non vuoto U di $\mathbb{V}$ è un sottospazio se, e solo se, è verificata la seguente proprietà:*

$$\text{per ogni } \boldsymbol{a},\, \boldsymbol{b} \in U \text{ e per ogni } \alpha,\, \beta \in \mathbb{F} \ \Rightarrow \ \alpha\boldsymbol{a} + \beta\boldsymbol{b} \in U. \tag{5.19}$$

Dimostrazione. Se vale la (5.19), le due proprietà che definiscono un sottospazio si ottengono per $(\alpha, \beta) = (1, 1)$ e $(\alpha, \beta) = (1, 0)$. La seconda parte della dimostrazione è immediata. $\square$

Corollario 5.6.13. *Un sottoinsieme non vuoto U di $\mathbb{V}$ è un sottospazio se, e solo se, contiene tutte le combinazioni lineari dei suoi vettori, cioè*

$$\mathcal{L}(U) = U\,. \tag{5.20}$$

Esempio 5.6.14. Le (5.20) e (5.11) assicurano che la chiusura lineare di un qualunque sottoinsieme di $\mathbb{V}$ è un sottospazio. $\square$

Proposizione 5.6.15. *L'intersezione di un insieme non vuoto $\mathcal{F}$ di sottospazi di $\mathbb{V}$ è un sottospazio di $\mathbb{V}$.*

Dimostrazione. L'intersezione di tutti i sottospazi in $\mathcal{F}$ è non vuota perchè il vettore nullo appartiene a ciascuno di questi sottospazi e, quindi, alla loro intersezione. Inoltre, si ha

$$\boldsymbol{a}, \boldsymbol{b} \in \bigcap_{U \in \mathcal{F}} U \quad \text{e} \quad \alpha, \beta \in \mathbb{F} \Rightarrow$$

$$\boldsymbol{a}, \boldsymbol{b} \in U, \text{ per ogni } U \in \mathcal{F} \Rightarrow$$

$$\alpha\boldsymbol{a} + \beta\boldsymbol{b} \in U, \text{ per ogni } U \in \mathcal{F} \text{ (}\textit{essendo ogni } U \textit{ un sottospazio}\text{)} \Rightarrow$$

$$\alpha\boldsymbol{a} + \beta\boldsymbol{b} \in U \in \bigcap_{U \in \mathcal{F}} U\,,$$

cioè la (5.19) e, quindi, l'asserto. $\square$

Osservazione 5.6.16. La proposizione precedente vale più in generale per le famiglie di sottospazi. La dimostrazione è esattamente la stessa. $\square$

Sia X un sottoinsieme di $\mathbb{V}$ e denotiamo con $\mathcal{S}_X$ l'insieme di tutti i sottospazi che contengono X, cioè

$$\mathcal{S}_X = \{U \supseteq X \;\; : \;\; U \leq \mathbb{V}\}. \tag{5.21}$$

L'insieme $\mathcal{S}_X$ è non vuoto perchè $\mathbb{V}$ è un suo elemento e, quindi, per la prop.5.6.15, l'intersezione degli elementi di $\mathcal{S}_X$ risulta un sottospazio di $\mathbb{V}$ che prende il nome di *sottospazio generato da* X e si denota con $< X >$. È chiaro che risulta

$$< X > \subseteq U\,, \text{ per ogni sottospazio } U \text{ contenente } X, \tag{5.22}$$

cioè, rispetto alla relazione di inclusione, $< X >$ è il minimo sottospazio contenente X[9]. Esso è caratterizzato dalla seguente proposizione.

[9] Questo significa che se U è un sottospazio tale che $X \subseteq U \subseteq < X >$, allora risulta $U =< X >$.

PROPOSIZIONE 5.6.17. *Un sottospazio U di $\mathbb{V}$ è il sottospazio generato da un insieme X se, e solo se, verifica le seguenti due proprietà:*

(i) *U contiene X,*

(ii) *ogni sottospazio contenente X contiene U.*

DIMOSTRAZIONE. Il sottospazio generato da X, essendo l'intersezione di tutti i sottospazi che contengono X, verifica le proprietà (i) e (ii). Se, viceversa, un sottospazio U verifica le (i) e (ii), oltre all'inclusione $< X >\subseteq U$, abbiamo anche $U \subseteq< X >$, cioè $U =< X >$. □

PROPOSIZIONE 5.6.18. *Il sottospazio generato da X coincide con la chiusura lineare di X, cioè*

$$\mathcal{L}(X) =< X > \ , \ \textit{per ogni } X \subseteq \mathbb{V}. \tag{5.23}$$

DIMOSTRAZIONE. È immediato provare che $\mathcal{L}(X)$ verifica le proprietà (i) e (ii) della proposizione precedente e da ciò segue l'asserto. □

5.7 Basi e dimensione

Siano $\mathbb{V}$ uno spazio vettoriale sul campo $\mathbb{F}$ e X un insieme di vettori. Si dice che X è un *generatore* di $\mathbb{V}$ se il sottospazio da esso generato coincide con $\mathbb{V}$, cioè se $< X >= \mathbb{V}$. Quando X è un generatore di V si dice anche che X *genera* $\mathbb{V}$. Se esiste un generatore finito (formato cioè da un numero finito di vettori) di $\mathbb{V}$ si dice che $\mathbb{V}$ è uno spazio vettoriale *finitamente generato*, o di *dimensione finita*.

ESEMPIO 5.7.1. Lo spazio vettoriale numerico $\mathbb{F}^n$ è finitamente generato. Posto, infatti,

$$\mathcal{B} = \{\boldsymbol{e}_1 = (1, 0, 0, \ldots, 0), \boldsymbol{e}_2 = (0, 1, 0, \ldots, 0), \ldots, \boldsymbol{e}_n = (0, 0, 0, \ldots, 1)\}\,,$$

risulta

$$\mathcal{L}(\mathcal{B}) =< \mathcal{B} >= \{(a_1, a_2, \ldots, a_n) \ : \ a_1, a_2, \ldots, a_n \in \mathbb{F}\} = \mathbb{F}^n.$$

□

ESEMPIO 5.7.2. Lo spazio vettoriale delle matrici $\mathbb{F}^{m,n}$ è finitamente generato. Infatti, se per $i = 1, 2, \ldots m$ e $j = 1, 2, \ldots n$, denotiamo con I_{ij} la matrice che presenta 1 nel posto (i, j) e 0 nei rimanenti e poniamo

$$\mathcal{B} = \{I_{ij} \ : \ i = 1, 2, \ldots m, \, j = 1, 2, \ldots n\},$$

risulta

$$\mathcal{L}(\mathcal{B}) =< \mathcal{B} >= \{(a_{ij}) \ : \ a_{ij} \in \mathbb{F}\} = \mathbb{F}^{m,n}.$$

□

Esempio 5.7.3. Lo spazio vettoriale $\mathbb{F}[x]$ dei polinomi in x a coefficienti in $\mathbb{F}$ non è finitamente generato. Consideriamo, infatti, i polinomi $a_1(x), a_2(x), \ldots,$ $a_h(x)$ di rispettivi gradi $n_1, n_2, \ldots, n_h$. Allora, in forza della (5.4), il grado di una qualsiasi combinazione lineare di $a_1(x), a_2(x), \ldots,$ $a_h(x)$ non può superare l'intero $n = n_1 + n_2 + \cdots + n_h$. Ne segue che un polinomio di grado maggiore di n non può essere combinazione lineare di $a_1(x), a_2(x), \ldots,$ $a_h(x)$ e, quindi, $\mathbb{F}[x]$ non ammette un insieme finito di generatori. Il sottospazio vettoriale $\mathbb{F}_n[x]$ di $\mathbb{F}[x]$ (*cfr.* esempio 5.6.7) è, invece, finitamente generato in quanto risulta $\mathbb{F}_n[x] = \mathcal{L}(1, x, x^2, \ldots, x^n)$ (*cfr.* esempio 5.3.6) , cioè

$$\mathbb{F}_n[x] =< 1, x, x^2, \ldots, x^n > \ . \tag{5.24}$$
□

Esempio 5.7.4. Il campo dei numeri complessi $\mathbb{C}$, considerato come spazio vettoriale sul campo reale $\mathbb{R}$, è finitamente generato. Risulta, infatti,

$$C =< 1, i >,$$

ove i è l'unità immaginaria ($i^2 = 1$), perchè ogni numero complesso è del tipo $a + ib$, con $a, b \in \mathbb{R}$. Si può, invece, dimostrare che $\mathbb{C}$ e $\mathbb{R}$, considerati come spazi vettoriali sul campo razionale $\mathbb{Q}$, non sono finitamente generati. □

Un insieme finito di vettori $\mathcal{B}$ si dice *base* di $\mathbb{V}$ se sono verificate le seguenti due proprietà:

1. $\mathcal{B}$ è un generatore di $\mathbb{V}$, cioè $< \mathcal{B} >= \mathbb{V}$,
2. $\mathcal{B}$ è un insieme indipendente.

Il numero di elementi di una base si chiama *ordine* della base.

Proposizione 5.7.5. *Uno spazio vettoriale $\mathbb{V}$ finitamente generato e non nullo contiene una base.*

Dimostrazione. Nelle nostre ipotesi esiste un insieme finito e non vuoto X di vettori che genera $\mathbb{V}$, cioè $< X >= \mathbb{V}$. Allora, eliminando da X, uno alla volta, i vettori che dipendono linearmenre dai rimanenti, si arriva ad un insieme indipendente che è ancora un generatore di $\mathbb{V}$, cioè una base. □

Convenzione 5.7.6. Nel seguito, tranne esplicito avviso, **supporremo sempre che $\mathbb{V}$ sia non nullo e finitamente generato**, anche se molte delle proposizioni che proveremo conservano la loro validità in un qualsiasi spazio vettoriale. □

Esempio 5.7.7. (Base canonica di $\mathbb{F}^n$) L'insieme

$$\mathcal{B} = \{\boldsymbol{e}_1 = (1, 0, 0, \ldots, 0), \boldsymbol{e}_2 = (0, 1, 0, \ldots, 0), \ldots, \boldsymbol{e}_n = (0, 0, 0, \ldots, 1)\}$$

è una base di $\mathbb{F}^n$, che si chiama *base canonica*, o *naturale*. Abbiamo già osservato che B è un generatore di $\mathbb{F}^n$ (*cfr.* esempio 5.7.1). D'altra parte, $\mathcal{B}$ è anche indipendente perchè, se $\alpha_1, \alpha_2, \ldots, \alpha_n$ sono scalari, risulta

$$\alpha_1 \boldsymbol{e}_1 + \alpha_2 \boldsymbol{e}_2 + \cdots + \alpha_n \boldsymbol{e}_n = (\alpha_1, \alpha_2, \ldots, \alpha_n) = \mathbf{0}$$
$$\Rightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0.$$
□

Esempio 5.7.8. (Base canonica di $\mathbb{F}^{m,n}$) L'insieme

$$\mathcal{B} = \{I_{ij} \ : \ i = 1, 2, \dots m, \ j = 1, 2, \dots n\},$$

ove I_{ij} é la matrice che presenta 1 nel posto (i, j) e 0 nei rimanenti, è una base di $\mathbb{F}^{m,n}$, che si chiama *base canonica,* o *naturale.* Abbiamo già osservato che B è un generatore di $\mathbb{F}^{m,n}$ (*cfr.* esempio 5.7.2). D'altra parte, $\mathcal{B}$ è anche indipendente perchè, se α_{ij} sono scalari, risulta

$$\sum_{ij} \alpha_{ij} I_{ij} = (\alpha_{ij}) = \mathbf{0} \ \Rightarrow \ \alpha_{ij} = 0\,, \text{ per ogni } i, j\,.$$

□

Esempio 5.7.9. (Base canonica di $\mathbb{F}_n[x]$) L'insieme

$$\mathcal{B} = \{1, x, x^2, \dots, x^n\}$$

è una base (*canonica*) dello spazio vettoriale $\mathbb{F}_n[x]$ (*cfr.* esempi 5.4.7, 5.6.7, 5.7.3). □

Esempio 5.7.10. (Base canonica di $\mathbb{F}_1[\boldsymbol{x}]$) L'insieme

$$\mathcal{B} = \{1, x_1, x_2, \dots, x_n\}$$

è una base (*canonica*) dello spazio vettoriale $\mathbb{F}_1[\boldsymbol{x}]$ dei polinomi lineari in $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$. □

Esempio 5.7.11. (Base canonica di $\mathbb{C}$ su $\mathbb{R}$) È facile rendersi conto che l'insieme $\{1, i\}$ è una base (detta *canonica*) del campo dei numeri complessi $\mathbb{C}$ considerato come spazio vettoriale sul campo reale $\mathbb{R}$. □

Proposizione 5.7.12. *Sia S un sistema lineare omogeneo e ridotto. Allora le soluzioni principali di S costituiscono una base del sottospazio vettoriale delle sue soluzioni.*

Dimostrazione. Supponiamo, cosa che non lede le generalità del problema, che le incognite principali siano $x_1, x_2, \dots, x_m$ e le libere $x_{m+1}, x_{m+2}, \dots, x_n$. Denotiamo con

$$\boldsymbol{c}_{m+1}, \ \boldsymbol{c}_{m+2}, \dots, \boldsymbol{c}_n$$

le soluzioni principali di S corrispondenti rispettivamente ai valori

$$(1, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, \dots, 0, 1)$$

dell'$(n-m)-$pla $(x_{m+1}, x_{m+2}, \dots, x_n)$ delle incognite libere. Allora, in forza della prop. 4.5.3, ogni soluzione di S è combinazione lineare di $\boldsymbol{c}_{m+1}, \boldsymbol{c}_{m+2}, \dots, \boldsymbol{c}_n$. D'altra parte, tali vettori sono indipendenti (*cfr.* esempio 5.4.9) e abbiamo così l'asserto. □

Proposizione 5.7.13. *Un insieme di vettori $\mathcal{B} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$ è una base di $\mathbb{V}$ se, e solo se, $\mathcal{B}$ è un insieme indipendente massimale*[10].

[10] Un insieme X di vettori è *indipendente massimale* se non è propriamente contenuto in un insieme indipendente, cioè se, per ogni vettore $\boldsymbol{a} \notin X$, l'insieme $X \cup \{\boldsymbol{a}\}$ è dipendente.

Dimostrazione. Se $\mathcal{B}$ è una base, per definizione, é indipendente e ogni vettore dipende linearmente da esso, cioè $\mathcal{B}$ è un insieme indipendente massimale. Se, viceversa, supponiamo $\mathcal{B}$ indipendente massimale, allora ogni vettore di $\mathbb{V}$ dipende linearmente da B e, quindi, $\mathcal{B}$ è una base. $\square$

Proposizione 5.7.14. *Un insieme di vettori $\mathcal{B} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$ è una base di $\mathbb{V}$ se, e solo se, $\mathcal{B}$ è un generatore minimale*[11] *di $\mathbb{V}$.*

Dimostrazione. Se $\mathcal{B}$ è una base, per definizione, risulta un generatore minimale di $\mathbb{V}$. Se, viceversa, supponiamo $\mathcal{B}$ generatore minimale, allora $\mathcal{B}$ è anche indipendente, altrimenti conterrebbe un vettore $\boldsymbol{a}$ dipendente dai rimanenti e si avrebbe $<\mathcal{B}>=<\mathcal{B} \setminus \{\boldsymbol{a}\}>$, un assurdo. Ne segue che $\mathcal{B}$ è una base. $\square$

Proposizione 5.7.15. (Lemma di Steinitz) *Siano $X = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_k\}$ un generatore di un sottospazio U e $Y = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_h\}$ un insieme indipendente contenuto in U, allora risulta $h \leq k$.*

Dimostrazione. Si osservi preliminarmente che nessun sottoinsieme proprio Z di Y può essere un generatore di U, altrimenti un vettore in $Y \setminus Z$ dipenderebbe linearmente da Z e Y sarebbe dipendente. Si tenga, inoltre, presente per tutto il corso della dimostrazione che Y, essendo indipendente, non contiene il vettore nullo. Ciò premesso, ragionando per assurdo, supponiamo $h > k$.

Poichè X genera U e $\boldsymbol{b}_1 \in U$, risulta

$$\boldsymbol{b}_1 = \alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2 + \cdots + \alpha_k\boldsymbol{a}_k\,,$$

con $\alpha_1, \alpha_2, \dots, \alpha_k$ scalari non tutti nulli e, senza perdita di generalità, possiamo supporre $\alpha_1 \neq 0$ e scrivere

$$\boldsymbol{a}_1 = \frac{1}{\alpha_1}\boldsymbol{b}_1 - \frac{\alpha_2}{\alpha_1}\boldsymbol{a}_2 - \cdots - \frac{\alpha_k}{\alpha_1}\boldsymbol{a}_k.$$

Ne segue che $X_1 = \{\boldsymbol{b}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_k\}$ è ancora un generatore di U, essendo $\boldsymbol{a}_1 \in < X_1 >$ e, quindi, $U =< X >=< X_1 >$.

Poichè X_1 genera U e $\boldsymbol{b}_2 \in U$, risulta

$$\boldsymbol{b}_2 = \beta_1\boldsymbol{b}_1 + \beta_2\boldsymbol{a}_2 + \cdots + \beta_k\boldsymbol{a}_k\,,$$

con $\beta_2, \beta_3, \dots, \beta_k$ scalari non tutti nulli, altrimenti $\boldsymbol{b}_1, \boldsymbol{b}_2$ sarebbero dipendenti, contro l'ipotesi che Y è indipendente. Senza perdita di generalità, possiamo supporre $\beta_2 \neq 0$ e scrivere

$$\boldsymbol{a}_2 = \frac{1}{\beta_2}\boldsymbol{b}_2 - \frac{\beta_1}{\beta_2}\boldsymbol{b}_1 - \frac{\beta_3}{\beta_2}\boldsymbol{a}_3 - \cdots - \frac{\beta_k}{\beta_2}\boldsymbol{a}_k.$$

[11] Un insieme X di vettori è un *generatore minimale* di $\mathbb{V}$ se è un suo generatore e se non contiene propriamente alcun altro generatore di $\mathbb{V}$, cioè se, per ogni vettore $\boldsymbol{a} \in X$, l'insieme $X \setminus \{\boldsymbol{a}\}$ non è un generatore di $\mathbb{V}$.

Ne segue che $X_2 = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{a}_3, \dots, \boldsymbol{a}_k\}$ è ancora un generatore di U, essendo $\boldsymbol{a}_2 \in < X_2 >$ e, quindi, $\mathbb{V} =< X_1 >=< X_2 >$.

A questo punto, iterando la precedente procedura e ricordando che stiamo supponendo $h > k$, arriviamo ad un insieme

$$X_k = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_k\}$$

che è un generatore di U e ciò è assurdo per l'osservazione iniziale. □

Corollario 5.7.16. *Siano $\mathcal{B} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$ una base di $\mathbb{V}$ con n elementi e $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_m$ m vettori. Allora, se è $n < m$, i vettori $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_m$ sono dipendenti.*

Dimostrazione. Se i vettori $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_m$ fossero indipendenti, dalla proposizione precedente, avremmo $m \leq n$, contro le ipotesi. □

Corollario 5.7.17. *Ogni base di $\mathbb{V}$ contiene lo stesso numero di vettori.*

Dimostrazione. Siano B e B' due basi di $\mathbb{V}$ d'ordine rispettivamente k e h. In forza della prop.5.7.15, poichè B è un generatore di $\mathbb{V}$ e B' è indipendente, abbiamo $h \leq k$. D'altra parte, invertendo i ruoli di B e B', abbiamo $k \leq h$. Ne segue che è $h = k$, cioè l'asserto. □

Il corollario precedente assicura che tutte le basi di uno spazio vettoriale finitamente generato $\mathbb{V}$ contengono lo stesso numero n di vettori. Tale intero si chiama *dimensione* di $\mathbb{V}$ e si denota con $dim\mathbb{V}$, si pone cioè

$$n = dim\mathbb{V}.$$

È chiaro che ogni sottospazio vettoriale di $\mathbb{V}$, essendo a sua volta uno spazio vettoriale finitamente generato, ha una sua dimensione. Nel seguito uno spazio vettoriale $\mathbb{V}$ di dimensione finita n sarà denotato con $\mathbb{V}_n$. Analogamente, per un sottospazio vettoriale U di $\mathbb{V}_n$ di dimensione h useremo la notazione U_h. In altre parole, un intero posto a indice della lettera che denota uno spazio o un sottospazio vettoriale sta ad indicare la sua dimensione.

Esempio 5.7.18. (Dimensioni di $\mathbb{F}^n$, $\mathbb{F}^{m,n}$, $\mathbb{F}_n[x]$, $\mathbb{F}_1[x_1, x_2, \dots, x_n]$) Per ogni campo $\mathbb{F}$, risulta

$$dim\mathbb{F}^n = n\,, \quad dim\mathbb{F}^{m,n} = mn\,,$$

$$dim\mathbb{F}_n[x] = n\,, \quad dim\mathbb{F}_1[x_1, x_2, \dots, x_n] = n + 1$$

perché le basi canoniche di $\mathbb{F}^n$, $\mathbb{F}^{m,n}$, $\mathbb{F}_n[x]$ e $\mathbb{F}_1[x_1, x_2, \dots, x_n]$ (*cfr.* esempi 5.7.7, 5.7.8, 5.7.9, 5.7.10) contengono rispettivamente n, nm, n e $n + 1$ elementi. □

Esempio 5.7.19. (Dimensione di $\mathbb{C}$ su $\mathbb{R}$) Il campo dei numeri complessi $\mathbb{C}$, considerato come spazio vettoriale sul campo reale $\mathbb{R}$, ha dimensione 2 (*cfr.* esempio 5.7.11). □

Esempio 5.7.20. (Dimensione dello spazio delle soluzioni di un sistema omogeneo ridotto) Sia S un sistema lineare in n incognite a coefficienti in un campo $\mathbb{F}$. Se S è ridotto ed ha m incognite principali, allora lo spazio vettoriale delle sue soluzioni ha dimensione $n - m$, una base essendo data dalle $n - m$ soluzioni principali (*cfr.* prop.5.7.12). □

Proposizione 5.7.21. *Sia $\mathbb{V}$ uno spazio vettoriale di dimensione n. Allora:*

- *Ogni insieme indipendente di n vettori è una base.*
- *Ogni sottospazio di $\mathbb{V}$ ha dimensione finita non superiore ad n.*
- *Un sottospazio U di $\mathbb{V}$ ha dimensione n se, e solo se, risulta $U = \mathbb{V}$.*

Dimostrazione. È una semplice conseguenza della prop.5.7.15 e si lascia per esercizio al Lettore. □

Sia X un insieme finito (risp. sistema) di vettori di uno spazio vettoriale $\mathbb{V}$ di dimensione finita. Si chiama *rango* di X, e si denota con $r(X)$, la dimensione del sottospazio vettoriale $< X >$ generato da X. Ovviamente, esiste almeno una base di $< X >$ contenuta in X e, in conseguenza delle prop.5.7.14, 5.7.17, si ha subito il seguente risultato.

Proposizione 5.7.22. *Per un insieme finito X di vettori di uno spazio vettoriale $\mathbb{V}$ sono equivalenti le seguenti affermazioni:*

- *X ha rango r;*
- *r è il massimo numero di vettori indipendenti di X;*
- *r è l'ordine di un sottoinsieme indipendente massimale di X.*
- *r è l'ordine di un generatore minimale di $< X >$.*

Uno spazio vettoriale che non sia finitamente generato si dice di *dimensione infinita*. Risulta, per esempio, di dimensione infinita lo spazio vettoriale $\mathbb{F}[x]$ dei polinomi a coefficienti in $\mathbb{F}$ (*cfr.* 5.7.3).

Osservazione 5.7.23. Le nozioni di dipendenza e indipendenza lineare in uno spazio vettoriale $\mathbb{V}$ sono state introdotte solo per insiemi e sistemi finiti X di vettori; nelle rispettive definizioni, infatti, intervengono combinazioni lineari di tutti gli elementi di X e queste hanno significato solo per un numero finito di vettori. Nel caso di insiemi arbitrari di vettori, tali nozioni si possono generalizzare nel seguente modo. Un insieme di vettori di $\mathbb{V}$ (anche infinito) si dice *libero,* se ogni suo sottoinsieme finito è linearmente indipendente. Per esempio, non è difficile rendersi conto che, nello spazio vettoriale dei polinomi $\mathbb{F}[x]$, l'insieme infinito

$$\{1, x, x^2, \dots, x^n, \dots\}$$

è libero (*cfr.* esempio 5.4.7). È chiaro che *un insieme finito è libero se, e solo se, è linearmente indipendente.* Naturalmente, ogni spazio vettoriale che contenga un insieme di vettori infinito e libero ha dimensione infinita.

Un insieme di vettori che non sia libero si dice *legato*. In altre parole, un insieme è legato se contiene un sottoinsieme finito di vettori linearmente dipendente. Anche qui si ha subito che *un insieme finito di vettori è legato se, e solo se, è linearmente dipendente*. Inoltre, ogni insieme infinito di vettori in uno spazio vettoriale di dimensione finita è legato. □

5.8 Riferimenti

In tutto il paragrafo riterremo fissato uno spazio vettoriale $\mathbb{V}_n$ di dimensione n sul campo $\mathbb{F}$.

Un sistema di n vettori distinti $\mathcal{R} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n)$ di $\mathbb{V}_n$ si chiama *base ordinata*, o *riferimento*, se l'insieme $\mathcal{B} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ è una base di $\mathbb{V}_n$ [12].

Se $\boldsymbol{a} \in \mathbb{V}$ e $\mathcal{R} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n)$ è un riferimento di $\mathbb{V}$, $\boldsymbol{a}$ dipende linearmente da $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ e, in forza della prop.5.4.16, sono univocamente determinati gli scalari $a_1, a_2, \ldots, a_n$ per cui è

$$\boldsymbol{a} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + \cdots + a_n\boldsymbol{e}_n. \tag{5.25}$$

Tali scalari prendono il nome di *componenti*, o *coordinate*, del vettore $\boldsymbol{a}$ nel riferimento $\mathcal{R}$ e il vettore numerico $(a_1, a_2, \ldots, a_n) \in \mathbb{F}^n$ si dice *vettore delle componenti*, o *delle coordinate* di $\boldsymbol{a}$ relativo a $\mathcal{R}$. La (5.25) definisce la funzione

$$\gamma_{\mathcal{R}} \ : \ \boldsymbol{a} \in \mathbb{V}_n \ \to \ (a_1, a_2, \ldots, a_n) \in \mathbb{F}^n \tag{5.26}$$

che prende il nome di *coordinazione di* $\mathbb{V}_n$ relativa al riferimento $\mathcal{R}$. Nel seguito spesso, per indicare che $a_1, a_2, \ldots, a_n$ sono le componenti del vettore $\boldsymbol{a}$ in $\mathcal{R}$, invece di scrivere $\gamma_{\mathcal{R}}(\boldsymbol{a}) = (a_1, a_2, \ldots, a_n)$, scriveremo $\boldsymbol{a} \equiv_{\mathcal{R}} (a_1, a_2, \ldots, a_n)$ o, se $\mathcal{R}$ è chiaro dal contesto, $\boldsymbol{a} \equiv (a_1, a_2, \ldots, a_n)$.

Proposizione 5.8.1. *La funzione* $\gamma_{\mathcal{R}}$ *definita dalla* (5.26) *è biunivoca.*

Dimostrazione. La funzione $\gamma_{\mathcal{R}}$ è iniettiva in forza della prop.5.4.16. D'altra parte, essa è anche suriettiva perché, per ogni $(a_1, a_2, \ldots, a_n) \in \mathbb{F}^n$, risulta

$$\gamma_{\mathcal{R}}(a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + \cdots + a_n\boldsymbol{e}_n) = (a_1, a_2, \ldots, a_n). \qquad \square$$

Osservazione 5.8.2. Una base di $\mathbb{V}_n$ si può ordinare in esattamente

$$n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1 \,^{13} \tag{5.27}$$

modi diversi. Ne segue che ogni base di $\mathbb{V}_n$ individua $n!$ riferimenti distinti.

Per esempio, la base canonica

$$\mathcal{B} = \{\boldsymbol{e}_1 = (1, 0, 0,)\,,\ \boldsymbol{e}_2 = (0, 1, 0,)\,,\ \boldsymbol{e}_3(0, 0, 1)\}$$

[12] Si noti che nella nozione di base non è importante l'ordine con cui si elencano i suoi vettori; quest'ordine è invece essenziale nella nozione di riferimento.

[13] Il prodotto dei primi n interi positivi si denota con $n!$ e si chiama *fattoriale* dell'intero n.

di $\mathbb{F}^3$ individua i 6 riferimenti

$$\mathcal{R}_1 = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)\ ,\ \ \mathcal{R}_2 = (\boldsymbol{e}_2, \boldsymbol{e}_1, \boldsymbol{e}_3)\ ,\ \ \mathcal{R}_3 = (\boldsymbol{e}_3, \boldsymbol{e}_2, \boldsymbol{e}_1)\ ,$$

$$\mathcal{R}_4 = (\boldsymbol{e}_1, \boldsymbol{e}_3, \boldsymbol{e}_2)\ ,\ \ \mathcal{R}_5 = (\boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_1)\ ,\ \ \mathcal{R}_6 = (\boldsymbol{e}_3, \boldsymbol{e}_1, \boldsymbol{e}_2)\,.$$

le componenti di un vettore $\boldsymbol{a} = (a_1, a_2, a_3)$ rispetto a tali riferimenti sono

$$\boldsymbol{a} \equiv_{\mathcal{R}_1} (a_1, a_2, a_3)\ ,\ \ \boldsymbol{a} \equiv_{\mathcal{R}_2} (a_2, a_1, a_3)\ ,\ \ \boldsymbol{a} \equiv_{\mathcal{R}_3} (a_3, a_3, a_3)\ ,$$

$$\boldsymbol{a} \equiv_{\mathcal{R}_4} (a_1, a_3, a_2)\ ,\ \ \boldsymbol{a} \equiv_{\mathcal{R}_5} (a_2, a_3, a_1)\ ,\ \ \boldsymbol{a} \equiv_{\mathcal{R}_6} (a_3, a_1, a_2)\,.$$ □

Esempio 5.8.3. (Riferimento canonico di $\mathbb{F}^n$) In $\mathbb{F}^n$, il riferimento

$$\mathcal{R} = (\boldsymbol{e}_1 = (1, 0, 0, \dots, 0), \boldsymbol{e}_2 = (0, 1, 0, \dots, 0), \dots, \boldsymbol{e}_n = (0, 0, 0, \dots, 1))$$

individuato dalla base canonica (*cfr.* esempio 5.7.7) si chiama *riferimento canonico.* In tale riferimento, le componenti di un vettore $\boldsymbol{a} = (a_1, a_2, \dots, a_n)$ sono esattamente $a_1, a_2, \dots, a_n$. □

Esempio 5.8.4. (Riferimento canonico di $\mathbb{F}^{m,n}$) In $\mathbb{F}^{m,n}$, il riferimento

$$\mathcal{R} = (I_{11}, I_{12}, .., I_{1n}, I_{21}, I_{22}, .., I_{2n}, ..., I_{m1}, I_{m2}, , .., I_{mn})$$

individuato dalla base canonica (*cfr.* esempio 5.7.8) si chiama *riferimento canonico.* In tale riferimento, il vettore delle componenti di una matrice $A = (a_{ij})$ è

$$(a_{11}, a_{12}, .., a_{1n}, a_{21}, a_{22}, .., a_{2n}, ..., a_{m1}, a_{m2}, , .., a_{mn}).$$ □

Esempio 5.8.5. (Riferimento canonico di $\mathbb{F}_n[x]$) In $\mathbb{F}_n[x]$ il riferimento

$$\mathcal{R} = (1, x, x^2, \dots, x^n)$$

individuato dalla base canonica (*cfr.* esempio 5.7.9) si chiama *riferimento canonico.* In tale riferimento, le componenti di un polinomio $a(x)$ sono esattamente $a_0, a_1, \dots, a_n$. □

Siano $\mathcal{R} = (\boldsymbol{e}_1,\, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n)$ e $\mathcal{R}' = (\boldsymbol{e}'{}_1,\, \boldsymbol{e}'_2, \dots, \boldsymbol{e}'_n)$ due riferimenti di $\mathbb{V}_n$ e sia $A = (a_{ij})$ la matrice le cui colonne sono rispettivamente le componenti dei vettori di $\boldsymbol{e}_1,\, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n$ nel riferimento $(\boldsymbol{e}'{}_1,\, \boldsymbol{e}'_2, \dots, \boldsymbol{e}'_n)$; si abbia cioè

$$\begin{cases} \boldsymbol{e}_1 = a_{11}\boldsymbol{e}'_1 + a_{21}\boldsymbol{e}'_2 + \cdots + a_{n1}\boldsymbol{e}'_n \\ \boldsymbol{e}_2 = a_{12}\boldsymbol{e}'_1 + a_{22}\boldsymbol{e}'_2 + \cdots + a_{n2}\boldsymbol{e}'_n \\ \qquad\qquad \vdots \\ \boldsymbol{e}_n = a_{1n}\boldsymbol{e}'_1 + a_{2n}\boldsymbol{e}'_2 + \cdots + a_{nn}\boldsymbol{e}'_n \end{cases} \tag{5.28}$$

o, equivalentemente,

$$\begin{pmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \\ \vdots \\ \boldsymbol{e}_n \end{pmatrix} = A^t \begin{pmatrix} \boldsymbol{e}'_1 \\ \boldsymbol{e}'_2 \\ \vdots \\ \boldsymbol{e}'_n \end{pmatrix} . \tag{5.29}$$

La matrice $A = (a_{ij})$, definita dalla (5.28), prende il nome di *matrice di passaggio dal riferimento* $\mathcal{R}$ *al riferimento* $\mathcal{R}'$. Notiamo esplicitamente che la matrice A è la trasposta di quella dei coefficienti delle relazioni (5.28).

Proposizione 5.8.6. *Siano* $\mathcal{R} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n)$ e $\mathcal{R}' = (\boldsymbol{e}'_1, \boldsymbol{e}'_2, \ldots, \boldsymbol{e}'_n)$ *due riferimenti di* $\mathbb{V}_n$ *e sia* $A = (a_{ij})$ *la matrice di passaggio da* $\mathcal{R}$ *a* $\mathcal{R}'$. *Se* $(x_1, x_2, \ldots, x_n)$ e $(x'_1, x'_2, \ldots, x'_n)$ *sono le componenti di un vettore* $\boldsymbol{x}$ *rispettivamente in* $\mathcal{R}$ e $\mathcal{R}'$, *risulta*

$$\begin{cases} x'_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ x'_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \qquad\qquad\vdots \\ x'_n = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{cases} \tag{5.30}$$

o, equivalentemente,

$$\begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} . \tag{5.31}$$

Dimostrazione. Risulta:

$$\boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{e}_i = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} a_{ji} \boldsymbol{e}'_j = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} a_{ji} x_i \right) \boldsymbol{e}'_j ,$$

da cui

$$\boldsymbol{x} = \sum_{j=1}^{n} x'_j \boldsymbol{e}'_j = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} a_{ji} x_i \right) \boldsymbol{e}'_j .$$

D'altra parte, sappiamo che $\boldsymbol{x}$ si scrive in modo unico come combinazione lineare di $\boldsymbol{e}'_1, \boldsymbol{e}'_2, \ldots, \boldsymbol{e}'_n$ e, quindi, abbiamo

$$x'_j = \sum_{i=1}^{n} a_{ji} x_i \ , \quad \text{per ogni } j = 1, 2, \ldots, n ,$$

cioè l'asserto. □

Le relazioni (5.30) e (5.31) si chiamano *formule di trasformazione delle coordinate,* o *formule di passaggio,* dal riferimento $\mathcal{R}$ al riferimento $\mathcal{R}'$. Esse, una volta nota la matrice A, permettono di calcolare le componenti di un vettore rispetto a $\mathcal{R}'$ in funzione delle componenti rispetto ad $\mathcal{R}$ dello stesso vettore.

Proposizione 5.8.7. *Se A è la matrice di passaggio da un riferimento $\mathcal{R}$ a uno $\mathcal{R}'$ di $\mathbb{V}_n$, allora A è invertibile e A^{-1} è la matrice di passaggio da $\mathcal{R}'$ a $\mathcal{R}$.*

Dimostrazione. Sia B la matrice di passaggio da $\mathcal{R}'$ a $\mathcal{R}$. Allora, se $(x_1, x_2, \ldots, x_n)$ e $(x'_1, x'_2, \ldots, x'_n)$ sono le componenti di un vettore $\boldsymbol{x}$ rispettivamente in $\mathcal{R}$ e $\mathcal{R}'$, risulta

$$\begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = AB \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix},$$

da cui ricaviamo che AB è la matrice identica, cioè $B = A^{-1}$. □

Esempio 5.8.8. In $\mathbb{F}^n$ la matrice A di passaggio da un riferimento $\mathcal{R}'$ al riferimento canonico (*cfr.* esempio 5.8.3) è quella che ha per colonne le $n-$ple che costituiscono gli elementi di $\mathcal{R}'$. Naturalmente, la matrice di passaggio dal riferimento canonico a $\mathcal{R}'$ sarà A^{-1}. Per esempio, in $\mathbb{R}^3$ la matrice di passaggio dal riferimento

$$\mathcal{R}' = ((0,1,1)\,,\,(1,0,1)\,,\,(1,1,0))$$

a quello canonico è

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

e le relative formule di trasformazione delle coordinate sono

$$\begin{cases} x'_1 = x_2 + x_3 \\ x'_2 = x_1 + x_3 \\ x'_n = x_1 + x_2 \end{cases}.$$

Di conseguenza, la matrice di passaggio dal riferimento canonico a $\mathcal{R}'$ e le relative formule di trasformazione delle coordinato sono

$$A^{-1} = \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}, \quad \begin{cases} x'_1 = -\frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 \\ x'_2 = \frac{1}{2}x_1 - \frac{1}{2}x_2 + \frac{1}{2}x_3 \\ x'_n = \frac{1}{2}x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_3 \end{cases}.$$

□

5.9 Somme e Somme dirette di sottospazi

Anche in questo paragrafo riterremo fissato uno spazio vettoriale $\mathbb{V}_n$ di dimensione n sul campo $\mathbb{F}$.

Siano U e W due sottospazi vettoriali di $\mathbb{V}_n$. Definiamo *somma* di U e W, e lo denotiamo con $U + W$, l'insieme dei vettori di $\mathbb{V}_n$ che risultano somma di un vettore di U e di uno di W, cioè

$$U + W = \{\boldsymbol{u} + \boldsymbol{w} \;:\; \boldsymbol{u} \in U\,,\, \boldsymbol{w} \in W\}\,. \tag{5.32}$$

Proposizione 5.9.1. *La somma $U + W$ di due sottospazi vettoriali di $\mathbb{V}_n$ è un sottospazio vettoriale. Tale sottospazio coincide con quello generato da $U \cup W$, cioè*

$$U + W =< U \cup W > \ . \tag{5.33}$$

Se i sottospazi U e W hanno in comune soltanto il vettore nullo, la loro somma $U + W$ si dice *somma diretta* e si denota con $U \oplus W$. In questo caso si dice anche che U e W sono *in somma diretta* o che sono *sommandi diretti*.

Proposizione 5.9.2. *La somma di due sottospazi U e W di $\mathbb{V}_n$ è diretta se, e solo se, ogni vettore $\boldsymbol{v} \in U + W$ si scrive in un unico modo come somma di un vettore di U e di uno di W.*

Proposizione 5.9.3. *Per ogni sottospazio U di $\mathbb{V}_n$ esiste un sottospazio W tale che $V_n = U \oplus W$.*

Proposizione 5.9.4. *Siano U e W sottospazi vettoriali di $\mathbb{V}_n$ con $U \cap W \neq (\mathbf{0})$ e siano*

$$\mathcal{R}_{U\cap W} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_i)$$

un riferimento di $U \cap W$,

$$\mathcal{R}_U = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_i, \boldsymbol{u}_{i+1}, \ldots, \boldsymbol{u}_h)$$

un riferimento di U contenente $\mathcal{R}_{U\cap W}$,

$$\mathcal{R}_W = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_i, \boldsymbol{w}_{i+1}, \ldots, \boldsymbol{w}_k)$$

un riferimento di U contenente $\mathcal{R}_{U\cap W}$. Allora,

$$\mathcal{R} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_i, \boldsymbol{u}_{i+1}, \ldots, \boldsymbol{u}_h, \boldsymbol{w}_{i+1}, \ldots, \boldsymbol{w}_k)$$

è un riferimento di $U + W$.

Corollario 5.9.5. (Relazione di Grassmann) *Se U e W sono sottospazi vettoriali di $\mathbb{V}_n$, risulta*

$$dim\, U + dim\, W = dim\,(U \cap W) + dim\,(U \cup W)\,. \tag{5.34}$$

5.10 Isomorfismi

Siano $\mathbb{V}$ e $\mathbb{W}$ due spazi vettoriali di dimensione finita su un campo $\mathbb{F}$. Una funzione $F : \mathbb{V} \to \mathbb{W}$ prende il nome di *isomorfismo* se sono verificate le seguenti due proprietà:

1. $F(\lambda\boldsymbol{a} + \mu\boldsymbol{b}) = \lambda F(\boldsymbol{a}) + \mu F(\boldsymbol{b})$, per ogni $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{V}$ e $\lambda, \mu \in \mathbb{F}$ (proprietà di *linearità)*;

2. F è biunivoca.

Quando esiste un isomorfismo di $\mathbb{V}$ su $\mathbb{W}$ si dice che $\mathbb{V}$ è *isomorfo* a $\mathbb{W}$ e si scrive $\mathbb{V} \sim \mathbb{W}$. Un isomorfismo di uno spazio vettoriale su se stesso si chiama anche *automorfismo*.

Osservazione 5.10.1. L'applicazione identica $i : \boldsymbol{a} \in \mathbb{V} \to \boldsymbol{a} \in \mathbb{V}$ è banalmente un isomorfismo, che si chiama *automorfismo identico* o *identità* di $\mathbb{V}$. □

Proposizione 5.10.2. *La funzione inversa di un isomorfismo è un isomorfismo.*

Dimostrazione. Sia $F : \mathbb{V} \to \mathbb{W}$ un isomorfismo e, per ogni $\boldsymbol{a}', \boldsymbol{b}' \in \mathbb{W}$ siano $\boldsymbol{a}, \boldsymbol{b}$ gli unici vettori di $\mathbb{V}$ tali che $F(\boldsymbol{a}) = \boldsymbol{a}'$ e $F(\boldsymbol{b}) = \boldsymbol{b}'$ o, equivalentemente, $\boldsymbol{a} = F^{-1}(\boldsymbol{a}')$ e $\boldsymbol{b} = F^{-1}(\boldsymbol{b}')$. Allora, la funzione $F^{-1} : \mathbb{W} \to \mathbb{V}$ è ovviamente biunivoca e risulta

$$\begin{aligned} F^{-1}(\lambda \boldsymbol{a}' + \mu \boldsymbol{b}') =& F^{-1}(\lambda F(\boldsymbol{a}) + \mu F(\boldsymbol{b})) \\ =& F^{-1}(F(\lambda \boldsymbol{a} + \mu \boldsymbol{b})) \\ =& \lambda \boldsymbol{a} + \mu \boldsymbol{b} \\ =& \lambda F^{-1}(\boldsymbol{a}') + \mu F^{-1}(\boldsymbol{b}'). \end{aligned}$$

L'asserto è così provato. □

Proposizione 5.10.3. *Siano* $\mathbb{V}, \mathbb{W}, \mathbb{U}$ *spazi vettoriali e* $F : \mathbb{V} \to \mathbb{W}$ *e* $G : \mathbb{W} \to \mathbb{U}$ *due isomorfismi. Allora la funzione composta* $G \circ F : \mathbb{V} \to \mathbb{U}$[14] *è un isomorfismo.*

Dimostrazione. La composta di due funzioni biunivoche è biunivoca e, per ogni $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{V}$ e $\lambda, \mu \in \mathbb{F}$, risulta

$$\begin{aligned} (G \circ F)(\lambda \boldsymbol{a} + \mu \boldsymbol{b}) &= G\left(F(\lambda \boldsymbol{a} + \mu \boldsymbol{b})\right) \\ &= G\left(\lambda F(\boldsymbol{a}) + \mu F(\boldsymbol{b})\right) \\ &= \lambda G\left(F(\boldsymbol{a})\right) + \mu G\left(F(\boldsymbol{b})\right) \\ &= \lambda (G \circ F)(\boldsymbol{a}) + \mu (G \circ F)(\boldsymbol{b}), \end{aligned}$$

cioè l'asserto. □

I precedenti risultati assicurano che la relazione "$\sim$" di "*essere isomorfo a*" tra spazi vettoriali è di equivalenza; cioè, se $\mathbb{U}, \mathbb{V}, \mathbb{W}$ sono spazi vettoriali, risulta

- $\mathbb{U} \sim \mathbb{U}$ ($\sim$ è riflessiva);
- $\mathbb{U} \sim \mathbb{V} \Leftrightarrow \mathbb{V} \sim \mathbb{U}$ ($\sim$ è simmetrica);
- $\mathbb{U} \sim \mathbb{V} \sim \mathbb{W} \Rightarrow \mathbb{U} \sim \mathbb{W}$ ($\sim$ è transitiva).

Due spazi vettoriali tra cui c'è un isomorfismo si dicono *isomorfi*.

[14] Ricordiamo che la funzione composta $G \circ F : \mathbb{V} \to \mathbb{U}$ è quella che ad ogni elemento $\boldsymbol{a} \in \mathbb{V}$ associa l'elemento $G(F(\boldsymbol{a})) \in \mathbb{U}$.

Proposizione 5.10.4. *Per un isomorfismo $F : \mathbb{V} \to \mathbb{W}$ tra due spazi vettoriali $\mathbb{V}$ e $\mathbb{W}$ valgono le seguenti proprietà:*

(1) *$F(\mathbf{0}_{\mathbb{V}}) = \mathbf{0}_{\mathbb{W}}$, ove $\mathbf{0}_{\mathbb{V}}$ e $\mathbf{0}_{\mathbb{W}}$ denotano rispettivamente il vettore nullo di $\mathbb{V}$ e quello di $\mathbb{W}$.*

(2) *F conserva la dipendenza e l'indipendenza lineare, nel senso che un insieme $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_h\}$ di vettori di $\mathbb{V}$ è linearmente dipendente (risp. indipentente) se, e solo se, è linearmente dipendente (risp. indipentente) l'insieme $\{F(\boldsymbol{a}_1), F(\boldsymbol{a}_2), \dots, F(\boldsymbol{a}_h)\}$.*

(3) *F trasforma basi e riferimenti di $\mathbb{V}$ rispettivamente in basi e riferimenti di $\mathbb{W}$.*

Dimostrazione.

(1) Per ogni vettore $\boldsymbol{a} \in \mathbb{V}$ risulta:

$$F(\boldsymbol{a}) = F(\boldsymbol{a} + \mathbf{0}_{\mathbb{V}}) = F(\boldsymbol{a}) + F(\mathbf{0}_{\mathbb{V}})$$

e da qui segue $F(\mathbf{0}_{\mathbb{V}}) = \mathbf{0}_{\mathbb{W}}$.

(2) Se $\lambda_1, \lambda_2, \dots, \lambda_h$ sono scalari non tutti nulli, risulta

$$\begin{aligned} &\lambda_1 \boldsymbol{a}_1 + \lambda_2 \boldsymbol{a}_2 + \cdots + \lambda_h \boldsymbol{a}_h = \mathbf{0}_{\mathbb{V}} \\ \Leftrightarrow\ &F(\lambda_1 \boldsymbol{a}_1 + \lambda_2 \boldsymbol{a}_2 + \cdots + \lambda_h \boldsymbol{a}_h) = F(\mathbf{0}_{\mathbb{V}}) \\ \Leftrightarrow\ &\lambda_1 F(\boldsymbol{a}_1) + \lambda_2 F(\boldsymbol{a}_2) + \cdots + \lambda_h F(\boldsymbol{a}_h) = \mathbf{0}_{\mathbb{W}} \end{aligned}$$

e questo significa che l'insieme $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_h\}$ è linearmente indipendente se, e solo se, lo è $\{F(\boldsymbol{a}_1), F(\boldsymbol{a}_2), \dots, F(\boldsymbol{a}_h)\}$.

(3) Questa proprietà è una immediata conseguenza della precedente. □

Dalla (3) della proposizione precedente si ha subito il seguente corollario.

Corollario 5.10.5. *Due spazi vettoriali isomorfi hanno la stessa dimensione.*

Osservazione 5.10.6. La prop. 5.10.4 e il suo corollario giustificano il fatto che spazi vettoriali isomorfi vengono considerati equivalenti da un punto di vista algebrico. Ciò si esprime anche dicendo che lo studio degli spazi vettoriali si fa a meno di isomorfismi. In altre parole, un isomorfismo fra due spazi vettoriali puó riguardarsi come un modo per identificare i vettori del primo con quelli del secondo, conservando le proprietà algebriche dei vettori stessi. □

Il risultato che segue, col suo corollario, prova che, a meno di isomorfismi, esiste un solo spazio vettoriale di dimensione n sul campo $\mathbb{F}$ e, di conseguenza, si può assumere $\mathbb{F}^n$ come modello di tali spazi.

Proposizione 5.10.7. (Teorema di rappresentazione degli spazi vettoriali di dimensione finita) *Ogni spazio vettoriale $\mathbb{V}_n$ di dimensione n sul campo $\mathbb{F}$ è isomorfo allo spazio vettoriale numerico $\mathbb{F}^n$.*

Dimostrazione. Se $\mathcal{R} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n)$ è un riferimento di $\mathbb{V}$, la coordinazione di $\mathbb{V}_n$ relativa al riferimento $\mathcal{R}$ (*cfr.* 5.26)

$$\boldsymbol{a} \in \mathbb{V}_n \ \rightarrow \ \gamma_{\mathcal{R}}(\boldsymbol{a}) \in \mathbb{F}^n$$

è un isomorfismo. Sappiamo, infatti, che $\gamma_{\mathcal{R}}$ è una funzione biunivoca (*cfr.* prop. 5.26); inoltre, se $\boldsymbol{a} \equiv (a_i)$, $\boldsymbol{b} \equiv (b_i)$ sono vettori di $\mathbb{V}$ e λ, μ scalari, risulta

$$\begin{aligned}\gamma_{\mathcal{R}}(\lambda \boldsymbol{a} + \mu \boldsymbol{b}) &= \gamma_{\mathcal{R}}(\lambda(a_i) + \mu(b_i)) \\ &= (\lambda a_i + \mu b_i) \\ &= \lambda(a_i) + \mu(b_i) \\ &= \lambda \gamma_{\mathcal{R}}(\boldsymbol{a}) + \mu \gamma_{\mathcal{R}}(\boldsymbol{b}),\end{aligned}$$

cioè l'asserto. □

Corollario 5.10.8. *Due spazi vettoriali sullo stesso campo e della stessa dimensione sono isomorfi.*

Dimostrazione. Siano $\mathcal{R}$ e $\mathcal{R}'$ due riferimenti rispettivamente degli spazi vettoriali $\mathbb{V}_n$ e $\mathbb{W}_n$ di dimensione n sul campo $\mathbb{F}$. Allora la funzione

$$\gamma_{\mathcal{R}'}^{-1} \circ \gamma_{\mathcal{R}} : \mathbb{V}_n \rightarrow \mathbb{W}_n\ [15],$$

essendo la composizione di due isomorfismi, è un isomorfismo. □

Osservazione 5.10.9. È bene mettere in evidenza che gli isomorfismi che intervengono nella prop.5.10.7 e nel corollario 5.10.8 possono realizzarsi solo dopo aver fissato dei riferimenti negli spazi vettoriali interessati. Questo fatto si esprime dicendo che tali isomorfismi sono *non canonici.* □

Chiudiamo questo paragrafo osservando che l'operazione di composizione tra funzioni induce un'operazione sull'insieme di tutti gli automorfismi di uno spazio vettoriale $\mathbb{V}$. Tale insieme si denota con $GL(\mathbb{V})$ e per esso vale il seguente risultato.

Proposizione 5.10.10. *Sia $\mathbb{V}$ uno spazio vettoriale sul campo $\mathbb{F}$. Allora l'insieme $GL(\mathbb{V})$ dei suoi automorfismi, rispetto all'operazione di composizione di funzioni, costituisce un gruppo* (il gruppo degli automorfismi di $\mathbb{V}$).

[15] Si noti che questo isomorfismo associa al vettore $\boldsymbol{a} \in \mathbb{V}_n$ di coordinate $a_1, a_2, \ldots, a_n$ in $\mathcal{R}$ l'unico vettore $\boldsymbol{a}' \in \mathbb{W}_n$ con le stesse coordinate in $\mathcal{R}'$.

Capitolo 6

Determinanti

6.1 Il gruppo simmetrico

Sia

$$\mathbb{N}_n = \{1, 2, \dots, n\}$$

l'insieme dei primi n interi positivi. Una funzione biunivoca di $\mathbb{N}_n$ su se stesso prende il nome di *permutazione* su $\mathbb{N}_n$. Se σ e τ sono permutazioni su $\mathbb{N}_n$, la funzione

$$\tau \circ \sigma \ : \ x \in \mathbb{N}_n \to \tau(\sigma(x)) \in \mathbb{N}_n \tag{6.1}$$

é ancora una permutazione e l'operazione definita dalla (6.1) induce una struttura di gruppo sull'insieme di tutte le permutazioni su $\mathbb{N}_n$. Tale gruppo si chiama *gruppo simmetrico* su n oggetti[1], o *di grado* n, e si denota con S_n. Ovviamente, l'unitá di S_n é la permutazione identica, cioé la funzione identitá su $\mathbb{N}_n$ ($m \in \mathbb{N}_n \ \to \ m \in \mathbb{N}_n$), che denoteremo col simbolo 1.

Nel seguito, date due permutazioni σ e τ, scriveremo $\sigma\tau$ invece di $\tau \circ \sigma$; porremo cioè

$$(\sigma\tau)(a) := \tau(\sigma(a)) \ , \ \text{per ogni } a \in \mathbb{N}_n \, .$$

A volte può essere comodo rappresentare una permutazione $\sigma \in S_n$ mediante la tabella

$$\begin{bmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{bmatrix}, \tag{6.2}$$

che ha nella prima riga gli interi da 1 ad n disposti in ordine crescente e sotto ognuno di essi l'intero corrispondente in σ. Ovviamente, con questa notazione, la matrice

$$\begin{bmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{bmatrix}$$

[1]Abbiamo definito il gruppo S_n a partire da $\mathbb{N}_n$, ma avremmo potuto riferirci ad un qualunque insieme finito X con n elementi. Una volta numerati da 1 a n gli elementi di X, infatti, è facile convincersi che le proprietà dei singoli elementi di X non giocano alcun ruolo particolare; tutto dipende solo dall'intero n.

rappresenta la permutazione identica, che è l'elemento neutro del gruppo S_n. A volte, invece della (6.2), può risultare utile la notazione

$$\begin{bmatrix} j_1 & j_2 & \dots & j_n \\ \sigma(j_1) & \sigma(j_2) & \dots & \sigma(j_n) \end{bmatrix},$$

ove la prima riga contiene gli interi $1, 2, \dots, n$ scritti in un ordine diverso da quello naturale. Con questa notazione, per esempio, la permutazione inversa di σ può scriversi

$$\sigma^{-1} = \begin{bmatrix} \sigma(1) & \sigma(2) & \dots & \sigma(n) \\ 1 & 2 & \dots & n \end{bmatrix}. \tag{6.3}$$

Esempio 6.1.1. Le permutazioni σ e τ di $\mathbb{N}_4 = \{1, 2, 3, 4\}$ definite da

$$\sigma(1) = 2, \sigma(2) = 3, \sigma(3) = 4, \sigma(4) = 1 \; ; \; \tau(1) = 3, \tau(2) = 4, \tau(3) = 1, \tau(4) = 2$$

si scrivono

$$\sigma = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{bmatrix} \quad , \quad \tau = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{bmatrix}.$$

In questo modo se, per esempio, vogliamo calcolare l'immagine di 2 nel prodotto $\sigma\tau$, basta leggere l'intero sotto 2 nella tabella di σ, nel nostro caso 3, e poi l'intero sotto 3 nella tabella di τ; abbiamo cosí $\sigma\tau(2) = \tau(\sigma(2)) = 1$. Osserviamo che per σ e τ possiamo anche scrivere

$$\sigma = \begin{bmatrix} 2 & 4 & 3 & 1 \\ 3 & 1 & 4 & 2 \end{bmatrix} \quad , \quad \tau = \begin{bmatrix} 4 & 3 & 2 & 1 \\ 2 & 1 & 4 & 3 \end{bmatrix}$$

e che risulta

$$\sigma^{-1} = \begin{bmatrix} 2 & 3 & 4 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{bmatrix},$$

$$\tau^{-1} = \begin{bmatrix} 3 & 4 & 1 & 2 \\ 1 & 2 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{bmatrix}$$

□

Esempio 6.1.2. Il gruppo simmetrico S_1 contiene un unico elemento: la permutazione identica.
Il gruppo simmetrico S_2 contiene due elementi:

$$\tau_1 = 1 = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \quad , \quad \tau_2 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}.$$

Si riportano di seguito tutti gli elementi del gruppo simmetrico S_3 :

$$\sigma_1 = 1 = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix} \quad , \quad \sigma_2 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} \quad , \quad \sigma_3 = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix} \quad ,$$

$$\sigma_4 = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix} \quad , \quad \sigma_5 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix} \quad , \quad \sigma_6 = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \quad .$$

□

Se m, n sono interi positivi con $m < n$, ad ogni $\sigma \in S_m$ possiamo associare la permutazione $\sigma' \in S_n$ che opera come σ sugli interi $1, 2, \ldots, m$ e trasforma in se stesso ogni altro elemento di $\mathbb{N}_n$. L'applicazione

$$f\,:\,\sigma \in S_m \to \sigma' \in S_n$$

è iniettiva e risulta

$$f(\sigma\tau) = f(\sigma)f(\tau)\,.$$

Per questo motivo, possiamo identificare S_m con $f(S_m)$ e, con abuso di notazione, denotiamo ancora con σ la permutazione $\sigma' = f(\sigma)$. Per esempio, con questa convenzione, la permutazione

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

considerata come elemento di S_6, è la permutazione

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 4 & 1 & 5 & 6 \end{bmatrix}.$$

Osserviamo ancora che, se σ e τ sono le due permutazioni su $\mathbb{N}_3$ definite da

$$\sigma = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}\ ,\ \tau = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix},$$

risulta

$$\sigma\tau = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix} \neq \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix} = \tau\sigma$$

e si ha cosí la seguente proposizione.

Proposizione 6.1.3. *Il gruppo simmetrico S_n non è abeliano per ogni intero $n > 2$.*

Calcolare l'ordine di S_n è un facile esercizio. Abbiamo infatti che, se si vuole costruire una permutazione σ su $\mathbb{N}_n$, vi sono n possibili scelte per $\sigma(1)$, $n-1$ per $\sigma(2)$, $n-3$ per $\sigma(3)$ e cosí via. Ne segue, per induzione, che

$$|S_n| = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1 = n!\,^2. \tag{6.4}$$

Diciamo che un elemento j di $\mathbb{N}_n$ è *unito* o *fisso* in una permutazione σ se risulta $\sigma(j) = j$. Una permutazione su n oggetti prende il nome di *trasposizione* se scambia fra di loro due elementi distinti di $\mathbb{N}_n$ e lascia fissi i rimanenti. La trasposizione che scambia fra loro gli interi s, t si denota con (s, t). Per esempio

$$(1,2) = (2,1) = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 3 & 4 & 5 \end{bmatrix}\ ,\ (2,5) = (5,2) = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 5 & 3 & 4 & 2 \end{bmatrix}$$

[2] Il prodotto dei primi n interi positivi si denota con $n!$ e si chiama *fattoriale* dell'intero n.

sono due trasposizioni in S_5. È chiaro che, per ogni trasposizione σ, risulta $\sigma = \sigma^{-1}$ o, equivalentemente, $\sigma^2 = 1$. Vale il seguente teorema di cui omettiamo la dimostrazione.

Proposizione 6.1.4. *Ogni permutazione su n oggetti è prodotto di trasposizioni.*

È da notare che la decomposizione in prodotto di trasposizioni di una permutazione non è unica. Allo scopo di trovare un invariante delle possibili fattorizzazioni di una permutazione in trasposizioni diamo le seguenti definizioni per una permutazione σ :

• Considerata una coppia (i, j), ove i, j sono elementi di N_n con $i < j$, si dice che σ presenta un'*inversione* su (i, j) se risulta $\sigma(i) > \sigma(j)$.

• Si dice che σ è una permutazione *pari* se presenta un numero pari di inversioni, *dispari* nel caso contrario.

• si definisce *segno* di σ, e si denota con $sgn(\sigma)$, l'intero 1 o -1 a seconda che σ sia rispettivamente pari o dispari.

Esempio 6.1.5. La permutazione

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{bmatrix}$$

presenta esattamente tre inversioni sulle coppie $(1, 4)$, $(2, 4)$, e $(3, 4)$. Essa è pertanto una permutazione dispari. □

Esempio 6.1.6. Si può provare che ogni trasposizione è una permutazione dispari e il suo segno è -1. La permutazione identica è ovviamente pari. □

Proposizione 6.1.7. *Una permutazione e la sua inversa hanno lo stesso segno.*

Dimostrazione. Dall'analisi delle tabelle (6.5) e (6.3) si ha subito che una permutazione σ e la sua inversa σ^{-1} presentano lo stesso numero di inversioni. □

Proposizione 6.1.8. *Il prodotto di due permutazioni dello stesso segno è pari e quello di segno diverso è dispari.*

Dimostrazione. Per ogni permutazione σ risulta

$$\prod_{\{i,j\} \subseteq \mathbb{N}_n} \frac{i-j}{\sigma(i)-\sigma(j)} = \prod_{i<j} \frac{i-j}{\sigma(i)-\sigma(j)} = sgn(\sigma)$$

e, per ogni due permutazioni σ e τ, risulta

$$sgn(\sigma\tau) = \prod_{i<j} \frac{i-j}{(\sigma\tau)(i)-(\sigma\tau)(j)} =$$

$$\left(\prod_{i<j}\frac{i-j}{\sigma(i)-\sigma(j)}\right)\left(\prod_{i<j}\frac{\sigma(i)-\sigma(j)}{\tau(\sigma(i))-\tau(\sigma(j))}\right)=$$

$$\left(\prod_{i<j}\frac{i-j}{\sigma(i)-\sigma(j)}\right)\left(\prod_{\sigma(i)<\sigma(j)}\frac{\sigma(i)-\sigma(j)}{\tau(\sigma(i))-\tau(\sigma(j))}\right)=$$

$$\left(\prod_{i<j}\frac{i-j}{\sigma(i)-\sigma(j)}\right)\left(\prod_{i<j}\frac{i-j}{\tau(i)-\tau(j)}\right)=sgn(\sigma)sgn(\tau).$$

L'asserto é, così. provato. □

Corollario 6.1.9. *Se $\sigma_1\sigma_2\cdots\sigma_s$ e $\tau_1\tau_2\cdots\tau_t$ sono due fattorizzazioni in trasposizioni di una stessa permutazione, allora gli interi s e t hanno la stessa parità, cioè sono entrambi pari o entrambi dispari.*

Dimostrazione. Segue dalle proposizioni 6.1.4 e 6.1.7, tenendo presente che ogni trasposizione è una permutazione dispari.

6.2 Definizione di determinante

Sia $A=(a_{ij})\in\mathbb{F}^{n,n}$ una matrice quadrata d'ordine n a coefficienti in un campo $\mathbb{F}$.

Si chiama *determinante* di A, e si denota con $\det A$, lo scalare definito da

$$\det A=\sum_{\sigma\in S_n} sgn(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}\,. \tag{6.5}$$

La funzione

$$\det\ :\ A\in\mathbb{F}^{n,n}\ \to\ \det A\in\mathbb{F} \tag{6.6}$$

prende il nome di *funzione determinante* o anche semplicemente *determinante*.

Osservazione 6.2.1. Nella definizione di determinante, ciascun addendo della somma (6.5) corrisponde biunivocamente ad un prodotto di n elementi di A a due a due non appartenenti ad una stessa linea; gli indici di questi elementi individuano una permutazione σ e il prodotto in questione è moltiplicato per 1 o -1 a seconda che σ sia una permutazione pari o dispari. □

Il determinante di A si denota anche con $|A|$ o, se si vogliono evidenziare gli elemeti di A, con

$$\det\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix}\quad\text{o}\quad\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{vmatrix}.$$

Con abuso di linguaggio, il determinante di una matrice A d'ordine n si chiama anche *determinante d'ordine n* e le linee di A si dicono anche *linee del determinante*. Più in generale, quando non vi è luogo ad equivoci, si estendono ai determinanti le definizioni relative alle matrici quadrate. Per esempio, dire che un determinante $\det A$ è triangolare equivale a dire che la matrice A è triangolare; dire, invece, che un determinante $\det A$ è nullo significa che lo scalare $\det A$ è zero e ciò, come avremo modo di vedere, non implica che la matrice A sia nulla.

Osserviamo che, nel caso $n = 1$, il gruppo simmetrico S_1 contiene un solo elemento (l'*identità*, che è una permutazione pari), la matrice $A = (a)$ si può identificare con l'elemento $a \in \mathbb{F}$ e, quindi, risulta $\det A = \det(a) = a$. Esaminiamo, nei due esempi che seguono, come si presenta l'espressione del determinante di una matrice nei casi particolari $n = 2, 3$.

Esempio 6.2.2. (Determinante d'ordine 2) Nel caso $n = 2$, abbiamo

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

e il gruppo simmetrico S_2 consta dei due elementi

$$\sigma_1 = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \quad , \quad \sigma_2 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix},$$

per i quali risulta $sgn(\sigma_1) = 1$ e $sgn(\sigma_2) = -1$. Ne segue che

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Per esempio, risulta

$$\begin{vmatrix} 5 & 1 \\ 1 & 2 \end{vmatrix} = 5 \cdot 2 - 1 \cdot 1 = 10 - 1 = 9$$

e

$$\begin{vmatrix} 5 & 5 \\ 1 & 1 \end{vmatrix} = 5 \cdot 1 - 5 \cdot 1 = 5 - 5 = 0.$$

L'ultimo esempio mostra che il determinante di una matrice non nulla può essere nullo. □

Esempio 6.2.3. (Determinante d'ordine 3) Nel caso $n = 3$, abbiamo

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

e il gruppo simmetrico S_6 consta dei 6 elementi

$$\sigma_1 = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix} \quad , \quad \sigma_2 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} \quad , \quad \sigma_3 = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix} \quad ,$$

$$\sigma_4 = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix}, \quad \sigma_5 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix}, \quad \sigma_6 = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix},$$

per i quali risulta

$$sgn(\sigma_1) = sgn(\sigma_2) = sgn(\sigma_3) = 1 \text{ e } sgn(\sigma_4) = sgn(\sigma_5) = sgn(\sigma_6) = -1.$$

Ne segue che

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

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33}.$$

Per esempio,

$$\begin{vmatrix} 1 & 3 & 1 \\ 2 & 1 & 2 \\ 4 & 2 & 1 \end{vmatrix}$$

$$\begin{aligned} &= 1 \cdot 1 \cdot 1 + 3 \cdot 2 \cdot 4 + 1 \cdot 2 \cdot 2 - 1 \cdot 2 \cdot 2 - 3 \cdot 2 \cdot 1 - 1 \cdot 1 \cdot 4 \\ &= 1 + 24 + 4 - 4 - 6 - 4 = 15 \end{aligned}$$

e

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

$$\begin{aligned} &= 1 \cdot 2 \cdot 3 + 3 \cdot 2 \cdot 0 + 1 \cdot 0 \cdot 0 - 1 \cdot 2 \cdot 0 - 3 \cdot 0 \cdot 3 - 1 \cdot 2 \cdot 0 \\ &= 6 + 0 + 0 - 0 - 0 - 0 = 6. \end{aligned}$$

Si osservi che l'ultimo determinante è triangolare ed è uguale al prodotto degli elementi della diagonale principale. □

6.3 Prime proprietà dei determinanti

Come il Lettore avrà senz'altro notato, calcolare un determinante d'ordine n applicando la definizione non è un'operazione facile da un punto di vista computazionale: bisogna calcolare $n!$ prodotti di n elementi e poi sommarli[3]. I risultati che seguono mettono in luce alcune proprietà dei determinanti che ci permetteranno di trovare metodi di calcolo più semplici. Vedremo che, anche in questo caso, l'algoritmo di Gauss 3.7.6 giocherà un ruolo importante.

Proposizione 6.3.1. *Il determinante di una matrice triangolare $T = (t_{ij})$, è uguale al prodotto degli elementi della diagonale principale, cioè*

$$\det T = t_{11}t_{12} \cdots t_{1n}.$$

Risulta, inoltre, che il determinante di T é non nullo se, e solo se, tutti gli elementi della diagonale principale di T sono non nulli.

[3]Si tenga presente che $n!$ cresce molto rapidamente al crescere di n. Per esempio, $10! = 3628800$ è dell'ordine di 10^7 e $20! = 1216451004088320000$ è dell'ordine di 10^{18}.

Dimostrazione. Un prodotto di n elementi di T a due a due non appartenenti ad una stessa linea e non appartenenti tutti alla diagonale principale contiene necessariamente un fattore nullo e, quindi, é uguale a zero. Allora, dalla definizione di determinante e dall'osservazione 6.2.1, segue subito l'asserto. □

Corollario 6.3.2. *Per una matrice diagonale* $D = (d_{ij})$, *risulta*

$$\det D = d_{11} d_{12} \cdots d_{1n} \,.$$

In particolare, per la matrice identica I *è*

$$\det I = 1 \,.$$

Dimostrazione. Le matrici diagonali sono particolari matrici triangolari. L'asserto é, dunque, una banale applicazione della prop.6.3.1. □

Proposizione 6.3.3. *La matrice* A *e la sua trasposta* A^t *hanno lo stesso determinante, cioè*

$$\det A^t = \det A. \tag{6.7}$$

Dimostrazione. Ricordiamo che $sgn(\sigma) = sgn(\sigma^{-1})$ per ogni permutazione σ (*cfr.* prop.6.1.7). Allora, poichè è $A^t = (a_{ji})$, dalla (6.5) e dall'osservazione 6.2.1, risulta

$$\begin{aligned} \det A^t &= \sum_{\sigma \in S_n} sgn(\sigma^{-1}) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} \\ &= \sum_{\sigma^{-1} \in S_n} sgn(\sigma^{-1}) a_{1\sigma^{-1}(1)} a_{2\sigma^{-1}(2)} \cdots a_{n\sigma^{-1}(n)} \\ &= \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \\ &= \det A. \end{aligned}$$

□

Osservazione 6.3.4. La proposizione precedente prova che ogni proprietà vera per le righe di un determinante è vera anche per le sue colonne, e viceversa. Per questo motivo nel seguito, quando esporremo qualche proprietà solo per le righe di un determinante, sottointenderemo che queste valgono anche per le colonne, senza notarlo esplicitamente di volta in volta. □

Proposizione 6.3.5. *Il determinante di una matrice* $A \in F^{n,n}$ *verifica le seguenti proprieà:*

(D1) *Se* A *contiene una linea nulla, allora* $\det A = 0$.

(D2) *Se* A' *è una matrice ottenuta scambiando due righe di* A, *allora* $\det A' = -\det A$.

(D3) *Se* A *contiene due righe uguali, allora* $\det A = 0$.

(D4) *Se A' è una matrice ottenuta moltiplicando una linea di A per uno scalare λ, allora risulta* $\det A' = \lambda \det A$.

(D5) *Se due righe di A sono proporzionali, allora* $\det A = 0$.

(D6) *Se l'i-esima riga di A è somma dei vettori $\boldsymbol{b} = (b_1, b_2, \ldots, b_n)$ e $\boldsymbol{c} = (c_1, c_2, \ldots c_n)$, cioè $a_{ij} = b_j + c_j$ per ogni $j = 1, 2, \ldots, n$, allora risulta*

$$\det A = \det B + \det C,$$

ove B, C sono le matrici ottenute da A sostituendo l'i-esima riga rispettivamente con i vettori $\boldsymbol{b}, \boldsymbol{c}$.

(D7) *Se una riga di A è combinazione lineare di altre righe, allora* $\det A = 0$.

(D8) *Se A' è una matrice ottenuta sommando ad una riga $\boldsymbol{a}_i$ di A un'altra riga $\boldsymbol{a}_j$, con $i \neq j$, allora* $\det A' = \det A$.

Dimostrazione.

(D1) Se A possiede una linea nulla, un prodotto di n elementi di A a due a due non appartenenti ad una stessa linea contiene necessariamente un fattore di questa linea nulla e, quindi, é uguale a zero. Allora, dalla definizione di determinante e dall'osservazione 6.2.1, segue subito che $\det A = 0$.

(D2) Se A' si ottiene da A scambiando le righe di posto s e t, con $s < t$, i determinanti di A e di A', a meno del segno, presentano gli stessi termini, essendo

$$a_{1j_1} a_{2j_2} \cdot \cdot a_{sj_s} \cdot \cdot a_{tj_t} \cdot \cdot a_{nj_n} = a_{1j_1} a_{2j_2} \cdot \cdot a_{tj_t} \cdot \cdot a_{sj_s} \cdot \cdot a_{nj_n}.$$

Allora la (D2) segue dal fatto che risulta

$$(s,t) \begin{bmatrix} 1 & 2 & .. & s & .. & t & .. & n \\ j_1 & j_2 & .. & j_s & .. & j_t & .. & j_n \end{bmatrix} = \begin{bmatrix} 1 & 2 & .. & s & .. & t & .. & n \\ j_1 & j_2 & .. & j_t & .. & j_s & .. & j_n \end{bmatrix}$$

e, in forza dell'esempio 6.1.6 e della prop. 6.1.8,

$$\begin{aligned} & sgn \begin{bmatrix} 1 & 2 & .. & s & .. & t & .. & n \\ j_1 & j_2 & .. & j_t & .. & j_s & .. & j_n \end{bmatrix} \\ & = -sgn \begin{bmatrix} 1 & 2 & .. & s & .. & t & .. & n \\ j_1 & j_2 & .. & j_s & .. & j_t & .. & j_n \end{bmatrix}. \end{aligned}$$

(D3) Lo scambio di due righe uguali in A produce una matrice uguale ad A. Allora, in forza della D2, abbiamo $detA = -detA$, cioè $detA = 0$.

(D4) Risulta:

$$\begin{aligned} detA' &= \sum_{\sigma \in S_n} sgn(\sigma) \lambda \, a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \\ &= \lambda \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \\ &= \lambda \, detA. \end{aligned}$$

(D5) Supponiamo che in A la riga di posto s si ottenga moltiplicando per uno scalare $\lambda \neq 0$ la riga di posto t, con $t \neq s$. Allora, la matrice A' ottenuta da A moltiplicando per λ^{-1} la riga di posto s possiede due righe uguali e, per la D3, ha il denerminante nullo. Dalla D4, segue allora che

$$0 = detA' = \lambda^{-1}\, detA\,,$$

cioè $detA = 0$.

(D6) Risulta:

$$\begin{aligned} detA &= \sum_{\sigma \in S_n} sgn(\sigma)\, a_{1\sigma(1)}a_{2\sigma(2)} \cdot \cdot a_{i\sigma(i)} \cdot \cdot a_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} sgn(\sigma)\, a_{1\sigma(1)}a_{2\sigma(2)} \cdot \cdot (b_{i\sigma(i)} + c_{i\sigma(i)}) \cdot \cdot a_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} sgn(\sigma)\, a_{1\sigma(1)}a_{2\sigma(2)} \cdot \cdot b_{i\sigma(i)} \cdot \cdot a_{n\sigma(n)} \\ &\quad + \sum_{\sigma \in S_n} sgn(\sigma)\, a_{1\sigma(1)}a_{2\sigma(2)} \cdot \cdot c_{i\sigma(i)} \cdot \cdot a_{n\sigma(n)} \\ &= detB + detC\,. \end{aligned}$$

(D7) Segue dalle proprietà D5 e D6.

(D8) Segue dalle proprietà D3 e D6. □

Esempio 6.3.6. Facciamo vedere che il determinante di una matrice antisimmetrica $A \in \mathbb{F}^{n,n}$, con n dispari, è nullo. Poiché A è del tipo

$$A = \begin{pmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ -a_{12} & 0 & a_{23} & \cdots & a_{2n} \\ -a_{13} & -a_{23} & 0 & \cdots & a_{3n} \\ & \vdots & & & \\ -a_{1n} & -a_{2n} & -a_{3n} & \cdots & 0 \end{pmatrix},$$

moltiplicando per -1 tulle le righe di A si ottiene la matrice A^t trasposta di A. Allora, in forza della (6.7) e della proprietà D4, abbiamo

$$|A| = |A^t| = (-1)^n |A| = -|A|\,,$$

da cui ricaviamo $|A| = 0$. □

Tenendo presente che il determinante della matrice identità è uguale ad 1, le proprietà D2, D4, D8 e la prop.3.5.3, si hanno subito i seguenti corollari della prop.6.3.5.

Corollario 6.3.7. *Per il determinante delle matrici elementari si ha:*

$$\det E_{ij} = -1\,, \quad \det E_i(\lambda) = \lambda\,, \quad \det E_{ij}(\lambda) = 1\,.$$

COROLLARIO 6.3.8. *Se $A \in \mathbb{F}^{n,n}$, risulta*

$$\det(E_{ij}A) = -\det A\,, \quad \det(E_i(\lambda)A) = \lambda \det A\,, \quad \det(E_{ij}(\lambda)A) = \det A\,.$$

Ne segue che, se $E, E_1, E_2, \dots E_h$ sono matrici elementari, si ha

$$\begin{cases} \det(EA) = \det E \det A\,, \\ \\ \det(E_1 E_2 \cdots E_h) = \det E_1 \det E_2 \cdots \det E_h\,. \end{cases} \tag{6.8}$$

I due corollari precedenti assicurano che, se $A = (a_{ij}) \in \mathbb{F}^{n,n}$ e A' è una matrice ridotta equivalente ad A tale che

$$A' = E_{i_1 j_1} \cdots E_{i_h j_h} E_{i_1}(\lambda_1) \cdots E_{i_k}(\lambda_k) E_{i_1 j_1}(\mu_1) \dots E_{i_t j_j}(\mu_t) A\,,$$

allora risulta

$$|A'| = (-1)^h \lambda_1 \cdots \lambda_k |A|.$$

Questa osservazione giustifica il seguente algoritmo per il calcolo di un determinante.

ALGORITMO 6.3.9. (CALCOLO DEL DETERMINANTE COL METODO DI GAUSS)

- INPUT: Una matrice $A \in \mathbb{F}^{n,n}$.

- PASSO 1: Calcolare con l'algoritmo di Gauss una matrice ridotta A' equivalente ad A e sia

  $$A' = E_{i_1 j_1} \cdots E_{i_h j_h} E_{i_1}(\lambda_1) \cdots E_{i_k}(\lambda_k) E_{i_1 j_1}(\mu_1) \dots E_{i_t j_j}(\mu_t) A$$

  con A' del tipo

  $$A' = \begin{pmatrix} a_1 & & M \\ & \ddots & \\ 0 & & a_n \end{pmatrix}.$$

- OUTPUT: $|A| = (-1)^h \lambda_1^{-1} \cdots \lambda_k^{-1} a_1 a_2 \cdots a_n$. □

OSSERVAZIONE 6.3.10. La validità dell'algoritmo 6.3.9 prova che la complessità computazionale del calcolo di un determinante d'ordine n è la stessa dell'algoritmo di Gauss 3.7.6 per portare in forma ridotta una matrice quadrata d'ordine n. □

ESEMPIO 6.3.11. Calcoliamo il determinante della matrice

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

col metodo di Gauss. A tale scopo, riduciamo A usando l'algoritmo di Gauss

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & -1 & 3 \\ 4 & 2 & 2 \end{pmatrix} \xrightarrow{R_2 \to -2R_1 + R_2} \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 1 \\ 4 & 2 & 2 \end{pmatrix}$$

$$\xrightarrow{R_3 \to -4R_1 + R_3} \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 1 \\ 0 & 2 & -2 \end{pmatrix}$$

$$\xrightarrow{R_2 \to -R_2} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 2 & -2 \end{pmatrix}$$

$$\xrightarrow{R_3 \to -2R_2 + R_3} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} = A'$$

Poiché A' presenta uno zero sulla diagonale principale, risulta $|A| = 0$.

Proposizione 6.3.12. *Il determinante di una matrice $A \in F^{n,n}$ verifica le seguenti proprieà:*

(1) $\det A = 0 \Leftrightarrow A$ *è equivalente ad una matrice completamente ridotta con almeno una riga nulla.*

(2) $\det A \neq 0 \Leftrightarrow A$ *è invertibile.*

Dimostrazione. Se J è una matrice completamente ridotta equivalente ad A, possiamo scrivere

$$A = E_1 E_2 \cdots E_h J\,,$$

ove ciascuna E_i è una matrice elementare e, in forza delle (6.8), risulta

$$\det(A) = \det(E_1 E_2 \cdots E_h J) = \det(E_1 E_2 \cdots E_h) \det J, \tag{6.9}$$

per cui, essendo $\det(E_1 E_2 \cdots E_h) \neq 0$, abbiamo

$$\det A = 0 \;\Leftrightarrow\; detJ = 0.$$

Osserviamo che, poichè A è quadrata, J è di forma diagonale e, quindi, l'annullarsi del determinante di J equivale all'esistenza di uno zero sulla diagonale principale (*cfr.* corollario 6.3.2), cioè all'esistenza di una riga nulla.

D'altra parte, l'invertibilità di A equivale a quella di J e, essendo quest'ultima di forma diagonale, la sua invertibilità equivale all'avere tutti gli elementi della diagonale principale diversi da zero, cioè $detJ \neq 0$. Allora, dalla (6.9), abbiamo

$$A \text{ invertibile } \;\Leftrightarrow\; detJ \neq 0 \;\Leftrightarrow\; detA \neq 0$$

e l'asserto è completamente provato. □

Le matrici quadrate con determinante uguale a zero si dicono *singolari*.

PROPOSIZIONE 6.3.13. (TEOREMA DI BINET) *Se A è B sono matrici quadrate dello stesso ordine, risulta*

$$\det(AB) = \det A \det B\,. \tag{6.10}$$

DIMOSTRAZIONE. Se A è invertibile possiamo scrivere $A = E_1 E_2 \cdots E_h$, ove ciascuna E_i è una matrice elementare (*cfr.* prop.3.9.5) e, in forza delle (6.8), abbiamo

$$\det(AB) = \det(E_1 E_2 \cdots E_h B) = \det(E_1 E_2 \cdots E_h) \det B = \det A \det B\,.$$

Se A non è invertibile, per la prop.6.3.12, risulta $detA = 0$ e dobbiamo soltanto provare che $\det(AB) = 0$. In queste ipotesi possiamo scrivere

$$A = E_1 E_2 \cdots E_h J,$$

ove ciascuna E_i è una matrice elementare e J è completamente ridotta e possiede una riga nulla (*cfr.* prop.3.8.5). Ora, osservato che $\det(JB) = 0$ perchè la matrice JB ha almeno una riga nulla, in forza delle (6.8) abbiamo

$$\det(AB) = \det(E_1 E_2 \cdots E_h JB) = \det(E_1 E_2 \cdots E_h) \det(JB) = 0$$

e l'asserto è completamente provato. □

6.4 Minori e cofattori

Sia $A = (a_{ij}) \in \mathbb{F}^{m,n}$ una matrice d'ordine $m \times n$ a coefficienti in un campo $\mathbb{F}$. Definiamo *minore d'ordine* k di A, o estratto da A, il determinante M di una qualsiasi sottomatrice quadrata d'ordine k di A.

ESEMPIO 6.4.1. Si consideri la matrice quadrata d'ordine 3 sul campo reale

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

I seguenti determinanti sono minori della matrice A.

$$\begin{vmatrix} -1 & 0 & 1 \\ 2 & 1 & 0 \\ 0 & -2 & 0 \end{vmatrix} = -4\,, \quad \begin{vmatrix} 1 & -1 & 1 \\ -1 & 2 & 0 \\ 0 & 0 & 0 \end{vmatrix} = 0\,, \quad \begin{vmatrix} 1 & -1 \\ -1 & 2 \end{vmatrix} = 1\,.$$

□

Il massimo ordine di un minore non nullo della matrice A si chiama *caratteristica* di A e si denota con $c(A)$, o con c se A é chiara dal contesto. L'intero c è, dunque, la caratteristica di A se sono verificate le seguenti due proprietà:

- A possiede un minore non nullo d'ordine c;
- tutti i minori di A d'ordine $c + 1$ sono nulli.

È chiaro dalla definizione che una matrice e la sua trasposta hanno la stessa caratteristica, cioè

$$c(A) = c(A^t)\,, \tag{6.11}$$

per ogni $A \in \mathbb{F}^{m,n}$. Un minore di A d'ordine massimo, cioè d'ordine c, si chiama anche *minore principale* di A.

Esempio 6.4.2. La matrice di tipo 3×4 dell'esempio precedente ha caratteristica 3 perché 3 è il massimo ordine dei suoi minori e possiede un minore non nullo d'ordine 3. □

Nell'ipotesi che A sia quadrata d'ordine n, cancellando da A le righe e le colonne che definiscono un suo minore M d'ordine k, si ottiene una sottomatrice di A il cui determinante è un minore d'ordine $n-k$, che si chiama *minore complementare di* M e si denota con M^c. Ovviamente M risulta il minore complementare di M^c, cioè $(M^c)^c = M$, e si dice che la coppia (M, M^c) è una *coppia di minori complementari.* Se le righe e le colonne di A che definiscono un minore M sono quelle di posti rispettivamente $i_1, i_2, \dots, i_k$, e $j_1, j_2, \dots, j_k$, poniamo

$$s_M = i_1 + i_2 + \cdots + i_k + j_1 + j_2 + \cdots + j_k,$$

e lo scalare

$$(-1)^{s_M} M^c$$

prende il nome di *cofattore* di M. In altre parole, il cofattore del minore M è uguale a M^c o a $-M^c$, a seconda che s_M sia pari o dispari.

In particolare, un minore di ordine 1 non è altro che un elemento a_{ij} di A e il suo minore complementare, che ha ordine $n-1$, si denota con $M_{ij}(A)$ o più semplicemente con M_{ij} quando A è chiara dal contesto. Il cofattore di a_{ij}, che si chiama anche *complemento algebrico* di a_{ij}, si denota con A_{ij} e per esso si ha

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

Esempio 6.4.3. Per la matrice quadrata d'ordine 3 sul campo reale

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

si ha

$$A_{11} = \begin{vmatrix} 1 & 0 \\ -2 & 1 \end{vmatrix} = 1\,, \quad A_{12} = -\begin{vmatrix} 2 & 0 \\ 0 & 1 \end{vmatrix} = -2\,, \quad A_{13} = \begin{vmatrix} 2 & 1 \\ 0 & -2 \end{vmatrix} = -4\,.$$

□

6.5 I teoremi di Laplace

In questo paragrafo vedremo come è possibile ridurre il calcolo di un determinante di un fissato ordine al calcolo di determinanti di ordini più piccoli.

PROPOSIZIONE 6.5.1. *Sia $A \in \mathbb{F}^{n,n}$ una matrice del tipo*

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}. \tag{6.12}$$

Allora risulta

$$\det A = \begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} = a_{11}A_{11} \tag{6.13}$$

DIMOSTRAZIONE. Sappiamo che, per definizione, é

$$\det A = \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

e che ogni addendo é un prodotto nel quale compare un unico fattore appartenente alla prima riga di A. Allora, nelle nostre ipotesi, sono nulli tutti i prodotti in cui non compare a_{11}, cioè i prodotti corrispondenti alle permutazioni σ per cui é $\sigma(1) \neq 1$. Nell'espressione di $\det A$ compaiono, dunque, soltanto gli addendi corrispondenti alle permutazioni per cui é $\sigma(1) = 1$ e, di conseguenza abbiamo

$$\begin{aligned} \det A &= \sum_{\sigma(1)=1} sgn(\sigma) a_{11} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \\ &= a_{11} \sum_{\sigma(1)=1} sgn(\sigma) a_{2\sigma(2)} a_{3\sigma(3)} \cdots a_{n\sigma(n)} \\ &= a_{11} \begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} \end{aligned},$$

cioè l'asserto. □

COROLLARIO 6.5.2. *Sia $A \in \mathbb{F}^{n,n}$ una matrice nella quale siano nulli tutti gli elementi della riga i−esima diversi dal primo a_{i1}, cioè*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Allora risulta

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{i1}(-1)^{i+1}M_{i1}(A) =$$

$$= a_{i1}(-1)^{i+1} \begin{vmatrix} a_{12} & a_{13} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{(i-1)2} & a_{(i-1)3} & \cdots & a_{(i-1)n} \\ a_{(i+1)2} & a_{(i+1)3} & \cdots & a_{(i+1)n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} = a_{i1}A_{i1}. \tag{6.14}$$

Dimostrazione. Consideriamo la matrice B ottenuta spostando in prima posizione l'$i-$esima riga di A e osserviamo che questa è il risultato di un numero di scambi di righe di A pari a $i-1$. Inoltre B è del tipo (6.12). Allora, in forza della proprietà D2 e della (6.13), abbiamo

$$\det A = (-1)^{i-1}\det B = (-1)^{i-1}a_{i1}M_{11}(B) = (-1)^{i-1}a_{i1}M_{i1}(A)$$

$$= (-1)^{i-1}a_{i1}\left((-1)^{i+1}A_{i1}\right) = (-1)^{2i}a_{i1}A_{i1} = a_{i1}A_{i1}$$

Corollario 6.5.3. *Sia $A \in \mathbb{F}^{n,n}$ una matrice nella quale siano nulli tutti gli elementi della riga $i-$esima diversi dall'elemento a_{ij}. Allora risulta*

$$\det A = a_{ij}A_{ij}. \tag{6.15}$$

Dimostrazione. Si costruisca la matrice B come nella dimostrazione del lemma precedente e poi si consideri la matrice C ottenuta dalla B spostando in prima posizione la colonna $j-$esima. Si osservi che C si ottiene da A con $i-1$ scambi di righe e $j-1$ scambi di colonne. Allora, usando la proprietà D2 e il lemma precedente, si ha:

$$\det A = (-1)^{i-1}(-1)^{j-1}\det C = (-1)^{i+j-2}a_{ij}M_{ij}(B)$$

$$= (-1)^{i+j-2}a_{ij}\left((-1)^{i+j}A_{ij}\right)(-1)^{2i+2j-2}a_{ij}A_{ij} = a_{ij}A_{ij}. \qquad \square$$

Proposizione 6.5.4. (Teorema di Laplace) *Sia $A \in \mathbb{F}^{n,n}$, e si fissi una linea di A. Allora il determinante di A é uguale alla somma dei prodotti degli elementi della linea fissata per i rispettivi complementi algebrici; risulta cioè*

$$\begin{aligned} \det A &= a_{j1}A_{j1} + a_{j2}A_{j2} + \cdots + a_{jn}A_{jn} \\ &= a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} \end{aligned} \tag{6.16}$$

per ogni intero j, $1 \leq j \leq n$.

DIMOSTRAZIONE. Per ogni intero t, $1 \leq t \leq n$, denotiamo con A_t la matrice che si ottiene A sostituendo il $j-$esimo vettore riga $\boldsymbol{a}_j$ col vettore $(0, ..., 0, a_{jt}, 0, ..., 0)$ avente la $t-$ esima componente uguale ad a_{jt} e le rimanenti uguali a 0. Scrivendo la $j-$esima riga della matrice A come

$$(a_{j1}, a_{j2}, ..., a_{in}) = (a_{j1}, 0, ..., 0) + \cdots + (0, a_{j2}, 0, ..., 0) + (0, ..., 0, a_{jn})$$

e applicando la proprietà D6 e il lemma 6.5.3, otteniamo

$$\begin{aligned}\det A &= \det A_1 + \det A_2 + \cdots + \det A_n \\ &= a_{j1}A_{j1} + a_{j2}A_{j2} + \cdots + a_{jn}A_{jn}.\end{aligned}$$

Resta così provata la prima uguaglianza. La seconda si prova con lo stesso ragionamento sulle righe della trasposta di A. □

COROLLARIO 6.5.5. *Sia $A \in \mathbb{F}^{n,n}$, e si fissino due indici di riga distinti i, j. Allora la somma dei prodotti degli elementi a_{ih} della riga $i-$esima per i complementi algebrici A_{jh} degli elementi corrispomdenti a_{jh} della riga $j-esima$ è uguale a zero; risulta cioè*

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn} = 0 \tag{6.17}$$

Analoga proprietà vale per le colonne di A, cioè

$$a_{1i}A_{1j} + a_{2i}A_{2j} + \cdots + a_{ni}A_{nj} = 0 \tag{6.18}$$

DIMOSTRAZIONE. Si consideri la matrice B ottenuta sostituendo la $j-$esima riga di A con la $i-$esima. Allora B, contenendo due righe uguali ha il determinante uguale a zero. Se si scrive tale determinante usando il teorema di Laplace applicato alla riga $j-$esima di B si ottiene esattamente la prima delle (6.17). □

Le relazioni (6.16) sono rispettivamente note come lo *sviluppo di Laplace del determinante secondo una riga* e *secondo una colonna.*

ESEMPIO 6.5.6. Abbiamo già visto che per la matrice quadrata d'ordine 3 sul campo reale

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

si ha

$$A_{11} = \begin{vmatrix} 1 & 0 \\ -2 & 1 \end{vmatrix} = 1\,, \quad A_{12} = -\begin{vmatrix} 2 & 0 \\ 0 & 1 \end{vmatrix} = -2\,, \quad A_{13} = \begin{vmatrix} 2 & 1 \\ 0 & -2 \end{vmatrix} = -4\,.$$

Allora, calcolando il determinante di A con lo sviluppo di Laplace secondo la prima riga, abbiamo

$$\det A = a_{11}A_{12} - a_{12}A_{12} + a_{13}A_{13} = -1 - 4 = -5.$$

□

Osservazione 6.5.7. Nella sua generalità, il teorema di Laplace riduce il calcolo di un determinante d'ordine n a quello di n determinanti d'ordine $n-1$ e questo, da un punto di vista computazionale, non è un grosso vantaggio. Se, però, una linea del determinante contiene pochi (rispetto ad n) elementi diversi da zero, allora l'applicazione congiunta di questo teorema e del metodo di Gauss può portare a notevoli semplificazioni di calcolo. □

Esempio 6.5.8. La prima riga della matrice

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

presenta un solo elemento diverso da zero. Allora, per il calcolo del suo determinante, conviene usare lo sviluppo di Laplace secondo la prima riga:

$$|A| = -\begin{vmatrix} 0 & 0 & 1 \\ -1 & 1 & 0 \\ -2 & -2 & 1 \end{vmatrix} = -\begin{vmatrix} -1 & 1 \\ -2 & -2 \end{vmatrix} = -(2+2) = -4.$$

□

Esempio 6.5.9. Calcoliamo il seguente determinante con lo sviluppo di Laplace secondo la terza colonna:

$$\begin{aligned} \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 0 & -1 \\ 2 & 3 & 1 & 0 \\ 2 & 1 & 0 & -2 \end{vmatrix} &= \begin{vmatrix} 1 & -1 & -1 \\ 2 & 3 & 0 \\ 2 & 1 & -2 \end{vmatrix} + \begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 2 & 1 & -2 \end{vmatrix} \\ &= (-6+0-2+0-4+6) + (2-2+1+1+2+2) \\ &= -6+6 = 0\,. \end{aligned}$$

□

Il teorema di Laplace ammette la seguente generalizzazione, di cui, per brevità, omettiamo la dimostrazione.

Proposizione 6.5.10. (Teorema di Laplace generalizzato) *Sia $A \in \mathbb{F}^{n,n}$ e si fissino k righe (risp. colonne) di A. Allora il determinante di A é uguale alla somma dei prodotti dei minori d'ordine k estratti dalle k righe (risp. colonne) fissate per i rispettivi cofattori.*

Esempio 6.5.11. Applicando il teorema di Laplace generalizzato alle prime due righe della matrice

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 12 & 7 & 0 & 1 & 0 & 0 \\ -54 & 2 & 0 & -1 & 0 & 1 \\ 3 & 66 & -1 & 2 & 1 & 0 \\ 5 & 20 & -2 & 1 & -2 & 1 \end{pmatrix},$$

abbiamo:

$$|A| = \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} \begin{vmatrix} 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ -1 & 2 & 1 & 0 \\ -2 & 1 & -2 & 1 \end{vmatrix} = - \begin{vmatrix} 0 & 0 & 1 \\ -1 & 1 & 0 \\ -2 & -2 & 1 \end{vmatrix} = - \begin{vmatrix} -1 & 1 \\ -2 & -2 \end{vmatrix} = -4.$$

Si osservi che, in questo caso, abbiamo ridotto il calcolo di un determinante d'ordine 6 a quello di due determinanti d'ordine 2. $\square$

Capitolo 7

Compatibilità e soluzioni dei sistemi lineari

7.1 Rango di una matrice

Sia $A \in \mathbb{F}^{m,n}$ una matrice con m righe e n colonne a coefficienti in un campo $\mathbb{F}$. Denotiamo con

$$\boldsymbol{a}_{(i)} = (a_{i1}, a_{i2}, \ldots, a_{1n})\,, \;\; i = 1, 2, \ldots, m,$$

i vettori riga e con

$$\boldsymbol{a}^{(j)} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}, \;\; j = 1, 2, \ldots, n,$$

i vettori colonna di A. Ricordiamo che con queste notazioni la matrice A può scriversi

$$A = \begin{pmatrix} \boldsymbol{a}_{(1)} \\ \boldsymbol{a}_{(2)} \\ \vdots \\ \boldsymbol{a}_{(m)} \end{pmatrix} = (\boldsymbol{a}^{(1)} \boldsymbol{a}^{(2)} \ldots \; \boldsymbol{a}^{(n)})\,.$$

Definiamo *rango* di A, e lo denotiamo con $r(A)$, il rango delle righe di A considerate come vettori di $\mathbb{F}^n$. In altre parole, il rango di A è il massimo numero di righe di A linearmente indipendenti (*cfr.* prop.5.7.22). Quando la matrice A risulta chiara dal contesto si può scrivere r in luogo di $r(A)$.

Proposizione 7.1.1. *Siano $A \in \mathbb{F}^{m,n}$ e B una sottomatrice quadrata di A d'ordine h a determinante diverso da zero. Allora le righe (risp. colonne) di A corrispondenti a quelle di B sono linearmente indipendenti.*

Dimostrazione. Posto $A = (a_{ij})$, non è restrittivo supporre che B sia individuata dalle prime h righe e dalle prime h colonne di A, cioè

$$B = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1h} \\ a_{21} & a_{22} & \cdots & a_{2h} \\ \vdots & \vdots & \ddots & \vdots \\ a_{h1} & a_{h2} & \cdots & a_{hh} \end{pmatrix} = \begin{pmatrix} \boldsymbol{b}_{(1)} \\ \boldsymbol{b}_{(2)} \\ \vdots \\ \boldsymbol{b}_{(h)} \end{pmatrix}.$$

Allora, dall'essere $|B| \neq 0$, abbiamo che le righe di B, $\boldsymbol{b}_{(1)}, \boldsymbol{b}_{(2)}, \ldots, \boldsymbol{b}_{(h)}$, sono vettori indipendenti di $\mathbb{F}^h$, per la (D7) della prop.6.3.5. Ne segue che, se λ_1,λ_2, $\ldots, \lambda_h$ sono scalari, risulta

$$\begin{aligned} &\lambda_1 \boldsymbol{a}_{(1)} + \lambda_2 \boldsymbol{a}_{(2)} + \cdots + \lambda_h \boldsymbol{a}_{(h)} = \mathbf{0} \\ \Rightarrow\ &\lambda_1 \boldsymbol{b}_{(1)} + \lambda_2 \boldsymbol{b}_{(2)} + \cdots + \lambda_h \boldsymbol{b}_{(h)} = \mathbf{0} \\ \Rightarrow\ &\lambda_1 = \lambda_2 = \cdots = \lambda_h = 0 \end{aligned}$$

e, così le prime h righe di A sono linearmente indipendenti. In modo analogo si ragiona per le colonne. □

Proposizione 7.1.2. *Il rango di una matrice ridotta è uguale al numero di righe non nulle.*

Dimostrazione. È chiaro che le righe nulle di una matrice non danno alcun contributo ai fini del calcolo del rango. Osserviamo, ora, che una matrice ridotta con h righe non nulle contiene una sottomatrice triangolare d'ordine h a determinante diverso da zero: basta prendere la sottomatrice individuata dalle h righe non nulle e dalle colonne contenenti i pivot. Ne segue che tali righe sono linearmente indipendenti (*cfr.* prop.7.1.1) e ciò prova il nostro asserto. □

Esempio 7.1.3. La matrice

$$A = \begin{pmatrix} 3 & -2 & 3,7 & 0 & 9 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 7 & -2 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

è ridotta ed ha 4 righe non nulle. Allora il suo rango è 4. Si noti che un minore principale M di A è individuato dalle righe non nulle e dalle colonne che contengono i pivot e, quindi, è

$$M = \begin{vmatrix} 3 & 3,7 & 0 & 9 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 7 & -2 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 21\,.$$

□

Osservazione 7.1.4. Il risultato della prop.7.1.2 è falso per le matrici che non siano ridotte. Per esempio, la matrice

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

ha rango 1 pur avendo tre righe non nulle. La seconda e la terza riga, infatti, si ottengono moltiplicando la prima rispettivamente per 2 e per -1. $\square$

Proposizione 7.1.5. *Siano $A, B \in \mathbb{F}^{m,n}$ due matrici equivalenti. Allora le righe di B sono linearmente dipendenti se, e solo se, lo sono le righe di A.*

Dimostrazione. Basta provare l'asserto nel caso che B si ottenga mediante una sola operazione elementare sulle righe di A. È chiaro che scambiando di posto dure rghe di A o moltiplicando una riga per uno scalare non nullo non si altera l'indipendenza o la dipendenza lineare delle righe di A. Vediamo, dunque, cosa succede se applichiamo ad A un'operazione elementare di terzo tipo. A tale scopo, senza perdita di generalitá, possiamo supporre che B si ottenga sostituendo la prima riga $\boldsymbol{a}_{(1)}$ di A con $\boldsymbol{b} = \boldsymbol{a}_{(1)} + \lambda\boldsymbol{a}_{(2)}$, $\lambda \in \mathbb{F}$. In queste ipotesi, denotati con $\lambda_1, \lambda_2, \ldots, \lambda_m$, scalari non tutti nulli, abbiamo

$$\begin{aligned} &\lambda_1\boldsymbol{a}_{(1)} + \lambda_2\boldsymbol{a}_{(2)} + \cdots + \lambda_m\boldsymbol{a}_{(m)} = \mathbf{0} \\ \Leftrightarrow\ &\lambda_1(\boldsymbol{b} - \lambda\boldsymbol{a}_{(2)}) + \lambda_2\boldsymbol{a}_{(2)} + \cdots + \lambda_m\boldsymbol{a}_{(m)} = \mathbf{0}\,, \\ \Leftrightarrow\ &\lambda_1\boldsymbol{b} + (-\lambda + \lambda_2)\boldsymbol{a}_{(2)} + \cdots + \lambda_m\boldsymbol{a}_{(m)} = \mathbf{0} \end{aligned}$$

da cui ricaviamo che le righe di B sono dipendenti se, e solo se, lo sono quelle di A. $\square$

I due corollari che seguono sono un'immediata conseguenza dell'ultima proposizione.

Corollario 7.1.6. *Siano $A, R \in \mathbb{F}^{m,n}$ due matrici equivalenti e supponiamo che R sia ridotta. Allora valgono le seguenti proprietà:*

- *righe di A che, mediante trasformazioni elementari, corrispondono a righe non nulle di R sono linearmente indipendenti;*
- *righe di R che, mediante trasformazioni elementari, corrispondono a righe di A linearmente indipendenti sono non nulle;*
- *il rango di A è uguale al rango di R e coincide col numero di righe non nulle di R.*

Corollario 7.1.7. *Matrici equivalenti hanno lo stesso rango.*

Osservazione 7.1.8. (Calcolo del rango di una matrice) Il corollario 7.1.6, in effetti, permette di calcolare il rango di una matrice A mediante l'algoritmo di Gauss (o di Gauss-Jordan). Calcolata, infatti, con tale algoritmo una matrice ridotta (o completamente ridotta) R equivalente ad A, si ha che il rango di A è uguale al numero di righe non nulle di R. $\square$

PROPOSIZIONE 7.1.9. (TEOREMA DEL RANGO) *Il rango di una matrice $A \in \mathbb{F}^{m,n}$ é uguale alla sua caratteristica, risulta cioè*

$$r(A) = c(A).$$

DIMOSTRAZIONE. Detta c la caratteristica di $A = (a_{ij})$, proviamo che c è il massimo numero di righe di A linearmente indipendenti. Possiamo supporre, senza perdita di generalità, che il minore individuato dalle prime c righe e dalle prime c colonne di A sia divero da zero, cioè

$$d = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1c} \\ a_{21} & a_{22} & \cdots & a_{2c} \\ \vdots & \vdots & \ddots & \vdots \\ a_{c1} & a_{c2} & \cdots & a_{cc} \end{vmatrix} \neq 0\,. \tag{7.1}$$

In queste ipotesi le prime c righe di A

$$\boldsymbol{a}_{(1)}, \boldsymbol{a}_{(2)}, \ldots, \boldsymbol{a}_{(c)}$$

sono linearmente indipendenti (*cfr.* prop.7.1.1). A questo punto fissiamo un indice di riga s maggiore di c e osserviamo che, per ogni $t = 1, 2, \ldots, n$, risulta

$$d_{st} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1c} & a_{1t} \\ a_{21} & a_{22} & \cdots & a_{2c} & a_{2t} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{c1} & a_{c2} & \cdots & a_{cc} & a_{ct} \\ a_{s1} & a_{s2} & \cdots & a_{sc} & a_{st} \end{vmatrix} = 0, \tag{7.2}$$

perchè:

- per $t \leq c$, il determinante d_{st} ha due colonne uguali,
- per $t > c$, il determinante d_{st} è un minore d'ordice $c + 1$ della matrice A di caratteristica c.

Osserviamo, ora, che nella matrice

$$D_{st} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1c} & a_{1t} \\ a_{21} & a_{22} & \cdots & a_{2c} & a_{2t} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{c1} & a_{c2} & \cdots & a_{cc} & a_{ct} \\ a_{s1} & a_{s2} & \cdots & a_{sc} & a_{st} \end{pmatrix}$$

gli elementi dell'ultima colonna $a_{1t}, a_{2t}, \ldots, a_{ct}, a_{st}$ hanno i rispettivi complementi algebrici $D_1, D_2, \ldots, D_c, D_s$ indipendenti da t, avendosi

$$D_j = (-1)^{j+c+1} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1c} \\ \vdots & \vdots & \ddots & \vdots \\ a_{(j-1)1} & a_{(j-1)2} & \cdots & a_{(j-1)c} \\ a_{(j+1)1} & a_{(j+1)2} & \cdots & a_{(j+1)c} \\ \vdots & \vdots & \ddots & \vdots \\ a_{c1} & a_{c2} & \cdots & a_{cc} \\ a_{s1} & a_{s2} & \cdots & a_{sc} \end{vmatrix}$$

per $j = 1, 2, \ldots, c$, e

$$D_s = d\,.$$

Allora, se calcoliamo il determinante $d_{st} = |D_{st}|$ con la regola di Laplace applicata all'ultima colonna di D_{st}, abbiamo

$$a_{1t}D_1 + a_{2t}D_2 + \cdots + a_{ct}D_c + a_{st}d = 0$$

e, essendo $d \neq 0$,

$$a_{st} = -\frac{D_1}{d}\,a_{1t} - \frac{D_2}{d}\,a_{2t} - \cdots - \frac{D_c}{d}\,a_{ct}\,, \tag{7.3}$$

per ogni $t = 1, 2, \ldots, n$. Ne segue che

$$\boldsymbol{a}_{(s)} = -\frac{D_1}{d}\,\boldsymbol{a}_{(1)} - \frac{D_2}{d}\,\boldsymbol{a}_{(2)} - \cdots - \frac{D_c}{d}\,\boldsymbol{a}_{(c)}\,,$$

cioè ogni riga di A è combinazione lineare delle prime c e, quindi, queste formano un insieme massimale di righe linearmente indipendenti. Ne segue che $c = r$ (*cfr.* prop.5.7.22), cioè l'asserto. $\square$

Nel corso dell'ultima dimostrazione, per provare che r(A)=c(A), non abbiamo utilizzato il fatto che tutti i minori di A d'ordine $c+1$ fossero nulli. Abbiamo usato soltanto i minori orlati[1] di un fissato minore non nullo d'ordine c. In altre parole, abbiamo implicitamente provato il seguente risultato.

Corollario 7.1.10. (Teorema dei minori orlati) *Sia M un minore non nullo d'ordine h di una matrice $A \in \mathbb{F}^{m,n}$. Se tutti i minori orlati di M sono nulli, allora il rango di A è uguale ad h.*

Il corollario precedente giustifica il seguente algoritmo per il calcolo del rango di una matrice.

Algoritmo 7.1.11. (Regola degli orlati per il calcolo del rango di una matrice)

[1]Se M è un minore di una matrice A, l'*orlato* di M mediante una riga e una colonna di A é il determinante della sottomatrice di A corrispondente alle righe e colonne di M cui si aggiungono la riga e la colonna fissata.

- INPUT: Una matrice A.
- PASSO 1: Si fissi un minore non nullo M[2] di A.
- PASSO 2: Si calcolino in successione i minori orlati di M e, appena se ne trova uno M' diverso da zero, si ritorni al PASSO 1 sostituendo M con M', ponendo cioè $M := M'$.
- OUTPUT: L'ordine del minore M (che è uguale al rango di A in forza del teorema dei minori orlati).

Esempio 7.1.12. La matrice

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 0 & -1 & 1 \\ 2 & 3 & 1 & 0 & -1 \\ 2 & 1 & 0 & -2 & -1 \end{pmatrix} \in \mathbb{R}^{4,5}$$

ha il minore (individuato dalle prime tre colonne e dalla prima, dalla seconda e dalla quarta riga)

$$M = \begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 2 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix} = 3$$

diverso da zero. Questo significa che la prima, la seconda e la quarta riga di A sono indipendenti. Analogamente, le prime tre colonne di A sono indipendenti. I minori orlati di M sono

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

e sono entrambi nulli. Ne segue che A ha rango 3. □

Osservazione 7.1.13. L'algoritmo (7.1.11), pur essendo interessante da un punto di vista teorico, è molto più lento dell'algoritmo di Gauss (*cfr.* osservazione 7.1.8) e, quindi, conviene usarlo solo in casi particolari. □

Corollario 7.1.14. *Il rango di una matrice $A \in \mathbb{F}^{m,n}$ é uguale a quello della sua trasposta, cioè*

$$r(A) = r(A^t). \tag{7.4}$$

In altre parole, nella matrice A, il massimo numero di righe linearmente indipendenti è uguale al massimo numero di colonne linearmente indipendenti.

Dimostrazione. Tenendo presente la (6.11), basta applicare la prop.7.1.9 alla matrice trasposta di A. □

[2] Si puó scegliere anche un elemento non nullo di A.

7.2 La matrice aggiunta e la regola di Cramer

Sia $A = (a_{ij}) \in \mathbb{F}^{n,n}$ una matrice quadrata d'ordine n a coefficienti in un campo $\mathbb{F}$ e ricordiamo che abbiamo denotato con A_{ij} i complementi algebrici degli elementi a_{ij} di A. La matrice trasposta di quella (A_{ij}) dei complementi algebrici degli elementi di A si chiama *matrice aggiunta*, o semplicemente *aggiunta*, di A e si denota con A^*; si pone cioé

$$A^* = (A_{ji}) = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}. \tag{7.5}$$

Esempio 7.2.1. Per la matrice $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, risulta

$$A_{11} = a_{22}\,,\ A_{12} = -a_{21}\,,\ A_{21} = -a_{12}\,,\ A_{22} = a_{11}$$

e, quindi,

$$A^* = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

□

Proposizione 7.2.2. *Sia $A \in \mathbb{F}^{n,n}$ una matrice quadrata d'ordine n su $\mathbb{F}$. Le matrici A e A^* commutano*[3] *e il loro prodotto è uguale alla matrice diagonale con gli elementi della diagonale principale tutti uguali a* $\det A$, *cioé*

$$AA^* = A^*A = \begin{pmatrix} |A| & 0 & \cdots & 0 \\ 0 & |A| & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & |A| \end{pmatrix} = |A|I\,. \tag{7.6}$$

In particolare, se A é singolare risulta $AA^* = A^*A = \mathbf{0}$.

Dimostrazione. Posto $AA^* = C = (c_{ij})$, in forza della prima delle (6.16) e della (6.17), per ogni i e per ogni j diverso da i, risulta

$$c_{ii} = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} = \det A$$

$$c_{ij} = a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn} = 0\,.$$

Allo stesso modo, usando però la seconda delle (6.16) e la (6.18), si vede che $A^*A = |A|I$. □

Come conseguenza immediata della proposizione precedente si ottiene una formula che, per una matrice invertibile A, esprime i coefficienti di A^{-1} mediante quelli di A.

[3]Si dice che due matrici quadrate A, B *commutano* se risulta $AB = BA$.

COROLLARIO 7.2.3. (TEOREMA DELLA MATRICE INVERSA) *Se* $A \in \mathbb{F}^{n,n}$ *è non singolare, risulta*

$$A^{-1} = \frac{1}{|A|} A^* = \begin{pmatrix} \frac{A_{11}}{|A|} & \frac{A_{21}}{|A|} & \dots & \frac{A_{n1}}{|A|} \\ \frac{A_{12}}{|A|} & \frac{A_{22}}{|A|} & \dots & \frac{A_{n2}}{|A|} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{A_{1n}}{|A|} & \frac{A_{2n}}{|A|} & \dots & \frac{A_{nn}}{|A|} \end{pmatrix}. \tag{7.7}$$

DIMOSTRAZIONE. Nelle nostre ipotesi A è invertibile in forza della prop. 6.3.12. Per ottenere la (7.7), basta dividere l'uguaglianza $AA^* = |A|I$ per $|A|$ e moltiplicarla a sinistra per A^{-1}. □

ESEMPIO 7.2.4. Applichiamo il teorema della matrice inversa alla seguente matrice

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

il cui determinante è $|A| = 9$. A tale scopo, costruiamo l'aggiunta di A; risulta

$$A_{11} = \begin{vmatrix} 1 & 1 \\ 2 & 3 \end{vmatrix}, \qquad A_{21} = -\begin{vmatrix} 2 & 0 \\ 2 & 3 \end{vmatrix}, \qquad A_{31} = \begin{vmatrix} 2 & 0 \\ 1 & 1 \end{vmatrix},$$

$$A_{12} = -\begin{vmatrix} -1 & 1 \\ 1 & 3 \end{vmatrix}, \quad A_{22} = \begin{vmatrix} 1 & 0 \\ 1 & 3 \end{vmatrix}, \quad A_{32} = -\begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix},$$

$$A_{13} = \begin{vmatrix} -1 & 1 \\ 1 & 2 \end{vmatrix}, \quad A_{23} = -\begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix}, \quad A_{33} = \begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix},$$

e, quindi,

$$A^* = \begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix} = \begin{pmatrix} 1 & -6 & 2 \\ 4 & 3 & -1 \\ -3 & 0 & 3 \end{pmatrix}.$$

In conclusione, abbiamo

$$A^{-1} = \frac{1}{|A|} A^* = \begin{pmatrix} \frac{1}{9} & \frac{-6}{9} & \frac{2}{9} \\ \frac{4}{9} & \frac{3}{9} & \frac{-1}{9} \\ \frac{-3}{9} & 0 & \frac{3}{9} \end{pmatrix}.$$

□

OSSERVAZIONE 7.2.5. La formula (7.7) è importante da un punto di vista teorico perché esprime i coefficienti dell'inversa di una matrice in funzione dei coefficienti della matrice stessa. Dal punto di vista computazionale, invece, è difficilmente utilizzabile. Per applicarla, infatti, ad una matrice A d'ordine n bisogna calcolare $n+1$ determinanti d'ordine n. Nella pratica, per il calcolo di A^{-1} conviene usare l'algoritmo 3.9.7, che è sostanzialmente equivalente al calcolo di un solo determinante d'ordine n. □

Proposizione 7.2.6. (Regola di Cramer) *Sia*

$$Ax = b \tag{7.8}$$

un sistema lineare di n equazioni in n incognite a coefficienti in un campo $\mathbb{F}$ con A non singolare. Per ogni $i = 1, 2, \ldots, n$ denotiamo con $A^{(i)}$ la matrice che si ottiene sostituendo col vettore colonna $\boldsymbol{b}$ l'$i-$esima colonna di A, cioè

$$A^{(i)} = \left(\boldsymbol{a}^{(1)}\, \boldsymbol{a}^{(2)} \,\ldots\, \boldsymbol{a}^{(i-1)}\, \boldsymbol{b}\, \boldsymbol{a}^{(i+1)} \,\cdots\, \boldsymbol{a}^{(n)} \right). \tag{7.9}$$

Allora il sistema (7.8) ha un'unica soluzione $\boldsymbol{c} = (c_1, c_2, \ldots, c_n)$ data da

$$\boldsymbol{c} = A^{-1}\boldsymbol{b} \tag{7.10}$$

e risulta

$$c_1 = \frac{|A^{(1)}|}{|A|},\ c_2 = \frac{|A^{(2)}|}{|A|}, \ldots,\ c_n = \frac{|A^{(n)}|}{|A|}. \tag{7.11}$$

In particolare, un sistema omogeneo di n equazioni in n incognite con matrice associata di rango n ammette soltanto la soluzione nulla.

Dimostrazione. Il vettore $A^{-1}\boldsymbol{b}$ è una soluzione del sistema (7.8) perché

$$A(A^{-1}\boldsymbol{b}) = (AA^{-1})\boldsymbol{b} = \boldsymbol{b}\,.$$

D'altra parte, se $\boldsymbol{c}$ è una soluzione del sistema (7.8) si ha

$$\begin{aligned} A\boldsymbol{c} = \boldsymbol{b} \ &\Leftrightarrow\ A^{-1}(A\boldsymbol{c}) = A^{-1}\boldsymbol{b} \\ &\Leftrightarrow (A^{-1}A)\boldsymbol{c} = A^{-1}\boldsymbol{b} \\ &\Leftrightarrow \boldsymbol{c} = A^{-1}\boldsymbol{b} \end{aligned}$$

e si ha così la prima parte dell'asserto. Per provare la (7.11), basta osservare che lo sviluppo di Laplace del determinante $|A^{(j)}|$ secondo la $j-$esima colonna è

$$|A^{(j)}| = A_{1j}b_1 + A_{2j}b_2 + \cdots + A_{nj}b_n$$

e, quindi, in forza della (7.7) risulta

$$c_j = \frac{1}{|A|}\,(A_{1j}b_1 + A_{2j}b_2 + \cdots + A_{nj}b_n) = \frac{|A^{(j)}|}{|A|}\,.$$

□

Un sistema lineare che verifica le ipotesi della proposizione precedente si chiama *sistema di Cramer* e le relazioni (7.11) si chiamano *formule di Cramer.*

Esempio 7.2.7. Il sistema lineare

$$\begin{cases} x & & + & z & = & 3 \\ & y & + & z & = & 2 \\ x & + & y & & = & 5 \end{cases}$$

ha il determinante della matrice dei coefficienti A diverso da zero, infatti

$$|A| = \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{vmatrix} = -2\,.$$

Per la sua unica soluzione, calcolata con la regola di Cramer, si ha

$$x = -\frac{1}{2}\begin{vmatrix} 3 & 0 & 1 \\ 2 & 1 & 1 \\ 5 & 1 & 0 \end{vmatrix} = 3\,,\ y = -\frac{1}{2}\begin{vmatrix} 1 & 3 & 1 \\ 0 & 2 & 1 \\ 1 & 5 & 0 \end{vmatrix} = 2\,,\ z = -\frac{1}{2}\begin{vmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 1 & 1 & 5 \end{vmatrix} = 0\,.$$

□

Osservazione 7.2.8. Le formule di Cramer sono importanti da un punto di vista teorico perché esprimono esplicitamente la soluzione di un sistema in funzione dei coefficienti e del termine noto del sistema stesso. Anche queste, però, dal punto di vista pratico sono difficilmente utilizzabili perché presentano le stesse difficoltà computazionali della (7.7). Così, per risolvere un sistema di n equazioni in n incognite, con matrice associata non singolare, conviene utilizzare il metodo di Gauss che, come già detto, è sostanzialmente equivalente al calcolo di un solo determinante d'ordine n. □

7.3 Compatibilità e soluzioni di un sistema lineare

In tutto il paragrafo supporremo fissato un sistema lineare

$$S = \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}, \tag{7.12}$$

di m equazioni in n incognite a coefficienti in un campo $\mathbb{F}$ che, come al solito, scriveremo anche in forma compatta

$$A\boldsymbol{x} = \boldsymbol{b}\,,\ \text{con } A = (a_{ij}) \text{ e } \boldsymbol{b} = (b_1, b_2, \ldots, \boldsymbol{b}_n)^t.$$

Di un tale sistema vogliamo determinare le condizioni di compatibilità e, quando esistono, le sue soluzioni in funzione dei coefficienti e del termine noto, come abbiamo fatto per i sistemi di Cramer.

Proposizione 7.3.1. (Teorema di Rouché-Capelli) *Un sistema lineare*

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

del tipo (7.12) *è compatibile se, e solo se, il rango della matrice dei coefficienti* A *è uguale al rango della matrice completa* $A' = (A|\boldsymbol{b})$.

Dimostrazione. Possiamo suppore che $\boldsymbol{b}$ non sia il vettore nullo, altrimenti il sistema ammette almeno la soluzione nulla e l'asserto è vero. In queste ipotesi, una eventuale soluzione $\boldsymbol{c}$ del il sistema (7.12) verifica la relazione

$$\boldsymbol{b} = A\boldsymbol{c} = c_1\boldsymbol{a}^{(1)} + c_2\boldsymbol{a}^{(2)} + \cdots + c_n\boldsymbol{a}^{(n)} \tag{7.13}$$

e, quindi, $\boldsymbol{b}$ è combinazione lineare delle colonne di A. Ne segue che l'ultima colonna di A' dipende dalle precedenti e, di conseguenza, risulta $r(A) = r(A')$. Viceversa, nell'ipotesi $r(A) = r(A')$, il vettore colonna $\boldsymbol{b}$ dipende linearmente dalle colonne di A, altrimenti avremmo $r(A') = r(A) + 1$, e dunque vale una relazione del tipo (7.13). Allora $\boldsymbol{b}$ è soluzione del nostro sistema e l'asserto è completamente provato. □

Esempio 7.3.2. Le matrici, quella dei coefficienti e quella completa, del sistema lineare

$$\begin{cases} x & & + & z & = & 1 \\ x & + & y & + & z & = & 1 \\ 2x & + & y & + & 2z & = & 1 \end{cases}$$

sono

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 2 & 1 & 2 \end{pmatrix} \quad \text{e} \quad A' = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 2 & 1 & 2 & 1 \end{pmatrix}.$$

Le prime due righe di A sono indipendenti e la terza è somma delle prime due; ne segue che è $r(A) = 2$. D'altra parte, il minore di A' individuato dalla prima, dalla seconda e dalla quarta colonna

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

è diverso da zero e, quindi, $r(A') = 3$. Allora, essendo $r(A) \neq r(A')$, il sistema non è compatibile. □

Proposizione 7.3.3. *Assegnato un sistema lineare compatibile del tipo* (7.12), *siano* r *il rango e* M *un minore d'ordine* r *della matrice* A *dei suoi coefficienti. Allora il sistema* S *è equivalente a quello* S' *che si ottiene considerando soltanto le equazioni corrispondenti alle righe di* A *che individuano il minore* M.

Dimostrazione. In forza delle proposizioni 7.1.1 e 7.1.9 le equazioni di S corrispondenti alle righe di A non individuate da M dipendono linearmente dalle equazioni di S' e così si ha l'asserto. □

OSSERVAZIONE 7.3.4. La proposizione precedente assicura che quando si studia un sistema compatibile si possono sempre trascurare delle sue equazioni in modo da ottenere un nuovo sistema, equivalente a quello iniziale, nel quale il numero delle equazioni uguaglia il rango della matrice dei coefficienti. È questo il motivo per cui, nel seguito, supporremo spesso che la matrice dei coefficienti del sistema (7.12) abbia rango m. □

Assegnati il sistema (7.12) e un minore principale (cioè d'ordine massimo) M della matrice A, diciamo *principali* (rispetto ad M) le incognite individuate dalle colonne di M. Le incognite non principali si dicono*libere,* o *secondarie,* o *parametri.* Notiamo che, in forza della prop.7.3.3, il secondo teorema di unicità per i sistemi lineari (*cfr.* prop.5.5.2) può enunciarsi nel seguente modo.

PROPOSIZIONE 7.3.5. *Assegnato il sistema lineare* (7.12)*, siano r il rango e M un minore d'ordine r della matrice A dei suoi coefficienti. Siano, inoltre, $x_{j_1}, x_{j_2}, \dots, x_{j_r}$ le incognite principali rispetto ad M. Allora ciascuna soluzione del sistema corrisponde biunivocamente ad un prefissato valore delle $n-r$ incognite libere $x_{t_1}, x_{t_2}, \dots, x_{t_{n-r}}$; in altre parole, fissati liberamente $n-r$ scalari $\alpha_{t_1}, \alpha_{t_2}, \dots, \alpha_{t_{n-r}}$, esiste un'unica soluzione del sistema $\boldsymbol{c} = (c_1, c_2, \dots, c_n)$ tale che*

$$c_{t_1} = \alpha_{t_1}\,,\ c_{t_2} = \alpha_{t_2}\,, \dots,\ c_{t_{n-r}} = \alpha_{t_{n-r}}\,.$$

Nel caso di un sistema omogeneo, una base dello spazio delle sue soluzioni può calcolarsi utilizzando il risultato seguente.

PROPOSIZIONE 7.3.6. (BASE DELLO SPAZIO DELLE SOLUZIONI DI UN SISTEMA OMOGENEO) *Nelle stesse ipotesi e con le stesse notazioni della proposizione precedente, supponiamo che il sistema lineare S sia omogeneo. Allora lo spazio $Sol(S)$ delle soluzioni di S ha dimensione $n-r$ e le soluzioni*

$$\boldsymbol{c}_1\,,\ \boldsymbol{c}_2\,, \dots,\ \boldsymbol{c}_{n-r}$$

corrispondenti rispettivamente ai valori

$$\begin{array}{c}(x_{t_1}, x_{t_2}, \dots, x_{t_{n-r}}) = (1, 0, \dots, 0)\\ (x_{t_1}, x_{t_2}, \dots, x_{t_{n-r}}) = (0, 1, \dots, 0)\\ \vdots \\ (x_{t_1}, x_{t_2}, \dots, x_{t_{n-r}}) = (0, 0, \dots, 1)\end{array}$$

delle incognite libere ne costituiscono una base e, di conseguenza, le soluzioni di S sono tutte e sole le combinazioni lineari di $\boldsymbol{c}_1, \boldsymbol{c}_2, \dots, \boldsymbol{c}_{n-r}$, cioè

$$Sol(S) = \{\lambda_1\boldsymbol{c}_1 + \lambda_2\boldsymbol{c}_2 + \dots + \lambda_{n-r}\boldsymbol{c}_{n-r}\ :\ (\lambda_1, \lambda_2, \dots, \lambda_{n-r}) \in \mathbb{F}^{n-r}\}.$$ □

DIMOSTRAZIONE. La matrice le cui righe sono $\boldsymbol{c}_1\,,\ \boldsymbol{c}_2\,, \dots,\ \boldsymbol{c}_{n-r}$ contiene I_{n-r} come sottomatrice e, quindi, tali vettori sono linearmente indipendenti (*cfr.* esempio 5.4.9). Inoltre, essendo $Sol(S)$ un sottospazio vettoriale di $\mathbb{F}^n$, ogni combinazione lineare di $\boldsymbol{c}_1$, $\boldsymbol{c}_2, \dots, \boldsymbol{c}_{n-r}$ è una soluzione di S. Ciò premesso, consideriamo una soluzione

$$\boldsymbol{b} = (b_1, \dots, b_r, b_{r+1}, \dots, b_n)$$

di S e osserviamo che questa è l'unica che corrisponde ai valori

$$(x_{r+1}, x_{r+2}, \dots, x_n) = (b_{r+1}, b_{r+2}, \dots, b_n) \tag{7.14}$$

delle incognite libere. D'altra parte, anche il vettore

$$b_{r+1}\boldsymbol{c}_1 + b_{r+2}\boldsymbol{c}_2 + \cdots + b_n\boldsymbol{c}_{n-r}$$

è una soluzione di S corrispondente ai valori (7.14) delle incognite libere e, di conseguenza, deve essere

$$\boldsymbol{b} = b_{r+1}\boldsymbol{c}_1 + b_{r+2}\boldsymbol{c}_2 + \cdots + b_n\boldsymbol{c}_{n-r}\,.$$

L'asserto è, dunque, completamente provato. □

Esempio 7.3.7. Consideriamo il sistema omogeneo

$$S = \begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 0 \\ x_1 - x_2 - x_4 + x_5 = 0 \\ 2x_1 + 3x_2 + x_3 - x_5 = 0 \\ 2x_1 + x_2 - 2x_4 - x_5 = 0 \end{cases}$$

e osserviamo che, nella matrice A dei coefficienti, il minore individuato dalle prime tre colonne e dalla prima, dalla seconda e dalla quarta riga

$$M = \begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 2 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix} = 3$$

è diverso da zero. Questo minore di A ha i due orlati

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

che sono entrambi nulli. Ne segue che il nostro sistema è equivalente a

$$\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 0 \\ x_1 - x_2 - x_4 + x_5 = 0 \\ 2x_1 + x_2 - 2x_4 - x_5 = 0 \end{cases},$$

ove x_1, x_2, x_3 sono le incognite principale e x_4, x_5 quelle libere. Allora, scriviamo per comodità il sistema nella forma

$$\begin{cases} x_1 + x_2 + x_3 = - x_4 - x_5 \\ x_1 - x_2 = + x_4 - x_5 \\ 2x_1 + x_2 = + 2x_4 + x_5 \end{cases}$$

e, con la regola di Cramer, calcoliamo le due soluzioni corrispondenti ai valori $(x_4, x_5) = (1, 0)$ e $(x_4, x_5) = (0, 1)$ delle incognite libere. Nel primo caso dobbiamo risolvere il sistema di Cramer

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

che ha per soluzione

$$\begin{aligned} x_1 =& \frac{1}{M} \begin{vmatrix} -1 & 1 & 1 \\ 1 & -1 & 0 \\ 2 & 1 & 0 \end{vmatrix} = \frac{1}{3}\, 3 = 1, \\ x_2 =& \frac{1}{M} \begin{vmatrix} 1 & -1 & 1 \\ 1 & 1 & 0 \\ 2 & 2 & 0 \end{vmatrix} = \frac{1}{3}\, 0 = 0, \\ x_3 =& \frac{1}{M} \begin{vmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ 2 & 1 & 2 \end{vmatrix} = \frac{1}{3}\, (-6) = -2. \end{aligned}$$

Nel secondo caso dobbiamo risolvere il sistema di Cramer

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

che ha per soluzione

$$\begin{aligned} x_1 =& \frac{1}{M} \begin{vmatrix} -1 & 1 & 1 \\ -1 & -1 & 0 \\ 1 & 1 & 0 \end{vmatrix} = \frac{1}{3}\, 0 = 0, \\ x_2 =& \frac{1}{M} \begin{vmatrix} 1 & -1 & 1 \\ 1 & -1 & 0 \\ 2 & 1 & 0 \end{vmatrix} = \frac{1}{3}\, 3 = 1, \\ x_3 =& \frac{1}{M} \begin{vmatrix} 1 & 1 & -1 \\ 1 & -1 & -1 \\ 2 & 1 & 1 \end{vmatrix} = \frac{1}{3}\, (-6) = -2. \end{aligned}$$

In conclusione le soluzioni del nostro sistema sono tutti e soli gli elementi del sottospazio vettoriale di $\mathbb{R}^5$ generato dai vettori $(1, 0, -2, 1, 0), (0, 1, -2, 0, 1)$, cioè

$$\begin{aligned} Sol(S) =& \{\lambda(1, 0, -2, 1, 0) + \mu(0, 1, -2, 0, 1)\ ,\ \lambda, \mu \in \mathbb{R}\} \\ =& \{(\lambda, \mu, -2(\lambda + \mu), \lambda, \mu)\ ,\ \lambda, \mu \in \mathbb{R}\}; \end{aligned}$$

e, quindi,

$$\begin{aligned} x_1 &= \lambda \\ x_2 &= \mu \\ x_3 &= -2(\lambda+\mu) \\ x_3 &= \lambda \\ x_4 &= \mu\,. \end{aligned}$$

□

I risultati precedenti giustificano il seguente algoritmo per lo studio di un sistema lineare.

Algoritmo 7.3.8. (Studio di un sistema lineare con la regola di Cramer)

- INPUT: Un sistema lineare $AX = \boldsymbol{b}$ di m equazioni in n incognite del tipo (7.12) di rango m.
- PASSO 1: Calcolare un minore M d'ordine massimo m della matrice A.
- PASSO 2: Calcolare una base $\boldsymbol{c}_1, \boldsymbol{c}_2, \dots, \boldsymbol{c}_{n-m}$ del sistema omogeneo $AX = 0$ associato ad S col metodo descritto nella prop.7.3.6
- PASSO 3: Calcolare una soluzione $\boldsymbol{c}$ di S assegnando un valore arbitrario alle incognite libere (rispetto ad M) e ricavando, con la regola di Cramer, i corrispondenti valori delle incognite principali.
- OUTPUT: L'insieme delle soluzioni di S

$$Sol(S) = \{\boldsymbol{c}+\lambda_1\boldsymbol{c}_1+\lambda_2\boldsymbol{c}_2+\cdots+\lambda_{n-m}\boldsymbol{c}_{n-m} \ : \ (\lambda_1,\lambda_2,\dots,\lambda_{n-m}) \in \mathbb{F}^{n-m}\}.$$

□

Esempio 7.3.9. Il sistema lineare

$$\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 1 \\ x_1 - x_2 - x_4 + x_5 = 1 \\ 2x_1 + 3x_2 + x_3 - x_5 = 1 \\ 2x_1 + x_2 - 2x_4 - x_5 = 1 \end{cases}$$

ha come sistema omogeneo associato quello studiato nell'esempio 7.3.7; la matrice A dei suoi coefficienti ha, quindi, rango 3 e un suo minore principale è (individuato dalle prime tre colonne e dalla prima, dalla seconda e dalla quarta riga)

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

L'unico minore orlato di M con la colonna dei termini noti è

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

ed è nullo perché la terza riga è uguale alla somma della prima, dell'opposta della seconda e della quarta. Allora anche la matrice completa ha rango 3 e, di conseguenza, il sistema è compatibile. In forza del teorema precedente, conoscendo già le soluzioni del sistema omogeneo associato, dobbiamo soltanto trovare una soluzione particolare del sistema stesso. A tale scopo osserviamo che il nostro sistema è equivalente a

$$\begin{cases} x_1 & + & x_2 & + & x_3 & = & - & x_4 & - & x_5 + & 1 \\ x_1 & - & x_2 & & & = & + & x_4 & - & x_5 + & 1 \\ 2x_1 & + & x_2 & & & = & + & 2x_4 & + & x_5 + & 1 \end{cases}$$

e calcoliamo l'unica soluzione che si ottiene assegnando il valore 0 alle incognite libere x_4, x_5. Dobbiamo, cioè, risolvere il sistema di Cramer

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

che ha per soluzione

$$\begin{aligned} x_1 =& \frac{1}{M} \begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 1 & 0 \end{vmatrix} = \frac{1}{3} 2 = \frac{2}{3}, \\ x_2 =& \frac{1}{M} \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 2 & 1 & 0 \end{vmatrix} = \frac{1}{3}(-1) = -\frac{1}{3}, \\ x_3 =& \frac{1}{M} \begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 2 & 1 & 1 \end{vmatrix} = \frac{1}{3} 2 = \frac{2}{3}. \end{aligned}$$

Concludendo, abbiamo che l'insieme delle soluzioni $Sol(S)$ si ottiene aggiungendo il vettore $(\frac{2}{3}, -\frac{1}{3}, \frac{2}{3}, 0, 0)$ alla soluzione generale (*cfr.* esempio 7.3.7) del sistema omogeneo associato, cioè

$$\begin{aligned} Sol(S) =& \left\{ \left(\frac{2}{3}, -\frac{1}{3}, \frac{2}{3}, 0, 0 \right) + \lambda(1, 0, -2, 1, 0) + \mu(0, 1, -2, 0, 1) \ , \ \lambda, \mu \in \mathbb{R} \right\} \\ =& \left\{ \left(\frac{2}{3} + \lambda \, , -\frac{1}{3} + \mu \, , \frac{2}{3} - 2(\lambda + \mu), \lambda, \mu \right) \ , \ \lambda, \mu \in \mathbb{R} \right\} \end{aligned}$$

e, quindi,

$$\begin{aligned} x_1 =& \frac{2}{3} + \lambda \\ x_2 =& -\frac{1}{3} + \mu \\ x_3 =& \frac{2}{3} - 2(\lambda + \mu) \\ x_3 =& \lambda \\ x_4 =& \mu \, . \end{aligned}$$

□